

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
імені В. Н. Каразіна

Кафедра хімічного матеріалознавства

УДК 54.057:547.83:547.792

До захисту допускаю



_____ Завідувач кафедри

“___” _____ 2024 р. д.х.н., проф. О.І. Коробов

ПОШУК НОВИХ ІНГІБІТОРІВ ЛІЗИН-СПЕЦИФІЧНОЇ ДЕМЕТИЛАЗИ-1 (LSD1)

Дипломна робота

студентки гр. ХМ-22

хімічного факультету

МОСІНОЇ АЛІНИ

ОЛЕКСАНДРІВНИ

Науковий керівник

д.х.н., проф.



В. В. Іванов

ХАРКІВ 2024

РЕФЕРАТ

Кваліфікаційна робота містить 35 сторінок, 4 таблиці та 15 рисунків. Використано 23 літературних джерела.

Об'єкт дослідження: лізин специфічна деметилаза LSD1.

Мета роботи: за допомогою докінгу знайти хімічні сполуки, які можуть слугувати в якості інгібітора LSD1.

Результати: проведено пошук інгібіторів LSD1 на основі сканування комерційно доступних баз даних і молекулярного докінгу структур з тієно[3,2-b]піррол-5-карбоксамідним каркасом. Знайдено нові структури які можуть бути перспективними інгібіторами LSD1.

Ключові слова: ДОКІНГ, DOCKING SCORE, ЛІЗИН СПЕЦИФІЧНА ДЕМЕТИЛАЗА (LSD1), ТІЄНО[3,2-В]ПІРРОЛ-5-КАРБОКСАМІД.

ABSTRACT

The qualification work contains 35 pages, 4 tables and 15 figures. 23 literary sources were used.

Object of research: lysine specific demethylase LSD1.

Objective: using docking to find chemical compounds that can serve as an LSD1 inhibitor.

Results: a large-scale search for LSD1 inhibitors was conducted based on a scan of commercially available databases; based on molecular docking, a number of structures with a thieno[3,2-b]pyrrole-5-carboxamide scaffold were found, which may be promising LSD1 inhibitors.

Keywords: DOCKING, DOCKING SCORE, LYSINE SPECIFIC DEMETHYLASE, LSD1, THIENO[3,2-B]PYRROLE-5-CARBOXAMIDE.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ	5
ВСТУП.....	6
1. ЛІТЕРАТУРНИЙ ОГЛЯД.....	7
1.1. Біологічні процеси	7
1.2. LSD1	8
1.3. LSD2 гомолог LSD1	10
1.4. Механізм дії LSD1	11
1.5. Інгібітори LSD1	12
1.6. Програмне забезпечення та методи	14
2 РОЗРАХУНКОВА ЧАСТИНА.....	18
ВИСНОВКИ	33
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	34

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

ADME	(англ. Absorption, Distribution, Metabolism, and Excretion) дескриптори, які кількісно визначають, як певний препарат взаємодіє з організмом і виводиться із організму
Docking Score	оціночна функція докінгу
ECIDAL	(англ. Essential Chemical Interaction Described for Analog Ligands) основні хімічні взаємодії, що описані для аналогових лігандів
<i>in vitro</i>	(лат. «у склі») технологія виконання експерименту поза живого організму
<i>in vivo</i>	(лат. «у (на) живому») технологія виконання експерименту на живому організмі
KDM1A	(англ. lysine (K)-specific demethylase 1A) лізин К-специфічна деметилаза 1A
KDM1B	(англ. lysine (K)-specific demethylase 1B) лізин К-специфічна деметилаза 1B
LSD1	лізин-специфічна деметилаза 1
PAINS	(англ. Pan-assay interference compounds) компоненти, що перешкоджають аналізу — хімічні сполуки, які часто дають хибнопозитивні результати
PLS	(англ. Partial least squares) часткові найменші квадрати
QSAR	(англ. Quantity structure-activity relationships) кількісні зв'язки структура-активність
vHTS	(англ. Virtual high throughput screening) віртуальний високопродуктивний скрінінг

ВСТУП

Лізин-специфічна деметилаза 1 (LSD1) була першою виявленою гістоновою (гістони — різновид білків, які необхідні для формування молекул ДНК у хроматині) деметилазою та членом-засновником сімейства флавін-залежних лізиндеметилаз (KDM1). У людини KDM1 включає KDM1A і KDM1B, які в першу чергу каталізують деметилювання гістону H3K4me1/2. Сімейство KDM1 бере участь в епігенетичній регуляції генів і відіграє важливу роль у різних біологічних процесах та процесах патогенезу захворювань, включаючи диференціювання клітин, ембріональний розвиток, передачу гормональних сигналів та канцерогенез. Порушення роботи багатьох епігенетичних регуляторів призводить до складних захворювань людини, включаючи рак.

Регулятори, такі як KDM1 стали потенційними терапевтичними мішенями через оборотність епігенетичного контролю функції геному. Справді, було розроблено кілька класів селективних низькомолекулярних інгібіторів KDM1, деякі з яких нині проходять клінічні випробування на лікування різних видів раку.

У даній роботі буде розглянуто молекулярні механізми дії, молекулярну структуру та біологію лізиндеметилаз сімейства KDM1. Зосередивши увагу на раку, також буде надано огляд нещодавно розроблених інгібіторів KDM1 та пов'язаних з ними доклінічних та клінічних досліджень, щоб забезпечити краще розуміння механізмів дії та застосування цих специфічних для KDM1 інгібіторів у терапевтичному лікуванні.

1. ЛІТЕРАТУРНИЙ ОГЛЯД

1.1. Біологічні процеси

Зміна ДНК спадкової матерії може суттєво впливати на розвиток пухлин, адже дефекти в генах чи їх експресія можуть сприяти неадекватному регулюванню клітинного росту і подібним процесам, що лежать в основі пухлин [1].

Епігенетична регуляція — це система регулювання генної експресії, яка не вимагає змін в самій послідовності ДНК. Замість цього, епігенетичні механізми контролюють доступність генів для транскрипції, тобто процесу переведення генетичної інформації з ДНК в РНК. Це відбувається за допомогою хімічних змін в хромосомах та їх асоційованих білках, таких як гістони.

Епігенетична регуляція, з іншого боку, відбувається без змін в самій послідовності ДНК, але впливає на те, як гени включаються або виключаються у конкретних клітинах. Епігенетична регуляція грає ключову роль у багатьох процесах, таких як розвиток, різноманітні фізіологічні функції, відповідь на середовищеві стреси та розвиток захворювань, включаючи рак. Існує три класи епігенетичної регуляції: метилювання ДНК, модифікація гістонів і дія некодуючої РНК [2].

Відомо багато видів модифікацій гістонів, таких як метилювання, фосфорилування, ацетилювання та убіквітування.

Метилювання додається до залишків лізину або аргініну гістонів Н3 і Н4 з різним впливом на транскрипцію. Метилювання аргініну сприяє активації транскрипції, тоді як метилювання лізину бере участь як в активації транскрипції, так і в репресії залежно від сайту метилювання. Цю гнучкість можна пояснити тим фактом, що це метилювання не змінює заряд гістону або безпосередньо не впливає на взаємодію гістон-ДНК, на відміну від ацетилювання [3].

Метилювання: гістонметилтрансферази (НМТ):

- Лізин:
 - Домен SET (гістонові хвости);
 - Не містить домен SET (гістонові ядра);

- Аргінін:
 - Родина PRMT (протеїн аргінін метилтрансферази);

Деметилування гістондеметилази:

- Лізин:
 - KDM1/LSD1 (лізин-специфічна деметилаза 1);
 - JmjC (з доменом Jumonji);
- Аргінін:
 - PAD4/PADI4.

Відома перша гістондеметилаза, LSD1 відіграє домінуючу роль у багатьох фізіологічних процесах.

Фосфорилування — процес який включає додавання фосфатних груп до гістонів, зазвичай на серинових, треонінових або тирозинових залишках. Фосфорилування може впливати на взаємодію гістонів з іншими білками та на їхню структуру, регулюючи таким чином доступність генів для транскрипції.

Під час ацетилювання ацетильні групи додаються до лізинових залишків гістонів. Цей процес часто пов'язаний з відкриттям хроматину та активною транскрипцією генів. Ацетилювання може змінювати електростатичні властивості гістонів та змінювати їхню взаємодію з ДНК та іншими білками.

Убіквітування включає приєднання до гістонів білкових молекул, які називають убіквітином. Цей процес може впливати на різноманітні аспекти генної експресії, включаючи деградацію білків та регулювання хроматинової структури.

1.2. LSD1

LSD1 (KDM1A; BHC110; AOF2) був першим білком, який, як повідомляється, виявляє активність гістондеметилази, і з тих пір було показано, що він відіграє багато важливих ролей у біології враховуючи його ферментативну активність і високий рівень експресії в багатьох злоякісних пухлинах людини [4].

Ферментативна активність LSD1 була вперше продемонстрована в основоположному дослідженні лабораторії Ши в 2004 році, де було виявлено, що він

деметилує моно- або диметиллізин 4 гістону H3 (H3K4m1/me2), але не триметил — H3K4 або метил — H3K9 [5].

