

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА  
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ВИСОКИХ ТЕХНОЛОГІЙ

Завідувач кафедри молекулярної біотехнології та біоінформатики

доц. Олексій Юрійович Нипорко

Протокол № \_\_\_\_ засідання кафедри

від “ \_\_\_\_ ” \_\_\_\_\_ 2023 р.

**ПОШУК ЕФЕКТОРІВ ДО ПРЕДСТАВНИКІВ РОДИНИ ВЕТ БІЛКІВ**

Випускна кваліфікаційна робота магістра  
студента спеціальності 091 Біологія  
ОП «Біоінформатика та структурна біологія»  
**Овдєнка Владислава Олеговича**

Науковий керівник від кафедри  
Професор кафедри молекулярної біотехнології  
та біоінформатики  
д.б.н. **Дзядевич Сергій Вікторович**

Робота виконана у відділі біоінформатики  
Інституту харчової біотехнології та геноміки НАН України  
під керівництвом д.б.н. **П. А. Карпова**

Оцінка захисту роботи

---

Київ – 2023р.

## АНОТАЦІЯ

Овдієнко В.О. Пошук ефекторів до представників родини ВЕТ білків. – Випускна кваліфікаційна робота магістра за спеціальністю 091 «Біологія» ОП «Біоінформатика та структурна біологія».

У роботі проведено пошук потенційних ефекторів до родини ВЕТ білків на основі референсного ліганду. Проведений молекулярний докінг з подальшим його аналізом і створення моделі фармакофору для лігандів з аналізом взаємодій речовин з амінокислотами активного сайту. Отримані результати можуть бути використані для подальших досліджень у сфері пошуку низькомолекулярних ефекторів проти родини білків ВЕТ з ціллю отримання лікарських засобів проти низки патологічних станів (наприклад ракових патологій), що опосередковуються цими білками.

**Ключові слова:** бромодомен; віртуальний скринінг; молекулярний докінг; Similarity Search

**ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ**

АРТ	–	Антиретровірусна терапія
БД	–	Бромодомен
РСА	–	Рентгеноструктурний аналіз
ТЗТ	–	Теломерна зворотня транскриптаза
ВЕТ	–	Bromo- and extraterminal domain (proteins) (білки з бромодоменом та екстратермінальним доменом)
DMSO	–	Dimethyl sulfoxide (диметилсульфоксид)
H3K27Ac	–	Histone 3 Lysine 27 Acetylated (Ацетильований 27й залишок лізину гістону H3)
H3K4Me1	–	Histone 3 Lysine 3 monomethylated (Монометильований 3й залишок лізину гістону H3)
H4R3me1	–	Histone 4 Arginine 3 monomethylated (Монометильований 4й залишок аргініну гістону H4)
H4R3me2	–	Histone 4 Arginine 3 dimethylated (Диметильований 3й залишок аргініну гістону H4)
LTR	–	Long terminal repeats (довгі кінцеві повтори)
p-TEFb	–	positive transcription elongation factor (позитивний фактор елонгації транскрипції)

## ЗМІСТ

<b>ВСТУП</b>	5
<b>РОЗДІЛ 1. Роль BRD4 у нормі та патології</b>	7
1.1. Роль BRD4 у нормі	9
1.2. Роль BRD4 у патології	13
1.3. Роль BRD4 у регуляції життєвого циклу ВІЛ	16
1.4. Роль та функції інших білків BET родини	17
<b>РОЗДІЛ 2. Матеріали та методи досліджень</b>	21
2.1. Програмне забезпечення та бази даних	21
2.2. Вибір референсного ліганду та формування бібліотеки сполук	21
2.3. Підготовка структур білків та лігандів	24
2.4. Проведення молекулярного докінгу бібліотеки речовин	24
2.5. Створення фармакофорної моделі	25
<b>РОЗДІЛ 3. Результати досліджень та їх обговорення</b>	26
3.1. Результати підготовки структур для докінгу	27
3.2. Результати докінгу	28
3.3. Аналіз фармакофорної моделі	30
<b>ВИСНОВКИ</b>	39
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</b>	40

## ВСТУП

Великий скачок у середній тривалості життя протягом останнього століття визначений значним розвитком біологічних та медичних наук. Серед них фармацевтика посідає особливе місце, бо постачає тисячі препаратів, що використовуються при терапії багатьох патологічних станів. На даний момент існують препарати синтетичного (малі органічні молекули) та біологічного (антитіла, ДНК-вакцини і т.д.) походження. Синтез низькомолекулярних органічних сполук є більш контрольованим та менш затратним процесом на відміну від отримання препаратів біологічної природи, тому він є основою сучасної фармацевтики. Методи тонкого органічного синтезу та медичної хімії у парі з хемо- та біоінформатикою допомагають отримувати речовини з вже заданими хімічними характеристиками та біологічною активністю. Ці сполуки являються основою лікарських препаратів, що використовуються при лікуванні багатьох патологічних станів.

BRD2, BRD3, BRD4, BRDT – білки, що відносяться до родини BET (bromo- and extraterminal) білків [1]. Основною відмінністю даних білків є наявність бромодоменів – білкових структурних одиниць, що можуть зв'язуватися з ацетильованим лізином гістонових та негістонових білків, наприклад факторів транскрипції [1]. Завдяки даній особливості ці білки можуть впливати на деякі етапи транскрипції різних генів, інгібуючи або активуючи їх експресію, окрім цього.

Зокрема, BRD4, який вивчений краще, ніж інші білки цієї родини, залучений, також, у процесі розвитку злоякісної трансформації [1]. Оскільки всі білки впливають на експресію генів і можуть бути драйверами патологічних процесів наразі дуже важливо шукати речовини, що є ефекторами даних білків. Оскільки білки є родиною, то ,можливо, деякі речовини будуть взаємодіяти з усіма представниками. Дана робота присвячена пошуку даних активних речовин, а також порівняння їх специфічності до

різних білків родини. Метою роботи є пошук речовин, які зв'язуються з білками родини ВЕТ методами структурної біоінформатики. Відповідно до мети були поставлені наступні задачі:

1. Із загальнодоступних баз даних хімічних сполук відібрати речовини на основі референсного ліганди для подальшого молекулярного докінгу проти всіх представників родини ВЕТ.
2. Отримати та оптимізувати структури білків та лігандів.
3. Провести молекулярний докінг з відібраними речовинами.
4. Розробити модель рецептор орієнтованого фармакофору лігандів під кожен домен.

## РОЗДІЛ 1

### РОЛЬ BRD4 У НОРМИ ТА ПАТОЛОГІЇ

BRD4 є найбільш дослідженим білком серед всієї BET родини, тому ми розглянемо його функції окремо.

Ацетилювання лізинових залишків в молекулах гістонів – одна з основних епігенетичних модифікацій. Окрім ацетилювання, гістони можуть піддаватися метилюванню та фосфорилуванню. Завдяки таким модифікаціям на епігенетичному рівні контролюється транскрипційна активність геному. Саме хвости гістонових білків є субстратами для даних модифікацій. Дані модифікації є по суті мітками, які впізнаються іншими білками: факторами транскрипції, регуляторами, ферментами тощо. Таким чином на гістонових хвостах можуть збиратися безліч білків, склад яких залежить від своєрідного розподілу модифікованих груп по хвостах. В свою чергу, білковий склад на гістонових хвостах визначає функціональні наслідки.

Співвідношення між розподілом модифікованих груп по хвостах (патерном модифікації) та набором певних білків називається гістоновим кодом. При ацетилюванні відбувається перенос залишку оцтової кислоти на залишок амінокислоти лізину. Подібна модифікація екранує позитивний заряд лізину гістонових білків, тим самим звільняючи нуклеосому ДНК, ( взаємодія ДНК з гістоновими білками має електростатичну природу) роблячи цю дану ділянку ДНК транскрипційно активною.

Ацетилювання гістонових хвостів відбувається за участю гістон-ацетилтрансфераз – одна з груп білків, що «пишуть» гістоновий код, тому ці білки ще називаються райтерами (англ. Writers). Деацетилювання гістонів відбувається за участі гістон-деацетилаз. Існують білки зчитувачі (readers), що можуть впізнавати ацетильовані гістони. Впізнають вони ці гістони через наявні в них бромодомени.

Бромодомен – консервативний білковий домен, що складається зі 110 амінокислотних залишків та впізнає ацетильовані залишки лізину як в

гістонових, так і в негістонових білках. В клітинах ссавців бромодомен сприяє активації транскрипції, зв'язуючись з ацетильованими залишками лізину на хвості гістону

На даний момент відомо 61 тип бромодоменів, які надалі можна згрупувати в 8 підтипів (наприклад родина білків BET) на основі гомології їх функцій та послідовностей. Білки, що мають бромодомен (BRDs) слугують епігенетичними зчитувачами ацетилювання. Ці білки можуть рекрутувати комплекси регуляції транскрипції до ділянки хроматину та зв'язувати ацетильовані гістони. BRD4 є найбільш вивченим членом родини білків BET(bromo and extraterminal) [1]. BET включає в себе BRD2, BRD3, BRD4 та BRDT.

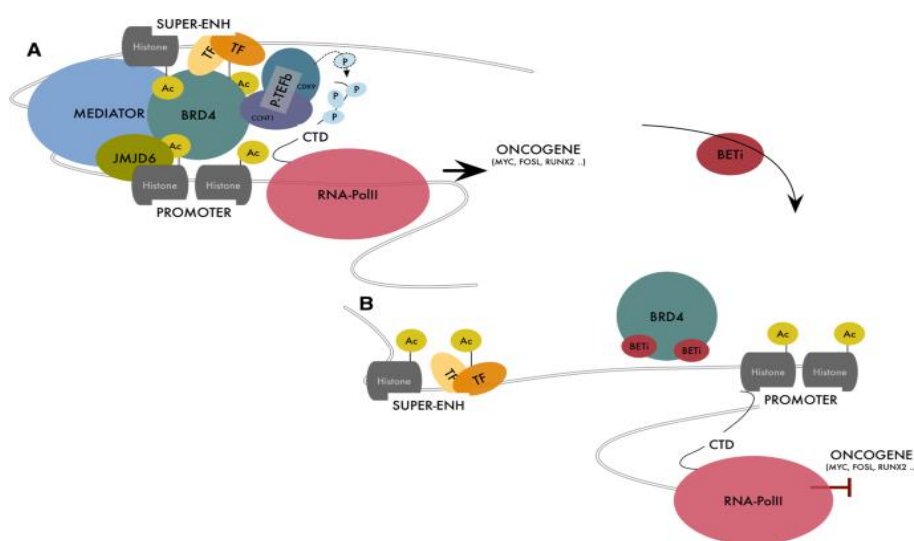
Ген BRD4 кодує, як мінімум, 2 ізоформи білку завдяки альтернативному сплайсингу – довгу та коротку. Довгий варіант BRD4(1 – 1362 амінокислотних залишків) має два консервативних бромодомена, термінальний кластерний домен, декілька консервативних мотивів (A, B, SEED) та С-термінальний домен (ET-домен) [2], а коротка форма (1-722 амінокислотних залишків) не має С-термінального домену, а також декількох ділянок, необхідних для формування зв'язку BRD4 з позитивним фактором елонгації транскрипції (3). BRD4 та інші білки BET зв'язуються з гіперацетильованими ділянками гістонів, оскільки білки BET мають більшу афінність до білків з кількома ацетильованими залишками. Таким чином BRD4 скупчується на транскрипційно активних елементах, сприяючи ініціації та елонгації транскрипції шляхом рекрутування позитивного фактору елонгації транскрипції (p-TEFb) [4]. p-TEFb фосфорилує серин 2 домену на С'-кінці РНК-полімерази 2, яка необхідна для елонгації транскрипції. Окрім епігенетичної регуляції, BRD4 бере участь у переході клітини з М-фази до G1 фази клітинного циклу. Також BRD4 бере участь реплікації та репарації ДНК, а також у ремоделюванні хроматину. Дисфункція білків родини BET, в тому числі BRD4 приводить до різних патологічних станів, серед запуск експресії

деяких онкогенів, таких як c-Мус, NF-κB, Aurora B та Bcl-2, залучаючи фактор транскрипції p-TEFb [5, 6]

### 1.1. Роль BRD4 у нормі

Ініціація транскрипції починається з рекрутування РНК-полімерази 2 до преініціаторного комплексу на промоторі відповідного гену, після чого відбувається фосфорилування серину у 5-му положенні. Збір преініціаторного комплексу піддається впливу з боку енхансеру і контролюється транскрипційними факторами. Енхансер – дистальна ділянка промотору для ДНК-полімерази 2, що споріднена до деяких факторів транскрипції, які сприяють збиранню преініціаторного комплексу та ініціації транскрипції. Зв'язок між транскрипційним факторами на енхансері та РНК-полімеразою забезпечує медіатор – білковий комплекс, що передає сигнал від енхансеру до РНК-полімерази.

На даний момент вважається, що BRD4 є «кофактором» для медіаторного комплексу [7, 8] (рис. 1.1).



**Рис. 1.1.** Схематична репрезентація функції BRD4 у ініціації транскрипції[4]

Було показано, що BRD4 та медіатор разом локалізуються в зонах енхансеру та супер енхансеру (ділянка геному, що включає безліч енхансерів, пов'язаних великою кількістю факторів транскрипції) і ця їх локалізація порушується інгібуванням BET. При введенні інгібіторів білків BET, медіатор відокремлюється від регуляторних факторів, з якими він був зв'язаний (рис. 1.1) [9, 7].

