ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ФРАНКА
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ДЕРЖАВНА УСТАНОВА «ІНСТИТУТ ХАРЧОВОЇ БІОТЕХНОЛОГІЇ ТА ГЕНОМІКИ НАЦІОНАЛЬНОЇ АКАДЕМІЇ НАУК УКРАЇНИ»
НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ

Кваліфікаційна наукова праця на правах наукової доповіді за сукупністю статей

Ющук Олександр Сергійович

УДК 575+577.21+579.2+579.6

ДИСЕРТАЦІЯ

Генетичні механізми регуляції біосинтезу пептидних та полікетидних антибіотиків

03.00.22 – молекулярна генетика
09 Біологія

Подається на здобуття наукового ступеня доктора біологічних наук

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.



О.С. Ющук

Львів – 2023

Присвячується моїм батькам

Автор хоче висловити подяку студентами і аспірантам кафедри генетики та біотехнології ЛНУ ім. І. Франка та Університету Інзубрії за їх внесок, зокрема І. Власюк, В. Гомонюк, А. Степанишин, К. Жукровській, С. Мельник, Е. Сальваде, А. Андрео-Відалю.

АНОТАЦІЯ

Ющук О.С. Генетичні механізми регуляції біосинтезу пептидних та полікетидних антибіотиків. – Кваліфікаційна наукова праця на правах наукової доповіді за сукупністю статей.

Дисертація на здобуття наукового ступеня доктора біологічних наук за спеціальністю 03.00.22 «Молекулярна генетика» (09 – Біологія). – Львівський національний університет імені Івана Франка, Львів. – ДУ «Інститут харчової біотехнології та геноміки НАН України» НАН України, Київ, 2023.

Розповсюдження штамів патогенних мікроорганізмів, резистентних до антибіотиків, створює нові виклики всесвітній системі охорони здоров'я, забираючи мільйони життів щороку. Гостро стоїть питання пошуку нових антибіотиків і покращення продукції вже відомих сполук, а також вивчення механізмів стійкості до них. Більшість антибіотиків природного походження, активних проти Грам-позитивних, Грам-негативних бактерій і патогенних грибів, походять з деяких груп актинобактерій, відомих під нетаксономічною збірною назвою актиноміцети. Актиноміцети є Грам-позитивними організмами, що мешкають в ґрунті у формі міцелію і здатні до морфологічної диференціації. Значні за розміром геноми актиноміцетів кодують десятки кластерів біосинтетичних генів (КБГ) антибіотиків, що є потенційним джерелом нових сполук для боротьби із мультирезистентними патогенами.

У вивченні біосинтезу антибіотиків ключове значення має дослідження регуляції біосинтезу цих сполук, бо саме функціонування регуляторних генів визначає як активно і за яких умов продукуватиметься той чи інший антибіотик, як координуватиметься біосинтез антибіотика з ростом і морфогенезом культури продуцента та буде забезпечуватися його стійкість до власного антибіотика. Вивчення регуляторних механізмів біосинтезу антибіотиків включає аналіз геномів продуцентів *in silico* в пошуках КБГ, експериментальне вивчення ролі генів глобальних і шлях-специфічних регуляторів, їх застосування для створення надпродуцентів антибіотиків.

Враховуючи те, що переважаюча більшість КБГ антибіотиків в геномах актиноміцетів є мовчазною, використання регуляторних генів часто є єдиним способом активації мовчазних КБГ. Успішне виконання досліджень, перелічених вище, потребує зручних та ефективних методів для генетичних і генно-інженерних маніпуляцій із бактеріями-продуцентами.

Одними з найважливіших антибіотиків, до продукції яких здатні актиноміцети, є пептидні і полікетидні антибіотики. Ці сполуки характеризуються значною різноманітністю хімічних структур, яку забезпечує модульна організація нерибосомних пептидсинтаз (НРПС) і полікетидсинтаз (ПКС) – ключових ферментів біосинтезу цих сполук. Серед пептидних і полікетидних антибіотиків є сполуки, що входять до списку критично-важливих для медицини антибіотиків, створеному Всесвітньою Організацією охорони здоров'я (ВООЗ). Зокрема, глікопептидні антибіотики (ГПА) – такі як ванкоміцин і тейкопланін – визнані критично-важливими антибіотиками із найвищим пріоритетом для їх вивчення і нагляду за формуванням стійкості до ГПА у патогенів. В цій роботі ми зосередили свою увагу на таких штамах-продуцентах пептидних і полікетидних (та деяких інших) антибіотиків: *Actinoplanes teichomyceticus* ATCC 31121 (продуцент тейкопланіну), *Actinoplanes rectilineatus* NRRL B-16090 (несе КБГ ГПА), *Nonomuraea gerenzanensis* ATCC 39727 (продуцент A40926), *Nonomuraea coxensis* DSM 45129 (продуцент A50926), *Streptomyces roseochromogenes* NRRL 3504 (продуцент хлоробіоцину, несе КБГ нового пептидного антибіотика), *Streptomyces cyanogenus* S136 (продуцент ландоміцину А і люцензоміцину).

Метою цієї роботи стало вивчення генетичної регуляції біосинтезу пептидних та полікетидних антибіотиків, створення нових систем експресії генів в актиноміцетах-продуцентах антибіотиків і дослідження механізмів та генів стійкості до глікопептидних антибіотиків (ГПА).

Спочатку ми вивчили плейотропні регулятори біосинтезу антибіотиків у модельних стрептоміцетів *S. albus* (=*albidoflavus*) J1074 і *S. coelicolor* A3(2). Це були елементи *adpA*-опосередкованого каскаду і транскрипційні регулятори

YtrA-підродини. Гени *adpA* є ключовими глобальними регуляторами біосинтезу антибіотиків у стрептоміцетів, експресія яких залежить від лейцинової тРНК$_{UAA}$, здатної транслювати рідкісний UUA-кодон (TTA-кодон в генах *adpA*). Ми дослідили поширення ортологів AdpA серед стрептоміцетів і виявили закономірності розміщення рідкісного TTA-кодона у відповідних генах. Нам вдалося показати, що видалення олігомеризаційного GATase-1 домена білків AdpA не призводить до втрати ДНК-зв'язувальних властивостей HTH_18-домена. Дослідження гена *bldA* в J1074 (який кодує тРНК$^{ЛЕЙ}_{UAA}$ здатну до трансляції рідкісного UUA-кодона *adpA*), показали, що *bldA* не є критично-важливим для морфологічної диференціації J1074, але є вирішальним для продукції таких антибіотиків як пауломіцини. Показано роль нових компонентів *adpA*-опосередкованого каскаду J1074 – генів ферментів модифікації тРНК$_{XXA}$ *miaA* і *miaB* – в регуляції вторинного метаболізму і морфогенезу J1074. Здійснено вивчення розповсюдження і філогенії плейотропних транскрипційних регуляторів YtrA-підродини у актиноміцетів, а також експериментально досліджено роль трьох генів YtrA-регуляторів модельного об'єкта *S. coelicolor* A3(2) – *SCO0823, SCO1728* і *SCO3812*. Нокаути цих генів слабо впливали на морфогенех і продукцію антибіотиків. Нокаут *SCO3812* вів до незначного зниження продукції антибіотика актинородину, тоді як гетерологічна експресія *SCO0823* мала вплив на продукцію інших антибіотиків в гетерологічних системах.

У підсумку, нам вдалося встановити закономірності розповсюдження і еволюції AdpA-подібних білків. Ми дослідили *adpA*-опосередкований каскад в *S. albus (=albidoflavus)* J1074 і гени YtrA-подібних регуляторів у *S. coelicolor* A3(2) встановивши їх роль в продукції антибіотиків і морфогенезі.

В наступній частині роботи ми зосередилися на створенні знарядь і підходів до генно-інженерних маніпуляцій із *A. rectilineatus* NRRL B-16090 і *S. roseochromogenes* NRRL 3504. Провівши аналіз послідовностей геномів обох штамів нам вдалося виявити нові КБГ пептидних антибіотиків. Обидва штами є складними об'єктами для генно-інженерних маніпуляцій. Тому, щоби зробити

подальші дослідження нових антибіотиків можливими, ми дослідили особливості і напрацювали методики кон'югаційного перенесення плазмід в NRRL B-16090 і NRRL 3504 з клітин *Escherichia coli* ET12567 pUZ8002^{+}. Дотепер NRRL B-16090 не був об'єктом генної інженерії. У випадку NRRL B-16090 нам вдалося продемонструвати, що плазміди на основі системи інтеграції актинофага φC31, а також pSG5-реплікона, переносяться із високою частотою з *E. coli* ET12567 pUZ8002^{+}. Ми визначили оптимальні умови культивування NRRL B-16090 для отримання спорулюючих газонів цього штаму. Ми перевірили активність 5 гетерологічних промоторів у клітинах NRRL B-16090 за допомогою кількісного і якісного аналізу експресії репортерного гена *gusA,* опосередкованої β-глюкуронідазною активністю. Виявилося, що промотори *aac(3)IVp, moeE5p* і *tei2p* є ефективними знаряддями для надекспресії генів в NRRL B-16090.

Хоча *S. roseochromogenes* NRRL 3504 є відомим продуцентом амінокумаринового антибіотика хлоробіоцину, знаряддя для генно-інженерних маніпуляцій із цим штамом є вкрай обмежені. Ми створили високоефективну методику кон'югативного перенесення плазмід в NRRL 3504, використовуючи спори штаму як клітини реципієнти, а клітини *E. coli* ET12567 pUZ8002^{+} як клітини донори. Ми продемонстрували, що плазміди на основі систем інтеграції фагів φBT1 і VWB1, і плазміди на основі pSG5-реплікона можна переносити і застосовувати для експресії генів в NRRL 3504. За допомогою суїцидного вектора pKC1132 ми створили штам NRRL 3504 не здатний до продукції хлоробіоцину. Також, гетерологічні гени *adpA* і *bldA* із *S. albus* (=*albidoflavus*) J1074 виявилися ефективними знаряддями для активування вторинного метаболому NRRL 3504. Використовуючи гени шлях-специфічних регуляторів КБГ хлоробіоцину і спорідненого антибіотика новобіоцину – *cloG* і *novG* як репортери, ми вивчили ефективність ряду платформ експресії генів в NRRL 3504. Виявилося, що плазміда pGCcymRP21, в якій експресія генів є кумат-індукованою, є найефективнішою платформою експресії генів у NRRL 3504: штами із надекспресіями *cloG* і *novG* продукували до 30 разів більше

хлоробіоцину, ніж штам дикого типу – більше 200 мг/л. Це найбільші показники продукції хлоробіоцину, описані в літературі.

Наступним кроком було вивчення регуляції біосинтезу клінічно-важливого ГПА A40926 в *N. gerenzanensis* ATCC 39727. Ми дослідили активність 11 гетерологічних і нативних промоторів у клітинах *N. gerenzanensis* ATCC 39727 за допомогою репортерного гена β-глюкуронідази *gusA*. Виявилося, що *aac(3)IVp* є найефективнішим промотором для експресії генів у ATCC 39727. Аналіз геномів інших представників роду *Nonomuraea* дав змогу відкрити новий КБГ ГПА в *N. coxensis* DSM 45129. Для DSM 45129 дотепер не існували методики генно-інженерних досліджень, тому ми розробили протокол для перенесення плазмід в клітини цього штаму, використовуючи спори як клітини-реципієнти. Ми вивчили активність 11 тих самих промоторів у DSM 45129, помітивши, що *aac(3)IVp* є найактивнішим і тут.

Ми застосували розроблену систему експресії генів з метою створення нових штамів-надпродуцентів A40926. Так, ми клонували гени шлях-специфічних регуляторів КБГ A40926 (названий *dbv*) – *dbv4* і *dbv3* – під контроль *aac(3)IVp* у складі інтегративних плазмід заснованих на системі інтеграції фага φC31. Отримані рекомбінантні штами ATCC 39727 із надекспресіями цих генів продукували більше A40926 за умов культивування в колбах Ерленмеєра та ферментері. Найбільший рівень продукції A40926 – прибл. 800 мг/л – спостерігали у штамів із одночасною надекспресією обох генів *dbv3* і *dbv4* у ферментері. Контрольний штам продукував прибл. 400 мг/мл антибіотика за тих же умов.

Ми секвенували геном DSM 45129 і охарактеризували новий КБГ ГПА, який назвали *noc*. Порівнявши КБГ *noc* і *dbv* ми виявили, що *noc* є повноцінним кластером, який містить всі необхідні біосинтетичні гени, а отже найімовірніше кодує продукцію нового ГПА. Підібравши оптимальні умови культивування, нам вдалося досягти продукції ГПА в DSM 45129, виділити його і вивчити його структуру за допомогою ВЕРХ-МС і МС/МС. Виявилося, що цей ГПА – названий A50926 – є подібним за структурою до A40926, але містить

некарбоксильований Glc*N*-ацильний залишок, приєднаний до 4 амінокислоти аглікона. Нам вдалося створити надпродуцентів A50926 за допомогою надекспресії гетерологічних регуляторних генів *dbv3* і *dbv4.* На останньому етапі ми створили альтернативного продуцента A40926 на основі DSM 45129 шляхом гетерологічної експресії гена гексозооксидази *dbv29.*

Отримані результати значно розширюють набір знарядь для генно-інженерних маніпуляцій із бактеріями роду *Nonomuraea* і описують застосування цих знарядь для створення надпродуцентів клінічно-важливого антибіотика A40926. Ми відкрили новий ГПА і описати його біосинтетичний шлях в *N. coxensis* DSM 45129, вказуючи на те, що рід *Nonomuraea* залишається джерелом нових ГПА.

A40926 і A50926 є типовими ГПА IV-го типу, що разом із ГПА типів I-III пригнічують ріст Грам-позитивних бактерії зв'язуючи ліпід II. Натомість, ГПА V-го типу мають альтернативну мішень – автолізини, чим і цікаві. Першим із виявлених ГПА V-го типу є антибіотик комплестатин, що продукується *Streptomyces lavendulae* SANK 60477. Наша реконструкція мультилокусної філогенії 18 КБГ ГПА продемонструвала, що КБГ комплестатину – *com* – є зовнішньою групою до решти КБГ ГПА типів I-IV. Далі ми проаналізували 2664 геномів стрептоміцетів в пошуках КБГ, подібних на КБГ комплестатину. Ми виявили такі КБГ в 53 геномах. Більше детальний аналіз дав змогу отримати 33 повні послідовності таких КБГ. Ми створили мультилокусну філогенію 33 віднайдених КБГ і виявили 5 окремих клад. Реконструкція *in silico* біосинтетичних шляхів, кодованих в знайдених КБГ, дала змогу припустити, що КБГ деяких клад можуть кодувати біосинтез нових аналогів комплестатину. Зокрема, КБГ клад IV і II ймовірно кодують *N*-малоніл-норкомплестатин, клади III – норкомплестатин, а клади I – *N*-ацетил-норкомплестатин. Знайдені КБГ заслуговують на подальше експериментальне вивчення.

Ми вивчали нові підходи до активації мовчазних КБГ на моделі продуцента полікетидного антибіотика ландоміцину A *S. cyanogenus* S136. Спочатку ми секвенували геном цього штаму і оцінили різноманіття КБГ

антибіотиків в ньому. Дослідивши властивості нативного алеля регуляторного гена *adpA* в S136 ми виявили, що він кодує нефункціональний білок. Натомість, надекспресія гетерологічних алелів *adpA* з різних стрептоміцетів, зокрема алеля з J1074, вела до багатократного збільшення продукції ландоміцину А (в шість разів – до 250 мг/л в рідкому середовищі SG), а також активувала продукцію цього антибіотика за умов, в яких S136 ландоміцин А не продукує.

Ми виявили, що, окрім продукції ландоміцину А, введення гетерологічних алелів *adpA* має активуючий вплив на інші компоненти вторинного метаболому *S. cyanogenus* S136. Зокрема, ми спостерігали активацію накопичення забарвлених сполук, ймовірно меланінів; появу активності проти Грам-позитивних бактерій і дріжджових грибів. ВЕРХ-МС аналіз екстрактів із штамів з надекспресією гетерологічних алелів *adpA* продемонстрував появу нових мас-піків. Один із них відповідав протигрибковому полікетиду люцензоміцину. Нам вдалося підібрати оптимальні умови для продукції люцензоміцину і виділити цей антибіотик для підтвердження його структури за допомогою МС/МС. Також, ми ідентифікували і анотували в геномі S136 дотепер невідомий КБГ люцензоміцину, назвавши його *lcm*. Ми встановили позитивну для продукції люцензоміцину роль шлях-специфічних регуляторних генів *lcmRII* і *lcmRIII*, створивши інтегровану *adpA*-опосередковану модель активації цього біосинтетичного шляху.

У підсумку, наші результати свідчать про те, що для успішної активації мовчазних біосинтетичних шляхів антибіотиків необхідний комплексний підхід, який включає: застосування шлях-специфічних і глобальних регуляторів, підбір оптимальних умов культивування, елімінацію інших біосинтетичних шляхів, що можуть конкурувати за речовини-попередники.

Остання частина цієї роботи описує функції генів стійкості до тейкопланіну в продуцента цього антибіотика – *A. teichomyceticus* ATCC 31121 і особливості розповсюдження та еволюції генів стійкості до ГПА у інших бактерій. Ми створили повну модель функціонування генів стійкості до ГПА в ATCC 31121. Вони згруповані у два оперони: *tei2-3-4* (ортологи *vanRS* + ген

дигідрофолат редуктази) і *tei7-6-5* (ортологи *vanHAX*). Експресія обох оперонів протягом всього життєвого циклу рівномірно забезпечена сильними і конститутивними промоторами, незалежними від шлях-специфічної регуляції біосинтезу тейкопланіну, а Tei2 розпізнає промотор *tei7p*. Сенсорна гістидинова кіназа Tei3 володіє конститутивною фосфорилазною активністю, а її сенсорний домен розпізнає не тейкопланін, а споріднений ГПА A40926.

Ми оцінили розповсюдження і еволюцію генів стійкості до ГПА серед всіх актинобактерій: проаналізувавши 7108 геномів актинобактерій вдалося продемонстрували, що *van*-гени є поширеними у бактерій типу *Actinobacteria* (зустрічаються з частотою приблизно 13% – знайдені принаймні в 900 геномах); при цьому, КБГ ГПА типів I-IV – стійкість до яких забезпечують *van*-гени – було знайдено лише в 39 геномах. Подальший скринінг вже геномів всіх бактерій, доступних в базі даних GenBank на 2022 рік, виявив, що гени стійкості до ГПА поширилися з актинобактерій до представників класів Грам-позитивних бактерій *Bacilli*, *Clostridia*, *Erysipelotrichia* і *Ktedonobacteria*, а також до представників Грам-негативних бактерій – класу *Anaerolineae*. Аналізуючи філогенію цих генів, ми описали ряд ймовірних подій горизонтального генного переносу, а також створили картину потенційних шляхів розповсюдження генів стійкості до ГПА серед бактерій.

У підсумку, гени стійкості до ГПА в *A. teichomyceticus* ATCC 31121 виявилися своєрідним «демонстратором» стратегій, які ведуть до високої стійкості цього штаму до ГПА: цілком ймовірно, що подібні стратегії можна буде спостерігати у патогенів. З іншого боку, гени стійкості до ГПА активно мігрують між різними групами *Terrabacteria*, створюючи загрозу розповсюдження нових генів стійкості до ГПА в патогенів.

Перед людством стоїть нагальна потреба пошуку нових антибіотиків, покращення продукції нових і нагляду за генами стійкості до них. В нашій роботі ми обґрунтовуємо застосування деяких глобальних позитивних регуляторів стрептоміцетів, зокрема AdpA, для покращення продукції відомих полікетидів (на прикладі ландоміцину А) і активації нових біосинтетичних

шляхів (на прикладі люцензоміцину). Іншим результатом цієї роботи стало значне розширення арсеналу знарядь і підходів для генетичних і генно-інженерних маніпуляцій із продуцентами пептидних антибіотиків, що походять не тільки із стрептоміцетів, а й з т.зв. «рідкісних родів» актиноміцетів: це дало змогу відкрити і описати новий ГПА, а також сприяти відкриттю інших нових пептидних антибіотиків. Не менш важливими є отримані дані про регуляцію стійкості до ГПА у *A. teichomyceticus* ATCC 31121, а також про еволюцію і поширення цих генів.

Ключові слова: актиноміцети, глікопептидні антибіотики, полікетидні антибіотики, регуляторні гени, гени стійкості до антибіотиків, тейкопланін, A40926, ландоміцин, люцензоміцин.

SUMMARY

Yushchuk O. Genetic mechanisms regulating the biosynthesis of peptide and polyketide antibiotics. – Qualifying scientific work in the form of a scientific report based on research articles.

Thesis for a degree of Doctor of Biological Sciences (D.Sc.) in Biology, specialty 03.00.22 – molecular genetics (09 – Biology). – Ivan Franko National University of Lviv, Lviv. – Institute of Food Biotechnology and Genomics of the National Academy of Sciences of Ukraine, Kyiv, 2023.

The emergence and spread of drug-resistant microorganisms critically endanger contemporary healthcare systems throughout the globe, taking millions of lives every year. In this light, the search for novel antimicrobials and continuous improvement of the production of known antibiotics are crucial. Surveillance of the antibiotic resistance genes spread and investigation of resistance mechanisms are important as well. The majority of the natural antibiotics active against Gram-positive, Gram-negative bacteria, and fungi comes from certain groups of actinobacteria, widely known under the non-taxonomic name of actinomycetes. Actinomycetes are soil-dwelling Gram-positive microorganisms existing in form of mycelium and able to

produce spores. Sizeable genomes of actinomycetes carry dozens of biosynthetic gene clusters (BGCs) and are a potential source of novel antibiotics against multidrug-resistant pathogens. The most important aspect of studying antibiotic biosynthesis is the investigation of regulatory mechanisms governing the biosynthesis of these compounds. Indeed, the activity and properties of regulatory genes determine how and under what conditions certain antibiotic would be produced. This also involves continuous screening of actinobacterial genomes *in silico* to identify novel BGCs, identification of pathway-specific and global regulatory genes, and their application to create overproducer strains. Considering BGCs coming from one organism are mostly silent, applying regulatory genes is often a viable way to activate them. Successful achievement of all these goals depends on the development of efficient methods for gene-engineering of antibiotic producer strains.

Peptide and polyketide antibiotics are among the most important antibiotics coming from actinomycetes. These compounds have extremely variable chemical structures due to the modularity of key biosynthetic enzymes – non-ribosomal peptide synthases and polyketide synthases. Some peptide and polyketide antibiotics belong to the World Health Organization lists of critically-important antimicrobials. These are, in particular, glycopeptide antibiotics (GPAs), *e.g.* vancomycin and teicoplanin; moreover, they were identified as the highest-priority antimicrobials for investigation and resistance surveillance. This work was focused on following actinomycetes able to produce peptide, polyketide, and some other antibiotics: *Actinoplanes teichomyceticus* ATCC 31121 (teicoplanin producer), *Actinoplanes rectilineatus* NRRL B-16090 (carries GPA BGC), *Nonomuraea gerenzanensis* ATCC 39727 (producer of A40926), *Nonomuraea coxensis* DSM 45129 (producer of A50926), *Streptomyces roseochromogenes* NRRL 3504 (clorobiocin producer, carries BGC of novel peptide antibiotic), *Streptomyces cyanogenus* S136 (landomycin A and lucensomycin producer).

The main goal of this work was to elucidate mechanisms of the genetic control behind the biosynthesis of peptide and polyketide antibiotics, design of novel gene-

expression systems for producer strains, and investigate functions and distribution of certain antibiotic-resistance genes.

First, we have investigated some pleiotropic regulators in model actinomycetes *S. albus* (=*albidoflavus*) J1074 and *S. coelicolor* A3(2). These were elements of the *adpA*-mediated regulatory cascade and transcriptional regulators of the YtrA-subfamily. Notably, *adpA* regulatory genes are the most important regulators of antibiotic biosynthesis in streptomycetes. We assessed the distribution of *adpA* orthologues among different streptomycetes and characterized the occurrence of rare TTA codons therein. Investigation *in vivo* allowed us to demonstrate that the oligomerization GATase-1 domain is not crucial for the proper functioning of AdpA. Studying the regulation of *adpA* on a translational level, we investigated the *bldA* gene (coding for tRNA able to decode rare UUA codon) in J1074. We demonstrate that *bldA* is not crucial for the morphogenesis of J1074, although the production of paulomycins depends on *bldA.* Two novel players of the *adpA*-mediated regulatory cascade were investigated and studied: these were genes for $tRNA_{XXA}$ modification enzymes MiaA and MiaB. Finally, we studied the spread and phylogeny of YtrA-like transcriptional regulators among actinobacteria. Three genes of YtrA-like regulators were studied in the *S. coelicolor* A3(2) model: *SCO0823, SCO1728,* and *SCO3812.* It appeared that these genes do not play a decisive role in antibiotic production and morphogenesis. In particular knockout of *SCO3812* led to a slight decrease in actinorhodin production, while heterologous expression of *SCO0823* influenced antibiotic production in some other streptomycetes.

Summarizing, herein we update the picture describing the distribution and evolution of AdpA-like regulators; we investigate *adpA*-mediated regulatory cascade in *S. albus (=albidoflavus)* J1074, introducing some new players (in particular, genes of the tRNA modification enzymes) and showing their role in antibiotic production and morphogenesis. The exact roles of YtrA-like regulators in *S. coelicolor* A3(2) were elusive under our experimental conditions, although gene deletions seemed to slightly impact antibiotic production and morphogenesis; this group of regulators indeed merits further investigations.

The next part of the work was dedicated to the design of tools and methods for gene-engineering of *A. rectilineatus* NRRL B-16090 and *S. roseochromogenes* NRRL 3504. Genomics analysis revealed both strains to be potential producers of novel peptide antibiotics. Since both seem to be challenging objects for gene-engineering manipulations, we had to develop efficient and straightforward methods to manipulate these strains. In the case of NRRL B-16090, we were able to create a protocol for conjugal transfers of plasmids based on the integration system of φC31 phage and pSG5-replicon. Unfortunately, we were unable to transfer φBT1- and VWB1-based plasmids into NRRL B-16090 under tested conditions. A prerequisite for developing the protocol for conjugal transfer of plasmids into NRRL B-16090 was finding optimal conditions for the sporulation of the strain. We also tested the activity of 5 heterologous promoters in NRRL B-16090 via qualitative and quantitative analysis of β-glucuronidase activity. Three of them, particularly *aac(3)IVp*, *moeE5p*, and *tei2p* appeared to be efficient tools for constitutive gene overexpression in NRRL B-16090.

Although *S. roseochromogenes* NRRL 3504 is a known producer of aminocoumarine antibiotic clorobiocin, the toolkit for genetic manipulations of this strain is limited. To improve this situation, we created a simple and efficient protocol for conjugal transfer using spores of NRRL 3504 as recipient cells. Using this protocol, we were able to successfully transfer a set of additional actinophage (φBT1 and VWB1) based plasmids, as well as plasmid derived from pSG5 replicon, into NRRL 3504. We also utilized suicide vector pKC1132 to perform insertional knockout of the clorobiocin biosynthetic cluster, generating a clorobiocin non-producing strain. Heterologous *adpA* and *bldA* genes from J1074 appeared to be an efficient tool to activate the secondary metabolome of NRRL 3504. Finally, we utilized pathway-specific regulators of clorobiocin biosynthesis – genes *cloG* and *novG* – as reporters to investigate the efficiency of a set of gene-expression platforms for NRRL 3504. pGUScymRP21 plasmid, having *cloG* and *novG* under the control of cumate-inducible promotor, appeared to be the most efficient tool for gene overexpression in NRRL 3504. Indeed, pGUScymRP21-mediated overexpression of

cloG and *novG* led to a dramatic increase in clorobiocin production reaching more than 200 mg/L (while the wild type produced less than 10 mg/L).

In the next step, we investigated the activity of 11 heterologous and native promoters in the producer of clinically-relevant GPA A40926 – *N. gerenzanensis* ATCC 39727 – using β-glucuronidase reporter gene *gusA*. We revealed *aac(3)IVp* as the most active promoter in ATCC 39727. Genome analysis made on other *Nonomuraea* spp. discovered putative novel GPA BGC in *N. coxensis* DSM 45129. Latter strain was never an object for gene engineering before, hence we have developed an efficient protocol to transfer plasmids into DSM 45129. Consequently, we investigated the activity of the same 11 promoters in DSM 45129 and observed that *aac(3)IVp* is the most active here as well.

We applied obtained information to create novel A40926 overproducers in ATCC 39727. To that goal, we have cloned pathway-specific regulatory genes of A40926 BGC (known as *dbv*) – *dbv3* and *dbv4* – under the control of *aac(3)IVp*. Generated recombinants tended to produce more A40926 than the wild type in Erlenmeyer flasks and bioreactors. The best A40926 production level was achieved in the strain overexpressing *dbv3* and *dbv4* simultaneously. This recombinant produced *ca.* 800 mg/L of the A40926 in the bioreactor, while the wild type accumulated only *ca.* 400 mg/L.

Returning to DSM 45129, we sequenced and completely assembled the genome of this strain, allowing us to identify and characterize novel GPA BGC, named *noc*. Comparing *noc* to *dbv* we realized that *noc* carries all genes necessary for the production of novel GPA. Therefore, we have identified the optimal conditions for GPA production in DSM 45129 and isolated this compound. Further chemical analysis by means of the HPLC-MS and MS/MS allowed concluding that this novel GPA – named A50926 – possesses a structure similar to A40926, although lacking carboxylic group at Glc*N*-acyl moiety attached to the amino acid 4 of the aglycone. Applying our experience from ATCC 39727, we utilized heterologous regulatory genes *dbv3* and *dbv4* to activate or increase A50926 production in DSM 45129.

Lastly, we have sculpted DSM 45129 into the alternative A40926 producer by expressing *dbv29* – a hexose oxidase gene.

Obtained results widen our arsenal of genetic tools for manipulating bacteria from the *Nonomuraea* genus and the application of those tools to create novel strains overproducing clinically-relevant GPA A40926. Moreover, we discovered a novel GPA from *N. coxensis* DSM 45129 and described its biosynthetic pathway, indicating that the genus *Nonomuraea* remains a promising source of novel GPAs.

A40926 and A50926 are representatives of type IV GPAs, which, as well as GPAs from types I-III inhibit cell-wall biosynthesis in Gram-positive bacteria via binding to the lipid II murein intermediate. Instead, GPAs of type V possess another molecular target – autolysins. An archetypical representative of type V GPAs is complestatin produced by *S. lavendulae* SANK 60477. We reconstructed the multi-locus phylogeny of 18 GPA BGCs, including complestatin BGC – *com* – showing that the latter significantly diverges from type I-IV BGCs. Further on, we have screened genomes of 2664 *Streptomyces* spp. and identified *com*-like BGCs in 53 genomes. The more detailed investigation allowed for retrieving 33 complete sequences of *com*-like BGCs. We created multilocus phylogeny of these 33 BGCs and identified 5 separate clades. When encoded biosynthetic pathways were reconstructed *in silico*, we observed that clade IV and II BGCs potentially encode the production of *N*-malonyl-norcomplestatin, clade III – norcomplestatin, while BGCs from clade I might be coding biosynthesis of *N*-acetyl-norcomplestatin. Discovered BGCs merit further experimental investigation.

We investigated novel approaches to activate silent BGCs using *S. cyanogenus* S136 as a model. Latter is a producer of polyketide antibiotic landomycin A. Initially we sequenced and assembled the genome of S136, assisting in further assessment of the antibiotic BGCs carried by S136. We have further studied the properties of the *adpA* allele of this strain, showing that it is probably coding for a non-functional protein. Indeed, the expression of heterologous *adpA*-alleles led to a several-fold increase in landomycin A production (up to 250 mg/L in

liquid SG medium). In addition, heterologous *adpA*-alleles triggered landomycin A production under conditions where production is not observed in the wild-type strain.

adpA-induced activation of the S136 secondary metabolome was not limited to landomycin A. In particular, we also observed activation of the production of several other secondary metabolites. One of them was identified as an antifungal polyene polyketide lucensomycin by means of LC-MS and MS/MS. We were able to identify optimal cultivation conditions for lucensomycin production. For the first time, we identified and characterized lucensomycin BGC named *lcm.* Studying the roles of *lcmRII* and *lcmRIII* pathway-specific regulators allowed us to deduce a model of *adpA*-mediated activation of this biosynthetic pathway.

Overall, our results indicate that successful activation of silent biosynthetic pathways requires a complex approach including the application of pathway-specific and global regulators, identification of the optimal cultivation conditions, and elimination of other competing pathways.

The last part of the work was dedicated to the investigation of GPA-resistance genes. First, we have studied GPA-resistance genes in teicoplanin producer *A. teichomyceticus* ATCC 31121. Extensive *in vivo* and *in silico* analysis allowed the creation of a detailed model of teicoplanin resistance in ATCC 31121. GPA-resistance genes in ATCC 31121 are organized into two operons: *tei2-3-4* (*vanRS*-orthologues + dihydrofolate reductase gene) and *tei7-6-5* (*vanHAX*-orthologues). Expression of both operons is stable throughout the whole lifecycle due to the activity of strong constitutive promoters, which are independent of pathway-specific regulation of teicoplanin biosynthesis. The Tei2 response regulator was shown to bind *tei7p in vivo.* By turn, sensor histidine kinase Tei3 exhibited constitutive phosphorylase activity. Paradoxically, the sensor domain of Tei3 was able to recognize not teicoplanin, but A40926. Going laterally, we also assessed the occurrence and evolution of GPA-resistance genes across actinobacteria. We have analyzed 7108 actinobacterial genomes and found GPA-resistance genes in *ca.* 13% of them. At the same time, we found BGCs for type I-IV GPAs only in 39 genomes. Further screening made in other bacterial lineages allowed to map the further spread

of GPA-resistance genes into other classes of Gram-positive bacteria – *Bacilli*, *Clostridia*, *Erysipelotrichia*, and *Ktedonobacteria*, as well as one class of Gram-negatives – *Anaerolineae*. Our phylogenetic analysis revealed several possible events of horizontal gene transfer, driving the spread of GPA-resistance genes from *Actinobacteria.*

In summary, GPA resistance genes in *A. teichomyceticus* ATCC 31121 appeared to be a sort of a "demonstration stand" showing multiple strategies this strain utilized to become GPA-resistant: it is likely that at least some of these strategies might be found in pathogens evolving GPA resistance. We also found evidence of constant migration of *van*-genes between different groups of *Terrabacteria,* which creates a constant threat of *van*-genes transmission to pathogens.

To sustain the continuous "arms race" between pathogens and mankind, we critically require to discover novel antibiotics, improve the production of used ones and surveil the spread of antibiotic resistance genes. In this work we demonstrate the potential of certain regulatory genes, for instance, *adpA*, to improve and activate the production of valuable antibiotics. The other outcome of this work was a significant improvement of our genetic toolkit for manipulations with the producers of peptide antibiotics, especially those coming from so-called rare genera (like *Actinoplanes* and *Nonomuraea*). This allowed us to discover and describe novel GPA and pave new venues for the discoveries of other peptide antibiotics. Of the same importance were our results about the regulation of GPA resistance genes in *A. teichomyceticus* ATCC 31121, as well as about the evolution and spread of *van*-genes.

Key words: actinomycetes, glycopeptide antibiotics, polyketide antibiotics, regulatory genes, antibiotic resistance genes, teicoplanin, A40926, landomycin, lucensomycin.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

Наукові праці, що розкривають основні наукові результати дисертації:

1. **Yushchuk O,** Zhukrovska K, Ostash B, Fedorenko V, Marinelli F (2022) Heterologous expression reveals ancient properties of Tei3 – a VanS ortholog from the teicoplanin producer *Actinoplanes teichomyceticus*. International journal of molecular sciences 23(24):15713. https://doi.org/10.3390/ijms232415713 **[IF 6,208; SCOPUS, WoS; SCImago Q2 – 2022 Molecular biology]** (*Особистий внесок здобувача: концептуалізація, планування і виконання експериментів, написання рукопису та підготовка статті до друку*).
2. **Yushchuk O,** Ostash B (2022) Glycopeptide antibiotics: genetics, chemistry, and new screening approaches. In: Rai, R.V., Bai, J.A. (eds.) Natural products from Actinomycetes. Springer, Singapore. https://doi.org/10.1007/978-981-16-6132-7_16 [**розділ монографії**] (*Особистий внесок здобувача: аналіз літератури, написання розділу, створення ілюстрацій, підготовка розділу до друку*).
3. **Ющук О,** Жукровська К, Федоренко В (2022) Філогенія білків-експортерів глікопептидних і деяких споріднених антибіотиків. Вісник Львівського університету. Серія біологічна 86:33-46. http://dx.doi.org/10.30970/vlubs.2022.86.03 [**фахове видання України**] (*Особистий внесок здобувача: концептуалізація, скринінг геномів і біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).
4. **Yushchuk O,** Zhukrovska K, Berini F, Fedorenko V, Marinelli F (2022) Genetics behind the glycosylation patterns in the biosynthesis of dalbaheptides. Frontiers in chemistry 10:858708. https://doi.org/10.3389/fchem.2022.858708 **[IF 5,545; SCOPUS, WoS; SCImago Q1 – 2022 Chemistry]** (*Особистий внесок здобувача: аналіз літератури і написання огляду, створення ілюстрацій, підготовка статті до друку*).

5. **Ющук О,** Жукровська К, Федоренко В (2022) Поширення кластерів генів біосинтезу комлестатину та подібних сполук у представників роду *Streptomyces*. Фактори експериментальної еволюції організмів 30:133-140. https://doi.org/10.7124/FEEO.v30.1474 [**фахове видання України**] (*Особистий внесок здобувача: концептуалізація, скринінг геномів і біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).

6. **Yushchuk O,** Binda E, Fedorenko V, Marinelli F (2022) Occurrence of *vanHAX* and related genes beyond the *Actinobacteria* phylum. Genes 13(11):1960. https://doi.org/10.3390/genes13111960 **[IF 4,141; SCOPUS, WoS; SCImago Q2 – 2022 Genetics]** (*Особистий внесок здобувача: концептуалізація, скринінг геномів і біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).

7. Melnyk S, Stepanyshyn A, **Yushchuk O,** Mandler M, Ostash I, Koshla O, Fedorenko V, Kahne D, Ostash B (2022) Genetic approaches to improve clorobiocin production in *Streptomyces roseochromogenes* NRRL 3504. Applied microbiology and biotechnology 106(4):1543–1556. https://doi.org/10.1007/s00253-022-11814-4 **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2022 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: участь в концептуалізації, розробка методики перенесення плазмід в NRRL 3504, аналіз продукції хлоробіоцину в рекомбінантних штамів, аналіз результатів ВЕРХ, написання рукопису та підготовка статті до друку*).

8. Andreo-Vidal A, Binda E, Fedorenko V, Marinelli F, **Yushchuk O** (2021) Genomic insights into the distribution and phylogeny of glycopeptide resistance determinants within the *Actinobacteria* phylum. Antibiotics (Basel, Switzerland) 10(12):1533. https://doi.org/10.3390/antibiotics10121533 **[IF 5,222; SCOPUS, WoS; SCImago Q2 – 2021 Infectious Diseases]** (*Особистий внесок здобувача: концептуалізація, скринінг геномів і біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).

9. Hrab P, Rückert C, Busche T, Ostash I, Kalinowski J, Fedorenko V, **Yushchuk O,** Ostash B (2021) Complete genome sequence of *Streptomyces cyanogenus* S136,

producer of anticancer angucycline landomycin A. 3 Biotech 11(6):282. https://doi.org/10.1007/s13205-021-02834-4 **[IF 2,893; SCOPUS, WoS; SCImago Q2 – 2021 Biotechnology]** (*Особистий внесок здобувача: концептуалізація і планування дослідження, верифікація результатів аналізу in silico, участь в підготовці рукопису до друку*).

10. **Yushchuk O,** Vior NM, Andreo-Vidal A, Berini F, Rückert C, Busche T, Binda E, Kalinowski J, Truman AW, Marinelli F (2021) Genomic-led discovery of a novel glycopeptide antibiotic by *Nonomuraea coxensis* DSM 45129. ACS Chemical Biology 16(5):915–928. https://doi.org/10.1021/acschembio.1c00170 **[IF 4,634; SCOPUS, WoS; SCImago Q1 – 2021 Biochemistry]** (*Особистий внесок здобувача: концептуалізація, аналіз in silico, характеристика кластера, виконання всіх генно-інженерних робіт, пошук оптимальних умов культивування DSM 45129 для продукції антибіотика, ферментація DSM 45129 і похідних, аналіз і інтерпретація даних ВЕРХ, написання рукопису та підготовка статті до друку*).

11. **Yushchuk O,** Ostash I, Mösker E, Vlasiuk I, Deneka M, Rückert C, Busche T, Fedorenko V, Kalinowski J, Süssmuth RD, Ostash B (2021) Eliciting the silent lucensomycin biosynthetic pathway in *Streptomyces cyanogenus* S136 via manipulation of the global regulatory gene *adpA*. Scientific reports 11(1):3507. https://doi.org/10.1038/s41598-021-82934-6 **[IF 4,996; SCOPUS, WoS; SCImago Q1 – 2021 Multidisciplinary]** (*Особистий внесок здобувача: концептуалізація, аналіз in silico, характеристика кластера, виконання всіх генно-інженерних робіт і аналізу отриманих рекомбінантів, пошук оптимальних умов культивування для продукції люцензоміцину, написання рукопису та підготовка статті до друку*).

12. **Yushchuk O,** Homoniuk V, Datsiuk Y, Ostash B, Marinelli F, Fedorenko V (2020) Development of a gene expression system for the uncommon actinomycete *Actinoplanes rectilineatus* NRRL B-16090. Journal of applied genetics 61(1):141–149. https://doi.org/10.1007/s13353-019-00534-7 **[IF 2,653; SCOPUS, WoS;**

SCImago Q2 – 2020 Genetics] (*Особистий внесок здобувача: концептуалізація, виконання досліджень, написання рукопису та підготовка статті до друку*).

13. **Yushchuk O,** Homoniuk V, Ostash B, Marinelli F, Fedorenko V (2020) Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. The Journal of antibiotics 73(4):255–259. https://doi.org/10.1038/s41429-019-0274-9 **[IF 3,424; SCOPUS, WoS; SCImago Q3 – 2020 Pharmacology]** (*Особистий внесок здобувача: концептуалізація, виконання досліджень, написання рукопису та підготовка статті до друку*).

14. **Yushchuk O,** Andreo-Vidal A, Marcone GL, Bibb M, Marinelli F, Binda E (2020) New molecular tools for regulation and improvement of A40926 glycopeptide antibiotic production in *Nonomuraea gerenzanensis* ATCC 39727 Frontiers in microbiology 11:8. https://doi.org/10.3389/fmicb.2020.00008 **[IF 6,064; SCOPUS, WoS; SCImago Q1 – 2020 Microbiology]** (*Особистий внесок здобувача: концептуалізація, виконання досліджень, написання рукопису та підготовка статті до друку*).

15. **Yushchuk O,** Binda E, Marinelli F (2020) Glycopeptide antibiotic resistance genes: distribution and function in the producer actinomycetes. Frontiers in microbiology 11:1173. https://doi.org/10.3389/fmicb.2020.01173 **[IF 6,064; SCOPUS, WoS; SCImago Q1 – 2020 Microbiology]** (*Особистий внесок здобувача: аналіз літератури і написання розділів присвячених продуцентам глікопептидів, створення ілюстрацій, підготовка статті до друку*).

16. **Yushchuk O,** Ostash B, Truman AW, Marinelli F, Fedorenko V (2020) Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes. Applied microbiology and biotechnology 104(8):3279–3291. https://doi.org/10.1007/s00253-020-10436-y **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2020 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: аналіз літератури і написання огляду, створення ілюстрацій, підготовка статті до друку*).

17. Koshla O, **Yushchuk O,** Ostash I, Dacyuk Y, Myronovskyi M, Jäger G, Süssmuth RD, Luzhetskyy A, Byström A, Kirsebom LA, Ostash B (2019) Gene *miaA*

for post-transcriptional modification of $tRNA_{XXA}$ is important for morphological and metabolic differentiation in *Streptomyces*. Molecular microbiology 112(1):249–265. https://doi.org/10.1111/mmi.14266 **[IF 3,979; SCOPUS, WoS; SCImago Q1 – 2019 Molecular biology]** (*Особистий внесок здобувача: створення репортерної системи на основі плазміди pSAGA, створення подвійного мутанта ΔmiaA ΔbldA, створення плазмід із різними алелями гена adpA, комплементація мутантів ΔmiaA і ΔmiaA ΔbldA різними алелями adpA, аналіз властивостей мутантів і рекомбінантів, участь в написанні рукопису і підготовці статті до друку*).

18. **Yushchuk O,** Kharel M, Ostash I, Ostash B (2019) Landomycin biosynthesis and its regulation in *Streptomyces*. Applied microbiology and biotechnology 103(4):1659–1665. https://doi.org/10.1007/s00253-018-09601-1 **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2019 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: аналіз літератури і розділів огляду, присвячених організації кластерів генів біосинтезу і регуляції продукції ландоміцинів, створення ілюстрацій, підготовка статті до друку*).

19. **Yushchuk O,** Horbal L, Ostash B, Marinelli F, Wohlleben W, Stegmann E, Fedorenko V (2019) Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. Applied microbiology and biotechnology 103(10):4089–4102. https://doi.org/10.1007/s00253-019-09789-w **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2019 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: концептуалізація, планування і виконання експериментів, написання рукопису та підготовка статті до друку*).

20. **Yushchuk O,** Ostash I, Vlasiuk I, Gren T, Luzhetskyy A, Kalinowski J, Fedorenko V, Ostash B (2018) Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions. Applied microbiology and biotechnology 102(19):8419–8428. https://doi.org/10.1007/s00253-018-9249-1 **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2018 Applied Microbiology and Biotechnology]** (*Особистий внесок*

здобувача: концептуалізація, планування і виконання експериментів, написання рукопису та підготовка статті до друку).

21. Rabyk M, **Yushchuk O,** Rokytskyy I, Anisimova M, Ostash B (2018) Genomic insights into evolution of adpa family master regulators of morphological differentiation and secondary metabolism in *Streptomyces*. Journal of molecular evolution 86(3-4):204–215. https://doi.org/10.1007/s00239-018-9834-z **[IF 3,973; SCOPUS, WoS; SCImago Q2 – 2018 Genetics]** (*Особистий внесок здобувача: участь в аналізі даних, виконання експериментальної частини, участь в написанні рукопису та підготовці статті до друку*).

22. Koshla O, Lopatniuk M, Rokytskyy I, **Yushchuk O,** Dacyuk Y, Fedorenko V, Luzhetskyy A, Ostash B (2017) Properties of *Streptomyces albus* J1074 mutant deficient in $tRNA^{Leu}_{UAA}$ gene *bldA*. Archives of microbiology 199(8):1175–1183. https://doi.org/10.1007/s00203-017-1389-7 **[IF 2,667; SCOPUS, WoS; SCImago Q3 – 2017 Genetics]** (*Особистий внесок здобувача: вивчення особливостей морфології і продукції антибіотиків в ΔbldA мутанта, участь в написанні рукопису та підготовці статті до друку*).

23. **Ющук ОС,** Осташ БО, Горбаль ЛО, Федоренко ВО (2017). Реконструкція філогенії кластерів генів біосинтезу глікопептидів. Фактори експериментальної еволюції організмів 20:109-115. [**фахове видання України**] (*Особистий внесок здобувача: концептуалізація, біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).

24. Tsypik O, **Yushchuk O,** Zaburannyi N, Flärdh K, Walker S, Fedorenko V, Ostash B (2016) Transcriptional regulators of GntR family in *Streptomyces coelicolor* A3(2): analysis *in silico* and *in vivo* of YtrA subfamily. Folia microbiologica 61(3):209–220. https://doi.org/10.1007/s12223-015-0426-7 **[IF 2,629; SCOPUS, WoS; SCImago Q3 – 2016 Microbiology]** (*Особистий внесок здобувача: вивчення розповсюдження генів YtrA-регуляторів серед актинобактерій, аналіз властивостей генів SCO0823 і SCO3812, участь в написанні рукопису та підготовці статті до друку*).

Патенти:

1. Патент України на винахід №119472 Україна, МПК (2006.01)C12P 1/06 C12P 19/64 C12N 15/80 C12R 1/465 Спосіб активування продукції сполук з антибіотичною та протигрибною властивостями у *Streptomyces cyanogenus* S136 / **Ющук О.,** Осташ Б.; заявник і власник Львівський національний ун-т ім. І. Франка. – № a201702642, заявл. 21 березня 2017 р., опубл. 25.06.2019, Бюл. № 12. [**патент України на винахід**] (*Особистий внесок здобувача: виконання експериментів, написання і підготовка тексту патенту*).

Публікації, що засвідчують апробацію матеріалів дисертації:

1. Stepanyshyn A., **Yushchuk O.,** Marinelli F., Kalinowski J., Fedorenko V. Expression of pathway-specific regulators from dalbaheptide biosynthetic gene clusters induces antimicrobial properties of *Amycolatopsis bartoniae* DSM 45807. XVIII international scientific conference for students and PhD students dedicated to the 195th anniversary from the birthday of Julius Planer «Youth and progress of biology»: abstracts, Lviv, 6–7 September 2022. Львів: Сполом, 2022. P. 72-73. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).
2. Duban R., Stepanyshyn A., Melnyk S., Fedorenko V., Ostash B., **Yushchuk O.** Approaching the activation of salinamide-like biosynthetic gene cluster in clorobiocin producer *Streptomyces roseochromogenes* NRRL 3504. IV International Scientific Conference: «Microbiology and Immunology – the Development»: book of abstracts, Kyiv, 22-23 September 2022. Kyiv: Taras Shevchenko National University of Kyiv, 2022. P. 65. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

3. Stepanyshyn A., Melnyk S., **Yushchuk O.,** Ostash I., Koshla O., Fedorenko V., Ostash B. Developing genetic instruments for manipulating *Streptomyces roseochromogenes* NRRL 3504 genome for increasing clorobiocin production. IV International Scientific Conference: «Microbiology and Immunology – the Development»: book of abstracts, Kyiv, 22-23 September 2022. Kyiv: Taras Shevchenko National University of Kyiv, 2022. P. 81. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

4. Andreo-Vidal A., **Yushchuk O.,** Berini F., Binda E., Marinelli F. Novel glycopeptide antibiotic by *Nonomuraea coxensis* DSM 45129. Online congress ExpoBAC Zaragoza 2020: abstracts, Zaragoza, 8-9 July 2021. Zaragoza: Spain, 2021. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

5. Andreo-Vidal A., Binda E., Marinelli F., **Yushchuk O.** Comparative genomics of *van*-genes in glycopeptide producing *Actinobacteria* and beyond. VAAM Workshop 2021: Abstract book, Tübingen, 22-23 September 2021. Tübingen: Germany, 2021. P. 49. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач здійснив онлайн-доповідь*).

6. Степанишин А., Мельник С., Осташ Б., **Ющук О.** Аналіз мовчазного кластера генів біосинтезу криптичного салінамід-подібного антибіотика у продуцента хлоробіоцину *S. roseochromogenes*. XVII Міжнародна конференція студентів і аспірантів «Молодь і поступ біології»: тези доповідей, Львів, 27–29 квітня 2021 р. Львів, 2021. С. 114-115. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

7. **Yushchuk O.,** Ostash I., Melnyk S., Kalinowski J., Süssmuth R., Ostash B. Revealing cryptic secondary metabolome of *Streptomyces* via manipulation of global regulatory gene *adpA*. 3rd International conference on natural products discovery and

development in the genomic era: abstracts, San Diego, 13 January 2020. https://sim.confex.com/sim/np2020/meetingapp.cgi/Paper/40268. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом*).

8. **Ющук О.,** Жукровська К.-О., Федоренко В. Сенсорний домен гістидинової кінази Tei3 не здатний реагувати на глікопептидні антибіотики. XV Міжнародна наукова конференція студентів і аспірантів «Молодь та поступ біології», присвячена 135 річниці від дня народження Я. Парнаса: тези доповідей, Львів, 9-11 квітня 2019 р. Львів, 2019. С. 37. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції і зробив усну доповідь*).

9. **Ющук О.,** Власюк І., Осташ І., Федоренко В., Осташ Б. Властивості базальних представників ортологічної групи AdpA-регуляторів. XV Міжнародна наукова конференція студентів і аспірантів «Молодь та поступ біології», присвячена 135 річниці від дня народження Я. Парнаса: тези доповідей, Львів, 9-11 квітня 2019 р. Львів, 2019. С. 36. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

10. Homoniuk V., **Yushchuk O.,** Ostash B., Fedorenko V. Pathway-specific regulator of teicoplanin biosynthesis *tei15** does not interact with *tei2* promotor in heterologous system. XIV International scientific conference of students and PhD students «Youth and progress of biology»: abstracts, Lviv, 10-12 April 2018. Lviv, 2018. P. 128-129. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції і зробив усну доповідь*).

11. Vlasiuk I., **Yushchuk O.,** Ostash B., Fedorenko B. Landomycin A production of AdpA-overexpressing *S. cyanogenus* S136 strains in extremely depleted media. XIV International scientific conference of students and PhD students «Youth and progress of biology»: abstracts, Lviv, 10-12 April 2018. Lviv, 2018. P. 132.

(*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

12. Vlasiuk I., Ostash B., **Yushchuk O.,** Ostash I., Kuzhyk Y., Melnyk S., Hrab P., Suessmuth R., Fedorenko V. Pleiotropic transcriptional factor AdpA as a tool to upregulate antibiotic production by *Streptomyces.* International Conference «Advances in microbiology and biotechnology»: abstract book, Lviv, 29-31 October 2018. Lviv, 2018. P. 44. (*Особистий внесок здобувача: тези і доповідь частково засновані на результатах отриманих здобувачем або під його керівництвом*).

13. Vlasiuk I., **Yushchuk O.** Heterologous AdpA activate cryptic lucensomycin biosynthetic gene cluster in *Streptomyces cyanogenus* S136. XIII international scientific conference of young scientists «Biology: from molecule to biosphere»: abstracts, Kharkiv, 28-30 November 2018. Kharkiv, 2018. P. 82-83. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

14. Жукровська К.-О., **Ющук О.,** Федоренко В. Дослідження індуцибельності сенсорної гістидинової кінази *tei3* з *Actinoplanes teichomyceticus.* XIV Міжнародна наукова конференція студентів і аспірантів «Молодь та поступ біології», присвячена 185 річниці від дня народження Б. Дибовського: тези доповідей, Львів, 10-12 квітня 2018. Львів, 2018. С. 133. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

15. Koshla O., **Yushchuk O.,** Ostash I, \ Kravets K., Suessmuth R., Jäger G., Bystrom A., Ostash B., Fedorenko V. Post-transcriptional modifications of tRNA modulate secondary metabolism of *Streptomyces*. International Conference «Advances in microbiology and biotechnology»: abstract book, Lviv, 29-31 October 2018. Lviv, 2018. P. 75. (*Особистий внесок здобувача: тези і доповідь частково засновані на результатах отриманих здобувачем або під його керівництвом*).

16. Vlasiuk I., **Yushchuk O.** Expression of different *adpA* genes increases landomycin A production in *Streptomyces cyanogenus* S136 // XII international scientific conference of young scientists «Biology: from molecule to biosphere»: abstracts, Kharkiv, 29 November – 1 December 2017. Kharkiv, 2017. P. 68. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції і зробив усну доповідь*).

17. Гомонюк В.В., **Ющук О.С.**, Горбаль Л.О. Експресія генів стійкості до глікопептидів у продуцента тейкопланіну *Actinoplanes teichomyceticus.* XII Міжнародна конференція молодих науковців «Біологія: від молекули до біосфери»: тези доповідей, Харків, 29 листопада – 1 грудня 2017 р. Харків, 2017. С. 64. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

18. Zhukrovska K.-O., **Yushchuk O.** Expression of *tei3* gene encoding VanS-like sensor histidine kinase from teicoplanin biosynthesis gene cluster leads to teicoplanin resistance in *S. coelicolor* M512. XII international scientific conference of young scientists «Biology: from molecule to biosphere»: abstracts, Kharkiv, 29 November – 1 December 2017. Kharkiv, 2017. P. 69. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

19. **Yushchuk O.** Evolution of glycopeptide and moenomycin antibiotic biosynthesis pathways in bacteria. International conference «Integrative Biology & Medicine»: abstracts, Kyiv, 2-7 October, 2017. Kyiv, 2017. P. 75. (*Особистий внесок здобувача: тези і доповідь засновані на результатах здобувача, здобувач здійснив усну доповідь*).

ЗМІСТ

Ст.

Перелік умовних скорочень…….. 1

Вступ…….. 2

РОЗДІЛ 1 Огляд літератури…….. 12

РОЗДІЛ 2 Матеріали і методи досліджень…….. 88

2.1 Матеріали досліджень……. 88

2.2 Методи досліджень……. 88

2.2.1 Мікробіологічні методи досліджень……... 88

2.2.2 Культивування штамів актиноміцетів задля продукції антибіотиків…… 90

2.2.3 Генетичні і генно-інженерні методи досліджень…….. 92

2.2.4 Біохімічні методи досліджень……. 97

2.2.5 Біоінформатичні методи досліджень……..... 99

РОЗДІЛ 3 Плейотропні регулятори морфогенезу і вторинного метаболізму стрептоміцетів, що належать до *adpA*-опосередкованого каскаду і YtrA-підродини, як знаряддя для контролю продукції антибіотиків……... 101

РОЗДІЛ 4 Підходи до генно-інженерних маніпуляцій із *Actinoplanes rectilineatus* NRRL B-16090 і *Streptomyces reoseochromogenes* NRRL 3504…….. 162

РОЗДІЛ 5 Гени, що контролюють продукцію глікопептидних антибіотиків в бактерій роду *Nonomuraea*…….. 192

5.1 Нові молекулярні знаряддя для експресії генів і покращення продукції ГПА в представниках роду *Nonomuraea*…… 193

5.2 A50926 – новий ГПА з *N. coxensis* DSM 45129……. 198

РОЗДІЛ 6 Особливості порівняльної філогеноміки кластерів генів біосинтезу глікопептидних антибіотиків…… 232

РОЗДІЛ 7 Регуляція продукції антибіотиків в *Streptomyces cyanogenus* S136 – моделі активації мовчазних кластерів біосинтетичних генів…… 249

7.1 Секвенування та характеристика генома *S. cyanogenus* S136……. 250

7.2 Вплив гетерологічних алелів гена *adpA* на продукцію ландоміцину А в *S. cyanogenus* S136…… 262
7.3 Зміни вторинного метаболому *S. cyanogenus* S136 у відповідь на експресію гетерологічних алелів гена *adpA*…… 276
РОЗДІЛ 8 Гени стійкості до глікопептидних антибіотиків в бактерій…. 306
8.1 Гени стійкості до глікопептидних антибіотиків в *A. teichomyceticus* ATCC 31121……. 306
8.2 Особливості розповсюдження та філогенії генів стійкості до глікопептидних антибіотиків в представників типу *Actinobacteria*….. 348
8.3 Особливості філогенії ABC-транспортерів, кодованих в кластерах біосинтетичних генів пептидних антибіотиків……. 381
8.4 Розповсюдження *van*-генів за межами типу *Actinobacteria*……. 396
Висновки……. 423
Список літературних джерел…… 426
Додаток Список публікацій здобувача…… 450

Перелік умовних скорочень

КБГ – кластер біосинтетичних генів;

п.н. – пар нуклеотидів;

млн.п.н. – мільйонів пар нуклеотидів;

ГПА – глікопептидні антибіотики;

СЕМ – сканувальна електронна мікроскопія;

кДНК – комплементарна ДНК;

ПЛР – полімеразна ланцюгова реакція, ЗТ-ПЛР – ПЛР із зворотною транскрипцією;

ТШХ – тонкошарова хроматографія;

ВЕРХ – високоефективна рідинна хроматографія, ВЕРХ-МС – ВЕРХ із мас-спектрометрією, МС/МС – ВЕРХ із тандемною мас-спектрометрією;

РБК – робочий банк клітин;

X-Gluc – 5-бромо-4-хлоро-3-індоліл-β-D-глюкуронід

НРПС – нерибосомна пептидсинтаза;

А-домен – аденілюючий домен;

ПКС – полікетидсинтаза;

Hpg – 4-гідроксифенілгліцин;

Dhpg – 3,5-дигідроксифенілгліцин;

Glc*N*Ac – *N*-ацетилглюкозамін;

ГГТ – горизонтальний генний трансфер;

МГЕ – мобільний генетичний елемент.

Вступ

Обґрунтування вибору теми дослідження. Антибіотики – це природні або синтетичні молекули, здатні пригнічувати або припиняти ріст і розмноження різноманітних мікроорганізмів. Антибіотики повністю змінили хід історії, давши змогу ефективно боротися із смертельними бактеріальними інфекціями. Хоча перші синтетичні антибіотики були введені в клінічне застосування Паулем Ерліхом наприкінці 19 століття, природні протимікробні сполуки з'явилися пізніше: бензилпеніцилін і граміцидин С (продукуються грибком *Penicillium chrysogenum* і Грам-позитивною бактерією *Bacillus brevis*) знайшли застосування в медицині в 1942 році [1]. Однак, справжній потенціал природних антибіотиків розкрився з відкриттям стрептоміцину (потрапив до клінічного використання в 1944 році) Селманом Ваксманом. Його роботи вперше показали, що актиноміцети – міцеліальні Грам-позитивні бактерії, що живуть у ґрунті – є фактично необмеженим джерелом природних антибіотиків [2]. Відтоді 24 класи природних антибактеріальних препаратів використовуються в клініці, а актиноміцети продукують представників 14 таких класів [1].

Цікаво, що більшість із цих антибіотиків є макромолекулярними структурами, біосинтетичні шляхи яких потребують десятків ферментів. Відповідно, ці ферменти кодовані генами, які колокалізуються на хромосомі, забезпечуючи синхронну експресію всіх необхідних компонентів біосинтезу. Такі асамблеї генів отримали назву кластерів біосинтетичних генів (КБГ). До складу КБГ, окрім біосинтетичних генів, також входять гени стійкості, експорту, біосинтезу попередників, та, щонайважливіше, регуляторні гени [3]. Потенціал актиноміцетів до продукції антибіотиків відображається в організації їх геномів: великі хромосоми (наприклад до 13 млн.п.н. у *Nonomuraea* sp. ATCC 55076 [4]) містять десятки КБГ для продукції антибіотиків. Зазвичай лише небагато таких КБГ активні за лабораторних умов; інші залишаються мовчазними з різних причин, наприклад внаслідок низької або відсутньої

експресії генів [5, 6]. Мовчазні біосинтетичні шляхи часто кодують біосинтез раніше невідомих – криптичних – антибіотиків [5].

Активне використання актиноміцетів (здебільшого *Streptomyces* spp.) розпочало «золоте століття» відкриття і використання антибіотиків [7]. Певні класи сполук почали також використовувалися у ветеринарії та як стимулятори росту сільськогосподарських тварин [8, 9]. Природне середовище забруднювалося антибіотиками, що надходили з лікарняних, комунальних і промислових стічних вод [10]. Таке широке, часто необґрунтоване використання антибіотиків мало серйозні наслідки. Як фактор природного добору, антибіотики сприяли змінам геномів факультативних і облігатних патогенів, а також еволюції метамобілому навколишнього середовища, що призвело до накопичення та поширення генів стійкості. Стійкість до протимікробних препаратів зробила багато успішних класів антибіотиків неефективними [11]. Передбачається, що множинно-резистентні штами *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* та *Enterobacter* spp. призведуть до 10 мільйонів смертей на рік до 2050 року [12]. Глобальна пандемія COVID-19, спричинена вірусом SARS-CoV-2, стимулює нераціональне використання антибіотиків [13–15]. Усе згадане вище робить пошук нових антибіотиків і вивчення біосинтезу вже відомих сполук важливими завданням біомедичної науки.

В цій роботі основний фокус зроблено на вивченні нерибосомних пептидних та полікетидних антибіотиків, які продукуються представниками різних родів актиноміцетів. Головну увагу приділено глікопептидним антибіотикам (ГПА). Ці сполуки є ліками вибору проти інфекцій, спричинених множинно-резистентними штамами золотистого стафілококу і *Enterococcus* spp. ГПА, що перебувають в клінічному використанні, включають як природні сполуки, такі як ванкоміцин і тейкопланін, так і напівсинтетичні антибіотики далбаванцин, орітаванцин і телаванцин [16]. Ці антибіотики входять до списку критично важливих для медицини препаратів Всесвітньої Організації Охорони Здоров’я [17]. ГПА інгібують ріст клітинної стінки Грам-позитивних бактерій зв’язуючи ліпід II [18]. Окрім клінічної значущості, ГПА цікаві завдяки

складним біосинтетичним шляхам: нерибосомні пептидні ядра цих сполук модифікуються шляхом приєднання залишків цукрів, сульфатуються, хлоруються, тощо [19, 20]. Нажаль, стійкі до ГПА стафілококи і ентерококи чимраз частіше трапляються в довкіллі [21, 22]. Тому гостро стоїть проблема пошуку нових ГПА та покращення продукції вже відомих.

Проблеми, згадані вище, вимагають зусиль, спрямованих у три основні напрямки, що і стали основними напрямками цієї робити. По-перше, ми вивчали генетичні знаряддя, які можна застосувати для маніпуляцій із відомими та потенційними продуцентами антибіотиків, з метою модифікації рівнів продукції цих сполук. Регуляція біосинтезу антибіотиків здійснюється на шлях-специфічному і глобальному рівнях [23]. Шлях-специфічні регулятори кодовані в межах відповідних КБГ, а їх регулон обмежується генами цього кластера [24–26]. Такі регулятори, здебільшого, є позитивними, а надекспресія їх генів веде до підвищення рівня продукції антибіотиків. Натомість, регулони глобальних регуляторів включають десятки генів, зокрема і шлях-специфічних регуляторів [27]; це дає глобальним регуляторам здатність одночасно контролювати шляхи біосинтезу різних антибіотиків. Мабуть, найцікавішим прикладом глобального позитивного регулятора біосинтезу антибіотиків у актиноміцетів є AdpA, що є перспективним знаряддям для покращення продукції антибіотиків [28]. Варто відмітити, що використання згаданих вище регуляторних генів із метою покращення рівнів продукції тих чи інших антибіотиків потребує ефективних платформ експресії цих генів, а також методів перенесення рекомбінантних молекул ДНК у штами продуценти. Зокрема, продуценти ГПА є часто дуже складними об’єктами для генно-інженерних маніпуляцій [29, 30], тому ряд підходів спочатку необхідно напрацювати на більш доступних моделях.

По-друге, потрібно розвивати підходи для активації мовчазного вторинного метаболому актиноміцетів. Для цього ми використовували глобальний позитивний регулятор AdpA.

По-третє, потрібно здійснювати постійний геномний скринінг потенційних КБГ з метою пошуку нових та цікавих кандидатів для експериментальної

роботи. Такі бази даних як MIBIG [3], а також алгоритми автоматизованого пошуку і анотації КБГ на кшталт antiSMASH [31] є ефективними знаряддями для ідентифікації нових КБГ.

Розуміння генетики регуляції та біосинтезу антибіотиків було б неповним без вивчення генів стійкості до цих сполук, пошуку первинних резервуарів генів стійкості та шляхів їх поширення.

Наша робота дає відповіді на ряд питань для кожного із напрямків, перелічених вище. У підсумку, ця робота демонструє нові механізми регуляції та активації біосинтезу антибіотиків, як у модельних об'єктах (таких як *Streptomyces albus, Streptomyces coelicolor, Streptomyces cyanogenus*), так і в продуцента клінічно важливого ГПА A40926 *Nonomuraea gerenzanensis* ATCC 39727. В світлі усього переліченого вище, вибрана тема дослідження є актуальною і відповідає сучасним напрямкам біомедичної науки.

Мета та завдання дослідження. *Метою* роботи було вивчити генетичні механізми регуляції біосинтезу пептидних та полікетидних антибіотиків. Для досягнення цієї мети поставлено такі *завдання*:

1) вивчити філогеноміку AdpA-подібних глобальних позитивних регуляторів і дослідити властивості інших компонентів *adpA*-опосередкованого регуляторного каскаду на моделі *Streptomyces albus (=albidoflavus)* J1074;
2) розробити набір методів для генно-інженерних маніпуляцій із *Streptomyces roseochromogenes* NRRL 3504 – потенційного продуцента салінамід-подібного пептидного антибіотика;
3) розробити набір методів для генно-інженерних маніпуляцій та експресії гетерологічних генів в *Actinoplanes rectilineatus* NRRL B-16090 – потенційного продуцента ГПА;
4) створити підходи для створення надпродуцентів промислово-важливого ГПА A40926 в *Nonomuraea gerenzanensis* ATCC 55076;
5) дослідити потенціал *Nonomuraea coxensis* DSM 45129 до продукції нових ГПА;

6) побудувати філогенетичну модель еволюції КБГ ГПА, з акцентом на пошук нових КБГ подібних до комплестатину;
7) вивчити *adpA*-опосередкований глобальний генетичний контроль продукції ландоміцину А та способи активації інших вторинних метаболітів у *Streptomyces cyanogenus* S136, як моделі активації мовчазних КБГ;
8) вивчити механізми стійкості до ГПА в продуцента тейкопланіну *Actinoplanes teichomyceticus* ATCC 31121;
9) здійснити порівняльний філогеномний аналіз розповсюдження генів стійкості до ГПА серед бактерій, з метою виявлення вихідного резервуару цих генів та вивчення шляхів їх розповсюдження.

Об'єкт дослідження – механізми регуляції біосинтезу пептидних та полікетидних антибіотиків, гени стійкості до ГПА.

Предмет дослідження – КБГ ГПА і полікетидних антибіотиків, їх шлях-специфічні регуляторні гени; регуляторні гени *adpA*-опосередкованого каскаду стрептоміцетів – *adpA, bldA, miaA, miaB*; гени регуляторів YtrA-підродини актинобактерій; системи для гетерологічної експресії генів; регуляторні гени біосинтезу хлоробіоцину *novG* і *cloG*; гени стійкості до тейкопланіну в *A. teichomyceticus*; гени стійкості до ГПА бактерій; гени експортерів пептидних антибіотиків.

Методи дослідження. В роботі було застосовано набір загальноприйнятих методів досліджень, зокрема:

- мікробіологічні методи досліджень – культивування мікроорганізмів, вивчення їх морфології методами світлової та сканувальної електронної мікроскопії (СЕМ), вивчення антибіотичних властивостей штамів методами пригнічення росту тест-культур;
- біотехнологічні методи досліджень – культивування штамів актиноміцетів-продуцентів антибіотиків в умовах, наближених до промислових, в колбах Ерленмеєра або лабораторних ферментерах, вивчення динаміки росту культур, накопичення біомаси і антибіотиків;

- генетичні й генно-інженерні методи досліджень – отримання і вивчення рекомбінантних штамів бактерій, генетична трансформація клітин *Escherichia coli*, кон'югаційне перенесення плазмід з донорських штамів *E. coli* до різних актиноміцетів; виділення плазмідної та хромосомної ДНК, виділення сумарної РНК та синтез кДНК, рестрикційний аналіз плазмідної ДНК, конструювання рекомбінантних молекул ДНК, горизонтальний гель-електрофорез ДНК, полімеразна ланцюгова реакція (ПЛР), секвенування ДНК, напівкількісна полімеразна ланцюгова реакція із зворотною транскрипцією.
- біохімічні методи досліджень – кількісний та якісний аналіз антибіотиків продукції антибіотиків за допомогою тонкошарової хроматографії (ТШХ), високоефективної рідинної хроматографії (ВЕРХ) та ВЕРХ із мас-спектрометрією (ВЕРХ-МС) чи тандемною мас-спектрометрією (МС/МС), екстракція антибіотиків, якісний і кількісний аналіз β-глюкуронідазної активності.
- біоінформатичні методи досліджень – аналіз нуклеотидних та амінокислотних послідовностей, анотація генів, пошук та характеристика КБГ, моделювання біосинтетичних шляхів *in silico*, філогенетичний аналіз амінокислотних та нуклеотидних послідовностей.

Наукова новизна отриманих результатів. Вперше здійснено філогеномний аналіз AdpA-подібних регуляторів та продемонстровано важливість окремих доменів цього білка; вперше вивчено інші компоненти *adpA*-опосередкованого регуляторного каскаду на моделі *S. albus* (=*albidoflavus*) J1074. Вперше вивчено функції YtrA-подібних регуляторів на моделі *S. coelicolor* A3(2). Розроблено ефективні методи для генно-інженерних маніпуляцій продуцента хлоробіоцину і потенційного продуцента салінамід-подібного пептидного антибіотика *S. roseochromogenes* NRRL 3504. Вперше розроблено набір знарядь для генно-інженерних маніпуляцій із потенційним продуцентом нового ГПА *A. rectilineatus* NRRL B-16090. Розроблено нові підходи до покращення біосинтезу клінічно важливого ГПА A40926 в *N. gerenzanensis* ATCC 55076. Відкрито та описано біосинтетичний шлях

нового ГПА – названого A50926 – в *N. coxensis* DSM 45129. Показана роль функціональних алелів *adpA* для продукції ландоміцину А в *S. cyanogenus* S136, а також їх важливість як знарядь активації мовчазного вторинного метаболому. Вперше описано КБГ полікетидного антибіотика люцензоміцину. Побудовано повноцінну модель функціонування генів стійкості до тейкопланіну в *A. teichomyceticus* ATCC 31121; вперше охарактеризовано глобальне розповсюдження генів стійкості до ГПА на основі даних порівняльної геноміки.

Практичне значення отриманих результатів. В ході виконання роботи автором створено надпродуцентів хлоробіоцину, які за рівнем продукції цього антибіотика не мають аналогів, описаних в літературі. Створено низку платформ для надекспресії генів в штамах продуцентах ГПА, таких як *Actinoplanes* spp. і *Nonomuraea* spp. Створено надпродуцентів промислово-важливого антибіотика A40926 на основі *N. gerenzanensis* ATCC 55076, що в рази переважають за продуктивністю інших описаних в літературі надпродуцентів A40926. Вперше створено альтернативний продуцент A40926 на основі *N. coxensis* DSM 45129 методом гетерологічної експресії гена *dbv29*. В ході виконання роботи, автор продемонстрував важливість різних алелів *adpA* для активації мовчазних КБГ, зокрема КБГ люцензоміцину в *S. cyanogenus* S136 – на ці результати автор отримав Патент України на винахід. Автором створено повний опис генів стійкості до ГПА серед усіх бактерій, що є важливим доповненням до нагляду за поширенням множинно-резистентних патогенів.

Особистий внесок здобувача. Роботи здобувача, представлені в цій дисертації, засновані на результатах отриманих ним особисто, або за його безпосередньої участі чи під його керівництвом. Мету і задачі дослідження, а також висновки було сформульовано разом із д.б.н., проф. В.О. Федоренком. Вивчення компонентів *adpA*-опосередкованого регуляторного каскаду в *S. albus* здійснювалося разом із к.б.н. О.Т. Кошлою (ЛНУ ім. І. Франка). Дослідження властивостей генів YtrA-регуляторів *S. coelicolor* було виконано разом із асп. О. Ципік (ЛНУ ім. І. Франка). Виділення та ВЕРХ-МС аналіз люцензоміцину було виконано разом із к.б.н. І.С. Осташ (ЛНУ ім. І. Франка) та групою проф. Р. Зюсмута (Технічний Університет Берліна, Німеччина). Секвенування та

анотація геномів актиноміцетів була здійснена у співпраці із групою проф. Й. Каліновські (Університет Білефельда, Німеччина). ВЕРХ-МС аналіз нового ГПА А50926 виконано у співпраці із доктором Е. Труманом і доктором Н.-М. Віор (Джон Іннес Центр, Великобританія). Вивчення продукції ГПА здійснювалося спільно з доктором Ф. Беріні, доктором Е. Біндою, доктором А. Андрео-Відалем (керівником дисертації доктора філософії – PhD – якого був автор) (Університет Інзубрії, Італія). Сканувальна електронна мікроскопія актиноміцетів виконана спільно із к.ф.-м.н. Ю.Р. Дацюком (ЛНУ ім. І. Франка). Планування експериментів, аналіз та обговорення отриманих результатів проведено спільно з д.б.н., проф. В.О. Федоренком, а також д.б.н. Б.О. Осташем (ЛНУ ім. І. Франка), проф. Ф. Марінеллі (Університет Інзубрії, Італія), проф. М. Бібом (Джон Іннес Центр, Великобританія), к.б.н. М. Рабик (ЛНУ ім. І. Франка), аспірантками С. Мельник і К. Жукровською (ЛНУ ім. І. Франка), з якими автор має спільні публікації. З наукових праць, опублікованих у співавторстві, в дисертаційній роботі використано тільки ті ідеї та здобутки, що являються особистим внеском здобувача.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційну роботу виконано в науково-дослідній лабораторії генетики, селекції та генетичної інженерії продуцентів біологічно активних речовин (НДЛ-42) кафедри генетики та біотехнології Львівського національного університету імені Івана Франка. Роботу виконано в межах держбюджетних тем БГ-41Нр «Універсальний генетичний механізм контролю продукції біологічно-активних речовин стрептоміцетами» (№ держреєстрації 0116U008070), БГ-46Ф «Нові гени актинобактерій, що контролюють продукцію і стійкість до антибіотиків-інгібіторів синтезу пептидоглікану» (№ держреєстрації 0117U001224), БГ-09Ф «Мутації стійкості актинобактерій до антибіотиків: джерело нових уявлень про механізми резистентності та біотехнологічних знарядь» (№ держреєстрації 0120U102039); теми Державного Фонду фундаментальних досліджень України у рамках співпраці з Японським т-вом сприяння науки (JSPS) Ф80-2018 «Посттранскрипційні модифікації тРНК як регулятори первинного й вторинного метаболізму в актинобактерій»; тем у

рамках спільної грантової програми Міністерства освіти і науки України й Федерального Міністерства Німеччини освіти і науки Німеччини (BMBF) М/18-2017 «Плейотропні транскрипційні регулятори родини AdpA як знаряддя відкриття нових природних біоактивних сполук», М/26-2018 «Плейотропні транскрипційні регулятори родини AdpA як знаряддя відкриття нових природних біоактивних сполук». Частину досліджень виконано під час наукового стажування на факультеті біотехнології та наук про життя Університету Інзубрії (Італія, 2018-2020, 2022 р.) за індивідуальними комерційними грантами та персональним грантом від Університету Інзубрії. За роботи, представлені в цій дисертації, автор був удостоєний Премії Верховної Ради України молодим ученим за 2021 рік (постанова ВРУ № 2833-IX від 13 грудня 2022 року).

Апробація результатів дисертації. Основні результати дисертації доповідалися та обговорювалися на Міжнародній конференції «Integrative Biology & Medicine», Київ (2017); XII Міжнародній конференції молодих науковців «Біологія: від молекули до біосфери», Харків (2017); Міжнародній конференції «Advances in Microbiology and Biotechnology», Львів (2018); XIV Міжнародній науковій конференції студентів і аспірантів «Молодь та поступ біології», присвяченій 185 річниці від дня народження Б. Дибовського, Львів (2018); XIII Міжнародній конференції молодих науковців «Біологія: від молекули до біосфери», Харків (2018); XV Міжнародній науковій конференції студентів і аспірантів «Молодь та поступ біології», присвяченій 135 річниці від дня народження Я. Парнаса, Львів (2019); 3rd International Conference on Natural Products Discovery and Development in the Genomic Era, San Diego (2020); XVII Міжнародній науковій конференції студентів і аспірантів «Молодь та поступ біології», Львів (2021); Міжнародному онлайн конгресі «VAAM Workshop 2021», Тюбінген (2021); Міжнародному онлайн конгресі «ExpoBAC Zaragoza 2020», Сарагоса (2021); IV Міжнародній науковій конференції «Microbiology and immunology – the development», Київ (2022); XVIII Міжнародній науковій конференції студентів і аспірантів «Молодь та поступ біології» присвяченій 195 річниці від дня народження Ю. Планера, Львів (2022); на звітних наукових

конференціях ЛНУ імені Івана Франка (2018-2022); звітних наукових конференціях факультету біотехнології та наук про життя Університету Інзубрії (2018-2020).

Публікації. За матеріалами дисертації опубліковано 44 наукові роботи, із них 3 роботи у наукових фахових виданнях України, 20 у зарубіжних наукових журналах, проіндексованих в міжнародних наукометричних базах даних Scopus і Web of Science, 19 тез доповідей на вітчизняних та закордонних наукових конференціях. Серед цих робіт: 7 експериментальних статей у журналах, що належать до Q1 SCImago; 6 експериментальних статей у журналах, що належать до Q2 SCImago; 3 експериментальні статті у журналах, що належать до Q3 SCImago; 4 літературні огляди у журналах, що належать до Q1 SCImago; 3 експериментальні статті у фахових виданнях України; один розділ монографії видавництва Springer і один патент України на винахід.

Структура та обсяг дисертації. Дисертаційна робота є кваліфікаційною науковою працею на правах наукової доповіді за сукупністю статей, викладена на 460 друкованих сторінках. Дисертація складається з наступних структурних елементів: вступу, огляду літератури, розділу матеріалів та методів досліджень, аналізу і узагальнення результатів досліджень (викладено в 5 розділах), висновків, переліку використаних джерел та додатку. Розділами дисертації є публікації здобувача наукового ступеня в виданнях високого рівня. Список використаної літератури включає 1549 позицій, з яких 1342 є в публікаціях здобувача. Робота містить 127 рисунків і 29 таблиць (в публікаціях здобувача).

РОЗДІЛ 1 Огляд літератури

Актиноміцети – збірна нетаксономічна назва ґрунтових актинобактерій, що ростуть у вигляді міцелію [32] – є відомими продуцентами антибіотиків, які широко застосовуються в медицині [1]. Антибіотики є лише частиною продуктів вторинного спеціалізованого метаболізму цих організмів. Біосинтез антибіотиків потребує десятків ферментів, гени яких колокалізовані на хромосомі. Такі сукупності генів отримали назви кластерів генів біосинтезу (КБГ). Координація експресії біосинтетичних генів здійснюється за допомогою шлях-специфічних регуляторів, гени яких також входять до складу КБГ. Окрім структурних і регуляторних, в складі КБГ можна знайти також гени експортерів і гени стійкості. Останні необхідні, бо продуценти потребують механізмів захисту від власних продуктів. Цікаво, що продукція антибіотиків і морфологічна диференціація актиноміцетів часто ко-регульована за допомогою транскрипційних регуляторів вищого (ніж шлях-специфічний) рівня.

Серед найважливіших антибіотиків є нерибосомні пептидні і полікетидні антибіотики. Ці сполуки синтезуються за допомогою великих мультидоменних ферментів, названих нерибосомними пептидсинтазами (НРПС) і полікетидсинтазами (ПКС). В подальших етапах біосинтезу цих антибіотиків присутні різноманітні модифікації, з яких найпоширенішими і найрізноманітнішими є реакції додавання цукрів. Клінічно-релевантним класом нерибосомних пептидних антибіотиків є глікопептидні антибіотики (ГПА). ГПА інгібують біосинтез клітинної стінки Грам-позитивних бактерій зв’язуючись із ліпідом II. Біосинтез ГПА забезпечується експресією складно-організованих КБГ, які можна знайти у різних груп актиноміцетів. Різноманіття генів КБГ ГПА веде до значного хімічного різноманіття цих сполук, що зустрічаються в природі. У подібний спосіб відбувається біосинтез полікетидних антибіотиків.

Усі вивчені аспекти біосинтезу ГПА, регуляторні механізми, а також механізми стійкості до цих антибіотиків повно висвітлені в ряді оглядових публікацій автора [33–35]. Генетика біосинтезу полікетидних антибіотиків розглянута на прикладі ландоміцину А, що описано в огляді автора [36].

Glycopeptide Antibiotics: Genetics, Chemistry, and New Screening Approaches 16

Oleksandr Yushchuk and Bohdan Ostash

Abstract

Glycopeptide antibiotics (GPAs) that target bacterial cell wall remain one of the most successful families of antibiotics. GPAs are produced by Actinobacteria, a class of Gram-positive microorganisms with complex life cycle and sizable genomes. The elucidation of biosynthesis of GPAs is important for several reasons. First, GPA producers are likely a primary source of respective antimicrobial resistance (AMR) genes; their study will inform us of possible ways for AMR evolution among pathogens. Second, the biosynthetic studies yield information about regulatory mechanisms governing the production of GPAs, which can be used in biotechnology. Third, manipulation of GPA biosynthetic genes is a viable approach to generate novel GPAs of practical value. Lastly, the information on GPA genes can be used to find new compounds through computational analyses of publicly available genomic information. In this chapter, we outline a comprehensive picture of GPAs, focusing on genetics and chemistry behind their biosynthetic pathways. Examples of the use of the gained knowledge for strain improvement and discovery of novel compounds are given. A concise outlook of bioactivities exhibited by GPAs will conclude the chapter.

Keywords

Glycopeptide antibiotics · Lipid II binders · Actinobacteria · Biosynthetic gene clusters · Non-ribosomal peptide synthetase · Transcriptional regulation · Resistance to glycopeptide antibiotics

O. Yushchuk · B. Ostash (✉)
Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, Lviv, Ukraine
e-mail: oleksandr.yushchuk@lnu.edu.ua; bohdan.ostash@lnu.edu.ua

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

R. V. Rai, J. A. Bai (eds.), *Natural Products from Actinomycetes*,
https://doi.org/10.1007/978-981-16-6132-7_16

16.1 Introduction

16.1.1 Bacterial Cell Wall: Organization, Biosynthesis, and Inhibitors

The cell membrane of the Gram-positive bacteria is covered with a thick (10–40 nm) layer of peptidoglycan (murein, PG), which is decorated with lipoteichoic acids (anchored in the cell membrane, LTA) and wall-teichoic acids (WTAs, submerged into PG). Such an organization of the cell wall is characteristic for many pathogens, such as staphylococci. The cell envelope of Gram-negative bacteria contains a thinner (3–6 nm) layer of PG, located between two membranes—the inner cytoplasmic membrane and the outer membrane. Unlike a symmetric phospholipid bilayer of the cytoplasmic membrane, the outer membrane is asymmetrically composed with PG-adjacent phospholipid layer and the external lipopolysaccharide (LPS) layer. Corynebacteria and mycobacteria possess much more complicated cell envelope, although belonging to a typically Gram-positive phylum Actinobacteria (Silhavy et al. 2010).

Chains of PG consist of repeating building blocks—*N*-acetylglucosamine (Glc*N*Ac) linked to *N*-acetylmuramic acid (Mur*N*Ac) via β 1 → 4 glycosidic bonds; pentapeptide side chains are attached to Mur*N*Ac. Most commonly, these side chains are formed with L-Ala–D-Gln–L-Lys–D-Ala–D-Ala in Gram-positive bacteria. PG biosynthesis is similar across bacteria; this topic is extensively reviewed (Barreteau et al. 2008; Vollmer et al. 2008; Silhavy et al. 2010; Liu and Breukink 2016) and summarized in Fig. 16.1.

PG biosynthesis is vital for bacteria and remains the most sought-after target for antibiotics (Sarkar et al. 2017). All inhibitors of PG biosynthesis fall into one of the two main groups: direct inhibitors of biosynthetic enzymes and binders of PG intermediates, such as lipid I and lipid II, and pentapeptide stems of PG monomers (see Fig. 16.1). All compounds that belong to the second group act on cell surface, and thus do not require mechanism(s) for cell entry.

16.1.2 Why Are GPAs So Important?

Among all PG-targeting antibiotics, glycopeptides (GPAs, *sensu* related to vancomycin) are the most thoroughly studied and clinically successful compounds (rivaled probably only by β-lactams). Resistance to GPAs among pathogens emerges very slowly (Zeng et al. 2016), perhaps due to the peculiar nature of their principal target, lipid II. The latter consists of Glc*N*Ac-Mur*N*Ac-pentapeptide attached to undecaprenyl lipid chain and acts as a molecular handle for PG-building units. GPAs bind D-Ala-D-Ala terminus of pentapeptide stem, forming five hydrogen bonds (Barna and Williams 1984). This interferes with the upstream transpeptidase and transglycosylase reactions, preventing the formation of cross-linked PG. The fraction of lipid II is extremely small among all cell membrane phospholipids (less than 1%, de Kruijff et al. 2008). At the same time, lipid II is heavily used; export of more than 5000 molecules of lipid II to the outer side of cell membrane is necessary

and in pathogens, triads of *van* genes usually form one operon. GPAs exhibit decreased affinity to D-Ala-D-Lac and D-Ala-D-Ser termini, forming only four hydrogen bonds, not enough for stable interaction. Alternatively, GPA resistance is achieved with the help of VanY, a D,D-carboxypeptidase cleaving the terminal D-Ala from pentapeptide stem. Such PG can still be cross-linked, but GPAs do not form stable complexes with it anymore. VanY is rather accessory GPA resistance factor in pathogens, but it might be a single resistance determinant in GPA-producing Actinobacteria (Marcone et al. 2010). Expression of *van* genes is usually regulated with VanRS two-component regulatory circuit, which consists of a sensor histidine kinase (VanS) and response regulator (VanR). VanS, upon sensing GPA, phosphorylates VanR, which upregulates *van* operon. VanS from different species often recognizes only certain GPA.

Today, two natural GPAs (the first generation)—vancomycin and teicoplanin, as well as three semisynthetic GPAs (second generation)—telavancin, oritavancin, and dalbavancin—are in clinical practice (Zeng et al. 2016). Vancomycin, produced by different *Amycolatopsis* (*Am.*) *orientalis* strains, was the first GPA marketed in 1958. Only 30 years later, the first vancomycin-resistant clinical isolate (*Enterococcus faecium* strain) was reported. Nowadays, unfortunately, vancomycin and teicoplanin resistance is widespread among enterococci and staphylococci. The spread of GPA resistance does not render GPAs worthless. This is due to several reasons. First, nature offers a remarkable chemical variety of GPAs (Nicolaou et al. 1999), which expands ever since the discovery of vancomycin and exploded in modern "post-genomic" era. Second, biosynthesis of GPAs is modular, which allows to use enzymes from one GPA biosynthetic pathways to modify another, leading to unprecedented possibilities for combinatorial biosynthesis. Therefore, VanS sensors, acting in pathogens, might be "bluffed" with minor modifications of the existing drugs with relative ease (Kalan et al. 2013; Yim et al. 2018). Finally, deep knowledge of genetics behind the GPA production allows to develop GPA overproducers much faster than before (Horbal et al. 2014; Lo Grasso et al. 2015; Yushchuk et al. 2020a). These advances allow to counter the spread of GPA-resistant pathogens with novel GPAs. Recent years were especially fruitful for new discoveries of unusual GPAs and underlying biosynthetic routes, which merit a detailed review. In this chapter, we portray an updated picture of biosynthesis of natural GPAs, going from genes to products.

16.1.3 Definitions

Historically, to define GPAs one took into account structure, origin, and mode of action. That is, GPAs are non-ribosomally synthesized molecules whose peptide core consists of seven amino acids, often non-proteinogenic; the core can be decorated with sugar moieties, halogen atoms, sulfate groups, or acyl chains; they are produced by Actinobacteria, and they bind D-Ala-D-Ala termini of PG of Gram-positive bacteria. In this chapter, we discuss GPA as they are denoted above. However, in a broad chemical sense, term "glycopeptide" can be attributed to any

glycosylated peptide molecule, such as bleomycin, that does not target PG (Hecht 2000).

All GPAs (*sensu* structurally related to vancomycin) are classified into five types according to their structural features (Nicolaou et al. 1999). Type I antibiotics have aliphatic amino acids in positions 1 and 3 of their aglyca (e.g., vancomycin); type II features amino acids with aromatic side chains at these positions, uncross-linked (e.g., keratinimicin); types III and IV have aromatic amino acids at 1 and 3 aglycone positions and they are cross-linked (e.g., ristocetin); finally, type V is different from all the rest by carrying tryptophan residue at the position 2 of aglycone (e.g., complestatin). Type I–IV GPAs bind D-Ala-D-Ala. Some GPAs, like A47934 (Boeck and Mertz 1986; Pootoolal et al. 2002), are D-Ala-D-Ala binders whose aglycones are not glycosylated. Recent discoveries of novel type V compounds (Cheng and Wuest 2020; Culp et al. 2020; Xu et al. 2020) revealed that at least some of them might have nine amino acid peptide core (or even more, Waglechner et al. 2019). Thus, historical definition of all these compounds as GPAs is quite confusing. There was another early term for types I-IV—dalbaheptides (Parenti and Cavalleri 1989; Cavalleri and Parenti 2000), derived from abbreviation of "***d-Ala***a-D-Ala-***b***inding ***a***ntibiotics with ***heptapeptide*** cores." This term includes both glycosylated and nonglycosylated members of types I–IV and might be further extended to glycodalbaheptides and lipoglycodalbaheptides. However, type V compounds are definitely not "dalba-" and often not "-heptides." Further on, we will often use term dalbaheptide referring to type I–IV GPAs. It is important, in our opinion, to propose a term, which would distinguish both dalbaheptides and complestatin-related antibiotics from the other "glycopeptides," *sensu lato*.

16.2 Structural Classification of GPAs with Recent Updates

GPAs are currently represented by dozens of natural compounds (Nicolaou et al. 1999). Majority of known GPAs are composed of oxidatively cross-linked heptapeptide aglyca, which may undergo various modifications. Seven amino acids are usually denoted as *aa-1*–*aa-7* from N- to C-terminus of the oligopeptide (Fig. 16.2). In rare cases (discussed further), the position of *aa-1* could be actually occupied with a keto acid. Vancomycin, the archetypical GPA (produced by *Am. orientalis* strains, Fig. 16.2), harbors amino acids with aliphatic side chains in *aa-1* and *aa-3* positions; positions 2 and 4 to 7 are occupied with aromatic amino acid. These five aryl rings (denoted with letters, A to E, Fig. 16.2) are involved in the formation of macrocycle structures by oxidative cross-linking, giving one biaryl (AB) and two bisaryl ethers (C-*O*-D and D-*O*-E). Such combination of features is characteristic for the other vancomycin-like GPAs (balhimycin, chloroeremomycin) and recently discovered type I GPA pekiskomycin (Thaker et al. 2013).

Aglyca of type II GPAs are composed exclusively of aromatic amino acids, thus adding F and G aryl rings at amino acid positions 1 and 3 (Fig. 16.2). They have the same macrocycles as type I GPAs (Fig. 16.2). Among antibiotics in this rather scanty type are recently discovered keratinimicins (Fig. 16.2, Xu et al. 2019), produced by

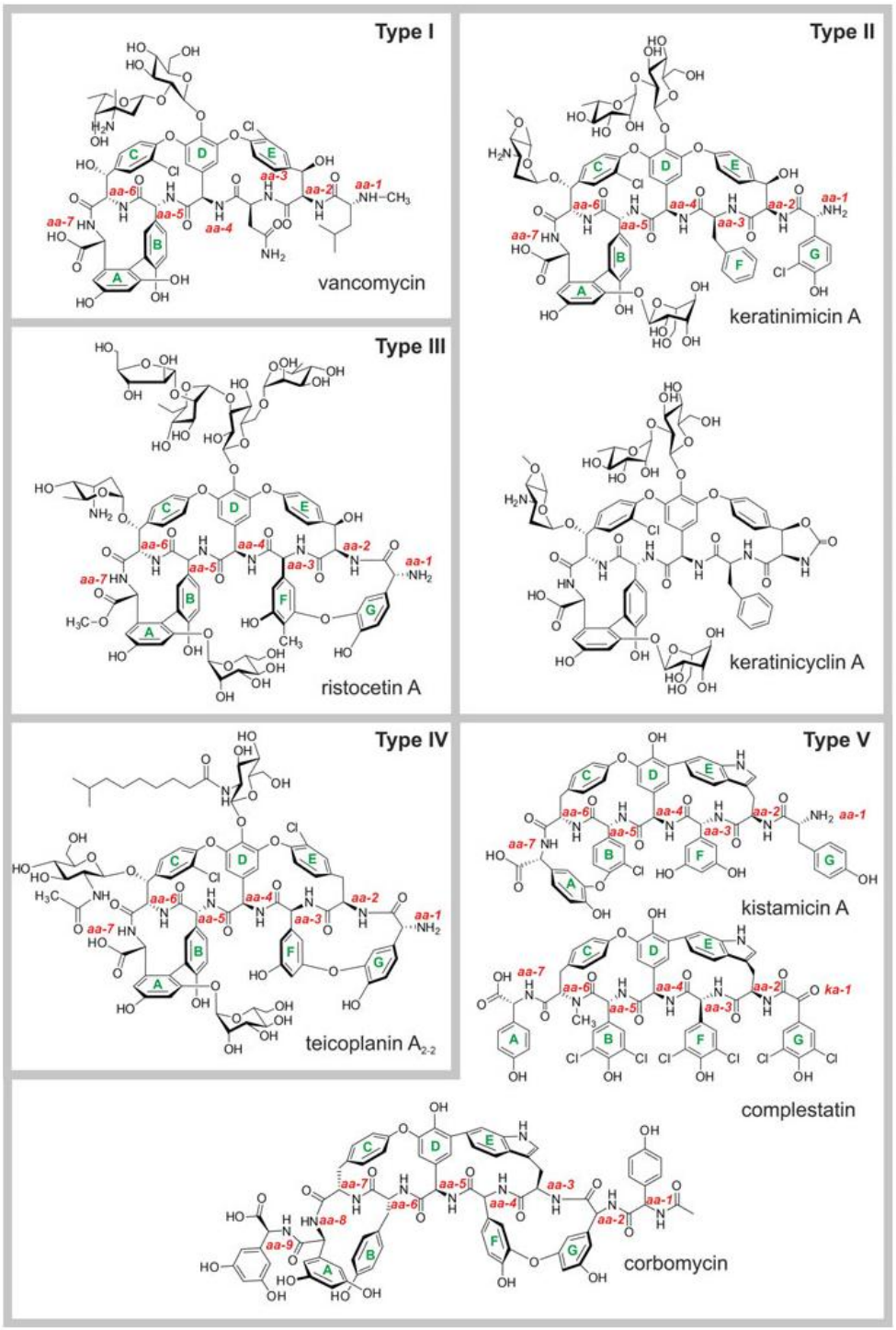


Fig. 16.2 Structural classification of GPAs into five types according to Nicolaou et al. (1999); each class shown with typical representative. Refer to the main text for more details

Am. keratiniphila NRRL B-24117. The latter also produces new natural derivatives of keratinimicins, known as keratinicyclines (minor fraction, Xu et al. 2019). In their aglycone, *aa-1* is substituted with N-terminal 2-oxazolidinone ring (Fig. 16.2). Thus, keratinicyclines combine features of dalbaheptides with the active site of oxazolidinone antibiotics that inhibit translation. However, such combination might not augment biological activity of these compounds, since targets of oxazolidinones are inside the cell, whereas dalbaheptides act on PG. Mechanisms behind the formation of keratinicyclines are not yet known. It is likely that the other type II GPAs also undergo such "-cyclinization."

Type III and IV GPAs feature an additional F-*O*-G macrocycle (Fig. 16.2). Type III GPAs are often hyperglycosylated (like ristocetin and actaplanin), some are aglyca (A47934). The presence of fatty acid side chain is a marker of type IV GPAs (such as teicoplanin and A40926, Fig. 16.2). Teicoplanin and A40926 are GPAs of clinical value (Alduina et al. 2018; Yushchuk et al. 2020c).

Type V GPAs are not limited to compounds with heptapeptide cores. For example, recently discovered GP6738 (Xu et al. 2020) and corbomycin (Culp et al. 2020) are nonapeptides. Oxidative cross-linking patterns are also variable in type V GPAs. Although D-E biaryl and C-*O*-D bisaryl ether bonds are present in all known type V GPAs (and might be the only cross-links, as in complestatin), they can also feature additional A-*O*-B (kistamicin) or A-B and F-*O*-G macrocycles (corbomycin). Type V GPAs are not glycosylated.

Members of all five types are exclusively Actinobacteria-derived antibiotics. GPAs are most abundant in "rare" actinobacterial families such as *Pseudonocardiaceae* (Adamek et al. 2018; Xu et al. 2020), *Micromonosporaceae, and Streptosporangiaceae* (Goldstein et al. 1987; Naruse et al. 1993a). Until very recently, GPAs were believed to be rather rare in genus *Streptomyces,* but now it becomes evident that streptomycetes might be abundant source of type I (Thaker et al. 2013), type III (Pootoolal et al. 2002), and especially type V GPAs (Waglechner et al. 2019; Culp et al. 2020; Xu et al. 2020).

16.3 Organization of GPA BGCs and Where (How) to Find Them

Multiple GPAs were discovered in course of microbiological screening programs, yielding remarkable chemical diversity of this class (Nicolaou et al. 1999). Genetic studies of producers of well-known GPAs yielded first GPA BGCs. These included BGCs for chloroeremomycin (*cep* from *Am. orientalis*), later shown to be in fact Kibdelosporangium aridum A82846, Waglechner et al. 2019), complestatin (*com* from *Streptomyces (S.) lavendulae* SANK 60477, Chiu et al. 2001), A47934 (*sta* from *S. toyocaensis* NRRL 15009, Pootoolal et al. 2002), teicoplanin (*tei* from *Actinoplanes (A.) teichomyceticus* NRRL B-16726, Sosio et al. 2003; Li et al. 2004), A40926 (*dbv* from *Nonomuraea (N.) gerenzanensis* ATCC 39727, Sosio et al. 2003), and balhimycin (*bal* from *Am. balhimycina* DSM 5908, Shawky et al. 2007). Later on, when genome sequencing became routine, *auk* and *kis* (kistamycin) BGCs were discovered in UK-68,597 producer *Actinoplanes* sp. ATCC 53533

(Yim et al. 2014) and *Nonomuraea* sp. ATCC 55076 (Nazari et al. 2017; Greule et al. 2019).

Later, analysis of metagenomic data showed the abundance of GPA-like BGCs. These were TEG and VEG BGCs (Banik and Brady 2008); CA37, CA878, and CA915 BGCs (Banik et al. 2010); GQ475282 (Kim et al. 2010) and esnapd15 (Owen et al. 2013). Not a single natural host of these BGCs was cultivated, and their final products are unknown (although predicted); however, these BGCs became a source of novel GPA tailoring enzymes. In the same time, accumulation of knowledge about the genetics of GPA production and resistance allowed to develop screening pipelines tailored specifically for discovering GPAs. One such pipeline yielded pekiskomycin (and *pek* BGCs coming from *Streptomyces* spp. WAC1420 and WAC4229, Thaker et al. 2013). Here, a collection of cultivable soil actinomycetes was "sifted" through a several "sieves." First, these strains were cultivated in the presence of vancomycin, and resistant colonies were selected. The vancomycin-resistant strains were assumed to be GPA producers. The presence of *van* genes in the selected resistant strains was verified via PCR. Strains that passed these two "filters" were further PCR-screened for the most conserved genes for cross-linking P450 monooxygenases involved in GPA biosynthesis. Positive strains were tested for GPA production and underwent shotgun genome sequencing to determine corresponding BGCs. Similar approach, involving GPA-sensitive microbiological assay system, was utilized to discover novel ristocetin BGC in *Amycolatopsis* sp. MJM2582 (Truman et al. 2014). Some of other GPA BGCs were silent and required activation (Spohn et al. 2014; Xu et al. 2019).

The recent dawn of post-genomic era had a significant impact on the search of novel GPA BGCs. It turned out that genomic databases harbor many GPA BGC-encoding genomes. This was revealed in a large phylogenomic attempt to reconstruct the evolution of GPA BGCs (Waglechner et al. 2019). Interestingly, one subset of GPA BGCs lacked typical GPA resistance genes, leading to BGC for a novel nonapeptide type V GPA—corbomycin (from *Streptomyces* sp. WAC01529, *crb* BGC, Culp et al. 2020). The latter has an unprecedented mode of action against Gram-positive bacteria (discussed further). A GPA BGC expression platform was then created, allowing to express in a heterologous host BGCs silent in native hosts (Xu et al. 2020).

GPA BGCs are large and have complex transcriptional ensembles. Certain common features of their organization could be deduced. In all cases, non-ribosomal peptide synthetase (NRPS) genes are oriented in one direction and immediately followed be cross-linking oxygenase genes (Fig. 16.3). A single exception is *dbv*, which is bizarrely rearranged for unknown reasons (Fig. 16.3g). Peculiarly, a recently described BGC (*noc*) for a novel A40926-related GPA from *N. coxensis* DSM 45129 is almost identical to *dbv*, but has normal arrangement of NRPS genes (Yushchuk et al. 2021). Majority of structural genes are located downstream of NRPS genes, at least in BGCs from *Actinoplanes* spp. and *Amycolatopsis* spp. BGCs from *Streptomyces* spp. (including type V GPA BGCs) tend to have structural genes both up- and downstream of NRPS genes (Fig. 16.3d, e, i). Another characteristic feature of all dalbaheptide BGCs is the presence of *van* genes (see above); their

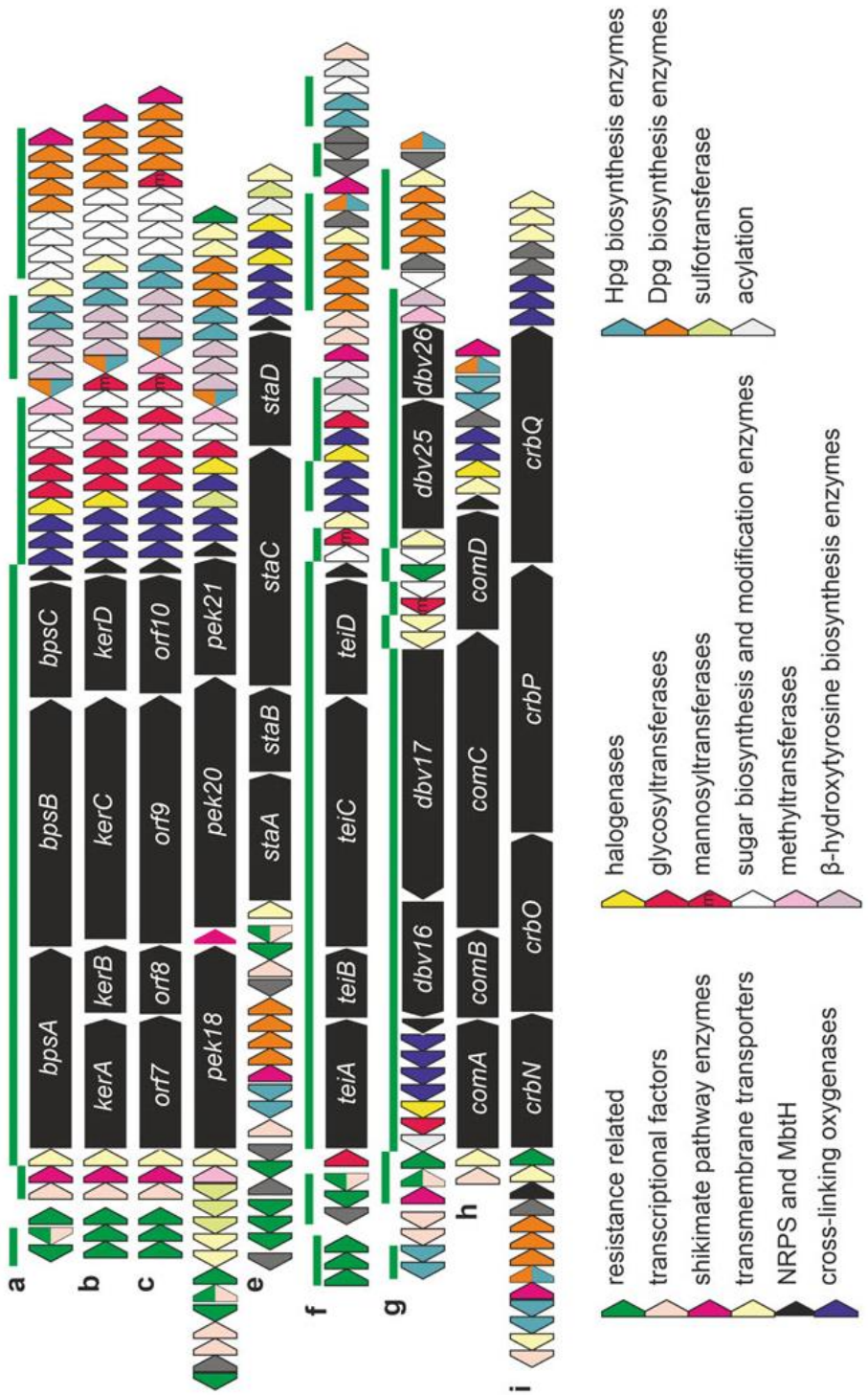


Fig. 16.3 Comparison of the genetic organization of some GPA BGCs, namely (**a**) *bal* for balhimycin; (**b**) *ker* for keratinimicin; (**c**) BGC for ristocetin; (**d**) *pek* for pekiskomycin; (**e**) *sta* for A47934; (**f**) *tei* for teicoplanin; (**g**) *dbv* for A40926; (**h**) *com* for complestatin; and (**i**) *crb* for corbomycin. For (**a**), (**f**), and (**g**), operon structure was established experimentally; operons are shown above the genes with solid green lines. Gene sizes are given not to scale

organization in different BGCs was recently discussed in detail (Yushchuk et al. 2020b).

Semi-quantitative reverse transcription PCR was used to elucidate transcriptional organization for *bal*, *tei,* and *dbv* BGCs. The first of them has rather straightforward organization with six polygenic transcriptional units (Fig. 16.3a, Shawky et al. 2007). Since other *Amycolatopsis*-derived GPA BGCs are rather syntenic with *bal*, it might be construed that they are organized similarly. In *tei* BGCs, nine polygenic transcriptional units were found (Fig. 16.3f, Yushchuk et al. 2019), 8—in *dbv* (Fig. 16.3g, Alduina et al. 2007). In all these examples, genes for Hpg and Dpg biosynthesis form operons.

16.4 Biosynthetic Machinery Behind the Production of GPAs

Biosynthesis of all GPAs consists of three stages: (I) biosynthesis of non-proteinogenic amino acids; (II) non-ribosomal synthesis of peptide core, together with the halogenation and oxidative cross-linking; and (III) further tailoring steps of the aglycone, which may include a combination of methylation, sulfation, acylation, or glycosylation reactions.

16.4.1 "Secondary Metabolism" of Tyrosine

GPA aglyca usually require a pool of certain non-proteinogenic amino acids for their biosyntheses. These are L-β-hydroxytyrosine (β-Ht), L-4-hydroxyphenylglycine (Hpg), and L-3,5-dihydroxyphenylglycine (Dpg). Respective enzymes could be found in all GPA BGCs. First two of these non-proteinogenic amino acids are directly derived from tyrosine (Tyr), which also acts as the amino donor in Hpg and Dpg biosyntheses.

A majority of characterized to date GPA BGCs contain genes for key enzymes of the shikimate pathway, leading to the tyrosine: 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthases, chorismate mutases (CMs), and the prephenate dehydrogenases (PDHs) (Thykaer et al. 2010; Waglechner et al. 2019; Yushchuk et al. 2020c). Roles of these enzymes are summarized in Fig. 16.4a. Knockout of the DAHP synthase gene from balhimycin BGC in *Am. balhimycina* DSM 5908 led to decreased antibiotic production, whereas overexpression had an opposite effect. Neither knockout nor overexpression of PDH gene influenced balhimycin production (Thykaer et al. 2010). Not a single GPA BGC-encoded CM was studied to date experimentally.

16.4.2 Formation of β-Hydroxytyrosine

All GPAs contain a β-hydroxytyrosine as a part of aglycone. In A40926 biosynthetic pathway of *N. gerenzanensis*, nonheme dioxygenase (β-hydroxylase) Dbv28 is

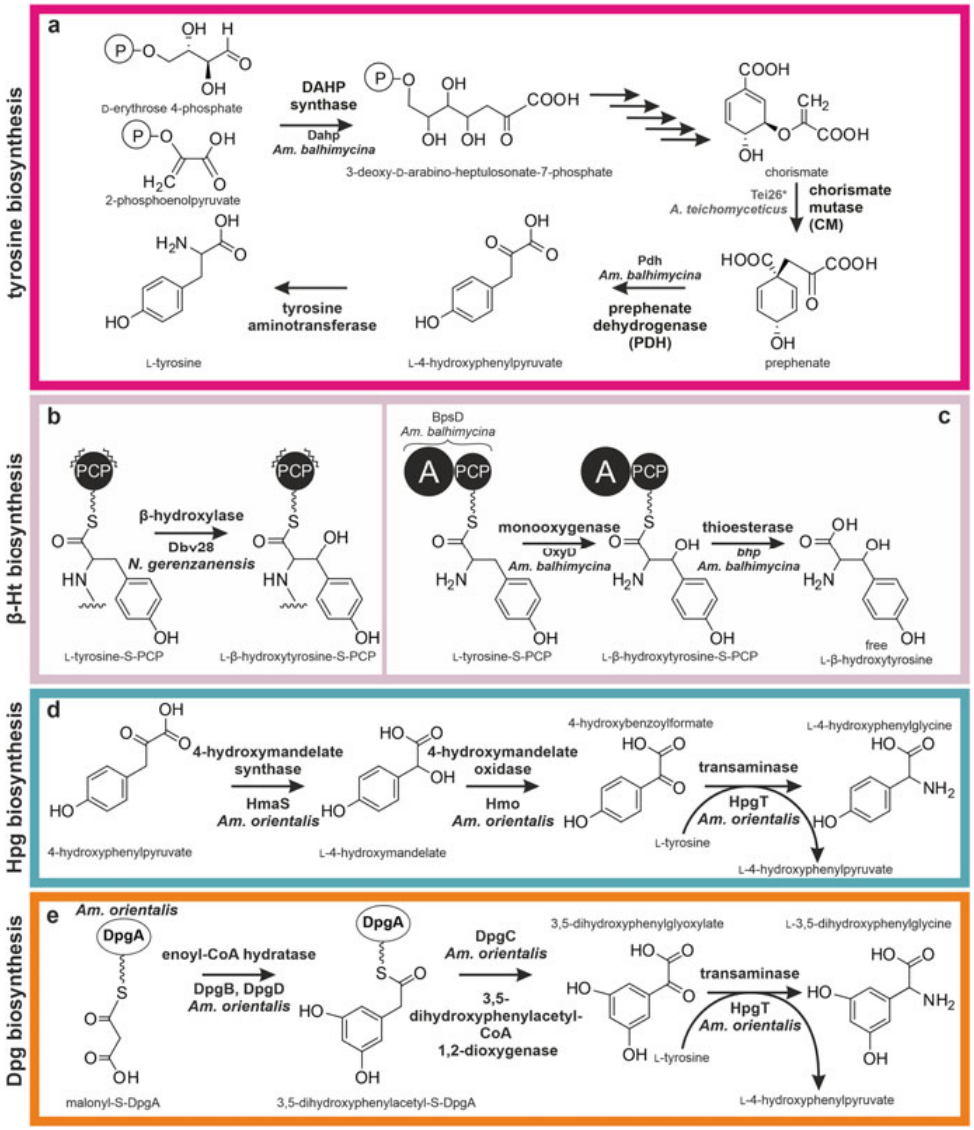


Fig. 16.4 Biosynthesis of tyrosine and non-proteinogenic amino acids, catalyzed by GPA BGC-encoded enzymes. The depicted enzymatic routes are taken from experimentally studied examples. (**a**) Shikimate pathway, at least three enzymes (DAHP synthase, PDH, and CM) are encoded within GPA BGCs. DAHP and PDH were studied experimentally in *Am. balhimycina*. CM Tei28* comes from *tei* BGC, but was not studied experimentally. There are two different pathways to L-β-hydroxytyrosine in A40926 biosynthesis (**b**) and chloroeremomycin biosynthesis (**c**). (**d**, **e**) Biosynthesis of Hpg and Dpg, respectively. Refer to the main text for more details

involved (Fig. 16.4b), acting on a tyrosine in a NRPS-bound aglycone. A different strategy is observed in balhimycin biosynthetic pathway, and most likely in the biosyntheses of the other related GPAs (Stegmann and Frasch 2010). Here, β-hydroxytyrosine is synthesized in free form in cytoplasm and only then is

incorporated in nascent peptide core (Mulyani et al. 2010). Three-gene operon—*bhp, bpsD,* and *oxyD*—is involved in this process (Mulyani et al. 2010; Recktenwald et al. 2002; Puk et al. 2004). L-tyrosine is first loaded on NRPS module BpsD and then hydroxylated by monooxygenase OxyD; thioesterase Bhp releases L-β-hydroxytyrosine from BpdD PCP domain (Fig. 16.4c).

16.4.3 Biosynthesis of Hpg and Dpg

Enzymes involved in Hpg biosynthesis were studied in *Am. orientalis*, chloroeremomycin producer (Hubbard et al. 2000; Li et al. 2001). HmaS converts 4-hydroxyphenylpyruvate (likely product of cluster-encoded PDH) into L-4-hydroxymandelate. The latter is oxidized by Hmo into L-4-hydroxybenzoylformate. HpgT uses an amino group of L-tyrosine to convert L-4-hydroxybenzoylformate into L-4-hydroxyphenylglycine (with L-4-hydroxyphenylpyruvate as a byproduct of this reaction, Fig. 16.4d).

Dpg biosynthesis was studied in chloroeremomycin and balhimycin producers (Chen et al. 2001; Pfeifer et al. 2001). Here, biosynthetic enzymes include DpgA, DpgB, DpgC, DpgD, and, once again, HpgT. DpgA is a type III polyketide synthase that forms a complex with the enoyl-CoA hydratases DpgB and DpgD, making 3,5-dihydroxyphenylacetate (tethered to DpgA) out of malonyl-CoA units (Fig. 16.4e). 3,5-dihydroxyphenylacetyl-CoA 1,2-dioxygenase DpgC converts 3,5-dihydroxyphenylacetate into 3,5-dihydroxyphenylglyoxylate. The final reaction of L-3,5-dihydroxyphenylglycine biosynthesis is catalyzed by a HpgT transaminase (Fig. 16.4e, Pfeifer et al. 2001; Sandercock et al. 2001). As in Hpg biosynthesis, L-tyrosine here serves as an amino donor and L-4-hydroxyphenylpyruvate is formed as a by-product (Fig. 16.4e). Some GPA BGCs, such as pekiskomycin ones, lack genes for DpgD (Thaker et al. 2013). This enzyme is probably dispensable for Dpg biosynthesis and is accessory to DpgB.

All aforementioned biosynthetic pathways are tightly intertwined. Products of shikimate pathway, L-tyrosine and L-4-hydroxyphenylpyruvate, serve either as direct precursors of non-proteinogenic amino acids, or take part in their biosynthesis. The final step of Dpg and Hpg biosyntheses also yields L-4-hydroxyphenylpyruvate, which can enter either the biosynthesis of L-tyrosine or Hpg. Finally, L-tyrosine itself is a part of the aglyca of GPAs such as teicoplanin and A40926. Tyrosine and Hpg/Dpg biosynthesis intermediates are also found within aglyca of some GPAs: 4-hydroxyphenylpyruvate in UK-68,597 (Yim et al. 2014) or 4-hydroxybenzoylformate in keratinimicins C and D (Xu et al. 2019) and complestatin (Chiu et al. 2001).

16.4.4 Non-ribosomal Biosynthesis of GPA Cores

NRPSs involved in the biosynthesis of dalbaheptides and type V GPAs exhibit typical modular organization (Fig. 16.5). Substrate specificities of corresponding

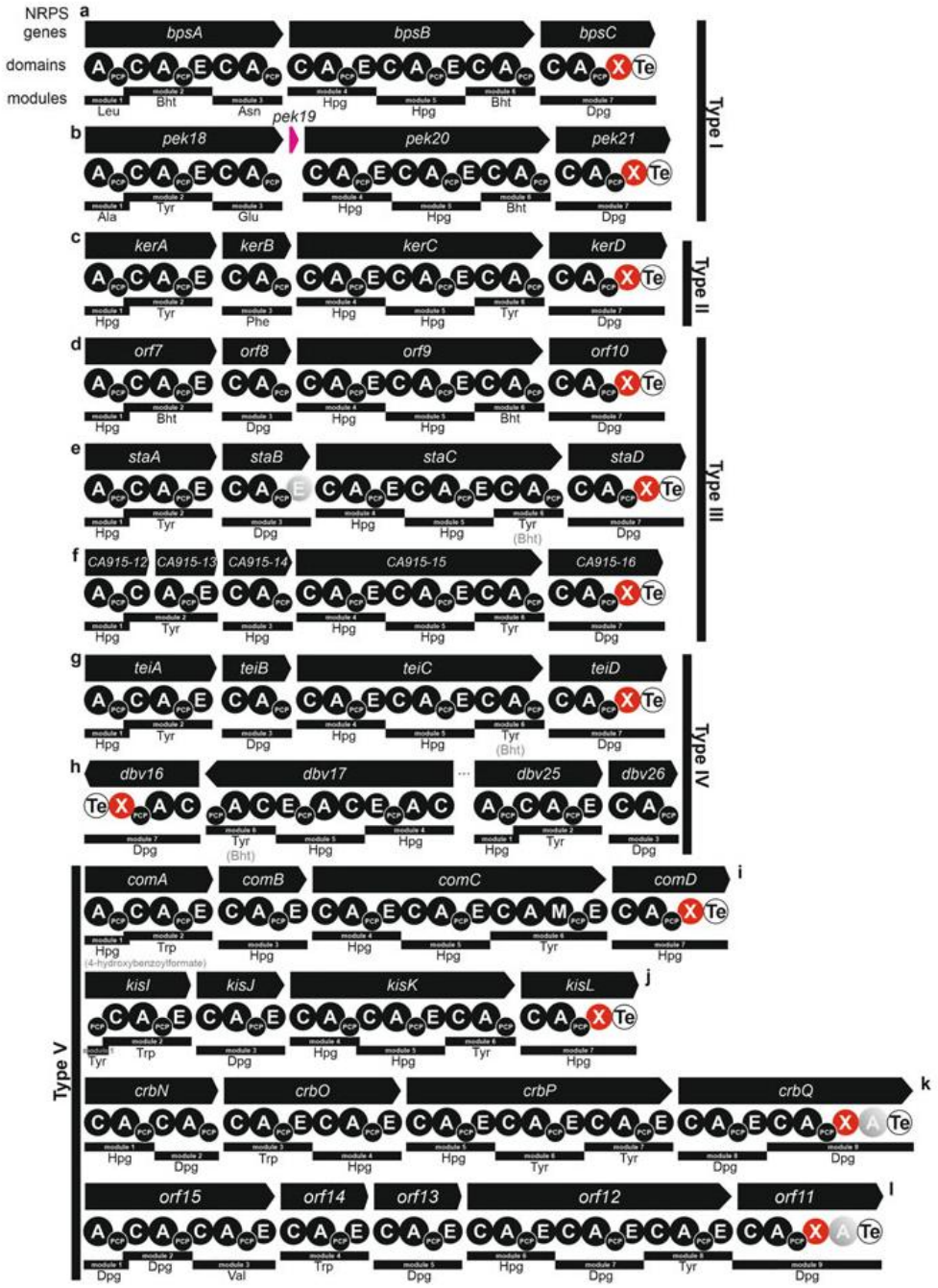


Fig. 16.5 Patterns in the organization of NRPSs (and corresponding genes) from GPA BGCs. Domains believed to be nonfunctional are shadowed. Refer to the main text for more details

adenylation (A) domains could be confidently predicted using in silico tools (Bachmann and Ravel 2009; Röttig et al. 2011).

Heptapeptide aglyca of dalbaheptides are genetically "imprinted" within seven modules of NRPS (Fig. 16.5a–h). These are encoded with three or four genes. In first case, e.g., in type I GPA balhimycin biosynthesis (Fig. 16.5a), genes *bpsA, bpsB,* and *bpsD* code 1st–3rd, 4th–6th, and 7th modules of NRPS, respectively (Shawky et al. 2007); in second, like in type II GPA keratinimicin biosynthesis (Fig. 16.5c), genes *kerA-D* (Xu et al. 2019) encode 1st–2nd, 3rd, 4th–6th, and 7th NRPS modules, respectively. One exception is the organization of NRPS genes in metagenome-derived CA915 BGC, where first three NRPS modules are encoded within single gene each (Fig. 16.5f, Banik et al. 2010). NRPS genes are located on one strand collinearly to the order of their function within NRPS assembly line. The *dbv* gene cluster is exception here: NRPS genes are on different strands and are separated by the other biosynthetic genes (Fig. 16.5h, Sosio et al. 2003). In *pek* BGC (Thaker et al. 2013), PDH gene (*pek19*) separates first two NRPS genes (*pek18* and *pek20*, Fig. 16.5b).

The domain organization of each module in dalbaheptide biosynthesis NRPSs remains quite conserved: 1st—(A)-(peptidyl carrier protein, PCP); 2nd—(condensation domain, C)-(A)-(PCP)-(epimerization domain, E); 3rd—(C)-(A)-(PCP); 4th—(C)-(A)-(PCP)-(E); 5th—(C)-(A)-(PCP)-(E); 6th—(C)-(A)-(PCP)-(E); and 7th—(C)-(A)-(PCP)-(X)-(thioesterase domain, Te).

Such an arrangement of E-domains should yield NH_2-L-D-L-D-D-L-L-COOH stereochemical configuration of the heptapeptide. However, this is not in agreement with experimentally solved NH_2-D-D-L-D-D-L-L-COOH configuration of all known dalbaheptides (e.g., Schäfer et al. 1996, 1998; Han et al. 2014). Such epimerization may arise from some orphan E-domain or epimerase, encoded outside of BGCs. However, recent in vitro experiments have shown that module 1 C-domain is likely involved in the epimerization of *aa-1* (Kaniusaite et al. 2021). A nonfunctional E-domain might sometimes be found in NRPS third module, like in A47934 NRPS StaB (Fig. 16.5e, Pootoolal et al. 2002).

Among all dalbaheptides, A-domain specificities in modules 2 (Tyr/Bht) and 4–7 (Hpg-Hpg-Tyr/Bht-Dpg) are the most conserved. On the contrary, A-domains of modules 1 and 3 might be responsible for the introduction of Leu and Asn (vancomycin, balhimycin), Ala and Gln (pekiskomycin), Hpg and Phe (keratinimicin), and Hpg and Dpg (type III and IV GPAs). In rare cases, *aa-1* position is occupied with a keto acid, whereas an amino acid is predicted in silico as a substrate for the corresponding A-domain. It might be due to substrate promiscuity of corresponding A-domain, or, alternatively, aglycone *aa-1* might be deaminated later on, although the exact mechanism is still to be unraveled.

Modular organizations of type V GPA NRPSs are more variable. For instance, E-domains are more common, giving different stereochemical patterns of aglyca in different compounds. An unusual truncated module 1, lacking A-domain, was found in kistamicin NRPS (Fig. 16.5j, Greule et al. 2019). It is so far unknown how the loading of first amino acid is achieved in kistamicin biosynthesis. Corbomycin NRPS starts with C-domain (Culp et al. 2020). Since corbomycin carries acetyl

group attached to N-terminus of its aglycone, it was proposed that as-yet-unknown polyketide synthase module initiates the biosynthesis of this antibiotic. It makes module 1 of CorA actually a hybrid PKS-NRPS, explaining the presence of C-domain. Complestatin NRPS features a unique methylation domain in module 6, responsible for the methylation of the amino group of tyrosine (Chiu et al. 2001). Finally, last modules of NRPSs of nonapeptide type V GPAs carry an additional, likely nonfunctional, A-domain (Fig. 16.5k, l).

16.4.5 Oxidative Cross-Linking and Halogenation of GPA Aglyca

Nascent GPA aglycone undergoes oxidative cross-linking while still being attached to the PCP domain of the last module of NRPS. There is invariant X-domain within all GPA NRPS (see Fig. 16.5) that plays a crucial role in this. It orchestrates the activity of P450 monooxygenases (Haslinger et al. 2015; Peschke et al. 2016b, c). For years, this domain was believed to be a nonfunctional C-domain. However, recent studies show that X-domain is required for the recruitment of cross-linking oxygenases to the non-ribosomal assembly line. The X-domain probably descended from an ancestral C-domain (Peschke et al. 2016b), which became catalytically inactive, but now serves as a docking site for the P450 monooxygenases (Haslinger et al. 2015; Peschke et al. 2016a, c). Indeed, Te domains of GPA NRPSs (Fig. 16.5) are specific toward cross-linked aglyca (Peschke et al. 2018), assisting the dissociation of fully processed aglycone from NRPS assembly line.

Type I and II GPAs feature three cross-links (Fig. 16.2). As investigated in *Am. balhimycina*, formation of these cross-links is achieved with P450 monooxygenases OxyA, OxyB, and OxyC (Fig. 16.6a, Stegmann et al. 2006). OxyB acts first, installing C-*O*-D cross-link, followed by OxyA (D-*O*-E) and OxyC (A-B). Type III and IV GPAs have fourth F-*O*-G cross-link, furnished by an additional P450 monooxygenase OxyE. Studies of oxygenases from *sta* and *tei* BGCs showed that OxyE acts after OxyB and before OxyA (Fig. 16.6b, Hadatsch et al. 2007; Cryle et al. 2011; Li et al. 2011). Thus, the number of cross-linking oxygenases encoded in dalbaheptide BGCs equals to the number of cross-links.

This, however, is different for type V GPAs: *kis* BGC, for instance, features two oxygenase genes, *kisN* and *kisO*, while the GPA has three cross-links (Nazari et al. 2017; Greule et al. 2019). KisN and KisO are related to dalbaheptide P450 oxygenases OxyA and OxyC, respectively (Fig. 16.6c). KisN installs D-E cross-link (although here it is biaryl bond, not bisaryl ether bond, Fig. 16.6c). KisO generates C-*O*-D and A-*O*-B cross-links (Fig. 16.6c). The sequence of the reactions is as follows: C-*O*-D > D-E > A-*O*-B (Greule et al. 2019).

Early works on *Am. balhimycina* (Puk et al. 2004) hinted that halogenation of GPA aglyca is coupled with non-ribosomal synthesis. This was supported with recent in vitro assays that have shown halogenases, such as Tei8* (responsible for the chlorination of *aa-2* and *aa-6*), acting only upon amino acids loaded onto the PCPs of the NRPS (Kittilä et al. 2017). This confirmed the aforementioned

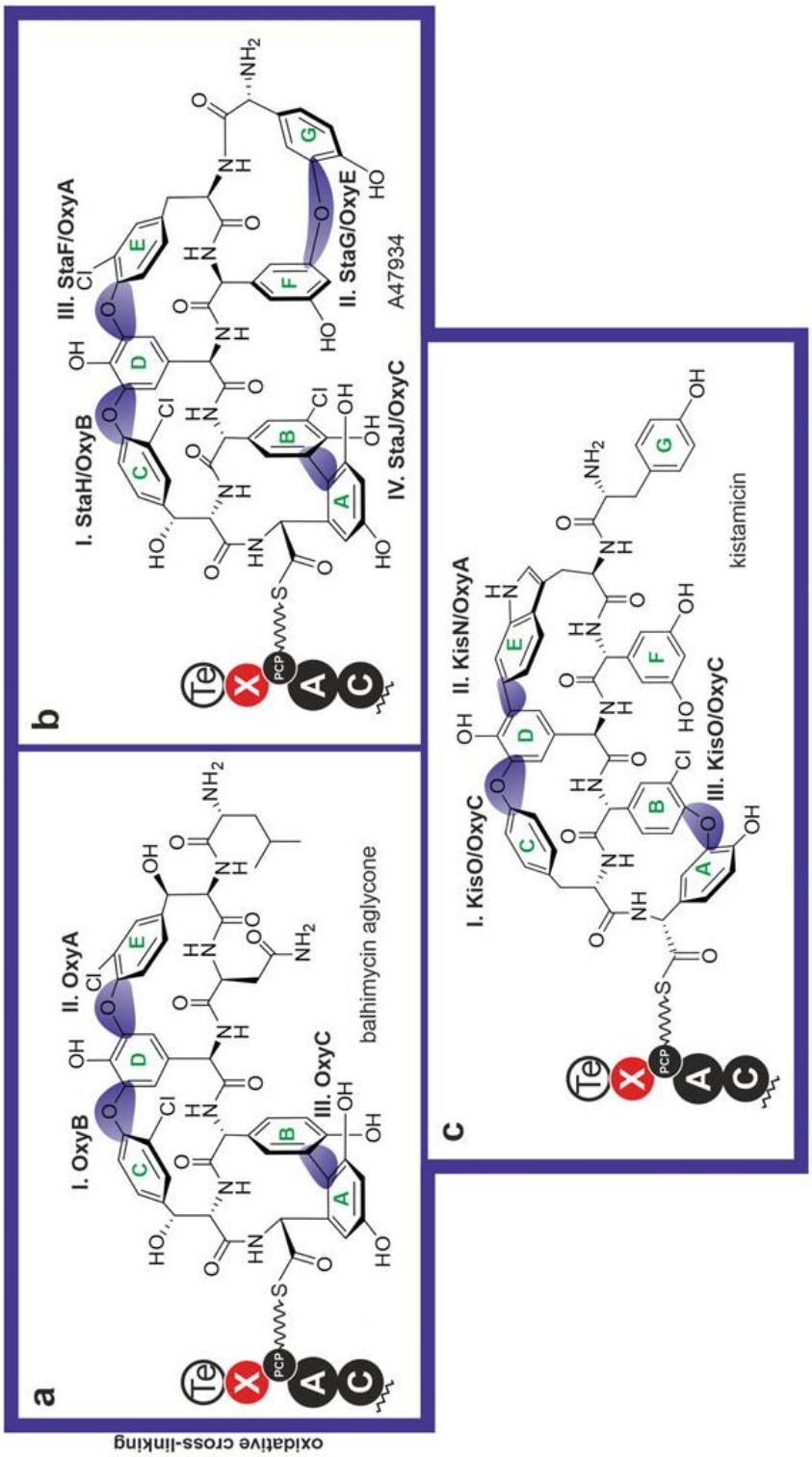


Fig. 16.6 Strategies of the oxidative cross-linking of NRPS-tethered GPA aglyca: (**a**) cross-linking of type I GPA, balhimycin; P450 monooxygenases act in following order: OxyB > OxyA > OxyC; (**b**) oxidative cross-linking of type III–IV GPAs exemplified with A47934, here one additional oxygenase acts between OxyB and OxyA; and (**c**) oxidative cross-linking of type V GPA kistamicin, featuring a unique dual-action P450 monooxygenase KisO

suggestion. Furthermore, a preference for chlorinated substrates was found for teicoplanin P450 monooxygenases (Peschke et al. 2017).

Most often among dalbaheptides, aglycones *aa-2* and *aa-6* are halogenated; corresponding BGCs carry a single halogenase gene responsible for this, like *bhaA* from *bal* (Puk et al. 2002) or *tei8** from *tei* (Li et al. 2004). The A40926 BGC carries a single *dbv10* halogenase gene sharing 81% of amino acid sequence identity to *bhaA* at protein level. Yet, A40926 carries chlorines on *aa-3* and *aa-6* (Sosio et al. 2003). A single halogenase gene within *pek* gene cluster (Thaker et al. 2013) is even more similar to BhaA (92%), but pekiskomycin is halogenated only at *aa-6*. The A47934 and *auk* BGCs carry two genes for halogenases, and corresponding antibiotics are halogenated at amino acids 2, 5, and 6 and 1, 2, 5, and 6, respectively. Recent work has shown that in A47934 biosynthetic pathway, StaK is inactive, and StaI alone carries out three chlorinations (Yim et al. 2018). Among type V GPAs, complestatin harbors six chlorines installed onto *aa-1*, *aa*-3, and *aa-5*. Such a "hyperchlorinated" pattern is likely achieved with a single BGC-encoded halogenase ComH. Other type V GPA, kistamicin, carries a single chlorine on *aa-5*. At the same time, *kis*-encoded halogenase—KisU—is 72% identical to ComH. Thus, specificities of GPA halogenases might be inferred from in silico analysis very cautiously.

GPA halogenases might also decorate GPA aglyca with halogens other than chlorine; feeding of *Am. balhimycina* with bromine yielded bromobalhimycin (Bister et al. 2003). Some GPAs (ristocetin, corbomycin, GP6738) lack halogen atoms. Corresponding BGCs have no halogenase genes (Spohn et al. 2014; Truman et al. 2014; Culp et al. 2020; Xu et al. 2020).

16.4.6 Methylation and Sulfation of GPA Aglyca

Aglyca of dalbaheptides are often, but not necessarily, methylated. Most commonly, N-terminal end of oligopeptide is methylated with an (*S*)-adenosyl-L-methionine (SAM)-dependent *N*-methyltransferases, initially studied on the basis of MtfA from chloroeremomycin BGC (O'Brien et al. 2000; Shi et al. 2009). MtfA has a very low affinity to a non-cross-linked heptapeptide, suggesting that *N*-methylation occurs on mature aglycone as a substrate (O'Brien et al. 2000; Brieke et al. 2016). Some *N*-methyltransferases methylate N-terminus of heptapeptide twice (like Pek28), or even three times (possibly in demethylvancomycin producer WAC4169, Thaker et al. 2013). *N*-methyltransferase genes are present in BGCs for vancomycin-like antibiotics, pekiskomycin and A40926. The *tei* and *sta* gene clusters, however, lack such genes. Yet, *N*-methyltransferases from *dbv, pek,* and *cem* BGCs are promiscuous enough to *N*-methylate A40926 and teicoplanin aglyca (Shi et al. 2009; Brieke et al. 2016). Interestingly, Pek28 was unable to methylate N-terminus of these aglyca twice, losing substrate specificity after the first round (Brieke et al. 2016). Metagenome-derived VEG BGC contains two genes for *N*-methyltransferases—*veg24* and *veg25* (Banik and Brady 2008).

Methylation of GPA aglyca is not limited to N-terminus of heptapeptide. Ristocetin aglycone features C-methylated *aa-3* (Dpg) and *O*-methylated C-terminus of the heptapeptide. Orf19 and Orf23 methyltransferases were predicted to methylate these sites (Truman et al. 2014). Function of Orf19 and Orf23 was confirmed later via heterologous expression in *S. coelicolor*, carrying *sta* gene cluster (Yim et al. 2016). Both enzymes are unique for GPAs; Orf19 has Veg18 as the only ortholog from metagenome-derived VEG BGC. Accordingly, Veg18 is able to methylate *aa-3* of A47934 in vitro (Banik and Brady 2008).

Sulfations of the GPA aglyca are much rarer than methylations. Among naturally occurring GPAs, only A47934, UK-68,897, and pekiskomycin are sulfated at *aa-1, aa-3,* and *aa-6,* respectively. Sulfotransferase genes are abundant in metagenome-derived GPA BGCs, mainly in TEG BGC (Banik and Brady 2008). Many of them were investigated in vitro and in vivo. StaL sulfates teicoplanin aglycone in vitro (Lamb et al. 2006) at the *aa-1* (as it naturally happens in A47934 biosynthesis). Auk20 sulfates A47934 at *aa-3* in vivo and vancomycin- and teicoplanin-derived compounds in vitro (Yim et al. 2014). Pek25 sulfates A47934 at *aa-6* in vivo (Yim et al. 2016). TEG-derived sulfotransferases offered more interesting case, being able to sulfate teicoplanin aglycone at *aa-3* (Teg12), *aa-4* (Teg14), and *aa-6* (Teg13). Crystal structures of StaL, Teg14, and Teg12 were published (Bick et al. 2010; Shi et al. 2012). We are not aware of type V GPAs carrying sulfate group naturally or as a result of genetic engineering.

16.4.7 Glycosylation and Acylation of Dalbaheptides

As mentioned above, eponymous tailoring steps are limited to type I-IV GPAs, and only type IV compounds carry aliphatic side chains. All dalbaheptides (with a very few exceptions, like A47934) are glycosylated. Each amino acid in aglycone could be glycosylated (except, probably, *aa-5,* Nicolaou et al. 1999). The astonishing variability of glycosylation patterns remains poorly understood at the genetic background. Nevertheless, glycosylation of some dalbaheptides was studied in great detail.

Among type I dalbaheptides, mechanisms of glycosylation are best known for chloroeremomycin, vancomycin, and balhimycin (Fig. 16.7a). All three antibiotics are glucosylated at *aa-4*. Chloroeremomycin also carries two residues of exotic amino sugar L-epivancosamine—one at *aa-6* and second attached to D-glucose; balhimycin has another amino sugar, L-4-oxovancosamine, at *aa-6*. The *aa-6* position of vancomycin is not glycosylated; L-vancosamine residue is appended to D-glucose (Fig. 16.7a).

Corresponding BGCs harbor different sets of glycosyltransferase genes: three in *cep* gene cluster—GtfA, GtfB, and GtfC; and three in *bal* gene cluster—BgtfA, BgtfB, and BgtfC. Vancomycin BGC encodes two glycosyltransferases—GtfD and GtfE. GtfB/GtfE transfers D-glucose residue onto *aa-4*, and then, GtfC/D attach L-epivancosamine/L-vancosamine to D-glucose, respectively (Fig. 16.7a, Losey et al. 2001, 2002; Mulichak et al. 2004). GtfA installs L-epivancosamine onto *aa-6*

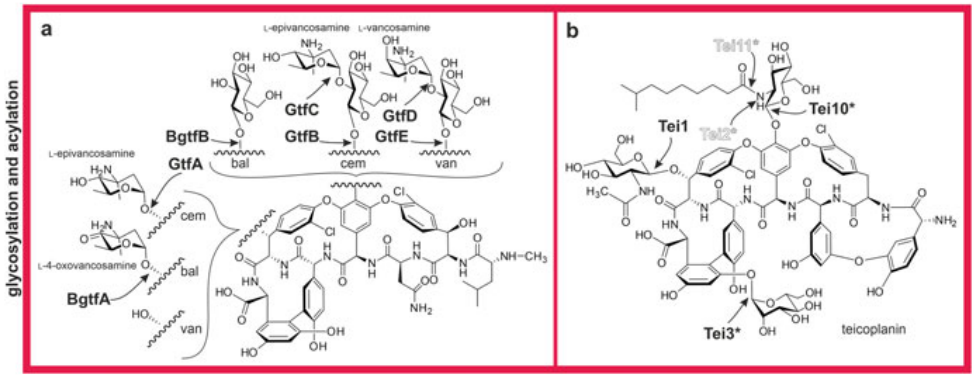


Fig. 16.7 Some of the best studied GPA glycosylation (and acylation) cascades: (**a**) glycosylation cascades for type I GPAs (balhimycin bal, chloroeremomycin cem, and vancomycin van); (**b**) coupled glycosylation/acylation of type IV GPA teicoplanin. Refer to the main text for more details

(Fig. 16.7a, Mulichak et al. 2003; Lu et al. 2004). All these enzymes are quite promiscuous with regard to carbohydrate donor substrates and acceptor aglyca, albeit retaining strict regiospecificity. This allows combinatorial glycosylation of vancomycin and teicoplanin aglyca. Function of BgtfB was studied in vivo. The *bgtfB* knockout mutant fails to produce any glycosylated GPA, implying that BgtfB catalyzes initial reaction of glycosylation cascade (Fig. 16.7a, Pelzer et al. 1999). BgtfC is practically unable to work on balhimycin scaffold (Pelzer et al. 1999; Stegmann and Frasch 2010).

Experimental data on glycosyltransferases from type **II** dalbaheptides are absent, although these GPAs are often decorated at unusual positions, like avoparcin (glycosylated at *aa-1* and *aa-2*), or carry exotic amino sugars, such as L-acosamine in actinoidin (Nicolaou et al. 1999).

Glycosylation is ubiquitous among type **III** dalbaheptides. A good example is ristocetin, decorated with six sugar moieties, including common D-arabinose, D-glucose, L-rhamnose, two D-mannose residues, and exotic amino sugar L-ristosamine (Truman et al. 2014). The ristocetin BGC carries six genes for glycosyltransferases (Truman et al. 2014). Functions of some genes were investigated in a heterologous strain *S. coelicolor* producing A47934 and desulfo-A47934 (Yim et al. 2016), showing that Orf22 is responsible for D-mannosylation of *aa-7*.

Glycosylation of type **IV** dalbaheptides, such a teicoplanin, is coupled with acylation. The teicoplanin aglycone carries three sugar residues and one lipid chain (Fig. 16.7b); correspondingly, three glycosyltransferases (Tei1, Tei3*, and Tei10*) and one acyltransferase (Tei11*) are encoded within *tei* BGC. Glycosylation and acylation cascade in teicoplanin biosynthesis begins with Tei10* transferring Glc*N*Ac to *aa-4* (Fig. 16.7b, Li et al. 2004). Glc*N*Ac is then deacetylated by Tei2* yielding a glucosaminyl moiety (Truman et al. 2006). Tei11* then transfers an acyl chain to the free amino group of this glucosaminyl moiety (Fig. 16.7b, Li et al. 2004; Kruger et al. 2005; Lyu et al. 2014). Tei11* is quite flexible toward the donor substrates, cell membrane lipids. It leads to the production of different teicoplanin

congeners, varying in length and branching of lipid chains (Lazzarini et al. 1997). Inactivation of *tei11** leads to accumulation of a teicoplanin derivative lacking *N*-acetylglucosamine on *aa-6* and D-mannose on *aa-7* (Yushchuk et al. 2016). In vitro characterization of Tei1 showed that it is responsible for transferring Glc*N*Ac moiety to *aa-6* (Fig. 16.7b, Li et al. 2004), while knockout of *tei3** showed that Tei3* attaches a D-mannose to *aa-7* (Fig. 16.7b, Yushchuk et al. 2016). Another type IV dalbaheptide, A40926, undergoes tailoring similarly to teicoplanin. Here, Dbv9 and Dbv20 are homologues of Tei10* and Tei3*, attaching Glc*N*Ac to *aa-4* and mannose to *aa-7*, respectively. Glc*N*Ac is then deacetylated by Dbv21 (Tei2* homologue) and acylated by Dbv8 (Tei11* homologue). However, two additional enzymes, absent in teicoplanin biosynthesis pathway, are involved in further modifications of A40926 sugars. These are Glc*N*Ac oxidase Dbv29 (which forms carboxylic group on Glc*N*Ac moiety), and Dbv23, which acetylates D-mannose residue.

The biosynthesis of unusual amino sugars (mentioned above) is directed by BGC-situated genes and was thoroughly elucidated on the example of L-epivancosamine production in chloroeremomycin pathway (Fig. 16.8, Chen et al. 2000). This pathway consists of five (EvaA-E) enzymes and serves as a reference to understand amino sugar biosynthesis for all GPAs (Fig. 16.8).

16.5 Transcriptional Regulation of the GPA Biosynthetic Gene Expression

All described to date BGCs for dalbaheptides and complestatin-like antibiotics possess a gene for a StrR-like transcriptional regulator (Waglechner et al. 2019; Yushchuk et al. 2019). These regulators act as a key transcriptional activators at lowest—pathway-specific—regulatory level. There is deep understanding of StrR homologues from teicoplanin, balhimycin, and A40926 BGCs (Shawky et al. 2007; Horbal et al. 2014; Lo Grasso et al. 2015). StrR family regulators are not the only transcriptional factors (TFs) encoded within GPA BGCs. The *tei* and *dbv* gene clusters carry also regulatory gene for a LuxR-like TF. The latter play central role in GPA biosynthesis. Type V GP6738 (Xu et al. 2020) BGC also contains a *lnbU*-like TF, overexpression of which increases the production of GP6738. Majority of GPA BGCs also harbor *vanR-vanS* genes involved only in regulation of self-resistance (not discussed here).

16.5.1 Pathway-Specific Regulation in Balhimycin BGC (and Related BGCs)

In *Am. balhimycina*, balhimycin BGC encodes StrR-like protein Bbr (Shawky et al. 2007). Bbr DNA-binding sites (BSs) are found within the promoter regions of *tba*, *oxyA*, *orf7*, *dvaA*, and *bbr* itself, and consensus sequence is **[GTCCAa** $(N)_{17}$**TtGGAC]** (base pairs corresponding to the inverted repeats here and further

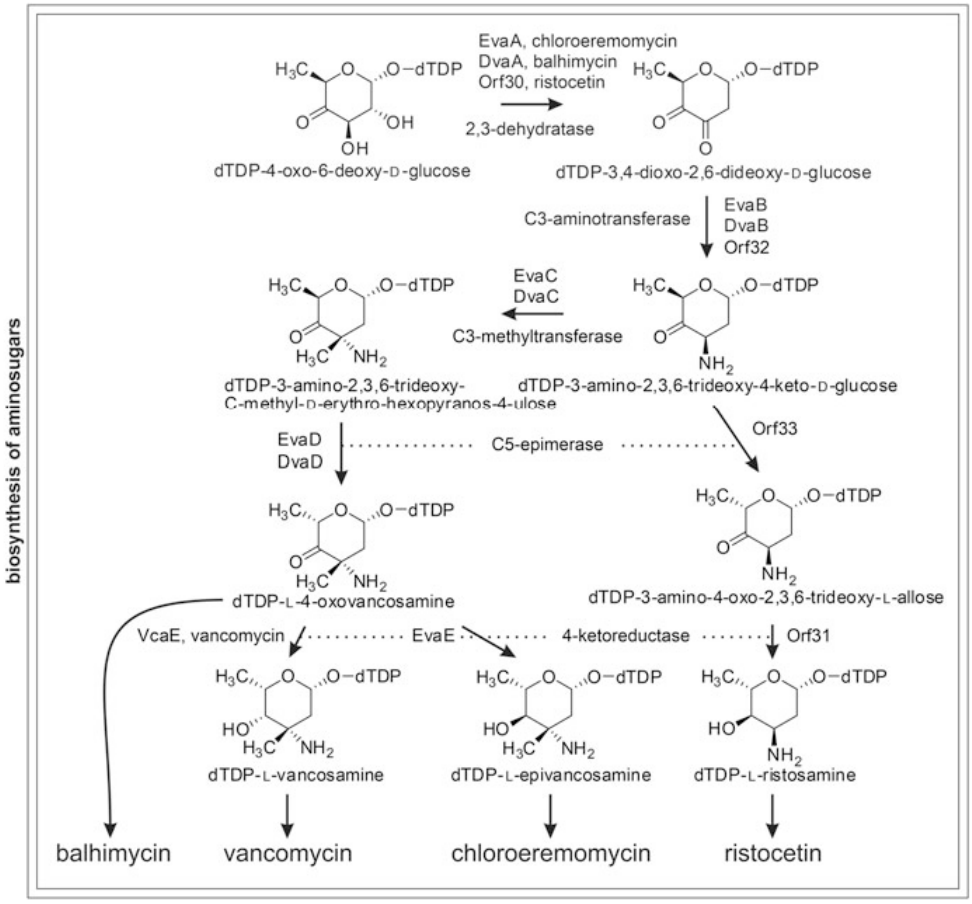


Fig. 16.8 Biosynthetic pathways that enable the production of some GPA amino sugars. The biosynthesis of L-epivancosamine was investigated experimentally in chloroeremomycin pathway; enzymes, required for the biosynthesis of L-ristosamine and others, were not studied experimentally, but their functions were predicted by analogy

are given in bold, Shawky et al. 2007). Bbr does not bind *vanY* promoter region, leaving the expression of biosynthetic and resistance genes without a link.

Balhimycin BGC is a model for studying the biosynthesis of clinically relevant vancomycin, since the producer of the latter poses formidable challenge for genetic manipulations. Recent work on the regulation of norvancomycin biosynthesis in *Am. orientalis* NCPC 2–48 (Li et al. 2021) provides collateral insight into Bbr-mediated regulation of balhimycin production. The norvancomycin BGC—*nvcm*—was described to be regulated by Bbr ortholog AoStrR1. Balhimycin BGC and *nvcm* are very similar. However, DNA-binding properties of AoStrR1 were drastically different from that of Bbr. Particularly, AoStrR1 was shown to bind at least 8 promoter regions within *nvcm*: *AoStrR1* (corresponding to *bbr* promoter region, where Bbr binds), *oxyA* (also recognized by Bbr), *oxyB* (Bbr binding to the homologous sequence was not observed), *vasA* (*dvaA* in balhimycin BGC, where Bbr

binds), *vanY* (not recognized by Bbr), *vhp* (corresponds to *bhp* in balhimycin BGC; Bbr does not bind), *hmaS* (same in balhimycin BGC; Bbr does not bind), and *ald* (corresponding to *dahp* in balhimycin BGC, where no binding was observed for Bbr). Some of them contain almost perfect BSs matching the consensus described for Bbr, some have rather degenerate sites. The consensus BS of AoStrR1 was thus much more degenerate than the one for Bbr—**[(G/T)NNCNN**$(N)_{17}$**(T/C)N(G/C)(G/C)(A/G)C**] (Li et al. 2021). Given the 84% amino acid sequence identity between Bbr and AoStrR1, the promiscuity of the latter is mind boggling. Perhaps, some AoStrR1 BSs are artificial (especially the most degenerate ones) and the binding occurs only under conditions of the in vitro assay. One of such nonphysiological BSs might be in the region upstream of *oxyB*, which is most likely a part of *oxyA-B-C-vhaA* polycistronic transcription unit (by analogy to the *oxyA-B-C-bhaA* operon in balhimycin BGC).

The same work (Li et al. 2021) also described another regulator, found on the outer boundary of *nvcm* BGC to be involved in norvancomycin biosynthesis. This TF contains a LuxR-like DNA-binding domain (completely different from the ones found in *dbv/tei*-derived LuxR-like TFs). Its overexpression led to the increase in norvancomycin production.

16.5.2 Pathway-Specific Regulation of tei BGC

Transcriptional regulation of teicoplanin biosynthesis in *A. teichomyceticus* was recently discussed in Yushchuk et al. (2020c). In brief, *tei* BGC contains three genes, coding for TFs: *tei15** (StrR-like), *tei16** (LuxR-like), and *tei31** (AfsR-like) (Li et al. 2004; Horbal et al. 2014; Yushchuk et al. 2019). Tei15* and Bbr are orthologs sharing 48% of amino acid sequence identity. Knockout of *tei15** canceled teicoplanin production (Horbal et al. 2014), and overexpression of the same gene improved teicoplanin titers up to 40-fold (Horbal et al. 2013, 2014). As Bbr, DNA-binding properties of Tei15* were studied in vitro. Tei15* BSs were found upstream *tei2–3* (*vanRS* orthologs), *teiA* (first NRPS gene), *tei2** (Glc*N*Ac deacetylase), *tei5** (OxyA), *tei16** (LuxR-like TF), *tei17** (DpgA), *tei23** (HpgT), *tei26** (type II CM), *tei28** (HmaS), *tei30** (type II thioesterase), and *tei31** (AfsR-like TF). Six of these promoter regions (*teiA, tei2*, tei16*, tei17*, tei27*,* and *tei31**) were bound in vitro by Tei15*; the others were not tested. Tei15* BS consensus sequence is **[(G/C)NN(C/G)NN**$(N)_{17}$**NT(G/C)G(A/G)C**] (Horbal et al. 2014), being more degenerate than the one for Bbr. Expression of *tei1, teiA-1** (one operon, Tei15* binding verified), *tei2*-3** (operon, Tei15* binding verified), *tei4*, tei5*-7** (operon, Tei15* BS predicted in silico), *tei8*-12** (operon, Tei15* BS predicted in silico), *tei13*, tei14*, tei24*,* and *tei28*-30** (operon, Tei15* BS predicted in silico) is decreased in Δ*tei15** mutant. At the same time, the expression of *tei2–4, tei17*-23** operons, and *tei16** was rather intact. Corresponding promoter regions contained either predicted in silico or verified in vitro Tei15* BS. Some genes having ample Tei15* BS in their promoter regions were not expressed at all in the wild type under teicoplanin production conditions. These results make the

picture of Tei15*-mediated regulation of teicoplanin biosynthesis quite puzzling, highlighting the gap in our understanding of DNA-binding properties of TFs under in vitro and in vivo conditions.

Tei16* is a second pathway-specific regulator of teicoplanin biosynthesis. The knockout of *tei16** completely abolishes the production of teicoplanin (Horbal et al. 2014); the overexpression of this gene boosts teicoplanin titers more that it was observed for *tei15** overexpression (Horbal et al. 2013, 2014). Studies of Tei16* DNA-binding properties in vitro shed no light on its targets within *tei* BGC (Horbal et al. 2014). Tei16* likely requires some partner proteins or ligands to bind DNA (like other LuxR-like TFs, de Schrijver and de Mot 1999). The transcriptional changes within *tei* gene cluster of Δ*tei16** and Δ*tei15** mutants were similar, including the complete cessation of *tei15** expression (Yushchuk et al. 2019). The *tei15** is likely to be a single target of Tei16*, thus placing the latter at the top of *tei* pathway-specific regulatory hierarchy. As Tei15* binds *tei16** promoter region in vitro, some feedback mechanism may exist, which so far remains unstudied. Knockouts of both *tei15** and *tei16** have no impact on *tei2–4* and *tei7–5* operons for teicoplanin resistance, similarly to what was reported for balhimycin BGC.

TF Tei31* appears to be dispensable for teicoplanin biosynthesis under teicoplanin production conditions (Yushchuk et al. 2019). However, Tei15* binds *tei31** promoter in vitro; so, Tei31* might have some effect on teicoplanin biosynthesis under certain conditions.

16.5.3 Pathway-Specific Regulation of dbv BGC

The A40926 (*dbv*) BGC of *N. gerenzanensis* encodes Dbv4 TF, which is Bbr/Tei15* ortholog (Sosio et al. 2003). LuxR-like Dbv3 is also found in *dbv* BGC (Sosio et al. 2003). A two-component regulatory system consisting of a sensor histidine kinase Dbv22 and response TF Dbv6 was very recently shown to be involved in A40926 production as well (Alduina et al. 2020).

Dbv4 possesses rather small regulon in *dbv* BGC as compared to Bbr and Tei15* regulons. Dbv4 directly controls the expression of *dbv30–35* operon (Dpg biosynthesis genes) and *dbv14–8* suboperon (cross-linking oxygenase and some tailoring enzyme genes). This was shown both by studying Dbv4 DNA-binding properties in vitro (Alduina et al. 2007) and by the analysis of the *dbv* gene expression in Δ*dbv4* mutant (Lo Grasso et al. 2015). DNA-binding properties of Dbv4 and Bbr are remarkably similar (Alduina et al. 2007), underscoring overall identity of these proteins. The promoter regions of *dbv30* and *dbv14* contain BSs identical to consensus BSs described for Bbr; moreover, Dbv4 was able to bind Bbr-regulated promoter regions within balhimycin BGC. Although the regulon of Dbv4 is smaller than the regulons of Bbr and Tei15*, it governs crucial A40926 biosynthetic genes, as *dbv4* knockout blocks A40926 production (Lo Grasso et al. 2015), while overexpression improves the latter (Yushchuk et al. 2020a).

Dbv3 is a second major TF in *dbv* BGC and a crucial one for A40926 biosynthesis, because *dbv3* knockout abolishes the production of this GPA (Lo Grasso et al.

2015), while overexpression improves the yield (Yushchuk et al. 2020a). Dbv3 targets within *dbv* BGC are known only through gene expression studies (Lo Grasso et al. 2015). Following *dbv* operons are not expressed in Δ*dbv3* mutant (Lo Grasso et al. 2015): *dbv1–2* (Hpg biosynthetic enzymes), *dbv15–17* (NRPS modules 4–7 and MbtH), *dbv21–21* (mannosyltransferase and mannose *O*-acetylase), and *dbv24–28* (NRPS modules 1–3, A40926 export, and some tailoring enzymes). The expression of *dbv29* (hexose oxidase), *dbv36* (type II thioesterase), and *dbv37* (HpgT) genes is absent in Δ*dbv3* mutant, as well as the expression of *dbv4*. Thus, unlike Tei16*, Dbv3 has a multiple *dbv* targets, except *dbv4*. It is not clear whether Dbv3 regulates the expression of *dbv* genes alone, or some yet-unknown intermediate player is involved.

A40926 BGC encodes a two-component regulatory system—*dbv6* and *dbv22*. Given its similarity to VanR/S regulatory circuit, Dbv6/22 was previously believed to be orthologs of the latter (Yushchuk et al. 2020b). When *dbv6* was first knocked out, no impact was observed on either A40926 production or other physiological properties of *N. gerenzanensis* (Lo Grasso et al. 2015). Dbv6 and Dbv22 were shown recently to be accessory regulators of A40926 biosynthesis (Alduina et al. 2020). Dbv6 binds promoter regions of *dbv4, dbv3,* and *dbv24–28* operons (also controlled by Dbv3), and *dbv22–23* operon (controlled by neither Dbv4 nor Dbv3). Moreover, it somehow binds *dbv7* upstream region, although the latter gene was previously shown to be part of polycistronic *dbv5–6–7* operon (Alduina et al. 2007). At least under some growth conditions (in RARE3 medium, but not in R3 medium, Lo Grasso et al. 2015; Alduina et al. 2020), Dbv6 and Dbv22 act as a A40926 biosynthesis repressors. At the same time, Δ*dbv22* mutant (but not Δ*dbv6*) was slightly more sensitive to A40926 and its precursor—*O*-acetyl-A40926. Additionally, A40926 (but not *O*-acetyl-A40926) was shown to stimulate the expression of *dbv* genes, maybe through Dbv22. In this scenario, Dbv22 acts as a sensor histidine kinase, which represses certain *dbv* genes (likely through Dbv6); Dbv22 displays greater affinity to A40926 than to *O*-acetyl-A40926 and loses its repressing function upon A40926 binding. Although solid, evidence-based scenario is still unavailable, one thing is certain—Dbv6/Dbv22 regulatory pair is involved in fine-tuning of the *dbv* gene expression. Being accessory to Dbv4 and Dbv3, Dbv6/Dbv22 likely couples antibiotic biosynthesis and self-resistance. Kistamicin BGC also carries *dbv6/dbv22* orthologs, in addition to the ones for *dbv3/dbv4* (Nazari et al. 2017).

To sum up, there currently are three "archetypal" models for pathway-specific regulation of dalbaheptide biosynthesis. The best understood model applies to *Amycolatopsis*-derived antibiotics. Here, a StrR family Bbr-like TF governs the biosynthesis of the majority of biosynthetic genes and operons. Its binding sites are firmly identified. All Bbr-like regulators are very similar and function across different biosynthetic pathways (Spohn et al. 2014). In contrast, the regulation of teicoplanin and, most likely, other *Actinoplanes*-derived GPAs (Yim et al. 2014) involves two regulators. One is StrR/Bbr type, exemplified by Tei15*, and the latter is a subject to action of LuxR-like TF Tei16*. Tei15* binds multiple promoters within *tei* BGC, recognizing, however, a more degenerate binding sequence than Bbr does. Gene *tei15** is likely a single target of Tei16*, but might upregulate the

expression of *tei16** as well. Factors contributing to the function of Tei16* are unknown. In *dbv* biosynthesis regulation, a StrR and LuxR regulators Dbv4 and Dbv3 are also focal. However, here LuxR-like TF Dbv3 is not an orthologue of Tei16*. Dbv4 is closely related to Bbr, but is responsible for the activation of only two *dbv* operons.

16.6 Biological Activities of GPAs: Beyond D-Ala-D-Ala Binding

In contrast to the other GPAs, type V GPAs, such as complestatin and corbomycin, do not bind D-Ala-D-Ala termini but are still active against Gram-positive bacteria, most likely because they target autolysins (Culp et al. 2020). Autolysins are a large group of lytic enzymes, involved in cell wall remodeling necessary for growth and division (Smith and Blackman 2000). Pharmacological properties of both dalbaheptides and complestatin-like antibiotics are not limited to the antimicrobial activities, as described below.

Anticancer properties Most of them show certain cytotoxic activities against cancer cells in vitro, but clinical prospects are uncertain (e.g., Balzarini et al. 2003; Xu et al. 2019). Vancomycin was found to potentiate antitumor immune response to the radiotherapy, thus inhibiting tumor growth (Uribe-Herranz et al. 2020). This effect was achieved not by vancomycin per se, but by depletion of vancomycin-sensitive gut microbiota.

Antiviral properties Complestatin inhibits the development of HIV-1 in different models in vitro (Momota et al. 1991). Complestatin appears to prevent the adsorption of viral particles on the cell surface. Antiviral properties were also described for kistamicin (Naruse et al. 1993a, b). Other type V GPAs, chloropeptins I and II, possess activity against HIV (Tanaka et al. 1997). Among dalbaheptides, only those carrying aliphatic side chain (type IV)—teicoplanin and A40926—exhibit antiviral properties (Balzarini et al. 2003). Heavily glycosylated dalbaheptides tended to lose antiviral potency and vice versa—vancomycin aglycone has weak antiviral effect (Balzarini et al. 2003). Teicoplanin inhibits the infection of cell lines with Ebolavirus in vitro (Wang et al. 2016; Zhou et al. 2016). Teicoplanin was effective only before viral entry, hinting that the cell surface, not the Ebola virus itself, is drug target. Teicoplanin is completely inefficient against non-enveloped human respiratory syncytial virus (hSCV) (Wang et al. 2016). It is now known that teicoplanin targets cathepsin L protease, which is located in the host cell membrane. Cathepsin L is involved in the proteolysis of the envelope glycoproteins, which is necessary for the entry of Ebolavirus into the host cell (e.g., Hood et al. 2010). Activity of teicoplanin against a number of other enveloped viruses was shown. Hydrophobic groups present in teicoplanin, dalbavancin, oritavancin, and telavancin are probably necessary for the proper anchoring of the antibiotic within host cell membrane, which therefore delivers the pharmacophore (probably located in aglycone) to cathepsin L. Activity of teicoplanin and related drugs against coronaviruses like SARS-CoV-1

and MERS-CoV is noticeable in times of pandemics caused by SARS-CoV-1 and SARS-CoV-2 (Wu et al. 2020).

Activity of dalbaheptides against SARS-CoV-2 might not be only cathepsin-mediated but also ACE2-dependent (Wang et al. 2021); the latter protein is the entry receptor for the SARS-CoV-2 (Shang et al. 2020). ACE2 tropism was shown for dalbavancin and not vancomycin. The presence of the aliphatic side chain in the dalbavancin is the most salient structural difference between abovementioned GPAs.

Usage of ristocetin in von Willebrand disease diagnostics Ristocetin induces thrombophlebitis and thrombocytopenia when administered as antibacterial, which prompted its withdrawal from clinical usage. These properties, however, find their use in the diagnostics of von Willebrand disease (James and Goodeve 2011), as well as rarer Bernard–Soulier syndrome (Lanza 2006). Ristocetin causes agglutination of thrombocytes mediated by von Willebrand factor multimers. Thus, addition of the ristocetin to a blood sample, lacking this factor (or corresponding receptor on the thrombocyte surface, as it is in the case of Bernard–Soulier syndrome), will not cause agglutination of thrombocytes, which could be detected (Coller and Gralnick 1977).

Some other "exotic properties" of complestatin Complestatin is one of the most potent inhibitors of sensitized erythrocytes hemolysis, caused by a complement system. This inhibition is irreversible and, most likely, is achieved due to the inhibition of C3 convertase (Kaneko et al. 1989). This GPA also prevents neuronal excitotoxicity caused by *N*-methyl-D-aspartate (NMDA) and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic (AMPA) acid. Thus, complestatin protects cortical neurons from prolonged oxygen and glucose starvation (Seo et al. 2001). Finally, complestatin has certain antiapoptotic properties (Kim et al. 2004).

16.7 Conclusions and Future Prospects

In this chapter, we have focused on a rather "mechanistic" description of the GPA family. However, many other aspects of the biology of this peculiar compounds merit detailed discussion and more research as well. One of such aspects is relationship of GPAs with other peptide inhibitors of PG synthesis, such as feglymycin (Gonsior et al. 2015) and ramoplanin (Farver et al. 2005). Feglymycin BGC—*feg*—has some genetic elements highly related to GPA BGCs, especially *tei*. These include two regulatory genes for StrR- and LuxR-like orthologs of Tei15* and Tei16*. Indeed, the discovery of corbomycin and GP6738 makes it easier to imagine how a GPA-like BGC might have evolved into *feg*-like one through the rearrangement of NRPS genes and loss of cross-linking oxygenases. Alternatively, *feg* might represent a sister evolutionary line to GPAs, coming from common ancestral pool of BGCs carrying basic NRPS genes, Dpg and Hpg biosynthetic genes, regulatory elements, and transporters. Ramoplanin BGC, at the same time, codes for a mannosyltransferase with a striking similarity to mannosyltransferases involved in

the GPA biosynthesis, as well as some GPA-like amino acid synthesis genes. Recent works on the phylogeny of GPAs (Waglechner et al. 2019) clearly highlight these links. However, a bigger picture of the evolution of GPAs together with related compounds still remains elusive. There is no doubt that further reconstruction of this "bigger picture" will yield novel classes of antibiotics related to GPAs, novel tailoring enzymes, and data on the biosynthetic regulatory mechanisms. Altogether, this will contribute to our arsenal against multidrug-resistant pathogens.

References

Adamek M, Alanjary M, Sales-Ortells H et al (2018) Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. BMC Genomics 19: 426. https://doi.org/10.1186/s12864-018-4809-4

Alduina R, Lo Piccolo L, D'Alia D et al (2007) Phosphate-controlled regulator for the biosynthesis of the dalbavancin precursor A40926. J Bacteriol 189:8120–8129. https://doi.org/10.1128/JB.01247-07

Alduina R, Sosio M, Donadio S (2018) Complex regulatory networks governing production of the glycopeptide A40926. Antibiotics (Basel) 7(2):30. https://doi.org/10.3390/antibiotics7020030

Alduina R, Tocchetti A, Costa S et al (2020) A two-component regulatory system with opposite effects on glycopeptide antibiotic biosynthesis and resistance. Sci Rep 10:6200. https://doi.org/10.1038/s41598-020-63257-4

Arthur M, Courvalin P (1993) Genetics and mechanisms of glycopeptide resistance in enterococci. Antimicrob Agents Chemother 37(8):1563–1571. https://doi.org/10.1128/aac.37.8.1563

Arthur M, Molinas C, Bugg TDH et al (1992) Evidence for *in vivo* incorporation of D-lactate into peptidoglycan precursors of vancomycin-resistant enterococci. Antimicrob Agents Chemother 36(4):867–869. https://doi.org/10.1128/aac.36.4.867

Bachmann BO, Ravel J (2009) Methods for *in silico* prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. Methods Enzymol 458:181–217. https://doi.org/10.1016/S0076-6879(09)04808-3

Balzarini J, Pannecouque C, De Clercq E et al (2003) Antiretroviral activity of semisynthetic derivatives of glycopeptide antibiotics. J Med Chem 46:2755–2764. https://doi.org/10.1021/jm0300882

Banik JJ, Brady SF (2008) Cloning and characterization of new glycopeptide gene clusters found in an environmental DNA megalibrary. Proc Natl Acad Sci U S A 105(45):17273–17277. https://doi.org/10.1073/pnas.0807564105

Banik JJ, Craig JW, Calle PY, Brady SF (2010) Tailoring enzyme-rich environmental DNA clones: a source of enzymes for generating libraries of unnatural natural products. J Am Chem Soc 132: 15661–15670. https://doi.org/10.1021/ja105825a

Barna JC, Williams DH (1984) The structure and mode of action of glycopeptide antibiotics of the vancomycin group. Annu Rev Microbiol 38:339–357. https://doi.org/10.1146/annurev.mi.38.100184.002011

Barreteau H, Kovač A, Boniface A et al (2008) Cytoplasmic steps of peptidoglycan biosynthesis. FEMS Microbiol Rev 32(2):168–207. https://doi.org/10.1111/j.1574-6976.2008.00104.x

Bick MJ, Banik JJ, Darst SA, Brady SF (2010) The 2.7 Å resolution structure of the glycopeptide sulfotransferase Teg14. Acta Crystallogr D Biol Crystallogr 66(Pt 12):1278–1286. https://doi.org/10.1107/S0907444910036681

Binda E, Marinelli F, Marcone GL (2014) Old and new glycopeptide antibiotics: action and resistance. Antibiotics (Basel) 3(4):572–594. https://doi.org/10.3390/antibiotics3040572

Bister B, Bischoff D, Nicholson GJ et al (2003) Bromobalhimycin and chlorobromobalhimycins—illuminating the potential of halogenases in glycopeptide antibiotic biosyntheses. ChemBioChem 4(7):658–662. https://doi.org/10.1002/cbic.200300619

Boeck LD, Mertz FP (1986) A47934, a novel glycopeptide-aglycone antibiotic produced by a strain of *Streptomyces toyocaensis* taxonomy and fermentation studies. J Antibiot (Tokyo) 39:1533–1540. https://doi.org/10.7164/antibiotics.39.1533

Brieke C, Yim G, Peschke M et al (2016) Catalytic promiscuity of glycopeptide: N-methyltransferases enables bio-orthogonal labelling of biosynthetic intermediates. Chem Commun (Camb) 52(94):13679–13682. https://doi.org/10.1039/c6cc06975d

Cavalleri B, Parenti F (2000) Glycopeptides (Dalbaheptides). In: Kirk-Othmer encyclopedia of chemical technology. https://doi.org/10.1002/0471238961.0712250303012201.a01

Chen H, Thomas MG, Hubbard BK et al (2000) Deoxysugars in glycopeptide antibiotics: enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis. Proc Natl Acad Sci U S A 97(22):11942–11947. https://doi.org/10.1073/pnas.210395097

Chen H, Tseng CC, Hubbard BK, Walsh CT (2001) Glycopeptide antibiotic biosynthesis: enzymatic assembly of the dedicated amino acid monomer (S)-3,5-dihydroxyphenylglycine. Proc Natl Acad Sci U S A 98(26):14901–14906. https://doi.org/10.1073/pnas.221582098

Cheng AV, Wuest WM (2020) Phylogeny-guided approach yields glycopeptides with unique action. Trends Pharmacol Sci 41(5):297–299. https://doi.org/10.1016/j.tips.2020.03.002

Chiu HT, Hubbard BK, Shah AN et al (2001) Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. Proc Natl Acad Sci U S A 98(15):8548–8553. https://doi.org/10.1073/pnas.151246498

Coller BS, Gralnick HR (1977) Studies on the mechanism of ristocetin induced platelet agglutination. Effects of structural modification of ristocetin and vancomycin. J Clin Invest 60:302–312. https://doi.org/10.1172/JCI108778

Courvalin P (2006) Vancomycin resistance in gram-positive cocci. Clin Infect Dis 42(Suppl 1): S25–S34. https://doi.org/10.1086/491711

Cryle MJ, Staaden J, Schlichting I (2011) Structural characterization of CYP165D3, a cytochrome P450 involved in phenolic coupling in teicoplanin biosynthesis. Arch Biochem Biophys 507 (1):163–173. https://doi.org/10.1016/j.abb.2010.10.017

Culp EJ, Waglechner N, Wang W et al (2020) Evolution-guided discovery of antibiotics that inhibit peptidoglycan remodelling. Nature 578:582–587. https://doi.org/10.1038/s41586-020-1990-9

de Kruijff B, van Dam V, Breukink E (2008) Lipid II: a central component in bacterial cell wall synthesis and a target for antibiotics. Prostaglandins Leukot Essent Fatty Acids 79 (3–5):117–121. https://doi.org/10.1016/j.plefa.2008.09.020

de Schrijver A, de Mot R (1999) A subfamily of MalT-related ATP-dependent regulators in the LuxR family. Microbiology 145:1287–1288. https://doi.org/10.1099/13500872-145-6-1287

Farver DK, Hedge DD, Lee SC (2005) Ramoplanin: a lipoglycodepsipeptide antibiotic. Ann Pharmacother 39(5):863–868. https://doi.org/10.1345/aph.1E397

Goldstein BP, Selva E, Gastaldo L et al (1987) A40926, a new glycopeptide antibiotic with anti-*Neisseria* activity. Antimicrob Agents Chemother 31:1961–1966. https://doi.org/10.1128/AAC.31.12.1961

Gonsior M, Mühlenweg A, Tietzmann M et al (2015) Biosynthesis of the peptide antibiotic feglymycin by a linear nonribosomal peptide synthetase mechanism. ChemBioChem 16 (18):2610–2614. https://doi.org/10.1002/cbic.201500432

Greule A, Izoré T, Iftime D et al (2019) Kistamicin biosynthesis reveals the biosynthetic requirements for production of highly crosslinked glycopeptide antibiotics. Nat Commun 10 (1):2613. https://doi.org/10.1038/s41467-019-10384-w

Hadatsch B, Butz D et al (2007) The biosynthesis of teicoplanin-type glycopeptide antibiotics: assignment of P450 mono-oxygenases to side chain cyclizations of glycopeptide A47934. Chem Biol 14(9):1078–1089. https://doi.org/10.1016/j.chembiol.2007.08.014

Li TL, Huang F, Haydock SF et al (2004) Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase. Chem Biol 11:107–119. https://doi.org/10.1016/S1074-5521(04)00002-X

Li Z, Rupasinghe SG, Schuler MA, Nair SK (2011) Crystal structure of a phenol-coupling P450 monooxygenase involved in teicoplanin biosynthesis. Proteins 79(6):1728–1738. https://doi.org/10.1002/prot.22996

Li X, Zhang C, Zhao Y et al (2021) Comparative genomics and transcriptomics analyses provide insights into the high yield and regulatory mechanism of norvancomycin biosynthesis in *Amycolatopsis orientalis* NCPC 2-48. Microb Cell Fact 20(1):28. https://doi.org/10.1186/s12934-021-01521-6

Liu Y, Breukink E (2016) The membrane steps of bacterial cell wall synthesis as antibiotic targets. Antibiotics (Basel) 5(3):28. https://doi.org/10.3390/antibiotics5030028

Lo Grasso L, Maffioli S, Sosio M et al (2015) Two master switch regulators trigger A40926 biosynthesis in *Nonomuraea* sp. strain ATCC 39727. J Bacteriol 197:2536–2544. https://doi.org/10.1128/JB.00262-15

Losey HC, Peczuh MW, Chen Z et al (2001) Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides. Biochemistry 40 (15):4745–4755. https://doi.org/10.1021/bi010050w

Losey HC, Jiang J, Biggins JB et al (2002) Incorporation of glucose analogs by GtfE and GtfD from the vancomycin biosynthetic pathway to generate variant glycopeptides. Chem Biol 9 (12):1305–1314. https://doi.org/10.1016/s1074-5521(02)00270-3

Lu W, Oberthür M, Leimkuhler C et al (2004) Characterization of a regiospecific epivancosaminyl transferase GtfA and enzymatic reconstitution of the antibiotic chloroeremomycin. Proc Natl Acad Sci U S A 101(13):4390–4395. https://doi.org/10.1073/pnas.0400277101

Lyu SY, Liu YC, Chang CY et al (2014) Multiple complexes of long aliphatic N-acyltransferases lead to synthesis of 2,6-diacylated/2-acyl-substituted glycopeptide antibiotics, effectively killing vancomycin-resistant enterococcus. J Am Chem Soc 136(31):10989–10995. https://doi.org/10.1021/ja504125v

Manat G, Roure S, Auger R et al (2014) Deciphering the metabolism of undecaprenyl-phosphate: the bacterial cell-wall unit carrier at the membrane frontier. Microb Drug Resist 20(3):199–214. https://doi.org/10.1089/mdr.2014.0035

Marcone GL, Beltrametti F, Binda E et al (2010) Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727. Antimicrob Agents Chemother 54:2465–2472. https://doi.org/10.1128/AAC.00106-10

Marcone GL, Binda E, Berini F, Marinelli F (2018) Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era. Biotechnol Adv 36(2):534–554. https://doi.org/10.1016/j.biotechadv.2018.02.009

Marshall CG, Broadhead G, Leskiw BK, Wright GD (1997) D-Ala-D-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB. Proc Natl Acad Sci U S A 94(12):6480–6483. https://doi.org/10.1073/pnas.94.12.6480

Marshall CG, Lessard IAD, Park IS, Wright GD (1998) Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. Antimicrob Agents Chemother 42(9):2215–2220. https://doi.org/10.1128/AAC.42.9.2215

Momota K, Kaneko I, Kimura S et al (1991) Inhibition of human immunodeficiency virus type-1-induced syncytium formation and cytopathicity by complestatin. Biochem Biophys Res Commun 179:243–250. https://doi.org/10.1016/0006-291X(91)91361-F

Mulichak AM, Losey HC, Lu W et al (2003) Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. Proc Natl Acad Sci U S A 100 (16):9238–9243. https://doi.org/10.1073/pnas.1233577100

Mulichak AM, Lu W, Losey HC et al (2004) Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. Biochemistry 43(18):5170–5180. https://doi.org/10.1021/bi036130c

Mulyani S, Egel E, Kittel C et al (2010) The thioesterase Bhp is involved in the formation of β-hydroxytyrosine during balhimycin biosynthesis in *Amycolatopsis balhimycina*. ChemBioChem 11:266–271. https://doi.org/10.1002/cbic.200900600

Naruse N, Oka M, Konishi M, Oki T (1993a) New antiviral antibiotics, kistamicins a and b II. Structure determination. J Antibiot (Tokyo) 46:1812–1818. https://doi.org/10.7164/antibiotics.46.1812

Naruse N, Tenmyo O, Kobaru S et al (1993b) New antiviral antibiotics, kistamicins a and b I. taxonomy, production, isolation, physico-chemical properties and biological activities. J Antibiot (Tokyo) 46:1804–1811. https://doi.org/10.7164/antibiotics.46.1804

Nazari B, Forneris CC, Gibson MI et al (2017) *Nonomuraea* sp. ATCC 55076 harbours the largest actinomycete chromosome to date and the kistamicin biosynthetic gene cluster. Medchemcomm 8:780–788. https://doi.org/10.1039/c6md00637j

Nicolaou KC, Boddy CNC, Bräse S, Winssinger N (1999) Chemistry, biology, and medicine of the glycopeptide antibiotics. Angew Chem Int Ed Engl 38(15):2096–2152. https://doi.org/10.1002/(sici)1521-3773(19990802)38:15<2096::aid-anie2096>3.0.co;2-f

O'Brien DP, Kirkpatrick PN, O'Brien SW et al (2000) Expression and assay of an N-methyltransferase involved in the biosynthesis of a vancomycin group antibiotic. Chem Commun 1:103–104. https://doi.org/10.1039/a907953j

Owen JG, Reddy BVB, Ternei MA et al (2013) Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. Proc Natl Acad Sci U S A 110(29):11797–11802. https://doi.org/10.1073/pnas.1222159110

Parenti F, Cavalleri B (1989) Proposal to name the vancomycin-ristocetin like glycopeptides as dalbaheptides. J Antibiot 42:1882–1883. https://doi.org/10.7164/antibiotics.42.1882

Pelzer S, Süßmuth R, Heckmann D et al (1999) Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908. Antimicrob Agents Chemother 43(7):1565–1573. https://doi.org/10.1128/AAC.43.7.1565

Peschke M, Brieke C, Cryle MJ (2016a) F-*O*-G ring formation in glycopeptide antibiotic biosynthesis is catalysed by OxyE. Sci Rep 6:35584. https://doi.org/10.1038/srep35584

Peschke M, Gonsior M, Süssmuth RD, Cryle MJ (2016b) Understanding the crucial interactions between cytochrome P450s and non-ribosomal peptide synthetases during glycopeptide antibiotic biosynthesis. Curr Opin Struct Biol 41:46–53. https://doi.org/10.1016/j.sbi.2016.05.018

Peschke M, Haslinger K, Brieke C et al (2016c) Regulation of the P450 oxygenation cascade involved in glycopeptide antibiotic biosynthesis. J Am Chem Soc 138:6746–6753. https://doi.org/10.1021/jacs.6b00307

Peschke M, Brieke C, Goode RJA et al (2017) Chlorinated glycopeptide antibiotic peptide precursors improve cytochrome p450-catalyzed cyclization cascade efficiency. Biochemistry 56(9):1239–1247. https://doi.org/10.1021/acs.biochem.6b01102

Peschke M, Brieke C, Heimes M, Cryle MJ (2018) The thioesterase domain in glycopeptide antibiotic biosynthesis is selective for cross-linked aglycones. ACS Chem Biol 13:110–120. https://doi.org/10.1021/acschembio.7b00943

Pfeifer V, Nicholson GJ, Ries J et al (2001) A polyketide synthase in glycopeptide biosynthesis. The biosynthesis of the non-proteinogenic amino acid (S)-3,5-dihydroxyphenylglycine. J Biol Chem 276:38370–38377. https://doi.org/10.1074/jbc.M106580200

Pootoolal J, Thomas MG, Marshall CG et al (2002) Assembling the glycopeptide antibiotic scaffold: the biosynthesis of A47934 from *Streptomyces toyocaensis* NRRL15009. Proc Natl Acad Sci U S A 99:8962–8967. https://doi.org/10.1073/pnas.102285099

Puk O, Huber P, Bischoff D et al (2002) Glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908: function of a halogenase and a haloperoxidase/perhydrolase. Chem Biol 9 (2):225–235. https://doi.org/10.1016/s1074-5521(02)00101-1

Puk O, Bischoff D, Kittel C et al (2004) Biosynthesis of chloro-β-hydroxytyrosine, a nonproteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics. J Bacteriol 186:6093–6100. https://doi.org/10.1128/JB.186.18.6093-6100.2004

Recktenwald J, Shawky R, Puk O et al (2002) Nonribosomal biosynthesis of vancomycin-type antibiotics: a heptapeptide backbone and eight peptide synthetase modules. Microbiology (Reading) 148(Pt 4):1105–1118. https://doi.org/10.1099/00221287-148-4-1105

Röttig M, Medema MH, Blin K et al (2011) NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res 39:W362–W367. https://doi.org/10.1093/nar/gkr323

Sandercock AM, Charles EH, Scaife W et al (2001) Biosynthesis of the di-meta-hydroxyphenylglycine constituent of the vancomycin-group antibiotic chloroeremomycin. Chem Biol 7(12):931–942. https://doi.org/10.1016/s1074-5521(00)00043-0

Sarkar P, Yarlagadda V, Ghosh C, Haldar J (2017) A review on cell wall synthesis inhibitors with an emphasis on glycopeptide antibiotics. Medchemcomm 8(3):516–533. https://doi.org/10.1039/c6md00585c

Schäfer M, Schneider TR, Shieldrick GM (1996) Crystal structure of vancomycin. Structure 4:1509–1515. https://doi.org/10.1016/S0969-2126(96)00156-6

Schäfer M, Sheldrick GM, Schneider TR, Vértesy L (1998) Structure of balhimycin and its complex with solvent molecules. Acta Crystallogr D Biol Crystallogr 54:175–183. https://doi.org/10.1107/S0907444997008895

Seo SY, Yun BS et al (2001) Complestatin is a noncompetitive peptide antagonist of N-methyl-D-aspartate and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid/kainate receptors: secure blockade of ischemic neuronal death. J Pharmacol Exp Ther 299(1):377–384

Shang J, Ye G, Shi K et al (2020) Structural basis of receptor recognition by SARS-CoV-2. Nature 581:221–224. https://doi.org/10.1038/s41586-020-2179-y

Shawky RM, Puk O, Wietzorrek A et al (2007) The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains *bbr*, encoding a StrR-like pathway-specific regulator. J Mol Microbiol Biotechnol 13:76–88. https://doi.org/10.1159/000103599

Shi R, Lamb SS, Zakeri B et al (2009) Structure and function of the glycopeptide N-methyltransferase MtfA, a tool for the biosynthesis of modified glycopeptide antibiotics. Chem Biol 16(4):401–410. https://doi.org/10.1016/j.chembiol.2009.02.007

Shi R, Munger C, Kalan L et al (2012) Sulfonation of glycopeptide antibiotics by sulfotransferase StaL depends on conformational flexibility of aglycone scaffold. Proc Natl Acad Sci U S A 109(29):11824–11829. https://doi.org/10.1073/pnas.1205377109

Silhavy TJ, Kahne D, Walker S (2010) The bacterial cell envelope. Cold Spring Harb Perspect Biol 2(5):a000414. https://doi.org/10.1101/cshperspect.a000414

Smith TJ, Blackman SA (2000) Autolysins of *Bacillus subtilis*: multiple enzymes with multiple functions. Microbiology 146:249–262. https://doi.org/10.1099/00221287-146-2-249

Sosio M, Stinchi S, Beltrametti F et al (2003) The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species. Chem Biol 10:541–549. https://doi.org/10.1016/S1074-5521(03)00120-0

Spohn M, Kirchner N, Kulik A et al (2014) Overproduction of ristomycin a by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. Antimicrob Agents Chemother 58:6185–6196. https://doi.org/10.1128/AAC.03512-14

Stegmann E, Frasch HJ (2010) Glycopeptide biosynthesis in the context of basic cellular functions. Curr Opin Microbiol 13:595–602. https://doi.org/10.1016/j.mib.2010.08.011

Stegmann E, Pelzer S, Bischoff D et al (2006) Genetic analysis of the balhimycin (vancomycin-type) oxygenase genes. J Biotechnol 124(4):640–653. https://doi.org/10.1016/j.jbiotec.2006.04.009

Tanaka H, Matsuzaki K, Nakashima H et al (1997) Chloropeptins, new anti-HIV antibiotics inhibiting GP120-CD4 binding from *Streptomyces* sp. I. Taxonomy, fermentation, isolation, and physico-chemical properties and biological activities. J Antibiot 50:58–65. https://doi.org/10.7164/antibiotics.50.58

Thaker MN, Wang W, Spanogiannopoulos P et al (2013) Identifying producers of antibacterial compounds by screening for antibiotic resistance. Nat Biotechnol 31:922–927. https://doi.org/10.1038/nbt.2685

Thykaer J, Nielsen J, Wohlleben W et al (2010) Increased glycopeptide production after overexpression of shikimate pathway genes being part of the balhimycin biosynthetic gene cluster. Metab Eng 12:455–461. https://doi.org/10.1016/j.ymben.2010.05.001

Truman AW, Robinson L, Spencer JB (2006) Identification of a deacetylase involved in the maturation of teicoplanin. ChemBioChem 7(11):1670–1675. https://doi.org/10.1002/cbic.200600308

Truman AW, Kwun MJ, Cheng J et al (2014) Antibiotic resistance mechanisms inform discovery: identification and characterization of a novel *Amycolatopsis* strain producing ristocetin. Antimicrob Agents Chemother 58:5687–5695. https://doi.org/10.1128/AAC.03349-14

Uribe-Herranz M, Rafail S, Beghi S et al (2020) Gut microbiota modulate dendritic cell antigen presentation and radiotherapy-induced antitumor immune response. J Clin Invest 130:466–479. https://doi.org/10.1172/JCI124332

Vollmer W, Blanot D, De Pedro MA (2008) Peptidoglycan structure and architecture. FEMS Microbiol Rev 32(2):149–167. https://doi.org/10.1111/j.1574-6976.2007.00094.x

Waglechner N, McArthur AG, Wright GD (2019) Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. Nat Microbiol 4:1862–1871. https://doi.org/10.1038/s41564-019-0531-5

Wang Y, Cui R, Li G et al (2016) Teicoplanin inhibits Ebola pseudovirus infection in cell culture. Antivir Res 125:1–7. https://doi.org/10.1016/j.antiviral.2015.11.003

Wang G, Yang ML, Duan ZL et al (2021) Dalbavancin binds ACE2 to block its interaction with SARS-CoV-2 spike protein and is effective in inhibiting SARS-CoV-2 infection in animal models. Cell Res 31:17–24. https://doi.org/10.1038/s41422-020-00450-0

Wu F, Zhao S, Yu B et al (2020) A new coronavirus associated with human respiratory disease in China. Nature 579:265–269. https://doi.org/10.1038/s41586-020-2008-3

Xu F, Wu Y, Zhang C et al (2019) A genetics-free method for high-throughput discovery of cryptic microbial metabolites. Nat Chem Biol 15:161–168. https://doi.org/10.1038/s41589-018-0193-2

Xu M, Wang W, Waglechner N et al (2020) GPAHex-A synthetic biology platform for Type IV–V glycopeptide antibiotic production and discovery. Nat Commun 11(1):5232. https://doi.org/10.1038/s41467-020-19138-5

Yim G, Kalan L, Koteva K et al (2014) Harnessing the synthetic capabilities of glycopeptide antibiotic tailoring enzymes: characterization of the UK-68, 597 biosynthetic cluster. ChemBioChem 15:2613–2623. https://doi.org/10.1002/cbic.201402179

Yim G, Wang W, Thaker MN et al (2016) How to make a glycopeptide: a synthetic biology approach to expand antibiotic chemical diversity. ACS Infect Dis 2(9):642–650. https://doi.org/10.1021/acsinfecdis.6b00105

Yim G, Wang W, Pawlowski AC, Wright GD (2018) Trichlorination of a teicoplanin-type glycopeptide antibiotic by the halogenase StaI evades resistance. Antimicrob Agents Chemother 62(12):e01540–e01518. https://doi.org/10.1128/AAC.01540-18

Yushchuk O, Ostash B, Pham TH et al (2016) Characterization of the post-assembly line tailoring processes in teicoplanin biosynthesis. ACS Chem Biol 11:2254–2264. https://doi.org/10.1021/acschembio.6b00018

Yushchuk O, Horbal L, Ostash B et al (2019) Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. Appl Microbiol Biotechnol 103:4089–4102. https://doi.org/10.1007/s00253-019-09789-w

Yushchuk O, Andreo-Vidal A, Marcone GL et al (2020a) New molecular tools for regulation and improvement of A40926 glycopeptide antibiotic production in *Nonomuraea gerenzanensis* ATCC 39727. Front Microbiol 11:8. https://doi.org/10.3389/fmicb.2020.00008

Yushchuk O, Binda E, Marinelli F (2020b) Glycopeptide antibiotic resistance genes: distribution and function in the producer actinomycetes. Front Microbiol 11:1173. https://doi.org/10.3389/fmicb.2020.01173

Yushchuk O, Ostash B, Truman AW et al (2020c) Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes. Appl Microbiol Biotechnol 104: 3279–3291. https://doi.org/10.1007/s00253-020-10436-y

Yushchuk O, Vior NM, Andreo-Vidal A et al (2021) Genomic-led discovery of a novel glycopeptide antibiotic by *Nonomuraea coxensis* DSM 45129. ACS Chem Biol 16(5):915–928. https://doi.org/10.1021/acschembio.1c00170

Zeng D, Debabov D, Hartsell TL et al (2016) Approved glycopeptide antibacterial drugs: mechanism of action and resistance. Cold Spring Harb Perspect Med 6(12):a026989. https://doi.org/10.1101/cshperspect.a026989

Zhou N, Pan T, Zhang J et al (2016) Glycopeptide antibiotics potently inhibit cathepsin l in the late endosome/lysosome and block the entry of ebola virus, middle east respiratory syndrome coronavirus (MERS-CoV), and severe acute respiratory syndrome coronavirus (SARS-CoV). J Biol Chem 291:9218–9232. https://doi.org/10.1074/jbc.M116.716100

Оригінал доступний за посиланням:
https://link.springer.com/chapter/10.1007/978-981-16-6132-7_16.

Анотація: *Глікопептидні антибіотики (ГПА), що порушують біосинтез клітинної стінки Грам-позитивних бактерій, залишаються однією із найуспішніших родин антибіотиків. ГПА продукуються Actinobacteria, класом Грам-позитивних мікроорганізмів зі складними життєвими циклами і великими геномами. Вивчення біосинтезу ГПА є важливим з кількох причин. По-перше, продуценти ГПА, ймовірно, є основним джерелом відповідних генів стійкості; їх дослідження дасть змогу передбачити можливі шляхи розвитку антибіотикорезистентності серед патогенів. По-друге, дослідження біосинтезу дають інформацію про регуляторні механізми, що контролюють продукцію ГПА, які можна використовувати в біотехнології. По-третє, маніпуляції з біосинтетичними генами ГПА є життєздатним підходом до створення нових ГПА із практичною цінністю. Нарешті, інформація про гени ГПА може бути використана для пошуку нових сполук шляхом біоінформатичного аналізу загальнодоступних геномних даних. У цьому розділі ми окреслюємо комплексну картину ГПА, зосереджуючись на генетиці та хімії, що стоять за їх біосинтезом. Наведено приклади використання отриманих знань для вдосконалення штамів-продуцентів та відкриття нових сполук.*

frontiers in Chemistry

MINI REVIEW
published: 24 March 2022
doi: 10.3389/fchem.2022.858708

Genetics Behind the Glycosylation Patterns in the Biosynthesis of Dalbaheptides

Oleksandr Yushchuk[1,2], *Kseniia Zhukrovska*[2], *Francesca Berini*[1], *Victor Fedorenko*[2] *and Flavia Marinelli*[1]*

[1]Department of Biotechnology and Life Sciences, University of Insubria, Varese, Italy, [2]Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, Lviv, Ukraine

Glycopeptide antibiotics are valuable natural metabolites endowed with different pharmacological properties, among them are dalbaheptides used to treat different infections caused by multidrug-resistant Gram-positive pathogens. Dalbaheptides are produced by soil-dwelling high G-C Gram-positive actinobacteria. Their biosynthetic pathways are encoded within large biosynthetic gene clusters. A non-ribosomally synthesized heptapeptide aglycone is the common scaffold for all dalbaheptides. Different enzymatic tailoring steps, including glycosylation, are further involved in decorating it. Glycosylation of dalbaheptides is a crucial step, conferring them specific biological activities. It is achieved by a plethora of glycosyltransferases, encoded within the corresponding biosynthetic gene clusters, able to install different sugar residues. These sugars might originate from the primary metabolism, or, alternatively, their biosynthesis might be encoded within the biosynthetic gene clusters. Already installed monosaccharides might be further enzymatically modified or work as substrates for additional glycosylation. In the current minireview, we cover recent updates concerning the genetics and enzymology behind the glycosylation of dalbaheptides, building a detailed and consecutive picture of this process and of its biological evolution. A thorough understanding of how glycosyltransferases function in dalbaheptide biosynthesis might open new ways to use them in chemo-enzymatic synthesis and/or in combinatorial biosynthesis for building novel glycosylated antibiotics.

Keywords: glycopeptide antibiotics, dalbaheptides, ramoplanin, teicoplanin, A40926, glycosyltransferase, biosynthetic gene cluster

OPEN ACCESS

Edited by:
Zhongping Tan,
Chinese Academy of Medical Sciences and Peking Union Medical College, China

Reviewed by:
Max Julian Cryle,
Monash University, Australia

***Correspondence:**
Flavia Marinelli
flavia.marinelli@uninsubria.it

Specialty section:
This article was submitted to Chemical Biology, a section of the journal Frontiers in Chemistry

Received: *20 January 2022*
Accepted: *21 February 2022*
Published: *24 March 2022*

Citation:
Yushchuk O, Zhukrovska K, Berini F, Fedorenko V and Marinelli F (2022) Genetics Behind the Glycosylation Patterns in the Biosynthesis of Dalbaheptides. Front. Chem. 10:858708. doi: 10.3389/fchem.2022.858708

INTRODUCTION

Among different bacterial phyla, the mycelia-forming members of actinobacteria-broadly known as actinomycetes-remain the best antibiotic providers (Bérdy, 2005; Hutchings et al., 2019). The biosynthesis of antibiotics involves many enzymes, which are encoded by co-localized genes-biosynthetic gene clusters (BGCs) (Medema et al., 2015). BGCs undergo modular evolution, often exchanging operons and single genes coding for biosynthetic and modification enzymes (Medema et al., 2014). Genes for glycosyltransferases (GTs) are one such example, being found in different BGCs, with corresponding proteins having relaxed substrate specificity and consequently being able to modify different natural scaffolds (Salas and Méndez, 2007).

The astonishing variability of glycosylation patterns in one group of antibiotics led to their eponymous description as glycopeptide antibiotics (GPAs, Nicolaou et al., 1999). Natural GPAs amalgamate five types of related compounds, differing in chemical structures, where, paradoxically, only types I-IV-also known as dalbaheptides (Parenti and Cavalleri, 1989)-are glycosylated (Nicolaou et al., 1999). All dalbaheptides possess a non-ribosomal heptapeptide aglycone differing in amino acid (aa) composition, cross-linking, and decoration (Nicolaou et al., 1999); they inhibit the growth of Gram-positive bacteria by blocking cell wall maturation (Binda et al., 2014; Yushchuk et al., 2020a). Glycosylation and acylation (a step depending on glycosylation) of dalbaheptides contribute to their antimicrobial activities, favoring dimerization and membrane localization at the site of action (Gerhard et al., 1993; Mackay et al., 1994; Beauregard et al., 1995; Snyder et al., 1998). On the other hand, excessive glycosylation does not always bring pharmacological benefits: it seemed to induce platelet aggregation in patients treated with ristocetin (Howard and Firkin, 1971; Coller and Gralnick, 1977), which was consequently withdrawn from the clinical use (Gangarosa et al., 1958).

Dalbaheptides are clinically used as drugs of last resort against multidrug-resistant Gram-positive pathogens (Marcone et al., 2018). First-generation GPAs-vancomycin and teicoplanin (produced by different *Amycolatopsis* spp. and *Actinoplanes teichomyceticus* ATCC 31121, respectively)-have a long and reliable history of clinical application (Jovetic et al., 2010; Binda et al., 2014; Marcone et al., 2018). In turn, natural GPAs served as precursors for three second-generation semisynthetic and clinically used GPAs (Binda et al., 2014; Butler et al., 2014): dalbavancin derived from A40926 (produced by *Nonomuraea gerenzanensis* ATCC 39727) (Crotty et al., 2016), telavancin from chloroeremomycin (from *Kibdelosporangium aridum* A82846) (Klinker and Borgert, 2015), and oritavancin from vancomycin (Crotty et al., 2016).

Although the chemical variety of glycosyl groups decorating dalbaheptide aglycones is quite remarkable (Nicolaou et al., 1999), many aspects of the genetics behind their biosynthesis and incorporation remain obscure. In this minireview, we focus on those dalbaheptides whose BGC sequences are nowadays available, and for which some experimental evidence about their glycosylation steps is reported in the literature. The model BGCs are *cep*, *bal*, *tei*, *vcm*, and *dbv*, responsible for the production of chloroeremomycin (in *K. aridum* A82846) (van Wageningen et al., 1998), balhimycin (in *Amycolatopsis balhimycina* DSM 5908) (Shawky et al., 2007), teicoplanin (Li et al., 2004), vancomycin (in *Amycolatopsis orientalis* HCCB10007) (Xu et al., 2014), and A40926 (Sosio et al., 2003), respectively. BGC from *Amycolatopsis* sp. MJM2582 (Truman et al., 2014) represents the ristocetin biosynthetic pathway and was found also in other *Amycolatopsis* spp. (Spohn et al., 2014; Liu et al., 2021). Glycosylation-related genes from more recently described BGCs for UK-68,597 (*auk* from *Actinoplanes* sp. ATCC 53533) (Yim et al., 2014a), pekiskomycin (*pek* from *Streptomyces* sp. WAC1420) (Thaker et al., 2013), keratinimicin (ker from *Amycolatopsis keratiniphila* NRRL B-24117) (Xu et al., 2019), and A50926 (from *Nonomuraea coxensis* DSM 45129) (Yushchuk et al., 2021) are also reviewed. Overall, multiple recent findings on dalbaheptide glycosylation updated the overall picture and merit a proper review, outlining what is known and why it is still worthy of further investigations.

Delineating Steps in Dalbaheptide Glycosylation

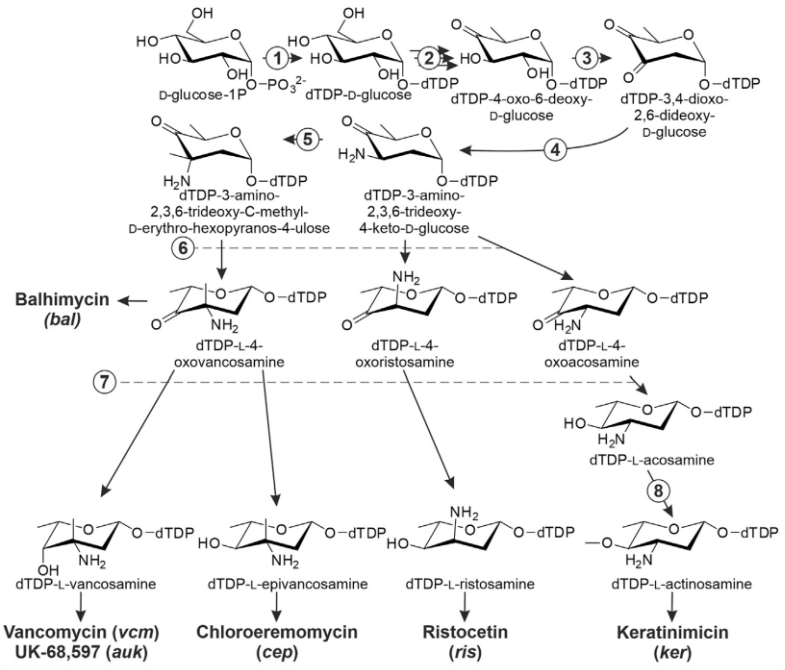
The biosynthesis of dalbaheptides is generally divided into three distinct stages (Yim et al., 2014b, 2016; Yushchuk et al., 2020b), that is, 1) generation of non-proteinogenic aa pool, further utilized in 2) non-ribosomal biosynthesis of the oligopeptide aglycones (coupled with the oxidative cross-linking); fully cross-linked aglycones are further 3) modified in a variety of tailoring steps. All dalbaheptide BGCs encode GTs, tailoring enzymes significantly contributing to the structural variety of these antibiotics (Nicolaou et al., 1999). Non-glycosylated dalbaheptide A47934 (from *Streptomyces toyocaensis* NRRL 15009) is the only exception here; consistently, the corresponding BGC lacks GT genes (Pootoolal et al., 2002).

More in detail, different steps might be defined in the glycosylation process of dalbaheptides, layer by layer "wrapping" the aglycone. The first step includes the biosynthesis of non-conventional sugar donors for aglycone decoration. Indeed, while some dalbaheptides are decorated with sugars deriving from primary metabolism (e.g., α-D-mannose and *N*-acetylglucosamine (GlcNAc) in teicoplanin or A40926), aglycones of vancomycin, balhimycin, and chloroeremomycin are decorated with the non-conventional monosaccharides L-vancosamine, L-4-oxovancosamine, and L-epivancosamine, respectively. In a similar manner to the biosynthesis of non-proteinogenic aa, enzymes required for the biosynthesis of such non-conventional monosaccharides are encoded within dalbaheptide BGCs. The second step (often the last one) consists of *O*-glycosylation of the aromatic aa forming the aglycone. In the third one, the installed sugars might be further modified in minor or major ways (e.g., α-D-mannose *O*-acetylation and GlcNAc deacetylation in A40926 biosynthesis).

Biosynthesis of Non-Conventional Monosaccharides was Required for the Glycosylation of Dalbaheptides

Conventional sugars in GPAs-from primary metabolism-are D-mannose, D-glucose, D-arabinose, GlcNAc, and L-rhamnose. Non-conventional sugar residues include L-vancosamine, L-epivancosamine, L-4-oxovancosamine, L-ristosamine, and L-actinosamine. In addition to the aforementioned examples, L-vancosamine is present in Substitute with UK-68,597, while L-ristosamine is characteristic for ristocetin, and L-actinosamine for keratinimicin (Xu et al., 2019).

Biosynthesis of L-epivancosamine was initially studied in chloroeremomycin producer (van Wageningen et al., 1998). *In vitro* experiments (Chen et al., 2000; Kirkpatrick et al., 2000) demonstrated how five enzymes encoded within *cep*-namely, EvaA-E-transformed



		chloroeremomycin vancomycin		balhimycin	ristocetin	keratinimicin	UK-68,597
		vcm	*cep*	*bal*	*ris*	*ker*	*auk*
1	glucose-1-phosphate thymidylylransferase	(*abs*)	(*abs*)	(*abs*)	(*abs*)	(*abs*)	Auk7
2	non-BGC-encoded enzymes	(*abs*)	(*abs*)	(*abs*)	(*abs*)	(*abs*)	(*abs*)
3	dTDP-4-dehydro-6-deoxy-α-D-glucopyranose 2,3-dehydratase	VcaA(*is*)	EvaA	DvaA(*is*)	Orf30(*is*)	KraA(*is*)	Auk13(*is*)
4	dTDP-3-amino-2,3,6-trideoxy-4-keto-D-glucose transaminase	VcaB(*is*)	EvaB	DvaB(*is*)	Orf32(*is*)	KraC(*is*)	Auk12(*is*)
5	SAM-dependent C3-methyltransferase	VcaC(*is*)	EvaC	DvaC(*is*)	(*abs*)	(*abs*)	Auk9(*is*)
6	C5-epimerase	VcaD(*is*)	EvaD	DvaD(*is*)	Orf33(*is*)	KraD(*is*)	Auk8(*is*)
7	C4-ketoreductase	VcaE(*is*)	EvaE	DvaE*	Orf31(*is*)	KraB(*is*)	(*abs*)
8	O-methyltransferase	(*abs*)	(*abs*)	(*abs*)	(*abs*)	KerM(?)	(*abs*)

FIGURE 1 | Enzymes involved in the biosynthesis of non-conventional aminosugars, decorating aglycones of some dalbaheptides. Biosynthetic pathway of L-epivancosamine serves as a model, since EvaA-B-C-D-E, coded within *cep*, are the only enzymes that were studied experimentally. Functions of all other enzymes were assigned by *in silico* comparison (*is*); (*abs*) indicates that the corresponding gene is absent from BGCs; (?) indicates that the assigned function is speculative, having no experimentally investigated prototype. Asterisk at DvaE indicates that this protein is mutated. Refer to the main text for more details.

dTDP-4-oxo-6-deoxy-D-glucose into L-epivancosamine (**Figure 1**). Biosynthetic routes for aminosugars decorating other dalbaheptides were deduced from this model pathway. Initially activated substrate dTDP-4-oxo-6-deoxy-D-glucose is commonly derived from D-glucose-1-phosphate by the action of non-BGC-encoded enzymes (**Figure 1**). One notable exception is the UK-68,597 biosynthesis, since *auk* contains a gene for glucose 1-phosphate thymidylylransferase-*auk7* (Yim et al., 2014a)-required for the D-glucose-1-phosphate activation. The presence of Auk7 likely positively contributes to the pool of L-vancosamine precursors in UK-68,597 biosynthesis.

vcm for vancomycin contains the orthologs of *evaA-E* genes, namely, *vcaA-E* (Xu et al., 2014) (**Figure 1**). The two sets of proteins are quite similar, sharing at least 75% of aa sequence identity (aa s.i.) in each pair (Xu et al., 2014). Nevertheless, minor differences in EvaE and VcaE seem to impact the function, making the first to convert dTDP-L-4-

oxovancosamine into dTDP-L-epivancosamine, while the second yields dTDP-L-vancosamine (**Figure 1**). Instead, *auk* contains the orthologs of only *evaA*, *evaB*, *evaC*, and *evaD* (*auk13*, *auk12*, *auk9*, and *auk8*, respectively) (Yim et al., 2014a), lacking the ortholog of *evaE*: probably a functional homolog (or analog) of *evaE* resides outside the *auk* borders, finally contributing to L-vancosamine production. More understandable is the case of *bal* (Shawky et al., 2007), where *evaA-E* orthologs are named *dvaA-E*. Here, *evaE* ortholog *dvaE* is truncated, coding only the 99 aa C-terminal part of C4-ketoreductase (Donadio et al., 2005). Such truncated protein is non-functional, and consequently, the aminosugar biosynthesis terminates at the stage of dTDP-L-4-oxovancosamine (**Figure 1**). Biosynthetic pathways for L-ristosamine and L-actinosamine, coded within *ris* and *ker*, are more diverged. They both missed the EvaC ortholog (Spohn et al., 2014; Truman et al., 2014; Xu et al., 2019; Liu et al., 2021), resulting in the lack of a methyl group at the C3 position (**Figure 1**). Thus, biosynthesis of L-ristosamine might be attributed to the orthologs of EvaA, EvaB, EvaD, and EvaE-Orf30, Orf32, Orf33, and Orf31, respectively (Truman et al., 2014). The same protein set is encoded within *ker*-KraA-D (Xu et al., 2019). The major difference between L-actinosamine and L-ristosamine is *O*-methylation of the C4-position. It is still unknown how this methylation is achieved; however, the annotation of *ker* reveals the presence of a gene coding for an *O*-methyltransferase-*kerM* (Xu et al., 2019). Since the aglycone of keratinimicin lacks any *O*-methylations, it seems plausible that KerM catalyzes the ultimate step of L-actinosamine biosynthesis (**Figure 1**).

In all cases described above the orthologs of *evaA-E-B-D* are most likely co-expressed, forming one operon, while the orthologs of *evaC* belong to a separate transcriptional unit (Shawky et al., 2007; Liu et al., 2021). *auk* is an exception, with *auk8-9* (*evaD-C*) and *auk12-13* (*evaA-B*) probably belonging to different operons (Yim et al., 2014a).

GTs Involved in Dalbaheptide Glycosylation

All GTs decorating aglycones of dalbaheptides belong to two families, according to the Carbohydrate-Active enZYmes Database (CAZy, http://www.cazy.org, Drula et al., 2022). GTs responsible for the installation of non-conventional aminosugars and conventional D-glucose, D-arabinose, Glc*N*Ac, and L-rhamnose, belong to the GT1 family. These GTs require sugar substrates to be either dTDP- or UDP-activated, and share a unique two-domain structure (the so-called GT-B fold, Lairson et al., 2008), having C- and N-terminal Rossmann-like domains connected by a flexible linker region (Zhang et al., 2020). Recognition sites for the donor NDP-activated sugars are located at C-terminal domains (Chang et al., 2011), while N-terminal domains contain the acceptor binding site for dalbaheptide aglycone (Chang et al., 2011; Zhang et al., 2020). GTs of the second-GT39-family are responsible for the installation of D-mannose and require D-mannosyl-1-phosphoundecaprenol as a donor substrate. Large hydrophobic GT39-GTs are predicted as membrane-associated, having a GT-C fold with multiple transmembrane helices and intracellular active sites (Lairson et al., 2008).

In the aforementioned dalbaheptides, GT1-GTs attach sugar residues preferentially at AA-4 (4-hydroxyphenylglycine, Hpg) and AA-6 (*β*-hydroxytyrosine, Bht) of the aglycone or add additional monosaccharides to already existing mono/di/trisaccharides at AA-4. Fully cross-linked aglycones serve as acceptor substrates for GT1-GTs under physiological conditions, albeit some GT1-GTs were able to recognize partially cross-linked aglycones under certain experimental conditions, for instance, in mutasynthesis approaches (Weist et al., 2004; Butz et al., 2008). GT39-GTs attach mannose at AA-7 (3,5-dihydroxyphenylglycine, Dpg). The presence of multiple GTs within one pathway might result in the production of mixtures of related congeners, differing in glycosylation patterns. Dalbaheptides glycosylated at other AA positions were also described (Nicolaou et al., 1999), indicating that glycosylation might also occur at 1) AA-2 (Bht) and AA-1 (Hpg) in type II aglycone, and 2) AA-1 (Hpg) and AA-3 (Bht) in type III aglycone. Unfortunately, we currently lack any genomic information on the producers of these molecules, which would merit further investigations.

Type I dalbaheptides chloroeremomycin, balhimycin, and vancomycin served as the first models for experimental investigation of GT functions. Corresponding BGCs encode slightly different sets of GT1-GTs: GtfA, GtfB, and GtfC in *cem*; BgtfA, BgtfB, and BgtfC in *bal*; and GtfD and GtfE in *vcm*. Orthologous GTs GtfA and BgtfA install L-epivancosamine and L-4-oxovancosamine at AA-6 of chloroeremomycin and balhimycin, respectively. *vcm* does not encode GtfA ortholog, explaining why vancomycin is not glycosylated at AA-6. Peculiarly, this particular difference between vancomycin and chloroeremomycin seems to augment the antimicrobial activity of the latter, implying that L-epivancosamine at AA-6 facilitates cell wall binding (Nagarajan, 1993; Allen et al., 2002). GtfB, BgtfB, and GtfE are orthologs, glucosylating AA-4 in the biosynthesis of all three antibiotics (**Figure 2A**, Pelzer et al., 1999; Losey et al., 2001; Mulichak et al., 2001). Then, GtfC and GtfD attach L-epivancosamine or L-vancosamine to D-glucose at AA-4 in the biosynthesis of chloroeremomycin and vancomycin, respectively (**Figure 2A**, Losey et al., 2001, 2002; Mulichak et al., 2004). GtfC, GtfD, and BgtfC are orthologous proteins, but balhimycin lacks a disaccharide at AA-4, which could be found only in balhimycin V (a congener produced in residual amounts) (Pelzer et al., 1999; Stegmann et al., 2010). This might be due to the low affinity of BgtfC for the donor substrate-L-4-oxovancosamine, produced as a consequence of DvaE mutation. More recently described type I dalbaheptide pekiskomycin is only glucosylated at AA-4, coherently with the only 1 GT encoded in *pek*: Pek28, which is a GtfB ortholog (**Figure 2A**) (Thaker et al., 2013).

So far, little is known about GTs decorating aglycones of type II dalbaheptides. *ker* (single type II BGC sequenced) carries three genes for GT1-GTs and one gene for GT39-GT (Xu et al., 2019). Among *ker*-encoded GT1-GTs, $gtfA_{ker}$, $gtfB_{ker}$, and $gtfC_{ker}$ (*ker* was added to distinguish them from *cep* genes) are orthologs of *gtfA*, *gtfB*, and *gtfC*, respectively (Xu et al., 2019). Thus, $GtfB_{ker}$

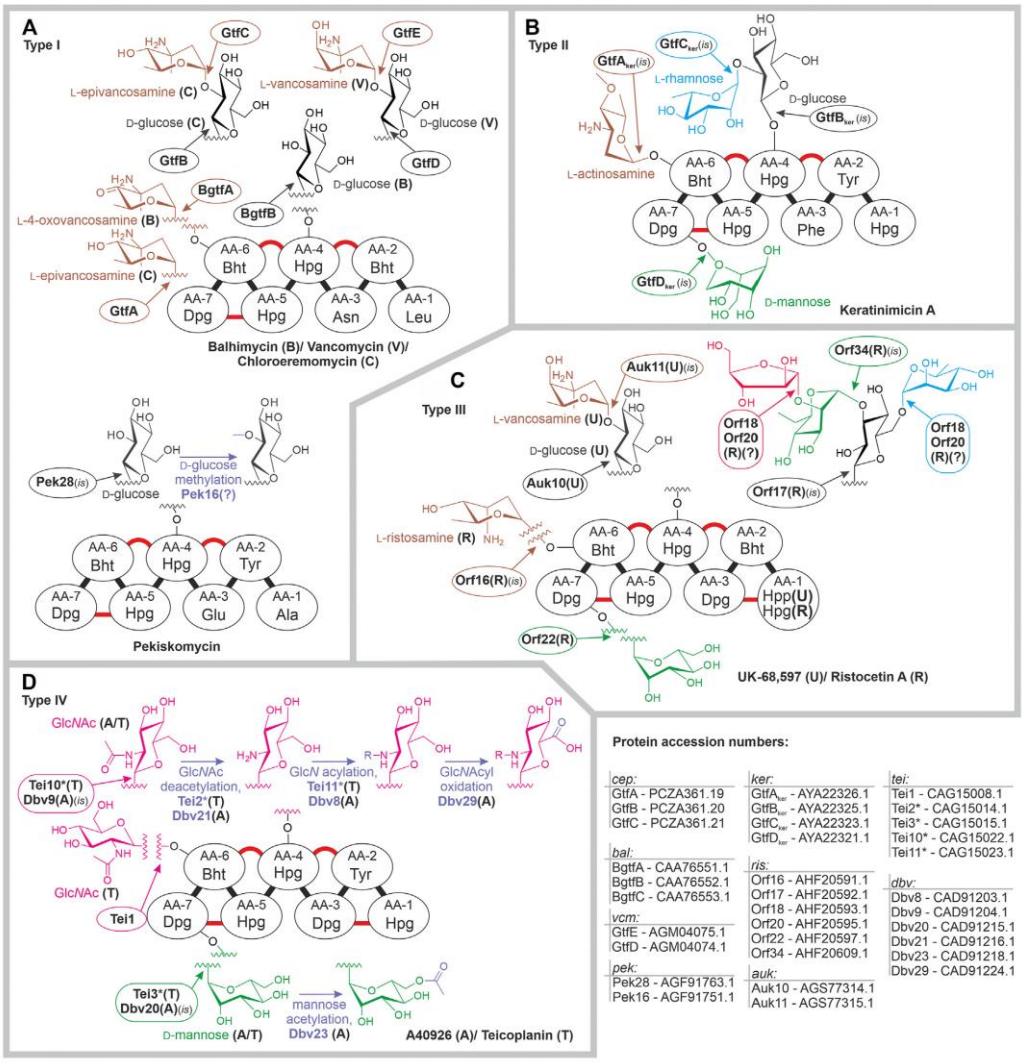


FIGURE 2 | Glycosylation patterns of type I **(A)**, type II **(B)**, type III **(C)**, and type IV **(D)** dalbaheptides and enzymes involved therein. Refer to the main text for more details. Dalbaheptide aglycones are depicted schematically with cross-links shown in red, and chlorination and sulfation sites are not shown. Aglycone amino acid abbreviations mean following: Leu-leucine; Asn-asparagine; Ala-alanine; Glu-glutamine; Phe-phenylalanine; Tyr-tyrosine; Bht-*β*-hydroxytyrosine; Dpg-3,5-dihydroxyphenylglycine; Hpg-4-hydroxyphenylglycine; Hpp-4-hydroxyphenypyruvate; Tyr-. GlcN stated for *N*-glucosamine, GlcNAcyl for N-acylglucosamine. All enzymes, whose functions were assigned by *in silico* comparison are marked with (*is*); (?) indicates that function was assigned *in silico* without experimentally investigated prototype or exact function remains unknown. For fast access to protein sequences mentioned in this figure, use following links: *cep*: GtfA-PCZA361.19, GtfB-PCZA361.20, GtfC-PCZA361.21; *bal*: BgtfA-CAA76551.1, BgtfB-CAA76552.1, BgtfC-CAA76553.1; *vcm*: GtfE-AGM04075.1, GtfD-AGM04074.1; *pek*: Pek28-AGF91763.1, Pek16-AGF91751.1; *ker*: $GtfA_{ker}$-AYA22326.1, $GtfB_{ker}$-AYA22325.1, $GtfC_{ker}$-AYA22323.1, $GtfD_{ker}$-AYA22321.1; *ris*: Orf16-AHF20591.1, Orf17-AHF20592.1, Orf18-AHF20593.1, Orf20-AHF20595.1; Orf22-AHF20597.1, Orf34-AHF20609.1; *auk*: Auk10-AGS77314.1, Auk11-AGS77315.1; *tei*: Tei1-CAG15008.1, Tei2*-CAG15014.1, Tei3*-CAG15015.1, Tei10*-CAG15022.1, Tei11*-CAG15023.1; *dbv*: Dbv8-CAD91203.1, Dbv9-CAD91204.1, Dbv20-CAD91215.1, Dbv21-CAD91216.1, Dbv23-CAD91218.1, Dbv29-CAD91224.1.

most likely installs D-glucose at AA-4, and L-rhamnose is then appended to D-glucose by $GtfC_{ker}$; this leaves $GtfA_{ker}$ responsible for the attachment of L-actinosamine at AA-6 (**Figure 2B**). Single *ker*-encoded GT39-GT $GtfD_{ker}$ most probably attaches D-mannose at AA-7 (**Figure 2B**).

More information is available on the GTs involved in the biosynthesis of type III dalbaheptides. Ristocetin BGC encodes 6 GTs, four of GT1 family and two belonging to the GT39 family (Truman et al., 2014). Phylogenetic reconstruction allowed to assign functions to 4 GTs, assuming that Orf16 attaches L-ristosamine at AA-6, Orf17-D-glucose at AA-4, then appended with D-mannose by Orf34; finally, a second mannosyltransferase-Orf22-was expected to act at AA-7 (**Figure 2C**, Truman et al., 2014), as later confirmed by its heterologous expression in *Streptomyces coelicolor* carrying *sta* (Yim et al., 2016). Functions of Orf18 and Orf20 were not assigned, since these proteins were distantly related to known GTs (Truman et al., 2014). Another type III dalbaheptide-UK-68,597-is decorated with 2-L-vancosaminyl-D-glucose disaccharide at AA-4 (Yim et al., 2014a), while *auk* carries three genes for GT1-GTs. *In vitro* assay showed that Auk10 (GtfB ortholog) is responsible for glucosylation; *in silico* analysis then suggested L-vancosamine to be installed by Auk11 (GtfC ortholog, **Figure 2C**) (Yim et al., 2014a). Although the third GT, Auk14, is a GtfA ortholog, UK-68,597 lacks sugar residues at AA-6. *In vitro* assay suggested that Auk14 might be inactive or possess a very low affinity to substrates available in UK-68,597 biosynthesis, as observed with BgtfA (Yim et al., 2014a).

Glycosylation events that take place in type IV dalbaheptide biosynthesis are well defined (**Figure 2D**). *tei* and *dbv* BGCs encode three and two GTs, respectively. Accordingly, both antibiotics carry a Glc*N*Ac moiety at AA-4 and D-mannose at AA-7, while teicoplanin aglycone is also decorated with another Glc*N*Ac at AA-6. GtfB ortholog-Tei10*-installs the Glc*N*Ac moiety at AA-4, as demonstrated from multiple *in vivo* and *in vitro* experiments (Li et al., 2004; Howard-Jones et al., 2007; Truman et al., 2009; Yushchuk et al., 2016, 2020b). Consistently, Dbv9-a Tei10* ortholog-is supposed to play the same role in A40926 glycosylation. Tei1 was shown to attach Glc*N*Ac to teicoplanin aglycone at AA-6 (Li et al., 2004), whereas *dbv* lacks any Tei1 ortholog, explaining why A40926 has no sugars at AA-6. Finally, Tei3* was shown to be responsible for the decoration of teicoplanin aglycone with D-mannose at AA-7 (Yushchuk et al., 2016), implying that Dbv20 (Tei3* ortholog) has the same function in A40926 biosynthesis (**Figure 2D**). Interestingly, ramoplanin BGC (*ramo*), recently shown to be genetically related to *tei* (Waglechner et al., 2019), encodes a homolog of Tei3*-Ramo29 (45% aa s.i., Chen et al., 2013). Ramoplanin is a clinically relevant peptide antibiotic produced by *Actinoplanes ramoplaninifer* ATCC 33076 (Marcone et al., 2017). Unlike dalbaheptides, it carries a 4-D-mannosyl-D-mannose disaccharide, instead of a single D-mannose residue. Ramo29 was shown to install the first D-mannose residue (Chen et al., 2013), but the GT responsible for the second mannosylation remains unknown. This merits further investigation, since such GT looks like a promising tool to further modify mannosylated dalbaheptides.

Concluding this section, it is interesting to report that genes for GTs tend to form one operon in type I-III BGCs (Shawky et al., 2007; Liu et al., 2021), being more scattered in type IV BGCs (Alduina et al., 2007; Yushchuk et al., 2019). It is also notable that genes for mannosyltransferases are present in different types of BGCs coming from distant actinobacterial lineages. This might indicate D-mannose residues at AA-7 to be an ancestral feature for all dalbaheptides.

Further Modification Occurring on Attached Sugar Residues

Some further modifications of attached sugars occur during the biosynthesis of dalbaheptides, although they are quite rare. The first example comes from pekiskomycin, having D-glucose methylated. *pek* encodes two methyltransferases. One of them-Pek30-was experimentally shown to methylate the N-terminus of A47934 aglycone (Yim et al., 2016), leaving the other-Pek16-possibly responsible for D-glucose methylation (**Figure 2A**). Another notable modification is the acylation of AA-4 Glc*N*Ac in type IV dalbaheptides, such as teicoplanin and A40926. To achieve this modification, Glc*N*Ac at AA-4 is first deacetylated with orthologous deacetylases Tei2*/Dbv21 (Ho et al., 2006; Truman et al., 2006), and *N*-glucosamine is then acylated with orthologous acyltransferases Tei11*/Dbv8 (**Figure 2D**) (Li et al., 2004; Kruger et al., 2005; Howard-Jones et al., 2007; Yushchuk et al., 2016). Peculiarly, orthologs of Tei2*/Dbv21 are present in many (if not in all) BGCs for non-acylated type I-III dalbaheptides. The one from *cep*-CepI-was studied *in vitro* and shown to be inactive due to a single aa substitution (Truman et al., 2008). The omnipresence of *tei2** orthologs in type I-III BGCs induces to speculate that the *N*-acylglucosamine moiety at AA-4 is an ancestral feature, lost or modified in many evolutionary lineages of dalbaheptides.

Modifications of A40926 sugars do not end with acylation. *N*-acylglucosamine moiety is further oxidized to *N*-acylaminoglucuronic acid group by Dbv29 (**Figure 2D**) (Li et al., 2007). The biological role of such modification is unclear, although it seems to reduce the A40926 antimicrobial activity (Malabarba et al., 1995). Peculiarly, *noc* BGC in *N. coxensis* DSM 45129 lacks an ortholog for *dbv29*, coding the biosynthesis of non-oxidized A40926 analog-dalbaheptide A50926 (Yushchuk et al., 2021). Finally, the D-mannose residue at AA-7 of A40926 is *O*-acetylated with Dbv23 (**Figure 2D**) (Sosio et al., 2010). This modification is unstable and fades away in the alkaline extraction of A40926 (Alt et al., 2019); once again, its biological role is unclear, although it might be important for the regulation of antibiotic export and self-resistance (Alduina et al., 2020).

CONCLUSION AND OUTLOOK

Further research of dalbaheptide glycosylation is important for several reasons. Understanding of GTs donor- (activated sugar) and acceptor- (aglycone) substrate specificities will allow further chemical derivatization of these scaffolds

using *in vitro* chemo-enzymatic synthesis (Nakayama et al., 2014) or *in vivo* combinatorial biosynthesis (Yim et al., 2016). While the first approach has been widely investigated in the past for generating novel hybrid GPAs by combining natural and synthetic aglycones and sugars (as reviewed in Marcone et al., 2018), *in vivo* combinatorial biosynthesis is promising, but still rather limited in its applications. Alternatively, the two-domain architecture of GT1-GTs might be exploited to create "chimaeras" with an expanded functional repertoire (Truman et al., 2009). Finally, a better comprehension of glycosylation mechanisms will contribute to tracing out a more complete picture of dalbaheptide evolution. Unlike other aspects of glycopeptide biosynthesis (Donadio et al., 2005; Waglechner et al., 2019; Andreo-Vidal et al., 2021), phylogenomics of GTs and sugar modification enzymes has not been studied yet. We believe that such reconstruction might open new scenarios on the evolution of antibiotic biosynthetic pathways.

AUTHOR CONTRIBUTIONS

KZ, FB, and OY collected data and articles and co-wrote the review; OY prepared the figures; and VF and FM supervised the work and co-wrote the review.

FUNDING

This work was supported by grant "Fondo di Ateneo per la Ricerca" to FM and FB and by the BG-09F grant of the Ministry of Education and Science of Ukraine to VF.

ACKNOWLEDGMENTS

We are grateful to Consorzio Interuniversitario per le Biotecnologie for supporting FB congresses attendances.

