

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 АВС-транспортерів. Основною функцією цих транспортерів є експортування пептидних антибіотиків із клітин продуцентів. КБГ і в інших актиномицетів-продуцентів антибіотиків також несуть гени АВС-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 АВС-транспортерів зі 102 КБГ ГПА та споріднених пептидних антибіотиків. Ми виявили, що АВС-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 АВС-транспортерів і мають N-термінальний трансмембранний домен із шістьма α -спіралями. Реконструкція філогенії цих АВС-транспортерів виявила низку клад, представники кожної з яких походять із КБГ ГПА специфічних типів. Філогенетична реконструкція АВС-транспортерів з КБГ ГПА та споріднених пептидних антибіотиків у контексті АВС-транспортерів, кодованих в КБГ інших груп сполук, підтвердила монофілетичне походження перших.

Ключові слова: кластери біосинтетичних генів, глікопептидні антибіотики, рамопланін, фегліміцин, АВС-транспортери

Глікопептидні антибіотики (далі ГПА) – це велика група природних сполук, що синтезуються актиномицетами, 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

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 ліпиду 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 ремодельовання клітинної стінки, в якій кінцевий D-Аланіл-D-Аланін бічного пептидного ланцюга ліпиду II замінюється 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 кодує 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 АВС-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 неродазоване множинне вирівнювання використано для встановлення оптимальної моделі амінокислотних замінів. 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 АВС-транспортерів використовували ТМНММ 2.0 [18], реконструкцію 2D-топології АВС-транспортера стосовно цитоплазматичної мембрани здійснювали за допомогою TMRPres2D [30].

Моделювання третинної та четвертинної структури АВС-транспортера Tei4.* Третинну структуру Tei4* моделювали на основі експериментально встановленої кристалічної структури АВС-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 синтениї генів білків-експортерів. У вибірці із 102 КБГ ГПА та споріднених пептидів було знайдено 95 кодованих АВС-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 АВС-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 АВС-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 АВС-транспортерів. Загалом, ці винятки не заперечують твердження, що кожен проаналізований КБГ ГПА та споріднених пептидних антибіотиків несе у своєму складі один ген АВС-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 АВС-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 АВС-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 АВС-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 АВС-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 синтениї розміщення гени АВС-транспортерів. Хоча в КБГ рамопланіну (DD002243.1) та черсинаміцину (WP_091321314.1), які також продукуються представниками порядку *Micromonosporales*, генне оточення генів АВС-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 АВС-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 АВС-транспортерів між геном StrR-подібного шлях-специфічного регулятора та геном перших модулів НРПС (рис. 1, д). Єдиним винятком у цій підмножині був КБГ VJHX01000001 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, е). Варто відзначити, що в КБГ CP011492 зі *Streptomyces* sp. CNQ-509 та CP025407 зі *Streptomyces* sp. SMB-StM0423 знайдено по два гени, що кодують ABC-транспортери. З кожної пари тільки один вписувався в загальну закономірність синтениї, 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. Виявилось, що всі вони належать до MdlB(MsbA)-подібних ABC транспортерів еукаріотичного типу підродиною ABC-B (COG1132) [15, 28]. Подібно до вже описаного Tba [24], ці білки поєднують N-термінальний трансмембранний і C-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 АKN86044.1 зі *Streptomyces* sp. CNQ-509 і AUN39259.1 *Streptomyces* sp. SMB-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-