Крім медіатору, багато білків ремоделювання хроматина та факторів транскрипції можуть зв'язуватися з BRD4 [10, 11]. Транскрипційні фактори рекрутують гістон-ацетилювальні ферменти р300/CBP, які ацетилюють гістони та/або негістонові білки (ті ж самі фактори транскрипції). BRD4 в свою чергу може чіплятися як на ацетильовані лізинові залишки гістонів, так і на фактори транскрипції. Нещодавня модель свідчить, що одночасне зв'язування BRD4 з факторами транскрипції та з гістонами стабілізує зв'язок транскрипційних факторів з медіатором, що сприяє підвищеній транскрипційній активності відповідної ділянки. Взаємодія BRD4 з факторами транскрипції та іншими білками може бути або опосередкована бромодоменами, тому ці взаємодії детермінуються рівнем ацетилювання мішеней BRD4 [11, 12], або за допомогою інших доменів BRD4, а отже на такі взаємодії не впливає лікування інгібіторами BET, бо структура багатьох BET-інгібіторів імітує структуру ацетильованого лізину, щоб конкурентно зв'язувати бромодомен [13]. Було показано, що ET-домен BRD4 може взаємодіяти з модифікаторами гістонів, такими як аргінін-деметилаза JMJD6, [14], метил-трансфераза NSD3 та білками ремоделювання хроматину SWIF/SNF та CDH4 [13]. Завдяки цим взаємодіям, BRD4 може впливати на активність транскрипції через регулювання декомпактизації хроматину.

Також було встановлено, що BRD4 взаємодіє з багатьма факторами транскрипції не через бромодомени, наприклад з p53, YY1, AP2, c-JUN, c-MYC. Окрім цього BRD4 бере участь у запуску програми запального

процесу, що керується NF-κB, шляхом зв'язування з ацетильованим RelA [11] та диацетильованим TWIST при раку молочної залози

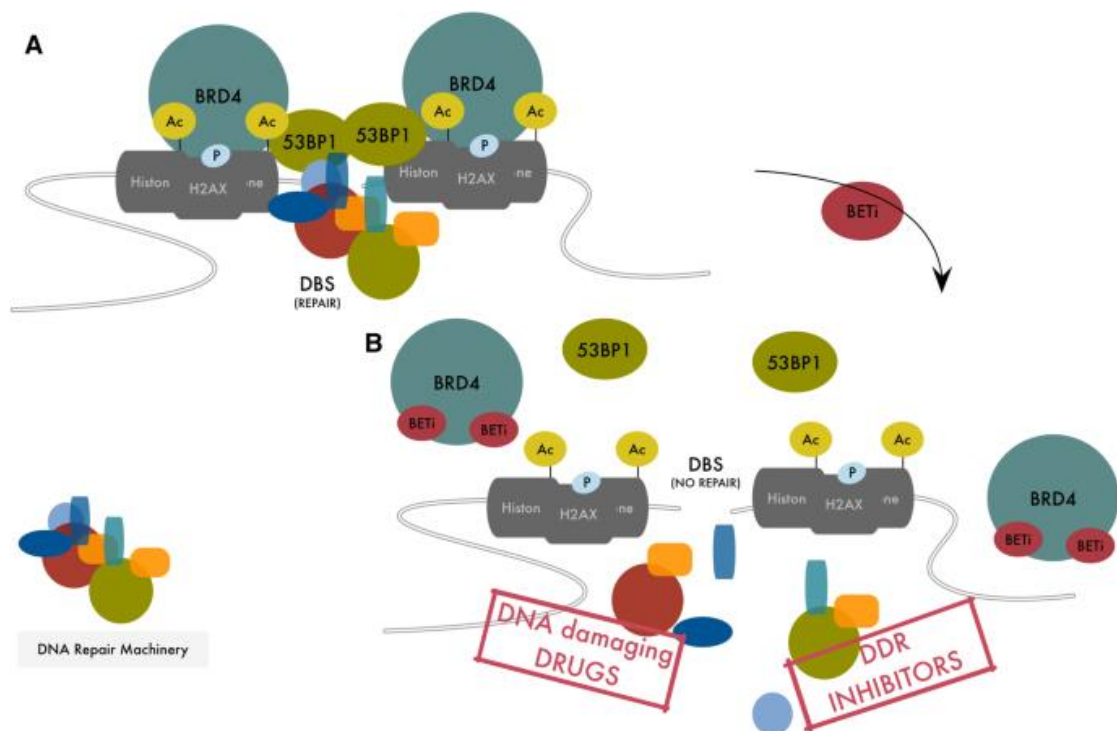
Зупинка РНК-полімерази в проксимальних ділянках її промотора являється лімітуючою стадією синтезу матричної РНК [15]. Для «активації» РНК-полімерази та її вступу в етап елонгації залучається комплекс p-TEFb, який складається з цикліну T1 (CCNT1) та циклінзалежної кінази 9 (CDK9). В своїй активній формі p-TEFb взаємодіє з факторами транскрипції та фосфорилує фактор (наприклад NELF), який заважає полімеразі продовжити нарощувати ланцюг РНК, тим самим знімаючи його гальмівний вплив. При цьому p-TEFb фосфорилує хвіст РНК-полімерази по серину у другому положенні, активуючи транскрипцію. Кіназна активність p-TEFb інгібується нуклеопротейновим комплексом 7SK/HEXIM (рис. 1.2) (HEXIM – білок, що зв'язується з циклінзалежною кіназою комплексу p-TEFb і тим самим інгібує її активність), але взаємодія BRD4 з p-TEFb попереджує зв'язування останнього з 7SK/HEXIM. Велике накопичення BRD4 на гіперацетильованих (а тому транскрипційно активних) сайтах є пунктом його зв'язування з p-TEFb, що в свою чергу веде до елонгації транскрипції [16].

Також BRD4 вибірково локалізується разом з JMJD6 на деяких енхансерах з високими рівнями H3K4Me1 та H3K27Ac (назви епігенетичної модифікації гістону H3 – метилювання лізину у 4му положенні та ацетилювання лізину у 27 положенні відповідно). BRD4 рекрутує JMJD6 до цих сайтів, яка в свою чергу деметилює ди-та монометиларгінінгістон гістон H4 (H4R3me1, H4R3me2), які є маркерами репресії транскрипції. Також JMJD6 деметилює 5'-метилфосфатний кеп 7SK RNA, що спричиняє її деградацію і активацію p-TEFb.

Порушення процесу репарації дволанцюгових розривів ДНК веде до нестабільності геному, що значно підвищує ризик виникнення раку [17]. Зміни в структурі хроматину, що в більшості випадків опосередковуються модифікуванням гістонів, являються ініціатором у запуску відповіді на пошкодження ДНК [18]. BRD4 являється важливою фігурою у запуску

програм репарації ДНК і робить він це декількома шляхами. BRD4 опосередковано може впливати на процеси репарації, регулюючи транскрипційну активність генів, що відповідають за відновлювання ДНК, але також має неопосередкований вплив на механізм виправлення пошкоджень ДНК [19].

Є дані, про роль BRD4 у негомологічному з'єднанні кінців [20]. BRD4 також необхідний для рекомбінації генів імуноглобулінів, а точніше – для переключення їх ізотипів [21]. Окрім цього, BRD4 потрібен для репарації дволанцюгових розривів, причиною яких стало іонізуюче випромінювання. При інгібуванні BRD4 відбувається фосфорилування гістону H2A, пошкодження ДНК набуває стійкого характеру, що веде до нестабільності геному (рис.1.2).



**Рис. 1.2.** Роль BRD4 у репарації дволанцюгових пошкоджень [4]

При дволанцюгових пошкодженнях ДНК, окрім фосфорилування гістону H2A, відбувається посилене ацетилювання гістону H4, тому в цих місцях локалізується багато молекул BRD4, які кріпляться до модифікованих гістонів

по обидві сторони від розриву, являються платформою для збірки машинерії репарації ДНК.

На сьогодні відомо, що одним з білків, що входять до комплексу репарації ДНК є p53. P53 не має власної ферментативної активності, але він прямує до місця розриву через специфічний гістоновий код і допомагає рекрутувати інші білки, що залучаються в програмі відповіді на пошкодження ДНК [22].

## 1.2. Роль BRD4 у патології

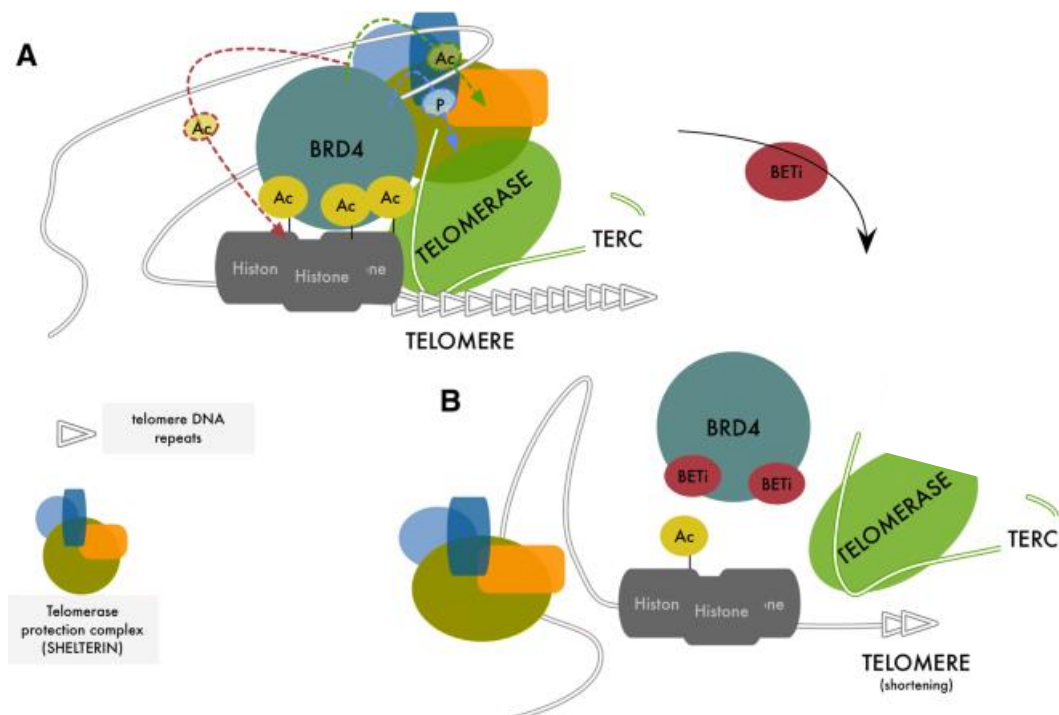
С кожним клітинним поділом довжина дволанцюгової ДНК скорочується. Оскільки це, знову ж таки, може призвести до нестабільності геному та, як наслідок, виникнення мутацій, на кінцях плечей кожної хромосоми знаходяться теломери – ділянки з послідовностями що повторюються. З кожним поділом клітини теломери скорочуються, і коли це скорочення досягає критичної точки, клітина перестає ділитися. В стовбурових та ракових клітинах наявний фермент теломераза, яка після кожного поділу добудовує втрачений фрагмент ДНК, тому ці типи клітин можуть ділитися постійно [10].

Теломерна зворотня транскриптаза (ТЗТ) є каталітичною субодиницею теломеразного комплексу. Є роботи, що вказують на певну роль BRD4 регуляції росту теломер при ракових захворюваннях, хоч молекулярні механізми цього явища ще не до кінця відомі. Але було виявлено, що BRD4 приймає участь в аберантній регуляції транскрипції гену теломерної зворотньої транскриптази при ракових захворюваннях. Були виявлені дві мутації в промоторі гену ТЗТ, асоційовані з підвищеним ризиком ракового захворювання агресивної природи та невтішними прогнозами [23, 24]. Ці мутації викликають появу сайтів зв'язування для транскрипційних факторів E-Twenty-Six (ETS), які підвищують експресію гену ТЗТ [25]. Один з транскрипційних факторів ETS, GABRA, специфічно зв'язується з мутантним промотором ТЗТ, регулюючи транскрипцію даного гену. До того ж GABRA

сприяє модифікуванню гістонів по шаблону, який полегшує транскрипцію на даній ділянці [26]. При цьому, автори даної роботи вважають, що BRD4 грає роль в цьому процесі як опосередковано, так і прямо. BRD4 прямує до промотору ТЗТ і зв'язується з гіперацетильованими гістонами і стабілізує взаємодії різних ділянок хроматину, ініційованих GABRA. Також BRD4 зв'язується з мутантним промотором ТЗТ, підтримуючи організацію компонентів необхідних для початку транскрипції.

Лікування інгібіторами BET, а також нокдаун гену BRD4 сприяє зниженню експресії гену ТЗТ порушенню теломеразної активності. Разом з цим знижується рівень модифікуванню гістонів по шаблону активації транскрипції [26].

Але при цьому всьому, на сьогодні не до кінця відомий механізм, за яким BRD4 підтримує постійний ріст теломер. На даний момент висунута лише модель (рис. 1.3).