Гістон-лізин-специфічна деметилаза 1 (LSD1) — це висококонсервативний флавінаденіндинуклеотид (ФАД)-залежний окисний фермент, здатність до експресії якого виявлена в різних пухлинних клітинах, включаючи нейробластому, рак молочної залози, рак легень, рак шлунку, гострий мієлолейкоз тощо [6-10]. LSD1 регулює зростання пухлин та метастазування за допомогою своєї диметилуючої функції, яка визначається молекулами, які він пов'язує. Комплекси LSD1/REST або LSD1/CO-REST рекрутуються на відповідну послідовність промотора і зв'язуються з нею, видаляючи метильну групу в цільовому гені [9, 10]. Структурно повнорозмірний LSD1 складається з 852 амінокислот з трьома ключовими структурними доменами: N-кінцевим доменом SWIRM (маленький альфа-спіральний домен, залишки 172-271), центральним виступаючим доменом TOWER (залишки 417-522) і залишки 271-417 і 522-833 (Рисунок 1.1) [11].

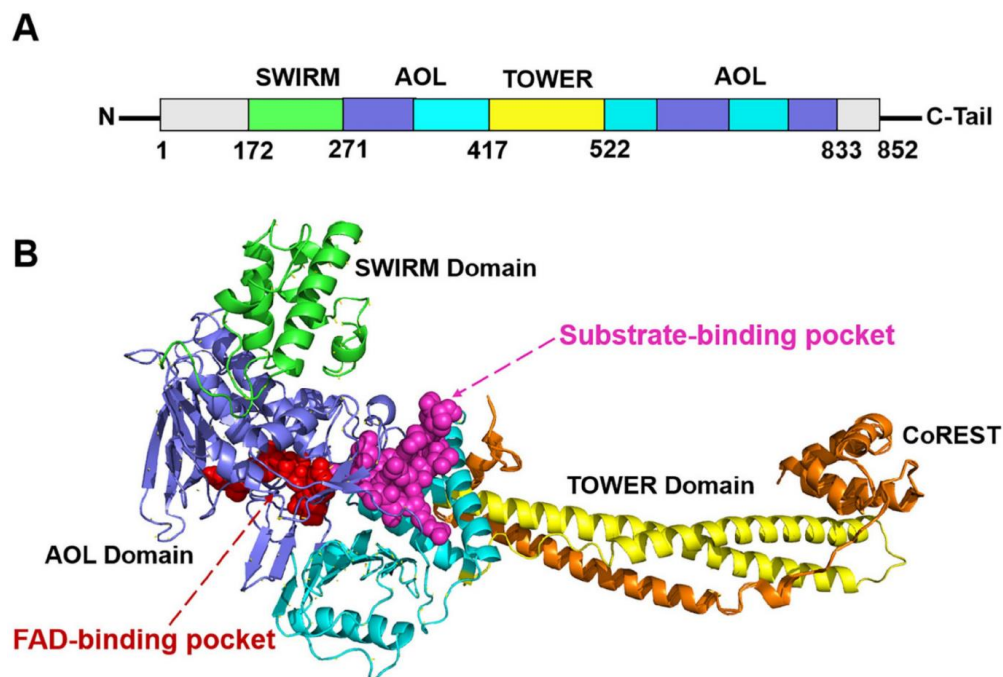


Рисунок 1.1 Схематичне зображення первинної послідовності (A) та третинної структури (B) білка LSD1 [12]. ФАД-зв'язувальна та субстрат-зв'язувальна кишені забарвлені відповідно у червоний та пурпуровий колір

Домен SWIRM сприяє стійкості білка та може брати участь у LSD1-опосередкованих білок-білкових взаємодіях, які необхідні для регуляції епігенетичних модифікацій [13]. Домен TOWER є важливим для підтримки просторової конфігурації домену AOL і може впливати на взаємодію з іншими білками, такими як CoREST. Домен AOL складається з домену, що зв'язує ФАД, і домену, що зв'язує субстрат, з двома активними центрами зв'язування між ними в точці кросовера. Ці два активні центри зв'язування спільно утворюють центр каталітичної активності [12].

1.3. LSD2 гомолог LSD1

Коли LSD1 була відкрита як перша лізиндеметилаза, LSD2 була ідентифікована як єдиний гомолог LSD1 у ссавців. Структура домену людської LSD2 показана на рис. 1.2. Як і LSD1, LSD2 деметилує моно- та диметильований H3K4 у FAD-залежний спосіб [14, 15].

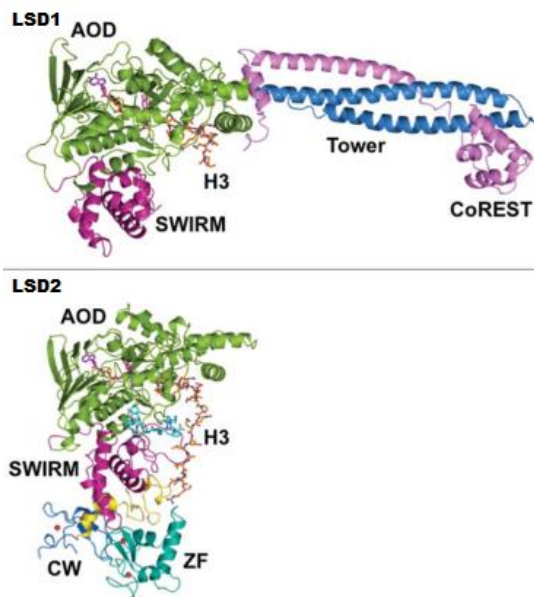


Рисунок 1.2. Структури доменів LSD1 та LSD2 [16]

На відміну від каталітичної активності LSD1, каталітична активність LSD2 досі залишається предметом дискусій. Доменні архітектури обох ферментів мають спільний AOD як каталітичний центр у С-кінцевій ділянці, а також SWIRM-домен на N-кінцевій

ділянці AOD. LSD2 деметилує нуклеосомний гістон H3 *in vivo*. Оскільки LSD2 не має нуклеосомно-зв'язуючої архітектури Tower та CoREST, деметилування нуклеосомного гістону H3 LSD2 має відбуватися структурно відмінним чином від деметилування LSD1. LSD2 має два типи цинкових доменів у N-кінцевій ділянці [16].

Основні відмінності полягають в структурі та активності деметилування. В структурі LSD1 відсутні два типи цинкових пальців, а в структурі LSD2 відсутні домени Tower та CoREST. Додаткова відмінність між цими двома структурами в тім що LSD1 специфічно деметилує гістон H3 на лізині 4 (H3K4) та на лізині 9 (H3K9), тоді як LSD2 здійснює тільки деметилування гістону H3 на лізині 9 (H3K9).

1.4. Механізм дії LSD1

LSD1 специфічно видаляє метильні групи з моно- та диметиллізину 4 або лізину 9 гістону 3 (H3K4me1/2 та H3K9me1/2), діючи як репресор або активатор експресії генів відповідно. Механічно, одноелектронне окислення моно- або диметильованого аміну в ϵ -положенні лізину призводить до утворення катіону імінію з одночасним відновленням FAD до FADH₂. Після відновлення FADH₂ до FAD під дією O₂ утворюється H₂O₂. Катіон імінію нестабільний і після легкого гідролізу дає амін з одним метилом менше і формальдегід як побічний продукт (Рисунок 1.3) [17].

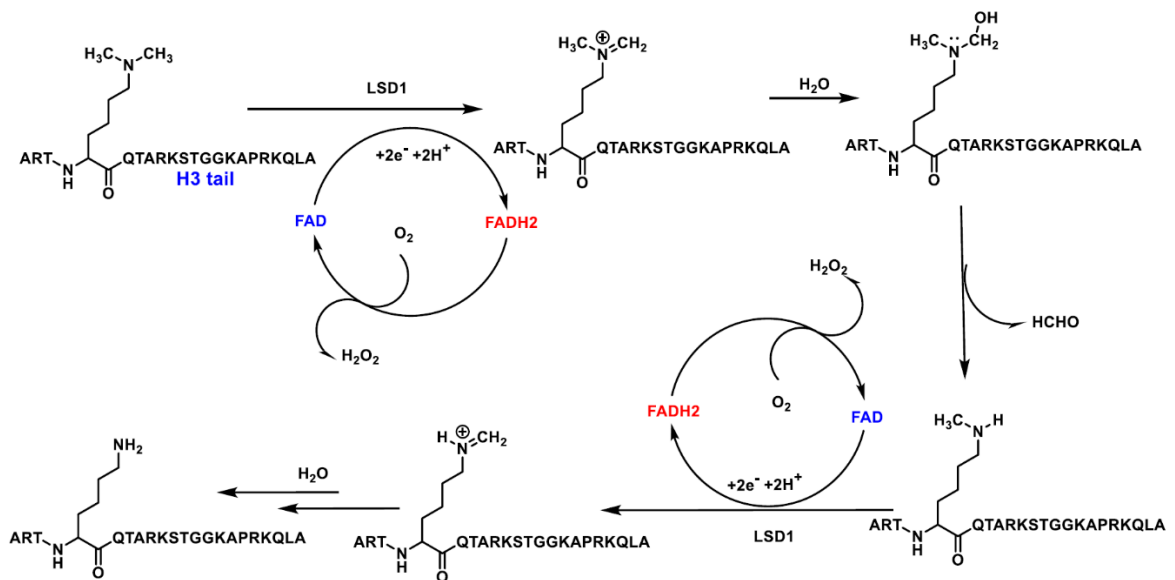


Рисунок 1.3 Каталітичний механізм деметилування LSD1

Ключовою особливістю, яка дозволяє цьому механізму відбуватися, є одиночна пара N-метильованого лізину. Отже, триметильований вид $N^3K^4me_3$ не є субстратом для LSD1, оскільки довга пара відсутня.

