

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

ІМЕНІ ТАРАСА ШЕВЧЕНКА

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ВИСОКИХ ТЕХНОЛОГІЙ

Завідувач кафедри молекулярної біотехнології та біоінформатики

Доцент Нипорко Олексій Юрійович

Протокол № \_\_\_\_ засідання кафедри

від “ \_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

## **Оцінка стійкості G-квадруплексів в геномі коронавірусу SARS-CoV-2**

Випускна кваліфікаційна робота магістра

студентки спеціальності 091 Біологія

(високі технології)

ОП «Біоінформатика та структурна

біологія»

**Гребінчук Вероніки Романівни**

Науковий керівник

доцент кафедри молекулярної

біотехнології та біоінформатики

**Нипорко Олексій Юрійович**

Оцінка захисту роботи

---

Київ – 2024 р.

## АНОТАЦІЯ

Гребінчук В.Р. Оцінка стійкості G-квадруплексів в геномі коронавірусу SARS-CoV-2 – Випускна кваліфікаційна робота магістра з спеціальністю 091 Біологія (високі технології), ОП «Біоінформатика та структурна біологія».

У роботі було проведено розрахунок молекулярної динаміки G-квадруплексів коронавірусу SARS-CoV-2. Отримані результати можуть бути використані для подальшого вивчення терапевтичного ефекту G-квадруплексів.

**Ключові слова:** G-квадруплекс, коронавірус SARS-CoV-2, молекулярна динаміка

## ABSTRACT

Veronika Hrebinchuk. Assessment of G-quadruplex Stability in the SARS-CoV-2 Genome - Master's Thesis in Biology (High Technologies), Specialization "Bioinformatics and Structural Biology"

This thesis presents a molecular dynamics simulation study of G-quadruplexes in the SARS-CoV-2 genome. The obtained results can be used for further investigation of the therapeutic potential of G-quadruplexes.

**Keywords:** G-quadruplex, SARS-CoV-2 coronavirus, molecular dynamics

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

COVID-19 - COronaVirus Disease 2019 – коронавірусна хвороба 2019

SARS-CoV-2 – Severe Acute Respiratory Syndrome Coronavirus 2 – Важкий Гострий Респіраторний Синдром Коронавірус 2

ORF – Open Reading Frames – відкрита рамка зчитування

nsp – nonstructursl protein – не структурні білки

RTC – Replication/Transcription Complex – комплекс реплікації та транскрипції

RMSD – Root Mean Square Deviation – це корінь середньоквадратичного відхилення

## ЗМІСТ

<b>ВСТУП</b> .....	6
<b>РОЗДІЛ 1. ОГЛЯД ЛІТЕРАТУРИ</b> .....	8
1.1. G-квадруплекси.....	8
1.1.1. Вплив катіонів та інших чинників на стабільність молекули G-квадруплексів.....	8
1.1.2. Подібності та відмінності між G-квадруплексами РНК і ДНК.....	9
1.1.3. Локалізація та роль G-квадруплексів в клітині.....	10
1.2. Коронавірус SARS-CoV-2 .....	11
1.2.1. Структура геному коронавірусу SARS-CoV-2.....	12
1.2.2. G-квадруплекси як терапевтичні мішені до SARS-CoV-2.....	13
1.3. Вивчення молекулярної динаміки за допомогою програмного забезпечення AMBER і її візуалізація за допомогою VMD.....	14
<b>РОЗДІЛ 2 МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕННЯ</b> .....	16
2.1. Підготовка молекул.....	16
2.2. Підготовка вихідних файлів у CHARMM-GUI за допомогою «Конструктору розчинів».....	16
2.3. Проведення молекулярної динаміки й інтерпретація результатів.....	18
<b>РОЗДІЛ 3 РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ</b> .....	19
3.1. Оцінка стійкості G-квадруплексу 3467 в розчинах NaCl і KCl при 20°C та 36,6°C.....	19
3.2. Оцінка стійкості G-квадруплексу 28903 в розчинах NaCl і KCl при 20°C та 36,6°C.....	23
<b>ВИСНОВКИ</b> .....	28
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</b> .....	29

## ВСТУП

Пандемія COVID-19, викликана вірусом SARS-CoV-2, вразила понад 200 країн і територій та стала глобальною загрозою для здоров'я та життя людей, а також викликом для світової охорони здоров'я. І саме тому, розуміння основних механізмів і розробка інноваційних методів лікування є надзвичайно актуальними.

Одними з перспективних терапевтичних мішеней було визнано G-квадруплекси. G-квадруплекси – це неканонічні спіральні структури, що утворені одним ланцюгом ДНК або РНК. Оскільки G-квадруплекси пригнічують експресію вірусних білків, мотиви цих молекул є одним з ключових елементів впливу на регуляцію життєвого циклу вірусів. Модуляція стабільності та конформації цієї молекули відіграє значну роль у багатьох інших патологічних, а також нейродегенеративних захворювань людини, і саме тому вивчення механізмів націлювання на G-квадруплекси потребує додаткового дослідження.

Дослідження молекулярної динаміки G-квадруплексів в геномі коронавірусу SARS-CoV-2 допоможе оцінити стійкість цієї молекули і допомогти у розробці терапевтичних засобів для лікування COVID-19.

**Мета роботи:** Розрахувати молекулярну динаміку G-квадруплексів коронавірусу SARS-CoV-2 в силовому полі AMBER в пакеті AMBER і оцінити стійкість цих молекул.

### **Завдання:**

1. Підготувати вхідні файли для симуляції молекулярної динаміки на веб-платформі CHARMM-GUI за допомогою вкладки «Конструктор розчинів» і налаштувати параметри для проведення симуляції.
2. Провести симуляцію молекулярної динаміки за допомогою програмного забезпечення AMBER.

3. Візуалізувати та інтерпретувати результати за допомогою програми VMD та оцінити стійкість G-квадруплексів в різних розчинах і за різних температур.

Вирішення поставлених завдань дозволить нам оцінити стійкість G-квадруплексів в геномі коронавірусу SARS-CoV-2.

## РОЗДІЛ 1 ОГЛЯД ЛІТЕРАТУРИ

### 1.1. G-квадруплекси

G-квадруплекси – це неканонічні структури ДНК або РНК, що утворюються повторюваними послідовностями, які багаті на гуанін [1]. Тетрада гуаніну утворюється асоціацією чотирьох основ гуаніну за допомогою водневих зв'язків Хугстіна у вигляді квадратної планарної структури, а дві і більше таких тетрад гуаніну можуть накладатися одна на одну, утворюючи G-квадруплекс [2].

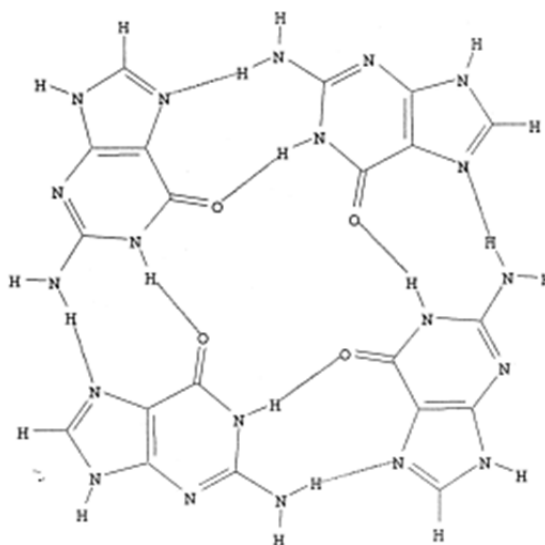


Рис. 1 G-тетрада утворена водневими зв'язками Хугстіна [2]

Залежно від орієнтації, структура G-квадруплексу може бути паралельною (з чотирма G-тетрадами в однаковому розташуванні), антипаралельною (з двома G-тетрадами у протилежному спрямуванні щодо двох інших) або гібридною (одна G-тетрада в протилежному спрямуванні по відношенню до трьох інших) і додатково можуть розрізнятися за своїм положенням петлі [1].