**Рис. 1.3.** Схема можливого механізму впливу BRD4 на ріст теломер [4]

За цією моделлю BRD4 слугує своєрідною платформою для рекрутування та стабілізації теломеразного комплексу та інших теломер-асоційованих комплексів, що сприяють росту теломер шляхом накопичення ацетильованих міток на хвостах гістонових білків [27].

Роль BRD4 у розвитку багатьох захворювань (в тому числі ракових) пов'язана з його регуляцією запального процесу. Так цей білок важливий для IL-1-індукованого запалення в епітеліальних клітинах повітряних шляхів людини [28]. Було показано, що інгібування BRD4 подавляє експресію прозапальних цитокінів при парадонтиті [29].

NF-κB – важливий транскрипційний фактор, активація якого запускає експресію генів багатьох прозапальних цитокінів [30]. Деякі дослідження показали, що роль BRD4 у запуску деяких програм запалення тісно пов'язана з сигнальним шляхом NF-κB [31]. Інгібування BRD4 веде до зниження продукції прозапальних цитокінів шляхом зниження загального рівня активованих факторів NF-κB [30].

В деяких випадках у якості триггеру ініціації канцерогенезу виступає запальний процес. Багато досліджень демонструють сильну асоціацію між хронічним запаленням метаболічної або інфекційної етіології та злоякісною трансформацією і початком канцерогенезу [32, 33]

В процесах очищення організму від інфекційних агентів, продуктів метаболізму паразитів та при звичайній регенерації тканин після їх пошкодження має місце хронічний запальний процес, який сприяє швидкій тканинній проліферації, реплікації ДНК, порушенню процесів її відновлення, що веде до злоякісної трансформації [34].

Було досліджено, що у хворих на цукровий діабет другого типу можуть виникати злоякісні новоутворення через розвиток запального процесу. BRD4 грає роль у цьому процесі, підсилюючи транскрипцію гену RORC, який є дуже важливим для розвитку адипоцитів, диференціації Th T-лімфоцитів. RORC підвищує транскрипцію генів прозапальних цитокінів – інтерлейкінів 17A, 17F та 22, які можуть ініціювати канцерогенез [34].

### 1.3. Роль BRD4 у регуляції життєвого циклу ВІЛ

Вірус імунодефіциту людини до сих пір є великою проблемою з глобальними наслідками. На даний момент інфіковано приблизно 38 мільйонів людей по всьому світу. Єдиним методом боротьби з даним захворюванням є антиретровірусна терапія (АРТ), що подавляє активну реплікацію вірусу і знижує концентрацію вірусних часток в плазмі крові до значень нижче чутливості діагностичних методів. Але при цьому АРТ не виліковує повністю людину від даного недугу, тому що вірус постійно створює стабільні латентні резервуари своєї генетичної інформації в геномі людини [35]. Припинення АРТ практично завжди приводить до вірусного рецидиву вже через 2-3 тижні [36].

Після інтеграції вірусної НК в геном людини, вірус знаходиться в стадії провірусу і його експресія регулюється епігенетичними механізмами хазяїна [37]. В останні роки нові потенційні шляхи лікування ВІЛ направлені на ці механізми, щоб викликати експресію цих ділянок геному або навпаки – заблокувати її (більш глибока латентність).

З 2005 року з'явилося багато досліджень, що підтвердили участь BRD4 у регуляції транскрипції генів ВІЛ, але самі результати різняться. Більш ранні дослідження показали, що BRD4 сприяє транскрипції генів ВІЛ [16], відносно недавні вказують на те, що BRD4 інгібує експресію вірусних генів [3]. В підтримку останніх було показано, що інгібування BRD4 інгібітором JQ1 призводить до активації транскрипції генів ВІЛ [38, 39]. Через свої бромодомени BRD4 може мобілізуватися до промотору ВІЛ через ацетильовані лізинові залишки гістонів. Для ініціації транскрипції вірусної генетичної інформації потрібен білок Тат (трансактиватор транскрипції), що кодується самим вірусом. Тат взаємодіє з Р-TEFb, що є компонентом Super

Elongation Complex (SEC) – комплекс факторів транскрипції для РНК полімерази 2. Tat використовує SEC у якості «кофактору» для транскрипції вірусного геному. BRD4 конкурує з Tat за обмежену кількість Р-TEFb і тим самим виступає інгібітором експресії вірусних генів [16].

Нещодавно було виявлено, що взаємодія BRD4 з гістонами H3 та H4 має різні ефекти на регуляцію транскрипції ВІЛ [40]. Було досліджено статус ацетилювання гістонів, що присутні на послідовностях LTR провірусу ВІЛ, які мали різні сайти інтеграції. Результати показують підвищений рівень BRD4 та ацетильованих гістонів H4 і різко знижений рівень ацетильованого гістону H3 при латентному періоді ВІЛ, що вказує на те, що ацетильований H4 є тим первинним маркером, що рекрутує BRD4 до промотеру та утримує його там [40].

Є три ацетилтрансферази, які модифікують гістон H4 і які відносяться до типу лізин ацетил трансфераз (ЛАТ) – ЛАТ5, ЛАТ 7 та ЛАТ8. Зіконг та співавт. дослідили, що інгібування саме ЛАТ5, але не ЛАТ7 та 8, знижує рівень ацетильованого H4 та BRD4 з LTR провірусу, що веде до активації Tat, залучення SEC, початку транскрипції та виходу з латентного режиму [40].

Але це не єдиний механізм впливу на експресію генів ВІЛ – було показано, що коротка ізоформа BRD4 може рекрутувати комплекс ремоделювання хроматину SWI/SNF на промотор ВІЛ та інгібувати транскрипцію з цього промотора [3].

#### **1.4. Роль та функції інших білків BET родини**

BRD2 є частиною h1lyb білків з бромо- та екстратермінальним доменами (BET), до якого входять Brd3, Brd4 та BrdT. BET розглядаються як посередники в розвитку низки видів ракових захворювань [41]. Відомо, що ці білки є головними регуляторами транскрипції і слугують для з'єднання

хроматинових міток (переважно ацетил-лізинів) з транскрипційним механізмом.

BRD2 може відігравати важливу роль у зв'язку між толл-подібним рецептором і TNF-сигналом та NF- $\kappa$ B-індукованою транскрипцією генів запалення. У цих дослідженнях гіпоморфні за BRD2 миші не мають інсулінорезистентності, пов'язаної із запаленням та ожирінням. Нарешті, існують докази того, що білки BET можуть також відігравати роль в аутоімунних захворюваннях, опосередкованих T-хелперами [42].

BRD2 необхідний для експресії *ACE2* у низці різних систем, пов'язаних з SARS-CoV-2. Ми також виявили, що лікування препаратом ABBV-744, інгібітором бромодому, може знизити концентрацію вірусної РНК SARS-CoV-2 у первинних клітинах носового епітелію людини та сирійських хом'яків. Ці дані свідчать про те, що фармакологічні інгібітори BRD2 можуть мати терапевтичну користь для запобігання або зменшення впливу інфекції SARS-CoV-2 [43].

За даними деяких досліджень, BRD2 є опосередкованим регулятором транскрипції *ACE2* і, таким чином, він регулює проникнення вірусу в клітини, релевантні до COVID-19 [43]. Це узгоджується з двома іншими нещодавніми дослідженнями, які показують, що інгібітори BET-домени можуть запобігати інфікуванню SARS-CoV-2 в культурі та знижувати експресію гена *ACE2* [44,45]. Наші дані свідчать про те, що BRD2 є найбільш вірогідною мішенню для інгібіторів BET-домени, які використовувалися в цих дослідженнях, на відміну від інших білків, що містять BET-домен, таких як BRD4.

Залучення BRD3 до раку показало його роль у певних транслокаціях NMC, і BRD3 може бути потенційно пов'язаний з онкопротеїнами злиття MLL у лейкемогенезі [46].

Дослідження показали, що BRD2, BRD3, BRD4 і BRDT виконують контекстно-залежні, відмінні і часто протилежні функції. Наприклад, у клітинах раку молочної залози було продемонстровано різні ролі білків BET у функції енхансера ER $\alpha$  та регуляції генів. У той час як BRD2, BRD3 і BRD4 виконують частково дублюючі функції в енхансерах ER $\alpha$  і транскрипції генів, виявлено більш унікальну роль BRD3 для ER $\alpha$ + раку молочної залози. BRD3 рекрутується до підмножини ER-зв'язуючих сайтів (ERBS), збагачених

енхансерами у кластерах, які, ймовірно, функціонують як супер-енхансери та асоціюються з генами, чутливими до високого рівня естрадіолу (E2) [47].

При потрійному негативному раку молочної залози (TNBC) BRD2 позитивно регулює епітеліально-мезенхімальний перехід (EMT), тоді як BRD3 і BRD4 функціонально пригнічують цю програму за допомогою різних транскрипційних програм [48].

Молекулярна основа відмінностей у функціях білків BET в межах однієї лінії залишалася невловимою. Нещодавнє дослідження провело інтерактомне профілювання взаємодіючих білків BRD2, BRD3, BRD4 та BRDT і отримало значні відомості про їхні селективні взаємодії. Як і очікувалося, автори підтвердили наявність деяких спільних інтеракторів для всіх білків BET. Зокрема, вони відзначили, що комплекс p-TEFb і NELF зазвичай асоціюється з BRDT і BRD4, компоненти негативних регуляторів транскрипції, такі як комплекс NURD, асоціюються з BRD3 і BRD4, тоді як субодиниці РНК Pol-II і субодиниці Mediator були ідентифіковані лише як такі, що переважно асоціюються з BRD4. Ці дані підкреслюють унікальну, але універсальну функцію BRD4, що свідчить про складну взаємодію між білками BET [49].

Нарешті, BRD2 та BRD3 мають подібну доменну структуру та функції до BRD4. Раннє дослідження з використанням флуоресцентного резонансного переносу енергії (FRET) для визначення взаємодії між BRD2, міченим блакитним флуоресцентним білком (CFP), та гістонами, міченими жовтим флуоресцентним білком (YFP), дало важливе розуміння специфічності розпізнавання ацетильованих гістонів BRD2. Зокрема, BRD2 взаємодіє з гістоном H4K5/K12 і меншою мірою з H2B, але не з іншими гістонами [50].

Дослідження показують, що тканинно-обмежений, хроматин-асоційований білок BRDT, що експресується в деяких сперматоцитах, відіграє істотну роль у сперматогенезі опосередковану першим бромодоменом, який зв'язує тетраацетильований амінотермінальний хвіст гістону 4 (H4Kac4) [51]. Структурні дослідження мишачого BRDT показали, що BRDT(1) зв'язує діацетильований пептид гістону 4 (H4K5ac8ac) частково через консервативний аспарагін.

BRDT діє як компонент механізму сплайсингу в пахітенових сперматоцитах і круглих сперматозоїдах. Необхідний для організації хромоцентру - структури, що складається з перицентромерного гетерохроматину [51].

Цей білок грає роль в деяких патологічних процесах, що включають безпліддя у чоловіків [51].

## РОЗДІЛ 2

### МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕНЬ

#### 2.1. Програмне забезпечення та бази даних

В нашому дослідженні було використано наступне програмне забезпечення: GOLD [52] для молекулярного докінгу, PyMol [53] для візуалізації білкових та білково-лігандних комплексів, LigandScout [54] – для побудови фармакофорної моделі, Discovery Studio Visualizer 2021 [55] для візуалізації лігандів, зміни формату файлів для збереження лігандів та оптимізації їх геометрії. Інформація про ліганди бралися з баз даних ZINC [56] та ChEMBL [57]. Структури білків були взяті з бази даних AlphaFold [58] та Protein Data Bank [59].

#### 2.2. Вибір референсного ліганду та формування бібліотеки сполук

Пошук речовин та формування скринінгових бібліотек для проведення молекулярного докінгу відбувалася за допомогою бібліотеки мови програмування Python, використовуючи бібліотеку RDKit [60]. База для формування бібліотеки – сполуки з бази даних ZINC.

Референсною сполукою виступав відомий ліганд BRD4 ZINC000060085058 (PDB код – 08K); його 3D структура була взята з бази даних ZINC.

