

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ

Інститут мікробіології і вірусології ім. Д.К. Заболотного НАН України

Кваліфікаційна наукова
праця на правах рукопису

Заремба Андрій Анатолійович

УДК: 578:577:615.1

ДИСЕРТАЦІЯ

**СТРУКТУРО ОРІЄНТОВАНА РОЗРОБКА І ПОШУК ТАРГЕТНИХ
АНТИВІРУСНИХ СПОЛУК ПРОТИ ВІРУСУ ЕПШТЕЙНА-БАРР ТА
SARS-COV-2**

091 – Біологія та біохімія

09 – Біологія

Подається на здобуття наукового ступеня доктора філософії

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

_____ Заремба А.А.

Науковий керівник: Загородня Світлана Дмитрівна, кандидат біологічних наук,
старший дослідник



Київ – 2026

АНОТАЦІЯ

Заремба А.А. Структуро орієнтована розробка і пошук таргетних антивірусних сполук проти вірусу Епштейна-Барр та SARS-CoV-2. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття ступеня доктора філософії за спеціальністю 091 «Біологія та біохімія». – Інститут мікробіології і вірусології ім. Д.К. Заболотного НАН України, Київ, 2026 р.

Дана робота є присвяченою пошуку і розробці ліків проти двох дуже різних у багатьох аспектах, однак подібних у значному впливі на людство, вірусів – вірусу Епштейна-Барр (ВЕБ) і SARS-CoV-2. Перший характеризується надзвичайним поширенням серед населення планети, складним життєвим циклом та здатністю стимулювати неопластичне перетворення уражених клітин. Другий є емерджентним патогеном, який попри скасування статусу пандемічного, знаходиться під постійним наглядом через високу швидкість еволюційних змін та пов'язаних з цим хвиль збільшення захворюваності в осінньо-зимовий період. З різних причин та попри значні зусилля, хвороби спричинені обома цими вірусами досі здебільшого піддаються лише симптоматичному лікуванню.

В ролі методологічної основи дисертації був вибраний структуро орієнтований дизайн ліків (СОДЛ) як найбільш сучасний підхід, що зарекомендував себе компромісним між швидкістю, затратами ресурсів і надійністю. Безпосередньо застосованими були віртуальний скринінг на основі напівгнучкого молекулярного докінгу та симуляція молекулярної динаміки у різних варіаціях. Траєкторія симуляції молекулярної динаміки ліганд-рецепторного комплексу також використовувалась для обрахунку вільної енергії зв'язування методами MM/PBSA та MM/GBSA. Крім того, для аналізу еволютивної лабільності SARS-CoV-2 як емерджентного патогену був використаний порівняльний аналіз зміни варіантного складу вірусу у відповідь на вплив позитивного відбору.

Безпосередніми цілями для застосування СОДЛ були вибрані антиапоптоичний білок BHRF1 та вірусний генوم для вірусу Епштейна-Барр і рецепторзв'язуючий домен (RBD) S-глікопротеїну для SARS-CoV-2. Кожна із зазначених цільових макромолекул є ключовою для життєдіяльності відповідного патогену. Зокрема, BHRF1 є BCL-2-подібним фактором ВЕБ, який виконує функцію дерегуляції внутрішнього шляху запуску апоптозу, забезпечує відтермінування смерті ураженої клітини та відповідне збільшення продуктивності інфекції, однак лише під час літичної стадії циклу вірусу. За успішного встановлення латентності, єдиною дійсно важливою структурно-функціональною компонентою патогену залишається його ДНК в епісомальній формі, що і обумовило вибір її як другої цілі для протидії ВЕБ. У випадку SARS-CoV-2, S-глікопротеїн, зокрема його RBD, безпосередньо забезпечують розпізнавання цим коронавірусом чутливої клітини. Останнє відбувається за рахунок взаємодії з поверхневою формою ангіотензинперетворюючого ферменту людини (hACE2). Переривання цього процесу дозволить протидіяти SARS-CoV-2 вже на стадії первинної взаємодії з чутливою клітиною.

Шляхом застосування ітераційного підходу, який опирався на симуляцію молекулярної динаміки ліганд-рецепторного комплексу як джерело даних щодо спорідненості конкретного варіанту ліганду з ціллю, була створена молекулярна конструкція здатна до стабільного комплексоутворення з BHRF1. Ця структура була названа EBAI від **E**pstein-**B**arr **A**ntiapoptotic **I**nhibitor. Ступінь її спорідненості до цілі був оцінений у двох паралельних симуляційних експериментах, де комплекс зазнав націленої дестабілізації. Було показано, що попри повну втрату контакту з поверхнею BHRF1, EBAI здатен самостійно відновлювати своє положення в межах ВНЗ-зв'язуючої ділянки цього білка. При чому найшвидше відновлюваними взаємодіями були водневі та іонні зв'язки з R100 та N61. В даному контексті слід зазначити, що цей аргінін і його позитивний заряд вважається вирішальним для виконання BHRF1 своєї антиапоптоичної функції. Це, а також дані про

терапевтичний потенціал інгібування BHRF1, почерпнуті з літературних джерел, говорить про значні подальші перспективи розвитку ЕВАІ як засобу протидії гострій інфекції ВЕБ.

Відповідно до зазначених вище особливостей життєвого циклу вірусу Епштейна-Барр і використовуючи подібний до застосованого під час розробки ЕВАІ підхід, було створено два покоління ДНК-інтеркаляторів, потенційно здатних розпізнавати фрагмент ДНК-дуплексу довжиною в 16 пар нуклеотидів. Ці структури були названі HASDI від **H**igh-**A**ffinity **S**elective **D**N**A** **I**ntercalator і HASDI-G2 від **H**igh-**A**ffinity **S**elective **D**N**A** **I**ntercalator – **G**eneration **2**. Як HASDI, так і HASDI-G2 мають модульну структуру, де кожен окремий модуль складається з двох хромофорів (феназин для HASDI та індазол для HASDI-G2) об'єднаних лінкером. Кожен лінкер може бути модифікованим в чотирьох місцях і, таким чином, направленим на розпізнавання двох пар нуклеїнових основ. Водночас хромофори, завдяки своїй ароматичній природі, здатні інтеркалювати в ДНК-дуплекс з обох сторін від пари нуклеїнових основ, що розпізнається.

В двох порівняльних симуляційних експериментах було показано, що HASDI здатен взаємодіяти з послідовністю націлювання з набагато вищою спорідненістю і стабільністю порівняно із випадковим фрагментом геному тієї ж довжини. Послідовністю націлювання був фрагмент гену EBNA1 як єдиного важливого для підтримки геному ВЕБ фактору, а випадковою послідовністю – фрагмент людського гену KCNH2. Однак було помічено, що, попри значну дестабілізацію водневих взаємодій, у останньому випадку інтеркаляція феназинових кілець була непорушеною, що свідчило про надлишковий внесок ненаправлених взаємодій у комплексоутворення і спонукало до розробки HASDI-G2.

Порівняно з HASDI, HASDI-G2 у ідентичному експерименті, разом зі значною дестабілізацією направлених взаємодій, продемонстрував швидку деінтеркаляцію частини індазольних кілець із одночасним частковим плавленням ДНК-дуплексу. Це свідчить на користь значного зниження внеску неселективної компоненти в

комплексоутворення і потенційно кращу селективність HASDI-G2. З метою визначення меж потенціалу HASDI-G2 було проведено два додаткових порівняльних симуляційних дослідження. В першому ця структура була націлена на місце зрощення філадельфійської хромосоми – основної причини хронічного мієлоїдного лейкозу, але інтеркальована не лише в порушений ген BCR_ABL1, а і до нативних BCR та ABL1 в місці потенційного зрощення. Було показано, що, попри набагато більший рівень подібності між цими ділянками порівняно з EBNA1 і KCNH2, у нецільових комплексах спостерігалась масова втрата водневих зв'язків та локальні плавлення ДНК-дуплексу, при чому не лише в зонах безпосередньої невідповідності, а і в цілком відповідних ділянках. Фінальний експеримент передбачав націлювання HASDI-G2 на KRAS_G12S – мутований ген, який відрізняється від нативного лише одним нуклеотидом. Результат цього симуляційного дослідження був подібним: поряд із цілком стабільною взаємодією з ДНК-дуплексом послідовності націлювання, комплекс HASDI-G2 з нецільовою послідовністю характеризувався низьким рівнем стабільності водневих взаємодій та навіть локальним плавленням подвійної спіралі.

Таким чином, було показано, що розроблені молекулярні конструкції, і особливо HASDI-G2, є здатними до високої за своїм потенціалом селективності у розпізнаванні ДНК-дуплексу певної попередньо запрограмованої послідовності. Останнє відкриває двері перед можливістю протидії таким патогенам як вірус Епштейна-Барр на недоступному нині рівні, хоча і не обмежується лише цим.

Особливості молекулярної біології SARS-CoV-2 через його емерджентну природу є дослідженими набагато менше порівняно з ВЕБ. Тож з метою розуміння швидкості пристосувальних змін до факторів позитивного відбору, яким в майбутньому, зокрема, може виступати і потенційний інгібітор, було проведено комплексний аналіз зміни варіантного складу патогену під впливом масової вакцинації. Як приклад різних популяцій людини були вибрані три країни з різним ступенем соціоекономічного розвитку – Індія, Німеччина та Україна.

Було показано, що варіант Дельта поряд з тим, що був присутній у популяції кожної з країн до початку наростання популяційного імунітету, набув здатності витіснити інші варіанти лише після початку вакцинації. Це свідчить на користь не лише здатності SARS-CoV-2 до надзвичайно швидкого пристосування, а і говорить про високу гетерогенність популяції цього патогену.

