

ПОЛІЩУК Л. В.

Інститут мікробіології і вірусології ім. Д. К. Заболотного НАН України,  
Україна, 03143, м. Київ, вул. Акад. Заболотного, 154, ORCID: 0000-0002-3159-5022,  
e-mail: LVPolishchuk@ukr.net

### БІОІНФОРМАТИЧНЕ ДОСЛІДЖЕННЯ КЕРАТИНАЗ ІЗ РОДИНИ S01 ШТАМУ *STREPTOMYCES MICROFLAVUS* DSM 40593

**Мета.** Визначення та дослідження первинних структур та доменної організації проферментів кератиназ (серинових протеаз з родини S01) *S. microflavus* DSM 40593 допомогою біоінформативних методів. **Методи.** У дослідженнях використана інформація, яка представлена у базах даних на сервері NCBI. Комп'ютеризований аналіз послідовностей проводили за допомогою програми BLAST на сервері NCBI. **Результати.** Проведено BLAST-аналіз послідовності геному штаму *S. microflavus* DSM 40593 з використанням в якості запиту послідовності проферменту серинової протеази AGK81825. Завдяки біоінформативному аналізу в геномі штаму *S. microflavus* DSM 40593 було виявлено 7 генів, що кодують проферменти серинових протеаз з родини S01. **Висновки.** У геномах стрептоміцетів може бути присутніми більш ніж один ген, що кодують проферменти кератиназ з родин S01. Встановлено, що амінокислотні послідовності ферментативних доменів проферментів кератиназ з родини S01 характеризуються більшими показниками подібності ніж послідовності доменів фолдаз.

**Ключові слова:** серинова протеаза, профермент, домен, стрептоміцет, подібність послідовностей.

Кератин є основною складовою різноманітних відходів сільського господарства та харчової промисловості, яка потрапляють у середовище у вигляді пір'я, рафотек дзьобів, волосся, шерсті, щетини та копит [1]. Всі кератини виробляються в епітеліальних клітинах вищих хребетних (рептилій, птахів і ссавців) і у людини [2].

Кератини характеризуються надзвичайною міцністю до дії фізичних, хімічних та біологічних факторів, що забезпечується структурою волокнистих білків. Встановлено, що лише деякі бактерії, актиноміцети та кератинофільні гриби, а з макроорганізмів личинки звичайної платтяної молі здатні використовувати кератин

як єдине джерело вуглецю, азоту, сірки та енергії [3]. Кератинолітичні види актиноміцетів, які інтенсивно руйнують природний кератин пір'я, волосся, нігтів і рогів, представлені в основному родом *Streptomyces*, включаючи *Streptomyces fradiae*, *Streptomyces pactum*, *Streptomyces thermoviolaceus* або інші роди, наприклад *Thermoactinomyces* [3, 4].

Була висунута гіпотеза, що деградації кератину під дією мікробних кератиназ – це двостадійний процес (кератинолізис). Перший етап – сульфолізис (руйнування дисульфідних зв'язків під дією дисульфідних редуктаз), другий етап – протеоліз (гідроліз пептидних зв'язків протеаза, а саме кератиназами). Кератинолітичні серинові протеази є в основному позаклітинними ферментами, синтез яких індукується присутністю субстрата у середовищі [2].

Більшість кератиназ, відомих на сьогоднішній день, були класифіковані як серинові протеази, а деякі як металопротеази. Аналіз, заснований на класифікації MEROPS, вказує на те, що кератинази *Bacillus* належать до родини серинових протеаз S08, а кератинази з *Actinobacteria*, наприклад *Nocardiosis*, *Actinomadura* і *Streptomyces*, належать переважно до родини S01 [3].

