

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА  
ІНСТИТУТ ВИСОКИХ ТЕХНОЛОГІЙ

Кваліфікаційна наукова  
праця на правах рукопису

**КРАВЕЦЬ ІРИНА ОЛЕКСАНДРІВНА**

УДК [54.057+519.688]:612.146:577.15

**ДИСЕРТАЦІЯ**

**ДИЗАЙН ТА ПОШУК ПОТЕНЦІЙНИХ ІНГІБІТОРІВ АНГІОТЕНЗИН-  
ПЕРЕТВОРЮВАЛЬНОГО ФЕРМЕНТУ 2**

102- ХІМІЯ

10 - ПРИРОДНИЧІ НАУКИ

ОНП «Молекулярний дизайн та синтез»

Подається на здобуття наукового ступеня доктора філософії

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

\_\_\_\_\_ (І.О.Кравець)

Науковий керівник:

Доктор хімічних наук, професор

**РЯБУХІН СЕРГІЙ ВІКТОРОВИЧ**

Київ – 2025

## АНОТАЦІЯ

Кравець І.О. Дизайн та пошук потенційних інгібіторів ангіотензин перетворювального ферменту<sup>2</sup> (АПФ<sup>2</sup> - англ. ACE<sup>2</sup>). - Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 102 Хімія (10 – Природничі науки). - Інститут високих технологій Київського національного університету ім.Т.Шевченко, Київ, 2025.

Дисертаційна робота присвячена дизайну та пошуку потенційних інгібіторів ангіотензин-перетворювального ферменту<sup>2</sup> (АПФ<sup>2</sup>).

У рамках роботи виконано такі завдання:

Ідентифіковано нові хемотипи серед структурно подібних аналогів відомих інгібіторів, що демонструють підтверджену селективну активність щодо АПФ<sup>2</sup>. Синтезовано та охарактеризовано молекули, з метою подальшої біологічної оцінки.

Оптимізовано низку відомих синтетичних шляхів та розроблено нові ефективні підходи до синтезу ключових сполук-зондів. Удосконалення процесів дозволило значно збільшити вихід цільових речовин (до десятків грамів), знизити витрати на реагенти та спростити апаратну реалізацію реакцій. Це стало можливим завдяки відмові від дорогих реагентів, впровадженню більш доступного лабораторного обладнання та загальній оптимізації умов синтезу.

Із літературно описаних інгібіторів АПФ<sup>2</sup> два були синтезовані в необхідних кількостях за валідованими процедурами.

Три сполуки подано для реєстрації на міжнародному порталі Chemical Probe Portal (<https://www.chemicalprobes.org/>) в якості хімічних зондів — авторитетному ресурсі для верифікації молекулярних інструментів у біомедичних дослідженнях.

Адаптований біохімічний методу для скринінгу інгібіторів АПФ<sup>2</sup>, а саме наявний комерційний набір Abcam (ab273373), оптимізувавши об'єми реакційної суміші та концентрації ключових компонентів (ферменту, субстрату, інгібітора),

що дозволило досягти надійного аналітичного сигналу навіть при зменшеному загальному об'ємі реакції (до 25 мкл).

Метод було валідувано шляхом визначення  $IC_{50}$  для референтних інгібіторів (MLN-4760 та DX600) в нових умовах. Це забезпечило можливість проведення високопродуктивного скринінгу (HTS) для 577 відібраних *in silico* сполук з подальшою побудовою дозозалежних кривих для підтверджених активних хітів.

Розроблено і валідовано нові моделі докінгу для мішеней АПФ2 та АПФ. Отримані *in silico* моделі продемонстрували високу точність при прогнозуванні активності відомих лігандів, що дало змогу оцінити селективність взаємодії з ферментами АПФ2/АПФ (англ. ACE) для створення набору низькомолекулярних хімічних зондів для специфічного інгібування АПФ2.

Реалізовано високопродуктивний підхід до віртуального скринінгу з фокусом на хемотипове різноманіття. Завдяки цьому підходу було опрацьовано бібліотеки молекул обсягом до 1,4 млрд. сполук із високою ефективністю щодо часу обчислень і ресурсів.

Запропоновано оптимізовані протоколи, що забезпечують до десятикратного зменшення обчислювальних витрат за рахунок раціонального комбінування швидких попередніх фільтрів і більш точних розрахунків.

Проведено віртуальний скринінг Enamine In-Stock колекції із використанням трьох незалежно валідованих робочих процесів. Вибрані комплекси були додатково проаналізовані методами молекулярної динаміки для верифікації стабільності зв'язування.

На основі результатів комп'ютерного скринінгу було відібрано та протестовано 577 сполук із Enamine In-Stock колекції в експериментальному ферментативному аналізі з використанням флуоресцентного методу (HTS-формат). Для двох еталонних сполук ("Abscam" та EN300-22203532) визначено  $IC_{50}$  значення — 185 нМ і 0,9 нМ відповідно.

Виявлено 7 нових хітів, які продемонстрували рівень інгібування >30%. Для цих семи активних сполук побудовані криві залежності доза-відповідь

(DRC), що підтвердило їхню біологічну релевантність. Одна зі сполук (Z1459912954) продемонструвала  $IC_{50} = 16,8$  мкМ та була рекомендована для подальшої структурної оптимізації.

Після виявлення активних хітів було виконано віртуальне подальше моделювання для розробки активніших та селективніших аналогів.

Сформовано новий інтегрований робочий процес для комплексного вивчення особливостей сайту зв'язування з урахуванням методик FTMove, FTMap, а також залученням QSAR-моделювання з використанням алгоритмів машинного навчання.

Окрему увагу приділено пошуку потенційних цинк-зв'язувальних сполук, здатних до координації з іоном  $Zn^{2+}$  у каталітичному центрі мішені. Реалізовано окремий докінг-напрямок, орієнтований на виявлення таких лігандів з бажаними координаційними властивостями.

Сумарно в результаті виконаної роботи отримано 224 сполук-кандидатів, рекомендованих для подальшої біологічної валідації.

**Наукова новизна отриманих результатів** полягає в застосуванні цілісного поєднання сучасних комп'ютерних (*in silico*) та експериментальних підходів, що дозволили реалізувати системний підхід до ідентифікації та валідації нових селективних інгібіторів АПФ2. Уперше реалізовано багаторівневу стратегію пошуку молекулярних зондів, яка поєднує докінг, фармакофорне моделювання, кластерний аналіз хемотипів, QSAR-моделювання з алгоритмами глибокого навчання, молекулярну динаміку, а також широкий спектр методів експериментальної перевірки (HTS-ферментативні аналізи, побудова кривих доза-відповідь). Запропоновано нові підходи до аналізу координаційної взаємодії лігандів з катіоном цинку у структурі мішені. Розроблені *in silico* моделі демонструють високу прогностичну здатність щодо селективності та афінності лігандів до АПФ2. Створено нові обчислювальні робочі процеси, які можуть бути використані для аналізу інших цинк-залежних ферментів. Отримано хімічні зонди з підтвердженою біологічною активністю та структурною новизною, три з яких подано на реєстрацію в Chemical Probe Portal.

**Практичне значення одержаних результатів** визначається створенням та оптимізацією хімічних сполук, які можуть бути застосовані як високоякісні молекулярні зонди для фундаментальних досліджень ролі АПФ2 у фізіології та патології, зокрема та його роль в розповсюдженні та наслідках COVID-19. Розроблені підходи дозволяють проводити скринінг і оптимізацію лігандів для широкого спектра біологічних мішеней. Запропоновані методиками можуть бути інтегровані в існуючі платформи комп'ютерного дизайну ліків, забезпечуючи швидкий перехід від первинного відбору до вдосконалення хітів та синтезу активних речовин у масштабах, придатних для біологічного тестування. Результати цієї роботи створюють основу для подальших розробок у сфері таргетної терапії та вивчення сигнальних шляхів, пов'язаних з АПФ2, що визначається можливістю створення нових інструментальних сполук — молекулярних зондів для фундаментального вивчення АПФ2. Запропоновані підходи можна застосовувати для дослідження інших біологічних мішеней і впровадження у сучасні платформи віртуального скринінгу та подальшої оптимізації активних сполук.

**Ключові слова:** ангіотензин-перетворювальний фермент 2 (АПФ2/АСЕ2), хімічні зонди, органічний синтез, covid-19, молекулярний докінг, *in silico* скринінг, QSAR-моделювання, молекулярна динаміка, медична хімія.

## SUMMARY

Kravets I.O. *Design and Discovery of Potential Inhibitors of Angiotensin-Converting Enzyme 2 (ACE2)*. — Qualification research manuscript. Dissertation for the degree of Doctor of Philosophy in the specialty 102 Chemistry (10 – Natural Sciences). — Institute of High Technologies, Taras Shevchenko National University of Kyiv, Kyiv, 2025.

The dissertation is dedicated to the design and discovery of potential inhibitors of angiotensin-converting enzyme 2 (ACE2).

Within the scope of this research, the following objectives were accomplished:

New chemotypes were identified among structurally related analogs of known inhibitors that demonstrated confirmed selective activity toward ACE2. Selected molecules were synthesized and characterized for further biological evaluation.

Several known synthetic routes were optimized, and new efficient strategies were developed for the synthesis of key probe compounds. Process improvements significantly increased the yield of target compounds (up to tens of grams), reduced reagent costs, and simplified the hardware implementation of reactions. This was achieved by eliminating the use of expensive reagents, implementing more accessible laboratory equipment, and overall optimization of synthesis conditions.

Two literature-described ACE2 inhibitors were synthesized in the required quantities using validated procedures.

Three compounds were submitted for registration in the international Chemical Probe Portal (<https://www.chemicalprobes.org/>) as chemical probes — an authoritative resource for verifying molecular tools in biomedical research.

A biochemical assay for ACE2 inhibitor screening was adapted by optimizing the use of the commercial kit from Abcam (ab273373). Reaction volumes and concentrations of key components (enzyme, substrate, inhibitor) were optimized, enabling reliable analytical signal detection even at reduced total reaction volumes (down to 25  $\mu\text{L}$ ).

The method was validated by determining the IC<sub>50</sub> values of reference inhibitors (MLN-4760 and DX600) under newly established conditions. This enabled high-throughput screening (HTS) of 577 *in silico* selected compounds, followed by the construction of dose-response curves for confirmed active hits.

New docking models for ACE2 and ACE were developed and validated. The resulting *in silico* models demonstrated high accuracy in predicting the activity of known ligands, allowing the assessment of binding selectivity between ACE2 and ACE for the purpose of designing a library of low-molecular-weight chemical probes for specific ACE2 inhibition.

A high-throughput virtual screening approach was implemented with a focus on chemotype diversity. This strategy enabled efficient processing of molecular libraries containing up to 1.4 billion compounds, with high computational efficiency in terms of time and resources.

Optimized protocols were proposed, achieving up to a tenfold reduction in computational costs by rationally combining fast pre-filtering steps with more accurate calculations.

Virtual screening of the Enamine In-Stock collection was performed using three independently validated workflows. Selected complexes were further analyzed via molecular dynamics simulations to verify binding stability.

Based on computational screening results, 577 compounds from the Enamine In-Stock collection were selected and tested in an experimental enzymatic assay using a fluorescent readout in HTS format. IC<sub>50</sub> values were determined for two reference compounds (Abcam reagent and EN300-22203532) — 185 nM and 0.9 nM, respectively.

Seven new hits were identified that demonstrated inhibition levels above 30%. For these seven active compounds, dose-response curves (DRC) were constructed, confirming their biological relevance. One compound (Z1459912954) showed an IC<sub>50</sub> of 16.8 μM and was recommended for further structural optimization.

Following the identification of active hits, virtual follow-up modeling was performed to develop more potent and selective analogs.

A new integrated workflow was developed for comprehensive investigation of the binding site, incorporating FTMove and FTMap methodologies, as well as QSAR modeling using machine learning algorithms.

Special attention was given to identifying potential zinc-binding compounds capable of coordinating with Zn<sup>2+</sup> ions in the catalytic center of the target. A separate docking strategy was implemented to identify ligands with desired coordination properties.

In total, 224 candidate compounds were obtained as a result of this work, recommended for further biological validation.

**Scientific novelty of the results** lies in the application of a comprehensive combination of modern computational (*in silico*) and experimental approaches that enabled a systematic strategy for identifying and validating novel selective ACE2 inhibitors. For the first time, a multi-level strategy was implemented for the discovery of molecular probes, combining docking, pharmacophore modeling, chemotype-based cluster analysis, QSAR modeling with deep learning algorithms, molecular dynamics simulations, and a broad spectrum of experimental validation techniques (HTS enzymatic assays, dose-response curve construction). New approaches to analyzing ligand–zinc ion coordination within the target structure were proposed. The developed *in silico* models exhibit high predictive capability in assessing ligand selectivity and affinity for ACE2. New computational workflows were created that can be applied to the analysis of other zinc-dependent enzymes. Several chemical probes with confirmed biological activity and structural novelty were identified, three of which have been submitted for registration in the Chemical Probe Portal.

**Practical significance of the obtained results** lies in the development and optimization of chemical compounds that can serve as high-quality molecular probes for fundamental research on the physiological and pathological roles of ACE2, including its involvement in the spread and consequences of COVID-19. The developed approaches enable ligand screening and optimization for a broad range of

biological targets. The proposed methodologies can be integrated into existing drug design platforms, ensuring a rapid transition from initial hit selection to hit optimization and synthesis of active compounds at scales suitable for biological testing. The results of this study provide a foundation for future developments in targeted therapy and the investigation of ACE2-related signaling pathways, through the creation of novel instrumental compounds — molecular probes for fundamental ACE2 research. The proposed approaches can also be applied to the study of other biological targets and integrated into modern platforms for virtual screening and further optimization of active compounds.

**Keywords:** angiotensin-converting enzyme 2 (ACE2), chemical probes, organic synthesis, COVID-19, molecular docking, *in silico* screening, QSAR modeling, molecular dynamics, medicinal chemistry.

## СПИСОК ПУБЛІКАЦІЙ

*Статті в іноземних та українських фахових журналах:*

1. **Kravets, I. O.**, Dudenko, D. V., Pashenko, A. E., Borisova, T. A., Tolstanova, G. M., Ryabukhin, S. V., & Volochnyuk, D. M. (2021). Virtual Screening in Search for a Chemical Probe for Angiotensin-Converting Enzyme 2 (ACE2). *Molecules*, 26(24), 7584. <https://doi.org/10.3390/molecules26247584>  
*Особистий внесок: збір літературних даних, побудування моделей віртуального скринінгу, побудова та оптимізація робочого процесу скринінгу, проведення віртуального скринінгу, підготовка всіх ілюстрацій та візуалізацій, написання статті.*
2. Rayevsky, A. V., Poturai, A. S., **Kravets, I. O.**, Pashenko, A. E., Borisova, T. A., Tolstanova, G. M., Volochnyuk, D. M., Borysko, P. O., Vadzyuk, O. B., Alieksieieva, D. O., Zabolotna, Y., Klimchuk, O., Horvath, D., Marcou, G., Ryabukhin, S. V., & Varnek, A. (2022). *In Vitro* Evaluation of *In silico* Screening Approaches in Search for Selective ACE2 Binding Chemical Probes. *Molecules* (Basel, Switzerland), 27(17), 5400. <https://doi.org/10.3390/molecules27175400>  
*Особистий внесок: збір літературних даних, побудування моделей віртуального скринінгу, проведення віртуального скринінгу, написання експериментальної частини статті.*
3. Andrii Kyrylchuk, **Iryna Kravets**, Anton Cherednichenko, Valentyna Tararina, Anna Kapeliukha, Dmytro Dudenko & Mykola Protopopov. Creation of targeted compound libraries based on 3D shape recognition. *Mol Divers* 27, 939–949 (2023). <https://doi.org/10.1007/s11030-022-10447-z>  
*Особистий внесок: збір літературних даних, побудування та валідування робочого процесу, виконання побудови таргетних бібліотек, написання експериментальної частини статті.*

*Тези доповідей на конференціях:*

1. **Грина О. Кравець**, Олексій В. Раєвський, Дмитро В. Дуденко, Юрій С. Мороз, Олександр Є. Пашенко, Дмитро М. Волочнюк, Сергій В. Рябухін. Нові хемоти́пи зв'язуючих речовин для АПФ2, знайдені за допомогою послідовних підходів *in silico* та *in vitro* скринінгу. XXVI УКРАЇНСЬКА КОНФЕРЕНЦІЯ З ОРГАНІЧНОЇ ТА БІООРГАНІЧНОЇ ХІМІЇ. Ужгород. 16-20 вересня 2024р.
2. Alexey Rayevsky, Dmytro Dudenko, **Iryna Kravets**, Yurii Moroz, Alexander Pashenko, Dmitriy Volochnyuk, Sergey Ryabukhin. In search for selective ACE2 chemical probes: *In silico* studies on “world joined screening collection” compounds dataset. ACS Fall 2021, Atlanta, GA, USA, August 22-26, 2021, PID 3593344.
3. O. V. Raevskii, D. V. Dudenko, **I. O. Kravets**, Y. S. Moroz, A. E. Pashenko, D. M. Volochnyuk, S. V. Ryabukhin. Novel chemotypes of angiotensin-converting enzyme 2 binders via successive *in silico* screening and *in vitro* evaluation approaches. EFMC-ASMC 2023, International Symposium on Advances in Synthetic and Medicinal Chemistry, Zagreb, Croatia, September 3-7, 2023, P028.

## ЗМІСТ

АНОТАЦІЯ .....	2
SUMMARY .....	6
СПИСОК ПУБЛІКАЦІЙ.....	10
ЗМІСТ .....	12
ПЕРЕЛІК УМОВНИК ПОЗНАЧЕНЬ .....	14
ВСТУП .....	18
РОЗДІЛ I. Огляд літератури .....	26
1.1. Епідеміологічні дані ВООЗ про COVID-19. Роль АПФ2 в інфекції SARS-CoV-2 .....	26
1.2. Фізіологічна роль АПФ2 та роль в інфекції ГРВІ-CoV-2 .....	28
1.3. Літературні данні щодо сайту зв'язування АПФ2 .....	36
1.4. Огляд Chemical Probes Portal .....	42
РОЗДІЛ II. ВІДБИР ТА СИНТЕЗ СЕЛЕКТИВНИХ ПОТЕНЦІЙНИХ ІНГІБІТОРІВ АПФ2 З ВІДОМОЮ АКТИВНІСТЮ .....	44
РОЗДІЛ III. СТВОРЕННЯ ХІМІЧНОГО ПРОСТОРУ ТА ПОШУК НОВИХ СЕЛЕКТИВНИХ ІНГІБІТОРІВ АПФ2 .....	61
3.1. Методи дослідження та програмне забезпечення .....	62
3.1.1. Методи UNISTRA .....	63
3.1.2. Методи групи Enamine.....	65
3.1.3. Методи групи ChemSpace.....	68
3.1.4. Методи другої ітерації .....	69
3.2. Збір та систематизація даних .....	72
3.2.1. Систематизація даних групи Chem-space.....	72
3.2.2. Набір даних UNISTRA та їх попередня обробка.....	75
3.3. РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ ВІРТУАЛЬНОГО СКРИНІНГУ .....	76
3.3.1. Загальний аналіз баз даних.....	76
3.3.2. Порівняння сайтів зв'язування групи ChemSpace.....	80
3.3.3. Молекулярний докінг групи ChemSpace. Верифікація моделей.....	85
3.3.4. Молекулярний докінг групи ChemSpace. Робочий процес скринінгу для Єнамін Ін-Сток колекції (3.2 мільйона сполук) .....	87
3.3.5. Молекулярний докінг групи ChemSpace. Робочий процес скринінгу для Enamine REAL Database (Enamine RDB, 1.4 мільярдів сполук) .....	91
3.3.6. Дослідження молекулярною динамікою результатів групи ChemSpace.....	96
3.3.7. Фармакофорний скринінг UNISTRA на основі структури.....	100
3.3.8. Фармакофорний скринінг UNISTRA на основі лігандів .....	102
3.3.9. Результати докінгу UNISTRA .....	103

3.3.10. Результати моделювання UNISTRA QSAR .....	105
3.3.11. Результати молекулярного докінгу ENAMINE .....	109
РОЗДІЛ IV. <i>IN VITRO</i> СКРИНІНГ АПФ2 ІНГІБІТОРІВ З КОЛЕКЦІЙ КОМЕРЦІЙНО ДОСТУПНИХ СПОЛУК .....	112
4.1. Методи та планування .....	115
4.2. Матеріали, реагенти та обладнання .....	116
4.3. Оптимізація методу .....	118
4.4. Скринінг винайдених в результаті докінгу хітів .....	121
РОЗДІЛ V. ВІРТУАЛЬНА ОПТИМІЗАЦІЯ СПОЛУК-ЛІДЕРІВ ТА ДОДАТКОВИЙ ВІРТУАЛЬНИЙ СКРИНІНГ REAL SPACE .....	124
5.1. Функціональний топологічний мапінг потенційних субпocketів .....	125
5.2. Пошук та дослідження активаторів .....	128
5.3. Подальша оптимізація 7 хітів знайдених в першому раунді віртуального скринінгу .....	134
ВИСНОВКИ .....	160
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....	162
Додаток А .....	178
Додаток Б .....	180
Додаток В .....	182
Додаток Г .....	187

## ПЕРЕЛІК УМОВНИК ПОЗНАЧЕНЬ

АПФ/АСЕ — ангіотензинперетворювальний фермент / Angiotensin-Converting Enzyme

АПФ2/АСЕ2 — ангіотензинперетворювальний фермент 2/ Angiotensin-Converting Enzyme 2

SARS — тяжкий гострий респіраторний синдром / Severe Acute Respiratory Syndrome

COVID-19 — коронавірусна хвороба 2019 / Coronavirus Disease 2019

PDB — банк даних білкових структур / Protein Data Bank

NEP — нейтральна ендопептидаза, неприлізин / Neutral Endopeptidase

PMID — ідентифікатор PubMed / PubMed Identifier

QSAR — кількісна структура-активність / Quantitative Structure–Activity Relationship

REAL — віртуальна бібліотека Enamine / REadily AccessibLe Database

UNISTRA — Університет Страсбурга / University of Strasbourg

ЯМР — ядерний магнітний резонанс / NMR - Nuclear Magnetic Resonance

РНК — рибонуклеїнова кислота / RNA - Ribonucleic Acid

RMSD — середньоквадратичне відхилення / Root Mean Square Deviation

СКЗ — середньоквадратичне відхилення

RBD — рецептор-зв'язуючий домен / Receptor Binding Domain

DRC — крива доза-відповідь / Dose–Response Curve

MD — молекулярна динаміка / Molecular Dynamics

HTS — високопродуктивний скринінг / High-Throughput Screening

PAINS — структурні інтерференти / Pan-Assay Interference Compounds

PAC — ренін-ангіотензинова система / Renin-Angiotensin System

RBM — мотив зв'язування рецептора / Receptor Binding Motif

NVIDIA — компанія NVIDIA / NVIDIA Corporation

VS — віртуальний скринінг / Virtual Screening

AUC — площа під кривою / Area Under the Curve

SAR — структура-активність / Structure–Activity Relationship

SD — середнє квадратичне відхилення / Standard Deviation

RF — випадковий ліс / Random Forest

MSE — середньоквадратична помилка / Mean Squared Error

ZBG — група зв'язування з цинком / Zinc Binding Group

GROMACS — пакет для моделювання молекулярної динаміки / GRONingen  
MACHINE for Chemical Simulations

Ang — ангіотензин

AT1R — рецептор ангіотензину II типу 1 / Angiotensin II Type 1 Receptor

SLC6A19 — білок-транспортер нейтральних амінокислот / Solute Carrier Family  
6 Member 19

ADME — абсорбція, розподіл, метаболізм і елімінація / Absorption, Distribution,  
Metabolism, and Excretion

DMPK — фармакокінетика та метаболізм лікарських засобів / Drug Metabolism  
and Pharmacokinetics

Exc/Em — довжини хвиль збудження та випромінювання / Excitation and  
Emission wavelengths

TMPRSS2 — трансмембранна серинова протеаза 2 / Transmembrane Protease,  
Serine 2

ГЕБ / BBB — гематоенцефалічний бар'єр / Blood–Brain Barrier

ARDS — гострий респіраторний дистрес-синдром / Acute Respiratory Distress  
Syndrome

hCoV-NL63 / NL63 — людина-коронавірус NL63 / Human Coronavirus NL63

Mas — рецептор Mas / Mas receptor

ChEMBL — база даних біоактивних малих молекул / ChEMBL – a database of  
bioactive drug-like small molecules

IC<sub>50</sub> — концентрація інгібітора, що знижує активність на 50% / Half-maximal  
inhibitory concentration

BA<sub>XV</sub> — збалансована точність класифікації / Balanced Accuracy

TIP4P — чотириточкова модель води / Transferable Intermolecular Potential with  
Four Points

NVT — ансамбль з постійною кількістю частинок, об'ємом і температурою /  
NVT ensemble – constant Number of particles, Volume, and Temperature

NPT — ансамбль з постійною кількістю частинок, тиском і температурою /  
NPT ensemble – constant Number of particles, Pressure, and Temperature

HTVS — високошвидкісний віртуальний скринінг / High-Throughput Virtual  
Screening

SP — стандартна точність / Standard Precision

XP — підвищена точність / Extra Precision

K<sub>i</sub> — константа інгібування / Inhibition constant

pK<sub>i</sub> — від'ємний десятковий логарифм константи інгібування /  
Negative logarithm of the inhibition constant

SMILES — спрощене позначення молекулярного вводу за допомогою рядка /  
Simplified Molecular Input Line Entry System

Q<sup>2</sup> — коефіцієнт детермінації при крос-валідації /  
Cross-validated coefficient of determination

LogP — логарифм розподільчого коефіцієнта / Logarithm of the partition  
coefficient

PSA — полярна поверхнева площа / Polar Surface Area

MW — молекулярна маса / Molecular Weight

HAC — кількість важких атомів / Heavy Atom Count

HBD — донори водневих зв'язків / Hydrogen Bond Donors

HBA — акцептори водневих зв'язків / Hydrogen Bond Acceptors

RotB — кількість обертових зв'язків / Number of Rotatable Bonds

ROC — крива робочих характеристик приймача / Receiver Operating  
Characteristic curve

logROC — логарифмічна ROC-крива / Logarithmic Receiver Operating  
Characteristic curve

LogAUC — площа під логарифмічною ROC-кривою / Logarithmic Area Under  
the Curve

AUAC — площа під кривою впорядкованої точності / Area Under the

Accumulation Curve

ЦП — центральний процесор / Central Processing Unit (CPU)

LB — фармакофор на основі ліганду / Ligand-Based pharmacophore

SB — фармакофор на основі структури / Structure-Based pharmacophore

$\Delta$ RFU - відсоткове зменшення зміни відносної флуоресценції зразка / percent decrease of the sample's relative fluorescence change

GTM — генеративне топографічне відображення / Generative Topographic Mapping

$pEC_{50}$  — від'ємний десятковий логарифм ефективної концентрації 50% / Negative logarithm of the half-maximal effective concentration ( $EC_{50}$ )

SMARTS — специфічна мова для позначення субструктур / SMiles ARbitrary Target Specification

CADD — комп'ютерний дизайн лікарських засобів / Computer-Aided Drug Design

## ВСТУП

**Актуальність роботи.** Незважаючи на те, що з моменту виявлення гострого респіраторного синдрому коронавірусу-2 (SARS-CoV-2) у грудні 2019 року в Ухані, Хубей, Китай, а також з моменту оголошення ВООЗ пандемії COVID-19 11 березня 2020 року, пройшло вже майже два роки на момент початку цього проекту, механізми дії вірусу та причини його впливу на різні мішені ще не до кінця з'ясовані [37]. Проте в науковій літературі з'являється все більше підтверджень гіпотези про ключову роль АПФ2 у зараженні та поширенні SARS-CoV-2 в організмі людини, що отримує дедалі більше підтримки з боку наукової спільноти [115,45].

Відкриття нових ліків — це складний і тривалий процес, який не може швидко адаптуватися до надзвичайних ситуацій, таких як остання світова санітарна криза, спричинена вірусом тяжкої гострої респіраторної синдрому коронавірусу 2 (SARS-CoV-2) [27]. Щоб мати змогу оперативно реагувати на подібні біомедичні виклики, людству необхідно заздалегідь формувати бібліотеки лікарських засобів та молекулярних зондів, здатних взаємодіяти з широким спектром біологічних мішеней — навіть до того моменту, коли ці мішені будуть остаточно ідентифіковані як ключові для перебігу певних захворювань.

На сьогоднішній день, завдяки безпрецедентному рівню глобальної наукової кооперації, життєвий цикл вірусу SARS-CoV-2 детально вивчений. Визначено низку критичних біомішеней, пов'язаних із цим вірусом, які вже доступні науковій спільноті для валідації та подальшого використання у розробці противірусних засобів.

Серед таких мішеней особливу увагу привертає ангіотензинперетворювальний фермент 2 (АПФ2) — цинк-залежний металофермент, локалізований на поверхні ендотеліальних клітин. Потрапляння SARS-CoV-2 у клітини здійснюється через взаємодію S-білка вірусу з позаклітинними доменами трансмембранного білка АПФ2, що, в свою чергу, спричиняє зниження експресії АПФ2 на клітинній поверхні внаслідок

інтерналізації. АПФ2 широко представлений в ендотелії судин, дихальному та кишковому епітелії, альвеолярних макрофагах і нейронах, що вказує на його системну роль у фізіології організму [100].

Багато джерел підтверджують, що епітелій шлунково-кишкового тракту є другим основним місцем вірусної інвазії. Вірусна РНК може залишатися в калі пацієнтів навіть після того, як тести на наявність РНК у дихальних шляхах дають негативний результат. Це вказує на можливість пролонгованої інфекції шлунково-кишкового тракту та потенційний механізм фекально-оральної передачі інфекції [122,121]. Клінічні дослідження показують, що діарея та інші розлади травлення спостерігаються приблизно у 50% випадків COVID-19 [82]. З'являються нові дані, що підтверджують проникнення SARS-CoV-2 в центральну нервову систему і розвиток неврологічних розладів [68]. Згідно з накопиченими даними, SARS-CoV-2 також чинить вплив на систему згортання крові. У пацієнтів з тяжким перебігом COVID-19, які перебувають на стаціонарному лікуванні, спостерігається підвищений ризик утворення тромбів, зокрема венозної тромбоемболії. Лабораторні показники та клінічні симптоми вказують на розвиток тромбоцитопенії та синдрому дисемінованого внутрішньосудинного згортання (ДВЗ-синдрому).

АПФ2 є важливим ферментом, що містить цинк у своєму активному центрі та локалізується на поверхні багатьох клітин, зокрема ендотеліальних. Його відкриття у 2000 році стало ключовим для переосмислення ролі ренін-ангіотензинової системи (РАС) у регуляції судинного тонуусу. Цей фермент функціонує як монокарбоксіпептидаза, розщеплюючи ангіотензин I (Ang I) до Ang 1–9, а ангіотензин II (Ang II) — до Ang 1–7, що має вазодилатуючу та протизапальну дію. Завдяки цим реакціям АПФ2 протидіє гіперактивності РАС і знижує патофізіологічний вплив Ang II через рецептор AT1R. Крім того, АПФ2 має незалежні функції: (i) регулює вбудовування в мембрану, експресію та активність транспортера нейтральних амінокислот SLC6A19 [64,15]; (ii) експресує антимікробні пептиди [68].

Попри важливу роль, яку відіграє АПФ2, експериментальні дослідження цього ферменту залишаються недостатньо розвинутими. Наявна інформація про його біологічні функції переважно є гіпотетичною або заснованою на теоретичних припущеннях, без чітких експериментальних доказів. Наразі більшість досліджень базується на використанні генетично модифікованих моделей тварин із повним «нокаутом» АПФ2 або непрямих методів, таких як інгібування класичного АПФ чи блокування ангіотензинових і мінералокортикоїдних рецепторів. Це значною мірою обмежує точність висновків і підкреслює необхідність нових, більш селективних інструментів дослідження [34].

Дані щодо використання хімічно-генетичних підходів для вивчення ролі АПФ2 через пряме пригнічення його каталітичної активності є обмеженими. Це частково пояснюється відсутністю доступних низькомолекулярних хімічних зондів для АПФ2. На даний момент єдиним комерційно доступним селективним інгібітором АПФ2 є поліпептид DX600 (Ac-GDYSHCSPLRYPWKCTYPDPEGGG-NH<sub>2</sub>), який має значні обмеження для використання *in vivo* через невідповідний профіль ADME/DMPK параметрів [69].

Тож питання про серйозні наслідки для органів і тканин при тимчасовому прямому пригніченні активності АПФ2 під час COVID-19 залишається невіршеним, особливо в контексті можливості летальних ускладнень.

Щоб прояснити цей аспект, ми виконали систематичний пошук у базі даних Reaxys із метою знайти речовини, здатні селективно інгібувати АПФ2, не впливаючи суттєво на активність АПФ. У результаті було ідентифіковано дев'ять потенційних хімічних зондів, які можуть бути ефективними інструментами для подальших досліджень у цьому напрямку [111,21]. Втім, усі знайдені сполуки наразі недоступні для придбання у необхідних кількостях і відомі лише на рівні аналітичних зразків (по кілька міліграмів). Крім того, їх ефективність і доцільність застосування як хімічних зондів поки що не отримали належного експериментального підтвердження.

**Зв'язок роботи з науковими програмами, планами, темами.**  
Дисертаційна робота виконувалась в рамках реалізації проекту НДФУ №2020.01/0480 «хіміко-генетичний підхід до вивчення наслідків пригнічення АПФ2 як фармакологічної моделі побічних ефектів covid-19 в нервовій, кишковій, серцево-судинній системах та системі гемостазу».

**Мета і завдання дослідження.** Метою дисертаційного дослідження є дизайн та пошук інгібіторів АПФ2 для використання як хімічних зондів для моделювання симптомів SARS-CoV-2.

Для досягнення поставленої мети необхідно було розв'язати наступні завдання:

-створити робочі процеси, що стосуються різноманітності хімічних речовин, що дозволить працювати з величезними колекціями сполук, що було застосовано на докінгу 1,4 мільярда сполук з високою виробничою та економічною ефективністю;

-для розробки робочого процесу для проведення *in silico* скринінгу провести детальне дослідження та порівняльне моделювання структур АПФ2 (як таргета) та АПФ та частково NEP у ролі антитаргетів;

-змодельовати фізіологічну взаємодію АПФ2 плюс субстрат;

-побудувати та валідувати загальноприйнятими показниками точності моделі молекулярного докінгу для АПФ2;

-провести віртуальний скринінг стокової колекції Єнамін (3.2 мільйони сполук на початок робіт) використовуючи побудовані моделі;

-провести віртуальний скринінг віртуальної REAL Database колекції Єнамін [36,112] (1.4 мільярд сполук на початок робіт) використовуючи побудовані моделі;

-провести дослідження MD для демонстрації стабільності взаємодії ліганд-білок для виявлення потенціалу сполук бути активними під час тестування *in vitro*;

-синтезувати відомі активні сполуки для розробки біологічного асею та використання в якості контрольних сполук;

-розробити та валідувати *in vitro* тестування активності сполук проти АПФ2;

-провести первинне *in vitro* тестування активності сполук проти АПФ2. Визначити IC<sub>50</sub> для активних сполук (DRC);

-оптимізувати/адаптувати моделі для скринінгу з урахуванням отриманих даних по *in vitro* активності;

-провести другий раунд віртуального скринінгу та хіт-ту-лід оптимізацію включно серед Enamine REAL Space (65B+ сполук);

-запропонувати сполуки зонди з підтвердженою біологічною активністю на реєстрацію в Chemical Probe Portal як потенційні хімічні зонди.

*Об'єкт дослідження:* ангіотензин-перетворювальний фермент 2 (АПФ2) та його роль у фізіологічних процесах і розвитку побічних ефектів COVID-19. Хімічні сполуки – потенційні селективні інгібітори АПФ2.

*Предмет дослідження:* структурно-функціональні особливості АПФ2 та його взаємодія з малими молекулами. Розробка і валідація моделей віртуального скринінгу для пошуку нових інгібіторів АПФ2, синтез і біологічне тестування отриманих хімічних зондів. Вплив тимчасового інгібування АПФ2 на фізіологічні функції організму.

*Методи дослідження:* молекулярний докінг, молекулярна динаміка, віртуальний скринінг, пошук аналогів по молекулярним фінгерпринтам, shape-скринінг, QSAR-моделювання, фармакофорний аналіз, FTMove, FTMap, органічний синтез сполук, очищення та ідентифікація синтезованих сполук, визначення **інгібуючої активності** за допомогою **флуоресцентного аналізу** (Exc/Em 320/420 нм), вивчення залежності активності сполук від дози (дозозалежні дослідження), спектроскопія ЯМР (1H, 13C, 31P), ас-спектрометрія, рентгеноструктурний аналіз.

**Наукова новизна отриманих результатів** полягає в застосуванні цілісного поєднання сучасних комп'ютерних (*in silico*) та експериментальних підходів, що дозволили реалізувати системний підхід до ідентифікації та валідації нових селективних інгібіторів АПФ2. Уперше реалізовано багаторівневу стратегію пошуку молекулярних зондів, яка поєднує докінг,

фармакофорне моделювання, кластерний аналіз хемотипів, QSAR-моделювання з алгоритмами глибокого навчання, молекулярну динаміку, а також широкий спектр методів експериментальної перевірки (HTS-ферментативні аналізи, побудова кривих доза-відповідь). Запропоновано нові підходи до аналізу координаційної взаємодії лігандів з катіоном цинку у структурі мішені. Розроблені *in silico* моделі демонструють високу прогностичну здатність щодо селективності та афінності лігандів до АПФ2. Створено нові обчислювальні воркфлоу, які можуть бути використані для аналізу інших цинк-залежних ферментів. Три сполуки зонди з підтвердженою біологічною активністю подано на реєстрацію в Chemical Probe Portal як потенційні хімічні зонди.

**Практичне значення отриманих результатів** визначається створенням та оптимізацією хімічних сполук, які можуть бути застосовані як високоякісні молекулярні зонди для фундаментальних досліджень ролі АПФ2 у фізіології та патології, зокрема у контексті COVID-19. Розроблені підходи дозволяють проводити скринінг і оптимізацію лігандів для широкого спектра біологічних мішеней. Запропоновані методика можуть бути інтегровані в існуючі платформи комп'ютерного дизайну ліків, забезпечуючи швидкий перехід від первинного скринінгу до хіт-ту-лід оптимізації та синтезу активних речовин у масштабах, придатних для біологічного тестування. Результати цієї роботи створюють основу для подальших розробок у сфері таргетної терапії та вивчення сигнальних шляхів, пов'язаних з АПФ2, а саме — молекулярних зондів для фундаментального вивчення АПФ2. Розроблені підходи можуть бути використані для дослідження інших біологічних мішеней, а також для інтеграції в сучасні платформи віртуального скринінгу та подальшої оптимізації хітів.

**Публікації за темою дисертації.** За темою дисертації опубліковано 3 статті у фахових журналах та тези 3-х доповідей на наукових конференціях.

**Апробація роботи.** Результати роботи були представлені на 3-х наукових конференціях: XXVI УКРАЇНСЬКА КОНФЕРЕНЦІЯ З ОРГАНІЧНОЇ ТА БІООРГАНІЧНОЇ ХІМІЇ, Ужгород, 2024; ACS Fall 2021, Atlanta, GA, USA, August 22-26, 2021; EFMC-ASMC 2023, International Symposium on Advances in Synthetic and Medicinal Chemistry, Zagreb, Croatia, September 3-7, 2023.

**Особистий внесок здобувача.** Основний обсяг експериментальної роботи, що стосується віртуального скринінгу групи ChemSpace та Enamine, включаючи додатковий скринінг, моделювання фізіологічного шарнірного руху, інтеграція моделей у робочий процес, візуальний огляд та візуальний відбір віртуальних «хітів», пошук і аналіз наукової літератури та формулювання висновків дисертаційної роботи, розробка робочих процесів та синтез потенційних інгібіторів ACE2 виконано особисто здобувачем. Постановка задачі, формулювання мети та обговорення отриманих результатів проведено спільно з науковим керівником д.х.н., проф. Рябухіним С. В. та д.х.н., проф. Волочнюком Д. М. Окремих етапів разом з к.б.н. Раєвським О.В. к.х.н., Пащенко О.Є. та к.б.н. Протопоповим М.В. Класичний фармакофорний докінг проведено у співробітництві з Стразбурзьким Університетом, групою проф. Варнека О. (Varnek A.), у складі Заболотна Ю., PhD Климчук О., PhD Марку Ж. та проф. Хорват Д. (Zabolotna, Y., Klimchuk, O., Marcou G., Horvath D.). Використання share-подібних бібліотек, їх дизайн та аналіз проведено в співробітництві з компанією КемСпейс, у складі: к.х.н. Кирильчук А.А., Чередніченко А.С., Тараріна В.В. та Капелюха А.О. Підбір методів класичного віртуального скринінгу, супровід докінгових розрахунків та інтеграція моделей у робочі процеси виконана у співробітництві з PhD Дуденком Д.М. Систематизація та оформлення результатів, структуризація даних та консультації з управлінням інформацією та її валідацією було зроблено разом з д.х.н. Морозом Ю.С. Деякі етапи синтетичної роботи зроблено у співробітництві з Потураєм А.С. Біологічна еквівалентність, валідація моделей, та експериментальна перевірка активності синтезованих сполук проведена за консультативної участі д.б.н., проф. Толстанової Г.М. та д.б.н., проф. Борисової Т.А. спільно з компанією БІСНТА, а саме к.б.н. Бориском П.О., Ваджук О.Б., Алексеєвою Д.О.

**Структура та обсяг роботи.** Дисертаційна робота викладена на 192 сторінках машинописного тексту, складається зі анотації, вступу, 5 розділів, висновків та переліку використаних джерел (128 найменування). Робота містить 5 схем, 12 таблиць та 83 малюнки. Перший розділ (літературний огляд) присвячено узагальненню літературних даних щодо фізіологічної ролі АПФ2, його участі в інфекції SARS-CoV-2, структурних особливостей сайту

зв'язування, а також аналізу наявних даних про селективні інгібітори АПФ2 та огляду міжнародної платформи Chemical Probes Portal.

У другому розділі розглянуто добір, синтез та структурну оптимізацію селективних потенційних інгібіторів АПФ2 на основі наявної літературної інформації, а також вивчення їхньої хімічної поведінки та біоактивності.

Третій розділ присвячено створенню хімічного простору і віртуальному скринінгу нових селективних інгібіторів АПФ2. Описано використані *in silico* методи (молекулярний докінг, фармакофорне моделювання, QSAR), процедури збору і фільтрації даних, валідацію моделей, аналіз байндінг-сайтів та результати віртуального скринінгу великих бібліотек.

Четвертий розділ є експериментальною частиною роботи та містить опис методик ферментативного скринінгу (HTS), умови проведення досліджень, дані про інгібування, визначення  $IC_{50}$  та побудову кривих доза-відповідь для активних хітів.

П'ятий розділ присвячений оптимізації активних сполук, а також додатковому віртуальному скринінгу та структурному аналізу з метою підвищення активності та селективності. Окрема увага приділена аналізу топології субпocketів, цинк-зв'язувальних властивостей лігандів та реалізації додаткових скринінгів в REAL Space.

## **РОЗДІЛ I. Огляд літератури**

### **1.1. Епідеміологічні дані ВООЗ про COVID-19. Роль АПФ2 в інфекції SARS-CoV-2**

11 березня 2020 року Всесвітня організація охорони здоров'я (ВООЗ) офіційно класифікувала спалах COVID-19, спричинений новим коронавірусом SARS-CoV-2, як глобальну пандемію. Це рішення підкреслило масштабність загрози та необхідність узгоджених дій на міжнародному рівні для стримування подальшого поширення вірусу [55]. Починаючи з грудня 2019 року, коли в китайському місті Ухань було вперше зафіксовано випадок зараження COVID-19, новий коронавірус SARS-CoV-2 стрімко поширився по всьому світу, спричинивши масштабні хвилі інфекції у всіх частинах планети. За офіційними даними, станом на жовтень 2023 року кількість підтверджених випадків захворювання перевищила 678 мільйонів, а загальна кількість летальних випадків склала понад 6,9 мільйона [20]. Ці цифри свідчать про безпрецедентний вплив пандемії на глобальне здоров'я та економіку.

COVID-19 спричинений вірусом SARS-CoV-2 — представником родини коронавірусів, що здатні викликати респіраторні захворювання у людини. За амінокислотною послідовністю SARS-CoV-2 має високу ступінь гомології з вірусом SARS-CoV, відповідальним за спалах тяжкого гострого респіраторного синдрому (SARS) у 2002–2003 роках [47]. Основний механізм передачі SARS-CoV-2 полягає в розповсюдженні респіраторних крапель, хоча зараження також можливе через контакт із контамінованими поверхнями. Інкубаційний період варіюється від 2 до 14 днів, що ускладнює епідеміологічний контроль через здатність вірусу передаватися навіть за відсутності симптомів.

Ключовим елементом у патогенезі SARS-CoV-2 виступає ангіотензинперетворювальний фермент 2 (АПФ2), який функціонує як головний рецептор для вірусного проникнення в клітини людини. АПФ2 — мембранозв'язаний фермент, що експресується в ряді критично важливих органів, зокрема в легенях, серці, нирках та кишечнику [47]. Шипоподібний білок вірусу (S-білок), що складається з двох функціональних субодиниць — S1

і S2, відіграє ключову роль у забезпеченні проникнення SARS-CoV-2 у клітину-господаря. Субодиниця S1 відповідає за впізнавання та зв'язування з рецептором АПФ2, тоді як S2 бере участь у процесі злиття мембран — вірусної та клітинної.

Процес інфікування стартує з прикріплення S1-домену до АПФ2 на поверхні клітини. Це зв'язування спричиняє структурні перебудови S-білка, які забезпечують його зближення з мембраною клітини. Подальше розщеплення S-білка здійснюється за участю серинової протеази TMPRSS2, що локалізована на мембрані клітини-господаря. Вона розділяє білок на субодиниці S1 та S2, активуючи механізм злиття мембран [47]. Цей процес розщеплення є критичним для активації S2-домену, який забезпечує злиття вірусної оболонки з мембраною клітини-господаря, що, у свою чергу, дозволяє вірусу проникнути в клітину шляхом ендоцитозу.

Наслідком взаємодії вірусу з рецептором є зменшення рівня експресії АПФ2, що порушує його фізіологічні функції. Таке зниження активності ферменту асоціюється з розвитком артеріальної гіпертензії, змінами скоротливої здатності міокарда та іншими патологічними станами [25]. Підвищена експресія АПФ2 у клітинах легень, судин і кишківника може опосередковано сприяти розвитку системних ускладнень, спричинених SARS-CoV-2, через порушення, викликані вірусом у цих тканинах. Згідно з останніми дослідженнями, встановлено тривожну залежність між спорідненістю взаємодії окремих варіантів SARS-CoV-2 з рецептором АПФ2 і клінічною тяжкістю захворювання. Зокрема, варіант Delta має низку мутацій у S-білку, які значно посилюють його здатність до зв'язування з АПФ2 у порівнянні з вихідним штамом [118]. У свою чергу, варіант Omicron демонструє значно нижчу афінність до цього рецептора [119]. Окрему увагу привертає той факт, що АПФ2 експресується мікрovasкулярними ендотеліальними клітинами кровоносно-мозкового бар'єра (КМБ, англ. BBB), що створює потенційну можливість для проникнення вірусу до центральної нервової системи. Це, ймовірно, пояснює його нейроінвазивні властивості та зв'язок з розвитком неврологічних порушень [99].

Підвищена експресія АПФ2 у клітинах, що вистилають кровоносні судини, а також у тканинах серця, нирок і кишківника, частково пояснює різноманіття клінічних проявів COVID-19 — від незначних респіраторних симптомів до тяжких системних ускладнень, таких як гострий респіраторний дистрес-синдром (ARDS), тромботичні події, ураження міокарда, гостра ниркова недостатність та патології шлунково-кишкового тракту.

Крім того, участь АПФ2 у функціонуванні ренін-ангіотензинової системи (RAS) має важливе значення для розуміння патогенезу захворювання. Зменшення рівня експресії цього ферменту під час інфекції SARS-CoV-2 може порушувати баланс у системі RAS, сприяючи розвитку гострих ушкоджень легеневої тканини та системного запалення, що ускладнює перебіг хвороби [47].

З огляду на важливість АПФ2 у патогенезі COVID-19, селективні інгібітори АПФ2 можуть мати значний потенціал як терапевтичні агенти або молекулярні зонди для дослідження механізмів інфекції та розробки нових стратегій лікування.

## **1.2. Фізіологічна роль АПФ2 та роль в інфекції ГРВІ-CoV-2**

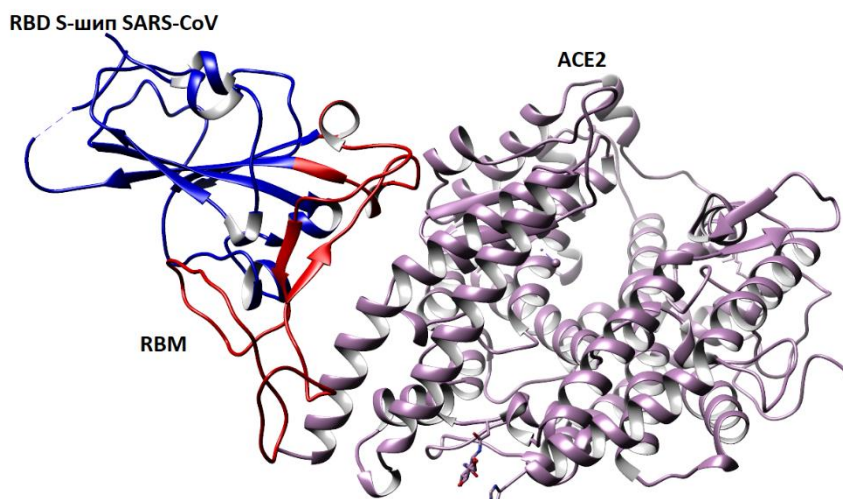
Ангіотензин-перетворювальний фермент 2 (АПФ2) відіграє ключову роль у проникненні COVID-19 в клітину, виступаючи рецептором для вірусного білка S. АПФ2 не тільки критично важливий у патогенезі SARS-CoV-2, але й має ширше значення, оскільки регулює функції серцево-судинної системи. Зокрема, АПФ2 є важливим компонентом ренін-ангіотензинової системи (RAS), яка регулює кров'яний тиск та забезпечення органів кров'ю. Порушення роботи АПФ2 може призвести до тяжких ускладнень, що ми спостерігаємо у пацієнтів з COVID-19 та супутніми серцево-судинними захворюваннями [35].

Епідеміологічні дані показують, що пацієнти із серцево-судинними захворюваннями та гіпертонічною хворобою мають підвищений ризик розвитку важких форм COVID-19. Це пов'язано з важливістю АПФ2 у регулюванні тиску та кровопостачанні органів. У зв'язку з цим, дисбаланс АПФ і АПФ2 може впливати на тяжкість перебігу COVID-19 [47]. Такий дисбаланс може сприяти гіпертензивним станам, а також загострювати вже існуючі серцево-судинні

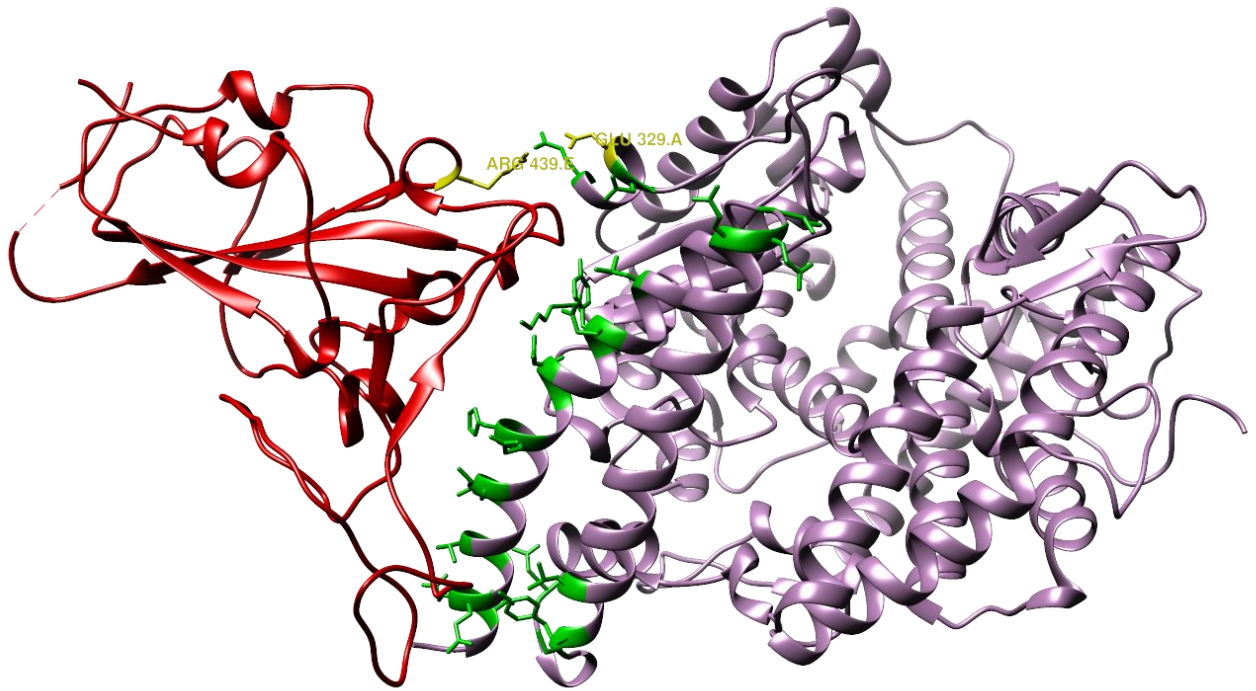
захворювання, що значно підвищує ризик летальних випадків у таких пацієнтів. АПФ2 також діє як рецептор для також людського коронавірусу NL63/НCoV-NL63 [118].

Зв'язування SARS-CoV з рецептором АПФ2 регулює клітинну експресію рецептора, а процес зв'язування індукує інтерналізацію [126]. АПФ2 не тільки сприяє інвазії та швидкому розмноженню SARS-CoV, але також використовується клітинною мембраною, таким чином пошкоджуючи ангіотензин II, що призводить до гострого пошкодження легневих тканин [17,53]. Оскільки легені є основними органами-мішенями для інфікування COVID-19, рання поява респіраторних симптомів є поширеною серед пацієнтів [17]. Є дослідження, що демонструють, що блокування сигнального шляху ренін-ангіотензин може полегшити важке гостре ураження легенів, викликане SARS-CoV-2 [58].

SARS-CoV-2 приєднується до АПФ2 людини шляхом зв'язування спайкових (S) білків [80,113]. Білок S SARS-CoV-2 містить субодиниці S1 і S2. Субодиниця S1 (Мал.1) складається з домену зв'язування рецепторів (RBD), який відповідає за зв'язування з АПФ2 господаря, а субодиниця S2 сприяє злиттю мембран у клітинах господаря [123,120].



*Мал.1. Структура рецептор-зв'язуючого домену (RBD) шипа S у SARS-CoV-2 (червоно-синя стрічки) у комплексі з АПФ2 людини. Червона стрічка позначає рецептор-зв'язуючий мотив (RBM) в амінокислотних залишках 424-494 або 438-504 [123,120].*



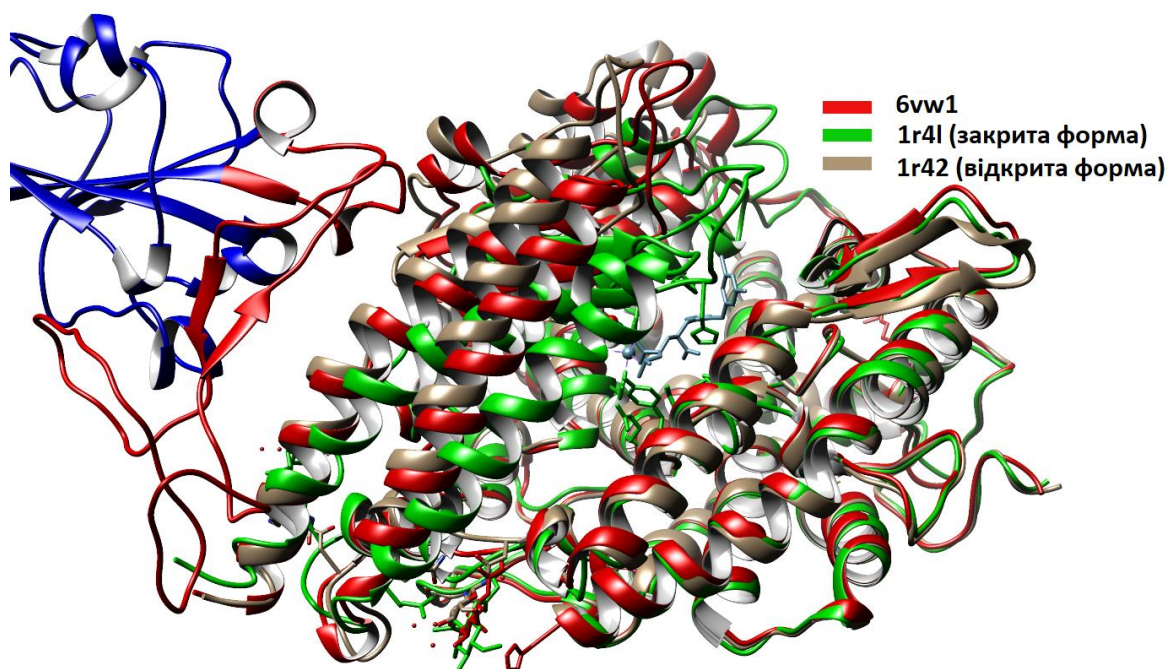
Мал.2. Активний центр АПФ2-RBD (зелений колір), який безпосередньо взаємодіє з RBD білка S SARS-CoV-2. Взаємодія між білком S SARS-CoV-2 і hАПФ2 стабілізується водневим зв'язком (жовтим) між Arg439 (білком S SARS-CoV-2) і Glu329 (hАПФ2). Фігуру було створено за допомогою Chimera шляхом візуалізації структури 6VW1 банку даних білків (PDB) [101].

RBD містить кишеню зв'язування петлі (залишок 424–494 або 438–506), яка називається мотивом зв'язування рецептора (RBM) [123]. RBM розщеплює рецептор АПФ2, щоб SARS-CoV міг проникнути в клітини господаря. Після того, як SARS-CoV зв'язується з АПФ2, субодиниця S2 сприяє злиттю мембран в ендосомальній плазмі через конформаційні зміни, тим самим вивільняючи генوم РНК у клітини-мішені. Після транскрипції та трансляції структурні та неструктурні білки CoV і геному РНК збираються у віріони, які транспортуються через везикули та вивільняються з клітин-мішеней.

Оглядом кристалічних структур виявлені залученими залишками у взаємодію, а саме GLN24, THR27, LYS31, HIS34, GLU37, ASP38, TYR41, GLN42, LEU45, LEU79, MET82, TYR83, ASP90, GLN325, GLU329, ASN330, LYS353 I GLY54, GLU22, GLU23, LYS26, ASP30, GLU35, GLU56 і GLU57 важливі у взаємодії (Мал.2). Примітно, що за даними літератури [43] LYS26 і ASP30 відіграють вирішальну роль у взаємодії білка RBD S SARS-CoV. Крім

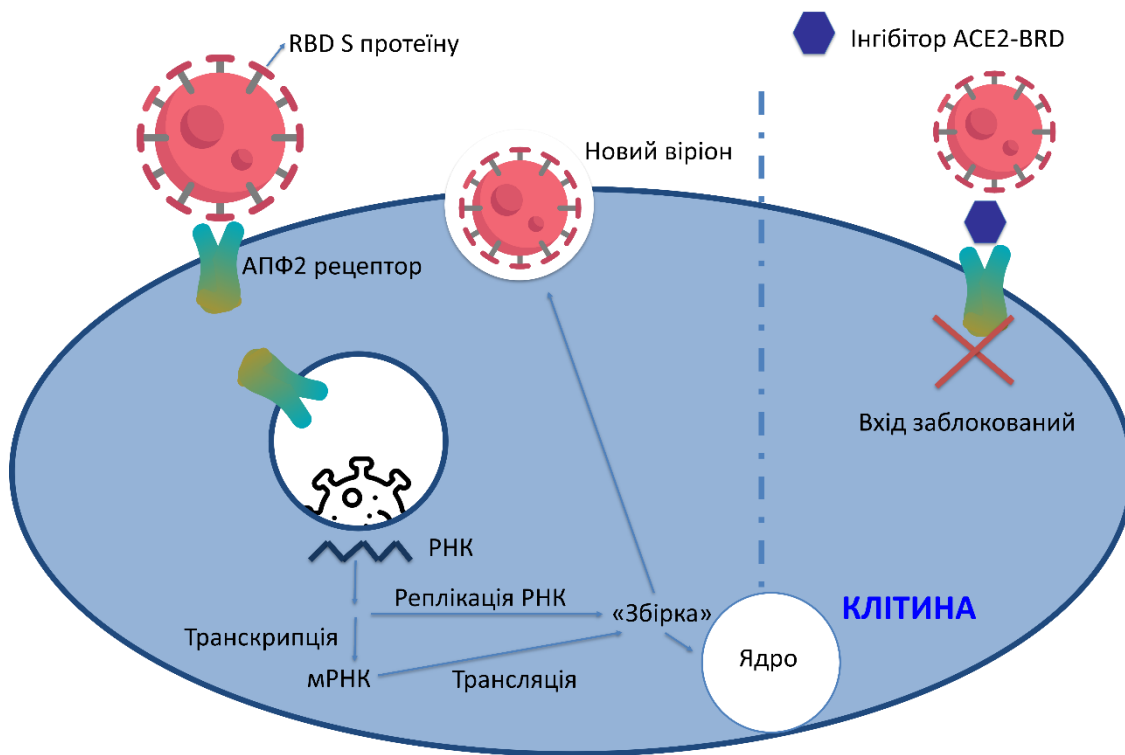
того, GLN325/GLU329 і ASP38/GLN42 АПФ2 є ключовими сайтами зв'язування, які утворюють водневі зв'язки з ARG426 і TYR436 білка S SARS-CoV-2 [127].

Цікавим фактор є те, що порівнявши структури 6VW1, 1R4L (закрита з інгібітором), 1R42 (відкрита апо-форма) виявилось, що структура химерного рецептор-зв'язуючого домену SARS-CoV-2 у комплексі з рецептором АПФ2 людини 6VW1 змінює відкриту апо-форму на напівзакриту. Можливо приєднання домену SARS-CoV-2 до АПФ2 заважає шарнірному руху для взаємодій з природними субстратами (Мал.3).



*Мал.3. Порівняння PDB структур 6VW1, 1R4L (закрита з інгібітором), 1R42 (відкрита апо-форма). Фігуру було створено за допомогою Chimera шляхом накладання та вирівнювання.*

На малюнку 4 показано життєвий цикл SARS-CoV, що ілюструє, як білок S SARS-CoV зв'язується з рецептором АПФ2, що дозволяє вірусу проникати у клітини господаря та вивільняти генوم РНК у клітини-мішені. Цей процес дозволяє вірусу швидко поширюватись і викликати інфекцію, підкреслюючи важливість АПФ2 у патогенезі вірусу.



Мал.4. Життєвий цикл ГРВІ-CoV. Білок (S) SARS-CoV зв'язується з рецептором ангіотензинперетворюючого ферменту 2 (АПФ2), щоб потрапити в клітини господаря і вивільнити геном РНК у клітини-мішені. Структурні та неструктурні білки CoV та геном РНК збираються у віріони, які виділяються з клітин-мішеней Адаптовано з [75].

Стаття [75] розглядає флавоноїди як потенційні інгібітори АПФ2, але вони націлені на пошук байндерів на поверхню зв'язування з RBD. Це дослідження підкреслює перспективність використання природних сполук, таких як флавоноїди та інші природні сполуки та сполуки синтетичного походження, для запобігання інфекції SARS-CoV-2 шляхом перешкоджання його взаємодії з клітинними рецепторами. Для розробки противірусних стратегій важливим є саме блокування взаємодії між рецептор-зв'язуючим доменом (RBD) спайкового білка SARS-CoV-2 та АПФ2, тому є декілька повідомлень [58,46,21,25,71] про знайдені інгібітори АПФ2, але вони є компетитивними лігандами до сайту взаємодії RBD-АПФ2.

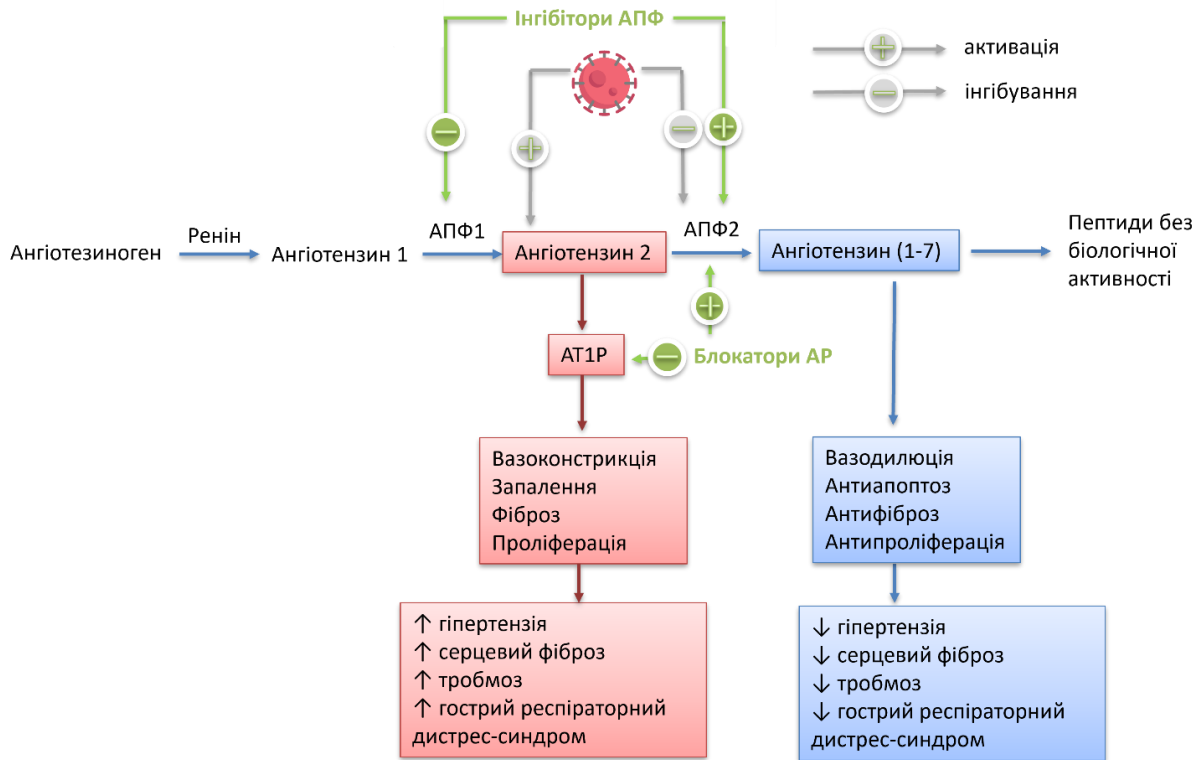
В нашій же роботі ми переслідуюємо іншу мету та маємо за ціль розробку молекулярних зондів, тому ми працювали на пошуком субстратних інгібіторів. Тому ця інформація про конкурентні RBD-АПФ2 дізраптери не може бути використана в побудуванні моделі в рамках даної роботи.

