

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

ІМЕНІ ТАРАСА ШЕВЧЕНКА

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ВИСОКИХ ТЕХНОЛОГІЙ

Зав. кафедри: к. б. н., доцент Нипорко О. Ю.

Протокол № __ засідання кафедри

від «__» травня 2022 р.

**ЗАСТОСУВАННЯ МЕТОДУ ЗОНТИЧНОЇ ВИБІРКИ ДЛЯ ОЦІНКИ
ЗМІНИ ВІЛЬНОЇ ЕНЕРГІЇ В ПРОЦЕСАХ БІОМОЛЕКУЛЯРНОГО
РОЗПІЗНАВАННЯ/ВЗАЄМОДІЇ ЗА УЧАСТЮ ДНК-ПОЛІМЕРАЗИ ЕТА**

Випускна кваліфікаційна робота

студентки ІІ курсу магістратури

091 Біологія

Федірко Тетяни Олегівни

Науковий керівник:

к. б. н., доцент

Нипорко О. Ю.

Оцінка захисту роботи

КИЇВ 2022

Анотація

Федірко Т. О. Застосування методу зонтичної вибірки для оцінки зміни вільної енергії в процесах біомолекулярного розпізнавання/взаємодії за участю днк-полімерази ета.

Мета цієї роботи була дослідити за допомогою методу зонтичної вибірки особливості змін вільної енергії в процесах взаємодії між комплексом білок-ДНК та лігандом, де в якості білка виступає полімераза ета, а ліганда – нуклеотид деоксиаденозинтрифосфат.

Різниця вільної енергії системи між різними її станами є однією з основних характеристик більшості біофізичних процесів. Але не достатньо знати лише різницю енергій між двома станами, між ними може бути енергетичний бар'єр, що в свою чергу впливає на швидкість проходження реакції, тобто швидкість переходу між цими двома станами.

Для опису такого переходу зручно ввести координату реакції, що дозволяє описати систему безперервно. Для цього використовували метод зонтичної вибірки (Umbrella Sampling) який дозволяє доволі точно описати стани системи і оцінити зміну вільної енергії при, наприклад, утворенні молекулярних комплексів. Особливістю цього методу є введення додаткового потенціалу. Без такого потенціалу цей метод неефективно використовувати в системах з бар'єром.

Були отримані цікаві результати, а саме:

- 1) в залежності від того, який нуклеотид знаходиться на матричному ланцюзі, до якого приєднується деоксиаденозинтрифосфат, профіль взаємодії значно змінюється;

- 2) матричний нуклеодит впливає не тільки на вільний нуклеотид безпосередньо в активному центрі, а також і на відстані. Цей вплив по суті – створення певних енергетичних бар'єрів. Важливо відзначити, що різні матричні нуклеотиди впливають на ліганд з індивідуальними, для кожного з чотирьох, характеристиками;
- 3) для даної полімерази, характерною рисою є наявність енергетичних бар'єрів на етапі приєднання вільного нуклеотиду в активному центрі.

Отримані результати можуть бути використані у подальших дослідженнях активного сайту полімерази ета та для досліджень з вивчення можливих шляхів уникнення хіміорезистентності при лікуванні раку.

Ключові слова: зонтична вибірка; полімераза ета; вільна енергія; нуклеотид; деоксиаденозинтрифосфат.

Вступ

Ракові клітини використовують кілька ліній механізмів самозахисту, щоб обійти хіміотерапію та променевою терапію. Одним із таких механізмів, що використовуються раковими клітинами, є транслезійний синтез ДНК (TLS), при якому спеціалізовані TLS-полімерази обходять пошкодження ДНК за допомогою моноубіквітинованого проліферуючого ядерного антигену клітини. Полімераза ета (Pol η) відповідає за обхід цисплатину та пошкоджень ДНК, викликаних УФ-випроміненням.

Дослідження TLS-полімераз, схильних до мутагенних помилок, можливо, зможе пояснити генетичні мутації в різних злоякісних пухлинах. Залучення полімераз TLS в місці пошкодження може сприяти появі нових мутацій, які в кінцевому підсумку призводять до хеморезистентності і, як наслідок, до виживання ракових клітин. У людей посттрансляційна модифікація білків, пов'язаних з TLS, у відповідь на УФ або інші агенти, що пошкоджують ДНК, підвищує спорідненість TLS-полімераз до пошкодженої ділянки порівняно з неушкодженою ДНК. Таким чином, ця модель дає можливе пояснення перемикання класів від високоточних реплікативних до полімераз TLS.

Актуальність даної роботи полягає в тому, що до цього часу достеменно не було відомо яким чином ДНК-полімераза ета підбирає відповідні до матричних нуклеотиди. Хоча це один із найбільш досліджених представників сімейства В, динаміка його процесивності у часі, та зміна вільної енергії за подію були не досліджені. Краще розуміння механізму TLS, опосередкованого Pol η , може бути корисним у визначенні потенційної фармакологічної мішені для розробки ліків проти раку.

Мета: дослідити за допомогою методу зонтичної вибірки особливості змін вільної енергії в процесах взаємодії між комплексом білок-ДНК та

лігандом, де в якості білка виступає полімераза ета, а ліганда – нуклеотид деокзиденозинтрифосфат.

Завдання: провести докінг потрібного комплексу ДНК-полімерази ета із лігандом та безпосередньо ДНК; створити бокс і додати до нього розчинник; врівноважити систему; провести пулінг та семплінг.

Зміст

Анотація	2
Вступ	4
Розділ 1. Характеристика днк-полімераз.....	7
1.1. Реплікативні ДНК-полімерази.....	11
1.1.1. В сімейство полімераз.....	12
1.2. Репаративні полімерази.....	13
1.2.1. Y сімейство полімераз.....	14
1.2.2. A сімейство ДНК-полімераз.....	15
1.2.3. PrimPol сімейство	16
1.2.4. X сімейство полімераз.....	16
Розділ 2. Метод зонтичної вибірки	17
Розділ 3. Практична частина.....	19
Висновки	32
Джерела	34

Розділ 1. Характеристика днк-полімераз

Біологічне значення ДНК полягає в тому, що вона відіграє роль носія інформації всередині клітин у вигляді закодованого сейфу, де знаходяться дані про специфіку та стабільність системи. Ці дані не дають жодного пояснення про властивості ДНК з точки зору стандартних хімічних характеристик високомолекулярних нерегулярних гетерополімерів. Щоб зрозуміти біологічне значення ДНК, потрібно більш детально зануритись на молекулярний рівень організації живого. І навіть тоді не вдається знайти жодної простої відповіді в самій ДНК, тому що окремі молекули можуть давати велику структурну та функціональну варіабельність і мати відповідно різний вплив. Навіть якщо вони ідентичні, за винятком обміну одним із 10^9 нуклеотидних фрагментів геному, більшість таких обмінів мають дуже незначний ефект. Це відразу підводить до висновку, що ДНК в організмах виконує інформаційну функцію [1] і що внутрішня ДНК-залежна динаміка клітин здійснює обробку інформації, тобто обчислення [2]. Можна сказати, що молекулярно-біологічні обчислення на основі ДНК контролюють весь спектр біохімічних реакцій, що відбуваються в клітинах [3].

Очевидний спосіб зберігання інформації в ДНК – це послідовність букв зі стандартного чотирибуквеного алфавіту нуклеотидів А, Т, Г, Ц, був зрозумілий з моменту відкриття її структури - подвійний лінійний полімер зі спареними основами, - та способу копіювання комплементарного ланцюга [4,5].

Механізми реплікації клітинної ДНК дуже консервативні. Ланцюг, який синтезується в тому ж напрямку, що і рухома реплікаційна вилка, подовжується безперервно, в той час як синтез в зворотному напрямку є дискретним [6]. Невеликі фрагменти, що знаходяться на відстаючому ланцюзі, називаються

фрагментами Окадзакі на честь Реджі Окадзакі та Тунеко Окадзакі [7], які вперше запропонували модель синтезу за їх допомогою в 1968 році. Термін дискретний не слід тлумачити занадто буквально. Зазвичай припускається, що лідируючий ланцюг реплікується безперервно, хоча включення неканонічних нуклеотидів ДНК-полімеразами, зокрема урацилом і рибонуклеотидами, з наступним їх вирізанням, створює враження, що лідируючий ланцюг також може реплікуватися дещо безперервно [8,9]. Насправді це була одна з моделей, спочатку запропонованих Окадзакі [7].

Крім цієї простої моделі реплікації, деталі значно різняться між царствами і навіть всередині них, зокрема всередині царств бактерій та архей. Притаманна вищій асиметрія накладає ітераційний механізм праймування на відстаючий ланцюг, який в еукаріотів поєднаний з механізмом полімеризації, відмінним від механізму на лідируючому ланцюзі [6].

ДНК-залежні ДНК-полімерази є ключовими наявними каталізаторами для реплікації ДНК, і, отже, вони відіграють центральну роль у розповсюдженні генетичної інформації. Вони були детально описані за 60 років [10] з використанням розроблених біохімічних та біофізичних методів широкого спектру, і додатково підкріплених значною структурною інформацією [11].

Полімерази є основою багатьох біотехнологічних та медичних застосувань, від ПЛР до відбору аптамерів, від секвенування ДНК до синтезу ДНК *de novo*. Функціональні вимоги, які спричинили природну еволюцію цих ферментів, створили неймовірні молекулярні машини, які можна додатково вдосконалити. Зворотне проектування цих машин, щоб по-справжньому зрозуміти їх функції, значно прискорить їх подальший розвиток, дозволяючи точно налаштувати функції для кожного бажаного застосування. Відбір нуклеотидів в активному центрі, зв'язування полімерази з подвійним ланцюгом

ДНК, баланс між активними центрами екзонуклеази та полімерази, а також процес транслокації – динамічні процеси, які можна оптимізувати, застосовуючи полімеразну інженерію [11].