REFERENCES

Alduina, R., Piccolo, L. L., D'Alia, D., Ferraro, C., Gunnarsson, N., Donadio, S., et al. (2007). Phosphate-controlled Regulator for the Biosynthesis of the Dalbavancin Precursor A40926. *J. Bacteriol.* 189, 8120–8129. doi:10.1128/JB.01247-07

Alduina, R., Tocchetti, A., Costa, S., Ferraro, C., Cancemi, P., Sosio, M., et al. (2020). A Two-Component Regulatory System with Opposite Effects on Glycopeptide Antibiotic Biosynthesis and Resistance. *Sci. Rep.* 10, 1–12. doi:10.1038/s41598-020-63257-4

Allen, N. E., LeTourneau, D. L., Hobbs, J. N., and Thompson, R. C. (2002). Hexapeptide Derivatives of Glycopeptide Antibiotics: Tools for Mechanism of Action Studies. *Antimicrob. Agents Chemother.* 46, 2344–2348. doi:10.1128/AAC.46.8.2344-2348.2002

Alt, S., Bernasconi, A., Sosio, M., Brunati, C., Donadio, S., and Maffioli, S. I. (2019). Toward Single-Peak Dalbavancin Analogs through Biology and Chemistry. *ACS Chem. Biol.* 14, 356–360. doi:10.1021/acschembio.9b00050

Andreo-Vidal, A., Binda, E., Fedorenko, V., Marinelli, F., and Yushchuk, O. (2021). Genomic Insights into the Distribution and Phylogeny of Glycopeptide Resistance Determinants within the *Actinobacteria* Phylum. *Antibiotics* 10, 1533–1630. doi:10.3390/antibiotics10121533

Beauregard, D. A., Williams, D. H., Gwynn, M. N., and Knowles, D. J. (1995). Dimerization and Membrane Anchors in Extracellular Targeting of Vancomycin Group Antibiotics. *Antimicrob. Agents Chemother.* 39, 781–785. doi:10.1128/AAC.39.3.781

Bérdy, J. (2005). Bioactive Microbial Metabolites. *J. Antibiot.* 58, 1–26. doi:10.1038/ja.2005.1

Binda, E., Marinelli, F., and Marcone, G. (2014). Old and New Glycopeptide Antibiotics: Action and Resistance. *Antibiotics* 3, 572–594. doi:10.3390/antibiotics3040572

Butler, M. S., Hansford, K. A., Blaskovich, M. A. T., Halai, R., and Cooper, M. A. (2014). Glycopeptide Antibiotics: Back to the Future. *J. Antibiot.* 67, 631–644. doi:10.1038/ja.2014.111

Butz, D., Schmiederer, T., Hadatsch, B., Wohlleben, W., Weber, T., and Süssmuth, R. D. (2008). Module Extension of a Non-ribosomal Peptide Synthetase of the Glycopeptide Antibiotic Balhimycin Produced by *Amycolatopsis balhimycina*. *ChemBioChem* 9, 1195–1200. doi:10.1002/cbic.200800068

Chang, A., Singh, S., Phillips, G. N., and Thorson, J. S. (2011). Glycosyltransferase Structural Biology and its Role in the Design of Catalysts for Glycosylation. *Curr. Opin. Biotechnol.* 22, 800–808. doi:10.1016/j.copbio.2011.04.013

Chen, H., Thomas, M. G., Hubbard, B. K., Losey, H. C., Walsh, C. T., and Burkart, M. D. (2000). Deoxysugars in Glycopeptide Antibiotics: Enzymatic Synthesis of TDP-L-Epivancosamine in Chloroeremomycin Biosynthesis. *Proc. Natl. Acad. Sci.* 97, 11942–11947. doi:10.1073/pnas.210395097

Chen, J.-S., Wang, Y.-X., Shao, L., Pan, H.-X., Li, J.-A., Lin, H.-M., et al. (2013). Functional Identification of the Gene Encoding the Enzyme Involved in Mannosylation in Ramoplanin Biosynthesis in *Actinoplanes* Sp. *Biotechnol. Lett.* 35, 1501–1508. doi:10.1007/s10529-013-1233-3

Coller, B. S., and Gralnick, H. R. (1977). Studies on the Mechanism of Ristocetin-Induced Platelet Agglutination. *J. Clin. Invest.* 60, 302–312. doi:10.1172/JCI108778

Crotty, M. P., Krekel, T., Burnham, C.-A. D., and Ritchie, D. J. (2016). New Gram-Positive Agents: the Next Generation of Oxazolidinones and Lipoglycopeptides. *J. Clin. Microbiol.* 54, 2225–2232. doi:10.1128/JCM.03395-15

Donadio, S., Sosio, M., Stegmann, E., Weber, T., and Wohlleben, W. (2005). Comparative Analysis and Insights into the Evolution of Gene Clusters for Glycopeptide Antibiotic Biosynthesis. *Mol. Genet. Genomics* 274, 40–50. doi:10.1007/s00438-005-1156-3

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The Carbohydrate-Active Enzyme Database: Functions and Literature. *Nucleic Acids Res.* 50, D571–D577. doi:10.1093/nar/gkab1045

Gangarosa, E. J., Landerman, N. S., Rosch, P. J., and Herndon, E. G. (1958). Hematologic Complications Arising during Ristocetin Therapy. *N. Engl. J. Med.* 259, 156–161. doi:10.1056/nejm195807242590402

Gerhard, U., Mackay, J. P., Maplestone, R. A., and Williams, D. H. (1993). The Role of the Sugar and Chlorine Substituents in the Dimerization of Vancomycin Antibiotics. *J. Am. Chem. Soc.* 115, 232–237. doi:10.1021/ja00054a033

Ho, J.-Y., Huang, Y.-T., Wu, C.-J., Li, Y.-S., Tsai, M.-D., and Li, T.-L. (2006). Glycopeptide Biosynthesis: Dbv21/Orf2* from *dbv/tcp* Gene Clusters Are *N*-Ac-Glm Teicoplanin Pseudoaglycone Deacetylases and Orf15 from *cep* Gene Cluster Is a Glc-1-P Thymidyltransferase. *J. Am. Chem. Soc.* 128, 13694–13695. doi:10.1021/ja0644834

Howard, M. A., and Firkin, B. G. (1971). Ristocetin - A New Tool in the Investigation of Platelet Aggregation. *Thromb. Haemost.* 26, 362–369. doi:10.1055/s-0038-1653684

Howard-Jones, A. R., Kruger, R. G., Lu, W., Tao, J., Leimkuhler, C., Kahne, D., et al. (2007). Kinetic Analysis of Teicoplanin Glycosyltransferases and Acyltransferase Reveal Ordered Tailoring of Aglycone Scaffold to Reconstitute Mature Teicoplanin. *J. Am. Chem. Soc.* 129, 10082–10083. doi:10.1021/ja0735857

Hutchings, M. I., Truman, A. W., and Wilkinson, B. (2019). Antibiotics: Past, Present and Future. *Curr. Opin. Microbiol.* 51, 72–80. doi:10.1016/j.mib.2019.10.008

Jovetic, S., Zhu, Y., Marcone, G. L., Marinelli, F., and Tramper, J. (2010). β-Lactam and Glycopeptide Antibiotics: First and Last Line of Defense? *Trends Biotechnol.* 28, 596–604. doi:10.1016/j.tibtech.2010.09.004

Kirkpatrick, P. N., Scaife, W., Spencer, J. B., Williams, D. H., Hallis, T. M., and Liu, H.-w. (2000). Characterisation of a Sugar Epimerase Enzyme Involved in the

Biosynthesis of a Vancomycin-Group Antibiotic. *Chem. Commun.* 17, 1565–1566. doi:10.1039/b004463f

Klinker, K. P., and Borgert, S. J. (2015). Beyond Vancomycin: the Tail of the Lipoglycopeptides. *Clin. Ther.* 37, 2619–2636. doi:10.1016/j.clinthera.2015.11.007

Kruger, R. G., Lu, W., Oberthür, M., Tao, J., Kahne, D., and Walsh, C. T. (2005). Tailoring of Glycopeptide Scaffolds by the Acyltransferases from the Teicoplanin and A-40,926 Biosynthetic Operons. *Chem. Biol.* 12, 131–140. doi:10.1016/j.chembiol.2004.12.005

Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008). Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* 77, 521–555. doi:10.1146/annurev.biochem.76.061005.092322

Li, T., Huang, F., Haydock, S., Mironenko, T., Leadlay, P., and Spencer, J. (2004). Biosynthetic Gene Cluster of the Glycopeptide Antibiotic Teicoplanin Characterization of Two Glycosyltransferases and the Key Acyltransferase. *Chem. Biol.* 11, 107–119. doi:10.1016/S1074-5521(04)00002-X

Li, Y.-S., Ho, J.-Y., Huang, C.-C., Lyu, S.-Y., Lee, C.-Y., Huang, Y.-T., et al. (2007). A Unique Flavin Mononucleotide-Linked Primary Alcohol Oxidase for Glycopeptide A40926 Maturation. *J. Am. Chem. Soc.* 129, 13384–13385. doi:10.1021/ja075748x

Liu, K., Hu, X.-R., Zhao, L.-X., Wang, Y., Deng, Z., and Tao, M. (2021). Enhancing Ristomycin A Production by Overexpression of ParB-like StrR Family Regulators Controlling the Biosynthesis Genes. *Appl. Environ. Microbiol.* 87, 1–19. doi:10.1128/AEM.01066-21

Losey, H. C., Jiang, J., Biggins, J. B., Oberthür, M., Ye, X.-Y., Dong, S. D., et al. (2002). Incorporation of Glucose Analogs by GtfE and GtfD from the Vancomycin Biosynthetic Pathway to Generate Variant Glycopeptides. *Chem. Biol.* 9, 1305–1314. doi:10.1016/S1074-5521(02)00270-3

Losey, H. C., Peczuh, M. W., Chen, Z., Eggert, U. S., Dong, S. D., Pelczer, I., et al. (2001). Tandem Action of Glycosyltransferases in the Maturation of Vancomycin and Teicoplanin Aglycones: Novel Glycopeptides,. *Biochemistry* 40, 4745–4755. doi:10.1021/bi010050w

Mackay, J. P., Gerhard, U., Beauregard, D. A., Maplestone, R. A., and Williams, D. H. (1994). Dissection of the Contributions toward Dimerization of Glycopeptide Antibiotics. *J. Am. Chem. Soc.* 116, 4573–4580. doi:10.1021/ja00090a005

Malabarba, A., Ciabatti, R., Scotti, R., Goldstein, B. P., Ferrari, P., Kurz, M., et al. (1995). New Semisynthetic Glycopeptides MDL 63,246 and MDL 63,042, and Other Amide Derivatives of Antibiotic A-40,926 Active against Highly Glycopeptide-Resistant VanA Enterococci. *J. Antibiot.* 48, 869–883. doi:10.7164/antibiotics.48.869

Marcone, G. L., Binda, E., Berini, F., and Marinelli, F. (2018). Old and New Glycopeptide Antibiotics: from Product to Gene and Back in the post-genomic Era. *Biotechnol. Adv.* 36, 534–554. doi:10.1016/j.biotechadv.2018.02.009

Marcone, G. L., Binda, E., Reguzzoni, M., Gastaldo, L., Dalmastri, C., and Marinelli, F. (2017). Classification of *Actinoplanes* Sp. ATCC 33076, an Actinomycete that Produces the Glycolipodepsipeptide Antibiotic Ramoplanin, as *Actinoplanes ramoplaninifer* Sp. Nov. *Int. J. Syst. Evol. Microbiol.* 67, 4181–4188. doi:10.1099/ijsem.0.002281

Medema, M. H., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J. B., Blin, K., et al. (2015). Minimum Information about a Biosynthetic Gene Cluster. *Nat. Chem. Biol.* 11, 625–631. doi:10.1038/nchembio.1890

Medema, M. H., Cimermancic, P., Sali, A., Takano, E., and Fischbach, M. A. (2014). A Systematic Computational Analysis of Biosynthetic Gene Cluster Evolution: Lessons for Engineering Biosynthesis. *Plos Comput. Biol.* 10, e1004016–12. doi:10.1371/journal.pcbi.1004016

Mulichak, A. M., Losey, H. C., Walsh, C. T., and Garavito, R. M. (2001). Structure of the UDP-Glucosyltransferase GtfB that Modifies the Heptapeptide Aglycone in the Biosynthesis of Vancomycin Group Antibiotics. *Structure* 9, 547–557. doi:10.1016/S0969-2126(01)00616-5

Mulichak, A. M., Lu, W., Losey, H. C., Walsh, C. T., and Garavito, R. M. (2004). Crystal Structure of Vancosaminyltransferase GtfD from the Vancomycin Biosynthetic Pathway: Interactions with Acceptor and Nucleotide Ligands,. *Biochemistry* 43, 5170–5180. doi:10.1021/bi036130c

Nagarajan, R. (1993). Structure-activity Relationships of Vancomycin-type Glycopeptide Antibiotics. *J. Antibiot.* 46, 1181–1195. doi:10.7164/antibiotics.46.1181

Nakayama, A., Okano, A., Peng, Y., Collins, J. C., Collins, K. C., Walsh, C. T., et al. (2014). Enzymatic Glycosylation of Vancomycin Aglycon: Completion of a Total Synthesis of Vancomycin and N- and C-Terminus Substituent Effects of the Aglycon Substrate. *Org. Lett.* 16, 3572–3575. doi:10.1021/OL501568T/SUPPL_FILE/OL501568T_SI_001.PDF

Nicolaou, K. C., Boddy, C. N. C., Bräse, S., and Winssinger, N. (1999). Chemistry, Biology, and Medicine of the Glycopeptide Antibiotics. *Angew. Chem. Int. Ed.* 38, 2096–2152. doi:10.1002/(sici)1521-3773(19990802)38:15<2096::aid-anie2096>3.0.co;2-f

Parenti, F., and Cavalleri, B. (1989). Proposal to Name the Vancomycin-Ristocetin like Glycopeptides as Dalbaheptides. *J. Antibiot.* 42, 1882–1883. doi:10.7164/antibiotics.42.1882

Pelzer, S., Süßmuth, R., Heckmann, D., Recktenwald, J., Huber, P., Jung, G., et al. (1999). Identification and Analysis of the Balhimycin Biosynthetic Gene Cluster and its Use for Manipulating Glycopeptide Biosynthesis in *Amycolatopsis mediterranei* DSM5908. *Antimicrob. Agents Chemother.* 43, 1565–1573. doi:10.1128/aac.43.7.1565

Pootoolal, J., Thomas, M. G., Marshall, C. G., Neu, J. M., Hubbard, B. K., Walsh, C. T., et al. (2002). Assembling the Glycopeptide Antibiotic Scaffold: The Biosynthesis of from *Streptomyces toyocaensis* NRRL15009. *Proc. Natl. Acad. Sci.* 99, 8962–8967. doi:10.1073/pnas.102285099

Salas, J. A., and Méndez, C. (2007). Engineering the Glycosylation of Natural Products in Actinomycetes. *Trends Microbiol.* 15, 219–232. doi:10.1016/j.tim.2007.03.004

Shawky, R. M., Puk, O., Wietzorrek, A., Pelzer, S., Takano, E., Wohlleben, W., et al. (2007). The Border Sequence of the Balhimycin Biosynthesis Gene Cluster from *Amycolatopsis balhimycina* Contains *bbr*, Encoding a StrR-like Pathway-specific Regulator. *J. Mol. Microbiol. Biotechnol.* 13, 76–88. doi:10.1159/000103599

Snyder, N. J., Cooper, R. D. G., Briggs, B. S., Zmijewski, M., Mullen, D. L., Kaiser, R. E., et al. (1998). Enzymatic Deacylation of Teicoplanin Followed by Reductive Alkylation: Synthesis and Antibacterial Activity of New Glycopeptides. *J. Antibiot.* 51, 945–951. doi:10.7164/antibiotics.51.945

Sosio, M., Canavesi, A., Stinchi, S., and Donadio, S. (2010). Improved Production of A40926 by *Nonomuraea* Sp. Through Deletion of a Pathway-specific Acetyltransferase. *Appl. Microbiol. Biotechnol.* 87, 1633–1638. doi:10.1007/s00253-010-2579-2

Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A., and Donadio, S. (2003). The Gene Cluster for the Biosynthesis of the Glycopeptide Antibiotic A40926 by *Nonomuraea* Species. *Chem. Biol.* 10, 541–549. doi:10.1016/S1074-5521(03)00120-0

Spohn, M., Kirchner, N., Kulik, A., Jochim, A., Wolf, F., Muenzer, P., et al. (2014). Overproduction of Ristomycin a by Activation of a Silent Gene Cluster in *Amycolatopsis japonicum* MG417-CF17. *Antimicrob. Agents Chemother.* 58, 6185–6196. doi:10.1128/AAC.03512-14

Stegmann, E., Frasch, H.-J., and Wohlleben, W. (2010). Glycopeptide Biosynthesis in the Context of Basic Cellular Functions. *Curr. Opin. Microbiol.* 13, 595–602. doi:10.1016/j.mib.2010.08.011

Thaker, M. N., Wang, W., Spanogiannopoulos, P., Waglechner, N., King, A. M., Medina, R., et al. (2013). Identifying Producers of Antibacterial Compounds by Screening for Antibiotic Resistance. *Nat. Biotechnol.* 31, 922–927. doi:10.1038/nbt.2685

Truman, A. W., Dias, M. V. B., Wu, S., Blundell, T. L., Huang, F., and Spencer, J. B. (2009). Chimeric Glycosyltransferases for the Generation of Hybrid Glycopeptides. *Chem. Biol.* 16, 676–685. doi:10.1016/j.chembiol.2009.04.013

Truman, A. W., Pan, Q., Röttgen, M., Stegmann, E., Leadlay, P. F., and Spencer, J. B. (2008). The Role of Cep15 in the Biosynthesis of Chloroeremomycin: Reactivation of an Ancestral Catalytic Function. *Chem. Biol.* 15, 476–484. doi:10.1016/j.chembiol.2008.03.019

Truman, A. W., Kwun, M. J., Cheng, J., Yang, S. H., Suh, J.-W., and Hong, H.-J. (2014). Antibiotic Resistance Mechanisms Inform Discovery: Identification and Characterization of a Novel *Amycolatopsis* Strain Producing Ristocetin. *Antimicrob. Agents Chemother.* 58, 5687–5695. doi:10.1128/AAC.03349-14

Truman, A. W., Robinson, L., and Spencer, J. B. (2006). Identification of a Deacetylase Involved in the Maturation of Teicoplanin. *ChemBioChem* 7, 1670–1675. doi:10.1002/cbic.200600308

Van Wageningen, A. A., Kirkpatrick, P. N., Williams, D. H., Harris, B. R., Kershaw, J. K., Lennard, N. J., et al. (1998). Sequencing and Analysis of Genes Involved in

the Biosynthesis of a Vancomycin Group Antibiotic. *Chem. Biol.* 5, 155–162. doi:10.1016/S1074-5521(98)90060-6

Waglechner, N., McArthur, A. G., and Wright, G. D. (2019). Phylogenetic Reconciliation Reveals the Natural History of Glycopeptide Antibiotic Biosynthesis and Resistance. *Nat. Microbiol.* 4, 1862–1871. doi:10.1038/s41564-019-0531-5

Weist, S., Kittel, C., Bischoff, D., Bister, B., Pfeifer, V., Nicholson, G. J., et al. (2004). Mutasynthesis of Glycopeptide Antibiotics: Variations of Vancomycin's AB-Ring Amino Acid 3,5-Dihydroxyphenylglycine. *J. Am. Chem. Soc.* 126, 5942–5943. doi:10.1021/ja0499389

Xu, F., Wu, Y., Zhang, C., Davis, K. M., Moon, K., Bushin, L. B., et al. (2019). A Genetics-free Method for High-Throughput Discovery of Cryptic Microbial Metabolites. *Nat. Chem. Biol.* 15, 161–168. doi:10.1038/s41589-018-0193-2

Xu, L., Huang, H., Wei, W., Zhong, Y., Tang, B., Yuan, H., et al. (2014). Complete Genome Sequence and Comparative Genomic Analyses of the Vancomycin-Producing *Amycolatopsis orientalis*. *BMC Genomics* 15, 1. doi:10.1186/1471-2164-15-363

Yim, G., Kalan, L., Koteva, K., Thaker, M. N., Waglechner, N., Tang, L., et al. (2014a). Harnessing the Synthetic Capabilities of Glycopeptide Antibiotic Tailoring Enzymes: Characterization of the UK-68,597 Biosynthetic Cluster. *ChemBioChem* 15, 2613–2623. doi:10.1002/cbic.201402179

Yim, G., Thaker, M. N., Koteva, K., and Wright, G. (2014b). Glycopeptide Antibiotic Biosynthesis. *J. Antibiot.* 67, 31–41. doi:10.1038/ja.2013.117

Yim, G., Wang, W., Thaker, M. N., Tan, S., and Wright, G. D. (2016). How to Make a Glycopeptide: a Synthetic Biology Approach to Expand Antibiotic Chemical Diversity. *ACS Infect. Dis.* 2, 642–650. doi:10.1021/acsinfecdis.6b00105

Yushchuk, O., Binda, E., and Marinelli, F. (2020a). Glycopeptide Antibiotic Resistance Genes: Distribution and Function in the Producer Actinomycetes. *Front. Microbiol.* 11, 1–9. doi:10.3389/fmicb.2020.01173

Yushchuk, O., Horbal, L., Ostash, B., Marinelli, F., Wohlleben, W., Stegmann, E., et al. (2019). Regulation of Teicoplanin Biosynthesis: Refining the Roles of Tei Cluster-Situated Regulatory Genes. *Appl. Microbiol. Biotechnol.* 103, 4089–4102. doi:10.1007/s00253-019-09789-w

Yushchuk, O., Ostash, B., Pham, T. H., Luzhetskyy, A., Fedorenko, V., Truman, A. W., et al. (2016). Characterization of the post-assembly Line Tailoring Processes in Teicoplanin Biosynthesis. *ACS Chem. Biol.* 11, 2254–2264. doi:10.1021/acschembio.6b00018

Yushchuk, O., Ostash, B., Truman, A. W., Marinelli, F., and Fedorenko, V. (2020b). Teicoplanin Biosynthesis: Unraveling the Interplay of Structural, Regulatory, and Resistance Genes. *Appl. Microbiol. Biotechnol.* 104, 3279–3291. doi:10.1007/s00253-020-10436-y

Yushchuk, O., Vior, N. M., Andreo-Vidal, A., Berini, F., Rückert, C., Busche, T., et al. (2021). Genomic-led Discovery of a Novel Glycopeptide Antibiotic by *Nonomuraea coxensis* DSM 45129. *ACS Chem. Biol.* 16, 915–928. doi:10.1021/acschembio.1c00170

Zhang, P., Zhang, Z., Zhang, L., Wang, J., and Wu, C. (2020). Glycosyltransferase GT1 Family: Phylogenetic Distribution, Substrates Coverage, and Representative Structural Features. *Comput. Struct. Biotechnol. J.* 18, 1383–1390. doi:10.1016/j.csbj.2020.06.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Yushchuk, Zhukrovska, Berini, Fedorenko and Marinelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Оригінал доступний за посиланням:
https://www.frontiersin.org/articles/10.3389/fchem.2022.858708/full.

Анотація: *Глікопептидні антибіотики є цінними природними метаболітами, що мають різні фармакологічні властивості: серед них далбагептиди, які використовуються для лікування різних інфекцій, викликаних мультирезистентними Грам-позитивними збудниками. Далбагептиди синтезуються Грам-позитивними актинобактеріями, ґрунтовими мікроорганізмами з високим вмістом ГЦ-пар в геномах. Біосинтетичні шляхи далбагептидів закодовані у великих кластерах біосинтетичних генів. Нерибосомний гептапептидний аглікон є спільним для всіх далбагептидів. Аглікон проходить через різні етапи модифікації, включаючи глікозилювання. Глікозилювання далбагептидів є останнім етапом, який надає їм специфічну біологічну активність. Глікозилювання досягається великою кількістю глікозилтрансфераз, закодованих у відповідних кластерах біосинтетичних генів. Глікозилтрансферази здатні приєднувати залишки різних цукрів. Ці цукри можуть походити з первинного метаболізму, або, альтернативно, їх біосинтез може бути закодований у кластерах біосинтетичних генів. Вже приєднані моносахариди можуть бути додатково модифіковані ферментативно або служити субстратами для додаткових етапів глікозилювання. У цьому огляді ми висвітлюємо останні доповнення, що стосуються генетики та ензимології глікозилювання далбагептидів, створюючи детальну та послідовну картину цього процесу і його біологічної еволюції. Глибоке розуміння того, як глікозилтрансферази функціонують у біосинтезі далбагептидів, може відкрити нові шляхи їх використання в хемоферментному синтезі й/або в комбінаторному біосинтезі для створення нових глікозильованих антибіотиків.*

frontiers in Microbiology

MINI REVIEW
published: 17 June 2020
doi: 10.3389/fmicb.2020.01173

Glycopeptide Antibiotic Resistance Genes: Distribution and Function in the Producer Actinomycetes

*Oleksandr Yushchuk†, Elisa Binda† and Flavia Marinelli**

Department of Biotechnology and Life Sciences, University of Insubria, Varese, Italy

Glycopeptide antibiotics (GPAs) are considered drugs of "last resort" for the treatment of life-threatening infections caused by relevant Gram-positive pathogens (enterococci, staphylococci, and clostridia). Driven by the issue of the never-stopping evolution of bacterial antibiotic resistance, research on GPA biosynthesis and resistance is developing fast in modern "post-genomic" era. It is today widely accepted that resistance mechanisms emerging in pathogens have been acquired from the soil-dwelling antibiotic-producing actinomycetes, which use them to avoid suicide during production, rather than being orchestrated *de novo* by pathogen bacteria upon continued treatment. Actually, more and more genomes of GPA producers are being unraveled, carrying a broad collection of differently arranged GPA resistance (named *van*) genes. In the producer actinomycetes, *van* genes are generally associated with the antibiotic biosynthetic gene clusters (BGCs) deputed to GPA biosynthesis, being probably transferred/arranged together, favoring a possible co-regulation between antibiotic production and self-resistance. GPA BGC-associated *van* genes have been also found mining public databases of bacterial genomic and metagenomic sequences. Interestingly, some BGCs for antibiotics, seemingly unrelated to GPAs (e.g., feglymycin), carry *van* gene homologues. Herein, we would like to cover the recent advances on the distribution of GPA resistance genes in genomic and metagenomics datasets related to GPA potential/proved producer microorganisms. A thorough understanding of GPA resistance in the producing microorganisms may prove useful in the future surveillance of emerging mechanisms of resistance to this clinically relevant antibiotic class.

Keywords: antimicrobial resistance, glycopeptide antibiotics, *van* genes, glycopeptide producers, biosynthetic gene clusters

OPEN ACCESS

Edited by:
Jose L. Martinez, Consejo Superior de Investigaciones Cientificas (CSIC), Spain

Reviewed by:
Kapil Tahlan, Memorial University of Newfoundland, Canada
Sonia Gullón, Centro Nacional de Biotecnología (CNB), Spain

Correspondence:
Flavia Marinelli
flavia.marinelli@uninsubria.it

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: *14 February 2020*
Accepted: *07 May 2020*
Published: *17 June 2020*

Citation:
Yushchuk O, Binda E and Marinelli F (2020) Glycopeptide Antibiotic Resistance Genes: Distribution and Function in the Producer Actinomycetes. Front. Microbiol. 11:1173. doi: 10.3389/fmicb.2020.01173

GPA MODE OF ACTION AND RESISTANCE GENES IN GRAM-POSITIVE PATHOGENS

According to a recent report (WHO, 2017), drug-resistant infections will kill more people than cancer in just over three decades: by 2050, 10 million people are going to die every year due to antimicrobial resistance (AMR). Consequently, it is mandatory to stimulate discovery and development of novel antibiotics to counteract AMR (O'Neill, 2016). Glycopeptide antibiotics (GPAs) are frequently used to treat life-threatening infections caused by multidrug-resistant Gram-positive pathogens, such as *Staphylococcus aureus*, *Enterococcus* spp., and *Clostridium difficile* (for a review on their discovery and development, see Marcone et al., 2018; on their

antimicrobial activity and clinical use, Zeng et al., 2016). GPAs inhibit bacterial cell wall synthesis in Gram-positive bacteria by binding to D-alanyl-D-alanine (D-Ala-D-Ala) dipeptide terminus of peptidoglycan (PG) precursors, sequestering the substrate from transpeptidation and transglycosylation reactions in the late extracellular stages of PG cross-linking. Thus, GPA action ultimately results in destabilizing cell wall integrity, causing bacterial cell death (Perkins and Nieto, 1974). Gram-negative microorganisms are intrinsically resistant to GPAs, because of their outer membrane, which prevents these molecules entering into the periplasm. In Gram-positive bacteria, the onset of vancomycin resistance was long-delayed in comparison to other antibiotic classes. The first vancomycin-resistant clinical isolate – an *Enterococcus faecium* strain – was reported in 1987, more than 30 years after the clinical introduction of vancomycin (Leclercq et al., 1988, Miller et al., 2016). Unfortunately, today a vast majority of *E. faecium* isolates harbor vancomycin resistance genes (*van*) (Vehreschild et al., 2019). The first vancomycin-resistant *S. aureus* (VRSA) isolate was reported in 2002 as a result of horizontal gene transfer from resistant enterococci (Bartley, 2002; Weigel et al., 2003); nowadays, 52 VRSA strains have been described worldwide (Cong et al., 2020).

The GPA resistance mechanisms in Gram-positive pathogens were intensively studied starting from the pioneering work published in the 1990s (Arthur et al., 1992, 1996). Gram-positive pathogens escape GPA action by reprogramming PG precursor biosynthesis, replacing the terminal D-Ala with D-lactate (D-Ala-D-Lac) or D-serine (D-Ala-D-Ser), thus reducing the affinity for cellular targets (Arthur et al., 1992, 1996; Courvalin, 2006). In enterococci, many different GPA-resistant phenotypes have been described according to their *van* gene operon organization (for a review, see Binda et al., 2014): in *vanA*, *vanB*, *vanD*, and *vanM* the key ligase determines the replacement of the terminus D-Ala with D-Lac, whereas in *vanC*, *vanE*, *vanG*, *vanL*, and *vanN* D-Ala. The D-Ala-D-Lac-type operons are located either on plasmids or on chromosomes, whereas the D-Ala-D-Ser-type ones are exclusively on the bacterial chromosome, except the case of *vanN* found on a plasmid in *E. faecium*. Operon expression could be inducible by GPAs (*vanA*, *vanB*, *vanG*, *vanE*, *vanL*, and *vanM*) or constitutive (*vanC*, *vanD*, and *vanN*) (Reynolds and Courvalin, 2005; Depardieu et al., 2007; Binda et al., 2014). The most clinically relevant manifestation of GPA resistance occurs in VanA enterococci and staphylococci, and in VanB enterococci. The first group is highly resistant to both vancomycin and teicoplanin, whereas the second group only to vancomycin. In both of them, resistance is mediated by the GPA-induced expression of the transposone-located *vanHAX* gene operon under the transcriptional control of the VanR/VanS two-component system (TCS). VanS is a membrane-associated sensor that in VanA bacteria is activated by the presence of either vancomycin or teicoplanin, whereas in VanB it is activated only by vancomycin. Consequently, VanB enterococci are sensitive to teicoplanin (Arthur et al., 1997, 1999; Arthur and Quintiliani, 2001). Activated VanS transfers a phosphoryl group to VanR, which is the response regulator that controls the co-transcription of the *vanH*, *vanA*, *vanX*, and *vanY* genes (Wright et al., 1993; Arthur et al., 1997, 1999; Arthur and Quintilliani, 2001). VanH is a dehydrogenase that reduces pyruvate to D-lactate; VanA is the key ligase that catalyzes the formation of the D-Ala-D-Lac resistant depsipeptide (Bugg et al., 1991; Arthur et al., 1992); VanX is a D,D-dipeptidase, which removes the intracellular pool of D-Ala-D-Ala produced by the native enterococcal ligase, ensuring that D-Ala-D-Lac is incorporated into PG precursors (Reynolds et al., 1994; Wu et al., 1995); and finally VanY has an ancillary role as a D,D-carboxypeptidase cleaving the last D-Ala from the residual pentapeptide PG precursors terminating in D-Ala-D-Ala (Arthur et al., 1998). Among the D-Ala-D-Ser-type operons, the better investigated was the *vanC*. It encodes for a racemase (VanT) that converts L-Ser to D-Ser, a ligase (VanC) that synthesizes D-Ala-D-Ser, and a bi-functional D,D-dipeptidase/D,D-carboxypeptidase (VanXYc) that cleaves the residual pools of D-Ala-D-Ala (Billot-Klein et al., 1994; Reynolds and Courvalin, 2005). In VanC phenotype, the TCS VanRcSc is located downstream the operon, but the resistance is constitutive due to mutations in the sensor VanSc (Healy et al., 2000; Hong et al., 2008; Koteva et al., 2010). VanC enterococci are intrinsically resistant to low levels of vancomycin, although they remain sensitive to teicoplanin.

Additional variants of these *van* gene operons were found in other Gram-positive pathogens including *Listeria* spp., streptococci, clostridia (Biavasco et al., 1996; Poyart et al., 1997; Peltier et al., 2013), and also in nonpathogenic Gram-positives, including *Bacillus circulans*, *Oerskovia* spp., *Corynebacterium* spp., and *Streptomyces coelicolor* (Power et al., 1995; Fontana et al., 1997; Hong et al., 2004). A novel vancomycin *vanF* operon ($vanY_FZ_FH_FFX_F$) was described in *Paenibacillus popilliae*, an environmental bacteria used as biopesticide to counteract beetle larvae that caused milky disease in Japan (Patel et al., 2000; Ahmed and Baptiste, 2018). The dissemination of GPA resistance more recently reached zoonotic pathogens such as the emergent *Streptococcus suis*, where the low level of vancomycin-resistance is due to the presence of a *vanG*-like operon (Huang et al., 2018). Herein, we focus our attention on *van* genes distribution and function in the GPA-producing actinomycetes, which are considered the putative primary source of the variety of GPA-resistant determinants occurring in environmental bacteria and pathogens (Marshall et al., 1998; Beltrametti et al., 2007; Marcone et al., 2010, 2014; Schäberle et al., 2011).

UPDATING THE GLYCOPEPTIDE RESISTANCE PARADIGM FOR THE GPA-PRODUCING STRAINS: *VAN* GENES AND THEIR ORGANIZATION IN KNOWN AND PUTATIVE GPA BGCs

Actinomycetes are Gram-positive soil-dwelling bacteria, which produce about two-thirds of the naturally derived antibiotics with clinical use (Bérdy, 2012; Barka et al., 2015), including GPAs (Nicolaou et al., 1999). Clinically relevant GPAs are produced by *Amycolatopsis orientalis* (vancomycin), *Actinoplanes*

teichomyceticus (teicoplanin), and *Nonomuraea gerenzanensis* (dalbavancin precursor – A40926) (Zeng et al., 2016; Marcone et al., 2018). GPA producers require self-resistance mechanisms to avoid suicide during antibiotic production and, like in pathogens, such resistance is due to *van* genes, whose description dates back to the end of the 1990s, one decade later than in pathogens (Marshall et al., 1997, 1998). Sequence and operon structure similarities of *van* genes between pathogens and GPA-producers are significant (Hong et al., 2008; Binda et al., 2014). The intriguing aspect is that in GPA producers, *van* genes are usually located within the GPA biosynthetic gene clusters (BGCs) deputed to the antibiotic biosynthesis (Pootoolal et al., 2002; Beltrametti et al., 2007; Marcone et al., 2010, 2014; Schäberle et al., 2011). In the last two decades, multiple novel GPA BGCs from actinomycetes were sequenced and annotated, and each of them (with few exceptions, see below) contains *van* genes (**Figure 1**).

Thus far, the majority of GPA BGCs were found in members of the genus *Amycolatopsis* (Adamek et al., 2018), which belongs to the *Pseudonocardiaceae* family. Besides the vancomycin producers, other *Amycolatopsis* spp. produce avoparcin, decaplanin, nogabecin, ristocetin, teicoplanin aglycone-like GPA, norvancomycin, balhimycin, and chloroeremomycin. In their corresponding BGCs, *van* genes were found just upstream the genes coding for the StrR-like pathway-specific regulators (orthologues of *bbr* from balhimycin BGC, **Figure 1**). We excluded from this comparison the chloroeremomycin BGC from *Amycolatopsis orientalis* PA-42867 (?) (van Wageningen et al., 1998), which apparently was not completely covered with sequencing. Thus, three patterns for the organization of *van* genes are recognizable (**Figure 1**) in *Pseudonocardiaceae* GPA producers. In the producers of avoparcin, decaplanin, nogabecin, ristocetin, and teicoplanin-like aglycone GPA, the GPA BGCs carry *vanHAX* orthologues, but not *vanY* or *vanRS* orthologues. In vancomycin and norvancomycin producers, *vanY* orthologues are clustered with *vanHAX* ones. Balhimycin producer – *Amycolatopsis balhimycina* – possesses a BGC with *vanRS* ($vlnR_{Ab}S_{Ab}$) and *vanY* ($vanY_{Ab}$) orthologues, but *vanHAX* orthologues ($vanH_{Ab}A_{Ab}X_{Ab}$) were actually found 2 kbp away from balhimycin BGC (Schäberle et al., 2011; Frasch et al., 2015). Indeed, no cluster-situated *van* genes were found sequencing the genome of *Kibdelosporangium aridum* – the producer of kibdelins (Shearer et al., 1986) – which also belongs to *Pseudonocardiaceae* family (**Figure 1**).

Other known GPA BGCs are from *Actinoplanes* spp. (family *Micromonosporaceae*) and *Nonomuraea* spp. (family *Streptosporangiaceae*). *Act. teichomyceticus* and *Actinoplanes* sp. ATCC 53533 produce teicoplanin (Bardone et al., 1978) and UK-68,597 (Skelton et al., 1990), respectively. Teicoplanin BGC (named *tei*) contains *vanHAX* and *vanRS* orthologues (*tei7-6-5* and *tei2-3*, respectively) organized in two separate operons, but none *vanY* orthologue (**Figure 1**; Li et al., 2004; Yushchuk et al., 2020b). In contrast, UK-68,597 BGC contains a *vanH*, not contiguous *vanR* and *vanS*, and a *vanY* orthologue (**Figure 1**; Yim et al., 2014). In the genus *Nonomuraea*, *N. gerenzanenis* ATCC 39727 and *Nonomuraea* sp. ATCC 55076 produce A40926 (Goldstein et al., 1987) and the type V glycopeptide kistamicin (Naruse et al., 1993), respectively. In A40926 BGC (named *dbv*) (Sosio et al., 2003), *vanR* and *vanS* homologues (*dbv6* and *dbv22*, respectively) are not contiguous and GPA resistance is due to the expression of *vanY* orthologue (*dbv7*, **Figure 1**; Marcone et al., 2010, 2014; Binda et al., 2012;). No *vanHAXY* genes are present in kistamicin BGC, although kistamicin BGC contains homologues of *vanS* and *vanR* named *kisG* and *kisB* (Nazari et al., 2017).

GPAs are also produced by few *Streptomyces* species (**Figure 1**). Interestingly, functional *van* genes were also found in *S. coelicolor*, which is not a GPA producer (Hong et al., 2004). A47934 BGC from *Streptomyces toyocaensis* NRRL 15009 contains $vanH_{St}A_{St}X_{St}$ and $vanR_{St}S_{St}$ operons, together with *staO* and *staP* orthologues to *S. coelicolor vanJ* and *vanK*, respectively (Pootoolal et al., 2002). Pekiskomycin BGC from *Streptomyces* sp. WAC1420 contains *vanY*, *vanJ* as well as *vanHAX* and *vanRS* homologues, but pekiskomycin BGC from *Streptomyces* sp. WAC4229 lacks *vanRS* homologues (Thaker et al., 2013). No homologues of *van* genes were found in complestatin (type V GPA) BGC from *S. lavendulae* SANK 60477 (?) (Chiu et al., 2001), although this antibiotic possesses a moderate antibacterial activity. However, complestatin was shown to inhibit the fatty acid biosynthesis in Gram-positive bacteria (Kwon et al., 2015), therefore the producer may require no cell wall remodeling for complestatin self-resistance. Finally, feglymycin BGC from *Streptomyces* sp. DSM11171 (**Figure 1**) encodes for a 13-mer peptide antibiotic acting on bacterial cell wall biosynthesis by inhibiting MurA and MurC. Albeit the structure and the mode of action of feglymycin differs from the ones of GPAs, feglymycin BGC shares a high level of similarity with GPA BGCs (Gonsior et al., 2015; Yushchuk et al., 2020a), including the presence of *vanRS*-like genes – *fegM* and *fegN*.

To conclude, CA915, CA37, and CA878 GPA BGCs (Banik et al., 2010), which were sequenced from metagenomics samples, contain *vanHAX*, *vanY* and *vanRS* homologues, whereas none *van* gene was found in other metagenome-derived GPA BGCs as TEG and VEG (Banik and Brady, 2008; **Figure 1**).

UPDATING ON WHAT IS KNOWN ABOUT THE *IN VIVO* FUNCTION OF *VAN* GENES IN GPA-PRODUCING STRAINS

Although *van* genes were found in multiple GPA BGCs, only for few of them the function was experimentally proven. Balhimycin resistance in *Am. balhimycina* is likely the most deeply investigated model among GPA producers (**Figure 2A**). $vanH_{Ab}A_{Ab}X_{Ab}$, that is located outside the BGC (**Figure 1**), was shown to be constitutively expressed through all the periods of growth and during balhimycin production (Schäberle et al., 2011). Deletion of $vanH_{Ab}A_{Ab}X_{Ab}$ genes makes *Am. balhimycina* significantly more sensitive to its own product, decreasing its MIC from 5 to 0.25 mg/ml, and causing an earlier expression of the BGC-situated $vanY_{Ab}$ (Frasch et al., 2015). However, $vanY_{Ab}$ itself does not play a decisive role in GPA-resistance since its deletion did not alter the GPA resistance phenotype

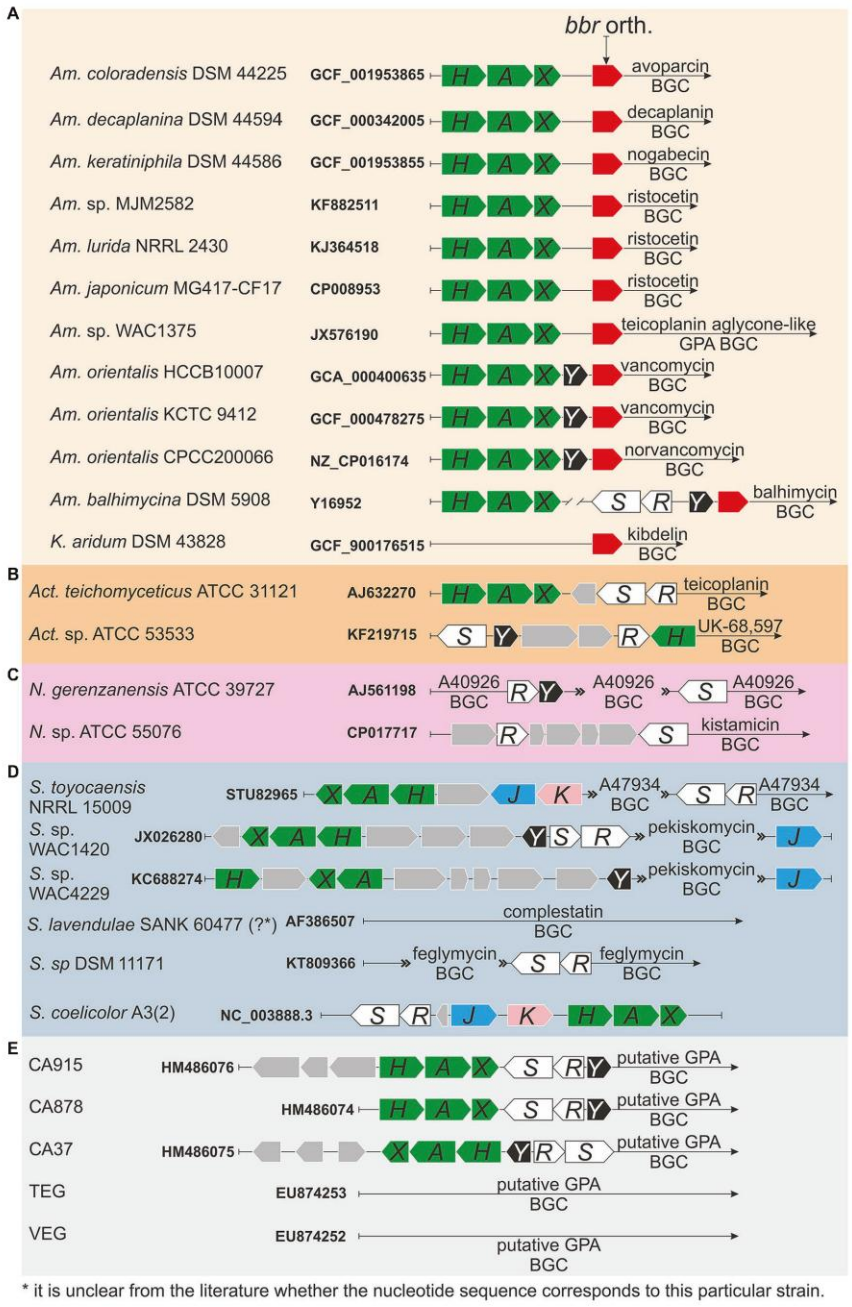


FIGURE 1 | (Continued)

FIGURE 1 | Organization of *van* genes orthologues in glycopeptide antibiotic (GPA) biosynthetic gene clusters (BGCs) of producers and metagenomics sequences: **(A)** family *Pseudonocardiaceae*: *Amycolatopsis coloradensis* DSM 44225 (GCF_001953865) – avoparcin producer (Kunstmann et al., 1968; Labeda, 1995); *Amycolatopsis decaplanina* DSM 44594 (GCF_000342005, Kaur et al., 2013) – decaplanin producer (Sanchez et al., 1992); *Amycolatopsis keratiniphila* subsp. *nogabecina* DSM 44586 (=*Amycolatopsis keratiniphila* subsp. *nogabecina* FH1893, GCF_001953855) – nogabecin producer (Wink et al., 2003); *Amycolatopsis* sp. MJM2582 (KF882511, Truman et al., 2014), *Amycolatopsis lurida* NRRL2430 (KJ364518, Truman et al., 2014), *Amycolatopsis japonicum* MG417-CF17 (CP008953, Spohn et al., 2014) – ristocetin producers; *Amycolatopsis* sp. WAC4169 (JX576190, Thaker et al., 2013) – producer of teicoplanin aglycone-like GPA; *Amycolatopsis orientalis* HCCB10007 (=*Am. keratiniphila*, GCA_000400635, Xu et al., 2014) and *Amycolatopsis orientalis* KCTC 9412 (=*Am. orientalis* DSM 40040, GCF_000478275, Jeong et al., 2013) – vancomycin producers; *Amycolatopsis orientalis* CPCC200066 (=*Am. orientalis* B-37, NZ_CP016174, Lei et al., 2015) – norvancomycin producer; *Amycolatopsis balhimycina* DSM 5908 (Y16952, Shawky et al., 2007) – balhimycin producer (Nadkarni et al., 1994; Wink et al., 2003); *Kibdelosporangium aridum* DSM 43828 (GCF_900176515) – the producer of kibdelins (Shearer et al., 1985), **(B)** genus *Actinoplanes* (please see main text for more details), **(C)** genus *Nonomuraea* (please see main text for more details), **(D)** genus *Streptomyces* (please see main text for more details); although *S. coelicolor* is not a GPA producer, organization of *S. coelicolor van* gene orthologues is also given, and **(E)** metagenomics sequences (please see main text for more details).

(Frasch et al., 2015). Double $vanH_{Ab}A_{Ab}X_{Ab}$ and $vanY_{Ab}$ knocked-out mutants showed the same GPA resistance phenotype as $\Delta vanH_{Ab}A_{Ab}X_{Ab}$. PG precursors ending in D-Ala-D-Lac were still found in the single $\Delta vanH_{Ab}A_{Ab}X_{Ab}$ and in the double $\Delta vanH_{Ab}A_{Ab}X_{Ab}$ $\Delta vanY_{Ab}$ mutants together with D-Ala-D-Ala ending PG precursors and tetrapeptides (Frasch et al., 2015). The residual GPA resistance in these mutants is probably due to an accessory $Ddl1_{Ab}$, a putative D-Ala-D-Lac ligase encoded in the genome of *Am. balhimycina*, which shares 72% of amino acid sequence identity with $VanA_{Ab}$ (Frasch et al., 2015). $Ddl1_{Ab}$ might add D-Lac to the tetrapeptide PG precursors generated by the D,D-carboxypetidase $VanY_{Ab}$ (although the presence of some other D-Ala-D-Ala carboxypeptidases encoded in the genome cannot be completely ruled out considering the resistant phenotype of the $\Delta vanY_{Ab}$ mutant). In the absence of $VanH_{Ab}$, D-Lac for this reaction is probably obtained from the primary metabolic pool. Expression of $vanH_{Ab}A_{Ab}X_{Ab}$ was demonstrated to be independent from the BGC-situated regulator $vlnR_{Ab}$ (Kilian et al., 2016). However, $VlnR_{Ab}$ is important for the activation of the BGC-situated $vanY_{Ab}$ expression (Kilian et al., 2016). Heterologous expression of $vlnR_{Ab}S_{Ab}$ in *S. coelicolor* Δ*vanRS* mutants indicated that both $VlnR_{Ab}$ and $VlnS_{Ab}$ are active and able to replace their counterparts VanR and VanS, which in *S. coelicolor* control the expression of *vanHAX* genes in response to vancomycin (Hong et al., 2004), restoring resistance to both balhimycin and teicoplanin in the complemented strains (Kilian et al., 2016). Overall, it seems that the BGC-situated $vlnR_{Ab}S_{Ab}$-$vanY_{Ab}$ regulatory circuit is functional, but does not play a major role in balhimycin resistance, which is mostly determined by $vanH_{Ab}A_{Ab}X_{Ab}$ expression. It would be interesting to test GPA resistance in $ddl1_{Ab}$ knocked-out mutant generated in *Am. balhimycina* $\Delta vanH_{Ab}A_{Ab}X_{Ab}$ $\Delta vanY_{Ab}$ to better understand the role of this accessory ligase and its connection with the D,D-carboxypetidase activity of $VanY_{Ab}$ (or of some other still-unknown carboxypeptidases).

Differently from *Am. balhimycina*, in *Act. teichomyceticus* *vanHAX* orthologues – *tei7-6-5* – are located within the *tei* BGC together with *vanRS* orthologues – *tei2-3* (**Figure 2B**). *tei7-6-5* expression determines the production of PG precursors ending in D-Ala-D-Lac, conferring a GPA-resistant phenotype to *Act. teichomyceticus* (Beltrametti et al., 2007; Binda et al., 2018). Interestingly, the expression of *tei7-6-5* operon is constant during the growth curve and in teicoplanin production conditions (Beltrametti et al., 2007; Yushchuk et al., 2019) and the VanX D,D-dipeptidase activity was detectable in cellular extracts independently from the addition of teicoplanin (Binda et al., 2018). One probable reason for the constitutive expression of *tei7-6-5* is the non-inducibility of the sensor histidine kinase Tei3, due to its point mutations previously known to confer a constitutive kinase activity to *S. coelicolor* VanS (Beltrametti et al., 2007). Also, the expression of *vanRS* orthologues – *tei2-3* – was also found constitutive under teicoplanin production conditions (Yushchuk et al., 2019) and these genes are co-expressed with *tei4* – coding for a dehydrofolate reductase with no obvious role in teicoplanin-resistance (Yushchuk et al., 2020b). Moreover, *tei2-3-4* expression is independent from *tei* cluster-encoded transcriptional regulators – Tei15* and Tei16* (Yushchuk et al., 2019, 2020b). Constitutive expression of *tei2-3-4* could be granted by *tei2* promoter, which was shown to be highly active in *Act. teichomyceticus*, starting from the very early stage of spore germination (Yushchuk et al., 2020b). More investigations are required for a complete understanding of teicoplanin-resistance in *Act. teichomyceticus*. Study of the Tei3 properties is among the most interesting tasks.

In *N. gerenzanensis* producing the teicoplanin-like A40926, *vanHAX* orthologues were not found neither in the BGC nor in the genome (D'Argenio et al., 2016). The only known mechanism of resistance relies on the action of $VanY_n$, whose coding gene (*dbv7*) is within the *dbv* BGC (**Figure 2C**) and whose knockout abolishes the resistance phenotype (Marcone et al., 2010, 2014; **Figure 2C**). $VanY_n$ is a D,D-carboxypeptidase that cleaves the last D-Ala from pentapeptide PG precursors generating tetrapeptides, drastically reducing GPA affinity for cellular targets (Binda et al., 2012). A L,D-transpeptidase (Ldt) then uses the tetrapeptide acyl donors supplied by $VanY_n$ to synthetize the mature cell wall (Hugonnet et al., 2014). The role of this protein and its features that assimilate/distinguish it from enterococcal VanY and from $VanY_{Ab}$ were investigated in detail (Marcone et al., 2010, 2014; Binda et al., 2012, 2013). Less clear is the regulatory circuit governing *dbv7* expression. Direct $VanY_n$ carboxypeptidase activity measurement in *N. gerenzanensis* growing with the addition of different GPAs, unambiguously showed that $VanY_n$ activity is induced by vancomycin, teicoplanin, and A40926 (Binda et al., 2018). *vanRS* homologues – *dbv6* and *dbv22* – are present in the *dbv* BGC, but the knockout of *dbv6* did not exert any influence on A40926 production and growth of

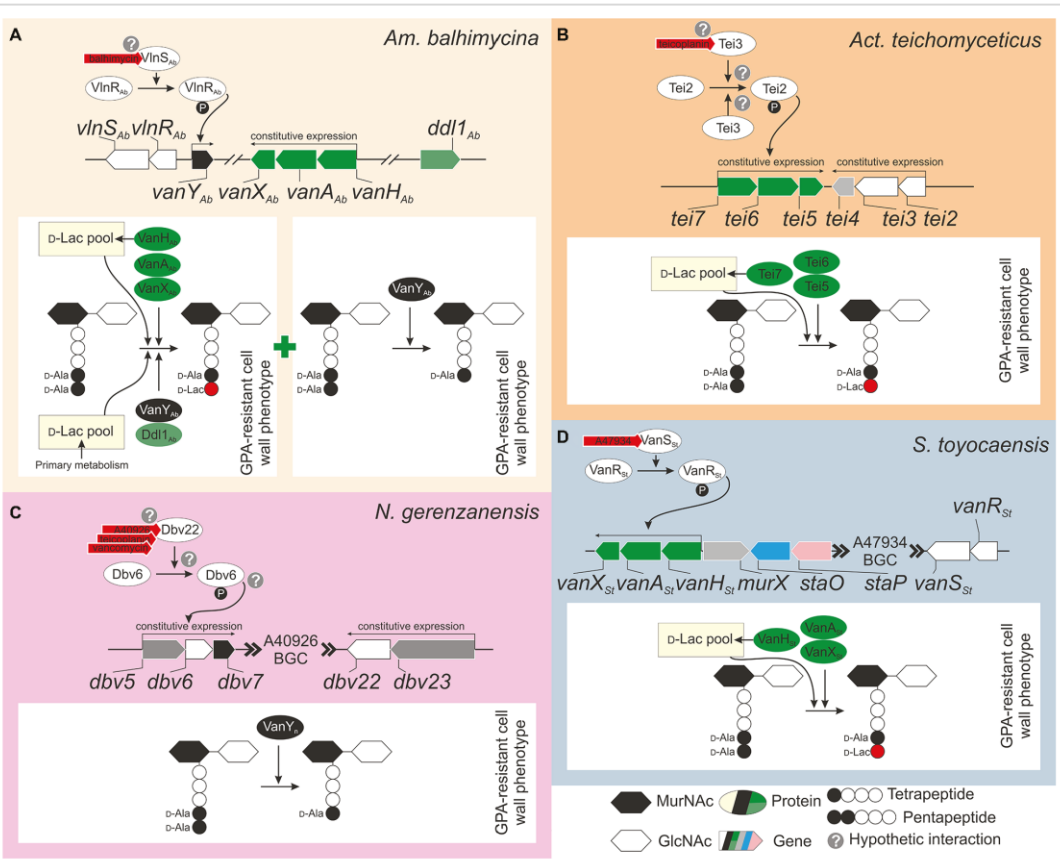


FIGURE 2 | Knowledge-based conceptual schemes for the GPA self-resistance mechanisms and their regulation in model GPA-producers: **(A)** *Am. balhimycina*, the producer of balhimycin, **(B)** *Act. teichomyceticus*, the producer of teicoplanin, **(C)** *N. gerenzanensis*, the producer of A40926, and **(D)** *S. toyocaensis*, the producer of A47934. Please refer to main text for more details on genes involved in GPA self-resistance. Please note that *tei4* in *Act. teichomyceticus*, *dbv5* and *dbv23* in *N. gerenzanensis*, and *murX* in *S. toyocaensis*, colored in light gray, are not *van* gene orthologues, but they are reported since they are close to *van* genes or part of their operons.

N. gerenzanensis (Lo Grasso et al., 2015). Unfortunately, the GPA resistance phenotype of this mutant was not described. On the other side, transcriptional analysis of *dbv* genes indicated that the expression of *dbv5-6-7* and *dbv23-22* operons is rather constitutive (Alduina et al., 2007). Although the presence of other GPA-sensitive TCS beyond the borders of A40926 BGC cannot be ruled out, role of Dbv6 and Dbv22 in *N. gerenzanensis* A40926 self-resistance merits further investigations.

Finally, *S. toyocaensis* possesses, perhaps, the most straightforward resistance mechanism among all the investigated GPA producers (**Figure 2D**), reminding the situation in *S. coelicolor* (Hong et al., 2004). The BGC-located $vanH_{St}A_{St}X_{St}$ operon was shown to be crucial for A47934 resistance and $vanA_{St}$ knockout made *S. toyocaensis* completely sensitive to A47934 (Pootoolal et al., 2002). At the same time, *vanHAX*-genes from the vancomycin producer *Am. orientalis* C329.2 were able to restore A47934 resistance phenotype in the knocked-out mutant (Pootoolal et al., 2002). Functions of $VanR_{St}$ and $VanS_{St}$ (both present in the A47934 BGC, **Figures 1, 2D**) were also studied in detail, showing that $VanS_{St}$ has a remarkable specificity for A47934 and it is unable to sense teicoplanin or vancomycin (Koteva et al., 2010; Novotna et al., 2016). Moreover, also the interaction between $VanR_{St}$ and $VanS_{St}$ was found to be very specific, since $VanR_{St}$ could not be phosphorylated by a non-cognate sensor-histidine kinase (Novotna et al., 2016). The roles (if there are any) of *staP* and *staO* (orthologues of *S. coelicolor vanK* and *vanJ*) in *S. toyocaensis* A47934 self-resistance were not investigated, thus the importance of these auxiliary resistance genes remains to be proved.

OUTLOOK

Soil GPA producers are considered the putative source of GPA resistance determinants, which might had been recruited and differently combined in pathogens. The goal of this mini review is to update the knowledge on the occurrence and role of *van* genes in producing microorganisms. It emerges that more *in silico*, *in vitro*, and *in vivo* investigations on their function and regulation are required to shed light on the intriguing issue of their origin and role. Overall, a detailed phylogenetic analysis would be useful to illuminate the evolution of GPA-resistant determinants in GPA producers and from them to pathogens. A recent pioneering work on the reconstruction of GPA BGC phylogeny (Waglechner et al., 2019) reported on the possible origin and evolution of GPA cluster-situated *van*-genes. According to these authors, *vanA* had likely originated within *Amycolatopsis* genus, whereas *vanH*, *vanX*, and *vanRS* within *Actinoplanes*; and *vanY* probably originated within genus *Nonomuraea* and it was then distributed among GPA BGCs by multiple transfer events. Combination of these genes in pathogens is today determining the urgent clinical need for new drugs to combat multi-drug resistant Gram-positive pathogens.

AUTHOR CONTRIBUTIONS

EB and OY collected data and papers and co-wrote the review. OY prepared the figures. FM and EB supervised the work.

FUNDING

This work was supported by grant "Fondo di Ateneo per la Ricerca" 2017, 2018 to FM, and Consorzio Italbiotec to EB.

ACKNOWLEDGMENTS

We are grateful to Consorzio Italbiotec for supporting EB.

REFERENCES

Adamek, M., Alanjary, M., Sales-Ortells, H., Goodfellow, M., Bull, A. T., Winkler, A., et al. (2018). Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC Genomics* 19:426. doi: 10.1186/s12864-018-4809-4

Ahmed, M. O., and Baptiste, K. E. (2018). Vancomycin-resistant enterococci: a review of antimicrobial resistance mechanisms and perspectives of human and animal health. *Microb. Drug Resist.* 24, 590–606. doi: 10.1089/mdr.2017.0147

Alduina, R., Piccolo, L. L., Dalia, D., Ferraro, C., Gunnarsson, N., Donadio, S., et al. (2007). Phosphate-controlled regulator for the biosynthesis of the dalbavancin precursor A40926. *J. Bacteriol.* 189, 8120–8129. doi: 10.1128/JB.01247-07

Arthur, M., Depardieu, F., Cabanié, L., Reynolds, P., and Courvalin, P. (1998). Requirement of the VanY and VanX D,D-peptidases for glycopeptide resistance in enterococci. *Mol. Mirobiol.* 30, 819–830. doi: 10.1046/j.1365-2958.1998.01114.x

Arthur, M., Depardieu, F., and Courvalin, P. (1999). Regulated interactions between partner and non-partner sensors and response regulators that control glycopeptide resistance gene expression in enterococci. *Microbiologica* 145, 1849–1858. doi: 10.1099/13500872-145-8-1849

Arthur, M., Depardieu, F., Gerbaud, G., Galimand, M., Leclercq, R., and Courvalin, P. (1997). The VanS sensor negatively controls VanR-mediated transcriptional activation of glycopeptide resistance genes of *Tn1546* and related elements in the absence of induction. *J. Bacteriol.* 179, 97–106. doi: 10.1128/JB.179.1.97-106.1997

Arthur, M., Molinas, C., Bugg, T. D., Wright, G. D., Walsh, C. T., and Courvalin, P. (1992). Evidence for in vivo incorporation of D-lactate into peptidoglycan precursors of vancomycin-resistant enterococci. *Antimicrob. Agents Chemother.* 36, 867–869. doi: 10.1128/aac.36.4.867

Arthur, M., and Quintiliani, R. (2001). Regulation of VanA- and VanB-type glycopeptide resistance in enterococci. *Antimicrob. Agents Chemother.* 45, 375–381. doi: 10.1128/aac.45.2.375-381.2001

Arthur, M., Reynolds, P., and Courvalin, P. (1996). Glycopeptide resistance in enterococci. *Trends Microbiol.* 4, 401–407. doi: 10.1016/0966-842X(96)10063-9

Banik, J. J., and Brady, S. F. (2008). Cloning and characterization of new glycopeptide gene clusters found in an environmental DNA megalibrary. *Proc. Natl. Acad. Sci. U. S. A.* 105, 17273–17277. doi: 10.1073/pnas.0807564105

Banik, J. J., Craig, J. W., Calle, P. Y., and Brady, S. F. (2010). Tailoring enzyme-rich environmental DNA clones: a source of enzymes for generating libraries of unnatural natural products. *J. Am. Chem. Soc.* 132, 15661–15670. doi: 10.1021/ja105825a

Bardone, M. R., Paternoster, M., and Coronelli, C. (1978). Teichomycins, new antibiotics from *Actinoplanes teichomyceticus* nov. sp. II. Extraction and chemical characterization. *J. Antibiot.* 31, 170–177. doi: 10.7164/antibiotics.31.170

Barka, E. A., Vatsa, P., Sanchez, L., Gavetas-Vaillant, N., Jacquard, C., Meier-Kolthoff, J. P., et al. (2015). Taxonomy, physiology, and natural products of Actinobacteria. *Microbiol. Mol. Biol. Rev.* 80, 1–43. doi: 10.1128/MMBR.00019-15

Bartley, J. (2002). First case of VRSA identified in Michigan. *Infect. Control Hosp. Epidemiol.* 23:480. doi: 10.1017/s0195941700082333

Beltrametti, F., Consolandi, A., Carrano, L., Bagatin, F., Rossi, R., Leoni, L., et al. (2007). Resistance to glycopeptide antibiotics in the teicoplanin producer is mediated by *van* gene homologue expression directing the synthesis of a modified cell wall peptidoglycan. *Antimicrob. Agents Chemother.* 51, 1135–1141. doi: 10.1128/AAC.01071-06

Bérdy, J. (2012). Thoughts and facts about antibiotics: where we are now and where we are heading. *J. Antibiot.* 65, 385–395. doi: 10.1038/ja.2012.27

Biavasco, F., Giovanetti, E., Miele, A., Vignaroli, C., Facinelli, B., and Varaldo, P. E. (1996). In vitro conjugative transfer of VanA vancomycin resistance between Enterococci and *Listeriae* of different species. *Eur. J. Clin. Microbiol. Infect. Dis.* 15, 50–59. doi: 10.1007/BF01586185

Billot-Klein, D., Blanot, D., Gutmann, L., and van Heijenoort, J. (1994). Association constants for the binding of vancomycin and teicoplanin to N-acetyl-D-alanyl-D-alanine and N-acetyl-D-alanyl-D-serine. *Biochem. J.* 304, 1021–1022. doi: 10.1042/bj3041021

Binda, E., Cappelletti, P., Marinelli, F., and Marcone, G. L. (2018). Specificity of induction of glycopeptide antibiotic resistance in the producing actinomycetes. *Antibiotics* 7, 36–46. doi: 10.3390/antibiotics7020036

Binda, E., Marcone, G. L., Berini, F., Pollegioni, L., and Marinelli, F. (2013). *Streptomyces* spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance. *BMC Biotechnol.* 13:24. doi: 10.1186/1472-6750-13-24

Binda, E., Marinelli, F., and Marcone, G. L. (2014). Old and new glycopeptide antibiotics: action and resistance. *Antibiotics* 3, 572–594. doi: 10.3390/antibiotics3040572

Binda, E., Marcone, G. L., Pollegioni, L., and Marinelli, F. (2012). Characterization of VanYn, a novel 485 D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727. *FEBS J.* 279, 3203–3213. doi: 10.1111/j.1742-487 4658.2012.08706.x

Bugg, T. D., Wright, G. D., Dutka-Malen, S., Arthur, M., Courvalin, P., and Walsh, C. T. (1991). Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by

vancomycin resistance proteins VanH and VanA. *Biochemist* 30, 10408–10415. doi: 10.1021/bi00107a007

Chiu, H.-T., Hubbard, B. K., Shah, A. N., Eide, J., Fredenburg, R. A., Walsh, C. T., et al. (2001). Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8548–8553. doi: 10.1073/pnas.151246498

Cong, Y., Yang, S., and Rao, X. (2020). Vancomycin resistant *Staphylococcus aureus* infections: a review of case updating and clinical features. *J. Adv. Res.* 21, 169–176. doi: 10.1016/j.jare.2019.10.005

Courvalin, P. (2006). Vancomycin resistance in gram-positive cocci. *Clin. Infect. Dis.* 42, S25–S34. doi: 10.1086/491711

D'Argenio, V., Petrillo, M., Pasanisi, D., Pagliarulo, C., Colicchio, R., Talà, A., et al. (2016). The complete 12 Mb genome and transcriptome of *Nonomuraea gerenzanensis* with new insights into its duplicated "magic" RNA polymerase. *Sci. Rep.* 6, 1–13. doi: 10.1038/s41598-016-0025-0

Depardieu, F., Podglajen, I., Leclercq, R., Collatz, E., and Courvalin, P. (2007). Modes and modulations of antibiotic resistance gene expression. *Clin. Microbiol. Rev.* 20, 79–114. doi: 10.1128/CMR.00015-06

Fontana, R., Ligozzi, M., Pedrotti, C., Padovani, E. M., and Cornaglia, G. C. (1997). Vancomycin-resistant *Bacillus circulans* carrying the vanA gene responsible for vancomycin resistance in enterococci. *Eur. J. Clin. Microbiol. Infect. Dis.* 16, 473–474. doi: 10.1007/BF02471915

Frasch, H.-J., Kalan, L., Kilian, R., Martin, T., Wright, G. D., and Stegmann, E. (2015). Alternative pathway to a glycopeptide-resistant cell wall in the balhimycin producer *Amycolatopsis balhimycina*. *ACS Infect. Dis.* 1, 243–252. doi: 10.1021/acsinfecdis.5b00011

Goldstein, B. P., Selva, E., Gastaldo, L., Berti, M., Pallanza, R., Ripamonti, F., et al. (1987). A40926, a new glycopeptide antibiotic with anti-*Neisseria* activity. *Antimicrob. Agents Chemother.* 31, 1961–1966. doi: 10.1128/aac.31.12.1961

Gonsior, M., Mühlenweg, A., Tietzmann, M., Rausch, S., Poch, A., and Süssmuth, R. D. (2015). Biosynthesis of the peptide antibiotic feglymycin by a linear nonribosomal peptide synthetase mechanism. *Chembiochem* 16, 2610–2614. doi: 10.1002/cbic.201500432

Healy, V. L., Lessard, I. A., Roper, D. I., Knox, J. R., and Walsh, C. T. (2000). Vancomycin resistance in enterococci: reprogramming of the D-ala-D-ala ligases in bacterial peptidoglycan biosynthesis. *Chem. Biol.* 7, R109–R119. doi: 10.1016/S1074-5521(00)00116-2

Hong, H. J., Hutchings, M. I., and Buttner, M. J. (2008). Vancomycin resistance VanS/VanR two-component systems. *Adv. Exp. Med. Biol.* 631, 200–213. doi: 10.1007/978-0-387-78885-2_14

Hong, H. J., Hutchings, M. I., Neu, J. M., Wright, G. D., Paget, M. S. B., and Buttner, M. J. (2004). Characterization of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (vanK) required for drug resistance. *Mol. Microbiol.* 52, 1107–1121. doi: 10.1111/j.1365-2958.2004.04032.x

Huang, J., Chen, L., Li, D., Wang, M., Du, F., Gao, Y., et al. (2018). Emergence of a vanG-n carrying and multidrug resistant ICE in zoonotic pathogen *Streptococcus suis*. *Vet. Microbiol.* 222, 109–113. doi: 10.1016/j.vetmic.2018.07.008

Hugonnet, J.-E., Haddache, N., Veckerlé, C., Dubost, L., Marie, A., Shikura, N., et al. (2014). Peptidoglycan cross-linking in glycopeptide-resistant actinomycetales. *Antimicrob. Agents Chemother.* 58, 1749–1756. doi: 10.1128/aac.02329-13

Jeong, H., Sim, Y. M., Kim, H. J., Lee, D.-W., Lim, S.-K., and Lee, S. J. (2013). Genome sequence of the vancomycin-producing *Amycolatopsis orientalis* subsp. *orientalis* strain KCTC 9412T. *Genome Announc.* 1, e00408–e00413. doi: 10.1128/genomea.00408-13

Kaur, N., Kumar, S., Bala, M., Raghava, G. P. S., and Mayilraj, S. (2013). Draft genome sequence of *Amycolatopsis decaplanina* strain DSM 44594T. *Genome Announc.* 1, e00138–e00113. doi: 10.1128/genomea.00138-13

Kilian, R., Frasch, H.-J., Kulik, A., Wohlleben, W., and Stegmann, E. (2016). The VanRS homologous two-component system VnlRSAb of the glycopeptide producer *Amycolatopsis balhimycina* activates transcription of the vanHAXSc genes in *Streptomyces coelicolor*, but not in *A. balhimycina*. *Microb. Drug Resist.* 22, 499–509. doi: 10.1089/mdr.2016.0128

Koteva, K., Hong, H.-J., Wang, X. D., Nazi, I., Hughes, D., Naldrett, M. J., et al. (2010). A vancomycin photoprobe identifies the histidine kinase VanSsc as a vancomycin receptor. *Nat. Chem. Biol.* 6, 327–329. doi: 10.1038/nchembio.350

Kunstmann, M. P., Mitscher, L. A., Porter, J. N., Shay, A. J., and Darken, M. A. (1968). LL-AV290, a new antibiotic. I. Fermentation, isolation, and characterization. *Antimicrob. Agents Chemother.* 8, 242–245.

Kwon, Y.-J., Kim, H.-J., and Kim, W.-G. (2015). Complestatin exerts antibacterial activity by the inhibition of fatty acid synthesis. *Biol. Pharm. Bull.* 38, 715–721. doi: 10.1248/bpb.b14-00824

Labeda, D. P. (1995). *Amycolatopsis coloradensis* sp., nov., the avoparcin (LL-AV290)-producing strain. *Int. J. Syst. Evol. Microbiol.* 45, 124–127. doi: 10.1099/00207713-45-1-124

Leclercq, R., Derlot, R., Duval, E. J., and Courvalin, P. (1988). Plasmid-mediated resistance to vancomycin and teicoplanin in *Enterococcus faecium*. *N. Engl. J. Med.* 319, 157–161. doi: 10.1056/NEJM198807213190307

Lei, X., Yuan, F., Shi, Y., Li, X., Wang, L., and Hong, B. (2015). Draft genome sequence of norvancomycin-producing strain *Amycolatopsis orientalis* CPCC200066. *Genome Announc.* 3, e00296–e00215. doi: 10.1128/genomea.00296-15

Li, T., Huang, F., Haydock, S., Mironenko, T., Leadlay, P., and Spencer, J. (2004). Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin characterization of two glycosyltransferases and the key acyltransferase. *Chem. Biol.* 11, 107–119. doi: 10.1016/j.chembiol.2004.01.001

Lo Grasso, L., Maffioli, S., Sosio, M., Bibb, M., Puglia, A. M., and Alduina, R. (2015). Two master switch regulators trigger A40926 biosynthesis in *Nonomuraea* sp. strain ATCC 39727. *J. Bacteriol.* 197, 2536–2544. doi: 10.1128/JB.00262-15

Marcone, G. L., Beltrametti, F., Binda, E., Carrano, L., Foulston, L., Hesketh, A., et al. (2010). Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother.* 54, 2465–2472. doi: 10.1128/AAC.00106-10

Marcone, G. L., Binda, E., Berini, F., and Marinelli, F. (2018). Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era. *Biotechnol. Adv.* 36, 534–554. doi: 10.1016/j.biotechadv.2018.02.009

Marcone, G. L., Binda, E., Carrano, L., Bibb, M., and Marinelli, F. (2014). Relationship between glycopeptide production and resistance in the actinomycete *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother.* 58, 5191–5201. doi: 10.1128/aac.02626-14

Marshall, C. G., Broadhead, G., Leskiw, B. K., and Wright, G. D. (1997). D-Ala-D-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB. *Proc. Natl. Acad. Sci. U. S. A.* 94, 6480–6483. doi: 10.1073/pnas.94.12.6480

Marshall, C. G., Lessard, I. A. D., Park, I.-S., and Wright, G. D. (1998). Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. *Antimicrob. Agents Chemother.* 42, 2215–2220. doi: 10.1128/AAC.42.9.2215

Miller, W. R., Murray, B. E., Rice, L. B., and Arias, C. A. (2016). Vancomycin-resistant enterococci: therapeutic challenges in the 21st century. *Infect. Dis. Clin. North Am.* 30, 415–439. doi: 10.1016/j.idc.2016.02.006

Nadkarni, S. R., Patel, M. V., Chatterjee, S., Vijayakumar, E. K. S., Desikan, K. R., Blumbach, J., et al. (1994). Balhimycin, a new glycopeptide antibiotic produced by *Amycolatopsis* sp. Y-86,21022. Taxonomy, production, isolation and biological activity. *J. Antibiot.* 47, 334–341. doi: 10.7164/antibiotics.47.334

Naruse, N., Oka, M., Konishi, M., and Oki, T. (1993). New antiviral antibiotics, Kistamicins A and B. II. Structure determination. *J. Antibiot.* 46, 1812–1818. doi: 10.7164/antibiotics.46.1812

Nazari, B., Forneris, C. C., Gibson, M. I., Moon, K., Schramma, K. R., and Seyedsayamdost, M. R. (2017). *Nonomuraea* sp. ATCC 55076 harbours the largest actinomycete chromosome to date and the kistamicin biosynthetic gene cluster. *Med. Chem. Commun.* 8, 780–788. doi: 10.1039/C6MD00637J

Nicolaou, K. C., Boddy, C. N. C., Bräse, S., and Winssinger, N. (1999). Chemistry, biology, and medicine of the glycopeptide antibiotics. *Angew. Chem. Int. Ed. Engl.* 38, 2096–2152. doi: 10.1002/(sici)1521-3773(19990802)38:15<2096::aid-anie2096>3.0.co;2-f

Novotna, G. B., Kwun, M. J., and Hong, H.-J. (2016). In vivo characterization of the activation and interaction of the VanR-VanS two-component regulatory system controlling glycopeptide antibiotic resistance in two related *Streptomyces* species. *Antimicrob. Agents Chemother.* 60, 1627–1637. doi: 10.1128/aac.01367-15

O'Neill, J. (ed.). (2016). *Tackling drug-resistant infections globally: final report and recommendations. The review on antimicrobial resistance.* Available at: http://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf

Patel, R., Piper, K., Cockerill, F. R., Steckelberg, J. M., and Yousten, A. A. (2000). The biopesticide *Paenibacillus popilliae* has a vancomycin resistance gene cluster homologous to the enterococcal VanA vancomycin resistance

gene cluster. *Antimicrob. Agents Chemother.* 44, 705–709. doi: 10.1128/aac.44.3.705-709.2000

Peltier, J., Courtin, P., Meouche, E. I., Catel-Ferreira, M., Chapot-Chartier, M. P., Lemée, L., et al. (2013). Genomic and expression analysis of the vanG-like gene cluster of *Clostridium difficile*. *Microbiologica* 159, 1510–1520. doi: 10.1099/mic.0.065060-0

Perkins, H. R., and Nieto, M. (1974). The chemical basis for the action of the vancomycin group of antibiotics. *Ann. N. Y. Acad. Sci.* 235, 348–363. doi: 10.1111/j.1749-6632.1974.tb43276.x

Pootoolal, J., Thomas, M. G., Marshall, C. G., Neu, J. M., Hubbard, B. K., Walsh, C. T., et al. (2002). Assembling the glycopeptide antibiotic scaffold: the biosynthesis of from *Streptomyces toyocaensis* NRRL15009. *Proc. Natl. Acad. Sci. U. S. A.* 99, 8962–8967. doi: 10.1073/pnas.102285099

Power, E. G. M., Abdulla, Y. H., Talsania, H. G., Spice, W., Aathithan, S., and French, G. L. (1995). vanA genes in vancomycin-resistant clinical isolates of *Oerskovia turbata* and *Arcanobacterium (Corynebacterium) haemolyticum*. *J. Antimicrob. Chemother.* 36, 595–606. doi: 10.1093/jac/36.4.595

Poyart, C., Pierre, C., Quesne, G., Pron, B., Berche, P., and Trieu-Cuot, P. (1997). Emergence of vancomycin resistance in the genus *Streptococcus*: characterization of a vanB transferable determinant in *Streptococcus bovis*. *Antimicrob. Agents Chemother.* 41, 24–29.

Reynolds, P. E., and Courvalin, P. (2005). Vancomycin resistance in enterococci due to synthesis of precursors terminating in D-alanyl-D-serine. *Antimicrob. Agents Chemother.* 49, 21–25. doi: 10.1128/AAC.49.1.21-25.2005

Reynolds, P. E., Depardieu, F., Dutka-Malen, S., Arthur, M., and Courvalin, P. (1994). Glycopeptide resistance mediated by enterococcal transposon *Tn1546* requires production of VanX for hydrolysis of D-alanyl-D-alanine. *Mol. Microbiol.* 13, 1065–1070. doi: 10.1111/j.1365-2958.1994.tb00497.x

Sanchez, M. L., Wenzel, R. P., and Jones, R. N. (1992). In vitro activity of decaplanin (M86-1410), a new glycopeptide antibiotic. *Antimicrob. Agents Chemother.* 36, 873–875. doi: 10.1128/aac.36.4.873

Schäberle, T. F., Vollmer, W., Frasch, H.-J., Hüttel, S., Kulik, A., Röttgen, M., et al. (2011). Self-resistance and cell wall composition in the glycopeptide producer *Amycolatopsis balhimycina*. *Antimicrob. Agents Chemother.* 55, 4283–4289. doi: 10.1128/aac.01372-10

Shawky, R. M., Puk, O., Wietzorrek, A., Pelzer, S., Takano, E., Wohlleben, W., et al. (2007). The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains *bbr*, encoding a StrR-like pathway-specific regulator. *J. Mol. Microbiol. Biotechnol.* 13, 76–88. doi: 10.1159/000103599

Shearer, M. C., Actor, P., Bowie, B. A., Grappel, S. F., Nash, C. H., Newman, D. J., et al. (1985). Aridicins, novel glycopeptide antibiotics. I. Taxonomy, production and biological activity. *J. Antibiot.* 38, 555–560. doi: 10.7164/antibiotics.38.555

Shearer, M. C., Giovenella, A. J., Grappel, S. F., Hedde, R. D., Mehta, R. J., Oh, Y. K., et al. (1986). Kibdelins, novel glycopeptide antibiotics. I. Discovery, production, and biological evaluation. *J. Antibiot.* 39, 1386–1394. doi: 10.7164/antibiotics.39.1386

Skelton, N. J., Williams, D. H., Monday, R. A., and Ruddock, J. C. (1990). Structure elucidation of the novel glycopeptide antibiotic UK-68,597. *J. Org. Chem. Res.* 55, 3718–3723. doi: 10.1021/jo00299a008

Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A., and Donadio, S. (2003). The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species. *Chem. Biol.* 10, 541–549. doi: 10.1016/S1074-5521(03)00120-0

Spohn, M., Kirchner, N., Kulik, A., Jochim, A., Wolf, F., Muenzer, P., et al. (2014). Overproduction of ristomycin A by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. *Antimicrob. Agents Chemother.* 58, 6185–6196. doi: 10.1128/aac.03512-14

Thaker, M. N., Wang, W., Spanogiannopoulos, P., Waglechner, N., King, A. M., Medina, R., et al. (2013). Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* 31, 922–927. doi: 10.1038/nbt.2685

Truman, A. W., Kwun, M. J., Cheng, J., Yang, S. H., Suh, J.-W., and Hong, H.-J. (2014). Antibiotic resistance mechanisms inform discovery: identification and characterization of a novel *Amycolatopsis* strain producing ristocetin. *Antimicrob. Agents Chemother.* 58, 5687–5695. doi: 10.1128/aac.03349-14

Vehreschild, M. J. G. T., Haverkamp, M., Biehl, L. M., Lemmen, S., and Fätkenheuer, G. (2019). Vancomycin-resistant enterococci (VRE): a reason to isolate? *Infection* 47, 7–11. doi: 10.1007/s15010-018-1202-9

van Wageningen, A. V., Kirkpatrick, P. N., Williams, D. H., Harris, B. R., Kershaw, J. K., Lennard, N. J., et al. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Biol.* 5, 155–162. doi: 10.1016/S1074-5521(98)90060-6

Waglechner, N., McArthur, A. G., and Wright, G. D. (2019). Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat. Microbiol.* 4, 1862–1871. doi: 10.1038/s41564-019-0531-5

Weigel, L. M., Clewell, D. B., Gill, S. R., Clark, N. C., McDougal, L. K., Flannagan, S. E., et al. (2003). Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. *Science* 302, 1569–1571. doi: 10.1126/science.1090956

WHO (2017). Antibacterial agents in clinical development: an analysis of the antibacterial clinical development pipeline, including tuberculosis. World Health Organization. WHO/EMP/IAU/2017.11

Wink, J. M., Kroppenstedt, R. M., Ganguli, B. N., Nadkarni, S. R., Schumann, P., Seibert, G., et al. (2003). Three new antibiotic producing species of the genus *Amycolatopsis*, *Amycolatopsis balhimycina* sp. nov., *A. tolypomycina* sp. nov., *A. vancoresmycina* sp. nov., and description of *Amycolatopsis keratiniphila* subsp. *keratiniphila* subsp. nov. and *A. keratiniphila* subsp. *nogabecina* subsp. nov. *Syst. Appl. Microbiol.* 26, 38–46. doi: 10.1078/072320203322337290

Wright, G. D., Holman, T. R., and Walsh, C. T. (1993). Purification and characterization of VanR and the cytosolic domain of VanS: a two-component regulatory system required for vancomycin resistance in *Enterococcus faecium* BM4147. *Biochemist* 32, 5057–5063. doi: 10.1021/bi00070a013

Wu, Z., Wright, G. D., and Walsh, C. T. (1995). Overexpression, purification, and characterization of VanX, a D,D-dipeptidase which is essential for vancomycin resistance in *Enterococcus faecium* BM4147. *Biochemist* 34, 2455–2463. doi: 10.1021/bi00008a008

Xu, L., Huang, H., Wei, W., Zhong, Y., Tang, B., Yuan, H., et al. (2014). Complete genome sequence and comparative genomic analyses of the vancomycin-producing *Amycolatopsis orientalis*. *BMC Genomics* 15:363. doi: 10.1186/1471-2164-15-363

Yim, G., Kalan, L., Koteva, K., Thaker, M. N., Waglechner, N., Tang, I., et al. (2014). Harnessing the synthetic capabilities of glycopeptide antibiotic tailoring enzymes: characterization of the UK-68,597 biosynthetic cluster. *Chembiochem* 15, 2613–2623. doi: 10.1002/cbic.201402179

Yushchuk, O., Homoniuk, V., Ostash, B., Marinelli, F., and Fedorenko, V. (2020b). Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. *J. Antibiot.* 73, 255–259. doi: 10.1038/s41429-019-0274-9

Yushchuk, O., Horbal, L., Ostash, B., Marinelli, F., Wohlleben, W., Stegmann, E., et al. (2019). Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. *Appl. Microbiol. Biotechnol.* 103, 4089–4102. doi: 10.1007/s00253-019-09789-w

Yushchuk, O., Ostash, B., Truman, A. W., Marinelli, F., and Fedorenko, V. (2020a). Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes. *Appl. Microbiol. Biotechnol.* 104, 3279–3291. doi: 10.1007/s00253-020-10436-y

Zeng, D., Debabov, D., Hartsell, T. L., Cano, R. J., Adams, S., Schuyler, J. A., et al. (2016). Approved glycopeptide antibacterial drugs: mechanism of action and resistance. *Cold Spring Harb. Perspect. Med.* 6:a026989. doi: 10.1101/cshperspect.a026989

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Yushchuk, Binda and Marinelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Оригінал доступний за посиланням:
https://www.frontiersin.org/articles/10.3389/fmicb.2020.01173/full.

Анотація: *Глікопептидні антибіотики (ГПА) вважаються препаратами «останньої інстанції» для лікування небезпечних для життя інфекцій, спричинених Грам-позитивними збудниками (ентерококами, стафілококами чи клостридіями). У сучасну «постгеномну» епоху активно досліджується біосинтез ГПА і механізми резистентності бактерій до цих антибіотиків. Вважається, що гени стійкості до ГПА були отримані патогенами від актиноміцетів – ґрунтових організмів, які продукують антибіотики. Гени стійкості необхідні актиноміцетам-продуцентам для того, щоб уникнути самогубства під час біосинтезу. Все більше і більше геномів продуцентів ГПА секвенуються, і в них знаходять різні гени стійкості до ГПА (названі van-генами). У актиноміцетів-продуцентів van-гени зазвичай асоційовані з кластерами генів біосинтезу (КБГ) ГПА, що, ймовірно, сприяє спільній регуляції біосинтезу антибіотиків і стійкості до них. Також відомо, що асоційовані з КБГ ГПА van-гени трапляються в метагеномних послідовностях. Цікаво, що деякі КБГ для антибіотиків, здавалося б, не пов'язаних з ГПА (наприклад, фегліміцину), містять гомологи van-генів. Тут ми хотіли б висвітлити останні досягнення присвячені поширенню генів стійкості до ГПА в геномах і метагеномах, що належать фактичним або потенційним мікроорганізмам-продуцентам ГПА. Глибоке розуміння механізмів стійкості до ГПА у мікроорганізмів-продуцентів може виявитися корисним у майбутньому для передбачення та відслідковування нових механізмів стійкості до представників цього клінічно значущого класу антибіотиків.*

Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-020-10436-y

MINI-REVIEW

Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes

Oleksandr Yushchuk[1,2] · Bohdan Ostash[1] · Andrew W. Truman[3] · Flavia Marinelli[2] · Victor Fedorenko[1]

Received: 14 November 2019 / Revised: 26 January 2020 / Accepted: 4 February 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
Teicoplanin (Tcp) is a clinically relevant glycopeptide antibiotic (GPA) that is produced by the actinobacterium *Actinoplanes teichomyceticus*. Tcp is a front-line therapy for treating severe infections caused by multidrug-resistant Gram-positive pathogens in adults and infants. In this review, we provide a detailed overview of how Tcp is produced by *A. teichomyceticus* by describing Tcp biosynthesis, regulation, and resistance. We summarize the knowledge gained from in vivo and in vitro studies to provide an integrated model of teicoplanin biosynthesis. Then, we discuss genetic and nutritional factors that contribute to the regulation of teicoplanin biosynthesis, focusing on those that have been successfully applied for improving teicoplanin production. A current view on teicoplanin self-resistance mechanisms in *A. teichomyceticus* is given, and we compare the Tcp biosynthetic gene cluster with other glycopeptide gene clusters from actinoplanetes and from unidentified isolates/metagenomics samples. Finally, we provide an outlook for further directions in studying Tcp biosynthesis and regulation.

Keywords Glycopeptide antibiotics · Teicoplanin · *Actinoplanes teichomyceticus* · Teicoplanin biosynthesis · Teicoplanin resistance · Regulation

Introduction

Initially known as teichomycin A_2, teicoplanin (Tcp) was co-purified together with moenomycin-like teichomycin (initially teichomycin A_1) from an uncommon sporangia-forming actinomycete *Actinoplanes teichomyceticus* ATCC 31121, which was isolated from a soil sample collected near Nimodi in India (Parenti et al. 1978; Bardone et al. 1978). Tcp is a glycopeptide antibiotic (GPA) that binds the D-Ala-D-Ala terminus of lipid II in the last (extracellular) step of peptidoglycan biosynthesis, thereby leading to bacterial death (Parenti 1986). Tcp is not the sole representative of *Actinoplanes*-derived glycopeptides (Nicolaou et al. 1999). Other examples of GPAs produced by actinoplanetes are UK-68,597 (Skelton et al. 1990) and actaplanin (Debono et al. 1984). However, these compounds (Fig. 1) have not been studied in as much detail as Tcp, which is the only one that was introduced into clinical practice.

Tcp is currently marketed in the EU and other countries to treat infections caused mainly by multidrug resistant (MDR) Gram-positive pathogens, including methicillin-resistant *Staphylococcus aureus*. As such, Tcp is currently one of a few remaining antibiotics that can be used to counter the threat of infections caused by MDR pathogens. One remarkable feature of Tcp, as a marketed drug, is that it is actually a mixture of closely related compounds with the five major components (A_{2-1}-A_{2-5}) differing in the length and branching of an aliphatic chain (Fig. 1). This attests to the chemical complexity of teicoplanin and to the challenges for its reproducible production and purification. There is a renewed interest in GPAs in general and Tcp in particular as a platform for developing next-generation antibiotics. For instance, teicoplanin has been recently conjugated with nanoparticles, where it was shown to have increased activity against biofilm-forming pathogens (Armenia et al. 2018). The chemistry of Tcp, its medical properties and the resistance mechanisms emerging to it in pathogens have been reviewed in recent years (Jung et al. 2009;

✉ Victor Fedorenko
viktor.fedorenko@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 4 Hrushevskoho st., Rm. 102, Lviv 79005, Ukraine

[2] Department of Biotechnology and Life Sciences, University of Insubria, Varese, Italy

[3] Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, UK

Published online: 19 February 2020

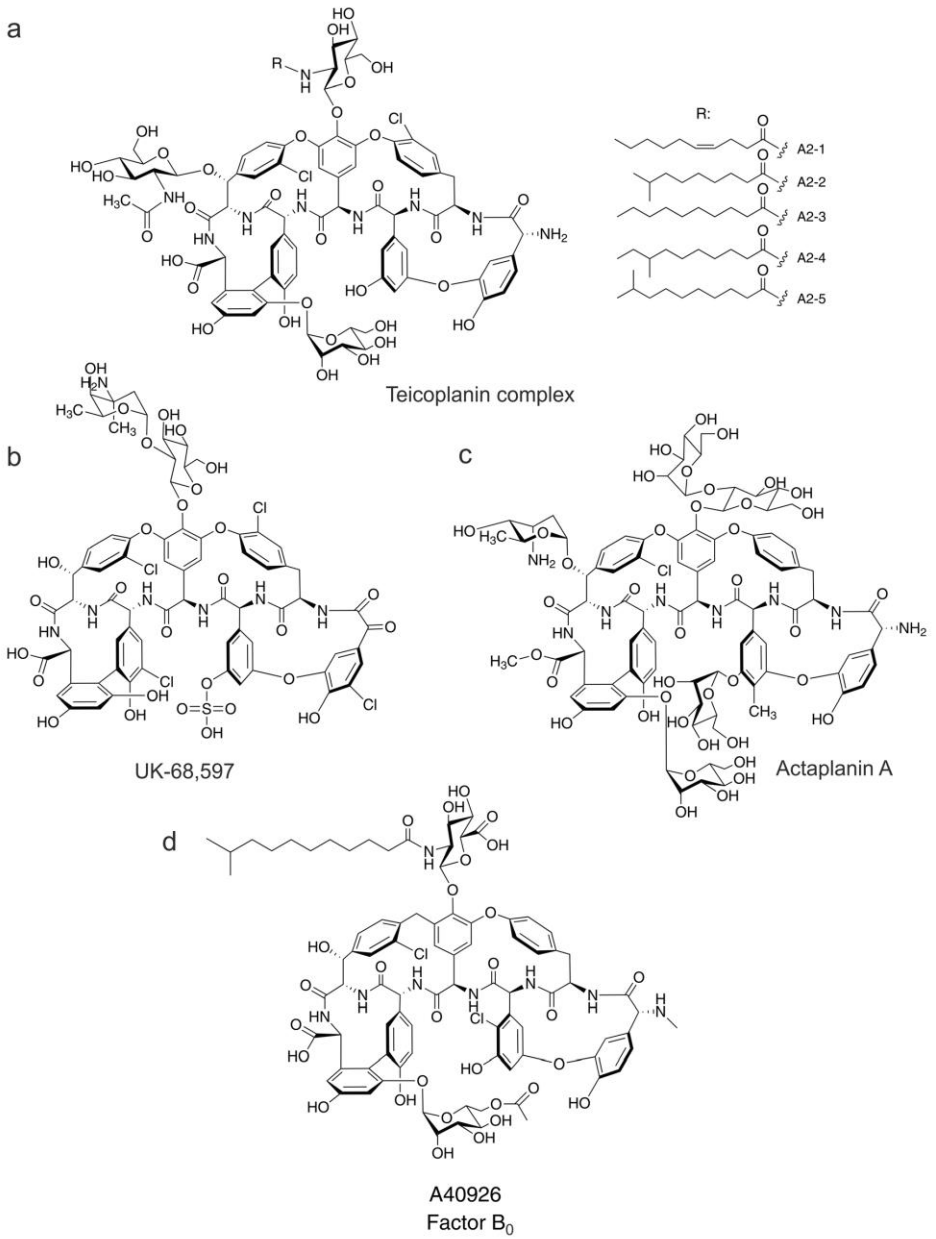


Fig. 1 Chemical structures of teicoplanin complex (**a**), two other *Actinoplanes*-derived GPAs: UK-68,597 (**b**) produced by *Actinoplanes* sp. ATCC 53533 and actaplanin A (**c**) produced by *Actinoplanes missouriensis* ATCC 23342, A40926 Factor B_0 produced by *Nonomuraea gerenzanensis* ATCC 39727 (**d**). Actaplanin A is actually a major compound of a multicomponent complex of related antibiotics, which differ in glycosylation pattern. Similarly, A40926 Factor B_0 is the major component of A40926 complex, where related compounds differ in the length and branching of aliphatic side chains

James et al. 2012; Roberts et al. 2014; Binda et al. 2014; Marcone et al. 2018). Here, we focus on Tcp biosynthesis in *A. teichomyceticus*, where we integrate recent advances in understanding the genes and enzymes for Tcp assembly, regulation and resistance.

Tcp biosynthesis: from genes to compound

The Tcp biosynthetic gene cluster (BGC) (*tei*) was sequenced and annotated independently by two groups in 2004 (Li et al. 2004; Sosio et al. 2004). Although some details concerning mainly the borders of the *tei* cluster differed, both annotations were largely consistent with each other. The *tei* cluster is believed to span approximately 78 kbp containing 42 ORFs organized in 18 transcriptional units, starting with *tei7* and ending with *tei31** (Yushchuk et al. 2019). The majority of *tei* genes are actively transcribed under Tcp producing conditions. The only exceptions are *tei25**, *tei26**, *tei27**, and *tei31** (encoding a metallohydrolase, a type II chorismate mutase, a siderophore-interacting protein and a transcriptional regulator, respectively), which were found to be silent under Tcp production conditions (Yushchuk et al. 2019). Functions of the majority of *tei* genes either were studied experimentally (Li et al. 2004; Truman et al. 2006; Horbal et al. 2014; Yushchuk et al. 2016) or could be deduced from the results obtained from other GPA pathways.

Biosynthesis of Tcp consists of three distinct parts: (a) biosynthesis of non-proteinogenic precursors; (b) non-ribosomal peptide synthetase (NRPS) assembly of the heptapeptide aglycone; (c) post-NRPS tailoring of the aglycone. Supply of precursor amino acids is crucial for teicoplanin biosynthesis. Out of the seven amino acids forming the Tcp aglycone, only the second and the sixth—chlorinated L-tyrosine and chlorinated β-hydroxytyrosine (Fig. 2)—enter non-ribosomal synthesis in form of the proteinogenic amino acid tyrosine (which undergoes chlorination and hydroxylation during non-ribosomal synthesis). Formation of β-hydroxytyrosine is likely carried out by nonheme dioxygenase (β-hydroxylase) Tei12*, an orthologue of the experimentally studied Dbv28 (85.1% amino acid sequence identity), which functions simultaneously with non-ribosomal synthesis or just after it, during the production of the A40926 glycopeptide in *Nonomuraea gerenzanensis* (Stinchi et al. 2006). All other amino acids are non-proteinogenic: three D-4-hydroxyphenylglycine moieties in positions 1, 4 and 5 together with two L-3,5-dihydroxyphenylglycine moieties in positions 3 and 7 (Fig. 2). Nevertheless, even these non-proteinogenic amino acids are directly or indirectly derived from tyrosine. Indeed, the importance of a tyrosine pool as a GPA precursor was previously shown for the balhimycin production in *Amycolatopsis balhimycina* (Thykaer et al. 2010). To satisfy the high demand in tyrosine, the balhimycin BGC carries two genes for the shikimate pathway enzymes: 3-deoxy-D-arabino-heptulosonate-7-phosphate (Dahp) synthase and prephenate dehydrogenase (Pdh). It is very likely that *tei*-encoded homologues of Dahp and Pdh, Tei14*, and Tei24*, respectively, play a similar role in supplying tyrosine to Tcp biosynthesis. Additionally, the *tei* cluster possesses a gene (*tei26**) coding for a type II chorismate mutase (CM) that could also be involved in tyrosine biosynthesis. Among other known GPA BGCs, Tei26* has orthologues encoded in esnapd15 and CA37 BGCs, which are putative GPAs retrieved from metagenomic sample genetic screening (Banik et al. 2010; Owen et al. 2013). The other *tei*-related BGCs—*auk* and CA915 (this first being a BGC for UK-68,597, Yim et al. 2014, and the second once again coming from an uncultivable actinobacterium, Banik et al. 2010)—also possess CM-encoding genes, but they belong to type I CMs (orthologues of Auk33). Three other *tei* encoded enzymes (dihydrofolate reductase Tei4, GTP cyclohydrolase Tei22*, and non-heme dioxygenase Tei25*) could be involved in tyrosine biosynthesis by hydroxylation of phenylalanine in the presence of Fe^{2+} (Li et al. 2004). However, neither *tei26** nor *tei25** was found to be actively expressed under Tcp-producing conditions (Yushchuk et al. 2019).

Although *tei*-encoded enzymes for the biosynthesis of non-proteinogenic amino acids were not investigated experimentally, their functions could be inferred from comparison to experimentally studied orthologues encoded within the other GPA BGCs. Thus, the biosynthesis of D-4-hydroxyphenylglycine (Hpg) is likely governed by Tei23*, Tei28*, and Tei29*, which are orthologues of experimentally studied HpgT (*p*-hydroxyphenylglycine transaminase, 47% amino acid sequence identity), HmaS (4-hydroxymandelate synthase, 51.7%), and Hmo (hydroxymandelate oxidase, 54%) from the chloroeremomycin BGC in *Amycolatopsis orientalis* (Hubbard et al. 2000; Choroba et al. 2000). Thus, biosynthesis of L-*p*-hydroxymandelate from *p*-hydroxyphenylpyruvate could be catalyzed by Tei28*, with the consequential transformation of L-*p*-hydroxymandelate into *p*-hydroxybenzoylformate catalyzed by Tei29*. Finally, formation of L-*p*-hydroxyphenylglycine from *p*-hydroxybenzoylformate could be catalyzed by Tei23*. Although L-*p*-hydroxyphenylglycine originates from tyrosine directly, the latter also might serve as an amino donor in the reaction catalyzed by Tei23*. Genes involved in the biosynthesis of L-3,5-dihydroxyphenylglycine could be inferred likewise. These are *tei17*-18*-19*-20**, orthologues of *dpgA-B-C-D* (77.2%, 52.3%, 63.9%, and 73.8% amino acid sequence identity, respectively) from the chloroeremomycin BGC (Chen et al. 2001). The DpgA orthologue, type III polyketide synthase Tei17*, is predicted to form a complex with the enoyl-CoA hydratases Tei18* and Tei20* and convert malonyl-CoA into 3,5-dihydroxyphenylacetyl-CoA. This would then be converted into 3,5-dihydroxyphenylglyoxylate (and, consequently, to L-3,5-dihydroxyphenylglycine) by

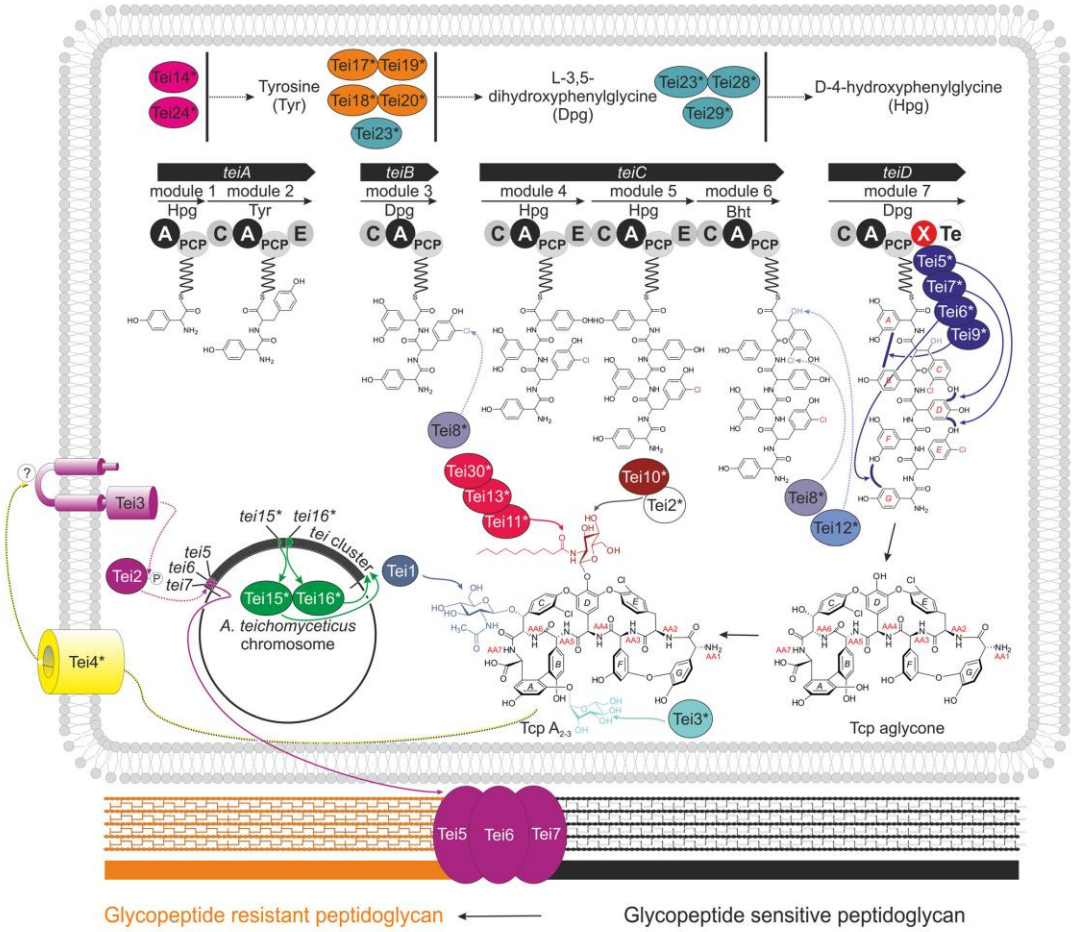


Fig. 2 Conceptual scheme, integrating Tcp biosynthesis, export, pathway-specific regulation and resistance. Presumed interactions are depicted with dash lines and solid lines show interactions supported with experimental evidence. Please see the main text for more details

Tei19* (3,5-dihydroxyphenylacetyl-CoA 1,2-dioxygenase). Finally, by analogy to balhimycin (Pfeifer et al. 2001) and chloroeremomycin (Sandercock et al. 2001) biosynthesis, the transformation of 3,5-dihydroxyphenylglyoxylate into L-3,5-dihydroxyphenylglycine would be catalyzed by the HpgT-Pgat orthologue Tei23*.

When the pool of precursor amino acids is synthesized, the non-ribosomal synthesis of heptapeptide aglycone begins. Four genes, *teiA-D* (Fig. 2), encode the NRPS in the *tei* cluster (Li et al. 2004). This NRPS is organized into seven modules, each responsible for the incorporation of one amino acid into the aglycone (Fig. 2). In all GPA NRPSs, such NRPS organization pattern provides characteristic stereochemical configuration of D-D-L-D-D-L-L amino acid residues in the aglycone. This configuration is crucial for the compound to adopt the correct conformation for successful antibacterial action (James et al. 2012). Intriguingly, the first modules of all GPA NRPSs lack an epimerization domain usually required for the generation of a D-amino acid. A unique feature of GPA NRPSs is the X-domain of module 7. This domain is a descendent of a C-domain from the ancestral proto-GPA NRPS (Peschke et al. 2016a). This recruits cytochrome P450 oxygenases involved in cyclization of the NRPS-bound heptapeptide (Haslinger et al. 2015; Peschke et al. 2016b), thereby providing an evolutionary rationale for retention of this catalytically inert domain.

The *tei* cluster encodes four oxygenases involved in linear heptapeptide cross-linking: Tei5* (OxyA), Tei6* (OxyE), Tei7* (OxyB), and Tei9* (OxyC). These P450s catalyze oxidative cross-linking of the aromatic side chains of Tcp,

forming O-C or C-C bonds. By analogy to A47934 biosynthesis in *Streptomyces toyocaensis* (Hadatsch et al. 2007), Tei oxygenases are predicted to catalyze the cross-linking reactions in the following order: Tei7* (*C-O-D*) → Tei6* (*F-O-G*) → Tei5* (*D-O-E*) → Tei9* (*A-B*) (Fig. 2). Additionally, recent in vitro experiments involving the module 7 X-domain, Tei5* and Tei7* (Schoppet et al. 2018; Peschke et al. 2016c) have shown complex dynamic protein-protein interactions underlying the oxidative cross-linking steps. Halogenation is another reaction tightly correlated with the non-ribosomal peptide synthesis. The *tei* cluster encodes a single halogenase, Tei8*, which chlorinates tyrosine and β-hydroxytyrosine residues that are incorporated into the aglycone at positions 2 and 6, respectively. Tei8* is active only on single amino acids loaded onto the peptidyl carrier protein (PCP) domain of the NRPS (Kittilä et al. 2017), meaning that chlorination occurs just before the incorporation of the amino acids into the growing oligopeptide. This evidence together with data revealing the preference of Tei5* and Tei7* for chlorinated substrates (Peschke et al. 2017) unambiguously couples chlorination event with non-ribosomal synthesis of the hexapeptide aglycone.

After the halogenated and cross-linked aglycone is assembled, it is decorated with three sugar residues and with an aliphatic side chain. Even though the aglycone is the key determinant of D-Ala-D-Ala binding, Tcp bioactivity is augmented by the further chemical derivatization of the aglycone. Snyder et al. showed that removal of the fatty acyl chain from Tcp reduced the antimicrobial activity against a range of enterococcal and staphylococcal strains (Snyder et al. 1998), while decoration of vancomycin with a teicoplanin-like N-acyl-glucosaminyl moiety increased its activity towards vancomycin-resistant enterococci (Thayer and Wong 2006). Second-generation lipidated GPAs as oritavancin and telavancin can also disrupt bacterial membranes and therefore provide an alternative antimicrobial mechanism (Yarlagadda et al. 2015). The sugars that decorate teicoplanin are products of primary metabolism, although many other GPAs contain sugars produced by specialized pathways encoded in the GPA BGC (e.g. Chen et al. 2000).

The teicoplanin aglycone is decorated with three sugars and one lipid chain to transform it into the mature lipoglycopeptide. Accordingly, three glycosyltransferases (Tei1, Tei3*, and Tei10*) and one acyltransferase (Tei11*) are encoded in the *tei* BGC. Multiple in vitro studies of the biosynthesis of Tcp (and the related GPA, A40926) provide a detailed understanding of the mechanism of these tailoring steps, and a gene inactivation study indicated a likely order of tailoring steps (Yushchuk et al. 2016).

These studies indicate that Tcp tailoring begins with Tei10*-catalyzed transfer of N-acetylglucosamine to the 4-hydroxyl group of 4-hydroxyphenylglycine (amino acid 4) using uridine diphospho-N-acetylglucosamine (UDP-GlcNAc) as the sugar donor (Li et al. 2004). This is then deacetylated by Tei2*, a pfam02585 family deacetylase, to yield a glucosaminyl moiety (Truman et al. 2006). Tei11* then transfers an acyl chain to the free amino group of glucosaminyl moiety, using acyl-coenzyme A as a donor. In vitro characterization of Tei11* made this the founding member of a rare family of acyltransferases (Li et al. 2004; Kruger et al. 2005) with a unique structural fold (Lyu et al. 2014). The substrate flexibility of Tei11* accounts for the production of multiple Tcp congeners that differ in their acyl chain length (Lazzarini et al. 1997; Kruger et al. 2005).

Genetic inactivation of *tei11** led to an accumulation of a Tcp derivative lacking N-acetylglucosamine on β-hydroxytyrosine (amino acid 6) and mannose on dihydroxyphenylglycine (amino acid 7) (Yushchuk et al. 2016). This provides strong evidence that these are the final biosynthetic steps in the pathway. Biochemical analysis demonstrated that Tei1 transfers the N-acetylglucosamine moiety to the β-hydroxyl group of β-hydroxytyrosine (Li et al. 2004), while genetic inactivation of *tei3** showed that Tei3* attaches a mannose group to the phenolic oxygen adjacent to the C-C cross-link on amino acid 7 (Yushchuk et al. 2016). Tei3* belongs to glycosyltransferase family 39 (inverting GT-C fold protein) and is predicted to use undecaprenyl-phosphomannose as the donor.

In addition to these essential tailoring proteins, two proteins were identified as being associated with Tcp biosynthesis but are not essential: Tei13* (acyl-CoA synthetase) and Tei30* (thioesterase). Inactivation of corresponding genes led to reduction in the proportion of C10 acyl chains on the N-acyl-glucosaminyl moiety on amino acid 4 (Yushchuk et al. 2016). However, a recent in vitro analysis of the Tcp NRPS indicates that Tei30* can alternatively function as a "proof-reading" type II thioesterase that removes long-lived PCP-bound amino acids that are not substrates for the NRPS (Kaniusaite et al. 2019). Unusually, there are no close homologues of Tei30* encoded in other GPA gene clusters.

Regulation of Tcp biosynthesis: genetic and nutritional factors

In streptomycetes, morphological differentiation and secondary metabolism are tightly coordinated (McCormick and Flärdh 2012; Jones and Elliot 2018) by sophisticated regulatory circuits responding to environmental signals. The components of these circuits can be used as tools for improving antibiotic production (Makitrynskyy et al. 2013; Yushchuk et al. 2018). Our understanding of global regulation of Tcp biosynthesis in *A. teichomyceticus* is still very fragmentary. Nevertheless, attempts to use *Streptomyces*-derived knowledge to improve Tcp production have been made. Pioneering works made on *Streptomyces coelicolor* (Okamoto et al. 2003) and *Streptomyces lividans* (Kim et al. 2003) showed the importance of intracellular S-adenosylmethionine (SAM)

for the onset of secondary metabolism. Manipulations of SAM synthase gene *metK* led to improved antibiotic production by different strains (Okamoto et al. 2003; Pang et al. 2015). Similarly, *metK* orthologue was detected in *A. teichomyceticus* and cloned under the control of *ermEp* (Kim et al. 2011). Overexpression of *metK* in *A. teichomyceticus* ATCC 31121 elevated Tcp production by more than twofold (Fig. 3). Likewise, overexpression of *A. teichomyceticus* (p)ppGpp synthetase *relA* orthologue increased Tcp production by threefold (Fig. 3) and had an even more significant impact on actinorhodin production when was heterologously expressed in *S. lividans* TK24 (Park et al. 2011). These two examples allow us to suggest that SAM- and (p)ppGpp-related regulatory mechanisms are basic enough to work similarly in distant genera such as *Actinoplanes* and *Streptomyces*. However, not all regulatory mechanisms work similarly in *Streptomyces* and *Actinoplanes*. In streptomycetes, AdpA orthologues are one of the most important global regulators of different secondary metabolic pathways (Ostash 2018). A set of *adpA* homologues was also identified in the genome of *A. teichomyceticus*, yet overexpression of the closest one to *adpA* did not exert any influence on Tcp production by *A. teichomyceticus* (Ostash et al. 2015).

So far, application of the heterologous *hrdB* genes (coding for main vegetative RNA-polymerase σ factor) to improve teicoplanin production is a single example of transcriptional engineering in *A. teichomyceticus* (Fig. 3; Wang et al. 2014). In this case, random mutagenesis was performed on a set of heterologous *hrdB* genes originating from *Micromonosporaceae*. This collection of mutant *hrdB* alleles was then expressed in *A. teichomyceticus* L-27 (industrial Tcp overproducer) under the control of *ermEp*. Among the resulting set of strains overexpressing mutant *hrdB* alleles in addition to the intact native one, some strains surpassed the production of Tcp in *A. teichomyceticus* L-27 by more than twofold (Wang et al. 2014).

Finally, after the development of the genetic toolkit to manipulate *A. teichomyceticus* (Truman et al. 2008; Horbal et al. 2012; Horbal et al. 2013), genetic factors that control Tcp production were investigated rationally. It was shown that two cluster-encoded regulators—Tei15* and Tei16*—are key players in the pathway-specific regulation of Tcp production (Figs. 2 and 3, Horbal et al. 2014). The third pathway-specific regulator encoded in *tei* cluster—Tei31*—appeared to be dispensable for Tcp production (Yushchuk et al. 2019). Both Tei15* and Tei16* together represented a regulatory circuit governing the expression of the majority of *tei*

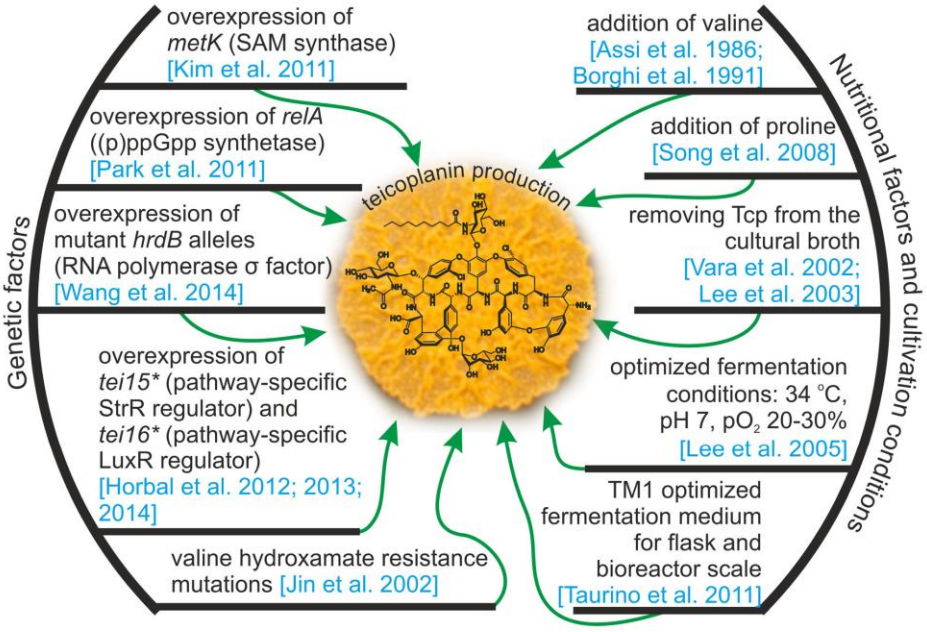


Fig. 3 Summary of the main genetic, nutritional factors and cultivation conditions that positively influence Tcp production in *A. teichomyceticus*. Products of the genes, mentioned on the figure, are given in brackets

transcriptional units (Horbal et al. 2014; Yushchuk et al. 2019). Genes *tei15** and *tei16** were overexpressed using different vectors, which led to the creation of strains significantly overproducing Tcp in the industrial TM1 medium, which was previously developed by Taurino et al. 2011 (see below). Integration of the knowledge of global and pathway-specific regulatory circuits looks like a promising way to further increase Tcp titers in industrial producers.

As in the case of many other antibiotics, Tcp production is heavily influenced by the cultivation conditions. Optimization of the fermentation conditions is another way to improve production of Tcp which is complementary to strain improvement (Marcone et al. 2018). The first breakthrough in Tcp production was the discovery that valine addition to the industrial medium enhanced dramatically the production of the main component carrying a branched ten-carbon saturated acid chain (Fig. 3; Assi et al. 1986). Following this early evidence and other pioneering work (Borghi et al. 1991), it became clear that each of the five major components of the teicoplanin complex could be modulated by adding oils containing linear fatty acids and/or branched amino acids, which are respectively the starters for the linear and branched fatty acids incorporated in Tcp (Taurino et al. 2011). If valine addition might double Tcp production and improve its complex composition in TM1 (Taurino et al. 2011), other authors reported that proline supplement increased Tcp production at fermenter scale by about fivefold (Song et al. 2008), and that in a glycerol-based rich medium an industrial mutant strain produced threefold more than in the non-glycerol medium (Fig. 3) (Park et al. 2009). Moreover, by means of conventional UV mutagenesis, (valine hydroxamate)-resistant *A. teichomyceticus* strains were isolated, which produced more Tcp requiring no addition of valine (Jin et al. 2002). However, the genetic background of these strains remains obscure.

The fermentation process to produce Tcp was optimized at fermenter scale by Jung et al. (2008), who indicated that the optimized parameters are 34 °C, pH 7.0 and dissolved oxygen at 20–30% of saturation, and by Lee et al. (2005) who tested different rich media during scale up (Fig. 3). Unfortunately, in these studies the strains tested were mutants obtained via different mutagenesis approaches. The genealogy of these mutants could not be clearly reconstructed and they possess different initial levels of Tcp production. Consequently, productivity results are not often comparable and easily transferable from one mutant to another one. For this reason, rational engineering of the wild-type strain (see above the work by Horbal et al. 2014) could be particularly useful since the different cultivation conditions can be compared using high producers, where single and controlled genetic modifications have been introduced into a clean background.

Another interesting feature of Tcp production is that the antibiotic inhibits the growth of the producing strain at low concentrations (around 25 mg/l) as reported first by Heydorn et al. (1999) and then by Beltrametti et al. (2007). The intriguing factor is that Tcp is an antibiotic active against growing Gram-positive bacteria and it is produced by a Gram-positive actinomycete, although it was proved that its action is not bactericidal to *A. teichomyceticus* even at 5000 mg/l (Beltrametti et al. 2007). Thus, the producing strain should adopt a self-resistance mechanism to avoid suicide during the early phases of GPA production; otherwise, cells stop growing (see next section). Removing the produced teicoplanin during the exponential growth phase by continuous culture (Vara et al. 2002) or by adding Tcp-entrapping resin during fermentation (Fig. 3) (Lee et al. 2003) facilitated antibiotic production and the transition from the exponential growth phase to the stationary one, where resistance is no longer required and the non-proliferating cells are tolerant to high concentrations of their own product.

Teicoplanin self-resistance and export in *A. teichomyceticus*

The self-resistance phenotype of *A. teichomyceticus* is provided by the expression of *vanHAX*-like genes *(tei7-6-5)*. These produce a peptidoglycan precursor terminating in D-Ala-D-Lac, which has a very reduced affinity for GPAs (Beltrametti et al. 2007). Thus far, these are the only known teicoplanin resistance determinants in *A. teichomyceticus*: the strain lacks genes coding for the D,D-carboxypeptidase VanY homologues, unlike *Am. balhimycina,* which possess both *vanHAX* and *vanY* genes (Kilian et al. 2016) and *N. gerenzanensis*, where the only known mechanism of resistance relies on the action of a VanY (Binda et al. 2012). The three genes *tei7-6-5* are transcribed as a unique operon and, differently from what was initially believed, their expression is not co-regulated with antibiotic production (Yushchuk et al. 2019; Yushchuk et al. 2020). Expression of *tei7-6-5* is constitutive in the TM1 medium and depends neither on growth time point nor on the presence/absence of teicoplanin in the medium (Beltrametti et al. 2007; Yushchuk et al. 2019). The results of the transcriptional analysis are consistent with studies of D,D-dipeptidase VanX activity in *A. teichomyceticus*, which is high and relatively stable through all periods of growth, thus protecting growing cells from bacteriostatic action of the antibiotic (Binda et al. 2018). Interestingly, the heterologous expression of *tei7-6-5* in *N. gerenzanensis* determines the production of the resistant peptidoglycan and increases the level of resistance to the A40926 (Marcone et al. 2014).

The *tei* cluster also contains orthologues of *vanRS* (*tei2–3*) coding for a response regulator and sensor histidine kinase respectively. Very recently, it was proved that Tei2 is essential for the *tei7p*-mediated gene expression of *tei7-6-5* operon (Yushchuk et al. 2020). *tei2* and *tei3* genes are co-expressed, forming a three gene operon together with *tei4*—coding for a

dihydrofolate reductase with no obvious function in the teicoplanin self-resistance (Yushchuk et al. 2020). Expression of *tei2-3-4* operon is also constitutive in *A. teichomyceticus* growing in TM1, resembling the expression pattern for *tei7-6-5* (Yushchuk et al. 2019). Additionally, *tei2p* in *A. teichomyceticus* is stronger in driving gene expression than *aac(3)IVp*—previously selected as the best heterologous promoter in *A. teichomyceticus* (Horbal et al. 2013)—and *tei2p*-driven expression occurs since the early spore-germination stage and does not depend on teicoplanin addition to the medium (Yushchuk et al. 2020). Overall, it could be assumed that the strong *tei2* promoter determines the constitutive expression of *tei2-3-4* operon, which consequently leads to the constitutive expression of *tei7-6-5*, granting the high levels of teicoplanin resistance to *A. teichomyceticus* (Yushchuk et al. 2020). One intriguing feature that remains to be further investigated is the putative non-inducibility of Tei3: certain mutations in its domain that is responsible for phosphorylation of Tei2 may be providing the kinase with a constant phosphorylase activity (Beltrametti et al. 2007). However, these potentially unique properties require further experimental verification.

The teicoplanin export system could be also contributing to self-resistance. A conceptual scheme of Tcp export can be proposed from studies on export of balhimycin in *Am. balhimycina*. In this strain, ABC-transporter Tba was shown to be involved in antibiotic export (Menges et al. 2007). The *tei* cluster possesses one gene coding for an ABC-transporter—*tei4**. Tei4* shares 78.9% of amino acid sequence identity with Tba, making it a candidate for Tcp export (Fig. 2), although its role has not been investigated yet.

Comparison of the *tei* cluster to other glycopeptide BGCs

Among all *Actinoplanes*-derived GPAs, BGCs are currently sequenced only for Tcp and UK-68,597 (the *auk* BGC, Yim et al. 2014). Considering the abundance of glycopeptide BGCs in the genus *Amycolatopsis* (Adamek et al. 2018), one can expect discoveries of new *Actinoplanes*-produced glycopeptides in the near future. Indeed, a recent study of GPA phylogeny describes new putative GPA BGCs from *Actinoplanes* (Waglechner et al. 2019). It is very likely that further study of the chemical diversity of GPAs from *Actinoplanes* spp. will yield new glycopeptide tailoring enzymes and thus power the generation of novel Tcp-like GPAs.

Albeit coming from species of one genus, the *tei* and *auk* BGCs are not fully syntenic and have undergone substantial genetic reorganization (Fig. 4), which is in contrast to the situation in abovementioned *Amycolatopsis* GPA BGCs (Adamek et al. 2018). Nevertheless, the similarity of the two BGCs is significant, which is not surprising regarding the structural similarity between Tcp and UK-68,597 (Fig. 1): both clusters possess 31 genes, which are close homologues of each other, differing mainly in genes for the aglycone modification enzymes (Fig. 4).

Other GPAs, especially A40926 (Fig. 1), also possess a high degree of structural similarity with Tcp. This obviously provokes the question about how structurally related scaffolds are related on the level of their respective BGCs. The first attempt to reconstruct the evolutionary relationship between GPA BGCs (including *tei*), was focused on building, protein by protein, the phylogeny of all key biosynthetic enzymes encoded within BGCs (Donadio et al. 2005). This was further integrated with comparisons of gene synteny and presence/absence of the putative regulatory sequences in promoters. Although the resolution of such analysis was quite low at that time (only 6 BGCs were available), *tei* was recognized as more related to the A47934 BGC from *S. toyocaensis* (Pootoolal et al. 2002) than to A40926 BGC from *N. gerenzanensis* (the *dbv* BGC, Sosio et al. 2003). In one of our recent works (Yushchuk et al. 2017), we extended such an approach to build multilocus phylogeny of 16 GPA BGCs to show that *tei* is in fact closely related with the *auk*, CA37, CA915 (Banik et al. 2010) and esnapd15 BGCs (Owen et al. 2013)—these latter three BGCs come from metagenomic samples. Moreover, *tei* was found to share (Yushchuk et al. 2017) a considerable number of closely related genes with the feglymycin BGC (the *feg* BGC, Gonsior et al. 2015). In contrast, *dbv* appeared to be more closely related to *Amycolatopsis*-derived clusters. Notably, in the paper describing the CA37 and CA915 BGCs (Banik et al. 2010), the authors speculated that CA37 and CA915 BGCs should encode the biosynthesis of actaplanin-like molecules, based on the tailoring enzymes found in these clusters. Since actaplanin is an *Actinoplanes*-derived antibiotic, our reconstruction was in line with this observation.

Similar phylogenetic relationships were also observed between StrR-like regulators (Yushchuk et al. 2019), where Tei15* formed a well-supported clade together with its orthologues encoded in the CA915, CA37, esnapd15, *auk*, GQ475282 (Kim et al. 2010) and *feg* BGCs. Interestingly, CA915, CA37, esnapd15, *auk*, GQ475282 and *feg* BGCs also encode close homologues of LuxR-family regulator Tei16* (Yushchuk et al. 2019). The *dbv* BGC encodes a LuxR-like regulator as well, but this seems unrelated to Tei16*-like regulators (Yushchuk et al. 2019).

Finally, a recently published study of GPA BGCs evolution (Waglechner et al. 2019) involved 71 GPA BGCs, including "old" BGCs, BGCs found in public databases, as well as in de novo sequenced actinobacterial genomes. This research covered BGCs for the biosynthesis of GPAs *sensu stricto* (I–V types (Nicolaou et al. 1999) with heptapeptide scaffolds), as well as other peptide antibiotics like feglymycin and ramoplanin (the *ramo* BGC, Farnet et al. 2002). Additionally, novel BGCs were also analyzed, including some

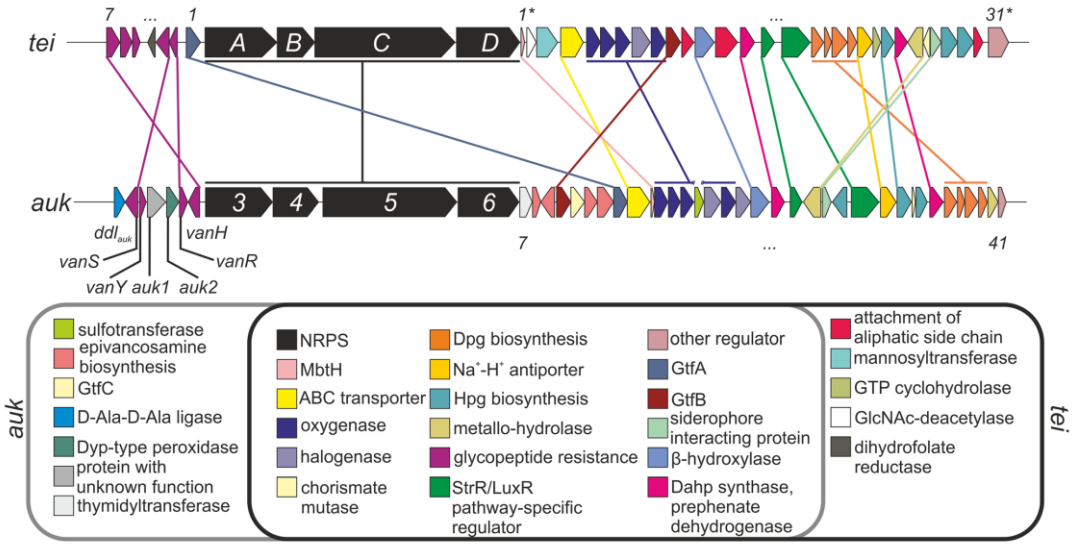


Fig. 4 Comparison of the genetic organization of the *tei* and *auk* BGCs, which are the only two BGCs identified in members of the genus *Actinoplanes* that produce known GPAs. It can be observed that, while there are obvious similarities, both clusters have undergone significant genetic reorganization, as highlighted by lines connecting orthologous genes between the clusters

predicted exotic scaffolds, such as hypothetic Gp1529, which is predicted to have a cross-linked nonapeptide aglycone. Although having such unusual aglyca, these *sensu lato* GPAs were clearly shown to be genetically related to *sensu stricto* GPAs. The study has used a novel approach to illuminate the natural history of this family of BGCs, correlating the timed phylogeny of the natural hosts with the phylogeny of all protein families encoded within GPA BGCs. This provides an insight into the main evolutionary events in the genesis of GPA BGCs, but also approximates the real age of these evolutionary events. Our analysis of all phylogenetic trees generated for the GPA BGC-encoded protein families (Waglechner et al. 2019, http://github.com/waglecn/GPA_evolution) reveals that clades containing *tei*-derived proteins, also feature homologues from *auk*, CA37, CA915, esnapd15, and esnapd16 (Owen et al. 2013) BGCs in the majority of cases. In a few cases, proteins also clade with the *feg* and *ramo* homologues, as well as some other putative BGCs. Notably, *dbv*-derived proteins are distantly related to the *tei* ones.

Thus, according to the data discussed above, the *tei* cluster seems to be most closely related to *auk*, CA915, CA37, esnapd15, and esnapd16 BGCs. Interestingly, *feg* also shares many elements in common with *tei*. Since *feg* encodes the biosynthesis of feglymycin, a non-cross-linked, non-glycosylated linear peptide (Gonsior et al. 2015) that is structurally distinct to glycopeptide scaffolds, the similarity of the *feg* and *tei* BGCs is thus quite remarkable. The *ramo* BGC, responsible for the biosynthesis of ramoplanin in *Actinoplanes ramoplaninifer* ATCC 33076 (Marcone et al. 2017), also shares a few common features with *tei* (Waglechner et al. 2019). It is also important to note, that, according to the discussed phylogenetic reconstructions (Donadio et al. 2005; Yushchuk et al. 2017; Yushchuk et al. 2019; Waglechner et al. 2019), the enzymes and regulators in the *dbv* BGC for A40926 are relatively distantly related to their *tei* homologs. This strongly suggests that the teicoplanin and A40926 BGCs are not closely related and are a result of convergent evolution within the GPA family, given the close structural similarity of the resulting GPAs (Fig. 1). Similarly, *tei* and ristocetin BGCs (Truman et al. 2014; Spohn et al. 2014) have converged on almost the same core aglycone despite even bigger differences in the phylogeny of biosynthetic enzymes.

Outlook

The biochemical logic behind the formation of Tcp is now well understood, and we are in good position to make use of this knowledge. Of the different possible areas for future development, we think that further studies around Tcp acylation and glycosylation would be most beneficial. Concerning acylation, it would be beneficial to unify the Tcp complex into a single compound for therapeutic use, although this would require engineering the substrate specificity of the Tei11* acyltransferase.

Further improvement of teicoplanin production in *A. teichomyceticus* also remains a task of high importance. Current attempts to improve Tcp titers were focused on defining optimal medium and fermentation conditions, classical

mutagenesis and, more recently, on cluster-situated regulatory genes. A new major direction in this regard would be to take a broader, genome-based look on the regulation of Tcp biosynthesis. Particularly, higher-order regulators of Tcp production and primary metabolic pathways feeding into Tcp pathway are yet to be elucidated. Thus far, our knowledge of global regulatory pathways of antibiotic biosynthesis is limited not only for *A. teichomyceticus,* but also for the vast majority of other non-*Streptomyces* actinomycetes. Therefore *A. teichomyceticus*, as a model test organism, could be a new key for opening the variety of natural products from uncommon actinobacteria. Moreover, it is of special interest to identify genes outside of *tei* cluster contributing to Tcp resistance of the producer strain. This will enable knowledge-based approaches to strain improvement for Tcp production and may prove useful in the future surveillance of emerging mechanisms of resistance in pathogens to clinically used glycopeptide antibiotics.

Finally, it is imperative to continue phylogenetic and experimental reconstructions of evolution of glycopeptide BGCs to more fully embrace chemical and biological spaces that can be accessed from these fascinating biosynthetic pathways.

Acknowledgments We want to thank all present and former members of groups studying teicoplanin in Lviv University, University of Insubria, and the John Innes Centre.

Funding information This work was supported by grant of Ministry of Education and Science of Ukraine Bg-46F (to V.F.).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

Adamek M, Alanjary M, Sales-Ortells H, Goodfellow M, Bull AT, Winkler A, Wibberg D, Kalinowski J, Ziemert N (2018) Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. BMC Genomics 19(1):426. https://doi.org/10.1186/s12864-018-4809-4

Armenia I, Marcone GL, Berini F, Orlandi VT, Pirrone C, Martegani E, Gornati R, Bernardini G, Marinelli F (2018) Magnetic nanoconjugated teicoplanin: a novel tool for bacterial infection site targeting. Front Microbiol 9:2270. https://doi.org/10.3389/fmicb.2018.02270

Assi F, Lancini G, Gianantonio A (1986) Method for selectively increasing the ratio of single major components of teicoplanin A2 complex. European Patent 0204179A1

Banik JJ, Craig JW, Calle PY, Brady SF (2010) Tailoring enzyme-rich environmental DNA clones: a source of enzymes for generating libraries of unnatural natural products. J Am Chem Soc 132(44): 15661–15670. https://doi.org/10.1021/ja105825a

Bardone MR, Paternoster M, Coronelli C (1978) Teichomycins, new antibiotics from *Actinoplanes teichomyceticus* nov. sp. II. Extraction and chemical characterization. J Antibiot (Tokyo) 31(3):170–177. https://doi.org/10.7164/antibiotics.31.170

Beltrametti F, Consolandi A, Carrano L, Bagatini F, Rossi R, Leoni L, Zennaro E, Selva E, Marinelli F (2007) Resistance to glycopeptide antibiotics in the teicoplanin producer is mediated by *van* gene homologue expression directing the synthesis of a modified cell wall peptidoglycan. Antimicrob Agents Chemother 51(4):1135–1141. https://doi.org/10.1128/AAC.01071-06

Binda E, Marcone GL, Pollegioni L, Marinelli F (2012) Characterization of VanYn, a novel D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727. FEBS J 279(17):3203–3213. https://doi.org/10.1111/j.1742-4658.2012.08706.x

Binda E, Marinelli F, Marcone GL (2014) Old and new glycopeptide antibiotics: action and resistance. Antibiotics (Basel) 3(4):572–594. https://doi.org/10.3390/antibiotics3040572

Binda E, Cappelletti P, Marinelli F, Marcone GL (2018) Specificity of induction of glycopeptide antibiotic resistance in the producing actinomycetes. Antibiotics (Basel) 7(2). pii: E36. https://doi.org/10.3390/antibiotics7020036

Borghi A, Edwards D, Zerilli LF, Lancini GC (1991) Factors affecting the normal and branched-chain acyl moieties of teicoplanin components produced by *Actinoplanes teichomyceticus*. J Gen Microbiol 137: 587–592. https://doi.org/10.1099/00221287-137-3-587

Chen H, Thomas MG, Hubbard BK, Losey HC, Walsh CT, Burkart MD (2000) Deoxysugars in glycopeptide antibiotics: enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis. PNAS 97(22):11942–11947. https://doi.org/10.1073/pnas.210395097

Chen H, Tseng CC, Hubbard BK, Walsh CT (2001) Glycopeptide antibiotic biosynthesis: enzymatic assembly of the dedicated amino acid monomer (S)-3,5-dihydroxyphenylglycine. PNAS 98(26):14901–14906. https://doi.org/10.1073/pnas.221582098

Choroba OW, Williams DH, Spencer JB (2000) Biosynthesis of the vancomycin group of antibiotics: involvement of an unusual dioxygenase in the pathway to (S)-4-hydroxyphenylglycine. JACS 122(22):5389–5390. https://doi.org/10.1021/ja000076v

Debono M, Merkel KE, Molloy MR, Barnhart M, Presti E, Hunt AH, Hamill RL (1984) Actaplanin, new glycopeptide antibiotics produced by *Actinoplanes missouriensis*. The isolation and preliminary chemical characterization of actaplanin. J Antibiot 37(2):85–95. https://doi.org/10.7164/antibiotics.37.85

Donadio S, Sosio M, Stegmann E, Weber T, Wohlleben W (2005) Comparative analysis and insights into the evolution of gene clusters for glycopeptide antibiotic biosynthesis. Mol Gen Genomics 274(1): 40–50. https://doi.org/10.1007/s00438-005-1156-3

Farnet CM, Staffa A, Zazopoulos E (2002) Gene cluster for ramoplanin biosynthesis. European Patent 1326983B1

Gonsior M, Mühlenweg A, Tietzmann M, Rausch S, Poch A, Süssmuth RD (2015) Biosynthesis of the peptide antibiotic feglymycin by a linear nonribosomal peptide synthetase mechanism. Chembiochem 16(18):2610–2614. https://doi.org/10.1002/cbic.201500432

Hadatsch B, Butz D, Schmiederer T, Steudle J, Wohlleben W, Süssmuth R, Stegmann E (2007) The biosynthesis of teicoplanin-type glycopeptide antibiotics: assignment of P450 mono-oxygenases to side chain cyclizations of glycopeptide A47934. Chem Biol 14(9):1078–1089. https://doi.org/10.1016/j.chembiol.2007.08.014

Haslinger K, Peschke M, Brieke C, Maximowitsch E, Cryle MJ (2015) X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. Nature 521(7550):105–109. https://doi.org/10.1038/nature14141

Heydorn A, Suhr-Jessen T, Nielsen J (1999) Growth and production kinetics of a teicoplanin producing strain of *Actinoplanes*

teichomyceticus. J Antibiot (Tokyo) 52:40–44. https://doi.org/10.7164/antibiotics.52.40

Horbal L, Zaburannyy N, Ostash B, Shulga S, Fedorenko V (2012) Manipulating the regulatory genes for teicoplanin production in *Actinoplanes teichomyceticus*. World J Microbiol Biotechnol 28(5):2095–2100. https://doi.org/10.1007/s11274-012-1013-6

Horbal L, Kobylyanskyy A, Yushchuk O, Zaburannyi N, Luzhetskyy A, Ostash B, Marinelli F, Fedorenko V (2013) Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus* – producer of teicoplanin, drug of last defense. J Biotechnol 168(4):367–372. https://doi.org/10.1016/j.jbiotec.2013.10.018

Horbal L, Kobylyanskyy A, Truman AW, Zaburranyi N, Ostash B, Luzhetskyy A, Marinelli F, Fedorenko V (2014) The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. Appl Microbiol Biotechnol 98(22):9295–9309. https://doi.org/10.1007/s00253-014-5969-z

Hubbard BK, Thomas MG, Walsh CT (2000) Biosynthesis of L-*p*-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. Chem Biol 7(12):931–942. https://doi.org/10.1016/s1074-5521(00)00043-0

James RC, Pierce JG, Okano A, Xie J, Boger DL (2012) Redesign of glycopeptide antibiotics: back to the future. ACS Chem Biol 7:797–804. https://doi.org/10.1021/cb300007j

Jin ZH, Wang MR, Cen PL (2002) Production of teicoplanin by valine analogue-resistant mutant strains of *Actinoplanes teichomyceticus*. Appl Microbiol Biotechnol 58(1):63–66. https://doi.org/10.1007/s00253-001-0872-9

Jones SE, Elliot MA (2018) 'Exploring' the regulation of *Streptomyces* growth and development. Curr Opin Microbiol 42:25–30. https://doi.org/10.1016/j.mib.2017.09.009

Jung HM, Kim SY, Prabhu P, Moon HJ, Kim IW, Lee JK (2008) Optimization of culture conditions and scale-up to plant scales for teicoplanin production by *Actinoplanes teichomyceticus*. Appl Microbiol Biotechnol 80(1):21–27. https://doi.org/10.1007/s00253-008-1530-2

Jung HM, Jeya M, Kim SY, Moon HJ, Kumar Singh R, Zhang YW, Lee JK (2009) Biosynthesis, biotechnological production, and application of teicoplanin: current state and perspectives. Appl Microbiol Biotechnol 84(3):417–428. https://doi.org/10.1007/s00253-009-2107-4

Kaniusaite M, Tailhades J, Marschall EA, Goode RJA, Schittenhelm RB, Cryle MJ (2019) A proof-reading mechanism for non-proteinogenic amino acid incorporation into glycopeptide antibiotics. Chem Sci 10:9466–9482. https://doi.org/10.1039/c9sc03678d

Kilian R, Frasch HJ, Kulik A, Wohlleben W, Stegmann E (2016) The VanRS homologous two-component system VnlRSAb of the glycopeptide producer *Amycolatopsis balhimycina* activates transcription of the $vanHAX_{Sc}$ genes in *Streptomyces coelicolor*, but not in *A. balhimycina*. Microb Drug Resist 22(6):499–509. https://doi.org/10.1089/mdr.2016.0128

Kim DJ, Huh JH, Yang YY, Kang CM, Lee IH, Hyun CG, Hong SK, Suh JW (2003) Accumulation of S-adenosyl-L-methionine enhances production of actinorhodin but inhibits sporulation in *Streptomyces lividans* TK23. J Bacteriol 185(2):592–600. https://doi.org/10.1128/jb.185.2.592-600.2003

Kim JH, Feng Z, Bauer JD, Kallifidas D, Calle PY, Brady SF (2010) Cloning large natural product gene clusters from the environment: piecing environmental DNA gene clusters back together with TAR. Biopolymers 93(9):833–844. https://doi.org/10.1002/bip.21450

Kim DY, Hwang YI, Choi SU (2011) Cloning of *metK* from *Actinoplanes teichomyceticus* ATCC31121 and effect of its high expression on antibiotic production. J Microbiol Biotechnol 21(12):1294–1298. https://doi.org/10.4014/jmb.1101.01018

Kittilä T, Kittel C, Tailhades J, Butz D, Schoppet M, Büttner A, Goode RJA, Schittenhelm RB, van Pee KH, Süssmuth RD, Wohlleben W, Cryle MJ, Stegmann E (2017) Halogenation of glycopeptide antibiotics occurs at the amino acid level during non-ribosomal peptide synthesis. Chem Sci 8(9):5992–6004. https://doi.org/10.1039/c7sc00460e

Kruger RG, Lu W, Oberthür M, Tao J, Kahne D, Walsh CT (2005) Tailoring of glycopeptide scaffolds by the acyltransferases from the teicoplanin and A-40,926 biosynthetic operons. Chem Biol 12(1):131–140. https://doi.org/10.1016/j.chembiol.2004.12.005

Lazzarini A, Borghi A, Zerilli LF, Ferrari P, Colombo L, Lancini GC (1997) Novel teicoplanins by directed biosynthesis. J Antibiot (Tokyo) 50(2):180–183. https://doi.org/10.7164/antibiotics.50.180

Lee JC, Park HR, Park DJ, Lee HB, Kim YB, Kim CJ (2003) Improved production of teicoplanin using adsorbent resin in fermentations. Lett Appl Microbiol 37(3):196–200. https://doi.org/10.1046/j.1472-765x.2003.01374.x

Lee JC, Min JW, Park DJ, Kwang HS, Yoon KH, Park HR, Park YS, Kwon MG, Lee JM, Kim CJ (2005) Large-scale fermentation for the production of teicoplanin from a mutant of *Actinoplanes teichomyceticus*. J Microbiol Biotechnol 15(4):787–791

Li TL, Huang F, Haydock SF, Mironenko T, Leadlay PF, Spencer JB (2004) Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase. Chem Biol 11(1):107–119. https://doi.org/10.1016/j.chembiol.2004.01.001

Lyu SY, Liu YC, Chang CY, Huang CJ, Chiu YH, Huang CM, Hsu NS, Lin KH, Wu CJ, Tsai MD, Li TL (2014) Multiple complexes of long aliphatic N-acyltransferases lead to synthesis of 2,6-diacylated/2-acyl-substituted glycopeptide antibiotics, effectively killing vancomycin-resistant enterococcus. JACS 136(31):10989–10995. https://doi.org/10.1021/ja504125v

Makitrynskyy R, Ostash B, Tsypik O, Rebets Y, Doud E, Meredith T, Luzhetskyy A, Bechthold A, Walker S, Fedorenko V (2013) Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. Open Biol 3(10):130121. https://doi.org/10.1098/rsob.130121

Marcone GL, Binda E, Carrano L, Bibb M, Marinelli F (2014) Relationship between glycopeptide production and resistance in the actinomycete *Nonomuraea* sp. ATCC 39727. Antimicrob Agents Chemother 58(9):5191–5201. https://doi.org/10.1128/AAC.02626-14

Marcone GL, Binda E, Reguzzoni M, Gastaldo L, Dalmastri C, Marinelli F (2017) Classification of *Actinoplanes sp.* ATCC 33076, an actinomycete that produces the glycolipodepsipeptide antibiotic ramoplanin, as *Actinoplanes ramoplaninifer* sp. nov. Int J Syst Evol Microbiol 67(10):4181–4188. https://doi.org/10.1099/ijsem.0.002281

Marcone GL, Binda E, Berini F, Marinelli F (2018) Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era. Biotechnol Adv 36(2):534–554. https://doi.org/10.1016/j.biotechadv.2018.02.009

McCormick JR, Flärdh K (2012) Signals and regulators that govern *Streptomyces* development. FEMS Microbiol Rev 36(1):206–231. https://doi.org/10.1111/j.1574-6976.2011.00317.x

Menges R, Muth G, Wohlleben W, Stegmann E (2007) The ABC transporter Tba of *Amycolatopsis balhimycina* is required for efficient export of the glycopeptide antibiotic balhimycin. Appl Microbiol Biotechnol 77(1):125–134. https://doi.org/10.1007/s00253-007-1139-x

Nicolaou KC, Boddy CNC, Bräse S, Winssinger N (1999) Chemistry, biology, and medicine of the glycopeptide antibiotics. Angew Chem Int Ed Engl 38(15):2096–2152. https://doi.org/10.1002/(sici)1521-3773(19990802)38:15<2096::aid-anie2096>3.0.co;2-f

Okamoto S, Lezhava A, Hosaka T, Okamoto-Hosoya Y, Ochi K (2003) Enhanced expression of S-adenosylmethionine synthetase causes overproduction of actinorhodin in *Streptomyces coelicolor* A3(2).

J Bacteriol 185(2):601–609. https://doi.org/10.1128/jb.185.2.601-609.2003

Ostash B (2018) Regulatory genes of AdpA subfamily in *Streptomyces*: function and evolution. https://doi.org/10.2478/9783110627770

Ostash B, Yushchuk O, Tistechok S, Mutenko H, Horbal L, Muryn A, Dacyuk Y, Kalinowski J, Luzhetskyy A, Fedorenko V (2015) The *adpA*-like regulatory gene from *Actinoplanes teichomyceticus*: *in silico* analysis and heterologous expression. World J Microbiol Biotechnol 31(8):1297–1301. https://doi.org/10.1007/s11274-015-1882-6

Owen JG, Reddy BV, Ternei MA, Charlop-Powers Z, Calle PY, Kim JH, Brady SF (2013) Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. PNAS 110(29):11797–11802. https://doi.org/10.1073/pnas.1222159110

Pang AP, Du L, Lin CY, Qiao J, Zhao GR (2015) Co-overexpression of *lmbW* and *metK* led to increased lincomycin A production and decreased byproduct lincomycin B content in an industrial strain of *Streptomyces lincolnensis*. J Appl Microbiol 119(4):1064–1074. https://doi.org/10.1111/jam.12919

Parenti F (1986) Structure and mechanism of action of teicoplanin. J Hosp Infect Suppl A:79-83. https://doi.org/10.1016/0195-6701(86)90011-3

Parenti F, Beretta G, Berti M, Arioli V (1978) Teichomycins, new antibiotics from *Actinoplanes teichomyceticus* Nov. Sp. I. Description of the producer strain, fermentation studies and biological properties. J Antibiot (Tokyo) 31(4):276–283. https://doi.org/10.7164/antibiotics.31.276

Park HR, Lee JC, Hwang JH, Park DJ, Kim CJ (2009) Glycerol affects the acyl moieties of teicoplanin components produced by *Actinoplanes teichomyceticus* MSI2210. Microbiol Res 164(5):588–592. https://doi.org/10.1016/j.micres.2007.05.007

Park SY, Kim DY, Hwang YI, Choi SU (2011) Cloning of *relA* from *Actinoplanes teichomyceticus* ATCC31121 and its application for increases in antibiotic production. J Korean Soc Appl Biol Chem 54(2):281–286. https://doi.org/10.3839/jksabc.2011.044

Peschke M, Gonsior M, Süssmuth RD, Cryle MJ (2016a) Understanding the crucial interactions between cytochrome P450s and non-ribosomal peptide synthetases during glycopeptide antibiotic biosynthesis. Curr Opin Struct Biol 41:46–53. https://doi.org/10.1016/j.sbi.2016.05.018

Peschke M, Brieke C, Cryle MJ (2016b) F-O-G ring formation in glycopeptide antibiotic biosynthesis is catalysed by OxyE. Sci Rep 6:35584. https://doi.org/10.1038/srep35584

Peschke M, Haslinger K, Brieke C, Reinstein J, Cryle MJ (2016c) Regulation of the P450 oxygenation cascade involved in glycopeptide antibiotic biosynthesis. JACS 138(21):6746–6753. https://doi.org/10.1021/jacs.6b00307

Peschke M, Brieke C, Goode RJ, Schittenhelm RB, Cryle MJ (2017) Chlorinated glycopeptide antibiotic peptide precursors improve cytochrome P450-catalyzed cyclization cascade efficiency. Biochemistry 56(9):1239–1247. https://doi.org/10.1021/acs.biochem.6b01102

Pfeifer V, Nicholson GJ, Ries J, Recktenwald J, Schefer AB, Shawky RM, Schröder J, Wohlleben W, Pelzer S (2001) A polyketide synthase in glycopeptide biosynthesis: the biosynthesis of the non-proteinogenic amino acid (S)-3,5-dihydroxyphenylglycine. J Biol Chem 276(42):38370–38377. https://doi.org/10.1074/jbc.M106580200

Pootoolal J, Thomas MG, Marshall CG, Neu JM, Hubbard BK, Walsh CT, Wright GD (2002) Assembling the glycopeptide antibiotic scaffold: the biosynthesis of A47934 from *Streptomyces toyocaensis* NRRL15009. PNAS 99(13):8962–8967. https://doi.org/10.1073/pnas.102285099

Roberts JA, Lipman J, Starr T, Wallis SC, Paul S, Ribas AM, De Waele JJ, DeCrop L, Spapen H, Wauters J, Dugernier T, Jorens P, Dapper I, De Backer D, Taccone FS, Rello J, Ruano L, Afonso E, Alvarez-Lerma F, Gracia-Arnillas MP, Fernández F, Feijoo N, Bardolet N, Rovira A, Garro P, Colon D, Castillo C, Fernado J, Lopez MJ, Fernandez JL, Arribas AM, Teja JL, Ots E, Montejo JC, Catalan M, Prieto I, Gonzalo G, Galvan B, Blasco MA, Meyer E, Nogal FD, Vidaur L, Sebastian R, Garde PM, Velasco Mdel M, Crespo RZ, Esperatti M, Torres A, Montravers P, Baldesi O, Dupont H, Mahjoub Y, Lasocki S, Constantin JM, Payen JF, Martin C, Albanese J, Malledant Y, Pottecher J, Lefrant JY, Jaber S, Joannes-Boyau O, Orban C, Ostermann M, McKenzie C, Berry W, Smith J, Lei K, Rubulotta F, Gordon A, Brett S, Stotz M, Templeton M, Rhodes A, Ebm C, Moran C, Kaukonen KM, Pettilä V, Dimopoulos G, Koulenti D, Xristodoulou A, Theodorou V, Kouliatsis G, Sertaridou E, Anthopoulos G, Choutas G, Rantis T, Karatzas S, Balla M, Papanikolaou M, Myrianthefs P, Gavala A, Fildisis G, Koutsoukou A, Kyriakopoulou M, Petrochilou K, Kompoti M, Michalia M, Clouva-Molyvdas FM, Gkiokas G, Nikolakopoulos F, Psychogiou V, Malliotakis P, Akoumianaki E, Lilitsis E, Koulouras V, Nakos G, Kalogirou M, Komnos A, Zafeiridis T, Chaintoutis C, Arvaniti K, Matamis D, Chaintoutis C, Kydona C, Gritsi-Gerogianni N, Giasnetsova T, Giannakou M, Soultati I, Chytas I, Antoniadou E, Antipa E, Lathyris D, Koukoubani T, Paraforou T, Spiropoulou K, Bekos V, Spring A, Kalatzi T, Nikolaou H, Laskou M, Strouvalis I, Aloizos S, Kapogiannis S, Soldatou O, Bassetti M, Adembri C, Villa G, Giarratano A, Raineri SM, Cortegiani A, Montalto F, Strano MT, Ranieri VM, Sandroni C, De Pascale G, Molin A, Pelosi P, Montagnani L, Urbino R, Mastromauro I, DeRosa FG, Ranieri VM, Cardoso T, Afonso S, Gonçalves-Pereira J, Baptista JP, Akova M, Özveren A (2014) Variability in protein binding of teicoplanin and achievement of therapeutic drug monitoring targets in critically ill patients: lessons from the DALI Study. Int J Antimicrob Agents 43(5):423–430. https://doi.org/10.1016/j.ijantimicag.2014.01.023

Sandercock AM, Charles EH, Scaife W, Kirkpatrick PN, O'Brien SW, Papageorgiou EA, Spencer JB, Williams DH (2001) Biosynthesis of the di-meta-hydroxyphenylglycine constituent of the vancomycin-group antibiotic chloroeremomycin. Chem Commun 14:1252–1253. https://doi.org/10.1039/b103668h

Schoppet M, Peschke M, Kirchberg A, Wiebach V, Süssmuth RD, Stegmann E, Cryle MJ (2018) The biosynthetic implications of late-stage condensation domain selectivity during glycopeptide antibiotic biosynthesis. Chem Sci 10(1):118–133. https://doi.org/10.1039/c8sc03530j

Skelton NJ, Williams DH, Monday RA, Ruddock JC (1990) Structure elucidation of the novel glycopeptide antibiotic UK-68,597. J Org Chem 55:3718–3723. https://doi.org/10.1021/jo00299a008

Snyder NJ, Cooper RD, Briggs BS, Zmijewski M, Mullen DL, Kaiser RE, Nicas TI (1998) Enzymatic deacylation of teicoplanin followed by reductive alkylation: synthesis and antibacterial activity of new glycopeptides. J Antibiot (Tokyo) 51(10):945–951. https://doi.org/10.7164/antibiotics.51.945

Song JM, Park JT, Lee HS, Kang JH, Kang DJ (2008) Production of teicoplanin from *Actinoplanes teichomyceticus* ID9303 by adding proline. Biosci Biotechnol Biochem 72(6):1635–1637. https://doi.org/10.1271/bbb.80143

Sosio M, Stinchi S, Beltrametti F, Lazzarini A, Donadio S (2003) The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by nonomuraea species. Chem Biol 10(6):541–549. https://doi.org/10.1016/s1074-5521(03)00120-0

Sosio M, Kloosterman H, Bianchi A, de Vreugd P, Dijkhuizen L, Donadio S (2004) Organization of the teicoplanin gene cluster in *Actinoplanes teichomyceticus*. Microbiology 150(1):95–102. https://doi.org/10.1099/mic.0.26507-0

Spohn M, Kirchner N, Kulik A, Jochim A, Wolf F, Muenzer P, Borst O, Gross H, Wohlleben W, Stegmann E (2014) Overproduction of ristomycin A by activation of a silent gene cluster in

Amycolatopsis japonicum MG417-CF17. Antimicrob Agents Chemother 58(10):6185–6196. https://doi.org/10.1128/AAC.03512-14

Stinchi S, Carrano L, Lazzarini A, Feroggio M, Grigoletto A, Sosio M, Donadio S (2006) A derivative of the glycopeptide A40926 produced by inactivation of the β-hydroxylase gene in *Nonomuraea* sp. ATCC39727. FEMS Microbiol Lett 256(2):229–235. https://doi.org/10.1111/j.1574-6968.2006.00120.x

Taurino C, Frattini L, Marcone GL, Gastaldo L, Marinelli F (2011) *Actinoplanes teichomyceticus* ATCC 31121 as a cell factory for producing teicoplanin. Microb Cell Factories 10:82. https://doi.org/10.1186/1475-2859-10-82

Thayer DA, Wong CH (2006) Vancomycin analogues containing monosaccharides exhibit improved antibiotic activity: a combined one-pot enzymatic glycosylation and chemical diversification strategy. Chem Asian J 1(3):445–452. https://doi.org/10.1002/asia.200600084

Thykaer J, Nielsen J, Wohlleben W, Weber T, Gutknecht M, Lantz AE, Stegmann E (2010) Increased glycopeptide production after overexpression of shikimate pathway genes being part of the balhimycin biosynthetic gene cluster. Metab Eng 12(5):455–461. https://doi.org/10.1016/j.ymben.2010.05.001

Truman AW, Robinson L, Spencer JB (2006) Identification of a deacetylase involved in the maturation of teicoplanin. Chembiochem 7(11):1670–1675. https://doi.org/10.1002/cbic.200600308

Truman AW, Fan Q, Röttgen M, Stegmann E, Leadlay PF, Spencer JB (2008) The role of *cep15* in the biosynthesis of chloroeremomycin: reactivation of an ancestral catalytic function. Chem Biol 15(5):476–484. https://doi.org/10.1016/j.chembiol.2008.03.019

Truman AW, Kwun MJ, Cheng J, Yang SH, Suh JW, Hong HJ (2014) Antibiotic resistance mechanisms inform discovery: identification and characterization of a novel amycolatopsis strain producing ristocetin. Antimicrob Agents Chemother 58(10):5687–5695. https://doi.org/10.1128/AAC.03349-14

Vara AG, Hochkoepple A, Nielsen J, Villadsen J (2002) Production of teicoplanin by *Actinoplanes teichomyceticus* in continuous fermentation. Biotechnol Bioeng 77:589–598. https://doi.org/10.1002/bit.10137

Waglechner N, McArthur AG, Wright GD (2019) Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. Nat Microbiol 4(11):1862–1871. https://doi.org/10.1038/s41564-019-0531-5

Wang H, Yang L, Wu K, Li G (2014) Rational selection and engineering of exogenous principal sigma factor (σ^{HrdB}) to increase teicoplanin production in an industrial strain of *Actinoplanes teichomyceticus*. Microb Cell Factories 13:10. https://doi.org/10.1186/1475-2859-13-10

Yarlagadda V, Samaddar S, Paramanandham K, Shome BR, Haldar J (2015) Membrane disruption and enhanced inhibition of cell-wall biosynthesis: a synergistic approach to tackle vancomycin-resistant bacteria. Angew Chem Int Ed Engl 54(46):13644–13649. https://doi.org/10.1002/anie.201507567

Yim G, Kalan L, Koteva K, Thaker MN, Waglechner N, Tang I, Wright GD (2014) Harnessing the synthetic capabilities of glycopeptide antibiotic tailoring enzymes: characterization of the UK-68,597 biosynthetic cluster. Chembiochem 15(17):2613–2623. https://doi.org/10.1002/cbic.201402179

Yushchuk O, Ostash B, Pham TH, Luzhetskyy A, Fedorenko V, Truman AW, Horbal L (2016) Characterization of the post-assembly line tailoring processes in teicoplanin biosynthesis. ACS Chem Biol 11(8):2254–2264. https://doi.org/10.1021/acschembio.6b00018

Yushchuk O, Ostash B, Horbal L, Fedorenko V (2017) Reconstructing the phylogeny of glycopeptide biosynthetic gene clusters. Factors in experimental evolution of organisms 20:109–115 (in Ukrainian)

Yushchuk O, Ostash I, Vlasiuk I, Gren T, Luzhetskyy A, Kalinowski J, Fedorenko V, Ostash B (2018) Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions. Appl Microbiol Biotechnol 102(19):8419–8428. https://doi.org/10.1007/s00253-018-9249-1

Yushchuk O, Horbal L, Ostash B, Marinelli F, Wohlleben W, Stegmann E, Fedorenko V (2019) Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. Appl Microbiol Biotechnol 103(10):4089–4102. https://doi.org/10.1007/s00253-019-09789-w

Yushchuk O, Homoniuk V, Ostash B, Marinelli F, Fedorenko V (2020) Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. J Antibiot (Tokyo) Epub ahead of print:1–5. https://doi.org/10.1038/s41429-019-0274-9

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s00253-020-10436-y.

Анотація: *Тейкопланін (Tcp) є клінічно-значущим глікопептидним антибіотиком (ГПА), який синтезується актинобактерією* Actinoplanes *teichomyceticus. Tcp є пріоритетним антибіотиком для лікування важких інфекцій, спричинених мультирезистентними Грам-позитивними мікроорганізмами у дорослих і немовлят. У цьому огляді ми надаємо детальний аналіз біосинтезу Tcp в A. teichomyceticus, описуючи етапи біосинтезу, регуляцію і стійкість до Tcp. Зокрема, ми узагальнюємо знання, отримані під час досліджень in vivo та in vitro, щоб створити інтегровану модель біосинтезу тейкопланіну. Потім ми обговорюємо генетичні та інші фактори, які сприяють регуляції біосинтезу тейкопланіну, зосереджуючись на тих, які були успішно застосовані для покращення продукції тейкопланіну. Також наведено сучасний погляд на механізми стійкості до тейкопланіну в A. teichomyceticus, і порівняно кластер біосинтетичних генів Tcp з іншими кластерами генів глікопептидів з актинопланет і неідентифікованих ізолятів/метагеномних зразків.*

Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-018-09601-1

MINI-REVIEW

Landomycin biosynthesis and its regulation in *Streptomyces*

Oleksandr Yushchuk[1] · Madan Kharel[2] · Iryna Ostash[1] · Bohdan Ostash[1]

Received: 21 November 2018 / Revised: 26 December 2018 / Accepted: 27 December 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract
This mini-review is centered on genetic aspects of biosynthesis of landomycins (La), a family of angucycline polyketides. From the very discovery in the 1990s, La were noted for unusual structure and potent anticancer properties. La are produced by a few actinobacteria that belong to genus *Streptomyces*. Biochemical logic behind the production of La aglycon and glycoside halves and effects of La on mammalian cells have been thoroughly reviewed in 2009–2012. Yet, the genetic diversity of La biosynthetic gene clusters (BGCs) and regulation of their production were not properly reviewed since discovery of La. Here, we aim to fill this gap by focusing on three interrelated topics. First, organization of known La BGCs is compared. Second, up-to-date scheme of biosynthetic pathway to landomycin A (LaA), the biggest (by molar weight) member of La family, is succinctly outlined. Third, we describe genetic and nutritional factors that influence La production and export. A summary of the practical utility of the gained knowledge and future directions to study La biosynthesis conclude this mini-review.

Keywords Angucyclines · Landomycins · Biosynthesis · Regulation · Antibiotic export

Introduction

Landomycins (La) are a sizable family of angucyclic polyketides (Kharel et al. 2012a) possessing antiproliferative properties (Ostash et al. 2009; Lehka et al. 2015; Panchuk et al. 2017). Landomycin A (LaA; Fig. 1)—the founding and also the biggest (by molar mass) member of the family—is produced by the bacterial species *Streptomyces cyanogenus* S136 (=DSM5087) (Henkel et al. 1990). *S. globisporus* 1912 is known to produce LaE (Fig. 1), a trisaccharide metabolite that can be considered an intermediate to LaA. LaE is not accumulated by the S136 strain and the 1912 strain is the only natural source of this compound (Ostash et al. 2009).

Interests in La biosynthetic investigation are two-fold. First, understanding of La biosynthesis allows us to rationally engineer pathways leading to the development of a new class of antitumor drugs. In addition, La and related analogues may serve as molecular probes to identify novel targets for chemotherapeutic intervention. Second, La biosynthesis is full of unusual biochemical transformations that are worth studying on their own (Kharel et al. 2012b). As observed in other microbial natural products, limited production of metabolite under standard culture conditions is a major challenge for generating a large quantity of La needed to conduct various studies. We herein focus mainly on discussing genetic and biochemical circuits involved in controlling the production of La. An optimal engineering of these circuits will potentially offer access to La for various synthetic/biosynthetic, mode-of-action, and structure-activity relationship studies.

La biosynthetic gene clusters: diversity and organization

Angucyclines constitute the largest group (over 150 members) of aromatic polyketides produced by iterative (type II) polyketide synthases (PKS). Two structural features put La apart from the rest of angucyclines: hydroxyl group at C11 and *O*-glycosylation at C8 positions. Therefore, any La biosynthetic gene cluster (BGC) should carry these genes (e.g., *lanZ4-lanZ5* and *lanGT* homologs) for respective modifications (Kharel et al. 2012a). There are only three BGCs with these La pathway–specific genes available in the public databases: LaA (*lan*) and LaE (*lnd*) BGCs from *S. cyanogenus* S136

✉ Bohdan Ostash
bohdan.ostash@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, Hrushevskoho St. 4, Rm. 102, Lviv 79005, Ukraine

[2] Department of Pharmaceutical Sciences, University of Maryland Eastern Shore, Somerset Hall 214, Princess Anne, MD 21853, USA

Published online: 11 January 2019

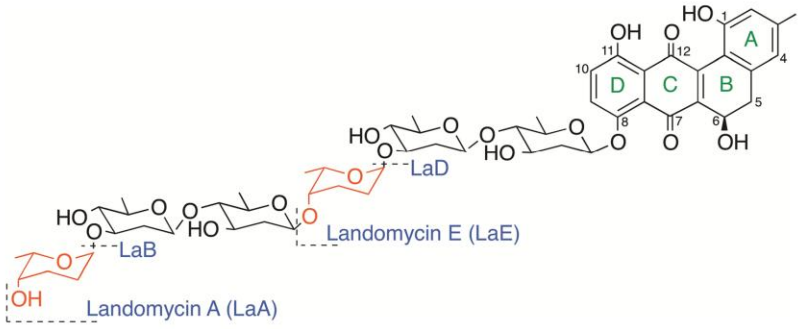


Fig. 1 Structure of landomycins A, B, D, E. L-rhodinose residues are shown in red. Carbon positions of La aglycon mentioned in the main text are numbered

(Westrich et al. 1999) and *S. globisporus* 1912 (Matselyukh et al. 2016; Fedorenko et al. 2000), respectively, and *lnd*-like BGC identified in the metagenomic DNA (Feng et al. 2011). All three BGCs share significant similarities, yet a number of notable differences exist in both structural and regulatory genes (Fig. 2). The glycosyltransferase (*lanGT3*) and the putative hexose synthase (*lanZ2*) genes that are present in the *lan* cluster are missing in the *lnd* cluster. Similarly, homolog of the TetR-like repressor gene *lanK* of the *lan* cluster is absent in the *lnd* BGC. At the same time, additional regulatory and export genes (such as *prx* (Dutko et al. 2006), *lndW-W2* (Ostash et al. 2008b), *lndY-lndYR* (Ostash et al. 2011)) that flank *lnd* BGC of *S. globisporus* 1912 are missing in the *lan* cluster. The implication of these differences in regulatory circuits and transcriptional organizations in La BGCs remains unknown.

Landomycin biosynthetic pathway

A La angucyclic scaffold is produced by PKS (LanABC), in conjunction with polyketide reductases and cyclases (Ostash et al. 2003; Kharel et al. 2010, 2012b). Once formed, the tetracyclic framework undergoes a cascade of oxidation, reduction, and sugar transfer reactions (Fig. 3), which became clear only recently once the powers of genetic analysis and in vitro assays were combined (Zhu et al. 2005; Baig et al. 2006; Kharel et al. 2012b; Paananen et al. 2013). La PKS utilizes acetate as the starter unit and malonate as an extender unit to assemble the angucycline intermediate. The cyclized product undergoes a sequence of modifications (hydrolysis, decarboxylation, and LanM2/LndM2-catalyzed dehydration) to produce the 2,3-dehydroUWM6, or prejadomycin. C-12 oxygenase (Lan/LndE) installs a hydroxyl group at the C-12 position of prejadomycin, which is followed by the reduction of a 6-keto group by Lan/LndV to produce 11-deoxylandomycinone. The latter is modified with the addition of a D-olivose by glycosyltransferase (GT) LanGT2 (Luzhetskyy et al. 2005a). Genes for production of D-olivose and the other deoxysugar used in the La pathway, L-rhodinose, are located within *lan*/*lnd* BGCs. The LanGT2-catalyzed glycosylation step is followed by C-11 oxygenation. This reaction is catalyzed by a pair of enzymes, flavin reductase LanZ4, and bifunctional oxygenase-dehydratase LanZ5, to produce LaI. Sequential actions of GTs (LanGT1 and LanGT4) facilitate the formation of the trisaccharide chain of LaE, the final product of La BGC in *S. globisporus* 1912 (Ostash et al. 2004). LaE biosynthesis is identical in 1912 and S136 strains. In the latter, the trisaccharide chain undergoes further extension via sequential action of GTs LanGT3, LanGT1, and LanGT4 to generate a hexasaccharide unit in the final product LaA (Luzhetskyy et al. 2004). A remarkable feature of LanGT1 and LanGT4 is that they are used twice

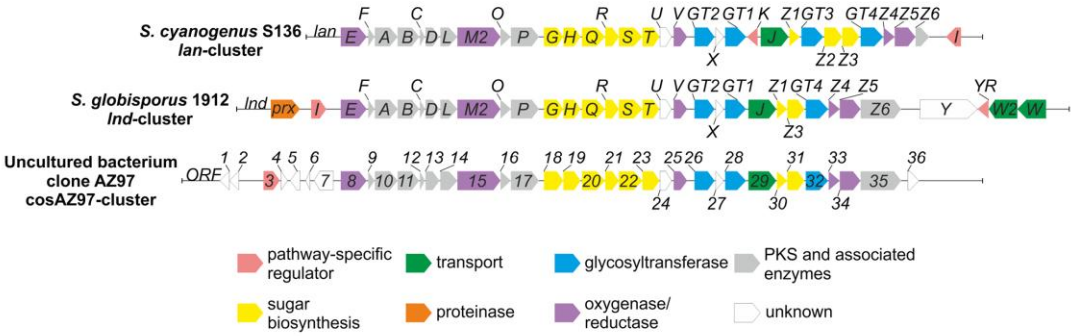


Fig. 2 Genetic organization of La BGCs. Please see the graphical legend for color code of the genes

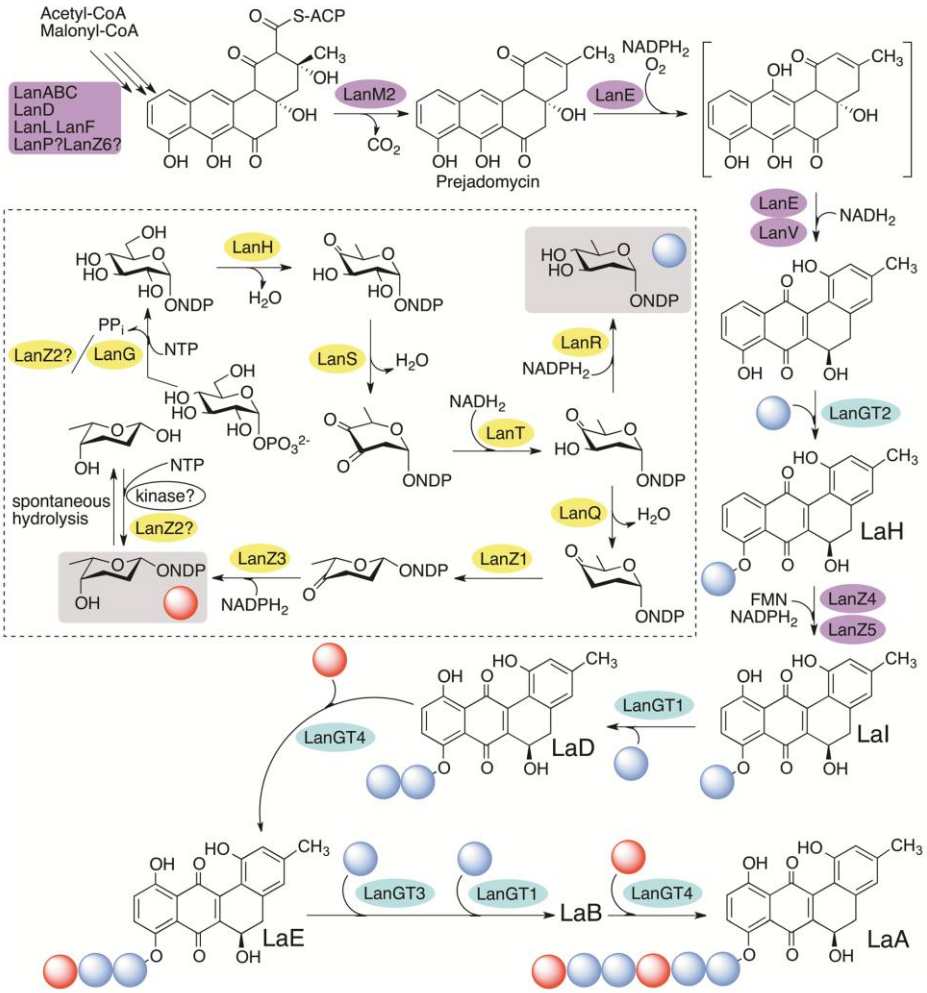


Fig. 3 Biosynthesis of LaA. Proteins involved in aglycon production and maturation are shown as violet ovals; glycosyltransferases—blue ovals; deoxysugar biosynthesis—yellow. Deoxysugar residues are sketched as colored spheres (blue for D-olivose and red for L-rhodinose). Enzymes designations, whose participation in certain step has no experimental or strong sequence homology–based evidence, are shown with interrogation sign (e.g. LanP?). Please see Fig. 1 and the main text for more details

during the LaA biosynthesis (Luzhetskyy et al. 2005b; Erb et al. 2009). LndGT1, the LanGT1 hologue in *S. globisporus* 1912, is not capable of functioning in an iterative manner. This inability was tracked to a few amino acid changes within N-terminal domain of LndGT1 (Krauth et al. 2009). Thus, the ability of the S136 strain to produce LaA is attributed to the presence of additional GT, LanGT3, and relaxed substrate specificity of the LanGT1 as compared to the LndGT1 of the 1912 counterpart. Lastly, hexose synthase gene *lanZ2* within LaA BGC is proposed to reactivate L-rhodinose which may be produced through a facile hydrolysis of NDP-L-rhodinose (Rohr et al. 1997; Luzhetskyy et al. 2005b). Such action of LanZ2 would not only sustain the sugar demand for LaA biosynthesis *S. cyanogenus*, but also increase the biosynthesis process by salvaging L-rhodinose produced through the multi-step enzymatic reactions. Yet, it cannot be ruled out at the moment that LanZ2 is just a spare copy of main nucleotidyltransferase LanG (Fig. 3).

Imbalanced expression of genes within La BGCs in native (Zhu et al. 2007; Erb et al. 2009; Shaaban et al. 2011) and heterologous (von Mulert et al. 2004; Myronovskyi et al. 2016) hosts and their genetic manipulations led to the isolation of over 30 new La; readers are referred to the recent review (Kharel et al. 2012a) where this topic was exhaustively covered.

discovered near LaE BGC (*prx*, *lndYR*) by sequencing *lnd* cluster flanks, or were revealed by transposon mutagenesis elsewhere in the 1912 genome (Dutko et al. 2006; Horbal et al. 2013). Both *lanI* and *lndI* genes carry the rare leucyl codon TTA (Fig. 4a), which limits the level and temporal window of translation of *lanI*/*lndI* mRNA (Rebets et al. 2005, 2006). Indeed, the introduction of extra copy of $tRNA^{Leu}_{UAA}$ *bldA* gene into *S. cyanogenus* enhanced LaA titers by 14-fold (Gessner et al. 2015). LaE-deficient 1912 strains also lost their ability to produce a diketopiperazine (DKP) family metabolite. The mechanism of interplay between biosynthesis of DKP-like metabolite and LaE biosynthesis is yet to be elucidated (Matselyukh et al. 2015).

Although La lack DNA intercalating properties (Korynevska et al. 2007), they are still toxic to their own producers and other streptomycetes (Rebets et al. 2003b; Ostash et al. 2005, 2008a, b). The mechanism of self-resistance against La in producer strains is not well understood. Similarly, molecular targets of La in bacteria and mammals remain elusive. La BGCs hint that active export might be a major self-resistance mechanism. In *S. cyanogenus* S136, *lanJ* encodes a proton-dependent antiporter protein. Overexpression of this gene confers resistance to LaA and shifts La production pattern of S136 strains towards the increased accumulation of La with shorter glycoside chains (e.g., LaD). The TetR-type repressor gene *lanK* is located divergently to *lanJ* allowing both genes to share the common promoter *lanKJp*. LanK binds *lanKJp* and inhibits *lanJ* transcription. La with saccharide chains of at least three residues are capable of interacting with LanK (Ostash et al. 2010) and thus prevent it from DNA binding. As a result, *lanJ* expression will be blocked until a certain amount of LaA is accumulated in the cell. The latter happens due to the basal read-through of *lanKJp* and expression of *lanGT3* and *lanGT4* genes. Once it exceeds the threshold level, LaA relieves *lanKJp* from LanK, and turns on a full-fledged expression of *lanJ-lanGT4* operon. This enables efficient La export and conversion of LaD into LaA (Fig. 4b). The *lanJ* was shown to control the export of not only LaA, but also its intermediates. Thus, the biological significance of temporal *lanJ* repression might be to prevent the early export of LaD resulting in its accumulation inside the cell and conversion into the fully assembled product LaA (Ostash et al. 2008a, b). Indeed, there is a *lanJ* functional counterpart-*lndJ* in the LaE BGC, but the *lanK* homolog is missing (Ostash et al. 2007). Perhaps, *lanK* is not needed for LaE production given its short triglycoside chain. Besides *lndJ*, the "right" arm of *lnd* cluster carries four genes, *lndY*, *lndYR*, *lndW2*, and *lndW* (Tsypik et al. 2013). Their role in LaE production is obscure (Ostash et al. 2011). Bioinformatics study revealed that *lndW* encodes an ATP-binding subunit of ABC transporters family protein. Overexpression of this gene in *S. lividans* conferred resistance against high concentrations of LaE (Ostash et al. 2008a, b).

Outlook

Since their discovery in 1990, La drew attention of chemists and biologists alike, which has translated into a large body of knowledge about their biosynthesis and biological activities. Elucidation of anticancer activities of various La has been a major focus of various groups in recent years (Korynevska et al. 2007; Panchuk et al. 2017). This feeds back into basic biological studies, aiming at the generation of novel La and improvement of their production. There is a solid understanding of most steps of the LaA biosynthetic pathway, yet the roles of a few genes remain enigmatic. Particularly, the role of LanO in La biosynthesis is largely unknown, and the functions of phosphopantheteinyltransferase LanZ6 and hexose synthase LanZ2 remain speculative. If the latter is involved in the activation of L-rhodinose salvage pathway, it would be an invaluable tool for chemoenzymatic production of deoxysugars. Next, the molecular basis for discrimination of La with different glycan chains by repressor LanK is yet to be understood (Ostash et al. 2010). Further insights into LanK-based regulatory circuit can pave the way for the design of a better reporter system for rapid in vivo screening of improved strains and novel La producers. Similarly, additional structural biology studies of Lan proteins would shed more light into the understanding of their molecular mechanisms, as demonstrated recently through the characterization of LanV and LanGT2 (Paananen et al. 2013; Tam et al. 2015).

Production of La through microbiological fermentation is the only fiscally viable route available at present. Although wild-type and some engineered La producers are available (Gromyko et al. 2004; Lavrinchuk et al. 2011; Yushchuk et al. 2018), they are not optimal sources to fulfill the growing demand of La. This highlights the need for the development of higher titer producer strains. La BGC offers a great potential for the use of genetic engineering for the strain improvement, as only a few components of La regulatory network have been uncovered to date. Draft genomes of S136 and 1912 strains have become recently available (Yushchuk et al. 2018; Prof. Matselyukh, accession # NZ_QWFA00000000.1) and their exploration can potentially offer additional help in strain engineering. Also, LaA BGC emerges as an interesting model of conditionally cryptic pathways. Elucidation of physiological and genetic reasons for its crypticity will allow us to devise novel approaches towards activation of silent secondary metabolome. Finally, elucidation of La resistance might also offer a fresh view on the mode of action and overproduction of La. We now know that expression of putative export genes (*lanJ*, *lndJ*, *lndW*) contribute to increased La resistance when these genes are overexpressed from plasmids. However, whether La export is the only mechanism of self-defense or other yet unknown mechanisms of resistance co-exist is still unclear. Further studies addressing this question may offer additional opportunities to engineer La overproducer strains.

Regulation of La production and export

La production is favored by a relatively high glucose level and trace amounts of cobalt salts (Yushchuk et al. 2018). At the same time, no LaA production is observed in many typical liquid media, such as tryptic-soy broth. This makes LaA biosynthesis an interesting case of a conditionally cryptic pathway. Several regulatory genes were shown to be implicated in the La production, and we will consider them first in *S. cyanogenus*. The gene *lanI* is situated at the very end of the BGC (Fig. 4a). LanI upregulates its own transcription and transcription of all structural *lan* genes, as well as the La exporter system *lanJ-lanK* (Rebets et al. 2008). Interestingly, *S. cyanogenus* lacks a functional gene for the AdpA—a global regulator of morphogenesis and secondary metabolism commonly found in *Streptomyces* (Yushchuk et al. 2018). Introduction of heterologous AdpA into the S136 strain led to increased LaA titers, presumably via upregulation of *lanI*. The LaE BGC of 1912 strain encodes LndI, a counterpart of LanI, which likely functions in the same way (Rebets et al. 2003a; Rebets et al. 2005). The role of AdpA in *lndI* expression remains unknown, while several other regulatory genes are proposed to influence *lndI* expression. These genes either were

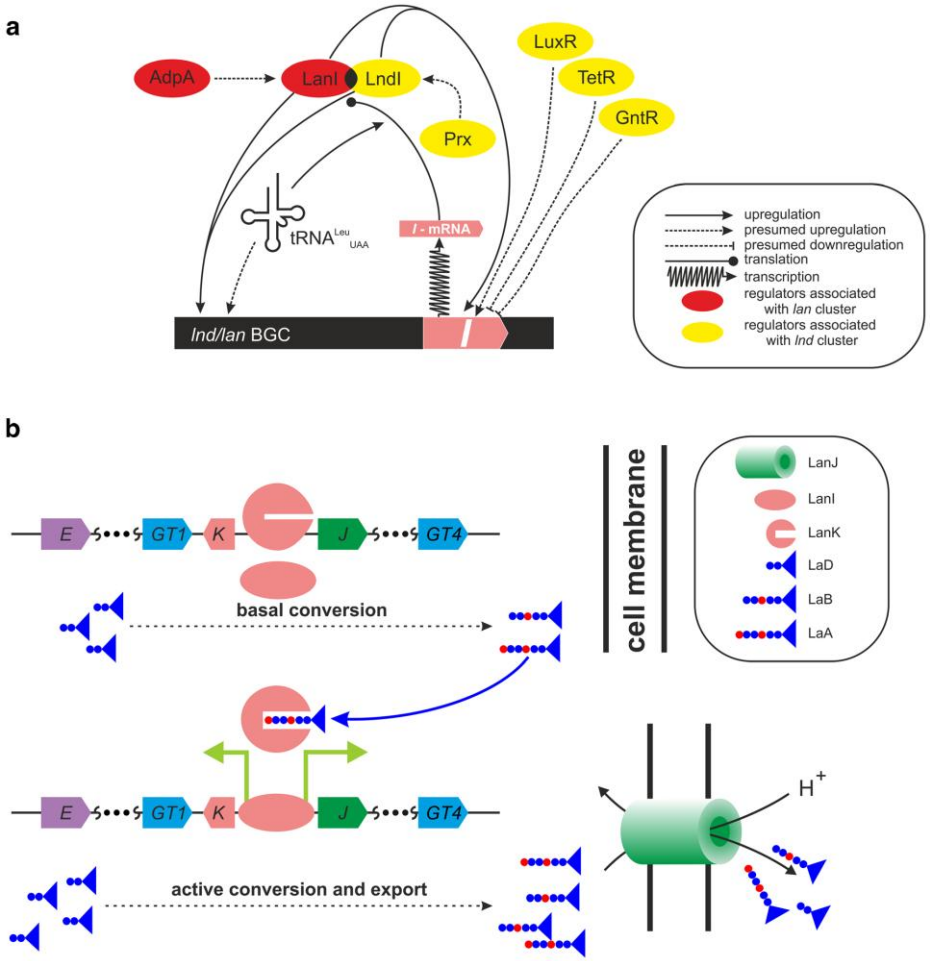


Fig. 4 Regulatory and export circuits in LaA production. **a** Gene *lanI/lndI* is believed to be a central hub that integrates regulatory influences of higher order regulators. The mRNA of the aforementioned gene is subject to TTA-mediated translational regulation. **b** Coordination of LaA export with its production. Interaction of LanK with *lanK-lanJ* intergenic region blocks both LanJ-mediated extrusion of La and conversion of LaD into LaA. Basal read-through of *lanJ-lanGT4* segment allows accumulation of small amounts of LaA; the latter cancels the repression and amplifies its own biosynthesis. Please see graphical legend and the main text for more details

Acknowledgements We thank all current and former members of Fedorenko group (Lviv University) who contributed to the studies of La. We also sincerely apologize to everyone who contributed to studies of La, and whose works were not cited due to page constraints.

Funding This work was supported by grants of Ministry of Education and Science of Ukraine M/26-2018 and Bg-41Nr (to B.O.).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical statement This article does not contain any studies with human participants or animals performed by any of the authors.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Baig I, Kharel M, Kobylyanskyy A, Zhu L, Rebets Y, Ostash B, Luzhetskyy A, Bechthold A, Fedorenko VA, Rohr J (2006) On the acceptor substrate of C-glycosyltransferase UrdGT2: three prejadomycin C-glycosides from an engineered mutant *of Streptomyces globisporus* 1912 Δ*lndE(urdGT2)*. Angew Chem Int Ed Engl 45(46):7842–7846. https://doi.org/10.1002/anie.200603176

Dutko L, Rebets Y, Ostash B, Luzhetskyy A, Bechthold A, Nakamura T, Fedorenko V (2006) A putative proteinase gene is involved in regulation of landomycin E biosynthesis in *Streptomyces globisporus* 1912. FEMS Microbiol Lett 255(2):280–285. https://doi.org/10.1111/j.1574-6968.2005.00085.x

Erb A, Krauth C, Luzhetskyy A, Bechthold A (2009) Differences in the substrate specificity of glycosyltransferases involved in landomycins A and E biosynthesis. Appl Microbiol Biotechnol 83(6):1067–1076. https://doi.org/10.1007/s00253-009-1993-9

Fedorenko V, Basiliya L, Pankevych K, Dubitska L, Ostash B, Luzhetskyy A, Gromyko O, Krugel H (2000) Genetic control of antitumor antibiotics-polyketides biosynthesis by actinomycetes. Bull Inst Agr Microbiol (Ukr) 8:27–31

Feng Z, Kallifidas D, Brady SF (2011) Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites. Proc Natl Acad Sci U S A 108(31):12629–12634. https://doi.org/10.1073/pnas.1103921108

Gessner A, Heitzler T, Zhang S, Klaus C, Murillo R, Zhao H, Vanner S, Zechel DL, Bechthold A (2015) Changing biosynthetic profiles by expressing *bldA* in *Streptomyces* strains. Chembiochem 16(15):2244–2252. https://doi.org/10.1002/cbic.201500297

Gromyko O, Rebets Y, Ostash B, Luzhetskyy A, Fukuhara M, Bechthold A, Nakamura T, Fedorenko V (2004) Generation of *Streptomyces globisporus* SMY622 strain with increased landomycin E production and it's initial characterization. J Antibiot 57(6):383–389. https://doi.org/10.7164/antibiotics.57.383

Henkel T, Rohr J, Beale JM, Schwenen L (1990) Landomycins, new angucycline antibiotics from *Streptomyces* sp. I. Structural studies on landomycins A-D. J Antibiot 43(5):492–503. https://doi.org/10.7164/antibiotics.43.492

Horbal L, Fedorenko V, Bechthold A, Luzhetskyy A (2013) A transposon-based strategy to identify the regulatory gene network responsible for landomycin E biosynthesis. FEMS Microbiol Lett 342(2):138–146. https://doi.org/10.1111/1574-6968.12117

Kharel MK, Pahari P, Lian H, Rohr J (2010) Enzymatic total synthesis of rabelomycin, an angucycline group antibiotic. Org Lett 12(12):2814–2817. https://doi.org/10.1021/ol1009009

Kharel MK, Pahari P, Shepherd MD, Tibrewal N, Nybo SE, Shaaban KA, Rohr J (2012a) Angucyclines: biosynthesis, mode-of-action, new natural products, and synthesis. Nat Prod Rep 29(2):264–325. https://doi.org/10.1039/c1np00068c

Kharel MK, Pahari P, Shaaban KA, Wang G, Morris C, Rohr J (2012b) Elucidation of post-PKS tailoring steps involved in landomycin biosynthesis. Org Biomol Chem 10(21):4256–4265. https://doi.org/10.1039/c2ob07171a

Korynevska A, Heffeter P, Matselyukh B, Elbling L, Micksche M, Stoika R, Berger W (2007) Mechanisms underlying the anticancer activities of the angucycline landomycin E. Biochem Pharmacol 74(12):1713–1726. https://doi.org/10.1016/j.bcp.2007.08.026

Krauth C, Fedoryshyn M, Schleberger C, Luzhetskyy A, Bechthold A (2009) Engineering a function into a glycosyltransferase. Chem Biol 16(1):28–35. https://doi.org/10.1016/j.chembiol.2008.12.003

Lavrinchuk B, Kotenko KA, Golembiovs'ka SL, Matseliukh BP (2011) Obtaining and characteristic of *Streptomyces globisporus* 1912 mutants with high biosynthesis of landomycin E. Mikrobiol Z (Ukr) 73(6):31–35

Lehka LV, Panchuk RR, Berger W, Rohr J, Stoika RS (2015) The role of reactive oxygen species in tumor cells apoptosis induced by landomycin A. Ukr Biochem J 87(5):72–82. https://doi.org/10.15407/ubj87.05.072

Luzhetskyy A, Liu T, Fedoryshyn M, Ostash B, Fedorenko V, Rohr J, Bechthold A (2004) Function of *lanGT3*, a glycosyltransferase gene involved in landomycin A biosynthesis. Chembiochem 5(11):1567–1570. https://doi.org/10.1002/cbic.200400123

Luzhetskyy A, Taguchi T, Fedoryshyn M, Dürr C, Wohlert SE, Novikov V, Bechthold A (2005a) LanGT2 catalyzes the first glycosylation step during landomycin A biosynthesis. Chembiochem 6(8):1406–1410. https://doi.org/10.1002/cbic.200500018

Luzhetskyy A, Fedoryshyn M, Dürr C, Taguchi T, Novikov V, Bechthold A (2005b) Iteratively acting glycosyltransferases involved in the hexasaccharide biosynthesis of landomycin A. Chem Biol 12(7):725–729. https://doi.org/10.1016/j.chembiol.2005.05.008

Matselyukh B, Mohammadipanah F, Laatsch H, Rohr J, Efremenkova O, Khilya V (2015) N-methylphenylalanyl-dehydrobutyrine diketopiperazine, an A-factor mimic that restores antibiotic biosynthesis and morphogenesis in *Streptomyces globisporus* 1912-B2 and *Streptomyces griseus* 1439. J Antibiot 68(1):9–14. https://doi.org/10.1038/ja.2014.86

Matselyukh BP, Polishchuk LV, Lukyanchuk VV, Golembiovska SL, Lavrenchuk VY (2016) Sequences of landomycin E and carotenoid biosynthetic gene clusters, and molecular structure of transcriptional regulator of *Streptomyces globisporus* 1912. Mikrobiol Z 78(6):60–70

Myronovskyi M, Brötz E, Rosenkränzer B, Manderscheid N, Tokovenko B, Rebets Y, Luzhetskyy A (2016) Generation of new compounds through unbalanced transcription of landomycin A cluster. Appl Microbiol Biotechnol 100(21):9175–9186. https://doi.org/10.1007/s00253-016-7721-3

Ostash B, Rebets Y, Yuskevich V, Luzhetskyy A, Tkachenko V, Fedorenko V (2003) Targeted disruption of *Streptomyces globisporus lndF* and *lndL* cyclase genes involved in landomycin E biosynthesis. Folia Microbiol (Praha) 48(4):484–488. https://doi.org/10.1007/BF02931329

Ostash B, Rix U, Rix LL, Liu T, Lombo F, Luzhetskyy A, Gromyko O, Wang C, Braña AF, Méndez C, Salas JA, Fedorenko V, Rohr J (2004) Generation of new landomycins by combinatorial biosynthetic manipulation of the LndGT4 gene of the landomycin E cluster in *S. globisporus*. Chem Biol 11(4):547–555. https://doi.org/10.1016/j.chembiol.2004.03.011

Ostash B, Kozarevska I, Fedorenko V (2005) Generation of landomycin D-producing strain *Streptomyces globisporus* LD3. Folia Microbiol (Praha) 50(1):19–23. https://doi.org/10.1007/BF02931289

Ostash I, Ostash B, Walker S, Fedorenko V (2007) Proton-dependent transporter gene *lndJ* confers resistance to landomycin E in *Streptomyces globisporus*. Genetika 43(8):1032–1037. https://doi.org/10.1134/S1022795407080030

Ostash I, Ostash B, Luzhetskyy A, Bechthold A, Walker S, Fedorenko V (2008a) Coordination of export and glycosylation of landomycins in *Streptomyces cyanogenus* S136. FEMS Microbiol Lett 285(2):195–202. https://doi.org/10.1111/j.1574-6968.2008.01225.x

Ostash I, Rebets Y, Ostash B, Kobylyanskyy A, Myronovskyy M, Nakamura T, Walker S, Fedorenko V (2008b) An ABC transporter encoding gene *lndW* confers resistance to landomycin E. Arch Microbiol 190(1):105–109. https://doi.org/10.1007/s00203-008-0367-5

Ostash B, Korynevska A, Stoika R, Fedorenko V (2009) Chemistry and biology of landomycins, an expanding family of polyketide natural products. Mini-Rev Med Chem 9(9):1040–1051. https://doi.org/10.2174/138955709788922593

Ostash B, Ostash I, Zhu L, Kharel MK, Luzhetskiĭ A, Bechthold A, Walker S, Rohr J, Fedorenko V (2010) Properties of *lanK*-based regulatory circuit involved in landomycin biosynthesis in *Streptomyces cyanogenus* S136. Genetika 46(5):604–609. https://doi.org/10.1134/S1022795410050030

Ostash B, Rebets Y, Myronovskyy M, Tsypik O, Ostash I, Kulachkovskyy O, Datsyuk Y, Nakamura T, Walker S, Fedorenko V (2011) Identification and characterization of the *Streptomyces globisporus* 1912 regulatory gene *lndYR* that affects sporulation and antibiotic production. Microbiology 157(Pt 4):1240–1249. https://doi.org/10.1099/mic.0.045088-0

Paananen P, Patrikainen P, Kallio P, Mäntsälä P, Niemi J, Niiranen L, Metsä-Ketelä M (2013) Structural and functional analysis of angucycline C-6 ketoreductase LanV involved in landomycin biosynthesis. Biochemistry 52(31):5304–5314. https://doi.org/10.1021/bi400712q

Panchuk RR, Lehka LV, Terenzi A, Matselyukh BP, Rohr J, Jha AK, Downey T, Kril IJ, Herbacek I, van Schoonhoven S, Heffeter P, Stoika RS, Berger W (2017) Rapid generation of hydrogen peroxide contributes to the complex cell death induction by the angucycline antibiotic landomycin E. Free Radic Biol Med 106:134–147. https://doi.org/10.1016/j.freeradbiomed.2017.02.024

Rebets Y, Ostash B, Luzhetskyy A, Hoffmeister D, Brana A, Mendez C, Salas JA, Bechthold A, Fedorenko V (2003a) Production of landomycins in *Streptomyces globisporus* 1912 and *S. cyanogenus* S136 is regulated by genes encoding putative transcriptional activators. FEMS Microbiol Lett 222(1):149–153. https://doi.org/10.1016/S0378-1097(03)00258-1

Rebets YU, Ostash B, Gromyko O, Mik U, Basiliya L, Fedorenko V (2003b) Resistance of recombinant *Streptomyces albus* Cest114 and *Streptomyces globisporus* 1912 strains to landomycin E. Visnyk Lviv Univ (Biol Ser) 33:47–54

Rebets Y, Ostash B, Luzhetskyy A, Kushnir S, Fukuhara M, Bechthold A, Nashimoto M, Nakamura T, Fedorenko V (2005) DNA-binding activity of LndI protein and temporal expression of the gene that upregulates landomycin E production in *Streptomyces globisporus* 1912. Microbiology 151(Pt 1):281–290. https://doi.org/10.1099/mic.0.27244-0

Rebets YV, Ostash BO, Fukuhara M, Nakamura T, Fedorenko VO (2006) Expression of the regulatory protein LndI for landomycin E production in *Streptomyces globisporus* 1912 is controlled by the availability of tRNA for the rare UUA codon. FEMS Microbiol Lett 256(1):30–37. https://doi.org/10.1111/j.1574-6968.2005.00087.x

Rebets Y, Dutko L, Ostash B, Luzhetskyy A, Kulachkovskyy O, Yamaguchi T, Nakamura T, Bechthold A, Fedorenko V (2008) Function of *lanI* in regulation of landomycin a biosynthesis in *Streptomyces cyanogenus* S136 and cross-complementation studies with *Streptomyces* antibiotic regulatory proteins encoding genes. Arch Microbiol 189(2):111–120. https://doi.org/10.1007/s00203-007-0299-5

Rohr J, Wohlert SE, Oelkers C, Kirschning A, Ries M (1997) Biosynthetic short activation of the 2,3,6-trideoxysugar L-rhodinose. Chem Commun 10:973–974. https://doi.org/10.1039/A702125I

Shaaban KA, Stamatkin C, Damodaran C, Rohr J (2011) 11-Deoxylandomycinone and landomycins X-Z, new cytotoxic angucyclin(on)es from a *Streptomyces cyanogenus* K62 mutant strain. J Antibiot 64(1):141–150. https://doi.org/10.1038/ja.2010.121

Tam HK, Härle J, Gerhardt S, Rohr J, Wang G, Thorson JS, Bigot A, Lutterbeck M, Seiche W, Breit B, Bechthold A, Einsle O (2015) Structural characterization of O- and C-glycosylating variants of the landomycin glycosyltransferase LanGT2. Angew Chem Int Ed Engl 54(9):2811–2815. https://doi.org/10.1002/anie.201409792

Tsypik O, Ostash B, Rebets'Iu, Fedorenko V (2013) Characterization of *Streptomyces globisporus* 1912 *lnd* cluster region containing *lndY, lndYR, lndW2* and *lndW* genes. Tsitol Genet (Ukr.) 47(1):11–16. https://doi.org/10.3103/S0095452713010106

von Mulert U, Luzhetskyy A, Hofmann C, Mayer A, Bechthold A (2004) Expression of the landomycin biosynthetic gene cluster in a PKS mutant of *Streptomyces fradiae* is dependent on the coexpression of a putative transcriptional activator gene. FEMS Microbiol Lett 230(1):91–97. https://doi.org/10.1016/S0378-1097(03)00861-9

Westrich L, Domann S, Faust B, Bedford D, Hopwood DA, Bechthold A (1999) Cloning and characterization of a gene cluster from *Streptomyces cyanogenus* S136 probably involved in landomycin biosynthesis. FEMS Microbiol Lett 170(2):381–387. https://doi.org/10.1111/j.1574-6968.1999.tb13398.x

Yushchuk O, Ostash I, Vlasiuk I, Gren T, Luzhetskyy A, Kalinowski J, Fedorenko V, Ostash B (2018) Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions. Appl Microbiol Biotechnol 102(19):8419–8428. https://doi.org/10.1007/s00253-018-9249-1

Zhu L, Ostash B, Rix U, Nur-E-Alam M, Mayers A, Luzhetskyy A, Mendez C, Salas JA, Bechthold A, Fedorenko V, Rohr J (2005) Identification of the function of gene *lndM2* encoding a bifunctional oxygenase-reductase involved in the biosynthesis of the antitumor antibiotic landomycin E by *Streptomyces globisporus* 1912 supports the originally assigned structure for landomycinone. J Org Chem 70(2):631–638. https://doi.org/10.1021/jo0483623

Zhu L, Luzhetskyy A, Luzhetska M, Mattingly C, Adams V, Bechthold A, Rohr J (2007) Generation of new landomycins with altered saccharide patterns through over-expression of the glycosyltransferase gene *lanGT3* in the biosynthetic gene cluster of landomycin A in *Streptomyces cyanogenus* S-136. Chembiochem 8(1):83–88. https://doi.org/10.1002/cbic.200600360

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s00253-018-09601-1.

Анотація: *Цей огляд зосереджений на висвітленні генетичних аспектів біосинтезу ландоміцинів (La) – ангуциклінових полікетидів. З самого відкриття в 1990-х роках La були відомі своєю незвичайною структурою та потужними протиракovими властивостями. La продукуються кількома актинобактеріями, які належать до роду Streptomyces. Біохімічна логіка, що лежить в основі біосинтезу аглікону La та подальшого глікозилювання, а також впливу La на клітини ссавців були ретельно розглянуті в 2009-2012 роках. Тим не менш, генетичне різноманіття кластерів біосинтетичних генів La (BGC) і регулювання їх біосинтезу не були належно розглянуті з моменту відкриття La. В цій роботі ми прагнемо заповнити цю прогалину, зосередившись на трьох взаємопов'язаних темах. По-перше, порівнюється організація відомих La кластерів. По-друге, лаконічно викладено сучасний погляд на особливості біосинтетичного шляху ландоміцину A (LaA), найбільшого (за молярною масою) представника родини La. По-третє, ми описуємо генетичні та інші фактори, які впливають на продукцію та експорт La.*

РОЗДІЛ 2 Матеріали і методи досліджень

2.1 Матеріали досліджень

Штами бактерій. Штам *Escherichia coli* DH5α [37] був використаний як господар для рутинного клонування, тоді як штам *E. coli* ET12567 pUZ8002^{+} використовували як донор для міжродових схрещувань [38]. В роботі використовували такі штами актиноміцетів, як: *S. coelicolor* A3(2), *S. cyanogenus* S136, *S. albus* (=*albidoflavus*) J1074, *S. sioyaensis* NRRL-B5408, *S. ghanaensis* (=*viridosporus*) ATCC14672, *S. globisporus* 1912, *N. gerenzanensis* ATCC 39727, *N. coxensis* DSM 45129, *A. rectilineatus* NRRL B-16090, *A. teichomyceticus* NRRL B-16726, та інші.

Плазміди. Набір різноманітних інтегративних та реплікативних плазмід було застосовано для конструювання похідних, залежно від конкретної мети. Інтегративні плазміди, використані в роботі, були засновані на трьох системах інтеграції актинофагів, а саме φC31 (похідні pSET152) [38]; φBT1 (похідні pRT801) [39]; VWB1 (похідні pSOK804) [40]. Реплікативні плазміди були похідними pKC1139, що несе термочутливий pSG5-реплікон [41]. Для генних нокаутів використовували суїцидну плазміду pKC1132 [38]. Репортерні плазміди, в яких різні промотори злиті з геном β-глюкуронідази *gusA*, створювали на основі pGUS [42].

2.2 Методи досліджень

2.2.1 Мікробіологічні методи досліджень

Культивування штамів бактерій. Всі компоненти живильних середовищ та антибіотики були вироблені Sigma-Aldrich, або іншими постачальниками.

Штами *E. coli* вирощували при 37 °C в бульйоні LB [37] або на середовищі LA [37], з додаванням за потреби 100 мкг/мл апраміцину сульфату, 50 мкг/мл

канаміцину сульфату, 25 мкг/мл хлорамфеніколу, 100 мкг/мл спектиноміцину, 100 мкг/мл гігроміцину.

Штами актиноміцетів вирощували при 30 °С як за умов культивування на чашках з агаризованими середовищами, так і за умов зануреного культивування. Для рутинного вирощування штамів стрептоміцетів і *Actinoplanes* spp. використовували середовище SFM або ISP3 [38] із додаванням 50 мкг/мл апраміцину сульфату; штамів *Nonomuraea* – середовище ISP3 із додаванням за потреби 50 мкг/мл апраміцину сульфату.

Для виділення геномної ДНК, штами вирощували: (1) *N. gerenzanensis* і *N. coxensis* в рідкому середовищі VSP [43]; (2) *A. teichomyceticus* і *A. rectilineatus* в рідкому середовищі ISP2 [44]; стрептоміцетів вирощували в рідкому середовищі TSB (Condalab) на орбітальному шейкері за 220 об./хв. при 30 °С протягом 72 год. Робочі банки клітин (РБК) для штамів актиноміцетів створювалися відповідно до опублікованого протоколу [45, 46]. Для виділення сумарної РНК штами *A. teichomyceticus* вирощували в рідкому середовищі TM1 [47], а штами стрептоміцетів в рідких середовищах TSB, YMPG і SG [24].

Тести на пригнічення росту тест-культур. Для визначення антимікробної активності тих чи інших штамів актинобактерій, або екстрактів антибіотиків, використовували тести пригнічення росту тест-культури методом дифузії антибіотиків в агар з агарового блока, паперовому диску Whatman (GE Healthcare), або лунки в агарі.

Як Грам-позитивні тест-культури використовувалися штами *Bacillus subtilis* ATCC 6633, *B. cereus* ATCC 19637, *Micrococcus luteus* ATCC 4698. Штами інкубували через ніч в бульйоні Мюллера-Хінтона II (Sigma Aldrich), та використовували для інокуляції свіжої культури (1% об./об.), яку підрощували до OD_{600}=0,6. Потім, 200 мкл цієї культури додавали до 25 мл 0,7% агару Мюллера-Хінтона (Condalab). Після затвердіння середовищ на поверхню агару поміщали агарові блоки, вирізані з чашок із газонами відповідних актиноміцетів, або паперові диски, що містили екстракти антибіотиків, або вирізали в агарі лунки (⌀ 5 мм), в які вносили зразки занурених культур актиноміцетів, або крапали зразки занурених культур

безпосередньо на поверхню агару. Альтернативно, використовували спори *B. subtilis* і *B. cereus* (приготовані, як описано в [48]) з розрахунку 10^6-10^7 спор на 25 мл агару Мюллера-Хінтона (Condalab). Чашки інкубували протягом 16 год при 37 °C, а тоді отримані результати документували. Як Грам-негативну тест-культуру використовували *E. coli* DH5α. Для приготування тесту на пригнічення росту, *E. coli* інкубували через ніч в бульйоні LB; нічну культуру використовували для інокуляції свіжої культури (5% об./об.), яку підрощували до OD_{600}=0,6. 100 мкл отриманої суспензії клітин додавали до 25-30 мл середовища LA. Отримані чашки із агаризованим середовищем, що містить клітини *E. coli*, використовували для приготування тестів на пригнічення росту у спосіб, описаний вище. Тест-культура *Debaryomyces hansenii* VKM Y-9 відображала дріжджові гриби (*Saccharomycetales*). *D. hansenii* культивували при 30 °C, а тести пригнічення росту виконували аналогічно до описаного для *E. coli*, однак для вирощування і для тесту використовували середовища TSB і TSA (Condalab), відповідно.

Світлова та сканувальна електронна мікроскопія. В роботі застосовувалися стандартні методи світлової мікроскопії для аналізу занурених культур актиноміцетів чи спорових суспензій (Carl Zeiss Primo Star). Мікроморфологія газонів актиноміцетів за потреби вивчалася за допомогою сканувального електронного мікроскопа JEOL JSM-T220a SEM (Jeol), як з напиленням тонкого шару міді на препарат, так і без нього.

2.2.2 Культивування штамів актиноміцетів задля продукції антибіотиків

В роботі культивували штами актиноміцетів в зануреній культурі задля вивчення продукції ними деяких антибіотиків. Зокрема це були ГПА A40926 і A50926, полікетиди ландоміцин А і люцензоміцин, типові методики для продукції яких подано нижче.

Продукція A40926 і A50926. Для старту ферментації, одну пробірку РБК (див. вище) додавали до ребристої колби Ерленмеєра (300 мл) (або колби із додаванням 10 скляних куль, ø 5 мм), що містила 50 мл вегетативного

середовища E26 [43]. Прекультури в колбах інкубували протягом 72 год на орбітальному шейкері при 220 об./хв., 30 °C, а потім використовували для інокуляції (10% об./об.) 500 мл ребристих колб Ерленмейєра, що містили 100 мл середовища FM2 [45] або для інокуляції (10% об./об.) 3-л скляного біореактора P-100 Applikon (висота 250 мм, ø 130 мм), оснащеного біоконтролером AD1030 і двигуном AD1032, що містить 2 л продукційного середовища FM2. Культивування в FM2 у колбах проводили при 30 °C і 220 об./хв. Ферментацію в біореакторі проводили при 30 °C, зі швидкістю перемішування 450 об./хв. (що відповідає швидкості лопаток 1,17 м/с) і швидкістю аерації 2 л/хв. Розчинений кисень (вимірюється як % pO2) контролювали за допомогою полярографічного кисневого електрода Ingold (Mettler Toledo). Значення pH культуральних бульйонів контролювали за допомогою pH-метра. Появу піни контролювали шляхом додавання піногасника Hodag (Hodag Chemical Corporation) через датчик піногасника. Зразки культуральної рідини збирали через регулярні інтервали культивування для оцінки накопичення біомаси (суха вага), споживання глюкози (палички Diastix, Bayer AG) і продукції A40926.

Продукцію A50926 вивчали в аналогічний спосіб, однак як продукційні середовища використовували ISP2 або TM1, а колби Ерленмеєра були не ребристі і з додаванням 10 скляних куль (ø 5 мм).

Продукція ландоміцину A. Задля продукції ландоміцину A, прекультури штамів *S. cyanogenus* засівали шляхом додавання фрагментів агару, вирізних із з газонів *S. cyanogenus* (168 год росту на агаризованому середовищі ISP3) до ребристих колб Ерленмеєра (250 мл) з 50 мл рідкого середовища TSB. Прекультури вирощували протягом 72 год на орбітальному шейкері при 220 об./хв., 30 °C. Потім, 3% (об./об.) прекультури переносили в ребристі колби Ерленмеєра (250 мл) з 50 мл середовища SG. Отримані у такий спосіб продукційні культури вирощували протягом 72 год за аналогічних умов.

Продукція люцензоміцину. Щоб розпочати ферментацію штамів *S. cyanogenus* з метою продукції люцензоміцину, в ребристі колби Ерленмеєра (250 мл), з 50 мл рідкого середовища TSB додавали фрагменти агару, вирізані з

газонів *S. cyanogenus*, вирощених на агаризованому середовищі ISP3 протягом 168 год. Прекультури інкубували протягом 72 год на орбітальному шейкері при 220 об./хв., 30 °С, а 1,5 мл отриманої прекультури потім переносили до ребристих колб Ерленмеєра (250 мл) з 50 мл середовища YMPG. Отримані культури інкубували до 120 год на орбітальному шейкері при 220 об./хв., 30 °С. З аналітичною метою, біомасу в такої продукційної культури відбирали після кожних 24 год.

2.2.3 Генетичні і генно-інженерні методи досліджень

Генетична трансформація клітин *E. coli.* Для трансформації клітин *E. coli* рекомбінантними молекулами ДНК, використовували або хімічно-компетентні клітини, або електро-компетентні клітини. Хімічно-компетентні клітини готували за традиційним протоколом з інкубуванням клітин на льоді в розчинах солей двовалентних металів (Ca^{2+} і Mg^{2+}) [37]. Електро-компетентні клітини використовували для електропорації та готували шляхом багатократного промивання крижаним 10% розчином гліцеролу [37].

Перенесення плазмід за допомогою міжродової кон'югації актиноміцетів з *E. coli*. Для перенесення плазмід у спорулюючі штами стрептоміцетів використовували типовий протокол [38]. Для цього, необхідні плазміди індивідуально переносили в неметилюючий штам *E. coli* ET12567 pUZ8002^{+}; отримані рекомбінанти використовували як донори для кон'югаційного перенесення, тоді як спорові суспензії стрептоміцетів використовували як клітини-реципієнти. Спорові суспензії готували із використання спорулюючих газонів стрептоміцетів, умови споруляції для яких підбиралися індивідуально. Прибл. 10^6 спор змішували з прибл. 10^9 клітин донора, а суміш висівали на добре підсушені чашки з агаризованими середовищами SFM або ISP3 з додаванням $MgCl_2$ (10 мМоль). Далі, після 12 год інкубування при 30 °С, кожну чашку заливали 1 мл стерильної води із селективними антибіотиками. З огляду на те, що *S. cyanogenus* є стрептоміцетом, не здатним до споруляції, протокол для кон'югаційного

перенесення плазмід в цей штам було модифіковано. Тут, як клітини-реципієнти використовували вегетативний міцелій, зібраний з газону, вирощеного на ISP3 протягом 120 год при 30 °С, а суміші із клітинами донора висівали виключно на чашки з агаризованим середовищем ISP3 з додаванням 10 мМоль $MgCl_2$.

Кон'югаційне перенесення плазмід в *N. gerenzanensis* виконували за методикою описаною раніше [43]. Для цього, всі рекомбінантні плазміди в індивідуальному порядку переносилися в неметилюючий штам *E. coli* ET12567 pUZ8002$^+$, а отримані похідні використовували як донори для кон'югаційного перенесення. Реципієнтом слугував свіжий вегетативний міцелій *N. gerenzanensis*, отриманий з культури, вирощеної в 50 мл середовища VSP (в 250 мл колбі Ерленмеєра з додаванням 10 ø 5 мм скляних куль) протягом 48 год на орбітальному шейкері при 220 об./хв. та 30 °С. Міцелій збирали за допомогою центрифугування (10 хв., 3220 g), двічі промивали із використанням 20% стерильного гліцеролу і ресуспендували в ньому же, доводячи до кінцевого об'єму 20 мл. Далі, 1 мл суспензії клітин вегетативного міцелію змішували із приблизно 10^9 клітин донорного штаму, а суміші висівали на добре підсушені чашки з агаризованим середовищем VM0.1 (+ 20 мМоль $MgCl_2$). Після 12-16 год інкубування при 30 °С, кожну чашку заливали 1 мл стерильної деіонізованої води, що містила 1,25 мг апраміцину сульфату та 750 мкг натрієвої солі налідиксової кислоти. Транскон'юганти відбиралися як стійкі до апраміцину сульфату в концентрації 50 мкг/мл.

У випадку *N. coxensis*, реципієнтом слугувала спорова суспензія, отримана з газонів, вирощених на середовищі ISP3 протягом 7 діб. Спори із однієї чашки збирали за допомогою деіонізованої стерильної води та фільтрували через один шар Miracloth (Merck KGaA, Німеччина), щоби позбутися фрагментів вегетативного міцелію. Потім спори осаджували із 50 мл фільтрованої суспензії за допомогою центрифугування (15 хв., 3220 g) і ресуспендували в 1 мл 15% гліцеролу. Для кон'югації, прибл. 10^6 спор змішували з 10^7 клітин донорних штамів *E. coli.* Висівання та подальша селекція транскон'югантів здійснювалася як і в випадку *N. gerenzanensis.*

При кон'югаційному перенесенні плазмід в *A. teichomyceticus* і *A. rectilineatus* реципієнтами слугували спори. Спорові суспензії готували у спосіб, описаний раніше [49]. Далі, прибл. 10^6 спор змішували з прибл. 10^9 клітин донора, а суміш висівали на чашки з агаризованими середовищами SFM або ISP3 з додаванням $MgCl_2$ (20-40 мМоль). Після 12 год інкубування при 30 °С, кожну чашку заливали 1 мл стерильної води із селективними антибіотиками.

Отриманих транскон'югантів у кожному із випадків перевіряли за допомогою ПЛР.

Горизонтальний агарозний гель-електрофорез та елюювання фрагментів ДНК з агарозного гелю. Електрофорез ДНК здійснювали в горизонтальному агарозному гелі в ТРІС-ацетатному буфері (ТАЕ) за напруги 100-150 В. До складу ТАЕ буферу входило: 40 мМоль Тріс, pH 8,0; 20 мМоль CH_3COONa; 1 мМоль Na_2ЕДТА; 25 мМоль CH_3COOH.

Для рутинного аналізу молекул ДНК в діапазоні 500-8000 п.н. використовували 1% агарозний гель; для розділення молекул ДНК з метою подальшого елюювання використовували 0,7% гель; для візуалізація малих фрагментів ДНК (>500 п.н., наприклад ампліконів отриманих в результаті ПЛР із зворотною транскрипцією – ЗТ-ПЛР) використовували 1,5-2% гель. 0,7% агарозний гель.

Для визначення розміру аналізованих фрагментів ДНК порівнювали їх рухливість із рухливістю комерційних маркерів ДНК із відомими розмірами фрагментів, наприклад GeneRuler 1 kb DNA Ladder (ThermoFisher Scientific). Візуалізацію фрагментів ДНК в гелі здійснювали з додаванням до гелю броміду етидію (1 мкг/мл) на трансілюмінаторах різних виробників.

Елюювання бажаних фрагментів ДНК здійснювали за допомогою комерційних наборів, наприклад GeneJET PCR Purification Kit (ThermoFisher Scientific), згідно із рекомендаціями виробника.

Ферментативна обробка ДНК. Очищену плазмідну ДНК чи амплікони, отримані за допомогою ПЛР, обробляли бажаними ендонуклеазами рестрикції (ThermoFisher Scientific або New England Biolabs) відповідно до рекомендацій

виробника. Лінеаризовані плазміди, які використовували як шасі для клонування генів, додатково обробляли за допомогою лужної фосфатази FastAP Thermosensitive Alkaline Phosphatase (ThermoFisher Scientific), щоб дефосфорилювати 5’-кінці та запобігти самолігування. Щоб позбутися виступаючих 5’- чи 3’-кінців використовували ДНК полімеразу фага Т4 (T4 DNA Polymerase, ThermoFisher Scientific). Лігування фрагментів ДНК здійснювали з використання ДНК-лігази фага Т4 (від різних виробників) згідно до рекомендацій виробника. Отримані рекомбінантні молекули ДНК перевіряли за допомогою картування із використанням ендонуклеаз рестрикції і комерційного секвенування.

Виділення плазмідної та хромосомної ДНК. Плазмідну ДНК з *E. coli* з аналітичною чи препаративною метою виділяли за допомогою методики лужного лізису [37], для секвенування – за допомогою GeneJET Plasmid Miniprep Kit (ThermoFisher Scientific). Хромосомну ДНК актиноміцетів виділяли за допомогою методу Кірбі [38].

Виділення сумарної РНК і синтез кДНК. У бажані часові точки, міцелій досліджуваних актиноміцетів збирали з 4 мл культурального середовища за допомогою центрифугування (14000 т.о., 5 хв, 4 °С). Отриману біомасу промивали 1 мл ТЕ буферу (pH 7,5) і ресуспендували в 2 мл ТЕ буферу із додаванням 5 мг/мл лізоциму. Після 15 хв інкубування зразків з лізоцимом при 37 °С, РНК виділяли за допомогою модифікованого методу Кірбі [38]. Отримані зразки нуклеїнових кислот піддавали двом послідовним обробкам ДНКазою, використовуючи RNase-free DNase I (ThermoFisher Scientific). Щоб виключити можливість присутності залишкових кількостей ДНК, отримані після обробок ДНКазою зразки РНК перевіряли за допомогою ПЛР. кДНК синтезували із однакових вихідних кількостей РНК із використанням First Strand cDNA Synthesis Kit (ThermoFisher Scientific). Отриману кДНК використовували для ПЛР із застосуванням Phusion High-Fidelity PCR Master Mix (ThermoFisher Scientific).

Полімеразна ланцюгова реакція. У випадку ПЛР фрагментів, призначених для подальшого субклонування або секвенування,

використовували високоточні термостабільні ДНК-полімерази Q5 High-Fidelity DNA Polymerase (New England Biolabs) або Phusion High-Fidelity DNA Polymerase (ThermoFisher Scientific), згідно із рекомендаціями виробника стосовно складу реакційного суміші та умов реакції. Для аналітичної ПЛР, де не потрібна точність ампліфікації, використовували Taq DNA Polymerase (ThermoFisher Scientific), згідно із рекомендаціями виробника. Для кращого плавлення ГЦ-багатої хромосомної ДНК актиноміцетів у суміш ПЛР додавали диметилсульфоксид (ДМСО) (10%). Для ПЛР використовувалися термоциклери різних моделей.

Секвенування геномів актиноміцетів. Геноми *N. coxensis* і *S. cyanogenus* були секвеновані за допомогою комбінації технологій HiSeq Illumina і GridION ONT у подібний спосіб. Дані Illumina були отримані з архіву послідовностей рідів, а для ONT бібліотека секвенування була приготована з використанням Ligation Sequencing Kit (Oxford Nanopore Technologies) відповідно до рекомендацій виробника та просеквенована з використанням сиквенсера GridION R9.4.1 (Oxford Nanopore Technologies). Поклик основ в сирих даних було здійснено за допомогою GUPPY-FOR-GRIDION v3.0.6. Збірка геномів була здійснена у спосіб, опублікований раніше [50], використовуючи canu v.1.8. Збирання контигів здійснювалося за допомогою NEWBLER v2.8. Після ручного вирішення дрібних конфліктів з використанням CONSED [51], було отримано повні геноми.

Геном *N. coxensis* DSM 45129 складався з однієї кільцевої хромосоми розміром 9073954 п.н. (ГЦ-склад 72.12%). Анотація генома була здійснена за допомогою PROKKA v1.11 [52]. В геномі анотовано 8398 кодуючих послідовностей, 5 оперонів рРНК, 73 гени тРНК і 5 генів некодуючих РНК. Анотований геном та бібліотеки сирих рідів було депоновано в DDBJ/ENA/GenBank з ідентифікатором BioProject PRJNA693185.

Отримана послідовність генома *S. cyanogenus* була лінійною хромосомою розміром 8773899 п.н., на якій було анотовано 7700 генів, з яких: 7590 потенційно кодують білки, 91 ген тРНК, 1 ген тмРНК і 6 оперонів генів 16S-5S-23S рРНК. Усереднений ГЦ склад генома становить 71,8%. Цікаво, що геном

S. cyanogenus S136 мав декілька аномалій. Зокрема, кількість генів тРНК є помітно більшою, ніж зазвичай спостерігається в геномах стрептоміцетів: 91 проти, скажімо, 65 генів тРНК у модельного об'єкта генетики стрептоміцетів *S. coelicolor* (при майже однакових лінійних розмірах обох геномів) [53]. Другою аномалією виявилося те, що один із шести генів 16S рРНК (а саме *S1361_18950*) виявився значно дивергований від решти п'яти, які були ідентичні. Ідентичність нуклеотидної послідовності *S1361_18950* з іншими генами 16S рРНК становила 97,4%, що нижче, ніж поріг для диференціації окремих видів [54]. Вцілому, певна варіабельність генів 16S рРНК стрептоміцетів у межах одного генома описана. Однак, йшлося про від однієї до десяти нуклеотидних замін (*sic!*) [55], коли *S1361_18950* відрізняється від решти 16S рРНК на 29 нуклеотидів. Послідовність була депонована в GenBank під номером CP071839.

2.2.4 Біохімічні методи досліджень

Кількісний та якісний тест на β-глюкуронідазну активність. Якісно, β-глюкуронідазна активність в рекомбінантних штамах актиноміцетів виявлялася шляхом додавання 5-бромо-4-хлоро-3-індоліл-β-D-глюкуроніду (X-Gluc, ThermoFisher Scientific) в концентрації 25-100 мг/мл (розчин в ДМСО) або до середовища культивування, або безпосередньо на поверхню газона. β-глюкуронідазна активність проявлялася як хромогенна конверсія X-Gluc в забарвлений у зелений колір 5,5′-дибромо-4,4′-дихлоро-індиго. Для кількісного вимірювання, культури актиноміцетів спочатку вирощували у оптимальних рідких середовищах за бажаних умов. Отриманий у такий спосіб міцелій використовували для приготування безклітинних лізатів, відповідно до методики, описаної раніше [29]. β-глюкуронідазна активність вимірювалася у лізатах з використанням спектрофотометричного аналізу для виявлення перетворення безбарвного *p*-нітрофеніл-β-D-глюкуроніду на забарвлений *p*-нітрофенол при λ=415 у спосіб, описаний раніше [29, 42]. Для цього використовували Unicam UV 500 UV-Visible Spectrometer (Thermo) або Infinite

200 PRO microplate reader (Tecan). β-глюкуронідазна активність була нормалізована до ваги сухої біомаси.

Екстракція і аналіз ландоміцинів. Із культур, вирощених на чашках з агаризованим середовищем, ландоміцини екстрагували з 15 мл подрібненого агару за допомогою 10 мл етилацетату (Sigma). Отримані екстракти випаровували за допомогою вакуумного концентратора (Labconco), а сухий антибіотик розчиняли в 1 мл метанолу (Sigma). У випадку занурених культур, 10 мл зразка екстрагували за допомогою 10 мл етилацетату; з екстрактом поводилися так само, як описано попередньо. Концентрація ландоміцину визначалася за допомогою спектрофотометра Unicam UV 500 UV-Visible Spectrometer (Thermo) при λ=445. Отримані значення екстинкції перераховували в концентрацію за допомогою калібраційної кривої, створеної для різних концентрацій ландоміцину А. Останній було виділено з продукційних культур штамів дикого типу (S136) за методикою, описаною в [56]. Для тонкошарової хроматографії, метанольні розчини ландоміцинів наносили на пластинки для ТШХ разом із стандартним зразком чистого ландоміцину А. ТШХ проводили в системі: нерухома фаза – пластинки із сілікагелем F60254 (Merck), рухома фаза – хлороформ (Sigma):метанол (9:1), протягом 20 хв при 22 °С.

Екстракція та аналіз A40926 і A50926. Для кількісного вимірювання, A40926 і A50926 екстрагували з культур *Nonomuraea* spp. за допомогою аналогічного об’єму борного буферу (100 мМоль H_3BO_3 (Sigma), 100 мМоль NaOH (Sigma), pH 12). Отримані екстракти аналізували за допомого ВЕРХ (VWR Hitachi diode array L-2455 HPLC), детекція при λ=254 нм згідно із протоколом, описаним в літературі [45, 46]. Високороздільча мас-спектрометрія і фрагментаційний аналіз A40926 і A50926 здійснювали за допомогою мас-спектрометра SYNAPT G2-Si mass spectrometer з Acquity UPLC (Waters) [57].

Екстракція та аналіз люцензоміцину. Для екстракції люцензоміцину, 1 г вологої біомаси, отриманої з продукуючої культури, промивався тричі за допомогою деіонізованої води, а антибіотик екстрагували за допомогою 2 мл метанолу. Отриманий екстракт концентрували за допомогою вакуумного

концентратора. Метанольні розчини люцензоміцину аналізували за допомогою ВЕРХ-МС на HiRes Extractive Oribtrap MassSpectrometer (ThermoFisher Scientific). Використовували систему розчинників: вода + 0,1% HFo (фаза А); ацетонітрил + 0.1% HFo (фаза Б); в градієнті: 95% А до 0% А за 8 хв., потім 1,5 хв при 0% А, потім знову до 95% А за 1,5 хв.

Очищення глікопептидних антибіотиків за допомогою D-Ala-D-Ala-афінної хроматографії. ГПА було очищено за допомогою афінної хроматографії з використанням полімерної смоли з приєднаними залишками D-Ala-D-Ala (мішень ГПА) та хроматографа FPLC AKTA Start (GE Healthcare). Активація колонки HiTrap NHS-Activated HP affinity column (5 мл, GE Healthcare) та зв'язування ГПА здійснювали відповідно до описаного протоколу [58] з певними модифікаціями. Зокрема, смолу було активовано за допомогою 30 мл 1 мМоль HCl, після чого додавали 200 мМоль розчин D-Ala-D-Ala, в 5 мл зв'язувального буферу (0.2 М $NaHCO_3$, pH 7,0). Після 30 хв інкубування, смолу тричі промивали за допомогою 0,5 Моль етаноламін гідрохлориду та 0,5 Моль NaCl (pH 4,0, 30 мл), потім послідовно за допомогою 0,1 Моль ацетату натрію та 0,5 мМоль NaCl (pH 4,0, по 30 мл кожного). Нарешті, смолу промивали за допомогою 50 мл зв'язувального буферу і залишали на 1 год перед використанням. ГПА екстрагували у спосіб, описаний вище, pH в отриманих екстрактах доводили до 7,5 за допомогою HCl та використовували для афінної хроматографії. Для цього, екстракти фільтрували через 0,45 мкм фільтри та завантажували на D-Ala-D-Ala-активовану колонку зі швидкістю 0,5 мл/хв. Після декількох раундів промивання колонки із використанням зв'язувального буферу, зв'язаний ГПА елюювався за допомогою 0,1 Моль NaOH, а отриманий елюат ліофілізувався.

2.2.5 Біоінформатичні методи досліджень

В роботі використано широкий спектр серверів, програм та підходів біоінформатичного аналізу, що були специфічно обрані для конкретних цілей. Основні підходи, які застосовувалися універсально, подано нижче.

Рутинний аналіз амінокислотних та нуклеотидних послідовностей здійснювали в GENEIOUS v4.8.5.[59] Пошук гомологів білків в онлайн базах даних здійснювався за допомогою алгоритму Blastp (Altschul et al., 1990). Множинні вирівнювання амінокислотних та нуклеотидних послідовностей, підбір оптимальних моделей еволюцій та побудова філогенетичних дерев, зокрема, відбувалася з використанням пакету програм MEGA 6, X, XI [60]. Попарні та множинні вирівнювання амінокислотних та нуклеотидних послідовностей також виконували за допомогою Clustal Omega [61]. Для пошуку КБГ в секвенованих геномах актиноміцетів застосовували, зокрема, antiSMASH 6.0 [31]; анотація здійснювалася відповідно до записів бази даних MIBIG [3]. Визначення доменної організації нерибосомних пептидсинтаз (НРПС) і полікетидсинтаз (ПКС), а також передбачення специфічності відповідних аденілазних та кетосинтазних доменів здійснювали за допомогою PKS/NRPS Analysis [62]. MultiGeneBlast, http://multigeneblast.sourceforge.net [63] було застосовано для пошуку гомологів колокалізованих генів. Для ідентифікації структурних та функціональних доменів і мотивів в амінокислотних послідовностях білків використовували CD-Search [64]. Для створення структурних формул хімічних речовин використовували ChemDraw Ultra 12.0 (CambridgeSoft).

РОЗДІЛ 3 Плейотропні регулятори морфогенезу і вторинного метаболізму стрептоміцетів, що належать до *adpA*-опосередкованого каскаду і YtrA-підродини, як знаряддя для контролю продукції антибіотиків

У цьому розділі представлено результати вивчення компонентів *adpA*-опосередкованого регуляторного каскаду в стрептоміцетів з метою застосування генів цього каскаду як знарядь для регуляції біосинтезу різних антибіотиків. Ген *adpA* було вперше досліджено на моделі *S. griseus*, в якому цей ген кодує глобальний регулятор процесів морфологічної диференціації і продукції антибіотика стрептоміцину [65]. У *S. griseus* експресія *adpA* залежна від низькомолекулярної гормоноподібної сполуки – γ-бутиролактона (відомого як А-фактор). Біосинтез останнього розпочинається, коли бактерійна культура досягає певного етапу росту; далі, А-фактор зв'язує рецепторний білок ArpA, знімаючи його репресорний вплив з експресії *adpA* [66, 67]. Регулон AdpA є значним та включає від 100 до 500 генів (як було показано на моделях *S. griseus* і *S. coelicolor*) [28]. Окрім регуляції на рівні транскрипції гена, продукція білка AdpA регулюється і на трансляційному рівні. Це опосередковується трансляцією рідкісного UUA кодона, який знаходиться в мРНК *adpA*; UUA є найменш вживаним кодоном в транскриптомі стрептоміцетів (та й інших актиноміцетів) [68], в чиїх геномах переважають ГЦ-пари. Відповідно, ген *bldA* кодує єдину в стрептоміцетів лейцинову тРНК (тРНК$^{\text{ЛЕЙ}}_{\text{UAA}}$), що забезпечує трансляцію UUA [68]. Функціонування лейцинової тРНК, кодованої *bldA,* в свою чергу може регулюватися пострансляційними модифікаціями цієї тРНК [69]. Очевидно, що всі ці компоненти *adpA*-опосередкованого каскаду відіграють важливу роль в регуляції продукції антибіотиків. Також у цьому розділі ми висвітлюємо роль інших плейотропних регуляторів – малодослідженої групи бактерійних транскрипційних репресорів YtrA-підродини.

Філогеноміка AdpA-регуляторів. Хоча функції багатьох AdpA-регуляторів в стрептоміцетів вивчено, їх філогенія досі не вивчалася. У цій роботі ми відшукали ортологів AdpA у стрептоміцетів, геноми яких були доступні в базі даних GenBank на момент виконання роботи (2017 рік) – знайдено 390

послідовностей. Ми також ідентифікували найближчі гомологи AdpA серед представників інших рядів класу *Actinobacteria*. Ортологічна група транскрипційних регуляторів AdpA у стрептоміцетів належить до білків AraC/XylS-родини, для яких характерна двохдоменна організація у складі GATase-1 домена (глутамінової амідотрансферази) і HTH_18 ДНК-зв'язувального домена [28]. Рідкісний TTA кодон модельних *adpA* розміщений між ділянками, що кодують ці два домени [70]; таке положення TTA-кодона виявилося консервативним в генах ідентифікованої ортологічної групи AdpA. Однак, ми виявили також декілька винятків – *adpA*-генів в яких були додаткові TTA-кодони. Зокрема, в деяких ортологів *adpA* додатковий TTA-кодон розміщувався в ділянці, що кодує GATase-1 домен; тоді як в *Streptomyces* sp. HmicA12 додатковий TTA-кодон було знайдено в ділянці, що кодує ДНК-зв'язувальний домен, а в *S. neyagawaensis* SL-387 додатковий TTA-кодон розміщувався біля стоп-кодона.

Цікаво, що в деяких актинобактерій ми виявили гени-гомологи ділянок, що кодують окремі домени AdpA. У цих випадках гени, що кодують глутамінові амідотрансферази та HTH_18-подібні транскрипційні регулятори були послідовно ко-локалізовані. Подібна організація дає змогу припустити, що група регуляторів AraC/XylS-родини виникла в результаті злиття двох окремих генів, що кодували GATase-1 і HTH_18-ДНК зв'язувальні домени.

В двохдоменній архітектурі AdpA (див. вище) GATase-1 домен, вочевидь, потрібний для димеризації білка, бо AdpA зв'язує свій операторний сайт у вигляді гомодимеру (а сам сайт зв'язування є інвертованим повтором) [66]. Ми спекулятивно припустили, що ДНК-зв'язувальний домен AdpA здатний функціонувати без димеризаційного; відсутність потреби в димеризації, в свою чергу, може вести до розширення репертуару зв'язування такого окремого домена в геномі. Щоби перевірити наше припущення, ми клонували цілі послідовності генів, що кодують ортологи *adpA* із низки модельних стрептоміцетів, а саме: із *S. albus* (=*albidoflavus*) J1074 (*XNR_4181*), *S. ghanaensis* (=*viridosporus*) ATCC 14672 ($adpA_{gh}$), *S. clavuligerus* ATCC 7064 ($adpA_{scl}$), *S. coelicolor* A3(2) ($adpA_{sco}$). Також ми клонували відповідні сегменти

цих генів, які кодують ДНК-зв'язувальні домени (*XNR_4181_dbd, adpA$_{gh}$_dbd, adpA$_{scl}$_dbd, adpA$_{sco}$_dbd*). Всі перелічені вище послідовності було клоновано у кумат-індуцибельний вектор експресії pGCymRP21 [71] і перенесено в штам *S. coelicolor* A3(2) з нокаутом гена *adpA$_{sco}$* – *S. coelicolor* M851 [70]. Цей штам має кондиційний «лисий» (**b**a**ld**, Bld) фенотип, при якому на середовищі R5 він не здатний до формування повітряного міцелію і споруляції. Ми очікували, що створені конструкти зможуть комплементувати такий фенотип у *S. coelicolor* M851.

Геноми стрептоміцетів містять багато генів, що кодують AraC/XylS-подібні транскрипційні регулятори; однак, із невідомих причин, лише *adpA* володіє розширеним функціоналом. Очевидно, це є результатом певних особливостей його ДНК-зв'язувального домену. Щоб підтвердити унікальність *adpA,* ми також клонували у плазміду pGCymRP21 ген *XNR_1761*, що кодує найближчий AraC/XylS-регулятор-гомолог XNR_4181 у геномі *S. albus* (=*albidoflavus*) J1074. Відповідну плазміду, pGC1761, також перенесли в *S. coelicolor* M851. XNR_1761 також належить до HTH_18-надродини транскрипційних регуляторів, має таку саму доменну архітектуру, що й AdpA, і володіє ідентичністю амінокислотної послідовності з XNR_4181 на рівні 60%. Враховуючи таку подібність послідовності XNR_1761 з AdpA, було цікаво встановити, чи здатний *XNR_1761* комплементувати Bld фенотип *S. coelicolor* M851.

Морфологія отриманих рекомбінантів була вивчена на середовищі R5. Штами M851 pGCymRP21^{+} не спорулювали (ознака відсутності гену *adpA*) на багатому середовищі R5 наприкінці 72 год культивування, тоді як дикий тип – *S. coelicolor* M145 – був здатний до споруляції. Введення повних генів *XNR_4181, adpA$_{gh}$, adpA$_{scl}$* і *adpA$_{sco}$* відновлювало утворення спор, вказуючи на те, що всі ці регулятори (будучи близькими гомологами) мають когерентні властивості. Введення *XNR_4181_dbd, adpA$_{gh}$_dbd, adpA$_{scl}$_dbd* і *adpA$_{sco}$_dbd* також відновлювало споруляцію M851. У свою чергу, введення *XNR_1761* не змогло відновити утворення спор в M851.

Отже, наша вихідна гіпотеза підтвердилася: наявність GATase-1 димеризаційного домена не є обов'язкова для функціонування AdpA, тоді як його найближчі гомологи не здатні комплементувати втрату цього гена. Отримані дані свідчать і на користь висунутого припущення про еволюційне походження AdpA-подібних регуляторів.

***Елементи adpA-опосередкованого регуляторного каскаду S. albus* (=*albidoflavus*) J1074.** J1074 відрізняється від інших стрептоміцетів доступністю для генно-інженерних маніпуляцій і природно-мінімізованим геномом (прибл. 7 млн.п.н.) [72]. Закономірно, J1074 розглядається як потенційний господар для гетерологічної експресії КБГ різних антибіотиків [73]. Варто відзначити, що описане тут дослідження здійснювалося на похідному J1074 – штамі SAM2 – який відрізняється від дикого типу видаленням псевдо-*attB*-сайту інтеграції φC31 [74].

У світлі того, що експресія більшості КБГ антибіотиків, які походять із стрептоміцетів, є *adpA*-опосередкована і лімітується трансляцією UUA кодонів, ми вирішили вивчити властивості гена *bldA* із SAM2 методом нокауту. Виявилося, що на відміну від інших модельних об'єктів [68], відсутність *bldA* не призводить в SAM2 до Bld фенотипу. Однак, Δ*bldA* штам відрізнявся від дикого типу за морфологією на ранніх етапах росту на деяких середовищах, а саме на агаризованих середовищах SFM і SMMS. Тут, після 72 год культивування мутант формував рідкі гіфи повітряного міцелію і нерегулярні спори, тоді як штам дикого типу бурхливо спорулював у цій часовій точці. Натомість, відсутність *bldA* призвела до зупинки продукції пауломіцину, антибіотика активного проти Грам-позитивних бактерій [75]; очевидно, причиною цього є наявність TTA-кодону в гені (*XNR_0585*) потенційного шлях-специфічного LuxR-подібного регулятора біосинтезу пауломіцину.

Для того, щоб досягти функціонального стану тРНК бактерій, а зокрема і стрептоміцетів, піддаються різноманітним посттрансляційним модифікаціям [69]. Перші кроки модифікації таких тРНК як $тРНК_{XXA}$, до яких належить і $тРНК^{ЛЕЙ}_{UAA}$ (кодована геном *bldA*), є найважливішими і включають модифікацію залишку аденіну в положенні 37 (A37), що розмішений в

антикодоновій петлі на 3’-кінці антикодона XXA (34-36) [76]. Дві послідовні реакції модифікації виконуються ферментами MiaA і MiaB. A37-N6-диметилалілтрансфераза MiaA каталізує утворення N6-ізопентеніл аденозину [77], а N6-ізопентеніл-A37-C2-метилтіотрансфераза MiaB забезпечує утворення 2-метилтіо-N6-ізопентеніладенозину [78]. Геном J1074 містить гени, які кодують MiaA і MiaB – *XNR_1074* (*miaA*) і *XNR_1078* (*miaB*), відповідно [79], даючи змогу припустити, що тРНК$^{\text{ЛЕЙ}}_{\text{UAA}}$ (продукт *bldA*) піддається цим модифікаціям. Щоб перевірити це припущенні, ми здійснили нокаути обох генів; ВЕРХ аналіз екстрактів тРНК мутантів продемонстрував, що припущення про функції *XNR_1074* і *XNR_1078* виявилося правильним.

Більш цікавим було вивчення впливу нокаутів генів *miaA* і *miaB* на морфологічну диференціацію та продукцію антибіотиків у SAM2. Зокрема, фенотип Δ*miaA* мутанта нагадував фенотип Δ*bldA* мутанта, описаний вище. Натомість, морфологія Δ*miaB* мутанта не відрізнялася від штаму дикого типу за умов вирощування на різних агаризованих середовищах. Подібною була і ситуація із продукцією пауломіцину: як і Δ*bldA*, Δ*miaA* мутант майже не продукував цього антибіотика, тоді як продукція пауломіцину в Δ*miaB* не відрізнялася від такої у дикого типу. Мабуть, спостережені ефекти спричинені тим, що модифікація тРНК$^{\text{ЛЕЙ}}_{\text{UAA}}$, яку забезпечує MiaA, є критичною для правильного розпізнавання UUA-кодону цією тРНК, і, відповідно, трансляції TTA-вмісних генів. Для того, щоб перевірити це припущення, ми створили плазміди pSAGA і pSAGAt. В обох плазмід два алелі гена β-глюкуронідази *gusA* [42] знаходяться під контролем сильного конститутивного промотора *aac(3)IVp* (промотор гена стійкості до апраміцину [38]). При цьому, у плазміді pSAGA використано вихідний ген *gusA*, в якому є лейциновий кодон CTG, розміщений відразу після старт-кодону. Натомість, в pSAGAt ми замінили CTG кодон на TTA, що не змінює амінокислотну послідовність білка, але впливає на трансляційну динаміку створеного алеля. Отримані плазміди перенесли в SAM2 і Δ*miaA*. Порівнявши β-глюкуронідазну активність отриманих рекомбінантних штамів, ми виявили, що вона порівняння у SAM2 pSAGA^{+}, SAM2 pSAGAt^{+} і Δ*miaA* pSAGAt^{+}, але практично відсутня в Δ*miaA* pSAGAt^{+}. Розвиваючи цю

лінію досліджень, ми спробували комплементувати фенотипи Δ*miaA* за допомогою надекспресії алеля *adpA,* в якому TTA-кодон було замінено на CTC; виявилося, що експресія такого алеля в Δ*miaA* дійсно позитивно впливає на морфогенез в останнього мутанта. Ці дані засвідчують, що вихідне припущення є, скоріше, правильним, а немодифікована тРНК$^{\text{ЛЕЙ}}_{\text{UAA}}$ не здатна до ефективної трансляції UUA-кодону.