АПФ2 є металопротеазою, що функціонує як карбоксипептидаза. Цей фермент складається з 805 амінокислотних залишків і містить трансмембранний домен, а також одиночну HEXXH послідовність, яка зв'язує цинк (положення амінокислотних залишків 374-378) - HEXXH: консервативний мотив металопротеїназ / Conserved metalloprotease motif (H – гістидин, E – глутамат, X – будь-яка амінокислота (дві позиції), H – гістидин).

. Каталітичний домен АПФ2 на 42% ідентичний домену АПФ [111]. Ця подібність дозволяє зрозуміти, чому АПФ2 також є важливим ферментом у регуляції ангіотензинової системи.

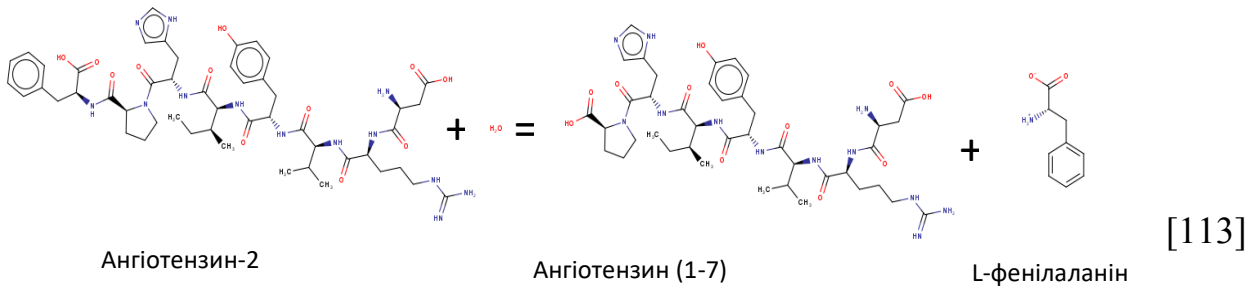
АПФ2 – це мембранний фермент з активним доменом на зовнішній поверхні клітин, що сприяє його ролі як рецептора для вірусів. Проте АПФ2 не лише виконує функцію вірусного рецептора, а й діє як судинорозширювальний засіб, що врівноважує ефекти АПФ. Ця функція є важливою для підтримки балансу між судинозвужуючими та судинорозширювальними ефектами, що дозволяє підтримувати оптимальний рівень артеріального тиску в організмі [34].

Збалансування активності АПФ та АПФ2 у RAS є критичним для підтримання гомеостазу організму. Як показано на малюнку 5, система RAS охоплює складну мережу ферментів, пептидних гормонів та рецепторів. АПФ каталізує перетворення ангіотензину I (Ang I) в ангіотензин II (Ang II), пептид, який викликає звуження судин та підвищення артеріального тиску. Надмірна активність АПФ може призвести до гіпертонії та інших серцево-судинних захворювань, тоді як АПФ2 каталізує перетворення Ang II в Ang (1-7), який діє через рецептор Mas, викликаючи розслаблення судин і зниження артеріального тиску [89]. Відомо, що Ang (1-7) посилює судинорозширювальну дію брадикініну [67], стимулює вивільнення NO та простагландинів і антагонізує дію Ang II [67].



Мал.5. Схематичне відображення регуляції ренінангіотензинової системи (RAS). Адаптовано з [38].

Реакція, що описує основне перетворення, яке каталізує АПФ2:



Додатково, АПФ2 розщеплює Ang I до Ang 1-9, що може знижувати активність АПФ, зменшуючи його субстрат і таким чином сприяючи регуляції кров'яного тиску. Цей процес є критичним для підтримки балансу між судинозвужуючими та судинорозширювальними ефектами, що дозволяє організму підтримувати оптимальний рівень артеріального тиску:



Також АПФ2 видаляє С-кінцевий залишок трьох інших вазоактивних пептидів, нейротензину, кінетензину та des-Arg брадикініну, але не діє на брадикінін:

- брадикінін (1-8) + H<sub>2</sub>O = брадикінін (1-7) + L-фенілаланін [28,111]. Ця реакція протікає в прямому напрямку;

- нейротензин + H<sub>2</sub>O = нейротензин-(1-12) + L-лейцин [28] (в прямому напрямку);

- нейротензин-(1-8) + H<sub>2</sub>O = нейротензин-(1-7) + L-аргінін [111] (в прямому напрямку);

- кінетензин + H<sub>2</sub>O = кінетензин-(1-8) + L-лейцин [28] (в прямому напрямку)

Є дані, що з високою ефективністю АПФ2 розщеплює інші біологічні пептиди, такі як апеліни (апелін-13, [Pyr1]апелін-13, апелін-17, апелін-36), казоморфіни (бета-казоморфін-7, неоказоморфін) і динорфін А:

- апелін-13 + H<sub>2</sub>O = апелін-12 + L-фенілаланін [111];

- [Pyr1]апелін-13 + H<sub>2</sub>O = [Pyr1]апелін-12 + L-фенілаланін [114,124];

- апелін-17 + H<sub>2</sub>O = апелін-16 + L-фенілаланін [114];

- бета-казоморфін-7 + H<sub>2</sub>O = бета-казоморфін-6 + L-ізолейцин [111];

- неоказоморфін + H<sub>2</sub>O = неоказоморфін-(1-5) + L-ізолейцин [111].

Додатково, С-кінцева частина АПФ2 є гомологічною до колектрину і відповідає за транспортування нейтрального амінокислотного транспортера SL6A19 до плазматичної мембрани епітеліальних клітин кишківника через пряму взаємодію, регулюючи його експресію на поверхні клітини та його каталізаторну активність [64,15].

Є дані щодо регулювання діяльності АПФ2 – регулюється хлоридом і фтором, але не бромідом [111], хлорид підвищує ангіотензин I і зменшує розщеплення ангіотензину II [92].

Детальний аналіз ролі АПФ2 у проникненні SARS-CoV-2 у клітину демонструє наступне:

- SARS-CoV і SARS-CoV-2 використовують АПФ2 як основний рецептор для проникнення у клітини;
- серинові протеази TMPRSS2 праймують S білок для забезпечення проникнення вірусу у клітину [47];

- дисбаланс АПФ та АПФ2 в ренін-ангіотензиновій системі може мати суттєвий вплив на тяжкість перебігу COVID-19, особливо у пацієнтів з уже існуючими серцево-судинними захворюваннями.

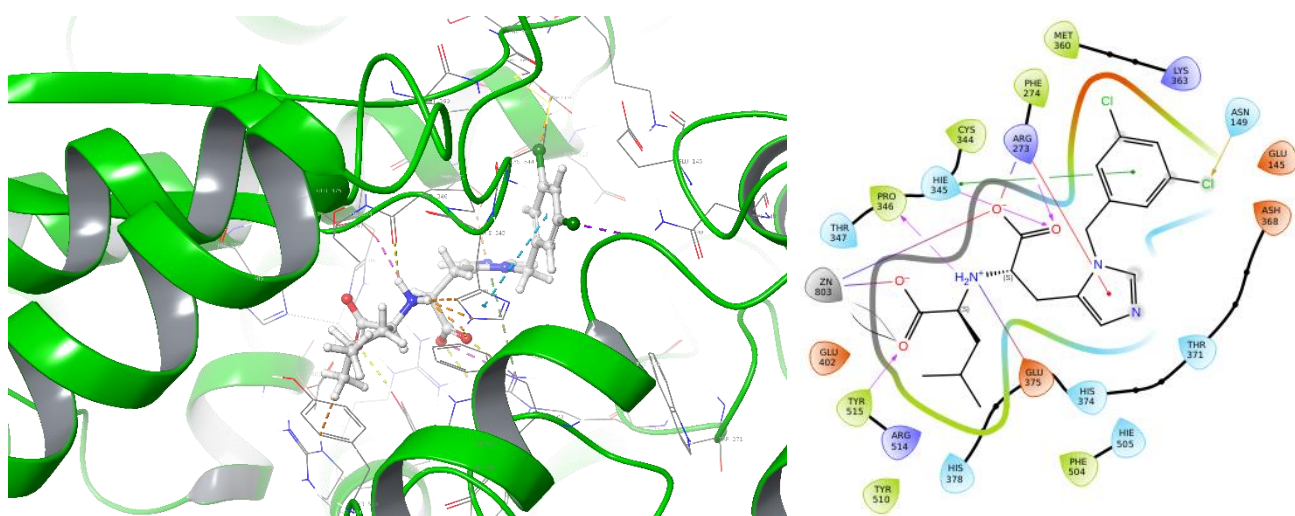
Таким чином, АПФ2 є не лише ключовим компонентом для інфекції SARS-CoV-2, але й критичним елементом у регуляції важливих фізіологічних процесів. Це робить його важливою мішенню для подальших досліджень і розробки нових терапевтичних стратегій, які можуть мати як негайні, так і довгострокові клінічні застосування.

Незважаючи на високу важливість АПФ2, проектів відкриття зв'язувачів АПФ2 значно менше порівняно з його гомологом АПФ [11]. Таким чином, все ще існує потреба у розумінні молекулярних механізмів дії АПФ2 (пов'язаних з SARS-CoV-2 чи ні) і моніторингу потенційного впливу тимчасового припинення дії АПФ2 на різні біологічні тканини і органи. Одним з інструментів, що дозволяють досягти такого розуміння, є хімічні зонди (далі – зонди) — малі молекули, які можуть індукувати бажані біологічні ефекти шляхом селективного зв'язування з певними молекулярними мішенями, дозволяючи дослідникам ставити механістичні та фенотипічні запитання щодо мішені в біохімічних, клітинних або тваринних дослідженнях [117,6,95]. Досі єдиний звіт про розробку зондів, пов'язаних з АПФ2, з експериментальною валідацією описує дослідження двох тіольних сполук, які функціонують як відновники і знижують зв'язування вірусного шипового глікопротеїну з АПФ2 за допомогою алостеричних модифікацій. Проте, замість АПФ2 вони атакують дисульфіди шипа SARS-CoV-2, і тому не можуть бути використані для дослідження впливу зниження рівня АПФ2 на різні біологічні процеси без присутності самого SARS-CoV-2. Тим не менш, існує лише невелика кількість сполук (108, потенційні дублікати включно—у 2020 році, на момент цього дослідження) з зареєстрованою активністю АПФ2 в базі даних ChEMBL [33].

### **1.3. Літературні данні щодо сайту зв'язування АПФ2**

Роль ангіотензин-перетворювального ферменту 2 (АПФ2) у взаємодії з різними інгібіторами була предметом значної кількості досліджень, особливо в

контексті його залучення до регуляції ренінгіотензинової системи (RAS) та взаємодії з патогенами, такими як SARS-CoV-2. Для детального розуміння структурних аспектів активного сайту АПФ2 використовувався комплекс АПФ2 з MLN-4760 (PDB ID: 1R4L). MLN-4760 [57] є селективним інгібітором АПФ2 з дуже низьким значенням  $IC_{50}$  ( $<1$  нМ), що свідчить про високу спорідненість до активного центру ферменту. Це інгібітор, який демонструє сильні та специфічні спрямовані взаємодії, що мають важливе значення для розуміння механізмів інгібування АПФ2 і, як наслідок, для розробки нових терапевтичних засобів [16,57] (Мал.6).

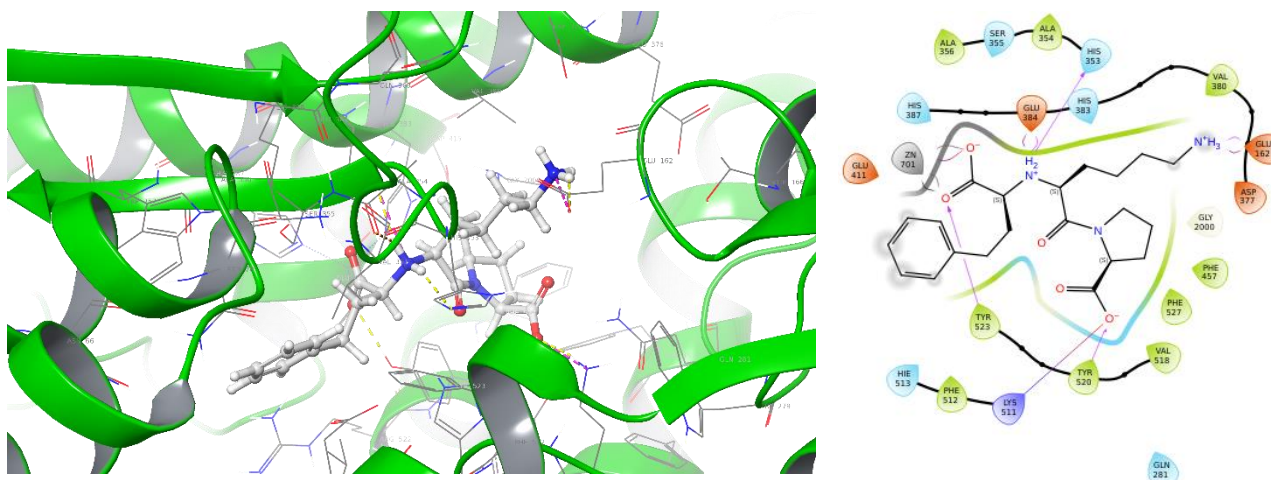


Мал.6. Ліворуч: Крупним планом кристалографічно визначений комплекс АПФ2 – MLN-4760 (PDB ID: 1R4L). Іонні взаємодії (сольові містки) та взаємодії заряду показані у вигляді пунктирних ліній кольору маджента. Водневі зв'язки показані у вигляді помаранчевих пунктирних ліній. Галогенні зв'язки показані у вигляді фіолетових пунктирних ліній.  $\pi$  –  $\pi$ -взаємодії відображаються у вигляді блакитних пунктирних ліній.

Праворуч: Двовимірний вигляд сприятливих взаємодій, що закріплюють кокрystalізований інгібітор в ділянці активного сайту ACE2 (PDB ID: 1R4L). Іонні взаємодії (сольові містки) та водневі взаємодії заряду показані у вигляді фіолетових стрілок. Галогенні зв'язки показані у вигляді помаранчевих стрілок.  $\pi$  –  $\pi$ -взаємодії відображаються у вигляді пурпурних та зелених ліній [40,75].

Також дуже важливе дослідження активного сайту АПФ, через високу спорідненість до АПФ2, особливо активних сайтів, тому була розглянута взаємодія АПФ з лізиноприлом, добре відомим інгібітором АПФ. Лізиноприл широко використовується для лікування артеріальної гіпертензії та серцевої недостатності і є прикладом того, що потрібно врахувати для пошуку саме

селективних інгібіторів АПФ2, але неактивними проти АПФ. На малюнку 7 показано структурні взаємодії між лізиноприлом і АПФ, де пунктиром позначені ключові взаємодії, що підтримують інгібування ферменту [16].



Мал.7. Активний сайт АСЕ у взаємозв'язку з лізиноприлом. Ліворуч: Крупним планом кристалографічно визначений комплекс АПФ – лізиноприл (PDB ID: 1O86). Іонні взаємодії (сольові містки) та взаємодії заряду показані у вигляді пунктирних ліній кольору маджента. Водневі зв'язки показані у вигляді помаранчевих пунктирних ліній. Галогенні зв'язки показані у вигляді фіолетових пунктирних ліній.  $\pi$  –  $\pi$ -взаємодії відображаються у вигляді блакитних пунктирних ліній.

Праворуч: Двовимірний вигляд сприятливих взаємодій, що закріплюють кокрystalізований інгібітор в ділянці активного сайту АСЕ (PDB ID: 1O86). Іонні взаємодії (сольові містки) та водневі взаємодії заряду показані у вигляді фіолетових стрілок. Галогенні зв'язки показані у вигляді помаранчевих стрілок.  $\pi$  –  $\pi$ -взаємодії відображаються у вигляді пурпурних та зелених ліній.

Щодо взаємодії лізиноприлу з амінокислотними залишками ферменту, він формує водневі та іонні зв'язки з негативно зарядженими залишками, такими як ASP415, GLU162, ASP377, GLU384, GLU143, та GLU376. Ці взаємодії є критично важливими для високої спорідненості лізиноприлу до АПФ. Крім того, лізиноприл утворює взаємодії з позитивно зарядженими залишками, такими як LYS511, HIS513, HIS353, та HIS383, що додатково стабілізує комплекс через водневі зв'язки, іонні взаємодії, а також ентропічно сприятливі взаємодії (Мал.7, праворуч) [16].

Ці структурні дослідження є надзвичайно важливими для розробки нових терапевтичних агентів, спрямованих на модулювання активності АПФ2,

особливо з огляду на його центральну роль у патогенезі різних захворювань, включаючи COVID-19. Також вони підкреслюють значення детального вивчення хімічних зв'язків та структурних особливостей сайту зв'язування АПФ2, що може бути корисним для розробки нових інгібіторів з покращеними фармакологічними характеристиками.

*Аналіз даних мутагенезу:*

Проаналізувавши всі доступні мутації в UniProt [105] (табл.1), ми виявили важливі мутації, які впливають на зміну активності АПФ2, а саме:

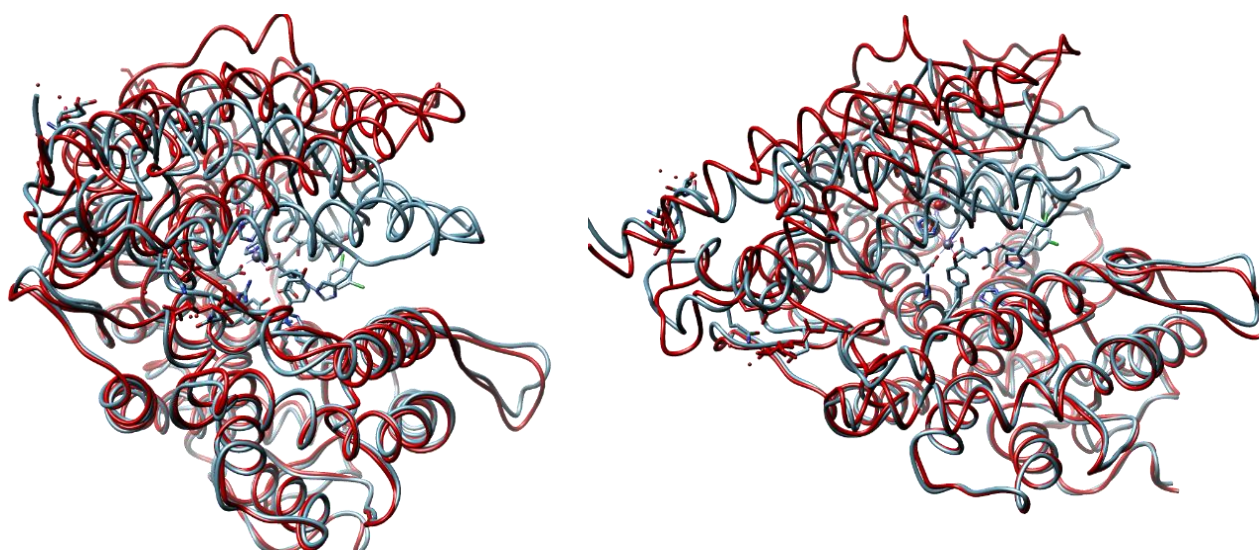
*Табл.1. Важливі мутації, які впливають на зміну активності АПФ2*

Позиція/ Заміна АК	Ефект
271 (W → Q)	Близько 95% втрати здатності до розщеплення ангіотензину I
273 (R → Q)	Повна втрата ферментативної активності. Не впливає на транспорт амінокислот білком SLC6A19
345 (H → A)	Повна втрата ферментативної активності
481 (K → Q)	Близько 80% втрати здатності до розщеплення ангіотензину I
505 (H → A)	Повна втрата ферментативної активності
514 (R → Q)	~50% втрати здатності до розщеплення ангіотензину I, але активність до ангіотензину II зростає вдвічі

Ці спостереження надалі будуть використані в побудові моделей.

Згідно з даними [107] металопептидазний домен АПФ2 можна далі розділити на два субдомени (I та II). Субдомен I, що містить N-кінець, і субдомен II, що містить C-кінець. Це визначення субдомену базується на переміщенні субдомену, яке спостерігалось при зв'язуванні інгібітора. Ділянка зв'язування цинку розташована біля дна та з одного боку великої щілини активного центру (субдомен I), майже посередині її довжини. Цинк координується HIS374, HIS378, GLU402 і однією молекулою води (у нативній структурі). Ці залишки в ділянці зв'язування цинку АПФ2 складають мотив HEXXH + E, збережений у сімействі цинкових металопептидаз.

*Шарнірно-згинальний рух субдомену, залежного від ліганду.* Існує чітка різниця між нативною структурою та структурою АПФ2, зв'язаною з інгібітором, щодо відстані, що розділяє два субдомени. Ці два субдомени зазнають значного залежного від інгібітора шарнірного згинання одного каталітичного субдомену відносно іншого ( $\sim 16^\circ$ ), що змушує глибоку відкриту щілину в нативній формі ферменту закриватися навколо інгібітора. Цей рух можна спостерігати, коли субдомен II з нативної та інгібіторно-зв'язаної структур АПФ2 накладаються (середнє квадратичне (середньоквадратичне) відхилення 1,41 Å для 409 залишків) [107], як показано на малюнку 8.

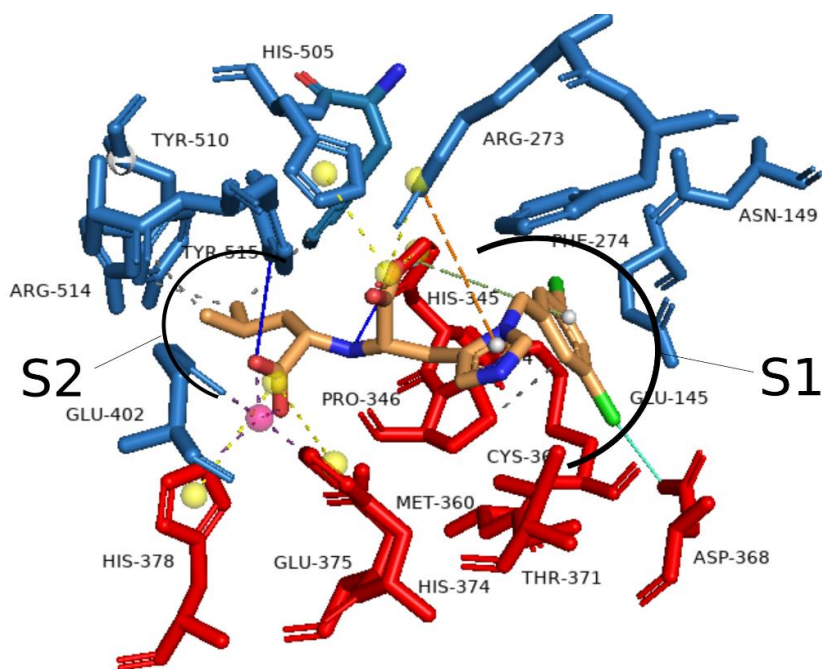


*Мал.8. Суперпозиція нативної та пов'язаної з інгібіторами структур АПФ2. Ліворуч - структури були накладені із середньоквадратичним відхиленням 1,41 Å. Нативний АПФ2 забарвлений у червоний колір, а АПФ2, пов'язаний з інгібіторами, – у блакитний колір. Іони цинку показано у вигляді сфер відповідних кольорів. Праворуч - (ракурс +45град).*

У такому вигляді субдомен II залишається по суті незмінним, але субдомен I рухається, щоб закрити проміжок, по суті імітуючи дію раковини молюска, що закривається. Атоми  $\alpha$ -вуглецю деяких залишків поблизу зовнішнього краю проміжку субдоменів переміщуються на  $\sim 13$  Å, тоді як залишки, що лежать поблизу або на осі шарніра (залишки 99 і 100, 284-293, 396 і 397, 409 і 410), 433 і 434, 539-548 і 564-568) майже нерухомі [107].

*Взаємодія зв'язування інгібіторів і вплив на субстратну специфічність і каталіз.* Обидва металопептидазні субдомени АПФ2 майже однаково залучені до зв'язування інгібітора MLN-4760. Перевірка взаємодії між MLN-4760 і АПФ2

виявила важливі залишки, відповідальні за зв'язування інгібіторів і, ймовірно, за зв'язування субстрату та каталіз. Інгібітор MLN-4760 має дві карбоксилатні групи, одна з яких зв'язується з атомом цинку шляхом витіснення зв'язаної молекули води, присутньої в нативній структурі АПФ2. Координаційна сфера цинку (HIS374, HIS378 і GLU402) є підгрупою з 21 залишку АПФ2, які розташовані в межах 4,5 Å від зв'язаного інгібітора та складають більшу частину активного центру. Шість найбільш важливих із цих залишків сприяють специфічній взаємодії Н-зв'язку з MLN-4760. Бічні ланцюги ARG273, HIS505 і HIS345 пов'язані Н-зв'язками з кінцевим карбоксилатом інгібітора. Карбонільний атом кисню PRO346 і атом азоту HIS345 знаходяться в межах відстані Н-зв'язку від вторинної аміногрупи інгібітора. THR371 знаходиться в межах відстані Н-зв'язку імідазольного кільця групи 3,5-дихлорбензилімідазолу MLN-4760. Фенольна група TYR515 віддає Н-зв'язок зв'язаній з цинком карбоксилатній групі інгібітора. Карбоксильна група GLU375 знаходиться на відстані Н-зв'язку від іншого атома кисню зв'язаного цинком карбоксилату MLN-4760, але, ймовірно, не протонується, доки не відбудеться гідроліз пептиду (Мал.9).



Мал.9. Взаємодія зв'язування інгібітора MLN-4760 з АПФ2. Малюнок зроблений з використанням [5] та PyMOL.

*Цинк як кофактор ензимів.* Цинк необхідний для діяльності >300 ферментів (<https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/enzyme>), що охоплюють усі шість класів ферментів. Місця зв'язування цинку в білках часто мають спотворену тетраедричну або тригонально-біпірамідну геометрію, що складається з сірки цистеїну, азоту гістидину або кисню аспартату та глутамату або їх комбінації. Цинк у білках може брати безпосередню участь у хімічному каталізі або бути важливим для підтримки структури та стабільності білка. У всіх каталітичних центрах іон цинку функціонує як кислота Льюїса.

Властивий хімічний потенціал і реакційна здатність цинку не є винятковими порівняно з іншими металами [61]. Однак, на відміну від інших перехідних металів першого ряду (наприклад,  $\text{Sc}^{2+}$ ,  $\text{Ti}^{2+}$ ,  $\text{V}^{2+}$ ,  $\text{Cr}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Fe}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$  і  $\text{Cu}^{2+}$ ), іон цинку ( $\text{Zn}^{2+}$ ) містить заповнену  $d$ -орбіталь ( $d^{10}$ ) і тому не бере участі в окисно-відновних реакціях, а скоріше функціонує як кислота Льюїса, щоб прийняти пару електронів [61]. Ця відсутність окисно-відновної активності робить  $\text{Zn}^{2+}$  стабільним іоном у біологічному середовищі, потенціал якого постійно змінюється. Таким чином, іон цинку є ідеальним кофактором металу для реакцій, які вимагають окислювально-відновного стабільного іона, щоб функціонувати як каталізатор типу кислоти Льюїса, таких як протеоліз і гідратація вуглекислого газу.

Тому в одній із наших майбутніх стратегій буде використана стратегія пошуку саме цинкових байндерів.

#### **1.4. Огляд Chemical Probes Portal**

Молекулярний зонд — це специфічна хімічна сполука, яка використовується для дослідження функцій біомолекул, таких як білки, нуклеїнові кислоти чи інші молекули в клітинах або тканинах. Молекулярні зонди допомагають виявити і дослідити молекулярні механізми, які лежать в основі біологічних процесів і захворювань. Вони можуть взаємодіяти з конкретними мішенями, такими як рецептори, ферменти чи білки, і часто використовуються для підтвердження ролі цих мішеней у патогенезі або в

реакціях на лікування. Основні характеристики молекулярних зондів включають:

- **Специфічність:** Здатність зв'язуватися з конкретними мішенями або структурами.
- **Активність:** Ефективність у зміні або модулюванні функцій мішеней.
- **Селективність:** Здатність вибірково взаємодіяти з однією або кількома мішенями без значної взаємодії з іншими.

Молекулярні зонди часто використовуються для розробки нових ліків, вивчення біологічних шляхів і механізмів хвороб, а також для проведення фундаментальних досліджень у біомедицині.

Chemical Probes Portal є цінним інструментом для дослідників, оскільки надає перевірену і надійну інформацію про хімічні зонди, що допомагає у виборі відповідних молекул для експериментальних досліджень і розробки нових терапевтичних підходів.

**Chemical Probes Portal** ([www.chemicalprobes.org](http://www.chemicalprobes.org)) — це безкоштовний публічний онлайн-ресурс, створений для підтримки біологічної дослідницької спільноти у виборі високоякісних хімічних інструментів, таких як інгібітори, активатори та деградатори, для використання в експериментах. Цей портал надає експертні огляди та оцінки хімічних проб, допомагаючи дослідникам обирати найбільш відповідні інструменти для *in vitro* та *in vivo* досліджень, а також ділиться рекомендаціями щодо найкращих практик використання цих проб [93].

На момент проведення літературного огляду жодного запису про молекулярні зонди проти АПФ2 на порталі зареєстровані не були.

## РОЗДІЛ II. ВІДБІР ТА СИНТЕЗ СЕЛЕКТИВНИХ ПОТЕНЦІЙНИХ ІНГІБІТОРІВ АПФ2 З ВІДОМОЮ АКТИВНІСТЮ

### 1.1. Підготовка реагентів і розчинників для синтетичних процедур.

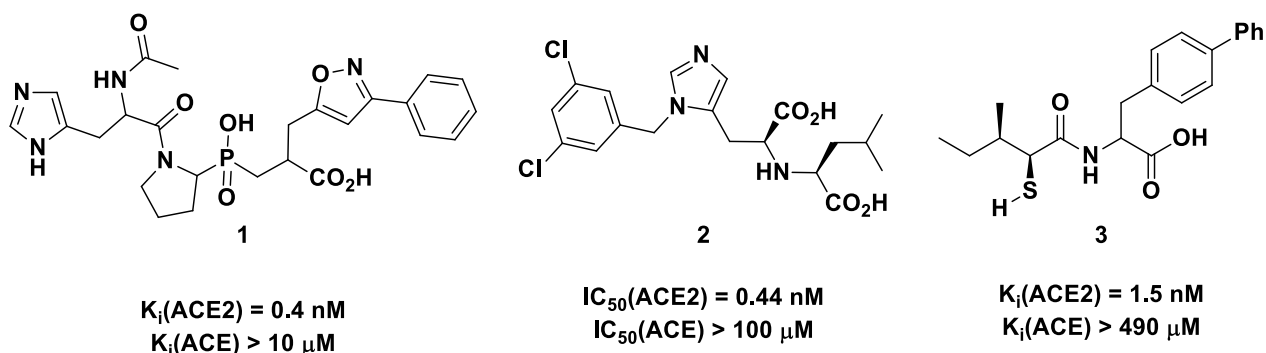
Розчинники очищали відповідно до встановлених стандартів [6]. Початкові реагенти, які були доступні комерційно або отримані в рамках проєкту, використовували без додаткового очищення. Структури отриманих сполук підтверджувалися аналітичними методами.

**1.2.** Температури плавлення визначали за допомогою автоматичної системи MPA100 OptiMelt. Колонкова хроматографія здійснювалася з використанням Kieselgel Merck 60 (230–400 mesh) як стаціонарної фази. Спектри ЯМР  $^1\text{H}$ ,  $^{13}\text{C}\{^1\text{H}\}$ ,  $^{31}\text{P}\{^1\text{H}\}$ ,  $^{19}\text{F}\{^1\text{H}\}$  записували на спектрометрі Bruker 170 Avance 500 (при 500 МГц для  $^1\text{H}$  ЯМР, 126 МГц для  $^{13}\text{C}\{^1\text{H}\}$  ЯМР, 202 МГц для  $^{31}\text{P}\{^1\text{H}\}$  ЯМР, 263 МГц для  $^{19}\text{F}\{^1\text{H}\}$  ЯМР) та Спектрометр Varian Unity Plus 400 (при 400 МГц для  $^1\text{H}$  ЯМР, 101 МГц для  $^{13}\text{C}\{^1\text{H}\}$  ЯМР, 162 МГц для  $^{31}\text{P}\{^1\text{H}\}$  ЯМР, 209 МГц для  $^{19}\text{F}\{^1\text{H}\}$  ЯМР). Хімічні зсуви у спектрах ЯМР наводяться у м.ч. (за шкалою  $\delta$ ) від найменшого до найбільшого з використанням TMS у якості зовнішнього стандарту. Допускаються залишкові піки розчинників для ЯМР при 7,26 та 77,16 м.ч. Константи спінової взаємодії (J) вказані у Гц. Представлення даних: хімічний зсув ( $\delta$ , м.ч.), кратність, інтегральна інтенсивність, J (Гц). Елементний аналіз проводився у лабораторії органічного аналізу Інституту органічної хімії НАН України. Мас-спектри отримували з використанням приладу Agilent 1100 LCMSD SL (режим хімічної іонізації).

**1.3. Відбір молекул для синтезу.** Формування набору молекул-кандидатів для створення моделі інгібування АПФ2 стало відправною точкою дослідження, спрямованого на ідентифікацію потенційних хімічних зондів. Було проведено ретельний аналіз даних з бази Reaxys Medicinal Chemistry з акцентом на сполуки з відомою активністю щодо АПФ2. На основі цього аналізу відібрано 392 сполуки з задокументованою здатністю інгібувати АПФ2.

До відбору були застосовані наступні критерії: показник активності (pX) не нижче 8, а також селективність щонайменше в один порядок між АПФ2 та

АПФ. Після фільтрації залишилось 59 молекул, які задовольняли ці умови. Серед них було виявлено три домінуючі хемоти́пи, що відповідали як критеріям активності, так і селективності. (див. Мал.10.) [21,74,95].



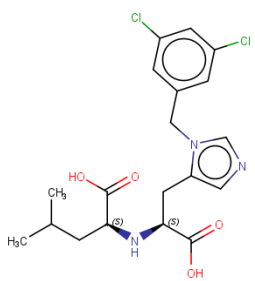
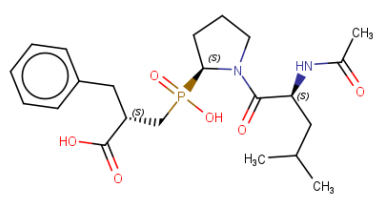
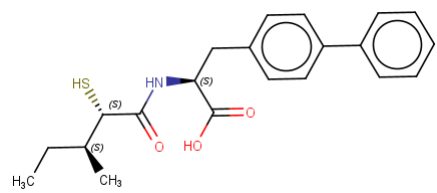
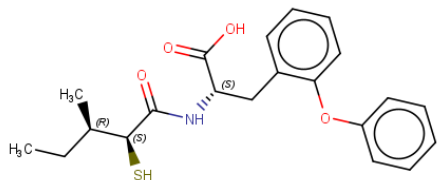
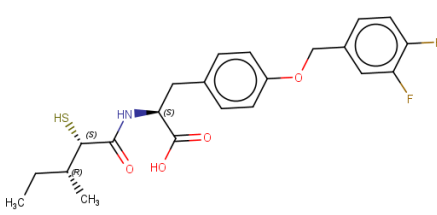
Мал.10. Представники основних хемотипів селективних інгібіторів АПФ2 згідно з даними ReaxysMedcinalChemistry.

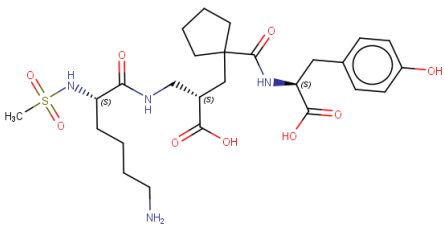
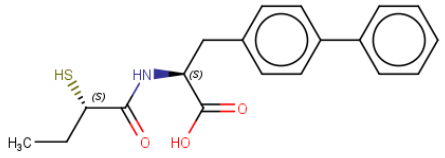
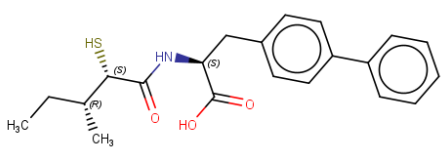
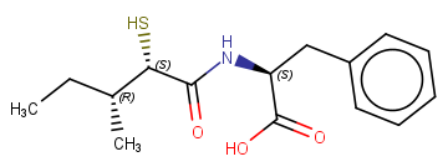
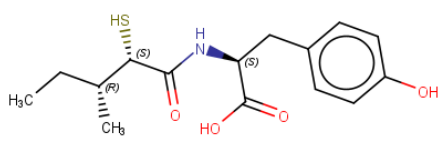
Із трьох виділених хемотипів було обрано 10 сполук, які, на нашу думку, найкраще відповідали критеріям потенційних хімічних зондів. Відбір ґрунтувався на таких ключових параметрах:

- синтетична доступність термодинамічно стабільних молекул;
- наномолярний рівень інгібуючої активності;
- висока селективність до АПФ2 порівняно з АПФ.

Незважаючи на перспективність обраних структур, усі ці сполуки виявилися недоступними у комерційному обігу. У науковій літературі наявна лише обмежена інформація щодо їх синтезу, переважно у вигляді аналітичних зразків у кількості кількох міліграмів. При цьому жодного з прикладів не супроводжено задокументованим підтвердженням використання зазначених речовин як молекулярних зондів (Табл.2).

Табл.2. Відомі молекули обрані для синтезу та подальшого використання в якості молекулярних зондів – селективних інгібіторів ACE2/ACE.

#	Структура	Мол. маса	Формула	Ki
I		428.31	$C_{19}H_{23}Cl_2N_3O_4$	CHEMBL429844 Ki(ACE2)=0.44nM Ki(ACE)>100nM
II		452.49	$C_{22}H_{33}N_2O_6P$	CHEMBL260273 Ki(ACE2)=0.13nM Ki(ACE)>10nM
III		371.50	$C_{21}H_{25}NO_3S$	CHEMBL269996 Ki(ACE2)=1.6nM Ki(ACE)=200nM
IV		387.49	$C_{21}H_{25}NO_4S$	CHEMBL251804 Ki(ACE2)=0.9nM Ki(ACE)=5000nM
V		437.50	$C_{22}H_{25}F_2NO_4S$	CHEMBL251808 Ki(ACE2)=0.85nM Ki(ACE)=1800nM

VI		584.69	$C_{26}H_{40}N_4O_9S$	CHEMBL42583 Ki(ACE2) = NA Ki(ACE) = 13.8nM
VII		343.44	$C_{19}H_{21}NO_3S$	CHEMBL257726 Ki(ACE2) = 1.4nM Ki(ACE) = 8.6nM
VIII		371.50	$C_{21}H_{25}NO_3S$	CHEMBL252391 Ki(ACE2) = 1.5nM Ki(ACE) = 490nM
IX		295.40	$C_{15}H_{21}NO_3S$	CHEMBL400526 Ki(ACE2) = 1.5nM Ki(ACE) = 570nM
X		311.40	$C_{15}H_{21}NO_4S$	CHEMBL253224 Ki(ACE2) = 1.3nM Ki(ACE) = 410nM

Із десяти відібраних молекул лише сполука I (GL1001, MLN-4760, ORE-1001, CHEMBL429844) була раніше протестована *in vivo* як потенційний лікарський засіб при експериментальному виразковому коліті, проте її не використовували як хімічний зонд для вивчення АПФ2 [14]. Беручи до уваги структурну різноманітність, рівень біологічної активності та селективність дії, ці

десять сполук було обрано для подальшого синтезу та біологічної оцінки. Варто зауважити, що сполука II, згідно з даними першоджерела, була представлена як суміш діастереомерів [74], що потенційно ускладнює розробку на її основі селективного хімічного зонду. З огляду на це, на першому етапі заплановано синтез суміші діастереомерів, а після первинного тестування — індивідуальний синтез усіх можливих діастереомерів з метою дослідження їх активності в діастереомерно чистій формі.

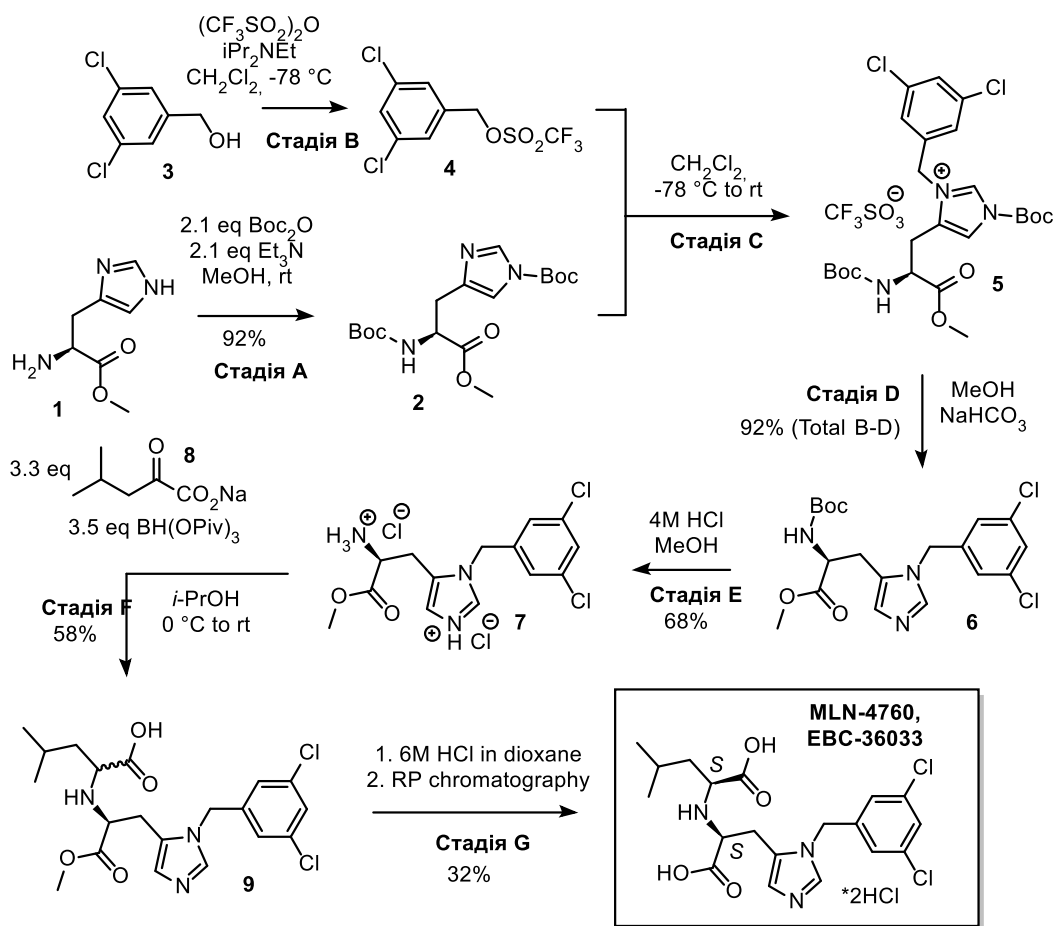
Нижче подано обрані шляхи синтезу разом із методиками, за якими були отримані відповідні сполуки. Усі синтезовані молекули були детально охарактеризовані з використанням сучасних фізико-хімічних методів, а їх структура підтверджена аналітичними даними.

#### **1.4. Сполуки I та II**

Масштабування синтезу сполуки I було здійснено на основі модифікованої оригінальної методики та патентної процедури [21], представлена на Схемі 1.4.1.

Синтез сполуки I (MLN-4760) було оптимізовано та масштабовано до 25 г за одну синтетичну сесію. Вихідними речовинами слугували комерційно доступні сполуки. Етапи B, C та D були реалізовані телескопічно, тобто без ізоляції проміжних продуктів. Сполука 9 була отримана у вигляді суміші двох діастереомерів і використовувалася на наступному етапі без додаткового очищення.

Кінцева сполука була отримана у формі дигідрохлориду та очищена за допомогою препаративної флеш-хроматографії в оберненій фазі. Спектральні характеристики отриманої речовини підтверджують її діастереомерну ідентичність і узгоджуються з раніше опублікованими даними.



**Схема 1.4.1** Синтетична схема синтезу сполуки I.

Для сполуки I найбільш придатною сольовою формою виявився дигідрохлорид.

**Стадія А.** (S)-трет-бутиловий 4-(2-((трет-бутоксикарбоніл)аміно)-3-метокси-3-оксопропіл)-1H-імідазол-1-карбоксилат (сполука 2). До розчину (S)-метилового ефіру гістидину гідрохлориду 1 (300 г, 1,24 моль, 1 екв.) у метанолі (2000 мл) та Et<sub>3</sub>N (360 мл, 2,6 моль, 2,1 екв.) при перемішуванні повільно додавали Boc<sub>2</sub>O (567 г, 2,6 моль, 2,1 екв.) через додаткову воронку. Реакційну суміш перемішували при кімнатній температурі протягом 24 годин. Потім реакційну суміш концентрували у вакуумі, залишок розчиняли у дихлорметані (2 л) та додавали воду (1,5 л). Після поділу фаз органічну фазу промивали сольовим розчином, висушували над Na<sub>2</sub>SO<sub>4</sub>, фільтрували та концентрували. Отримували сполуку 2 у вигляді білого твердого продукту (421 г, 1,14 моль, 92% вихід), яка використовувалась у наступних етапах без додаткового очищення. <sup>1</sup>H NMR (400 МГц, CDCl<sub>3</sub>) δ 7.98 (с, 1H), 7.14 (с, 1H),

5.70 (с, 1H), 4.58 (с, 1H), 3.73 (с, 3H), 3.46 – 2.92 (м, 2H), 1.73 – 1.52 (м, 9H), 1.51 – 1.34 (м, 9H).

**Стадії В, С, D. (S)-метиловий 2-((трет-бутоксикарбоніл)аміно)-3-(1-(3,5-дихлорбензил)-1H-імідазол-5-іл)пропаноат (сполука 6).** Розчин спирту 3,5-дихлорбензилу 3 (150 г, 847 ммоль, 1,1 екв.) та DIPEA (110 г, 852 ммоль, 1,1 екв.) у  $\text{CH}_2\text{Cl}_2$  (1000 мл) повільно додавали до охолодженого до  $-78^\circ\text{C}$  розчину трифторометансульфонового ангідриду  $\text{Tf}_2\text{O}$  (240 г, 851 ммоль, 1,1 екв.) у  $\text{CH}_2\text{Cl}_2$  (2 мл) під аргоном — утворюючи розчин трифлату 4. Через 30 хвилин до цієї суміші повільно додавали розчин сполуки 2 (300 г, 813 ммоль, 1 екв.) у  $\text{CH}_2\text{Cl}_2$  (500 мл). Реакційну суміш залишали при кімнатній температурі на ніч, отримуючи проміжну сполуку 5. Через 20 годин суміш концентрували у вакуумі, залишок розчиняли в MeOH (2000 мл) і нагрівали до кипіння протягом 1 години. Потім суміш концентрували та повторно розчиняли в EtOAc (2 л). Органічну фазу промивали насиченим розчином  $\text{NaHCO}_3$  ( $2 \times 1$  л) та сольовим розчином (1 л), сушили над  $\text{Na}_2\text{SO}_4$ , фільтрували та випаровували до сухого залишку. Отримували алкільований по положенню N-3 імідазольний похідний 6 (322 г, 752 ммоль, 92,5% вихід, сирий продукт), який використовувався без очищення в наступному етапі.  $^1\text{H NMR}$  (500 МГц,  $\text{CDCl}_3$ )  $\delta$  7.59 – 7.41 (м, 1H), 7.35 – 7.18 (м, 1H), 7.02 – 6.60 (м, 3H), 5.99 – 5.73 (ш, 1H), 5.16 – 4.95 (м, 2H), 4.52 – 4.38 (м, 1H), 3.81 – 3.62 (м, 3H), 3.17 – 3.05 (м, 1H), 3.00 – 2.89 (м, 1H), 1.41 (с, 9H).

**Стадія E. (S)-метиловий 2-аміно-3-(1-(3,5-дихлорбензил)-1H-імідазол-5-іл)пропаноат дигідрохлорид (сполука 7).** Сполуку 6 (322 г, 752 ммоль) розчиняли в MeOH (1000 мл) та додавали 4M розчин HCl у діоксані (1 л) при перемішуванні. Отриману суміш перемішували при кімнатній температурі 16 годин, осад фільтрували, промивали діоксаном (500 мл), потім MTBE ( $3 \times 500$  мл) та сушили, отримуючи білий твердий продукт 7 у вигляді дигідрохлориду (204 г, 509 ммоль, 67,7% вихід).  $^1\text{H NMR}$  (400 МГц,  $\text{DMSO}-d_6$ )  $\delta$  9.29 (с, 1H), 9.00 (с, 2H), 7.79 – 7.59 (м, 2H), 7.50 (с, 2H), 5.60 (с, 2H), 4.39 – 4.19 (м, 1H), 3.73 (с, 3H), 3.31 – 3.19 (м, 2H).

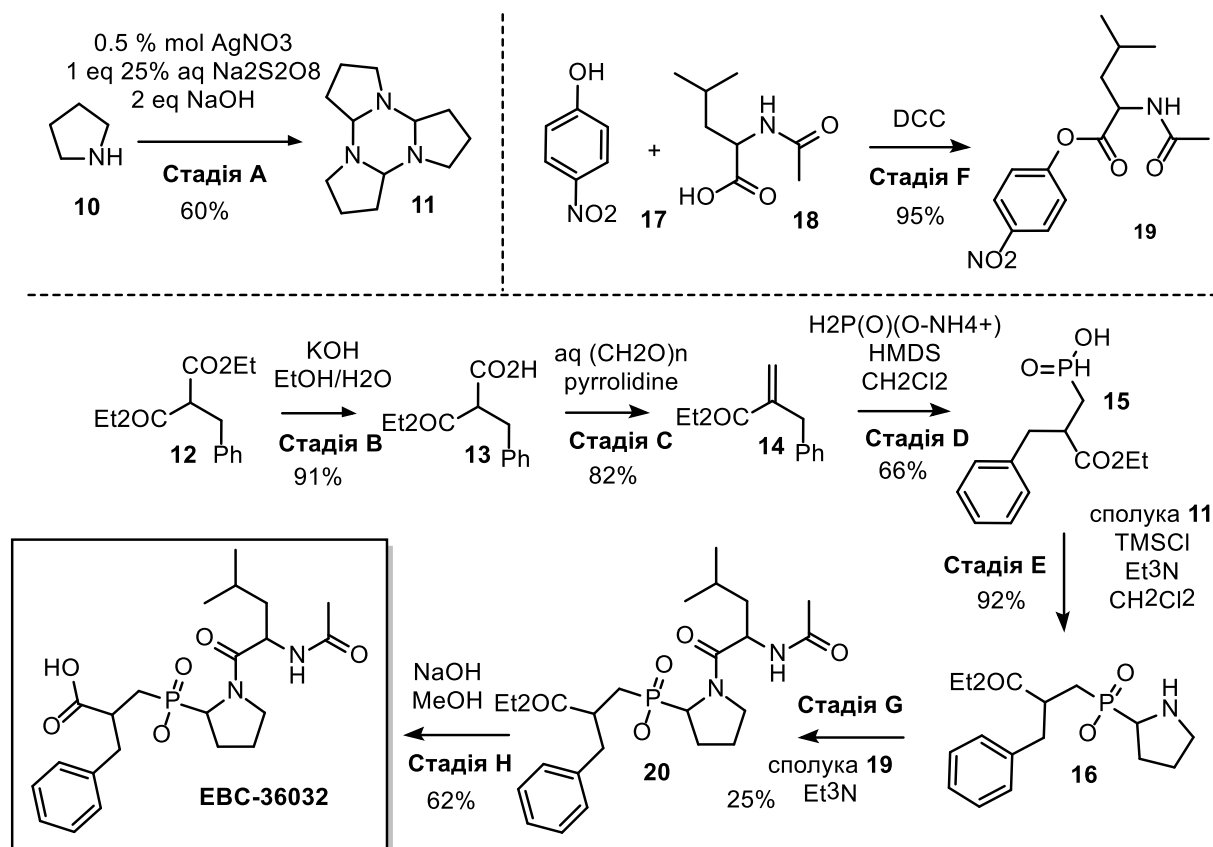
**Стадія F. 2-(((S)-3-(1-(3,5-дихлорбензил)-1H-імідазол-5-іл)-1-метокси-**

**1-оксопропан-2-іл)аміно)-4-метилпентанова кислота (сполука 9).** Півалову кислоту (307,8 г, 3,02 моль, 11 екв.) розчиняли в THF (1500 мл), охолоджували до 0 °С та порціями додавали NaBH<sub>4</sub> (36,5 г, 961 ммоль, 3,5 екв.). Суміш залишали на ніч при кімнатній температурі. Потім розчинник випаровували, залишок розчиняли в ізопропанолі (600 мл), охолоджували до 0 °С та додавали сполуку 7 (110 г, 275 ммоль, 1 екв.). Потім порціями додавали натрієву сіль 4-метил-2-оксопентанової кислоти 8, яку готували з 105 г (905 ммоль, 3,3 екв.) кислоти та 36,4 г (910 ммоль) NaOH у MeOH (500 мл). Отриману суміш залишали при кімнатній температурі на 16 годин. Після випаровування розчинника залишок розчиняли в 2М водному розчині HCl (2,5 л) та екстрагували EtOAc (1 л). Органічну фазу відокремлювали, промивали 2М HCl (500 мл) і відкидали. Водну фазу повторно екстрагували EtOAc (500 мл), підлугували до рН 6–6,5 твердим карбонатом калію та екстрагували EtOAc (3 × 500 мл). Органічні фази об'єднували, сушили над сульфатом натрію, фільтрували та концентрували до сирого продукту 9 (71 г, 161 ммоль, 58,5% вихід) у вигляді суміші двох діастереомерів, що використовувалась у наступному етапі без очищення. <sup>1</sup>H NMR (400 МГц, DMSO-*d*<sub>6</sub>) δ 8.00 (с, 1H), 7.55 (кв, *J* = 3.5, 2.7 Гц, 1H), 7.16 (т, *J* = 2.2 Гц, 2H), 6.92 (с, 1H), 6.87 (с, 2H), 5.41 – 5.22 (м, 2H), 4.09 – 3.97 (м, 1H), 3.63 – 3.49 (м, 3H), 3.43 (т, *J* = 6.8 Гц, 1H), 3.31 – 2.60 (м, 2H), 1.71 – 1.59 (м, 1H), 1.30 (кв, *J* = 7.2 Гц, 1H), 1.16 (тд, *J* = 7.1, 1.5 Гц, 1H), 0.93 – 0.67 (м, 6H).

**Стадія G. (S)-2-(((S)-1-карбокси-2-(1-(3,5-дихлорбензил)-1H-імідазол-5-іл)етил)аміно)-4-метилпентанова кислота дигідрохлорид (MLN-4760).** Сполуку 9 (71 г, 161 ммоль) розчиняли в 6М водному розчині HCl (300 мл) та діоксані (300 мл). Суміш кип'ятили під зворотним холодильником протягом 40 годин, потім концентрували до утворення сирого залишку. Продукт очищували препаративною флеш-хроматографією в оберненій фазі з використанням градієнта H<sub>2</sub>O/CH<sub>3</sub>CN (деталі описується окремо), одержуючи цільову сполуку MLN-4760 у вигляді дигідрохлориду (25,36 г, 50,8 ммоль, 31,6% вихід). <sup>1</sup>H NMR (400 МГц, D<sub>2</sub>O) δ 8.56 (с, 1H), 7.41 (д, *J* = 9.5 Гц, 2H), 7.15 (с, 2H), 5.31 (с, 2H), 3.61 (т, *J* = 6.9 Гц, 1H), 3.47 (т, *J* = 6.9 Гц, 1H), 3.13 (с, 2H), 1.75 – 1.62 (м, 1H),

1.64 – 1.53 (м, 2H), 0.89 – 0.76 (м, 6H) (NH, OH сигнал не спостерігається через обмін). LC/MS (CI):  $m/z = 428 [M+H]^+$ .

Масштабування синтезу сполуки **II** було здійснене на основі модифікованої оригінальної методики [74], представленої на Схемі 1.4.2. Синтез **II** було оптимізовано та масштабовано до 15 г за одну синтетичну сесію. Синтез розпочинався з використання загальнодоступних комерційних вихідних речовин. Сполуки 16 та 19 були отримані у вигляді суміші стереоізомерів і використовувалися на наступному етапі без додаткового очищення. Кінцева сполука також була отримана у вигляді суміші стереоізомерів, аналогічно до оригінальної публікації, і застосовувалася в біологічних дослідженнях без подальшого розділення.



**Схема 1.4.2.** Синтетична схема синтезу сполуки **II**.

**Стадія А.** Додекагідротрипіроло[1,2-а:1',2'-с:1'',2''-е][1,3,5]триазин, **сполука 11**. У круглодонну колбу, обгорнуту алюмінієвою фольгою, до розчину піролідину (175 мл, 149,1 г, 2,10 моль) та твердого NaOH (168 г, 4,2 моль) у воді (2 л) додавали AgNO<sub>3</sub> (1,764 г, 10,36 ммоль). До отриманої суміші при

перемішуванні та температурі 0 °С повільно додавали 25% водний розчин Na<sub>2</sub>S<sub>2</sub>O<sub>8</sub> (2 л, 500 г, 2,10 моль). Реакційну суміш перемішували при кімнатній температурі протягом 3 годин, після чого додавали NaCl до насичення та екстрагували суміш дихлорметаном (2 × 2 л). Органічну фазу висушували над сульфатом натрію та концентрували у вакуумі. Отриманий залишок розчиняли в діетиловому ефірі (200 мл), фільтрували через нейтральну алюміна та видаляли ефір, одержуючи продукт 11 у вигляді оранжевої олії (87 г, 60%). <sup>1</sup>H NMR (500 МГц, DMSO-*d*<sub>6</sub>) δ 3.83 – 3.71 (м, 3H), 3.02 – 2.89 (м, 3H), 2.53 – 2.41 (м, 3H), 1.90 – 1.76 (м, 3H), 1.78 – 1.58 (м, 9H). <sup>13</sup>C NMR (126 МГц, CDCl<sub>3</sub>) δ 81.5, 45.4, 27.4, 19.8.

**Стадія В. 2-Бензил-3-етокси-3-оксoproпано́ва кислота, сполука 12.**

До розчину сполуки 12 (250,3 г, 1,0 моль) в етанолі/воді (2,5 л / 0,3 л) додавали розчин КОН (56,0 г, 1,0 моль) у воді (1 л) крапельно при перемішуванні. Суміш перемішували при кімнатній температурі протягом 16 годин, потім випаровували, розбавляли водою (1,5 л) та екстрагували метилтретбутиловим ефіром (1 л). Водну фазу підкислювали до рН 1 розчином HCl (3 M) і екстрагували етилацетатом (1,5 л). Органічну фазу висушували, концентрували та отримували сполуку 13 (203 г, 91%). <sup>1</sup>H NMR (500 МГц, CDCl<sub>3</sub>) δ 9.27 (с, 1H), 7.33 – 7.25 (м, 2H), 7.25 – 7.18 (м, 3H), 4.16 (кв, *J* = 5.9, 4.8 Гц, 2H), 3.70 (тд, *J* = 7.8, 2.6 Гц, 1H), 3.23 (д, *J* = 7.8 Гц, 2H), 1.20 (тд, *J* = 7.2, 2.5 Гц, 3H).

**Стадія С. Етил-2-бензилакрилат, сполука 14.** До охолодженої до 0 °С суміші сполуки 13 (203 г, 0,91 моль) і формаліну (120 мл) додавали піролідин (120 мл) крапельно. Суміш перемішували при кімнатній температурі 16 годин, після чого розбавляли водою (1 л) і EtOAc (1 л), обережно підкислювали 6 M HCl. Органічну фазу відокремлювали, промивали NaHCO<sub>3</sub>, сушили та концентрували, отримуючи сполуку 14 (142 г, 82%). <sup>1</sup>H NMR (400 МГц, CDCl<sub>3</sub>) δ 7.91 – 6.19 (м, 1H), 7.31 – 7.16 (м, 3H), 7.12 (д, *J* = 7.6 Гц, 2H), 4.08 (квд, *J* = 7.2, 2.6 Гц, 2H), 3.03 (дт, *J* = 15.9, 7.6 Гц, 2H), 2.86 (дт, *J* = 13.7, 5.7 Гц, 1H), 2.16 – 2.00 (м, 1H), 1.90 – 1.73 (м, 1H), 1.16 (тд, *J* = 7.2, 2.4 Гц, 3H).

**Стадія Д. (2-Бензил-3-етокси-3-оксoproпіл)фосфінова кислота,**

**сполука 15.** Суспензію сухого  $\text{H}_2\text{P}(\text{O})(\text{O}^-\text{NH}_4^+)$  (118 г, 1,42 моль) з HMDS (360 мл, 277,2 г, 1,72 моль) нагрівали до 110 °С протягом 2 годин під аргоном. Після охолодження додавали сухий  $\text{CH}_2\text{Cl}_2$  (2 л) і при 0 °С вводили розчин сполуки 14 (142 г, 0,75 моль) у  $\text{CH}_2\text{Cl}_2$  (500 мл) протягом 30 хвилин. Реакційну суміш перемішували 16 годин при кімнатній температурі, потім додавали EtOH (400 мл), випаровували у вакуумі та розчиняли залишок у 5%  $\text{NaHCO}_3$  (2 л). Після промивання з MTBE, підкислення до pH 1 (6 М HCl), та екстракції EtOAc (1,5 л) отримували сполуку 15 у вигляді безбарвної в'язкої рідини (126 г, 66%).  $^1\text{H NMR}$  (400 МГц, DMSO-*d*6)  $\delta$  9.55 (ш, 2H), 7.31 – 7.19 (м, 2H), 7.21 – 7.04 (м, 3H), 4.07 – 3.83 (м, 2H), 3.59 (т,  $J = 6.1$  Гц, 1H), 3.28 – 2.65 (м, 5H), 2.11 – 1.45 (м, 6H), 1.01 (кв,  $J = 6.8$  Гц, 3H).

**Стадія Е. (2-Бензил-3-етокси-3-оксопропіл)піролідин-2-іл)фосфінова кислота, сполука 16.** До розчину сполуки 15 (70 г, 0,27 моль) у сухому DCM (1,2 л) при 0 °С додавали  $\text{Et}_3\text{N}$  (154 мл, 112 г, 1,1 моль) і TMSCl (140 мл, 119,6 г, 1,1 моль). Через 2 години додавали розчин сполуки 11 (70 г, 0,34 моль) у DCM (200 мл). Суміш перемішували 12 годин, додавали EtOH (200 мл), випаровували, обробляли  $\text{Et}_2\text{O}$ , фільтрували, промивали теплим THF (4 л), випаровували та отримували сирий продукт 7 (82 г, 92%), який використовували без очищення. Вихід 82 г (252,04 ммоль, 92%).

**Стадія Ф. 4-Нітрофеніловий ефір 2-ацетамідо-4-метилпентанової кислоти, сполука 19.** До розчину сполук 17 (72 г, 0,52 моль) і 9 (90 г, 0,52 моль) у DCM (700 мл) додавали DCC (98 г, 0,54 моль) порціями при 0 °С. Реакційну суміш залишали на ніч при кімнатній температурі, фільтрували, промивали DCM (100 мл) і випаровували. Отриманий залишок (сполука 19) використовували в наступному етапі без очищення (144 г, 95%).  $^1\text{H NMR}$  (500 МГц,  $\text{CDCl}_3$ )  $\delta$  8.29 – 8.22 (м, 2H), 7.28 (д,  $J = 9.1$  Гц, 2H), 6.03 (д,  $J = 7.7$  Гц, 1H), 5.05 – 4.60 (м, 1H), 2.07 (с, 3H), 1.87 – 1.75 (м, 2H), 1.77 – 1.64 (м, 1H), 1.02 (т,  $J = 6.6$  Гц, 6H).

**Стадія Г. (1-(2-Ацетамідо-4-метилпентаноїл)піролідин-2-іл)(2-бензил-3-етокси-3-оксопропіл)фосфінова кислота, сполука 20.** До суміші сполук 16 (82 г, 252,04 ммоль) і 19 (144 г, 489,2 ммоль) у DMF (700 мл) додавали  $\text{Et}_3\text{N}$  (160

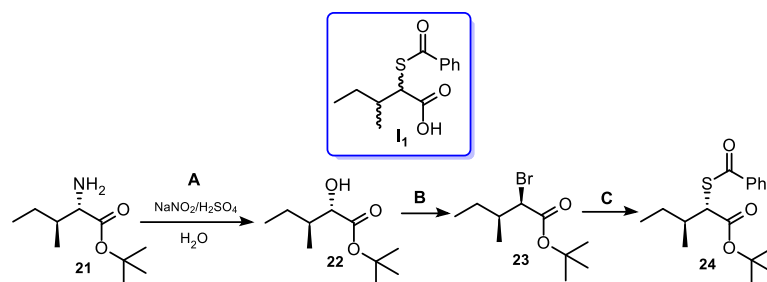
мл, 117,2 г, 1,14 моль). Перемішували при кімнатній температурі 16 годин, додавали NaHSO<sub>4</sub> (2 л), екстрагували EtOAc (2 × 800 мл), промивали водою, сушили та концентрували. Отриманий залишок очищували флеш-хроматографією (MTBE/MeOH), одержуючи сполуку 20 (30 г, 25%) у вигляді суміші стереоізомерів. <sup>1</sup>H NMR (500 МГц, CDCl<sub>3</sub>) δ 7.25 – 7.06 (м, 5H), 4.85 (с, 2H), 3.99 (с, 4H), 3.47 (с, 3H), 3.11 – 2.72 (м, 4H), 2.50 – 1.35 (м, 11H), 1.04 (с, 3H), 0.89 (с, 5H). LC/MS (CI): m/z = 481 [M+H]<sup>+</sup>.