#### 1.1.1. Вплив катіонів та інших чинників на стабільність молекули G-квадруплексів

Підтверджено, що в G-квадруплексах існує центральний канал, який є універсальною та унікальною особливістю всіх квадруплексних структур, і в

якому можуть міститись молекули води або катіони. Катіони стабілізують структуру G-квадруплексу [3]. Найкраще стабілізувати молекулу будуть іони калію. Декілька інших іонів також можуть стабілізувати G-квадруплекси, наприклад, іони натрію, а також в стабілізації можуть брати участь іони двовалентних важких металів, таких як рубідій і цезій, стронцій, талій, кальцій, свинець і барій, проте всі вони стабілізуватимуть молекулу в меншій мірі за калій. Іони металу в каналі завжди знаходяться в кристалічних структурах. Зв'язування катіонів з негативно зарядженими фосфатними групами зменшуватиме електростатичне відштовхування і таким чином забезпечуватиме стабільність складеної конформації. [4]. Присутність певних іонів у каналі також може змінити конформацію молекули та її згортання [5].

Окрім катіонів, стабільність G-квадруплексів значною мірою залежить від кількості G-квартетів, сольватації і довжини петлі. Оскільки стекінг-взаємодії частково пояснюють енергетичний приріст, збільшення кількості квартетів є енергетично вигідним і забезпечуватиме більшу стабільність. Щодо сольватації, то молекули води між лігандами та G-квадруплексами будуть утворювати більшу частину водневих зв'язків і таким чином підвищуватимуть стабільність структури [6]. G-квадруплекси, що мають коротшу петлю є більш стабільними, і ця тенденція спостерігається як у G-квадруплексах ДНК, так і в G-квадруплексах РНК з двома або трьома G-квартетами [7].

### **1.1.2. Подібності та відмінності між G-квадруплексами РНК і ДНК**

Найбільш фундаментальними відмінностями між G-квадруплексами РНК і ДНК є ті ж самі відмінності що і між самими РНК і ДНК. До них відносяться наявність урацилу замість тиміну і моносахариди рибози замість дезоксирибози. Наявність 2'-гідроксильної групи в рибозі дозволяє більше внутрішньо-молекулярних взаємодій і призводить до підвищення стабільності молекули. Заміна тиміну на урацил у петлі олігонуклеотидів

робить G4 більш стабільним та менш гідратованим в межах петлі навколо жолобків.

Також G-квадруплекси ДНК і РНК демонструють чіткі відмінності в специфічності зв'язування катіонів, які впливають на стабілізацію молекули, що було зазначено вище. Катіон калію стабілізує G-квадруплекси як ДНК так і РНК, а катіон натрію мав сильніший вплив лише на G-квадруплекс ДНК.

Таким чином, назважаючи на численні подібності G-квадруплекси РНК є більш компактними, менш гідратованими та часто більш термодинамічно стабільнішими ніж G-квадруплекси ДНК [8].

### **1.1.3. Локалізація та роль G-квадруплексів в клітині**

Насправді, біологічне значення та молекулярні механізми, які лежать в основі G-квадруплексів не до кінця вивчені. Дослідження на рівні транскриптомів демонструють, що G-квадруплекси РНК здебільшого розгортаються *in vivo*. Це може свідчити про те, що G-квадруплекси можуть існувати лише за певних умов, наприклад, під час певних стадій клітинного циклу або захворювань [8].

Також підтверджено, що внутрішньо-молекулярні G-квадруплекси РНК відіграють різні ролі в таких основних біологічних процесах: регуляція трансляції, термінація транскрипції, альтернативний сплайсинг. Дуже цікаво, що G-квадруплекси відіграють роль у локалізації мРНК. Транспорт мРНК до субклітинних компартментів є одним з ключових пост транскрипційних механізмів, який буде забезпечувати локалізованих синтез білка і є надзвичайно важливим для асиметричних клітин, таких як нейрони, де транскрибовані мРНК долають великі відстані до місць трансляції, наприклад, кінчик зростаючого аксона або синапс. [8].

Теломери захищають хромосомну ДНК від деградації і запобігають різним формам абераційної рекомбінації. Довжина теломерів корелює зі здатністю клітин до великої кількості клітинних поділів, адже зазвичай теломерна ДНК при кожному поділі скорочується на 50-200 нуклеотидів, що

може контролювати проліферативну здатність нормальних соматичних клітин. Проте цього не відбувається в пухлинних клітинах через високу активність теломерази. Теломераза – це фермент, який відповідає за підтримку довжини теломер і синтезує втрачені теломерні послідовності шляхом додавання теломерних повторів. Звичайні соматичні клітини позбавлені активності теломерази, тому високий рівень її експресії безпосередньо пов'язаний з пухлинними клітинами, а пригнічення активності цього ферменту викликає старіння ракових клітин з їх подальшою загибеллю [9]. G-квадруплекси ДНК було знайдено в теломерній області хромосом, оскільки саме вона містить багаті на гуанін послідовності ДНК. G-квадруплекси відіграють важливу роль у регуляції довжини теломер через регуляцію активності теломерази [10].

Окрім людини, G-квадруплекси РНК також існують у деяких бактеріях і вірусах, включаючи вірус імунодефіциту людини (ВІЛ), вірус простого герпесу, коронавірус SARS-CoV-2, вірус папіломи людини, вірус Епштейна-Барра і вірус гепатиту С [11].

## **1.2. Коронавірус SARS-CoV-2**

SARS-CoV-2 – це надзвичайно заразний респіраторний вірус, який викликає у дорослих атипову пневмонію COVID-19 із важким гострим респіраторним синдромом. Окрім легень, цей вірус вражає серцево-судинну, нервову системи, а також печінку і нирки [11].

SARS-CoV-2 має одноланцюговий геном позитивної РНК (+РНК) з 29 882, 29 891 або 29 903 нуклеотидів і демонструє значний генетичний зсув від різних ізолятів. Після проникнення у сприятливі клітини, геном вірусу безпосередньо функціонує як мРНК для трансляції двох поліпротеїнів з регіону ORF1a та ORF1b, які розщеплюються двома вірусними протеазами на шістнадцять не структурних білків (nsp1-16), щоб ініціювати реплікацію та транскрипцію вірусного геному [12].

### 1.2.1. Структура геному коронавірусу SARS-CoV-2

Геном упакований білками вірусного нуклеокапсиду (N) у вигляді великого рибонуклеопротеїдного комплексу і оточений мембраною з ліпідами і вірусними білками S, M та E. Геном цього вірусу демонструє значне генетичне різноманіття з моменту свого відкриття і продемонстрував понад 7000 унікальних нуклеотидних мутацій. Геном SARS-CoV-2 нестабільний при підвищеній температурі через високозбагачений вміст аденіну+урацилу (62%) і низький вміст гуаніну+цитозину (38%) [12]. Геном SARS-CoV-2 має структуру mG-cap, mGpppA1, на 5'-кінці геному та хвіст полі-А завдовжки ~30–60 нуклеотидів (середнє значення 47 нуклеотидів) на 3'-кінці для стабільності вірусного геному та запобігання клітинному розщепленню екзорибонуклеазою [13]. 5'-нетрансльована область (UTR) геному SARS-CoV-2 має довжину 265 нуклеотидів і містить послідовність регулятора транскрипції та кілька інших цис-елементів для регулювання трансляції вірусу, синтезу субгенома та упаковки вірусного геному [14].

3' UTR геному SARS-CoV-2 має довжину 337 нуклеотидів. Вірусна 3' UTR містить сайт зв'язування комплексу реплікації та транскрипції (RTC), важливого для ініціації реплікації та транскрипції проміжної негативної-РНК (-РНК) [13].

Геном SARS-CoV-2 кодує 16 неструктурних, 4 структурних і 6 додаткових білків. Усі 16 неструктурних білків, що беруть участь у транскрипції, реплікації та ухиленні від імунітету вірусної РНК, є продуктами розщеплення двох поліпротеїнів, кодованих ORF1a та ORF1b, які разом займають приблизно 70% вірусного геному з 5'-кінця. Структурні білки S, E, M і N для утворення віріонів і додаткові білки (3a, 6, 7a, 7b, 8 і 9b) з невідомою функцією кодуються разом рештою 30% вірусного геному на 3'-кінці (рис. 2) [12].