В підсумку, ми вивчили властивості генів *adpA* і *bldA S. albus* (=*albidoflavus*) J1074 і продемонстрували їх важливість для продукції антибіотиків цим штамом. Проте, *adpA* і *bldA*, схоже, не відіграють вирішальної ролі для морфогенезу в J1074, на відміну від інших модельних стрептоміцетів. Ми також описали новий компонент *adpA*-опосередкованого регуляторного каскаду J1074, продемонструвавши регуляторний вплив гена *miaA*, який кодує фермент модифікації тРНК$^{\text{ЛЕЙ}}_{\text{UAA}}$.

Філогенія та властивості YtrA-подібних регуляторів. Важливою родиною плейотропних регуляторів продукції антибіотиків і морфогенезу стрептоміцетів є родина GntR [80]. Подібно до архітектури AdpA (див. вище), білки цієї родини складаються з N-термінального ДНК-зв'язувального HTH-домена і С-термінального олігомеризаційного домена. GntR-родина поділяється на сім підродин, члени яких відрізняються за структурою С-термінального домена [80]. Серед цих підродин є YtrA-підродина, білки якої мають найкоротший олігомеризаційний домен, що складається лише із двох α-спіралей; регулятори цієї підродини мало вивчені для актиноміцетів. Представники YtrA-підродини зацікавили нас, бо один із них – LndYR – є залученим в регуляції біосинтезу полікетидного антибіотика ландоміцину Е в *S. globisporus* 1912 [81]. Тому, ми вирішили вивчити особливості розповсюдження та філогенії YtrA-регуляторів в актинобактерій, а також їх властивості у модельного стрептоміцета *S. coelicolor* M145.

Використовуючи *ytrA Bacillus subtilis* як запит, ми проаналізували всі доступні в базі даних GenBank на момент виконання роботи (2016 рік) геноми актинобактерій на наявність гомологів. Всього ми виявили 17 гомологів в 6 геномах підкласу *Coriobacteridae,* 3 гомолога в 2 геномах підкласу

Rubrobacteridae і 76 гомологів в 80 геномах підкласу *Actinobacteridae.* Паразитичні актинобактерії з малими геномами не містили гомологів *ytrA,* як і високоспеціалізовані представники підкласів *Acidimicrobidae* і *Nitriliruptoridae.* Вивчення геномного контексту та філогенії знайдених генів показало, що ці гени найчастіше колокалізовані з генами ABC-транспортерів, і поділяються на 7 клад. В модельного штаму *S. coelicolor* було знайдено три гени YtrA-регуляторів, а саме *SCO0823, SCO1728* і *SCO3812*, що належали до різних клад. Дослідивши мутанти *S. coelicolor* за генами *SCO0823, SCO1728* і *SCO3812,* ми не виявили радикального впливу відсутності цих генів на продукцію антибіотиків та морфогенез, лише певні маргінальні зміни. Це дає нам змогу припустити, що YtrA-регулятори стрептоміцетів забезпечують адаптивну мінливість організмів до невстановлених чинників зовнішнього середовища, але не відіграють ключової ролі.

У підсумку, нам вдалося оновити уявлення про розповсюдження і еволюцію AdpA-подібних білків; ми дослідили *adpA*-опосередкований каскад в *S. albus (=albidoflavus)* J1074, описавши нові його компоненти, а саме гени ферментів модифікації тРНК, їх роль в продукції антибіотиків та морфологічній диференціації. Точну роль YtrA-подібних регуляторів за наших лабораторних умов визначити не вдалося, хоча і делеції їх генів у *S. coelicolor* мали слабкий вплив на продукцію антибіотиків і морфологічну диференціацію; ця група транскрипційних регуляторів заслуговує подальшого вивчення.

Основні результати, отримані у цьому розділі, опубліковано в [82–85].

Journal of Molecular Evolution
https://doi.org/10.1007/s00239-018-9834-z

ORIGINAL ARTICLE

Genomic Insights into Evolution of AdpA Family Master Regulators of Morphological Differentiation and Secondary Metabolism in *Streptomyces*

Mariia Rabyk[1] · Oleksandr Yushchuk[1] · Ihor Rokytskyy[1] · Maria Anisimova[2,3] · Bohdan Ostash[1]

Received: 6 October 2017 / Accepted: 9 March 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract
The AdpA protein from a streptomycin producer *Streptomyces griseus* is a founding member of the AdpA family of pleiotropic regulators, known to be ubiquitously present in streptomycetes. Functional genomic approaches revealed a huge number of AdpA targets, leading to the claim that the AdpA regulon is the largest one in bacteria. The expression of *adpA* is limited at the level of translation of the rare leucyl UUA codon. All known properties of AdpA regulators were discovered on a few streptomycete strains. There are open questions about the true abundance and diversity of AdpA across actinobacterial taxa (and beyond) and about the possible evolutionary forces that shape the AdpA orthologous group in *Streptomyces*. Here we show that, with respect to the TTA codon, streptomycete *adpA* is more diverse than has been previously thought, as the genes differ in presence/position of this codon. Reciprocal best hits to AdpA can be found in many actinobacterial orders, with a domain organization resembling that of the prototypical AdpA, but other configurations also exist. Diversifying positive selection was detected within the DNA-binding (AraC) domain in *adpA* of *Streptomyces* origin, most likely affecting residues enabling AdpA to recognize a degenerate operator. Sequence coding for putative glutamine amidotransferase (GATase-1) domain also shows signs of positive selection. The two-domain organization of AdpA most likely arose from a fusion of genes encoding separate GATase-1 and AraC domains. Indeed, we show that the AraC domain retains a biological function in the absence of the GATase-1 part. We suggest that acquisition of the regulatory role by TTA codon is a relatively recent event in the evolution of AdpA, which coincided with the rise of the *Streptomycetales* clade and, at present, is under relaxed selective constraints. Further experimental scrutiny of our findings is invited, which should provide new insights into the evolution and prospects for engineering of an AdpA-centered regulatory network.

Keywords AdpA · Rare codons · Actinobacteria · Positive and negative selection · Evolutionary analysis

Introduction

The *adpA* gene was first shown in *Streptomyces griseus* to be a pleiotropic regulator influencing processes as diverse as morphological differentiation and streptomycin production (Ohnishi et al. 1999). In *S. griseus* the expression of *adpA* is dependent on small signaling molecule, γ-butyrolactone, referred to as A-factor. The latter is produced in a growth-dependent manner and turns on the transcription of *adpA* by binding to the A-factor receptor protein (ArpA), which has bound the *adpA* promoter, and dissociating the ArpA from the DNA (Yamazaki et al. 2004; Ohnishi et al. 2005; Horinouchi 2007). In *S. coelicolor adpA* expression it has been shown to be controlled by *cis*-antisense RNA (Šetinová et al. 2018). Recently the AdpA regulon was demonstrated to encompass from 100 to 500 genes, depending

Mariia Rabyk and Oleksandr Yushchuk have contributed equally to the work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00239-018-9834-z) contains supplementary material, which is available to authorized users.

✉ Bohdan Ostash
bohdan.ostash@lnu.edu.ua

1 Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, Hrushevskoho St. 4, Lviv 79005, Ukraine

2 Institute of Applied Simulations, School of Life Sciences and Facility Management, Zürich University of Applied Sciences, Einsiedlerstrasse 31a, 8820 Wädenswil, Switzerland

3 Swiss Institute of Bioinformatics, Lausanne, Switzerland

Published online: 13 March 2018

on streptomycete species. (Higo et al. 2011, 2012; Guyet et al. 2014). AdpA was also shown to regulate replication (Wolański et al. 2011, 2014). The *adpA* deletion in streptomycetes affects the aerial mycelium formation and, often, antibiotic production as well (Ohnishi et al. 1999; Nguyen et al. 2003; Takano et al. 2003; Zhao et al. 2007; Pan et al. 2009; Lopez-García et al. 2010; Xu et al. 2017). All available experimental and genomic data attest to the focal role of AdpA in *Streptomyces* biology.

In addition to being controlled by the hormone-like molecule, asRNA and the repressor protein ArpA, the AdpA expression is subject to a remarkable regulation at the level of translation of UUA codon within the *adpA* transcript. The UUA codon is the rarest one within GC-rich genomes of *Actinobacteria* (Chater and Chandra 2008; Zaburannyy et al. 2009). Gene *bldA* encodes the only $tRNA^{Leu}_{UAA}$ that is able to read the UUA codon efficiently, and its delayed occurrence during a cell cycle limits the expression of UUA-containing genes (Lawlor et al. 1987; Chater and Chandra 2008; Hackl and Bechthold 2015). In the latter, TTA codon(s) are usually located close to a start codon of a gene. It is thought to lower the cost of abortive translation until the $tRNA^{Leu}_{UAA}$ concentration in cells exceeds a certain threshold level (Chater and Chandra 2008). So far, a biased intragenic position of TTA was documented for *Streptomyces* genes (Ventura et al. 2007). The TTA-mediated regulation of gene expression is well documented in a number of streptomycetes, but the role for this codon has not been investigated beyond the genus *Streptomyces*. Only recently, the TTA codon-containing genes in *Frankia* were analyzed (Sen et al. 2012). *Frankia* genomes were shown to contain a higher number of TTA^+ genes compared to most of the other actinomycetes. However, this codon does not seem to play a regulatory role in *Frankia* (Sen et al. 2012).

Aforementioned observations suggest that the AdpA protein and its codon sequence have undergone adaptive evolution in *Streptomyces*, although no attempts were undertaken to address this question directly. Our understanding of the abundance, structural, and functional diversity of members of the AdpA orthologous family within and beyond *Streptomyces* is currently based on rather modest number of genomes (Chandra and Chater 2014), as compared to what is currently available. The present work is an attempt to address aforementioned questions via a bioinformatics analysis of available genomic data. We demonstrate that the AdpA orthologous group is diverse (regarding its genetic architecture) and abundant within *Actinobacteria*; significant structural diversity was found even for the best studied (and previously thought to be homogeneous) group of streptomycete *adpA* sequences. Several AdpA-like sequences were identified beyond actinobacterial taxa, hinting at the ancient evolutionary origin of this fascinating group of regulators. Here, for the first time, through the analysis of AdpA orthologous group of *Streptomyces* origin we present evidence that certain sites within DNA-binding helix-turn-helix motif are subject to positive selection. In contrast to previous reports, we revealed no signals of selection on TTA codon. Finally, we suggest that a typical two-domain AdpA arose as a result of a fusion of genes that originally encoded these domains.

Materials and Methods

Gene Sequences

A set of all probable AdpA orthologs was collected from GenBank using a reciprocal best hit (RBH) strategy based on an *E* value threshold of 10 and BLOSUM62 similarity matrix (Moreno-Hagelsieb and Latimer 2008). *S. griseus* $AdpA_{gr}$ (Ohnishi et al. 2005), *S. coelicolor* $AdpA_{sc}$ (Wolański et al. 2011), and *S. clavuligerus* $AdpA_{cl}$ (Lopez-García et al. 2010) were initially tested as BLASTP queries. When necessary, manual annotation of genes was conducted by examining the database hits of BLASTP with KEGG (Tanabe and Kanehisa 2012) and CDD (Marchler-Bauer et al. 2011) databases, and the results of FRAMEPLOT analysis (Ishikawa and Hotta 1999).

The codon sequences of identified RBHs were aligned using the ProGraphMSA program (Szalkowski 2012) with options -m --codon -f -w -F, for amino acid alignment options were -f -m -A. The codon alignment was further refined by deleting columns in alignment where gaps constituted more than 90% of the content. Alignments of genes that contain 'classical' TTA codons and those which do not were extracted from the initial alignment.

Phylogenetic Analysis

A phylogenetic tree was inferred by maximum likelihood using CodonPhyML (Gil et al. 2013). The following models were used and compared: codon models M0 and M0+Γ (with among-site rate variation), and amino acid model LG+Γ+F (Le and Gascuel 2008) with amino acid frequencies estimated empirically from data. Model M0 assumes constant selection pressure on the protein across the sites, as measured by $\omega = dN/dS$ ratio of non-synonymous to synonymous substitution rates (Goldman and Yang 1994; Nielsen and Yang 1998). When sequences evolve neutrally, under no selection, it is considered that $\omega = 1$, whereas $\omega < 1$ indicates purifying selection, and $\omega > 1$ positive selection.

Selection Pressure Analysis

To estimate ω, random-sites models (REL) which allow for across-site variation in dN and dS (Pond and Muse

2005) were applied using the HyPhy software package (Kosakovsky Pond et al. 2005) (version 2.2.4). The phylogenetic trees constructed under the best fitting model in CodonPhyML analysis were used for the analysis of selection pressure. Most REL models pre-specify a number of ω classes and fit the values for each class as well as the fraction of sites belonging to each class (Nielsen and Yang 1998; Yang et al. 2000; Kosakovsky Pond and Frost 2005). Probabilities of each individual site to fall into one of the classes are inferred via an empirical Bayes approach. Traditional codon models with ω parameter estimate dN and dS as a ratio and assume a constant dS (Nielsen and Yang 1998; Suzuki and Gojobori 1999; Yang et al. 2000; Yang and Bielawski 2000). However, in some cases, synonymous substitutions could be under selection (Sharp et al. 1995). To account for the effect of variation of synonymous substitution rates across sites, we performed codon selection analysis using the dNdSRateAnalysis program from HyPhy. The 'Nonsynonymous' model assumes a constant dS (= 1), but samples dN values from a given rate distribution. The 'Dual' model estimates dS and dN independently, sampling both from a given rate distribution (Pond and Muse 2005). In this study, both models were fitted using the MG94 × REV core rate matrix with rates from the general discrete distribution (GDD) and three classes. All models were fitted several times using randomized initial values to find the global optimum. To test for site-to-site variation in dS across both AdpA datasets, Dual and Non-synonymous models were compared using a likelihood ratio test (LRT).

To detect amino acid sites under positive selection, we used the Bayesian approach (Pond and Muse 2005) as implemented in dNdSResultProcessor of the HyPhy software. We selected every site that had a Bayes factor (BF) for positive selection greater than 100 (Kass and Raftery 1995).

Plasmid Construction and Complementation Studies for Functional Experiments

The coding sequence of *Streptomyces albus* J1074 *adpA* orthologue (*XNR_4181*) together with its putative ribosome-binding site was amplified using primers xnr4181_XbaIup and xnr4181_EcoRIrp yielding 1250 bp amplicon (these and other primer sequences mentioned throughout the work are listed in Table 1). This PCR product was digested with restriction endonucleases XbaI and EcoRI, and cloned into respective sites of integrative, cumate-inducible expression vector pGCymRP21 (Horbal et al. 2014) to give pGC4181. The *adpA* genes from *S. coelicolor* M145 ($adpA_{sc}$), *S. clavuligerus* ATCC27064 ($adpA_{cl}$), *S. ghanaensis* ATCC14672 ($adpA_{gh}$), as well as *XNR_1761* gene, were cloned likewise using appropriate primer pairs listed in Table 1, thus leading to plasmids pGCSCO, pGCSCLA, pOOB103c, and pGC1761, respectively. Part of the *XNR_4181* encoding the AraC–HTH DNA-binding domain (700–1245 bp) was amplified with xnr4181araC_up and xnr4181_EcoRIrp primers yielding 601 bp amplicon. The forward primer contained artificially introduced start codon and the ribosome-binding site. The PCR product was digested with XbaI and EcoRI, and cloned into respective sites of pGCymRP21

Table 1 Primer sequences used in this work

Name	Sequence, 5′→3′	Purpose
xnr4181_XbaIup	AAATCTAGAGGGGGGCTTAGTCACATG	*XNR_4181* cloning
xnr4181araC_up	AAATCTAGACGAGCAAC*GGAGGT*ACG GACA**TG**CCGGAGGAGATCGGCGCCGAC	*XNR_4181_dbd, adpAsco_dbd, adpAgh_dbd* cloning
xnr4181_EcoRIrp	AAAGAATTCGGAGCTGTCCTCTCTCAGAC	*XNR_4181*/*XNR_4181_dbd* cloning
SCOADPA_F	AAATCTAGAGGGGGGCTTAGCAGTATG	*adpAsco* cloning
SCOADPA_R	AAAGAATTCGCCGTCTGCTCACCTCACGG	*adpAsco*/*adpAsco_dbd* cloning
SCLAVADPA_F	AAATCTAGAGGGGGGCTTAGCGTCATG	*adpAscla* cloning
SCLAd_F	AAATCTAGACGAGCAAC*GGAGGT*ACG GACA**TG**CCAGAGGAAATCGGGTCGGAC	*adpAscla_dbd* cloning
SCLAVADPA_R	AAAGAATTCCATGCGACTACCTTATGG	*adpAscla*/*adpAscla_dbd* cloning
adpA_gh_upKpnIXbaI	AAAGGTACCTCTAGAAACGTGCAGGAGCGACACAAC	*adpAgh* cloning
adpA_gh_revXbaIEcoRI	AAATCTAGAATTCTACGTCACGGCGCGCTGCG	*adpAgh*/*adpAgh_dbd* cloning
XNR_1761_F	AAATCTAGAGACGACAGGACTGACCCCATG	*XNR_1761* cloning
XNR_1761_R	AAAGAATTCGCACCTAATGCCCCCTCA	*XNR_1761* cloning
amR_F	CGGGGTCTGACGCTCAGTGGA	*aac3(IV)* amplification
amR_R	AGCGTCTGCTCCGCCATTCG	*aac3(IV) amplification*

The recognition sites of restriction endonucleases are underlined, the artificially introduced RBSs are in italics and the artificially introduced start codons are in bold

giving pGC4181d. The same approach was used to clone coding sequences of DNA-binding domains of $adpA_{sc}$, $adpA_{cl}$, and $adpA_{gh}$, generating pGCSCOd, pGCSCLAd, pGCSGHd.

All plasmids (together with pGCymRP21) were transferred conjugally from *E. coli* ET12567 (pUZ8002) into *S. coelicolor* M851 (Δ*adpA* mutant; Takano et al. 2003), as described elsewhere (Ostash et al. 2015). Apramycin-resistant transconjugants were verified with PCR, using primers specific to the vector marker gene *aac3(IV)*. The transconjugants and control strains (M851 pGCymRP21$^+$, its initial M145 strain) were grown for 3–5 days on R5 agar plates (Kieser et al. 2000) supplemented with 50 μM cumate to induce the expression of the cloned genes.

Results and Discussion

Abundance and Diversity of AdpA Orthologous Family Within Phylum *Actinobacteria*

We explored amino acid sequences of three experimentally studied AdpA proteins—$AdpA_{gr}$, $AdpA_{sc}$, and $AdpA_{cl}$—as BLASTP queries to identify putative orthologs across *Actinobacteria*. Our initial searches (in 2013) yielded 323 unique protein sequences in total. The number of reciprocal best hits (RBHs) retrieved for each of the three queries ranged from 114 to 138 and to 323 (Fig. 1), respectively, and included representatives of most actinobacterial orders (*Streptomycetales, Corynebacteriales, Bifidobacteriales* etc) (Fig. S1, Electronic Supplementary Materials; ESM). For example, $AdpA_{cl}$ has led to 177 RBHs unidentified when using either $AdpA_{gr}$ or $AdpA_{sc}$ as queries. (Here and further down in this paper we label the 323 AdpA RBHs identified by $AdpA_{cl}$ as "initial dataset").

As mentioned above, AdpA consists of ThiJ/PfpI/DJ-1-like (also known as GATase-1) dimerization domain and an AraC/XylS-type DNA-binding domain at its N- and C-termini, respectively (Ohnishi et al. 2005). Searching across the *Actinobacteria* class, we discovered that certain species have no full-length AdpA RBHs. Rather, orthology was revealed between the query and proteins featuring either single AraC or GATase-1 domains. This prompted us to analyze a more extensive set of actinobacterial genomes in order to study a diversity of domain organization of AdpA regulators within the class. As a source of genomes we used the PATRIC database (Wattam et al. 2014) and the full-length $AdpA_{cl}$, $GATase\text{-}1_{cl}$, or $AraC_{cl}$ domains of $AdpA_{cl}$ as the queries. Our updated search yielded 1394 unique RBHs, which spanned 17 orders of class *Actinobacteria* class (Fig. 2). With respect to available genome sequences, streptomycetes constitute the most deeply covered single genus within the phylum

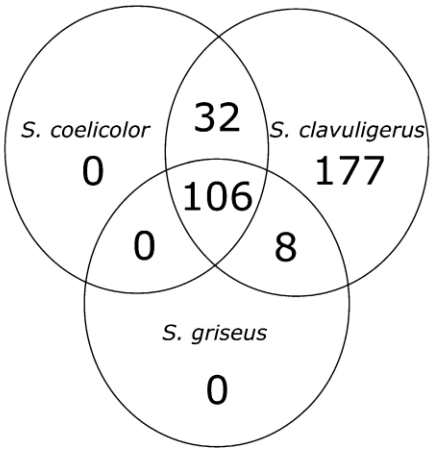


Fig. 1 Venn diagram for orthologs detected against GenBank database using AdpA queries from three *Streptomyces* species shown on the figure. BLAST searches were carried out against GenBank for all proteins of 323 species using BLOSUM62 matrices at default *E* value cut-off

Actinobacteria, with the largest (390) set of probable AdpA orthologs (Fig. 2).

In terms of domain organization, the most diverse AdpA RBHs were revealed in orders *Corynebacteriales, Bifidobacteriales,* and *Micrococcales* (Fig. 2). For example, in *Corynebacteriales*, the 3-Mbp genome of *Corynebacterium freneyi DNF00450* carries one gene *HMPREF1650_RS07820*-encoding ThiJ/PfpI protein orthologous to $AdpA_{cl}$/$GATase\text{-}1_{cl}$, and no genes for AraC family regulators. *Corynebacterium humireducens* NBRC 106098 (2.68 Mbp) encodes only three AraC-family transcriptional regulators (one of them—AdpA RBH) and no GATase-1 proteins. Other orders usually possessed AdpA RBHs having typical two-domain architecture. Our results suggest that *adpA* could originate from the gene fusion. Indeed, pairs of adjacent genes encoding ThiJ- and AraC domain-containing proteins are found in numerous bacterial genomes. Particularly, such gene pairs (paralogous to *adpA*) were found in *S. coelicolor* (SCO6746/6747), *S. griseus* (SGR_981/982), *Saccharopolyspora erythraea* (SACE_6420/SACE_6421, SACE_4268/4269), *Actinosynnema mirum* (Amir_4315/4316), *Salinispora arenicola* (Sare_0089/0090), *Kribbella flavida* (Kfla_5799/Kfla_5798), and even in chloroflexi bacterium *Herpetosiphon aurantiacus* (Haur_2399/Haur_2400).

Streptomycete *adpA* genes fell into several groups with regard to the position of TTA codon (Fig. 3). The absolute majority had TTA codon in the 'classical' position, as it is in *S. griseus adpA* gene, between GATase-1 and HTH domains.

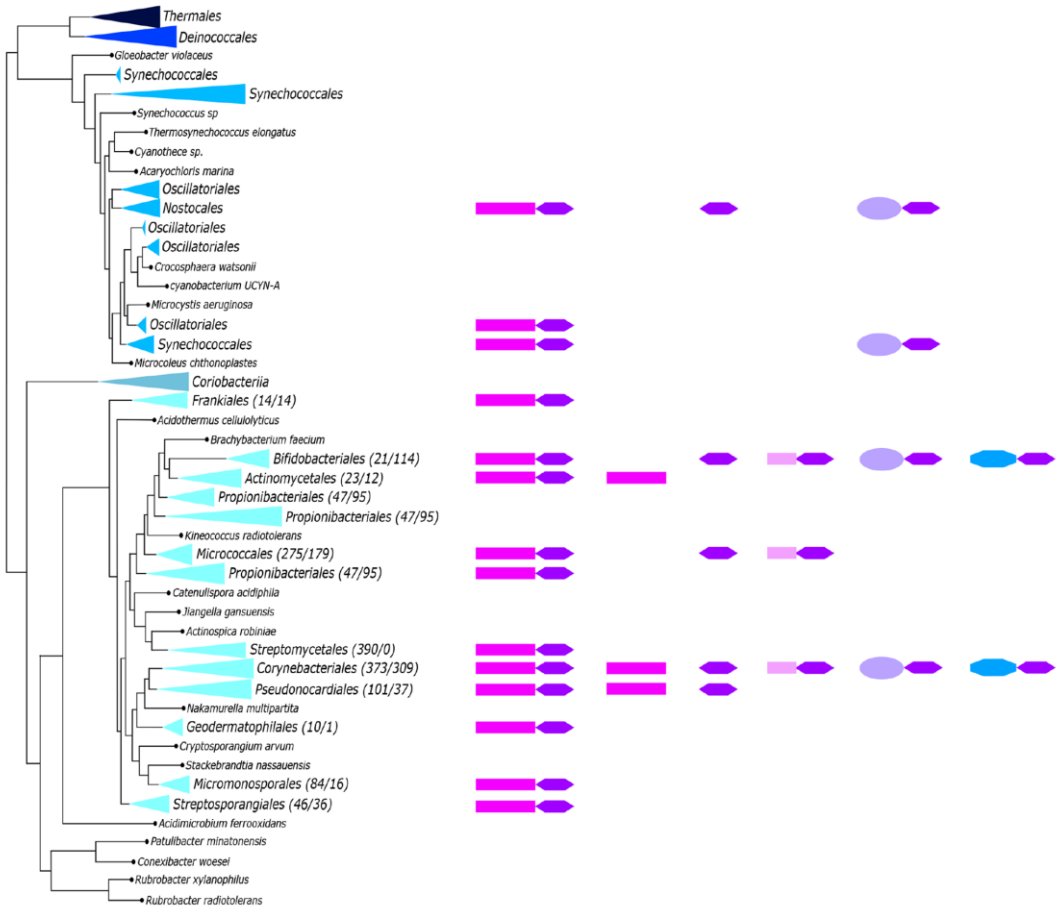


Fig. 2 The unrooted tree of *Actinobacteria* and *Cyanobacteria* phyla reconstructed using PhyloPhlAn phylogenetic pipeline. Collapsed clades (at the level of orders) are color-highlighted and labeled with the order name. The latter is followed by round brackets showing two numbers: first one is for quantity of retrieved AdpA orthologs, and second one—number of genomes where no $AdpA_{cl}$ ortholog was found. The types of domain organization of AdpA orthologs are shown next to corresponding orders. Pink rectangular denotes GATse domain; violet hexagon—AraC HTH domain; light pink rectangular—truncated GATse domain; blue octagon—arabinose-binding domain of AraC transcription regulator; gray oval—domain of unknown function. (Color figure online)

In a few *adpA* orthologs TTA codon was found within the GATase-1 domain, closer to the start of the gene; in *S.* sp. HmicA12 TTA was near the beginning of AraC domain; and in *S. neyagawaensis* TTA codon was located close to the stop codon of *adpA*. Finally, we identified *adpA* genes (from *S.* sp. *C, S.* sp. *Mg1, S.* sp. *ScaeMP-e10*, and *S. lysosuperificus* ATCC 31396) that have no TTA at all.

An average actinobacterial AdpA protein had the length of 320–345 aa. The largest protein of 705 aa was from *Kitasatospora cheerisanensis* KCTC 2395; the increased size of the protein appears to result from incorporation of low complexity sequences of unknown function at the C- and N-termini (Fig. S3, ESM). As compared to the other actinobacterial AdpA RBHs, streptomycete AdpA proteins had a longer C-terminal part, and its length was within the 252–496 aa range.

In our first dataset we revealed probable AdpA orthologs in 27 mycobacterial species. However, *Mycobacterium tuberculosis* H37Rv appeared to lack full-length two-domain AdpA counterpart. The only strain from *Mycobacterium tuberculosis* complex that returned AdpA RBH—MarA2—was *M. tuberculosis* 401416. Using the latter protein as a query against H37Rv proteome, protein Rv1931c (AdpA RBH) was identified as

```
neyagawaensis        --------------------MLKNVAAVLLDNVHPFELGVVCEVFGIDRSEEGLPAYDFA
sp._HmicA12          MSHDSTTAPEAAARKLSGRRRREIVAVLLFSGGPIFESSIPLSVFGIDRQDAGVPRYRLL
somaliensis          MSQDS--AAPETGRRLSGRRRREVVAVMLFSGGPIFESSIPLSVFGVDRQDAGVPRYRLL
griseus              MSQDSA-AATEAARKLTGRRRREVVAVLLFSGGPIFESSIPLSVFGIDRQDAGVPRYRLL
lysosuperificus      MSQDSTAVVADAGRKLAGRRRREIVAVLLFSGGPIFESSIPLSVFGIDRQDAGVPRYRLL
                                         : **.:*:..   ** .:  .***:**.: *:* * :

neyagawaensis        LVAAEASSVRAVPGFTISTPYGLDRLEEADLIAIPVGDDFHTRDFPPALLEALRQAVARG
sp._HmicA12          VCAGEEGPLRTTGGLELTAPHGLEAISRAGTVVVPAWRSIT-SPPPEEALDAIRRAHEEG
somaliensis          VCAGEEGPLRTTGGLELSAPYGLEAISRAGTVVVPAWRSIT-SPPPAEALDALRRAHEEG
griseus              VCGGEEGPLRTTGGLELTAPYGLEAISRAGTVVVPAWRSIT-SPPPAEALDALRRAHEEG
lysosuperificus      VCAGEDGPLRTTGGLELTAPYGLEAIARAGTVVVPAWRSIT-SPPPPEALDALRLAHEEG
                     : ..* . :*:. *: :::*:**: : .*. :.:*.  .:     *   *:*:* *  .*

neyagawaensis        ARVVSVCVGAFVLGAAGLLDGRRCTTHWRYAAALTERYPKAQVEPGVLYVDEDPILTSAG
sp._HmicA12          ARIVGLCTGAFVLAAAGLLDGRPATTHWMYAPTLAKRYPSVHVDPRELFVDDGDVLTSAG
somaliensis          ARIVGLCTGAFVLA-AGLLDGRPATTHWMYAPTLAKRYPSVHVDPRELFVDDGDVLTSAG
griseus              ARIVGLCTGAFVLAAAGLLDGRPATTHWMYAPTLAKRYPSVHVDPRELFVDDGDVLTSAG
lysosuperificus      ARIVGLCTGAFVLAAAGLLDGRPATTHWMYAPTLAKRYPSVHVDPRELFVDDGDVLTSAG
                     **:*.:*.*****. ******* .**** ** :*::***..:*:*  *:**:. :*****

neyagawaensis        TSAGIDACLHLVRKIQGSDVANAIARRMVVPPHREGGQAQYIERPVA-HVGEDGVREVLA
sp._HmicA12          TAAGIDLCLHIVRTDHGNEAAGALARRLVVPPRRSGGQERYLDRSLPEEIGADPLAEVVA
somaliensis          TAAGIDLCLHIVRTDHGTEAAGALARRLVVPPRRSGGQERYLDRSLPEEIGADPLAEVVA
griseus              TAAGIDLCLHIVRTDHGTEAAGALARRLVVPPRRSGGQERYLDRSLPEEIGSDPLAEVVA
lysosuperificus      TAAGIDLCLHIVRTDHGSEAAGALARRLVVPPRRTGGQERYLDRSLPEEIGADPLAEVVA
                     *:**** ***:**. :*.:.*.*:***:****:* *** :*::* :  .:* * : **:*

neyagawaensis        WAERHLDQEISVEQLSARACMSPRTFARRFRMETGTTPYRWLLAQRVLSAQRLLENTDDT
sp._HmicA12          WALEHLHEQFDVETLAARAYMSRRTFDRRFRSLTGSAPLQWLITQRVLQAQRLLETSDYS
somaliensis          WALEHLHEQFDVETLAARAYMSRRTFDRRFRSLTGSAPLQWLITQRVLQAQRLLETSDYS
griseus              WALEHLHEQFDVETLAARAYMSRRTFDRRFRSLTGSAPLQWLITQRVLQAQRLLETSDYS
lysosuperificus      WALEHLHEQFDVETLAARAYMSRRTFDRRFRSLTGSAPLQWLITQRVLQAQRLLETSDYS
                     ** .**.:::.** *:*** ** *** ****  **::* :**::****.******.:* :

neyagawaensis        MEMVAARTGFANAAALRYQFVRSLNTTPNAYRRAFRGRLPDTTTH----------------
sp._HmicA12          VDEVAGRCGFRSPVALRGHFRRQLGSSPAAYRAAYRARRPQTGEGPGSNGSGGHPGKLSE
somaliensis          VDEVAGRCGFRSPVALRGHFRRQLGSSPAAYRAAYRARRPQGDGAAPAAE----------
griseus              VDEVAGRCGFRSPVALRGHFRRQLGSSPAAYRAAYRARRPQGVAESAATV----------
lysosuperificus      VDEVAGRCGFRSPVALRGHFRRQLGSSPAAYRSAYRARRPQADVAQVTEM----------
                     :: **.* ** . .*** :* *.*.::* *** *:*.* *:

neyagawaensis        ------------------------------------------------------------
sp._HmicA12          KPSVELTSVPGPVDRDPAGAQGPGPVPTQSRRTAAASSPLGAHAAALQTGPGGPTGHTEG
somaliensis          ------PAVPSQSGGGPGPALR---------------------------RTAAGVGGGEP
griseus              ------VETMVPSQGPPSGRRG----------STLSS---------AAVAVAASVGSGEL
lysosuperificus      ------STGPVPHQRTPQPHRA----------AAALA----------------AAGPTVT


neyagawaensis        ----------------------
sp._HmicA12          GKPPSDAFAPGRASLPGPRTAP
somaliensis          GRQHVEALASGRASLPGQRSAP
griseus              SLPGPDAYVPGRPALPGQRSAP
lysosuperificus      E-----LYAPNRVLREHA----
```

Fig. 3 Amino acid alignment of streptomycetes AdpA proteins from *S. griseus*, S. *somaliensis, S.* sp. *HmicA12, S. lysosuperificus*, and *S. neyagawaensis* showing the position of the TTA codon-encoded leucine amino acid (yellow box) along with the position of the HTH motifs (box) of DNA-binding domain. Conserved (asterisk) and homologous (colon) amino acids are marked underneath. (Color figure online)

a MarA2 ortholog. Interestingly, its gene is adjacent to the one encoding a ThiJ-like protein Rv1930c. Although Rv1931 was not among RBHs for $AdpA_{cl}$ or any other streptomycete AdpA, we decided to take a closer look at mycobacterial AdpA-like proteins given the medical importance of mycobacteria and one report (Frota et al. 2004) showing that *rv1931c* plays a role in the virulence of *M. tuberculosis*.

Table 2 Maximum log-likelihood scores and, in parentheses, AIC for the 3 different models for *adpA* gene from 91, 232, or 323 species

Dataset model	TTA$^+$ *adpA* genes	TTA$^-$ *adpA* genes	All *adpA* genes
M0	−22928.1 (46218.2)	−78081.5 (157089.0)	−89770.1 (180830.2)
M0+Γ	**−21934.7 (44233.4)**	**−75257.2 (151442.4)**	**−86863.3 (175018.6)**
LG+Γ+F	−32075.4 (64510.8)	−112,952 (226828.0)	−123,272 (247832.0)

Values for the best fitting model for each dataset are in bold. *AIC* the value of the Akaike Information Criterion (Akaike 1973) model selection index. Low AIC values indicate better models

We constructed position-specific weight matrix from known AdpA operator sequences (Fig. S4, ESM) and searched for the presence of such operators in the promoters of genes for mycobacterial AdpA orthologs. PREDetector (Hiard et al. 2007) and MEME (Bailey et al. 2009) identified putative AdpA operators upstream of 18 mycobacterial genes for AdpA RBHs. These operators had *E* values (at least 10^{-10}) and *p* values (see Table S1 for examples, ESM) low enough to consider discovered motifs significant. Also, we revealed a high degree of gene conservation around *adpA* in these *Mycobacterium* species (Fig. S5, ESM).

We carried out genome-wide screening of mycobacterial genomes for AdpA operators using PreDetector. The operators have been identified in the promoters of *whiB* genes from genomes of *M. smegmatis* JS623, *M. rhodesiae* NBB3, *M. abscessus* subsp. *bolletii* 50594, and *M. tuberculosis* H37Rv (Table S2, ESM). Some of the WhiB-encoding genes are regulated by AdpA in *S. griseus* (Higo et al. 2012). WhiB proteins are well known for their diverse roles in actinobacterial morphogenesis, cell division, virulence, primary and secondary metabolism, and intrinsic antibiotic resistance. Our findings suggest that AdpA-like regulators in mycobacteria may play an important function in regulating different cell processes. It is of interest to investigate whether AdpA-like proteins can be pleiotropic regulators of diverse cellular processes in mycobacteria, or whether such regulatory model has evolved only in certain genera, such as *Streptomyces*.

Distribution of Probable AdpA Orthologs Beyond Phylum *Actinobacteria*

We decided to analyze the closest neighbors of *Actinobacteria* on the bacterial tree. Transcriptional regulators of AraC family constitute one of the most abundant functional classes encoded in genomes from all domains of life (Aziz et al. 2010). However, AraC domain is not listed as having one of "the 100 most ubiquitous functional roles in genomes." Abundance without ubiquity is believed to be an indicator of adaptive, organism-specific, or habitat-specific functionality. Different approaches towards phylogenetic reconstruction (Ciccarelli et al. 2006; Jolley et al. 2012; Segata et al. 2013) suggest that cyano- and actinobacteria are one of the most deeply branched phyla that share the closest last common ancestor. According to literature-based evolutionary time-scale server TimeTree (http://www.timetree.org/), cyano- and actinobacteria diverged around 2850 Mya, which is less than divergence time between actinobacteria and Firmicutes, such as *Bacillus* (3070 Mya). Indeed, using the PhyloPhlAn method (Segata et al. 2013), we confirmed previous results for our extended set of actinobacterial genomes (Fig. 2). PhyloPhlAn is a computational pipeline for reconstructing highly accurate and resolved phylogenetic trees based on whole-genome sequence information. This method assigns microbial phylogeny and putative taxonomy using > 400 proteins optimized from among 3737 genomes. Also the method measures the sequence diversity of all clades, classifies genomes from deep-branching candidate divisions through closely related subspecies, and improves consistency between phylogenetic and taxonomic groupings. We used PhyloPhlAn because it implements insertion operations, so we could incorporate new genomes of interest into pre-built microbial tree of life.

After an analysis of 198 genomes, we identified nine AdpA$_{cl}$ RBHs within cyanobacterial orders of *Nostocales, Synechococcales*, and *Oscillatoriales*. Interestingly, only three of them are full-length orthologs with two-domain organization: those were from *Acaryochloris* sp. CCMEE 5410, *Oscillatoriales cyanobacterium 'hensonii'* and *Hassallia byssoidea* VB512170. Others contained only C-terminal AraC domain. Thus, our genome mining efforts yielded a great diversity of genetic architectures within a group of AdpA-like proteins of *Actinobacteria* and *Cyanobacteria*, as summarized in Fig. 2.

Estimation of Selective Pressure on the Protein

To infer *adpA* codon positions under selection on the protein level, we choose to examine the initial (smaller) dataset because it contains the most extensively studied and validated sequences. We analyzed the entire 323 *adpA* gene dataset, only TTA$^+$ genes (91; mostly *Streptomyces*), and 232 TTA$^-$ genes. Here TTA$^+$ genes were those having TTA codon within the interdomain region (as in *adpA$_{cl}$*); the TTA$^-$ dataset consisted of genes that either lack TTA or carry it in positions other than in *adpA$_{cl}$*.

After performing the analyses with CodonPhyML program, we compared the fit of three different models for three

datasets using AIC (Akaike 1973). Table 2 indicates that the codon model M0+Γ gives a better fit to *adpA* datasets compared to other models. Thus, phylogenetic trees constructed under M0+Γ model were used for further selection pressure analysis. We evaluated the selection pressure at the protein level and examined the effect of among-site variation in dS on the estimation of ω in our datasets. For all three datasets the site-to-site variation in dS was significant (p value $\ll 0.01$; Table 3).

For the group of 91 TTA+*adpA* genes, highest synonymous substitutions rates (sites with fast inferred dS rate categories) can be observed at the C-terminus downstream of a DNA-binding motif (Fig. 4). Two small regions with slowest a posteriori dS rates locate in GATase domain and another two in the interdomain part and HTH domain of the protein. Interestingly, one of these regions contains TTA codon, whose adjacent positions correspond to the inferred lowest dS rate categories. At the same time, three regions with sites having highest values of dN/dS can be found in GATase domain and one in HTH domain (Fig. 4). We identified 93 positively selected sites in this dataset, 66 of them were located in the region involved in dimerization and 21—in the DNA-binding motif. Two amino acids under positive selection were located next to TTA codons:proline downstream L224 ('classical' TTA codon in *S. griseus*), and glycine upstream of Leu (which is encoded by TTA codon in *S. aurantiacus* JA 4570, *S. lysosuperificus* ATCC 31396, and *S. violaceusniger* SPC6). However, according to our analysis TTA codon is under purifying selection.

In the DNA-binding motif sites under positive selection included the following amino acid residues: A255, R261, R262, R267, R269, A275, L277, R320 (Fig. S6, ESM). From the crystal structure of AdpA DNA-binding domain it could be inferred that the guanidinium groups of Arg-261, Arg-267, and Arg-269 form salt bridges with the phosphate groups of C4′, T1, and C6′ nucleotides, respectively, of the 10-bp consensus AdpA-binding sequence (5′-TGGCSNG-WWY; S:G or C; W:A T; Y:T or C; and N:any nucleotide). C4′ is one of the most highly conserved nucleotides in the AdpA operator. G2 is recognized by R262—the residue essential for the recognition of the consensus sequence (Yao et al. 2013). The Cβ atom of Ala-275, via van der Waals forces, interacts with the 2′-deoxyriboses of C5′ nucleotide. In the HTH2 motif, the side-chain guanidinium of Arg-320 forms electrostatic interactions with the phosphate groups of C10 nucleotide (Yao et al. 2013). Thus, the ability of AdpA to recognize degenerate sequence appears to be due to the ongoing evolutionary pressure for innovation that is taking place in the *Streptomyces* clade.

In the 232 (TTA^-) and 323 (all AdpA) datasets no sites under positive selection were identified. Sites with elevated dS and ω were spread evenly along the gene length in both datasets (Fig. 4). Overall, more sites with elevated ω values were observed in the TTA^+ set compared to the 232 and 323 sets.

Functional Significance of AraC Domain: Experimental Test

The aforementioned genomic analysis pointed to the possibility that two-domain architecture of AdpA (as well as other HTH_18 superfamily proteins) could be a result of fortuitous gene fusion. We tested the hypothesis that isolated AraC domain of AdpA should still exhibit biological function in the absence of GATase-1. In the two-domain organization of AdpA, AraC domain is thought to play a principal role in transcriptional regulation, whereas GATase-1 mediates protein dimerization (Yamazaki et al. 2004; Kato et al. 2005). The importance (structural or functional) of dimerization is not supported experimentally.

We cloned the entire sequences of *adpA*-encoding genes from *Streptomyces albus* J1074 (*XNR_4181*), *S. ghanaensis* ATCC14672 ($adpA_{gh}$), *S. clavuligerus* ATCC7064 ($adpA_{cl}$), *S. coelicolor* M145 ($adpA_{sc}$) as well as their AraC domain-encoding segments (*XNR_4181_dbd*, $adpA_{gh}$*_dbd*, $adpA_{cl}$*_dbd*, $adpA_{sc}$*_dbd*) into cumate-inducible expression vector pGCymRP21. The resulting plasmids, referred to as pGC4181, pOOB103c, pGCSCLA, pGCSCO, pGC4181d, pGCSGHd, pGCSCLAd, and pGCSCOd (see Materials and Methods), were introduced into $adpA_{sc}$ knockout strain *S. coelicolor* M851 (Takano et al. 2003). We have also generated and transferred into M851 plasmid pGC1761, which

Table 3 Likelihood analyses of *adpA* datasets

Model dataset	MG94×REV Non-synonymous GDD 3		MG94×REV dual GDD 3×3		Likelihood ratio test statistic
	Log L	AIC	Log L	AIC	
TTA^+ *adpA* genes	−21861.71	44101.42	**−21544.91**	**43475.82**	633.6
TTA^- *adpA* genes	−76204.42	153350.85	**−75903.37**	**152756.74**	602.1
All *adpA* genes	−88054.06	177414.12	**−87382.14**	**176078.27**	1343.8

Log L is the natural logarithm of the maximum likelihood value for a model. All likelihood ratio tests are significant with p values $<<0.01$ (degrees of freedom = 4). Values for the best fitting model are in bold. *AIC* the value of the Akaike Information Criterion (Akaike 1973) model selection index. Low AIC values indicate better models

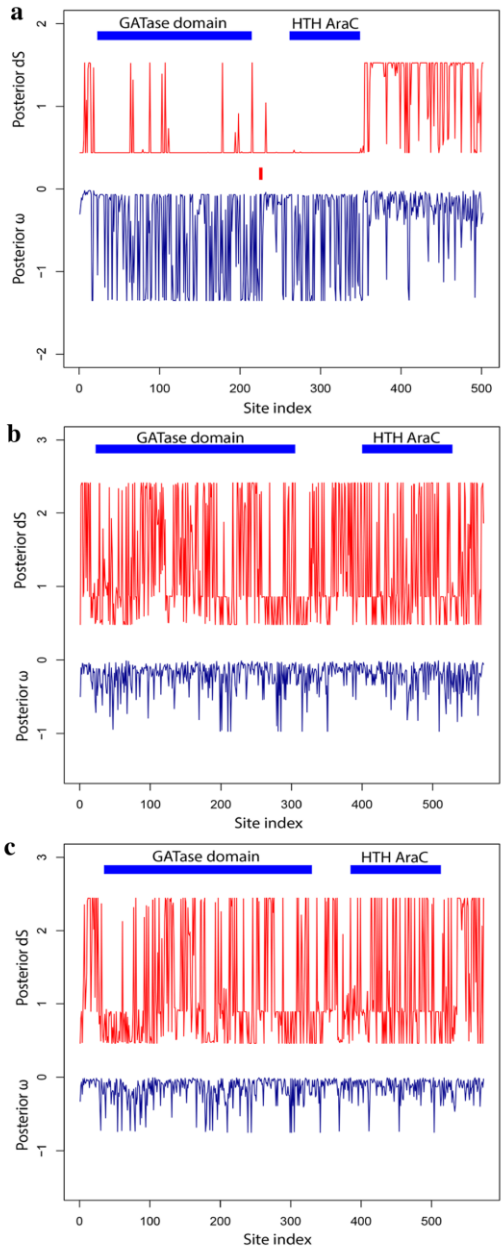


Fig. 4 Posterior dS and ω at each amino acid site across the *adpA* gene for dual rate variation model: **a** TTA^{+} set; **b** TTA^{-} set; **c** all AdpA set. TTA codon is shown as the red box on the *x* axis of the plot. A schematic of AdpA domain structure is shown above the plot. (Color figure online)

is pGCymRP21 carrying *XNR_1761*, the closest paralog of *XNR_4181* in *S. albus* genome. Xnr_1761 also belongs to HTH_18 superfamily of transcriptional regulators, possesses the same domain architecture as AdpA proteins do, and shares 60% aa sequence identity with Xnr_4181. It was interesting to see whether *XNR_1761* is able to complement Bld phenotype of *S. coelicolor* M851.

As could be seen from Fig. 5, M851 pGCymRP21^{+} strains failed to produce spores (a hallmark of *adpA* deficiency) on rich R5 medium, while introduction of full *XNR_4181*, *adpA*$_{gh}$, *adpA*$_{cl}$, and *adpA*$_{sc}$ restored spore formation to M851. The introduction of *XNR_4181_dbd*, *adpA*$_{gh}$*_dbd*, *adpA*$_{cl}$*_dbd*, *adpA*$_{sc}$*_dbd* also led to sporulation of M851. In contrast, introduction of *XNR_1761* did not influence M851 morphogenesis, suggesting that effects of *adpA* genes and their DNA-binding domains on the *adpA* mutant morphology are rather specific.

Discussion

In this work we combined computational and experimental approaches to investigate the diversity, evolution, and function of AdpA family regulatory proteins in *Streptomyces* and class *Actinobacteria*. The obtained data expand and complement our understanding of this family, which was previously based on a few model streptomycetes.

It remained an open question whether AdpA, as a two-domain protein of high pleiotropicity, was "invented" by *Streptomyces*, or whether it has originated before the *Streptomyces* branched off the actinobacterial phylogenetic tree. Chandra and Chater (2014) addressed this issue on the basis of 100 actinobacterial genomes. Their work suggested that AdpA orthologs are ubiquitously present in *Streptomyces*, and corresponding *adpA* genes carry a TTA codon in the "classical" region. Importantly, probable AdpA orthologs were sparsely found in some other actinobacterial taxa, suggesting AdpA origin before the speciation of *Streptomyces*. After mining 2215 actinobacterial genomes, we came to similar conclusions, albeit with a few important amendments. The first one concerns the *Streptomyces* clade (see Fig. 2). The latter invariably features AdpA orthologs, yet there is diversity with regard to codon TTA. In addition to the "classic" location of TTA between the two domains, we found a few "TTA-free" *adpA*, as well as orthologs in which this codon is located within the GATse domain or close to the stop codon. The second amendment concerns the presence of AdpA orthologs on a substantially larger phylogenetic scale. Here we show that AdpA$_{cl}$ RBHs exist in many actinobacteria and GATase-1-AraC two-domain configuration prevails. Other variants (such as single GATase-1/AraC domains, or presence of other domains) could be found, mostly in orders *Corynebacteriales* and *Bifidobacteriales*.

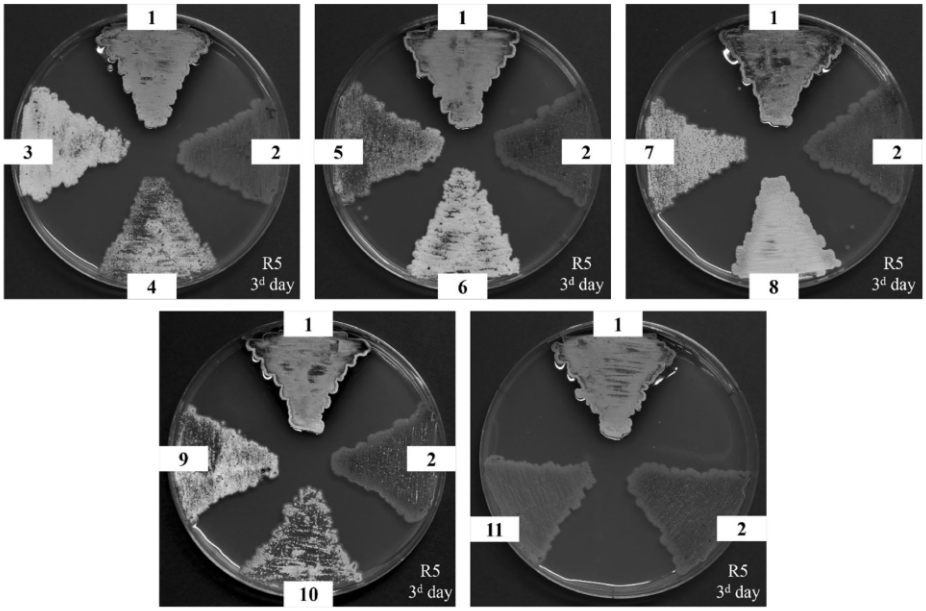


Fig. 5 Three-day lawns of control [*S. coelicolor* M145 (*1*) and *S. coelicolor* M851 pGCymRP21$^+$(*2*)] and complemented *S. coelicolor* strains [*S. coelicolor* M851 pGCSCOd$^+$ (*3*), pGCSCO$^+$(*4*), pGCSCLAd$^+$ (*5*), pGCSCLA$^+$ (*6*), pGC4181d$^+$ (*7*), pGC4181$^+$ (*8*), pGCSGHd$^+$ (*9*), pOOB103c$^+$ (*10*), and pGC1761$^+$ (*11*)] on R5 agar plates. The *adpA* mutant-carrying pGCymRP21 (M851 pGCymRP21$^+$) did not produce aerial mycelium. Formation of aerial mycelium and spores could be observed for all complemented strains except *S. coelicolor* M851 pGC1761$^+$

We also identified AdpA$_{cl}$ orthologs in a few cyanobacteria. In our work ortholog identification was based on BLASTP RBHs assuming that protein domains are units of orthology (Gabaldon and Koonin 2013). These criteria are less strict as compared to those used by Chandra and Chater (2014), and so might lead to ortholog misidentification in case of multiparalogous protein families, such as AdpA. Nevertheless, our approach agrees with previous estimates for a number of actinobacterial taxa; greater abundance and diversity of AdpAs could be a natural result of deeper sampling of all taxonomic groups. We believe that our dataset outlines a general picture of AdpA counterparts across long evolutionary distances, creating a starting point for the identification of genuine orthologs beyond *Streptomyces* and their functional characterization.

The ability to recognize a degenerate operator is the most salient trait of the AdpA function in Streptomycetes. In contrast, nothing is known about the mechanism and the role of probable AdpA orthologues in non-streptomycete species. Here we reveal that *Streptomyces* AdpA is under positive selection. Positively selected sites are grouped in regions involved in DNA-binding and protein dimerization. The precise location of codon sites under selection could not be identified reliably within other *adpA* genes. Thus, AdpA, as it is known from the studies of a few model cases, is most likely the *Streptomyces*-specific innovation, and this knowledge could not be directly extrapolated onto the other genera. Our recent studies of the AdpA-like protein from uncommon actinomycete *Actinoplanes teichomyceticus* support this suggestion (Ostash et al. 2015).

Our work reaffirms previous finding that only streptomycete *adpA* genes are associated with the rare TTA codon; nevertheless, a few TTA-free *adpA* sequences from *Streptomyces* were revealed. Although previously implicitly assumed (Chater and Chandra 2008; Chandra and Chater 2014), our data revealed no evidence of strong selection for TTA codon in *adpA*. The TTA-containing genes in streptomycetes are characterized by an extreme positional bias of TTA codon towards the start of genes (Zaburannyy et al. 2009). This is speculated to decrease the energy waste that is proportional to the number of codons that will be translated until the UUA is encountered (Fuglsang 2005). In *Streptomyces adpA* genes TTA is usually located in the middle of coding sequences, clearly violating the above-mentioned rule. However, such a TTA codon placement is easier to understand if two-domain AdpA arose from gene fusion; in

this case TTA would be at the beginning of AraC-encoding gene. Although aforementioned explanation looks theoretically plausible, it is not supported by available evidence. First, there are no separate GATase-1 or AraC-encoding genes across actinobacteria carrying TTA at C- or N-termini, respectively; second, we have found several streptomycete TTA-free *adpA* as well as several *adpA* carrying TTA outside of linker sequence; third, there are no signs of positive selection on codon position where TTA is usually located. It is therefore likely that TTA emerged in *adpA* after two-domain-encoding sequence has been formed, although evolutionary and functional reasons behind this codon remain obscure.

The gene fusion scenario of AdpA evolution is supported by the presence of the adjacent GATase-1 and AraC-encoding genes (*adpA* paralogues) in actinomycete genomes and heterologous complementation of *adpA* mutant *S. coelicolor* M851 with AraC domains derived from native and several heterologous AdpA-encoding genes. In the latter experiment we tested the ability of the isolated AraC domain to complement the most obvious defect of $adpA_{sc}$ mutant, e.g., its inability to erect aerial hyphae. The AdpA is proposed to bind its operator sequences in either monovalent or bivalent mode, namely either one or two AraC domains of the AdpA dimer contact the operators (Yamazaki et al. 2004). Hence a separate AraC domain of AdpA would still be able to function in a monovalent fashion; only targets requiring bivalent AdpA binding will be affected. At present we assume that GATase-1 domain is not absolutely required for AdpA to control morphogenesis. Notably, none of the *adpA-l*/*adpA_dbd*-expressing constructs fully complemented the loss of an actinorhodin production by the *S. coelicolor* ΔadpA (M851) strain. Although reasons for that can be different for different genes, we observed that vector-expressed cumate repressor CymR affected actinorhodin biosynthesis. Also, recent report (Xu et al. 2017) showed that full-length AdpA could interfere with antibiotic production. It will be of interest to examine the effect of AraC domain on the other traits known to be influenced by the full-length AdpA (Higo et al. 2012). This should deepen the understanding of the dimerization domain in an overall function of AdpA proteins, which, in turn, will further illuminate the evolutionary trajectory of this fascinating family of regulators.

Acknowledgements The work was supported by grant BG-41Nr from the Ministry of Education and Science of Ukraine (to B.O.).

References

Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F (eds) Second international symposium on information theory. Akademiai Kiado, Budapest, pp 267–281

Aziz RK, Breitbart M, Edwards RA (2010) Transposases are the most abundant, most ubiquitous genes in nature. Nucleic Acids Res 38(13):4207–4217

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME suite: tools for motif discovery and searching. Nucleic Acids Res 37(Suppl 2):202–208

Chandra G, Chater KF (2014) Developmental biology of *Streptomyces* from the perspective of 100 actinobacterial genome sequences. FEMS Microbiol Rev 38(3):345–379

Chater KF, Chandra G (2008) The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in *Streptomyces*. J Microbiol 46(1):1–11

Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P (2006) Toward automatic reconstruction of a highly resolved tree of life. Science 311(5765):1283–1287

Frota CC, Papavinasasundaram KG, Davis EO, Colston MJ (2004) The AraC family transcriptional regulator Rv1931c plays a role in the virulence of *Mycobacterium tuberculosis*. Infect Immun 72(9):5483–5486

Fuglsang A (2005) Intragenic position of UUA codons in streptomycetes. Microbiology 151(Pt 10):3150–3152

Gabaldon T, Koonin EV (2013) Functional and evolutionary implications of gene orthology. Nat Rev Genet 14(5):360–366

Gil M, Zanetti MS, Zoller S, Anisimova M (2013) CodonPhyML: fast maximum likelihood phylogeny estimation under codon substitution models. Mol Biol Evol 30(6):1270–1280

Goldman N, Yang Z (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol 11:725–736

Guyet A, Benaroudj N, Proux C, Gominet M, Coppée JY, Mazodier P (2014) Identified members of the *Streptomyces lividans* AdpA regulon involved in differentiation and secondary metabolism. BMC Microbiol 14:81

Hackl S, Bechthold A (2015) The Gene bldA, a regulator of morphological differentiation and antibiotic production in *Streptomyces*. Arch Pharm 348(7):455–462

Hiard S, Maree R, Colson S, Hoskisson PA, Titgemeyer F, van Wezel GP, Joris B, Wehenkel L, Rigali S (2007) PREDetector: a new tool to identify regulatory elements in bacterial genomes. Biochem Biophys Res Commun 357(4):861–864

Higo A, Horinouchi S, Ohnishi Y (2011) Strict regulation of morphological differentiation and secondary metabolism by a positive feedback loop between two global regulators AdpA and BldA in *Streptomyces griseus*. Mol Microbiol 81(6):1607–1622

Higo A, Hara H, Horinouchi S, Ohnishi Y (2012) Genome-wide distribution of AdpA, a global regulator for secondary metabolism and morphological differentiation in *streptomyces*, revealed the extent and complexity of the AdpA regulatory network. DNA Res 19(3):259–273

Horbal L, Fedorenko V, Luzhetskyy A (2014) Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. Appl Microbiol Biotechnol 98(20):8641–8655

Horinouchi S (2007) Mining and polishing of the treasure trove in the bacterial genus *streptomyces*. Biosci Biotechnol Biochem 71(2):283–299

Ishikawa J, Hotta K (1999) FramePlot: a new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. FEMS Microbiol Rev 174(2):251–253

Jolley KA, Bliss CM, Bennett JS, Bratcher HB, Brehony C, Colles FM, Wimalarathna H, Harrison OB, Sheppard SK, Cody AJ, Maiden MCJ (2012) Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. Microbiology 158(4):1005–1015

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90(430):773–795
Kato JY, Ohnishi Y, Horinouchi S (2005) Autorepression of AdpA of the AraC/XylS family, a key transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. J Mol Biol 350(1):12–26
Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical streptomyces genetics. John Innes Foundation, Norwich
Kosakovsky Pond SL, Frost SDW (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 22(5):1208–1222
Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. Bioinformatics 21(5):676–679
Lawlor EJ, Baylis HA, Chater KF (1987) Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). Genes Dev 1(10):1305–1310
Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol 25:1307–1320
López-García MT, Santamarta I, Liras P (2010) Morphological differentiation and clavulanic acid formation are affected in a *Streptomyces clavuligerus adpA*-deleted mutant. Microbiology 156(8):2354–2365
Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH (2011) CDD: a conserved domain database for the functional annotation of proteins. Nucleic Acids Res 39(Suppl 1):225–229
Moreno-Hagelsieb G, Latimer K (2008) Choosing BLAST options for better detection of orthologs as reciprocal best hits. Bioinformatics 24(3):319–324
Nguyen KT, Tenor J, Stettler H, Nguyen LT, Nguyen LD, Thompson CJ (2003) Colonial differentiation in *Streptomyces coelicolor* depends on translation of a specific codon within the *adpA* Gene. J Bacteriol 185(24):7291–7296
Nielsen R, Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148(3):929–936
Ohnishi Y, Kameyama S, Onaka H, Horinouchi S (1999) The A-factor regulatory cascade leading to streptomycin biosynthesis in *Streptomyces griseus*: identification of a target gene of the A-factor receptor. Mol Microbiol 34(1):102–111
Ohnishi Y, Yamazaki H, Kato JY, Tomono A, Horinouchi S (2005) AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. Biosci Biotechnol Biochem 69(3):431–439
Ostash B, Yushchuk O, Tistechok S, Mutenko H, Horbal L, Muryn A, Dacyuk Y, Kalinowski J, Luzhetskyy A, Fedorenko V (2015) The *adpA*-like regulatory gene from *Actinoplanes teichomyceticus*: in silico analysis and heterologous expression. World J Microbiol Biotechnol 31(8):1297–1301
Pan Y, Liu G, Yang H, Tian Y, Tan H (2009) The pleiotropic regulator AdpA-L directly controls the pathway-specific activator of nikkomycin biosynthesis in *Streptomyces ansochromogenes*. Mol Microbiol 72(3):710–723
Pond SK, Muse SV (2005) Site-to-site variation of synonymous substitution rates. Mol Biol Evol 22(12):2375–2385
Rodrigue N, Lartillot N, Philippe H (2008) Bayesian comparisons of codon substitution models. Genetics 180(3):1579–1591
Segata N, Börnigen D, Morgan XC, Huttenhower C (2013) PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. Nat Commun 4:2304
Sen A, Thakur S, Bothra AK, Sur S, Tisa LS (2012) Identification of TTA codon containing genes in *Frankia* and exploration of the role of tRNA in regulating these genes. Arch Microbiol 194(1):35–45
Šetinová D, Šmídová K, Pohl P, Musić I, Bobek J (2018) RNase III-binding-mRNAs revealed novel complementary transcripts in *Streptomyces*. Front Microbiol 8:2693
Sharp PM, Averof M, Lloyd AT, Matassi G, Peden JF (1995) DNA sequence evolution: the sounds of silence. Philos Trans R Soc Lond B 349(1329):241–247
Suzuki Y, Gojobori T (1999) A method for detecting positive selection at single amino acid sites. Mol Biol Evol 16(10):1315–1328
Szalkowski AM (2012) Fast and robust multiple sequence alignment with phylogeny-aware gap placement. BMC Bioinform 13:129
Takano E, Tao M, Long F, Bibb MJ, Wang L, Li W, Buttner MJ, Bibb MJ, Deng ZX, Chater KF (2003) A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. Mol Microbiol 50(2):475–486
Tanabe M, Kanehisa M (2012) Using the KEGG database resource. Curr Protoc Bioinform 38:1.12:1.12.1–1.12.43
Ventura M, Canchaya C, Tauch A, Chandra G, Fitzgerald GF, Chater KF, van Sinderen D (2007) Genomics of *Actinobacteria*: tracing the evolutionary history of an ancient phylum. Microbiol Mol Biol Rev 71(3):495–548
Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJ, Yoo HS, Zhang C, Zhang Y, Sobral BW (2014) PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res 42(D1):581–591
Wolański M, Donczew R, Kois-Ostrowska A, Masiewicz P, Jakimowicz D, Zakrzewska-Czerwińska J (2011) The level of AdpA directly affects expression of developmental genes in *Streptomyces coelicolor*. J Bacteriol 193(22):6358–6365
Wolański M, Jakimowicz D, Zakrzewska-Czerwińska J (2014) Fifty years after the replicon hypothesis: cell-specific master regulators as new players in chromosome replication control. J Bacteriol 196(16):2901–2911
Xu J, Zhang J, Zhuo J, Li Y, Tian Y, Tan H (2017) Activation and molecular mechanism of a cryptic oviedomycin biosynthetic gene cluster via the disruption of a global regulatory gene *adpA* in *Streptomyces ansochromogenes*. J Biol Chem. https://doi.org/10.1074/jbc.M117.809145
Yamazaki H, Tomono A, Ohnishi Y, Horinouchi S (2004) DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. Mol Microbiol 53(2):555–572
Yang Z, Bielawski JP (2000) Statistical methods for detecting molecular adaptation. Trends Ecol Evol 15(12):496–503
Yang Z, Nielsen R, Goldman N, Pedersen AM (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155(1):431–449
Yao MD, Ohtsuka J, Nagata K, Miyazono KI, Zhi Y, Ohnishi Y, Tanokura M (2013) Complex structure of the DNA-binding domain of AdpA, the global transcription factor in *Streptomyces griseus*, and a target duplex DNA reveals the structural basis of its tolerant DNA sequence specificity. J Biol Chem 288(43):31019–31029
Zaburannyy N, Ostash B, Fedorenko V (2009) TTA Lynx: a web-based service for analysis of actinomycete genes containing rare TTA codon. Bioinformatics 25(18):2432–2433
Zhao J, Wen Y, Chen Z, Song Y, Li J (2007) An adpA homologue in *Streptomyces avermitilis* is involved in regulation of morphogenesis and melanogenesis. Chinese Sci Bull 52(5):623–630

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s00239-018-9834-z.
Електронний додаток до статті доступний за посиланням:
https://static-content.springer.com/esm/art%3A10.1007%2Fs00239-018-9834-z/MediaObjects/239_2018_9834_MOESM1_ESM.pdf.

Анотація: *Білок AdpA з продуцента стрептоміцину Streptomyces griseus є засновником родини плейотропних регуляторів AdpA, які, як відомо, повсюдно присутні в геномах стрептоміцетів. Функціональні геномні підходи виявили величезну кількість мішеней AdpA, що дало змогу охарактеризувати регулон AdpA як найбільший у бактерій. Експресія adpA обмежена на рівні трансляції наявністю рідкісного лейцинового кодону UUA. Усі відомі властивості регуляторів AdpA були досліджені на кількох штамах стрептоміцетів. Існують відкриті питання щодо справжньої розповсюдженості та різноманітності AdpA серед інших таксонів актинобактерій (і не тільки), а також щодо можливих еволюційних процесів, які формують ортологічну групу AdpA у Streptomyces. В цій роботі ми демонструємо, що положення TTA-кодона в генах ортологічної групи adpA є варіабельнішим, ніж вважалося раніше. Ми виявили гомологів AdpA в представників багатьох рядів актинобактерій, які мають організацію доменів, що нагадує організацію AdpA, але існують також інші конфігурації. Диверсифікуючий позитивний добір був виявлений в межах ділянки adpA, що кодує ДНК-зв'язувальний домен (AraC) у Streptomyces. Цей добір, швидше за все, впливає на амінокислоти, що дають змогу AdpA розпізнавати вироджену операторну послідовність. Послідовність, що кодує передбачуваний домен глутамінової амідотрансферази (GATase-1), також демонструє ознаки позитивного добору. Двохдоменна організація AdpA, швидше за все, виникла в результаті злиття генів, що кодують окремі домени GATase-1 і AraC. Дійсно, виявлено, що домен AraC зберігає біологічну функцію за відсутності GATase-1 домена. Ми припускаємо, що регуляторна роль кодону TTA є результатом відносно недавньої події в еволюції AdpA, яка збіглася з розвитком клади Streptomycetales і в даний час знаходиться під послабленими селективним тиском. Запрошується подальше експериментальне дослідження наших висновків, яке має дати нове розуміння і перспектив еволюції AdpA-опосередкованого регуляторного каскаду.*

Author's personal copy

Arch Microbiol
DOI 10.1007/s00203-017-1389-7

ORIGINAL PAPER

Properties of *Streptomyces albus* J1074 mutant deficient in $tRNA^{Leu}_{UAA}$ gene *bldA*

Oksana Koshla[1] · Maria Lopatniuk[1,2] · Ihor Rokytskyy[1] · Oleksandr Yushchuk[1] · Yuriy Dacyuk[3] · Victor Fedorenko[1] · Andriy Luzhetskyy[2] · Bohdan Ostash[1]

Received: 30 March 2017 / Revised: 6 May 2017 / Accepted: 16 May 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract *Streptomyces albus* J1074 is one of the most popular and convenient hosts for heterologous expression of gene clusters directing the biosynthesis of various natural metabolic products, such as antibiotics. This fuels interest in elucidation of genetic mechanisms that may limit secondary metabolism in J1074. Here, we report the generation and initial study of J1074 mutant, deficient in gene *bldA* for $tRNA^{Leu}_{UAA}$, the only tRNA capable of decoding rare leucyl TTA codon in *Streptomyces*. The *bldA* deletion in J1074 resulted in a highly conditional Bld phenotype, with depleted formation of aerial hyphae on certain solid media. In addition, *bldA* mutant of J1074 was unable to produce endogenous antibacterial compounds and two heterologous antibiotics, moenomycin and aranciamycin, whose biosynthesis is directed by TTA-containing genes. We have employed a new TTA codon-specific β-galactosidase reporter system to provide genetic evidence that J1074 *bldA* mutant is impaired in translation of TTA. In addition, we have discussed the possible reasons for differences in the phenotypes of *bldA* mutants described here and in previous studies, providing knowledge to study *bldA*-based regulation of antibiotic biosynthesis.

Keywords *Streptomyces albus* · tRNA · Antibiotics · β-galactosidase

Communicated by Erko Stackebrandt.

Electronic supplementary material The online version of this article (doi:10.1007/s00203-017-1389-7) contains supplementary material, which is available to authorized users.

✉ Bohdan Ostash
b.ostash@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 4 Hrushevskoho st., Lviv 79005, Ukraine

[2] Actinobacteria Metabolic Engineering Group, Department of Pharmaceutical Biotechnology, Helmholtz Institute for Pharmaceutical Research Saarland, Saarland University, UdS Campus C2 3, 66123 Saarbrucken, Germany

[3] Department of Physics of Earth, Ivan Franko National University of Lviv, 4 Hrushevskoho st., Lviv 79005, Ukraine

Introduction

The actinobacterium *Streptomyces albus* J1074 was isolated in 1980 as a *S. albus* G derivative deficient in the *Sal*GI restriction-modification system (Chater and Wilde 1980). Later, J1074 served as a model to study phage restriction phenomena in streptomycetes, and in 1990s it has gained prominence as a host for heterologous expression of antibiotic biosynthesis genes (Rodriguez et al. 1993, 2000). Fast, highly dispersed growth and genetic amenability make J1074 a strain of first choice for expression of heterologous gene clusters from actinobacteria (Baltz 2010). Previous studies (Lopatniuk et al. 2014; Iqbal et al. 2016) portray J1074 as an optimal starting point to develop an efficient platform for heterologous expression (metagenomic) purposes. It is expected that such a platform would greatly increase our ability to discover bioactive natural products encrypted in environmental DNA and silent gene clusters of actinobacteria. In this context, it is essential to elucidate genetic mechanisms that control secondary metabolism in J1074 to find limiting factors and the ways to circumvent them. It has to be mentioned that regulation of antibiotic production has been relatively well studied in few models of streptomycetes (Liu et al. 2013), although it is not yet understood to what extent this knowledge can be extrapolated onto new strains, such as J1074.

Published online: 20 May 2017

Author's personal copy

Transcriptional control of gene expression is arguably the most important regulatory layer in bacteria, and antibiotic production is no exception. In addition, streptomycetes possess a highly unusual regulatory switch based on leucyl (TTA) codon and its cognate $tRNA^{Leu}_{UAA}$, encoded by gene *bldA*. TTA codons are located exclusively in genes for auxiliary functions such as morphogenesis and antibiotic production, and *bldA* deficiency is not lethal. There is a considerable temporal gap between occurrence of UUA-containing transcripts and their subsequent translation (Gramajo et al. 1993; Pope et al. 1996; Rebets et al. 2006). Failure to develop aerial mycelium and spores on top of substrate mycelium is a hallmark of all *bldA*-mutants (the "Bald" phenotype) studied to date. The latter is thought to be the result of blocked translation of TTA-containing gene for master regulator AdpA (Takano et al. 2003; Nguyen et al. 2003). The influence of *bldA* deficiency on antibiotic biosynthesis remains more controversial, as well as the exact chain of events linking *bldA* to delayed antibiotic production (Guthrie and Chater 1990; Gramajo et al. 1993; Chater and Chandra 2008; Makitrynskyy et al. 2013).

The effects of *bldA* deletion on J1074 have not been studied so far, and it remained unknown to what extent colony development and antibiotic production will be affected by the former. Here, we show that on certain (minimal) solid media *bldA*-deficient mutant J1074 produces aerial hyphae and spores less densely than the parental strain; in the same time, under many culture conditions parental and mutant strains are phenotypically indistinguishable. In line with many other reports, our *bldA* mutant's ability to express TTA codon-containing antibiotic biosynthesis gene clusters is strictly dependent on $tRNA^{Leu}_{UAA}$. Finally, we describe a novel reporter system that was used to study TTA codon-based regulatory mechanisms in *S. albus*.

Materials and methods

Microorganisms, vectors, and culture conditions

Streptomyces albus SAM2, a derivative of J1074 carrying single $attB^{\varphi C31}$ site (Bilyk and Luzhetskyy 2014), was used throughout the study. *Escherichia coli* DH5α was used for routine subcloning. *E. coli* ET12567 (pUB307) was used to perform intergeneric conjugation from *E. coli* to *S. albus*. *Escherichia coli* strains were grown under standard conditions (Sambrook and Russell 2001). *Bacillus cereus* ATCC19637 and *Debaryomyces hansenii* VKM Y-9 were used as antibiotic-sensitive test cultures (growth temperature −30 °C). RedET-mediated gene replacements (Datsenko and Wanner 2000) were carried out with the help of REDIRECT system (Gust et al. 2003). All vectors, cosmids, plasmids, and BAC DNAs are listed in Table 1. For intergeneric matings, *S. albus* strains were grown on soy-flour mannitol (SFM) agar (Kieser et al. 2000) at 30 °C. To reveal endogenous antibacterial and antifungal activities, *S. albus* strains were grown on SG2 agar (in g/L: glucose 20, yeast extract 5, soytone 10, pH 7.2 prior to autoclaving) and other solid media as detailed in Electronic Supplementary Materials (ESM). For submerged production of moenomycin and aranciamycin *S. albus* strains were grown in liquid medium SG1 (SG2 supplemented with 2 g/L $CaCO_3$) at 28 °C for 5 days. For phenotypic examinations, *S. albus* were grown on SFM and SMMS agar (Kieser et al. 2000), as well as other media as described in ESM. Wherever needed, strains were maintained in the presence of apramycin (25 μg/mL); chromogenic substrate X-Gal was added to the media to a final concentration of 50 μg/mL; cumate (50 mM) was used to induce *SCO3479* expression from plasmids pRV4 and pRV3.

Plasmids and BACs construction

Bacterial artificial chromosome (BAC) BAC_1N12 was used to prepare *bldA* (*XNR_1995*) knockout BAC. For this purpose, apramycin resistance cassette *aac(3)IV-oriT* flanked with P-GG and B-CC sites was amplified from plasmid patt-saac-oriT (Myronovskyi et al. 2014) with primers Am-bldAdelF and Am-bldAdelR (all primer sequences are given in Table s1, ESM). The resulting 1378-bp PCR product was used to replace *bldA* gene in BAC_1N12 using recombineering. The resulting BAC was named 1N12d_bldA. This and other constructs generated in course of the work were verified by sequencing. Plasmid pTOSbldA (complementation of the *bldA* deletion). The *bldA* gene together with putative promoter region was amplified with primers bldA-Hind-for and bldA-Xba-rev yielding a product 675 bp in length. The PCR product was digested with enzymes *Hin*dIII–*Xba*I and cloned into respective sites of vector pTOS to give pTOSbldA. Plasmid pTESsco3479 (expression of β-galactosidase gene *SCO3479* under promoter *ermEp**). Gene *SCO3479* was amplified with primers Sco3479upXbaI and Sco3479rpMfeIBgl from the chromosome of *S. coelicolor*. The PCR product, 3016 bp in length, was digested with restriction endonucleases *Xba*I and *Bgl*II, and then ligated to *Xba*I–*Bgl*II sites of vector pTES. Plasmids pOOB109 and pOOB110. The 3.3-kb fragment of *S. coelicolor* M145 genome that encompassed the entire coding sequence of *SCO3479* plus 0.3-kb putative promoter region was amplified with primers Sco3479upBglII and 3479rpMfeIBglII. The resulting amplicon was treated with restriction endonuclease *Bgl*II and cloned into *Bam*HI + *Bgl*II-digested vector pTES. Two clones were selected, referred to as pOOB109 and pOOB110, which contained *SCO3479* in two orientations within vector pTES. In pOOB109, *SCO3479* is under control of its own

Author's personal copy

Table 1 Plasmids, cosmids, and BACs used in this work

Name	Description	Source
BAC_1N12	Cm^r; pSMART-based bacterial artificial chromosome (BAC) carrying a segment of *S. albus* genome with gene *bldA*	This work
1N12_ bldA-I	BAC 1N12 derivative, replacement of *bldA* with IMES-flanked cassette *aac(3)IV-oriT*	This work
moeno38	SuperCos1-based, carries entire *moe* cluster 1 except *moeR5moeS5*, Km^rAp^r	Ostash et al. (2013)
moeno38-6	moeno38, replacement of *neo* with $int^{\varphi C31}$-$oriT^{RK2}$-*hyg* cassette (Hy^r), Δ*moeH5* (Km^r)	Ostash et al. (2013)
pOJ436ara	pOJ436-based cosmid, carries entire aranciamycin biosynthesis gene (*ara*) cluster	Luzhetskyy et al. (2007)
pTES	Am^r; phiC31-based *Streptomyces* integrative vector; expression of cloned gene from *ermEp**	Herrmann et al. (2012)
pTOS	Am^r; VWB-based *Streptomyces* integrative vector	Herrmann et al. (2012)
pIJ773	Source of FRT-flanked *aac(3)IV-oriT* cassette	Gust et al. (2003)
patt-saac-oriT	Source of IMES-flanked *aac(3)IV-oriT* cassette	Myronovskyi et al. (2014)
pALFLP	pSG5-based plasmid expressing recombinase Flp; Hy^r	Fedoryshyn et al. (2008)
pUWLint31	pIJ101-based plasmid expressing Int-phiC31; Hy^r	Myronovskyi et al. (2014)
pUWLDre	pIJ101-based, expresses SSR Dre to evict pTOS vector sequences	Herrmann et al. (2012)
pGCymRP21	Am^rSp^r; phiC31-based *Streptomyces* integrative vector; expression of a cloned gene from cumate-inducible promoter *cmt*-P21	Horbal et al. (2014)
pTESsco3479	pTES carrying *SCO3479* under control of *ermEp**	This work
pOOB109	pTES carrying *SCO3479* under control of tandem promoter—truncated *ermEp* + full-length *SCO3479p*	This work
pOOB110	pTES carrying *SCO3479* under control of full-length *SCO3479p*	This work
pOOB114	pTES carrying "dead" version of *SCO3479* (stop codon in 19th position) under control of *ermEp**	This work
pRV3	pGCymRP21 carrying mutant (TTA codon in 19th position) version of *SCO3479* under control of cmt-P21	This work
pRV4	pGCymRP21 carrying native version of SCO3479 under control of *cmt*-P21	This work
pTOSbldA	pTOS derivative carrying *bldA* gene along with upstream 0.6-kb region	This work

promoter as well as truncated ermEp*; and in pOOB110, *SCO3479* is under control of its own promoter. Plasmid pOOB114. 3017-bp fragment containing entire coding sequence of *SCO3479* plus 29-bp putative promoter was amplified from plasmid pOOB109 with primers lacZ_XbaI_TGA2 and 3479rpMfeIBglII. In this way, leucyl codon CTG in 19th position of *SCO3479* was replaced with stop codon TGA. The resulting amplicon was treated with XbaI and MfeI and cloned into XbaI–EcoRI sites of pTES to give pOOB114. Plasmid pRV3. 3017-bp fragment containing entire coding sequence of *SCO3479* plus 29-bp putative promoter was amplified from plasmid pOOB109 with primers lacZ_XbaI_TTA1 and Sco3479_6Hisrp. In this way, leucyl codon CTG in 8th position of *SCO3479* was replaced with synonymous codon TTA and six histidine codons CAC were appended upstream of stop codon. The resulting amplicon was treated with *Xba*I and *Mfe*I and cloned into *Spe*I–*Eco*RI sites of pGCymRP21 to give pRV3. In the latter, *SCO3479* replaced *gusA* gene. Plasmid pRV4. 3017-bp fragment containing entire coding sequence of *SCO3479* plus 29-bp putative promoter was amplified from plasmid pOOB109 with primers Sco3479upXbaI and Sco3479_6Hisrp. In this way, native sequence of *SCO3479* was appended upstream of hexahistidine-stop codon sequence $[(CAC)_6$-TGA]. The resulting amplicon was treated with *Xba*I and *Mfe*I and cloned into *Spe*I–*Eco*RI sites of pGCymRP21 to give pRV4.

Generation and verification of the *S. albus* recombinant strains

All constructs were transferred into *S. albus* conjugally, as described elsewhere; gene replacements in *S. albus* were generated as described in Myronovskyi et al. (2014). Apramycin resistance marker flanked with B-GG and P-GG sites was evicted from the genome of *bldA* mutant with recombinase Int-phiC31 (plasmid pUWLint31). pTOS vector sequences from *bldA* complementation strain were evicted with recombinase Dre (plasmid pUWLDre). PCR was employed to confirm the presence of the plasmids, cosmids, and expected gene replacements in the chromosomes of streptomycetes.

Determination of antibiotic production

Moenomycin production was analyzed in *S. albus* strains carrying cosmid moeno38-6. The latter leads to accumulation of two moenomycins in *S. albus*—nosokomycin

Author's personal copy

A_2 and its precursor lacking terminal galacturonic residue (Ostash et al. 2013; Lopatniuk et al. 2014). Strains were grown in SG1. Quantity of moenomycins was determined by antibiotic disc (Ø 5 mm, Whatman) diffusion assay against spores (10^7 per plate) of moenomycin-sensitive *Bacillus cereus* ATCC19637 as described in Makitrynskyy et al. (2010). Productivity index (PI) was calculated according to the equation: PI = (Ø growth inhibition zone, mm/Ø disc, mm) − 1. PI = 0 is the lowest possible value and corresponds to antibiotic production below detection limit of the bioassay (˜2% of production level of the control strain). The PIs were referred back to equal amounts of dry biomass (10 mg) in different strains.

For aranciamycin production, control and pOJ436ara-carrying strains were grown for 24 h in preculture medium (tryptic soy broth, TSB; Himedia) at 30 °C and 1 mL of the preculture was used to inoculate 50 mL (in 300-mL flasks with glass beads) of main medium SG1. After 5 days of growth, biomass was separated from supernatant via centrifugation. Spent medium (10–20 mL) was extracted with equal volume of ethylacetate. Ethylacetate extract was dried in vacuo, then reconstituted in methanol, and subjected to TLC (immobile phase: silicagel plates F60, Merck; mobile phase: chloroform:methanol = 95:5) and disc diffusion assay against *B. cereus* ATCC19637 spores (10^7 per plate) on modified minimal agar (g/L: KH_2PO_4 3, K_2HPO_4 7, sodium citrate × $4H_2O$ 0.5, $MgSO_4$ × $7H_2O$ 0.1, $(NH_4)_2SO_4$ 1, glucose 2, bacto peptone 0.3, agar 16) (Anagnostopoulos and Spizizen 1961).

Aranciamycin production and native antibiotic activity of *S. albus* strains were also monitored using agar plug antibiotic diffusion assay. Briefly, strains were grown on SG1 agar for 5 days. Then, agar plugs (Ø 5 mm) were cut off the lawn and stacked on top of TSB agar plates with test culture *D. hansenii* spread immediately prior to the experiments or *B. cereus* spores as it is described above. Halos of growth inhibition around the plugs were measured after 18 h of incubation.

Scanning electron microscopy (SEM)

Small pieces of 3-day-old lawns were cut off the OM or SFM agar plate samples, vacuum-dried, and directly analyzed on a Jeol JSM-T220A scanning microscope.

Results

Generation and growth characteristics of *S. albus* strains carrying *bldA* knockout

For *S. albus* J1074 *bldA* gene knockout, we prepared BAC 1N12_bldA-I, where *bldA* was replaced with *aac(3)IV-oriT* cassette flanked with B-GG and P-GG sites for recombinase PhiC31. With this BAC, we successfully generated respective *S. albus* mutant (ΔbldA-I) and then its markerless derivative (upon expression of PhiC31 from plasmid pUWLInt31). The markerless *bldA* mutant was referred to as OK3. OK3 genotype was confirmed by PCR (see Fig. S1, Electronic Supplementary Materials, ESM). We also introduced *S. albus* SAM2 *bldA* gene on vector pTOS back into OK3, and, after pTOS sequences eviction, generated markerless complemented strain OK3 + bldA.

In submerged culture, OK3 and the parent (SAM2) strains exhibited similar dynamics of growth and biomass accumulation (Fig. S2). The strains were also incubated on a number of solid media and analyzed after 3 and 5 days. On most media, OK3 was macroscopically and microscopically indistinguishable from the parent (Figs. S3, S4). Nevertheless, we have found that certain minimal media did result in visible changes in lawn morphology of OK3 (see Fig. S4). The most pronounced morphological differences were seen on SFM and SMMS (Figs. 1 and S5). On these media, OK3 produced sparse aerial hyphae and spore chains without significant delay as compared to the parental strain.

Antibiotic production by OK3 strain

Genomic potential of *S. albus* J1074 for secondary metabolism is significant (Zaburannyi et al. 2014), yet, to date, its expression required genetic interventions (Olano et al. 2014; Brana et al. 2014). Compounds of paulomycin family are the only known antibacterially active metabolites to be accumulated by J1074, although their production pattern was unstable (Olano et al. 2014; Gonzalez et al. 2016) We tested over a dozen of the most common agar media used for *Streptomyces* cultivation, and found those that reproducibly elicit the production of antibiotic compounds by J1074 (Figs. S6, S7). Particularly, *S. albus* accumulated antibacterial and (to a lesser extent) antifungal compounds when growing on SG2 agar. Using SG2, we compared antibiotic activity of SAM2, *bldA* mutant OK3 and OK3 + bldA. The results are summarized in Fig. 2. OK3 did not accumulate detectable amounts of antibacterial compounds and introduction of *bldA* restored antibiotic production to the mutant. The antifungal activity of OK3 was not impaired and even slightly increased as compared to SAM2. Using HPLC–MS (as detailed in ESM), we confirmed that production of candicidin-like compounds by OK3 is increased as compared to parent strain, while paulomycins were not found in OK3 (Fig. S8).

Next, we compared the ability of SAM2 and OK3 to heterologously produce antibiotics whose biosynthesis depend on TTA codon-containing genes. These were moenomycin and aranciamycin biosynthesis gene clusters, carried on

Author's personal copy

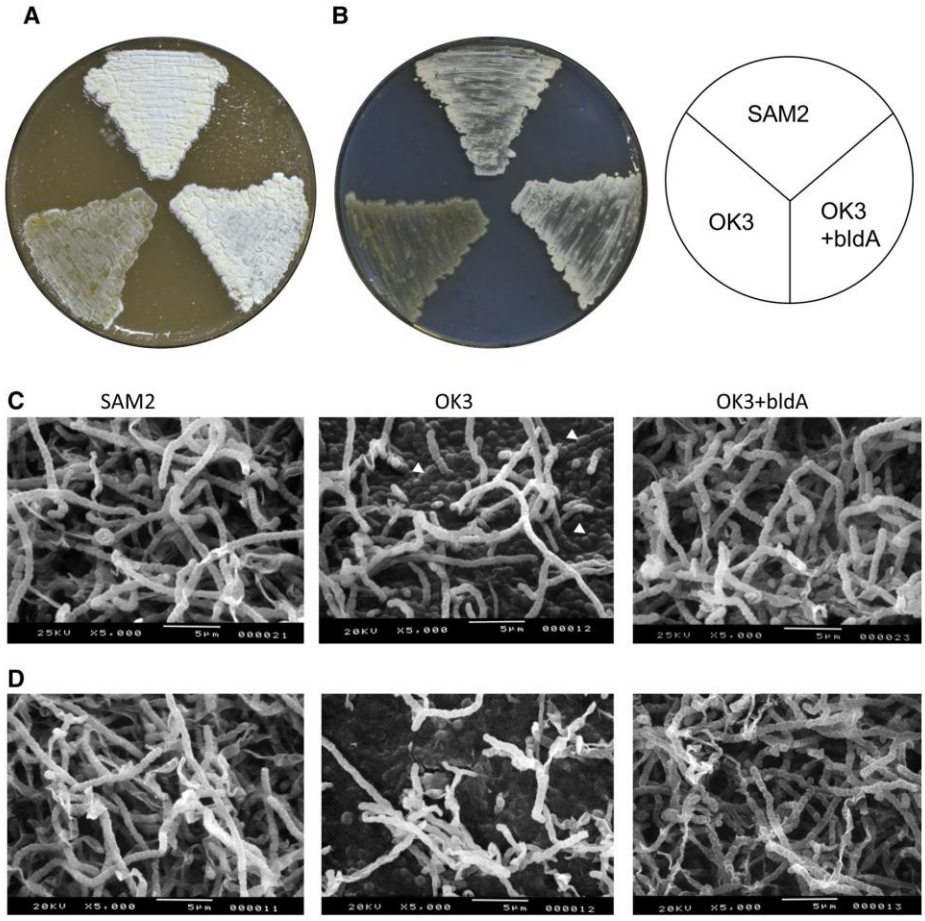


Fig. 1 *Streptomyces albus* J1074 *bldA* mutant phenotype. Parental (SAM2), mutant and complemented strains were grown for 3 days on SFM (**a**) and SMMS (**b**). The same lawns were used to obtain scanning electron microscopic views of colonial surfaces of SFM (**c**) and SMMS-grown cultures (**d**). The substrate mycelium of OK3 strain on SFM (**c**) appeared to be fragmented into spore-like chains (*white triangles*)

cosmids moeno38-6 (Lopatniuk et al. 2014) and pOJ436ara (Luzhetskyy et al. 2007), respectively. As can be judged from bioactivity and TLC-based assays, moeno38-6^{+} and pOJ436ara^{+} transconjugants of the parental strain abundantly produced respective antibiotics, while OK3 strain did not (Figs. 3, S9–S10).

The β-galactosidase reporter system to detect and elucidate *bldA*-based regulation

The above-described results imply that, in line with what is known for the other *Streptomyces* (Chater and Chandra 2008; Liu et al. 2013), failure to translate codon UUA in transcripts of secondary metabolism genes is the primary reason for the inability of OK3 to produce endogenous and heterologous antibiotics. Yet, with two exceptions (Wang et al. 2009; Makitrinskyy et al. 2013), direct evidence for involvement of *bldA* in regulation at the level of UUA translation is absent for all studied cases. In *S. albus* J1074, the *bldA* mutation and the observed phenotype (lack of antibacterial activity of the strain) are bridged by a complicated cascade of regulatory and structural genes. Codon-swap reporters based on β-galactosidase and luciferase genes are used in enterobacteria and yeast to peer into regulatory events at the level of translation of a certain codon (Urbonavicius et al. 2003; Lamichhane et al. 2013). We set out

Author's personal copy

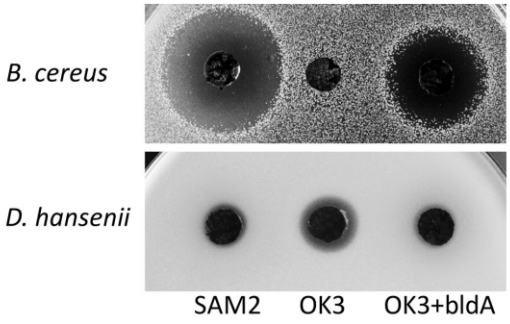


Fig. 2 The *bldA* knockout in *S. albus* abolished the production of antibacterial compound(s), but not antifungal one(s). For the bioassay, agar plugs were cut off 5-day lawns of *S. albus* strains grown on SG2 agar, and stacked on top of plates with test cultures. Note the small halos of *D. hansenii* growth inhibition around SAM2 and OK3 + *bldA* plugs, and increased halo around OK3 plug

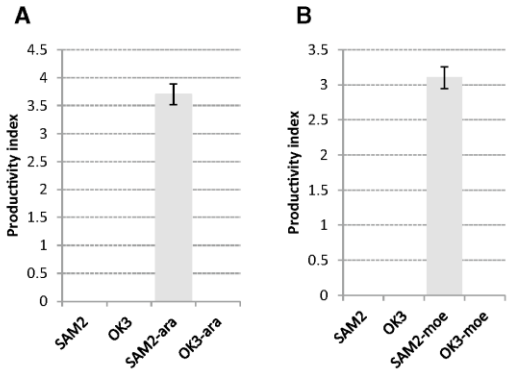


Fig. 3 Aranciamycin (**a**) and moenomycin (**b**) production is blocked in *bldA* mutant of *S. albus*. Strains were grown in SG1 medium, and productivity indices were calculated from disc diffusion assays as described in "Materials and methods". SAM2, OK3—strains carrying empty vector pSET152; SAM2/OK3-ara, SAM2/OK3-moe—*S. albus* strains carrying either cosmid pOJ436ara or cosmid moeno38-6, respectively

to develop an analogous reporter system for *Streptomyces albus* J1074, which would enable straightforward phenotypic detection of translation of UUA. Here, we took advantage of the fact that J1074 is one of the rare streptomycetes having no detectable intra- or extracellular enzymatic activity towards lactose analog, X-Gal (King and Chater 1986). Indeed, our analysis confirmed that none of the seven annotated *S. coelicolor* β-galactosidases had counterparts within J1074. Using reciprocal best BLAST hit strategy (Kuzniar et al. 2008), *S. coelicolor* M145 protein Sco3479 (referred hereafter to as LacZ_{sc}) was identified as an ortholog of *E. coli* LacZ. Several *lacZ*$_{sc}$-carrying plasmids were prepared (Fig. 4a) and expression of functional β-galactosidase from all of them (except for pOOB114 which carries in-frame stop codon in position 19 of *lacZ*$_{sc}$) was confirmed. Particularly, gene *lacZ*$_{sc}$ conferred *lacZ*-deficient *E. coli* to the ability to hydrolyze X-Gal; in liquid and solid media, different *lacZ*$^{+}_{sc}$ *S. albus* strains produced blue color of different intensity (Fig. S11). The latter most likely reflects different strength of transcription of *lacZ*$_{sc}$.

We introduced TTA, CTC, and stop codon-containing *lacZ*$_{sc}$ reporter plasmids (pRV3, pRV4, pOOB114, respectively) individually into OK3 strain. As expected, the plasmids did not cause the production of blue color (X-Gal hydrolysis) in the absence of inducer (cumate). In the presence of 50 mM cumate only the strain carrying CTC-specific reporter (pRV4) turned blue (Fig. 4). The ability of OK3-pRV3 to hydrolyze X-Gal was restored upon introduction of *bldA* (Fig. S12). All available data suggest that *bldA* deletion directly and specifically impaired the translation of TTA codon in *lacZ*$_{sc}$ reporter gene.

Discussion

In this work, we have carried out a set of experiments aimed to study the effects of *bldA* deletion on *S. albus* J1074 and to develop a simple reporter system for further investigation of *bldA*-mediated regulation in J1074.

Perhaps, the most striking property of Bld phenotype (displayed by *bldA* mutants of streptomycetes, vide supra) is its conditionality. That is, the degree of morphological and antibiotic synthesis defects strongly depend on growth conditions and level of transcription TTA codon-containing (TTA^{+}) genes (Guthrie and Chater 1990; Leskiw et al. 1993; Gramajo et al. 1993; Pope et al. 1996; Xu et al. 2008). The conditionality of Bld phenotype of *S. albus* OK3 has extreme manifestation. Indeed, we were able to identify media, where OK3 exhibits impaired morphological development, but under no conditions did we observe complete cessation of aerial mycelium and spore formation. To the best of our knowledge, there is no experimentally validated information on gene networks that govern morphogenesis in *S. albus* J1074. However, as far as we can judge from in silico analysis of J1074 genome (Zaburannyy et al. 2014), the *bld*- and *whi*-cascades are essentially the same in J1074 and other model streptomycetes. Gene *XNR4181* encoding ortholog of *S. coelicolor* AdpA (also known as BldH), a master regulator of colonial differentiation and (often) antibiotic biosynthesis (Chater and Chandra 2008), contains TTA codon, and thus should be under *bldA* control. Hence, either UUA within *adpA* mRNA is mistranslated efficiently enough to allow colonial development, or other (as-yet-unknown) genes compensate/bypass the loss of AdpA. In

Author's personal copy

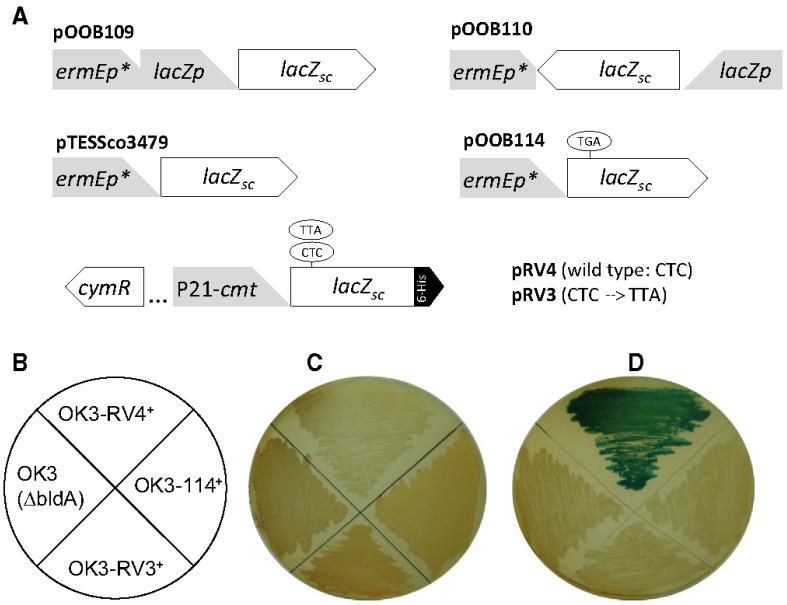


Fig. 4 Codon-specific β-galactosidase reporter for *S. albus* visualizes arrested TTA codon translation in *bldA* mutant OK3. All plasmids (**a**) carrying *S. coelicolor* M145 β-galactosidase gene *SCO3479* ($lacZ_{sc}$) contain actinophage ϕC31 *int-attP* integration module and apramycin resistance marker *aac(3)IV*. They differ in the type of promoter (*shaded gray*) that drives $lacZ_{sc}$ transcription, and presence of distinct codons (shown as *ovals*) in positions 8 and 19 of $lacZ_{sc}$ ORF (if not shown, then CTC codon is present). In pOOB109 and pOOB110, *ermEp* is rendered nonfunctional through truncation of its last 30 nts. In pRV3 and pRV4 $lacZ_{sc}$, transcription is cumate-inducible (Horbal et al. 2014). Four OK3 derivatives (**b**) were streaked onto TSA X-Gal-containing plates in the absence (**c**) and in presence (**d**) of cumate. CTC^+ reporter is expressed in OK3, while TTA^+ is not. The latter is expressed in wild type (SAM2) strain (see ESM, Fig. S11)

Streptomyces, there are several reports on mistranslation of various UUA^+ mRNAs, including *adpA*, in the absence of *bldA*-encoded $tRNA^{Leu}_{UAA}$ (Leskiw et al. 1991; Makitrinskyy et al. 2013). Although existence of a subset of TTA^+ genes capable of translation in *bldA* mutants appears to be paradoxical, an explanation for this was offered (Trepanier et al. 2002). Briefly, propensity for mistranslation is attributed to the nature of the first nucleotide downstream of the TTA codon. If either A or G follow TTA, then the ribosome is able to mistranslate the codon; otherwise (C/T downstream of TTA) translation is aborted. Thus, in the absence of $tRNA^{Leu}_{UAA}$, only mRNA containing the quadruplets TTAC/T will not be translated, and so regulatory effect will be attained. There is TTAC quadruplet in case of *S. albus* J1074 *adpA*; thus, codon context hypothesis does not seem to apply. The position of TTA within the coding sequence could be important. In most streptomycete genes, it is located close to the start codon (Zaburannyy et al. 2009), while in *adpA* gene family, it is positioned roughly in the middle of the ORF. Also, enhancement of UUA translation in *adpA* mRNA in the presence of *S*-adenosylmethionine was reported (Xu et al. 2008). These facts underscore the importance of environmental (rather than genetic) factors in translation of *adpA* mRNA, and they might play a crucial role in shaping the morphology of OK3 strain.

The *bldA* knockout abolished production of two heterologous antibiotics, moenomycin and aranciamycin, the biosynthesis of which is controlled by TTA-containing genes. Moreover, OK3 no longer produced endogenous antibacterial compounds. According to all available data, glycosylated antibiotics, paulomycins, are the principal source of antibacterial activity exhibited by *S. albus* J1074 (Majer and Chater 1987; Olano et al. 2014). We also believe that paulomycins accounted for most (if not all) of the antibacterial activity of SAM2, given the characteristic reddish coloration of *S. albus* lawns on SG2 agar, and the absence of that color in OK3, although detailed analysis was not pursued. Cessation of their production by OK3 strain is most likely due to presence of TTA codon within regulatory gene *SSGH_05342* (*plm30*) encoding LuxR family regulator essential for paulomycin production (Gonzalez et al. 2016). At the same time, level of antifungal activity of OK3 is not affected as compared to SAM2, and even appears to be elevated. *S. albus*

Author's personal copy

genome carries several gene clusters for biosynthesis of compounds that might target eukaryotic cells (Zaburannyy et al. 2014). One of these clusters (around gene *XNR_RS29020*), directing the production of candicidin-like compound, is free of TTA codons, and so could be expressed in *bldA*-minus background.

Our work has once again highlighted all the challenges associated with interpretation of complex, conditional, and composite phenotypes of *bldA* mutants. For example, since TTA codons are embedded into genes for pleiotropic and pathway-specific transcriptional factors, any observable defect of *bldA* mutant would be the net result of deregulated transcription and translation. It is thus difficult to account for and disentangle the relative contribution of each genetic and environmental factor to a final *bldA*-minus phenotype. This calls for development of a simple reporter system where effects of *bldA* knockout on translation can be reliably detected and studied in the most direct way. Several such systems (mostly based on antibiotic resistance genes) were employed in the past, and they gathered valuable information about *bldA* function (Leskiw et al. 1991; Pope et al. 1996). Nevertheless, available reporter genes are not without shortcomings, such as unnatural (for *Streptomyces*) codon usage biases; poor control of transcription; and absence, in most cases, of TTA-free version of the reporter gene. We developed, for *S. albus* J1074, a specialized codon reporter system based on *S. coelicolor* M145 β-galactosidase gene *SCO3479* ($lacZ_{sc}$). It allows tight transcriptional control over *SCO3479* expression, ease of phenotypic detection (in presence of chromogenic substrates), and His-tagged protein tracking. The simplicity of our reporter system lends itself to a rigorous elucidation of the regulatory role of TTA codons at the level of translation, as our initial experiments (vide supra) suggest.

A number of recent studies suggest that it is not the delayed expression of *bldA* gene but formation of mature (translationally competent) tRNA holds the key to its regulatory function (Leskiw et al. 1991; Xu et al. 2008; Pettersson and Kirsebom 2011). Yet, besides initial in silico predictions (Rokytskyy et al. 2016) nothing is known about the genetics of tRNA maturation in *Streptomyces*. Likewise, there is a growing body of data on the crucial role of environmental factors in tRNA biology (Laxman et al. 2013; Moukadiri et al. 2014; Sakai et al. 2016), although this area remains unexplored in case of *bldA* gene. We envision that the above-described reporter system will facilitate the screening of growth conditions and genes that modulate UUA translation efficiency, both in wild type and *bldA*-minus *S. albus* backgrounds.

Acknowledgements This work was supported by grant BG-41Nr from the Ministry of Education and Science of Ukraine (to B.O.).

References

Anagnostopoulos C, Spizizen J (1961) Requirements for transformation of *Bacillus subtilis*. J Bacteriol 81(5):741–746

Baltz RH (2010) *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. J Ind Microbiol Biotechnol 37(8):759–772. doi:10.1007/s10295-010-0730-9

Bilyk B, Luzhetskyy A (2014) Unusual site-specific DNA integration into the highly active pseudo-*attB* of the *Streptomyces albus* J1074 genome. Appl Microbiol Biotechnol 98(11):5095–5104. doi:10.1007/s00253-014-5605-y

Braña AF, Rodríguez M, Pahari P, Rohr J, García LA, Blanco G (2014) Activation and silencing of secondary metabolites in *Streptomyces albus* and *Streptomyces lividans* after transformation with cosmids containing the thienamycin gene cluster from *Streptomyces cattleya*. Arch Microbiol 196:345–355. doi:10.1007/s00203-014-0977-z

Chater K, Chandra G (2008) The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in *Streptomyces*. J Microbiol 46:1–11

Chater KF, Wilde LC (1980) *Streptomyces albus* G mutants defective in the SalGI restriction-modification system. J Gen Microbiol 116(2):323–334

Datsenko K, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97:6640–6645

Fedoryshyn M, Petzke L, Welle E, Bechthold A, Luzhetskyy A (2008) Marker removal from actinomycetes genome using Flp recombinase. Gene 419(1–2):43–47. doi:10.1016/j.gene.2008.04.011

González A, Rodríguez M, Braña AF, Méndez C, Salas JA, Olano C (2016) New insights into paulomycin biosynthesis pathway in *Streptomyces albus* J1074 and generation of novel derivatives by combinatorial biosynthesis. Microb Cell Fact 21:56. doi:10.1186/s12934-016-0452-4

Gramajo H, Takano E, Bibb M (1993) Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor*A3(2) is transcriptionally regulated. Mol Microbiol 7(6):837–845

Gust B, Challis GL, Fowler K, Kieser T, Chater KF (2003) PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci USA 100:1541–1546

Guthrie EP, Chater KF (1990) The level of a transcript required for production of a *Streptomyces coelicolor* antibiotic is conditionally dependent on a tRNA gene. J Bacteriol 172:6189–6193. doi:10.1128/jb.172.11.6189-6193.1990

Herrmann S, Siegl T, Luzhetska M, Petzke L, Jilg C, Welle E, Erb A, Leadlay PF, Bechthold A, Luzhetskyy A (2012) Site-specific recombination strategies for engineering actinomycete genomes. Appl Environ Microbiol 78(6):1804–1812. doi:10.1128/AEM.06054-11

Horbal L, Fedorenko V, Luzhetskyy A (2014) Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. Appl Microbiol Biotechnol 98(20):8641–8655. doi:10.1007/s00253-014-5918-x

Iqbal HA, Low-Beinart L, Obiajulu JU, Brady SF (2016) Natural product discovery through improved functional metagenomics in *Streptomyces*. J Am Chem Soc 138(30):9341–9344. doi:10.1021/jacs.6b02921

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical *Streptomyces* genetics. John Innes Foundation, Norwich

Author's personal copy

King AA, Chater KF (1986) The expression of the *Escherichia coli lacZ* gene in *Streptomyces*. J Gen Microbiol 132(6):1739–1752. doi:10.1099/00221287-132-6-1739

Kuzniar A, van Ham RC, Pongor S, Leunissen JA (2008) The quest for orthologs: finding the corresponding gene across genomes. Trends Genet 24:539–551. doi:10.1016/j.tig.2008.08.009

Lamichhane TN, Blewett NH, Crawford AK, Cherkasova VA, Iben JR, Begley TJ, Farabaugh PJ, Maraia RJ (2013) Lack of tRNA modification isopentenyl-A37 alters mRNA decoding and causes metabolic deficiencies in fission yeast. Mol Cell Biol 33:2918–2929. doi:10.1128/MCB.00278-13

Laxman S, Sutter BM, Wu X, Kumar S, Guo X, Trudgian DC, Mirzaei H, Tu BP (2013) Sulfur amino acids regulate translational capacity and metabolic homeostasis through modulation of tRNA thiolation. Cell 154:416–429. doi:10.1016/j.cell.2013.06.043

Leskiw BK, Lawlor EJ, Fernandez-Abalos JM, Chater KF (1991) TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. Proc Natl Acad Sci USA 88:2461–2465

Leskiw BK, Mah R, Lawlor EJ, Chater KF (1993) Accumulation of *bldA*-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2). J Bacteriol 175:1995–2005. doi:10.1128/jb.175.7.1995-2005.1993

Liu G, Chater KF, Chandra G, Niu G, Tan H (2013) Molecular regulation of antibiotic biosynthesis in *Streptomyces*. Microbiol Mol Biol Rev 77(1):112–143. doi:10.1128/MMBR.00054-12

Lopatniuk M, Ostash B, Luzhetskyy A, Walker S, Fedorenko V (2014) Generation and study of the strains of streptomycetes—heterologous hosts for production of moenomycin. Russ J Genet 50(4):360–365

Luzhetskyy A, Mayer A, Hoffmann J, Pelzer S, Holzenkämper M, Schmitt B, Wohlert SE, Vente A, Bechthold A (2007) Cloning and heterologous expression of the aranciamycin biosynthetic gene cluster revealed a new flexible glycosyltransferase. ChemBioChem 8(6):599–602

Majer J, Chater KF (1987) *Streptomyces albus* G produces an antibiotic complex identical to paulomycins A and B. J Gen Microbiol 133:2503–2507. doi:10.1099/00221287-133-9-2503

Makitrynskyy R, Rebets Y, Ostash B, Zaburannyi N, Rabyk M, Walker S, Fedorenko V (2010) Genetic factors that influence moenomycin production in streptomycetes. J Ind Microbiol Biotechnol 37:559–566

Makitrynskyy R, Ostash B, Tsypik O, Rebets Y, Doud E, Meredith T, Luzhetskyy A, Bechthold A, Walker S, Fedorenko V (2013) Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. Open Biol 3:130121. doi:10.1098/rsob.130121

Moukadiri I, Garzón MJ, Björk GR, Armengod ME (2014) The output of the tRNA modification pathways controlled by the *Escherichia coli* MnmEG and MnmC enzymes depends on the growth conditions and the tRNA species. Nucleic Acids Res 42:2602–2623. doi:10.1093/nar/gkt1228

Myronovskyi M, Rosenkränzer B, Luzhetskyy A (2014) Iterative marker excision system. Appl Microbiol Biotechnol 98(10):4557–4570. doi:10.1007/s00253-014-5523-z

Nguyen KT, Tenor J, Stettler H, Nguyen LT, Nguyen LD, Thompson CJ (2003) Colonial differentiation in *Streptomyces coelicolor* depends on translation of a specific codon within the *adpA* gene. J Bacteriol 185:7291–7296. doi:10.1128/JB.185.24.7291-7296.2003

Olano C, García I, González A, Rodriguez M, Rozas D, Rubio J, Sánchez-Hidalgo M, Braña AF, Méndez C, Salas JA (2014) Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. Microb Biotechnol 7:242–256. doi:10.1111/1751-7915.12116

Ostash B, Campbell J, Luzhetskyy A, Walker S (2013) MoeH5: a natural glycorandomizer from the moenomycin biosynthetic pathway. Mol Microbiol 90:1324–1338. doi:10.1111/mmi.12437

Pettersson BM, Kirsebom LA (2011) tRNA accumulation and suppression of the *bldA* phenotype during development in *Streptomyces coelicolor*. Mol Microbiol 79:1602–1614. doi:10.1111/j.1365-2958.2011.07543.x

Pope MK, Green BD, Westpheling J (1996) The *bld* mutants of *Streptomyces coelicolor* are defective in the regulation of carbon utilization, morphogenesis and cell–cell signalling. Mol Microbiol 19:747–756. doi:10.1046/j.1365-2958.1996.414933.x

Rebets YV, Ostash BO, Fukuhara M, Nakamura T, Fedorenko VO (2006) Expression of the regulatory protein LndI for landomycin E production in *Streptomyces globisporus* 1912 is controlled by the availability of tRNA for the rare UUA codon. FEMS Microbiol Lett 256(1):30–37

Rodriguez L, Oelkers C, Aguirrezabalaga I, Braña AF, Rohr J, Méndez C, Salas JA (2000) Generation of hybrid elloramycin analogs by combinatorial biosynthesis using genes from anthracycline-type and macrolide biosynthetic pathways. J Mol Microbiol Biotechnol 2(3):271–276

Rodríguez A, Olano C, Vilches C, Méndez C, Salas JA (1993) *Streptomyces antibioticus* contains at least three oleandomycin-resistance determinants, one of which shows similarity with proteins of the ABC-transporter superfamily. Mol Microbiol 8(3):571–582

Rokytskyy I, Koshla O, Fedorenko V, Ostash B (2016) Decoding options and accuracy of translation of developmentally regulated UUA codon in *Streptomyces*: bioinformatics analysis. Springerplus 5:982. doi:10.1186/s40064-016-2683-6

Sakai Y, Miyauchi K, Kimura S, Suzuki T (2016) Biogenesis and growth phase-dependent alteration of 5-methoxycarbonylmethoxyuridine in tRNA anticodons. Nucleic Acids Res 44:509–523. doi:10.1093/nar/gkv1470

Sambrook J, Russell DW (2001) Molecular cloning, a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, New-York

Takano E, Tao M, Long F, Bibb MJ, Wang L, Li W, Buttner MJ, Bibb MJ, Deng ZX, Chater KF (2003) A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. Mol Microbiol 50:475–486

Trepanier N, Jensen S, Alexander D, Leskiw B (2002) The positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus* is mistranslated in a *bldA* mutant. Microbiology 148:643–656

Urbonavicius J, Stahl G, Durand JM, Ben Salem SN, Qian Q, Farabaugh PJ, Björk GR (2003) Transfer RNA modifications that alter +1 frameshifting in general fail to affect −1 frameshifting. RNA 9:760–768. doi:10.1261/rna.5210803

Wang J, Schully KL, Pettis GS (2009) Growth-regulated expression of a bacteriocin, produced by the sweet potato pathogen *Streptomyces ipomoeae*, that exhibits interstrain inhibition. Appl Environ Microbiol 75:1236–1242. doi:10.1128/AEM.01598-08

Xu D, Kwon HJ, Suh JW (2008) S-Adenosylmethionine induces BldH and activates secondary metabolism by involving the TTA-codon control of *bldH* expression in *Streptomyces lividans*. Arch Microbiol 189:419–426. doi:10.1007/s00203-007-0336-4

Zaburannyi N, Rabyk M, Ostash B, Fedorenko V, Luzhetskyy A (2014) Insights into naturally minimised *Streptomyces albus* J1074 genome. BMC Genom 15:97. doi:10.1186/1471-2164-15-97

Zaburannyy N, Ostash B, Fedorenko V (2009) TTA Lynx: a web-based service for analysis of actinomycete genes containing rare TTA codon. Bioinformatics 25:2432–2443. doi:10.1093/bioinformatics/btp402

Springer

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s00203-017-1389-7.
Електронний додаток до статті доступний за посиланням:
https://static-content.springer.com/esm/art%3A10.1007%2Fs00203-017-1389-7/MediaObjects/203_2017_1389_MOESM1_ESM.pdf.

Анотація: *Streptomyces albus J1074 є одним із найпопулярніших і найзручніших господарів для гетерологічної експресії генних кластерів, що контролюють біосинтез різних природних метаболітів, таких як антибіотики. Це підживлює інтерес до з'ясування генетичних механізмів, які можуть регулювати вторинний метаболізм у J1074. В цій роботі ми повідомляємо про створення та дослідження мутанта J1074, із нокаутом гена bldA для мРНК$^{Leu}_{UAA,}$ єдиної мРНК, здатної декодувати рідкісний лейциновий TTA кодон у Streptomyces. Делеція bldA в J1074 призвела до умовного* Bld *фенотипу із порушеним утворенням повітряних гіфів і спор на певних агаризованих середовищах. Крім того, bldA мутант J1074 не зміг продукувати ендогенні антибактеріальні сполуки та два гетерологічних антибіотики, моеноміцин і аранціаміцин, біосинтез яких контролюється TTA-вмісними генами. Ми використали нову репортерну систему на основі β-галактозидази, специфічну для кодону TTA, щоб довести, що мутант J1074 bldA має порушення трансляції TTA-кодону. Крім того, ми обговорили можливі причини відмінностей у фенотипах мутантів bldA, описаних тут і в попередніх дослідженнях, надаючи нову інформацію для вивчення bldA-опосередкованої регуляції біосинтезу антибіотиків.*

Molecular Microbiology (2019) 112(1), 249–265

doi:10.1111/mmi.14266
First published online 6 May 2019

Gene *miaA* for post-transcriptional modification of $tRNA_{XXA}$ is important for morphological and metabolic differentiation in *Streptomyces*

Oksana Koshla,[1] Oleksandr Yushchuk,[1] Iryna Ostash,[1] Yuriy Dacyuk,[2] Maksym Myronovskyi,[3] Gunilla Jäger,[4] Roderich D. Süssmuth,[5] Andriy Luzhetskyy,[3] Anders Byström,[4] Leif A. Kirsebom[6] and Bohdan Ostash[1]*

[1] *Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 4 Hrushevskoho st., Lviv 79005, Ukraine.*
[2] *Department of Physics of Earth, Ivan Franko National University of Lviv, 4 Hrushevskoho st., Lviv 79005, Ukraine.*
[3] *Helmholtz Institute for Pharmaceutical Research, Saarland Campus, Building C2.3, Saarbrucken, 66123, Germany.*
[4] *Department of Molecular Biology, Umeå University, 6K och 6L, Sjukhusområdet, Umeå, 90197, Sweden.*
[5] *Institut für Chemie, Technische Universität Berlin, Straβe des 17 Juni 124/TC2, Berlin, 10623, Germany.*
[6] *Uppsala Biomedicinska Centrum BMC, Uppsala University, Husargatan 3, Box 596, Uppsala, 75124, Sweden.*

Summary

Members of actinobacterial genus *Streptomyces* possess a sophisticated life cycle and are the deepest source of bioactive secondary metabolites. Although morphogenesis and secondary metabolism are subject to transcriptional co-regulation, streptomycetes employ an additional mechanism to initiate the aforementioned processes. This mechanism is based on delayed translation of rare leucyl codon UUA by the only cognate $tRNA^{Leu}_{UAA}$ (encoded by *bldA*). The *bldA*–based genetic switch is an extensively documented example of translational regulation in *Streptomyces*. Yet, after five decades since the discovery of *bldA*, factors that shape its function and peculiar conditionality remained elusive. Here we address the hypothesis that post-transcriptional tRNA modifications play a role in tRNA-based mechanisms of translational control in *Streptomyces*. Particularly, we studied two *Streptomyces albus* J1074 genes, *XNR_1074* (*miaA*) and *XNR_1078* (*miaB*), encoding tRNA (adenosine(37)-N6)-dimethylallyltransferase and tRNA (N6-isopentenyl adenosine(37)-C2)-methylthiotransferase respectively. These enzymes produce, in a sequential manner, a hypermodified ms^2i^6A37 residue in most of the A36-A37-containing tRNAs. We show that *miaB* and especially *miaA* null mutant of *S. albus* possess altered morphogenesis and secondary metabolism. We provide genetic evidence that *miaA* deficiency impacts translational level of gene expression, most likely through impaired decoding of codons UXX and UUA in particular.

Accepted 20 April, 2019. *For correspondence: E-mail: bohdan.ostash@lnu.edu.ua; Tel. (+1) 380-032-2394407; Fax (+1) 380-032-3294475.

Introduction

The potential role of post-transcriptional tRNA modifications (PTTMs) as a regulatory device is rapidly gaining interest in biology (Agris *et al*., 2017). Sheer biochemical and genetic complexity of PTTMs pose a formidable challenge to understand them (El Yacoubi *et al*., 2012; Shigi, 2014; Shepherd and Ibba, 2015), even though mutants deficient in PTTMs were first reported almost 50 years ago (Chang *et al*., 1971). Although transcriptional regulation was framed into a straightforward scheme of interaction between *trans*-acting (TF) and *cis*-acting (DNA, operator sequence) elements (Benos *et al*., 2002), a unifying concept of PTTM-mediated regulation has not emerged yet. Particularly, it is often difficult to explain how tRNA could serve as a regulator of specific cellular process if it has to decode thousands of codons in hundreds of genes (Björk *et al*., 1999). Recent developments of genomic and computational biology approaches provide new insights

© 2019 John Wiley & Sons Ltd

and more detailed picture of PTTM-mediated processes in various organisms. To, summarize, the plasticity (in terms of time and/or chemical nature) of modifications of certain tRNAs and non-random usage of codon(s) it reads are suggested to provide necessary level of specificity to a PTTM-based regulatory mechanism (Moukadiri *et al.*, 2014; Deng *et al.*, 2015). In analogy to the transcriptional level, tRNA would be a *trans*-acting element, whereas codon(s) are *cis*-acting ones. The effects of such regulation on the phenotype are subtle, conditional and most likely depend on the degree of enrichment of a coding sequence in *cis*-acting elements (codons) (Aubee *et al.*, 2016). Currently, PTTM-based effects on microbial physiology are far from being fully understood even for the most extensively studied cases (Aubee *et al.*, 2017). More tractable models are desirable, which would help to establish clear causal links between codon usage biases (CUB), PTTMs and the phenotype.

Modification of adenosine residue in position 37 (A37), adjacent to A36 within the anticodon of $tRNA_{XXA}$, was among the first described PTTMs (Schweizer *et al.*, 2017). In *Escherichia coli* modification of A37 is a two-step process. The first one is controlled by MiaA, tRNA (adenosine(37)-N6)-dimethylallyltransferase, and results in N6-isopentenyladenosine (i^6A). The second step is catalyzed by the MiaB protein, tRNA (N6-isopentenyl adenosine(37)-C2)-methylthiotransferase, which produces thiomethylated derivative of i^6A, 2-methylthio-N6-isopentenyladenosine (ms^2i^6A). This is the terminal hypermodified product in *E. coli*, whereas in *Salmonella typhimurium*, ms^2i^6A is a substrate for tRNA 2-methylthio-N6-isopentenyl adenosine (37) hydroxylase, MiaE, that oxidizes the terminal methyl of the isopentenyl chain (Persson and Björk, 1993); MiaE can recognize both i^6A and ms^2i^6A as substrates and in *S. typhimurium* tRNAs both *N6*-(*cis*-hydroxyisopentenyl) adenosine (io^6A) and 2-methylthio-N6-(*cis*-hydroxyisopentenyl) adenosine (ms^2io^6A) could be present, depending on growth conditions. Bacteria deficient in *mia* genes are viable, yet possess certain phenotypic changes. For example, *miaA* mutants of *S. typhimurium* carrying unmodified A37 exhibit decreased growth rates and translation elongation rates (Ericson and Björk, 1986). The *miaE* mutant of *S. typhimurium*, blocked only in last PTTM step (hence A37 is in the ms^2i^6A state), is unable to grow on dicarboxylic acids succinate, fumarate or malate (Persson *et al.*, 1998). In *E. coli miaA* mutation affects translation of several UUX-rich genes. Absence of the MiaA-controlled modification of A37 of $tRNA_{XXA}$ either directly impairs decoding of UUX codons, or it blocks the other PTTMs essential for UUX recognition by the tRNA (Aubee *et al.*, 2016; 2017).

Genetic studies of the members of actinobacterial genus *Streptomyces* perhaps provide the earliest and phenotypically striking example of translational regulation. In 1967, Hopwood mapped the position of genetic marker S48 in *S. coelicolor* chromosome associated with 'bald' (Bld) phenotype. The latter refers to smooth colonies unable to erect aerial mycelium and, often, produce secondary metabolites, in contrast to fuzzy, deeply pigmented wild type colonies (Hopwood, 1967). The mutant allele later was renamed to *bldA1* and shown to affect $tRNA^{Leu}_{UAA}$ gene *bldA* (Merrick, 1976; Lawlor *et al.*, 1987). Most *Streptomyces bldA* null strains exhibit a phenotype similar to the one described above (Hackl and Bechthold, 2015). The Bld phenotype was shown to come from the inability to translate UUA-containing transcripts in the absence of a functional $tRNA^{Leu}_{UAA}$ isoacceptor. Interestingly, only dispensable genes, mostly involved in morphogenesis or secondary metabolism, contain TTA codons, the rarest one within *Streptomyces* genomes (Li *et al.*, 2007; Chater and Chandra, 2008). Translation of those genes is delayed until the culture accumulates sufficient amounts of mature $tRNA^{Leu}_{UAA}$, which coincides with the onset of stationary phase. Hence $tRNA^{Leu}_{UAA}$ might be considered a regulatory device that triggers morphological and metabolic transitions. Both processes are under the control of pleiotropic TF AdpA (Higo *et al.*, 2012). In a vast majority of *Streptomyces* strains the *adpA* gene carries TTA codon (Rabyk *et al.*, 2018), and so translation of the *adpA* transcript might be a focal point of *bldA*-dependent control of morphogenesis and antibiotic production (Higo *et al.*, 2011). However, it remains unknown what factors limit the accumulation of mature $tRNA^{Leu}_{UAA}$. As $tRNA^{Leu}_{UAA}$ transcripts were detected at early stages of growth (Pettersson and Kirsebom, 2011), *bldA* transcription does not seem to be a bottleneck for expression of endogenous Streptomyces genes. Downstream ribonucleolytic 5'-processing of nascent *bldA* transcript might contribute to temporal dynamics of *bldA* accumulation (Leskiw *et al.*, 1993). Although hypotheses on the role of PTTMs in *bldA*-mediated regulation were proposed immediately after the discovery that *bldA* is a tRNA gene (Lawlor *et al.*, 1987), no further studies in this direction have been undertaken. Recently, we have shown that several *Streptomyces* strains carry an extensive repertoire of PTTM genes, including those for A37 hypermodification (Rokytskyy *et al.*, 2016); we also provided initial evidence that reduced level of MiaB-mediated PTTM affects *Streptomyces* biology (Sehin *et al.*, 2018). The latter work has been carried out on transposon mutant of *Streptomyces ghanaensis* ATCC14672 and left many questions unanswered because of low genetic tractability of ATCC14672. This prompted us to perform deeper analysis of PTTMs in *Streptomyces* using *Streptomyces albus* J1074 (also known as *Streptomyces albidoflavus*; Labeda *et al.*, 2014). The J1074 strain has been chosen as an experimental model because of its superior growth characteristics, genetic amenability and popularity as a

platform for natural products discovery (Olano *et al.*, 2014; Lopatniuk *et al.*, 2017; Myronovskyi *et al.*, 2018). Here we show that deficiency in the first step of Mia-controlled pathway leads to significant changes in morphogenesis of *Streptomyces* and its secondary metabolome. These changes are most likely caused by less efficient translation of UUX codons; specifically, we provide genetic evidence for less efficient translation for one of these codons, UUA. As genes for Mia-based A37 PTTMs are widespread in *Actinobacteria*, we predict that similar mechanisms may operate among other members of this class.

Results

Deletions of S. albus miaA *(XNR_1074) and* miaB *(XNR_1078) genes lead to expected changes in the status of A37 PTTM*

Based on previous data (Rokytskyy *et al.*, 2016) we focused on *S. albus* genes *XNR_1074* and *XNR_1078* (hereafter referred to as *miaA* and *miaB* respectively) that encode orthologs of the MiaA and MiaB proteins from proteobacteria. Location of *miaA* and *miaB* in *S. albus* genome is shown in Fig. 1. We replaced the entire coding sequence of *miaA* and *miaB* individually with a hygromycin resistance cassette, and then generated markerless derivatives of knockout strains with the help of site-specific recombinase PhiC31 (Myronovskyi *et al.*, 2014a), as detailed in the *Experimental procedures* section. The generated strains were labelled as Δ*miaA* and Δ*miaB*, their identity on molecular level was confirmed using diagnostic PCR (Fig. S1 in Supporting Information). Disruption of *miaA* is expected to abolish all A37 PTTMs, as it is known (see Introduction) that MiaA-catalyzed A37 isopentenylation is a prerequisite for MiaB-dependent thiomethylation. It was not known at the beginning of our study as to whether *S. albus* possesses MiaE-assisted

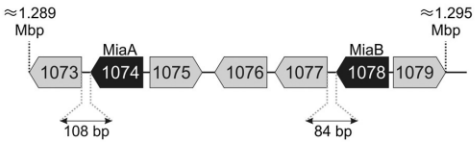


Fig. 1. Genetic organization of *S. albus* J1074 genes for Mia pathway. *XNR_1074* and *XNR_1078* encode functional orthologues of MiaA and MiaB respectively. Numbers above the dashed black lines refer to chromosomal location, numbers below the grey dashed lines – to intergenic region size. *XNR_1073* encodes short (148 aa) hypothetical protein with three transmembrane helices at the C-terminus. According to HHPred structural homology server, XNR_1073 could be amino acid transporter. XNR_1073 orthologues are ubiquitous across *Streptomyces* (blastp results). *XNR_1075* encodes short (110 aa) antitoxin component of MT0933 toxin-antitoxin family. There is no synteny in the gene arrangement around *XNR_1074* and *miaA* in *S. albus* and *E. coli* respectively. Symbols of genes are not drawn to scale.

modification of A37; no orthologs of proteobacterial MiaE protein were identified in BLASTP searches (Rokytskyy *et al.*, 2016). HPLC analysis of tRNA hydrolysates from the parental SAM2 strain (derivative of J1074 carrying single *attB*$^{\varphi C31}$, see *Experimental procedures*) revealed the presence of two major fractions – ms^{2}io^{6}A and io^{6}A (Fig. 2), pointing to the presence of MiaE activity. In agreement with what is known for enterobacteria, total tRNA hydrolyzates from Δ*miaA* were devoid of isopentenylated and thiomethylated adenosine, whereas Δ*miaB* samples predominantly contained io^{6}A (Fig. 2). Based on the *in silico* and HPLC data, we propose a pathway to hypermodified A37 in *Streptomyces albus* tRNA$_{XXA}$ (Fig. 3).

Morphology and growth of ΔmiaA *and* ΔmiaB *differs from the parental strain*

The mutants were incubated on a number of rich and minimal media and analysed during 4 days of growth. For the Δ*miaA* strain visible growth defects were observed (Figs S2 and 4). As compared to SAM2, aerial hyphae formation was retarded in Δ*miaA* on TSB, SFM, ISP3 and MM agars, which, as a consequence, delayed spore formation. While SAM2 started forming white aerial mycelium at the end of the first day of growth on SFM medium, Δ*miaA* mutant did so 24 h later. Chains of spores were observed on the second day of growth for SAM2, while Δ*miaA* did not form spores even after 3 days. Morphological defects of Δ*miaA* were restored to the wild type upon introduction of a plasmid carrying a single copy of *miaA* under control of constitutive promoter *ermEp* (plasmid pTES1074, Fig. 4B). This ruled out the possibility that the observed phenotype is due to unanticipated polar effects or rearrangements due to *miaA* deletion. We observed no significant phenotypic differences between Δ*miaB* and SAM2 strains when growing them on various media, including SFM (Fig. 4A). On R5 agar, however, SAM2 formed a confluent lawn of aerial mycelium and spores after 3 days of growth, while Δ*miaB* showed a patchy pattern of sporulation; on the fourth day of growth both strains looked indistinguishable.

Growth pattern of the *mia* mutants in liquid culture (TSB) was consistent with what we observed on agar plates (Fig. 4C). Particularly, Δ*miaA* exhibited a delayed peak of biomass accumulation (~72 h) as compared to SAM2 (peaked at 48 h). The Δ*miaB* mutant and SAM2 growth curves were essentially the same.

Genes *miaA* and *miaB* were introduced individually into SAM2 genome (plasmids pTES1074 and pTES1078 respectively) under the control of *ermE* promoter. The generated transconjugants did not differ from parental strain in their morphology and antibiotic activity against

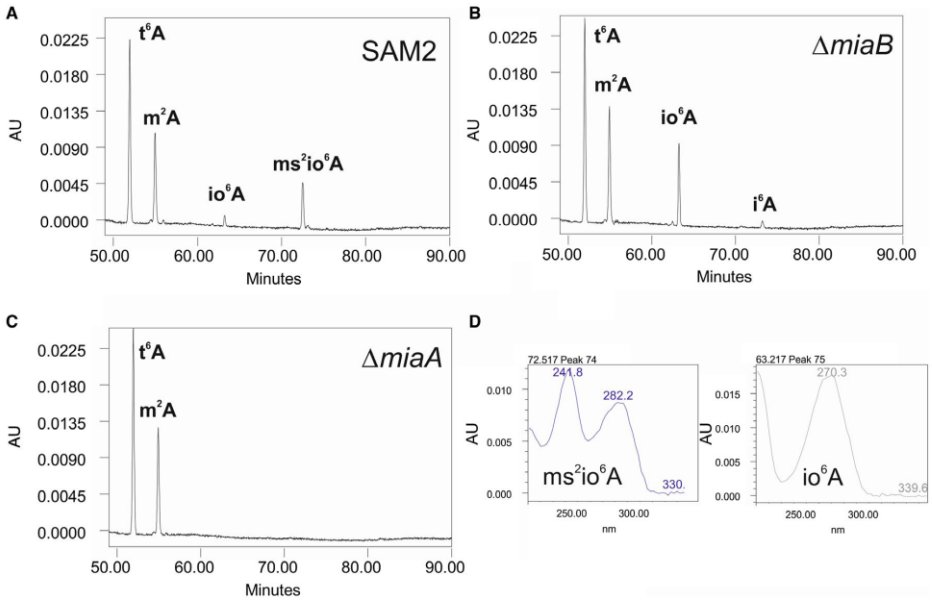


Fig. 2. Knockouts of *XNR_1074* and *XNR_1078* abolish Mia-dependent tRNA modifications. HPLC traces of tRNA hydrolysates from *S. albus* SAM2, Δ*miaB* and Δ*miaA* strains.
A. SAM2 carries final products of A37 Mia-dependent modifications, *N6*-(*cis*-hydroxy-isopentenyl) adenosine (io^6A) and 2-methylthio-*N6*-(*cis*-hydroxyisopentenyl) adenosine (ms^2io^6A). No free i^6A or ms^2i^6A are seen probably because of their complete conversion to oxidized forms (see above).
B. Knockout of *XNR_1078* stops methylthiolation of i^6A and causes enrichment in io^6A.
C. *XNR_1074* null mutant is devoid A37 PTTMs. Traces of the other A37 modifications are present, such as *N6*-threonylcarbamoyladenosine (t^6A) and 2-methyladenosine (m^2A). The tRNA hydrolyzates were run on commercially available io^6A- and ms^2io^6A-calibrated column and modifications were identified by comparing their retention time and UV spectra with the published data (see D and Persson *et al.*, 1998).

Bacillus cereus and *Debaryomyces hansenii* (data not shown).

The mia mutations affect secondary metabolism of S. albus

S. albus SAM2 genome harbours more than 20 gene clusters for the production of secondary metabolites, of which a minority is expressed to a level high enough to be detected by bioassays or HPLC-MS. The antibacterial activity of SAM2 can be attributed mainly to accumulation of paulomycins (Olano *et al.*, 2014), while candicidins, antimycins and macrolide tetramates account for the activity of this strain against eukaryotes, such as fungi (Ahmed *et al.*, 2017). SAM2 was grown along with both *mia* mutants on appropriate agar plates, where antibiotic activities could be detected with a simple agar plug method. The Δ*miaB* and SAM2 did not differ significantly in their activity against *B. cereus* in contrast to Δ*miaA*, which showed almost no antibacterial activity. At the same time, both *mia* mutants exhibited a significant increase in antifungal activity against *D. hansenii* (Fig. 5A). The bioassays suggested that the secondary metabolism of *mia* mutants has undergone profound changes. To gain more insight into the latter, we cultivated the SAM2, Δ*miaA* and Δ*miaB* strains in liquid media SG2 and analysed the metabolomes by HPLC-MS. In the supernatant and biomass extracts we reproducibly observed molecular mass peaks corresponding to six classes of compounds known to be accumulated by *S. albus*; a typical LC-MS result is shown on Figs 5B and S3. Interestingly, among those six classes we detected surugamides, previously described as cryptic metabolites (Xu *et al.*, 2017). Fractions corresponding to antimycins were increased in Δ*miaA* as compared to SAM2 strain, while candicidin amounts produced by the mutants did not exceed that accumulated by SAM2 (Fig. S4). Novel mass-peaks, not observed in SAM2, accounted in Δ*miaA* extracts for fractions with retention times of 6.49 and 6.78 min (Fig. S5).

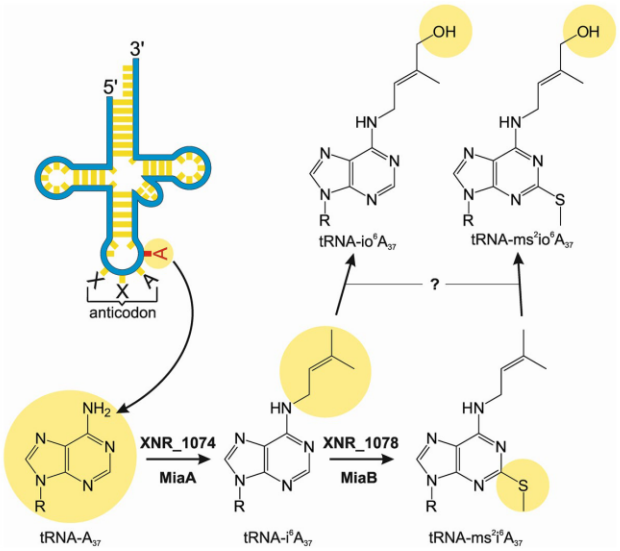


Fig. 3. Mia-controlled pathway of A37 tRNA modifications in *S. albus*. Mia-proteins are involved in modification of A37 position of most tRNAs that feature A36A37. First reaction is controlled by XNR_1074 (MiaA orthologue) and the product is *N6*-isopentenyladenosine (i^6A). The second step is formation of 2-methylthio-*N6*-isopentenyladenosine (ms^2i^6A) via action of MiaB orthologue XNR_1078. Genetic control of hydroxylation of i^6A and ms^2i^6A in *S. albus* remains elusive. [Colour figure can be viewed at wileyonlinelibrary.com]

The effect on secondary metabolism due to deleting *miaB,* and in particular *miaA,* resembled $tRNA^{Leu}_{UAA}$-deficient mutant, OK3, as previously described (Koshla *et al.*, 2017), suggesting an involvement of UUA codon. Furthermore, we introduced TTA codon-carrying gene clusters for biosynthesis of aranciamycin (cosmid pOJ436ara) and moenomycin (cosmid moeno38-6) into the SAM2 and *mia* mutants. Compared to SAM2 the production of both antibiotics was reduced in Δ*miaB* and undetectable in the Δ*miaA* strain (Fig. 6).

Response of mia *mutants to elevated temperature and antibiotic stress*

SAM2 and the *mia* mutants were challenged with two kinds of stress: elevated temperature and exposure to antibiotics. Particularly, we monitored the growth *S. albus* strains on different solid media at 30, 37 and 42°C. It has to be noted that growth of the parental strain was markedly affected at 42°C on many media. Nevertheless, on R5 medium Δ*miaA* growth was retarded compared to SAM2, while on several minimal agars the mutant grew where SAM2 growth was not observed (Fig. S6). We also observed an accumulation of red pigment by Δ*miaA* while SAM2 produced no coloured compounds. Using an agar plug assay (*vide supra*), we revealed that the red pigment had no antibiotic activity, and most probably reflects the production of carotenoids, the known cryptic secondary metabolites of *S. albus* (Myronovskyi *et al.*, 2014b). No significant differences between SAM2 and Δ*miaB* growth were observed under the conditions of this study. Furthermore, there were no significant differences between SAM2 and the mutants in their antibiotic resistance profiles (Table S1 in Supporting Information).

Codon-specific reporters support the idea that expression of TTA-containing genes in mia *mutants is impaired at the translation level*

The important target for $tRNA^{Leu}_{UAA}$ (*bldA*) is the pleiotropic transcriptional factor *adpA* transcript, which carries single UUA codon (Nguyen *et al.*, 2003; Takano *et al.*, 2003). To investigate whether translation of the *adpA* mRNA is impaired in the Δ*miaA* and Δ*miaB* strains, we used an *adpA-gusA* (β-glucuronidase) reporter system (Makitrynskyy *et al.*, 2013). The data revealed no difference in *adpA* transcription between SAM2 and Δ*miaA*/Δ*miaB* mutants. Translation of *adpA* was reduced by ~30% in Δ*miaA* compared to SAM2 (Fig. 7). Hence, *miaA* deficiency affects *adpA* expression at the level of translation.

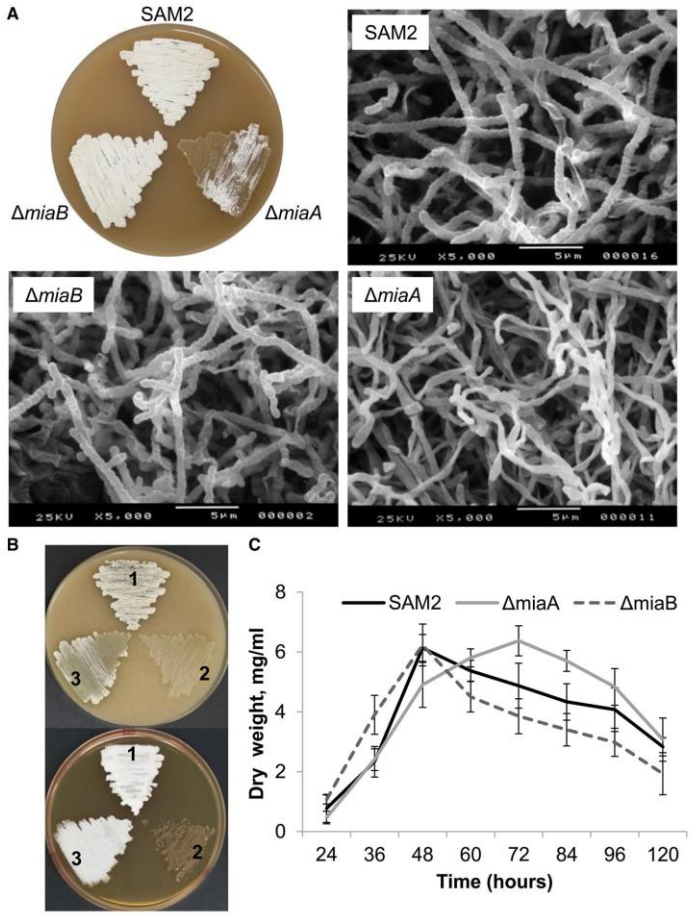


Fig. 4. Growth characteristics of *S. albus mia*-mutants.
A. Knockout of *miaA* caused impairment of sporulation. Parental strain (SAM2), Δ*miaA* and Δ*miaB* mutants were grown on SFM plate for 2 days at 30°C; the same lawns were used to obtain scanning electron microscopic images.
B. *XNR_1074* complements Δ*miaA* morphological phenotype. Strains were grown on ISP3 (upper plate) and TSB (lower plate) agar for 2 days at 30°C. 1 – SAM2 pTES; 2 – Δ*miaA* pTES; 3 – Δ*miaA* pTES1074 (complemented strain).
C. Δ*miaA* has prolonged early stages of growth and later reaches biomass accumulation peak, compared to the parental strain, while Δ*miaB* showed no significant difference. Growth curve (via dried weight accumulation) was determined in TSB medium at 30°C; $n = 4$. Cultures started from same amount of spores (approx. 10^7 per 30 ml of medium). Error bars indicate standard deviation; p-value < 0.05. [Colour figure can be viewed at wileyonlinelibrary.com]

Causal link between inability of Δ*miaA* mutant to express certain genes and presence of TTA codon in them has been further addressed using the *gusA*-based TTA-specific reporter. For this purpose, we constructed the plasmid pSAGA, where *gusA* was fused to the strong promoter of the apramycin resistance gene *aac(3)IV*. In pSAGA, the second leucyl codon CTG was replaced with TTA to generate pSAGAt. While β-glucuronidase (GusA) activity of SAM2 pSAGAt$^+$ was lower than that of SAM2 pSAGA$^+$ on early stages of growth (19 and 25 h), GusA levels reached a plateau during the late stage of growth. For the *miaA* mutant the onset of TTA$^+$ GusA expression was significantly delayed and reduced in comparison to the CTG$^+$ reporter, both on plates and in liquid culture (Fig. 8).

The effects of mia *mutations on* S. albus *cannot be explained solely by their impact on UUA codons*

The Mia-pathway is known to control PTTM of a large family of tRNAs (Schweizer *et al.*, 2017), yet not all of them

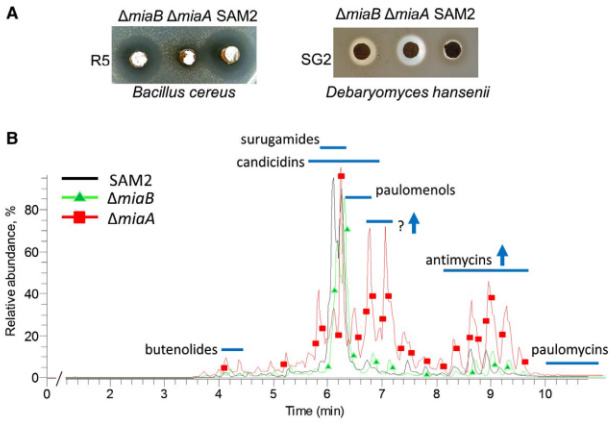


Fig. 5. Absence of Mia-controlled PTTMs affects secondary metabolism of *S. albus*.
A. Agar plug bioassay of *B. cereus* and *D. hansenii* growth inhibition by endogenous antibiotics produced by parental strain (SAM2) and its *mia* mutants. Streptomycetes were grown on R5 and SG2 solid media for 5 days prior to the assay.
B. Overlaid HPLC-MS traces of butanol extracts from supernatant of *S. albus* strains grown in SG2 broth at 30°C for 5 days. Retention time range of identified families of natural products and two unknown fractions (labeled with interrogation sign) is indicated with thick blue lines. Poor peak overlap within 5.5-6.5 min range is because different candicidin species are produced in different quantitites by the studied strains. Blue arrows mark reproducibly and significantly more abundant fractions in *miaA* mutant as compared to SAM2. Please see Figs. S3 and S4 for more information on these fractions. The images in this figure represent typical result of four independent bioassay and LC-MS experiments. [Colour figure can be viewed at wileyonlinelibrary.com]

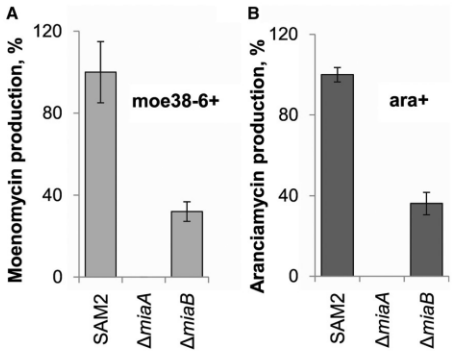


Fig. 6. Accumulation of moenomycin (A) and aranciamycin (B) is abolished in Δ*miaA* and decreased in Δ*miaB*. For A and B 100% indicate wild type (SAM2) production levels (approx. 0.5 mg of nosokomycin A2 per 1L, for moeno38-6+ SAM2 and aranciamycin productivity index of 1.8 for pOJ436ara$^+$ SAM2; see Koshla *et al.*, 2017). Antibiotic production is normalized against the same amounts of biomass. Error bars represent standard deviation; p-value < 0.01 (n = 4).

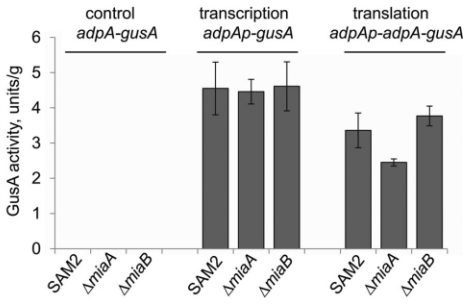


Fig. 7. Expression of TTA-codon containing *adpA* gene is affected in *miaA*-minus mutant at the level of translation. Glucuronidase (GusA) activity of cell-free extracts of *S. albus* strains carrying padpAcontrol (promoterless *adpA-gusA* fusion), padpAscript (gene *gusA* under control of *adpAp*) and padpAtransl (*adpAp-adpA-gusA* fusion) constructs. Strains were grown in TSB for 48 h. Error bars represent standard deviation based on three biological replicates; p-value < 0.01 (one-way ANOVA test).

appear to be equally affected in their decoding capacities in hypomodified state (Aubee *et al.*, 2016). Rare codons (such as UUA in *Streptomyces*) and/or those having single Mia-dependent cognate tRNA would suffer more in the absence of MiaA. Out of eight MiaA-modified anticodons (UAA, CAA, GCA, UGA, CGA, GAA, CCA and QUA) UAA is the only one composed exclusively of U/A bases and has the weakest energy of codon recognition among all aforementioned anticodons (Grosjean and Westhof, 2016). UAA would be especially vulnerable to the loss of PTTMs known to increase the specificity and efficiency of codon–anticodon interactions. We therefore looked for the evidence that codons other than UUA are sensitive to the loss of A37 PTTM. This issue was approached in two ways. If the $tRNA^{Leu}_{UAA}$ gene *bldA* is the only means by which Mia pathway affects morphogenesis, then *bldA*

deletion in Δ*miaA* background should not cause more severe morphological defects than single mutations do. We generated Δ*miaA*Δ*bldA* double mutant of *S. albus* (Fig. S7) and found that it fails to erect aerial hyphae even after 120 h of incubation. In contrast, the initial single mutants were capable, to varying extent, of escaping morphological arrest in the substrate mycelium phase of growth (Fig. 9).

Expression of the *adpA* gene in *Streptomyces* leads to a simple visual phenotype in the form of the production of white aerial hyphae. We took advantage of this fact to provide one more line of evidence that the UUA codon is not the only target of Mia-pathway in our study. For this purpose we designed a TTA-free version of SAM2 *adpA* gene (plamid pGM4181tta carrying TTACTC substitution, driven from the strong *adpA*-regulated promoter *moeE5p* (Makitrynskyy *et al.*, 2013)) and expressed it, in parallel with native *adpA* (pGM4181), in the Δ*miaA* strain. If the morphological defects of the latter stem from inefficient translation of UUA in *adpA* mRNA, then we should be able to complement the Δ*miaA* mutant with a TTA-free version of *adpA*. Indeed, we observed the partial restoration of aerial mycelium formation in Δ*miaA*, while a plasmid-borne copy of native *adpA* had only slight effects on the mutant (Fig. 10). Taken together, results of this section imply that less efficient translation of codons other than UUA also contributes to the Δ*miaA* mutant phenotype.

Discussion

Although the structure and molecular mechanisms of MiaA, MiaB and MiaE enzymes are thoroughly studied (Pierrel *et al.*, 2004; Mathevon *et al.*, 2007; Schweizer *et al.*, 2017), the exact chain of cause-to-consequence events linking certain PTTM (or its absence) to a physiological effect are less well understood (Persson *et al.*, 1998; Armengod *et al.*, 2014; Aubee *et al.*, 2016). Actinobacteria, members of the largest bacterial taxon of great industrial and medical value, are also virtually not studied in this regard (Sehin *et al.*, 2018). Here, we identify genes for hypermodification of A37 residue of $tRNA_{XXA}$ in *Streptomyces albus* J1074 and show that failure to install these PTTMs lead to specific perturbations of morphology and antibiotic production. Below we attempt to put our findings into the context of current knowledge of the biology of PTTMs.

Mia-controlled A37 PTTMs are widespread in bacterial and eukaryotic tRNAs, although the degree of modification (and underlying enzymatic machinery) differs. Only the first step (i^6A) occurs in yeast; two-step modification leads to ms^2i^6A in *E. coli*; three enzymes in *Salmonella* result in hypermodified ms^2io^6A, where the oxygenase MiaE catalyzes the terminal (hydroxylation) step. As in *S. ghanaensis*, ms^2io^6A PTTM is present in *S. albus*. A recent metabolomic study also confirmed the presence of io^6A in *Mycobacterium* (Chionh *et al.*, 2016), reinforcing

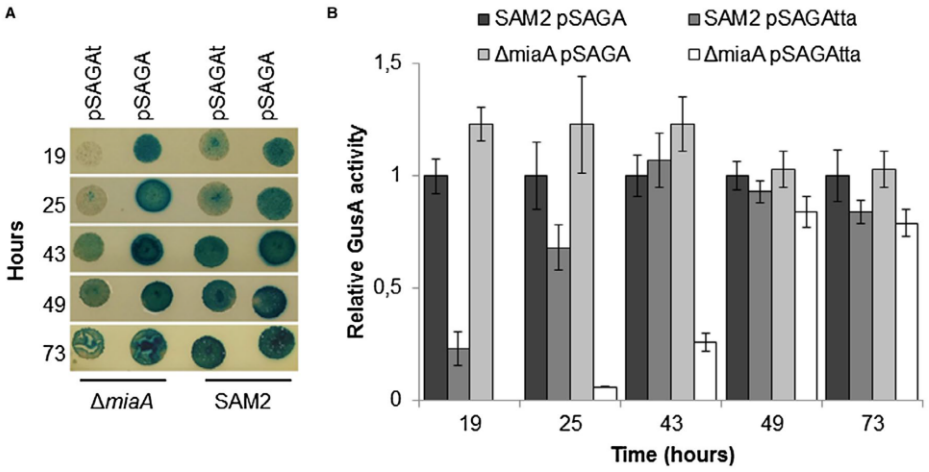


Fig. 8. Expression of TTA-specific reporter gene is affected in Δ*miaA* strain.
A. Spot test of glucuronidase activity of Δ*miaA* and parental strains. Equal amounts of spores of SAM2 and Δ*miaA*, carrying pSAGAt (TTA+) and pSAGA, were plated on TSA medium, incubated for specified time, then X-gluc solution was applied on top of the spot. Blue colour indicates GusA activity.
B. Relative GusA activity of Δ*miaA* carrying pSAGAt (TTA+) is noticebly lower at early stages of growth. Strains SAM2 and Δ*miaA*, carrying pSAGAt (TTA+) and pSAGA, were grown in liquid TSB culture. GusA activity of SAM2 $pSAGA^+$ strain was taken as 1 for all time points. Error bars represent standard deviation based on three biological replicates; *p*-value < 0.05 (one-way ANOVA test). [Colour figure can be viewed at wileyonlinelibrary.com]

© 2019 John Wiley & Sons Ltd,

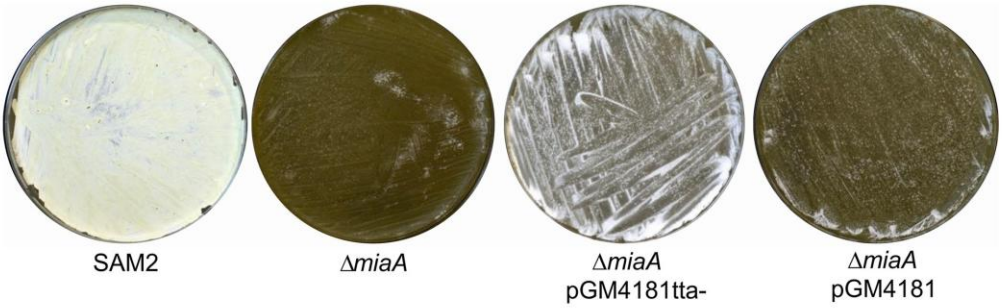


Fig. 9. Double *miaA-bldA* knockout completely derails aerial hyphae formation. Equal amounts of spores (approx. 10^6) were used to start the lawns. Strains were grown for 120 h on agar media. A longer incubation time (compared to Fig. 10) has been used for a better visibility of the terminal phenotype. Photographs represent typical result of four independent cultivations. [Colour figure can be viewed at wileyonlinelibrary.com]

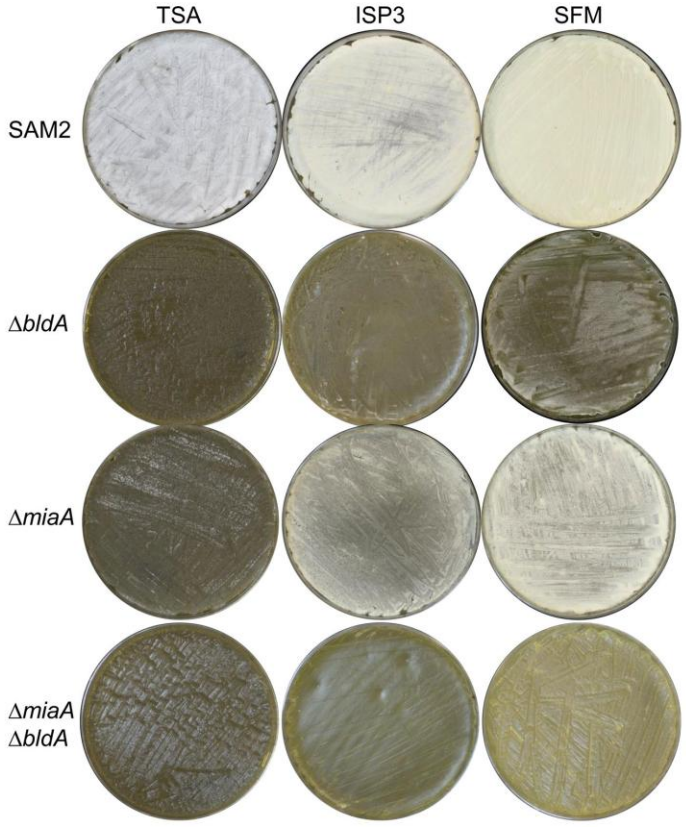


Fig. 10. Codon TTA of *adpA* gene significantly contributes to morphological defects of Δ*miaA* strain. Equal amounts of spores (approx. 10^6) were used to start the lawns. Strains were grown for 45 h on SFM agar. Photographs represent typical result of four independent cultivations. [Colour figure can be viewed at wileyonlinelibrary.com]

the suggestion that MiaE-controlled hypermodification is typical for actinobacteria. However, the genetic control of ms^2i^6A hydroxylation remains enigmatic.

While the reduced level of MiaB-controlled thiomethylation led to significant phenotypic perturbations in *S. ghanaensis* (Sehin *et al.*, 2018), in *S. albus miaB* deletion has minimal effects. We do not understand this difference at the moment. Both different nature of mutations and host's genotypes may account for the observed differences, emphasizing the importance of further studies of PTTM effects under different conditions. In *S. albus* the *miaA* and *bldA* mutants are phenotypically quite similar (Koshla *et al.*, 2017). Both are incapable of a timely and complete formation of sporulating hyphae and they basically do not produce antibiotics whose genes carry UUA codons. Morphogenesis in streptomycetes is under the control of the TF AdpA (Nguyen *et al.*, 2003; Takano *et al.*, 2003), whose gene also carries single UUA codon and, as our reporter systems demonstrated, is poorly translated in the Δ*miaA* strain. Nevertheless, complementation of *miaA* mutant with TTA-free version of *adpA* and analysis of double Δ*miaA*Δ*bldA* mutant shows that codons other than TTA also contribute to the Δ*miaA* phenotype. In *E. coli*, the MiaA-mediated PTTM is known to be present in A36-A37-containing tRNAs for Cys, Leu, Ser, Phe, Trp, Tyr (Dal Magro *et al.*, 2018). All of these tRNAs would be devoid of $ms^2io^6A_{37}$ modification in *S. albus* Δ*miaA*, although it is not known whether all aforementioned tRNA are equal contributors to Δ*miaA* phenotype. As the only tRNA with U/A-based anticodon among MiaA-dependent tRNAs, *bldA* might be especially susceptible to the loss of PTTM; similar losses for anticodons featuring at least one 'strong' (G/C) base would be less detrimental for codon–anticodon recognition (Grosjean and Westhof, 2016). In *E. coli*, an elevated TTX *ratio* (a fraction of TTX among all leucyl codons) was proposed to be a marker for MiaA dependency (Aubee *et al.*, 2016). We believe that the mentioned hypothesis does not fully apply to the *Streptomyces* case due to several observations. First, and probably the most illustrious argument comes from analysis of candicidin and antimycin biosynthetic gene clusters of *S. albus*, lacking TTA codons. The 32-kb gene *XNR_RS29035* encodes giant candicidin polyketide synthase, one of the key enzymes involved in antibiotic biosynthesis, containing 55 TTG codons. Several antimycin biosynthesis genes carry from 1 to 4 TTG. Apparently, decreased translation of such a large number of codons will eventually lead to significant decrease, or even complete cessation of antibiotic accumulation. Yet, the production of respective antibiotics by Δ*miaA* is not decreased, and even increased in case of antimycins. Second, although the TTX ratio for the *adpA* gene is above the background level (0.022), it is due to the presence of only one TTA codon and no TTG. Its translation is affected in Δ*miaA*. See also Fig. S8 for more data on TTX usage in *S. albus* genes mentioned in this study. We therefore assume that not all MiaA-dependent tRNAs are equally impaired in reading their cognate codons in Δ*miaA* background. The decoding of TTA codon appears to be MiaA-tunable, although other as-yet-undefined codons clearly are read less efficiently. As codon usage is influenced by GC content, the latter has to be taken into consideration when applying the UUX codon model for MiaA sensitivity across the prokaryotic domain. More research that directly measures translation efficiency is needed to define true extent of tRNAs and codons affected in the absence of *miaA* in *Streptomyces*.

It is unclear why Δ*miaA* produces more antimycins and novel compounds as compared to parental SAM2 strain. At present we are not aware of any reports, where deficiency in modification of natural tRNA would lead to improved expression of a certain trait (however, precedents are known for heterologous tRNA that mediate incorporation of nonstandard amino acids in *E. coli*; Crnković *et al.*, 2018). As in the related species *S. albus* S4, antimycin and candicidin gene clusters are grouped together in J1074 genome (see Fig. S3); organization of genes within these gene clusters is the same in S4 and J1074. For the S4 strain antimycin and candicidin clusters are co-regulated by a LuxR family regulator (McLean *et al.*, 2016), and a similar scenario most likely applies to J1074. Thus, if the effects of *miaA* mutation on antimycin biosynthesis are mediated by upregulation of the LuxR regulator XNR_RS29090 (=XNR_5869), then it would be reasonable to expect increased amounts of candicidins as well. This is not the case. A tentative suggestion would be that *miaA* deficiency impairs or arrests the expression of certain MiaA-dependent repressors of secondary metabolism, which leads to observed changes in antibiotic production.

Dynamic and often the reversible nature of PTTMs allows for a continuum of regulatory opportunities for the cell via selective translation of transcripts (Deng *et al.*, 2015). PTTMs are therefore considered as a part of the mechanisms that shape conditional phenotypes in microorganisms (Chionh *et al.*, 2016; Jaroensuk *et al.*, 2016; Yu *et al.*, 2017). Molecular mechanisms underlying these phenomena are beginning to emerge (Sakai *et al.*, 2016). For example, MnmEG complex may utilize either ammonium or glycine for modification of U34 of tRNA decoding NNA codons, and, in *E. coli*, the pathway that MnmEG would take is dependent on growth conditions (Moukadiri *et al.*, 2014). Similar processes in *Streptomyces* could also contribute to a lagged accumulation of translationally competent $tRNA^{Leu}_{UAA}$.

We suggest that *S. albus* is a simple and genetically tractable model to study biological roles of PTTMs in the ways hardly attainable in other experimental systems. For example, extreme rarity of TTA codon and the dispensability of TTA-containing genes allow for the deletion of

$tRNA^{Leu}_{UAA}$ gene and, in principle, for a complete removal of TTA codon from the genome. As TTA is present only 78 times in the whole 6.8-Mbp *S. albus* genome, it is technically simpler as compared to the formidable task of reassignment/removal of hundreds of codons in other species. Defects in PTTMs are associated with clearly observable phenotypes, either simple (such as those provided by reporter genes), or complex (antibiotic production or morphological transitions).

Results presented in this work raise a number of new questions. First, it is not known as to whether impaired A37 isopentenylation is the direct cause of the observed defects in *mia* mutants. The i^6A is a prerequisite for U34 ribose methylation by TrmL (Hori, 2017); absence of the latter modification in *miaA* mutant of *E. coli* is among the true reasons for reduced expression of *rpoS* (Aubee *et al.*, 2017). HPLC analysis revealed no methylated U34 in *S. albus* (data not shown), yet U34m is a minor modification and might be missed. Besides TrmL, MnmEG (*vide supra*) might be important for the function of $tRNA^{Leu}_{UAA}$. MnmEG-catalyzed modifications were not observed as well, probably for the same reasons as above. *Streptomyces* orthologs of TrmL and MnmEG could not be identified through simple BLASTP searches, necessitating the use of more sophisticated computational and experimental tools. Genome mining for the missing PTTM enzymes, more sensitive MS-based assays of *Streptomyces* tRNome and structural elucidation of $tRNA^{Leu}_{UAA}$ are thus considered primary research topics by us.

Experimental procedures

Microorganisms and culture conditions

Microorganisms and plasmids used in this study are listed in Table 1. *Streptomyces albus* SAM2, a derivative of J1074 with deletion of pseudo-$attB^{\varphi C31}$ (thus enabling single copy integration of actinophage φC31-based vectors) (Bilyk and Luzhetskyy, 2014) was used to prepare knockouts of *XNR_1074* and *XNR_1078*. *Escherichia coli* DH5α used for routine subcloning; WM6026 – for conjugal plasmid transfer to *S. albus*; BW25113 (pKD46) – for construction of knockout plasmids via recombineering (Datsenko and Wanner, 2000; Gust *et al.*, 2003). For all procedures with *E. coli* strains standard growth conditions were used (Sambrook and Russell, 2001). For intergeneric mating *S. albus* strains were grown on soy-flour mannitol (SFM) agar at 30°C (Kieser *et al.*, 2000). *Bacillus cereus* ATCC19637 and *Debaryomyces hansenii* VKM Y-9 were used as antibiotic sensitive test cultures (grown at 30°C). For examination of growth phenotype and antibiotic activity, *S. albus* strains were grown on media: SFM, MM (with glucose – MMG; with mannitol – MMM), SMMS, R5 (Kieser *et al.*, 2000), SG1, SG2, GYM, ISP3 and ISP5 (Koshla *et al.*, 2017). For moenomycin and aranciamycin production, moeno38-6+ and pOJ436ara+ *S. albus* strains, respectively, were grown in liquid medium SG1 at 30 °C for 5 days. Where needed, strains were maintained in presence of apramycin (25 μg ml^{-1}), ampicillin (100 μg ml^{-1}), kanamycin (50 μg ml^{-1}), thiostrepton (100 μg ml^{-1}), hygromycin (100 μg ml^{-1}) and phosphomycin (25 μg m^{-1}).

DNA constructs, construction of plasmids and generation of S. albus *recombinant strains*

All primers used in this study are listed in Table 2. BAC 1N12d_bldA (Koshla *et al.*, 2017) was used for *bldA* knockout. Suicide vector pKC1132 was used to prepare knockout plasmids for *miaA* (*XNR_1074 = XNR_RS05320*) and *miaB* (*XNR_1078 = XNR_RS05340*). Aforementioned genes were amplified with approximately 2-kb flanking arms from *S. albus* SAM2 chromosome with primers 1074_XbaI_up and 1074_EcoRI_rp (4.7-kb product), 1078_XbaI_up and 1078_EcoRI_rp (5.4-kb amplicon) respectively. The amplicons were digested with XbaI and EcoRI restriction endonucleases and cloned into XbaI-EcoRI sites of pKC1132 yielding plasmids pKC1074 and pKC1078. Hygromycin resistance cassette *hyg*, flanked with P-GG and B-CC sites, was amplified from patt-shyg (Myronovskyi *et al.*, 2014a) with primers 1074_uprd and 1074_rprd, 1078_uprd and 1078_rprd. Resulting 1359-bp PCR products were used to replace *miaA* and *miaB* genes in pKC1074 and pKC1078, respectively, using recombineering. As a result, plasmids pKC1074::hyg and pKC1078::hyg were constructed. These and other constructs generated in course of the work were verified by PCR and sequencing. To construct plasmids pTES1074 and pTES1078 (*miaA* and *miaB* genes, respectively, under control of strong promoter *ermEp*), genes *XNR_1074* and *XNR_1078* were amplified with 1074_XbaI_upcl and 1074_EcoRI_rpcl (1042-bp product), 1078_XbaI_upcl and 1078_EcoRI_rpcl (1742-bp product) primers, respectively, from SAM2 chromosome. Resulting amplicons were digested with XbaI and EcoRI restriction endonucleases and cloned downstream of *ermEp* into pTES. To construct pSAGA, *gusA* was amplified from pGUS (Myronovskyi *et al.*, 2011) with gusA_F and gusA_R primers and cloned into EcoRI and EcoRV sites of pSETPAm. pSAGAt was constructed similarly, however, the forward primer (gusA_tta_F) was designed to replace second leucyl codon CTG of *gusA* with synonymous one TTA.

The *adpA*-expressing construct pGM4181 is described in (Yushchuk *et al.*, 2018). Its TTA-free derivative pGM4181tta- was constructed as follows. A synthetic *XNR_4181* gene was ordered (General Biosystems Inc., USA), where a TTA codon (673-675 bp) was replaced with the synonymous CTC codon. Synthetic gene was amplified with primers xnr4181_XbaIup and xnr4181_EcoRIrp (Rabyk *et al.*, 2018), digested with XbaI and EcoRI restriction endonucleases and subcloned into SpeI/EcoRI recognition sites of pmoeE5script, yielding $pGM4181tta^-$. The plasmid was verified with restriction mapping and sequencing.

Intergeneric conjugation *E. coli* – *Streptomyces* (Luzhetskii *et al.*, 2001) was used for plasmid transfer into *S. albus*. The *miaA* and *miaB* knockout strains of *S. albus* were selected for hygromycin resistance (gene replacement with *hyg*) and apramycin sensitivity (loss of vector sequences). Hygromycin resistance cassette flanked with B-GG and

Table 1. Strains and plasmids used in this work.

Strain/Plasmid	Characteristics	Source/Reference
Strains		
E. coli DH5α	*supE44* Δ*lacU169(Φ80ΔlacZM15) glnV44 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	NEB
E. coli WM6026	*lacI*^{q} *rrnB3* Δ*lacZ4787 hsdR514* Δ*araBAD567* Δ*rhaBAD568 rph-1 attλ*::pAE12 (Δ*oriR6K-cat*::Frt5) Δ*endA*::Frt *uidA*(ΔMluI)::*pir attHK*::pJK1006Δ (*oriR6K-cat*::Frt5 *trfA*::Frt)	Luo *et al.* (2013)
E. coli BW25113 plJ790	Δ(*araD-araB*)*567* Δ*lacZ4787*(::*rrnB-4*) *lacIp-4000*(*lacI*^{q}) λ-*rpoS369*(*Am*) *rph-1* Δ(*rhaD-rhaB*)*568 hsdR514* λ-RED (*gam, bet, exo*), *cat, araC, rep101*^{ts}	Gust *et al.* (2003)
S. albus SAM2	*S. albus* J1074 derivative; deletion of φC31 *pseB4*	Bilyk and Luzhetskyy (2014)
S. albus OK3	*bldA* null mutant of SAM2	Koshla *et al.* (2017)
S. albus 1074::hyg	Hyg^{R}; SAM2 with *XNR1074::hyg* replacement	This work
S. albus 1078::hyg	Hyg^{R}; SAM2 with *XNR1078::hyg* replacement	This work
S. albus Δ*miaA*	eviction of *hyg* from *1074::hyg*; *miaA* null mutant	This work
S. albus Δ*miaA* Δ*bldA*	Am^{R}; Δ*miaA* with *bldA::aac(3)IV-oriT* replacement	This work
S. albus Δ*miaB*	eviction of *hyg* from *1078::hyg*; *miaB* null mutant	This work
S. albus Δ*miaA* pTES	Am^{R}; Δ*miaA*, carries integrated empty vector pTES	This work
S. albus Δ*miaA* pTES1074	Am^{R}; Δ*miaA*, carries integrated pTES1074; complemented Δ*miaA* strain	This work
S. albus pTES	Am^{R}; SAM2, carries integrated empty vector pTES	This work
S. albus pTES1074	Am^{R}; SAM2, carries integrated pTES1074	This work
S. albus pTES1078	Am^{R}; SAM2, carries integrated pTES1078	This work
Bacillus cereus ATCC19637	Test culture to detect antibacterial compounds; moenomycin sensitive culture	ATCC
Debaryomyces hansenii VKM Y-9	Test culture to detect antifungal compounds	Koshla *et al.* (2017)
Cosmids		
pOJ436ara	Am^{R}; pOJ436; carries aranciamycin biosynthetic gene cluster from *S. echinatus*	Luzhetskyy *et al.* (2007)
moeno38-6	Km^{R}, Am^{R}, Hyg^{R}; SuperCos1-based, carries entire *moe* cluster 1 except *moeR-5moeS5*, intφC31-oriTRK2-hyg cassette, Δ*moeH5*	Ostash *et al.* (2013)
Plasmids		
pKC1132	Am^{R}; suicidal *Streptomyces* vector	Bierman *et al.* (1992)
pKC1074	Am^{R}; pKC1132, carrying 4.7-kb fragment with *XNR_1074* cloned into XbaI/EcoRI sites	This work
pKC1078	Am^{R}; pKC1132, carrying 5.4-kb fragment with *XNR_1078* cloned into XbaI/EcoRI sites	This work
pKC1074::hyg	Am^{R}, Hyg^{R}; pKC1074, with Δ*XNR_1074::hyg* replacement (*hyg* is from patt-shyg)	This work
pKC1078::hyg	Am^{R}, Hyg^{R}; pKC1078, with Δ*XNR_1078::hyg* replacement (*hyg* is from patt-shyg)	This work
patt-shyg	Am^{R}, Hyg^{R}; contains a synthetic *hyg*R6Kγ, flanked with B-CC and P-GG sites for φC31 integrase	Myronovskyi *et al.* (2014a)
pUWLint31	Ap^{R}, Th^{R}; expression of φC31 integrase (*int* gene) from high copy number vector pUWL101	Myronovskyi *et al.* (2014a)
pTES	Am^{R}; pSET152 derivative; *attP* flanked with *loxP* sites, *ermEp* flanked by *tfd* terminators	Herrmann *et al.* (2012)
pTES1074	Am^{R}; pTES with *XNR_1074* cloned into XbaI/EcoRI sites	This work
pTES1078	Am^{R}; pTES with *XNR_1078* cloned into XbaI/EcoRI sites	This work
padpAscript	Am^{R}, Sp^{R}; pGUS, contains *adpAghp-gusA* fusion	Makitrynskyy *et al.* (2013)
padpAtransl	Am^{R}, Sp^{R}; pGUSHL4aadA, contains *adpAghp-adpAgh-gusA* fusion with HL4 linker	Makitrynskyy *et al.* (2013)
padpAcontrol	Am^{R}, Sp^{R}; pGUSHL4aadA, contains promoterless *adpAgh-gusA* fusion with HL4 linker	Makitrynskyy *et al.* (2013)
pSAGA	Am^{R}; pSETpAm with *gusA* from pGUS cloned into EcoRI/EcoRV sites, under *aac(3)IVp*	This work
pSAGAt	Am^{R}; pSETpAm with *gusA* (second CTG codon substituted with TTA), cloned into EcoRI/EcoRV sites, under *aac(3)IVp*	This work
pRV3	Am^{R}, Sp^{R}; pGCymRP21 carrying TTA+ version of *SCO3479 (lacZ)* under control of cmt-P21	Koshla *et al.* (2017)
pGM4181	Am^{R}, Sp^{R}; pGUS contains TTA+ version of XNR_4181 (*adpA*) under control of *moeE5p*	Yushchuk *et al.* (2018)
pGM4181tta-	Am^{R}, Sp^{R}; pGUS with TTA-free synthetic version of *XNR_4181* (*adpA*) under control of *moeE5p*	This work

Hyg^{R}; hygromycin resistant; Am^{R}, apramycin resistant; Km^{R}, kanamycin resistant; Ap^{r}, ampicillin resistant; Sp^{R}; spectinomycin resistant; Th^{R}; thiostrepton resistant.

Table 2. Primers used.

Name	Sequence, 5′→3′	Purpose
1074_XbaI_up	AAATCTAGATTAGGCTCCGGGCATGCTTG	To generate pKC1074 and pKC1078
1074_EcoRI_rp	AAAGAATTCGATGCGGTGGCGGGGACG	
1078_XbaI_up	AAATCTAGACGCCACCACCCACGACATC	
1078_EcoRI_rp	AAAGAATTCGGCGTTGACGACCTCGCC	
1074_XbaI_upcl	AAATCTAGACCCCTCGGCACCCCTCG	To generate pTES1074, pTES1078 and verify knockouts of *XNR_1074, XNR_1078*
1074_EcoRI_rpcl	AAAGAATTCGTCCCGATGCCATCACGTG	
1078_XbaI_upcl	AAATCTAGACCATGCGGGTGATCCTGCC	
1078_EcoRI_rpcl	AAAGAATTCAAGCATGCCCGGAGCCTAAC	
1074_uprd	GCCGCGCCCCCGCCCGCCGTTTGCGAGACTGGGACGGTGTTCCGGGGATCCGTCGACCC	To amplify *hyg* resistance cassettes from patt-shyg for *XNR_1074* and *XNR_1078* genes replacements
1074_rprd	CGATGACGGGTGGAGCGTCCCGATGCCATCACGTGATCATGTAGGCTGGAGCTGCTTCG	
1078_uprd	GACGGAGCGCTCCGCCCCAGCCCTTACTCTGGACGGATGTTCCGGGGATCCGTCGACCC	
1078_rprd	CGGGCGGCGCGTGGCCCGGTCCGGGGGCGTCTCCGGTCATGTAGGCTGGAGCTGCTTCG	
pattshyg_hyg_rp	GTCCGGCAGCCAGTCTTCCAG	Diagnostic primer to verify presence of *hyg* resistance cassette
gusA_F	TTTGATATCGGAGGTACGGACATGCTG	To construct pSAGA and pSAGAt
gusA_tta_F	TTTGATATCGGAGGACGGACATGTTACGGCCCGTCGAAACCCCG	
gusA_R	TTTGAATTCACTGCTTCCCGCCCTGCT	

Sites for restriction endonucleases are underlined in primer sequence.

P-GG sites was evicted from genomes of *miaA::hyg* and *miaB::hyg* mutants with recombinase Int-phiC31 (expressed from pUWLint31) (Myronovskyi *et al.*, 2014a). Replacement of *bldA* with apramycin resistance cassette in Δ*miaA* genome was accomplished with the *bldA*-targetting BAC as described in Koshla *et al.* (2017). All genome rearrangements for *S. albus* chromosome were confirmed with diagnostic PCR; appropriate primers were used as mentioned in Fig S1.

Total tRNA isolation from S. albus *strains and HPLC analysis*

Strains were grown from preculture in liquid TSB medium at 30°C for 2 days. Total tRNA fraction was isolated using modified protocol of (Cayama *et al.*, 2000). Biomass from 100 ml of culture was separated via centrifugation, washed two times in 25 ml of Wash buffer (50 mM of Tris-HCl, 0.9% of NaCl, pH 7.6) and resuspended in 1.5V of Resuspension buffer (1 mM of Tris-HCl, 10 mM of $MgCl_2$, pH 7.2). 1.5 vol of Aquaphenol, saturated with the above buffer, was added. After mixing at 4°C for 1 h, samples were spun down at 4°C for 20 min (6000 rpm) and aqueous layer was transferred to the new tube. Nucleic acids were precipitated using absolute ethanol, dissolved in 0.05 M of NaCl, treated with LiCl (for purification from high molecular weight DNA and rRNA) and then precipitated with isopropanol. Quality and quantity of the samples were checked with NanoDrop and agarose gel electrophoresis.

For tRNA digestion to nucleosides nuclease P1 and alkaline phosphatase were used. The resulting hydrolysates were run on HPLC on a Develosil C-30 column with a Waters Alliance HPLC system column essentially as in 1989(1989).

Scanning electron microscopy

Small pieces of *S. albus* lawns were cut off from agar lawns, vacuum-dried and directly analysed on a Jeol JSM-T220A scanning microscope.

Analysis of secondary metabolites produced by S. albus *strains*

LC-MS-based moenomycin quantification has been carried out as in Makitrynskyy *et al.* (2010). In LC-MS, we monitored nosokomycin A2 anion (1501.6 Da $[M-H]^-$) (Ostash *et al.*, 2013) and mean value of its mass peak area in moeno38-6+ SAM2 strain was taken for 100%. The extracts were spiked with equal amounts of moenomycin A (1580.6 Da $[M-H]^-$; to a final concentration of 1 μM) as an internal standard.

For aranciamycin production analysis, pOJ436ara+ strains were grown for 24 h in TSB at 30°C. Five hundred microlitres of the preculture was used to inoculate 25 ml of main medium SG1. After 5 days of growth biomass was separated from 10 ml of culture via centrifugation. Aranciamycin from spent medium was extracted with equal volume of ethyl acetate. Ethyl acetate extracts were dried *in vacuo*, dry residue was dissolved in 50 μL of methanol. The extracts were subjected to TLC and disc diffusion assay against lawn of *B. cereus* ATCC19637 (10^7 spores per plate). TLC and productivity index (PI) measurement was carried out as described in Koshla *et al.* (2017). PI of pOJ436ara+ SAM2 strain was taken for 100%.

Endogenous antibiotic activity of *S. albus* strains (as well as aranciamycin production of pOJ436ara+ strains) was monitored using agar plug antibiotic diffusion assay. Strains were grown on agar plates for 6 days, Agar plugs (Ø 5 mm) were cut off the lawns after 6 days of growth and stacked on top of test cultures. As test cultures *D. hansenii* (spread immediately prior to the experiments on TSA plate) and *B. cereus* spores (as described above) were used. For untargeted HPLC-MS-based analysis of secondary metabolome strains were grown in SG2 media for 120 h at 30°C. Biomass from 35-ml fermentation (approx. 2.3 ± 0.2 g, wet weight) was extracted with methanol-acetonitrile (3:1) while 10-ml samples of spent medium were extracted with equal volume of butanol. Further extract processing and analysis was carried out essentially as in Ahmed *et al.* (2017).

Screening of and measuring GusA activity

To visualize β-glucuronidase (GusA) activity in solid culture, lawns of *gusA*-expressing *S. albus* strains were grown in the presence of chromogenic substrate X-gluc (5-bromo-4-chloro-3-indolyl-8-D-glucuronide). GusA activity in cell-free extracts was measured as described in Myronovskyi *et al.* (2011). For studying *adpA* transcriptional and translational activities, padpAscript^r, padpAtransl$^+$ and padpAcontrol$^+$ strains were grown in TSB medium for 48 h at 30°C, 200 rpm. GusA activity was calculated in Units per 1 g of dried biomass. For studying of TTA+ *gusA* gene expression efficiency, pSAGA$^+$ and pSAGAt$^+$ strains were grown in TSB medium from equal amounts of spores (approximately 1–1.5 × 10^7 spores in 30 ml of medium) at 30°C, 200 rpm. Samples were taken at 19, 25, 43, 49 and 73 h, and GusA activity was measured in Units/g. Final data shown in Fig. 9 are calculated as ratio of GusA activity of experimental sample to GusA activity of pSAGA$^+$ *S. albus* SAM2 at a given time point (the latter was taken for 1 at all time points).

Stress response study

To monitor response to high temperature stress, *S. albus* strains were grown on agar plates from equal amounts of spores (approx. 10^5 spores per lawn) at 30, 37 and 42°C for 144 h. In antibiotic resistance studies, *S. albus* strains were grown on ISP3, antibiotic disks were placed on the lawns and plates incubated for 48 h at 30°C prior to measurement of diameters of halos of *S. albus* growth inhibition (in mm).

Acknowledgements

This work was supported by grant BG-41Nr from Ministry of Education and Science of Ukraine (to BO). Publication in part is based on the research provided by the grant support of the State Fund for Fundamental Research of Ukraine (project F80/2-2018, to BO). OK was supported with Swedish Institute Visby scholarship for PhD studies (in Uppsala University, Sweden). Work in Umeå was supported by Swedish Research Council (621-2016-03949 to AB); and Karin and Harald Silvanders Foundation/Insamlingsstiftelsen Umeå universitet (FS 2.1.6-1870-16 to AB). R.D.S. and B.O. thank the Bundesministerium für Bildung und Forschung (BMBF) project: 021DK17013 AWAKEDRUGS for funding. Authors have no conflict of interest to declare.

Author contributions

Conceived and designed the experiments: OK, OY, IO, MM, AL and BO; performed the experiments: OK, YD, OY, IO, GJ and AB; contributed new analytical tools: LAK, MM and RDS; analysed the data: OK, GJ, AB, LAK and BO; wrote the paper: OK, LAK, RDS and BO.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Agris, P.F., Narendran, A., Sarachan, K., Väre, V.Y.P. and Eruysal, E. (2017) The importance of being modified: the role of RNA modifications in translational fidelity. In: Chanfreau, G.F. (Ed.) *The Enzymes*. Vol. **41**, Cambridge, MA: Academic Press, pp. 1–50.

Ahmed, Y., Rebets, Y., Tokovenko, B., Brötz, E. and Luzhetskyy, A. (2017) Identification of butenolide regulatory system controlling secondary metabolism in *Streptomyces albus* J1074. *Scientific Reports*, **7**, 9784.

Armengod, M.E., Meseguer, S., Villarroya, M., Prado, S., Moukadiri, I., Ruiz-Partida, R., et al. (2014) Modification of the wobble uridine in bacterial and mitochondrial tRNAs reading NNA/NNG triplets of 2-codon boxes. *RNA Biology*, **11**, 1495–1507.

Aubee, J.I., Olu, M. and Thompson, K.M. (2016) The i^6 A37 tRNA modification is essential for proper decoding of UUX-Leucine codons during *rpoS* and *iraP* translation. *RNA*, **22**, 729–742.

Aubee, J., Olu, M. and Thompson, K. (2017) TrmL and TusA are necessary for *rpoS* and MiaA is required for *hfq* expression in *Escherichia coli*. *Biomolecules*, **7**, 39.

Benos, P.V., Lapedes, A.S. and Stormo, G.D. (2002) Is there a code for protein-DNA recognition? Probab(ilistical)ly. *BioEssays*, **24**, 466–475.

Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N. and Schoner, B.E. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene*, **116**, 43–49.

Bilyk, B. and Luzhetskyy, A. (2014) Unusual site-specific DNA integration into the highly active pseudo-*attB* of the *Streptomyces albus* J1074 genome. *Applied Microbiology and Biotechnology*, **98**, 5095–5104.

Björk, G.R., Durand, J.M.B., Hagervall, T.G., Leipuviene, R., Lundgren, H.K., Nilsson, K., *et al.* (1999) Transfer RNA modification: influence on translational frameshifting and metabolism. *FEBS Letters*, **452**, 47–51.

Cayama, E., Yépez, A., Rotondo, F., Bandeira, E., Ferreras, A.C. and Triana-Alonso, F.J. (2000) New chromatographic and biochemical strategies for quick preparative isolation of tRNA. *Nucleic Acids Research*, **28**, E64.

Chang, G.W., Roth, J.R. and Ames, B.N. (1971) Histidine regulation in *Salmonella typhimurium*. 8. Mutations of the *hisT* gene. *Journal of Bacteriology*, **108**, 410–414.

Chater, K.F. and Chandra, G. (2008) The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in *Streptomyces*. *The Journal of Microbiology*, **46**, 1–11.

Chionh, Y.H., McBee, M., Babu, I.R., Hia, F., Lin, W., Zhao, W., *et al.* (2016) tRNA-mediated codon-biased translation in mycobacterial hypoxic persistence. *Nature Communications*, **7**, 13302.

Crnković, A., Vargas-Rodriguez, O., Merkuryev, A. and Söll, D. (2018) Effects of heterologous tRNA modifications on the production of proteins containing noncanonical amino acids. *Bioengineering*, **5**, 11.

Datsenko, K. and Wanner, B. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, **97**, 6640–6645.

Deng, W., Babu, I.R., Su, D., Yin, S., Begley, T.J. and Dedon, P.C. (2015) Trm9-catalyzed tRNA modifications regulate global protein expression by codon-biased translation. *PLoS Genetics*, **11**, e1005706.

Ericson, J.U. and Björk, G.R. (1986) Pleiotropic effects induced by modification deficiency next to the anticodon of tRNA from *Salmonella typhimurium* LT2. *Journal of Bacteriology*, **166**, 1013–1021.

Gehrke, C. and Kuo, K. (1989) Ribonucleoside analysis by reversed-phase high-performance liquid chromatography. *Journal of Chromatography*, **471**, 3–36.

Grosjean, H. and Westhof, E. (2016) An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Research*, **30**, 8020–8040.

Gust, B., Challis, G.L., Fowler, K., Kieser, T. and Chater, K.F. (2003) PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proceedings of the National Academy of Sciences*, **100**, 1541–1546.

Hackl, S. and Bechthold, A. (2015) The gene *bldA*, a regulator of morphological differentiation and antibiotic production in *Streptomyces*. *Archiv der Pharmazie*, **348**, 455–462.

Herrmann, S., Siegl, T., Luzhetska, M., Jilg, L.P., Welle, E., Erb, A., *et al.* (2012) Site-specific recombination strategies for engineering actinomycete genomes. *Applied and Environment Microbiology*, **78**, 1804–1812.

Higo, A., Horinouchi, S. and Ohnishi, Y. (2011) Strict regulation of morphological differentiation and secondary metabolism by a positive feedback loop between two global regulators AdpA and BldA in *Streptomyces griseus*. *Molecular Microbiology*, **81**, 1607–1622.

Higo, A., Hara, H., Horinouchi, S. and Ohnishi, Y. (2012) Genome-wide distribution of AdpA, a global regulator for secondary metabolism and morphological differentiation in streptomyces, revealed the extent and complexity of the AdpA regulatory network. *DNA Research*, **19**, 259–273.

Hopwood, D.A. (1967) Genetic analysis and genome structure in *Streptomyces coelicolor*. *Bacteriological Reviews*, **31**, 373–403.

Hori, H. (2017) Transfer RNA methyltransferases with a SpoU-TrmD (SPOUT) fold and their modified nucleosides in tRNA. *Biomolecules*, **7**, 23.

Jaroensuk, J., Atichartpongkul, S., Chionh, Y.H., Hwa Wong, Y., Liew, C.W., McBee, M.E., *et al.* (2016) Methylation at position 32 of tRNA catalyzed by TrmJ alters oxidative stress response in *Pseudomonas aeruginosa*. *Nucleic Acids Research*, **44**, 10834–10848.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A. (2000) *Practical Streptomyces genetics*. Norwich: The John Innes Foundation.

Koshla, O., Lopatniuk, M., Rokytskyy, I., Yushchuk, O., Dacyuk, Y., Fedorenko, V., *et al.* (2017) Properties of *Streptomyces albus* J1074 mutant deficient in $tRNA^{Leu}_{UAA}$ gene *bldA*. *Archives of Microbiology*, **199**, 1175–1183.

Labeda, D.P., Doroghazi, J.R., Ju, K.S. and Metcalf, W.W. (2014) Taxonomic evaluation of *Streptomyces albus* and related species using multilocus sequence analysis and proposals to emend the description of *Streptomyces albus* and describe *Streptomyces pathocidini* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, **64**, 894–900.

Lawlor, E.J., Baylis, H.A. and Chater, K.F. (1987) Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). *Genes & Development*, **1**, 1305–1310.

Leskiw, B.K., Mah, R., Lawlor, E.J. and Chater, K.F. (1993) Accumulation of *bldA*-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2). *Journal of Bacteriology*, **175**, 1995–2005.

Li, W., Wu, J., Tao, W., Zhao, C., Wang, Y., He, X., *et al.* (2007) A genetic and bioinformatic analysis of *Streptomyces coelicolor* genes containing TTA codons, possible targets for regulation by a developmentally significant tRNA. *FEMS Microbiology Letters*, **266**, 20–28.

Lopatniuk, M., Myronovskyi, M. and Luzhetskyy, A. (2017) *Streptomyces albus*: a new cell factory for non-canonical amino acids incorporation into ribosomally synthesized natural products. *ACS Chemical Biology*, **12**, 2362–2370.

Luo, Y., Huang, H., Liang, J., Wang, M., Lu, L., Shao, Z., *et al.* (2013) Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nature Communications*, **4**, 2894.

Luzhetskii, A.N., Ostash, B.E. and Fedorenko, V.A. (2001) Intergeneric conjugation *Escherichia coli – Streptomyces globisporus* 1912 using integrative plasmid pSET152 and its derivatives. *Russian Journal of Genetics*, **37**, 1123–1129.

Luzhetskyy, A., Mayer, A., Hoffmann, J., Pelzer, S., Holzenkämper, M., Schmitt, B., *et al.* (2007) Cloning and heterologous expression of the aranciamycin biosynthetic gene cluster revealed a new flexible glycosyltransferase. *ChemBioChem*, **8**, 599–602.

Dal Magro, C., Keller, P., Kotter, A., Werner, S., Duarte, V., Marchand, V., *et al.* (2018) A vastly increased chemical variety of RNA modifications containing a thioacetal structure. *Angewandte Chemie International Edition*, **57**, 7893–7897.

Makitrynskyy, R., Rebets, Y., Ostash, B., Zaburannyi, N., Rabyk, M., Walker, S. and Fedorenko, V. (2010) Genetic factors that influence moenomycin production in streptomycetes. *Journal of Industrial Microbiology and Biotechnology*, **37**, 559–66.

Makitrynskyy, R., Ostash, B., Tsypik, O., Rebets, Y., Doud, E., Meredith, T., *et al.* (2013) Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. *Open Biology*, **3**, 130121.

Mathevon, C., Pierrel, F., Oddou, J.-L., Garcia-Serres, R., Blondin, G., Latour, J.-M., *et al.* (2007) tRNA-modifying MiaE protein from *Salmonella typhimurium* is a nonheme diiron monooxygenase. *Proceedings of the National Academy of Sciences*, **104**, 13295–13300.

McLean, T.C., Hoskisson, P.A. and Seipke, R.F. (2016) Coordinate regulation of antimycin and candicidin biosynthesis. *mSphere*, **1**.

Merrick, M.J. (1976) A morphological and genetic mapping study of bald colony mutants of *Streptomyces coelicolor*. *Journal of General Microbiology*, **96**, 299–315.

Moukadiri, I., Garzón, M.-J., Björk, G.R. and Armengod, M.-E. (2014) The output of the tRNA modification pathways controlled by the *Escherichia coli* MnmEG and MnmC enzymes depends on the growth conditions and the tRNA species. *Nucleic Acids Research*, **42**, 2602–2623.

Myronovskyi, M., Welle, E., Fedorenko, V. and Luzhetskyy, A. (2011) Beta-glucuronidase as a sensitive and versatile reporter in actinomycetes. *Applied and Environment Microbiology*, **77**, 5370–5383.

Myronovskyi, M., Rosenkränzer, B. and Luzhetskyy, A. (2014a) Iterative marker excision system. *Applied Microbiology and Biotechnology*, **98**, 4557–4570.

Myronovskyi, M., Tokovenko, B., Brötz, E., Rückert, C., Kalinowski, J. and Luzhetskyy, A. (2014b) Genome rearrangements of *Streptomyces albus* J1074 lead to the carotenoid gene cluster activation. *Applied Microbiology and Biotechnology*, **98**, 795–806.

Myronovskyi, M., Rosenkränzer, B., Nadmid, S., Pujic, P., Normand, P. and Luzhetskyy, A. (2018) Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metabolic Engineering*, **49**, 316–324.

Nguyen, K.T., Tenor, J., Stettler, H., Nguyen, L.T., Nguyen, L.D. and Thompson, C.J. (2003) Colonial differentiation in *Streptomyces coelicolor* depends on translation of a specific codon within the *adpA* gene. *Journal of Bacteriology*, **185**, 7291–7296.

Olano, C., García, I., González, A., Rodriguez, M., Rozas, D., Rubio, J., *et al.* (2014) Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. *Microbial Biotechnology*, **7**, 242–256.

Ostash, B., Campbell, J., Luzhetskyy, A. and Walker, S. (2013) MoeH5: a natural glycorandomizer from the moenomycin biosynthetic pathway. *Molecular Microbiology*, **90**, 1324–1338.

Persson, B.C. and Björk, G.R. (1993) Isolation of the gene (miaE) encoding the hydroxylase involved in the synthesis of 2-methylthio-cis-ribozeatin in tRNA of Salmonella typhimurium and characterization of mutants. *Journal of Bacteriology*, **175**, 7776–7785.

Persson, B.C., Olafsson, Ó., Lundgren, H.K., Hederstedt, L. and Björk, G.R. (1998) The ms2io6A37 modification of tRNA in *Salmonella typhimurium* regulates growth on citric acid cycle intermediates. *Journal of Bacteriology*, **180**, 3144–3151.

Pettersson, B.M.F. and Kirsebom, L.A. (2011) tRNA accumulation and suppression of the bldA phenotype during development in *Streptomyces coelicolor*. *Molecular Microbiology*, **79**, 1602–1614.

Pierrel, F., Douki, T., Fontecave, M. and Atta, M. (2004) MiaB protein is a bifunctional radical- *S* -adenosylmethionine enzyme involved in thiolation and methylation of tRNA. *Journal of Biological Chemistry*, **279**, 47555–47563.

Rabyk, M., Yushchuk, O., Rokytskyy, I., Anisimova, M. and Ostash, B. (2018) Genomic insights into evolution of AdpA family master regulators of morphological differentiation and secondary metabolism in *Streptomyces*. *Journal of Molecular Evolution*, **86**, 204–215.

Rokytskyy, I., Koshla, O., Fedorenko, V. and Ostash, B. (2016) Decoding options and accuracy of translation of developmentally regulated UUA codon in *Streptomyces*: bioinformatic analysis. *Springerplus*, **5**, 982.

Sakai, Y., Miyauchi, K., Kimura, S. and Suzuki, T. (2016) Biogenesis and growth phase-dependent alteration of 5-methoxycarbonylmethoxyuridine in tRNA anticodons. *Nucleic Acids Research*, **44**, 509–523.

Sambrook, J. and Russell, W.D. (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schweizer, U., Bohleber, S. and Fradejas-Villar, N. (2017) The modified base isopentenyladenosine and its derivatives in tRNA. *RNA Biology*, **14**, 1197–1208.

Sehin, Y., Koshla, O., Dacyuk, Y., Zhao, R., Ross, R., Myronovskyi, M., *et al.* (2018) Gene ssfg_01967 (miaB) for tRNA modification influences morphogenesis and moenomycin biosynthesis in *Streptomyces ghanaensis* ATCC14672. *Microbiology*, **01967**, 1–13.

Shepherd, J. and Ibba, M. (2015) Bacterial transfer RNAs. *FEMS Microbiology Reviews*, **39**, 280–300.

Shigi, N. (2014) Biosynthesis and functions of sulfur modifications in tRNA. *Frontiers in Genetics*, **5**, 67.

Takano, E., Tao, M., Long, F., Bibb, M.J., Wang, L., Li, W., *et al.* (2003) A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. *Molecular Microbiology*, **50**, 475–486.

Xu, F., Nazari, B., Moon, K., Bushin, L.B. and Seyedsayamdost, M.R. (2017) Discovery of a cryptic antifungal compound from *Streptomyces albus* J1074 using high-throughput elicitor screens. *Journal of the American Chemical Society*, **139**, 9203–9212.

Yacoubi, B.El, Bailly, M. and de Crécy-Lagard, V. (2012) Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annual Review of Genetics*, **46**, 69–95.

Yu, J.M., Wang, D., Pierson, L.S. and Pierson, E.A. (2017) Disruption of MiaA provides insights into the regulation of phenazine biosynthesis under suboptimal growth conditions in *Pseudomonas chlororaphis* 30–84. *Microbiology*, **163**, 94–108.

Yushchuk, O., Ostash, I., Vlasiuk, I., Gren, T., Luzhetskyy, A., Kalinowski, J., *et al.* (2018) Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions. *Applied Microbiology and Biotechnology*, **102**, 8419–8428.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article



© 2019 John Wiley & Sons Ltd, *Molecular Microbiology*, **112**, 249–265

Оригінал доступний за посиланням:
https://onlinelibrary.wiley.com/doi/10.1111/mmi.14266.
Електронний додаток до статті доступний за посиланням:
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fmmi.14266&file=mmi14266-sup-0001-Supinfo.pdf.

Анотація: *Представники актинобактерій роду Streptomyces мають складний життєвий цикл і є найбільшим джерелом біоактивних вторинних метаболітів. Хоча морфогенез і вторинний метаболізм підлягають транскрипційній корегуляції, стрептоміцети використовують додатковий механізм для активації вищезгаданих процесів. Цей механізм заснований на відстроченій трансляції рідкісного лейцинового кодону UUA єдиною спорідненою мРНК$^{Leu}_{UAA}$ (кодується геном bldA). Генетичний перемикач на основі bldA є добре вивченим прикладом регуляції трансляції у Streptomyces. Проте, після п'яти десятиліть з моменту відкриття bldA, деякі його особливості залишаються невідомими. В цій роботі ми розглядаємо гіпотезу про те, що посттранскрипційні модифікації мРНК відіграють роль у мРНК-опосередкованих механізмах трансляційного контролю в Streptomyces. Зокрема, ми досліджуємо два гени Streptomyces albus J1074, XNR_1074 (miaA) і XNR_1078 (miaB), що кодують мРНК (аденозин(37)-N6)-диметилалілтрансферазу і мРНК (N6-ізопентеніладенозин(37)-C2)-метилтіотрансферазу, відповідно. Ці ферменти послідовно здійснюють гіпермодифікацію залишку A37 у більшості мРНК, що містять A36-A37. Ми демонструємо, що miaB і особливо miaA мутанти S. albus мають змінений морфогенез і вторинний метаболізм. Ми надаємо генетичні докази того, що дефіцит miaA впливає на трансляційний рівень експресії генів, швидше за все, через порушення декодування кодонів УХХ і зокрема UUA.*

Folia Microbiol (2016) 61:209–220
DOI 10.1007/s12223-015-0426-7

Transcriptional regulators of GntR family in *Streptomyces coelicolor* A3(2): analysis in silico and in vivo of YtrA subfamily

O. Tsypik[1] · O. Yushchuk[1] · N. Zaburannyi[1] · K. Flärdh[2] · S. Walker[3] · V. Fedorenko[1] · B. Ostash[1,3]

Received: 20 May 2015 / Accepted: 27 September 2015 / Published online: 3 October 2015
© Institute of Microbiology, Academy of Sciences of the Czech Republic, v.v.i. 2015

Abstract Transcriptional factors of the GntR family regulate numerous physiological and morphological processes in response to the nutrient state of bacterial cells. The number of GntR transcriptional factors in genomes of soil-dwelling actinomycetes is one of the highest among bacteria, reflecting both the large size of their chromosomes and the complex ecological niche that they occupy. However, very little is known about the roles of GntRs in actinomycete biology. Here, we analyzed the genome of model actinomycete, *Streptomyces coelicolor* A3(2), in an attempt to gain new insights into the function of GntR family. All 56 GntR proteins of M145 strain were classified into FadR, HutC, MocR, YtrA, and DevA subfamilies according to their secondary structure. We then checked for the presence of GntR orthologs in six other sequenced *Streptomyces* and one *Kitasatospora* genomes, revealing that 12 GntRs were conserved in all analyzed strains. Genomic analysis of the less studied YtrA type regulators revealed 160 sequences present in 88 members of *Coriobacteridae*, *Rubrobacteridae*, and *Actinobacteridae* subclasses. These proteins form seven dense clusters on the consensus phylogenetic tree and their genes are usually co-located with the genes for transport proteins. Probable operator sites were identified for orthologous groups of Sco0823 and Sco3812 proteins. All *S. coelicolor* YtrA-like regulatory genes (*SCO0823*, *SCO1728*, *SCO3812*) were analyzed at transcriptional level, knocked out, and introduced on moderate copy number plasmid in M145 strain. Also, gene *SCO0824*, a part of putative *SCO0823* operon, was studied. Results of these experiments are discussed here.

Introduction

Streptomycetes are soil-dwelling mycelial bacteria with a complex life cycle. Members of this genus have large genomes and are well known for their ability to produce a broad spectrum of biologically active substances, in particular, many known antibiotics (Nett et al. 2009). Streptomycetes undergo a complex morphological and physiological differentiation, and both processes are orchestrated by a very complex regulatory network (McCormick and Flärdh 2012; van Wezel and McDowall 2011). Nevertheless, because the major portion of genes for transcriptional factors have no defined function even in model streptomycetes, the larger picture of these regulatory pathways is incomplete (Liu et al. 2013).

One of the most abundant and widely distributed groups of transcriptional factors in actinobacteria is the GntR family of regulators (GntRs). Except the KorSA protein (involved in conjugative plasmid transfer in *Streptomyces*), all GntRs studied to date have complex, pleiotropic effects on actinobacterial biology, influencing morphogenesis and antibiotic production (Hoskisson and Rigali 2009). The structure of GntRs consists of a conserved N-terminal helix-turn-helix motif and a variable C-terminal effector binding/oligomerization domain (Rigali et al. 2002). GntRs are divided into seven subfamilies based on the secondary structure of the C-terminal domain—

Electronic supplementary material The online version of this article (doi:10.1007/s12223-015-0426-7) contains supplementary material, which is available to authorized users.

✉ B. Ostash
b.ostash@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, Hrushevskoho st. 4, Lviv 79005, Ukraine

[2] Department of Biology, Lund University, Lund 22362, Sweden

[3] Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA

FadR, HutC, MocR, PlmA, DevA, AraR, and YtrA. YtrA is the largest of the seven subfamilies but has been studied minimally in bacteria. These proteins all possess a reduced C-terminal domain with only two α-helices; YtrA-encoding genes are usually part of operon with genes for ATP-binding cassette (ABC) transporters (Hoskisson and Rigali 2009). In *Bacillus subtilis*, YtrA mediates the antibiotic stress response via regulation of the *ytrA* and *ywoB* operons (Salzberg et al. 2011). LndYR protein of *Streptomyces globisporus* 1912 is the only extensively studied representative of YtrA subfamily in *Streptomyces*. LndYR was shown to affect sporulation and landomycin E production by *S. globisporus* 1912 and to bind the promoter region of the ABC transporter genes *lndW-W2* (Ostash et al. 2011).

Elucidating the functions of GntRs in streptomycetes may clarify the regulation of morphogenesis and secondary metabolism. To this end, we explored various aspects of the biology of GntRs in model strain *Streptomyces coelicolor* M145 using bioinformatics tools as well as performed more detailed functional studies on genes for the YtrA subfamily regulators.

Materials and methods

In silico analysis Nucleotide and amino acid sequences of annotated GntRs were downloaded from *Streptomyces* Database StrepDB (http://strepdb.streptomyces.org.uk). We also searched for GntRs in annotated *S. coelicolor* and *Actinobacteria* genomes using DELTA-BLAST (Boratyn et al. 2012). Secondary structure prediction was carried out using JPred (Cole et al. 2008), PhyreII (Kelley and Sternberg 2009), and Psipred (Buchan et al. 2010). PRALINE web server (Simossis and Heringa 2003), MUSCLE, and PROBCONS were used for multiple sequence alignments. Protein trees were calculated with TNT (maximum parsimony algorithm) and PhyML 3.0 (maximum-likelihood algorithm, http://www.phylogeny.fr; Dereeper et al. 2008). Reciprocal BLAST has been performed to identify orthologs.

Putative palindromic operator sites were discovered using RegPredict server (Novichkov et al. 2011) using de novo regulon inference option.

Bacterial strains and culture conditions Strains used in this work are listed in Table 1. *Escherichia coli* DH5α was used for routine subcloning. *E. coli* ET12567 (pUZ8002) was used to perform intergeneric conjugation from *E. coli* to *Streptomyces* species. *E. coli* BW25113 with pIJ790 was used to carry out RedET-mediated gene replacement (Gust et al. 2004). Solid oatmeal (g/L: oat flour 34, agar 18) and SM media (Kieser et al. 2000) were used to harvest *Streptomyces* spores suspension, maintain strains, and plate *E. coli–Streptomyces* matings. Biomass and actinorhodin (Act) production was assayed in yeast extract-malt extract (YEME) (17 % sucrose), soy-mannitol medium (SMM) (Kieser et al. 2000), and yeast extract-malt extract-peptone-glucose (YMPG) (g/L: yeast extract 4, peptone 1, malt extract 10, glucose 10, $MgCl_2$ 6H2O 2) liquid media according to described procedures (Kieser et al. 2000). For phenotype examinations, the following solid media were used: R2YE, oatmeal agar, tryptic soy agar (TSA), MM, MMGT (Kieser et al. 2000), Bennett (g/L: glucose 10, peptone 1, yeast extract 1, tryptone 2, agar 20), modified Evans media (Lewis et al. 2011). Detailed procedures of phenotypic analysis of *S. coelicolor* strains are given in Electronic Supplementary Materials (ESM; legends to the figures S5, S8, S19, Table S3). *E. coli* strains were grown as described elsewhere.

Nucleic acid manipulations Genomic DNA from *Streptomyces* and plasmid DNA from *E. coli* were isolated using standard protocols (Kieser et al. 2000; Sambrook and Russell 2001). Total RNA was isolated from YMPG-grown mycelia of *S. coelicolor* using RNAeasy RNA isolation kit (Qiagen, Germany) according to instructions of the supplier. Complete removal of DNA contamination in RNA samples was achieved using RNAfree kit (Ambion, USA). Equal amounts of RNA (3 μg) and cDNA (200 ng) were used for RT and second strand synthesis steps of analysis, respectively. *E. coli* transformation and intergenetic *E. coli–Streptomyces* matings were performed as described by Kieser et al. (2000). Gene knockouts in *S. coelicolor* were verified via PCR analysis; these data are presented in ESM, Fig. S1–S3. *SCO0823 disruption construct.* A DNA segment of 2.5 kb containing the gene and 1-kb flanking regions was amplified from *S. coelicolor* chromosome using SCO0823_for and SCO0823_rev primers (primer sequences are listed in Table 2). The PCR product was cloned into XbaI, BamHI sites of pKC1139 giving pKCSCO0823. RedET technology (Gust et al. 2004) was then used to replace *SCO0823* within pKCSO0823 with kanamycin resistance gene *kan* from pKD4 (Datsenko and Wanner 2000). For this purpose, primers SCO0823Am_for and SCO0823Am_rev were used. The final construct was labeled as pKCSCO0823::Km. This plasmid and the other constructs described in this work were verified by PCR, restriction mapping or sequencing. *SCO3812 and SCO1728 disruption constructs*. Genes *SCO3812* and *SCO1728* were replaced (via RedET-mediated engineering) within cosmids StGD3 and StI11 (kindly provided by M. Buttner, JIC), respectively, with apramycin resistance cassette *aac(3)IV-oriT* amplified from pIJ774. Primer pairs SCO1728Am_for and SCO1728Am_rev, SCO3812Am_for and SCO3812Am_rev were used for this purpose. The final constructs were named as StGD3::Am and StI11::Am. All three constructs were transferred conjugally into *S. coelicolor*, and double

Table 1 Strains and plasmids used in the work

Strain or plasmid/cosmid	Description	Source or reference
Streptomyces coelicolor M145	SCP1$^-$, SCP2$^-$ derivative of A3(2); Act and Red producer	Kieser et al. (2000)
S. coelicolor 0823::Km	SCO0823 insertional inactivation in M145	This work
S. coelicolor 1728::Am	SCO1728 replacement in M145	This work
S. coelicolor 3812::Am	SCO3812 replacement in M145	This work
S. coelicolor pKC1728E	M145 carrying plasmid pKC1728E	This work
S.coelicolor pKC0823	M145 carrying plasmid pKC0823	This work
S.coelicolor pKC3812	M145 carrying plasmid pKC3812	This work
S. coelicolor p GUS0823	M145 carrying plasmid pGUSSCO0823	This work
S. coelicolor 0823::Km pGUS0823	S. coelicolor 0823::Km carrying plasmid pGUS0823	This work
S. coelicolor Δ0824	Replacement of SCO0824 with a Tn-disrupted allele (from cosmid SCF43A.1.C05), in M145	This work
S. globisporus 1912	Wild type, landomycin E producer	Ostash et al. (2011)
S. ghanaensis ATCC14672	Wild type, moenomycin producer	ATCC
S. sioyaensis NRRL-B5408	Wild type, siomycin producer	NRRL
Escherichia coli DH5α	Routine cloning host	Life Technologies
E. coli ET12567 (pUZ8002)	Host for conjugative DNA transfer	Kieser et al. (2000)
E. coli BW25113 (pIJ790)	Host for recombineering experiments	Gust et al. (2004)
Bacillus cereus	Moenomycin production test culture	ATCC
Micrococcus luteus	Siomycin production test culture	ATCC
pBluescriptIIKS+	General purpose cloning vector, Ap^r	MBI Fermentas
pKC1139	pSG5 ts-replicon; Am^r shuttle vector	Muth et al. (1989)
pKC0823	pKC1139 carrying SCO0823	This work
pKC3812	pKC1139 carrying SCO3812	This work
pKCSCO0823::Km	Insertion of kan (from pKD4) into SCO0823	This work
pKC1218E	SCP2-based vector carrying ermEp, Am^r	Ostash et al. (2011)
pKC1728E	pKC1218E carrying SCO1728	This work
pIJ774	aac(3)IV-oriT cassette flanked with loxP sites	JIC
pKD4	kan cassette flanked with FRT sites	Datsenko and Wanner (2000)
pGUS	gusA-based promoter-probe vector; Am^{r}Sp^r	Myronovskyi et al. (2011)
pGUS0823	pGUS carrying sSCO0823p upstream of gusA	This work
StI11	SuperCos1-based cosmid carrying SCO1728	John Innes Center
StI11::Am	Replacement of SCO1728 with aac(3)IV-oriT in StI11	This work
StGD3	SuperCos1-based cosmid carrying SCO3812	John Innes Center
StGD3::Am	Replacement of SCO1728 with aac(3)IV-oriT in StGD3	This work
SCF43A.1.C05	Cosmid SCF43A carrying Tn insertion into coding sequence of SCO0824	P. Dyson, University of Swansea

crossover mutants were selected as described elsewhere (Kieser et al. 2000; Gust et al. 2004). PCR analysis confirmed that the expected gene replacements occurred in *S. coelicolor* genome (data not shown). *Constructs for SCO0823, 1728, and 3812 gene expression. SCO0823* and *SCO3812* genes with their own promoter were amplified from *S. coelicolor* chromosome using primers SCO0823_for and SCO0823_rev and SCO3812_for and SCO3812_rev. The amplicons were inserted into EcoRV site of moderate copy number vector pKC1139 via TA-cloning, yielding pKC0823 and pKC3812. The DNA fragment containing *SCO1728* along with its promoter region was amplified from cosmid StI11 with primer pair SCO1728_for and SCO1728_rev and cloned into BamHI/XbaI sites of pBlueScriptIIKS (Thermo Scientific Inc). Then, *SCO1728* was moved as a PstI/XbaI fragment from pBluescriptIIKS into respective sites of low copy number vector pKC1218E, to give pKC1728E. In this construct, *SCO1728* is under transcriptional control of two promoters, *sco1728p* (comes from the amplicon) and *ermEp* (from the vector). *Plasmid for SCO0823 promoter activity studies*. The 0.5-kb promoter region of *SCO0823*

Table 2 Primers used in this work

Primer name	Sequence	Purpose
SCO1728Am_for SCO1728Am_rev	CGGTCATGGGGGCGCACGCTACGGCGGCCGGGCGGTCGGATGATTCCGGGGATCCGTCGACC GATCTGCGGCCGGGCCGCGGCCCCGATCCCGACGGCGCTCA TGTAGGCTGGAGCTGCTTC	To replace SCO1728 with aac(3)IV-oriT
SCO1728_for[a] SCO1728_rev	ATTGAAAGGATCCCCGATCACCTCCCCGTCCTCCT ATTGAAATCTAGAGGCCGCATACGAGCGACGTCTA	To amplify SCO1728
SCO0823Km_for SCO0823Km_rev	ATCTGATCATGCTTCTACTAATGAATTAATGAAAGCATGTGTAGGCTGGAGCTGCTTC TGCTCGTGTCGCTCGTCGTGCTCGTCGAGGGCCGCCGTGCATATGAATATCCTCCTTAG	To replace SCO0823 with kan
SCO0823_for SCO0823_rev	AAATCTAGAGCACTTGAGGTGGACCTCG AAAGGATCCGATCAGTGCCCGTTGCTGG	To amplify SCO0823
SCO0824check	GCATCACGGCCCTGGTC	
SCO3812Am_for SCO3812Am_rev	ATGCTGCTGAGTCTGGACAAGACCGACACCCGCCCCCTGCATTCCGGGGATCCGTCGACC TCCCCCTGATCATGTCGATGAGATCGGTCCTGCTGTAGCCTGTAGGCTGGAGCTGCTTC	To replace SCO3812 with aacn3)IV-oriT
SCO3812_for SCO3812_rev	AAATCTAGACGCCGCACCGCCACGAAGGG AAAGATATCGCGGCGAGCACGGCTGTGAC	To amplify SCO3812
GUS0823_for GUS0823_rev	TCTGAATCTAGAACGGCGGTGACGTCCAGGTT TCTGAAGGTACCGCATGATCAGATCCCGTCCC	To clone SCO0823 promoter in pGUS
rt0823_up rt0823_rp	ACTCCCGACCGCCAAGGAG GTCGCTCGTCGTGCTCGTC	RT-PCR analysis of sco0823 gene
rt0824_up	TGTCCTCGCACGTCCTGCC	RT-PCR analysis of sco0824 gene
rt0824_rp	CCTGGAGGTGTCCGAGCAG	
rt0825_up rt0825_rp	GTACGTCAGCGACGGCTGG GCAGCAGGAGGACCGTGAC	RT-PCR analysis of sco0825 gene
rt1728_up rt1728_rp	GATCTCCGAACAGGCGCGC CGCAACGCATCCCGAACCG	RT-PCR analysis of sco1728 gene
rt1729_up rt1729_rp	GGACGAACCGAGCGACGTC GGTGTCGTTGCCGGTGAGC	RT-PCR analysis of sco1729 gene
rt1730_up rt1730_rp	TGCTGCACGACGACGGCAC CTCGTCCAGACGCCGTACG	RT-PCR analysis of sco1730 gene
rt3812_up rt3812_rp	CCTGCACGAACAGGTGGCC CGGTCCTGCTGTAGCCCAG	RT-PCR analysis of sco3812 gene
rt3813_up rt3813_rp	CTGCGCGGTCTTCTCCTCG GCGTTCGGCGTCGTCGATG	RT-PCR analysis of sco3813 gene
rt_hrdB_up rt_hrdB_rp	GCTGGCCAAGGAGCTCGAC CGTCGAGGGTCTTCGGCTG	RT-PCR analysis of hrdB gene

[a] Restriction sites are italicized

(*sco0823p*) was amplified from *S. coelicolor* chromosome with primers GUS0823_for and GUS0823_rev. The PCR product was digested with XbaI/KpnI and cloned into the respective sites of pGUS (Myronovskyi et al. 2011), giving pGusSCO0823. In this plasmid, transcription of the *gusA* gene is under the control of *sco0823p*.

Analysis of antibiotic production Strains were grown for 5 days in liquid media to measure antibiotic production. Moenomycin production by *S. ghanaensis* ATCC14672 was studied in tryptic soy broth (TSB) as described by Makitrynskyy et al. (2010). Siomycin production by *S. sioyaensis* NRRL-B5408 was determined in medium SG1 according to Myronovskyy et al. (2009). Landomycin production by *S. globisporus* was studied in SG medium as described by Ostash et al. (2011). Actinorhodin production was assayed as described by Kieser et al. (2000). In all cases, experiments were repeated three times, data averaged, and referred back to equal amounts of dry biomass.

Analysis of β-glucuronidase activity *S. coelicolor* strains carrying pGUS-based reporter plasmids were grown for 120 h in TSB. Every 24 h, 5 mL of the culture were examined for *SCO0823* expression profile according to protocol described by Myronovskyi et al. (2011).

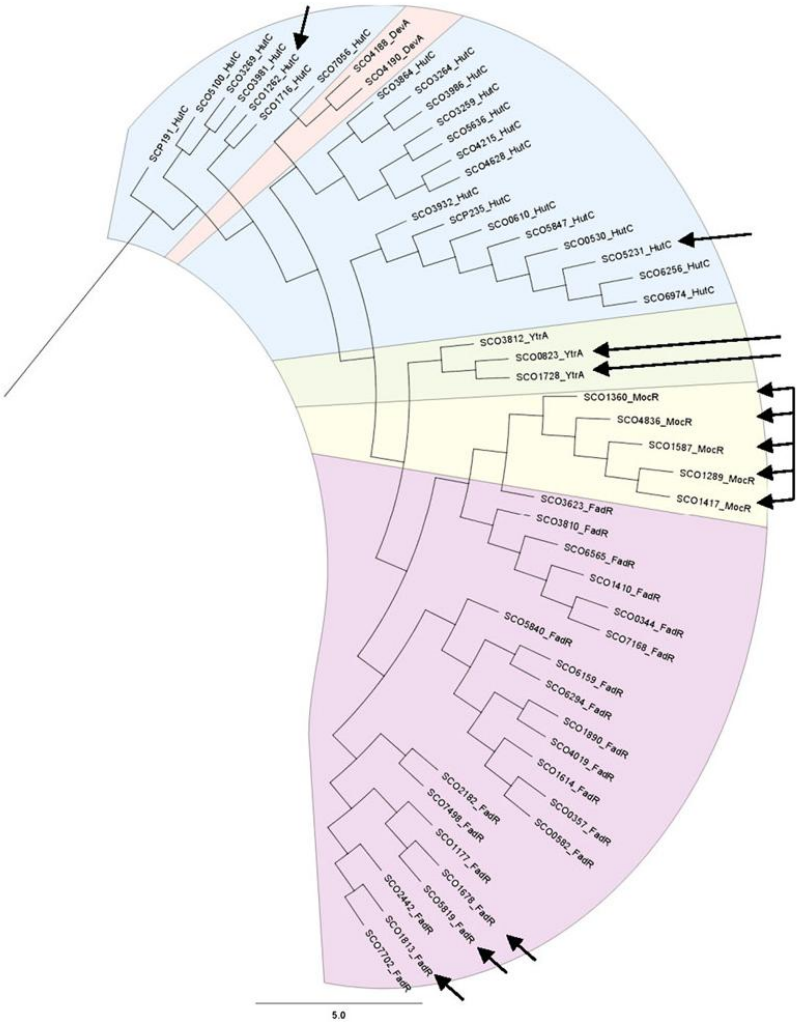


Fig. 1 Unrooted maximum parsimony tree schematically shows the relationships within a family of GntR-type regulators of S. coelicolor. Entire aa sequences of the proteins were used for analysis. The scale bar represents 5 aa substitution per aa position. Members of 5 subfamilies could be observed (highlighted with different colors). SCO5607 and SCO2845 form separate branches on the tree and were excluded from analysis (see main text). Arrows point to core GntRs

Results

Classification of putative GntR family proteins encoded by the *S. coelicolor* genome

According to StrepDB, there are 60 putative GntR-like regulatory genes in the *S. coelicolor* genome (chromosome plus plasmids). However, using several protein secondary structure prediction tools, we were able to confirm that only 56 of them (54 on chromosome and 2 on plasmid SCP1) possess the fold typical for certain subfamilies of GntRs. The other four proteins are more similar to members of the Mar or LacI families and are, most likely, mislabeled as GntR. We classified the 56 corresponding translation products into FadR, HutC, MocR, YtrA, and DevA subfamilies (these data are presented in Electronic Supplementary Materials (ESM), Table S1). None of the proteins were ascribed to the AraR and PlmA subfamilies. Two small regulators, Sco5607 and Sco2845, could not be unambiguously classified into any of the known subfamilies since they lacked completely or possessed only a truncated C-terminal domain, respectively. These proteins could be either founding members of new subfamilies or represent a process of gene decay. Members of different subfamilies clustered together on the unrooted maximum parsimony tree of GntR proteins (Fig. 1), in agreement

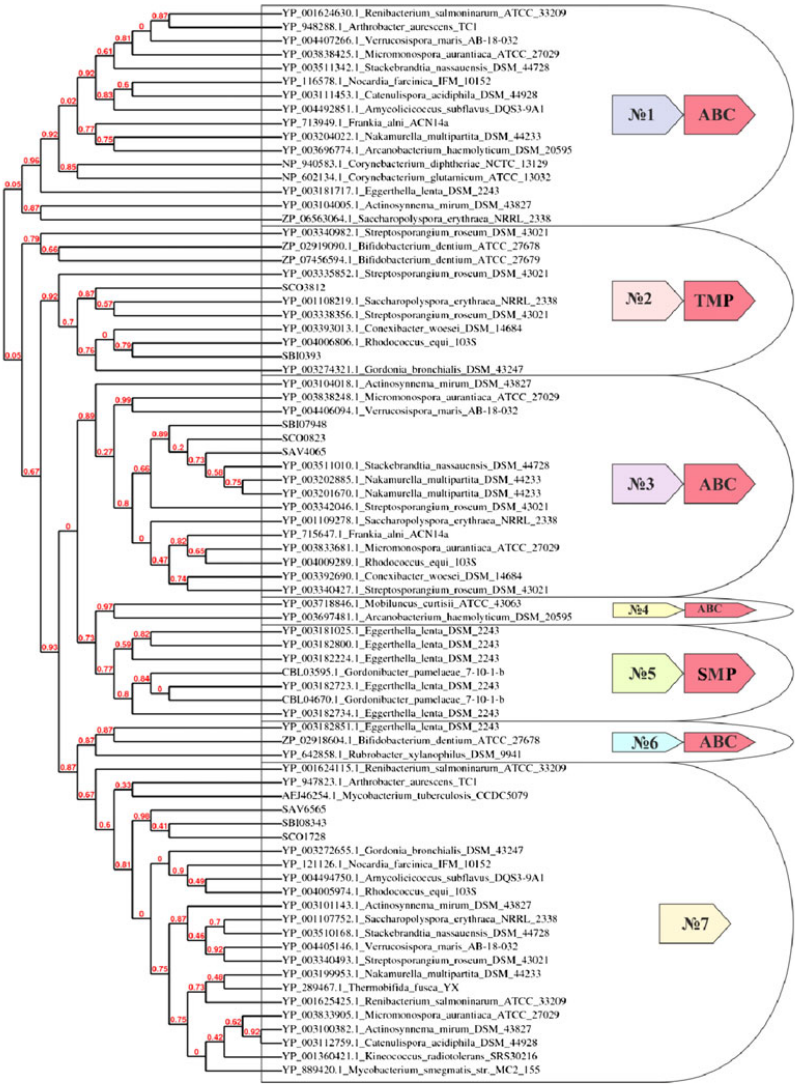


Fig. 2 Unrooted maximum-likelihood cladogram showing the relationships within of 78 YtrA-like proteins, each from different actinobacterial species. The tree was built with the help of phylogeny.fr v.2 server (Dereeper et al. 2008) using default parameters. Confidence indices aLRT are shown on the nodes. Cladogram shows clustering of YtrAs into seven clusters with genomic context of respective genes being in co-occurrence with: genes of ABC transporters, No. 1, 3, 4, 6; genes of conservative transmembrane proteins of unknown functions, o. 2; genes of stomatin-like membrane proteases, no. 5; localized singly, no. 7

with protein structure prediction. These sequence and structural comparisons suggest that DevA proteins have evolved from the members of HutC subfamily.

Of the 56 GntRs, 22 proteins fell into the FadR subfamily. The C-terminal domain of this subfamily consists of seven or six α-helices and can be further classified in FadR (9 proteins) and VanR (13 protein) subgroups. Sco0344 is much bigger than the other FadR-like proteins and includes additional α-helices upstream of HTH motif. Protein WhiH is the only studied FadR-like regulator; it is known to regulate septation of aerial hyphae and spore formation in *S. coelicolor* (Persson et al. 2013).

Twenty-two of the other GntRs were identified as the members of HutC subfamily. The C-terminal domain of HutC is a unique combination of α-helical and β-sheet structures (Hoskisson and Rigali 2009). These proteins regulate diverse processes such as amino acid uptake, plasmid transfer, and effector molecule responses. DasR, KorSA, and AglR3 are the studied HutC members; KorSA is involved in plasmid transfer (Sezonov et al. 2000). DasR and Agl3R are carbon

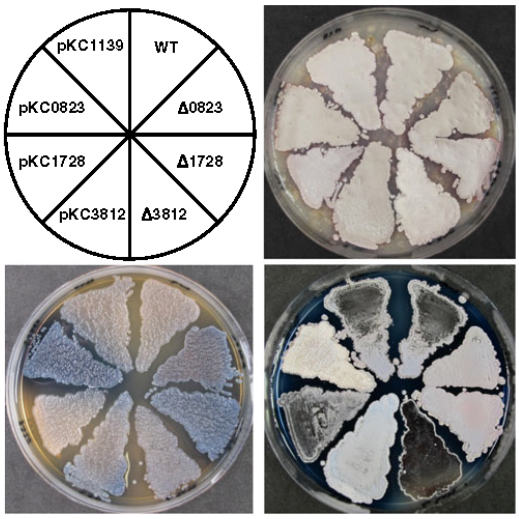


Fig. 3 Lawns of S. coelicolor knockout strains and those carrying additional copies on the plasmid. Abbreviations on the scheme of streaks placement (upper left corner): pKC0823, pKC1728, pKC3812—M145 derivatives carrying genes SCO0823, SCO1728, SCO3812, respectively, on replicative vector; pKC1139—M145 carrying empty vector pKC1139; WT, Δ0823, Δ1728, Δ3812—wild type (M145) and strains with knockouts of SCO0823, SCO1728, and SCO3812, respectively. Shown are the plates after 120 h of incubation (clockwise from the scheme): oatmeal agar, R2YE, TSB agar

source-responsive regulators required for morphogenesis and antibiotic production by *S. coelicolor* (Rigali et al. 2008; Hillerich and Westpheling 2006).

The third subfamily, MocR, is easily recognized by its long (about 350 amino acids) C-terminal domain showing homology to class I of aminotransferases. The MocR requires pyridoxal 5′-phosphate as a cofactor (Rigali et al. 2002). Only five *S. coelicolor* GntRs fell into this subfamily. One of them, Sco1417, was homologous to the *S. venezuelae* PdxR protein which is involved in the regulation of pyridoxal phosphate synthesis (Magarvey et al. 2001).

Three proteins were classified as members of the YtrA subfamily. The E-b/O domain of YtrA proteins is very short; it contains only two α-helices, raising doubts that this domain is capable of binding a ligand (Rigali et al. 2002). No *S. coelicolor* YtrA subfamily proteins have been studied to date.

Two DevA subfamily proteins are encoded in the *S. coelicolor* genome, and they have been shown to participate in *S. coelicolor* development (Hoskisson and Rigali 2009). As evident from structure predictions and phylogenetic analysis, DevA proteins are most similar to HutC proteins.

Ortholog prediction

Using reciprocal BLAST searches, a number of orthologs of *S. coelicolor* GntRs were identified in several sequenced streptomycete genomes (ESM, Table S2). Twelve *S. coelicolor* GntRs have orthologs in all the analyzed members of the family *Streptomycetaceae* (*S. griseus*, *S. venezuelae*, *S. clavuligerus*, *S. avermitilis*, *S. scabiei*, *S. bingchenggensis*, and *Kitasatospora setae*), and here, we refer to them as "core" GntRs. These genes are as follows:

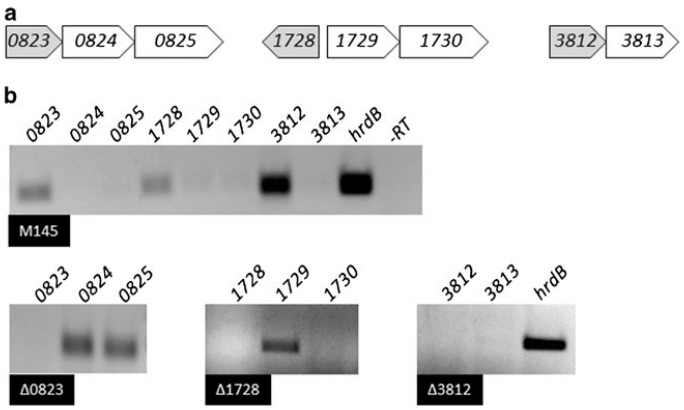


Fig. 4 Transcriptional analysis of selected ytrA-like genes and their neighbors in S. coelicolor after 72 h of growth in YMPG. **a** Genetic localization of sets of S. coelicolor genes analyzed in this study. YtrA-encoding genes are shown in gray. All RT-PCR primers were designed to produce approximately 290-bp amplicons in the vicinity of start codon of respective genes/mRNAs. **b** Semiquantitative RT-PCR analysis of selected sco genes in wild type (M145), sco0823, sco1728 and sco3812 knockout (Δ0823, Δ1728, and Δ3812, respectively) backgrounds. Gene hrdB was used as positive control. No hrdB amplicon was observed when M145 RNA in the absence of RT step (RT) was used as a template. The gel photos shown on this figure were inverted and uniformly contrasted using standard settings of AlphaImager software

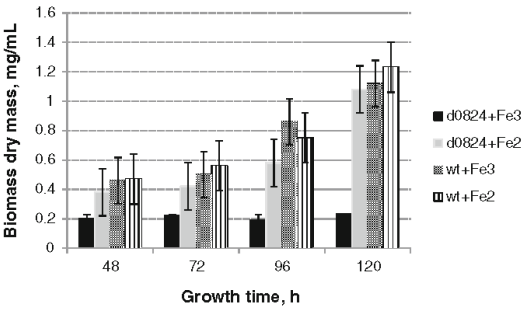


Fig. 5 Biomass accumulation by M145 (M145) and Δ0824 (d0824) strains during the growth in minimal glucose-containing medium supplemented with either ferrous (FeII) or ferric (FeIII) sulfate. Results are mean values of four independent experiments, error bars are ±2σ. More details on workup conditions can be found in ESM

SCO0823, SCO1262, SCO1289, SCO1360, SCO1417, SCO1587, SCO1678, SCO1728, SCO1813, SCO4836, SCO5231, SCO5819. Interestingly, one of them, SCO1728, has an ortholog in *Mycobacterium tuberculosis*, the only precedent for the entire GntR family of *S. coelicolor*. All five MocR subfamily proteins belong to the core GntRs, indicating their essential role in streptomycete biology. Two core GntRs, WhiH and DasR from the FadR and HutC subfamilies, respectively, have already been extensively investigated and shown to play major role in *Streptomyces* biology (Persson et al. 2013; Rigali et al. 2008). The functions of the ten other core GntRs remain unknown. Given their omnipresence in streptomycete genomes, we suggest that the core GntRs have important cellular role(s).

In silico analysis of actinobacterial YtrA-like regulators

Our in silico analysis showed that the YtrA subfamily of GntR-type regulators is one of the smallest and the least studied in actinomycetes. At the same time, two YtrA-like regulators, Sco0823 and Sco1728, are core GntRs, pointing to their importance in *Streptomyces* biology. We therefore decided to study YtrA subfamily in more detail. We used the sequence of *Bacillus subtilis* YtrA protein (the prototype for the subfamily) as a template in BlastP searches for homologues across actinobacterial genomes. After refinement, we obtained 17 sequences in 6 genomes of *Coriobacteridae*, 3 in 2 genomes of the *Rubrobacteridae* subclass, and 76 in 80 genomes of the *Actinobacteridae* subclass. Parasitic actinobacteria with small genomes lacked such sequences, as did highly specialized members of *Acidimicrobidae* and *Nitriliruptoridae* subclasses. We saw no correlation between genome size and the number of discovered sequences.

We have chosen several species from each actinobacterial family (78 in total) and used their YtrA-like protein sequences to build a maximum-likelihood consensus tree (Fig. 2). On the resulting tree, seven clades could be distinguished that coincide with certain patterns of genomic context. Particularly, the genes of actinobacterial YtrA-like proteins are most likely to be a part of operons with ABC transporter genes. This was observed for groups no. 1, 3, 4, and 6 (Fig. 2). Genes of group no. 2 are co-localized with genes for proteins of unknown function. Their secondary structure is enriched with hydrophobic regions, suggesting that they are membrane-spanning proteins. Almost all genes for group no. 5 proteins were found in genomes of genus *Eggerthella* and are localized near genes for stomatin-like extracellular proteases. Finally, the members of the group no. 7 did not appear to have specific genomic context. *S. coelicolor* YtrA representatives Sco3812, Sco0823, and Sco1728 clustered with groups no. 2, 3, and 7 of the consensus tree, respectively.

GntRs are often autoregulatory proteins, and analysis of promoters of their genes may provide additional information about the operator sequences that these GntRs recognize. We therefore screened for the conservative palindromic sequences in the putative promoter regions of the YtrA-encoding streptomycete genes using RegPredict web server (Novichkov et al. 2011). The promoter regions of *SCO3812* and its orthologs contained an imperfect palindrome 5′-cATtGTtcTAgTtnCATg-3′. Promoters of *SCO0823* orthologous group featured sequence 5′-CACtAnnCgATTAGTG-3′. The high-scoring copies of the latter were also found in a number of promoter regions of *S. coelicolor* genes encoding putative regulatory proteins. No conservative palindromes in promoters of the group of *SCO1728* orthologs were discovered with the bioinformatics tools being used. DNA logos of the identified operator sequences and list of genes whose upstream regions contains motifs similar the one found upstream of *SCO0823* are given in ESM Fig. S4.