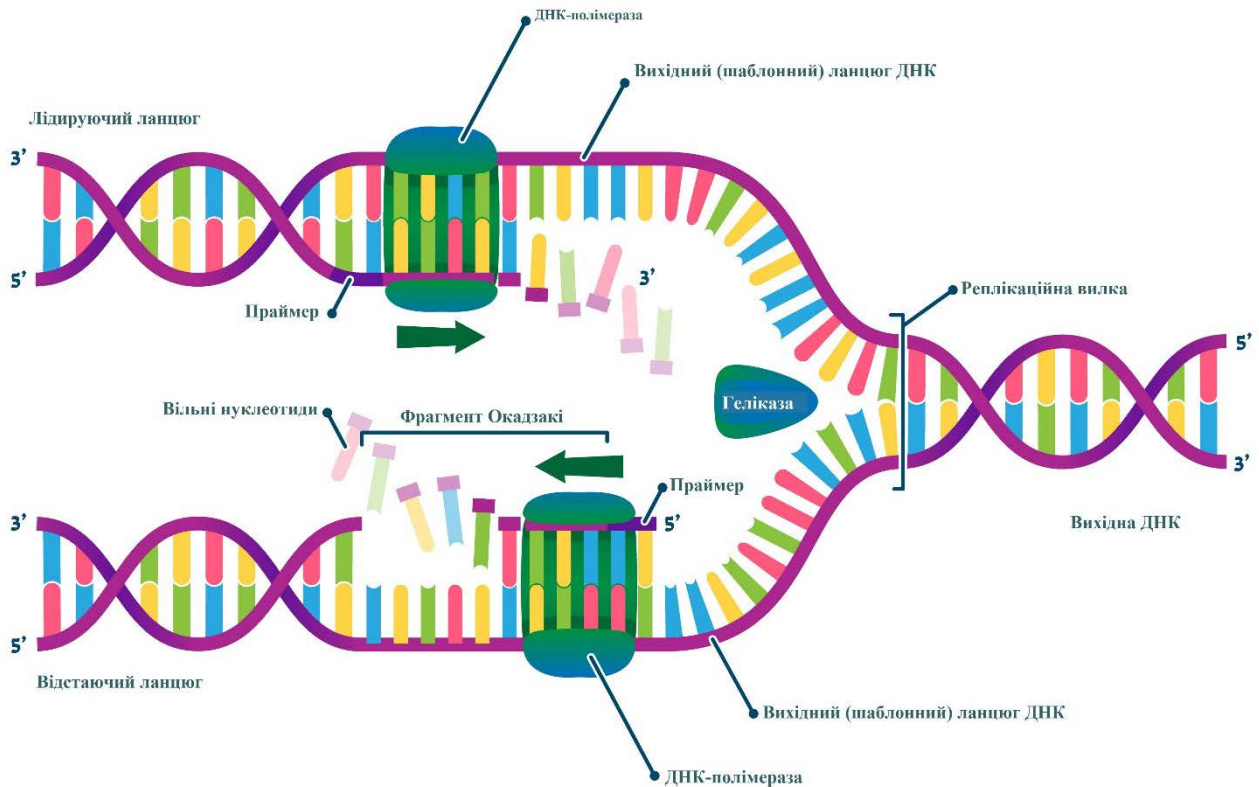


Рисунок 1.1. Загальна схема процесу реплікації ДНК.

Час перебування нуклеотиду в закритому полімеразному комплексі є ключовим фактором для його безпосереднього приєднання. Охарактеризовані мутації (зокрема для архейних гіпертермофільних полімераз В-сімейства), підтримують модель індукованої відповідності, запропоновану Цзаєм та Джонсоном [12], де бінарний комплекс полімераз зв'язує правильні та неправильні нуклеотиди з мінімальним виродженням. Правильний нуклеотид, швидше за все, стабілізує закриту конформацію, що призводить до включення останнього в ланцюг. Порушення у синтезі дочірнього ланцюга, викликані

вибором неправильного нуклеотиду, менш імовірно стабілізують закриту конформацію, а також можуть перешкоджати каталізу – обидва процеси підвищують точність полімерази. Таким чином, полімераза уособлює такий механізм з правильним нуклеотидом, що відбувається майже необоротний хімічний процес. Якщо помилкове включення все-таки відбувається, це може призвести до зупинки полімерази та корекції активним екзонуклеазним доменом (у репаративних полімеразах).

Існують деякі стійкі парадигми відносно ДНК-полімераз. Наприклад, всі ДНК-полімерази діють з двома каталітичними іонами магнію в їх активному центрі, і що ці білки можна найкраще згрупувати за схожістю їх послідовностей і консервативними мотивами у вісім описаних сімейств: А, В, С, D, Е, Х, Y і RT.

Структура активного центру ДНК-полімераз із різних сімейств дуже схожа (Рисунок 1.1). Вона поділяється зазвичай на три домени: великий палець (thumb), рухомі пальці (fingers) та долоню (palm). Домен великого пальця зв'язується із маленьким жолобком, домен рухливих пальців накриває частину ДНК, де безпосередньо приєднується новий деоксинуклеотидтрифосфат, домен долоні несе в собі каталітичні залишки.

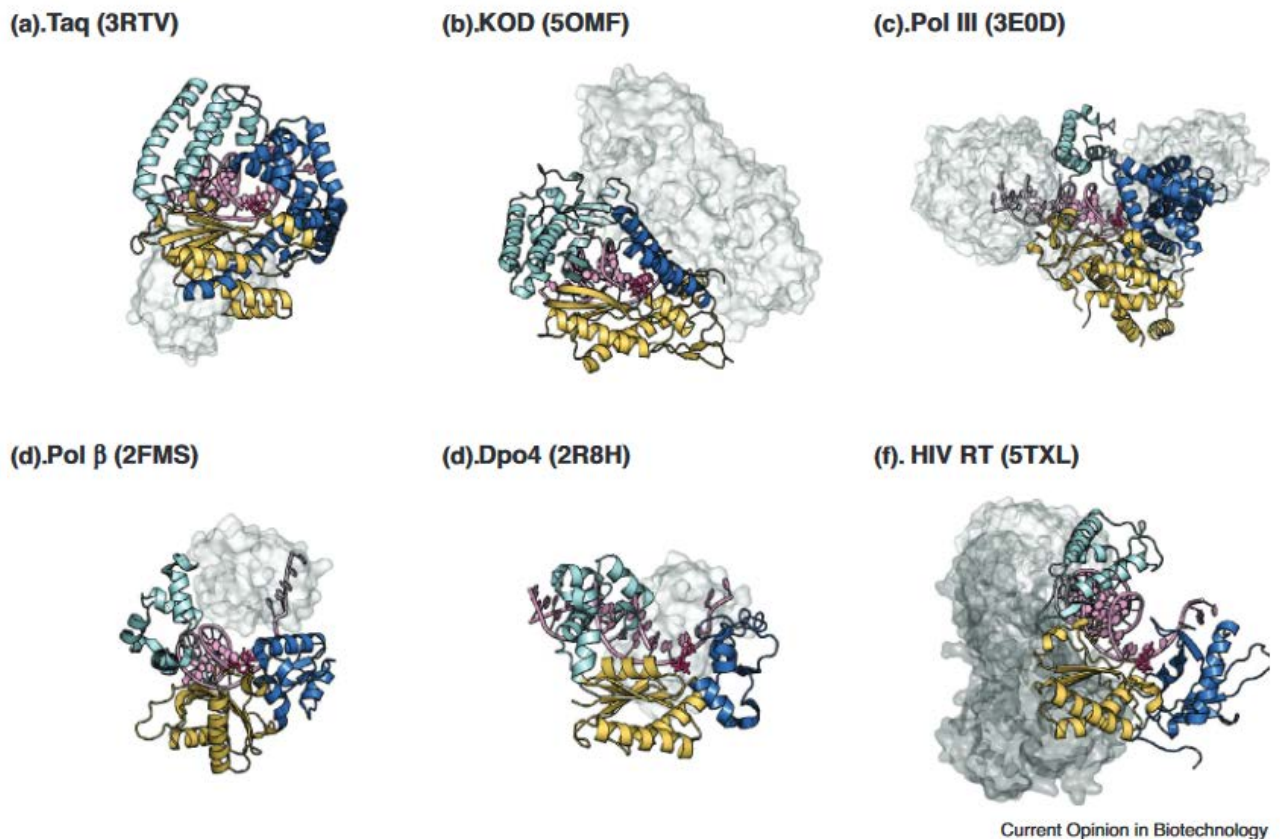


Рисунок 1.1. Структурна схожість полімераз із різних сімейств. Жовтим представлений домен «долоні», синім – домен «пальців», а блакитним – домен «великого пальця».

Механізм реплікації ДНК у еукаріотів схожий як у дріжджів, так і у людини та вимагає задіяних кількох ДНК-полімераз. На додаток до реплікативних ДНК-полімераз для дублювання лідируючих і відстаючих ланцюгів ДНК, інша група спеціалізованих полімераз необхідна для репарації ДНК та/або синтезу ДНК із транслезією (TLS).

1.1. Реплікативні ДНК-полімерази

Збірка реплісоми еукаріотів це багатоетапний процес. Ключовим протагоністом збірки є комплекс із 11 субодиниць CMG, що містить Cdc45, Mcm2–7 та білки GINS [13,14]. Mcm2–7 включають АТФ-залежні субодиниці,

які розкручують матричну дволанцюгову ДНК для синтезу лідируючого та відстаючого ланцюгів. Оскільки синтез ДНК відбувається в напрямку від 5' до 3', лідируючий ланцюг реплікується безперервно, а відстаючий ланцюг - періодично. На відстаючому ланцюзі Pol α -праймаза багаторазово синтезує короткі праймери РНК-ДНК для подовження за допомогою Pol δ . На лідируючому ланцюзі РНК-ДНК-праймер потрібен лише один. Є докази того, що Pol ϵ [15] та Pol δ [16] – полімерази провідного ланцюга.

1.1.1. В сімейство полімераз

Pol α -праймаза, Pol ϵ і Pol δ складаються із декількох субодиниць [14]. Номенклатура для дріжджів і для людини - різна. Кристалічні структури цих полімераз виявляють характерну особливість сімейства В. В даному сімействі кожен фермент складається з п'яти доменів: домен долоні, домен рухливих пальців, домен великого пальця, домен екзонуклеази (активний у Pol ϵ і Pol δ , але неактивний у Pol α -праймази) і N-кінцевий домен (NTD) [17]. С-кінцевий домен (CTD), що зв'язується із металом, знаходиться на кінці каталітичних субодиниць і є основою для залучення структурно спорідненої регуляторної субодиниці.

