

СЛІЩУК Г. І.^{1,2}, МЕЩЕРЯКОВ В. І.², ВОЛКОВА Н. Е.¹, ЗАХАРОВА О. А.¹¹ ТОВ «КОТЕКНА УКРАЇНА ЛІМІТЕД»,

Україна, 65114, м. Одеса, вул. Люстдорфська дорога, 140 а, ORCID: 0000-0003-4245-8557, 0000-0002-9333-4872, 0000-0003-1828-9044

² Одеський державний екологічний університет,

Україна, 65016, м. Одеса, вул. Львівська, 15, ORCID: 0000-0003-0499-827X

✉ george.slishchuk@cotecna.com

БІОІНФОРМАТИЧНЕ ДОСЛІДЖЕННЯ ПРОТЕЇНУ СYP719A14 АРГЕМОНИ МЕКСИКАНСЬКОЇ (*ARGEMONE MEXICANA* L.)

Мета. Дослідження структури протеїну СYP719A14 і взаємодії з лігандами біоінформатичними методами. **Методи.** Гомологічне моделювання тримірної структури ензиму. Молекулярний докінг. **Результати.** За результатами гомологічного моделювання побудовано модель ензиму СYP719A14, яка проявляє хейліантіфолін синтазну активність. Методом молекулярного докінгу досліджено взаємодію побудованої моделі з ймовірними субстратами-похідними бензилізохиноліну: (S)-Скоулеріном, (S)-Хейліантіфоліном, (S)-Тетрагідраколумбаїном, (S)-Ретикуліном, побудовані моделі і розраховані термодинамічні характеристики взаємодії. Ідентифіковано потенційний канал доступу ліганда – потенційний активний центр ензиму. **Висновки.** Ідентифіковано потенційний канал доступу ліганду ензиму СYP719A14, яким є порожнина, що формується амінокислотними залишками PHE57-GLY69, ALA285-ARG310, THR369-HSD390, що дозволить використовувати отримані дані у дослідженнях інженерінгу алкалоїдів з метою удосконалення алкалоїдного профілю аргемони мексиканської.

Ключові слова: аргемона, СYP719A14, специфічність субстрату, структура.

Аргемона мексиканська (*Argemone mexicana* L.), відома також як «мексиканський колючий мак», cardosanto або chicalote – вид рослини родини Макові (*Papaveraceae*). Він походить із тропічної Америки, зокрема із Мексики, але зараз поширений у тропічних і субтропічних регіонах світу. Вид є однорічною, прямостоячою трав'янистою рослиною заввишки біля 100 см, вкритою колючками. Листя блакитно-зелене із більш-менш великими білими та зеленими зонами, вкрите колючками. Квіти одиночні, жовті, діаметром 4–5 см, без аромату. Плід є капсулою, вкритою колючками,

овальної форми довжиною біля 3 см. Насіння сферичне, чорне та блискуче [1].

Аргемона мексиканська широко використовується як лікарська рослина. Відомі її антибактеріальні, антидіабетичні, антиоксидантні та гепатопротекторні властивості. Вона є перспективною для дослідження вторинних метаболітів як потенційно лікарських субстанцій [2]. Ідентифіковано більше ніж 20 алкалоїдів, серед яких особливе значення мають похідні бензилізохиноліну, зокрема берберин, протоберберин, сангвінарін [3, 4].

Охарактеризовано численні ензими, які залучені у біосинтезі похідних бензилізохиноліну, зокрема СYP719A14, що проявляє хейліантіфолінсинтазну активність і формує метилендіоксовий міст [5].

Інженерінг алкалоїдів є новим та перспективним напрямом досліджень, метою якого є вивчення структури та функцій ензимів, залучених у біосинтезі алкалоїдів, з метою підвищення біосинтезу цільового алкалоїду або створення нового для майбутніх фармакологічних розробок [6]. Дослідження структури та функції цитохромів є актуальним та важливим, особливо аналіз взаємодії ензиму із лігандом для ідентифікації каналів доступу до лігандів, що, у свою чергу, можуть грати критичну роль у модуляції активності ензиму [7].