Опираючись на отримані дані, було проведено дослідження з пошуку консервативної кишені розташованої в межах біологічно важливої поверхні рецепторзв'язуючого фактору SARS-CoV-2 і при цьому потенційно здатної до зв'язування малої хімічної сполуки. Шляхом аналізу симуляції молекулярної динаміки RBD варіантів Ухань, P.1 і Кластер 5, поряд з особливостями змін у їх взаємодії з hACE2, було помічено, що у центральній частині hACE2-зв'язуючої ділянки розташовується топологічно відносно консервативний покот потенційно здатний зв'язувати малу хімічну сполуку.

Шляхом застосування ітераційного підходу, який опирався на симуляцію молекулярної динаміки ліганд-рецепторного комплексу, була *de novo* розроблена мала молекула здатна до стабільної взаємодії з RBD в зоні відкритої кишені одразу чотирьох варіантів SARS-CoV-2: Ухань, Омикрон, Дельта та Кластер 5. Ця речовина утворювала зіставні за своєю стабільністю комплекси з кожним із перелічених варіантів, попри наявні у випадку варіантів Кластер 5 і Омикрон амінокислотні заміни безпосередньо в зоні покоту. Це, окрім того, що свідчить на користь високого потенціалу розробленої молекули як інгібітора взаємодії RBD/hACE2 і відповідного її фармацевтичного потенціалу, також говорить про подальші перспективи відкритої кишені як такої. В сфері СОДЛ і особливо в контексті таких варіабельних патогенів як віруси, рідкісним є явище хоча і обмежено, але стійкої до впливу мутацій кишені.

Розвиваючись в цьому напрямку, було проведено додаткове дослідження із застосування класичного віртуального скринінгу бібліотеки затверджених FDA (Food and Drug Administration) лікарських сполук для пошуку речовин, потенційно здатних до взаємодії з RBD у ділянці відкритого покоту. Після трьох послідовних

етапів оцінки і відсіву, зокрема із застосуванням симуляції молекулярної динамки різної тривалості, антиандроген кетодаролутамід був ідентифікований як сполука, здатна до стабільної взаємодії з RBD у його hACE2-зв'язуючій ділянці. Хоча потенціал цього лікарського засобу до безпосереднього лікування COVID-19 (coronavirus disease 2019) є обмеженим через потужний антиандрогенний ефект, він може бути використаний в ролі скефолду для розробки більш перспективних речовин. Не менш важливим результатом цього дослідження є також розкриття потенційного механізму впливу кетодаролутаміду на перебіг COVID-19 під час лікування андроген залежного раку передміхурової залози. Звіти про цей ефект були помічені під час підготовки цього рукопису.

Ключові слова: ВЕБ, SARS-CoV-2, BHRF1, RBD, варіант Омікрон, варіант Дельта, віруси, вакцинація, епідеміологія, розробка антивірусних ліків, віртуальний скринінг, молекулярна динаміка, молекулярний докінг, інгібування взаємодії вірус-господар, ДНК-інтеркалятори.

ABSTRACT

Zaremba A.A. Structure-based design and search for targeted antiviral compounds against Epstein-Barr virus and SARS-CoV-2. – Qualification scientific work in the form of a manuscript.

The thesis for the degree of Doctor of Philosophy in specialty 091 “Biology and Biochemistry”. – D.K. Zabolotny Institute of Microbiology and Virology of the NAS of Ukraine, Kyiv, 2026.

This work is devoted to the search and design of drugs against two viruses, namely Epstein-Barr virus (EBV) and SARS-CoV-2, that are very different in many aspects, but similar in their significant impact on humanity. The first one is characterized by an extraordinary spread among the world population, a complex life cycle and the ability to stimulate neoplastic transformation of infected cells. The second one is an emerging pathogen, which, despite the abolition of pandemic status, is under constant surveillance due to the high rate of evolutionary changes and the associated waves of increased morbidity in the autumn-winter period. For various reasons and despite significant efforts, the diseases caused by both of these viruses are still mostly treatable only symptomatically.

As the methodological basis of the thesis, structure-based drug design (SBDD) was chosen as the most modern approach, which has proven to be a compromise between speed, resource consumption and reliability. Virtual screening based on semi-flexible molecular docking and molecular dynamics simulation in various variations were used. The trajectory of the molecular dynamics simulation of the ligand-receptor complex was also used to calculate the free binding energy by the MM/PBSA and MM/GBSA methods. The comparative analysis of the change in the variant composition of the virus in response to the influence of positive selection was also used to analyze the evolutionary lability of

SARS-CoV-2 as an emerging pathogen.

The genome Epstein-Barr virus and its antiapoptotic protein BHRF1, as well as the receptor-binding domain (RBD) of the SARS-CoV-2 S-glycoprotein were selected as direct targets for the application of SBDD. Each of the specified target macromolecules is essential for the activity of the corresponding pathogen. In particular, BHRF1 is a BCL-2-like factor of EBV, which performs the function of deregulating the intrinsic pathway of apoptosis and provides a delay in the death of an infected cell and a corresponding increase in the productivity of the infection. However, it occurs only during the lytic stage of the virus cycle. If latency is successfully established, viral DNA in a form of episome remains as the only truly important structural and functional component of the pathogen, which is why it was chosen as the second target for combating EBV. In the case of SARS-CoV-2, the S-glycoprotein and in particular its RBD directly ensure the recognition of a sensitive cell by this coronavirus. The latter occurs due to interaction with the surface form of human angiotensin converting enzyme (hACE2). Interrupting this process will allow counteracting SARS-CoV-2 already at the stage of primary reception and entry into a cell.

By using an iterative approach, which was based on the molecular dynamics simulation of the ligand-receptor complex as a source of data on the affinity of a particular ligand variant for the target, a molecular structure capable of stable complexation with BHRF1 was created. This structure was named EBAI from **Epstein-Barr Antiapoptotic Inhibitor**. The degree of its affinity for the target was assessed in two parallel simulation experiments where the complex underwent targeted destabilization. It was shown that despite the complete loss of contact with the BHRF1 surface, EBAI is able to independently restore its position within the BH3-binding region of this protein. Moreover, the most rapidly restored interactions were hydrogen and ionic bonds with R100 and N61. In this context, it should be noted that this arginine and its positive charge are considered crucial for BHRF1 to perform its anti-apoptotic function. This, as well as

data on the therapeutic potential of BHRF1 inhibition obtained from literature sources, suggests significant further prospects for the development of EBAI as a means of counteracting acute EBV infection.

In accordance with the above-mentioned features of the EBV life cycle and using an approach similar to that used during the development of EBAI, two generations of DNA intercalators were created, potentially capable of recognizing a fragment of a DNA duplex of 16 nucleotide pairs in length. These structures were named HASDI from **H**igh-**A**ffinity **S**elective **D**N**A** **I**ntercalator and HASDI-G2 from **H**igh-**A**ffinity **S**elective **D**N**A** **I**ntercalator – **G**eneration **2**. Both HASDI and HASDI-G2 have a modular structure, where each individual module consists of two chromophores (phenazine for HASDI and indazole for HASDI-G2) connected by a linker. Each linker can be modified at four places and, thus, directed to the recognition of two pairs of nucleic bases. At the same time, the chromophores, due to their aromatic nature, are able to intercalate into the DNA duplex on both sides of the recognized nucleic base pair.

In two comparative simulation experiments, it was shown that HASDI is able to interact with a targeting sequence with much higher affinity and stability compared to a random genome fragment of the same length. The targeting sequence was a fragment of the EBNA1 gene as the only factor important for the maintenance of the EBV genome, and the random sequence was a fragment of the human KCNH2 gene. However, it was observed that despite significant destabilization of hydrogen bonds in the latter case, the intercalation of the phenazine rings was intact, which indicated an excessive contribution of non-directional interactions to complex formation and prompted the development of HASDI-G2.

Compared to HASDI, HASDI-G2 in an identical experiment, along with significant destabilization of directional interactions, demonstrated rapid deintercalation of part of the indazole rings with simultaneous partial melting of the DNA duplex. This suggests a

significant reduction in the contribution of the non-selective component to complex formation and potentially better selectivity of HASDI-G2. Two additional comparative simulation studies were performed to determine the limits of the HASDI-G2 potential. In the first, this structure was targeted to the fusion site of the Philadelphia chromosome, which is a major cause of chronic myeloid leukemia, but intercalated not only into the disrupted BCR_ABL1 gene, but also into native BCR and ABL1 at the potential fusion site. It was shown that, despite the much higher level of similarity between these regions compared to EBNA1 and KCNH2, massive loss of hydrogen bonds and local melting of the DNA duplex were observed in the non-targeted complexes, not only in areas of direct mismatch, but also in completely matched regions. The final experiment involved targeting HASDI-G2 to KRAS_G12S, a mutated gene that differs from the native one by only one nucleotide. The result of this simulation study was similar: along with a completely stable interaction with the DNA duplex of the targeting sequence, the HASDI-G2 complex with the non-target sequence was characterized by a low level of stability of hydrogen interactions and even local melting of the double helix.

Thus, it was shown that the developed structures and especially HASDI-G2 are capable of high potential selectivity in recognizing a DNA duplex of a certain pre-programmed sequence. The latter opens the door to the possibility of counteracting pathogens such as the Epstein-Barr virus at a currently unavailable level, although it is not limited to this.