Investigation of function of YtrA-encoding genes *SCO3812*, *SCO1728*, and *SCO0823*

We generated *S. coelicolor SCO3812*, *SCO1728*, and *SCO0823* knockout strains and M145 derivatives carrying these genes on replicative vectors: either pKC1139 (for *SCO0823* and *SCO3812*) or pKC1218E (for *SCO1728*; see

"Materials and Methods" and ESM). Wild type and the generated strains did not differ in biomass accumulation when growing in liquid media YEME, YMPG, and SMM over the period of 120 h; undecylprodigiosin production was also essentially the same for knockout strains (ESM, Fig. S5–S7). All strains were prototrophic and did not differ from the control ones in resistance to antibiotics. We observed more variable results when Act production was assayed in the recombinant strains under conditions of submerged cultivation. When compared to control strain (M145 harboring empty vector pKC1139), additional copies of *SCO3812* significantly increased Act yields starting from 72 h of growth, irrespective of the medium being used (ESM, Fig. S8). In YMPG, plasmid-borne *SCO0823* increased Act production, while *SCO1728* decreased it (Fig. S8). The observed differences lied within ±1.5-fold range of average Act production by the control strain. Deletion of *SCO3812* decreased Act production in YEME and YMPG, while no difference from the parental strain was observed in SMM (ESM, Fig. S9). The *SCO0823*-deficient strain produced roughly two times more Act than M145 strain in YMPG and SMM media (Fig. S9). Certain differences in sporulation and Act production were observed between the parental and recombinant strains on agar plates (Fig. 3 and ESM Figs S10–S12). Taking into account the data from both liquid and solid media, a general trend can be inferred in two cases with regard to Act production. Namely, the latter is improved upon either *SCO0823* disruption or introduction of additional *SCO3812* copies in M145. The other data (especially those related to *SCO1728*) were more difficult to generalize, probably due to complex nutrient-dependent nature of the mechanisms underlying changes in Act production. As to morphological development, we observed that sporulation of *SCO3812*-deficient mutant was impaired on most solid media, although the underlying reason (e.g., whether the defect lies at the stage of either spore or aerial mycelium formation) remains unknown. Also, the strain carrying pKCSCO0823 appeared to sporulate more abundantly than the initial (wild type) strain (Fig. 3 and ESM Figs. S10–S12).

We analyzed, in semiquantitative manner, the transcriptional activity of *sco0823*, *sco1728*, and *sco3812* genes as well as their neighbors, whose expression the aforementioned *ytrA*-like genes might control. Here, we focused on a single time point, 72 h of growth, and one liquid medium (YMPG); those were conditions where changes in Act production and morphology were documented (vide supra). The results of these experiments are summarized on Fig. 4. The data allowed us to conclude that *i. ytrA*-like genes are transcriptionally active in wild type strain and that *ii. sco0823* and *sco1728* downregulate the transcription of respective nearby genes. Gene *sco3812* is abundantly transcribed in M145 strain (Fig. 4). Nevertheless, its knockout did not lead to increased expression of downstream *sco3813* gene, as would be expected if Sco3812 was repressor of *sco3813* transcription.

Gaining deeper insight into function of *SCO0823*

Gene SCO0823: promoter-probe and heterologous expression experiments Of the three studied genes, *SCO0823* appeared the most interesting to us for several reasons. Firstly, unlike *SCO3812*, *SCO0823* encodes core GntR and it negatively regulates expression of downstream genes. Secondly, and in contrast to *SCO1728*, the genomic context of *SCO0823* is highly conserved across actinomycete genomes; putative Sco0823 operator sequence can be predicted; and effects of *SCO0823* mutation on Act production are straightforward. The gene *SCO0823* encodes a typical YtrA-like protein. It consists of conserved N-terminal DNA-binding domain and a small C-terminal E-b/o domain with only two α-helices. In the chromosome, *SCO0823* forms a putative operon with adjacent genes encoding ABC transporter ATP-binding and transmembrane subunits.

We attempted to investigate the *SCO0823* temporal expression profile in the wild type and *SCO0823* disrupted strains using the *gusA* reporter system. However, β-glucuronidase activity was low (as compared to the studied precedents; see Horbal et al. 2013; Makitrynskyy et al. 2013) and roughly the same in both strains over the entire incubation time (Fig. S13) suggesting that *SCO0823* promoter is not highly transcribed under our experimental conditions. We also undertook several attempts to overexpress Sco0823 protein using commercially available systems, such as pTYB2 (fusion to chitin-binding domain), pET24b (C-terminal His-tag), pET44a (N-terminal Nus-, Strep- and His-tag; C-terminal His-tag) and pTB146 (N-terminal His- and SUMO-tags). However, in no case were we able to purify soluble Sco0823 protein, which precluded biochemical studies of its function.

SCO0823 was also expressed in the streptomycete strains, producers of polyketide landomycin E *Streptomyces globisporus* 1912 (Ostash et al. 2011), phosphoglycolipid moenomycin *S. ghanaensis* ATCC14672 (Makitrynskyy et al. 2010) and thiopeptide siomycin, *S. sioyaensis* (Myronovskyy et al. 2009). Plasmid pKC0823 was introduced into these strains, and antibiotics production by the plasmid-borne and initial strains has been compared. We revealed that *SCO0823* repressed antibiotic production by *S. globisporus* and *S. ghanaensis* (approximately 1.5- and 2-fold decrease, respectively). In *S. sioyaensis*, *SCO0823* caused 2-fold increase in synthesis of siomycins (Fig. S14).

Disruption of SCO0824 The knockouts of Sco0823-controlled genes could shed more light on *SCO0823* function. Our transcriptional studies strongly suggested that the ABC transporter genes downstream of *SCO0823* are the first candidates for such studies. *SCO0824* encodes putative ABC

transporter ATP-binding protein, and *SCO0825* codes for transmembrane component of the transporter. Like *SCO0823*, *SCO0824-0825* genes are ubiquitous in actinomycete genomes, although specific roles for respective proteins in actinomycete metabolism could not be inferred from available data or from pairwise sequence alignments, such as BLASTP. Nevertheless, a structural homology search program HHPred (Söding et al. 2005) revealed the similarity of Sco0824-Sco0825 proteins to those involved in uptake of metals (particularly ferric ion and cobalt), aminoacids, or carbohydrates.

SCO0824 was disrupted using the available *S. coelicolor* transposon insertion knockout library (Bishop et al. 2004), as detailed in Fig. S15. The generated mutant (Δ0824) grew as well as the wild type (M145) did on various solid and liquid media (see also Fig. S16). Its resistance to antibiotics, cobalt, and zinc did not differ from that of initial strain (data not shown). In liquid medium YEME, Δ0824 produced roughly two times more Act than M145 did (Fig. S17). The enhanced ability of Δ0824 to produce Act was also observed on certain solid media, such as tryptic soy agar (Fig. S18). It also appeared that Δ0824 started sporulating earlier and produced spores more abundantly in comparison to M145 (as evident from colony counts of serial dilutions plated on SM).

Guided by the bioinformatic predictions of Sco0824-Sco0825 functions, we set up a number of tests to reveal the phenotypic difference between M145 and Δ0824. Firstly, we used a disc diffusion test to see the influence of Fe^{3+}, Co^{2+}, Zn^{2+} Ni^{2+} Mn^{2+}, and Cu^{2+} ions on these strains. Strains did behave differently in presence of zinc and copper, and these data are shown on Fig. S19. Particularly, zinc ions induced early and enhanced Act production by the mutant. Also, for Δ0824, we observed curious phenomenon of metal precipitation on mycelia in the area between discs soaked in $CuSO_4$ and $CoCl_2$. Secondly, If Sco0824 protein is required for uptake of certain essential compound, then the growth of the Δ0824 should be arrested or severely impaired in minimal media containing this compound. To this end, we focused on different carbon sources and two iron sources, ferrous and ferric sulfates. The growth of M145 and Δ0824 was similar on different aminoacids and carbohydrates (full list is given in Table S3, ESM). In minimal medium supplemented with ferric ions and glucose wild type produced significantly more biomass than Δ0824 (Fig. 5). Furthermore, and in contrast to M145, the growth of Δ0824 was planktonic and clumpy (Fig. S20).

Discussion

The family of GntR-type regulators is extensively represented in actinomycetes, although its significance for actinomycete biology remains largely unknown. Here, we report the bioinformatic analysis of 56 putative GntRs encoded within the genome of model actinomycete *S. coelicolor* as well as a more detailed examination of the YtrA subfamily proteins from this bacterium, which should facilitate future studies of these regulators.

What would be the evolutionary reasons for streptomycetes to keep such an extensive array of GntRs? Generally, there are many more transcriptional factors present in the genomes of soil-dwelling actinobacteria as compared to pathogenic ones. For instance, the 4.4 Mb genome of *M. tuberculosis* contains only 7 GntR-like regulators (Vindal et al. 2007a), while genomes of its saprophytic relatives *M. smegmatis* (≈7 Mb) and *S. coelicolor* (8.6 Mb), contain 62 and 56, respectively (Vindal et al. 2007b; Hoskisson and Rigali 2009). A similar situation is observed for pathogenic *Corynebacterium diphtheriae* (2.4 Mb; 5 GntR's) and free-living *C. glutamicum* (2.4 Mb; 11 GntR's) and other actinobacteria, as well as for other groups of regulatory genes. It is likely that more dynamic environments, such as soil, select for microorganisms having more regulatory systems for adaptating to fast-changing conditions. Nevertheless, not all 56 *S. coelicolor* GntRs are universally conserved in streptomycetes. Our analysis revealed only 12 GntRs that have orthologs in all six *Streptomycetaceae* being analyzed. We refer to them as core GntRs in this work, to highlight their likely essential role in *Streptomyces* biology. Since the total number of GntRs encoded within different *Streptomyces* genomes are comparable (data not shown), it suggests that evolution of GntRs within this genus was mainly driven by duplications and divergence in response to different environmental and cellular factors for different species. This eventually would lead to loss of orthology and the occurrence of functionally new proteins and subfamilies, as in the case of the DevA subfamily.

According to the results of secondary protein structure prediction, almost all GntRs of *S. coelicolor* can be classified into five different subfamilies. Very few studies have been done on any of these proteins, and this is especially true for the YtrA subfamily. In *S. coelicolor,* this subfamily is represented by only three proteins; nevertheless, two of them are core GntRs. Extensive in silico analysis showed that YtrA regulators are abundant in *Actinobacteria*, and they are different in terms of phylogeny and genomic context. Although our analysis supports a paradigmatic view of YtrA proteins as regulators of transcription of transporter genes, it also shows that other scenarios exist. For instance, in the genomes of intestinal symbionts of the genus *Eggerthella,* YtrA-encoding genes are located near protease genes, while *SCO1728* and its homologues do not have conserved neighbors. Furthermore, the presence of the conserved palindromic sequence upstream of core *SCO0823* and several regulatory genes provides indirect evidence for a wider role of YtrAs in streptomycete biology. In this context, it should be noted that YtrA proteins are traditionally viewed as transcriptional repressors. Our RT-PCR

data are in agreement with this notion for Sco0823 and Sco1728, whereas molecular function of Sco3812 remains murky. The simplest explanation would be that (despite overlap of *sco3812* and *sco3813* stop and start codons) Sco3812 probably controls the transcription of genes other than *sco3813*. It would be needed to test wider set of genes and growth conditions to reach better understanding of function of Sco3812.

A consensus recognition sequence for the YtrA subfamily $N_{(x)}$**GTN**$_{(y)}$TAN$_{(z)}$TAN$_{(y)}$**AC**N$_{(x)}$ has been proposed (Rigali et al. 2002; Hoskisson and Rigali 2009), and palindromic sequences similar to the cited above were found in promoter regions of *SCO3812* and *SCO0823*. The promoter regions of the *SCO1728* orthologous group do not seem to possess over-represented palindromic sites. Taking into account the variable genomic context of *SCO1728*, it implies that this protein has a role in actinomycete biology unrelated to control of transport processes.

S. coelicolor transposon mutant deficient in *SCO3812* gene was shown recently to possess impaired Act production on certain solid media (Bilyk et al. 2013). Here, we further show that *SCO3812* mutant produces less Act than the wild type on plates and in submerged culture. More interestingly, the mutant displayed altered morphology on most, although not all, tested solid media. All available data point to the possibility that Sco3812 responds to nutritional status. For the other two mutants, we failed to observe a specific phenotype other than modest changes in Act production. The reasons for that can be numerous, although we think that inability to find suitable selective conditions for phenotype expression is a major one. Under our experimental conditions, the transcriptional activity of the *SCO0823* promoter was below the level where the reliable difference can be detected by glucuronidase reporter. Because GntRs (and YtrA subfamily in particular) are often repressors, it is conceivable that their knockout might be not a loss-of-function mutation. Rather, it would lead to certain (probably moderate, given the results of promoter-probe experiments) increase in expression of responsive genes. We note that Sco0823, Sco1728, and Sco3812 proteins are similar to each other and may be functionally redundant, thus masking the effects of individual knockouts. It will be a challenge for researchers to devise efficient strategies to elucidate the functions of these highly specialized, cryptic, and often redundant genes within actinomycete genomes. Nevertheless, our studies demonstrate that once introduced in high copy number, YtrA regulatory genes do lead to certain changes in sporulation. Furthermore, in three streptomycetes, Sco0823 led to changes in antibiotic titers. Although the effects of Sco0823 were different for different strains, it is not uncommon to observe such a behavior for regulatory genes. For example, TetR-like regulator AveI represses avermectin production by *S. avermitilis* and enhances actinorhodin synthesis in *S. coelicolor* (Chen et al. 2008). Therefore, a final phenotype produced by transcriptional factor is dependent on species-differential peculiarities of interaction with overall regulatory network and location of cognate operator sequences.

It is necessary to identify and study the genes under control of YtrA regulators. Our initial studies of *SCO0824* gene clearly demonstrate this point, and allow us to propose tentatively, that it, and *SCO0823*, are involved in metal homeostasis in *S. coelicolor*. Most likely, *sco0823-0824-0825* operon is involved in ferric ion uptake, by analogy to a number of described ABC transport systems in Gram-positive and -negative bacteria (Koster 2005). More detailed studies of *SCO0823-SCO0825* genes are underway in our laboratory to better understand their roles. Our work demonstrates that in-depth investigation of the GntRs, and YtrA subfamily proteins in particular, is a worthwhile effort.

Acknowledgments This work was supported by grants M/256-2014 from the State Agency on Science, Innovation and Informatization of Ukraine (to B.O.) and Bg-98f from the Ministry of Education and Science of Ukraine (V.F.). Part of the project was carried out within the frame of the NIH grant R03TW009424 (to S.W. and VF). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Publication was partially based on the research provided by the grant support of the State Fund For Fundamental Research (project № F60/2-2015). O.T. was supported by Visby and DAAD fellowships. We thank Emma Doud (Northwestern University, USA) for editing of our manuscript; Prof. Paul Dyson (Swansea University, UK) is thanked for cosmids from *S.coelicolor* transposon library.

References

Bilyk B, Weber S, Myronovskyi M, Bilyk O, Petzke L, Luzhetskyy A (2013) *In vivo* random mutagenesis of streptomycetes using mariner-based transposon Himar1. Appl Microbiol Biotechnol 97: 351–359. doi:10.1007/s00253-012-4550-x

Bishop A, Fielding S, Dyson P, Herron P (2004) Systematic insertional mutagenesis of streptomycete genome: a link between osmoadaptation and antibiotic production. Genome Res 14:893–900. doi:10.1101/gr.1710304

Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL (2012) Domain enhanced lookup time accelerated BLAST. Biol Direct 7:12. doi:10.1186/1745-6150-7-12

Buchan DW, Ward SM, Lobley AE, Nugent TC, Bryson K, Jones DT (2010) Protein annotation and modelling servers at University College London. Nucl Acids Res 38(Suppl):W563–W568

Chen L, Lu Y, Chen J, Zhang W, Shu D, Qin Z, Yang S, Jiang W (2008) Characterization of a negative regulator AveI for avermectin biosynthesis in *Streptomyces avermitilis* NRRL8165. Appl Microbiol Biotechnol 80:277–286. doi:10.1007/s00253-008-1545-8

Cole C, Barber JD, Barton GJ (2008) The Jpred 3 secondary structure prediction server. Nucleic Acids Res 36:W197–201. doi:10.1093/nar/gkn238

Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97:6640–6645

Dereeper A, Guignon V, Blanc G, Audic S et al (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36:W465–9. doi:10.1093/nar/gkn180

Gust B, Chandra G, Jakimowicz D, Yuqing T, Bruton CJ, Chater KF (2004) Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. Adv Appl Microbiol 54:107–128

Hillerich B, Westpheling J (2006) A new GntR family transcriptional regulator in *Streptomyces coelicolor* is required for morphogenesis and antibiotic production and controls transcription of an ABC transporter in response to carbon source. J Bacteriol 188(21): 7477–87

Horbal L, Kobylyanskyy A, Yushchuk O, Zaburannyi N, Luzhetskyy A, Ostash B, Marinelli F, Fedorenko V (2013) Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus*—producer of teicoplanin, drug of last defense. J Biotechnol 168:367–372. doi:10.1016/j.jbiotec.2013.10.018

Hoskisson PA, Rigali S (2009) Chapter 1: variation in form and function the helix-turn-helix regulators of the GntR superfamily. Adv Appl Microbiol 69:1–22. doi:10.1016/S0065-2164(09)69001-8

Kelley LA, Sternberg MJ (2009) Protein structure prediction on the web: a case study using the phyre server. Nat Protoc 4(3):363–71. doi:10.1038/nprot.2009.2

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical Streptomyces genetics. John Innes Foundation, Norwich

Koster W (2005) Cytoplasmic membrane iron permease systems in the bacterial cell envelope. Front Biosci 10:462–477. doi:10.2741/1542

Lewis RA, Shahi SK, Laing E, Bucca G, Efthimiou G, Bushell M, Smith CP (2011) Genome-wide transcriptomic analysis of the response to nitrogen limitation in Streptomyces coelicolor A3(2). BMC Res Notes 4:78. doi:10.1186/1756-0500-4-78

Liu G, Chater KF, Chandra G, Niu G, Tan H (2013) Molecular regulation of antibiotic biosynthesis in Streptomyces. Microbiol Mol Biol Rev 77(1):112–43. doi:10.1128/MMBR.00054-12

Magarvey N, He J, Aidoo KA, Vining LC (2001) The pdx genetic marker adjacent to the chloramphenicol biosynthesis gene cluster in Streptomyces venezuelae ISP5230: functional characterization. Microbiology 147(Pt 8):2103–12

Makitrynskyy R, Rebets Y, Ostash B, Zaburannyi N, Rabyk M, Walker S, Fedorenko V (2010) Genetic factors that influence moenomycin production in streptomycetes. J Ind Microbiol Biotechnol 37:559–566

Makitrynskyy R, Ostash B, Tsypik O, Rebets Y, Doud E, Meredith T, Luzhetskyy A, Bechthold A, Walker S, Fedorenko V (2013) Pleiotropic regulatory genes bldA, adpA and absB are implicated in production of phosphoglycolipid antibiotic moenomycin. Open Biol 3:130121. doi:10.1098/rsob.130121

McCormick JR, Flärdh K (2012) Signals and regulators that govern Streptomyces development. FEMS Microbiol Rev 36:206–231

Muth G, Nussbaumer B, Wohlleben W, Pühler A (1989) A vector system with temperature-sensitive replication for gene disruption and mutational cloning in streptomycetes. Mol Gen Genet 6:1–8

Myronovskyi M, Welle E, Fedorenko V, Luzhetskyy A (2011) Beta-glucuronidase as a sensitive and versatile reporter in actinomycetes. Appl Environ Microbiol 77(15):5370–83. doi:10.1128/AEM.00434-11

Myronovskyy M, Ostash B, Ostash I, Fedorenko V (2009) A gene cloning system for the siomycin producer Streptomyces sioyaensis NRRL-B5408. Folia Microbiol 54:91–96

Nett M, Ikeda H, Moore BS (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. Nat Prod Rep 26(11): 1362–84. doi:10.1039/b817069j

Novichkov PS, Rodionov DA, Stavrovskaya ED, Novichkova ES, Kazakov AE, Gelfand MS, Arkin AP, Mironov AA, Dubchak I (2011) RegPredict: an integrated system for regulon inference in prokaryotes by comparative genomics approach. Nucleic Acids Res 38:W299–307. doi:10.1093/nar/gkq531

Ostash B, Rebets Y, Myronovskyy M, Tsypik O et al (2011) Identification and characterization of the Streptomyces globisporus 1912 regulatory gene lndYR that affects sporulation and antibiotic production. Microbiology 157(Pt 4):1240–9. doi:10.1099/mic.0.045088-0

Persson J, Chater KF, Flärdh K (2013) Molecular and cytological analysis of the expression of Streptomyces sporulation regulatory gene whiH. FEMS Microbiol Lett 341(2):96–105. doi:10.1111/1574-6968.12099

Rigali S, Derouaux A, Giannotta F, Dusart J (2002) Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies. J Biol Chem 277(15):12507–15

Rigali S, Titgemeyer F, Barends S, Mulder S, Thomae AW, Hopwood DA, van Wezel GP (2008) Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by Streptomyces. EMBO Rep 9(7):670–5. doi:10.1038/embor.2008.83

Salzberg LI, Luo Y, Hachmann AB, Mascher T, Helmann JD (2011) The Bacillus subtilis GntR family repressor YtrA responds to cell wall antibiotics. J Bacteriol 193(20):5793–801. doi:10.1128/JB.05862-11

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor

Sezonov G, Possoz C, Friedmann A, Pernodet JL, Guérineau M (2000) KorSA from the Streptomyces integrative element pSAM2 is a central transcriptional repressor: target genes and binding sites. J Bacteriol 182(5):1243–1250

Simossis VA, Heringa J (2003) The PRALINE online server: optimising progressive multiple alignment on the web. Comput Biol Chem 27(4–5):511–9

Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:W244–W248

van Wezel GP, McDowall KJ (2011) The regulation of the secondary metabolism of Streptomyces: new links and experimental advances. Nat Prod Rep 28(7):1311–33. doi:10.1039/c1np00003a

Vindal V, Suma K, Ranjan A (2007a) GntR family of regulators in Mycobacterium smegmatis: a sequence and structure based characterization. BMC Genomics 23:289–301

Vindal V, Ranjan S, Ranjan A (2007b) In silico analysis and characterization of GntR family of regulators from Mycobacterium tuberculosis. Tuberculosis 87:242–247

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s12223-015-0426-7.
Електронний додаток до статті доступний за посиланням:
https://static-content.springer.com/esm/art%3A10.1007%2Fs12223-015-0426-7/MediaObjects/12223_2015_426_MOESM1_ESM.pdf.

Анотація: *Транскрипційні фактори родини GntR задіяні в бактерій для регуляції чисельних фізіологічних та морфологічних процесів у відповідь на сигнали наявності факторів живлення. Кількість генів, що кодують GntR-подібні транскрипційні фактори, у актиноміцетів є найвищою серед бактерій. Це відображає великі геноми актиноміцетів і складні екологічні ніші, які займають актиноміцети. Однак, існує мало даних про роль GntR-подібних регуляторів у біології актиноміцетів. В цій роботі ми проаналізували геном модельного стрептоміцета Streptomyces coelicolor A3(2) задля отримання нових даних про роль GntR-регуляторів. Всього було знайдено 56 генів, що кодують такі регулятори, які були класифіковані в підродини FadR, HutC, MocR, YtrA і DevA, відповідно до вторинної структури білків. Потім ми перевірили присутність ортологів цих 56 генів в 6 геномах стрептоміцетів та одному геномі Kitasatospora, знайшовши 12 консервативних регуляторів. Порівняльний геномний аналіз розповсюдження YtrA-подібних регуляторів в представників підкласів Coriobacteridae, Rubrobacteridae і Actinobacteridae виявив послідовності 96 таких білків, кодовані в 88 геномах. Філогенетично, ці білки утворювали сім щільних кластерів, а їх гени були найчастіше колокалізованими з генами транспортерів. Ймовірні операторні сайти були знайдені для білків з ортологічних груп SCO0823* і *SCO3812. Три гени YtrA регуляторів S. coelicolor (SCO0823, SCO1728, SCO3812) було проаналізовано на транскрипційному рівні, нокаутовано та надекспресовано. Також було вивчено ген SCO0824, потенційний компонент оперона SCO0823. Результати всіх цих експериментів подано і обговорено в статті.*

РОЗДІЛ 4 Підходи до генно-інженерних маніпуляцій із *Actinoplanes rectilineatus* NRRL B-16090 і *Streptomyces roseochromogenes* NRRL 3504

Публічно-доступні геноми актиноміцетів дають змогу проаналізувати потенціал цих бактерій до продукції цікавих вторинних метаболітів і обрати пріоритетних кандидатів для експериментального вивчення. У цій роботі, нашу увагу привернули такі організми як *A. rectilineatus* NRRL B-16090 і *S. roseochromogenes* NRRL 3504, в геномах яких ми знайшли потенційні КБГ пептидних антибіотиків. В NRRL B-16090 це був КБГ ГПА подібний до КБГ ристоцетину [86], а в NRRL 3504 – КБГ, подібний до КБГ депсипептидного антибіотика салінаміду [87]. Як перший крок до подальшого вивчення цих перспективних КБГ, необхідно було розробити ефективні методи генно-інженерних маніпуляцій із NRRL B-16090 і NRRL 3504. Доступність NRRL B-16090 для генно-інженерних маніпуляцій дотепер не вивчалася, бо штам не був об'єктом генетичних досліджень. NRRL 3504 же, хоча і є відомим продуцентом амінокумаринового антибіотика хлоробіоцину [88], був описаний як надзвичайно складний об'єкт для генно-інженерних маніпуляцій [89, 90].

***Створення систем експресії генів в A. rectilineatus* NRRL B-16090.** Актинобактерії, які належать до роду *Actinoplanes*, унікальні тим, що формують спори у спорангіальних структурах [49], а також є складними об'єктами для генно-інженерних маніпуляцій [29]. Для того, щоб знайти оптимальні умови для формування спорангіїв, і, відповідно, спор, ми вивчили морфологію NRRL B-16090 за умов росту на різних агаризованих середовищах. За допомогою СЕМ ми виявили, що спорангії формуються на середовищі ISP3 за умов пролонгованого культивування (>240 год). Використовуючи культури, вирощені за таких умов, нам вдалося відпрацювати методику отримання концентрованих спорових суспензій, що є важливою передумовою для перенесення плазмід.

Інтегративні вектори, засновані на системах інтеграції актинофагів широко використовуються в генетиці актиноміцетів як знаряддя для експресії генів [38]. Це вектори, що використовують *int-attP* модулі актинофагів φC31, VWB1 і

φBT1 [39, 91–93]. Додатково використовуються реплікативні вектори на основі pSG5-реплікона [41]. Ми вирішили перевірити, чи доступний NRRL B-16090 для перенесення таких плазмід методом міжродової кон'югації з використанням клітин *E. coli* ET12567 pUZ8002^{+} як донорів, а спор NRRL B-16090 – як клітин-реципієнтів.

Щоби полегшити скринінг транскон'югантів, ми створили похідні pSET152 (система інтеграції φC31), pRT801 (система інтеграції φBT1), pSOK804 (система інтеграції VWB1) і pKC1139 (pSG5-реплікон) – pSAGA (див. **Розділ 3**), pRAGA1, pSoAGA2 і pKAGA1, відповідно, що б несли репортерний ген β-глюкуронідази *gusA* під контролем конститутивного промотора *aac(3)IVp*. NRRL B-16090 виявився не здатним до β-глюкуронідазної активності, тому отримані транскон'юганти можна було б легко верифікувати за здатністю до хромогенної конверсії безбарвного субстрату 5-бромо-4-хлоро-3-індоліл-β-D-глюкуроніду в забарвлений 5,5′-дибромо-4,4′-дихлоро-індиго.

В результаті, нам вдалося перенести в клітини NRRL B-16090 лише дві плазміди: pSAGA – на основі системи інтеграції фага φC31 і pKAGA1 – на основі pSG5 реплікона. Оптимальними умовами для кон'югаційного перенесення виявилося середовище SFM із додаванням 40 мМоль $MgCl_2$, на якому частота перенесення pSAGA становила прибл. 10^{-3}, а pKAGA1 – прибл. 10^{-5}. Плазміди на основі систем інтеграції фагів φBT1 і VWB1 в NRRL B-16090 перенести не вдалося за жодних досліджених умов.

Отже, плазміди на основі системи інтеграції φC31 виявилися ефективним знаряддям для перенесення генів у NRRL B-16090. Розвиваючи цей напрямок, ми вирішили перевірити активність ряду гетерологічних промоторів, щоб обрати найефективніший для експресії генів у NRRL B-16090. Для цього, ми використали такі плазміди: pSAGA, pTEGA (несе *gusA* під контролем конститутивного промотора *ermEp*), pGUSmoeE5script (несе *gusA* під контролем конститутивного промотора *moeE5p*), pGT2p (несе *gusA* під контролем конститутивного промотора *tei2p*, див. **Розділ 8** цієї роботи) і pGCymRP21 (несе *gusA* під контролем кумат-індуцибельного промотора *P21*). Кількісний та якісний аналіз β-глюкуронідазної активності отриманих рекомбінантів дав

змогу визначити, що промотори *aac(3)IVp, moeE5p* і *tei2p* забезпечують сильну експресію гена *gusA*, а отже є ефективними знаряддями для надекспресії генів в NRRL B-16090.

Отримані результати роблять *A. rectilineatus* NRRL B-16090 – представника т.зв. «рідкісних актиноміцетів» – доступним для генно-інженерних маніпуляцій і спрощують подальше вивчення продукції нового ГПА цим штамом.

Створення набору знарядь для маніпулювання вторинним метаболомом S. roseochromogenes NRRL 3504. На даний момент, для *S. roseochromogenes* NRRL 3504 описане використання лише векторів на основі системи інтеграції фага φC31 [94]. Володіючи послідовністю генома NRRL 3504 [95], ми перевірили її на наявність *attB*-сайтів актинофагів φC31, VWB1 і φBT1, які були знайдені.

В наступному кроці ми вирішили перевірити, чи можна перенести плазміди на основі згаданих вище систем інтеграції і pSG5 реплікона в клітини NRRL 3504 за допомогою міжродової кон'югації з *E. coli* ET12567 pUZ8002^{+}. Дотепер, основним методом перенесення плазмід в клітини NRRL 3504 була трансформація протопластів; успішна спроба міжродової кон'югації також була описана, однак в цьому випадку саме вегетативний міцелій штаму використовувався як реципієнт [94]. Обидва підходи є складними у виконанні та не підходять для рутинних робіт. Тому, ми вирішили використати спори NRRL 3504 в якості клітин-реципієнтів. Підібравши оптимальні умови для споруляції NRRL 3504 (вирощування на ISP3-агарі протягом 168 год), ми розробили методику для приготування концентрованих спорових суспензій NRRL 3504. Використовуючи спори як клітини-реципієнти, нам вдалося успішно перенести в NRRL 3504 ряд плазмід, а саме: pSAGA (система інтеграції φC31), pRAGA1 (система інтеграції φBT1), pSoAGA2 (система інтеграції VWB1) і pKAGA1 (pSG5-реплікон). Усі ці плазміди несуть репортерну касету *aac(3)IVp-gusA*, яка дала змогу виявити, що всі рекомбінантні штами проявляють β-глюкуронідазну активність на відміну від штаму дикого типу – NRRL 3504.

Для подальшого вивчення вторинного метаболому NRRL 3504 необхідний штам, не здатний до продукції хлоробіоцину. Тому, нас зацікавило, чи можна застосувати суїцидні плазміди для інсерційного нокауту генів в NRRL 3504 за один етап гомологічної рекомбінації. Для цього ми створили плазміду pKOClo на основі суїцидного вектора pKC1132 [38], клонувавши в нього внутрішню ділянку оперона *cloH-I* (ключовий для біосинтезу хлоробіоцину оперон). Використавши описаний вище протокол для міжродової кон'югації, ми перенесли pKOClo в NRRL 3504. Отримані рекомбінантні штами виявилися не здатними до продукції хлоробіоцину, що випливало із результатів тестів на пригнічення росту Грам-позитивної тест-культури *M. luteus* ATCC 4698 і ВЕРХ-МС.

У **Розділі 3**, ми описали елементи *adpA*-опосередкованого регуляторного каскаду, а саме гени *adpA* і *bldA*. Ми перевірили, чи здатні гетерологічні гени *adpA* і *bldA*, отримані із модельного об'єкта *S. albus* (=*albidoflavus*) J1074, позитивно впливати на вторинний метаболізм NRRL 3504 на прикладі шляху біосинтезу хлоробіоцину. Рекомбінантні штами NRRL 3504 з надекспресіями генів *adpA* і *bldA* продукували прибл. в 3 і 2,5 рази більше хлоробіоцину, відповідно, як випливало із даних тестів на пригнічення росту *M. luteus* ATCC 4698 і ВЕРХ-МС.

В останньому кроці, ми перевірили ефективність ряду платформ для надекспресії генів у NRRL 3504 на прикладі надекспресії гена шлях-специфічного регулятора біосинтезу хлоробіоцину *cloG* [96], а також надекспресії його ортолога із КБГ новобіоцину (спорідненого амінокумаринового антибіотика) – *novG* [97]. Для цього, регуляторні гени клонували в такі вектори: pmoeE5script (система інтеграції φC31, під контроль конститутивного промотора *moeEp* [98]); pTES (система інтеграції φC31, під контроль конститутивного промотора *ermEp* [99]); pGCymRP21 (система інтеграції φC31, під контроль кумат-індуцибельного промотора *P21* [100]); pIJ6902 (система інтеграції φC31, під контроль тіострептон-індуцибельного промотора *tipAp* [101]); pOOB83e (pSG5-реплікон, під контроль конститутивного *ermEp* [102]). Всі отримані штами надпродукували

хлоробіоцин, демонструючи високу ефективність обраних платформ експресії. Найвищі показники продукції антибіотика продемонстрували рекомбінанти, які несли *cloG* і *novG* у складі pGCymRP21 (за умов індукції куматом), продукуючи в 20 і 32 рази більше хлоробіоцину, ніж вихідний NRRL 3504. В кількісному перерахунку це становило 169 і 239 мг/л, що є найвищим зареєстрованим в літературі показником продукції хлоробіоцину.

Отримані результати значно оновлюють та розширюють інструментарій для генно-інженерних маніпуляцій з *S. roseochromogenes* NRRL 3504, відкриваючи шлях для подальшого вивчення цього штаму в якості продуцента нового депсипептидного антибіотика.

Основні результати, отримані в цьому розділі, опубліковані в [103, 104].

Journal of Applied Genetics
https://doi.org/10.1007/s13353-019-00534-7

MICROBIAL GENETICS • ORIGINAL PAPER

Development of a gene expression system for the uncommon actinomycete *Actinoplanes rectilineatus* NRRL B-16090

Oleksandr Yushchuk[1,2] · Vitalina Homoniuk[1] · Yurij Datsiuk[3] · Bohdan Ostash[1] · Flavia Marinelli[2] · Victor Fedorenko[1]

Received: 13 August 2019 / Revised: 13 October 2019 / Accepted: 27 December 2019
© Institute of Plant Genetics, Polish Academy of Sciences, Poznan 2020

Abstract
The urgent need for discovering new bioactive metabolites prompts exploring novel actinobacterial taxa by developing appropriate tools for their genome mining and rational genetic engineering. One promising source of new bioactive natural products is the genus *Actinoplanes*, a home to filamentous sporangia-forming actinobacteria producing many important specialized metabolites such as teicoplanin, ramoplanin, and acarbose. Here we describe the development of a gene expression system for a new *Actinoplanes* species, *A. rectilineatus* (NRRL B-16090), which is a potential producer of moenomycin-like antibiotics. We have determined the optimal conditions for spore formation in *A. rectilineatus* and a plasmid transfer procedure for its engineering via intergeneric *E. coli–A. rectilineatus* conjugation. The φC31- and pSG5-based vectors were successfully transferred into *A. rectilineatus,* but φBT1- and VWB-based vectors were not transferable. Finally, using the glucuronidase reporter system, we assessed the strength of several heterologous promoters for gene expression in *A. rectilineatus.*

Keywords *Actinoplanes* · Sporangia · Glucuronidase assay · Intergeneric conjugation · Promoter activity

Introduction

For many decades, both fundamental and applied research on mycelium-forming Actinobacteria has been mostly restricted to one genus: *Streptomyces*. Among the soil-inhabiting filamentous actinomycetes, streptomycetes are relatively easy to isolate and cultivate. Consequently, specialized genetic engineering toolkits have been developed to manipulate the representatives of the aforementioned genus. However, the other less easy-to-handle genera of Actinobacteria (so-called uncommon or rare actinomycetes) have attracted an increasing interest in recent years (Marcone et al. 2010a). Among them, genus *Actinoplanes* is standing out for two main reasons: first, *Actinoplanes* strains possess a highly sophisticated life cycle that includes differentiating sporangia and motile spores; second, actinoplanetes produce a plethora of structurally complex and pharmaceutically important secondary metabolites. *Actinoplanes* spp. have entered the molecular era more than 15 years ago with the works on friulimycin biosynthesis in *Actinoplanes friuliensis* (Heinzelmann et al. 2003; Heinzelmann et al. 2005; Müller et al. 2007). Today, however, only a few other members of *Actinoplanes* genus are accessible for genetic manipulations, including producers of acarbose (*Actinoplanes* sp. SE50/110, Gren et al. 2016; Wolf et al. 2016; Wolf et al. 2017), teicoplanin (*Actinoplanes teichomyceticus*, Horbal et al. 2012; Horbal et al. 2014a; Yushchuk et al. 2016a; Yushchuk et al. 2019), and ramoplanin (*Actinoplanes ramoplaninifer*, Pan et al. 2013). Other actinoplanetes remain a kind of "dark matter" as it is unknown whether the available described genetic methods are directly applicable to them. A promising diversity of valuable natural compounds could be lurking in this "dark matter"—hardly accessible without proper genetic toolkit.

Actinoplanes rectilineatus NRRL B-16090 (Lechevalier and Lechevalier 1975) attracted our interest in the course of

Communicated by: Agnieszka Szalewska-Palasz

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13353-019-00534-7) contains supplementary material, which is available to authorized users.

✉ Victor Fedorenko
viktor.fedorenko@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 4 Hrushevskoho St, Lviv 79005, Ukraine

[2] Department of Biotechnology and Life Sciences, University of Insubria, via J.H. Dunant 3, 21100 Varese, Italy

[3] Department of Physics of Earth, Ivan Franko National University of Lviv, 4 Hrushevskoho st, Lviv 79005, Ukraine

Published online: 07 January 2020

ongoing work on morphogenesis and secondary metabolism of *A. teichomyceticus*. In contrast to *A. teichomyceticus* possessing globose-sporangia, *A. rectilineatus* possesses unique cylinder-shaped sporangia. We believe that a comparative study of *A. rectilineatus* and *A. teichomyceticus* could give valuable new insights into our knowledge of sporangia formation. An *A. rectilineatus* draft genome (although not the best quality) is available (ASM96282v1). Our initial analysis revealed the presence of a biosynthetic gene cluster for a moenomycin-like compound (ESM Fig. 1). Moenomycins are phosphoglycolipid antibiotics with great therapeutic potential (Ostash and Walker 2010). The list of moenomycin producers is currently limited to a few taxonomically related *Streptomyces* species and *A. teichomyceticus* (Horbal et al. 2016). In this work, we describe the development of a basic genetic toolkit for *A. rectilineatus*. Particularly, we studied the morphology of this organism growing on different solid media and determined optimal conditions for spore production. Then we optimized a protocol for conjugal transfer of replicative and integrative vectors into this strain. The φC31-based integrative vectors as well as pSG5-replicon-based replicative vector were transferred into *A. rectilineatus*. Finally, we introduced a GusA-based reporter system in *A. rectilineatus* for studying of the activity of several constitutive and inducible promoters in this strain. Our results make *A. rectilineatus*—a rare actinomycete with an interesting morphology and potential to produce valuable secondary metabolites—accessible for genetic engineering.

Materials and methods

Bacterial strains, plasmids, and culture conditions Bacterial strains and plasmids, used in this work, are listed in Tables 1 and 2. For morphological studies, *A. rectilineatus* strains were cultivated on ISP2 (Marcone et al. 2017), ISP3, TSA (Koshla et al. 2017), SM, SMMS, and R5 (Kieser et al. 2000) media. All the media components were purchased from Sigma-Aldrich, except the TSA (Condalab, Spain), soybean meal (Hensel, Germany), and finely ground whole oats (Kozub, Ukraine). ISP3 medium was used to induce the formation of sporulating lawns. Liquid ISP2 was used for growth under submerged conditions on a rotary shaker with 220 rpm at 30 °C. *Escherichia coli* strains were grown in either liquid or solid LB-medium. *A. rectilineatus* and derivatives were grown at 30 °C, *E. coli* strains—at 37 °C.

***A. rectilineatus* spore suspension preparation** The lawns of *A. rectilineatus* were grown for 12 days on ISP3. Sporangia were flushed from the surface of the lawn with distilled water. Motile spores were released from sporangia after 10 min of incubation on a rotary shaker at 200 rpm, 30 °C. The suspensions of swimming spores, as well as sporangiferous hyphae and sporangial coating debris, were filtered through a cotton wool. Spores were pelleted from final suspensions at 9.000 rpm for 5 min, washed with distilled water, resuspended in 10.3% (*w*/*v*) sucrose and stored at − 80 °C. For estimation of spore titers, serial dilutions were plated on ISP2 and the number of colony forming units (CFUs) was calculated.

Intergeneric conjugation with *E. coli* *E. coli* ET12567 pUZ8002$^+$ was used as a donor for plasmid conjugal transfer into *A. rectilineatus*. 10^6 spores of *A. rectilineatus* were mixed with 10^7 *E. coli* cells carrying the different vectors and plated on either SM or ISP3 agar containing 10 or 40 mM $MgCl_2$. After 15–20 h of incubation, plates were overlaid with 1 ml of distilled water containing 1.25 mg of apramycin-sulfate, 1.25 mg of 5-bromo-4-chloro-3-indolyl β-D-glucuronide

Table 1 Bacterial strains used in the study

Name	Description	Source
A. rectilineatus NRRL B-16090	Initial strain	NRRL
E. coli DH5α	General cloning host	MBI Fermentas
E. coli ET12567 pUZ8002$^+$	(*dam*-13::Tn9 *dcm*-6), pUZ8002+ (Δ*oriT*), used for conjugative transfer of DNA	Kieser et al. (2000)
A. rectilineatus pSAGA$^+$	*A. rectilineatus* NRRL B-16090 derivative carrying pSAGA vector	This work
A. rectilineatus pTEGA$^+$	*A. rectilineatus* NRRL B-16090 derivative carrying pTEGA vector	This work
A. rectilineatus pGUSmoeE5script$^+$	*A. rectilineatus* NRRL B-16090 derivative carrying pGUSmoeE5script vector	This work
A. rectilineatus pGT2p$^+$	*A. rectilineatus* NRRL B-16090 derivative carrying pGT2p vector	This work
A. rectilineatus pGCymRP21$^+$	*A. rectilineatus* NRRL B-16090 derivative carrying pGCymRP21 vector	This work

Table 2 Plasmid vectors used in the study

Name	Description	Source
pTES	pSET152 derivative, carrying: φC31 *attP* flanked by *loxP* sites; *ermEp* flanked by *tfd* terminator sequences; *aac(3)IV* apramycin resistance gene; φC31 integrase gene	Herrmann et al. (2012)
pTEGA	pTES derivative, carrying *gusA* under the control of *ermEp*	This work
pSAGA	pSET152 derivative, carrying: φC31 *attP*; *gusA* under the control of *aac(3)IVp*; *aac(3)IV* apramycin resistance gene; φC31 integrase gene	Koshla et al. (2019)
pRT801	φBT1-based integrative vector, carrying: φBT1 *attP*; *aac(3)IV* apramycin resistance gene; φBT1 integrase gene	Gregory et al. (2003)
pRAGA1	pRT801 derivative, carrying *gusA* under the control of *aac(3)IVp*	This work
pSOK804	VWB-based integrative vector, carrying: VWB *attP*; *aac(3)IV* apramycin resistance gene; VWB integrase gene	Sekurova et al. (2004)
pSoAGA2	pSOK804 derivative, carrying *gusA* under the control of *aac(3)IVp*	This work
pKC1139	pSG5-replicon based replicative vector, carrying: replication initiator protein gene, *aac(3)IV* apramycin resistance gene	Kieser et al. (2000)
pKAGA1	pKC1139 derivative, carrying *gusA* under the control of *aac(3)IVp*	This work
pGUSmoeE5script	pGUS derivative, carrying: φC31 *attP*; *gusA* under the control of *moeE5p*; *aac(3)IV* apramycin resistance gene; *aadA* streptomycin/spectinomycin resistance gene; φC31 integrase gene	Makitrynskyy et al. (2013)
pGT2p	pGUS derivative, carrying: φC31 *attP*; *gusA* under the control of *tei2p*; *aac(3)IV* apramycin resistance gene; *aadA* streptomycin/spectinomycin resistance gene; φC31 integrase gene	Dr. O. Yushchuk, unpublished
pGCymRP21	pGUS derivative carrying: φC31 *attP*; *gusA* under the control of CymR operator fused with *P21p*; *aac(3)IV* apramycin resistance gene; *aadA* streptomycin/spectinomycin resistance gene; φC31 integrase gene; *cymR* gene	Horbal et al. (2014b)

(X-Gluc), and 750 μg of nalidixic acid. Transconjugants were observed as colored colonies resistant to 50 μg/ml apramycin-sulfate after 5 days of growth.

Recombinant vector construction To construct pTEGA vector, pSAGA (Koshla et al. 2019) was digested with *Eco*RI and *Eco*RV restriction endonucleases, and the resulting 1830 bp restriction fragment containing *gusA* was cloned into respective sites of pTES (Herrmann et al. 2012). To generate pRAGA1, pSoAGA2 and pKAGA1 vectors, pSAGA was digested with *Pvu*II; the resulting 2469 bp fragment containing *aac(3)IVp-gusA* was then cloned into *Eco*RV recognition site of pRT801 (Gregory et al. 2003), pSOK804 (Sekurova et al. 2004), and pKC1139 (Kieser et al. 2000), respectively.

PCR verification of the recombinant strains Genomic DNA from the recombinant *A. rectilineatus* strains was isolated according to the Kirby mix protocol (Kieser et al. 2000). Integration of pSAGA vector was verified using PCR with attB_arec_F/R and pSET152_int_R primers (ESM Table 1). Recombinant strains, carrying pTEGA, pGT2p, pGUSmoeE5script, and pGCymRP21 were verified by amplification 1000 bp internal fragment of *gusA* using gusA_ver_F/R primer pair (ESM Table 1). For the total DNA isolation from the strain, carrying replicative pKAGA1 vector, protocol described in (Demydchuk et al. 1998) was applied. Obtained total DNA samples were used to transform competent *E. coli* DH5α cells; plasmid was re-isolated from the transformants and verified by restriction mapping. Prior to DNA isolation, *A. rectilineatus* strains were grown in 250-ml baffled Erlenmeyer flasks containing 50 ml of liquid ISP2 medium for 96 h on a rotary shaker (30 °C, 220 rpm).

Scanning electron microscopy (SEM) To prepare the samples for SEM, thin slices of agar were cut from the surfaces of the lawns at the desired time points. Then obtained samples were sprayed with thin layers of copper in VUP-5 m (USSR) vacuum evaporator and analyzed using JEOL JSM-T220a SEM (Jeol, Japan) with 25 kV electron beam.

Glucuronidase assay For qualitative glucuronidase activity estimation, desired strains were streaked on ISP3 agar plates as sectors and cultivated for 5 days at 30 °C. Then, 10 μl drops of X-Gluc (50 mg/ml DMSO solution) were added to the surfaces of the lawns. Chromogenic conversion of X-Gluc into 5,5′-dibromo-4,4′-dichloro-indigo was observed after 1 h and

24 h. Glucuronidase activity was analyzed quantitatively in a spectrophotometric assay measuring the transformation of colorless *p*-nitrophenyl-β-D-glucuronide into colored *p*-nitrophenol at 415 nm using an Infinite 200 PRO microplatereader (Tecan, Switzerland). The assay was carried out as described previously (Myronovskyi et al. 2011; Horbal et al. 2013). Glucuronidase activity was normalized to dry biomass weight. For quantitative glucuronidase assay, *A. rectilineatus* strains were grown in liquid ISP2 medium. To prepare a preculture, 500-ml baffled Erlenmeyer flasks containing 80 ml of ISP2 were inoculated with 10^7 spores of *A. rectilineatus* strains and grown for 62 h at 220 rpm and 30 °C. Then 4 ml of preculture was inoculated into a 500-ml baffled Erlenmeyer flasks containing 80 ml of ISP2. Glucuronidase activity was assayed after 62 h of growth at 200 rpm and 30 °C in cultural cell lysates. To induce *cmtP21p*-driven *gusA*-expression, cumate was added at the final concentration of 50 μM into *A. rectilineatus* pGCymRP21$^+$ cultures after 48 h of growth.

Results

Defining optimal medium for spore production in *A. rectilineatus*

High spore titers are a prerequisite for an effective conjugal transfer. Defining the optimal conditions for spore formation in *A. rectilineatus*, we studied morphology of this strain growing on different solid media. Taking into account our previous experience with *A. teichomyceticus* (Yushchuk et al. 2016b), we tested SM and ISP3 agar media. In contrast to *A. teichomyceticus* and *Actinoplanes* sp. SE50/110 (Gren et al. 2016), SM did not support sporangia and spore formation in *A. rectilineatus*. In fact, the strain exhibited very poor growth on SM and displayed bald phenotype after 12 days of growth (Fig. 1a). However, like described for *A. teichomyceticus*, sporangia and spores were formed abundantly on ISP3, although after a more extended cultivation time (10 days) compared to *A. teichomyceticus* (5 days) (Fig. 1b). Interestingly, SEM revealed a characteristic cylindrical form for the sporangia of *A. rectilineatus*, which was previously described by means of bright-field microscopy (Lechevalier and Lechevalier1975). ISP2 agar supported an abundant vegetative growth of *A. rectilineatus*. However, only rare sporangiferous hyphae were formed on the surfaces of the lawns, grown on ISP2, and sporangia were never observed (Fig. 1c). In accordance with this, similar results were previously obtained for *A. teichomyceticus* (Yushchuk et al. 2016b) and *A. ramoplaninifer* (Marcone et al. 2017), revealing that ISP2 agar supports good vegetative growth but not the formation of sporangia in *Actinoplanes* spp. We have also tested a set of other media, commonly used for growth and maintenance of Actinobacteria (SMMS, R5, TSA etc.); however, sporangia formation was never observed on any of these media and the vegetative growth was inferior to the one achieved in ISP2 (data not shown). Unlike in other *Actinoplanes* spp. (Uchida et al. 2011, Marcone et al. 2017), *A. rectilineatus* spore release from sporangia was easily achieved: swimming spores were observed after 10 min of incubation in distilled water.

Antibiotic susceptibility of *A. rectilineatus*

The resistance profile of *A. rectilineatus* to common selective antibiotics was investigated, in order to use the proper antibiotic resistance markers. Thus, 10^7 spores were plated on ISP3 and ISP2 agar containing one of the following antibiotics: apramycin-sulfate (50 μg/ml), kanamycin-sulfate (50 μg/ml), hygromycin (50 μg/ml), streptomycin (100 μg/ml), nalidixic acid (100 μg/ml), and ampicillin (150 μg/ml). At these concentrations, *A. rectilineatus* was sensitive to all of the antibiotics mentioned above except nalidixic acid, to which *A. rectilineatus* was fully resistant.

Conjugal transfer of different vectors into *A. rectilineatus*

Intergeneric conjugation is probably the most straightforward and often the only available method for transferring recombinant DNA molecules into rare Actinobacteria, including members of the genus *Actinoplanes* (Ostash et al. 2018; Ha et al. 2008; Gren et al. 2016). Nevertheless, obtaining transconjugants is often challenging since protocols have to be customized for individual species before achieving acceptable conjugation efficiencies (Marcone et al. 2010b). We adapted the protocol of the intergeneric conjugation used for the other *Actinoplanes* spp. (Ha et al. 2008; Gren et al. 2016) to *A. rectilineatus* utilizing the methylation-deficient *E. coli* ET12567 pUZ8002$^+$ as a donor strain. Conjugal transfer was tested with derivatives of different integrative vectors such as pSET152 (φC31-based integration system, Kieser et al. 2000), pRT801 (φBT1-based integration system, Gregory et al. 2003), pSOK804 (VWB-based integration system, Sekurova et al. 2004), and with a derivative of the replicative vector pKC1139 (pSG5-replicon based, Kieser et al. 2000) (Table 2).

To improve the efficiency of the selection procedure for transconjugants, in each of the above listed vectors, the *gusA* reporter gene was introduced under the control of *aac(3)IV* promoter. These resulting vectors were named pSAGA (derivative of pSET152), pRAGA1 (derivative of pRT801), pSoAGA2 (derivative of pSOK804), and pKAGA1 (derivative of pKC1139), respectively (Table 2). Like other *Actinoplanes* spp. (Horbal et al. 2013; Gren et al. 2016), *A. rectilineatus* showed no basal glucuronidase activity

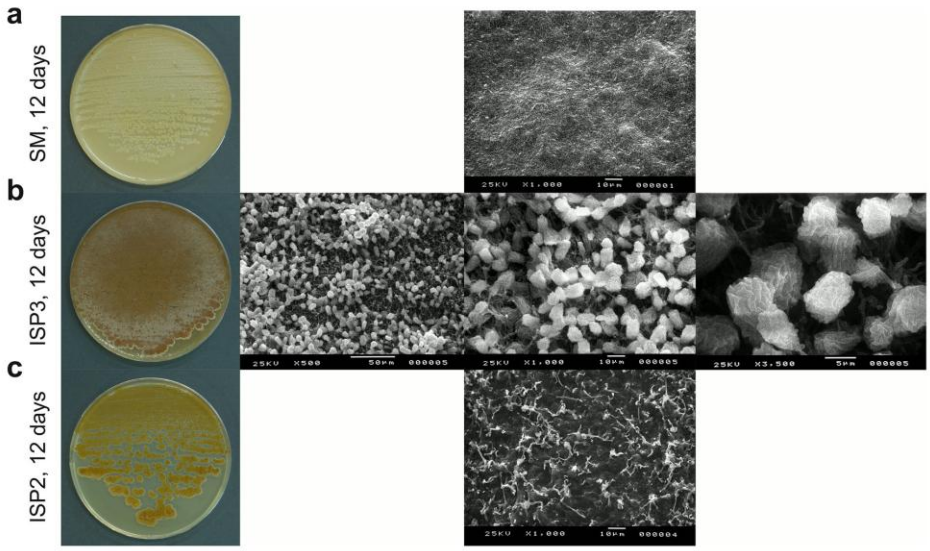


Fig. 1 Morphology of *A. rectilineatus* grown for 12 days on different solid media. **a** SM. **b** ISP3. **c** ISP2. Sporangia are exclusively formed on ISP3 (**b**). A characteristic cylindrical shape of sporangia could be noted

(Fig. 2a). Therefore, a dual selection of transconjugants was performed: for resistance to 50 μg/ml of apramycin-sulfate and for the ability to convert X-Gluc into a blue colored compound (5,5′-dibromo-4,4′-dichloro-indigo) (ESM Fig. 2). This strategy eliminated the risk of picking up false positives due to selection of spontaneous apramycin-resistant mutants. Overall, we observed no apramycin-resistant transconjugants that would display no chromogenic conversion of X-Gluc. To verify the integration of pSAGA, we have picked three random GusA-positive apramycin-resistant transconjugants and analyzed them with PCR (ESM Fig. 3): all of them appeared to carry pSAGA in an integrated form, as expected.

The protocol for the conjugal plasmid transfer from *E. coli* to *Streptomyces* spp. involves a step of spore heat shocking at 50 °C (Kieser et al. 2000). However, exposure of *A. rectilineatus* spores to elevated temperature of 50 °C abrogated the plasmid transfer.

As reported for other uncommon actinomycetes (Ha et al. 2008; Marcone et al. 2010b), the efficiency of conjugation depended on the medium and on the type of plasmid (Table 2). Overall, we succeeded in transferring only pSET152- and pKC1139-derived vectors (pSAGA and pKAGA1 respectively) into *A. rectilineatus*. The conjugation frequencies were around 10^{-3} for the integrative vector pSAGA and 10^{-5} for the replicative pKAGA1 (see Table 3). Interestingly, more transconjugants were observed for both pSAGA and pKAGA1 when plated on SM medium with high concentration of $MgCl_2$ (40 mM), although in our previous experiments it was observed, that SM medium does not support spore formation and abundant growth of *A. rectilineatus*. A similar positive impact of $MgCl_2$ on the conjugation efficiency was observed before for the other uncommon actinomycetes: *A. teichomyceticus* and *Nonomuraea gerenzanensis* (Ha et al. 2008; Marcone et al. 2010b). However, its role has not been investigated in detail and remains unclear. We did not observe any correlation between the timing of overlay and conjugation efficiency. Finally, we failed to transfer the pRT801- and pSOK804-based plasmids (pRAGA1, pSoAGA2) into *A. rectilineatus*. Further modification of the basic protocol, such as increasing the amount of spores and donor cells (up to $10^8/10^9$) and prolonging the time before overlay up to 24 h gave no positive results.

Activity of different heterologous promoters in *A. rectilineatus*

A set of φC31-based promoter-probe vectors was transferred into *A. rectilineatus,* in which *gusA* was expressed under the control of different heterologous constitutive and inducible promoters. The efficiency of conjugation was comparable to the one reported above with pSAGA, where *gusA* is placed under the control of *aac(3)IVp*. We used pTEGA (carrying *gusA* under the control of *ermEp*), pGUSmoeE5script (carrying *gusA* under the control of *moeE5p*), pGT2p (carrying *gusA* under the control of *tei2p*, which is, according to our unpublished data, very active in *A. teichomyceticus*), and pGCymRP21 (carrying *gusA* under the control of cumate-inducible *cmt-P21p*) promoter-probe vectors (Table 2, ESM Table 2). All recombinant strains were verified by means of PCR (ESM Fig.4). Glucuronidase activity of the recombinant

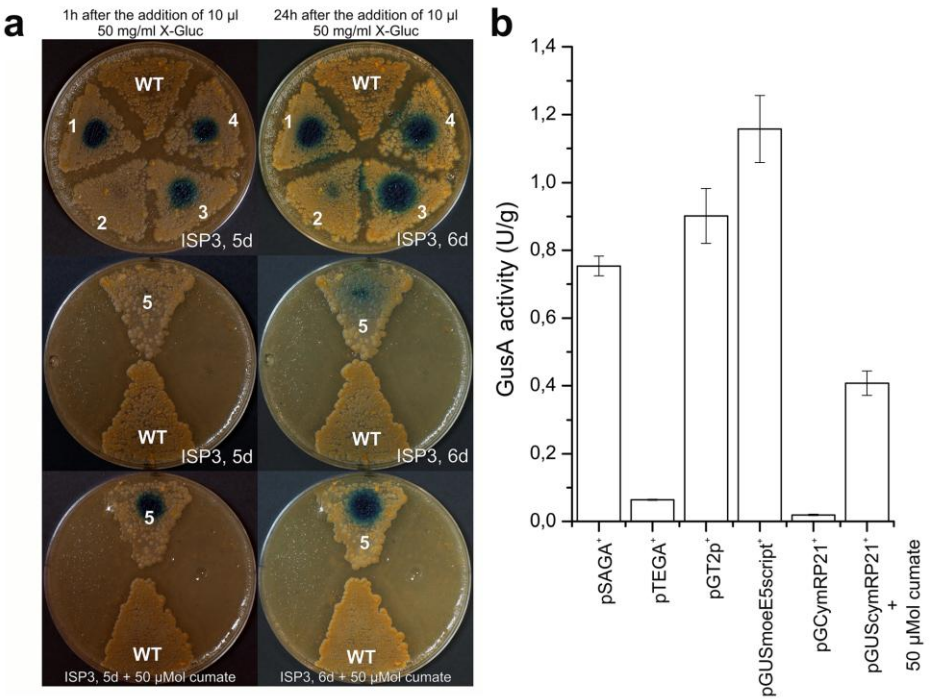


Fig. 2 Qualitative (**a**) and quantitative (**b**) analysis of glucuronidase activity in *A. rectilineatus* recombinant strains carrying vectors pSAGA (1), pTEGA (2), pGUSmoeE5script (3), pGT2p (4), pGCymRP21 (5). Recombinant strains in (**a**) were plated on ISP3 agar medium, and in (**b**) they were cultivated in ISP2 liquid medium; results of GusA-activity quantitative measurements represent the mean values of three independent biological replicates; error bars represent standard deviations

strains was tested qualitatively (on plate) and quantitatively (in cell lysates obtained from mycelium grown in liquid medium). Inspection of the X-Gluc chromogenic conversion on plates revealed that *aac(3)IVp*, *moeE5p*, and *tei2p* possessed a comparable strength in *A. rectilineatus* strains (Fig. 2a). The *ermEp* was almost nonfunctional, with a residual chromogenic conversion of X-Gluc appearing on the surface of the *A. rectilineatus* pTEGA$^+$ lawn only after 24 h of incubation. The *cmt-P21p*-controlled glucuronidase activity was strongly induced by adding 50 µM of cumate into the solid medium. However, the activity of *cmt-P21p* was still leaky, since a weak chromogenic conversion of X-Gluc was observed on the plates without cumate. The quantitative measures of the relative promoter strength (Fig. 2b) confirmed the results described above: *moeE5p* appeared to be the strongest tested promoter, followed by *tei2p* and *aac(3)IVp*. Activity of *ermEp* was hardly detectable. Finally, addition of the cumate to liquid ISP2 medium induced glucuronidase activity in *A. rectilineatus* pGCymRP21$^+$ cultures. At the same time, parallel *A. rectilineatus* pGCymRP21$^+$ cultures grown in the absence of cumate exhibited a basal glucuronidase activity.

Table 3 Frequency of transconjugants occurrence in *A. rectilineatus*

Vector (integrase/replicon)	Conjugation efficiency			
	ISP3 $MgCl_2$ 10 mM	ISP3 $MgCl_2$ 40 mM	SM $MgCl_2$ 10 mM	SM $MgCl_2$ 40 mM
pSAGA (φC31)	$(1.800 \pm 0.075) \times 10^{-3}$	$(2.800 \pm 0.215) \times 10^{-3}$	$(3.450 \pm 0.092) \times 10^{-3}$	$(6.800 \pm 0.308) \times 10^{-3}$
pSoAGA2 (VWB)	0	0	0	0
pRAGA1 (φBT1)	0	0	0	0
pKAGA1 (pSG5)	$(0.600 \pm 0.012) \times 10^{-5}$	$(1.300 \pm 0.028) \times 10^{-5}$	$(1.250 \pm 0.110) \times 10^{-5}$	$(3.100 \pm 0.086) \times 10^{-5}$

Discussion

More and more *Actinoplanes* species are becoming attractive targets for genetic engineering and genome mining due to their potential to produce valuable secondary metabolites. This paves the road to a better understanding of this peculiar genus. Using a species-tailored protocol, we demonstrated that *A. rectilineatus* could be genetically manipulated via intergeneric conjugation. The frequency of 10^{-3} transconjugants, using derivatives of the φC31-integrative plasmid pSET152, is comparable to the data reported for the same family of vectors in other uncommon actinomycetes, such as *A. teichomyceticus* (Ha et al. 2008; Horbal et al. 2012) and *N. gerenzanensis* (Marcone et al. 2010b). Due to the high titers of spores obtained in *A. rectilineatus*, we were able to transfer five φC31-based vectors into this strain. The analysis of *A. rectilineatus* draft genome sequence confirmed the presence of a conserved $attB^{\varphi C31}$ site in the *A. rectilineatus WH85_RS13100* gene (coding for a putative chromosome condensation protein), which is orthologous to the *Streptomyces coelicolor* gene *SCO3798* (Combes et al. 2002) (ESM Fig. 5a). Our PCR analysis of the *A. rectilineatus* pSAGA^{+} strains confirmed this integration site (ESM Fig. 3). Differently from what reported in *A. teichomyceticus* (Horbal et al. 2012) and *Actinoplanes* sp. SE50/110 (Gren et al. 2016), we did not succeed in the conjugal transfer of φBT1- and VWB-based integrative plasmids into *A. rectilineatus*. In *S. coelicolor*, the $attB^{\varphi BT1}$ integration site was identified in *SCO4848* encoding a putative integral membrane protein (Gregory et al. 2003). The attachment sites of the VWB-phage were identified in genes, coding for a tRNA-Arg in *Streptomyces ghanaensis* (*trnR*, Ostash et al. 2009) and in *Streptomyces venezuelae* (*SVEN_t29* (*AQF52_t0029*), van Mellaert et al. 1998), respectively. We were able to identify the orthologues of these genes in *A. rectilineatus* genome. However, the putative $attB^{\varphi BT1}$ sequence was degenerate and lacked the crossover site (ESM Fig. 5b), whereas the $attB^{VWB}$ sequence presented several mismatches when aligned with those from *S. ghanaensis* and *S. venezuelae* (ESM Fig. 5c). Probably, in the latter case, even these minor differences in $attB^{VWB}$ could prevent the integration of the VWB-based vector. A comparison to the integration sites of VWB- and φBT1-based plasmids in *A. teichomyceticus* was not possible, since these are still unknown. Concerning *Actinoplanes* sp. SE50/110, it was reported that the integration sites for φBT1- and VWB-based vectors are localized within intergenic region between ORFs coding for ACPL_7730/7731 (dihydrolipoamide dehydrogenase and γ-glutamylcyclotransferase) and ORF coding for ACPL_7821 (a hypothetical protein) respectively (Gren et al. 2016). We have found orthologues for all these genes in the genome of *A. rectilineatus* (ESM Fig. 6): *WH85_RS02595* (orthologue of *ACPL_7730*), *WH85_RS02600* (orthologue of *ACPL_7731*), and *WH85_RS03025* (orthologue of *ACPL_7821*). However, neither *Actinoplanes* sp. SE50/110 nor *A. rectilineatus* carry conventional $attB^{\varphi BT1}$ and $attB^{VWB}$ in the abovementioned loci. Perhaps, φBT1- and VWB-integrases possess dramatically different recognition patterns in the genomes of *Streptomyces* and *Actinoplanes* sp. SE50/110. Whatever real scenario is, the behavior of site-specific recombinases/integrases in uncommon actinobacteria is a topic worth more detailed research.

Finally, we compared the activity of four constitutive promoters in *A. rectilineatus,* and revealed that *aac(3)IVp* functions very well, as observed for *A. teichomyceticus* (Horbal et al. 2014a), *A. friuliensis* (Wagner et al. 2009), and *Actinoplanes* sp. SE50/110 (Schaffert et al. 2019). Both *moeE5p* and *tei2p* were stronger than *aac(3)IVp,* in accordance with the results described for *A. teichomyceticus* (Horbal et al. 2013; unpublished). The fourth constitutive promotor—*ermEp*, which is the golden standard for driving gene expression in streptomycetes—exhibited only low activity in *A. rectilineatus*, similarly to the situation observed in *A. friuliensis* (Wagner et al. 2009). In contrast, in *A. teichomyceticus ermEp* strength was comparable to *aac(3)IVp* (Horbal et al. 2013). Finally, *cmt-P21p* inducible system appeared to work as expected in *A. rectilineatus,* giving strong inducible gene expression and a low basal expression without induction. Based on these results, *aac3(IV)p* could be recommended as a promoter for the strong constitutive gene expression in *A. rectilineatus*. At the same time, *cmt-P21p* could be a preferential solution when induced gene expression is required. Overall, our results pave the way for the investigation of rich secondary metabolome and peculiar morphology of new uncommon actinomycete—*A. rectilineatus.*

Funding information This work was supported by grant Bg-46F from the Ministry of Education and Science of Ukraine (to V. F.).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

Combes P, Till R, Bee S, Smith MC (2002) The *Streptomyces* genome contains multiple pseudo-*attB* sites for the φC31-encoded site-specific recombination system. J Bacteriol 184(20):5746–5752

Demydchuk J, Oliynyk Z, Fedorenko V (1998) Analysis of a kanamycin resistance gene (*kmr*) from *Streptomyces kanamyceticus* and a mutant with increased aminoglycoside resistance. J Basic Microbiol 38(4):231–239

Gregory MA, Till R, Smith MC (2003) Integration site for *Streptomyces* phage φBT1 and development of site-specific integrating vectors. J Bacteriol 185(17):5320–5323

Gren T, Ortseifen V, Wibberg D, Schneiker-Bekel S, Bednarz H, Niehaus K, Zemke T, Persicke M, Pühler A, Kalinowski J (2016) Genetic engineering in *Actinoplanes* sp. SE50/110 – development of an intergeneric conjugation system for the introduction of actinophage-based integrative vectors. J Biotechnol 232:79–88. https://doi.org/10.1016/j.jbiotec.2016.05.012

Ha H, Hwang Y, Choi S (2008) Application of conjugation using φC31 att/int system for *Actinoplanes teichomyceticus*, a producer of teicoplanin. Biotechnol Lett 30(7):1233–1238. https://doi.org/10.1007/s10529-008-9671-z

Heinzelmann E, Berger S, Müller C, Härtner T, Poralla K, Wohlleben W, Schwartz D (2005) An acyl-CoA dehydrogenase is involved in the formation of the Delta cis3 double bond in the acyl residue of the lipopeptide antibiotic friulimicin in *Actinoplanes friuliensis*. Microbiology 151(Pt 6):1963–1974. https://doi.org/10.1099/mic.0.27844-0

Heinzelmann E, Berger S, Puk O, Reichenstein B, Wohlleben W, Schwartz D (2003) A glutamate mutase is involved in the biosynthesis of the lipopeptide antibiotic friulimicin in *Actinoplanes friuliensis*. Antimicrob Agents Chemother 47(2):447–457. https://doi.org/10.1128/aac.47.2.447-457.2003

Herrmann S, Siegl T, Luzhetska M, Petzke L, Jilg C, Welle E, Erb A, Leadlay PF, Bechthold A, Luzhetskyy A (2012) Site-specific recombination strategies for engineering actinomycete genomes. Appl Environ Microbiol 78(6):1804–1812. https://doi.org/10.1128/AEM.06054-11

Horbal L, Fedorenko V, Luzhetskyy A (2014b) Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. Appl Microbiol Biotechnol 98(20):8641–8655. https://doi.org/10.1007/s00253-014-5918-x

Horbal L, Kobylyanskyy A, Truman AW, Zaburranyi N, Ostash B, Luzhetskyy A, Marinelli F, Fedorenko V (2014a) The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. Appl Microbiol Biotechnol 98(22):9295–9309. https://doi.org/10.1007/s00253-014-5969-z

Horbal L, Kobylyanskyy A, Yushchuk O, Zaburannyi N, Luzhetskyy A, Ostash B, Marinelli F, Fedorenko V (2013) Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus* – producer of teicoplanin, drug of last defense. J Biotechnol 168(4):367–372. https://doi.org/10.1016/j.jbiotec.2013.10.018

Horbal L, Ostash B, Luzhetskyy A, Walker S, Kalinowski J, Fedorenko V (2016) A gene cluster for the biosynthesis of moenomycin family antibiotics in the genome of teicoplanin producer *Actinoplanes teichomyceticus*. Appl Microbiol Biotechnol 100(17):7629–7638. https://doi.org/10.1007/s00253-016-7685-3

Horbal L, Zaburannyy N, Ostash B, Shulga S, Fedorenko V (2012) Manipulating the regulatory genes for teicoplanin production in *Actinoplanes teichomyceticus*. World J Microbiol Biotechnol 28(5):2095–2100. https://doi.org/10.1007/s11274-012-1013-6

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical *Streptomyces* genetics. John Innes Foundation, Norwich

Koshla O, Lopatniuk M, Rokytskyy I, Yushchuk O, Dacyuk Y, Fedorenko V, Luzhetskyy A, Ostash B (2017) Properties of *Streptomyces albus* J1074 mutant deficient in tRNA (Leu)(UAA) gene *bldA*. Arch Microbiol 199(8):1175–1183. https://doi.org/10.1007/s00203-017-1389-7

Koshla O, Yushchuk O, Ostash I, Dacyuk Y, Myronovskyi M, Jäger G, Süssmuth RD, Luzhetskyy A, Byström A, Kirsebom LA, Ostash B (2019) Gene *miaA* for post-transcriptional modification of tRNA (XXA) is important for morphological and metabolic differentiation in *Streptomyces*. Mol Microbiol 112(1):249–265. https://doi.org/10.1111/mmi.14266

Lechevalier MP, Lechevalier HA (1975) Actinoplanete with cylindrical sporangia, *Actinoplanes rectilineatus* sp. nov. Int J Syst Bacteriol 25:371–376

Makitrynskyy R, Ostash B, Tsypik O, Rebets Y, Doud E, Meredith T, Luzhetskyy A, Bechthold A, Walker S, Fedorenko V (2013) Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. Open Biol 23;3(10):130121. doi: https://doi.org/10.1098/rsob.130121

Marcone GL, Foulston L, Binda E, Marinelli F, Bibb MJ, Beltrametti F (2010b) Methods for the genetic manipulation of *Nonomuraea* sp. ATCC 39727. J Ind Microbiol and Biotech 37:1097–1110

Marcone GL, Binda E, Reguzzoni M, Gastaldo L, Dalmastri C, Marinelli F (2017) Classification of *Actinoplanes* sp. ATCC 33076, an actinomycete that produces the glycolipodepsipeptide antibiotic ramoplanin, as *Actinoplanes ramoplaninifer* sp. nov. Int J Syst Evol Microbiol 2017(67):4181–4188. https://doi.org/10.1099/ijsem.0.002281

Marcone GL, Carrano L, Marinelli F, Beltrametti F (2010a) Protoplast preparation and reversion to the normal filamentous growth in antibiotic-producing uncommon actinomycetes. J Antibiot 63:83–88

Müller C, Nolden S, Gebhardt P, Heinzelmann E, Lange C, Puk O, Welzel K, Wohlleben W, Schwartz D (2007) Sequencing and analysis of the biosynthetic gene cluster of the lipopeptide antibiotic friulimicin in *Actinoplanes friuliensis*. Antimicrob Agents Chemother 51(3):1028–1037. https://doi.org/10.1128/AAC.00942-06

Myronovskyi M, Welle E, Fedorenko V, Luzhetskyy A (2011) β-Glucuronidase as a sensitive and versatile reporter in actinomycetes. Appl Environ Microbiol 77(15):5370–5383. https://doi.org/10.1128/AEM.00434-11

Ostash B, Makitrinskyy R, Walker S, Fedorenko V (2009) Identification and characterization of *Streptomyces ghanaensis* ATCC14672 integration sites for three actinophage-based plasmids. Plasmid 61(3):171–175. https://doi.org/10.1016/j.plasmid.2008.12.002

Ostash B, Walker S (2010) Moenomycin family antibiotics: chemical synthesis, biosynthesis, and biological activity. Nat Prod Rep 27(11):1594–1617. https://doi.org/10.1039/c001461n

Ostash I, Kolvenbach B, Corvini PF, Fedorenko V, Ostash B, Cichocka D (2018) Gene cloning system for sulfonamide-mineralizing *Microbacterium* sp. strain BR1. J Appl Genet 59(1):119–121. https://doi.org/10.1007/s13353-017-0427-0

Pan HX, Li JA, Shao L, Zhu CB, Chen JS, Tang GL, Chen DJ (2013) Genetic manipulation revealing an unusual N-terminal region in a stand-alone non-ribosomal peptide synthetase involved in the biosynthesis of ramoplanins. Biotechnol Lett 35(1):107–114. https://doi.org/10.1007/s10529-012-1056-7

Schaffert L, März C, Burkhardt L, Droste J, Brandt D, Busche T, Rosen W, Schneiker-Bekel S, Persicke M, Pühler A, Kalinowski J (2019) Evaluation of vector systems and promoters for overexpression of the acarbose biosynthesis gene *acbC* in *Actinoplanes* sp. SE50/110. Microb Cell Factories 18(1):114. https://doi.org/10.1186/s12934-019-1162-5

Sekurova ON, Brautaset T, Sletta H, Borgos SE, Jakobsen ØM, Ellingsen TE, Strøm AR, Valla S, Zotchev SB (2004) In vivo analysis of the regulatory genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455 reveals their differential control over antibiotic biosynthesis. J Bacteriol 186:1345–1354

Uchida K, Jang MS, Ohnishi Y, Horinouchi S, Hayakawa M, Fujita N, Aizawa S (2011) Characterization of *Actinoplanes missouriensis* spore flagella. Appl Environ Microbiol 77(7):2559–2562. https://doi.org/10.1128/AEM.02061-10

Van Mellaert L, Mei L, Lammertyn E, Schacht S, Anné J (1998) Site-specific integration of bacteriophage VWB genome into *Streptomyces venezuelae* and construction of a VWB-based integrative vector. Microbiology 144(Pt12):3351–3358

Wagner N, Osswald C, Biener R, Schwartz D (2009) Comparative analysis of transcriptional activities of heterologous promoters in the rare actinomycete *Actinoplanes friuliensis*. J Biotechnol 142(3–4):200–204. https://doi.org/10.1016/j.jbiotec.2009.05.002

Wolf T, Droste J, Gren T, Ortseifen V, Schneiker-Bekel S, Zemke T, Pühler A, Kalinowski J (2017) The MalR type regulator AcrC is a transcriptional repressor of acarbose biosynthetic genes in *Actinoplanes* sp. SE50/110. BMC Genomics 18(1):562. https://doi.org/10.1186/s12864-017-3941-x

Wolf T, Gren T, Thieme E, Wibberg D, Zemke T, Pühler A, Kalinowski J (2016) Targeted genome editing in the rare actinomycete *Actinoplanes* sp. SE50/110 by using the CRISPR/Cas9 System. J Biotechnol 231:122–128. https://doi.org/10.1016/j.jbiotec.2016.05.039

Yushchuk O, Horbal L, Datsyuk J, Stegmann E, Fedorenko V (2016b) Peculiarities of *Actinoplanes teichomyceticus* NRRL-B16726 morphology and life cycle. Visnyk of the Lviv University. Series Biology 71:1626–1136

Yushchuk O, Horbal L, Ostash B, Marinelli F, Wohlleben W, Stegmann E, Fedorenko V (2019) Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. Appl Microbiol Biotechnol 103(10):4089–4102. https://doi.org/10.1007/s00253-019-09789-w

Yushchuk O, Ostash B, Pham TH, Luzhetskyy A, Fedorenko V, Truman AW, Horbal L (2016a) Characterization of the post-assembly line tailoring processes in teicoplanin biosynthesis. ACS Chem Biol 11(8):2254–2264. https://doi.org/10.1021/acschembio.6b00018

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s13353-019-00534-7.
Електронний додаток до статті доступний за посиланням:
https://static-content.springer.com/esm/art%3A10.1007%2Fs13353-019-00534-7/MediaObjects/13353_2019_534_MOESM1_ESM.docx.

Анотація: *Нагальна потреба у відкритті нових біологічно-активних метаболітів спонукає до вивчення нових таксонів актинобактерій шляхом створення відповідних інструментів для аналізу їх геному та раціональної генної інженерії. Одним з багатообіцяючих джерел нових біоактивних природних сполук є рід Actinoplanes, домівка для актинобактерій, що здатні утворювати спорангії і продукують багато важливих спеціалізованих метаболітів, таких як тейкопланін, рамопланін і акарбоза. В цій роботі ми описуємо розробку системи експресії генів для нового виду Actinoplanes,* A. rectilineatus *(NRRL B-16090), який є потенційним продуцентом цікавих антибіотиків. Визначено оптимальні умови спороутворення в A*. rectilineatus *та процедуру перенесення плазмід за допомогою міжродової кон'югації. Вектори на основі φC31 і pSG5 були успішно перенесені в A. rectilineatus, але вектори на основі φBT1 та VWB1 не були перенесені. Нарешті, використовуючи GusA репортерну систему, ми оцінили силу кількох гетерологічних промоторів для експресії генів у A.* rectilineatus.

Applied Microbiology and Biotechnology (2022) 106:1543–1556
https://doi.org/10.1007/s00253-022-11814-4

APPLIED GENETICS AND MOLECULAR BIOTECHNOLOGY

Genetic approaches to improve clorobiocin production in *Streptomyces roseochromogenes* NRRL 3504

Sofia Melnyk[1] · Anastasia Stepanyshyn[1] · Oleksandr Yushchuk[1] · Michael Mandler[2] · Iryna Ostash[1] · Oksana Koshla[1] · Victor Fedorenko[1] · Daniel Kahne[2] · Bohdan Ostash[1]

Received: 25 December 2021 / Revised: 27 January 2022 / Accepted: 29 January 2022 / Published online: 11 February 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Streptomyces roseochromogenes NRRL 3504 is best known as a producer of clorobiocin, a DNA replication inhibitor from the aminocoumarin family of antibiotics. This natural product currently draws attention as a promising adjuvant for co-application with other antibiotics against Gram-negative multidrug-resistant pathogens. Herein, we expand the genetic toolkit for NRRL 3504 by showing that a set of integrative and replicative vectors, not tested previously for this strain, could be conjugally transferred at high frequency from *Escherichia coli* to NRRL 3504. Using this approach, we leverage a cumate-inducible expression of cluster-situated regulatory gene *novG* to increase clorobiocin titers by 30-fold (up to approximately 200 mg/L). To our best knowledge, this is the highest level of clorobiocin production reported so far. Our findings set a working ground for further improvement of clorobiocin production as well as for the application of genetic methods to illuminate the cryptic secondary metabolome of NRRL 3504.

Key Points

- *Efficient system for conjugative transfer of plasmids into NRRL 3504 was developed.*
- *Expression of regulatory genes in NRRL 3504 led to increase in clorobiocin titer.*
- *Secondary metabolome of NRRL 3504 becomes an accessible target for genetic manipulations using the expanded vector set and improved intergeneric conjugation protocol.*

Keywords *Streptomyces roseochromogenes* NRRL 3504 · Clorobiocin · Antibiotics · Cluster-situated regulatory genes · Conjugative transfer · Integrative plasmids

Introduction

Streptomyces roseochromogenes was first described in the 1970s as a producer of a unique natural product clorobiocin (Clo) (Mancy et al. 1974). Clo belongs to the aminocoumarin family of antibiotics (Fig. 1a) which, via inhibition of DNA gyrases (Flatman et al. 2006), interfere with DNA replication in many Gram-positive bacteria, including multidrug-resistant cocci (Galm et al. 2004). However, clorobiocin has not reached clinical application, unlike its structurally related counterpart novobiocin (Nov). The latter was an FDA-approved drug to treat infections caused by *Staphylococus aureus* (Walsh et al. 1993; Raad et al. 1995), until the withdrawal in 2011. Clo, however, is a more potent in vitro inhibitor of DNA gyrase than Nov (IC_{50} against *Escherichia coli* GyrB are 0.08 and 0.9 μM for Clo and Nov, respectively). Moreover, Clo and Nov (as well as their derivatives) also inhibit topoisomerase IV (Flatman et al. 2006). The clinical relevance of Nov fueled much interest in the family of aminocoumarin antibiotics. Biosynthetic gene clusters (BGCs) for Clo (*clo*) and Nov (*nov*) biosynthesis were sequenced and annotated (Fig. 1b) (Steffensky et al. 2000; Pojer et al. 2002; Li and Heide 2006). Later on, the whole genome sequences were reported for NRRL 3504 and

Sofia Melnyk and Anastasia Stepanyshyn contributed equally to this work.

✉ Bohdan Ostash
bohdan.ostash@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, Hrushevskoho st. 4, Rm. 102, Lviv 79005, Ukraine

[2] Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

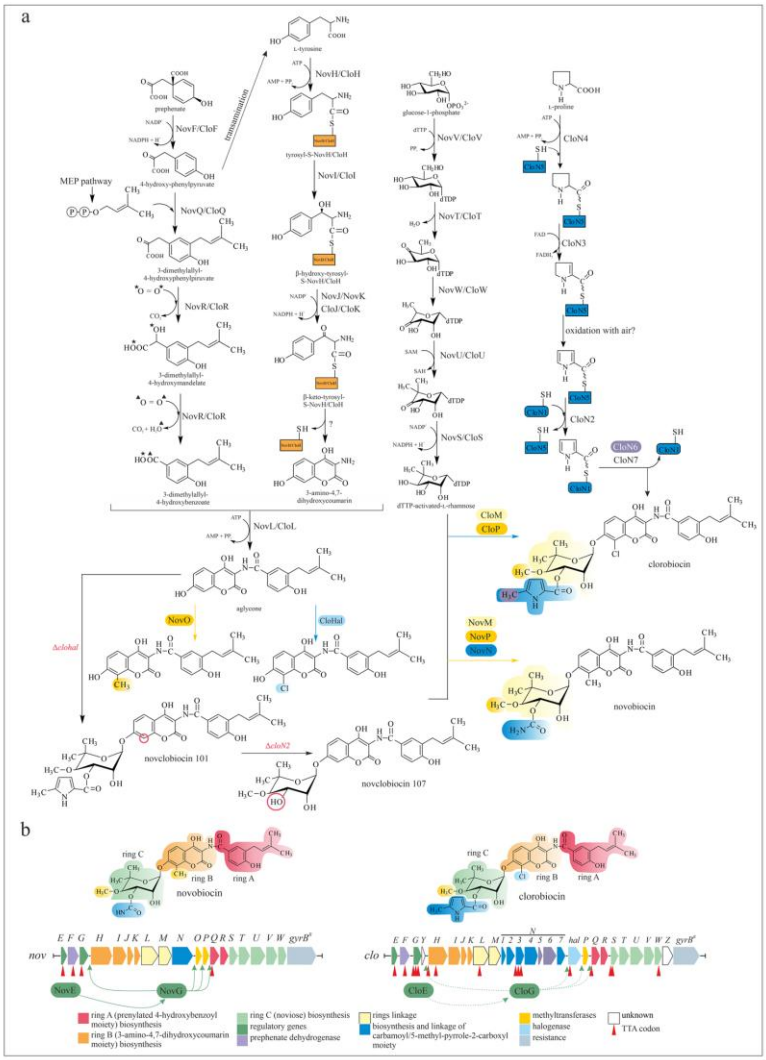


Fig. 1 Biosynthetic pathways of clorobiocin (Clo) and novobiocin (Nov). a) Biosynthetic pathways for both Nov and Clo differ by certain tailoring enzymes, encoded within *nov* and *clo* BGCs (b). Different portions of clorobiocin and novobiocin and the enzymes responsible for their assembly are color-coded appropriately. Putative regulatory schemes for *nov* and *clo* gene expression are shown; green arrows point to the promoters that are/expected to be targeted by the cluster-encoded transcriptional activators NovG and CloG

novobiocin producer *Streptomyces niveus* NCIMB 11,891 (Flinspach et al. 2014; Rückert et al. 2014).

Early identification of *clo* and *nov* BGCs allowed investigating the precise functions of almost all structural genes in a series of genetic and biochemical experiments (see extensive review in Heide 2009 and Pojer et al. 2003a, b). In fact, biosynthetic routes leading to Clo and Nov remain among the most studied secondary metabolic pathways (Fig. 1a). Some investigations were performed in the native producer (Eustáquio et al. 2003; Westrich et al. 2003; Xu et al. 2003), yet the most of our current understanding of aminocoumarin assembly comes from the studies of the heterologous hosts—*Streptomyces coelicolor* derivatives carrying *clo* BGC on a cosmid (Freitag et al. 2005a, b, 2006; Wolpert et al. 2007). The native host was manipulated via either protoplast transformation or by *E. coli—Streptomyces* conjugation to achieve both gene knockouts (Pojer et al. 2002) and knockout complementations (Eustaquio et al. 2003). Plasmids based on φC31 phage integrase or pIJ101 replicon were used for these experiments. Nevertheless, genetic manipulations of NRRL 3504 appeared to be a challenge grave enough to shift most of the work to heterologous hosts (Eustaquio et al. 2005).

More recently, both Nov and Clo were shown to enhance lipopolysaccharide transport in *E. coli* and now might be considered promising adjuvants for polymyxins and colistins to overcome multiple antibiotic resistance in Gram-negative pathogens (May et al. 2017; Mandler et al. 2018; Mattingly et al. 2020). This makes Clo an interesting target for further investigations, demanding significant amounts of the compound. Purification of Clo from fermentation broths of streptomycetes is currently the only feasible and scalable way to produce this compound. Several strains of *S. roseochromogenes* have been reported in the literature, some of which resulted from conventional selection for increased Clo accumulation, such as *S. roseochromogenes* subsp. *oscitans* DS 12.976, known also as NRRL 3504 (Mancy 1974). The latter was reported to produce 25 mg/L of Clo in an optimized fermentation medium (Eustaquio et al. 2005). Heterologous expression of *clo* BGC in the engineered *S. coelicolor* strains is another viable approach (Flinspach et al. 2010). Here, it was possible to reach up to cumulative 160 mg/L of Clo (16%) and its derivatives (Flinspach et al. 2010). While simplifying genetic studies of Clo biosynthesis, heterologous hosts do not offer ample advantages for Clo production as compared to the native strain; moreover, they may suffer from the instability of integrated cosmids. There are no works aimed at improving the titers of Clo in NRRL 3504, which might be naturally best adapted for Clo overproduction. We set out to explore the use of integrative and replicative plasmids for the expression of different regulatory genes in NRRL 3504, which would increase Clo production rates in the native host. An extended toolkit for the manipulation of NRRL 3504 would also facilitate the investigations of the other secondary metabolic pathways in this strain. Here, we show for the first time that one replicative (pSG5-based) and two integrative (actinophage-based) vectors (in addition to φC31-vectors reported previously) can be conjugally transferred and stably maintained in NRRL 3504. We also show that insertional inactivation of genes in NRRL 3504 is possible using non-replicative plasmid pKC1132. The introduction of different regulatory genes into this strain has been achieved, and some of them, particularly *novG* and *adpA*, led to quantitative and qualitative changes in aminocoumarin production. This report sets the working ground for further use of genetic tools to manipulate antibiotic production in NRRL 3504.

Materials and methods

Bacterial strains and growth media

S. roseochromogenes NRRL 3504 was used throughout the study. *Streptomyces sphaeroides* (=*niveus*) NRRL 2449 served as a source of genomic DNA to clone gene *novG*. *Micrococcus luteus* ATCC 4698 was used as a Clo-sensitive test culture in bioassays. *E. coli* DH5α was used for routine cloning procedures. *E. coli* ET12567 (pUZ8002) was used as a donor strain for intergeneric conjugations. *E. coli* strains were grown at 37 °C in liquid or agar LB medium, supplemented with 100 μg/mL of apramycin sulfate, 50 μg/mL of kanamycin sulfate, and 25 μg/mL of chloramphenicol when appropriate (Sambrook and Russell 2001). *Streptomyces* strains were routinely maintained on SFM agar (Kieser et al. 2000) at 30 °C, while ISP3 (Koshla et al. 2017) was the medium of choice to obtain sporulating lawns of *S. roseochromogenes*. To reveal endogenous antibiotic activity, *S. roseochromogenes* strains were grown in liquid and solid GYM (Koshla et al. 2017) medium for up to 120 h at 30 °C. For genomic DNA isolation, *S. roseochromogenes* and *S. sphaeroides*, strains were cultivated in TSB for 48–96 h. Recombinant *Streptomyces* strains were cultivated in presence of 50 μg/mL of apramycin sulfate, when appropriate; cumate (9 μg/mL) and thiostrepton (5 μg/mL) were used to induce *novG*/*cloG* expression from plasmids pGCymRP21 and pIJ6902, respectively. See Electronic Supplementary Materials (ESM) for compositions of all media used in this study.

Generation of recombinant plasmids

All vectors and plasmids, used in this study, are listed in Table 1. Oligonucleotide primers used for the recombinant plasmids generation are given in Table 2. Genomic DNA for cloning purposes was isolated according to salting out procedure no.4 or Kirby procedure as described in (Hopwood et al. 1985; Pospiech and Neumann 1995; Kieser et al. 2000). All recombinant vectors were verified with restriction mapping and sequencing.

Recombinant plasmids for *novG*/*cloG* expression were generated according to a single scheme. First, coding sequences (along with the putative RBS) of *novG* (1,012 bp) and *cloG* (1,022 bp) were amplified from the chromosomal DNA of *S. sphaeroides* and *S. roseochromogenes* using novG_XbaI_up/novG_EcoRI_rev and cloG_XbaI_up/cloG_EcoRI_RP primer pairs (Table 2), respectively. The amplicons were digested with *Eco*RI and *Xba*I restriction endonucleases, and cloned into the following integrative expression vectors: pmoeE5script (via *Eco*RI/*Spe*I sites), to obtain pNOVG101 and pCLOG101 plasmids (the *novG* and *cloG* would be under the control of promoter *moeEp*); pTES (via *Eco*RI/*Xba*I sites), yielding pNOVG102 and pCLOG102 (where *novG* and *cloG* would be under the control of *ermEp*); pGCymRP21 (via *Eco*RI/*Spe*I sites), to give pNOVG103 and pCLOG103 (the *novG* and *cloG* would be under the control of cumate-inducible P21 promoter); pIJ6902 (via *Eco*RI/*Xba*I sites) generating pNOVG104 and pCLOG104 plasmids (the *novG* and *cloG* would be under the control of thiostrepton-inducible promoter *tipAp*). Finally,

Table 1 Plasmids and bacterial strains and used in this work

Name	Characteristic	Reference
Plasmids		
pTES	Am^r; φC31-based *Streptomyces* integrative vector; expression of cloned gene from *ermEp*	(Herrmann et al. 2012)
pIJ6902	Am^r; φC31-based *Streptomyces* integrative vector; expression of cloned gene from thiostrepton-inducible promoter *tipAp*	(Huang et al. 2005)
pRT801	Am^r; φBT1-based *Streptomyces* integrative vector	(Gregory et al. 2003)
pKC1139	Am^r; pSG5-based *Streptomyces* replicative vector	(Bierman et al. 1992)
pOOB83e	Am^r; pKC1139E derivative carrying *Streptomyces ghanaensis* gene *moeH5* under the control of *ermEp*	(Ostash et al. 2013)
pTOSbldA	Am^r; pTOS (Herrmann et al. 2012) derivative carrying *bldA* gene from *Streptomyces albidoflavus* J1074 under the control of its native promoter	(Koshla et al. 2017)
pGCymRP21	Am^r; pGUS (Myronovskyi et al. 2011) derivative carrying *gusA* under the control of CymR operator fused with *P21p*;	(Horbal et al. 2014)
pmoeE5script	Am^r; pGUS (Myronovskyi et al. 2011) derivative carrying *gusA* under the control of *S. ghanaensis* promoter *moeE5p*	(Makitrynskyy et al. 2013)
pSAGA	Am^r; pSET152 (Kieser et al. 2000) derivative carrying *gusA* under the control of *aac(3)IVp*	(Koshla et al. 2019)
pRAGA1	Am^r; pRT801 derivative carrying *gusA* under the control of *aac(3)IVp*	(Yushchuk et al. 2020)
pKAGA1	Am^r; pKC1139 derivative carrying *gusA* under the control of *aac(3)IVp*	(Yushchuk et al. 2020)
pSoAGA2	Am^r; pSOK804 (Sekurova et al. 2004) derivative carrying *gusA* under the control of *aac(3)IVp*	(Yushchuk et al. 2020)
pKC1132	Am^r; suicide vector	(Bierman et al. 1992)
pKCBA	Am^r; pKC1139 derivative, carrying *bldA* gene from *S. albidoflavus* J1074 under the control of its native promoter	This work
pGM4181	Am^r; pmoeE5script derivative carrying *adpA* gene of *S. albidoflavus* J1074 under the control of promoter *moeE5p*	(Yushchuk et al. 2018)
pMoRT4181	Am^r; pRT801 derivative, carrying *moeE5p-adpA* fusion cloned from pGM4181	This work
pKOClo	Am^r; pKC1132 derivative, carrying 2,300 bp internal fragment of *cloH-I*	This work
pNOVG101	Am^r; pmoeE5script derivative carrying *novG* under the control of *moeE5p*	This work
pCLOG101	Am^r; pmoeE5script derivative carrying *cloG* under the control of *moeE5p*	This work
pNOVG102	Am^r; pTES derivative carrying *novG* under the control of *ermEp*	This work
pCLOG102	Am^r; pTES derivative carrying *cloG* under the control of *ermEp*	This work
pNOVG103	Am^r; pGCymRP21 derivative carrying *novG* under the control of *P21p*	This work
pCLOG103	Am^r; pGCymRP21 derivative carrying *cloG* under the control of *P21p*	This work
pNOVG104	Am^r; pIJ6902 derivative carrying *novG* under the control of *tipAp*	This work
pCLOG104	Am^r; pIJ6902 derivative carrying *cloG* under the control of *tipAp*	This work
pNOVG105	Am^r; pKC1139E derivative carrying *novG* under the control of *ermEp*	This work
pCLOG105	Am^r; pKC1139E derivative carrying *cloG* under the control of *ermEp*	This work
Bacterial strains		
E. coli DH5α	General cloning host	MBI Fermentas
E. coli ET12567 (pUZ8002)	(*dam*-13::Tn9 *dcm*-6), (pUZ8002), (Δ*oriT*), a donor strain for conjugative transfer of vector DNA	(Kieser et al. 2000)

Table 1 (continued)

Name	Characteristic	Reference
M. luteus ATCC 4698	Gram-positive bacterium, test strain for the clorobiocin antimicrobial activity assays	American Type Culture Collection
S. sphaeroides (=*niveus*) NRRL 2449	Wild type, novobiocin producer	ARS Culture Collection (NRRL)
S. roseochromogenes NRRL 3504	Wild type, clorobiocin producer	ARS Culture Collection (NRRL)
S. roseochromogenes pMoRT4181$^+$	NRRL 3504 derivative carrying pMoRT4181	This work
S. roseochromogenes pKCBA$^+$	NRRL 3504 derivative carrying pKCBA	This work
S. roseochromogenes pSAGA$^+$	NRRL 3504 derivative carrying pSAGA	This work
S. roseochromogenes pKAGA1$^+$	NRRL 3504 derivative carrying pKAGA1	This work
S. roseochromogenes pRAGA1$^+$	NRRL 3504 derivative carrying pRAGA1	This work
S. roseochromogenes pSoAGA2$^+$	NRRL 3504 derivative carrying pSoAGA2	This work
S. roseochromogenes pNOVG101$^+$	NRRL 3504 derivative carrying pNOVG101	This work
S. roseochromogenes pCLOG101$^+$	NRRL 3504 derivative carrying pCLOG101	This work
S. roseochromogenes pNOVG102$^+$	NRRL 3504 derivative carrying pNOVG102	This work
S. roseochromogenes pCLOG102$^+$	NRRL 3504 derivative carrying pCLOG102	This work
S. roseochromogenes pNOVG103$^+$	NRRL 3504 derivative carrying pNOVG103	This work
S. roseochromogenes pCLOG103$^+$	NRRL 3504 derivative carrying pCLOG103	This work
S. roseochromogenes pNOVG104$^+$	NRRL 3504 derivative carrying pNOVG104	This work
S. roseochromogenes pCLOG104$^+$	NRRL 3504 derivative carrying pCLOG104	This work
S. roseochromogenes pNOVG105$^+$	NRRL 3504 derivative carrying pNOVG105	This work
S. roseochromogenes pCLOG105$^+$	NRRL 3504 derivative carrying pCLOG105	This work
S. roseochromogenes cloH-I::pKOClo	NRRL 3504 mutant with the insertional inactivation of *cloH-I-J-K* operon	This work

Table 2 Oligonucleotide primers, used in this work

Name	Sequence (5'-3')*	Purpose
cloG_XbaI_up	AAATCTAGACCGAATCATTATGTCCTAC	*cloG* cloning into
cloG_EcoRI_rp	AAAGAATTCAGACCGTATTTCTCCGTC	the expression vectors
novG_XbaI_up	AAATCTAGACCAGACCATTAAGTCCTATG	*novG* cloning into the
novG_EcoRI_rp	AAAGAATTCAGTTGAACGTCAGGCGGTG	expression vectors
nov-clo-inKO_F	TTTTCTAGATGATCAGCGTGCTTGCCGTG	Cloning of the 2,300 bp
nov-clo-inKO_R	TTTGAATTCGTCGCCGCCGATCATCAGCC	internal fragment of *cloH-I* operon
aac(3)IV_F	ATCGACTGATGTCATCAGCG	Diagnostic primers for
aac(3)IV_R	CGAGCTGAAGAAAGACAAT	*aac(3)IV* region amplification (911 bp)
gusA_ver_F	GGCGGCTACACGCCCTTCGA	Diagnostic primers for the
gusA_ver_R	TGATGGGCCGGGTGGGGTC	amplification of 1,000 bp *gusA* internal fragment

* restriction endonuclease recognition sites are underlined

*Eco*RI/*Xba*I-treated *novG* and *cloG* amplicons were cloned into *Eco*RI/*Xba*I-digested replicative vector pOOB83e, to generate pNOVG105 and pCLOG105, where both genes would be under the control of *ermEp*.

To generate *adpA* expression plasmid pMoRT4181, pGM4181 (carrying *adpA* allele from *Streptomyces albidoflavus* (=*albus*) J1074 under the control of *moeE5p*; see Yushchuk et al. 2018) was digested with *Bam*HI/*Xho*I restriction endonucleases. 3,555-bp fragment, carrying *moeE5p-adpA* was then ligated with 3,362 bp fragment of pRT801 (Gregory et al. 2003) vector, treated with the same restriction endonucleases. To generate a replicative vector carrying an additional copy of *bldA* (leucyl(UUA)-tRNA gene from *S. albidoflavus* J1074), pTOSbldA (Table 1) was digested with *Hind*III and *Xba*I. The 777-bp fragment, carrying *bldA* and its putative promoter region, was subcloned into pKC1139 recognition sites of the same restriction endonucleases. The resulting plasmid was labeled pKCBA.

The plasmid for one-step insertional inactivation of the *cloH-I-J-K* operon was generated as follows. 2,300-bp fragment of the *cloH-I* (covering last 1,550 bp of *cloH* and first 747 bp of *cloI*) was amplified using nov-clo-inko_F/R primer

pair (Table 2). The amplicon was digested with *Xba*I/*Eco*RI restriction endonucleases and cloned into *Xba*I/*Eco*RI-digested pKC1132 suicide vector, yielding pKOClo.

Preparation of *S. roseochromogenes* spore suspensions

To obtain spore suspensions, *S. roseochromogenes* lawns were grown on ISP3 agar for up to 240 h. Each lawn was flooded with 10 mL of 15% (v/v) glycerol and the spores were scraped off the surface of the lawn with a spatula. Obtained suspensions were vigorously vortexed and filtered through sterile cotton wool to eliminate vegetative mycelia and agar debris. Filtered suspensions were examined microscopically to evaluate the homogeneity and the cotton wool filtration step was repeated in case some mycelial fragments were still present in the suspension. Homogeneous spore suspensions were then concentrated by centrifuging for 15 min at 8,000 rcf. Finally, spore titers were estimated by plating serial tenfold dilutions of the spore suspension on TSA plates, incubating for 96 h at 30 °C, and counting the number of colonies. Spore suspensions with defined spore titers were stored at -80 °C in 15% (v/v) glycerol.

E. coli—S. roseochromogenes conjugation

A standard intergeneric mating protocol (Kieser et al. 2000) was used to transfer vector DNA from *E. coli* ET12567 (pUZ8002) to germinating spores of *S. roseochromogenes*. In brief, germination of 10^8 *S. roseochromogenes* spores was induced by incubating spores at 50 °C for 10 min. Then, spores were cooled down at room temperature and mixed with 10^9 cells of an overnight culture of donor *E. coli* ET12567 (pUB8002), which carried plasmids mentioned in Table 1. Mixed cell suspensions were plated on well-dried SFM or ISP3 plates, optionally supplemented with $MgCl_2$ to 10 mM. After 14–18 h of incubation at 30 °C, plates were overlaid with 1 mL of sterile distilled water containing 1.25 mg of apramycin sulfate and 750 µg of nalidixic acid sodium salt. Emerging transconjugants, selected for apramycin resistance, were analyzed after 120 h. Conjugation efficiency was calculated as a ratio of the total transconjugants number per one plate to the number of spores applied to conjugation. Transconjugants were verified by isolating total DNA with Kirby procedure (Kieser et al. 2000) and amplifying *aac(3)IV* gene or internal fragment of *gusA* gene.

Production and extraction of Clo

To initiate the fermentation, 10^7 spores of *S. roseochromogenes* strains were inoculated into 50 mL Erlenmeyer flasks with 20 mL of TSB and 6 glass beads (ø5 mm) supplemented with 50 µg/mL apramycin sulfate when appropriate. These pre-cultures were incubated for 48 h at 30 °C on an orbital shaker (220 rpm). Clo-producing cultures were then initiated by inoculating 2.8% (v/v) of pre-culture into a 500 mL Erlenmeyer flask containing 125 mL of GYM medium and 12 glass beads (ø5 mm). Producing cultures were incubated at 30 °C on an orbital shaker (160 rpm) and sampled every 24 h for the downstream determination of biomass accumulation and Clo production. Biomass accumulation was measured as follows. 5 mL of whole cultural broth were collected and the cells were spun down (15 min at 8,000 rcf), the pellet was washed with 15% (v/v) glycerol and left to desiccate for 48 h at 60 °C; finally, the weight of the pellet was determined.

To extract Clo, the pH of the samples, containing 4 mL of whole cultural broth (containing cells and medium, harvested at regular intervals from GYM Clo-producing *S. roseochromogenes* submerged cultures) was adjusted to 4.0 with concentrated HCl and the samples were vortexed for 30 s. Then, pH was restored to 8.0 by adding NaOH (10 N) and samples were mixed with the equal volume of ethyl acetate. Mixtures were left for 90 min at the orbital shaker (220 rpm) at room temperature, then fractionated by centrifuging 10 min at 8,000 rcf. The organic (upper) phase, containing Clo, was collected. Finally, 2 mL of the organic phase were evaporated *in vacuo* and the dry pellet was re-suspended in 200 µL of methanol.

Bioassays to determine the antibiotic activity of *S. roseochromogenes* strains

Bioassays aimed to detect the antimicrobial activity of *S. roseochromogenes* culture broth samples (well diffusion assay, WDA) or extracts (Kirby–Bauer disk diffusion assay, DDA) were carried out using a similar approach and involved *M. luteus* ATCC 4698 as a Clo-sensitive test culture. Basic plates for both types of bioassays were made out of two agar layers: the bottom layer was 20 mL of TSA (2% (w/v) agar) and the upper was 10 mL of TSA (0.7% (w/v) agar) where 200 µL of fresh *M. luteus* cell suspension ($OD_{600}=0.6$) was re-suspended. Then, in the case of DDA, ø5 mm Whatman paper disks soaked with 10 µL of methanol-dissolved ethyl acetate extracts were placed on the surfaces of bioassay plates. In the case of WDA, ø7 mm agar plugs were removed out of the upper agar layer and 10 µL of whole cultural broth were poured into the well. Plates with DDA and WDA were incubated for 24 h at 37 °C, and *M. luteus* growth inhibition zones were examined and documented.

LC–MS analysis of *S. roseochromogenes* extracts

Ethyl acetate extracts were obtained as described above, dried *in vacuo,* and re-dissolved in 200 µL of methanol

before analysis. 1 µL of methanol solution was separated on C18 Phenomenex columns (100 × 2.1 mm, 1.7 µm) using Dionex Ultimate 3000 HPLC–DAD system coupled to MaXis Impact HD LC-Q-TOF mass-spectrometer (Bruker Daltonics). The LC runtime was 22 min, at a flow rate of 0.6 mL/min. Solvent system: water + 0.1% formic acid (HFo; solvent A), acetonitrile + 0.1% HFo (solvent B); from 95% A to 0% A in 12 min, then 3 min at 0% A, then reversion to 45% A within the remaining 7 min. Ionization was performed in positive and negative modes. Analysis of LC–MS data was performed using the program Compass Data Analysis 4.2 (Bruker Daltonics). The quantitative analysis of Clo production was done by comparing areas of extracted Clo mass-peak (697.2 Da, $[M+H]^+$) present in NRRL 3504 ethyl acetate extracts to the areas of 697.2-Da mass peak produced by the known amounts of authentic Clo run through the machine. A reference calibration curve was created using the values of areas for 697.2-Da mass peaks obtained when 6 different amounts of Clo were injected and run under conditions mentioned above (see Table S1, Fig. S1). See ESM for the access to raw LC–MS data.

Scanning electron microscopy (SEM)

Before SEM, NRRL 3504 lawns were grown on ISP3 agar for 216 h. Then, thin slices of the lawn surfaces were cut, sprayed with thin layers of copper *in vacuo,* and imaged with JEOL JSM-T220a SEM (Jeol, Japan) using a 25 kV electron beam.

Qualitative glucuronidase assay

A qualitative glucuronidase assay was utilized to test whether recombinant strains of *S. roseochromogenes* exhibit *gusA*-mediated conversion of 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-Gluc) into 5,5'-dibromo-4,4'-dichloro-indigo. For this, 2 µL of X-Gluc (25 mg/mL DMSO solution) was added to the surfaces of the lawns; plates were left for 2 h at 30 °C and then examined for chromogenic conversion of X-Gluc.

Results

Conditions and rates of conjugative transfer of different vectors into *S. roseochromogenes* using spores as recipients

Integrative, actinophage-based vectors are widely used in *Streptomyces* genetics because of their efficiency as a tool for stable expression of the cloned genes (Kieser et al. 2000). Currently, vectors based on the *int-attP* module of actinophages φC31, VWB, and φBT1 are the most popular ones (Bierman et al. 1992; Van Mellaert et al. 1998; Gregory et al. 2003; Ostash et al. 2009). Replicative vectors are a powerful additional tool for gene expression and among them are widely used pSG5 replicon-based moderate copy number (20–50 per genome) plasmids (Bierman et al. 1992; Muth 2018). We, therefore, decided to test the utility of the aforementioned expression systems for Clo producer. Although φC31-based plasmids are known to be transferable into NRRL 3504 (Eustáquio et al. 2003), the other vectors were not tested. NRRL 3504 genome sequence showed that the latter possesses a typical *attB* site for φC31, as well as intact *attB* sites for φBT1 and VWB (Fig. S2a, b, and c, respectively, ESM).

We choose the intergeneric conjugation with *E. coli* ET 12567 (pUZ8002) as a method to deliver recombinant plasmids into cells of NRRL 3504. This method was previously proved efficient for NRRL 3504 (Eustáquio et al. 2003); however, in that case, the recipient was applied to the mating as vegetative mycelium. Such an approach has certain drawbacks, such as low efficiency and reproducibility, as well as a high labor burden. Thus, we decided to test the efficiency of intergeneric conjugation using spores of NRRL 3504 as a recipient material.

First, we tested the growth of NRRL 3504 on a range of common agar media, aiming to detect the optimal sporulation conditions (data not shown). We found out that ISP3 supports abundant sporulation in NRRL 3504 (Fig. S3), allowing to prepare concentrated spore suspensions (see Materials and Methods section).

To test if conjugation with spores works in NRRL 3804 and how efficient it is for different plasmids, we used a previously designed approach (Yushchuk et al. 2020), involving the derivatives of pSET152 (φC31-based integrative vector), pRT801 (φBT1-based integrative vector), pSOK804 (VWB-based integrative vector), and pKC1139 (pSG5-replicon replicative vector)—pSAGA, pRAGA1, pSoAGA2, and pKAGA1, respectively. These plasmids all carry a reporter cassette *aac(3)IVp-gusA*, which allows to instantly rule out false-positive colonies among transconjugants by adding X-Gluc to the overlay. Thus, we tested the transfer efficiency of four vectors mentioned above from *E. coli* ET12567 (pUZ8002) to *S. roseochromogenes.* The pSAGA vector (φC31-based) exhibited the highest transfer efficiency, followed by pRAGA1 (φBT1-based) and pSoAGA2 (VWB-based) (Table 3). The lowest transfer efficiency was shown for a replicative vector pKAGA1 (pSG5-replicon based, Table 3). All transconjugants were able to convert X-Gluc, proving the absence of spontaneous apramycin-resistant colonies (Fig. S4a). Furthermore, we tested three random transconjugants for each platform by PCR and were successfully able to amplify a 1000-bp internal fragment of *gusA* (Fig. S4b).

We also noticed that conjugative transfer efficiency was medium-dependent: pSAGA- and pRAGA1-carrying transconjugants occurred on ISP3 agar more frequently than on SFM, while for pSoAGA2 and pKAGA SFM turned out to be more favorable (Table 3). Finally, we observed that Mg^{2+} ions have a positive impact on the conjugative transfer efficiency in *S. roseochromogenes*, although to a different extent (Table 3). These results are in agreement with many other described cases, where Mg^{2+} ions improve conjugation efficiencies of different *Streptomyces* spp.; however, the explanation of this fact remains unknown (Galm et al. 2008; Kim et al. 2008). All platforms mentioned above did not affect the growth, morphology, or antimicrobial activity of NRRL 3504 (data not shown).

Insertional inactivation of *cloH-I-J-K* operon in *S. roseochromogenes*

We focused on the combination of well and disk diffusion assays (WDA and DDA, see Materials and Methods section) as surrogates for measuring Clo production. First, WDA would be used to screen the antimicrobial properties of different recombinants at different growth time points and replicates. Then, Clo was extracted (see Materials and Methods section) and tested in DDA. Finally, Clo would be measured quantitatively in these extracts by LC–MS.

We first cultivated NRRL 3504 in GYM medium for 120 h. WDA showed that antimicrobial activity appears in the samples of NRRL 3504 cultural broth starting from 96 h, and reaches the peak at 120 h (Fig. S5a). At the same time, biomass accumulation reaches a peak at 96 h (Fig. S6a). The results of the DDA (Fig. 2a) were in agreement. Finally, we tested the extracts of NRRL 3504 cultural broth samples (collected at 120 h) by LC–MS to detect the presence of Clo (observed *m/z* 697.2165, calculated Clo $[M+H]^+$ 697.2164 Da, 0.14 ppm difference, Fig. 3).

The total antimicrobial activity revealed by the bioassays could be a net result of the production of Clo as well other uncharacterized compounds. Indeed, streptomycetes are known to produce different antibiotics with similar antimicrobial properties and *S. roseochromogenes* genome of carries numerous BGCs for specialized metabolic pathways (Rückert et al. 2014). To rule out this possibility, we decided to construct a strain with blocked Clo production. Using the established conjugation protocol, *cloH-I-J-K* operon (central to Clo production, see Fig. 1b) of NRRL 3504 was insertionaly inactivated with plasmid pKOClo. The bioassays showed the complete cessation of antimicrobial activity of *cloH-I*::pKOClo mutant against *M. luteus* (Fig. S5b, Fig. 2b), while its growth rate (measured as dry biomass) was at the slightly lower level than NRRL 3504 (Fig. S6b). Accordingly, LC–MS analysis revealed the absence of Clo mass peak in the extracts of *cloH-I*::pKOClo strain (Fig. 3). Hence, antimicrobial activity of NRRL 3504 is indeed caused by accumulation of Clo; WDA and DDA are sensitive and reliable methods for the detection of Clo production.

Properties of recombinant *S. roseochromogenes* strains carrying extra copies of *adpA* and *bldA* genes

Global positive regulator AdpA is known to improve production of different antibiotics in *Streptomyces* spp. (Ohnishi et al. 2005; Rabyk et al. 2018; Yushchuk et al. 2018). Moreover, AdpA action is tightly intertwined with the expression of *bldA*—a gene for leu(TTA)-tRNA, involved in the translation of rare TTA codon (Chater and Chandra 2008; Makitrynskyy et al. 2013; Koshla et al. 2017). Within the G-C rich genomes of *Streptomyces* spp., TTA codon often occurs in genes for morphogenesis and antibiotic biosynthesis, playing a regulatory role (McCormick and Flärdh 2012). The presence of a large number of TTA codons (sixteen, which is four times as many as in *nov* BGC, see Fig. 1b) in *clo* genes and multiple presumed operators for pleiotropic regulator AdpA within several intergenic regions of *clo* BGC (Fig. S7b) prompted us to investigate the effects of additional copies of *bldA* and *adpA* genes on Clo production. For this purpose, we have used heterologous genes from *S. albidoflavus* J1074. This was justified, because AdpA from NRRL 3504 (coded by the *M878_RS76560* locus) shares 84% of amino acid sequence identity with

Table 3 Efficiencies of conjugal transfer of different plasmids from *E. coli* ET12567 (pUZ8002) to NRRL 3504 spores. Significance of the differences in obtained data was accessed via Welch's *t* test, please see ESM Table S2

Vector	Conjugative transfer efficiency			
	ISP3	ISP3 10 mM $MgCl_2$	SFM	SFM 10 mM $MgCl_2$
pSAGA	$(2.32 \pm 0.06) \times 10^{-5}$	$(1.95 \pm 0.05) \times 10^{-5}$	$(0.83 \pm 0.12) \times 10^{-5}$	$(1.01 \pm 0.1) \times 10^{-5}$
pSoAGA2	$(0.5 \pm 0.1) \times 10^{-6}$	$(2.6 \pm 0.19) \times 10^{-6}$	$(0.63 \pm 0.13) \times 10^{-6}$	$(9.7 \pm 0.1) \times 10^{-6}$
pRAGA1	$(5.3 \pm 0.21) \times 10^{-6}$	$(5.96 \pm 0.23) \times 10^{-6}$	$(2.7 \pm 0.09) \times 10^{-6}$	$(3.73 \pm 0.11) \times 10^{-6}$
pKAGA1	$(0.145 \pm 0.032) \times 10^{-7}$	$(0.285 \pm 0.047) \times 10^{-7}$	$(0.255 \pm 0.051) \times 10^{-7}$	$(1.6 \pm 0.22) \times 10^{-7}$

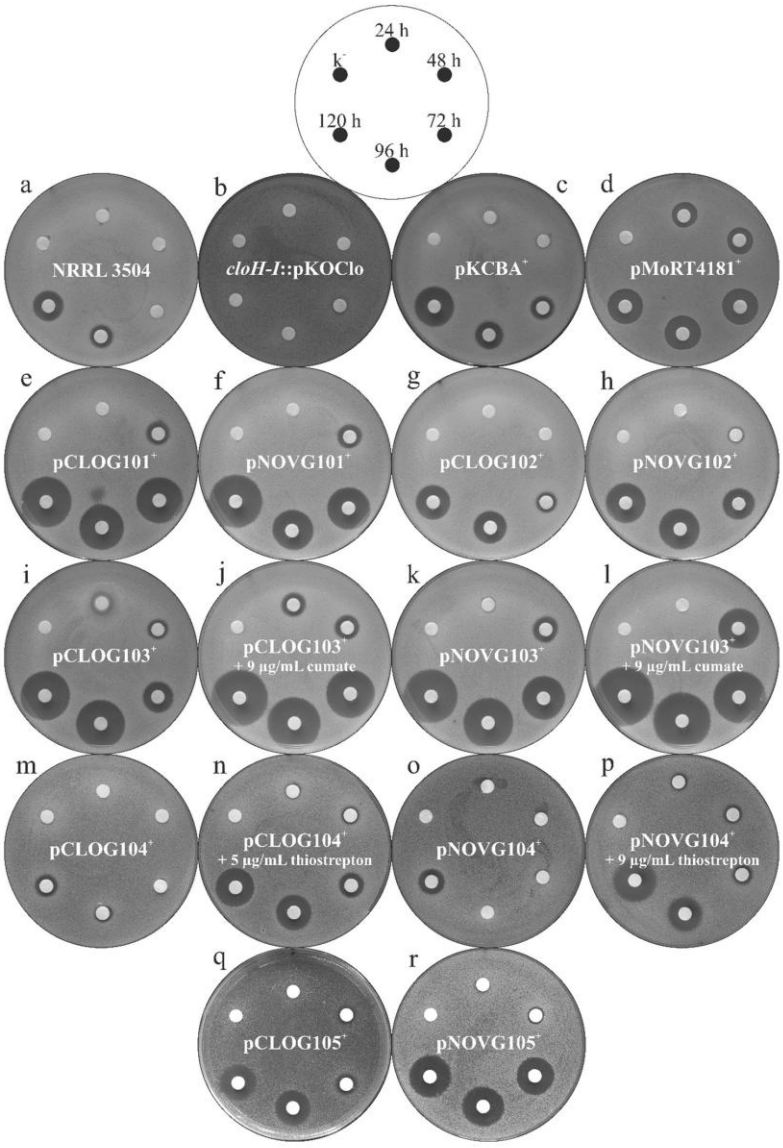


Fig. 2 Results of disk diffusion assays showing antimicrobial activity of ethyl acetate extracts obtained from the cultures of NRRL 3504 and recombinant strains. All strains were cultivated in GYM medium and samples for the extraction were collected at five time points from 24 to 120 h. k^- was always a disk soaked with 10 µL of solvent—methanol. Plates here show typical results of three replicates, for the numerical information on growth inhibition zones diameters and statistics, please see Table S4

its counterpart from J1074 (having identical DNA binding domains), while *bldA* genes of J1074 and NRRL 3504 (*M878_RS74880*) share 94% of nucleotide sequence identity. Thus, *bldA* under the control of native promotor was transferred into NRRL 3504 using pSG5-based low copy number replicative plasmid (pKCBA), while *adpA* gene under the control of *moeE5*-promotor was introduced on actinophage φBT1-based integrative plasmid pMoRT4181 (see also Table 1). Both $pKCBA^+$ and $pMoRT4181^+$ recombinants produced more Clo in comparison to the parent strain, as evident from the bioassays and MS measurements (Fig. 2c and d; Fig. S5c and d; Fig. 4). At the same time, biomass accumulation was comparable to the parent strain (Fig. S6c and d). Interestingly, in addition to Clo, $pMoRT4181^+$ accumulated two compounds whose masses fit known Clo intermediates: novclobiocins 101 (observed *m/z* 661.24, calculated $[M—H]^-$ 661.2391, 1.36 ppm difference) and 107 (observed *m/z* 554.2030,

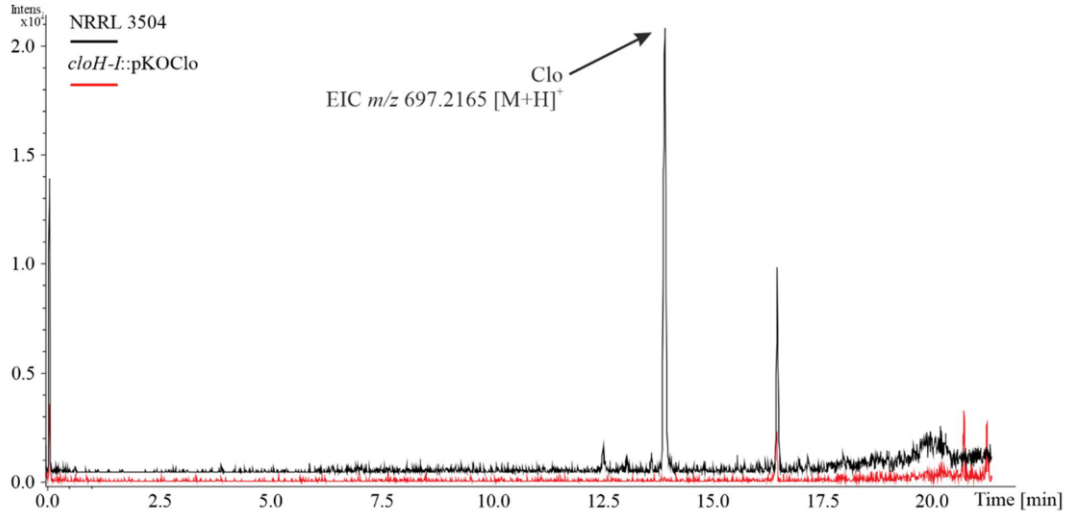


Fig. 3 Extracted ion chromatograms (EICs) of the masses corresponding to Clo (*m/z* 697.2165) in ethyl acetate extracts from cultivation broths of NRRL 3504 (black) and *cloH-I*::pKOClo mutant (red); desired mass is absent in the EIC corresponding to *cloH-I*::pKOClo extract

calculated $[M—H]^-$ 554.2021, 1.62 ppm difference; see also ESM Fig. S8; see Fig. 1a for chemical structures). We did not observe significant amounts of these intermediates in either NRRL 3504 or $pKCBA^+$ strain, or any other NRRL 3504 derivatives described below (data not shown).

Overexpression of cluster-situated regulatory genes *novG* and *cloG* using different expression platforms

The *nov* cluster-encoded transcriptional activator NovG is essential for the initiation and sustained production of Nov by *S. sphaeroides* (Eustáquio et al. 2005a, b). The same functional importance for Clo biosynthesis is suggested for CloG, the NovG ortholog (Heide 2009). Both NovG and CloG belong to the group of StrR-like transcriptional regulators often found to control the biosynthesis of antibiotics (van der Heul et al. 2018). Both proteins share 85% of amino acid sequence identity and so might be functionally interchangeable. Gene *cloG* contains three TTA codons within the reading frame, while only one TTA is present in *novG*. We speculated that the presence of three TTA codons might decrease the efficiency of *cloG* as a tool for Clo overproduction.

We cloned *novG* and *cloG* into replicative pKC1139-based vector and a series of actinophage-based vectors so that the expression of these genes would be under the control of different promoters (see Table 1). All *novG*/*cloG*-expressing plasmids were transferred into NRRL 3504 and the Clo production by the resulting recombinants and the parental strain was

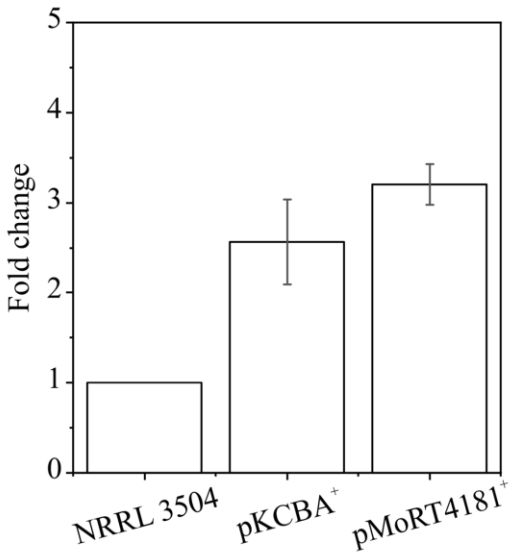


Fig. 4 Production of Clo in the recombinant strains carrying additional copies of *bldA* and *adpA* genes from *S. albidoflavus*, as judged from quantitative analysis of MS data. The diagram was derived from MS data as follows. Mass peaks of 697.2 Da, that corresponds to Clo cation ($[M+H]^+$), were extracted from the total ion chromatograms and ratios of the area of $pKCBA^+$ and $pMoRT4181^+$ strains to the wild type (NRRL 3504) were determined and normalized against the same amount of the biomass (dry weight from 5 mL). Results represent mean values (three repeats); error bars represent SD

first compared using bioassays. Almost all recombinant strains had better antimicrobial activity as compared to the parental strain (Fig. S5, Fig. 2) while showing no significant decrease in biomass accumulation (Fig. S6). Among them, pIJ6902-derived expression vectors (pCLOG104 and pNOVG104, where regulatory genes are under the control of thiostrepton-inducible promoters) exerted negative effect on antimicrobial activity of recombinants in the absence of inducer, but still increased antimicrobial activity when inducer was added (Fig. S5m-p, Fig. 2m-p). Notably, *novG*-overexpressing strains had more potent antimicrobial properties than *cloG*$^+$ strains (Fig. S5, Fig. 2). Then, we have estimated Clo production levels in cultural extracts (120 h time point) by means of LC–MS. Obtained data were again consistent with the WDA and DDA results, showing that *novG*$^+$ and *cloG*$^+$ strains specifically overproduced Clo (Fig. 5, Tables S3 and S4). The highest increase in Clo titers (up to 32-fold, Fig. 5, Fig. S9) was observed upon cumate-inducible expression of *novG* from *P21* promoter of vector pGCymRP21 (plasmid pNOVG103, see Table 1); at the same time, expression of *cloG* gave only 20-fold increase of Clo production under the same conditions (Fig. 5). Such productivity corresponded to 239 and 169 mg/L of Clo, respectively. Although overproducing Clo significantly, pNOVG103$^+$ and pCLOG103$^+$ strains accumulated biomass at the levels comparable with the parent strain (Fig. S6a and j).

Discussion

The natural producer of Clo, *S. roseochromogenes* NRRL 3504, is considered poorly amenable to genetic manipulation and genetic approaches to improve Clo titers in NRRL 3504 were not pursued. Here we report that intergeneric *E. coli—Streptomyces* conjugation, using spores of NRRL 3504 as a recipient, can be used to transfer integrative and replicative plasmids into *S. roseochromogenes* and to make gene knockouts in this strain. Overall, we have successfully transferred 12 integrative plasmids (φC31-, φBT1-, and VWB-based) and 4 replicative (pSG5-based) ones, demonstrating high conjugation efficiency. The utility of integrative vectors was demonstrated through the expression of pleiotropic and pathway-specific regulators of secondary metabolism, some of which led to significant enhancements in Clo production levels. This work therefore expands the genetic toolkit for *S. roseochromogenes* and facilitates the access to large quantities of Clo.

Additional plasmid-borne copies of *bldA* and regulatory gene *adpA* led to increased Clo titers. One of the possible explanations for the case of *bldA* might be the anomalously high frequency of rare TTA codon in *clo* genes (16 TTA codons across 28-gene BGC), which may limit the Clo production level. Likewise, the presence of putative AdpA operators within *clo* BGC motivated us to check the effects of *adpA* on NRRL 3504. Of note, *adpA* expression boosted the production of not only Clo but also novclobiocins 101 and 107. We tentatively suggest that this is caused by unbalanced activation of different *clo* operons by AdpA, although this idea requires further experimental verification. Alternatively, such effect might be due to the exhausted pool of Leu(TTA)-tRNA. Activation of some other biosynthetic pathways in NRRL 3504 by *adpA*-overexpression is also not excluded; this might explain diffused edges of *M. luteus* growth inhibition halos caused with pMoRT4181$^+$ culture broth samples (Fig. S5d). In fact, our recent experience in *Streptomyces cyanogenus* S136

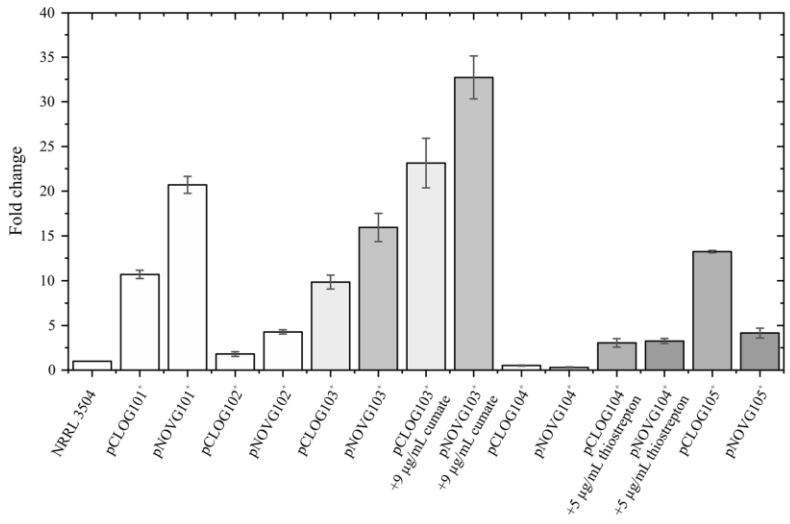


Fig. 5 Production of Clo in the recombinant strains, carrying additional copies of *novG* and *cloG* cluster-situated transcriptional regulatory genes under the control of different promoters. More than 30-fold increase in Clo production was achieved when *novG* was expressed from pGCymRP21 platform in presence of 9 μg/mL of cumate as an inducer. The diagram was built following the approach explained in the legend for Fig. 4; exact values of Clo peak areas for different measurements are given in Table S5

(Yushchuk et al. 2021), which used the same *adpA*-expression platform, makes this explanation very likely. The work to test these possibilities is underway in our laboratories. While manipulations of global regulatory genes positively influenced Clo production, introducing extra copies of either *cloG* or *novG*, which are pathway-specific regulators of biosynthesis of Clo and novobiocin, respectively, into NRRL 3504 led to even greater yields of Clo. The cumate-inducible expression of *novG*/*cloG* from a weak *P21* promoter yielded the best results. Both the bioassay and MS measurements agree with the estimate that under the described fermentation conditions the parental strain produced Clo at the level of *ca.* 7 mg/L (Table S5). Consequently, the aforementioned *novG*-expressing strain would yield around 200 mg/L of Clo, which is the highest titer for this natural product reported so far (Flinspach et al. 2010). Notably, we performed our Clo production experiments in the "generic" liquid medium—GYM—which lacks many properties of industrial production media. Nevertheless, Clo titers obtained here were already significantly higher than in *S. coelicolor* strains carrying *clo* gene cluster, which were fermented under highly optimized conditions. Therefore, we believe that recombinant strains overexpressing either *novG* or *cloG* might perform even better under more optimal production conditions.

Peculiarly, we noted that *novG* was more efficient than *cloG* in terms of increasing Clo production. Perhaps this is caused by the fact that *novG* is less saturated with TTA codons than *cloG* is, although we cannot exclude some other reasons. There is ample room for further improvements to Clo production level in NRRL 3504, either via fine-tuning the *cloG*/*novG* expression, or a combination of additional copies of pathway-specific and global regulators within one strain. The availability of three different actinophage-based integrative plasmids and a replicative one for manipulation of NRRL 3504 sets a reliable working ground for such endeavors.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00253-022-11814-4.

Acknowledgements Prof. Andriy Luzhetskyy (Saarland University) and Prof. Margaret Smith (University of York) are thanked for kind provision of some vectors used in this study.

Author contribution BO and OY conceived and designed research, with conceptual input from OK, VF, and DK. SM, AS, and IO performed research and analyzed data. MM provided mass-spectrometry assistance and critical reagents. BO, OY, AS, and DK wrote the paper.

Funding This project has received funding from the Ministry of Education and Science of Ukraine under grants BG-80F, M-26 (to B.O.), and BG-09F (to V.F.).

Data availability All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Declarations

Ethical approval This article does not contain studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare no competing interests.

References

Bierman M, Logan R, O'Brien K, Seno ET, Nagaraja Rao R, Schoner BE (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene 116:43–49. https://doi.org/10.1016/0378-1119(92)90627-2

Chater KF, Chandra G (2008) The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in *Streptomyces*. J Microbiol 46:1–11. https://doi.org/10.1007/s12275-007-0233-1

Eustáquio AS, Gust B, Luft T, Li SM, Chater KF, Heide L (2003) Clorobiocin biosynthesis in *Streptomyces*: identification of the halogenase and generation of structural analogs. Chem Biol 10:279–288. https://doi.org/10.1016/S1074-5521(03)00051-6

Eustáquio AS, Gust B, Galm U, Li SM, Chater KF, Heide L (2005a) Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. Appl Environ Microbiol 71:2452–2459. https://doi.org/10.1128/AEM.71.5.2452-2459.2005

Eustáquio AS, Li S-M, Heide L (2005b) NovG, a DNA-binding protein acting as a positive regulator of novobiocin biosynthesis. Microbiology 151:1949–1961. https://doi.org/10.1099/MIC.0.27669-0

Flatman RH, Eustaquio A, Li SM, Heide L, Maxwell A (2006) Structure-activity relationships of aminocoumarin-type gyrase and topoisomerase IV inhibitors obtained by combinatorial biosynthesis. Antimicrob Agents Chemother 50:1136–1142. https://doi.org/10.1128/AAC.50.4.1136-1142.2006

Flinspach K, Westrich L, Kaysser L, Siebenberg S, Gomez-Escribano JP, Bibb M, Gust B, Heide L (2010) Heterologous expression of the biosynthetic gene clusters of coumermycin A1, clorobiocin and caprazamycins in genetically modified *Streptomyces coelicolor* strains. Biopolymers 93:823–832. https://doi.org/10.1002/BIP.21493

Flinspach K, Rückert C, Kalinowski J, Heide L, Apel AK (2014) Draft genome sequence of *Streptomyces niveus* NCIMB 11891, producer of the aminocoumarin antibiotic novobiocin. Genome Announc 2(1):e01146-e1213. https://doi.org/10.1128/GENOMEA.01146-13

Freitag A, Rapp H, Heide L, Li SM (2005a) Metabolic engineering of aminocoumarins: Inactivation of the methyltransferase gene *cloP* and generation of new clorobiocin derivatives in a heterologous host. ChemBioChem 6:1411–1418. https://doi.org/10.1002/cbic.200500019

Freitag A, Wemakor E, Li SM, Heide L (2005b) Acyl transfer in clorobiocin biosynthesis: involvement of several proteins in the transfer of the pyrrole-2-carboxyl moiety to the deoxysugar. ChemBioChem 6:2316–2325. https://doi.org/10.1002/cbic.200500252

Freitag A, Li SM, Heide L (2006) Biosynthesis of the unusual 5,5-gem-dimethyl-deoxysugar noviose: Investigation of the C-methyltransferase gene cloU. Microbiology 152:2433–2442. https://doi.org/10.1099/mic.0.28931-0

Galm U, Heller S, Shapiro S, Page M, Li SM, Heide L (2004) Antimicrobial and DNA gyrase-inhibitory activities of novel clorobiocin derivatives produced by mutasynthesis. Antimicrob Agents Chemother 48:1307–1312. https://doi.org/10.1128/AAC.48.4.1307-1312.2004

Galm U, Wang L, Wendt-Pienkowski E, Yang R, Liu W, Tao M, Coughlin JM, Shen B (2008) *In vivo* manipulation of the bleomycin biosynthetic gene cluster in *Streptomyces verticillus* ATCC15003 revealing new insights into its biosynthetic pathway. J Biol Chem 283:25236–28245. https://doi.org/10.1074/jbc.M804971200

Gregory MA, Till R, Smith MCM (2003) Integration site for *Streptomyces* phage φBT1 and development of site-specific integrating vectors. J Bacteriol 185:5320–5323. https://doi.org/10.1128/JB.185.17.5320-5323.2003

Heide L (2009) The aminocoumarins: biosynthesis and biology. Nat Prod Rep 26:1241–1250. https://doi.org/10.1039/b808333a

Herrmann S, Siegl T, Luzhetska M, Jilg LP, Welle E, Erb A, Leadlay PF, Bechthold A, Luzhetskyy A (2012) Site-specific recombination strategies for engineering actinomycete genomes. Appl Environ Microbiol 78:1804–1812. https://doi.org/10.1128/AEM.06054-11

Horbal L, Fedorenko V, Luzhetskyy A (2014) Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. Appl Microbiol Biotechnol 98:8641–8655. https://doi.org/10.1007/s00253-014-5918-x

Huang J, Shi J, Molle V, Sohlberg B, Weaver D, Bibb MJ, Karoonuthaisiri N, Lih CJ, Kao CM, Buttner MJ, Cohen SN (2005) Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*. Mol Microbiol 58:1276–1287. https://doi.org/10.1111/j.1365-2958.2005.04879.x

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical *Streptomyces* genetics. John Innes Foundation, Norwich

Kim M-K, Ha H-S, Choi S-U (2008) Conjugal transfer using the bacteriophage φC31 *att/int* system and properties of the *attB* site in *Streptomyces ambofaciens*. Biotechnol Lett 30:695–699. https://doi.org/10.1007/s10529-007-9586-0

Koshla O, Lopatniuk M, Rokytskyy I, Yushchuk O, Dacyuk Y, Fedorenko V, Luzhetskyy A, Ostash B (2017) Properties of *Streptomyces albus* J1074 mutant deficient in $tRNALeu_{UAA}$ gene *bldA*. Arch Microbiol 199:1175–1183. https://doi.org/10.1007/s00203-017-1389-7

Koshla O, Yushchuk O, Ostash I, Dacyuk Y, Myronovskyi M, Jäger G, Süssmuth RD, Luzhetskyy A, Byström A, Kirsebom LA, Ostash B (2019) Gene *miaA* for post-transcriptional modification of $tRNA_{XXA}$ is important for morphological and metabolic differentiation in *Streptomyces*. Mol Microbiol 112:249–265. https://doi.org/10.1111/mmi.14266

Li SM, Heide L (2006) The biosynthetic gene clusters of aminocoumarin antibiotics. Planta Med 72:1093–1099. https://doi.org/10.1055/s-2006-946699

Makitrynskyy R, Ostash B, Tsypik O, Rebets Y, Doud E, Meredith T, Luzhetskyy A, Bechthold A, Walker S, Fedorenko V (2013) Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. Open Biol 3:130121. https://doi.org/10.1098/rsob.130121

Mancy D, Ninet L, Preud'Homme J (1974) Antibiotic 18631 RP. U.S. patent. 3(793):147

Mandler MD, Baidin V, Lee J, Pahil KS, Owens TW, Kahne D (2018) Novobiocin enhances polymyxin activity by stimulating lipopolysaccharide transport. J Am Chem Soc 140:6749–6753. https://doi.org/10.1021/JACS.8B02283

Mattingly AE, Cox KE, Smith R, Melander RJ, Ernst RK, Melander C (2020) Screening an established natural product library identifies secondary metabolites that potentiate conventional antibiotics. ACS Infect Dis 6:2629–2640. https://doi.org/10.1021/ACSINFECDIS.0C00259

May JM, Owens TW, Mandler MD, Simpson BW, Lazarus MB, Sherman DJ, Davis RM, Okuda S, Massefski W, Ruiz N, Kahne D (2017) The antibiotic novobiocin binds and activates the ATPase that powers lipopolysaccharide transport. J Am Chem Soc 139:17221–17224. https://doi.org/10.1021/jacs.7b07736

McCormick JR, Flärdh K (2012) Signals and regulators that govern *Streptomyces* development. FEMS Microbiol Rev 36:206–231. https://doi.org/10.1111/j.1574-6976.2011.00317.x

Muth G (2018) The pSG5-based thermosensitive vector family for genome editing and gene expression in actinomycetes. Appl Microbiol Biotechnol 102:9067–9080. https://doi.org/10.1007/s00253-018-9334-5

Myronovskyi M, Welle E, Fedorenko V, Luzhetskyy A (2011) β-Glucuronidase as a sensitive and versatile reporter in actinomycetes. Appl Environ Microbiol 77:5370–5383. https://doi.org/10.1128/AEM.00434-11

Ohnishi Y, Yamazaki H, Kato JY, Tomono A, Horinouchi S (2005) AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. Biosci Biotechnol Biochem 69:431–439. https://doi.org/10.1271/bbb.69.431

Ostash B, Makitrinskyy R, Walker S, Fedorenko V (2009) Identification and characterization of *Streptomyces ghanaensis* ATCC14672 integration sites for three actinophage-based plasmids. Plasmid 61:171–175. https://doi.org/10.1016/j.plasmid.2008.12.002

Ostash B, Campbell J, Luzhetskyy A, Walker S (2013) MoeH5: a natural glycorandomizer from the moenomycin biosynthetic pathway. Mol Microbiol 90:1324–1338. https://doi.org/10.1111/mmi.12437

Pojer F, Li S-M, Heide L (2002) Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics. Microbiology 148:3901–3911. https://doi.org/10.1099/00221287-148-12-3901

Pojer F, Kahlich R, Kammerer B, Li SM, Heide L (2003a) CloR, a bifunctional non-heme iron oxygenase involved in clorobiocin biosynthesis. J Biol Chem 278:30661–30668. https://doi.org/10.1074/JBC.M303190200

Pojer F, Wemakor E, Kammerer B, Chen H, Walsh CT, Li S-M, Heide L (2003b) CloQ, a prenyltransferase involved in clorobiocin biosynthesis. Proc Natl Acad Sci 100:2316–2321. https://doi.org/10.1073/PNAS.0337708100

Raad I, Darouiche R, Hachem R, Sacilowski M, Bodey GP (1995) Antibiotics and prevention of microbial colonization of catheters. Antimicrob Agents Chemother 39:2397. https://doi.org/10.1128/AAC.39.11.2397

Rabyk M, Yushchuk O, Rokytskyy I, Anisimova M, Ostash B (2018) Genomic insights into evolution of AdpA family master regulators of morphological differentiation and secondary metabolism in *Streptomyces*. J Mol Evol 86:204–215. https://doi.org/10.1007/s00239-018-9834-z

Rückert C, Kalinowski J, Heide L, Apel AK (2014) Draft genome sequence of *Streptomyces roseochromogenes* subsp *oscitans* DS 12.976, producer of the aminocoumarin antibiotic clorobiocin. Genome Announc 2(1):e01147-13. https://doi.org/10.1128/GENOMEA.01147-13

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, New-York

Sekurova ON, Brautaset T, Sletta H, Borgos SEF, Jakobsen ØM, Ellingsen TE, Strøm AR, Valla S, Zotchev SB (2004) *In vivo* analysis of the regulatory genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455 reveals their differential control over antibiotic biosynthesis. J Bacteriol 186:1345–1354. https://doi.org/10.1128/JB.186.5.1345-1354.2004

Steffensky M, Mühlenweg A, Wang ZX, Li SM, Heide L (2000) Identification of the novobiocin biosynthetic gene cluster of *Streptomyces spheroides* NCIB 11891. Antimicrob Agents Chemother 44:1214–1222. https://doi.org/10.1128/AAC.44.5.1214-1222.2000

van der Heul HU, Bilyk BL, McDowall KJ, Seipke RF, van Wezel GP (2018) Regulation of antibiotic production in *Actinobacteria*: new perspectives from the post-genomic era. Nat Prod Rep 35:575–604. https://doi.org/10.1039/C8NP00012C

van Mellaert L, Mei L, Lammertyn E, Schacht S, Anne J (1998) Site-specific integration of bacteriophage VWB genome into *Streptomyces venezuelae* and construction of a VWB-based integrative vector. Microbiology 144:3351–3358. https://doi.org/10.1099/00221287-144-12-3351

Walsh TJ, Standiford HC, Reboli AC, John JF, Mulligan ME, Ribner BS, Montgomerie JZ, Goetz MB, Mayhall CG, Rimland D, Stevens DA, Hansen SL, Gerard GC, Ragual RJ (1993) Randomized double-blinded trial of rifampin with either novobiocin or trimethoprim-sulfamethoxazole against methicillin-resistant *Staphylococcus aureus* colonization: prevention of antimicrobial resistance and effect of host factors on outcome. Antimicrob Agents Chemother 37:1334. https://doi.org/10.1128/AAC.37.6.1334

Westrich L, Heide L, Li SM (2003) CloN6, a novel methyltransferase catalysing the methylation of the pyrrole-2-carboxyl moiety of clorobiocin. ChemBioChem 4:768–773. https://doi.org/10.1002/cbic.200300609

Wolpert M, Gust B, Kammerer B, Heide L (2007) Effects of deletions of *mbtH*-like genes on clorobiocin biosynthesis in *Streptomyces coelicolor*. Microbiology 153:1413–1423. https://doi.org/10.1099/MIC.0.2006/002998-0

Xu H, Kahlich R, Kammerer B, Heide L, Li SM (2003) CloN2, a novel acyltransferase involved in the attachment of the pyrrole-2-carboxyl moiety to the deoxysugar of clorobiocin. Microbiology 149:2183–2191. https://doi.org/10.1099/mic.0.26314-0

Yushchuk O, Ostash I, Vlasiuk I, Gren T, Luzhetskyy A, Kalinowski J, Fedorenko V, Ostash B (2018) Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions. Appl Microbiol Biotechnol 102:8419–8428. https://doi.org/10.1007/s00253-018-9249-1

Yushchuk O, Homoniuk V, Datsiuk Y, Ostash B, Marinelli F, Fedorenko V (2020) Development of a gene expression system for the uncommon actinomycete *Actinoplanes rectilineatus* NRRL B-16090. J Appl Genet 61:141–149. https://doi.org/10.1007/s13353-019-00534-7

Yushchuk O, Ostash I, Mösker E, Vlasiuk I, Deneka M, Rückert C, Busche T, Fedorenko V, Kalinowski J, Süssmuth RD, Ostash B (2021) Eliciting the silent lucensomycin biosynthetic pathway in *Streptomyces cyanogenus* S136 via manipulation of the global regulatory gene *adpA*. Sci Rep 11:3507. https://doi.org/10.1038/s41598-021-82934-6

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s00253-022-11814-4.
Електронний додаток до статті доступний за посиланням:
https://static-content.springer.com/esm/art%3A10.1007%2Fs00253-022-11814-4/MediaObjects/253_2022_11814_MOESM1_ESM.pdf.

Анотація: *Streptomyces roseochromogenes NRRL 3504 найбільш відомий як продуцент хлоробіоцину, інгібітора реплікації ДНК із родини амінокумаринових антибіотиків. Цей натуральний продукт наразі привертає увагу як багатообіцяючий ад'ювант для спільного застосування з іншими антибіотиками проти Грам-негативних мультирезистентних патогенів. В цій роботі, ми розширюємо набір генетичних інструментів для NRRL 3504, демонструючи, що деякі інтегративні і реплікативні вектори, які раніше не використовувалися для цього штаму, можуть бути перенесені з високою частотою від Escherichia coli до NRRL 3504. Використовуючи такий підхід, ми застосували кумат-індуковану експресію шлях-специфічного регуляторного гена novG для збільшення титрів хлоробіоцину в 30 разів (приблизно до 200 мг/л). Наскільки нам відомо, це найвищий опублікований рівень продукції хлоробіоцину. Наші результати створюють робочу основу для подальшого вдосконалення продукції хлоробіоцину, а також для застосування генетичних методів з метою подальшого висвітлення загадкового вторинного метаболому NRRL 3504.*

РОЗДІЛ 5 Гени, що контролюють продукцію глікопептидних антибіотиків в бактерій роду *Nonomuraea*

Nonomuraea є родом так званих «рідкісних» актиноміцетів, чий потенціал до продукування вторинних метаболітів досі вивчений лише поверхнево [4, 105]. Нещодавно опубліковані повні геноми видів цього роду здебільшого дещо більші, ніж референсні геноми стрептоміцетів. Так, середній розмір геномів *Nonomuraea*, визначений на основі трьох доступних повних збірок [4, 106], становить близько 12 млн.п.н., тоді як середній розмір генома стрептоміцетів, обрахований на основі 251 повних збірок, доступних в GenBank, становить 8,6 млн.п.н. Великі геноми *Nonomuraea* spp. несуть десятки потенційних КБГ [4, 106, 107]. Представники роду *Nonomuraea* визнані доволі складними об'єктами для генно-інженерних маніпуляцій [43].

Найбільш важливим вторинним метаболітом, що походить з *Nonomuraea*, є ГПА типу IV A40926 [108], який продукується бактерією *N. gerenzanensis* ATCC 39727. A40926 структурно споріднений із іншим клінічно-важливим ГПА – тейкопланіном, який виділено з *A. teichomyceticus* ATCC 31121, а також з ристоцетином, виділеним з різних представників роду *Amycolatopsis* (*Am. lurida* NRRL 2430, *Am. japonicum* MG417-CF17 і *Amycolatopsis* sp. MJM2582) [86, 109, 110]. Як і тейкопланін, A40926 продукується у вигляді суміші споріднених сполук. Основними компонентами комплексу A40926 є фактори A40926 B і A40926 A, що відрізняються за структурою бічного аліфатичного ланцюга. A40926 є попередником напівсинтетичного клінічного ГПА другого покоління – далбаванцину. Далбаванцин використовують в клініці для лікування інфекцій шкіри, спричинених стійкими до антибіотиків патогенами [111]. Відповідно, вивчення продукції A40926 та створення його надпродуцентів є важливою задачею.

У цьому розділі ми поставили за мету розробити комплекс знарядь для генно-інженерних маніпуляцій із *N. gerenzanensis* ATCC 39727 і створити надпродуцентів A40926. Також, ми дослідили іншого представника роду *Nonomuraea* – *N. coxensis* DSM 45129 – на предмет продукції нового ГПА, спорідненого до A40926.

5.1 Нові молекулярні знаряддя для експресії генів і покращення продукції ГПА в представниках роду *Nonomuraea*

Порівняння активності гетерологічних і нативних промоторів в бактерій роду Nonomuraea. Дотепер, системи для надекспресії генів у *N. gerenzanensis* ATCC 39727 раціонально не вивчалися; зокрема, використовувався загальноприйнятий для надекспресії генів у стрептоміцетів промотор гена стійкості до еритроміцину – *ermEp.* Базуючись на досвіді, отриманому нами для продуцента іншого ГПА – а саме для *A. teichomyceticus* ATCC 31121 – *ermEp* не є оптимальним вибором для надекспресії генів [29] в т.зв. «рідкісних» актиноміцетів. Тому, ми перевірили активність ряду нативних і гетерологічних промоторів в *N. gerenzanensis* ATCC 39727 і *N. coxensis* DSM 45129. Для цього, ми створили та перенесли в ці два штами набір інтегративних плазмід, що несли ген β-глюкуронідази *gusA* під контролем серії нативних і гетерологічних промоторів. У цій роботі ми успішно застосували частину плазмід, які були створені в **Розділах 3, 4 і 8** цієї роботи. З огляду на те, що методи перенесення плазмід до *N. coxensis* DSM 45129 дотепер невідомі, ми розробили методику кон'югаційного перенесення плазмід до *N. coxensis* DSM 45129, використовуючи спори цього штаму як реципієнт, а клітини *E. coli* ET12567 pUZ8002^{+} як донор [112].

β-глюкуронідазну активність рекомбінантних штамів *Nonomuraea* оцінювали якісно на чашках з агаризованими середовищами і кількісно в клітинних лізатах, отриманих з міцелію, вирощеного в рідких середовищах. Обидва штами дикого типу *N. gerenzanensis* ATCC 39727 і *N. coxensis* DSM 45129 не проявляли β-глюкуронідазної активності [112]. Ми протестували такі гетерологічні промотори: *aac(3)IVp* (за допомогою плазміди pSAGA) – промотор гена апраміцинової ацетилтрансферази, отриманий з pSET152A [29]; *ermEp* (за допомогою плазміди pTEGA) – промотор гена стійкості до еритроміцину з pTES [99]; *moeE5p* (за допомогою плазміди pGUSmoeE5script) – промотор гена біосинтезу моеноміцину *S. ghanaensis moeE5* [98]; *tei2p* (за допомогою плазміди pGT2p) – промотор гена стійкості до

тейкопланіну *tei2* з *A. teichomyceticus*; *P21p*, синтетичний промотор, злитий з куматним оператором *cmt*-оперону *Pseudomonas putida* F1 (за допомогою плазміди pGCymRP21) [71]. Паралельно також перевіряли активність шести нативних промоторів, що належать генам домашнього господарства *N. gerenzanensis*: $gyrB_{ng}p$ – промотор гена β-субодиниці ДНК-гірази (за допомогою плазміди pGBP); $hrdB_{ng}p$ – промотор гена σ-фактора РНК-полімерази RpoD (за допомогою плазміди pHBP); $ssb_{ng}p$ – промотор гена білка зв'язування одноланцюгові ДНК (за допомогою плазміди pSBP); $rpsL_{ng}p$ – промотор гена S12p рибосомного білка SSU (за допомогою плазміди pRLP); $rpoBR_{ng}p$ – промотор гена β-субодиниці РНК-полімерази, стійкої до рифампіцину (за допомогою плазміди pRBP2); $rpoB_{ng}p$ – промотор гена β-субодиниці РНК-полімерази, чутливої до рифампіцину (за допомогою плазміди pRBP1). Рекомбінанти *N. gerenzanensis* pSAGA$^+$ і *N. coxensis* pSAGA$^+$, що несуть *aac(3)IVp-gusA*, демонстрували найінтенсивнішу β-глюкуронідазну активність, як за умов вирощування на чашках, так і в зануреній культурі. Натомість, у *N. gerenzanensis* pTEGA$^+$ і *N. coxensis* pTEGA$^+$, що несуть *ermEp-gusA,* β-глюкуронідазна активність спостерігалася на низькому рівні. У штамів, що несуть плазміду pGCymRP21, β-глюкуронідазна активність була індукована додаванням 50 мкМоль кумату, що підтверджує індуцибельну експресію гена *gusA*. Однак, у *N. gerenzanensis* pGCymRP21$^+$ базальний рівень GusA-активності був виявлений навіть за відсутності кумату. На відміну від цього, у *N. coxensis* pGCymRP21$^+$ не було виявлено базального рівня експресії *cmt-P21p-gusA*. У підсумку, *aac(3)IVp* поводився як сильний промотор як у *N. gerenzanensis*, так і в *N. coxensis*, а активнішим за нього був лише *cmt-P21p*, за умов індукції шляхом додаванням 50 мкМоль кумату. Це робить обидві системи експресії потенційними кандидатами для надекспресії генів в обох досліджених штамах *Nonomuraea* [112].

Використання шлях-специфічних регуляторних генів для створення надпродуцентів A40926. Наступним кроком, ми вирішили застосувати створені системи експресії генів для дизайну надпродуцентів A40926 в *N. gerenzanensis* ATCC 39727 Найефективнішим існуючим способом для

покращення продукції важливих ГПА є маніпуляції з шлях-специфічними регуляторами відповідних КБГ. Їх функції вже вивчалися на продуценті баліміцину *Am. balhimycina* [113], A40926 *N. gerenzanensis* [25] і тейкопланіну *A. teichomyceticus* [26]. Надекспресія генів шлях-специфічних регуляторів біосинтезу тейкопланіну *tei15**, що кодує StrR-подібний транскрипційний регулятор і *tei16**, що кодує LuxR-подібний транскрипційний регулятор, в *A. teichomyceticus* помітно збільшувала рівень синтезу тейкопланіну [26, 114].

КБГ для A40926, названий *dbv*, містить два гени, що кодують транскрипційні регулятори: *dbv3* і *dbv4* [25]. Ген *dbv3* кодує регулятор типу LuxR, який, однак, не є ортологічним до LuxR-регулятора з *tei* – Tei16*. Ген *dbv4* кодує StrR-подібний регулятор транскрипції, який має ортологів в кожному КБГ ГПА. Спроба створення надпродуцентів A40926 за допомогою надекспресії *dbv3* вже здійснювалася, але використання неоптимальної системи експресії і середовища вирощування не дає підстав вважати її успішною [25].

Оскільки *aac(3)IVp* було ідентифіковано як найактивніший конститутивний промотор для *Nonomuraea* spp., ми використали його для надекспресії регуляторних генів *dbv3* і *dbv4* у *N. gerenzanensis* ATCC 39727. Обидва гени були клоновані в інтегративний вектор pSET152A [29], що дало рекомбінантні плазміди pSAD3 і pSAD4, відповідно. Користуючись послідовним розташуванням генів *dbv3* і *dbv4*, ми також клонували їх разом, створивши плазміду pSAD3-4. Крім того, *dbv3*-подібний регуляторний ген *nocRI* з *N. coxensis* DSM 45129 був також клонований в pSET152A, даючи плазміду pSAR1. Всі створені плазміди було перенесено в *N. gerenzanensis* ATCC 39727. Отримані рекомбінантні штами вирощували протягом 120 год паралельно з контрольним штамом, що несе pSET152A, а продукцію A40926 вимірювали за допомогою ВЕРХ.

Продукція A40926 у *N. gerenzanensis*, що містить pSET152A, досягла рівня в майже 250 мг/л після 120 год культивування, а всі рекомбінанти мали підвищений рівень продукції A40926. Продукція A40926 у *N. gerenzanensis* pSAD3^{+} і pSAR1^{+} була на рівні прибл. 500 мг/л, pSAD3-4^{+}, продукував прибл.

550 мг/л. Найкращим продуцентом за наших умов виявився штам із надекспресією *dbv4*, який продукував більше 650 мг/л A40926 [112].

Порівняння динаміки росту і продукції A40926 штамами-надпродуцентами в колбах Ерленмеєра і 3 л ферментерах. Далі ми в деталях вивчили динаміку росту і продукції A40926 рекомбінантними штамами. Для цього, штами *N. gerenzanensis*, що несуть плазміди pSET152A, pSAD3, pSAD4 або pSAD3-4, паралельно вирощували протягом 192 год в промисловому середовищі FM2 в колбах Ерленмеєра, вивчаючи динаміку накопичення сухої біомаси, змін pH, споживання глюкози і продукції A40926 (кожні 24 год). Усі три рекомбінантні штами починали накопичувати помітну кількість A40926 раніше, ніж контрольний штам. Продукція A40926 досягала піку після 120 год росту в *N. gerenzanensis* pSAD3$^+$ (прибл. 500 мг/л) і pSAD4$^+$ (майже 650 мг/л), тоді як у pSAD3-4$^+$ максимальна продуктивність (прибл. 600 мг/л) досягалася на 144 год культивування. Контрольний штам продукував близько 300 мг/л через 144-168 год після інокуляції. Глюкоза споживалася швидше рекомбінантними штамами, що містять надекспресовані регуляторні гени, у порівнянні з контрольним; ці ж штами накопичували менше біомаси, ніж *N. gerenzanensis* pSET152A$^+$.

Також ми вивчили, як рекомбінантні штами, що надпродукують A40926, поводитимуть себе в ферментерах на 3 л, що містять 2 л робочого об’єму FM2. Виявилося, що усі штами продукували значно більше A40926 і росли краще, ніж у в колбі, за винятком *N. gerenzanensis* pSAD3$^+$. Контрольний штам досягав піку продукції A40926 – майже 400 мг/л – через 120 год інкубування. Рекомбінантний штам pSAD4$^+$ досягав піку продукції антибіотика – 700 мг/л – через 168 год. Найкращі показники щодо продукції A40926 у ферментері продемонстрував штам pSAD3-4$^+$, який продукував максимальну кількість антибіотика – 800 мг/л – після 168 год культивування. Штам pSAD3$^+$ слабо накопичував біомасу, а продукція A40926 починалася дуже пізно (після 168 год культивування) і ніколи не перевищувала 100 мг/л

У підсумку, результати, отримані в цьому розділі роботи, показують створення і ефективне застосування нових систем експресії генів в бактерій

роду *Nonomuraea*. Окрім деяких повідомлень про генно-інженерні маніпуляції з продуцентом A40926 *N. gerenzanensis* [25, 43] і продуцентом кистаміцину *Nonomuraea* sp. ATCC 55076 [115], будь-які інші спроби розробити ефективні генетичні знаряддя для маніпулювання штамами *Nonomuraea* були відсутні. Тому, ми випробували набори гетерологічних і нативних промоторів (останні походять з генів домашнього господарства в *N. gerenzanensis*) як на *N. gerenzanensis*, так і на *N. coxensis*, ставлячи за мету відібрати найефективніші для надекспресії генів. Як і у випадку *Actinoplanes* spp. [29], найсильнішим гетерологічним промотором виявився *aac(3)IVp*, хоча деякі нативні промотори *N. gerenzanensis*, такі як $hrdB_{ng}p$ і $rpsL_{ng}p$, мали еквівалентну ефективність і заслуговують на подальше застосування. Використавши *aac(3)IVp* для надекспресії регуляторних генів *dbv3* і *dbv4* з КБГ *dbv* і *nocRI* (ортолога *dbv3*) з КБГ *noc* в *N. gerenzanensis,* нам вдалося створити надпродуцентів A40926. Ми також ефективно застосували надекспресію генів *dbv3* і *dbv4* для підвищення продукції нового ГПА A50926 в *N. coxensis* DSM 45129.

У попередній роботі [25] повідомлялося, що надекспресія *dbv3* (під контролем тіострептон-індукованого промотору *tipA** в інтегративній плазміді pIJ8600) у *N. gerenzanensis* збільшила продукцію A40926 з 13 до 27 мг/л в лабораторному середовищі R3. У нашій роботі ми перевірили реальний промисловий потенціал надекспресії не лише *dbv3*, але й *dbv4*, а також *dbv3* і *dbv4* разом, клонуючи їх під сильним конститутивним промотором *aac(3)IVp*, та вивчаючи продуктивність відповідних рекомбінантних штамів у ферментері з використанням промислового середовища FM2 [45]. У ферментері найкращий рівень продукції A40926 продемонстрував штам, що надекспресує гени *dbv3* і *dbv4* одночасно. Цей штам продукував майже 800 мг/л A40926, що вдвічі більше, ніж у вихідного штаму, вирощеного в тих же умовах.

5.2 A50926 – новий ГПА з *N. coxensis* DSM 45129

Характеристика нового КБГ ГПА в N. coxensis DSM 45129 і відповідного біосинтетичного шляху. Як вже згадувалося вище, ми знайшли потенційний

КБГ ГПА в *N. coxensis* DSM 45129. Однак, проект генома *N. coxensis,* доступний в базі даних GenBank, містив послідовність низької якості. Як наслідок, потенційний КБГ ГПА був фрагментований на декількох контигах, а деякі важливі частини КБГ і взагалі були відсутні. Тому, ми секвенували і повністю зібрали геном *N. coxensis* DSM 45129 використавши комбінацію HiSeq Illumina і GridION ONT. Аналіз генома за допомогою antiSMASH 5.0 [116] виявив 26 потенційних КБГ. При цьому лише деякі з них були більше, ніж на 20%, подібні до вже відомих КБГ, включно з КБГ терпеноїдів, десферіоксаміну, полікетидів, мірубактину, акарвіостатину і ГПА.

КБГ ГПА ми назвали *noc* (від ***No****nomuraea* ***c****oxensis*) і перевірили анотацію, отриману за допомогою antiSMASH. До складу *noc* входить 36 відкритих рамок зчитування, з яких 35 гомологічні до генів *dbv* (негомологічний ген кодував ймовірну транспозазу). Враховуючи відомості про шлях біосинтезу A40926 [117], ми передбачили біосинтетичний шлях потенційного ГПА з *N. coxensis* DSM 45129. Набори генів, необхідних для біосинтезу непротеїногенних амінокислот-попередників 4-гідроксифенілгліцину (Hpg), 3,5-дигідроксифенілгліцину (Dpg) і β-гідрокситирозину, виявилися однаковими в КБГ *noc* і *dbv*. НРПС, закодована у *noc,* мала таку саму організацію і специфічність аденілюючих (А)-доменів, що й НРПС *dbv*. Усі інші гени, відповідальні за етапи формування поперечних зшивок та модифікації аглікона, виявилися гомологічними у *noc* і *dbv*. Помітна відмінність полягала в тому, що *noc* не мав ортолога *dbv29*. Dbv29 є гексозооксидазою, відповідальною за окислення Glc*N*-ацильної частини A40926 [118]. Відповідно, ми передбачили, що шлях *noc* кодує біосинтез аналога A40926, який не має карбоксильної групи на Glc*N*-ацильному залишку, нагадуючи в цьому тейкопланін.

Окрім біосинтетичних генів, *noc* і *dbv* містять аналогічний набір регуляторних генів. Два головних регулятори біосинтезу A40926 – LuxR-подібний Dbv3 і StrR-подібний Dbv4 – мають ортологів кодованих в *noc* – NocRI (94.4% ідентичності амінокислотної послідовності) і NocRII (86.3% ідентичності амінокислотної послідовності), відповідно. Також, *noc* кодує єдиний білок стійкості до ГПА – це Noc27, близький (87,3% ідентичності

амінокислотної послідовності) ортолог Dbv7 (VanYn), який є D,D-карбоксипептидазою, що формує стійкість до А40926 у *N. gerenzanensis* ATCC 39727 [119, 120].

Хоча біосинтетичні, регуляторні гени і гени стійкості майже однакові в *dbv* і *noc*, організація обох КБГ значно відрізняється. Поки що майже всі КБГ ГПА мають гени НРПС, розташовані на одному ланцюзі в порядку, який є колінеарним до порядку модулів НРПС-конвеєра. Єдиним винятком є *dbv*, де гени НРПС розміщені на різних ланцюгах і розділені іншими біосинтетичними генами [121]. КБГ *noc*, хоча й схожий з *dbv*, має організацію генів НРПС типову для всіх інших ГПА [122].

Ідентифікація нового ГПА з N. coxensis DSM 45129 – А50926. *N. coxensis* DSM 45129 було вперше описано в 2007 році [123], але здатність цього штаму до продукції антибіотиків ніколи не перевірялася. Ми вивчили ріст *N. coxensis* DSM 45129 і продукцію антибіотика на середовищах які були оптимальними для інших продуцентів ГПА, а саме FM2, TM1 (оптимальне для продукції тейкопланіну в *A. teichomyceticus*) [47], R5 (оптимальне для продукції баліміцину в *Am. balhimycina*) [124], середовищ VM0.1 і ISP2l. Антимікробна активність спостерігалася у культур, вирощених в середовищах ISP2l і TM1. Ми застосували D-Ala-D-Ala-афінну хроматографію для того, щоб виділити потенційний комплекс ГПА з продукційних культур *N. coxensis* DSM 45129, вирощених в середовищах ISP2lm і TM1m. Аналіз отриманих зразків за допомогою ВЕРХ-МС показав присутність сполук, що утворюють іони з *m/z* 852,3 і 859,3 ($[M+2H]^{2+}$), на 28 і 14 Да, відповідно, меншим ніж стандарт А40926 ($[M+2H]^{2+}$=866,3, що відповідає А40926 В). Отриманий комплекс було названо А50926. Подальший аналіз за допомогою МС/МС дав змогу локалізувати цю відмінність від А40926 у залишку глюкозаміну. Тобто, в А50926 залишок Glc*N*-ацилу містить 6-гідроксильну групу. Це корелює з відсутністю гомолога *dbv29* в *noc*, бо цей ген кодує фермент відповідальний за окислення С-6 гідроксильної групи Glc*N*-ацильного залишку в карбоксильну групу в біосинтезі А40926 [122].

Створення штамів-надпродуцентів A50926 за допомогою надекспресії гетерологічних шлях-специфічних регуляторів. Виходячи з того, що надекспресія генів *dbv4* і *dbv3* мала позитивний вплив на продукцію A40926, ми вирішили перевірити, чи можна застосувати ці гени для покращення продукції A50926 у *N. coxensis* DSM 45129. Ми перенесли плазміди pSAD4 і pSAD3 в клітини *N. coxensis* DSM 45129 та вивчили продукцію A50926 в отриманих рекомбінантних штамах. Виявилося, що в продукційних середовищах ISP2lm, і TM1m, штами *N. coxensis* pSAD3$^+$ і pSAD4$^+$ продукували більше A50926, ніж дикий тип. У ISP2lm *N. coxensis* pSAD3$^+$ досягав піку продукції антибіотика на 192 год, продукуючи прибл. 45 мг/л, що перевищує продуктивність як дикого типу (прибл. 20 мг/л), так і *N. coxensis* pSAD4$^+$ (прибл. 30 мг/л). У середовищі TM1m продукція антибіотика в штамі *N. coxensis* pSAD3$^+$ була трохи вищою, досягаючи прибл. 50 мг/л на 192 год культивування. У той самий момент часу в TM1m дикий тип і *N. coxensis* pSAD4$^+$ продукували прибл. 16 і 22 мг/л антибіотика, відповідно [122].

Створення альтернативного продуцента A40926 шляхом надекспресії гетерологічного гена dbv29 в N. coxensis DSM 45129. Враховуючи, що *N. gerenzanensis* є єдиним відомим продуцентом A40926 – попередника клінічно-релевантного ГПА далбаванцину – а також те, що ми створили ефективні системи експресії генів в *N. coxensis* DSM 45129, ми вирішили перевірити, чи можна створити нового продуцента A40926 на основі *N. coxensis* DSM 45129. Ми припустили, що можна перетворити *N. coxensis* DSM 45129 на продуцента A40926 шляхом надекспресії гена *dbv29*, який кодує гексозооксидазу, необхідну для окислення C-6 гідроксильної групи Glc*N*-ацилу у відповідну карбонову кислоту, з *N. gerenzanensis*. Для цього ми клонували ген *dbv29* в pSET152A, даючи плазміду pSAD29, яка потім була перенесена в *N. coxensis* DSM 45129. *N. coxensis* pSAD29$^+$ вирощували в середовищі ISP2lm протягом 168 год, а отриманий комплекс ГПА очищали за допомогою афінної хроматографії з D-Ala-D-Ala.

Аналіз за допомогою ВЕРХ-МС показав, що створений штам *N. coxensis* pSAD29$^+$ здатний продукувати молекулу з ідентичним часом виходу та МС-

спектром, як і у A40926 (спостережене значення *m/z* 1731,5181, розраховане для A40926 $[M+H]^+$ *m/z* 1731,5107). MC/MC продукту *N. coxensis* $pSAD29^+$ продемонструвала, що він має фрагментаційний спектр ідентичний до стандарту A40926, включаючи фрагмент з *m/z* 358,22, що відповідає карбоксильованому залишку Glc*N*-ацилу [122]. Рівні продукції A40926 у *N. coxensis* $pSAD29^+$ були порівнянними з тими, які спостерігалися для A50926 в *N. coxensis* DSM 45129. Отже, хоча в наших умовах *N. coxensis* $pSAD29^+$ і продукував менше A40926, ніж *N. gerenzanensis* дикого типу, це перший успішний крок до створення альтернативних продуцентів A40926 шляхом генної інженерії.

У підсумку, ми ідентифікували продукцію нового ГПА, який було названо A50926, в *N. coxensis* DSM 45129. Секвенувавши геном *N. coxensis* DSM 45129 нам вдалося анотувати відповідний КБГ – *noc* – і детально описати відповідний біосинтетичний шлях у порівнянні із біосинтетичним шляхом A40926. Тоді ми знайшли оптимальні умови для продукції ГПА в *N. coxensis* DSM 45129, та виділили цей антибіотик. Детальний МС-аналіз продемонстрував, що A50926 відрізняється від A40926 відсутністю карбоксильної групи в Glc*N*-ацильному фрагменті, приєднаному до AA4 аглікону. Різниця в хімічній структурі між A50926 і A40926 виявилася пов'язаною з відсутністю ортолога *dbv29* в КБГ A50926 – *noc*. Надекспресія *dbv29* в *N. coxensis* DSM 45129 привела до продукції A40926.

Результати, отримані в **Розділі 5**, оновлюють уявлення про представників роду *Nonomuraea* як про продуцентів ГПА і створюють передумови для подальшого вивчення бактерій цього роду з метою пошуку нових ГПА, а також використання вже відомих продуцентів ГПА для комбінаторного біосинтезу нових похідних.

Основні результати цього розділу опубліковано в [112, 122].

frontiers in Microbiology

ORIGINAL RESEARCH
published: 21 January 2020
doi: 10.3389/fmicb.2020.00008

New Molecular Tools for Regulation and Improvement of A40926 Glycopeptide Antibiotic Production in *Nonomuraea gerenzanensis* ATCC 39727

Oleksandr Yushchuk[1], Andres Andreo-Vidal[1], Giorgia Letizia Marcone[1], Mervyn Bibb[2], Flavia Marinelli[1] and Elisa Binda[1]*

[1] Department of Biotechnology and Life Sciences, University of Insubria, Varese, Italy, [2] Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich, United Kingdom

OPEN ACCESS

Edited by:
Yvonne Mast,
German Collection of Microorganisms and Cell Cultures GmbH (DSMZ), Germany

Reviewed by:
Max Julian Cryle,
Monash University, Australia
Marta Vaz Mendes,
University of Porto, Portugal

***Correspondence:**
Flavia Marinelli
flavia.marinelli@uninsubria.it

Specialty section:
This article was submitted to Antimicrobials, Resistance and Chemotherapy, a section of the journal Frontiers in Microbiology

Received: 11 November 2019
Accepted: 06 January 2020
Published: 21 January 2020

Citation:
Yushchuk O, Andreo-Vidal A, Marcone GL, Bibb M, Marinelli F and Binda E (2020) New Molecular Tools for Regulation and Improvement of A40926 Glycopeptide Antibiotic Production in Nonomuraea gerenzanensis ATCC 39727. Front. Microbiol. 11:8. doi: 10.3389/fmicb.2020.00008

Genome sequencing has revealed that *Nonomuraea* spp. represent a still largely unexplored source of specialized metabolites. *Nonomuraea gerenzanensis* ATCC 39727 is the most studied representative species since it produces the glycopeptide antibiotic (GPA) A40926 – the precursor of the clinically relevant antibiotic dalbavancin, approved by the FDA in 2014 for the treatment of acute skin infections caused by multi-drug resistant Gram-positive pathogens. The clinical relevance of dalbavancin has prompted increased attention on A40926 biosynthesis and its regulation. In this paper, we investigated how to enhance the genetic toolkit for members of the *Nonomuraea* genus, which have proved quite recalcitrant to genetic manipulation. By constructing promoter-probe vectors, we tested the activity of 11 promoters (heterologous and native) using the GusA reporter system in *N. gerenzanensis* and in *Nonomuraea coxensis*; this latter species is phylogenetically distant from *N. gerenzanesis* and also possesses the genetic potential to produce A40926 or a very similar GPA. Finally, the strongest constitutive promoter analyzed in this study, *aac(3) IVp*, was used to overexpress the cluster-situated regulatory genes controlling A40926 biosynthesis (*dbv3* and *dbv4* from *N. gerenzanensis* and *nocRI* from *N. coxensis*) in *N. gerenzanensis*, and the growth and productivity of the best performing strains were assessed at bioreactor scale using an industrial production medium. Overexpression of positive pathway-specific regulatory genes resulted in a significant increase in the level of A40926 production in *N. gerenzanensis*, providing a new knowledge-based approach to strain improvement for this valuable glycopeptide antibiotic.

Keywords: A40926, *Nonomuraea*, glycopeptide antibiotics, pathway-specific regulators, strain improvement

INTRODUCTION

Research on glycopeptide antibiotics (GPAs) – drugs of "last resort" for treating severe infections caused by multi-drug resistant Gram-positive pathogens – has experienced a "renaissance" over the last decade (Marcone et al., 2018). Clinically important GPAs include two natural products (vancomycin and teicoplanin) and three second generation antibiotics (telavancin,

dalbavancin, and oritavancin), which are semisynthetic derivatives of natural products endowed with an increased antimicrobial potency and superior pharmacokinetic properties. The urgent need for new potent antibiotics has driven much recent interest in GPAs. For example, chemical synthesis recently resulted in the generation of a plethora of vancomycin derivatives with novel modifications that show superior antimicrobial activities (Okano et al., 2017; Wu and Boger, 2019), while teicoplanin has been conjugated with nanoparticles resulting in increased activity against biofilm-forming pathogens (Armenia et al., 2018).

In contrast, genetic manipulation of GPA producers to yield novel potent derivatives is in its infancy. Recent work (Haslinger et al., 2015; Peschke et al., 2017; Schoppet et al., 2019) has revealed new details of the specificity and timing of non-ribosomal peptide synthesis, including chlorination and cross-linking steps, suggesting that the use of combinatorial biosynthesis to generate GPAs with completely novel oligopeptide scaffolds should be possible. In parallel, heterologous expression of enzymes involved in later stages of GPA biosynthesis (glycosylation, sulfation, acylation etc.) in known producers or *in vitro* has already generated novel GPA derivatives that could not be prepared easily by chemical synthesis (Banik and Brady, 2008; Banik et al., 2010; Yim et al., 2014).

In the meantime, genome sequencing has revealed the organization of GPA biosynthetic gene clusters (BGCs) in industrially valuable actinobacteria (D'Argenio et al., 2016), including long known GPA-producers (Vongsangnak et al., 2012; Truman et al., 2014; Kusserow and Gulder, 2017; Nazari et al., 2017; Adamek et al., 2018) as well as in novel producing strains (Thaker et al., 2013; Stegmann et al., 2014). Although the global regulation of GPA biosynthesis is still largely unexplored, the pathway-specific regulation controlling the expression of BGCs is being elucidated in model systems (Bibb, 2013). The roles of cluster-situated regulatory genes have been investigated in *Amycolatopsis balhimycina* (Shawky et al., 2007), *Nonomuraea gerenzanensis* (Lo Grasso et al., 2015; Alduina et al., 2018), and *Actinoplanes teichomyceticus* (Horbal et al., 2014b; Yushchuk et al., 2019), producing balhimycin, A40926 (the natural precursor of dalbavancin), and teicoplanin, respectively. Overexpression of the teicoplanin cluster-situated regulatory genes (*tei15**, coding for a StrR-like transcriptional regulator, and *tei16**, coding for a LuxR-type regulator) in *A. teichomyceticus* markedly increased teicoplanin production in the wild type strain, representing one of the most successful examples of using molecular tools for improving antibiotic production (Horbal et al., 2012, 2014b). In this work, we investigated the potential of molecular tools to improve the production of A40926, the dalbavancin precursor.

The BGC for A40926, named *dbv*, contains two genes encoding transcriptional regulators: *dbv3* and *dbv4* (Lo Grasso et al., 2015). *dbv3* encodes a LuxR-type regulator, which, however, is non-orthologous to the *tei* cluster encoded LuxR-regulator – Tei16* (Yushchuk et al., 2019). *dbv4* codes for a StrR-like transcriptional regulator with close homologues in every GPA BGC (Yushchuk et al., 2019). The A40926-producing strain, recently re-classified as *N. gerenzanensis* (Dalmastri et al., 2016), belongs to a still poorly investigated genus of actinobacteria that was only recently identified as an untapped source of novel antibiotics and other bioactive metabolites (Sungthong and Nakaew, 2015). More recently, fully sequenced genomes of *N. gerenzanensis* (D'Argenio et al., 2016) and of the kistamicin producer *Nonomuraea* sp. ATCC 55076, previously classified as *Actinomadura parvosata* subsp. *kistnae* S382–8 (Kusserow and Gulder, 2017; Nazari et al., 2017), have confirmed the hidden potential of these uncommon actinomycetes as prolific producers of specialized metabolites. In this paper, we report that another member of this genus, *Nonomuraea coxensis* DSM 45129, which was isolated in Bangladesh in 2007 (Ara et al., 2007), has the genetic potential to produce A40926 or a very similar GPA; its BGC contains two regulatory genes, *nocRI* and *nocRII*, which are close homologs of *dbv3* and *dbv4*, respectively. Thus, we first developed the molecular tools to manipulate both *N. gerenzanensis* and *N. coxensis*, we then selected the strongest heterologous promoter to drive gene expression in *Nonomuraea* spp., and finally we overexpressed both native and heterologous cluster-specific regulatory genes in *N. gerenzanensis*, assessing the best performers at flask and bioreactor scale in industrial media. The overexpression of the positive pathway-specific regulators significantly increased the level of A40926 production in *N. gerenzanensis*, paving the way for knowledge-based strain improvement for the production of this valuable GPA.

MATERIALS AND METHODS

Plasmids, Bacterial Strains, Antibiotics, and Culture Conditions

Plasmids and bacterial strains used in this work are summarized in **Table 1**. Compositions of media are given in ESM. Unless otherwise stated, all media components and antibiotics were supplied by Sigma-Aldrich, St. Louis, MO, United States. For routine maintenance, actinobacterial strains were cultivated on ISP3 or VM0.1 agar media supplemented with 50 μg/ml apramycin-sulfate when appropriate. For genomic DNA isolation, *N. gerenzanensis* ATCC 39727 and *N. coxensis* DSM 45129 were grown in 250 ml Erlenmeyer flasks containing 10 glass beads (ø5 mm) with 50 ml of liquid VSP medium on an orbital shaker at 220 rpm and at 30°C. Working cell banks (WCB) of *Nonomuraea* spp. were prepared as described previously (Marcone et al., 2014). *Escherichia coli* DH5α was used as a routine cloning host and *E. coli* ET12567 pUZ8002 as a donor for intergeneric conjugations. *E. coli* strains were grown at 37°C in LB liquid or agar media supplemented with 100 μg/ml of apramycin-sulfate, 50 μg/ml of kanamycin-sulfate and 25 μg/ml of chloramphenicol when appropriate.

Generation of Recombinant Plasmids

Construction of Promoter-Probe Vectors

To test the activity of different native *N. gerenzanensis* promoters, pSAGA (Koshla et al., 2019), where *gusA* (Myronovskyi et al., 2011) is expressed from *aac(3)IVp*, was chosen as a chassis. Genomic DNA, extracted from *N. gerenzanensis* according to the Kirby procedure (Kieser et al., 2000), was used as a template

to amplify the putative promoter regions of: *BN4615_P641* (coding for the DNA gyrase B subunit – $GyrB_{ng}$) 342 bp; *BN4615_P8899* (coding for the RNA polymerase sigma factor RpoD – $HrdB_{ng}$), 524 bp; *BN4615_P604* (coding for the single-stranded DNA-binding protein – Ssb_{ng}), 207 bp; *BN4615_P1543* (coding for the SSU ribosomal protein S12p – $RpsL_{ng}$), 339 bp; *BN4615_P7269* (coding for the rifamycin-resistant RNA polymerase subunit β – $RpoB_{Rng}$), 558 bp; and *BN4615_P1539* (coding for the rifamycin-sensitive RNA polymerase subunit β – $RpoB_{ng}$), 493 bp. Amplicons were generated using Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, United States) according to the supplier's protocol and the oligonucleotide primers listed in **Table 2**. All of the amplicons were digested with *Bam*HI and *Eco*RV and cloned in pSAGA cleaved with the same endonucleases, thus replacing *aac(3)IVp* in front of *gusA* with each of the amplified promoter regions. The resulting recombinant plasmids were named pGBP (carrying $gyrB_{ng}p$), pHBP ($hrdB_{ng}p$), pSBP ($ssb_{ng}p$), pRLP ($rpsL_{ng}p$), pRBP1 ($rpoB_{ng}p$), and pRBP2 ($rpoB_{Rng}p$).

Construction of the *dbv3*, *dbv4*, and *nocRI* Overexpression Plasmids

The coding sequences of *dbv3* (2,635 bp) and *dbv4* (1,006 bp) were amplified from the A40Y cosmid (**Table 1**; Marcone et al., 2010a) using Q5 High-Fidelity DNA Polymerase and the dbv3_F/R or dbv4_F/R primer pairs (**Table 2**). The obtained amplicons were digested with *Eco*RI and *Eco*RV and cloned into pSET152A cleaved with the same enzymes. The resulting plasmids were named pSAD3 and pSAD4. To generate a vector for the overexpression of both *dbv3* and *dbv4*, the regions containing the coding sequences of both genes were amplified using the dbv4_F/dbv3_R primer pair and cloned into pSET152A in a similar fashion, generating pSAD3–4.

To construct the vector for overexpression of *nocRI* (the *dbv3* ortholog from *N. coxensis*), the coding sequence of *A3G7_RS0138355* was amplified from the genomic DNA of *N. coxensis* isolated using the Kirby procedure (Kieser et al., 2000) using Q5 High-Fidelity DNA Polymerase and the orfR1_F/R primer pair (**Table 2**). The obtained amplicon (2,661 bp) was digested with *Eco*RI and *Eco*RV and cloned into pSET152A cut with the same enzymes to generate pSAR1. All of the generated recombinant plasmids were verified by restriction endonuclease mapping and sequencing at BMR Genomics (University of Padua, Italy).

Conjugative Transfer of Plasmids Into *Nonomuraea* spp. and Verification of the Recombinant Strains

Conjugative transfer of plasmids into *N. gerenzanensis* was performed essentially as described previously (Marcone et al., 2010c). All recombinant plasmids were transferred individually into the non-methylating *E. coli* ET12567 pUZ8002 and the resulting derivatives used as donor strains for intergeneric conjugation. To prepare fresh vegetative mycelium of *N. gerenzanensis* prior to conjugal transfer, one vial of WCB was inoculated into 50 ml of VSP medium (250 ml Erlenmeyer

TABLE 1 | Bacterial strains and plasmids used in this work.

Name	Description	Source of reference
N. gerenzanensis	Wild type, A40926 producer	ATCC 39727
N. coxensis	Wild type	DSM 45129
N. gerenzanensis pSET152A$^+$	Wild type derivative carrying pSET152A	This work
N. gerenzanensis pSAD3$^+$	Wild type derivative carrying pSAD3	This work
N. gerenzanensis pSAD4$^+$	Wild type derivative carrying pSAD4	This work
N. gerenzanensis pSAD3–4$^+$	Wild type derivative carrying pSAD3–4	This work
N. gerenzanensis pSAR1$^+$	Wild type derivative carrying pSAR1	This work
N. gerenzanensis pSAGA$^+$	Wild type derivative carrying pSAGA	This work
N. gerenzanensis pTEGA$^+$	Wild type derivative carrying pTEGA	This work
N. gerenzanensis pGUSmoeE5script$^+$	Wild type derivative carrying pGUSmoeE5script	This work
N. gerenzanensis pGCymRP21$^+$	Wild type derivative carrying pGCymRP21	This work
N. gerenzanensis pGT2p$^+$	Wild type derivative carrying pGT2p	This work
N. gerenzanensis pGBP$^+$	Wild type derivative carrying pGBP	This work
N. gerenzanensis pHBP$^+$	Wild type derivative carrying pHBP	This work
N. gerenzanensis pSBP$^+$	Wild type derivative carrying pSBP	This work
N. gerenzanensis pRLP$^+$	Wild type derivative carrying pRLP	This work
N. gerenzanensis pRBP1$^+$	Wild type derivative carrying pRBP1	This work
N. gerenzanensis pRBP2$^+$	Wild type derivative carrying pRBP2	This work
N. coxensis pSAGA$^+$	Wild type derivative carrying pSAGA	This work
N. coxensis pTEGA$^+$	Wild type derivative carrying pTEGA	This work
N. coxensis pGUSmoeE5script$^+$	Wild type derivative carrying pGUSmoeE5script	This work
N. coxensis pGCymRP21$^+$	Wild type derivative carrying pGCymRP21	This work
N. coxensis pGT2p$^+$	Wild type derivative carrying pGT2p	This work
N. coxensis pGBP$^+$	Wild type derivative carrying pGBP	This work
N. coxensis pHBP$^+$	Wild type derivative carrying pHBP	This work
N. coxensis pSBP$^+$	Wild type derivative carrying pSBP	This work
N. coxensis pRLP$^+$	Wild type derivative carrying pRLP	This work
N. coxensis pRBP1$^+$	Wild type derivative carrying pRBP1	This work
N. coxensis pRBP2$^+$	Wild type derivative carrying pRBP2	This work
E. coli DH5α	General cloning host (*dam-13*::Tn9 *dcm-6*),	MBI Fermentas, USA
E. coli ET12567 pUZ8002	pUZ8002$^+$ (Δ*oriT*), used for conjugative transfer of DNA	Kieser et al. (2000)

(Continued)

TABLE 1 | Continued

Name	Description	Source of reference
A40Y	SuperCos1 derivative, including 22 kb of *dbv* cluster (*dbv1-dbv17*)	Marcone et al. (2010a)
pGUS	pSET152 derivative, containing promoterless *gusA*	Myronovskyi et al. (2011)
pSET152A	pSET152 derivative, containing *aac(3)IVp* from pIJ773	Horbal et al. (2013)
pSAD3	pSET152A derivative, containing *dbv3* under the control of *aac(3)IVp*	This work
pSAD4	pSET152A derivative, containing *dbv4* under the control of *aac(3)IVp*	This work
pSAD3–4	pSET152A derivative, containing *dbv4* together with *dbv3*	This work
pSAR1	pSET152A derivative, containing *nocRI* under the control of *aac(3)IVp*	This work
pSAGA	pSET152A derivative, containing *gusA* under the control of *aac(3)IVp*	Koshla et al. (2019)
pTEGA	pTES derivative, containing *gusA* under the control of *ermEp*	Yushchuk et al. (2020a)
pGUSmoeE5script	pGUS derivative, containing *gusA* under the control of *moeE5p*	Makitrynskyy et al. (2013)
pGCymRP21	pGUS derivative, containing CymR operator, *P21* promoter and *cymR* gene	Horbal et al. (2014a)
pGT2p	pGUS derivative, containing *gusA* under the control of *tei2p*	Yushchuk et al. (2020b)
pGBP	pSAGA derivative, containing *gusA* under the control of $gyrB_{ng}p$	This work
pHBP	pSAGA derivative, containing *gusA* under the control of $hrdB_{ng}p$	This work
pSBP	pSAGA derivative, containing *gusA* under the control of $ssb_{ng}p$	This work
pRLP	pSAGA derivative, containing *gusA* under the control of $rpsL_{ng}p$	This work
pRBP1	pSAGA derivative, containing *gusA* under the control of $rpoB_{ng}p$	This work
pRBP2	pSAGA derivative, containing *gusA* under the control of $rpoB_{Rng}p$	This work

flask with 10 ø5 mm glass beads) and incubated for 48 h on the orbital shaker at 220 rpm and at 30°C. The mycelium was collected by centrifugation (10 min, 3,220 × *g*), washed twice with sterile 20% v/v glycerol, resuspended in the same solution to a final volume of 20 ml, and stored at −80°C. 1 ml of mycelial suspension was mixed with approximately 10^9 of donor *E. coli* cells and the mixtures were plated on well dried VM0.1 agar plates supplemented with 20 mM of $MgCl_2$. After 12–16 h of incubation at 30°C, each plate was overlaid with 1 ml of sterile deionized water containing 1.25 mg of apramycin-sulfate and 750 μg of nalidixic acid sodium salt. Transconjugants were selected as resistant to 50 μg/ml of apramycin-sulfate.

Spore suspensions of *N. coxensis* were prepared from lawns grown on ISP3 agar for 7 days. Spores from one plate were collected in deionized water and filtered through one layer of Miracloth (Merck KGaA, Darmstadt, Germany) to remove vegetative mycelial fragments. Then, spores were pelleted from a 50 ml suspension by centrifugation (15 min, 3,220 × *g*), resuspended in 1 ml of 15% v/v glycerol, and stored at −80°C. For conjugation, approx. 10^6 spores were mixed with 10^7 *E. coli* donor cells and plated on VM0.1 agar plates supplemented with 20 mM of $MgCl_2$. The overlay for the selection of transconjugants was performed as described previously for *N. gerenzanensis*.

To verify the integration of promoter-probe vectors, an ~1 kbp internal fragment of *gusA* was amplified from the genomic DNA of recombinant *N. gerenzanensis* or *N. coxensis* strains using the gusA_ver_F/R primer pair (**Table 2**). To verify the integration of pSET152A, *aac(3)IV* was amplified using the aac(3)IV_F/R primer pair (**Table 2**). To verify the integration of pSAD4, an ~1 kbp fragment of pSAD4 was amplified with the PAM_seq_F/dbv4_R primer pair (**Table 2**), in which PAM_seq_F anneals within the *aac(3)IVp* sequence. Verification of pSAD3 and pSAD3–4 integration was made by amplification of an ~2 kbp fragment (for pSAD3) or an ~3 kbp fragment (for pSAD3–4), using the PAM_seq_F/dbv3_seq_R primer pair (**Table 2**), in which dbv3_seq_R anneals in the middle of *dbv3*. Finally, to verify the integration of pSAR1 an ~2 kbp fragment was amplified using the PAM_seq_F/orfR1mid_EcoRV_R (**Table 2**) primer pair, in which orfR1mid_EcoRV_R anneals in the middle of *nocRI*.

β-Glucuronidase Activity Assay

β-Glucuronidase (GusA) activity in *Nonomuraea* strains growing on VM0.1 agar medium was assessed by adding, after 6 days of cultivation at 30°C, 10 μl drops of 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-Gluc, Thermo Fisher Scientific, Waltham, MA, United States) 50 mg/ml in DMSO to the surfaces of the lawns. Chromogenic conversion of X-Gluc into the blue-colored 5,5′-dibromo-4,4′-dichloro-indigo was monitored after 1 h of incubation. For the quantitative measurements of GusA activity, *Nonomuraea* strains were grown in liquid media. One WCB vial of each of the strains was inoculated into a baffled 500 ml Erlenmeyer flask containing 100 ml of E26 (*N. gerenzanensis* strains) or of VSP (*N. coxensis* strains). After 72 h of cultivation, 10% v/v of this preculture was transferred into a baffled 500 ml Erlenmeyer flask containing 100 ml of FM2 (*N. gerenzanensis* strains) or ISP2 (*N. coxensis* strains). To induce *P21-cmt*-driven *gusA*-expression, cumate was added at the final concentration of 50 μM to cultures carrying pGCymRP21 24 h after inoculation. After 120 h of cultivation, mycelial lysates were prepared as previously reported by Horbal et al., 2013. Glucuronidase activity was measured as previously described (Myronovskyi et al., 2011; Horbal et al., 2013) using

TABLE 2 | Oligonucleotide primers used in this work.

Primer	Nucleotide sequence (5′-3′)*	Purpose
Dbv3_F	TTTGATATCGGAGGGCAAGAGTGCTGTTCGGGC	Cloning of *dbv3* into pSET152A
Dbv3_R	TTTGAATTCCTCTACAGCCGCACTGCCT	
dbv4_F	TTTGATATCGGAGGGGCTAGGTGGACCCGACGG	Cloning of *dbv4* into pSET152A
dbv4_R	TTTGAATTCTCCACTCGTGCTCATCCAG	
PAM_seq_F	GATGTCATCAGCGGTGGAG	Verification of recombinant strains
dbv3_seq_R	CCAGCGCTGGACCGCCTGC	
orfR1nid_EcoRV_rev	TTTGATATCGCAAGGGGCCTCCCCGCCG	
orfR1_F	TTTGATATCGGAGGACTGCGTTGACGAACCGCT	Cloning of *nocRI* into pSET152A
orfR1_R	TTTGAATTCGCGTCATGGGACCACCGCC	
hrdBp_F	TTTGGATCCACCGAAGCGCCGCCTGAGG	Cloning of $hrdB_{Ng}p$
hrdBp_R	TTTGATATCGAAGGCCTGACGGACATCC	
rpoB1p_F	TTTGGATCCTCTCGCTGGCTGGTGGCCG	Cloning of $rpoB_{Ng}p$
rpoB1p_R	TTTGATATCTCGCGGCGGACTGACTACA	
rpoB2p_F	TTTGGATCCTGTCGTACTGCTC	Cloning of $rpoB_{RNg}p$
rpoB2p_R	TTTGATATCATACGAAGGCGAGGGAGGG	
rpsLp_F	TTTGGATCCATGGACGGCGGAGCTGTAG	Cloning of $rpsL_{Ng}p$
rpsLp_R	TTTGATATCTTGGCCGGTGTTACGTCA	
ssbp_F	TTTGGATCCAAGTCCGAAGGCATCTACG	Cloning of $ssb_{Ng}p$
ssbp_R	TTTGATATCTGCACGCCTTCGCTTGGGT	
gyrBAp_F	TTTGGATCCAGGCTTCGCACAGTAACGG	Cloning of $gyrB_{Ng}p$
gyrBAp_R	TTTGATATCGCGGACACGCGGCGGGGGA	
aac(3)IV_F	ATCGACTGATGTCATCAGCG	Amplification of *aac(3)IV*
aac(3)IV_R	CGAGCTGAAGAAAGACAAT	
gusA_ver_F	GGCGGCTACACGCCCTTCGA	Amplification of *gusA* internal fragment
gusA_ver_R	TGATGGGCCGGGTGGGGTC	

*Restriction sites are underlined in primer sequence.

a spectrophotometric assay following the conversion of the colorless *p*-nitrophenyl-β-D-glucuronide (Thermo Fisher Scientific, Waltham, MA, United States) into the colored *p*-nitrophenol at 415 nm using an Infinite 200 PRO microplate reader (Tecan, Switzerland). Glucuronidase activity was normalized to dry biomass weight as previously reported (Marcone et al., 2014). One unit of activity is defined as the amount of enzyme that is able to convert 1 μM of substrate in 1 min.

A40926 Production

One WCB vial was inoculated into 300 ml baffled flasks containing 50 ml of vegetative medium E26 with 10 glass beads (ø5 mm). Flask cultures were incubated for 72 h on a rotary shaker at 220 rpm and 30°C and then used to inoculate (10% v/v) 500 ml baffled Erlenmeyer flasks containing 100 ml of FM2 medium or a 3-l P-100 Applikon glass reactor (height 250 mm, ø130 mm) equipped with a AD1030 Biocontroller and AD1032 motor, and containing 2 l of the same production medium. Cultivations in FM2 in shake-flasks were conducted at 30°C and 220 rpm. Bioreactor fermentations were conducted at 30°C, with stirring at 450 rpm (corresponding to 1.17 m/s of tip speed) and 2 l/min aeration rate. Dissolved oxygen (measured as % pO2) was monitored using an Ingold polarographic oxygen electrode. The pH values of culture broths were monitored using a pH meter. Foam production was controlled by adding Hodag antifoam (Hodag Chemical Corporation, Chicago, IL, United States) through an antifoam sensor. Samples were collected at regular cultivation time intervals and analyzed to estimate biomass (dry weight), glucose consumption (Diastix sticks, Bayer AG, Leverkusen, Germany), and A40926 production.

HPLC Analysis of Culture Extracts

A40926 was extracted from *Nonomuraea* spp. cultures as previously reported (Marcone et al., 2014). Chromatography was performed with a VWR Hitachi diode array L-2455 HPLC system with detection at 254 nm. The A40926 titers in the batch cultivations were estimated by injecting 50 μl of sample onto a 5 μm-particle-size Ultrasphere ODS (Beckman) HPLC column (4.6 by 250 mm) and eluting at a flow rate of 1 ml/min with a 30 min linear gradient from 15 to 64% of phase B. Phase A was 32 mM $HCOONH_4$ (pH 7) – CH_3CN [90:10 (vol/vol)], and phase B was 32 mM $HCOONH_4$ (pH 7) – CH_3CN [30:70 (vol/vol)]. A volume of 50 μl of a pure sample of 200 μg/ml A40926 (Sigma-Aldrich, St. Louis, MO, United States) was used as an internal standard.

Tools for the Bioinformatics Analysis

Blastp was used to search for homologs (Altschul, 1990); protein sequence alignments were performed with Clustal Omega (EMBL-EBI, Sievers et al., 2011).

RESULTS

Genetic Manipulation of *Nonomuraea coxensis*, a Novel Putative Producer of a A40926-Like Molecule

According to the 16S rRNA gene-based reconstruction of *Nonomuraea* phylogeny (Dalmastri et al., 2016), *N. coxensis* occupies a relatively distant position from *N. gerenzanensis*. Conversely, mining the partially sequenced genome of *N. coxensis*

(ASM37988v1), we found a close homolog of the *N. gerenzanensis* regulatory gene *dbv3* (locus *A3G7_RS0138355* in the *N. coxensis* genome, coding for WP_020547054.1 with 86.6% predicted amino acid sequence identity with Dbv3, ESM **Figure 1**). Dbv3 is a unique LuxR-regulator controlling the expression of the A40926 BGC and it is not closely related to the better-characterized family of Tei16*-like regulators controlling teicoplanin biosynthesis, sharing only 32% of amino acid sequence identity with Tei16* (Yushchuk et al., 2019). This *dbv3*-like gene (we named it *nocRI*) was found on a short *N. coxensis* contig (NZ_KB904006) flanking a gene coding for a putative StrR-like transcriptional regulator apparently orthologous to *dbv4* (locus *A3G7_*RS0138355, coding for WP_026215141.1 with 94.39% of amino acid sequence identity with Dbv4); *dbv4* is the other known cluster-situated regulatory gene in the *dbv* BGC and we named the *N. coxensis* homolog *nocRII* (ESM **Figure 2**). AntiSMASH (Blin et al., 2019) analysis of *N. coxensis* genome revealed the presence of three short contigs (NZ_KB904006, NZ_KB903995, NZ_KB903969, ESM **Figure 3**) covering the majority of a *dbv*-like BGC, suggesting that *N. coxensis* might produce A40926 or a very similar GPA. Thus, we considered *N. coxensis* an interesting candidate for developing *Nonomuraea*-targeted genetic tools.

Initially, we tried to transfer *φC31*-based integrative plasmids (**Table 1**) into *N. coxensis* by using the protocol of intergeneric conjugation optimized for conjugal transfer from a DNA-non-methylating *E. coli* donor strain to *N. gerenzanensis* vegetative mycelium (Marcone et al., 2010c). Transconjugants were obtained at a very low frequency (*ca.* 1×10^{-7}) and only for the relatively small pSAGA-based promoter-probe vectors (approx. 6 kbp). Increasing the amount of donor and recipient cells, changing the time of overlay, and adjusting medium composition (including increasing or decreasing $MgCl_2$ concentration) did not allow the transfer of the larger pGUS-based promoter-probe vectors (such as pGUSmoeE5script or pGCymRP21, both more than 9 kbp). Since, unlike *N. gerenzanensis*, *N. coxensis* sporulates abundantly when grown on ISP3 agar medium (ESM **Figure 4**), we tried to use spores for conjugal transfer. This resulted in transfer rates of approximately 1×10^{-3} when 10^6 spores were mixed with 10^7 *E. coli* donor cells, regardless of plasmid size.

Using the GusA-Reporter System to Assess the Activity of Native and Heterologous Promoters in *Nonomuraea* Species

The set of *φC31*-based integrative plasmids transferred to *N. gerenzanensis* and *N. coxensis* were promoter-probe vectors utilizing the GusA reporter system and carrying a selection of native and heterologous promoters, the latter having been used previously to drive gene expression in streptomycetes or actinoplanetes (Horbal et al., 2013, 2014a; Makitrynskyy et al., 2013). Glucuronidase activity of the recombinant *Nonomuraea* strains was assessed qualitatively (on agar plates) and quantitatively (in cell lysates obtained from mycelium grown in liquid medium). Both of the *N. coxensis* and *N. gerenzanensis* wild type strains did not display any glucuronidase activity (**Figure 1**). The heterologous promoters tested were: *aac(3)IVp* (in pSAGA) – the apramycin acetyltransferase gene promoter, derived from pSET152A (Horbal et al., 2013); *ermEp* (in pTEGA, Yushchuk et al., 2020a) – the erythromycin resistance gene promoter from pTES (Herrmann et al., 2012); *moeE5p* (in pGUSmoeE5script) – the *S. ghanaensis* moenomycin biosynthesis gene *moeE5* promoter (Makitrynskyy et al., 2013); *tei2p* (pGT2p, Yushchuk et al., 2020b) – the *A. teichomyceticus* teicoplanin resistance gene *tei2* promoter; and *P21*, a synthetic promoter fused with the cumate inducible *Pseudomonas putida* F1 *cmt* operon operator (in pGCymRP21, Horbal et al., 2014a). In parallel, we tested the activity of six native promoters derived from *N. gerenzanensis* house-keeping genes: $gyrB_{ng}p$ (in pGBP – the promoter of the DNA gyrase B subunit gene); $hrdB_{ng}p$

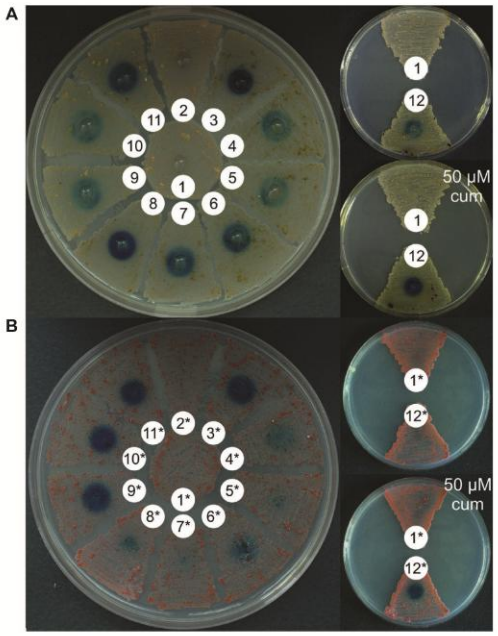


FIGURE 1 | Comparison of GusA activity in recombinant strains of *N. gerenzanensis* **(A)** and *N. coxensis* **(B)**, carrying the following promoter-probe vectors: 2, 2* – pTEGA (*gusA* under the control of *ermEp*); 3, 3* – pSAGA (*gusA* under the control of *aac(3)IVp*); 4, 4* – pHBP (*gusA* under the control of $hrdB_{ng}p$); 5, 5* – pRBP2 (*gusA* under the control of $rpoB_{2ng}p$); 6, 6* – pRBP1 (*gusA* under the control of $rpoB_{ng}p$); 7, 7* – pSBP (*gusA* under the control of $ssb_{ng}p$); 8, 8* – pRLP (*gusA* under the control of $rpsL_{ng}p$); 9, 9* – pGUSmoeE5script (*gusA* under the control of *moeE5p*); 10, 10* – pGT2p (*gusA* under the control of *tei2p*); 11, 11* – pGBP (*gusA* under the control of $gyrB_{ng}p$); 12, 12* – pGCymRP21 (*gusA* under the control of *P21-cmt*). The control parental strains (1, *N. gerenzanensis* and 1*, *N. coxensis*) do not display chromogenic conversion of X-Gluc. Strains were cultivated for 6 days on VM0.1 medium.

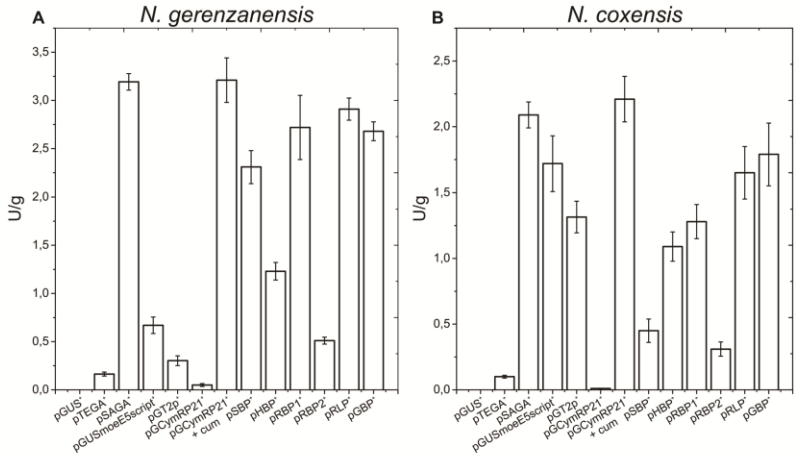


FIGURE 2 | Quantitative measurement of GusA activity (U/gram of dry biomass) in *N. gerenzanensis* **(A)** and *N. coxensis* **(B)** recombinant strains grown in liquid media for 120 h (see growth curves in ESM **Figure 5**), carrying different promoter-probe plasmids. Control strains carried the promoterless pGUS vector and exhibited no GusA activity. Activities are the mean values of three independent experiments. Error bars represent standard deviations.

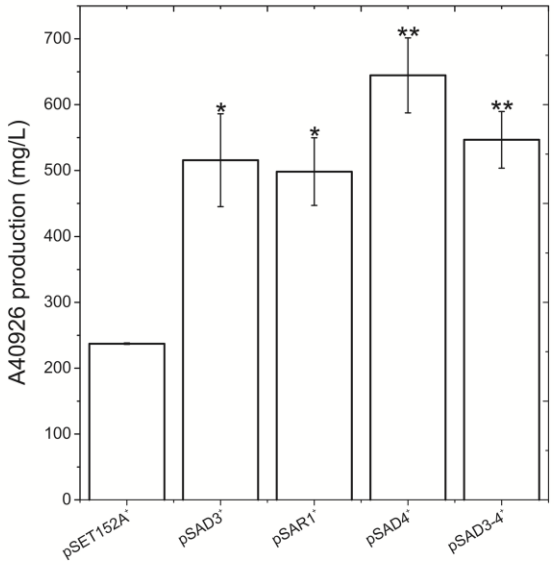


FIGURE 3 | A40926 production in *N. gerenzanensis* carrying the empty vector pSET152A^{+} and in its recombinant strains overexpressing *dbv3* (pSAD3), *nocRI* (pSAR1), *dbv4* (pSAD4), and co-expressing *dbv3* and *dbv4* genes (pSAD3–4) cultivated in FM2 industrial medium at 500 ml Erlenmeyer flasks-scale. A40926 production was measured after 120 h of cultivation. The results given represent three independent fermentations, error bars represent standard deviations. Statistical significance of the differences in A40926 production between the control and the recombinant strains was estimated using Welch's *t*-test: $^{*}p < 0.05$; $^{**}p < 0.01$.

(in pHBP – the promoter of the RNA polymerase σ-factor RpoD gene); $ssb_{Ng}p$ (in pSBP – the promoter of the single-stranded DNA-binding protein); $rpsL_{Ng}p$ (in pRLP – the promoter of the SSU ribosomal protein S12p gene); $rpoB_{Rng}p$ (in pRBP2 – the promoter of the rifamycin-resistant RNA polymerase subunit β gene); and $rpoB_{Ng}p$ (in pRBP1 – the

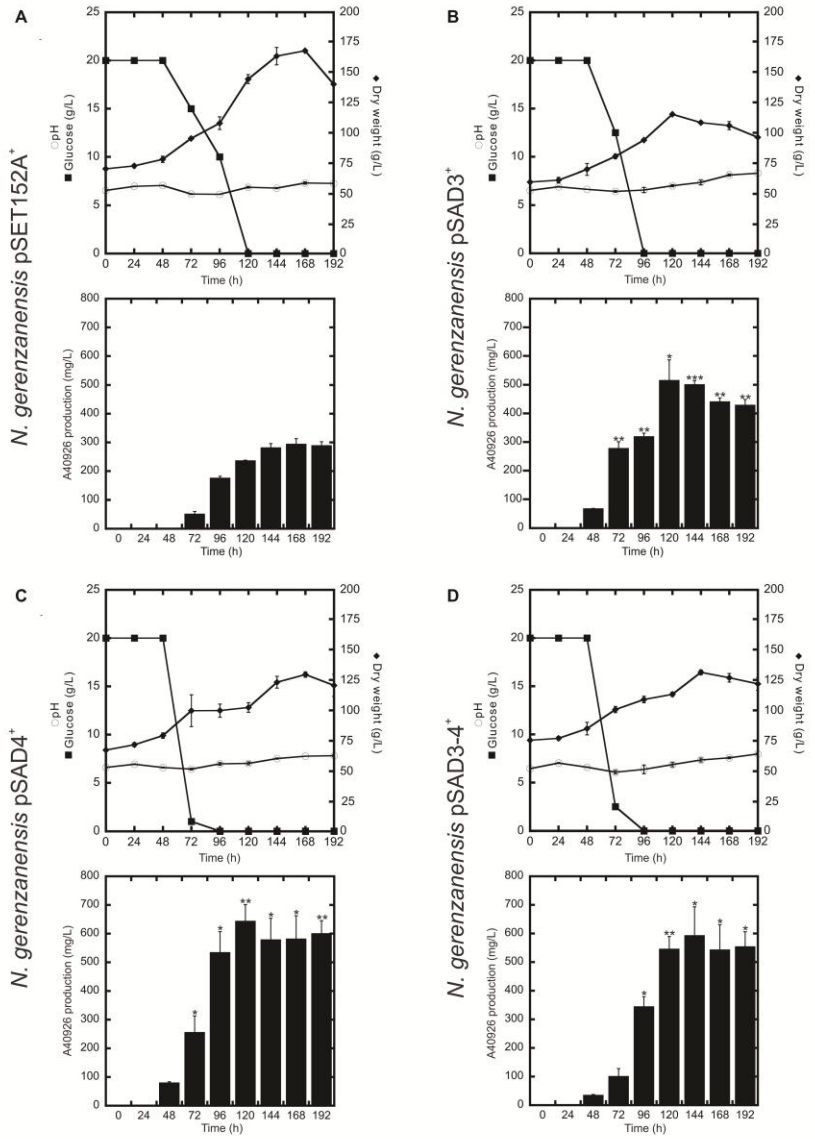


FIGURE 4 | Time courses of *N. gerenzanensis* pSET152A+ **(A)** and recombinant strains overexpressing *dbv3* **(B)**, *dbv4* **(C)** and co-expressing both genes **(D)** cultivated in FM2 industrial medium in 500 ml Erlenmeyer flasks. Glucose consumption (filled rectangles), biomass accumulation (filled rhombi), pH (circles), and A40926 production were monitored every 24 h. Results given are mean values of three independent experiments. Error bars represent standard deviations. Statistical significance of the differences in A40926 production between the control and the recombinant strains was estimated using Welch's *t*-test: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

promoter of the rifamycin-sensitive RNA polymerase subunit β gene). In contrast to the wild type strains, all of the recombinant derivatives grown on VM0.1 agar plates converted X-Gluc to its colored derivative 5,5′-dibromo-4,4′-dichloro-indigo, albeit to different extents (**Figure 1**). *N. gerenzanensis* pSAGA+ and *N. coxensis* pSAGA+, carrying *aac(3)IVp*, produced

the most intensive color (**Figures 1A,B**), while the chromogenic conversion of X-Gluc in *N. gerenzanensis* pTEGA⁺ and *N. coxensis* pTEGA⁺, carrying the *ermEp*, was only slightly visible. In strains carrying the inducible pGCymRP21, GusA activity was induced by the addition of 50 μM of cumate, proving inducible gene expression (**Figures 1A,B**). However, in *N. gerenzanensis* pGCymRP21⁺, a basal level of GusA activity was detected even in the absence of cumate (**Figure 1A**). Contrary to this, no basal level of expression from *P21-cmt* was detected in *N. coxensis* pGCymRP21⁺ (**Figure 1B**).

When the glucuronidase activity present in cell lysates obtained from liquid cultures grown for 120 h (late exponential/early stationary growth phase, see ESM **Figure 5**) was measured using a spectrophotometric assay, overall these quantitative results (normalized for the dry weight of the differently growing strains) correlated with those observed on agar plates: *aac(3)IVp* behaved as a strong promoter in both *N. gerenzanensis* and *N. coxensis*, surpassed only by *cmt-P21p* when induced by the addition of 50 μM cumate (**Figures 2A,B**). Also $rpsL_{Ng}p$ and $hrdB_{Ng}p$ proved to be strong promoters in both of the *Nonomuraea* spp. whereas the weakest was the *ermEp* (**Figures 2A,B**). Interestingly, the activity of $rpoB_{Ng}P$ was higher than that of $rpoB_{Rng}p$ (**Figures 1**, **2**), consistent with previously reported data about differences in transcription levels of the two alleles (Vigliotta et al., 2004).

Knowledge-Based Generation of A40926 Overproducing Strains

Since *aac(3)IVp* was identified as the strongest constitutive promoter for *Nonomuraea* spp., we used it to overexpress *dbv3* and *dbv4* in *N. gerenzanensis*. Both genes were cloned into the integrative pSET152A vector (Horbal et al., 2013) yielding the recombinant vectors pSAD3 and pSAD4, respectively (**Table 1**). Benefiting from the neighboring positions of *dbv3* and *dbv4*, we also cloned them together, generating pSAD3–4, where *dbv4* was directly under the control of *aac(3)IVp*, but *dbv3* remained under the control of its native promoter (**Table 1**). Additionally, the *dbv3*-like *nocRI* from *N. coxensis* was cloned into pSET152A, generating pSAR1, and transferred to *N. gerenzanensis* to determine if it could improve A40926 production, and to assess possible cross-talk between the regulators from the two different *Nonomuraea* spp. All the *N. gerenzanensis* recombinant strains were grown for 120 h in parallel with the control strain carrying the empty vector and A40926 production was measured (**Figure 3**). The production of A40926 in *N. gerenzanensis* carrying pSET152A reached almost 250 mg/l after 120 h of cultivation (**Figure 3**). The four recombinant strains overexpressing the cluster-situated regulatory genes produced more antibiotic than the control strain (**Figure 3**). A40926 production in *N. gerenzanensis* pSAD3⁺ and pSAR1⁺ was comparable (*ca.* 500 mg/l), proving that *nocRI* had a similar impact in *N. gerenzanensis* as *dbv3*. The recombinant strain carrying pSAD3–4 vector produced slightly more, around 550 mg/l, whereas the best producer in these conditions was the strain overexpressing *dbv4*, which produced more than 650 mg/l.

Time Courses of A40926 Production in *Nonomuraea gerenzanensis* Recombinant Strains at Flask and at Bioreactor Scale

Although further investigations will be devoted to the expression of the heterologous *nocRI* and *nocRII in N. gerenzanensis*, in our strain improvement work we then focused on *N. gerenzanensis* strains overexpressing native regulators. Consequently, *N. gerenzanensis* strains containing pSET152A, pSAD3, pSAD4, or pSAD3–4 were grown for 192 h in parallel with the control strain carrying the empty vector using the previously optimized industrial medium FM2 (Marcone et al., 2010a, 2014) at flask scale. Samples for the analysis of dry weight, pH, glucose consumption, and A40926 production were collected at regular 24 h intervals. All three recombinant strains expressing the regulatory genes from *aac(3)IVp* accumulated detectable amounts of A40926 earlier than the empty vector control (**Figure 4**). A40926 production reached its peak after 120 h of growth for *N. gerenzanensis* pSAD3⁺ (*ca.* 500 mg/l) (**Figure 4B**) and pSAD4⁺ (nearly 650 mg/l) (**Figure 4C**), whereas in pSAD3–4⁺ the maximum productivity (*ca.* 600 mg/l) was delayed to 144 h (**Figure 4D**). The control strain produced *ca.* 300 mg/l after 144–168 h from inoculation (**Figure 4A**). Although glucose consumption was faster in the recombinant strains containing the cloned regulatory genes in comparison to the control, they accumulated less biomass than *N. gerenzanensis* pSET152A⁺ (**Figure 4**). Maximum biomass production was around 125 g/l (dry weight) in the overexpression strains versus the 175 g/l produced by the empty vector control strain. These data suggest that part of the consumed glucose was used by the strains carrying the regulatory genes under the control of the strong constitutive promoter *aac(3)IVp* to support antibiotic production at the detriment of biomass formation.

When the cultivation of recombinant strains was scaled up in 3 l vessel-bioreactors containing 2 l working volume of FM2, all of the strains produced significantly more A40926 and grew better than in flask culture with the exception of *N. gerenzanensis* pSAD3⁺ (**Figure 5**). The control strain with the empty vector grew and produced the antibiotic faster than at the flask level (**Figure 5A**). Maximum biomass (300 g/l dry weight, more than the double of that achieved in flasks) and A40926 production (nearly 400 mg/l) were reached after 120 h from inoculation; glucose was completely consumed within 96 h versus the 120 h needed at flask level. The recombinant strain pSAD4⁺ grew more (maximum biomass production of 220 g/l after 144 h of growth) than in flask culture, although less than the control strain in the bioreactor. Glucose was consumed faster than in the control strain and glucose concentration tended to zero at 48 h of fermentation (**Figure 5C**), and antibiotic production reached a peak of 700 mg/l after 168 h. The best performance in terms of A40926 productivity in the bioreactor was shown by the pSAD3–4⁺ strain, which grew better (nearly 250 g/l dry weight biomass) than in flask culture and produced the maximum concentration of the antibiotic (800 mg/l) after 168 h of cultivation (**Figure 5D**). Conversely, the pSAD3⁺ strain showed a reduced biomass production in comparison to all of the other strains, with consumption of glucose markedly delayed and A40926 production

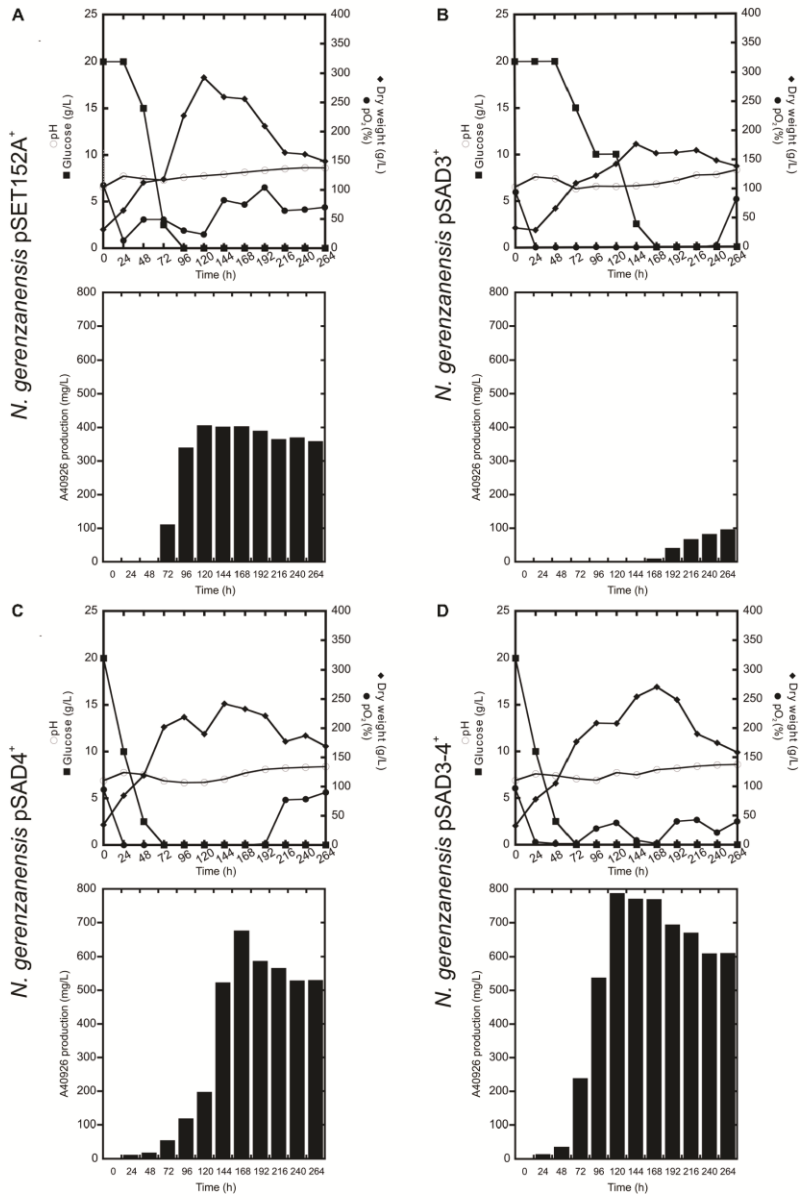


FIGURE 5 | Time courses of *N. gerenzanensis* pSET152A+ **(A)** and recombinant strains, overexpressing *dbv3* **(B)**, *dbv4* **(C)** and co-expressing both genes **(D)** cultivated in 2 l FM2 industrial medium in a 3 l fermenter. Glucose consumption (filled rectangles), biomass accumulation (filled rhombi), pH (circles), O_2 level (filled circles), and A40926 production were monitored every 24 h.

starting very late (after 168 h of cultivation) and never exceeding 100 mg/l (**Figure 5B**). Microscopical observation of pSAD3+ strain showed that the mycelium was highly fragmented (data not shown), suggesting some kind of physiological stress resulting from *dbv3* overexpression. Some fragmentation of the pSAD3+ strain (as well as of the pSAR1 strain carrying the *dbv3*-like

nocRI from *N. coxensis*) was also observed in flask culture, in contrast to the other strains which produced dense mycelial pellets; the fragmentation was less pronounced, indicating that scaling up in the bioreactor dramatically enhanced this effect. Interestingly, in the best performer at the bioreactor scale – the pSAD3–4^{+} strain – *dbv4* was expressed from the strong constitutive *aac(3)IVp* but *dbv3* was left under its endogenous promoter, suggesting that the balance between the expression level of the two cluster-situated regulatory genes is important for optimal improvement of A40926 production.

DISCUSSION

Apart from some reports on cultivating and manipulating the industrially valuable A40926 producer *N. gerenzanensis* (Stinchi et al., 2003, 2006; Marcone et al., 2010a,b,c; Alt et al., 2019) and the kistamicin producer *Nonomuraea* sp. ATCC 55076 (Greule et al., 2019), we are not aware of any other attempt to develop genetic tools for manipulating *Nonomuraea* spp. However, some other glycopeptide producers, like *A. teichomyceticus*, already possess well-developed toolkits for genetic manipulation which has greatly simplified investigations in these strains (Horbal et al., 2013; Yushchuk et al., 2016). Therefore, our first goal in this work was to develop a set of genetic tools for manipulating diverse species of *Nonomuraea*. To do this, we decided to work in parallel with the better-known A40926 producer, recently re-classified as *N. gerenzanensis* (Dalmastri et al., 2016), and with the little investigated *N. coxensis*, which was isolated in Bangladesh in 2007 (Ara et al., 2007). Although the available *N. coxensis* genome sequence is still incomplete, we could identify three contigs covering most of a *dbv*-like gene cluster including the *dbv*-like cluster-situated regulatory genes, which we named *nocRI* (*dbv3* homolog) and *nocRII* (*dbv4* homolog). The next step of our work will be additional sequencing to yield a properly annotated *N. coxensis* genome. Interestingly, during the course of our investigations, Waglechner et al. (2019), systematically screening the available sequences in genomic databases, also reported the presence of a BGC encoding for a A40926-like GPA in the genome of *N. coxensis*. In addition, the same authors reported that a newly isolated *Nonomuraea* sp. WAC01424 possess a BGC which could produce another A40926-related compound (Waglechner et al., 2019). Our preliminary analysis of this BGC suggests that it could be a sulfated A40926-like GPA, lacking the aliphatic side chains.

In the meantime, we tested both in *N. gerenzanensis* and *N. coxensis*, a set of heterologous and native promoters (the latter derived from a set of house-keeping genes in *N. gerenzanensis*) with the final goal of using them for driving gene expression in these strains. Besides the practical outcome of this screening (all of the generated promoter-probe vectors could be easily used as expression vectors simply by exchanging *gusA* for a gene of interest, offering a set of variable tools for gene expression), it is interesting to observe that the studied promoters had similar strengths in the two phylogenetically distant *Nonomuraea* species and that the native promoters from *N. gerenzanensis* worked similarly in *N. coxensis*. Consistently, we detected a comparable difference in the two strains between the strength of promoters driving the expression of rifamycin-sensitive and rifamycin-resistant *rpoB* alleles of *N. gerenzanensis* ($rpoB_{ng}p$ and $rpoB_{Rng}p$). Since we found the two *rpoB* alleles in *N. coxensis* draft genome (on a short genomic scaffolds KB904038 and KB904038), we might suppose that their manipulation could improve antibiotic production in this strain as already reported for *N. gerenzanensis* (Vigliotta et al., 2004). More generally, these vectors should be useful for the genetic manipulation of other members of the genus *Nonomuraea*, which have the potential to produce novel valuable specialized metabolites (Sungthong and Nakaew, 2015; Nazari et al., 2017).

As in the case of *Actinoplanes* spp. (Horbal et al., 2013), the strongest heterologous promoter was *aac(3)IVp*, although some *N. gerenzanensis* native promoters like $hrdB_{ng}p$ and $rpsL_{ng}p$ appeared to have comparable strength and merit further investigations. When we used *aac(3)IVp* to overexpress the cluster-situated regulatory genes *dbv3* and *dbv4* from the *N. gerenzanensis dbv* gene cluster and *nocRI* (*dbv3*-like) from the A40926-like BGC of *N. coxensis*, the recombinant *N. gerenzanensis* strains produced significantly more A40926 than the parental strain. The evidence that the heterologous expression of *nocRI* increased A40926 production in *N. gerenzanensis* confirmed its role in regulating the expression of a A40926-like BGC in *N. coxensis*. Additionally, it represents another case of cross-talk between regulators controlling GPA BGCs in producing actinomycetes (Spohn et al., 2014). Our next goal will be to investigate if and how (in which cultivation conditions) *N. coxensis* produces A40926 or a A40926-like molecule.

Previous work (Lo Grasso et al., 2015) reported that overexpression of *dbv3* (under the control of the thiostrepton-inducible *tipA** promoter in the integrative plasmid pIJ8600) in *N. gerenzanensis* increased A40926 production from 13 to 27 mg/l using the laboratory medium R3 (Lo Grasso et al., 2015). In this paper, we tested the real industrial potential of overexpressing not only *dbv3*, but also *dbv4*, and *dbv3* and *dbv4* together, cloning them under the strong constitutive *aac(3)IVp* promoter, scaling up their cultivation at bioreactor scale and using a previously optimized industrial medium where A40926 is produced in hundreds of milligrams per liter (Marcone et al., 2010a, 2014). At the bioreactor level, where strains could have a different performance from the flask-cultivation due to different mixing and mass transfer rates of nutrients and oxygen, the best performer was the strain carrying both *dbv4* under *aac(3)IVp* and *dbv3* under its own endogenous promoter. This strain produced nearly 800 mg/l of A40926, which is twice that of the parental strain grown under the same conditions. Conversely, the strain carrying only *dbv3* expressed from *aac(3)IVp* showed reduced production capacity and an altered phenotype particularly after scaling up from flask to bioreactor; in contrast, the *dbv4* overexpressing recombinant grew similarly under both conditions. It is widely recognized that any potentially higher producing mutant or derivative needs to be validated in a bioreactor-scale fermentation since unpredictable discrepancies

in strain performance can occur during scaling up from flask culture (Lee and Kim, 2015). In the case of the *dbv3* overexpressing recombinant, we believe that the fragmented mycelial phenotype, which was much more apparent in the bioreactor, could be a specific consequence of the overexpression of the *dbv7* gene encoding a D,D-carboxypeptidase known as VanYn (Binda et al., 2012). The level of VanYn activity in cell extracts from the *dbv3*-carrying recombinant cultivated at bioreactor scale was found to be much higher than in the parental strain (unpublished data). Consistent with this, overexpression of VanYn altered the mycelial morphology in *N. gerenzanensis* as well as in heterologous hosts such as streptomycetes strains (Binda et al., 2013).

In conclusion, only a few GPAs are used in clinical practice and those produced by semi-synthesis from natural products, such as dalbavancin derived from A40926, are still quite expensive. Dalbavancin is the first antibiotic designated as a Qualified Infection Diseases Product by the FDA because of its potency, extended dosing interval, and unique dose regimen, but its cost largely exceeds that of first-generation GPAs and consequently its use in hospitals is still limited (Chiasson and White, 2016; Agarwal et al., 2018). Improving A40926-producing strains might lead to a decrease in the cost of dalbavancin. As demonstrated in this paper, A40926 production could be significantly enhanced by manipulating the expression of *dbv* cluster-situated regulators. An important and often-neglected aspect is testing the genetic stability and productivity of the selected recombinant strains in a fully developed industrial process at bioreactor level, which mimics the conditions of antibiotic large scale production. We were able to demonstrate here that the improvements we made to A40926 production levels in shake flasks were also achieved in the bioreactor, indicating the relevance of this approach to industrial-scale strain improvement.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

OY, EB, MB, and FM conceived and designed the experiments and wrote the paper. OY, AA-V, GM, and EB performed the experiments. OY, GM, and EB analyzed the data.

FUNDING

This work was supported by public grants "Fondo di Ateneo per la Ricerca" 2018 and 2019 to FM and GM and by Federation of European Microbiological Societies (FEMS) Research Fellowship 2015 to EB.

ACKNOWLEDGMENTS

Consorzio Interuniversitario per le Biotecnologie is acknowledged for its support to traveling and attending congresses for EB and GM.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.00008/full#supplementary-material

REFERENCES

Adamek, M., Alanjary, M., Sales-Ortells, H., Goodfellow, M., Bull, A. T., Winkler, A., et al. (2018). Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC Genomics* 19, 1–15. doi: 10.1186/s12864-018-4809-4

Agarwal, R., Bartsch, S., Kelly, B., Prewitt, M., Liu, Y., Chen, Y., et al. (2018). Newer glycopeptide antibiotics for treatment of complicated skin and soft tissue infections: systematic review, network meta-analysis and cost analysis. *Clin. Microbiol. Infect.* 24, 361–368. doi: 10.1016/j.cmi.2017.08.028

Alduina, R., Sosio, M., and Donadio, S. (2018). Complex regulatory networks governing production of the glycopeptide A40926. *Antibiotics* 7, pii: E30. doi: 10.3390/antibiotics7020030

Alt, S., Bernasconi, A., Sosio, M., Brunati, C., Donadio, S., and Maffioli, S. I. (2019). Toward single-peak dalbavancin analogs through biology and chemistry. *ACS Chem. Biol.* 14, 356–360. doi: 10.1021/acschembio.9b00050

Altschul, S. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Ara, I., Kudo, T., Matsumoto, A., Takahashi, Y., and Omura, S. (2007). *Nonomuraea bangladeshensis* sp. nov. and *Nonomuraea coxensis* sp. nov. *Int. J. Syst. Evol. Microbiol.* 57, 1504–1509. doi: 10.1099/ijs.0.65054-0

Armenia, I., Marcone, G. L., Berini, F., Orlandi, V. T., Pirrone, C., Martegani, E., et al. (2018). Magnetic nanoconjugated teicoplanin: a novel tool for bacterial infection site targeting. *Front. Microbiol.* 9, 1–17. doi: 10.3389/fmicb.2018.02270

Banik, J. J., and Brady, S. F. (2008). Cloning and characterization of new glycopeptide gene clusters found in an environmental DNA megalibrary. *Proc. Natl. Acad. Sci.* 105, 17273–17277. doi: 10.1073/pnas.0807564105

Banik, J. J., Craig, J. W., Calle, P. Y., and Brady, S. F. (2010). Tailoring enzyme-rich environmental DNA clones: a source of enzymes for generating libraries of unnatural natural products. *J. Am. Chem. Soc.* 132, 15661–15670. doi: 10.1021/ja105825a

Bibb, M. J. (2013). Understanding and manipulating antibiotic production in actinomycetes. *Biochem. Soc. Trans.* 41, 1355–1364. doi: 10.1042/BST20130214

Binda, E., Marcone, G. L., Berini, F., Pollegioni, L., and Marinelli, F. (2013). *Streptomyces* spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance. *BMC Biotechnol.* 13, 1–14. doi: 10.1186/1472-6750-13-24

Binda, E., Marcone, G. L., Pollegioni, L., and Marinelli, F. (2012). Characterization of VanYn, a novel D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727. *FEBS J.* 279, 3203–3213. doi: 10.1111/j.1742-4658.2012.08706.x

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310

Chiasson, J., and White, R. (2016). Comparison of dalbavancin, oritavancin and vancomycin pharmacodynamics using clinical targets against ABSSSI pathogens. *Open Forum Infect. Dis.* 3:S515. doi: 10.1093/ofid/ofw172.1506

Dalmastri, C., Gastaldo, L., Marcone, G. L., Binda, E., Congiu, T., and Marinelli, F. (2016). Classification of *Nonomuraea* sp. ATCC 39727, an actinomycete that produces the glycopeptide antibiotic A40926, as *Nonomuraea gerenzanensis* sp. nov. *Int. J. Syst. Evol. Microbiol.* 66, 912–921. doi: 10.1099/ijsem.0.000810

D'Argenio, V., Petrillo, M., Pasanisi, D., Pagliarulo, C., Colicchio, R., Talà, A., et al. (2016). The complete 12Mb genome and transcriptome of *Nonomuraea gerenzanensis* with new insights into its duplicated "magic" RNA polymerase. *Sci. Rep.* 6, 1–13. doi: 10.1038/s41598-016-0025-0

Greule, A., Izoré, T., Iftime, D., Tailhades, J., Schoppet, M., Zhao, Y., et al. (2019). Kistamicin biosynthesis reveals the biosynthetic requirements for production of highly crosslinked glycopeptide antibiotics. *Nat. Commun.* 10, 1–15. doi: 10.1038/s41467-019-10384-w

Haslinger, K., Peschke, M., Brieke, C., Maximowitsch, E., and Cryle, M. J. (2015). X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. *Nature* 521, 105–109. doi: 10.1038/nature14141

Herrmann, S., Siegl, T., Luzhetska, M., Petzke, L., Jilg, C., Welle, E., et al. (2012). Site-specific recombination strategies for engineering actinomycete genomes. *Appl. Environ. Microbiol.* 78, 1804–1812. doi: 10.1128/AEM.06054-11

Horbal, L., Fedorenko, V., and Luzhetskyy, A. (2014a). Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. *Appl. Microbiol. Biotechnol.* 98, 8641–8655. doi: 10.1007/s00253-014-5918-x

Horbal, L., Kobylyanskyy, A., Truman, A. W., Zaburranyi, N., Ostash, B., Luzhetskyy, A., et al. (2014b). The pathway-specific regulatory genes, tei15* and tei16*, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. *Appl. Microbiol. Biotechnol.* 98, 9295–9309. doi: 10.1007/s00253-014-5969-z

Horbal, L., Kobylyanskyy, A., Yushchuk, O., Zaburannyi, N., Luzhetskyy, A., Ostash, B., et al. (2013). Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus* – producer of teicoplanin, drug of last defense. *J. Biotechnol.* 168, 367–372. doi: 10.1016/j.jbiotec.2013.10.018

Horbal, L., Zaburannyy, N., Ostash, B., Shulga, S., and Fedorenko, V. (2012). Manipulating the regulatory genes for teicoplanin production in *Actinoplanes teichomyceticus*. *World J. Microbiol. Biotechnol.* 28, 2095–2100. doi: 10.1007/s11274-012-1013-6

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical streptomyces genetics*. Norwich: John Innes Foundation.