На основі референсного ліганду було проведено Similarity Search, використовуючи морганівські фінгерпрінти. Для цього була використана бібліотека мови програмування Python RDKit. Через цей код пройшло більше 2 мільйонів сполук з бази даних ZINC.

```
[ ] import pandas as pd
import numpy as np
import seaborn as sn
#----- RDKit packages
from rdkit.Chem import AllChem
from rdkit.Chem import rdMolDescriptors
from rdkit.Chem.Draw import IPythonConsole
from rdkit.Chem import Draw
from rdkit import DataStructs
from rdkit.ML.Cluster import Butina
#----- progress bar
from tqdm import tqdm
#----- hide warning
import warnings
warnings.filterwarnings('ignore')
from rdkit import Chem

[ ] chembl_mols=pd.read_csv('/content/drive/MyDrive/cheminformatics/allchembl.csv',sep=';', on_bad_lines='skip')
chembl_mols.head()
df=chembl_mols[['CHEMBL ID','#RO5 Violations (Lipinski)','Smiles']]

[ ] df['Smiles']=df['Smiles'].astype(str)

[ ] from rdkit.Chem import PandasTools
PandasTools.AddMoleculeColumnToFrame(df,'smiles', 'Structure')

[ ] problematicSmiles = df[df['Structure'].map(lambda x: x is None)]

df = df.drop(problematicSmiles.index,axis=0)

lead = AllChem.MolFromSmiles('Cc1nnc2n1-c1ccc(Cl)cc1C(c1ccccc1)=NN2C')
lead_fps = AllChem.GetMorganFingerprintAsBitVect(lead, 2, nBits=4096)
def SimSearch(data):
    ids = []
    for smile in data:
        mol = Chem.MolFromSmiles(smile)
        if mol!= None:
            Mfpts = AllChem.GetMorganFingerprintAsBitVect(mol, 2, nBits=4096)
            Tanimoto_similarity = DataStructs.FingerprintsSimilarity(lead_fps, Mfpts, metric=DataStructs.TanimotoSimilarity)
            if Tanimoto_similarity >=0.5:
                ids.append([smile,Tanimoto_similarity])
    print(ids)
```

**Рис. 2.1.** Приклад коду для проведення Similarity Search

Спочатку були імпортовані потрібні бібліотеки. Потім загрузка файлу з речовинами до коду. В сирому файлі міститься певна к-сть речовин в яких невалідні SMILES або їх там взагалі нема. Тому наступним кроком було видалення всіх проблемних рядків з сирих даних з речовинами. Далі були загружені SMILES референсної речовини та отримані її морганівські фінгерпринти. Після цього була створена функція, що приймала на вході речовину в форматі SMILES, конвертувала її в формат mol, отримувала морганівський фінгерпринт цієї сполуки та, порівнюючи її з референсом,

шукала коефіцієнт подібності (Tanimoto coefficient/ Tanimoto value) і, якщо цей коефіцієнт  $\geq 0.5$ , то функція додавала SMILES цих сполук та їх коефіцієнти до списку. Основуючись на даних наукових публікацій, для того, щоб бути впевненим, що сполука має схожі хімічні властивості, коефіцієнт має бути  $\geq 0.85$  [61]. Знайти сполуки з коефіцієнтом подібності більше 0.8 надзвичайно важко, тому було прийнято рішення зробити прохідним значення 0.5. Було відібрано на основі коду 18 сполуки, в яких коефіцієнт був вище або дорівнював 0.5. Всі вони були перевірені в базах даних ZINC та ChEMBL на наявність таргету до цієї речовини (Таблиця 2.1.). Загалом нас цікавили чи є серед мішеней цих лігандів білки родини ВЕТ людини.

**Таблиця 2.1.** Список речовин, що пройшли cut-off

zinc_id	tanimoto_values	Target existence
ZINC73295906	0.77	registered target
ZINC73295908	0.63	registered target
ZINC00000903	0.61	registered target
ZINC903 (Alprazolam)	0.61	registered target
ZINC3802399	0.60	no target
ZINC27721698 /ZINC27721706	0.60	no target
ZINC05161502	0.58	no target
ZINC02570882	0.58	no target
ZINC22046889	0.52	no target
ZINC22046885	0.52	no target
ZINC39247027	0.51	no target
ZINC27727533 /ZINC27727540	0.51	no target
ZINC27727522	0.5	no target
ZINC39207453	0.5	no target
ZINC27727522/ZINC27727527	0.5	no target

Загалом 14 сполук за даними ChEMBL та ZINC на момент 27 березня 2023 року не мали інформацію про таргет. Вони були вибрані для подальшого докінгу з 8 бромодоменами 4-ьох білків.

### 2.3. Підготовка структур білків та лігандів

Структури лігандів були взяті з бази даних ZINC у форматі SDF. Вони були конвертовані в формат mol2 та потім оптимізовані в програмному забезпеченні Discovery Studio Visualizer. Конвертація у формат mol2 є необхідною умовою для проведення докінгу, тому що програмне забезпечення GOLD приймає на вхід тільки mol2 та pdb формати.

Структури білків були взяті з бази даних AlphaFold. Причинами взяти структури саме звідти, а не з Protein Data Bank є те, що X-ray структури не мають атомів водню, також в кристалічних структурах іноді не достає деяких ланцюгів. Тому було прийнято рішення взяти структури з AlphaFold. Повні структури були візуалізовані в PyMol, де з них були вирізані окремі бромодомени, які потім були оптимізовані в програмному забезпеченні Ligand Scout в силовому полі MMFF94, а потім протонізовані. Коли структури білків були оптимізовані, вони були вирівняні на референс. Референсом послугувала кристалічна структура комплексу першого бромодому BRD4 з лігандом ZINC000060085058, на основі якого ми робили Similarity Search.

### 2.4. Проведення молекулярного докінгу бібліотеки речовин

Відібрана бібліотека речовин була продокована проти всіх восьми бромодоменів чотирьох білків. Був відпрацьований протокол докінгу за амінокислотними залишками в сайті зв'язування. Для цього спочатку був візуалізований активний сайт в PyMol для кожного бромодому. Після цього файл зі списком амінокислот, що знаходяться на відстані 5 Å від референсного

ліганду, був створений та використаний при формуванні протоколу для молекулярного докінгу. Для скорингу використовувалася функція CHEMPLP, для рескорингу – ASP. CHEMPLP – фітнес-функція, яка оцінює на скільки добре зі стеричної точки зору ліганд позиціонується в активному сайті білку. ASP це фітнес-функція, «базована на знаннях», вона оцінює взаємодії різних пар атомів, використовуючи дані про структуру молекули та схожі взаємодії з Protein Data Bank. Тип докінгу – напівригідний. Ліганди були налаштовані на максимальну лабільність, тоді як в структурі білка лабільними були лише бічні ланцюги залишків амінокислот. Результатами докінгу були по 10 рішень для кожної молекули, які в подальшому аналізувалися на предмет пози в активному сайті а також сортувалися за скорингом.

## **2.5. Створення фармакофорної моделі**

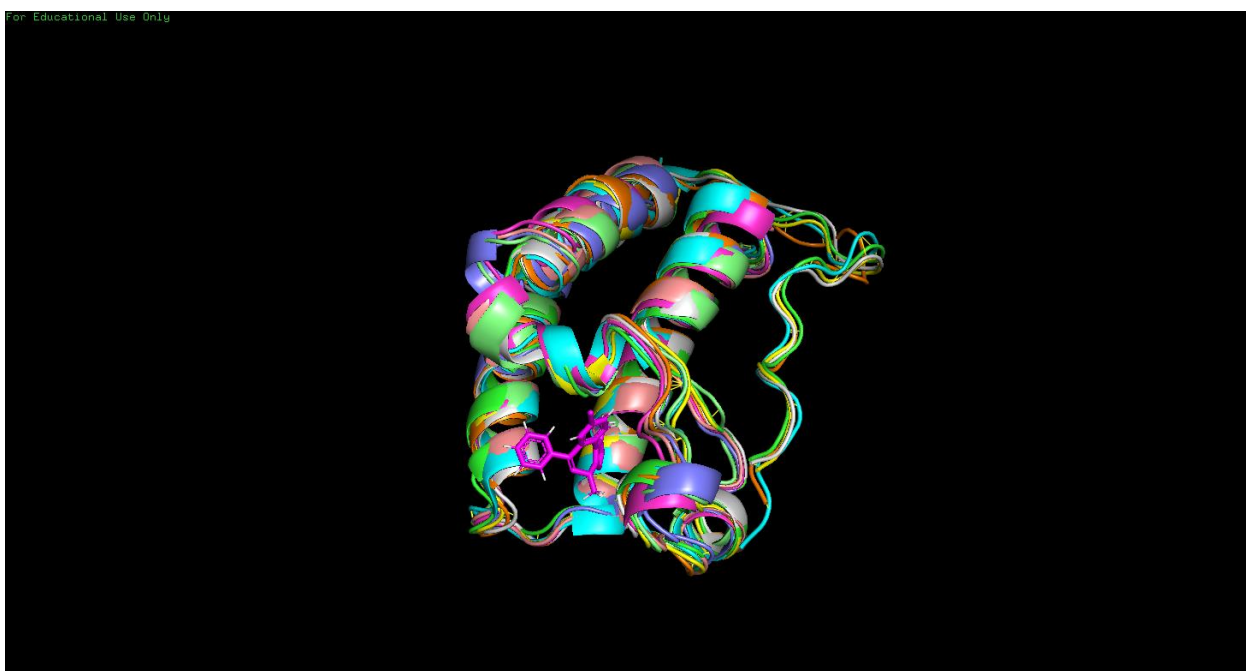
Для створення фармакофорної моделі використовувалося програмне забезпечення LigandScout. В програму були завантажені білкові структури бромодоменів проти яких ліганди, в межах постановки молекулярного докінгу, показали якісь взаємодії. Для кожного з цих бромодоменів був завантажений ліганд з найбільшим скором і був побудований рецептор-орієнтований фармакофор. Із всіх інших лігандів, що за результатами докінгу зв'язуються з даним доменом, була створена бібліотека, яка потім використалась у pharmacophore-based віртуальному скринінгу. Це було зроблено для того, щоб побачити, чи розділяються усі ліганди до цього домену спільні фармакофорні властивості.

## РОЗДІЛ 3

### РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ ТА ЇХ ОБГОВОРЕННЯ

#### 3.1. Результати підготовки структур для докінгу

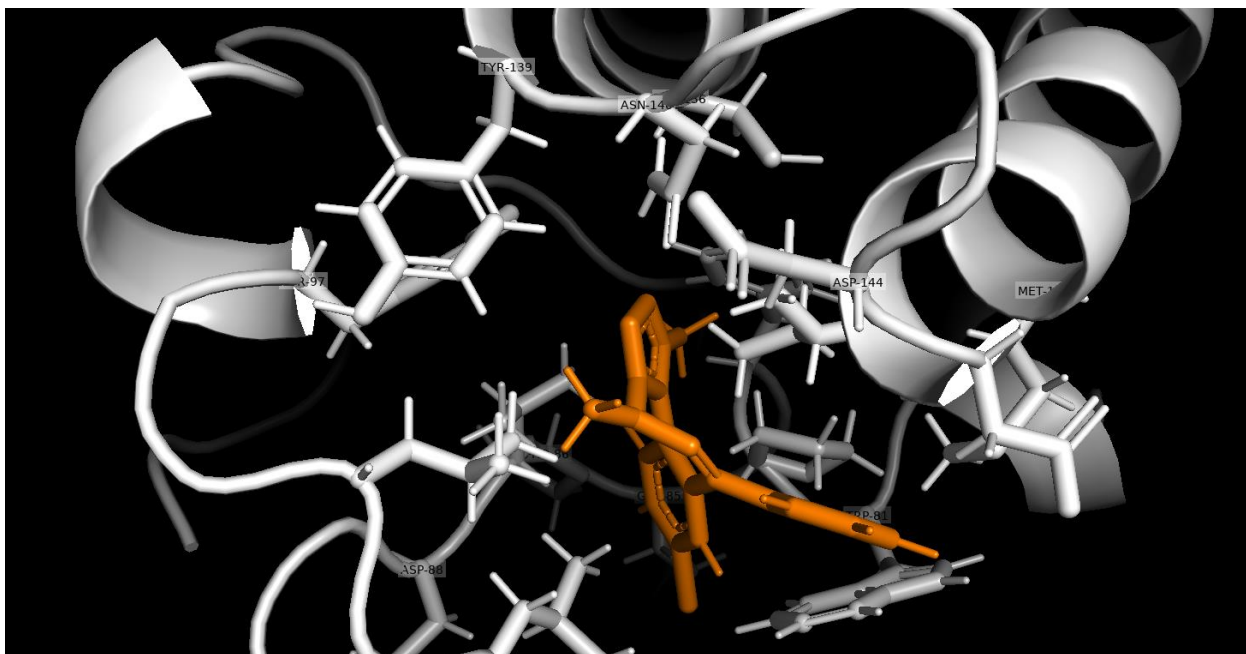
Структури білків були взяті з бази даних AlphaFold. Вони були завантажені та візуалізовані в програмному забезпеченні PyMol. Далі кожен домен усіх чотирьох білків був вирівняний на референсну кристалічну структуру першого бромодомену BRD4 (PDB 3U5L) з референсною речовиною для того, щоб “урівняти” їх координати для подальшої постановки молекулярного докінгу, RMSD коливалися від 0.7 до 1.2 ангстрем.