під час збирання віріону через його взаємодію з вірусним геномом і мембранним білком (M). Окрім цього, білок N має вирішальне значення для підвищення ефективності транскрипції субгеномної вірусної РНК, а також реплікації вірусу. Специфічні для G4 сполуки, такі як PDP можуть стабілізувати цей G-квадруплекс і значно знизити рівень білка SARS-CoV-2 (N) шляхом інгібування його трансляції як *in vitro* так й *in vivo*. Це відкриття також дає нові знання про розробку нових противірусних препаратів проти COVID-19 [11].

### **1.3. Вивчення молекулярної динаміки за допомогою програмного забезпечення AMBER і її візуалізація за допомогою VMD**

Молекулярна динаміка дозволяє вивчати біологічні та хімічні системи на рівні атомів в часових масштабах від фемтосекунд до мілісекунд. Існують численні пакети програмного забезпечення для проведення симуляцій молекулярної динаміки, серед яких найпопулярнішими є AMBER та GROMACS.

AMBER стосується не лише пакета молекулярної динаміки, а включає в себе набір численних програм, які працюють разом для налаштування, виконання аналізу і моделювання молекулярної динаміки, від підготовки необхідних вхідних файлів до аналізу результатів. Також назва AMBER використовується до серії класичних силових полів молекулярної механіки, призначених для моделювання біологічних молекул [17]. Однак код і силове поле є окремими: кілька інших комп'ютерних пакетів реалізують силові поля AMBER, а також інші силові поля можуть бути реалізовані за допомогою пакету AMBER [18].

Набір програмного забезпечення AMBER розділений на дві частини: AmberTools, який складається з кількох незалежно розроблених пакетів, що добре працюють як окремо, так і з Amber здебільшого під ліцензією GPL, і Amber, яка базується на AmberTools і зосереджена навколо програми моделювання *pmemd*. При поєднанні AmberTools і Amber можна виконувати

багато стимуляційних завдань, а при їхньому поєднанні можна виконувати більш широке моделювання.

Частиною AmberTools є базова програма для мінімізації енергії та молекулярної динаміки – sander. Вона релаксує структуру, ітеративно переміщаючи атоми вздовж градієнту енергії до отримання достатньо низького середнього градієнту. Молекулярна динаміка генерує конфігурації системи, інтегруючи рівняння руху Ньютона і буде досліджувати більше конфігураційного простору, ніж мінімізація, та дозволить структурі перетнути невеликі бар'єри потенційної енергії. За допомогою модуля sander можна виконати більш складні пошуки конформацій і моделювання динаміки. Як зазначалось вище, pmemd (Particle Mesh Ewald Molecular Dynamics) є частиною Amber і це оптимізована для швидкості та паралельного масштабування версія sander. Версія pmemd.cuda працює на графічних процесорах [19].

VMD (Visual Molecular Dynamics) – це програма молекулярної графіки, яка була розроблена для інтерактивного графічного відображення молекулярних систем, зокрема біополімерів, таких як білки та нуклеїнові кислоти. VMD було розроблено під егідою головного дослідника Клауса Шультена в групі теоретичної та обчислювальної біофізики в Інституті передових наук і технологій Бекмана Університету Іллінойсу в Урбана-Шампейн і початкова версія програми була випущена у 1995 році.

Це програмне забезпечення надає повний графічний інтерфейс користувача для керування програмою, а також текстовий інтерфейс Tcl із використанням вбудованого синтаксичного аналізатора для створення складних скриптів із підстановкою змінних, циклами керування та викликами функцій. VMD був також розроблений спеціально із можливістю анімації траєкторій симуляції молекулярної динаміки. Зокрема він містить функції для аналізу молекулярних структур і їх траєкторій і обчислення середньоквадратичного відхилення або кореляційних функцій траєкторії динаміки [20].

## РОЗДІЛ 2 МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕННЯ

### 2.1. Підготовка молекул

В роботі було розглянуто молекулярну динаміку G-квадруплексів з такою послідовністю нуклеотидів:

Структура №3467: GGAGGAGGUGUUGCAGG

Структура №28903: GGCUGGCAAUGGCGG

Як вже було згадано в огляді літератури, структури G-квадруплексів стабілізуватимуть іони калію. До кожного з них за допомогою програмного забезпечення Avogadro було додано іони калію до G-квадруплексів. Після додавання калію, структури мали наступний вигляд:

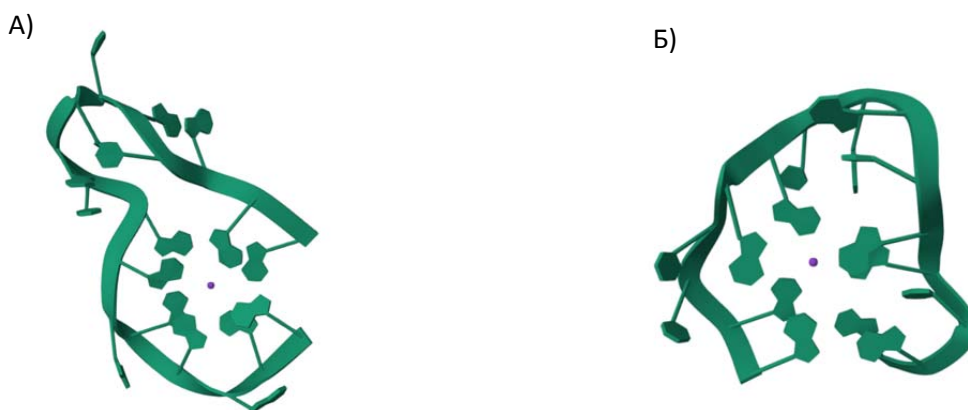


Рис.3 Структури досліджуваних G-квадруплексів. А) Структура №3467, Б) Структура №28903

### 2.2. Підготовка вихідних файлів у CHARMM-GUI за допомогою «Конструктору розчинів»

CHARMM-GUI – веб-платформа із зручним інтерфейсом для користувача, яка готує складні біомолекулярні системи до молекулярного моделювання. Ця платформа створює вхідні файли для ряду важливих програм, таких як CHARMM, NAMD, GROMACS, AMBER, GENESIS тощо.

На платформі CHARMM-GUI є розділ «Конструктор розчинів», який допомагає у створенні низки вхідних файлів для моделювання молекулярної динаміки молекул у середовищах водного розчинника. Дуже зручно, що іони можна додати у вікно моделювання, вказавши іони та їх концентрацію, а їх кількість автоматично визначається на платформі.

За допомогою цього розділу, наступним чином було створено вхідні файли для програм AMBER у силовому полі AMBER. Після завантаження файлу G-квадруплексу у форматі pdb і обравши необхідні опції, такі як pH системи додаємо іони. Було обрано розчин KCl у концентрації 0,15M та NaCl з такою ж концентрацією.

**Add Ions:**

Include Ions

Ion Placing Method:

Basic Ion Types

More Ion Types

Formula	Cation	Anion	Concentration	Neutralizing
KCl	K <sup>+</sup>	Cl <sup>-</sup>	0.15	<input type="radio"/>

Ion	Count
K <sup>+</sup>	37
Cl <sup>-</sup>	22

**Add Ions:**

Include Ions

Ion Placing Method:

Basic Ion Types

More Ion Types

Formula	Cation	Anion	Concentration	Neutralizing
NaCl	Na <sup>+</sup>	Cl <sup>-</sup>	0.15	<input checked="" type="radio"/>

Ion	Count
Na <sup>+</sup>	37
Cl <sup>-</sup>	22

Рис. 4 Приклад додавання іонів у вкладці «Конструктор розчинів» на веб-платформі CHARMM-GUI

Після скачування архіву з файлами, потрібно налаштувати відповідні параметри, а саме:

- у файлі «README» змінити `rtemd` на `rtemd.cuda`, адже другий показує більшу продуктивність для великих і складних симуляцій;
- у файлі `production.mdin` задати параметр `nstlim`. Він визначає загальну кількість кроків стимуляції молекулярної динаміки, які будуть виконані. Було обрано 5 000 000 кроків із `dt` (тривалість одного кроку симуляції) = 0,002 пс.;
- також у файлі `production.mdin` задати параметр `ntrj`. Він визначає частоту запису енергетичних даних у файл виводу. Було обрано 50 000 кроків;

- у файлі production.mdin задати параметр temp0. Він задає цільову температуру системи. Для G-квадруплексів було обрано дві температури – 293K (20 °C) і 309,6K (36,6 °C).