Мета наших досліджень полягала у дослідженні структури протеїну СYP719A14 і взаємодії з лігандами біоінформатичними методами.

Матеріали і методи

Матеріалом слугувала нуклеотидна послідовність мРНК гена *CYP719A14* аргемони мексиканської, представлена у базі даних Національного центру біотехнологічної інформації (National Center for Biotechnology Information, NCBI), номер доступу EF451152.1 [8]. Для пошуку тривимірних структур лігандів використо-

ували базу даних PubChem [9–12].

Гомологічне моделювання тримірної структури ензиму CYP719A14, що проявляє хейліантіфолінсинтазну активність виконували за допомогою серверу SWISS-MODEL [13–17]. Молекулярний докінг проводили за допомогою веб-сервісу SwissDock [18]. Візуалізацію результатів гомологічного моделювання протеїнів та докінгу проводили за допомогою UCSF Chimera [19].

Результати та обговорення

За результатами гомологічного моделювання побудовано модель ензиму CYP719A14, який проявляє хейліантіфолінсинтазну активність (рис. 1). Моделлю для гомологічного моделювання був протеїн 6b82.1.B Cytochrome P450, family 17, subfamily A, polypeptide 1 Zebra Fish CYP-450 17A1 Mutant Abiraterone Complex. Показники якості моделі: стандартний показник точності реконструкції (Global Model Quality Estimate, GMQE) дорівнював 0,60, аналіз якості енергії моделі (Qualitative Model Energy Analysis, QMEANDisCo Global) – $0,63 \pm 0,05$, що відповідало критеріям прийнятності.

Побудовану методом гомологічного моделювання структуру ензиму CYP719A14 порівняли із структурою, реконструйованою за допомогою штучного інтелекту AlphaFold B1NF20. Структури були майже тотожні, відрзнялася суттєво лише конформація ділянки протеїну

MDETIWLIISTVHIVLGIKFLGKSSSSSLSTM EWPVGPCK (рис. 2, 3)

Враховуючи той факт, що для цитохромів є характерним наявність порожнини, що є кристою від розчинника, а також те, що активний центр протеїну розташований глибоко в корі протеїну, проводили ідентифікацію порожнини, що проявляє афінитет до потенційного субстрату та продукту – каналу доступу ліганда. Для ідентифікації порожнини, яка проявляє афінитет до потенційного субстрату та продукту – каналу доступу ліганда, проводили молекулярний докінг моделі ензиму із ймовірними субстратами-похідними бензилізохиноліну: (S)-Скоулеріном, (S)-Хейліантіфоліном, (S)-Тетрагідраколумбаїном, (S)-Ретикуліном. (S)-Скоулерін є субстратом для хейліантіфолінсинтазної активності, (S)-Хейліантіфолін є продуктом, а (S)-Тетрагідраколумбаїн та (S)-Ретикулін є субстратами споріднених ензимів.

Для (S)-Скоулеріну ідентифіковано 30 кластерів докінгу. Показник FullFitness варіював від -2526,65 до -2550,23 ккал/моль, вільна енергія Гібса ΔG варіювала від -7,88 до -5,92 ккал/моль. Більшість кластерів докінгу групувалася у порожнині, яка формувалася амінокислотними залишками PHE57-GLY69, ALA285-ARG310, THR369-HSD390, що, ймовірно, формують потенційний канал доступу ліганду (рис. 4, А). У потенційного каналу доступу лігандів показник FullFitness варіював від -2536,62 до -2550,23, вільна енергія Гібса ΔG варіювала від -7,88 до -6,95.