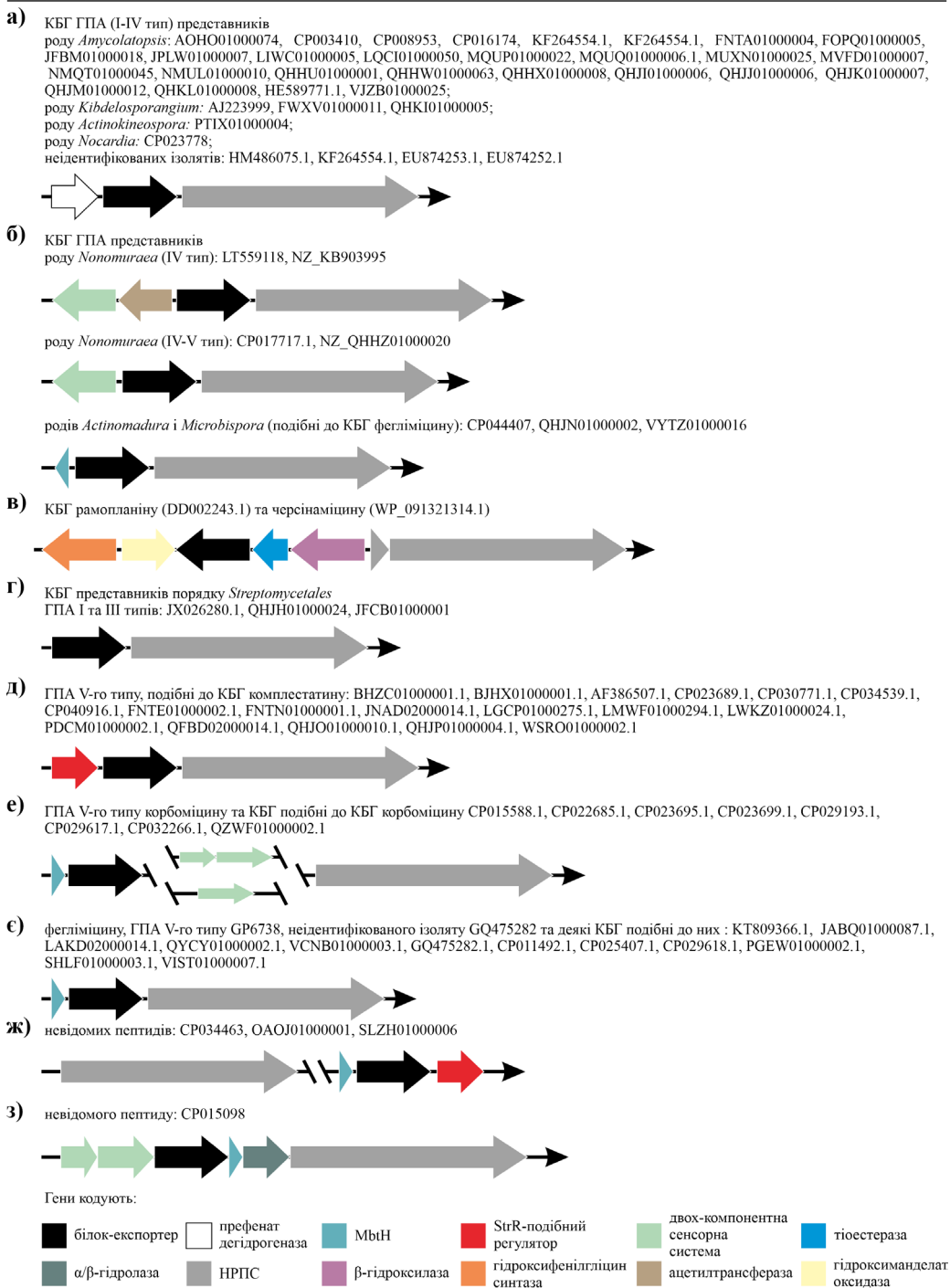


Рис. 1. Особливості розміщення генів ABC-транспортерів у різних кластерах генів біосинтезу глікопептидів і споріднених пептидних антибіотиків

ність кореляції між структурою і особливостями декорування ГПА та структурою АВС-транспортерів. Моделювання трансмембранних ділянок показало, що всі білки клади I мають шість трансмембранних α -спіралей, а С-термінальні АТФ-азні домени мають повний набір мотивів [33], необхідних для їхнього функціонування (рис. 3, а). Це, зокрема, мотив Волкера А (Р-петля), Q-петля, мотив Волкера В, D-петля, H-петля та сигнатурний мотив АВС-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 С-термінального домена. Отже, АВС-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 АВС-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 С-термінальні домени всіх 16 білків володіли набором характерних мотивів (рис. 3, б).

Варто відзначити, що АВС-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 АВС-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 – клади 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 АВС-транспортери клади V (рис. 2) кодовані в ряді різноманітних КБГ. По-перше, це були КБГ ліпидепсипептидів, 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 АВС-транспортери, закодовані в КБГ фегліміцину та GQ475282. До цієї клади також належали і АВС-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 АВС-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 АВС-транспортер, який кодовано в КБГ es-parp15 – 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.