Завдяки цьому кроку маємо усі необхідні вхідні файли з правильними параметрами для проведення симуляції молекулярної динаміки.

### **2.3. Проведення молекулярної динаміки й інтерпретація результатів**

Молекулярна динаміка проводилась за допомогою пакета AMBER, який зосереджений на rmemd.cuda. Інтерпретація результатів проводилась за допомогою програми VMD на операційній системі UBUNTU. У VMD було завантажено файл топології з розширенням parm7 та траєкторій з розширенням .nc. За допомогою вкладки Extensions було проведено аналіз RMSD (Root Mean Square Deviation) за допомогою RMSD Trajectory Tool. Він виводить на екран середнє значення, значення стандартного відхилення, мінімальне і максимальне значення RMSD і записує значення на кожному відрізьку часу у файл з розширенням .dat, який потім використовувався для побудови графіків.

## РОЗДІЛ 3 РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ

### 3.1. Оцінка стійкості G-квадруплексу 3467 в розчинах NaCl і KCl при 20°C та 36,6°C.

RMSD – це корінь середньоквадратичного відхилення атомів або залишків молекул від їхньої початкової позиції протягом певного часу. Цей параметр вказуватиме на стабільність молекули протягом симуляції молекулярної динаміки.

Таблиця 1 Значення RMSD для всіх траєкторій G-квадруплексу 3467 в розчинах NaCl і KCl при 36,6°C та 20°C.

Структура	Середнє значення (Å)	Стандартне відхилення (Å)	Min. (Å)	Max. (Å)
3467 в NaCl при 36,6°C	33,818	8,08	2,549	53,440
3467 в NaCl при 20°C	30,801	9,29	1,532	53,183
3467 в KCl при 36,6°C	30,181	10,433	3,436	47,820
3467 в KCl при 20°C	30.487	6.538	3.813	45.522

Низькі значення RMSD будуть вказувати на стабільність молекули, тоді як високі значення – на значні зміни у структурі. Після початкової фази нормально мати високі значення, адже це фаза адаптації молекули, проте, якщо після цієї фази значення цього параметру постійно змінюється – це вказуватиме на нестабільність структури, або значні зміни в ній.

Стандартне відхилення RMSD показуватиме варіабельність значень протягом симуляції. Низьке значення цього параметру вказує на те, що молекула має стабільну структуру з невеликими коливаннями навколо середнього значення.

З отриманих результатів було побудовано графіки значення RMSD для всіх отриманих траєкторій, який показуватиме як структура змінюється протягом симуляції відносно початкової структури.

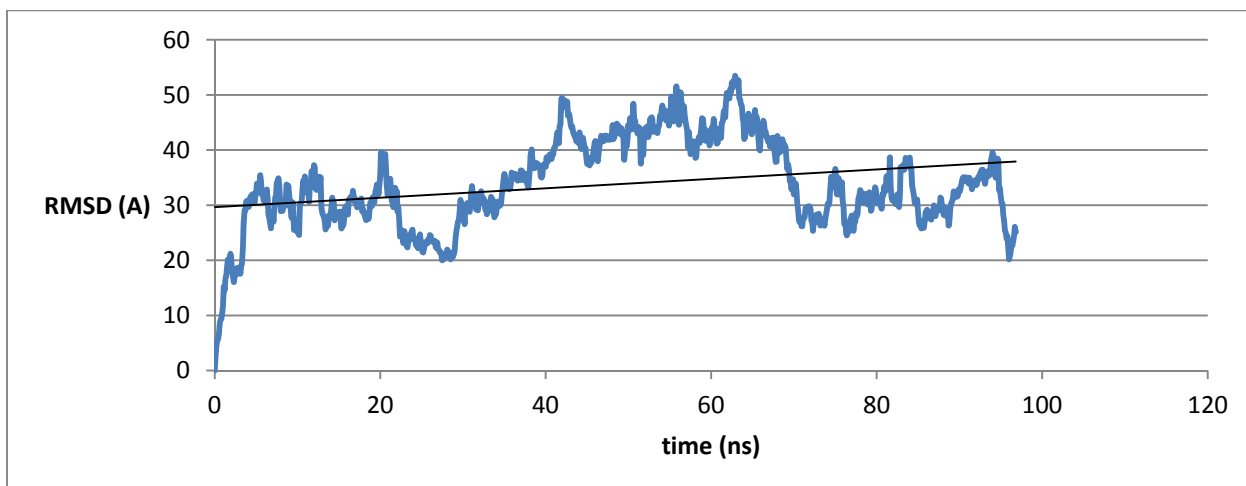


Рис. 5 Графік RMSD для всіх траекторій структури 3467 в розчині NaCl при температурі 36,6°C

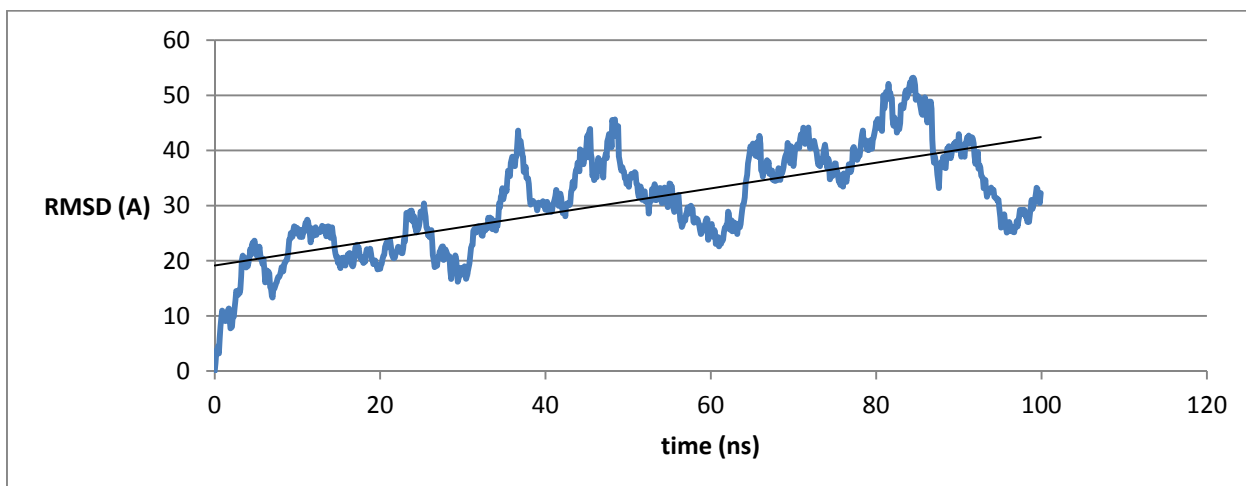


Рис. 6 Графік RMSD для всіх траекторій структури 3467 в розчині NaCl при температурі 20°C

Бачимо, що структура 3467 в розчині NaCl при температурі 36,6°C демонструє на графіку менші флуктуації (часті підйоми і спади), що свідчитиме про те, що G-квадруплекс буде більш стабільним саме за цих умов. Також це буде підтверджувати таблиця 1, з якої видно, що ця молекула буде мати нижче значення стандартного відхилення. Порівняння початкового і останнього положення G-квадруплексу 3467 в розчині NaCl при температурі 36,6°C зображено на рис. 7.