Pol α -праймаза – гетеротетрамер із двох праймазних субодиниць та двох полімеразних. На реплікаційній вилці Pol α -праймаза взаємодіє з гомотримерним білком Ctf4, який, у свою чергу, з'єднується з комплексом GINS [18]. Згідно з дослідженнями з використанням електронної мікроскопії різних комплексів CMG, Pol α -праймаза і Pol ϵ знаходяться на протилежних сторонах комплексу CMG [19]. Таким чином, Pol α -праймаза займає близьке до розкрученого відстаючого ланцюга положення [20,21].

Pol δ загалом складається із трьох субодиниць, але в людини є ще одна невелика додаткова четверта субодиниця.

Pol ϵ складається із чотирьох субодиниць. Каталітична субодиниця Pol ϵ приблизно в два рази більша за таку саму у Pol δ . Її полімеразна та екзонуклеазна активності обмежені N-кінцем, в той час як C-кінець відіграє більше структурну роль. Pol ϵ більш процесивна (число нуклеотидів, що додаються за подію зв'язування більше) ніж Pol δ . Дослідження криоелектронної мікроскопії показує, що у даної полімерази мобільність між субодиницями є значною [22].

Впродовж останніх років було виявлено багато різних мутацій в екзонуклеазних доменах полімераз сімейства В людини при ракових захворюваннях, що пов'язані із «гіпермутованими» та «ультрамутованими» утвореннями [23, 24]. Через великі значення мутаційного навантаження ці пухлини легше піддаються імунотерапії. До того ж активні центри полімераз позбавлені мутаційного впливу, що робить їх шкідливими для пухлин.

Окремо в цьому сімействі треба виділити полімеазу ζ . Вона належить до класу репаративних полімераз і задіяна в транслезійному синтезі.

1.2. Репаративні полімерази

Генотоксичні матеріали пошкоджують ДНК, а репаративні ДНК-полімерази відновлюють структуру ланцюга у випадках, коли звичайні реплікативні полімерази просто зупинялися б [25-27]. Представниками таких білків є сімейство полімераз Y, один представник із сімейства описаного вище – Pol ζ , а також Pripol. Також у процесі транслезійного синтезу можуть бути задіяні сімейства X та A полімераз. Дослідження TLS-полімераз, схильних до мутагенних помилок, можливо, зможе пояснити генетичні мутації в різних типах злоякісних новоутворень. Ракові клітини можуть використовувати механізм TLS, щоб обійти дію хіміотерапевтичних засобів. Залучення полімераз TLS в місці пошкодження може сприяти появі нових мутацій, які в кінцевому

підсумку призводять до хеморезистентності і, як наслідок, до виживання ракових клітин. У людей посттрансляційна модифікація білків, пов'язаних з TLS, у відповідь на УФ або інші агенти, що пошкоджують ДНК, підвищує спорідненість TLS-полімераз до пошкодженої ділянки порівняно з неушкодженою ДНК. Таким чином, ця модель дає можливе пояснення перемикання класів від високоточних реплікативних до схильних до помилок полімераз TLS [28].

1.2.1. Y сімейство полімераз

Pol η – найперший описаний представник ДНК-полімераз Y сімейства [29]. Складається із звичних трьох доменів (долоні, пальців та великого пальця) і додатково із ще одного унікального домену PAD. Особливістю цієї полімерази є наявність при каталітичному центрі цієї полімерази третього іону Mg^{2+} [30], котрий не тільки стабілізує вхідний пірофосфат, але також є надважливим для проходження самої реакції приєднання. Тобто наявність вище зазначеного металу накладає певні обмеження на включення нуклеотидів у ланцюг ДНК.

Pol ι – найбільш ефективно і точно приєднує піримідини, ніж пурини [25-27]. Також ця полімераза є найбільш мутагенною і зв'язує нуклеотиди за принципом Хугстинівських пар. Каталітичними іонами радше виступають Mn^{2+} , а не Mg^{2+} [31].

Pol κ – трансфераза, яка приймає участь у багатьох процесах: альтернативному сплайсингу, синтезі ДНК, транслезійному синтезі, репарації ДНК. Домен долоні пов'язаний із каталізом [32]. Особливість цієї полімерази в тому, що її ортологи існують у бактерій та архей, на відміну від інших представників сімейства Y [33].

Rev1 примітний своєю здатністю виконувати некаталітичні функції. Він виконує роль каркасного білка для збирання транслезійного комплексу [27].

Інактивація полімерази ϵ та призводить до генетичного захворювання, а саме – пігментної ксеродерми [34,35]. Позитивна регуляція та інгібування Pol ι пов'язані із різними видами раку. Крім того, активність Pol ι та інших неканонічних полімераз людини має небажані побічні ефекти, а також впливає на токсичність нуклеозидних інгібіторів зворотної транскриптази, які використовуються в противірусній терапії проти ВІЛ та гепатиту В [36,37]. Пов'язані із раковими пухлинами точкові мутації в Pol κ впливають на функції останньої по-різному. Ті мутації, які понижують активність ензиму, роблять клітини більш чутливими до різноманітних негативних впливів на ДНК, зокрема до хіміотерапії. Ті ж, які змінюють точність цієї полімерази, впливають на геномну стабільність ракових клітин [33].

1.2.2. А сімейство ДНК-полімераз

Pol γ , Pol ν , Pol θ – представники сімейства А, що представлені у людини. Pol γ має високу точність, локалізується в мітохондріях і, відповідно, виконує репаративну та реплікативну функції над мітохондріальними ДНК [38]. Pol ν , Pol θ – точність низька, ядерні полімерази, приймають участь в процесах транслезійного синтезу та репарації. В даному сімействі представлені такі домени: великого пальця, рухливих пальців, екзонуклеази, долоні. Екзонуклеазний домен неактивний. У Pol ν характерна риса виражена атипичним відкритим доменом рухливих пальців, що дає їй зв'язувати вхідний, але не матричний нуклеотид. Тобто це змінює класичну динаміку процесу утворення нової пари основ.

1.2.3. PrimPol сімейство

Це окрема полімераза, що радикально відрізняється від усіх інших. Локалізується і у архей, і у еукаріотів. У неї виражена не тільки транслезійна активність, але також і праймазна [39-41]. Дана особливість допомагає уникати пошкоджень на ДНК через ТЛС, а також надає можливість цьому білку пропускати пошкоджені основи та почати синтезувати ДНК знову. Ще одна характерна риса PrimPol – приєднувати дезоксинуклеотидтрифосфат для відновлення синтезу, на відміну від інших праймаз. Нещодавні дослідження кристалічної структури полімерази показали, що в неї є два модулі, а саме N-кінцевий ModN і C-кінцевий ModC [42]. Перший взаємодіє із матричними ланцюгом і основою, в той час як другий – із парою основ, що утворюється.

1.2.4. X сімейство полімераз

Переважаючі представники даного сімейства беруть участь у репарації ДНК, проте Pol β , Pol μ та Pol λ можуть також сприяти транслезійному синтезу [27].

Pol β заповнює короткі нуклеотидні проміжки, що утворюються під час ексцизійної репарації основ. Саме ця полімераза була першою серед усіх, які на даний момент відомі, охарактеризована та досліджена [43]. Нещодавно було відкрито перехідний третій іон металу, який знаходиться майже в тому ж місці, де і у Pol η [44]. Проте в полімеразі β він з'являється після утворення ковалентного зв'язку.

За допомогою кристалографії було виявлено перехідний третій іон також і у Pol μ . Так само як і в Pol β , він з'являється після утворення ковалентного зв'язку, проте в полімеразі μ він може бути виявленим тільки коли реакція проводиться з Mn^{2+} , а не з Mg^{2+} [45].

Pol λ пов'язана із ексцезійною репарацією основ – процесом видалення із ДНК пошкодженого фрагменту ланцюга. Також ця полімераза може приймати участь в транслейному синтезі ДНК.

Розділ 2. Метод зонтичної вибірки

Однією з головних характеристик для більшості біофізичних процесів є різниця вільної енергії між різними станами системи. В цілому, різниця енергій між двома станами не показує повної картини процесів, що відбувається в системі. Між станами може існувати енергетичний бар'єр, що впливає на константи швидкості реакції між цими станами.

Щоб описати перехід системи між двома станами зручно ввести координату реакції ξ , що дозволяє безперервно описувати зміни системи. В такому випадку вільну енергію Гіббса для системи можна записати в такому вигляді:

$$G(\xi) = -kT \ln P(\xi), \quad (1)$$

де $P(\xi)$ – густина ймовірності; k – стала Больцмана; T – температура.

Таким чином, теоретично можливо знайти профіль вільної енергії, отримавши розподіл ймовірностей зі звичайної молекулярно-динамічної симуляції. Однак, є недолік. Ділянки конформаційного простору, що знаходяться далеко від мінімуму потенційної енергії, будуть представлені неналежним чином у результативній вибірці, або можуть бути не досягнуті взагалі (якщо різниця енергій набагато більша за kT). Для вирішення цієї проблеми було запропоновано різні методи прискореної вибірки: метадинаміка, термодинамічне інтегрування і зонтична вибірка. У всіх цих методах вводиться

додатковий потенціал, який знижує потенціальні бар'єри, що дозволяє системі легше обходити конформаційний простір. В ході аналізу результатів ефект від даного потенціалу вираховується з отриманого профілю вільної енергії.

В цьому методі в якості зовнішнього потенціалу використовується гармонічний потенціал, прикладений до координати реакції. Моделювання методом зонтичної вибірки проводиться в три етапи: підготовчий, основний та аналіз даних. На першому етапі проводиться генерація початкових станів для подальших досліджень, обираються фрейми. На другому етапі до кожного із обраних на попередньому кроці фреймів прикладається зонтичний потенціал і проводиться симуляція. Цей крок найдовший, через певні особливості і складнощі підбору коректного часу симуляції. Далі, наступний етап, це побудова профілю вільної енергії за результатами зонтичної вибірки. Існує декілька варіантів методів для аналізу отриманих даних, а саме: метод зважених гістограм, оціночна функція Бенетта для багатьох станів, зонтичне інтегрування. Найпоширенішим є метод зважених гістограм. Він дає достатньо точну оцінку за короткий проміжок часу.

Дослідження зміни вільної енергії вздовж координати реакції в процесі міграції нуклеотиду з каріоплазми до активного центру фермента проводили методом зонтичної вибірки. Umbrella Sampling дозволяє досить повно описати стани системи і оцінити зміну вільної енергії в процесах асоціації та дисоціації молекулярних комплексів.