**Стадія Н. 3-((1-(2-Ацетамідо-4-метилпентаноїл)піролідин-2-іл)(гідрокси)фосфорил)-2-бензилпропанова кислота, ЕВС-36032.** До розчину сполуки 16 (30 г, 62,4 ммоль) у MeOH (70 мл) додавали водний розчин NaOH (14 г, 350 ммоль, 70 мл). Реакційну суміш перемішували 16 годин, випаровували метанол у вакуумі, розбавляли водою (400 мл), промивали iPrOAc (300 мл), DCM (300 мл), підкислювали розчином NaHSO<sub>4</sub> (300 мл) до рН 1 та екстрагували EtOAc (400 мл). Органічну фазу висушували, випаровували та сушили у вакуумі (1 мм рт. ст.), отримуючи цільову сполуку ЕВС-36032 (17,4 г, 62%). <sup>1</sup>H NMR (400 МГц, CDCl<sub>3</sub>) δ 11.66 (с, 2H), 7.17 (п, *J* = 7.2, 6.8 Гц, 5H), 4.83 – 4.68 (м, 1H), 4.40 – 4.15 (м, 1H), 3.95 – 3.59 (м, 1H), 3.45 (с, 1H), 3.19 – 2.93 (м, 2H), 2.86 (с, 1H), 2.13 (дд, *J* = 57.0, 34.8 Гц, 3H), 1.91 (д, *J* = 10.7 Гц, 6H), 1.75 – 1.20 (м, 4H), 0.89 (кв, *J* = 7.5, 7.0 Гц, 6H). <sup>31</sup>P NMR (202 МГц, CDCl<sub>3</sub>) δ 51.10, 50.61, 49.71. LC/MS (CI): m/z = 453 [M+H]<sup>+</sup>.

### 1.5. Синтез сполук III, IV, V, VII, VIII, IX та X

Ключовим проміжним продуктом для синтезу всіх цільових сполук є прекурсор I<sub>1</sub>, як показано на Схемі 1.5.1. Єдине, що змінювалося в цьому прекурсорі при переході від однієї сполуки до іншої – це конфігурація відповідних карбокситіолів. Саме тому на початку ми вирішили валідувати та масштабувати існуючі підходи, використовуючи відому послідовність трансформації доступних амінокислот через діазотування в гідрокси кислоти і вже далі в відповідні тіокислоти. Для валідації в якості моделі ми взяли природний ізолейцин. Проте, протягом подальших досліджень нами було встановлено, що перетворення інтермедіату 21 в інтермедіат 22 в

мультиграмових масштабах є неефективним [94]. Це привело до відмови від подальшої роботи саме над цією серією сполук через їх недостатню синтетичну доступність.



**Схема 1.5.1.** Синтетична схема синтезу прекурсору 24 на модельному природному ізолейцині.

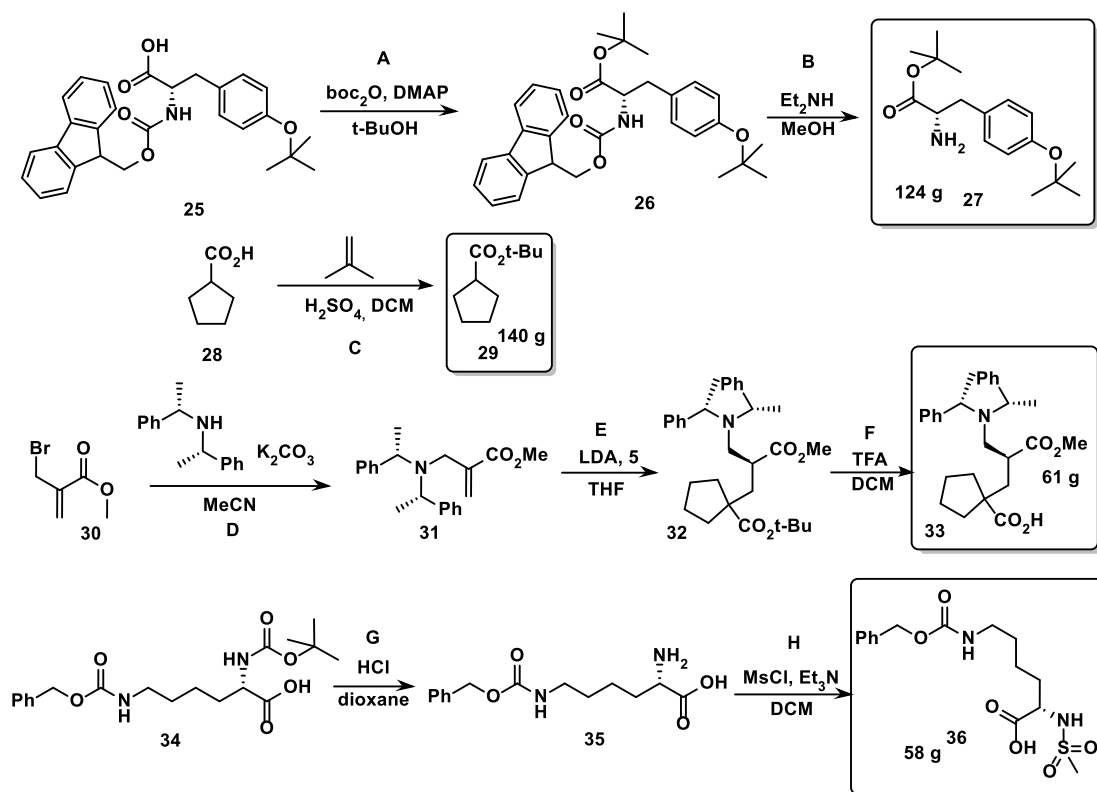
**Стадія А.** До колби з круглим дном додавали амінокислоту (1,0 екв.) у сірчаній кислоті (0,5 М у воді; 0,25 М відносно амінокислоти). Суміш охолоджували до 0 °С, після чого крапельно додавали водний розчин натрій нітриту (NaNO<sub>2</sub>) (5,0 екв.; 0,67 М відносно амінокислоти). Реакційну суміш залишали при кімнатній температурі на ніч для повного завершення реакції. Після завершення реакції суміш розбавляли водою та тричі екстрагували етилацетатом (EtOAc). Об'єднані органічні шари промивали насиченим розчином хлориду натрію, висушували над безводним сульфатом магнію (MgSO<sub>4</sub>), після чого розчинник видаляли під зниженим тиском. Отриманий сирий продукт використовували на наступному етапі без додаткового очищення.

Сирий продукт розчиняли в метанолі (MeOH, 1 М), до розчину крапельно додавали тіонілхлорид (SOCl<sub>2</sub>, 1,2 екв.). Реакційну суміш нагрівали до 65 °С протягом 1 години. Після охолодження до кімнатної температури додавали воду, а реакційну суміш тричі екстрагували діетиловим ефіром (Et<sub>2</sub>O). Об'єднані органічні фази промивали насиченим сольовим розчином, висушували над MgSO<sub>4</sub> та упарювали розчинник під зниженим тиском. Отриману суміш очищували колонковою хроматографією на силікагелі (SiO<sub>2</sub>) із використанням заданої комбінації розчинників.

## 1.6. Синтез сполуки VI.

Для синтезу сполуки 6 нами були успішно адаптовані, масштабовані та

валідовані низка літературних методів, що привело до синтезу ключових інтермедіатів 27, 29, 33 та 36 (Схема 1.6.1) у кількостях понад 100 г та понад 50 г відповідно [26,77,76,22].



**Схема 1.6.1.** Синтетична схема синтезу прекурсорів 27, 29, 33, 36.

Надалі ми намагалися використати отримані вихідні сполуки в синтезі кінцевої сполуки (Схема 1.6.2.), проте, оптимізувати стадію К з використанням прекурсору 36 нам не вдалося. Тому подальші дослідження сполуки VI були припинені через синтетичні труднощі та відсутність перспектив їх подолання.

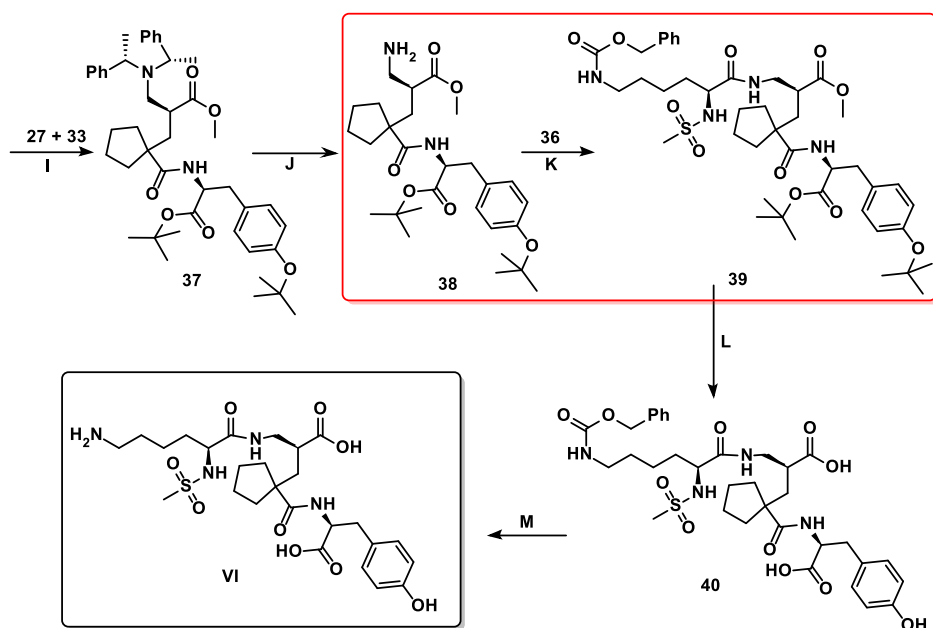


Схема 1.6.2. Синтетична схема синтезу сполуки VI.

### Синтез сполуки VI (вдалі спроби синтезу інтермедіатів).

**Стадія А.** Речовину **25** (300 г, 0.653 моль, 1 екв) розчинили у *t*-BuOH (1.5 л), додали DMAP (79.8 г, 0.653 моль, 1 екв), охолодили до 0 °С та порціями додали  $\text{Woc}_2\text{O}$  (157 г, 0.718 моль, 1.1 екв), підтримуючи температуру в діапазоні 0–5 °С. Отриману суміш залишили перемішуватись за кімнатної температури протягом 16 годин, після чого частково сконцентрували у вакуумі, розбавили водою (2 л), та екстрагували  $\text{CH}_2\text{Cl}_2$  (2 × 2 л). Об'єднані органічні екстракти промили 10% водним розчином HCl (1 л), висушили над  $\text{Na}_2\text{SO}_4$  та випарили у вакуумі. Вихід 246 г (73%) [26].  $^1\text{H NMR}$  (500 МГц,  $\text{CDCl}_3$ )  $\delta$  7.81 – 7.52 (м, 4H), 7.38 (т,  $J = 7.4$  Гц, 2H), 7.29 (кв,  $J = 6.4, 5.6$  Гц, 2H), 7.03 (д,  $J = 8.1$  Гц, 2H), 6.88 (д,  $J = 7.9$  Гц, 2H), 5.32 (д,  $J = 8.1$  Гц, 1H), 4.59 (с, 1H), 4.54 – 4.45 (м, 1H), 4.37 (дт,  $J = 44.8, 9.8$  Гц, 1H), 4.24 – 4.06 (м, 1H), 3.02 (д,  $J = 7.0$  Гц, 2H), 1.45 (с, 4H), 1.38 (с, 5H), 1.30 (с, 4H), 1.25 (с, 5H).

**Стадія В.** Речовину **26** (246 г, 0.477 моль, 1 екв) розчинили у MeOH (2 л) та додали  $\text{Et}_2\text{NH}$  (69.8 г, 0.954 моль, 2 екв), після чого перемішували за кімнатної температури протягом 16 годин. Реакційну суміш сконцентрували у вакуумі та очистили за допомогою флеш-хроматографії (градієнт гексан-EtOAc). Вихід 124 г (89%) [26].  $^1\text{H NMR}$  (400 МГц,  $\text{CDCl}_3$ )  $\delta$  7.11 – 7.05 (м, 2H), 6.95 – 6.84 (м, 2H), 3.55 (дд,  $J = 7.4, 6.0$  Гц, 2H), 3.06 – 2.75 (м, 1H), 1.46 – 1.39 (м, 2H), 1.37 (с, 9H),

1.30 (с, 9H).

**Стадія С.** Речовину **28** (150 г, 1.31 моль, 1 екв) розчинили у  $\text{CH}_2\text{Cl}_2$  (2 л) та додали  $\text{H}_2\text{SO}_4$  (5 мл), після чого розчин продували газоподібним ізобутеном протягом 2 год. Реакційну суміш промили насиченим водним розчином  $\text{Na}_2\text{CO}_3$ , висушили над  $\text{Na}_2\text{SO}_4$  та випарили у вакуумі. Вихід 140 г (63%) [22].  $^1\text{H NMR}$  (500 МГц,  $\text{CDCl}_3$ )  $\delta$  2.68 – 2.55 (м, 1H), 1.89 – 1.77 (м, 2H), 1.78 – 1.67 (м, 2H), 1.70 – 1.60 (м, 2H), 1.58 – 1.49 (м, 2H), 1.42 (с, 6H), 1.42 (с, 3H).

**Стадія D.** Речовину **30** (100 г, 0.559 моль, 1 екв) розчинили у MeCN (1 л) та додали  $\text{K}_2\text{CO}_3$  (154 г, 1.12 моль, 2 екв). Біс(1-фенілетил)амін (138 г, 0.615 моль, 1.1 екв) додали у декілька порцій, після чого розчин перемішували при  $50^\circ\text{C}$  протягом 8 год. Реакційну суміш сконцентрували у вакуумі, розбавили *t*-BuOMe (1 л), осад відфільтрували. Фільтрат висушили над  $\text{Na}_2\text{SO}_4$  та випарили у вакуумі. Вихід 132 г (73%) [22].  $^1\text{H NMR}$  (400 МГц,  $\text{CDCl}_3$ )  $\delta$  7.35 – 7.16 (м, 10H), 6.25 (кв,  $J = 1.9$  Гц, 1H), 6.18 (кв,  $J = 2.1$  Гц, 1H), 3.93 (кв,  $J = 6.9$  Гц, 2H), 3.69 (с, 3H), 3.68 – 3.62 (м, 1H), 3.26 (дт,  $J = 17.7$ , 1.8 Гц, 1H), 1.30 (д,  $J = 6.9$  Гц, 6H).

**Стадія E.** Розчин речовини **31** (69.5 г, 0.408 моль, 1 екв) у ТГФ (500 мл) порціями додали до 1 М розчину LDA (428 мл, 0.428 моль, 1.05 екв) при  $-78^\circ\text{C}$  та перемішували протягом 1 год. Речовину **7** (132 г, 0.408 моль, 1 екв) додали по краплях при  $-78^\circ\text{C}$ , після чого розчин перемішували при  $-78^\circ\text{C}$  протягом 2 год. Реакційну суміш обережно нейтралізували насиченим розчином  $\text{NH}_4\text{Cl}$  та частково сконцентрували у вакуумі. Отриману суспензію додатково розбавили водою (1 л) та екстрагували *t*-BuOMe ( $2 \times 1$  л), висушили над  $\text{Na}_2\text{SO}_4$  та випарили у вакуумі. Вихід 103 г (51%) [22].  $^1\text{H NMR}$  (400 МГц,  $\text{CDCl}_3$ )  $\delta$  7.30 – 7.14 (м, 10H), 3.96 (кв,  $J = 7.0$  Гц, 2H), 3.41 (с, 3H), 2.82 – 2.65 (м, 3H), 2.06 – 1.97 (м, 2H), 1.92 – 1.78 (м, 2H), 1.68 – 1.52 (м, 4H), 1.42 (с, 9H), 1.38 (с, 3H), 1.36 (с, 3H), 1.34 – 1.19 (м, 2H).

**Стадія F.** Речовину **32** (103 г, 0.209 моль, 1 екв) розчинили у  $\text{CH}_2\text{Cl}_2$  (1 л) та додали порціями 50 мл  $\text{CF}_3\text{CO}_2\text{H}$  після чого розчин перемішували за кімнатної температури протягом 16 год. Реакційну суміш випарили у вакуумі. Вихід 61 г (67%) [22].  $^1\text{H NMR}$  (500 МГц,  $\text{CDCl}_3$ )  $\delta$  7.42 – 7.10 (м, 10H), 3.93 (кв,  $J = 7.0$  Гц,

2H), 3.36 (с, 3H), 2.77 – 2.63 (м, 3H), 2.04 – 1.91 (м, 4H), 1.47 – 1.35 (м, 4H), 1.35 (с, 3H), 1.33 (с, 3H), 1.31 – 1.25 (м, 2H) (Сигнал ОН-групи не спостерігається через обмін).

**Стадія G, H.** Речовину **34** (150 г, 0.394 моль, 1 екв) додали порціями до розчину HCl у 1,4-діоксані та перемішували протягом 16 год. Реакційну суміш випарили у вакуумі та перекристалізували з *t*-BuOMe. Вихід 76 г (69%). Речовину **35** (76 г, 0.271 моль, 1 екв) суспендували у CH<sub>2</sub>Cl<sub>2</sub> (1 л), додали Et<sub>3</sub>N (41.1 г, 0.407 моль, 1.5 екв), охолодили до 0 °С та по краплях додали MsCl (34.2 г, 0.298 моль, 1.1 екв), підтримуючи температуру в діапазоні 0–5 °С. Отриману суміш залишили перемішуватись за кімнатної температури протягом 16 годин, після чого розбавили водою (1 л), та екстрагували CH<sub>2</sub>Cl<sub>2</sub> (2 × 1 л). Об'єднані органічні екстракти промили 10% водним розчином HCl (1 л), висушили над Na<sub>2</sub>SO<sub>4</sub> та випарили у вакуумі. Вихід 58 г (60%). <sup>1</sup>H NMR (400 МГц, CDCl<sub>3</sub>) δ 7.33 (с, 5H), 6.46 (с, 3H), 4.10 (с, 1H), 3.19 (с, 2H), 2.94 (с, 3H), 1.94 – 1.68 (м, 2H), 1.56 – 1.44 (м, 4H) (NH та ОН-групи перебувають в обмінному процесі).

**Стадія I.** Речовину **27** (5.00 г, 11.4 ммоль, 1 екв) розчинили у CH<sub>2</sub>Cl<sub>2</sub> (50 мл), додали Et<sub>3</sub>N (10 мл), EDC (2.63 г, 0.137 моль, 1.2 екв), HOBT (0.309 г, 2.29 ммоль, 0.2 екв) та речовину речовину **33** (3.35 г, 11.4 ммоль, 1 екв). Реакційну суміш промили 10% водним розчином HCl (50 мл), висушили над Na<sub>2</sub>SO<sub>4</sub> та випарили у вакуумі. Вихід 3.3 г (40%)

**Стадія J.** Речовину **37** (3 г, 4.21 ммоль, 1 екв) розчинили у MeOH (50 мл), додали Pd/C (0.5 г) та залишили перемішуватись під тиском водню (50 атм) протягом 16 год. Каталізатор відфільтрували, фільтрат випарили під вакуумом. Вихід 1.7 г (80%)

В результаті виконаної синтетичної роботи в мультиграмових масштабах вдалося одержати лише дві цільові сполуки — **I** (MLN-4760) та **II**. Незважаючи на обмежену кількість масштабованих синтезів, отримані об'єми є достатніми для подальшої біологічної валідації, проведення детальних фармакологічних досліджень і підтвердження концепції селективного інгібування АПФ2.

### РОЗДІЛ III. СТВОРЕННЯ ХІМІЧНОГО ПРОСТОРУ ТА ПОШУК НОВИХ СЕЛЕКТИВНИХ ІНГІБІТОРІВ АПФ2

Дослідження сучасних біологічних наслідків інгібування каталітичної активності АПФ2 є складним через обмежені дані про зв'язування малих органічних молекул з АПФ2. Ця частина дослідження була спрямована на ідентифікацію найбільш перспективних молекул-кандидатів, які відповідають сучасним критеріям хімічного зонда для АПФ2-рецепторів. Для цього було застосовано віртуальний скринінг *in silico*, поєднуючи методи молекулярного докінгу та молекулярної динаміки. Використовувалися комерційно доступні колекції, такі як Enamine Screening Collection з 3,2 мільйона сполук, та база даних синтетично доступних віртуальних сполук Enamine REAL Database (RDB) [112], що містить 1,4 мільярда сполук.

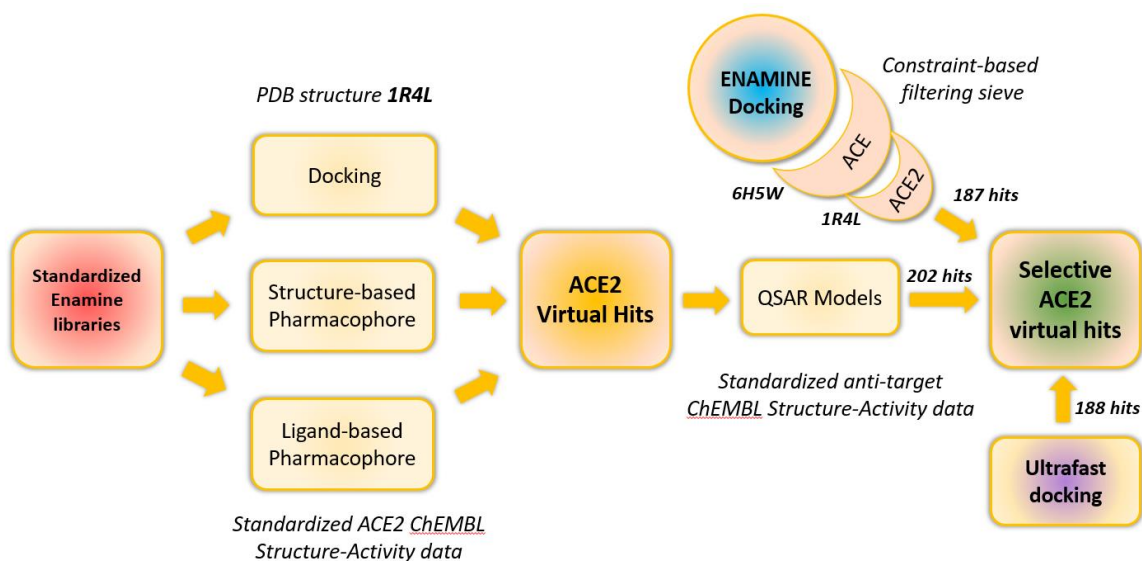
Для моделювання блокувальної дії АПФ2 як ферменту необхідно заблокувати його активний центр, і найкращою відправною точкою для цього слугує рентгеноструктурна модель із ко-кристалізованим інгібітором у активному сайті. У нашому дослідженні ми використали відповідні PDB-структури мішеней: АПФ2 (PDB ID: 1R4L) та АПФ (PDB ID: 1O86).

Незважаючи на високу гомологію між АПФ2 та АПФ — зокрема, приблизно 42% ідентичності їх каталітичних доменів — функціонально ці білки виконують протилежні ролі. АПФ2 не інгібується класичними інгібіторами АПФ, такими як лізиноприл, каптоприл чи еналаприлат [106]. Функціонально АПФ сприяє утворенню ангіотензину II, тоді як АПФ2 відповідає за його перетворення в ангіотензин 1–7, тим самим стримуючи надмірну активність РААС. Отже, АПФ2 діє як фізіологічний антагоніст АПФ.

У цьому контексті ми сфокусувалися на пошуку високоселективних інгібіторів саме для АПФ2, оскільки вони можуть чинити ефекти, прямо протилежні до тих, які спостерігаються при інгібуванні АПФ. Блокада обох білків водночас може спотворити модель дії SARS-CoV-2, тому АПФ у нашому дослідженні було визначено як анти-мішень, з точки зору селективності лігандів.

Для дослідження взаємодії сполук з білковими мішенями та виявлення активних кандидатів було застосовано три незалежні обчислювальні підходи (див. Мал.12):

- Група **UNISTRA** реалізувала комбінований підхід, що включав фармакофорне моделювання, молекулярний докінг і QSAR-аналіз. Це дозволило інтегрувати фармакофорні характеристики, здійснити структурний докінг та провести кількісну оцінку взаємодій на основі прогнозованої активності.
- Група **ENAMINE** використала метод молекулярного докінгу із застосуванням адаптованої структури зв'язувальної кишені, попередньо оптимізованої за допомогою молекулярної динаміки. Такий підхід забезпечив урахування конформаційної гнучкості мішені та дозволив точніше змоделювати взаємодії лігандів.
- Група **ChemSpace** застосувала підхід “ультрашвидкого” докінгу, який поєднує моделі з різними рівнями точності. Завдяки верифікації докінгових комплексів за допомогою молекулярної динаміки цей метод дозволяє оперативно оцінити ймовірність формування стабільних комплексів та забезпечити достатній рівень достовірності результатів



Мал.12. Схематичне зображення трьох незалежних обчислювальних підходів для дослідження цільових молекул та ідентифікації потенційно-активних сполук.

### 3.1. Методи дослідження та програмне забезпечення

Група ENAMINE застосували програмне забезпечення ICM для проведення докінгу, використовуючи уточнену геометрію активного сайту «апо» структури АПФ2, отриману шляхом молекулярної динаміки без ліганду, але з

включенням водного середовища. У другій ітерації скринінгу був застосований пакет Schrödinger2024 (Maestro).

UNISTRA (Мал.12.) виконала скринінг фармакофорів на основі структурного підходу за допомогою LigandScout [116] і докінг за допомогою PLANTS, орієнтуючись на наявну рентгенівську структуру АПФ2 [107] (1R4L у PDB). Одночасно вони створювали фармакофори на основі лігандів АПФ2 (LigandScout) [116] і розробляли QSAR-моделі, використовуючи дескриптори ISIDA [110] для АПФ і NEP. Ці підходи були побудовані на основі відомих інгібіторів АПФ2 (та АПФ і NEP) з бази даних ChEMBL26 [33].

Група ChemSpace для створення вихідного набору сполук використовувала стандартні методи фільтрування за структурними критеріями, розрахунок дескрипторів, фільтрацію за допомогою коефіцієнта Танімото, генерування "decoys", молекулярний докінг і молекулярну динаміку. Дослідження проводилися за допомогою різних доступних програмних пакетів: Chimera [125] та PyMOL [97] для просторового зіставлення та 3D-візуалізації, Python у середовищі Jupyter з пакетом RDKit для роботи з бібліотеками і хемоінформатичних фільтрів, а також для генерування "decoys" (<http://dude.docking.org/generate>) [78]. Скринінг *in silico* здійснювався за допомогою пакету Molsoft 3.8.6 [1,2] у поєднанні з програмою візуалізації Chimera 1.14.

### **3.1.1. Методи UNISTRA**

**3.1.1.1. Група UNISTRA** використовувала PLANTS [63] для виконання докінгу на основі структури PDB 1R4L ферменту, визначаючи центр активного сайту як геометричний центр співкристалізованого ліганду та використовуючи радіус обмеження 12 Å. Оцінка ChemPLP використовувалася на "швидкості 1" (найнижча швидкість - найвища точність). Попередній досвід з PLANTS/ChemPLP показав, що ChemPLP залежить від розміру сполуки (більші молекули систематично отримують більш негативні оцінки) і що ця залежність найкраще може бути емпірично охарактеризована припущенням, що ChemPLP пропорційний  $N^{-0.35}$ , де  $N$  — кількість важких атомів у молекулі. Тому хіти

докінгу були вибрані та ранжовані за модифікованою, незалежною від розміру ChemPLP  $\times N^{-0.35}$ , а не за замовчуванням ChemPLP.

**3.1.1.2. Фармакофорна модель на основі лігандів UNISTRA.** За допомогою інструменту LigandScout було проведено накладення 51 звіту з бази ChEMBL щодо зв'язувачів АПФ2, які мали значення pKi або pIC<sub>50</sub> вище 7. У результаті було виділено кластер із семи структурно споріднених молекул, серед яких також були відомі селективні інгібітори АПФ2 порівняно з АПФ.

На основі цього кластеру був сформований фармакофор, який надалі використовувався для віртуального скринінгу. Перед запуском моделі були навмисно виключені фармакофорні ознаки, що відповідають за хелатування іонів цинку, а також аніонні функції — вони були переосмислені як прості акцептори водневого зв'язку.

Унаслідок такої спрощеної конструкції фармакофору його селективність, ймовірно, знизилася, що може призвести до підвищення частоти помилкових негативних результатів. Водночас такий підхід дозволяє виявляти нові потенційні зв'язувачі з меншою кількістю зарядів — а отже, з потенційно кращими фармакокінетичними властивостями, включно з проникністю через гематоенцефалічний бар'єр (BBB) [83].

### **3.1.1.3. Структурна фармакофорна модель UNISTRA**

Фармакофорна модель була автоматично згенерована в LigandScout на основі рентгенівської структури комплексу АПФ2 з інгібітором (PDB ID: 1R4L). Як і в попередньому випадку, модель була модифікована шляхом видалення фармакофорних ознак, відповідальних за хелатування цинку, а також заряджених функціональних груп. Співкристалізований ліганд MLN-4760 містив дві карбоксилатні групи, які в спрощеній моделі були представлені як звичайні акцептори водневих зв'язків, а також протоновану вторинну аміногрупу, інтерпретовану як донор водневого зв'язку. Така спрощена інтерпретація дозволила створити фармакофор, орієнтований на більш загальні взаємодії, з потенційною перевагою для пошуку молекул з поліпшеними фармакокінетичними властивостями.

#### **3.1.1.4. Категоріальні моделі QSAR проти цілей UNISTRA**

QSAR-моделі класу «структура–активність» були побудовані для антимішеней АПФ та NEP відповідно до стандартної еволюційної процедури моделювання [48], що базується на алгоритмі машини підтримки векторів (SVM, Support Vector Machines), зокрема його класифікаційній версії. Методологія включала оптимізацію простору дескрипторів шляхом відбору найінформативніших схем фрагментації з-поміж сотні варіантів, запропонованих у рамках платформи ISIDA.

Найкращі індивідуальні моделі були відібрані за допомогою еволюційного алгоритму, а їх якість оцінювалася за допомогою інтегрального балу, що відображав середню збалансовану точність ( $BA_{XV}$ ) у 12-разовій схемі перехресної валідації типу leave-1/3-out. Найбільш точні моделі, побудовані на основі різних наборів ISIDA-дескрипторів, були інтегровані в консенсусні предиктори для кожної з антимішеней.

Ці консенсусні моделі використовувалися як фінальний фільтр для виключення зі списку кандидатів тих сполук, які передбачалися як активні відносно АПФ або NEP. Таким чином, здійснювалося уточнення (тобто підвищення селективності) вихідного списку потенційних лігандів АПФ2, відібраних на етапах фармакофорного скринінгу та докінгу.

Для всіх моделей на основі дескрипторів фрагментів ISIDA використовувалася критерій "обмежувальної рамки" як проста область застосування (O3) [110]. Це означає, що для можливості надійного прогнозування зовнішня сполука повинна мати значення фрагментних підрахунків у межах діапазонів, що охоплюються навчальним набором.

#### **3.1.2. Методи групи Enamine**

У цьому дослідженні були використані рентгенівські структури мішеней АПФ2 (PDB ID: 1R4L) та АПФ (PDB ID: 1O86). На першому етапі білкові структури було імпортовано у формат ICM як хімічні об'єкти [2], а потім видалили всі молекули води. До білків було емпірично додано атоми водню, після чого їх положення було оптимізовано разом зі структурою білка за

допомогою вбудованого інструменту енергетичної мінімізації. Окрему увагу приділено орієнтації залишків HIS, PRO, ASN, GLN та CYS. Для гістидину було перевірено три варіанти протонування, а також два можливі оберти імідазольного кільця. Залежно від оптимізованого стану, залишок перейменовували відповідно до таутомерної форми: HIE ( $\epsilon$ -протонований) або HIP (позитивно заряджений). Аналогічно, орієнтації амідних груп у залишках аспарагіну (ASN) та глутаміну (GLN) були скориговані для забезпечення коректного відображення водневих зв'язків у подальшому моделюванні.

### 3.1.2.1. Молекулярна динаміка

Пакет Gromacs [4] 2019 року використовувався для проведення MD та подальшого аналізу результатів. Енергетична мінімізація білок-лігандних комплексів здійснювалася з використанням повноатомного силового поля AMBER99 за алгоритмом градієнтного спуску (20 000 кроків) [86]. Етап еквилібрування середовища проводився за умов фіксації положень білкових атомів, при цьому молекули розчинника (вода TIP3P) залишалися вільно рухомими протягом 100 пс симуляції. Кожну систему розміщували в ортогональну коробку з водним шаром товщиною 10 Å, після чого проводили нейтралізацію та іонне балансування додаванням  $\text{Na}^+$  та  $\text{Cl}^-$  до концентрації 150 мМ.

Симуляції здійснювались в умовах періодичних граничних умов. Для підтримки сталої температури (310 K) застосовувався термостат v-rescale, а для контролю тиску (1 бар) — баростат Parrinello–Rahman [84,13]. Довгодіючі електростатичні взаємодії розраховувалися за методом Particle-Mesh Ewald [3]. Топологічні параметри лігандів були згенеровані за допомогою модуля Antechamber із пакета AmberTools18. Для MD-аналізу використовувалися як зв'язані комплекси лігандів з АПФ2 (PDB ID: 1R4L) та АПФ (PDB ID: 6H5W), так і відповідні апо-форми (ті ж білки без лігандів). Кожну з систем моделювали протягом 150 нс.

Аналіз конформаційної поведінки проводився шляхом кластеризації траєкторій, використовуючи функціонал GROMACS [88]. Для кластеризації

було розраховано середньоквадратичні відхилення (RMSD) атомів Ca, із заданою відстанню між кластерами 0,15 нм.

### **3.1.2.2. Докінг Enamine до кількох конформерів активного сайту**

Кристалічні структури білків АПФ і АПФ2 були отримані з бази даних RCSB PDB. Вирівнювання амінокислотних послідовностей і надструктурний аналіз проводилися за допомогою вбудованих інструментів платформи ICM Molsoft, призначених для біоінформатичних задач.

Для молекулярного докінгу були обрані білкові конформації, отримані в результаті MD-симуляцій, які підготовлювались як активні сайти в середовищі ICM. Було побудовано грід-карти (grid maps) у межах прямокутної області з розміром сітки 0,5 Å, орієнтованої на центр зв'язувального сайту. Референтні ліганди були зафіксовані у відповідних положеннях, а вся бібліотека для віртуального скринінгу (Enamine In-Stock Collection, 3,2 мільйона сполук) була перетворена у формат індексованої бази даних із 3D-конформерами.

Гнучкий докінг здійснювався методом глобальної оптимізації Монте-Карло, який оперує внутрішніми координатами ліганду в просторі попередньо розрахованих грід-карт [2]. Зважаючи на наявність металевого іона у зв'язувальній кишені, рівень точності моделювання було підвищено до 3, що дозволило покращити адекватність результатів, хоч і збільшило тривалість розрахунків.

Оцінка та ранжування положень зв'язування здійснювались згідно з енергетичною функцією ICM scoring function, яка враховує низку параметрів, що визначають якість комплексу:

- Ван-дер-Ваальс взаємодії між лігандом і мішенню;
- Внутрішня енергія ліганду, яка впливає на його здатність до зв'язування;
- Зміна конформаційної енергії ліганду при зв'язуванні;
- Енергія водневих зв'язків, які стабілізують комплекс;
- Енергія десольватації донорно-акцепторних груп у водневих зв'язках;
- Електростатичні ефекти сольватації;

- Гідрофобний ефект — приріст вільної енергії за рахунок гідрофобної взаємодії;
- Корекція на розмір молекули (ligand efficiency), що нормалізує енергію зв'язування на кількість важких атомів у ліганді.

Сукупність цих параметрів дозволяє детально оцінити не лише енергію зв'язування, а й потенційну стабільність білок-лігандного комплексу [79].

### **3.1.3. Методи групи ChemSpace**

**3.1.3.1. Молекулярну динаміку** розраховували за допомогою програмного забезпечення Gromacs 2019 (версія 2019.5) [4]. Для моделювання розчинника використовувалася модель TIP4P, форма коробки була ромбічною з розмірами  $10 \text{ \AA} \times 10 \text{ \AA} \times 10 \text{ \AA}$ , а силове поле — charmm36. Нейтралізацію системи здійснювали шляхом додавання іонів  $\text{Na}^+$ . Час моделювання склав 100 нс, кількість кадрів — 1000. Модельна система була розслаблена з часом релаксації 1 пс. Ансамбль використовувався типу NPT з температурою 300 К і тиском 1,01325 бар. Метод термостата — ланцюг Нозе-Гувера. Радіус відсікання для взаємодій становив 9  $\text{ \AA}$ .

В другій частині дослідження використовувався пакет Desmond від Schrödinger [12]. Для моделювання розчинника використовувалася модель TIP4P, форма коробки була ромбічною з розмірами  $10 \text{ \AA} \times 10 \text{ \AA} \times 10 \text{ \AA}$ , а силове поле. Застосовувалася постідійний протокол який включає поетапний “підігрів” системи - 100нс ансамбль NVT при 250К, 100нс ансамбль NPT при 250К, 100нс ансамбль NPT при 300К, 100нс ансамбль NVT при 300К.

### **3.1.3.2. Молекулярний докінг групи ChemSpace**

Рентгеноструктурну модель цільового білка було обрано для докінгу на основі кількох ключових критеріїв: висока роздільна здатність (у межах 1,5–2,5  $\text{ \AA}$ ), наявність зв'язаного ліганду в кристалічній структурі, приналежність до організму *Homo sapiens*, актуальність дати публікації та авторитетність джерела (рейтинг журналу).

Підготовка структури розпочалася з конвертації білка у внутрішній формат ICM — специфічний формат молекулярного об'єкта програмного

середовища MOLSOFT. Після видалення всіх молекул води було застосовано інструмент "Optimize Hydrogen", що дозволив уточнити орієнтацію та протонування залишків HIS, PRO, ASN, GLN і CYS. Протонування структури проводилося для умов фізіологічного середовища (pH 7,4 ± 0,2).

Подальша енергетична мінімізація виконувалася з використанням силового поля Merck Molecular Force Field (MMFF), при цьому допустиме зміщення важких атомів не перевищувало 0,3 Å. На основі підготовленої білкової моделі були побудовані та оптимізовані докінгові об'єкти для білків АПФ2 та АПФ із урахуванням особливих параметрів, необхідних для подальших віртуальних скринінгів.

### 3.1.4. Методи другої ітерації

#### 3.1.4.1. USRCAT

USRCAT (**Ultrafast Shape Recognition with CAlytic constraints**) [96] використовується для надшвидкого молекулярного порівняння за формою і є ефективним інструментом для віртуального скринінгу великих хімічних бібліотек. Це безвирівнювальний підхід, який обчислює дескриптори форми на основі розподілу атомів у просторі, що робить його суттєво швидшим порівняно з традиційними методами, які потребують попереднього вирівнювання молекул.

USRCAT особливо ефективний у завданнях scaffold hopping, де необхідно виявити нові структури зі схожою просторовою конфігурацією, але відмінною хімічною будовою. Порівняно з базовим алгоритмом USR, версія USRCAT доповнена фармакоформними ознаками, що надає порівнянню додаткового біологічного контексту.

У попередніх дослідженнях цей підхід успішно застосовувався нами для генерації таргетних бібліотек [66] із великого віртуального хімічного простору. Тут основною метою було **знайти сполуки з подібною 3D-формою до відомих активних лігандів**, водночас забезпечуючи хімічну різноманітність і виявлення нових потенційно активних скелетів. Тому цей метод також спробували використати і для пошуку інгібіторів АПФ2.

#### 3.1.4.2. Докінг Schrodinger

Докінг у Schrödinger реалізується за допомогою програмного комплексу Glide, який є одним із провідних інструментів для молекулярного докінгу. Glide використовує багатоетапний алгоритм, що включає попередню генерацію конформацій лігандів, точне визначення зв'язувального сайту та розміщення лігандів у просторі білка з подальшим енергетичним оцінюванням. Оцінка зв'язування здійснюється з урахуванням ван-дер-ваальсових взаємодій, електростатичних ефектів і впливу розчинника. Glide підтримує кілька режимів докінгу: HTVS, Standard Precision (SP), High Precision (XP), що дозволяє враховувати конформаційні зміни білка під час взаємодії з лігандом. Однією з ключових особливостей є використання функцій оцінки Glide scoring function, які оптимізуються на основі експериментальних даних, що забезпечує високу точність прогнозування.

#### 3.1.4.3. AutoQSAR

AutoQSAR – це автоматизований інструмент для розробки моделей кількісного співвідношення «структура-активність» (QSAR), який входить до складу програмного пакету Schrödinger. Даний метод дозволяє швидко створювати та оптимізувати предиктивні моделі на основі великих наборів молекулярних дескрипторів. AutoQSAR використовує різні алгоритми машинного навчання, включаючи випадкові ліси, опорні вектори та нейронні мережі, для виявлення закономірностей між хімічною структурою та біологічною активністю. Метод автоматично відбирає найбільш інформативні дескриптори, оптимізує параметри моделі та оцінює її якість за допомогою крос-валідації. Однією з ключових переваг AutoQSAR є його здатність обробляти великі набори даних та генерувати високоточні прогностичні моделі для віртуального скринінгу нових біологічно активних сполук. У нашому дослідженні була спроба застосування цього підходу для аналізу та ранжування потенційних інгібіторів АПФ2 на основі наявних даних в базі ChEMBL, але через однотипність даних, ми не отримали достатньо якісної модулі, яку можна було б застосовувати для передбачення активності сполук для подальших експериментальних перевірок.

#### 3.1.4.4. FTMap

FTMap (Functional Topology Mapping) — це інструмент для аналізу взаємодій між молекулами в рамках комп'ютерного молекулярного моделювання. Його основна мета полягає в ідентифікації функціональних елементів на поверхні молекул, що можуть бути потенційними мішенями для лігандів. FTMap допомагає виявляти потенційні місця зв'язування лігандів, використовуючи інформацію про топологію молекули та її поверхневі характеристики [56].

Інструмент працює за допомогою сканування молекул, враховуючи різні фізико-хімічні властивості поверхні, такі як гідрофобність, електростатичні взаємодії, і можливі водневі зв'язки. Він застосовує алгоритми для виявлення найбільш інтерактивних ділянок на молекулі, які можуть взаємодіяти з іншими молекулами, такими як ліганди або інші макромолекули.

FTMap є корисним інструментом для пошуку молекулярних зондів, оскільки дозволяє виявити найбільш перспективні ділянки молекул, які можуть бути використані для подальшої модифікації або оптимізації з метою підвищення їх ефективності. Цей інструмент може застосовуватися як на етапах віртуального скринінгу, так і для вдосконалення вже знайдених сполук [65].

#### 3.1.4.5. FTMove

FTMove (Functional Topology Move) [29] — це автоматизований веб-сервер, розроблений для виявлення та аналізу криптичних та алостеричних сайтів зв'язування на поверхні білків шляхом обробки множинних структурних конформацій. Основна мета FTMove полягає у визначенні ділянок білка, які можуть змінювати свою конформацію та ставати доступними для зв'язування лігандів, що є важливим для розробки нових лікарських засобів.

#### **Принцип роботи FTMove:**

1. **Вхідні дані:** Користувач надає структуру білка у форматі PDB або вводить відповідний PDB-код та ідентифікатор ланцюга;

2. **Пошук подібних структур:** Сервер ідентифікує всі доступні в PDB структури з подібною послідовністю ( $\geq 90\%$  ідентичності та покриття) (FTMove.bu.edu);

3. **Мапування гарячих точок:** Кожна з відібраних структур аналізується за допомогою алгоритму FTMap для виявлення гарячих точок зв'язування — регіонів, що сприяють зв'язуванню лігандів;

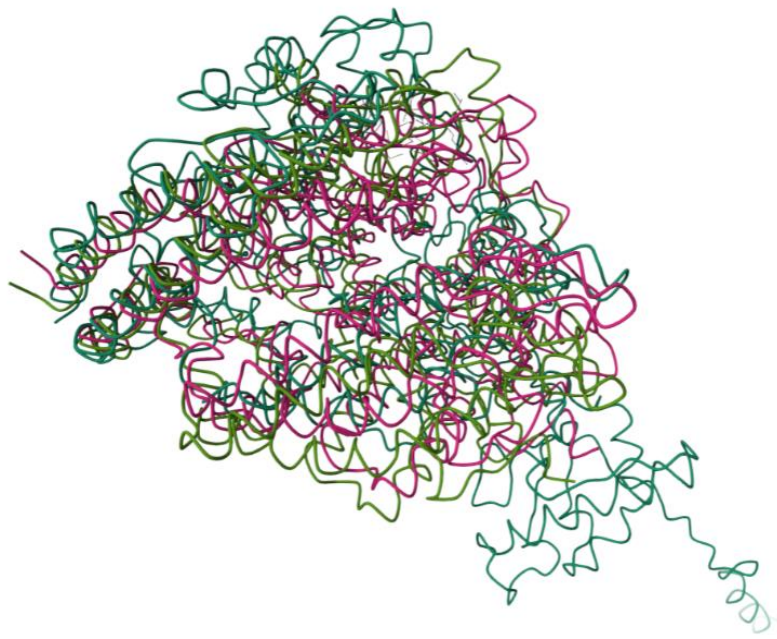
4. **Аналіз конформаційних змін:** Отримані результати об'єднуються для визначення консенсусних сайтів зв'язування, що дозволяє оцінити, як різні конформації білка впливають на доступність та силу сайтів зв'язування [29].

3.1.4.6. DiffDock [19] — це сучасна генеративна модель, розроблена для прогнозування тривимірної структури комплексу білок-ліганд, що є критичним етапом у процесі відкриття ліків. Вона використовує дифузійні моделі для передбачення зв'язування малих молекул з білками з високою точністю. DiffDock складається з двох моделей: Score та Confidence. Score-модель генерує потенційні пози зв'язування ліганду з білком, а Confidence-модель оцінює та ранжує ці пози за ймовірністю правильності. Цей підхід дозволяє ефективно ідентифікувати потенційні сайти зв'язування без попереднього знання про активні центри білка. Розробка та впровадження DiffDock здійснені командою дослідників з Массачусетського технологічного інституту (MIT) у співпраці з NVIDIA.

## 3.2. Збір та систематизація даних

### 3.2.1. Систематизація даних групи Chem-space

В базах <https://www.ebi.ac.uk/> та <https://swissmodel.expasy.org/> було знайдено 298 експериментальних структур АПФ2

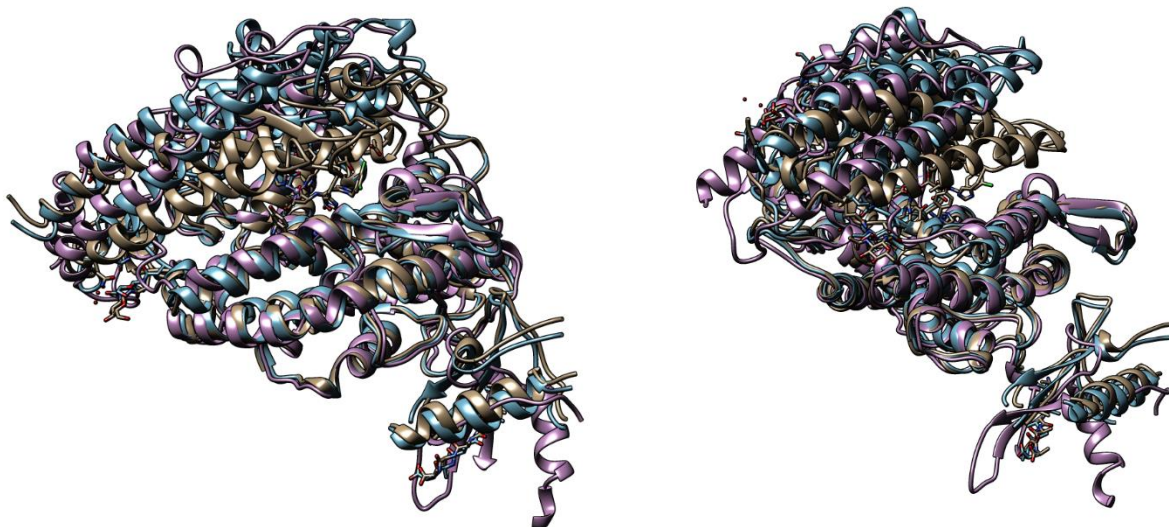


Мал.13. Кластеризація 3d структур протеїну АПФ2 на платформі PDBe-KB: 1r41 – рожевий колір, 1r42 - колір трави, 6m1d – смарагдовий (<https://www.ebi.ac.uk/pdbe/pdbe-kb/proteins/Q9BYF1>).

Платформа PDBe-KB виділяє 2 кластери (Мал.13.) - перший кластер містить всього 2 ланцюжка однієї кристалічної структури. Представник першого кластеру структура з PDB кодом 6m1d. Другий кластер містить всі інші 419 ланцюгів з 316 структур. Ми виділили два представники, які мають для роботи важливе значення - 1r41 (містить відомий активний ліганд MLN-4760, закрита форма) та 1r42 (відкрита форма). 1r41 та 1r42 були обрані як ключові представники серед доступних структур ACE2, оскільки на момент проведення дослідження вони були найбільш якісними і репрезентативними прикладами закритої та відкритої форм мішені відповідно.

- 1r41 (роздільна здатність 3.00 Å) — єдина структура ACE2 у закритій конформації, що містить ко-кристалізований відомий інгібітор MLN-4760;
- 1r42 (роздільна здатність 2.20 Å) — структура у відкритій формі з найліпшою роздільною здатністю, яка забезпечує високу якість та деталізацію активного сайту без присутності ліганду. Це дозволяє моделювати гнучкість білка та оцінювати доступність кишені для взаємодії з новими сполуками.

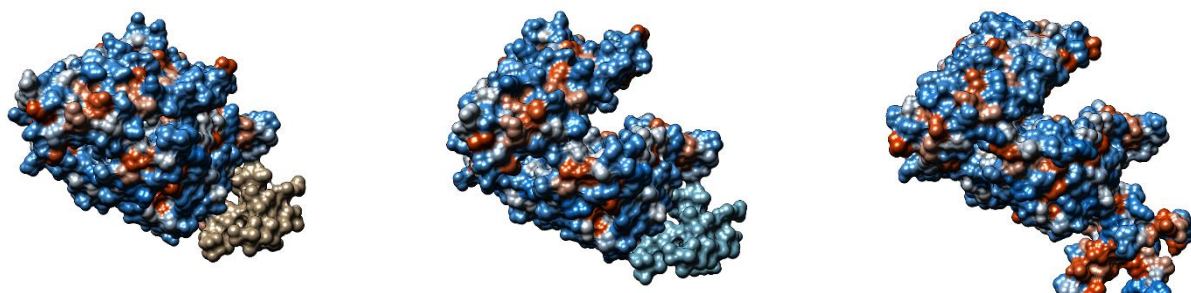
Також було проаналізовано та проведено порівняння перелічених трьох кристалічних структур в Chimera (Мал.14).



Мал.14.а) Порівняльне вирівнювання PDB структур Ir41 (в бежевому кольорі), Ir42 (в блакитному кольорі), 6t1d (в рожевому кольорі).

Мал.14. б) Проекція після горизонтальний розвороту на 45°.

Візуалізація та порівняння поверхні різних структур демонструє наявність двох форм протеїну АПФ2 (Мал.15)



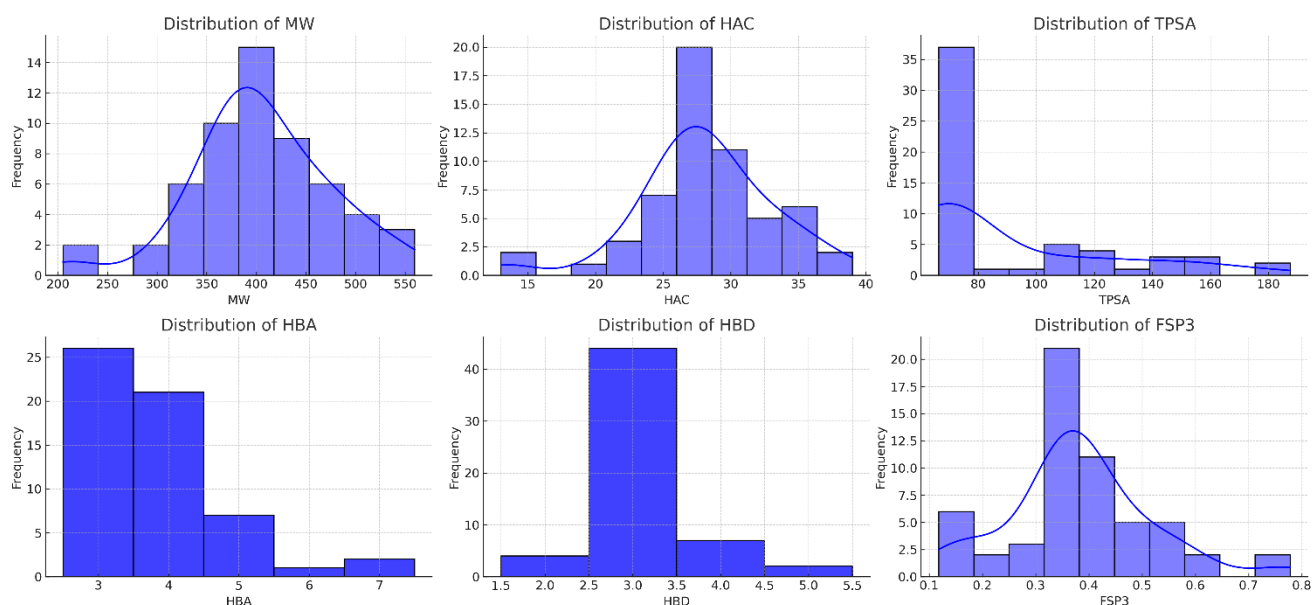
Мал.15. а) Ir41 (закрита форма)

б) Ir42 (відкрита форма)

в) 6t1d (відкрита форма)

Основним, але не єдиним джерелом даних для еталонних лігандів у цьому дослідженні є база даних ChEMBL (реліз 26) [33]. Наукові групи використовували три підгрупи сполук з відомими значеннями рКі або IC<sub>50</sub> щодо наступних мішеней: АПФ2 (65 унікальних структур з вказаними термодинамічними константами інгібування Кі), АПФ (580 унікальних структур) і Neprilysin (NEP) (678 структур).

Очевидно, що АПФ2 досліджено набагато менше, ніж два інші згадані ферменти. На схемах розподілу дескрипторів проаналізовані найшіжливіши медико-хімічні параметри відомих АПФ2 інгібіторів (Мал.16).



Мал.16. Схеми розподілу дескрипторів найшіжливіших медико-хімічних параметрів відомих АПФ2 інгібіторів.

### 3.2.2. Набір даних UNISTRA та їх попередня обробка

Кристалічні структури білків АПФ (PDB ID: 6H5W) та АПФ2 (PDB ID: 1R4L) були завантажені з бази RCSB Protein Data Bank та попередньо оброблені за допомогою інструменту VegaZZ [85]. На стадії "промивання" зі структур видалялися кристалізаційні молекули води, а відсутні атоми водню були додані відповідно до протонного стану амінокислот при фізіологічному рН.

Основним джерелом інформації про референтні ліганди та значення активності для білків АПФ2, АПФ і NEP слугувала база ChEMBL v26 [33]. Дані завантажувалися у вигляді окремих файлів і проходили стандартизацію згідно з оприлюдненим протоколом [50]. Після первинної обробки усувалися дублікати як із врахуванням стереохімії, так і без неї (залежно від подальшого застосування).

Для побудови фармакофорних моделей і проведення молекулярного докінгу розглядалися лише точні стереохімічні дублікати (включно з

енантіомерами) для молекул, щодо яких у базі ChEMBL були доступні значення термодинамічних констант інгібування ( $K_i$ ). Якщо для одного стереоізомеру існувало декілька експериментальних записів, активність розраховувалася як середнє значення  $pK_i$  за умови узгодженості даних (допустима розбіжність —  $\leq 0,5$  логарифмічної одиниці). За наявності більшої варіації, усі відповідні записи виключались із аналізу.

Для побудови QSAR-моделей категоріального типу застосовували фрагментні дескриптори ISIDA [110], які не враховують стереохімічну інформацію. Тому дублікати визначались на основі стандартизованих SMILES без стереохімії — тобто всі стереоізомери вважались ідентичними структурами. Такий підхід було однаково реалізовано для двох антимішеней — АПФ та NEP — з аналогічним протоколом обробки.

Категоризацію сполук за активністю здійснювали відповідно до автоматизованого алгоритму, який повторює процедуру, застосовану в ChEMBL для структури–активності-зв'язків [102]. Для АПФ, який має два функціональні активні центри (N- та C-домен), класифікація за схемою “активний/неактивний” не розрізняла локалізацію зв'язування — усі дані об'єднувались. У разі, якщо SMILES-структури були віднесені до різних класів активності в різних записах, такі сполуки виключались з подальшого аналізу.

### **3.3. РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ ВІРТУАЛЬНОГО СКРИНІНГУ**

#### **3.3.1. Загальний аналіз баз даних**

Пошук у базі ChEMBL v26 за мішенню людський АПФ2 (ChEMBL3736) виявив лише 102 сполуки, для яких доступні 108 експериментальних записів. Для порівняння:

- людський АПФ (ChEMBL1808) має асоційовані 1032 сполуки та 1346 тестових значень;
- неприлізин (NEP) (ChEMBL1944) — 673 сполуки та 806 тестів.

Ці цифри свідчать про суттєво менший обсяг досліджень, проведених щодо АПФ2, порівняно з іншими двома мішенями.

Після стандартизації структур та видалення дублікатів залишилось лише 60 унікальних записів, які можуть бути використані для побудови QSAR-моделей. Такий обмежений набір даних не дозволяє забезпечити достатню надійність прогнозів при використанні методів машинного навчання.

Хоча була створена модель SVM-регресії для передбачення значень  $pK_i$  по АПФ2, яка показала хорошу якість за результатами крос-валідації ( $Q^2 = 0.87$ ), її область застосування обмежується виключно хімічними класами, представленими в базі ChEMBL. Через це модель не використовувалась у процесі віртуального скринінгу, де пріоритетом є здатність до генералізації на нові хімічні простори.

**Проблема селективності та підхід до її вирішення.** Безпосереднє навчання моделі для прогнозування селективності між АПФ2/АПФ або АПФ2/NEP є неможливим у зв'язку з відсутністю повних експериментальних даних щодо активності лігандів одночасно проти обох мішеней. Для створення кількісної моделі селективності необхідно мати відповідні значення активності щодо кожної мішені, які можна було б перетворити на числовий показник селективності. У більшості випадків такі дані відсутні або неповні.

З огляду на це, проблему селективності було вирішено за допомогою альтернативного підходу:

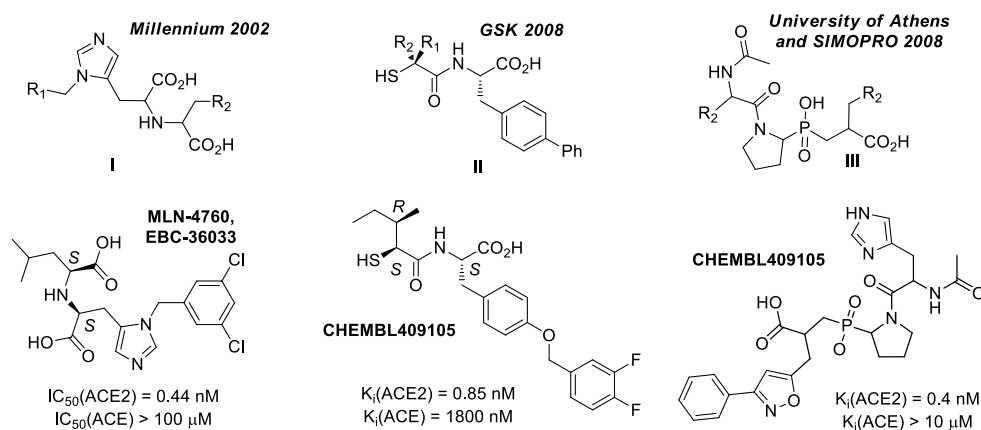
- Було побудовано окремі QSAR-моделі для АПФ та NEP на основі великих валідованих наборів експериментальних даних;
- Це дало змогу проводити негативний відбір, виключаючи з подальшого розгляду ті сполуки, які прогнозувалися як активні щодо АПФ або NEP, зберігаючи при цьому потенційні селективні ліганди для АПФ2.

**Відомі селективні зв'язувачі АПФ2.** Серед добре відомих селективних інгібіторів АПФ2 можна відзначити:

- ChEMBL429844 (MLN-4760, EBC-36033, EN300-22203532) – 0.44 нМ;
- ChEMBL251808 – 0.85 нМ;
- ChEMBL409105 – 0.4 нМ [24].

Аналіз показав, що всі ці сполуки походять із трьох незалежних дослідницьких проєктів, що представляють три різні хемотипи (Мал.17):

1. Поліпептидний міметик на основі гістидину (I) – розроблений Millennium Pharmaceuticals у 2002 році [21];
2.  $\alpha$ -Тіоламідний інгібітор (II) – запропонований GlaxoSmithKline у 2008 році [24];
3. Фосфінові пептиди, що містять псевдопролін (III) – розроблені в межах академічного співробітництва між Університетом Афін і SIMOPRO [74].



Мал.17. Основні хемотипи відомих селективних інгібіторів АПФ2 з відповідними хемотипами (I-III) та прикладами структур.

## Обмеження для застосування селективних АПФ2 інгібіторів *in vivo*

Попри високу селективність щодо АПФ2, багато з відомих інгібіторів містять високополярні функціональні групи, зокрема карбоксильні, які значно знижують їх здатність проникати через гематоенцефалічний бар'єр (BBB). Це робить такі сполуки непридатними для дослідження нейроінвазивного потенціалу SARS-CoV-2 та пов'язаних з ним неврологічних ускладнень. Варто зазначити, що на цьому етапі група ChemSpace не ставила за мету ідентифікацію саме BBB-проникних молекул, а фокусувалася на загальній селективності до мішені АПФ2.

**Прогнозування пасивної проникності через гематоенцефалічний бар'єр.** У випадку пасивної дифузії (тобто без участі активних транспортерів),

здатність молекули проникати через гематоенцефалічний бар'єр (BBB) може бути емпірично оцінена на основі таких параметрів:

- Молекулярна маса  $\leq 400$  Да;
- Загальна кількість донорно-акцепторних водневих зв'язків — менше 8;
- Ліпофільність (LogP) в межах 1.5–2.7;
- Полярна поверхнева площа (PSA) у діапазоні 60–70 Å<sup>2</sup>.

Дотримання цих критеріїв вважається сприятливим для пасивного проходження молекул через BBB [100,81].

Три з відібраних хімічних сполук з відомою активністю були подані для реєстрації на міжнародному ресурсі Chemical Probe Portal (<https://www.chemicalprobes.org/>), який є авторитетною базою даних молекулярних зондів:

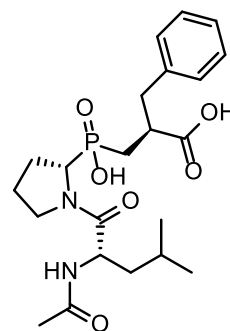
#### **EBC-36032**

Структура:

CC(C)CC(NC(=O)C)C(=O)N1CCCC1P(=O)(O)CC(CC=2C=CC=CC2)C(=O)O

Ідентифікатор ChEMBL: ChEMBL260273

Активність:  $K_i = 0.13$  nM

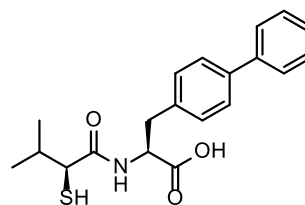


#### **EBC-44380**

Структура: CC(C)[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O

Ідентифікатор ChEMBL: ChEMBL269997

Активність:  $K_i = 1.5$  nM



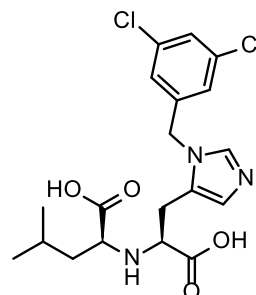
#### **ORE-1001**

Структура:

CC(C)C[C@H](N[C@@H](Cc1cncn1Cc1cc(Cl)cc(Cl)c1)C(=O)O)C(=O)O

Ідентифікатор ChEMBL: ChEMBL429844

Активність:  $K_i = 0.44$  nM



Ці зонди мають високий рівень активності ( $K_i$  в субнаномолярному діапазоні) та відповідають строгим вимогам щодо потенціалу використання як

молекулярних інструментів у біологічних дослідженнях. Після проходження відповідної експертної оцінки, зазначені сполуки можуть бути включені до бази даних Chemical Probe Portal, що підтвердить їхню цінність як наукових інструментів для подальших досліджень.

### 3.3.2. Порівняння сайтів зв'язування групи ChemSpace

Після детального аналізу сайтів зв'язування в АПФ2 і АПФ були виявлені специфічні структурні відмінності, які можуть бути використані для досягнення селективності і специфічності [40]. Обидва білки є цинковмісними ферментами, чутливими до аніонної активації [40,59], але, на відміну від АПФ, АПФ2 функціонує як карбоксипептидаза і не підлягає інгібуванню традиційними інгібіторами АПФ [59].

Порівняння сайтів зв'язування білків і лігандів може надати цінну інформацію про нетривіальні подібності між різними білками, які не завжди очевидні з точки зору первинної послідовності або загальної структури. Такі інсайти особливо корисні в контексті обчислювальної біології, структурної біоінформатики та на етапах оптимізації молекул у хімічній біології.

Наприклад, виявлення альтернативних або криптичних сайтів зв'язування, а також прогноз поліфармакологічних ефектів через структурні подібності сайтів, може мати важливе значення у проєктах, пов'язаних із таргетною селективністю та профілюванням побічних взаємодій. І хоча існують бази даних, що надають інструменти для такого аналізу [62], ручні або специфічні порівняння залишаються актуальними, особливо при роботі з новими або слабо охарактеризованими білками [30].

Під час поглибленого аналізу активних центрів АПФ2 та АПФ було виявлено дві критично важливі відмінності:

- TYR510 в АПФ2, який відповідає VAL518 в АПФ;
- та, найбільш суттєво, ARG273 у АПФ2 замість GLN281 у АПФ.

Разом з тим, тринадцять амінокислотних залишків зберігаються в обох ферментах, зокрема: GLU145/GLU162, CYS344/CYS352, HIS345/HIS353, CYS361/CYS370, ASP368/ASP377, HIS374/HIS383, GLU375/GLU384,

HIS378/HIS387, GLU402/GLU411, PHE504/PHE512, HIS505/HIS513, ARG514/ARG522, TYR515/TYR523.