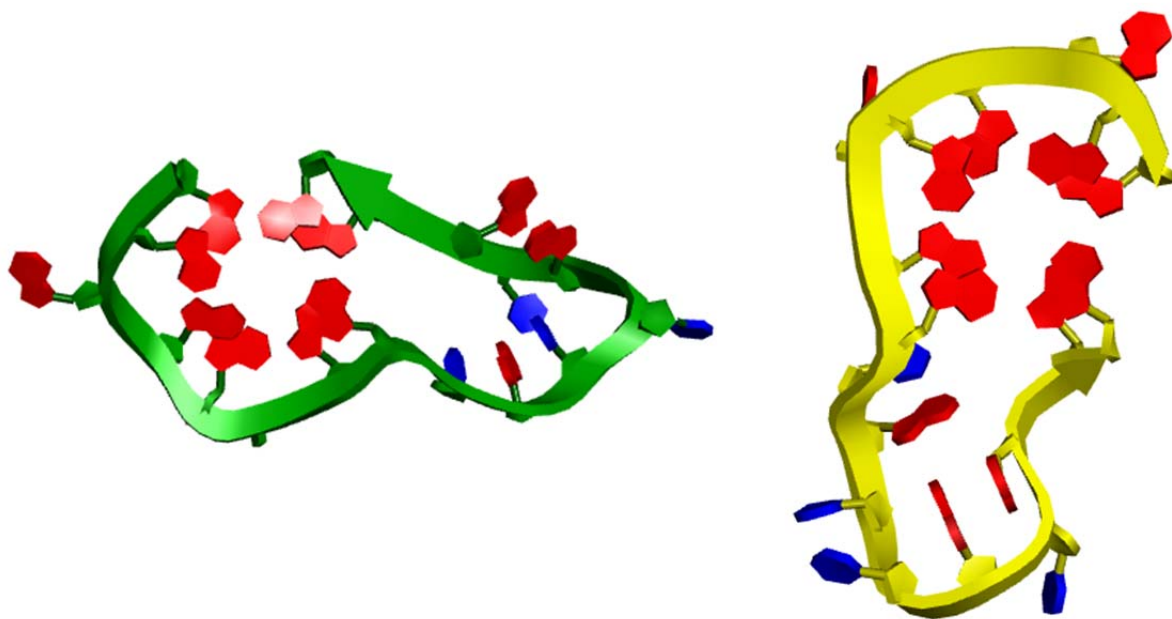


Рис. 7 Початкове (позначено зеленим ланцюгом) і останнє (позначено жовтим ланцюгом) положення G-квадруплексу 34 67 в розчині NaCl при температурі 36,6°C. Червоним кольором позначено пуринові основи.

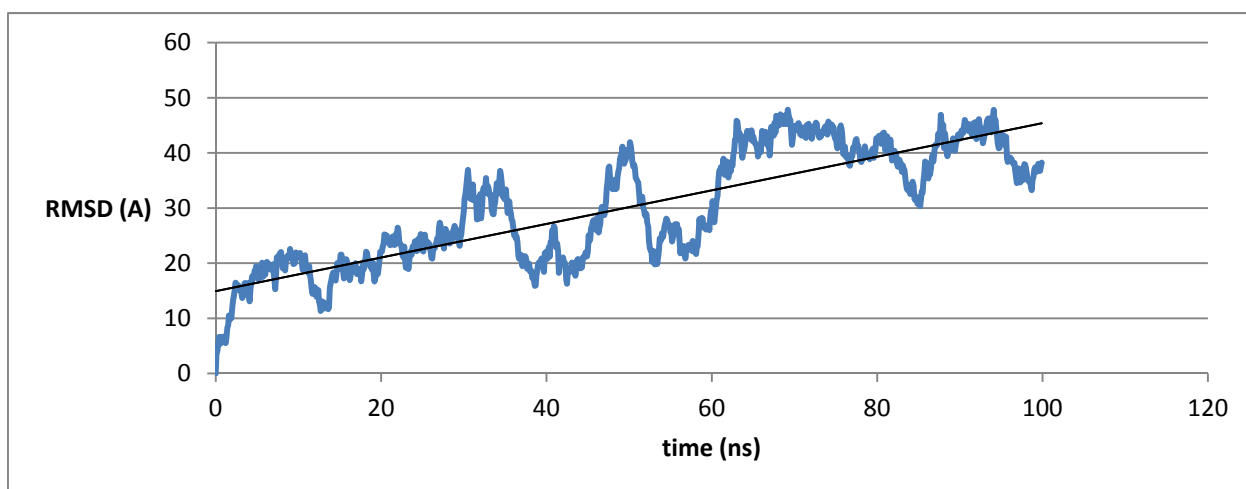


Рис. 8 Графік RMSD для всіх траєкторій структури 3467 в розчині KCl при температурі 36,6°C

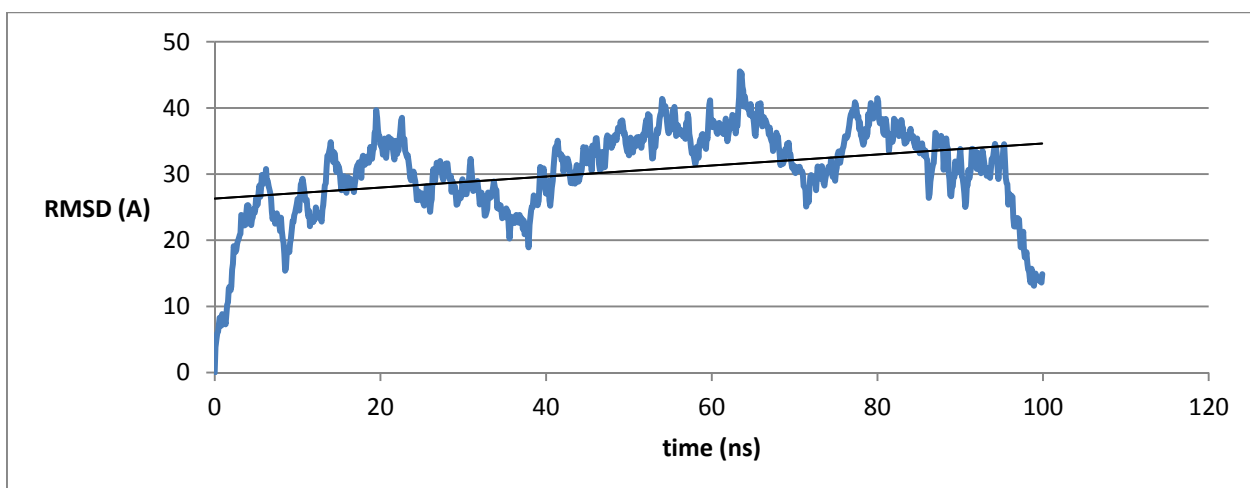


Рис. 9 Графік RMSD для всіх траєкторій структури 3467 в розчині KCl при температурі 20°C

На відміну від попереднього випадку, структура 3467 в розчині KCl буде більш стабільною за температури 20°C, так як демонструє менші флуктуації та нижче значення стандартного відхилення.