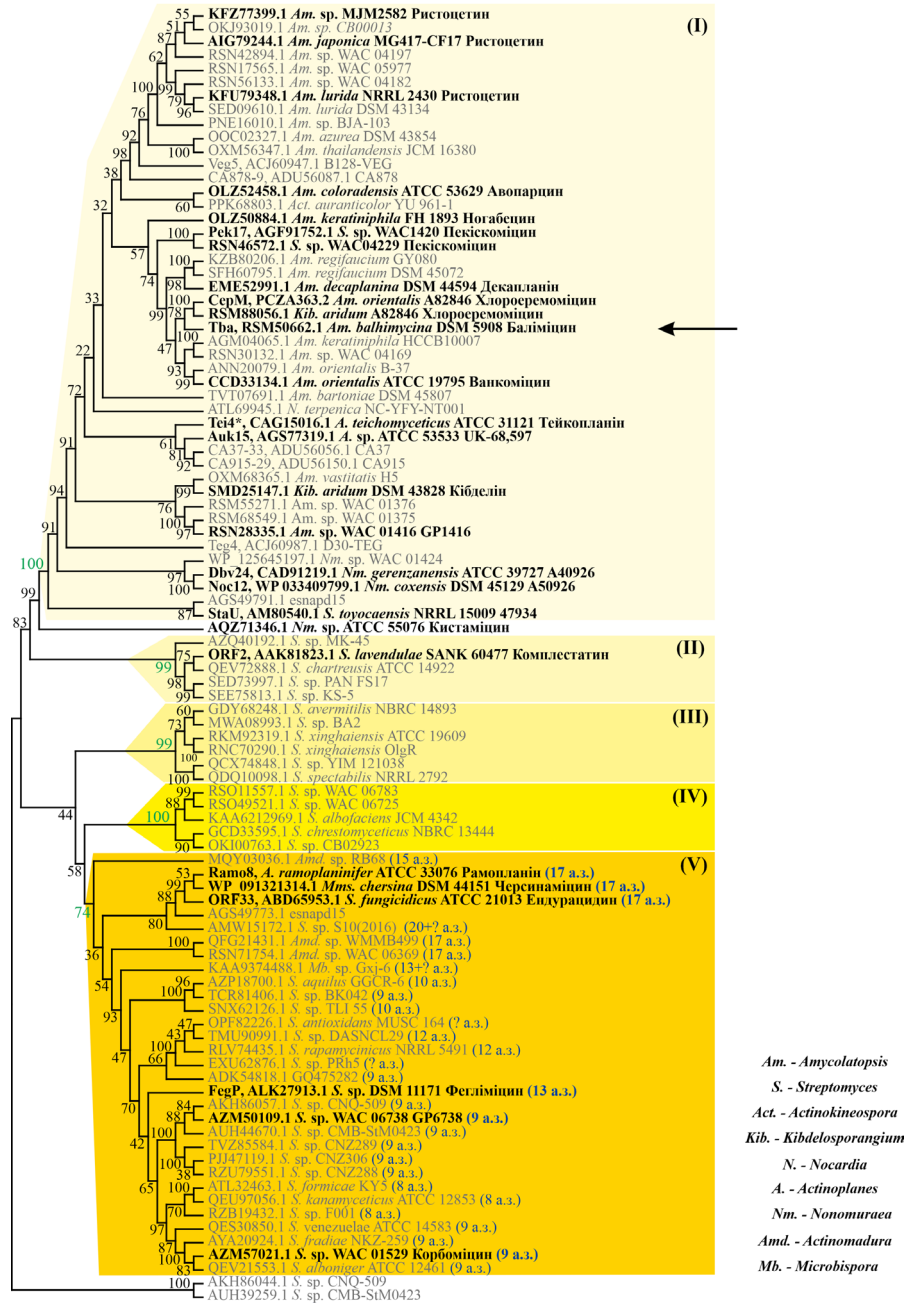


Рис. 2. Філогенетичне дерево (у вигляді кладограми) амінокислотних послідовностей АВС-транспортерів, що кодовані в КБГ ГПА та споріднених пептидів – вибірки № 1. На дереві було виділено п'ять основних клад із високою бутстреп-підтримкою; єдиний АВС-транспортер, вивчений експериментально – Tva з КБГ баліміцину – 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»