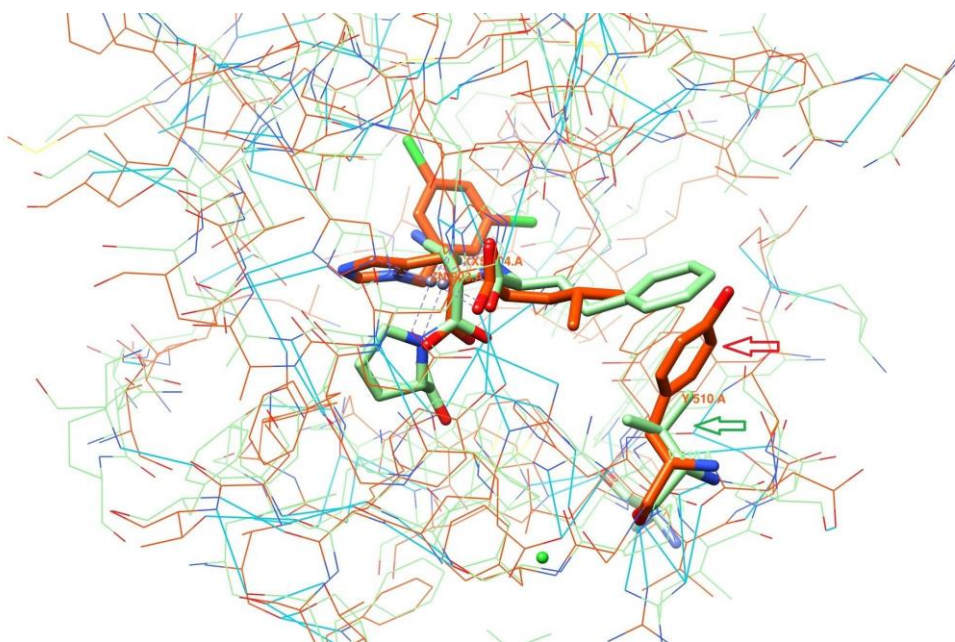
Ця висока консервативність активних центрів забезпечує практично ідентичне розташування зв'язувальних сайтів для лігандів XX5 (PDB код структури MLN-4760) та LPR (лізіноприл) при їх порівнянні.

У випадку АПФ2, залишки TYR510 і THR347 формують гідрофобний "вхід" у субсайт, який обмежує доступ великих бічних ланцюгів і створює перевагу для лігандів малого або середнього розміру — що добре узгоджується з експериментально підтвердженими перевагами субстратів.

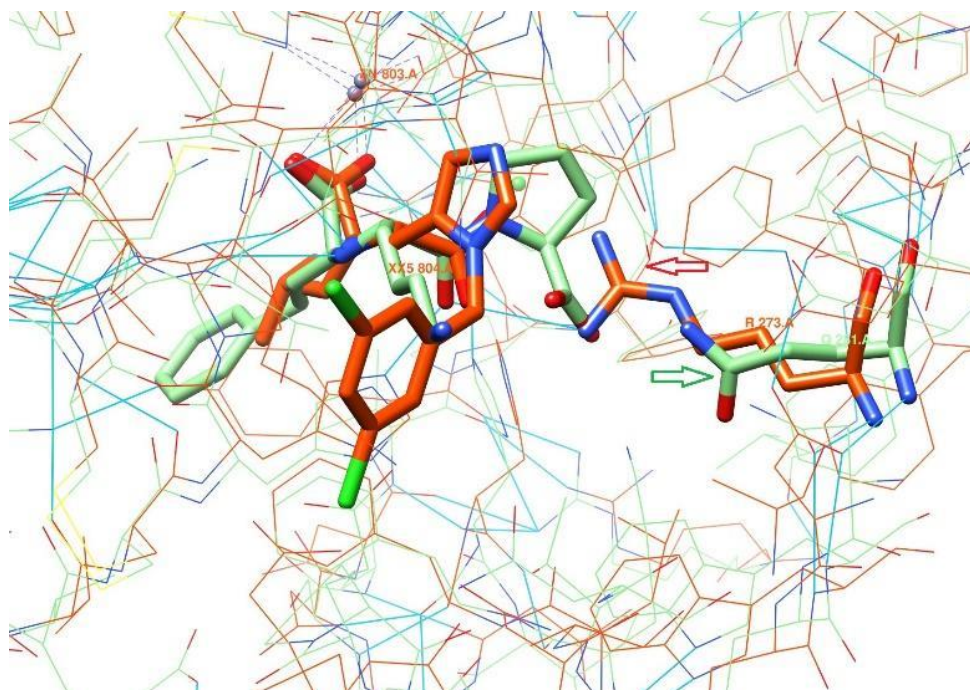
Під час накладення структур АПФ2, зв'язаного з ORE-1001, та АПФ, зв'язаного з лізіноприлом, було виявлено, що фенольне кільце TYR510 в АПФ2 займає простір, у якому в АПФ локалізується фенілпропілова частина лізіноприлу. Це свідчить про важливу роль TYR510 у формуванні селективного профілю зв'язування.

Ще однією істотною відмінністю між АПФ2 і АПФ є заміна ARG273 на GLN281 (Мал.18,19). Гуанідинієва група залишку ARG273 в АПФ2 утворює сольовий міст із термінальною карбоксильною групою інгібітора ORE-1001, що відіграє ключову роль у стабілізації комплексу.

Важливо, що гіпотеза щодо участі ARG273 у зв'язуванні лігандів була експериментально підтверджена. Проведення сайт-спрямованого мутагенезу, зокрема заміна аргініну на глютамін (мутація R273Q), призвело до значного зниження спорідненості відомих інгібіторів до рецептора. Це свідчить про критичне значення ARG273 у формуванні специфічної взаємодії між АПФ2 та його лігандами [40].



Мал.18. Порівняння сайтів зв'язування АПФ2 і АПФ. Зеленими позначені фрагменти АПФ з лігандом лізиноприлом (PDB: 1O86/LPR), оранжевими — АПФ2 з лігандом ORE-1001 (PDB: 1R4L/XX5). Синіми лініями показані водневі зв'язки. Зелені та оранжеві стрілки вказують на VAL518 в АПФ та TYR510 в АПФ2 відповідно.



Мал.19. Порівняння сайтів зв'язування АПФ2 і АПФ. Зеленими показані фрагменти АПФ з лігандом лізиноприлом (PDB: 1O86/LPR), оранжевими — АПФ2 з лігандом ORE-1001 (PDB: 1R4L/XX5). Синіми лініями позначені водневі зв'язки. Зелені та оранжеві стрілки вказують на GLN281 в АПФ та ARG273 в АПФ2 відповідно.

Ця істотна різниця між залишками в гомологічно подібних ферментах допомагає пояснити, чому потужні інгібітори АПФ, такі як лізиноприл,

еналаприлат і каптоприл, не є активними проти АПФ2 [103,106]. Виявлені взаємодії між ключовими амінокислотними залишками в активних сайтах АПФ2 та АПФ суттєво впливають на відмінності в субстратній специфічності та профілях зв'язування інгібіторів для цих ферментів.

Для АПФ2 було сформовано дві моделі з відповідними просторовими та хімічними обмеженнями:

#### **Модель 1:**

- ARG273: наявність обов'язкового водневого зв'язку;
- Область TYR510: позиційне (дистанційне) обмеження у вигляді сфери з радіусом 2 Å;
- Іонна взаємодія з  $Zn^{2+}$  (координація металу).

#### **Модель 2:**

- ARG273: два позиційно-дистанційні обмеження у вигляді сфер радіусом 1 Å для точного позиціонування обох атомів кисню карбоксильної групи ліганду;
- Область TYR510: таке саме сферичне обмеження (радіус 2 Å);
- Координація з іоном  $Zn^{2+}$ .

Для АПФ модель включала такі обмеження:

- Іонна взаємодія з  $Zn^{2+}$ ;
- LYS511: формування обов'язкового водневого зв'язку.

Візуалізацію обох докінгових моделей представлено на малюнках 20 та 21.

Для валідації побудованих моделей було використано набір лігандів із бази ChEMBL (<https://www.ebi.ac.uk/chembl/>, доступ від 21 липня 2021 року), для яких наявні експериментальні значення  $K_i$  або  $IC_{50}$  щодо відповідних мішеней. Загальна кількість таких лігандів склала 60 (Додаток Б).

Для калібрування докінгових моделей ми використали безкоштовний онлайн-ресурс DUD (Directory of Useful Decoys) [78], який містить каталоги структурно подібних, але неактивних сполук («приманок»), відібраних на основі фізико-хімічної схожості з відомими активними лігандами. Для кожного референтного інгібітора було підібрано по 100 відповідних "приманок", що дало

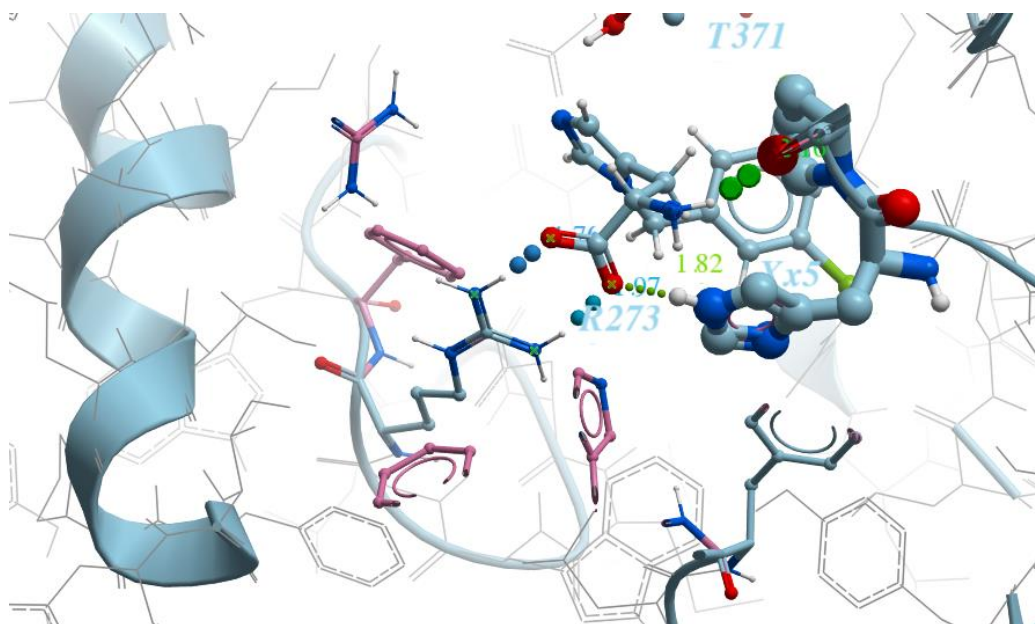
у сумі 6000 молекул. Усі побудовані моделі докінгу були протестовані в режимі *in silico*, шляхом скринінгу на наборах еталонних інгібіторів (активних сполук) та відповідних приманок. Це дозволило сформувавши різні варіанти моделей, які надалі оцінювалися з точки зору ефективності розмежування активних і неактивних молекул.

Для цього проводився аналіз кривих збагачення, зокрема: ROC-кривих (Receiver Operating Characteristic), logROC-кривих, LogAUC (логарифмічна площа під кривою).

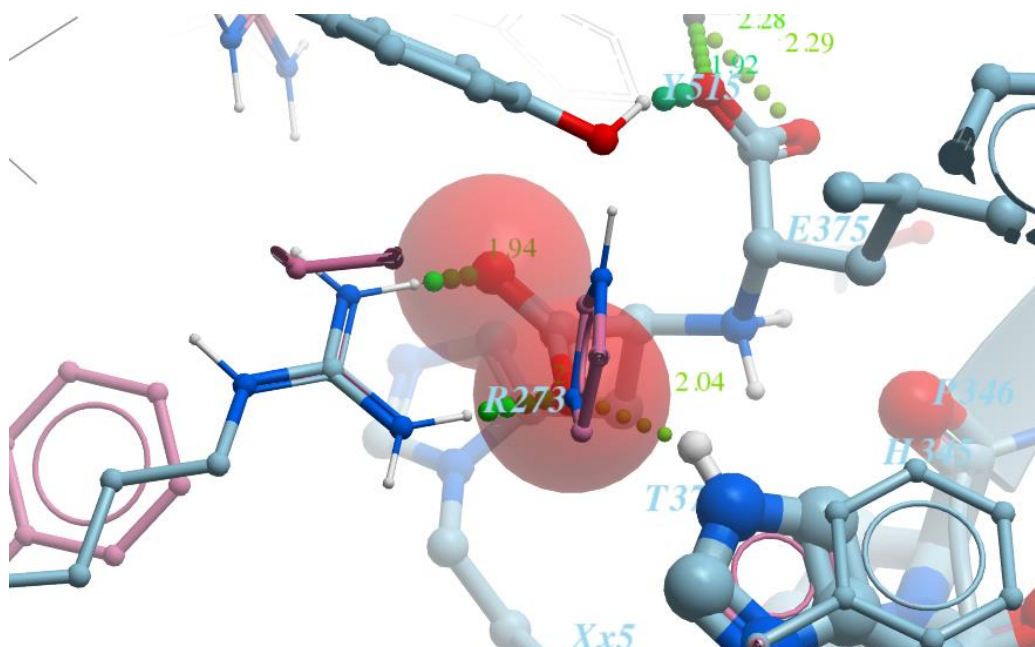
Ці метрики дозволяють оцінити здатність моделі пріоритезувати справді активні ліганди над неактивними аналогами. Найчастіше головний фокус зосереджується саме на активних молекулах, а не на приманках, що подібні за властивостями, але не зв'язуються з мішенню. Загальноприйнятим вважається підхід, згідно з яким, якщо модель демонструє високу здатність розділяти активні ліганди від приманок у тестовому наборі, то вона буде надійною і для подальшого віртуального скринінгу великих хімічних бібліотек [108,10].

Оцінка збагачення лігандів у порівнянні з приманками була використана як метод для кількісного визначення ефективності докінгової моделі щодо заданої мішені (або групи мішеней). Приманки — це молекули, які ймовірно не мають здатності зв'язуватися з ціллю, однак спеціально підібрані таким чином, щоб максимально збігатися з референтними активними лігандами за основними фізико-хімічними параметрами (молекулярна маса, LogP, кількість донорів/акцепторів водневих зв'язків тощо). Якщо під час докінгу референтні ліганди отримують суттєво кращі оцінки (scoring) порівняно з приманками, це вважається ознакою предиктивної здатності моделі, тобто її здатності ефективно відрізняти активні молекули від неактивних, навіть за умов високої фізико-хімічної подібності.

Таким чином, високий рівень збагачення слугує непрямим, але інформативним критерієм якості докінгового підходу.



Мал.20. Модель сольового зв'язування ARG273 і карбоксильною групою (перша модель АПФ2).



Мал.21. Обмеження відстані за допомогою двох сфер радіусом 1 Å, розміщених навпроти атомів азоту ARG273 (друга модель АПФ2).

### 3.3.3. Молекулярний докінг групи ChemSpace. Верифікація моделей

Для оцінки ефективності побудованих моделей (див. підрозділ 2.3) було виконано молекулярний докінг із використанням як низького, так і стандартного рівня зусиль. Моделювання проводилося з урахуванням трьох основних просторових/функціональних обмежень:

- Гідрофобна взаємодія поблизу TYR510 — зіткнення з будь-яким гідрофобним атомом (20 патернів);

- Іонна взаємодія з  $Zn^{2+}$  — контакт із будь-яким акцептором водневого зв'язку (30 патернів);
- ARG273:
  - Модель 1 (Grid-1) — взаємодія з будь-яким атомом-акцептором Н-зв'язку;
  - Модель 2 (Grid-2) — додаткові позиційні обмеження на обидва атоми кисню в карбоксильній групі.

Було протестовано чотири конфігурації параметрів докінгу:

Модель-1 і рівень зусилля докінгу 1;

Модель-2 і рівень зусилля докінгу 1;

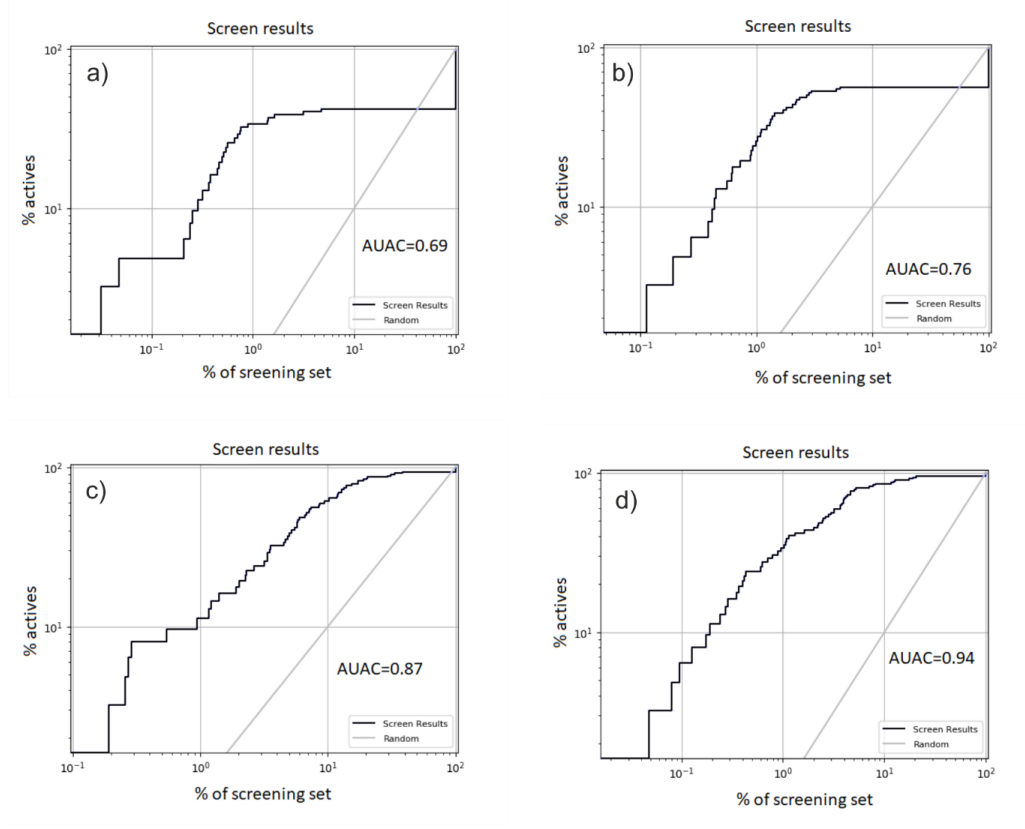
Модель-1 і рівень зусилля докінгу 5;

Модель-2 і рівень зусилля докінгу 5.

Для оцінки продуктивності докінгу було розраховано коефіцієнт збагачення, який дозволяє кількісно охарактеризувати здатність моделі відрізняти активні ліганди від неактивних (приманок). Серед використовуваних метрик — площа під ROC-кривою (AUC ROC) та площа під кривою накопичення (AUAC). Зокрема, AUAC інтерпретується як ймовірність того, що активна сполука, вибрана зі списку, ранжованого за оцінкою докінгу, опиниться вище за випадкову молекулу з рівномірного розподілу. Значення цього показника варіюється від 0 до 1, де 1 означає ідеальне збагачення [10].

У нашому дослідженні розраховане значення AUAC = 0.94, що свідчить про високу якість побудованих моделей. Серед чотирьох варіантів конфігурацій найкращі результати були отримані з Моделлю 2, яка використовує позиційно-дистанційні обмеження поблизу ARG273, а не лише Н-зв'язки. Це спостерігалось незалежно від рівня зусиль докінгу (1 або 5), що підтверджує вищу точність Моделі 2 у розмежуванні активних та неактивних сполук.

На основі цих результатів Модель 2 була обрана як основна для подальших масштабних віртуальних скринінгів (див. Мал.22).



Мал.22. Криві ROC для чотирьох результатів стикування. (а) Перший набір: Модель-1 і рівень зусилля докінгу 1. Площа під кривою накопичення: 0,69. (б) Другий набір: Модель-2 і рівень зусилля докінгу 1. Площа під кривою накопичення: 0,76. (с) Модель-1 і рівень зусилля докінгу 5. Площа під кривою накопичення: 0,87. (д) Модель-2 і рівень зусилля докінгу 5. Площа під кривою накопичення: 0,94.

### 3.3.4. Молекулярний докінг групи ChemSpace. Робочий процес скринінгу для Єнамін Ін-Сток колекції (3.2 мільйона сполук)

Інтеграція надвеликих хімічних бібліотек у віртуальний скринінг відкриває нові перспективи у пошуку лігандів, але водночас супроводжується суттєвими технічними викликами. Молекулярний докінг у такому масштабі потребує оцінки взаємодії кожної молекули з білковою мішенню, що включає генерацію та аналіз сотень тисяч або мільйонів конфігурацій для кожного з'єднання. Для забезпечення здійсненності скринінгу бібліотеки розміром понад мільярд молекул, як у випадку з Enamine REAL Space (1,4 млрд сполук), час докінгу має бути обмежений до не більше ніж 1 секунда на молекулу на одне обчислювальне ядро, особливо при використанні кластерів середнього масштабу (500–1000 ядер). Це, в свою чергу, виключає можливість застосування детальних

енергетичних функцій і складних моделей взаємодій, необхідних для досягнення "хімічної точності" [52]. З огляду на ці обмеження, дана частина дослідження була присвячена розробці та апробації ефективного обчислювального конвеєра, здатного обробляти бібліотеки середнього розміру з урахуванням ресурсоспоживання, швидкодії та збереження адекватної точності. Метою є формування основи для масштабування докінгових протоколів до бібліотек надвеликих розмірів.

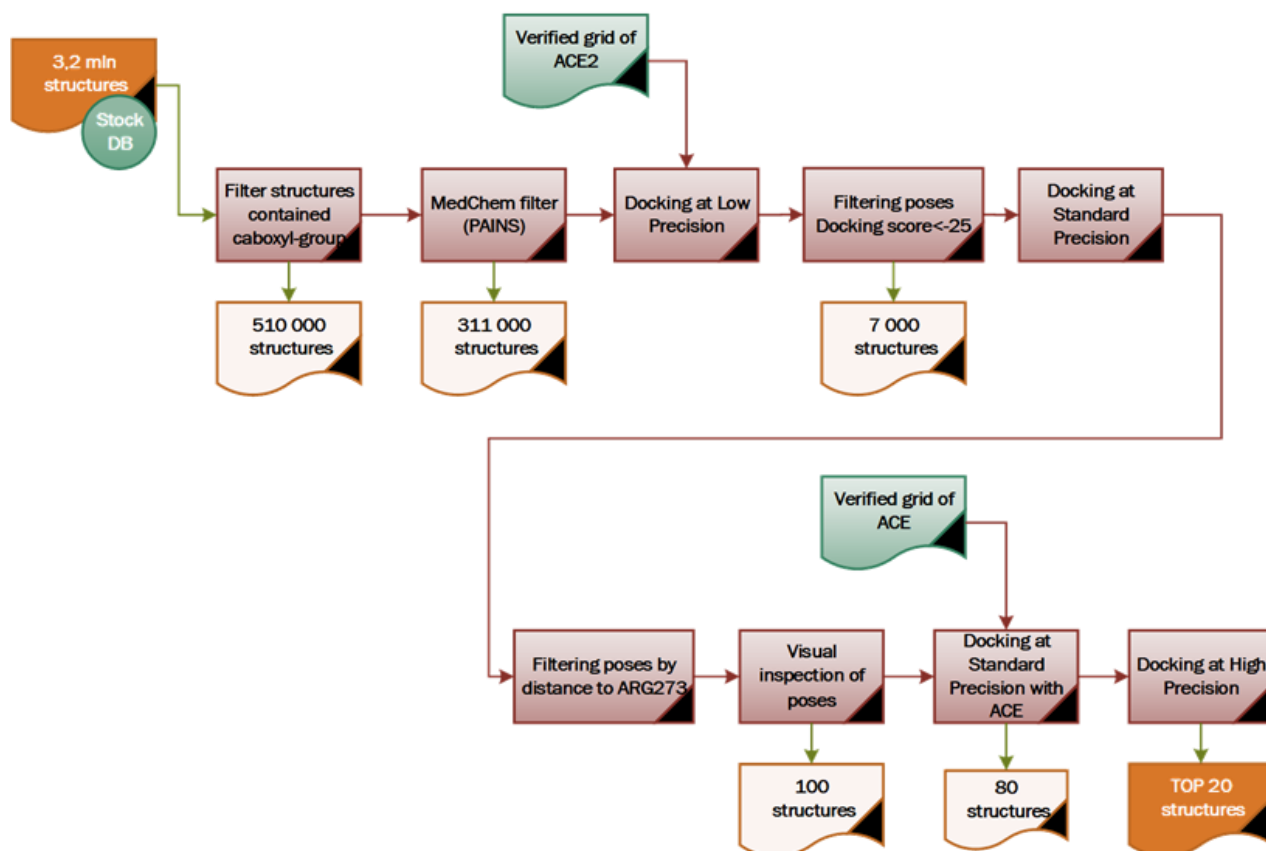
Для цього ми адаптували та модифікували підхід, нещодавно запропонований Halgren та співавторами, який було створено спеціально для докінгу з великими обсягами даних. Наша модифікація орієнтована на практичне застосування у високопродуктивних середовищах, а також на валідацію ефективності при скринінгу Enamine REAL бази [42]. Тому для скринінгу Enamine SCC ми використали наступний покроковий алгоритм (Мал.23.) . Процес фільтрації та відбору кандидатів для докінгу складався з кількох послідовних етапів:

- Початкова SMARTS-фільтрація бази даних із метою виділення молекул, що містять принаймні одну карбоксильну групу;
- Застосування стандартного набору фільтрів медичної хімії для усунення небажаних сполук:
  - PAINS — виключення фрагментів, що мають неспецифічну активність;
  - BRENK — відбір сполук без потенційно токсичних субструктур;
  - NIH, ZINC — виключення структур із небажаними функціональними групами;
  - LILLY — фільтрація потенційно реактивних або непридатних для подальшої розробки молекул [8,31];
- Проведення початкового докінгу з мінімальним рівнем точності з фокусом на оцінку взаємодії з ключовими залишками мішені;

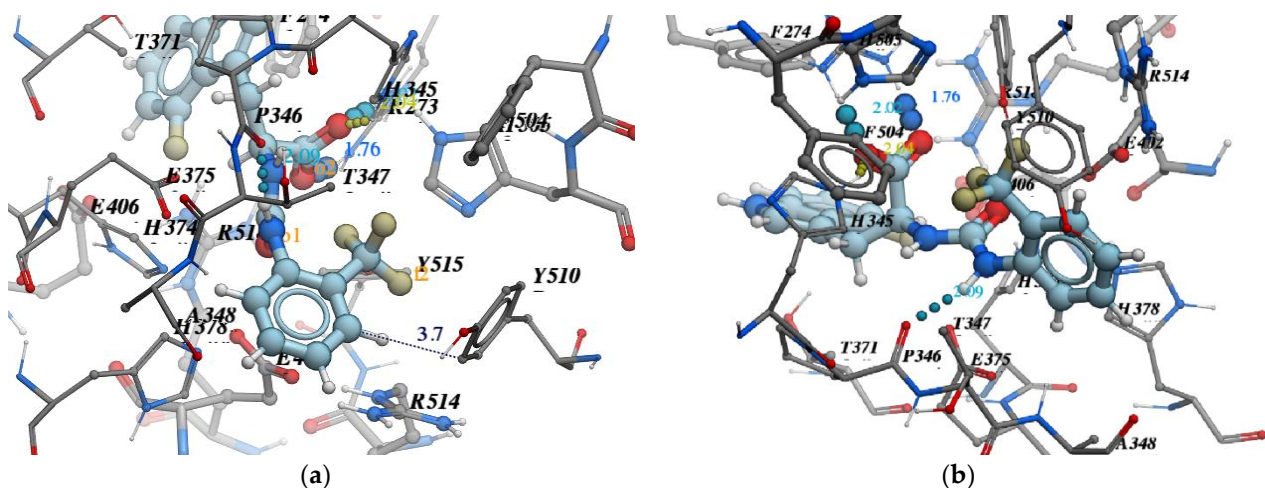
- Подальший докінг із середнім рівнем точності, після чого результати були ранжовані за енергією взаємодії з критично важливими амінокислотами активного сайту;
- Візуальний аналіз положень лігандів у кишені зв'язування. На основі цього було відібрано 100 молекул з найбільш релевантними позами.

Під час остаточного відбору найперспективніших кандидатів особливу увагу приділяли таким критеріям:

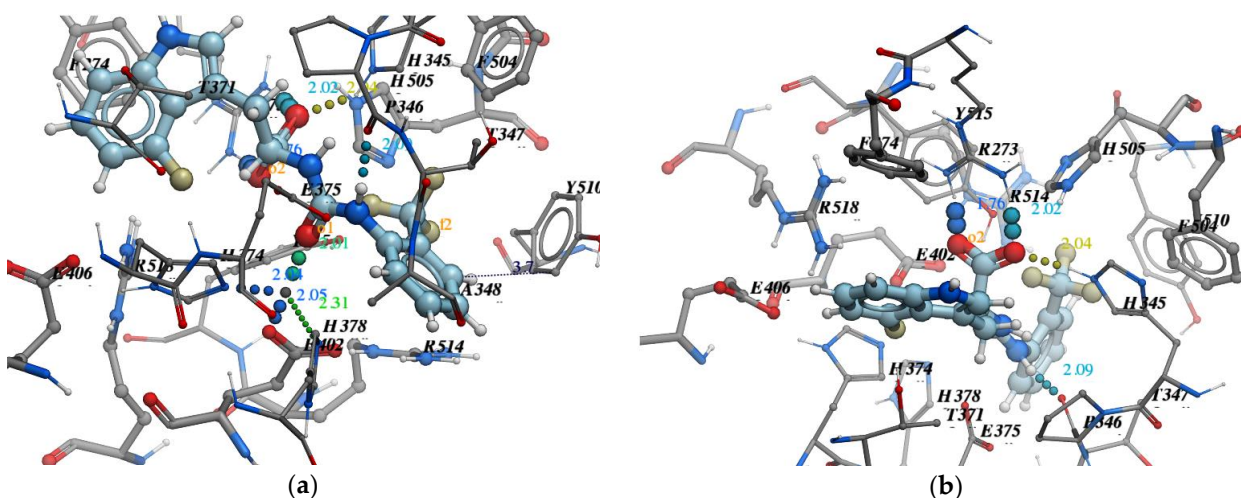
- Точність позиціонування та відстань до залишку ARG273;
- Розташування відносно іону цинку ( $Zn^{2+}$ ) — як показник потенційної координаційної взаємодії;
- Локалізація гідрофобної частини ліганду поблизу TYR510, що відповідає за формування обмеженого субсайту (див. Мал.24 і 25);
- Наявність стабільного водневого зв'язку з PRO346 — як додатковий фактор стабілізації комплексу.



Мал.23. Робочий процес віртуального скринінгу для Enamine In-Stock Collection (3,2 мільйона сполук).



Мал.24. Важливі критерії відбору найкращих кандидатів: (а) положення гідрофобної частини ліганду поблизу TYR510; (b) наявність водневого зв'язку з PRO346.



Мал.25. Важливі критерії відбору найкращих кандидатів: (а) точність відносно ARG273; (b) точність відносно  $Zn^{2+}$ .

У межах робочого процесу віртуального скринінгу, проведеного з використанням бібліотеки Enamine In-Stock Collection (3,2 мільйона сполук), було відібрано 100 найкращих молекул, які демонстрували перспективну взаємодію з мішенню АПФ2. Наступний етап передбачав оцінку селективності шляхом додаткового циклу молекулярного докінгу на високому рівні точності. Отримані комплекси було піддано візуальному аналізу, що дозволило критично оцінити просторову орієнтацію лігандів та їхню взаємодію з ключовими амінокислотними залишками. У результаті цього багатоетапного фільтраційного процесу було сформовано остаточний список із 20 сполук-кандидатів, які рекомендовані для подальшої експериментальної перевірки (Додаток В).

### 3.3.5. Молекулярний докінг групи ChemSpace. Робочий процес скринінгу для Enamine REAL Database (Enamine RDB, 1.4 мільярдів сполук)

Методи віртуального скринінгу (VS) сьогодні є ключовим етапом у сучасному процесі відкриття лікарських засобів [39]. Застосування широкомасштабного VS стало стандартною практикою перед проведенням лабораторних експериментів для первинного відбору потенційних кандидатів [128,97,104]. Завдяки стрімкому розвитку обчислювальних технологій з'явилась можливість здійснювати високопродуктивний скринінг за участю від тисяч до мільярдів молекул. Однак для дослідників, які не спеціалізуються на обчислювальних методах, оцінка ефективності різних технічних підходів та їх впливу на результати може становити суттєву складність [128]. Важливо оптимально використовувати час обчислювальних команд і наявні обчислювальні ресурси. У зв'язку з цим виникає потреба в раціональному управлінні ресурсами, зокрема ефективному плануванні часу роботи обчислювальних вузлів і розподілу завдань між методами різної складності.

Організація VS-процесу включає адміністрування великих масивів даних, як структурованих (наприклад, бібліотеки молекул), так і неструктурованих (звіти, параметри моделювання тощо), і має на меті досягнення високої результативності при помірних обчислювальних витратах.

У нашому підході ми зосередились на оптимізації VS-робочого процесу, орієнтуючись на об'єднання обчислювально ефективних (але менш точних) і більш ресурсоемних (але точніших) методів. Основна мета — знайти оптимальний компроміс між продуктивністю та точністю.

З метою попередньої оцінки можливостей масштабування нашої докінгової платформи, ми провели тестування на контрольному наборі з 250 молекул, підготовлених у форматі 3D-SDF. Для кожної молекули було виконано докінг із використанням одного ядра ЦП, а також трьох рівнів зусиль:

- низький (рівень 1),
- стандартний (рівень 5),
- високий (рівень 10).

Результати продуктивності для кожного режиму докінгу представлено в Табл.2.

Табл.2. Результати вимірювань обчислювальних можливостей<sup>1</sup>.

Level of Effort	CPU	Number of Ligands	Time, min	Min Per Ligand
1	1	250	31	0.124
5	1	250	110	0.44
10	1	250	220	0.88

<sup>1</sup> Технічна інформація робочої станції для тестування, на якій проводилися вимірювання: 1. ЦП: Intel (R) Core (TM) i5-9600K CPU @ 3,70 ГГц; 2. 16 ГБ оперативної пам'яті; 3. Операційна система: Ubuntu 20.04.3 LTS.

Крім того, ми спланували та оцінили час процесора, необхідний для реалізації кожного з трьох планів скринінгу. Детальний опис дорожніх карт RM1, RM5 та RM10 наведено в таблицях 3-5.

У таблиці 6 порівнюються переваги правильного планування скринінгової програми.

Як показали розрахунки, правильне проєктування робочого процесу скринінгу та раціональне управління ресурсами дозволяють досягти значної економії процесорного часу — до 2000-кратного скорочення обчислювальних витрат. На основі аналізу ми обрали протокол RM3 як найбільш ефективний за співвідношенням обчислювальні витрати / результативність. Використовуючи RM3, стало можливим масштабувати скринінг до бібліотек надвеликих розмірів — до 40 мільярдів молекул. Для порівняння: Enamine REAL Database (1,4 млрд сполук) була повністю оброблена за 17 годин (тільки розрахунковий час) на стандартному обчислювальному вузлі з 60 ядрами. Таким чином, скринінг 40 млрд сполук за цим самим протоколом оцінюється в 486 години (приблизно 20 днів), що є цілком прийнятним часом для високопродуктивних кампаній віртуального скринінгу.

Для контрасту, застосування інших протоколів на такому ж обладнанні дало б суттєво вищі часові витрати: RM1 — орієнтовно 41 349 години ( $\approx$  57 місяців), RM2 — близько 4 141 годин ( $\approx$  5,75 місяців).

Варто зазначити, що ці оцінки є теоретичними й не враховують людський фактор, паралельні завдання, непередбачені збої системи чи інші технічні ризики, які можуть вплинути на загальний час виконання.

Табл.3. Базова дорожня карта RM1.

Description of the Step	Number of Structures Left	CPU time (min)
INPUT	1.41 bln	-
MedChem filter (PAINS, etc.)	1.2 bln	1800
Ligand docking at low effort (1.)	100000	148,800,000
Ligand docking at standard effort (5.)	10000	44,000
Ligand docking at standard effort (5.) with ACE	9000	4400
Ligand docking at high effort(10.)	200	7920
Sort pose by docking score and visual inspection	20	-
Total CPU time (min)	-	148,858,120

Табл.4. Поглиблена дорожня карта RM2.

Description of the Step	Number of Structures Left	CPU time (min)
INPUT	1.41 bln	-
Diversification 1/10	141 mln	1800
MedChem filter (PAINS, etc.)	120 mln	180
Ligand docking at low effort (1.)	-	14,880,000
Sort poses by distance to ARG273, TYR510, PRO346	50,000	-
Ligand docking at standard effort (5.)	-	22,000
Sort poses by distance to ARG273, TYR510, PRO346	5000	-
Ligand docking at standard effort (5.) with ACE	4500	2200
Ligand docking at high effort(10.)	100	3960
Sort poses by docking score and visual inspection	20	-
Total CPU time (min)	-	14,910,140

Табл.5. Оптимальна дорожня карта RM3.

Description of the Step	Number of Structures Left	CPU time (min)
INPUT	1.41 bln	-
Filter structures contained carboxyl-group	29 mln	600
Diversification	30,000	600

Ligand docking at low effort (1.)	-	3720
Sort poses by distance to ARG273, TYR510, PRO346	5,000	-
Visual inspection of poses	50	-
Similarity searching by Morgan FP	121,000	2250
Ligand docking at standard effort (5.)	28,000	53,240
Sort poses by distance to ARG273, TYR510, PRO346	-	-
Sort poses by docking score and visual inspection	175	-
MedChem filter (PAINS, etc.)	105	-
Ligand docking at standard effort (5.) with ACE		46.2
Ligand docking at the level of effort (10.)	20	92.4
Total CPU time (min)	-	60,548

Табл.6. Переваги правильного планування скринінгової програми.

Comparison of the Estimations of CPU Time Required for Roadmaps	Difference (Number of Times)
RM1 vs. RM2	10
RM1 vs. RM3	2458
RM2 to RM3	246

Скринінг бази даних Enamine RDB було проведено за алгоритмом, що мав найбільшу оцінку ефективності та включав кілька послідовних етапів (Мал.26). Спочатку застосували SMARTS-фільтрацію для виявлення сполук з карбоксильною групою, отримавши в результаті близько 29 мільйонів молекул. Через значний обсяг цього набору додатково застосували метод хімічного різноманіття, заснований на розрахунку коефіцієнта подібності Танімото з використанням молекулярних відбитків Моргана [91], скоротивши кількість сполук до 30 тисяч. Для цих 30 тисяч сполук було проведено молекулярний докінг із подальшим візуальним аналізом найкращих поз. У результаті цього кроку відібрали 50 молекул із найбільш перспективними значеннями докінг-оцінки та чітко вираженими взаємодіями з цільовими залишками.

На наступному етапі провели пошук молекул із подібністю форми, використовуючи метод USRCAT (Ultrafast Shape Recognition with Credo Atom

Types) [96]. Враховуючи сполуки з коефіцієнтом подібності Tanimoto не нижче 0,4, сформували додатковий набір молекул для подальшого скринінгу, що збільшило їх кількість до приблизно 121 тисячі.

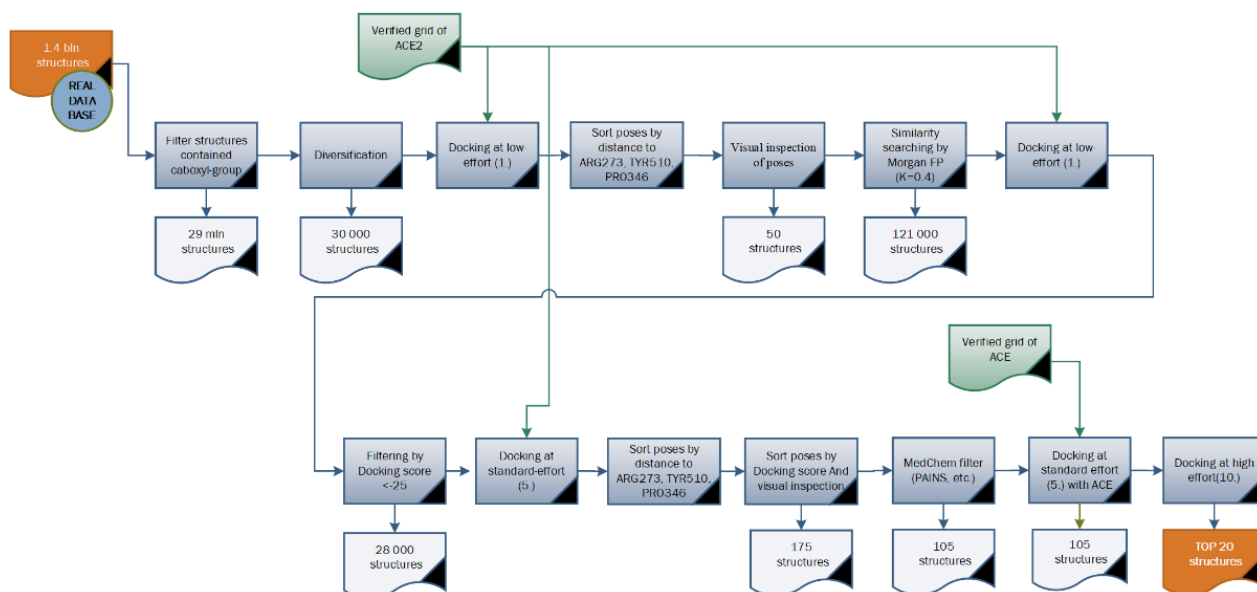
На наступному етапі дослідження відібраний набір сполук було піддано повторному молекулярному докінгу з використанням моделі з низьким рівнем зусиль (рівень 1). Критерієм відбору слугувало порогове значення докінг-оцінки  $< -25$ , що дозволило звузити вибір до 28 000 кандидатів.

Далі ці молекули були протестовані в умовах стандартного рівня зусиль (рівень 5). Для підвищення продуктивності скринінгу додатково було застосовано фільтр просторової наближеності до ключових амінокислотних залишків мішені: ARG273, TYR510 та PRO346. Після цього результати було відсортовано за значенням енергії зв'язування, використовуючи ті ж порогові значення, що й для докінгу стандартної бази.

Отриманий набір був додатково відфільтрований за допомогою ChemFilters, включаючи правила PAINS та LILLY, що дозволило виключити потенційно непридатні для подальшої розробки молекули. Після фільтрації залишилось 105 кандидатів.

Для перевірки селективності цих лігандів щодо АПФ2 було проведено контрольний докінг до ферменту АПФ. Лише одна сполука мала оцінку понад 25, але її зв'язування в активному центрі виявилось нефізіологічним за просторовою орієнтацією, тому жодна молекула не була виключена на цьому етапі.

На завершальному етапі було виконано докінг із високим рівнем зусиль (рівень 10) для всіх 105 сполук із використанням перевіреної моделі. Особлива увага приділялась експертній оцінці зв'язування та специфічності взаємодій з ключовими залишками білка. У підсумку було обрано 20 перспективних молекул-кандидатів, які демонстрували стабільні й специфічні взаємодії у кишені зв'язування.



Мал.26. Робочий процес віртуального скринінгу для Enamine RDB (1,4 мільярда сполук).

### 3.3.6. Дослідження молекулярною динамікою результатів групи ChemSpace

На завершальному етапі дослідження для трьох найбільш перспективних лігандів було проведено молекулярно-динамічне (MD) моделювання комплексів із білковою мішенню протягом 10 нс. Отримані результати засвідчили високу стабільність комплексів "господар-гість", що дозволяє зробити припущення щодо потенційної *in vitro* активності зазначених сполук.

Параметри MD-симуляцій, зокрема кількість частинок, часовий крок та тривалість моделювання, були адаптовані під наявні обчислювальні ресурси, забезпечуючи оптимальний баланс між точністю відображення природних процесів та ефективністю розрахунків. Загальна тривалість симуляцій відповідала часовим межах, необхідним для статистично обґрунтованого аналізу кінетики взаємодій.

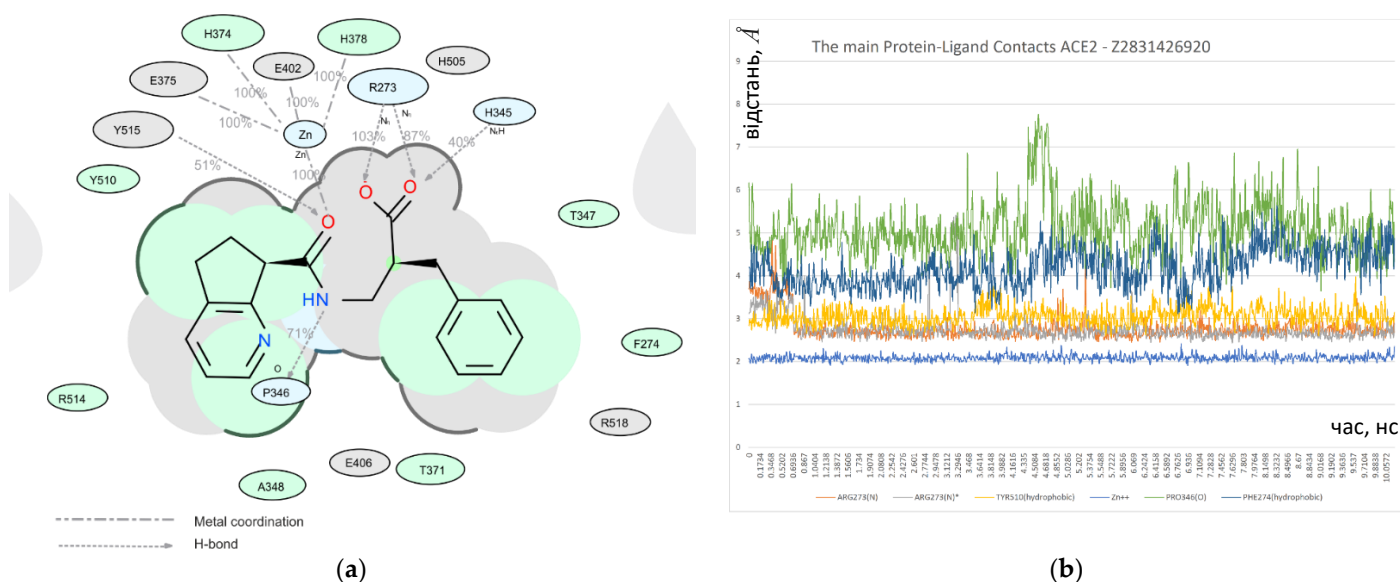
Аналіз MD-траєкторій був зосереджений на таких ключових параметрах:

- Середньоквадратичне відхилення (RMSD) положення ліганду протягом симуляції;
- Кількість ключових контактів ліганд-білок у кожен момент часу;
- Еволюція специфічних взаємодій між лігандом і білком протягом усього MD-інтервалу.

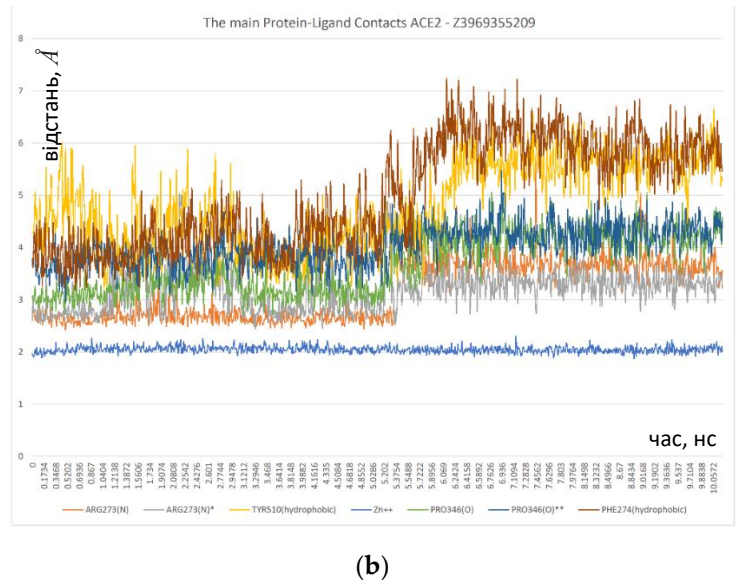
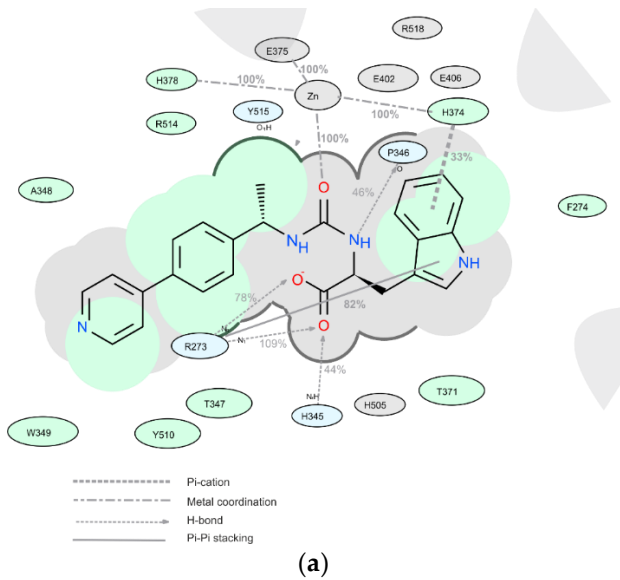
Окрема увага приділялася чотирьом типам взаємодій: водневим зв'язкам, координації з іонами металів,  $\pi$ -катіонним контактам,  $\pi$ - $\pi$  стекингу.

Для кількісного представлення результатів були побудовані нормалізовані стовпчасті діаграми, де, наприклад, значення 0,7 означає, що певна взаємодія зберігалася протягом 70% часу моделювання. Значення вище 1,0 можуть вказувати на множинні контакти одного типу, сформовані тим самим амінокислотним залишком.

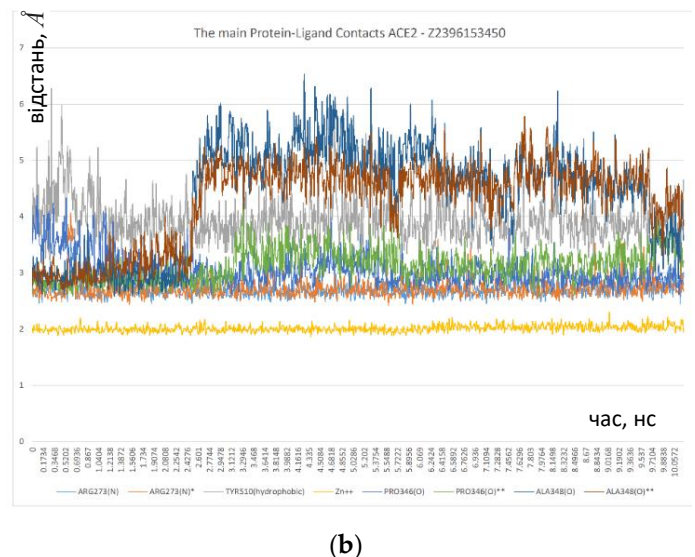
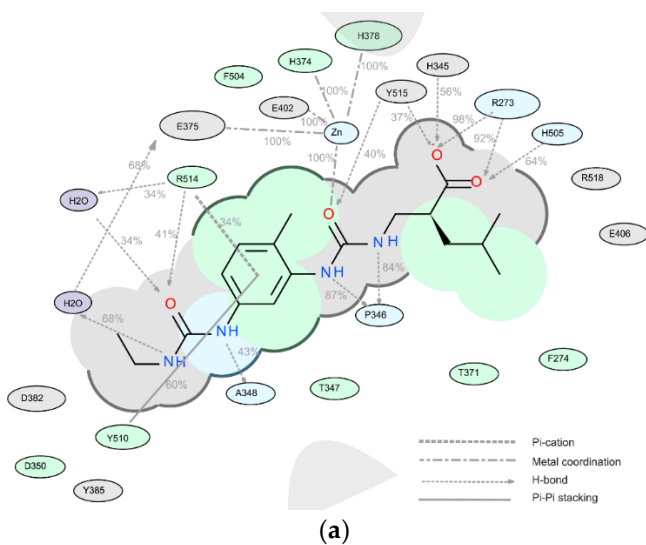
Деталізовані схеми атомних взаємодій між лігандами та ключовими залишками білка наведено на рисунках 27–29. Загалом, у вибраних траєкторіях було зафіксовано, що стійкі ліганд-білкові контакти підтримувалися щонайменше 30% часу симуляції (0–10 нс) [72].



Мал.27. Детальна схема взаємодії атомів ліганду (Z2831426920) із білковими залишками: (а) узагальнене графічне зображення взаємодій і контактів (Н-зв'язки, гідрофобні, іонні, водневих містки); (б) схема, що показує еволюцію відстані (в Å) між лігандом і залишками.

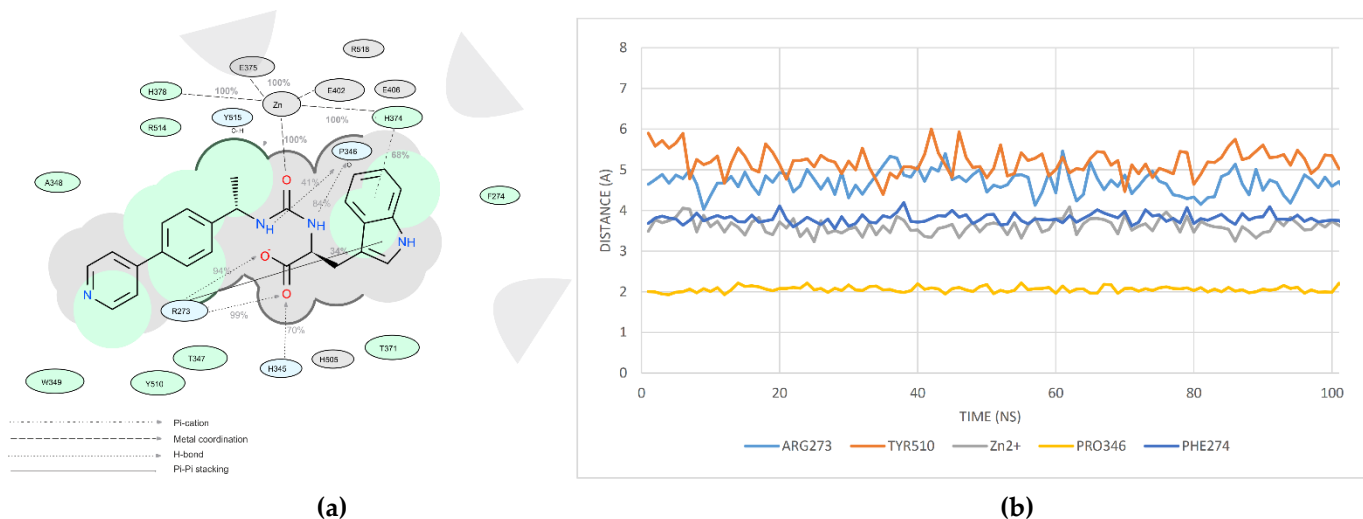


Мал.28. Детальна схема взаємодії атомів ліганду (Z3969355209) із білковими залишками: (a) узагальнене графічне зображення взаємодій та контактів (H-зв'язки, гідрофобні, іонні, водні містки); (b) схема, що показує еволюцію відстані (в Å) між лігандом і залишками.



Мал.29. Детальна схема взаємодії атомів ліганду (Z2396153450) із білковими залишками: (a) узагальнене графічне зображення взаємодій і контактів (H-зв'язки, гідрофобні, іонні, водні містки); (b) схема, що показує еволюцію відстані (в Å) між лігандом і залишками.

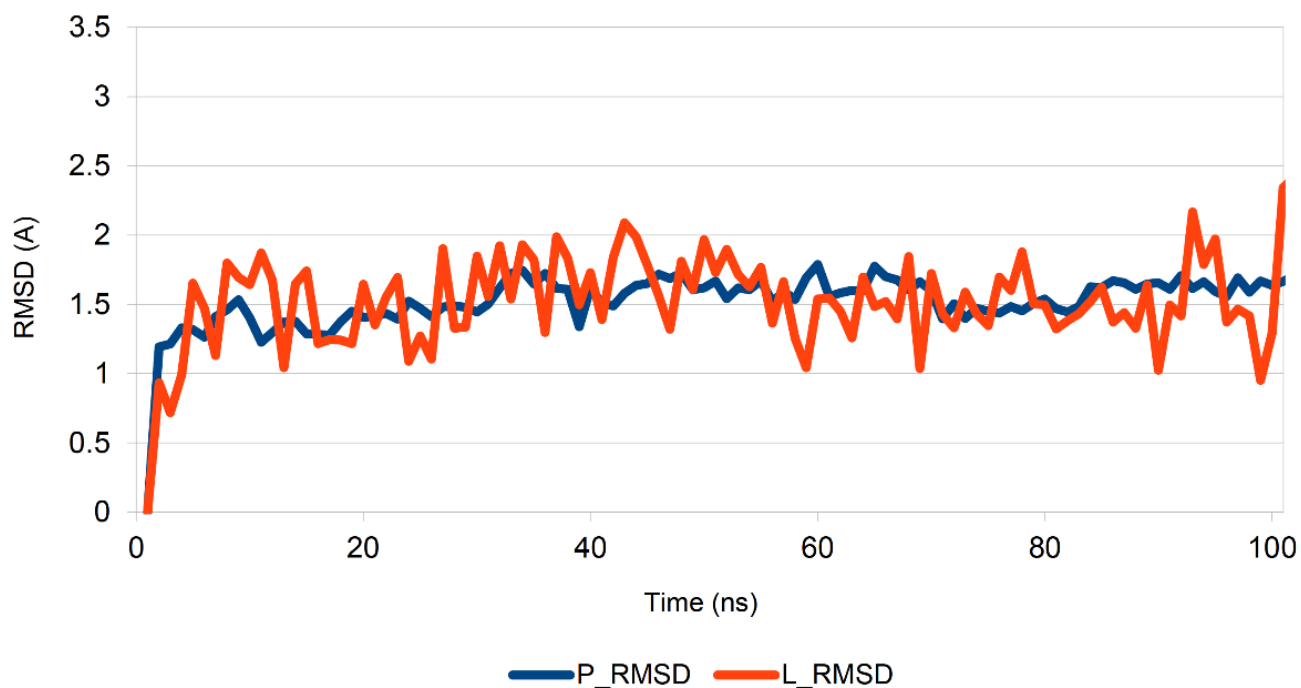
Ми вирішили підтвердити обґрунтованість нашого вибору для тривалості моделювання молекулярної динаміки обраних лігандів. Для цього ми провели аналогічні дослідження тривалістю 100 нс з використанням одного з топ-лігандів (Z3969355209) (Мал.30).



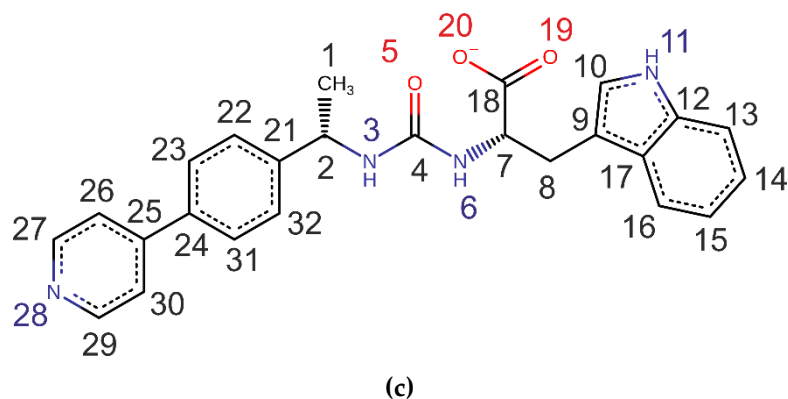
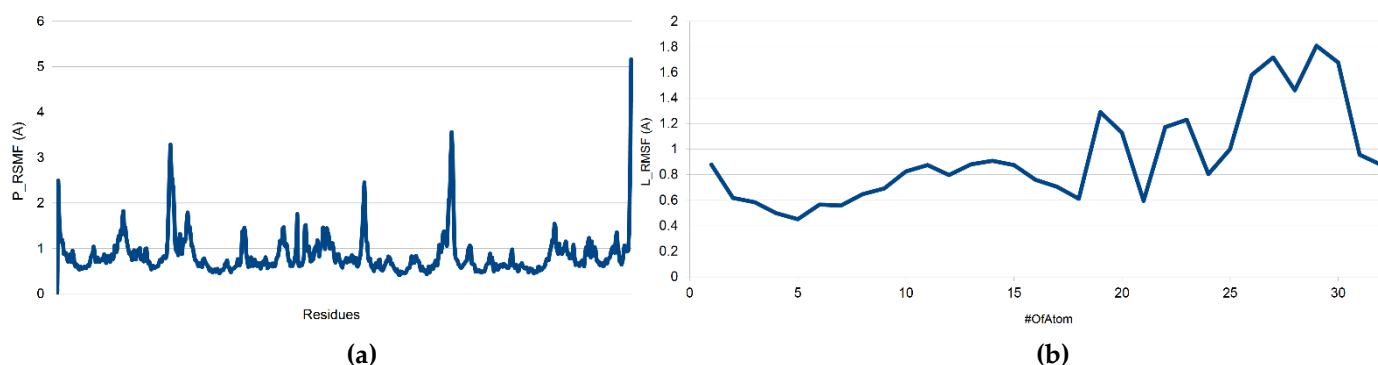
Мал.30. Детальна схема взаємодії атомів ліганду (Z3969355209) із білковими залишками (100 нс): (а) узагальнене графічне зображення взаємодій та контактів (Н-зв'язки, гідрофобні, іонні, водневі містки); (б) схема, що показує еволюцію відстані (в Å) між лігандом і залишками.

Крім того, середньоквадратичне відхилення (RMSD) певних атомів у молекулі відносно еталонної структури було розраховано за допомогою підгонки структури до еталонної структури за методом найменших квадратів. Синій графік (P\_RMSD) показує еволюцію RMSD білків. Усі білкові каркаси в першу чергу були вирівняні з каркасом, а потім було обчислено RMSD. Зміни порядку 1-3 Å цілком прийнятні для малих глобулярних білків. Ліганд RMSD (L\_RMSD) вказує, наскільки стабільний ліганд щодо білка та його зв'язуючої кишені. Якщо спостережувані значення значно перевищують RMSD білка, то, ймовірно, ліганд дифундував зі свого вихідного сайту зв'язування. У нашому випадку значення RMSD білка та ліганду не перевищували 2,5-3 Å від початкової рамки, що свідчить про стабільність комплексу (Мал.31) [72].

Крім того, ми розрахували середньоквадратичну флуктуацію (RMSF) як корисний індикатор локальних змін у положенні атомів ліганду та вздовж білкового ланцюга. Було побудовано стандартне відхилення (RMSF). Піки на цій діаграмі вказують на ділянки білка, які найбільше коливались під час моделювання [88]. Ліганд RMSF пояснює, як фрагменти ліганду взаємодіють з білком та їхню ентропійну роль у події зв'язування. На малюнку 32 показано ці зміни в білках (а) і лігандах (б) відповідно до атомної нумерації в ліганді.



Мал.31. Ділянки білка та ліганду RMSD комплексу (АПФ2-Z3969355209).



Мал.32. Графіки RSMF для (а) білка, (б) ліганду Z3969355209 та (в) відповідного перерахування атомів у ліганді.

### 3.3.7. Фармакофорний скринінг UNISTRA на основі структури

З огляду на обмеження точності при моделюванні зв'язування металів, було вирішено імітувати координацію  $Zn^{2+}$ -ліганд шляхом включення до фармакофорної моделі двох направляючих функцій донорів водневого зв'язку,

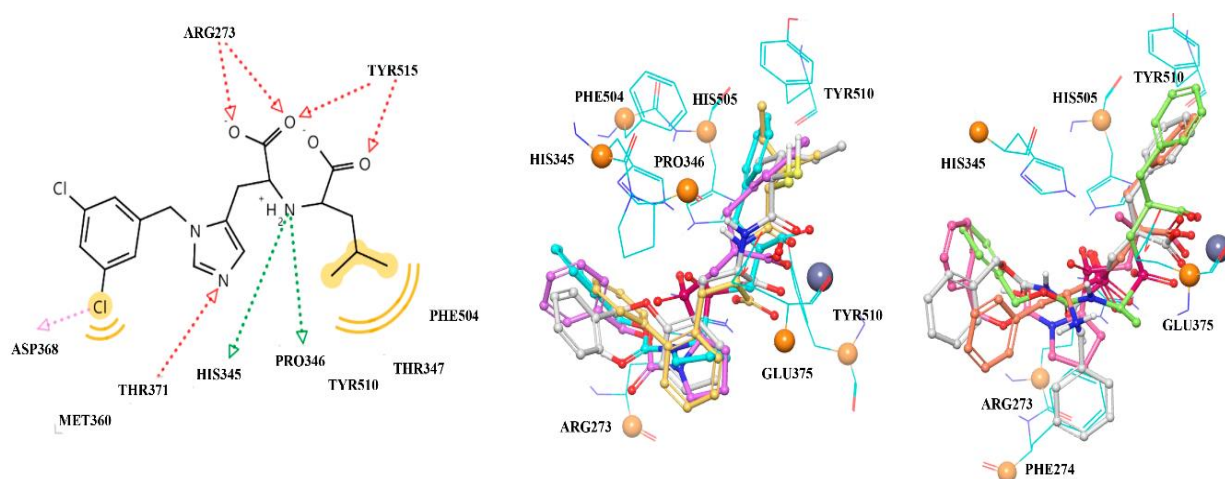
розташованих на відстані 4,1 Å, що відповідає геометрії координаційного зв'язування. Такий підхід дозволяє наблизити модель до фізіологічних умов, компенсуючи недосконалість енергетичних функцій у традиційному докінгу.

Крім того, для збільшення ймовірності виявлення лігандів з потенційною проникністю через гематоенцефалічний бар'єр (ГЕБ), із моделей виключено заряджені фармакофорні функції, а натомість збережено ключові гідрофобні елементи, які можуть сприяти пасивній дифузії. Це було враховано як при побудові фармакофора на основі структури (structure-based, SB), так і у лігандно-орієнтованій моделі (ligand-based, LB).

Фармакофор SB був сформований на основі рентгенівської структури 1R4L, у комплексі з співкристалізованим лігандом ChEMBL429844. Отримана модель включала 10 фармакофорних ознак, а саме: 2 донори водневого зв'язку (H), 5 акцепторів (HA), 2 гідрофобні області (HA), 1 негативний донор (XD).

З метою розширення охоплення та зниження ризику пропуску потенційно активних структур, фармакофорна модель була гнучкою: під час віртуального скринінгу допускалося виконання 9 із 10 ознак, що дозволяло включити до пошуку сполуки, які лише частково відповідають референсному профілю.

Цей підхід був застосований для скринінгу бібліотеки Enamine, що забезпечило високу чутливість пошуку без суттєвого компромісу з вибірковістю.



Мал.33. Зліва проворуч: 1. Взаємодія між MLN-4760 і АПФ2 (за винятком взаємодії  $Zn^{2+}$ -ліганду). Найбільш активні сполуки з контрольного набору добре підходять до фармакофорної моделі (2), тоді як менш активні не повністю відповідали фірмакофорній моделі (3).

Малюнок 33 ілюструє ключові типи взаємодій між інгібітором MLN-4760 та білком АПФ2, акцентуючи увагу на незаряджених ділянках та взаємодіях, що не включають координацію з іоном цинку.

Частина А демонструє деталізоване зв'язування ліганду в активному центрі, де чітко видно, як MLN-4760 імітує природний пептидний субстрат, що сприяє його специфічному впізнаванню ферментом.