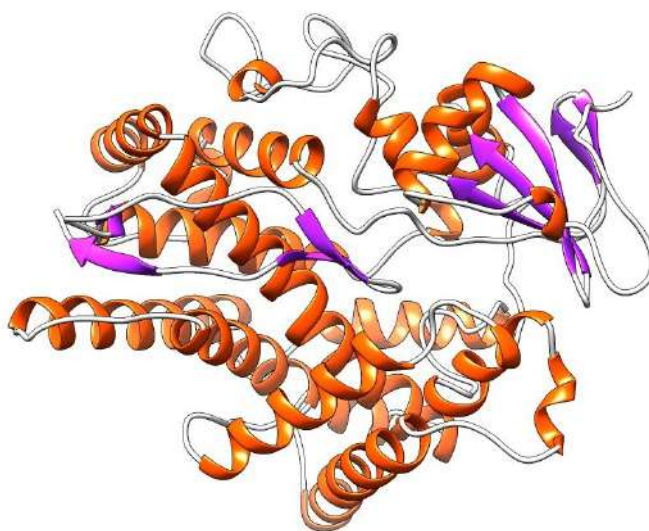


Рис. 1. Модель ензиму CYP719A14, який проявляє хейліантіфолінсинтазну активність.

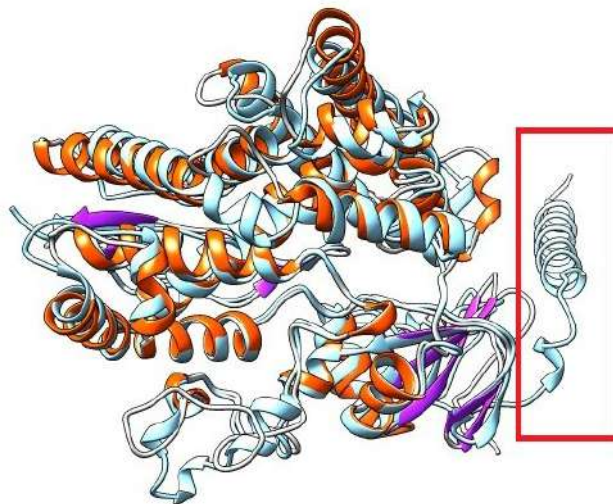


Рис. 2. Порівняння структур, побудованих методом гомологічного моделювання і AlphaFold. Регіон, який відрізнявся, помічений прямокутником.

	1	11	21	31	41
RMSD: ca					
model.pdb, chain B	35 E				
AF-B1NF20-F1-m...v4.pdb, chain A	1 MDETIIWLII	STVIIVLGI A	KFLLGKSS	SLSTM	WVPGPKKLPPIIGN
	51	61	71	81	91
RMSD: ca					
model.pdb, chain B	49 LHLQLG	GD	VFHVYLANL	AKVYGSVFTI	WVGS
AF-B1NF20-F1-m...v4.pdb, chain A	46 LNLHQL	GD	DVFHVVLANL	AKVYGSVFTI	WVGS
	101	111	121	131	141
RMSD: ca					
model.pdb, chain B	94 VLVNKSSDYS	ARDMPDITKI	ISANWKNISC	SDSGPFWHNL	RKGLQ
AF-B1NF20-F1-m...v4.pdb, chain A	94 VLVNKSSDYS	ARDMPDITKI	ISANWKNISC	SDSGPFWHNL	RKGLQ
	151	161	171	181	191
RMSD: ca					
model.pdb, chain B	142 LTPLV	ASQYHLQERD	MKNLINS	MY	KD
AF-B1NF20-F1-m...v4.pdb, chain A	144 PLN	VA	SOYHLQERD	MKNLINS	MY
	201	211	221	231	241
RMSD: ca					
model.pdb, chain B	181 YLKEETVRL	SRLIF	GGDFGDEKL	VVG	MHHAL
AF-B1NF20-F1-m...v4.pdb, chain A	181 YLKEETVRL	SRLIF	GGDFGDEKL	VVG	MHHAL
	251	261	271	281	291
RMSD: ca					
model.pdb, chain B	222 ASLADA	FK	F	CENLPSH	
AF-B1NF20-F1-m...v4.pdb, chain A	222 ASLADA	FK	F	CENLPSH	
	301	311	321	331	341
RMSD: ca					
model.pdb, chain B	283 PPTN		ITVLY	FLKI	IQ
AF-B1NF20-F1-m...v4.pdb, chain A	283 PPTN		ITVLY	FLKI	IQ

Рис. 3. Порівняння структур, побудованих методом гомологічного моделювання і AlphaFold – амінокислотне вирівнювання.