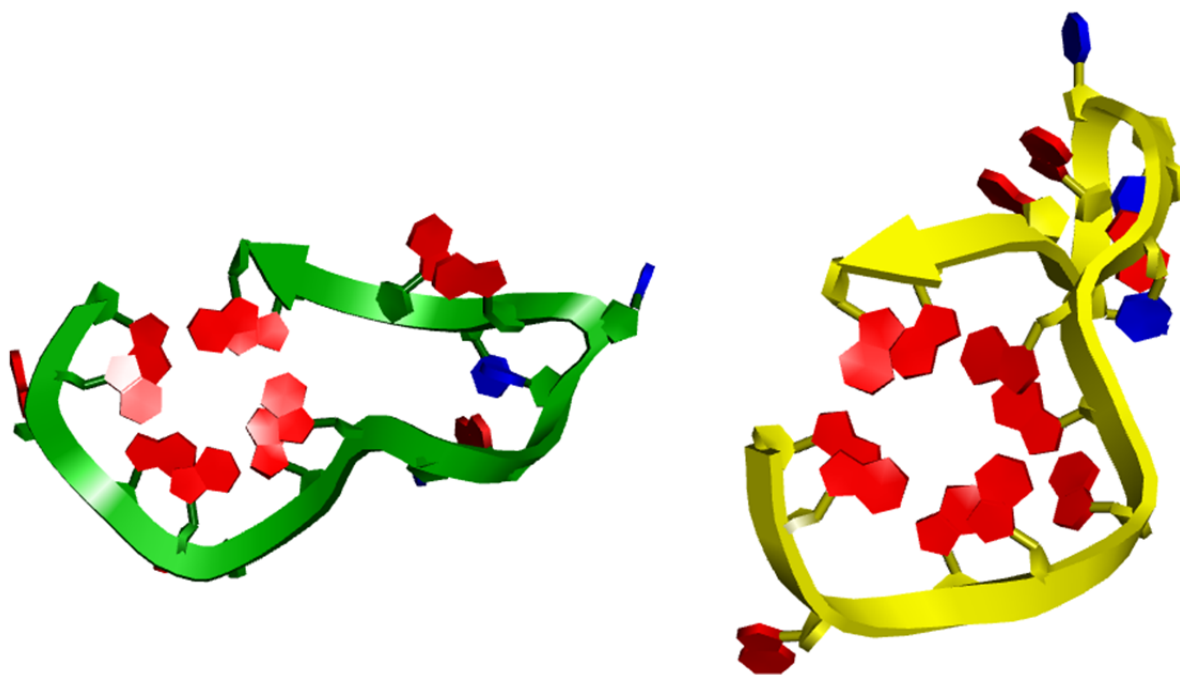


Рис. 10 Початкове (позначено зеленим ланцюгом) і останнє (позначено жовтим ланцюгом) положення G-квадруплексу 3467 в розчині KCl при температурі 20°C. Червоним кольором позначено пуринові основи.

Хоча відмінності між середнім значенням, стандартним відхиленням, мінімумом і максимумом параметру RMSD є не сильно великими для G-квадруплексу 3467, як в розчинні NaCl при температурі 20°C і 36,6°C так і в розчині KCl при таких же температурах, але завдяки графікам RMSD можна зробити висновок про стабільність молекули.

### 3.2. Оцінка стійкості G-квадруплексу 28903 в розчинах NaCl і KCl при 20°C та 36,6°C.

Таблиця 2 Значення RMSD для всіх траєкторій G-квадруплексу 28903 в розчинах NaCl і KCl при 36,6°C та 20°C.

Структура	Середнє значення (А)	Стандартне відхилення (А)	Min. (А)	Max. (А)
28903 в NaCl при 36,6°C	25,274	7,115	4,121	37,940
28903 в NaCl при 20°C	27,591	5,377	4,538	37,554
28903 в KCl при 36,6°C	26,905	5,292	3,734	40,719
28903 в KCl при 20°C	27,892	6,064	3,668	44,281

З таблиці 2 видно, що G-квадруплекс 28903 в порівнянні з структурою 3467 має нижчі середні і максимальні значення RMSD та значення стандартного відхилення. Це вказуватиме на те, що молекула 28903 загалом буде стабільнішою за 3467.

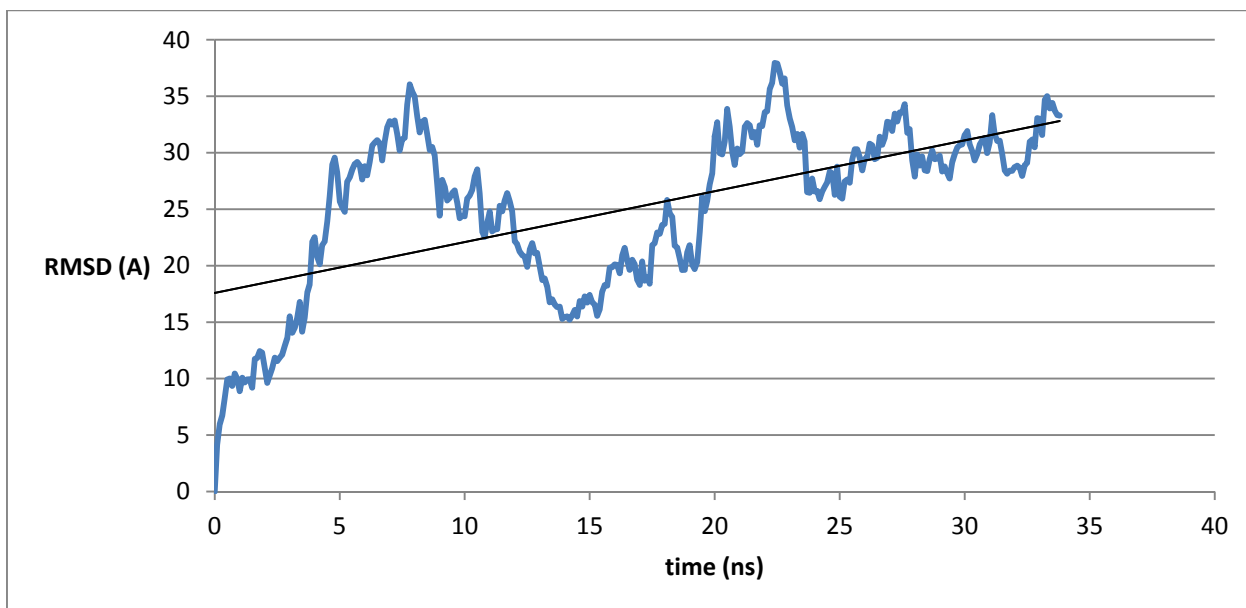


Рис. 11 Графік RMSD для всіх траєкторій структури 3467 в розчині NaCl при температурі 36,6°C

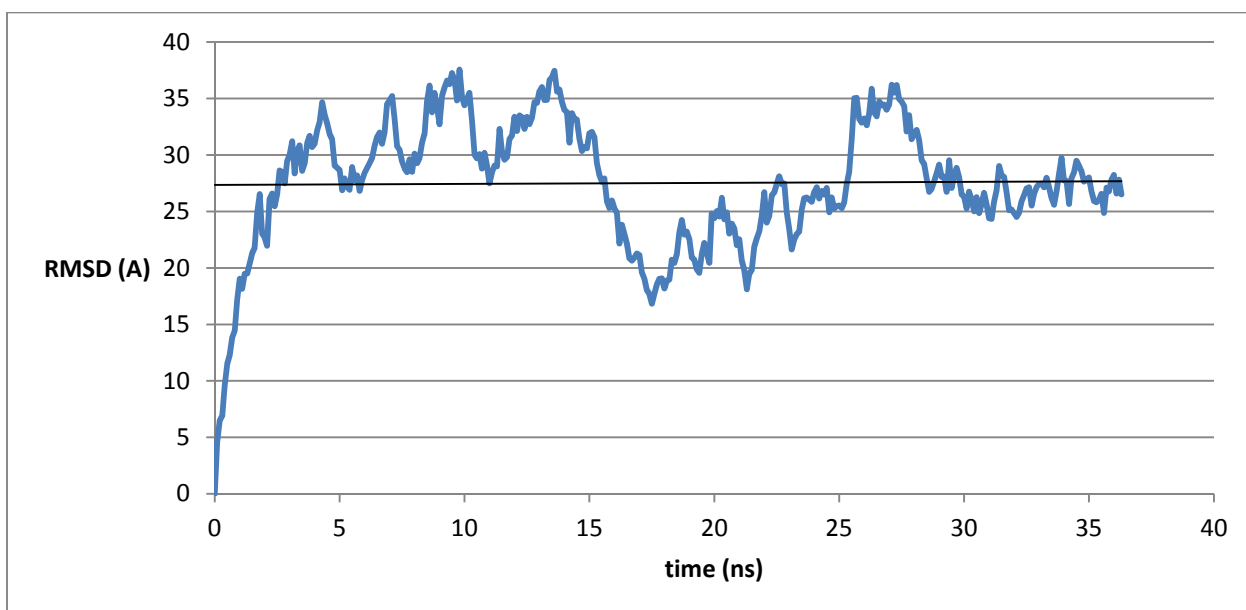


Рис. 12 Графік RMSD для всіх траєкторій структури 28903 в розчині NaCl при температурі 20°C

Як видно з графіків на рис. 11 та рис. 12 структура 28903 в розчині NaCl при температурі 20°C демонструє на графіку менші флуктуації (часті підйоми і спади), і можна зробити висновок, що молекула буде більш стабільною саме за цих умов. Значення стандартного відхилення з таблиці 2 також буде це підтверджувати.