Переважно свої зусилля дослідники витрачають на пошук нових ензимів з стрептоміцетів, здатних ферментувати кератин та умов культивування їх продуцентів; виділенню виявлених кератиназ, організації, фізико-хімічним властивостям та оптимізації умов ферментативних процесів; можливості використання ензимів у різних галузях промисловості [1, 5–9]. Первинні та надвторинні (доменні) структури молекул визначені для незначної кількості проферментів кератиназ. У джерелах літератури повідомлено про доменну організації молекул ряду проферментів стрептоміцетів, що детермінують кератинази: MG765531, KY368946, MBT3158181, SAN05006, AGK81825 та встановлено, що всі вони складаються з сигнальної

© ПОЛІЩУК Л. В.

послідовності та 2 доменів: про-домену (фолдазу) та ферментативного – кератинази (рис. 1). Так, зимогени кератиназ S01 (MG765531, KY368946, MBT3158181, САН05006, AGK81825) містять домени CDD:427092 (фолдаза) та CDD:411050 (кератиназа), які відносять відповідно до надродин c108385 і c142389. Схема доменної організації молекули серинових протеаз родин S01 представлена на рис. 1.

Однією з найбільш різнобічно вивчених кератиназ стрептоміцетів є серинова протеаза 2 з родини S01 САН05008.1 (307 а. к.) стрептоміцету *S. fradiae* var. K11. Встановлена амінокислотна послідовність молекули проферменту, його доменна та третинна структури, механізм утворення каталітичноактивного ферменту та нуклеотидна послідовність, що кодує протеїн зимогену (AJ784940, GenBank) [10].

У джерелах літератури повідомляється про визначення сиквенсів тільки незначної кількості детермінант, що визначають структури кератиназ стрептоміцетів [3, 10, 11].

### Матеріали і методи

У дослідженнях використана інформація щодо нуклеотидних та амінокислотних послідовностей, яка представлена у базах даних на сервері NCBI (The National Center for Biotechnology Information) [https://blast.ncbi.nlm.nih.gov] (табл. 1).

Також у дослідженнях використовували послідовність геному штаму *S. microflavus* DSM 40593 (NC\_021177.1 – 7905758 п. н.).

Комп'ютеризований аналіз послідовностей стрептоміцетів проводили за допомогою програми BLASTN [www.ncbi.nlm.nih.gov/blast] на сервері NCBI.

### Результати та обговорення

Значна увага приділяється дослідниками практичним аспектам вивчення кератиназ. Дослідження спрямовані на розширення можливостей обробки кератинвмісної сировини з метою забезпечення безпека та екологічність виробництв шляхом створення безвідходних і мало-відходних технологій з максимальним залученням побічних продуктів переробки в основному виробництві [12, 13]. Однак, останнім часом проводяться успішне біоінформативне передбачення та дослідження генів, що кодують вірогідні кератинази та їх продуктів [3].

Цікавим було визначити ступінь подібності послідовностей як проферментів, так і окремих доменів. Як запит (Query) використовували послідовність САН05006. BLASTN-аналізом встановлено, що послідовності, які кодують кератинази мають більший ступінь подібності (табл. 2).



Рис. 1. Схема доменної організації проферменту серинової протеази САН05008 з родини S01. Позначення: 1 – послідовність проферменту, 2 – домен серинової протеази (кератинази), 3 – домен фолдази, 4 – амінокислотні залишки, що утворюють активний центр кератинази.

Таблиця 1. Проферменти кератиназ з родин S01 актиноміцетів та гени, що їх детермінують.

№ доступу амінокислотної послідовності, Genbank	№ доступу нуклеотидної послідовності, Genbank	Штам – джерело проферменту
AYM48028	MG765531	<i>S. albidoflavus</i> TBG-S13A5
AQX39246	KY3668946	<i>S. albidoflavus</i> Fea-10
MBT3158181	G11C_01512	<i>Streptomyces</i> sp. G11C
САН05008	AJ784940	<i>S. fradiae</i> var. K11
AGK81825	sprA	<i>S. microflavus</i> DSM 40593
AAO06113	AY151208	<i>Nokardiopsis</i> sp. TOA-1
AMH86070	KU550701	<i>Actinomadura viridilutea</i> DZ50
ASU91959	KY914477	<i>Actinomadura karatinilytica</i> Cpt20