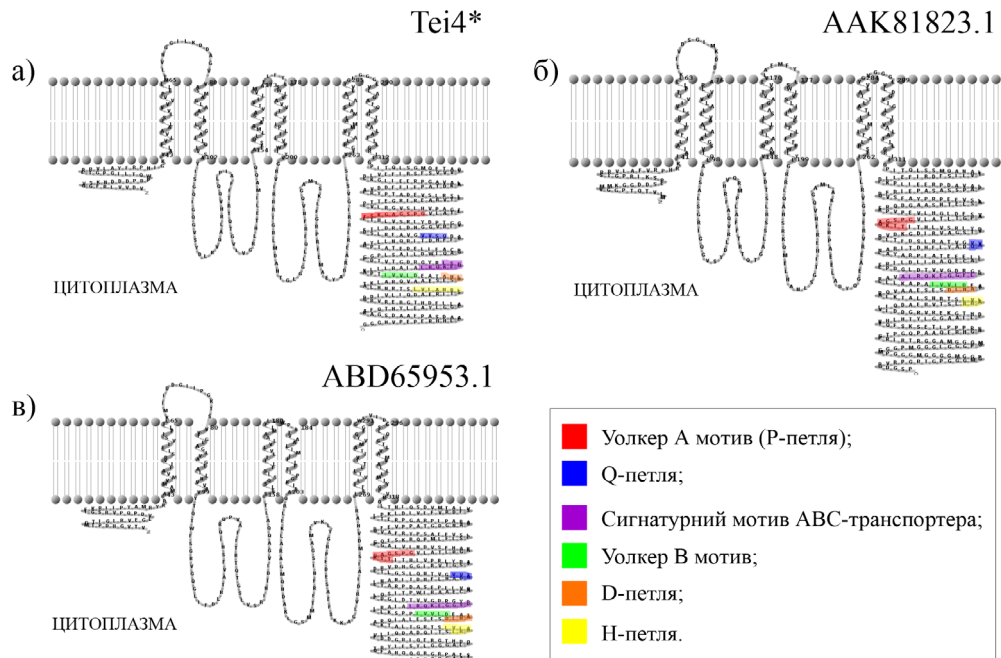


Рис. 3. Вторинна структура АВС-транспортёрів, що належать до класу I, II-IV та V на прикладах АВС-транспортёра, кодованого в КБГ тейкопланіну – Tei4* (а), комплестатину – ААК81823.1 (б) та ендурацидину – ABD65953.1 (в), відповідно. Передбачення трансмембранних α -спіралей і розміщення білків у клітинній 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-транспортёри, 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-транспортёри функціонують у вигляді гомодимерів. Щоб оцінити це припущення, ми спершу передбачили третинну структуру АВС-транспортёра, що закодований в КБГ тейкопланіну – Tei4*, а потім спробували змодельовувати 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-модельювання показало, що дві молекули Tei4* цілком імовірно можуть формувати димер (рис. 4, б).

Філогенія АВС-транспортёрів, закодованих в КБГ ГПА та споріднених пептидних антибіотиків у контексті АВС-транспортёрів інших природних сполук. Нарешті цікаво було дослідити, з якими АВС-транспортёрами, відповідальними за викачування інших груп природних сполук, найбільш споріднені АВС-транспортёри ГПА. Для цього було реконструйовано філогенію вибірки № 2 (див. розділ «Методи»). По-перше, виявилось (рис. 5), що всі АВС-транспортёри ГПА та споріднених пептидних антибіотиків належать до однієї класу дерева на основі вибірки № 2. Це свідчить про монофілію усіх цих АВС-транспортёрів – незважаючи на суттєві відмінності у структурі ГПА, феєліміцину

чи ліпидепсипептидів. Схоже, що всі їхні КБГ еволюціонували від одного предкового протокластера, який вже містив ген АВС-транспортера. По-друге, клади 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 ліпепептидним 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.

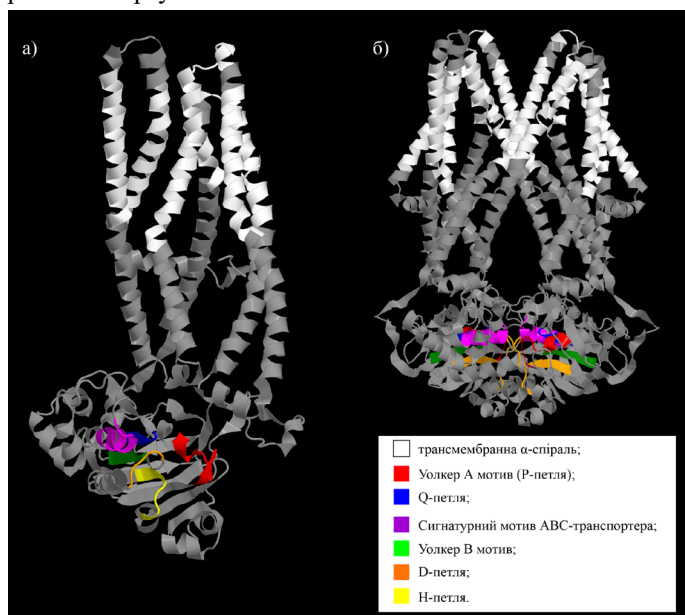


Рис. 4. Моделі третинної (а) та четвертинної (б) структури АВС-транспортера, закодованого в КБГ тейкопланіну – Tei4*. Моделі побудовано, як описано в розділі «Методи». Структурні елементи виділено кольорами