The features of the molecular biology of SARS-CoV-2 due to its emergent nature have been studied much less compared to EBV. Therefore, in order to understand the speed of adaptive changes to positive selection factors, which in the future may also act as a potential inhibitor, a comprehensive analysis of changes in the variant composition of the pathogen under the influence of mass vaccination was conducted. As an example of different human populations, three countries with different levels of socio-economic

development were chosen - India, Germany and Ukraine.

It was shown that the Delta variant, in addition to being present in the population of each country before the start of the increase in population immunity, acquired the ability to displace other variants only after the start of vaccination. This indicates not only the ability of SARS-CoV-2 to adapt extremely quickly, but also the high heterogeneity of the population of this pathogen.

Taking into account the obtained data, a study was conducted to search for a conserved pocket located within the biologically important surface of the receptor-binding factor of SARS-CoV-2 and potentially capable of binding a small chemical compound. By analyzing the molecular dynamics simulation of the RBD of the Wuhan, P.1 and Cluster 5 variants, along with the features of changes in their interaction with hACE2, it was observed that in the central part of the hACE2-binding site there is a topologically relatively conservative pocket potentially capable of binding a small chemical compound.

By using an iterative approach based on molecular dynamics simulation of the ligand-receptor complex, a small molecule capable of stable interaction with the RBD in the discovered pocket area of four SARS-CoV-2 variants – Wuhan, Omicron, Delta and Cluster 5 – was *de novo* designed. This substance formed complexes with each of the listed variants comparable in stability despite the amino acid substitutions present in the case of the Cluster 5 and Omicron variants directly in the pocket area. This, in addition to testifying to the high potential of the developed molecule as an inhibitor of the RBD/hACE2 interaction and its corresponding pharmaceutical potential, also speaks about the further prospects of the discovered pocket as such. In the field of SBDD and especially in the context of such variable pathogens as viruses, the phenomenon of a pocket that is resistant to mutations, albeit to a limited extent, is rare.

In this direction, an additional study was conducted using a classical virtual screening of a library of FDA-approved drug compounds to search for substances with

potential interaction with RBD in the discovered pocket region. After three successive stages of evaluation and screening, including the use of molecular dynamics simulations of various durations, the antiandrogen ketodarolutamide was identified as a compound capable of stably interacting with RBD in its hACE2-binding site. Although the potential of this drug for the direct treatment of COVID-19 (coronavirus disease 2019) is limited due to its potent antiandrogenic effect, it can be used as a scaffold for the development of more promising substances. No less important result of this study is also the disclosure of a potential mechanism of the effect of ketodarolutamide on the course of COVID-19 during the treatment of androgen-dependent prostate cancer. Reports of this effect were noticed during the preparation of this manuscript.

Keywords: EBV, SARS-CoV-2, BHRF1, RBD, Omicron variant, Delta variant, viruses, antiviral drug development, virtual screening, molecular dynamics, molecular docking, inhibition of virus-host interaction, DNA intercalators.

СПИСОК ПУБЛІКАЦІЙ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові статті за основною темою наукової роботи

(* – особистий внесок здобувача)

1. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2025). A thorough insight into the life cycle of the Epstein-Barr virus. From the molecular to the organismal level. *Current Research in Microbial Sciences* 9, 100505. <https://doi.org/10.1016/j.crmicr.2025.100505>.

(*проаналізовано весь спектр наукових робіт щодо особливостей життєвого циклу вірусу Ештейна-Барр депонованих у PubMed, сформовано і письмово викладено повноцінний механістичний погляд на життєвий цикл цього патогену)

2. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2025). In silico development of HASDI-G2 as a novel agent for selective recognition of the DNA sequence. *Sci Rep* 15, 8577. <https://doi.org/10.1038/s41598-025-89967-1>.

(*розроблено HASDI-G2, проведено ряд симуляційних експериментів де підтверджено покращений ступінь його селективності щодо ДНК-послідовності)

3. Zaremba, A.A., Zaremba, P.Y., and Zahorodnia, S.D. (2023). In silico study of HASDI (high-affinity selective DNA intercalator) as a new agent capable of highly selective recognition of the DNA sequence. *Sci Rep* 13, 5395.

<https://doi.org/10.1038/s41598-023-32595-4>. (*розроблено HASDI, проведено ряд симуляційних експериментів де підтверджено наявність селективності щодо ДНК-послідовності)

4. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2023). De novo designed inhibitor has high affinity to four variants of the RBD of S-glycoprotein of SARS-CoV-2 - an in silico study. *Journal of Biomolecular Structure and Dynamics* 41, 9389–9397.

<https://doi.org/10.1080/07391102.2022.2141886>. (*розроблено малу молекулу здатну зв'язуватись з RBD SARS-CoV-2, проведено ряд симуляційних експериментів де підтверджено можливість її зв'язування з чотирма варіантами SARS-CoV-2)

5. Zaremba, A.A., Zaremba, P.Y., and Platonov, M.O. (2023). De novo designed EBAI as a potential inhibitor of the viral protein BHRF1. Research in silico. Journal of Biomolecular Structure and Dynamics 41, 3680–3685. <https://doi.org/10.1080/07391102.2022.2053746>. (*розроблено EBAI, проведено два паралельних симуляційних експерименти де підтверджено здатність до комплексоутворення з BHRF1 вірусу Епштейна-Барр)
6. Zaremba, A., Zaremba, P., Muchnyk, F.V., Baranova, G., and Zahorodnia, S. (2021). In silico Identification of a Viral Surface Glycoprotein Site Suitable for the Development of Low Molecular Weight Inhibitors for Various Variants of the SARS-CoV-2. Mikrobiolohichnyi Zhurnal 84, 34–43. <https://doi.org/10.15407/microbiolj84.01.034>. (*проведено ряд симуляційних експериментів, визначено амінокислотний склад інтерфейсу взаємодії RBD/hACE2 і ідентифіковано кишеню придатну для СОДЛ)
7. Zaremba, A., Zaremba, P., Budzanivska, I., and Zahorodnia, S. (2022). Patterns of the influence of vaccination on the dynamics of different SARS-CoV-2 variants spread. Two-year analysis. Bulletin of Taras Shevchenko National University of Kyiv Series Biology 89, 39–45. <https://doi.org/10.17721/1728.2748.2022.89.39-45>. (*проаналізовано дані з відкритих джерел щодо різних епідемічних і соціальних аспектів поширення SARS-CoV-2 у Індії, Німеччині та Україні, визначено залежність між ступенем вакцинованості населення і часткою варіанту Дельта)

Наукові праці, які засвідчують апробацію матеріалів дисертації

8. Zaremba, A., Zaremba, P., and Zahorodnia, S. In silico identification of a new potential drug-binding pocket on the surface of the receptor-binding domain of the SARS-CoV-2 S-glycoprotein. The 1st International Electronic Conference on Medicinal Chemistry and Pharmaceutics, 1–30 November 2025, Basel, Switzerland, online.
9. Zaremba A., and Zahorodnia S. In silico study of ketodarolutamide as a potential

inhibitor of the SARS-CoV-2 RBD interaction with human ACE2. The V Scientific Conference “Youth and Modern Problems of Microbiology and Virology”, 19-20 November 2024, Kyiv, Ukraine, P. 48.

10. Zaremba A., and Zahorodnia S. 3D structure data validation of the SARS-CoV-2 protein E transmembrane domain pentamer form as a potential target for drug development. Modern aspects of microbiology, virology and biotechnology in wartime and post-war period, 15-16 November 2023, Kyiv, Ukraine, P. 277-278.

11. Zaremba, A., Shalimov, O., Onys'ko, P., and Zahorodnia, S. In silico identification of a potential inhibitor of the SARS-CoV-2 S-glycoprotein receptor-binding domain interaction with human ACE2. 9th International Electronic Conference on Medicinal Chemistry, 1–30 November 2023, Basel, Switzerland, online.

12. Zaremba, A., Zaremba, P., and Zagorodnya, S. Selective DNA intercalation of massive molecules as a new method of highly specific inhibition of transcription. 6th International Electronic Conference on Medicinal Chemistry, 1–30 November 2020, Basel, Switzerland, online.

ЗМІСТ

Перелік умовних позначень.....	19
ВСТУП	22
РОЗДІЛ 1 Вірус Епштейна-Барр і SARS-CoV-2, їх життєві цикли. Сучасний стан таргетної антивірусної терапії.....	31
1.1 Вірус Епштейна-Барр.....	31
1.1.1 Будова віріону і геному вірусу Епштейна-Барр.....	31
1.1.2 Життєвий цикл вірусу Епштейна-Барр.....	32
1.1.3 Хвороби, пов'язані з вірусом Епштейна-Барр.....	46
1.1.4 Сучасний стан у розробці і пошуку антивірусних препаратів проти вірусу Епштейна-Барр.....	49
1.2 Вірус SARS-CoV-2.....	50
1.2.1 Будова віріону вірусу SARS-CoV-2. Структурні білки.....	50
1.2.2 Геном вірусу SARS-CoV-2.....	51
1.2.3 Життєвий цикл вірусу SARS-CoV-2.....	52
1.2.4 Сучасний стан у розробці і пошуку антивірусних препаратів проти вірусу SARS-CoV-2.....	58
РОЗДІЛ 2 Структуро орієнтований дизайн ліків.....	62
РОЗДІЛ 3 Методологія роботи.....	72
3.1 Джерела вхідних даних і їх підготовка.....	72
3.1.1 Отримання статистичних даних.....	72
3.1.2 Ліганди і їх бібліотеки.....	72
3.1.3 Цілі і їх підготовка.....	73
3.2 Молекулярний докінг і віртуальний скринінг.....	73
3.3 Симуляція молекулярної динаміки.....	74