Для (S)-Хейліантіфоліну ідентифіковано 33 кластери докінгу. Показник FullFitness варіював від -2531,56 до -2553,06 ккал/моль, вільна енергія Гібса ΔG варіювала від -7,61 до -5,58 ккал/моль. Більшість кластерів докінгу групувалася також у потенційному каналі доступу лігандів, як і у випадку із (S)-Скоулеріном (рис. 4, Б).

Для (S)-Тетрагідроклумбаміну ідентифіковано 30 кластерів докінгу. Показник FullFitness варіював від -2515,15 до -2537,23 ккал/моль, вільна енергія Гібса ΔG варіювала

від -8,18 до -6,31 ккал/моль. Більшість кластерів докінгу групувалася також у потенційному каналі доступу лігандів, як і у випадку із (S)-Скоулеріном (рис. 4, В).

Для Ретикуліну ідентифіковано 26 кластерів докінгу. Показник FullFitness варіював від -2517,81 до -2542,82 ккал/моль, вільна енергія Гібса ΔG варіювала від -8,18 до -6,31 ккал/моль. Всі кластери докінгу групувалася лише у потенційному каналі доступу лігандів, як і у випадку із (S)-Скоулеріном (рис. 4, Г).

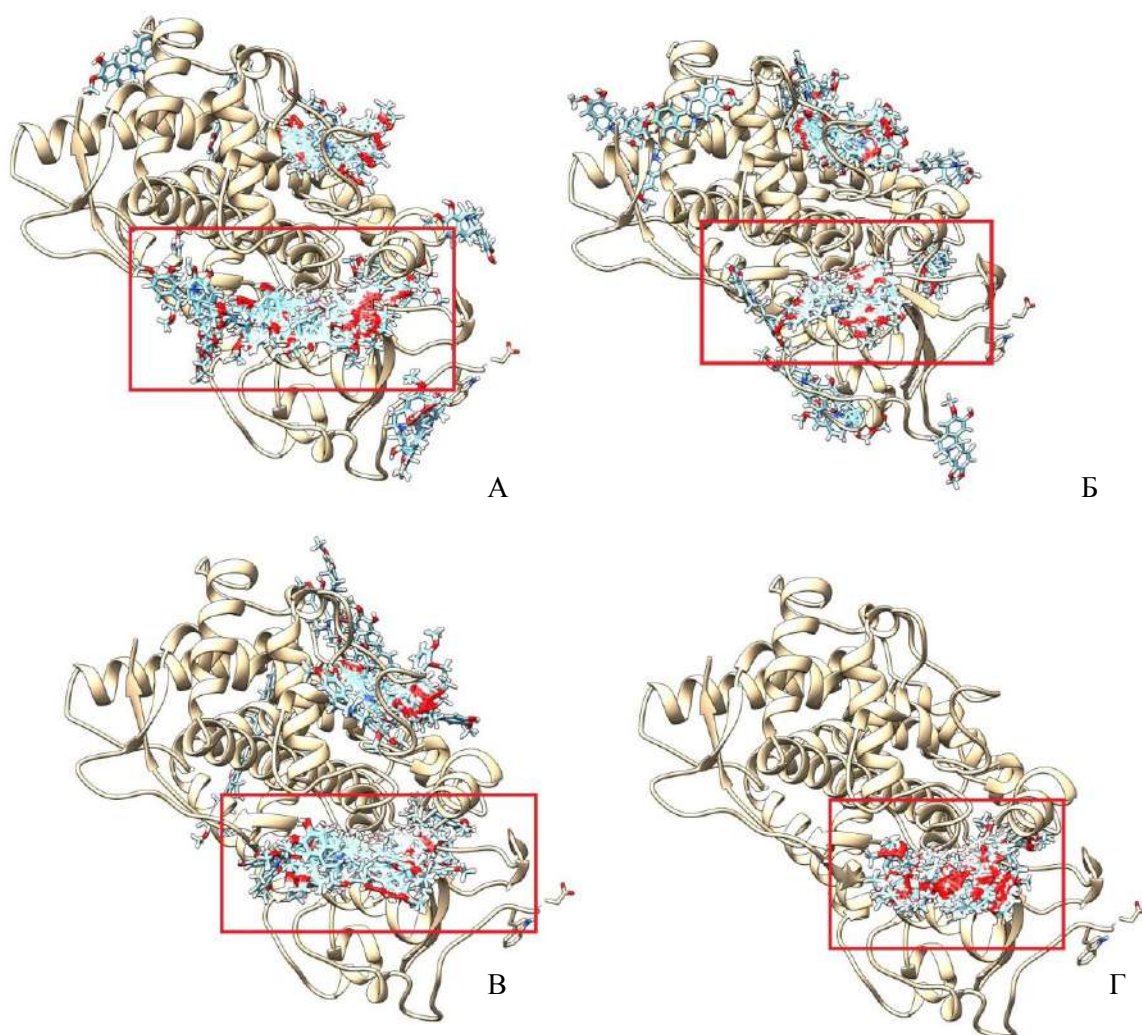


Рис. 4. Кластери молекулярного докінгу: А – (S)-Скоулеріну, Б – (S)-Хейліантіфоліну, В – (S)-Тетрагідраколумбаміну, Г – Ретикуліну. Потенційний канал доступу ліганду позначено прямокутником.