Частина В показує відповідність найактивніших сполук з референтного набору до побудованої фармакофорної моделі, підтверджуючи, що саме ці взаємодії є критичними для селективного розпізнавання АПФ2.

Частина С ілюструє приклади малоефективних сполук, що мають специфічні просторові конфігурації або неklasичні патерни зв'язування, які можуть знижувати їхню афінність або селективність, але водночас надають цінну інформацію для подальшої оптимізації.

З метою розширення охоплення потенційних хітів, було введено гнучке правило відповідності: під час скринінгу одна з десяти фармакофорних ознак вважалася необов'язковою. Таким чином, до вибірки включалися усі сполуки, що відповідали щонайменше дев'яти з десяти характеристик. Застосування цього критерію до бібліотеки Enamine дозволило ідентифікувати 72 потенційні хіти, що демонструють високу ступінь відповідності ключовим ознакам зв'язування в моделі.

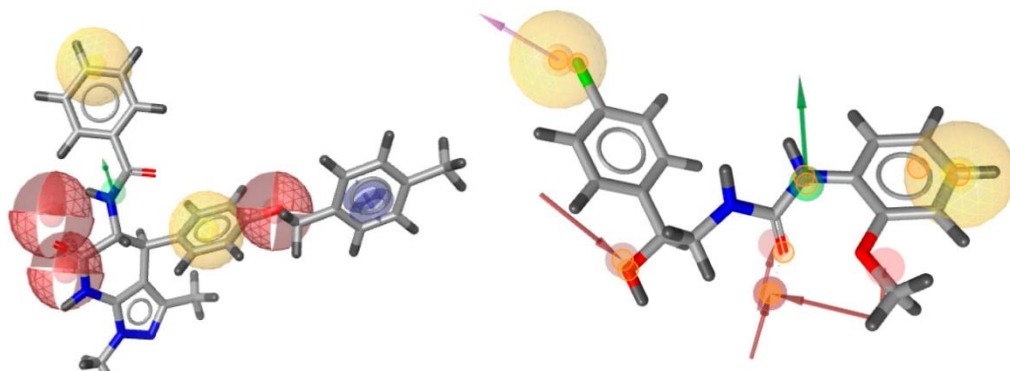
### **3.3.8. Фармакофорний скринінг UNISTRA на основі лігандів**

Для побудови фармакофору, орієнтованого на ліганди (ligand-based, LB), було використано набір із 51 сполуки, активної щодо АПФ2, з  $pK_i > 7$ , серед яких 35 молекул мали  $pK_i > 8$ , що свідчить про їх високу афінність. За допомогою програмного забезпечення LigandScout ці сполуки були кластеризовані на основі спільних фармакофорних характеристик.

В результаті було отримано сім кластерів, з яких три включали більше трьох молекул — саме ці кластери представляли основні хемотипи, притаманні відомим інгібіторам АПФ2. Особливу увагу було приділено одному кластеру з семи сполук, які, окрім високої активності, демонстрували селективність до

АПФ2 порівняно з АПФ. На основі цього кластера було побудовано селективний фармакофор, що включав сім консервативних ознак, спільних для всіх його представників. Таким чином, вдалося виокремити фармакофорну модель, яка відображає специфічні вимоги до селективного зв'язування з АПФ2.

Віртуальний скринінг бібліотеки Enamine із застосуванням цієї моделі дозволив відібрати 91 хіт. Аналіз фізико-хімічних властивостей відібраних сполук показав, що виключення заряджених функціональних груп із моделі та фокус на гідрофобні взаємодії призвели до зростання значень  $\log P$  у відібраних хітах. Якщо для більшості референсних лігандів  $\log P$  мав негативні або близькі до нуля значення, то серед нових хітів він перемістився у діапазон від 1 до 4, що є більш сприятливим для проникнення через ГЕБ (див. Мал.34).



Мал.34. Фармакофори на основі ліганду (ліворуч) і структури (праворуч). Сфери червоного кольору – ознаки акцептора Н-зв'язку, зелені – ознаки донора Н-зв'язку, жовті – ліпофільні ознаки, сині – ароматичне кільце.

Для віртуального скринінгу бібліотеки Enamine було застосовано обидва типи фармакофорних моделей — структурно-орієнтовану (SB) та лігандно-орієнтовану (LB).

Використання фармакофору SB дозволило ідентифікувати 72 потенційну сполуку, тоді як модель LB забезпечила відбір 91 хітів.

Таким чином, об'єднане застосування двох комплементарних підходів сприяло розширенню охоплення хімічного простору та збільшенню ймовірності виявлення активних та селективних інгібіторів АПФ2.

### 3.3.9. Результати докінгу UNISTRA

*Оцінка протоколу докінгу PLANTS перед віртуальним скринінгом.*

Перед застосуванням докінгового протоколу PLANTS для віртуального скринінгу було здійснено його попередню валідацію на здатність розпізнавати активні інгібітори АПФ2. Для цього використовували референтний набір з бази ChEMBL, який включав 61 сполуку з відомими константами інгібування  $pK_i > 7$ .

Однак, ефективність докінгу на основі такого набору має обмеження, оскільки:

- Сполуки у тестовому наборі мають подібний рівень активності;
- Докінгові алгоритми зазвичай не здатні точно розмежувати молекули з близькими  $pK_i$  значеннями, особливо в межах нано- та мікромольного діапазону.

Для підвищення чутливості оцінки, до тестового набору було додано 200 випадкових декой-сполук з ChEMBL, що не мають відомої активності щодо АПФ2. У цьому контексті протокол PLANTS продемонстрував адекватну здатність до ранжування:

- При використанні стандартної функції оцінювання ChemPLP, значення ROC AUC склало 0.71, що свідчить про помірну здатність відрізнити активні сполуки від неактивних.

Проте, ранжування на основі ефективності ліганду (визначеної як ChemPLP/ $N$ , де  $N$  — кількість важких атомів у молекулі) виявилось менш ефективним:

- ROC AUC = 0.59, що свідчить про знижену селективність при корекції на розмір молекули.

Застосування альтернативного нормалізованого критерію — ChemPLP  $\times N^{-0.35}$ , що враховує логарифмічну залежність функції ChemPLP від розміру молекули (за спостереженнями групи UNISTRA), дозволило суттєво покращити роздільну здатність:

- ROC AUC збільшився до 0.73.

Коефіцієнт 0.35 був визначений незалежно на основі спостережень UNISTRA, які показали, що значення ChemPLP корелює з кількістю важких атомів логарифмічно:  $\log(-\text{ChemPLP}) = 0.35 \log(N) + \text{константа}$

**Визначення хітів у віртуальному скринінгу бібліотеки Enamine.** Для ідентифікації перспективних хітів у рамках віртуального скринінгу бібліотеки Enamine було застосовано триетапний фільтр, який базувався на порівнянні з референтним лігандом з кристалічної структури 1R4L.

Критерієм відбору була одночасна перевага сполуки над референтом за трьома показниками:

1. Абсолютне значення ChemPLP — оцінка енергії взаємодії ліганду з мішенню;
2. ChemPLP/N — нормалізована ефективність зв'язування, скоригована на кількість важких атомів у молекулі (лігандна ефективність);
3.  $\text{ChemPLP} \times N^{-0,35}$  — скор, що усуває залежність від розміру молекули, з урахуванням логарифмічного характеру зростання оцінки зі збільшенням кількості атомів.

Тільки сполуки, які одночасно перевищували значення всіх трьох метрик для рідного ліганду, вважалися потенційними хітами.

У результаті було відібрано 245 сполук, які задовольняють цим критеріям, для подальшого візуального аналізу, відбору та експертної оцінки.

### **3.3.10. Результати моделювання UNISTRA QSAR**

Для вирішення двох основних завдань були використані методи моделювання зв'язку між структурою та властивістю (SAR) і кількісного зв'язку «структура-властивість» (QSAR): а) Передбачення активності проти АПФ2 для перевірки результатів скринінгу; б) Виявлення потенційних неселективних зв'язувань з антицілями АПФ і NEP.

Процедура створення консенсусної моделі для кожної конкретної цілі включала кілька етапів. Спочатку генерувалися ISIDA-дескриптори для набору навчальних даних, після чого модель навчалася за допомогою обраної функції ядра і алгоритму навчання.

Для оцінки точності та стабільності QSAR (регресійних) моделей використовувався класичний коефіцієнт детермінації  $R^2$ . Для оцінки моделей

SAR (класифікації) застосовувався показник збалансованої точності (BA – balanced accuracy).

Моделювання QSAR зіткнулося з кількома основними проблемами. По-перше, була обмежена кількість перевірених лігандів, що ускладнювало побудову точних моделей. По-друге, існувала проблема однорідності відомих сполук для АПФ2 та NEP, оскільки перевірені хемотипи були недостатньо різноманітними.

Ще однією суттєвою проблемою було відсутність даних про неактивні сполуки для АПФ2 у базі даних ChEMBL DB, що унеможливило створення ефективної класифікаційної моделі для АПФ2.

Крім того, структурна особливість АПФ у вигляді двох активних доменів створювала додаткові труднощі при побудові моделей. Незважаючи на високий рівень подібності між N- і C-доменами цього білка, їх функціональна активність суттєво відрізняється. Для деяких лігандів різниця в афінності до доменів може перевищувати дві логарифмічні одиниці концентрації, що критично для побудови коректної регресійної моделі.

Ускладнює ситуацію і те, що в більшості експериментальних джерел не вказується, до якого саме домену належить виміряна активність, через що тренувальні набори даних залишаються неоднозначними для точного машинного навчання.

Процедура побудови консенсусних моделей для кожної з мішеней включала кілька послідовних етапів:

- Генерацію ISIDA-дескрипторів на основі структур навчального набору;
- Навчання моделей із використанням відповідної функції ядра та методу оптимізації;
- Побудову QSAR-моделей (регресійних), для яких точність оцінювали за допомогою коефіцієнта детермінації  $R^2$ ;
- Побудову SAR-моделей (класифікаційних), які оцінювали за метрикою збалансованої точності (Balanced Accuracy, BA).

Основні показники таких моделей:

- Модель класифікації для АПФ (поріг активності = 25 нМ): BA = 0.97;
- Модель класифікації для NEP (поріг активності = 50 нМ): BA = 0.91;
- Регресійна модель для АПФ2:  $R^2 = 0.82$ .

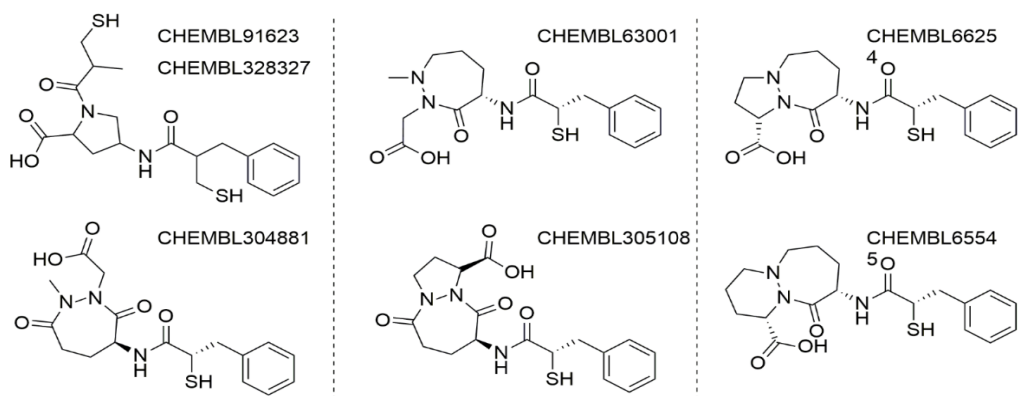
Всі ці моделі доступні на веб-сервері Страсбурга за адресою <http://infochim.u-strasbg.fr/webserv/VSEngine.html> і були використані для аналізу різноманітних хітів.

Хоча використана схема класифікації не є надмірно жорсткою, обраний емпіричний поріг активності 25 нМ для віднесення сполук до класу "активних" дозволив досягти майже ідеального розділення класів — збалансована точність (BA), перевірена методом крос-валідації, становила 0.97.

Такі високі значення точності зазвичай свідчать про те, що задача класифікації була "занадто простою". Це часто трапляється, коли набір активних сполук складається з хімічно однорідної серії, для якої алгоритм машинного навчання легко знаходить характерний структурний підпис у просторі дескрипторів.

Таким чином, побудована модель для АПФ здебільшого навчилася ідентифікувати типові хемотипи, найбільш часто представлені серед активних сполук у базі ChEMBL. У цьому контексті, якщо модель класифікує певну нову молекулу як "неактивну", це радше означає, що вона не належить до домінуючих хімічних серій відомих інгібіторів АПФ.

Крім того, всі такі висновки стосуються молекул, які знаходяться в межах області застосовності моделі (Applicability Domain, AD), тобто містять субструктурні фрагменти, що вже були представлені серед молекул, протестованих на АПФ (див. Мал.35).



Мал.35. Представники структурних сімей найкращих зв'язувачів АПФ

Область застосування регресійної моделі для АПФ2 виявилася вузькою — вона не покривала весь хімічний простір потенційних лігандів. Це пов'язано з використанням протоколу «контролю фрагментів», який позначає як «поза моделлю» будь-яку сполуку, що містить фрагменти, відсутні у навчальному наборі. Через обмежену кількість активних сполук у навчальних даних, модель не змогла дати достовірні прогнози щодо активності жодного з нових кандидатів.

Натомість класифікаційна модель для АПФ показала значно кращі результати: 52,5% хітів (123 сполуки) були класифіковані з високою достовірністю. На основі консенсусного прогнозу для кожної молекули було оцінено ризик побічної активності за шкалою від 0 до 7, де: 0 означає найнижчий рівень ризику (висока впевненість у неактивності), 7 — найвищий (висока ймовірність активності з підтвердженням обома частинами моделі).

Модель NEP показала низьку роздільну здатність, продемонструвавши лише поодинокі передбачення з достатньою впевненістю щодо "неактивності". Це частково пояснюється високим ступенем структурної подібності між NEP і АПФ: молекули, що активні для АПФ, можуть бути ще більш активними для NEP. До того ж, високий поріг для класифікації як "активна" може призводити до того, що навіть ліганди з  $K_i \approx 30$  нМ можуть бути помилково визначені як неактивні. Однак невелика кількість хітів, визнаних активними для АПФ або NEP, може свідчити про хімічну новизну цих молекул порівняно з уже відомими хемотипами інгібіторів. Лише 20 з них задовольняли усім емпіричним критеріям для оптимальної проникності через гематоенцефалічний бар'єр (ГЕБ).



Щоб зрозуміти специфічність АПФ2 та фармакофорні особливості нових інгібіторів АПФ2, також був проаналізований структурно збережений активний центр АПФ [90]. Загальна гомологія між доменами АПФ2 і АПФ становить приблизно 32% ідентичності послідовності, тоді як домен пептидази АПФ2 має 42% ідентичності послідовності (або 75% схожість) з N-доменом та 42% ідентичності (або 76% схожість) з C-доменом АПФ [70].

Цікаво, що під час молекулярно-динамічного (MD) моделювання комплекс «АПФ2–ліганд» демонстрував високу стабільність — орієнтація бічних ланцюгів залишалася практично незмінною протягом усієї симуляції. Для аналізу конформаційної гнучкості активного сайту без присутності ліганду ми провели 150-нс MD-симуляцію апо-форми АПФ2 з координуваним атомом цинку в центрі зв'язування. На першому етапі аналізу проводилася кластеризація на основі значень RMSD для атомів бічних ланцюгів у межах 10 Å від центру сайту зв'язування. Конформації з мінімальними відхиленнями (до 0,05 нм) від еталонного стану були включені до кластерів, що дозволило простежити еволюцію структури активного сайту з високою роздільною здатністю.

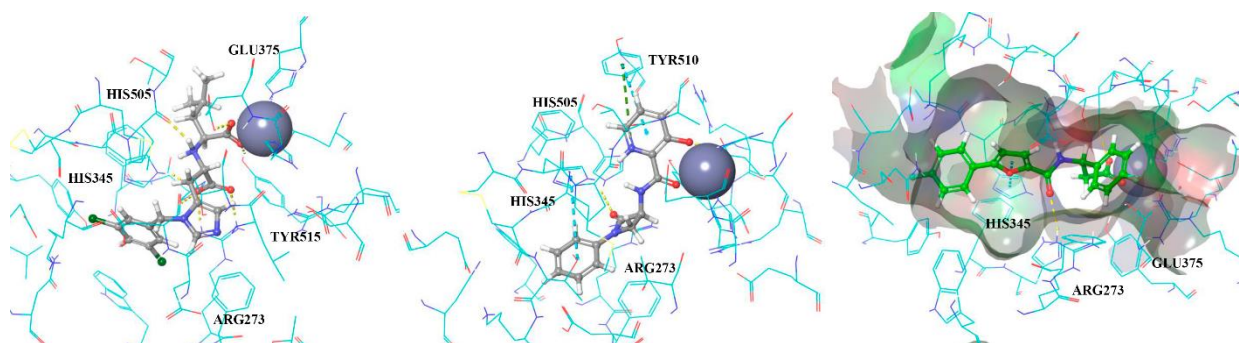
Після 120 нс розширення розщілини зупинилося, однак помітних змін у ключових амінокислотних залишках, відповідальних за взаємодію з лігандом, зафіксовано не було. Водночас, молекули води, які координуються з білком у кристалічних структурах АПФ2, залишалися стабільно закріпленими протягом усієї траєкторії. Це важливий момент, оскільки такі молекули можуть обмежувати здатність селективних інгібіторів АПФ2 взаємодіяти з активним сайтом АПФ і навпаки.

Подібна поведінка спостерігалася і для апо-форми АПФ. Після 150-нс симуляції жодні з координуваних молекул води, що були присутні у стартовій структурі, не змінили свого положення, що видно з аналізу траєкторій (див. Мал.36).

Ми застосували докінг для прогнозування стабільних конформацій і орієнтацій лігандів у сайті зв'язування білка, але не для оцінки їхньої спорідненості [24]. Найефективніша модель докінгу для АПФ2 включала

молекули води, які взаємодіють з залишками HIS345, HIS505 I TYR510 через Pi-Pi взаємодії; формували водневі зв'язки з ARG273; і також взаємодіяли з TYR515 або GLN375. Крім того, модель враховувала доступність до донора електронів поблизу атома металу (Мал.37). Для моделі докінгу АПФ також були враховані координаційні молекули води, виявлені у кристалічних структурах. В результаті 39 з 52 селективних інгібіторів, протестованих за цією моделлю, виявилися здатними зв'язуватися з сайтом АПФ.

Такий підхід забезпечив більш точне моделювання зв'язувального сайту та дозволив ідентифікувати 166 сполук із розрахунковими значеннями функції оцінки, подібними до тих, що були отримані для природного ліганду.



*Мал.37. Сокристалізований MLN-4760 утворює кілька стабільних конформацій у сайті зв'язування (А). Подібні розташування та взаємодії спостерігалися у двох сполук з найвищими оцінками (Z3488516360 і Z85905794) за результатами високопродуктивного скринінгу.*

Як результат три незалежні групи підібрали більше 600 сполук, після видалення перетинів та перевірки на наявність в Енамін Ін-Сток колекції отримали 577 сполук, які були напрувлені на біологічну валідацію активності проти АПФ2.

## РОЗДІЛ IV. *IN VITRO* СКРИНІНГ АПФ2 ІНГІБІТОРІВ З КОЛЕКЦІЙ КОМЕРЦІЙНО ДОСТУПНИХ СПОЛУК

Попри наявність комерційно доступних наборів для скринінгу інгібіторів АПФ2, зокрема Abscam (ab273373), оптимізовані та загальноприйняті протоколи для проведення HTS-досліджень наразі не опубліковані. Це підкреслює потребу у подальшій валідації таких підходів для забезпечення їхньої надійності та відтворюваності [32]. Аналіз базувався на ферментативній активності ангіотензинперетворювального ферменту 2 (АПФ2), здатного гідролізувати синтетичний пептидний субстрат із флуорофорною міткою на основі МСА, що супроводжується вивільненням вільного флуорофору.

У межах проведеної скринінгової кампанії були послідовно реалізовані наступні етапи:

- (i) оптимізація об'єму розчинів тестованих сполук без втрати якості аналітичного сигналу;
- (ii) оптимізація загального об'єму реакційної суміші та концентрацій ключових реагентів;
- (iii) визначення  $IC_{50}$  для відомого інгібітора MLN-4760 (EN300-22203532), а також референтного інгібітора з набору (структура не опублікована) у нових валідованих умовах;
- (iv) первинний скринінг 577 відібраних сполук;
- (v) підтвердження активності хітів;
- (vi) побудова дозозалежних кривих (dose-response curves, DRC) для відібраних активних сполук.

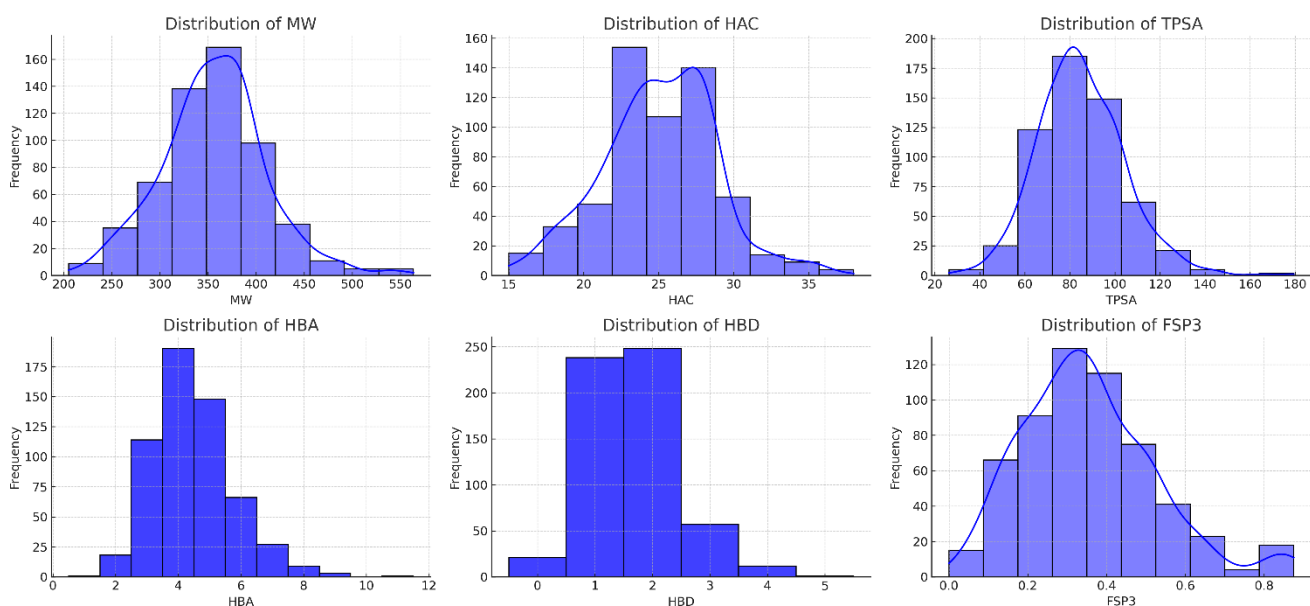


Рис.38. Графіки розподілу медико-хімічних параметрів 577 віртуальних хітів

### Біохімічний скринінг інгібіторів АПФ2

Для проведення дослідження використовували комерційний набір Angiotensin II Converting Enzyme (АПФ2) Inhibitor Screening Kit від Abcam (ab273373) та 384-лункові планшети Corning (3544) з низьким робочим об'ємом, чорним корпусом і прозорим дном. Підготовка всіх реагентів здійснювалася згідно з інструкцією виробника.

### Приготування розчинів ферменту та субстрату

Для приготування робочого розчину ангіотензинперетворюючого ферменту 2 (АПФ2):

- 20 мкл вихідного концентрованого розчину ферменту (концентрація та молекулярна маса не розкриваються) було розведено в 198 мкл буфера для розведення АПФ2 (надається в наборі);
- Отриманий розчин аліквотно розподіляли у пробірки по 25 мкл та зберігали при температурі  $-70\text{ }^{\circ}\text{C}$  до подальшого використання.

Робочий розчин субстрату для однієї проби готували безпосередньо перед використанням, змішуючи:

- 0,5 мкл вихідного розчину субстрату;
- 9,5 мкл буфера для аналізу АПФ2.

Розчин субстрату також зберігали при температурі  $-70\text{ }^{\circ}\text{C}$ . Референтний інгібітор з набору (5 мкл; структура та молекулярна маса не розкриваються виробником) не підлягав попередньому аліквотуванню. Безпосередньо перед аналізом до нього додавали 50 мкл буфера для аналізу АПФ2.

#### *Проведення аналізу*

Робочу суміш готували в об'ємі 25 або 40 мкл, залежно від формату планшета, використовуючи відповідно 0,5 або 0,8 мкл розчину ферменту АПФ2 та 12 або 19,2 мкл буфера для аналізу. У лунки для негативного контролю (без ферменту) вносили тільки буфер, у той час як у контрольні лунки (позитивний контроль) або лунки з інгібітором додавали 2,5 або 4 мкл буфера чи інгібітора відповідно. Після додавання всіх компонентів проводили інкубацію при кімнатній температурі протягом 15 хвилин. Далі до всіх лунок додавали субстратну суміш АПФ2 (10 або 16 мкл), після чого здійснювали вимірювання флуоресценції в кінетичному режимі протягом однієї години при кімнатній температурі ( $E_{xc} = 320\text{ нм}$ ,  $E_m = 420\text{ нм}$ ).

#### *Оптимізація концентрацій ферменту та субстрату*

Для знаходження оптимального співвідношення концентрацій ферменту та субстрату обидва розчини розводили в буфері 1:1, 1:2, 1:4, 1:8:

- Ферментні розчини додавали у відповідні лунки планшета в різних концентраціях;
- У негативні контролі додавали буфер;
- Планшет інкубували 15 хв при  $25\text{ }^{\circ}\text{C}$ ;
- У всі експериментальні лунки додавали субстрат у відповідних розведеннях;
- Планшет центрифугували і проводили кінетичне вимірювання ( $E_{xc}/E_m\ 320/420\text{ нм}$ ).

#### *Розрахунок $IC_{50}$ для інгібіторів*

Для перевірки отриманих умов на еталонному інгібіторі та сполуці EN300-22203532 проводили додатковий аналіз:

- Ферментний розчин додавали у лунки в розведенні 1:2;
- У негативні контролі додавали чистий буфер АПФ2;
- Обидва інгібітори додавали по 2,5 мкл у 10-кратній концентрації у відповідні лунки;
- Виконували 3-кратні серійні розведення;
- Планшет інкубували 15 хв при 25 °С;
- Додавали розчин субстрату (1:2);
- Планшет центрифугували і проводили флуоресцентний аналіз (Ехс/Ет 320/420 нм).

#### *Первинний скринінг інгібіторів АПФ2:*

- Під час первинного скринінгу використовували два інгібітори як еталонні;
- Розчини ферментів і субстратів готували та розводили 1:2;
- З кінетичних кривих обирали дві часові точки, для яких розраховували значення RFU (відносна флуоресценція);
- Відношення  $\Delta\text{RFU}/\Delta t$  визначали для кожної кривої, після чого інгібування розраховували за формулою:  

$$\text{Інгібування}(\%) = 100\% - \left( \frac{\Delta\text{RFU зразка}}{\Delta\text{RFU ферментного контролю}} \times 100\% \right);$$
- Dose-response curves (DRC) будували в GraphPad Prism, використовуючи сигмоїдальну модель з варіабельним нахилом.

### **4.1. Методи та планування**

#### *Скринінг інгібіторів АПФ2*

На цьому етапі проекту було здійснено скринінг 577 сполук, відібраних на основі *in silico* прогнозування, з метою оцінки їхньої здатності інгібувати активність ангіотензинперетворюючого ферменту 2 (АПФ2). Для експериментальної перевірки використовували комерційний набір Abcam (ab273373), придбаний у рамках проекту.

Принцип дії тест-системи ґрунтується на ферментативному розщепленні синтетичного пептидного субстрату, міченого метилкумариновим флуорофором (MCA). При розщепленні вивільняється флуоресцентна мітка, інтенсивність світіння якої пропорційна активності ферменту. Вимірювання проводили за допомогою мікропланшетного флуоресцентного зчитувача, що дозволяло здійснювати кількісну оцінку.

Перед початком основного скринінгу була проведена оптимізація об'ємів реакції (25 або 40 мкл), що дозволило зберегти високу чутливість методу. Методика була валідувана з урахуванням обраних концентрацій реагентів. У межах валідації було визначено значення  $IC_{50}$  для синтезованої в межах проєкту сполуки EBC-36033 (MLN-4760), а також для еталонного інгібітора DX600 (Ac-GDYSHCSPLRYPWKCTYPDPEGGG-NH<sub>2</sub>) у нових умовах.

Після цього проведено первинний скринінг усіх 577 сполук, відібрано хіти, які пройшли подальшу валідацію, включно з побудовою дозозалежних кривих (DRC) для підтверджених активних молекул.

#### **4.2. Матеріали, реагенти та обладнання**

**Тип плати:** 3544 Corning, low volume, black clear bottom

**Об'єм:** 25/40  $\mu$ л

**Концентрації:** *невідомі*, всі розчини були приготовані у відповідності до інструкції з набору

**Буферні розчини:** отримані від виробника (склад не розголошений)

**Температура:** 25°C

**Час інкубації:** 15 хв

**Вимірювання флуоресценції (Exc/Em 320/420nm) у кінетичному варіанті:**

1 година

**Приготування реагентів:**

**Ензим** (*концентрація та молекулярна маса у отриманому розчині не розголошена*): розчинити 20  $\mu$ л отриманого розчину ензиму у 198  $\mu$ л

відповідного буферного розчину для розведення АПФ2. Відбиралися аліквоти по 25  $\mu\text{л}$  та заморожувалися до  $-70^{\circ}\text{C}$ .

**Субстрат** (концентрація та молекулярна маса у отриманому розчині не розголошена): Відбиралися аліквоти по 10  $\mu\text{л}$  та заморожувалися до  $-70^{\circ}\text{C}$ .

**Для однієї лунки відбиралося:** 0.5  $\mu\text{л}$  розчину субстрату + 9.5  $\mu\text{л}$  приготованого розчину ензиму.

**Референтний інгібітор:** (DX600, концентрація не розголошена): 5  $\mu\text{л}$ . Додати 50  $\mu\text{л}$  відповідного буферного розчину для розведення АПФ2 до вимірювань.

### **Протокол для вимірювань:**

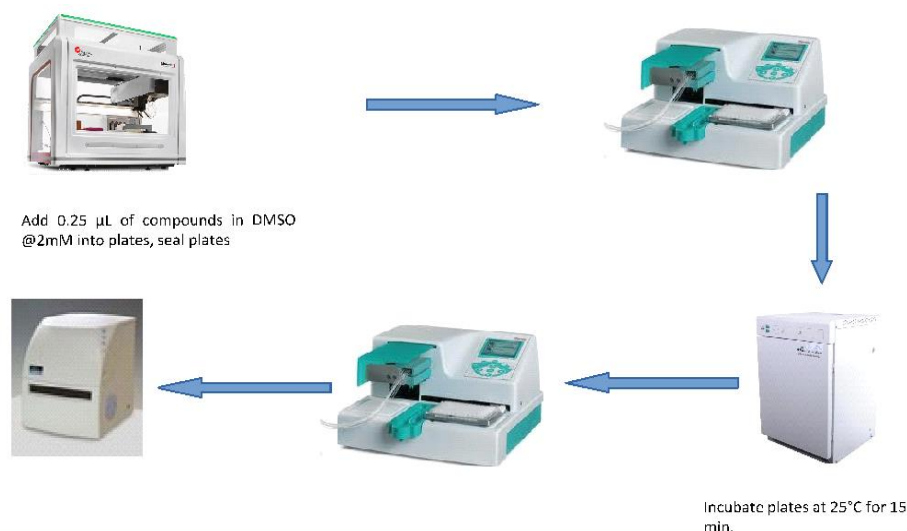
**Для 25/40  $\mu\text{л}$ :**

- 1) Додати 0.5/0.8  $\mu\text{л}$  розчину ензиму до 12/19.2  $\mu\text{л}$  відповідного буферного розчину для розведення АПФ2, ретельно перемішати.
- 2) Використовувати 12.5/20  $\mu\text{л}$  відповідного буферного розчину для розведення АПФ2 як негативний контроль.
- 3) Додати 2.5/4  $\mu\text{л}$  розчину досліджуваної сполуки у відповідну лунку

### **Позитивний контроль (референтний інгібітор):**

- 4) Додати 2.5/4  $\mu\text{л}$  відповідного буферного розчину для розведення АПФ2 у лунку для ензиму (позитивний контроль);
- 5) Інкубувати 15 хв при кімнатній температурі;
- 6) Додати 10/16  $\mu\text{л}$  розчину субстрату до лунок **негативний контроль, позитивний контроль, контроль інгібітора (ЕВС-36033);**
- 7) Виміряти флуоресценцію (Ехс 320 нм, Ем 420 нм) у кінетичному варіанті: 1 година при кімнатній температурі.

Загалом дизайн експерименту можна зобразити наступною схемою (Мал.39.):



Мал.39. Схематизована послідовність дій для вимірювання інгібуючих властивостей щодо ACE2 досліджуваних сполук.

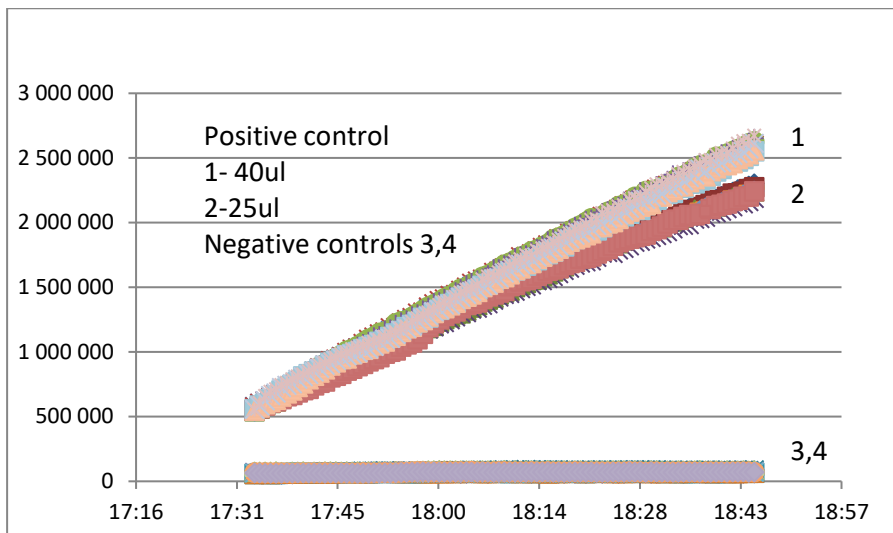
### 4.3. Оптимізація методу

#### *Результати попередніх оптимізаційних експериментів*

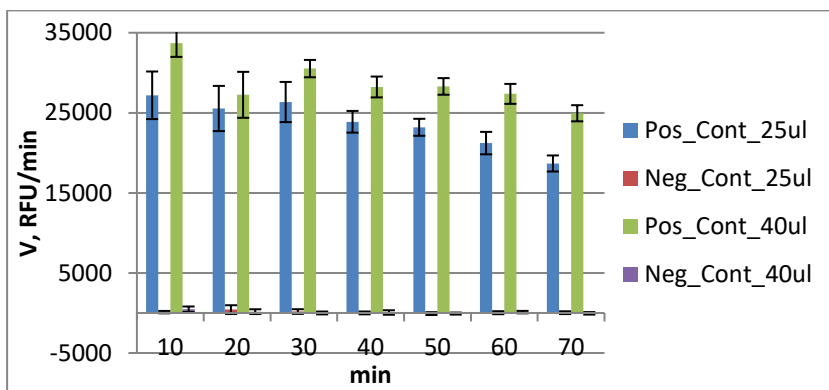
У дослідженні використано здатність ферменту АПФ2 розщеплювати синтетичний пептидний субстрат на основі метилкумаринового флуорофора (МСА), у результаті чого вивільняється флуоресцентний сигнал. Цей сигнал зростає в процесі реакції та піддається кількісному вимірюванню за допомогою флуоресцентного мікропланшетного рідера.

На етапі оптимізації було підбрано об'єм реакційної суміші, що забезпечував високу чутливість та відтворюваність методу без втрати якості аналізу (див. Мал.40, Табл. 6). Отримані дані підтверджують ефективність скринінгової системи: Z-приймальне значення, обчислене на основі швидкості реакції, варіювалося в межах 0,6–0,8, що відповідає прийнятним критеріям якості для високопродуктивного скринінгу. Співвідношення сигнал/фон (S/B), розраховане за відносною флуоресценцією, становило 25–30 при об'ємі 25 мкл та 40–50 при 40 мкл (Мал.41).

Водночас під час інкубації було зафіксовано поступове зниження швидкості реакції, ймовірно спричинене підвищенням температури з 25 °C до 27 °C упродовж 70 хвилин, що підкреслює важливість контролю температурного режиму під час проведення вимірювань.



Мал.40. Кінетичні криві позитивного (1 – 40  $\mu$ L; 2 - 25  $\mu$ L) та негативного (3 – 40  $\mu$ L; 4 - 25  $\mu$ L) контролю.



Мал.41. Залежність відносної швидкості реакції (RFU/t) від часу реакції для позитивного та негативного контролів при 40  $\mu$ L та 25  $\mu$ L.

Табл.6. Значення стандартних відхилень та коефіцієнтів варіацій відносної швидкості реакції (RFU/t) при різних об'ємах та часі реакції.

Time, min	Pos_Cont_25ul			Neg_Cont_25ul		Pos_Cont_40ul			Neg_Cont_40ul	
	Average	SD	CV	Average	SD	Average	SD	CV	Average	SD
10	27186.5	2972.5	10.9	105.6	150.6	33693.1	1705.9	5.1	466.0	342.4
20	25538.0	2822.9	11.1	414.4	549.4	27246.3	2869.6	10.5	163.7	298.0
30	26348.2	2509.5	9.5	177.9	291.1	30522.3	1083.2	3.5	0.5	178.6
40	23878.1	1356.4	5.7	5.1	175.7	28227.1	1303.6	4.6	59.6	280.5
50	23191.4	1062.0	4.6	-79.7	173.5	28295.7	1036.8	3.7	-74.6	126.4
60	21215.8	1402.6	6.6	32.6	186.5	27364.4	1246.4	4.6	105.9	152.7
70	18673.5	1000.4	5.4	34.8	173.5	24943.7	1008.6	4.0	-44.2	157.8

Наступним кроком було вивчення можливості зменшення кількості ензиму та флуорогенного субстрату за рахунок розведення ензиму та субстрату. Результати цих досліджень наведені у таблиці 7.

Табл. 7. Оптимізація концентрації (тобто розведення) ензиму та субстрату.

Substrate	1:1		1:2		1:4		1:8	
Enzyme	S/B	Z-Prime	S/B	Z-Prime	S/B	Z-Prime	S/B	Z-Prime
1:1	24.29	0.92	13.14	0.87	7.26	0.80	4.44	0.82
1:2	13.61	0.90	7.68	0.94	4.46	0.90	2.97	0.79
1:4	6.77	0.94	4.09	0.87	2.62	0.78	1.92	0.82
1:8	4.56	0.89	2.87	0.80	1.96	0.73	1.56	0.60

У всіх тестованих умовах параметри якості відповідали критеріям, встановленим для високопродуктивного скринінгу ( $Z\text{-prime} > 0,5$ ). Водночас при зниженні концентрації ферменту або субстрату було зафіксовано суттєве зменшення співвідношення сигналу до фону (S/B), що вказує на зниження чутливості методу. Тому оптимальним виявилось використання розведень не більше ніж 1:4 як для ферменту, так і для субстрату.

Для перевірки надійності методики визначення  $IC_{50}$  було використано два референтні інгібітори. Перший — інгібітор з набору, для якого виробником заявлене  $IC_{50}$  на рівні 33 нМ (визначене в рекомендованих умовах). Другий — MLN-4760 (EN300-22203532, з колекції Enamine), для якого в базі даних PubChem зазначено значення  $IC_{50} = 0.44$  нМ. Обидві сполуки використовувалися як еталонні контрольні інгібітори в процесі первинного скринінгу.

У ході тестування:

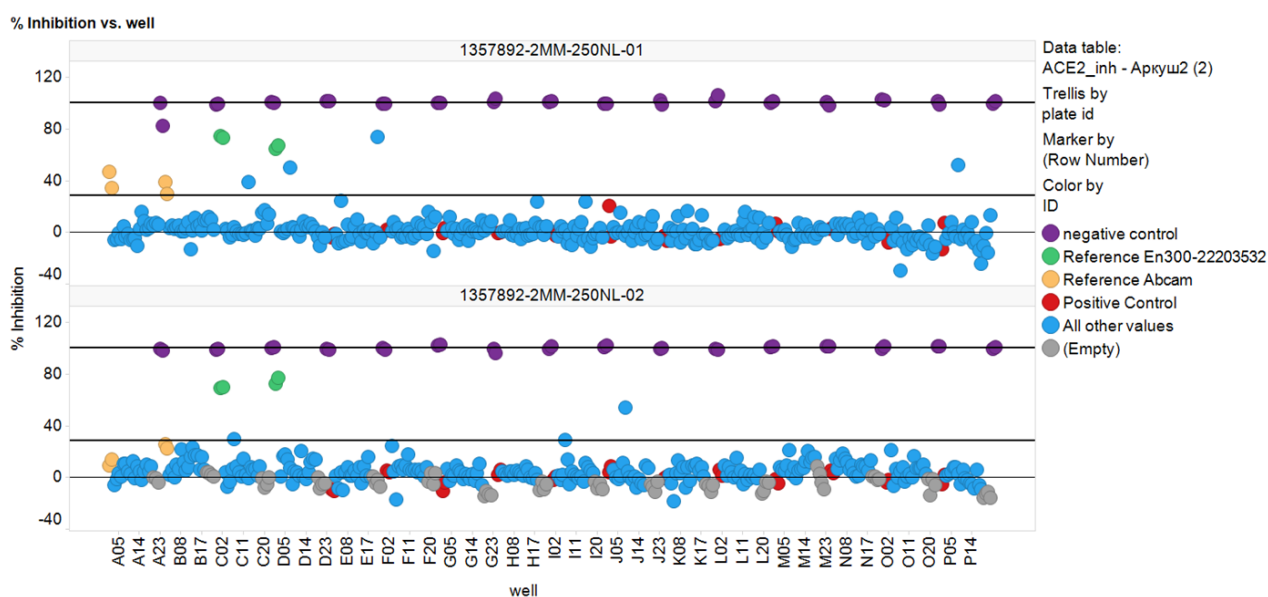
- Розчини ферменту та субстрату готувалися в розведенні 1:2;
- Для кожної кінетичної кривої обиралися дві часові точки, за якими розраховували значення відносної флуоресценції (RFU);
- Показник  $\Delta RFU/\Delta t$  використовували для оцінки швидкості реакції;
- Відсоток інгібування розраховували за формулою:

$$\text{Інгібування (\%)} = 100\% - (\Delta RFU_{\text{сполуки}} / \Delta RFU_{\text{контролю}}) \times 100\%,$$

відповідно до інструкцій виробника тест-набору.

#### 4.4. Скринінг винайдених в результаті докінгу хітів

Отже з використанням оптимізованих параметрів було проведено скринінг 577 сполук, результати якого наведені на Мал.42.



Мал.42. Скринінг 577 сполук щодо інгібуючої активності до ACE2.

Параметри «якості» при порівнянні результатів двох плат задовольняли сучасним критеріям та наведені у таблиці 8.

Табл.8. Порівняльні характеристики скринінгових плат.

Parameter	Плата 1	Плата 2
Z''	0.84	0.82
S/B	5.54	5.2
%inh_DX600	37.5	17.9
SD_%inh DX600	7.2	7.4
%inh_Ref EBC-36033	69.9	72.2
SD_%inh EBC-36033	4.8	3.7

За результатами скринінгу 577 сполук було відібрано 7, що відповідали критеріям хітів ( $Inh\% > 3 \cdot SD + Avg$ , що дорівнювало 28.2 (перша плата) та 28.8 (Друга плата). Дані наведені у таблиці 9. Слід відмітити, що 5 сполук з 7 мали свою інтенсивну флуоресценцію при 320exc/420em, однак це не завадило розрахунку %inhibition тому що додатково використовувався критерій  $\Delta RFU/\Delta t$  та параметри «якості» задовольняли критерії скринінгу ( $Z\text{-prime} < 0.5$ ).

Табл.9. Сполуки, що були ідентифіковані як хіті згідно критерію  $Inh\% > 3*SD+Avg$ .

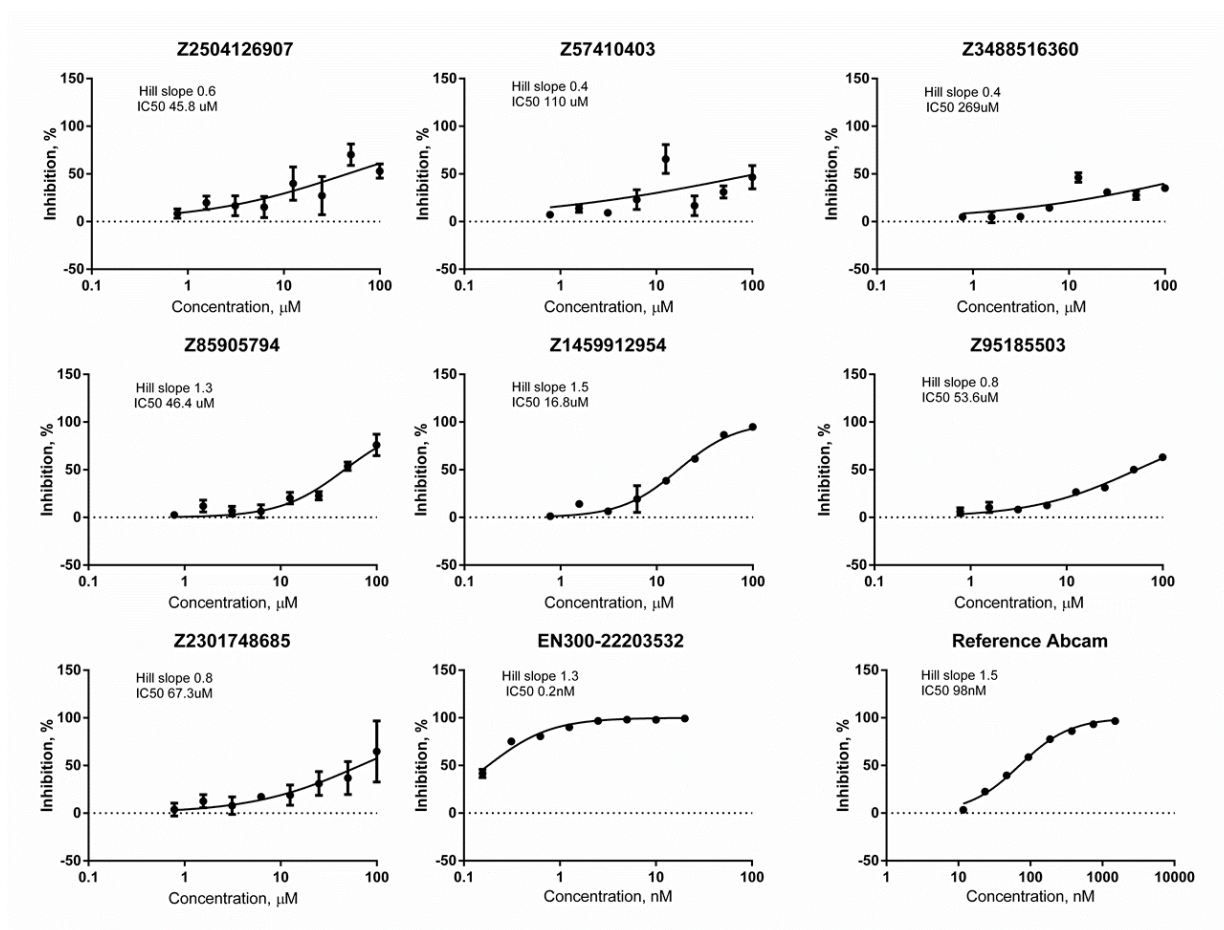
Плата	Лунка	ID сполуки	Inhibition%
1	E21	Z57410403	74.1
2	J08	Z1459912954	54
1	P08	Z3488516360	52.1
1	D07	Z2301748685	50.1
1	C13	Z2504126907	38.8
2	C07	Z95185503	29.8
2	I06	Z85905794	28.9

Останнім етапом цього блоку дослідження було визначення дозозалежних кривих для всіх сполук, які були попередньо ідентифіковані як хіти. Результати цих вимірювань наведено на Малюнку 43.

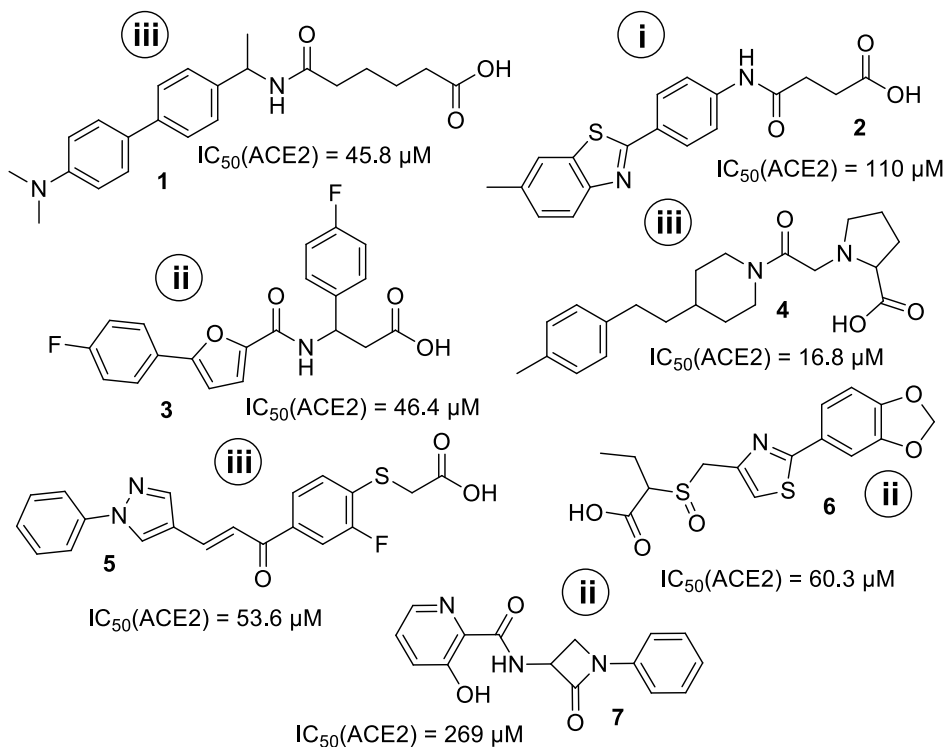
Структури відібраних хітів представлено нижче. Варто зазначити, що підхід віртуального скринінгу, орієнтований на хімічне різноманіття (i), який було спеціально розроблено для цього проєкту та детально описано в цьому звіті, дозволив виявити лише один хіт з 200 протестованих сполук.

Натомість два інші підходи — більш традиційні — які ґрунтуються на 3D фармакофорному моделюванні (ii, 177 сполук) та методі generative topographic mapping (GTM) (iii, 200 сполук), продемонстрували вищу ефективність у виявленні активних сполук. (див. Мал.44).

Водночас, через відносно невеликий масштаб проведеної скринінгової кампанії, робити остаточні висновки щодо переваг того чи іншого підходу наразі передчасно.



Мал.43. Дозозалежні криві винайдених хітів та референсних сполук.

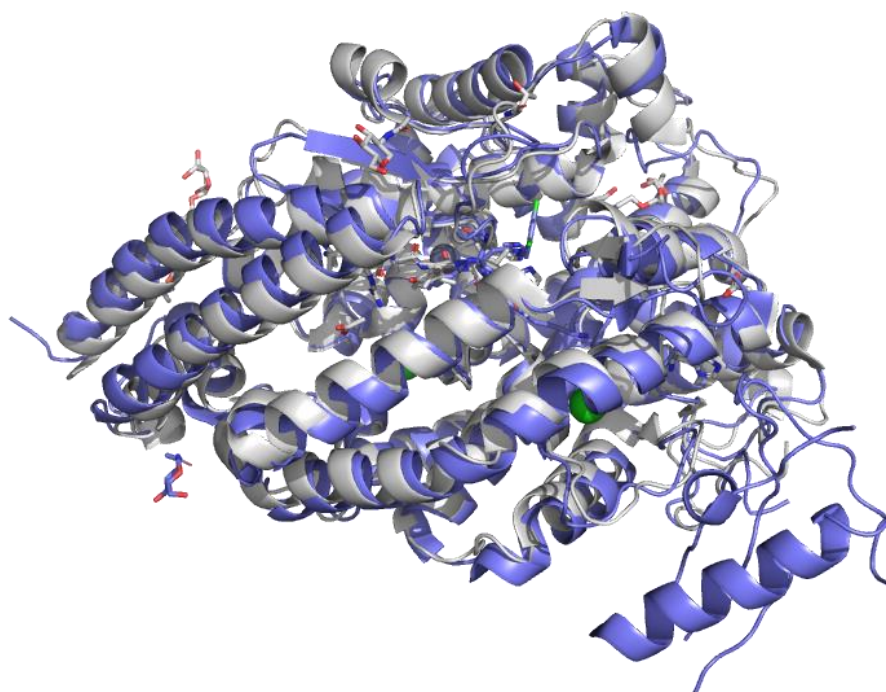


Мал.44. Структури винайдених хітів.

## РОЗДІЛ V. ВІРТУАЛЬНА ОПТИМІЗАЦІЯ СПОЛУК-ЛІДЕРІВ ТА ДОДАТКОВИЙ ВІРТУАЛЬНИЙ СКРИНІНГ REAL SPACE

Перш за все ми поставили за мету більш глибоко дослідити ізоформи АПФ2 і змоделювати використовуючи молекулярну динаміку шарнірно-згинальний рух субдомену, залежного від ліганду.

Кристалічної структури АПФ2 з природнім субстратом Ангіотензином 2 (Ang2) в банку кристалічних структур PDB немає. Але провівши пошук схожих сполук, була знайдена структура 4APH [73] – це кристалічна структура АПФ1 в комплексі з Ангіотензином 2. Цікавим фактом є те, що ця структура має закритую форму, та досить точно співпадає із закритою формою АПФ2 (1R4L) (Мал.45).

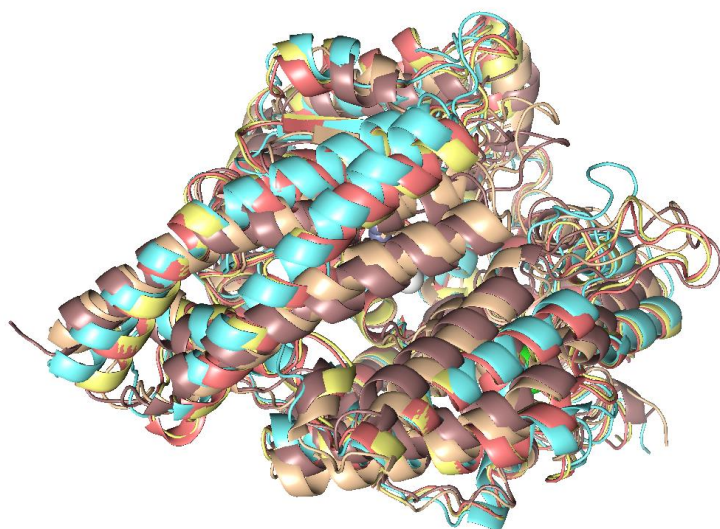


Мал.45. Порівняльне вирівнювання двох PDB структур - 1r4l синім кольором, 4aph – сірим.

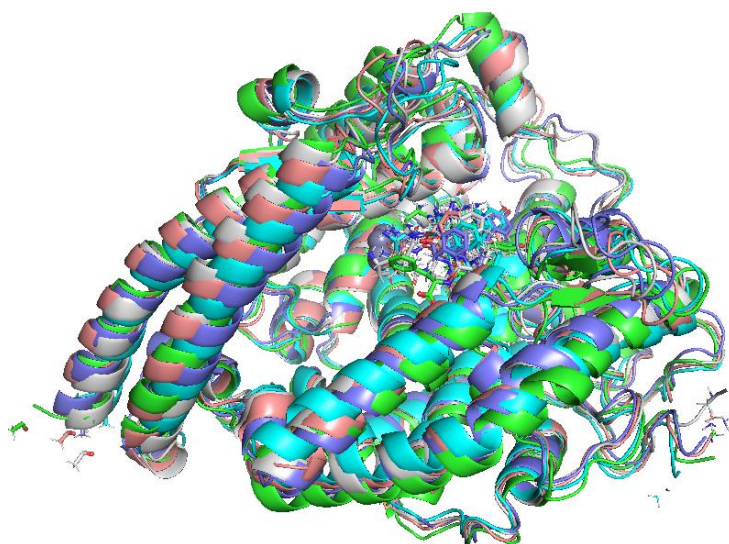
Таким чином, ми можемо зробити висновок, що реакція перетворення біологічного субстрату (ангіотензину 2) АПФ1 та АПФ2 відбувається в закритій конформації, але для входу субстрату, дуже важливим є шарнірний рух, який був описаний раніше.

Ми провели молекулярний докінг Ангіотензину 2 у відкриту та закритую форму АПФ2 з подальшою молекулярною динамікою за наступним протоколом: 100нс NVT ансамбль при 250К, 100нс NPT ансамбль при 250К, 100нс NPT ансамбль при 300К та ще 100нс NVT ансамбль при 300К комплексів 1r4l+Ang2

та 1r42+Ang2. Далі провели порівняння впливу Ang2 на зміну конформації АПФ2. На малюнках 46-47 ми можемо спостерігати, що природній субстрат Ang2 призводить до відкриття закритої структури під час молекулярної динаміки, в той же час відкрита форма майже не змінюється. Це ще раз підтверджує важливість переходу відкрита-закрита форма для функціонування протеїну. Але звісно для повної симуляції шарнірно-згинальний рух потрібно проводити набагато довші МД симуляції.



*Мал.46. Ілюстрація зміни конформації структури 1R4L(закрита форма)+Ang2 під час проведення МД симуляції. (коричневий – 0нс, бежевий – 100нс, жовтий – 200нс, кораловий – 300нс, бірюзовий – 400нс.*

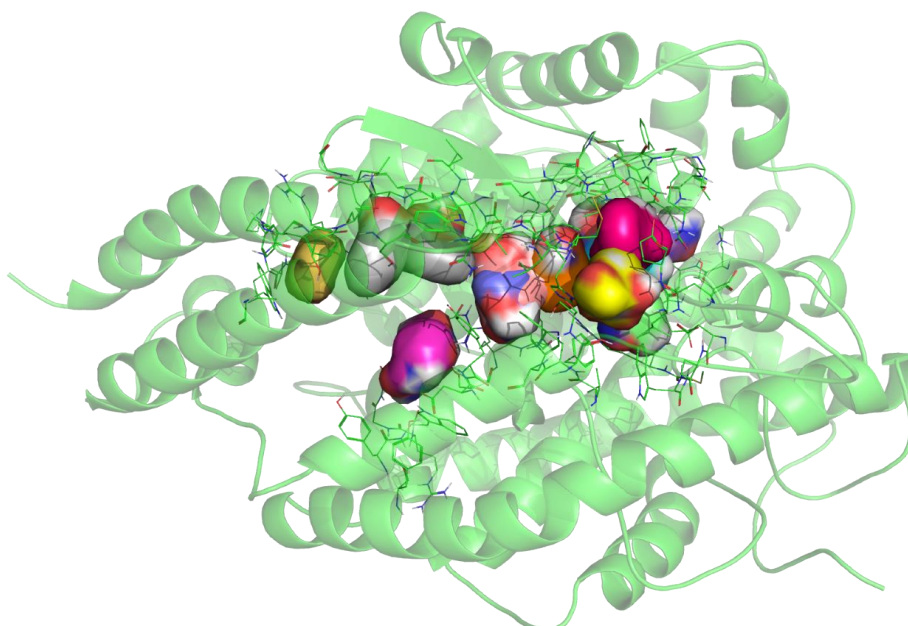


*Мал.47. Ілюстрація зміни конформації структури 1R42 (відкрита форма)+Ang2 під час проведення МД симуляції. (салатовий – 0нс, бірюзовий – 100нс, кораловий – 200нс, сірий – 300нс, світло-фіолетовий – 400нс.*

### **5.1. Функціональний топологічний мапінг потенційних субпocketів**

FTMap — це обчислювальний сервер, який ідентифікує гарячі точки зв'язування на поверхні макромолекул, тобто регіони, що вносять найбільший вклад у вільну енергію зв'язування ліганда. Для роботи з FTMap користувачі

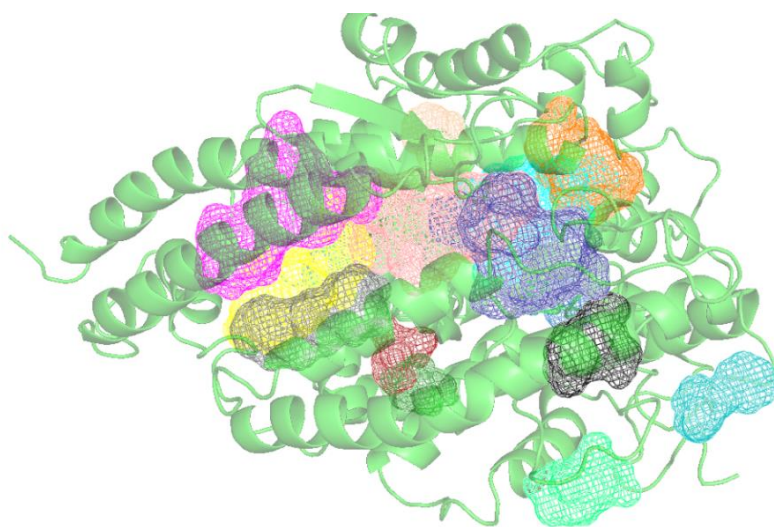
завантажують структуру білка, ДНК або РНК у форматі PDB. Сервер виконує вибірку мільярдів положень малих органічних молекул, які використовуються в якості зондів, і оцінює їх за допомогою детальної енергетичної функції. Гарячі точки зв'язування визначаються на основі зв'язування кількох типів зондів, що добре узгоджується з експериментальними даними. FTMap дозволяє картувати сайти зв'язування лігандів, враховувати гнучкість бічних ланцюгів білка, параметризувати додаткові зонди та працювати з ансамблями білкових структур (Мал.48).



Мал.48. Сфери доступні для розташування лігандів та вигідних взаємодій – зроблено FTMap.

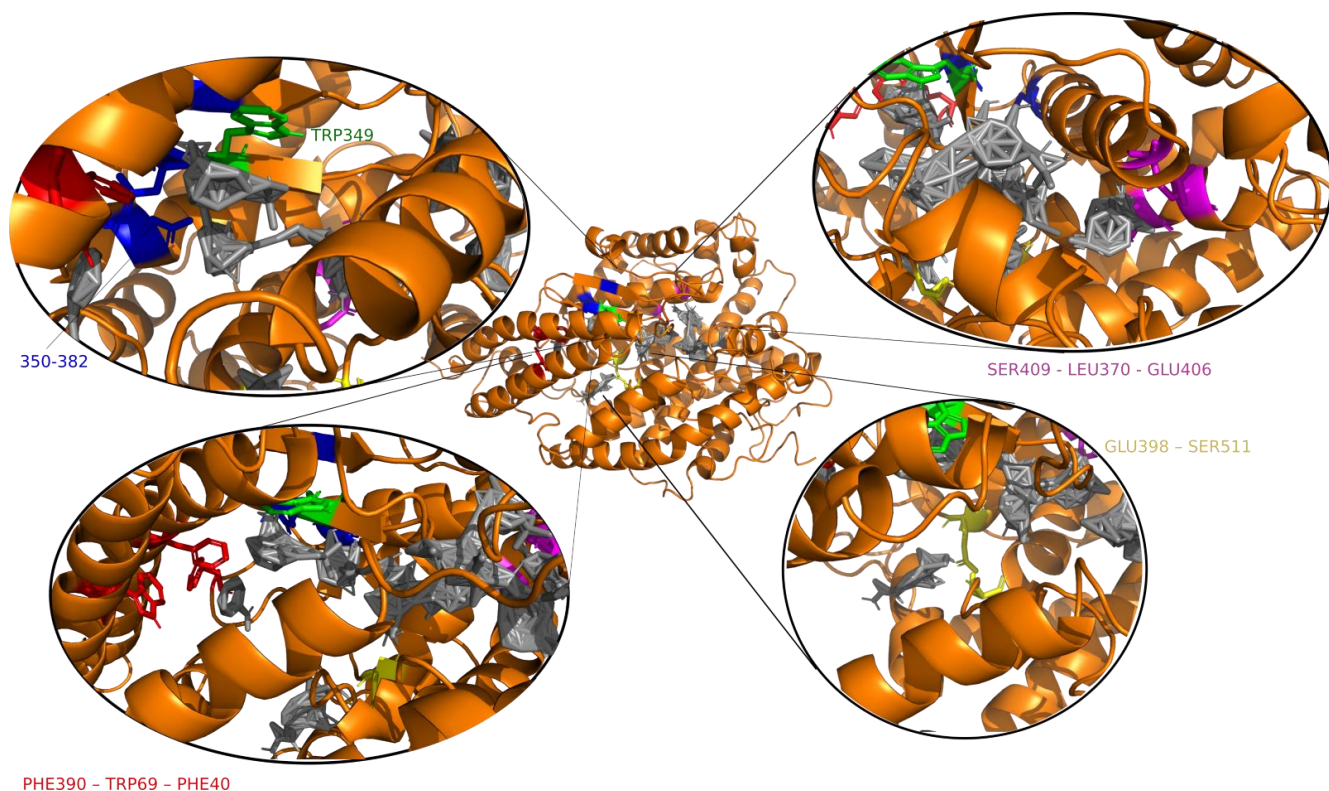
Цей метод використовують для визначення здатності білків до зв'язування з лігандами, ідентифікації важливих фрагментів лігандів, пошуку найбільш зв'язуваних конфігурацій у ансамблях білкових структур та для фрагментного дизайну ліків. FTMap є більш точним, ніж класичні методи картування, і значно швидшим за новітні підходи, засновані на молекулярній динаміці.

Ідентифікувавши гарячі точки зв'язування в структурі АПФ2 (PDB код – 1R4L) ми виявили наступні потенційні «зони росту» молекул для підвищення афінності хітів та більш точного пошуку потенційних байндерів з інших набагато більших баз Enamine. Також ми провели дослідження інструментом FTMove для виявленні нових сайтів зв'язування, особливо тих, що з'являються внаслідок конформаційних змін білка (Мал.49).



Мал.49. FTMove карта потенційних місць зв'язування.