1.5. Інгібітори LSD1

З моменту відкриття LSD1 у 2004 році в літературі описано багато інгібіторів LSD1. Інгібітори LSD1 можна поділити на ковалентні та нековалентні. Кожна група включає деякі гібридні сполуки, здатні інгібувати LSD1 на додаток до інших мішеней одночасно. На сьогоднішній день 9 інгібіторів LSD1 пройшли клінічні випробування для лікування різних видів раку. Сім з них (tranylcypromine, iadademstat (ORY-1001), bomedemstat (IMG-7289), GSK-2879552, INCB059872, JBI-802 і Phenelzine) ковалентно зв'язують кофактор FAD, а два є нековалентними інгібіторами LSD1 (pulrodemstat (CC-90011) і seclidemstat (SP-2577)). Ще один подвійний інгібітор LSD1/MAO-B на основі TCP, vafidemstat (ORY-2001), знаходиться в клінічних дослідженнях для лікування хвороби Альцгеймера та розладів особистості (Рисунок 1.4) [18].

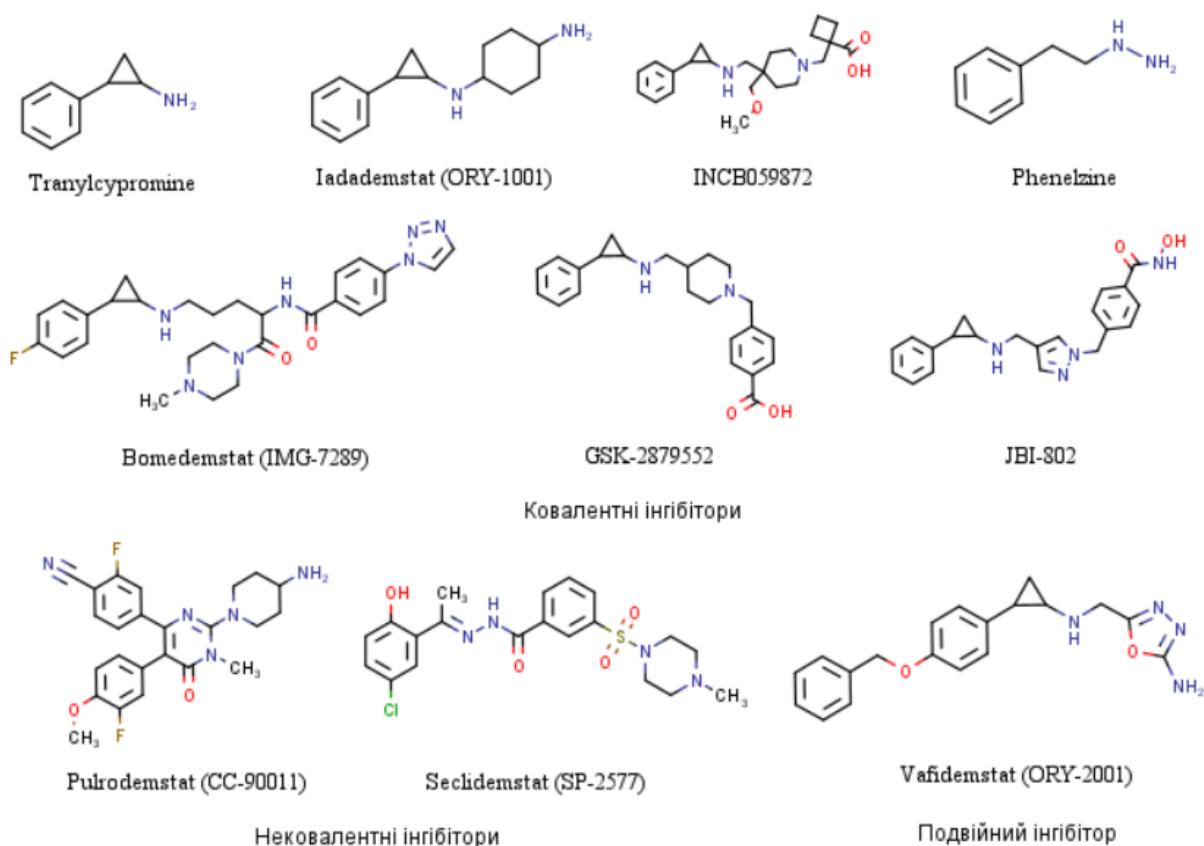


Рисунок 1.4 Інгібітори LSD1

Ковалентні інгібітори LSD1 діють шляхом незворотного зв'язування з ферментом, зазвичай націлені на кофактор FAD у складі ферменту для пригнічення його деметилазної активності, що має вирішальне значення при раку та інших захворюваннях. Ці інгібітори часто демонструють потужну активність і можуть забезпечувати стійкий ефект завдяки стабільній взаємодії. З іншого боку, нековалентні інгібітори оборотно зв'язуються з LSD1, забезпечуючи більш контрольований терапевтичний ефект з потенційно меншою кількістю побічних ефектів. Ці інгібітори зазвичай вступають у транзиторну взаємодію з LSD1, що робить їх корисними в умовах, коли бажана модуляція, а не повне пригнічення LSD1 [18, 19].

Субстрат-зв'язувальна кишеня може регулювати обертання домену аміноксидази щодо домену TOWER, тим самим впливаючи на загальну гнучкість рецептора. Таким

чином, вплив на кишеню, що зв'язує субстрат, може стати життєво важливою стратегією розробки інгібіторів LSD1.

У 2017 році було проведено дослідження ряду молекул тієно[3,2-*b*]піррол-5-карбоксамідів [22].

Представлені інгібітори можуть блокувати функцію LSD1, що може призвести до змін у генній експресії, зупинки росту пухлини або зниження її агресивності. Нещодавні дослідження показали обіцяні результати щодо ефективності інгібіторів LSD1 у клінічних випробуваннях на пацієнтах з різними типами раку. Потрібно проводити дослідження та розробляти нові інгібітори LSD1, які були б не лише більш ефективними у блокуванні функції цього білка, але й менш токсичними для організму. Це важливо для того, щоб забезпечити оптимальну ефективність лікування, одночасно мінімізуючи негативний вплив на здоров'я пацієнтів.

Нові інгібітори повинні мати вищу специфічність до цільового білка, щоб уникнути небажаних побічних ефектів. Крім того, вони повинні мати покращену фармакокінетику та фармакодинаміку, щоб забезпечити оптимальну концентрацію в організмі та тривалість дії при мінімальному впливі на інші системи.

1.6. Програмне забезпечення та методи

Протягом дуже тривалого часу процес відкриття ліків був повільним, трудомістким і надзвичайно капіталомістким. Традиційні процеси зазвичай займають 5-6 років, щоб отримати заявку на застосування ліків, і понад 60% випробувань відкриття зазнають невдачі. Через усі ці неефективності отримання цієї програми в середньому коштує 35 мільйонів доларів. Проте платформа Schrodinger, передбачає критичні властивості молекул, які можуть використовувати розробники. У рецензованому дослідженні Schrodinger порівняли з традиційними методами виявлення для відбору міцно зв'язуваних молекул. Це дослідження призвело до восьмикратного збільшення кількості молекул із бажаною спорідненістю.

Завдяки машинному навчанню Schrodinger набагато швидше знаходить і оцінює молекули, що призводить до швидшого часу відкриття. Оскільки Schrodinger відкриває

більш оптимальні молекули, ці молекули не можуть бути використані лише для розробки ліків (імовірно, для розробки кількох ліків), але вони також можуть бути використані для застосування та винаходу матеріалів для вирощування.

Окремої уваги заслуговує Maestro, який є критично важливим компонентом трансформаційної платформи відкриття ліків, яка використовується для оптимізації циклу проектування-вироблення-тестування-аналізу, що призводить до швидшого отримання сполук кращої якості. Maestro — це спрощений портал для структурної візуалізації та доступу до найсучаснішого прогностного обчислювального моделювання та робочих процесів машинного навчання для відкриття маломолекулярних ліків. Перевагами Maestro є зрозуміла візуалізація структур, моделювання та інтерпретація молекул в різних форматах, що полегшує молекулярні дослідження та пришвидшує процес оцінки молекулярної структури [12]. Для дослідження обрані модулі Glide, LigPrep, 3d-QSAR, детальніше кожен з них описаний далі.

Молекулярний докінг - це важливий обчислювальний метод у структурній біології та дизайні ліків, який використовується для прогнозування переважної орієнтації однієї молекули відносно другої, коли вони зв'язуються одна з одною, утворюючи стабільний комплекс. Дані, отримані в результаті докінгу, можуть бути використані для прогнозування сили і типу взаємодій, які можуть виникати між молекулами, таких як водневі зв'язки, гідрофобні взаємодії та електростатичні взаємодії [20].

Кількісні співвідношення структура-активність (QSAR) давно були способом для дослідників оптимізувати біологічноактивні сполуки. Однак традиційні моделі QSAR зазвичай включають лише грубі наближення 3D-структури. Field-Based QSAR усуває це обмеження. Починаючи з набору вирівняних лігандів із відомою активністю, Field-Based QSAR здатний зробити висновок про те, як електростатичні, гідрофобні та стеричні поля ліганду призводять до біологічної активності. Він здатний швидко перетворювати наявні набори даних у корисні моделі QSAR, синтетичні блокпости та обмеження ADME. Field-Based QSAR дає точний ступінь контролю над побудовою моделі, завдяки налаштувати найважливіші параметри побудови [21].

Також він дозволяє легко перевірити прогностичну силу вашої моделі. Використовуючи ліганди з відомою активністю, Field-Based QSAR може випадковим чином розділити їх на тренувальні та тестові набори. Окрім коефіцієнтів кореляції, Field-Based QSAR звітує про всі широко використовувані статистичні показники та пропонує інструменти двовимірного побудови. На додаток до його очевидної корисності як кількісного інструменту, Field-Based QSAR також допомагає визначати загальні якісні характеристики активних лігандів. Контурна візуалізація QSAR показує, де розміщувати стеричний об'єм, гідрофобні групи та позитивно чи негативно заряджені функціональні групи [21].