Koshla, O., Yushchuk, O., Ostash, I., Dacyuk, Y., Myronovskyi, M., Jäger, G., et al. (2019). Gene *miaA* for post-transcriptional modification of tRNA XXA is important for morphological and metabolic differentiation in *Streptomyces*. *Mol. Microbiol.* 112, 249–265. doi: 10.1111/mmi.14266

Kusserow, K., and Gulder, T. A. M. (2017). Complete genome sequence of *Actinomadura parvosata* subsp. *kistnae*, a rich source of novel natural product (bio-)chemistry. *J. Genomics* 5, 75–76. doi: 10.7150/jgen.19673

Lee, S. Y., and Kim, H. U. (2015). Systems strategies for developing industrial microbial strains. *Nat. Biotechnol.* 33, 1061–1072. doi: 10.1038/nbt.3365

Lo Grasso, L. L., Maffioli, S., Sosio, M., Bibb, M., Puglia, A. M., and Alduina, R. (2015). Two master switch regulators trigger A40926 biosynthesis in *Nonomuraea* sp. strain ATCC 39727. *J. Bacteriol.* 197, 2536–2544. doi: 10.1128/JB.00262-15

Makitrynskyy, R., Ostash, B., Tsypik, O., Rebets, Y., Doud, E., Meredith, T., et al. (2013). Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. *Open Biol.* 3:130121. doi: 10.1098/rsob.130121

Marcone, G. L., Beltrametti, F., Binda, E., Carrano, L., Foulston, L., Hesketh, A., et al. (2010a). Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother.* 54, 2465–2472. doi: 10.1128/AAC.00106-10

Marcone, G. L., Binda, E., Berini, F., and Marinelli, F. (2018). Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era. *Biotechnol. Adv.* 36, 534–554. doi: 10.1016/j.biotechadv.2018.02.009

Marcone, G. L., Binda, E., Carrano, L., Bibb, M., and Marinelli, F. (2014). Relationship between glycopeptide production and resistance in the actinomycete *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother.* 58, 5191–5201. doi: 10.1128/AAC.02626-14

Marcone, G. L., Carrano, L., Marinelli, F., and Beltrametti, F. (2010b). Protoplast formationa and reversion to normal growth in antibiotic-producing uncommon actinomycetes. *J. Antibiot.* 63, 83–88. doi: 10.1038/ja.2009.127

Marcone, G. L., Foulston, L., Binda, E., Marinelli, F., Bibb, M., and Beltrametti, F. (2010c). Methods for the genetic manipulation of *Nonomuraea* sp. ATCC 39727. *J. Ind. Microbiol. Biotechnol.* 37, 1097–1103. doi: 10.1007/s10295-010-0807-5

Myronovskyi, M., Welle, E., Fedorenko, V., and Luzhetskyy, A. (2011). β-Glucuronidase as a sensitive and versatile reporter in actinomycetes. *Appl. Environ. Microbiol.* 77, 5370–5383. doi: 10.1128/AEM.00434-11

Nazari, B., Forneris, C. C., Gibson, M. I., Moon, K., Schramma, K. R., and Seyedsayamdost, M. R. (2017). *Nonomuraea* sp. ATCC 55076 harbours the largest actinomycete chromosome to date and the kistamicin biosynthetic gene cluster. *MedChemComm* 8, 780–788. doi: 10.1039/C6MD00637J

Okano, A., Isley, N. A., and Boger, D. L. (2017). Peripheral modifications of [Ψ[CH_2NH]Tpg4]vancomycin with added synergistic mechanisms of action provide durable and potent antibiotics. *Proc. Natl. Acad. Sci.* 114, E5052–E5061. doi: 10.1073/pnas.1704125114

Peschke, M., Brieke, C., Goode, R. J. A., Schittenhelm, R. B., and Cryle, M. J. (2017). Chlorinated glycopeptide antibiotic peptide precursors improve cytochrome P450-catalyzed cyclization cascade efficiency. *Biochemistry* 56, 1239–1247. doi: 10.1021/acs.biochem.6b01102

Schoppet, M., Peschke, M., Kirchberg, A., Wiebach, V., Süssmuth, R. D., Stegmann, E., et al. (2019). The biosynthetic implications of late-stage condensation domain selectivity during glycopeptide antibiotic biosynthesis. *Chem. Sci.* 10, 118–133. doi: 10.1039/C8SC03530J

Shawky, R. M., Puk, O., Wietzorrek, A., Pelzer, S., Takano, E., Wohlleben, W., et al. (2007). The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains *bbr*, encoding a StrR-like pathway-specific regulator. *J. Mol. Microbiol. Biotechnol.* 13, 76–88. doi: 10.1159/000103599

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* 7, 1–6. doi: 10.1038/msb.2011.75

Spohn, M., Kirchner, N., Kulik, A., Jochim, A., Wolf, F., Muenzer, P., et al. (2014). Overproduction of ristomycin a by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. *Antimicrob. Agents Chemother.* 58, 6185–6196. doi: 10.1128/AAC.03512-14

Stegmann, E., Albersmeier, A., Spohn, M., Gert, H., Weber, T., Wohlleben, W., et al. (2014). Complete genome sequence of the actinobacterium *Amycolatopsis japonica* MG417-CF17T (=DSM 44213T) producing (S,S)-N,N′-ethylenediaminedisuccinic acid. *J. Biotechnol.* 189, 46–47. doi: 10.1016/j.jbiotec.2014.08.034

Stinchi, S., Azimonti, S., Donadio, S., and Sosio, M. (2003). A gene transfer system for the glycopeptide producer *Nonomuraea* sp. ATCC39727. *FEMS Microbiol. Lett.* 225, 53–57. doi: 10.1016/S0378-1097(03)00490-7

Stinchi, S., Carrano, L., Lazzarini, A., Feroggio, M., Grigoletto, A., Sosio, M., et al. (2006). A derivative of the glycopeptide A40926 produced by inactivation of the β-hydroxylase gene in *Nonomuraea* sp. ATCC39727. *FEMS Microbiol. Lett.* 256, 229–235. doi: 10.1111/j.1574-6968.2006.00120.x

Sungthong, R., and Nakaew, N. (2015). The genus *Nonomuraea*: a review of a rare actinomycete taxon for novel metabolites. *J. Basic Microbiol.* 55, 554–565. doi: 10.1002/jobm.201300691

Thaker, M. N., Wang, W., Spanogiannopoulos, P., Waglechner, N., King, A. M., Medina, R., et al. (2013). Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* 31, 922–927. doi: 10.1038/nbt.2685

Truman, A. W., Kwun, M. J., Cheng, J., Yang, S. H., Suh, J.-W., and Hong, H.-J. (2014). Antibiotic resistance mechanisms inform discovery: identification and characterization of a novel *Amycolatopsis* strain producing ristocetin. *Antimicrob. Agents Chemother.* 58, 5687–5695. doi: 10.1128/AAC.03349-14

Vigliotta, G., Tredici, S. M., Damiano, F., Montinaro, M. R., Pulimeno, R., Summa, R. D., et al. (2004). Natural merodiploidy involving duplicated *rpoB* alleles affects secondary metabolism in a producer actinomycete. *Mol. Microbiol.* 55, 396–412. doi: 10.1111/j.1365-2958.2004.04406.x

Vongsangnak, W., Figueiredo, L. F., Förster, J., Weber, T., Thykaer, J., Stegmann, E., et al. (2012). Genome-scale metabolic representation of *Amycolatopsis balhimycina*. *Biotechnol. Bioeng.* 109, 1798–1807. doi: 10.1002/bit.24436

Waglechner, N., Mcarthur, A. G., and Wright, G. D. (2019). Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat. Microbiol.* 4, 1862–1871. doi: 10.1038/s41564-019-0531-5

Wu, Z.-C., and Boger, D. L. (2019). Exploration of the site-specific nature and generalizability of a trimethylammonium salt modification on vancomycin: A-ring derivatives. *Tetrahedron* 75, 3160–3165. doi: 10.1016/j.tet.2019.02.008

Yim, G., Kalan, L., Koteva, K., Thaker, M. N., Waglechner, N., Tang, I., et al. (2014). Harnessing the synthetic capabilities of glycopeptide antibiotic tailoring enzymes: characterization of the UK-68,597 biosynthetic cluster. *ChemBioChem* 15, 2613–2623. doi: 10.1002/cbic.201402179

Yushchuk, O., Homoniuk, V., Datsiuk, Y., Ostash, B., Marinelli, F., and Fedorenko, V. (2020a). Development of a gene expression system for the uncommon actinomycete *Actinoplanes rectilineatus* NRRL B-16090. *J. Appl. Genet.* doi: 10.1007/s13353-019-00534-7 [Epub ahead of print].

Yushchuk, O., Homoniuk, V., Ostash, B., Marinelli, F., and Fedorenko, V. (2020b). Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. *J. Antibiot.* (in press). doi: 10.1038/s41429-019-0274-9

Yushchuk, O., Horbal, L., Ostash, B., Marinelli, F., Wohlleben, W., Stegmann, E., et al. (2019). Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. *Appl. Microbiol. Biotechnol.* 103, 4089–4102. doi: 10.1007/s00253-019-09789-w

Yushchuk, O., Ostash, B., Pham, T. H., Luzhetskyy, A., Fedorenko, V., Truman, A. W., et al. (2016). Characterization of the post-assembly line tailoring processes in teicoplanin biosynthesis. *ACS Chem. Biol.* 11, 2254–2264. doi: 10.1021/acschembio.6b00018

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Yushchuk, Andreo-Vidal, Marcone, Bibb, Marinelli and Binda. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Оригінал доступний за посиланням:
https://www.frontiersin.org/articles/10.3389/fmicb.2020.00008/full.
Електронний додаток до статті доступний за посиланням:
https://www.frontiersin.org/articles/10.3389/fmicb.2020.00008/full#supplementary-material.

Анотація: *Секвенування геномів показало, що Nonomuraea spp. залишаються все ще невивченим джерелом спеціалізованих метаболітів. Nonomuraea gerenzanensis ATCC 39727 є найбільш вивченим видом роду, оскільки є продуцентом глікопептидного антибіотика (ГПА) A40926 – попередника клінічно-значущого антибіотика далбаванцину, схваленого FDA у 2014 році для лікування гострих шкірних інфекцій, спричинених мультирезистентними Грам-позитивними збудниками. Клінічна значущість далбаванцину привернула увагу до вивчення механізмів біосинтезу A40926 та його регуляції. У цій статті ми досліджуємо, як покращити генетичний інструментарій для представників роду Nonomuraea, що виявилися досить складними об'єктами для генетичних маніпуляцій. Конструюючи вектори із різними промоторами, ми протестували активність 11 промоторів (гетерологічних і нативних) за допомогою репортерної системи GusA у N. gerenzanensis і Nonomuraea coxensis; цей останній вид є філогенетично далеким від N. gerenzanensis і також має генетичний потенціал продукувати A40926 або дуже схожий ГПА. Нарешті, найсильніший конститутивний промотор, проаналізований у цьому дослідженні, aac(3)IVp, використовувався для надекспресії розташованих у кластері регуляторних генів, що контролюють біосинтез A40926 (dbv3 і dbv4 з N. gerenzanensis і nocRI з N. coxensis) у N. gerenzanensis, а динаміку росту і продуктивність найкращих штамів оцінювали в масштабі біореактора з використанням промислового середовища. Надекспресія регуляторних генів привела до значного підвищення рівня продукції A40926 у N. gerenzanensis, забезпечивши новий, заснований на знаннях, підхід до покращення продуцента цього цінного ГПА.*

pubs.acs.org/acschemicalbiology

Articles

Genomic-Led Discovery of a Novel Glycopeptide Antibiotic by *Nonomuraea coxensis* DSM 45129

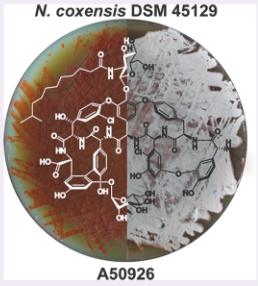
Oleksandr Yushchuk, Natalia M. Vior, Andres Andreo-Vidal, Francesca Berini, Christian Rückert, Tobias Busche, Elisa Binda, Jörn Kalinowski, Andrew W. Truman,* and Flavia Marinelli*

Cite This: *ACS Chem. Biol.* 2021, 16, 915−928

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Glycopeptide antibiotics (GPAs) are last defense line drugs against multidrug-resistant Gram-positive pathogens. Natural GPAs teicoplanin and vancomycin, as well as semisynthetic oritavancin, telavancin, and dalbavancin, are currently approved for clinical use. Although these antibiotics remain efficient, emergence of novel GPA-resistant pathogens is a question of time. Therefore, it is important to investigate the natural variety of GPAs coming from so-called "rare" actinobacteria. Herein we describe a novel GPA producer—*Nonomuraea coxensis* DSM 45129. Its *de novo* sequenced and completely assembled genome harbors a biosynthetic gene cluster (BGC) similar to the *dbv* BGC of A40926, the natural precursor to dalbavancin. The strain produces a novel GPA, which we propose is an A40926 analogue lacking the carboxyl group on the *N*-acylglucosamine moiety. This structural difference correlates with the absence of *dbv29*—coding for an enzyme responsible for the oxidation of the *N*-acylglucosamine moiety. Introduction of *dbv29* into *N. coxensis* led to A40926 production in this strain. Finally, we successfully applied *dbv3* and *dbv4* heterologous transcriptional regulators to trigger and improve A50926 production in *N. coxensis*, making them prospective tools for screening other *Nonomuraea* spp. for GPA production. Our work highlights genus *Nonomuraea* as a still untapped source of novel GPAs.



Downloaded via 178.95.0.215 on January 16, 2023 at 14:09:32 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

1. INTRODUCTION

Nonomuraea is a genus of so-called "rare" actinomycetes whose potential to produce specialized (secondary) metabolites is still rather poorly explored.[1,2] Recently sequenced genomes of *Nonomuraea* species appear to be generally larger than the reference *Streptomyces* ones. The mean genome size of *Nonomuraea* (based on the three available complete assemblies[2,3]) is around 12 Mbp, whereas the mean genome size of *Streptomyces* (calculated on 251 fully assembled genomes available in GenBank) equals 8.6 Mbp. The larger genomes of *Nonomuraea* spp. encode dozens of putative biosynthetic gene clusters (BGCs).[2−4] *Nonomuraea* spp. were initially found to be recalcitrant to commonly used genetic engineering manipulations, but new tools are now being developed for this genus.[5−7] This paves the way for unravelling the huge hidden biosynthetic potential of these organisms.

Probably the most important bioactive metabolite produced by a *Nonomuraea* species is the type IV[8] glycopeptide antibiotic (GPA) A40926[9] (Figure 1) produced by *Nonomuraea gerenzanensis* ATCC 39727. Like other GPAs, A40926 acts as a selective and potent inhibitor of cell-wall biosynthesis in Gram-positive bacteria. A40926 is structurally related to the clinically relevant GPA teicoplanin (Figure 1), produced by *Actinoplanes teichomyceticus* ATCC 31121[10,11] and to ristocetin (Figure 1), previously isolated from numerous *Amycolatopsis* spp. (i.e., *A. lurida* NRRL 2430, *A. japonicum* MG417-CF17, and *Amycolatopsis* sp. MJM2582).[12−14] Like teicoplanin, A40926 is produced as a mixture of related compounds (major components are A40926 B and A40926 A factors), which differ in the length and branching of an aliphatic side chain (Figure 1). It was recently clarified that *N. gerenzanensis* produces the GPA in the form of *O*-acetyl-A40926 (with an *O*-acetylated mannose residue), but the acetyl group is lost during the alkaline extraction of the antibiotic.[15,16] Since it was this deacetylated GPA that was initially named A40926, we will refer to it as A40926 hereafter.

A40926 is the precursor of the second-generation semisynthetic GPA dalbavancin (Figure 1), which is currently applied in clinics to treat severe infections caused by multidrug-resistant Gram-positive pathogens.[17] Dalbavancin (marketed in Europe and USA under the trade names xydalba and dalvance, respectively) is the first antibiotic designated as a qualified infectious disease product by FDA because of its potency, extended dosing interval, and unique dose regimen (once-a-week), but its cost still largely exceeds that of first-generation GPAs.[10] Therefore, improvement of A40926 production by recombinant engineering of *N. gerenzanensis*

Received: March 8, 2021
Accepted: April 14, 2021
Published: April 29, 2021



ACS Publications

© 2021 The Authors. Published by American Chemical Society

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

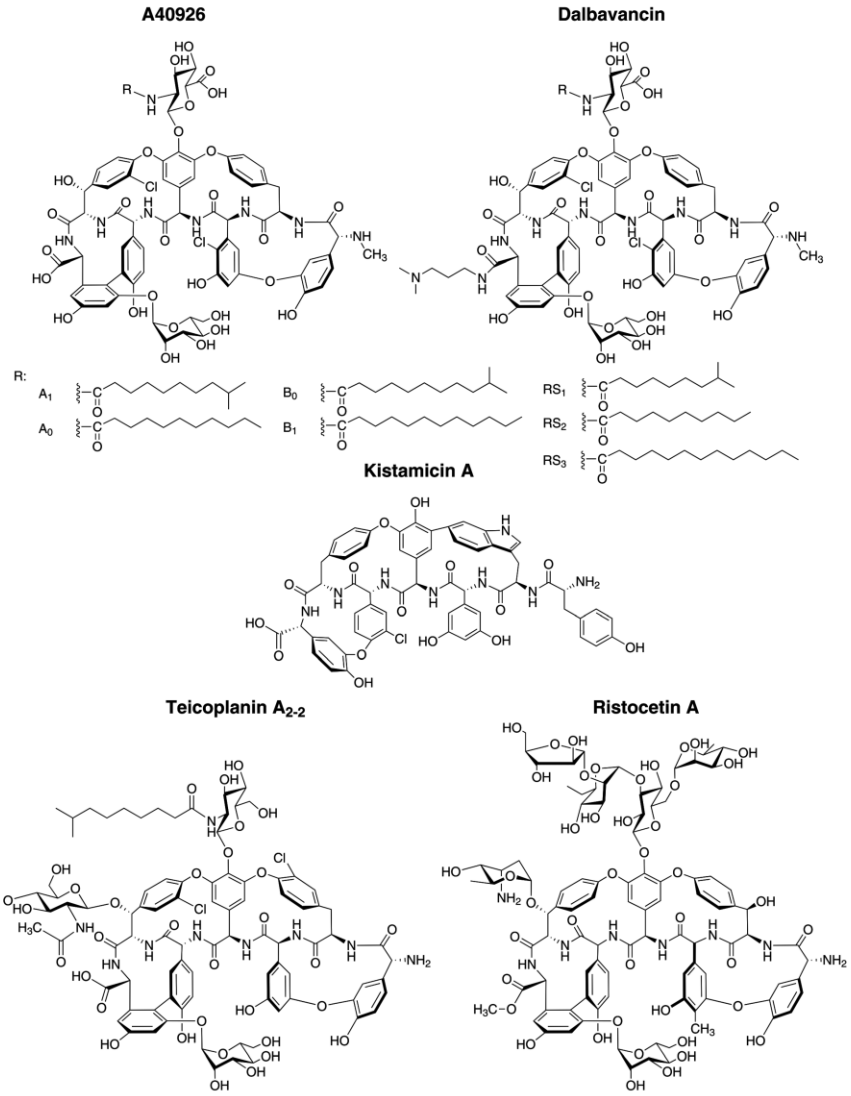


Figure 1. Structures of the GPAs found in genus *Nonomuraea*: type IV A40926 and type V kistamicin. Clinically used dalbavancin is obtained from A40926 by conversion of the C-terminal carboxyl group into a (3-dimethylamino)-1-propylamide. Type IV teicoplanin and type III ristocetin are shown due to their structural similarities with A40926. For teicoplanin, the main factor (TA_{2-2}) of the complex produced by *A. teichomyceticus* is shown, the other factors are differing by the length and branching of the lipid chain. Ristocetin is produced by numerous *Amycolatopsis* species.

has become increasingly relevant.[6,16] Following the sequencing of the A40926 BGC (*dbv*) almost two decades ago,[18] multiple aspects of A40926 biosynthesis were investigated, including nonribosomal aglycone assembly and tailoring steps,[15,19,20] self-resistance,[21,22] and pathway-specific regulation of its production.[6,23,24] *N. gerenzanensis* was also engineered to produce A40926 derivatives that are better suited for downstream chemical modification to dalbavancin.[16] Another GPA produced by a *Nonomuraea* species is the type V GPA kistamicin (Figure 1) from *Nonomuraea* sp. ATCC 55076, which was reported to exhibit potent antiviral activity as well as mild antibiosis against Gram-positive bacteria.[2,25] Its structure contains an unusual indole−phenol cross-link which makes this GPA unique among those already known.[5,26]

Genome mining has recently shown that other species from the genus *Nonomuraea* also possess BGCs for GPAs,[27] as in the cases of *Nonomuraea* sp. WAC 01424 and *Nonomuraea coxensis* DSM 45129. Notwithstanding the low quality of the available

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

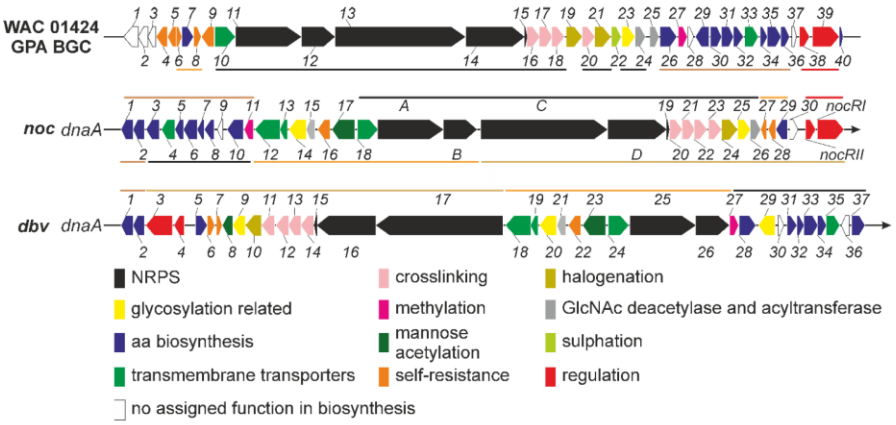


Figure 2. Comparison of BGCs from *N. gerenzanensis* (*dbv*), *N. coxensis* (*noc*), and *Nonomuraea* sp. WAC 01424. Colored lines indicate the homology segments among the BGCs. For *dbv* and *noc*, the orientation of the BGC genes is in relation to the orientation of the *dnaA* gene of the chromosome. This orientation was not possible for the WAC 01424 GPA BGC, since the corresponding genome is fragmented across multiple contigs. Details on gene function and homology are given in Table 1 and in the main text.

draft genomic data, we recently showed that *N. coxensis* DSM 45129 carries a BGC remarkably similar to *dbv*.[6] We found that this BGC contains a putative regulatory gene orthologous to *dbv3*, which encodes the pathway-specific regulator of LuxR-type in *N. gerenzanensis*.[6] The heterologous expression of this gene from *N. coxensis* (named *nocRI*) led to A40926 overproduction in *N. gerenzanensis*, indicating that it might be functional in *N. coxensis* as well. Thus, in this paper we present the fully assembled genome of *N. coxensis*, which has allowed us to properly describe the putative GPA BGC (called *noc*). Additionally, we report that *N. coxensis* produces a novel GPA complex, which we named A50926. Structural characterization of A50926 by liquid chromatography−mass spectrometry (LC-MS) and tandem MS (MS/MS) showed it has high similarity to A40926, although A50926 lacks the carboxyl group on the *N*-acylglucosamine (GlcN-Acyl) moiety. Consistently, the *noc* BGC lacks an orthologue of *dbv29*, which in *N. gerenzanensis* encodes the enzyme oxidizing the GlcN-Acyl moiety to an *N*-acylaminoglucuronic group.[19] Introduction of *dbv29* into *N. coxensis* changed the GPA production profile of this strain to A40926. Finally, we have introduced *dbv3* and *dbv4* pathway-specific regulatory genes in *N. coxensis* to trigger and overproduce A50926 by regulatory gene cross-talking. In conclusion, our results describe the biosynthesis of a novel GPA, which may have superior properties to A40926[28] and thus may contribute to developing a platform for the combinatorial biosynthesis of third generation lipo-GPAs.

2. RESULTS AND DISCUSSION

2.1. Complete Assembly of *N. coxensis* Genome Reveals the Presence of a Novel GPA BGC. The presence of a novel GPA BGC in the genome of *N. coxensis* was recently anticipated.[6,27] However, due to the poor quality of the available draft, fragments of the BGC were found on different contigs and did not cover the full expected sequence of the BGC. Therefore, we sequenced and fully assembled the genome of *N. coxensis* DSM 45129 using a combination of HiSeq Illumina and GridION ONT technologies. The circular chromosome of *N. coxensis* was found to have a smaller size in comparison to the other two previously published *Nonomuraea* genomes—only 9.07 Mbp compared to 11.85 Mbp in *N. gerenzanensis*[3] and 13.05 Mbp in *Nonomuraea* sp. ATCC 55076.[2] The average GC-content was 71.8%. Annotation of the *N. coxensis* genome revealed 8398 predicted protein coding sequences, five operons for 16S-23S-5S rRNA, and 73 tRNA genes. Genome analysis by antiSMASH 5.0,[29] a specialized metabolite BGC identification tool, led to the discovery of 27 putative BGCs when used in the "relaxed" search mode. However, only a few BGCs showed more than 20% similarity to known BGCs (Table S1).

We thus focused our attention on the GPA-like BGC, which we denoted as *noc* (from ***N****onomuraeacoxensis*). The *noc* BGC is the fourth GPA BGC described from *Nonomuraea* genus, following the *dbv* BGC from *N. gerenzanensis*,[18] a putative GPA BGC from *Nonomuraea* sp. WAC 01424[27] and the type V GPA kistamicin (*kis*) BGC from *Nonomuraea* sp. ATCC 55076.[2] Overall, *noc* contains 36 open reading frames (ORFs) with 35 among them homologous to *dbv* genes (the nonhomologous *noc* gene encoding for a putative transposase) and 32 being homologous to genes in the *Nonomuraea* sp. WAC 01424 GPA BGC (Figure 2, Table 1). The *kis* BGC differed from *noc* most significantly (data not shown).

2.2. Comparative Genomics of *Nonomuraea* GPA Producers. At the time of writing, genomic information for 34 *Nonomuraea* species was available in GenBank, although there are only three complete assemblies (Table S2). Along with the four reported *Nonomuraea* GPA BGCs, we found a *kis*-like BGC in the draft genome of *Nonomuraea* sp. NN258 (Figure S1). We have then reconstructed the multilocus phylogeny (MLP) of all *Nonomuraea* species with available genomic data using conserved house-keeping proteins (Table S3). It revealed *N. coxensis* to be most closely related to *N. wenchangensis* CGMCC 4.5598, *N. polychroma* DSM 43925, and *N. turkmeniaca* DSM 43926 (Figure S2). None of these species have GPA BGCs in their genomes. *N. gerenzanensis* is most closely related to *Nonomuraea* sp. FMUSA5−5 and to the kistamicin producer *Nonomuraea* sp. ATCC 55076, whereas

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

Table 1. Characterization of *noc* BGC genes and their comparison to the *dbv* and WAC 01424 GPA BGCs from *Nonomuraea* spp.

noc BGC genes	homologues from *dbv* BGC (aa identity of protein product with *noc* homologue, %)	homologues from WAC 01424 GPA BGC (numbered as in Figure 2) (aa identity of protein product with *noc* homologue, %)	encoded protein
noc1	*dbv1* (90.6%)	*DMB42_RS42735* (31) (60%)	hydroxymandelate oxidase (Hmo)
noc2	*dbv2* (89.3%)	*DMB42_RS42740* (30) (62%)	hydroxymandelate synthase (HmaS)
noc3	*dbv37* (90.9%)	*DMB42_RS42745* (29) (83%)	hydroxyphenylglycine aminotransferase (HpgT)
noc4	*dbv35* (90.9%)	*DMB42_RS42730* (32) (63%)	Na^+−H^+ antiporter
noc5	*dbv34* (93.9%)	*DMB42_RS42710* (36) (87%)	enoyl-CoA hydratase (DpgD)
noc6	*dbv33* (89.2%)	*DMB42_RS42715* (35) (84%)	dihydroxyphenylacetyl-CoA dioxygenase (DpgC)
noc7	*dbv32* (85.1%)	*DMB42_RS42720* (34) (74%)	enoyl-CoA hydratase (DpgB)
noc8	*dbv31* (94.3%)	*DMB42_RS42725* (33) (91%)	type III polyketide synthase (DpgA)
noc9	*dbv30* (83.5%)	*DMB42_RS42750* (28) (69%)	4HB-CoA thioesterase
noc10	*dbv28* (92.4%)	*DMB42_RS42760* (26) (86%)	β-hydroxylase
noc11	*dbv27* (91.8%)	*DMB42_RS42755* (27) (58%)	methyltransferase
noc12	*dbv18* (87.3%)	[a]	ABC transporter
noc13	*dbv19* (92.2%)	[a]	ABC transporter
noc14	*dbv20* (89.7%)	[a]	mannosyltransferase
noc15	*dbv21* (86.6%)	*DMB42_RS42765* (25) (64%)	deacetylase
noc16	*dbv22* (92.3%)	*DMB42_RS42850* (9) (77%)	sensory histidine kinase
noc17	*dbv23* (88.1%)	[a]	acetyltransferase
noc18	*dbv24* (92.4%)	*DMB42_RS42845* (10) (81%)	ABC transporter
nocA	*dbv25* (88.7%)	*DMB42_RS42840* (11) (76%)	NRPS modules 1−2
nocB	*dbv26* (91%)	*DMB42_RS42835* (12) (78%)	NRPS module 3
nocC	*dbv17* (89.6%)	*DMB42_RS42830* (13) (77%)	NRPS modules 4−5−6
nocD	*dbv16* (91.7%)	*DMB42_RS42825* (14) (79%)	NRPS module 7
noc19	*dbv15* (94.2%)	*DMB42_RS42820* (15) (93%)	MbtH-like protein
noc20	*dbv14* (91.8%)	*DMB42_RS42815* (16) (78%)	cross-linking oxygenase (OxyA)
noc21	*dbv13* (89.8%)	*DMB42_RS42810* (17) (77%)	cross-linking oxygenase (OxyC)
noc22	*dbv12* (93.5%)	*DMB42_RS42805* (18) (77%)	cross-linking oxygenase (OxyB)
noc23	*dbv11* (91.9%)	*DMB42_RS42795* (20) (78%)	cross-linking oxygenase (OxyE)
noc24	*dbv10* (94.1%)	*DMB42_RS42790* (21) (87%)	halogenase
noc25	*dbv9* (90.4%)	*DMB42_RS42780* (23) (74%)	glycosyltransferase (GtfB)
noc26	*dbv8* (87.5%)	*DMB42_RS42775* (24) (77%)	acyltransferase
noc27	*dbv7* (87.3%)	*DMB42_RS42865* (6) (78%)	VanY-carboxypeptidase
noc28	*dbv6* (95.9%)	*DMB42_RS42855* (8) (92%)	response regulator
noc29	*dbv5* (92.8%)	*DMB42_RS42860* (7) (85%)	prephenate dehydrogenase (Pdh)
noc30	[a]	[a]	putative transposase
nocRII	*dbv4* (94.4%)	*DMB42_RS42700* (38) (85%)	StrR-like transcriptional regulator
nocRI	*dbv3* (86.3%)	*DMB42_RS42695* (39) (70%)	LuxR-like transcriptional regulator

[a]Homologue is absent.

Nonomuraea sp. WAC 01424 is distantly related to both *N. coxensis* and *N. gerenzanensis* (Figure S2). Thus, GPA-producing *Nonomuraea* species do not form a single phylogenetic group, which is different from what occurs in the majority of *Amycolatopsis* spp. producing GPAs.[30]

Since *N. gerenzanensis* and *Nonomuraea* sp. ATCC 55076 are closely related and their genomes had been completely assembled, we compared their sequences using the MAUVE genome alignment tool.[31] We found that the two genomes are very similar, having few rearranged homologous segments (Figure S3A). Interestingly, the regions flanking the *dbv* BGC in *N. gerenzanensis* show synteny in *Nonomuraea* sp. ATCC 55076, but in this genome, they flank a miscellaneous assemblage of GPA-unrelated genes instead of the *dbv* genes. No *dbv*-like BGC is present in *Nonomuraea* sp. ATCC 55076. Similarly, no *kis*-like BGC is in the *N. gerenzanensis* genome, but the regions flanking the *kis* BGC in *Nonomuraea* sp. ATCC 55076 have their homologous counterparts in the *N.*

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

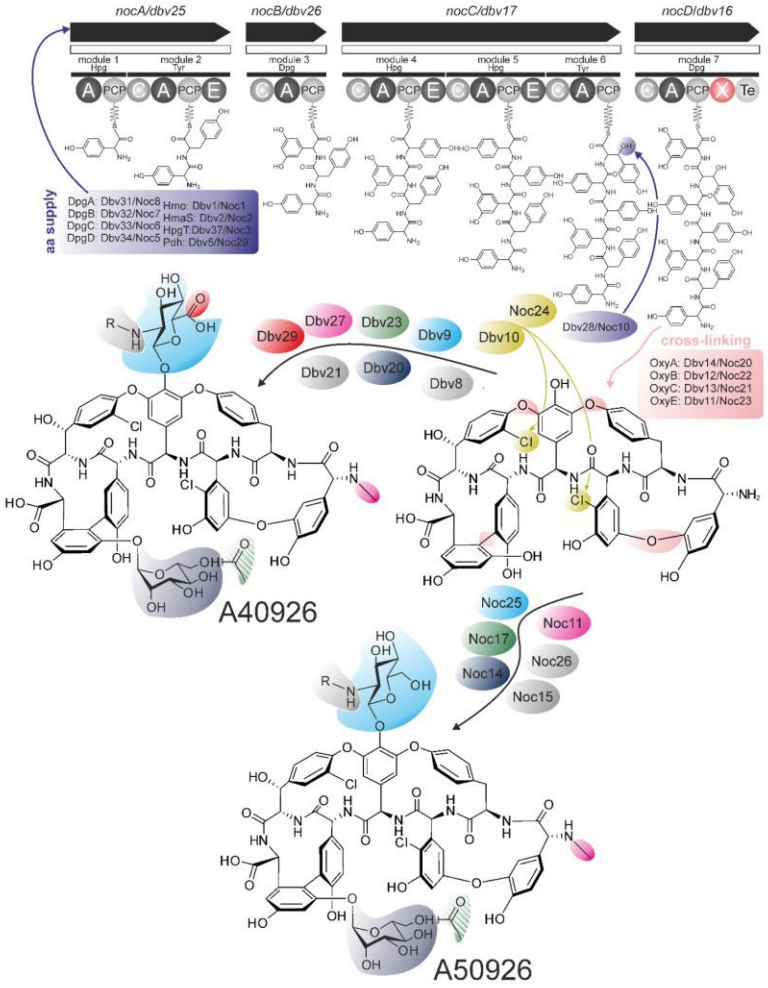


Figure 3. Conceptual scheme of the biosynthesis of A40926 and of the GPA (named A50926) from *N. coxensis*. Please note the dashed acyl group at the mannose residue, which is installed by Dbv23/Noc17 but consequentially lost during antibiotic extraction. For more details and encoded protein names, please refer to the main text and Table 1.

gerenzanensis genome (Figure S3A). Dot plots of *N. gerenzanensis* and *Nonomuraea* sp. ATCC 55076 confirm the high homology between the two strains (Figure S3B). A possible explanation is that *Nonomuraea* sp. ATCC 55076 and *N. gerenzanensis* genomes might have acquired different GPA BGCs independently through horizontal gene transfer (HGT) events from other *Nonomuraea* (or not) species.

Dot plots of *N. coxensis* and *Nonomuraea* sp. ATCC 55076 genomes (Figure S3C) as well as of *N. coxensis* and *N. gerenzanensis* (Figure S3D) indicate that *N. coxensis* is more distantly related to the other GPA producing species. Unfortunately, it was impossible to compare the genome of *N. coxensis* with its closest relatives *N. wenchangensis* CGMCC 4.5598, *N. polychroma* DSM 43925, and *N. turkmeniaca* DSM 43926 (Figure S2), due to the incompleteness of their genome assemblies. Overall, it seems that the position of GPA BGCs is not conserved within *Nonomuraea* genomes, which contrasts to what was observed in most *Amycolatopsis* spp.[30]

2.3. Comparing *noc* and *dbv* Biosynthetic Pathways: From Genes to Products. The biosynthesis of A40926 is well understood (Figure 3). The heptapeptide core of this antibiotic is synthesized by a nonribosomal peptide synthetase (NRPS) assembly line involving Dbv25, Dbv26, Dbv17, and Dbv16 proteins. The linear peptide is cross-linked by four monooxygenases (Dbv14, Dbv12, Dbv13, and Dbv11) and halogenated by Dbv10, giving the core aglycone. This aglycone is further modified with the glycosyltransferases Dbv9 and Dbv20, which attach *N*-acetyl glucosamine (GlcNAc) and mannose, respectively.[32] Then, GlcNAc is oxidized by Dbv29, deacetylated by Dbv21, and acylated by Dbv8. Finally, the

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

mannose moiety is acetylated by Dbv23, giving *O*-acetyl-A40926.

Considering the A40926 pathway, it was possible to predict the biosynthetic pathway of the putative GPA from *N. coxensis* (Figure 3). Sets of genes required for the biosynthesis of the nonproteinogenic precursor amino acids 4-hydroxyphenylglycine (Hpg), 3,5-dihydroxyphenylglycine (Dpg), and β-hydroxytyrosine (further used as substrates for NRPS) are the same in *noc* and *dbv* BGCs (Table 1, Figures 2 and 3). Next, the NRPS, encoded within *noc* BGC, was found to have the same organization and A-domain specificities as the *dbv* NRPS (Figure S4, Table S4). All other genes, responsible for the cross-linking and tailoring steps, were identical in both the *noc* and *dbv* pathways (Table 1, Figures 2 and 3). However, one notable difference between *dbv* and *noc* was the absence of a *dbv29* orthologue in the latter. As mentioned above, Dbv29 is a hexose oxidase responsible for the oxidation of the Glc*N*-Acyl moiety of A40926.[19] On this basis, we predicted that the *noc* pathway might produce an A40926 analogue lacking the carboxylic group on the Glc*N*-Acyl residue and therefore resembling teicoplanin in this moiety (Figures 1 and 3).

Beyond the biosynthetic genes, *noc* and *dbv* feature homologous regulatory genes. Two master regulators of A40926 biosynthesis—LuxR-like Dbv3 and StrR-like Dbv4—have orthologues coded within *noc*—NocRI (94% aa sequence identity) and NocRII (86% aa sequence identity), respectively.[6] In *N. gerenzanensis*, both Dbv3 and Dbv4 are crucial for biosynthesis activation.[23] Dbv4 was shown to bind the promoter regions of operons *dbv30-35* (mainly coding for Dpg biosynthesis enzymes) and *dbv14-8* (including the genes coding for cross-linking monooxygenases), and its binding sites were identified.[33] Our *in silico* analysis indicates that identical binding sites are present in the promoter regions of *noc20* and *noc8*, orthologues of *dbv14* and *dbv30*, respectively (Figure S5). DNA-binding sites of Dbv3 remain uncharacterized, but its regulon was defined from gene expression analysis and includes other biosynthetic genes and Dbv4.[23] Given all these similarities, we presume that NocRI/NocRII have functions identical to Dbv3/Dbv4 and both regulatory pairs might cross-talk between these species. Our previous results,[6] where heterologous expression of *nocRI* in *N. gerenzanensis* improved A40926 production, support this assumption. The single GPA resistance determinant encoded within *noc* is Noc27, a close (87%) orthologue of Dbv7 (VanY$_n$), which is a D,D-carboxypeptidase involved in A40926 self-resistance.[21,34]

Although the biosynthetic, regulatory, and resistance genes are apparently shared by the *dbv* and *noc* BGCs, their genetic organization is different. So far, almost all GPA BGCs have NRPS genes located on one strand in an order that is colinear to the order of the modules in the NRPS assembly line. The only exception is the *dbv* BGC, where the NRPS genes are coded on different strands and are separated by other biosynthetic genes.[18] The *noc* BGC, although sharing a remarkable similarity with *dbv*, features an organization of NRPS genes that is typical of all the other GPAs. Interestingly, only two chromosomal inversion events are needed to rearrange *noc* into *dbv* (Figure S6), indicating how a *dbv*-like gene arrangement might have derived from a *noc*-like BGC in a common ancestor of *N. coxensis* and *N. gerenzanensis* (or in an ancestralprotocluster).

The putative GPA BGC in *Nonomuraea* sp. WAC 01424 (Figure 2) differs more substantially from both *noc* and *dbv*. It lacks a *noc14/dbv20* homologue encoding for a mannosyltransferase, as well as a *noc17/dbv23* homologue encoding for a mannose-*O*-acetyltransferase (Table 1). Instead, WAC 01424 GPA BGC contains a close homologue of *staL* (Figure S7), which encodes for a sulfotransferase involved in the biosynthesis of A47934 from *Streptomyces toyocaensis* NRRL 15009.[35] Additionally, the WAC 01424 BGC-encoded halogenases seem more related to the ones from the A47934 BGC than to Noc24 and Dbv8 (Figure S7). Thus, we suggest that WAC 01424 GPA is a nonmannosylated, but sulfated, A40926 analogue, putatively with a halogenation pattern different from A40926 (Figure S8).

2.4. Optimization of GPA-Producing Conditions for *N. coxensis*. *N. coxensis* was first described in 2007,[36] but as far as we know, it was never tested for the production of antimicrobials. Considering the predicted similarity between the putative GPA produced by *N. coxensis* with A40926, we first applied to *N. coxensis* the cultivation and A40926 production conditions that we had previously optimized for *N. gerenzanensis*.[22,37] In these conditions (namely a vegetative preculture in E26 medium and a GPA production step in FM2 medium using baffled flasks), *N. coxensis* tended to grow poorly, and no antimicrobial activity was detectable throughout the 168h cultivation from inoculum. Thus, we further screened different media and fermentation conditions previously used for growing other GPA producing strains, such as TM1 used for teicoplanin production by *A. teichomyceticus*[38] and R5 adopted for balhimycin production in *Amycolatopsis balhimycina*,[39] as well as VM0.1 and ISP2l previously employed for the vegetative cultivation of *N. coxensis*[6] (media composition detailed in the Supporting Information). The production of antimicrobial activity toward *Bacillus subtilis* ATCC 6633 was observed only in TM1 and ISP2l media when glass beads were added to favor dispersed growth (Figure S9). Indeed, adding glass beads to E26 medium cultures allowed us to use it for a successful vegetative preculture step (Figure S10A). Interestingly, routine analysis of glucose consumption in all media described above indicated that *N. coxensis* did not visibly consume glucose during growth (data not shown). We thus tested the glucose-lacking E26 (named E27), TM1 (TM1m), and ISP2l (ISP2lm) media variants for *N. coxensis* growth and putative GPA production. We found that biomass accumulation was similar in E26 and E27 (Figure S10A) and that biomass and antimicrobial production were equivalent in TM1 and ISP2l as well as in their glucose lacking variants TM1m and ISP2lm (Figure S10B and C). Currently, it is impossible to say why *N. coxensis* fails to use glucose throughout cultivation given that all necessary genes are present within its genome (Figure S10D). Thus, for all the following work with *N. coxensis*, E27, TM1m, and ISP2lm were used.

2.5. Expression of VanY-like Activity in *N. coxensis*. As already mentioned, the *noc* BGC encodes a Dbv7 orthologue—Noc27. We therefore tested whether D,D-carboxypeptidase activity could be detected in GPA-producing cultures of *N. coxensis*. This was measured in membrane extracts as previously reported for *N. gerenzanensis* and its mutant strains.[22] D,D-carboxypeptidase activity was measurable in *N. coxensis* extracts, although at an inferior level than in *N. gerenzanensis* (Figure S11). This indicated that Noc27 is functional and its expression correlates with the antimicrobial producing conditions. These results corroborate the hypothesis that *noc* genes are expressed and a novel GPA active versus *B. subtilis* is produced by *N. coxensis*. As in *dbv*[40] and WAC 01424

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

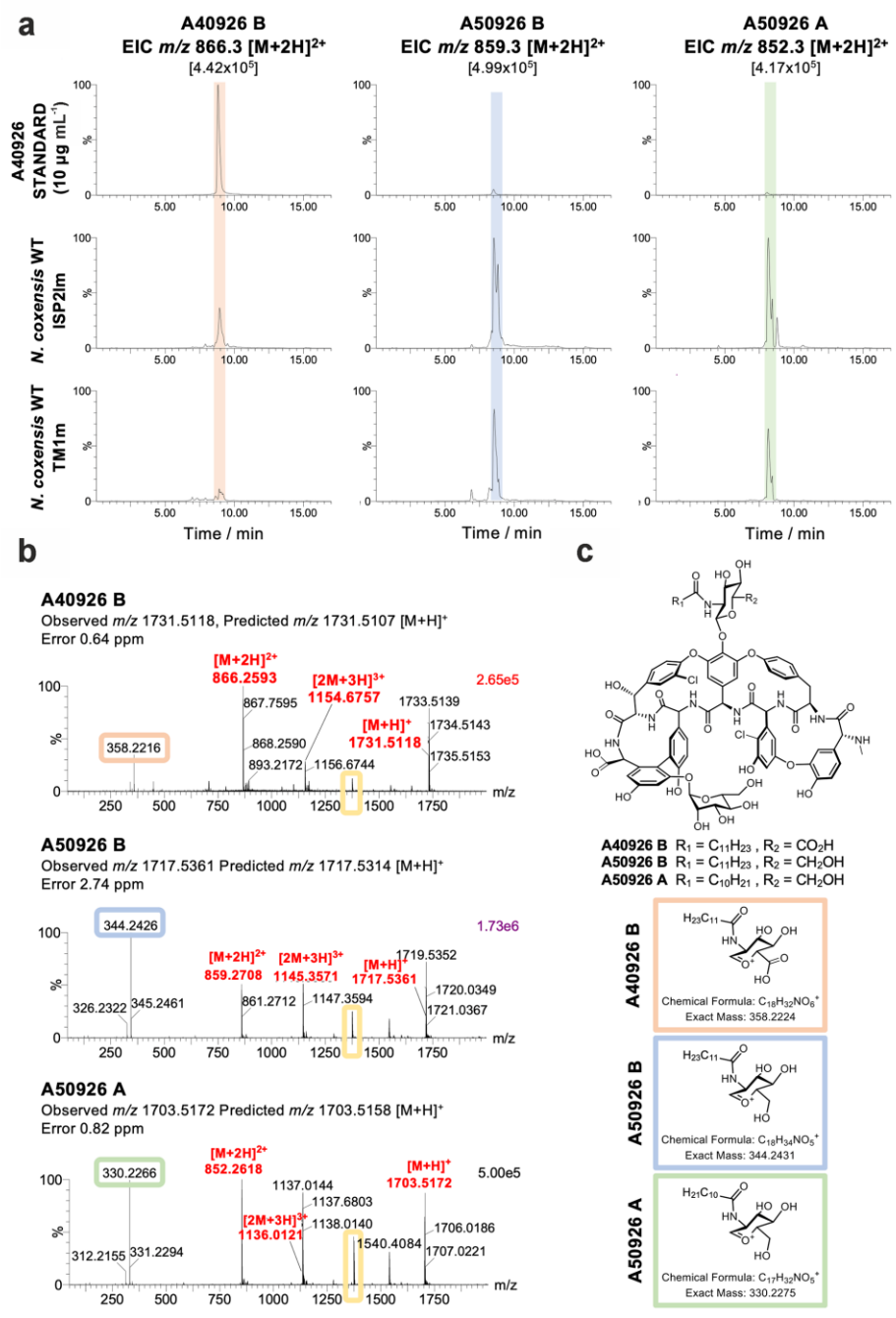


Figure 4. MS characterization of novel GPA complex produced by wild type *N. coxensis* grown in ISP2lm and TM1m media for 7 days. (a) Extracted ion chromatograms (EICs) of masses corresponding to A40926 B (left column) and the major components of the A50926 complex produced by *N. coxensis* WT, A50926 B (m/z 859.3, second column), and A50926 A (m/z 852.3, third column). The top row corresponds to a

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

Figure 4. continued

commercial standard of A40926 and the middle and bottom rows to culture extracts from ISP2lm and TM1m, respectively. For each mass, peak heights are normalized relative to the intensity of the largest peak in the sample set, shown in brackets at the top of each column. (b) MS spectra for A40926 B, A50926 B, and A50926 A. Peak heights are normalized to the intensity of the top peak in each spectrum, shown on the top right corner of each plot. Signature in-source fragments for each of the analyzed molecules are circled in pink, blue, and green, respectively, whereas the fragment corresponding to the mannosylated aglycone common to all of them is highlighted in yellow. (c) Proposed structure for the A50926 molecules. The top schematic represents a generic proposed structure common to A40926 and A50926 while the insets below represent the differential fragments for each of the analyzed molecules, as inferred from MS and MS/MS data.

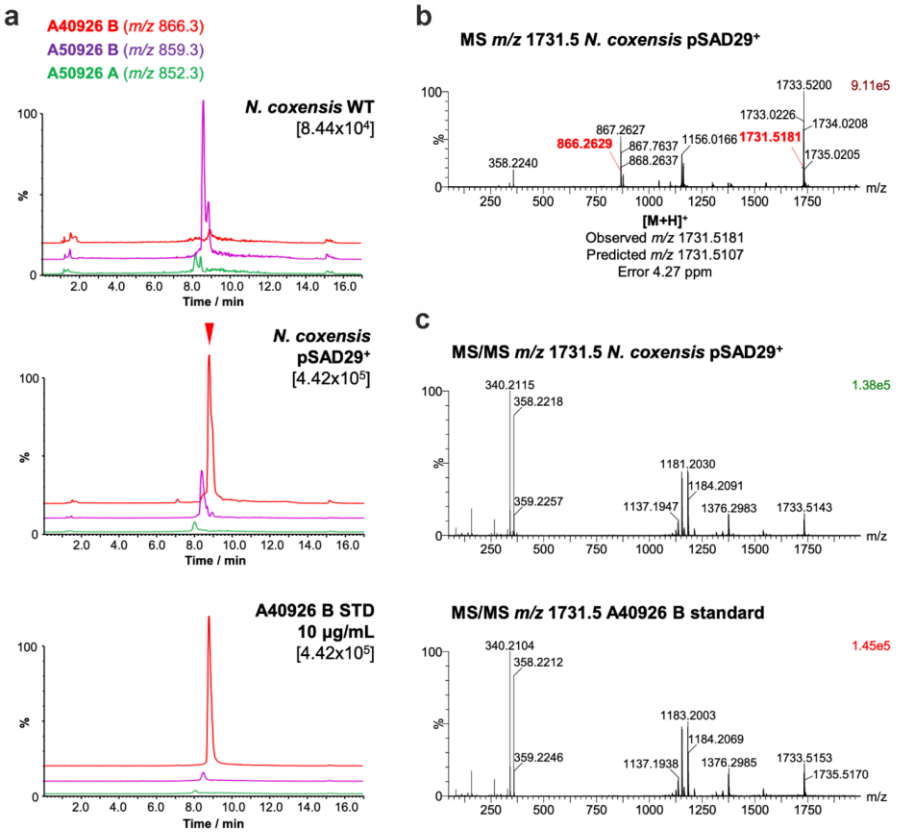


Figure 5. Production of A40926 in *N. coxensis* pSAD29⁺ grown in ISP2lm for 7 days. (a) EICs for masses corresponding to A40926 B (red trace), A50926 B (purple), and A50926 A (green) in purified extracts of *N. coxensis* pSAD29⁺ (top chromatogram) and *N. coxensis* WT (middle) in comparison to an A40926 commercial standard. The intensity for the top peak in each chromatogram is shown in brackets under the sample name. (b) MS spectrum of A40926 B from *N. coxensis* pSAD29⁺ cultures. Monoisotopic masses corresponding to $[M + 2H]^{2+}$ and $[M + H]^{+}$ adducts are highlighted in red, and the deviation between the observed accurate mass and the predicted mass for A40926 is represented in parts per million. (c) MS/MS spectra of A40926 B produced by *N. coxensis* pSAD29⁺ and an A40926 B commercial standard.

BGCs, a *vanY* gene seems to be the only cluster-situated determinant of self-resistance in *N. coxensis*.

2.6. Purification and Identification of the Novel Glycopeptide Complex Produced by *N. coxensis*. D-Alanine-D-Alanine (D-Ala-D-Ala) affinity resin chromatography was used to capture the putative GPA from cultures of *N. coxensis* grown in ISP2lm and TM1m media. ISP2lm appeared to be the most suitable medium for GPA purification, since the rich composition and high viscosity of TM1m interfered with affinity chromatography. Analyzed by HPLC, the affinity resin eluates contained two major peaks with the characteristic UV spectra of the commercially available A40926 standard, but with a different retention time (Figure S12). LC-MS analysis of these peaks revealed they corresponded to ions with m/z 852.3 and 859.3 ($[M + 2H]^{2+}$), 28 and 14 Da smaller respectively than an A40926 standard ($[M + 2H]^{2+}$ = 866.3, corresponding to A40926 B). We therefore tentatively named this new GPA complex A50926 (Figure 4a and c).

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

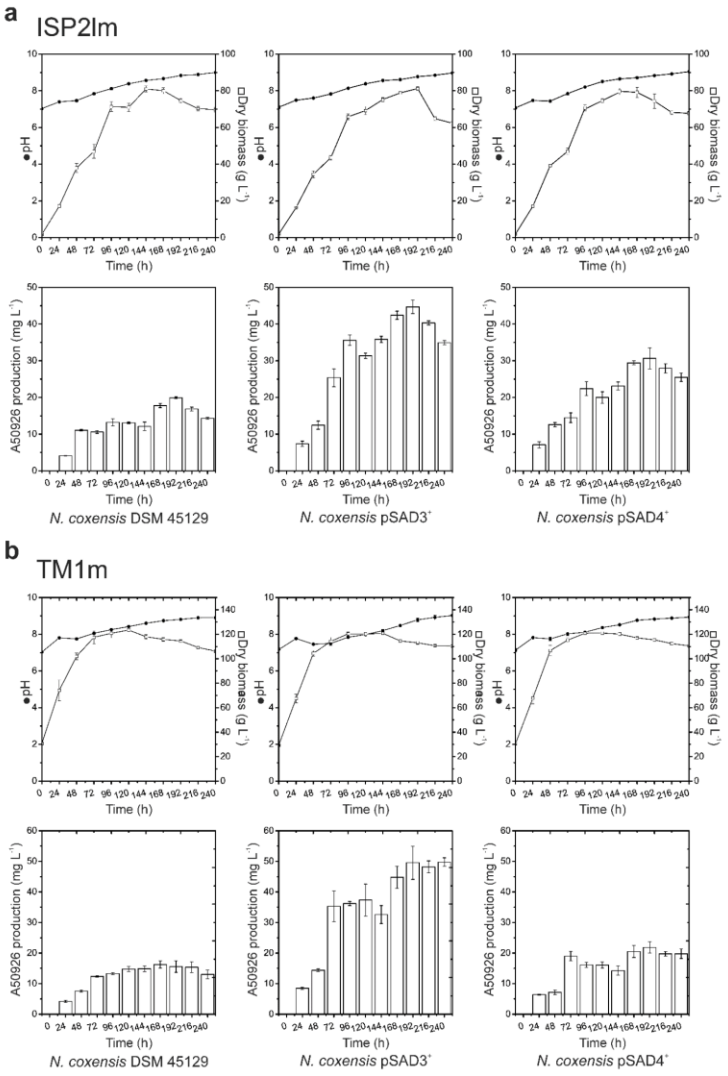


Figure 6. Time courses of *N. coxensis* wild type and of the recombinant strains overexpressing *dbv3* (pSAD3$^+$) and *dbv4* (pSAD4$^+$) cultivated in ISP2lm (a) or TM1m (b) in 500 mL Erlenmeyer flasks. pH (filled circles), biomass accumulation (empty squares), and A50926 production were monitored every 24 h. Results given are mean values of three independent experiments, and error bars represent standard deviations.

All three molecules showed similar MS spectra with single, double, and triple charge proton adducts as well as in-source fragments corresponding to the aglycone carrying the mannose moiety and the Glc*N*-Acyl moiety (Figure 4b,c). The mannosylated aglycone fragment (m/z 1374.3) was common to all three peaks (Figures 4b and S13), indicating that they share the same aglycone structure and mannose decoration. In contrast, the in-source fragment corresponding to the acylated sugar carried the signature mass difference for each molecule (Figures 4b, S14, and S15): the main A50926 peak ($[M + H]^+$ = 1717.5361) had a fragment with m/z 344.2, whereas the A40926 standard had a fragment with m/z 358.22 (Figures 4b, S14, and S15). Further MS and MS/MS analyses of these fragments (Figures S14 and S16) allowed us to assign this 14 Da mass difference to the glucosamine moiety. The masses are consistent with this sugar featuring a regular 6-hydroxyl group in A50926 versus being carboxylated in A40926 (Figures 1, 3, 4c, S14, S15, and S16). This correlates with the lack of a homologue of *dbv29* in the *noc* BGC, as it encodes the enzyme responsible for the oxidation of the C-6 hydroxyl group of Glc*N*-Acyl into a carboxylic acid in A40926. The second A50926 peak ($[M + H]^+$ = 1703.5172) had a further 14 Da

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

mass difference in the GlcN-Acyl moiety (Figures 4b and S14), but in this case MS/MS showed this difference to be in the acyl chain (Figure S16), which is consistent with an A50926 congener with a C11 acyl chain instead of a C12 acyl chain. This is equivalent to the A and B series of congeners in the A40926 complex.[41] Based on this analysis and accurate mass data (Figure 4b), we named the compound with $[M + H]^+$ = 1717.54 A50926 B (Figure 4c) and the compound with $[M + H]^+$ = 1703.52 A50926 A (Figure 4c).

2.7. Single Gene Expression Leads to A40926 Production in *N. coxensis*. To support our MS-based characterization of A50926, we hypothesized that we could convert *N. coxensis* into an A40926 producer by overexpression of the *dbv29* gene from *N. gerenzanensis*, which encodes the hexose oxidase required for oxidation of the C-6 hydroxyl group of GlcN-Acyl into the corresponding carboxylic acid. To achieve this, we used the pSET152A expression platform, which has proven to be very effective for gene overexpression in both *N. coxensis* and *N. gerenzanensis*.[6] *dbv29* was cloned into pSET152A to generate pSAD29, which was then introduced into *N. coxensis* by conjugation from *Escherichia coli*. *N. coxensis* pSAD29$^+$ was grown in ISP2lm medium for 168 h, and the resulting GPA complex was purified using D-Ala-D-Ala affinity resin. LC-MS analysis determined that *N. coxensis* pSAD29$^+$ was able to produce a molecule with an identical retention time and MS spectrum to that of A40926 (observed m/z 1731.5181, calculated A40926 $[M + H]^+$ 1731.5107, 4.27 ppm difference) (Figure 5a and b).

MS/MS analysis of the molecule showed it also had an identical fragmentation pattern to the A40926 standard, including the in-source fragment with m/z 358.22 characteristic of the carboxylated GlcN-Acyl moiety (Figures 5c and S17). Traces of A50926 could also be detected in the extract of the complemented strain, indicating that while complementation was very efficient, conversion from A50926 to A40926 was not complete (Figure 5a). Alongside the BGC homology (Figure 2), this provides strong evidence that A50926 is chemically identical to A40926 with the exception of the carboxylated GlcN-Acyl. However, we cannot completely rule out small differences, such as acyl chain branching.

2.8. Heterologous Expression of Transcriptional Regulators *dbv3* and *dbv4* to Enhance the Production of A50926 in *N. coxensis*. In previous work, we overexpressed the two *dbv* BGC situated master regulators in *N. gerenzanensis* (*dbv4* and *dbv3*) to successfully improve A40926 production.[6] Therefore, hereby we used the previously constructed expression vectors pSAD4 and pSAD3 carrying *dbv4* and *dbv3*, respectively, in *N. coxensis* to trigger and improve A50926 production. First, we observed that *N. coxensis* pSAD3$^+$ and pSAD4$^+$ recombinant strains grown in the E27 and VSP vegetative media produced an antimicrobial activity against *B. subtilis* (Figure S18A), whereas their parental wild type strain did not exhibit any antimicrobial activity in these media. Overexpression of *dbv3* also triggered antimicrobial activity on VM0.1 and ISP2 solid media, whereas the wild type was not active (Figure S18B). Consistently, in both ISP2lm and TM1m production media *N. coxensis* pSAD3$^+$ and pSAD4$^+$ produced more antibiotic than the wild type (Figures S18C and 6a and b). In ISP2lm (Figure 6a), at 192 h *N. coxensis* pSAD3$^+$ reached the maximum production of approximately 45 $\mu g\ mL^{-1}$, exceeding both wild type (approximately 20 $\mu g\ mL^{-1}$) and *N. coxensis* pSAD4$^+$ (approximately 30 $\mu g\ mL^{-1}$) productivities. In TM1m medium (Figure 6b), *N. coxensis* pSAD3$^+$ produced approximately 50 $\mu g\ mL^{-1}$ after 192 h of cultivation. At the same time point in TM1m the wild type and *N. coxensis* pSAD4$^+$ produced approximately 16 and 22 $\mu g\ mL^{-1}$ of antibiotic, respectively. The control strain carrying the "empty" pSET152A vector performed exactly as the wild type (data not shown). No significant differences between biomass accumulation or pH were observed among the recombinant strains, or in comparison with the parental *N. coxensis* wild type strain. Thus, overexpression of *dbv3* and *dbv4* regulatory genes triggered or improved the production of A50926 in *N. coxensis* under different cultivation conditions.

3. CONCLUSIONS

A novel GPA, A50926, was identified from *N. coxensis* DSM 45129. Detailed MS and MS/MS analysis indicates that A50926 differs from the previously characterized A40926 GPA by lacking the carboxyl group on the GlcN-Acyl moiety attached to Hpg4 of the GPA aglycone, resembling teicoplanin in this part of the molecule. A compound with the same chemical structure was described 25 years ago as a chemically prepared derivative of A40926 (named RA[28]). Extensive study of antibacterial activities of RA *in vitro*[28] indicated that RA has slightly better antimicrobial activity than A40926: minimal inhibitory concentrations (MICs) of RA were 2−4 times lower against different staphylococcal and enterococcal strains when compared to A40926. The difference of chemical structure between the newly described A50926 and A40926 correlates with the absence of *dbv29* orthologue in the A50926 BGC (*noc*). Consistently, when *dbv29* was introduced into *N. coxensis*, we obtained A40926 production in the recombinant strain. Otherwise, both *noc* and *dbv* BGCs share all biosynthetic genes, which are closely related. Heterologous expression of A40926 regulatory genes *dbv3* and *dbv4* in *N. coxensis* improved A50926 production.

Although the majority of *noc* and *dbv* genes are orthologous, the *dbv* BGC is significantly rearranged in comparison to the *noc* BGC, as well as all other characterized GPA BGCs. We have proposed a series of genetic inversions that could have occurred in a common *Nonomuraea* ancestor to explain these different genetic architectures. Both BGCs are quite similar to the putative GPA BGC from *Nonomuraea* sp. WAC 01424. The latter lacks genes required for the addition of mannose, but possesses a gene encoding a sulfotransferase and an additional gene encoding a halogenase. Thus, the putative nonmannosylated GPA from *Nonomuraea* sp. WAC 01424 might be sulfated and have a different chlorination pattern than A40926/A50926. Consequently, *Nonomuraea* sp. WAC 01424 GPA BGC seems an attractive source for new tailoring genes to obtain A40926 derivatives with altered pharmacological properties. Notwithstanding the GPA BGC similarity, multilocus phylogeny of *Nonomuraea* spp. shows that GPA producers are not clustered together: GPA producers are found in distinct clades within the genus. Our analysis indicates that type IV and V GPA BGCs are common in *Nonomuraea* spp., which is in contrast to how rare these BGCs were believed to be. This is comparable to studies that show that BGCs for types I−III−IV GPAs are common in *Amycolatopsis*, and type V GPAs in *Streptomyces*.[27,42,43] This highlights how rare actinomycete genera, such as *Nonomuraea*, may represent a rich untapped source of novel GPAs, as well as GPA tailoring enzymes for the diversification of existing GPA scaffolds.

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

4. METHODS

4.1. Bacterial Strains and Cultivation Conditions. Bacterial strains and plasmids used in this work are summarized in Table S6. Compositions of all the media used for cultivation and GPA production are also given in the Supporting Information. All media components and antibiotics were supplied by Sigma-Aldrich, unless otherwise stated. For routine maintenance, *N. gerenzanensis* and *N. coxensis* strains were cultivated on ISP3 agar medium supplemented with 50 μg mL^{-1} apramycin-sulfate when appropriate. For genomic DNA isolation, *N. gerenzanensis* and *N. coxensis* strains were cultivated in liquid VSP medium on an orbital shaker at 220 rpm and at 30 °C. The working cell banks (WCBs) for *N. gerenzanensis* and *N. coxensis* strains were prepared as described previously.[22,37] *E. coli* DH5α was used as a routine cloning host, and *E. coli* ET12567 pUZ8002 was used as a donor for intergeneric conjugations. *E. coli* strains were cultivated at 37 °C in LB liquid or agar media supplemented with 100 μg mL^{-1} of apramycin-sulfate, 50 μg mL^{-1} of kanamycin-sulfate, and 25 μg mL^{-1} of chloramphenicol when appropriate.

4.2. Plasmid Construction and Generation of Recombinant *N. coxensis* Strains. To construct the pSAD29 expression vector, the coding sequence of *dbv29* (1601 bp) was amplified from the genomic DNA of *N. gerenzanensis* using dbv29_F/R primer pair (Table S7) and Q5 high-fidelity DNA polymerase (New England Biolabs). The resulting amplicon was digested with *Eco*RI and *Eco*RV restriction endonucleases and cloned into pSET152A[44,45] cleaved at the same binding sites. The resulting plasmid was verified by endonuclease restriction mapping and sequencing at BMR Genomics.

pSAD29, as well as pSAD3,[6] pSAD4,[6] and pSET152A,[44] were transferred to *N. coxensis* conjugatively, as described previously.[6] Transconjugants were selected as resistant to 50 μg mL^{-1} of apramycin-sulfate. Obtained strains were verified by PCR. To verify the integration of pSAD29, a ~1.1 kbp fragment of pSAD29 was amplified using the dbv29_seq_int/PAM_seq_R (Table S7) primer pair, in which dbv29_seq_int anneals within *dbv29* and PAM_seq_R anneals upstream the *Eco*RV cleavage site of pSET152A. To verify the integrations of pSAD4 and pSAD3, ~1 kbp and ~2 kbp fragments were amplified respectively using PAM_seq_F/dbv4_R and PAM_seq_F/dbv3_seq_R primer pairs (Table S7). Finally, the integration of pSET152A was verified by amplifying *aac(3)IV* with the aac(3)IV_F/R primer pair (Table S7). In all cases, genomic DNA was isolated using the Kirby procedure.[46]

4.3. *N. coxensis* Cultivation for A50926 Production. To initiate the cultivation of *N. coxensis*, one WCB vial was inoculated into a 250 mL Erlenmeyer flask with 50 mL of VSP reactivation medium containing 6 glass beads (ø5 mm). After 72 h of incubation on a rotary shaker at 220 rpm, 30 °C the culture was used to inoculate (10% v/v) 500 mL Erlenmeyer flasks containing 100 mL of E27 vegetative medium and 12 glass beads (ø 5 mm). Following 72 h of incubation on a rotary shaker at 220 rpm, 30 °C this culture was used to inoculate (10% v/v) 500 mL Erlenmeyer flasks with 100 mL of ISP2lm or TM1m production media containing 12 glass beads (ø 5 mm). A50926 production cultures were then incubated up to 240 h on a rotary shaker at 220 rpm, 30 °C. Samples were collected at regular time points to estimate biomass accumulation (dry weight), pH, and A50926 production.

4.4. VanY-Related Activity Measurement. D,D-carboxypeptidase activity in *Nonomuraea* spp. was measured in FM2 production medium (*N. gerenzanensis*) and ISP2lm (*N. coxensis*) at 24, 48, 72, 96, 120, and 144 h time points. Mycelial lysates were prepared as described previously.[34] The enzyme activity releasing D-Ala from the tripeptide *N*-Acetyl-L-Lys-D-Ala-D-Ala (10 mM) was followed spectrophotometrically by a D-amino acid oxidase/peroxidase coupled reaction that oxidizes the colorimetric substrate 4-aminoantipyrine to chinonemine. D,D-carboxypeptidase activity was normalized to dry biomass weight, as previously reported.[22] One unit is defined as the amount of enzyme that is able to convert 1 μmol of substrate in 1 min.

4.5. HPLC and LC-MS Analysis of GPAs. For quantitative measurement, A40926 and A50926 were extracted from *N. coxensis* cultures with equal volumes of borate buffer composed of 100 mM H_3BO_3 (Sigma-Aldrich) and 100 mM NaOH (Sigma-Aldrich), pH 12. During this extraction the *O*-acetylated forms were converted in the corresponding deacetylated GPAs A40926 and A50926. A40926 and A50926 were analyzed using HPLC as previously reported.[6,22,37] In all cases the injection volumes of studied samples and standards were the same (50 μL). Concentration of A50926 was estimated as follows:

$$\text{A50926 concentration}\left(\frac{\text{mg}}{\text{L}}\right) = \frac{C(\text{A40926 std}) \times A(\text{A50926})}{A(\text{A40926 std})} \times 2$$

Where, C(A40926 std) is the concentration of the commercial A40926 sample; A(A50926) is the area sum of the peaks corresponding to A50926 B; A(A40926 std) is the area of the peak corresponding to the standard A40926 factor B_0; and 2 is the dilution factor.

High resolution liquid chromatography−mass spectrometry (LC-MS) and fragmentation (MS/MS) analysis of A40926 and A50926 was carried out on a SYNAPT G2-Si mass spectrometer equipped with an Acquity UPLC (Waters). Samples were injected onto a Waters Acquity UPLC BEH 1.7 μm, 1 × 100 mm C18 column, and eluted with a gradient of (B) acetonitrile/0.1% formic acid in (A) water/0.1% formic acid with a flow rate of 0.08 mL min^{-1} at 45 °C. The concentration of B was kept at 1% for 2 min followed by a gradient up to 40% B over 9 min, ramping to 99% B in 1 min, kept at 99% B for 2 min and re-equilibrated at 1% B for 4 min. MS data were collected in positive mode with the following parameters: resolution mode, scan time 0.5 s, mass range m/z 50−2000 calibrated with sodium iodide, capillary voltage = 2.5 kV; cone voltage = 40 V; source temperature = 120 °C; desolvation temperature = 350 °C. Leu-enkephalin peptide was used to generate a lock-mass calibration with m/z 556.2766 for positive mode, measured every 90 s during the run. For MS/MS fragmentation, a data directed analysis (DDA) method was used with the following parameters: precursor selected from the 4 most intense ions; MS/MS threshold 5000; scan time 2 s; no dynamic exclusion. Collision energy (CE) was ramped between 8 and 35 at low mass (m/z 50) and 10−70 at high mass (m/z 1200).

4.6. Purification of GPAs Using D-Ala-D-Ala Based Affinity Resin. GPAs were purified by affinity chromatography with a D-Alanine-D-Alanine (D-Ala-D-Ala) based resin. Activation of 5 mL HiTrap NHS-activated HP affinity columns (GE Healthcare) and ligand binding was conducted as described before[47] with modifications. Briefly, the resin was activated with 30 mL of 1 mM HCl, followed by injection of 200 mM D-Ala-D-Ala dipeptide, dissolved into 5 mL of coupling buffer (0.2 M $NaHCO_3$, pH 7.0). After 30 min incubation, the resin was washed with three cycles of 0.5 M ethanolamine hydrochloride, 0.5 M NaCl (pH 4.0, 30 mL), followed by 0.1 M sodium acetate, 0.5 mM NaCl (pH 4.0, 30 mL), alternately. Finally, the resin was washed with 50 mL coupling buffer and left to equilibrate for at least 1 h before use.

N. coxensis cultures were extracted in borate buffer as reported above, the pH in the obtained extracts was adjusted to 7.5 with HCl, and they were applied to the affinity chromatography system. Thus, extracts in borate buffer, coming from *N. coxensis* strains cultivated in TM1m or ISP2lm media, were filtered with 0.45 μm cutoff and loaded onto a D-Ala-D-Ala column at a flow rate of 0.5 mL min^{-1}. After extensive washing with coupling buffer, the bound GPA was eluted with 0.1 M NaOH and the eluate was lyophilized.

4.7. Bioassays for the Detection of A50926. Agar plug or Whatman paper disc (GE Healthcare) antibiotic diffusion assays were used to determine antimicrobial activities. An overnight *B. subtilis* ATCC 6633 culture in Mueller-Hinton broth II (cation adjusted, Sigma-Aldrich) was used to inoculate (1% v/v) a fresh culture, which was grown to OD_{600} = 0.6. A 200 μL portion of this culture was then added to 25 mL of 0.7% (w/v) Mueller-Hinton agar (Condalab) and plated. After solidification of the media, agar plugs cut from the plates with *N. coxensis* lawns, or Whatman paper discs containing GPAs,

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

were placed on the agar surface. Bioassay plates were incubated for 16 h at 37 °C before examination.

4.8. Sequencing and Annotation of the *N. coxensis* Genome. The genome of *N. coxensis* was sequenced using a combination of HiSeq Illumina and GridION ONT technologies. The Illumina data was obtained from SRA (PRJNA165411), while for the ONT data, a sequencing library (SQK-LSK109) was prepared using the Ligation Sequencing Kit (Oxford Nanopore Technologies) according to the manufacturer's instructions and run on a GridION sequencer in an R9.4.1 flowcell (both Oxford Nanopore Technologies). Base-calling of the raw data was performed with GUPPY-FOR-GRIDION v3.0.6. The assembly and polishing were performed as described previously,[48] using canu v.1.8 instead of v.1.6. The ONT data was assembled into 5 contigs, while the Illumina data were assembled into 87 scaffolds containing 310 contigs using NEWBLER v2.8. After manual curation using CONSED,[49] the complete genome of *N. coxensis* DSM 45129, consisting of one circular chromosome of 9,073,954 bp (72.12% G + C) was obtained. Annotation was performed using PROKKA v1.11[50] resulting in the prediction of 8,398 coding sequences (CDS), 5 rRNA operons, 73 tRNAs, and 5 noncoding RNA elements. The annotated genome and ONT raw data were deposited at DDBJ/ENA/GenBank under the BioProject accession number PRJNA693185.

4.9. *In Silico* Analysis Tools and Approaches. Routine analysis of nucleotide and amino acid sequences was performed in GENEIOUS v4.8.5.[51] Multiple sequence alignments, selection of the best models for the phylogenetic reconstruction and phylogenetic reconstruction itself were done with the MEGA X package.[52] To reconstruct the multilocus phylogeny of *Nonomuraea*, orthologues of 30 *S. coelicolor* house-keeping proteins (Table S3,[53]) were identified within the genomes of 34 *Nonomuraea* spp. (Table S2) using reciprocal best hit (RBH) BLAST. Sequences of these proteins from each *Nonomuraea* spp. were concatenated, and these concatenates were used for the upstream phylogenetic reconstruction.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.1c00170.

Tables S1−S7 and Figures S1−S18 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Andrew W. Truman − *Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, United Kingdom*; orcid.org/0000-0001-5453-7485; Email: andrew.truman@jic.ac.uk

Flavia Marinelli − *Department of Biotechnology and Life Sciences, University of Insubria, 21100 Varese, Italy*; orcid.org/0000-0001-9195-6777; Email: flavia.marinelli@uninsubria.it

Authors

Oleksandr Yushchuk − *Department of Biotechnology and Life Sciences, University of Insubria, 21100 Varese, Italy*; orcid.org/0000-0003-3296-8297

Natalia M. Vior − *Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, United Kingdom*

Andres Andreo-Vidal − *Department of Biotechnology and Life Sciences, University of Insubria, 21100 Varese, Italy*

Francesca Berini − *Department of Biotechnology and Life Sciences, University of Insubria, 21100 Varese, Italy*

Christian Rückert − *Technology Platform Genomics, CeBiTec, Bielefeld University, 33615 Bielefeld, Germany*

Tobias Busche − *Technology Platform Genomics, CeBiTec, Bielefeld University, 33615 Bielefeld, Germany*

Elisa Binda − *Department of Biotechnology and Life Sciences, University of Insubria, 21100 Varese, Italy*

Jörn Kalinowski − *Technology Platform Genomics, CeBiTec, Bielefeld University, 33615 Bielefeld, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.1c00170

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the public grant "Fondo di Ateneo per la Ricerca" 2019 to F.M., by the Biotechnology and Biological Sciences Research Council MfN Institute Strategic Programme grant (BBS/E/J/000PR9790) for the John Innes Centre (JIC) and by a Royal Society University Research Fellowship for A.W.T. A.A.-V. is a Ph.D. student of the "Life Science and Biotechnology" course at Università degli Studi dell'Insubria. We thank graduate student E. Salvadè of University of Insubria for her contribution to this research and Consorzio Interuniversitario per le Biotecnologie for supporting E.B. and F.B. We thank C. de Oliveira Martins (JIC) for technical support with mass spectrometry.

■ REFERENCES

(1) Sungthong, R., and Nakaew, N. (2015) The genus *Nonomuraea*: a review of a rare actinomycete taxon for novel metabolites. *J. Basic Microbiol. 55*, 554−565.

(2) Nazari, B., Forneris, C. C., Gibson, M. I., Moon, K., Schramma, K. R., and Seyedsayamdost, M. R. (2017) *Nonomuraea* sp. ATCC 55076 harbours the largest actinomycete chromosome to date and the kistamicin biosynthetic gene cluster. *MedChemComm 8*, 780−788.

(3) D'Argenio, V., Petrillo, M., Pasanisi, D., Pagliarulo, C., Colicchio, R., Talà, A., De Biase, M. S., Zanfardino, M., Scolamiero, E., Pagliuca, C., Gaballo, A., Cicatiello, A. G., Cantiello, P., Postiglione, I., Naso, B., Boccia, A., Durante, M., Cozzuto, L., Salvatore, P., Paolella, G., Salvatore, F., and Alifano, P. (2016) The complete 12 Mb genome and transcriptome of *Nonomuraea gerenzanensis* with new insights into its duplicated "magic" RNA polymerase. *Sci. Rep. 6*, 18.

(4) Narsing Rao, M. P., Dong, Z. Y., Jiao, J. Y., Zhou, Y., Zhao, J., Xiao, M., and Li, W. J. (2020) Genome sequence and comparative analysis of DRQ-2, the type strain of *Nonomuraea indica*. *Genomics 112*, 2842−2844.

(5) Greule, A., Izoré, T., Iftime, D., Tailhades, J., Schoppet, M., Zhao, Y., Peschke, M., Ahmed, I., Kulik, A., Adamek, M., Goode, R. J. A., Schittenhelm, R. B., Kaczmarski, J. A., Jackson, C. J., Ziemert, N., Krenske, E. H., De Voss, J. J., Stegmann, E., and Cryle, M. J. (2019) Kistamicin biosynthesis reveals the biosynthetic requirements for production of highly crosslinked glycopeptide antibiotics. *Nat. Commun. 10*, 2613.

(6) Yushchuk, O., Andreo-Vidal, A., Marcone, G. L., Bibb, M., Marinelli, F., and Binda, E. (2020) New molecular tools for regulation and improvement of A40926 glycopeptide antibiotic production in *Nonomuraea gerenzanensis* ATCC 39727. *Front. Microbiol. 11*, 8.

(7) Yue, X., Xia, T., Wang, S., Dong, H., and Li, Y. (2020) Highly efficient genome editing in *N. gerenzanensis* using an inducible CRISPR/Cas9−RecA system. *Biotechnol. Lett. 42*, 1699−1706.

(8) Nicolaou, K. C., Boddy, C. N. C., Bräse, S., and Winssinger, N. (1999) Chemistry, biology, and medicine of the glycopeptide antibiotics. *Angew. Chem., Int. Ed. 38*, 2096−2152.

(9) Goldstein, B. P., Selva, E., Gastaldo, L., Berti, M., Pallanza, R., Ripamonti, F., Ferrari, P., Denaro, M., Arioli, V., and Cassani, G. (1987) A40926, a new glycopeptide antibiotic with anti-*Neisseria* activity. *Antimicrob. Agents Chemother. 31*, 1961−1966.

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

(10) Marcone, G. L., Binda, E., Berini, F., and Marinelli, F. (2018) Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era. *Biotechnol. Adv. 36*, 534−554.
(11) Yushchuk, O., Ostash, B., Truman, A. W., Marinelli, F., and Fedorenko, V. (2020) Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes. *Appl. Microbiol. Biotechnol. 104*, 3279−3291.
(12) Grundy, W. E., Sinclair, A. C., Theriault, R. J., Goldstein, A. W., Rickher, C. J., Warren, H. B., Oliver, T. J., and Sylvester, J. C. (1956) Ristocetin, microbiologic properties. *Antibiot. Annu.*, 687−692.
(13) Truman, A. W., Kwun, M. J., Cheng, J., Yang, S. H., Suh, J. W., and Hong, H. J. (2014) Antibiotic resistance mechanisms inform discovery: identification and characterization of a novel *Amycolatopsis* strain producing ristocetin. *Antimicrob. Agents Chemother. 58*, 5687−5695.
(14) Spohn, M., Kirchner, N., Kulik, A., Jochim, A., Wolf, F., Muenzer, P., Borst, O., Gross, H., Wohlleben, W., and Stegmann, E. (2014) Overproduction of ristomycin a by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. *Antimicrob. Agents Chemother. 58*, 6185−6196.
(15) Sosio, M., Canavesi, A., Stinchi, S., and Donadio, S. (2010) Improved production of A40926 by *Nonomuraea* sp. through deletion of a pathway-specific acetyltransferase. *Appl. Microbiol. Biotechnol. 87*, 1633−1638.
(16) Alt, S., Bernasconi, A., Sosio, M., Brunati, C., Donadio, S., and Maffioli, S. I. (2019) Toward single-peak dalbavancin analogs through biology and chemistry. *ACS Chem. Biol. 14*, 356−360.
(17) Soriano, A., Rossolini, G. M., and Pea, F. (2020) The role of dalbavancin in the treatment of acute bacterial skin and skin structure infections (ABSSSIs). *Expert Rev. Anti-Infect. Ther. 18*, 415−422.
(18) Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A., and Donadio, S. (2003) The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species. *Chem. Biol. 10*, 541−549.
(19) Li, Y. S., Ho, J. Y., Huang, C. C., Lyu, S. Y., Lee, C. Y., Huang, Y. T., Wu, C. J., Chan, H. C., Huang, C. J., Hsu, N. S., Tsai, M. D., and Li, T. L. (2007) A unique flavin mononucleotide-linked primary alcohol oxidase for glycopeptide A40926 maturation. *J. Am. Chem. Soc. 129*, 13384−13385.
(20) Kaniusaite, M., Tailhades, J., Kittilä, T., Fage, C. D., Goode, R. J. A., Schittenhelm, R. B., and Cryle, M. J. (2021) Understanding the early stages of peptide formation during the biosynthesis of teicoplanin and related glycopeptide antibiotics. *FEBS J. 288*, 507−529.
(21) Binda, E., Marcone, G. L., Pollegioni, L., and Marinelli, F. (2012) Characterization of VanYn, a novel D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727. *FEBS J. 279*, 3203−3213.
(22) Marcone, G. L., Binda, E., Carrano, L., Bibb, M., and Marinelli, F. (2014) Relationship between glycopeptide production and resistance in the actinomycete *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother. 58*, 5191−5201.
(23) Lo Grasso, L., Maffioli, S., Sosio, M., Bibb, M., Puglia, A. M., and Alduina, R. (2015) Two master switch regulators trigger A40926 biosynthesis in *Nonomuraea* sp. strain ATCC 39727. *J. Bacteriol. 197*, 2536−2544.
(24) Alduina, R., Tocchetti, A., Costa, S., Ferraro, C., Cancemi, P., Sosio, M., and Donadio, S. (2020) A two-component regulatory system with opposite effects on glycopeptide antibiotic biosynthesis and resistance. *Sci. Rep. 10*, 6200.
(25) Naruse, N., Tenmyo, O., Kobaru, S., Hatori, M., Tomita, K., Hamagishi, Y., and Oki, T. (1993) New antiviral antibiotics, kistamicins A and B I. Taxonomy, production, isolation, physico-chemical properties and biological activities. *J. Antibiot. 46*, 1804−1811.
(26) Naruse, N., Oka, M., Konishi, M., and Oki, T. (1993) New Antiviral Antibiotics, Kistamicins A and B II. Structure determination. *J. Antibiot. 46*, 1812−1818.
(27) Waglechner, N., McArthur, A. G., and Wright, G. D. (2019) Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat. Microbiol. 4*, 1862−1871.
(28) Malabarba, A., Ciabatti, R., Scotti, R., Goldstein, B. P., Ferrari, P., Kurz, M., Andreini, B. P., and Denaro, M. (1995) New semisynthetic glycopeptides MDL 63,246 and MDL 63,042, and other amide derivatives of antibiotic A-40,926 active against highly glycopeptide-resistant VanA enterococci. *J. Antibiot. 48*, 869−883.
(29) Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H., and Weber, T. (2019) AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res. 47*, W81−W87.
(30) Adamek, M., Alanjary, M., Sales-Ortells, H., Goodfellow, M., Bull, A. T., Winkler, A., Wibberg, D., Kalinowski, J., and Ziemert, N. (2018) Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC Genomics 19*, 426.
(31) Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. (2004) MAUVE: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res. 14*, 1394−1403.
(32) Alduina, R., Sosio, M., and Donadio, S. (2018) Complex regulatory networks governing production of the glycopeptide A40926. *Antibiotics (Basel, Switz.) 7*, 30.
(33) Alduina, R., Piccolo, L. L., D'Alia, D., Ferraro, C., Gunnarsson, N., Donadio, S., and Puglia, A. M. (2007) Phosphate-controlled regulator for the biosynthesis of the dalbavancin precursor A40926. *J. Bacteriol. 189*, 8120−8129.
(34) Binda, E., Marcone, G. L., Berini, F., Pollegioni, L., and Marinelli, F. (2013) *Streptomyces* spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance. *BMC Biotechnol. 13*, 24.
(35) Pootoolal, J., Thomas, M. G., Marshall, C. G., Neu, J. M., Hubbard, B. K., Walsh, C. T., and Wright, G. D. (2002) Assembling the glycopeptide antibiotic scaffold: the biosynthesis of A47934 from *Streptomyces toyocaensis* NRRL15009. *Proc. Natl. Acad. Sci. U. S. A. 99*, 8962−8967.
(36) Ara, I., Kudo, T., Matsumoto, A., Takahashi, Y., and Omura, S. (2007) *Nonomuraea bangladeshensis* sp. nov. and *Nonomuraea coxensis* sp. nov. *Int. J. Syst. Evol. Microbiol. 57*, 1504−1509.
(37) Marcone, G. L., Beltrametti, F., Binda, E., Carrano, L., Foulston, L., Hesketh, A., Bibb, M., and Marinelli, F. (2010) Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother. 54*, 2465−2472.
(38) Taurino, C., Frattini, L., Marcone, G. L., Gastaldo, L., and Marinelli, F. (2011) *Actinoplanes teichomyceticus* ATCC 31121 as a cell factory for producing teicoplanin. *Microb. Cell Fact. 10*, 82.
(39) Puk, O., Bischoff, D., Kittel, C., Pelzer, S., Weist, S., Stegmann, E., Süssmuth, R. D., and Wohlleben, W. (2004) Biosynthesis of chloro-β-hydroxytyrosine, a nonproteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics. *J. Bacteriol. 186*, 6093−6100.
(40) Yushchuk, O., Binda, E., and Marinelli, F. (2020) Glycopeptide antibiotic resistance genes: distribution and function in the producer actinomycetes. *Front. Microbiol. 11*, 1.
(41) Zerilli, L. F., Edwards, D. M. F., Borghi, A., Gallo, G. G., Selva, E., Denaro, M., and Lancini, G. C. (1992) Determination of the acyl moieties of the antibiotic complex A40926 and their relation with the membrane lipids of the producer strain. *Rapid Commun. Mass Spectrom. 6*, 109−114.
(42) Xu, M., Wang, W., Waglechner, N., Culp, E. J., Guitor, A. K., and Wright, G. D. (2020) GPAHex - a synthetic biology platform for type IV−V glycopeptide antibiotic production and discovery. *Nat. Commun. 11*, 5232.
(43) Culp, E. J., Waglechner, N., Wang, W., Fiebig-Comyn, A. A., Hsu, Y. P., Koteva, K., Sychantha, D., Coombes, B. K., Van Nieuwenhze, M. S., Brun, Y. V., and Wright, G. D. (2020) Evolution-guided discovery of antibiotics that inhibit peptidoglycan remodelling. *Nature 578*, 582−587.

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

(44) Horbal, L., Kobylyanskyy, A., Yushchuk, O., Zaburannyi, N., Luzhetskyy, A., Ostash, B., Marinelli, F., and Fedorenko, V. (2013) Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus* - producer of teicoplanin, drug of last defense. *J. Biotechnol. 168*, 367−372.
(45) Horbal, L., Kobylyanskyy, A., Truman, A. W., Zaburranyi, N., Ostash, B., Luzhetskyy, A., Marinelli, F., and Fedorenko, V. (2014) The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. *Appl. Microbiol. Biotechnol. 98*, 9295−9309.
(46) Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000) *Practical Streptomyces genetics*; John Innes Foundation, Norwich, England.
(47) Holding, A. N., and Spencer, J. B. (2008) Investigation into the mechanism of phenolic couplings during the biosynthesis of glycopeptide antibiotics. *ChemBioChem 9*, 2209−2214.
(48) Bekiesch, P., Zehl, M., Domingo-Contreras, E., Martín, J., Pérez-Victoria, I., Reyes, F., Kaplan, A., Rückert, C., Busche, T., Kalinowski, J., and Zotchev, S. B. (2020) Viennamycins: lipopeptides produced by a *Streptomyces* sp. *J. Nat. Prod. 83*, 2381−2389.
(49) Gordon, D., and Green, P. (2013) Consed: A Graphical editor for next-generation sequencing. *Bioinformatics 29*, 2936−2937.
(50) Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics 30*, 2068−2069.
(51) Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and Drummond, A. (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics 28*, 1647−1649.
(52) Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol. 35*, 1547−1549.
(53) Gao, B., and Gupta, R. S. (2012) Phylogenetic framework and molecular signatures for the main clades of the phylum Actinobacteria. *Microbiol. Mol. Biol. Rev. 76*, 66−112.

https://doi.org/10.1021/acschembio.1c00170
ACS Chem. Biol. 2021, 16, 915−928

Оригінал і електронний додаток до статті доступні за посиланням: https://pubs.acs.org/doi/10.1021/acschembio.1c00170.

Анотація: *Глікопептидні антибіотики (ГПА) є препаратами останньої лінії захисту від мультирезистентних Грам-позитивних патогенів. Природні ГПА тейкопланін і ванкоміцин, а також напівсинтетичні оритаванцин, телаванцин і далбаванцин, наразі дозволені для клінічного використання. Хоча ці антибіотики залишаються ефективними, поява нових патогенів, стійких до них, є питанням часу. Тому важливо досліджувати природне різноманіття ГПА, що походять від так званих «рідкісних» актинобактерій. Тут ми описуємо новий продуцент ГПА – Nonomuraea coxensis DSM 45129. Його повністю секвенований і зібраний геном містить кластер біосинтетичних генів (КБГ), подібний до dbv – КБГ A40926, природного попередника далбаванцину. Штам продукує новий ГПА, який є аналогом A40926, у якому відсутня карбоксильна група у залишку N-ацилглюкозаміну. Ця структурна відмінність корелює з відсутністю ортолога dbv29, що кодує фермент, відповідальний за окислення залишку N-ацилглюкозаміну. Експресія dbv29 у N. coxensis веде до продукції A40926 у цьому штамі. Нарешті, ми успішно застосували гетерологічні регулятори транскрипції dbv3 і dbv4 для активації та покращення продукції A50926 у N. coxensis, що робить їх перспективними інструментами для скринінгу інших видів Nonomuraea. Наша робота висвітлює рід Nonomuraea як досі не вичерпане джерело нових ГПА.*

РОЗДІЛ 6 Особливості порівняльної філогеноміки кластерів генів біосинтезу глікопептидних антибіотиків

Досліджені в попередньому розділі A40926 і A50926 є типовими ГПА типу IV [19], що, як і представники типів I-III, активні проти Грам-позитивних бактерій шляхом зв'язування ліпіду II. На противагу, нещодавно було виявлено, що ГПА типу V, такі як комплестатин [125], мають іншу мішень дії, а саме автолізини [126] – ферменти утилізації компонентів клітинної стінки Грам-позитивних бактерій [127]. На відміну від ГПА інших типів, КБГ ГПА типу V зустрічаються майже виключно в представників роду *Streptomyces* [128].

Прототипом ГПА типу V є комплестатин – потужний інгібітор системи комплементу у людини [125], який також має антимікробну активність (діє на автолізини, див. вище). На даний момент хімічна різноманітність типу V є вивченою недостатньо, що спонукало нас в цьому розділі спочатку вивчити філогенію 17 модельних КБГ ГПА, включно із КБГ біосинтезу комплестатину (єдиний КБГ ГПА типу V вибірки). Для цього, ми дослідили філогенію окремих найконсервативніших генів, наявних в усіх 17 КБГ, а саме генів шлях-специфічних регуляторів, генів ферментів біосинтезу амінокислот *dahp, pdh, dpgA, hpgT,* генів найконсервативніших модулів НРПС, а потім збудували мультилокусну філогенію всіх КБГ одночасно. В результаті, виявилося, що гени *com* є, скоріше, зовнішньою групою для всіх інших генів біосинтезу ГПА, даючи змогу ствердити, що КБГ ГПА типу V еволюційно відмежовані від КБГ ГПА типів I-IV.

Наступним кроком, ми проаналізували 2664 геномів стрептоміцетів в пошуках КБГ, подібних на КБГ комплестатину, за допомогою Multi Gene Blast [63] та виявили такі КБГ в 53 геномах. Більше детальний аналіз дав змогу отримати 33 повні послідовності таких КБГ (інші виявилися просеквенованими лише частково). Ми створили мультилокусну філогенію 33 віднайдених КБГ і виявили 5 окремих клад. КБГ в межах однієї клади мали однакову генетичну організацію. Змоделювавши *in silico* біосинтетичні шляхи потенційних ГПА, подібних до комплестатину, які закодовані в КБГ цих 5 клад, ми виявили, що

КБГ деяких клад потенційно можуть кодувати біосинтез нових аналогів комплестатину. Зокрема, КБГ клади V кодують біосинтез комплестатину, клад IV і II – гіпотетичного *N*-малоніл-норкомплестатину, клади III – норкомплестатину, а клади I – *N*-ацетил-норкомплестатину.

Цікаво, що в біосинтетичних генах проаналізованих КБГ було знайдено численні TTA-кодони, а в промоторних ділянках деяких генів (зокрема генів шлях-специфічних регуляторів і НРПС) ми знайшли потенційні сайти зв'язування глобального позитивного регулятора AdpA. Разом, це вказує на ймовірність AdpA-опосередкованого контролю біосинтезу комплестатин-подібних антибіотиків у стрептоміцетів. Ця інформація важлива у світлі даних, отриманих в **Розділі 3**.

Отримані у цьому розділі результати вказують на стрептоміцетів як на джерело нових ГПА антибіотиків типу V, подібних до комплестатину, що заслуговують подальшого експериментального вивчення.

Основні результати цього розділу опубліковано в [129, 130].

УДК 575.87:577.182.82:579.873

ЮЩУК О.С.[1], ОСТАШ Б.О.[1], ГОРБАЛЬ Л.О.[2], ФЕДОРЕНКО В.О.[1]
[1] Львівський національний університет імені Івана Франка,
Україна, 79005, м. Львів, вул. Грушевського, 4, e-mail: v_fedorenko@franko.lviv.ua
[2] Helmholtz-Institute for Pharmaceutical Research Saarland,
Germany, Saarbrucken
✉*v_fedorenko@franko.lviv.ua, +38 (032) 239-44-07, +38 (032) 239-44-75*

РЕКОНСТРУКЦІЯ ФІЛОГЕНІЇ КЛАСТЕРІВ ГЕНІВ БІОСИНТЕЗУ ГЛІКОПЕПТИДІВ

Клас глікопептидних антибіотиків представлений сполуками, що сьогодні складають останню лінію боротьби проти мультирезистентних Грам-позитивних патогенів [1]. Глікопептиди ванкоміцин та тейкопланін є перевіреними часом клінічними сполуками. Медичну важливість глікопептидних антибіотиків важко переоцінити, що й стимулює інтерес до дослідження цього класу природних сполук. Добре дослідженими глікопептидами, кластери генів біосинтезу яких секвеновано, є лише тейкопланін, ванкоміцин, баліміцин, рістоцетин, хлороеремоміцин, сполуки A40926, A47934 та UK-68597. У базах даних нуклеотидних послідовностей присутні анотовані генні кластери, які потенційно можуть кодувати біосинтез глікопептидів. Проте, ці кластери знайдено в секвенованих зразках метагеномної ДНК чи в геномах неідентифікованих бактерій.

Кластери генів біосинтезу глікопептидів (КБГ) досить подібні між собою, що вкладається в модель модульної еволюції кластерів генів біосинтезу природних сполук [2]; продукти окремих генів одних кластерів можуть брати на себе роль своїх гомологів з інших кластерів [3]. Це становить особливий інтерес, якщо врахувати, що штучні комбінації генів біосинтезу глікопептидів можуть потенційно давати нові сполуки. Такий підхід можна застосувати, створюючи нові глікопептиди шляхом комбінаторного біосинтезу. Очевидно, що комбінаторний біосинтез нових глікопептидних сполук ґрунтуються на розумінні філогенії компонентів КБГ у контексті їхніх біологічних функцій. Спроби побудувати філогенію КБГ вже здійснювалися [4], проте, для формування цілісного погляду на цю проблему на той момент було замало секвенованих генів біосинтезу глікопептидів. На сьогодні є значно більше інформації про секвеновані КБГ та більше експериментальних даних. Це дає змогу побудувати цілеснішу картину філогенії КБГ. У цій роботі ми проаналізували філогенетичні взаємозв'язки окремих компонентів КБГ і на основі отриманих даних запропонували консенсусну модель філогенії цих генних кластерів.

Матеріали і методи

Вихідні нуклеотидні послідовності кластерів біосинтезу глікопептидних антибіотиків отримано із бази даних GenBank; у роботі використано такі послідовності (скорочення, використане в роботі/номер депонування в GB): *tei*/AJ632270 (тейкопланін); UK68597/KF192710 (UK-68,597); *bal*/Y16952 (баліміцин); *van*/HE589771 (ванкоміцин); *risL, risJ*/KJ364518, CP008953 (рістоцетини); *cem*/AJ223998-9 (хлороеремоміцин); *VEG*/EU874252; *TEG*/EU874253.1; *A40926*/AJ561198 (A40926, попередник дальбаванцину); *A47934*/STU82965 (A47934); *CA37*/HM486074; *CA915*/HM486076; *esnapd_15*/KF264554; *pek_WAC1420, pek_WAC4229*/JX026280.1, KC688274.1 (пекіскоміцини); *com*/AF386507.1 (комплестатин); *feg*/KT809366.1 (фегліміцин). Первинний біоінформатичний аналіз амінокислотних та нуклеотидних послідовностей здійснено за допомогою програми Geneious 4.8.5 [5]. Побудову множинних вирівнювань амінокислотних та нуклеотидних послідовностей, а також конструювання філогенетичних дерев здійснено за допомогою програми MEGA6 [6]. Для побудови множинних вирівнювань використовували MUSCLE із методом ієрархічної кластеризації UPGMA/UPGMB. Дерева у цій роботі створено за допомогою методу Maximum Likelihood, із використанням моделі Джукса-Кантора для нуклеотидних замін. Топологію фінального дерева будували на основі 500 бутстрепів.

Результати та обговорення

Як вже зазначалося, сьогодні в базах да-

© ЮЩУК О.С., ОСТАШ Б.О., ГОРБАЛЬ Л.О., ФЕДОРЕНКО В.О.

них нуклеотидних послідовностей можна знайти вже багато кластерів, що кодують біосинтез як відомих, так і потенційних глікопептидів. Для проведення філогенетичного аналізу ми відібрали 18 вище перелічених кластерів. Серед «глікопептидних» кластерів присутній і кластер генів біосинтезу лінійного пептидного антибіотика фегліміцину (*feg*), побудованого із 13 амінокислотних залишків (на відміну від 7 таких залишків у глікопептидів), який має багато спільного із «глікопептидними» кластерами. Для обраної вибірки кластерів нами визначено групи ортологічних генів, представники яких присутні у переважній більшості кластерів. Це, зокрема, такі: гени шлях-специфічних регуляторів родини StrR, гени біосинтезу непротеїногенних амінокислот L-3,5-дигідроксифенілгліцину (Dpg) та D-4-гідроксифенілгліцину (Hpg), гени ферментів шикіматного шляху, гени галогеназ і сульфотрансфераз та гени найконсервативніших модулів нерибосомної пептидсинтетази (НРПС4-7). Вони, можливо, були присутні у предковому кластері генів, з якого походять усі сучасні КБГ, і їх можна використати у побудові мультилокусної філогенії останніх. Гени, що присутні лише у деяких кластерах та забезпечують унікальні модифікації (ацилювання, глікозилювання тощо), у цій роботі не розглядалися.

Філогенетичні взаємозв'язки генів шлях-специфічних регуляторів біосинтезу глікопептидів. Усі досліджені генні кластери містять ген, що кодує транскрипційний регулятор родини StrR. Експериментально доведено, що StrR-регулятори задіяні в синтезі баліміцину [7], рістоміцину [3], тейкопланіну [8] та A40926 [9]. Усі вони описані як такі, що мають ключове значення для біосинтезу. Ми використали нуклеотидні послідовності відповідних генів для побудови загального філогенетичного дерева 1А (рис. 1 А), використавши як аутгруп послідовність гена *strR* кластера генів біосинтезу стрептоміцину.

На отриманому дереві можна розрізнити дві основні клади, топологія яких була підтверджена максимальними бутстреп-індексами: перша (I) представлена генами з шляхів біосинтезу баліміцину, ванкоміцину, рістоміцинів, пекіскоміцинів, A40926 та з кластерів *TEG* та *WEG*. У другу (II) увійшов ген шлях-специфічного регулятора кластера генів біосинтезу фегліміцину. Гени StrR-регуляторів з кластерів біосинтезу A47934 та комплестатину зайняли проміжне положення між двома основними кладами.

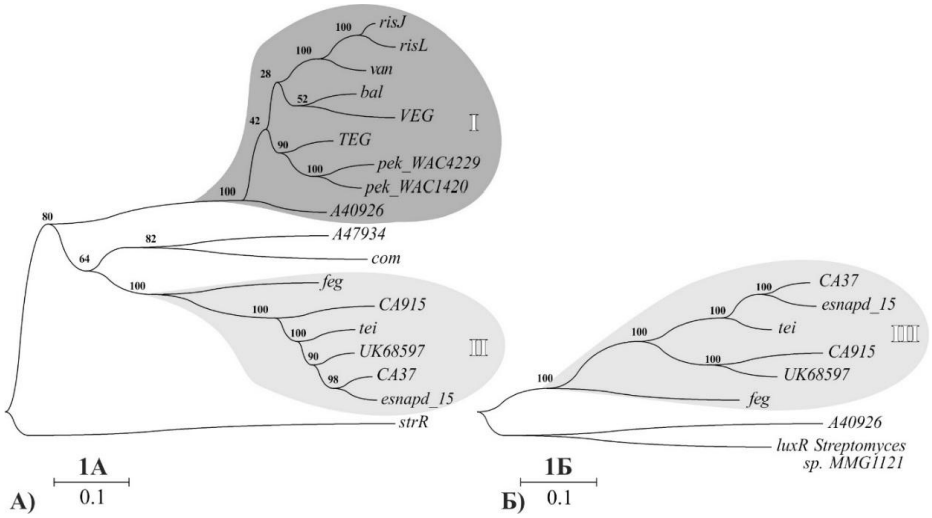


Рис. 1. Філогенетичні дерева, побудовані на основі нуклеотидних послідовностей генів StrR-подібних регуляторів (А) та LuxR-подібних регуляторів (Б) дослідженої вибірки КБГ. Планка масштабу (тут і далі) вказує на кількість нуклеотидних замін на один сайт.

У деяких КБГ присутній другий шлях-специфічний регулятор, що належить до родини LuxR. Вони маловивчені, але відомо, що нокаути генів LuxR-регуляторів блокують біосинтез A40926 та тейкопланіну [8, 9]. Ми спробували відшукати інші LuxR-регулятори у нашій вибірці кластерів. Виявлено ортологи *tei16** (гена LuxR-регулятора з *tei*-кластера) у послідовностях усіх кластерів, що належали до клади II дерева 1А. Отже, StrR-регулятори «глікопептидних» кластерів можна поділити на дві еволюційно віддалені підгрупи. Це корелює із нашими експериментальними даними, які свідчать, що експресія *tei15** в штамі *Amycolatopsis japonicum* не впливала на синтез рістоміцину, тоді як надекспресії гена власного StrR, чи гена StrR з кластера генів біосинтезу баліміцину – активували цей біосинтез [3]. Аналогічно побудовано філогенетичне дерево 1Б (рис. 1 Б) для генів усіх LuxR-регуляторів. Як аутгруп для цього дерева використано ген недослідженого LuxR-регулятора з *Streptomyces* sp. MMG1121. Тут помітно, що всі LuxR-регулятори з кластерів генів біосинтезу тейкопланіну, UK-68,597, фегліміцину і кластерів *CA37*, *CA915*, *esnapd_15* формують одну чітко відділену кладу (III). Ген LuxR-регулятора біосинтезу A40926 до неї не входить, проте її аутгрупує разом із геном з *Streptomyces* sp. MMG1121. Цікавою є топологія клади III дерева 1Б, що практично повторює топологію клади II дерева 1А. Імовірно, що спільний предок («протокластер») клад II/III вже мав як StrR-, так і LuxR-регулятор. Водночас, спільний предок клади II мав лише StrR-регулятор, а LuxR-регулятор зі шляху біосинтезу A40926, мабуть, має походження, незалежне від решти досліджених LuxR.

Філогенія генів, що забезпечують біосинтез попередників глікопептидів. До складу агліконів глікопептидних антибіотиків входять непротеїногенні ароматичні амінокислоти, які походять із тирозину. Велику потребу в тирозині як попередникові глікопептидів відображено в структурі генних кластерів: у більшості КБГ є гени, які кодують ферменти шикіматного шляху біосинтезу ароматичних амінокислот. Це гени префенат-дегідрогеназ (*pdh*) та дезокси-арабіногептулозонат-фосфат-синтаз (*dahp*). Для кластера біосинтезу баліміцину роль таких генів доведено експериментально [10]. Ми виявили, що гени *pdh* присутні у всіх 18 досліджених кластерах, натомість, *dahp* – лише у 12. Для того, щоб оцінити взаємозв'язки між генами *pdh,* ми побудували дерево 2А (рис. 2 А). Як аутгруп ми використали ген *pdh* первинного метаболізму *A. orientalis* (*AORI_RS42670*).

На дереві 2А можна розрізнити дві головні клади (IV та VI) та одну мінорну (V). Вміст клад IV та VI збігається із вмістом клад I та II дерева 1А. Аналогічний розподіл спостерігався при аналізі філогенії *dahp* (рис. 2 Б). І тут вміст клад VII та VIII дерева 2Б відповідав вмісту клад IV і VI дерева 2А відповідно.

Філогенія генів, продукти яких задіяні в біосинтезі непротеїногенних амінокислот. До складу агліконів глікопептидних антибіотиків часто входять непротеїногенні амінокислоти, такі як Dpg і Hpg. Кожен кластер містить у своєму складі набори консервативних генів, що контролюють їхній біосинтез. Для філогенетичного аналізу обрано гени *hpgT* та *dpgA*, продукти яких забезпечують біосинтез Hpg та Dpg. Їхні ортологи знайдено у всіх генних кластерах. Винятки – відсутність ортолога *dpgA* у *com*-кластері та *TEG*-кластері. Втім, гептапептидний попередник комплестатину не містить залишків Dpg, а продукт *teg*-кластера невідомий. Дерева ортологів *hpgT* (рис. 3 А) та *dpgA* (рис. 3 Б) мали топологію, характерну і для попередніх дерев (рис. 1, 2). На кожному помітно дві основні клади, одна з яких містить гени з кластерів *tei*, *UK68597*, *CA37*, *CA915*, *esnapd_15* (клади IX та XII дерев 3А і 3Б), а друга – з генів кластерів *bal*, *van*, *ris*, *pek*, *A40926*, *TEG* та *VEG* (клади X та XI). Ген *hpgT* з *com*-кластера займав проміжне положення між двома основними кладами дерева 3А, тяжіючи до клади X. Гени *dpgA* з кластерів *A47934* та *feg* виявилися близькими до клади XII.

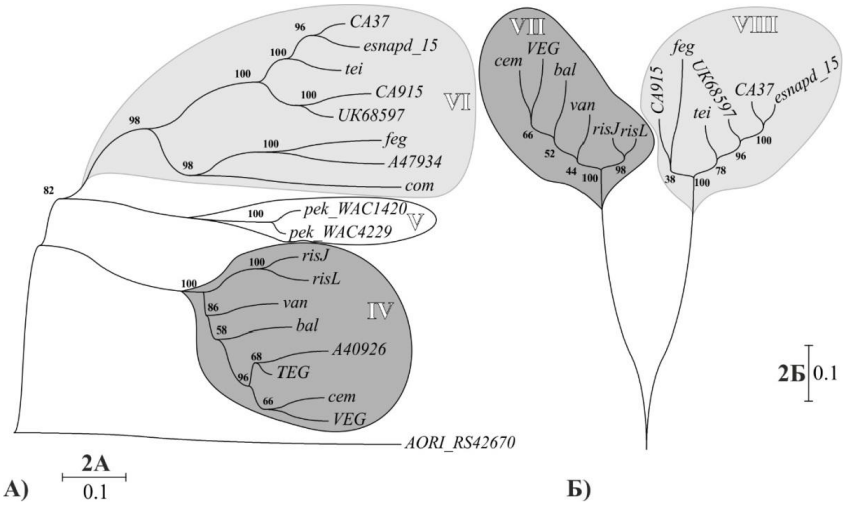


Рис. 2. Філогенетичне дерево, побудоване на основі нуклеотидних послідовностей генів *pdh* (А) та невкорінене дерево генів *dahp* (Б) з дослідженої вибірки КБГ.

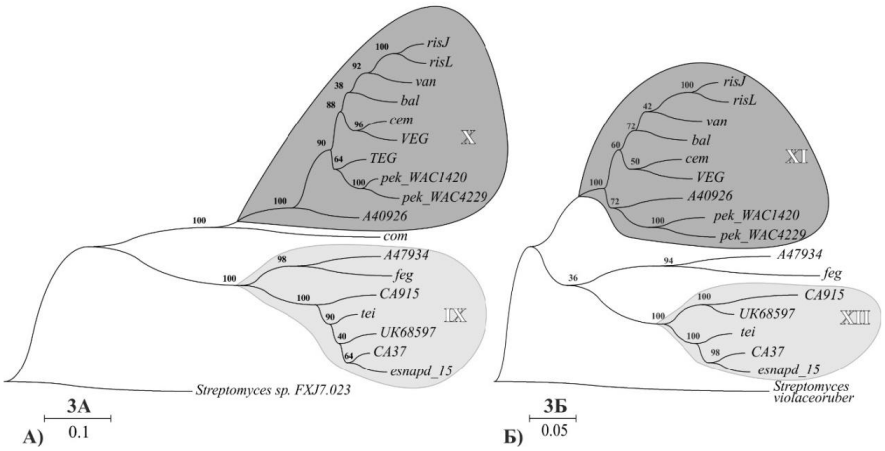


Рис. 3. Дерева, побудовані на основі нуклеотидних послідовностей генів *hpgT* (А) та *dpgA* (Б) дослідженої вибірки КБГ. Як аутгруп для першого було використано послідовність недослідженого гена *hpgT S.* sp. FXJ7.023, для другого – *dpgA* кластера генів біосинтезу кендоміцину *S. violaceusruber*.

Філогенія генів, продукти яких задіяні в хлоруванні та сульфурилюванні. Серед хімічно вивчених глікопептидів усі, окрім ристоцетинів, аглікони «декоровано» атомами хлору. Найчастіше ним модифіковано 2 і 6 амінокислотні залишки (ак. з.) агліконів (у ванкоміцині, баліміцині, хлоролеремоміцині, тейкопланіні). Крім того, у A47934 хлорованим є 5 ак.з. аглікону, а в UK-68,597 – перший. У A40926 хлоровані 6 і 3 ак.з, а у пекіскоміцині – 6 ак.з. Продуценти сполук, що містять атом хлору у положенні, додатковому до стандартних (2 та 6 ак.з.), у своїх генних кластерах містять додатковий ген галогенази. Усі гени галогеназ консервативні, що робить побудову дерев із достовірною топологією неможливою. Єдиною кладою, що підтверджена бутстреп індексом 100, є та, що сформована генами галогеназ кластерів *pek, bal, VEG* та *van,* що відповідає кладам I, VII, X, XI попередніх дерев. Оскільки кластери генів

біосинтезу рістоцетинів завжди групувалися із генами *pek, bal, VEG* та *van,* можна припустити, що *ris*-кластери спочатку мали ген галогенази, подібний до генів галогеназ кластерів *pek, bal, VEG* та *van*, який пізніше втратили. Отже, найімовірніше, первісна галогеназа була здатна включати атом хлору саме в стандартних положеннях. Ми також припускаємо, що гени галогеназ, які приєднують хлор у нестандартних положеннях, виникли унаслідок дуплікації первісного гена та дальшої дивергенції. Сульфурилювання глікопептидів відбувається за допомогою сульфотрансфераз, що належать до I надродини. Відповідні гени, знайдені у кластерах *A40926, A47934, UK68597, TEG* та *pek*, консервативні й формують одну щільну кладу на загальному дереві сульфотрансфераз актинобактерій. Експериментальні дані [11] показують, що всі вони мають різну регіоспецифічність дії. Корелювати цю регіоспецифічність із філогенією неможливо.

Філогенетичні взаємозв'язки генів 4–7 модулів НРПС. Біосинтез пептидної частини глікопептидів забезпечується модульними НРПС. Відповідні гени консервативні і мають багато ділянок подібності. Тому різні генні кластери (не тільки всередині родини глікопептидних антибіотиків) здатні обмінюватися окремими генами (їхніми фрагментами), що кодують модулі НРПС. У практично всіх досліджених випадках з нашої вибірки найконсервативнішими є модулі 4–7 НРПС, продукти яких забезпечують включення 4, 5, 6 та 7 ак. з. у гептапептид: двох Hpg, β-гідрокситирозину та Dpg відповідно. Винятками є НРПС фегліміцину (принципово інша структура кору) та комплестатину (замість β-гідрокситирозину та Dpg включається тирозин та лейцин). Ми побудували філогенетичне дерево на основі генів 4–7 модулів НРПС 17 кластерів нашої вибірки. Топологія отриманого дерева нагадує ту, що спостерігали вище (рис. 1): дві чітко відділені кладі, сформовані генами із *tei*, *UK68597*, *CA37*, *CA915*, *esnapd_15* та генами із кластерів *bal, van, ris, pek, A40926, TEG ma VEG*.

Підсумовуючи всі дані, нами побудовано мультилокусну філогенію КБГ. Для цього ми конкатенували нуклеотидні послідовності найконсервативніших груп ортологів: *pdh, dpgA, hpgT*, *strR*-регуляторів та 7 модуля НРПС. Гени *feg*-кластера не враховувалися. Фінальне дерево та коментарі до нього зображено на рис. 4.

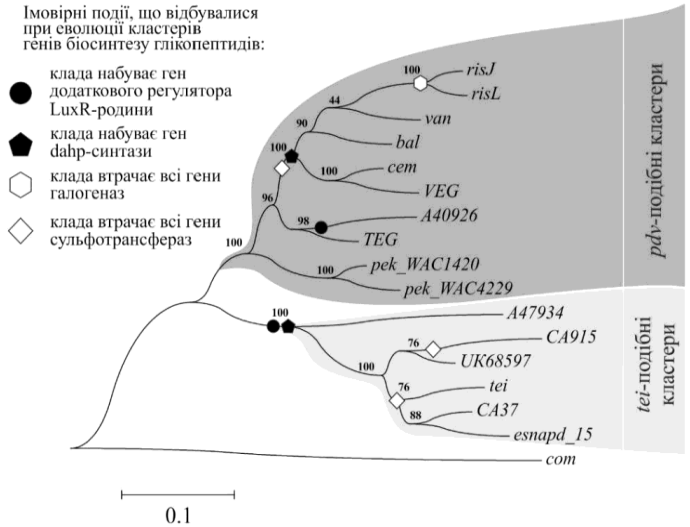


Рис. 4. Філогенетичне дерево, побудоване на основі конкатенатів генів *pdh, dpgA, hpgT*, *strR* та 7 модуля НРПС вибірки КБГ. На дереві можна розрізнити дві чітко відділені кладі, сформовані *tei*-(тейкопланін)-подібними кластерами та *pdv*-(пекіскоміцин, дальбаванцин (A40926), ванкоміцин)-подібними кластерами. Перша клада об'єднує представників I, III, IV структурних типів глікопептидів [12], друга – III та IV. Послідовність генів *com*-кластера (V тип структури) аутгрупує обидві кладі. Легенда описує основні еволюційні події, що вписуються у це дерево.

Висновки

Як на основі мультилокусного порівняння, так і на основі порівнянь окремих консервативних генів, можна стверджувати, що всі кластери генів біосинтезу глікопептидів поділяються на дві основні групи – *tei*-подібні та *pdv*-подібні (рис. 4). Перша об'єднує кластери, які кодують відомі глікопептиди тейкопланін, UK-68,597 та A47934, а також кластери *esnapd_15*, *CA37*, *CA915* із невідомими продуктами. Більшість генів біосинтезу нестандартних сполук – фегліміцину та комплестатину – надзвичайно близькі до першої групи. До другої належать кластери, що кодують біосинтез рістоцетинів, пекіскоміцинів, ванкоміцину, баліміцину, хлороеремоміцину, A40926 та кластери VEG і TEG. Філогенетичний розподіл кластерів генів цілком не збігається із структурним розподілом глікопептидів на V типів. Ми припускаємо, що структурна подібність глікопептидів, кластери генів яких філогенетично віддалені, є прикладом конвергентної еволюції. Отримане уявлення про філогенетичні зв'язки КБГ можна використати для планування дальших експериментів із комбінаторного біосинтезу нових сполук.

Література

1. Binda E., Marinelli F., Marcone G.L. Old and new glycopeptide antibiotics: action and resistance // Antibiotics (Basel). – 2014. – V. 3, N 4. – P. 572–594.
2. Medema M., Cimermancic P., Sali A., Takano E., Fischbach M.A. A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis // PLoS Comput. Biol. – 2014. – V. 10, N 12. – P. 1–12.
3. Spohn M., Kirchner N., Kulik A., Jochim A., Wolf F., Muenzer P., Borst O., Gross H., Wohlleben W., Stegmann E. Overproduction of Ristomycin A by activation of a silent gene cluster in *A. japonicum* MG417-CF17 // Antimicrob. Ag. Chemother. – 2014. – V. 58 N 10. – P. 6185–6196.
4. Donadio S., Sosio M., Stegmann E., Weber T., Wohlleben W. Comparative analysis and insights into the evolution of gene clusters for glycopeptide antibiotic biosynthesis // Mol. Genet. Genomics. – 2005. – V. 274, N 1. – P. 40–50.
5. Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Meintjes P., Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data // Bioinformatics. – 2012. – V. 28, N 12. – P. 1647–1649.
6. Tamura K., Stecher G., Peterson D., Filipski A., Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0 // Mol. Biol. Evol. – 2013. – V. 30, N 12. – P. 2725–2729.
7. 7.Shawky R., Puk O., Wietzorrek A., Pelzer S., Takano E., Wohlleben W., Stegmann E. The border sequence of the balhimycin biosynthesis gene cluster from *A. balhimycina* contains *bbr*, encoding a StrR-like pathway-specific regulator // J. Mol. Microbiol. Biotechnol. – 2007. – V. 13, N 1–3. – P. 76–88.
8. Horbal L., Kobylyanskyy A., Truman A.W., Zaburranyi N., Ostash B., Luzhetskyy A., Marinelli F., Fedorenko V. The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *A. teichomyceticus* // Appl. Microbiol. Biotechnol. – 2014. – V. 98, N 22. – P. 9295–9309.
9. Lo Grasso L., Maffioli S., Sosio M., Bibb M., Puglia A.M., Alduina R. Two master switch regulators trigger A40926 biosynthesis in *Nonomuraea* sp. strain ATCC 39727 // J. Bacteriol. – 2015. – V. 197, N 15. – P. 2536–2544.
10. Thykaer J., Nielsen J., Wohlleben W., Weber T., Gutknecht M., Lantz A.E., Stegmann E. Increased glycopeptide production after overexpression of shikimate pathway genes being part of the balhimycin biosynthetic gene cluster // Metab. Eng. – 2010. – V. 12, N 5. – P. 455–461.
11. Bischoff D., Bister B., Bertazzo M., Pfeifer V., Stegmann E., Nicholson GJ., Keller S., Pelzer S., Wohlleben W., Süssmuth R.D. The biosynthesis of vancomycin-type glycopeptide antibiotics – a model for oxidative side-chain cross-linking by oxygenases coupled to the action of peptide synthetases // Chembiochem. – 2005. – V. 6, N 2. – P. 267–272.
12. Nicolaou K., Boddy C.N., Bräse S., Winssinger N. Chemistry, Biology, and Medicine of the Glycopeptide Antibiotics // Angew. Chem. Int. Ed. Engl. – 1999. – V. 38 N 15. – P. 2096–2152.

YUSHCHUK O.[1], OSTASH B.O.[1], HORBAL L.[2], FEDORENKO V.[1]
[1] *Ivan Franko National University of Lviv,*
Ukraine, 79005, Lviv, Hrushevskoho str., 4, e-mail: v_fedorenko@franko.lviv.ua
[2] *Helmholtz-Institute for Pharmaceutical Research Saarland,*
Germany, Saarbrucken

RECONSTRUCTING THE PHYLOGENY OF GLYCOPEPTIDE BIOSYNTHETIC GENE CLUSTERS

Aim. In current study we have endeavored to analyze the phylogeny of 18 sequenced glycopeptide biosynthetic gene clusters. ***Methods.*** Standard methods of phylogenetic reconstruction were used in this research. ***Results.*** We have created sets of conservative genes, present in all selected clusters and analyzed their phylogeny. In most cases two main clades were formed on final trees. One clade always included genes from biosynthetic clusters of teicoplanin-like com-

pounds, the second – genes from biosynthetic clusters of vancomycin-like compounds. Finally, we have built a multilocus phylogeny (MLP) of selected clusters. Once again, two groups of glycopeptide gene clusters could be recognized according to the results of MLP: biosynthetic clusters of teicoplanin-like compounds and biosynthetic clusters of vancomycin-like compounds. These groups do not correlate with accepted structural classification of glycopeptides. Biosynthetic cluster of unusual antibiotic feglymycin appeared to be closely related to biosynthetic clusters of teicoplanin-like compounds. ***Conclusions.*** Based on the obtained results we proposed a possible scenario of the evolution of glycopeptide biosynthetic gene clusters.
Keywords: gene cluster, glycopeptide antibiotic, phylogeny.

Оригінал доступний за посиланням:
http://utgis.org.ua/journals/index.php/Faktory/article/view/744.

УДК 575+579.6+604 https://doi.org/10.7124/FEEO.v30.1474

ЮЩУК О.С.[1], ЖУКРОВСЬКА К.А.[1], ФЕДОРЕНКО В.О.[1]✉
Львівський національний університет імені Івана Франка,
Україна, 79005, м. Львів, вул. Грушевського, 4
✉ *viktor.fedorenko@lnu.edu.ua, +38(032) 239-44-07, +38(032) 239-44-75*

ПОШИРЕННЯ КЛАСТЕРІВ ГЕНІВ БІОСИНТЕЗУ КОМЛЕСТАТИНУ ТА ПОДІБНИХ СПОЛУК У ПРЕДСТАВНИКІВ РОДУ *STREPTOMYCES*

Мета. У цій роботі проаналізовано 2664 геномів представників роду *Streptomyces* (послідовності яких доступні в базі даних GenBank) з метою пошуку кластерів біосинтетичних генів (КБГ), подібних до КБГ комплестатину. На основі послідовностей знайдених КБГ було реконструйовано біосинтетичні шляхи, що дозволило оцінити структурну різноманітність комплестатин-подібних сполук, які походять із природніх джерел. ***Методи***. В роботі застосовано набір загальновживаних програм для пошуку КБГ та філогенетичної реконструкції. ***Результати***. Комплестатин-подібні КБГ було знайдено в 53 геномах, хоча лише 33 локалізовувались в межах одного контігу – інші КБГ були неповні або фрагментовані. Реконструкція мультилокусної філогенії 33 знайдених повних КБГ та КБГ комплестатину дозволила виокремити п'ять філогенетичних підгруп. Представники кожної підгрупи мали характерну генетичну організацію КБГ. Реконструкція потенційних біосинтетичних шляхів, кодованих у знайдених КБГ, показала, що ці КБГ можуть кодувати біосинтез деяких нових похідних комплестатину, а саме норкомплестатину, *N*-малоніл-норкомплестатину та *N*-ацетил-норкомплестатину. ***Висновки***. Комплестатин-подібні КБГ є розповсюдженими серед стрептоміцетів та можуть кодувати біосинтез нових похідних комплестатину, що заслуговує подальшого експериментального вивчення.

Ключові слова: комплестатин, кластери генів біосинтезу, стрептоміцети, майнинг геномів.

Глікопептидні антибіотики (ГПА) відомі як потужні інгібітори синтезу клітинної стінки Грам-позитивних бактерій, що зв'язують ліпід II [1]. Сьогодні в клініці використовуються два природні ГПА, а також три напівсинтетичні сполуки. До перших належить ванкоміцин і тейкопланін, що десятками років ефективно використовуються в клініці для боротьби із мультирезистентними Грам-позитивними патогенами [1]. ГПА не вичерпуються сполуками, що використовуються в медицині як протибактерійні препарати. Описано десятки різноманітних ГПА, які мали обмежене медичне застосування, або взагалі не використовувалися [1]. Усі ГПА класифікують на п'ять типів залежно від особливостей їхньої структури [1]. До типу I належать глікозильовані гептапептиди, що містять аліфатичні амінокислоти у 1 та 3 положеннях аглікону і три поперечні зшивки між ароматичними кільцями амінокислот у положеннях 2 і 4, 4 і 6, 5 і 7. Представники типу II – це глікозильовані сполуки, в яких, на відміну від ГПА типу I, всі сім амінокислот їхніх агліконів ароматичні. В агліконах ГПА типу III всі сім амінокислот ароматичні і характеризуються одною додатковою поперечною зшивкою між ароматичними кільцями амінокислот у 1 та 3 положеннях. До цього типу належать як глікозильовані, так і неглікозильовані сполуки. За структурою аглікону ГПА типу IV аналогічні до типу III, однак містять аліфатичний бічний ланцюг. ГПА типу V не глікозильовані, можуть бути як гептатак і нона-пептидами, містять різні поперечні зшивки. У строгому сенсі, ГПА типу V не є глікопептидами. Більше того, ці сполуки не здатні зв'язувати ліпід II, а крім антимікробних, для них характерні й інші біологічні активності, важливі з фармакологічної точки зору [1].

Одним із цікавих ГПА V-го типу є комплестатин, вперше виділений з *Streptomyces lavendulae* SANK 60477 [2]. Комплестатину властива слабка антимікробна активність. Донедавна вважалося, що він інгібує біосинтез жирних кислот Грам-позитивних бактерій, але більш детальний аналіз показав, що молекулярними мішенями комплестатину є аутолізини – ферменти ремоделювання та утилізації муреїну бактерій [3]. Комплестатин здатний ефективно пригнічувати формування біоплівок Грам-негативними бактеріями через пригнічення нітритредуктаз [4]. Однак, найважливішим є те, що комплестатин виявився потужним інгібітором

© ЮЩУК О.С., ЖУКРОВСЬКА К.А., ФЕДОРЕНКО В.О.

гемолізу еритроцитів, сенсибілізованих під впливом системи комплементу [2]. Комплестатин також здатний *in vitro* інгібувати розвиток вірусу імунодефіциту людини HIV-1 [5]. Варто також відзначити, що комплестатин є потужним протектором нейронів від ексайтотоксичності, що спричинена *N*-метил-D-аспартатом (NMDA) та α-аміно-3-гідрокси-5-метил-4-ізоксазолепропіоновою кислотою (AMPA). Така нейропротекторна дія комплестатину є наслідком блокування транспорту іонів Ca^{2+} після активації NMDA- та AMPA-рецепторів. Більше того, комплестатин демонстрував здатність захищати нейрони кори головного мозку від кисневого та глюкозного голодування [6]. Для комплестатину також описані певні протиракові активності [7]. Кластер біосинтетичних генів (КБГ) комплестатину з *S. lavendulae* SANK 60477 (*com*-кластер) секвеновано одним із перших серед генів ГПА [8]. Згодом показано наявність КБГ комплестатину в *S. chartreusis* AN1542 [9], *Streptomyces* sp. WAC01325 і деяких інших стрептоміцетів [10]. Можна припустити, що КБГ комплестатину є доволі поширеними серед представників роду *Streptomyces*. Однак, дотепер не проводився порівняльний геномний аналіз *com*-подібних КБГ, їх поширення та особливостей організації. Тому, метою нашої роботи був аналіз нуклеотидних послідовностей 2664 геномів стрептоміцетів, доступних сьогодні в базі даних GenBank, на наявність КБГ, подібних до *com*-кластера. Порівняння послідовностей КБГ дасть змогу оцінити різноманітність ГПА V-го типу, подібних до комплестатину, та спростить пошук нових антибіотиків.

Матеріали і методи

Скринінг геномів стрептоміцетів у пошуках КБГ, подібних до комплестатину. Нуклеотидні послідовності 2664 геномів *Streptomyces* spp. було отримано із репозиторію GenBank у вигляді файлів *.gb. Перелік усіх цих послідовностей можна знайти онлайн у Електронному Додатку 1 (ЕД1). Для пошуку *com*-подібних КБГ у відібраних нуклеотидних послідовностях було використано наступний алгоритм дій: 1) за допомогою програми MultiGeneBlast, http://multigeneblast.sourceforge.net [11] створено offline базу даних, що включала нуклеотидні послідовності з ЕД1; 2) конкатеновані послідовності всіх генів *com* (AF386507, крім генів нерибосомних пептидсинтетаз – НРПС) використано як «запит» для пошуку проти створеної бази даних в MultiGeneBlast; 3) нуклеотидні послідовності, в яких було виявлено ко-локалізовані гомологи *com*-генів, проаналізовано на наявність КБГ антибіотиків за допомогою antiSMASH 6.0 [12]; 4) отримані послідовності *com*-подібних кластерів перевірено вручну і за потреби переанотовано .

Філогенетична реконструкція. Пакет програм MEGA 10 [13] використано для створення множинних вирівнювань конкатенованих амінокислотних послідовностей білків, закодованих в 33 із знайдених КБГ, що є ортологами семи білків закодованих в *com*, а саме: AAK81822, AAK81823, AAK81829, AAK818230, AAK81831 та AAK81832. Ці послідовності можна знайти онлайн в Електронному Додатку 2 (ЕД2). MEGA 10 також використали для визначення оптимальної моделі філогенії, застосованої для реконструкції філогенетичного дерева методом максимальної правдоподібності. Такою виявилася модель Jones-Taylor-Thornton (JTT) [14] з оцінкою γ-розподілу варіацій частот між сайтами (5 дискретних γ-категорій, G5) та емпіричними частотами амінокислот (F). Достовірність топології філогенетичного дерева, побудованого за цією моделлю, була оцінена за допомогою бутстреп-тесту (1000 повторів).

Для не згаданих вище рутинних маніпуляцій та аналізів амінокислотних та нуклеотидних послідовностей використовували GENEIOUS 4.8.5 [15].

Результати та обговорення

Розповсюдження com-подібних КБГ в геномах стрептоміцетів. Ми проаналізували нуклеотидні послідовності 2664 геномів представників роду *Streptomyces*. Серед них були як повністю зібрані геноми, так і чернетки різної якості. КБГ, подібні до *com*, знайдено в 53 геномах. Так як було проаналізовано нуклеотидні послідовності різної якості, лише у 33 випадках ми отримали повні послідовності КБГ (див. Електронний Додаток 3, ЕД3) – розміщені у межах одного контігу. У решти 20 випадків КБГ були фрагментовані, або містили помилки секвенування (ЕД3). У цілому, можна вважати, що *com*-подібні КБГ є доволі поширеними у геномах стрептоміцетів. Для дальшого аналізу ми використали тільки ті 33 КБГ, які знайдені в межах одного контігу.

Особливості організації знайдених com-подібних КБГ. Для того, щоб оцінити різноманіт-

тність організації 34 *com*-подібних КБГ (включно із *com*), ми спершу вирішили реконструювати мульти-локусну філогенію цих КБГ, щоб виявити кореляцію такого філогенетичного фреймворку з особливостями їх організації. З'ясовано, що всі 34 КБГ діляться на 5 добре відділених філогенетичних підгруп (рис. 1). Аналіз організації КБГ кожної з підгруп показав, що початкове припущення було правильним, і кожна з філогенетичних підгруп включає в себе кластери з характерною організацією. Далі охарактеризовано КБГ підгруп у порядку від базальної підгрупи V до кронової підгрупи I (рис. 1).

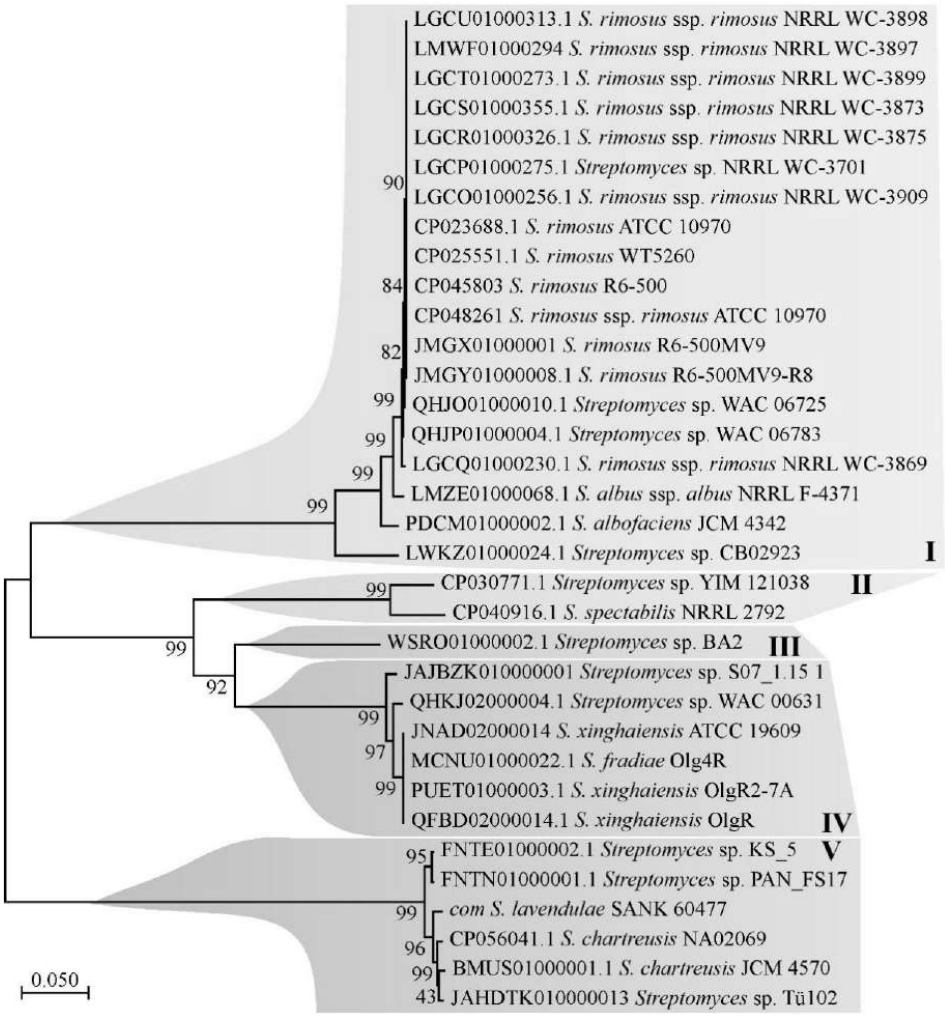


Рис. 1. Філогенетичне дерево конкатенованих амінокислотних послідовностей семи білків (див. розділ «Методи»), що закодовані в 33 *com*-подібних КБГ ГПА, а також в самому *com*-кластері. На дереві виділено 5 клад із високою бутстреп-підтримкою, що відповідають п'яти філогенетичним підгрупам КБГ. Алгоритм побудови цього дерева описано в розділ «Методи». Планка масштабу показує кількість амінокислотних замін на один сайт.

Підгрупа V. До підгрупи V входить *com*-кластер. Усі кластери цієї підгрупи організовані подібно до нього (рис. 2а). Виявилося, що КБГ підгрупи V містять у своєму складі 6 генів, раніше не описаних в кластері *com*. Це два гени двох-субодиничного ABC-транспортера, два гени двох-компонентної регуляторної системи, а також два гени, що кодують білки, які не задіяні у біосинтезі ГПА.

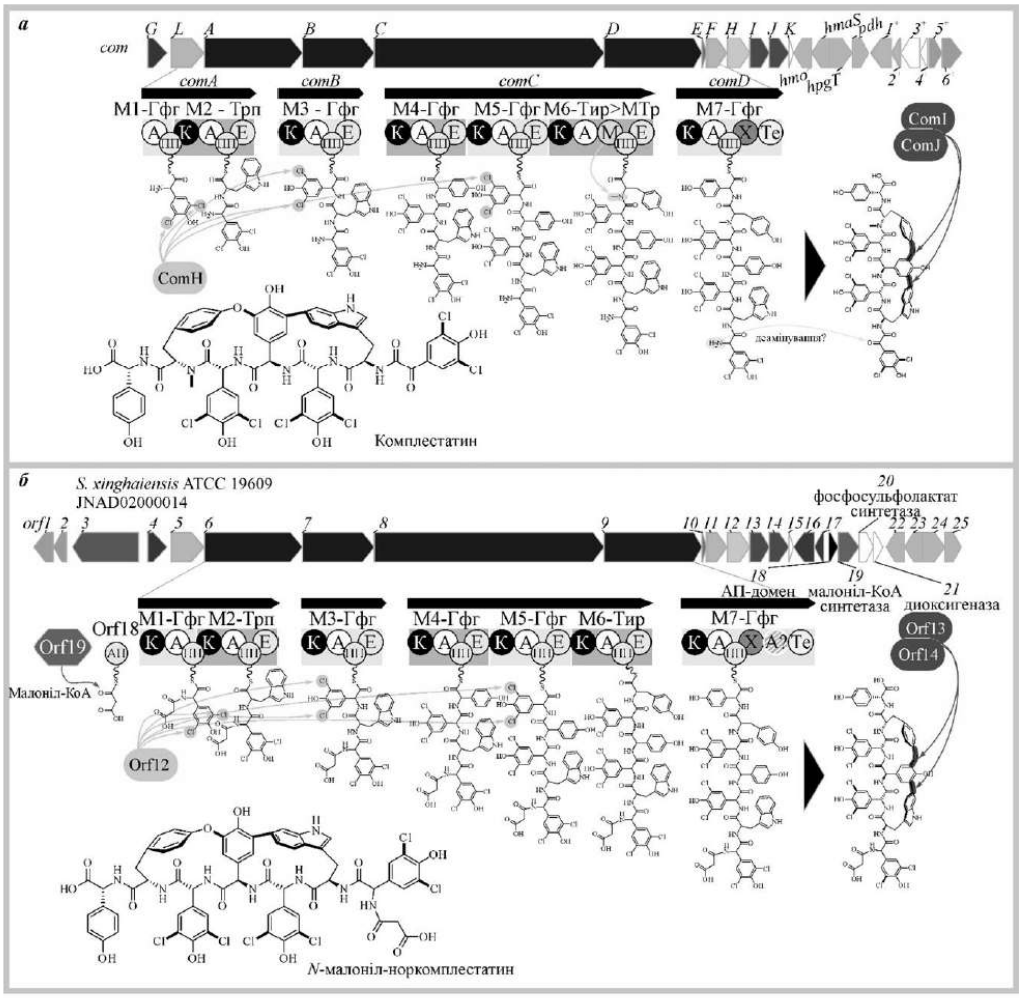


Рис. 2. Інтегровані моделі КБГ та біосинтетичних шляхів комплестатину (а) і *N*-малоніл-норкомплестатину (б), що належать до філогенетичних підгруп V та IV КБГ ГПА, подібних до комплестатину. Умовні позначення розшифровано в легенді до рисунку, деталі обговорюються у тексті статті.

Послідовність *com*-кластера, опублікована раніше [8], містила довгі неанотовані 5'- та 3'-фланги. Анотувавши їх за допомогою antiSMASH, ми виявили в кластері *com* ці ж 6 додаткових генів, які назвали *com1*$^{+}$-*6*$^{+}$ (рис. 2а). Це показує гомогенність структурної організації КБГ підгрупи V, зручність створеного філогенетичного фреймворку і дає змогу зробити важливе доповнення до розуміння організації кластера *com*.

Підгрупа IV. КБГ підгрупи IV (рис. 1) мають організацію, відмінну від *com*-кластера. Зокрема, гени двокомпонентної регуляторної системи містяться не на 3'-кінці кластера, а на його 5'-кінці, а гени двосубодиничного ABC-транспортера відсутні (рис. 2б). Важливо, що у складі КБГ з'явилося сім додаткових генів: ген LuxR-подібного транскрипційного регулятора, два гени транспозаз, ген окремого домена ацил-переносника, гени малоніл-КоА-синтетази, фосфосульфолактонат-синтетази та диоксигенази (рис. 2б). Відмінності спостерігалися і в організації генів НРПС. Зокрема, в гені *comC,* який кодує модулі 4-5-6 комплестатинової НРПС, міститься ділянка, яка кодує унікальний метилтрансферазний домен модуля 6. Функціонування цього домену зумовлює те, що тирозин у шостому положенні гептапептиду комплестатину перебуває у метильованій формі (рис. 2а) [8]. Натомість, ортологи *comC* в кластерах підгрупи IV кодують модуль 6 НРПС, в якому відсутній метилтрансферазний домен. Важливо, що перший модуль НРПС, закодованих в КБГ підгрупи IV, починається із додаткового конденсаційного домена, нагадуючи структуру НРПС іншого ГПА V-го типу – корбоміцину [3]. На кінець, модуль 7 НРПС має у своєму складі додатковий – нефункціональний – аденілазний домен (рис. 2б). Сукупність цих особливостей організації КБГ підгрупи IV дає змогу передбачити структуру ймовірного продукту, який мав би помітно відрізнятися від комплестатину. Зокрема, наявність генів окремого домена ацил-переносника та малоніл-КоА-синтетази, а також присутність додаткового конденсаційного домена в модулі 1 НРПС, дозволяє, за аналогією із корбоміцином [3], припустити, що продукти КБГ підгрупи IV є гібридними НРПС-ПКС молекулами, початковою одиницею у біосинтезі яких виступає малоніл-КоА (рис. 2б). Це, а також відсутність закодованого метилтрансферазного домена, дозволяє говорити про продукт КБГ підгрупи IV як про *N*-малоніл-норкомплестатин. Подібного роду ГПА досі описано не було, тому цей аспект заслуговує подальшого експериментального вивчення.

Підгрупа III. Ця підгрупа представлена єдиним КБГ із *Streptomyces* sp. BA2 (рис. 1). Цікаво, що він більш подібний до *com*-кластера, ніж до кластерів підгрупи V. Відміни – це розміщення генів двох-компонентної регуляторної системи (ортологів *com5*$^{+}$-*6*$^{+}$) на 5'-кінці кластера і поява двох додаткових генів: LuxR-регулятора – ортолога генів LuxR-регуляторів КБГ підгрупи IV, і 3-деокси-D-арабіногептулозонат 7-фосфат (ДАГП) синтази (рис. 3а). Гени, подібні до гена цієї синтази, поширені в КБГ ГПА I-IV типів, але раніше не описувалися у складі *com*-подібних кластерів [1]. Цей кластер кодує НРПС, що не містить метилтрансферазного домена в модулі 6. Ці особливості дають підставу припустити, що він мав би кодувати біосинтез норкомплестатину (рис. 3а).

Підгрупа II. Два представники цієї підгрупи, а саме кластери зі *Streptomyces sp.* YIM 121038 і *S. spectabilis* NRRL 2792 (рис. 1), подібні до КБГ підгрупи IV. Вони також кодують біосинтез гіпотетичного *N*-малоніл-норкомплестатину (рис. 3б). Серед відмінностей можна назвати відсутність генів транспозаз, наявність генів двох-компонентного ABC-транспортера, які містяться також у КБГ підгруп V та III, а також певні різниці в орієнтації деяких генів (рис. 3б).

Підгрупа I. КБГ філогенетичної підгрупи I (рис. 1) демонструють ще один різновид організації. У цілому схожі до *com*-кластера, вони не містять ортологів генів *com1*$^{+}$-*6*$^{+}$, але несуть ортологи генів LuxR-подібного транскрипційного регулятора, присутні також в КБГ підгруп IV-II, а також ДАГП синтази, знайденого в КБГ підгрупи III (рис. 3в). Цікаво, що НРПС, закодовані в КБГ підгрупи I, містять додатковий перший конденсаційний домен в модулі 1, хоча і не мають генів окремого домена ацил-переносника та малоніл-КоА-синтетази, які наявні в КБГ підгруп IV та II. Цей випадок також нагадує ситуацію з біосинтезом корбоміцину [3]. Екстраполюючи його, можна припустити, що біосинтез ГПА, кодований кластерами підгрупи I, потребує для ініціації окремого домена ацил-переносника, що не закодований в КБГ та функціонує *in trans*. Тому, КБГ підгрупи I найімовірніше кодують ще один новий ГПА, а саме *N*-ацетил-норкомплестатин (рис. 3в).

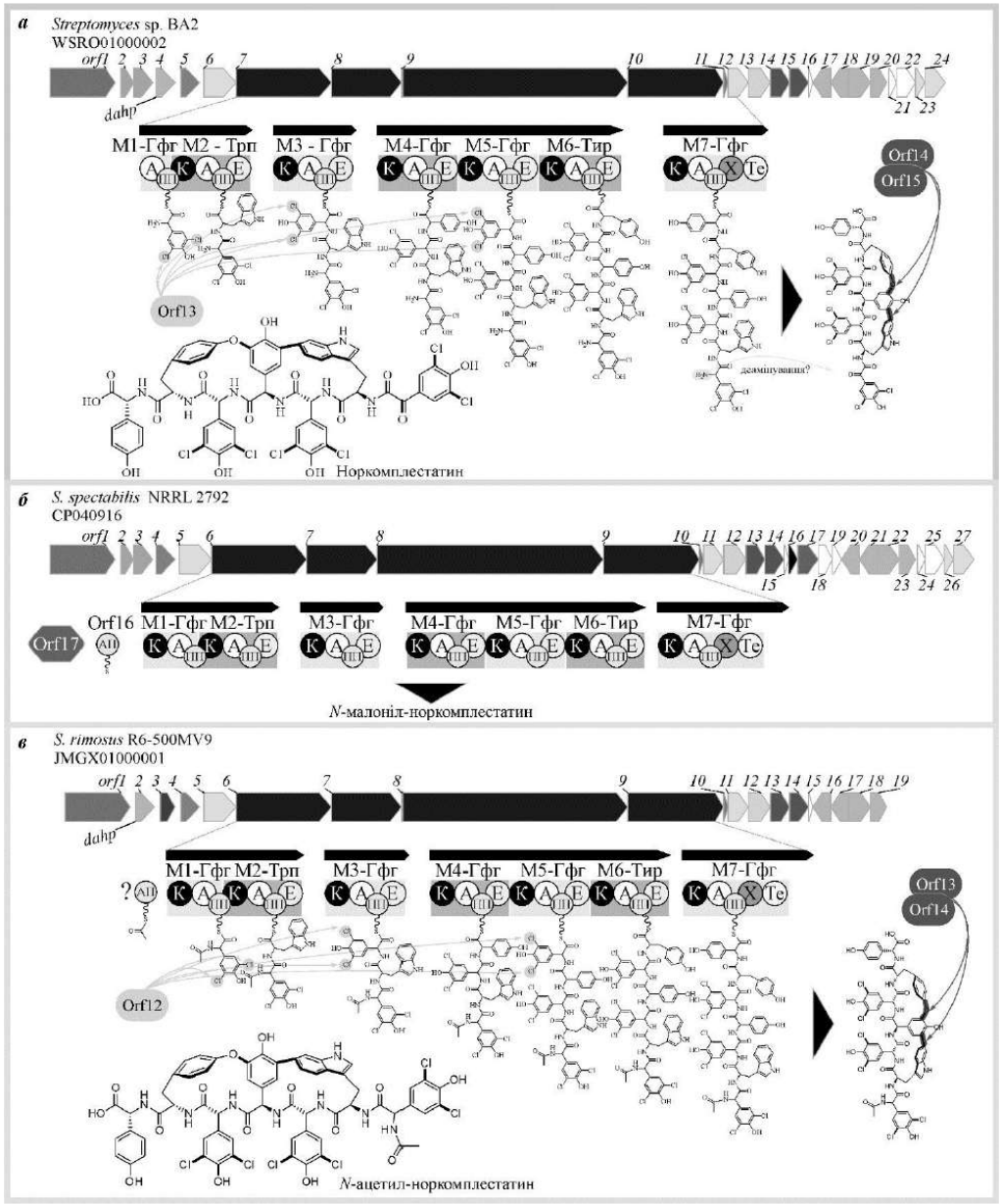


Рис. 3. Інтегровані моделі КБГ та біосинтетичних шляхів норкомплестатину (а), *N*-малоніл-норкомплестатину (б) та *N*-ацетил-норкомплестатину, що належать до філогенетичних підгруп III, II та I КБГ ГПА, подібних до комплестатину. Див. легенду до рис. 2; деталі рисунку обговорюються у тексті статті.

Висновки

У роботі проведено скринінг та аналіз КБГ стрептоміцетів, подібних до кластера біосинтезу комплестатину. Виявлено 33 повних нових кластерів, які ділять на п'ять філогенетичних підгруп. Кожна з цих підгруп характеризується особливою організацією кластерів. Деякі знахідки заслуговують окремого обговорення.

По-перше, передбачення біосинтетичних шляхів, закодованих в знайдених КБГ, вказує на те, що, ймовірно, існує декілька аналогів комплестатину, а саме норкомплестатин, *N*-малоніл-норкомплестатин і *N*-ацетил-норкомплестатин. Таке теоретичне передбачення спонукає до подальшого експериментального підтвердження продукції цих сполук відповідними штамами. *По-друге*, цікавим є факт наявності гена LuxR-подібного транскрипційного регулятора в КБГ усіх підгруп, крім першої, до якої належить і *com*-кластер. Взагалі, основним шлях-специфічним регулятором біосинтезу ГПА є StrR-подібний регулятор, що закодований в кожному без винятку КБГ ГПА, у даному випадку це ComG та його ортологи (рис. 2, 3). Однак, деякі КБГ ГПА, зокрема тейкопланіну та A40926 [1], містять додатковий ген транскрипційного регулятора LuxR-родини. Привабливим виглядає припущення, що LuxR-подібні регулятори КБГ підгруп I-IV також беруть участь в регуляції відповідних біосинтетичних шляхів. *По-третє*, порівняння організації КБГ різних підгруп дає змогу висловити припущення про еволюцію цих КБГ. Зокрема, ймовірним є те, що протокластер підгруп I-IV вже мав ген ДАГП синтази, втрачений у КБГ підгруп II та IV. Також схоже, що гени малоніл-КоА-синтази та окремого домена ацил-переносника потрапили до протокластера підгруп I-IV внаслідок горизонтального генного переносу, про що свідчить наявність генів транспозаз відразу поруч із ним, а пізніше були втрачені в деяких підгрупах. На кінець, наявність унікального метилтрансферазного домена в модулі 6 НРПС, закодованих в КБГ підгрупи V, є, мабуть, унікальним еволюційний здобутком саме цієї підгрупи. У цілому, проведений тут аналіз спонукає нас до подальшого експериментального вивчення комплестатин-подібних КБГ із різних філогенетичних підгруп, що дасть можливість описати нові ГПА типу V.

Електронні додатки. Файли електронних додатків, згаданих у статті (ЕД1, ЕД2, ЕД3), доступні онлайн: https://data.mendeley.com/datasets/h443z9ztz5/1, doi: 10.17632/h443z9ztz5.1.

Ця робота була виконана в межах держбюджетної теми Міністерства освіти і науки України БГ-09Ф.

References

1. Yushchuk O., Ostash B. Glycopeptide antibiotics: genetics, chemistry, and new acreening approaches. in *Natural Products from Actinomycetes: Diversity, Ecology and Drug Discovery*. 2022. P. 411–444. doi: 10.1007/978-981-16-6132-7_16.
2. Kaneko I., Kamoshida K., Takahashi S. Complestain, a potent anti-complement substance produced by *Streptomyces lavendulae*. I. Fermentation, isolation and biological characterization. *J. Antibiot. (Tokyo)*. 1989. Vol. 42. P. 236–241. doi: 10.7164/antibiotics.42.236.
3. Culp E.J., Waglechner N., Wang W., Fiebig-Comyn A.A., Hsu Y.P., Koteva K., Sychantha D., Coombes B.K., van Nieuwenhze M.S., Brun Y.V. et al. Evolution-guided discovery of antibiotics that inhibit peptidoglycan remodelling. *Nature*. 2020. Vol. 578. P. 582–587. doi: 10.1038/s41586-020-1990-9.
4. Park Ji-Su, Choi Ha-Young, Kim Won-Gon. The nitrite transporter facilitates biofilm formation via suppression of nitrite reductase and is a new antibiofilm target in *Pseudomonas aeruginosa*. *mBio*. 2020. Vol. 11. P. e00878-20. doi: 10.1128/mBio.00878-20.
5. Momota K., Kaneko I., Kimura S., Mitamura K., Shimada K. Inhibition of human immunodeficiency virus type-1-induced syncytium formation and cytopathicity by complestatin. *Biochem. Biophys. Res. Commun.* 1991. Vol. 179. P. 243–250. doi: 10.1016/0006-291X(91)91361-F.
6. Seo S.Y., Yun B.S., Ryoo I.J., Choi J.S., Joo C.K., Chang S.Y., Chung J.M., Oh S., Gwag B.J., Yoo I.D. Complestatin is a noncompetitive peptide antagonist of N-methyl-D-aspartate and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid/kainate receptors: secure blockade of ischemic neuronal death. *J. Pharmacol. Exp. Ther.* 2001. Vol. 299. P. 377–384.
7. Kim E.C., Yun B.S., Ryoo I.J., Min J.K., Won M.H., Lee K.S., Kim Y.M., Yoo I.D., Kwon Y.G. Complestatin prevents apoptotic cell death: inhibition of a mitochondrial caspase pathway through AKT/PKB activation. *Biochem. Biophys. Res. Commun.* 2004. Vol. 313. P. 193–204. doi: 10.1016/j.bbrc.2003.11.104.
8. Chiu H.T., Hubbard B.K., Shah A.N., Eide J., Fredenburg R.A., Walsh C.T., Khosla C. Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. *Proc. Natl. Acad. Sci. U. S. A.* 2001. Vol. 98. P. 8548–8553. doi: 10.1073/pnas.151246498.
9. Park O.K., Choi H.Y., Kim G.W., Kim W.G. Generation of new complestatin analogues by heterologous expression of the complestatin biosynthetic gene cluster from *Streptomyces chartreusis* AN1542. *ChemBioChem*. 2016. Vol. 17. P. 1725–1731. doi: 10.1002/cbic.201600241.
10. Waglechner N., McArthur A.G., Wright G.D. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat. Microbiol.* 2019. Vol. 4. P. 1862–1871. doi: 10.1038/s41564-019-0531-5.
11. Medema M.H., Takano E., Breitling R. Detecting sequence homology at the gene cluster level with multigeneblast. *Mol. Biol. Evol.* 2013. Vol. 30. P. 1218–1223. doi: 10.1093/molbev/mst025.
12. Blin K., Shaw S., Kloosterman A.M., Charlop-Powers Z., Van Wezel G.P., Medema M.H., Weber T. AntiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 2021. Vol. 49. P. W29–W35. doi: 10.1093/nar/gkab335.

13. Kumar S., Stecher G., Li M., Knyaz C., Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 2018. Vol. 35. P. 1547–1549. doi: 10.1093/molbev/msy096.
14. Jones D.T., Taylor W.R., Thornton J.M. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 1992. Vol. 8. P. 275–282. doi: 10.1093/bioinformatics/8.3.275.
15. Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012. Vol. 28. P. 1647–1649. doi: 10.1093/bioinformatics/bts199.

YUSHCHUK O.S., ZHUKROVSKA K.A., FEDORENKO V.O.
Ivan Franko National University of Lviv,
Ukraine, 79005, Lviv, Hrushevskoho str., 4

DISTRIBUTION OF BIOSYNTHETIC GENE CLUSTERS FOR COMPLESTATIN AND RELATED COMPOUNDS IN *STREPTOMYCES* SPP.

Aim. In current work we have analyzed 2664 genomes of *Streptomyces* spp. (publically available from GenBank) searching for complestatin-like biosynthetic gene clusters (BGCs). Then we aimed to reconstruct *in silico* putative biosynthetic pathways, encoded within these BGCs, estimating the chemical variability of complestatin-like compounds from natural sources. ***Methods.*** Widely accepted genome-mining tools and approaches for phylogenetic reconstruction were utilized in this work. ***Results.*** 53 genomes of *Streptomyces* spp. were found to contain complestatin-like BGCs, although only 33 BGCs were found within one contig – others were partial or highly fragmented. Reconstruction of multi-locus phylogeny for 33 found BGCs and complestatin BGC allowed to divide all these BGCs into five phylogenetic subgroups. Representatives of each subgroup exhibited characteristic organization of corresponding BGCs. Reconstruction of putative biosynthetic pathways allowed us to predict that discovered BGCs might potentially code the biosynthesis of new complestatin derivatives: norcomplestatin, *N*-malonyl-norcomplestatin, and *N*-acetyl-norcomplestatin. ***Conclusions.*** Complestatin-like BGCs are widely distributed among *Streptomyces* spp. and might encode novel complestain derivatives, which merits further experimental investigation.
Keywords: complestatin, biosynthetic gene clusters, *Streptomyces* spp., genome-mining.

Оригінал доступний за посиланням:
https://utgis.org.ua/journals/index.php/Faktory/article/view/1474.
Електронні додатки до статті доступні за посиланням:
https://data.mendeley.com/datasets/h443z9ztz5/1.

РОЗДІЛ 7 Регуляція продукції антибіотиків в *Streptomyces cyanogenus* S136 – моделі активації мовчазних кластерів біосинтетичних генів

У попередньому розділі ми отримали докази того, що стрептоміцети потенційно є багатим джерелом ГПА типу V, подібних до комплестатину. Тим не менше, ці антибіотики досі не були описані, даючи змогу припустити, що знайдені КБГ є мовчазними. КБГ актиноміцетів залишаються мовчазними із різних причин, зокрема через низьку або відсутню транскрипцію відповідних генів [110, 131]. Відповідно, необхідно вивчати підходи до активації мовчазних КБГ.

У цьому розділі ми вивчали продукцію антибіотиків у *Streptomyces cyanogenus* S136, продуценті перспективного протиракового полікетидного антибіотика ландоміцину А [56, 132]. КБГ ландоміцину – *lan* – відомий, а особливості шлях-специфічної регуляції та механізмів його біосинтезу вже вивчалися [24, 133, 134]. *S. cyanogenus* привабив нас як модель вивчення мовчазних КБГ із низки причин: (1) даний штам є доступним для генно-інженерних маніпуляцій, він невибагливий до умов культивування та швидко росте; (2) попередні роботи давали певні натяки, що цей штам здатний продукувати певні вторинні метаболіти, крім ландоміцинів [24]; (3) нарешті, штам *S. cyanogenus* S136 дикого типу не здатний до споруляції, що може свідчити про зміни в глобальних регуляторних механізмах біосинтезу антибіотиків і морфогенезу.

Цільовим знаряддям, роль якого для активації продукції антибіотиків у *S. cyanogenus* вивчалася, став глобальний позитивний регуляторний ген *adpA*. Він привернув нашу увагу, оскільки є присутнім в геномах багатьох стрептоміцетів і часто впливає на транскрипцію генів біосинтезу антибіотиків, як безпосередньо [98, 135, 136], так і опосередковано [67, 70, 137–140]. AdpA виглядає як привабливе знаряддя і в тому стосунку, що здатен зв’язувати дуже вироджені операторні послідовності [66] і, в такий спосіб, має величезний регулон [141]. Ми припустили, що надекспресія такого транскрипційного фактора може привести до зв’язування з численними (ймовірно, неспецифічними за фізіологічних умов) ділянками хромосоми, впливаючи на

експресію багатьох генів і оперонів. Тобто, надекспресія генів *adpA* може бути використана для високопродуктивної активації різних КБГ у штамах з невідомим геномним фоном.

7.1 Секвенування та характеристика генома *S. cyanogenus* S136

Щоби отримати уявлення, які саме КБГ несе хромосома *S. cyanogenus* S136, ми вирішили секвенувати геном цього штаму. Секвенування генома було здійснено за допомогою комбінації Illumina HiSeq і Oxford Nanopore Technologies (ONT) GridION. Отримана послідовність репрезентувала лінійну хромосому розміром 8773899 п.н., на якій було анотовано 7700 генів, з яких: 7590 потенційно кодують білки, 91 ген тРНК, 1 ген тмРНК і 6 оперонів генів 16S-5S-23S рРНК. Усереднений ГЦ склад генома становить 71,8%.

Для того, щоб оцінити потенціал *S. cyanogenus* до продукції вторинних метаболітів ми проаналізували отриману послідовність геному за допомогою antiSMASH 5.0 [116]. В результаті, було ідентифіковано 41 КБГ для різноманітних вторинних метаболітів (включно із КБГ ландоміцину А). Цікаво, що геном ніс КБГ спорового пігмента – сполуки полікетидної природи, що інкрустує спори стрептоміцетів, хоча, як вже зазначалося, сам *S. cyanogenus* до споруляції не здатний.

В підсумку, геном *S. cyanogenus* S136 має розмір, типовий для інших стрептоміцетів і володіє значним потенціалом до продукції різноманітних вторинних метаболітів, зокрема полікетидів.

Результати цього підрозділу опубліковано в [142].

3 Biotech (2021) 11:282
https://doi.org/10.1007/s13205-021-02834-4

ORIGINAL ARTICLE

Complete genome sequence of *Streptomyces cyanogenus* S136, producer of anticancer angucycline landomycin A

Pavlo Hrab[1] · Christian Rückert[2] · Tobias Busche[2] · Iryna Ostash[1] · Jörn Kalinowski[2] · Victor Fedorenko[1] · Oleksandr Yushchuk[1] · Bohdan Ostash[1]

Received: 3 April 2021 / Accepted: 7 May 2021 / Published online: 21 May 2021
© King Abdulaziz City for Science and Technology 2021

Abstract
Streptomyces cyanogenus S136 is the only known producer of landomycin A (LaA), one of the founding members of angucycline family of aromatic polyketides. LaA displays potent anticancer activities which has made this natural product a target of numerous chemical and cell biological studies. Little is known about the potential of S136 strain to produce other secondary metabolites. Here we report complete genome sequence of LaA producer and how we used this sequence to evaluate for this species its phylogenetic position and diversity of gene clusters for natural product biosynthesis.

Keywords Landomycin A · *Streptomyces cyanogenus* · Genomics

Introduction

Streptomyces cyanogenus S136 (=DSM5087) has first become known in 1990 as a producer of orange-colored angucycline compounds of the landomycin family (Henkel et al. 1990). Since that time S136 remains the only natural source of landomycin A (LaA), the largest member of the family, as well as one of the most potent angucyclines in terms of activity against a range of cancer cell lines (Ostash et al. 2009). An in-depth understanding of genetics of *S. cyanogenus* is believed to provide more facile access to landomycins for various cell biology studies (Yushchuk et al. 2019). Besides the published sequences of biosynthetic gene clusters for LaA (Westrich et al. 1999; Rebets et al. 2003), and, recently, polyene lucensomycin (Yushchuk et al. 2021), nothing is known about the genetic capacity of *S. cyanogenus* S136 for the production of other secondary metabolites. Furthermore, the relatedness of S136 to the other streptomycetes and information on global regulators of LaA biosynthesis remain vestigial. As a necessary step towards addressing these issues, here we report generation and bioinformatics analysis of complete S136 genome. Our data reveal that S136 strain possesses an unusual set of 16S rRNA genes and a significant number of natural product biosynthetic gene clusters (BGCs). The latter is in line with what is observed for other *Streptomyces* spp. (Nett et al. 2009). The generated sequence was used as a test case to compare different computational tools for BGC mining. Results of our findings are presented and discussed here.

✉ Bohdan Ostash
bohdan.ostash@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 4 Hrushevskoho st., Rm 102, Lviv 79005, Ukraine

[2] Technology Platform Genomics, CeBiTec, Bielefeld University, Sequenz 1, 33615 Bielefeld, Germany

Materials and methods

Genomic DNA manipulations, sequencing and initial annotation

Streptomyces cyanogenus S136 (DSM5087) was kindly provided by Prof. A. Bechthold (Freiburg University, Germany). Its genomic DNA was purified from a 24 h culture grown at 30 °C in tryptic soy broth according to salting out procedure N4 as described in (Kieser et al. 2000). DNA concentrations and quality were determined using Trinean Xpose (Gentbrugge, Belgium) and Agilent RNA Nano 6000 kit on Agilent 2100 Bioanalyzer (Agilent Technologies, Böblingen, Germany). The DNA sample that has passed all quality control checks was used to make two WGS libraries, one for Illumina (TruSeq DNA PCR-Free

Kit) and one for Oxford Nanopore Technologies (ONT) sequencing (SQK-LSK109 ligation sequencing kit). The Illumina library was sequenced on HiSeq platform, and ONT library was subjected to two GridION runs. Further data handling and hybrid assembly was carried out exactly as described in (Tippelt et al. 2020). The assembled S136 genome was annotated with Prokka v1.11 (Seeman et al. 2014) and deposited into GenBank under accession number CP071839.

Agar plug assay of antibacterial activity

S. cyanogenus S136 and its LaA-nonproducing mutant ΔlanI7 (Rebets et al. 2008) were grown for 5 days at 30 °C on following agar plates: SMMS, R5, SFM (Kieser et al. 2000) and ISP3 (g/L: oatmeal flour - 34, agar - 18, tap water to 1 L, pH prior to sterilization 8.0). The agar plugs were cut off the lawns and assayed against *Bacillus cereus* ATCC19637 as described in (Yushchuk et al. 2021).

Phylogenetic reconstruction

Phylogenetic analysis (maximum likelihood algorithm) was performed on concatenated amino acid sequences of conserved proteins using IQtree software on CIPRES server [www.phylo.org]. List of conserved *S. coelicolor* genes used in this work (Fig. S1, Electronic Supplementary Materials, ESM), was obtained from work of Gao and Gupta (2012). Using reciprocal best BLASTP hit (RBH) strategy, we identified orthologues in 82 complete *Streptomyces* genomes (species and corresponding GenBank accession numbers are given in Table S1), as well as in *S. cyanogenus* genome. BLAST v 2.9.0 minimum parameters to consider RBH orthologues were as follows: E-value $\leq 1e-10$, identity percentage $\geq 30\%$, coverage per high-scoring pair, no less than 50%. Evolutionary model LG + I + F + G4 was selected as the most optimal one for our dataset on the basis of minimal BIC score. Reliability of topology of the resulting trees was estimated with SH-aLRT test. For 16S phylogenetic tree GTR model was used for single gene nucleotide sequences.

Additional tools for *S. cyanogenus* genome annotation

To find *oriC* and replication-related genes, Ori-Finder2 server (Luo et al. 2019) was used. Pseudofinder (Syberg-Olsen et al. 2020) was used to predict pseudogenes. BlastKOALA and GhostKOALA (Kanehisa et al. 2016) were applied for K number assignment. Average nucleotide identity (ANI) calculator was accessed from http://enveomics.ce.gatech.edu/ani/.

Tools used for analysis of secondary metabolite BGCs

Analysis of secondary metabolite BGC was performed with antiSMASH (Blin et al. 2019) with subsequent manual inspection. If region contained several putative BGCs (further referred to as subclusters in this paragraph), each of them was extracted from the region and re-annotated with antiSMASH. If similarity score of a subcluster to known BGC was higher than the score of the region, the subcluster was considered an independent BGC. If the similarity score was high (> 60%), BGC borders assignment was done using comparison to the known BGCs from MiBIG (Kautsar et al. 2020). For those high-scoring BGCs additional BLAST run was performed with reciprocal hits retrieval. There were cases when similarity score was below 60% or similarity cannot be inferred on the basis of resemblance to the known BGC. In these cases, Cluster Blast results were applied to retrieve the shared BGC block with an assumption, that this block is genuine BGC, taking into the synteny between the latter and the other automatically annotated regions. For the BGCs that can be compared to the known ones, manual inspection was performed to highlight missing genes, using publicly available articles where the reference BGCs were described.

Default setting were applied when we run S136 genome through DeepBGC, SEMPI, PRISM for BGC annotation, RRE-Finder, RiPPMiner, DeepRiPP, BAGEL4, NeuRiPP for ribosomally produced peptides (RiPP) annotation, ARTS and BiG-FAM for additional annotation. For RRE-Finder and NeuRiPP runs the S136 genome was re-annotated with Prodigal short. The DeepBGC data was filtered on such variables: deepbgc score > 50, cluster_type score > 50, activity score > 50, # domains > 5, # biodomains > 1, # proteins > 1. For NeuRiPP analysis we used the CNN-parallel architecture, as it yielded the best results, according to the published data (Los 2019).

Results

General structure of *S. cyanogenus* S136 genome

A single high-quality (Phred score > 30) 8,773,899-bp long contig of *S. cyanogenus* S136 genome was generated as a result of sequencing and assembly strategy outlined in the Methods section. The contig was subsequently analyzed with BUSCO software (Seppey et al. 2019) to assess the completeness of the generated sequence. This analysis

showed that the sequence truly belongs to the *Streptomycetales* lineage with two missing and one fragmented nearly-universal single copy orthologues (out of 1579; Fig. S2). Given the result, the sequence can be confidently considered a complete genome. Computed features of S136 genome are presented in Table 1 in comparison with several selected *Streptomyces* species. We included *S. reticuli* into the comparison because it is the closest neighbor of S136 strain in the phylogenetic tree (see next paragraphs).

Several peculiarities of S136 genome are worth further comments. Regardless of the search algorithm, S136 genome appears to possess tRNA gene set larger than in the other *Streptomyces* spp. genomes (Table 1). In the same time, differences between Prokka and tRNAscan-SE results show that this issue is worth studying in more detail. For example, after cursory analysis we revealed that some of the tRNA genes annotated by Prokka actually are parts of the protein-coding sequences. Out of 7590 annotated S136 CDS, 2534 fell into one of the KEGG categories. Results of BlastKOALA and GhostKOALA are similar with 33.4% and 32.2% of annotation success rate (Fig. S3, ESM). There was no observed GC skew in 960-bp *oriC* region (GC = 66.35%); nine genes involved in replication are scattered over the genome. We found 295 ORFs too short to be classified as genes, and 93 ORFs interrupted with stop codons (Syberg-Olsen et al. 2020). In total 388 elements could be classified as pseudogenes, which are evenly spread over the genome (Fig. S4, ESM).

Phylogenetic position of S136

Six genes for 16S rRNA were found within the genome of *S. cyanogenus* S136. Five of these genes were identical, the sixth one (namely S1361_18950) diverged significantly, sharing only 97.4% of nucleotide sequence identity with the other homologues (Fig. 1). Notably, such identity is lower than general taxonomic threshold, often used for species discrimination (98.7%, Yarza et al. 2014).

Table 1 Features of *S. cyanogenus* genome in comparison with model *Streptomyces*

Organism	Genome length (Mbp)	GC (%)	CDS	Average CDS length (bp)	Average coding density (%)	rRNA operons	tRNA	tRNA (tRNAscan-SE)[a]
S. cyanogenus	8.8	71.84	7590	992.09	87.07	6	90[a]	79
S. coelicolor	8.7	72.12	7767	992.03	88.9	6	65	72
S. bingchenggensis	11.9	70.76	10,096	1046.8	88.54	5	66	76
S. albidoflavus	6.8	73.32	5895	1014.93	87.45	7	66	69
S. xiamenensis	5.92	72.02	5348	987.43	88.58	5	56	61
S. reticuli	8.3	72.72	7411	990.29	87.85	6	72	74

[a]As annotated with barrnap in Prokka

[a]As annotated with barrnap in Prokka

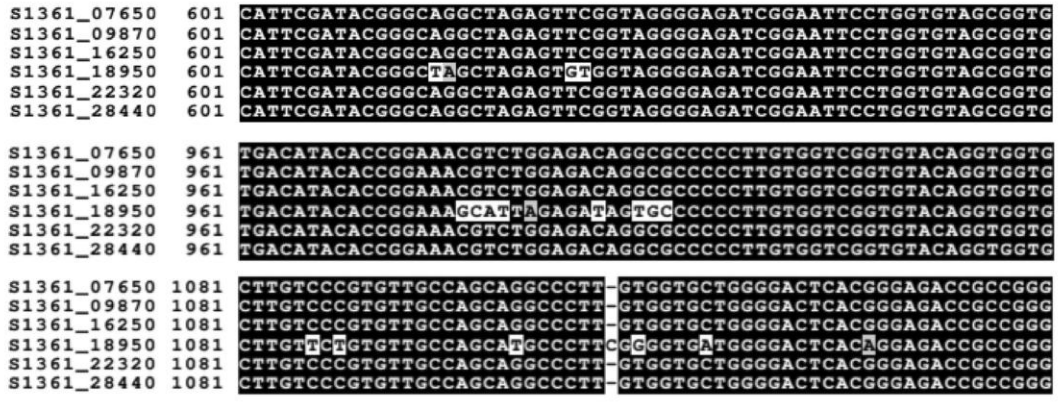


Fig. 1 Alignment of 480-bp segments of six 16S rRNA genes of *S. cyanogenus* encompassing variable V4-V6 region of 16S rRNA gene. Gene S1361_18950 is the most divergent one

We then first reconstructed the phylogeny of S136 strain and 82 other *Streptomyces* spp. (Table S1) using nucleotide sequence of S1361_07650 (one of the five identical 16S rRNA genes). The other tree was reconstructed using nucleotide sequence of the divergent S136 16S rRNA gene S1361_18950. The phylogenetic position of S136 appeared to be different in both trees, clearly depending on the 16S rRNA being used for the reconstruction (Fig. 2). *S. punicis-cabiei* TW1S1 appeared to be a sister clade for S136 strain when S1361_18950 was used for reconstruction, whereas *S. pluripotens* MUSC135 was found to be the closest to S136 in the other reconstruction. Indeed, the 16S rRNA gene

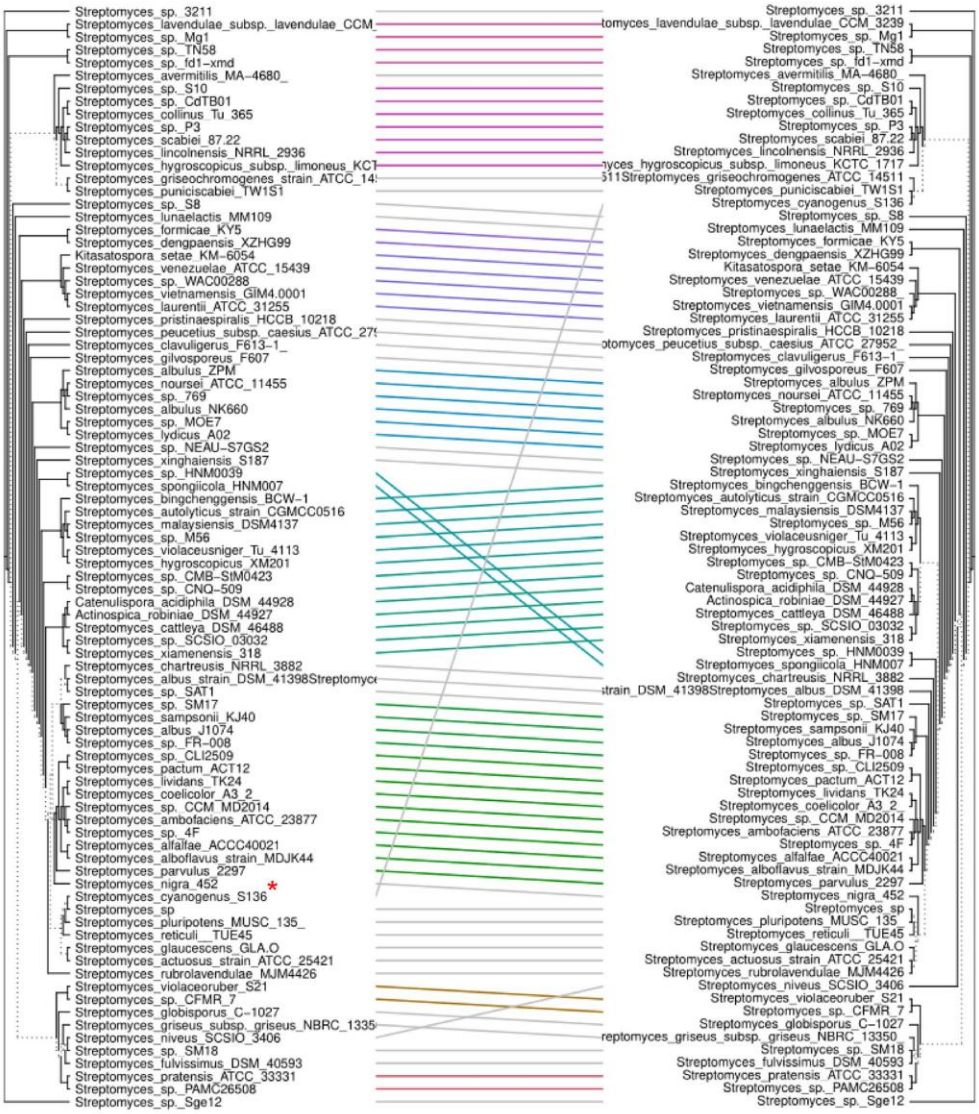


Fig. 2 Comparison of topologies of the tree (tanglegram) built on the dataset containing typical (left) and atypical (right) *S. cyanogenus* 16S rRNA genes. S136 position is marked with red asterisk. Multiple sequence alignment for these trees was prepared with MAFFT (maxiter = 1000, ginsi method)

set of *S. puniciscabiei* aligned well to S1361_18950 (Fig. S5A, ESM), sharing 98.5% nucleotide sequence identity. Hence, S136 genome is a rare case of presence of disparate 16S rRNA gene whose origin and functional consequences deserve further investigations. Neither of the two reconstructions placed S136 strain in one clade with the other 4 *S. cyanogenus* strains, whose 16S rRNA genes are available in GenBank. (Fig. S5b, ESM).

Given that 16S rRNA genes might be suboptimal proxy to phylogeny of *S. cyanogenus*, we decided to compare the phylogenetic position of S136 on the trees based on protein and 16S markers. We followed the approach of Gao and Gupta who proposed a set of highly conserved proteins for phylogenetic reconstruction within class Actinobacteria (Gao and Gupta 2012). Unfortunately, genomic records for the other *S. cyanogenus* species were unavailable and thus were not included into reconstruction. There was a strong discordance in topology of 16S- and protein-based trees (Fig. S6). Given that a set of protein markers provides a more robust phylogenetic signal to infer intrageneric relationships (Kämpfer et al. 2012), we assume that the protein tree (Fig. 3) represents the most likely phylogenetic position of *S. cyanogenus* within *Streptomyces* genus. In this case *S. reticuli* TUE45 came out as the closest relative of *S. cyanogenus*. An average nucleotide identity between S136 and TUE45 genomes (using two-way ANI approach; Goris et al. 2007) was 91.44% which is below 95% threshold level for the members of one species. Therefore, *S. cyanogenus* S136 likely represents new genomic species of *Streptomyces*.

Secondary metabolome of S136

Within S136 genome antiSMASH annotated 33 secondary metabolite BGCs with default settings. Highest score was observed for several BGCs directing the production of ectoine, spore pigment, desferrioxamine B, albaflavenone, lagmysin, landomycin, geosmin, citrulassin E, hopene (more than 80% confidence). Only landomycin BGC is known to be actively expressed in S136 (Yushchuk et al. 2019). The number of BGCs predicted by antiSMASH should be treated with caution due to known problems with correct identification of BGC borders and inefficient recognition of carbohydrate-based BGCs (Olano et al. 2014). An example of artificial fusion of disparate BGCs into one region was clearly present in S136 genome, where well-studied LaA BGC was lumped together with the other two putative BGCs (Fig. S7, ESM) Furthermore, the alternative tools for BGC annotation, such as PRISM and DeepBGC, predict somewhat different sets of BGCs in S136 genome. We therefore manually compared the outputs of different software tools in an attempt to assess, as precisely as possible, the true number of different secondary metabolite BGCs in S136 genome (see Methods section and Fig. S8, ESM). As a result, a total of 41 BGCs were revealed (Table 2). Almost half of these BGCs (19 out of 41) significantly resemble known BGCs. 30% of all S136 BGCs are distinct from the BGCs families present in BiG-FAM database, a vast (over 1.2 million entries) collection of BGCs from different taxa (Kautsar et al. 2021). Particularly, their distance from BiG-FAM families was more than 900, a proposed threshold above which the BGC of interest is considered novel. The novelty of natural products encoded within S136 genome could also be viewed from a perspective of BGC percentage that belong to small BiG-FAM families (e.g. those that consist of less than 100 BGCs). In S136 genome 12.5% of the BGCs fell into this category. This could be interpreted as an additional layer of confidence in BGC prioritization, especially if those families have no experimentally characterized BGCs. A more detailed description of how different software tools portray genetic blueprint of the secondary metabolic landscape of S136 strain can be found in fig. S8–17, ESM, and the text that accompanies the figures.

Given the large complement of BGCs revealed in *S. cyanogenus* genome, we wondered as to whether all of them are unconditionally silent. Variation of cultivation conditions is one of the simplest ways to induce the expression of specialized pathways in *Streptomyces*; indeed, even landomycin A production by S136 is medium-dependent (Yushchuk et al. 2018, 2019). We therefore explored how several common agar media influence the production of antibacterial compounds by two *S. cyanogenus* strains, wild type (S136) and its landomycin-nonproducing mutant ΔlanI7. While the wild type showed no activity under our test conditions, the ΔlanI7 accumulated some antibacterial compounds when grown on SMMS, ISP3 and especially R5 agars (Fig. 4). The level of antibacterial activity observed under these conditions appeared to be higher than the one observed by us previously under different cultivation conditions (Yushchuk et al. 2021).

Discussion

In this work we report complete sequence of *Streptomyces cyanogenus* S136 genome, the only known producer of anticancer polyketide landomycin A. Two features stand out in this genome as compared to the genomes of the other *Streptomyces* species: increased number of tRNA genes and presence of highly divergent 16S rRNA gene. While the fact and relevance of inflated number of tRNA genes for S136 physiology await experimental scrutiny, the existence of one dissimilar 16S rRNA gene in S136 genome readily raises the issues of taxonomical kind. Indeed, we showed in this work that the use of different S136 rRNA genes in phylogenetic reconstruction led to incongruent species trees. One therefore has to be careful

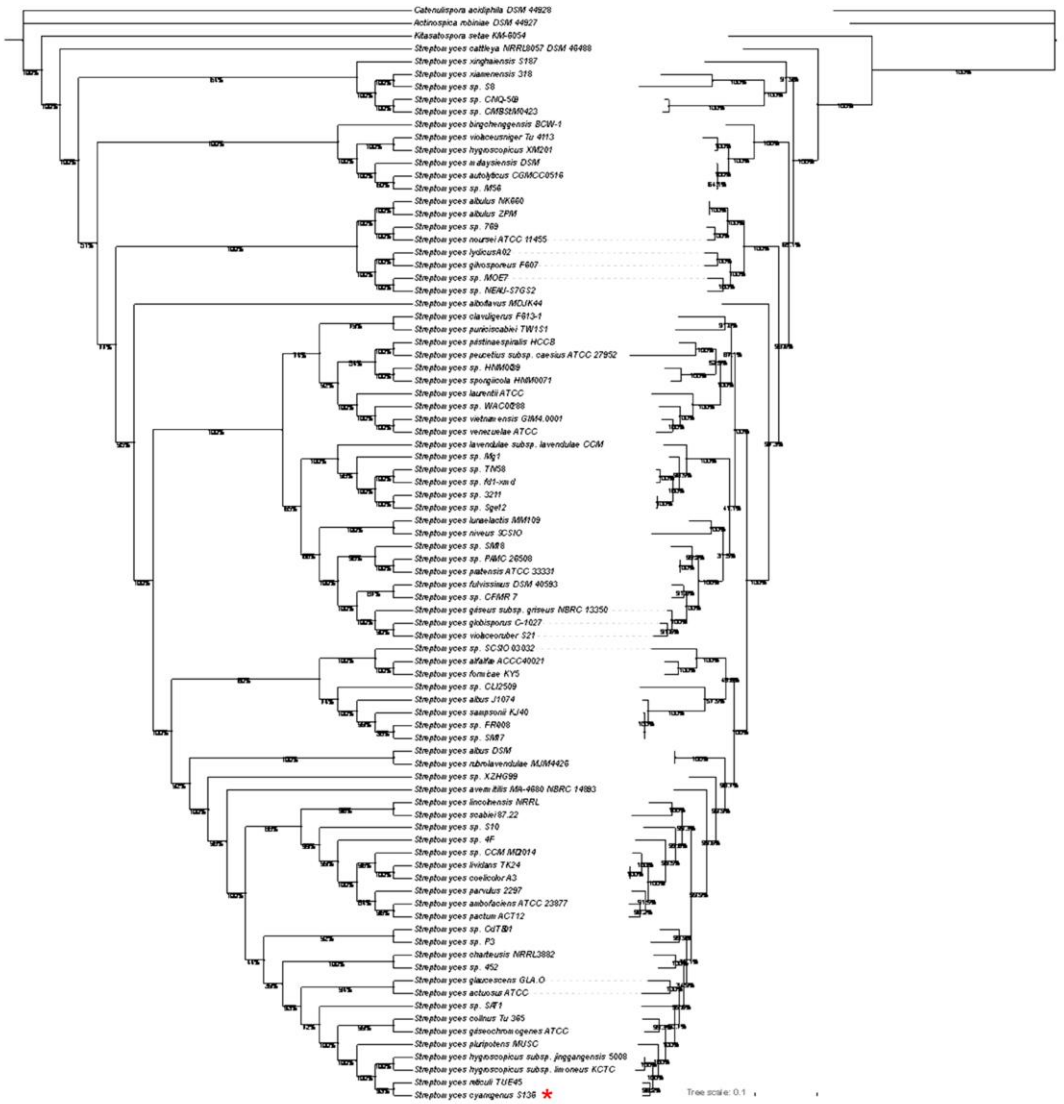


Fig. 3 *Streptomyces* species tree built on the basis of conservative proteins. *S. cyanogenus* is at the bottom of the tree (left, red asterisk). ML algorithm was used for the tree reconstruction. On the left is the cladogram, on the right-phylogram of conserved proteins tree. Branch support values (%) are computed with the help of SH-alRT test

in the choice of the molecular markers to infer phylogeny for such strains. We recommend to use a set of protein-encoding genes in this case, as they would be less susceptible to reconstruction artifact. In all, our phylogenetic efforts led us to believe that *S. cyanogenus* S136 is a new genomic species, and its closest known relative is *S. reticuli* TUE45. 16S rRNA phylogeny reconstructions also showed that *S. cyanogenus* S136 is distantly related to four other *S. cyanogenus* strains.

The genome sequence was subjected to a number of bioinformatics tools that identify natural product BGCs, and a round of manual curation of the output data. This approach yielded a list of 41 BGCs, a significant portion of which overlap with the previously experimentally characterized

Table 2 Biosynthetic Gene Clusters in *Streptomyces cyanogenus* S136 genome

BGCs #	Coordinates (left–right edges)	Size, bp	Chemical core	Most similar (%, if available) BGC directs the production of	Notes
1	246,653–289,035	42,382	Polyketide, type I PKS	Macbecin (39)	Some genes for transport, regulation, precursor supply and polyketide modification are missing
2	335,166–341,095	5929	Saccharide	–	–
3	738,785–742,279	3494	Polyphenols, melanin	Melanin (57)	No transport gene in the BGC as compared to the reference
4	1,016,423–1,017,610	1187	Pyranone, type 3 PKS	Germicidin (100)	–
5	1,027,563–1,051,082	23,519	Peptide + nucleoside	–	–
6	1,121,707–1,212,402	90,695	Polyketide, type I PKS	Chlorothricin (55)	Similar to chlorothricin BGC; several structural genes absent
7	1,237,690–1,273,267	35,577	Peptide, NRPS	Acyldepsipeptide 1 (15)	Only core is similar to acyldepsipeptide
8	1,473,681–1,476,488	2807	Naphthquinone, type III PKS	Flaviolin (75)	Absent regulatory gene
9	1,620,953–1,667,528	46,575	Polyketide, type I PKS	–	–
10	2,301,274–2,304,405	3131	Amino acid (ectoine)	Ectoine (100)	Known ectoine BGC
11	2,700,072–2,709,366	9294	Polyketide, type II PKS	spore pigment (83)	100% similarity to curamycin BGC (BGC0000215)
12	3,187,463–3,191,286	3823	Terpene	–	No known BGC; clusterblast reveals BGCs with similar terpene subBGCs
13	3,323,050–3,324,394	1344	Polyphenols, melanin	Melanin (100)	Known melanin BGC
14	3,411,660–3,417,647	5987	Hydroxamate, NRPS	Desferrioxamin B (83)	Known siderophore BGC with missing transport gene
15	4,115,479–4,164,034	48,555	RiPP	Linaridin	Identified only by RiPPMiner
16	4,509,957–4,535,739	25,782	Peptide, NRPS	–	No similar BGCs in MiBIG
17	4,554,282–4,563,340	9058	Other, lincomycin	–	Core gene similar to LmbU from lincomycin BGC. No other hits
18	4,671,291–4,678,868	8480	Peptide, bacteriocin		Novel bacteriocin BGC
19	4,997,513–5,003,746	6233	RiPP	–	Core similar to BGCs for thiopeptide biosynthesis
20	5,735,976–5,738,290	2314	Terpene	Albaflavenone (100)	–
21	6,374,219–6,384,127	9908	Other, siderophore	–	–
22	6,453,759–6,520,752	66,993	Peptide, NRPS	–	Likely two BGCs; first one carries core genes and second one accessory (+ MeAsp subBGC). Considered one BGC until further verification
23	6,522,403–6,527,344	4941	RiPP	Lagmysin (80)	Highly similar to lagmysin BGC
24	6,552,992–6,560,907	7915	Terpene	Isorenieratene (75)	No regulatory gene and geranylgeranyl diphosphate synthesis genes
25	6,588,268–6,597,540	9272	Peptide, bacteriocin	–	–
26	6,606,883–6,632,271	25,388	Peptide, NRPS	–	Often located near bacteriocin BGC
27	6,643,964–6,676,435	32,471	Polyketide, type II PKS	landomycin A (100)	Active BGC
28	6,731,997–6,741,890	9893	Peptide, bacteriocin	–	–
29	6,770,569–6,772,731	2162	Terpene	Geosmin (100)	
30	6,952,316–6,959,533	7217	Other, siderophore	–	–

Table 2 (continued)

BGCs #	Coordinates (left–right edges)	Size, bp	Chemical core	Most similar (%, if available) BGC directs the production of	Notes
31	6,985,105–6,991,275	6170	Butyrolactone	–	Butyrolactone BGC present in different species; no known hits
32	7,019,743–7,027,247	7504	Betalactone	–	–
33	7,293,687–7,298,086	4399	lanthipeptide, RiPP	citrulassin E (100)	Similar to citrulassin BGCs with different degree of gene losses
34	7,298,435–7,307,214	8779	lassopeptide, RiPP	–	No known BGC hit, but shared synteny across several species
35	7,427,159–7,440,932	13,773	Terpene	Hopene (92)	
36	7,480,261–7,485,226	4965	Lassopeptide, RiPP	–	–
37	7,571,920–7,665,592	93,672	Polyketide, type I PKS	Naphthomycin A (68)	Naphtomycin BGC
38	7,705,775–7,792,970	87,195	Polyketide, type I PKS	Lucensomycin	(Yushchuk et al. 2021)
39	8,000,207–8,077,410	77,203	bacteriocin + NRP	–	–
40	8,414,168–8,439,406	25,238	Polyketide, type I PKS	–	–
41	8,461,058–8,490,615	29,557	Saccharide	–	Carbohydrate synthesis subBGC (42% similarity to D-olivose pathway from clorothricin BGC -)

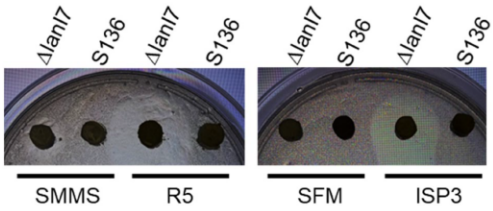


Fig. 4 *S. cyanogenus* produces antibacterially active compounds under certain conditions. Agar plugs of S136 and ΔlanI7 were grown and assayed against *B. cereus* as described in Methods. Agar media used to grow the strains are denoted on the bottom of the figure. Halos of *B. cereus* growth inhibition are visible around agar plugs cut off the lawns of ΔlanI7 grown on SMMS, R5 and ISP3 media. Photos represent typical results of three independent expreriments

BGCs present in MiBIG database. Some S136 BGCs lack MiBIG hit; in this case their identity was inferred from combined data annotation approach, which is graphically summarized in ESM with the help BGCViz software (Fig. S9, ESM). Therefore, even in absence of comparable MiBIG data, novel BGCs can be pinpointed by taking into account the combined annotation results from a set of search tools other than antiSMASH.

According to our findings (see ESM, Fig. S9–17), BGCs #9, 16, 22, 26, 39, 40 and 41 deserve special attention due to large distance between them and BGCs deposited in BiG-FAM database, implying their novelty. Nevertheless, we cannot rule out at the moment that our assumption of novelty can be undermined by the incompleteness of the BGCs within S136 genome. Attempts to induce the expression of these BGCs in S136 genome are underway in our laboratories. Genome of S136 also carries two BGCs similar, yet not identical, to well-studied gene clusters whose activation could be achieved in a relatively straightforward way given the presence of all necessary cluster-situated regulatory genes. These are BGCs #6 (for chlorothricin-like compound) and #37 (naphthomycin A). Indeed, our recent efforts in this regard allowed us to activate the BGC #38 leading to production of expected polyene compound lucensomycin (Yushchuk et al. 2021). Several BGCs seem to be silent because of the absence of regulatory or transport genes. These are BGCs #1 (macbecin), #3 (melanin), #8 (flaviolin), #14 (siderophore) and #24 (isorenieratene). Respective or similar BGCs have been studied in other species, and they carry regulatory or transport genes that can be expressed heterologously in S136 to induce the aforementioned BGCs.

Specialized RiPP-finding tools reveal a swath of RiPP "dark matter" potentially encoded within S136 genome. RiPPMiner and BAGEL4 were the most promising tools displaying good balance between ability to identify putative BGCs within novel genomic regions and identification reliability. Full list of potential RiPP BGCs, as predicted by the aforementioned tools, are given in ESM, fig. S15–17; below we discuss a few of them, which could be confidently described. BGCs #19, 34 and 36 encode novel RiPPs and

are identified by multiple bioinformatic tools (see Table 2). The cluster #15, initially detected by RippMiner, wasn't pinpointed by any other search tools, except NeuRiPP, and so it can be considered a novel RiPP. BGC #1 from DeepRiPP (Fig. S16) intersects with NRPS clusters of PRISM and DeepBGC. Given the high reliability of PRISM identification of NRPS clusters, this RiPP hit should be treated with caution. The antiSMASH-identified BGC (#7) does not overlap with this gene cluster.

To conclude, *S. cyanogenus* S136 likely represents novel species that displays a large potential to produce various natural products besides landomycins and lucensomycin. Its genome harbors numerous BGCs directing the biosynthesis of reduced polyketides, peptides of ribosomal and non-ribosomal origin, terpenes and saccharides. Some of them are worth further experimental efforts due to high likelihood of chemical novelty of the corresponding natural products. Our initial bioassays readily reveal that some of these BGCs are expressed and lead to copious production of some as-yet-unknown antibacterially active compounds. This emphasizes biotechnological potential of *S. cyanogenus* S136.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13205-021-02834-4.

Funding The work was supported by grants BG-80F and BG-09F from the Ministry of Education and Science of Ukraine (to B.O. and V.F.).

Declarations

Conflict of interest The authors declare that they have no conflict of interests.

References

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T (2019) antiSMASH 50: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 47(W1):W81–W87. https://doi.org/10.1093/nar/gkz310

de Los SE (2019) NeuRiPP: neural network identification of RiPP precursor peptides. Sci Rep 9:13406. https://doi.org/10.1038/s41598-019-49764-z

Gao B, Gupta RS (2012) Phylogenetic framework and molecular signatures for the main clades of the phylum actinobacteria. Microbiol Mol Biol Rev 76:66–112. https://doi.org/10.1128/MMBR.05011-11

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. https://doi.org/10.1099/ijs.0.64483-0

Henkel T, Rohr J, Beale JM, Schwenen L (1990) Landomycins, new angucycline antibiotics from Streptomyces sp. I. Structural studies on landomycins A-D. J Antibiot 43:492–503

Kämpfer P, Glaeser SP (2012) Prokaryotic taxonomy in the sequencing era–the polyphasic approach revisited. Environ Microbiol 14:291–317. https://doi.org/10.1111/j.1462-2920.2011.02615.x

Kanehisa M, Sato Y, Morishima K (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006

Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC et al (2020) MIBiG 2.0: a repository for biosynthetic gene clusters of known function. Nucleic Acids Res 48:D454–D458. https://doi.org/10.1093/nar/gkz882

Kautsar SA, Blin K, Shaw S, Weber T, Medema MH (2021) BiG-FAM: the biosynthetic gene cluster families database. Nucleic Acids Res 49:D490–D497. https://doi.org/10.1093/nar/gkaa812

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical *Streptomyces* genetics. John innes foundation, Norwich, United Kingdom

Luo H, Quan CL, Peng C, Gao F (2019) Recent development of Ori-Finder system and DoriC database for microbial replication origins. Brief Bioinf 20:1114–1124. https://doi.org/10.1093/bib/bbx174

Nett M, Ikeda H, Moore BS (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. Nat Prod Rep 26:1362–1384. https://doi.org/10.1039/b817069j

Olano C, García I, González A, Rodriguez M, Rozas D, Rubio J, Sánchez-Hidalgo M, Braña AF, Méndez C, Salas JA (2014) Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. Microb Biotechnol 7:242–256. https://doi.org/10.1111/1751-7915.12116

Ostash B, Korynevska A, Stoika R, Fedorenko V (2009) Chemistry and biology of landomycins, an expanding family of polyketide natural products. Mini Rev Med Chem 9:1040–1051. https://doi.org/10.2174/138955709788922593

Rebets Y, Ostash B, Luzhetskyy A, Hoffmeister D, Brana A, Mendez C, Salas JA, Bechthold A, Fedorenko V (2003) Production of landomycins in Streptomyces globisporus 1912 and S. cyanogenus S136 is regulated by genes encoding putative transcriptional activators. FEMS Microbiol Lett 222:149–153. https://doi.org/10.1016/S0378-1097(03)00258-1

Rebets Y, Dutko L, Ostash B, Luzhetskyy A, Kulachkovskyy O, Yamaguchi T, Nakamura T, Bechthold A, Fedorenko V (2008) Function of *lanI* in regulation of landomycin a biosynthesis in *Streptomyces cyanogenus* S136 and cross-complementation studies with *Streptomyces* antibiotic regulatory proteins encoding genes. Arch Microbiol 189:111–120. https://doi.org/10.1007/s00203-007-0299-5

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Seppey M, Manni M, Zdobnov EM (2019) BUSCO assessing genome assembly and annotation completeness. In: Kollmar M (ed) Gene Prediction. Methods in Molecular Biology, vol 1962. Springer, New York

Syberg-Olsen M, Garber A, Keeling P, McCutcheon J, Husnik F (2020) Pseudofinder, GitHub repository: https://github.com/filip-husnik/pseudofinder/. Accessed 20 Dec 2020

Tippelt A, Busche T, Rückert C, Nett M (2020) Complete genome sequence of the cryptophycin-producing cyanobacterium *Nostoc* sp. strain ATCC 53789. Microbiol Resour Announ 9:e00040-e120. https://doi.org/10.1128/MRA.00040-20

Westrich L, Domann S, Faust B, Bedford D, Hopwood DA, Bechthold A (1999) Cloning and characterization of a gene cluster from *Streptomyces cyanogenus* S136 probably involved in landomycin biosynthesis. FEMS Microbiol Lett 170:381–387. https://doi.org/10.1111/j.1574-6968.1999.tb13398.x

Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R (2014) Uniting the classification of cultured and uncultured bacteria

and archaea using 16S rRNA gene sequences. Nat Rev Microbiol 12(9):635–645. https://doi.org/10.1038/nrmicro3330

Yushchuk O, Ostash I, Vlasiuk I, Gren T, Luzhetskyy A, Kalinowski J, Fedorenko V, Ostash B (2018) Heterologous AdpA transcription factors enhance landomycin production in Streptomyces cyanogenus S136 under a broad range of growth conditions. Appl Microbiol Biotechnol 102(19):8419–8428. https://doi.org/10.1007/s00253-018-9249-1

Yushchuk O, Kharel M, Ostash I, Ostash B (2019) Landomycin biosynthesis and its regulation in *Streptomyces*. Appl Microbiol Biotechnol 103:1659–1665. https://doi.org/10.1007/s00253-018-09601-1

Yushchuk O, Ostash I, Mösker E, Vlasiuk I, Deneka M, Rückert C, Busche T, Fedorenko V, Kalinowski J, Süssmuth RD, Ostash B (2021) Eliciting the silent lucensomycin biosynthetic pathway in *Streptomyces cyanogenus* S136 via manipulation of the global regulatory gene *adpA*. Sci Rep 202111:3507. https://doi.org/10.1038/s41598-021-82934-6

مدينة الملك عبدالعزيز للعلوم والتقنية KACST Springer

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s13205-021-02834-4.
Електронні додатки до статті доступні за посиланням:
https://static-content.springer.com/esm/art%3A10.1007%2Fs13205-021-02834-4/MediaObjects/13205_2021_2834_MOESM1_ESM.pptx.

Анотація: *Streptomyces cyanogenus S136 є єдиним відомим продуцентом ландоміцину А (LaA), одного із засновників родини ароматичних ангуциклінових полікетидів. LaA демонструє потужні протиракові властивості, що робить цей природний продукт об'єктом численних хімічних і клітинно-біологічних досліджень. Мало відомо про потенціал штаму S136 продукувати інші вторинні метаболіти. Тут ми повідомляємо про повну послідовність геному продуцента LaA та про те, як ми використали цю послідовність для оцінки цього виду, його філогенетичного положення та різноманітності кластерів генів для біосинтезу антибіотиків.*

7.2 Вплив гетерологічних алелів гена *adpA* на продукцію ландоміцину А в *S. cyanogenus* S136

Володіючи послідовністю генома *S. cyanogenus* S136, ми ідентифікували в ній гени, який кодує ортолог AdpA (названий $AdpA_{scy}$). Нуклеотидна послідовність КБГ *lan* містила потенційні сайти зв'язування для AdpA на відстані 9 п.н. від старт-кодону гена *lanX* (кодує ацил-КоА мутазу), а також на відстані 1000 і 800 п.н. від 5'-кінця гена *lanI*, який кодує ключовий шлях-специфічний активатор транскрипції всіх структурних генів *lan*. Це дає змогу припустити, що експресія *lan*-генів може контролюватися за допомогою AdpA.

В подальшому, ми дослідили властивості $adpA_{scy}$ експресувавши його в *S. coelicolor* M851 (системі, яка добре зарекомендувала себе в наших дослідженнях, див. **Розділ 3**), а також в нативному штамі. Отже, повнорозмірний ген $adpA_{scy}$, а також його фрагмент $adpA_{scy}_dbd$, який кодує лише ДНК-зв'язувальний домен, були надекспресовані в *S. cyanogenus* S136. У цьому експерименті ми хотіли з'ясувати, чи справляє ДНК-зв'язувальний домен $AdpA_{scy}$ якийсь вплив на морфологію та продукцію ландоміцину за відсутності димеризаційного домена. Створені рекомбінантні похідні *S. cyanogenus* S136 не відрізнялися від дикого типу за продукцією ландоміцину А [143].

Відсутність впливу додаткової копії $adpA_{scy}$ на морфологію і продукцію ландоміцину А в *S. cyanogenus* можна пояснити двома способами. По-перше, AdpA дійсно може не бути активатором продукції ландоміцину, а знайдені нам *in silico* в КБГ *lan* операторні сайти можуть бути нерелевантними. Друге пояснення припускає, що AdpA може бути нефункціональним в *S. cyanogenus* S136 (що може мати сенс в контексті Bld фенотипу цього штаму) [143].

Щоби підтвердити чи спростувати ці припущення, ми спробували комплементувати Bld фенотип *S. coelicolor* M851 за допомогою генів $adpA_{scy}$ і $adpA_{scy}_dbd$. Виявилося, що Bld фенотип не комплементується у отриманих рекомбінантних штамів, підтверджуючи припущення, згідно з яким ген $adpA_{scy}$ кодує нефункціональний білок [143].

Наявність кількох операторних ділянок для AdpA в КБГ *lan* (див. вище) дала змогу нам припустити, що цей транскрипційний фактор був частиною

регуляторної мережі біосинтезу ландоміцину А до того, як втратив свою функціональність. Якщо так, то гетерологічні гени *adpA* можуть впливати на морфогенез *S. cyanogenus* і продукцію антибіотиків. Ми вибрали набір *adpA*-генів з кількох добре вивчених видів *Streptomyces*, зокрема *S. coelicolor* M145, *S. ghanaensis* (=*viridosporus*) ATCC 14672, *S. clavuligerus* ATCC 27064 і *S. albus* (=*albidoflavus*) J1074. Ортологи AdpA цих видів представляють різні клади філогенії AdpA *Streptomyces*. Відповідні гени було надекспресовано в *S. cyanogenus* S136 [143].

Виявилося, що гетерологічні *adpA*-гени справляли позитивний вплив на продукцію ландоміцину А в *S. cyanogenus*. А саме, штами накопичували більше ландоміцину А на різних агаризованих середовищах. У рідкому середовищі SG рекомбінантні штами продукували до 250 мг/л ландоміцину А, тоді як рівень продукції антибіотика в штамі дикого типу був прибл. у 6 разів нижчий. У мінімальному середовищі ISP5l продукція ландоміцину А рекомбінантними штамами становила в середньому 275 мг/л, що перевищувало продукцію дикого типу більше ніж у 15 разів. Рекомбінантні штами не відрізнялися рівнем накопичення біомаси у зазначених середовищах порівняно з диким типом [143].

Ми також виявили, що продукція ландоміцину А рекомбінантними штамами індукується за умов, коли дикий тип не здатний продукувати цей антибіотик, зокрема за умов культивування в середовищі TSB [143].

Беручи до уваги те, що наш аналіз *in silico* передбачив декілька операторних послідовностей для AdpA в КБГ *lan*, цілком ймовірно, що позитивний ефект надекспресії *adpA* опосередковується геном *lanI*, який кодує ключовий шлях-специфічний регулятор біосинтезу ландоміцину А [24]. Для того, щоб перевірити це припущення, ми надекспресували згадані вище алелі *adpA* в штамі *S. cyanogenus*, мутантному за геном *lanI* – *S. cyanogenus* ΔlanI7. Отримані рекомбінантні штами виявилися нездатними до продукції ландоміцину А за умов культивування в середовищах SG, ISP5l і TSB. Отже, AdpA-опосередкована індукція продукції ландоміцину А в *S. cyanogenus* S136 залежить від присутності гена *lanI* [143].

Щоби чітко прослідкувати зв'язок між фенотипами та експресією генів, ми порівняли експресію гена *lanI*, а також деяких інших генів з КБГ *lan* (зокрема, *lanA*, *lanJ* і *lanG*), у *S. cyanogenus* S136 і *S. cyanogenus* pGM4181^{+} (рекомбінантний штам, що надекспресує ген *adpA* з *S. albus* J1074) за допомогою ЗТ-ПЛР. Ми проаналізували експресію генів у культурах, вирощених у середовищах TSB та SG протягом 45 год. У випадку SG, *lanI* експресувався сильніше в *S. cyanogenus* pGM4181^{+}; відповідно, експресія структурних генів *lan* була помітно вищою. За умов культивування в середовищі TSB, ми не виявили експресії генів *lanI*, *lanA*, *lanJ*, *lanG* у *S. cyanogenus* S136, тоді як ці гени активно експресувалися в *S. cyanogenus* pGM4181^{+} за тих же умов культивування [143].

Отже, отримані нами результати демонструють, що функціональні алелі *adpA*, які походять з гетерологічних господарів, є цінними знаряддями для підвищення рівнів продукції ландоміцину А в *S. cyanogenus*. Що більш важливо, відбувається AdpA-опосередкована активація продукції цього антибіотика за умов, в яких дикий типу до продукції не здатний. Це говорить на користь використання поточного підходу для спроб активації мовчазних КБГ.

Основні результати цього підрозділу опубліковані в [143].

Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-018-9249-1

APPLIED GENETICS AND MOLECULAR BIOTECHNOLOGY

Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions

Oleksandr Yushchuk[1] · Iryna Ostash[1] · Iryna Vlasiuk[1] · Tetiana Gren[2] · Andriy Luzhetskyy[3] · Joern Kalinowski[2] · Victor Fedorenko[1] · Bohdan Ostash[1]

Received: 16 May 2018 / Revised: 9 July 2018 / Accepted: 14 July 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Streptomyces cyanogenus S136 is the only known producer of landomycin A (LaA), one of the largest glycosylated angucycline antibiotics possessing strong antiproliferative properties. There is rising interest in elucidation of mechanisms of action of landomycins, which, in turn, requires access to large quantities of the pure compounds. Overproduction of LaA has been achieved in the past through manipulation of cluster-situated regulatory genes. However, other components of the LaA biosynthetic regulatory network remain unknown. To fill this gap, we elucidated the contribution of AdpA family pleiotropic regulators in landomycin production via expression of *adpA* genes of different origins in *S. cyanogenus* S136. Overexpression of the native *S. cyanogenus* S136 *adpA* ortholog had no effect on landomycin titers. In the same time, expression of several heterologous *adpA* genes led to significantly increased landomycin production under different cultivation conditions. Hence, heterologous *adpA* genes are a useful tool to enhance or activate landomycin production by *S. cyanogenus*. Our ongoing research effort is focused on identification of mutations that render *S. cyanogenus* AdpA nonfunctional.

Keywords *Streptomyces* · Landomycin A · Regulation · AdpA family regulators

Introduction

Streptomyces cyanogenus S136 produces a family of glycosylated angucycline secondary metabolites known as landomycins (Henkel et al. 1990; Ostash et al. 2009). Landomycin A (LaA; Fig. 1) is the major and largest (by molecular mass) member within the family; it is also the compound most active against various tumor cell lines, as compared to the other landomycins (Ostash et al. 2004; Kharel et al. 2012). A number of studies suggest that anticancer activities of landomycins, and LaA in particular, are mediated by novel mechanisms of action (Korynevska et al. 2007; Lehka et al. 2015; Panchuk et al. 2017). This fuels interest in identification of molecular targets of LaA action and further development of novel LaA-based anticancer drugs.

Access to large quantities of LaA is key to its transformation into a drug. As total synthesis of LaA remains economically unfeasible, a biological (fermentation-based) route is currently the only one to generate landomycins. Besides *S. cyanogenus*, no other LaA producer has been described to date. Over the last 15 years, we studied a number of regulatory genes involved in LaA biosynthesis and searched for heterologous production hosts (Ostash et al. 2009; Myronovskyi et al. 2016). All our efforts so far have been focused on pathway-specific (cluster-situated) regulatory genes; no information is available about a global layer of landomycin production control. Recent advances in *Streptomyces* genomics unveiled a number of pleiotropic regulators that coordinate morpholo

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00253-018-9249-1) contains supplementary material, which is available to authorized users.

✉ Bohdan Ostash
bohdan.ostash@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, Hrushevskoho St. 4, Rm. 102, Lviv 79005, Ukraine

[2] Technology Platform Genomics, CeBiTec, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany

[3] Department of Pharmaceutical Biotechnology, Helmholtz Institute for Pharmaceutical Research Saarland, Actinobacteria Metabolic Engineering Group, Saarland University, UdS Campus C2 3, 66123 Saarbrucken, Germany

Published online: 28 July 2018

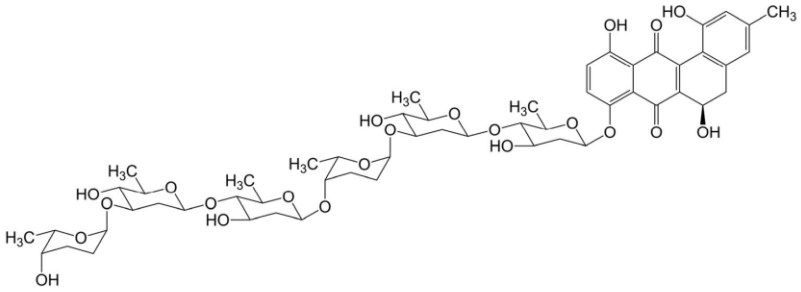


Fig. 1 Structure of landomycin A (LaA), final product of *lan* cluster of *S. cyanogenus* S136

gical transitions and antibiotic production (Martín and Liras 2010; Liu et al. 2013). Out of a dozen global regulators studied to date, the transcriptional factor of the AraC/XylS family, known as AdpA (Higo et al. 2012), attracted our attention because it is universally present in *Streptomyces* genomes (Rabyk et al. 2018) and frequently influences the transcription of antibiotic biosynthesis genes, both directly (Wang et al. 2009; Claesen and Bibb 2011; Makitrynskyy et al. 2013) and indirectly (Takano et al. 2003; Ohnishi et al. 2005; Higashi et al. 2007; Pan et al. 2009; López-García et al. 2010; Du et al. 2011; Xu et al. 2017; Yu et al. 2018). We therefore decided to study whether the *adpA* gene is involved in LaA biosynthesis and can be utilized to improve the LaA titers. Here, we report identification of *adpA* within the *S. cyanogenus* chromosome and its expression, along with a set of heterologous *adpA* genes, in the LaA producer. Our results suggest that *adpA* is not an essential component of the LaA biosynthesis regulatory network. Expression of several heterologous *adpA* genes leads to significant increases in LaA titers under different fermentation conditions, including those where this antibiotic is not typically produced. We discuss our findings in a context of LaA biosynthesis regulation, as well as prospects of *adpA* genes for activation of cryptic secondary metabolic pathways.

Materials and methods

Microorganisms, vectors, and culture conditions

Streptomyces cyanogenus wild type (S136; DSM5087) and its *lanI*-minus mutant ΔlanI7 with completely blocked LaA production (Rebets et al. 2008) were used throughout the work. The *adpA*-minus mutant *S. coelicolor* M851 (Takano et al. 2003) was used to express an *adpA* ortholog from *S. cyanogenus*. *Escherichia coli* DH5α was used for routine cloning procedures. *E. coli* ET12567 (pUZ8002) was used to perform intergeneric conjugation from *E. coli* to *Streptomyces*. *E. coli* strains were grown under standard conditions (Sambrook and Russell 2001). All vectors and plasmids used in this work are listed in Table 1. For intergeneric matings, *S. cyanogenus* strains were grown on ISP3 or SM agar at 30 °C. To reveal landomycin production, *S. cyanogenus* strains were grown on a number of different solid media (please see Electronic Supplementary Materials (ESM) for media composition). Equivalent amounts of *S. cyanogenus* mycelia ($(0.8–1) \times 10^7$ cfu) were plated on agar media to produce confluent lawns for analysis of morphology and antibiotic production. For submerged production of LaA, *S. cyanogenus* strains were grown in SG, ISP5l, TSB, SWS, and other media listed in ESM. Where needed, the strains were maintained in the presence of apramycin (50 μg/ml).

Plasmids construction

The coding sequence of $adpA_{scy}$ gene (*SCY4743*), together with its putative RBS, was amplified using a SCY_F/SCY_R primer pair; a part of $adpA_{scy}$ coding for DNA-binding domain ($adpA_{scy_dbd}$) was amplified using a xnr4181araC_up/SCY_R primer pair. These amplicons (1260 bp for $adpA_{scy}$ and 593 bp for $adpA_{scy_dbd}$) were subcloned into *Spe*I/*Eco*RI sites of pmoeE5script and pGCymRP21 generating pGMSCY, pGCSCY, pGMSCYd, and pGCSCYd, respectively. Similarly, *adpA* genes from *S. coelicolor* M145 ($adpA_{sco}$), *S. clavuligerus* ATCC27064 ($adpA_{scl}$), *S. ghanaensis* ATCC14672 ($adpA_{gh}$), and *S. albus* J1074 (*XNR4181*) were amplified using primer pairs listed in Table 2 and subcloned into *Spe*I/*Eco*RI sites of pmoeE5script, generating pGMSCO, pGMSCL, pOOB95d, and pGM4181, respectively. All constructed vectors were verified by restriction mapping and sequencing.

RNA isolation for semi-quantitative RT-PCR analysis

At the chosen time points, *S. cyanogenus* mycelia were collected from 4 ml of culture broth by centrifugation (14,000, 5 min, 4 °C). Biomass was washed with 1 ml of TE buffer (pH = 7.5), then resuspended in 2 ml of TE buffer supplemented with 5 mg/ml lysozyme. After 15 min of incubation of the samples in lysozyme solution at 37 °C, RNA was isolated according to modified Kirby Mix protocol (Kieser et al. 2000). The samples of total nucleic acids underwent two

Table 1 Plasmids used in this work

Name	Description	Source or reference
pmoeE5script	pGUS derivative containing *moeE5p* cloned upstream of *gusA* (Am^r Sp^r)*	Makitrynskyy et al. 2013
pGCymRP21	pGUS derivative, cumate-inducible expression system (cymR-P21; Am^r Sp^r Ap^r)	Horbal et al. 2014
pGM4181	pmoeE5script with *XNR_4181* cloned instead of *gusA*	This work
pGMSCLA	pmoeE5script with $adpA_{scl}$ cloned instead of *gusA*	This work
pGMSCO	pmoeE5script with $adpA_{sco}$ cloned instead of *gusA*	This work
pOOB95d	pmoeE5script with $adpA_{gh}$ cloned instead of *gusA*	This work
pGMSCY	pmoeE5script with $adpA_{scy}$ cloned instead of *gusA*	This work
pGMSCYd	pmoeE5script with $adpA_{scy\ dbd}$ cloned instead of *gusA*	This work
pGCSCY	pGCymRP21 with $adpA_{scy}$ cloned instead of *gusA*	This work
pGCSCYd	pGCymRP21 with $adpA_{scy\ dbd}$ cloned instead of *gusA*	This work

*Am^r, Sp^r, and Ap^r are marker genes for apramycin, spectinomycin, and ampicillin resistance respectively

DNase treatments, using RNase-free DNase I (Thermo Fisher Scientific). To rule out the possibility of DNA contamination, the purity of final RNA samples was checked by means of PCR. cDNA was synthesized from equal amounts of RNA using First Strand cDNA Synthesis Kit from Thermo Fisher Scientific. PCR was performed using Phusion High-Fidelity PCR Master Mix (Thermo Fisher Scientific) and primer pairs specific to each individual gene (Table 2). As a positive control for gene expression, the *hrdB* primer pair specific to *S. cyanogenus* was used. PCR products were analyzed by electrophoresis in 2% agarose gel.

Generation and verification of the recombinant *Streptomyces* strains

All constructs were transferred into *S. cyanogenus* strains or *S. coelicolor* M851 by conjugation, as described elsewhere (Kieser et al. 2000). *S. cyanogenus* and *S. coelicolor* transconjugants carrying pmoeE5script- and pGCymRP21-derived vectors were selected on plates overlaid with 50 μg/ml apramycin. Apramycin-resistant transconjugants were verified with PCR and primers complementary to marker *aac3(IV)* gene.

Table 2 Primers used in this work

Primer	Nucleotide sequence (5′-3′)*	Purpose
xnr4181_XbaIup	AAA*TCTAGA*GGGGGGCTTAGTCACATG	*XNR_4181* cloning
xnr4181_EcoRIrp	AAA*GAATTC*GGAGCTGTCCTCTCTCAGAC	
SCOADPA_F	AAA*TCTAGA*GGGGGGCTTAGCAGTATG	$adpA_{sco}$ cloning
SCOADPA_R	AAA*GAATTC*GCCGTCTGCTCACCTCACGG	
SCLAVADPA_F	AAA*TCTAGA*GGGGGGCTTAGCGTCATG	$adpA_{scla}$ cloning
SCLAVADPA_R	AAA*GAATTC*CATGCGACTACCTTATGG	$adpA_{scla}$/$adpA_{scla\ dbd}$ cloning
SCY_F	AAA*GAATTC*TCAGCTTCCGCTGCGCGG	$adpA_{scy}$ cloning
Xnr4181araC_up	AAATCTAGACGAGCAACGGAGGTACGGACA TGCCGGAGGAGATCGGCGCCGAC	$adpA_{scy\ dbd}$ cloning
SCY_R	TTT*TCTAGA*GGGGGGCTTAGCAGTATG	
amR_F	CGGGGTCTGACGCTCAGTGGA	*aac3(IV)* amplification
amR_R	AGCGTCTGCTCCGCCATTCG	
rt_hrdB_fp	GCTGGCCAAGGAGCTCGAC	278-bp fragment of *hrdB* (RT-PCR)
rt_hrdB_rp	CGTCGAGGGTCTTCGGCTG	
rt_lanI_up	CCGTACGGCATCCGTGAGC	296-bp fragment of *lanI* (RT-PCR)
rt_lanI_rp	CGCAGGCTGCTGACATGCG	
rt_lanA_up	CAACCCCGAGGCCATCGAC	295-bp fragment of *lanA* (RT-PCR)
rt_lanA_rp	GTCCGTGGTGTGCTCACGG	
rt_lanG_up	CACGCTCATCGGCACGGTC	298 fragment of *lanG* (RT-PCR)
rt_lanG_rp	GGTGTGGTCACCGACGACG	
rt_lanJ_up	GGCGATGGAGAGCGTCGAG	272 fragment of *lanJ* (RT-PCR)
rt_lanJ_rp	CCGAGACGGTCAGGAAGGC	

*Restriction sites are italicized in primer sequence

In silico analysis and *S. cyanogenus* genome sequencing

Search for conserved motifs within the *lan* cluster was carried out with the help of the MEME suite (Bailey et al. 2009) and a set of experimentally validated operators for AdpA (Higo et al. 2012). The reciprocal best BLASTP hit strategy (Kuzniar et al. 2008) was used to identify orthologs. *S. cyanogenus* genome was sequenced using Illumina approach (Metzker 2010) at the Center of Biotechnology of Bielefeld University (CeBiTec) using the Illumina-compatible Nextera DNA Sample Prep Kit (Epicentre, WI, USA) according to the manufacturer's protocol. The resulting sequencing data provided a single contig (approx. 8.8 Mbp) with 56× coverage of the genome; it is estimated that less than 1% of the genome is located in sequence gaps. Detailed description of the *S. cyanogenus* sequencing project will be published elsewhere. The accession number of the nucleotide sequence of *SCY4743* (*adpA* of *S. cyanogenus* S136; $adpA_{scy}$) and its flanking regions is MH178374.

Landomycin A production under submerged cultivation

Precultures of *S. cyanogenus* strains were prepared by inoculation in 250-ml flasks containing 50 ml of TSB medium with agar slices cut from 7-day-old lawns grown on ISP3 medium. Precultures were grown in TSB media for 72 h on a rotary shaker at 30 °C, 200 rpm; 3% of culture broth was transferred into new 250-ml flasks containing 50 ml of certain fermentation media, as detailed in the main text. Main cultures were grown for 72 h in case of cultivation in SG medium, and for 120 h in case of ISP5l and SWS media (see ESM for the recipes), both at 30 °C, unless otherwise stated. Landomycins were immediately extracted at the end of fermentation from the whole culture broths as detailed below.

Extraction, quantification, and thin-layer chromatography of landomycins

In the case of solid cultures, landomycins were extracted from 15 ml of meshed agar lawns with 10 ml of ethylacetate. Extracts were evaporated using a vacuum concentrator; dry landomycin residue was dissolved in 1 ml of methanol. In the case of liquid cultures, 10-ml samples were extracted with 10 ml of ethylacetate and processed as described above. LaA was quantified in methanol solutions spectrophotometrically at a wavelength of 445 nm. Extinction values were converted into LaA concentrations with the help of a calibration curve prepared from pure LaA. The latter was purified from the wild-type (S136) strain according to Henkel et al. (1990); the curve exhibited linear dependency between LaA concentration and extinction values within the 0.5–5.0 μg/ml range (see Fig. S1, ESM). Amounts of LaA were referred back to equal amounts of dry *S. cyanogenus* cell weight. Data are presented as mean values ± SEM calculated from at least three independent experiments.

For thin-layer chromatography (TLC), LaA was extracted from 1.5 cm^3 of agar using 1 ml of ethylacetate. Extracts were evaporated in vacuo and dissolved in 50 μl of methanol. Three to six microliters of extracted methanol solution was spotted onto TLC plates along with pure landomycin A as a TLC standard. TLC conditions were as follows: immobile phase—silica gel plates $F60_{254}$ (Merck), mobile phase—chloroform:methanol (9:1), duration of TLC run (at 22 °C)—20 min.

Results

Defining cultivation conditions that support LaA production

Increased amounts of glucose (2%) and cobalt chloride are the peculiar features of soytone-based medium SG (Rebets et al. 2003) routinely used to produce LaA. The rationale behind the choice of this medium for *S. cyanogenus* S136 is obscure. We compared LaA production by *S. cyanogenus* S136 in SG, SG without $CoCl_2$, SG with 4% glucose, and a number of conventional liquid media used in *Streptomyces* genetics, namely TSB, YEME, R5, MM, SMM (Kieser et al. 2000), ISP2, and ISP5 (see ESM for the composition). S136 was grown for 72 h in this experiment. Rather uniform growth rate of S136 was observed in the aforementioned media, as could be judged from biomass dry weight values (within 18–22 mg range per 10 ml); the only exceptions were minimal medium MM, where biomass yield was considerably lower (11–12 mg) and SG with 4% glucose (25–28 mg/10 ml). Our results confirmed that SG was superior as a medium for LaA production—its titers in SG exceeded by at least twofold the highest LaA titers observed in the other tested media (Fig. 2a). There was no significant change in LaA production in the presence of 4% glucose. In TSB, S136 did not produce even trace amounts of landomycins; this observation, inferred from spectrophotometric data, was further confirmed via TLC of concentrated ethylacetate extracts of TSB-grown S136 (data not shown). We performed semi-quantitative RT-PCR analysis of several key *lan* genes when S136 was grown in either SG or TSB. It was reasonable to expect that in TSB, *lan* genes are not transcribed. However, transcription of the *lan* genes was not completely turned off in TSB after 72-h cultivation, but rather decreased as compared to SG (Fig. S2, ESM).

We have also studied landomycin production in different solid media. LaA was actively produced on SM, SGa, and ISP3 media, but not in ISP2, ISP5, and SMMS (Fig. 2b).

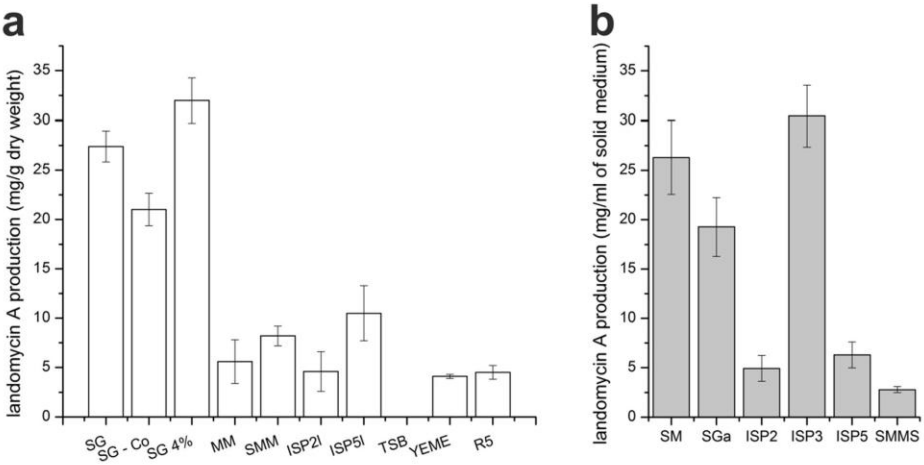


Fig. 2 Normalized LaA production levels in different liquid and solid media. In liquid culture (**a**), *S. cyanogenus* S136 was grown for 72 h, and the antibiotic amounts were referred back to equal amounts of dry cell weight. The LaA production on agar plates (**b**) was analyzed after 120 h of growth. Data shows mean values ± SEM. For media abbreviations, please refer to the main text

Elements of AdpA-based regulation in *S. cyanogenus*: operator sequences in *lan* cluster and identification of *adpA* ortholog in its genome

We screened the nucleotide sequence of the *lan* cluster for the presence of potential AdpA operator sequences. When setting the *p* value to a stringent threshold of less than 10^{-4}, the single AdpA operator could be identified 9 nt downstream of a start codon of the putative acyl-CoA mutase gene *lanX* (Fig. 3). Under relaxed search conditions ($p \leq 10^{-3}$), more operators can be found, including ones 1000 and 800 bp upstream of the gene *lanI*, that encodes the key pathway-specific transcriptional activator of all structural *lan* genes.

We generated a draft sequence of the *S. cyanogenus* S136 genome (as described in the "Materials and methods" section) and have identified an ortholog of AdpA ($AdpA_{scy}$) using the amino acid sequence of *S. coelicolor* $AdpA_{sco}$ as a query. $AdpA_{scy}$ and $AdpA_{sco}$ shared 88% of amino acid sequence identity. The DNA-binding and dimerization domains were almost identical in both proteins, although the C-terminal regions were rather variable (Fig. S3, ESM). Pfam-assisted analysis of $AdpA_{scy}$ pointed to possible impairment of the

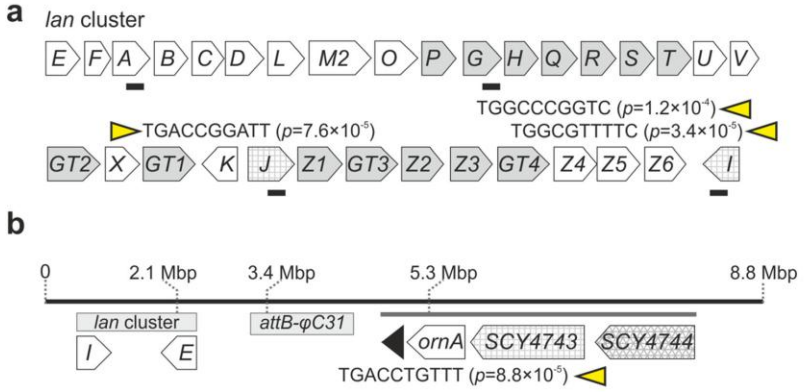


Fig. 3 Genetic organization of LaA biosynthesis gene cluster (**a**) and its relative location in *S. cyanogenus* genome (**b**). Color code for gene function: white—polyketide scaffold biosynthesis; gray—deoxysugar biosynthesis and transfer to polyketide moiety; gray squares—regulation and efflux; black—$tRNA^{His}_{GUG}$ gene; gray triangles—*SCY4744*, gene for protein of unknown function. Black lines under the *lan* gene symbols mark fragments amplified in RT-PCR experiment. Putative AdpA operator sequences are shown as triangles along with the sequence of identified operator and its *p* value. On the scheme of S136 genome relative locations of *lan* cluster, $attB^{\phi C31}$ site (where *adpA* expression constructs were integrated) and *adpA* ortholog *SCY4743* (= $adpA_{scy}$) are shown. The gene symbols and genome scheme are drawn not to scale

dimerization domain. The $AdpA_{scy}$-encoding gene *SCY4743* is located near the genes for oligoribonuclease and histidinyl tRNA (Fig. 3b). This is the usual gene neighborhood for all streptomycete *adpA* genes. A MEME search revealed single potential AdpA operator 370 bp upstream of *SCY4743* gene ($p \leq 10^{-4}$). According to the available *S. cyanogenus* contig map, the *lan* cluster is located 2.1 Mbp from the relative "left" end of the chromosome, and *SCY4743* resides approximately two 2 Mbp to the right of the *lan* cluster.

We decided to probe the function of $adpA_{scy}$ in a series of expression studies. The DNA-binding domain of AdpA from several streptomycetes was recently shown to restore sporulation to the *adpA* null mutant of *S. coelicolor* M851 (Rabyk et al. 2018). Therefore, the full-length $adpA_{scy}$ gene and $adpA_{scy_dbd}$ fragment that encodes only the DNA-binding domain were both introduced into *S. cyanogenus* S136 on the actinophage ϕC31-based integrative expression plasmid pmoeE5script (plasmids pGMSCY and pGMSCYd, respectively; see the "Materials and methods" section). In the plasmid, $adpA_{scy}$ expression would be under the control of the strong promoter *moeE5p* (Makitrynskyy et al. 2013). In this experiment, we wanted to find out whether the DNA-binding domain of $AdpA_{scy}$ exerts certain gross effects on streptomycetes in the absence of a presumably deficient dimerization domain. The generated *S. cyanogenus* S136 pGMSCY$^+$ and pGMSCYd$^+$ transconjugants did not differ in morphology or pigmentation (an indicator of altered LaA accumulation) from the control strain (carrying the empty vector) during the growth on different solid media (Fig. 4a). LaA production by the transconjugants and control strains was also essentially the same in SG and defined ISP5 media (Fig. 4b).

The absence of effects of an additional copy of $adpA_{scy}$ on *S. cyanogenus* morphology and landomycin production suggested that $AdpA_{scy}$ indeed might be nonfunctional. We endeavored to complement the Bld phenotype of *S. coelicolor* M851 using pGCSCY and pGCSCYd vectors, where expression of subcloned $adpA_{scy}$ and $adpA_{scy_dbd}$ genes, respectively, is cumate inducible. It turned out that *S. coelicolor* pGCSCY$^+$ and pGCSCYd$^+$ strains were unable to erect aerial hyphae and sporulate on R5 medium (Fig. 4c), further lending support to the assumption that $adpA_{scy}$ encodes nonfunctional protein. At present, we cannot specify what changes to primary structure of $AdpA_{scy}$ render it inactive; however, multiple sequence alignment (Fig. S3, ESM) suggests that it is C-terminal part of the protein that accumulated many changes (see also the "Discussion" section).

Expression of heterologous *adpA* genes in *S. cyanogenus* S136

The presence of several AdpA operators within the *lan* cluster (see above) allowed us to suggest that this transcriptional factor was a part of the LaA production regulatory network before $AdpA_{scy}$ lost its functionality. If so, then heterologous *adpA* genes might influence *S. cyanogenus* morphogenesis and antibiotic production. We have picked a set of *adpA* genes from several well-studied *Streptomyces* species, particularly *S. coelicolor* M145, *S. ghanaensis* ATCC14672, *S. clavuligerus* ATCC27064, and *S. albus* J1074. The AdpA orthologs from these species represent distinct clades of a *Streptomyces*-wide AdpA phylogeny (Rabyk et al. 2018). Genes from *S. coelicolor*, *S. ghanaensis*, *S. clavuligerus*, and *S. albus* were cloned into pmoeE5script vectors yielding pGMSCO, pOOB95d, pGMSCLA, and pGM4181 respectively. We analyzed landomycin production by the recombinant strains in a number of solid media, and the liquid media SG and ISP5. We also monitored the morphology of recombinant strains on different solid media.