3.4 Розрахунок вільної енергії зв'язування.....	76
3.5 Аналіз результатів.....	77
3.6 Ітераційний дизайн малих лігандів.....	77
3.7 Обрахунок фармакокінетичної прийнятності.....	79
РОЗДІЛ 4 Розробка інгібіторів репродукції вірусу Епштейна-Барр шляхом застосування структуро орієнтованого дизайну ліків.....	80
4.1 <i>De novo</i> розробка ЕВАІ як потенційного інгібітора антиапоптозного фактору вірусу Епштейна-Барр білка BHRF1.....	80
4.2 <i>De novo</i> розробка селективних ДНК-поліінтеркаляторів з метою протидії вірусу Епштейна-Барр на всіх стадіях життєвого циклу.....	85
РОЗДІЛ 5 Аналіз мутаційного потенціалу SARS-COV-2 і розробка/пошук інгібіторів його проникнення шляхом застосування структуро орієнтованого дизайну ліків.....	125
5.1 Аналіз статистично значимих змін у поширенні різних варіантів SARS-CoV-2, пов'язаних з поствакцинальним імунітетом в період пандемії.....	125
5.2 Ідентифікація консервативної кишені в межах ділянки взаємодії RBD з hACE2, придатної для зв'язування низькомолекулярної сполуки.....	135
5.3 <i>De novo</i> розробка інгібітора RBD S-глікопротеїну, потенційно активного проти чотирьох штамів SARS-CoV-2.....	143
5.4 Пошук і ідентифікація потенційного інгібітора RBD S-глікопротеїну SARS-CoV-2 серед затверджених FDA лікарських сполук.....	153
РОЗДІЛ 6 Обговорення результатів роботи.....	161
ВИСНОВКИ	179
Перелік використаних джерел.....	181
Додатки.....	218

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

а.з.	– амінокислотних залишків
ВЕБ	– вірус Епштейна-Барр
ВООЗ	– Всесвітньої організації охорони здоров'я
длДНК	– дволанцюгова ДНК
ЛБ	– лімфоми Беркитта
ЛХ	– лімфома Ходжкіна
НК	– назофарингеальна карцинома
олДНК	– одноланцюгова ДНК
ПП	– пірол-імідазол поліаміди
п.о.	– пар основ
ПТЛЗ	– посттрансплантаційне лімфопроліферативне захворювання
РТК	– реплікативно-транскрипційний комплекс
СОДЛ	– структуро орієнтований дизайн ліків
ЯПК	– ядерно-поровий комплекс
ADMET	– absorption, distribution, metabolism, excretion, toxicity
BCR	– B-cell antigen receptor
BINDI	– BHRF1-INhibiting Design acting Intracellularly
CATC	– capsid-associated tegument complexes
CDK	– cyclin-dependent kinase
COVID-19	– coronavirus disease 2019
DS	– dyad symmetry

EBAI	– Epstein-Barr antiapoptotic inhibitor
FDA	– Food and drug administration
FR	– family of repeats
H	– helix
hACE2	– human angiotensin-converting enzyme 2
HASDI	– High-Affinity Selective DNA Intercalator
HASDI-G2	– High-Affinity Selective DNA Intercalator – Generation 2
HLA	– human leukocyte antigen
IRES	– internal ribosome entry site
ITAM	– immunoreceptor tyrosine-based activation motif
LMPs	– latent membrane proteins
NEC	– nuclear egress complex
NSP	– non-structural proteins
NTD	– N-terminal domain
ORF	– open reading frame
PIC	– preinitiation complex
pp	– polyprotein
PRF	– programmed ribosomal frameshifting
RBD	– receptor-binding domain
RdRp	– RNA-dependent RNA polymerase
RC	– replication compartments

- RMSD – root mean square deviation
- RMSF – root mean square fluctuation
- SARS-CoV-2 – severe acute respiratory syndrome coronavirus 2
- sgRNA – subgenomic RNA
- SL – Stem-Loop
- TRS – transcription regulatory sequence
- UTR – untranslated region
- vPIC – viral preinitiation complex

ВСТУП

Актуальність роботи

Віруси є дивовижним проявом організації живої матерії, єдиним здатним виявляти усі ознаки живого виключно у складі клітинних форм життя. У позаклітинному просторі вони за своєю суттю є лише макромолекулою генетичного матеріалу захищеною від зовнішнього середовища білковою або білок-ліпідною оболонкою. Хоча деякі обмежуються лише особливою за свою стабільністю формою РНК. Нині точаться дискусії щодо того коли саме виникли віруси [1]. Можливо, вони були присутніми ще до появи перших клітин, а може є наслідком їх деградації і спеціалізації. На ці питання ще буде потрібно відповісти. Однак дещо ми знаємо точно: попри очевидний позитивний внесок вірусів в еволюційному плані, переважна більшість з них є патогенами, які щороку забирають мільйони життів [2].

Дана робота є присвяченою пошуку і розробці ліків проти двох дуже різних у багатьох аспектах, однак подібних у значному впливі на людство вірусів – вірусу Епштейна-Барр (ВЕБ) і SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2). Перший, як і всі герпесвіруси, характеризується надзвичайним поширенням серед населення планети, що є наслідком запрограмованої на рівні геному латентної стадії у його життєвому циклі [3]. На цьому етапі вірус Епштейна-Барр є невидимим для імунної системи, він не продукує вірусних білків і майже не впливає на клітинний метаболізм. Однак цьому кінцевому стану ідеального маскування передують складні процеси, коли ВЕБ крок за кроком, використовуючи в тому числі і механізми імунної системи господаря, підкоряє уражену клітину. Ці надзвичайно точні механізми є вивіреніми мільйонами років коеволюції патогену і господаря [4]. Здебільшого вони дозволяють успішно їм співіснувати, однак значне поширення ВЕБ у людській популяції виводить на передній план саме випадки помилок у цій еволюційно сформованій парі. Нині вірус Епштейна-Барр відносять до нечисленноі, але впливової, групи вірусів асоційованих з раком.

BHRF1 є BCL-2-подібним білком цього патогену, який виконує функцію дерегуляції внутрішнього шляху запуску апоптозу і забезпечує йому відтермінування смерті ураженої клітини та відповідне збільшення виходу віріонів [5]. Однак апоптотична смерть клітини не тільки є ключовим способом протидії вірусній інфекції, а і забезпечує чіткий контроль над гіперпроліферуючими клітинами, такими як ракові. Наразі доведеною є вирішальна роль BHRF1 у патогенезі і лікарській резистентності таких поширених і часто ВЕБ-індукованих ракових захворювань як лімфома Беркітта (ЛБ) [6] та назофарингеальна карцинома (НК) [7]. Таким чином, цей білковий фактор розглядається на роль перспективної цілі для розробки препаратів не тільки здатних протидіяти ВЕБ, а і спрощувати лікування опосередкованих цим вірусом видів раку.

Однак через згадані особливості життєвого циклу вірусу Епштейна-Барр, інгібування будь-якого білкового фактору є перспективним лише з точки зору лікування порушень пов'язаних з літичною стадією його життєвого циклу. У випадку спроби повноцінного очищення від патогену, яка, до прикладу, буде необхідною при трансплантації від ВЕБ-позитивного донора до ВЕБ-негативного реципієнта [8], єдиним фактором придатним для націлювання є епісома ВЕБ, присутня в майже інтактному стані у складі пулу уражених В-лімфоцитів пам'яті.

SARS-CoV-2, натомість, є емерджентним патогеном. Тобто таким, для якого людина є новим господарем. Відсутність пристосованості до нашої вродженої імунної відповіді детермінувало для нього лише один спосіб виживання – максимально швидку репродукцію і високу контагіозність. Це призвело до надзвичайно високого рівня захворюваності і смертності на початку пандемії. З часом SARS-CoV-2 еволюціонував і пристосовувався до людського організму. Зниження захворюваності і найголовніше смертності дозволило зняти пандемічний статус з цього патогену [9]. І все ж цей вірус досі знаходиться під постійним наглядом через високу швидкість еволюційних змін та пов'язаних з цим хвиль загострення епідемічного статусу в осінньо-зимовий період [10]. Сумарно це може

свідчити на користь того, що SARS-CoV-2 починає займати нішу сезонного збудника із постійним щорічним негативним впливом на людство [11].

S-глікопротеїн SARS-CoV-2 і зокрема його рецепторзв'язуючий домен (RBD від *receptor-binding domain* (анг.)) виконують ключову роль у життєвому циклі цього вірусу забезпечуючи первинну рецепцію із поверхневою формою ангіотензинперетворюючого ферменту людини (hACE2 від *human angiotensin-converting enzyme 2* (анг.)). Останній є особливо рясно присутнім на поверхні епітеліоцитів дихальних шляхів людини. Через свою ключову роль, RBD нині вважається однією із найбільш перспективних цілей для розробки антикоронавірусних ліків [12].

Таким чином, як ВЕБ, так і SARS-CoV-2 є патогенами, розробка ліків проти яких є одним із першочергових завдань світової науки і фармакології.

Найбільш сучасним і таким, що зарекомендував себе як компроміс між швидкістю і затратами ресурсів є структуро орієнтований дизайн ліків (СОДЛ) [13]. Націленість на конкретний фактор життєвого циклу патогену дозволяє не лише чітко контролювати процес розробки, використовуючи біохімічні методи, а і в значній мірі оминати стадію скринінгу великої кількості хімічних речовин з надією відшукати достатньо активну молекулу. Це стало можливим останні десятиліття завдяки розвитку обчислювальної хімії і біології. Зокрема, застосування молекулярного докінгу і інших підходів дозволяють в тисячі разів пришвидшити скринінг бібліотек хімічних сполук розміром в десятки мільйонів молекул і в подібному ступені здешевити цей процес. Отримане таким чином зменшення вибірки дозволяє досліджувати *in vitro* лише тисячі чи сотні сполук з найбільшим біологічним потенціалом. Одночасно методологія розробки лігандів виключно на основі структури цілі *de novo* надає додаткову можливість принципово вийти за межі відомого хімічного простору, де останнє є важливим через деяку широко відому однотипність наявних баз хімічних речовин, які опираються на сучасні підходи до хімічного синтезу і вже синтезовані сполуки [14].