Модуль LigPrep вирішує проблему ефективного та точного перетворення 2D у 3D. Окрім простого структурного перетворення «один-до-одного», не менш важливо створювати науково обґрунтовані молекулярні моделі, які перераховують різні структурні та хімічні можливості, які може взяти ліганд, оскільки ці варіації можуть призвести до різко різних результатів у наступних обчисленнях. LigPrep створює точні тривимірні молекулярні структури з мінімізованою енергією а також застосовує складні правила для виправлення структур Льюїса та усунення помилок у лігандах, щоб зменшити обчислювальні помилки, що відбуваються далі.

Цей модуль додатково розширює таутомерні та іонізаційні стани, конформації кільця та стереоізомери для створення широкого хімічного та структурного різноманіття з однієї вхідної структури. LigPrep застосовує фільтри для усунення сполук, які не відповідають заданим користувачем критеріям, що дозволяє створити повністю налаштовану бібліотеку лігандів. Особливістю LigPrep є налаштування, спеціально налаштовані для генерації вхідних структур для Glide і Phase та швидкість обробки лігандів, приблизно один ліганд за секунду, що дає змогу конвертувати цілі бази даних одночасно [22].

Широке використання комбінаторної хімії та високопродуктивного скринінгу (HTS) у фармацевтичній та біотехнологічній промисловості означає, що велику кількість сполук тепер можна регулярно досліджувати на біологічну активність. Проте скринінг великих хімічних бібліотек залишається дорогим і трудомістким процесом із значними показниками як хибнопозитивних, так і хибнонегативних результатів.

Glide пропонує повний спектр варіантів залежності швидкості від точності, від режиму HTVS (високопродуктивний віртуальний скринінг) для ефективного збагачення мільйонів бібліотек сполук до режиму SP (стандартна точність) для надійного приєднання десятків до сотень тисяч лігандів із високою точністю, до режиму XP (надточна точність), де подальше усунення хибних спрацьовувань досягається більш широкою вибіркою та вдосконаленим оцінюванням, що призводить до ще більшого збагачення. Також цей модуль забезпечує раціональний робочий процес для віртуального скринінгу від HTVS до SP до XP, збагачуючи дані на кожному рівні таким чином, що лише на порядок менше сполук потрібно вивчати на наступному вищому рівні точності.

Glide надійно знаходить правильні режими прив'язки для великої кількості тестів. Він перевершує інші програми док-станції в досягненні менших середньоквадратичних відхилень від нативних співкристалізованих структур та демонструє чудову точність стикування та високе збагачення для різноманітних типів рецепторів [23].

2 РОЗРАХУНКОВА ЧАСТИНА

Останнім часом особливо цікавими для дослідження стали інгібітори тієно[3,2-*b*]піррол-5-карбоксаміди, які були відкриті у 2017 році. Ці сполуки є біологічно активними і виявили потенціал у блокуванні дії лізин диметилази 1A (LSD1). Основна структура цих інгібіторів була взята з Protein Data Bank за кодом O60341, який відповідає LSD1. За допомогою кристалографічних досліджень, було встановлено структурні властивості цих інгібіторів. Кристали 5LGT, 5LHG, 5LGU, 5LHN та 5LHI представляють собою візуалізацію взаємодії між цими сполуками та білком LSD1 (Рисунок 2.1).

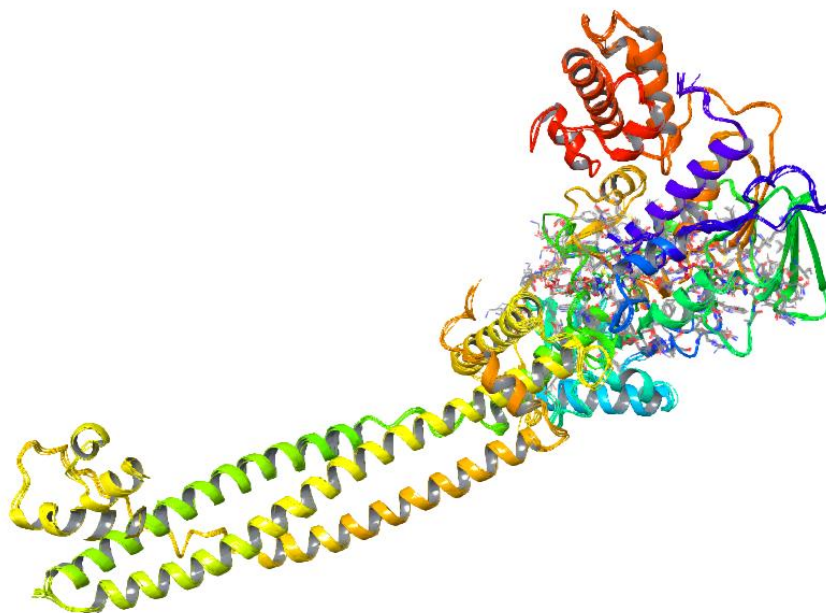







Рисунок 2.1. Візуалізація зв'язування LSD1 (дано сукупне представлення відомих кристалічних форм)

Основні параметри та характеристики структур, такі як роздільна здатність (Resolution) та оцінка якості ліганда (Ligand Structure Quality Assessment) наведені в Таблиці 2.1.

Роздільна здатність вказує на міру якості або точності структурного розв'язку. Даний параметр вимірюється в ангстремах (Å) і представляє собою мінімальну відстань між двома атомами в білку, яку можна розпізнати на рентгенограмі із заданою точністю. Низьке значення роздільної здатності вказує на високу точність. Ligand Structure Quality Assessment — це процес оцінки якості структури ліганду у білкових комплексах. Оцінка якості структури ліганду визначається різними метриками, такими як узгодженість атомів, стереохімічна якість та інші. Ці параметри дозволяють зробити висновок про достовірність та якість визначеної структури ліганду. Краще значення даного параметру вважається коли показник наближається до одиниці.

Таблиця 2.1. Основні характеристики кристалів

Identifier	Resolution	Ligand Structure Quality Assessment
5LGT	3.00 Å	 Ligand structure goodness of fit to experimental data
5LHG	3.34 Å	 Ligand structure goodness of fit to experimental data
5LGU	3.20 Å	 Ligand structure goodness of fit to experimental data
5LHN	3.05 Å	 Ligand structure goodness of fit to experimental data
5LHI	3.40 Å	 Ligand structure goodness of fit to experimental data

Після аналізу структур та допоміжних сайтів, свою увагу залишили на кристалі 5LHN. За даними Таблиці 2.1, даний кристал має достатньо низьку роздільну здатність, у порівнянні з іншими кристалами, та значення його показника оцінки якості ліганду більше заходить у синю межу.

Наступним етапом є перевірка як рідний ліганд буде докуватися в структуру. На Рисунку 2.2 наведена схема зв'язування рідного ліганду в кристалі, яка була отримана в результаті проведення процедури редокінгу. Для цього спершу було використано процедуру LigPrep для переформатування файлу з лігандом, забезпечуючи його оптимізацію та підготовку до подальшого докінгу. Потім була застосована процедура грід для визначення потенційних місць зв'язування ліганду з активною субстратзв'язуючою кишенею білка LSD1.

Згідно з літературою, активна субстратзв'язуюча кишеня LSD1 розташована на поверхні структурного домену AOL і містить 22 амінокислотні залишки. Ці амінокислоти включають Gly330, Met332, Val333, Thr335, Ile356, Gln358, Phe538, Ala539, Asn540, Trp552, Asp555, His564, Lys661, Leu659, Leu677, Trp695, Leu706, Tyr761, Ser762, Pro808, Ala809 та Thr810 [12].

На основі отриманих результатів можна зробити висновок про те, як ліганд взаємодіє з активною субстратзв'язуючою кишенею білка LSD1.

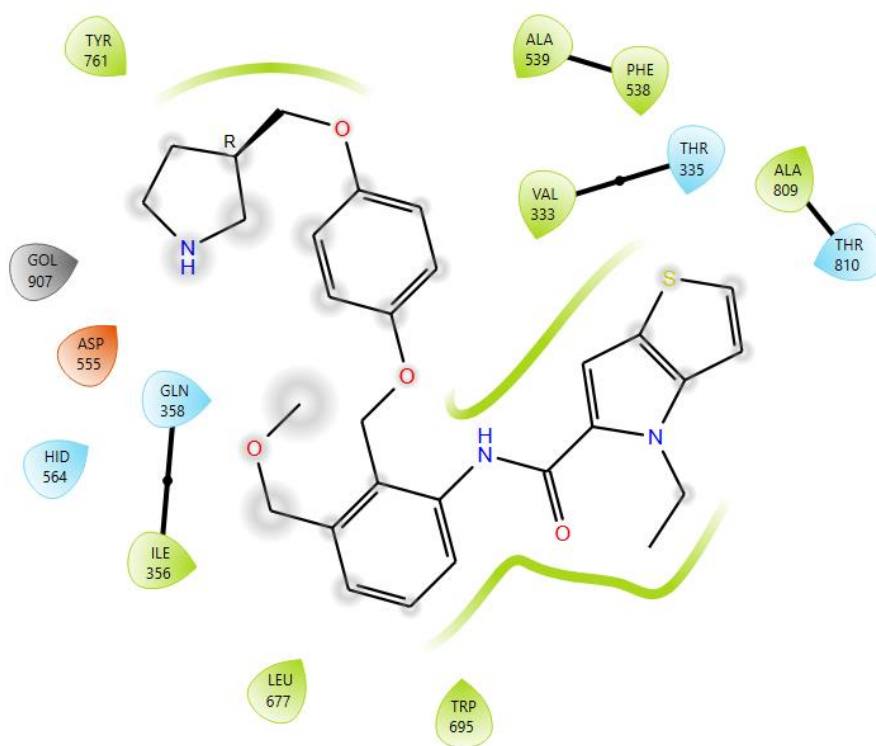


Рисунок 2.2. Схема зв'язування ліганда із білком 5LHN

Після повторного докінгу виявлено, що ліганди мають однаковий скафолд, але результати проведеного докінгу майже співпали з попередніми результатами, як показано на Рисунку 2.3. Оцінка docking score складає -9.294, що вказує на задовільне розташування. Цей параметр відображає, наскільки ефективно ліганд зв'язується зі своєю цільовою макромолекулою. Для програми Maestro бажано, щоб значення цього параметра було нижче -6, що вказує на більш сильне зв'язування ліганду з білком.