Different heterologous *adpA* genes exerted quantitatively similar positive effect on landomycin production by *S. cyanogenus*. Namely, recombinant strains accumulated more LaA on different solid media (Fig. 5; Fig. S4, Fig. S5, ESM). We observed no effect of the introduced plasmids on *S. cyanogenus* lawn morphology. In liquid SG medium, *S. cyanogenus* strains pGMSCO$^+$, pOOB95d$^+$, pGMSCLA$^+$, and pGM4181$^+$ produced up to 250 mg/l LaA, whereas the parent strain reached approximately six-fold lower titers (40 mg/l; Fig. 5b) under the same conditions. In the minimal ISP5l medium, landomycin production by *S. cyanogenus* pGMSCO$^+$, pOOB95d$^+$, pGMSCLA$^+$, and pGM4181$^+$ strains was, on average, 275 mg/l, which exceeded wild-type production more than 15 times (Fig. 5b). The recombinant strains did not differ in biomass accumulation in the aforementioned media as compared to the wild type (in particular, *S. cyanogenus* S136 accumulated up to 3.8 ± 0.196 mg/ml of dry biomass in SG and 3.4 ± 0.236 mg/ml in ISP5l, while *S. cyanogenus* pGM4181$^+$ accumulated 3.6 ± 0.184 mg/ml and 3.18 ± 0.287 mg/ml of dry biomass in SG and ISP5l respectively). We have also found that LaA production by *S. cyanogenus* pGMSCO$^+$, pOOB95d$^+$, pGMSCLA$^+$, and pGM4181$^+$ strains was induced under conditions where *S. cyanogenus* S136 is typically incapable of producing this antibiotic, particularly in TSB (Fig. S6, ESM).

Taking into account our predictions of AdpA operator sequences within the *lan* cluster, it is likely that the positive effect of *adpA* overexpression is mediated by the *lanI* gene which is encoding a crucial pathway-specific regulator of landomycin biosynthesis (Rebets et al. 2008). We transferred each aforementioned *adpA* expression vector (pGMSCY, pGMSCO, pGMSCLA, pGM4181, and pOOB95d) individually into a *lanI* null mutant *S. cyanogenus* ΔlanI7. Landomycin production by pGMSCY$^+$, pGMSCO$^+$, pGMSCLA$^+$, pGM4181$^+$, and pOOB95d$^+$ ΔlanI7 strains was studied in SG, ISP5,

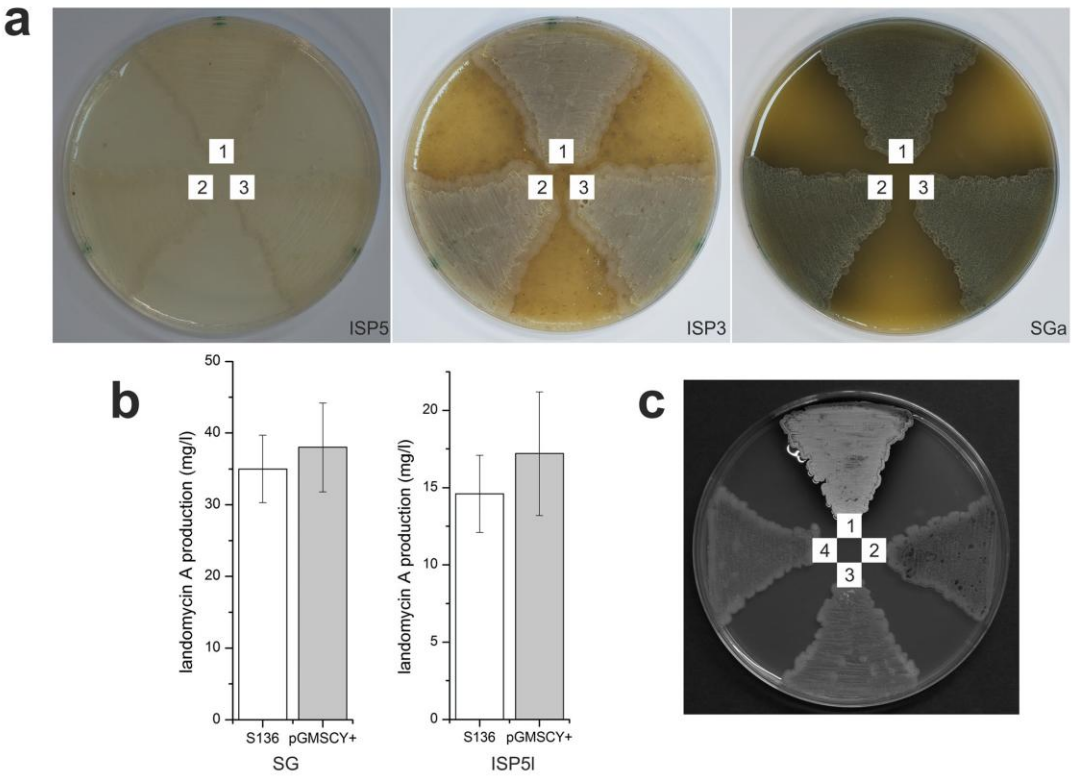


Fig. 4 Exploring the function of *S. cyanogenus* S136 *adpA* ortholog, *SCY4743*. **a** Morphology of *S. cyanogenus* S136 (1) in comparison to *S. cyanogenus* pGMSCY$^+$ (2) and pGMSCYd$^+$ (3) on the fifth day of growth on different solid media. No differences in growth and landomycin production were observed. **b** The LaA production by *S. cyanogenus* S136 and *S. cyanogenus* pGMSCY$^+$ in ISP5l and SG liquid media. Mean values ± SEM are shown. **c** Heterologous $adpA_{scy}$ and $adpA_{scy_dbd}$ (4—*S. coelicolor* M851 pGMSCY$^+$; 3—*S. coelicolor* M851 pGMSCYd$^+$; *S. coelicolor* M145) failed to complement Bld phenotype of *S. coelicolor* M851 (2) on R5 medium (3 days of growth)

and TSB media, and under no conditions, we did detect landomycins (Fig. S6b, ESM). Therefore, the AdpA-mediated boost of LaA production by *S. cyanogenus* S136 is dependent on *lanI* gene.

To further corroborate this finding, we have compared the expression of *lanI*, as well as some other genes from *lan* cluster (particularly, *lanA*, *lanJ*, and *lanG*), in *S. cyanogenus* S136 and *S. cyanogenus* pGM4181$^+$ by means of semi-quantitative RT-PCR. We have analyzed gene expression in cultures grown in TSB and SG media for 45 h. In the case of SG, *lanI* was expressed stronger in *S. cyanogenus* pGM4181$^+$; consequently, the expression of structural *lan* genes was higher (Fig. 5c). Modest expression of *lan* genes was observed in *S. cyanogenus* S136 grown in TSB for 72 h (see the first section of the "Results" section). Here, at an earlier time point, we have detected no expression of *lanI*, *lanA*, *lanJ*, and *lanG* in *S. cyanogenus* S136 grown in TSB (Fig. 5c), whereas these genes were abundantly expressed in *S. cyanogenus* pGM4181$^+$ under the same conditions.

Defined medium for landomycin A production by *adpA*-overexpressing *S. cyanogenus* strains

The *adpA*-enhanced landomycin production is not strictly medium dependent under submerged conditions, as it is the case for the wild type. We therefore decided to test a few minimal media in hope to find those that would support LaA production at or above the level observed for the wild type grown in SG. After some trials, we came up with a suitable medium referred to as SWS. Its composition was as follows (g/l): sucrose—20 g, NaCl—4 g, $(NH_4)_2SO_4$, trace element solution, same as the one used to prepare ISP5. We failed to detect landomycins in extracts from SWS-grown *S. cyanogenus* S136 even using TLC of highly concentrated samples (Fig. S7, ESM). In contrast, heterologous *adpA*-carrying strains produced almost 70 mg/l, which exceeded by 15% the LaA yields achieved by *S. cyanogenus* S136 in SG medium.

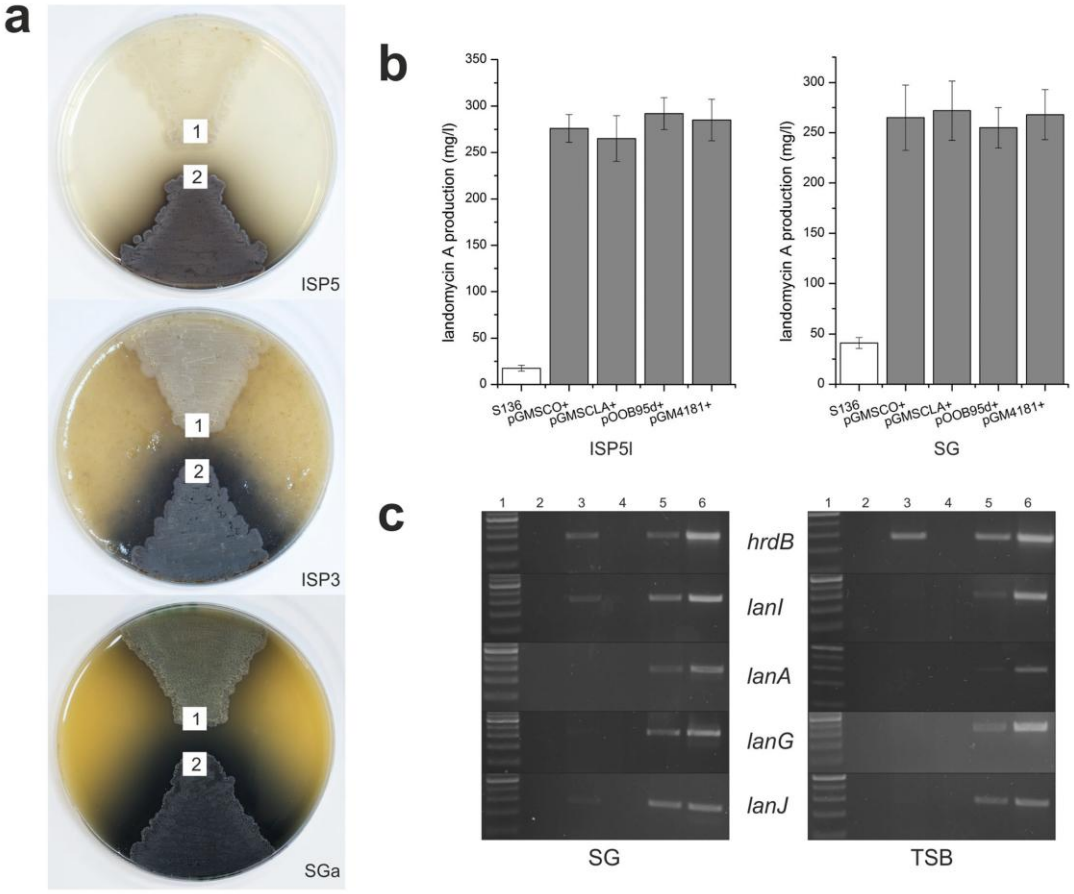


Fig. 5 Heterologous *adpA* genes boost LaA production by *S. cyanogenus* under different cultivation conditions. **a** Solid cultures of *S. cyanogenus* S136 (1) and its derivative, *S. cyanogenus* pGM4181^{+}, carrying *S. albus* J1074 *adpA* ortholog, *XNR_4181* (2). Photos were taken on the fifth day of growth. **b** LaA production titers of *S. cyanogenus* S136 carrying different heterologous *adpA* genes in two liquid media, ISP5 and SG. Time of cultivation, 72 h. Data is represented as mean values ± SEM. **c** RT-PCR analysis of *lanI*, *lanJ*, *lanG*, and *lanA* genes expression in cultures grown in SG and TSA media for 45 h. Lanes: 1—GeneRuler DNA Ladder Mix; 2—PCR using RNA isolated from *S. cyanogenus* S136 as a template (negative control); 3—PCR using cDNA synthesized from RNA isolated from *S. cyanogenus* S136 as a template, 4—PCR using RNA isolated from *S. cyanogenus* pGM4181^{+} as a template (negative control), 5—PCR using cDNA synthesized from RNA isolated from *S. cyanogenus* pGM4181^{+} as a template, 6—PCR using genomic DNA of *S. cyanogenus* S136 (positive control)

Discussion

Landomycin A is one of the archetypal representatives of angucycline antibiotics, the largest subfamily of aromatic polyketides (Kharel et al. 2012). From the very beginning, research efforts were mainly focused on elucidation of structural *lan* genes, leading to a rich repertoire of landomycin analogs. Besides the roles of cluster-situated genes coding for transcriptional factors LanI and LanK (Rebets et al. 2008; Ostash et al. 2008), little is known about the global physiological and genetic factors involved in LaA production. In this work, we provide initial insights into these issues through analysis of LaA accumulation in different liquid and solid media, and elucidation of the role of global regulatory genes of *adpA* family in LaA biosynthesis.

Our results confirm the empirical notion that high-level LaA production requires soytone-based liquid medium SG; the cobalt amendment increases LaA titers, although is not necessary for the production. Finally, we revealed that 4% glucose does not repress LaA biosynthesis, yet it also does not lead to increases in relative LaA productivity. LaA accumulation is very low or completely absent in many liquid media, and TSB is an illustrious example of the latter. Nevertheless, such a phenotypic crypticity is not a result of a complete lack of transcribed *lan*

genes, as our RT-PCR analysis showed. Out of more than 30 *lan* genes, only four were profiled here, and we cannot rule out the scenario where the transcription of some key *lan* genes is completely arrested in TSB; thus, the net result would be derailed landomycin biosynthesis. It is also likely that the observed decrease in *lan* genes transcription could be enough to block LaA production. Indeed, recent studies underline the fact that the cryptic secondary metabolism gene clusters of *Streptomyces* are not transcriptionally silent (Ahmed et al. 2017). We believe that LaA production is an interesting and valuable model of the conditionally cryptic biosynthetic pathway. Its further studies may lead to new ideas of how to activate the silent metabolic pathways of actinobacteria.

We sequenced the *S. cyanogenus* S136 genome and used the generated sequence to identify gene for the global transcriptional regulator AdpA. Genes of this family are known to influence antibiotic biosynthesis, and, from the presence of AdpA-like operator sequences in the *lan* cluster, we anticipated that the *S. cyanogenus* AdpA would be involved in LaA biosynthesis as well. Contrary to expectations, the AdpA-encoding gene *SCY4743* (as well as *SCY4143_dbd* expressed separately) had no impact on the morphology and antibiotic production of the native *S. cyanogenus* and the heterologous host *adpA*-minus mutant *S. coelicolor* M851. As the sequence encoding DNA-binding domain had no effects on the morphology and secondary metabolism of *S. cyanogenus* and *S. coelicolor*, it might be that causal mutations reside in the former. Multiple sequence alignment of $AdpA_{scy}$ and several best studied (and functional) AdpA proteins reveals a number of substitutions and indels scattered across C-terminal part of AdpA (see ESM). Insertion of valine codon is perhaps the most prominent among them, as it is embedded into second helix-turn-helix motif (HTH-2) of DNA-binding domain. However, it is difficult to count for this insertion as the reason for inactive $AdpA_{scy}$ given that HTH-2 has no decisive role in DNA recognition (Yao et al. 2013). More detailed analysis of *SCY4743* is underway in our laboratories, which should help us to identify aberrant mutation(s) within this gene. To the best of our knowledge, *SCY4743* is the first case of a naturally occurring gene for a nonfunctional AdpA ortholog. This finding notwithstanding, we think that AdpA-mediated regulation *was* important for LaA biosynthesis. Two points of evidence speak in favor of our hypothesis: first, the *lan* cluster still carries potential AdpA operators; second, the heterologous *adpA* genes indeed lead to dramatic increases in LaA production. Moreover, the heterologous *adpA* genes induced the production of landomycin A in TSB media, which in fact is the case of activation of cryptic (at least under our experimental conditions) biosynthetic gene cluster. We believe that AdpA regulates *lan* cluster expression via *lanI*: the identification of putative AdpA operators within *lanI* promoter region, RT-PCR data along with the absence of LaA production in Δ*lanI adpA*-overexpressing strain support this assumption. Thus, the "crypticity" of the *lan* cluster under different cultivation conditions could be considered a consequence of *S. cyanogenus* having an *adpA*-deficient background. It is also interesting whether the expression of heterologous *adpA* genes leads to activation of other cryptic gene clusters in *S. cyanogenus*. Investigations along this line are ongoing in our laboratories.

Using the heterologous *adpA*, we generated a number of S136 derivatives that routinely demonstrated tenfold increase in LaA titers over the wild type when grown in the optimal medium SG. This is the highest level of LaA productivity reported so far (Rebets et al. 2008; Ostash et al. 2009), attesting to the practical value of *adpA* as a tool for the manipulation of secondary metabolome. Furthermore, we tested LaA productivity of heterologous, *adpA*-expressing strains not only in SG but also in the minimal liquid ISP5l medium, where S136 produced only residual amounts of LaA. The high LaA production by $adpA^+$ strains in ISP5l encouraged us to examine even more depleted production media. We have designed one of the media, called SWS, with sucrose and ammonium sulfate as carbon and nitrogen sources, respectively. LaA production by the chosen $adpA^+$ recombinant strains in SWS exceeded the level observed for parent (S136) strain cultivated in SG. The SWS medium is based on inexpensive components and might be used to dissect the effects of different nutritional and environmental factors on landomycin production and the secondary metabolism of *S. cyanogenus* S136.

Acknowledgements We thank Simon Shaw (Technical University of Denmark, Copenhagen) for careful proofreading of the manuscript.

Funding This work was supported by grants Bg-46F (to V. F.) and M/26-2018 and Bg-41Nr (to B.O.) from the Ministry of Education and Science of Ukraine.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

Ahmed Y, Rebets Y, Tokovenko B, Brötz E, Luzhetskyy A (2017) Identification of butenolide regulatory system controlling secondary metabolism in *Streptomyces albus* J1074. Sci Rep 7(1):9784. https://doi.org/10.1038/s41598-017-10316-y

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37(Web Server issue):W202–W208. https://doi.org/10.1093/nar/gkp335

Claesen J, Bibb MJ (2011) Biosynthesis and regulation of grisemycin, a new member of the linaridin family of ribosomally synthesized peptides produced by *Streptomyces griseus* IFO 13350. J Bacteriol 193(10):2510–2516. https://doi.org/10.1128/JB.00171-11

Du YL, Li SZ, Zhou Z, Chen SF, Fan WM, Li YQ (2011) The pleiotropic regulator $AdpA_{ch}$ is required for natamycin biosynthesis and morphological differentiation in *Streptomyces chattanoogensis*. Microbiology 157(Pt 5):1300–1311. https://doi.org/10.1099/mic.0.046607-0

Henkel T, Rohr J, Beale JM, Schwenen L (1990) Landomycins, new angucycline antibiotics from *Streptomyces sp.* I. Structural studies on landomycins A-D. J Antibiot 43(5):492–503

Higashi T, Iwasaki Y, Ohnishi Y, Horinouchi S (2007) A-factor and phosphate depletion signals are transmitted to the grixazone biosynthesis genes via the pathway-specific transcriptional activator GriR. J Bacteriol 189(9):3515–3524. https://doi.org/10.1128/JB.00055-07

Higo A, Hara H, Horinouchi S, Ohnishi Y (2012) Genome-wide distribution of AdpA, a global regulator for secondary metabolism and morphological differentiation in *Streptomyces*, revealed the extent and complexity of the AdpA regulatory network. DNA Res 19(3):259–273. https://doi.org/10.1093/dnares/dss010

Horbal L, Fedorenko V, Luzhetskyy A (2014) Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. Appl Microbiol Biotechnol 98(20):8641–8655. https://doi.org/10.1007/s00253-014-5918-x

Kharel MK, Pahari P, Shepherd MD, Tibrewal N, Nybo SE, Shaaban KA, Rohr J (2012) Angucyclines: biosynthesis, mode-of-action, new natural products, and synthesis. Nat Prod Rep 29(2):264–325. https://doi.org/10.1039/c1np00068c

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical Streptomyces genetics. John Innes Foundation, Norwich

Korynevska A, Heffeter P, Matselyukh B, Elbling L, Micksche M, Stoika R, Berger W (2007) Mechanisms underlying the anticancer activities of the angucycline landomycin E. Biochem Pharmacol 74(12):1713–1726. https://doi.org/10.1016/j.bcp.2007.08.026

Kuzniar A, van Ham RC, Pongor S, Leunissen JA (2008) The quest for orthologs: finding the corresponding gene across genomes. Trends Genet 24(11):539–551. https://doi.org/10.1016/j.tig.2008.08.009

Lehka LV, Panchuk RR, Berger W, Rohr J, Stoika RS (2015) The role of reactive oxygen species in tumor cells apoptosis induced by landomycin A. Ukr Biochem J 87(5):72–82

Liu G, Chater KF, Chandra G, Niu G, Tan H (2013) Molecular regulation of antibiotic biosynthesis in *Streptomyces*. Microbiol Mol Biol Rev 77(1):112–143. https://doi.org/10.1128/MMBR.00054-12

López-García MT, Santamarta I, Liras P (2010) Morphological differentiation and clavulanic acid formation are affected in a *Streptomyces clavuligerus adpA*-deleted mutant. Microbiology 156(8):2354–2365. https://doi.org/10.1099/mic.0.035956-0

Makitrynskyy R, Ostash B, Tsypik O, Rebets Y, Doud E, Meredith T, Luzhetskyy A, Bechthold A, Walker S, Fedorenko V (2013) Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. Open Biol 3(10):130121. https://doi.org/10.1098/rsob.130121

Martín JF, Liras P (2010) Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in *Streptomyces*. Curr Opin Microbiol 13(3):263–273. https://doi.org/10.1016/j.mib.2010.02.008

Metzker ML (2010) Sequencing technologies - the next generation. Nat Rev Genet 11(1):31–46. https://doi.org/10.1038/nrg2626

Myronovskyi M, Brötz E, Rosenkränzer B, Manderscheid N, Tokovenko B, Rebets Y, Luzhetskyy A (2016) Generation of new compounds through unbalanced transcription of landomycin A cluster. Appl Microbiol Biotechnol 100(21):9175–9186. https://doi.org/10.1007/s00253-016-7721-3

Ohnishi Y, Yamazaki H, Kato JY, Tomono A, Horinouchi S (2005) AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. Biosci Biotechnol Biochem 69(3):431–439. https://doi.org/10.1271/bbb.69.431

Ostash B, Rix U, Rix LL, Liu T, Lombo F, Luzhetskyy A, Gromyko O, Wang C, Braña AF, Méndez C, Salas JA, Fedorenko V, Rohr J (2004) Generation of new landomycins by combinatorial biosynthetic manipulation of the LndGT4 gene of the landomycin E cluster in *S. globisporus*. Chem Biol 11(4):547–555. https://doi.org/10.1016/j.chembiol.2004.03.011

Ostash I, Ostash B, Luzhetskyy A, Bechthold A, Walker S, Fedorenko V (2008) Coordination of export and glycosylation of landomycins in *Streptomyces cyanogenus* S136. FEMS Microbiol Lett 285(2):195–202. https://doi.org/10.1111/j.1574-6968.2008.01225.x

Ostash B, Korynevska A, Stoika R, Fedorenko V (2009) Chemistry and biology of landomycins, an expanding family of polyketide natural products. Mini Rev Med Chem 9(9):1040–1051

Pan Y, Liu G, Yang H, Tian Y, Tan H (2009) The pleiotropic regulator AdpA-L directly controls the pathway-specific activator of nikkomycin biosynthesis in *Streptomyces ansochromogenes*. Mol Microbiol 72(3):710–723. https://doi.org/10.1111/j.1365-2958.2009.06681.x

Panchuk RR, Lehka LV, Terenzi A, Matselyukh BP, Rohr J, Jha AK, Downey T, Kril IJ, Herbacek I, van Schoonhoven S, Heffeter P, Stoika RS, Berger W (2017) Rapid generation of hydrogen peroxide contributes to the complex cell death induction by the angucycline antibiotic landomycin E. Free Radic Biol Med 106:134–147. https://doi.org/10.1016/j.freeradbiomed.2017.02.024

Rabyk M, Yushchuk O, Rokytskyy I, Anisimova M, Ostash B (2018) Genomic insights into evolution of AdpA family master regulators of morphological differentiation and secondary metabolism in *Streptomyces*. J Mol Evol 86:204–215. https://doi.org/10.1007/s00239-018-9834-z

Rebets Y, Ostash B, Luzhetskyy A, Hoffmeister D, Brana A, Mendez C, Salas JA, Bechthold A, Fedorenko V (2003) Production of landomycins in *Streptomyces globisporus* 1912 and *S. cyanogenus* S136 is regulated by genes encoding putative transcriptional activators. FEMS Microbiol Lett 222(1):149–153

Rebets Y, Dutko L, Ostash B, Luzhetskyy A, Kulachkovskyy O, Yamaguchi T, Nakamura T, Bechthold A, Fedorenko V (2008) Function of *lanI* in regulation of landomycin A biosynthesis in *Streptomyces cyanogenus* S136 and cross-complementation studies with *Streptomyces* antibiotic regulatory proteins encoding genes. Arch Microbiol 189(2):111–120. https://doi.org/10.1007/s00203-007-0299-5

Sambrook J, Russell DW (2001) Molecular cloning, a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, New-York

Takano E, Tao M, Long F, Bibb MJ, Wang L, Li W, Buttner MJ, Bibb MJ, Deng ZX, Chater KF (2003) A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. Mol Microbiol 50(2):475–486

Wang J, Schully KL, Pettis GS (2009) Growth-regulated expression of a bacteriocin, produced by the sweet potato pathogen *Streptomyces ipomoeae*, that exhibits interstrain inhibition. Appl Environ Microbiol 75(5):1236–1242. https://doi.org/10.1128/AEM.01598-08

Xu J, Zhang J, Zhuo J, Li Y, Tian Y, Tan H (2017) Activation and molecular mechanism of a cryptic oviedomycin biosynthetic gene cluster via the disruption of a global regulatory gene *adpA* in *Streptomyces ansochromogenes*. J Biol Chem 292(48):19708–19720. https://doi.org/10.1074/jbc.M117.809145

Yu P, Bu QT, Tang YL, Mao XM, Li YQ (2018) Bidirectional regulation of AdpA(ch) in controlling the expression of *scnRI* and *scnRII* in the natamycin biosynthesis of *Streptomyces chattanoogensis* L10. Front Microbiol 157:1300–1311. https://doi.org/10.3389/fmicb.2018.00316

Yao MD, Ohtsuka J, Nagata K, Miyazono KI, Zhi Y, Ohnishi Y, Tanokura M (2013) Complex structure of the DNA-binding domain of AdpA, the global transcription factor in *Streptomyces griseus*, and a target duplex DNA reveals the structural basis of its tolerant DNA sequence specificity. J Biol Chem 288:31019–31029. https://doi.org/10.1074/jbc.M113.473611

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s00253-018-9249-1.
Електронні додатки до статті доступні за посиланням:
https://static-content.springer.com/esm/art%3A10.1007%2Fs00253-018-9249-1/MediaObjects/253_2018_9249_MOESM1_ESM.pdf.

Анотація: *Streptomyces cyanogenus S136 є єдиним відомим продуцентом ландоміцину А (LaA), одного з найбільших глікозильованих ангуциклінових антибіотиків, що володіє сильними антипроліферативними властивостями. Зростає інтерес до з'ясування механізмів дії ландоміцинів, що, у свою чергу, вимагає доступу до великої кількості чистої речовини. Надпродукції LaA було досягнуто в минулому шляхом маніпулювання регуляторними генами, розташованими в кластері. Однак, інші компоненти регуляторної мережі біосинтезу LaA залишаються невідомими. Щоб заповнити цю прогалину, ми з'ясували внесок плейотропних регуляторів родини AdpA у продукцію ландоміцину через експресію генів adpA різного походження в S. cyanogenus S136. Надекспресія нативного ортолога adpA S. cyanogenus S136 не впливала на продукцію ландоміцину. У той же час, експресія кількох гетерологічних генів adpA привела до значного збільшення продукції ландоміцину за різних умов культивування. Отже, гетерологічні гени adpA є корисним інструментом для посилення або активації продукції ландоміцину S. cyanogenus. Подальші зусилля зосереджені на ідентифікації мутацій, які роблять S. cyanogenus AdpA нефункціональним.*

7.3 Зміни вторинного метаболому *S. cyanogenus* S136 у відповідь на експресію гетерологічних алелів гена *adpA*

Як випливає із наших результатів, в *S. cyanogenus* S136 ген $adpA_{scy}$ кодує нефункціональний білок, а експресія гетерологічних алелів *adpA* (перелічених в **Підрозділі 7.2**) активує біосинтез ландоміцину А в умовах, коли S136 зазвичай не продукує цей антибіотик. Також, ми спостерігали накопичення невідомої сполуки із темним забарвленням, яка з'являлася в *S. cyanogenus* ΔlanI7, де експресуються ці алелі. Ці спостереження спонукали нас глибше вивчити вторинний метаболом штамів *S. cyanogenus*, що несуть гетерологічний алель *adpA* з *S. albus* (=*albidoflavus*) J1074. Для цих експериментів ми використали згаданий вище мутант *S. cyanogenus* ΔlanI7, який не здатний до продукції ландоміцину [24]. Цей мутантний штам був обраний і для того, щоб усунути конкуренцію за попередники (ацил-КоА, цукри) біосинтетичного шляху ландоміцину А з потенційними активованими КБГ, що кодують біосинтез споріднених сполук (наприклад, полікетидів). Для досліджень ми використали штами *S. cyanogenus* ΔlanI7 pGM4181d$^+$ (надекспресує ділянку *adpA S. albus,* що кодує ДНК-зв'язувальний домен) і ΔlanI7 pGM4181$^+$. Ми культивували ці штами у порівнянні із диким типом на різноманітних агаризованих середовищах та перевіряли їх антимікробні властивості методом тестів на пригнічення росту тест-культур Грам-позитивних, Грам-негативних бактерій і дріжджових грибів.

Виявилося, що *S. cyanogenus* ΔlanI7 pGM4181$^+$, але не pGM4181d$^+$, накопичує темно-коричневу сполуку при вирощуванні на середовищі TSA або в середовищі TSB. Ми не змогли виділити цю сполуку з культуральних бульйонів або агаризованого середовища звичайними розчинниками, але припускаємо, що це можуть бути полімери на основі тирозину – меланіни. Подальші дослідження виявили, що *S. cyanogenus* ΔlanI7 виявляє слабку активність проти Грам-позитивних бактерій при вирощуванні на агарі ISP5. Така активність була значно підвищена у *S. cyanogenus* ΔlanI7 pGM4181$^+$ і pGM4181d$^+$. Натомість, ні вихідний, ні рекомбінантні штами не виявили

активності проти *E. coli*. Також, експресія *adpA* в *S. cyanogenus* ΔlanI7 викликала сильну протигрибкову активність, відсутню в контрольному штамі. Розвиваючи цей напрямок, ми перевірили низку агаризованих середовищ і виявили, що багато з них сприяють продукції невідомої протигрибкової сполуки і в *S. cyanogenus* ΔlanI7 pGM4181d$^+$. Середовище YMPG сприяло найкращій продукції невідомої сполуки із протигрибковою активністю.

Далі ми вирощували ΔlanI7 та його похідне pGM4181$^+$ у рідкому середовищі YMPG і аналізували метанольні екстракти з біомаси та супернатанту за допомогою ВЕРХ-МС; те саме було зроблено паралельно для дикого типу (S136) та його похідного, що несе pGM4181. Піки ландоміцину А та його попередників домінували на хроматограмах екстрактів S136 і S136 pGM4181$^+$; не було виявлено якісних відмінностей між хроматограмами ВЕРХ цих штамів. Натомість, принаймні чотири нові мас-піки з'явилися на хроматограмах екстрактів ΔlanI7 pGM4181$^+$. Значення *m/z*, що відповідають менш інтенсивним мас-пікам, не мали збігів у базах даних природних сполук. Однак, молекулярна маса найбільш помітного «активованого» іонного піку ([M+H]$^+$=708,35 Да) збігалася з молекулярною масою протигрибкового полієну люцензоміцину [144].

Ми визначили оптимальні умови для продукції люцензоміцину в *S. cyanogenus* ΔlanI7 pGM4181$^+$. Для цього ми використали рідке середовище YMPG. Для виявлення люцензоміцину, ми екстрагували за допомогою метанолу зразки біомаси *S. cyanogenus* ΔlanI7 pGM4181$^+$ зібрані на 24, 48, 72, 96 та 120 год культивування. Отримані екстракти використовували для тестів пригнічення росту *D. hansenii* методом дифузії антибіотика із паперового диску в агар. Продукція протигрибкової сполуки стала помітною вже на 48 год росту та досягла свого піку через 72 год. Поява зон пригнічення росту *D. hansenii* навколо дисків, просякнутих метанольними екстрактами, корелювала з відносною інтенсивністю мас-піку 708,35 Да на хроматограмах ВЕРХ тих самих екстрактів. Ми не виявили жодної протигрибкової активності в екстрактах біомаси ΔlanI7, зібраних на ті самі часові точки культивування в середовищі YMPG.

Використовуючи оптимізовані умови для культивування *S. cyanogenus* ΔlanI7 pGM4181$^+$, нам вдалося очистити прибл. 40 мг цільової протигрибкової сполуки з культур, вирощених в 2 л колбах Ерленмеєра із середовищем YMPG. Аналіз отриманої сполуки за допомогою ВЕРХ-МС і МС/МС продемонстрував маси (очікувана *m/z* [M-H]$^-$ для люцензоміцину = 706,3444 Да, спостережені *m/z* [M-H]$^-$=706,3447, *m/z* [M+H]$^+$=708,3586) та фрагментаційний спектр, що повністю співпадає зі структурою люцензоміцину.

Наш аналіз генома *S. cyanogenus* продемонстрував (див. **Підрозділ 7.1**), що один КБГ має значну подібність до КБГ полікетиду I-го типу пімарицину (також відомого як натаміцин) (*pim*) [145–147]. Оскільки люцензоміцин і пімарицин є структурними гомологами [148], ми припустили, що активація цього конкретного КБГ веде до накопичення люцензоміцину в *S. cyanogenus* ΔlanI7 pGM4181$^+$ і pGM4181d$^+$. Ми назвали цей КБГ *lcm*, а переанотована вручну послідовність цього КБГ була задепонована в GenBank під номером MW175861. Всього, ми анотували 19 ORF в межах *lcm*, серед яких 18 виявили схожість з генами, знайденими в інших КБГ поліенових антибіотиків. Ми скористалися наявними відомостями про біосинтез пімарицину та структурною подібністю між ним і люцензоміцином, щоб реконструювати біосинтетичний каскад для останнього.

Щоб зрозуміти механізм AdpA-опосередкованої активації біосинтезу люцензоміцину, ми проаналізували транскрипцію вибраних *lcm* генів у штамах *S. cyanogenus* S136, S136 pGM4181$^+$, ΔlanI7 і ΔlanI7 pGM4181$^+$ за допомогою ЗТ-ПЛР. Вибрані гени кодують ГТФ-D-манозо 4,6-дегідратазу (*lcm2*), монооксигеназу (*lcm6*), ABC-транспортер (*lcm8*), холестерол оксидаза (*lcm11*), ключовий шлях-специфічний регулятор (*lcmRIII*) і ПКС (*lcmC*). Продукти цих генів беруть участь в різних етапах біосинтезу, регуляції та експорту люцензоміцину, а отже досить повно відображають експресію генів КБГ

Спочатку ми перевірили експресію обраних *lcm* генів в культурах *S. cyanogenus* ΔlanI7 і *S. cyanogenus* ΔlanI7 pGM4181$^+$, вирощених у середовищі YMPG протягом 72 год. *lcm* транскрипти були відсутніми в *S. cyanogenus* ΔlanI7, але присутніми в штамі ΔlanI7 pGM4181$^+$. Далі ми

проаналізували біосинтез люцензоміцину та експресію *lcm* генів у штамів-продуцентів ландоміцину А *S. cyanogenus* S136 та S136 pGM4181$^+$, культивованих протягом 72 год у середовищі SG (оптимальному для продукції ландоміцину А) і YMPG (оптимальному для продукції люцензоміцину). Ми не виявили ознак накопичення люцензоміцину в S136 при культивуванні на агаризованих середовищах SG і YMPG, але слабка протигрибкова активність була виявлена у випадку S136 pGM4181$^+$, вирощеного на YMPG. Хоча й продукцію люцензоміцину в культурах S136 pGM4181$^+$ у SG неможливо було виявити за допомогою тестів пригнічення росту *D. hansenii*, *lcm* гени тут експресувалися.

Коли ми кількісно визначили накопичення люцензоміцину в екстрактах різних штамів *S. cyanogenus* вирощених у рідкому YMPG за допомогою ВЕРХ-МС, то виявилося, що сигнал люцензоміцину є ледь помітним в S136 і ΔlanI7 (прибл. 10^3 та 10^4 ум.од., відповідно). У S136 pGM4181$^+$ і ΔlanI7 pGM4181$^+$ інтенсивність сигналу люцензоміцину збільшилася приблизно на 2 і на 3 порядки, відповідно, порівняно з їх батьківськими штамами.

Отже, у дикому типі (S136) AdpA посилює експресію генів як люцензоміцину, так і ландоміцину А, хоча і продукція люцензоміцину спостерігається лише в кількостях достатніх для детекції методом ВЕРХ з мас-спектрометрією, але не методом пригнічення росту *D. hansenii*. Якщо інактивувати шлях біосинтезу ландоміцину А (штам ΔlanI7), надекспресія *adpA* веде до інтенсивної продукції люцензоміцину.

Наведені вище результати демонструють, що AdpA необхідний для експресії генів *lcm* і продукції люцензоміцину в *S. cyanogenus*. Усі описані на сьогоднішній день КБГ протигрибкових полієнів містять в своєму складі гени транскрипційних регуляторів, чиї продукти беруть участь в регуляції експресії генів КБГ BGC несуть гени регуляторів транскрипції [149]. КБГ *pim*, наприклад, містить два таких гени, *pimR* і *pimM* [150, 151], а КБГ ністатину чотири: *nysRI*, *nysRII*, *nysRIII* і *nysRIV* [40]. PimR – це великий мультидоменний білок, який, як відомо, має єдину мішень у КБГ *pim*: *pimM*. PimM складається з N-кінцевого ефекторного PAS-подібного домена та С-кінцевого LuxR-

подібного ДНК-зв'язувального домена. PimM активує експресію структурних генів *pim*. У біосинтезі ністатину NysRI, NysRII і NysRIII є паралогами, кооротологічними PimR, які відіграють некумулятивну роль в активації *nysRIV*, ортолога *pimM* [40]. Вищезгадані шлях-специфічні регулятори є вирішальними для біосинтезу відповідних сполук, бо нокаут відповідних генів веде до припинення біосинтезу поліенових антибіотиків. Кластер *lcm* містить три гени для шлях-специфічних транскрипційних регуляторів: *lcmRI*, *lcmRII* і *lcmRIII*. LcmRI є гомологом NysRIV, а LcmRII і LcmRIII є гомологами PimM та PimR.

Щоб визначити можливі мішені AdpA, ми проаналізували кластер *lcm in silico* в пошуках потенційних сайтів зв'язування AdpA. Два суміжних сайти зв'язування (TGGCGGAAAC, TGGCCGATAG; $p < 10^{-4}$) були знайдені в ділянці промотора гена *lcmRIII*. Це, разом з даними про транскрипційну активність генів *lcm* (див. вище), свідчить про те, що AdpA активує транскрипцію генів *lcm* шляхом безпосередньої активації *lcmRIII*. Як і у випадку біосинтезу пімарицину, LcmRIII повинен потім активувати транскрипцію або *lcmRI*, або *lcmRII*, або обох. Оскільки LcmRI і LcmRII обидва є PimM/NysRIV-подібними регуляторами, важко зробити висновок, який з них відповідає за активацію генів *lcm*.

Для уточнення ролі шлях-специфічних регуляторних генів з *lcm* КБГ, ми надекспресували кожен із них індивідуально під контролем *ermEp* у *S. cyanogenus* ΔlanI7 і визначили рівні продукції люцензоміцину в отриманих штамах. Активація продукції люцензоміцину спостерігалася в рекомбінантних *S. cyanogenus* ΔlanI7 *lcmRIII*$^+$ та *lcmRII*$^+$, але не в *S. cyanogenus* ΔlanI7 *lcmRI*$^+$. Це дало змогу припустити, що AdpA не має мішеней у КБГ *lcm*, окрім *lcmRII* і *lcmRIII*. Ген *lcmRI*, ймовірно, кодує нефункціональний паралог *lcmRII*.

Якщо низький рівень продукції люцензоміцину, що спостерігався в штамі S136 pGM4181$^+$, є наслідком паралельної надпродукції ландоміцину А, то надекспресія шлях-специфічних регуляторів біосинтезу люцензоміцину повинна би привести до накопичення останнього в штамі S136. Дійсно, після надекспресії *lcmRIII* в S136 ми змогли виявити помітну протигрибкову активність рекомбінантного штаму при вирощуванні на агаризованих

середовищах, які сприяють продукції ландоміцину А (SG) або люцензоміцину (YMPG), а також на середовищі TSB, несприятливому для обох біосинтетичних шляхів. У той же час, надекспресія *lcmRIII* не вплинула на продукцію ландоміцину А в S136.

У підсумку, результати, отримані у цьому підрозділі, показують, що різні алелі гена *adpA* є ефективним знаряддям для активації мовчазних КБГ антибіотиків у *S. cyanogenus* S136, а, ймовірно, і в нших стрептоміцетів. Зокрема, нам вдалося активувати мовчазний КБГ полікетидного антибіотика люцензоміцину. Хоча сам антибіотик є відомим [148], наші результати вперше дали змогу ідентифікувати відповідний КБГ та повністю охарактеризувати біосинтетичний шлях люцензоміцину.

Результати, отримані в цьому підрозділі опубліковано в [152, 153].

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

Eliciting the silent lucensomycin biosynthetic pathway in *Streptomyces cyanogenus* S136 via manipulation of the global regulatory gene *adpA*

Oleksandr Yushchuk[1], Iryna Ostash[1], Eva Mösker[2], Iryna Vlasiuk[1], Maksym Deneka[1], Christian Rückert[3], Tobias Busche[3], Victor Fedorenko[1], Jörn Kalinowski[3], Roderich D. Süssmuth[2]✉ & Bohdan Ostash[1]✉

Actinobacteria are among the most prolific sources of medically and agriculturally important compounds, derived from their biosynthetic gene clusters (BGCs) for specialized (secondary) pathways of metabolism. Genomics witnesses that the majority of actinobacterial BGCs are silent, most likely due to their low or zero transcription. Much effort is put into the search for approaches towards activation of silent BGCs, as this is believed to revitalize the discovery of novel natural products. We hypothesized that the global transcriptional factor AdpA, due to its highly degenerate operator sequence, could be used to upregulate the expression of silent BGCs. Using *Streptomyces cyanogenus* S136 as a test case, we showed that plasmids expressing either full-length *adpA* or its DNA-binding domain led to significant changes in the metabolome. These were evident as changes in the accumulation of colored compounds, bioactivity, as well as the emergence of a new pattern of secondary metabolites as revealed by HPLC-ESI-mass spectrometry. We further focused on the most abundant secondary metabolite and identified it as the polyene antibiotic lucensomycin. Finally, we uncovered the entire gene cluster for lucensomycin biosynthesis (*lcm*), that remained elusive for five decades until now, and outlined an evidence-based scenario for its *adpA*-mediated activation.

Over decades, prokaryotes were one of the most prolific sources of antibiotics, rivaled only by lower fungi[1]. Among antibiotic-producing prokaryotes, the classes Actinobacteria and Myxobacteria currently dominate[2,3], although exploration of the other taxa could change this view significantly in the near future[4–6]. The actinobacterial genus *Streptomyces* probably still is the most successful provider of drugs for healthcare and agriculture[1]. Beginning with seminal works of Selman Waksman[7,8], the majority of known antibiotic compounds from *Streptomyces* were isolated through classical screening approaches. However, due to high rediscovery rates, this approach has gradually lost its efficiency[9]. More recently, the sequencing of *Streptomyces* spp. genomes revealed a large hidden potential for the production of antibiotics and other secondary metabolites. Exemplary studies of the fully sequenced genomes of *S. coelicolor* and *S. avermitilis* revealed many more biosynthetic gene clusters (BGCs) than compounds previously known to be produced by these species[10,11]. Ever since, each new sequenced *Streptomyces* genome gave rise to the idea that most of the actinobacterial BGCs are silent, i.e. that they are not associated with the production of certain compounds under typical cultivation conditions, most likely due to low or zero transcription[12,13].

Today's relative simplicity of in silico identification of BGCs within genomic sequences[14,15] prompted the search for approaches to activate silent BGCs. Certain mutations in genes for ribosomal proteins and RNA polymerases are capable of eliciting the expression of the silent secondary metabolome[16,17]. Extracellular triggers, like xenobiotics, rare earth elements or co-cultivation with other bacteria were also shown to activate the

[1]Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 4 Hrushevskoho st., Rm. 102, Lviv 79005, Ukraine. [2]Institut für Chemie, Technische Universität Berlin, Straße des 17. Juni 124, 10623 Berlin, Germany. [3]Technology Platform Genomics, CeBiTec, Bielefeld University, Universitätsstraße 27, 33615 Bielefeld, Germany. ✉email: suessmuth@chem.tu-berlin.de; bohdan.ostash@lnu.edu.ua

production of novel antibiotics by streptomycetes[18–22]. Manipulations of either regulatory elements within the putative BGCs[23,24], or BGC-situated regulatory genes[23,25,26] often lead to the activation of silent BGCs. As antibiotic production in streptomycetes is tightly intertwined with morphogenesis, global regulators that coordinate both aforementioned processes can be useful tools for activation of silent BGCs for known or cryptic biosynthetic pathways[27–32]. Such an approach could be especially useful for activation of the BGCs lacking cluster-situated regulatory genes.

The AraC family transcriptional factor (TF) AdpA is a global regulator of secondary metabolism and morphogenesis, and was successfully used to activate secondary metabolism[29,30]. AdpA appears as an appealing tool in this regard because it binds to a highly degenerate operator sequence[33] and thus possesses an enormous regulon[34]. The overexpression of such a promiscuous TF could enforce binding to numerous (probably unrecognizable by AdpA under physiological conditions) sites on the chromosome, impacting the expression of multiple genes and operons. The biosynthesis of many antibiotics in streptomycetes is AdpA-regulated and AdpA TFs are almost identical across different *Streptomyces* species and therefore likely display the same properties[35]. AdpA-binding sites are common in silent BGCs and overexpression of *adpA* genes could be used for high-throughput activation of such BGCs in strains of unknown genomic background.

In this work we used *Streptomyces cyanogenus* S136 as a proof-of-concept to probe the above conjecture. This strain is known to produce deeply colored angucyclines, polyketides from the landomycin family, in a medium-dependent fashion; furthermore, sequencing revealed the presence of > 30 BGCs in its genome. We show that the secondary metabolome of *S. cyanogenus* actively responds to the overexpression of heterologous *adpA* genes with the production of novel secondary metabolites. Here we focused on the genetic basis of the production of the most abundant compound, lucensomycin, known for its antifungal activity. From our results we deduced a scheme for AdpA-mediated regulation of lucensomycin biosynthesis. Finally, we explored the functional significance of different parts of the AdpA protein for its ability to induce lucensomycin production. Our results expand the current knowledge of the secondary metabolome of *S. cyanogenus* and offer new insights into the ways of activation of silent BGCs in *Streptomyces*.

Results

Genome of *Streptomyces cyanogenus* encodes sizable secondary metabolome that responds to *adpA*. In the genome of *S. cyanogenus* S136, in addition to the *lan* cluster coding for landomycin biosynthesis, at least 33 putative BGCs were found under the default antiSMASH search conditions. From these, 16 BGCs shared significant similarities to known BGCs (ESM Fig. S1). These were the BGCs for informatipeptin, pimaricin, naphtomycin, chlorothricin and different RiPPs, furthermore for terpenoids hopene, isorenieratene and albaflavenone, the siderophore desferrioxamine B, ectoine, melanins and the spore pigment of polyketide nature. It is worth mentioning that most of the antibiotic BGCs putatively encode the biosyntheses of polyketide compounds (e.g. pimaricin, naphtomycin, chlorothricin).

In *S. cyanogenus* S136, the gene *SCY4743* codes for a nonfunctional AdpA, and expression of heterologous *adpA* genes activated landomycin biosynthesis under conditions where S136 normally does not produce this antibiotic[36]. This observation prompted us to look deeper into the secondary metabolome of *S. cyanogenus* strains carrying heterologous versions of *adpA*. Initially, we chose a landomycin-nonproducing mutant ΔlanI7[37] for these experiments. The mutant strain was chosen in order to eliminate the competition for precursors (acyl-CoA, carbohydrates) of the landomycin pathway with potentially activated BGCs directing the production of biogenetically related compounds. Plasmids pGM4181 and pGM4181d carrying full *adpA* of *S. albus* J1074 and its DNA-binding domain (DBD), respectively, under the control of *moeE5p* promoter, were introduced into *S. cyanogenus* ΔlanI7. This promoter allows strong constitutive gene expression in *S. cyanogenus*[36].

S. cyanogenus ΔlanI7 pGM4181$^+$, but not pGM4181d$^+$, started producing a dark brown compound when grown on TSA plates or in TSB media (Fig. 1a). However, we were not able to extract this compound from culture broths or solid agar media with common solvents. Likewise, mass spectrometry of the culture broth failed to identify distinct mass peaks. We suspected that the observed pigmentation was due to accumulation of tyrosine-based polymers, melanins, because in several *Streptomyces* species production of the latter is known to be upregulated by AdpA[38]. The structure elucidation of these polymers remains elusive due to their heterogeneity and poor solubility[39,40]. We therefore tentatively suggest that melanin-type compounds are accumulated in *S. cyanogenus* in presence of *adpA*. Indeed, its genome carries two *melC* homologs responsible for melanin biosynthesis, and their expression is activated upon pGM4181 introduction (ESM Fig. S2).

S. cyanogenus ΔlanI7 displayed weak activity against Gram-positive bacteria when grown on ISP5 agar. Such an activity was significantly increased in *S. cyanogenus* ΔlanI7 pGM4181$^+$ and pGM4181d$^+$ (Fig. 1b). Neither initial nor recombinant strains showed activity against *E. coli* (data not shown). The *adpA* expression in *S. cyanogenus* ΔlanI7 elicited strong antifungal activity, apparently absent in the control strain (Fig. 1c). Encouraged by this finding, we tested a number of agar media and found that many of them (but not all, see ESM Fig. S3) supported the production of antifungal compound(s) by *S. cyanogenus* ΔlanI7 pGM4181$^+$ and pGM4181d$^+$. We noted that for certain media the expression of DBD of AdpA is sufficient to induce the occurrence of antifungal activity (ESM Fig. S3), but the full-length gene was found to be more efficient. YMPG medium supported the highest level of antifungal activity, and was used for further analysis of *adpA* effects on the *S. cyanogenus* metabolome.

We grew ΔlanI7 and its pGM4181$^+$ derivative in liquid YMPG medium and analyzed the biomass and supernatant extracts with LC–MS; the same was done in parallel for wild type (S136) and its pGM4181$^+$ derivative. The peaks for landomycin A and its precursors dominated the chromatograms of the extracts from S136 and S136 pGM4181$^+$; no qualitative differences were found between the HPLC chromatograms of these strains (ESM Fig. S4). In contrast, at least four new mass peaks occurred in the HPLC chromatograms of ΔlanI7 carrying pGM4181 (Fig. 1d). The *m/z* values corresponding to the less abundant mass peaks had no matches in the

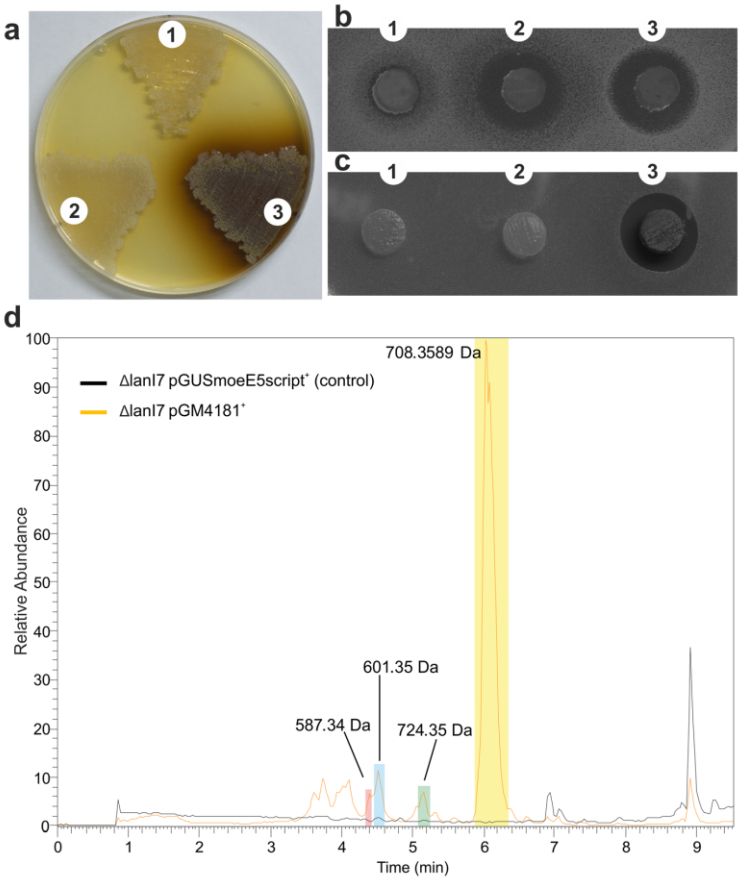


Figure 1. The *adpA* gene activates silent secondary metabolome of *S. cyanogenus*. (**a**) Phenotypes of *S. cyanogenus* ΔlanI7 (1), *S. cyanogenus* ΔlanI7 pGM4181d$^+$ (2; DBD of *adpA*) and *S. cyanogenus* ΔlanI7 pGM4181$^+$(3; full-length *adpA*) after 168 h of growth on TSB agar; *B. cereus* (**b**) and *D. hansenii* (**c**) growth inhibition assays using agar plugs taken from 120 h old ISP5 lawns of *S. cyanogenus* ΔlanI7 (1), *S. cyanogenus* ΔlanI7 pGM4181d$^+$ (2) and *S. cyanogenus* ΔlanI7 pGM4181$^+$(3). (**d**) LC–MS analysis of the secondary metabolites produced in liquid YMPG medium by *S. cyanogenus* ΔlanI7 pGUSmoeE5script$^+$ (control) and *S. cyanogenus* ΔlanI7 pGM4181d$^+$; four new mass peaks appeared in the chromatogram for the latter strain.

databases of natural compounds. The molecular mass of the most prominent "activated" ion peak (708.35 Da, $[M+H]^+$) matched to that of the polyene antibiotic lucensomycin (Lcm)[41,42], and thus we focused our attention on this peak. Lcm is known as antifungal compound, and its molecular mass peak is strongly prevailing in the HPLC–MS analytics of the metabolome of the ΔlanI7 pGM4181$^+$ strain. We suppose therefore that Lcm is responsible for the antifungal activity exhibited by this strain. Nevertheless, we cannot exclude the contribution of the other, as yet unknown, compounds accumulated by ΔlanI7 pGM4181$^+$ to antifungal activities of its extracts and agar plugs, although this contribution (if present) is likely insignificant.

Optimized conditions for expression of antifungal activity in *S. cyanogenus* ΔlanI7. We set out to identify the optimal conditions for the production of the antifungal compound by *S. cyanogenus* ΔlanI7 pGM4181$^+$ prior to its detailed structural verification. In liquid YMPG the production of an antifungal compound became detectable after 48 h of growth and reached its peak at 96 h (Fig. 2a). The bioassay data correlated with the abundance of the 708.35 Da peak in HPLC chromatograms of the same extracts (Fig. 2b). No antifungal activity was observed in the extracts of ΔlanI7 biomass harvested at 24 h, 48 h, 72 h, 96 h and 120 h (Fig. 2a).

Subsequently, we introduced into *S. cyanogenus* ΔlanI7 a set of full-length *adpA* genes and corresponding DBDs—*adpAdbd*—from different *Streptomyces* species. This was carried out to determine whether and to what extent different *adpA* genes induce the antifungal activity. All heterologous *adpA* genes were shown to be

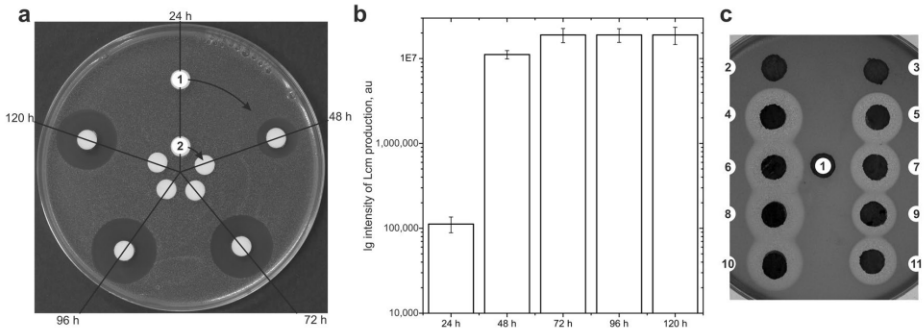


Figure 2. Antifungal activity of *S. cyanogenus* ΔlanI7 pGM4181$^+$ as a function of cultivation time (**a**, **b**) and origin of *adpA* gene (**c**). For *D. hansenii* growth inhibition assay (**a**) the extracts were prepared from YMPG-grown ΔlanI7 pGM4181$^+$ (1) and ΔlanI7 (2) at timepoints shown in the figure. The initial strain ΔlanI7 (2) showed no activity at all investigated time points. LC–MS quantification of the Lcm mass peak in the same extracts (**b**) was in agreement with the results of the bioassay. Data represent mean values of three independent experiments ± 2SD. (**c**) *D. hansenii* growth inhibition assay showing that *adpA* genes (and corresponding DNA-binding domains, DBD), other than *XNR_4181,* were able to activate Lcm production in *S. cyanogenus* ΔlanI7: (4,5)—ΔlanI7 pGMSCO$^+$ and pGMSCOd$^+$, overexpressing *S. coelicolor adpA* and its DBD, respectively; (6,7)—ΔlanI7 pGMSCLA$^+$ and pGMSCLAd$^+$, *S. clavuligerus adpA* and its DBD, respectively; (8,9)—ΔlanI7 pGM4181$^+$ and pGM4181d$^+$, *S. albus adpA* and its DBD, respectively; (10,11)—ΔlanI7 pOOB95d$^+$ and pGMSGHd$^+$, *S. ghanaensis adpA* and its DBD, respectively. Initial strain (1) as well as strains carrying native *adpA* from *S. cyanogenus* (ΔlanI7 pGMSCY$^+$) and its DNA-binding domain (ΔlanI7 pGMSCYd$^+$) do not inhibit growth of *D. hansenii.* Agar plugs were cut from the lawns grown on YMPG agar for 120 h.

strong inducers, and *adpAdbd*'s were less active in this regard (Fig. 2c). The overexpression of neither native *adpA* (*SCY4743*) nor its *adpAdbd* led to the activation of antifungal activity, once again pointing to the non-functionality of SCY4743 in *S. cyanogenus.*

The AdpA-induced compound corresponds to the polyene antibiotic lucensomycin. Using the optimized production conditions, we were able to purify 40 mg of target antifungal compound from ΔlanI7 pGM4181$^+$ grown in 2 L of YMPG (shaken flasks). In disk diffusion assay the compound showed activity only against fungal species (ESM Table S2, ESM Fig. S5). High-resolution mass spectrometric (MS) measurements (calculated for Lcm: 706.3444 Da ($[M-H]^-$), 708.3590 Da ($[M+H]^+$); observed: 706.3447 ($[M-H]^-$) and 708.3586 ($[M+H]^+$); mass tolerance ≤ 5 mmu) and the fragmentation pattern of the compound (ESM Fig. S6) are in full agreement with the structure of Lcm[42,43].

Description of the lucensomycin (*lcm*) BGC in *S. cyanogenus*. A survey of BGCs within the *S. cyanogenus* genome (see above) readily led us to recognize that one BGC shared significant similarity to the pimaricin (also known as natamycin) BGC (*pim* cluster)[44–46]. As Lcm and pimaricin are structural homologs[41], we suggested that activation of this particular BGC results in Lcm accumulation by *S. cyanogenus* ΔlanI7 pGM4181$^+$ and pGM4181d$^+$. We therefore refer to this BGC as the *lcm* (**lucenso**mycin) BGC (accession number MW175861). Besides the *lcm* BGC, a few other genes for type I polyketide synthases were found within the genome of *S. cyanogenus* (ESM Fig. S1). There was an overall synteny between the *lcm* BGC and the BGC already proposed to control Lcm production in *S. achromogenes* NRRL 3125[47]. However, the Lcm BGC from NRRL 3125 appeared to be incomplete (see below) and its nucleotide sequence is absent in publicly available databases; this precludes their detailed comparison. Nineteen ORFs were annotated within the *lcm* BGC (Table 1; Fig. 3a), among which 18 are similar to genes found in the other polyene BGCs (Table 1). We took advantage of detailed knowledge of the biosynthesis of pimaricin and structural similarities between the latter and Lcm to propose a tentative biochemical pathway leading to Lcm.

Assembly of Lcm aglycone. There are five PKS genes within the *lcm* BGC, four of which (*lcmB, lcmC, lcmD* and *lcmE*) are colinear to *pimS1, pimS2, pimS3* and *pimS4*. It is therefore reasonable to propose analogous functions for *pim* and *lcm* homologs in the polyene aglycone assembly (Fig. 3b). However, the *lcmA* gene encodes for a PKS in which the initiation module is apparently fused to the first PKS module. This kind of organization is unprecedented in the biosynthesis of polyenes, although widely known for other type I PKS-derived products, such as erythromycin[48]. Most likely, this reflects the unique structure of Lcm featuring a *n*-propyl side chain. From analysis of the LcmA domain order, we propose that KS_0 would activate methylmalonyl-CoA as the starter substrate, and the second module (featuring full reductive loop) would incorporate malonyl-CoA, thus leading to a fully reduced C5 unit. Indeed, by a number of in silico tools KS_0 and KS_1 are reliably predicted to bind methylmalonyl- and malonyl-CoA, respectively, lending support for our conjecture. The identified type II thioesterase Lcm3, a homolog of FscTE from the biosynthesis of polyene candicidin[49], would be assigned responsible for the proofreading in Lcm aglycone formation.

lcm	The closest homologue from antibiotic BGCs	AA sequence identity(%)/E-value	Putative product	Putative function in Lcm biosynthesis
lcm1	N/A		–	–
lcm2	AviE3 (avilamycin BGC)	61.92/5e^{-111}	GDP-D-mannose 4,6-dehydratase (COG1089)	Transformation of GDP-D-mannose into GDP-4-keto-6-deoxy-D-mannose
lcm3	FscTE (candicidin BGC)	60,08/5e^{-70}	Type II thioesterase (COG3208)	Removal of aberrant intermediates
lcmB			PKS modules 2–5	Lcm aglycone assembly line
lcmA			PKS initiation and 1 modules	Lcm aglycone assembly line
lcm4	AmphDI (amphotericin B BGC)	58,89/4e^{-156}	GDP-mycosamine glycosyltransferase (COG1819)	Glycosylation of Lcm aglycone with mycosamine
lcm5	FscMII (candicidin BGC)	74,43/3e^{-154}	GDP-3-keto-6-deoxy-D-mannose C-3 aminotransferase (COG0399)	Transformation of GDP-3-keto-6-deoxy-D-mannose into GDP-mycosamine
lcm6	PimG (pimaricin BGC)	64,78/5e^{-153}	Cytochrome P450 monooxygenase (COG2124)	Formation of carboxyl group at C-12
lcm7	RimH (rimocidin BGC)	66,13/8e^{-13}	Ferredoxin (COG1141)	Electron transfer in P450 system
lcmE			PKS module 12	Lcm aglycone assembly line
lcmD			PKS module 11	Lcm aglycone assembly line
lcmC			PKS modules 5–10	Lcm aglycone assembly line
lcm8	PimA (pimaricin BGC)	57,71/3e^{-147}	ABC transporter (COG1132)	Export of Lcm
lcm9	AmphG (amphotericin B BGC)	56,20/1e^{-171}	ABC transporter (COG1132)	Export of Lcm
lcm10	NysL (nystatin BGC)	42,46/3e^{-81}	Cytochrome P450 monooxygenase (COG2124)	Epoxidation of Lcm aglycone at C4-C5
lcmRI	NysRIV (nystatin BGC)	39,38/1e^{-29}	HTH-LuxR domain containing TF	Regulation of Lcm biosynthesis
lcmRII	PimM(pimaricin BGC)	40,1/2e^{-32}	HTH-LuxR domain containing TF	Regulation of Lcm biosynthesis
lcm11	PimE	79,34/0	Cholesterol oxidase (COG2303)	"Fungal sensor" for ergosterol detection
lcmRIII	PimR	77,9/0	Transcriptional regulator	Regulation of Lcm biosynthesis

Table 1. Description of the ORFs discovered in *lcm* gene cluster.

Post-PKS modifications of Lcm aglycone. This is achieved in three sequential tailoring reactions. First, a carboxy group is installed at position C-12 (Fig. 3), most probably via action of monooxygenase Lcm6 (Table 1). In the next tailoring reaction aminosugar mycosamine is attached. Genes responsible for the biosynthesis of mycosamine and its attachment to the aglycone are present within all polyene BGCs, and the *lcm* gene cluster is no exception to this finding. Gene *lcm2* encodes a GDP-D-mannose 4,6-dehydratase, most probably responsible for the transformation of GDP-D-mannose into GDP-4-keto-6-deoxy-D-mannose. The latter undergoes spontaneous isomerization forming GDP-3-keto-6-deoxy-D-mannose (GDPDM). GDPDM is converted by aminotransferase Lcm5 into GDP-mycosamine. From the latter, the mycosamine residue is transferred onto polyene scaffold by glycosyltransferase Lcm4. Epoxidation is the last tailoring reaction, and Lcm10 is the most likely candidate for catalysis of this reaction[50].

Pathway-specific regulation and export. All antifungal polyene BGCs described to date carry genes for transcriptional regulators[45]. The *pim* cluster, for instance, contains two such genes, PimR and PimM[51,52], the nystatin BGC four: NysRI, NysRII, NysRIII and NysRIV[53]. PimR is a large multidomain protein known so far to have a single target within the *pim* cluster: *pimM*. PimM consists of a N-terminal effector PAS-like domain and a C-terminal LuxR-like DNA binding domain. PimM upregulates structural *pim* genes. In nystatin biosynthesis NysRI, NysRII and NysRIII are paralogs co-orthologous to PimR that play a redundant role in the activation of *nysRIV*, a *pimM* ortholog. The aforementioned pathway-specific regulators are crucial for the biosynthesis of respective compounds. Of note is that regulatory genes apparently are not within the sequenced part of the Lcm BGC of *S. achromogenes* NRRL 3125[47]. In our case the flank of the *lcm* gene cluster possesses three genes for the pathway-specific transcriptional regulators: *lcmRI*, *lcmRII* and *lcmRIII* (Table 1). LcmRI is a NysRIV homolog, LcmRII and LcmRIII are homologs of PimM and PimR. There are also genes for ABC-transporters within the *lcm* cluster, *lcm8* and *lcm9*, encoding homologs of known ABC transporters involved in the export of pimaricin and amphotericin[54]. Finally, *lcm11* encodes a PimE homologue, which is functionally assigned to a cholesterol oxidase. The role of PimE remains obscure. It is crucial for pimaricin production and is believed to be a sensor which detects ergosterol in fungal membrane and somehow triggers the polyene export[55–57].

Relationships between AdpA, lucensomycin and landomycin biosynthetic pathways. To understand the mechanism of AdpA-mediated activation of Lcm biosynthesis, we analyzed the transcription of selected *lcm* genes in *S. cyanogenus* strains S136, S136 pGM4181^{+}, ΔlanI7 and ΔlanI7 pGM4181^{+} by means of semi-quantitative RT-PCR. Chosen genes code for GDP-D-mannose 4,6-dehydratase (*lcm2*), monooxygenase (*lcm6*), ABC transporter (*lcm8*), fungal sensor (*lcm11*), key pathway-specific regulator (*lcmRIII*) and PKS (*lcmC*). Their products are involved in different steps of Lcm biosynthesis.

First, we checked *lcm* gene expression in *S. cyanogenus* ΔlanI7 and *S. cyanogenus* ΔlanI7 pGM4181^{+} grown in YMPG for 72 h. The *lcm* transcripts were absent in *S. cyanogenus* ΔlanI7 and present in ΔlanI7 pGM4181^{+} strain (Fig. 4a). Next we analyzed the Lcm biosynthesis and *lcm* gene expression in landomycin-producing

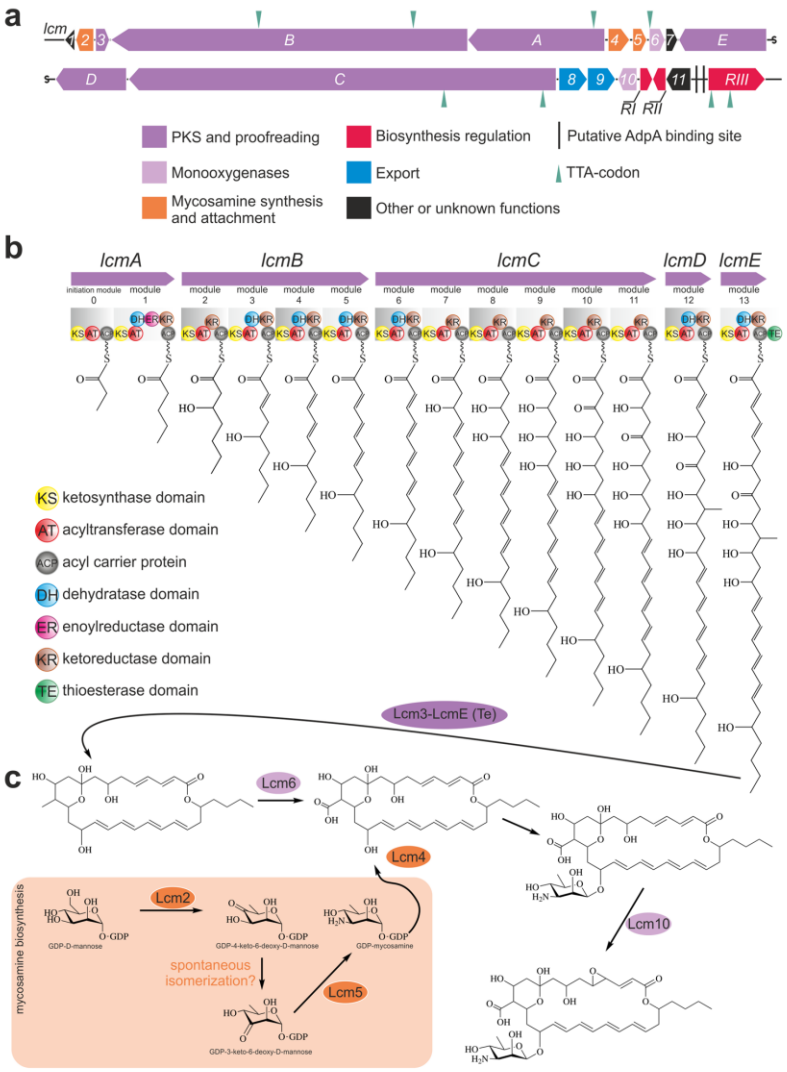


Figure 3. A scheme of Lcm biosynthetic gene cluster (*lcm* BGC) (**a**) and proposed sequence for the Lcm PKS assembly line (**b**) with the tailoring reactions (**c**). For more details, please see the main text and Table 1.

strains *S. cyanogenus* S136 and S136 pGM4181$^+$ cultivated for 72 h in SG medium (optimal for landomycin A production) and YMPG (supports Lcm synthesis). In bioassays, strain S136 showed no signs of Lcm accumulation when grown on either SG or YMPG agar; weak antifungal activity was detected in case of YMPG-grown S136 pGM4181$^+$ (ESM Fig. S7a). Although Lcm production by S136 pGM4181$^+$ in SG was below detection levels (as evident from data in ESM Fig. S7b) the *lcm* genes were transcribed in the latter (Fig. 4a). We quantified the accumulation of Lcm by mass spectrometry in the extracts of different *S. cyanogenus* strains grown in liquid YMPG. The Lcm molecular mass signal was barely detectable in S136 and ΔlanI7. In S136 pGM4181$^+$ and ΔlanI7 pGM4181$^+$ Lcm increased roughly 100-fold (to $(5.2 \pm 0.5) \times 10^5$ au) and 450-fold (to $(1.9 \pm 0.6 \times 10^7$ au), respectively, as compared to their parental strains (Fig. 4b). Hence, in the wild type (S136) AdpA transcriptionally upregulates both Lcm (see above) and landomycin[36] BGCs, yet only traces of Lcm can be detected. When the landomycin biosynthesis pathway is shut off (strain ΔlanI7), AdpA leads to abundant production of Lcm. Substantial quantitative differences in the levels of activation of Lcm production by S136 and ΔlanI7 most likely underlie qualitative differences in the bioassay results.

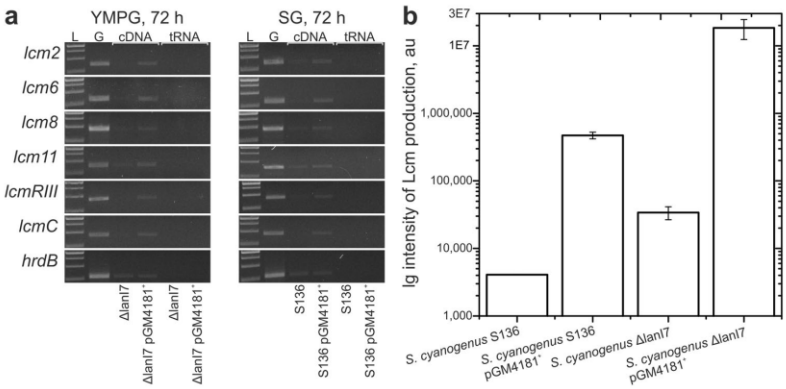


Figure 4. (**a**) Transcription of *lcm* genes involved in different stages of Lcm production (as well as house-keeping gene *hrdB*) in *S. cyanogenus* ΔlanI7 and ΔlanI7 pGM4181$^+$ grown in YMPG medium for 72 h; *S. cyanogenus* S136 and S136 pGM4181$^+$ grown in landomycin production SG medium. Genomic DNA isolated from S136 strain served as a positive control for the PCR-reactions (G). Total RNA samples from ΔlanI7, ΔlanI7 pGM4181$^+$, S136 and S136 pGM4181$^+$ were negative controls of the PCR reactions; (L) was 1 kbp DNA Ladder (Thermo Fisher Scientific, USA). (**b**) Quantification of Lcm production by *S. cyanogenus* S136, S136 pGM4181$^+$, ΔlanI7 and ΔlanI7 pGM4181$^+$ grown in YMPG for 120 h. Areas of Lcm mass peaks (708.35 Da (M + H)$^+$) were integrated and represented as arbitrary units (au); please note logarithmic scale of the *y*-axis. Data represent mean values of three independent experiments ± 2SD.

Mechanism of AdpA-mediated regulation of lucensomycin production. The above findings show that AdpA is required for *lcm* gene expression and Lcm production in *S. cyanogenus*. To identify possible targets of AdpA, we analyzed the *lcm* cluster in silico for putative AdpA operator sites. Two adjacent high-scoring binding sites (TGGCGGAAAC, TGGCCGATAG; $p < 10^{-4}$) were found within the *lcmRIII* promoter region. This, along with data on transcriptional activity of *lcm* genes (see above), allows to suggest that AdpA is likely to activate transcription of *lcm* genes by directly turning on *lcmRIII*. As in the case of pimaricin biosynthesis, LcmRIII should activate transcription of either *lcmRI* or *lcmRII*, or both. Since LcmRI and LcmRII are PimM/NysRIV-like regulators, at present it is difficult to deduce which one is responsible for the activation of *lcm* genes.

To clarify the role of regulatory *lcm* genes, we overexpressed them individually under the control of *ermEp* in *S. cyanogenus* ΔlanI7 and determined Lcm production levels in the resulting strains. Lcm was observed in *S. cyanogenus* ΔlanI7 *lcmRIII*$^+$ and *lcmRII*$^+$, but not in *S. cyanogenus* ΔlanI7 *lcmRI*$^+$ (Fig. 5a, b). We therefore assume that AdpA has no targets in the *lcm* cluster other than *lcmRII* and *lcmRIII*. This assumption remains speculative and requires further investigations on AdpA binding to the *lcmRIII* promoter region. Gene *lcmRI* likely encodes a nonfunctional *lcmRII* paralog.

If low Lcm titers observed in the *adpA*-carrying S136 strain are caused by the concurrent production of landomycins, then a pathway-specific activation of the Lcm BGC should provide a competitive advantage for Lcm production, irrespective of the cultivation medium. Indeed, upon introduction of a *lcmRIII* expression construct into S136 we were able to detect significant antifungal activity of the latter when grown in media that support either landomycin (SG) or Lcm (YMPG) production, as well as on tryptic soy agar unfavorable for either biosynthetic route (Fig. S8a). At the same time, introduction of *lcmRIII* did not impact landomycin production by S136 (Fig. S8b).

Variable termini of AdpA and a rare TTA codon are involved in high-level activation of Lcm biosynthesis. Results presented here and previously[35] show that the isolated DNA binding domain of AdpA is capable of turning on gene expression. Yet, the contribution of different structural parts of AdpA to the efficiency of its function as an "activation tool" was not quantified. We have decided to build and test the allele of *adpA* lacking variable termini. Amino acid sequences of XNR_4181 (AdpA ortholog from *S. albus* J1074; 406 amino acids) and its 45 orthologs from the other *Streptomyces* were aligned, allowing for the identification of the most conserved region spanning from 15 to 350th aa residue of XNR_4181 (Fig. S9). We have cloned the sequence of *XNR_4181* coding the 15-350-aa stretch (XNR_4181i; Fig. S10) into pmoeE5script generating pGM4181i. Next, we cloned synthetic version of *XNR_4181* with TTA → CTC substitution[58] into pmoeE5script generating pGM4181$_{tta}$-. The recombinant *S. cyanogenus* ΔlanI7 strains carrying engineered *adpA* alleles produced Lcm on solid media as judged by the bioassays. Quantitative analysis of liquid cultures showed that removal of variable termini (pGM4181i) led to over tenfold decrease in Lcm production in comparison to ΔlanI7 pGM4181$^+$. The Lcm production level of ΔlanI7 pGM4181$_{tta}$-$^+$ increased fivefold as compared to ΔlanI7 pGM4181i$^+$, yet it still was two times lower than in ΔlanI7 pGM4181$^+$ (Fig. S11).

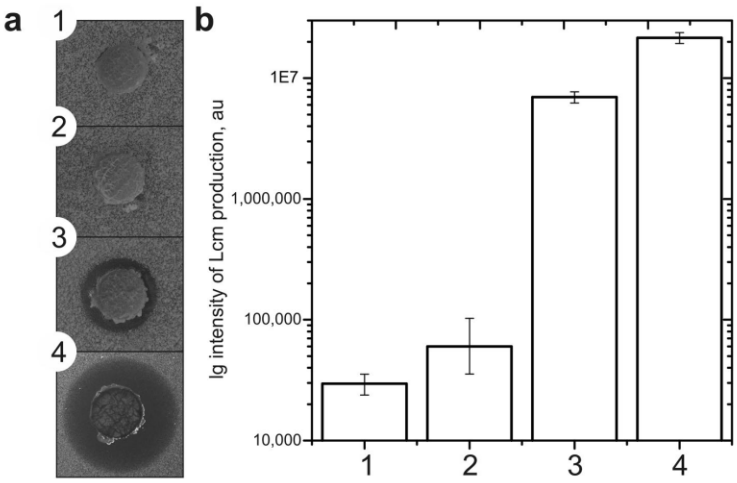


Figure 5. *D. hansenii* growth inhibition assay (**a**) revealed the activation of Lcm production by *S. cyanogenus* ΔlanI7 upon introduction of extra copy of *lcm* cluster-situated transcriptional regulators. Agar plugs were cut from the lawns of *S. cyanogenus* ΔlanI7 (1), ΔlanI7 *lcmRI*$^{+}$ (2), ΔlanI7 *lcmRII*$^{+}$ (3) and ΔlanI7 *lcmRIII*$^{+}$ (4) grown on YMPG agar for 120 h. Lcm production by the same strains in liquid YMPG was quantified via LC–MS (**b**) in the three recombinant strains as compared to the initial one. Data represent mean values of three independent experiments, error bars are mean values ± 2SD.

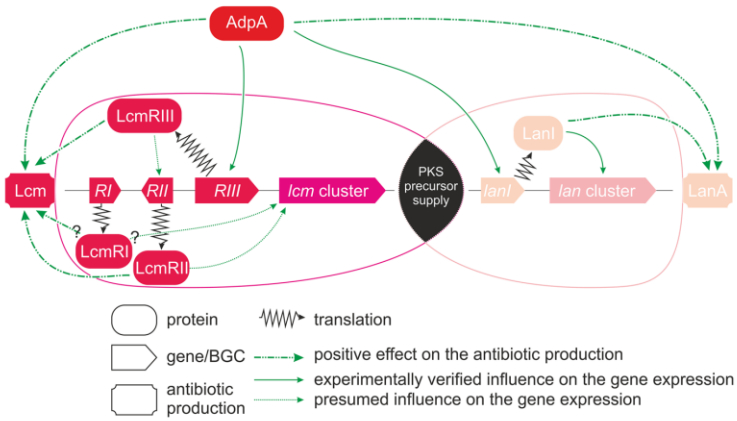


Figure 6. Proposal for the AdpA-mediated regulation of Lcm and the landomycin A biosynthetic pathways in *S. cyanogenus*. Both biosynthetic pathways compete for PKS precursor supply. AdpA, presumably through direct binding to the *lcmRIII* promoter region, activates the expression of this *lcm* cluster-situated transcriptional regulator. Both LcmRII and LcmRIII are positive regulators of Lcm production. The role of LcmRI in the regulation of the Lcm biosynthesis is unclear and awaits further investigations.

Discussion

Over the last two decades of investigations it has become clear that the vast majority of secondary metabolite BGCs within actinobacterial genomes is silent[59,60]. As summarized in the introductory part, the obstructed transcription appears to be the key reason for the BGCs to be silent, and researchers proposed numerous approaches to overcome this obstacle. There is a recent report on the activation of silent BGCs in actinomycetes via knockouts of genes for active secondary metabolic pathways, implying that precocious production of one metabolite diminishes the chances for the other to be produced[61]. Our study, summarized in Fig. 6, describes a more complex case

of successful activation of a silent pathway requiring its transcriptional activation, elimination of the competing active pathway, and appropriate cultivation medium. A few aspects of our findings merit further comments.

Our previous results[36] showed that overexpression of heterologous *adpA* genes in *S. cyanogenus* S136 enhances the transcription of *lan* genes and improves landomycin A production. The above mentioned study and the results reported here support the idea that native AdpA of *S. cyanogenus* is non-functional. The landomycin A biosynthetic route benefits from functional AdpA, but is not dependent on it. In contrast, the Lcm biosynthetic pathway remains silent in the absence of functional *adpA*. RT-PCR analysis shows that in the wild type *S. cyanogenus* S136 (as well as in *S. cyanogenus* Δ*lanI7*) the transcription of *lcm* genes is enhanced upon introduction of functional *adpA*. Yet in case of *S. cyanogenus* S136 (and unlike in Δ*lanI7*) this did not cause Lcm production high enough to suppress fungal growth in the bioassays. This phenomenon might be explained by the competition between landomycin and Lcm biosynthetic pathways for the common pool of acyl-CoA units. If the landomycin biosynthetic route is silenced (as it is the case for mutant Δ*lanI7*), the precursor pool becomes available for the Lcm pathway. Such scenario, however, does not explain the absence of Lcm production in *S. cyanogenus* S136 pGM4181$^+$, where *lcm* genes are expressed at the same level as in Δ*lanI7* pGM4181$^+$ and Lcm is actively produced. The reasons for and mechanisms of such "preferences" for the landomycin biosynthetic route in *S. cyanogenus* at the expense of the other secondary metabolic pathways remain obscure and require further investigations. These findings demonstrate that transcription of genes for secondary metabolism might be inconsequential to the production of secondary metabolites themselves. Altogether our results caution against overt reliance on transcriptome and translatome data as a proxy to *phenotypic* expression of actinobacterial BGCs[13,62,63].

AdpA-activated expression of *lcm* BGC is mediated with the activation of *lcmRIII*, coding for a *lcm* cluster-situated transcriptional regulator. Transcription of this gene appears to be strictly dependent on *adpA* and putative AdpA binding sites were found within its promoter region. We note that AdpA was also shown to be involved in regulation of two other polyene antibiotics, pimaricin and candicidin[64,65]. Nevertheless, the location of AdpA binding sites relative to coding sequences cannot serve as a predictor of its involvement into or a sign of effects on secondary metabolic pathways[66,67]. Thus, more experiments are required to unambiguously prove that *lcmRIII* is a major and direct target of AdpA in context of Lcm biosynthesis. These experiments are underway in our laboratories.

Besides the importance of a DNA-binding domain for AdpA function, little is known about other structural features of AdpA that influence its regulatory capabilities. Here we show that removal of N- and C-terminal variable regions of AdpA affected its ability to upregulate Lcm production. One possibility is that AdpA with shortened termini is not able to form dimers. This inability would not block regulatory properties of AdpA (as the expression of separate DNA-binding domain of AdpA showed), but may lower the affinity of its binding to operator sequences. Also, the shortened AdpA might not be able to efficiently recruit RNA polymerase to form the open complex at the promoter regions. The *adpA* allele featuring synonymous TTA → CTC substitution was also less efficient in the activation of Lcm biosynthesis. The TTA codon is thought to be a part of a regulatory mechanism that limits the level of AdpA protein at the translational level[68]. For a number of TTA-containing regulatory genes of *Streptomyces* the substitution of this codon for a more abundant synonym caused earlier and more abundant production of respective regulatory protein[69]. Nevertheless, the role of the TTA codon in *adpA* remains elusive. AdpA is not produced in mutants devoid of tRNA for the UUA codon[70–72], yet strains expressing TTA$^+$ and TTG$^+$ *adpA* alleles did not differ in temporal patterns of AdpA accumulation[34]. The rather detrimental effect of TTA → CTC substitution on Lcm production suggests that TTA could be more than a switch to control the onset/level of AdpA production. For example, rare codons within interdomain linkers of bacterial mRNAs could slow down protein synthesis of the ribosome and thus optimize co-translational protein folding[73].

Activation of Lcm BGC in *S. cyanogenus* led to the first description of the entire biosynthetic machinery behind the assembly of this natural product. It is similar to other antifungal polyene (e.g. pimaricin, nystatin, rimocidin etc.) BGCs in many respects, making it straightforward to ascribe the function to most *lcm* genes. The organization of initiation module LcmA protein is a distinguishing feature of Lcm PKS. We tentatively suggest that this module is responsible for the generation of the side propyl chain of the lucensomycin aglycone. Further investigations are in progress to verify the proposed mechanism and to understand as to whether LcmA can be used to modify the other polyene antibiotics.

The identity of the other novel mass peaks observed in *adpA*-expressing Δ*lanI7* strain remains unknown. This emphasizes one another formidable challenge in the area of natural product research: signals of novel chemical compounds detected in biological samples are often not pursued if there are no ways for their robust production in reasonable quantities for downstream analysis. Our report stands in a long line of the other "success stories" where a single product was activated in a particular strain (in response to a certain procedure), leaving out of focus the factors that limit the expression of the rest BGCs in that strain. Nevertheless, our work provides a rationale for a new research venue to solve this issue. Particularly, much like the elimination of expression of the landomycin BGC helped us to produce and characterize Lcm, the engineered silencing of Lcm BGC may lead to increased production of the other as-yet-unknown metabolites. We believe that a more systematic investigation of the effects of active and activated BGCs on the silent metabolome of single *Streptomyces* strain is worth the effort. This will yield new insights into complex interactions between the pathways at genetic and metabolic levels, eventually leading to a fuller exploitation of chemical diversity hidden in bacterial genomes.

Methods

Bacterial strains, plasmids and cultivation conditions. Bacterial strains and plasmids, used in this work, are described in ESM Table S1. For intergeneric conjugations, *S. cyanogenus* strains were grown on ISP3 or SFM agar[74] at 30 °C. Characterization of antimicrobial activities of recombinant *S. cyanogenus* Δ*lanI7* derivatives was done on different solid media described in ESM. The same amounts of *S. cyanogenus* mycelia ($(0.8 \div 1) \times 10^7$

c.f.u.) were plated to produce lawns for the bioassays. For submerged production of Lcm, *S. cyanogenus* strains were grown in YMPG medium (please see ESM). Where required, the strains were maintained in presence of apramycin sulphate (50 μg/mL).

Bioassays. Agar plug antibiotic diffusion assay was used to determine antimicrobial activities of *S. cyanogenus* strains grown on agar media listed in ESM. In all cases plates contained 30 mL of solid medium. To assay antifungal activity, agar plugs (Ø 5 mm) were cut from *S. cyanogenus* lawns and placed on the surface of TSB agar plates with freshly seeded 10^6 cells of *D. hansenii*. The antibacterial activity was similarly tested; here plugs were stacked onto modified minimal medium (g/L: KH_2PO_4—3, K_2HPO_4—7, sodium citrate × $4H_2O$—0.5, $MgSO_4 \times 7H_2O$—0.1, $(NH_4)_2SO_4$—1, glucose—2, bacto peptone—0.3, agar—16;[75]) with *B. cereus* ATCC 19637 spores (10^7 per plate); or LB agar with *E. coli* DH5α cells (10^7 per plate). Disc diffusion assay was also employed. Here 5 mm Whatman paper discs were soaked into methanol extracts containing lucensomycin and placed on the surface of TSB agar plates with 10^6 cells of *D. hansenii*. Unless otherwise stated, in all cases the extracts were obtained from the same amount of biomass (400 mg, wet weight). Bioassay plates were incubated at 37 °C (30 °C in the case of *D. hansenii*), halos of growth inhibition around the plugs or discs were analyzed after 20 h of incubation.

Submerged production, purification and LC–MS of Lcm. 250-mL flasks containing 50 mL of TSB were inoculated with agar slices cut from 168 h old *S. cyanogenus* lawns grown on ISP3 agar. The flasks were incubated for 72 h on the rotary shaker at 30 °C, 200 rpm. 1.5 mL of the resulting culture was transferred into 250-mL flasks containing 50 mL of YMPG medium. Main cultures were grown for up to 120 h at 30 °C, 200 rpm. For analytical purposes, 1 g of wet biomass was washed three times with deionized water and extracted with 2 mL of methanol. To determine optimal fermentation time for Lcm production, biomass was collected from the main culture at five time points. Extracts were concentrated to the volume of 100 μL and 40 μL of were subjected to paper disc diffusion assay (see above).

Scaled-up purification of Lcm was done as follows. Four 2-L flasks with steel springs containing 500 mL of YMPG were inoculated with 8 mL of preculture (*S. cyanogenus* ΔlanI7 pGM418$^+$; prepared as described above) and grown for 72 h on the rotary shaker at 30 °C, 200 rpm. Biomass and spent medium were separated with centrifugation. Cells from 2 L of the cultures were washed with water and resuspended in 160 mL of methanol:acetone (1:1). The resulting mixture was sonicated for 15 min, spun down; supernatant was filtered through the 3 mm paper. The spent medium was mixed with ethyl acetate (1:1) by inversion a few times, then the mixture was left to allow for the phases to separate. The organic extracts from cells and supernatant were combined and evaporated *in vacuo*. Dry residue was dissolved in methanol and separated on preparative HPLC column (conditions were the same as for analytical LC–MS, see below). Fractions having an absorption maximum at λ = 310 nm (characteristic for polyene compounds) eluted at 7.4- ÷ 8.4 min. The presence of the 708.35 Da (Lcm) mass peak in this fraction was verified with LC–MS. The collected fraction was evaporated to dryness, giving approximately 100 mg of pale pink solid. Upon addition of 4 mL methanol a white precipitate was formed, which was shown via LC–MS to consist of pure 708.35 Da compound. As a result, we were able to purify 40 mg of Lcm from 2-L flask fermentation. Whenever possible, all plastic- and glassware was wrapped in aluminum foil to eliminate light-induced degradation of Lcm.

Methanol-dissolved Lcm samples were analyzed with HPLC–MS on a HiRes Extractive Oribtrap MassSpectrometer (Thermo Fisher Scientific). The solvent system was water + 0.1% HFo (solvent A); acetonitrile + 0.1% HFo (solvent B); 95% A to 0% A in 8 min, then 1.5 at 0% A, then back to 95% A in 1.5 min. Data analysis was carried out using Xcalibur 5.1 software (Thermo Fisher Scientific).

Plasmid construction. Fragment of the *XNR_4181* coding for DNA-binding domain (700–1245 bp) was amplified using xnr4181araC_up and xnr4181_EcoRIrp primers and cloned into SpeI/EcoRI sites of pmoeE5script[76] yielding pGM4181d. Gene fragments coding for DNA-binding domains of *adpAsco*, *adpAscla* and *adpAgh* were cloned similarly using primers described in Table 2, generating pGMSCOd, pGMSCLAd, pGMSGHd. To construct pGM4181i, fragment of *XNR_4181* coding sequences from 46 to 1019 bp was amplified using iAdpA_F/R primer pair from pGM4181 and cloned into SpeI/EcoRI sites of pmoeE5script. $pGM4181_{tta-}$ was constructed in a similar fashion, using xnr4181_XbaIup/xnr4181_EcoRIrp primer pair to amplify XNR_4181_{tta-} coding sequence from pUCXNR. Pathway-specific regulatory genes *lcmRI, lcmRII* and *lcmRIII* were amplified from the chromosome DNA of *S. cyanogenus* S136 using primers listed in Table 2 and cloned into XbaI/EcoRI sites of pTES[77], yielding pTES22, pTES23 and pTES25 respectively.

Generation of recombinant *Streptomyces* strains. All constructs were transferred into *S. cyanogenus* ΔlanI7 by intergeneric conjugation with *E. coli* ET12567 pUZ8002$^+$, as described elsewhere[74]. *S. cyanogenus* and transconjugants carrying pmoeE5script- and pTES-based plasmids were selected on plates overlaid with 50 μg/mL of apramycin sulphate and 30 μg/mL of nalidixic acid. Apramycin-resistant transconjugants were verified with PCR and primers complementary to marker *aac(3)IV* gene (Table 2).

Total RNA isolation and semi-quantitative RT-PCR analysis. Samples of *S. cyanogenus* mycelium were collected from 5 mL of culture broth by centrifugation (6000 rpm, 5 min, 4 °C). RNA isolation and cDNA synthesis was performed as described previously[36]. PCR was performed using Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific) and primer pairs specific to each individual gene (Table 2). As a positive control for gene expression, the primer pair specific to *S. cyanogenus* house-keeping gene *hrdB* was used. PCR products were analyzed by electrophoresis in 2% TAE agarose gel.

Primer	Nucleotide sequence (5′–3′), recognition sites for restriction endonucleases are underlined in primer sequence	Purpose
iAdpA_F	AAATCTAGACGAGCAACGGAGGTACGGACATGAAACTCTCCGGGCGGCGCC	*XNR_4181i* cloning
iAdpA_R	AAAGAATTCTCACGGACGGCGGGCCCGGTAGG	-//-
xnr4181_XbaIup	AAATCTAGAGGGGGGCTTAGTCACATG	XNR_4181_{tta}- cloning
xnr4181_EcoRIrp	AAAGAATTCGGAGCTGTCCTCTCTCAGAC	-//-
Xnr4181araC_up	AAATCTAGACGAGCAACGGAGGTACGGACATGCCGGAGGAGATCGGCGCCGAC	$adpA_{sco}dbd$/ $adpA_{sgh}dbd$/ $adpA_{scl}dbd$ cloning
SCLAd_F	AAATCTAGACGAGCAACGGAGGTACGGACATGCCAGAGGAAATCGGGTCGGAC	$adpA_{scl}dbd$ cloning
SCLAVADPA_R	AAAGAATTCCATGCGACTACCTTATGG	$adpA_{scl}dbd$ cloning
SCOADPA_R	AAAGAATTCGCCGTCTGCTCACCTCACGG	$adpA_{sco}dbd$ cloning
SGHADPA_R	AAAGATATCGCCTCCGGCCCCGTCCGGTGT	$adpA_{sgh}dbd$ cloning
gbc22_F	TTTTCTAGAGGAGGGTCGGAATGGCGGCGCTGG	*lcmRI* cloning
gbc22_R	TTTGAATTCGGGGTTTCAGCCGGACTCG	-//-
gbc23_F	TTTTCTAGAGGAGGCACACCGTGATTGACGAAAC	*lcmRII* cloning
gbc23_R	TTTGAATTCTCCGCTCCCTGTCGTCCTC	-//-
gbc25_F	TTTTCTAGAGGAGGACCGTCATGCCCGTACCCC	*lcmRIII* cloning
gbc25_R	TTTGAATTCGACCGCGCGAGCGGTGGCTC	-//-
amR_F	CGGGGTCTGACGCTCAGTGGA	*aac(3)IV* amplification
amR_R	AGCGTCTGCTCCGCCATTCG	-//-
rt_hrdB_fp	GCTGGCCAAGGAGCTCGAC	278 bp fragment of *hrdB* (RT-PCR)
rt_hrdB_rp	CGTCGAGGGTCTTCGGCTG	-//-
L2_F	ACGCCGGTCCCGATCACGTA	355 bp fragment of *lcm2* (RT-PCR)
L2_R	TCGGGAAGGTCGCGGAGACT	-//-
L6_F	CGAGTCGCGCACCAAAACGC	312 bp fragment of *lcm6* (RT-PCR)
L6_R	TAGCGCAGCAGCTCGTCCAC	-//-
L8_F	GGTCGGCCGCGATCTGAGGT	446 bp fragment of *lcm8* (RT-PCR)
L8_R	CCCTGCGGGTCAGCACGAAC	-//-
L11_F	CTGCGCTGCCGGATCGTCTT	380 bp fragment of *lcm11* (RT-PCR)
L11_R	TTCGAGGAGGTGCTGCCCCA	-//-
LRIII_F	CGTCACGGATCGCGGGCAC	430 bp fragment of *lcmRIII* (RT-PCR)
LRIII_R	CGTGCTTTCCAGACGTGCGG	-//-
lPKS_F	TGTTGCCCGAGAGCCTCACG	420 bp fragment of *lcmC* (RT-PCR)
lPKS_R	CTACGGCCGCAACCGCCCC	-//-

Table 2. Primers used in this work. Recognition sites of the restriction endonucleases used for the cloning are underlined.

DNA sequencing and in silico analysis. Initially, *S. cyanogenus* genome was sequenced using Illumina approach as described in[36]. However, the initial sequence of *lcm* cluster contained large gaps in the one of the *lcm* PKS genes, known to be challenging target to sequence due to their highly repetitive nature. Therefore, in this work *S. cyanogenus* S136 genome was re-sequenced using a combination of HiSeq Illumina and GridION ONT technologies, essentially as described in[78]. Final S136 genome (8773 kbp) was manually curated and annotated using Prokka annotation pipeline[79]. Full sequence of S136 genome will be published separately.

Annotation of putative BGCs on the chromosome of *S. cyanogenus* was achieved with antiSMASH[14]. Comparison of the *lcm* cluster with already described polyene antibiotic BGCs was made with the assistance of DoBISCUIT database[80]. PKS modules were predicted using PKS/NRPS Analysis website[81]. Search for conserved motifs within the *lcm* cluster was carried out with the help of the MEME suite[82].

Received: 10 November 2020; Accepted: 25 January 2021
Published online: 10 February 2021

References

1. Bérdy, J. Thoughts and facts about antibiotics: Where we are now and where we are heading. *J. Antibiot.* **65**, 385–395. https://doi.org/10.1038/ja.2012.54 (2012).
2. Genilloud, O. Actinomycetes: Still a source of novel antibiotics. *Nat. Prod. Rep.* **34**, 1203–1232. https://doi.org/10.1039/c7np00026j (2017).
3. Diez, J. *et al.* Myxobacteria: Natural pharmaceutical factories. *Microb. Cell. Fact.* **11**, 52. https://doi.org/10.1186/1475-2859-11-52 (2011).
4. Letzel, A.-C., Pidot, S. J. & Hertweck, C. A genomic approach to the cryptic secondary metabolome of the anaerobic world. *Nat. Prod. Rep.* **30**, 392–428. https://doi.org/10.1039/c2np20103h (2013).

5. Wilson, M. C. *et al.* An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* **506**, 58–62. https://doi.org/10.1038/nature12959 (2014).
6. Challinor, V. L. & Bode, H. B. Bioactive natural products from novel microbial sources. *Ann. NY Acad. Sci.* **1354**, 82–97. https://doi.org/10.1111/nyas.12954 (2015).
7. Waksman, S. A. & Woodruff, H. B. The soil as a source of microorganisms antagonistic to disease-producing bacteria. *J. Bacteriol.* **40**, 581–600. https://doi.org/10.1128/JB.40.4.581-600.1940 (1940).
8. Waksman, S. A. & Woodruff, H. B. *Actinomyces antibioticus*, a new soil organism antagonistic to pathogenic and non-pathogenic bacteria 1. *J. Bacteriol.* **42**, 231–249. https://doi.org/10.1128/JB.42.2.231-249.1941 (1941).
9. Lewis, K. Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* **12**, 371–387. https://doi.org/10.1038/nrd3975 (2013).
10. Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147. https://doi.org/10.1038/417141a (2002).
11. Omura, S. *et al.* Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. USA* **98**, 12215–12220. https://doi.org/10.1073/pnas.211433198 (2001).
12. Spohn, M. *et al.* Overproduction of ristomycin A by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. *Antimicrob. Agents. Chemother.* **58**, 6185–6196. https://doi.org/10.1128/AAC.03512-14 (2014).
13. Ahmed, Y., Rebets, Y., Tokovenko, B., Brötz, E. & Luzhetskyy, A. Identification of butenolide regulatory system controlling secondary metabolism in *Streptomyces albus* J1074. *Sci. Rep.* **7**, 9784. https://doi.org/10.1038/s41598-017-10316-y (2017).
14. Weber, T. *et al.* antiSMASH 30-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic. Acids Res.* **43**, W237–W243. https://doi.org/10.1093/nar/gkv437 (2015).
15. Alanjary, M. *et al.* The Antibiotic Resistant Target Seeker (ARTS), an exploration engine for antibiotic cluster prioritization and novel drug target discovery. *Nucleic Acids. Res.* **45**, W42–W48. https://doi.org/10.1093/nar/gkx360 (2017).
16. Hosaka, T. *et al.* Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat. Biotechnol.* **27**, 462–464. https://doi.org/10.1038/nbt.1538 (2009).
17. Tanaka, Y. *et al.* Activation and products of the cryptic secondary metabolite biosynthetic gene clusters by rifampin resistance (*rpoB*) mutations in actinomycetes. *J. Bacteriol.* **195**, 2959–2970. https://doi.org/10.1128/JB.00147-13 (2013).
18. Riedlinger, J. *et al.* Auxofuran, a novel metabolite that stimulates the growth of fly agaric, is produced by the mycorrhiza helper bacterium *Streptomyces* strain AcH 505. *Appl. Environ. Microbiol.* **72**, 3550–3557. https://doi.org/10.1128/AEM.72.5.3550-3557.2006 (2006).
19. Tanaka, Y., Hosaka, T. & Ochi, K. Rare earth elements activate the secondary metabolite-biosynthetic gene clusters in *Streptomyces coelicolor* A3(2). *J. Antibiot. Tokyo* **63**, 477–481. https://doi.org/10.1038/ja.2010.53 (2010).
20. Onaka, H., Mori, Y., Igarashi, Y. & Furumai, T. Mycolic acid-containing bacteria induce natural-product biosynthesis in *Streptomyces* species. *Appl. Environ. Microbiol.* **77**, 400–406. https://doi.org/10.1128/AEM.01337-10 (2011).
21. Onaka, H. Novel antibiotic screening methods to awaken silent or cryptic secondary metabolic pathways in actinomycetes. *J. Antibiot. Tokyo* **70**, 865–870. https://doi.org/10.1038/ja.2017.51 (2017).
22. Okada, B. K. & Seyedsayamdost, M. R. Antibiotic dialogues: induction of silent biosynthetic gene clusters by exogenous small molecules. *FEMS Microbiol. Rev.* **41**, 19–33. https://doi.org/10.1093/femsre/fuw035 (2016).
23. Olano, C. *et al.* Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. *Microb. Biotechnol.* **7**, 242–256. https://doi.org/10.1111/1751-7915.12116 (2014).
24. Zhang, M. M. *et al.* CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.* **13**, 607–609. https://doi.org/10.1038/nchembio.2341 (2017).
25. Takano, E. *et al.* A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol. Microbiol.* **56**, 465–479. https://doi.org/10.1111/j.1365-2958.2005.04543.x (2005).
26. Chen, R. *et al.* Genome mining and activation of a silent PKS/NRPS gene cluster direct the production of totopotensamides. *Org. Lett.* **19**, 5697–5700. https://doi.org/10.1021/acs.orglett.7b02878 (2017).
27. Gao, C., Hindra, A., Mulder, D., Yin, C. & Elliot, M. A. Crp is a global regulator of antibiotic production in *Streptomyces*. *mBio* **3**, e00407-12. https://doi.org/10.1128/mBio.00407-12 (2012).
28. Huang, H. *et al.* Activation of a plasmid-situated type III PKS gene cluster by deletion of a *wbl* gene in deepsea-derived *Streptomyces somaliensis* SCSIO ZH66. *Microb. Cell. Factor.* **15**, 116. https://doi.org/10.1186/s12934-016-0515-6 (2016).
29. Xu, J. *et al.* Activation and mechanism of a cryptic oviedomycin gene cluster via the disruption of a global regulatory gene, adpA, in *Streptomyces ansochromogenes*. *J. Biol. Chem.* **292**, 19708–19720. https://doi.org/10.1074/jbc.M117.809145 (2017).
30. Becerril, A. *et al.* Uncovering production of specialized metabolites by *Streptomyces argillaceus*: activation of cryptic biosynthesis gene clusters using nutritional and genetic approaches. *PLoS ONE* **13**, e0198145. https://doi.org/10.1371/journal.pone.0198145 (2018).
31. Rigali, S. *et al.* Feast or famine: The global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. *EMBO Rep.* **9**, 670–675. https://doi.org/10.1038/embor.2008.83 (2008).
32. Gessner, A. *et al.* Changing biosynthetic profiles by expressing *bldA* in *Streptomyces* strains. *Chembiochem* **16**, 2244–2252. https://doi.org/10.1002/cbic.201500297 (2015).
33. Yamazaki, H., Tomono, A., Ohnishi, Y. & Horinouchi, S. DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. *Mol. Microbiol.* **53**, 555–572. https://doi.org/10.1111/j.1365-2958.2004.04153.x (2004).
34. Wolanski, M. *et al.* The level of AdpA directly affects expression of developmental genes in *Streptomyces coelicolor*. *J. Bacteriol.* **193**, 6358–6365. https://doi.org/10.1128/JB.05734-11 (2011).
35. Rabyk, M., Yushchuk, O., Rokytskyy, I., Anisimova, M. & Ostash, B. Genomic insights into evolution of AdpA family master regulators of morphological differentiation and secondary metabolism in *Streptomyces*. *J. Mol. Evol.* **86**, 204–215. https://doi.org/10.1007/s00239-018-9834-z (2018).
36. Yushchuk, O. *et al.* Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions. *Appl. Microbiol. Biotechnol.* **102**, 8419–8428. https://doi.org/10.1007/s00253-018-9249-1 (2018).
37. Rebets, Y. *et al.* Function of *lanI* in regulation of landomycin A biosynthesis in *Streptomyces cyanogenus* S136 and cross-complementation studies with *Streptomyces* antibiotic regulatory proteins encoding genes. *Arch. Microbiol.* **189**, 111–120. https://doi.org/10.1007/s00203-007-0299-5 (2008).
38. Zhu, D., He, X., Zhou, X. & Deng, Z. Expression of the *melC* operon in several *Streptomyces* strains is positively regulated by AdpA, an AraC family transcriptional regulator involved in morphological development in *Streptomyces coelicolor*. *J. Bacteriol.* **187**, 3180–3187. https://doi.org/10.1128/JB.187.9.3180-3187.2005 (2005).
39. Pralea, I. E. *et al.* From extraction to advanced analytical methods: the challenges of melanin analysis. *Int. J. Mol. Sci.* **20**, 3943. https://doi.org/10.3390/ijms20163943 (2019).
40. Nosanchuk, J. D., Stark, R. E. & Casadevall, A. Fungal melanin: what do we know about structure?. *Front. Microbiol.* **6**, 1463. https://doi.org/10.3389/fmicb.2015.01463 (2015).
41. Gaudiano, G., Bravo, P., Quilico, A., Golding, B. T. & Rickards, R. W. The structure of lucensomycin. Part II. *Tetrahedron. Lett.* **30**, 3567–3571 (1966).

42. Kim, J. D., Kang, J. E. & Kim, B. S. Postharvest disease control efficacy of the polyene macrolide lucensomycin produced by *Streptomyces plumbeus* strain CA5 against gray mold on grapes. *Postharvest. Biol. Technol.* **162**, 111115. https://doi.org/10.1016/j.postharvbio.2019.111115 (2020).
43. Lin, N. *et al.* Structure elucidation of polyene macrolide antibiotic lucensomycin. *Chin. J. Magn. Reson.* **25**, 514–522 (2008).
44. Aparicio, J. F., Colina, A. J., Ceballos, E. & Martín, J. F. The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin. A new polyketide synthase organization encoded by two subclusters separated by functionalization genes. *J. Biol. Chem.* **274**, 10133–10139. https://doi.org/10.1074/jbc.274.15.10133 (1999).
45. Aparicio, J. F., Caffrey, P., Gil, J. A. & Zotchev, S. B. Polyene antibiotic biosynthesis gene clusters. *Appl. Microbiol. Biotechnol.* **61**, 179–188. https://doi.org/10.1007/s00253-002-1183-5 (2003).
46. Martín, J. F. & Aparicio, J. F. Enzymology of the polyenes pimaricin and candicidin biosynthesis. *Methods Enzymol.* **459**, 215–242. https://doi.org/10.1016/S0076-6879(09)04610-2 (2009).
47. Dejong, C. A. *et al.* Polyketide and nonribosomal peptide retro-biosynthesis and global gene cluster matching. *Nat. Chem. Biol.* **12**, 1007–1014. https://doi.org/10.1038/nchembio.2188 (2016).
48. Fischbach, M. A. & Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496. https://doi.org/10.1021/cr0503097 (2006).
49. Zhou, Y. *et al.* Selective removal of aberrant extender units by a type II thioesterase for efficient FR-008/candicidin biosynthesis in *Streptomyces* sp. strain FR-008. *Appl. Environ. Microbiol.* **74**, 7235–42. https://doi.org/10.1128/AEM.01012-08 (2008).
50. Mendes, M. V., Antón, N., Martín, J. F. & Aparicio, J. F. Characterization of the polyene macrolide P450 epoxidase from *Streptomyces natalensis* that converts de-epoxypimaricin into pimaricin. *Biochem. J.* **386**, 57–62. https://doi.org/10.1042/BJ20040490 (2005).
51. Antón, N., Mendes, M. V., Martín, J. F. & Aparicio, J. F. Identification of PimR as a positive regulator of pimaricin biosynthesis in *Streptomyces natalensis*. *J. Bacteriol.* **186**, 2567–2575. https://doi.org/10.1128/JB.186.9.2567-2575.2004 (2004).
52. Antón, N. *et al.* PimM, a PAS domain positive regulator of pimaricin biosynthesis in *Streptomyces natalensis*. *Microbiology* **53**, 3174–3183. https://doi.org/10.1099/mic.0.2007/009126-0 (2007).
53. Sekurova, O. N. *et al. In vivo* analysis of the regulatory genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455 reveals their differential control over antibiotic biosynthesis. *J. Bacteriol.* **186**, 1345–1354. https://doi.org/10.1128/jb.186.5.1345-1354.2004 (2004).
54. Mendez, C. & Salas, J. A. The role of ABC transporters in antibiotic producing organisms: Drug secretion and resistance mechanisms. *Res. Microbiol.* **152**, 341–350. https://doi.org/10.1016/S0923-2508(01)01205-0 (2001).
55. de Kruijff, B. & Demel, R. A. Polyene antibiotic-sterol interactions in membranes of Acholeplasma laidlawii cells and lecithin liposomes. 3 .Molecular structure of the polyene antibiotic-cholesterol complexes. *Biochim. Biophys. Acta.* **339**, 57–70. https://doi.org/10.1016/0005-2736(74)90332-0 (1974).
56. Hammond, S. M. Biological activity of polyene antibiotics. *Prog. Med. Chem.* **14**, 105–179. https://doi.org/10.1016/s0079-6468(08)70148-6 (1977).
57. Mendes, M. V. *et al.* Cholesterol oxidases act as signaling proteins for the biosynthesis of the polyene macrolide pimaricin. *Chem. Biol.* **14**, 279–290. https://doi.org/10.1016/j.chembiol.2007.01.010 (2007).
58. Koshla, O. *et al.* Gene *miaA* for post-transcriptional modification of $tRNA_{XXA}$ is important for morphological and metabolic differentiation in *Streptomyces*. *Mol. Microbiol.* **112**, 249–265. https://doi.org/10.1111/mmi.14266 (2019).
59. Ward, A. C. & Allenby, N. E. Genome mining for the search and discovery of bioactive compounds: The *Streptomyces* paradigm. *FEMS Microbiol. Lett.* **365**, fny240. https://doi.org/10.1093/femsle/fny240 (2018).
60. Hoskisson, P. A. & Seipke, R. F. Cryptic or silent? The known unknowns, unknown knowns, and unknown unknowns of secondary metabolism. *mBio* **11**, e02642-20. https://doi.org/10.1128/mBio.02642-20 (2020).
61. Culp, E. J. *et al.* Hidden antibiotics in actinomycetes can be identified by inactivation of gene clusters for common antibiotics. *Nat. Biotechnol.* **37**, 1149–1154. https://doi.org/10.1038/s41587-019-0241-9 (2019).
62. Jeong, Y. *et al.* The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2). *Nat. Commun.* **7**, 11605. https://doi.org/10.1038/ncomms11605 (2016).
63. Bucca, G. *et al.* Translational control plays an important role in the adaptive heat-shock response of *Streptomyces coelicolor*. *Nucleic. Acids Res.* **46**, 5692–5703. https://doi.org/10.1093/nar/gky335 (2018).
64. Du, Y. L. *et al.* The pleitropic regulator AdpAch is required for natamycin biosynthesis and morphological differentiation in *Streptomyces chattanoogensis*. *Microbiology* **157**, 1300–1311. https://doi.org/10.1099/mic.0.046607-0 (2011).
65. Liu, X. *et al.* Dynamic changes of metabolomics and expression of candicidin biosynthesis gene cluster caused by the presence of a pleiotropic regulator AdpA in *Streptomyces* ZYJ-6. *Bioprocess Biosyst. Eng.* **42**, 1353–1365. https://doi.org/10.1007/s00449-019-02135-4 (2019).
66. Pan, Y., Liu, G., Yang, H., Tian, Y. & Tan, H. The pleiotropic regulator AdpA-L directly controls the pathway-specific activator of nikkomycin biosynthesis in *Streptomyces ansochromogenes*. *Mol. Microbiol.* **72**, 710–723. https://doi.org/10.1111/j.1365-2958.2009.06681.x (2009).
67. Higo, A., Hara, H., Horinouchi, S. & Ohnishi, Y. Genome-wide distribution of AdpA, a global regulator for secondary metabolism and morphological differentiation in *Streptomyces*, revealed the extent and complexity of the AdpA regulatory network. *DNA Res.* **19**, 259–273. https://doi.org/10.1093/dnares/dss010 (2012).
68. Takano, E. *et al.* A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. *Mol. Microbiol.* **50**, 475–486. https://doi.org/10.1046/j.1365-2958.2003.03728.x (2003).
69. Rebets, Y. V., Ostash, B. O., Fukuhara, M., Nakamura, T. & Fedorenko, V. O. Expression of the regulatory protein LndI for landomycin E production in *Streptomyces globisporus* 1912 is controlled by the availability of tRNA for the rare UUA codon. *FEMS Microbiol. Lett.* **256**, 30–37. https://doi.org/10.1111/j.1574-6968.2005.00087.x (2006).
70. Higo, A., Horinouchi, S. & Ohnishi, Y. Strict regulation of morphological differentiation and secondary metabolism by a positive feedback loop between two global regulators AdpA and BldA in *Streptomyces griseus*. *Mol. Microbiol.* **81**, 1607–1622. https://doi.org/10.1111/j.1365-2958.2011.07795.x (2011).
71. Nguyen, K. T. *et al.* Colonial differentiation in *Streptomyces coelicolor* depends on translation of a specific codon within the *adpA* gene. *J. Bacteriol.* **185**, 7291–7296. https://doi.org/10.1128/jb.185.24.7291-7296.2003 (2003).
72. López-García, M. T., Santamarta, I. & Liras, P. Morphological differentiation and clavulanic acid formation are affected in a *Streptomyces clavuligerus adpA*-deleted mutant. *Microbiology* **156**, 2354–2365. https://doi.org/10.1099/mic.0.035956-0 (2010).
73. Hanson, G. & Coller, J. Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* **19**, 20–30. https://doi.org/10.1038/nrm.2017.91 (2018).
74. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. *Practical Streptomyces genetics genetics* (Norwich, John Innes Foundation, 2000).
75. Anagnostopoulos, C. & Spizizen, J. Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* **81**, 741–746. https://doi.org/10.1128/JB.81.5.741-746.1961 (1961).
76. Makitrynskyy, R. *et al.* Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. *Open Biol.* **3**, 130121. https://doi.org/10.1098/rsob.130121 (2013).
77. Herrmann, S. *et al.* Site-specific recombination strategies for engineering actinomycete genomes. *Appl. Environ. Microbiol.* **78**, 1804–1812. https://doi.org/10.1128/AEM.06054-11 (2012).

78. Bekiesch, P. *et al.* Viennamycins: lipopeptides produced by a *Streptomyces* sp. *J. Nat. Prod.* **83**, 2381–2389. https://doi.org/10.1021/acs.jnatprod.0c00152 (2020).
79. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153 (2014).
80. Ichikawa, N. *et al.* DoBISCUIT: A database of secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res.* **41**, D408–D414. https://doi.org/10.1093/nar/gks1177 (2013).
81. Bachmann, B. O. & Ravel, J. Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Methods Enzymol.* **458**, 181–217. https://doi.org/10.1016/S0076-6879(09)04808-3 (2009).
82. Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208. https://doi.org/10.1093/nar/gkp335 (2009).

Acknowledgements
This work was supported by grants Bg-80F, M/26-2018 (to B.O.) and BG-09F (to V.F.), all from Ministry of Education and Science of Ukraine. B.O. was supported by DAAD fellowship. R.D.S. and B.O. thank the Bundesministerium für Bildung und Forschung (BMBF) for funding the project 021DK17013 AWAKEDRUGS. B.O. thanks staff of the Collection of microorganisms-producers of antibiotics at Lviv University for the provision of fungal strains.

Author contributions
O.Y., J.K., V.F., R.D.S. and B.O. conceived the experiments, O.Y., I.O., I.V., E.M. and M.D. conducted the experiments, O.Y., B.O. and R.D.S. wrote the manuscript, all authors reviewed and approved the manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-82934-6.

Correspondence and requests for materials should be addressed to R.D.S. or B.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021

Оригінал доступний за посиланням:
https://www.nature.com/articles/s41598-021-82934-6.
Електронні додатки до статті доступні за посиланням:
https://static-content.springer.com/esm/art%3A10.1038%2Fs41598-021-82934-6/MediaObjects/41598_2021_82934_MOESM1_ESM.pdf.

Анотація: *Актинобактерії є одними з найбільш багатих джерел важливих для медицини та сільського господарства сполук, отриманих з їхніх кластерів біосинтетичних генів (КБГ) для спеціалізованих (вторинних) шляхів метаболізму. Геноміка засвідчує, що більшість актинобактеріальних КБГ є мовчазними, швидше за все, через їх низьку або нульову транскрипцію. Багато зусиль докладено для пошуку підходів до активації мовчазних КБГ, оскільки вважається, що це пожвавить відкриття нових природних продуктів. Ми припустили, що глобальний транскрипційний фактор AdpA, завдяки його сильно виродженій операторній послідовності, може бути використаний для підвищення експресії мовчазних КБГ. Використовуючи Streptomyces cyanogenus S136 як взірець, ми показали, що плазміди, які експресують або повнорозмірний adpA, або його ДНК-зв'язуючий домен, ведуть до значних змін у вторинному метаболомі. Це було очевидно із змін в накопиченні забарвлених сполук, біоактивності, а також у появі нових вторинних метаболітів, які виявлено за допомогою ВЕРХ-МС. Далі ми зосередилися на найбільш явному вторинному метаболіті та ідентифікували його як полієновий антибіотик люцензоміцин. Нарешті, ми анотували повний кластер генів для біосинтезу люцензоміцину (lcm), який залишався невловимим протягом п'яти десятиліть, і окреслили заснований на доказах сценарій його adpA-опосередкованої активації.*

УКРАЇНА

(19) UA (11) 119472 (13) C2

(51) МПК
C12P 1/06 (2006.01)
C12P 19/64 (2006.01)
C12N 15/80 (2006.01)
C12R 1/465 (2006.01)

МІНІСТЕРСТВО ЕКОНОМІЧНОГО РОЗВИТКУ І ТОРГІВЛІ УКРАЇНИ

(12) ОПИС ДО ПАТЕНТУ НА ВИНАХІД

(21) Номер заявки: **a 2017 02642**
(22) Дата подання заявки: **21.03.2017**
(24) Дата, з якої є чинними права на винахід: **25.06.2019**
(41) Публікація відомостей про заявку: **25.09.2017, Бюл.№ 18**
(46) Публікація відомостей про видачу патенту: **25.06.2019, Бюл.№ 12**

(72) Винахідник(и):
Ющук Олександр Сергійович (UA),
Осташ Богдан Омелянович (UA)

(73) Власник(и):
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ФРАНКА,
вул. Університетська, 1, м. Львів, 79000 (UA)

(56) Перелік документів, взятих до уваги експертизою:
UA 94608 U, 25.11.2014
UA 79968 U, 13.05.2013
MAKITRYNSKYY, R. et. al. Pleiotropic regulatory genes bldA, adpA and absB are implicated in production of phosphoglycolipid antibiotic moenomycin. Open biology. 2013, V. 3, № 10, P. 1-13
KLYMYSHIN, D. et. al. Expression of pleiotropic regulatory genes in streptomycetes. Abstracts of the 8th Parnas Conference, Warsaw, Poland, August 27-31, 2011, C. 18
ТОДОСІЙЧУК, Т. С. и др. Сучасний стан і перспективи біотехнологічного виробництва антибіотиків. Біологічні Студії/Studia Biologica. 2011, Т. 5, № 1, С. 159-172
BRAUTASET T. et. al. Positively regulated bacterial expression systems. Microbial biotechnology. 2009, V. 2, № 1, P. 15-30
AraC-family transcriptional regulator with DNA-binding transcription factor activity from Streptomyces albus J1074, July 13, 2010 [online] < https://www.uniprot.org/uniprot/D6B298 >
Streptomyces albidoflavus: XNR_4181, July 13, 2010 [online] < https://www.genome.jp/dbget-bin/www_bget?salb:XNR_4181>
OSTASH, B. et. al. The adpA-like regulatory gene from Actinoplanes teichomyceticus: in silico analysis and heterologous expression. World Journal of Microbiology and Biotechnology. 2015, V. 31, № 8, P. 1297-1301

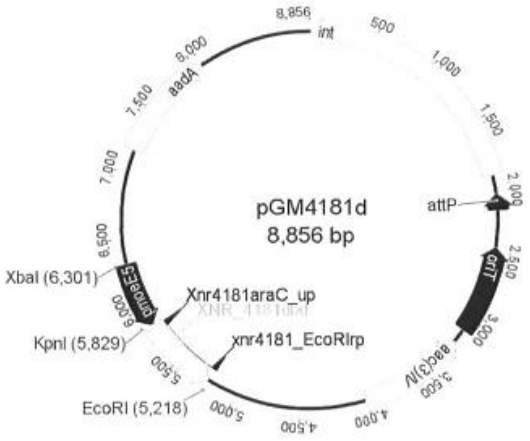
UA 119472 C2

(54) СПОСІБ АКТИВУВАННЯ ПРОДУКЦІЇ СПОЛУК З АНТИБІОТИЧНОЮ ТА ПРОТИГРИБНОЮ ВЛАСТИВОСТЯМИ У STREPTOMYCES CYANOGENUS S136

(57) Реферат:

Винахід стосується генетики бактерій та біотехнології і може бути використаний для одержання сполук з антибіотичною (ландоміцин А) та протигрибною властивостями трансформованими бактеріями виду Streptomyces cyanogenus. Спосіб активування продукції сполук з антибіотичною та протигрибною властивостями у Streptomyces cyanogenus S136 ґрунтується на уведенні додаткових копій регуляторних елементів, причому як регуляторний елемент використовують ділянку гена плейотропного регулятора транскрипції XNR_4181 зі Streptomyces

albus J1074, клоновану під контролем промотору moeE5p у складі інтегративної плазміди pGM4181d, при цьому ця ділянка гена кодує ДНК-зв'язувальний домен на С-кінці білка AdpA. Використання способу дасть змогу розширити спектр сполук, які продукуються бактеріями виду Streptomyces cyanogenus S136.



Фіг. 2

Винахід стосується генетики бактерій та біотехнології і може бути використаний для одержання сполук з антибіотичною (ландоміцину А) та протигрибною активностями трасформованими бактеріями виду Streptomyces cyanogenus.

Відомий спосіб, за яким суспензії клітин чи спор бактерій порядку Streptomycetales (стрептоміцетів) або обробляють, або не обробляють хімічним чи фізичним мутагеном та висівають на агаризоване середовище із додаванням перелічених нижче антибіотиків: стрептоміцину, рифампіцину, тіострептону, або висівають на такі середовища. У вибірці колоній, що виростають на цих середовищах, за способами біологічних тестів та аналітичної хімії виявляють продукцію нових сполук [Hosaka T., Ohnishi-Kameyama M., Muramatsu H., Murakami K., Tsurumi Y., Kodani S., Yoshida M., Fujie A., Ochi K. Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12 // Nat. Biotechnol. - 2009. - Vol. 27. - P. 462-464].

Однак, цей спосіб вимагає великої кількості розхідних матеріалів на етапі вирощування та скринінгу колоній. Він тривалий і вимагає індивідуального підбору умов мутагенезу. Первинний відбір колоній часто складно раціоналізувати, якщо сполука, синтез якої активують, не має забарвлення або сильної антибактерійної дії. Отримані мутанти часто нестабільні і потребують підтримувальної селекції.

Відомий спосіб, за яким проводять повне секвенування геномів стрептоміцетів, комп'ютерну анотацію усіх генів вторинного метаболізму та надекспресію окремих генів шлях-специфічних регуляторів, які можуть активувати транскрипцію відповідних структурних генів біосинтезу антибіотиків. Нові речовини виявляють за способами біотестів та аналітичної хімії [Laureti L., Song L., Huang S., Corre C., Leblond P., Challis G.L., Aigle B. Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens II Proc. Natl. Acad. Sci. USA-2011. - Vol. 108. P. 6258-6263].

Однак, цей спосіб економічно невиправданий у великих масштабах, враховуючи тривалість і вартість усіх етапів конструювання рекомбінантних молекул ДНК для надекспресії генів. Автоматична анотація геномів не завжди дає змогу ідентифікувати необхідні шлях-специфічні регуляторні гени, що збільшує ризик негативних результатів. Крім того, цей підхід не спрацює для тих кластерів генів вторинного метаболізму, що не кодують шлях-специфічні регулятори.

Відомий спосіб, за яким здійснюють секвенування геномів актиноміцетів, ідентифікацію кластерів генів вторинного метаболізму і видалення (нокаут) генів репресорного типу, що призводить до синтезу нових сполук [Metsä-Ketelä M., Ylihonko K., Mäntsälä P. Partial activation of a silent angucycline-type gene cluster from a rubromycin beta producing Streptomyces sp. PGA64 // J. Antibiot. - 2004. - Vol. 57. - P. 502-510].

Проте, недоліками цього способу є неточність передбачення генів репресорного типу, а також необхідність розробки способів генетичних маніпуляцій з геномами досліджуваних штамів, що не завжди можна досягнути для певних видів стрептоміцетів.

Відомий спосіб, за яким проводять часткове секвенування геномів стрептоміцетів, біоінформатичне передбачення усіх кластерів генів вторинного метаболізму; вирощування стрептоміцетів, що, згідно даних біоінформатики, містять нові кластери генів, у великому наборі поживних середовищ; вичерпний аналіз спектра синтезованих сполук за різних умов культивування. Сполуки ідентифікують за способами біологічних тестів і порівняння мас-спектрів з комерційно доступними хімічними базами даних [McAlpine J.B., Bachmann B.O., Piraee M., Tremblay S., Alarco A.M., Zazopoulos E., Farnet C.M. Microbial genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel antifungal agent, as an example // J. Nat. Prod. -2005.-Vol. 68. - P. 493-496].

Однак, цей підхід ресурсо- і трудомісткий, вимагає складної апаратури і платного доступу до комерційних баз даних.

Відомий спосіб, який ґрунтується на надекспресії ієна фосфатази absA1, що походить зі Streptomyces coelicolor, у клітинах актиноміцетів [McKenzie N.L., Thaker M., Koteva K., Hughes D. W., Wright G.D., Nodwell J.R. Induction of antimicrobial activities in heterologous streptomycetes using alleles of the Streptomyces coelicolor gene absAl // J. Antibiot. - 2010. - Vol. 63. - P. 177-182]. У геноми стрептоміцетів інтегрують ген absAl під контролем сильного конститутивного промотора гена стійкості до еритроміцину ermEp. Порівнюють спектри синтезованих вторинних метаболітів вихідним і рекомбінантним штамами за допомогою біотестів і мас-спектрометрії. Відбирають ті штами, що мають змінений спектр синтезу, очищують нові сполуки і встановлюють їхню структуру за допомогою загальновживаних способів аналітичної хімії.

Проте, цей підхід є вузьким, бо базується на дефосфорилуванні репресора AbsA2 фосфатазою AbsAl, а AbsA2 у вихідному стані має невелике коло операторних ділянок. Тому і

коло генів вторинного метаболізму, який він контролює, не є широким. Неможливо індукувати синтез тих нових сполук, які не підлягають репресії з боку AbsA2.

Відомий спосіб, за яким виділяють з ґрунту чи іншого екотопу високомолекулярну геномну ДНК, її клонують у бактерійні штучні хромосоми (БШХ) та експресію БШХ в "універсальному" актиноміцетному штамі-хазяїні [Iqbal H., Low-Beinart L., Obiajulu J., Brady S. Natural product discovery through improved functional metagenomics in Streptomyces // J. Am. Chem. Soc. -2016. - Vol. 138, № 30. - P. 9341-9344]. Продукцію нових сполук виявляють за зміною забарвлення, способами біологічних тестів та мас-спектрометрії.

Однак, спосіб є трудомістким і тривалим, оскільки вимагає постійного конструювання рекомбінантних ДНК in vitro. На сьогодні невідомо, чи буде успішно експресуватися метагеномна ДНК різного походження в одному гетерологічному хазяїні з роду Streptomyces.

Найближчим за технічною суттю - прототипом - є спосіб, що ґрунтується на надекспресії в клітинах стрептоміцетів гена adpA19, що кодує гомолога плейотропного транскрипційного фактора AdpA з рідкісної актинобактерії Actinoplanes teichomyceticus [Ostash B., Yushchuk O., Tistechok S., Mutenko H., Horbal L., Muryn A., Dacyuk Y., Kalinowski J., Luzhetskyy A., Fedorenko V. The adpA-like regulatory gene from Actinoplanes teichomyceticus: in silico analysis and heterologous expression // World J. Microbiol. Biotechnol. - 2015. - Vol. 31, № 8. - P. 1297-1301]. Ген adpA19 клонують у реплікативний вектор pKC 1139 і переносять у різні штами актиноміцетів за допомогою міжродової кон'югації [Flett F., Mersinias V., Smith C.P. High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl-DNA-restricting Streptomycetes // FEMS Microbiol. Lett-1997. vol.155. P.223-229]. Рекомбінантні й батьківські штами порівнюють за забарвленням та рівнем антибіотичної активності, і так виявляють ті, що продукують нові антибіотики, або надсинтезують відомі сполуки.

Проте, у цьому підході ген adpA19 знаходиться у складі реплікативного вектора і тому він втрачається у штамах, якщо їх вирощувати за відсутності маркерного антибіотика апраміцину. Продукт трансляції гена adpA19 містить глутамін-амідотрансферазний домен, що може обмежувати здатність білка AdpA19 активувати транскрипцію генів вторинного метаболізму.

В основу винаходу поставлено задачу удосконалити спосіб активування продукції сполук з антибіотичною та протигрибною властивостями у Streptomyces cyanogenus S136 шляхом уведення додаткових копій С-кінцевої ділянки гена XNR_4181, що кодує ДНК-зв'язувальний домен білка AdpA Streptomyces albus J1074, на основі інтегративної плазміди pGM4181d, що дасть змогу розширити спектр сполук, які продукуються бактеріями виду Streptomyces cyanogenus S136.

Поставлена задача вирішується тим, що у способі активування продукції сполук з антибіотичною та протигрибною властивостями у Streptomyces cyanogenus S136, який ґрунтується на уведенні додаткових копій регуляторних елементів, при чому як регуляторний елемент використовують ділянку гена плейотропного регулятора транскрипції XNR_4181 зі Streptomyces albus J1074, клоновану під контролем промотора moeE5p у складі інтегративної плазміди pGM4181d, при цьому ця ділянка гена кодує ДНК-зв'язувальний домен на С-кінці білка AdpA.

Більшість відомих протимікробних та протигрибкових препаратів сьогодні є неефективними проти збудників інфекційних захворювань. 21 вересня 2016 193 країни 0011 підписали документ, в якому зобов'язалися позбавити світ від так званих "супербактерій" штамів, стійких до більшості чи навіть усіх наявних у медичному арсеналі антибіотиків і створити нові препарати, що володіють новими механізмами дії та модифікувати вже існуючі. Стрептоміцети - група бактерій, що є джерелом багатьох медично важливих антибіотиків, які виявляли переважно за допомогою традиційних способів мікробіологічного та/або хімічного скринінгів. Аналіз сотень геномів стрептоміцетів засвідчив, що їхній потенціал щодо продукції біоактивних сполук значно вищий, ніж це показує мікробіологічний скринінг. Так, стрептоміцетний геном розміром 6-13 млн. п.н. може містити 20-48 кластерів генів біосинтезу вторинних метаболітів, з яких лише 10-20 % мають вияв на рівні продукції відповідної сполуки. Кластери генів вторинного метаболізму, що не експресуються або експресуються дуже слабо отримали назву криптичних, або мовчазних. Активування експресії мовчазних генів може призвести до ідентифікації нових класів природних сполук, які досі не вдалося виявити традиційними способами.

Авторами вперше запропоновано використати плазміду pGM4181d, яка містить як регуляторний елемент ділянку гена плейотропного регулятора транскрипції XNR_4181 зі Streptomyces albus J1074, яка кодує ДНК-зв'язувальний домен на С-кінці білка AdpA, та промотор moeE5p для розширення спектра сполук з антибіотичною та протигрибною активностями, які продукуються Streptomyces cyanogenus S136. Запропонований підхід має низку переваг порівняно із відомими способами активування кластерів генів вторинного

метаболізму. Біоінформатичний аналіз продуктів трансляції генів родини AdpA показує, що всі вони містять два домени: ДНК-зв'язувальний (HTH) на С-кінці та глутамін-амідотрансферазний (GAT) на N-кінці [Chater K.F., Chandra G. The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in Streptomyces // J. Microbiol. - 2008. - Vol. 46, № 1. - P. 1-11]. Вважають, що N-кінцевий GAT домен забезпечує димеризацію білка AdpA, і саме як димер AdpA виконує свою типові фізіологічні функції [Yamazaki H., Tomono A., Ohnishi Y., Horinouchi S. DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in Streptomyces griseus // Mol. Microbiol. - 2004. - Vol. 53, № 2. - P. 555-572]. Унікальність AdpA полягає у здатності розпізнавати дегенеровану декануклеотидну послідовність 5-TGGCSNGWWY-3', що, згідно розрахунків, зустрічається кілька десятків тисяч разів у хромосомі модельного об'єкта, продуцента стрептоміцину Streptomyces griseus. Експериментальні дані свідчать, що AdpA контролює експресію більше 500 генів - це найбільший регулон із описаних наразі в біології [Higo A., Hara H., Horinouchi S., Ohnishi Y. Genome-wide distribution of AdpA, a global regulator for secondary metabolism and morphological differentiation in Streptomyces, revealed the extent and complexity of the AdpA regulatory network // DNA Res. - 2012. - Vol. 19. - P. 259-273]. Серед AdpA-залежного регулону виявлено велику кількість генів вторинного метаболізму, і в абсолютній більшості описаних випадків AdpA є позитивним регулятором цих генів [Makitrinskyy R., Ostash B., Tsypik O., Rebets Y., Doud E., Meredith T., Luzhetskyy A., Bechthold A., Walker S., Fedorenko V. Pleiotropic regulatory genes bldA, adpA and absB are implicated in production of phosphoglycolipid antibiotic moenomycin // Open Biol. - 2013. - Vol. 3. - P. 1-13]. Тому транскрипційні фактори родини AdpA - найперспективніші для активації мовчазних генів, оскільки мають велику кількість операторних послідовностей і їхнє використання базуватиметься на надекспресії, а не на технічно складніших генних заміщеннях. Однак, природний білок AdpA, як знаряддя активації мовчазних генів, має два недоліки. По-перше, присутність GAT-домену може суттєво зменшувати коло операторных ділянок-мішеней, що розпізнаються AdpA. По-друге, лінкерний район, що з'єднує GAT- та HTH-домени, містить рідкісний лейциновий кодон TTA, що гальмує, а за деяких умов блокує накопичення білка AdpA, що буде зменшувати рівень синтезу цільових продуктів - біоактивних природних речовин [Takano E., Tao M., Long F., Bibb M.J., Wang L., Li W., Buttner M.J., Bibb M.J., Deng Z.X., Chater K.F. A rare leucine codon in adpA is implicated in the morphological defect of bldA mutants of Streptomyces coelicolor // Mol. Microbiol. - 2003. - Vol. 50, № 2ю - P. 475-486]. Вперше запропоновано використати для активування генів вторинного метаболізму актиноміцетів не цілий ген adpA, а лише С-кінцеву частину цього гена, що кодує ДНК-зв'язувальний домен (HTH-AdpA). Це дасть змогу уникнути неефективної трансляції білка, пов'язаної із рідкісним кодоном TTA у лінкері, а також його облігатної димеризації. Також для експресії запропоновано використати сильний промотор moeE5p, що походить з кластера генів біосинтезу моеноміцину [Ostash B., Saghatelian A., Walker S.A streamlined metabolic pathway for the biosynthesis of moenomycin A // Chem. Biol. - 2007. - Vol. 14, № 3. - P. 257-267], та інтегративну плазміду на основі системи int-attP актинофага φC31. Використання промотора moeE5p дасть змогу синхронізувати експресію HTH-AdpA з початком вторинного метаболізму; рекомбінантна конструкція на основі актинофага буде стабільно успадковуватися, без додавання у ростові середовища селективного антибіотика.

Фіг. 1 Схема гена XNR 4181 та доменна організація білка XNR_4181 (AdpA), де:

697-1221 п.н. - ділянка гена XNR_4181, що кодує ДНК-зв'язувальний домен; GATase-1 (25-210 а.з.) - GAT-домен, що ймовірно бере участь у димеризації білка XNR_4181;

LR (210-242 а.з.) лінкерна ділянка, що з'єднує GAT-домен та ДНК-зв'язувальний домен білка XNR_4181;

HTH (242-281 а.з.) ДНК-зв'язувальний домен XNR_4181 типу "петля-поворот-петля".

Фіг. 2 Схема плазміди pGM4181d, де:

aac(3)IV - ген аміноглікозид ацетилтрансферази, що забезпечує стійкість до апраміцину;

aadA - ген аміноглікозид аденілілтрансферази, що забезпечує стійкість до спектиноміцину та стрептоміцину;

attP - ділянка att фага φC31;

int - ген інтегрази фага φC31;

oriT- ділянка кон'югаційного переносу плазміди RK2;

moeE5p промоторна ділянка гена гексозоізмерази moeE5 з кластера біосинтезу моеноміцину Streptomyces ghanaensis;

XNR_4181dbd - ділянка гена XNR_4181, що кодує ДНК-зв'язувальний домен, клонована з хромосоми Streptomyces albus J1074;

xnr4181 araC_up та xnr4181_EcoRIrp - ділянки, що відповідають праймерам, використаним для клонування XNR_4181dbd.

Фіг. 3. Виявлення продукції антибіотиків способом тесту на пригнічення росту B. cereus ATCC 19637, де 1 - агаровий блок, вирізаний із газону S. cyanogenus SI36; 2 агаровий блок, вирізаний із газону S. cyanogenus pGM4181d.

Фіг. 4. Результат ТШХ, що ілюструє появу ландоміцину А в екстракті з газону S. cyanogenus pGM4181d (2) та його відсутність в екстракті з газону S. cyanogenus S136 (1); 3 хроматографічний свідок чистий препарат ландоміцину А.

Фіг. 5. Тест на пригнічення росту B. cereus ATCC 19637, де 1 - агаровий блок, вирізаний із газону S. cyanogenns AlanI; 2 - блок, вирізаний із газону S. cyanogenus AlanI pGM4181 d.

Фіг. 6. Тест на пригнічення росту Debaryomyces hansenii BKM Y-9, де 1 - агаровий блок, вирізаний із газону S. cyanogenus AlanI; 2 - блок, вирізаний із газону S. cyanogenus AlanI pGM4181 d.

Спосіб можна проілюструвати прикладами:

Для надекспресії частини гена XNR 4181, що кодує ДНК-зв'язувальний домен ортолога AdpA в S, albus J1074, конструюють плазміду pGM4181d. За допомогою полімеразної ланцюгової реакції із використанням полімерази KOD HotStart з хромосомної ДНК S. albus J1074 ампліфікують послідовність ДНК розміром 560 п.н., що кодує ДНК-зв'язувальний домен білка XNR_4181. Усі маніпуляції виконують згідно стандартних методик [Sambrook J., Russell D. W. Molecular cloning, a laboratory manual. 3rd ed. // Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 2001. 450 P.]. Для цього використовують пару праймерів xnr41 81 araC_up (AAATCTAGACGAGCAACGGAGGTACGGACATGCCGGAGGAGATCGGCGCCGAC) та Xnr4181_EcoRIrp (AAAGAATTCGGAGCTGTCCTCTCTCAGAC). Перший праймер (xnr4181 araC_up) містить у своєму складі сайт зв'язування рибосоми та старт кодон для того, щоб забезпечити трансляцію ДНК-зв'язувального домена білка XNR_4181. Амплікон елююють з агарозного гелю відповідно до стандартних методик [Sambrook J., Russell D. W. Molecular cloning, a laboratory manual. 3rd ed. // Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. - 2001. - 450 P.]; елюйований амплікон обробляють ендонуклеазами рестрикції XbaI та EcoRJ. Далі вектор pGUSmoeE5script [Makitrinskyy R., Ostash B., Tsypik O., Rebets Y., Doud E., Meredith T., Luzhetskyy A., Bechthold A., Walker S., Fedorenko V. Pleiotropic regulatory genes bldA, adpA and absB are implicated in production of phosphoglycolipid antibiotic moenomycin // Open Biol. - 2013. - Vol. 3. - P. 1-13] обробляють ендонуклеазами рестрикції SpeI та EcoRI. Отримані лінійні фрагменти вектора та досліджуваного гена лігують між собою, далі отриманою сумішшю трансформують компетентні клітини E. coli і відбирають трансформантів за стійкістю до апраміцину у концентрації 100 мкг/мл за стандартними методиками [Sambrook J., Russell D. W. Molecular cloning, a laboratory manual. 3rd ed. // Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. - 2001. - 450 P.]. У сконструйованій плазміді pGM4181d транскрипція фрагмента гена XNRj418I знаходиться під контролем промотора moeE5p, як зображено на фіг. 2. Отриманий вектор перевіряють за допомогою рестрикційного картування.

Плазмідою pGM4181d трансформують штам E. coli ET12567 (pUZ8002), який за рахунок tra-генів плазміди plJZ8002 забезпечує кон'югативне перенесення корезидентних плазмід [Paget M., Chamberlin L., Atrih A., Foster S., Buttner M. Evidence that the extracytoplasmic function sigma factor σ^E is required for normal cell wall structure in Streptomyces coelicolor A3(2) // J. Bacteriol. 1999. - vol. 181. - P.204-211.]. Трансформацію здійснюють за стандартними методиками [Sambrook J., Russell D. W. Molecular cloning, a laboratory manual. 3rd ed. // Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. - 2001. - 450 P.], трансформантів E. coli ET12567 (pUZ8002, pGM4181d) відбирають за стійкістю до апраміцину у концентрації 100 мкг/мл на середовищі із додаванням Km у концентрації 50 мкг/мл, що необхідно для підтримання плазміди pUZ8002.

Плазміду pGM4181d у штам-продуцент ландоміцину A Streptomyces cyanogenus S136 та в Streptomyces cyanogenus Δlan1 із зруйнованим геном шлях-специфічного регулятора кластера біосинтезу ландоміцину, що не продукує ландоміцину, переносять шляхом міжродової кон'югації зі штамом E. coli ET 12567 (pUZ8002) pGM4181 d.

Вегетативні клітини міцелію штамів S136 та Δlan1 висівають на вівсяне середовище (г/л: вівсяне борошно - 30, агар - 18, вода водопровідна - до 1 л, pH до стерилізації - 7,0; після стерилізації додають розчин хлориду магнію до кінцевої концентрації 10 мМ) та вирощують 7 діб при 30 °C для отримання газону клітин. Штам E. coli ET12567 (pUZ8002, pGM4181d) висівають на чашку із середовищем LA зі 100 мкг/мл апраміцину і 50 мкг/мл канаміцину та вирощують 18 год. при 37 °C. Готують суспензію клітин у 4 мл середовища LB. Одночасно готують суспензію вегетативного міцелію S. cyanogenus у 7 мл стерильної дистильованої води. Суспензію клітин і вегетативного міцелію S. cyanogenus S136 (S. cyanogenus Δlanl) осаджують центрифугуванням

5 хв при 9000 об/хв., зливають надосадову рідину, розчиняють у 100 мкл середовища LB, змішують між собою та висівають на чашки з вівсяним середовищем. Чашки інкубують 12-16 год. при 30 °С і заливають 1 мл водного розчину 50 мкг апраміцину та 30 мкг налідиксової кислоти.

При виконанні досліду використовують: для визначення температури - ртутний термометр з похибкою показів ± 0,5 °С; об'єму рідин - мірні циліндри з похибкою показів ± 0,5 мл та лабораторні дозатори з похибкою показів ± 0,05-0,5 мкл; маси речовин - аналітичну вагу "Sartorius" з похибкою показів ± 0,00005 г та лабораторну вагу AXIS з похибкою показів ± 0,05 г; час - електронним годинником з похибкою показів ± 0,5 сек.

Спочатку перевіряють вплив надекспресії фрагмента гена XNR 4/81, що кодує ДНК-зв'язувальний домен, на біосинтез ландоміцину А. Використовують середовище GYM (г/л: глюкоза - 4, дріжджовий екстракт - 4, мальтозний екстракт - 10, $CaCO_3$ - 2, агар -16), на якому S. cyanogenus S136 не продукує ландоміцину А; тоді відповідний біосинтетичний шлях стає моделлю криптичного кластера генів біосинтезу антибіотика. Для цього на чашки із середовищем GYM наносять по 100 мг вегетативного міцелію штамів S. cyanogenus SI36 і рекомбінантного S. cyanogenus pGM4181d та розподіляють біомасу по усій поверхні чашки шпателем Дригальського. Чашки інкубують упродовж 5 діб, при 30 °С. Після цього, рівень продукції ландоміцину А оцінюють за допомогою тесту на пригнічення росту Bacillus cereus ATCC 19637 та тонкошарової хроматографії. Для тесту на пригнічення росту B. cereus ATCC 19637 у 30 мл спеціального збідненого агару (г/л: KH_2PO_4 - 3, K_2HPO_4 - 7, NaCyt x$4H_2O$ - 0.5, $MgSO_4 \times 7H_2O$ - 0.1, $(NH_4)_2SO_4$ - 1, глюкоза - 2, бактопептон - 0,3, агар - 16) додають 10^6 спор B. cereus ATCC 19637 та розливають на чашку Петрі. Коли агар зі спорами застигає, на його поверхню ставлять блоки діаметром 4 мм, вирізані блокорізом із чашок, на яких вирощувалися відповідні штами S. cyanogenus. Через 16-20 год. навколо блока спостерігають зону пригнічення росту тест-культури B. cereus. Для тонкошарової хроматографії ландоміцину А антибіотик екстрагують із 3,5 $см^3$ агару за допомогою 2 мл етилового ефіру оцтової кислоти протягом 2 год.; суміш центрифугують і відділяють надосадову рідину. Далі надосадову рідину випаровують із використанням вакуумного концентратора, а отриманий осад, в якому міститься ландоміцин А, перерозчиняють у 50 мкл метанолу. Далі 6 мкл концентрованого екстракту наносять на алюмінієву пластинку для ТШХ вкриту силікагелем та розділяють у системі хлороформ:метанол (9:1) протягом 20 хв. Отримані дані проілюстровано на фіг. 3 та фіг. 4. На фіг. 3 помітно, що S. cyanogenus S136 не пригнічує ріст B. cereus ATCC 19637, на відміну від S. cyanogenus pGM4181 d. Дані ТШХ на фіг. 4 узгоджуються з результатами тесту на пригнічення росту тест-культури, показуючи, що S. cyanogenus SI36 практично не продукує ландоміцину А, у той час як S. cyanogenus pGM4181d продукує його у кількостях, що легко визначаються способами біотесту й ТШХ.

Щоби продемонструвати індукцію біосинтезу інших вторинних метаболітів використовують штам S. cyanogenus ΔlanI (що не продукує ландоміцину). Застосовують середовище ISP5 (г/л: L-аспарагін 1, гліцерол 10, K_2HPO_4 - 1, $Fe_2SO_4 \times 7H_2O$ - 0.001, $MnCl_2 \times 4H_2O$ - 0.001, $ZnSO_4 \times 7H_2O$ - 0.001, агар - 20). На чашки із середовищем ISP5 наносять по 100 мг вегетативного міцелію штамів S. cyanogenus ΔlanI і його похідного ΔlanI-pGMA181d та розподіляють біомасу по поверхні чашки шпателем Дригальського. Чашки інкубують упродовж 5 діб, при 30 °С. Після цього антибіотичну активність штамів оцінюють за допомогою тесту на пригнічення росту B. cereus ATCC 1963 7. На фіг. 5 проілюстровано індукцію біосинтезу нової невідомої сполуки із антибіотичною дією, що продукується штамом S. cyanogenus ΔlanI pGM4181d. Помітно, що навколо блока, вирізаного із газону ΔlanI-pGM4181d з'являється зона пригнічення росту B. cereus ATCC 19637, у той час як навколо блока з газону ΔlanI її немає.

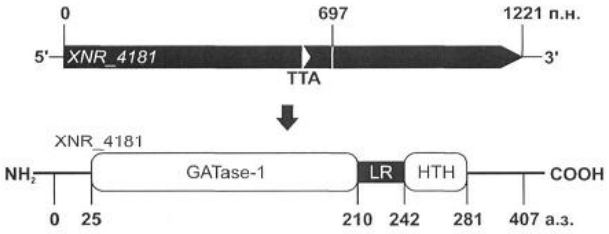
Також, демонструють індукцію біосинтезу вторинних сполук із антигрибковими властивостями. Застосовують середовище YMPG (г/л: дріжджовий екстракт - 4, пептон - 1, мальтозний екстракт - 10, глюкоза 10, MgCb x$6H_2O$ - 2, агар 20). На чашки із середовищем YMPG наносять по 100 мг вегетативного міцелію штамів S. cyanogenus ΔlanI і його похідного ΔlanI-pGM4181d та розподіляють біомасу по поверхні чашки шпателем Дригальського. Чашки інкубують упродовж 5 діб, при 30 °С. Для демонстрації індукції біосинтезу сполук із антигрибковими властивостями використовують тест на пригнічення росту Debaryomyces hansenii (Candida famata) BKM Y-9. Для тесту на пригнічення росту D. hansenii BKM Y-9 у 30 мл середовища м'якого TSA (г/л: триптон-соєвий бульйон - 30, агар - 10) додають 1 мл суспензії клітин D. hansenii BKM Y-9 із OD_{600}=0.3 та розливають на чашку Петрі. Коли агар із клітинами застигає, на його поверхню ставлять блоки діаметром 4 мм, вирізані блокорізом із чашок, на яких вирощувалися відповідні штами S. cyanogenus. Через 16-20 год. навколо блока спостерігають зону пригнічення росту тест-культури D. hansenii. На фіг. 6 проілюстровано

індукцію біосинтезу нової невідомої сполуки із здатністю пригнічувати ріст D. hansenii, що продукується штамом S. cyanogenus ΔlanI pGM4181d. Помітно, що навколо блока, вирізаного із газону ΔlanI-pGM4181d з'являється зона пригнічення росту D. hansenii ВКМ Y-9, у той час як навколо блока з газону ΔlanI така зона не утворюється.

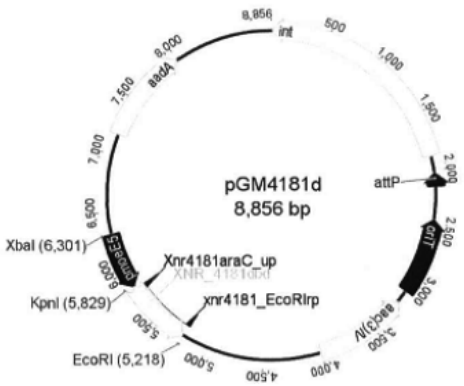
Отже, уведення вектора pGM4181d, який містить С-кінцеву ділянку гена XNR_4181, що кодує ДНК-зв'язувальний домен білка AdpA, у штами S. cyanogenus S136 та ΔlanI активує синтез ландоміцину А та сполуки з протигрибною активністю. Це підтверджує отримання передбачуваного результату.

ФОРМУЛА ВИНАХОДУ

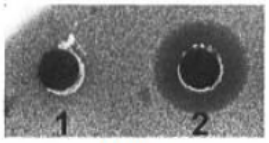
Спосіб активування продукції сполук з антибіотичною та протигрибною властивостями у Streptomyces cyanogenus S136, який ґрунтується на уведенні додаткових копій регуляторних елементів, який **відрізняється** тим, що як регуляторний елемент використовують ділянку гена плейотропного регулятора транскрипції XNR_4181 зі Streptomyces albus J1074, клоновану під контролем промотору moeE5p у складі інтегративної плазміди pGM4181d, при цьому ця ділянка гена кодує ДНК-зв'язувальний домен на С-кінці білка AdpA.



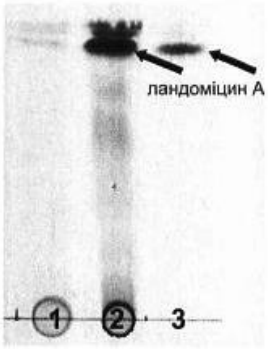
Фіг. 1



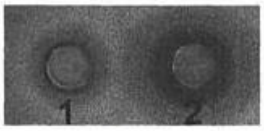
Фіг. 2



Фіг. 3



Фіг. 4



Фіг. 5



Фіг. 6

Комп'ютерна верстка С. Чулій

Міністерство економічного розвитку і торгівлі України, вул. М. Грушевського, 12/2, м. Київ, 01008, Україна

ДП "Український інститут інтелектуальної власності", вул. Глазунова, 1, м. Київ – 42, 01601

РОЗДІЛ 8 Гени стійкості до глікопептидних антибіотиків в бактерій

У попередніх розділах роботи продемонстровано вивчення генетичних механізмів регуляції та біосинтезу різних полікетидних та пептидних антибіотиків. Очевидно, що найбільш важливим із практичної точки зору було вивчення саме глікопептидних антибіотиків (ГПА), які застосовуються в клініці [16]. Саме те, що ці сполуки активно застосовуються, веде до появи та поширення стійких штамів патогенів, а саме ентерококів та стафілококів. З одного боку, нагляд за поширенням стійких до ГПА штамів ведеться доволі прискіпливо [21, 22, 154], однак загальна картина явища залишається загадковою: як на рівні особливостей функціонування генів стійкості до ГПА – *van*-генів, так і на рівні глобальної філогеноміки розповсюдження та еволюції цих генів. Зокрема, невідомо чи відбувається корегуляції біосинтезу ГПА і стійкості до них у продуцентів. Прикладів повного вивчення регуляторних механізмів експресії генів стійкості до ГПА також недостатньо.

У цьому завершальному розділі дисертаційної роботи ми поставили за мету створити детальну модель функціонування *van*-генів у модельного продуцента клінічно-важливого ГПА тейкопланіну – *A. teichomyceticus*, і, зокрема перевірити чи існує взаємозв'язок між механізмами регуляції біосинтезу тейкопланіну і стійкості до нього; оцінити, як розповсюджені *van*-гени у актинобактерій та інших груп бактерій, і прослідкувати особливості і шляхи їх розповсюдження та еволюції. Також ми вивчили особливості розповсюдження та філогенії ABC-транспортерів, кодованих в КБГ ГПА, які можуть відігравати додаткову роль у формуванні стійкості.

8.1 Гени стійкості до глікопептидних антибіотиків в *A. teichomyceticus* ATCC 31121

Отже, відомо, що гени *vanHAXRS* забезпечують стійкість до ГПА у (1) актинобактерій, що продукують ГПА (див **Розділ 1**), (2) актинобактерій не здатних до продукції ГПА [155], (3) патогенів, стійких до ГПА [156], та (4) інших ґрунтових бактерій [157]. У більшості відомих випадків двокомпонентна регуляторна система

VanRS використовується для детекції позаклітинних ГПА [158], або, можливо, комплексу ГПА з ліпідом II [159] і, як наслідок, активації експресії структурних *van*-генів. Патогени та актинобактерії, що не продукують ГПА, зазвичай мають індуцибельні *van*-резистентні фенотипи, оскільки конститутивна експресія *van*-генів веде до зниження загального фітнесу клітин [160, 161].

Фенотип стійкості до ГПА є конститутивним у продуцентів ГПА, як, наприклад, в продуцента баліміцину *Am. balhimycina*. В *Am. balhimycina*, ортологи *vanHAX* є основними факторами стійкості до ГПА [162, 163]. Цікаво, що гени *vanHAX* не розмішені в межах КБГ баліміцину, як це зазвичай є в інших актиноміцетів, що продукують ГПА, і вони не колокалізовані з ортологами *vanRS* [163]. Отже, експресія *vanHAX* у цього штаму позбавлена *vanRS*-опосередкованої регуляції та є конститутивною. Примітно, що *Am. balhimycina* також має додаткові механізми стійкості до ГПА, які не є *van*-опосередкованими [164]. Генетичні механізми стійкості до ГПА у інших представників роду *Amycolatopsis*, мабуть, аналогічні таким в *Am. balhimycina*. У представників іншого роду актиноміцетів-продуцентів ГПА, а саме *Nonomuraea*, стійкість до ГПА визначається функціонуванням гена *vanY*, що кодує D-Ala-D-Ala карбоксипептидазу; гени *vanHAXRS* відсутні [45, 165, 166].

Модельний об'єкт третього архетипового роду (*Actinoplanes*) продуцентів ГПА – *A. teichomyceticus* ATCC 31121 – містить гени *vanHAX* (*tei7-6-5*) та *vanRS* (*tei2-3*) колокалізовані та розташовані в КБГ. На даний момент, функції цих генів вивчені недостатньо. Відомо, що принаймні *tei7-6-5* формують один оперон, а стосовно сенсорної гістидинової кінази Tei3 припускається, що вона може функціонувати конститутивно і не потребувати ліганда для активації *vanHAX* [167].

У цьому розділі ми вирішили прояснити деякі аспекти стійкості до тейкопланіну в *A. teichomyceticus* ATCC 31121, створюючи інтегровану модель. Спочатку, ми перевірили транскрипційну організацію *van*-генів в *A. teichomyceticus*, та виявили, що *tei2-3* (ортологи *vanRS*) транскрибуються спільно з *tei4*, який кодує потенційну дигідрофолатредуктазу [168]. Важко сказати, чи впливає Tei4 на стійкість до тейкопланіну. Також, було підтверджено, що гени *tei7-6-5* (ортологи *vanHAX*) котранскрибуються, але *tei8*

(кодує потенційний MurF-подібний білок) становить окрему транскрипційну одиницю [168]. Подібно до *Am. balhimycina*, експресія генів *tei2-3* і *tei7-6-5* виявилася конститутивною і не залежала від фази росту чи середовища культивування [168, 169]. Важливо, що експресія *van*-генів виявилася незалежною і від шлях-специфічного регулятора біосинтезу тейкопланіну Tei15* [169]. Реконструкція регуляторної ланки Tei15* → *tei2-3* у *S. albus* (=*albidoflavus*) J1074 продемонструвала, що промотор *tei2*, найімовірніше, не залежить від регуляторного впливу Tei15* [168]. Також, промотори генів *tei2-3-4* і *tei7-6-5* виявилися дуже активними в клітинах *A. teichomyceticus* ATCC 31121, а для Tei2 ми *in vivo* показали зв'язування з *tei7p* [168].

Із літератури відомо, що в *S. coelicolor* A3(2) певні мутації в АТФазному домені $VanS_{Sc}$ можуть надавати йому конститутивну кіназну активність [155]. Порівняння послідовностей Tei3 і $VanS_{Sc}$ показує, що аналогічні точкові мутації також присутні в АТФ-азному домені Tei3 [167]. Зокрема, у $VanS_{Sc}$ це заміни L216P і G271V [155], гомологічними сайтами до яких в Tei3 є N215 і R270. Ми вирішили перевірити, чи дійсно Tei3 функціонує конститутивно. Тому, ми реконструювали Tei3-опосередкований шлях стійкості до ГПА в гетерологічній системі *S. coelicolor* M512 і отримали *in vivo* докази того, що Tei3 функціонує як конститутивна сенсорна кіназа. Примітно, що Tei3 зміг активувати резистентність до тейкопланіну та A40926 у *S. coelicolor*, ймовірно, шляхом фосфорилювання неспорідненого регулятора відповіді – $VanR_{Sc}$ [170]. Враховуючи те, що АТФ-азні домени Tei3 і $VanS_{Sc}$ значно відрізняються й окрім згаданих вище замін, ми припускаємо, що конститутивна кіназна активність, скоріше, є наслідком значної еволюційної дивергенції цих білків, а не лише декількох замін. Експресія вкорочених варіантів *tei3* та $vanS_{Sc}$, у яких відсутні сенсорні домени, продемонструвала, що окремих АТФазних доменів недостатньо для надання стійкості [170].

Найбільш важливим результатом нашої роботи було те, що сенсорний домен Tei3 не реагував на тейкопланін, але розпізнавав A40926: експресія гібридної сенсорної кінази, що несе сенсорний домен Tei3, злитий з АТФазним доменом $VanS_{Sc}$, змінила конститутивний фенотип резистентності до ГПА в

S. coelicolor J3200, зробивши штам чутливим до тейкопланіну та ванкоміцину, але стійким до A40926 [170]. Це спонукало нас зосередитися на відповідних відмінностях хімічної структури A40926 і тейкопланіну (наявність залишку Glc*N*Ac у тейкопланіні, який відсутній в A40296), а також на можливій еволюції моделі глікозилювання тейкопланіну та всього *tei*. Реконструкція філогенії глікозилтрансфераз, закодованих у відомих КБГ ГПА, дала змогу нам побудувати можливі сценарії, що б пояснювали отримані результати. Зокрема виявилося, що предковий КБГ до *tei*, найімовірніше, кодував біосинтез дез-Glc*N*Ac-тейкопланіну, який структурно нагадує A40926 (що виявився субстратом сенсорного домена Tei3). Ми припускаємо, що Tei3 у продуценті дез-Glc*N*Ac-тейкопланіну стала неіндуцибельною внаслідок випадкової мутації, відповідно давши КБГ простір для подальшої еволюції та структурних змін продукованого ГПА. Відповідно, мутації, що ведуть до неіндуцибельності Tei3, стали преадаптацією, яка дала змогу тейкопланіну з'явитися в його нинішній формі [170].

У підсумку, наші результати дають змогу створити доволі повну модель функціонування генів стійкості до ГПА в *A. teichomyceticus*. Вони згруповані у два оперони: *tei2-3-4* (ортологи *vanRS* + ген дигідрофолат редуктази) і *tei7-6-5* (ортологи *vanHAX*). Експресія обох оперонів протягом всього життєвого циклу рівномірно забезпечена сильними і конститутивними промоторами, незалежними від шлях-специфічної регуляції біосинтезу тейкопланіну, а Tei2 розпізнає промотор *tei7p*. Сенсорна гістидинова кіназа Tei3 володіє конститутивною фосфорилазною активністю, а її сенсорний домен розпізнає не тейкопланін, а споріднений ГПА A40926. Важливо відмітити і декілька практично-цінних результатів, отриманих у цьому розділі. Зокрема, ми показали промотори *tei2p* і *tei7p* як потенційні інструменти для надекспресії генів в *A. teichomyceticus*. Іншим важливим результатом стало створення тест-системи на основі *S. coelicolor*, здатної до виявлення ванкоміцину в нг-діапазоні.

Основні результати, отримані в цьому підрозділі дисертації опубліковані в таких роботах автора: [168–170].

Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-019-09789-w

APPLIED GENETICS AND MOLECULAR BIOTECHNOLOGY

Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes

Oleksandr Yushchuk[1] · Liliya Horbal[2] · Bohdan Ostash[1] · Flavia Marinelli[3] · Wolfgang Wohlleben[4] · Evi Stegmann[4] · Victor Fedorenko[1]

Received: 14 February 2019 / Revised: 15 March 2019 / Accepted: 19 March 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Teicoplanin is a frontline glycopeptide antibiotic produced by *Actinoplanes teichomyceticus*. It is used to treat complicated cases of infection, including pediatric ones, caused by Gram-positive pathogens. There is a steady interest in elucidating the genetic mechanisms determining teicoplanin production, as they would help overproduce known teicoplanins and discover novel glycopeptides. Herein, we investigate the transcriptional organization of the *tei* biosynthetic gene cluster and the roles of the cluster-situated regulatory genes in controlling teicoplanin production and self-resistance in *A. teichomyceticus*. We demonstrate that the *tei* cluster is organized into nine polygenic and nine monogenic transcriptional units. Most of *tei* biosynthetic genes are subjected to StrR-like Tei15* control, which, in turn, appears to be regulated by LuxR-type Tei16*. Expression of the genes conferring teicoplanin self-resistance in *A. teichomyceticus* is not co-regulated with antibiotic production. The gene *tei31**, coding for a putative DNA binding protein, is not expressed under teicoplanin producing conditions and is dispensable for antibiotic production. Finally, phylogenesis reconstruction of the glycopeptide cluster-encoded regulators reveals two main clades of StrR-like regulators. Tei15* and close orthologues form one of these clades; the second clade is composed by orthologues of Bbr and Dbv4, governing the biosynthesis of balhimycin and teicoplanin-like A40926, respectively. In addition, the LuxR-type Tei16* appears unrelated to the LuxR-like Dbv3, which is controlling A40926 biosynthesis. Our results shed new light on teicoplanin biosynthesis regulation and on the evolution of novel and old glycopeptide biosynthetic gene clusters.

Keywords *Actinoplanes* · Glycopeptide antibiotics · Glycopeptide resistance · Teicoplanin · Cluster-situated regulatory genes · *tei* cluster

Introduction

Transcriptional regulators encoded in antibiotic biosynthetic gene clusters (BGCs) often are the most important players in antibiotic biosynthesis regulation, switching it on and off in response to changing environmental conditions (Liu et al. 2013). Consequently, cluster-situated regulatory genes are widely manipulated in order to improve the biosynthetic potential of antibiotic producers. Such approaches have been successfully applied to enhance production of glycopeptide antibiotics, with a particular focus on those actually in clinical practice and with a relevant industrial value (Spohn et al. 2014; Marcone et al. 2018). They include first-generation natural products (teicoplanin and vancomycin) and second-generation semisynthetic molecules derived from natural products (dalbavancin, oritavancin, and telavancin). Teicoplanin is particularly important, because it is considered a drug of "last resort" for treating severe infections by

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00253-019-09789-w) contains supplementary material, which is available to authorized users.

✉ Victor Fedorenko
viktor.fedorenko@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, Rm. 102, 4 Hrushevskoho St., Lviv 79005, Ukraine

[2] Department of Pharmaceutical Biotechnology, Helmholtz Institute for Pharmaceutical Research Saarland, Actinobacteria Metabolic Engineering Group, Saarland University, UdS Campus C2 3, 66123 Saarbrucken, Germany

[3] Department of Biotechnology and Life Sciences, University of Insubria, via J.H. Dunant 3, 21100 Varese, Italy

[4] Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Microbiology/Biotechnology, University of Tübingen, Auf der Morgenstelle 28, 72076 Tubingen, Germany

Published online: 01 April 2019

multiresistant Gram-positive pathogens, including the methicillin-resistant *Staphylococcus aureus* (MRSA) and the anaerobe *Clostridioides difficile*, with a therapeutic value superior to vancomycin and comparable to novel semisynthetic second-generation glycopeptides (Binda et al. 2014). Suitability for pediatric usage remains an important advantage of teicoplanin over semisynthetic glycopeptides (Bassetti and Cruciani 1990; Fuchs et al. 2018). Very recently, teicoplanin has been used to functionalize magnetic nanoparticles providing a novel promising tool for treating deep *S. aureus* biofilms (Armenia et al. 2018).

Teicoplanin is produced by the rare actinobacterium *Actinoplanes teichomyceticus* (Somma et al. 1984) as a complex of natural compounds, varying only in the length and branching of aliphatic side chain (Fig. 1). Because of the importance of teicoplanin, many aspects of its biosynthesis (Li et al. 2004; Yushchuk et al. 2016), regulation (Horbal et al. 2013; Horbal et al. 2014), self-resistance (Beltrametti et al. 2007), and optimal production conditions (Taurino et al. 2011) have been intensively studied in the last two decades. Annotation of the teicoplanin BGC was reported simultaneously by two independent groups and the cluster was named as *tcp* (Sosio et al. 2004) or *tei* cluster (Li et al. 2004), respectively. Following what described in previous papers (Horbal et al. 2013; Horbal et al. 2014; Yushchuk et al. 2016), hereby we refer to the *tei* BCG annotation (Li et al. 2004). According to this annotation, all the genes before NRPS-encoding ones received numeric indexes (e.g., *tei1*), the ones after, numeric indexes with asterisks (e.g., *tei1**); finally, the genes coding for NRPS received alphabetic indexes (e.g., *teiA*). The cluster was reported to span approximately 89 kb and include 49 open reading frames (ORFs) that were predicted to participate in teicoplanin biosynthesis, including genes responsible for the assembly of the antibiotic, resistance, export, and regulation of its synthesis. However, real borders of the cluster were not clearly identified at that time and the role of many genes was not investigated further.

The *tei* cluster contains five putative regulatory genes (Horbal et al. 2014). *tei2* and *tei3* genes code for a response regulator and a kinase, respectively, which show high level of homology to the VanR/VanS system of *Streptomyces coelicolor* (Hutchings et al. 2006) and were supposed to be involved in regulation of resistance gene expression (Beltrametti et al. 2007; Horbal et al. 2014). The other three cluster-situated regulatory genes are *tei15**, *tei16**, and *tei31**. Two of them, *tei15** and *tei16**, encode regulatory proteins belonging to StrR and LuxR families, respectively. The last one, *tei31**, encodes a putative DNA-binding protein of unknown function. Recent studies demonstrated that Tei15* and Tei16* are key positive regulators of teicoplanin biosynthesis (Horbal et al. 2014). Overexpression of *tei15** and *tei16** led to strong overproduction of teicoplanin, and their knockout completely abolished antibiotic production. Additionally, in vitro and in silico analyses of the DNA-binding properties of Tei15* allowed to identify putative binding sites in 12 promoter regions within *tei* cluster, including the promoter regions of *tei2* and of *tei31** (Horbal et al. 2014).

In this paper, we investigated the transcriptional organization of *tei* BGC using RT-PCR and its dependence on cluster-situated regulatory genes *tei15** and *tei16**. We also demonstrated that *tei31** was dispensable for teicoplanin production

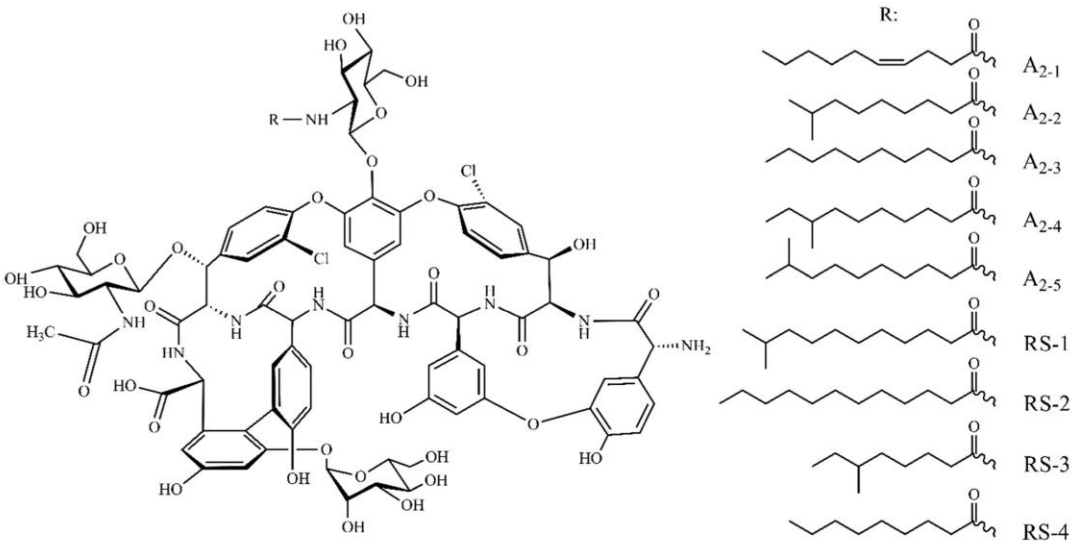


Fig. 1 Chemical structure of the teicoplanin complex produced by *A. teichomyceticus* NRRL-B16726. The five components of the A_2 group prevail in the teicoplanin complex and are the most therapeutically important

and it was not expressed in teicoplanin-producing cultures. Additionally, self-resistance was definitively proved not to be co-regulated with antibiotic production. Finally, we investigated the positions of *tei* cluster-encoded regulators in the reconstructed phylogeny of the regulators from the 21 glycopeptide BGCs today available. Our results shed new light on teicoplanin biosynthesis regulation as well as on evolution of regulatory circuits controlling glycopeptide antibiotic production, providing a novel knowledge-based approach to group old and novel glycopeptide BGCs.

Materials and methods

Strains and growth conditions The list of the bacterial strains used in this work is reported in Table 1. Lysogeny broth (LB) medium was used to grow *Escherichia coli* strains; antibiotics were added to cultures when necessary at the following concentrations (μg/ml): apramycin 100, kanamycin 50, and chloramphenicol and hygromycin 100. All medium components and antibiotics were purchased from Sigma-Aldrich (St. Louis, MO, USA), except the finely ground whole oats for ISP3 (tolokno; Kozub®, Ukraine). *A. teichomyceticus* NRRL-B16726 (= ATCC 31121) and its recombinant strains were grown on ISP3 (g/L: finely ground whole oats – 30, agar – 18, pH 8) or SM (soya flour mannitol medium, Kieser et al. 2000) for routine maintenance with the addition of 50 μg/ml apramycin. To collect spore suspensions needed for the conjugation, *A. teichomyceticus* strains were grown on ISP3 medium; conjugation was performed as described previously (Ha et al. 2008) on SM agar plates supplemented with 40 mM $MgCl_2$. Morphology of *A. teichomyceticus* strains was observed when the strains were grown on ISP2 (g/L: yeast extract – 4, malt extract – 10, dextrose – 4, agar – 20, pH 7.2) and ISP3 media. For DNA isolation, *A. teichomyceticus* strains were grown in ISP2 liquid medium. *E. coli* strains were cultivated at 37 °C, *A. teichomyceticus* at 30 °C.

tei31* **knockout and overexpression** The coding sequence of the *tei31** gene was replaced with the *aac(3)IV* cassette within the cosmid JB7hyg (Table 1) using the λRed-mediated recombination (Gust et al. 2002), giving JB7hyg*tei31**::*aac3(IV)*. The tei31REDF/tei31REDR primer pair was used for the replacement (Table 2). The gene disruption cosmid was delivered to *A. teichomyceticus* NRRL-B16726 via intergeneric conjugation with *E. coli* ET12567 pUZ8002. Transconjugants were selected on plates overlaid with 1 ml of distilled water containing 1.5 μg of apramycin and 0.9 μg of nalidixic acid. To generate *A. teichomyceticus* Δ*tei31**, apramycin- and hygromycin-resistant single-crossover transconjugants were screened for the double-crossover event, which led to the loss of hygromycin resistance. Replacement of *tei31** in double-crossover mutants was confirmed using PCR with 311/312 primer pair. Such PCR reaction should generate a fragment of 1885 bp from the chromosomal DNA of the wild type and 1541 bp amplicon using the *A. teichomyceticus* Δ*tei31** chromosomal DNA as a template (Fig. S1, ESM). To overexpress *tei31**, the coding sequence was cloned together with the putative *tei31** promoter region, using the 311/312 primer pair. Generated amplicon (1885 bp) was blunt-end-ligated into the *Eco*RV-digested pSET152 vector, giving pSETtei31* (Table 1). The recombinant plasmid was verified by restriction mapping and sequencing. pSETtei31* was transferred into *A. teichomyceticus* NRRL-B16726 via intergeneric conjugation with *E. coli* ET12567 pUZ8002. Transconjugants carrying pSETtei31* were selected as described above. Recombinant strains were verified by amplification of *aac3(IV)* gene using Am_F/Am_R primer pair (Table 2).

Teicoplanin production and quantification To analyze the production of teicoplanin and to collect biomass for RNA

Table 1 Bacterial strains and plasmids used in this work

Name	Description	Source of reference
A. teichomyceticus	Wild type, teicoplanin producer	NRRL-B16726
A. teichomyceticus Δ*tei15**	*A. teichomyceticus* derivative with *tei15** regulatory gene inactivated	Horbal et al. (2014)
A. teichomyceticus Δ*tei16**	*A. teichomyceticus* derivative with *tei16** regulatory gene inactivated	Horbal et al. (2014)
A. teichomyceticus Δ*tei31**	*A. teichomyceticus* derivative with *tei31** regulatory gene inactivated	This work
A. teichomyceticus pSETtei31*$^+$	*A. teichomyceticus* derivative carrying pSETtei31* vector	This work
E. coli DH5α	General cloning host	MBI Fermentas, USA
E. coli ET12567 (pUZ8002$^+$)	(*dam-13*::Tn9 *dcm-6*), pUZ8002$^+$ (ΔoriT), used for conjugative transfer of DNA	Kieser et al. (2000)
pSET152	ϕC31-based integrative vector, Am^r	Kieser et al. (2000)
pSETtei31*	pSET152 derivative, carrying *tei31** gene together with its putative promoter region, Am^r	This work
JB7hyg	JB7 cosmid derivative, Hg^r, Ap^r	Yushchuk et al. (2016)
JB7hygtei31*::aac3(IV)	JB7hyg cosmid derivative having the coding sequence of *tei31** exchanged for *aac3(IV)*	This work

Table 2 Primers used in this work

Primer	Nucleotide sequence (5′–3′)	Purpose
Am_F	CGGGGTCTGACGCTCAGTGGA	*aac3(IV)* amplification
Am_R	AATCAGCGCGACCTTGCCCC	
31*1	GACGGAGGTCCAGCAGGGCAC	*tei31* cloning*
31*2	CACCTGGCATGTCGGTCGCA	
tei31REDF	GAAGGCACTCCAGCGGTCCTTGGGAGGCA CATTCGGGTGATTCCCGGGGATCCGTCGACC	*aac3(IV)* amplification in order to knockout *tei31**
tei31REDR	CCGCCTGGCGGGTACACCTGGCATGTCGG TCGCACATCATGTAGGCTGGAGCTGCTTC	
tei5_RT_F	GGCGACCTGCGACTTCCTC	*tei5* reverse transcript identification
tei5_RT_R	GCGCCCATGACGAACGCGG	
tei2_RT_F	CGGGTCCGCGTTCTCGTCC	*tei2*reverse transcript identification
tei2_RT_R	CGCCTACAGCAGGCCACCT	
tei1_RT_F	GTGCCGATGGTGCCGATCG	*tei1* reverse transcript identification
tei1_RT_R	CAGGGGCGCTCGGTGAAGGC	
teiA_RT_F	CTTAGCCCCGGCACTGGTGG	*teiA* reverse transcript identification
teiA_RT_R	GCCGGATCCACCGGGACGAA	
tei3*_RT_F	GGCCAAGATGGTGCAGGCG	*tei3** reverse transcript identification
tei3*_RT_R	TTGACCTCCTCGGCGCCGG	
tei4*_RT_F	GGCCTTGGTGGACTCCTCC	*tei4** reverse transcript identification
tei4*_RT_R	AGCGCGACCGCACCGGGACG	
tei5*_RT_F	ACGCCATCGAGGCGCTCGGC	*tei5** reverse transcript identification
tei5*_RT_R	GCCCATGACGAACGCGGCG	
tei8*_RT_F	CTCGGTCACCGGGGTGATC	*tei8** reverse transcript identification
tei8*_RT_R	TGTCGCTCAGCGGGATGTA	
tei12*_RT_F	ATCCGGCCCACCACCCTGAC	*tei12** reverse transcript identification
tei12*_RT_R	CGGTCGGGCAGGTCGTTGTA	
tei13*_RT_F	TCGGCGATCCGGCGGGCTGG	*tei12** reverse transcript identification
tei13*_RT_R	CGGCTATCGTCCGCAGCCAG	
tei14*_RT_F	TACCAGGGGTTGGGGGAGCG	*tei14** reverse transcript identification
tei14*_RT_R	GCGGCGCTGAGCCATTCCT	
tei15*_RT_F	CGGTGACCGGTATCTCGCCG	*tei15** reverse transcript identification
tei15*_RT_R	TGAGCAGCCGATTGCGGACG	
tei16*_RT_F	TGGGCGAGCTGAGCCAGGA	*tei15** reverse transcript identification
tei16*_RT_R	AGGTGACCACGGCCATCGGA	
tei17*_RT_F	CGTGACGTCGACCGGCTTCC	*tei15** reverse transcript identification
tei17*_RT_R	CGCCGAAGAGGCTGTTGACCA	
tei24*_RT_F	TGTTCGCCGGCAAGACCTGG	*tei15** reverse transcript identification
tei24*_RT_R	TGACGTCACGGACCCCCTGG	
tei25*_RT_F	GTCCGGCTCGCCGTGCTCGTC	*tei15** reverse transcript identification
tei25*_RT_R	TTCATCAGCGCCGACAACCG	
tei27*_RT_F	CACCCGCACCTACACCGTGC	*tei15** reverse transcript identification
tei27*_RT_R	GTCCACCTCCACCAGGGCCA	
tei29*_RT_F	GGACATCCGCAACGAGTTC	*tei15** reverse transcript identification
tei29*_RT_R	CAGGATGCCCTTCACCAC	
tei30*_RT_F	TGCTCGGCCCGGTCCCGCTG	*tei15** reverse transcript identification
tei30*_RT_R	AGCGGTAGCCGTCGAGCAGC	
tei31*_RT_F	CCAGCCCGGCGCATGTCGG	*tei15** reverse transcript identification
tei31*_RT_R	GCAGGCGGCACAGCGCATC	
teiB-C_RT_F	GCGCGCGAGAGGTGTTCCAG	*teiB-C* intergenic region reverse transcript identification

Table 2 (continued)

Primer	Nucleotide sequence (5′–3′)	Purpose
teiB-C_RT_R	GTCCACCGGGCCGTCGAGAT	
CD_F	GACCCCGTGGCAGGTGTTCG	*teiC-D* intergenic region reverse transcript identification
CD_R	GCATGCCGGTCCACCAGGG	
teiD-1*_RT_F	TGCCGGTGTCCGAGGCGATA	*teiD-1** intergenic region reverse transcript identification
teiD-1*_RT_R	GCCGAGACAATCCTGCCGGG	
tei2*-3*_RT_F	TCGACCGGCTGGACGAGCAT	*tei2*-3** intergenic region reverse transcript identification
tei2*-3*_RT_R	CCAGGCATACAGCACCGCCG	
tei5*-6*_RT_F	GATTCCGGCCCTGGCACACC	*tei5*-6** intergenic region reverse transcript identification
tei5*-6*_RT_R	GTCCGGTGACCAGCCACTGC	
6*7*_F	CGGTTCCCGTCTTTGCGGCT	*tei6*-7** intergenic region reverse transcript identification
6*7*_R	CCCGCCGATCTCGTCACGTC	
8*9*_F	GCAGGTGATGCAGGAGGGCG	*tei8*-9** intergenic region reverse transcript identification
8*9*_R	CACCGAGGATCTGCCGCACG	
9*10*_F	CCGGCGCCACACATCGGATT	*tei9*-10** intergenic region reverse transcript identification
9*10*_R	GTAGCGATGGCCTCGGCGTC	
10*11*_F	CCCGACCGTCGAGTCCCTGT	*tei10*-11** intergenic region reverse transcript identification
10*11*_R	GATCACGTGCCTCCTCGCCG	
11*12*_F	CTTGCGCGGTTCTGGACGCT	*tei11*-12** intergenic region reverse transcript identification
11*12*_R	CCGGAGCCGTCGCGTACATC	
tei19*-20*_RT_F	CCGCCGGAGGAGATGGACGA	*tei19*-20** intergenic region reverse transcript identification
tei19*-20*_RT_R	ATGGCGTTGAGCACGTCGGG	
tei20*-21*_RT_F	TCGGTCCGGGCGATCAAGGA	*tei20*-21** intergenic region reverse transcript identification
tei20*-21*_RT_R	GTGGCCAGCTCACCCACGAC	
tei22*-23*_RT_F	GGTCGGCGCGAACACCATCA	*tei22*-23** intergenic region reverse transcript identification
tei22*-23*_RT_R	GCGAACGAGATCGCCTCGGG	
tei23*-24*_RT_F	GCGGTACTTCCACCCGGACG	*tei23*-24** intergenic region reverse transcript identification
tei23*-24*_RT_R	CGACGAGCGCGAGATCCACC	
29*30*_F	GGTGTTCGTGGACAGCGGCA	*tei29*-30** intergenic region reverse transcript identification
29*30*_R	CCCCGGCGTACGGGAAACAG	
rpoB_RT_F	GCGCCGAGGTCGACTACATGG	$rpoB_{AT}$ reverse transcript identification
rpoB_RT_R	TCGAGCGGCGGAACTTGTGC	

isolation in teicoplanin producing conditions, 1 ml of stocked working cell bank (WCB) of *A. teichomyceticus* strains, prepared as previously described (Taurino et al. 2011), was inoculated into 50 ml of E25 liquid seed medium and incubated for 72 h at 28 °C on the rotary shaker (220 rpm). Then, 2.5 ml of these pre-cultures was transferred into 300-ml baffled Erlenmeyer flasks containing 50 ml of TM1 production medium and grown for 120 h at 28 °C on the rotary shaker (220 rpm). Seed (E25) and production (TM1) medium were prepared as previously described (Taurino et al. 2011). For RNA extraction, biomass was collected at regular intervals (see below) and teicoplanin production was measured after 120 h of cultivation in TM1. For teicoplanin quantification, one volume of fermentation broth (1 ml) was mixed with an equal volume of borate buffer (100 mM H_3BO_3, 100 mM NaOH, pH 12) and left in the thermomixer (36 °C, 600 rpm) for 30 min. Then, mixtures were centrifuged and the supernatants were used for HPLC. HPLC analyses and calculations were performed as previously described (Taurino et al. 2011). Teicoplanin production level for each recombinant strain was calculated as a mean value of three independent HPLC analyses from three independent fermentations.

Total RNA isolation and semi-quantitative RT-PCR For RNA extraction, biomass was collected from 5 ml of cultures grown in TM1 at four time points (48, 72, 96, and 120 h). Mycelium was recovered by centrifugation and then washed with TE-buffer (pH 7.5). Pellets were frozen at − 80 °C or directly used for RNA isolation. To isolate RNA, samples were transferred into Precellys screw cap tubes with 200 μl of glass beads (Ø 0.17–0.18), then 0.75 ml of Kirby mix (Kieser et al. 2000) was added, and mycelium was lysed in a Precellys 24 bead beater (Bertin

instruments, France) for 30 s at 6000 rpm. Of phenol-chloroform-isoamyl alcohol (PCI), 0.5 ml was added to the sample, followed by two additional rounds of homogenization in the Precellys 24 bead beater under the same conditions. RNA was extracted with PCI according to (Kieser et al. 2000), and the integrity of the extracted RNA was checked by horizontal gel electrophoresis (HGE) in agarose gel. DNA digestion was performed in two steps: first, 50 μl of RNA sample was treated with 4 U of RNase free DNase I (Thermo Fisher Scientific, USA) for 90 min, the RNA integrity was checked via HGE, and RNA concentration was measured by NanoDrop (Thermo Fisher Scientific, USA). Second, 60 μg of RNA was treated with 8 U of DNase I for 2 h. Obtained samples were checked for DNA contamination with PCR. If DNA traces were still present, RNA samples were further incubated with DNase I for additional 2 h and then PCR was repeated. cDNA was synthesized using RevertAid First Strand cDNA synthesis Kit (Thermo Fisher Scientific, USA), according to the supplier's instructions. Four hundred nanograms of cDNA was used as a template in a standard PCR reaction aimed to amplify fragments of transcripts of interest. The cycling conditions were following: 95 °C–5 min (95 °C–20 s; 59 °C–20 s; 72 °C–30 s) × 24. Obtained amplicons were visualized using HGE in a 2% agarose gel.

In silico analysis

Amino acid sequences of proteins homologous to *tei* regulatory genes were retrieved using Blast (Altschul et al. 1990). Conserved domain search (CD-search, Marchler-Bauer and Bryant 2004) was utilized for comparing conserved domains in amino acid sequences. Search for proteins with similar domain architecture was performed using CDART (Geer et al. 2002). Phylogenetic analysis was performed using Mega 6 (Tamura et al. 2013).

Results

Transcriptional organization of the *tei* cluster

BGCs for different glycopeptides vary in transcriptional organization and number of ORFs. Additionally, they are generally not collinear, making difficult to predict their transcriptional organization by comparison. For instance, balhimycin BGC (*bal*) possesses 36 ORFs, presumably united into 6 polygenic and 5 monogenic transcriptional units (Pfeifer et al. 2001; Puk et al. 2004; Shawky et al. 2007). By turn, the 37 ORFs of the A40926 BGC (*dbv*) are grouped into 8 polygenic and 5 monogenic transcriptional units (Alduina et al. 2007). Teicoplanin BGC (Fig. 2a) is one of the largest BGCs involved in the production of glycopeptide antibiotics. Initially, 49 ORFs were considered to compose the *tei* cluster (Li et al. 2004; Horbal et al. 2014), but analyzing the bordering regions, we could reasonably exclude genes *tei11*, *tei10*, *tei9*, *tei8*, *tei32**, *tei33**, and *tei34** (indeed not reported in Fig. 2a) from the *tei* cluster, as they encode proteins which cannot be correlated with teicoplanin biosynthesis or resistance. Moreover, these genes have no orthologues in other glycopeptide BGCs (Table S1, ESM). Therefore, in this paper, we proposed new boundaries of the *tei* cluster starting from *tei7* (coding for a D-lactate dehydrogenase orthologous to VanH) and ending with *tei31** (encoding a putative DNA-binding protein and possessing a Tei15* operator in its promoter region). Of the remaining 42 ORFs constituting the *tei* BGC, the transcriptional organization was previously studied only for genes *tei7-6-5* (Fig. 2a) that were reported to form a single transcriptional unit (Beltrametti et al. 2007). These genes are orthologues of *vanHAX*—encoding a D-lactate dehydrogenase, a D-alanine–D-lactate ligase, and a D-alanine–D-alanine-dipeptidase respectively—and confer self-resistance to teicoplanin in *A. teichomyceticus* (Beltrametti et al. 2007; Binda et al. 2018). Similarly, the *vanRS* orthologues *tei3-2* (Fig. 2a) (coding for a putative sensor histidine kinase and a putative response regulator, respectively, that commonly control *vanHAX* expression) were supposed to form one operon, probably involving also *tei4*, coding for a dehydropholate reductase of unknown function (Beltrametti et al. 2007).

For elucidating the transcriptional organization of the *tei* BGC, we started from the observation that in the *bal* and *dbv* BGCs, the size of intergenic regions between the genes grouped in operons never exceeds the 156 bp (Alduina et al. 2007; Shawky et al. 2007). Consequently, we assumed that in *tei* cluster all the intergenic regions longer than 156 bp could be indicative of different adjacent transcriptional units (Fig. 2a). We also took into account that overlapping ORFs necessarily fall into single transcriptional units; on the contrary, adjacent ORFs positioned on different strands cannot be read as polycistronic mRNA. Following these assumptions, we then checked experimentally by semi-quantitative RT-PCR if the 16 intergenic regions (from *tei1* to *tei31**) having a size ranging from 9 to 143 bp could be amplified (Fig. 2b). In these experiments, RNA was always isolated after 72 h of growth in the rich production TM1 medium, where teicoplanin was produced (Taurino et al. 2011) and thus genes from the *tei* cluster should be strongly expressed. Indeed, in these conditions, we observed amplicons for almost all of the selected intergenic regions (Fig. 2b), indicating that neighboring *tei* genes were co-expressed. The only exceptions were the intergeneric regions between *tei23**-*24** (Fig. 2b) and between *tei25**-*26** (data not shown). Later, it was shown (see next section) that *tei25** and *tei26** were not expressed under these cultivation conditions.

As illustrated in Fig. 2, we could conclude that the *tei* cluster consists of nine monogenic (*tei1*, *tei4**, *tei13**, *tei14**, *tei15**, *tei16**, *tei24**, *tei27**, *tei31**) and nine

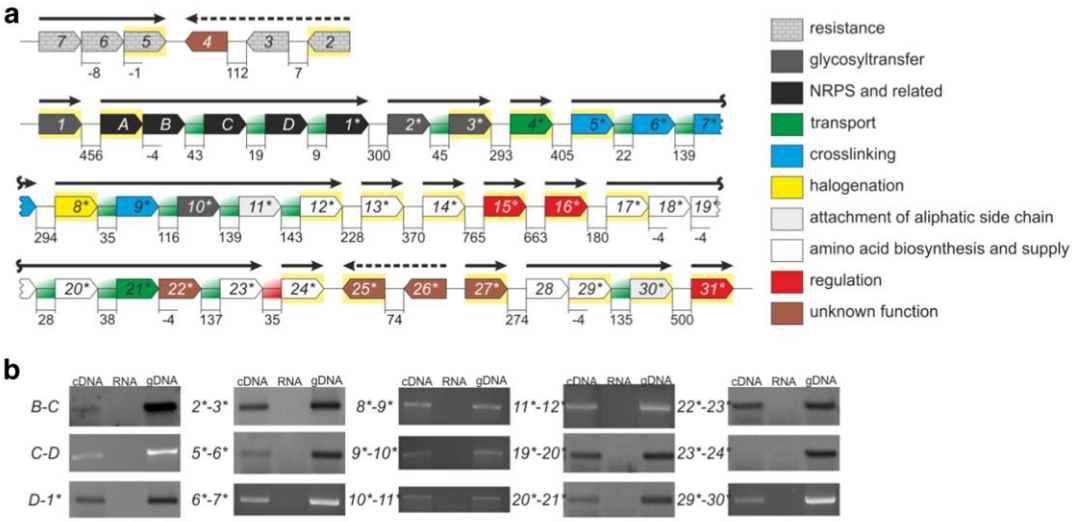


Fig. 2 Transcriptional organization of the teicoplanin biosynthetic gene cluster (*tei*). **a** The *tei* cluster is composed of nine polygenic and nine monogenic transcriptional units; these units are represented by the black arrows above the genes. Solid arrows represent the operons verified experimentally in this work or by other authors (Beltrametti et al. 2007); dashed arrows represent the operons deduced by sequence analysis. Sizes of the genes are not given to scale. Below and between the genes, the lengths of the intergenic regions in base pairs are reported. The expression of the genes highlighted in yellow was studied in this work (see Fig. 3). **b** Expression analysis of the 15 selected intergenic regions (with size ranging from 9 to 143 bp) from the *tei* cluster. For all the intergenic regions (shadowed in green in **a**), except the one between *tei 23*-tei24** (shadowed in red in **a**), amplicons were obtained, meaning that neighboring genes are co-expressed. The intergeneric region between *tei25** and *tei26** was not amplified since these genes were not expressed under teicoplanin production conditions (see "Results")

polygenic transcriptional units (Fig. 2). The latter are composed as follows: *tei7-6-5* (resistance), *tei4-3-2* (regulation of resistance), *teiA-B-C-D-1** (peptide synthesis), *tei2*-3** (glycosylation), *tei5*-6*-7** (cross-linking), *tei8*-9*-10*-11*-12** (tailoring reactions), *tei17*-18*-19*-20*-21*-22*-23** (amino acid supply), *tei26*-25** (unknown function), and *tei28*-29*-30** (amino acid supply).

Expression of the majority of *tei*-genes is Tei15*-Tei16* dependent

Previous studies in vitro showed that Tei15* is able to bind certain promoter regions in the *tei* BGC, particularly promoters of *teiA*, *tei2**, *tei17**, *tei27**, and *tei31** (Horbal et al. 2014). In this paper, Tei15* and Tei16* regulons were studied by following the transcription of the *tei* genes by means of semi quantitative RT-PCR in *A. teichomyceticus* Δ*tei15** and *A. teichomyceticus* Δ*tei16** knockout mutants in comparison to the wild type (Fig. 3). In these experiments, RNA was isolated at four time points (48, 72, 92, and 120 h of growth) from the strains cultivated in TM1 medium. Overall, the expression of 20 genes (*tei5*, *tei2*, *tei1*, *teiA*, *tei3**, *tei4**, *tei5**, *tei8**, *tei12**, *tei13**, *tei14**, *tei15**, *tei16**, *tei17**, *tei24**, *tei25**, *tei27**, *tei29**, *tei30**, *tei31**), selected to represent all the transcriptional units in the *tei* BGC (Fig. 2), was analyzed.

Based on the obtained results, we can conclude that transcription of the majority of genes in the wild type started already at the second day of cultivation, reaching maximum expression after 72–92 h of cultivation (Fig. 3). One hundred and twenty hours after inoculation, expression of the *tei*-genes tended to decrease. Consistently, the expression of *tei* genes in the wild-type strain forestalled teicoplanin production, which occurred under these cultivation conditions on the third day of growth and reached its maximum production after 120–144 h of cultivation (Taurino et al. 2011).

As anticipated above, *tei25**, *tei26**, and *tei27**—encoding a putative dioxygenase, a protein with unknown function and a putative siderophore-reductase, respectively—were not expressed neither in the wild type nor in both knockout mutants (Fig. 3p, q). Surprisingly, the putative regulatory gene *tei31** was also not expressed throughout the whole period of growth neither in *A. teichomyceticus* Δ*tei15**, *A. teichomyceticus* Δ*tei16**, nor in the wild type (Fig. 3t) and this unexpected result merits further investigations (see the following section).

By comparing the level of the *tei* gene expression between the wild type and the knocked out mutants, it was confirmed that the genes *tei5* and *tei2*, which represent the resistance operons (*tei7-6-5–vanHAX* and *tei2-3–vanRS*), were expressed constitutively in the wild type and their expression remained intact in *A. teichomyceticus* Δ*tei15** and *A. teichomyceticus* Δ*tei16** (Fig. 3a, b). Similarly, the expression of the *tei17** gene, initiating

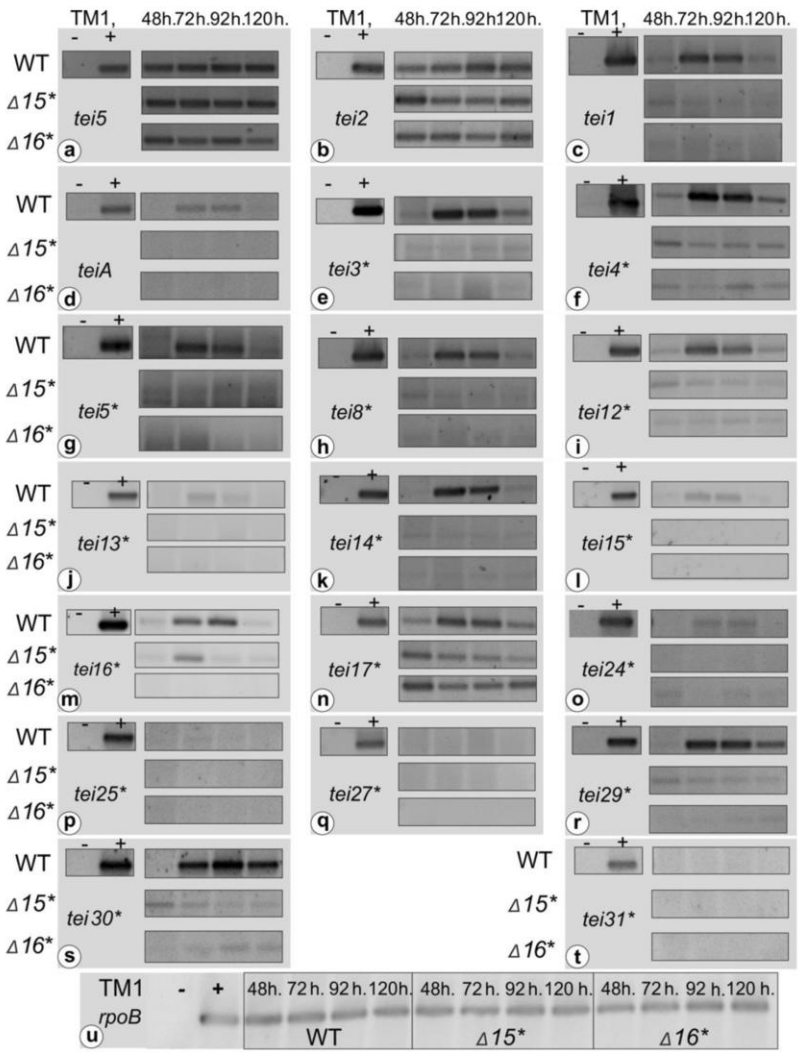


Fig. 3 Gene expression analysis of the 20 *tei* genes (from **a** to **t**) representing the main transcriptional units of the *tei* cluster in the *A. teichomyceticus* wild type and in the *A. teichomyceticus* Δ*tei15** and *A. teichomyceticus* Δ*tei16** knockout mutants. All strains were grown in parallel in TM1 teicoplanin production medium. PCR reactions were performed without a template (minus sign), with chromosome DNA as a template (plus sign), or with cDNA synthesized from RNA isolated from biomass at different time points (48 h, 72 h, 92 h, 120 h). As a control, the expression of the housekeeping gene *rpoB* (coding for the RNA polymerase β subunit) was analyzed (**u**)

the long polygenic operon *tei17*-18*-19*-20*-21*-22*-23**, was approximately the same in *A. teichomyceticus* Δ*tei15** and *A. teichomyceticus* Δ*tei16** mutants as in the wild type (Fig. 3n).

Expression of the other transcriptional units was different in the *A. teichomyceticus* Δ*tei15** and *A. teichomyceticus* Δ*tei16** mutants in comparison to the wild type (Fig. 3c–m, o, r, s). Transcription of the *tei1*, *tei3**, *tei4**, *tei5**, *tei8**, *tei12**, *tei14**, *tei24**, *tei29**, and *tei30** genes was not completely abolished in the knockout mutants but rather lowered to the basal level. Alternatively, gene expression of *teiA* and *tei13** was completely abolished in the knockout mutants (Fig. 3d, j). Additionally, we demonstrated that in *A. teichomyceticus* Δ*tei16*, *tei15** (coding for StrR-like cluster-situated regulator) was not expressed (Fig. 3l); at the same time, in *A. teichomyceticus* Δ*tei15**, transcript for *tei16** (encoding a LuxR-like regulator) was indeed present, although its level seemed to be lower than in *A. teichomyceticus* wild type at 92 h of growth (Fig. 3m).

*tei31** is not involved in regulation of teicoplanin biosynthesis

As reported above (Fig. 3t), we were unable to detect expression of the putative regulatory *tei31** gene (which presumably encodes a DNA-binding protein) in teicoplanin-producing cultures. According to amino acid sequence analysis, Tei31*

could be associated to a family of regulators with unknown function found in some species of *Micromonosporaceae* and *Pseudonocardiaceae* and annotated as AfsR-like transcriptional regulators (Fig. S2, ESM). In *S. coelicolor*, AfsR is a pleiotropic regulator of secondary metabolism (Lee et al. 2002). The closest Tei31* homologue—OLE27884—was found in *Actinobacteria sp.* 13_1_20CM_3_71_11 (soil metagenomic sample) sharing 39% of amino acid sequence identity with Tei31*. Other two close Tei31* homologues were WP_084297973 from *Actinoplanes friuliensis* and WP_052407872 from *Allokutzneria albata* that shared 37% and 33% of amino acid sequence identity to Tei31*, respectively (Fig. S2, ESM).

To investigate if Tei31* could play a role as a pleiotropic regulator in *A. teichomyceticus*, we generated a *tei31** knockout strain where the coding sequence of *tei31** was replaced with *aac3(IV)* cassette (*A. teichomyceticus tei31*::aac3(IV)*). We also created a *tei31**-overexpressing strain using the plasmid pSETtei31*, where *tei31** was cloned under the control of native promoter (*A. teichomyceticus* pSETtei31*$^{+}$). Both *A. teichomyceticus tei31*::aac3(IV)* and *A. teichomyceticus* pSETtei31*$^{+}$ exhibited no differences in teicoplanin production in comparison to the wild type (Fig. 4a), implying that *tei31** has no impact on teicoplanin biosynthesis. Moreover, knockout and overexpression of *tei31** had no impact on the morphology of *A. teichomyceticus* grown on different media (Fig. 4b). This ruled out the possibility that Tei31* is a pleiotropic regulator of morphogenesis, like AfsR. Indeed, further investigations by CD-search comparing the molecular structure of Tei31* with those of abovementioned close homologues (OLE27884, WP_084297973, WP_052407872) revealed that they share the presence of a helix-turn-helix (HTH) DNA-binding domain and of an ATP-ase domain (P-loop-NTPase domain), but Tei31* lacks the following tetratricopeptide-repeat (TPR) containing C-terminal region (Figs. S2 and S3, ESM). In *S. coelicolor* AfsR, this region is important for phosphorylation (and activation) by the cognate serine threonine kinase AfsK (Umeyama et al. 2002). CDART search revealed the occurrence of only few bacterial proteins with a molecular architecture similar to Tei31*, particularly transcriptional regulators of unknown function from *Nostocaceae*, for instance WP_012410910 from *Nostoc punctiforme*, that shares only 16% of amino acid sequence identity with Tei31* (Fig. S3, ESM).

Phylogeny of cluster-situated regulators from glycopeptide BGCs

Mining the sequences of the 21 glycopeptide BGCs nowadays available (listed in Table S1, ESM), we retrieved a total of 24 amino acid sequences of StrR-like transcriptional regulators (including Tei15*). These amino acid sequences were used to infer maximum-likelihood phylogenetic tree using the StrR sequence from streptomycin BGC as an outgroup (Fig. 5a). The topology of the final tree revealed two main clades of StrR regulators. The first included Bbr (from *bal* BGC), Dbv4 (from *dbv*), and regulators encoded in BGCs for vancomycin, decaplanin, avoparcin, nogabecin, ristocetin, and pekiskomycin as well as in the metagenome-originated TEG and VEG BGCs; we named it a *dpb* (dalbavancin–pekiskomycin–balhimycin)-clade (Fig. 5a). The second clade grouped Tei15*, Auk26 (UK-68,597 BGC), FegB

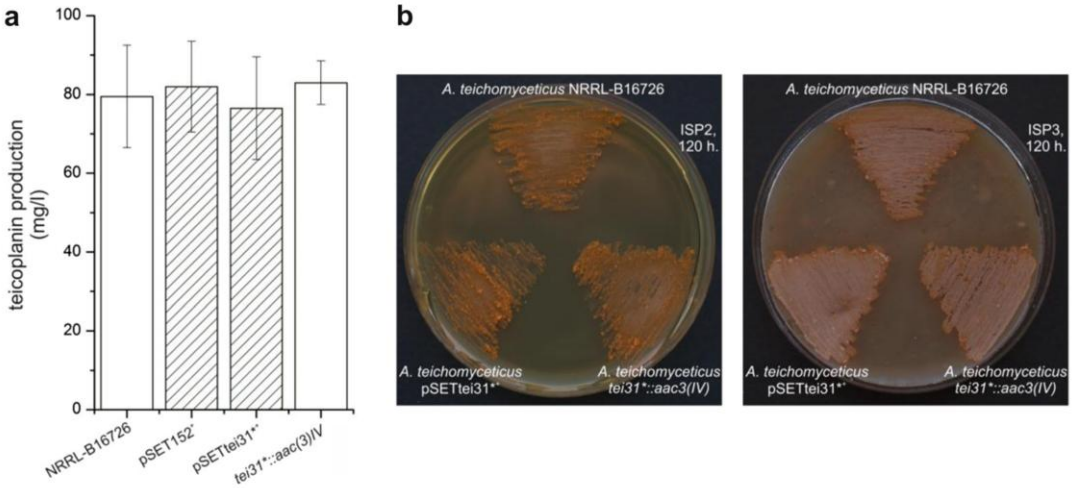


Fig. 4 Teicoplanin production in TM1 medium (**a**) and morphogenesis on ISP2 and ISP3 agar media (**b**) in the *A. teichomyceticus* strains with the knockout (*tei31*::aac3(IV)*) and overexpression (pSETtei31*$^{+}$) of the *tei31** gene in comparison to the wild type (NRRL-B16726) and the strain containing the empty vector (pSET152^{+})

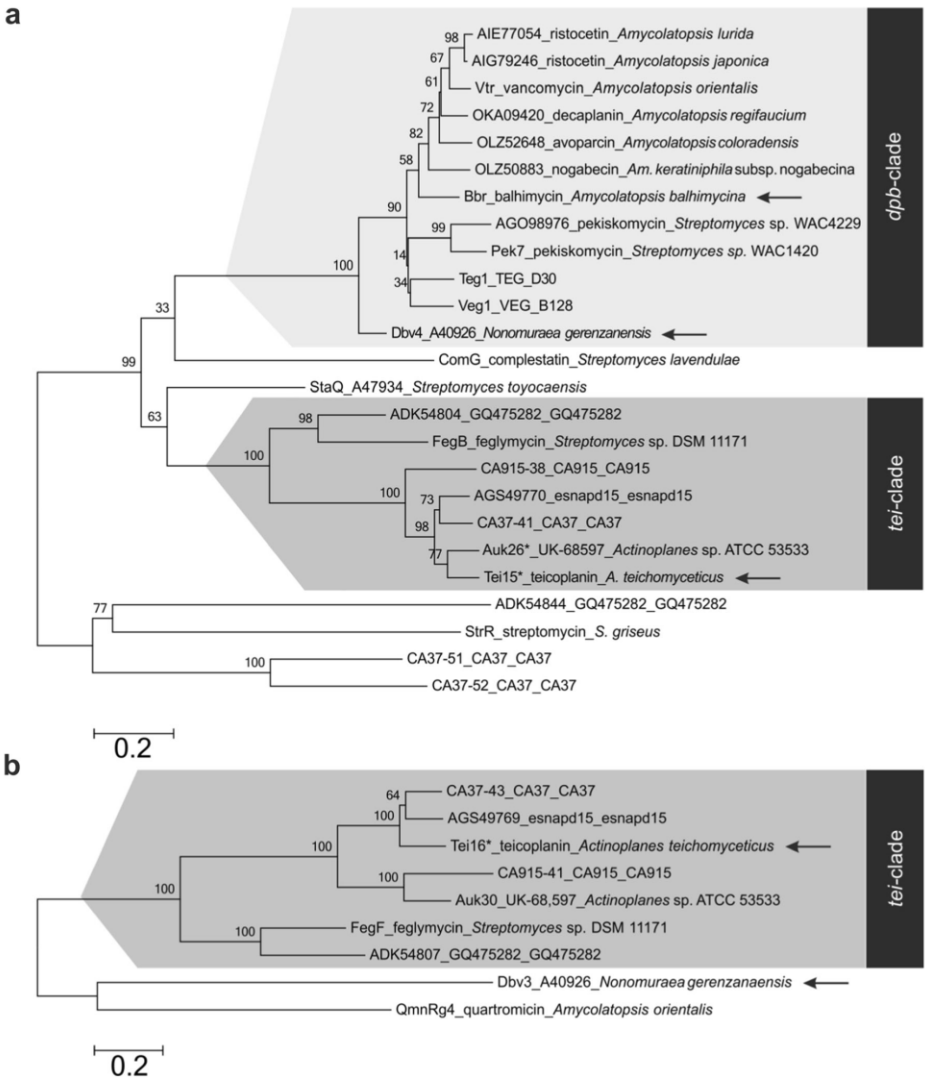


Fig. 5 Trees representing the phylogeny of the StrR (**a**) and LuxR (**b**) transcriptional regulators encoded in glycopeptide BGCs. Experimentally studied regulators of teicoplanin, balhimycin, and A40926 biosynthetic pathways are indicated with arrows. The final topology was derived from 1000 bootstraps

(feglymycin BGC), and related proteins encoded in the BGCs obtained from uncultured bacteria (Fig. 5a; Table S1, ESM); we referred to it as a *tei*-clade. Albeit the regulators encoded in *com* (complestatin) and *sta* (A47934) BGCs tended to group with *dpd*-clade and *tei*-clade respectively, the branch support values did not allow to include them into these clades. Structural similarity of A40926 and teicoplanin notwithstanding, Dbv4 and Tei15* appeared to be located in different clades, with Dbv4 being more related to Bbr. Interestingly, FegB (from the BGC for production of a linear nonglycosylated peptide; Gonsior et al. 2015) fell into *tei*--clade appearing to be closely related to Tei15*. Three StrR-like transcriptional regulators were found to be encoded in CA37 BGC from an uncultured bacterium, but only one (CA37-41) appeared to be related to Tei15*, and the other two outgrouped both *dpb*- and *tei*-clades together with StrR (Fig. 5a). Similarly, two StrR-like regulators were present in the GQ475282 BGC (derived from uncultured bacterium), but

only one (ADK54804) grouped to the *tei*-clade, and the other outgrouped both clades (Fig. 5a).

Next, we retrieved a total of only eight amino acid sequences of LuxR regulators encoded within the eight BGCs for A40926, teicoplanin, UK-68,597, CA37, CA915, GQ475282, esnapd15, and feglymycin. These amino acid sequences together with the sequence of QmnRg4, a LuxR regulator from quartromicin biosynthetic gene cluster (He et al. 2012) as an outgroup, were used to construct a maximum-likelihood phylogenetic tree (Fig. 5b). The final tree contained a well separated clade corresponding to the *tei*-clade of the tree of StrR regulators. Once again, the feglymycin BGC possessed a gene (*fegF*) encoding a putative LuxR regulator very similar to Tei16*. Finally, Dbv3 appeared not to be related to LuxR regulators from *tei*-clade, outgrouping it together with QmnRg4. Indeed, Dbv3 and Tei16* shared only 32% of amino acid sequence identity.

None *tei31** orthologue was found mining the sequences of the 21 glycopeptide BGCs (Table S1, ESM). It was only present in the *tei* BGC.

Discussion

A. teichomyceticus was inaccessible for genetic engineering until recently (Horbal et al. 2012; Horbal et al. 2013), when StrR- and LuxR-like transcriptional regulators (Tei15* and Tei16*, respectively) encoded within the *tei* cluster were shown to be crucial for teicoplanin biosynthesis. Moreover, they represent an important tool for improving teicoplanin production by the wild-type strain (Horbal et al. 2014). The third *tei* cluster-encoded regulator—Tei31*—was studied in this paper. Although Tei15* was reported to bind *tei31** promoter region in vitro (Horbal et al. 2014), our transcriptional analyses unequivocally indicated that *tei31** was not expressed in teicoplanin producing cultures (Fig. 6). Accordingly, neither knockout nor overexpression of *tei31** had impact on teicoplanin production or morphogenesis in *A. teichomyceticus*. Tei31* appears to be a nonfunctional AfsR-like transcriptional regulator, lacking TPRs and consequently being unable to form complexes with kinases as a prerequisite for phosphorylation (Umeyama et al. 2002). Actually, if the *tei31** does belong or not to *tei* BGC could be disputed: this gene was not expressed in teicoplanin producing conditions and has no homologues in the BGCs of other glycopeptides. The only reason to consider *tei31**, although probably nonfunctional, still as a part of the *tei* BCG remains the evidence that Tei15* binds in vitro to its promoter region (Horbal et al. 2014). To this purpose, it is interesting to note that other three genes (*tei25**, *tei26** and *tei27**) were not expressed in *A. teichomyceticus* grown under teicoplanin producing conditions, too (Fig. 6). These genes, according to (Li et al. 2004), have no obvious role in teicoplanin biosynthesis.

To our best knowledge, the presence of unexpressed genes is a particular feature of the *tei* BGC, being uncommon for other glycopeptide BGCs. Indeed, it was shown that all transcriptional units were expressed in *dbv* cluster under A40926 production conditions (Alduina et al. 2007). The majority of *bal* cluster genes were also expressed in balhimycin producing cultures, and the rest were reported to encode enzymes crucial for balhimycin biosynthesis (Shawky et al. 2007; Kilian et al. 2016).

Alternatively, the group of the genes conferring self-resistance to teicoplanin was expressed constitutively in the wild-type strain under teicoplanin production conditions. We in fact demonstrated that the transcription of both the *vanHAX*-like operon (genes *tei7-6-5*) and the *vanS/vanR*-like regulatory operon (*tei2-3*) remained at the same level in parent strain, *A. teichomyceticus* Δ*tei15** and *A. teichomyceticus* Δ*tei16** mutants, providing an important advance to the current model of self-resistance (Beltrametti et al. 2007; Binda et al. 2018). The constitutive expression of the putative two-component regulatory signal transduction system (*tei2-3*) and of the genes responsible for cell wall remodeling (*tei7-6-5*) confirms that self-resistance is not co-regulated with teicoplanin production. These data are in accord with recent results which demonstrated that *A. teichomyceticus* is phenotypically resistant to all glycopeptides (Binda et al. 2018), producing an intrinsically resistant cell wall where the peptidoglycan precursors invariably terminate with the D-Ala-D-Lac depsipeptide (Marcone et al. 2010). Differently from most of the known antibiotics, in glycopeptide-producing strains, the expression of biosynthetic genes and of self-resistance genes is not controlled by the same regulatory circuits. This is the case of the A40926 producer *Nonomuraea gerenzanensis* in which self-resistance is induced by its own product (Marcone et al. 2014; Binda et al. 2018) or of the balhimycin producer *Amycolatopsis balhimycina*, whose multilayered resistance mechanisms are not under the control of the cluster-situated positive regulatory gene *bbr* (Schäberle et al. 2011; Frasch et al. 2015). However, there may be anyhow a connection between glycopeptide production and self-resistance, which in many cases is still to be elucidated.

Also, the long polygenic operon from *tei17** to *tei23** (containing the genes responsible for the biosynthesis of dihydroxyphenylglycine) appeared almost equally expressed in the wild type as in *A. teichomyceticus* Δ*tei15** and *A. teichomyceticus* Δ*tei16** knockout mutants. This could indicate the importance of dihydroxyphenylglycine supply for teicoplanin biosynthesis. However, the expression of the transcriptional unit containing the genes responsible for dihydroxyphenylglycine biosynthesis in *N. gerenzanensis* (*dbv30-35*) was shown to be crucially dependent on Dbv4 (lo Grasso et al. 2015).

The above exceptions notwithstanding, most of the transcriptional units involved in teicoplanin assembling and

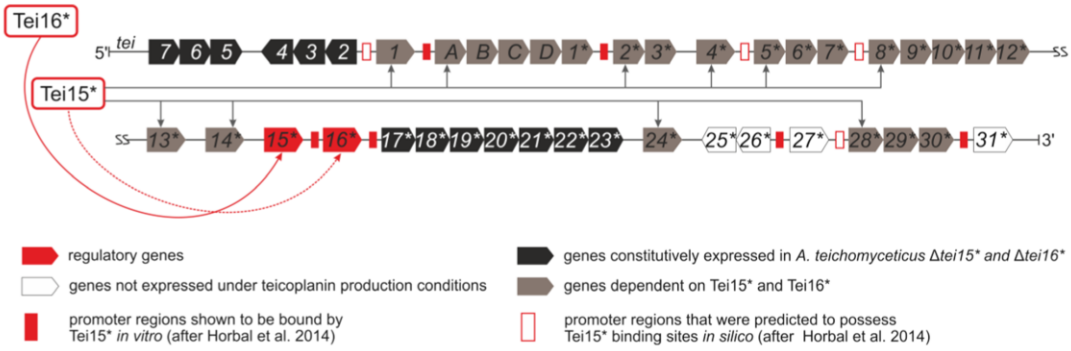


Fig. 6 Proposed scheme of the pathway-specific regulation of gene expression in *tei* cluster. The details are discussed in the main text

decoration (*tei1*, *teiABCD-1**, *tei2*-3**, *tei5*-7**, *tei8*-12**, *tei28*-29**, *tei30**) were under the positive control of Tei15* and Tei 16* (Fig. 6). In the knockout mutants, the expression of the most of these genes was decreased to basal level or in rare cases completely abolished. Thus, the Tei15* and Tei16* regulators in most cases did not initiate the *tei* gene transcription but rather significantly enhanced it. The obtained data are in agreement with the previous information about the DNA-binding properties of Tei15* (Horbal et al. 2014): in spite of some exceptions, transcriptional units possessing Tei15* binding sites (shown either in vitro or in silico) are differentially expressed in *A. teichomyceticus* Δ*tei15** in comparison to the wild type (Fig. 6).

Finally, we observed a hierarchical interplay between Tei15* and Tei16*: the latter appeared to be crucial for *tei15** expression. Thus, we suggest that both Tei15* and Tei16* regulate the biosynthesis of teicoplanin and are dependent on each other. Tei16* initiates the expression of *tei15** (probably in response to some kind of ligands like many other LuxR transcriptional regulators (Santos et al. 2012)) and Tei15* in its turn upregulates *tei16** expression (Fig. 6). Altogether, our data explain the dramatic increase in teicoplanin production observed earlier (Horbal et al. 2012; Horbal et al. 2014) in the *tei15**/*tei16**-overexpressing strains: by increasing the pool of one regulator, we consequently increased the pool of the other one and thus multiplied the positive effect. In conclusion, both regulators seem to activate the majority of teicoplanin biosynthesis genes in a synergistic manner (Fig. 6). One intriguing hypothesis is that *tei15** might be the single target of Tei16*, although the precise identification of Tei15* and Tei16* regulons requires further investigation.

The regulatory model that emerges from our work resembles the one of A40926 biosynthesis, where LuxR-type transcriptional factor Dbv3 governs the expression of *dbv4* (Io Grasso et al. 2015). However, Tei15* and Tei16* are distant from the Dbv4 and Dbv3. Dbv4 is more related to StrR-like regulatory genes involved in the regulation of the vancomycin-family of glycopeptides (balhimycin, vancomycin, and ristocetin) than to Tei15*. Genes *tei16** and *dbv3* seem to have a different evolutionary origin, and the *dbv3* gene appears as a singular example of convergent evolution. Interestingly, there is a remarkable similarity between the regulatory genes encoded in feglymycin BGC and the ones from *tei* cluster. This is quite surprising since feglymycin is a linear uncrosslinked, nonglycosylated peptide antibiotic produced by *Streptomyces* sp. DSM 11171 (Vértesy et al. 1999).

In conclusion, it seems that vancomycin family BGCs have a simpler organization with less transcriptional units, whose expression is regulated by a single transcriptional regulator belonging to StrR-family, closely related to Bbr (*dpb*-clade of StrR regulators). More complex glycopeptide BGCs consist of many transcriptional units (like *tei*). They are controlled by a regulatory gene pair encoding StrR- and LuxR-like regulatory proteins. The *dbv* cluster stands apart, encoding a Tei16* unrelated LuxR regulator Dbv3 in addition to a Bbr-related Dbv4. Further efforts in transcriptional analysis of glycopeptide BGCs and in genome/metagenome mining of novel sequences (like recently described *auk* and *pek* clusters, Thaker et al. 2013; Yim et al. 2014) will surely improve our current comprehension of the driving forces in the evolution of glycopeptide antibiotics. Since slight differences in glycopeptide biosynthesis determine relevant variations in their mode of action and resistance (Marcone et al. 2018; Binda et al. 2018), such studies might contribute to the future development of this clinically important class of antibiotics.

Funding This work was supported by grant Bg-46F from the Ministry of Education and Science of Ukraine to V. F., personal grants from CIB (Consorzio Interuniversitario per le Biotecnologie) and DAAD (German Academic Exchange Service) 57048249 to O.Y., and support from the DFG (SFB 766) to E.S.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

Alduina R, Piccolo LL, D'alia D, Ferraro C, Gunnarsson N, Donadio S, Puglia AM (2007) Phosphate-controlled regulator for the biosynthesis of the dalbavancin precursor A40926. J Bacteriol 189(22):8120–8129. https://doi.org/10.1128/JB.01247-07

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410

Armenia I, Marcone GL, Berini F, Orlandi VT, Pirrone C, Martegani E, Gornati R, Bernardini G, Marinelli F (2018) Magnetic nanoconjugated teicoplanin: a novel tool for bacterial infection site targeting. Front Microbiol 9:2270. https://doi.org/10.3389/fmicb.2018.02270

Bassetti D, Cruciani M (1990) Teicoplanin therapy in children: a review. Scand J Infect Dis Suppl 72:35–37

Beltrametti F, Consolandi A, Carrano L, Bagatini F, Rossi R, Leoni L, Zennaro E, Selva E, Marinelli F (2007) Resistance to glycopeptide antibiotics in the teicoplanin producer is mediated by van gene homologue expression directing the synthesis of a modified cell wall peptidoglycan. Antimicrob Agents Chemother 51(4):1135–1141. https://doi.org/10.1128/AAC.01071-06

Binda E, Marinelli F, Marcone GL (2014) Old and new glycopeptide antibiotics: action and resistance. Antibiotics (Basel) 3(4):572–594. https://doi.org/10.3390/antibiotics3040572

Binda E, Cappelletti P, Marinelli F, Marcone GL (2018) Specificity of induction of glycopeptide antibiotic resistance in the producing actinomycetes. Antibiotics (Basel) 7(2):36. https://doi.org/10.3390/antibiotics7020036

Frasch HJ, Kalan L, Kilian R, Martin T, Wright GD, Stegmann E (2015) Alternative pathway to a glycopeptide-resistant cell wall in the balhimycin producer *Amycolatopsis balhimycina*. ACS Infect Dis 1(6):243–252. https://doi.org/10.1021/acsinfecdis.5b00011

Fuchs A, Bielicki J, Mathur S, Sharland M, Van Den Anker JN (2018) Reviewing the WHO guidelines for antibiotic use for sepsis in neonates and children. Paediatr Int Child Health 38(sup1):S3–S15. https://doi.org/10.1080/20469047.2017.1408738

Geer LY, Domrachev M, Lipman DJ, Bryant SH (2002) CDART: protein homology by domain architecture. Genome Res 12(10):1619–1623

Gonsior M, Mühlenweg A, Tietzmann M, Rausch S, Poch A, Süssmuth RD (2015) Biosynthesis of the peptide antibiotic feglymycin by a linear nonribosomal peptide synthetase mechanism. Chembiochem 16(18):2610–2614. https://doi.org/10.1002/cbic.201500432

Gust B, Kieser T, Chater K (2002) PCR targeting system in *Streptomyces coelicolor* A3(2). The John Innes Foundation, Norwich

Ha H, Hwang Y, Choi S (2008) Application of conjugation using ϕC31 att/int system for *Actinoplanes teichomyceticus*, a producer of teicoplanin. Biotechnol Lett 30(7):1233–1238. https://doi.org/10.1007/s10529-008-9671-z

He HY, Pan HX, Wu LF, Zhang BB, Chai HB, Liu W, Tang GL (2012) Quartromicin biosynthesis: two alternative polyketide chains produced by one polyketide synthase assembly line. Chem Biol 19(10):1313–1323. https://doi.org/10.1016/j.chembiol.2012.07.024

Horbal L, Zaburannyy N, Ostash B, Shulga S, Fedorenko V (2012) Manipulating the regulatory genes for teicoplanin production in *Actinoplanes teichomyceticus*. World J Microbiol Biotechnol 28(5):2095–2100. https://doi.org/10.1007/s11274-012-1013-6

Horbal L, Kobylyanskyy A, Yushchuk O, Zaburannyi N, Luzhetskyy A, Ostash B, Marinelli F, Fedorenko V (2013) Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus*—producer of teicoplanin, drug of last defense. J Biotechnol 168(4):367–372. https://doi.org/10.1016/j.jbiotec.2013.10.018

Horbal L, Kobylyanskyy A, Truman AW, Zaburranyi N, Ostash B, Luzhetskyy A, Marinelli F, Fedorenko V (2014) The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. Appl Microbiol Biotechnol 98(22):9295–9309. https://doi.org/10.1007/s00253-014-5969-z

Hutchings MI, Hong HJ, Buttner MJ (2006) The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*. Mol Microbiol 59(3):923–935. https://doi.org/10.1111/j.1365-2958.2005.04953.x

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical *Streptomyces* genetics. John Innes Foundation, Norwich

Kilian R, Frasch HJ, Kulik A, Wohlleben W, Stegmann E (2016) The VanRS homologous two-component system VnlRSAb of the glycopeptide producer *Amycolatopsis balhimycina* activates transcription of the *vanHAXSc* genes in *Streptomyces coelicolor*, but not in *A. balhimycina*. Microb Drug Resist 22(6):499–509. https://doi.org/10.1089/mdr.2016.0128

Lee P, Umeyama T, Horinouchi S (2002) *afsS* is a target of AfsR, a transcriptional factor with ATPase activity that globally controls secondary metabolism in *Streptomyces coelicolor* A3(2). Mol Microbiol 43(6):1413–1430

Li TL, Huang F, Haydock SF, Mironenko T, Leadlay PF, Spencer JB (2004) Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase. Chem Biol 11(1):107–119. https://doi.org/10.1016/j.chembiol.2004.01.001

Liu G, Chater KF, Chandra G, Niu G, Tan H (2013) Molecular regulation of antibiotic biosynthesis in *Streptomyces*. Microbiol Mol Biol Rev 77(1):112–143. https://doi.org/10.1128/MMBR.00054-12

lo Grasso L, Maffioli S, Sosio M, Bibb M, Puglia AM, Alduina R (2015) Two master switch regulators trigger A40926 biosynthesis in *Nonomuraea* sp. strain ATCC 39727. J Bacteriol 197(15):2536–2544. https://doi.org/10.1128/JB.00262-15

Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res 32(Web Server issue):W327–W331

Marcone GL, Carrano L, Marinelli F, Beltrametti F (2010) Protoplast preparation and reversion to the normal filamentous growth in antibiotic-producing uncommon actinomycetes. J Antibiot (Tokyo) 63(2):83–88. https://doi.org/10.1038/ja.2009.127

Marcone GL, Binda E, Carrano L, Bibb M, Marinelli F (2014) Relationship between glycopeptide production and resistance in the actinomycete *Nonomuraea* sp. ATCC 39727. Antimicrob Agents Chemother 58(9):5191–5201. https://doi.org/10.1128/AAC.02626-14

Marcone GL, Binda E, Berini F, Marinelli F (2018) Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era. Biotechnol Adv 36(2):534–554. https://doi.org/10.1016/j.biotechadv.2018.02.009

Pfeifer V, Nicholson GJ, Ries J, Recktenwald J, Schefer AB, Shawky RM, Schröder J, Wohlleben W, Pelzer S (2001) A polyketide synthase in glycopeptide biosynthesis: the biosynthesis of the non-proteinogenic amino acid (s)-3,5-dihydroxyphenylglycine. J Biol Chem 276(42):38370–38377. https://doi.org/10.1074/jbc.M106580200

Puk O, Bischoff D, Kittel C, Pelzer S, Weist S, Stegmann E, Süssmuth RD, Wohlleben W (2004) Biosynthesis of chloro-β-hydroxytyrosine, a nonproteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics. J Bacteriol 186(18):6093–6100. https://doi.org/10.1128/JB.186.18.6093-6100.2004

Santos CL, Correia-Neves M, Moradas-Ferreira P, Mendes MV (2012) A walk into the LuxR regulators of *Actinobacteria*: phylogenomic distribution and functional diversity. PLoS One 7(10):e46758. https://doi.org/10.1371/journal.pone.0046758

Schäberle TF, Vollmer W, Frasch HJ, Hüttel S, Kulik A, Röttgen M, von Thaler AK, Wohlleben W, Stegmann E (2011) Self-resistance and cell wall composition in the glycopeptide producer *Amycolatopsis balhimycina*. Antimicrob Agents Chemother 55(9):4283–4289. https://doi.org/10.1128/AAC.01372-10

Shawky RM, Puk O, Wietzorrek A, Pelzer S, Takano E, Wohlleben W, Stegmann E (2007) The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains *bbr*, encoding a StrR-like pathway-specific regulator. J Mol Microbiol Biotechnol 13(1–3):76–88. https://doi.org/10.1159/000103599

Somma S, Gastaldo L, Corti A (1984) Teicoplanin, a new antibiotic from *Actinoplanes teichomyceticus* nov. sp. Antimicrob Agents Chemother 26(6):917–923

Sosio M, Kloosterman H, Bianchi A, de Vreugd P, Dijkhuizen L, Donadio S (2004) Organization of the teicoplanin gene cluster in *Actinoplanes teichomyceticus*. Microbiology 150(Pt 1):95–102. https://doi.org/10.1099/mic.0.26507-0

Spohn M, Kirchner N, Kulik A, Jochim A, Wolf F, Muenzer P, Borst O, Gross H, Wohlleben W, Stegmann E (2014) Overproduction of Ristomycin A by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. Antimicrob Agents Chemother 58(10):6185–6196. https://doi.org/10.1128/AAC.03512-14

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12):2725–2729. https://doi.org/10.1093/molbev/mst197

Taurino C, Frattini L, Marcone GL, Gastaldo L, Marinelli F (2011) *Actinoplanes teichomyceticus* ATCC 31121 as a cell factory for producing teicoplanin. Microb Cell Factories 10:82. https://doi.org/10.1186/1475-2859-10-82

Thaker MN, Wang W, Spanogiannopoulos P, Waglechner N, King AM, Medina R, Wright GD (2013) Identifying producers of antibacterial compounds by screening for antibiotic resistance. Nat Biotechnol 31(10):922–927. https://doi.org/10.1038/nbt.2685

Umeyama T, Lee P-C, Horinouchi S (2002) Protein serine/threonine kinases in signal transduction for secondary metabolism and morphogenesis in *Streptomyces*. Appl Microbiol Biotechnol 59(4–5):419–425. https://doi.org/10.1007/s00253-002-1045-1

Vértesy L, Aretz W, Knauf M, Markus A, Vogel M, Wink J (1999) Feglymycin, a novel inhibitor of the replication of the human immunodeficiency virus. Fermentation, isolation and structure elucidation. J Antibiot (Tokyo) 52(4):374–382

Yim G, Kalan L, Koteva K, Thaker MN, Waglechner N, Tang I, Wright GD (2014) Harnessing the synthetic capabilities of glycopeptide antibiotic tailoring enzymes: characterization of the UK-68,597 biosynthetic cluster. Chembiochem 15(17):2613–2623. https://doi.org/10.1002/cbic.201402179

Yushchuk O, Ostash B, Pham TH, Luzhetskyy A, Fedorenko V, Truman AW, Horbal L (2016) Characterization of the post-assembly line tailoring processes in teicoplanin biosynthesis. ACS Chem Biol 11(8):2254–2264. https://doi.org/10.1021/acschembio.6b00018

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Оригінал доступний за посиланням:
https://link.springer.com/article/10.1007/s00253-019-09789-w.
Електронні додатки до статті доступні за посиланням:
https://static-content.springer.com/esm/art%3A10.1007%2Fs00253-019-09789-w/MediaObjects/253_2019_9789_MOESM1_ESM.pdf.