Мета, об'єкт і предмет роботи

Метою даної дисертаційної роботи є розробка *de novo* і пошук хімічних сполук потенційно здатних націлено протидіяти життєдіяльності вірусу Епштейна-Барр і SARS-CoV-2, а також дослідження біологічних особливостей відповідних факторів життєвого циклу цих патогенів, необхідних для успішного застосування структуро орієнтованого дизайну ліків.

Таким чином, **об'єктом** даної роботи є фактори життєвого циклу вірусу Епштейна-Барр та SARS-CoV-2 придатні для їх використання в ролі цілей для структуро орієнтованого дизайну ліків – RBD S-глікопротеїну (SARS-CoV-2), BHRF1 (БЕБ) та вірусний геном (БЕБ).

Предметом досліджень, відповідно, є *de novo* розроблені і ідентифіковані хімічні сполуки потенційно здатні втручатись у функціонування RBD S-глікопротеїну SARS-CoV-2, BHRF1 та геному вірусу Епштейна-Барр, а також важливі для успішного застосування СОДЛ біологічні особливості цих вірусних факторів.

Завдання роботи і використані методи

1. *De novo* розробити малу молекулу потенційно здатну до націленого інгібування антиапоптотичного фактору вірусу Епштейна-Барр – білка BHRF1. Як методологічну основу використати ітераційний підхід з методом симуляції молекулярної динаміки в ролі основного джерела даних щодо стабільності ліганд-рецепторного комплексу.

2. *De novo* розробити молекулярну конструкцію здатну високоселективно взаємодіяти з геном транскрипційного фактору вірусу Епштейна-Барр EBNA1. В ролі структурної основи використати поліінтеркалятори, як клас сполук з доведеною здатністю до взаємодії з ДНК.

3. Проаналізувати мутаційний потенціал SARS-CoV-2 шляхом пошуку залежностей між ступенем вакцинованості і поширеністю різних варіантів патогену серед населення Німеччини, Індії та України.

4. Ідентифікувати, спільну для різних варіантів SARS-CoV-2, кишеню розташовану в межах інтерфейсу взаємодії рецепторзв'язуючого домену його S-глікопротеїну з hACE2 і придатну для пошуку та розробки антикоронавірусних ліків. В ролі методу для дослідження топології поверхні цілі використати симуляцію молекулярної динаміки.

5. *De novo* розробити малу молекулу здатну до стабільної взаємодії з відкритою кишенею (п. 4) і таким чином потенційно придатну до застосування в ролі селективного інгібітора RBD S-глікопротеїну різних варіантів SARS-CoV-2. Як методологічну основу використати ітераційний підхід з методом симуляції молекулярної динаміки в ролі основного джерела даних щодо стабільності ліганд-рецепторного комплексу.

6. Дослідити затверджену FDA (Food and Drug Administration) базу лікарських сполук на предмет наявності потенційно активних інгібіторів RBD S-глікопротеїну SARS-CoV-2. Як методологічну основу використати класичний метод віртуального скринінгу з подальшою валідацією результатів методом симуляції молекулярної динаміки ліганд-рецепторного комплексу.

Таким чином, задля успішного виконання зазначених завдань передбачалось застосування ітераційного підходу та таких *класичних методів структуроорієнтованого дизайну ліків* як віртуальний скринінг на основі напівгнучкого молекулярного докінгу та симуляція молекулярної динаміки ліганд-рецепторних комплексів. При цьому, окрім класичних параметрів стабільності комплексу, як от середньоквадратичне відхилення та залежність кількості міжмолекулярних взаємодій від часу симуляції, траєкторія симуляції молекулярної динаміки ліганд-рецепторного комплексу також використовувалась для обрахунку вільної енергії зв'язування методами MM/PBSA та MM/GBSA.

У випадку менш дослідженого вірусу SARS-CoV-2, задля визначення ступеня еволютивної лабільності патогену, був використаний порівняльний аналіз як метод загальної оцінки швидкості пристосувальних змін популяції вірусу до позитивного

відбору, одним з факторів якого в майбутньому може стати розроблений в межах даної роботи інгібітор.

Зв'язок роботи з науковими програмами, планами, темами, грантами

Дисертаційна робота була виконана в межах наукових проєктів відділу репродукції вірусів Інституту мікробіології і вірусології ім. Д.К. Заболотного НАН України: «Особливості репродукції ДНК-вмісних вірусів людини за впливу потенційних противірусних речовин» (реєстраційний номер 0220U000575, 2020-2024 рр.), «Прогнозування біоактивності та створення сучасних підходів до синтезу гетероциклічних сполук для потреб фарміндустрії» (реєстраційний номер 0123U10137, 2023-2024 рр.) та «Дослідження топологічних особливостей функціональної поверхні факторів життєвого циклу SARS-CoV-2 на предмет їх потенціалу до взаємодії з низькомолекулярними хімічними сполуками» (реєстраційний номер 0125U002921, 2025-2026 рр.).

Внесок результатів роботи у світову систему наукових знань (новизна роботи)

В межах дисертаційної роботи було розроблено та ідентифіковано ряд молекулярних конструкцій, потенційно здатних пригнічувати життєдіяльність вірусу Епштейна-Барр та SARS-CoV-2. Зокрема:

Розроблено хімічну структуру, потенційно здатну виступити першим у своєму роді низькомолекулярним інгібітором антиапоптотичного фактору вірусу Епштейна-Барр білка BHRF1 – EBAl. Таким чином, на принциповому рівні було доведено можливість розробки малої хімічної речовини здатної взаємодіяти з ВНЗ-зв'язуючою ділянкою BCL-2-подібних білків вірусного походження.

Розроблено дві модульні молекулярні конструкції (HASDI і HASDI-G2) здатні до інтеркаляції та високоселективного розпізнавання певної послідовності ДНК гену EBNA1 – ключового транскрипційного фактору ВЕБ. Подібні структури є першими у своєму роді. Їх розробка доводить принципову можливість дизайну ДНК-поліінтеркаляторів, здатних не лише до невибіркової взаємодії з ДНК, а і до

націленого зв'язування з цією макромолекулою. Останнє відкриває двері до розробки агентів, здатних пригнічувати транскрипційну активність певних регіонів ДНК, до прикладу, патогенного походження. Це особливо важливо у випадку ДНК-вмісних вірусів, здатних до встановлення високорегульованої латентної інфекції.

Визначено, що постійно циркулююча у певній популяції господаря когорта варіантів SARS-CoV-2 характеризується високою гетерогенністю складу. На прикладі змін у співвідношенні різних варіантів під впливом масової вакцинації населення показано, що ця особливість призводить до швидкої адаптації SARS-CoV-2 до факторів позитивного відбору. Опираючись на ці дані, досліджено hACE2-зв'язуючий інтерфейс рецепторзв'язуючого домену S-глікопротеїну варіантів Ухань, P.1 і Кластер 5 цього патогену. Отримані дані дозволили ідентифікувати до того невідому кишеню в межах hACE2-зв'язуючого інтерфейсу RBD, придатну до взаємодії з потенційним лігандом і, що найважливіше, присутню в різних варіантів SARS-CoV-2. Останнє може вказувати на її відносно консервативний характер.

Розроблено хімічну структуру, потенційно здатну виступити низькомолекулярним інгібітором RBD S-глікопротеїну SARS-CoV-2. Показано, що отримана молекула є здатною до співставного за силою зв'язування з варіантами Ухань, Омикрон, Дельта та Кластер 5 цього патогену у ділянці відкритої кишені. Це дозволило як додатково підтвердити відносну консервативність цього покету та його придатність для взаємодії з малим лігандом, так і показує потенціал розробленої конструкції до впливу на різні варіанти SARS-CoV-2, що особливо важливо в контексті високої гетерогенності його популяції.

Під час поглиблених досліджень з пошуку інгібіторів RBD SARS-CoV-2 також визначно, що затверджений FDA антиандроген кетодаролутамід має потенціал до зв'язування RBD в зоні його взаємодії з hACE2. Цей потенціал прояснює раніше відомі звіти в яких хворі на рак простати, що отримували антиандрогенну терапію мали нижчий рівень захворюваності на COVID-19 (coronavirus disease 2019) [15].

Також слід зазначити, що успішні результати, отримані в межах дисертаційної роботи за рахунок інтеграції ітеративного підходу до класичної методології СОДЛ, свідчать на користь високого потенціалу подібної комбінації у сфері фармакологічних досліджень.

Практичне значення отриманих результатів

В межах дисертаційної роботи було розроблено ряд молекулярних конструкцій, націлених на такі економічно і соціально важливі патогени як вірус Епштейна-Барр і SARS-CoV-2. Подальший їх розвиток як основи для створення ряду таргетних препаратів дозволить протидіяти цим вірусам і з ними пов'язаними порушеннями на недоступному нині рівні.

Методологічні підходи використані в цій роботі і їх відпрацьоване поєднання також можуть бути основою для створення протоколів швидкої розробки таргетних антивірусних препаратів, зокрема в умовах пандемії, спричиненої емерджентним патогеном.