Розділ 3. Практична частина

Методи *in silico* широко використовуються в сучасному світі і мають ряд переваг на відміну від методів *in vivo* та *in vitro*. Наприклад, вони дають змогу поррахувати такі мікроскопічні параметри, які за використання інших методів визначити неможливо; також вони потребують меншої кількості часу і матеріалів.

Різниця вільних енергій системи між різними її станами являють своєю одну з основних характеристик більшості біофізичних процесів. Але не достатньо знати лише різницю енергій між двома станами, між ними може бути енергетичний бар'єр, що в свою чергу впливає на швидкість проходження реакції, тобто швидкість переходу між цими двома станами. Для опису такого переходу вводять координату реакції ζ , що дозволяє описати систему безперервно.

В даній роботі використовувався метод зонтичної вибірки (Umbrella Sampling). Він дозволяє доволі точно описати стани системи і оцінити зміну вільної енергії при утворенні і дисоціації комплексів. Особливістю цього методу є введення додаткового (зонтичного) потенціалу. Без такого потенціалу цей метод не можна було б використовувати в системах з бар'єром.

Під час даного експерименту використовувались такі ресурси : Protein Data Bank (PDB), Swiss Param, Excel. Та наступні програми: Discovery Studio Visualization (DS Visualization), Visual Molecular Dynamics (VMD) та GROMACS.

За основу був взятий потрібний комплекс Pol η -ДНК людини (5L9X) взятий із Protein Data Bank. Він складається із ДНК, трьох іонів мангану, безпосередньо полімерази, а також деоксиаденозинтрифосфату (Рис. 3.1).

Для створення порівняльної вибірки було проведено маніпуляцію на матричній ДНК. Досліджуваний матричний оригінальний нуклеотид замінювали за допомогою функції мутування в DS Visualizer на інші – G, A, C (Рис. 3.2 – 3.5).

Для експериментів було обране силове поля CHARMM27. Для початку створюються топологія та параметри для ліганду, які відсутні в даному силовому полі. Для цього використали можливості ресурсу Swiss Param [46], куди було завантажено файли у форматі .mol2. Після опрацювання запиту ми отримали файл топології для ліганду у форматі .itp. Наступним кроком було об'єднання комплексу з опрацьованим лігандом. Усі подальші розрахунки проводилися з використанням GROMACS.

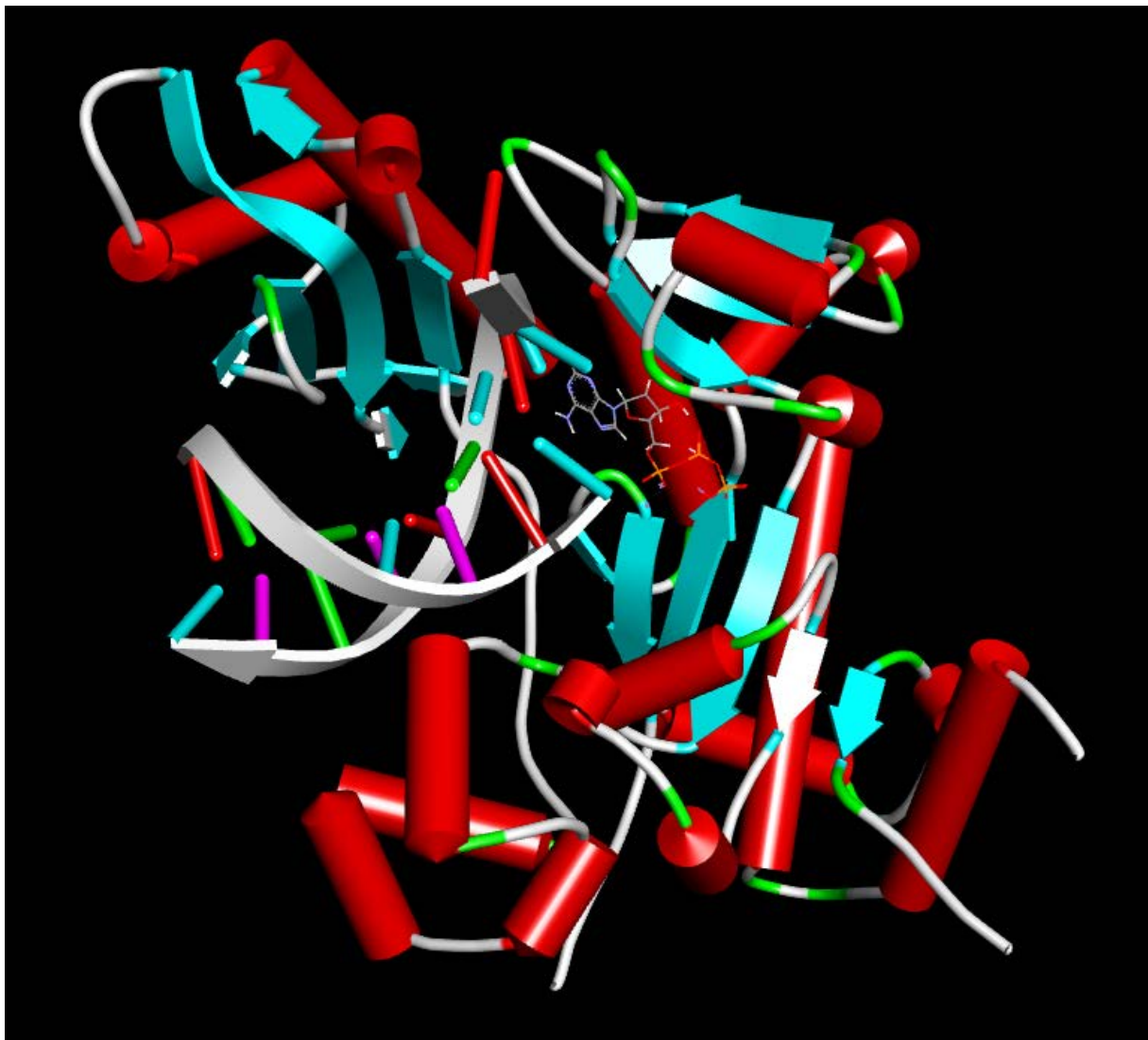


Рисунок 3.1. Комплекс 5L9X візуалізований за допомогою DS Visualizer

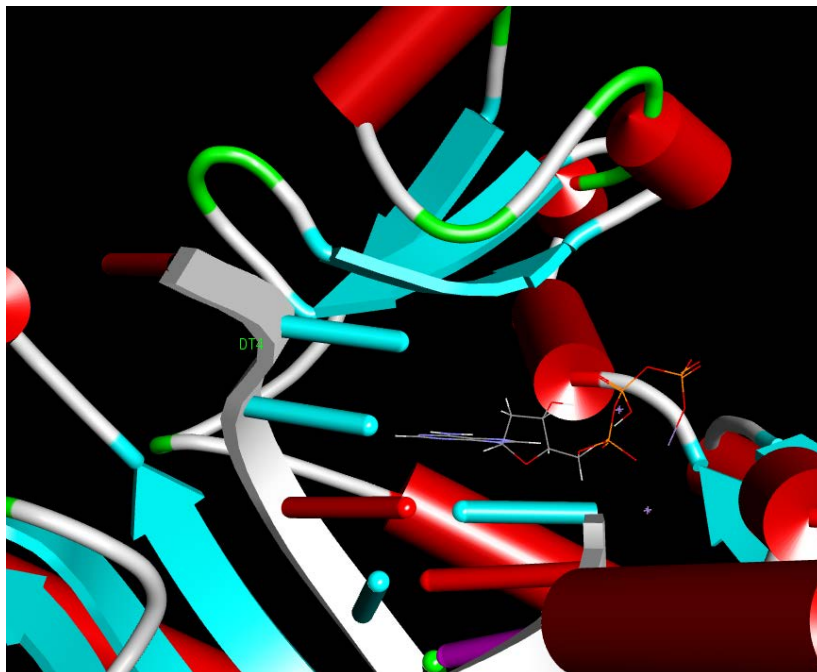


Рисунок 3.2. Комплекс із матричним деокситиміном (вихідний нуклеотид) в активному сайті.

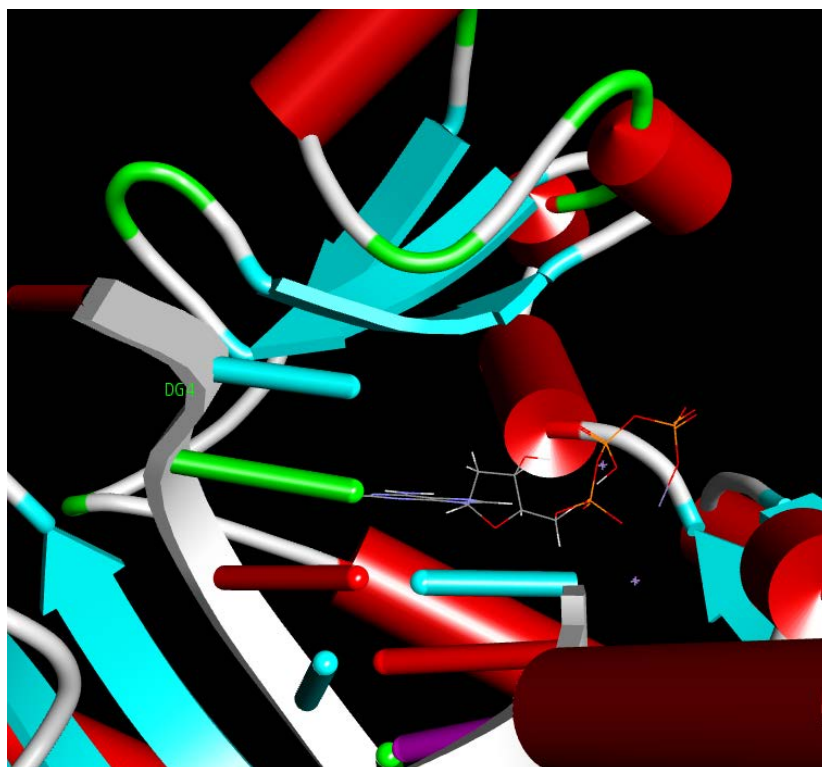


Рисунок 3.3. Комплекс із матричним деоксигуанозином (мутований) в активному сайті.