Таблиця 2. Подібність послідовностей проферментів серинових протеаз актиноміцетів з родини S01 послідовності запиту (САН05006)

Проферменти актиноміцетів	Показники подібності амінокислотних послідовностей до послідовності запиту		
	профермента	домена фолдази	домена кератинази
МВТ3158181	Qc=85 % I=74 % P=83 %	Qc=98 % I=61 % P=77 %	Qc=100 % I=80 % P=86 %
АГК81825	Qc=85 % I=96 % P=78 %	Qc=98 % I=55 % P=71 %	Qc=100 % I=74 % P=81 %
АУМ48028	Qc=80 % I=60 % P=72 %	Qc=82 % I=52 % P=61 %	Qc=99 % I=65 % P=77 %
АQX39246	Qc=80 % I=60 % P=72 %	Qc=82 % I=52 % P=61 %	Qc=99 % I=65 % P=77 %
АМН86070	Qc=81 % I=38 % P=53 %	В	Qc=93 % I=43 % P=59 %
ASU91959	Qc=77 % I=40 % P=53 %	В	Qc=86 % I=45 % P=59 %
ААО06113	Qc=81 % I=37 % P=52 %	В	Qc=93 % I=43 % P=60 %

Примітки: В – низькі показники подібності послідовності структури запиту. Qc (query coverage) – покриття запитом, I (identities) – ідентичні, P (positives) – кількості амінокислот, які або однакові між запитом і послідовністю суб'єкта або мають подібні хімічні властивості.

Повідомлено про спорідненість виявлених серинових протеаз S01 актиноміцетів (*Nocardia opsis* sp. TOA-1 – ААО06113 (AY151208, 1557 п.н.), *Actinomadura viridilutea* DZ50 – АМН86070 (KU550701, 1557 п.н.), *Actinomadura karatinilytica* Cpt20 – ASU91959 (KY914477, 1557 п. н.), *S. fradiae* var. k11 – САН05008 [3]. Аналізом баз даних NCBI встановлено, що молекули вказаних проферментів організовані за однією схемою (рис. 1). Крім того, послідовності проферментів 3 актиноміцетів мають достовірну подібність послідовності стрептоміцетного зариту, однак показники їх подібності менші ніж у послідовностей стрептоміцетів (табл. 2). Також треба відмітити, що послідовності доменів кератиназ мають більший ступінь подібності, ніж послідовності доменів фолдаз. Крім того, первинна структура домену фолдаз трьох актиноміцетів значно відрізнялися від послідовності запиту.

Відомо, що актиноміцети, особливо стрептоміцети, продукують низку різних протеаз у живильне середовище [14]. Так повідомляється, що штам *S. albidoflavus* K1-02 продукує 6 кератинолітичних протеаз [14], у той час, як *Streptomyces* sp. G11C, крім кератинази МВТ3158181, вірогідно здатен синтезувати ще понад 10 серинових протеаз-кератиназ 2 [3].

На даний час, зі штамів продуцентів кератиназ *S. albidoflavus* TBG-S13A5, *S. albidoflavus*

Fea-10, *S. fradiae* var. K11, *S. microflavus* DSM 40593 тільки нуклеотидна послідовність хромосоми штаму *S. microflavus* DSM 40593 (NC\_021177.1 – 7905758 п. н.) визначена у повному об'ємі (complete sequence). Було вирішено визначити присутність у геномі цього штаму інших генів, що можуть детермінувати кератинази та визначити їх подібність з послідовністю проензиму АГК81825.