Проаналізувавши ми виявили наступні амінокислотні залишки, які можуть бути потенційно задіяні у взаємодії з потенційними інгібіторами (див Мал.50.), а саме - гідрофобний субпocket PHE390 – TRP69 – PHE40, полярний субпocket 350-382, стекінг - TRP349, полярний субпocket GLU398 – SER511, полярний субпocket SER409 - LEU370 – GLU406.



Мал.50. Аналіз додаткових субпocketів - гідрофобний субпocket PHE390 – TRP69 – PHE40 (позначено червоним), полярний субпocket 350-382 (позначено синім), стекінг - TRP349 (зеленим), полярний субпocket GLU398 – SER511 (жовтим), полярний субпocket SER409 - LEU370 – GLU406 (рожевим).

Ці знахідки будуть застосовані для аналізу докінгових поз та відбору потенційних хітів для додаткової стадії віртуального скринінгу.

## 5.2. Пошук та дослідження активаторів

Цікавий факт, що в попередньому огляді та під час пошуку відомих активних сполук проти АПФ2 в базі ChEMBL не було знайдено жодного активатора. В той же час на платформі Guide to Pharmacology [44] було знайдено цікаві данні, а саме окрім інгібіторів, були знайдені записи про активатори АПФ2. Для побудови та поліпшення моделі інформація як про інгібітори так і про активатори може бути дуже корисною для розуміння механізму функціонування протеїну в природі, а також механізму інгібування/активування. Тому ми дослідили ці дані, та спробували знайти спосіб зв'язування біологічно підтверджених активаторів. З точки зору інгібіторів АПФ2, ці результати є важливими, оскільки ідентифікація специфічних взаємодій може бути використана для створення антагоністів, що блокують функціональну активність ферменту. Дослідження [87] ґрунтується на молекулярному докінгу та структурному аналізі потенційних лігандів, відібраних з бібліотеки NCI/DTP.

*Для молекулярного докінгу в дослідженні [87] було обрано три структурні ділянки АПФ2: Ділянка 1 – це структурна зона АПФ2 (PDBID 1R42), яка була визначена для докінгу в дослідженні. Вона розташована в шарнірній зоні між двома субдоменами ферменту і знаходиться у відкритій конформації (substrate-free). Саме ця ділянка виявилася найбільш придатною для зв'язування лігандів серед трьох досліджених.*

Це означає, що молекули, здатні зв'язуватися з Ділянкою 1, можуть змінювати конформацію ферменту, що може бути використано як для активації, так і для потенційного інгібування його активності залежно від типу взаємодії. Ділянка 2 знаходиться віддалено від активного центру та вибиралася через структурні відмінності між відкритою (1R42) та закритою (1R4L) конформацією АПФ2. Ділянка 3 знаходиться на межі субдоменів 1 і 2 біля входу в активний центр. Нажаль у статті не окреслені детально амінокислотні залишки зазначених ділянок.

Стосовно біологічних досліджень, авторами було проведено тести на тваринних моделях (Wistar-Kyoto та спонтанно гіпертензивні щури, SHR). Найбільш перспективними виявилися сполуки XNT (CAS# 86456-22-6) та резорцинолнафталеїн, які посилювали активність АПФ2 у 1.8-2.2 рази. Встановлено, що гостре введення XNT викликало дозозалежне зниження артеріального тиску, особливо виражене у SHR (~71 мм рт. ст. при 10 мг/кг). Хронічне введення XNT сприяло зменшенню фіброзу серця та нирок з  $рEC_{50} = 4,7$ . Резорцинолнафталеїн також посилював активність АПФ2 залежно від дози [41]. *In vivo*: резорцинолнафталеїн викликав значну активацію АПФ2 як у здорових, так і у хворих щурів через 7 днів після лікування. Лікування резорцинолнафталеїном запобігало високому РАР, гіпертрофії правого шлуночка та утворенню неоінтими. Резорцинолнафталеїн викликав покращену ендотеліально-залежну вазорелаксацію та зниження рівня прозапальних цитокінів, таких як TNF- $\alpha$ , MCP-1, IL-6, і підвищення рівня протизапального цитокіну IL-10 на ранній стадії патогенезу [41].

Табл.10. Структура найактивніших активаторів.

	<p>SMILES (XNT) -  <chem>OCC1=CC=C(C=C1)OS(=O)(=O)C2=CC=C3C(=O)OC(=O)C=C3C=C2CNCCN(C)C</chem>          В ChEMBL структура не знайдена, що ще раз доводить необхідність пошуку у декількох публічних базах.</p>
	<p>SMILES (Резорцинолнафталеїн) –  <chem>Oc1ccc2c(c1)Oc1cc(O)ccc1C12OC(=O)c2cccc3ccccc1c23</chem>          В ChEMBL структура не знайдена</p>

Патент WO2008066770A2 [51] описує сполуки-активатори АПФ2 та методи їх застосування для лікування серцево-судинних та легневих

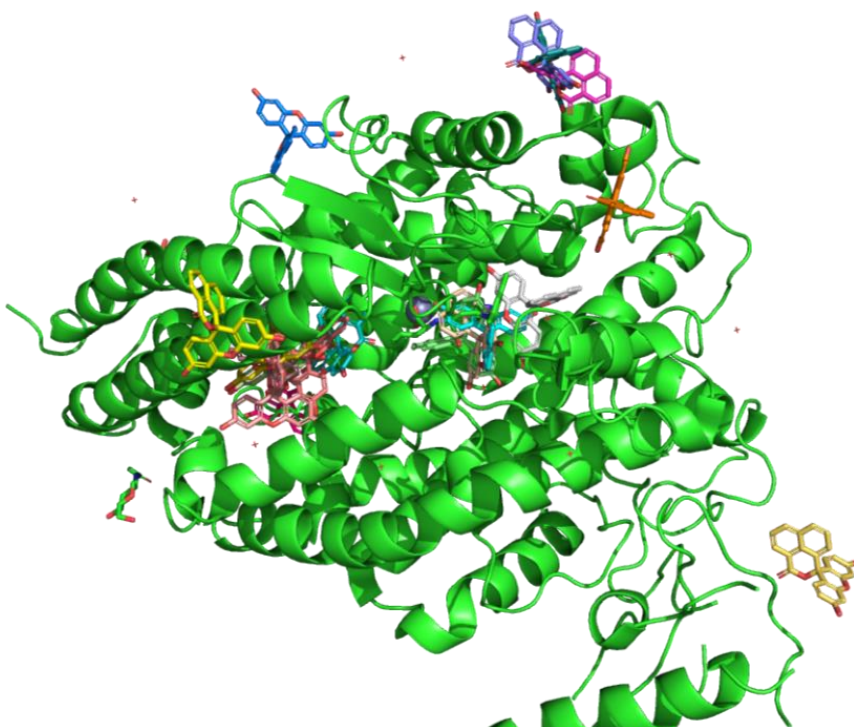
захворювань. Ці сполуки сприяють активації ангіотензинперетворювального ферменту 2 (АПФ2), що може бути корисним у терапії таких станів, як гіпертонія та легенева артеріальна гіпертензія.

У патенті детально описано хімічну структуру цих сполук, способи їх синтезу та результати доклінічних досліджень, які демонструють їх ефективність у моделюванні відповідних захворювань. Також розглядаються можливі фармацевтичні композиції та режими дозування для клінічного застосування.

Через те, що у нас немає достовірних доказів сайту зв'язування для обох ідентифікованих активаторів, було проведено моделювання для виявлення нового або підтвердженню запропонованого сайту зв'язування.

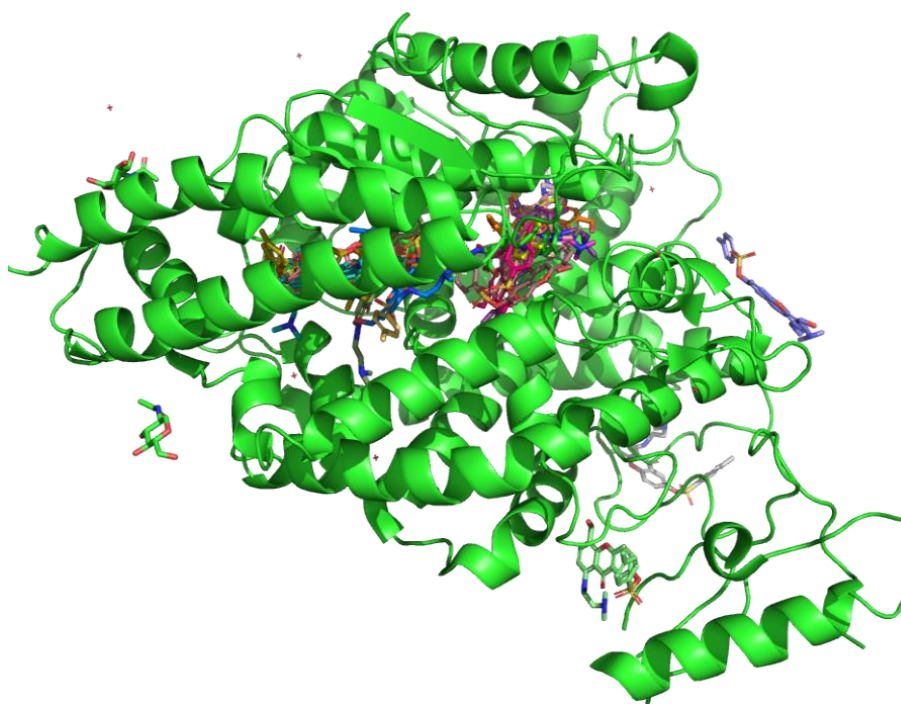
*DiffDock* – це новітній підхід до молекулярного докінгу, заснований на глибокому навчанні, розроблений командою NVIDIA. Основна відмінність цього методу від традиційних докінгових алгоритмів полягає у використанні дифузійних моделей (diffusion models), що дозволяють прогнозувати взаємодію лігандів із білками з вищою точністю. Цей метод значно підвищує точність передбачення зв'язування молекул та може бути особливо корисним для ідентифікації потенційних сайтів зв'язування в дослідженнях інгібіторів АПФ2.

За передбаченням Резорцинолнафталеїн (RF) має 8 потенційних сайтів зв'язування, але найбільш вирогідними є дві області в порожнині активного сайту, в місті активаційного повороту. В цій позиції із двадцяти запропанованих поз 11 знаходиться саме в цій області, але найвирогідніша співпадає з областю взаємодії відомого інгібітора MLN-4760 (Мал.51).



Мал.51. Передбачення поз активатору резорцинолнафталеїну в АПФ2.

DiffDock передбачив потенційні місця зв'язування для ХНТ також з найбільшим рівнем впевненості в активному сайті та загалом 17 поз із 20 в області між доменами (Мал.52)



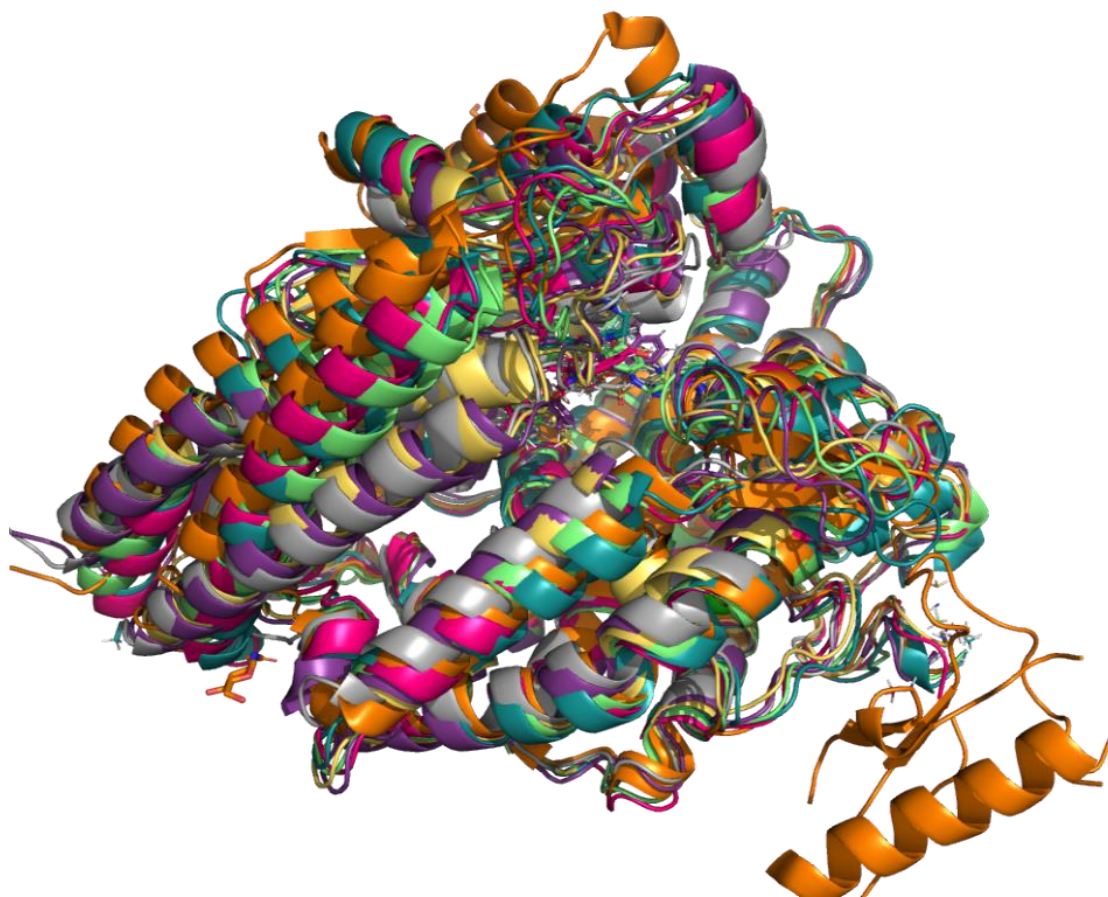
Мал.52. Передбачені пози ХНТ за допомогою DiffDock (NVIDIA).

Для подальшого дослідження вважається доцільним провести молекулярну динаміку з використанням поз передбачених DiffDock.

### *Молекулярна динаміка комплексів з активаторами*

Проведена симуляція молекулярною динамікою по протоколу поетапної релаксації: 100нс NVT 250К, 100нс NPT 250К, 100нс NPT 300К, 100нс NVT 300К, 100нс NPT 300К. Молекулярній динаміці (МД) піддавались наступні комплекси: 1R4L+XNT, 1R4L+RF, 1R4L+MNL-4760.

Фрейми траєкторії молекулярної динаміки комплексу 1R4L+XNT на 0, 100, 200, 300, 400 та 500 нс були накладені для порівняння, також для порівняння було накладені кристалічна структура 1R42 (відкрита форма) (Мал.53).



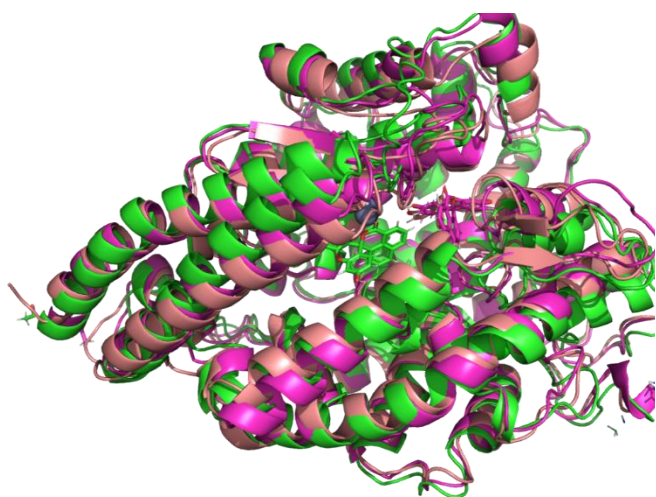
*Мал.53. Фрейми траєкторії молекулярної динаміки комплексу 1R4L+XNT на 0нс (сірий), 100нс (фіолетовий), 200нс (жовтий), 300нс (салатовий), 400нс (зелений) та 500 нс (рожевий), 1R42 – помаранчевий.*

Для комплексу 1R4L+XNT з 400 нс (структура в зеленому кольорі) ми можемо побачити явний рух в бік відкритої конформації, до 500 нс дистанція між відповідними лупами стала ще більше, але маємо зазначити, що за 500 нс (структура рожевого кольору) конформація майже "відкрилась" до форми кристалічної структури 1L42 (в помаранчевому кольорі). Важливим питанням є

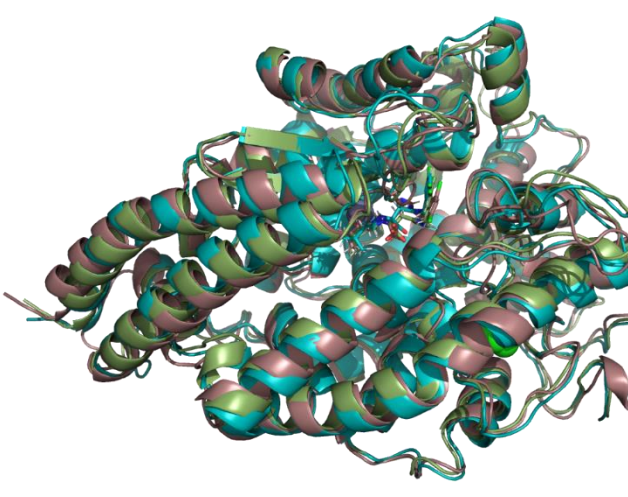
незрозумілий механізм активації окрім впливу на шарнірний рух. Досі немає пояснення механізму активації, а саме як знаходячись у активному сайті та фізично займаючи каталітичний сайт, сполука може активувати фермент. Ми також провели молекулярно-динамічну симуляцію пози збоку каталітичного сайту (це була альтернативна поза передбачена DiffDock) та отримали, набагато менший рух субдоменів один відносно одного, що показує, що це питання може бути досліджене в майбутньому, але не в рамках цієї роботи.

Для комплексу 1R4L+RF також спостерігалось розкриття структури, але не так динамічно, як з XNT (Мал.54.). Але в даному випадку ми спостерігали зміщення активатора в область «альтернативної» пози RF, тобто він «звільнив» доступ до каталітичної ділянки надаючи можливість субстрату підійти до каталітичного центру ферменту АПФ2.

Також проводилась МД симуляція комплексу 1R4L з інгібітором MNL-4760 по тому ж протоколу для порівняння поведінки у часі. Та розкриття структури білка не спостерігалось зовсім (Мал.55). Тому можемо зробити висновок, що активаційний механізм двох активаторів, скоріш за все пов'язаний зі зміною конформацій протеїну, та відповідне поліпшення входу субстрату, але на даний момент не має чіткого пояснення цього процесу.



Мал.54. Фрейми МД комплексу 1R4L+RF: на 0нс (кораловий), 300нс (рожевий), 500нс (салатовий)



Мал.55. Фрейми МД комплексу 1R4L з інгібітором MNL-4760: 0нс (кораловий), 300нс (колір трави), 500нс (лазурний)

### 5.3. Подальша оптимізація 7 хітів знайдених в першому раунді віртуального скринінгу

Хіт-ту-лід оптимізація є ключовим етапом у процесі створення лікарських засобів, спрямованим на покращення активності, селективності та фармакокінетичних властивостей вихідних сполук, що були ідентифіковані на стадії первинного віртуального скринінгу. Загальний підхід полягає у систематичному аналізі структури та активності хітів, моделюванні їх взаємодії з мішенню, а також раціональному дизайні похідних сполук для підвищення потенціалу кандидатів для подальших досліджень.

Стандартний воркфлоу, розроблений для виконання проєктів молекулярного моделювання та віртуального скринінгу у Відділі науково-розрахункових сервісів компанії «Єнамін», починається з пошуку аналогів у бібліотеці REAL Space. На момент проведення цієї частини дослідження бібліотека містить близько 65 мільярдів сполук, що є найбільшою на сьогодні колекцією комерційно доступних сполук типу «make-on-demand». Молекули в REAL Space генеруються за допомогою понад 167 перевірених на практиці протоколів паралельного органічного синтезу, що використовують більш ніж 170 202 ретельно відібраних реагенти та будівельних блоків. Дані синтетичні методики включають як стандартні, так і просунуті однореакторні («one-pot») процедури, що різняться кількістю етапів, типом очищення, умовами роботи з речовинами та, відповідно, рівнем складності отримання кінцевих продуктів.

*Пошук аналогів в Enamine In-Stock Collection та REAL Space (65+ мільярдів сполук).* Пошук аналогів у бібліотеці REAL Space здійснювався з використанням власних Python-скриптів, що були розроблені на основі бібліотеки RDKit. Для кожної сполуки попередньо згенеровано молекулярні відбитки (fingerprints, FP), що дозволяє швидко порівнювати структурну схожість хітів з величезною базою молекул. У процесі пошуку аналогів використовувалися алгоритми обчислення танімото-коефіцієнту (Tanimoto similarity), реалізовані в RDKit, для ефективної ідентифікації найбільш близьких структурних аналогів. Для прискорення обчислень та ефективної роботи з великими обсягами даних було застосовано

паралельні розрахунки на багатопроцесорних обчислювальних кластерах, що суттєво скоротило час виконання даного етапу. Відсікання проводили по значенню 0.5, отримано 1142 сполуки з In-Stock Collection та 499017 сполук з REAL Space. Після підготовки 3d конформерів цей сет далі буде використаний для декількох підходів відбору:

- двостадійний молекулярний докінг з різним ступенем точності для оптимального використання ресурсів зі збереженням якості;
- 3d shape пошук по моделі згенерованої по MLN-4670, з наступним молекулярним стикуванням та візуальною інспекцією;
- побудова QSAR моделі, натренованої на даних з ChEMBL, з подальшим застосуванням моделі для передбачення значень  $-\log IC_{50}$  (що відповідає pChEMBL value для речовин в тренувальному сеті), топові сполуки також піддавали молекулярному стикуванню з послідуочим візуальним оглядом.

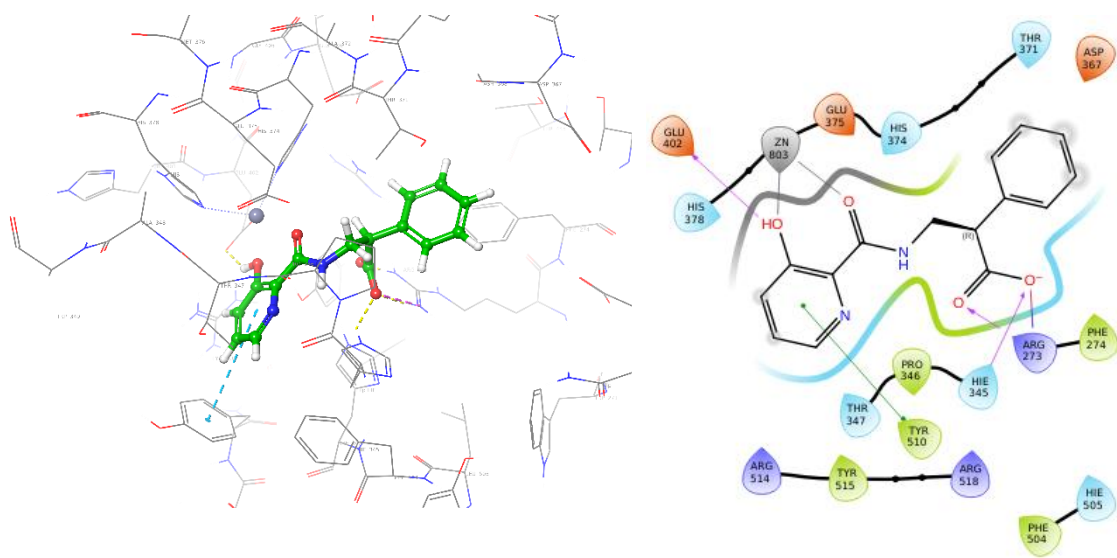
*Двостадійне молекулярне стикування знайдених аналогів.* Двостадійне молекулярне стикування (докінг) знайдених аналогів було здійснено за класичним робочим процесом, який добре зарекомендував себе у відділі науково-розрахункових сервісів компанії «Єнамін». Першим етапом став докінг високої пропускної здатності (high-throughput virtual screening, HTVS), який дозволяє швидко оцінити мільйони сполук і відібрати найперспективніші з них за докінг-скорингом. Цей етап є важливим для зменшення початкової кількості кандидатів до оптимального числа для подальшого, більш точного аналізу.

На другій стадії було проведено докінг із стандартною точністю (standard precision, SP), під час якого було детально проаналізовано топ-100 тисяч сполук, попередньо відібраних у ході HTVS-етапу. Завдяки застосуванню SP-докінгу, який характеризується точнішими параметрами та детальнішим опрацюванням конформаційних варіантів, вдалось значно підвищити надійність прогнозу щодо потенційної взаємодії молекул із мішенню.

У результаті такого комплексного двостадійного підходу було відібрано 44 перспективні сполуки для подальших досліджень, з яких 19 сполук є доступними в стоковій колекції, а ще 17—з REAL Space. Деякі з цих сполук







Мал.59. Зліва направо - 3d огляд взаємодій ліганду в кишені, 2d діаграма взаємодій  
 s\_188690\_\_\_7326850\_\_\_25747122

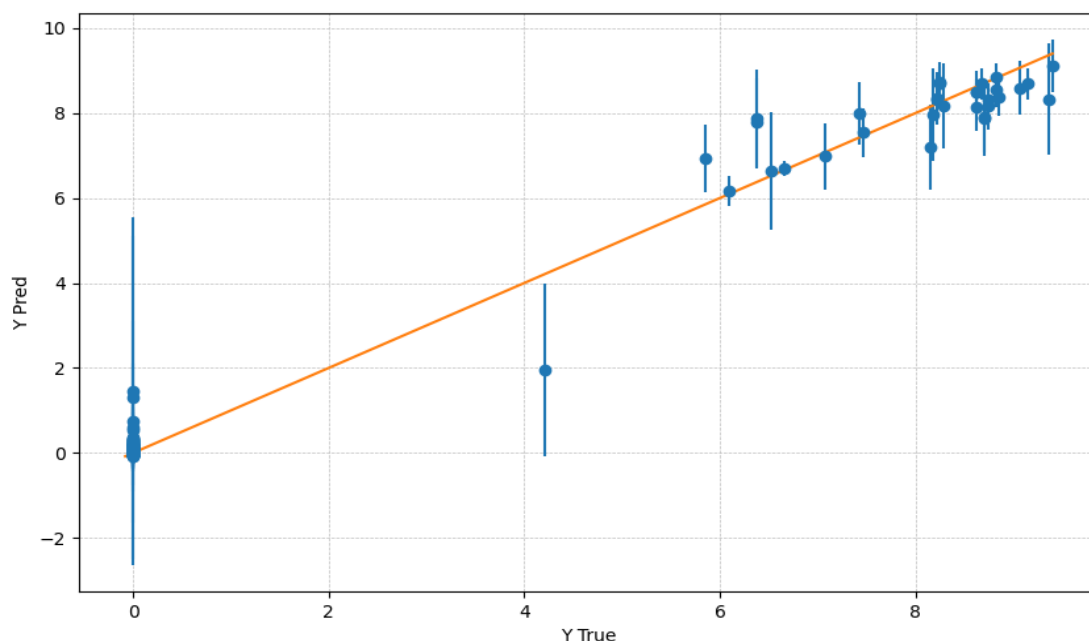
*QSAR моделювання та скринінг з наступним докінгом.* На наступному етапі дослідження було виконано QSAR-моделювання з подальшим скринінгом та молекулярним стикуванням. Для побудови моделі було використано програмний модуль DeepAutoQSAR [49] від Schrödinger, який дозволяє автоматично генерувати QSAR-моделі на основі алгоритмів глибокого навчання (deep learning). В якості навчального набору використовували активні сполуки з бази даних ChEMBL (зі значеннями активності, представленими як pChEMBL value, які були детально описані вище), а також додатково відібрані неактивні сполуки. Весь наявний набір даних по активності був розділений випадковим чином на тренувальний (75%) і тестовий (25%) набори для забезпечення незалежної перевірки точності прогнозування моделі.

Отримана регресійна модель продемонструвала високу точність прогнозування, що підтверджується графіком кореляції між фактичними та передбаченими значеннями активності для тестового набору (Мал.60.). Статистичні характеристики моделі мали високі значення: коефіцієнт детермінації ( $r^2 = 0.9887$ ) та середньоквадратична помилка ( $MSE = 0.1112$ ).

**Коефіцієнт детермінації ( $r^2$ )** характеризує ступінь відповідності між прогнозованими та експериментальними даними, де значення близькі до одиниці

свідчать про високу надійність моделі. Отримане значення  $r^2 = 0.9887$  вказує на відмінну якість моделі та її здатність до точного прогнозування активності.

**Середньоквадратична помилка (MSE)** визначає середню величину квадратів помилок передбачення моделі. Значення  $MSE = 0.1112$  є невеликим, що також підтверджує високу точність моделі.



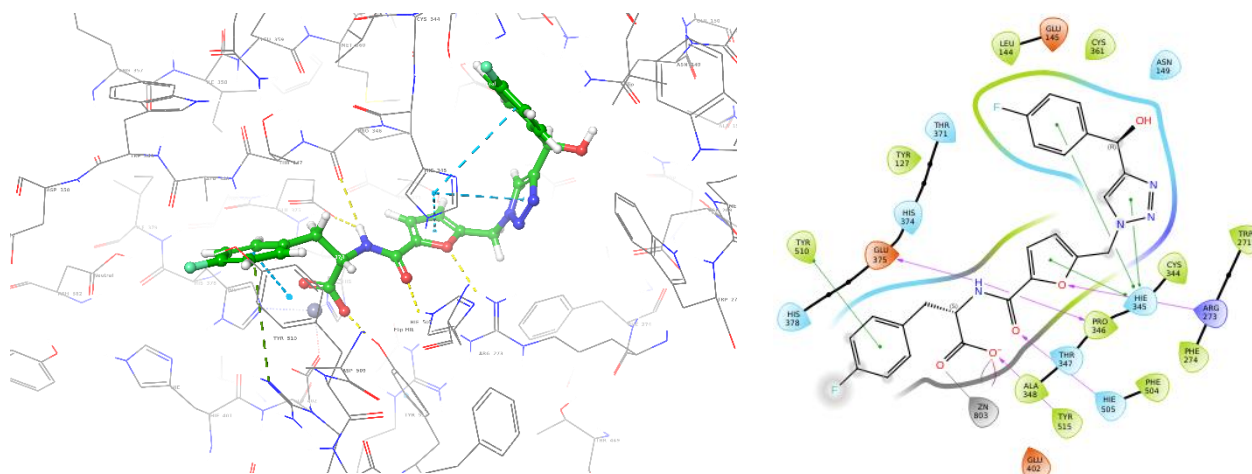
Мал.60. Візуалізація кореляції між фактичними значеннями активності (pChEMBL) та передбаченими значеннями, отриманими за допомогою QSAR-моделі DeepAutoQSAR. На графіку чітко помітна висока узгодженість між прогнозами моделі та експериментальними даними тестового набору, що підтверджує якість і достовірність створеної регресійної моделі.

Використовуючи отриману QSAR-модель, було проведено прогнозування активності для сету аналогів із бібліотеки. На основі отриманих прогнозів було відібрано тільки ті сполуки, для яких передбачене значення pChEMBL було вище ніж 2. В результаті було отримано 1668 перспективних сполук, які на наступному етапі були піддані докінгу з використанням стандартної точності (SP docking).

Після проведення SP-докінгу та детального аналізу докінг-поз і скорингових значень було остаточно відібрано **18 віртуальних хітів**, що можуть слугувати основою для створення молекулярних зондів. Приклад однієї з перспективних молекул, відібраних за результатами QSAR-скрінінгу з подальшим SP-докінгом, наведено на малюнку 61. Показана докінг-поза демонструє ефективні взаємодії з

ключовими амінокислотами активного сайту мішені, що додатково підтверджує потенціал цієї сполуки як можливого молекулярного зонду.

Запропонований підхід дозволив ефективно поєднати QSAR-аналіз із молекулярним стикуванням, що суттєво підвищило якість і надійність відбору сполук-кандидатів для подальших експериментальних досліджень.



Мал.61. Зліва направо - 3d огляд взаємодій ліганду в кишені, 2d діаграма взаємодій  
т\_280190\_\_\_19288966\_\_\_18077368\_\_\_19345860

Звісно, QSAR-моделювання, порівняно з іншими підходами, дозволяє знайти фармакофори, максимально наближені до вже відомих активних сполук. Це пояснюється тим, що QSAR-методологія базується на аналізі наявних експериментальних даних та відомих хімічних каркасів (скеффолдів). Відповідно, відібрані за допомогою QSAR сполуки часто мають високу структурну подібність до відомих активних молекул, що підвищує ймовірність їхньої експериментальної активності та робить цей метод ефективним для оптимізації та розширення існуючих фармакофорних груп.

Таким чином, завдяки комплексному використанню трьох описаних вище методів (двостадійний докінг із попереднім пошуком аналогів, фармакофорний 3D shape-скрінг із докінгом SP та QSAR-моделювання з подальшим SP-докінгом) було відібрано загалом **70 нових потенційних кандидатів**, що є перспективними для подальшого етапу біологічної валідації. Комбінація різних підходів дозволила охопити ширший хімічний простір, забезпечуючи як

структурну новизну отриманих сполук, так і достатню впевненість у їхньому біологічному потенціалі для використання як молекулярних зондів.

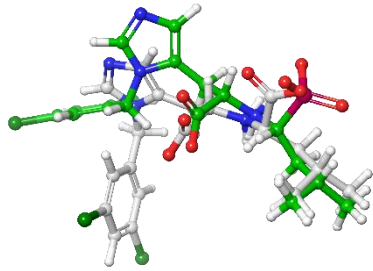
#### 5.4. Додатковий скринінг Enamine In-Stock колекції та REAL Space

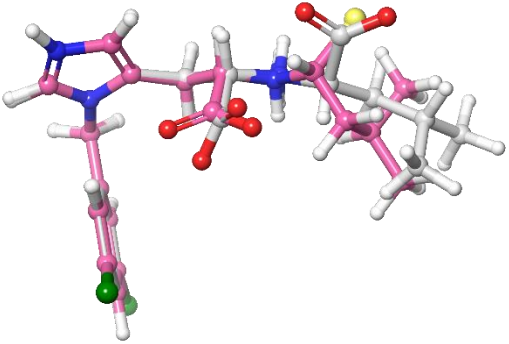
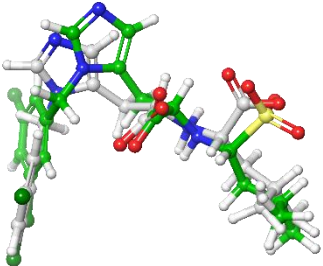
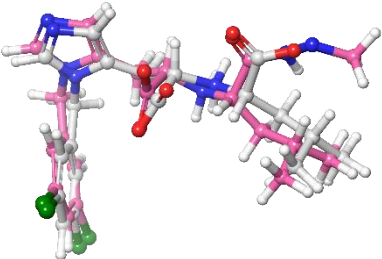
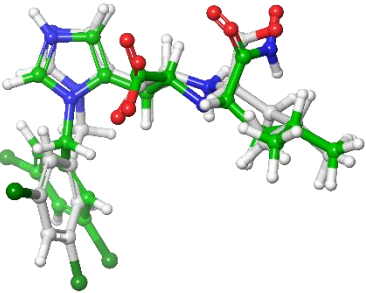
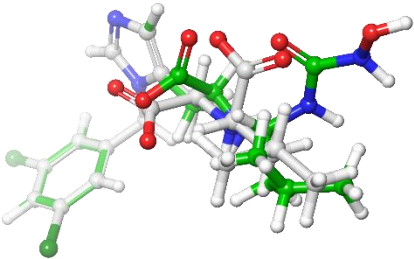
Наступним кроком є зміна стратегії пошуку, щоб знайти нові хемотипи та збільшити шанси отримати наномолярні інгібітори.

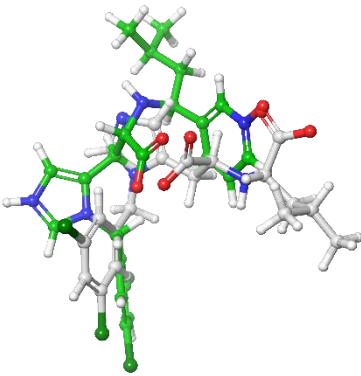
АПФ2 є цинк-залежною монокарбоксіпептидазою, іон цинку знаходиться в активному сайті та відіграє важливу каталітичну роль, крім того, всі відомі інгібітори, це карбонові кислоти, які є найчастішими представниками Zn-байндерів.

Для хемоінформатичного пошуку функціональних замінників карбоксильної групи в ролі цинкового байндера ми провели віртуальну заміну карбоксильної групи, що зв'язується з  $Zn^{2+}$  на різноманітні групи. Потім було розраховано подібність по USRCAT відбиткам там само як було запропоновано в [66] (300\*300 конформерів) для створення таргетних бібліотек сполук, де основною метою було знайти сполуки з подібною 3D-формою до відомих активних лігандів. Тут ми застосували цей метод для пошуку груп, подібних за формою та розташуванням донорів/акцепторів, водночас забезпечуючи хімічну різноманітність і виявлення нових потенційно активних замісників. Групи, які мали найвищі значення USRCAT подібності визначені в таблиці 11.

Табл.11. Значення USRCAT подібності різних подібних структур з різними функціональними групами.

Хімічна група	USRCAT similarity	3d накладені структури
Фосфонати	0.7794	

Тіолати	0.7189	
Сульфонати	0.704	
Гідразиди	0.7028	
Гідроксамові кислоти	0.6478	
Гідроксисечовина	0.6142	

N-оксиди	0.4141	
----------	--------	--

Попередня стратегія була побудована на карбоксильних кислотах здебільшого. Тому було вирішено також розширити це обмеження, та знайти інших представників груп сполук, які є потенційними цинковими байдерами. Знайдені за USRCAT подібністю структури далі будуть використані в наступному пошуку, так само як і додаткові, знайдені з літератури.

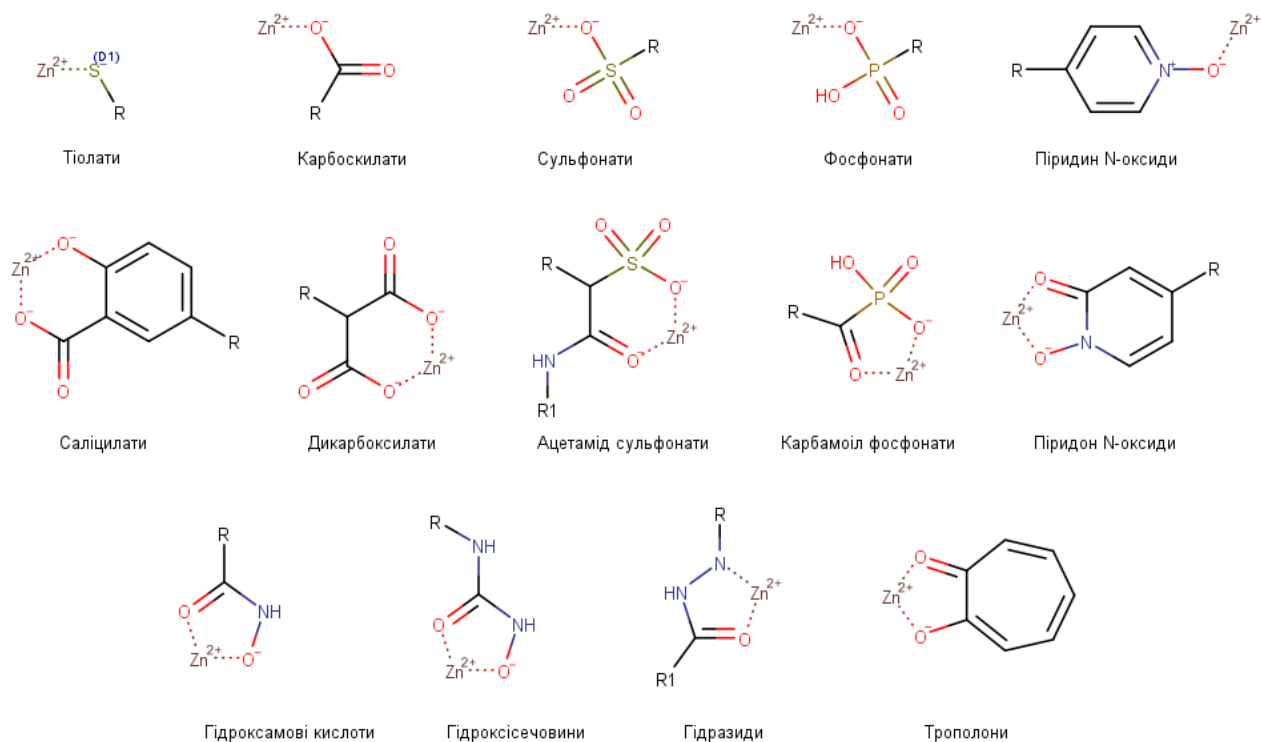
#### ***Огляд Zn-зв'язуючих груп (ZBG) для інгібування металопротеїназ***

Одним із ключових підходів до інгібування металопротеїназ є використання Zn-зв'язуючих груп (ZBG). У статті [54], розглядаються новітні підходи до дизайну селективних інгібіторів металопротеїназ, що включають:

- Гідроксамати – найпоширеніша Zn-зв'язуюча група, яка ефективно координує  $Zn^{2+}$  у каталітичному центрі металопротеїназ;
- Карбоксилати, тіоли, фосфонати – альтернативні групи, що демонструють потенціал у підвищенні селективності та біодоступності інгібіторів;
- Нітроген-вмісні ZBG – такі як піримідин-2,4,6-триони, які забезпечують стабільну координацію з цинком і дозволяють створювати інгібітори з покращеними фармакокінетичними характеристиками;
- Похідні бензамідину та сульфонамідів – перспективні альтернативи, які можуть підвищити селективність та знизити токсичність;
- Координаційні Zn-зв'язуючі групи – такі як N-оксиди, що діють як електродонори у взаємодії з цинком, а також азол-вмісні структури, які можуть сприяти стабільному зв'язуванню та пролонгованій дії інгібітора.

Було вирішено сконцентруватись на відборі наступних груп: тіолати, сульфонати, фосфонати, піридин N-оксиди, саліцилати, дікарбоксилати,

ацетаміди сульфонати, карбамоіл фосфонати, гідроксамові кислоти, гідроксісечовини, гідразини, трополони (Мал.62).



Мал.62. Zn-байндери.

### Пошук відповідних груп в Enamine In-Stock колекції та REAL Space

В пергу чергу ми мали прописати SMARTS [23] відповідних груп, та врахувати всі тонкощі для вірного пошуку. В результаті були написані наступні SMARTS для пошуку:

- [#16D1]-C – тіолати\*;
- [OD1]C(=O)[#6] – карбоксилати;
- [#6]#16(=O)=O – сульфонати;
- [#8D1]P([#8D1])([#6])=[OX1] – фосфонати;
- [#8H0]-[n+]1cccc1 – піридин-N-оксиди;
- [OD1]-#6-c1cccc1-[OD1] – саліцилати;
- [#8;H1]-#6-[#6]-#6=O – дікарбоксилати;
- [OD1]S(=O)(=O)[C]C(=O)-[#7] – ацетаміди сульфонати;
- [OD1]P([OD1])(=O)#6=O – карбамоіл фосфонати;
- [#8D1]-[N]1-#6-[#6]=[#6]-[#6]=[#6]-1 – піридон-N-оксиди;

- [#8D1]-[#7H1]-#6=O – гідроксамові кислоти;
- [#8D1]-[#7H1]-#6-[#7]-[#6] – гідроксисечовини;
- [#6][NH][NH]C([#6])=O – гідразиди;
- [#8D1]C1=CC=CC=O – трополони.

\*У SMARTS-нотації символ D використовується для позначення ступеня атома (degree), тобто кількості зв'язків, які має атом (включаючи одинарні, подвійні, потрійні та ароматичні зв'язки).

Приклад: [#16D1]-C: #16 – атом сірки (S); D1 – означає, що цей атом сірки має лише один зв'язок; -C – атом вуглецю, з яким ця сірка зв'язана.

Таким чином, цей SMARTS-фрагмент описує сірку, яка має один зв'язок, зв'язану з атомом вуглецю. Це може відповідати тіолам (-SH) або іншим подібним сполукам.

Після проведення підструктурного пошуку знайшли наступну кількість сполук, що зазначена в таблиці 12.

Табл.12. Кількість сполук після проведення підструктурного пошуку.

Назва групи	Кількість сполук, знайдених в Ін-сток Колекції	Кількість сполук, знайдених в REAL SPACE)
Ацетамід сульфонати	2	0
Карбамоїл сульфонати	0	0
Дикарбоксилати	59	246 тис
Гідразиди	33.3 тис	41.4 мільонів
Гідроксамові кислоти	207	1.8 мільонів
Гідроксисечовини	4	0
Фосфонати	240	11.4тис
Піридин N-оксиди	3 тис	8.6 мільонів
Піридон N-оксиди	0	0
Саліцилати	864	282 тис
Сульфонати	322	45.5 тис
Тіолати	534	276 тис
Трополони	0	0

Пошук був проведений власним скриптом, який написаний групою CADD компанії Енамін. Групи, які налічили мільони сполук, були диверсифіковані до сотен тисяч з використанням скрипта який використовує MinMaxPicker підхід [8].

Сполуки були підготовлені із застосуванням LigPrep [98] від Schrodinger. LigPrep – це програмний модуль компанії Schrödinger, призначений для підготовки молекул перед молекулярним докінгом або динамічним моделюванням. Цей метод дозволяє генерувати біологічно релевантні конформації лігандів, оптимізуючи їхню геометрію та зарядові стани.

У цьому дослідженні ми використовували LigPrep для підготовки лігандів перед молекулярним докінгом. Зокрема, було застосовано автоматичну генерацію протонних станів при фізіологічному рН (7.4), враховано можливі протомерні та стереоізомерні форми молекул. Для оптимізації структур використовували силові поля OPLS. Також було активовано функцію Metal States для коректної обробки металоорганічних комплексів, що дозволило врахувати специфічні координаційні взаємодії між лігандами та іонами металів у активному центрі ферменту.

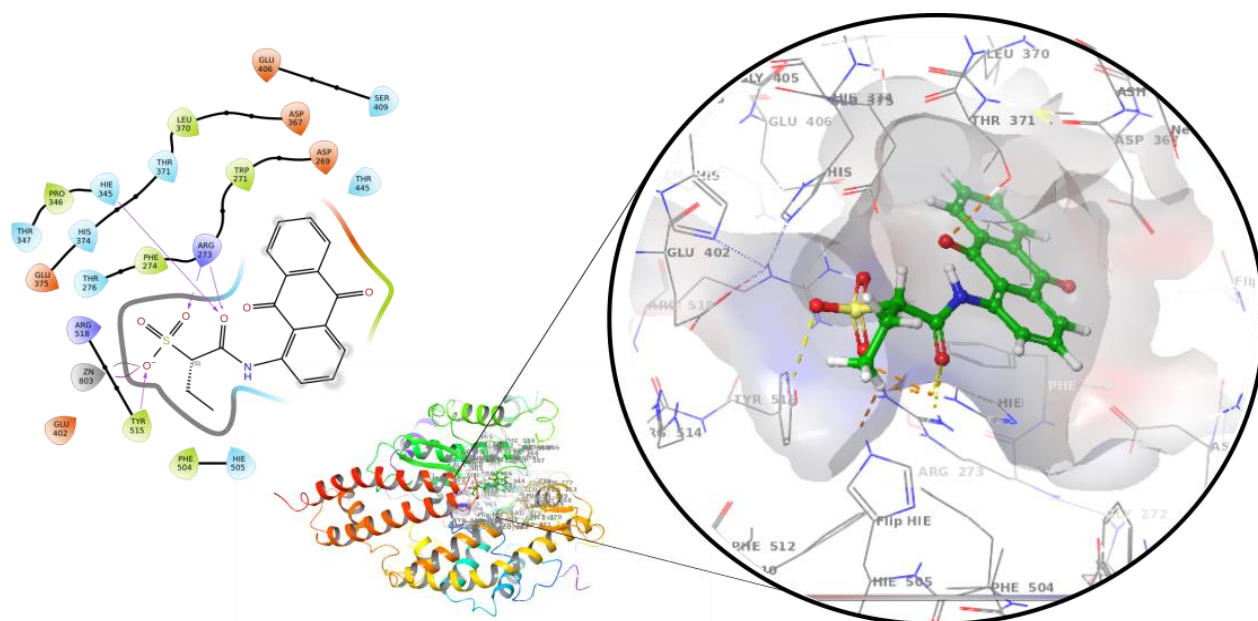
Докінгова модель (грід) була побудована з урахуванням висновків зроблених в нових дослідженнях та було прийнято рішення спростити обмеження, та залишити єдиний констрейнс, а саме координацію з металом цинку. Взаємодії з важливими амінокислотними залишками, які були виявлені під час дослідження FTMove, FTMap та позою природнього субстрату, будуть враховані як фактори, які підвищують “привабливість” лігандів під час візуального огляду поз стикування.

### ***Результати молекулярного докінгу, візуального огляду поз та «чері пікінгу»***

#### ***1. Ацетамід сульфонати***

Не дивлячись на те, що було знайдено тільки дві сполуки, одна із них показала досить гарну позу (Z56777340) (Мал.63.). На діаграмі взаємодій бачимо, що такий фрагмент чудово підходить в даному випадку як цинковий байндер, тому що сульфонова група займає таку геометричну позицію, що дозволяє як

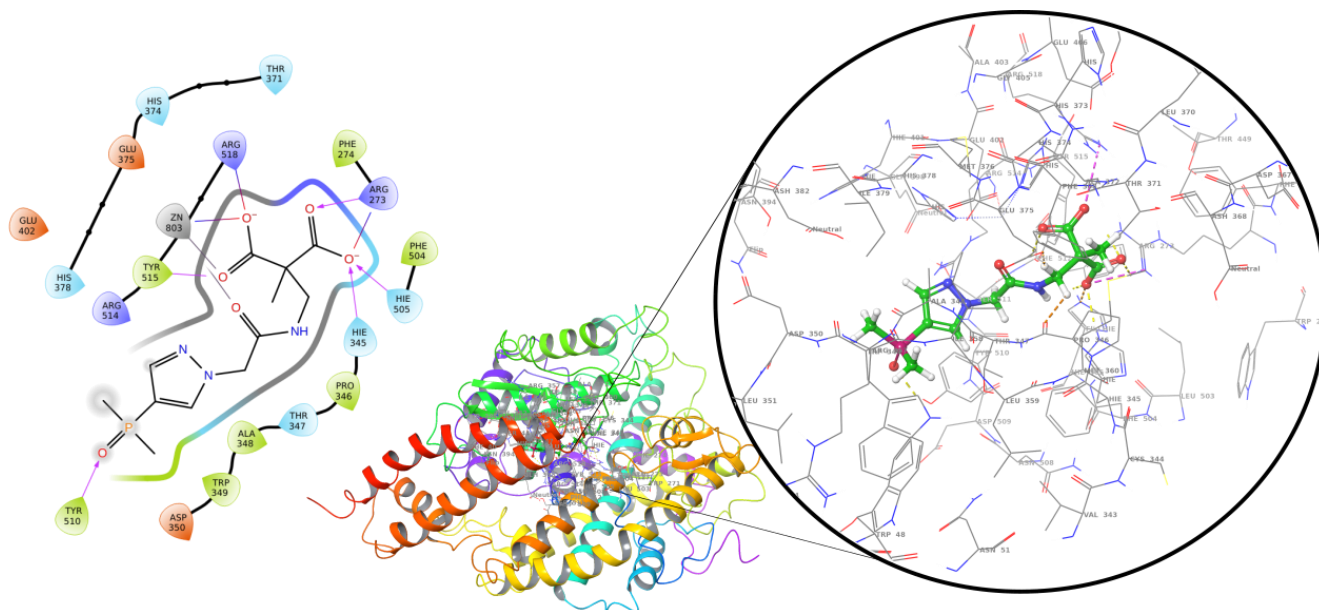
координувати іон цинку, так і одночасно створювати водневий зв'язок з ARG273 та TYR515. Кисень ацетамідної групи також чудово доповнює афінність водневим зв'язком з ARG273 та HIE345.



Мал.63. Зліва направо - 2d діаграма взаємодій Z56777340, загальний вигляд структури протеїну АПФ2 із стикованою позою ліганду, 3d огляд взаємодій ліганду в кишені.

## 2. Дікарбоксилати

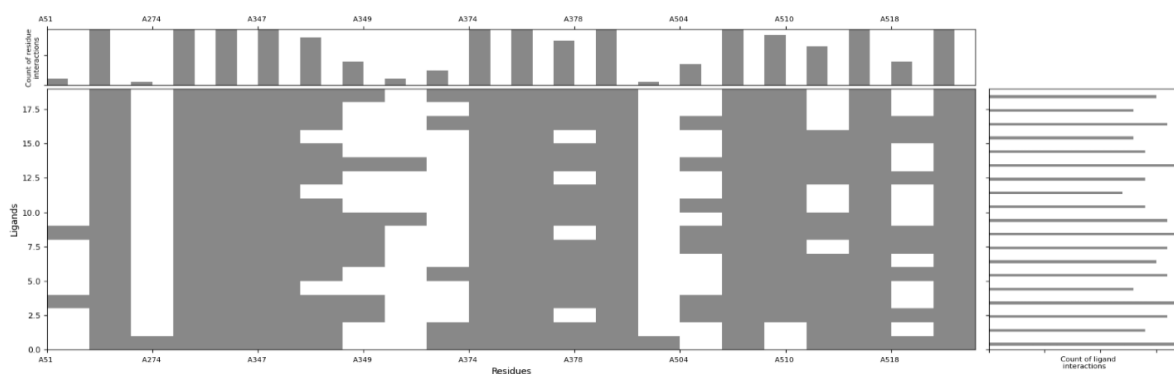
Загалом було відібрано 19 сполук.



Мал.64. Зліва направо - 2d діаграма взаємодій t\_188690\_18313080\_21804122, загальний вигляд структури протеїну АПФ2 із стикованою позою ліганду, 3d огляд взаємодій ліганду в кишені.

Для огляду та демонстрації надаємо позу одного представника цієї групи `m_188690__18313080__21804122` (Мал.64.). Цікавий факт, що всі топові молекули зайняли інший субпocket в порівнянні з MLN-4760. Геометрія дікарбоксилатного фрагменту дозволяє створити координацію іону цинку, а також водневі зв'язки та сольові місточки з ARG273, ARG518, HIE505, HIE345, TYR515. Також значимо, те що займаючи субпocket ліворуч від іону цинку, молекула може створити рі-рі стекінг та/або водневий зв'язок з TYR510, який є одним із двох важливих відмінностей між послідовностями АПФ та АПФ2 (VAL518 не зможе створити водневий зв'язок з атомом кисню), що може мати ключовий вплив на селективність сполук.

Для всіх 19 сполук побудована матриця взаємодій (Мал.65.), де можна побачити, що кожна сполука має більше 10 взаємодій з амінокислотними залишками таргету.

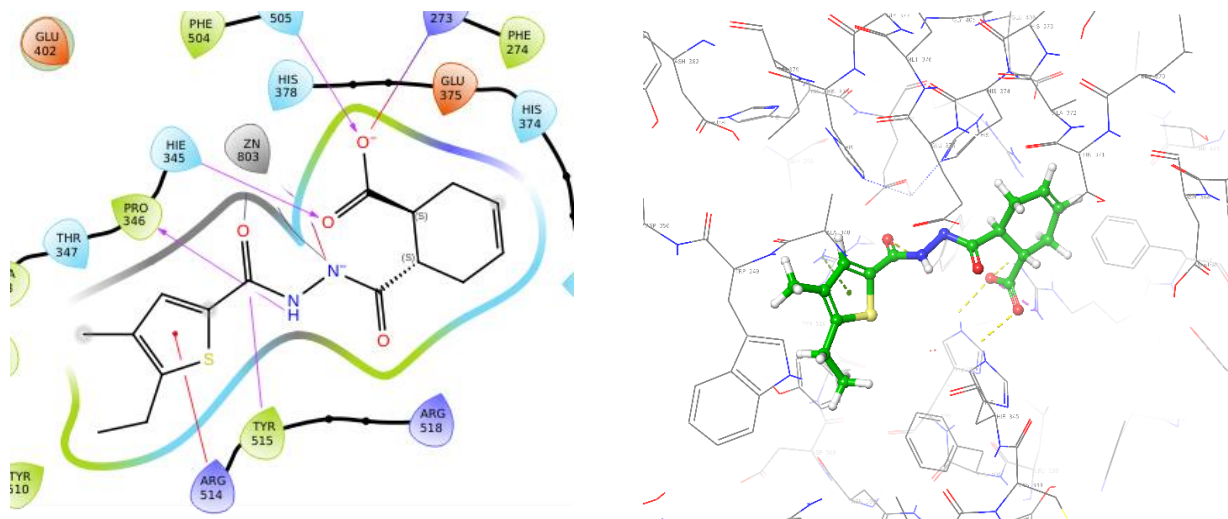


Мал.65. Матриця взаємодій відібраних дікарбоксилатів.

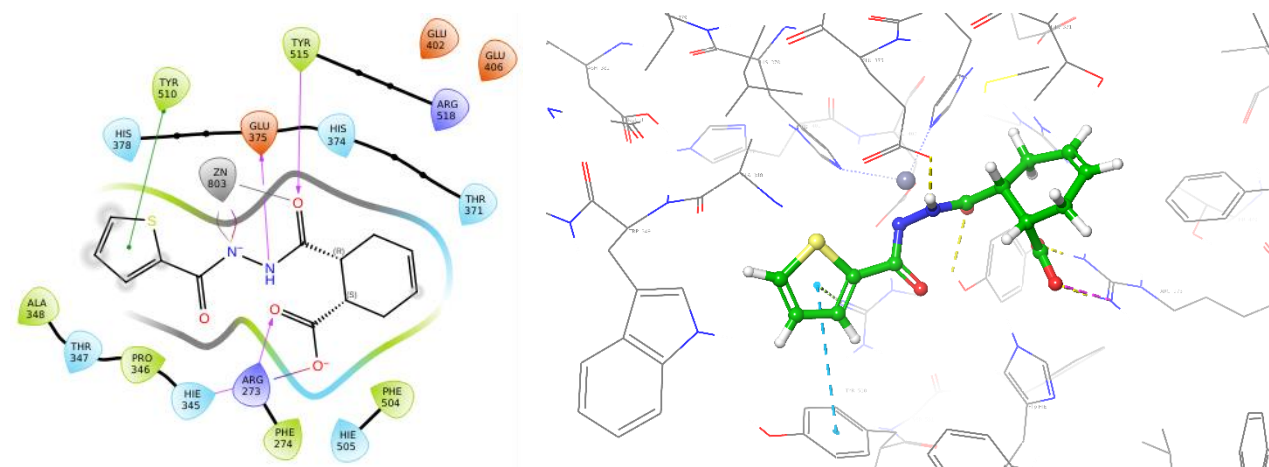
### 3. Гідразиди

Гідразиди координують іон цинку ( $Zn^{2+}$ ) через свої функціональні групи, утворюючи стабільні хелатні комплекси. Зокрема, карбонільний кисень ( $C=O$ ) та азот аміногрупи ( $-NH-$ ) гідразидної частини взаємодіють з цинком, формуючи бідентатний зв'язок. Це призводить до утворення п'яти- або шестичленного хелатного циклу, який стабілізує комплекс. Така координація може інгібувати активність цинк-залежних ферментів, блокуючи їх каталітичні центри. Зокрема, гідразиди відомі як інгібітори металопротеїназ, оскільки їх зв'язування з цинком порушує нормальну функцію цих ферментів.

Всього відібрано 45 сполук. Докінг скор від -9.5 до -11.7. Під час кластеризації виявлено, що дану кількість можна сгрупувати в 6 кластерів, та відібрати по декілька з кластера, в залежності від обсягів доступної кількості для біологічної валідації.



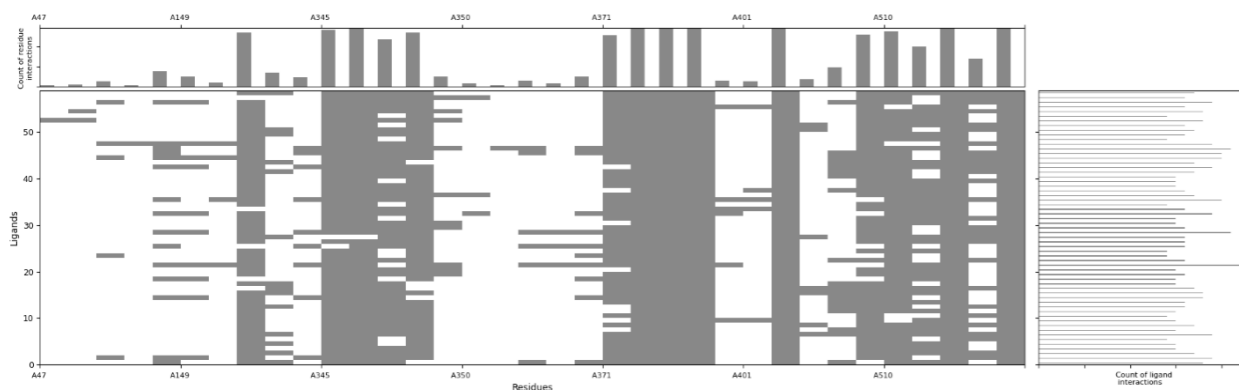
Мал.66. Зліва направо - 2d діаграма взаємодій Z169903020, 3d огляд взаємодій ліганду в кишені.



Мал.67. Зліва направо - 2d діаграма взаємодій Z169884988, 3d огляд взаємодій ліганду в кишені.

Аналізуючи пози (Мал.66,67) важливо зазначити, що ця група добре координує цинк утворюючи п'ятичленний хелатний цикл. Розташовуючись посередині молекули (а не термінально, як, наприклад, дікарбосилати), похідні гідразидів можуть заповнювати як правий так і лівий субпocket. З наявністю поряд додаткової карбоксильної групи, вона може утворювати водневі зв'язки, а також і сольовий мостик з ARG273. Ароматичні замісники створюють гарний стекінг з TYR510, що важливо для селективності проти АПФ.

Також показана матриця взаємодій (Мал.68).

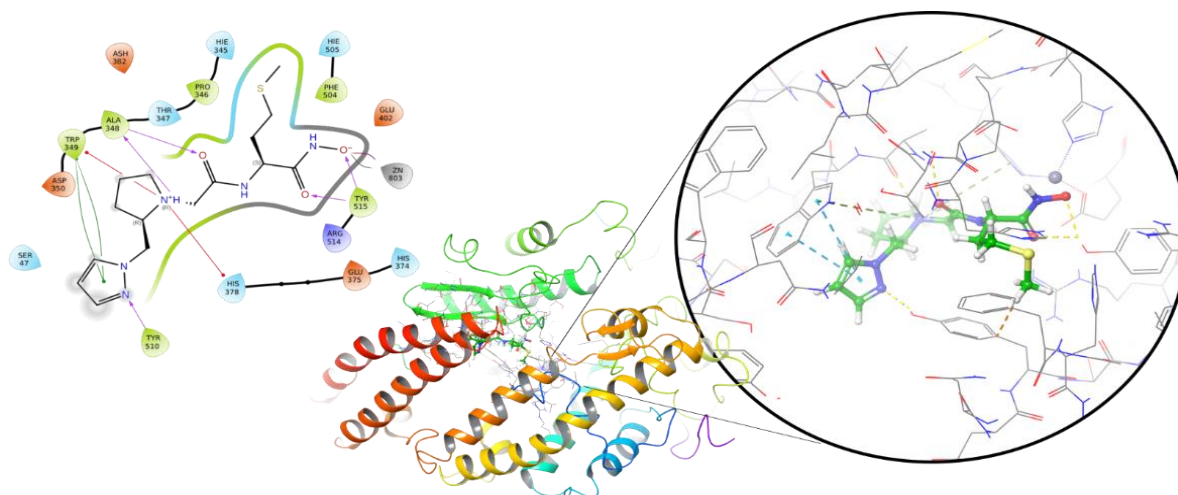


Мал.68. Матриця взаємодій відібраних гідрозидів.

#### 4. Гідроксамові кислоти

З обох - і стокової і віртуальної бази Енаміну вибрано 14 сполук похідних гідроксамової кислоти (Мал.69).

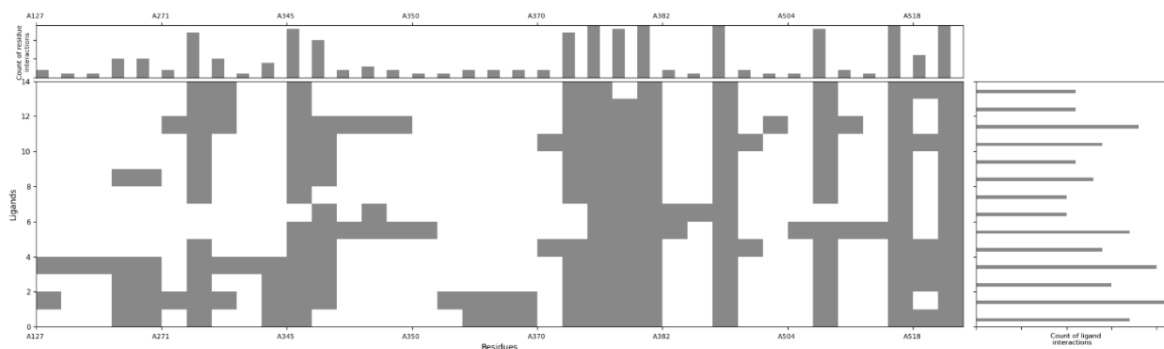
Ворхед гідроксамової кислоти чудово координує іон цинку. Здебільшого ці похідні займають лівий субпocket, тому часто не взаємодіють з ARG273, але мають додаткові взаємодії з амінокислотними залишками ліворуч, які під час минулих пошуків рідко були задіяні, бачимо стекінг з TRP349, який теж визначений як важливий методами FTMap та FTMove.



Мал.69. Зліва направо - 2d діаграма взаємодій Z8716415636, загальний вигляд структури протеїну АПФ2 із стикованою позою ліганду, 3d огляд взаємодій ліганду в кишені.

Матриця взаємодій (Мал.70) показує меншу кількість взаємодій, але їх достатньо для того, щоб спробувати цю групу у біологічній валідації. Окрім того,

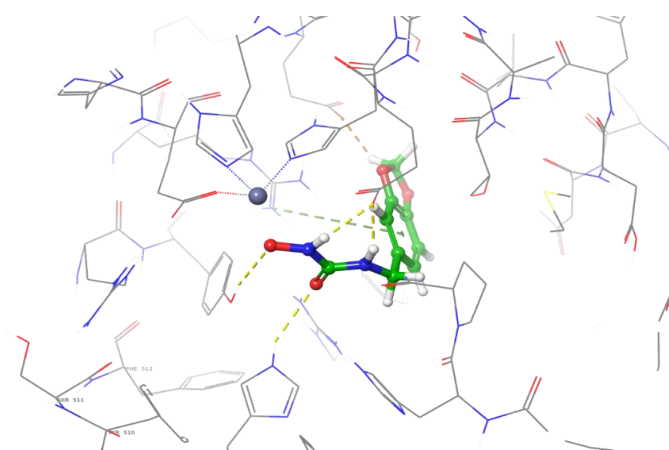
гідроксамові кислоти досить розповсюджені представники Zn-байндерів, і є відомі запатентовані байндери інших цинкових протеаз.