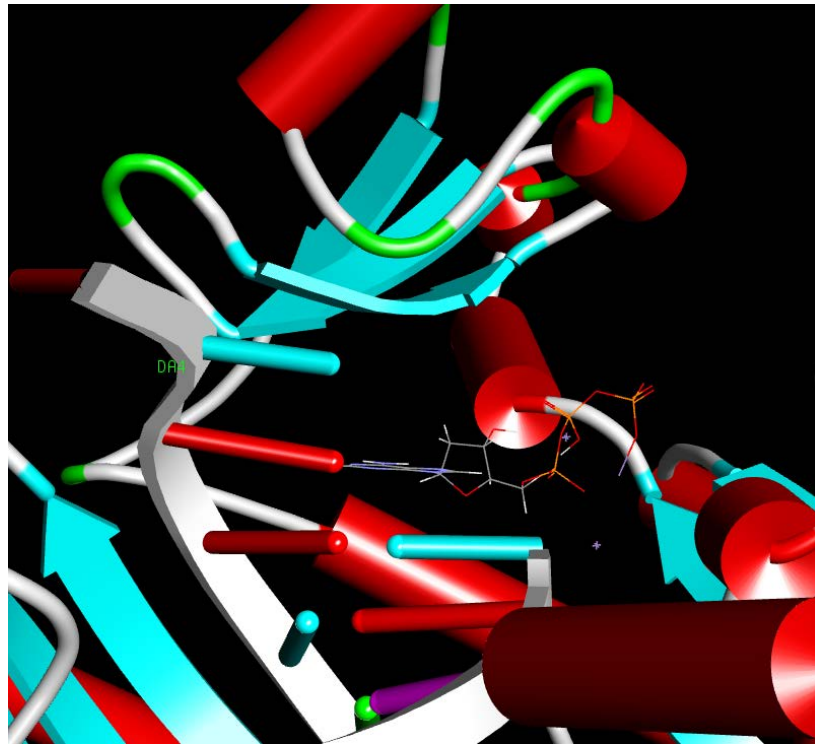


Рисунок 3.4. Комплекс із матричним дезоксиаденозином (мутований) в активному сайті.

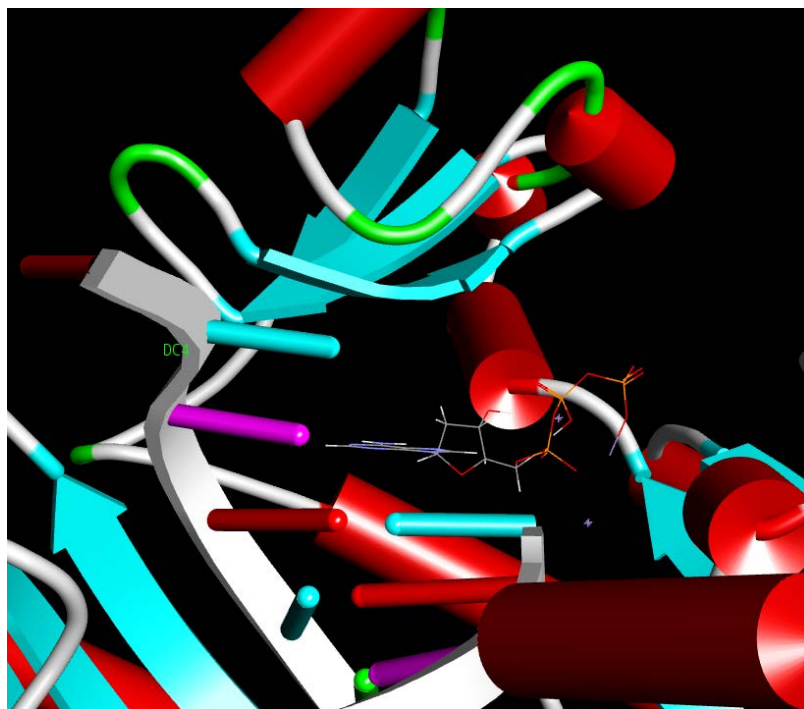


Рисунок 3.5. Комплекс із матричним деоксицитозином (мутований) в активному сайті.

Для будь яких обрахунків в молекулярній динаміці потрібно задати стандартні параметри для атомів та зав'язків між ними, це дасть представлення енергії системи. Для цього було обране силове поле CHARMM27.

Аби отримати результати максимально близькі до реальності, необхідно додати розчинник. Це значно сповільнить розрахунки, тому що молекул води має бути багато, аби повністю, з усіх сторін оточити молекулу. У даному дослідженні було використано трьохточкову жорстку модель води, що не має ефекту поляризації - TIP3P (Transferable Intermolecular Potential with 3 Points). Ця модель дозволяє отримувати досить точні результати розрахунків для моделювання білкових систем. Було створено попередньо бокс довкола молекули, а потім його заповняли розчинником.

Для того, щоб врівноважити систему по заряду до системи додатково потрібно додати іони. Вони замінюють деяку кількість молекул води (у випадку даного експерименту ця кількість дорівнює $N_a = 72$, а $Cl = 64$). Далі потрібно врівноважити отриману модель за температурою та тиском, які повинні бути максимально наближені до умов всередині клітини.

Наступний крок – проведення пулінгу. Загалом можливих напрямків 7. Був обраний напрямок YYY, оскільки попередні дослідження полімераз показали, що це один з найвдаліших випадків і білок може рухатись по всім осям. Тривалість пулінгу складала 0,5 нс, а результати розрахунків записувались кожні 0,2 нс. В результаті було отримано файли із геометрією. Із них обрали 24 файли із геометрією на відстані $\sim 0,1$ нм для кожного із чотирьох експериментів з різними матричними нуклеотидами.

Після чого з на цих геометрії був проведений аналіз методом зонтичної вибірки який використовується для систем з енергетичним бар'єром.

Отримані в результаті геометрії аналізувалися методом зважених гістограм, звідки були отримані файли .xvg. Далі було побудовано профілі вільної енергії для кожного матричного нуклеотиду окремо на основі даних із цих файлів (Рис. 3.6 – 3.8). На даних графіках можемо бачити криві потенціалу середньої сили для трьох матричних нуклеотидів Т, А та Г, а також відповідні гістограми. Отримані результати свідчать про наявність потенційних бар'єрів. Тобто бар'єр є прешекодою для ліганду, яку він має пройти для того, щоб приєднатися до матричного нуклеотиду в активному центрі, або не приєднатися. Для нуклеотиду А майже немає потенційного бар'єру (Рис3.7), тоді як для двох інших матричних нуклеотидів маємо бар'єри на відстані приблизно 1 нм для Т і на відстані 1,4 нм для Г. На Рисунку 3.7 А на відстані 1,5 нм наявний певний провал прямої до 0, проте ця область не є бар'єром, а тільки

показує, що на даній відстані була перенасиченість вибірки, через що програма з нею не впоралась. Так само і на Рисунку 3.6 А в діапазоні від 3,8 до 4,2 маємо певні аномалії кривої, що говорить про перенасиченість даними.

На основі отриманих після основного етапу моделювання методом зонтичної вибірки файлів .edr було побудовано графіки залежності вільної енергії від відстані, на яку віддалявся ліганд від активного центру. Визначено лінії тенденції, які чітко показують, що деоксиаденозинтрифосфат з'єднується із матричним тиміном, а також можливе утворення зв'язку із аденіном. При тому що до матричної основи А він приєднується швидше і з меншими енергетичними витратами на проходження бар'єру, проте імовірність того, що вхідний нуклеотид залишиться в такому положенні і утворить нову пару основ – невелика, що можна побачити із Рисунку 3.10, оскільки вільна енергія значно зростає із відстанню, а у випадку нуклеотида Т (Рисунок 3.11) - вільна енергія прямує до нуля, що каже про високу ймовірність приєднання ліганду в сайті зв'язування. На Рисунках 3.9 та 3.12 з ліній тенденцій, можна побачити, що ймовірність приєднання вільного нуклеотиду в такі сайти зв'язування - дуже низька, оскільки система не прямує до нуля, що говорить про відсутність стабілізації.