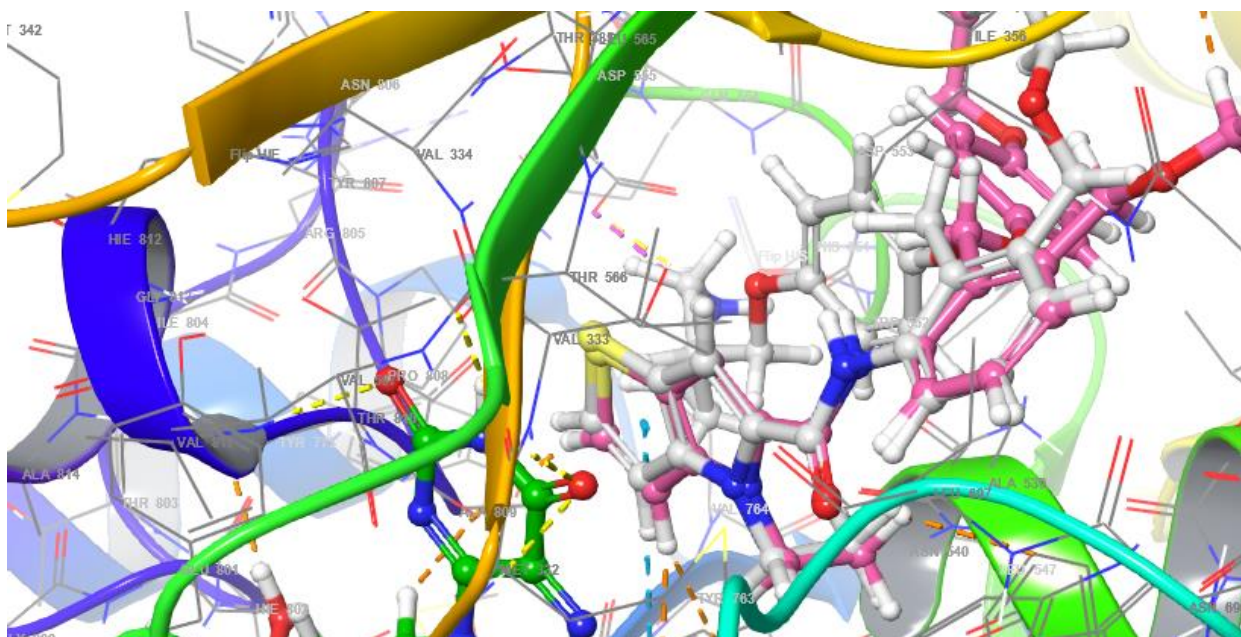


Рисунок 2.3 Результати повторного докінгу (білий — ліганд в кристалі, рожевий — положення ліганду після повторного докінгу)

Схожа ситуація спостерігається і в інших структурах, де ліганди мають спільний тієно[3,2-*b*]піррол-5-карбоксамідний каркас з 75% співпадіння, як показано на Рисунку 2.4. Це вказує на те, що цей каркас є ключовим елементом у структурі та взаємодії лігандів з білком LSD1. Цей фрагмент може мати значення для дизайну нових інгібіторів або вдосконалення існуючих, оскільки він відображає ключові взаємодії з активною субстратзв'язуючою кишенею білка.

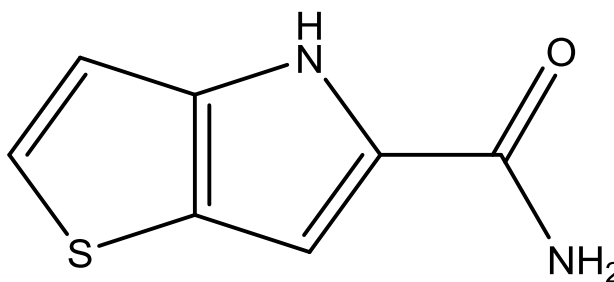


Рисунок 2.4 Структурна формула тієно[3,2-*b*]піррол-5-карбоксамідного каркасу

Використовувався набір даних, що складався з 22 сполук. Ці сполуки були обрані на основі їхньої активності в інгібуванні LSD1 та описані у статті [11]. Основним критерієм вибору було наявність значень pIC_{50} для кожної сполуки, які вказують на їхню потужність в інгібуванні цільового білка. Крім того, для кожної з цих сполук також використовувалося правило Ліпінського, яке дозволяє визначити, наскільки молекула є потенційно проникливою через біологічні мембрани та наскільки вона може бути ефективною як лікарський засіб.

pIC_{50} — це міра концентрації або активності лікарської сполуки у біологічному тесті, яка використовується для визначення її потужності як інгібітора чи агоніста взаємодії з біомолекулою, зазвичай рецептором або ферментом.

Правило Ліпінського (Lipinski's Rule) — це набір емпіричних правил для оцінки "лікарської податливості" сполук, які мають високий потенціал для стаціонарного дослідження на етапі розробки лікарських засобів. Це правило допомагає визначити, що дана сполука може бути успішною як лікарський препарат. Основні принципи правила Ліпінського включають:

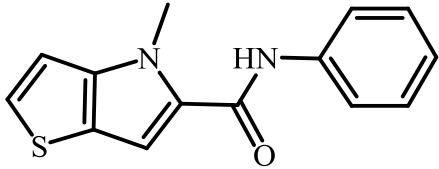
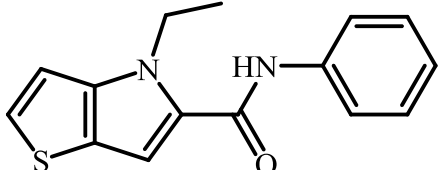
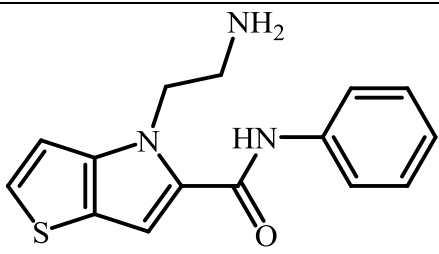
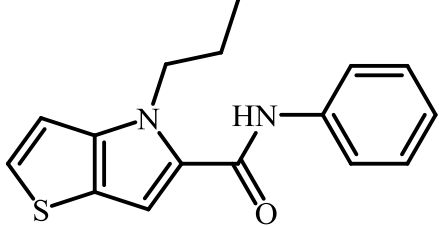
- молекули з молекулярною масою (MW) менше 500 добре проникають через клітинну мембрану;
- коефіцієнт розподілу між октанолом і водою ($\log P$) не повинен перевищувати 5, бо високі значення можуть свідчити про погану розчинність у воді або надмірну ліпофільність;

– кількість акцепторів водню (H-acceptors) не повинно бути більше 10, бо велика кількість H-acceptors може сприяти високій афінності до води і поганій оральній біодоступності;

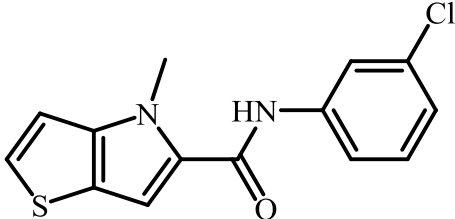
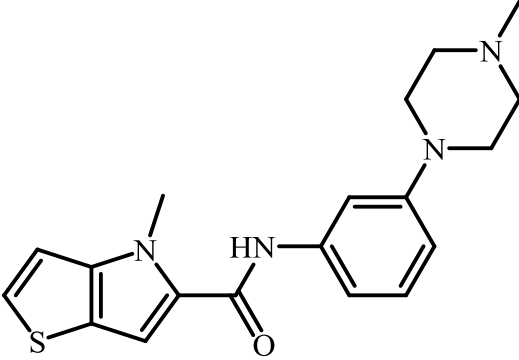
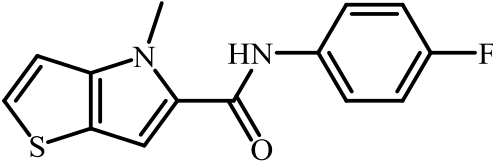
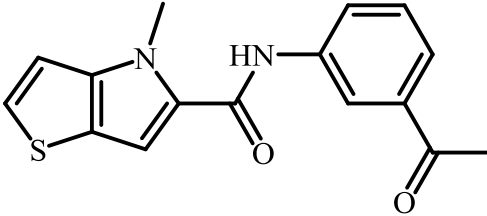
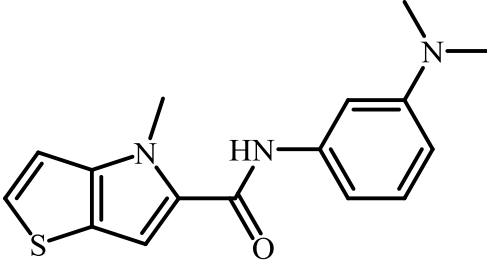
– кількість донорів водню (H-donors) не повинно бути більше 5, бо велика кількість H-donors може сприяти високій розчинності у воді.