Зважаючи на те, що більшість кластерів для всіх лігандів групувалося саме у порожнині, яка формувалася амінокислотними залишками PHE57-GLY69, ALA285-ARG310, THR369-HSD390, можна зробити припущення, що саме цю порожнину ідентифіковано як потенційний канал доступу ліганду і вказує на адекватність побудованої моделі ензиму.

Висновки

Ідентифіковано потенційний канал доступу ліганду ензиму CYP719A14, яким є порожнина, яка формується амінокислотними залишками PHE57-GLY69, ALA285-ARG310, THR369-HSD390, що дозволить використовувати отримані дані у дослідженнях інженерингу алкалоїдів з метою удосконалення алкалоїдного профілю аргемони мексиканської.

References

1. Brahmachari G., Gorai D., Roy R. Argemone mexicana: Chemical and pharmacological aspects. *Revista Brasileira de Farmacognosia*. 2013. Vol. 23 (3). P. 559–575. doi: 10.1590/S0102-695X2013005000021.
2. Priya C. L., Rao K. V. B Ethanobotanical and current ethanopharmacological aspects of argemone mexicana linn: an overview. *International Journal Of Pharmaceutical Sciences And Research*. 2012. Vol. 3 (7). P. 2143–2148. doi: 10.13040/IJPSR.0975-8232.3(7).2143-48.
3. Mirón-López G. Tissue distribution of transcripts involved in biosynthesis of benzyloisoquinoline alkaloid in mature plants of *Argemone mexicana* L. (*Papaveraceae*). *Revista Fitotecnia Mexicana*. 2018. Vol. 41 (1). P. 13–21. doi: 10.35196/rfm.2018.1.13-21.