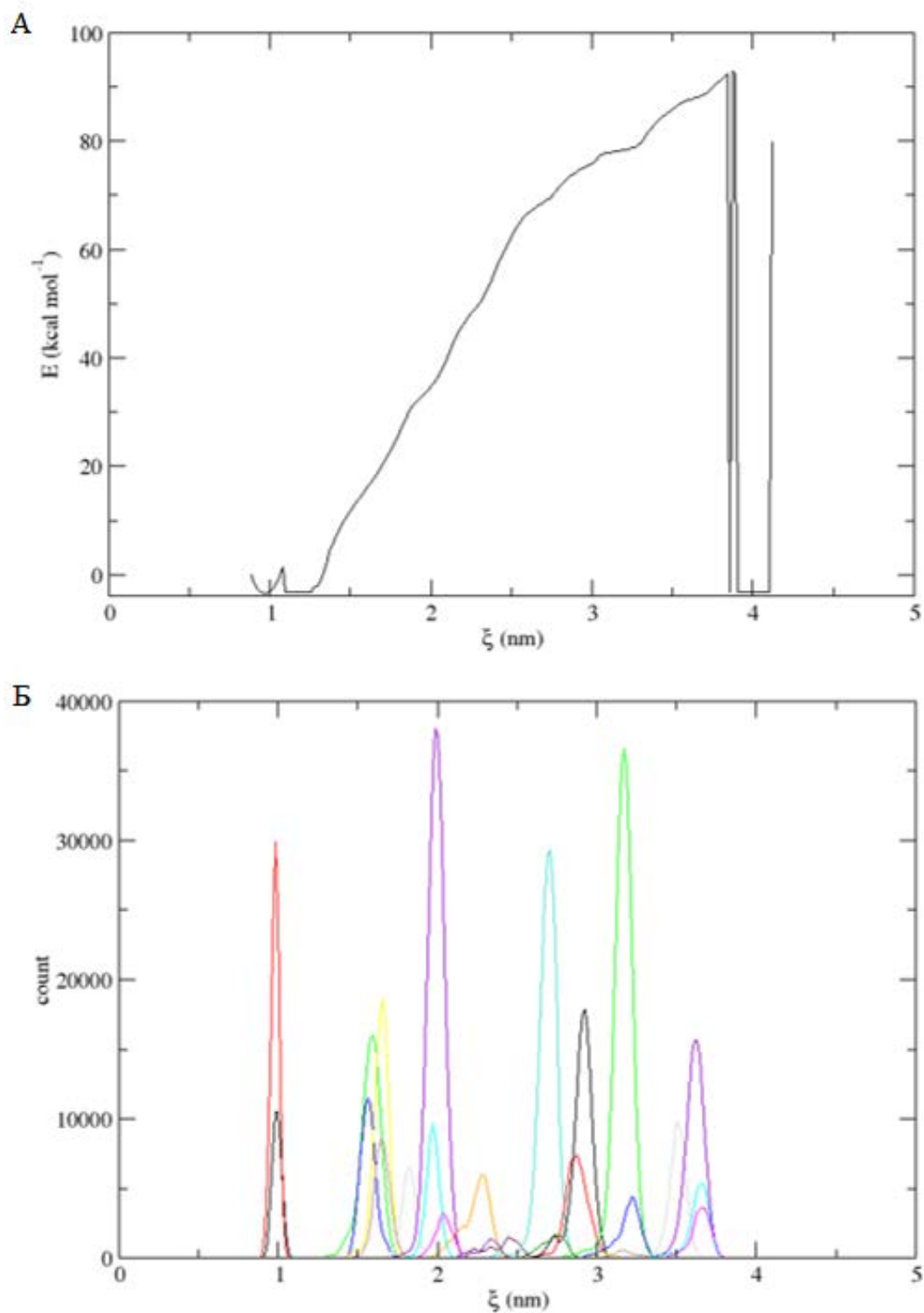


Рисунок 3.6. А - графік залежності потенціалу середньої сили від відстані між центрами мас у комплексі 5L9X з оригінальним матричним нуклеотидом тиміном; Б - гістограма повноти охоплення фазового простору.

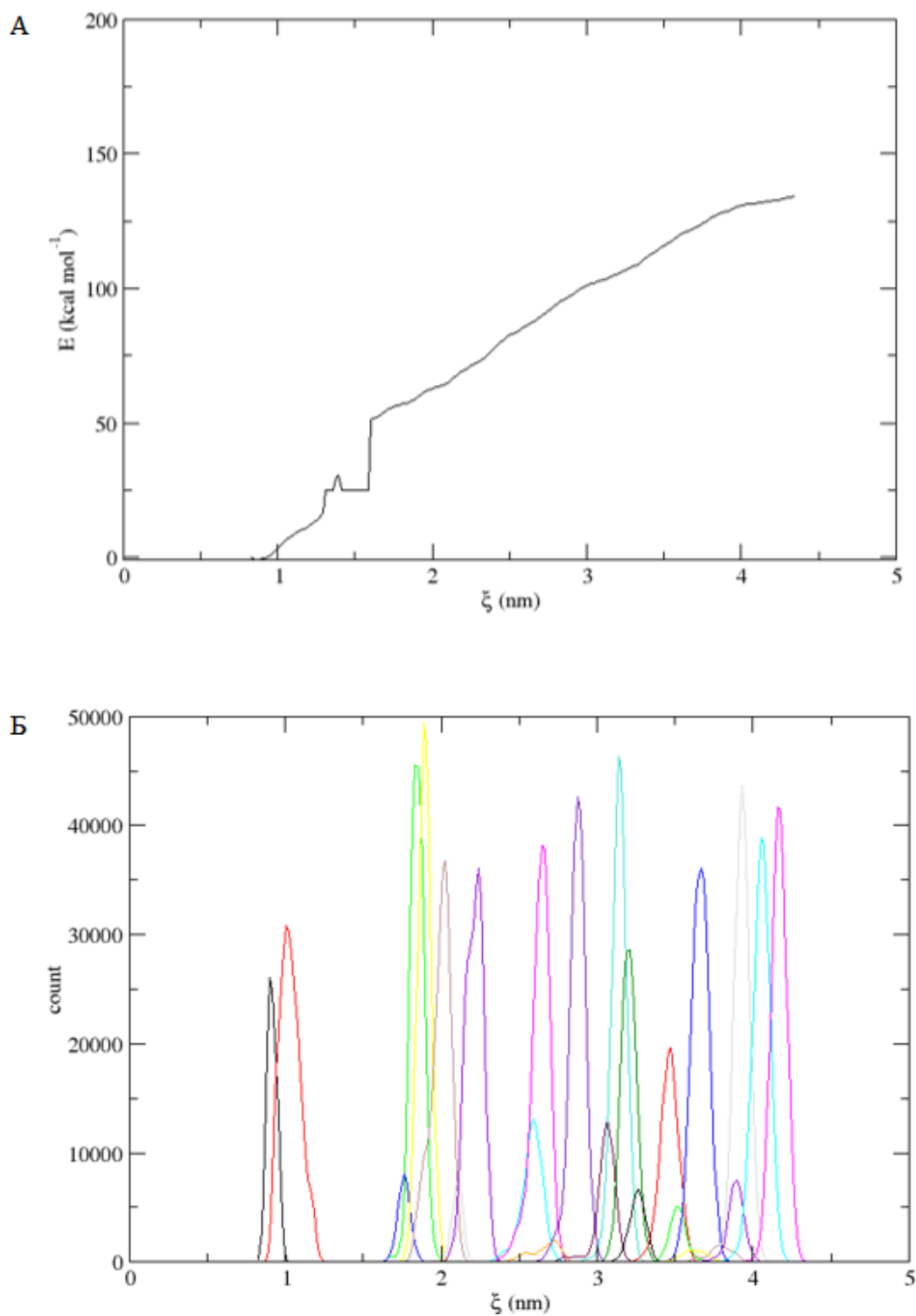


Рисунок 3.7. графік залежності потенціалу середньої сили від відстані між центрами мас у комплексі 5L9X з промтованим матричним нуклеотидом аденіном; Б - гістограма повноти охоплення фазового простору.

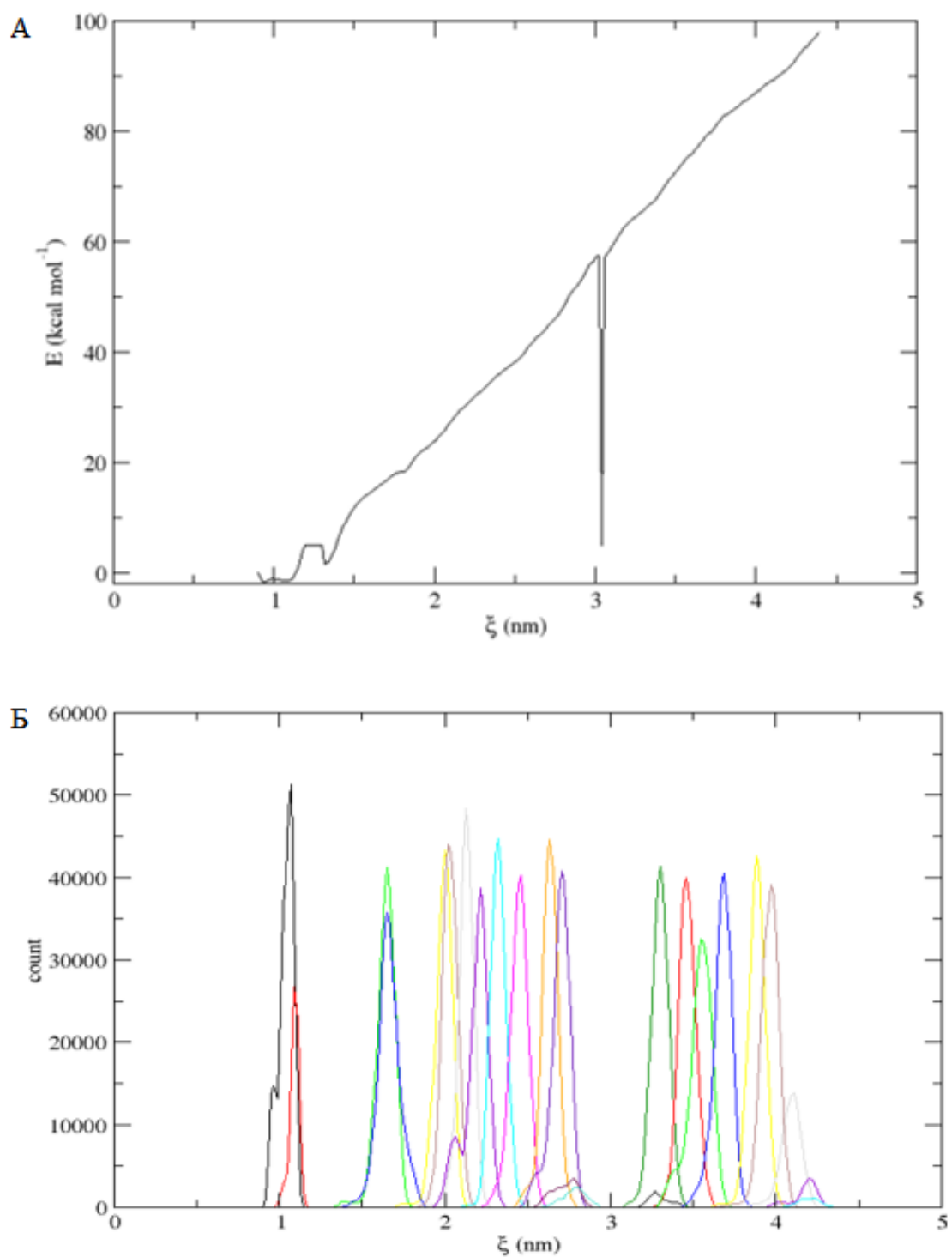


Рисунок 3.8. А - графік залежності потенціалу середньої сили від відстані між центрами мас у комплексі 5L9X з промтованим матричним нуклеотидом гуаніном; Б - гістограма повноти охоплення фазового простору.