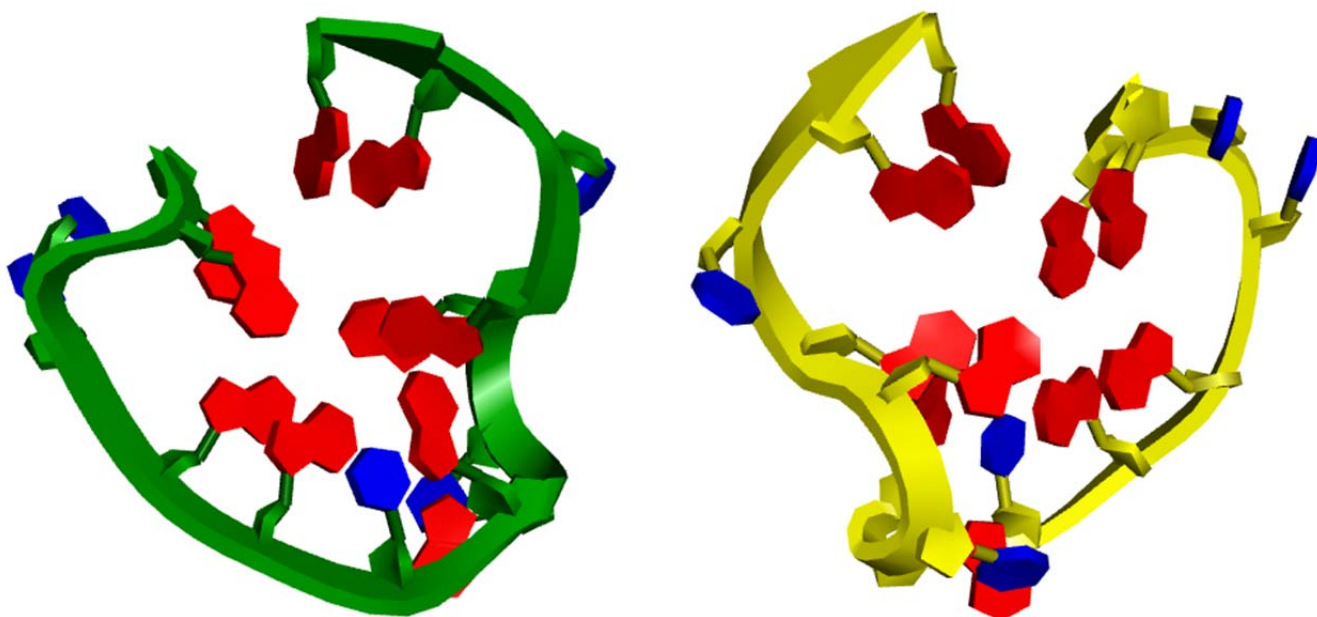


Рис. 13 Початкове ( позначено зеленим ланцюгом) і останнє (позначено жовтим ланцюгом) положення G-квадруплексу 28903 в розчині NaCl при температурі 20°C. Червоним кольором позначено пуринові основи.

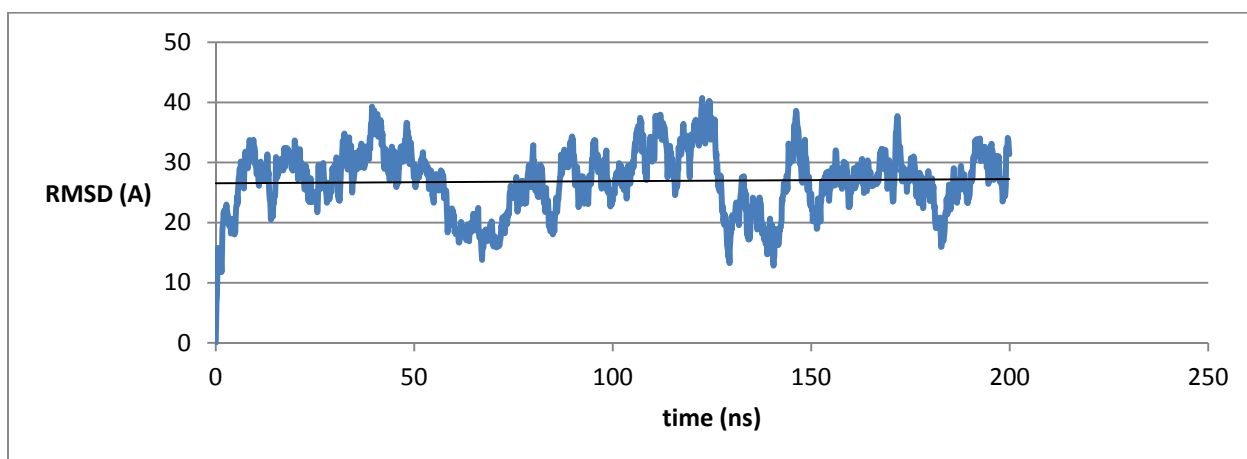


Рис. 14 Графік RMSD для всіх траєкторій структури 28903 в розчині KCl при температурі 36,6°C

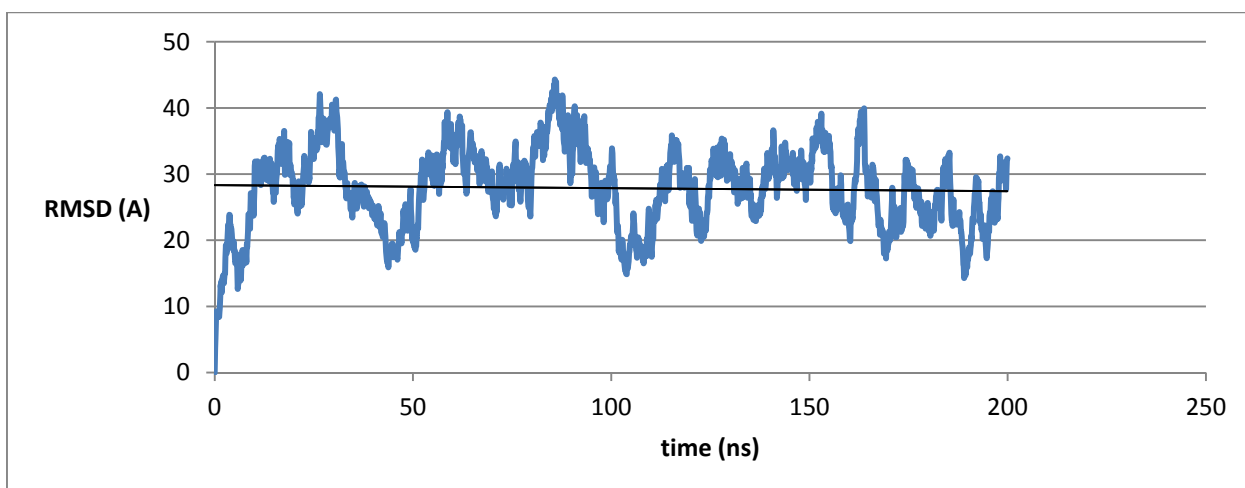


Рис. 15 Графік RMSD для всіх траєкторій структури 28903 в розчині KCl при температурі 20°C

Структура G-квадруплексу 28903 в розчині KCl при температурі 36,6°C хоч і показує різку флуктуацію RMSD у проміжках 55,5-80 нс. та 129,4-153 нс., проте все ж має трішки кращі значення стандартного відхилення і середнє значення, що зазначається в таблиці 2. Це робитиме її стабільнішою.