BLASTN-аналізом (tblastn) з використанням у якості запиту послідовність проферменту АГК81825 було виявлено в геномі штаму *S. microflavus* DSM 40593 сім генів, послідовності яких подібні послідовності запиту (табл. 3). Всі сім детермінованих ними проферментів складаються з доменів CDD:427092 (фолдаза) та CDD:411050 (кератиназа), які відносяться відповідно до надродин c108385 і c142389. Схема доменної організації молекул цих серинових протеаз родин S01 представлена на рис. 1.

Було визначено ступінь подібності послідовностей як проферментів, так і окремих доменів семи виявлених зимогенів. Як запит (Query) використовували послідовність АГК81825. BLASTN-аналізом послідовностей встановлено, що послідовності, які кодують кератинази мають більший ступінь подібності (табл. 4).

Таблиця 3. Гени *S. microflavus* DSM 40593 та зимогени серинових протеаз з родини S01, які вони кодують

Гени, що кодують проферменти	Проферменти	Гени, що кодують проферменти	Проферменти
SFUL_RS34270	WP_015613116	SFUL_RS15325	WP_015609388
SFUL_RS06315	WP_180288864	SFUL_RS32965	WP_015612852
SFUL_RS06350	WP_015607634	SFUL_RS13990	WP_015609132
SFUL_RS26545	WP_015611582		

Таблиця 4. Подібність послідовностей проферментів протеаз з родини S01 *S. microflavus* DSM 40593 послідовності запиту (AGK81825)

Зимогени кера- тиназ	Показники подібності амінокислотних послідовностей		
	проферментів	домени фолдаз	доменів кератиназ
WP_015613116	Qc=100 % I=100 %	Qc=100 % I=100 %	Qc=100 % I=100 %
WP_180288864	Qc=80 % I=65 % P=75 %	Qc=60 % I=52 % P=63 %	Qc=100 % I=65 % P=73 %
WP_015607634	Qc=99 % I=61 % P=72 %	Qc=100 % I=58 % P=76 %	Qc=99 % I=62 % P=75 %
WP_015611582	Qc=85 % I=47 % P=66 %	Qc=87 % I=39 % P=67 %	Qc=100 % I=52 % P=69 %
WP_015609388	Qc=77 % I=38 % P=52 %	Qc=56 % I=35 % P=45 %	Qc=96 % I=48 % P=64 %
WP_015612852	Qc=85 % I=39 % P=53 %	Qc=65 % I=38 % P=58 %	Qc=95 % I=44 % P=57 %
WP_015609132	Qc=81 % I=31 % P=47 %	B	Qc=57 % I=40 % P=60 %

Примітки: B – низькі показники подібності послідовності структури запиту. Qc (query coverage) – покриття запитом, I (identities) - ідентичні, P (positives) – кількості амінокислот, які або однакові між запитом і послідовністю суб'єкта або мають подібні хімічні властивості

Показано, що послідовності доменів кератиназ 6 проферментів штаму *S. microflavus* DSM 40593 більш подібні до відповідної послідовності запиту, ніж загальні послідовності проферментів. Найбільш відмінною виявилась послідовність проферменту WP\_015609132. Цей результат підтверджено з використанням в якості запитів й інших послідовностей, зокрема WP\_180288864 та WP\_015609132 – результати не представлено).

Використавши інформацію баз даних NCBI та KEGG встановили, що проферменти організовані за схемою, характерною іншим проферментам з кератиназною активністю (рис. 1). Лише виявлений профермент WP\_015609388 містить додатковий 3 домен – домен зв'язування з глікозидами (ChiA1\_BD, CDD:213177 cl00046).

Каталітична активність ферменту забезпечується незначною частиною ферменту – активним центром. Наприклад, кератиназа, утворена з проферменту WP\_015613116 складається з 183 амінокислотних залишків, а її активний центр – це залишки 3 амінокислот: гістидина, аспарагіна та серина.

Було визначено за допомогою інструменту BLASTP (Multiple alignment) наявність і розташування відмінностей у первинній структурі проферментів кератиназ S01 штаму *S. microflavus* DSM 40593. Встановлено, що, не дивлячись на значну відмінність послідовностей доменів кератиназ, амінокислотні залишки, що утворюють ферментативний центр у всіх проферментів з родини S01, збережені (рис. 2).