Особистий внесок здобувача

Здобувач здійснював попередній аналіз літературних даних в процесі підготовчого етапу кожного із досліджень, разом з науковим керівником він склав чіткий план як кожного експерименту, так і роботи загалом. Здобувач особисто проводив підготовку програмного забезпечення для подальших обрахунків, написання і тестування відповідних скриптів, та відбір необхідних вхідних даних. Він здійснив всі необхідні обрахунки, проаналізував отримані результати і сформував на основі останніх повноцінну дисертаційну роботу, яка повністю висвітлює його наукові ідеї, гіпотези і методологічні підходи. У співавторстві та під керівництвом наукового керівника здобувач представив як результати аналізу літературних даних, так і безпосередньо отримані експериментально дані широкій науковій спільноті в межах семи наукових статей опублікованих в фахових періодичних виданнях світового класу.

Апробація результатів роботи

Основні положення дисертаційної роботи були оприлюднені на: міжнародній електронній конференції «6th International Electronic Conference on Medicinal Chemistry» (1-30 листопада, 2020 р); міжнародній електронній конференції «9th International Electronic Conference on Medicinal Chemistry» (1-30 листопада, 2023 р); міжнародній науково-практичній конференції «Сучасні аспекти мікробіології, вірусології та біотехнології у воєнний та післявоєнний час» (15-16 листопада, 2023 р); V науково-практичній конференції «Молодь та сучасні проблеми мікробіології і вірусології» (19–20 листопада, 2024 р); міжнародній електронній конференції «The 1st International Electronic Conference on Medicinal Chemistry and Pharmaceuticals» (1-30 листопада, 2025 р).

Публікації за результатами дисертаційної роботи

Результати роботи представлено загалом у 12-ти публікаціях. Серед них 7 наукових статей у фахових міжнародних і вітчизняних виданнях: 3 статті опубліковано у журналах Q1, 2 статті – в журналі Q2, 1 стаття – в журналі Q4, 1 стаття – у фаховому періодичному виданні категорії «Б», а також 5 тез на міжнародних і вітчизняних конференціях.

Структура та обсяг дисертації

Загальний обсяг дисертаційної роботи становить 231 сторінок комп'ютерного тексту. Рукопис складається з анотації, вступу, двох розділів огляду літератури, розділу методології роботи, двох розділів з результатами оригінальних досліджень, розділу обговорення цих результатів, висновків, списку використаних джерел та додатків. Роботу ілюстровано 49 рисунками, 8 таблицями та доповнено 5 додатками. Перелік використаних джерел містить 299 найменувань.

РОЗДІЛ 1

ВІРУС ЕПШТЕЙНА-БАРР І SARS-COV-2, ЇХ ЖИТТЄВІ ЦИКЛИ. СУЧАСНИЙ СТАН ТАРГЕТНОЇ АНТИВІРУСНОЇ ТЕРАПІЇ

1.1 Вірус Епштейна-Барр

1.1.1 Будова віріону і геному вірусу Епштейна-Барр

ВЕБ є класичним представником родини *Herpesviridae*, підродини *Gammapherpesvirinae*. Він характеризується складним капсидом, розміром віріону близько 125 нм та стандартною для членів цього таксону тришаровою будовою. Остання включає поверхневий ліпідний бішар із інтегрованими в нього вірусними глікопротеїнами, тегумент та псевдоікосаедричний нуклеокапсид. Вірусні глікопротеїни представлені п'ятьма білками: gp350/220, gp42, BMRF2, gH/gL і gB [16]. Водночас, нуклеокапсид складається з двох основних білків: великого і малого об'єднаних в 150 гексамерів і 11 пентамерів (T=16), а також малих білків Tri1 і Tri2, які формують триплекси (Tri1+Tri2A+Tri2B) і розміщуються в проміжках між пентамерами та гексамерами. Пентамери розміщуються у всіх вершинах капсиду, окрім зайнятої порталним білком. Між нуклеокапсидом і мембраною розміщується плеоморфний тегумент, який складається із 20–40 різних вірусних білків [17, 18].

У внутрішній частині капсиду знаходиться геномом вірусу Епштейна-Барр – дволанцюгова лінійна ДНК розміром 170–180 кб. Вона кодує від 85 до понад 100 вірусних білків і 44 міРНК [19, 20]. У своїй лінійній формі длДНК (дволанцюгова ДНК) характеризується наявністю двох комплементарних регулярних ділянок на кінцях дуплексу (TR) довжиною 538 пар основ (п.о.). В подальшому, під час становлення інфекції ВЕБ, спарювання цих областей забезпечує успішне лігування кінців длДНК і утворення епісомальної форми геному. Окрім термінальних повторів, геном ВЕБ містить чотири внутрішні повторювані ділянки (IR), які не мають гомології з кінцевими повторами і кодують деякі із вірусних факторів. Кількість власне повторів є специфічною ознакою і залежить від конкретного

ізоляту [21]. Кінцеві та внутрішні повтори ділять геном ВЕБ на 5 неповторюваних частин – U1-U5, які також кодують вірусні фактори. Загалом, особливості будови, різноманіття, регуляції геному ВЕБ, а також альтернативного сплайсингу є відкритим питанням, яке активно досліджується [22, 23].

1.1.2 Життєвий цикл вірусу Епштейна-Барр

Передача ВЕБ здійснюється шляхом контакту сприйнятливою організму з рідинами ураженого організму. Інколи, через те, що найчастіше безпосереднім носієм патогену у достатній концентрації є слина інфікованої особи, інфекційний мононуклеоз навіть називають «хворобою поцілунків». Відповідно, канонічними вхідними воротами для даного патогену є ротоглотка. Нині домінує теорія, що первинними клітинами, які інфікуються при цьому є епітелій ротоглотки. Це підтверджується високою тропністю патогену до цих клітин. Однак існує і суміжна теорія, яка також має ряд підтверджень. Припускається, що віріони патогену з током слини потрапляють на піднебінні мигдалики і безпосередньо у їх крипти [24], де перетинають тонкий шар епітелію і уражають наївні В-лімфоцити. В такому випадку ураження епітеліоцитів розглядається як вторинне явище [25]. Не виключений і третій варіант, за якого репродукція патогену на перших стадіях інфікування відбувається в рівній мірі і в лімфоїдній, і в епітеліальній тканині ротоглотки [26]. За сучасними даними на молекулярному рівні процес проникнення і репродукції вірусу Епштейна-Барр значно відрізняється між цими двома типами клітин [27].

У випадку епітелію взаємодія патогену з клітиною починається із зв'язування поверхневого глікопротеїну BMRF2 з інтегрином $\alpha V\beta 1$ [28]. Безпосередньо за взаємодію з лігандом відповідає специфічний інтегринзв'язуючий структурний мотив із трьох амінокислот: аргініну, гліцину та аспартату (RGD), присутній у складі BMRF2. gH має у своєму складі подібний мотив – KGD. Цей глікопротеїн, в межах комплексу gH/gL, після первинного зближення ліпідних бішарів клітини і

віріону здатен додатково розпізнавати інтегрини $\alpha V\beta 5$, $\alpha V\beta 6$ і $\alpha V\beta 8$ [29]. На цій стадії у зону максимального наближення клітинної на вірусної ліпідних мембран залучається gB (фузійний білок III класу) [30].

У випадку В-лімфоцитів рецептором прикріплення виступає глікопротеїн gp350/220. Він представлений у двох ізоформах – 350 кДа та 220 кДа, які є результатом альтернативного сплайсингу і виконують ідентичну функцію: розпізнають або CD35 (CR1), або CD21 (CR2). Після первинної взаємодії відбувається ендоцитоз іmobilізованого віріона [31]. Далі С-кінцевий домен gp42 специфічно взаємодіє з β -ланцюгами HLA (human leukocyte antigen) класу II. Gp42, при цьому, знаходиться в складі комплексу gH/gL/gp42, де gH виступає основною структурною одиницею [32]. Зв'язування gp42 з HLA II запускає ряд конформаційних перебудов, які призводять до взаємодії gL з gB, відповідних конформаційних перебудов та проникнення вірусу.

Слід зазначити, що gH є здатним до взаємодії як з gp42 – фактором необхідним для успішного інфікування В-клітин, так і з інтегринами епітеліоцитів. Відповідно, у випадку достатньої кількості gp42 у складі мембрани віріону, gH є недоступними для взаємодії з інтегринами епітеліоцитів – тропність таких віріонів до епітеліоцитів буде зниженою. У зворотному ж випадку gp42 буде зв'язаний з HLA класу II отриманим від материнської мембрани – тропність буде зниженою вже до В-лімфоцитів. Таким чином субпопуляція ВЕБ напрацьована у В-лімфоцитах буде володіти порівняно вищою спорідненістю до епітеліоцитів носоглотки. І навпаки, напрацьована у епітеліоцитах субпопуляція ВЕБ буде більш спорідненою саме до В-лімфоцитів [33, 16].