**Рис. 3.1.** Структурне вирівнювання бромодоменів до референсної структури з наявним лігандом

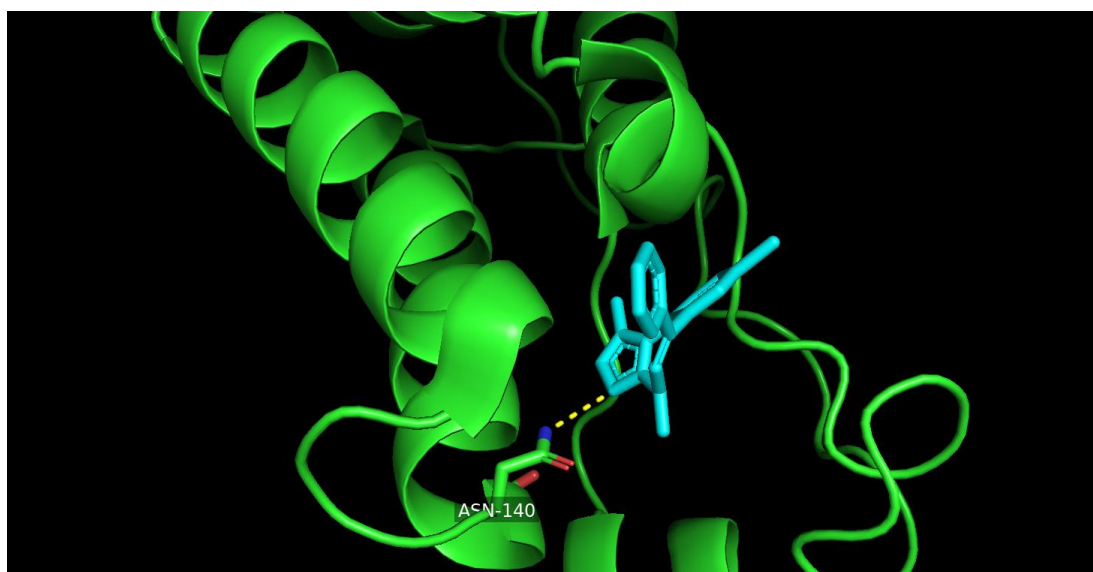
### 3.2. Результати докінгу

Перша спроба провести молекулярний докінг з використанням позицій референсного ліганду, як відправної точки, завершилася невдачею. Ліганди не “сідали” в активний сайт домену в тій же позі чи хоча б близькій до неї, що й референсний ліганд. Тому, було прийнято рішення визначити сайт зв’язування ліганду списком амінокислотних залишків в радіусі 5 Ангстрем від ліганду. Для цього в програмному забезпеченні PyMol кожен бромодомен був вирівняний з структурою першого бромодомену з референсним лігандом в ньому. Далі були “відмальовані” амінокислотні залишки відповідного бромодомену від ліганду на відстані 5 А (Рис.3.2).



**Рис.3.2.** Візуалізація амінокислотних залишків активного сайту brd4(bd1)

По кожному лігаду для всіх 6 доменів GOLD видав по 10 рішень, які були відсортовані за принципом схожості пози даного рішення з позою референсу у кристалічній структурі. Далі будуть розглянуті лише ліганди, які взаємодіють з білками подібно до референсної сполуки. А саме, відповідно до статті [62], триазольна група референсного ліганду (PDB код 08K) формує водневий зв'язок з амідною групою аспарагінового залишку (Рис.3.3).



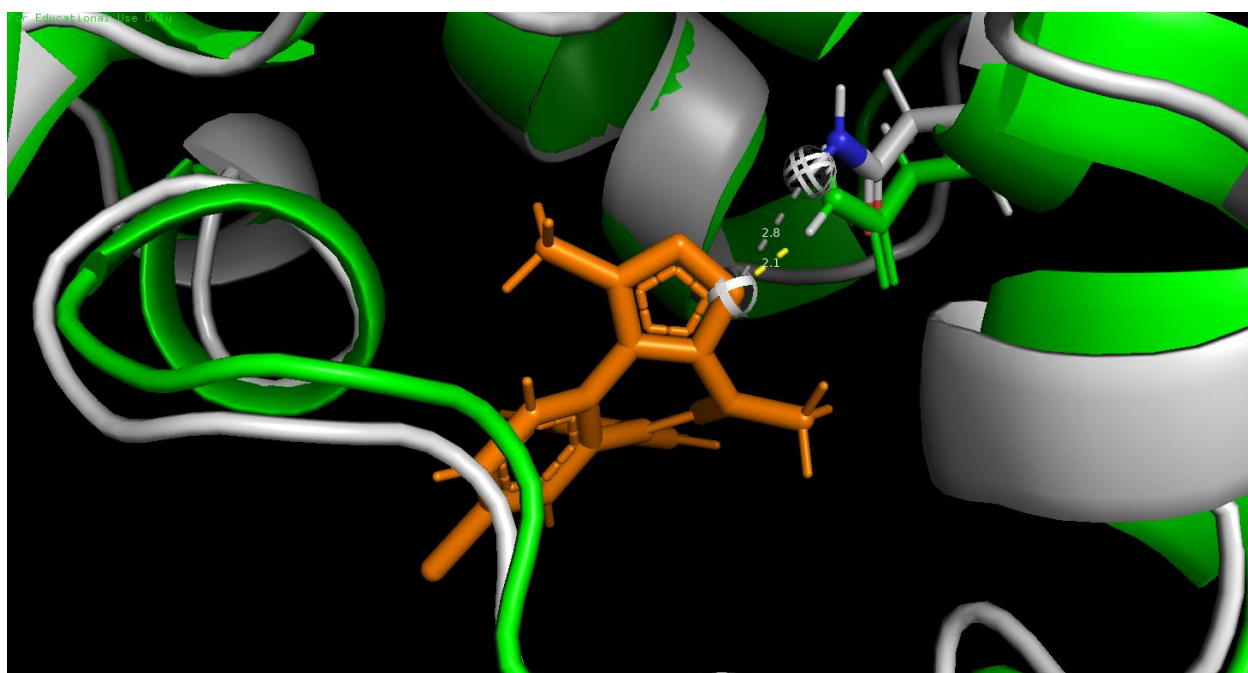
**Рис.3.3.** Візуалізація BRD4(bd1) з референсним лігандом 08K

Оскільки бромодомени ВЕТ родини доволі консервативні за послідовністю та структурою, є гіпотеза, що ця взаємодія з аспарагіновим залишком присутня в усіх бромодоменах родини.

**Таблиця 3.1.** Результати докінгу

Ligand	brd2_bd1		bdr3_bd1		brd4_bd1		brd4_bd2		brdt_bd1		brdt_bd2	
	Chemplp	ASP	Chemplp	ASP	Chemplp	ASP	Chemplp	ASP	Chemplp	ASP	Chemplp	ASP
08K_model	45.58	29.22	53.46	32.14	ni	ni	ni	ni	ni	ni	ni	ni
ZINC3802399	49.945	26.12	ni	ni	ni	ni	ni	ni	ni	ni	ni	ni
ZINC5161502	48.22	29.84	ni	ni	ni	ni	ni	ni	49.8	24.98	ni	ni
ZINC22046885	49.87	36.67	51.58	31.18	40.22	20.9	ni	ni	ni	ni	ni	ni
ZINC22046889	49.04	29.57	ni	ni	ni	ni	ni	ni	ni	ni	50.98	31.86
ZINC27721698	45.85	30.46	45.16	27.26	ni	ni	36.88	22.75	ni	ni	49.98	30.47
ZINC27721706	45.79	32.43	52.39	31.07	ni	ni	41.17	25.36	ni	ni	44.19	25.7
ZINC27727522	46.24	26.03	ni	ni	ni	ni	ni	ni	ni	ni	ni	ni
ZINC27727540	49.04	29.33	ni	ni	ni	ni	ni	ni	ni	ni	ni	ni
ZINC2570882	ni	ni	ni	ni	ni	ni	38.77	25.99	ni	ni	ni	ni

Загалом таким лігандів 9, три з них являються специфічні. Ліганди ZINC3802399 та ZINC27727527 (середній скоринг 49.945 та 46.235 відповідно) зв'язується подібним до референса чином тільки до brd2(bd1), ZINC2570882 чіпляється тільки до brd4(bd2) (скоринг 38.77). Щодо референсної сполуки, то посадка в активний сайт була зафіксована тільки в двох випадках – brd2(bd1) та brd3(bd1), за скорингом ці два рішення мають 45.58 та 53.46 відповідно, тобто референсна сполука краще розміщується в сайті brd3(bd1). Хоча референсна сполука підбиралася таким чином, щоб взаємодіяти з brd4(bd1), за результатами докінгу, вона на це білок не сіла. Причина цього може критися в тому, що для моделювання були вибрані структури з альфафолду, бо кристалічна структура референсного ліганду, зв'язаного з таргетним доменом є тільки для структури бромодомену brd4. Візуалізація та вирівнювання структур з PDB та AlphaFold показала, що є розбіжності в координатах залишку Аспарагіну 140 – ключової молекули, для взаємодії бромодомену з лігандом(Рис. 3.4)



**Рис. 3.4.** Моделювання взаємодії brd4(bd1) з референсною сполукою

Зеленим показана структура першого бромодому brd4 з Protein Data Bank, а сірим – структуру з AlphaFold. Як видно з рисунку, бічний ланцюг амінокислотного залишку з кристалічної структури знаходиться ближче до ліганду ( 2.1 Å ). Залишок тієї самої амінокислоти в структурі за AlphaFold знаходиться на відстані 2.8 Å. Ліганд ZINC5161502 за результатами докінгу сідає на brd2(bd1) та brdt(bd1), при чому зв'язування є сильніше з brdt – скоринг 49.8, а для brd2(bd1) - 48.22.

ZINC22046885 зв'язується з brd2\_bd1 та bdr3\_bd1, при чому оціночні функції показують, що ця молекула більш специфічна для brd3(bd1) (скоринг - 51.58), ніж для brd2\_bd1 (скоринг - 49.87)

ZINC27721698 сідає на 4 бромодомени різних білків, найбільш специфічний ця молекула для brdt(bd2), середній скоринг для всіх рішень - 49.975, найменш специфічний – для brd4(bd2) - 36.88. Для brd2\_bd1 та bdr3\_bd1 середній скоринг становить ~45

ZINC27721706 сідає на ті самі бромодомени, що і попередня молекула ZINC27721698. Але найбільш специфічно вона сідає на bdr3\_bd1 (середній скоринг 52.38667), найменш специфічним таргетом є brd4\_bd2 (скоринг 41.17), brd2\_bd1 та brdt\_bd2 знаходяться по середині (середній скоринг 45.8 та 44.19 відповідно).

ZINC000027727540 сідає на brd2\_bd1 з середнім скорингом 49.038 та на brdt\_bd2 (середній скоринг 55.11), тобто ліганд є більш специфічним саме для brdt\_bd2.

Загалом, молекула ZINC22046885 та ZINC5161502 за результатами докінгу є специфічна до перших бромодомени білків BET.

Хоча референсна сполука була специфічна до першого бромодому BRD4, не було знайдено жодної сполуки серед сформованої бібліотеки, що взаємодіяла б тільки з цим білком (доменом). Слід зазначити, що ці результати не підкріплені методами мокрої біології, тому потребують подальших досліджень.

Ліганди та бромодомени, наведені в таблиці 3.1, надалі було відібрано для побудови моделі рецептор-орієнтованого фармакофору.

### 3.3. Аналіз фармакофорної моделі

На основі даних докінгу були побудовані рецептор-орієнтовані фармакофорні моделі доменів білків ВЕТ родини, з якими за результатами докінгу зв'язувалися ліганди бібліотеки. Для кожного домену після результатів докінгу, була обрана молекула з найбільшим скорингом, на основі якої був побудований рецептор-орієнтований фармакофор. З інших сполук, що взаємодіють з цим бромодоменом, була створена бібліотека, яка потім була протестована на предмет схожості фармакофорних властивостей з молекулою з найбільшим скорингом. Це було зроблено для того, щоб побачити, чи молекули мають спільні властивості і взаємодіють за схожим механізмом зі своїм таргетом. Побудова фармакофорної моделі та скринінг на її базі був проведений, використовуючи програмне забезпечення LigandScout, зі стандартними налаштування для проведення віртуального скринінгу, а саме скринінг був налаштований на пошук схожих (shared) фармакофорних властивостей, і алгоритм мав представити лише найкращу модель, відхиляючи інші.

Нижче будуть наведені рецептор-орієнтовані фармакофорні моделі для різних доменів та їх лігандів. Рисунки будуть включати в себе просторову орієнтацію моделі (3D), а також 2D-діаграми, де будуть зображені амінокислотні залишки в активному центрі білка, що беруть участь у формуванні сайту зв'язування.

Як ми бачимо, ключовою амінокислотою для взаємодії з лігандами є Asn156, яка взаємодіє з Нітрогеном триазольної групи сполук, де нітроген триазольної групи є донором водневого зв'язку, а гідроген амідної групи – акцептором (Рис

3.5). Окрім неї, активний сайт складають залишки фенілаланіну 99, триптофану 97, лейцину 108, валіну 103 та ізолейцину 162. Ці амінокислотні залишки беруть участь у формуванні гідрофобного карману для взаємодій з неполярними групами лігандів.