### Висновки

У геномі штаму *S. microflavus* DSM присутні більш, ніж по одному гену, що кодують проферменти кератиназ з родини S01 (SFUL\_RS34270, SFUL\_RS06315, SFUL\_RS06350, SFUL\_RS26545, SFUL\_RS15325, SFUL\_RS32965, SFUL\_RS13990). Встановлено, що амінокислотні послідовності ферментативних доменів проферментів кератиназ з родини S01 штаму *S. microflavus* DSM 40593 характеризуються більшими показниками подібності, ніж загальні послідовності проферментів.

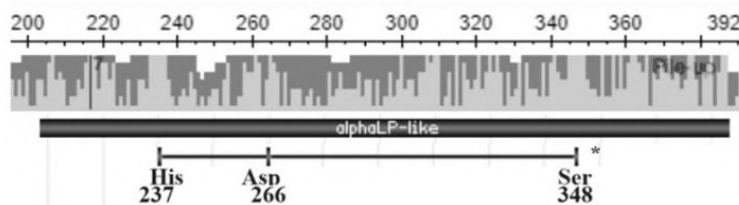


Рис. 2. Розташування відмінностей послідовностей 7 кератиназ штаму *S. microflavus* DSM 40593 виявлене за допомогою інструменту BLASTP (Multiple alignment). Як запит використано послідовність проферменту WP\_180288864. \* – амінокислотні залишки, що утворюють каталітичний центр.