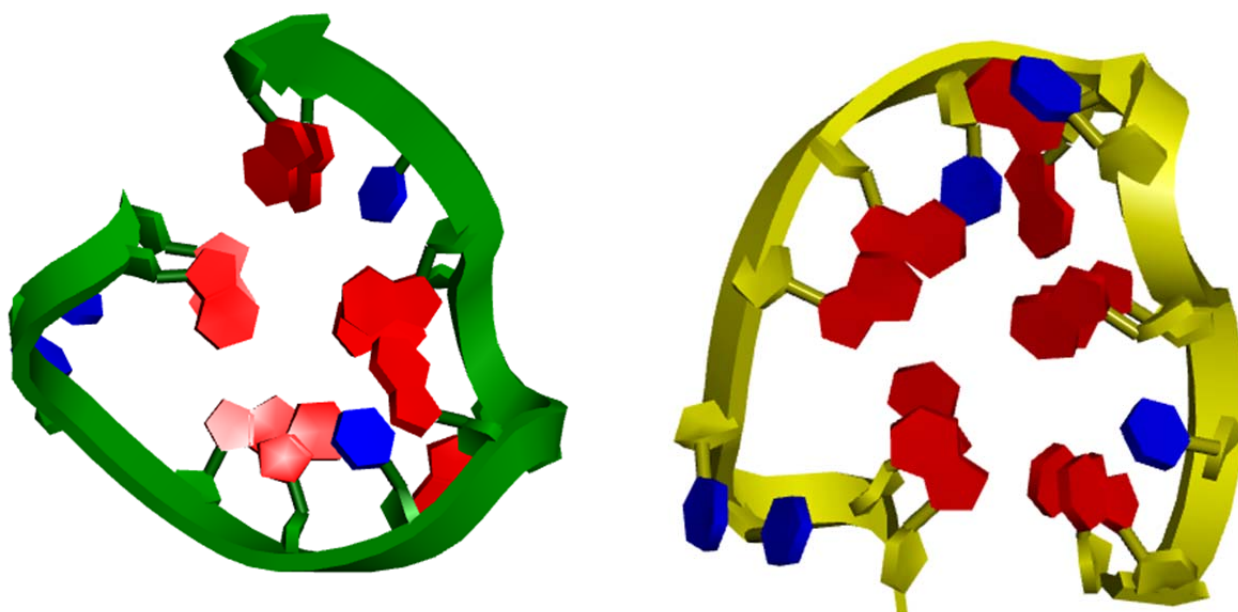


Рис. 16 Початкове ( позначено зеленим ланцюгом) і останнє (позначено жовтим ланцюгом) положення G-квадруплексу 28903 в розчині KCl при температурі 36,6°C. Червоним кольором позначено пуринові основи.

Цікавим є те, що G-квадруплекс 28903, на відміну від структури 3467, буде стабільнішим в розчинні NaCl при температурі 20°C, а в розчинні KCl при температурі 36,6°C.

## ВИСНОВКИ

1. За допомогою методу молекулярної динаміки проведено оцінку стійкості G-квадруплексів в геномі корона вірусу SARS-CoV-2 в розчинах NaCl і KCl при температурах 20°C і 36,6°C, побудовано і проаналізовано графіки RMSD для кожного випадку.
2. G-квадруплекси 3467 стабільніші в розчині NaCl при температурі 36,6°C, а в розчині KCl – при 20°C. G-квадруплекси 28903, навпаки, стабільніші в розчині NaCl при температурі 20°C, а в розчині KCl – при 36,6°C.
3. Встановлено, що найбільш стійкими є G-квадруплекси 28903, так як вони показують нижчі значення параметру RMSD в порівнянні з G-квадруплексами 3467.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Matsumoto, S., Tateishi-Karimata, H., Takahashi, S., Ohyama, T., & Sugimoto, N. (2020). «Effect of molecular crowding on the stability of RNA G-quadruplexes with various numbers of quartets and lengths of loops». *Biochemistry*.
2. Jiande Gu a,b, Jerzy Leszczynski, Manju Bansal «A new insight into the structure and stability of Hoogsteen hydrogen-bonded G-tetrad: an ab initio SCF study»
3. Exploiting hydrogen bonding interactions to probe smaller linear and cyclic diamines binding to G-quadruplexes: a DFT and molecular dynamics study
4. N. H. Campbell and S. Neidle, «G-Quadruplexes and Metal Ions», Springer, 2012, ch. 4, pp. 119–134
5. S. Neidle, «Therapeutic Applications of Quadruplex Nucleic Acids» . Academic Press, San Diego, USA, 2011.
6. Neidle S. «Structured waters mediate small molecule binding to G-quadruplex nucleic acids». *Pharmaceuticals*. 2022. V.15. P. 7.
7. Pandey, S., Agarwala, P., and Maiti, S. (2013) «Effect of loops and G-quartets on the stability of RNA G-quadruplexes», *J. Phys. Chem. B* 117, 6896-6905.
8. Fay, M. M., Lyons, S. M., & Ivanov, P. (2017). «RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms». *Journal of Molecular Biology*, 429(14), 2127–2147. doi:10.1016/j.jmb.2017.05.017
9. V. V. Negrutka, L. V. Dubey, M. M. Ilchenko, I. Ya. Dubey «Design and study of telomerase inhibitors based on G-quadruplex ligands» *Biopolymers and Cell*. 2013. Vol. 29. N 3. P. 169–176 doi: 10.7124/bc.000817
10. Zahler, A. M., Williamson, J. R., Cech, T. R., & Prescott, D. M. (1991). «Inhibition of telomerase by G-quartet DMA structures». *Nature*, 350(6320), 718–720. doi:10.1038/350718a0

11. Zhao, C., Qin, G., Niu, J., Wang, Z., Wang, C., Ren, J., & Qu, X. (2020). «Targeting RNA G-Quadruplex in SARS-CoV-2: A Promising Therapeutic Target for COVID-19?» *Angewandte Chemie International Edition*, 60(1), 432–438. doi:10.1002/anie.202011419
12. Brant, A. C., Tian, W., Majerciak, V., Yang, W., & Zheng, Z.-M. (2021). «SARS-CoV-2: from its discovery to genome structure, transcription, and replication». *Cell & Bioscience*, 11(1). doi:10.1186/s13578-021-00643-z
13. Viswanathan T, Arya S, Chan SH, Qi S, Dai N, Misra A, Park JG, Oladunni F, Kovalskyy D, Hromas RA, et al. «Structural basis of RNA cap modification by SARS-CoV-2». *Nat Commun*. 2020;11(1):3718.
14. Miao Z, Tidu A, Eriani G, Martin F. «Secondary structure of the SARSCoV-2 5'-UTR. RNA» *Biol*. 2021;18(4):447–56
15. Sanjib K. Mukherjee , Jim-Marcel Knop, and Roland Winter «Modulation of the Conformational Space of SARS-CoV-2 RNA Quadruplex RG-1 by Cellular Components and the Amyloidogenic Peptides  $\alpha$ -Synuclein and hIAPP»
16. Perrone, R., Butovskaya, E., Daelemans, D., Palù, G., Pannecouque, C. and Richter, S.N. (2014) «Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19». *J. Antimicrob. Chemother.*, 2014, 69, 3248-3258.
17. Salomon-Ferrer, R., Case, D. A., & Walker, R. C. (2012). «An overview of the Amber biomolecular simulation package». *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(2), 198–210. doi:10.1002/wcms.1121
18. D. A. Case; T. Cheatham; T. Darden; H. Gohlke; R. Luo; K. M. Merz, Jr.; A. Onufriev; C. Simmerling; B. Wang; R. Woods. «The Amber biomolecular simulation programs». *J. Computat. Chem.*, 2005, 26, 1668–1688.
19. D.A. Case, H.M. Aktulga, K. Belfon et al. «Amber Reference Manual», 2021

20. Humphrey, W., Dalke, A., & Schulten, K. (1996). «VMD: Visual molecular dynamics». *Journal of Molecular Graphics*, 14(1), 33–38. doi:10.1016/0263-7855(96)00018-5