Найперше, після проникнення ВЕБ у цитозоль і позбавлення ліпідної оболонки, відбувається дисоціація білків зовнішнього тегументу. Вони виконують функцію первинного впливу на метаболізм господаря, зменшуючи його власний білковий синтез та протидіючи стимуляції апоптозу. До прикладу, BNRF1 взаємодіє з білком Daхх, чим порушує утворення комплексу Daхх-ATRХ, що призводить до

інгібування метилювання ДНК і дозволяє зберігати транскрипційну активність геному ВЕБ. Білки внутрішнього тегументу і комплексу асоційованого з тегументом (САТС від capsid-associated tegument complexes (анг.)), зокрема BPLF1, взаємодіють з динеїнами, а також білками асоційованими з плюс кінцем мікротрубочок (+TIPs) [34]. Після асоціації з клітинним цитоскелетом нуклеокапсид ВЕБ ретроградно транспортується до розміщеної біля ядра centrosomi зі швидкістю близько 1-2 мкм/с. Далі BPLF1, завдяки сигналу ядерної локалізації, забезпечує переміщення нуклеокапсидів до ядерної мембрани. Цей же білок володіє спорідненістю до Nup358 — цитоплазматичного компонента ядерно-порового комплексу (ЯПК) [34]. Зв'язування BVRF1, ще одного компонента САТС, з нуклеопоринами Nup214 і hCG1 додатково стабілізує цю взаємодію. BVRF1 формує так званий кеп (cap – ковпачок, кришечка (анг.)) на порталній вершині нуклеокапсиду, який, внаслідок взаємодії з нуклеопоринами основи ЯПК, покидає своє положення. Таким чином, порталний комплекс, який формує канал діаметром 36Å, опиняється відкритим. Надалі за подібним до бактеріофагів механізмом вірусна ДНК покидає віріон і, проходячи через ЯПК, потрапляє у ядро [35, 36].

Після цього вірусний геном циклізується і завдяки клітинній машинерії гомологічної рекомбінації переходить в епісомальну форму [37]. Паралельно із циклізацією він взаємодіє з гістонами. При цьому, через те, що інфікованими ВЕБ є здебільшого клітини у стані спокою, класичний гістон H3, який синтезується тільки в S-фазі клітинного циклу, є заміненим на гістон H3.3, який є пов'язаним із швидким метилюванням геному [38]. Білок зовнішнього тегументу BNRF1 втручається в цей процес іммобілізуючи один з компонентів машинерії, яку запускає H3.3. Відсутність метилювання на цьому етапі призводить до слабoreгульованого синтезу літичних і латентних факторів життєвого циклу ВЕБ [39]. Цей період називають прелатентно абортивно літичною фазою або просто прелатентною фазою. Вважається, що він є необхідним для виживання ураженої клітини, оскільки тоді синтезуються такі білки, як BALF1 і BHRF1 – антиапоптотичні BCL2-подібні білки,

а також BNLF2A, BGLF5, BILF1 і BDLF3 – інгібітори експресії HLA [40].

Натомість скоординована транскрипція геному ВЕБ починається з активації кластеру промоторів W_p , що контролює транскрипцію кластеру генів EBNA-2 і EBNA-LP. Результатом цього є напрацювання однієї біцистронної мРНК [41]. Остання в подальшому зазнає процесингу і транслюється у зрілі EBNA-2 і EBNA-LP. Ці білки функціонально є дуже пов'язаними між собою коактиваторами транскрипції, які спільно впливають на активність клітинного і вірусного геномів [42]. Попри це, спектр їх активностей за сучасними даними дещо відрізняється. Вважається, що EBNA2 реалізує свій вплив здебільшого за рахунок взаємодії з таким транскрипційним фактором як RBPJ. Останній є учасником Notch-сигналізації та забезпечує репресію цільових генів після активації Notch рецептора [43]. Інша активність EBNA2 передбачає рекрутування коактиваторів EP300/CBP, факторів ремоделювання хроматину SNF/SWI, а також базальних факторів транскрипції TFIID, TFIIB і TFIIE. Натомість EBNA-LP володіє лише здатністю до іммобілізації клітинних репресорів транскрипції, зокрема NCOR та HDAC4 [44].

Одним із найкраще досліджених є механізм активації транскрипції протоонкогену MYC під час інфекції ВЕБ [45]. Передбачається, що EBNA-2 взаємодіє з енхансером на ~556-428 кб вище за послідовністю від гену MYC [46]. Залучення інших клітинних факторів дозволяє наблизити цей енхансер до промоторної ділянки MYC формуючи таким чином петлеподібну структуру. Слід зазначити, що порушення функціонування цього онкогену є одним із найпоширеніших маркерів неопластичного переродження клітини [47]. Вважається, що на більш глобальному рівні період стимульованого EBNA-2 накопичення MYC під час інфекції ВЕБ дозволяє найбільш швидко відвести уражену клітину від апоптозу [48].

Вплив EBNA-2 не обмежується клітинним геномом: EBNA-2 є сильним стимулятором промотору S_p , який знаходиться вище за послідовністю від W_p і забезпечує ефективну експресією всіх EBNA в межах одного поліцистронного

транскрипту [49]. Власне програма транскрипції вірусного геному з Ср в ролі промотора і EBNA-1, EBNA-2, EBNA-LP та EBNA-3s в ролі продуктів носить назву програми латентності Пб.

EBNA-3A, EBNA-3B та EBNA-3C є паралогами, які виникли внаслідок серії послідовних дуплікацій в процесі еволюції ВЕБ [49]. Гомологічна ділянка, яка формує схожий між ними домен, відповідає за взаємодію з згаданим раніше у контексті EBNA-2 транскрипційним фактором RBPJ [50]. Нині найбільш поширеною є гіпотеза про наявність антагоністичних взаємозв'язків між EBNA-2 і родиною EBNA-3. Пряма конкуренція дозволяє поступово, відповідно до ступеня накопичення EBNA-3, послабити надлишковий синтез MYC, стимульований EBNA-2, і відвести уражену клітину від апоптозу, спричиненого занадто довготривалою високою проліферативною активністю [51]. Однак взаємодія цих віропротеїнів далеко не обмежується RBPJ, а подібно до EBNA-2 охоплює значну частину геному як господаря, так і патогену [52].

EBNA-3A та EBNA-3C є здатними до пригнічення накопичення двох альтернативних транскриптів гену CDKN2A – білків p16^{INK4a} та p14^{ARF} [53]. Де p16^{INK4a} є негативним регулятором клітинного циклу, а p14^{ARF} інгібітором фактору MDM2 і, відповідно, стимулятором апоптозу. Одночасно EBNA-3A і EBNA-3C інгібують накопичення BIM – проапоптичного фактору родини BCL-2 [54].

EBNA-3C є єдиним серед EBNA-3s фактором здатним до активації транскрипції. Він, разом з EBNA2, є коактиватором промотора LMP-1 [55]. Також EBNA-3C активує експресію генів CXCR4 і CXCL12 [56]. Де CXCL12 є хемоатрактантом для всіх клітин, що посилено експресують CXCR4. Зв'язування CXCR4 з CXCL12 запускає опосередкований G-білками сигнальний каскад стимуляції проліферації. EBNA-3A, водночас, за рахунок прямого зв'язування з міжгенними енансерами пригнічує накопичення CXCL10 і CXCL9 – двох з трьох стимульованих IFN- γ хемокінів. Цьому процесу парадоксально протидіє EBNA3B, для якого доведеною є здатність до стимуляції CXCL10 [57].

У випадку програм латентності IIb і III ще одним білком, який напрацьовується, є EBNA-1. Окрім Cp, EBNA-1 регулюється своїм власним конститутивно активним промотором Q [58]. Його активність забезпечується транскрипційними факторами «домашнього господарства», але негативно ауторегулюється EBNA-1. Таким чином, в період високої активності Cp промотор Q є пригніченим достатньо, щоб стверджувати про відсутність його активності [59].

Основною функцією EBNA-1 є забезпечення реплікації вірусного геному синхронно з геномом господаря. Виконання функцій цього фактору починається з взаємодії його C-кінцевого домену з oriP ВЕБ, а точніше з FR (family of repeats) і DS (dyad symmetry) в межах oriP [60]. Зв'язування димерів EBNA1 в ділянці DS призводить до рекрутування білків компонентів передреплікативного комплексу і, відповідно, є пов'язаним з ініціацією реплікації геному ВЕБ [61]. Натомість взаємодія з FR-областю призводить до фізичної прив'язки вірусного геному до однієї з хромосом в локусах багатих на АТ-пари. Узгоджене таким чином напрацювання геному ВЕБ дозволяє рівномірно розділити новосинтезований вірусний генетичний матеріал між дочірніми клітинами.

Подібно до того, як EBNA-2 активує транскрипцію кластеру всіх EBNA з Cp, цей ранній латентний фактор також є потужним позитивним регулятором родини мембранних віропротеїнів (LMP) – LMP1, LMP2a та LMP2b. Концентрація цих білків, попри стимулюючий вплив EBNA-2, починає наростати через близько 2 доби після початку інфекції і лише на 28 день інфекції поступово досягає піку [62]. Власне початок масового напрацювання білків родини LMP вважається переходом паттерну транскрипційної активності геному ВЕБ від програми латентності IIb до програми латентності III.

Така кінетика накопичення факторів LMP пояснюється інгібуючим впливом MYC онкогену, який реалізується за рахунок Mad1 і Max – членів каскаду MYC-опосередкованої регуляції транскрипції. Ці фактори є здатними зв'язуватись з ДНК вище за течією від одного з енансерів транскрипції LMP1. Ця взаємодія призводить

до значного пригнічення транскрипції як LMP1, так і LMP2b [63].

LMP1 складається з шести трансмембранних спіралей та С-кінцевої цитоплазматичної термінальної області довжиною в 200 амінокислотних залишків [64]. Трансмембранна частина LMP1 забезпечує конститутивну тримеризацію цього фактору та відповідну активацію ряду сигнальних шляхів. Наразі найбільш важливими вважаються сигнальні шляхи, пов'язані із стимуляцією NF-κB – одного із ключових факторів формування імунної відповіді. На основі високої подібності механізмів активації цього сигнального шляху LMP1 вважається прямим аналогом рецептору CD40 – основного корецептора В-лімфоцитів [65].