Мал.70. Матриця взаємодій відібраних гідроксамових кислот.

Стосовно кластеризації, то сполуки виявились досить різноманітними, тільки дві сполуки мають досить високу схожість та можуть бути поміщені в один кластер.

5. Представників класу гідроксисечовини у каталогах Єнамін всього 4, та вони мають досить маленький розмір, тому жодна з молекул не може бути потенційним байндером серед знайдених. Хоча потрібно зазначити, що як група, вона доволі непогано може взаємодіяти з субпocketом біля Zn (Мал.71), тому на майбутнє можна розглянути такі кастомні модифікації.

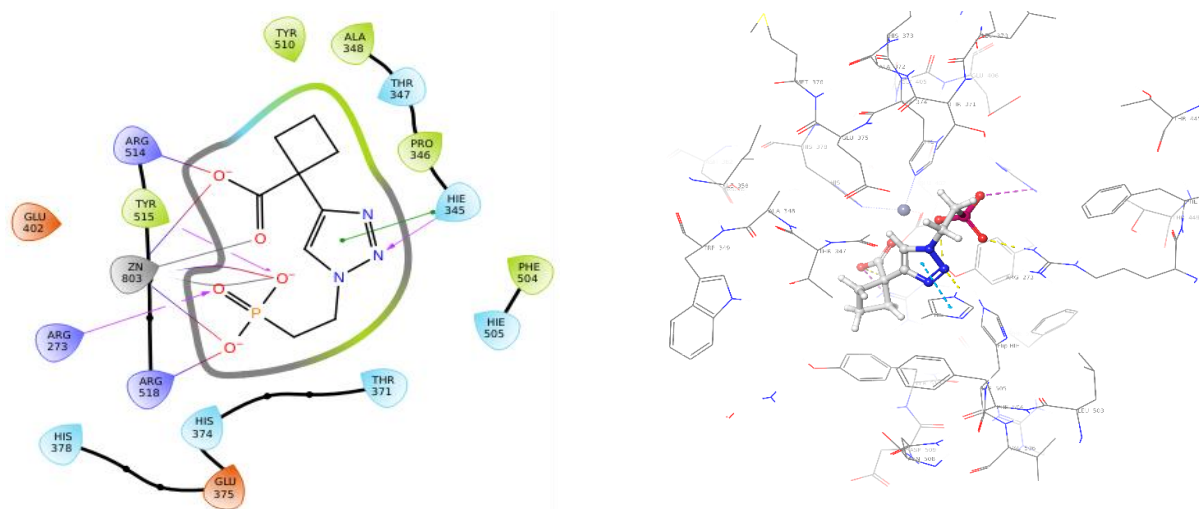


Мал.71. Взаємодія гідроксисечовини в кишени.

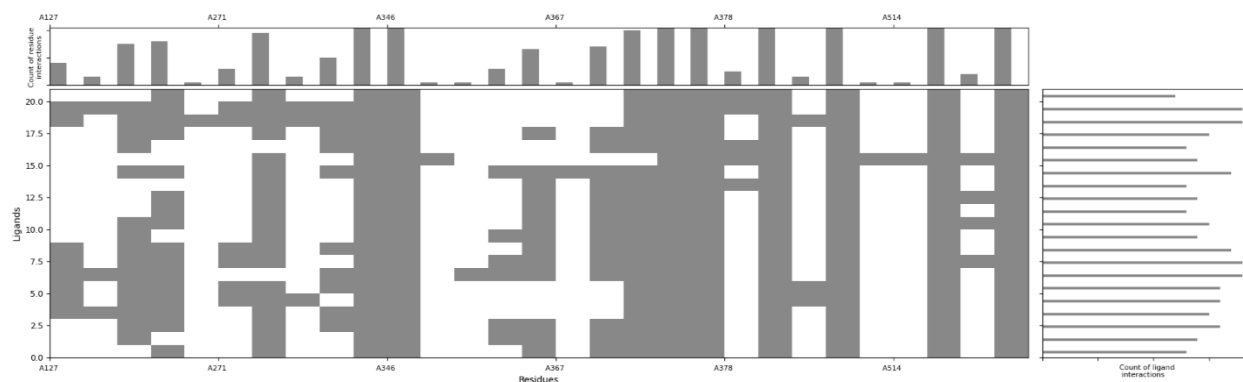
## 6. Фосфонати

Після молекулярного стикування, аналізу та візуального огляду ми відібрали 21 сполуку.

Фосфонатна група чудово взаємодіє одночасно як з іоном цинку, так і з ARG273, крім того в той же час може утворювати водневі зв'язки з HIS505 та HIS345, що є дуже важливим (Мал.72,73).



Мал.72. Зліва направо - 2d діаграма взаємодій  $m_{273910}$   $m_{21676896}$   $m_{21505024}$ , 3d огляд взаємодій ліганду в кишені.



Мал.73. Матриця взаємодій відібраних фосфонатів.

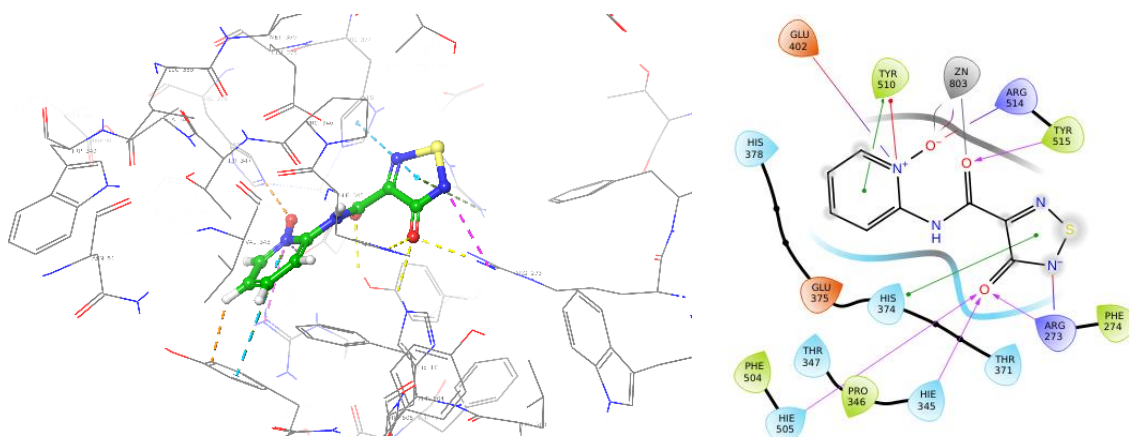
Ліганди цієї групи займають субпocket відомого інгібітора MLN-4760, але сполуки мають можливість "дотягнутись" ще і до LYS363, що теж важливо, та додає різноманіття в байндінг моуд відібраних нами сполук.

## 7. Піридин-N-Оксиди

Після молекулярного стикування, аналізу та візуального огляду ми відібрали 35 сполук.

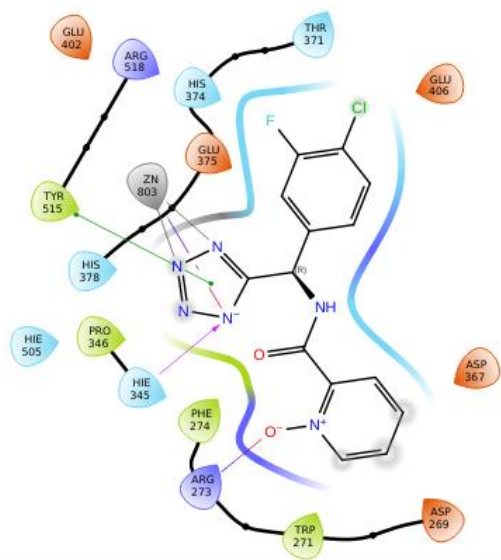
Піридин-N-оксидна група може займати дуже вигідну позицію, коли кисень одночасно координуючі  $Zn^{++}$ , утворює сьоловий мостик з ARG514 та водневий зв'язок з TYR515, в свою ж чергу N(-) утворіє йонний зв'язок з GLU375

та GLU402, також можемо спостерігати *pi-pi* стекінг з TYR510, що є дуже важливим для селективності (Мал.74).

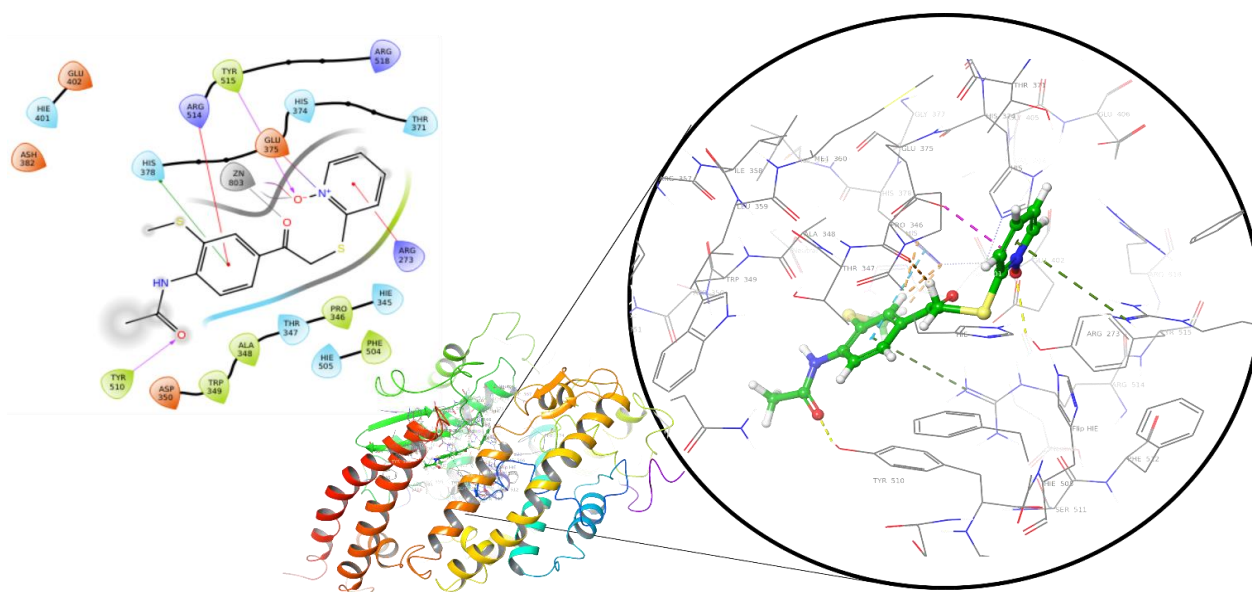


Мал.74. Зліва направо - 3d діаграма взаємодій s\_22\_20813372\_11730510, 2d огляд взаємодій ліганду в кишені.

Хочеться зазначити, що тут був виявлений гарний приклад (Мал.75), де координатором Zn виступив тетразольний замісник, який хоча і не є класичним Zn-байндером, але маючи риси біоізостера карбоксильної групи, на наш погляд може чудово координувати Zn, в той же час залишаючи N-оксидну групу для взаємодії з ARG273.

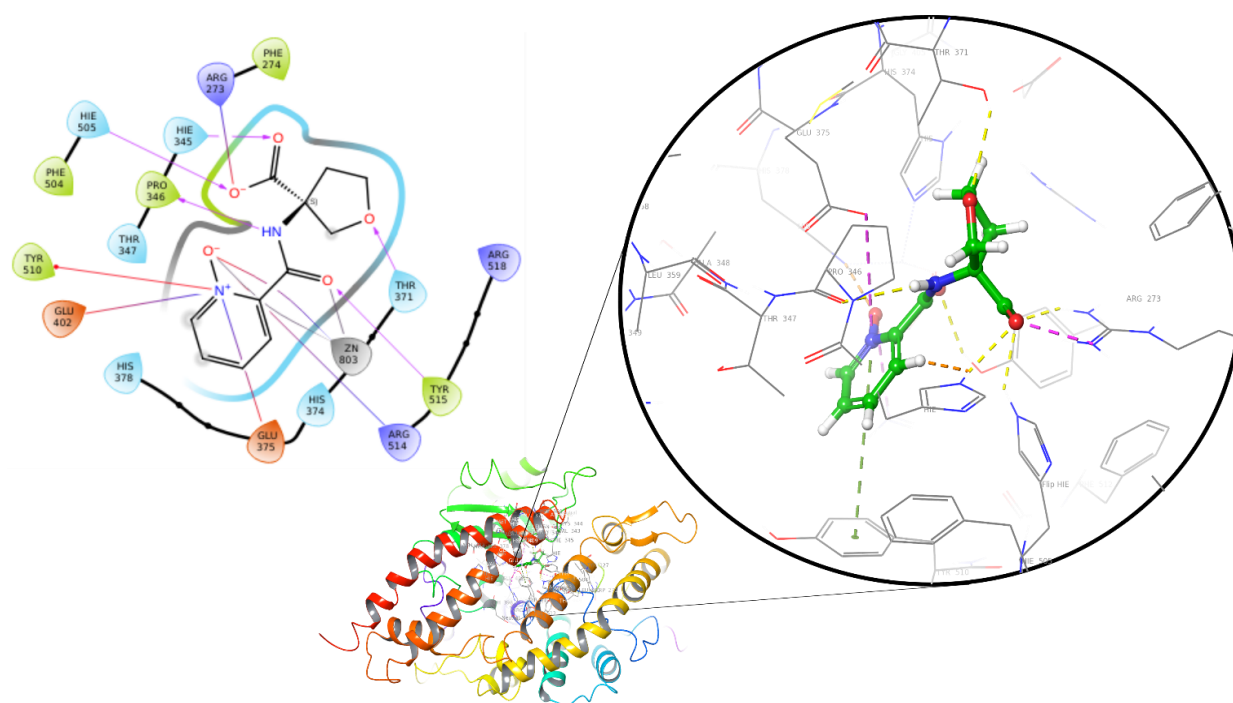


Мал.75. 2d огляд взаємодій ліганду Z1713510693 в кишені.



Мал.76. Зліва направо - 2d діаграма взаємодій Z448102166 , загальний вигляд структури протеїну АПФ2 із стикованою позою ліганду, 3d огляд взаємодій ліганду в кишені.

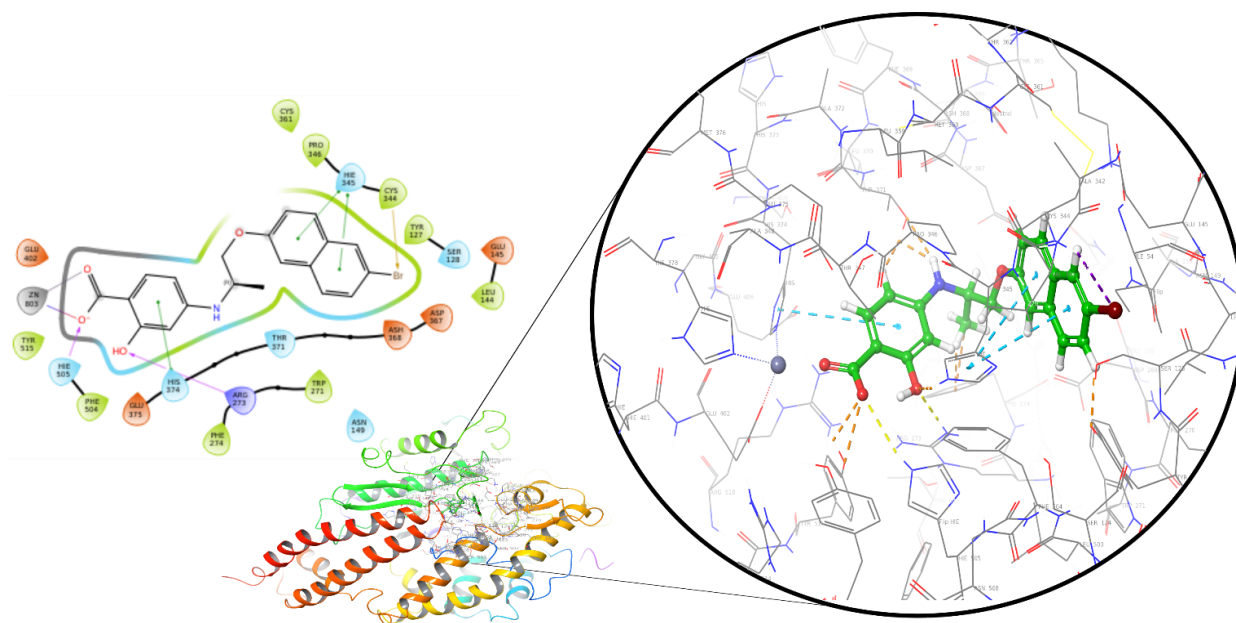
Також цікавим фактом є те, що ця група може окупувати як класичний субпocket відомого інгібітора MLN-4760 (дотягуючись до LYS363) так і субпocket ліворуч цинку, утворюючи важливі зв'язки з THR349 (Мал.76,77).



Мал.77. Зліва направо - 2d діаграма взаємодій s\_269946\_11987984\_7302472 , загальний вигляд структури протеїну АПФ2 із стикованою позою ліганду, 3d огляд взаємодій ліганду в кишені.

## 8. Саліцилати

Функціональна група саліцилової кислоти доволі непогано проявив себе у зв'язування металу цинку в комбінації з водневими зв'язками з одні із двор залишків аргініну (в залежності від геометрії) - ARG274 та ARG518. Похідні саліцилової кислоти зайняли стандартний субпocket потягуючись до LYS363 (Мал.78).

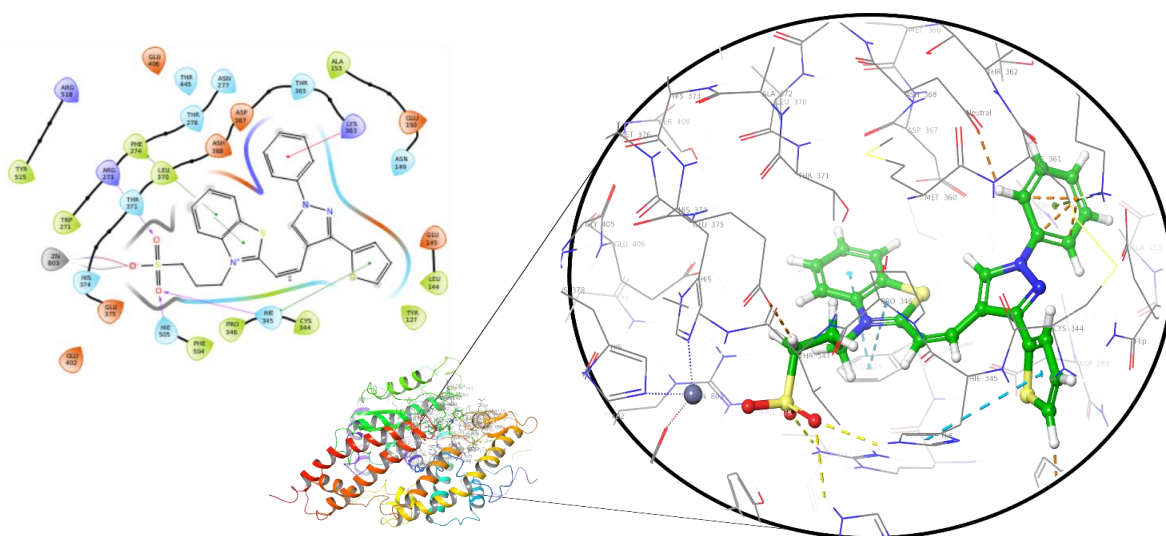


Мал.78. Зліва направо - 2d діаграма взаємодій s\_274370\_\_\_25905060\_\_\_13089394, загальний вигляд структури протеїну АПФ2 із стикованою позою ліганду, 3d огляд взаємодій ліганду в кишені.

Всього зі стоку та з REAL Space відібрано 12 сполук.

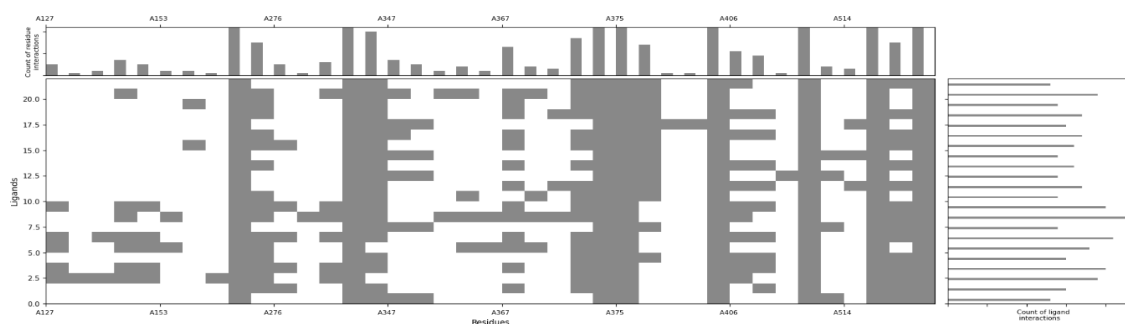
## 9. Сульфонати

Сульфонатна група також проявляє себе як гарний байндер для цинку, в той же час одночасно три атоми кисню сульфонатної групи можуть створювати водневі зв'язки з HIS345, HIS505 та ARG273. Термінальне розташування групи дає можливість молекулі ще мати такі замісники, щоб утворювати стекінги з PHE274 та LYS363. Також ацетамід сульфонати мають шанси проявити активність, комплементуючі взаємодію додатковим киснем та NH (Мал.79).



Мал.79. Зліва направо - 2d діаграма взаємодій Z56766107, загальний вигляд структури протеїну АПФ2 із стикованою позою ліганду, 3d огляд взаємодій ліганду в кишені.

Після кластеріації вібранили всього 19 сполук із цієї групи, що мають шанси бути ангіботрами АПФ2. Матриця взаємодій демонструє зв'язки відібраних сполук (Мал.80).



Мал.80. Матриця взаємодій відібраних сульфонатів.

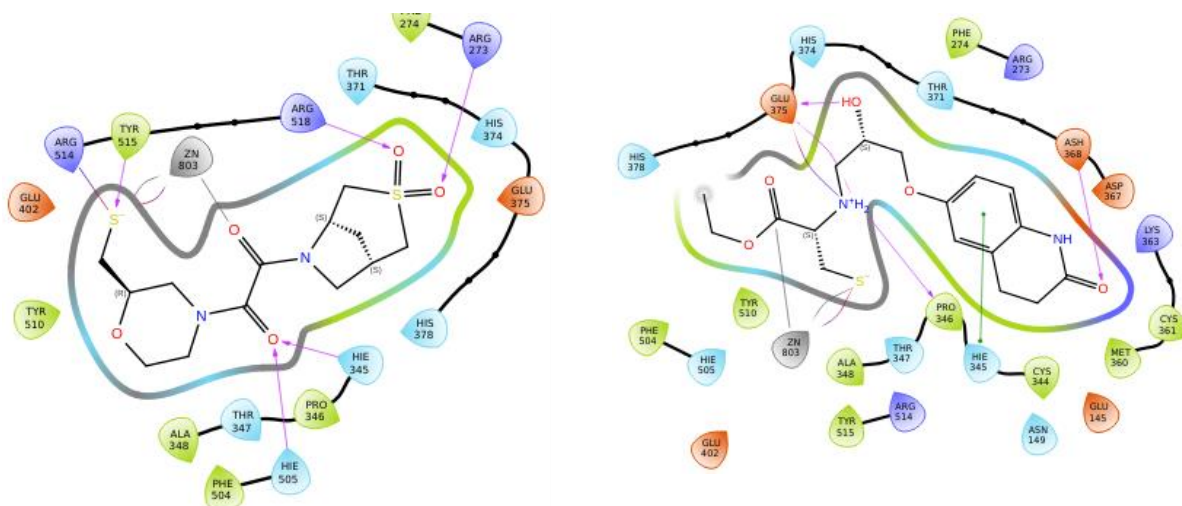
## 10. Тіолати

Стосовно тіолатів як інгібіторів ACE2, можна провести аналогію з каптоприлом та іншими відомими тіолатними лікарськими засобами. Так, каптоприл, класичний інгібітор АПФ, містить тіольну групу, яка ефективно координується з іоном цинку в активному центрі ферменту, порушуючи його каталітичну активність. Аналогічно, інші лікарські препарати на основі тіолатів (наприклад, зофеноприл або похідні пеніциламіну) використовують тіольну функціональну групу для утворення міцних, оборотних координаційних зв'язків із центром цинку. Цей механізм підкреслює перспективність застосування

тіолатних груп для створення нових інгібіторів АСЕ2, використовуючи аналогічні взаємодії із цинк-зв'язуючими центрами.

Тіолатна група ( $-S^-$ ) є саме дисоційованою (депротонованою) формою тіольної групи ( $-SH$ ). У фізіологічних умовах ( $pH \sim 7,4$ ) багато тіолів ( $-SH$ ) частково дисоціюють, утворюючи тіолат-аніон ( $-S^-$ ). Саме тіолатна форма групи проявляє високу нуклеофільність і здатна ефективно взаємодіяти (координуватись) із каталітичними центрами ферментів, що містять цинк, зокрема АПФ (АСЕ) або АСЕ2. Таким чином, коли йдеться про взаємодію тіолів із цинк-вмісними ферментами (як у випадку каптоприлу), найчастіше активною формою є саме дисоційована тіолатна форма ( $-S^-$ ).

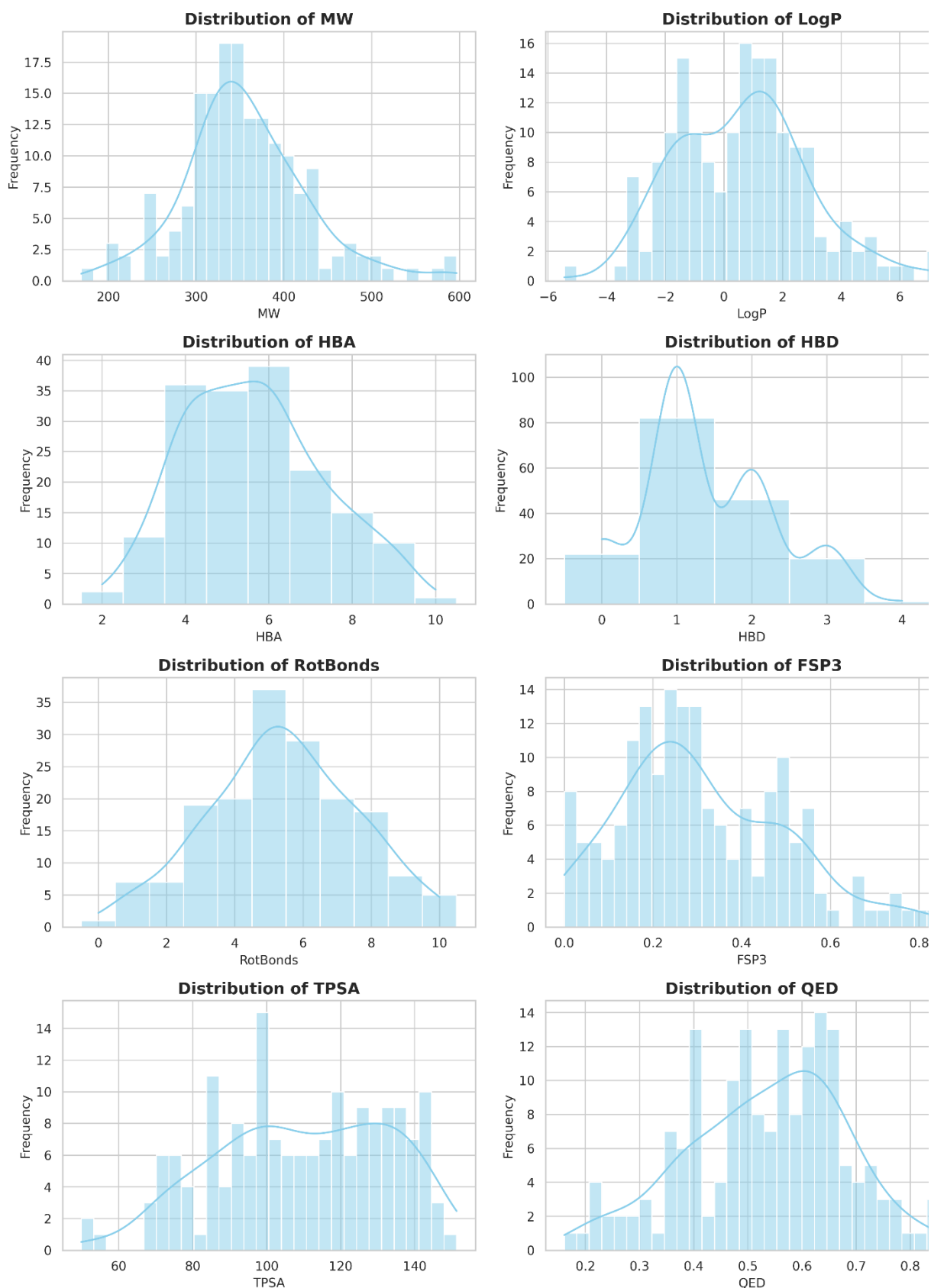
З обох баз Єнаміну ми відібрали 10 сполук - похідних тіолатів. Два приклади наведені на малюнку 81.



Мал.81. Ліворуч - 2d діаграма взаємодій з *t*\_271949\_20608416\_21050494, праворуч - 2d діаграма взаємодій з *s*\_273610\_20326850\_11516418.

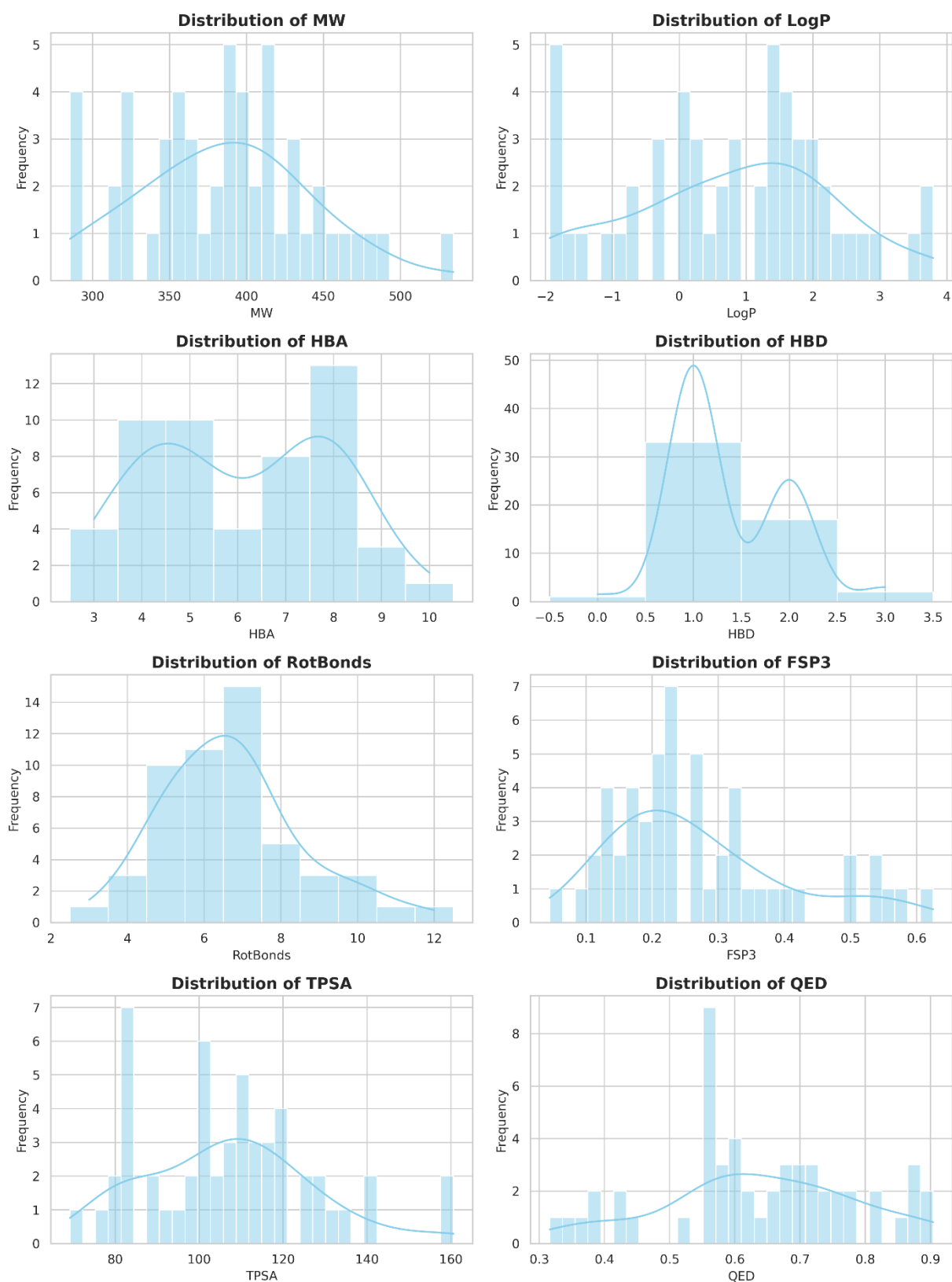
Разом зі всіх груп було відібрано 173 сполук, що доповнять 70 сполук знайдених у хіт-ту-лід стадії, після видалення перехресних дублікатів залишилось 171 та 53 сполуки відповідно. На малюнках 82 та 83 наведені графіки розподілу фізико-хімічних параметрів всіх 224 сполуки (Додаток Г). Ми сподіваємось, що ці сполуки будуть мати шанс бути біологічно валідовані в наступному раунді, та проявити активність *in vitro*.

*a*



*Мал.82. Розподіл фізико-хімічних параметрів 171 сполуки, відібраних шляхом скринінгу цинкових байндерів.*

6



Мал.83. Розподіл фізико-хімічних параметрів 53 сполук, знайдених за рахунок оптимізації хімії.

## ВИСНОВКИ

У дисертаційній роботі представлено цілісний підхід до дизайну та пошуку нових селективних інгібіторів ангіотензин-перетворювального ферменту 2 (АПФ2) із використанням сучасних методів комп'ютерного моделювання, органічного синтезу та біологічної валідації. За результатами виконаного дослідження сформульовано наступні основні висновки:

1. **Проаналізовані відомі хемоти́пи** відомих інгібіторів АПФ2. Здійснено синтез референтних сполук для подальших біологічних досліджень, а саме розробки асею та використанню у якості контролю;
2. **Оптимізовано та розроблено нові синтетичні маршрути**, які забезпечили високі виходи цільових речовин, спростили лабораторну реалізацію синтезу, зменшили витрати на реагенти та підвищили масштабованість процесів;
3. Серед десяти відомих літературних інгібіторів **два були повністю синтезовані** та оптимізовані у формі солей;
4. **Три нові хімічні зонди** подано на реєстрацію в міжнародний ресурс *Chemical Probe Portal* як потенційні молекулярні інструменти для дослідження АПФ2;
5. **Розроблено та валідовано біологічний асей** (флуоресцентному ферментативному HTS-аналіз) для проведення біологічних досліджень наявності інгібуючого ефекту сполук проти АПФ2;
6. **Розроблено і валідовано нові *in silico* моделі докінгу** для мішеней АПФ2 та АПФ, що продемонстрували високу точність у прогнозуванні активності відомих лігандів та дозволили оцінити селективність зв'язування;
7. Проведено **віртуальний скринінг Enamine In-Stock колекції** з використанням трьох незалежно валідованих воркфлоу, результати якого були перевірені за допомогою молекулярної динаміки;
8. **Реалізовано масштабний віртуальний скринінг** бібліотек обсягом до 1,4 млрд сполук із фокусом на хемотипове різноманіття та раціональне

- використання обчислювальних ресурсів шляхом впровадження багаторівневих фільтраційних протоколів;
9. За результатами комп'ютерного скринінгу **відібрано 577 сполук**, протестованих у флуоресцентному ферментативному HTS-аналізі. Виявлено **7 нових активних хітів**, що продемонстрували >30% інгібування, одна з яких (Z1459912954) виявила  $IC_{50} = 16,8$  мкМ та рекомендована для подальшої оптимізації;
  10. **Проведено віртуальну оптимізацію хітів** на основі отриманих хітів із метою підвищення активності та селективності;
  11. **Розроблено інтегрований робочий процес**, що включає гнучке моделювання байндинг-сайту (FTMove, FTMap), QSAR-моделювання на основі алгоритмів машинного навчання, та виявлення потенційних цинк-зв'язувальних сполук із відповідними координаційними властивостями;
  12. **Сформовано фінальний набір із 224 сполук-кандидатів**, рекомендованих для подальшої біологічної перевірки як потенційних інгібіторів АПФ2;
  13. Отримані результати засвідчили **ефективність поєднання комп'ютерних і експериментальних методів**, а розроблені робочі процеси можуть бути застосовані для дослідження інших цинк-залежних ферментів та інтегровані у сучасні платформи віртуального скринінгу та оптимізації хітів.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Abagyan, R. A.; Totrov, M. M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **1994**, *235*, 983–1002. <https://doi.org/10.1006/jmbi.1994.1052>
2. Abagyan, R. A.; Totrov, M. M.; Kuznetsov, D. N. ICM – a new method for protein modeling and design. Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **1994**, *15*, 488–506. <https://doi.org/10.1002/jcc.540150503>
3. Abraham, M. J.; Gready, J. E. Optimization of parameters for molecular dynamics simulation using smooth particle-mesh Ewald in GROMACS 4.5. *J. Comput. Chem.* **2011**, *32*, 2031–2040. <https://doi.org/10.1002/jcc.21773>
4. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001>
5. Adasme M. F., Linnemann K. L., Bolz S. N., Kaiser F., Salentin S., Haupt V. J., Schroeder M. PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Research*, **2021**, *49*(W1), W530–W534. <https://doi.org/10.1093/nar/gkab294>
6. Armarego, W. L. F.; Chai, C. *Purification of Laboratory Chemicals*, 5th ed.; Elsevier: Oxford, 2003
7. Arrowsmith, C. H.; Audia, J. E.; Austin, C.; Baell, J.; Bennett, J.; Blagg, J.; Bountra, C.; Brennan, P. E.; Brown, P. J.; Bunnage, M. E.; et al. The promise and peril of chemical probes. *Nat. Chem. Biol.* **2015**, *11*, 536–541. <https://doi.org/10.1038/nchembio.1867>
8. Ashton, M.; Barnard, J.; Casset, F.; Charlton, M.; Downs, G.; Gorse, D.; Holliday, J.; Lahana, R.; Willett, P. Identification of diverse database subsets using property-based and fragment-based molecular descriptions. *Quant. Struct.-Act. Relatsh.* **2002**, *21* (6), 598–604. <https://doi.org/10.1002/qsar.200290002>

9. Baell, J. B.; Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **2010**, *53*, 2719–2740. <https://doi.org/10.1021/jm901137j>
10. Bender, B. J.; Gahbauer, S.; Lutgens, A.; Lyu, J.; Webb, C. M.; Stein, R. M.; Fink, E. A.; Balius, T. E.; Carlsson, J.; Irwin, J. J.; et al. A practical guide to large-scale docking. *Nat. Protoc.* **2021**, *16*, 4799–4832. <https://doi.org/10.1038/s41596-021-00518-1>
11. Bhalla, V.; Blish, C. A.; South, A. M. A historical perspective on ACE2 in the COVID-19 era. *J. Hum. Hypertens.* **2021**, *35*, 935–939. <https://doi.org/10.1038/s41371-020-00429-2>
12. Bowers K. J., Chow E., Xu H., Dror R. O., Eastwood M. P., Gregersen B. A., Klepeis J. L., Kolossváry I., Moraes M. A., Sanfordoti F., Salmon J. K., Shan Y., Shaw D. E. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, Tampa, FL, USA, **2006**, pp. 43–43. <https://doi.org/10.1109/SC.2006.54>
13. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101. <https://doi.org/10.1063/1.2408420>
14. Byrnes J. J., Gross S., Ellard C., Connolly K., Donahue S., Picarella D. *Inflammation Research*, **2009**, *58*(11), 819–827. <https://doi.org/10.1007/s00011-009-0052-4>
15. Camargo S. M., Singer D., Makrides V., Huggel K., Pos K. M., Wagner C. A., Kuba K., Danilczyk U., Skovby F., Kleta R., Penninger J. M., Verrey F. Tissue-specific amino acid transporter partners ACE2 and collectrin differentially interact with Hartnup mutations. *Gastroenterology*, **2009**, *136*(3), 872–882. <https://doi.org/10.1053/j.gastro.2008.10.055>
16. Carpenter L., Kingsbury G., Tierney P., Dippel J., Hendrickson W., Hunt S., Sem D. S. Selective Binding of Lisinopril in Angiotensin-Converting Enzyme

- Domains. *Poster presentation*, Concordia University School of Pharmacy, Mequon, WI, USA
17. Chen R., Yu J., Wang K., Chen Z., Wen C., Xu Z. The spatial and cell-type distribution of SARS-CoV-2 receptor ACE2 in human and mouse brain. *BioRxiv*, **2020**, 1–26. <https://doi.org/10.3389/fneur.2020.573095>
  18. Choi S.-R., Breugst M., Houk K. N., Poulter C. D. *Journal of Organic Chemistry*, **2014**, 79(8), 3572–3580. <https://doi.org/10.1021/jo500131r>
  19. Corso G., Stärk H., Jing B., Barzilay R., Jaakkola T. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. *International Conference on Learning Representations (ICLR)*, **2023**. <https://doi.org/10.48550/arXiv.2210.01776>
  20. COVID-19 Map – Johns Hopkins Coronavirus Resource Center. <https://coronavirus.jhu.edu/map.html>
  21. Dales N. A., Gould A. E., Brown J. A., Calderwood E. F., Guan B., Minor C. A., Gavin J. M., Hales P., Kaushik V. K., Stewart M., Tummino P. J., Vickers C. S., Ocain T. D., Patane M. A. Substrate-based design of the first class of angiotensin-converting enzyme-related carboxypeptidase (ACE2) inhibitors. *Journal of Biological Chemistry*, **2002**. <https://doi.org/10.1074/jbc.M207383200>
  22. Davies S. G., Ichihara O., Lenoir I., Walters I. A. S. *Journal of the Chemical Society, Perkin Transactions I*, **1994**, 11, 1411–1416. <https://doi.org/10.1039/P19940001411>
  23. Daylight Chemical Information Systems. SMARTS – A Language for Describing Molecular Patterns. *Daylight Theory Manual*, **2008**
  24. Deaton D. N., Graham K. P., Gross J. W., Miller A. B. Thiol-based angiotensin-converting enzyme 2 inhibitors: P1' modifications for the exploration of the S1' subsite. *Bioorganic & Medicinal Chemistry Letters*, **2008**, 18, 1681–1687. <https://doi.org/10.1016/j.bmcl.2007.11.048>
  25. Deng, Y.F.; Aluko, R.E.; Jin, Q.; Zhang, Y.; Yuan, L.J. Inhibitory activities of baicalin against renin and angiotensin-converting enzyme. *Pharm. Biol.* **2012**, 50, 401–406. <https://doi.org/10.3109/13880209.2011.608076>

26. Di Gioia, M. L.; Costanzo, P.; De Nino, A.; Maiuolo, L.; Nardi, M.; Olivito, F.; Procopio, A. Simple and efficient Fmoc removal in ionic liquid. *RSC Advances*, **2017**, *7*(58), 36482–36491. <https://doi.org/10.1039/C7RA04425A>
27. Dong, E.; Du, H.; Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **2020**, *20*, 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
28. Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, Donovan M, Woolf B, Robison K, Jeyaseelan R, Breitbart RE, Acton S. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circ Res.* 2000 Sep 1;87(5):E1-9. doi: 10.1161/01.res.87.5.e1
29. Egbert M., Jones G., Collins M.R., Kozakov D., Vajda S. "FTMove: A Web Server for Detection and Analysis of Cryptic and Allosteric Binding Sites by Mapping Multiple Protein Structures." *Journal of Molecular Biology*, 2022, 434(11):167587. doi:10.1016/j.jmb.2022.167587
30. Ehrt C, Brinkjost T, Koch O (2018) A benchmark driven guide to binding site comparison: An exhaustive evaluation using tailor-made data sets (ProSPECCTs). *PLoS Comput Biol* 14(11): e1006483. <https://doi.org/10.1371/journal.pcbi.1006483>
31. Fischer A., Smieško M., Sellner M., Lill M. A. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. *Journal of Medicinal Chemistry*, **2021**, *64*, 2489–2500. <https://doi.org/10.1021/acs.jmedchem.0c02227>
32. Fu H., Fan Y., Zhang X., Lan H., Yang T., Shao M., Li S. Rapid Discrimination for Traditional Complex Herbal Medicines from Different Parts, Collection Time, and Origins Using High-Performance Liquid Chromatography and Near-Infrared Spectral Fingerprints with Aid of Pattern Recognition Methods. *Journal of Analytical Methods in Chemistry*, **2015**, *2015*, 727589. <https://doi.org/10.1155/2015/727589>

33. Gaulton A., Bellis L. J., Bento A. P., Chambers J., Davies M., Hersey A., Light Y., McGlinchey S., Michalovich D., Al-Lazikani B., et al. ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, **2012**, *40*, D1100–D1107. <https://doi.org/10.1093/nar/gkr777>
34. Gheblawi, M. Wang K., Viveiros A., Nguyen Q., Zhong J.-C., Turner A. J., Raizada M. K., Grant M. B., Oudit G. Y. Angiotensin-Converting Enzyme 2: SARS-CoV-2 Receptor and Regulator of the Renin-Angiotensin System. *Circulation Research*, **2020**, *126*, 1456–1474. <https://doi.org/10.1161/CIRCRESAHA.120.317015>
35. Groß S., Jahn C., Cushman S., Bär C., Thum T. SARS-CoV-2 receptor ACE2-dependent implications on the cardiovascular system: From basic science to clinical implications. *Journal of Molecular and Cellular Cardiology*, **2020**, *144*, 47–53. <https://doi.org/10.1016/j.yjmcc.2020.04.031>
36. Grygorenko, O.O.; Radchenko D. S., Dziuba I., Chuprina A., Gubina K. E., Moroz Y. S. Generating Multibillion Chemical Space of Readily Accessible Screening Compounds. *iScience*, **2020**, *23*, 101681. <https://doi.org/10.1016/j.isci.2020.101681>
37. Guan, W.-J., Ni Z.-Y., Hu Y., Liang W.-H., Ou C.-Q., He J.-X., et al. Clinical characteristics of coronavirus disease 2019 in China. *The New England Journal of Medicine*, **2020**, *382*, 1708–1720. <https://doi.org/10.1056/NEJMoa2002032>.
38. Guang C, Phillips R.D., Jiang B, Milani F. Three key proteases — angiotensin-I-converting enzyme (ACE), ACE2 and renin — within and beyond the renin-angiotensin system. *Archives of Cardiovascular Diseases*, **2012**, *105*(6–7), 373–385. <https://doi.org/10.1016/j.acvd.2012.02.010>
39. Guerrero, G.D.; Perez-S, H.E.; Cecilia, J.M.; Garcia, J.M. Parallelization of Virtual Screening in Drug Discovery on Massively Parallel Architectures. *Proceedings of the 2012 20th Euromicro International Conference on Parallel, Distributed and Network-based Processing*, Garching, Germany, **2012**, pp. 588–595. <https://doi.org/10.1109/PDP.2012.26>

40. Guy, J.L.; Jackson, R.M.; Acharya, K.R.; Sturrock, E.D.; Hooper, N.M.; Turner, A.J. Angiotensin-Converting Enzyme-2 (ACE2): Comparative Modeling of the Active Site, Specificity Requirements, and Chloride Dependence. *Biochemistry*, **2003**, *42*, 13185–13192. <https://doi.org/10.1021/bi0342793>
41. Haber, P.K.; Ye M., Wysocki J., et al. Angiotensin-converting enzyme 2–independent action of presumed angiotensin-converting enzyme 2 activators. *Hypertension*, **2014**, *63*(4), 774–782. <https://doi.org/10.1161/HYPERTENSIONAHA.113.02856>
42. Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of Medicinal Chemistry*, **2004**, *47*, 1750–1759. <https://doi.org/10.1021/jm030644s>
43. Han, D.P.; Penn-Nicholson A., Cho M. W. Identification of critical determinants on ACE2 for SARS-CoV entry and development of a potent entry inhibitor. *Virology*, **2006**, *350*, 15–25. <https://doi.org/10.1016/j.virol.2006.01.029>
44. Harding SD, Armstrong JF, Faccenda E, Southan C, Alexander SPH, Davenport AP, Spedding M, Davies JA. (2023) The IUPHAR/BPS Guide to PHARMACOLOGY in 2024. *Nucleic Acids Research*, **2023**, *52*(D1), D1438–D1449. <https://doi.org/10.1093/nar/gkad944>
45. Heyman, S. N.; Kinaneh, S.; Abassi, Z. The Duplicitous Nature of ACE2 in COVID-19 Disease. *EBioMedicine*, **2021**, *67*, 103356. <https://doi.org/10.1016/j.ebiom.2021.103356>
46. Ho, T.-Y.; Wu, S.-L.; Chen, J.-C.; Li, C.-C.; Hsiang, C.-Y. Emodin blocks the SARS coronavirus spike protein and angiotensin-converting enzyme 2 interaction. *Antivir. Res.* **2007**, *74*, 92-101. <https://doi.org/10.1016/j.antiviral.2006.04.014>
47. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Müller MA, Drosten C, Pöhlmann S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by

- a Clinically Proven Protease Inhibitor. *Cell*, **2020**, *181*(2), 271–280.e8. <https://doi.org/10.1016/j.cell.2020.02.052>
48. Horvath, D.; Brown, J.; Marcou, G.; Varnek, A. An evolutionary optimizer of *libsvm* models. *Challenges* **2014**, *5*, 450–472. <https://doi.org/10.3390/challe5020450>
49. Horvath, D.; Marcou, G.; Varnek, A. Predicting the Predictability: A Unified Approach to the Applicability Domain Problem of QSAR Models. *Journal of Chemical Information and Modeling*, **2009**, *49*, 1762–1776. <https://doi.org/10.1021/ci9000579>
50. <http://infochim.u-strasbg.fr/webserv/VSEngine.html> (станом на жовтень 2020 року)
51. <https://worldwide.espacenet.com/patent/search/family/039468477/publication/WO2008066770A2>
52. Huang, N.; Shoichet, B.K.; Irwin, J.J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789–6801. <https://doi.org/10.1021/jm0608356>.
53. Imai, Y.; Kuba, K.; Penninger, J.M. The discovery of angiotensin-converting enzyme 2 and its role in acute lung injury in mice. *Exp. Physiol.* **2008**, *93*, 543–548. <https://doi.org/10.1113/expphysiol.2007.040048>
54. Jacobsen J.A., Major Jourden J.L., Miller M.T., Cohen S.M. To bind zinc or not to bind zinc: An examination of innovative approaches to improved metalloproteinase inhibition. *Biochimica et Biophysica Acta*, **2010**, *1803*(1), 72–94. <https://doi.org/10.1016/j.bbamcr.2009.09.005>
55. Jiang F, Deng L, Zhang L, Cai Y, Cheung CW, Xia Z. Review of the Clinical Characteristics of Coronavirus Disease 2019 (COVID-19). *Journal of General Internal Medicine*, **2020**, *35*(5), 1545–1549. <https://doi.org/10.1007/s11606-020-05762-w>
56. Jones G, Jindal A, Ghani U, Kotelnikov S, Egbert M, Hashemi N, Vajda S, Padhorny D, Kozakov D. Elucidation of protein function using computational docking and hotspot analysis by ClusPro and FTMap. *Acta Crystallographica*

- Section D: Structural Biology*, **2022**, 78(Pt 6), 690–697.  
<https://doi.org/10.1107/S2059798322002741>
57. Joshi S, Balasubramanian N, Vasam G, Jarajapu YP. Angiotensin converting enzyme versus angiotensin converting enzyme-2 selectivity of MLN-4760 and DX600 in human and murine bone marrow-derived cells. *European Journal of Pharmacology*, **2016**, 774, 25–33. <https://doi.org/10.1016/j.ejphar.2016.01.007>
  58. Joshi, R.S.; Jagdale, S.S.; Bansode, S.B.; Shankar, S.S.; Tellis, M.B.; Pandya, V.K.; Chugh, A.; Giri, A.P.; Kulkarni, M.J. Discovery of Potential Multi-Target-Directed Ligands by Targeting Host-specific SARS-CoV-2 Structurally Conserved Main Protease. *J. Biomol. Struct. Dyn.* **2020**, 1–16. <https://doi.org/10.1080/07391102.2020.1760137>
  59. Jullien N., Makritis A., Georgiadis D., Beau F., Yiotakis A., Dive V. Phosphinic tripeptides as dual angiotensin-converting enzyme c-domain and endothelin-converting enzyme-1 inhibitors. *Journal of Medicinal Chemistry*, **2009**, 53(1), 208–220. <https://doi.org/10.1021/jm9010803>
  60. Karuppusamy, V.; Ilangovan, A. *Org. Lett.*, **2020**, 22(18), 7147–7151.
  61. Keith A. McCall, Huang C.-C., Fierke C. A. Function and Mechanism of Zinc Metalloenzymes. *The Journal of Nutrition*, **2000**, 130(5), 1437S–1446S. <https://doi.org/10.1093/jn/130.5.1437S>
  62. Konc, J.; Cesnik, T.; Konc, J.T.; Penca, M.; Janežič, D. ProBiS-database: Precalculated binding site similarities and local pairwise alignments of PDB structures. *J. Chem. Inf. Model.* **2012**, 52, 604–612. <https://doi.org/10.1021/ci2005687>
  63. Korb, O.; Stutzle, T.; Exner, T.E. Empirical scoring functions for advanced protein-ligand docking with PLANTS. *J. Chem. Inf. Modeling* **2009**, 49, 84–96. <https://doi.org/10.1021/ci800298z>
  64. Kowalczyk S., Bröer A., Tietze N., Vanslambrouck J. M., Rasko J. E., Bröer S. A protein complex in the brush-border membrane explains a Hartnup disorder allele. *FASEB Journal*, **2008**, 22(8), 2880–2887. <https://doi.org/10.1096/fj.08-107300>

65. Kozakov D., Grove L. E., Hall D. R., Bohnuud T., Mottarella S. E., Luo L., Xia B., Beglov D., Vajda S. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nature Protocols*, **2015**, *10*(5), 733–755. <https://doi.org/10.1038/nprot.2015.043>
66. Kravets I., Cherednichenko A., Tararina V., Kapeliukha A., Dudenko D., Protopopov M. Creation of targeted compound libraries based on 3D shape recognition. *Molecular Diversity*, **2023**, *27*(2), 939–949. <https://doi.org/10.1007/s11030-022-10447-z>
67. Lambert D.W., Hooper N. M., Turner A. J. Angiotensin-converting enzyme 2 and new insights into the renin-angiotensin system. *Biochemical Pharmacology*, **2008**, *75*(4), 781–786. <https://doi.org/10.1016/j.bcp.2007.08.012>
68. Li, Y.-C.; Bai W.-Z., Hashikawa T. The neuroinvasive potential of SARS-CoV-2 may play a role in the respiratory failure of COVID-19 patients. *Journal of Medical Virology*, **2020**, *92*, 552–555. <https://doi.org/10.1002/jmv.25728>
69. Liao, K. Sikkema, D., Wang, C., & Lee, T. N. (2013). Development of an enzymatic assay for the detection of neutralizing antibodies against therapeutic angiotensin-converting enzyme 2 (ACE2). *Journal of immunological methods*, *389*(1-2), 52–60. <https://doi.org/10.1016/j.jim.2012.12.010>
70. Lubbe L., Cozier G. E., Oosthuizen D., Acharya K. R., Sturrock E. D. ACE2 and ACE: structure-based insights into mechanism, regulation and receptor recognition by SARS-CoV. *Clinical Science*, **2020**, *134*(21), 2851–2871. <https://doi.org/10.1042/CS20200899>
71. Ma, L.L.; Liu, H.M.; Liu, X.M.; Yuan, X.Y.; Xu, C.; Wang, F.; Lin, J.Z.; Xu, R.C.; Zhang, D.K. Screening S protein - ACE2 blockers from natural products: Strategies and advances in the discovery of potential inhibitors of COVID-19. *Eur J. Med. Chem* **2021**, *226*, 113857. <https://doi.org/10.1016/j.ejmech.2021.113857>
72. Margreitter, C.; Oostenbrink, C. MDplot: Visualise Molecular Dynamics. *R J.* **2017**, *9*, 164–186
73. Masuyer, G., Schwager S., Sturrock E., et al. Molecular recognition and regulation of human angiotensin-I converting enzyme (ACE) activity by natural

- inhibitory peptides. *Scientific Reports*, **2012**, *2*, 717. <https://doi.org/10.1038/srep00717>
74. Mores A.; Matziari M.; Beau F.; Cuniasse P.; Yiotakis A.; Dive V. Development of potent and selective phosphinic peptide inhibitors of angiotensin-converting enzyme 2. *J. Med. Chem.* 2008, *51* (7), 2216–2226. <https://doi.org/10.1021/jm701275z>
75. Muchtaridi, M., Fauzi M., Khairul Ikram N. K., Mohd Gazzali A., Wahab H. A. Natural Flavonoids as Potential Angiotensin-Converting Enzyme 2 Inhibitors for Anti-SARS-CoV-2. *Molecules*, **2020**, *25*(17), 3980. <https://doi.org/10.3390/molecules25173980>
76. Mues H., Kazmaier U. A Straightforward Approach Towards Substituted cis Hydroxyprolines. *Synlett*, **2000**, (7), 1004–1006. <https://doi.org/10.1055/s-2000-6652>
77. Mues, U. The Asymmetric Chelate-Claisen Rearrangement as a Key Step in the Syntheses of Non-Proteinogenic Amino Acids. *Synthesis*, **2001**, *3*, 487–498. <https://doi.org/10.1055/s-2001-11451>
78. Mysinger, M.M.; Carchia, M.; Irwin, J.J.; Shoichet, B.K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–6594. <https://doi.org/10.1021/jm300687e>
79. Neves M., Totrov M., Abagyan R. Docking and scoring with ICM: the benchmarking results and strategies for improvement. *Journal of Computer-Aided Molecular Design*, **2012**, *26*(6), 675–686. <https://doi.org/10.1007/s10822-012-9547-0>
80. Othman, H.; Bouslama, Z.; Brandenburg, J.-T.; Da Rocha, J.; Hamdi, Y.; Ghedira, K.; Abid, N.-S.; Hazelhurst, S. *In silico* study of the spike protein from SARS-CoV-2 interaction with ACE2: Similarity with SARS-CoV, hot-spot analysis and effect of the receptor polymorphism. *BioRxiv* **2020**, 527, 702–708. <https://doi.org/10.1101/2020.03.04.976027>

81. Pajouhesh, H. & Lenz, G. R. (2005). Medicinal chemical properties of successful central nervous system drugs. *NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics*, 2(4), 541–553. <https://doi.org/10.1602/neurorx.2.4.541>
82. Pan, L., Mu, M., Yang, P., Sun, Y., Wang, R., Yan, J., Li, P., Hu, B., Wang, J., Hu, C., Jin, Y., Niu, X., Ping, R., Du, Y., Li, T., Xu, G., Hu, Q., & Tu, L. (2020). Clinical Characteristics of COVID-19 Patients With Digestive Symptoms in Hubei, China: A Descriptive, Cross-Sectional, Multicenter Study. *The American journal of gastroenterology*, 115(5), 766–773. <https://doi.org/10.14309/ajg.0000000000000620>
83. Pardridge, W.M. Transport of small molecules through the blood-brain barrier: Biology and methodology. *Adv. Drug Deliv. Rev.* **1995**, 15, 5–36
84. Parrinello M., Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, **1981**, 52, 7182–7190. <https://doi.org/10.1063/1.328693>
85. Pedretti, A.; Villa, L.; Vistoli, G. VEGA—An open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. *J. Comput.-Aided Mol. Des.* **2004**, 18, 167–173. <https://doi.org/10.1023/b:jcam.0000035186.90683.f2>
86. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T.E. 3rd; Laughton, C.A.; Orozco, M. Refinement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers. *Biophys. J.* **2007**, 92, 3817–3829. <https://doi.org/10.1529/biophysj.106.097782>
87. Prada J. A. H., Ferreira A. J., Katovich M. J., et al. Structure-based identification of small-molecule angiotensin-converting enzyme 2 activators as novel antihypertensive agents. *Hypertension*, **2008**, 51(5), 1312–1317. <https://doi.org/10.1161/HYPERTENSIONAHA.107.108944>
88. Rakhshani, H., Dehghanian E., Rahati A. Enhanced GROMACS: Toward a better numerical simulation framework. *Journal of Molecular Modeling*, **2019**, 25, 355. <https://doi.org/10.1007/s00894-019-4232-z>

89. Raz A., Gamliel-Lazarovich A., Bogner I., Strigevsky A., Keidar S. [The importance of ACE2 in regulating the cardiovascular system]. *Harefuah*, **2007**, *146*(9), 703–706, 733
90. Rella M., Rushworth C. A., Guy J. L., Turner A. J., Langer T., Jackson R. M. Structure-based pharmacophore design and virtual screening for novel angiotensin converting enzyme 2 inhibitors. *Journal of Chemical Information and Modeling*, **2006**, *46*(2), 708–716. <https://doi.org/10.1021/ci0503614>
91. Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754. <https://doi.org/10.1021/ci100050t>
92. Rushworth C. A., Guy J. L., Turner A. J. Residues affecting the chloride regulation and substrate selectivity of the angiotensin-converting enzymes (ACE and ACE2) identified by site-directed mutagenesis. *FEBS Journal*, **2008**, *275*(23), 6033–6042. <https://doi.org/10.1111/j.1742-4658.2008.06733.x>
93. Sanfelice D., Antolin A. A., Crisp A., Chen Y., Bellenie B., Brennan P. E., Edwards A., Müller S., Al-Lazikani B., Workman P. The Chemical Probes Portal – 2024: update on this public resource to support best-practice selection and use of small molecules in biomedical research. *Nucleic Acids Research*, **2025**, *53*(D1), D1663–D1669. <https://doi.org/10.1093/nar/gkae1062>
94. Schäfer M., Stünkel T., Daniliuc C. G., Gilmour R. Regio- and enantioselective intermolecular aminofluorination of alkenes via iodine(I)/iodine(III) catalysis. *Angewandte Chemie International Edition*, **2022**, *61*(32), e202205508. <https://doi.org/10.1002/anie.202205508>
95. Schreiber, S.L.; Kotz, J.D.; Li, M.; Aube, J.; Austin, C.P.; Reed, J.C.; Rosen, H.; White, E.L.; Sklar, L.A.; Lindsley, C.W.; et al. Advancing Biological Understanding and Therapeutics Discovery with Small-Molecule Probes. *Cell* **2015**, *161*, 1252–1265. <https://doi.org/10.1002/anie.202205508>
96. Schreyer A. M., Blundell T. USRCAT: Real-time ultrafast shape recognition with pharmacophoric constraints. *Journal of Cheminformatics*, **2012**, *4*, 27.
97. Schrödinger LLC. The PyMOL Molecular Graphics System, Version 3.0. New York, NY, USA

98. Schrödinger Release 2021-3: LigPrep. Schrödinger, LLC, New York, NY, **2021**
99. Sepehrinezhad, A.; Shahbazi, A.; Negah, S.S. COVID-19 virus may have neuroinvasive potential and cause neurological complications: A perspective review. *J. Neurovirol.* **2020**, *26*, 324–329. <https://doi.org/10.1007/s13365-020-00851-2>
100. Shang J., Wan Y., Luo C., Ye G., Geng Q., Auerbach A., Li F. Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences of the USA*, **2020**, *117*, 11727–11734. <https://doi.org/10.1073/pnas.2003138117>
101. Shang J., Ye G., Shi K., Wan Y., Luo C., Aihara H., Geng Q., Auerbach A., Li F. Structural basis of receptor recognition by SARS-CoV-2. *Nature*, **2020**, *581*, 221–224. <https://doi.org/10.1038/s41586-020-2179-y>
102. Sidorov, P.; Gaspar, H.; Marcou, G.; Varnek, A.; Horvath, D. Mappability of drug-like space: Towards a polypharmacologically competent map of drug-relevant compounds. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 1087–1108. <https://doi.org/10.1007/s10822-015-9882-z>
103. Sui H., Yu Q., Zhi Y., Geng G., Liu H., Xu H. Effects of apigenin on the expression of angiotensin-converting enzyme 2 in kidney in spontaneously hypertensive rats. *Journal of Hygiene Research*, **2010**, *39*, 693–696
104. Tanrikulu, Y.; Krüger, B.; Proschak, E. The holistic integration of virtual screening in drug discovery. *Drug Discov. Today* **2013**, *18*, 358–364. <https://doi.org/10.1016/j.drudis.2013.01.007>
105. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, **2023**, *51*(D1), D523–D531. <https://doi.org/10.1093/nar/gkac1052>
106. Tipnis S. R., Hooper N. M., Hyde R., Karran E., Christie G., Turner A. J. A human homolog of angiotensin-converting enzyme. Cloning and functional expression as a captopril-insensitive carboxypeptidase. *Journal of Biological Chemistry*, **2000**, *275*(43), 33238–33243. <https://doi.org/10.1074/jbc.M002615200>

107. Towler P., Staker B., Prasad S. G., Menon S., Tang J., Parsons T., Ryan D., Fisher M., Williams D., Dales N. A., Patane M. A., Pantoliano M. W. ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *Journal of Biological Chemistry*, **2004**, 279(17), 17996–18007. <https://doi.org/10.1074/jbc.M311191200>
108. Truchon J. F., Bayly C. I. Evaluating virtual screening methods: Good and bad metrics for the “early recognition” problem. *Journal of Chemical Information and Modeling*, **2007**, 47, 488–508. <https://doi.org/10.1021/ci600426e>
109. Utomo, R.Y.; Ikawati, M.; Meiyanto, E. Revealing the Potency of Citrus and Galangal Constituents to Halt SARS-CoV-2 Infection. *Preprints* **2020**, 214, 1–8. <https://doi.org/10.20944/preprints202003.0214.v1>
110. Varnek, A.; Fourches, D.; Horvath, D.; Klimchuk, O.; Gaudin, C.; Vayer, P.; Solov'ev, V.; Hoonakker, F.; Tetko, I.V.; Marcou, G. ISIDA - Platform for virtual screening based on fragment and pharmacophoric descriptors. *Curr. Comput.-Aided Drug Des.* **2008**, 4, 191–198. <https://doi.org/10.2174/157340908785747465>
111. Vickers C., Hales P., Kaushik V., Dick L., Gavin J., Tang J., Godbout K., Parsons T., Baronas E., Hsieh F., Acton S., Patane M., Nichols A., Tummino P. Hydrolysis of biological peptides by human angiotensin-converting enzyme-related carboxypeptidase. *Journal of Biological Chemistry*, **2002**, 277(17), 14838–14843. <https://doi.org/10.1074/jbc.M200581200>
112. Volochnyuk, D. M., Ryabukhin, S. V., Moroz, Y. S., Savych, O., Chuprina, A., Horvath, D., Zabolotna, Y., Varnek, A., & Judd, D. B. (2019). Evolution of commercially available compounds for HTS. *Drug discovery today*, 24(2), 390–402. <https://doi.org/10.1016/j.drudis.2018.10.016>
113. Wang G., Yang M. L., Duan Z. L., et al. 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 Research*, **2021**, 31, 17–24. <https://doi.org/10.1038/s41422-020-00450-0>

114. Wang W., McKinnie S. M., Farhan M., Paul M., McDonald T., McLean B., Llorens-Cortes C., Hazra S., Murray A. G., Vederas J. C., Oudit G. Y. Angiotensin-Converting Enzyme 2 Metabolizes and Partially Inactivates Pyr-Apelin-13 and Apelin-17: Physiological Effects in the Cardiovascular System. *Hypertension*, **2016**, *68*(2), 365–377. <https://doi.org/10.1161/HYPERTENSIONAHA.115.06892>
115. Williams B. Renin Angiotensin System Inhibition as treatment for Covid-19? *EClinicalMedicine*, **2021**, *37*, 101023. <https://doi.org/10.1016/j.eclinm.2021.101023>
116. Wolber, G.; Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Modeling* **2005**, *45*, 160–169. <https://doi.org/10.1021/ci049885e>
117. Workman, P.; Collins, I. Probing the probes: Fitness factors for small molecule tools. *Chem. Biol.* **2010**, *17*, 561–577. <https://doi.org/10.1016/j.chembiol.2010.05.013>
118. Wu K., Li W., Peng G., Li F. Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. *Proceedings of the National Academy of Sciences of the USA*, **2009**, *106*(47), 19970–19974. <https://doi.org/10.1073/pnas.0908837106>
119. Wu, L.; Zhou, L.; Mo, M.; Liu, T.; Wu, C.; Gong, C.; Lu, K.; Gong, L.; Zhu, W.; Xu, Z. SARS-CoV-2 Omicron RBD shows weaker binding affinity than the currently dominant Delta variant to human ACE2. *Signal Transduct. Target. Ther.* **2022**, *7*, 8. <https://doi.org/10.1038/s41392-021-00863-2>
120. Xia, S.; Zhu, Y.; Liu, M.; Lan, Q.; Xu, W.; Wu, Y.; Ying, T.; Liu, S.; Shi, Z.; Jiang, S.; et al. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. *Cell Mol. Immunol.* **2020**, *17*, 765–767. <https://doi.org/10.1038/s41423-020-0374-2>
121. Xiao F., Tang M., Zheng X., Liu Y., Li X., Shan H. Evidence for Gastrointestinal Infection of SARS-CoV-2. *Gastroenterology*, **2020**, *158*(4), 1831–1833. <https://doi.org/10.1053/j.gastro.2020.02.055>

122. Xiao, F.; Tang, M.; Zheng, X.; Liu, Y.; Li, X., Shan, H. Evidence for Gastrointestinal Infection of SARS-CoV-2 // *Gastroenterology* – 2020 – V. 158 – P. 1831-1833. №4
123. Xu, Y.; Lou, Z.; Liu, Y.; Pang, H.; Tien, P.; Gao, G.F.; Rao, Z. Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. *J. Biol. Chem.* **2004**, *279*, 49414–49419. <https://doi.org/10.1074/jbc.M408782200>
124. Yang P., Kuc R. E., Brame A. L., Dyson A., Singer M., Glen R. C., Cheriyan J., Wilkinson I. B., Davenport A. P., Maguire J. J. [Pyr1]Apelin-13(1–12) Is a Biologically Active ACE2 Metabolite of the Endogenous Cardiovascular Peptide [Pyr1]Apelin-13. *Frontiers in Neuroscience*, **2017**, *11*, 92. <https://doi.org/10.3389/fnins.2017.00092>
125. Yang Z., Lasker K., Schneidman-Duhovny D., Webb B., Huang C. C., Pettersen E. F., Goddard T. D., Meng E. C., Sali A., Ferrin T. E. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *Journal of Structural Biology*, **2012**, *179*(3), 269–278. <https://doi.org/10.1016/j.jsb.2011.09.006>
126. Yeo C., Kaushal S., Yeo D. Enteric involvement of coronaviruses: Is faecal–oral transmission of SARS-CoV-2 possible? *The Lancet Gastroenterology & Hepatology*, **2020**, *5*, 335–337. [https://doi.org/10.1016/S2468-1253\(20\)30048-0](https://doi.org/10.1016/S2468-1253(20)30048-0)
127. Yu J.-w., Wang L., Bao L.-d. Exploring the Active Compounds of Traditional Mongolian Medicine in Intervention of Novel Coronavirus (COVID-19) Based on Molecular Docking Method. *Journal of Functional Foods*, **2020**. <https://doi.org/10.1016/j.jff.2020.104016>
128. Zhu, T.; Cao, S.; Su, P.-C.; Patel, R.; Shah, D.; Chokshi, H.B.; Szukala, R.; Johnson, M.E.; Hevener, K.E. Hit Identification and Optimization in Virtual Screening: Practical Recommendations Based on a Critical Literature Analysis. *J. Med. Chem.* **2013**, *56*, 6560–6572. <https://doi.org/10.1021/jm301916b>

## СПИСОК ПУБЛІКАЦІЙ

*Статті в іноземних та українських фахових журналах:*

1. **Kravets, I. O.**, Dudenko, D. V., Pashenko, A. E., Borisova, T. A., Tolstanova, G. M., Ryabukhin, S. V., & Volochnyuk, D. M. (2021). Virtual Screening in Search for a Chemical Probe for Angiotensin-Converting Enzyme 2 (ACE2). *Molecules*, 26(24), 7584. <https://doi.org/10.3390/molecules26247584>.