## References

- Hassan M. A., Abol-Fotouh D., Omer A. M., Tamer T. M., Abbas E. Comprehensive insights into microbial keratinases and their implication in various biotechnological and industrial sectors: A review. *International Journal of Biological Macromolecules*. 2020. Vol. 154. P. 567–583. doi: [10.1016/j.ijbiomac.2020.03.116](https://doi.org/10.1016/j.ijbiomac.2020.03.116).
- Korniłowicz-Kowalska T., Bohacz J. Biodegradation of keratin waste: Theory and practical aspects. *Waste management*. 2011. Vol. 31 (8). P. 1689–1701. doi: [10.1016/j.wasman.2011.03.024](https://doi.org/10.1016/j.wasman.2011.03.024).
- Valencia R., González V., Undabarrena A., Zamora-Leiva L., Ugalde J. A., Cámara B. An integrative bioinformatic analysis for keratinase detection in marine-derived *Streptomyces*. *Marine Drugs*. 2021. Vol. 19 (6). P. 286. doi: [10.3390/md19060286](https://doi.org/10.3390/md19060286).
- Qiu J., Barrett K., Wilkens C., Meyer A. S. Bioinformatics based discovery of new keratinases in protease family M36. *New biotechnology*. 2020. Vol. 68. P. 19–27. doi: [10.1016/j.nbt.2022.01.004](https://doi.org/10.1016/j.nbt.2022.01.004)
- Li Q. Structure, application, and biochemistry of microbial keratinases. *Frontiers Microbiologie*. 2021. Vol. 12. e674345. doi: [10.3389/fmicb.2021.674345](https://doi.org/10.3389/fmicb.2021.674345).
- Li Q. Progress in microbial degradation of feather waste. *Frontiers microbiology*. 2019. Vol. 10. e2717. doi: [10.3389/fmicb.2019.02717](https://doi.org/10.3389/fmicb.2019.02717).
- Vidmar B., Vodovnik M. Microbial keratinases: enzymes with promising biotechnological applications. *Food technology and biotechnology*. 2018. Vol. 56 (3). P. 312–328. doi: [10.17113/ftb.56.03.18.5658](https://doi.org/10.17113/ftb.56.03.18.5658).
- Qiu J., Wilkens C., Barrett K., Meyer A.S. Microbial enzymes catalyzing keratin degradation: classification, structure, function. *Biotechnology advances*. 2020. Vol. 44. e107607. doi: [10.1016/j.biotechadv.2020.107607](https://doi.org/10.1016/j.biotechadv.2020.107607).
- Dmitrieva A. I., Alexeyenko A., Belova D., Piskaeva N., Tereshchuk L. *Streptomyces* and *Bacillus* keratinases: properties and uses. *Food processing techniques and technology*. 2020. Vol. 50 (4). P. 602–615. doi: [10.21603/2074-9414-2020-4-602-615](https://doi.org/10.21603/2074-9414-2020-4-602-615).
- Wu B., Shi P., Li J., Wang Y., Meng K., Bai Y., Luo H., Yang P., Zhou Z., Yao B. A new aminopeptidase from the keratin-degrading strain *Streptomyces fradiae* var. k11. *Applied biochemistry biotechnology*. 2010. Vol. 160. P. 730–739. doi: [10.1007/s12010-009-8537-8](https://doi.org/10.1007/s12010-009-8537-8)
- Ma Y., Ke X., Li X., Shu W., Yang W., Liu Y., Yan X., Jia L., Yan H. Expression and characterization of a keratinase encoding gene gm2886 in *Streptomyces pactum* ACT12 strain. *Chinese journal of biotechnology*. 2017. Vol. 33(12). P. 1968–1978. doi: [10.13345/j.cjb.170081](https://doi.org/10.13345/j.cjb.170081) [in Chinese]
- Avdiyuk K. V., Roy A. O. Selection of optimal conditions for cultivation of *Bacillus megaterium* UCM B-5710 – producer of keratinase. *Mikrobiolohichnyi zhurnal*. 2021. Vol. 83 (6). P. 32–40.
- Tatineni R., Doddapaneni K. K., Potumarthi R. C., Mangamoori L. N. Optimization of keratinase production and enzyme activity using response surface methodology with *Streptomyces* sp. 7. *Applied Biochemistry and Biotechnology*. 2007. Vol. 141 (2–3). P. 187–201. doi: [10.1007/BF02729061](https://doi.org/10.1007/BF02729061).
- Bressollier P., Letourneau F., Urdaci M., Verneuil B. Purification and characterization of a keratinolytic serine proteinase from *Streptomyces albidoflavus*. *Applied environmental microbiology*. 1999. Vol. 65 (6). P. 2570–2576. doi: [10.1128/AEM.65.6.2570-2576.1999](https://doi.org/10.1128/AEM.65.6.2570-2576.1999).

## POLISHCHUK L.V.

Zabolotny' Institute of Microbiology and Virology of the National Academy of Sciences of Ukraine, Ukraine, 03143, Kyiv, Acad. Zabolotny str., 154

## BIOINFORMATIC STUDY OF KERATINASES FROM THE S01 FAMILY FROM *STREPTOMYCES MICROFLAVUS* STRAIN DSM 40593

**Aim.** Determination and study of the primary structures and domain organization of keratinase precursors (serine proteases from the S01 family) of *S. microflavus* DSM 40593 using bioinformatic methods. **Methods.** In the research information presented in databases on the NCBI server was analyzed. Computerized sequence analysis was performed using the BLAST program on the NCBI server. **Results.** BLAST analysis of the genome sequence of the strain *S. microflavus* DSM 40593 was carried out using the sequence of the precursor AGK81825 as a query. Thanks to bioinformatic analysis, seven genes encoding precursors of serine proteases from the S01 family were found in the genome of the strain *S. microflavus* DSM 40593. **Conclusions.** In the genomes of streptomycetes, there may be more than one gene encoding precursors of keratinases from the S01 family. It was established that the amino acid sequences of the enzymatic domains of keratinase precursors from the S01 family are characterized by higher similarity indices than the sequences of foldase domains.

**Keywords:** serine protease, precursor, domain, streptomycete, sequence similarity.