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. Виявилось, що АВС-транспортери з КБГ ГПА близько споріднені з АВС-транспортерами таких антибіотиків як фегліміцин і ліпидепсипептиди – 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

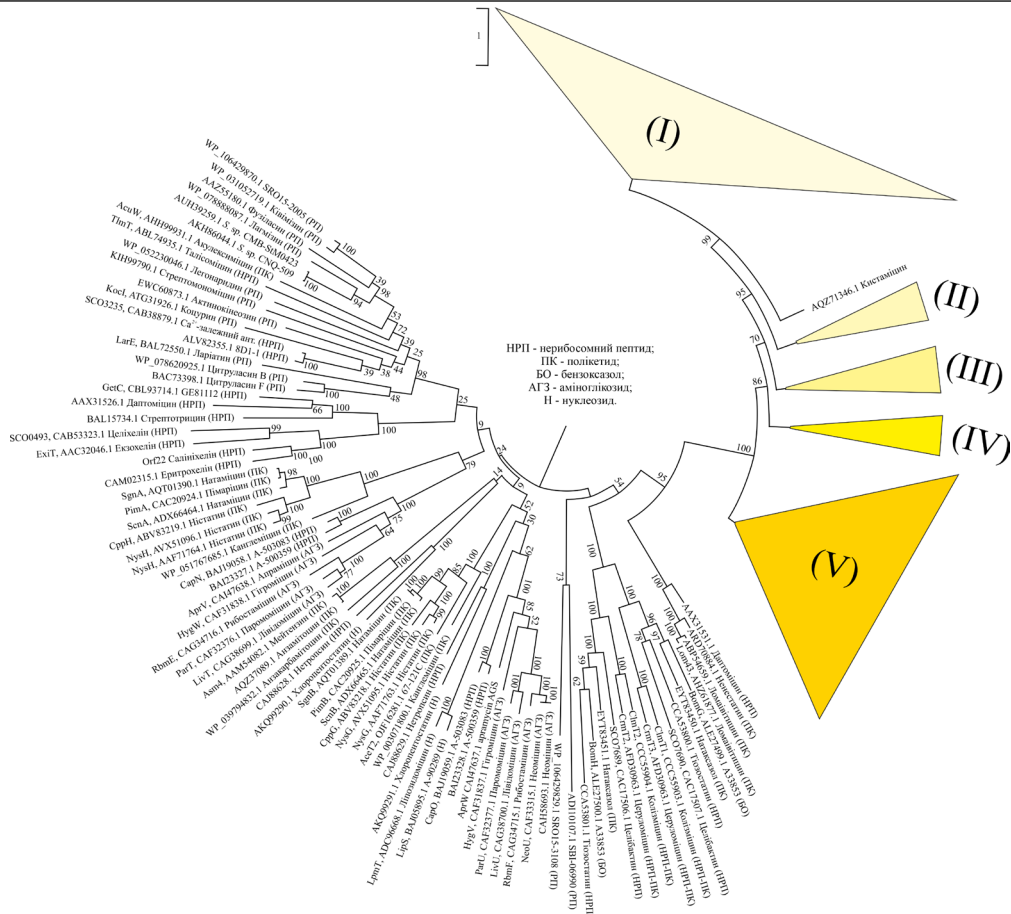


Рис. 5. Філогенетичне дерево, що репрезентує філогенетичні зв'язки між білками ABC-транспортерами вибірок № 1 та № 2. Помітно, що всі білки ABC-транспортерів ГПА та споріднених антибіотиків мають монофілетичне походження, формуючи одну кладу. Білки ABC-транспортерів, що закодовані в КБГ даптоміцину, 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 ABC-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 ABC-транспортери, їхні послідовності дуже консервативні. Це свідчить про те, що ці ABC-транспортери універсальні та здатні експортувати різні пептидні антибіотики незалежно від їхньої структури. Усі ABC-транспортери з КБГ ГПА і споріднених пептидних антибіотиків мають монофілетичне походження. Результати, отримані в роботі, дають змогу краще зрозуміти особливості ABC-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 non-ribosomal 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/REFERENCES
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, Hrushevskyyi 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 MdB(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