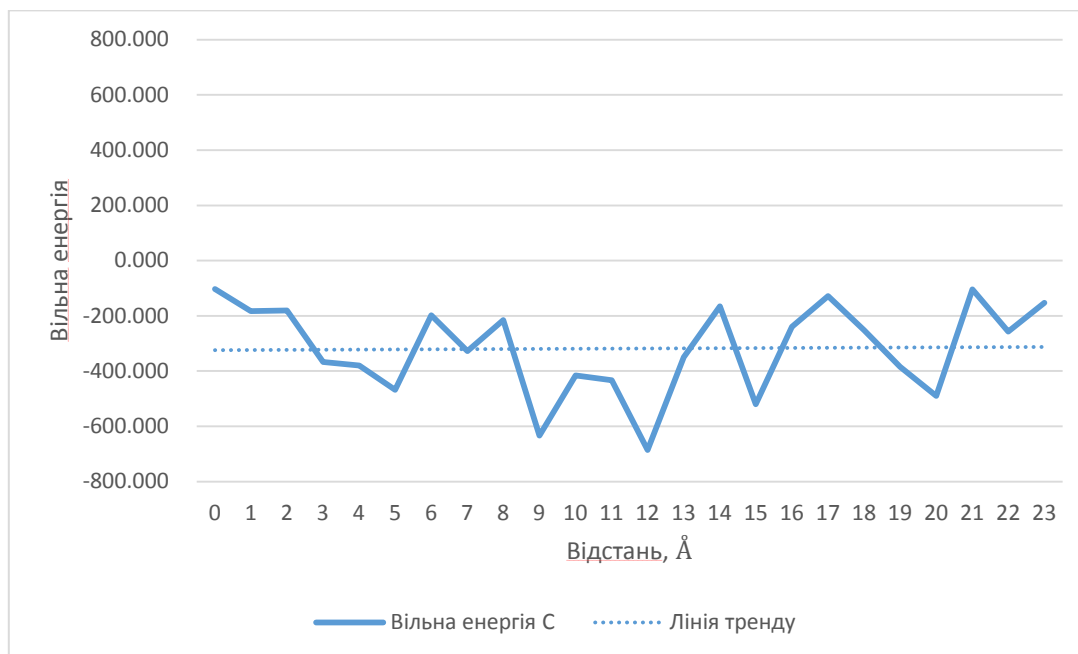


Рисунок 3.9. Графік залежності вільної енергії від відстані між деоксиаденозинтрифосфатом та активним сайтом із матричним нуклеотидом С.

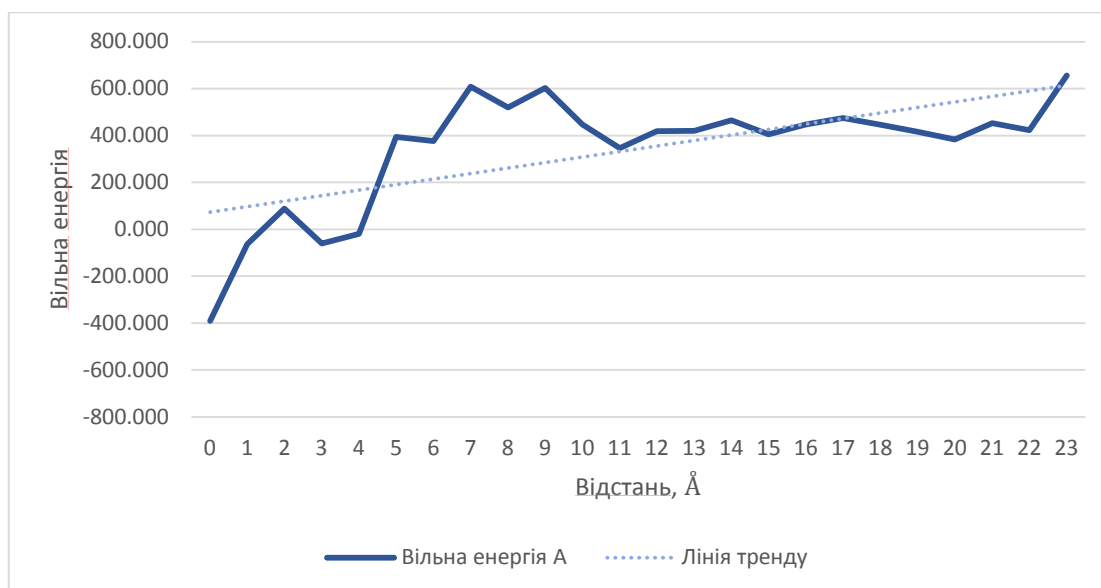


Рисунок 3.10. Графік залежності вільної енергії від відстані між деоксиаденозинтрифосфатом та активним сайтом із матричним нуклеотидом А.

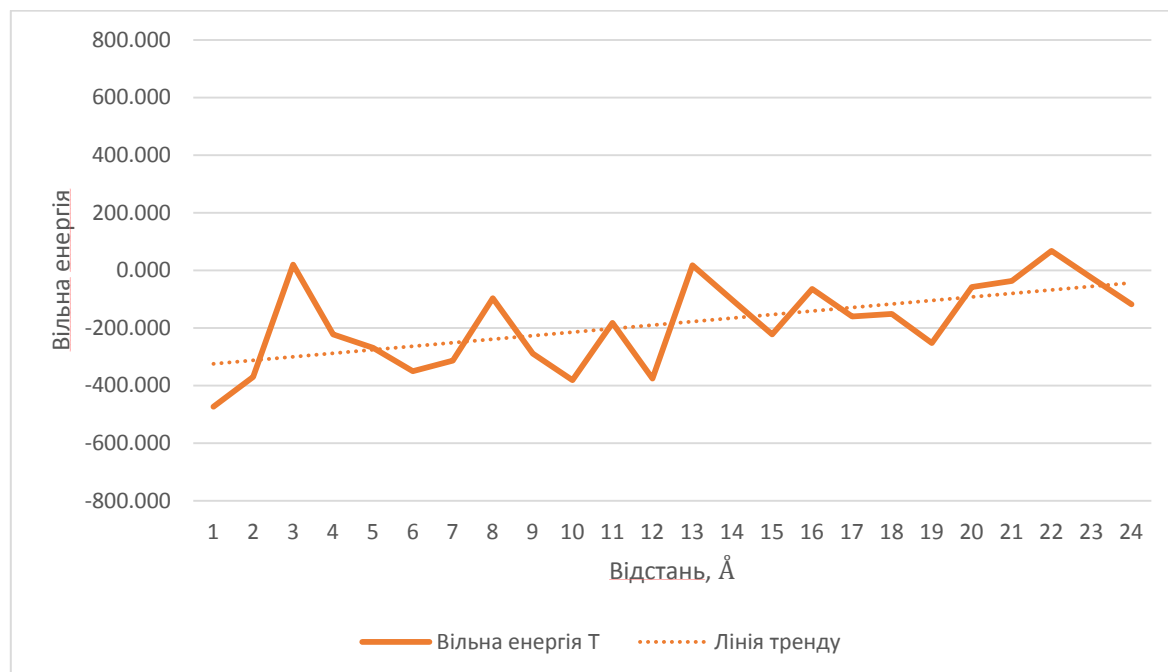


Рисунок 3.11. Графік залежності вільної енергії від відстані між деоксиадозинтрифосфатом та активним сайтом із матричним нуклеотидом Т.

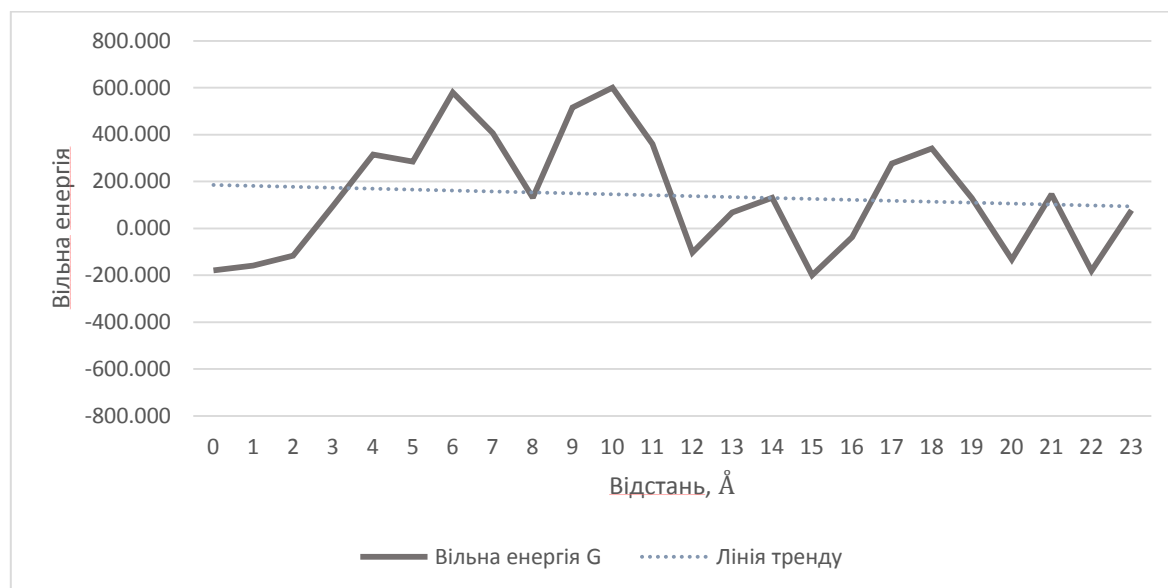


Рисунок 3.12. Графік залежності вільної енергії від відстані між деоксиадозинтрифосфатом та активним сайтом із матричним нуклеотидом Г.

Висновки

В результаті проведеної роботи можна зробити декілька важливих висновків:

1. В залежності від того, який нуклеотид знаходиться на матричному ланцюзі, до якого приєднується деоксиаденозинтрифосфат, профіль взаємодії значно змінюється. У випадку, коли нуклеотид надходить не з розчину, а проходить деякий шлях уздовж полімерази, він витрачає більше енергії для входу в активний центр, де зв'язується/не зв'язується.
2. Матричний нуклеотид впливає не тільки на вільний нуклеотид безпосередньо в активному центрі, а також і на відстані. Цей вплив це фактично створення певних енергетичних бар'єрів. Важливо відзначити, що різні матричні нуклеотиди впливають на ліганд з індивідуальними, для кожного з чотирьох, характеристиками. Для матричних нуклеотидів G, C ліганд буде знаходитися практично в рівноважних позиціях, що показує лінія тренду на графіках зміни вільної енергії.
3. Приєднання деоксиаденозинтрифосфату в активному центрі до матричних нуклеотидів T і A, виходячи із отриманих результатів, які відображені на графіках, є найбільш вірогідним. При тому що до матричної основи A він приєднується швидше і з меншими енергетичними витратами через відсутність там бар'єру, проте імовірність того, що вхідний нуклеотид залишиться в такому положенні і утворить нову пару основ – невелика, що можна побачити із графіків залежності вільної енергії від відстані.