За допомогою цих правил в статті були обрані дані сполуки з даним каркасом. Нижче наведена таблиця 2.2 зі сполуками їх значенням pIC50 та результатом докінгу.

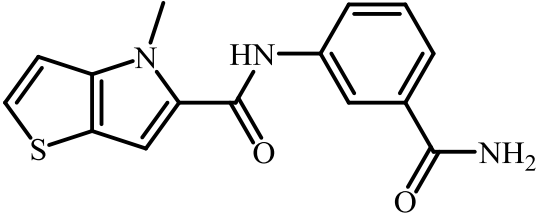
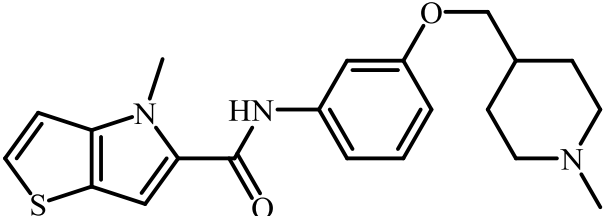
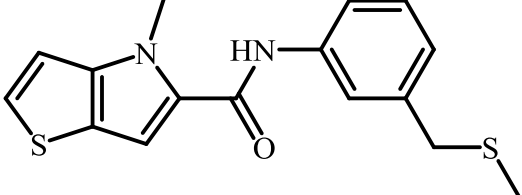
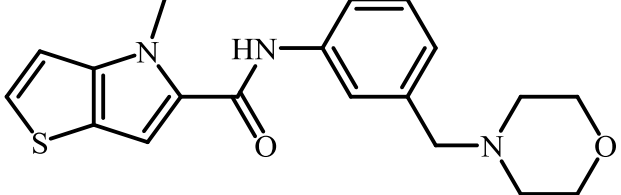
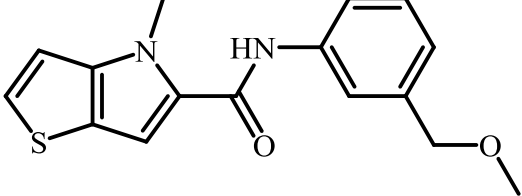
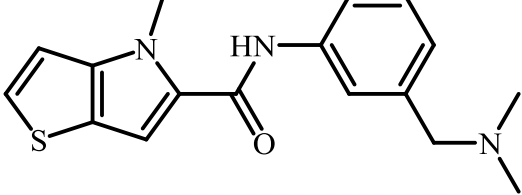
Таблиця 2.2 Сполуки для докінгу та результати докінгу за скафолдом

№	Сполука	pIC50	Docking score
1		4,59	-6.877
2		5,032	-7.310
3		4,318	-6.921
4		4,27	-7.084

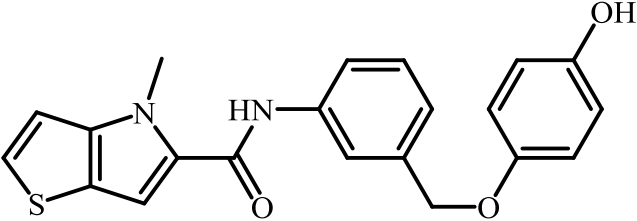
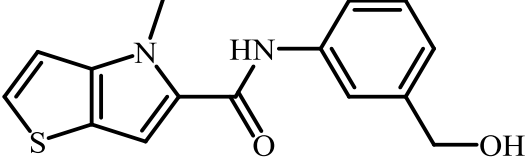
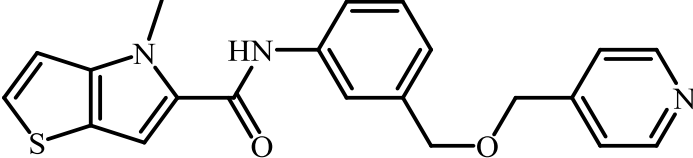
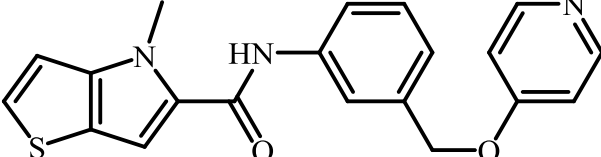
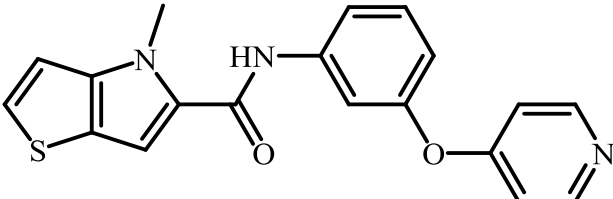
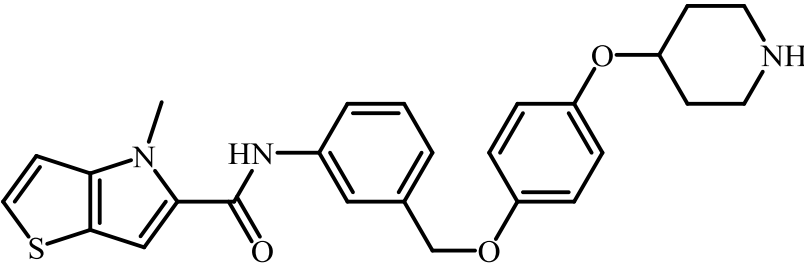
Продовження Таблиці 2.2

№	Сполука	pIC50	Docking score
5		4,728	-6.761
6		4,668	-7.124
7		4,654	-6.886
8		4,577	-7.485
9		4,529	-7.345

Продовження Таблиці 2.2

№	Сполука	pIC50	Docking score
10		4,012	-8.249
11		5,658	-7.110
12		5,131	-7.015
13		4,955	-7.413
14		4,873	-6.985
15		4,777	-7.115

Продовження Таблиці 2.2

№	Сполука	pIC50	Docking score
16		4,599	-7.423
17		4,281	-7.352
18		5,143	-7.584
19		5,027	-7.183
20		4,046	-7.485
21		6,79	-7.730

Під час докінгу було важливо, щоб всі сполуки лягали рівно каркасом, тому на стадії створення грид боксу застосовувалися обмеження, як показано на Рисунку 2.5. Обмеження в докінгу використовуються для контролю і підвищення точності докінгу, а також для

моделювання конкретних умов або взаємодій, які можуть бути важливими у біологічних системах.

Існують різні типи обмежень, які можна використовувати, такі як фіксація позиції атома, фіксація відстані та фіксація зв'язку з амінокислотою. У даному випадку було застосовано обмеження на фіксацію позиції атома, щоб забезпечити рівне лягання каркасу усіх сполук під час докінгу.

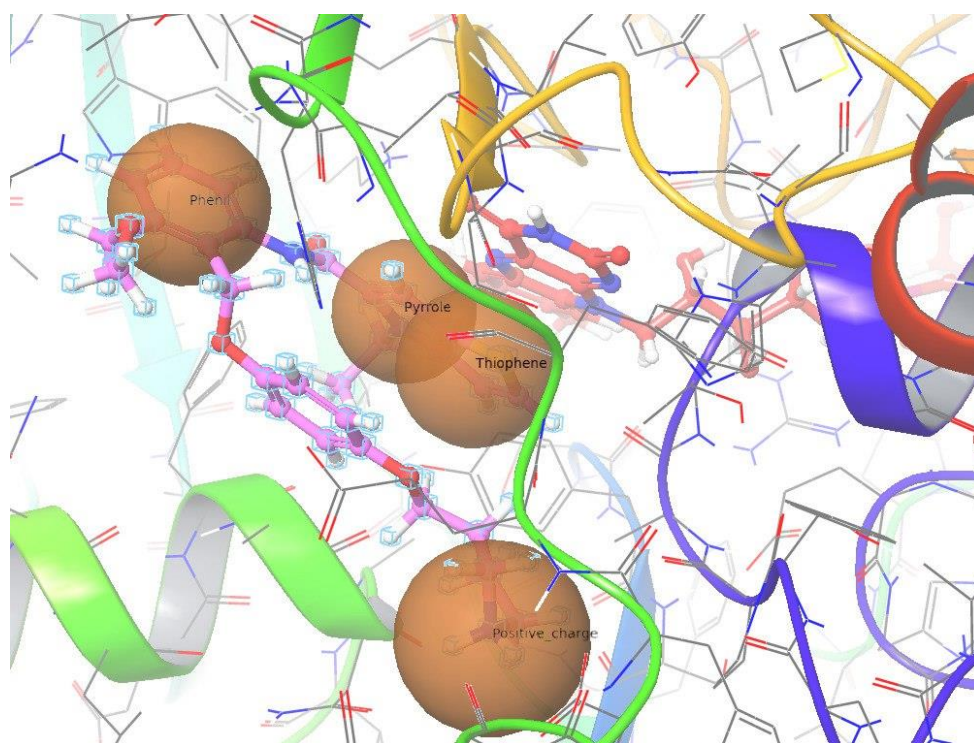


Рисунок 2.5 Обмеження для докінгу

Після докінгу було виявлено що всі доковані ліганди лягли однаковим скафолдом, що викликає зацікавленість. Величини оціночної функції (docking score) вказані в Таблиці 2.2, а на Рисунку 2.6, можна побачити, що структури «лягли» в одну позицію, яка відповідає ефективному зв'язуванню.