4. Laines-Hidalgo J. I., Muñoz-Sánchez J. A., Loza-Müller L., Vázquez-Flota F. An Update of the Sanguinarine and Benzophenanthridine Alkaloids' Biosynthesis and Their Applications. *Molecules*. 2022. Vol. 27 (4). P. 1378. doi: 10.3390/molecules27041378.
5. Díaz Chávez M. L., Rolf M., Gesell A., Kutchan T. M. Characterization of two methylenedioxy bridge-forming cytochrome P450-dependent enzymes of alkaloid formation in the Mexican prickly poppy *Argemone mexicana*. *Arch. Biochem. Biophys.* 2011. Vol. 507 (1). P. 186–193. doi: 10.1016/j.abb.2010.11.016.
6. Yamada Y., Sato F. Transcription Factors in Alkaloid Engineering. *Biomolecules*. 2021. Vol. 11 (11). P. 1719. doi: 10.3390/biom11111719.
7. Urban P., Lautier T., Pompon D., Truan G. Ligand Access Channels in Cytochrome P450 Enzymes: A Review. *Int. J. Mol. Sci.* 2018. Vol. 19 (6). P. 1617. doi: 10.3390/ijms19061617.
8. National Center for Biotechnology Information Database. Retrieved from: <http://www.ncbi.nlm.nih.gov/>.
9. National Center for Biotechnology Information (2023). PubChem Compound Summary for CID 439654, (S)-Scoulerine. Retrieved from: https://pubchem.ncbi.nlm.nih.gov/compound/S_-Scoulerine.
10. National Center for Biotechnology Information (2023). PubChem Compound Summary for CID 440229, (S)-Tetrahydrocolumbamine. Retrieved from: https://pubchem.ncbi.nlm.nih.gov/compound/S_-Tetrahydrocolumbamine.
11. National Center for Biotechnology Information (2023). PubChem Compound Summary for CID 439653, Reticuline. Retrieved from: <https://pubchem.ncbi.nlm.nih.gov/compound/Reticuline>.
12. National Center for Biotechnology Information (2023). PubChem Compound Summary for CID 440582, (S)-Cheilanthifoline. Retrieved from: https://pubchem.ncbi.nlm.nih.gov/compound/S_-Cheilanthifoline.
13. Waterhouse A., Bertoni M., Bienert S. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018. Vol. 46 (W1). P. W296–W303. doi: 10.1093/nar/gky427.
14. Bienert S., Waterhouse A., de Beer T. et al. The SWISS-MODEL Repository – new features and functionality. *Nucleic Acids Res.* 2017. Vol. 45. P. D313–D319. doi: 10.1093/nar/gkw1132.
15. Guex N., Peitsch M., Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*. 2009. Vol. 30. P. S162–S173. doi: 10.1002/elps.200900140.
16. Benkert P., Biasini M., Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*. 2011. Vol. 27. P. 343–350. doi: 10.1093/bioinformatics/btq662.
17. Bertoni M., Kiefer F., Biasini M., Bordoli L., Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* 2017. Vol. 7 (1). P. 10480. doi: 10.1038/s41598-017-09654-8.
18. Grosdidier A., Zoete V., Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* 2011. Vol. 39. P. W270–W277. doi: 10.1093/nar/gkr366.
19. Pettersen E. F., Goddard T. D., Huang C. C., Couch G. S., Greenblatt D. M., Meng E. C., Ferrin T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* 2004. Vol. 25 (13). P. 1605–1612. doi: 10.1002/jcc.20084.

SLISHCHUK H.^{1,2}, MESHCHERYAKOV V.², VOLKOVA N.¹, ZAKHAROVA O.¹

¹ LLC COTECNA UKRAINE LIMITED,

Ukraine, 65114, Odesa, Lustdorfs'ka doroga str., 140 a

² Odessa state environmental university,

Ukraine, 65016, Odesa, Lvivska str., 15

MEXICAN ARGEMONE (*ARGEMONE MEXICANA* L.) CYP719A14 PROTEIN BIOINFORMATIC ANALYSIS

Aim. CYP719A14 structure analysis and its interaction with ligands via bioinformatic methods. **Methods.** Homologous modeling of three-dimensional structure of enzyme. Molecular docking. **Results.** CYP719A14 enzyme, which shows cheilanthifoline synthase activity, model was build based on results of homologous modeling. Interaction of built model with possible substrates – benzyloquinoline compounds: (S)-Scoulerine, (S)-Cheilanthifoline, (S)-Tetrahydrocolumbamine, (S)-Reticuline was investigated using molecular docking method, thermodynamic characteristics of interaction were calculated. Potential Ligand Access Chanel – potential active center was identified. **Conclusions.** Potential Ligand Access Chanel CYP719A14 enzyme, which is formed by aminoacids residues PHE57-GLY69, ALA285-ARG310, THR369-HSD390 researched, allowing to use discovered data in alkaloid engineering research with aim to improve Mexican argemone alkaloid profile.

Keywords: argemone, CYP719A14, substrate specificity, structure.