4. Підсумовуючи, для даної полімерази, характерною рисою є наявність енергетичних бар'єрів на етапі приєднання вільного нуклеотиду в активному центрі.
5. Загалом, результати щодо того, який матричний нуклеотид має найбільші шанси зв'язатися із деоксиаденозинтрифосатом, є очікуваними, оскільки таким нуклеотидом виявився Т.

Дана полімераза приймає участь в транслейному синтезі, що є важливим фактором утворення помилок на синтезованих дочірніх ланцюгах. Це призводить до хеморезистентності при лікуванні раку. Подальше дослідження полімерази ета може виявити нову фармакологічну мішень, яка стане основою для вирішення проблеми хеморезистентності.

Джерела

1. Schrödinger E. 1944 What is life? Cambridge, UK: Cambridge University Press.
2. Von Neumann J. 1949 Theory and organization of complicated automata. In Theory of self-reproducing automata (ed. AW Burks), pp. 29–87. Urbana, IL: University of Illinois Press.
3. Wills PR. 2016 DNA as information. *Phil. Trans. R. Soc. A* 374: 20150417.
4. Watson JD, Crick FHC. 1953 Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature* 171, 737–738. (doi:10.1038/171737a0)
5. Watson JD, Crick FHC. 1953 Genetical implications of the structure of deoxyribose nucleic acid. *Nature* 171, 964–967. (doi:10.1038/171964b0)
6. Burgers M.J. P., Kunkel A. Th. 2017 Eukaryotic DNA Replication Fork. *Annual Reviews Volume 86*, pp 417-438
7. Okazaki R, Okazaki T, Sakabe K, Sugimoto K, Kainuma R, et al. In vivo mechanism of DNA chain growth. *Cold Spring Harb Symp Quant Biol.* 1968;33:129–43.
8. Tye BK, Nyman PO, Lehman IR, Hochhauser S, Weiss B. Transient accumulation of Okazaki fragments as a result of uracil incorporation into nascent DNA. *PNAS.* 1977;74:154–57.
9. Nick McElhinny SA, Watts BE, Kumar D, Watt DL, Lundstrom EB, et al. Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *PNAS.* 2010;107:4949–54.

10. I.R. Lehman et al. Enzymatic synthesis of deoxyribonucleic acid: I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli* *J Biol Chem* (1958)
11. Pinheiro B V. Engineering-driven biological insights into DNA polymerase mechanism *Current Opinion in Biotechnology* Volume 60, December 2019, Pages 9-16
12. Tsai YC, Johnson KA: A new paradigm for DNA polymerase specificity. *Biochemistry* 2006, 45:9675-9687.
13. O'Donnell M, Li H: The eukaryotic replisome goes under the microscope. *Curr Biol* 2016, 26:R247-256.
14. Burgers PMJ, Kunkel TA: Eukaryotic DNA replication fork. *Annu Rev Biochem* 2017, 86:417-438.
15. Pursell ZF, Isoz I, Lundstrom EB, Johansson E, Kunkel TA: Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* 2007, 317:127-130.
16. Johnson RE, Klassen R, Prakash L, Prakash S: A major role of DNA polymerase delta in replication of both the leading and lagging DNA strands. *Mol Cell* 2015, 59:163-175.
17. Doublet S, Zahn KE: Structural insights into eukaryotic DNA replication. *Front Microbiol* 2014, 5:444.
18. Simon AC, Zhou JC, Perera RL, van Deursen F, Evrin C, Ivanova ME, Kilkenny ML, Renault L, Kjaer S, Matak-Vinkovic D et al.: A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. *Nature* 2014, 510:293-297.
19. Sun J, Shi Y, Georgescu RE, Yuan Z, Chait BT, Li H, O'Donnell ME: The architecture of a eukaryotic replisome. *Nat Struct Mol Biol* 2015, 22:976-982.

20. Georgescu R, Yuan Z, Bai L, de Luna Almeida Santos R, Sun J, Zhang D, Yurieva O, Li H, O'Donnell ME: Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. *Proc Natl Acad Sci U S A* 2017, 114:E697-E706
21. Zhou JC, Janska A, Goswami P, Renault L, Abid Ali F, Kotecha A, Diffley JFX, Costa A: CMG-Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis
22. Asturias FJ, Cheung IK, Sabouri N, Chilkova O, Wepplo D, Johansson E: Structure of *Saccharomyces cerevisiae* DNA polymerase epsilon by cryo-electron microscopy. *Nat Struct Mol Biol* 2006, 13:35-43.
23. Rayner E, van Gool IC, Palles C, Kearsley SE, Bosse T, Tomlinson I, Church DN: A panoply of errors: polymerase proofreading domain mutations in cancer. *Nat Rev Cancer* 2016, 16:71-81.
24. Campbell BB, Light N, Fabrizio D, Zatzman M, Fuligni F, de Borja R, Davidson S, Edwards M, Elvin JA, Hodel KP et al.: Comprehensive analysis of hypermutation in human cancer. *Cell* 2017, 171:1042-1056 e1010
25. Prakash S, Johnson RE, Prakash L: Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu Rev Biochem* 2005, 74:317-353.
26. Washington MT, Carlson KD, Freudenthal BD, Pryor JM: Variations on a theme: eukaryotic Y-family DNA polymerases. *Biochim Biophys Acta* 2010, 1804:1113-1123.
27. Vaisman A, Woodgate R: Translesion DNA polymerases in eukaryotes: what makes them tick? *Crit Rev Biochem Mol Biol* 2017, 52:274-303

- 28.Saha P, Mandal T, Talukdar AD, Kumar D. et al: DNA polymerase eta: A potential pharmacological target for cancer therapy. *J Cell Phys* 2020; 236(6): 4106- 4120.
- 29.Trincao J, Johnson RE, Escalante CR, Prakash S, Prakash L, Aggarwal AK: Structure of the catalytic core of *S. cerevisiae* DNA polymerase eta: implications for translesion DNA synthesis. *Mol Cell* 2001, 8:417-426.
- 30.Nakamura T, Zhao Y, Yamagata Y, Hua YJ, Yang W: Watching DNA polymerase eta make a phosphodiester bond. *Nature* 2012, 487:196-201.
- 31.McIntyre J: Polymerase iota - an odd sibling among Y family polymerases. *DNA Repair* 2020, 86:102753
- 32.Goodman M.F., Woodgate R. Translesion DNA polymerases. *Cold Spring Harb Perspect Biol.* 2013;5:a010363. doi: 10.1101/cshperspect.a010363
- 33.Stern H.R., Sefcikova J., Chaparro V.E., Beuning P.J. Mammalian DNA Polymerase Kappa Activity and Specificity. *Molecules.* 2019 24(15): 2805.
- 34.Johnson RE, Kondratik CM, Prakash S, Prakash L: hRAD30 mutations in the variant form of xeroderma pigmentosum. *Science* 1999, 285:263-265.
- 35.Masutani C, Kusumoto R, Yamada A, Dohmae N, Yokoi M, Yuasa M, Araki M, Iwai S, Takio K, Hanaoka F: The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature* 1999, 399:700-704.
- 36.J.A. Brown, L.R. Pack, J.D. Fowler, Z. Suo: Pre-Steady-State Kinetic Analysis of the Incorporation of Anti-HIV Nucleotide Analogs Catalyzed by Human X- and Y-Family DNA Polymerases. *Antimicrob Agents Chemother*, 55 (1) (2011), pp. 276-283.
- 37.J.A. Brown, L.R. Pack, J.D. Fowler, Z. Suo. Presteady state kinetic investigation of the incorporation of anti-hepatitis B nucleotide analogues

- catalyzed by noncanonical human DNA polymerases. *Chem Res Toxicol*, 25 (1) (2012), pp. 225-233
- 38.Szymanski MR, Kuznetsov VB, Shumate C, Meng Q, Lee YS, Patel G, Patel S, Yin YW: Structural basis for processivity and antiviral drug toxicity in human mitochondrial DNA replicase. *EMBO J* 2015, 34:1959-1970
- 39.1. Bianchi J, Rudd SG, Jozwiakowski SK, Bailey LJ, Soura V, Taylor E, Stevanovic I, Green AJ, Stracker TH, Lindsay HD et al.: PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. *Mol Cell* 2013, 52:566-573.
- 40.Garcia-Gomez S, Reyes A, Martinez-Jimenez MI, Chocron ES, Mouron S, Terrados G, Powell C, Salido E, Mendez J, Holt IJ et al.: PrimPol, an archaic primase/polymerase operating in human cells. *Mol Cell* 2013, 52:541-553.
- 41.Wan L, Lou J, Xia Y, Su B, Liu T, Cui J, Sun Y, Lou H, Huang J: hPrimpol1/CCDC111 is a human DNA primase-polymerase required for the maintenance of genome integrity. *EMBO Rep* 2013, 14:1104-1112
- 42.Rechkoblit O, Gupta YK, Malik R, Rajashankar KR, Johnson RE, Prakash L, Prakash S, Aggarwal AK: Structure and mechanism of human PrimPol, a DNA polymerase with primase activity. *Sci Adv* 2016, 2:e1601317.
- 43.Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J: Structures of ternary complexes of rat DNA polymerase beta, a DNA template–primer, and ddCTP [see comments]. *Science* 1994, 264:1891-1903.
- 44.Freudenthal BD, Beard WA, Shock DD, Wilson SH: Observing a DNA polymerase choose right from wrong. *Cell* 2013, 154:157-166
- 45.Jamsen JA, Beard WA, Pedersen LC, Shock DD, Moon AF, Krahn JM, Bebenek K, Kunkel TA, Wilson SH: Time-lapse crystallography snapshots

of a double-strand break repair polymerase in action. *Nat Commun* 2017, 8:253

46.V. Zoete, m. A. Cuendet, a. Grosdidier, o. Michielin: Swissparam, a fast force field generation tool for small organic molecules, *J. Comput. Chem*, 2011, 32(11), 2359-68