Анотація: *Тейкопланін є глікопептидним антибіотиком, що продукується Actinoplanes teichomyceticus. Тейкопланін застосовується для лікування складних випадків інфекцій, у тому числі дитячих, спричинених Грам-позитивними патогенами. Існує постійний інтерес до вияснення генетичних механізмів, що детермінують продукцію тейкопланіну, оскільки це сприятиме створенню надпродуцентів тейкопланіну і відкриттю нових глікопептидів. В цій роботі ми досліджуємо транскрипційну організацію кластера біосинтетичних генів tei і роль розташованих у кластері регуляторних генів у контролі продукції тейкопланіну та стійкості до нього в A. teichomyceticus. Ми демонструємо, що кластер tei організований у дев'ять полігенних і дев'ять моногенних транскрипційних одиниць. Більшість біосинтетичних генів tei піддаються контролю StrR-подібним Tei15*, який, у свою чергу, регулюється Tei16*, регулятором типу LuxR. Експресія генів, що забезпечують стійкість до тейкопланіну в A. teichomyceticus, не корегулюється разом із біосинтезом антибіотика. Ген tei31*, який кодує ймовірний ДНК-зв'язуючий білок, не експресується в умовах продукції тейкопланіну, і не є необхідним для продукції. Нарешті, реконструкція філогенії регуляторів, кодованих глікопептидними кластерами, виявляє дві основні групи StrR-подібних регуляторів. Tei15* і близькі ортологи утворюють одну з цих клад; друга клада складається з ортологів Bbr і Dbv4, що керують біосинтезом баліміцину та тейкопланіноподібного A40926, відповідно. Крім того, Tei16* типу LuxR не пов'язаний з LuxR-подібним Dbv3, який контролює біосинтез A40926.*

The Journal of Antibiotics
https://doi.org/10.1038/s41429-019-0274-9

JARA Japan Antibiotics Research Association
The Society for Actinomycetes Japan

BRIEF COMMUNICATION

Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*

Oleksandr Yushchuk[1] · Vitalina Homoniuk[1] · Bohdan Ostash[1] · Flavia Marinelli[2] · Victor Fedorenko[1]

Received: 12 July 2019 / Revised: 25 November 2019 / Accepted: 20 December 2019
© The Author(s), under exclusive licence to the Japan Antibiotics Research Association 2020

Abstract

Actinoplanes teichomyceticus NRRL B-16726 is the only known producer of the clinically important glycopeptide antibiotic teicoplanin. The producing strain is highly self-resistant to teicoplanin. Although the biosynthesis of teicoplanin has been investigated, much of our understanding of self-resistance in the producing strain is based on the extrapolation of existing data about glycopeptide resistance (mediated by the expression of *vanRS-vanHAX* genes) in other actinomycetes and cocci. The organization of the operons carrying putative *van* orthologues in *A. teichomyceticus* differs from known precedents, further adding up to the uncertainty about teicoplanin self-resistance mechanisms. Here, we determined operon structure of the teicoplanin resistance genes in *A. teichomyceticus*. Although Tei15* is necessary to activate teicoplanin biosynthetic genes, the expression of *van* orthologues was shown to be independent of Tei15*. We further showed that *tei7* promoter driving the expression of *vanHAX* orthologues is dependent on Tei2 (VanR). Finally, we demonstrate the utility of the *tei2* promoter as a new tool to achieve strong constitutive expression in *A. teichomyceticus*.

Introduction

Cell wall is essential for bacterial cell shape and integrity. Consequently, peptidoglycan (PG) biosynthesis is an optimal target for antimicrobial agents. Many clinically successful antibiotics target PG biosynthesis [1]. Actinomycetes produce myriads of cell wall biosynthesis inhibitors. Among the latter, glycopeptide antibiotics (GPAs) occupy a special position since they are considered drugs of last resort against infections caused by multidrug resistant Gram-positive pathogens [2]. The GPA-producing actinomycetes are themselves Gram-positive bacteria and need to develop specific self-resistance mechanisms to prevent suicide during GPA production. Like pathogenic cocci, GPA resistance in most of the producing actinomycetes is mediated by the expression of *vanRS-vanHAX* genes [3, 4]. It is widely assumed that clinically relevant GPA-resistant pathogens have acquired resistance genes from the GPA-producing actinomycetes by horizontal gene transfer [4]. In resistant enterococci and staphylococci, VanS first detects vancomycin in the environment and then transduces the signal to VanR, which turns on *vanHAX* operon expression [5, 6]. *vanHAX*-encoded proteins remove terminal D-Ala-D-Ala from the pentapeptide stems of PG and install D-Ala-D-Lac, decreasing the affinity of vancomycin to PG precursors about 1000-fold [4]. Activity of teicoplanin against some vancomycin-resistant cocci relies on a still not completely elucidated mechanism, likely originating from the inability of the respective VanS protein to recognize teicoplanin and activate cell wall remodeling [7]. In addition, an exotic mechanism of self-resistance in the teicoplanin-like A40926 producer *Nonomuraea gerenzanensis* involves a single peptidase/carboxypeptidase (VanYn), which removes the last D-Ala from the pentapeptide stem [8]. In this work, we investigate the genetic basis for teicoplanin resistance in the producing *Actinoplanes teichomyceticus* NRRL B-16726. It is known that *vanHAX* orthologues in *A. teichomyceticus*—*tei7-6-5*—are part of the *tei* biosynthetic gene cluster (BGC), although, differently from the most of the biosynthetic *tei*

Supplementary information The online version of this article (https://doi.org/10.1038/s41429-019-0274-9) contains supplementary material, which is available to authorized users.

✉ Victor Fedorenko
viktor.fedorenko@lnu.edu.ua

[1] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 4 Hrushevskoho st., Lviv 79005, Ukraine

[2] Department of Biotechnology and Life Sciences, University of Insubria, Via J.H. Dunant 3, 21100 Varese, Italy

Published online: 17 January 2020

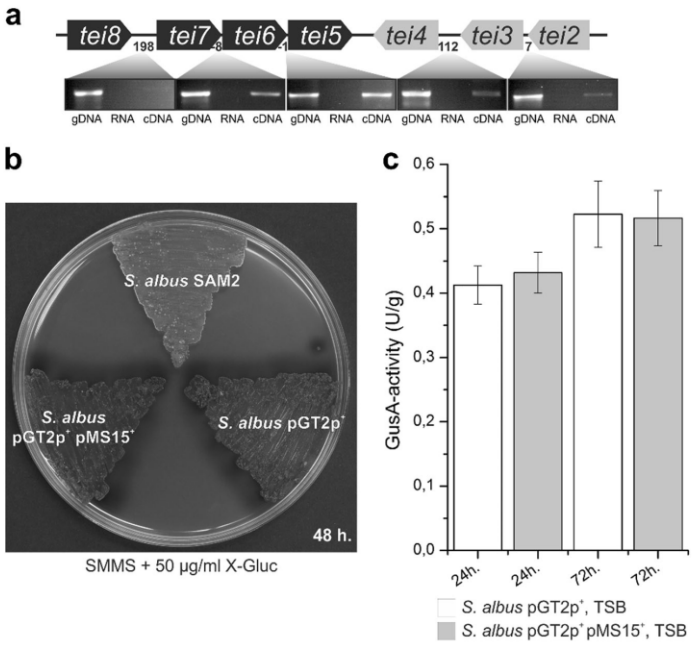


Fig. 1 **a** Transcriptional organization of teicoplanin self-resistance genes in *A. teichomyceticus*. Genes *tei7-6-5* are *vanHAX* orthologues, *tei8* codes for a putative MurF-like protein, *tei4* for a putative dihydrofolate reductase, and *tei3-2* are *vanS/vanR* orthologues. Sizes of the genes are not given to scale. Below and between the genes having the same orientation, the lengths of the intergenic regions are reported in bp. Expression analysis was performed for five selected intergenic regions (with sizes ranging from −8 (overlapping genes) to 198 bp) between *tei* resistance genes. For all the intergenic regions (shadowed in green), except the one between *tei8-tei7* (shadowed in red), amplicons were obtained, meaning that neighboring genes are co-expressed. The intergenic region between *tei5* and *tei4* was not studied since the neighboring genes are located on the opposite strands. **b** Qualitative GusA assays of *S. albus* SAM2 and its derivatives carrying pGT2p and pGT2p together with pMS15 grown on SMMS agar for 48 h. The medium was supplemented with 50 µg ml^{-1} of X-Gluc (5-Bromo-4-chloro-3-indolyl-β-D-glucuronide), and the chromogenic conversion of the latter to 5,5′-dibromo-4,4′-dichloro-indigo indicates the GusA activity of the strains. No significant differences between GusA activity of *S. albus* pGT2p$^+$ and *S. albus* pGT2p$^+$ pMS15$^+$ could be observed. **c** Results of the quantitative GusA assay (based on the conversion of *p*-nitrophenyl-β-D-glucuronide by cell lysates into *p*-nitrophenol with consequent spectrophotometric measurement of the latter). GusA activity in the cell lysates of *S. albus* pGT2p$^+$ and *S. albus* pGT2p$^+$ pMS15$^+$ grown in TSB medium was comparable at the early and late growth periods

genes, their transcription is constitutive in teicoplanin-producing cultures [9, 10]. The *vanRS* orthologues *tei2-3* are supposed to form one operon, whose expression was reported to be controlled by the pathway-specific regulator Tei15*, which indeed controls teicoplanin production [11]. On the contrary to this assumption, our recent results have shown that *tei2-3* expression remained intact in *A. teichomyceticus* Δ*tei15** [10]. Thanks to the recent advances in the genetic manipulation of the teicoplanin producer [10–12], hereby we refine the operon structure of *tei2-3-4* and *tei7-6-5* and show that *tei2* promoter (*tei2p*) is not regulated by Tei15*. At the same time, gene expression mediated by *tei7p* depends on *tei2*. *tei2* gene is strongly constitutively expressed due to the strength of *tei2p*, contributing to the high level of teicoplanin self-resistance reported in *A. teichomyceticus* [8, 9].

Materials and methods

A. teichomyceticus NRRL B-16726 and *Streptomyces albus* SAM2 derivatives were grown on MS medium [13] at 30 °C for routine maintenance. Qualitative chromogenic GusA assays were performed on SMMS [13], ISP2 and ISP3 [10] media; for quantitative GusA assays, *S. albus* strains were grown in TSB medium (Laboratorios Conda, Madrid, Spain), and *A. teichomyceticus* strains in TM1 [14]. Teicoplanin minimum inhibitory concentrations (MICs) for *A. teichomyceticus* were measured plating spores on ISP2 and ISP3 agar media [10], containing increasing concentrations of teicoplanin. The number of colonies was counted after 7 days of incubation at 30 °C. Qualitative and quantitative GusA assays were performed as described previously [12, 15]. *Escherichia coli* strains were grown at 37 °C in

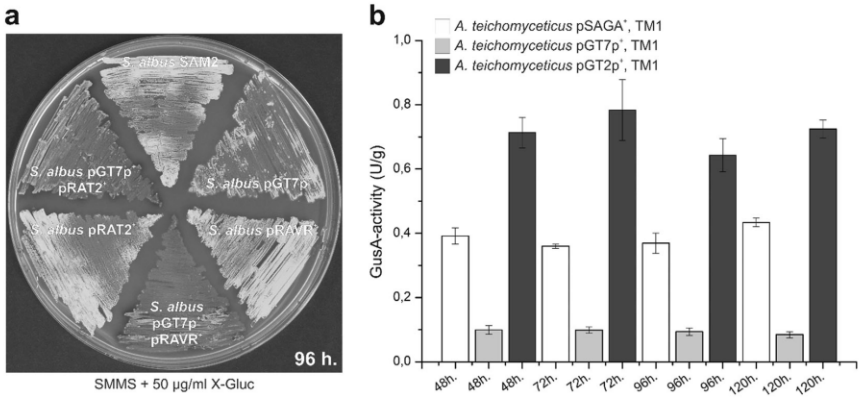


Fig. 2 **a** Qualitative assay of GusA activity of *S. albus* SAM2 derivatives, carrying pGT7p promoter-probe vector and pRAVR or pRAT2 vectors for the expression of *vanRsc* and *tei2*, respectively, and their combinations. Strain *S. albus* pGT7p$^+$ (as well as the control *S. albus* SAM2) was unable to convert X-Gluc. Both *S. albus* pGT7p$^+$ pRAT2$^+$ and *S. albus* pGT7p$^+$ pRAVR$^+$ supported the chromogenic conversion of X-Gluc, meaning that *tei7p* drives *gusA* expression only in the presence of either Tei2 or VanRsc. The strains were grown on SMMS agar for 96 h. **b** Comparison of the GusA activity of cell lysates of *A. teichomyceticus* strains carrying pSAGA (*aac(3)IVp*), pGT7p (*tei7p*) and pGT2p (*tei2p*) promoter-probe vectors grown in TM1 media at different time points. *tei2p*-mediated GusA activity in *A. teichomyceticus* pGT2p$^+$ was higher than in *A. teichomyceticus* pSAGA$^+$ and in *A. teichomyceticus* pGT7p$^+$

lysogeny broth/agar supplemented with apramycin sulfate (100 µg ml^{-1}), kanamycin sulfate (50 µg ml^{-1}), or hygromycin (100 µg ml^{-1}), when appropriate. All antibiotics used in this study were purchased from Sigma-Aldrich (St. Louis, MO, USA).

To construct pGT2p and pGT7p vectors, promoter regions of *tei2* and *tei7* were amplified using tei2p_F/R and tei7p_F/R (Supplementary Table 1) primer pairs, respectively. The resulting amplicons were cloned into *Kpn*I/*Spe*I sites of pGUS [15]. To generate pMS15, pSET152Atei15 vector [11] was digested with *Pvu*II, and the 1754 bp digestion product containing *aac(3)IVp-tei15** was ligated into pMS82 [16] *Eco*RV site. Finally, to generate pRAVR and pRAT2, *tei2* and *Streptomyces coelicolor vanR* (*vanRsc*) were amplified using T2_F/R and RSCO_F/R (Supplementary Table 1) primer pairs, respectively, and cloned into *Eco*RI/*Eco*RV sites of pSET152A [12]. The obtained pSAVR and pSAT2 vectors were *Pvu*II-digested, 1333 bp *aac(3)IVp-tei2* and 1345 bp *aac(3)IVp-vanRsc* fragments were cloned into *Eco*RV site of pRT801 [16]. This gave plasmids pRAT2 and pRAVR. All the constructs were verified by restriction mapping and sequencing. Plasmid maps are given in Supplementary Fig. 1. Plasmids were transferred conjugally from *E. coli* into *S. albus* SAM2 or *A. teichomyceticus* NRRL B-16726; transconjugants were selected for resistance to apramycin sulfate (50 µg ml^{-1}), hygromycin (50 µg ml^{-1}), or spectinomycin dihydrochloride (100 µg ml^{-1}). Transconjugants carrying pGUS derivatives were verified using Am_R/F primer pair to amplify *aac(3)IV* (Supplementary Table 1); the ones carrying plasmid-borne *tei2*, *vanRsc* and *tei15** were verified amplifying respective genes with T2_F/R, RSCO_F/R (Supplementary Table 1) and tei15*F/R [11] primer pairs.

For transcription analysis, total RNA was isolated from *A. teichomyceticus* NRRL B-16726, grown in TM1 for 96 h (teicoplanin production conditions [10, 14]); culture inoculation and cultivation conditions were as described previously [14]. Total RNA isolation, cDNA synthesis, and RT-PCR were performed as reported previously [10]. RT-PCR primers are listed in Supplementary Table 1.

Results and discussion

Similar to GPA-resistant enterococci and staphylococci [3], *tei7-6-5* (*vanHAX* orthologues) are co-transcribed in *A. teichomyceticus* [9]. However, no information about the transcription pattern of *tei2-3* has been reported. According to the published sequence of *tei* BGC [17], *tei4* gene is located on the same strand as *tei2-3*, 112 bp downstream of *tei3*. *tei4* codes for a putative dihydrofolate reductase with no obvious role in glycopeptide resistance [10]. Gene *tei8*, coding for a putative MurF-like protein, resides on the same strand as *tei7-6-5*, 198 bp upstream of *tei7*. In one of our recent works [10], we speculated that *tei8* is not co-transcribed with *tei7-6-5* and is probably not a part of *tei* cluster. However, this was not demonstrated experimentally. For elucidating the transcriptional organization and the operon structure of *vanRS*/*vanHAX* orthologues in *A. teichomyceticus*, we amplified the intergenic regions in between *tei8-7*, *tei7-6*, *tei6-5*, *tei2-3*, and *tei3-4* gene pairs, using semiquantitative RT-PCR. For all the intergenic

regions, except the one between *tei8-tei7*, amplicons were obtained, meaning that neighboring genes are co-expressed (Fig. 1a). Thus, genes *tei7-6-5* were confirmed to be co-transcribed, but not with *tei8*. Genes *tei2-3* were co-transcribed with *tei4*, forming a three-gene operon (Fig. 1a).

We previously [10] showed that the expression of *tei2-3* and *tei7-6-5* in *A. teichomyceticus* wild type was constitutive in teicoplanin production conditions, and it remained unaffected in *A. teichomyceticus* mutants where the key positive regulatory genes *tei15** and *tei16** were knocked out. These results are in contrast to the previously reported in vitro and in silico analyses of DNA-binding properties of Tei15*, which identified putative binding site of this regulatory protein within the *tei2p* region [11]. In this paper we attempted to solve this issue by reconstructing a putative Tei15* → *tei2-3* regulatory cascade in *S. albus* SAM2. This host carries neither GPA resistance genes nor teicoplanin-specific regulatory determinants and it is highly susceptible to teicoplanin and vancomycin [18]. A novel pGUS-based promoter-probe vector was constructed carrying the reporter gene *gusA* [15] under the control of *tei2p* (pGT2p) and transferred to *S. albus* SAM2. Next, we generated a recombinant *S. albus* strain carrying both pGT2p vector and pMS15 plasmid. The latter is pMS82-derivative carrying *tei15** under the control of the strong *aac(3)IV* promoter [11]. Glucuronidase activity in these recombinant strains was compared visually (Fig. 1b) and quantitatively (Fig. 1c), revealing no differences. Hence, our data support the idea that *tei2p* is Tei15*-independent. Following the same approach, Tei2 → *tei7p* regulatory cascade was reconstructed in *S. albus* SAM2, using plasmids pGT7p (carrying the reporter gene *gusA* under the control of *tei7p*) and pRAT2 (carrying *tei2* under the control of *aac(3)IVp*). Glucuronidase activity analysis in *S. albus* pGT7p$^+$ strain (Fig. 2a) revealed that *tei7p* is not functional in *S. albus* alone. However, introduction of pRAT2 vector into *S. albus* pGT7p$^+$ induced the GusA activity, indicating, that *tei7p* driven expression is dependent on *A. teichomyceticus tei2*. *tei2* and *tei3* products show high levels of homology to VanR and VanS of *S. coelicolor*, in which VanR was shown to act as a transcription activator even in the absence of the kinase activity of VanS [19]. Thus, in parallel, *vanR* from *S. coelicolor* M145 (*vanRsc*) was cloned under the *aac(3)IVp* generating pRAVR vector, which was transferred to *S. albus* pGT7p$^+$. As in the case of *S. albus* pGT7p$^+$ pRAT2$^+$, the detection of glucuronidase activity in *S. albus* pGT7p$^+$ pRAVR$^+$ (Fig. 2a) confirmed that VanRsc is able to cross-activate *tei7p*.

Since our data portray *tei2p* as a strong and constitutive promoter, we decided to compare promoter strengths of *tei2p* and *tei7p* with *aac(3)IVp*; the latter was previously revealed to be the most effective heterologous promoter for *A. teichomyceticus* [12]. For this purpose, pGT2p, pGT7p, and pSAGA [20] vectors were transferred into *A. teichomyceticus* NRRL B-16726 and GusA activity was measured in recombinant strains grown in teicoplanin production TM1 medium [14]. Indeed, *tei2p* activity was at least two times higher than that of *aac(3)IVp* (Fig. 2b) and *tei7p* was four times less active than *aac(3)IVp* (Fig. 2b).

Thus, the high teicoplanin resistance (20 μg ml^{-1} MIC) of vegetative *A. teichomyceticus* cells growing in liquid medium reported in [8, 9] is likely due to the *tei2p*-driven strong expression of *tei2-3*, which consequently leads to *tei7-6-5* constitutive expression. To corroborate this hypothesis, we tested the germination rates of *A. teichomyceticus* spores exposed to different concentrations of teicoplanin in solid media. On ISP2 agar medium, the MIC of teicoplanin for *A. teichomyceticus* germinating spores was confirmed to be around 20 μg ml^{-1} (Supplementary Fig. 2). Interestingly, on ISP3 agar medium, teicoplanin inhibited the germination of spores at much higher concentration (*ca.* 500 μg ml^{-1}) (Supplementary Fig. 2). This solid medium-dependent effect of teicoplanin needs further investigation, although we cannot rule out that it might be due to the physical adsorption/sequestration of the antibiotic on ISP3 insoluble components (ground oats). Notwithstanding these different responses to teicoplanin, when we tested the glucuronidase activity of *A. teichomyceticus* pGT2p$^+$ germinating spores on ISP2 and ISP3 agar plates added with the increasing concentrations of teicoplanin, we observed that the recombinant strain was able to convert X-Gluc under all conditions (Supplementary Fig. 3). After 36 h from plate inoculation, young colonies of *A. teichomyceticus* pGT2p$^+$ showed the same intensity of X-Gluc chromogenic conversion in the plates added with sub-inhibitory concentrations of teicoplanin as well as in those without the addition of teicoplanin (Supplementary Fig. 3). These results indicate that the *tei2p*-driven expression of the *gusA* reporter gene is independent from teicoplanin addition and that *tei2p* is active from the early phase of spore germination, pointing to the deeply entrenched resistance strategy built into *A. teichomyceticus* to counter the activity of its own product.

In summary (Supplementary Fig. 4), transcriptional organization of resistance genes in *A. teichomyceticus* was clarified, revealing that *tei2-3* (*vanRS* orthologues) are co-transcribed together with *tei4*, coding for a putative dihydrofolate reductase. It remains to verify whether Tei4 exerts some influence on teicoplanin self-resistance, considering that the latter might be a net result of multiple factors. The *tei2-3-4* genes are strongly expressed, independently of pathway-specific regulators and of the addition of different teicoplanin concentrations. Their expression leads to the expression of *tei7-6-5*, due to Tei2 binding to *tei7p*. Putative constitutive kinase activity of Tei3 [9] could also be an additional contributing factor to

the high level of self-resistance of *A. teichomyceticus*. Indeed, the role of Tei4 and the properties of Tei3 represent interesting topics worth of further investigations. Finally, we have identified *tei2p* as a constitutive and strong promoter, which could be useful for driving gene overexpression in *A. teichomyceticus* and maybe in other *Actinoplanes* species.

Acknowledgements We thank Prof. Margaret Smith for providing us with pMS82 vector. This work was supported by grant Bg-46F from the Ministry of Education and Science of Ukraine (to VF).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Jovetic S, Zhu Y, Marcone GL, Marinelli F, Tramper J. β-lactam and glycopeptide antibiotics: first and last line of defense? Trends Biotechnol. 2010;28:596–604.
2. Nicolaou KC, Boddy CN, Bräse S, Winssinger N. Chemistry, biology, and medicine of the glycopeptide antibiotics. Angew Chem Int Ed Engl. 1999;38:2096–152.
3. Marcone GL, Binda E, Berini F, Marinelli F. Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era. Biotechnol Adv. 2018;36:534–54.
4. Pootoolal J, Neu J, Wright GD. Glycopeptide antibiotic resistance. Annu Rev Pharm Toxicol. 2002;42:381–408.
5. Koteva K, Hong HJ, Wang XD, Nazi I, Hughes D, Naldrett MJ, et al. A vancomycin photoprobe identifies the histidine kinase VanSsc as a vancomycin receptor. Nat Chem Biol. 2010;6:327–9.
6. Hughes CS, Longo E, Phillips-Jones MK, Hussain R. Characterization of the selective binding of antibiotics vancomycin and teicoplanin by the VanS receptor regulating type A vancomycin resistance in the enterococci. Biochim Biophys Acta Gen Subj. 2017;1861:1951–9.
7. Kwun MJ, Hong HJ. The activity of glycopeptide antibiotics against resistant bacteria correlates with their ability to induce the resistance system. Antimicrob Agents Chemother. 2014;58:6306–10.
8. Binda E, Cappelletti P, Marinelli F, Marcone GL. Specificity of induction of glycopeptide antibiotic resistance in the producing actinomycetes. Antibiotics. 2018;7:36.
9. Beltrametti F, Consolandi A, Carrano L, Bagatin F, Rossi R, Leoni L, et al. Resistance to glycopeptide antibiotics in the teicoplanin producer is mediated by *van* gene homologue expression directing the synthesis of a modified cell wall peptidoglycan. Antimicrob Agents Chemother. 2007;51:1135–41.
10. Yushchuk O, Horbal L, Ostash B, Marinelli F, Wohlleben W, Stegmann E, et al. Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. Appl Microbiol Biotechnol. 2019;103:4089–102.
11. Horbal L, Kobylyanskyy A, Truman AW, Zaburranyi N, Ostash B, Luzhetskyy A, et al. The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. Appl Microbiol Biotechnol. 2014;98:9295–309.
12. Horbal L, Kobylyanskyy A, Yushchuk O, Zaburannyi N, Luzhetskyy A, Ostash B, et al. Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus*—producer of teicoplanin, drug of last defense. J Biotechnol.2013;168:367–72.
13. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. Practical *Sreptomyces* genetics. Norwich: John Innes Foundation; 2000.
14. Taurino C, Frattini L, Marcone GL, Gastaldo L, Marinelli F. *Actinoplanes teichomyceticus* ATCC 31121 as a cell factory for producing teicoplanin. Micro Cell Fact. 2011;10:82.
15. Myronovskyi M, Welle E, Fedorenko V, Luzhetskyy A. β-glucuronidase as a sensitive and versatile reporter in actinomycetes. Appl Environ Microbiol. 2011;77:5370–83.
16. Gregory MA, Till R, Smith MC. Integration site for *Streptomyces* phage φBT1 and development of site-specific integrating vectors. J Bacteriol. 2003;185:5320–3.
17. Li TL, Huang F, Haydock SF, Mironenko T, Leadlay PF, Spencer JB. Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase. Chem Biol. 2004;11:107–19.
18. Zaburannyi N, Rabyk M, Ostash B, Fedorenko V, Luzhetskyy A. Insights into naturally minimised *Streptomyces albus* J1074 genome. BMC Genomics. 2014;15:97.
19. Hutchings MI, Hong HJ, Buttner MJ. The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*. Mol Microbiol. 2006;59:923–35.
20. Koshla O, Yushchuk O, Ostash I, Dacyuk Y, Myronovskyi M, Jäger G, et al. Gene *miaA* for post-transcriptional modification of tRNA inferior is important for morphological and metabolic differentiation in *Streptomyces*. Mol Microbiol.2019;112:249–65.



SPRINGER NATURE

Оригінал доступний за посиланням:
https://www.nature.com/articles/s41429-019-0274-9.
Електронні додатки до статті доступні за посиланням:
https://static-content.springer.com/esm/art%3A10.1038%2Fs41429-019-0274-9/MediaObjects/41429_2019_274_MOESM1_ESM.docx.

Анотація: *Actinoplanes teichomyceticus NRRL B-16726 є єдиним відомим продуцентом клінічно важливого глікопептидного антибіотика тейкопланіну. Штам-продуцент має високу стійкість до тейкопланіну. Незважаючи на те, що біосинтез тейкопланіну досліджувався, більша частина нашого розуміння механізмів стійкості у штамі-продуценті базується на екстраполяції існуючих даних про стійкість до глікопептидів (опосередковану експресією генів vanRS-vanHAX) в інших актиноміцетах і коках. Організація оперонів, які несуть передбачувані ортологи van в A. teichomyceticus, відрізняється від відомих прецедентів, що ще більше додає невизначеності щодо механізмів авторезистентності до тейкопланіну. В цій роботі ми визначили оперонну структуру генів стійкості до тейкопланіну в A. teichomyceticus. Хоча Tei15* необхідний для активації генів біосинтезу тейкопланіну, було показано, що експресія ортологів van не залежить від Tei15*. Крім того, ми показали, що промотор tei7, який керує експресією ортологів vanHAX, залежить від Tei2 (VanR). Нарешті, ми демонструємо корисність промотора tei2 як нового інструменту для досягнення сильної конститутивної експресії в A. teichomyceticus.*

International Journal of
Molecular Sciences

MDPI

Article

Heterologous Expression Reveals Ancient Properties of Tei3—A VanS Ortholog from the Teicoplanin Producer *Actinoplanes teichomyceticus*

Oleksandr Yushchuk [1,2], Kseniia Zhukrovska [2], Bohdan Ostash [2], Victor Fedorenko [2] and Flavia Marinelli [1,*]

[1] Department of Biotechnology and Life Sciences, University of Insubria, 21100 Varese, Italy
[2] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 79005 Lviv, Ukraine
* Correspondence: flavia.marinelli@uninsubria.it; Tel.: +39-0332-42-1546

Abstract: Glycopeptide antibiotics (GPAs) are among the most clinically successful antimicrobials. GPAs inhibit cell-wall biosynthesis in Gram-positive bacteria via binding to lipid II. Natural GPAs are produced by various actinobacteria. Being themselves Gram-positives, the GPA producers evolved sophisticated mechanisms of self-resistance to avoid suicide during antibiotic production. These self-resistance genes are considered the primary source of GPA resistance genes actually spreading among pathogenic enterococci and staphylococci. The GPA-resistance mechanism in *Actinoplanes teichomyceticus*—the producer of the last-resort-drug teicoplanin—has been intensively studied in recent years, posing relevant questions about the role of Tei3 sensor histidine kinase. In the current work, the molecular properties of Tei3 were investigated. The setup of a GPA-responsive assay system in the model *Streptomyces coelicolor* allowed us to demonstrate that Tei3 functions as a non-inducible kinase, conferring high levels of GPA resistance in *A. teichomyceticus*. The expression of different truncated versions of *tei3* in *S. coelicolor* indicated that both the transmembrane helices of Tei3 are crucial for proper functioning. Finally, a hybrid gene was constructed, coding for a chimera protein combining the Tei3 sensor domain with the kinase domain of VanS, with the latter being the inducible Tei3 ortholog from *S. coelicolor*. Surprisingly, such a chimera did not respond to teicoplanin, but indeed to the related GPA A40926. Coupling these experimental results with a further in silico analysis, a novel scenario on GPA-resistance and biosynthetic genes co-evolution in *A. teichomyceticus* was hereby proposed.

Keywords: *Actinoplanes teichomyceticus*; VanS; glycopeptide antibiotic resistance; teicoplanin; A40926

check for updates

Citation: Yushchuk, O.; Zhukrovska, K.; Ostash, B.; Fedorenko, V.; Marinelli, F. Heterologous Expression Reveals Ancient Properties of Tei3—A VanS Ortholog from the Teicoplanin Producer *Actinoplanes teichomyceticus*. *Int. J. Mol. Sci.* **2022**, *23*, 15713. https://doi.org/10.3390/ijms232415713

Academic Editor: Michał Arabski

Received: 7 November 2022
Accepted: 6 December 2022
Published: 11 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Glycopeptide antibiotics (GPAs) [1], particularly dalbaheptides [2], belong to a clinically relevant group of natural compounds produced by soil-dwelling and mycelium-forming Gram-positive actinobacteria, traditionally named actinomycetes [3,4]. As with other natural compounds, genes responsible for the biosynthesis of GPAs are grouped in large biosynthetic gene clusters (BGCs) [5]. GPAs are successfully utilized to treat severe infections caused by multi-drug-resistant (MDR) staphylococcal and enterococcal strains. Two natural and three semisynthetic GPAs are currently approved for clinical use [3,6,7]. Natural first-generation GPAs are vancomycin and teicoplanin, produced by the actinobacteria *Amycolatopsis orientalis* (various strains) [8] and *Actinoplanes teichomyceticus* ATCC 31121 [9], respectively. Vancomycin was introduced in clinics first (in 1958), followed by teicoplanin (1988) in Europe and then in Japan (1998) [3]. Semisynthetic second-generation GPAs include telavancin—a vancomycin derivative [10], oritavancin—a derivative of the natural GPA chloroeremomycin [11], and dalbavancin—a derivative of the other natural GPA, A40926 [12].

Although GPAs of both generations are still clinically useful, resistant pathogens are inevitably emerging [13–16]. Peculiarly, resistance mechanisms in GPA producers and pathogens are very similar [17]. It is likely that pathogens acquired GPA resistance genes from GPA producers, or more generally from non-producing actinobacteria, where GPA resistance determinants are abundant [18,19]. GPAs interact with D-alanyl-D-alanine (D-Ala-D-Ala) termini of the nascent peptidoglycan (PG), impeding upstream transpeptidation and transglycosylation reactions and consequently blocking cell wall biosynthesis [20]. The resistance mechanisms shared between GPA producers and pathogens lead to the modification of the D-Ala-D-Ala termini, reducing the affinity of GPAs for their target. D-Ala-D-Ala termini in PG precursors are replaced by D-Ala-D-lactate (D-Lac) [17,21] or truncated to a single D-Ala residue [22–24]. The first mechanism requires a three-gene operon—*vanHAX*, where *vanH* codes for a D-Lac dehydrogenase responsible for generating a D-Lac pool, *vanX*—for a D,D-dipeptidase whose role is reducing the intracellular pool of D-Ala-D-Ala [25,26], and *vanA*—for a D-Ala-D-Lac ligase [21,27]. The second involves a single D,D-carboxypeptidase named VanY, which trims the terminal D-Ala residue. In both cases, PG precursors still undergo transpeptidation reactions, but this process is not affected by GPAs [22–24]. The expression of *vanY* or *vanHAX* might be constitutive or inducible. The latter requires a two-component regulatory system consisting of a sensor histidine kinase (SHK) and a transcriptional response regulator (RR), also known as VanS and VanR, respectively [28–31]. VanS can specifically recognize extracellular GPAs and phosphorylate VanR, which, in turn, activates the expression of other *van* genes [29–33].

Actinobacterial GPA producers such as *A. teichomyceticus*, *Amycolatopsis balhimycina* DSM 5908 (balhimycin producer), *Nonomuraea gerenzanensis* ATCC 39727 (A40926 producer), *Streptomyces toyocaensis* NRRL 15009 (A47934 producer), as well as the GPA non-producing *Streptomyces coelicolor* A3(2) represent well-studied models of GPA resistance mechanisms (recently reviewed in [34]), although certain aspects are not fully elucidated yet and merit further investigation. In this work, the properties of Tei3—the VanS ortholog from the teicoplanin producer *A. teichomyceticus*—were investigated in vivo. Notably, *A. teichomyceticus* is resistant to very high concentrations of teicoplanin, likely due to the strong constitutive expression of the *vanHAX* orthologs—*tei7-6-5*, which was previously reported during the transcription analyses of the *tei* BGC [35–37]. It was speculated that Tei3 acts as a constitutive SHK [35]. We aimed to experimentally investigate whether Tei3 really acts as a constitutive SHK and if any GPA is eventually recognized by its sensor domain (SD). Conducting a series of experiments in *S. coelicolor* as a heterologous host, it was found out that Tei3 does not require any ligand to function. Strikingly, the SD of Tei3 appeared to recognize not teicoplanin but the structurally related GPA A40926. Finally, reconstruction of the phylogeny of glycosyltransferases (GTFs) coded within GPA BGCs allowed us to build a possible scenario explaining the obtained results.

2. Results

The main goals of this work were to test if Tei3 acts as a non-inducible phosphorylase and to investigate which ligand could be sensed by Tei3. Considering *A. teichomyceticus* a challenging microorganism for gene-engineering manipulations, a series of genetic experiments were designed and performed in heterologous hosts such as *S. coelicolor* M512 [38] and *S. coelicolor* J3200 [31]. As anticipated in the introduction, *S. coelicolor* has a full set of *van* genes conferring vancomycin resistance, although it does not produce any GPAs [39]. *S. coelicolor* M512 does not produce the pigmented antibiotics actinorhodin and undecylprodigiosin [38], permitting the use of a β-glucuronidase (GusA)-based reporter assay, as described below. The second strain—J3200—is the Δ$vanS_{Sc}$ mutant, in which the host *vanS* gene was knocked out [31]. Thus, the experimental steps described below aimed: (1) to test if the heterologous Tei3 can cross-phosphorylate the host $VanR_{Sc}$; (2) to build a *gusA*-based reporter *S. coelicolor* strain, able to convert X-Gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronide) to 5,5′-dibromo-4,4′-dichloro-indigo following GPA induction; (3) to utilize the created reporter strain to show if Tei3 needs the presence of GPAs or not; (4) to

replace the SD of the host $VanS_{Sc}$ with its counterpart from Tei3 and test which GPAs may act as ligands for Tei3 SD.

2.1. Heterologous Expression of Tei3 SHK Leads to Teicoplanin and A40926 Resistance in S. coelicolor M512

VanRS-like two-component regulatory pairs are conserved enough among actinobacteria to show a certain degree of cross-talking [33,40]. For example, it was shown that $VanR_{Sc}$ (coming from *S. coelicolor*) could be phosphorylated in vivo by $VanS_{St}$ (from A47934 producer *S. toyocaensis*), but not vice versa, implying that $VanR_{Sc}$ is accessible for non-cognate SHKs [33]. In this case, SHKs and RRs both came from *Streptomyces* spp. Instead, it was necessary to check if Tei3 SHK from *A. teichomyceticus* (order *Micromonosporales*) could phosphorylate $VanR_{Sc}$. Overall, Tei2 and Tei3 share a high percentage of aa sequence identity with $VanR_{Sc}$ and $VanS_{Sc}$: 91% and 67%, respectively. Tei3 and $VanS_{Sc}$ are collinear, sharing highly similar transmembrane helix (TMHs) regions and conserved putative autophosphorylation sites (Figure 1a). The most divergent region between Tei3 and $VanS_{Sc}$ is the extracytoplasmic sensory loop (ESL), implying that quite different ligands should be recognized by the two proteins (Figure 1a). At the same time, Tei2 and $VanR_{Sc}$ were almost identical (Figure 1b). Thus, it seems plausible that Tei3 would be able to phosphorylate $VanR_{Sc}$ in vivo.

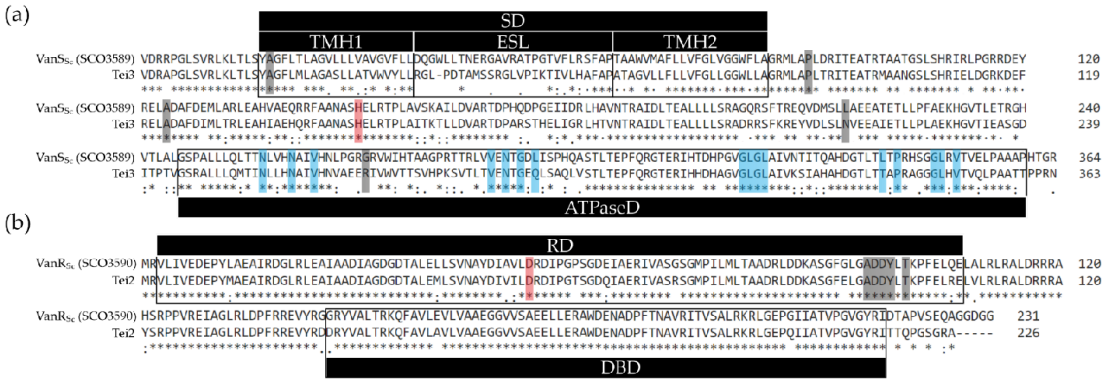


Figure 1. Clustal Omega [41] pairwise alignments of aa sequences of Tei3 and $VanS_{Sc}$ (**a**) as well as of Tei2 and $VanR_{Sc}$ (**b**). Domains and sequence features were annotated according to [31] and CD-Search tool [42]; SD—sensor domain, TMH1/2—first and second transmembrane α-helices, ESL—extracytoplasmic sensory loop, ATPaseD—ATPase domain, RD—response domain, DBD—DNA binding domain. In (**a**): mutation sites that were shown to impair $VanS_{Sc}$ function [31] are highlighted in gray; the autophosphorylation site is highlighted in red; conserved ATP-binding sites are in blue. In (**b**): residues putatively involved in dimerization are highlighted in gray, while the phosphorylation site is in red.

To test if Tei3 can phosphorylate the non-cognate RR—$VanR_{Sc}$, *tei3* was cloned into the pSET152A vector, giving pGP101, and then transferred into *S. coelicolor* M512. The obtained recombinant strain—*S. coelicolor* M101—gained teicoplanin and A40926 resistance, in addition to the endogenous vancomycin resistance (Figure 2). Consequently, it could be concluded that Tei3 phosphorylates $VanR_{Sc}$ in vivo. However, this experiment did not clarify whether Tei3 acts as a constitutive phosphorylase, considering that teicoplanin and A40926 added to the plates could serve as its activators.

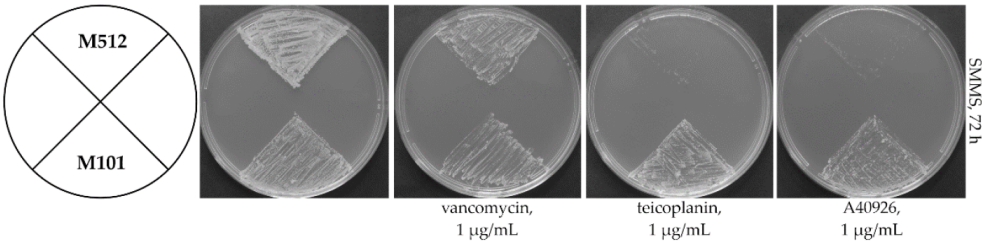


Figure 2. Expression of *tei3* led to teicoplanin and A40926 resistance in *S. coelicolor* M512. A number of 10^6 spores of *S. coelicolor* M512 and M101 were inoculated at each sector of SMMS agar plates added with vancomycin, teicoplanin, and A40926. Plates were examined after 72 h of incubation. *S. coelicolor* M512 was resistant to vancomycin and sensitive to teicoplanin and A40926, as previously reported [43], whereas M101 was resistant to vancomycin, A40926, and teicoplanin.

2.2. *Tei3 Acts as a Non-Inducible Phosphorylase*

Expression of *tei2-3-4* and *tei7-6-5* operons remains stable throughout the life cycle of *A. teichomyceticus*, independently from the teicoplanin concentration [35,37], pathway-specific regulation [36], and growth phase [37]. Previous papers speculated about the need of a strong constitutive promoter driving *tei2-3-4* expression [37] or of the presence of specific mutations in Tei3 that might explain its function as a constitutive phosphorylase, working independently from the presence of any extracellular GPA [35]. Protein sequence comparison indicated that Tei3 carries single aa substitutions in two sites where analogous mutations transform the vancomycin-inducible $VanS_{Sc}$ into a constitutive phosphorylase [35]. Specifically, these are L216P and G271V substitutions [31]. Homologous sites in Tei3 are N215 and R270 (Figure 1a). However, there are multiple other sites within the putative ATPase domain (ATPaseD) of Tei3, which significantly diverged from $VanS_{Sc}$ (see Figure 1a).

To experimentally verify the GPA inducibility of Tei3, a bioassay responding to the inducers of *van* genes was first developed in *S. coelicolor*. Other authors previously described a reporter *S. coelicolor* strain, where the endogenous *vanJ* promoter (*vanJp*) was cloned into a multicopy plasmid, fused with the kanamycin/neomycin resistance gene—*neo*—in a way that induction of *vanJp* by vancomycin conferred resistance to both neomycin and kanamycin [39]. Using this experience, either *vanJp* or the other endogenous vancomycin-responsive *S. coelicolor vanHAX* promoter (*vanHp*) was fused with the *gusA* gene, coding for a β-glucuronidase in the pGUS chassis [44]. *vanJp* or *vanHp* activation by GPAs in the reporter strain should activate the chromogenic conversion of the X-Gluc substrate into the green-colored 5,5′-dibromo-4,4′-dichloro-indigo [44]. Thus, plasmids pGHp (carrying *vanHp-gusA*) and pGJp (carrying *vanJp-gusA*) were transferred to *S. coelicolor* M512 by means of intergeneric conjugation with *Escherichia coli* ET12567 pUZ8002$^+$. The inducibility of the two generated reporter strains *S. coelicolor* pGHp$^+$ and pGJp$^+$ was first tested in liquid medium and then in solid plates. When vancomycin was added at 10 μg/mL to 50 h old cultures in TSB liquid medium, the basal glucuronidase activity of the mycelia was increased by, ca., twenty- and forty-fold, in *S. coelicolor* pGHp$^+$ and pGJp$^+$, respectively (Figure S1), indicating that *vanJp* seems more responsive to vancomycin than *vanHp*. In solid media containing 25 μg/mL of X-Gluc, the presence of vancomycin induced *vanJp*-mediated glucuronidase activity in recombinant strains, yielding green halos around the Whatman discs soaked in antibiotic solution (Figures 3 and S2). *S. coelicolor* pGJp$^+$ acted very well as a reporter, giving a detectable chromogenic conversion in response to very low concentrations of vancomycin (250 ng, Figure S2), whereas *S. coelicolor* pGHp$^+$ resulted as much less responsive to vancomycin (data not shown) and, thus, it was not further used.

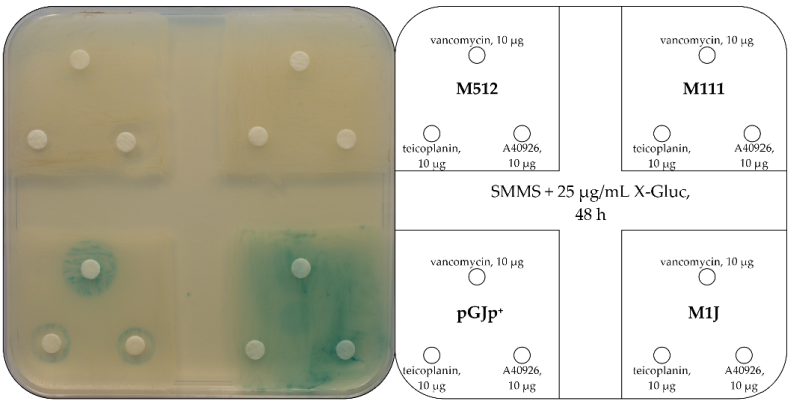


Figure 3. *gusA*-based GPA-inducible reporter assay revealed a constitutive phosphorylase activity of Tei3. *S. coelicolor* strains pGJp$^+$ and M1J were inoculated on SMMS agar added with 25 µg/mL of X-Gluc along with their *gusA*-control strains, i.e., the teicoplanin/A40926-sensitive M512 and the teicoplanin/A40926-resistant M111. Notably, in pGJp$^+$, the chromogenic conversion of X-Gluc (observed as green halos) occurred only upon GPA induction, while the background of M1J remained completely green, implying that Tei3 did not require GPA induction of its activity. A number of 10^7 spores of each strain were used for inoculation and the plate was examined after 48 h of incubation.

The next step was transferring Tei3 into *S. coelicolor* pGJp$^+$ to test if its activity is inducible by GPAs. As either pGJp or pGP101 is a distant derivative of pSET152 [45], both plasmids use the φC31 *attB* site for integration and are not compatible. To solve this issue, the pRT801 plasmid [46] was used to create a φBT1-based derivative of pGP101, named pGP111. The M512 derivative carrying pGP111 (carrying *tei3*) was named M111, while the strain carrying both pGP111 and pGJp was named M1J. It was expected that if Tei3 functions as a constitutive phosphorylase, M1J would express a constitutive glucuronidase activity in the presence of X-Gluc, not depending on the presence or absence of GPAs. This assumption was correct as M1J converted X-Gluc independently from the presence of any GPA (Figure 3), confirming in vivo that Tei3 functions as a constitutive phosphorylase. As a control, it was shown that no induction was observed in those strains not carrying *vanJp-gusA* (M512 and M111), whereas the GPA induction was observed in *S. coelicolor* pGJp$^+$ due to the action of the endogenous VanS$_{Sc}$.

2.3. Loss of Extracytoplasmic Sensory Loop and Transmembrane Helices Renders Tei3 and VanS$_{Sc}$ Nonfunctional

Considering the non-inducible properties of Tei3, we were wondering whether TMHs and ESL are still necessary for its function, or whether ATPaseD would be the only required domain. To answer this question, a series of plasmids was created carrying truncated versions of *tei3* or of *vanS$_{Sc}$* lacking (i) TMH1 and ESL (*tei3′*, *vanS$_{Sc}$′*); (ii) TMH1, ESL, and TMH2 (*tei3″*, *vanS$_{Sc}$″*) (Figure S3). These plasmids were named pGP104 (*tei3′*), pGP105 (*vanS$_{Sc}$′*), pGP106 (*tei3″*), and pGP107 (*vanS$_{Sc}$″*). Plasmids were transferred into *S. coelicolor* M512 (resistant to vancomycin but sensitive to teicoplanin and A40926) and *S. coelicolor* J3200 (this last one—Δ*vanS$_{Sc}$*—is constitutively resistant to all GPAs), generating M104-107 and J104-107 strains. It was expected that if ATPaseD of Tei3 would be able to function alone, the expression of *tei3′* and *tei3″* might lead to constitutive GPA resistance in M512. On the contrary, the absence of the SD in VanS$_{Sc}$ might render J3200 constitutively sensitive to GPAs. However, the obtained results indicated that ATPaseDs of neither Tei3 nor VanS$_{Sc}$ can function alone: the phenotypes of M104-107 and J104-107 strains in the presence of vancomycin, teicoplanin, and A40926 were the same as the parental M512 and J3200 strains, respectively (Figure 4).

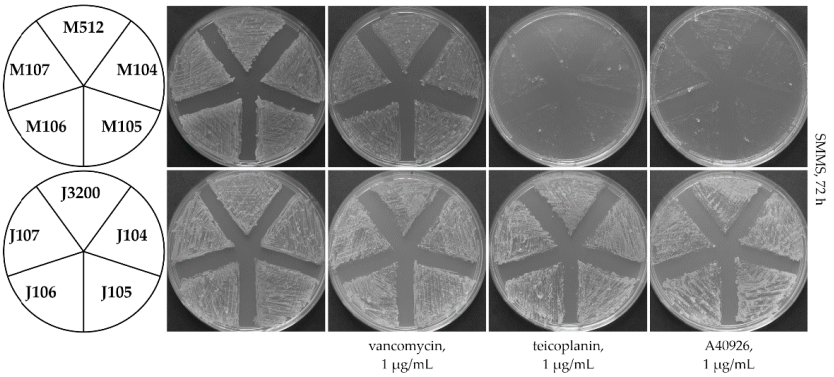


Figure 4. Loss of the gene fragments coding for transmembrane helices and extracytoplasmic sensory loops makes *tei3* and $vanS_{Sc}$ non-functional. A number of 10^6 spores were inoculated at each sector of SMMS agar plates added with vancomycin, teicoplanin, or A40926; plates were examined after 72 h of incubation. All the recombinant strains deriving from M512 were vancomycin-resistant and teicoplanin- and A40926-sensitive, whereas all the recombinant strains deriving from J3200 were resistant to all the tested GPAs.

2.4. Tei3 Sensor Domain Is Sensitive to A40926 but Not to Teicoplanin and Vancomycin

Logically, teicoplanin should be the original ligand for Tei3. However, direct experimental verification of this assumption is impossible due to the non-inducible properties of Tei3. It is possible that maybe Tei3 became a constitutive phosphorylase in the course of the evolution, but clues for Tei3 sensitivity might have remained "fossilized" within its SD. To answer this question, a hybrid gene coding for a chimeric SHK combining the SD of Tei3 with the ATPaseD of $VanS_{Sc}$ was created. $vanS_{Sc}$ and *tei3* (Figure S4) were collinear, facilitating such an exchange. Luckily, a unique *Pae*I recognition site was found only 4 bp after the $vanS_{Sc}$ region coding for SD. This allowed us to use this site for replacing the $vanS_{Sc}$ region coding for the SD with the corresponding one from *tei3* (see Section 4 for details). The obtained hybrid gene—*tei3*-$vanS_{Sc}$ (Figure S4)—was cloned into the pSET152A plasmid generating pGP103. Next, pGP103 was transferred into M512 and J3200 generating the recombinant strains M103 and J103, respectively. In the first case, a merodiploid strain was generated, carrying the native $vanS_{Sc}$ together with the hybrid *tei3*-$vanS_{Sc}$ allele, while in J3200, the knockout of $vanS_{Sc}$ was complemented by the added *tei3*-$vanS_{Sc}$. GPA resistance phenotypes of these recombinants showed that both M103 and J103 became sensitive to vancomycin, implying that this GPA is not an inducer for the Tei3 SD (Figure 5).

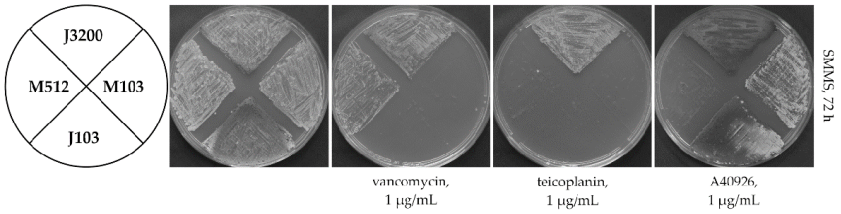


Figure 5. Hybrid SHK containing SD from Tei3 and ATPaseD from $VanS_{Sc}$ made *S. coelicolor* J3200 sensitive to teicoplanin and vancomycin, and *S. coelicolor* M512 sensitive to vancomycin, but resistant to A40926, implying that Tei3 SD recognized A40926 as a ligand. A number of 10^6 spores were inoculated at each sector of SMMS agar plates added with vancomycin, teicoplanin, or A40926; plates were examined after 72 h of incubation.

Surprisingly, M103 and J103 were also sensitive to teicoplanin, excluding its possible role as a ligand for the Tei3 SD. Adding a final detail to these puzzling results, both M103 and J103 were resistant to A40926, implying that the Tei3 SD recognizes A40926 as a ligand.

2.5. Establishing a Link between the Evolution of Glycosylation Pattern of Teicoplanin and Properties of Tei3 SD

Teicoplanin and A40926 are chemically similar GPAs that likely emerged in the course of convergent evolution (Figure S5) [47–49]. Differences lie in the chlorination and methylation pattern; notably, A40926 also lacks a *N*-acetyl glucosamine (Glc*N*Ac) moiety attached to the aglycone of teicoplanin at the aa position 6 (AA6) (Figure S5). The glycosylation pattern might be important for binding VanS, in accordance with previous results reporting that VanS from *S. toyocaensis* was unable to recognize vancomycin, sensing only the non-glycosylated A47934 [33]. Hence, it could be speculated that the responsiveness of Tei3 to a GPA lacking a Glc*N*Ac residue might be an ancient property from the times when the ancestral teicoplanin BGC did not carry a gene for the attachment of the Glc*N*Ac moiety at aglycone AA6. To understand this better, the phylogeny of glycosyl transferases (GTFs) coming from experimentally studied GPA BGCs was reconstructed (Figure 6).

Five clades (A–D) might be delineated on the obtained tree (Figure 6) and they seem to correspond to the regiospecificity of GTFs well. The regiospecificity of clades (A–C) GTFs could be predicted with high confidence as, for many members of these clades, experimental evidence exists. Thus, clade (A) comprehended GTFs attaching either D-glucose or Glc*N*Ac to AA4 of the GPA aglycone (see [50] for the review of such GTFs); clade (B) GTFs are responsible for the attachment of L-aminosugars to AA4 D-glucose [50]; clade (C) included GTFs from ristocetin BGCs likely attaching D-arabinose to AA4 D-mannosyl-D-glucose [51] (Figure 6). The substrate- and regiospecificity of GTFs from clades (D) and (E) were dubious: clade (D) GTFs probably attach L-rhamnose to the AA4 D-glucose, while clade (E) GTFs may attach L-aminosugars to aglycone AA6. To our surprise, Tei1—known to attach the Glc*N*Ac moiety at the teicoplanin aglycone AA6 [52]—was found deep in clade (A), being a sister branch to Tei10* (known to attach Glc*N*Ac at AA4). One other GPA—GP1416 from *Amycolatopsis* sp. WAC01416 [53]—is known to closely resemble the teicoplanin structure, bearing the Glc*N*Ac moiety at AA6. However, the GTF that is likely responsible for this did not belong to any of the clades and was located far from either Tei1 or Tei10* (Figure 6). The presence of the Glc*N*Ac moiety at AA6 of teicoplanin and GP1416 could, thus, be considered an example of the convergent evolution of GPAs. Considering all mentioned above, one could reasonably assume that Tei1 is a recent product of the Tei10* duplication/divergence event. In the course of evolution, it is likely that Tei1 changed its regiospecificity but retained the substrate specificity. A direct ancestor of teicoplanin BGC likely coded the biosynthesis of des-Glc*N*Ac-teicoplanin, which structurally resembles A40926 and was recognized by the Tei3 SD (see Figure 7 and the discussion below for a possible reconstruction of co-evolution of GTFs encoded in teicoplanin BGC and Tei3).

3. Discussion

vanHAXRS genes were experimentally shown to provide GPA resistance in (i) GPA-producing actinobacteria [34], (ii) actinobacterial GPA non-producers [31], (iii) GPA-resistant pathogens [28], and (iv) other soil bacteria [19,54]. In the majority of the known cases, the VanRS two-component regulatory system is utilized to sense extracellular GPAs [29] (or perhaps the GPA-lipid II complex [30]) and activate the expression of functional *van* genes. Pathogens and GPA-non-producing actinobacteria generally have inducible *van*-resistant phenotypes, because the constitutive expression of *van* genes comes at a cost of decreased fitness of the cells [55,56].

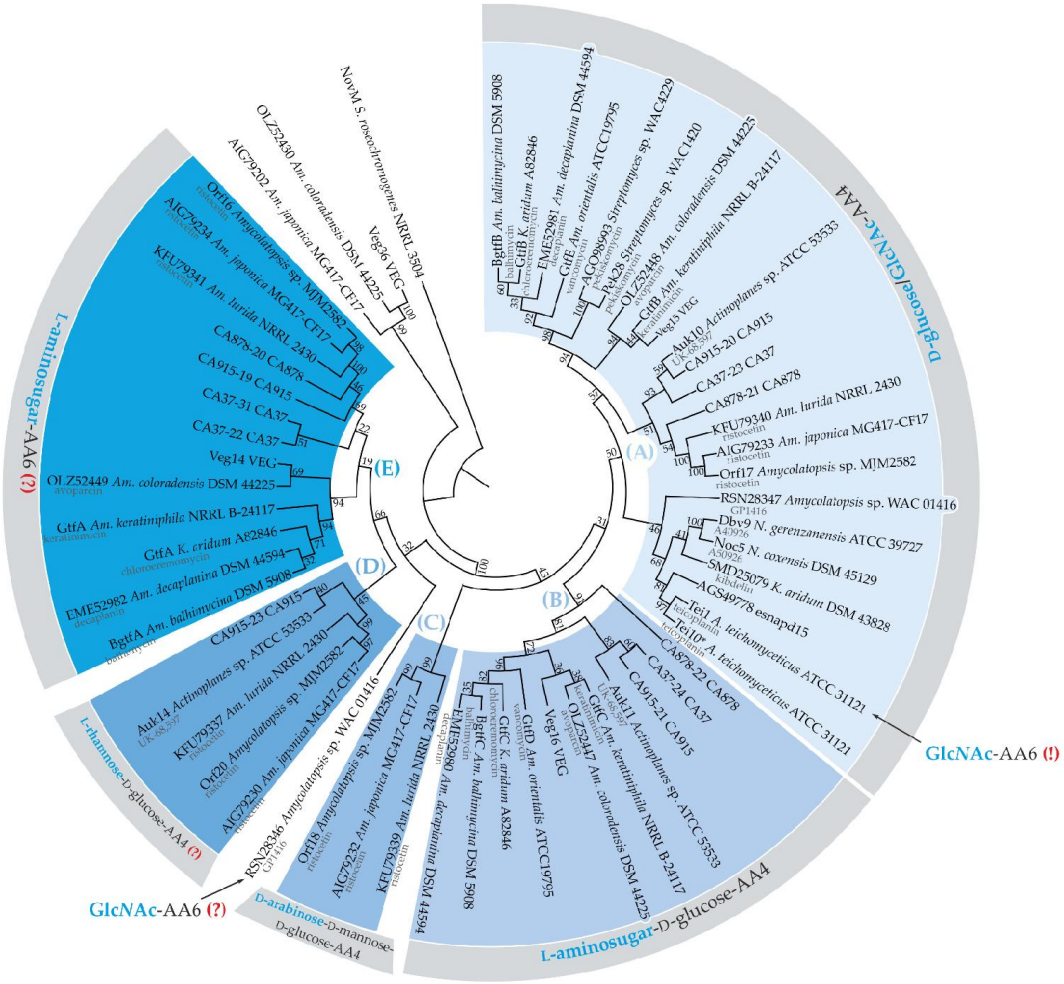


Figure 6. Maximum-likelihood phylogenetic tree of 59 GTFs coded within or near experimentally studied GPA BGCs and NovM from clorobiocin BGC [57], used as an outgroup (the tree is not drawn to scale). MEGA 11 was used for the analysis [58]. Evolution was inferred by using the JTT matrix-based model [59]. A discrete γ distribution was used to model evolutionary rate differences among sites (5 categories). Numbers at nodes indicate bootstrap-support values derived from 1000 replications. Five clades were delineated on this tree (**A**–E); clades correlated with substrate- and regiospecificity of GTFs. Veg36-OLZ52430-AIG79202 seemed to be a clade of few additional irrelevant GTFs coded nearby the corresponding BGCs. Notably, Tei1 (marked with "GlcNAc-AA6 (!)" label) appeared in the clade (**A**) grouping GTFs attaching either D-glucose or GlcNAc to aglycone AA4, while the enzyme was experimentally shown to attach GlcNAc to aglycone AA6. "(?)" signs indicate presumed regiospecificity (lack of experimental evidence).

In contrast, the GPA resistance phenotype is commonly constitutive in GPA producers, as it was shown in *A. teichomyceticus* and *Am. balhimycina* [34]. In *Am. balhimycina*, *vanHAX* orthologs are the main contributors to GPA resistance [40,60]. However, *vanHAX* are not situated within the borders of balhimycin BGC (as commonly occurs in the other actinomycetes producing GPAs), and they are not co-localized with *vanRS* orthologs [40]. Hence,

the expression of *vanHAX* is devoid of *vanRS*-mediated regulation and is constitutive. It should be noted that *Am. balhimycina* also possesses additional GPA-resistance mechanisms that are not *van*-mediated [61].

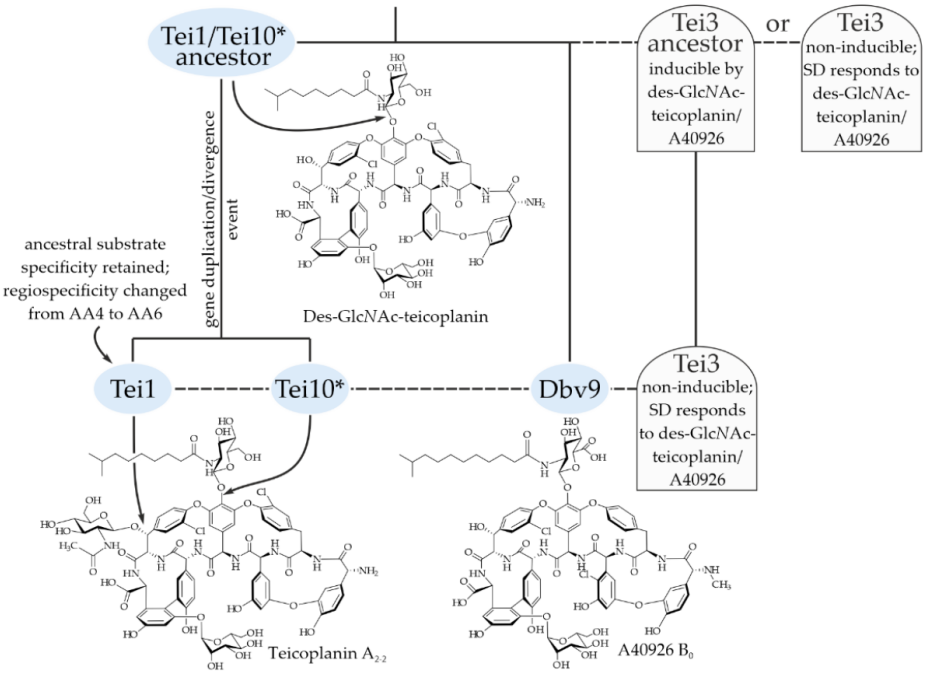


Figure 7. A scheme illustrating possible scenarios for the co-evolution of GTFs encoded in teicoplanin BGC and Tei3. Please refer to the main text for more detail.

The most striking outcome of our work was that the SD of Tei3 did not respond to teicoplanin, but recognized A40926: the expression of the hybrid SHK carrying the Tei3 SD fused with the catalytic part from $VanS_{Sc}$ changed the constitutive GPA-resistance phenotype of *S. coelicolor* J3200, making the strain teicoplanin- and vancomycin-sensitive, but A40926-resistant. This prompted us to focus on the chemical differences in the structure of A40926 and teicoplanin (the presence of a GlcNAc residue in teicoplanin, which is absent in A40296) and on the possible evolution of the teicoplanin glycosylation pattern and of the whole *tei* BGC. Reconstruction of the phylogeny of GTFs coded within the known GPA BGCs allowed us to build possible scenarios explaining the obtained results (Figure 7).

If our reconstruction of the evolution of Tei1/Tei10* is correct, the ancestor of the teicoplanin BGC coded the biosynthesis of des-GlcNAc-teicoplanin, which structurally resembles A40926 and was recognized by the Tei3 SD. Tei3 SHK initially responded to des-GlcNAc-teicoplanin (modeled with A40926 in the experiment presented in this work) (Figure 7) and it was probably inducible. A further two alternative scenarios are possible. The first implies that the non-inducibility of Tei3 evolved as an adaptation to the appearance of Tei1 because *A. teichomyceticus* was required to rapidly gain constitutive GPA resistance as Tei3 was unable to sense teicoplanin. The second scenario implies that Tei3 in the des-GlcNAc-teicoplanin producer became non-inducible at the first place, consequently giving the BGC room for further evolution and structural changes of the produced GPA. Considering that, in the first hypothesis, fewer mutations would probably be required to give Tei3 constitutive properties than to remodel its SD, the second scenario is more likely in our opinion. Hence, mutations leading to the non-inducibility of Tei3 might have served

as a preadaptation, which allowed teicoplanin to emerge in its current form, carrying the GlcNAc residue at AA6 (Figure 7).

To conclude, our results proved the non-inducibility of Tei3, showed that its SD is able to respond to A40926, but not to teicoplanin or vancomycin, and demonstrated a novel example of how BGCs, antibiotic structures, and antibiotic resistance genes might have co-evolved. Further investigations will include mutational analysis of Tei3 to elucidate which particular aa changes in its ATPaseD lead to non-inducible properties. Finally, as a byproduct, a *S. coelicolor*-based chromogenic assay for vancomycin detection was developed capable of sensing vancomycin at the ng range.

4. Materials and Methods

4.1. Plasmids, Bacterial Strains, and Cultivation Conditions

All plasmids and bacterial strains utilized or generated in the course of the current study are summarized in Table 1. For routine maintenance, *A. teichomyceticus* was cultivated on ISP3 agar [62] at 30 °C, and *S. coelicolor* strains were cultivated on SFM [45] agar at 30 °C. For genomic DNA isolation, *A. teichomyceticus* was cultivated in ISP2 liquid medium [47] in 50 mL baffled flasks on an orbital shaker at 200 rpm and 30 °C; *S. coelicolor* strains for genomic DNA isolation were cultivated under the same conditions in TSB medium [62]. Antibiotic susceptibility tests with *S. coelicolor* strains were performed on SMMS medium [45]. Apramycin sulfate (50 µg/mL), spectinomycin hydrochloride (50 µg/mL), and nalidixic acid (30 µg/mL) were used for the selection and maintenance of *S. coelicolor* recombinant strains. Vancomycin, teicoplanin, and A40926 were added to solid or liquid cultures at the desired concentrations reported in Section 2. *E. coli* DH5α was used as a general cloning host, while *E. coli* ET12567 pUZ8002^{+} was used as a donor for intergeneric matings. *E. coli* strains were cultivated at 37 °C in lysogeny broth or agar media supplemented with 100 µg/mL of apramycin sulfate, 50 µg/mL of kanamycin sulfate, and 25 µg/mL of chloramphenicol when appropriate. All antibiotics were purchased from Sigma-Aldrich (Merck Group, Darmstadt, Germany).

Table 1. Plasmids and bacterial strains used or generated in this study.

Name	Characteristic	Reference
Plasmids:		
pSET152A	φC31-based integrative plasmid, pSET152 derivative carrying *aac(3)IVp* from pIJ773, Am^{r}	[63]
pKC1132	suicide vector, Am^{r}	[45]
pRT801	φBT1-based integrative plasmid, Am^{r}	[46]
pGUS	φC31-based integrative plasmid, pSET152 derivative containing promoterless GusA, Am^{r}, Sp^{r}	[44]
pGP101	pSET152A derivative carrying *tei3*	this work
pGP111	pRT801 derivative carrying *tei3*	this work
pGP102	pSET152A derivative carrying SCO3589 (*vanS$_{Sc}$*)	this work
pKC1132-102	pKC1132 derivative carrying SCO3589 (*vanS$_{Sc}$*)	this work
pKC1132-103	pKC1132 derivative carrying hybrid SHK gene with *tei3*-SD and *vanS$_{Sc}$*-ATPaseD	this work
pGP103	pSET152A derivative carrying hybrid SHK gene with *tei3*-SD and *vanS$_{Sc}$*-ATPaseD	this work
pGP104	pSET152A derivative carrying *tei3′*	this work
pGP105	pSET152A derivative carrying SCO3589′	this work
pGP106	pSET152A derivative carrying *tei3″*	this work
pGP107	pSET152A derivative carrying SCO3589″	this work
pGJp	pGUS derivative carrying *vanJp*	this work
pHJp	pGUS derivative carrying *vanHp*	this work
Strains:		
A. teichomyceticus ATCC 31121	wild type, teicoplanin producer	ATCC

Table 1. *Cont.*

Name	Characteristic	Reference
S. coelicolor M512	A3(2) derivative, Δ*redD* Δ*actII-ORF4* SCP1$^-$ SCP2$^-$	[38]
S. coelicolor J3200	A3(2) derivative, Δ*SCO3589* (*vanS*$_{Sc}$)	[31]
S. coelicolor M101	M512 derivative carrying pGP101	this work
S. coelicolor M111	M512 derivative carrying pGP111	this work
S. coelicolor pGJp$^+$	M512 derivative carrying pGJp	this work
S. coelicolor pGHp$^+$	M512 derivative carrying pGHp	this work
S. coelicolor M1J	M512 derivative carrying pGJp and pGP111	this work
S. coelicolor M103	M512 derivative carrying pGP103	this work
S. coelicolor M104	M512 derivative carrying pGP104	this work
S. coelicolor M105	M512 derivative carrying pGP105	this work
S. coelicolor M106	M512 derivative carrying pGP106	this work
S. coelicolor M107	M512 derivative carrying pGP107	this work
S. coelicolor J103	J3200 derivative carrying pGP103	this work
S. coelicolor J104	J3200 derivative carrying pGP104	this work
S. coelicolor J105	J3200 derivative carrying pGP105	this work
S. coelicolor J106	J3200 derivative carrying pGP106	this work
S. coelicolor J107	J3200 derivative carrying pGP107	this work

4.2. Generation of Recombinant Plasmids for Gene Expression and Promoter-Probe Vectors

In all cases described below, Q5 High-Fidelity DNA Polymerase (NEB, Ipswich, MA, USA) was used as a DNA polymerase of choice in PCRs according to the recommendations of the supplier. Restriction endonucleases and T4 DNA ligase from Thermo Fisher Scientific (Waltham, MA, USA) were always used for DNA digestion and ligation according to the supplier's recommendations. Chromosome DNA, utilized as a PCR template, was isolated according to the Kirby procedure [45].

pGP101, pGP111. The coding sequence of the *tei3* gene was amplified from the chromosome DNA of *A. teichomyceticus* using the tei3_F/R primer pair (Table 2); the obtained 1133 bp amplicon was digested with *Eco*RI/*Eco*RV restriction endonucleases and cloned into the pSET152A vector digested with the same restriction endonucleases yielding pGP101. Further, pGP101 was digested with *Bam*HI/*Xho*I restriction endonuclease, and a 3247 bp DNA fragment carrying *aac(3)IVp-tei3* was purified; in parallel, the pRT801 plasmid was also digested with *Bam*HI/*Xho*I and the 3362 bp DNA fragment was purified. Both obtained fragments were then ligated, generating pGP111.

Table 2. Oligonucleotide primers used in this study.

Name	Sequence (5′-3′) *	Purpose
tei3_sdomain_PaeI	CGAGCATGCGACCGGCCAG	Reverse primer for cloning the *tei3* region coding for the sensor domain
tei3_F	TTTGATATCGGAGGGAGACCGTGGACCGAGCCC	Cloning of the *tei3* and its truncated versions
tei3′_F	TTTGATATCGGAGGGAGACCGTGTTCGCCCCGGCGACGG	-/ /-
tei3″_F	TTTGATATCGGAGGAGACCGTGGGTCGGATGCTCGCTCCTCT	-/ /-
tei3_R	TTTGAATTCGCGGTGGGCGGTTCAGTTT	-/ /-
SCO3589_F	TTTGATATCGGAGGGCGACGGTGGATAGGCGCC	Cloning of the *SCO3589* (*vanS*$_{Sc}$) and its truncated versions
SCO3589′_F	TTTGATATCGGAGGGAGACCGTGCTTCGCAGTTTCGCCC	-/ /-
SCO3589″_F	TTTGATATCGGAGGGCGACGGTGGGACGCATGCTCGCCCCCCT	-/ /-
SCO3589_R	TTTGAATTCTGGCGCTCACCTGCCGGTG	-/ /-
vanHp_KpnI_F	TTTGGTACCTACGTCCACACCGCCGAGC	Cloning of *vanH*$_{Sc}$ promoter region into pGUS
vanHp_SpeI_R	TTTACTAGTGCCGTCCCGTATGCGCTTT	-/ /-
vanJp_KpnI_F	TTTGGTACCACACTCAGCAGCTCCAACG	Cloning of *vanJ*$_{Sc}$ promoter region into pGUS
vanJp_SpeI_R	TTTACTAGTCTGGCGCCGGTGCGGCCGA	-/ /-
aac(3)IV_F	ATCGACTGATGTCATCAGCG	Diagnostic primers for the amplification of *aac(3)IV* gene
aac(3)IV_R	CGAGCTGAAGAAAGACAAT	-/ /-
gusA_ver_F	GGCGGCTACACGCCCTTCGA	Diagnostic primers for the amplification of 1000 bp internal region of *gusA*
gusA_ver_R	TGATGGGCCGGGTGGGGTC	-/ /-

* artificial ribosome binding sites are highlighted in red and artificial start codons are underlined.

pGP102, pKC1132-102, pKC1132-103, and pGP103. The coding sequence of *vanSsc* (*SCO3589*) was amplified using the chromosome DNA of *S. coelicolor* M512 as a template with the SCO3859_F/R primer pair; the obtained 1130 bp amplicon was digested with *Eco*RI/*Eco*RV restriction endonucleases and cloned into pSET152A via the same sites, giving pGP102. pGP102 was digested with *Pvu*II and the 1753 bp fragment (carrying *aac(3)IVp-SCO3589*) was cloned into pKC1132 digested with the same restriction endonuclease, generating pKC1132-102. Then, a fragment of *SCO3589* coding for the SD was exchanged for the corresponding fragment of *tei3*. The fragment of the *tei3* sequence coding for the SD was amplified from pGP101 using the tei3_F and tei3_sdomain_PaeI primers pair; the obtained 285 bp amplicon was digested with *Eco*RV and *Pae*I restriction endonucleases. In parallel, pKC1132-102 was also digested with *Eco*RV and *Pae*I and a 4289 bp fragment was selected; this latter fragment was ligated with the *Pae*I-digested *tei3-SD* amplicon, yielding pKC1132-103. Finally, the hybrid ORF of *SCO3589* with the SD-coding fragment exchanged with *tei3-SD* was excised from pKC1132-103 using *Eco*RV and *Eco*RI restriction endonucleases, and cloned into pSET152A via the same recognition sites generating pGP103.

pGP104, pGP105, pGP106, and pGP107. These vectors carried truncated versions of *tei3* and *SCO3589* lacking parts of the sequence coding for (i) TMh1 (*tei3′*—pGP104; *SCO3589′*—pGP105) and (ii) whole SD (*tei3′′*—pGP106; *SCO3589′′*—pGP107). All these coding sequences were amplified using primers listed in Table 2, generating: *tei3′*—956 bp; *tei3′′*—879 bp; *SCO3589′*—959 bp; *SCO3589′′*—878 bp. Amplicons were digested with *Eco*RV/*Eco*RI and cloned into pSET152A via the same recognition sites.

pGJp and pGHp. DNA fragments including promoter regions of *S. coelicolor vanJ (SCO3592)* and *vanH (SCO3594)* genes were amplified from the chromosome DNA of *S. coelicolor* M512 using vanJp_KpnI_F/vanJp_SpeI_R and vanHp_KpnI_F/vanHp_SpeI_R primers pairs (Table 2). The obtained amplicons for *vanJp* (321 bp) and *vanHp* (539 bp) were digested with *Kpn*I/*Spe*I restriction endonucleases and ligated with pGUS digested by the same enzymes, yielding pGJp and pGHp promoter-probe vectors.

All generated plasmids were verified by restriction mapping and sequencing.

4.3. Conjugal Transfer of Recombinant Plasmids to S. coelicolor Strains

A standard protocol for the conjugal transfer of plasmids to *S. coelicolor* strains was utilized [45]. All necessary plasmids were individually transferred into the non-methylating *E. coli* ET12567 pUZ8002$^+$ and the resulting derivatives were used as donor strains for inter-generic conjugation, while spores of *S. coelicolor* were used as acceptors. Circa 10^6 spores were mixed with, ca., 10^9 donor cells and plated on ISP3 agar supplemented with $MgCl_2$ (10 mM); after 12 h of incubation at 30 °C, each plate was overlaid with 1 mL of sterile water with 1.25 mg of apramycin-sulfate and 750 µg of nalidixic acid. Transconjugants were selected as resistant to 50 µg/mL of apramycin sulfate or 50 µg/mL of spectinomycin hydrochloride. All recombinant strains were tested with PCR using chromosome DNA isolated according to the Kirby procedure [45]. *aac(3)IV* was amplified from chromosome DNA of transconjugants carrying pSET152A and pRT801 derivatives with the aac(3)IV_F/R primer pair. Transconjugants carrying promoter-probe vectors were tested by amplifying the 1000 bp internal region of *gusA* with the gusA_ver_F/R primer pair (Table 2).

4.4. Qualitative and Quantitative Glucuronidase Assays

Qualitatively, the β-glucuronidase (GusA) activity in *S. coelicolor* was assessed by adding 25 mg/mL of 5-bromo-4-chloro-3-indolyl-β-D-glucuronide (X-Gluc, Thermo Fisher Scientific, Waltham, MA, USA) to SMMS agar. The chromogenic conversion of X-Gluc into the green-colored 5,5′-dibromo-4,4′-dichloro-indigo was then monitored. Quantitative measurement was performed as follows. A number of 10^7 spores of *S. coelicolor* pGJp$^+$ and pGHp$^+$ were inoculated into a baffled 300 mL Erlenmeyer flask containing 50 mL of TSB and incubated without and with vancomycin as an inducer. Mycelium obtained in this way was used to prepare cell-free lysates as reported previously [63]. Glucuronidase activity was measured in cell-free lysates as described previously [44,63] utilizing a spectrophotometric

assay to detect the conversion of the colorless *p*-nitrophenyl-β-D-glucuronide (Thermo Fisher Scientific, Waltham, MA, USA) into the colored *p*-nitrophenol at 415 nm using a Unicam UV 500 UV-Visible Spectrometer (Thermo, Waltham, MA, USA). Glucuronidase activity was normalized to the weight of dry biomass, and one unit of activity was considered as the amount of enzyme able to convert 1 µM of the substrate in 1 min.

4.5. Tools for In Silico Analysis

Clustal Omega (EMBL-EBI) [41] was used for pairwise alignment of aa and nucleic acid sequences. CD-Search was used to identify conserved domain regions [42]. GENEIOUS 4.8.5 was utilized for routine analysis of aa and nucleic acid sequences [64]. MEGA11 (v.11.0.13) was used to perform phylogenetic reconstruction [58].

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms232415713/s1.

Author Contributions: Conceptualization, O.Y., B.O., V.F. and F.M.; formal analysis, O.Y.; investigation, O.Y. and K.Z.; writing—original draft, O.Y. and F.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the public grant "Fondo di Ateneo per la Ricerca" 2020 and 2021 to F.M., and by the BG-09F grant of the Ministry of Education and Science of Ukraine to V.F.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data are available from the corresponding author upon reasonable request.

Acknowledgments: We thank Mark Buttner (John Innes Centre) for kindly providing us with *S. coelicolor* M512 and J3200 strains and Margaret Smith (University of York) for kindly providing us with the pRT801 vector.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Nicolaou, K.C.; Boddy, C.N.C.; Bräse, S.; Winssinger, N. Chemistry, biology, and medicine of the glycopeptide antibiotics. *Angew. Chem. Int. Ed.* **1999**, *38*, 2096–2152. [CrossRef]
2. Parenti, F.; Cavalleri, B. Proposal to name the vancomycin-ristocetin like glycopeptides as dalbaheptides. *J. Antibiot.* **1989**, *42*, 1882–1883. [CrossRef] [PubMed]
3. Marcone, G.L.; Binda, E.; Berini, F.; Marinelli, F. Old and new glycopeptide antibiotics: From product to gene and back in the post-genomic era. *Biotechnol. Adv.* **2018**, *36*, 534–554. [CrossRef] [PubMed]
4. Binda, E.; Marinelli, F.; Marcone, G.L. Old and new glycopeptide antibiotics: Action and resistance. *Antibiotics* **2014**, *3*, 572–594. [CrossRef] [PubMed]
5. Thaker, M.N.; Wright, G.D. Opportunities for synthetic biology in antibiotics: Expanding glycopeptide chemical diversity. *ACS Synth. Biol.* **2015**, *4*, 195–206. [CrossRef] [PubMed]
6. Butler, M.S.; Hansford, K.A.; Blaskovich, M.A.T.; Halai, R.; Cooper, M.A. Glycopeptide antibiotics: Back to the future. *J. Antibiot.* **2014**, *67*, 631–644. [CrossRef]
7. Van Bambeke, F. Lipoglycopeptide antibacterial agents in gram-positive infections: A comparative review. *Drugs* **2015**, *75*, 2073–2095. [CrossRef]
8. McCormick, M.H.; McGuire, J.M.; Pittenger, G.E.; Pittenger, R.C.; Stark, W.M. Vancomycin, a new antibiotic. I. Chemical and biologic properties. *Antibiot. Annu.* **1955**, *3*, 606–611.
9. Malabarba, A.; Strazzolini, P.; Depaoli, A.; Landi, M.; Berti, M.; Cavalleri, B. Teicoplanin, antibiotics from *Actinoplanes teichomyceticus* nov. sp. VI. Chemical degradation: Physico-chemical and biological properties of acid hydrolysis products. *J. Antibiot.* **1984**, *37*, 988–999. [CrossRef]
10. Leadbetter, M.R.; Adams, S.M.; Bazzini, B.; Fatheree, P.R.; Karr, D.E.; Krause, K.M.; Lam, B.M.T.; Linsell, M.S.; Nodwell, M.B.; Pace, J.L.; et al. Hydrophobic vancomycin derivatives with improved ADME properties: Discovery of telavancin (TD-6424). *J. Antibiot.* **2004**, *57*, 326–336. [CrossRef]

11. Cooper, R.D.; Snyder, N.J.; Zweifel, M.J.; Staszak, M.A.; Wilkie, S.C.; Nicas, T.I.; Mullen, D.L.; Butler, T.F.; Rodriguez, M.J.; Huff, B.E.; et al. Reductive alkylation of glycopeptide antibiotics: Synthesis and antibacterial activity. *J. Antibiot.* **1996**, *49*, 575–581. [CrossRef] [PubMed]
12. Steiert, M.; Schmitz, F.-J. Dalbavancin (Biosearch Italia/Versicor). *Curr. Opin. Investig. Drugs* **2002**, *3*, 229–233. [PubMed]
13. Melegh, S.; Nyul, A.; Kovács, K.; Kovács, T.; Ghidán, Á.; Dombrádi, Z.; Szabó, J.; Berta, B.; Lesinszki, V.; Pászti, J.; et al. Dissemination of VanA-type *Enterococcus faecium* isolates in Hungary. *Microb. Drug Resist.* **2018**, *24*, 1376–1390. [CrossRef] [PubMed]
14. Gold, H.S. Vancomycin-resistant enterococci: Mechanisms and clinical observations. *Clin. Infect. Dis.* **2001**, *33*, 210–219. [CrossRef]
15. Boneca, I.G.; Chiosis, G. Vancomycin resistance: Occurrence, mechanisms and strategies to combat it. *Expert Opin. Ther. Targets* **2003**, *7*, 311–328. [CrossRef]
16. Pootoolal, J.; Neu, J.; Wright, G.D. Glycopeptide antibiotic resistance. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 381–408. [CrossRef]
17. Marshall, C.G.; Broadhead, G.; Leskiw, B.K.; Wright, G.D. D-Ala-D-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6480–6483. [CrossRef]
18. Andreo-Vidal, A.; Binda, E.; Fedorenko, V.; Marinelli, F.; Yushchuk, O. Genomic insights into the distribution and phylogeny of glycopeptide resistance determinants within the *Actinobacteria* phylum. *Antibiotics* **2021**, *10*, 1533. [CrossRef]
19. Yushchuk, O.; Binda, E.; Fedorenko, V.; Marinelli, F. Occurrence of *vanHAX* and related genes beyond the *Actinobacteria* phylum. *Genes* **2022**, *13*, 1960. [CrossRef]
20. Williams, D.H. The glycopeptide story—How to kill the deadly "superbugs". *Nat. Prod. Rep.* **1996**, *13*, 469–477. [CrossRef]
21. Arthur, M.; Molinas, C.; Bugg, T.D.H.; Wright, G.D.; Walsh, C.T.; Courvalin, P. Evidence for in vivo incorporation of D-lactate into peptidoglycan precursors of vancomycin-resistant enterococci. *Antimicrob. Agents Chemother.* **1992**, *36*, 867–869. [CrossRef] [PubMed]
22. Marcone, G.L.; Beltrametti, F.; Binda, E.; Carrano, L.; Foulston, L.; Hesketh, A.; Bibb, M.; Marinelli, F. Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother.* **2010**, *54*, 2465–2472. [CrossRef]
23. Marcone, G.L.; Binda, E.; Carrano, L.; Bibb, M.; Marinelli, F. Relationship between glycopeptide production and resistance in the actinomycete *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother.* **2014**, *58*, 5191–5201. [CrossRef]
24. Binda, E.; Marcone, G.L.; Pollegioni, L.; Marinelli, F. Characterization of $VanY_n$, a novel D,D-peptidase/D,D- carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727. *FEBS J.* **2012**, *279*, 3203–3213. [CrossRef] [PubMed]
25. Reynolds, P.E.; Depardieu, F.; Dutka-Malen, S.; Arthur, M.; Courvalin, P. Glycopeptide resistance mediated by enterococcal transposon Tn1546 requires production of VanX for hydrolysis of D-alanyl-D-alanine. *Mol. Microbiol.* **1994**, *13*, 1065–1070. [CrossRef] [PubMed]
26. Wu, Z.; Wright, G.D.; Walsh, C.T. Overexpression, purification, and characterization of VanX, a D-,D-dipeptidase which is essential for vancomycin resistance in *Enterococcus faecium* BM4147. *Biochemistry* **1995**, *34*, 2455–2463. [CrossRef] [PubMed]
27. Bugg, T.D.; Wright, G.D.; Dutka-Malen, S.; Arthur, M.; Courvalin, P.; Walsh, C.T. Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: Biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA. *Biochemistry* **1991**, *30*, 10408–10415. [CrossRef]
28. Arthur, M.; Quintiliani, R., Jr. Regulation of VanA- and VanB-type glycopeptide resistance in enterococci. *Antimicrob. Agents Chemother.* **2001**, *45*, 375–381. [CrossRef]
29. Koteva, K.; Hong, H.J.; Wang, X.D.; Nazi, I.; Hughes, D.; Naldrett, M.J.; Buttner, M.J.; Wright, G.D. A vancomycin photoprobe identifies the histidine kinase $VanS_{Sc}$ as a vancomycin receptor. *Nat. Chem. Biol.* **2010**, *6*, 327–329. [CrossRef]
30. Kwun, M.J.; Novotna, G.; Hesketh, A.R.; Hill, L.; Hong, H.J. In vivo studies suggest that induction of VanS-dependent vancomycin resistance requires binding of the drug to D-Ala-D-Ala termini in the peptidoglycan cell wall. *Antimicrob. Agents Chemother.* **2013**, *57*, 4470–4480. [CrossRef]
31. Hutchings, M.I.; Hong, H.J.; Buttner, M.J. The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*. *Mol. Microbiol.* **2006**, *59*, 923–935. [CrossRef] [PubMed]
32. Hong, H.J.; Paget, M.S.B.; Buttner, M.J. A signal transduction system in *Streptomyces coelicolor* that activates expression of a putative cell wall glycan operon in response to vancomycin and other cell wall-specific antibiotics. *Mol. Microbiol.* **2002**, *44*, 1199–1211. [CrossRef] [PubMed]
33. Novotna, G.B.; Kwun, M.J.; Hong, H.J. In vivo characterization of the activation and interaction of the VanR-VanS two-component regulatory system controlling glycopeptide antibiotic resistance in two related *Streptomyces* species. *Antimicrob. Agents Chemother.* **2016**, *60*, 1627–1637. [CrossRef] [PubMed]
34. Yushchuk, O.; Binda, E.; Marinelli, F. Glycopeptide antibiotic resistance genes: Distribution and function in the producer actinomycetes. *Front. Microbiol.* **2020**, *11*, 1173. [CrossRef]
35. Beltrametti, F.; Consolandi, A.; Carrano, L.; Bagatin, F.; Rossi, R.; Leoni, L.; Zennaro, E.; Selva, E.; Marinelli, F. Resistance to glycopeptide antibiotics in the teicoplanin producer is mediated by *van* gene homologue expression directing the synthesis of a modified cell wall peptidoglycan. *Antimicrob. Agents Chemother.* **2007**, *51*, 1135–1141. [CrossRef]

36. Yushchuk, O.; Horbal, L.; Ostash, B.; Marinelli, F.; Wohlleben, W.; Stegmann, E.; Fedorenko, V. Regulation of teicoplanin biosynthesis: Refining the roles of *tei* cluster-situated regulatory genes. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 4089–4102. [CrossRef]
37. Yushchuk, O.; Homoniuk, V.; Ostash, B.; Marinelli, F.; Fedorenko, V. Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. *J. Antibiot.* **2020**, *73*, 255–259. [CrossRef]
38. Floriano, B.; Bibb, M. *afsR* is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.* **1996**, *21*, 385–396. [CrossRef]
39. Hong, H.J.; Hutchings, M.I.; Neu, J.M.; Wright, G.D.; Paget, M.S.B.; Buttner, M.J. Characterization of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (*vanK*) required for drug resistance. *Mol. Microbiol.* **2004**, *52*, 1107–1121. [CrossRef]
40. Kilian, R.; Frasch, H.J.; Kulik, A.; Wohlleben, W.; Stegmann, E. The VanRS homologous two-component system $VnlRS_{Ab}$ of the glycopeptide producer *Amycolatopsis balhimycina* activates transcription of the $vanHAX_{Sc}$ genes in *Streptomyces coelicolor*, but not in *A. balhimycina*. *Microb. Drug Resist.* **2016**, *22*, 499–509. [CrossRef]
41. Sievers, F.; Higgins, D.G. Clustal omega. *Curr. Protoc. Bioinform.* **2014**, *48*, 1–33. [CrossRef] [PubMed]
42. Marchler-Bauer, A.; Bryant, S.H. CD-Search: Protein domain annotations on the fly. *Nucleic Acids Res.* **2004**, *32*, 327–331. [CrossRef] [PubMed]
43. Binda, E.; Cappelletti, P.; Marinelli, F.; Marcone, G.L. Specificity of induction of glycopeptide antibiotic resistance in the producing actinomycetes. *Antibiotics* **2018**, *7*, 36. [CrossRef]
44. Myronovskyi, M.; Welle, E.; Fedorenko, V.; Luzhetskyy, A. β-Glucuronidase as a sensitive and versatile reporter in *Actinomycetes*. *Appl. Environ. Microbiol.* **2011**, *77*, 5370–5383. [CrossRef] [PubMed]
45. Kieser, T.; Bibb, M.J.; Buttner, M.J.; Chater, K.F.; Hopwood, D.A. *Practical Streptomyces Genetics*; John Innes Foundation: Norwich, UK, 2000.
46. Gregory, M.A.; Till, R.; Smith, M.C.M. Integration site for *Streptomyces* phage φBT1 and development of site-specific integrating vectors. *J. Bacteriol.* **2003**, *185*, 5320–5323. [CrossRef] [PubMed]
47. Yushchuk, O.; Vior, N.M.; Andreo-Vidal, A.; Berini, F.; Rückert, C.; Busche, T.; Binda, E.; Kalinowski, J.; Truman, A.W.; Marinelli, F. Genomic-led discovery of a novel glycopeptide antibiotic by *Nonomuraea coxensis* DSM 45129. *ACS Chem. Biol.* **2021**, *16*, 915–928. [CrossRef]
48. Yushchuk, O.; Ostash, B.; Truman, A.W.; Marinelli, F.; Fedorenko, V. Teicoplanin biosynthesis: Unraveling the interplay of structural, regulatory, and resistance genes. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 3279–3291. [CrossRef]
49. Donadio, S.; Sosio, M.; Stegmann, E.; Weber, T.; Wohlleben, W. Comparative analysis and insights into the evolution of gene clusters for glycopeptide antibiotic biosynthesis. *Mol. Genet. Genom.* **2005**, *274*, 40–50. [CrossRef]
50. Yushchuk, O.; Zhukrovska, K.; Berini, F.; Fedorenko, V.; Marinelli, F. Genetics behind the glycosylation patterns in the biosynthesis of dalbaheptides. *Front. Chem.* **2022**, *10*, 858708. [CrossRef]
51. Truman, A.W.; Kwun, M.J.; Cheng, J.; Yang, S.H.; Suh, J.W.; Hong, H.J. Antibiotic resistance mechanisms inform discovery: Identification and characterization of a novel *Amycolatopsis* strain producing ristocetin. *Antimicrob. Agents Chemother.* **2014**, *58*, 5687–5695. [CrossRef]
52. Yushchuk, O.; Ostash, B.; Pham, T.H.; Luzhetskyy, A.; Fedorenko, V.; Truman, A.W.; Horbal, L. Characterization of the post-assembly line tailoring processes in teicoplanin biosynthesis. *ACS Chem. Biol.* **2016**, *11*, 2254–2264. [CrossRef] [PubMed]
53. Xu, M.; Wang, W.; Waglechner, N.; Culp, E.J.; Guitor, A.K.; Wright, G.D. GPAHex—A synthetic biology platform for type IV–V glycopeptide antibiotic production and discovery. *Nat. Commun.* **2020**, *11*, 5232. [CrossRef] [PubMed]
54. Kruse, T.; Levisson, M.; de Vos, W.M.; Smidt, H. *vanI*: A novel D-Ala-D-Lac vancomycin resistance gene cluster found in *Desulfitobacterium hafniense*. *Microb. Biotechnol.* **2014**, *7*, 456–466. [CrossRef]
55. Foucault, M.-L.; Courvalin, P.; Grillot-Courvalin, C. Fitness cost of VanA-type vancomycin resistance in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **2009**, *53*, 2354–2359. [CrossRef] [PubMed]
56. Foucault, M.-L.; Depardieu, F.; Courvalin, P.; Grillot-Courvalin, C. Inducible expression eliminates the fitness cost of vancomycin resistance in enterococci. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16964–16969. [CrossRef]
57. Heide, L. The aminocoumarins: Biosynthesis and biology. *Nat. Prod. Rep.* **2009**, *26*, 1241–1250. [CrossRef]
58. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. [CrossRef]
59. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* **1992**, *8*, 275–282. [CrossRef]
60. Schäberle, T.F.; Vollmer, W.; Frasch, H.J.; Hüttel, S.; Kulik, A.; Röttgen, M.; Von Thaler, A.K.; Wohlleben, W.; Stegmann, E. Self-resistance and cell wall composition in the glycopeptide producer *Amycolatopsis balhimycina*. *Antimicrob. Agents Chemother.* **2011**, *55*, 4283–4289. [CrossRef]
61. Frasch, H.J.; Kalan, L.; Kilian, R.; Martin, T.; Wright, G.D.; Stegmann, E. Alternative pathway to a glycopeptide-resistant cell wall in the balhimycin producer *Amycolatopsis balhimycina*. *ACS Infect. Dis.* **2016**, *1*, 243–252. [CrossRef]
62. Koshla, O.; Lopatniuk, M.; Rokytskyy, I.; Yushchuk, O.; Dacyuk, Y.; Fedorenko, V.; Luzhetskyy, A.; Ostash, B. Properties of *Streptomyces albus* J1074 mutant deficient in $tRNA^{Leu}_{UAA}$ gene *bldA*. *Arch. Microbiol.* **2017**, *199*, 1175–1183. [CrossRef] [PubMed]

63. Horbal, L.; Kobylyanskyy, A.; Yushchuk, O.; Zaburannyi, N.; Luzhetskyy, A.; Ostash, B.; Marinelli, F.; Fedorenko, V. Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus*—Producer of teicoplanin, drug of last defense. *J. Biotechnol.* **2013**, *168*, 367–372. [CrossRef] [PubMed]
64. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef] [PubMed]

Оригінал доступний за посиланням:
https://www.mdpi.com/1422-0067/23/24/15713.
Електронні додатки до статті доступні за посиланням:
https://www.mdpi.com/article/10.3390/ijms232415713/s1.

Анотація: *Глікопептидні антибіотики (ГПА) є одним із найуспішніших класів клінічних антибіотиків. ГПА інгібують біосинтез клітинної стінки у Грам-позитивних бактерій шляхом зв'язування ліпіду II. Природні ГПА є продуктами різних актинобактерій, які і самі є Грам-позитивними бактеріями. Продуценти виробили складні механізми авторезистентності, щоб уникнути загибелі при продукції власних антибіотиків. Ці гени стійкості, скоріше за все, є первинним джерелом генів стійкості до ГПА, які поширилися до патогенних стафілококів та ентерококів. Механізми стійкості до ГПА в продуценті тейкопланіну – Actinoplanes teichomyceticus – активно вивчалися останніми роками, викликаючи закономірні запитання стосовно ролі сенсорної гістидинової кінази TeiS. У цій роботі, ми вивчаємо молекулярні властивості TeiS. Створення чутливої до ГПА сенсорної системи в модельному об'єкті Streptomyces coelicolor дало змогу продемонструвати, що TeiS функціонує як неіндуцибельна кіназа, забезпечуючи високий рівень стійкості до ГПА в A. teichomyceticus. Експресія різних вкорочених варіантів гена teiS в S. coelicolor вказала на те, що обидві трансмембранні α-спіралі TeiS необхідні для нормального функціонування. Нарешті, було створено гібридний ген, який кодував химерний білок із сенсорним доменом TeiS та кіназним доменом VanS. Неочікувано, такий химерний білок не був здатний до індукції тейкопланіном, а лише спорідненим ГПА A40926. Об'єднуючи ці експериментальні дані разом з подальшими дослідженнями in silico, ми запропонували новий сценарій для коеволюції генів стійкості до ГПА та генів біосинтезу тейкопланіну в A. teichomyceticus.*

8.2 Особливості розповсюдження та філогенії генів стійкості до глікопептидних антибіотиків в представників типу *Actinobacteria*

Як ми вже зазначали раніше, гени стійкості до ГПА зустрічаються не тільки в продуцентів ГПА, а й в актиноміцетів, не здатних до продукції ГПА. Останній випадок (наприклад *S. coelicolor*) завжди розглядався як виняток, і було незрозумілим, для чого цьому організму необхідні гени стійкості до ГПА [155]. Користуючись величезною кількістю доступної до аналізу інформації про послідовності геномів різноманітних актинобактерій, у цьому підрозділі ми вирішили дати оцінку поширенню та філогенії *van*-генів, вивчити їх розповсюдження і геномний контекст у представників різних рядів типу *Actinobacteria.*

Всього ми проаналізували 7108 геномів актинобактерій і продемонстрували, що *van*-гени є поширеними у бактерій типу *Actinobacteria* (зустрічаються з частотою приблизно 13% – знайдені принаймні в 900 геномах); при цьому, КБГ ГПА типів I-IV – стійкість до яких забезпечують *van*-гени – було знайдено лише в 39 геномах. Наш скринінг включив геноми представників 28 рядів типу *Actinobacteria*, а *van*-гени було знайдено у представників 22 рядів. Пошук був націлений на виявлення *vanY*- і *vanHAX*-подібних послідовностей, колокалізованих з *vanRS*-подібними двохкомпонентними регуляторними парами, з подальшим аналізом генетичного контексту знайдених генів. Найбільше було знайдено гомологів генів *vanY*, колокалізованих з *vanRS*, або розташованих поодиноко. Велика кількість і контекстна різноманітність вказують на тип *Actinobacteria* як на вихідне джерело *van*-генів.

Наш порівняльний геномний аналіз бактерій з різних рядів типу *Actinobacteria* показав, що *van*-гени не обов’язково колокалізовані з КБГ ГПА типів I-IV; крім того, у більшості випадків *van*-гени були виявлені у непродуцентів ГПА, а також у продуцентів ГПА типу V. Отже, раціонально припустити, що *van*-гени, розташовані не в КБГ, існували незалежно від КБГ ГПА і насправді можуть вважатися ознакою преадаптації, яка потім сприяла поширенню КБГ ГПА серед бактерій типу *Actinobacteria*. Далі, *van*-гени, розташовані в КБГ ГПА, не виявилися монофілетичною групою, а це означає,

що різні КБГ ГПА, ймовірно, отримали ці гени з різних джерел, які походять зі спільного пулу *van*-генів актинобактерій.

Нам видається малоймовірним, що така різноманітність *van*-генів у актинобактерій, які не продукують ГПА, необхідна їм для захисту від продуцентів ГПА, які є в 30 разів менш поширеними. Ми припускаємо, що *van*-гени можуть мати свої природні функції у ремоделюванні клітинної стінки у відповідь на чинники навколишнього середовища та/або в процесі морфогенезу. Наприклад, VanY-подібні M15B карбоксипептидази надзвичайно поширені в актинобактерій, а наш філогенетичний аналіз показує, що вони утворюють різні еволюційні лінії; таке розповсюдження сприяє думці, що принаймні деякі з цих білків задіяні у ремоделюванні клітинної стінки носіїв протягом життєвого циклу – подібна функція, дійсно була показана для деяких інших D-Ala-D-Ala карбоксипептидаз [171]. На користь такій ідеї свідчить знайдений в *Micromonosporales* потенційний оперон, що складається з генів PALP-подібної серинової/ треонінової дегідратази, Ddl-лігази та VanX-подібної дипептидази, колокалізованих з *vanRS*-подібною регуляторною парою. Така організація, можливо, забезпечує якийсь альтернативний механізм ремоделювання попередників пептидоглікану, який нагадує введення кінців D-Ala-D-Ser дипептидів у клітинну стінку деяких ентерококів [172]. Цей випадок дійсно заслуговує подальшого експериментального дослідження.

У підсумку, наші результати зображують актинобактерій як резервуар, в якому міститься дуже багато потенційних генів стійкості до ГПА, які рано чи пізно можуть передатися патогенним мікроорганізмам та посприяти поширенню лікарської резистентності.

Отримані в цьому підрозділі результати опубліковано в [173].

antibiotics

MDPI

Article

Genomic Insights into the Distribution and Phylogeny of Glycopeptide Resistance Determinants within the *Actinobacteria* Phylum

Andrés Andreo-Vidal [1], Elisa Binda [1], Victor Fedorenko [2], Flavia Marinelli [1,*] and Oleksandr Yushchuk [1,2]

[1] Department of Biotechnology and Life Sciences, University of Insubria, 21100 Varese, Italy; aandreovidal@uninsubria.it (A.A.-V.); elisa.binda@uninsubria.it (E.B.); oleksandr.yushchuk@uninsubria.it (O.Y.)
[2] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 79005 Lviv, Ukraine; viktor.fedorenko@lnu.edu.ua
* Correspondence: flavia.marinelli@uninsubria.it; Tel.: +39-0332-42-1546

Abstract: The spread of antimicrobial resistance (AMR) creates a challenge for global health security, rendering many previously successful classes of antibiotics useless. Unfortunately, this also includes glycopeptide antibiotics (GPAs), such as vancomycin and teicoplanin, which are currently being considered last-resort drugs. Emerging resistance towards GPAs risks limiting the clinical use of this class of antibiotics—our ultimate line of defense against multidrug-resistant (MDR) Gram-positive pathogens. But where does this resistance come from? It is widely recognized that the GPA resistance determinants—*van* genes—might have originated from GPA producers, such as soil-dwelling Gram-positive actinobacteria, that use them for self-protection. In the current work, we present a comprehensive bioinformatics study on the distribution and phylogeny of GPA resistance determinants within the *Actinobacteria* phylum. Interestingly, *van*-like genes (*vlgs*) were found distributed in different arrangements not only among GPA-producing actinobacteria but also in the non-producers: more than 10% of the screened actinobacterial genomes contained one or multiple *vlgs*, while less than 1% encoded for a biosynthetic gene cluster (BGC). By phylogenetic reconstructions, our results highlight the co-evolution of the different *vlgs*, indicating that the most diffused are the ones coding for putative VanY carboxypeptidases, which can be found alone in the genomes or associated with a *vanS/R* regulatory pair.

Keywords: glycopeptide antibiotics; multidrug resistance; antibiotic resistance; *van* genes

check for updates

Citation: Andreo-Vidal, A.; Binda, E.; Fedorenko, V.; Marinelli, F.; Yushchuk, O. Genomic Insights into the Distribution and Phylogeny of Glycopeptide Resistance Determinants within the *Actinobacteria* Phylum. *Antibiotics* **2021**, *10*, 1533. https://doi.org/10.3390/antibiotics10121533

Academic Editor: Yuji Morita

Received: 4 November 2021
Accepted: 10 December 2021
Published: 14 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Starting with the discovery of penicillin [1], humanity has been involved in a never-ending arms race between life-threatening bacterial pathogens and antibiotics, either natural, semisynthetic or completely synthetic. An estimation made by the United Nations Interagency Coordination Group on Antimicrobial Resistance in 2014 predicted that the spread of antimicrobial resistance (AMR) will cause up to 10 million deaths per year by 2050 [2]. However, recent events might make this number even more grim: the world-wide health crisis caused by SARS-CoV-2 has led to an increase in antibiotic use and misuse, which, in turn, is likely to further accelerate AMR diffusion [3]. AMR has rendered many previously successful groups of antibiotics non-functional, leaving us hiding behind the "thin red line" of last-resort drugs, capable to combat multidrug-resistant (MDR) pathogens. Glycopeptide antibiotics (GPAs) are a class of non-ribosomally synthesized, highly cross-linked, halogenated and glycosylated natural products, which are considered frontline drugs against Gram-positive MDR pathogens such as *Staphylococcus aureus*, *Enterococcus* spp., *Clostridioides difficile*, etc. [4].

GPAs are produced by high G-C content soil-dwelling actinobacteria and could be divided into five types according to their chemical structures and molecular targets [5].

Types I–IV are compounds having a cross-linked heptapeptide core highly glycosylated and/or lipidated; their molecular target in pathogens is lipid II [6], a highly conserved macromolecule among bacteria that is essential for cell-wall biosynthesis [7]. Types I–IV GPAs form five hydrogen bonds with the D-alanyl-D-alanine (D-Ala-D-Ala) terminus of the lipid II pentapeptide stem (Figure 1a) [8–11]. Such binding terminates upstream transpeptidation and transglycosylation reactions, preventing the formation of mature peptidoglycan (PG). An old term describes Types I–IV GPAs very precisely: dalbaheptides [12], meaning D-Ala-D-Ala-binding antibiotics with heptapeptide structures. Lipid II is obviously involved in the cell wall biosynthesis of dalbaheptide producers as well. Thus, to avoid suicide during GPA production, producing strains need self-resistance mechanisms. This topic was recently revised in detail [13,14]. In brief, two main mechanisms of cell-wall remodeling exist in dalbaheptide-producers as well as in the resistant pathogens [15,16]. The first involves three genes—*vanHAX*—coding for: (i) an α-ketoacid dehydrogenase, VanH, which stereospecifically reduces pyruvate to D-lactate (D-Lac); (ii) a D-Ala-D-Lac ligase, VanA; and (iii) a D,D-dipeptidase, VanX. VanX removes the D-Ala-D-Ala termini of lipid II pentapeptide stems, while VanA prepares a pool of D-Ala-D-Lac dipeptides, which MurF, a UDP-*N*-acetylmuramoyl-tripeptide ligase, installs instead of the terminal D-Ala-D-Ala. As a result, GPAs form four instead of five hydrogen bonds with such D-Ala-D-Lac termini, and the repulsion of lone electron pairs between oxygen atoms contributes to make these GPA-lipid II complexes unstable (Figure 1b). The second resistance mechanism requires the expression of a D,D-carboxypeptidase VanY, which cleaves the terminal D-Ala residues of lipid II pentapeptide stems, hampering GPA complex formation with such truncated lipid II derivatives (Figure 1c). So, the formed tetrapeptide-carrying lipid II molecules are still able to enter the transpeptidation and transglycosylation steps, yielding a 3-3 crosslinked mature peptidoglycan [17]. Expression of either *vanHAX* or *vanY* (known overall as *van* genes) is often inducible and regulated by a VanRS two-component regulatory pair. There, a sensor histidine kinase VanS (de)phosphorylates transcriptional regulator VanR in response to the extracellular presence of GPAs, and phosphorylated VanR activates the expression of the *van* genes involved in cell-wall remodeling. In dalbaheptide producers, *van* genes are always localized inside the borders of GPA biosynthetic gene clusters (BGCs), which group genes devoted to antibiotic biosynthesis, transport and regulation, guaranteeing co-regulation of self-resistance with antibiotic production. The only exception is the chloroeremomycin BGC in *Kibdelosporangium aridum* A82846, which apparently does not include *van* genes [18].

Type V GPAs are not glycopeptides *sensu stricto*, since they are not glycosylated; moreover, they are not dalbaheptides, since they include also nonapeptide antibiotics (such as corbomycin and GP6738 [19,20]), and they target autolysins (named also murein hydrolases) instead of lipid II. Autolysins are enzymes breaking the bonds within the peptidoglycan to allow bacterial growth and cell division [21]; thus, type V GPAs block cell-wall remodeling, arresting cell division. Consistently, type V GPA BGCs lack the cluster-situated *van* genes needed to remodel lipid II termini [19,22]. *van* genes are also missed in the related BGC of the uncrosslinked non-glycosylated peptide antibiotic known as feglymycin [23], which inhibits peptidoglycan synthesis targeting MurA (the UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase, catalyzing phosphoenolpyruvate transfer to UDP-*N*-acetylglucosamine) and MurB (the UDP-*N*-acetylenolpyruvoylglucosamine reductase, catalyzing the last step of the formation of UDP-*N*-acetylmuramic acid), both acting at the cytoplasmic step of cell-wall biosynthesis. However, type V GPA and feglymycin BGCs carry genes for a two-component regulatory pair, consisting of a sensor histidine kinase and a response regulator, which remind VanS and VanR of dalbaheptide BGCs, but whose functions remains unclear.

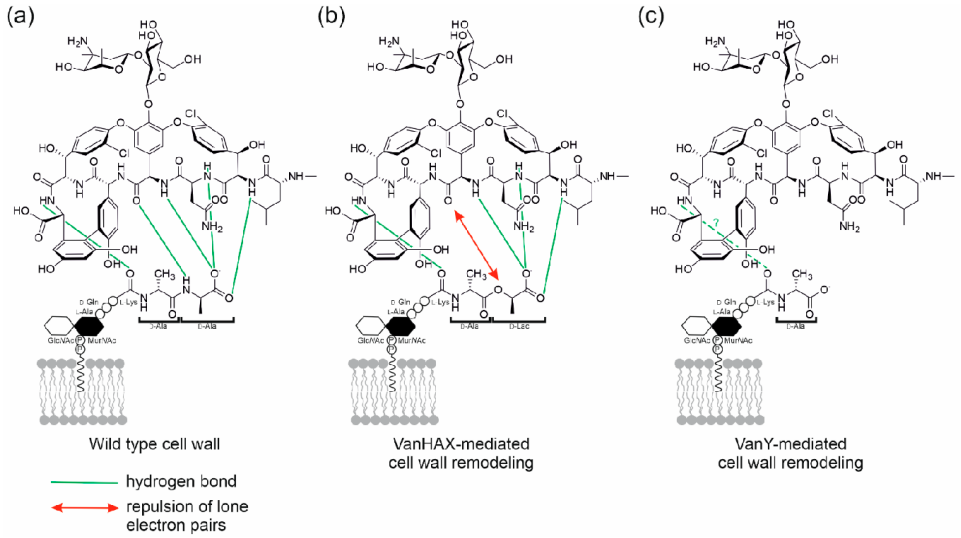


Figure 1. Schematic representation of (**a**) vancomycin (as a model GPA), interacting with the D-Ala-D-Ala terminus of the lipid II pentapeptide stem of a Gram-positive cell wall, forming five hydrogen bonds [10]; (**b**) in the cell wall remodeled by the action of VanHAX, the D-Ala-D-Lac termini of lipid II pentapeptide stem interacts with a significantly lower (1000 fold lower) affinity with vancomycin, due to the formation of only four hydrogen bonds and the repulsion of lone electron pairs between oxygen atoms [11]; (**c**) in the cell wall remodeled by the action of VanY D,D-carboxypeptidase, the lipid II pentapeptide stem is truncated by the excision of the terminal D-Ala, and vancomycin affinity for this target appears significantly reduced, although to which extent has not been investigated yet.

Intriguingly, functional *van* genes are also found in non-producing actinobacteria such as in *Streptomyces coelicolor* A3(2) [24], as well as in various other environmental non-infectious low G-C content bacteria (such as *Paenibacillus popilliae* [25] or *Bacillus circulans* [26]). As a consequence, different hypotheses have been proposed to explain *van* gene distribution and evolution in bacteria. The most accredited one is that pathogens such as enterococci (and the GPA non-producing environmental bacteria) might have acquired *van* genes from dalbaheptide producers through a series of horizontal gene transfer events, likely promoted by the selective pressure exerted by antibiotic environmental contamination [27]. An alternative hypothesis suggests that pathogens received *van* genes from low G-C soil Gram-positives (*Firmicutes* phylum), as the above-mentioned *Pnb. popilliae* and *Bac. circulans* [28], implying that in these bacteria *van* genes evolved independently.

To get an insight into the phylogeny of *van* genes, in this work we analyze their distribution and arrangement within the different orders belonging to the *Actinobacteria* phylum, using the genomic data available in public databases. We have investigated more than 7000 actinobacterial genomes and found *van*-like genes (defined hereafter *vlgs*) in more than one tenth of them. Thus, we can confirm that the presence of *vlgs* is not limited to the dalbaheptide producers, since these genes are also widely distributed among genomes of Type V GPA producers and in non-producing actinobacterial taxa, which do not need them for self-resistance. In addition, phylogenetic reconstructions made for VanY-like proteins as well as for the VanHAX triads and for the two-component VanS/VanR regulatory system, highlight the evolutionary independent stories of the corresponding gene acquisitions. Thanks to this comparative genomic analysis, novel transposon-like mobile elements carrying *vlgs* are here for the first time described, originating from poorly investigated orders such as *Eggerthellales* and *Coriobacteriales*. Finally, as a control, the same bioinformatic analysis was applied to more than 2000 *Bacillales* complete genome assemblies, which yielded only a few *vlgs*, often adjacent to transposase-like open reading frames (ORFs).

Taken altogether, these data reveal that the phylum *Actinobacteria* is an incredibly vast source of variable GPA resistance determinants, which might potentially continue to move towards pathogens, contributing to the alarming diffusion of AMR. Their study might help the surveillance of AMR spread in compliance with the One-Health approach [29].

2. Results

2.1. Organization of *vlgs* in GPA Producers and Beyond

Until now, *vlgs* in actinobacteria were reported as BGC-situated in more than 20 "classical" dalbaheptide-producers [13]. The case of the GPA non-producer actinobacterium *S. coelicolor* A3(2)—having a whole set of *vlgs* [30,31]—was rather interpreted as an exception. Thanks to the abundance of genomic data on actinobacteria today, our aim is to prove or disprove the assumption that *vlgs* are peculiar to GPA producers, and to clarify how these genes are eventually distributed and organized among different orders belonging to *Actinobacteria* phylum. Thus, we screened all the genome assemblies available for actinobacteria in GenBank at the moment of this work preparation (April 2020, Supplementary Excel File S1). This search covered 28 orders of the *Actinobacteria* phylum (Table 1), including two "candidate" ones (namely *candidatus* Actinomarinales and Nanopelagicales). We searched for *vanY* and *vanHAX* sequences, co-localized with *vanRS*-like two-component regulatory pairs, and then we analyzed the genetic context of these genes.

Table 1. List of orders belonging to the *Actinobacteria* phylum, which genome assemblies were available in GenBank at the moment of this work preparation (April 2020, Supplementary Excel File S1) and a summary of the *vlgs* found in them.

Order	Number of Genome Assemblies Analyzed	Number of Genome Assembling Containing at Least One *Vlg*	Occurrence (%)	*vanY*-like	*vanR*-like	*vanS*-like	*vanH*	*vanA*	*vanX*
Acidimicrobiales	216	0	0	-	-	-	-	-	-
Actinomycetales	223	1	0.45	1	1	1	-	-	-
Actinopolysporales	10	0	0	-	-	-	-	-	-
Bifidobacteriales	1028	0	0	-	-	-	-	-	-
candidatus Actinomarinales	214	0	0	-	-	-	-	-	-
candidatus Nanopelagicales	26	1	3.85	1	-	-	-	-	-
Catenulisporales	3	1	34	1	1	1	1	1	1
Coriobacteriales	217	2	0.9	1	2	2	2	2	2
Corynebacteriales	707	110	15.6	122	23	25	14	14	14
Cryptosporangiales	3	2	67	-	1	1	1	1	1
Eggerthellales	106	1	0.9	-	1	1	1	1	1
Egibacterales	3	0	0	-	-	-	-	-	-
Frankiales	46	7	15	-	6	7	8	8	8
Gaiellales	3	0	0	-	-	-	-	-	-
Geodermatophilales	60	11	18	12	10	9	-	-	-
Glycomycetales	12	11	92	10	10	10	8	8	8
Jiangellales	11	10	91	12	11	10	7	7	7
Kineosporiales	12	1	8.4	1	1	1	-	-	-
Micrococcales	1741	15	0.86	2	13	13	15	15	14
Micromonosporales	200	83	42	14	86	88	5	57	56
Nakamurellales	6	6	100	6	1	1	-	-	-
Nitriliruptorales	6	1	17	1	-	-	-	-	-
Propionibacteriales	593	16	2.7	13	16	16	16	16	16
Pseudonocardiales	243	135	56	141	129	139	105	105	105
Rubrobacterales	8	1	12.5	2	1	1	-	-	-
Solirubrobacterales	45	4	8.9	6	6	6	2	2	2
Streptomycetales	1138	418	37	126	414	429	93	92	94
Streptosporangiales	228	63	28	52	67	63	7	7	6

At least one *vlg* sequence was found in the majority of orders (22, Table 1); only the orders *Acidimicrobiales, candidatus* Actinomarinales, *Actinopolysporales, Bifidobacteriales, Egibacteriales* and *Gaiellales* lacked any *vlgs*. *vlgs* were most abundant (found in more than 10% of the genomic records per each order) in orders *Catenulisporales, Corynebacteriales, Cryptosporangiales, Frankiales, Geodermatophilales, Glycomycetales, Jiangellales, Micromonosporales, Nakamurellales, Nitriliruptorales, Pseudonocardiales, Rubrobacterales, Streptomycetales* and *Streptosporangiales* (Table 1, Supplementary Excel File S2). Since the known GPA producers belong to *Pseudonocardiales, Streptosporangiales, Micromonosporales* and *Streptomycetales* [5], we started to investigate *vlgs* organization in these orders in detail, to move then into the analysis of the still unexplored taxa.

2.1.1. Order *Pseudonocardiales*

Order *Pseudonocardiales* is the most abundant source of Types I–IV GPAs [5,32]. So far, GPA BGCs were described only in the *Amycolatopsis* and *Kibdelosporangium* genera. In our screening of 243 genomes from *Pseudonocardiales* spp., we found at least one *vlgs* sequence in 135 assemblies (Table 1, Supplementary Excel Files S1 and S2). Only 30 assemblies contained GPA BGCs (Supplementary Excel File S2). Besides the known GPA producing genera, *Amycolatopsis* and *Kibdelosporangium* [18], a GPA BGC was, for the first time, found in a species belonging to the genus *Actinokineospora*. Overall, no correlation between the quality and quantity of *vlgs* in GPA producers and non-producers was observed. The repertoire of *vlgs* was quite different in each strain. The following combinations were found: *vanYRS, vanHAXRS, vanHAX, vanYHAX* and *vanYHAXRS*. A significant portion of the *vanY*-like genes was found to be "orphan" (meaning not co-localized with any other *vlgs*). Sequences similar to the *vanJ* and *vanZ* genes, which were previously sporadically reported as involved in GPA resistance, but apparently without an essential role [14], were rare and always co-localized with the *vanRS*-like pair (except cases in *Pseudonocardia* sp. CNS-139—*vanHAXZ* arrangement; and in *Pseudonocardia cypriaca* DSM 45511—*vanHAXJ* arrangement).

To study the arrangement of *vlgs* in detail, we focused on the 30 genomes of the known GPA producers and, as a control, on 25 genomes of never previously investigated non-producing *Pseudonocardiales* spp. (Figure 2). Thus, in the majority of the GPA-producing *Amycolatopsis*, the *vanHAX* operon was found just upstream of the *bbr*-orthologue (coding for a StrR-like cluster-situated-pathway-specific regulator of GPA biosynthesis [33]), in rare cases having a *vanY*-like gene in between (Figure 2). At the same time, genomes of these *Amycolatopsis* GPA producers contained two copies of *vanY*-like genes located outside the GPA BGCs (Figure 2). One of these copies was always co-localized with *vanRS*-like two-component regulatory genes. *Amycolatopsis balhimycina* DSM 5908 (balhimycin producer) and *Amycolatopsis* sp. H5 were notable exceptions: belonging to a different clade than other *Amycolatopsis* spp. GPA producers, they had a *vanSRY*-genes cluster-situated and a *vanHAX* operon outside the BGC (Figure 2). In *Amycolatopsis bartoniae* DSM 45807, only a *vanY*-like gene was found upstream of the *bbr*-orthologue, while a *vanHAX* operon coupled with *vanRS*-like two-component regulatory genes was placed somewhere else on the chromosome (Figure 2). However, according to 16S rRNA gene phylogeny (see phylogenetic framework on Figure 2), *Am. bartoniae* appeared to outgroup all other *Amycolatopsis* spp. together with *Prauserella muralis* DSM 45305, *Tamaricihabitans halophyticus* DSM 45765 and *Amycolatopsis* sp. KNN50.9b. Thus, it is likely that *Am. bartoniae* (as well as *Amycolatopsis* sp. KNN50.9b) might not belong to the *Amycolatopsis* genus at all. GPA-producing *Kibdelosporangium* spp. had no cluster-situated *vlgs*, but *vanHAX* operons and multiple copies of *vanY*-like genes were found somewhere else on the chromosome (Figure 2). Finally, in *Actinokineospora auranticolor* YU 961-1, *vlgs* were only GPA cluster-situated—a set of *vanHAXRSY* genes was found upstream of the *bbr*-orthologue (Figure 2).

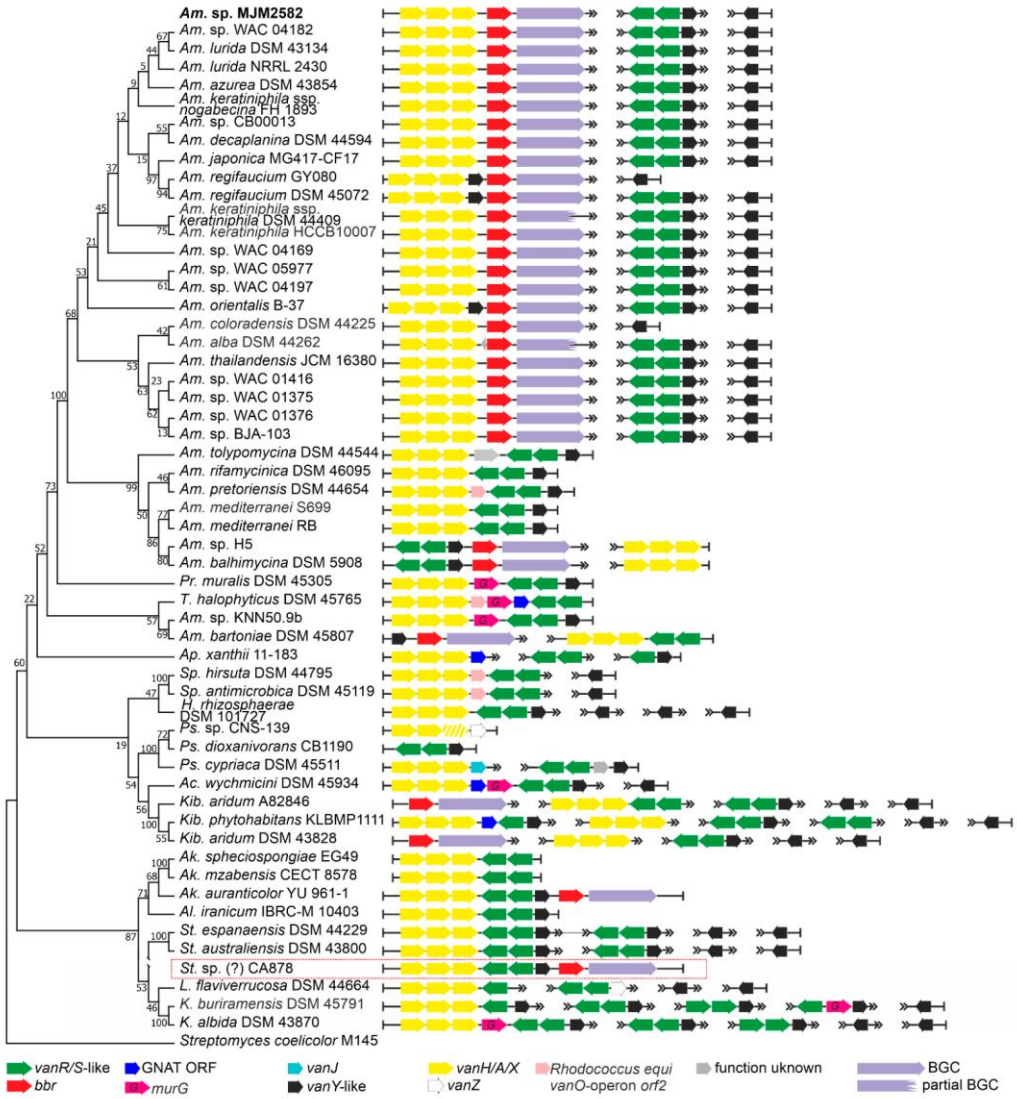


Figure 2. Organization of *vlgs* found within the set of 55 chosen genomes from *Pseudonocardiales* spp. Maximum Likelihood phylogenetic tree of the 16S rRNA genes of the corresponding species (see Section 4 for details) served as a phylogenetic framework for the scheme. *Amycolatopsis* sp. MJM2582 (in bold at the top of the figure) is outside the framework due to the lack of a full 16S rRNA gene in the corresponding genome assembly. *vlgs* from the metagenomic CA878 BGC (highlighted with red frame) were arbitrary introduced in this scheme since the retrieved sequences likely belong to *Saccharothrix* spp. Names of the *Pseudonocardiales* genera were abbreviated according to ESM Table S1. The legend below the figure explains the color-coding of the scheme. Please refer to the text for the role of the single genes. Pseudogenes are shaded.

Notably, in some cases (e.g., *Saccharopolyspora hirsuta* DSM 44795), the *vanHAX* operon was followed by a homologue of *orf2* (a gene coding for a protein with unknown function) previously identified in the *Rhodococcus equi* S7B *vanO* operon [34]. As it does occur in the *vanO* operon, the *orf2* homologue was in some species followed by a *murG*-like gene (e.g., *T. halophyticus* DSM 45765), coding for an essential peptidoglycan glyco-

syltransferase [35]. It is so far unknown how exactly MurG contributes to GPA resistance, but it might be assisting *van*-mediated cell-wall remodeling. In a few species, such as *Pr. muralis* DSM 45305, *murG* was following *vanHAX* directly. Genes coding for a GCN5-related *N*-acetyltransferases (GNATs) were also often found co-localized with *Pseudonocardiales* spp. *vlgs*.

Finally, we introduced in our phylogenetic framework the only metagenomics-derived GPA BGC—CA878 [36]—which was recently shown to be rather closely related to BGCs from *Amycolatopsis* spp. [18]. We speculated that this BGC might also belong to some unknown species of *Pseudonocardiales*—the organization of cluster-situated *vlgs* here seemed identical to the one from *Ak. auranticolor* YU 961-1. Fortunately, at the moment of its discovery, CA878 BGC was sequenced together with unannotated DNA flanks (*ca.* 26 and 3 kbp). This allowed us to annotate 9 ORFs upstream and 6 ORFs downstream of the borders of CA878 BGC (ESM Figure S1). It appeared that the vast majority of these ORFs coded for proteins with orthologues in *Saccharothrix* spp. (ESM Figure S1). In our opinion, it is possible that CA878 comes from an unknown species belonging to the *Saccharothrix* genus, further expanding the list of GPA-producing *Pseudonocardiales* genera.

2.1.2. Order *Streptosporangiales*

Order *Streptosporangiales* is the source of valuable lipidated dalbaheptides, such as A40926 from *Nonomuraea gerenzanensis* ATCC 39727, which is the precursor of second-generation dalbavancin [4,37]. In addition, *N. gerenzanensis* was the first model for studying the role of VanY-like carboxypeptidases in self-resistance [38,39]. Other known GPA producers are *Nonomuraea coxensis* DSM 45129, recently reported to produce the lipoglycopeptide A50926 [40], and *Nonomuraea* sp. ATCC 5507, which produces the type V kistamicin [41]. Here, we analyzed the genomic records from 228 *Streptosporangiales* spp. (Supplementary Excel File S1). *vlgs* were found in 63 out of them (Table 1, Supplementary Excel File S2). *vanY*-like genes were the most abundant, in almost all cases being co-localized with *vanRS*-like regulatory pairs (although few "orphan" ones were also found). *vanHAX* operons were not found so often, coming exclusively from *Actinomadura* spp.; anyhow, the *vanHAX* operon was never found associated with *vanY*-like genes. Accessory *vlgs*, such as *vanJ*, were found rarely and tended to be co-localized with the *vanRS*-like regulatory pairs. Only in one case—in *Actinomadura* sp. H3C3—a *vanZ* pseudogene was discovered, co-localized with *vanRSY*.

Going into more detail, we analyzed the genomes from the three known GPA producers, those from five strains carrying putative GPA-like BGCs, including *Nonomuraea* sp. WAC 01424 [18], together with 17 genomes of other *Streptosporangiales* spp. lacking any GPA BGCs (Figure 3). We found that the two Type IV GPA producers—*N. gerenzanensis* ATCC 39727 and *N. coxensis* DSM 45129—carried one BGC-situated *vanY*-like gene and an additional *vanY*-like gene, co-localized with a *vanRS*-like regulatory pair, distantly from the BGCs. The similarity of both distant- and cluster-encoded VanY-carboxypeptidases (amino acid sequence identity of 82.4% in *N. coxensis* and of 80.7% in *N. gerenzanensis*) was remarkable. The *vanRSY*-triad was found also in other *Nonomuraea* spp., lacking any GPA BGCs, such as in *Nonomuraea fuscirosea* CGMCC 4.7104 (Figure 3); it also was present on the chromosome of the kistamicin producer *Nonomuraea* sp. ATCC 55076 away from its BGCs, although this Type V GPA probably targets autolysins (such as corbomycin and complestatin [19]), thus not requiring *van* genes for self-resistance. Peculiarly, in the putative Type IV GPA producer *Nonomuraea* sp. WAC 01424 [18], GPA BGC seemed to be localized just downstream the *vanRSY* triad (which indeed was not cluster-situated in the other three GPA-producing *Nonomuraea* spp., Figure 3). WAC 01424 BGC carried another two-component regulatory pair, but without an additional copy of *vanY*.

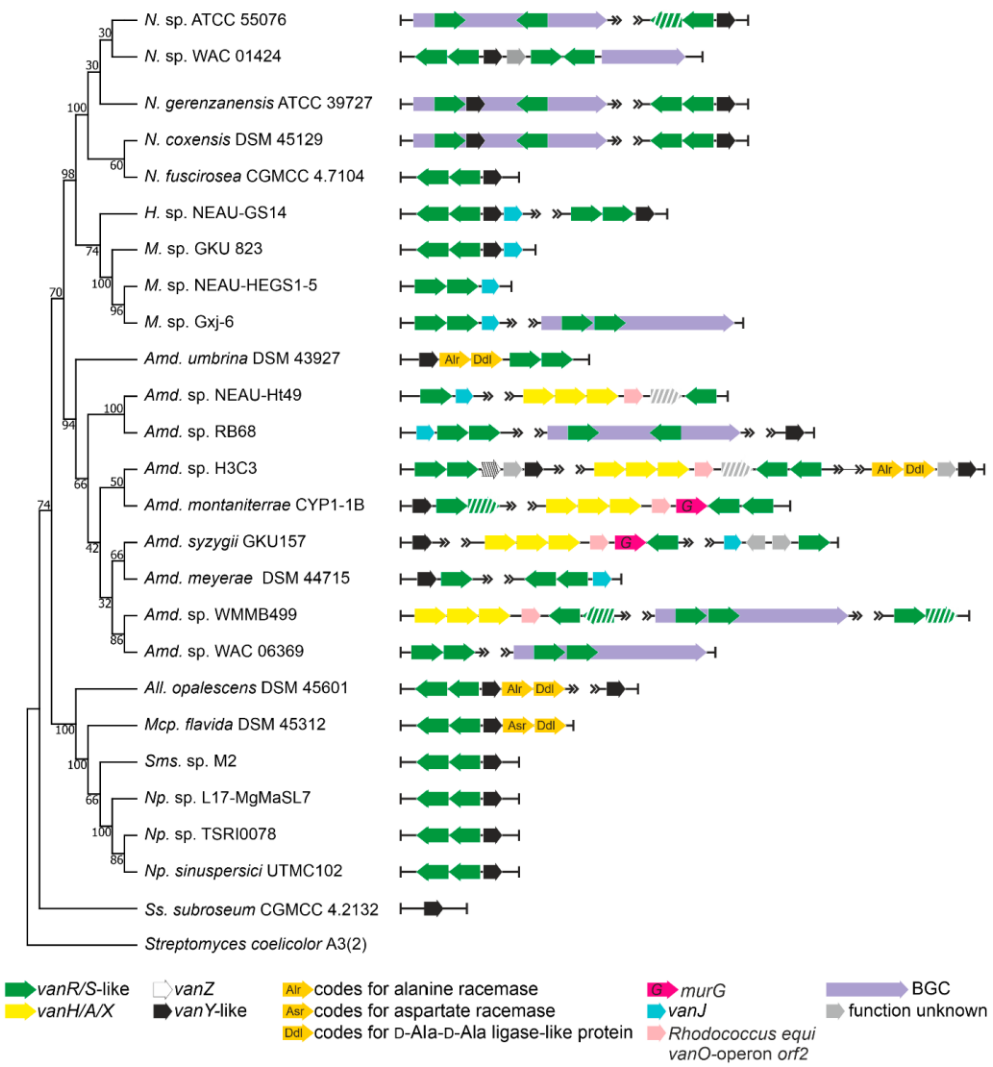


Figure 3. Organization of *vlgs* found within the set of 25 chosen genomes from *Streptosporangiales* spp. Maximum Likelihood phylogenetic tree of the 16S rRNA genes of the corresponding species (see Section 4 for details) served as a phylogenetic framework for the scheme. Names of *Streptosporangiales* genera were abbreviated according to ESM Table S1. The legend below the figure explains the color-coding of the scheme. Please refer to the text for the role of the single genes. Pseudogenes are shaded.

Another two findings are worth mentioning. Similar to what was observed for *Pseudonocardiales*, we found that all *vanHAX* operons (present exclusively in *Actinomadura* spp. among *Streptosporangiales*) had a homologue of the *vanO* operon *orf2* (and sometimes a *murG*-like gene, too) downstream of *vanX* (Figure 3). Finally, the *vanRSY* triad was in rare cases co-localized with genes coding for alanine/aspartate racemase- and D-Ala-D-Ala-ligase (Ddl)-like proteins (e.g., *Allonocardiopsis opalecscens* DSM 45601 and *Murinocardiopsis flavida* DSM 45312, Figure 3), whose role would merit further investigation, indicating a possible alternative mechanism of cell-wall remodeling, such as the one reported in few enterococci based on the incorporation of D-Ala-D-Ser termini in the resistant peptidoglycan precursors [14].

2.1.3. Order *Micromonosporales*

The order *Micromonosporales* is rich in GPA producers coming from the genus *Actinoplanes* [42]. These are: (i) the clinically relevant lipo-GPA teicoplanin, coming from *Actinoplanes teichomyceticus* ATCC 31121 [43]; (ii) the sulfated GPA UK-68,597 [44], coming from *Actinoplanes* sp. ATCC 53533; and (iii) the hyperglycosylated GPA actaplanin from *Actinoplanes missouriensis* ATCC 23342 [45]. We screened 200 genome assemblies (Supplementary Excel File S1) and we found *vlgs* in 83 of them (Table 1, Supplementary Excel File S2). Complete sets of *vanHAXRS* genes were found only in GPA producers *Apl. teichomyceticus* (these *vlgs* were already studied experimentally [46]) and *Actinoplanes* sp. ATCC 53533, as well as in the GPA non-producer *Actinoplanes derwentensis* DSM 43941 (Figure 4).



Figure 4. Organization of *vlgs* found within the set of 28 chosen genomes of *Micromonosporales* spp. Maximum Likelihood phylogenetic tree of the 16S rRNA genes of the corresponding species (see Section 4 for details) served as a phylogenetic framework for the scheme. Names of the *Micromonosporales* genera were abbreviated according to ESM Table S1. *vlgs* from the metagenomic CA915 and CA37 BGCs (highlighted with red frame) were arbitrarily introduced within the scheme since the retrieved sequences likely belong to some *Actinoplanes* spp. The legend below the figure explains the color-coding of the scheme. Please refer to the text for the role of the single genes. Pseudogenes are shaded.

We chose a total of 26 genomes (including only two genomes of GPA producers, since the genome of *Apl. missouriensis* ATCC 23342 is not yet available) for more detailed examination. *vanY*-like genes were found in *Actinoplanes* spp., although they were not co-localized with the *vanRS*-like regulatory pairs. Instead, the *vanRS*-like regulatory pairs were often found co-localized with *vanZ* genes (like in *Actinoplanes missouriensis* 431, Figure 4) or found without any other close *vlg* (such as in *Actinoplanes italicus* DSM 43146). However, one peculiarity specifically attracted our attention. In the course of our screenings, a particular gene arrangement was found to occur very often in the genomes of different *Micromonosporales* spp., especially in those belonging to the *Micromonospora* genus (Figure 4, Supplementary Excel File S2). This arrangement included a triad of genes coding for a PALP (pyridoxal-phosphate dependent)-like serine-threonine dehydratase, a Ddl-like protein and a VanX-like dipeptidase (further referred to as a *pdx* operon). Such a triad was accompanied with a VanRS-like regulatory pair. Overall, such an arrangement strikingly resembled typical *vanHAX-vanRS* operons, although the gene for lactate dehydrogenase was replaced by a serine-threonine dehydratase.

Finally, two metagenome-derived GPA-BGCs were described as related to the *Actinoplanes*-derived ones [18]. These were CA915 and CA37 [47]. Both of them were submitted to GenBank with rather long unannotated flanking regions. As in the case of CA878 (see above), we annotated the genes present on these flanks. The majority of the BGCs flanking genes seemed to be homologous to *Actinoplanes* spp. genes (ESM Figure S2). Thus, CA915 and CA37 most likely belong to some unknown *Actinoplanes* spp.

2.1.4. Order *Streptomycetales*

Although multiple genomes of *Streptomyces* spp. were sequenced (definitively more than in the other orders belonging to *Actinobacteria* phylum), only few Type I–IV GPA BGCs are known for this genus. These are A47934 BGCs from *Streptomyces toyocaensis* NRRL 15009 [48] and pekiskomycin BGCs from *Streptomyces* spp. WAC 04229 (WAC4229) and WAC1420 [49]. Our current analysis involved 1138 genome assemblies of *Streptomycetales* spp. (Supplementary Excel File S1), but no other novel BGC for Types I–IV GPA was found (Supplementary Excel File S2). On the contrary, BGCs for Type V GPAs and *feg*-like BGCs were found in 44 assemblies—all *Streptomyces* spp. except two *Kitasatospora* spp. Some of these BGCs were already reported [18], but we identified new ones (see Supplementary Excel File S2). *vlgs* were found exceptionally widespread in *Streptomycetales* spp.: more than one third of the analyzed genomes (418) contained *vlgs* (Table 1, Supplementary Excel File S2). Once again, we observed no correlation between the distribution of *vlgs* and of the GPA-like BGCs. More detailed analysis of 76 *Streptomycetales* spp. genomes (including 46 genomes carrying GPA- and *feg*-like BGCs) showed several different combinations, where certain strains carried a putative Type V GPA BGC and no *vlgs* (e.g., *Streptomyces fradiae* NKZ-259, Figure 5) along with strains carrying Type V GPA BGCs and a full set of *vlgs* (e.g., *Streptomyces* sp. NRRL WC-3897, Figure 5). Additionally, different combinations of *vlgs* were found in strains carrying no GPA-like BGCs. One peculiar trait of *Streptomycetales* spp. carrying the canonical *vanHAX-vanRS* operons was the presence of *vanK*, coding for an enzyme belonging to the Fem family, which adds the branch amino acid(s) to the stem pentapeptide of the peptidoglycan precursors carrying the D-Ala-D-Lac termini [50].

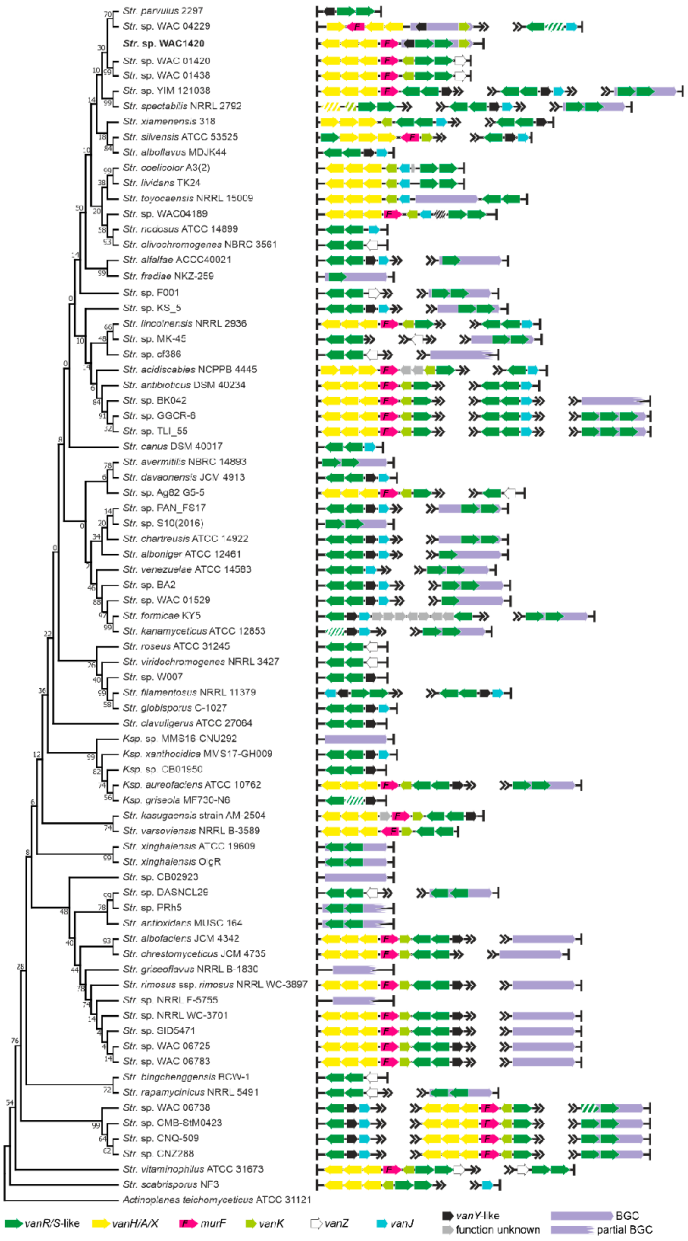


Figure 5. Organization of *vlgs* found within the set of 74 chosen genomes of *Streptomycetales* spp. Maximum Likelihood phylogenetic tree of the 16S rRNA genes of the corresponding species (see Section 4 for details) served as a phylogenetic framework for the scheme. Names of the *Streptomycetales* genera were abbreviated according to ESM Table S1. We were unable to detect the BGC for pekiskomycin within the published genome assembly of *Streptomyces* sp. WAC 01420, although the assembly contained *vlgs*; thus, *vlgs* in pekiksomycin BGC (as published in [49]) are given outside of the phylogenetic framework (highlighted in bold). Please refer to the text for the role of the single genes. Legend below the figure explains the color-coding of the scheme. Pseudogenes are shaded.

2.1.5. Occurrence of *vlgs* in GPA Non-Producing Groups

Order Actinomycetales. Although being known for various opportunistic animal and human pathogens, *Actinomycetales* spp. did not carry multiple *vlgs*. The only taxon (out of the 223 genome assembles screened, Supplementary Excel Files S1 and S2) carrying *vanYRS* genes was *Actinomycetales* bacterium JB111 (Table 1, Figure 6a).

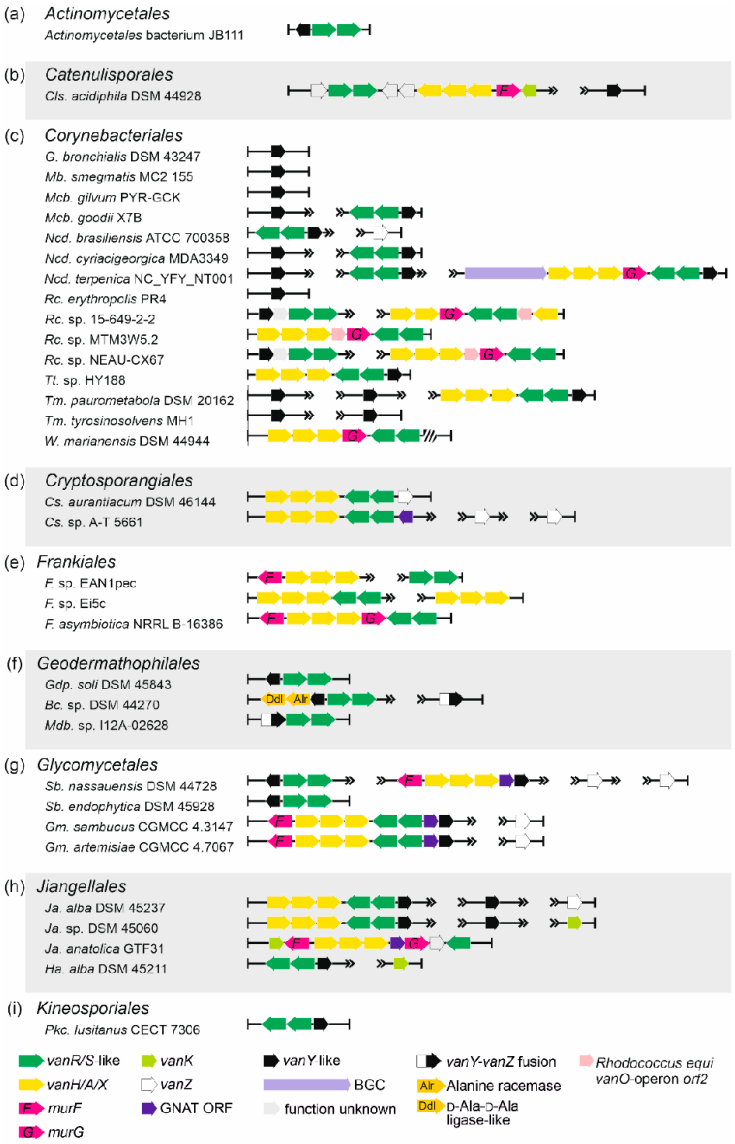


Figure 6. Arrangements of *vlgs* discovered in representative genomes from the orders (**a**) *Actinomycetales*, (**b**) *Catenulisporales*, (**c**) *Corynebacteriales*, (**d**) *Cryptosporangiales*, (**e**) *Frankiales*, (**f**) *Geodermathophilales*, (**g**) *Glycomycetales*, (**h**) *Jiangellales* and (**i**) *Kineosporiales*. Please refer to the main text for more details; genus names were abbreviated according to the ESM Table S1. The legend below the figure explains the color-coding of the scheme. Pseudogenes are shaded.

Order Catenulisporales. Only three genomes of *Catenulisporales* spp. were available, and in one of them—*Catenulispora acidiphila* DSM 44928—a complete set of *vanHAXRSY* was found (Table 1, Figure 6b, Supplementary Excel Files S1 and S2).

Order Coriobacteriales. Only two genomes out of the 217 available for *Coriobacteriales* spp. contained a complete set of *vanHAXRSY* (Table 1, Supplementary Excel Files S1 and S2; see the next paragraph for a more detailed description).

Order Corynebacteriales. *vlgs* appeared to be quite common in *Corynebacteriales* spp. (707 genomes available for screening) (Table 1, Figure 6c, Supplementary Excel Files S1 and S2). Remarkably, a full set of *vlgs* (*vanHAXRS*) was discovered within the genome of *Williamsia marianensis* DSM 44944—a species isolated from Mariana trench (10.898 m below the sea level) in 1998 [51]. Other *vlgs* combinations included *vanY*-like genes paired with a *vanRS*-like regulatory pair; "orphan" *vanY*-like genes; and *vanHAXRSY* genes, often accompanied with *murG* genes and homologues of the *vanO* operon *orf2* (see typical examples on Figure 6c). Peculiarly, during this analysis, a putative unknown GPA BGC was found in the genome of *Nocardia terpenica* NC_YFY_NT001, which is a clinical isolate derived from human cerebrospinal fluid (see CP023778 genome assembly information).

Order Cryptosporangiales. *vlgs* were found in two out of the three available genome assemblies of *Cryptosporangiales* spp. and were arranged as *vanHAXRS/HAXRSZ* (Table 1, Supplementary Excel Files S1 and S2); indeed, multiple copies of "orphan" *vanZ*-like genes were also present in the genome of *Cryptosporangium.* sp. A-T5661 (Figure 6d). In the genome of the latter, a gene coding for a GNAT was co-localized with *vanHAXRS*-genes (Figure 6d).

Order Eggerthellales. Only in one genome out of the 106 available for *Eggerthellales*, a *vlg* was found (Table 1, Supplementary Excel Files S1 and S2; see the next paragraph).

Order Frankiales. Approximately 15% of the analyzed *Frankiales* spp. genomes (46 in total, Supplementary Excel Files S1 and S2) possessed *vlgs*, arranged most often as *vanHAXRS* (and sometimes co-localized with *murF*- and *murG*-like genes) (Table 1, Figure 6e).

Order Geodermathophilales. A large portion of the analyzed *Geodermathophilales* spp. genomes (60, Supplementary Excel Files S1 and S2) carried *vlgs*, namely, *vanY*-like genes co-localized with *vanRS*-like regulatory pairs (Table 1). Rarely, *vanYRS*-like genes were found together with genes coding for a Ddl and for an alanine-racemase (as it was observed in *Alr. opalescens* DSM 45601, Figure 6f). Another unique feature (discovered only in *Geodermathophilales* spp.) was the presence of genes coding for putative VanY–VanZ fusion proteins (e.g., in *Modestobacter* sp. I12A-02628, Figure 6f).

Order Glycomycetales. *vlgs* were ubiquitously found in *Glycomycetales* spp. genomes (12 available in total), organized as *vanYRS*, *vanHAXY* or *vanHAXRSY* (Table 1, Supplementary Excel Files S1 and S2, Figure 6g). Latter arrangements were coupled with the genes coding for MurF and GNAT. Multiple copies of "orphan" *vanZ*-like genes were also found (Figure 6g).

Order Jiangellales. *vlgs* were found in more than 90% of the eleven analyzed *Jiangellales* spp. genomes (Table 1, Supplementary Excel Files S1 and S2). The most frequent arrangement was *vanHAXRSY,* although in *Jiangella anatolica* GTF31 *vanHAXS* genes were co-localized with *vanK* and with genes coding for MurF-, MurG- and GNAT-like proteins (Figure 6h).

Order Kineosporiales. A single set of *vlgs* in this order—*vanRSY*—was found in *Pseudokineococcus lusitanus* CECT 7306 (Figure 6i) among the twelve genomes analyzed (Table 1, Supplementary Excel Files S1 and S2).

Order Micrococcales. *vlg*s were found in less than 1% of the analyzed *Micrococcales* spp. genomes (a total of 1741, Supplementary Excel Files S1 and S2) and were represented mainly as *vanHAX, vanHAXRS* or *vanHAXRSY* arrangements (Table 1, Figure 7a). *vanZ*-like genes, coupled with *vanRS*-like regulatory pairs, were also observed as well as "orphan" *vanY*-like genes co-localized with *vanZ* genes (Figure 7a).

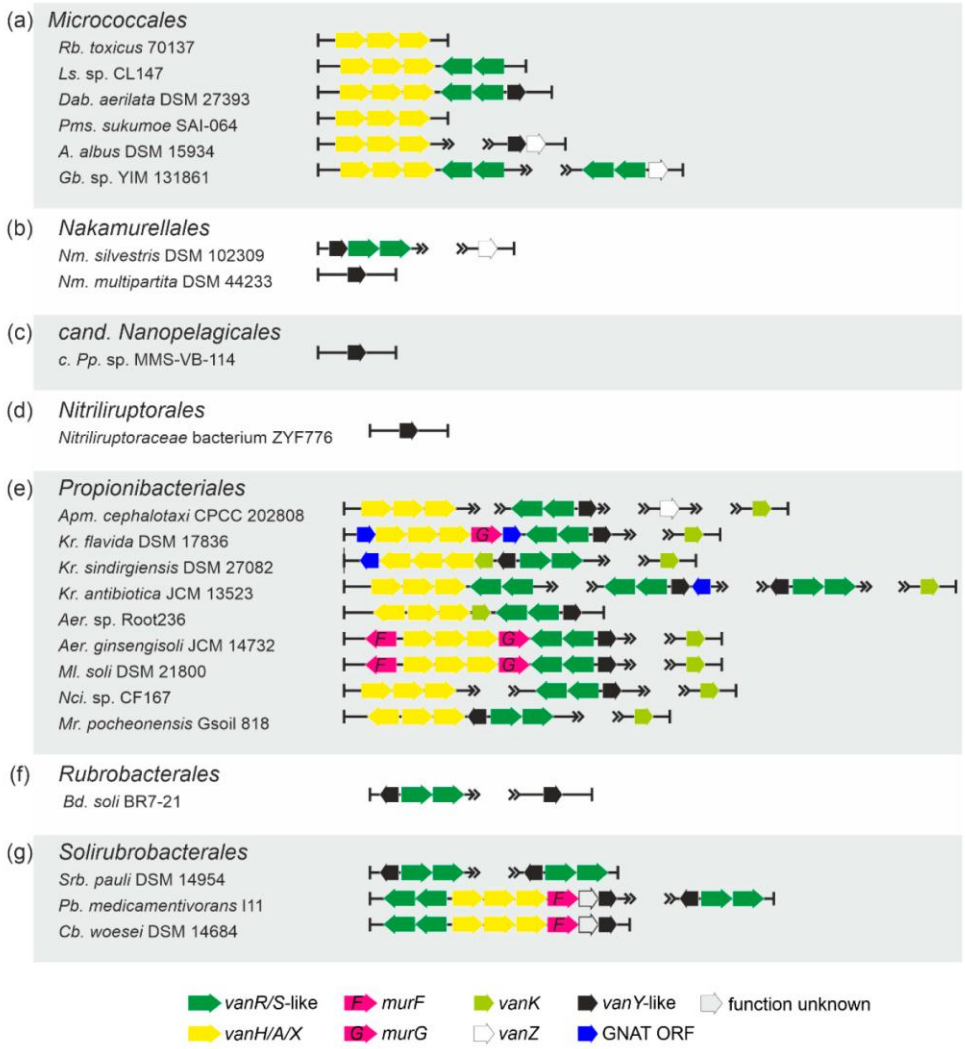


Figure 7. Arrangements of *vlgs* discovered in representative genomes from the orders (**a**) *Micrococcales*, (**b**) *Nakamurellales*, (**c**) *candidatus* Nanopelagicales, (**d**) *Nitriliruptorales*, (**e**) *Propionibacteriales*, (**f**) *Rubrobacterales* and (**g**) *Solirubrobacterales*. Please refer to the main text for more details; genus names were abbreviated according to the ESM Table S1. The legend below the figure explains the color-coding of the scheme.

Order Nakamurellales. *vlgs* were found within all the 6 genome assemblies of *Nakamurella* spp., either as *vanYRS* or as "orphan" *vanY*-like and *vanZ* genes (Table 1, Figure 7b, Supplementary Excel Files S1 and S2).

Order cand. Nanopelagicales and order Nitriliruptorales. In the few genome assemblies belonging to the species of both orders, only "orphan" *vanY*-like genes were rarely found (Table 1, Supplementary Excel Files S1 and S2, Figure 7c, d).

Order Propionibacteriales. *vlgs* were found in the genome assemblies belonging to few genera of *Propionibacteriales* (Table 1, Supplementary Excel File S2), although it was possible to analyze 593 genomes (Supplementary Excel File S1). There, *vlgs* exhibited

different arrangements, summarized in Figure 7e. Genes coding for MurF, MurG and GNAT proteins were often co-localized with *vlgs*.

Orders Rubrobacterales and Solirubrobacterales. Only a small portion of species, belonging to both orders, carried *vlgs* within their genomes (Table 1, Supplementary Excel Files S1 and S2). *vlgs* mainly were arranged as either *vanYRS* or *vanRSHAXY* (sometimes co-localized with a *murF* gene, Figure 7f, g).

2.1.6. Putatively Novel Transposable Elements Carrying *vlgs* in *Eggerthellales* and *Coriobacteriales* spp.

Analyzing the genomes of actinobacteria belonging to orders *Eggerthellales* and *Coriobacteriales*, we found *vlgs* in *Enterorhabdus mucosicola* NM66_B29, *Parvibacter caecicola* DSM 22242 and *Atopobium minutum* 10063974. When we examined the genetic neighborhood of these genes, it emerged that they might belong to the family of transposon-like mobile genetic elements (MGEs), involving multiple genes deputed to DNA transfer (Figure 8). Moreover, these *vlgs*-carrying putative MGEs were almost identical in *Er. mucosicola* NM66_B29 (order *Eggerthellales*) and *Pb. caecicola* DSM 22242 (order *Coriobacteriales*), while the MGE from *Atp. minutum* 10063974 (order *Coriobacteriales*) significantly differed from both (Figure 8). Transposons and other MGEs are believed to be one of the main sources of *van* genes dissemination throughout pathogens [52]. Thus, we decided to check whether MGEs from the three abovementioned species corresponded to those already known. For this, we compared putative integrases/recombinases from *Er. mucosicola* NM66_B29, *Pb. caecicola* DSM 22242 and *Atp. minutum* 10063974 with integrases found in known MGEs carrying *van* genes (ESM Table S2). It came out that the putative integrase from *Atp. minutum* 10063974 (EMZ42128) was identical to the integrase from *Enterococcus faecalis* transposon Tn1549 [53]. Further comparison of Tn1549 genes with the genes from *Atp. minutum* 10063974 MGE confirmed that they were identical. However, both integrases from *Er. mucosicola* NM66_B29 and *Pb. caecicola* DSM 22242 (EMZ42128 and MVX60893, respectively) were only slightly related to the transposases/integrases from Tn1549 and *Enterococcus faecium* insertion sequence IS1216V [54], sharing only an 18.6% and 14.6% amino acid sequence identity, respectively. Thus, MGE found in *Er. mucosicola* NM66_B29 and *Pb. caecicola* DSM 22242 might represent a novel MGE carrying *van* genes. Notably, this last putative MGE seemed to code for a VanYD protein [55], which is a D-Ala-D-Ala carboxypeptidase belonging to the penicillin-binding protein family, structurally unrelated to the VanY M15 peptidases (see below).

2.1.7. Occurrence of *vlgs* in *Bacillales* spp. (*Firmicutes* phylum)

Low G-C soil Gram-positives and in particular bacilli were considered as a one of possible sources of *vlgs* for pathogens, as it was considered that the ancestral *vanHAX* cluster might have evolved into one such species (e.g., those belonging to the *Paenibacillus* genus) and was then disseminated to pathogens in a transposon-mediated fashion [28]. We decided to test such a hypothesis by screening 2379 full genome assemblies of *Bacillales* spp. (Supplementary Excel File S3), available in GenBank, searching for *vlgs*. The results indicated that *vlgs* were quite rare in *Bacillales* spp. genomes: *vanHAXRS* were found in the genomes assemblies of *Brevibacillus laterosporus* E7593-50, *Thermoactinomyces vulgaris* and *Paenibacillus sonchi* LMG 24727 (Figure 9, Supplementary Excel File S3). *vanAX*-pseudogenes were also found in *Paenibacillus yonginensis* DCY84, while *vanHA*-genes (degraded to different extents) were found in 7 strains of *Paenibacillus larvae* (Figure 9, Supplementary Excel File S3). Only one *vanY*-like gene was found in *Bbac. laterosporus* E7593-50. Most of the *vlgs* were found adjacent to transposase-related genes, except the cases of *Pnb. yonginensis* DCY84, *Pnb. sonchi* LMG 24727 and *Tam. vulgaris* 2H; indeed, the latter strain is known to be naturally competent for exogenous DNA [56]. Overall, such findings indicated that the occurrence of *vlgs* in bacilli is not comparable to their distribution in most of the orders belonging to actinobacteria, making it highly unlikely for *van* genes to arise independently in bacilli.

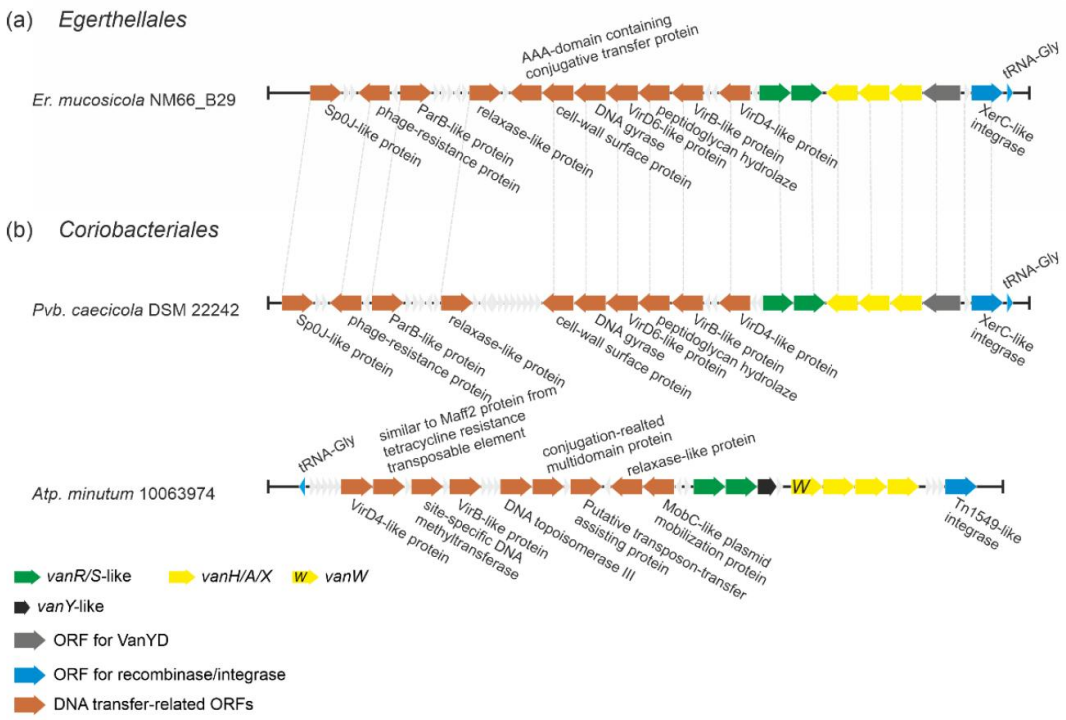


Figure 8. A scheme of putative MGEs found in (**a**) *Er. mucosicola* (*Eggerthellales*) and (**b**) *Coriobacteriales* spp. *Pvb. caecicola* and *Atp. minutum* (note that this MGE contains *vanW* gene – coding for a protein of unknown function). MGEs from *Er. mucosicola* and *Pvb. caecicola* seem to be almost identical (identical genes are joined with gray dashed lines), while MGE from *Atp. minutum* differs from both. The legend below the figure explains the color-coding of the scheme.

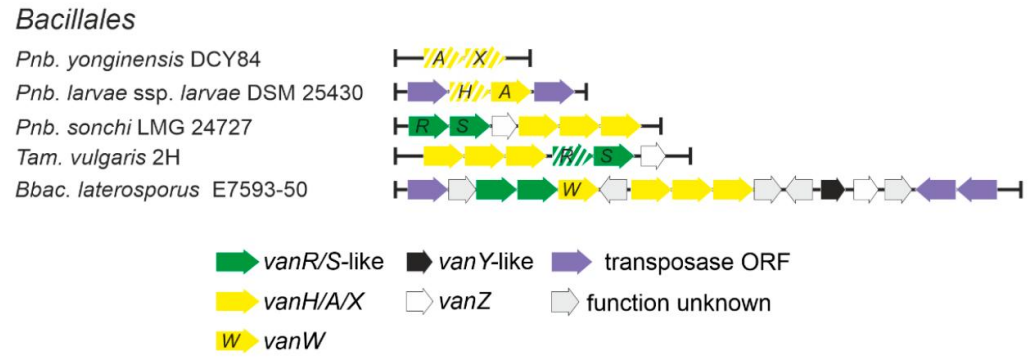


Figure 9. Arrangements of *vlgs* discovered in species belonging to *Bacillales* order. Please refer to the main text for more details; genus names were abbreviated according to ESM Table S1. The legend below the figure explains the color-coding of the scheme. Pseudogenes are shaded.

2.2. *Phylogeny of VanY-like Carboxypeptidases*

Surprisingly, *vanY*-like genes were the most common *vlgs* found in actinobacteria. We decided to reconstruct the phylogeny of the VanY-like proteins to comprehend such a variety and its relation to similar proteins described in low G-C Gram-positives, including pathogens such as GPA-resistant enterococci. For this purpose, we selected 251 proteins (see Section 4) coming either from actinobacteria or from low G-C Gram-positives, which, according to the MEROPS peptidase database [57,58], belonged to the M15B subfamily of the M15 family of metallopeptidases (mostly carboxypeptidases and dipeptidases). Other subfamilies are M15A, composed of a specific group of the so-called *Streptomyces*-type zinc-D-Ala-D-Ala carboxypeptidases [59], and M15D, to which VanX D-Ala-D-Ala dipeptidases belong (see the following section) [60,61]. To check that our selected proteins were VanY-like and no other carboxypeptidases, we controlled their sequence by CD-Search [58] and excluded those sharing the putative peptidoglycan-binding domain on the N-terminal region, which is typical of M15A proteins.

Reconstruction of the Maximum Likelihood phylogeny of VanY-like proteins from our dataset yielded a tree, where 5 distinct clusters were differentiated (Figure 10, ESM Figure S3). Cluster Y1 (Figure 10, ESM Figures S3 and S4) was the outgroup of the tree and contained VanY-like proteins originating from enterococci and other low G-C Gram-positives, consistently with previous reports [37]. Differently from what described a decade before [38], this cluster included also VanY-like proteins coming from different *Actinoplanes* spp., such as *Bd. soli* BR7-21, *H. rhizosphaerae* DSM 101727 and *All. opalescens* DSM 45601 (Figure 10, ESM Figure S4). This discrepancy is due to the increasing number of genomes that have become accessible in the meantime. Notably, all the genes corresponding to Y1 proteins in actinobacteria were 'orphan' (i.e., not co-localized with any other *vlgs*; see Figures 3, 4 and 7, and Supplementary Excel File S2).

In contrast, the VanY-like proteins from the large Y2 cluster were almost exclusively encoded by genes adjacent to the *vanRS*-like pairs and/or other *vlgs* (Figure 10, ESM Figure S5, Supplementary Excel File S2). They were found in the genomes of multiple orders of actinobacteria (Figure 10, ESM Figure S5, Supplementary Excel File S2). Importantly, Y2 included the VanY-like proteins from *Nonomuraea* GPA BGCs (including the most-studied $VanY_n$ from *N. gerenzanensis* ATCC 39727 [62]) and pekiskomycin BGCs (*Streptomyces* spp. WAC1420, WAC4229, WAC 04229 [49]).

The next clusters—Y3 and Y4 (Figure 10, ESM Figure S6)—were composed of VanY-like proteins coming from *Pseunocardiales* and *Corynebacteriales*, respectively, with the corresponding genes again being mostly 'orphan'.

Finally, Y5 was the last cluster defined within the tree of the VanY-like proteins. It was the biggest and most separated from the others, although the internal branching pattern was not completely clear, often lacking a trustable bootstrap support (Figure 10, ESM Figure S6). Nevertheless, genes corresponding to Y5 proteins were found either co-localized with different other *vlgs* or 'orphan'. The VanY-like proteins coded by *Pseudonocardiales* spp. GPA BGCs, including the most-studied $VanY_{Ab}$ from the balhimycin producer *Am. balhimycina* [63], formed a distinct subclade within Y5. The VanY-like proteins encoded within CA878, CA37, CA915, *auk* and *Ncd. terpenica* NC_YFY_NT001 BGCs were also found in Y5.

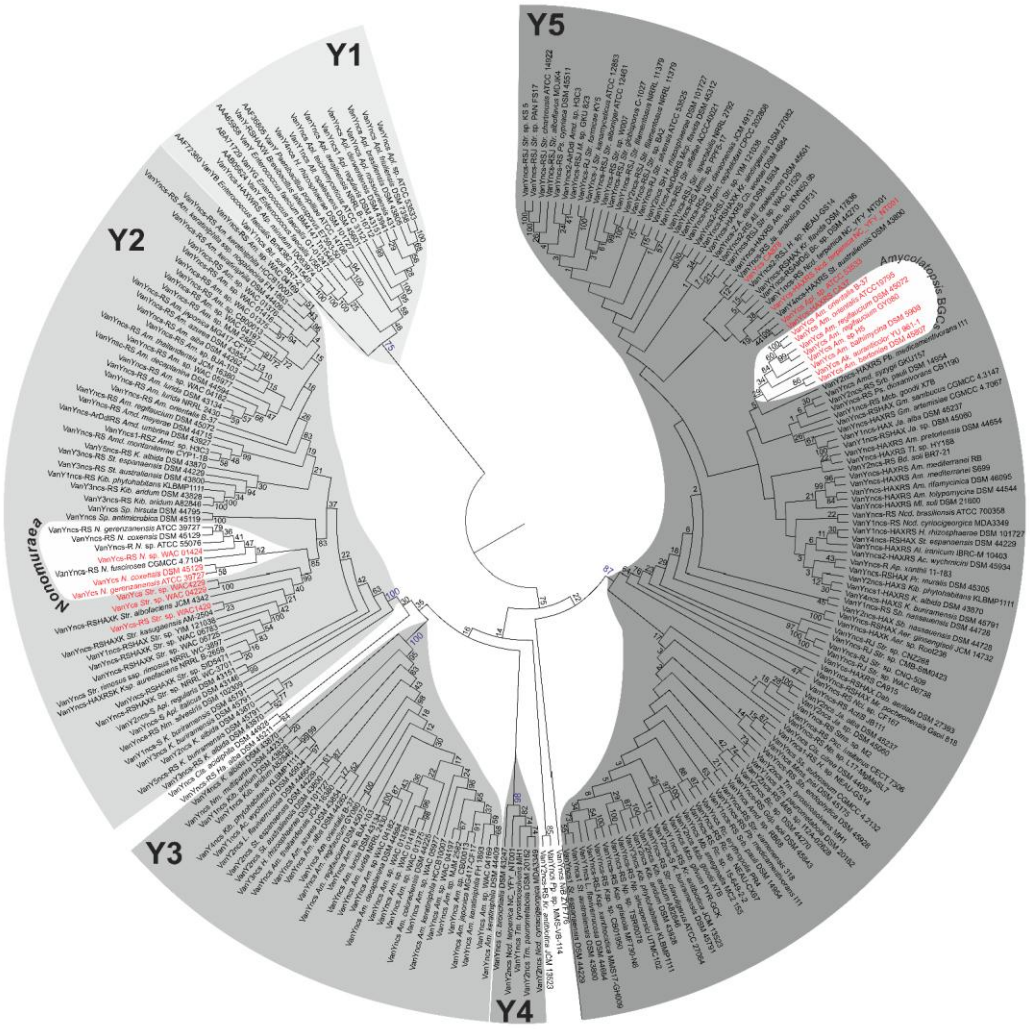


Figure 10. Maximum Likelihood phylogenetic tree of 251 VanY-like M15B carboxypeptidases. To show the topology of the tree better, branch lengths were ignored; the same tree drawn to scale is given in ESM Figures S3–S7. Five well supported clusters—Y1 to Y5—were distinguished on the tree. "cs/ncs" abbreviations in the label at the tip of each branch mean "cluster-situated/non-cluster-situated". BGC-encoded proteins are given in red. Importantly, the most-studied $VanY_n$ (Dbv7) from the A40926-producing *N. gerenzanensis* ATCC 39727 [62] belongs to Y2, while $VanY_{Ab}$ (coming from the balhimycin producer *Am. balhimycina* DSM 5908 [63]) belongs to the Y5 cluster.

2.3. Phylogeny of VanHAX

VanH lactate dehydrogenases. For the phylogenetic reconstruction of VanH, we used a dataset of 156 VanH proteins from actinobacteria and low G-C Gram-positives, with SCO2118 lactate dehydrogenase serving as an outgroup (see Section 4). It appeared that VanH proteins were quite conserved, and a branching pattern with completely reliable bootstrap support was difficult to obtain (ESM Figure S8). Nevertheless, few features could be presumed with certainty. First, VanH proteins coded within the putative MGEs of *Er. mucosicola* NM66_B29 and *Pvb. caecicola* DSM 22242 reliably out-grouped all other proteins (ESM Figure S8). Then, a well-defined cluster (named VH1, ESM Figure S8) was

composed of VanH proteins from pathogens and from low G-C soil Gram-positives. VH1 was not that close to the base of the tree, with multiple actinobacterial VanH proteins outgrouping it (ESM Figure S8). Another well-supported cluster—VH2 (ESM Figure S8)—was composed of VanH proteins coming from different *Streptomyces* spp., including those from A47934 BGC (*Str. toyocaensis* NRRL 15009) and pekiskomycin BGC (*Streptomyces* sp. WAC 04229). A third big cluster (VH3) was formed by VanH proteins coming from different GPA non-producing *Streptomyces* spp. Finally, VanH proteins coded within *tei*, *auk*, CA37 and CA915 also grouped together on the tree (ESM Figure S8).

VanA* D-*Ala*-D-*Lac-ligases. Overall, from previous works, it seems plausible that VanA D-Ala-D-Lac-ligases are a specialized evolutionary branch of D-Ala-D-Ala ligases (Ddl) involved in the "primary metabolism" of the cell wall in actinobacteria [64]. As reported above, Ddl-like proteins were also found encoded in metagenomic-sourced CA37, CA915 and in the *auk* BGCs; additionally, in the *Micromonosporales* we found a peculiar putative *pdx* operon (see Section 2.1.3) composed of genes for a PALP threonine dehydratase, a Ddl-like ligase and a VanX-like dipeptidase, adjacent to a *vanRS*-like regulatory pair. Thus, here we decided to test the phylogeny of the VanA ligases together with the above-mentioned Ddl-like proteins on a background of 'house-keeping' Ddl-ligases from the main actinobacterial orders. The protein set (see Section 4) used for this phylogenetic reconstruction contained 153 VanA ligases from actinobacteria and low G-C Gram-positives, 12 Ddl-like ligases from *Micromonosporales* spp., the 3 Ddl-like ligases from CA37, CA915 and *auk* BGCs, as well as 81 'house-keeping' actinobacterial Ddl-ligases.

In the resulting phylogenetic tree (ESM Figure S9), the 'house-keeping' Ddl-ligases formed a number of distinct clades that corresponded to the orders of origin (ESM Figure S9). There, Ddl-like ligases coded in CA915, CA37 and *auk* BGCs grouped together with "house-keeping" Ddl-ligases from order *Eggerthellales*, while Ddl-like ligases from *Micromonosporales* (coded in putative *pdx*-operons) formed a distinct clade among the other clades of 'house-keeping' Ddl-ligases (ESM Figures S9 and S10). Most strikingly, clades for VanA ligases from low G-C Gram-positives (ESM Figure S11) and actinobacteria (ESM Figure S12) (both out-grouped by VanA ligases coded within the putative MGEs of *Er. mucosicola* NM66_B29 and *Pvb. caecicola* DSM 22242) clustered together with 'house-keeping' Ddl-ligases from the order *Coriobacteriales* (ESM Figure S9).

Composition of the main actinobacterial VanA-clade (ESM Figures S9 and S12) required further comments. The resolution of this clade was high enough to distinguish four well-supported clusters—VA1–4. VA1 was formed by the VanA ligases from GPA non-producing streptomycetes together with the one coded in A47934 GPA BGC from *Str. toyocaensis* NRRL 15009. Other VanA ligases from GPA non-producing streptomycetes were composed of VA2, while VA3 contained proteins from different groups of actinobacteria. Finally, VA4 was formed by VanA ligases coded within GPA BGCs of *Amycolatopsis* spp. VanA ligases from CA915, CA37 and *tei* BGCs grouped together on the tree. Notably, VanA ligases from pekiskomycin BGCs (from *Str.* sp. WAC 04229 and *Str.* sp. WAC1420) grouped together with VanA from *Str. varsoviensis* NRRL B-3589, which does not carry any GPA BGC.

VanX M15D dipeptidases. According to the MEROPS database, VanX dipeptidases belong to the same M15 family as VanY carboxypeptidases, but group in the M15D subfamily. The dataset used for the phylogenetic reconstruction contained 155 VanX dipeptidases coded within different *vanHAX* operons from actinobacteria and low G-C Gram-positives and 12 VanX-like dipeptidases coded in the putative *pdx* operon found in *Micromonosporales*. The obtained phylogenetic tree revealed that VanX-like dipeptidases coded in the *pdx* operon and VanX proteins from low G-C Gram-positives formed well-supported clusters (VX1 and VX2, respectively, ESM Figure S13). All other actinobacterial VanX-like dipeptidases were out-grouped by the VanX-like dipeptidases coded within the putative MGEs of *Er. mucosicola* NM66_B29 and *Pvb. caecicola* DSM 22242 (ESM Figure S13). Unfortunately, internal topology of the latter clade had non-optimal bootstrap support, since the VanX-proteins appeared well conserved (ESM Figure S14). However, it was possible to

distinguish two clusters—VX3 and VX4. Interestingly, the composition of VX3 and VX4 corresponded to the clusters VA4 and VA3, respectively, of th eVanA-like ligases tree (ESM Figure S12). Finally, the VanX dipeptidases coded in pekiskomycin BGCs (from *Str.* sp. WAC 04229 and *Str.* sp. WAC1420) again grouped together with VanX from *Str. varsoviensis* NRRL B-3589, which is not a GPA producer.

The *van* genes from pathogenic enterococci often code a peculiar group of bifunctional D,D-dipeptidases/D,D-carboxypeptidases, known as VanXY [65]. In 2014, a comparative structural study of VanXY, VanX and VanY [57] assumed VanXY peptidases to be a specialized evolutionary branch of VanY carboxypeptidases in pathogens. This assumption was supported with a phylogenetic reconstruction of the M15 family dipeptidases [57]. However, rather few protein sequences were available at that time. Thus, we decided to check the phylogeny of the VanXY peptidases in relation to our VanX and VanY datasets simultaneously. We used a dataset of 425 proteins, including 7 VanXY-peptidases. The topology of the resulting tree correlated with the topologies of the trees received for the VanX and VanY datasets, separately. At the same time, the VanXY peptidases appeared to root deeply within the VanY clade (ESM Figure S15) from low G-C Gram-positives (corresponding to the Y1 cluster of VanY carboxypeptidases, Figure 10). Therefore, our large-scale reconstruction was in line with the previously made assumption about the origin of VanXY [57].

2.4. Phylogeny of VanRS-like Two-Component Regulatory Pairs

In the course of our screen for *vlgs*, many of them were found co-localized with *vanRS*-like regulatory pairs. *vanRS*-like regulatory pairs were also found adjacent to (i) a putative *pdx* operon in *Micromonosporales*; (ii) putative operons formed with genes coding for alanine/aspartate racemase, Ddl and VanY-like carboxypeptidase (such as in *All. opalescens* DSM 45601 and *Mcp. flavida* DSM 45312, respectively, Figure 3); or (iii) genes for β-lactamases (such as in *Xa. phaseoli* DSM 45730, Figure 4). Additionally, multiple BGCs for Type V GPAs and *feg*-like BGCs from *Streptomyces, Microbispora* and *Actinomadura* spp., as well as for Type IV GPAs from *Nonomuraea* spp., were found to carry *vanRS*-like regulatory pairs. To clarify if, how and to what extent all these VanR-like response regulators and VanS-like sensor histidine kinases are related to each other and to the VanRS proteins from low G-C Gram-positives, we reconstructed separate phylogenies of the two datasets: one for the VanR- and another for the VanS-like proteins. The first dataset contained 295 proteins, while the second was composed of 313 proteins. Such a discrepancy in numbers derives from the fact that the *vanRS*-like pairs often lacked one of the genes or had it impaired (as pseudogene). Overall, the final topology of both trees was quite coherent, implying the co-evolution of the VanR- and VanS-like proteins coded within one gene pair (Figure 11); this allowed us to define a set of well-supported clusters (named VS in the VanS and VR in VanR trees). Both trees showed a set of clusters formed as basal clades as well as defined crown groups. Basal clades formed VR1–4 clusters plus a Dbv6-like cluster in the VanR phylogenetic tree, and VS1–3 clusters plus a Dbv22-like cluster in the VanS phylogenetic tree. These data were consistent with what was previously reported on Dbv6/Dbv22, which is a cluster-situated two-component regulatory pair in A40926 BGC from *N. gerenzanensis*, not grouping with the "classic" VanRS-like proteins from other GPA BGCs [66].

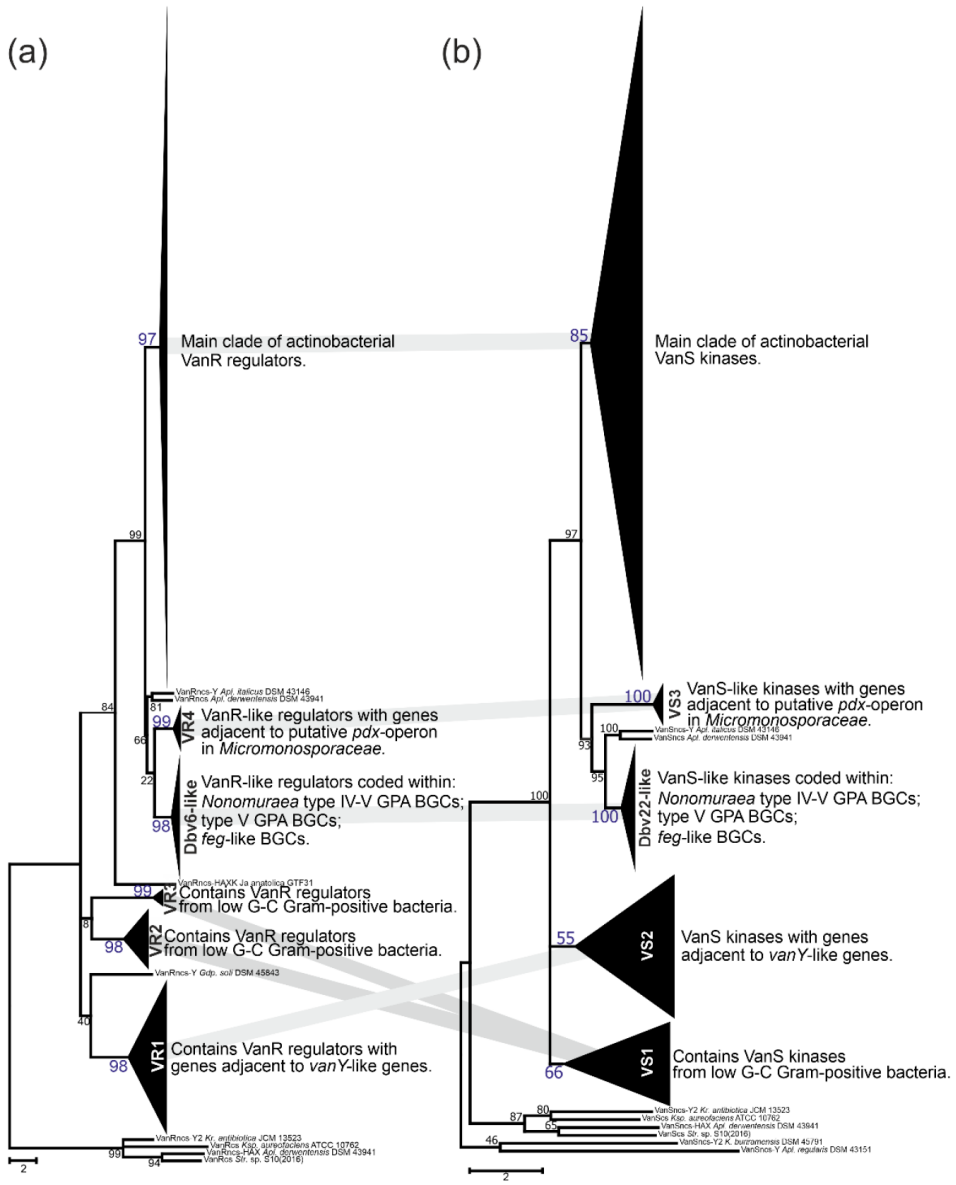


Figure 11. Phylogenetic trees of actinobacterial VanR-like response regulators (**a**) and VanS-sensor histidine kinases (**b**). Defined clusters on each tree were collapsed; for the expanded versions, please refer to ESM Figures S16–31. Coherent clusters are joined with thick grey lines. Please see main text for more details. Scale bar represents the number of substitutions per site.

Proteins from coherent clusters VR1 (VanR-tree, Figure 11a, ESM Figure S16) and VS2 (VanS-tree, Figure 11b, ESM Figure S17) had genes co-localized with *vanY*-like genes, coding for Y2-cluster VanY-like proteins (Figure 10); the overall topology of Y2 was found similar to both VR1 and VS2 (ESM Figures S4, S16 and S17). Then, the VR2+3 and VS1 clusters were also coherent and contained VanS- and VanR-like proteins found in low G-C Gram-positives, as well as the ones coded within the MGEs from *Coriobacteriales* and

Eggerthellales spp. (Figure 11, ESM Figures S18 and S19). Next pair of coherent clusters contained VanS- and VanR-like proteins coded within *feg*-like BGCs, Type V GPA BGCs and Type IV GPA BGCs from *Nonomuraea* spp. (Figure 11). Basically, both clades were formed with orthologues of either Dbv6 or Dbv22 from *dbv* BGCs, thus they received the corresponding naming (ESM Figures S20 and S21). Finally, the last pair of coherent basal clusters were VR4 and VS3 (Figure 11), formed by VanR- and VanS-like proteins coded adjacent to the putative *pdx* operon from *Micromonosporales* (ESM Figures S22 and S23).

Similar to what was previously observed for VanX, VanA and VanH phylogenies, the resolution of VanR and VanS trees crown groups was not perfect (ESM Figures S24 and S25). Nevertheless, we defined three additional clusters in the VanR-phylogenetic tree—VR5–7 (ESM Figures S24 and S26). VanR regulators from clusters VR5–6 were coded adjacent to *vanY*-like genes in different actinomycetes, while VR7 was composed of VanR-regulators coded adjacent to *vanHAXYK* in *Streptomyces* spp. VR7 contained VanR coded within pekiskomycin BGC from *Str.* sp. WAC1420. The resolution of the VanS tree crown group was better, allowing to distinguish five additional clusters—VS4–8 (ESM Figure S25). Here, the VS4 and VS5 clusters were formed by VanS-kinases coming from different actinobacterial orders and coded adjacent to the *vanY*-like genes and *vanZ*-like genes, respectively (ESM Figures S27 and S28). Then, proteins forming the VS6 cluster were coded adjacent to the *vanHAX* genes in different actinobacteria (ESM Figure S29), while VS7 was formed by VanS kinases coded adjacent to *vanHAXYK* genes from *Streptomyces* spp., including the one from pekiskomycin BGC (*Str.* sp. WAC1420, ESM Figure S30). Finally, proteins from VS8 were coded adjacent to *vanHAXK* genes in streptomycetes; VS8 also included VanS from model *Str. coelicolor* (ESM Figure S31).

Some other notable features in the VanS and VanR crown groups phylogenies require comments. First, both reconstructions placed VanS and VanR proteins encoded within *tei* and A47934 BGCs together (ESM Figures S25 and S26). Second, both trees showed evidence of a possible evolution of the VanRS-like regulatory pair, expanding its regulon control from *van* genes to some other genes: although the VanRS pair from *All. opalescens* DSM 45601 and *Mcp. flavida* DSM 45312 were related to the VanRS proteins coded adjacent to the *vanHAX* genes, the corresponding *vanRS*-like gene pairs actually were co-localized with genes for alanine or aspartate racemases, together with *ddl* and *vanY*-like genes.

3. Discussion

In the current work we aimed to address certain unclear issues about *van* genes, such as their distribution and phylogeny. Although we are aware that our results might risk generating more questions than answers, we tried in the following section to summarize what are, in our opinion, the most relevant findings.

Actinobacteria are the most likely primary sources of vlgs. First of all, the results of our screens, covering more than 7000 actinobacterial genomes and 2000 *Bacillales* genomes, revealed that *vlgs* are abundant within the *Actinobacteria* phylum (with an incidence of ca. 13%), while vanishingly rare in bacilli and related species belonging to the *Firmicutes* phylum. This disproves the idea of *van*-like genes and operons emerging independently in soil-dwelling actinobacteria and bacilli [28]. Abundance and context variability of actinobacterial *vlgs* point to the *Actinobacteria* phylum as to the original source of *van* genes. At the same time, the assumption that ubiquitous low G-C soil Gram-positives served as a bridge for *vlgs* to arrive in pathogens [28] seems likely. Such a transfer was probably achieved via MGEs, which often were proved to carry *vlgs* in low G-C soil Gram-positives and pathogens [52]. In fact, we found the classical Tn1549 transposon in the actinobacterium *Atp. minutum* 10063974, even though until now Tn1549-like transposons were described only in enterococci [53]. It is hard to say whether *Atp. minutum* 10063974 might represent the original actinobacterial source of Tn1549, or if this is an example of reverse HGT event: G-C content (estimated from *vanRSYWHAX* genes) of *Atp. minutum* 10063974 Tn1549 is the same as in *Enterococcus faecalis* BM4382 Tn1549—47%. However, this is comparable to the overall genome G-C content of *Atp. minutum* 10063974, which is ca.

48%. At the same time, two novel, putative MGEs (very similar to each other) were also found in the genomes of the actinobacteria *Er. mucosicola* NM66_B29 and *Pvb. caecicola* DSM 22242. These MGEs coded unusual transposases, carrying *vlgs* coding proteins that in our phylogenetic reconstructions did not cluster with Van proteins derived from low G-C Gram-positives. Once again, it is not clear if these MGEs might represent the 'original' actinobacterial *vlg*-carrying elements. The G-C content of MGEs from *Er. mucosicola* NM66_B29 and *Pvb. caecicola* DSM 22242 (as estimated from the G-C content of their *vanRSHAXY-D*-genes) is ca. 55%. This is higher than the usual G-C content of MGEs from low G-C Gram-positives, but lower than the overall genome G-C content of *Er. mucosicola* NM66_B29 and *Pvb. caecicola* DSM 22242 (64.6% and 62.4%, respectively). Thus, the study of *vlgs* in the soil mobilome requires more detailed and focused research, which could in future contribute to a better understanding of how *vlgs* were disseminated from actinobacteria.

vlgs are distributed in actinobacteria without any evident strict correlation to GPA BGCs distribution, although a complex co-evolution with BGCs likely occurred. Our comparative genomic analysis of the different orders belonging to the *Actinobacteria* phylum showed that *vlgs* are not necessarily co-localized with Type I–IV GPA BGCs; moreover, in the majority of the cases, *vlgs* were found in GPA non-producers as well as in Type V GPA producers. Consequently, it is reasonable to assume that non-cluster-situated *vlgs* existed independently from GPA BGCs and might actually be considered a preadaptation feature, which then facilitated the spread of GPA BGCs within the *Actinobacteria* phylum. Next, GPA cluster-situated *vlgs* are not monophyletic, meaning that different GPA BGCs likely acquired these genes from different sources of the vast actinobacterial pool. All our phylogenetic reconstructions in fact showed that cluster-situated *vlgs* emerged randomly on the trees, surrounded by non-cluster-situated *vlgs*. One of the most prominent lines of evidence came from the VanY phylogeny, where *Nonomuraea* spp. and *Amycolatopsis* spp. GPA cluster-encoded VanY carboxypeptidases belonged to distant clusters (Y2 and Y5, Figure 9), separated by multiple other non-cluster-encoded VanY proteins.

Moreover, *vlgs* and GPA BGCs seem to have a shared, complex co-evolution pattern. We were able to reconstruct one of the most obvious ones, observed for the self-resistance phenotype in the A40926 producer *N. gerenzanensis*, which relied on the expression of the cluster-situated *vanY*-like gene—*dbv7*. In the corresponding *dbv7* knockout mutant, GPA resistance was significantly reduced, but not completely abolished [62]. Screening the full genome of *N. gerenzanensis* [67], the reason became evident: an additional *vanY* allele was found far away the *dbv* BGC. Interestingly, this allele was co-localized with a *vanRS*-like regulatory pair, whereas *dbv7* was not. The same situation was observed for *N. coxensis* DSM 45129—a recently described producer of A40926-like Type IV GPA [40,68]: a cluster-situated *dbv7* homologue (almost identical) and a non-cluster-situated *vanRSY*-triad were identified in its genome. The two other screened genomes of *Nonomuraea* spp.—the GPA non-producing *N. fuscirosea* CGMCC 4.7104 and the kistamicin producer *Nonomuraea* sp. ATCC 55076—contained close homologues of the *vanRSY*-triad. In addition, the genome of the other putative GPA producer *Nonomuraea* sp. WAC 01424 presented the GPA BGC just upstream the *vanRSY*-triad, while no *vanY*-allele was found in the BGC itself. Finally, in our reconstructed phylogeny of VanY-like carboxypeptidases, the *dbv* and *noc* cluster-encoded VanY-proteins out-grouped (Y1 cluster) all other ones (Figure 10). We believe that such a picture is as a result of a series of HGT events. In our hypothesis, summarized in Figure 12, first, *noc* and *dbv* protoclusters recruited the *vanY* gene from the *vanRSY* triad present in the common ancestor of *Nonomuraea* spp., whereas the WAC 01424 and kistamycin ancestral BGCs did not acquire it. Then, *noc* and *dbv,* through another HGT event, were introduced in the *Nonomuraea* ancestors already carrying the *vanRSY* triad. Another HGT event delivered WAC 01424 BGC to *Nonomuraea* sp. WAC 01424 and kistamycin BGC to *Nonomuraea* sp. ATCC 55076, explaining why the *vanRSY* triad is the only resistance determinant in these strains. Finally, no such HGT event occurred in *N. fuscirosea* CGMCC 4.7104, leaving this strain without any GPA BGCs but still carrying the *vanRSY* triad.

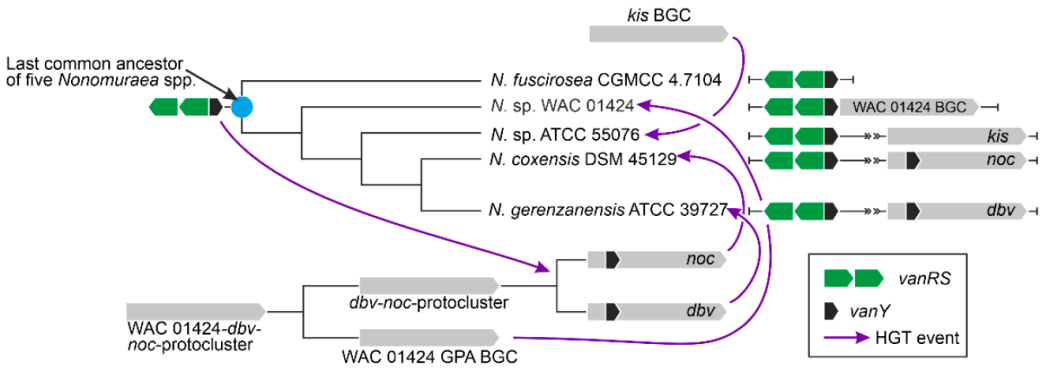


Figure 12. Scheme for the discussed scenario for *vanY*-like genes and GPA BGCs co-evolution in *Nonomuraea* spp. Please refer to the main text for more details.

Is the GPA resistance the original function of vlgs? As reported above, actinobacteria are an extremely abundant source of *vlgs*. It is quite difficult to consider that such a variety of *vlgs* in GPA non-producing actinobacteria is needed to protect them versus the GPA producers, which definitively seemed quite rare. Alternatively, *vlgs* might have natural functions in cell-wall remodeling in response to environmental and/or developmental triggers. For instance, VanY-like M15B carboxypeptidases are extremely abundant in actinobacteria, forming diverged evolutionary lineages; it is tempting to suspect their role in, for instance, cell-wall remodeling during the life cycle (a function indeed shown for some other D-Ala-D-Ala carboxypeptidase [69]). Our finding of the so-called *pdx* operon, common for *Micromonosporales*, contributed to such an idea. This putative operon is composed of a PALP-like serine-threonine dehydratase, a Ddl ligase and a VanX-like dipeptidase genes, co-localized with a *vanRS*-like regulatory pair. This organization suggests an alternative mechanism of peptidoglycan precursor remodeling, which reminds the introduction of D-Ala-D-Ser termini in the cell wall of some enterococci, showing a low level of GPA resistance [14]. It is unknown whether the *pdx* operon is functional and if its expression leads to some cell wall remodeling, but such a case indeed merits further experimental evaluation.

To conclude, we must say that the current work only scratched the surface of actinobacterial *vlgs*, leaving many issues without a complete final evaluation and understanding. However, we hope that this should provoke other in silico, in vitro and in vivo studies, which will shed light on such an important question such as GPA resistance. Our findings portray actinobacteria as a Pandora's box, hosting myriads of putative GPA resistance genes that might be transferred sooner or later to pathogens, significantly contributing to AMR. A thorough understanding of GPA resistance in actinobacteria may prove useful in the future surveillance of emerging mechanisms of resistance to clinically used GPAs. Although further experiments are necessary to show that the discovered in silico putative *vlgs* have a real function in vivo conferring GPA resistance, their study may reveal new insights into their biological functions in actinobacteria, augmenting our comprehension of this remarkable phylum.

4. Methods

4.1. Routine Analysis of Nucleic and Amino Acid Sequences

All routine analytic work with nucleic acid and amino acid sequences was performed using Geneious 4.8.5 [70], Mega X [71]; routine amino acid sequence alignments were performed with Clustal Omega (EMBL-EBI) [72].

4.2. *vlgs* Search Pipeline

To perform the search for *van*-like genes (named *vlgs*), all genome assemblies of *Actinobacteria*—available at the time of work preparation (April 2020) as either full sequences or uncompleted ones—were retrieved from the GenBank database (in few exceptions, genome assemblies were retrieved from RefSeq database), for a total of 7108 assemblies from species belonging to 26 established and 2 tentative orders within the *Actinobacteria* phylum, represented by 653,111 corresponding nucleic acid sequences. A full list of the genome assemblies and corresponding nucleic acid sequences is given in Supplementary Excel Table S1. The MultiGeneBlast tool [73] was utilized to screen the chosen set of genome assemblies for *vlgs*. To do that, the chosen set of genome assemblies belonging to each order was downloaded from GenBank (or RefSeq) in a genomic GenBank format (*.gbff). These files were then used to create offline MultiGeneBlast custom databases for each order belonging to the *Actinobacteria* phylum. MultiGeneBlast was run in a "homology" mode with the default settings, which included 25% minimal sequence coverage of the BLAST hits and 30% minimal amino acid sequence identity of BLAST hits. The maximum distance between genes in locus was increased to 40 kb (considering that *vlgs* inside a BGC might be separated by some other genes—such as in A47934 BGC [48]). Two input files were used as queries for the MultiGeneBlast search: one included *van* genes from *Str. coelicolor*—*SCO3589-3590-3594-3595-3596* (*vanSRHAX*); and the other included balhimycin BGC-situated *van* genes from *Am. balhimycina*—*DMA12_00360, DMA12_00365* and *DMA12_00370* (*vanS*, *vanR* and *vanY* respectively). The first input files allowed to detect *vanHAX* orthologues co-localized with the *vanRS*-like regulatory pairs, while the second helped to detect cases when the *vanY*-like genes were co-localized with *vanRS* but lacked other *van* genes in their genetic neighborhood. MultiGeneBlast outputs were then carefully examined, and the amino acid sequences of the proteins coded with the so-identified *vlgs* were extracted. The information about these *vlgs*, including the corresponding protein accession numbers, nucleic acid accession numbers and taxa, are summarized in Supplementary Excel Table S2. To refine this initial screening, chosen sets of genomes for each order (highlighted in red in Supplementary Excel Table S2) were manually reexamined for *vlgs* using BLASTP [74] with SCO3589 (VanS), SCO3590 (VanR), SCO3592 (VanJ), SCO3593 (VanK), SCO3594 (VanH), SCO3595 (VanA), SCO3596 (VanX), CAG25753 (VanY) and ELS50663 (VanZ) as queries. These last sets of genomes were chosen to cover all the genera within a certain order and to include all the genomes carrying known and putative Type I–V GPA BGCs as well as *feg*-like BGCs. Selected hits were tested for orthology with queries using the Reciprocal Best Hit BLAST approach. Information received here was used to build Figures 2–9.

4.3. Search for Putative GPA-like BGCs

The MultiGeneBlast tool [73] was used to screen all the genome assemblies for GPA-like BGCs. We utilized the same offline custom databases created for *vlgs* screening. MultiGeneBlast was also run in "homology" mode with the default settings; however, the maximum distance between genes in one locus was increased to 60 kb. An input file was composed of teicoplanin BGC [75] genes—*tei4**, *tei8**, *tei15**, *tei17**, *tei23**, *tei24**, *tei28** and *tei29**—coding for an ABC transporter, teicoplanin halogenase, StrR-like transcriptional regulator, prephenate dehydrogenase, L-4-hydroxyphenylglycine biosynthesis enzyme HpgT, type III polyketide synthase DpgA (involved in the biosynthesis of both L-4-hydroxyphenylglycine and L-3,5-dihydroxyphenylglycine), HmaS and Hmo (L-3,5-dihydroxyphenylglycine biosynthesis enzymes), respectively. All these genes have their orthologues in BGCs for Type I–V GPAs and in *feg*. MultiGeneBlast outputs were manually examined and the nucleic acid sequences containing MultiGeneBlast hits were applied for upstream antiSMASH [76] analysis. A list of the obtained putative GPA- and *feg*-like BGCs is given in Supplementary Excel Table S2.

4.4. Phylogenetic Reconstruction

Since the screening revealed hundreds of *vlgs*, it was difficult to use all the sequences for a comprehensive phylogenetic analysis. Therefore, phylogenies were reconstructed for sets of proteins coded with *vlgs* from the chosen genomes for each order (highlighted in red in Supplementary Excel Table S2). The final protein datasets for the reconstruction of the phylogenies of the VanY-like carboxypeptidases—VanH; VanA and Ddl; VanX and VanX-like dipeptidases; VanY-like carboxypeptidases, VanX and VanX-like dipeptidases; VanR-like regulators; and VanS-like sensor histidine kinases—are given in Supplementary FASTA Files S1–S7, respectively. Some additional information was coded in the name of each protein sequence to indicate (i) whether this protein is BGC-encoded or not; and (ii) whether this protein is coded with an 'orphan' gene or the corresponding gene is co-localized with other *vlgs*. For instance, "VanYncs-HAXRS_Tt_sp_HY188" indicates that this is VanY-like non-BGC-encoded peptidase, with the corresponding gene co-localized with *vanHAXRS*, coming from *Tomitella* sp. HY188.

The Mega X [71] package was used to perform the phylogenetic reconstructions. On the road to the representative phylogenetic tree, we always followed the next algorithm. First, multiple amino acid sequence alignments for each dataset were generated using Muscle; the obtained alignments were manually curated to preserve as much meaningful data as possible. Then the curated multiple sequence alignments were analyzed using the Mega X model finder to discover the most appropriate evolutionary models, and the best-scoring models were applied to generate the Maximum Likelihood phylogenies for each protein dataset. A similar approach was used to generate the 16S rRNA gene phylogenetic trees, which can be found in Figures 2–5. Final topologies of either the protein or gene trees were derived from 500 bootstraps.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/antibiotics10121533/s1: Supplementary Excel File S1, Summary of actinobacterial genome assemblies with corresponding nucleic acid sequences analyzed in this work. Supplementary Excel File S2, Accession numbers and locations of Van-like proteins and GPA BGCs, discovered in this work for actinobacteria. Supplementary Excel File S3, Summary of *Bacillales* genome assemblies analyzed with corresponding nucleic acid sequences; accession numbers of Van-like proteins discovered for *Bacillales* spp. Supplementary FASTA Files S1–S7, containing protein datasets for reconstructing phylogenies of different Van-like proteins, see Section 4.4. Supplementary Tables and Figures File, containing ESM Table S1, Abbreviations of *Actinobacteria* and *Bacillales* genera names, used throughout the work; ESM Table S2, List of transposases from known MGEs carrying *vlgs* used for comparing with the transposases encoded within *Atopobium minutum* 10063974, *Enterorhabdus mucosicola* NM66_B29 and *Parvibacter caecicola* DSM 22242 MGEs; ESM Figure S1, Defining the genes annotated within previously unannotated regions flanking the CA878 GPA BGC; ESM Figure S2, Defining the genes annotated within previously unannotated flanks of CA915 and CA37 GPA BGCs; ESM Figure S3, Phylogenetic tree showing the overall phylogeny of VanY-like M15B carboxypeptidase dataset; ESM Figure S4, Expanded clade corresponding to the Y1 cluster from ESM Figure S3; ESM Figure S5, Expanded clade corresponding to the Y2 cluster from ESM Figure S3; ESM Figure S6, Expanded clades corresponding to the Y3 and Y4 clusters from ESM Figure S3; ESM Figure S7, Expanded clade corresponding to the Y5 cluster from ESM Figure S3; ESM Figure S8, Phylogenetic tree showing the overall phylogeny of VanH-dataset; ESM Figure S9, Phylogenetic tree showing the overall phylogeny of the dataset composed with VanA proteins and Ddl-ligases; ESM Figure S10, Expanded clade of Ddl-ligases coded in putative pdx-operons; ESM Figure S11, Expanded clade containing VanA-ligases from low G-C Gram-positive bacteria; ESM Figure S12, Expanded crown group of VanA-phylogenetic tree; ESM Figure S13, Phylogenetic tree showing the overall phylogeny of the dataset composed with VanX-like proteins; ESM Figure S14, Expanded version of the crown group of actinobacterial VanX M15D dipeptidases; ESM Figure S15, Phylogeny of combined datasets for VanY-like M15B carboxypeptidases and VanX M15D dipeptidases, including also VanXY-proteins; ESM Figure S16, Expanded version of the VR1 cluster from the phylogenetic tree of actinobacterial VanR-like response regulators; ESM Figure S17, Expanded version of the VS2 cluster from the phylogenetic tree of actinobacterial VanS-like sensor histidine kinases; ESM Figure S18, Expanded versions of the VR2 and VR3 clusters from the phylogenetic tree of actinobacterial VanR-like response regulators; ESM

Figure S19, Expanded version of the VS1 cluster from the phylogenetic tree of actinobacterial VanS-like sensor histidine kinases; ESM Figure S20, Expanded version of the Dbv6-like cluster from the phylogenetic tree of actinobacterial VanR-like response regulators; ESM Figure S21, Expanded version of the Dbv22-like cluster from the phylogenetic tree of actinobacterial VanS-like sensor histidine kinases; ESM Figure S22, Expanded version of the VR4 cluster from the phylogenetic tree of actinobacterial VanR-like response regulators; ESM Figure S23, Expanded version of the VS3 cluster from the phylogenetic tree of actinobacterial VanS-like sensor histidine kinases; ESM Figure S24, Expanded version of the clade corresponding to the crown group of bacterial VanR response regulators; ESM Figure S25, Expanded version of the crown group of bacterial VanS-like sensor histidine kinases; ESM Figure S26, Expanded versions of the VR5-7 clusters of the phylogenetic tree from ESM Figure S24; ESM Figure S27, Expanded version of the VS4 cluster of the phylogenetic tree from ESM Figure S25; ESM Figure S28, Expanded version of the VS5 cluster of the phylogenetic tree from ESM Figure S25; ESM Figure S29, Expanded version of the VS6 cluster of the phylogenetic tree from ESM Figure S25; ESM Figure S30, Expanded version of the VS7 cluster of the phylogenetic tree from ESM Figure S25; ESM Figure S31, Expanded version of the VS8 cluster of the phylogenetic tree from ESM Figure S25.

Author Contributions: Conceptualization, V.F., F.M. and O.Y.; formal analysis, E.B. and O.Y.; investigation, A.A.-V. and E.B.; writing—original draft, F.M. and O.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the public grant "Fondo di Ateneo per la Ricerca" 2019 and 2020 to F.M. and by the BG-09F grant of Ministry of Education and Science of Ukraine to V.F.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data are available from the corresponding author upon reasonable request.

Acknowledgments: A.A.-V. is a Ph.D. student of the "Life Science and Biotechnology" course at Università degli Studi dell'Insubria.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bull. World Health Organ.* **2001**, *79*, 780–790. [PubMed]
2. O' Neil, J. Review on Antibiotic resistance. Antimicrobial resistance: Tackling a crisis for the health and wealth of nations. *Health Wealth Nations* **2014**, 1–16. Available online: https://wellcomecollection.org/works/rdpck35v (accessed on 21 October 2021).
3. Ghosh, S.; Bornman, C.; Zafer, M.M. Antimicrobial resistance threats in the emerging COVID-19 pandemic: Where do we stand? *J. Infect. Public Health* **2021**, *14*, 555–560. [CrossRef]
4. Marcone, G.L.; Binda, E.; Berini, F.; Marinelli, F. Old and new glycopeptide antibiotics: From product to gene and back in the post-genomic era. *Biotechnol. Adv.* **2018**, *36*, 534–554. [CrossRef] [PubMed]
5. Nicolaou, K.C.; Boddy, C.N.C.; Bräse, S.; Winssinger, N. Chemistry, biology, and medicine of the glycopeptide antibiotics. *Angew. Chem. Int. Ed.* **1999**, *38*, 2096–2152. [CrossRef]
6. Sarkar, P.; Yarlagadda, V.; Ghosh, C.; Haldar, J. A review on cell wall synthesis inhibitors with an emphasis on glycopeptide antibiotics. *Medchemcomm* **2017**, *8*, 516–533. [CrossRef] [PubMed]
7. Vollmer, W.; Blanot, D.; De Pedro, M.A. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **2008**, *32*, 149–167. [CrossRef] [PubMed]
8. Perkins, H.R. Specificity of combination between mucopeptide precursors and vancomycin or ristocetin. *Biochem. J.* **1969**, *111*, 195–205. [CrossRef] [PubMed]
9. Nitanai, Y.; Kikuchi, T.; Kakoi, K.; Hanamaki, S.; Fujisawa, I.; Aoki, K. Crystal structures of the complexes between vancomycin and cell-wall precursor analogs. *J. Mol. Biol.* **2009**, *385*, 1422–1432. [CrossRef] [PubMed]
10. Williams, D.H. The glycopeptide story-how to kill the deadly "Superbugs". *Nat. Prod. Rep.* **1996**, *13*, 469–477. [CrossRef]
11. Marschall, E.; Cryle, M.J.; Tailhades, J. Biological, chemical, and biochemical strategies for modifying glycopeptide antibiotics. *J. Biol. Chem.* **2019**, *294*, 18769–18783. [CrossRef]
12. Parenti, F.; Cavalleri, B. Proposal to name the vancomycin-ristocetin like glycopeptides as dalbaheptides. *J. Antibiot.* **1989**, *42*, 1882–1883. [CrossRef] [PubMed]

13. Yushchuk, O.; Binda, E.; Marinelli, F. Glycopeptide antibiotic resistance genes: Distribution and function in the producer actinomycetes. *Front. Microbiol.* **2020**, *11*, 1173. [CrossRef] [PubMed]
14. Binda, E.; Marinelli, F.; Marcone, G.L. Old and new glycopeptide antibiotics: Action and resistance. *Antibiotics* **2014**, *3*, 572–594. [CrossRef] [PubMed]
15. Arthur, M.; Quintiliani, R.J. Regulation of VanA-and VanB-type glycopeptide resistance in enterococci. *Antimicrob. Agents Chemother.* **2001**, *45*, 375–381. [CrossRef] [PubMed]
16. Arthur, M.; Molinas, C.; Bugg, T.D.H.; Wright, G.D.; Walsh, C.T.; Courvalin, P. Evidence for *in vivo* incorporation of D-lactate into peptidoglycan precursors of vancomycin-resistant enterococci. *Antimicrob. Agents Chemother.* **1992**, *36*, 867–869. [CrossRef] [PubMed]
17. Stegmann, E.; Frasch, H.J.; Kilian, R.; Pozzi, R. Self-resistance mechanisms of actinomycetes producing lipid II-targeting antibiotics. *Int. J. Med. Microbiol.* **2015**, *305*, 190–195. [CrossRef] [PubMed]
18. Waglechner, N.; McArthur, A.G.; Wright, G.D. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. *Nat. Microbiol.* **2019**, *4*, 1862–1871. [CrossRef] [PubMed]
19. Culp, E.J.; Waglechner, N.; Wang, W.; Fiebig-Comyn, A.A.; Hsu, Y.P.; Koteva, K.; Sychantha, D.; Coombes, B.K.; Van Nieuwenhze, M.S.; Brun, Y.V.; et al. Evolution-guided discovery of antibiotics that inhibit peptidoglycan remodelling. *Nature* **2020**, *578*, 582–587. [CrossRef]
20. Xu, M.; Wang, W.; Waglechner, N.; Culp, E.J.; Guitor, A.K.; Wright, G.D. GPAHex-A synthetic biology platform for Type IV–V glycopeptide antibiotic production and discovery. *Nat. Commun.* **2020**, *11*, 5232. [CrossRef]
21. Mitchell, S.J.; Verma, D.; Griswold, K.E.; Bailey-Kellogg, C. Building blocks and blueprints for bacterial autolysins. *PLoS Comput. Biol.* **2021**, *17*, e1008889. [CrossRef] [PubMed]
22. Cheng, A.V.; Wuest, W.M. Phylogeny-guided approach yields glycopeptides with unique action. *Trends Pharmacol. Sci.* **2020**, *41*, 297–299. [CrossRef] [PubMed]
23. Gonsior, M.; Mühlenweg, A.; Tietzmann, M.; Rausch, S.; Poch, A.; Süssmuth, R.D. Biosynthesis of the peptide antibiotic feglymycin by a linear nonribosomal peptide synthetase mechanism. *Chem. Biol. Chem.* **2015**, *16*, 2610–2614. [CrossRef] [PubMed]
24. Hong, H.J.; Paget, M.S.B.; Buttner, M.J. A signal transduction system in *Streptomyces coelicolor* that activates expression of a putative cell wall glycan operon in response to vancomycin and other cell wall-specific antibiotics. *Mol. Microbiol.* **2002**, *44*, 1199–1211. [CrossRef]
25. Patel, R.; Piper, K.; Cockerill, F.R.; Steckelberg, J.M.; Yousten, A.A. The biopesticide *Paenibacillus popilliae* has a vancomycin resistance gene cluster homologous to the enterococcal VanA vancomycin resistance gene cluster. *Antimicrob. Agents Chemother.* **2000**, *44*, 705–709. [CrossRef] [PubMed]
26. Fontana, R.; Ligozzi, M.; Pedrotti, C.; Padovani, E.M.; Cornaglia, G. Vancomycin-resistant *Bacillus circulans* carrying the *vanA* gene responsible for vancomycin resistance in enterococci. *Eur. J. Clin. Microbiol. Infect. Dis.* **1997**, *16*, 473–474. [CrossRef] [PubMed]
27. Marshall, C.G.; Broadhead, G.; Leskiw, B.K.; Wright, G.D. D-Ala-D-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6480–6483. [CrossRef] [PubMed]
28. Aminov, R.I.; Mackie, R.I. Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol. Lett.* **2007**, *271*, 147–161. [CrossRef] [PubMed]
29. One Health. Available online: https://www.who.int/news-room/q-a-detail/one-health (accessed on 21 October 2021).
30. Hutchings, M.I.; Hong, H.J.; Buttner, M.J. The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*. *Mol. Microbiol.* **2006**, *59*, 923–935. [CrossRef]
31. Hong, H.J.; Hutchings, M.I.; Neu, J.M.; Wright, G.D.; Paget, M.S.B.; Buttner, M.J. Characterization of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (*vanK*) required for drug resistance. *Mol. Microbiol.* **2004**, *52*, 1107–1121. [CrossRef] [PubMed]
32. Adamek, M.; Alanjary, M.; Sales-Ortells, H.; Goodfellow, M.; Bull, A.T.; Winkler, A.; Wibberg, D.; Kalinowski, J.; Ziemert, N. Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC Genom.* **2018**, *19*, 1–15. [CrossRef]
33. Shawky, R.M.; Puk, O.; Wietzorrek, A.; Pelzer, S.; Takano, E.; Wohlleben, W.; Stegmann, E. The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains *bbr*, encoding a StrR-like pathway-specific regulator. *J. Mol. Microbiol. Biotechnol.* **2007**, *13*, 76–88. [CrossRef] [PubMed]
34. Gudeta, D.D.; Moodley, A.; Bortolaia, V.; Guardabassi, L. VanO, a new glycopeptide resistance operon in environmental *Rhodococcus equi* isolates. *Antimicrob. Agents Chemother.* **2014**, *58*, 1768–1770. [CrossRef]
35. Mengin-Lecreulx, D.; Texier, L.; Rousseau, M.; Van Heijenoort, J. The *murG* gene of *Escherichia coli* codes for the UDP-*N*-acetylglucosamine: *N*-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol *N*-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. *J. Bacteriol.* **1991**, *173*, 4625–4636. [CrossRef] [PubMed]
36. Banik, J.J.; Craig, J.W.; Calle, P.Y.; Brady, S.F. Tailoring enzyme-rich environmental DNA clones: A source of enzymes for generating libraries of unnatural natural products. *J. Am. Chem. Soc.* **2010**, *132*, 15661–15670. [CrossRef] [PubMed]
37. Goldstein, B.P.; Selva, E.; Gastaldo, L.; Berti, M.; Pallanza, R.; Ripamonti, F.; Ferrari, P.; Denaro, M.; Arioli, V.; Cassani, G. A40926, a new glycopeptide antibiotic with anti-*Neisseria* activity. *Antimicrob. Agents Chemother.* **1987**, *31*, 1961–1966. [CrossRef] [PubMed]

38. Binda, E.; Marcone, G.L.; Pollegioni, L.; Marinelli, F. Characterization of VanYn, a novel D, D-peptidase/D, D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727. *FEBS J.* **2012**, *279*, 3203–3213. [CrossRef]
39. Binda, E.; Marcone, G.L.; Berini, F.; Pollegioni, L.; Marinelli, F. *Streptomyces* spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance. *BMC Biotechnol.* **2013**, *13*, 24. [CrossRef]
40. Yushchuk, O.; Vior, N.M.; Andreo-Vidal, A.; Berini, F.; Rückert, C.; Busche, T.; Binda, E.; Kalinowski, J.; Truman, A.W.; Marinelli, F. Genomic-led discovery of a novel glycopeptide antibiotic by *Nonomuraea coxensis* DSM 45129. *ACS Chem. Biol.* **2021**, *16*, 915–928. [CrossRef]
41. Nazari, B.; Forneris, C.C.; Gibson, M.I.; Moon, K.; Schramma, K.R.; Seyedsayamdost, M.R. *Nonomuraea* sp. ATCC 55076 harbours the largest actinomycete chromosome to date and the kistamicin biosynthetic gene cluster. *Medchemcomm* **2017**, *8*, 780–788. [CrossRef]
42. Yushchuk, O.; Ostash, B.; Truman, A.W.; Marinelli, F.; Fedorenko, V. Teicoplanin biosynthesis: Unraveling the interplay of structural, regulatory, and resistance genes. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 3279–3291. [CrossRef]
43. Bardone, M.R.; Paternoster, M.; Coronelli, C. Teichomycins, new antibiotics from *Actinoplanes teichomyceticus* nov. sp. II. Extraction and chemical characterization. *J. Antibiot.* **1978**, *31*, 170–177. [CrossRef] [PubMed]
44. Yim, G.; Kalan, L.; Koteva, K.; Thaker, M.N.; Waglechner, N.; Tang, I.; Wright, G.D. Harnessing the synthetic capabilities of glycopeptide antibiotic tailoring enzymes: Characterization of the UK-68, 597 biosynthetic cluster. *Chem. Biol. Chem.* **2014**, *15*, 2613–2623. [CrossRef]
45. Debono, M.; Merkel, K.E.; Molloy, R.M.; Barnhart, M.; Prestí, E.; Hunt, A.H.; Hamill, R.L. Actaplanin, new glycopeptide antibiotics produced by *Actinoplanes missouriensis* the isolation and preliminary chemical characterization of actaplanin. *J. Antibiot.* **1984**, *37*, 85–95. [CrossRef] [PubMed]
46. Yushchuk, O.; Homoniuk, V.; Ostash, B.; Marinelli, F.; Fedorenko, V. Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. *J. Antibiot.* **2020**, *73*, 255–259. [CrossRef]
47. Owen, J.G.; Reddy, B.V.B.; Ternei, M.A.; Charlop-Powers, Z.; Calle, P.Y.; Kim, J.H.; Brady, S.F. Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11797–11802. [CrossRef]
48. Pootoolal, J.; Thomas, M.G.; Marshall, C.G.; Neu, J.M.; Hubbard, B.K.; Walsh, C.T.; Wright, G.D. Assembling the glycopeptide antibiotic scaffold: The biosynthesis of A47934 from *Streptomyces toyocaensis* NRRL15009. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8962–8967. [CrossRef] [PubMed]
49. Thaker, M.N.; Wang, W.; Spanogiannopoulos, P.; Waglechner, N.; King, A.M.; Medina, R.; Wright, G.D. Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* **2013**, *31*, 922–927. [CrossRef]
50. Hong, H.J.; Hutchings, M.I.; Hill, L.M.; Buttner, M.J. The role of the novel fem protein VanK in vancomycin resistance in *Streptomyces coelicolor*. *J. Biol. Chem.* **2005**, *280*, 13055–13061. [CrossRef]
51. Pathom-Aree, W.; Nogi, Y.; Sutcliffe, I.C.; Ward, A.C.; Horikoshi, K.; Bull, A.T.; Goodfellow, M. *Williamsia marianensis* sp. nov., a novel actinomycete isolated from the Mariana Trench. *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 1123–1126. [CrossRef]
52. Hegstad, K.; Mikalsen, T.; Coque, T.M.; Werner, G.; Sundsfjord, A. Mobile genetic elements and their contribution to the emergence of antimicrobial resistant *Enterococcus faecalis* and *Enterococcus faecium*. *Clin. Microbiol. Infect.* **2010**, *16*, 541–554. [CrossRef] [PubMed]
53. Garnier, F.; Taourit, S.; Glaser, P.; Courvalin, P.; Galimand, M. Characterization of transposon Tn1549, conferring VanB-type resistance in *Enterococcus* spp. *Microbiol.* **2000**, *146*, 1481–1489. [CrossRef]
54. Handwerger, S.; Skoble, J. Identification of chromosomal mobile element conferring high-level vancomycin resistance in *Enterococcus faecium*. *Antimicrob. Agents Chemother.* **1995**, *39*, 2446–2453. [CrossRef]
55. Reynolds, P.E.; Ambur, O.H.; Casadewall, B.; Courvalin, P. The VanYD DD-carboxypeptidase of *Enterococcus faecium* BM4339 is a penicillin-binding protein. *Microbiol.* **2001**, *147*, 2571–2578. [CrossRef]
56. Hopwood, D.A.; Wright, H.M. Transformation in *Thermoactinomyces vulgaris*. *J. Gen. Microbiol.* **1972**, *71*, 383–398. [CrossRef]
57. Meziane-Cherif, D.; Stogios, P.J.; Evdokimova, E.; Savchenko, A.; Courvalin, P. Structural basis for the evolution of vancomycin resistance D,D-peptidases. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5872–5877. [CrossRef]
58. Rawlings, N.D.; Waller, M.; Barrett, A.J.; Bateman, A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **2014**, *42*, D503–D509. [CrossRef] [PubMed]
59. Charlier, P.; Wery, J.-P.; Dideberg, O.; Frère, J.-M. *Streptomyces albus* G D-Ala-D-Ala carboxypeptidase. *Handb. Met.* **2006**. [CrossRef]
60. Marshall, C.G.; Lessard, I.A.D.; Park, I.S.; Wright, G.D. Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. *Antimicrob. Agents Chemother.* **1998**, *42*, 2215–2220. [CrossRef] [PubMed]
61. Bussiere, D.E.; Pratt, S.D.; Katz, L.; Severin, J.M.; Holzman, T.; Park, C.H. The structure of VanX reveals a novel amino-dipeptidase involved in mediating transposon-based vancomycin resistance. *Mol. Cell* **1998**, *2*, 75–84. [CrossRef]
62. Marcone, G.L.; Beltrametti, F.; Binda, E.; Carrano, L.; Foulston, L.; Hesketh, A.; Bibb, M.; Marinelli, F. Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother.* **2010**, *54*, 2465–2472. [CrossRef] [PubMed]
63. Kilian, R.; Frasch, H.J.; Kulik, A.; Wohlleben, W.; Stegmann, E. The VanRS homologous two-component system $VnlRS_{Ab}$ of the glycopeptide producer *Amycolatopsis balhimycina* activates transcription of the $vanHAX_{Sc}$ genes in *Streptomyces coelicolor*, but not in *A. balhimycina*. *Microb. Drug Resist.* **2016**, *22*, 499–509. [CrossRef]

64. Mainardi, J.-L.; Villet, R.; Bugg, T.D.; Mayer, C.; Arthur, M. Evolution of peptidoglycan biosynthesis under the selective pressure of antibiotics in Gram-positive bacteria. *FEMS Microbiol. Rev.* **2008**, *32*, 386–408. [CrossRef] [PubMed]
65. Podmore, A.H.B.; Reynolds, P.E. Purification and characterization of VanXYc, a D,D-dipeptidase/D,D-carboxypeptidase in vancomycin-resistant *Enterococcus gallinarum* BM4174. *Eur. J. Biochem.* **2002**, *269*, 2740–2746. [CrossRef] [PubMed]
66. Alduina, R.; Tocchetti, A.; Costa, S.; Ferraro, C.; Cancemi, P.; Sosio, M.; Donadio, S. A two-component regulatory system with opposite effects on glycopeptide antibiotic biosynthesis and resistance. *Sci. Rep.* **2020**, *10*, 6200. [CrossRef]
67. D'Argenio, V.; Petrillo, M.; Pasanisi, D.; Pagliarulo, C.; Colicchio, R.; Talà, A.; De Biase, M.S.; Zanfardino, M.; Scolamiero, E.; Pagliuca, C.; et al. The complete 12 Mb genome and transcriptome of *Nonomuraea gerenzanensis* with new insights into its duplicated "magic" RNA polymerase. *Sci. Rep.* **2016**, *6*, 18. [CrossRef] [PubMed]
68. Yushchuk, O.; Andreo-Vidal, A.; Marcone, G.L.; Bibb, M.; Marinelli, F.; Binda, E. New molecular tools for regulation and improvement of A40926 glycopeptide antibiotic production in *Nonomuraea gerenzanensis* ATCC 39727. *Front. Microbiol.* **2020**, *11*, 8. [CrossRef]
69. Rioseras, B.; Yagüe, P.; López-Garciá, M.T.; Gonzalez-Quiñonez, N.; Binda, E.; Marinelli, F.; Manteca, A. Characterization of SCO4439, a D-alanyl-D-alanine carboxypeptidase involved in spore cell wall maturation, resistance, and germination in *Streptomyces Coelicolor*. *Sci. Rep.* **2016**, *6*, 21659. [CrossRef] [PubMed]
70. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef]
71. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef] [PubMed]
72. Sievers, F.; Higgins, D.G. Clustal Omega. *Curr. Protoc. Bioinform.* **2014**, *2014*, 3.13.1–3.13.16. [CrossRef]
73. Medema, M.H.; Takano, E.; Breitling, R. Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol. Biol. Evol.* **2013**, *30*, 1218–1223. [CrossRef] [PubMed]
74. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
75. Li, T.L.; Huang, F.; Haydock, S.F.; Mironenko, T.; Leadlay, P.F.; Spencer, J.B. Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: Characterization of two glycosyltransferases and the key acyltransferase. *Chem. Biol.* **2004**, *11*, 107–119. [CrossRef] [PubMed]
76. Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S.Y.; Medema, M.H.; Weber, T. AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47*, W81–W87. [CrossRef]

Оригінал доступний за посиланням:
https://www.mdpi.com/2079-6382/10/12/1533.
Електронні додатки до статті доступні за посиланням:
https://www.mdpi.com/article/10.3390/antibiotics10121533/s1.

Анотація: *Поширення резистентності до протимікробних препаратів створює виклик глобальній системі охорони здоров'я, роблячи багато раніше успішних класів антибіотиків неефективними. На жаль, цей перелік також включає глікопептидні антибіотики (ГПА), такі як ванкоміцин і тейкопланін, які наразі вважаються препаратами вибору. Виникнення резистентності до ГПА загрожує обмежити клінічне використання цього класу антибіотиків – нашої останньої лінії захисту від мультирезистентних (МЛР) Грам-позитивних патогенів. Широко визнано, що детермінанти резистентності до ГПА – van-гени– могли походити від продуцентів ГПА, таких як ґрунтові Грам-позитивні актинобактерії, де використовуються для самозахисту. У поточній роботі ми здійснюємо комплексне біоінформатичне дослідження поширення та філогенії детермінантів стійкості до ГПА в межах типу Actinobacteria. Цікаво, що van-подібні гени (vlg) були розподілені в різних порядків актинобактерій, і не лише серед тих, що продукують ГПА, але й у не-продуцентів: більше 10% проаналізованих геномів актинобактерій містили один або кілька vlg, тоді як менше 1% мали кластери біосинтетичних генів ГПА. За допомогою ряду філогенетичних реконструкцій наші результати показують коеволюцію різних vlg, вказуючи на те, що найбільш розповсюдженими є ті, які кодують карбоксипептидази типу VanY і які знайдені поодиноко або поруч з регуляторною парою vanS/R.*

8.3 Особливості філогенії ABC-транспортерів, кодованих в кластерах біосинтетичних генів пептидних антибіотиків

Важливим компонентом КБГ ГПА та подібних антибіотиків є гени трансмембранних експортерів. Їх продукти щонайперше задіяні у експорті антибіотиків. Цілком ймовірно, це не єдина їх функція: КБГ інших пептидних антибіотиків також кодують експортери, що часто відіграють роль у формуванні резистентності продуцента до власного продукту, як наприклад для випадку лантибіотика нізину [174]. На жаль, цей аспект до сьогодні не розглядався для продуцентів ГПА. Схоже, що гени експортерів присутні в усіх секвенованих КБГ ГПА [175]. Однак, на даний момент лише Tba з КБГ баліміцину в *Am. balhimycina* DSM 5908 [176]. З'ясовано, що нокаут гена *tba* веде до зниження позаклітинної концентрації баліміцину, але до підвищення його внутрішньоклітинної концентрації; вплив нокаута на стійкість до ГПА не вивчався [176].

У цьому підрозділі ми вирішили проаналізувати розповсюдження та особливості філогенії ABC-транспортерів кодованих в КБГ ГПА і споріднених пептидних антибіотиків, таких як рамопланін чи фегліміцин, а також потенційних КБГ, виявлених в ході скринінгу, представленому в попередньому **підрозділі 7.2**. Всього ми проаналізували 102 КБГ ГПА та споріднених пептидів, та виявили ортологи *tba* у 95 із них; інші, очевидно, також містять ортологи *tba,* але їх неможливо було ідентифікувати через погану якість нуклеотидних послідовностей. На наш подив, послідовності всіх відповідних білків виявилися дуже консервативними, незалежно від структурного різноманіття сполук, в експорті яких вони ймовірно задіяні. Також, всі ці білки мали монофілетичне походження. Схоже, що досліджені ABC-транспортери є доволі універсальними та здатні експортувати різні пептидні антибіотики незалежно від їх структури.

Отримані в цьому підрозділі результати є основою для подальшого експериментального вивчення цих транспортерів, зокрема їх потенційної ролі в формуванні стійкості до ГПА.

Результати підрозділу опубліковано в [177].

ISSN 0206-5657. Вісник Львівського університету. Серія біологічна. 2022. Випуск 86. С. 33–46
Visnyk of the Lviv University. Series Biology. 2022. Issue 86. P. 33–46

ГЕНЕТИКА

УДК 573:575:579.6 HTTPS://DOI.ORG/10.30970/VLUBS.2022.86.03

ФІЛОГЕНІЯ БІЛКІВ-ЕКСПОРТЕРІВ ГЛІКОПЕПТИДНИХ І ДЕЯКИХ СПОРІДНЕНИХ АНТИБІОТИКІВ

О. Ющук, К. Жукровська, В. Федоренко*

Львівський національний університет імені Івана Франка
вул. Грушевського, 4, Львів 790005, Україна
**e-mail: viktor.fedorenko@lnu.edu.ua*

Глікопептидні антибіотики (ГПА) – це група сполук, які продукуються актинобактеріями – Грам-позитивними ґрунтовими бактеріями з Г-Ц-багатими геномами. Серед глікопептидів є ефективні антибіотики, які застосовуються у клініці як «препарати вибору» в боротьбі з мультирезистентними Грам-позитивними патогенами. Нещодавні дослідження показали, що такі пептиди як рамопланін і фегліміцин генетично споріднені з ГПА, незважаючи на значні відмінності в їхній хімічній структурі. Біосинтез ГПА закодований у кластерах біосинтетичних генів (КБГ), які мають багато спільних рис. Однією з них є наявність генів, що кодують ABC-транспортери. Основною функцією цих транспортерів є експортування пептидних антибіотиків із клітин продуцентів. КБГ і в інших актиноміцетів-продуцентів антибіотиків також несуть гени ABC-транспортерів, які відіграють важливу роль у стійкості продуцентів до власних токсичних продуктів. Не менш важливим є і те, що маніпуляції з генами транспортерів є перспективним біотехнологічним знаряддям для збільшення продукції антибіотиків. Обидва цих аспекти вивчені мало для ГПА і споріднених антибіотиків. Доступність значної кількості секвенованих геномів актинобактерій дає змогу зібрати переконливий масив КБГ як відомих, так і потенційних ГПА, а також споріднених із ними антибіотиків. У цій роботі ми поставили за мету проаналізувати *in silico* особливості розповсюдження, структури та філогенії ABC-транспортерів зі 102 КБГ ГПА та споріднених пептидних антибіотиків. Ми виявили, що ABC-транспортери з КБГ ГПА, фегліміцину і рамопланіну, а також багатьох КБГ невідомих сполук, мають подібну архітектуру й амінокислотні послідовності. Усі вони належать до групи MdlB(MsbA)-подібних ABC-транспортерів і мають N-термінальний трансмембранний домен із шістьма α-спіралями. Реконструкція філогенії цих ABC-транспортерів виявила низку клад, представники кожної з яких походять із КБГ ГПА специфічних типів. Філогенетична реконструкція ABC-транспортерів з КБГ ГПА та споріднених пептидних антибіотиків у контексті ABC-транспортерів, кодованих в КБГ інших груп сполук, підтвердила монофілетичне походження перших.

Ключові слова: кластери біосинтетичних генів, глікопептидні антибіотики, рамопланін, фегліміцин, ABC-транспортери

Глікопептидні антибіотики (далі ГПА) – це велика група природних сполук, що синтезуються актиноміцетами, ґрунтовими Грам-позитивними бактеріями із високим вмістом Г-Ц-пар у геномах [27]. ГПА ділять на п’ять типів, залежно від особливостей структури. Усі ГПА є пептидними антибіотиками, які синтезуються за участю нерибосомних пептид синтетаз (НРПС). Виниклі лінійні олігопептиди дозрівають із утворенням поперечних зшивок між фенольними кільцями ароматичних амінокислот, що входять до складу олігопептиду, та декорування атомами хлору [22]. Цим завершується біосинтез ГПА V типу. Натомість ГПА типів I-IV в подальшому можуть глікозилюватись, ацилюватись, а також декоруватися

© Ющук О., Жукровська К., Федоренко В., 2022

сульфатними групами. Отже, *глікопептидами* у строгому сенсі є лише сполуки I-IV типів, а представники V типу є неглікозильованими пептидами. Інакшими, ніж у типу V, є й антимікробні властивості ГПА I-IV типів. Так, ГПА I-IV типів селективно зв'язують кінцевий D-Аланіл-D-Аланін бічного пептидного ланцюга ліпіду II Грам-позитивних бактерій. Це блокує подальші реакції трансглікозилювання і транспептидування, порушуючи біосинтез клітинної стінки [36]. Натомість, ГПА V типу діють на автолізини, групу ферментів, задіяних у рециклінгу пептидоглікану Грам-позитивних бактерій [6, 26]. Якщо оцінити ефективність ГПА як антибактерійних агентів, то можна дійти висновку, що ліпід II є надзвичайно вдалою мішенню їхньої дії. Саме вплив на цю мішень забезпечив успіх ГПА I-IV типів як важливих клінічних сполук. Це, зокрема, такі клінічні ГПА першого покоління як ванкоміцин [14] і тейкопланін [36], а також напівсинтетичні ГПА другого покоління – далбаванцин, оритаванцин і телаванцин [4].