*Особистий внесок: збір літературних даних, побудування моделей віртуального скринінгу, побудова та оптимізація робочого процесу скринінгу, проведення віртуального скринінгу, підготовка всіх ілюстрацій та візуалізацій, написання статті.*

2. Rayevsky, A. V., Poturai, A. S., **Kravets, I. O.**, Pashenko, A. E., Borisova, T. A., Tolstanova, G. M., Volochnyuk, D. M., Borysko, P. O., Vadzyuk, O. B., Alieksieieva, D. O., Zabolotna, Y., Klimchuk, O., Horvath, D., Marcou, G., Ryabukhin, S. V., & Varnek, A. (2022). *In Vitro* Evaluation of *In silico* Screening Approaches in Search for Selective ACE2 Binding Chemical Probes. *Molecules* (Basel, Switzerland), 27(17), 5400. <https://doi.org/10.3390/molecules27175400>.

*Особистий внесок: збір літературних даних, побудування моделей віртуального скринінгу, проведення віртуального скринінгу, написання експериментальної частини статті.*

3. Andrii Kyrylchuk, **Iryna Kravets**, Anton Cherednichenko, Valentyna Tararina, Anna Kapeliukha, Dmytro Dudenko & Mykola Protopopov. Creation of targeted compound libraries based on 3D shape recognition. *Mol Divers* 27, 939–949 (2023). <https://doi.org/10.1007/s11030-022-10447-z>

*Особистий внесок: збір літературних даних, побудування та валідування робочого процесу, виконання побудови таргетних бібліотек, написання експериментальної частини статті.*

*Тези доповідей на конференціях:*

1. **Ірина О. Кравець**, Олексій В. Раєвський, Дмитро В. Дуденко, Юрій С. Мороз, Олександр Є. Пашенко, Дмитро М. Волочнюк, Сергій В. Рябухін. Нові хемотипи зв'язуючих речовин для АПФ2, знайдені за допомогою послідовних підходів *in silico* та *in vitro* скринінгу. XXVI УКРАЇНСЬКА КОНФЕРЕНЦІЯ З ОРГАНІЧНОЇ ТА БІООРГАНІЧНОЇ ХІМІЇ. Ужгород. 16-20 вересня 2024р.
2. Alexey Rayevsky, Dmytro Dudenko, **Iryna Kravets**, Yurii Moroz, Alexander Pashenko, Dmitriy Volochnyuk, Sergey Ryabukhin. In search for selective ACE2 chemical probes: *In silico* studies on “world joined screening collection” compounds dataset. ACS Fall 2021, Atlanta, GA, USA, August 22-26, 2021, PID 3593344.
3. O. V. Raevskii, D. V. Dudenko, **I. O. Kravets**, Y. S. Moroz, A. E. Pashenko, D. M. Volochnyuk, S. V. Ryabukhin. Novel chemotypes of angiotensin-converting enzyme 2 binders via successive *in silico* screening and *in vitro* evaluation approaches. EFMC-ASMC 2023, International Symposium on Advances in Synthetic and Medicinal Chemistry, Zagreb, Croatia, September 3-7, 2023, P028.

Список лігандів з бази даних ChEMBL з експериментально вимірною активністю Кі або IC<sub>50</sub> проти АПФ2.

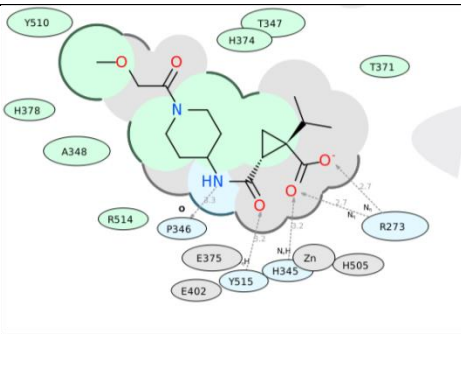
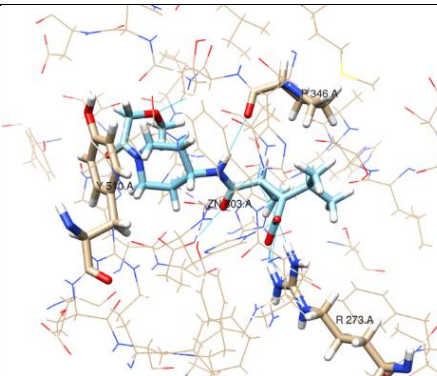
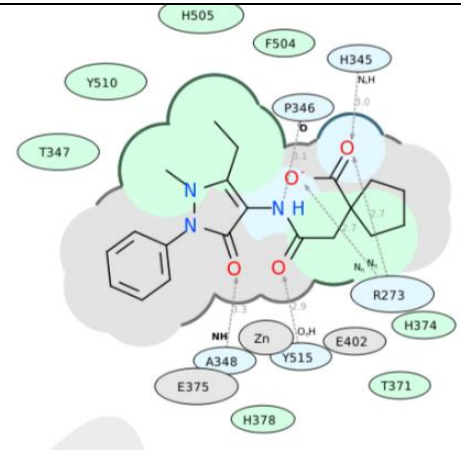
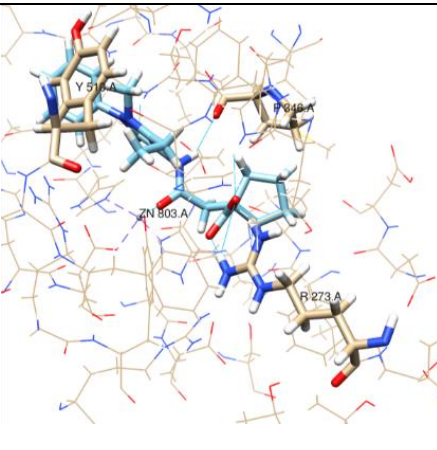
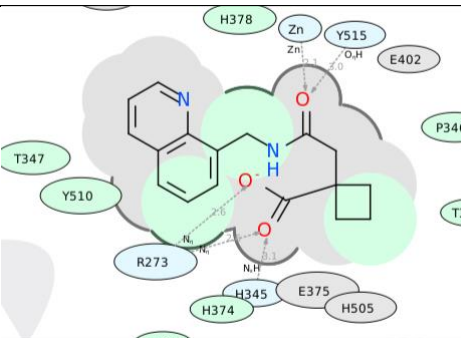
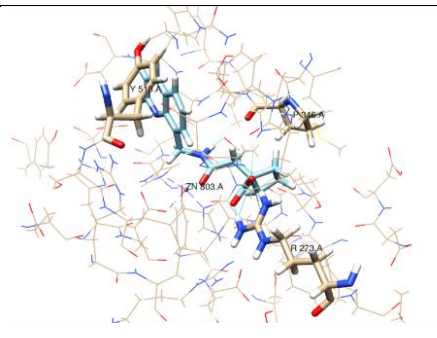
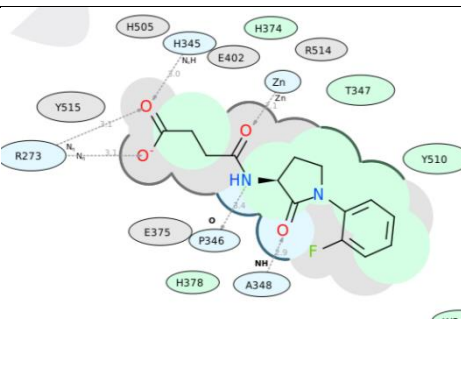
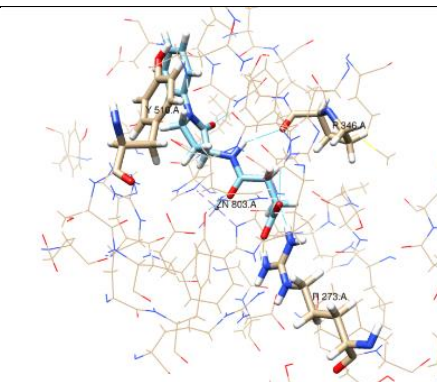
Smiles	Molecule ChEMBL ID
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL253224
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(Oc2ccccc2)c1)C(=O)O</chem>	CHEMBL252003
<chem>CCCC[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL257270
<chem>CC[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL257726
<chem>O=C(S)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL258333
<chem>O=C(O)[C@H](Cc1ccc(-c2ccccc2)cc1)NC(=O)[C@H](S)Cc1ccccc1</chem>	CHEMBL163454
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL254493
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL400527
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL398771
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL254911
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](C)C(=O)O</chem>	CHEMBL254900
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL252391
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL252009
<chem>CC[C@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL269996
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)c1)C(=O)O</chem>	CHEMBL401397
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL400526
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL254495
<chem>CC(C)C[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL271225
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL398545
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL254282
<chem>O=C(O)[C@H](Cc1ccc(-c2ccccc2)cc1)NC(=O)[C@@H](S)C1CCCC1</chem>	CHEMBL404117
<chem>O=C(N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O)[C@@H](S)Cc1ccccc1</chem>	CHEMBL350414
<chem>O=C(N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O)[C@@H](S)CC1CCCC1</chem>	CHEMBL271224
<chem>CC(C)(C)C[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL271223
<chem>O=C(O)[C@H](Cc1ccc(-c2ccccc2)cc1)NC(=O)[C@@H](S)C1CCCC1</chem>	CHEMBL257026
<chem>CC(=O)N[C@@H](Cc1c[nH]cn1)C(=O)N[C@@H](CC(C)C)P(=O)(O)CC(Cc1cc(-c2ccccc2)no1)C(=O)O</chem>	CHEMBL408182
<chem>CC(=O)N[C@@H](CC(C)C)C(=O)N1CCC[C@@H]1P(=O)(O)CC(Cc1cc(-c2ccccc2)no1)C(=O)O</chem>	CHEMBL261121
<chem>CC(C)C[C@@H](N)P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O</chem>	CHEMBL405232
<chem>CC(C)(S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL404044
<chem>C[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O</chem>	CHEMBL257727
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL251809
<chem>CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)O</chem>	CHEMBL251804
<chem>CC(C)C[C@@H](NC(=O)OCc1ccccc1)P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O</chem>	CHEMBL264665

O=C(N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O)[C@@H](S)Cc1ccc2ccccc2c1 CHEMBL272925  
O=C(O)[C@H](Cc1ccc(-c2ccccc2)cc1)NC(=O)[C@@H](S)c1ccccc1 CHEMBL437595  
O=C(O)[C@H](Cc1ccccc1)CP(=O)(O)[C@H]1CCCN1 CHEMBL261033  
CC(=O)N[C@@H](Cc1c[nH]cn1)C(=O)N1CCC[C@@H]1P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O  
CHEMBL261423  
CC(=O)N[C@@H](CCCCN)C(=O)N1CCC[C@@H]1P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O CHEMBL411942  
C[C@@H](N)P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O CHEMBL405913  
N[C@H](Cc1ccccc1)P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O CHEMBL412123  
CC(=O)N[C@@H](CC(C)C)C(=O)N1CCC[C@@H]1P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O CHEMBL260273  
C[C@H](CP(=O)(O)[C@H]1CCCN1C(=O)OCc1ccccc1)C(=O)O CHEMBL409713  
CC(=O)N[C@@H](CCC(=O)O)C(=O)N1CCC[C@@H]1P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O  
CHEMBL263663  
CC(=O)N[C@@H](Cc1c[nH]cn1)C(=O)N[C@H](CC(C)C)P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O  
CHEMBL408448  
CC(=O)N[C@H](C(=O)N1CCC[C@@H]1P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O)C(C)C CHEMBL411052  
CC(=O)N[C@@H](Cc1ccc(O)cc1)C(=O)N1CCC[C@@H]1P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O  
CHEMBL410509  
CC(C)C[C@H](CP(=O)(O)[C@H]1CCCN1C(=O)OCc1ccccc1)C(=O)O CHEMBL411298  
CC(=O)N[C@@H](Cc1c[nH]cn1)C(=O)N1CCC[C@@H]1P(=O)(O)CC(Cc1cc(-c2ccccc2)no1)C(=O)O  
CHEMBL409105  
CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(OCc2ccccc2)cc1)C(=O)O CHEMBL254703  
CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc2ccccc2c1)C(=O)O CHEMBL253428  
CC[C@@H](C)[C@H](S)C(=O)N[C@@H](Cc1ccc(OCc2ccc(F)c(F)c2)cc1)C(=O)O CHEMBL251808  
CC[C@@H](C)[C@H](S)C(=O)NCC(=O)O CHEMBL401086  
CC(=O)N[C@@H](CC(C)C)C(=O)N[C@H](Cc1ccccc1)P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O  
CHEMBL408705  
CC(=O)N[C@@H](Cc1ccccc1)C(=O)N1CCC[C@@H]1P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O  
CHEMBL258698  
O=C(N[C@H](Cc1ccccc1)P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O)OCc1ccccc1 CHEMBL258464  
CC(=O)N[C@@H](C)C(=O)N1CCC[C@@H]1P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O CHEMBL409721  
O=C(O)[C@H](Cc1ccccc1)CP(=O)(O)[C@H]1CCCN1C(=O)OCc1ccccc1 CHEMBL260677  
C[C@@H](NC(=O)OCc1ccccc1)P(=O)(O)C[C@@H](Cc1ccccc1)C(=O)O CHEMBL258683  
O=C(N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O)[C@@H](S)CCc1ccccc1 CHEMBL402987  
O=C(O)[C@H](Cc1ccc(-c2ccccc2)cc1)NC(=O)[C@@H](S)C1CCC1 CHEMBL257229  
CC(C)[C@H](S)C(=O)N[C@@H](Cc1ccc(-c2ccccc2)cc1)C(=O)O CHEMBL269997  
Cc1nnc(S/C(=C/c2ccc(-c3cc(Cl)ccc3C)o2)C(=O)O)[nH]1 CHEMBL252417  
CC(C)C[C@H](N[C@@H](Cc1cncn1Cc1cc(Cl)cc(Cl)c1)C(=O)O)C(=O)O CHEMBL429844  
CNC(=O)[C@H](Cc1c[nH]cn1)NC(=O)CN(CCCc1ccccc1)CC(=O)O CHEMBL3235416  
NCCNCCN1CC1 CHEMBL398940

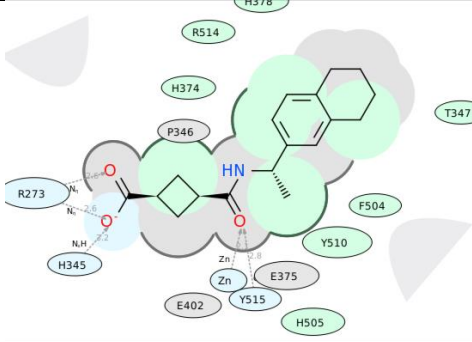
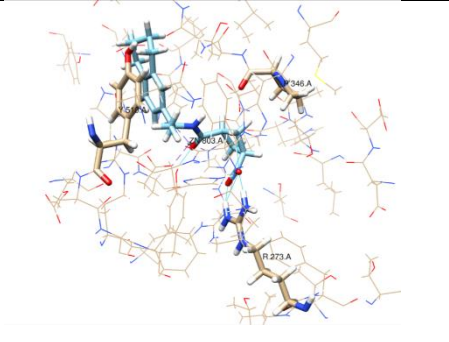
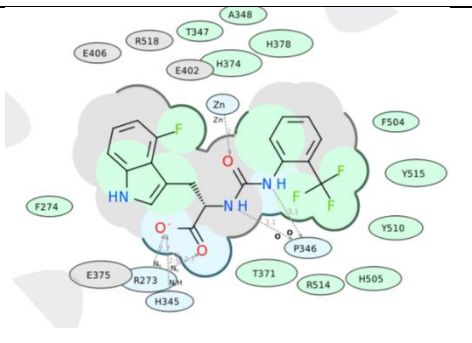
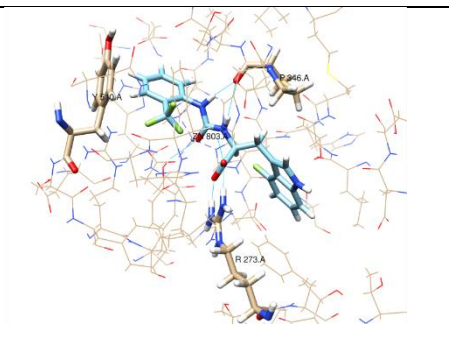
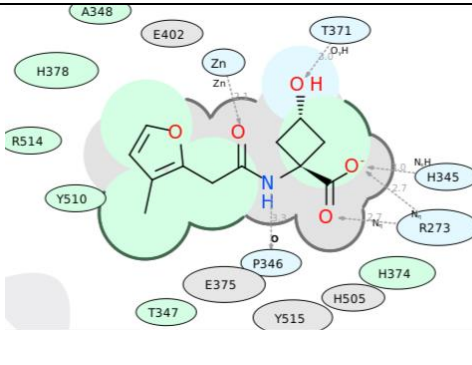
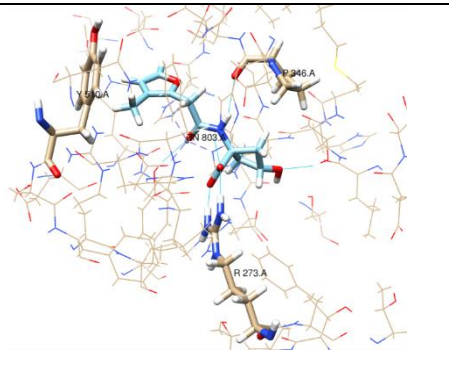
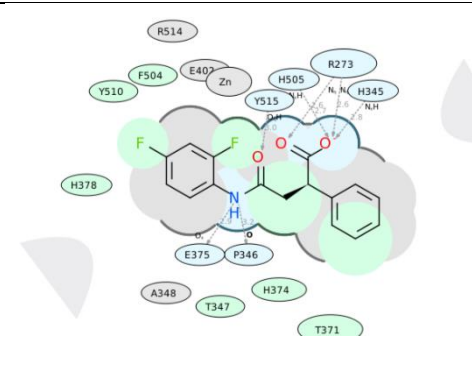
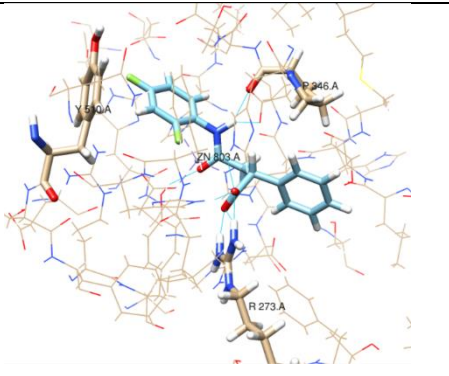
20 структур-лідерів з бази Enamine Ін-Сток колекції, з продемонстрованою візуалізацію розташування у кишені для АПФ2.

#	ID	Score	2d diagram of interaction	3d visualization of interaction
1	Z2831426920	-35,38		
2	Z3681209542	-26,96		
3	Z3681208436	-32,95		
4	Z3681332013	-26,15		

5	Z3969355209	-38,31		
6	Z3681736600	-34,15		
7	Z3681207877	-30,48		
8	Z3681593370	-27,76		

9	Z3681369109	-25,08		
10	Z4188337437	-33,82		
11	Z3037294184	-30,6		
12	Z4549268235	-34,46		

13	Z4361984504	-33.6		
14	Z3681806363	-35,03		
15	Z3681799936	-22,59		
16	Z4602008618	-29		

17	Z3466179106	-26,43		
18	Z4221819990	-40,66		
19	Z4561748167	-32,92		
20	Z4119544425	-34,1		

Перелік сполук, відібраних в додатковому райунді скринінгу Ін-Сток колекції та REAL Space.

А) Сполуки знайдені як зв'язувачі іону цинку:

SMILES ID  
c1ccnc1-c(c2C)cccc2C(=O)NC[C@](C)(C([O-])=O)C(=O)O Z4879191050  
[O-]C(=O)[C@](C)(C(=O)O)CNC(=O)c1ncnc(c12)CCC2 Z4900258160  
O=C([O-])C1(C(=O)[O-])C[C@H](O1)C[N@H+](C)[C@H](C)c2cccc(c2)N(S3(=O)=O)CCC3  
s\_270122\_17677636\_22563296  
O=C([O-])C(C)(C(=O)[O-])C[NH2+][C@@H](C)C1CN(C1)C(=O)OCc2ccccc2  
s\_274370\_25796444\_25885648  
Cc1noc(n1)[C@H]2C[C@H](C2)[NH2+][C@H](C)(C(=O)[O-])C(=O)[O-]  
s\_274370\_25796444\_14032724  
O=C([O-])C(C)(C(=O)[O-])CNC(=O)Cn1cc(en1)P(=O)(C)C m\_188690\_18313080\_21804122  
COC(OC)Cn1cc(nn1)CCCC(C)(C(=O)[O-])C(=O)[O-] m\_282512\_24323246\_24203238  
CS(=O)(=O)c(cc1)ccc1[C@H](C)C[NH2+][C@H](C)C(C)(C(=O)[O-])C(=O)[O-]  
s\_274370\_25845300\_13089860  
Cc1cccc(c1)-c(mn2)n(c23)C[C@H](CC3)C(=O)NCC(C)(C(=O)[O-])C(=O)[O-]  
s\_188690\_18313080\_21724816  
CC(C)(C)OC(=O)NC[C@H]1CC[C@H](C1)[NH2+][C@H](C)(C(=O)[O-])C(=O)[O-]  
m\_274370\_25796444\_25884786  
CC(=O)N[C@H]1CC[C@@H](CC1)[NH2+][C@@H](C)C(F)(C(=O)[O-])C(=O)[O-]  
m\_274370\_25845296\_12192406  
CNC(=O)N1CC[NH+](CC1)C[C@H](O2)CC2(C(=O)[O-])C(=O)[O-] m\_270122\_22550990\_22563296  
O=C([O-])C(F)(C(=O)[O-])[C@H](C)[NH2+][C1cccc(c1)OCCn2ccnc2  
s\_274370\_25845296\_25891064  
Br(c1)cccc1-n(nc2C)cc2C[NH2+][C@H](C)(C(=O)[O-])C(=O)[O-] s\_274370\_25796444\_13089450  
O=C([O-])C(F)(C(=O)[O-])[C@@H](C)[NH2+][C@H](CC1(C)C)[C@@H]1OCc2ccccc2  
s\_274370\_25845296\_25886638  
O=C([O-])C(F)(C(=O)[O-])[C@H](C)[NH2+][C1cccc(c(F)c1)-n2ccnc2 s\_274370\_25845296\_25892716  
CCC(C(=O)[O-])(C(=O)[O-])C[NH2+][C1cccc(c1)-c2[nH]ncc2 s\_274370\_25796844\_25890190  
Cc1cc(CO)cc(n1)C[NH2+][C@@H](C)C(C)(C(=O)[O-])C(=O)[O-] m\_274370\_25845300\_25893416  
CCC(C(=O)[O-])(C(=O)[O-])C[NH2+][C1ccc(S(=O)(=O)C)c(o1)C s\_274370\_25796844\_25892102  
c1cccc([n+][I][O-])C(=O)NNC(=O)COc2c(cccc2)-c3cccc3 Z454147888  
OCc(c1)nc(nc1C)N[N-]C(=O)C(=O)Nc2nn(cc2)Cc3cccc(c34)cccn4 Z4898005905  
c1cccc1[C@@H](C([O-])=O)CC(=O)N[N-]C(=O)c([nH]n2)c(c23)CCCC3 Z232396430  
c1cccc([n+][I][O-])C(=O)N[N-]C(=O)Cc2cccc(c23)cccc3 Z753354148  
c1cccc(c1C)C(=O)[N-]NC(=O)C[C@H](C([O-])=O)c2ccccc2 Z169872436  
[O-]C(=O)[C@H]1CC=CC[C@H]1C(=O)N[N-]C(=O)c2cccs2 Z169884988  
C1OCC[C@@H]1C[C@H]([NH3+])C(=O)N[N-]C(=O)c2c([O-])ccc(Cl)c2 Z9109944436  
c1cccc1C[C@@H]([NH+](C)C)C(=O)[N-]NC(=O)c2c([O-])ccc2 Z1143406678  
c1cccc1[C@](C)(C([O-])=O)CC(=O)[N-]NC(=O)c2cccs2 Z169884984  
Cc1c(Br)ccc(c1)C(=O)[N-]NC(=O)/C=C\C([O-])=O Z57413397  
COc1cc(c(Cl)cc1)N[N-]C(=O)c2[n+](c2)cccc2 Z1285177106  
N#CCC(=O)N[N-]C(=NS1(=O)=O)c(c12)cccc2 Z56789320  
CCc1c(C)cc(s1)C(=O)N[N-]C(=O)[C@H]2CC=CC[C@@H]2C([O-])=O Z169903020  
C1C=CCC1(C([O-])=O)C(=O)N[N-]C(=O)c2c(cccc2)-n3cccc3 Z6480536693  
C1COCCN1c(c2[N+](c2)O)ccc(c2)C(=O)OCC(=O)[N-]NC(=O)c3cccs3 Z19692561  
C1OCC[C@@H]1C[C@H](N)C(=O)N[N-]C(=O)c2c([O-])ccc(Cl)c2 Z9109944436  
c1cccc([n+][I][O-])C(=O)NNC(=O)c(c2)sc(c23)CCCCC3 Z461843780  
Cc1cc(c(o1)C)C(=O)NNC(=O)c2[n+](c2)cccc2 Z425215446  
[O-][n+]1cc(ccc1)C(=O)[N-]NC(=O)c2ccc(cc2)CN(C3=O)c(c4c35)cccc4ccc5 Z27150422  
CC(C)Oc1ccc(cc1)C(=O)N[N-]C(=O)[C@@H]([C@H]23)[C@H](C([O-])=O)[C@@H](C2)C=C3  
Z221345228  
c1cccc(c12)c(c[nH]2)C[C@H](NC(=O)C)C(=O)[N-]NC(=O)c3c([O-])cccc3 Z71625791  
C[C@H]1C[C@@H](C)CN(C1)S(=O)(=O)c2ccc(cc2)C(=O)[N-]NC(=O)c3ccc(cc3)COc4c(cccc4)-c5cccc5  
Z27164873

[O-]C(=O)[C@H]1CC=CC[C@H]1C(=O)[N-]Nc2ncnc(c23)cccc3 Z57036690  
c1cc(=O)[nH]nc1C(=O)[N-]NC(=O)c2ccc(cc2)COc3c(cccc3)-c4cccc4 Z27163579  
[O-]C(=O)[C@@H]1CC=CC[C@H]1C(=O)N[N-]C(=O)c2c(F)cccc2 Z169873290  
c1cccc(c12)n(c(=O)cn2)CC(=O)[N-]NC(=O)c3ccc(cc3)COc(c4)ccc(c45)cccc5 Z27101710  
C1C[C@H](C)Cc(c12)cc(s2)C(=O)N[N-]C(=O)[C@@H]([C@@H]34)[C@H](C([O-])=O)[C@H](C3)C=C4  
Z221345060  
c1cccc(c12)n(C)nc2NC(=O)N[N-]C(=O)c3cnc(N)cc3 Z8864996618  
OCc(c1)nc(nc1C)N[N-]C(=O)C(=O)Nc2cc(Cl)c(cc2)OCc3cc(C#N)ccc3 Z4898005906  
Cc1nc(cs1)-c2c[nH]c(c2)C(=O)N[N-]C(=O)c3c([O-])cccc3 Z116548836  
[O-]C(=O)CCC(=O)N[N-]C(=O)COc1c(Cl)cc(Cl)cc1 Z57201901  
[O-]C(=O)[C@H]1CC=CC[C@H]1C(=O)N[N-]C(=O)c2ccc(F)cc2 Z169899940  
c1cc(C(F)(F)F)cc(Cl)c1N[C@H](C(C)C)C(=O)[N-]NC(=O)c2ccc(cc2)CN(C3=O)c(c4c35)cccc4ccc5 Z27150766  
Clc(c1)cc(Cl)cc1[C@H](N(CC)CC)C(=O)[N-]NC(=O)C(=O)NC2CC2 Z1437045818  
c1cccc(c1[O-])C(=O)N[N-]C(=O)c2c[n+](c1[O-])ccc2 Z169959104  
c1cccc1Cc(cccc2)c2OCC(=O)[N-]NC(=O)c3ccc(cc3)COc4c(cccc4)-c5cccc5 Z27165505  
o1nc(C)c(c1C)COc2ccc(cc2)C(=O)[N-]NC(=O)c3c([O-])cccc3 Z68347958  
c1cccc(Cl)c1C(=O)N[C@H](CCSC)C(=O)[N-]NC(=O)c2c([O-])cccc2 Z68348137  
[O-]C(=O)[C@H]([C@@H]12)[C@H]([C@H](C1)C=C2)C(=O)[N-]NC(=O)c3ccc(F)cc3 Z221344984  
COc1cc[nH+]c(c1[O-])C(=O)N[N-]C(=O)c2ccc(F)cc2Cl s\_22\_675730\_15065692  
CSelcccc1NC(=O)N[N-]C(=O)c2c[n+](c1[O-])ccc2 s\_2430\_11687504\_22496164  
CCc1cc(=O)[nH]c(n1)N[N-]C(=O)C(=O)Nc2ncc(s2)Cc3cccc3Cl s\_2718\_16954944\_8260734  
c1cc[n+](c1[O-])cc1C(=O)[N-]NC(=O)[C@@]23[C@H](COC3)[NH2+]CCC2 |&1;.6,11|  
Felcc(Cl)ccc1Cc(c2C)sc(n2)NC(=O)C(=O)[N-]Nc3nc(C)c(C)c(=O)[nH]3 s\_2718\_16947538\_11675860  
Cc1cc(CO)nc(n1)N[N-]C(=O)C(=O)Nc2ccc(cc2)O[C@H](C)c3cccc3 s\_2718\_16905032\_14523036  
c1cccc1-c(n[nH]2)cc2C(=O)[N-]NC(=O)C(=O)Nc3ccc(cc3)Sc4cccc4Cl s\_2718\_8797748\_2009410  
Clc(c1)c(Cl)ccc1-n2nc(cc2C)NC(=O)C(=O)[N-]NC(=O)Cc3n[nH]cc3 s\_2718\_8798194\_8332292  
Clc1cc(Cl)ccc1C(=O)[N-]NC(=O)[C@@]2(C(=O)[O-])CC[NH2+]C2 s\_240690\_21564624\_16783964  
Cc1ccc(cc1)C(=O)[N-]NC(=O)C[N@@H+](C2)[C@H](C(=O)[O-])c(c23)enn3C  
s\_270122\_22542064\_22570700  
Cc1ccc(c1C)c1C(=O)N[N-]C(=O)c2c(c(=O)[O-])ccc23 s\_58\_97882\_20871682  
Clc1cccc1C(=O)[N-]NC(=O)[C@H]2[C@@H](C(=O)[O-])[C@@H]2C(=O)[O-]  
s\_188690\_21785974\_19714506  
Cc1ccc(cc1)C[NH+]2CCN(CC2)CC(=O)[N-]NC(=O)c3c([O-])c(C)n[nH]3 s\_22\_7288444\_7632488  
O1COc(c12)cc(cc2)C(=O)[N-]NC(=O)[C@@]3(Br)CC[C@H]3C(=O)[O-] s\_58\_97744\_23378788  
c1cc(Cl)ccc1CNC(=O)C(=O)N[N-]C(=O)[C@@]2(C(=O)[O-])CCCCO2 s\_188690\_21784786\_17802836  
C1CC=C[C@]1(C(=O)[O-])C(=O)N[N-]C(=O)CSec2cccc2O s\_188690\_21787902\_21600142  
CC(C)[C@H](C)Nc1cccc(c1C)C(=O)N[N-]C(=O)c2cc[n+](c1[O-])cc2 s\_527\_21811732\_5726572  
c1cccc(c12)[C@](C(=O)[O-])(CC2)CC(=O)N[N-]C(=O)[C@@H](CO3)Oc(c34)cccc4  
s\_58\_97747\_12864554  
[O-]NC(=O)C(C)(C)NC(=O)/C=C/c1cccc(c1)OCc2cccc2 s\_22\_20976584\_215157  
[O-]NC(=O)CCNC(=O)[C@]1(OC)CCN(C1)C(=O)/C(C)=C/c2c[nH]nc2  
m\_274552\_23520110\_24995818\_25056620  
Cc1cc(n[nH]1)CCc2cn(nn2)-c3ccc(C(=O)N[O-])c([O-])c3 m\_282512\_24314996\_24221394  
[O-]NC(=O)c1ccc(cc1[O-])n2cc(nn2)C[C@H](C(=O)[O-])C3CC3 m\_282512\_24314996\_24221856  
ONC(=O)c1c(F)c([O-])c(F)c(F)c1 Z321037342  
[O-]NC(=O)c1c[n-]c(=O)[nH]c1=O Z203401972  
[O-]NC(=O)CN(C1=O)C(=O)[N-][C@]1(C)c(c2)ccc(c23)cccc3 Z59183341  
[O-]NC(=O)c(c1[O-])ccc(c12)cccc2 Z244466702  
[O-]NC(=O)c(c([O-])c1)cc(c12)cccc2 Z90125059  
CC(C)(C)OC(=O)Nc1ccc(cc1)-c2cc(no2)C(=O)NCCCCC(=O)N[O-] Z2739897696  
CSCC[C@@H](C(=O)N[O-])NC(=O)C[N@H+]1CCC[C@@H]1Cn2cccn2 Z8716415636  
[O-]NC(=O)[C@@H]1[C@H]([C@H]12)c3c(C2)cc(Cl)cc3 Z2724723409  
c1cccc(c12)N(C[C@H](C2)C(=O)N)Cc3cn(nn3)CCP(=O)([O-])O  
m\_274952\_12993010\_12611784\_26213330  
Cc1c(C(=O)[O-])ccc1-c2cn(nn2)CCCP(=O)([O-])[O-] s\_273910\_21676774\_25422098  
O=C([O-])c1cc(F)cc(c1)-c2cn(nn2)CCCP(=O)([O-])[O-] s\_273910\_21676774\_25422088  
c1cccc1[C@@H](CC(=O)[O-])c2cn(nn2)CCP(=O)([O-])[O-] m\_273910\_21676896\_25421604  
n1cccc1[C@]2(C#N)CCCN(C2)Cc3cn(nn3)CCP(=O)([O-])[O-]  
m\_274952\_13006666\_12611784\_26213330  
O=C([O-])c1cccc1OCc2cn(nn2)CCP(=O)([O-])[O-] s\_273910\_21676896\_25807160  
c1cccc(c12)c(C(=O)[O-])cn2Cc3cn(nn3)CCP(=O)([O-])[O-] s\_273910\_21676896\_25421742  
Cc1cc(C(=O)[O-])ccc1-c2cn(nn2)CCP(=O)([O-])O s\_273910\_21676896\_25421892

Cc1c(C(=O)[O-])cccc1-c2cn(nn2)CCP(=O)([O-])[O-] s\_273910 21676896 25422098  
 O=C([O-])c1cccc(c1F)-c2cn(nn2)CCP(=O)([O-])[O-]s\_273910 21676896 25422094  
 Cc1c(Cl)cccc1CN(CC(=O)[O-])Cc2cn(nn2)CCP(=O)([O-])[O-] s\_273910 21676896 25422560  
 Clc1c(Cl)cccc1[C@@H](CO)N(C)Cc2cn(nn2)CCCP(=O)([O-])[O-]  
 m\_274952 13006012 12611784 26213238  
 O=P([O-])([O-])CCn1cc(nn1)C[N@@H+]2C[C@H](O)C[C@H]2c3cccc3  
 m\_274952 13004060 12611784 26213330  
 [O-]P(=O)(O)[C@@](O)(P([O-])([O-])=O)CC[NH2+][C@H](SC)Nc1cccc1 Z56754446  
 O=P([O-])(O)CCn1cc(nn1)COc2cccc2F s\_273910 21676896 11992710  
 O=C([O-])c1ccc(cc1)OCc2cn(nn2)CCP(=O)([O-])[O-] s\_273910 21676896 11994356  
 N#CCCC1(OCCO1)CCc2cn(nn2)CCP(=O)([O-])[O-] s\_273910 21676896 14590144  
 c1cccc(c12)n(c(C(=O)[O-])c2)Cc3cn(nn3)CCCP(=O)([O-])[O-] s\_273910 21676774 25421602  
 c1cccc1C(C)(c2cccc2)NCc3cn(nn3)CCP(=O)([O-])[O-] m\_274952 12996114 12611784 26213330  
 COc1c(SC)ccc(c1)CN(C)Cc2cn(nn2)CCCP([O-])(=O)O Z8821317264  
 C1CCC1(C(=O)[O-])c2cn(nn2)CCP(=O)([O-])[O-] m\_273910 21676896 21505024  
 [O-][n+]1cccc1SCC(=O)N[C@H](CCC)c2cn(nn2)-c3cccc3-c4cccc4  
 m\_276010 17281718 17265472 24752038  
 CCc1c[nH]nc1C(=O)N(C)Cc2ccc(cc2)CNC(=O)c3cccc[n+][3][O-]  
 m\_275592 15865082 15840532 19065872  
 Cc1cccc1C[C@H](C(=O)[O-])NC(=O)c2cc(Cl)cc[n+][2][O-] m\_269946 25345840 22839532  
 COc1cc[n+](c1)[O-]c(c1)C(=O)NCc2c(nco2)Cn3cc(nn3)-c4cn(C(C)C)nc4  
 m\_276010 17281598 17279596 13624758  
 c1n[nH]cc1[C@H](C)CCNC(=O)[C@](C)(C2CCC2)NC(=O)c3cccc[n+][3][O-]  
 m\_274552 23519528 24974668 15982916  
 c1cc(C)cc(c12)sc(c2Cl)C(=O)N(CC[C@H]3C)C[C@H]3NC(=O)CSc4[n+](c1)cccc4 Z7746706793  
 [O-][n+]1cccc1Clcc1C(=O)N[C@@H](C(=O)[O-])Cc2cnco2 m\_269946 23796736 22839532  
 COc1cc([n+](c1)[O-])cc1C(=O)N[C@@H]2C[C@H](C)N(C2)C(=O)[C@@H](O)COc3cccc3 Z5129500870  
 c1cccc([n+][1][O-])C(=O)NCC2CCN(CC2)C(=O)c3cc(ccc3)-c4cccc4 Z9110447068  
 c1cc[nH]c(c1)[O-]=N/C(=O)c2[n+](c1)cccc2 Z459590058  
 Cc1c(C#N)[n-]cc1C(=O)NCc2ccc(c(F)c2)CNC(=O)c3cccc[n+][3][O-]  
 m\_275592 14115244 15840532 25327220  
 Oc1cc([O-])c(cc1)/C=N/[N-]C(=O)c2c[n+](c1)cccc2 Z56774426  
 [O-][n+]1cccc1NC(=O)CCc2nncc2[O-] s\_22 20813372 23005286  
 c1cccc(c1)[O-]C(=O)NNC(=O)c2[n+](c1)cccc2 Z425227852  
 Cc1cc([O-])c(cc1)C(=O)[N-]NC(=O)c2cc[n+](c1)cccc2 Z757433736  
 c1cccc(c12)c(es2)[C@@H](C)n3cc(nn3)C(C)(C)NC(=O)c4cccc[n+][4][O-]  
 m\_276010 17282592 17268416 17280772  
 COc1cc[n+](c1)[O-]c(c1)C(=O)N[C@@H](C(=O)[O-])Cc2ccc(C#N)nc2 s\_269946 26258404 21440630  
 [O-][n+]1cccc1C(=O)NC(C)(C)C#Cc2cn(nn2)[C@@H](CC3)c(c34)cccc4Cl  
 m\_276010 17282460 17268416 24752062  
 COc1cc[n+](c1)[O-]c(c1)C(=O)N[C@@H](C(=O)[O-])Cc2csc(Cl)n2 s\_269946 21814810 21440630  
 Cc1nc(Cl)ccc1Cn2cc(nn2)C#CC(C)(C)NC(=O)c3cccc[n+][3][O-]  
 m\_276010 17281620 17268416 24752062  
 O=c1[n-]snc1C(=O)Nc2cccc[n+][2][O-] s\_22 20813372 11730510  
 FC(F)(F)C[C@H](C(=O)[O-])NC(=O)c1cccc[n+][1][O-] s\_269946 16809500 7302472  
 [O-][n+]1cccc1C(=O)NCCC[C@@H]2CN(CCO2)C(=O)c3ccc(cc3)CC4CC4  
 m\_275592 18896730 15840532 15864406  
 CSC1(CC1)CNC(=O)Nc2cccc[n+][2][O-] s\_2430 22525722 9448358  
 [O-][n+]1cccc1C(=O)NCCCc2cn(nn2)Cc3ccc(Cl)cc3Cl m\_276010 17281438 17268416 17281140  
 c1cccc([n+][1][O-])C(=O)N[C@@H]2CN(C(=O)C2)[C@@H]3C[C@H]3c4cccc4 Z1275889388  
 C1[C@@H](C2)C[C@H]3[C@@H]2[C@@]13NC(=O)Nc4cccc[n+][4][O-] m\_2430 22525722 17843074  
 CCc(c1)ccc(c12)cc([nH]2)C(=O)N3CC[C@@H]([C@H](C)C3)CNC(=O)c4[n+](c1)cccc4 Z9224597611  
 c1cccc([n+][1][O-])C(=O)[N-][C@@H](c2nnn[n-]2)c3cc(F)c(Cl)cc3 Z1713510693  
 C1OCCC12[C@H](CCC2)NC(=O)Nc3cccc[n+][3][O-] s\_2430 22525722 9820720  
 C1OCC[C@]1(C(=O)[O-])NC(=O)c2cccc[n+][2][O-] s\_269946 11987984 7302472  
 CC[C@@H](C)N(CC([O-])=O)C(=O)c1[n+](c1)cccc1 Z1497304817  
 CS1c(NC(=O)C)ccc(c1)C(=O)CSc2[n+](c1)cccc2 Z448102166  
 COc1cc[n+](c1)[O-]c(c1)C(=O)NCc2cc(P(=O)(C)C)nc2 m\_11 22181110 22643102  
 Cc1cc(c(Br)o1)C(=O)Nc2cccc[n+][2][O-] s\_11 22175572 23796740  
 [O-][n+]1cccc1NC(=O)NCCc2c(cc23)cccc3 s\_2430 22525722 741724  
 CC(C)=C(Br)CNC(=O)Nc1cccc[n+][1][O-] s\_2430 22525722 15074884  
 NC(=O)c(c1)c([O-])ccc1[C@@H](C)NC(=O)Nc2cccc[n+][2][O-] s\_2430 22525722 20817218

c1csccl[C@H](C)C(=O)Nc2cccc[n+][O-] s\_11 22175572 5350584  
Cc1nsc(Cl)c1C(=O)NCC[C@@H]2CC[C@@H](C2)NC(=O)c3cccc[n+][O-]  
m\_275592 18896748 15840532 25328916  
[O-][n+]1cccc1C(=O)N[C@@H]2CCC[C@@H](CC2)C(=O)NCc3nc(F)ccc3  
m\_274552 23520486 15955794 15982916  
CC(C)n1nnnc1[C@@H](c(ccc2)c(c23)oc(=O)[nH]3)NCc4c[n+][O-]ccc4  
m\_281630 20798284 21629548 20772828  
CCn1nnnc1[C@@H](c(nc2)n(c23)cccc3)NCc4c[n+][O-]ccc4  
m\_281630 20798284 21623250 20648542  
O=C([O-])c1cc(ccc1O)/C=C/c(cc[n+][O-])c(c23)cccc3 s\_6 15608778 9043926  
O=C([O-])c1cc(ccc1O)[C@H]2CC[N@@H+](CC2)Cc3cc(O)cc(c34)cccc4  
s\_274370 25858292 25889894  
O=C([O-])c1cc(ccc1O)[C@H]2CC[N@@H+](CC2)Cc3cc(CO)cc(c34)cccc4  
s\_274370 25858292 25893658  
O=C([O-])c1cc(ccc1O)[C@H]2CC[N@@H+](CC2)Cc3cc(Br)cc(c34)[nH]nc4  
s\_274370 25858292 25888956  
O=C([O-])c(c(O)c1)ccc1N[C@H](C)COc(cc2)cc(c23)ccc(Br)c3 s\_274370 25905060 13089394  
O=C([O-])c1cc(c(O)c(C(=O)[O-])c1)NCc2ccc(o2)-c3cccc3Cl m\_271304 16785080 9304784  
NC(=O)c1cc(ccc1[O-])C(=O)/C=C/c2ccc(C(=O)[O-])c(O)c2 s\_6 20499924 13207838  
C1CCCC1C(=O)C(\C#N)=C/c2ccc(C(=O)[O-])c(O)c2 s\_6 7345454 13207838  
c1cccc1[C@H](C)[C@H](CO)NC(=O)c2c(O)c(C(=O)[O-])ccc2 s\_188690 21785168 20761338  
OCCc1nc(s1)/C=C/c2cc(O)c(C([O-])=O)cc2 Z5557939342  
O=C([O-])c1cccc(c1O)C(=O)N[C@@H](CO)c2cccc2Br s\_188690 25719116 20761338  
O=C([O-])c1cccc(c1O)C(=O)N(C2)C[C@H](CO)C23CCC3 s\_188690 14657228 20761338  
[O-]S(=O)(=O)CC(=O)Nc1cccc1 Z4081408523  
CC[C@H](S([O-])(=O)=O)C(=O)Nc1cccc(c12)C(=O)c3c(C2=O)cccc3 Z56777340  
[O-]S(=O)(=O)CCC[n+](c(c12)cccc2)c(s1)/C=C/c3ccc(o3)-c4ccc(SC(F)F)cc4 Z56782166  
c1csc(c12)c3c(cc2OC)sc([n+][O-])3CCCS([O-])(=O)=O)/C=C/c4ccc(C#N)cc4 Z46025302  
[O-]S(=O)(=O)C[C@H](C([O-])=O)[NH2+]Cc1ccc(cc1)Oc2cc(Br)ccc2 Z9080742184  
[O-]S(=O)(=O)CCC[n+](c(c12)cccc2)c(s1)/C=C/c3cn(nc3-c4cccs4)-c5cccc5 Z56766107  
CCc1ccc(cc1)/C=C/c([n+][O-])2CCCS([O-])(=O)=O)sc(c23)cc(OC)c4c3ccc4 Z46025312  
[O-]S(=O)(=O)C(F)(F)c(n1)[nH]c(c12)cccc2 Z56796614  
[O-]S(=O)(=O)CCCN(c12)cccc2/C(S1)=C/c([n+][O-])3CCCS([O-])(=O)=O)sc(c34)cccc5c4cccc5 Z2831484117  
c1csc(c12)c3c(cc2OC)sc([n+][O-])3CCCS([O-])(=O)=O)/C=C/c4ccc(Br)cc4 Z46025244  
O=S(=O)([O-])CC[C@H](C(=O)[O-])NC(=O)/C=C/Sc1cccc1 m\_272610 22846852 18218596  
O=S(=O)([O-])C[C@@H](C(=O)[O-])NC(=O)c1ccc(Cl)cc1Cl m\_272610 19635080 17767432  
O=S(=O)([O-])CC[C@H](C(=O)[O-])NC(=O)[C@H](C1)Oc(c12)cc(F)cc2  
m\_272610 22846852 18318212  
O=S(=O)([O-])C[C@H](C(=O)[O-])[NH2+]Cc(cc1)cc(c12)cco2 m\_271304 19635078 16775762  
O=S(=O)([O-])CC[C@H](C(=O)[O-])NC(=O)Cc1nn(cc1)-c2ccc(F)cc2m\_272610 22846852 17773044  
O1CC(C)C)c(c12)cccc2C(=O)N[C@@H](C(=O)[O-])CS(=O)(=O)[O-]  
m\_272610 19635080 17773128  
CC[NH+](CC)Cc1ccc(o1)C(=O)N[C@@H](C(=O)[O-])CS(=O)(=O)[O-]m\_272610 19635080 17771534  
O=S(=O)([O-])C[C@H](C(=O)[O-])NC(=O)c(c1n)n(c12)CCCC2 m\_272610 19635080 17774424  
O=S(=O)([O-])C[C@@H](C(=O)[O-])NC(=O)CCn1cc[nH+]c1 m\_272610 19635080 17771552  
O=S(=O)([O-])C[C@H](C(=O)[O-])NC(=O)c(ccc1OC)cc1-c2cccc2 m\_272610 19635080 18210158  
O=S(=O)([O-])CC[C@H](C(=O)[O-])NC(=O)[C@H](C1)c(c12)cccc2 m\_272610 22846852 17770202  
C[NH+](C)CCN(CC[S-])C(=O)C(=O)N(CC1)Cc(c12)cc(N)cc2 s\_273498 20865610 14586814  
C[NH+](C)CCN(CC[S-])C(=O)C(=O)N1CCO[C@H](C1)Cn2cnen2 s\_271949 20865584 24395212  
CCOC(=O)[C@@H](C[S-])[NH2+]C[C@H](O)COc(cc1)cc(c12)CCC(=O)N2  
s\_273610 20326850 11516418  
[S-]C[C@@H](C1)OCCN1C(=O)C(=O)N(C)CCc2cccc2C#N m\_271949 16728660 20608418  
O=S(=O)(C1)C[C@H](C2)N(C[C@@H]12)C(=O)C(=O)N3CCO[C@@H](C[S-])C3  
m\_271949 20608416 21050494  
c1ccc(C)c(c1C)N(C(C)C)C(=O)CNC(=O)C(=O)N2CCO[C@@H](C[S-])C2  
m\_271948 14480030 20608414  
NC(=O)c1cnn(c12)c(C)c(c(n2)C)CCC(=O)N3CCO[C@@H](C[S-])C3 m\_527 20608398 859882  
[S-]C[C@@H](C1)OCCN1C(=O)C(=O)N2CC[C@@H](C2)Sc3cc(Cl)ccc3  
m\_271949 20608416 10847760  
CCOC(=O)[C@@H](C[S-])[NH2+]C[C@H](O)COc(cc1)cc(c12)OCO2  
s\_273610 20326850 11515800  
CCOC(=O)[C@@H](C[S-])NC(=O)c(c[nH]1)c(=O)c(c12)encc2 s\_527 20326838 9963286

Б) Сполуки, отримані оптимізацією отриманих хітів:

SMILES	ID
<chem>c1cc(F)ccc1C[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)Cn3cc(nn3)[C@H](O)c4ccc(F)cc4</chem>	m_280190_19288966_18077368_19345860
<chem>CP(=O)(C)c1ccc(cc1)C[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)-c3ccc(F)cc3</chem>	m_188690_25683436_7314346
<chem>Cn1c[nH+]cc1C[C@@H](C(=O)[O-])NC(=O)Cc2csc(n2)-c(cc3)cc(c34)OCO4</chem>	s_188690_27072420_7317202
<chem>c1cc(F)ccc1C[C@@H](C(=O)[O-])NC(=O)CCc2ncc(o2)-c3ccc(F)cc3</chem>	s_188690_13157912_5367204
<chem>c1cc(F)ccc1C[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)-c3cn(nn3)[C@H](C(=O)NC4CC4)c5ccc(F)cc5</chem>	m_280190_19288966_18083544_19344500
<chem>O=C([O-])[C@H]([NH3+])CC(=O)N1CCC(CC1)CCc2ccc(O)cc2</chem>	s_240690_7353512_19619820
<chem>CN(C)c1ccc(cc1)-c2en(nn2)CCCC(=O)N[C@H](C(=O)[O-])Cc3ccc(F)cc3</chem>	m_280190_19288966_19326012_19344658
<chem>c1ccccc1C[C@@H](C(=O)[O-])NC(=O)Cc2csc(n2)-c(cc3)cc(c34)OCO4</chem>	s_269946_13424932_7311712
<chem>[nH+]1c[nH]c(c1C)C[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)-c3ccc(F)cc3</chem>	m_272610_18687326_18206866
<chem>COc1ccc(cc1)C[C@@H](C(=O)[O-])NC(=O)Cc2csc(n2)-c(cc3)cc(c34)OCO4</chem>	m_272610_21474256_18210248
<chem>Cn1ncc1C[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)-c3ccc(F)cc3</chem>	s_188690_27072420_7314346
<chem>Cc1ccccc1C[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)-c3ccc(F)cc3</chem>	m_272610_17766350_18206866
<chem>[nH+]1c[nH]c(c1C)C[C@@H](C(=O)[O-])NC(=O)Cc2csc(n2)-c(cc3)cc(c34)OCO4</chem>	m_272610_18687326_18210248
<chem>CCc1ccc(cc1)-c2ccc(o2)C(=O)N[C@H](C(=O)[O-])Cc3ccc(C)cc3</chem>	m_272610_21474274_17773272
<chem>c1cc(F)ccc1C[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)-c3ccccc3Cl</chem>	s_188690_13157912_7314350
<chem>Cn1c[nH+]cc1C[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)-c3ccc(F)cc3</chem>	s_188690_22386314_7314346
<chem>c1cc(C)cc(c12)sc(n2)-c3ccc(cc3)NC(=O)CN[C@H](C(=O)[O-])Cc4cncn4C</chem>	s_270122_22386318_22572466
<chem>O=C([O-])C[C@H](C(=O)[O-])NC(=O)c1csc(n1)-c(c2)ccc(c23)OCO3</chem>	Z4859725735
<chem>O=C([O-])[C@@H]1CCC[N@@H+]1CC(=O)[N-]Cc2ccc(cc2)C[NH+]3CCCC3</chem>	Z1459911788
<chem>c1cc(F)ccc1[C@@H](CC(=O)[O-])NC(=O)c2cccc(c23)cn[nH]3</chem>	Z929744756
<chem>CC[C@H](C(=O)[O-])[S@](=O)Cc1csc(n1)-c(c2)ccc(c23)OCO3</chem>	Z2301748685
<chem>o1ccccc1C[N-]C(=O)[C@H]2CCC[N@@H+]2CC(=O)N3CCC(CC3)Cc4ccccc4</chem>	Z335367154
<chem>c1ccccc1[C@@H](P(=O)(C)C)NC(=O)c2c([O-])ccc[nH+]2</chem>	Z9192436281
<chem>O=C([O-])C[C@H](C)NC(=O)c1ccc(o1)-c2ccc(F)cc2</chem>	Z85972354
<chem>c1cc(F)ccc1[C@@H](CC(=O)[O-])NC(=O)c(c2)ccc(c23)sen3</chem>	Z6256863474
<chem>c1ccccc1N2N=C(C[C@H]2C(=O)N)C(=O)/N=c3c([O-])ccc[nH]3</chem>	Z332713566
<chem>C1OCCC[C@H]1[C@@H](C(=O)[O-])NC(=O)Cc2csc(n2)-c(c3)ccc(c34)OCO4</chem>	Z1870348262
<chem>Cc(c1)ccc(c12)nc(s2)-c3ccc(cc3)NC(=O)[C@H]4CCCC[C@H]4C(=O)[O-]</chem>	Z57248568
<chem>c1ccccc1[C@H](C[S@](=O)C)NC(=O)c2c([O-])ccc[nH+]2</chem>	Z1688557789
<chem>Cc(c1)ccc(c12)nc(s2)[N-]C(=O)CCC(=O)[O-]</chem>	Z57598631
<chem>c1cc(F)ccc1[C@H](CC(=O)[O-])NC(=O)c(c2)ccc(c23)OCCO3</chem>	Z85905321
<chem>Cc1ccc(cc1)[C@@H](C2CC2)[N-]C(=O)C[N@H+]3CCC[C@H]3C(=O)[O-]</chem>	Z1459912106
<chem>c1cc(F)ccc1[C@@H](CC(=O)[O-])NC(=O)[C@@H](C2)Oc(c23)c(F)ccc3</chem>	Z1443671099
<chem>c1ccccc12[C@@H](C[N@@H+](C)C2)[N-]C(=O)c3c([O-])ccc[nH+]3</chem>	Z8546707139
<chem>Cc(c1)ccc(c12)nc(s2)-c3ccc(cc3)NC(=O)c4c(C(=O)[O-])c([N+](=O)[O-])ccc4</chem>	Z283604232
<chem>c1ccccc1[C@@H](CC(=O)[O-])NC(=O)c2ccc(F)cc2</chem>	Z98660411
<chem>c1cc(F)ccc1[C@@H](CC(=O)[O-])NC(=O)c2cnc2</chem>	Z1238812144
<chem>O=C([O-])c1nnc([n-]c1=O)SCc2csc(n2)-c(cc3)cc(c34)OCO4</chem>	s_7_14601776_2757950
<chem>[O-]c1ccc[nH+]c1C(=O)NC[C@H](C(=O)[O-])c2ccccc2</chem>	s_188690_7326850_25747122
<chem>[O-]c1ccc[nH+]c1C(=O)N[C@H](C2)C[C@]2(C(=O)[O-])c3ccccc3</chem>	s_188690_21998030_25747122
<chem>c1ccccc1-n(nc2C)c(n2)[N-]C(=O)CNC(=O)c3[nH+]ccc3[O-]</chem>	s_527_6821762_19450810
<chem>c1cc(C)cc(c12)sc(n2)[N-]C(=O)C[N@@H+](CCC(=O)[O-])Cc3ccccc(Br)c3</chem>	s_270122_11726754_22570398
<chem>[O-]c1ccc[nH+]c1C(=O)N[C@H](C[C@@H]2C(=O)[O-])c(c23)ccccc3</chem>	s_188690_12902118_25747122
<chem>c1cc(F)ccc1[C@H](CC(=O)[O-])NC(=O)c(c2[O-])ccc(c23)ccccc3</chem>	s_188690_2735052_16887472
<chem>c1ccccc1-n(nc2)cc2C(=O)N[N-]C(=O)c3[nH+]ccc3[O-]</chem>	s_527_4839312_2559366
<chem>Cc1ccc(cc1)CCC(=O)N2CCC(CC2)CCNC(=O)c3c(C)[nH]c[n+]3[O-]</chem>	m_275592_14114834_14139410_19059808
<chem>CN(C)c(cc1)ccc1[C@H](C)NC(=O)CC[C@@H](C(=O)[O-])NC(=O)OCc2ccccc2</chem>	s_269946_21823040_21548810

COc1ccc(cc1)[C@H](O)[C@@H](C(=O)[O-])NC(=O)CCc2ccc(N(C)C)cc2  
s\_188690\_\_\_21756120\_\_\_13231950  
O=C([O-])[C@@]1(O)CC[N@H+](C[C@@H]1O)Cc2csc(n2)-c(cc3)cc(c34)OCO4  
s\_270122\_\_\_18341848\_\_\_11728256  
CCc1ccc(cc1)-c2cn(nn2)Cc3ccc(o3)C(=O)N[C@H](CC(=O)[O-])c4ccc(SC)cc4  
m\_280190\_\_\_19274204\_\_\_18077368\_\_\_19344348  
Cn1c[nH+]c(c1)C[C@H](C(=O)[O-])NC(=O)Cc2csc(n2)-c(cc3)cc(c34)OCO4  
m\_272610\_\_\_17766912\_\_\_18210248  
s1cccc1C(=O)C[C@@H](C(=O)[O-])NC(=O)Cc2csc(n2)-c(cc3)cc(c34)OCO4  
m\_272610\_\_\_17766330\_\_\_18210248  
Cn1cc[nH+]c1C[C@@H](C(=O)[O-])[NH2+]Cc2csc(n2)-c(cc3)cc(c34)OCO4  
s\_274370\_\_\_25845108\_\_\_13088668  
O=C([O-])C1(C(=O)[O-])CC[NH+](CC1)Cc2csc(n2)-c(cc3)cc(c34)OCO4  
s\_270122\_\_\_13587202\_\_\_11728256  
O=S(=O)([O-])C[C@H](C(=O)[O-])NC(=O)Cc1csc(n1)-c(cc2)cc(c23)OCO3  
m\_272610\_\_\_19635080\_\_\_18210248  
O=C([O-])C(F)(C(=O)[O-])[C@H](C)[NH2+]Cc1csc(n1)-c(cc2)cc(c23)OCO3  
s\_274370\_\_\_25845296\_\_\_13088668  
O=C([O-])[C@H]([NH3+])CC(=O)N1CCC(CC1)CCc2ccc(O)cc2 s\_240690\_\_\_7353512\_\_\_19621912  
Cc1ccc(cc1)CC2CCN(CC2)C(=O)Cn3cc(nn3)C[C@@H](O)C(=O)[O-]  
m\_280190\_\_\_19275734\_\_\_19320720\_\_\_28050116  
COc1ccc(cc1)[C@H](O)[C@@H](C(=O)[O-])NC(=O)c2ccc(o2)-c3cc(F)ccc3  
s\_188690\_\_\_21756120\_\_\_21753562  
c1cc(F)ccc1-c(cc2)ccc2[C@H](C)NC(=O)CCCCn3cc(nn3)C[C@@H](O)C(=O)[O-]  
m\_280190\_\_\_28060258\_\_\_18077364\_\_\_28050116  
O=C([O-])CC[C@H](C(=O)[O-])[NH2+]Cc1csc(n1)-c(cc2)cc(c23)OCO3  
s\_274370\_\_\_25844750\_\_\_13088668  
CCn1cc[nH+]c1[C@@H](C(=O)[O-])NC(=O)Cc2csc(n2)-c(cc3)cc(c34)OCO4 s\_188690\_\_\_21771310\_\_\_7317202  
[O-]c1ccc[nH+]c1C(=O)NCC(=O)NCc2nncn2-c3ccccc3 s\_527\_\_\_636098\_\_\_19450810