Ці результати свідчать про те, що всі досліджувані сполуки мають схожу взаємодію з білком LSD1, що може вказувати на наявність специфічних взаємодій між цими

лігандами та цільовим білком. Отримані docking score також дозволяють оцінити потенційну активність кожного з лігандів в інгібуванні білка LSD1.

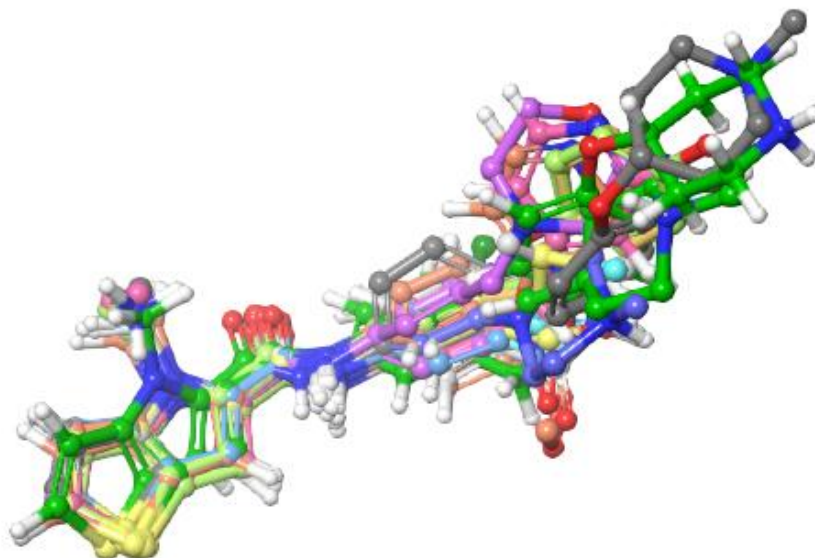


Рисунок 2.6 Результат докінгу: ліганди зв'язані скафолдом

Для порівняння, на Рисунку 2.7 показані ліганди з кристалів, що були згадані вище, а також структури, які викликали зацікавленість своїм спільним каркасом.

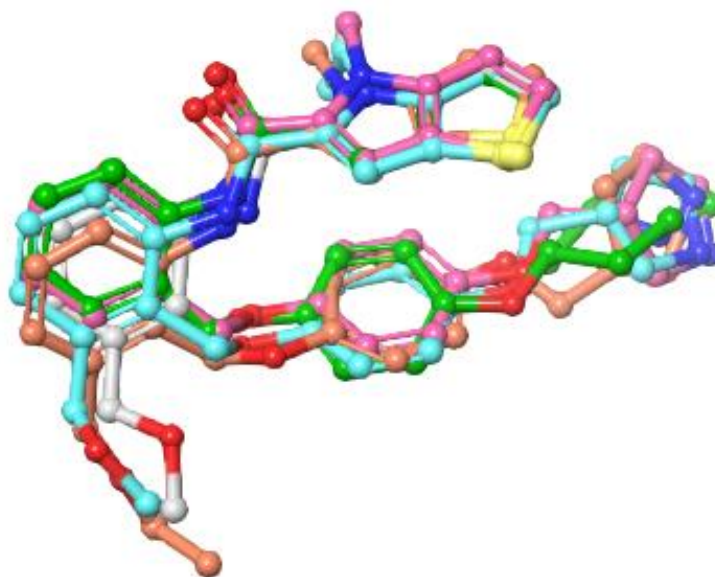


Рисунок 2.7 Ліганди з кристалів

Докінговий експеримент був проведений з метою побудови QSAR моделі та передбачення активності для нових сполук. Нові сполуки були знайдені у базі ZINC за схожістю з експериментально встановленими лігандами, які мали однаковий каркас. За моїм пошуком було виявлено 97 сполук з таким же каркасом. Наступним кроком було перевірка, чи вони відповідають правилам Ліпінського, PEINS, BRENK та LILLY.

Компоненти, що впливають на пантест (PAINS), — це хімічні сполуки, які часто дають хибнопозитивні результати на високопродуктивних екранах. PAINS, як правило, неспецифічно реагує на численні біологічні мішені, а не на одну потрібну мішень. Ряд руйнівних функціональних груп спільні для багатьох PAINS. Правило Ліпінського використовується для оцінки потенційної пермеабельності сполук через біологічні мембрани, BRENK визначає, чи сполука має потенціал стати лікарським засобом, а LILLY оцінює ймовірність, що сполука буде мати лікарські властивості.

Після перевірки залишилось 89 сполук. Далі для сполук які залишились проводився докінг. Результати вийшли гарними, найвищий docking score -7.488 для сполуки з кодом ZINC000006804679 (Рисунок 2.8 та Рисунок 2.9).

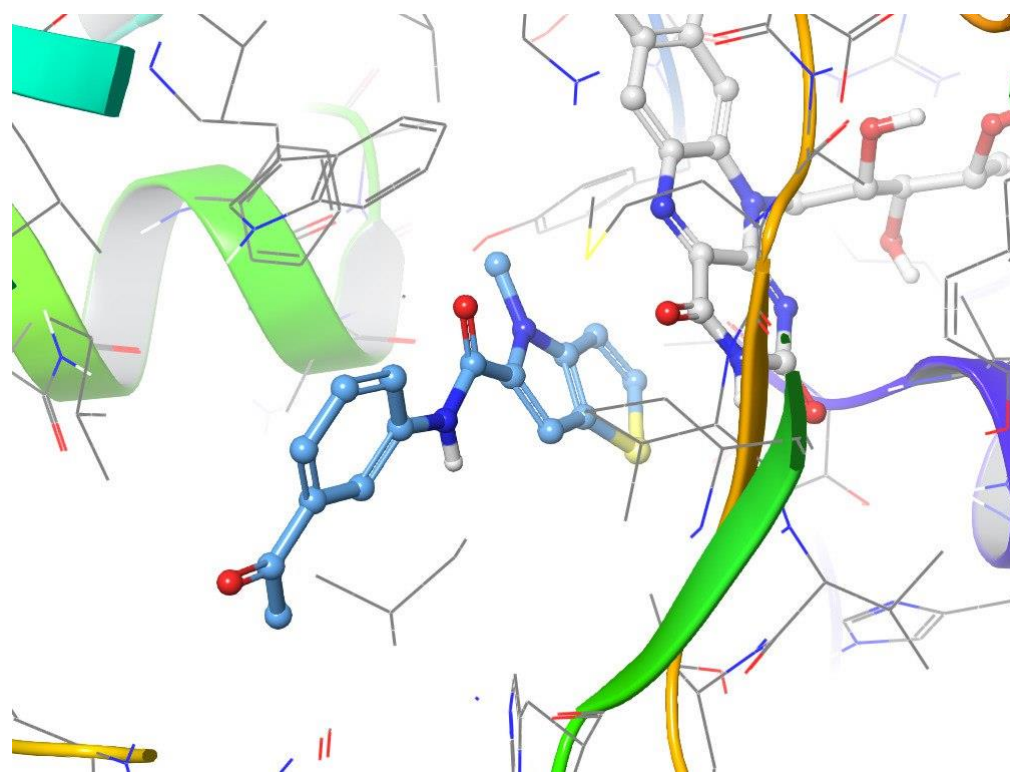


Рисунок 2.8 ZINC000006804679 в структурі білку за результатами докінгу

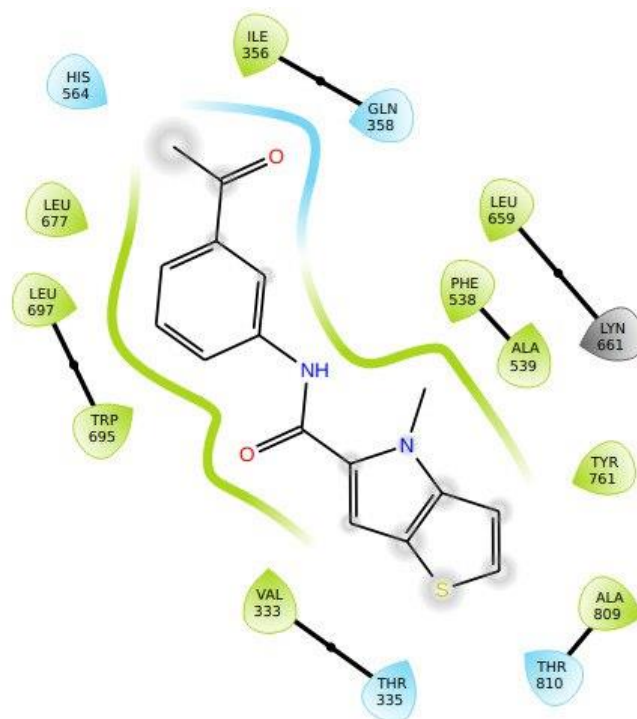


Рисунок 2.9 Діаграма взаємодії ZINC000006804679 в кармані зв'язування

Після докінгу за допомогою 3-D Field-Based (CoMFA, Comparative Molecular Field Analysis) побудовано модель для прогнозу значення pIC_{50} у нових сполук. Field-Based QSAR дає змогу оцінити точність побудованої моделі. Спираючись на інформацію щодо активності відомих сполук побудовано модель PLS з урахуванням до п'яти факторів. Отже отримано 5 моделей, і як можна побачити з Таблиці 2.3., то вже гарний коефіцієнт детермінації (R^2) починається з урахуванням 4 факторів.