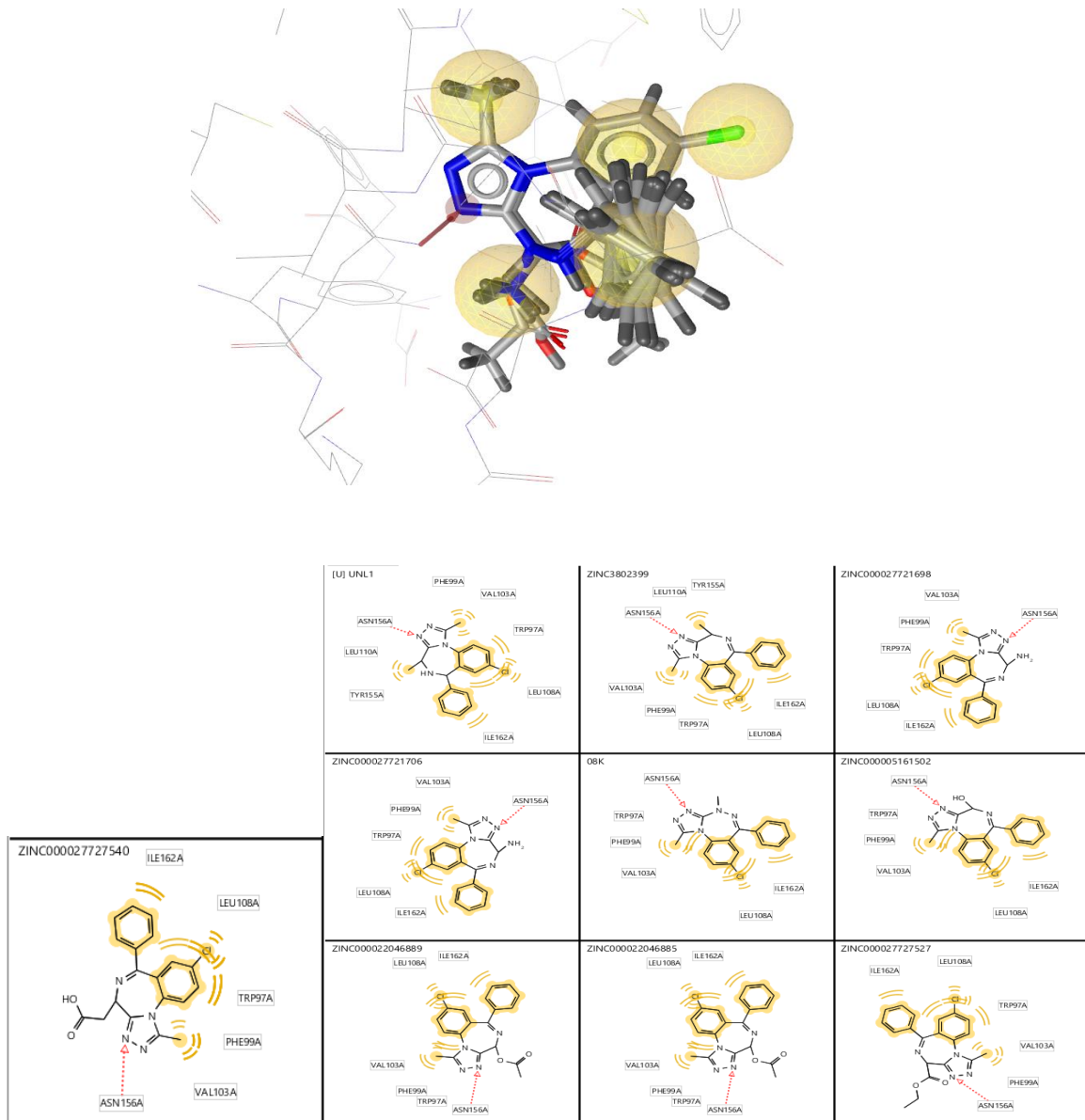
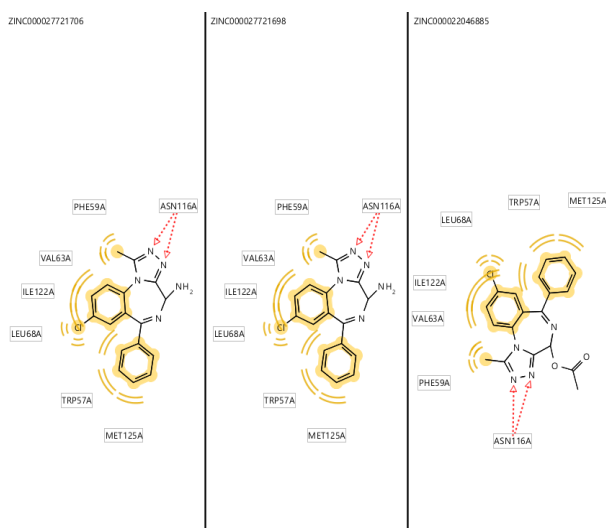
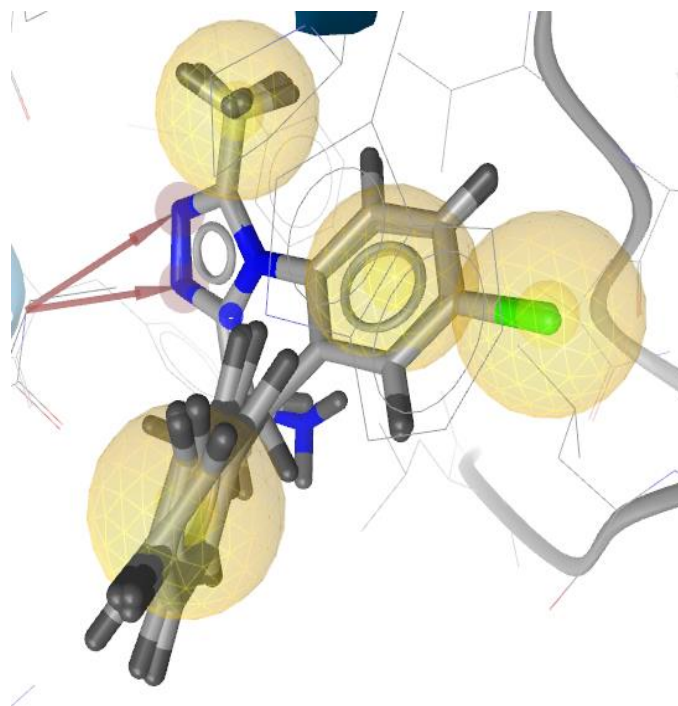
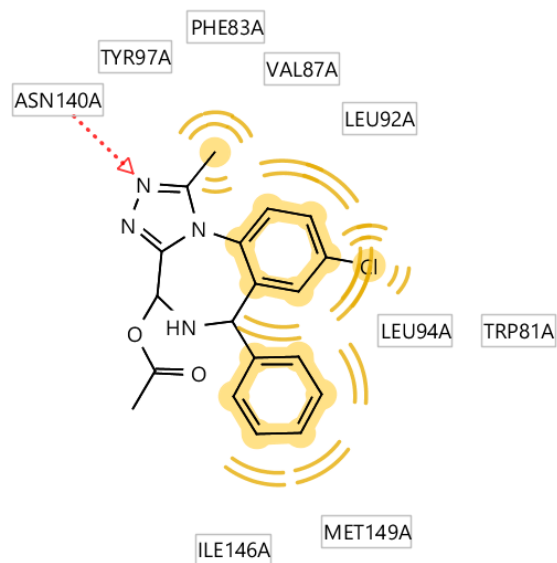
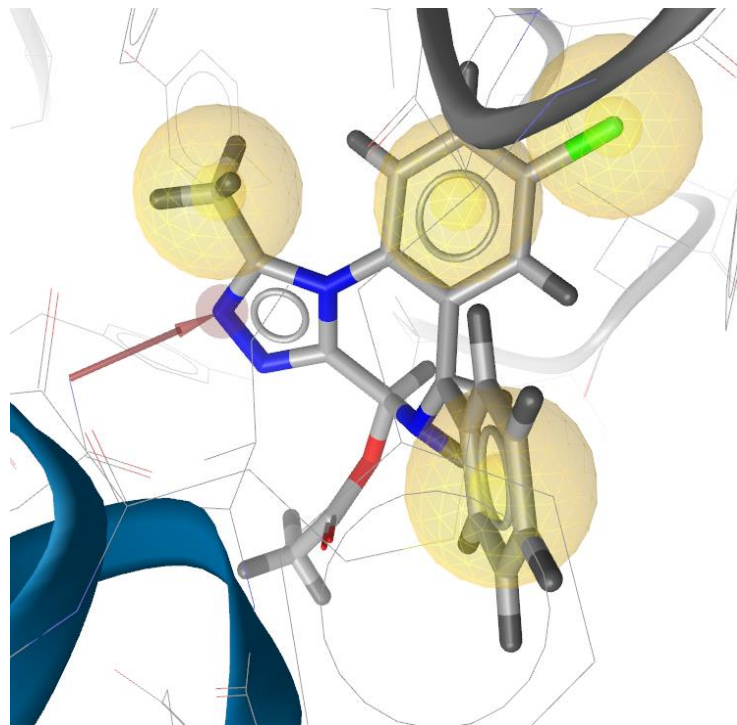


Рис. 3.5. Рецептор-орієнтований фармакофор сету лігандів для brd2(bd1)

Для brd3(bd1) ми бачимо схожий патерн взаємодій, але в даному випадку маємо 2 водневі зв'язки і за скорингом 2/3 сполук показали кращі результати, ніж ті самі речовини в інших доменах (Таб. 3.1). Також можна підкреслити схожий патерн амінокислотних залишків в активному сайті бромодому.



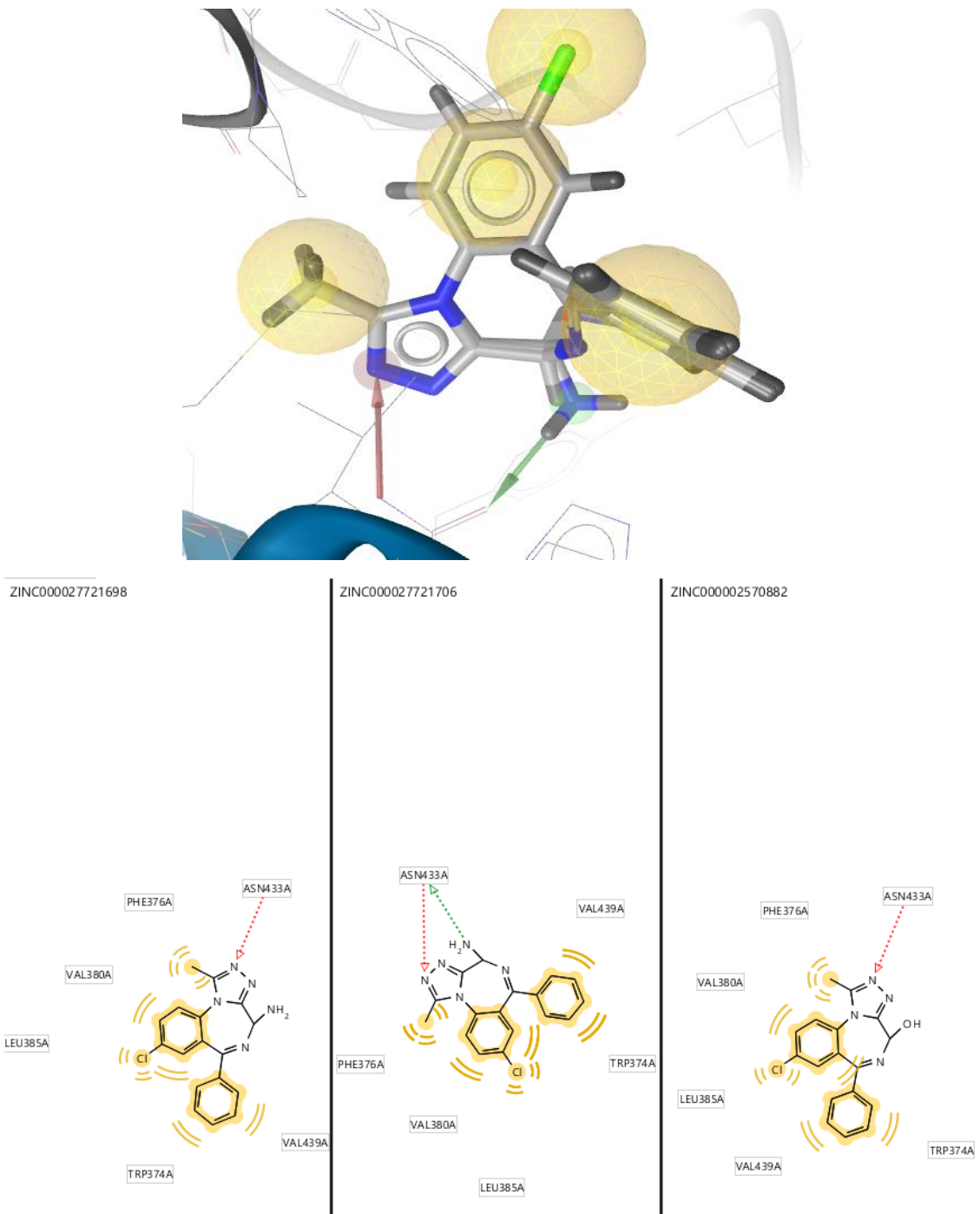
**Рис. 3.6.** Рецептор-орієнтований фармакофор сету лігандів для brd3(bd1)