Біосинтез ГПА закодовано у складних мультиоперонних кластерах біосинтетичних генів (КБГ) розміром 60–100 тисяч пар нуклеотидів [34]. До складу КБГ входять біосинтетичні та регуляторні гени, а також гени, що забезпечують стійкість продуцентів до власного антибіотика. Серед останніх є т. зв. *van*-гени, що зумовлюють ремоделювання клітинної стінки, в якій кінцевий D-Аланіл-D-Аланін бічного пептидного ланцюга ліпіду II заміняється на D-Аланіл-D-Лактат [35]. Такий модифікований ліпід II вже не зв'язується ГПА, роблячи клітини продуцента нечутливими до власного продукту. Важливим компонентом КБГ ГПА та подібних антибіотиків є гени трансмембранних експортерів. Їхні продукти найперше задіяні в експорті антибіотиків. Однак це може бути не єдина їхня функція. КБГ інших пептидних антибіотиків також кодують експортери, що часто відіграють роль у формуванні резистентності продуцента до власного продукту, як, наприклад, для випадку антибіотика нізину [17]. Не менш важливою є перспектива застосування генів експортерів як біотехнологічних знарядь із метою створення надпродуцентів антибіотиків. Класичним прикладом тут є біосинтез авермектину в *Streptomyces avermitilis* NRRL 8165, де надекспресія генів ABC-транспортерів *avtA-B* веде до значного підвищення продукції антибіотика. На жаль, обидва цих аспекти ніколи не вивчалися для генів експортерів ГПА та подібних антибіотиків. Схоже, що гени експортерів наявні в усіх секвенованих КБГ ГПА [7]. Тим не менше, до сьогодні експериментально було вивчено лише один із них – *tba* з КБГ баліміцину в *Amycolatopsis balhimycina* DSM 5908 [24]. Tba описано як ABC-транспортер, що об'єднує трансмембранний і АТФ-азний домени в одному білку та, ймовірно, функціонує як гомодимер [24]. З'ясовано, що нокаут гена *tba* веде до зниження позаклітинної концентрації баліміцину, але до підвищення його внутрішньоклітинної концентрації. Системний аналіз розповсюдження, еволюції та структурних особливостей ABC-транспортерів ортологів Tba відсутній. На момент дослідження *tba* у доступі були нуклеотидні послідовності кількох КБГ глікопептидів. На сьогодні кількість доступних послідовностей КБГ незрівнянно більша [1]. Також існуючий консенсус щодо еволюції ГПА трактує деякі інші пептидні антибіотики, такі як рамопланін і фегліміцин, а також низку КБГ із невідомими продуктами, як споріднені до ГПА [1, 10, 31]. Наявність такого масиву даних дає змогу проаналізувати особливості філогенії ABC-транспортерів ГПА, що належать до різних типів, а також споріднених пептидних антибіотиків.

Метою цієї роботи є характеристика розповсюдження ABC-транспортерів, які закодовані в КБГ ГПА I-V типу і споріднених антибіотиків, встановлення філогенетичних зв'язків цих білків у контексті білків-експортерів інших груп природних сполук актиноміцетів, а також оцінка їхньої спорідненості. Досліджені кластери проаналізовано на наявність і розміщення генів ABC-транспортерів. Дані, отримані в роботі, важливі як для

розуміння особливостей походження, генетичного контролю і функціонування ABC-транспортерів ГПА, так і для їхнього подальшого експериментального вивчення.

Методи

Створення вибірок амінокислотних послідовностей білків-експортерів антибіотиків. У Geneious 4.8.5 [16] було створено базу даних, що налічувала 102 КБГ ГПА та споріднених пептидних антибіотиків, а також базу даних, яка налічувала всі КБГ вторинних метаболітів, що продукуються актинобактеріями та депоновані в репозиторії MIBIG [23] станом на травень 2021 року (всього 856 КБГ) (див. електронний додаток 1 – ЕД1). Обидві бази даних було проаналізовано на наявність кодованих гомологів білка-експортера КБГ тейкопланіну – Tei4* (CAG15016.1) – за допомогою алгоритму tblastn [9]. У результаті скринінгу першої бази даних отримано набір із 95 (див. електронний додаток 2 – ЕД2) послідовностей, які сформували вибірку № 1 для філогенетичної реконструкції. Скринінг другої бази даних дав 80 амінокислотних послідовностей білків-експортерів (див. електронний додаток 3 – ЕД3); ці послідовності разом із послідовностями вибірки № 1 склали вибірку № 2 для філогенетичної реконструкції. Файли електронних додатків можна знайти онлайн за посиланням: https://data.mendeley.com/datasets/9jvx4fv2f6/1, doi: 10.17632/9jvx4fv2f6.1.

Філогенетична реконструкція. Пакет програм Mega 10 [19] використовували для створення множинних вирівнювань амінокислотних послідовностей, визначення оптимальних моделей філогенії та для побудови філогенетичних дерев.

Філогенія вибірки № 1 була реконструйована так. Спочатку послідовності вибірки № 1 було вирівняно за допомогою алгоритму Muscle [8]; отримане нередаговане множинне вирівнювання використано для встановлення оптимальної моделі амінокислотних замін. Нею виявилася модель Le-Gascuel (LG) [20] з оцінкою γ-розподілу варіацій частот між сайтами (5 дискретних γ-категорій, G5) та емпіричними частотами амінокислот (F) як така, що показала найменше значення критерію Шварца. Цю модель застосовано для реконструкції філогенетичного дерева білків вибірки № 1 за методом максимальної правдоподібності. Достовірність топології філогенетичного дерева оцінювали за допомогою бутстреп-тесту (500 повторів). Філогенію вибірки № 2 реконструювали за тим самим алгоритмом.

Передбачення КБГ і встановлення організації та субстратної специфічності НРПС. Для передбачення потенційних КБГ використовували платформу antiSMASH 6.0.1 [3]. Визначення доменної організації НРПС та передбачення специфічності відповідних аденілазних доменів здійснювали за допомогою PKS/NRPS Analysis [2].

Передбачення доменної організації білків-експортерів і моделювання їхнього розташування у клітинній мембрані. Для ідентифікації структурних і функціональних доменів та мотивів в амінокислотних послідовностях білків використовували CD-Search [21]. Для визначення трансмембранних α-спіралей в амінокислотних послідовностях ABC-транспортерів використовували TMHMM 2.0 [18], реконструкцію 2D-топології ABC-транспортера стосовно цитоплазматичної мембрани здійснювали за допомогою TMRPres2D [30].

*Моделювання третинної та четвертинної структури ABC-транспортера Tei4**. Третинну структуру Tei4* моделювали на основі експериментально встановленої кристалічної структури ABC-транспортера 3QF4 з *Thermotoga maritima* [11] із використанням сервера Swiss-Model [32]. Отриману модель мономера Tei4* використовували для моделювання четвертинної структури – гомодимера Tei4* – за допомогою симетричного докінгу в SymmDock [29]. Із ряду отриманих моделей димерів вибрали ту, що мала найбільшу площу білок-білкового інтерфейсу.

Результати і їхнє обговорення

Поширення й особливості синтенії генів білків-експортерів. У вибірці із 102 КБГ ГПА та споріднених пептидів було знайдено 95 кодованих ABC-транспортерів. Невідповідність цих цифр пояснюється низкою причин. По-перше, в трьох КБГ, а саме: CP015588 зі *Streptomyces alfalfae* ACCC40021, LGCP01000275 зі *Streptomyces* sp. NRRL WC-3701 та LMWF01000294 зі *Streptomyces rimosus* ssp. *rimosus* NRRL WC-3897 гени ABC-транспортерів анотовано як псевдогени, що не кодують продукту. По-друге, КБГ CP011492 зі *Streptomyces* sp. CNQ-509, CP025407 зі *Streptomyces* sp. CMB-StM0423 та КБГ метагеномного зразка esnapd15 (KF264554.1) несли по два гени ABC-транспортера. Нарешті, КБГ FNHV01000033 зі *Streptomyces* sp. Cf386, LAKD02000014 зі *Streptomyces antioxidans* MUSC 164 і LQMT02000011 з *Amycolatopsis keratiniphila* ssp. *keratiniphila* DSM 44409 перебували на коротких контігах, що не покривали всю дійсну послідовність КБГ. У цьому разі були втрачені частини кластерів, де очікувано мають міститися гени ABC-транспортерів. Загалом, ці винятки не заперечують твердження, що кожен проаналізований КБГ ГПА та споріднених пептидних антибіотиків несе у своєму складі один ген ABC-транспортера, а іноді і два. Цікаво, що анотовані псевдогени в трьох випадках можуть бути помилками секвенування та потребують подальших досліджень. Не виключено також, що ці гени насправді містять мутації, а відповідні КБГ можуть бути неактивними.

Виявлено певні закономірності розміщення генів ABC-транспортерів у КБГ. Їхні позиції доволі консервативні в різних КБГ, що походять із різних родів актинобактерій і кодують різні типи сполук. Зокрема, в 37 КБГ, що належать головно представникам порядку *Pseudonocardiales*, включно з одним КБГ з роду *Nocardia* та 4 КБГ з неідентифікованих ізолятів (рис. 1, а), ген ABC-транспортера завжди містився на початку КБГ, між геном префенат дегідрогенази і геном перших модулів НРПС (рис. 1, а). Ці КБГ кодують як відомі, так і невідомі ГПА I-IV типів. КБГ метагеномного зразка esnapd15 (KF264554.1) несе два гени ABC-транспортерів, однак лише один міститься між геном префенат дегідрогенази і геном перших модулів НРПС. Подібна тенденція зберігалася для семи КБГ з порядку *Streptosporangiaceae*, в яких ген ABC-транспортера завжди розміщений перед генами НРПС (рис. 1, б). На противагу цьому, у КБГ ГПА представників родини *Micromonosporales* – тейкопланіну (AJ632270.1) і UK-68,597 (KF192710.1), а також у КБГ неідентифікованих ізолятів CA37 (HM486074.1) і CA915 (HM486076.1) не виявлено синтенії розміщення гени ABC-транспортерів. Хоча в КБГ рамопланіну (DD002243.1) та черсинаміцину (WP_091321314.1), які також продукуються представниками порядку *Micromonosporales*, генне оточення генів ABC-транспортерів збігалося (рис. 1, в).

Найбільша кількість проаналізованих КБГ ГПА та споріднених пептидних антибіотиків походить із представників порядку *Streptomycetales*, включаючи КБГ неідентифікованого ізолята GQ475282 (таксономічна приналежність якого не визначена). Усього проаналізовано 42 КБГ зі *Streptomycetales*. Із них 3 кодують біосинтез ГПА I та III типів, де ген ABC-транспортера виявився розміщений перед генами НРПС (рис. 1, г).

Ще 39 КБГ *Streptomycetales* кодують біосинтез ГПА V типу, таких як комплестатин, корбоміцин і GP6738, фегліміцин, а також ендурацидин, і невідомих пептидних антибіотиків. Серед них для найчисленнішої підмножини – КБГ ГПА V типу комплестатину та подібних йому (18 КБГ) – характерним було розміщення генів ABC-транспортерів між геном StrR-подібного шлях-специфічного регулятора та геном перших модулів НРПС (рис. 1, д). Єдиним винятком у цій підмножині був КБГ BJHX01000001 зі *Streptomyces avermitilis* NBRC 14893, в якому виявився додатковий ген, що кодує білок із невідомою

функцією, розміщений геном StrR-подібного регулятора та геном ABC-транспортера. Наступна підмножина (8 КБГ) включала в себе КБГ ГПА V типу корбоміцину та низку подібних до нього КБГ, продукти яких невідомі, і характеризувалася розміщенням гена ABC-транспортера між геном MbtH-шаперона, задіяного в функціонування НРПС, і генами двокомпонентної регуляторної системи із генами НРПС (рис. 1, е). Гени двокомпонентної регуляторної системи кодували сенсорну гістидинову кіназу та регулятор відповіді, причому ген регулятора відповіді міг бути відсутнім (рис. 1, е).

Третя підмножина складалась із КБГ ГПА V типу GP6738, КБГ фегліміцину (KT809366.1), GQ475282 і 9 КБГ, подібних до них. У цих КБГ ген ABC-транспортера розміщений між геном MbtH і геном перших модулів НРПС (рис. 1, є). Варто відзначити, що в КБГ CP011492 зі *Streptomyces* sp. CNQ-509 та CP025407 зі *Streptomyces* sp. CMB-StM0423 знайдено по два гени, що кодують ABC-транспортери. З кожної пари тільки один вписувався в загальну закономірність синтенії, тоді як інший містився на 5'-краю потенційного КБГ. Ймовірно, останні взагалі не входять до складу КБГ. Розміщення генів ABC-транспортера ще в трьох КБГ суттєво відрізнялося від усіх згаданих вище. У цьому разі ці гени віддалені від генів НРПС, але оточені генами MbtH та StrR-подібного регулятора (рис. 1, ж).

Нарешті, КБГ CP015098 із *Streptomyces* sp. S10(2016) характеризувався розміщенням гена ABC-транспортера, не подібним до описаних вище. Цікаво, що цей КБГ також суттєво відрізнявся від усіх інших зі *Streptomycetales*, зокрема, його НРПС потенційно може кодувати пептид, що складається принаймні з 20 амінокислотних залишків (а. з.) У ньому гени двокомпонентної регуляторної системи були перед геном ABC-транспортера, а ген MbtH – після нього (рис. 1, з).

Аналізуючи вищеописані дані, можна зробити висновок, що гени ABC-транспортерів найчастіше ко-локалізовані з генами НРПС. Ймовірно, таке розташування гарантує синхронізацію експресії цих генів, забезпечуючи правильне узгодження біосинтезу й експорту антибіотиків у часі.

Порівняння амінокислотних послідовностей ABC-транспортерів, що кодовані в КБГ ГПА, та споріднених пептидних антибіотиків. Спочатку проаналізовано амінокислотні послідовності 95 ABC-транспортерів із вибірки № 1, віднайдених у КБГ ГПА та споріднених антибіотиків. Виявилося, що всі вони належать до MdlB(MsbA)-подібних ABC транспортерів еукаріотичного типу підродини ABC-B (COG1132) [15, 28]. Подібно до вже описаного Tba [24], ці білки поєднують N-термінальний трансмембранний і С-термінальний АТФ-зв'язувальний домени, що наявні у всіх білків вибірки в інтактному вигляді.

Далі здійснено реконструкцію філогенії для амінокислотних послідовностей вибірки № 1 (див. Методи). З отриманого філогенетичного дерева видно, що дві послідовності, а саме AKH86044.1 зі *Streptomyces* sp. CNQ-509 і AUH39259.1 *Streptomyces* sp. CMB-StM0423, виступали як зовнішні групи щодо всіх інших (рис. 2). Як згадувалося вище, ці ABC-транспортери були кодовані додатковими генами, що лежали на 5'-кінці відповідних КБГ. Ми припустили, що ці гени взагалі не входять до складу КБГ, а отримана реконструкція філогенії свідчить на користь такого припущення. Далі ми виділили на дереві 5 основних клад, які отримали найкращу бутстреп-підтримку. Зокрема, клада I сформована послідовностями ABC-транспортерів, що походили із КБГ ГПА I-IV типів (39 послідовностей), а також з КБГ ГПА неідентифікованих ізолятів TEG, VEG, CA37, CA915, CA878 та enspd15. Усі ці послідовності надзвичайно консервативні, що вказує на відсут-

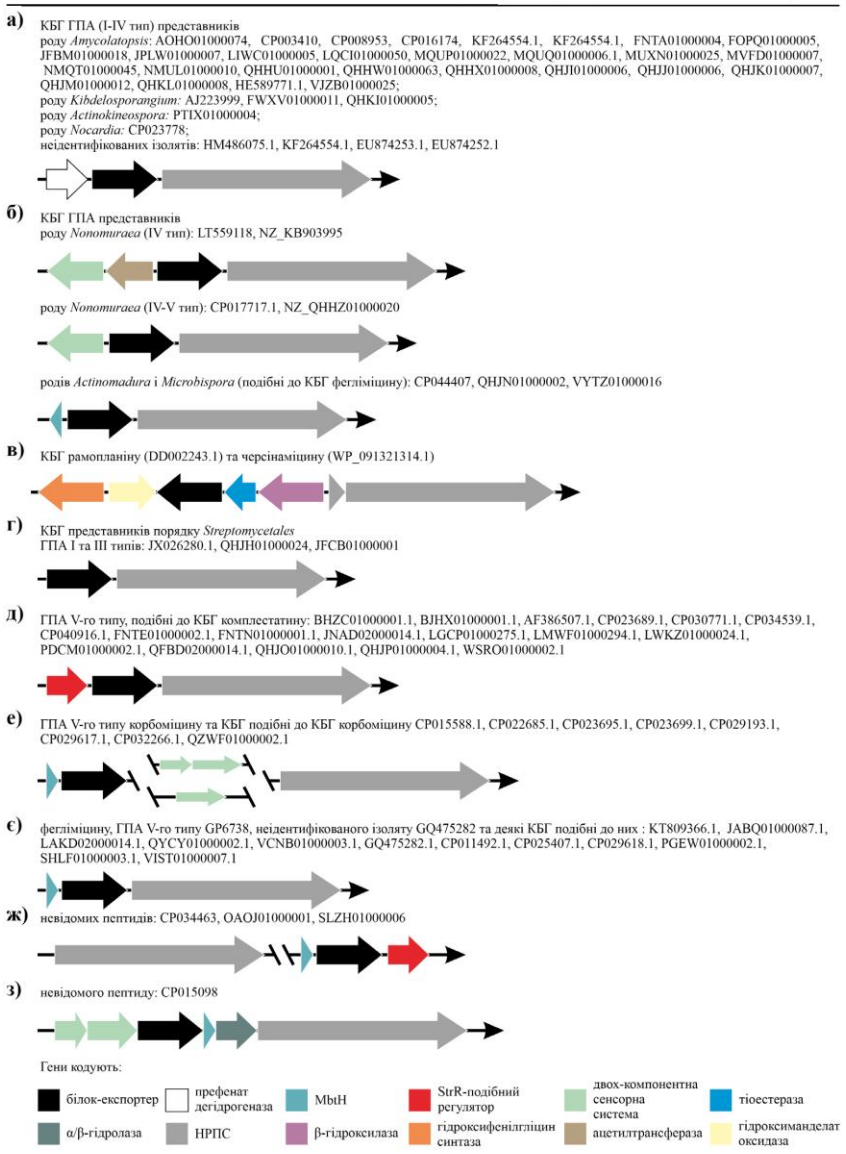


Рис. 1. Особливості розміщення генів ABC-транспортерів у різних кластерів генів біосинтезу глікопептидів і споріднених пептидних антибіотиків

ність кореляції між структурою і особливостями декорування ГПА та структурою ABC-транспортерів. Моделювання трансмембранних ділянок показало, що всі білки клади I мають шість трансмембранних α-спіралей, а C-термінальні АТФ-азні домени мають повний набір мотивів [33], необхідних для їхнього функціонування (рис. 3, а). Це, зокрема, мотив Волкера A (P-петля), Q-петля, мотив Волкера B, D-петля, H-петля та сигнатурний мотив ABC-транспортера. Амінокислотні послідовності всіх шести трансмембранних α-спіралей були майже ідентичними у всіх 45 послідовностей клади I. Те саме можна було сказати і про набір мотивів АТФ-азного C-термінального домена. Отже, ABC-транспортери клади I, найімовірніше, є функціональними та володіють набором структур, необхідних для цього.

Клади II, III та IV сформовані послідовностями білків, кодованих у КБГ комплестатину, та подібних до нього. Важко пояснити, чому ці ABC-транспортери формують три добре відділені клади, хоч і походять із однотипних КБГ. Більш детальний аналіз цих КБГ показав, що потенційні продукти найімовірніше мають однакові особливості галогенування та формування поперечних зшивок. Єдина відмінність полягала в тому, що білки клади III походять із КБГ, що кодують комплестатин-подібні ГПА V типу, в яких наш аналіз передбачає наявність D-тирозину в 6 положенні олігопептидного кору, тоді як білки клад II та IV походять з КБГ, які кодують олігопептиди із D-метил-тирозину в 6 положенні (що, власне, і характерне для комплестатину) [5]. Подібно до представників клади I, всі 16 послідовностей білків клад II, III та IV були надзвичайно подібними. Їх об'єднували 6 консервативних трансмембранних α-спіралей в N-термінальному домені (рис. 3, б). АТФ-азні C-термінальні домени всіх 16 білків володіли набором характерних мотивів (рис. 3, б).

Варто відзначити, що ABC-транспортер із КБГ кистаміцину (з *Nonomuraea* sp. ATCC 55076) – також ГПА V типу – не належав до жодної із цих клад. Натомість, він займав своєрідне «перехідне» положення, будучи зовнішньою групою клади I. При цьому цей ABC-транспортер мав усі структурні особливості, характерні для білків клад II-IV.

Передбачення доменної організації всіх без винятку НРПС, що кодовані в КБГ, відповідних до клад I-IV, вказувало на гептапептидну структуру всіх продуктів. Це не дивно, враховуючи, що ГПА I-IV типів – клада I, а також ГПА V типу, подібні до комплестатину, – клади II-IV – всі без винятку мають гептапептидну структуру.

Клада V була радикально відмінною в цьому аспекті. Виявилося, що ABC-транспортери клади V (рис. 2) кодовані в ряді різноманітних КБГ. По-перше, це були КБГ ліподепсипептидів, таких як рамопланін, черсинаміцин і ендурацидин. По-друге, ряд КБГ із продуктами, подібними до нонапептидів, що належать до ГПА V типу, таких як корбоміцин та GP6738. До клади V належали і ABC-транспортери, закодовані в КБГ фегліміцину та GQ475282. До цієї клади також належали і ABC-транспортери, що кодовані в екзотичних КБГ, лиш віддалено подібних до КБГ фегліміцину чи нонапептидних ГПА V типу. Що ж об'єднувало таку різноманітність КБГ, в яких кодовано ABC-транспортери клади V?

Передбачення доменної організації відповідних НРПС показало, що всі ці КБГ кодують біосинтез пептидних антибіотиків, які складаються із восьми та більше а. з. Найчастіше це було 8 або 9 а. з. в імовірних сполуках, КБГ яких були найбільше схожими до КБГ корбоміцину та GP6738. А от КБГ із *Streptomyces* sp. S10(2016) кодував НРПС, здатну до біосинтезу пептиду із щонайменше 20 а. з. Аналіз імовірної структури білків клади V показав, що всі вони мають 6 трансмембранних α-спіралей і функціональні АТФ-азні домени (рис. 3, в). Цікаво, що другий ABC-транспортер, який кодовано в КБГ esnapd15 – ADK54818.1 – також потрапив до клади V. Однак у послідовності цього білка була відсутня ділянка, що відповідає першій трансмембранній α-спіралі. Це може свідчити про нефункціональність цього транспортера.

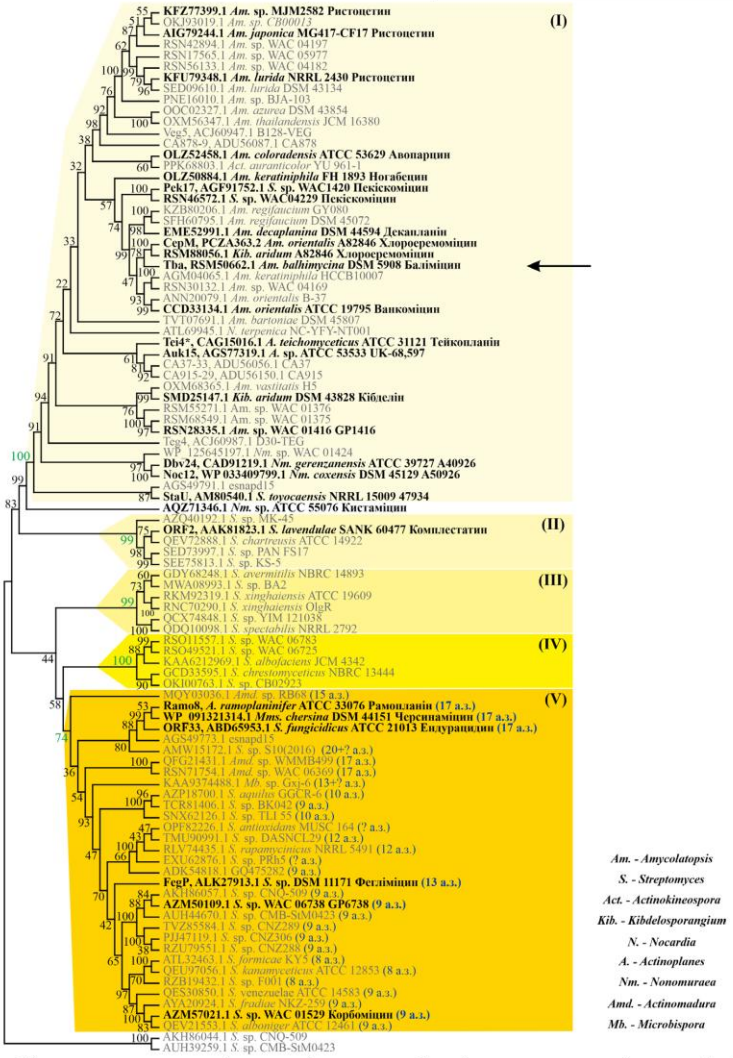


Рис. 2. Філогенетичне дерево (у вигляді кладограми) амінокислотних послідовностей ABC-транспортерів, що кодовані в КБГ ГПА та споріднених пептидів – вибірки № 1. На дереві було виділено п'ять основних клад із високою бутстреп-підтримкою; єдиний ABC-транспортер, вивчений експериментально – Tba з КБГ баліміцину – показаний стрілкою; послідовності, що походять із КБГ відомих сполук, виділені жирним шрифтом. Особливості білків, що належать до кожної з клад, обговорюються в тексті. Алгоритм побудови цього дерева описано в розділі «Методи»

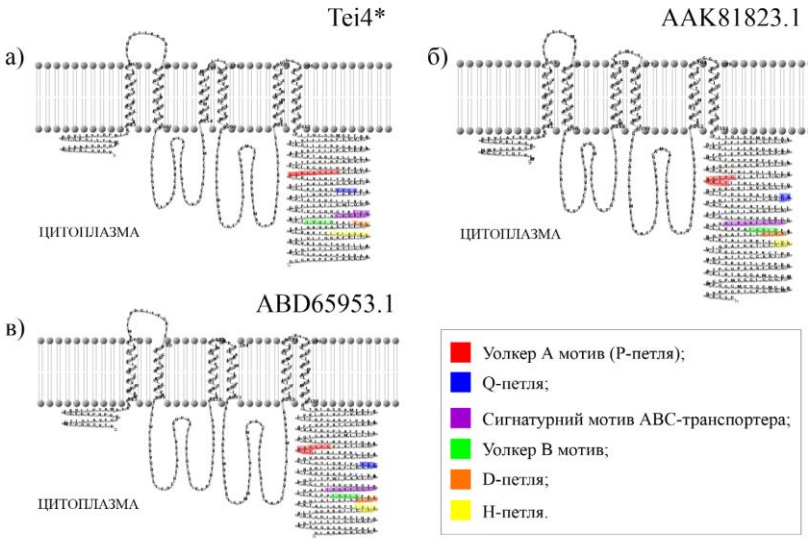


Рис. 3. Вторинна структура АВС-транспортерів, що належать до клад I, II-IV та V на прикладах АВС-транспортера, кодованого в КБГ тейкопланіну – Tei4* (а), комплестатину – AAK81823.1 (б) та ендурацидину – ABD65953.1 (в), відповідно. Передбачення трансмембранних α-спіралей і розміщення білків у клітинній мембрані було здійснено, як описано в розділі «Методи». Різними кольорами виділено мотиви С-термінальних АТФ-азних доменів, необхідні для функціонування білків

Отже, з отриманих результатів можна зробити висновок, що АВС-транспортери, закодовані в КБГ ГПА та споріднених пептидних антибіотиків, є доволі консервативними: всі вони мають шість трансмембранних α-спіралей та функціональні С-термінальні АТФ-азні домени. Як вже згадувалося, існує припущення [24], що такі АВС-транспортери функціонують у вигляді гомодимерів. Щоб оцінити це припущення, ми спершу передбачили третинну структуру АВС-транспортера, що закодований в КБГ тейкопланіну – Tei4*, а потім спробували змоделювати четвертинну структуру – гомодимер Tei4*. Отримані моделі мономеру Tei4* показували правдоподібне розміщення усіх структурних елементів – трансмембранних α-спіралей і компонентів активного сайту АТФ-азного С-термінального домену (рис. 4, а), тоді як симетричне докінг-моделювання показало, що дві молекули Tei4* цілком імовірно можуть формувати димер (рис. 4, б).

Філогенія АВС-транспортерів, закодованих в КБГ ГПА та споріднених пептидних антибіотиків у контексті АВС-транспортерів інших природних сполук. Нарешті цікаво було дослідити, з якими АВС-транспортерами, відповідальними за викачування інших груп природних сполук, найбільш споріднені АВС-транспортери ГПА. Для цього було реконструйовано філогенію вибірки № 2 (див. розділ «Методи»). По-перше, виявилося (рис. 5), що всі АВС-транспортери ГПА та споріднених пептидних антибіотиків належать до однієї клади дерева на основі вибірки № 2. Це свідчить про монофілію усіх цих АВС-транспортерів – незважаючи на суттєві відмінності у структурі ГПА, фегліміцину

чи ліподепсипептидів. Схоже, що всі їхні КБГ еволюціонували від одного предкового протокластера, який вже містив ген АВС-транспортера. По-друге, клада АВС-транспортерів, кодованих в КБГ ГПА та споріднених пептидних антибіотиків, виявилася сестринською до клади, що об'єднувала 4 АВС-транспортери з КБГ даптоміцину [25], ненестатину [13] та ломаівітицинів [12] (рис. 5). Хоча даптоміцин також є ліпопептидним антибіотиком, що синтезується нерибосомно, ненестатин і ломаівітицини є полікетидами. Спорідненість усіх цих АВС-транспортерів вкотре вказує на те, для них не характерна особлива специфічність стосовно субстратів експортування.

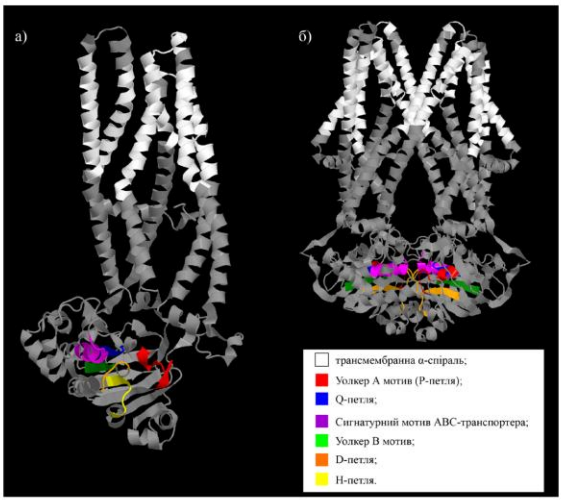


Рис. 4. Моделі третинної (а) та четвертинної (б) структури АВС-транспортера, закодованого в КБГ тейкопланіну – Tei4*. Моделі побудовано, як описано в розділі «Методи». Структурні елементи виділено кольорами

У цій роботі ми проаналізували *in silico* особливості розповсюдження, структури та філогенії АВС-транспортерів, що походять з КБГ ГПА та споріднених пептидних антибіотиків. З отриманих результатів можна зробити кілька важливих висновків. Виявилося, що АВС-транспортери з КБГ ГПА близько споріднені з АВС-транспортерами таких антибіотиків як фегліміцин і ліподепсипептиди – рамопланін, ендурацидин і черсинаміцин. Усі проаналізовані КБГ містять гени АВС-транспортерів. Розміщення цих генів доволі консервативне – найчастіше вони містяться з 5'-флангу від генів НРПС. Усі кодовані ними білки належать до MdlB(MsbA)-подібних АВС-транспортерів і мають шість консервативних трансмембранних α-спіралей в N-термінальному домені. АТФ-азні С-термінальні домени цих білків не менш консервативні та несуть усі необхідні для функціонування набори структурних мотивів. Моделювання третинної та четвертинної структури одного з цих АВС-транспортерів – Tei4* – вказує на те, що вони цілком можуть функціонувати як гомодимери. Філогенія АВС-транспортерів, що походять з КБГ ГПА та споріднених пептидних антибіотиків, дає змогу розподілити їх на п'ять клад. До клади I належать АВС-транспортери ГПА I–IV типів, тоді як клади II–IV об'єднують АВС-транспортери гептапептидних ГПА V типу, подібних до комплестатину. Останню кладу

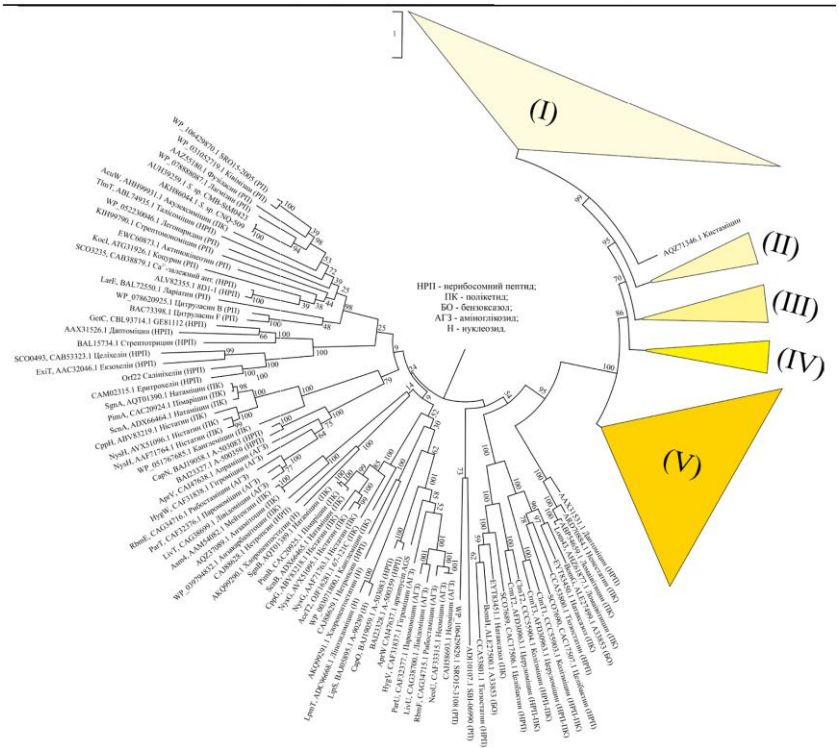


Рис. 5. Філогенетичне дерево, що репрезентує філогенетичні зв'язки між білками АВС-транспортерами вибірок № 1 та № 2. Помітно, що всі білки АВС-транспортерів ГПА та споріднених антибіотиків мають монофілетичне походження, формуючи одну кладу. Білки АВС-транспортерів, що закодовані в КБГ даптоміцину, ненестатину та ломаівітицину виступають зовнішньою групою до всіх білків вибірки № 1. Планка масштабу вказує на кількість амінокислотних замін на один сайт. Алгоритм побудови цього дерева описано в розділі «Методи»

V формують АВС-транспортери ліподепсипептидів, фегліміцину, нонапептидних ГПА V типу та множини невідомих антибіотиків, які складаються з восьми та більше а. з. Незважаючи на різноманітність сполук, що їх імовірно експортують ці АВС-транспортери, їхні послідовності дуже консервативні. Це свідчить про те, що ці АВС-транспортери універсальні та здатні експортувати різні пептидні антибіотики незалежно від їхньої структури. Усі АВС-транспортери з КБГ ГПА і споріднених пептидних антибіотиків мають монофілетичне походження. Результати, отримані в роботі, дають змогу краще зрозуміти особливості АВС-транспортерів такої важливої групи сполук, як ГПА, і є основою для їхнього подальшого експериментального вивчення.

Фінансування

Ця робота була виконана в рамках держбюджетної теми БГ-09Ф.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Andreo-Vidal A., Binda E., Fedorenko V.* et al. Genomic insights into the distribution and phylogeny of glycopeptide resistance determinants within the actinobacteria phylum // Antibiotics. 2021. Vol. 10. P. 1533. doi.org/10.3390/antibiotics10121533
2. *Bachmann B. O., Ravel J.* Methods for *in silico* prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from dna sequence data // Meth. Enzymol. 2009. Vol. 458. P. 181–217. doi.org/10.1016/S0076-6879(09)04808-3
3. *Blin K., Shaw S., Kloosterman A. M.* et al. AntiSMASH 6.0: improving cluster detection and comparison capabilities // Nucleic Acids Res. 2021. Vol. 49. P. W29–W35. doi.org/10.1093/nar/gkab335
4. *Butler M. S., Hansford K. A., Blaskovich M. A.* et al. Glycopeptide antibiotics: back to the future // J. Antibiot. 2014. Vol. 67. P. 631–644. doi.org/10.1038/ja.2014.111
5. *Chiu H. T., Hubbard B. K., Shah A. N.* et al. Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster // Proc. Natl. Acad. Sci. U.S.A. 2001. Vol. 98. P. 8548–8553. doi.org/10.1073/pnas.151246498
6. *Culp E. J., Waglechner N., Wang W.* et al. Evolution-guided discovery of antibiotics that inhibit peptidoglycan remodelling // Nature. 2020. Vol. 578. P. 582–587. doi.org/10.1038/s41586-020-1990-9
7. *Donadio S., Sosio M., Stegmann E.* et al. Comparative analysis and insights into the evolution of gene clusters for glycopeptide antibiotic biosynthesis // Mol. Genet. Genom. 2005. Vol. 274. P. 40–50. doi.org/10.1007/s00438-005-1156-3
8. *Edgar R. C.* MUSCLE: multiple sequence alignment with high accuracy and high throughput // Nucleic Acids Res. 2004. Vol. 32. P. 1792–1797. doi.org/10.1093/nar/gkh340
9. *Gerts E. M., Yu Y. K., Agarwala R.* et al. Composition-based statistics and translated nucleotide searches: improving the TBLASTN module of BLAST // BMC Biol. 2006. Vol. 4. P. 1–14. doi.org/10.1186/1741-7007-4-41
10. *Gonsior M., Mühlenweg A., Tietzmann M.* et al. Biosynthesis of the peptide antibiotic feglymycin by a linear nonribosomal peptide synthetase mechanism // ChemBioChem. 2015. Vol. 16. P. 2610–2614. doi.org/10.1002/cbic.201500432
11. *Hohl M., Briand C., Grütter M. G., Seeger M. A.* Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation // Nat. Struct. Mol. Biol. 2012. Vol. 19. P. 395–402. doi.org/10.1038/nsmb.2267
12. *Janso J. E., Haltli B. A., Eustáquio A. S.* et al. Discovery of the lomaiviticin biosynthetic gene cluster in *Salinispora pacifica* // Tetrahedron. 2014. Vol. 70. P. 4156. doi.org/10.1016/J.TET.2014.03.009
13. *Jiang X., Fang Z., Zhang Q.* et al. Discovery of a new asymmetric dimer nenestatin B and implications of a dimerizing enzyme in a deep sea actinomycete // Org. Biomol. Chem. 2021. Vol. 19. P. 4243–4247. doi.org/10.1039/D1OB00310K
14. *Jovetic S., Zhu Y., Marcone G. L.* et al. β-Lactam and glycopeptide antibiotics: first and last line of defense? // Trends in Biotechnol. 2010. Vol. 28. P. 596–604. doi.org/10.1016/j.tibtech.2010.09.004
15. *Karow M., Georgopoulos C.* The essential *Escherichia coli msbA* gene, a multicopy suppressor of null mutations in the *htrB* gene, is related to the universally conserved family of ATP-dependent translocators // Mol. Microbiol. 1993. Vol. 7. P. 69–79. doi.org/10.1111/j.1365-2958.1993.tb01098.x
16. *Kearse M., Moir R., Wilson A.* et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data // Bioinformatics. 2012. Vol. 28. P. 1647–1649. doi.org/10.1093/bioinformatics/bts199

17. *Khosa S., Lagedroste M., Smits S. H. J.* Protein defense systems against the lantibiotic nisin: function of the immunity protein NisI and the resistance protein NSR // Front. Microbiol. 2016. Vol. 7. P. 504. doi.org/10.3389/fmicb.2016.00504
18. *Krogh A., Larsson B., Von Heijne G., Sonnhammer E. L. L.* Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes // J. Mol. Biol. 2001. Vol. 305. P. 567–580. doi.org/10.1006/jmbi.2000.4315
19. *Kumar S., Stecher G., Li, M., Knyaz C., Tamura K.* MEGA X: molecular evolutionary genetics analysis across computing platforms // Mol. Biol. Evol. 2018. Vol. 35. P. 1547–1549. doi.org/10.1093/molbev/msy096
20. *Le S. Q., Gascuel O.* An improved general amino acid replacement matrix // Mol. Biol. Evol. 2008. Vol. 25. P. 1307–1320. doi.org/10.1093/molbev/msn067
21. *Marchler-Bauer A., Bryant S. H.* CD-Search: protein domain annotations on the fly // Nucleic Acids Res. 2004. Vol. 32. P. 327–331. doi.org/10.1093/nar/gkh454
22. *Marcone G. L., Binda E., Berini F., Marinelli F.* Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era // Biotechnol. Adv. 2018. Vol. 36. P. 534–554. doi.org/10.1016/j.biotechadv.2018.02.009
23. *Medema M. H., Kottmann R., Yilmaz P.* et al. Minimum information about a biosynthetic gene cluster // Nat. Chem. Biol. 2015. Vol. 11. P. 625–631. doi.org/10.1038/nchembio.1890
24. *Menges R., Muth G., Wohlleben W., Stegmann E.* The ABC transporter Tba of *Amycolatopsis balhimycina* is required for efficient export of the glycopeptide antibiotic balhimycin // Appl. Microbiol. Biotechnol. 2007. Vol. 77. P. 125–134. doi.org/10.1007/s00253-007-1139-x
25. *Miao V., Coëffet-LeGal M. F., Brian P.* et al. Daptomycin biosynthesis in *Streptomyces roseosporus*: cloning and analysis of the gene cluster and revision of peptide stereochemistry // Microbiology. 2005. Vol. 151. P. 1507–1523. doi.org/10.1099/MIC.0.27757-0/CITE/REFWORKS
26. *Mitchell S. J., Verma D., Griswold K. E., Bailey-Kellogg C.* Building blocks and blueprints for bacterial autolysins // PLoS Computat. Biol. 2021. Vol. 17. P. e1008889. doi.org/10.1371/JOURNAL.PCBI.1008889
27. *Nicolaou K. C., Boddy C. N. C., Bräse S., Winssinger N.* Chemistry, biology, and medicine of the glycopeptide antibiotics // Angew. Chem. 1999. Vol. 38. P. 2096–2152. doi.org/10.1002/(sici)1521-3773(19990802)38:15<2096::aid-anie2096>3.0.co;2-f
28. *Nishino K., Yamaguchi A.* Analysis of a complete library of putative drug transporter genes in *Escherichia coli* // J. Bacteriol. 2001. Vol. 183. P. 5803–5812. doi.org/10.1128/JB.183.20.5803-5812.2001
29. *Schneidman-Duhovny D., Inbar Y., Nussinov R., Wolfson H. J.* PatchDock and SymmDock: servers for rigid and symmetric docking // Nucleic Acids Res. 2005. Vol. P. W363. doi.org/10.1093/nar/gki481
30. *Spyropoulos I. C., Liakopoulos T. D., Bagos P. G., Hamodrakas S. J.* TMRPres2D: high quality visual representation of transmembrane protein models // Bioinformatics. 2004. Vol. 20. P. 3258–3260. doi.org/10.1093/bioinformatics/bth358
31. *Waglechner N., McArthur A. G., Wright G. D.* Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance // Nat. Microbiol. 2019. Vol. 4. P. 1862–1871. doi.org/10.1038/s41564-019-0531-5
32. *Waterhouse A., Bertoni M., Bienert S.* et al. SWISS-MODEL: homology modelling of protein structures and complexes // Nucleic Acids Res. 2018. Vol. 46. P. W296–W303. doi.org/10.1093/nar/gky427

33. *Westfahl K. M., Merten J. A., Buchaklian A. H., Klug C. S.* Functionally important ATP binding and hydrolysis sites in *Escherichia coli* MsbA // Biochem. 2008. Vol. 47. P. 13878–13886. doi.org/10.1021/bi801745u
34. *Yim G., Thaker M. N., Koteva K., Wright G.* Glycopeptide antibiotic biosynthesis // J. Antibiot. 2014. Vol. 67. P. 31–41. doi.org/10.1038/ja.2013.117
35. *Yushchuk O., Binda E., Marinelli F.* Glycopeptide antibiotic resistance genes: distribution and function in the producer actinomycetes // Front. Microbiol. 2020. Vol. 11. P. 1173. doi.org/10.3389/fmicb.2020.01173
36. *Yushchuk O., Ostash B., Truman A. W.* et al. Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes // Appl. Microbiol. Biotechnol. 2020. Vol. 104. P. 3279–3291. doi.org/10.1007/s00253-020-10436-y

Стаття надійшла до редакції 09.05.22

доопрацьована 17.05.22

прийнята до друку 18.05.22

INSIGHTS INTO THE PHYLOGENY OF TRANSPORTERS CODED WITHIN BIOSYNTHETIC GENE CLUSTERS FOR GLYCOPEPTIDES AND RELATED ANTIBIOTICS

O. Yushchuk, K. Zhukrovska, V. Fedorenko

Ivan Franko National University of Lviv
4, Hrushevskyi St., Lviv 79005, Ukraine
e-mail: viktor.fedorenko@lnu.edu.ua

Glycopeptide antibiotics (GPAs) represent one of the most important classes of natural antibiotics coming from actinomycetes – high GC soil-dwelling Gram-positive bacteria. Among GPAs are important clinical compounds, such as vancomycin and teicoplanin, being "last defense line" against multidrug resistant Gram-positive pathogens. Recent works demonstrated, that peptide antibiotics like ramoplanin and feglymycin, although having rather distinct structure, are genetically related to GPAs. Biosynthesis of all these compounds is coded within large gene assemblages – biosynthetic gene cluster (BGCs). BGCs of GPAs, ramoplanin, feglymycin and other related peptide antibiotics share multiple common features. One of them is the presence of genes coding for ABC-transporters. Most obvious role of these ABC-transporters is export of antibiotics. However, certain role of ABC-transporters in the auto-resistance cannot be excluded as well. Multiple genomes of actinomycetes were sequenced and are fully available today, allowing to build a significant collection of BGCs for GPAs and related peptide antibiotics. Therefore, in this work we aimed to investigate *in silico* distribution, structural features and phylogeny of ABC-transporters, encoded within 102 BGC of GPAs and related peptide antibiotics. We found out, that ABC-transporters from GPA BGCs are very similar to ABC-transporters from ramoplanin and feglymycin BGCs, as well as to ABC-transporters coded within BGCs of putative compounds. All these proteins belonged to MdlB(MsbA)-like ABC-transporters, possessing N-terminal transmembrane domain with 6 α-helices. Phylogenetic reconstruction revealed that these ABC-transporters fall into several clades, which might be correlated with specific types of peptide antibiotics. Finally, a wider phylogenetic reconstruction allowed to conclude the monophyly of ABC-transporters, encoded within BGCs of GPAs and other related peptide antibiotics.

Keywords: biosynthetic gene clusters, glucopeptide antibiotics, ramoplanin, feglymycin, ABC-transporters

Оригінал доступний за посиланням:
http://publications.lnu.edu.ua/bulletins/index.php/biology/article/view/11479.
Електронні додатки до статті доступні за посиланням:
https://data.mendeley.com/datasets/9jvx4fv2f6/1.

8.4 Розповсюдження *van*-генів за межами типу *Actinobacteria*

Результати нашого філогеномного аналізу розповсюдження *van*-генів серед актинобактерій (викладені у **підрозділі 8.2**) доволі безальтернативно вказують на цей таксон як на вихідне джерело детермінантів стійкості до ГПА.

Вивчення поширення *van*-генів серед умовно-патогенних мікроорганізмів, таких як ентерококи або стафілококи, є актуальною темою через їх клінічну значимість [22, 178–197]. З іншого боку, передача *vanHAXRS* генів до нейтральних ґрунтових бактерій з низьким ГЦ-вмістом на сьогодні досліджувалася недостатньо, хоча саме вони, найімовірніше, служать мостом для поширення *van*-генів від актинобактерій до небезпечних патогенів. Зокрема, в літературі існують повідомлення про поширення *van*-генів серед двох класів ґрунтових Грам-позитивних бактерій, а саме класів *Bacilli* і *Clostridia* [154, 157, 198–200], на додачу до актиноміцетів. Загалом, гени *vanHAXRS* знаходили в штамах стафілококів і ентерококів, деяких інших бактерій із класу *Bacilli,* а також зрідка в клостридій (напр. *Clostridioides difficile* AI0499). У багатьох випадках *van*-гени колокалізовувалися з генами МГЕ: транспозонів і плазмід, які, ймовірно, сприяють розповсюдженню. У цьому підрозділі ми вирішили детально дослідити як і наскільки *van*-гени розповсюджені за межами типу *Actinobacteria*. Ми перевірили загальнодоступні бази даних геномних послідовностей в пошуках генів *vanHAXRS*.

Наш скринінг виявив присутність *vanHAXRS* та споріднених генів у чотирьох класах Грам-позитивних бактерій за межами типу *Actinobacteria*, а саме в класах *Bacilli*, *Clostridia*, *Erysipelotrichia* та *Ktedonobacteria*, а також в одному класі Грам-негативних бактерій – *Anaerolineae*, усі належать до надтипу *Terrabacteria* [201]. Наскільки нам відомо, *van*-гени раніше не описували в представників останніх трьох класів. Щоб інтегрувати та інтерпретувати цю інформацію, ми створили надійне філогенетичне дерево, що представляло б еволюційні взаємозв’язки Van білків, кодованих різними класами бактерій і перевірили, чи корелює генетичний контекст *van*-генів (тобто колокалізація генів, пов’язаних з МГЕ) з їх філогенією.

Аналіз філогенії виявив нещодавню подію ГГТ, яка привела до передачі *van*-генів від актинобактерій до представників класу *Ktedonobacteria* та деяких бактерій класу *Clostridia*. Відповідні білки не сильно відрізнялися від актинобактеріальних і були знайдені глибоко в коровій кладі *Actinobacteria*. Більшість із цих ктедонобактерій та клостридій були виділені з ґрунтових біотопів [202], що уможливлює їхнє співіснування з актинобактеріями – джерелом *van*-генів.

Драйвери цієї події ГГТ не вдалося встановити, хоча ми виявили, що численні гени транспозаз і резолваз/рекомбіназ/інтеграз були колокалізовані з *van*-локусами у цих *Ktedonobacteria* і *Clostridia* spp. Можливо, тут ми спостерігаємо певну «ерозію» геномного ландшафту навколо *van*-генів, коли гени МГЕ втрачалися після ГГТ події, унеможливлюючи їх ідентифікацію на даний момент. Проте, IS-елементи таки могли відіграти певну роль в перенесенні *van*-генів до ктедонобактерій, бо ці мікроорганізми мають вражаючу кількість генів IS-транспозаз та інтактних IS-елементів поруч із *van*-генами. Примітно, що *van*-локуси *Ktedonobacteria* spp. піддалися цікавим модифікаціям: в усіх випадках гени, що кодують VanX, були втрачені, тоді як нові гени пептидаз різних типів були набуті. Згідно з нашими результатами, принаймні деякі з цих генів пептидаз були набуті як гени-пасажири IS-елементів. Ктедонобактерії мають спосіб життя подібний до актиноміцетного [202]. Відповідно, ці мікроорганізми могли рекрутувати *van*-гени для захисту від продуцентів ГПА (з якими вони ділять біотопи), авторезистентності до якихось ендогенних метаболітів або для ремоделювання клітинної стінки в процесі морфогенезу. Кейс *van*-генів в ктедонобактерій є найцікавішим та вимагає подальшого експериментального вивчення.

Окрім глибоко вкорінених в корову кладу *Actinobacteria* послідовностей VanHA, консенсусне дерево мало п’ять чітко визначених клад. Ми інтерпретували філогенію Van-білків виходячи із двох припущень: (1) кожна з цих п’яти клад виникла в результаті однієї події ГГТ від *Actinobacteria*; (2) базальні таксони кожної клади є вихідними точками, куди *van*-локуси були

передані первинною подією ГГТ, а далі поширилися до інших таксонів клади. Це дало змогу створити сценарій розповсюдження *van*-генів.

Зокрема, у кладі (I) первинна подія ГГТ, ймовірно, перенесла *van*-локуси до деяких клостридій. Клада (II) є ілюстрацією кількох інших подій ГГТ на додаток до первинної. Більшість таксонів цієї клади належить до класу *Bacilli*. Послідовності, що знайдені в бактерії класу *Anaerolineae* розташовувалися далеко від основи клади (II), даючи змогу інтерпретувати це як доказ події ГГТ, яка перенесла *van*-гени від *Bacilli* spp. до *Anaerolineae*. Останній клас є Грам-негативними бактеріями [203], носії *van*-генів серед яких раніше не документувалися. Той самий сценарій справедливий також для двох *Clostridia* spp., знайдених в кладі (II). Принаймні в одному таксоні клади (II) *van*-локуси виявилися на плазміді, даючи змогу припустити, що носій, який переніс *van*-гени до клади (II), міг бути плазмідою.

Усі таксони, що утворюють кладу (III), належать до *Clostridia* spp. за винятком одного. Винятками є послідовності VanHA, що знайдені в представника класу *Erysipelotrichia*, які були вкорінені глибоко в кладі (III). Така топологія, швидше за все, вказує на те, що *Erysipelotrichia* отримали *van*-локуси від деяких *Clostridia* spp. в результаті додаткової події ГГТ. Важливо також відзначити, що *van*-локуси в базальному таксоні клади (III) розташовані поруч із ділянкою деградованого профага. Отже, саме фаги могли перенести *van*-гени до організмів клади (III).

Клада (IV) складалася з послідовностей VanHA, що походять із класів *Bacilli*, *Erysipelotrichia*, *Clostridia* і типу *Actinobacteria*. Базальна гілка дерева була представлена послідовностями з бактерій, виділених із ґрунту [204] та стічних вод [205], відповідно. Ніякі гени, пов'язані з МГЕ, не супроводжували *van*-локуси у цих організмів. Однак у всіх інших таксонах клади (III) *van*-гени були знайдені у складі Tn*1549*-подібних транспозонів.

На відміну від усіх інших клад, клада (V) складалася виключно з *Bacilli* spp. Базальні таксони клади (V) були переважно мешканцями ґрунту і не демонстрували жодних чітких повторюваних патернів у пов’язаних з МГЕ генах, колокалізованих спільно з *van*-локусами. Наше найбільше зацікавлення

викликала субклада (V'), яка складалася переважно з послідовностей VanHA, закодованих на плазмідах, що несуть Tn*1546*-подібні транспозони з різних клінічних ізолятів стафілококів і ентерококів. Базальною гілкою субклади (V') є послідовності VanHA з клінічного ізоляту умовно-патогенної бактерії *B. circulans* VR0709, що живе в ґрунті [200]. Загалом така топологія клади (V) свідчить про те, що невідомий носій переніс *van*-локуси до *B. circulans* VR0709 із генофонду, який існує в базальних таксонах клади (V). Як наслідок, плазміда, що містить Tn*1546*, мобілізувала *van*-локуси з *B. circulans* VR0709 і опосередкувала подальше їх розповсюдження в нозокоміальних середовищах серед стафілококів і ентерококів. Це останнє спостереження фактично підтверджує попередній висновок, зроблений на основі філогенетичного аналізу, проведеного більше десяти років тому на менш розширеній вибірці доступних послідовностей *van* генів [206].

Наші результати переконливо демонструють, що *van*-гени з бактерій типу *Actinobacteria* є цікавою та розповсюдженою частиною мобілому *Terrabacteria*. Перенесення *van*-генів є активним процесом, який іде в реальному часі. Це демонструють результати нашого філогенетичного аналізу стосовно ктедонобактерій. *van*-гени переносяться різними носіями з одного таксону бактерій до іншого, і назад, створюючи складну мережу послідовних подій ГГТ. Бактерії з класів *Bacilli* і *Clostridia*, схоже, служать джерелом *van*-генів для представників класів *Anaerolineae* і *Erysipelotrichia*. В підсумку, наші результати створюють загальну картину розповсюдження *van* генів та намічають шляхи до подальшого експериментального вивчення цієї проблеми.

Результати цього підрозділу опубліковано в [207].

genes

MDPI

Article

Occurrence of *vanHAX* and Related Genes beyond the *Actinobacteria* Phylum

Oleksandr Yushchuk [1,2], Elisa Binda [1], Victor Fedorenko [2] and Flavia Marinelli [1,*]

[1] Department of Biotechnology and Life Sciences, University of Insubria, 21100 Varese, Italy
[2] Department of Genetics and Biotechnology, Ivan Franko National University of Lviv, 79005 Lviv, Ukraine
* Correspondence: flavia.marinelli@uninsubria.it; Tel.: +39-0332-42-1546

Abstract: Clinically relevant glycopeptide antibiotics remain among the most successful classes of natural antibacterials. This success, however, is endangered by the spread of glycopeptide resistance genes, also known as *van* genes. Thus, it is important to trace and comprehend possible routes of *van* gene dissemination. In the current work, we present a comprehensive bioinformatic analysis aimed at mapping the occurrence of *van* genes beyond the *Actinobacteria* phylum—the most likely natural reservoir of *van* genes. We show that two additional classes of Gram-positive bacteria, *Erysipelotrichia* and *Ktedonobacteria*, as well as one class of Gram-negative bacteria, *Anaerolineae*, carry *van* genes. Additionally, we demonstrate that various new genera belonging to the classes *Clostridia* and *Bacilli* also carry *van* genes. The majority of discovered *van* loci are co-localized with MGE-related genes of various types. Finally, we propose a phylogeny-based scenario for the spread of *van* genes, unraveling a network of consequential horizontal gene transfer events linking the phylum *Actinobacteria* with the five other bacterial classes carrying *van* genes.

Keywords: antibiotic resistance; glycopeptide resistance; *van* genes; mobile genetic elements

check for updates

Citation: Yushchuk, O.; Binda, E.; Fedorenko, V.; Marinelli, F. Occurrence of *vanHAX* and Related Genes beyond the *Actinobacteria* Phylum. *Genes* **2022**, *13*, 1960. https://doi.org/10.3390/genes13111960

Academic Editors: Ruichao Li, Ning Dong and Cemil Kürekci

Received: 23 September 2022
Accepted: 23 October 2022
Published: 27 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Glycopeptide antibiotics—more precisely, dalbaheptides—represent a large group of natural products produced by soil-dwelling, high-G-C-content, and Gram-positive actinomycetes [1,2]. Dalbaheptides are products of sophisticated biosynthetic machinery involving the non-ribosomal synthesis of the heptapeptide aglycone, followed by numerous and variable modifications. Genes encoding these pathways are co-localized, co-expressed, and organized in large biosynthetic gene clusters (BGCs). Dalbaheptide biosynthesis has been extensively studied and reviewed in recent years [3–9].

Although biosynthetic and regulatory mechanisms behind the production of dalbaheptides are interesting *per se* and represent a valuable model for studying other natural products, the real importance of dalbaheptides comes from their clinical relevance [10–13]. Today, two natural (vancomycin and teicoplanin) and three semi-synthetic (dalbavancin, telavancin, and oritavancin) dalbaheptides are used to treat infections caused by multidrug-resistant (MDR) pathogens [8]. As potent inhibitors of Gram-positive cell-wall biosynthesis, dalbaheptides strongly and selectively bind lipid II [14], essential for the biosynthesis of the bacterial cell wall [15]. They specifically interact with the D-alanyl-D-alanine (D-Ala-D-Ala) termini of the lipid II pentapeptide side stems by forming five defined hydrogen bonds, thus impeding the following transpeptidation and transglycosylation steps [9,16,17].

As Gram-positive bacteria, actinobacterial dalbaheptide producers require the expression of self-resistance mechanisms to avoid suicide during antibiotic biosynthesis and excretion. Resistance mechanisms in model dalbaheptide producers have been investigated in detail in recent years [18–21]. The most relevant dalbaheptide resistance mechanism involves at least five genes—known as *van* genes (i.e., *van*comycin resistance genes)—commonly organized in two operons [22,23]. The first operon consists of three genes—*vanHAX*—coding

for structural enzymes required for cell-wall remodeling. Here, VanH is an α-ketoacid dehydrogenase, transforming pyruvate to D-lactate (D-Lac); VanA is a D-Ala-D-Lac ligase, while VanX is a D,D-dipeptidase [24]. The second operon—*vanRS*—encodes a two-component regulatory system tuning *vanHAX* expression in response to the extracellular presence of a glycopeptide antibiotic. Dalbaheptide-producing actinobacteria carry *vanHAXRS* genes adjacent to or within the corresponding glycopeptide BGCs.

Based on extensive genomic analysis, we recently reported [25] that *vanHAXRS* genes are also frequently present in dalbaheptide non-producing actinobacteria, indicating that there is no unique correlation between the presence of glycopeptide BGCs and *van* genes. However, it remains unclear why dalbaheptide non-producing actinobacteria carry *vanHAXRS* genes so frequently. Actinobacteria—either dalbaheptide-producing or not—are the most likely primary environmental reservoir of *van* genes, further spreading to other bacterial taxa [25,26]. The spread of *van* genes among opportunistic pathogens such as enterococci or staphylococci represents a hot topic of investigation due to their clinical relevance [27–47]. On the other hand, the migration of *vanHAXRS* genes to non-harmful soil-dwelling bacteria other than *Actinobacteria* has not attracted much immediate attention, although they likely serve as a bridge to distribute *van* genes from *Actinobacteria* spp. to dangerous pathogens [26]. Previous reports highlighted that *van* genes are also present in two additional classes of Gram-positive soil bacteria, namely, the class *Bacilli* and the class *Clostridia* [48–52]. As reported in ESM Table S1, we analyzed in detail the precedent literature and the genomic data available on the bacterial species carrying *vanHAXRS* genes other than those belonging to the *Actinobacteria* phylum. It was observed that *vanHAXRS* genes were mainly found in enterococci (*Enterococcus* (*Ecc.*) *fecalis* and *Enterococcus faecium*) and *Staphylococcus* (*Scc.*) *aureus* strains, followed by *Paenibacillus* spp. and a few other species belonging to the *Bacilli* and *Clostridia* classes (for example, in *Clostridioides* (*Cld.*) *difficile* AI0499). In many cases, genes were found to be associated with mobile genetic elements (MGEs): transposons and plasmids (ESM Table S1), which are probably involved in their interspecies migration.

Considering the current progress in bacterial genome sequencing, in this work, we screened the vast amount of current publicly available genomic records to investigate *vanHAXRS* gene distribution in species other than those belonging to the *Actinobacteria* phylum. Our findings demonstrate for the first time that *vanHAXRS* and related genes might be found in *Anaerolineae* [53], *Erysipelotrichia* [54], and *Ktedonobacteria* [55] classes. Additional information on *van* loci spread within *Bacilli* and *Clostridia* classes is also reported. Further phylogenetic analysis of Van proteins following the one published earlier on actinobacterial Van proteins [25] adds more details to the overall picture of how *vanHAXRS* genes spread among bacterial taxa.

2. Methods

2.1. *van* Genes Discovery

Amino acid (aa) sequences of SCO3594, SCO3595, and SCO3596 (VanH, VanA, and VanX) from *Streptomyces* (*S.*) *coelicolor* A3(2) were used as queries for BLASTP [56] against non-redundant protein sequence databases for each bacterial class listed in the NCBI Taxonomy database [57]. The BLASTP search used default algorithm parameters, but the maximum number of aligned sequences to display was set to 5000. The analysis was performed in May 2022. All obtained hits (with no regard to the corresponding e-values) were further investigated manually and mapped to the corresponding genomic records retrieved from either GenBank [58] or RefSeq [59] databases. Every bacterial taxon in which genes coding for SCO3594-95-96 homologs were found to be co-localized was further investigated. In total, we obtained 1 set of *van* genes for *Anaerolineae*, 48 for *Bacilli*, 23 for *Clostridia*, 2 for *Erysipelotrichia*, and 7 for *Ktedonobacteria*; the accession numbers for the discovered nucleic acid and aa sequences are summarized in Supplementary Excel File S1. Metagenome-assembled genomes (MAGs) were not considered in further analyses. Degraded *van* loci (lacking either *vanH*, *vanA*, or *vanX* homologs), such as those previously

observed in *Desulfitobacterium* (*Dsf.*) *hafniense* strains [52] and in some *Bacilli* spp. [25], were omitted to avoid overcomplicating the general picture.

Routine analysis with all amino and nucleic acid sequences was conducted using GENEIOUS 4.8.5 [60].

2.2. Mapping MGE-Related Genes

To discover and map MGE-related genes co-localized with *van* loci, aa sequences coded by ORFs from ca. 50 kbp (or less if not available) regions up- and downstream of *van* loci were analyzed with either CD-Search or Batch CD-Search [61,62]. This allowed us to identify proteins involved in DNA transfer and recombination or prophage-related proteins. Proteins that carried conserved domains typical for IS transposases were further used as queries for BLASTP (default algorithm parameters) against the ISfinder database [63], allowing us to identify those belonging to different IS families. If necessary, inverted repeat regions were identified using the EMBOSS einverted tool [64] with the default algorithm parameters, although the "maximum extent of repeats" parameter was manually adjusted for each search to cover the whole region of interest. Putative prophage regions were inspected using PHASTER [65] and PHISDetector [66] web servers under default search conditions.

The accession numbers for MGE-related genes and proteins, as well as additional information are given in Supplementary Excel File S1.

2.3. Phylogenetic Reconstruction

MEGA 11 (v.11.0.13) [67] software was used for the phylogenetic analysis. All trees were generated using the neighbor-joining method [68]. The confidence of the branching was estimated using an interior-branch test (1000 replicates) [69]. The evolutionary distances were computed using the JTT matrix-based method [70]. Ambiguous positions were removed with the pairwise deletion option. The rate variation among sites was modeled with a 5-parameter σ distribution. The phylogeny of VanH proteins was built using 269 aa sequences, including SCO2118 (putative D-Lac dehydrogenase from *S. coelicolor* A3(2)) as an outgroup (Supplementary FASTA File S1); the phylogeny of VanA was built using 265 aa sequences proteins, including SCO5560 (putative D-Ala-D-Ala ligase from *S. coelicolor* A3(2)) as an outgroup (Supplementary FASTA File S2); the phylogeny of VanHA concatenates was constructed by using 262 aa sequences and included concatenated SCO2118-SCO5560 as an outgroup (Supplementary FASTA File S3).

3. Results

Our screening of publicly available genomic records revealed the presence of *vanHAXRS* and related genes in four classes of Gram-positive bacteria beyond the phylum *Actinobacteria*, i.e., *Bacilli*, *Clostridia*, *Erysipelotrichia*, and *Ktedonobacteria*, as well as in one class of Gram-negative bacteria—*Anaerolineae* (Table 1)—all belonging to the *Terrabacteria* superphylum [71]. To our best knowledge, *van* genes have not been previously reported in representatives of the last three classes. Notably, we did not include *van* loci found in MAGs in this analysis. For example, at least one MAG containing *van* loci was also found in the classes *Acidobacteria*, *Fusobacteria*, *Negativicutes*, and *Thermotogae*. However, the analysis of genes up- and downstream of these *van* loci showed them to be identical or homologous to genes from different *Actinobacteria* spp. (namely, micrococci). This allowed us to reasonably suspect anomalies in the MAG assembly, presuming that the final genome reconstruction was "contaminated" with nucleic acid sequences from other bacteria.

In the following subsections, the organization and genetic background of *vanHAXRS* for each taxonomical class is further investigated.

Table 1. List of bacterial strains from *Anaerolineae*, *Bacilli*, *Clostridia*, *Erysipelotrichia*, and *Ktedonobacteria* classes found to carry *van* loci; more detailed information, as well as accession numbers for nucleic and aa sequences, is given in Supplementary Excel File S1.

Classes:	Genera:	Bacterial Strains Carrying *van* Genes:
Anaerolineae	*Aggregatilinea* (*Al.*)	*Aggregatilinea lenta* MO-CFX2
Bacilli	*Alicyclobacillus* (*Alic.*)	*Alicyclobacillus shizuokensis* NBRC 103103
	Alkalihalobacillus	*Alkalihalobacillus* sp. EGI L200015
	Anoxybacillus (*Anox.*)	*Anoxybacillus sediminis* PCH 117
	Bacillus	*Bacillus* sp. JCA
	Brevibacillus (*Bba.*)	*Brevibacillus halotolerans* J5TS2; *Brevibacillus laterosporus* ACRRF; *Bba. laterosporus* G25-128; *Bba. laterosporus* OSY-I1; *Bba. laterosporus* SAM19; *Bba. laterosporus* B9 (plasmid); *Brevibacillus* sp. MCWH; *Brevibacillus* sp. NL20B1; *Brevibacillus* sp. SKDU10; *Brevibacillus* sp. VP
	Cohnella (*Coh.*)	*Cohnella zeiphila* CBP-2801
	Neobacillus (*Nba.*)	*Neobacillus jeddahensis* MGYG-HGUT-01469; *Nba. jeddahensis* JCE
	Paenibacillus (*Pba.*)	*Paenibacillus apiarius* MW-14; *Paenibacillus dendritiformis* J6TS7; *Paenibacillus elgii* MER 157; *Pba. elgii* SWL-W8; *Paenibacillus* sp. DXL2; *Paenibacillus hemerocallicola* KCTC 33185; *Paenibacillus jilunlii* DSM 23019; *Paenibacillus macerans* 3CT49; *Pba. macerans* DSM 24746; *Pba. macerans* HMSSN-036; *Pba. macerans* FSL R5-0527; *Pba. macerans* MER 50.2; *Paenibacillus mesophilus* SYSU K30004; *Paenibacillus nasutitermitis* CGMCC 1.15178; *Paenibacillus oralis* KCOM 3021; *Paenibacillus physcomitrellae* CGMCC 1.15044; *Pba. physcomitrellae* XB; *Paenibacillus piri* MS74; *Paenibacillus sonchi* IIRRBNF1; *Pba. sonchi* X19-5; *Paenibacillus* sp. cl123; *Paenibacillus* sp. SM 69; *Paenibacillus* sp. UNCCL117; *Paenibacillus* sp. YIM B00362; *Paenibacillus* sp. YIM B00624; *Paenibacillus thiaminolyticus* BO5; *Paenibacillus zanthoxyli* JH29; *Paenibacillus flagellatus* DXL2
	Thermoactinomyces	*Thermoactinomyces* sp. CICC 10735
	Thermobacillus (*Tba.*)	*Thermobacillus xylanilyticus* XE
	Weizmannia (*W.*)	*Weizmannia ginsengihumi* Gsoil 114
Clostridia	Dubious genus	[*Clostridium*] *indolis* DSM 755; [*Clostridium*] *methoxybenzovorans* SR3; [*Clostridium*] *scindens* MSK.5.24; [*Clostridium*] *symbiosum* MSK.7.21
	Unidentified genus	*Lachnospiraceae* bacterium M18-1
	Aminipila (*Am.*)	*Aminipila terrae* CBA3637
	Anaerobacterium (*An.*)	*Anaerobacterium chartisolvens* DSM 27016
	Bariatricus (*Bar.*)	*Bariatricus massiliensis* DFI.1.181
	Beduini (*Bed.*)	*Beduini massiliensis* GM1
	Blautia (*Bl.*)	*Blautia producta* ER3; *Blautia pseudococcoides* SCSK; *Blautia* sp. NBRC 113351
	Candidatus Formimonas	*Candidatus* Formimonas warabiya DCMF
	Clostridium (*Cli.*)	*Clostridium argentinense* 113/29; *Clostridium* sp. M3/9; *Clostridium tagluense* CS002
	Cuneatibacter	*Cuneatibacter* sp. NSJ-177
	Extibacter (*Ext.*)	*Extibacter muris* DSM 28560
	Lachnotalea	*Lachnotalea* sp. AF33-28
	Oxobacter (*Ox.*)	*Oxobacter pfennigii* DSM 3222
	Ruminococcus (*Rcc.*)	*Ruminococcus gauvreauii* DSM 19829; *Ruminococcus gauvreauii* MGYG-HGUT-01690
	Ruthenibacterium (*Rba.*)	*Ruthenibacterium lactatiformans* 668

Table 1. *Cont.*

Classes:	Genera:	Bacterial Strains Carrying *van* Genes:
Erysipelotrichia	Unidentified genus	*Erysipelotrichaceae* bacterium 66202529
	Longicatena (*Lct.*)	*Longicatena caecimuris* 3BBH23
Ktedonobacteria	Unidentified genus	*Ktedonobacteria* bacterium brp13
	Dictyobacter (*Db.*)	*Dictyobacter alpinus* Uno16
	Ktedonobacter (*Kb.*)	*Ktedonobacter racemifer* DSM 44963; *Ktedonobacter* sp. SOSP1-52
	Ktedonosporobacter (*Ksb.*)	*Ktedonosporobacter rubrisoli* SCAWS-G2
	Reticulobacter (*Rb.*)	*Reticulibacter mediterranei* 150040
	Thermosporothrix	*Thermosporothrix* sp. COM3

3.1. *van* Genes in Class *Anaerolineae*

A single set of *van* genes was discovered in one of the species from a still poorly investigated class of Gram-negative bacteria—*Anaerolineae* [53]. Here, the organization of *van* genes in *Al. lenta* MO-CFX2 SCF03 (*vanRSWHAX*, Figure 1a, Supplementary Excel File S1 "Anaerolineae" worksheet) resembled the one described in the *Ecc. faecalis* BM4382 pIP834^{+} plasmid-borne Tn*1549* transposon, although it lacked *vanY* homolog. In addition, a set of genes coding for putative transposases, recombinases, and integrases was found near the *van* locus (see Supplementary Excel File S1 "Anaerolineae MGERG"). However, none of them showed significant similarity to the counterparts from previously published MGEs carrying *van* genes (summarized in ESM Table S1).

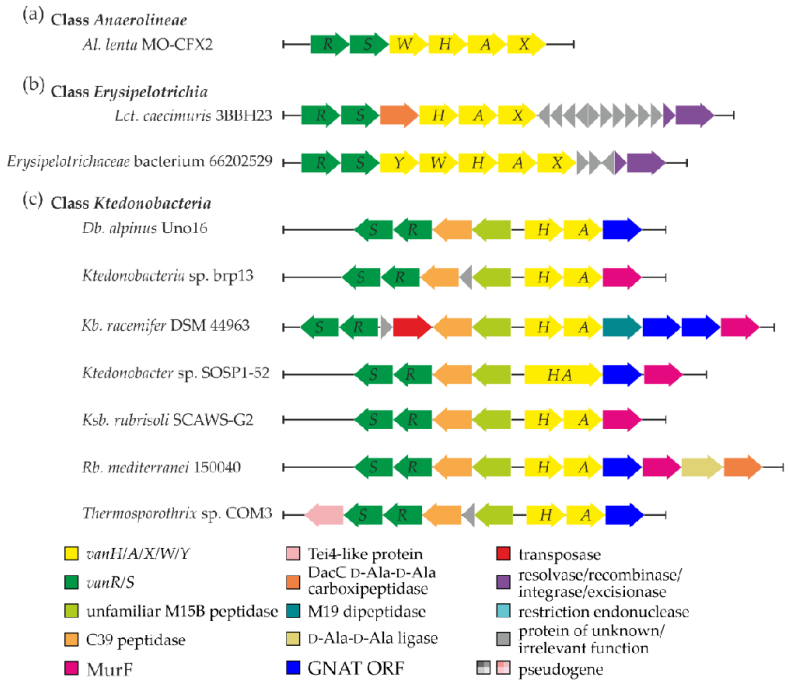


Figure 1. Graphical summary of *van* and other relevant co-localized genes discovered in *Anaerolineae* (**a**), *Erysipelotrichia* (**b**), and *Ktedonobacteria* (**c**) classes. The color coding of the schemes is explained in the legend of the figure. Please see the main text for more details.

3.2. *van* Genes in Class *Erysipelotrichia*

Most studied representatives of the class *Erysipelotrichia* belong to the gut microbiota, while the class itself was recently delineated as a sister taxon of the *Clostridia* class within the phylum *Firmicutes* [72]. Our screen indicated the presence of *van* genes in the genomes of *Lct. caecimuris* 3BBH23 (isolated from healthy human feces [73]) and of unclassified clinical *Erysipelotrichaceae* isolate 66202529 (CP046174, Supplementary Excel File S1 "Erysipelotrichia"). Both *vanRS* and *vanHAX* operons were found in *Lct. caecimuris* 3BBH23, with a gene coding for a VanYD DacC-like carboxypeptidase [74] in between (Figure 1b). The chromosome of *Erysipelotrichaceae* isolate 66202529 carried a set of *vanRSYWHAX* genes (Figure 1b), like in the *Ecc. faecium* MLG229 Tn1549-like transposon. In both *Lct. caecimuris* 3BBH23 and isolate 66202529, we found a pair of genes coding for excisionase and a tyrosine-based site-specific integrase downstream of *van* loci (Figure 1b), as well as multiple other MGE-related genes (Supplementary Excel File S1 "Erysipelotrichia"). Peculiarly, the integrase coded downstream of *van* loci in *Lct. caecimuris* 3BBH23 lacked any significant similarity to homologs in previously published MGEs carrying *van* genes (summarized in ESM Table S1). Thus, *van* loci in *Lct. caecimuris* 3BBH23 are likely to be a part of a novel MGE. On the contrary, the integrase from isolate 66202529 appeared to share 100% aa sequence identity with the integrases found downstream of *van* loci in *Ecc. faecium* MLG229 and *Clostridium* sp. MLG245 Tn*1549*-like transposons, implying that isolate 66202529 carries the latter MGE.

3.3. *Peculiarities of Organization and Genetic Context of van Genes in Class Ktedonobacteria*

For the first time, our screen revealed that *van* genes are present in representatives of the class *Ktedonobacteria* (phylum *Chloroflexi*) (Table 1, Supplementary Excel File S1 "Ktedonobacteria"). These enigmatic organisms belong to an unusual group of filamentous Gram-positive soil dwellers. *Ktedonobacteria* spp. resemble filamentous sporulating actinomycetes in certain aspects of morphology and lifestyle [55,75], implying that a sophisticated machinery for cell-wall rearrangement might exist there.

The arrangements and genetic contexts of *van* genes in *Ktedonobacteria* spp. were unprecedented: in all of them, we observed a tendency to lose *vanX* (Figure 1c, Supplementary Excel File S1 "Ktedonobacteria"). However, this loss was correlated with the gain of a novel gene pair positioned between *vanRS* and *vanHA* (Figure 1c). Remarkably, these two genes were found to code for peptidases. One peptidase belonged to the MEROPS [76] M15B subfamily of metallopeptidases, while the second resembled peptidases of the MEROPS C39 family. Typical VanX proteins also belong to the M15B subfamily but are quite different from M15B metallopeptidases from the *Ktedonobacteria* spp. mentioned above. VanX from *S. coelicolor* A3(2) (taken as a reference) shares only ca. 15% aa sequence identity with any of the *Ktedonobacteria* spp. M15B metallopeptidases, and it is shorter (202 aa vs. ca. > 320 aa). Moreover, *Kb. racemifer* DSM 44963 contained an additional gene coding for a MEROPS M19 family dipeptidase just after *vanA* (thus directly replacing the canonical position of *vanX*, Figure 1c). Other relevant genes were co-localized with *van* loci in *Ktedonobacteria* spp., including genes for MurF (a UDP-*N*-acetylmuramoyl-tripeptide ligase), a GCN5-related N-acetyltransferase (GNAT), a VanYD-like DacC D-Ala-D-Ala carboxypeptidase, and a D-Ala-D-Ala ligase (Figure 1c, Supplementary Excel File S1). Another interesting feature was observed in *Ktedonobacter* sp. SOSP1-52, where *vanH* and *vanA* were fused (Figure 1c); however, it remains unknown whether the fusion protein is produced or whether what we observed is just an error of genome sequencing and annotation.

Further analysis of the close genomic neighborhood of *van* loci in *Ktedonobacteria* spp. (except *Db. alpinus* Uno16) revealed a striking number of genes coding for transposases (often incomplete or frameshifted, Supplementary Excel File S1 "Ktedonobacteria MGERG") belonging to different insertion sequence (IS) families. Most of them had DDE catalytic residues [77]. In addition, we discovered multiple short and long inverted repeats. Genes coding for recombinase/integrase-like enzymes were also found (Supplementary Excel

File S1 "Ktedonobacteria MGERG"), although they did not show repetitive patterns across different *Ktedonobacteria* spp.

To study the last aspect more comprehensively, we created a detailed map of the genomic neighborhood of *van* loci in *Kb. racemifer* DSM 44963 (given in Figure 2). The examined region (ca. 50 kbp up- and downstream of *van* loci) contained 21 genes for transposases (4 of them were pseudogenes), 2 genes for integrase-like enzymes, and one gene for a restriction endonuclease. Three segments surrounded by large inverted repeat regions were found upstream of the *van* loci (Figure 2): in two cases, these inverted repeats were duplicated genes for transposases, while the third case involved the duplication and inversion of the *vanRS* locus. Moreover, two intact ISKra4-like IS elements and one ISKra7-like IS element were found up- and downstream of the *van* loci, respectively. Another putative IS element composed of a gene for an IS1380-family transposase with 36 bp inverted repeats was found downstream of *van* loci; however, neither the transposase nor the repeat regions were significantly similar to any entries of the ISfinder database [63]. Finally, two peptidase genes (M15B and C39, see above) might also belong to an unknown IS element in *Kb. racemifer* DSM 44963. Both genes were co-localized with an IS200/IS605-family transposase gene and flanked with 58 bp inverted repeats (Figure 2). Interestingly, the same inverted repeats were found flanking M15B- and C39-peptidase genes in other *Ktedonobacteria* spp., although the IS200/IS605-family transposase gene was present only in *Kb. racemifer* DSM 44963.

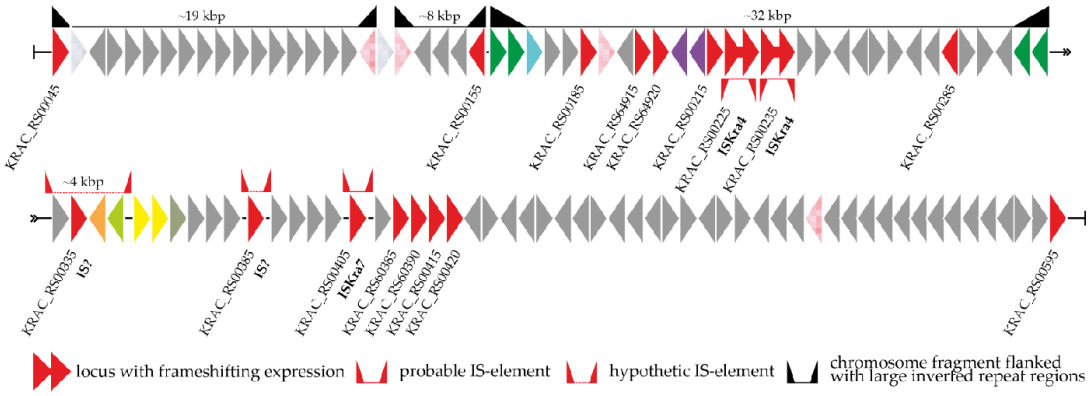


Figure 2. Detailed map of the genomic context of *van* loci in *Kb. racemifer* DSM 44963. The color coding of the scheme is explained in the Figure 1 legend. Please see the main text for more details.

3.4. Updates on vanHAXRS Gene Distribution in Bacilli Class

Bacteria belonging to the *Bacilli* class are well known to carry *vanHAXRS* genes; these are either opportunistic pathogens such as staphylococci and enterococci or non-harmful soil dwellers (see summary in ESM Table S1). Since our previous bioinformatic analysis [25], multiple novel genomic records for *Bacilli* spp. have become available in GenBank and RefSeq. This allowed us to significantly update the picture of the *vanHAXRS* distribution in the *Bacilli* class: *vanHAXRS* genes were discovered in 48 novel strains from this class (Table 1, Supplementary Excel File S1"Bacilli").

The arrangements of *van* genes varied across different strains and are summarized in Figure 3. The simplest one—only *vanHAX* genes with no MGE-related genes nearby—was discovered in *Pba. dendritiformis* J6TS7 (Figure 3a, Supplementary Excel File S1 "Bacilli"). An additional *vanS* pseudogene was present in *Pba. flagellatus* DXL2 (Figure 3b, Supplementary Excel File S1 "Bacilli"); however, *vanS* was found at the 5′-end of the available contig, suggesting that it could be intact and likely followed by a *vanR* gene. No adjacent MGE-related genes were discovered in *Pba. flagellatus* DXL2, either. The *van* genes in *Brevibacillus* sp. SKDU10 also lacked the *vanRS* locus, and they were arranged as *vanHAXYZ*. A *vanH* pseu-

dogene, another copy of *vanZ*, and a transposase pseudogene (Figure 3c, Supplementary Excel File S1 "Bacilli", "Bacilli MGERG") were discovered further downstream.

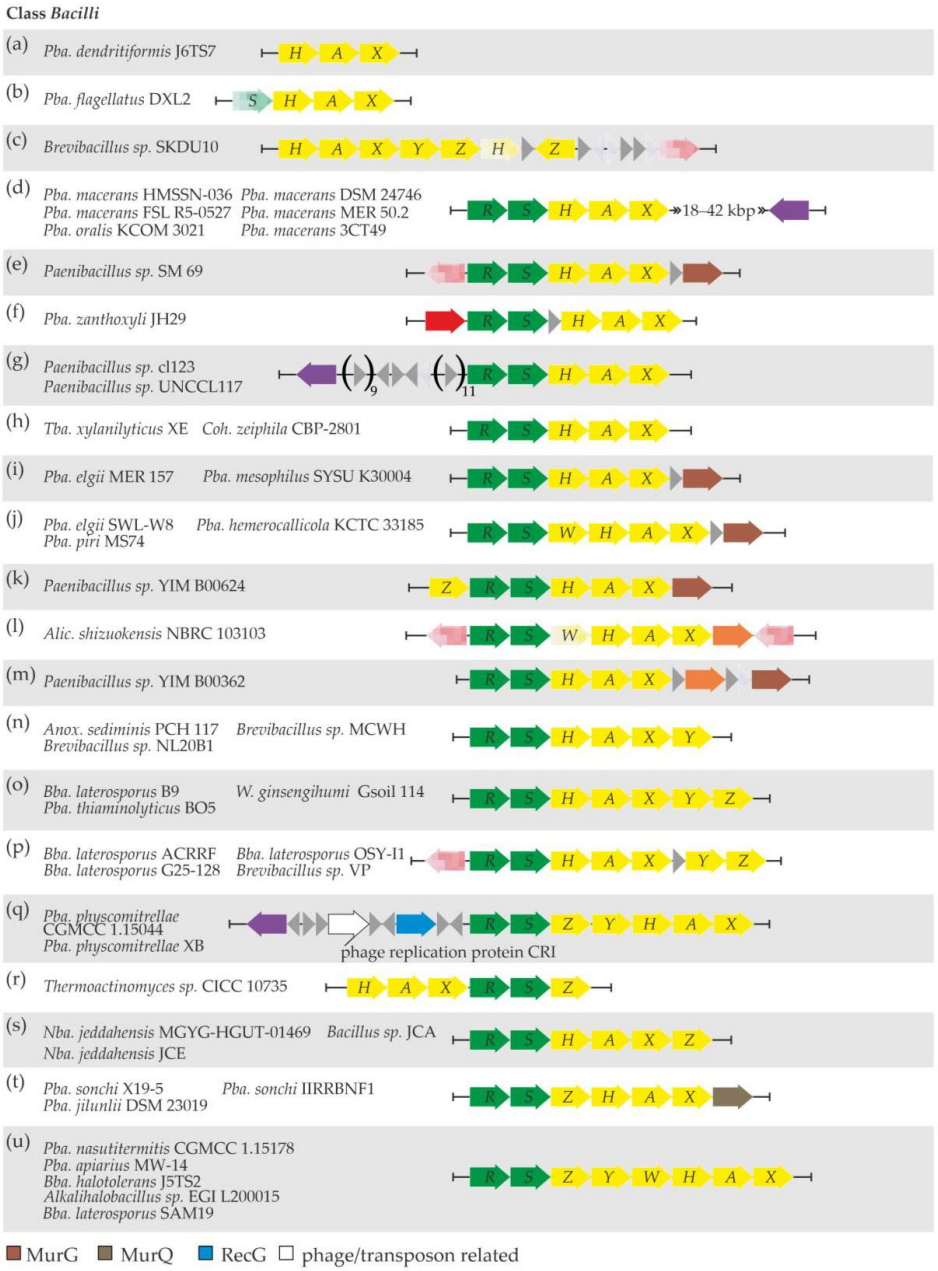


Figure 3. Schematic representation of the arrangement patterns observed in the *van* and relevant co-localized genes from the *Bacilli* class. Twenty one arrangement patterns (**a**–**u**) were determined and are discussed in the main text. To interpret the color coding of Figure 3, please also refer to the legend in Figure 1.

A set of five *Pba. macerans* strains, together with *Pba. oralis* KCOM 3021 (Figure 3d), carried *vanRSHAX* genes with a highly conserved gene for a tyrosine-type site-specific recombinase lying downstream. This recombinase did not share any similarity with the recombinases already reported to be co-localized with *van* loci. Some other MGE-related genes were also discovered up- and downstream of *van* loci in these six strains, although they were not strictly conserved (Supplementary Excel File S1 "Bacilli MGERG"). Furthermore, *van* genes were organized as *vanRSHAX* in eight additional strains (Figure 3e–i, Supplementary Excel File S1 "Bacilli"). In addition, *van* loci were co-localized with transposase genes in *Paenibacillus* sp. SM 69 and *Pba. zanthoxyli* JH29 and with a gene for an unprecedented tyrosine-type site-specific recombinase in *Paenibacillus* sp. cl123 and *Paenibacillus* sp. UNCCL117 (Supplementary Excel File S1 "Bacilli MGERG"). Notably, in *Paenibacillus* sp. SM 69, *Pba. elgii* MER 157, and *Pba. mesophilus* SYSU K30004, a gene for UDP-*N*-acetylglucosamine-*N*-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol *N*-acetylglucosamine transferase (MurG) was found downstream of *van* loci (Figure 3e,i). A slightly different arrangement—*vanRSWHAX*—was observed in three other *Paenibacillus* spp. (Figure 3j). Here, no MGE-related genes were discovered nearby, but a gene for MurG was again found downstream of *van* loci (Supplementary Excel File S1 "Bacilli").

In *Paenibacillus* sp. YIM B00624, *van* genes were organized as *vanZRSHAX* (Figure 3k), again with no MGE-related genes nearby, but followed by a gene for MurG (Supplementary Excel File S1 "Bacilli").

Only two strains, namely, *Alic. shizuokensis* NBRC 103103 and *Paenibacillus* sp. YIM B00362, were found to carry a gene for a DacC-like VanYD carboxypeptidase within *van* loci, organized as *vanRS(W pseudogene)HAX(YD)* and *vanRSHAX(YD)*, respectively (Figure 3l,m). In *Alic. shizuokensis* NBRC 103103, *van* loci were surrounded by two transposase pseudogenes, while *Paenibacillus* sp. YIM B00362 had no MGE-related genes near *van* loci (Supplementary Excel File S1 "Bacilli MGERG"). As in some of the other abovementioned *Paenibacillus* spp., *Paenibacillus* sp. YIM B00362 had a gene for MurG just downstream of *van* genes (Supplementary Excel File S1 "Bacilli").

Next, in *Anox. sediminis* PCH 117, *Brevibacillus* sp. MCWH, and *Brevibacillus* sp. NL20B1 *van* genes were organized as *vanRSHAXY* (Figure 3n). No MGE-related genes were found to co-localize with *van* genes in these three strains. One additional gene—*vanZ*—was found downstream of *vanRSHAXY* in *Bba. laterosporus* B9, *Pba. thiaminolyticus* BO5, and *W. ginsengihumi Gsoil* 114 (Figure 3o). Notably, in *Bba. laterosporus* B9, *van* loci were carried on a plasmid. Another group of four *Brevibacillus* spp. had a similar arrangement of *van* genes: *vanRSHAX* and *vanYZ* loci separated by a non-correlated gene (Figure 3p), also sharing a transposase pseudogene just upstream of *vanR* (Supplementary Excel File S1 "Bacilli MGERG").

Pba. physcomitrellae CGMCC 1.15044 and *Pba. physcomitrellae* XB represented a rather unique case among *Bacilli* spp., where *van* genes were organized as *vanRSZYHAX*, and a set of three MGE-related genes was found upstream (Figure 3q). These three genes code for an ATP-dependent DNA helicase RecG, a CRI phage replication protein, and an unprecedented serine-type recombinase (Supplementary Excel File S1 "Bacilli MGERG"). The RecG and recombinase proteins shared high similarity with counterparts in the genomes of phages of the order *Caudovirales* [78].

Three distinct combinations of *vanH*, *A*, *X*, *R*, *S*, and *Z* genes were observed in the following strains: *vanHAXRSZ* in *Thermoactinomyces* sp. CICC 10735 (Figure 3r); *vanRSHAXZ* in *Nba. jeddahensis* MGYG-HGUT-01469, *Nba. jeddahensis* JCE, and *Bacillus* sp. JCA (Figure 3s); and *vanRSZHAX* in *Pba. sonchi* X19-5, *Pba. sonchi* IIRRBNF1, and *Pba. jilunlii* DSM 23019 (Figure 3t). In *Thermoactinomyces* sp. CICC 10735, *Nba. jeddahensis* MGYG-HGUT-01469, *Nba. jeddahensis* JCE, *Pba. sonchi* X19-5, and *Pba. jilunlii* DSM 23019, *van* loci were not accompanied by MGE-related genes, while the genomic neighborhood of *van* genes in *Pba. sonchi* IIRRBNF1 and *Bacillus* sp. JCA was enriched with genes and pseudogenes for transposases of different types (Supplementary Excel File S1 "Bacilli MGERG").

In *Pba. sonchi* X19-5, *Pba. sonchi* IIRRBNF1, and *Pba. jilunlii* DSM 23019, a gene coding for an *N*-acetylmuramic acid 6-phosphate etherase (MurQ) was found just downstream of *vanX*.

Finally, the *vanRSZYWHAX* arrangement was discovered in *Pba. nasutitermitis* CGMCC 1.15178, *Pba. apiarius* MW-14, *Bba. halotolerans* J5TS2, *Alkalihalobacillus* sp. EGI L200015, and *Bba. laterosporus* SAM19 (Figure 3u). In each of these strains, *van* loci were accompanied by variable sets of genes and pseudogenes for transposases of different types (Supplementary Excel File S1 "Bacilli MGERG"). Notably, a gene for an unprecedented tyrosine-type site-specific integrase was found downstream of *van* loci in *Pba. apiarius* MW-14.

3.5. Updating the Picture of vanHAXRS Gene Distribution in Clostridia Class

To date, at least two *Clostridia* spp. carrying *vanHAX* and *vanRS* loci as a part of a Tn*1549*-like transposon have been reported (ESM Table S1). Bizarrely recombined *van* loci lacking *vanH* orthologs were also discovered in different strains of *Dsf. hafniense* [52] (ESM Table S1). Our screen added new results to previous data: *vanHAX*, *vanRS*, and related genes were found within the genomic records of 23 strains from the *Clostridia* class (Table 1, Supplementary Excel File S1 "Clostridia"). The arrangements of *vanHAX*, *vanRS*, and related genes were found to be quite variable.

The simplest one—the *vanHAX* locus only—was observed in *Am. terrae* CBA3637 and *Bl. producta* ER3 (Figure 4a, Supplementary Excel File S1 "Clostridia"). Additionally, in *Am. terrae* CBA3637, genes for UDP-*N*-acetylmuramoyl-L-alanyl-D-glutamate-2,6-diaminopimelate ligase (MurE) and a serine-based recombinase were found upstream of *vanHAX* (Figure 4a, Supplementary Excel File S1 "Clostridia" and "Clostridia MGERG"). This recombinase did not share any similarity with other *van*-loci-associated recombinases published previously.

The *vanHAXRS* arrangement was found in *Bed. massiliensis* GM1 (Figure 4b, Supplementary Excel File S1 "Clostridia"), although no MGE-related genes were found nearby.

Next, in *Cli. tagluense* CS002, *van* genes were arranged as *vanHAXYRS* with additional genes for MurE and undecaprenyl-diphosphate phosphatase downstream, and no MGE-related genes were found nearby (Figure 4c, Supplementary Excel File S1 "Clostridia").

The *vanRSHAX* arrangement was observed in *Cuneatibacter* sp. NSJ-177, with an upstream gene coding for a tyrosine-based site-specific recombinase/integrase (Figure 4d, Supplementary Excel File S1 "Clostridia" and "Clostridia MGERG"). Again, this recombinase shared no similarity to the previously reported ones associated with *van* loci.

Cli. argentinense 113/29 exhibited the *vanRSHAXYZ* arrangement of *van* genes, with no MGE-related genes nearby (Figure 4e, Supplementary Excel File S1 "Clostridia").

In *An. chartisolvens* DSM 27016, we found the *vanRS* locus, followed by *vanK*, *vanYD* (coding for a DacC-like carboxypeptidase), and *vanHAX* (Figure 4f, Supplementary Excel File S1 "Clostridia"); the *van* genes were surrounded by genes coding for a RecG-like DNA helicase and a restriction endonuclease. Finally, a gene coding for a serine-based recombinase was found downstream of *van* loci; as in the cases above, this recombinase shared no similarity with previously published *van*-loci-associated recombinases (Supplementary Excel File S1 "Clostridia MGERG").

A *vanRS(YD)HAX* arrangement was found within the genomic records of several *Clostridia* spp. The first group of them included *Lachnospiraceae* bacterium M18-1, *Rcc. gauvreauii* DSM 19829, *Rcc. gauvreauii* MGYG-HGUT-01690, [*Clostridium*] *scindens* MSK.5.24, *Bar. massiliensis* DFI.1.181, and *Ext. muris* DSM 28560 (Figure 4g, Supplementary Excel File S1 "Clostridia"). All six of these bacterial strains shared a common gene for an unknown highly conserved tyrosine-based site-specific recombinase/integrase just after *vanX*. In addition, the genomic neighborhood of van loci in *Lachnospiraceae* bacterium M18-1 and *Ext. muris* DSM 28560 contained a rich repertoire of putative MGE-related genes (Supplementary Excel File S1 "Clostridia MGERG"). Although many of them were homologous to genes found in phages from the order *Caudovirales* [78], neither PHASTER nor PHISDetector tools were able to identify prophage regions near *van* loci in M18-1 and *Ext. muris* DSM 28560.

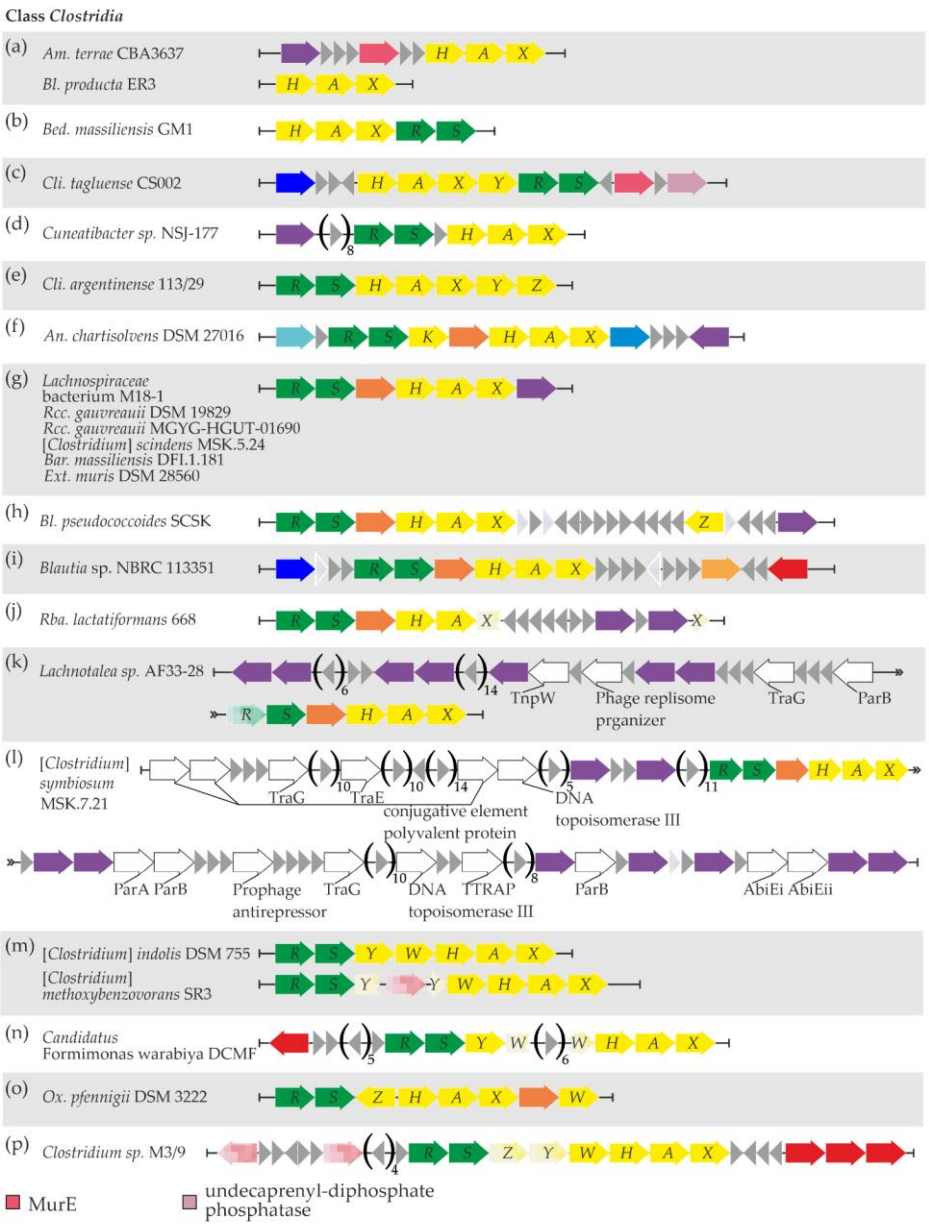


Figure 4. Sixteen (**a**–**p**) arrangement patterns of *van* genes and relevant neighboring genes were discovered in bacteria from the *Clostridia* class. The contents of the figure are discussed in the main text. To interpret the color coding of Figure 4, please also refer to the legends in Figures 1 and 3.

The same arrangement of *van* genes was observed in *Bl. pseudococcoides* SCSK and *Blautia* sp. NBRC 113351 (Figure 4h,i). In the first strain, an additional *vanZ* gene was found downstream of the main *van* loci (Figure 4h, Supplementary Excel File S1 "Clostridia"), as well as a gene for an unprecedented serine-based recombinase (Supplementary Excel File S1 "Clostridia MGERG"). At the same time, a gene for an IS110-family transposase

was downstream of *van* loci in *Blautia* sp. NBRC 113351 (Supplementary Excel File S1 "Clostridia MGERG"). A peculiar case was observed for *Rba. lactatiformans* 668, where the *vanRS(YD)HAX* arrangement was interrupted by the insertion of ten genes into the open reading frame (ORF) of *vanX* (Figure 4j, Supplementary Excel File S1 "Clostridia"). Two of these genes code for a MobM-like relaxase and a serine-based recombinase, respectively (Supplementary Excel File S1 "Clostridia MGERG"); neither of them was similar to the MGE-related genes found before in association with *van* loci. Finally, the *vanRS(YD)HAX* arrangement was also found in *Lachnotalea* sp. AF33-28 and [*Clostridium*] *symbiosum* MSK.7.21 (Figure 4k,l). The genetic context of *van* loci in both of these strains was quite interesting. Multiple genes coding for MGE-related proteins were found upstream of *van* loci in *Lachnotalea* sp. AF33-28 (Supplementary Excel File S1 "Clostridia MGERG"). All of these proteins lacked counterparts in previously published MGEs carrying *van* genes but appeared similar ($\geq$56% aa sequence identity) to proteins encoded in genomes of phages from the *Caudovirales* order [78]. When prophage identification analysis was performed on *Lachnotalea* sp. AF33-28 using PHASTER, an incomplete prophage region was indeed identified upstream of *van* loci. The most similar phage for this region was identified as *Faecalibacterium* phage FP_Brigit (NC_047909). A similar situation was observed in [*Clostridium*] *symbiosum* MSK.7.21 (Supplementary Excel File S1 "Clostridia MGERG"). However, neither PHASTER nor PHISDetector prophage detection tools were able to discover prophages near *van* loci in [*Clostridium*] *symbiosum* MSK.7.21.

Another arrangement—*vanRSYWHAX*—was discovered in [*Clostridium*] *indolis* DSM 755, [*Clostridium*] *methoxybenzovorans* SR3 (Figure 4m, Supplementary Excel File S1 "Clostridia"), and *Candidatus* Formimonas warabiya DCMF (Figure 4n). Notably, in [*Clostridium*] *methoxybenzovorans* SR3, the *vanY* ORF was interrupted by a transposase pseudogene, while in *Candidatus* Formimonas warabiya DCMF, the *vanW* ORF contained a six-gene insertion. An IS1634-family transposase gene was also found upstream of *van* loci in the latter strain (Supplementary Excel File S1 "Clostridia MGERG").

Ox. pfennigii DSM 3222 carried a *vanRSZHAX(YD)W* arrangement (Figure 4o), with no MGE-related genes nearby.

Finally, a *vanRSZYWHAX* arrangement was observed in *Clostridium* sp. M3/9, with *vanZ* and *vanY* genes containing mutations that interrupt ORFs (Figure 4p, Supplementary Excel File S1 "Clostridia"). A set of pseudo- and functional genes for transposases was found up- and downstream of *van* loci (Supplementary Excel File S1 "Clostridia MGERG").

3.6. Building a Consensus Scheme for Phylogenetic Relations between Newly Discovered van Proteins and Those from Phylum Actinobacteria

Thus, our screen revealed multiple novel *van* genes from representatives of five bacterial classes. To integrate this information, we further aimed to fulfill two goals: (a) create a reliable phylogenetic tree representing the evolutionary interrelationships among Van proteins encoded by different bacterial classes; (b) see whether the genetic context of *van* genes (i.e., co-localized MGE-related genes) correlates with their phylogeny.

To achieve these goals, we first reconstructed separate phylogenies for VanH and VanA (i) encoded within the genomes of *Actinobacteria* spp. (using datasets from our previous work [25]), (ii) from nucleic acid sequences of pathogens and non-harmful soil bacteria published before (see ESM Table S1), and (iii) discovered in the current work. Separate phylogenetic reconstructions for the two proteins were coherent (see ESM Figures S1 and S2), allowing us to reasonably assume that *vanHA* (and corresponding proteins) evolved as a single unit (with insignificant exceptions). We moved onward to the creation of a representative phylogenetic tree using concatenated sequences of VanH and VanA proteins, encoded in a single locus. Unfortunately, we could not also use VanX sequences, since all of the representatives of the *Ktedonobacteria* class lack the corresponding genes (see Section 3.3). Sequences from *Dsf. hafniense* strains were also not included, since corresponding *van* loci lack *vanH* orthologs [52].

The final tree is shown in Figure 5. An analysis of its topology allowed us to make some important observations. At first glance, a very recent horizontal gene transfer (HGT) event involving *van* genes from *Actinobacteria* seemed evident: VanHA sequences from all *Ktedonobacteria* spp. and from some *Clostridia* spp. (namely, *An. chartisolvens* DSM 27016, *Cli. tagluense* CS002, *Ox. pfennigii* DSM 3222, and *Candidatus* Formimonas warabiya DCMF) rooted deep in the core *Actinobacteria* clade (further referred to as "Core *Actinobacteria*") (Figure 5), being a sister clade to VanHA from *Stackebrandtia nassauensis* DSM 44728. Notably, VanA sequences from *Dsf. hafniense* strains also belonged to an analogous clade of the VanA tree (ESM Figure S2). The repertoire of MGE-related genes co-localized with *van* loci in all of these species was variable, hindering the comprehension of transmission vectors (Figure 5, Supplementary Excel File S1 "Clostridia MGERG" and "Ktedonobacteria MGERG"). However, in *Ktedonobacteria* spp., multiple IS transposase genes were co-localized with *van* loci, as well as with seemingly intact IS elements (Supplementary Excel File S1 "Ktedonobacteria MGERG").

In addition to the "Core *Actinobacteria*", five major clades were defined within the tree. Clade (I) (Figure 5) appeared to be a sister clade of the "Core *Actinobacteria*". It consisted of sequences from 12 strains of *Paenibacillus*, *Tba. xylanilyticus* XE (all 13 belonging to the class *Bacilli*), and *Am. terrae* CBA3637 (class *Clostridia*)—the last one being a basal taxon of clade (I). In all of these species, MGE-related genes co-localized with *van* loci were quite different (Figure 5, Supplementary Excel File S1 "Clostridia MGERG" and "Bacilli MGERG"). Clade (II) was composed of VanHA sequences mainly from *Bacilli* spp. These were six strains of the genus *Paenibacillus*, four strains of *Brevibacillus*, *Thermoactinomyces* sp. CICC 10735, two strains of the genus *Clostridium*, and *Al. lenta* MO-CFX2 (class *Anaerolineae*) (Figure 5). VanHA sequences from *Pba. physcomitrellae* strains formed the basal branch of clade (II). Again, here, we did not observe any reproducible pattern of MGE-related genes co-localized with *van* loci (Figure 5, Supplementary Excel File S1). Notably, it appeared that one of the clade (II) species, namely, *Bba. laterosporus* B9, carries *van* loci on a plasmid. Unfortunately, the quality of genome assemblies of another clade (II) species did not allow us to understand whether they carry *van* loci on plasmids as well.

VanHA sequences from *Clostridia* spp. (with the single exception of *Lct. caecimuris* 3BBH23, class *Erysipelotrichia*) formed clade (III) (Figure 5). The basal taxon of clade (III) was *Lachnotalea* sp. AF33-28. In contrast to clade (II), we observed a certain similarity in the repertoire of MGE-related genes co-localized with *van* loci in strains forming clade (III) (Figure 5). In some of the strains, most notably in *Lachnotalea* sp. AF33-28, [*Clostridium*] *symbiosum* MSK.7.21, *Lachnospiraceae* bacterium M18-1, and *Ext. muris* DSM 28560, large assemblages of putative MGE-related genes were associated with *van* loci (see Section 3.5). These assemblages resembled prophages, although an incomplete prophage region was identified only in *Lachnotalea* sp. AF33-28.

The last two clades—(IV) and (V)—deserve special attention. Clade (IV) was composed of VanHA sequences from [*Clostridium*] *methoxybenzovorans* SR3, [*Clostridium*] *indolis* DSM 755, *Erysipelotrichaceae* bacterium 66202529 (class *Erysipelotrichia*), and a set of strains known to carry *van* loci as a part of Tn*1549*-like transposons, namely, *Atopobium* (*Atpb.*) *minutum* 10063974 (phylum *Actinobacteria*), *Clostridium* sp. MLG245, *Cld. difficile* AI0499, *Ecc. faecium* E525, *Ecc. faecium* RBWH1, *Ecc. faecium* MLG229, *Ecc. faecalis* BM4382, and *Ecc. faecium* E155 (see ESM Table S1 for references) (Figure 5). While *Erysipelotrichaceae* bacterium 66202529 clearly carried genes from the Tn*1549*-like transposon near *van* loci (see Section 3.2), no MGE-related genes were discovered near *van* loci in [*Clostridium*] *methoxybenzovorans* SR3 and [*Clostridium*] *indolis* DSM 755.

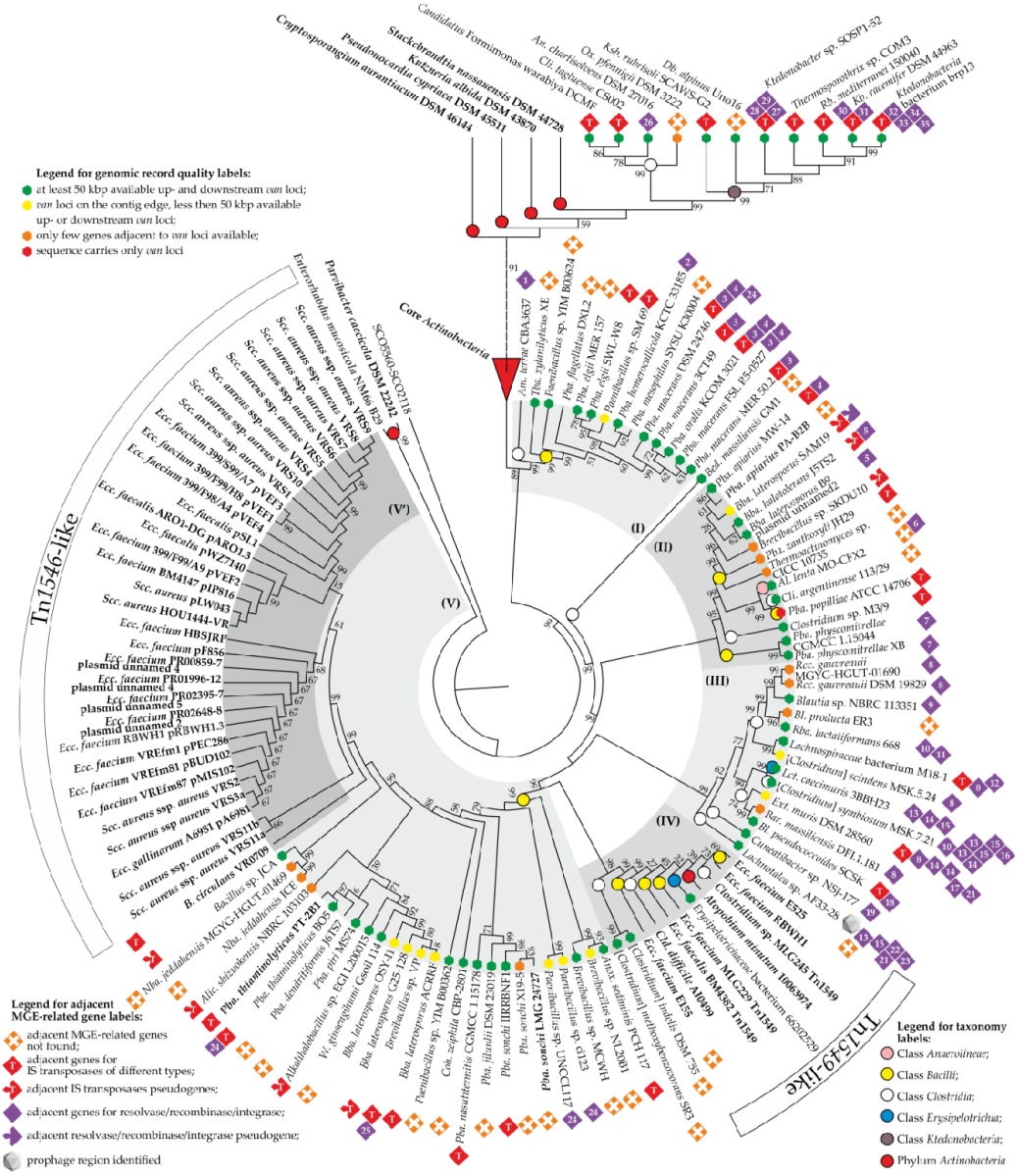


Figure 5. Rooted neighbor-joining phylogenetic tree of 261 VanHA sequences from *Actinobacteria*, *Anaerolineae*, *Bacilli*, *Clostridia*, *Erysipelotrichia*, and *Ktedonobacteria*, together with the sequence of SCO5560-SCO2118 used as an outgroup. The tree was generated as described in Section 2. For its better representability, branch lengths were ignored, while the "Core *Actinobacteria*" clade was collapsed; the same tree completely expanded and drawn to scale is given in ESM Figure S3. Numbers at nodes indicate bootstrap-support percentages derived from 1000 replications (shown only for ≥15% values). Five clades (I–V) discussed in the main text are highlighted in gray, while the labels are interpreted above and below the tree. Taxa known to carry *van* loci from previous works are given

in bold. The following MGE-related enzymes are numbered 1-35 (if a numbered label is found near few taxa, the accession number would be given only for "prototype", implying that its counterparts are close homologs sharing ≥ 50% aa sequence identity): (1) WP_162363028—serine-based recombinase/integrase; (2) WP_139603211—relaxase/mobilisation nuclease; (3) WP_124332115—tyrosine-based recombinase/integrase; (4) WP_036622915—tyrosine-based recombinase/integrase; (5) AKF95793—tyrosine-based recombinase/integrase; (6) WP_119072093—serine-based recombinase/integrase; (7) WP_124332115—serine-based recombinase/integrase; (8) WP_028530137—tyrosine-based recombinase/integrase; (9) WP_225305221—serine-based recombinase/integrase; (10) WP_016316877—relaxase/mobilisation nuclease; (11) WP_016316878—serine-based recombinase/integrase; (12) EOS38773—serine-based recombinase/integrase; (13) WP_213540055 – relaxase/mobilisation nuclease; (14) WP_213540032—tyrosine-based recombinase/integrase; (15) WP_213540058—relaxase/mobilisation nuclease; (16) WP_173892454—serine-based recombinase/integrase; (17) WP_132278971—relaxase/mobilisation nuclease; (18) WP_065543677—tyrosine-type recombinase/integrase; (19) WP_244802434—tyrosine-type recombinase/integrase; (20) WP_002569238—relaxase/mobilisation nuclease; (21) WP_117962491—serine-based recombinase/integrase; (22) WP_117761817—serine-based recombinase/integrase; (23) WP_117962497—relaxase/mobilisation nuclease; (24) WP_072327346—tyrosine-based recombinase/integrase; (25) WP_161555111—tyrosine-based recombinase/integrase; (26) WP_170138094—serine-based recombinase/integrase; (27) GHO64028—tyrosine-based recombinase/integrase; (28) GHO64041—phage integrase; (29) GHO64148—serine-based recombinase/integrase; (30) WP_007903552—tyrosine-based recombinase/integrase; (31) WP_075164122—tyrosine-based recombinase/integrase; (32) BCL79139—tyrosine-based recombinase/integrase; (33) BCL79135—serine-based recombinase/integrase; (34) BCL79150—tyrosine-based recombinase/integrase; (35) BCL79151—tyrosine-based recombinase/integrase.

Clade (V) included a large portion of VanHA sequences from different *Bacilli* spp.; the repertoire of MGE-related genes co-localized with *van* loci was rather poor in species belonging to this clade (Figure 5). Finally, a subclade of clade (V)—(V')—appeared to be composed of VanHA sequences from plasmids carrying Tn*1546*-like transposons, isolated from clinically relevant enterococci and staphylococci from all over the world (Figure 5, see also ESM Table S1 for references). *Bacillus circulans* VR0709 was a basal taxon of (V') (Figure 5), where a single MGE-related gene was found adjacent to the *van* loci [50,79]. This was a gene for an IS200/IS605-family (Y1) transposase (CAB61226) that had no homologs in other clade (V') strains. Peculiarly, *Ecc. faecium* RBWH1 pRBWH1.3^{+} [32] appeared to carry two sets of *van* genes simultaneously: one as a part of the Tn*1546*-like transposon on the plasmid (pRBWH1.3), and another as a part of the Tn*1549*-like transposon integrated into the chromosome.

Concluding the analysis of the tree, it seems interesting to note that VanHA sequences encoded in putative MGEs from two actinobacteria—*Parvibacter* (*Prv.*) *caecicola* DSM 22242 and *Enterorhabdus* (*Enr.*) *mucosicola* NM66_B29—formed the basal branch of the whole tree (Figure 5).

4. Discussion

The results obtained in the current work indicate that *van* genes are more widely distributed among bacteria than expected before: for the first time, *van* genes were discovered in the genomic records of organisms from still poorly studied classes such as *Anaerolineae*, *Erysipelotrichia*, and *Ktedonobacteria*. Most likely, the spread of *van* genes among these different bacteria was determined by multiple HGT events. This idea is not new [26,52,80,81], but the abundance of genomic data available now allows us to outline some of these HGT events more precisely. Although the overall picture is quite puzzling and we are only at the beginning of such types of investigations, in the discussion, we propose a speculative scenario of how *van* genes might have disseminated (Figure 6).

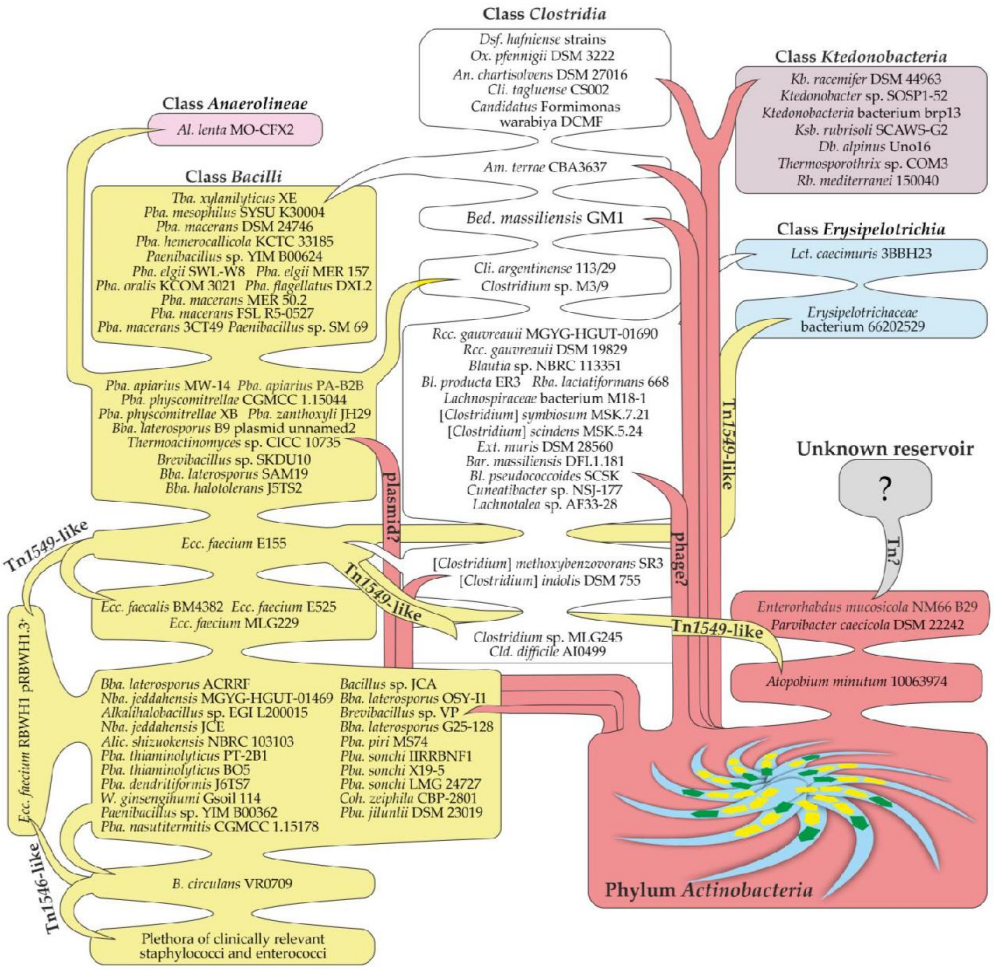


Figure 6. A scheme depicting a possible scenario for the transmission of *vanHAX* genes from *Actinobacteria* spp. to representatives of other classes. The scheme is discussed in the main text.

As noted in the previous section, the most recent HGT event may have led to the transmission of *van* genes from actinobacteria to the *Ktedonobacteria* class and to some members of the *Clostridia* class (see scheme in Figure 6). The corresponding proteins did not diverge much from the actinobacterial ones and were found deep in the "core *Actinobacteria*" clade (Figure 5). However, the G-C content of these *van* loci (<50%) is definitely lower than that of actinobacteria (>60%), indicating that the corresponding genes were eventually adapted to the translation machinery of the novel hosts. Most of these *Ktedonobacteria* and *Clostridia* spp. were isolated from soil biotopes [55], likely sharing their ecological niches with *Actinobacteria* spp.—the source of *van* loci. Notable exceptions were *Ox. pfennigii* DSM 3222, isolated from cattle rumen [82] (a biotope also inhabited by different actinobacteria [83]), and *Cli. tagluense* CS002—a causative agent of meat spoilage [84]. Although no MGE-related genes were discovered near *van* loci in *Ox. pfennigii* DSM 3222 and only a single ISNCY-family transposase was encoded near *van* loci in *Cli. tagluense* CS002, these examples chart new possible routes for the further spread of *van* genes in antropocenoses. The vehicle/vehicles of such putative HGT events remain

elusive, although multiple genes for transposases and resolvases/recombinases/integrases were found to co-localize with *van* loci in these *Ktedonobacteria* and *Clostridia* spp. It seems that in all of these cases, *van* loci were located on the chromosome; consequently, it is also possible that MGE-related genes might have been partially or completely lost upon integration, eroding the genomic landscapes of putative vehicles carrying *van* genes.

IS-like elements might have played some role in the HGT of *van* loci into *Ktedonobacteria* spp., since, in these microbes, an immense number of transposase genes and intact IS elements are found nearby. Notably, *van* loci in *Ktedonobacteria* spp. underwent peculiar modifications: genes coding for VanX were lost, while novel genes for other types of peptidases were gained. Some evidence exists that at least some of these new genes were introduced as "passenger genes" in IS-like elements (see Section 3.3). As already mentioned above, *Ktedonobacteria* spp. have an actinomycete-like lifestyle [55]. Accordingly, these microbes might have recruited *van* loci for protection against dalbaheptide producers (inhabiting the same biotopes), self-immunity against certain endogenous secondary metabolites, or morphogenesis-related cell-wall remodeling. Overall, the case of *van* genes in *Ktedonobacteria* spp. is of utmost interest and requires further experimental investigation.

Aside from VanHA sequences rooted deep in "Core *Actinobacteria*", our reference tree demonstrated the presence of five well-defined clades (99% bootstrap support) and a monotaxon branch (*Bed. massiliensis* GM1). For further interpretation, we would like to make two assumptions: (a) each of these clades (and *Bed. massiliensis* GM1 branch) emerged as a result of a single HGT event from *Actinobacteria*; (b) the basal taxa of each clade represent a starting point in which *van* loci were delivered by an initial HGT event (before further spreading to the other taxa of the clade). Following these assumptions, an interesting speculative yet plausible scenario for *van* gene dissemination can be proposed (Figure 6).

In clade (I), a HGT event likely introduced *van* loci to some *Clostridia* spp. since the basal taxon of this clade is *Am. terrae* CBA3637, belonging to the class *Clostridia*. *Am. terrae* CBA3637 was isolated from river sediments [85], which are also inhabited by a plethora of actinobacteria [86]. Further, *van* genes spread to other clade (I) taxa—*Tba. xylanilyticus* XE and *Paenibacillus* spp. (Figure 6).

Clade (II) points to several HGT events in addition to the putative initial one. Here, the majority of taxa belong to the *Bacilli* class. However, sequences from an *Anaerolineae* class bacterium—*Al. lenta* MO-CFX2—are positioned far from the base of clade (II), suggesting a putative HGT event delivering *van* genes from clade (II) *Bacilli* spp. to *Al. lenta* MO-CFX2. Notably, *Al. lenta* MO-CFX2 is the first reported Gram-negative bacterium [87] carrying *van* loci. The same scenario also works for the two *Clostridia* spp. found in clade (II)—*Clostridium* sp. M3/9 and *Cli. argentinense* 113/29. In at least one clade (II) taxon, *Bba. laterosporus* B9, *van* loci are located on a plasmid, implying that a possible vehicle bringing *van* genes to clade (II) might have been a plasmid.

Except for one, all taxa forming clade (III) belong to *Clostridia* spp. An exception is the VanHA sequences of *Lct. caecimuris* 3BBH23 from the *Erysipelotrichia* class, rooted deep in clade (III). Such a topology most likely indicates that *Lct. caecimuris* 3BBH23 received *van* loci from some clade (III) *Clostridia* spp. as a result of an additional HGT event (Figure 6). It is also important to note that *van* loci in the basal taxon of the clade (III)—*Lachnotalea* sp. AF33-28—are located downstream of an incomplete prophage region. Thus, the vehicles that delivered *van* genes to clade (III) organisms might have been some phages.

VanHA sequences from the classes *Bacilli*, *Erysipelotrichia*, *Clostridia* and the phylum *Actinobacteria* composed clade (IV). The basal branch of clade (IV) was represented by sequences from [*Clostridium*] *indolis* DSM 755 and [*Clostridium*] *methoxybenzovorans* SR3, isolated from soil [88] and wastewater [89], respectively. No MGE-related genes accompanied *van* loci in these two organisms. However, in all of the other taxa of clade (III), *van* genes were found as a part of Tn*1549*-like transposons. Such evidence induces us to speculate that an unknown vehicle delivered *van* genes to [*Clostridium*] *indolis* DSM 755 and [*Clostridium*] *methoxybenzovorans* SR3 (or ancestral organisms) in an initial HGT

event, leading to the emergence of clade (III) (Figure 6). In the following HGT events, *van* loci were mobilized by a Tn*1549*-like transposon and then spread to other clade (III) taxa. These taxa included several clinical and human isolates: an organism from the *Erysipelotrichia* class (*Erysipelotrichaceae* bacterium 66202529), *Cld. difficile* AI0499, *Clostridium sp.* MLG245, several strains of enterococci, and the actinobacterium *Atpb. minutum* 10063974 (Figure 6). Overall, we might conclude that two bacteria from the *Erysipelotrichia* class obtained *van* genes in two independent HGT events, while the Tn*1549*-like transposon mediated the reverse transition of *van* genes to the *Actinobacteria* phylum in the case of *Atpb. minutum* 10063974.

Unlike all other clades, clade (V) was composed exclusively of *Bacilli* spp. The basal taxa of clade (V) were mostly soil dwellers and did not show any clear repetitive patterns of MGE-related genes co-localized with *van* loci (Figure 5). Herein, subclade (V′) was the most interesting, being composed principally of VanHA sequences encoded on plasmids carrying *Tn1546*-like transposons from different clinical isolates of staphylococci and enterococci. At the basal branch of subclade (V′), there are VanHA sequences from a clinical isolate of the soil-dwelling opportunistic human pathogen *B. circulans* VR0709 [50]. Overall, the topology of clade (V) suggests that an unknown vehicle delivered *van* loci to *B. circulans* VR0709 from a gene pool existing in basal clade (V) taxa. Consequentially, a Tn*1546*-containing plasmid mobilized *van* loci from *B. circulans* VR0709 and mediated further dissemination in nosocomial environments among staphylococci and enterococci (Figure 6). This last observation actually corroborates a previous conclusion made from a phylogenetic analysis conducted more than ten years ago on a less extensive sample of available *van* gene sequences [80].

Finally, we would also like to comment on *van* genes from the two actinobacterial strains—*Prv. caecicola* DSM 22242 and *Enr. mucosicola* NM66_B29. VanHA sequences from these two organisms were the most divergent, being an outgroup for our reference tree (Figure 5). In our previous work, we demonstrated that *van* loci in *Prv. caecicola* DSM 22242 and *Enr. mucosicola* NM66_B29 are co-localized with an assemblage of MGE-related genes resembling a transposon of unknown type [25]. Notably, both strains were isolated from the intestines of laboratory mice [90,91]. Keeping in mind the case of *Atpb. minutum* 10063974 (see above), it is tempting to suppose that *Prv. caecicola* DSM 22242 and *Enr. mucosicola* NM66_B29 represent another example of a reverse HGT event that brought *van* genes back to the *Actinobacteria* phylum. However, the source of the putative transposons found in *Prv. caecicola* DSM 22242 and *Enr. mucosicola* NM66_B29 remains unknown (Figure 6).

Our results demonstrate that *van* genes from the *Actinobacteria* phylum represent an interesting and widespread part of the *Terrabacteria* mobilome. The flow of *van* genes is a continuous process that occurs actively, as demonstrated by the example of *Ktedonobacteria* spp. Likely carried by different vehicles, *van* genes spread from one taxon to another, creating a puzzling network of consequential HGT events (Figure 6). Bacteria from the *Bacilli* and *Clostridia* classes play an important role in the spread of *van* genes and might represent their source for classes such as *Anaerolineae* and *Erysipelotrichia*. We are aware that our scenario depicting the spread of *vanHAX* genes is speculative and that it will probably change when more genomic information becomes available. However, we believe that at the current point, it provides a useful interpretation of the overall picture to start from in future investigations.

5. Conclusions

Through extensive bioinformatic work, we discovered that *van* genes are present in three classes of eubacteria—*Anaerolineae*, *Erysipelotrichia*, and *Ktedonobacteria*—as well as in new genera belonging to the *Clostridia* and *Bacilli* classes. The majority of these *van* genes tended to be co-localized with MGE-related genes of various types. Based on phylogenetic reconstruction, we could assume that the class *Anaerolineae* recently obtained *van* genes from *Bacilli* spp., while two recent independent HGT events delivered *van* genes to *Erysipelotrichia* from *Clostridia* and *Bacilli* spp., respectively. In addition, the most recent HGT event spread

van genes from *Actinobacteria* directly to species belonging to the *Ktedonobacteria* class. *Ktedonobacteria* tended to lose *vanX* and gain novel genes coding for carboxypeptidases. Finally, a Tn*1549*-like transposon likely mobilized *van* genes from *Clostridia* spp., while a Tn*1546*-like transposon mobilized *van* genes from *Bacilli* spp.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/genes13111960/s1.

Author Contributions: Conceptualization, O.Y.; Investigation, O.Y. and E.B.; Supervision, F.M. and V.F.; Writing—Original Draft, O.Y.; Writing—Review and Editing, F.M., E.B. and V.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the public grant "Fondo di Ateneo per la Ricerca" 2020 and 2021 to F.M. and by the BG-09F grant of the Ministry of Education and Science of Ukraine to V.F.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data obtained in the study are presented in either the main text or Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nicolaou, K.C.; Boddy, C.N.C.; Bräse, S.; Winssinger, N. Chemistry, biology, and medicine of the glycopeptide antibiotics. *Angew. Chem. Int. Ed.* **1999**, *38*, 2096–2152. [CrossRef]
2. Parenti, F.; Cavalleri, B. Proposal to name the vancomycin-ristocetin like glycopeptides as dalbaheptides. *J. Antibiot.* **1989**, *42*, 1882–1883. [CrossRef] [PubMed]
3. Kaniusaite, M.; Tailhades, J.; Kittilä, T.; Fage, C.D.; Goode, R.J.A.; Schittenhelm, R.B.; Cryle, M.J. Understanding the early stages of peptide formation during the biosynthesis of teicoplanin and related glycopeptide antibiotics. *FEBS J.* **2021**, *288*, 507–529. [CrossRef]
4. Stegmann, E.; Frasch, H.J.; Wohlleben, W. Glycopeptide biosynthesis in the context of basic cellular functions. *Curr. Opin. Microbiol.* **2010**, *13*, 595–602. [CrossRef] [PubMed]
5. Thaker, M.N.; Wright, G.D. Opportunities for synthetic biology in antibiotics: Expanding glycopeptide chemical diversity. *ACS Synth. Biol.* **2015**, *4*, 195–206. [CrossRef]
6. Yushchuk, O.; Ostash, B. Glycopeptide antibiotics: Genetics, chemistry, and new screening approaches. In *Natural Products from Actinomycetes*; Springer: Singapore, 2022; pp. 411–444. [CrossRef]
7. Yushchuk, O.; Zhukrovska, K.; Berini, F.; Fedorenko, V.; Marinelli, F. Genetics behind the glycosylation patterns in the biosynthesis of dalbaheptides. *Front. Chem.* **2022**, *10*, 858708. [CrossRef]
8. Marcone, G.L.; Binda, E.; Berini, F.; Marinelli, F. Old and new glycopeptide antibiotics: From product to gene and back in the post-genomic era. *Biotechnol. Adv.* **2018**, *36*, 534–554. [CrossRef]
9. Marschall, E.; Cryle, M.J.; Tailhades, J. Biological, chemical, and biochemical strategies for modifying glycopeptide antibiotics. *J. Biol. Chem.* **2019**, *294*, 18769–18783. [CrossRef]
10. Müller, A.; Klöckner, A.; Schneider, T. Targeting a cell wall biosynthesis hot spot. *Nat. Prod. Rep.* **2017**, *34*, 909–932. [CrossRef]
11. Jovetic, S.; Zhu, Y.; Marcone, G.L.; Marinelli, F.; Tramper, J. β-Lactam and glycopeptide antibiotics: First and last line of defense? *Trends Biotechnol.* **2010**, *28*, 596–604. [CrossRef]
12. Zhanel, G.G.; Calic, D.; Schweizer, F.; Zelenitsky, S.; Adam, H.; Lagacé-Wiens, P.R.S.; Rubinstein, E.; Gin, A.S.; Hoban, D.J.; Karlowsky, J.A. New lipoglycopeptides: A comparative review of dalbavancin, oritavancin and telavancin. *Drugs* **2010**, *70*, 859–886. [CrossRef] [PubMed]
13. Binda, E.; Marinelli, F.; Marcone, G.L. Old and new glycopeptide antibiotics: Action and resistance. *Antibiotics* **2014**, *3*, 572–594. [CrossRef] [PubMed]
14. Sarkar, P.; Yarlagadda, V.; Ghosh, C.; Haldar, J. A review on cell wall synthesis inhibitors with an emphasis on glycopeptide antibiotics. *Medchemcomm* **2017**, *8*, 516–533. [CrossRef]
15. Vollmer, W.; Blanot, D.; De Pedro, M.A. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **2008**, *32*, 149–167. [CrossRef] [PubMed]
16. Nitanai, Y.; Kikuchi, T.; Kakoi, K.; Hanamaki, S.; Fujisawa, I.; Aoki, K. Crystal structures of the complexes between vancomycin and cell-wall precursor analogs. *J. Mol. Biol.* **2009**, *385*, 1422–1432. [CrossRef]
17. Williams, D.H. The glycopeptide story—How to kill the deadly "superbugs." *Nat. Prod. Rep.* **1996**, *13*, 469–477. [CrossRef]
18. Binda, E.; Marcone, G.L.; Pollegioni, L.; Marinelli, F. Characterization of VanYn, a novel D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance in *Nonomuraea* sp. ATCC 39727. *FEBS J.* **2012**, *279*, 3203–3213. [CrossRef]

19. Marcone, G.L.; Beltrametti, F.; Binda, E.; Carrano, L.; Foulston, L.; Hesketh, A.; Bibb, M.; Marinelli, F. Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727. *Antimicrob. Agents Chemother.* **2010**, *54*, 2465–2472. [CrossRef]
20. Yushchuk, O.; Homoniuk, V.; Ostash, B.; Marinelli, F.; Fedorenko, V. Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. *J. Antibiot.* **2020**, *73*, 255–259. [CrossRef]
21. Kilian, R.; Frasch, H.J.; Kulik, A.; Wohlleben, W.; Stegmann, E. The VanRS homologous two-component system $VnlRS_{Ab}$ of the glycopeptide producer *Amycolatopsis balhimycina* activates transcription of the $vanHAX_{Sc}$ genes in *Streptomyces coelicolor*, but not in *A. balhimycina*. *Microb. Drug Resist.* **2016**, *22*, 499–509. [CrossRef]
22. Yushchuk, O.; Binda, E.; Marinelli, F. Glycopeptide antibiotic resistance genes: Distribution and function in the producer actinomycetes. *Front. Microbiol.* **2020**, *11*. [CrossRef] [PubMed]
23. Stegmann, E.; Frasch, H.J.; Kilian, R.; Pozzi, R. Self-resistance mechanisms of actinomycetes producing lipid II-targeting antibiotics. *Int. J. Med. Microbiol.* **2015**, *305*, 190–195. [CrossRef] [PubMed]
24. Hong, H.J.; Hutchings, M.I.; Neu, J.M.; Wright, G.D.; Paget, M.S.B.; Buttner, M.J. Characterization of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (*vanK*) required for drug resistance. *Mol. Microbiol.* **2004**, *52*, 1107–1121. [CrossRef] [PubMed]
25. Andreo-Vidal, A.; Binda, E.; Fedorenko, V.; Marinelli, F.; Yushchuk, O. Genomic insights into the distribution and phylogeny of glycopeptide resistance determinants within the actinobacteria phylum. *Antibiotics* **2021**, *10*, 1533. [CrossRef] [PubMed]
26. Marshall, C.G.; Broadhead, G.; Leskiw, B.K.; Wright, G.D. D-Ala-D-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6480–6483. [CrossRef]
27. Rushton-Green, R.; Darnell, R.L.; Taiaroa, G.; Carter, G.P.; Cook, G.M.; Morgan, X.C. Agricultural origins of a highly persistent lineage of vancomycin-resistant *Enterococcus faecalis* in New Zealand. *Appl. Environ. Microbiol.* **2019**, *85*, e00137-19. [CrossRef]
28. Lim, S.K.; Tanimoto, K.; Tomita, H.; Ike, Y. Pheromone-responsive conjugative vancomycin resistance plasmids in *Enterococcus faecalis* isolates from humans and chicken feces. *Appl. Environ. Microbiol.* **2006**, *72*, 6544–6553. [CrossRef]
29. Mello, S.S.; Van Tyne, D.; Lebreton, F.; Silva, S.Q.; Nogueira, M.C.L.; Gilmore, M.S.; Camargo, I.L.B.C. A mutation in the glycosyltransferase gene *lafB* causes daptomycin hypersusceptibility in *Enterococcus faecium*. *J. Antimicrob. Chemother.* **2020**, *75*, 36–45. [CrossRef]
30. Ballard, S.A.; Pertile, K.K.; Lim, M.; Johnson, P.D.R.; Grayson, M.L. Molecular characterization of *vanB* elements in naturally occurring gut anaerobes. *Antimicrob. Agents Chemother.* **2005**, *49*, 1688–1694. [CrossRef]
31. Kinnear, C.L.; Hansen, E.; Morley, V.J.; Tracy, K.C.; Forstchen, M.; Read, A.F.; Woods, R.J. Daptomycin treatment impacts resistance in off-target populations of vancomycin-resistant *Enterococcus faecium*. *PLoS Biol.* **2020**, *18*, e3000987. [CrossRef]
32. Bohlmann, L.; De Oliveira, D.M.P.; El-Deeb, I.M.; Brazel, E.B.; Harbison-Price, N.; Ong, C.L.Y.; Rivera-Hernandez, T.; Ferguson, S.A.; Cork, A.J.; Phan, M.D.; et al. Chemical synergy between ionophore PBT2 and zinc reverses antibiotic resistance. *MBio* **2018**, *9*, e02391-18. [CrossRef] [PubMed]
33. Khan, S.; Sung, K.; Marasa, B.; Min, S.; Kweon, O.; Nawaz, M.; Cerniglia, C. Draft genome sequence of multidrug-resistant *Enterococcus faecium* clinical isolate VRE3, with a sequence type 16 pattern and novel structural arrangement of Tn*1546*. *Genome Announc.* **2015**, *3*, 1000000. [CrossRef] [PubMed]
34. Melegh, S.; Nyul, A.; Kovács, K.; Kovács, T.; Ghidán, Á.; Dombrádi, Z.; Szabó, J.; Berta, B.; Lesinszki, V.; Pászti, J.; et al. Dissemination of VanA-type *Enterococcus faecium* isolates in Hungary. *Microb. Drug Resist.* **2018**, *24*, 1376–1390. [CrossRef] [PubMed]
35. Eshaghi, A.; Shahinas, D.; Li, A.; Kariyawasam, R.; Banh, P.; Desjardins, M.; Melano, R.G.; Patel, S.N. Characterization of an *Enterococcus gallinarum* isolate carrying a dual *vanA* and *vanB* cassette. *J. Clin. Microbiol.* **2015**, *53*, 2225–2229. [CrossRef]
36. Zaheer, R.; Cook, S.R.; Barbieri, R.; Goji, N.; Cameron, A.; Petkau, A.; Polo, R.O.; Tymensen, L.; Stamm, C.; Song, J.; et al. Surveillance of *Enterococcus* spp. reveals distinct species and antimicrobial resistance diversity across a One-Health continuum. *Sci. Rep.* **2020**, *10*, 3937. [CrossRef]
37. Weigel, L.M.; Clewell, D.B.; Gill, S.R.; Clark, N.C.; McDougal, L.K.; Flannagan, S.E.; Kolonay, J.F.; Shetty, J.; Killgore, G.E.; Tenover, F.C. Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. *Science* **2003**, *302*, 1569–1571. [CrossRef] [PubMed]
38. Panesso, D.; Planet, P.J.; Diaz, L.; Hugonnet, J.E.; Tran, T.T.; Narechania, A.; Munita, J.M.; Rincon, S.; Carvajal, L.P.; Reyes, J.; et al. Methicillin-susceptible, vancomycin-resistant *Staphylococcus aureus*, Brazil. *Emerg. Infect. Dis.* **2015**, *21*, 1844–1848. [CrossRef]
39. Zhu, W.; Murray, P.R.; Huskins, W.C.; Jernigan, J.A.; McDonald, L.C.; Clark, N.C.; Anderson, K.F.; McDougal, L.K.; Hageman, J.C.; Olsen-Rasmussen, M.; et al. Dissemination of an *Enterococcus Inc18*-like *vanA* plasmid associated with vancomycin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **2010**, *54*, 4314–4320. [CrossRef]
40. Kos, V.N.; Desjardins, C.A.; Griggs, A.; Cerqueira, G.; Van Tonder, A.; Holden, M.T.G.; Godfrey, P.; Palmer, K.L.; Bodi, K.; Mongodin, E.F.; et al. Comparative genomics of vancomycin-resistant *Staphylococcus aureus* strains and their positions within the clade most commonly associated with methicillin-resistant *S. aureus* hospital-acquired infection in the United States. *MBio* **2012**, *3*, e00112-12. [CrossRef]
41. Garnier, F.; Taourit, S.; Glaser, P.; Courvalin, P.; Galimand, M. Characterization of transposon Tn*1549*, conferring VanB-type resistance in *Enterococcus* spp. *Microbiology* **2000**, *146*, 1481–1489. [CrossRef]

42. De Been, M.; Van Schaik, W.; Cheng, L.; Corander, J.; Willems, R.J. Recent recombination events in the core genome are associated with adaptive evolution in *Enterococcus faecium*. *Genome Biol. Evol.* **2013**, *5*, 1524–1535. [CrossRef] [PubMed]
43. Sletvold, H.; Johnsen, P.J.; Wikmark, O.G.; Simonsen, G.S.; Sundsfjord, A.; Nielsen, K.M. Tn*1546* is part of a larger plasmid-encoded genetic unit horizontally disseminated among clonal *Enterococcus faecium* lineages. *J. Antimicrob. Chemother.* **2010**, *65*, 1894–1906. [CrossRef] [PubMed]
44. Halvorsen, E.M.; Williams, J.J.; Bhimani, A.J.; Billings, E.A.; Hergenrother, P.J. Txe, an endoribonuclease of the enterococcal Axe-Txe toxin-antitoxin system, cleaves mRNA and inhibits protein synthesis. *Microbiology* **2011**, *157*, 387–397. [CrossRef] [PubMed]
45. Szakacs, T.A.; Kalan, L.; McConnell, M.J.; Eshaghi, A.; Shahinas, D.; McGeer, A.; Wright, G.D.; Low, D.E.; Patel, S.N. Outbreak of vancomycin-susceptible *Enterococcus faecium* containing the wild-type *vanA* gene. *J. Clin. Microbiol.* **2014**, *52*, 1682–1686. [CrossRef] [PubMed]
46. Sletvold, H.; Johnsen, P.J.; Simonsen, G.S.; Aasnæs, B.; Sundsfjord, A.; Nielsen, K.M. Comparative DNA analysis of two *vanA* plasmids from *Enterococcus faecium* strains isolated from poultry and a poultry farmer in Norway. *Antimicrob. Agents Chemother.* **2007**, *51*, 736–739. [CrossRef]
47. Sletvold, H.; Johnsen, P.J.; Hamre, I.; Simonsen, G.S.; Sundsfjord, A.; Nielsen, K.M. Complete sequence of *Enterococcus faecium* pVEF3 and the detection of an ω-ε-ζ toxin-antitoxin module and an ABC transporter. *Plasmid* **2008**, *60*, 75–85. [CrossRef]
48. Guardabassi, L.; Perichon, B.; Van Heijenoort, J.; Blanot, D.; Courvalin, P. Glycopeptide resistance *vanA* operons in *Paenibacillus* strains isolated from soil. *Antimicrob. Agents Chemother.* **2005**, *49*, 4227–4233. [CrossRef]
49. Fraimow, H.; Knob, C.; Herrero, I.A.; Patel, R. Putative VanRS-like two-component regulatory system associated with the inducible glycopeptide resistance cluster of *Paenibacillus popilliae*. *Antimicrob. Agents Chemother.* **2005**, *49*, 2625–2633. [CrossRef]
50. Ligozzi, M.; Lo Cascio, G.; Fontana, R. *vanA* gene cluster in a vancomycin-resistant clinical isolate of *Bacillus circulans*. *Antimicrob. Agents Chemother.* **1998**, *42*, 2055–2059. [CrossRef]
51. Knight, D.R.; Androga, G.O.; Ballard, S.A.; Howden, B.P.; Riley, T.V. A phenotypically silent *vanB2* operon carried on a Tn*1549*-like element in *Clostridium difficile*. *Msphere* **2016**, *1*, e00177-16. [CrossRef]
52. Kruse, T.; Levisson, M.; de Vos, W.M.; Smidt, H. *vanI*: A novel D-Ala-D-Lac vancomycin resistance gene cluster found in *Desulfitobacterium hafniense*. *Microb. Biotechnol.* **2014**, *7*, 456–466. [CrossRef] [PubMed]
53. Yamada, T.; Sekiguchi, Y. Cultivation of uncultured *Chloroflexi subphyla: Significance* and ecophysiology of formerly uncultured *Chloroflexi "subphylum I"* with natural and biotechnological relevance. *Microbes Environ.* **2009**, *24*, 205–216. [CrossRef] [PubMed]
54. Davis, J.J.; Xia, F.; Overbeek, R.A.; Olsen, G.J. Genomes of the class *Erysipelotrichia clarify* the firmicute origin of the class Mollicutes. *Int. J. Syst. Evol. Microbiol.* **2013**, *63*, 2727–2741. [CrossRef]
55. Yabe, S.; Sakai, Y.; Abe, K.; Yokota, A. Diversity of *Ktedonobacteria* with actinomycetes-like morphology in terrestrial environments. *Microbes Environ.* **2017**, *32*, 61–70. [CrossRef]
56. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
57. Schoch, C.L.; Ciufo, S.; Domrachev, M.; Hotton, C.L.; Kannan, S.; Khovanskaya, R.; Leipe, D.; McVeigh, R.; O'Neill, K.; Robbertse, B.; et al. NCBI Taxonomy: A comprehensive update on curation, resources and tools. *Database* **2020**, *2020*, baaa062. [CrossRef]
58. Clark, K.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2016**, *44*, D67–D72. [CrossRef]
59. O'Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufo, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [CrossRef]
60. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* **2012**, *28*, 1647–1649. [CrossRef]
61. Marchler-Bauer, A.; Bryant, S.H. CD-Search: Protein domain annotations on the fly. *Nucleic Acids Res.* **2004**, *32*, 327–331. [CrossRef]
62. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268. [CrossRef]
63. Siguier, P.; Perochon, J.; Lestrade, L.; Mahillon, J.; Chandler, M. ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Res.* **2006**, *34*, 32–36. [CrossRef]
64. Rice, P.; Longden, L.; Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **2000**, *16*, 276–277. [CrossRef]
65. Arndt, D.; Grant, J.R.; Marcu, A.; Sajed, T.; Pon, A.; Liang, Y.; Wishart, D.S. PHASTER: A better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* **2016**, *44*, W16–W21. [CrossRef]
66. Zhou, F.; Gan, R.; Zhang, F.; Ren, C.; Yu, L.; Si, Y.; Huang, Z. PHISDetector: A tool to detect diverse in silico phage–host interaction signals for virome studies. *Genom. Proteom. Bioinform.* **2022**, *in press*. [CrossRef]
67. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. [CrossRef]
68. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [CrossRef]

69. Sitnikova, T.; Rzhetsky, A.; Nei, M. Interior-branch and bootstrap tests of phylogenetic trees. *Mol. Biol. Evol.* **1995**, *12*, 319–333. [CrossRef]
70. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics.* **1992**, *8*, 275–282. [CrossRef]
71. Battistuzzi, F.U.; Feijao, A.; Hedges, S.B. A genomic timescale of prokaryote evolution: Insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* **2004**, *4*, 44. [CrossRef]
72. Schleifer, K.H. Phylum XIII. *Firmicutes* Gibbons and Murray 1978, 5 (*Firmacutes* [sic] Gibbons and Murray 1978, 5). In *Bergey's Manual® of Systematic Bacteriology*; Springer: Berlin/Heidelberg, Germany, 2009; Volume 3, p. 1450.
73. Ogata, Y.; Sakamoto, M.; Ohkuma, M.; Hattori, M.; Suda, W. Complete genome sequence of *Longicatena caecimuris* strain 3BBH23, isolated from healthy Japanese feces. *Microbiol. Resour. Announc.* **2021**, *10*, e00282-21. [CrossRef]
74. Reynolds, P.E.; Ambur, O.H.; Casadewall, B.; Courvalin, P. The VanYD D,D-carboxypeptidase of *Enterococcus faecium* BM4339 is a penicillin-binding protein. *Microbiology* **2001**, *147*, 2571–2578. [CrossRef]
75. Cavaletti, L.; Monciardini, P.; Bamonte, R.; Schumann, P.; Ronde, M.; Sosio, M.; Donadio, S. New lineage of filamentous, spore-forming, Gram-positive bacteria from soil. *Appl. Environ. Microbiol.* **2006**, *72*, 4360–4369. [CrossRef]
76. Rawlings, N.D.; Waller, M.; Barrett, A.J.; Bateman, A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **2014**, *42*, D503–D509. [CrossRef]
77. Siguier, P.; Gourbeyre, E.; Varani, A.; Ton-Hoang, B.; Chandler, M. Everyman's guide to bacterial insertion sequences. *Microbiol. Spectr.* **2015**, *3*, 1–35. [CrossRef]
78. Maniloff, J.; Ackermann, H.W. Taxonomy of bacterial viruses: Establishment of tailed virus genera and the order *Caudovirales*. *Arch. Virol.* **1998**, *143*, 2051–2063. [CrossRef]
79. Fontana, R.; Ligozzi, M.; Pedrotti, C.; Padovani, E.M.; Cornaglia, G. Vancomycin-resistant *Bacillus circulans* carrying the *vanA* gene responsible for vancomycin resistance in enterococci. *Eur. J. Clin. Microbiol. Infect. Dis.* **1997**, *16*, 473–474. [CrossRef]
80. Aminov, R.I.; Mackie, R.I. Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol. Lett.* **2007**, *271*, 147–161. [CrossRef]
81. Heaton, M.P.; Discotto, L.F.; Pucci, M.J.; Handwerger, S. Mobilization of vancomycin resistance by transposon-mediated fusion of a VanA plasmid with an *Enterococcus faecium* sex pheromone-response plasmid. *Gene* **1996**, *171*, 9–17. [CrossRef]
82. Krumholz, L.R.; Bryant, M.P. *Clostridium pfennigii* sp. nov. uses methoxyl groups of monobenzenoids and produces butyrate. *Int. J. Syst. Bacteriol.* **1985**, *35*, 454–456. [CrossRef]
83. Šuľák, M.; Sikorová, L.; Jankuvová, J.; Javorský, P.; Pristaš, P. Variability of *Actinobacteria*, a minor component of rumen microflora. *Folia Microbiol.* **2012**, *57*, 351–353. [CrossRef]
84. Wambui, J.; Stevens, M.J.A.; Cernela, N.; Stephan, R. Unraveling the genotypic and phenotypic diversity of the psychrophilic *Clostridium estertheticum* complex, a meat spoilage agent. *Front. Microbiol.* **2022**, *13*, 856810. [CrossRef]
85. Kim, Y.B.; Kim, J.Y.; Kim, J.; Song, H.S.; Whon, T.W.; Lee, S.H.; Yoo, S.R.; Myoung, J.; Son, H.S.; Roh, S.W. *Aminipila terrae* sp. nov., a strictly anaerobic bacterium isolated from river sediment. *Arch. Microbiol.* **2021**, *203*, 3163–3169. [CrossRef]
86. Passari, A.K.; Leo, V.V.; Chandra, P.; Kumar, B.; Nayak, C.; Hashem, A.; Abd Allah, E.F.; Alqarawi, A.A.; Singh, B.P. Bioprospection of actinobacteria derived from freshwater sediments for their potential to produce antimicrobial compounds. *Microb. Cell Fact.* **2018**, *17*, 68. [CrossRef]
87. Nakahara, N.; Nobu, M.K.; Takaki, Y.; Miyazaki, M.; Tasumi, E.; Sakai, S.; Ogawara, M.; Yoshida, N.; Tamaki, H.; Yamanaka, Y.; et al. *Aggregatilinea lenta* gen. nov., sp. nov., a slow-growing, facultatively anaerobic bacterium isolated from subseafloor sediment, and proposal of the new order *Aggregatilineales* ord. nov. within the class *Anaerolineae* of the phylum *Chloroflexi*. *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, 1185–1194. [CrossRef]
88. Biddle, A.S.; Leschine, S.; Huntemann, M.; Han, J.; Chen, A.; Kyrpides, N.; Markowitz, V.; Palaniappan, K.; Ivanova, N.; Mikhailova, N.; et al. The complete genome sequence of *Clostridium indolis* DSM 755T. *Stand. Genom. Sci.* **2015**, *9*, 1089–1104. [CrossRef]
89. Mechichi, T.; Labat, M.; Patel, B.K.C.; Woo, T.H.; Thomas, P.; Garcia, J. New aromatic *O*-demethylating homoacetogen from an olive mill wastewater treatment digester. *Int. J. Syst. Bacteriol.* **1997**, *49*, 1201–1209. [CrossRef]
90. Danylec, N.; Stoll, D.A.; Huch, M. Draft genome sequences of type strains of *Gordonibacter faecihominis*, *Paraeggerthella hongkongensis*, *Parvibacter caecicola*, *Slackia equolifaciens*, *Slackia faecicanis*, and *Slackia isoflavoniconvertens*. *Microbiol. Resour. Announc.* **2019**, *36*, e01532-18. [CrossRef]
91. Clavel, T.; Charrier, C.; Braune, A.; Wenning, M.; Blaut, M.; Haller, D. Isolation of bacteria from the ileal mucosa of TNF^{deltaARE} mice and description of *Enterorhabdus mucosicola* gen. nov., sp. nov. *Int. J. Syst. Evol. Microbiol.* **2009**, *59*, 1805–1812. [CrossRef]



Оригінал доступний за посиланням:
https://www.mdpi.com/2073-4425/13/11/1960.
Електронні додатки до статті доступні за посиланням:
https://www.mdpi.com/article/10.3390/genes13111960/s1.

Анотація: *Клінічно-значущі глікопептидні антибіотики залишаються одним з найуспішніших класів природних антибактеріальних засобів. Цей успіх, однак, знаходиться під загрозою через поширення генів стійкості до глікопептидів, також відомих як van-гени. Важливо простежити та зрозуміти можливі шляхи поширення van-генів. У поточній роботі ми представляємо комплексний біоінформатичний аналіз, спрямований на відстеження van-генів за межами типу Actinobacteria – найбільш вірогідного природного резервуару van-генів. Ми демонструємо, що два додаткових класи Грам-позитивних бактерій, Erysipelotrichia і Ktedonobacteria, а також один клас Грам-негативних бактерій, Anaerolineae, несуть van-гени. Крім того, ми демонструємо, що різні нові роди, які належать до класів Clostridia та Bacilli, також несуть van-гени. Більшість виявлених van-локусів колокалізовані з генами транспозаз різних типів. Нарешті, ми пропонуємо заснований на філогенії сценарій поширення van-генів, розкриваючи мережу послідовних подій горизонтального переносу генів, що зв'язують тип Actinobacteria з п'ятьма іншими класами бактерій, які несуть van-гени.*

Висновки

Отримані в роботі результати дозволяють вирішувати нагальні проблеми пошуку і створення антибіотиків нового покоління та нових механізмів генетичної детермінації синтезу антибіотиків. У результаті проведеної роботи за допомогою методів молекулярної генетики, геноміки і біоінформатики встановлено взаємозв'язки між організацією і експресією структурних генів, генів плейотропних і шлях-специфічних регуляторів, та біосинтезом антибіотиків. На підставі глобального філогеномного аналізу геномів бактерій, доступних в базі даних GenBank, встановлено закономірності поширення і еволюції генів стійкості до глікопептидних антибіотиків. Розроблено підходи до активації мовчазного вторинного метаболому актиноміцетів із використанням генів плейотропних регуляторів. Оптимізовано методи перенесення плазмідної ДНК і надекспресії генів у актиноміцетів-продуцентів пептидних і полікетидних антибіотиків. В ході роботи отримано надпродуцентів клінічно-важливого антибіотика A40926, перспективного протиракового антибіотика ландоміцину А, а також відкрито новий глікопептидний антибіотик.

1. Представників ортологічної групи AdpA-регуляторів знайдено в 91 геномі стрептоміцетів, вказуючи на важливість відповідних генів для біології цих бактерій. Двох-доменна архітектура AdpA, ймовірно, постала внаслідок злиття двох окремих генів, кожен із яких кодував білок із функціями окремих доменів: GATase-1-подібного N-термінального і ДНК-зв'язувального доменів. Експресія окремого ДНК-зв'язувального домена AdpA здатна комплементувати відсутність повного білка.

2. У *S. albus* (=*albidoflavus*) J1074 відсутність гена *bldA* (і як наслідок правильної експресії *adpA*) сповільнює морфологічну диференціацію і призводить до зменшення продукції вторинних метаболітів. тРНК$^{\text{ЛЕЙ}}_{\text{UAA}}$ (кодована *bldA*) в процесі дозрівання модифікується ферментами MiaA і MiaB, гени яких є описаними вперше компонентами *adpA*-опосередкованого регуляторного каскаду.

3. Плазміди на основі систем інтеграції фагів φC31, φBT1, VWB1, і pSG5-реплікона можна із високою частотою переносити в клітини *A. rectilineatus* NRRL B-16090, *S. roseochromogenes* NRRL 3504, *N. coxensis* DSM 45129, за допомогою кон'югації з *E. coli*, використовуючи спори цих штамів актиноміцетів як реципієнти. Гетерологічні промотори *aac(3)IVp, moeE5p* і *tei2p* є ефективними знаряддями для надекспресії генів в NRRL B-16090. Вектор pGCymRP21 є ефективною платформою для експресії генів в NRRL 3504: надекспресія генів шлях-специфічних регуляторів хлоробіоцину і новобіоцину із використанням pGCymRP21 привела до збільшення рівня продукції хлоробіоцину.

4. У продуцента глікопептидного антибіотика A40926 *N. gerenzanensis* ATCC 39727 найактивнішим конститутивним промотором є *aac(3)IVp*. Надекспресія генів шлях-специфічних регуляторів біосинтезу A40926 *dbv3* і *dbv4* за допомогою *aac(3)IVp* веде до зростання продукції A40926 рекомбінантними штамами.

5. Секвеновано геном *N. coxensis* DSM 45129. Цей штам здатний до продукції досі не описаного глікопептидного антибіотика, названого A50926; відповідний кластер генів біосинтезу – названий *noc* – анотовано і охарактеризовано. A50926 подібний за структурою до A40926, однак містить некарбоксильований Glc*N*-ацильний залишок, приєднаний до 4 амінокислоти аглікона. Гетерологічні регулятори *dbv3* і *dbv4* збільшують рівень продукції A50926, а експресія гена гексозооксидази *dbv29* веде до продукції A40926 в DSM 45129.

6. Кластери біосинтетичних генів, подібні до кластера генів біосинтезу глікопептидного антибіотика типу V комплестатину поширені в стрептоміцетів і можуть кодувати біосинтез норкомплестатину, *N*-малоніл-норкомплестатину і *N*-ацетил-норкомплестатину.

7. Секвеновано геном продуцента ландоміцину A *S. cyanogenus* S136. *adpA* алель S136 нефункціональний, тоді як експресія гетерологічних алелів *adpA* веде до багатократного зростання продукції ландоміцину A в S136. Експресія гетерологічних алелів *adpA* веде до активації біосинтезу полієнового

полікетиду люцензоміцину, а його КБГ – названий *lcm* – описано і охарактеризовано. Активація *lcm* опосередкована активністю генів шлях-специфічних регуляторів *lcmRII* і *lcmRIII*.

8. Вивчено механізми стійкості до тейкопланіну в продуцента *A. teichomyceticus* ATCC 31121. Висока стійкість ATCC 31121 до власного продукту забезпечена комбінацією факторів: конститутивною експресією оперонів стійкості *tei7-6-5* і *tei2-3*, незалежною від шлях-специфічної регуляції біосинтезу тейкопланіну; високою активністю промоторів *tei2p* і *tei7p*; конститутивною фосфорилазною активністю Tei3. Сенсорний домен Tei3 не здатний розпізнавати тейкопланін, але реагує на споріднений ГПА A40926.

9. Гени стійкості до ГПА – *van*-гени – поширені в актинобактерій, які є їх природним резервуаром. Аналіз філогенії і розповсюдження *van*-генів актинобактерій демонструє, що присутність цих генів не корелює з присутністю КБГ ГПА, а самі ці гени є преадаптацією для отримання КБГ шляхом горизонтального генного переносу. За межами актинобактерій *van*-гени поширилися до класів Грам-позитивних бактерій *Bacilli*, *Clostridia*, *Erysipelotrichia* і *Ktedonobacteria*, а також до Грам-негативних бактерій – класу *Anaerolineae*.

Список літературних джерел

1. Hutchings M, Truman A, Wilkinson B (2019) Antibiotics: past, present and future. Curr Opin Microbiol 51:72–80. https://doi.org/10.1016/j.mib.2019.10.008
2. Waksman SA, Woodruff HB (1940) The soil as a source of microorganisms antagonistic to disease-producing bacteria. J Bacteriol 40(4): 581–600. https://doi.org/10.1128/jb.40.4.581-600.1940
3. Medema MH, Kottmann R, Yilmaz P, et al (2015) Minimum Information about a biosynthetic gene cluster. Nat Chem Biol 11:625–631. https://doi.org/10.1038/nchembio.1890
4. Nazari B, Forneris CC, Gibson MI, et al (2017) *Nonomuraea* sp. ATCC 55076 harbours the largest actinomycete chromosome to date and the kistamicin biosynthetic gene cluster. Medchemcomm 8:780–788. https://doi.org/10.1039/c6md00637j
5. Hoskisson PA, Seipke RF (2020) Cryptic or silent? The known unknowns, unknown knowns, and unknown unknowns of secondary metabolism. MBio 11:1–5. https://doi.org/10.1128/mBio.02642-20
6. Reen FJ, Romano S, Dobson ADW, O'Gara F (2015) The sound of silence: activating silent biosynthetic gene clusters in marine microorganisms. Mar Drugs 13:4754–4783. https://doi.org/10.3390/md13084754
7. Davies J (2006) Where have all the antibiotics gone? Can J Infect Dis Med Microbiol 17:287–290. https://doi.org/10.1155/2006/707296
8. Graham DW, Bergeron G, Bourassa MW, et al (2019) Complexities in understanding antimicrobial resistance across domesticated animal, human, and environmental systems. Ann N Y Acad Sci 1441:17–30. https://doi.org/10.1111/nyas.14036
9. Scott HM, Acuff G, Bergeron G, et al (2019) Critically important antibiotics: criteria and approaches for measuring and reducing their use in food animal agriculture. Ann N Y Acad Sci 1441:8–16. https://doi.org/10.1111/nyas.14058
10. Hassoun-Kheir N, Stabholz Y, Kreft JU, et al (2020) Comparison of antibiotic-

resistant bacteria and antibiotic resistance genes abundance in hospital and community wastewater: a systematic review. Sci Total Environ 743:140804. https://doi.org/10.1016/j.scitotenv.2020.140804

11. Naylor NR, Atun R, Zhu N, et al (2018) Estimating the burden of antimicrobial resistance: a systematic literature review. Antimicrob Resist Infect Control 7:58. https://doi.org/10.1186/s13756-018-0336-y
12. O' Neil J (2014) Review on Antibiotic resisitance. Antimicrobial Resistance: tackling a crisis for the health and wealth of nations. 1–16. https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf
13. Ansari S, Hays JP, Kemp A, et al (2021) The potential impact of the COVID-19 pandemic on global antimicrobial and biocide resistance: an amr insights global perspective. JAC-Antimicrobial Resist 3(2):dlab038. https://doi.org/10.1093/jacamr/dlab038
14. Ghosh S, Bornman C, Zafer MM (2021) Antimicrobial resistance threats in the emerging Covid-19 pandemic: where do we stand? J Infect Public Health 14:555–560. https://doi.org/10.1016/j.jiph.2021.02.011
15. Rodríguez-Baño J, Rossolini GM, Schultsz C, et al (2021) Antimicrobial resistance research in a post-pandemic world: insights on antimicrobial resistance research in the COVID-19 pandemic. J Glob Antimicrob Resist 25:5–7. https://doi.org/10.1016/j.jgar.2021.02.013
16. Marcone GL, Binda E, Berini F, Marinelli F (2018) Old and new glycopeptide antibiotics: from product to gene and back in the post-genomic era. Biotechnol Adv 36:534–554. https://doi.org/10.1016/j.biotechadv.2018.02.009
17. WHO list of critically important antimicrobials for human medicine (WHO CIA list) 6th Rev. https://www.who.int/publications/i/item/9789241515528
18. Williams DH (1996) The glycopeptide story – how to kill the deadly "Superbugs." Nat Prod Rep 13:469–477. https://doi.org/10.1039/NP9961300469
19. Nicolaou KC, Boddy CNC, Bräse S, Winssinger N (1999) Chemistry, biology,

and medicine of the glycopeptide antibiotics. Angew Chemie - Int Ed 38:2096–2152. https://doi.org/10.1002/(sici)1521-3773(19990802)38:15<2096::aid-anie2096>3.0.co;2-f

20. Yim G, Thaker MN, Koteva K, Wright G (2014) Glycopeptide antibiotic biosynthesis. J Antibiot (Tokyo) 67:31–41. https://doi.org/10.1038/ja.2013.117
21. Bartley J (2002) First case of VRSA identified in Michigan. Infect Control Hosp Epidemiol 23(8):480. https://doi.org/10.1017/s0195941700082333.
22. Sletvold H, Johnsen PJ, Hamre I, et al (2008) Complete sequence of *Enterococcus faecium* pVEF3 and the detection of an ω-ε-ζ toxin-antitoxin module and an ABC transporter. Plasmid 60:75–85. https://doi.org/10.1016/j.plasmid.2008.04.002
23. Liu G, Chater KF, Chandra G, et al (2013) Molecular regulation of antibiotic biosynthesis in *Streptomyces*. Microbiol Mol Biol Rev 77:112–143. https://doi.org/10.1128/mmbr.00054-12
24. Rebets Y, Dutko L, Ostash B, et al (2008) Function of *lanI* in regulation of landomycin A biosynthesis in *Streptomyces cyanogenus* S136 and cross-complementation studies with *Streptomyces* antibiotic regulatory proteins encoding genes. Arch Microbiol 189:111–120. https://doi.org/10.1007/s00203-007-0299-5
25. Lo Grasso L, Maffioli S, Sosio M, et al (2015) Two master switch regulators trigger A40926 biosynthesis in *Nonomuraea* sp. strain ATCC 39727. J Bacteriol 197:2536–2544. https://doi.org/10.1128/JB.00262-15
26. Horbal L, Kobylyanskyy A, Truman AW, et al (2014) The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. Appl Microbiol Biotechnol 98:9295–9309. https://doi.org/10.1007/s00253-014-5969-z
27. McCormick JR, Flärdh K (2012) Signals and regulators that govern *Streptomyces* development. FEMS Microbiol Rev 36:206–231. https://doi.org/10.1111/j.1574-6976.2011.00317.x
28. Ostash B (2018) Regulatory genes of *adpA* subfamily in *Streptomyces*: function and evolution. https://doi.org/doi:10.2478/9783110627770

29. Horbal L, Kobylyanskyy A, Yushchuk O, et al (2013) Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus* – producer of teicoplanin, drug of last defense. J Biotechnol 168:367–372. https://doi.org/10.1016/j.jbiotec.2013.10.018
30. Stegmann E, Pelzer S, Bischoff D, et al (2006) Genetic analysis of the balhimycin (vancomycin-type) oxygenase genes. J Biotechnol 124:640–653. https://doi.org/10.1016/j.jbiotec.2006.04.009
31. Blin K, Shaw S, Kloosterman AM, et al (2021) AntiSMASH 6.0: Improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. https://doi.org/10.1093/nar/gkab335
32. Ventura M, Canchaya C, Tauch A, et al (2007) Genomics of *Actinobacteria*: tracing the evolutionary history of an ancient phylum. Microbiol Mol Biol Rev 71:495–548. https://doi.org/10.1128/mmbr.00005-07
33. Yushchuk O, Zhukrovska K, Berini F, et al (2022) Genetics behind the glycosylation patterns in the biosynthesis of dalbaheptides. Front Chem 10:1–9. https://doi.org/10.3389/fchem.2022.858708
34. Yushchuk O, Ostash B (2022) Glycopeptide antibiotics: genetics, chemistry, and new screening approaches. Nat Prod from Actinomycetes 411–444. https://doi.org/10.1007/978-981-16-6132-7_16
35. Yushchuk O, Ostash B, Truman AW, et al (2020) Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes. Appl Microbiol Biotechnol 104:3279–3291. https://doi.org/10.1007/s00253-020-10436-y
36. Yushchuk O, Kharel M, Ostash I, Ostash B (2019) Landomycin biosynthesis and its regulation in *Streptomyces*. Appl Microbiol Biotechnol 103:1659–1665. https://doi.org/10.1007/s00253-018-09601-1
37. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, New-York
38. Kieser T, Bibb MJ, Buttner MJ, et al (2000) Practical *Streptomyces* genetics. John Innes Found Norwich, Engl
39. Gregory MA, Till R, Smith MCM (2003) Integration site for *Streptomyces*

phage φBT1 and development of site-specific integrating vectors. J Bacteriol 185:5320–5323. https://doi.org/10.1128/JB.185.17.5320-5323.2003

40. Sekurova ON, Brautaset T, Sletta H, et al (2004) *In vivo* analysis of the regulatory genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455 reveals their differential control over antibiotic biosynthesis. J Bacteriol 186:1345–1354. https://doi.org/10.1128/JB.186.5.1345-1354.2004
41. Muth G (2018) The pSG5-based thermosensitive vector family for genome editing and gene expression in actinomycetes. Appl Microbiol Biotechnol 102:9067–9080. https://doi.org/10.1007/s00253-018-9334-5
42. Myronovskyi M, Welle E, Fedorenko V, Luzhetskyy A (2011) β-Glucuronidase as a sensitive and versatile reporter in *Actinomycetes*. Appl Environ Microbiol 77:5370–5383. https://doi.org/10.1128/AEM.00434-11
43. Marcone GL, Foulston L, Binda E, et al (2010) Methods for the genetic manipulation of *Nonomuraea* sp. ATCC 39727. J Ind Microbiol Biotechnol 37:1097–1103. https://doi.org/10.1007/s10295-010-0807-5
44. Marcone GL, Binda E, Reguzzoni M, et al (2017) Classification of *Actinoplanes* sp. ATCC 33076, an actinomycete that produces the glycolipodepsipeptide antibiotic ramoplanin, as *Actinoplanes ramoplaninifer* sp. Nov. Int J Syst Evol Microbiol 67:4181–4188. https://doi.org/10.1099/ijsem.0.002281
45. Marcone GL, Beltrametti F, Binda E, et al (2010) Novel mechanism of glycopeptide resistance in the A40926 producer *Nonomuraea* sp. ATCC 39727. Antimicrob Agents Chemother 54:2465–2472. https://doi.org/10.1128/AAC.00106-10
46. Marcone GL, Binda E, Carrano L, et al (2014) Relationship between glycopeptide production and resistance in the actinomycete *Nonomuraea* sp. ATCC 39727. Antimicrob Agents Chemother 58:5191–5201. https://doi.org/10.1128/AAC.02626-14
47. Taurino C, Frattini L, Marcone GL, et al (2011) *Actinoplanes teichomyceticus* ATCC 31121 as a cell factory for producing teicoplanin. Microb Cell Fact

10:82. https://doi.org/10.1186/1475-2859-10-82

48. Monteiro SM, Clemente JJ, Henriques AO, et al (2005) A procedure for high-yield spore production by *Bacillus subtilis*. Biotechnol Prog 21:1026–1031. https://doi.org/10.1021/bp050062z
49. Yushchuk O, Horbal L, Datsyuk J, et al (2016) Peculiarities of *Actinoplanes teichomyceticus* NRRL-B 16726 morphology and life cycle. Вісник Львівського університету. Серія біологічна 71:126–136.
50. Bekiesch P, Zehl M, Domingo-Contreras E, et al (2020) Viennamycins: lipopeptides produced by a *Streptomyces* sp. J Nat Prod 83:2381–2389. https://doi.org/10.1021/acs.jnatprod.0c00152
51. Gordon D, Green P (2013) CONSED: A graphical editor for next-generation sequencing. Bioinformatics 29:2936–2937. https://doi.org/10.1093/bioinformatics/btt515
52. Seemann T (2014) PROKKA: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153
53. Bentley SD, Chater KF, Cerdeño-Tárraga A-MM, et al (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417:141–147. https://doi.org/10.1038/417141a
54. Yarza P, Yilmaz P, Pruesse E, et al (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12(9):635-45. https://doi.org/10.1038/nrmicro3330.
55. Ueda K, Seki T, Kudo T, et al (1999) Two distinct mechanisms cause heterogeneity of 16S rRNA. J Bacteriol 181:78–82. https://doi.org/10.1128/JB.181.1.78-82.1999
56. Henkel T, Rohr J, Beale JM, Schwenen L (1990) Landomycins, new angucycline antibiotics from *Streptomyces* sp. I. Structural studies on landomycins A-D. J Antibiot (Tokyo) 43:492–503. https://doi.org/10.7164/antibiotics.43.492
57. Yushchuk O, Ostash B, Pham TH, et al (2016) Characterization of the post-assembly line tailoring processes in teicoplanin biosynthesis. ACS Chem Biol 11:2254–2264. https://doi.org/10.1021/acschembio.6b00018

58. Holding AN, Spencer JB (2008) Investigation into the mechanism of phenolic couplings during the biosynthesis of glycopeptide antibiotics. ChemBioChem 9:2209–2214. https://doi.org/10.1002/cbic.200800303
59. Kearse M, Moir R, Wilson A, et al (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199
60. Kumar S, Stecher G, Li M, et al (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096
61. Sievers F, Higgins DG (2014) Clustal Omega. Curr Protoc Bioinforma 2014:3.13.1-3.13.16. https://doi.org/10.1002/0471250953.bi0313s48
62. Röttig M, Medema MH, Blin K, et al (2011) NRPSpredictor2 – a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res 39(Web Server issue):W362-7. https://doi.org/10.1093/nar/gkr323
63. Medema MH, Takano E, Breitling R (2013) Detecting sequence homology at the gene cluster level with MultiGeneBlast. Mol Biol Evol 30:1218–1223. https://doi.org/10.1093/molbev/mst025
64. Marchler-Bauer A, Bryant SH (2004) CD-Search: Protein domain annotations on the fly. Nucleic Acids Res 32:327–331. https://doi.org/10.1093/nar/gkh454
65. Ohnishi Y, Kameyama S, Onaka H, Horinouchi S (1999) The A-factor regulatory cascade leading to streptomycin biosynthesis in *Streptomyces griseus* : identification of a target gene of the A-factor receptor. Mol Microbiol 34:102–111. https://doi.org/10.1046/j.1365-2958.1999.01579.x
66. Yamazaki H, Tomono A, Ohnishi Y, Horinouchi S (2004) DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. Mol Microbiol 53:555–572. https://doi.org/10.1111/j.1365-2958.2004.04153.x
67. Ohnishi Y, Yamazaki H, Kato JY, et al (2005) AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. Biosci

Biotechnol Biochem 69:431–439. https://doi.org/10.1271/bbb.69.431

68. Chater KF, Chandra G (2008) The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in *Streptomyces*. J Microbiol 46:1–11. https://doi.org/10.1007/s12275-007-0233-1
69. Thompson KM, Gottesman S (2014) The MiaA tRNA modification enzyme is necessary for robust RpoS expression in *Escherichia coli*. J Bacteriol 196(4):754-61. https://doi.org/10.1128/JB.01013-13.
70. Takano E, Tao M, Long F, et al (2003) A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. Mol Microbiol 50:475–486. https://doi.org/10.1046/j.1365-2958.2003.03728.x
71. Horbal L, Fedorenko V, Luzhetskyy A (2014) Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. Appl Microbiol Biotechnol 98:8641–8655. https://doi.org/10.1007/s00253-014-5918-x
72. Zaburannyi N, Rabyk M, Ostash B, et al (2014) Insights into naturally minimised *Streptomyces albus* J1074 genome. BMC Genomics 15:1–11. https://doi.org/10.1186/1471-2164-15-97
73. Myronovskyi M, Rosenkränzer B, Nadmid S, et al (2018) Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. Metab Eng 49:316–324. https://doi.org/10.1016/j.ymben.2018.09.004
74. Bilyk B, Luzhetskyy A (2014) Unusual site-specific DNA integration into the highly active pseudo-attB of the *Streptomyces albus* J1074 genome. Appl Microbiol Biotechnol 98:5095–5104. https://doi.org/10.1007/s00253-014-5605-y
75. González A, Rodríguez M, Braña AF, et al (2016) New insights into paulomycin biosynthesis pathway in *Streptomyces albus* J1074 and generation of novel derivatives by combinatorial biosynthesis. Microb Cell Fact 15:1–16. https://doi.org/10.1186/s12934-016-0452-4

76. Schweizer U, Bohleber S, Fradejas-Villar N (2017) The modified base isopentenyladenosine and its derivatives in tRNA. RNA Biol 14:1197–1208. https://doi.org/10.1080/15476286.2017.1294309

77. Zhao J, Leung HE, Winkler ME (2001) The *miaA* mutator phenotype of *Escherichia coli* K-12 requires recombination functions. J Bacteriol 183:1796–1800. https://doi.org/10.1128/JB.183.5.1796-1800.2001

78. Esberg B, Leung HC, Tsui HC, et al (1999) Identification of the *miaB* gene, involved in methylthiolation of isopentenylated A37 derivatives in the tRNA of *Salmonella typhimurium* and *Escherichia coli*. J Bacteriol 181:7256–7265. https://doi.org/10.1128/JB.181.23.7256-7265.1999

79. Rokytskyy I, Koshla O, Fedorenko V, Ostash B (2016) Decoding options and accuracy of translation of developmentally regulated UUA codon in *Streptomyces*: bioinformatic analysis. Springerplus 5:982. https://doi.org/10.1186/s40064-016-2683-6

80. Rigali S, Derouaux A, Giannotta F, Dusart J (2002) Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies. J Biol Chem 277:12507–12515. https://doi.org/10.1074/jbc.M110968200

81. Ostash B, Rebets Y, Myronovskyy M, et al (2011) Identification and characterization of the *Streptomyces globisporus* 1912 regulatory gene *lndYR* that affects sporulation and antibiotic production. Microbiology 157:1240–1249. https://doi.org/10.1099/mic.0.045088-0

82. Koshla O, Lopatniuk M, Rokytskyy I, et al (2017) Properties of *Streptomyces albus* J1074 mutant deficient in $tRNA^{Leu}_{UAA}$ gene *bldA*. Arch Microbiol 199:1175–1183. https://doi.org/10.1007/s00203-017-1389-7

83. Koshla O, Yushchuk O, Ostash I, et al (2019) Gene *miaA* for post-transcriptional modification of $tRNA_{XXA}$ is important for morphological and metabolic differentiation in *Streptomyces*. Mol Microbiol 112:249–265. https://doi.org/10.1111/mmi.14266

84. Tsypik O, Yushchuk O, Zaburannyi N, et al (2016) Transcriptional regulators of GntR family in *Streptomyces coelicolor* A3(2): analysis *in silico* and *in vivo*

of YtrA subfamily. Folia Microbiol (Praha) 61:209–220. https://doi.org/10.1007/s12223-015-0426-7

85. Rabyk M, Yushchuk O, Rokytskyy I, et al (2018) Genomic insights into evolution of AdpA family master regulators of morphological differentiation and secondary metabolism in *Streptomyces*. J Mol Evol 86:204–215. https://doi.org/10.1007/s00239-018-9834-z
86. Truman AW, Kwun MJ, Cheng J, et al (2014) Antibiotic resistance mechanisms inform discovery: identification and characterization of a novel *Amycolatopsis* strain producing ristocetin. Antimicrob Agents Chemother 58:5687–5695. https://doi.org/10.1128/AAC.03349-14
87. Ray L, Yamanaka K, Moore BS (2016) A peptidyl-transesterifying type I thioesterase in salinamide biosynthesis. Angew Chemie - Int Ed 55:364–367. https://doi.org/10.1002/anie.201508576
88. Heide L (2009) Chapter 18: Aminocoumarins. Mutasynthesis, chemoenzymatic synthesis, and metabolic engineering. Methods Enzymol 459:437-55. https://doi.org/10.1016/S0076-6879(09)04618-7.
89. Freitag A, Wemakor E, Li SM, Heide L (2005) Acyl transfer in clorobiocin biosynthesis: Involvement of several proteins in the transfer of the pyrrole-2-carboxyl moiety to the deoxysugar. ChemBioChem 6:2316–2325. https://doi.org/10.1002/cbic.200500252
90. Freitag A, Rapp H, Heide L, Li SM (2005) Metabolic engineering of aminocoumarins: inactivation of the methyltransferase gene *cloP* and generation of new clorobiocin derivatives in a heterologous host. ChemBioChem 6:1411–1418. https://doi.org/10.1002/cbic.200500019
91. Bierman M, Logan R, O'Brien K, et al (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene 116:43–49. https://doi.org/10.1016/0378-1119(92)90627-2
92. Van Mellaert L, Mei L, Lammertyn E, et al (1998) Site-specific integration of bacteriophage VWB genome into *Streptomyces venezuelae* and construction of a VWB-based integrative vector. Microbiology 144:3351–3358. https://doi.org/10.1099/00221287-144-12-3351

93. Ostash B, Makitrinskyy R, Walker S, Fedorenko V (2009) Identification and characterization of *Streptomyces ghanaensis* ATCC14672 integration sites for three actinophage-based plasmids. Plasmid 61:171–175. https://doi.org/10.1016/j.plasmid.2008.12.002
94. Eustáquio AS, Gust B, Luft T, et al (2003) Clorobiocin biosynthesis in *Streptomyces*: identification of the halogenase and generation of structural analogs. Chem Biol 10:279–288. https://doi.org/10.1016/S1074-5521(03)00051-6
95. Rückert C, Kalinowski J, Heide L, Apel AK (2014) Draft genome sequence of *Streptomyces roseochromogenes* subsp. *oscitans* DS 12.976, producer of the aminocoumarin antibiotic clorobiocin. Genome Announc 2(1):e01147-13. https://doi.org/10.1128/genomea.01147-13
96. Flinspach K, Westrich L, Kaysser L, et al (2010) Heterologous expression of the biosynthetic gene clusters of coumermycina, clorobiocin and caprazamycins in genetically modified *Streptomyces coelicolor* strains. Biopolymers 93:823–832. https://doi.org/10.1002/bip.21493
97. Eustáquio AS, Li SM, Heide L (2005) NovG, a DNA-binding protein acting as a positive regulator of novobiocin biosynthesis. Microbiology 151:1949–1961. https://doi.org/10.1099/mic.0.27669-0
98. Makitrynskyy R, Ostash B, Tsypik O, et al (2013) Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. Open Biol 3:130121. https://doi.org/10.1098/rsob.130121
99. Herrmann S, Siegl T, Luzhetska M, et al (2012) Site-specific recombination strategies for engineering actinomycete. Appl Environ Microbiol 78:1804–1812. https://doi.org/10.1128/AEM.06054-11
100. Horbal L, Luzhetskyy A (2016) Dual control system – a novel scaffolding architecture of an inducible regulatory device for the precise regulation of gene expression. Metab Eng 37:11–23. https://doi.org/10.1016/j.ymben.2016.03.008
101. Huang J, Shi J, Molle V, et al (2005) Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*. Mol Microbiol

58:1276–1287. https://doi.org/10.1111/j.1365-2958.2005.04879.x

102. Ostash B, Campbell J, Luzhetskyy A, Walker S (2013) MoeH5: A natural glycorandomizer from the moenomycin biosynthetic pathway. Mol Microbiol 90:1324–1338. https://doi.org/10.1111/mmi.12437
103. Yushchuk O, Homoniuk V, Datsiuk Y, et al (2020) Development of a gene expression system for the uncommon actinomycete *Actinoplanes rectilineatus* NRRL B-16090. J Appl Genet 61:141–149. https://doi.org/10.1007/s13353-019-00534-7
104. Melnyk S, Stepanyshyn A, Yushchuk O, et al (2022) Genetic approaches to improve clorobiocin production in *Streptomyces roseochromogenes* NRRL 3504. Appl Microbiol Biotechnol 106:1543–1556. https://doi.org/10.1007/s00253-022-11814-4
105. Sungthong R, Nakaew N (2015) The genus *Nonomuraea*: a review of a rare actinomycete taxon for novel metabolites. J. Basic Microbiol. 55(5):554-65. https://doi.org/10.1002/jobm.201300691.
106. D'Argenio V, Petrillo M, Pasanisi D, et al (2016) The complete 12 Mb genome and transcriptome of *Nonomuraea gerenzanensis* with new insights into its duplicated "magic" RNA polymerase. Sci Rep 6(1):18. https://doi.org/10.1038/s41598-016-0025-0
107. Narsing Rao MP, Dong ZY, Jiao JY, et al (2020) Genome sequence and comparative analysis of DRQ-2, the type strain of *Nonomuraea indica*. Genomics 112:2842–2844. https://doi.org/10.1016/j.ygeno.2020.03.023
108. Goldstein BP, Selva E, Gastaldo L, et al (1987) A40926, a new glycopeptide antibiotic with anti-*Neisseria* activity. Antimicrob Agents Chemother 31:1961–1966. https://doi.org/10.1128/AAC.31.12.1961
109. Grundy WE, Sinclair AC, Theriault RJ, et al (1956) Ristocetin, microbiologic properties. Antibiot Annu 687–692.
110. Spohn M, Kirchner N, Kulik A, et al (2014) Overproduction of ristomycin a by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. Antimicrob Agents Chemother 58:6185–6196. https://doi.org/10.1128/AAC.03512-14

111. Soriano A, Rossolini GM, Pea F (2020) The role of dalbavancin in the treatment of acute bacterial skin and skin structure infections (ABSSSIs). Expert Rev. Anti. Infect. Ther. 18(5):415-422. https://doi.org/10.1080/14787210.2020.1746643
112. Yushchuk O, Andreo-Vidal A, Marcone GL, et al (2020) New molecular tools for regulation and improvement of A40926 glycopeptide antibiotic production in *Nonomuraea gerenzanensis* ATCC 39727. Front Microbiol 11:. https://doi.org/10.3389/fmicb.2020.00008
113. Shawky RM, Puk O, Wietzorrek A, et al (2007) The border sequence of the balhimycin biosynthesis gene cluster from *Amycolatopsis balhimycina* contains *bbr*, encoding a StrR-like pathway-specific regulator. J Mol Microbiol Biotechnol 13:76–88. https://doi.org/10.1159/000103599
114. Horbal L, Zaburannyy N, Ostash B, et al (2012) Manipulating the regulatory genes for teicoplanin production in *Actinoplanes teichomyceticus*. World J Microbiol Biotechnol 28:2095–2100. https://doi.org/10.1007/s11274-012-1013-6
115. Greule A, Izoré T, Iftime D, et al (2019) Kistamicin biosynthesis reveals the biosynthetic requirements for production of highly crosslinked glycopeptide antibiotics. Nat Commun 10:2613. https://doi.org/10.1038/s41467-019-10384-w
116. Blin K, Shaw S, Steinke K, et al (2019) AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 47:W81–W87. https://doi.org/10.1093/nar/gkz310
117. Alduina R, Sosio M, Donadio S (2018) Complex regulatory networks governing production of the glycopeptide A40926. Antibiotics (Basel) 7(2):30. https://doi.org/10.3390/antibiotics7020030.
118. Li YS, Ho JY, Huang CC, et al (2007) A unique flavin mononucleotide-linked primary alcohol oxidase for glycopeptide A40926 maturation. J Am Chem Soc 129:13384–13385. https://doi.org/10.1021/ja075748x
119. Binda E, Marcone GL, Pollegioni L, Marinelli F (2012) Characterization of VanYn, a novel D,D-peptidase/ D,D-carboxypeptidase involved in glycopeptide

antibiotic resistance in *Nonomuraea* sp. ATCC 39727. FEBS J 279:3203–3213. https://doi.org/10.1111/j.1742-4658.2012.08706.x

120. Binda E, Marcone GL, Berini F, et al (2013) *Streptomyces* spp. as efficient expression system for a D,D-peptidase/D,D-carboxypeptidase involved in glycopeptide antibiotic resistance. BMC Biotechnol 13:24. https://doi.org/10.1186/1472-6750-13-24
121. Sosio M, Stinchi S, Beltrametti F, et al (2003) The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species. Chem Biol 10:541–549. https://doi.org/10.1016/S1074-5521(03)00120-0
122. Yushchuk O, Vior NM, Andreo-Vidal A, et al (2021) Genomic-led discovery of a novel glycopeptide antibiotic by *Nonomuraea coxensis* DSM 45129. ACS Chem Biol 16:915–928. https://doi.org/10.1021/acschembio.1c00170
123. Ara I, Kudo T, Matsumoto A, et al (2007) *Nonomuraea bangladeshensis* sp. nov. and *Nonomuraea coxensis* sp. nov. Int J Syst Evol Microbiol 57:1504–1509. https://doi.org/10.1099/ijs.0.65054-0
124. Puk O, Bischoff D, Kittel C, et al (2004) Biosynthesis of chloro-β-hydroxytyrosine, a nonproteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics. J Bacteriol 186:6093–6100. https://doi.org/10.1128/JB.186.18.6093-6100.2004
125. Kaneko I, Kamoshida K, Takahashi S (1989) Complestain, a potent anti-complement substance produced by streptomyces lavendulae. I. Fermentation, isolation and biological characterization. J Antibiot (Tokyo) 42:236–241. https://doi.org/10.7164/antibiotics.42.236
126. Culp EJ, Waglechner N, Wang W, et al (2020) Evolution-guided discovery of antibiotics that inhibit peptidoglycan remodelling. Nature 578:582–587. https://doi.org/10.1038/s41586-020-1990-9
127. Smith TJ, Blackman SA, Foster SJ (2000) Autolysins of *Bacillus subtilis*: multiple enzymes with multiple functions. Microbiology 146:249–262
128. Cheng A V., Wuest WM (2020) Phylogeny-guided approach yields glycopeptides with unique action. Trends Pharmacol. Sci. 41:297–299
129. Ющук ОС, Осташ БО, Горбаль ЛО, Федоренко ВО (2017) Реконструкція

філогенії кластерів генів біосинтезу глікопептидів. Фактори експериментальної еволюції організмів 20:109–115

130. Ющук ОС, Жукровська КА, Федоренко ВО (2022) Поширення кластерів генів біосинтезу комплестатину та подібних сполук у представників роду *Streptomyces*. Фактори експериментальної еволюції організмів 30:133–140
131. Lim FY, Sanchez JF, Wang CCC, Keller NP (2012) Toward awakening cryptic secondary metabolite gene clusters in filamentous fungi. Methods Enzymol 517:303–324. https://doi.org/10.1016/B978-0-12-404634-4.00015-2
132. Ostash B, Korynevska A, Stoika R, Fedorenko V (2009) Chemistry and biology of landomycins, an expanding family of polyketide natural products. Mini-Reviews Med Chem 9:1040–1051. https://doi.org/10.2174/138955709788922593
133. Ostash B, Rix U, Rix LLR, et al (2004) Generation of new landomycins by combinatorial biosynthetic manipulation of the *lndGT4* gene of the landomycin E cluster in *S. globisporus*. Chem Biol 11:547–555. https://doi.org/10.1016/j.chembiol.2004.03.011
134. Ostash I, Ostash B, Luzhetskyy A, et al (2008) Coordination of export and glycosylation of landomycins in *Streptomyces cyanogenus* S136. FEMS Microbiol Lett 285:195–202. https://doi.org/10.1111/j.1574-6968.2008.01225.x
135. Wang J, Schully KL, Pettis GS (2009) Growth-regulated expression of a bacteriocin, produced by the sweet potato pathogen *Streptomyces ipomoeae*, that exhibits interstrain inhibition. Appl Environ Microbiol 75:1236–1242. https://doi.org/10.1128/AEM.01598-08
136. Claesen J, Bibb MJ (2011) Biosynthesis and regulation of grisemycin, a new member of the linaridin family of ribosomally synthesized peptides produced by *Streptomyces griseus* IFO 13350. J Bacteriol 193:2510–2516. https://doi.org/10.1128/JB.00171-11
137. López-García MT, Santamarta I, Liras P (2010) Morphological differentiation and clavulanic acid formation are affected in a *Streptomyces clavuligerus adpA*-deleted mutant. Microbiology 156:2354–2365. https://doi.org/10.1099/mic.0.035956-0

138. Du YL, Li SZ, Zhou Z, et al (2011) The pleitropic regulator $AdpA_{ch}$ is required for natamycin biosynthesis and morphological differentiation in *Streptomyces chattanoogensis*. Microbiology 157:1300–1311. https://doi.org/10.1099/mic.0.046607-0
139. Xu J, Zhang J, Zhuo J, et al (2017) Activation and mechanism of a cryptic oviedomycin gene cluster via the disruption of a global regulatory gene, *adpA*, in *Streptomyces ansochromogenes*. J Biol Chem 292:19708–19720. https://doi.org/10.1074/jbc.M117.809145
140. Sun Y, Wang H, Ma K, Shao W (2018) Construction and characterization of the GFAT gene as a novel selection marker in *Aspergillus nidulans*. Appl Microbiol Biotechnol 102:7951–7962. https://doi.org/10.1007/s00253-018-9185-0
141. Wolański M, Donczew R, Kois-Ostrowska A, et al (2011) The level of AdpA directly affects expression of developmental genes in *Streptomyces coelicolor*. J Bacteriol 193:6358–6365. https://doi.org/10.1128/JB.05734-11
142. Hrab P, Rückert C, Busche T, et al (2021) Complete genome sequence of *Streptomyces cyanogenus* S136, producer of anticancer angucycline landomycin A. 3 Biotech 11:1–10. https://doi.org/10.1007/s13205-021-02834-4
143. Yushchuk O, Ostash I, Vlasiuk I, et al (2018) Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions. Appl Microbiol Biotechnol 102:8419–8428. https://doi.org/10.1007/s00253-018-9249-1
144. Kim J Do, Kang JE, Kim BS (2020) Postharvest disease control efficacy of the polyene macrolide lucensomycin produced by *Streptomyces plumbeus* strain CA5 against gray mold on grapes. Postharvest Biol Technol 162:111115. https://doi.org/10.1016/J.POSTHARVBIO.2019.111115
145. Martín JF, Aparicio JF (2009) Chapter 10: Enzymology of the polyenes pimaricin and candicidin biosynthesis. Methods Enzymol 459:215–242. https://doi.org/10.1016/S0076-6879(09)04610-2
146. Aparicio JF, Colina AJ, Ceballos E, Martín JF (1999) The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin: a new

polyketide synthase organization encoded by two subclusters separated by functionalization genes. J Biol Chem 274:10133–10139. https://doi.org/10.1074/jbc.274.15.10133

147. Aparicio JF, Barreales EG, Payero TD, et al (2016) Biotechnological production and application of the antibiotic pimaricin: biosynthesis and its regulation. Appl Microbiol Biotechnol 100:61–78. https://doi.org/10.1007/s00253-015-7077-0
148. Gaudiano S, Bravo P, Quilico A, et al (1966) The structure of lucensomycin. Part II. Tetrahedron Lett 7:3567–3571. https://doi.org/10.1016/S0040-4039(01)82829-X
149. Aparicio JF, Caffrey P, Gil JA, Zotchev SB (2003) Polyene antibiotic biosynthesis gene clusters. Appl Microbiol Biotechnol 61:179–188. https://doi.org/10.1007/s00253-002-1183-5
150. Antón N, Santos-Aberturas J, Mendes M V, et al (2007) PimM, a PAS domain positive regulator of pimaricin biosynthesis in *Streptomyces natalensis*. Microbiology 153:3174–3183. https://doi.org/10.1099/mic.0.2007/009126-0
151. Antón N, Mendes M V, Martín JF, Aparicio JF (2004) Identification of PimR as a positive regulator of pimaricin biosynthesis in *Streptomyces natalensis*. J Bacteriol 186:2567–2575. https://doi.org/10.1128/JB.186.9.2567-2575.2004
152. Yushchuk O, Ostash I, Mösker E, et al (2021) Eliciting the silent lucensomycin biosynthetic pathway in *Streptomyces cyanogenus* S136 via manipulation of the global regulatory gene *adpA*. Sci Rep 11(1):3507. https://doi.org/10.1038/s41598-021-82934-6
153. Патент України на винахід №119472 Україна, МПК (2006.01)C12P 1/06 C12P 19/64 C12N 15/80 C12R 1/465 Спосіб активування продукції сполук з антибіотичною та протигрибною властивостями у *Streptomyces cyanogenus* S136 / Ющук О., Осташ Б.; заявник і власник Львівський національний ун-т ім. І. Франка. – № a201702642, заявл. 21 березня 2017 р., опубл. 25.06.2019, Бюл. № 12.
154. Knight DR, Androga GO, Ballard SA, et al (2016) A phenotypically silent *vanB2* operon carried on a Tn*1549*-like element in *Clostridium difficile*.

mSphere 1(4):e00177-16. https://doi.org/10.1128/msphere.00177-16

155. Hutchings MI, Hong HJ, Buttner MJ (2006) The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*. Mol Microbiol 59:923–935. https://doi.org/10.1111/j.1365-2958.2005.04953.x
156. Arthur M, Quintiliani R. J (2001) Regulation of *vanA*- and *vanB*-type glycopeptide resistance in enterococci. Antimicrob Agents Chemother 45:375–381. https://doi.org/10.1128/AAC.45.2.375-381.2001
157. Kruse T, Levisson M, de Vos WM, Smidt H (2014) *vanI*: A novel D-Ala-D-Lac vancomycin resistance gene cluster found in *Desulfitobacterium hafniense*. Microb Biotechnol 7:456–466. https://doi.org/10.1111/1751-7915.12139
158. Koteva K, Hong HJ, Wang XD, et al (2010) A vancomycin photoprobe identifies the histidine kinase $VanS_{sc}$ as a vancomycin receptor. Nat Chem Biol 6:327–329. https://doi.org/10.1038/nchembio.350
159. Kwun MJ, Novotna G, Hesketh AR, et al (2013) *In vivo* studies suggest that induction of VanS-dependent vancomycin resistance requires binding of the drug to D-Ala-D-Ala termini in the peptidoglycan cell wall. Antimicrob. Agents Chemother. 57:4470–4480. https://doi.org/10.1128/AAC.00523-13
160. Foucault M-L, Courvalin P, Grillot-Courvalin C (2009) Fitness cost of *vanA*-type vancomycin resistance in methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 53:2354–2359. https://doi.org/10.1128/AAC.01702-08
161. Foucault M-L, Depardieu F, Courvalin P, Grillot-Courvalin C (2010) Inducible expression eliminates the fitness cost of vancomycin resistance in enterococci. Proc Natl Acad Sci USA 107:16964–16969. https://doi.org/10.1073/pnas.1006855107
162. Schäberle TF, Vollmer W, Frasch HJ, et al (2011) Self-resistance and cell wall composition in the glycopeptide producer *Amycolatopsis balhimycina*. Antimicrob Agents Chemother 55:4283–4289. https://doi.org/10.1128/AAC.01372-10
163. Kilian R, Frasch HJ, Kulik A, et al (2016) The VanRS homologous two-component system $VnlRS_{Ab}$ of the glycopeptide producer amycolatopsis

balhimycina activates transcription of the *vanHAX*$_{Sc}$ genes in *Streptomyces coelicolor*, but not in *A. balhimycina*. Microb Drug Resist 22:499–509. https://doi.org/10.1089/mdr.2016.0128

164. Frasch HJ, Kalan L, Kilian R, et al (2016) Alternative pathway to a glycopeptide-resistant cell wall in the balhimycin producer *Amycolatopsis balhimycina*. ACS Infect Dis 1:243–252. https://doi.org/10.1021/acsinfecdis.5b00011
165. Binda E, Cappelletti P, Marinelli F, Marcone GL (2018) Specificity of induction of glycopeptide antibiotic resistance in the producing actinomycetes. Antibiot (Basel) 7(2):36. https://doi.org/10.3390/antibiotics7020036
166. Stegmann E, Frasch HJ, Kilian R, Pozzi R (2015) Self-resistance mechanisms of actinomycetes producing lipid II-targeting antibiotics. Int J Med Microbiol 305:190–195. https://doi.org/10.1016/j.ijmm.2014.12.015
167. Beltrametti F, Consolandi A, Carrano L, et al (2007) Resistance to glycopeptide antibiotics in the teicoplanin producer is mediated by *van* gene homologue expression directing the synthesis of a modified cell wall peptidoglycan. Antimicrob Agents Chemother 51:1135–1141. https://doi.org/10.1128/AAC.01071-06
168. Yushchuk O, Homoniuk V, Ostash B, et al (2020) Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. J Antibiot (Tokyo) 73:255–259. https://doi.org/10.1038/s41429-019-0274-9
169. Yushchuk O, Horbal L, Ostash B, et al (2019) Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. Appl Microbiol Biotechnol 103:4089–4102. https://doi.org/10.1007/s00253-019-09789-w
170. Yushchuk O, Zhukrovska K, Ostash B, et al (2022) Heterologous expression reveals ancient properties of Tei3 – a VanS ortholog from the teicoplanin producer *Actinoplanes teichomyceticus*. Int J Mol Sci 23(24):15713. https://doi.org/10.3390/ijms232415713
171. Rioseras B, Yaguë P, López-Garciá MT, et al (2016) Characterization of SCO4439, a D-alanyl-D-alanine carboxypeptidase involved in spore cell wall

maturation, resistance, and germination in *Streptomyces coelicolor*. Sci Rep 6:21659. https://doi.org/10.1038/srep21659

172. Binda E, Marinelli F, Marcone GL (2014) Old and new glycopeptide antibiotics: action and resistance. Antibiotics (Basel) 3:572–594
173. Andreo-Vidal A, Binda E, Fedorenko V, et al (2021) Genomic insights into the distribution and phylogeny of glycopeptide resistance determinants within the *Actinobacteria* phylum. Antibiotics (Basel) 10:1533. https://doi.org/10.3390/antibiotics10121533
174. Khosa S, Lagedroste M, Smits SHJ (2016) Protein defense systems against the lantibiotic nisin: function of the immunity protein NisI and the resistance protein NSR. Front Microbiol 7:1–9. https://doi.org/10.3389/fmicb.2016.00504
175. Donadio S, Sosio M, Stegmann E, et al (2005) Comparative analysis and insights into the evolution of gene clusters for glycopeptide antibiotic biosynthesis. Mol Genet Genomics 274:40–50. https://doi.org/10.1007/s00438-005-1156-3
176. Menges R, Muth G, Wohlleben W, Stegmann E (2007) The ABC transporter Tba of *Amycolatopsis balhimycina* is required for efficient export of the glycopeptide antibiotic balhimycin. Appl Microbiol Biotechnol 77:125–134. https://doi.org/10.1007/s00253-007-1139-x
177. Ющук О, Жукровська К, Федоренко В (2022) Філогенія білків-експортерів глікопептидних і деяких споріднених антибіотиків. Вісник Львівського університету. Серія біологічна 86:33–46. https://doi.org/10.14341/cbai-2022-79
178. Rushton-Green R, Darnell RL, Taiaroa G, et al (2019) Agricultural origins of a highly persistent lineage of vancomycin-resistant *Enterococcus faecalis* in New Zealand. Appl Environ Microbiol 85(13):e00137-19. https://doi.org/10.1128/AEM.00137-19
179. Mello SS, Van Tyne D, Lebreton F, et al (2020) A mutation in the glycosyltransferase gene *lafB* causes daptomycin hypersusceptibility in *Enterococcus faecium*. J Antimicrob Chemother 75:36–45. https://doi.org/10.1093/jac/dkz403

180. Ballard SA, Pertile KK, Lim M, et al (2005) Molecular characterization of *vanB* elements in naturally occurring gut anaerobes. Antimicrob Agents Chemother 49:1688–1694. https://doi.org/10.1128/AAC.49.5.1688-1694.2005
181. Kinnear CL, Hansen E, Morley VJ, et al (2020) Daptomycin treatment impacts resistance in off-target populations of vancomycin-resistant *Enterococcus faecium*. PLoS Biol 18:1–21. https://doi.org/10.1371/journal.pbio.3000987
182. Bohlmann L, De Oliveira DMP, El-Deeb IM, et al (2018) Chemical synergy between ionophore PBT2 and zinc reverses antibiotic resistance. MBio 9(6):e02391-18. https://doi.org/10.1128/mBio.02391-18
183. Khan S, Sung K, Marasa B, et al (2015) Draft genome sequence of multidrug-resistant *Enterococcus faecium* clinical isolate VRE3, with a sequence type 16 pattern and novel structural arrangement of Tn*1546*. Genome Announc 3:1000000. https://doi.org/10.1128/genomeA.00871-15
184. Melegh S, Nyul A, Kovács K, et al (2018) Dissemination of VanA-Type *Enterococcus faecium* isolates in Hungary. Microb Drug Resist 24:1376–1390. https://doi.org/10.1089/mdr.2017.0296
185. Eshaghi A, Shahinas D, Li A, et al (2015) Characterization of an *Enterococcus gallinarum* isolate carrying a dual *vanA* and *vanB* cassette. J Clin Microbiol 53:2225–2229. https://doi.org/10.1128/JCM.03267-14
186. Zaheer R, Cook SR, Barbieri R, et al (2020) Surveillance of *Enterococcus* spp. reveals distinct species and antimicrobial resistance diversity across a One-Health continuum. Sci Rep 10:1–16. https://doi.org/10.1038/s41598-020-61002-5
187. Weigel LM, Clewell DB, Gill SR, et al (2003) Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. Science 302:1569–1571. https://doi.org/10.1126/science.1090956
188. Panesso D, Planet PJ, Diaz L, et al (2015) Methicillin-susceptible, vancomycin-resistant *Staphylococcus aureus*, Brazil. Emerg Infect Dis 21:1844–1848. https://doi.org/10.3201/eid2110.141914
189. Lim SK, Tanimoto K, Tomita H, Ike Y (2006) Pheromone-responsive conjugative vancomycin resistance plasmids in *Enterococcus faecalis* isolates

from humans and chicken feces. Appl Environ Microbiol 72:6544–6553. https://doi.org/10.1128/AEM.00749-06

190. Hospital-acquired AMS, Kos VN, Desjardins C a, et al (2012) Comparative genomics of vancomycin-resistant *Staphylococcus aureus* strains and their positions within the clade most commonly infection in the United States. MBio 3:1–9. https://doi.org/10.1128/mBio.00112-12.Editor
191. Zhu W, Murray PR, Huskins WC, et al (2010) Dissemination of an *Enterococcus* Inc18-like *vanA* plasmid associated with vancomycin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother 54:4314–4320. https://doi.org/10.1128/AAC.00185-10
192. Garnier F, Taourit S, Glaser P, et al (2000) Characterization of transposon Tn*1549*, conferring VanB-type resistance in *Enterococcus* spp. Microbiology 146:1481–1489. https://doi.org/10.1099/00221287-146-6-1481
193. De Been M, Van Schaik W, Cheng L, et al (2013) Recent recombination events in the core genome are associated with adaptive evolution in *Enterococcus faecium*. Genome Biol Evol 5:1524–1535. https://doi.org/10.1093/gbe/evt111
194. Sletvold H, Johnsen PJ, Wikmark OG, et al (2010) Tn*1546* is part of a larger plasmid-encoded genetic unit horizontally disseminated among clonal *Enterococcus faecium* lineages. J Antimicrob Chemother 65:1894–1906. https://doi.org/10.1093/jac/dkq219
195. Halvorsen EM, Williams JJ, Bhimani AJ, et al (2011) Txe, an endoribonuclease of the enterococcal Axe-Txe toxin-antitoxin system, cleaves mRNA and inhibits protein synthesis. Microbiology 157:387–397. https://doi.org/10.1099/mic.0.045492-0
196. Szakacs TA, Kalan L, McConnell MJ, et al (2014) Outbreak of vancomycin-susceptible *Enterococcus faecium* containing the wild-type *vanA* gene. J Clin Microbiol 52:1682–1686. https://doi.org/10.1128/JCM.03563-13
197. Sletvold H, Johnsen PJ, Simonsen GS, et al (2007) Comparative DNA analysis of two *vanA* plasmids from *Enterococcus faecium* strains isolated from poultry and a poultry farmer in Norway. Antimicrob Agents Chemother 51:736–739. https://doi.org/10.1128/AAC.00557-06

198. Guardabassi L, Perichon B, Van Heijenoort J, et al (2005) Glycopeptide resistance *vanA* operons in *Paenibacillus* strains isolated from soil. Antimicrob Agents Chemother 49:4227–4233. https://doi.org/10.1128/AAC.49.10.4227-4233.2005
199. Fraimow H, Knob C, Herrero IA, Patel R (2005) Putative VanRS-like two-component regulatory system associated with the inducible glycopeptide resistance cluster of *Paenibacillus popilliae*. Antimicrob Agents Chemother 49:2625–2633. https://doi.org/10.1128/AAC.49.7.2625-2633.2005
200. Ligozzi M, Lo Cascio G, Fontana R (1998) *vanA* gene cluster in a vancomycin-resistant clinical isolate of *Bacillus circulans*. Antimicrob Agents Chemother 42:2055–2059. https://doi.org/10.1128/aac.42.8.2055
201. Battistuzzi FU, Feijao A, Hedges SB (2004) A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. BMC Evol Biol 4:1–14. https://doi.org/10.1186/1471-2148-4-44
202. Yabe S, Sakai Y, Abe K, Yokota A (2017) Diversity of *Ktedonobacteria* with actinomycetes-like morphology in terrestrial environments. Microbes Environ 32:61–70. https://doi.org/10.1264/jsme2.ME16144
203. Nakahara N, Nobu MK, Takaki Y, et al (2019) *Aggregatilinea lenta* gen. nov., sp. nov., a slow-growing, facultatively anaerobic bacterium isolated from subseafloor sediment, and proposal of the new order *Aggregatilineales* ord. nov. within the class *Anaerolineae* of the phylum *Chloroflexi*. Int J Syst Evol Microbiol 69:1185–1194. https://doi.org/10.1099/ijsem.0.003291
204. Biddle AS, Leschine S, Huntemann M, et al (2015) The complete genome sequence of *Clostridium indolis* DSM 755T. Stand Genomic Sci 9:1089–1104. https://doi.org/10.4056/sigs.5281010
205. Mechichi T, Labat M, Patel BKC, et al (1997) *Clostridium methoxybenzovorans* sp. nov., a new aromatic *O*-demethylating homoacetogen from an olive mill wastewater treatment digester. Int J Syst Bacteriol 49(3):1201–1209. https://doi.org/10.1099/00207713-49-3-1201
206. Aminov RI, Mackie RI (2007) Evolution and ecology of antibiotic resistance

genes. FEMS Microbiol Lett 271:147–161. https://doi.org/10.1111/j.1574-6968.2007.00757.x

207. Yushchuk O, Binda E, Fedorenko V, Marinelli F (2022) Occurrence of *vanHAX* and related genes beyond the *Actinobacteria* phylum. Genes (Basel) 13(11):1960. https://doi.org/10.3390/genes13111960

Додаток

Список публікацій здобувача

Наукові праці, що розкривають основні наукові результати дисертації:

1. **Yushchuk O,** Zhukrovska K, Ostash B, Fedorenko V, Marinelli F (2022) Heterologous expression reveals ancient properties of Tei3 – a VanS ortholog from the teicoplanin producer *Actinoplanes teichomyceticus*. International journal of molecular sciences 23(24):15713. https://doi.org/10.3390/ijms232415713 **[IF 6,208; SCOPUS, WoS; SCImago Q2 – 2022 Molecular biology]** (*Особистий внесок здобувача: концептуалізація, планування і виконання експериментів, написання рукопису та підготовка статті до друку*).
2. **Yushchuk O,** Ostash B (2022) Glycopeptide antibiotics: genetics, chemistry, and new screening approaches. In: Rai, R.V., Bai, J.A. (eds.) Natural products from Actinomycetes. Springer, Singapore. https://doi.org/10.1007/978-981-16-6132-7_16 [**розділ монографії**] (*Особистий внесок здобувача: аналіз літератури, написання розділу, створення ілюстрацій, підготовка розділу до друку*).
3. **Ющук О,** Жукровська К, Федоренко В (2022) Філогенія білків-експортерів глікопептидних і деяких споріднених антибіотиків. Вісник Львівського університету. Серія біологічна 86:33-46. http://dx.doi.org/10.30970/vlubs.2022.86.03 [**фахове видання України**] (*Особистий внесок здобувача: концептуалізація, скринінг геномів і біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).
4. **Yushchuk O,** Zhukrovska K, Berini F, Fedorenko V, Marinelli F (2022) Genetics behind the glycosylation patterns in the biosynthesis of dalbaheptides. Frontiers in chemistry 10:858708. https://doi.org/10.3389/fchem.2022.858708 **[IF 5,545; SCOPUS, WoS; SCImago Q1 – 2022 Chemistry]** (*Особистий внесок здобувача: аналіз літератури і написання огляду, створення ілюстрацій, підготовка статті до друку*).

5. **Ющук О,** Жукровська К, Федоренко В (2022) Поширення кластерів генів біосинтезу комлестатину та подібних сполук у представників роду *Streptomyces*. Фактори експериментальної еволюції організмів 30:133-140. https://doi.org/10.7124/FEEO.v30.1474 [**фахове видання України**] (*Особистий внесок здобувача: концептуалізація, скринінг геномів і біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).

6. **Yushchuk O,** Binda E, Fedorenko V, Marinelli F (2022) Occurrence of *vanHAX* and related genes beyond the *Actinobacteria* phylum. Genes 13(11):1960. https://doi.org/10.3390/genes13111960 **[IF 4,141; SCOPUS, WoS; SCImago Q2 – 2022 Genetics]** (*Особистий внесок здобувача: концептуалізація, скринінг геномів і біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).

7. Melnyk S, Stepanyshyn A, **Yushchuk O,** Mandler M, Ostash I, Koshla O, Fedorenko V, Kahne D, Ostash B (2022) Genetic approaches to improve clorobiocin production in *Streptomyces roseochromogenes* NRRL 3504. Applied microbiology and biotechnology 106(4):1543–1556. https://doi.org/10.1007/s00253-022-11814-4 **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2022 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: участь в концептуалізації, розробка методики перенесення плазмід в NRRL 3504, аналіз продукції хлоробіоцину в рекомбінантних штамів, аналіз результатів ВЕРХ, написання рукопису та підготовка статті до друку*).

8. Andreo-Vidal A, Binda E, Fedorenko V, Marinelli F, **Yushchuk O** (2021) Genomic insights into the distribution and phylogeny of glycopeptide resistance determinants within the *Actinobacteria* phylum. Antibiotics (Basel, Switzerland) 10(12):1533. https://doi.org/10.3390/antibiotics10121533 **[IF 5,222; SCOPUS, WoS; SCImago Q2 – 2021 Infectious Diseases]** (*Особистий внесок здобувача: концептуалізація, скринінг геномів і біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).

9. Hrab P, Rückert C, Busche T, Ostash I, Kalinowski J, Fedorenko V, **Yushchuk O,** Ostash B (2021) Complete genome sequence of *Streptomyces cyanogenus* S136, producer of anticancer angucycline landomycin A. 3 Biotech 11(6):282.

https://doi.org/10.1007/s13205-021-02834-4 **[IF 2,893; SCOPUS, WoS; SCImago Q2 – 2021 Biotechnology]** (*Особистий внесок здобувача: концептуалізація і планування дослідження, верифікація результатів аналізу in silico, участь в підготовці рукопису до друку*).

10. **Yushchuk O,** Vior NM, Andreo-Vidal A, Berini F, Rückert C, Busche T, Binda E, Kalinowski J, Truman AW, Marinelli F (2021) Genomic-led discovery of a novel glycopeptide antibiotic by *Nonomuraea coxensis* DSM 45129. ACS Chemical Biology 16(5):915–928. https://doi.org/10.1021/acschembio.1c00170 **[IF 4,634; SCOPUS, WoS; SCImago Q1 – 2021 Biochemistry]** (*Особистий внесок здобувача: концептуалізація, аналіз in silico, характеристика кластера, виконання всіх генно-інженерних робіт, пошук оптимальних умов культивування DSM 45129 для продукції антибіотика, ферментація DSM 45129 і похідних, аналіз і інтерпретація даних ВЕРХ, написання рукопису та підготовка статті до друку*).

11. **Yushchuk O,** Ostash I, Mösker E, Vlasiuk I, Deneka M, Rückert C, Busche T, Fedorenko V, Kalinowski J, Süssmuth RD, Ostash B (2021) Eliciting the silent lucensomycin biosynthetic pathway in *Streptomyces cyanogenus* S136 via manipulation of the global regulatory gene *adpA*. Scientific reports 11(1):3507. https://doi.org/10.1038/s41598-021-82934-6 **[IF 4,996; SCOPUS, WoS; SCImago Q1 – 2021 Multidisciplinary]** (*Особистий внесок здобувача: концептуалізація, аналіз in silico, характеристика кластера, виконання всіх генно-інженерних робіт і аналізу отриманих рекомбінантів, пошук оптимальних умов культивування для продукції люцензоміцину, написання рукопису та підготовка статті до друку*).

12. **Yushchuk O,** Homoniuk V, Datsiuk Y, Ostash B, Marinelli F, Fedorenko V (2020) Development of a gene expression system for the uncommon actinomycete *Actinoplanes rectilineatus* NRRL B-16090. Journal of applied genetics 61(1):141–149. https://doi.org/10.1007/s13353-019-00534-7 **[IF 2,653; SCOPUS, WoS; SCImago Q2 – 2020 Genetics]** (*Особистий внесок здобувача: концептуалізація, виконання досліджень, написання рукопису та підготовка статті до друку*).

13. **Yushchuk O,** Homoniuk V, Ostash B, Marinelli F, Fedorenko V (2020) Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*. The Journal of antibiotics 73(4):255–259. https://doi.org/10.1038/s41429-019-0274-9 **[IF 3,424; SCOPUS, WoS; SCImago Q3 – 2020 Pharmacology]** (*Особистий внесок здобувача: концептуалізація, виконання досліджень, написання рукопису та підготовка статті до друку*).

14. **Yushchuk O,** Andreo-Vidal A, Marcone GL, Bibb M, Marinelli F, Binda E (2020) New molecular tools for regulation and improvement of A40926 glycopeptide antibiotic production in *Nonomuraea gerenzanensis* ATCC 39727 Frontiers in microbiology 11:8. https://doi.org/10.3389/fmicb.2020.00008 **[IF 6,064; SCOPUS, WoS; SCImago Q1 – 2020 Microbiology]** (*Особистий внесок здобувача: концептуалізація, виконання досліджень, написання рукопису та підготовка статті до друку*).

15. **Yushchuk O,** Binda E, Marinelli F (2020) Glycopeptide antibiotic resistance genes: distribution and function in the producer actinomycetes. Frontiers in microbiology 11:1173. https://doi.org/10.3389/fmicb.2020.01173 **[IF 6,064; SCOPUS, WoS; SCImago Q1 – 2020 Microbiology]** (*Особистий внесок здобувача: аналіз літератури і написання розділів присвячених продуцентам глікопептидів, створення ілюстрацій, підготовка статті до друку*).

16. **Yushchuk O,** Ostash B, Truman AW, Marinelli F, Fedorenko V (2020) Teicoplanin biosynthesis: unraveling the interplay of structural, regulatory, and resistance genes. Applied microbiology and biotechnology 104(8):3279–3291. https://doi.org/10.1007/s00253-020-10436-y **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2020 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: аналіз літератури і написання огляду, створення ілюстрацій, підготовка статті до друку*).

17. Koshla O, **Yushchuk O,** Ostash I, Dacyuk Y, Myronovskyi M, Jäger G, Süssmuth RD, Luzhetskyy A, Byström A, Kirsebom LA, Ostash B (2019) Gene *miaA* for post-transcriptional modification of $tRNA_{XXA}$ is important for morphological and metabolic differentiation in *Streptomyces*. Molecular microbiology 112(1):249–265. https://doi.org/10.1111/mmi.14266 **[IF 3,979; SCOPUS, WoS; SCImago Q1 – 2019**

Molecular biology] (*Особистий внесок здобувача: створення репортерної системи на основі плазміди pSAGA, створення подвійного мутанта ΔmiaA ΔbldA, створення плазмід із різними алелями гена adpA, комплементація мутантів ΔmiaA і ΔmiaA ΔbldA різними алелями adpA, аналіз властивостей мутантів і рекомбінантів, участь в написанні рукопису і підготовці статті до друку*).

18. **Yushchuk O,** Kharel M, Ostash I, Ostash B (2019) Landomycin biosynthesis and its regulation in *Streptomyces*. Applied microbiology and biotechnology 103(4):1659–1665. https://doi.org/10.1007/s00253-018-09601-1 **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2019 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: аналіз літератури і розділів огляду, присвячених організації кластерів генів біосинтезу і регуляції продукції ландоміцинів, створення ілюстрацій, підготовка статті до друку*).

19. **Yushchuk O,** Horbal L, Ostash B, Marinelli F, Wohlleben W, Stegmann E, Fedorenko V (2019) Regulation of teicoplanin biosynthesis: refining the roles of *tei* cluster-situated regulatory genes. Applied microbiology and biotechnology 103(10):4089–4102. https://doi.org/10.1007/s00253-019-09789-w **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2019 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: концептуалізація, планування і виконання експериментів, написання рукопису та підготовка статті до друку*).

20. **Yushchuk O,** Ostash I, Vlasiuk I, Gren T, Luzhetskyy A, Kalinowski J, Fedorenko V, Ostash B (2018) Heterologous AdpA transcription factors enhance landomycin production in *Streptomyces cyanogenus* S136 under a broad range of growth conditions. Applied microbiology and biotechnology 102(19):8419–8428. https://doi.org/10.1007/s00253-018-9249-1 **[IF 5,560; SCOPUS, WoS; SCImago Q1 – 2018 Applied Microbiology and Biotechnology]** (*Особистий внесок здобувача: концептуалізація, планування і виконання експериментів, написання рукопису та підготовка статті до друку*).

21. Rabyk M, **Yushchuk O,** Rokytskyy I, Anisimova M, Ostash B (2018) Genomic insights into evolution of adpa family master regulators of morphological differentiation and secondary metabolism in *Streptomyces*. Journal of molecular

evolution 86(3-4):204–215. https://doi.org/10.1007/s00239-018-9834-z **[IF 3,973; SCOPUS, WoS; SCImago Q2 – 2018 Genetics]** (*Особистий внесок здобувача: участь в аналізі даних, виконання експериментальної частини, участь в написанні рукопису та підготовці статті до друку*).

22. Koshla O, Lopatniuk M, Rokytskyy I, **Yushchuk O,** Dacyuk Y, Fedorenko V, Luzhetskyy A, Ostash B (2017) Properties of *Streptomyces albus* J1074 mutant deficient in $tRNA^{Leu}_{UAA}$ gene *bldA*. Archives of microbiology 199(8):1175–1183. https://doi.org/10.1007/s00203-017-1389-7 **[IF 2,667; SCOPUS, WoS; SCImago Q3 – 2017 Genetics]** (*Особистий внесок здобувача: вивчення особливостей морфології і продукції антибіотиків в ΔbldA мутанта, участь в написанні рукопису та підготовці статті до друку*).

23. **Ющук ОС,** Осташ БО, Горбаль ЛО, Федоренко ВО (2017). Реконструкція філогенії кластерів генів біосинтезу глікопептидів. Фактори експериментальної еволюції організмів 20:109-115. [**фахове видання України**] (*Особистий внесок здобувача: концептуалізація, біоінформатичний аналіз, написання рукопису та підготовка статті до друку*).

24. Tsypik O, **Yushchuk O,** Zaburannyi N, Flärdh K, Walker S, Fedorenko V, Ostash B (2016) Transcriptional regulators of GntR family in *Streptomyces coelicolor* A3(2): analysis *in silico* and *in vivo* of YtrA subfamily. Folia microbiologica 61(3):209–220. https://doi.org/10.1007/s12223-015-0426-7 **[IF 2,629; SCOPUS, WoS; SCImago Q3 – 2016 Microbiology]** (*Особистий внесок здобувача: вивчення розповсюдження генів YtrA-регуляторів серед актинобактерій, аналіз властивостей генів SCO0823 і SCO3812, участь в написанні рукопису та підготовці статті до друку*).

Патенти:

1. Патент України на винахід №119472 Україна, МПК (2006.01)C12P 1/06 C12P 19/64 C12N 15/80 C12R 1/465 Спосіб активування продукції сполук з антибіотичною та протигрибною властивостями у *Streptomyces cyanogenus* S136 / **Ющук О.,** Осташ Б.; заявник і власник Львівський національний ун-т ім. І.

Франка. – № a201702642, заявл. 21 березня 2017 р., опубл. 25.06.2019, Бюл. № 12. [**патент України на винахід**] (*Особистий внесок здобувача: виконання експериментів, написання і підготовка тексту патенту*).

Публікації, що засвідчують апробацію матеріалів дисертації:

1. Stepanyshyn A., **Yushchuk O.,** Marinelli F., Kalinowski J., Fedorenko V. Expression of pathway-specific regulators from dalbaheptide biosynthetic gene clusters induces antimicrobial properties of *Amycolatopsis bartoniae* DSM 45807. XVIII international scientific conference for students and PhD students dedicated to the 195th anniversary from the birthday of Julius Planer "Youth and progress of biology": abstracts, Lviv, 6–7 September 2022. Львів: Сполом, 2022. P. 72-73. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).
2. Duban R., Stepanyshyn A., Melnyk S., Fedorenko V., Ostash B., **Yushchuk O.** Approaching the activation of salinamide-like biosynthetic gene cluster in clorobiocin producer *Streptomyces roseochromogenes* NRRL 3504. IV International Scientific Conference: "Microbiology and Immunology – the Development": book of abstracts, Kyiv, 22-23 September 2022. Kyiv: Taras Shevchenko National University of Kyiv, 2022. P. 65. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).
3. Stepanyshyn A., Melnyk S., **Yushchuk O.,** Ostash I., Koshla O., Fedorenko V., Ostash B. Developing genetic instruments for manipulating *Streptomyces roseochromogenes* NRRL 3504 genome for increasing clorobiocin production. IV International Scientific Conference: "Microbiology and Immunology – the Development": book of abstracts, Kyiv, 22-23 September 2022. Kyiv: Taras Shevchenko National University of Kyiv, 2022. P. 81. (*Особистий внесок*

здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції).

4. Andreo-Vidal A., **Yushchuk O.,** Berini F., Binda E., Marinelli F. Novel glycopeptide antibiotic by *Nonomuraea coxensis* DSM 45129. Online congress ExpoBAC Zaragoza 2020: abstracts, Zaragoza, 8-9 July 2021. Zaragoza: Spain, 2021. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

5. Andreo-Vidal A., Binda E., Marinelli F., **Yushchuk O.** Comparative genomics of *van*-genes in glycopeptide producing *Actinobacteria* and beyond. VAAM Workshop 2021: Abstract book, Tübingen, 22-23 September 2021. Tübingen: Germany, 2021. P. 49. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач здійснив онлайн-доповідь*).

6. Степанишин А., Мельник С., Осташ Б., **Ющук О.** Аналіз мовчазного кластера генів біосинтезу криптичного салінамід-подібного антибіотика у продуцента хлоробіоцину *S. roseochromogenes*. XVII Міжнародна конференція студентів і аспірантів „Молодь і поступ біології”: тези доповідей, Львів, 27–29 квітня 2021 р. Львів, 2021. С. 114-115. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

7. **Yushchuk O.,** Ostash I., Melnyk S., Kalinowski J., Süssmuth R., Ostash B. Revealing cryptic secondary metabolome of *Streptomyces* via manipulation of global regulatory gene *adpA*. 3rd International conference on natural products discovery and development in the genomic era: abstracts, San Diego, 13 January 2020. https://sim.confex.com/sim/np2020/meetingapp.cgi/Paper/40268. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом*).

8. **Ющук О.,** Жукровська К.-О., Федоренко В. Сенсорний домен гістидинової кінази Tei3 не здатний реагувати на глікопептидні антибіотики. XV Міжнародна наукова конференція студентів і аспірантів «Молодь та поступ

біології», присвячена 135 річниці від дня народження Я. Парнаса: тези доповідей, Львів, 9-11 квітня 2019 р. Львів, 2019. С. 37. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції і зробив усну доповідь*).

9. **Ющук О.,** Власюк І., Осташ І., Федоренко В., Осташ Б. Властивості базальних представників ортологічної групи AdpA-регуляторів. XV Міжнародна наукова конференція студентів і аспірантів «Молодь та поступ біології», присвячена 135 річниці від дня народження Я. Парнаса: тези доповідей, Львів, 9-11 квітня 2019 р. Львів, 2019. С. 36. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

10. Homoniuk V., **Yushchuk O.,** Ostash B., Fedorenko V. Pathway-specific regulator of teicoplanin biosynthesis *tei15** does not interact with *tei2* promotor in heterologous system. XIV International scientific conference of students and PhD students “Youth and progress of biology”: abstracts, Lviv, 10-12 April 2018. Lviv, 2018. P. 128-129. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції і зробив усну доповідь*).

11. Vlasiuk I., **Yushchuk O.,** Ostash B., Fedorenko B. Landomycin A production of AdpA-overexpressing *S. cyanogenus* S136 strains in extremely depleted media. XIV International scientific conference of students and PhD students “Youth and progress of biology”: abstracts, Lviv, 10-12 April 2018. Lviv, 2018. P. 132. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

12. Vlasiuk I., Ostash B., **Yushchuk O.,** Ostash I., Kuzhyk Y., Melnyk S., Hrab P., Suessmuth R., Fedorenko V. Pleiotropic transcriptional factor AdpA as a tool to upregulate antibiotic production by *Streptomyces*. International Conference “Advances in microbiology and biotechnology”: abstract book, Lviv, 29-31 October

2018. Lviv, 2018. P. 44. (*Особистий внесок здобувача: тези і доповідь частково засновані на результатах отриманих здобувачем або під його керівництвом*).

13. Vlasiuk I., **Yushchuk O.** Heterologous AdpA activate cryptic lucensomycin biosynthetic gene cluster in *Streptomyces cyanogenus* S136. XIII international scientific conference of young scientists «Biology: from molecule to biosphere»: abstracts, Kharkiv, 28-30 November 2018. Kharkiv, 2018. P. 82-83. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

14. Жукровська К.-О., **Ющук О.,** Федоренко В. Дослідження індуцибельності сенсорної гістидинової кінази *tei3* з *Actinoplanes teichomyceticus.* XIV Міжнародна наукова конференція студентів і аспірантів «Молодь та поступ біології», присвячена 185 річниці від дня народження Б. Дибовського: тези доповідей, Львів, 10-12 квітня 2018. Львів, 2018. С. 133. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

15. Koshla O., **Yushchuk O.,** Ostash I, \ Kravets K., Suessmuth R., Jäger G., Bystrom A., Ostash B., Fedorenko V. Post-transcriptional modifications of tRNA modulate secondary metabolism of *Streptomyces*. International Conference “Advances in microbiology and biotechnology”: abstract book, Lviv, 29-31 October 2018. Lviv, 2018. P. 75. (*Особистий внесок здобувача: тези і доповідь частково засновані на результатах отриманих здобувачем або під його керівництвом*).

16. Vlasiuk I., **Yushchuk O.** Expression of different *adpA* genes increases landomycin A production in *Streptomyces cyanogenus* S136 // XII international scientific conference of young scientists «Biology: from molecule to biosphere»: abstracts, Kharkiv, 29 November – 1 December 2017. Kharkiv, 2017. P. 68. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції і зробив усну доповідь*).

17. Гомонюк В.В., **Ющук О.С.**, Горбаль Л.О. Експресія генів стійкості до глікопептидів у продуцента тейкопланіну *Actinoplanes teichomyceticus.* XII

Міжнародна конференція молодих науковців «Біологія: від молекули до біосфери»: тези доповідей, Харків, 29 листопада – 1 грудня 2017 р. Харків, 2017. С. 64. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

18. Zhukrovska K.-O., **Yushchuk O.** Expression of *tei3* gene encoding VanS-like sensor histidine kinase from teicoplanin biosynthesis gene cluster leads to teicoplanin resistance in *S. coelicolor* M512. XII international scientific conference of young scientists «Biology: from molecule to biosphere»: abstracts, Kharkiv, 29 November – 1 December 2017. Kharkiv, 2017. P. 69. (*Особистий внесок здобувача: тези і доповідь засновані на результатах отриманих здобувачем або під його керівництвом, здобувач написав тези конференції*).

19. **Yushchuk O.** Evolution of glycopeptide and moenomycin antibiotic biosynthesis pathways in bacteria. International conference “Integrative Biology & Medicine”: abstracts, Kyiv, 2-7 October, 2017. Kyiv, 2017. P. 75. (*Особистий внесок здобувача: тези і доповідь засновані на результатах здобувача, здобувач здійснив усну доповідь*).