Таблиця 2.3. Статистика QSAR

Factors	SD	R^2	RMSE	Q^2
1	0.68	0.71	0.97	0.33
2	0.44	0.88	0.88	0.45
3	0.36	0.93	0.79	0.55
4	0.26	0.96	0.81	0.53
5	0.22	0.97	0.79	0.55

В Таблиці 2.3 величини Q^2 — коефіцієнти детермінації, що були отримані за процедурою LOO-CV (leave-one-out cross validation). Графіки фактичної та прогнозованої активності за процедурою 3-D Field-Based представлено на Рисунок 2.10.

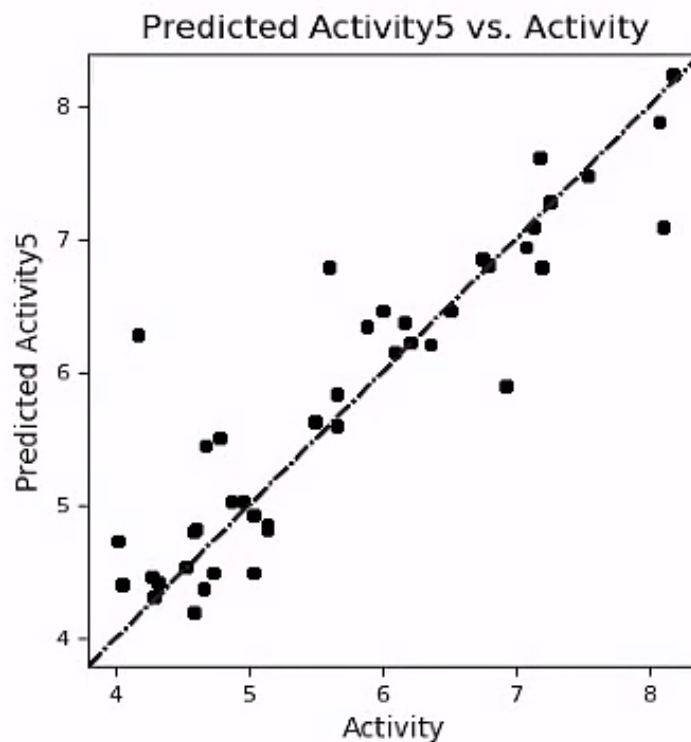


Рисунок 2.10 Фактична та прогнозована активність моделі QSAR 3-D Field-Based 5

Далі вже прогнозуємо pIC_{50} за п'ятою моделлю. Результати наведені в Таблиці 2.4. для 20 кращих сполук.

Таблиця 2.4 Прогнозовані результати

ID	predicted activity 5
ZINC000004970920	6.073
ZINC000004970997	5.973
ZINC000004970998	5.894
ZINC000040786565	5.888
ZINC000040798817	5.817
ZINC000040786565	5.783

Продовження таблиці 2.4

ID	predicted activity 5
ZINC000096363840	5.579
ZINC000040798654	5.565
ZINC000075487890	5.553
ZINC000040287737	5.496
ZINC000005036602	5.464
ZINC000005036535	5.458
ZINC000096363847	5.430
ZINC000075487786	5.420
ZINC000044920761	5.396
ZINC000040786270	5.394
ZINC000096363839	5.380
ZINC000075487833	5.354
ZINC000040786525	5.325
ZINC000075487798	5.307

Таким чином, згідно нашої прогностичної моделі, найкращою можна визнати сполуку ZINC000004970920 з найвищим значенням активності (Рисунок 2.11). Надалі необхідно провести біологічні дослідження для визначення ефективності даної сполуки.

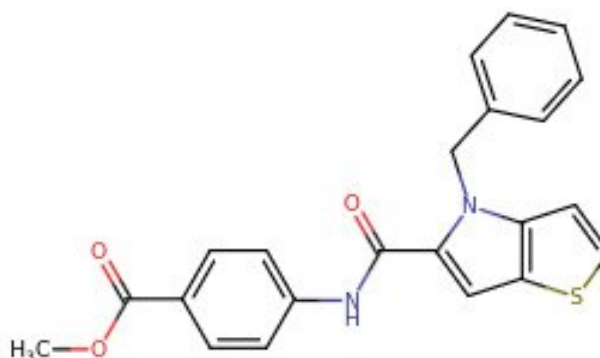


Рисунок 2.11 ZINC000004970920

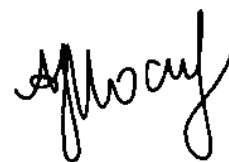
ВИСНОВКИ

1) Проведено масштабний пошук інгібіторів LSD1 на основі сканування комерційно доступних баз даних. Серед них Enamine STOCK, та бази ChEMBL і PubChem.

2) На основі молекулярного докінгу знайдено ряд структур які можуть бути перспективними інгібіторами LSD1. Серед них серія тієно[3,2-*b*]піррол-5-карбоксамідів. Моделі взаємодії білок-ліганд були ідентифіковані та визначені за допомогою основних хімічних взаємодій що описані для аналогових лігандів.

3) Репрезентативні результати молекулярного докінгу малих молекул (лігандів) були використані для пояснення відмінностей у біологічній активності. В рамках методу молекулярних полів отримано ряд моделей які мають непогану прогностичну здатність згідно параметрів внутрішньої валідації.

4) Отримані результати і моделі можна використати для прогнозування активності цієї серії сполук, і тим самим зменшити втрати «сліпих інвестицій» в експеримент.



ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Hogg, S.J., et al., *Targeting the epigenetic regulation of antitumour immunity*. Nat Rev Drug Discov, 2020. **19**(11): p. 776-800.
2. Dawson, M.A. and T. Kouzarides, *Cancer epigenetics: from mechanism to therapy*. Cell, 2012. **150**(1): p. 12-27.
3. Greer, E.L. and Y. Shi, *Histone methylation: a dynamic mark in health, disease and inheritance*. Nat Rev Genet, 2012. **13**(5): p. 343-57.
4. Maiques-Diaz, A. and T.C. Somerville, *LSD1: biologic roles and therapeutic targeting*. Epigenomics, 2016. **8**(8): p. 1103-16.
5. Shi, Y., et al., *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1*. Cell, 2004. **119**(7): p. 941-53.
6. Jenuwein, T. and C.D. Allis, *Translating the histone code*. Science, 2001. **293**(5532): p. 1074-80.
7. Anand, R. and R. Marmorstein, *Structure and mechanism of lysine-specific demethylase enzymes*. J Biol Chem, 2007. **282**(49): p. 35425-9.
8. Luka, Z., et al., *Crystal structure of the histone lysine specific demethylase LSD1 complexed with tetrahydrofolate*. Protein Sci, 2014. **23**(7): p. 993-8.
9. Forneris, F., et al., *Histone demethylation catalysed by LSD1 is a flavin-dependent oxidative process*. FEBS Lett, 2005. **579**(10): p. 2203-7.
10. Ciccone, D.N., et al., *KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints*. Nature, 2009. **461**(7262): p. 415-8.
11. Xu, Y., et al., *3D-QSAR, molecular docking, and molecular dynamics simulation study of thieno[3,2-b]pyrrole-5-carboxamide derivatives as LSD1 inhibitors*. RSC Adv, 2020. **10**(12): p. 6927-6943.
12. Wright, W.C., et al., *Clobetasol Propionate Is a Heme-Mediated Selective Inhibitor of Human Cytochrome P450 3A5*. J Med Chem, 2020. **63**(3): p. 1415-1433.

13. Niwa, H., et al., *Crystal Structure of LSD1 in Complex with 4-[5-(Piperidin-4-ylmethoxy)-2-(p-tolyl)pyridin-3-yl]benzotrile*. *Molecules*, 2018. **23**(7).
14. Karytinios, A., et al., *A novel mammalian flavin-dependent histone demethylase*. *J Biol Chem*, 2009. **284**(26): p. 17775-82.
15. Fang, R., et al., *Human LSD2/KDM1b/AOF1 Regulates Gene Transcription by Modulating Intragenic H3K4me2 Methylation*. *Molecular Cell*, 2010. **39**(2): p. 222-233.
16. Nagaoka, K., et al., *Lysine-Specific Demethylase 2 Suppresses Lipid Influx and Metabolism in Hepatic Cells*. *Molecular and Cellular Biology*, 2015. **35**(7): p. 1068-1080.
17. Hayward, D. and P.A. Cole, *Chapter Eleven - LSD1 Histone Demethylase Assays and Inhibition*, in *Methods in Enzymology*, R. Marmorstein, Editor. 2016, Academic Press. p. 261-278.
18. Noce, B., et al., *LSD1 inhibitors for cancer treatment: Focus on multi-target agents and compounds in clinical trials*. *Front Pharmacol*, 2023. **14**: p. 1120911.
19. Yin, Z., et al., *LSD1-Based Reversible Inhibitors Virtual Screening and Binding Mechanism Computational Study*. *Molecules*, 2023. **28**(14).
20. Dnyandev, K.M., et al., *A Review on Molecular Docking*. *International Research Journal of Pure and Applied Chemistry*, 2021. **22**(3): p. 60-68.
21. Cappel, D., et al., *Exploring conformational search protocols for ligand-based virtual screening and 3-D QSAR modeling*. *J Comput Aided Mol Des*, 2015. **29**(2): p. 165-82.
22. Van Den Driessche, G. and D. Fourches, *Adverse drug reactions triggered by the common HLA-B*57:01 variant: a molecular docking study*. *J Cheminform*, 2017. **9**: p. 13.
23. Friesner, R.A., et al., *Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes*. *J Med Chem*, 2006. **49**(21): p. 6177-96.