**Рис. 3.7.**Рецептор-орієнтований фармакофор ліганду ZINC22046885 для brd4(bd1)

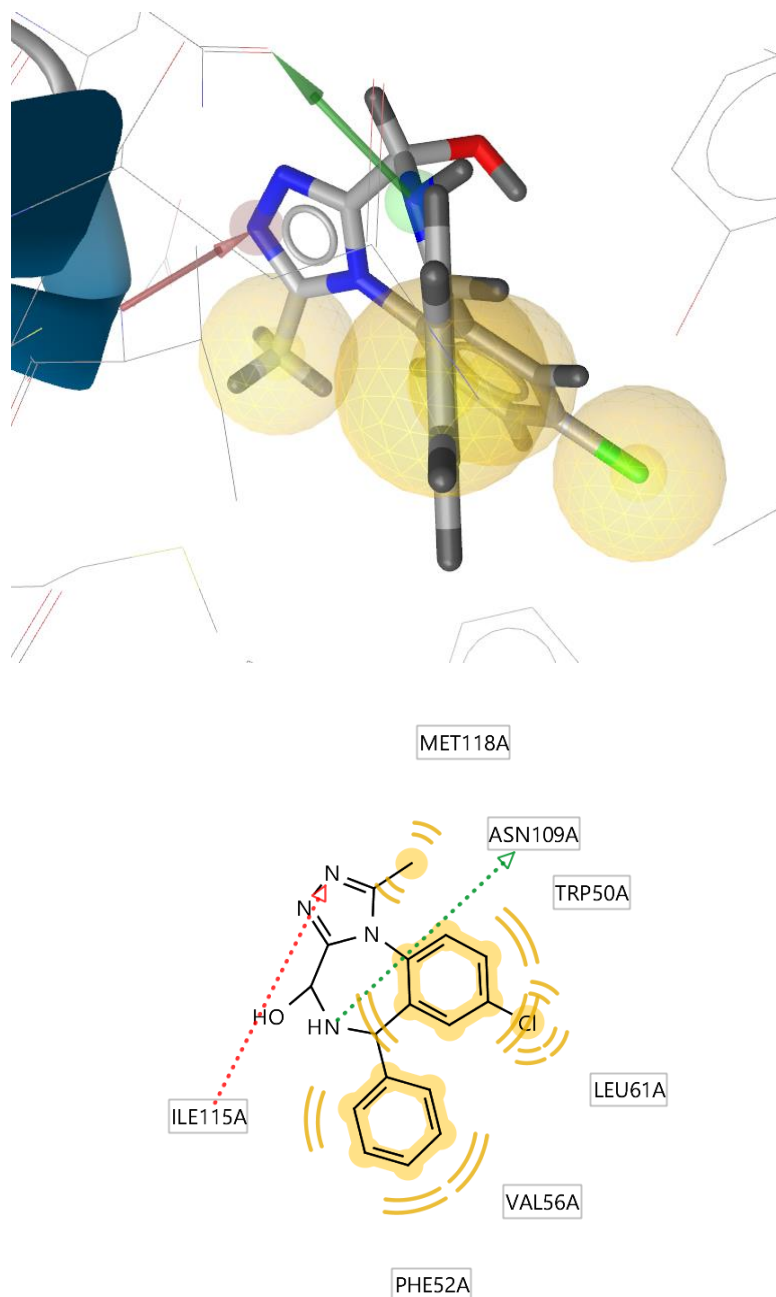
На Рис. 3.8. речовина ZINC27721706 формує додатковий водневий зв'язок з аспарагіновим залишком в активному центрі. І якщо порівнювати цю сполуку з іншими двома за їх скорингом, то ми побачимо, що він

більший, тобто можна зробити гіпотезу, що дана сполука закріплюється в активному сайті білка краще за дві інші.



**Рис. 3.8.** Рецептор-орієнтований фармакофор сету лігандів для brd4(bd2)

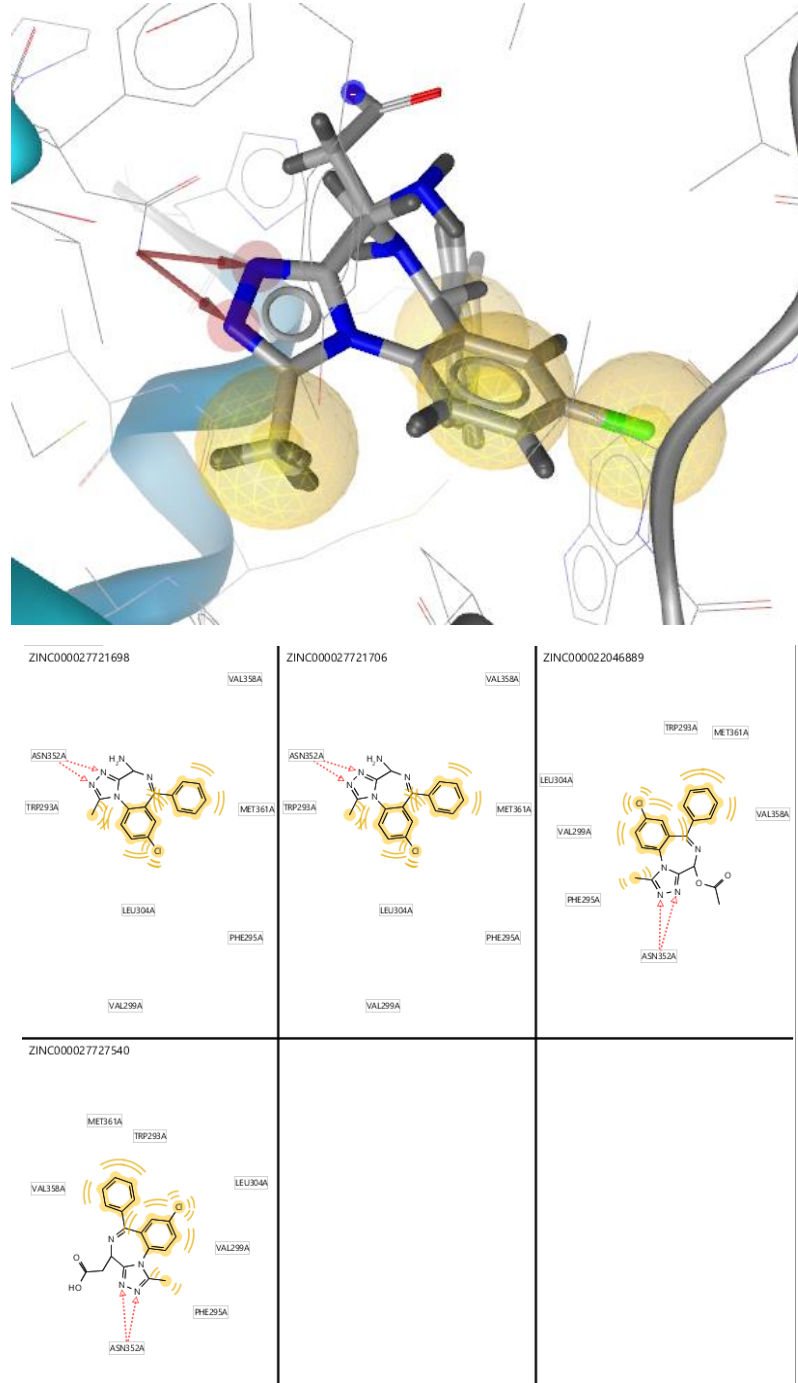
Як видно на Рис. 3.9. речовина ZINC5161502, хоч і зв'язується з ключовим амінокислотним залишком аспарагінової кислоти, він зв'язується не нітрогеном триазольної групи, але іншого гетероциклу.



**Рис. 3.9.** Рецептор-орієнтований фармакофор ліганду ZINC5161502 для brdt(bd1)

На Рис.3.10. видно, що залишок Аспарагіну формує два зв'язки з триазольними групами сполук, але якщо порівнювати їх за скорингом з тими

же сполуками, але в сайті brd2(bd2), то ми не побачимо суттєвих відмінностей. На ці взаємодії треба в подальшому дивитись в контексті молекулярної динаміки, бо хоч алгоритм докінгу показує декілька поз ліганду в кативному центрі, ми не можемо точно сказати, чи існують дані взаємодії насправді, чи це просто «вдачне» позиціонування в активному сайті.



**Рис. 3.10.** Рецептор-орієнтований фармакофор сету лігандів для brdt(bd2)

Отже, всі розглянуті бромодомени та їх ліганди взаємодіють та розміщуються в активному сайті подібним чином за винятком деяких випадків (але навіть там взаємодії відбуваються шляхом створення з аспарагіном водневого зв'язку). Серед бромодоменів зберігається патерн амінокислот, які формують активний центр. Ключовою серед них є амінокислота Аспарагін, яка зв'язується своєю амідною групою з триазольною групою ліганду. Інші амінокислоти активного центру формують гідрофобний карман для взаємодії з гетероциклами та бензольними кільцями лігандів.

Загалом, на основі референсного ліганду було знайдено 9 молекул, що потенційно можуть зв'язуватися з різними бромодоменами родини ВЕТ. З них є специфічними для якогось одного бромодомену, 2 зв'язуються лише с першим типом, а решта мають декілька різних таргетів. Подальші дослідження є необхідними для підтвердження гіпотези активності обраних молекул щодо їх потенційних таргетів. Зокрема, слід провести молекулярну динаміку для аналізу поведінки ліганду в активному сайті. Оскільки в цій роботі розглядалися лише ті ліганди, що позиціонувалися в активному сайті білка відповідно до референсної речовини, залишаючи позаду сполуки, що були близькі до референсної пози, але не достатньо, що уловити ключову взаємодію, цей тип аналізу стане в нагоді. Після цього треба буде провести біохімічне дослідження та встановити факт зв'язування речовини з таргетом. Це може бути зроблено шляхом детекції самого зв'язування наприклад методами TSA та SPR або експериментами, що оцінюють відсоток інгібування сполуки.

## ВИСНОВКИ

1. На основі Similarity Search були вибрані 14 сполук з бази даних ZINC для подальшого молекулярного докінгу.
2. Були оптимізовані структури доменів білків родини ВЕТ, завантажених з бази даних AlphaFold та лігандів.
3. Був проведений молекулярний докінг 14-ти сполук проти 8-ми бромодоменів 4-ох білків. В результаті, 9 сполук показали, що можуть позиціонуватися в активному сайті різних 6 доменів. 3 сполуки показали абсолютну специфічність - лише до якогось одного з доменів, ще 2 показали специфічність лише до одного типу доменів.
4. Побудова рецептор-орієнтованої фармакофорної моделі та її аналіз показав, що ліганди мають спільні фармакофорні властивості та взаємодіють подібним чином. Ключовою хімічною групою для взаємодії з рецептором у цих лігандів є триазольна група (як і в референсного ліганду), а ключовим амінокислотним залишком, приймаючим участь у цих взаємодіях - аспарагіновий залишок у всіх 6 бромодоменах.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Suarez-Alvarez, B., Rodriguez, R.M., Ruiz-Ortega, M. and Lopez-Larrea, C. (2017). BET Proteins: An Approach to Future Therapies in Transplantation. *American Journal of Transplantation*, 17(9), pp.2254–2262.
2. Lu, Tian; Lu, Wenchao, Luo, Cheng (2019). A patent review of BRD4 inhibitors (2013-2019). *Expert Opinion on Therapeutic Patents*, pp. 1–25.
3. Conrad, R.J., Fozouni, P., Thomas, S., Sy, H., Zhang, Q., Zhou, M.-M. and Ott, M. (2017). The Short Isoform of BRD4 Promotes HIV-1 Latency by Engaging Repressive SWI/SNF Chromatin-Remodeling Complexes. *Molecular Cell*, 67(6), pp.1001-1012.e6.
4. Donati, Benedetta, Lorenzini, Eugenia, Ciarrocchi, Alessia (2018). BRD4 and Cancer: going beyond transcriptional regulation. *Molecular Cancer*, 17(1), pp. 164–172
5. Li, G.-Q., Guo, W.-Z., Zhang, Y., Seng, J.-J., Zhang, H.-P., Ma, X.-X., Zhang, G., Li, J., Yan, B., Tang, H.-W., Li, S.-S., Wang, L.-D. and Zhang, S.-J. (2015). Suppression of BRD4 inhibits human hepatocellular carcinoma by repressing MYC and enhancing BIM expression. *Oncotarget*, 7(3), pp. 51-67
6. Wu, X., Liu, D., Tao, D., Xiang, W., Xiao, X., Wang, M., Wang, L., Luo, G., Li, Y., Zeng, F. and Jiang, G. (2016). BRD4 Regulates EZH2 Transcription through Upregulation of C-MYC and Represents a Novel Therapeutic Target in Bladder Cancer. *Molecular Cancer Therapeutics*, 15(5), pp.1029–1042.
7. Bhagwat, Anand S., Roe, J.-S., Mok, Beverly Y.L., Hohmann, Anja F., Shi, J. and Vakoc, Christopher R. (2016). BET Bromodomain Inhibition Releases the Mediator Complex from Select cis -Regulatory Elements. *Cell Reports*, 15(3), pp.519–530.
8. Donner, A.J., Ebmeier, C.C., Taatjes, D.J. and Espinosa, J.M. (2010). CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nature Structural & Molecular Biology*, 17(2), pp.194–201.

9. Lovén, J., Hoke, Heather A., Lin, Charles Y., Lau, A., Orlando, David A., Vakoc, Christopher R., Bradner, James E., Lee, T. and Young, Richard A. (2013). Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. *Cell*, 153(2), pp.320–334.
10. Shi, J. and Vakoc, Christopher R. (2014). The Mechanisms behind the Therapeutic Activity of BET Bromodomain Inhibition. *Molecular Cell*, 54(5), pp.728–736.
11. Huang, B., Yang, X.-D., Zhou, M.-M., Ozato, K. and Chen, L.-F. (2008). Brd4 Coactivates Transcriptional Activation of NF- $\kappa$ B via Specific Binding to Acetylated RelA. *Molecular and Cellular Biology*, 29(5), pp.1375–1387.
12. Zou, Z., Huang, B., Wu, X., Zhang, H., Qi, J., Bradner, J., Nair, S. and Chen, L.-F. (2013). Brd4 maintains constitutively active NF- $\kappa$ B in cancer cells by binding to acetylated RelA. *Oncogene*, 33(18), pp.2395–2404.
13. Rahman, S., Sowa, M.E., Ottinger, M., Smith, J.A., Shi, Y., Harper, J.W. and Howley, P.M. (2011). The Brd4 Extraterminal Domain Confers Transcription Activation Independent of pTEFb by Recruiting Multiple Proteins, Including NSD3. *Molecular and Cellular Biology*, 31(13), pp.2641–2652.
14. Liu, W., Ma, Q., Wong, K., Li, W., Ohgi, K., Zhang, J., Aggarwal, Aneel K. and Rosenfeld, Michael G. (2013). Brd4 and JMJD6-Associated Anti-Pause Enhancers in Regulation of Transcriptional Pause Release. *Cell*, 155(7), pp.1581–1595.
15. Jonkers, I. and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. *Nature Reviews Molecular Cell Biology*, [online] 16(3), pp.167–177.
16. Yang, Z., Yik, J.H.N., Chen, R., He, N., Jang, M.K., Ozato, K. and Zhou, Q. (2005). Recruitment of P-TEFb for Stimulation of Transcriptional Elongation by the Bromodomain Protein Brd4. *Molecular Cell*, 19(4), pp.535–545.
17. Mani, R.-S. and Chinnaiyan, A.M. (2010). Triggers for genomic rearrangements: insights into genomic, cellular and environmental influences. *Nature Reviews Genetics*, 11(12), pp.819–829.

18. Misteli, T. and Soutoglou, E. (2009). The emerging role of nuclear architecture in DNA repair and genome maintenance. *Nature Reviews Molecular Cell Biology*, 10(4), pp.243–254.
19. Donati, B., Lorenzini, E. and Ciarrocchi, A. (2018). BRD4 and Cancer: going beyond transcriptional regulation. *Molecular Cancer*, 17(1), pp. 324-329
20. Li, X., Baek, G., Ramanand, S.G., Sharp, A., Gao, Y., Yuan, W., Welti, J., Rodrigues, D.N., Dolling, D., Figueiredo, I., Sumanasuriya, S., Crespo, M., Aslam, A., Li, R., Yin, Y., Mukherjee, B., Kanchwala, M., Hughes, A.M., Halsey, W.S. and Chiang, C.-M. (2018). BRD4 Promotes DNA Repair and Mediates the Formation of TMPRSS2-ERG Gene Rearrangements in Prostate Cancer. *Cell Reports*, 22(3), pp.796–808.
21. Stanlie, A., Yousif, Ashraf S., Akiyama, H., Honjo, T. and Begum, Nasim A. (2014). Chromatin Reader Brd4 Functions in Ig Class Switching as a Repair Complex Adaptor of Nonhomologous End-Joining. *Molecular Cell*, 55(1), pp.97–110.
22. Schultz, L.B., Chehab, N.H., Malikzay, A. and Halazonetis, T.D. (2000). P53 Binding Protein 1 (53bp1) Is an Early Participant in the Cellular Response to DNA Double-Strand Breaks. *Journal of Cell Biology*, 151(7), pp.1381–1390.
23. Gandolfi, G., Ragazzi, M., de Biase, D., Visani, M., Zanetti, E., Torricelli, F., Sancisi, V., Gugnoni, M., Manzotti, G., Braglia, L., Cavuto, S., Merlo, D.F., Tallini, G., Frasoldati, A., Piana, S. and Ciarrocchi, A. (2017). Genome-wide profiling identifies the THYT1 signature as a distinctive feature of widely metastatic Papillary Thyroid Carcinomas. *Oncotarget*, 9(2), pp.1813–1825.
24. Gandolfi, G., Ragazzi, M., Frasoldati, A., Piana, S., Ciarrocchi, A. and Sancisi, V. (2015). TERT promoter mutations are associated with distant metastases in papillary thyroid carcinoma. *European Journal of Endocrinology*, 172(4), pp.403–413.
25. Horn, S., Figl, A., Rachakonda, P.S., Fischer, C., Sucker, A., Gast, A., Kadel, S., Moll, I., Nagore, E., Hemminki, K., Schadendorf, D. and Kumar, R.

- (2013). TERT Promoter Mutations in Familial and Sporadic Melanoma. *Science*, 339(6122), pp.959–961.
26. Akincilar, S.C., Khattar, E., Boon, P.L.S., Unal, B., Fullwood, M.J. and Tergaonkar, V. (2016). Long-Range Chromatin Interactions Drive Mutant TERT Promoter Activation. *Cancer Discovery*, [online] 6(11), pp.1276–1291.
27. Bandaria, Jigar N., Qin, P., Berk, V., Chu, S. and Yildiz, A. (2016). Shelterin Protects Chromosome Ends by Compacting Telomeric Chromatin. *Cell*, 164(4), pp.735–746.
28. Ren, Y., Zhang, Y., Wang, Z., Wang, C., Zhang, H., Wang, Y. and Zhao, Z. (2019). Role of Brd4 in the production of inflammatory cytokines in mouse macrophages treated with titanium particles. *Canadian Journal of Physiology and Pharmacology*, 97(11), pp.1028–1034.
29. Meng, S., Zhang, L., Tang, Y., Tu, Q., Zheng, L., Yu, L., Murray, D., Cheng, J., Kim, S.H., Zhou, X. and Chen, J. (2014). BET Inhibitor JQ1 Blocks Inflammation and Bone Destruction. *Journal of Dental Research*, 93(7), pp.657–662.
30. Xiao, Y., Liang, L., Huang, M., Qiu, Q., Zeng, S., Shi, M., Zou, Y., Ye, Y., Yang, X. and Xu, H. (2015). Bromodomain and extra-terminal domain bromodomain inhibition prevents synovial inflammation via blocking I $\kappa$ B kinase–dependent NF- $\kappa$ B activation in rheumatoid fibroblast-like synoviocytes. *Rheumatology*, 55(1), pp.173–184.
31. Hajmirza, A., Emadali, A., Gauthier, A., Casasnovas, O., Gressin, R. and Callanan, M. (2018). BET Family Protein BRD4: An Emerging Actor in NF $\kappa$ B Signaling in Inflammation and Cancer. *Biomedicines*, 6(1), p.16.
32. Grivennikov, S.I. and Karin, M. (2010). Inflammation and oncogenesis: a vicious connection. *Current opinion in genetics & development*, [online] 20(1), p.65.
33. Andrieu, G.P., Shafran, J.S., Deeney, J.T., Bharadwaj, K.R., Rangarajan, A. and Denis, G.V. (2018). BET proteins in abnormal metabolism, inflammation,

- and the breast cancer microenvironment. *Journal of Leukocyte Biology*, 104(2), pp.265–274.
34. White, M.E., Fenger, J.M. and Carson, W.E. (2019). Emerging roles of and therapeutic strategies targeting BRD4 in cancer. *Cellular Immunology*, 337, pp.48–53.
  35. Dahabieh, M.S., Battivelli, E. and Verdin, E. (2015). Understanding HIV Latency: The Road to an HIV Cure. *Annual Review of Medicine*, 66(1), pp.407–421.
  36. Davey, R.T., Bhat, N., Yoder, C., Chun, T.-W. ., Metcalf, J.A., Dewar, R., Natarajan, V., Lempicki, R.A., Adelsberger, J.W., Miller, K.D., Kovacs, J.A., Polis, M.A., Walker, R.E., Falloon, J., Masur, H., Gee, D., Baseler, M., Dimitrov, D.S., Fauci, A.S. and Lane, H.C. (1999). HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *Proceedings of the National Academy of Sciences*, [online] 96(26), pp.15109–15114.
  37. Mori, L. and Valente, S.T. (2020). Key Players in HIV-1 Transcriptional Regulation: Targets for a Functional Cure. *Viruses*, 12(5), p.529.
  38. Li, Z., Guo, J., Wu, Y. and Zhou, Q. (2012). The BET bromodomain inhibitor JQ1 activates HIV latency through antagonizing Brd4 inhibition of Tat-transactivation. *Nucleic Acids Research*, 41(1), pp.277–287.
  39. Zhu, J., Gaiha, Gaurav D., John, Sinu P., Pertel, T., Chin, Christopher R., Gao, G., Qu, H., Walker, Bruce D., Elledge, Stephen J. and Brass, Abraham L. (2012). Reactivation of Latent HIV-1 by Inhibition of BRD4. *Cell Reports*, 2(4), pp.807–816.
  40. Li, Z., Mbonye, U., Feng, Z., Wang, X., Gao, X., Karn, J. and Zhou, Q. (2018). The KAT5-Acetyl-Histone4-Brd4 axis silences HIV-1 transcription and promotes viral latency. *PLOS Pathogens*, 14(4), p.e1007012.
  41. Shi J & Vakoc CR The Mechanisms behind the Therapeutic Activity of BET Bromodomain Inhibition. *Mol. Cell* 54, 728–736 (2014)

42. French, C. A. (2016). *Small-Molecule Targeting of BET Proteins in Cancer. Advances in Cancer Research*, 21–58. doi:10.1016/bs.acr.2016.04.001
43. Samelson, A.J., Tran, Q.D., Robinot, R. *et al.* BRD2 inhibition blocks SARS-CoV-2 infection by reducing transcription of the host cell receptor ACE2. *Nat Cell Biol* 24, 24–34 (2022).
44. Qiao, Y. *et al.* Targeting transcriptional regulation of SARS-CoV-2 entry factors *ACE2* and *TMPRSS2*. *Proc. Natl Acad. Sci. USA* 118, e2021450118 (2020).
45. Gilham, D. *et al.* Bromodomain and extraterminal protein inhibitor, apabetalone (RVX-208), reduces ACE2 expression and attenuates SARS-CoV-2 infection in vitro. *Biomedicines* 9, 437 (2021).
46. Belkina AC, Nikolajczyk BS and Denis GV. BET protein function is required for inflammation: Brd2 genetic disruption and BET inhibitor JQ1 impair mouse macrophage inflammatory responses. *The Journal of Immunology*, 2013, 190(7): 3670-3678
47. Murakami, S., Li, R., Nagari, A., Chae, M., Camacho, C. V., and Kraus, W. L. (2019). Distinct Roles for BET Family Members in Estrogen Receptor  $\alpha$  Enhancer Function and Gene Regulation in Breast Cancer Cells. *Mol. Cancer Res.* 17, 2356–2368.
48. Andrieu, G. P., and Denis, G. V. (2018). BET Proteins Exhibit Transcriptional and Functional Opposition in the Epithelial-To-Mesenchymal Transition. *Mol. Cancer Res.* 16, 580–586
49. Lambert, J.-P., Picaud, S., Fujisawa, T., Hou, H., Savitsky, P., Uusküla-Reimand, L., *et al.* (2019). Interactome Rewiring Following Pharmacological Targeting of BET Bromodomains. *Mol. Cel.* 73, 621–638.
50. Kanno, T., Kanno, Y., Siegel, R. M., Jang, M. K., Lenardo, M. J., and Ozato, K. (2004). Selective Recognition of Acetylated Histones by Bromodomain Proteins Visualized in Living Cells. *Mol. Cel.* 13, 33–43.

51. Matzuk, M. M., McKeown, M. R., Filippakopoulos, P., Li, Q., Ma, L., Agno, J. E., ... Bradner, J. E. (2012). *Small-Molecule Inhibition of BRDT for Male Contraception*. *Cell*, 150(4), 673–684.
52. *Development and Validation of a Genetic Algorithm for Flexible Docking* G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 267, 727-748, 1997
53. Schrödinger, L. & DeLano, W., 2020. *PyMOL*, Available at: <http://www.pymol.org/pymol>.
54. Wolber, G., & Langer, T. (2005). *LigandScout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters*. *Journal of Chemical Information and Modeling*, 45(1), 160–169.
55. BIOVIA, Dassault Systèmes, [Discovery Studio Visualizer, [2021], San Diego: Dassault Systèmes, [2021].
56. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., & Coleman, R. G. (2012). *ZINC: A Free Tool to Discover Chemistry for Biology*. *Journal of Chemical Information and Modeling*, 52(7), 1757–1768.
57. A. Gaulton, L. Bellis, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, R. Akhtar, A.P. Bento, B. Al-Lazikani, D. Michalovich, & J.P. Overington (2012) ‘ChEMBL: A Large-scale Bioactivity Database For Chemical Biology and Drug Discovery’ *Nucleic Acids Res. Database Issue*, **40** D1100-1107
58. Varadi, M *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* (2021).
59. The Protein Data Bank H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne (2000) *Nucleic Acids Research*, **28**: 235-242.
60. Landrum, G. 2010. “RDKit.” Q2. <https://www.rdkit.org/>.

61. Patterson DE, Cramer RD, Ferguson AM, Clark RD, Weinberger LE: Neighborhood behavior: A useful concept for validation of "molecular diversity" descriptors. *J Med Chem* 1996, 39:3049-3059
62. Filippakopoulos, P., Picaud, S., Fedorov, O., Keller, M., Wrobel, M., Morgenstern, O., ... Knapp, S. (2012). *Benzodiazepines and benzotriazepines as protein interaction inhibitors targeting bromodomains of the BET family. Bioorganic & Medicinal Chemistry*, 20(6), 1878–1886.