Власне NF-κB, подібно до MYC, регулює клітинний цикл клітини [66]. Цей фактор посилює гліколіз, окислення жирних кислот, а також синтез NADPH, що дозволяє нормально активованим В-лімфоцитам виживати в темній та світлій зонах гермінативного центру під час афінного дозрівання. Слід зазначити, що одним із стимульованих цим транскрипційним фактором генів є MYC [67]. Поряд з цим існують дослідження, де постулюється несумісність високої активності NF-κB з надекспресією MYC [68]. Рознесення в часі опосередкованої вірусом надстимуляції MYC і NF-κB дозволяє ВЕБ уникнути небезпечної активації двох взаємовиключних сигнальних шляхів стимуляції росту. Однак експресія лише LMP1 без експресії інших вірусних факторів призводить до унеможливлення утворення гермінативного центру у відповідь на імунізацію [69]. Чого, однак, не спостерігається під час повноцінної інфекції.

Нині вважається, що під час інфекції роль регулятора активованого LMP1 сигнального каскаду виконує ще один поверхневий віропротеїн – LMP2A [70]. LMP2A також складається з мембранної та цитоплазматичної частин. Остання містить як С-кінець, який забезпечує гомотипові взаємодії, так і N-кінець, який забезпечує основну функціональну активність і містить ITAM мотив (Immunoreceptor Tyrosine-based Activation Motif). ITAM в нормі є присутнім у складі BCR (B-cell antigen receptor) [71]. На відміну від BCR ITAM, тирозини LMP2A

ITAM є постійно доступними для фосфорилування, а такі стимульовані BCR сигнальні шляхи як PLC γ 2, PI3K/Akt та MAPK, відповідно, є конститутивно активними.

Останнім членом родини LMP є білок LMP2B. LMP1 та LMP2B знаходяться під контролем одного двонаправленого промотора. Цей промотор знаходиться в межах гена LMP2A нижче за послідовністю. Таким чином, ініціація транскрипції з LMP1/LMP2B призводить до напрацювання LMP1 та його усіченої форми – LMP2A. Ця усічена форма містить повноцінний трансмембранний домен і С-кінець. Таким чином, припускається, що ймовірною (але не обмежуючись) функцією LMP2B є регулювання функції останнього. Опосередковано це підтверджується локалізацією LMP2A та LMP2B в одних і тих же клітинних компартментах та інгібуванням фосфорилування LMP2A при його коекспресії з LMP2B [72]. У випадку інтактних лімфоцитів патерн експресії геному ВЕБ стабілізується на рівні програми латентності III, іморталізуючи уражені клітини [73]. Однак *in vivo*, після проникнення у фолікул і локалізації в зоні формування гермінативного центру, уражена клітина змінює паттерн експресії геному ВЕБ [74]. Зокрема всі фактори ВЕБ, які залежать від активності С ρ (EBNA-2, EBNA-LP, EBNA3s) є відсутніми в таких лімфоцитах. Цей паттерн транскрипційної активності геному ВЕБ отримав назву програми латентності IIa, або в більш ранній літературі – програми латентності II. Перехід від однієї транскрипційної програми латентності до іншої є результатом прямої дії специфічного мікрооточення гермінативного центру.

Було виявлено, що більшість цитокінів гермінативного центру впливали на транскрипційну активність ВЕБ, хоча і по-різному. Всі інтерлейкіни фолікулярних дендритних клітин (FDC) – IL-6, IL-15 та IL-27 пригнічують EBNA-2 і, відповідно, EBNA-3s не впливаючи на решту латентних білків. Інтерлейкіни IL-2 і IL-4 CD4⁺ Т-лімфоцитів впливають подібним чином. Натомість інтерлейкіни IL-10 та IL-21, які також виділяються CD4⁺ Т-лімфоцитами, поряд із вираженим пригніченням С ρ призводять до явного посилення експресії LMP1, особливо IL-21.

Вплив інтерлейкінів, як і більшості цитокінів, є опосередкованим JAK/STAT сигнальним каскадом [75]. У цьому дослідженні вчені припустили, що синергетичний вплив цитокінів на уражені ВЕБ лімфоцити може бути опосередкований саме за рахунок STAT і їх комбінації. Подальші їх дослідження підтвердили це припущення.

Після переходу транскрипційної програми до Па, основну роль починають відігравати LMPs (latent membrane proteins). Наразі немає даних щодо того, яким чином відбувається регуляція цієї родини білків в період такого високорегульованого процесу як дозрівання афінності. Однак на основі значно зниженої кількості уражених лімфоцитів – 3-4 клітини на весь гермінативний центр, а також збільшеної концентрації активованих цитотоксичних Т-клітин в уражених ВЕБ мигдаликах, можна припустити значно посилену елімінацію уражених В-клітин [76, 77, 78]. Враховуючи складні перехресні взаємодії на рівні сигналізації як в межах родини LMPs, так і CD40/BCR, виглядає ймовірно, що виживають лише клітини, які мають достатньо жорстковрегульований рівень взаємодії вірусних і клітинних регуляторних білків, а всі інші елімінуються.

Після виходу за межі гермінативного центру, уражена ВЕБ В-клітина пам'яті, подібно до природно диференційованих В-лімфоцитів, з часом депонується у селезінці, червоному кістковому мозку або ж периферичних тканинах [79]. Частина також залишається циркулювати з кровотоком. В цей же період відбувається майже повне припинення їх проліферативної активності. Уражені ВЕБ лімфоцити серед вірусних білків або експресують лише EBNA1 (під час рідкісної події поділу), або взагалі не синтезують віропротеїнів. Ці два варіанти патерну експресії вірусного геному називають програмою латентності I і програмою латентності 0, відповідно [80]. Пригнічення LMPs на даному етапі ймовірно відбувається через відсутність активуючої дії цитокінів за межами гермінативного центру.

В цей період широко присутні в диференційованих В-клітинах транскрипційні фактори YY1, E2-2, MEF-2D і ZEBs пригнічують надранні літичні гени BZLF1 та

BRLF1 за рахунок зв'язування безпосередньо в межах Z_p, або ж вище за послідовністю від R_p [81], де Z_p і R_p є промоторами для BZLF1 BRLF, відповідно. Продуктами цих генів є транскрипційні фактори Zta і Rta – трансактиватори промоторів літичних білків ВЕБ. Регуляція цього процесу також відбувається на рівні білкових продуктів. Зокрема, транскрипційні фактори PAX5 і Oct-2, які беруть активну участь на всіх стадіях розвитку В-лімфоцитів, окрім стадії В-плазмоциту, взаємодіють з Zta, що призводить до зниження його транскрипційної активності [82, 83].

Попри чітку супресивну дію вже самого репертуару білків спочиваючих В-клітин пам'яті, деяка частина таких клітинних резервуарів вірусу постійно зазнає літичної реактивації патогену. Існують дослідження, які постулюють значний внесок у цей процес клітинної сигналізації BCR. Зокрема, активація BCR призводить до швидкої і потужної активації таких сигнальних шляхів як JNK, MAPK/p38, ERK, PKC, PKD та PI3K/AKT [81]. MEF-2D, згаданий раніше як інгібітор транскрипції BZLF1, є одним із цілей функціонування цих сигнальних шляхів. Зокрема, було показано, що його дефосфорильована активна форма є здатною стимулювати Z_p [84]. TGF-β і його канонічний сигнальний каскад також беруть участь в реактивації ВЕБ. Основним посередником передачі сигналу в даному випадку виступають білки Smad, які після фосфорилування зв'язуються із Z_p у щонайменше п'яти сайтах і рекрутують фактори транскрипції [85].

Zta і Rta є стимуляторами транскрипції ранніх генів літичного циклу, один одного, а також безпосередньо самих себе [86]. Серед ранніх літичних факторів BALF5, BMRF1, BALF2, BBLF4, BSLF1 та BBLF2/3 є компонентами вірускованої літичної реплікативної машинерії ВЕБ. Зокрема, BALF5 кодує ДНК-полімеразу, BMRF1 – длДНК-зв'язуючий білок та фактор посилення процесивності BALF5, BALF2 – зв'язуючий одноланцюгову ДНК (олДНК) білок, BBLF4 – геліказу, BSLF1 – праймазу, BBLF2/3 – лінкерний білок геліказо-праймазного комплексу. Ген BKRF3 натомість кодує урацил специфічну ДНК-глікозилазу, BGLF4 –

протеїнкіназу, BARF1 – вірорецептор для M-CSF, а продуктами BALF0/1 і BHRF1 є два однойменних BCL2 подібних антиапоптотичних фактори. BDLF3.5, BDLF4, BVLF1, BGLF3, BFRF2 і BcRF1 є білками вірусного преініціаторного комплексу (vPIC від viral preinitiation complex (анг.)), необхідного для синтезу пізніх факторів ВЕБ [87].

Основним процесом, який ініціюється на цій стадії життєвого циклу ВЕБ є реплікація його геному. На першому етапі білок Zta взаємодіє з відповідними йому сайтами в межах oriLyt [88]. Далі цей фактор рекрутує геліказу BBLF4, яка в свою чергу входить до геліказо-праймазного комплексу BBLF4/BSLF1/BBLF2/3 [89]. На даному етапі відбувається індуковане плавлення ДНК і залучення BALF2 для захисту олДНК від дії ендонуклеаз. BMRF1 також володіє спорідненістю до Zta, долучаючись до комплексу і рекрутуючи BALF5 [90]. Таким чином, Zta можна розглядати в ролі основного організатора вірусоспецифічної реплікативної машинерії ВЕБ.

Внаслідок збірки реплісоми і відповідного формування класичної реплікативної вилки, напрацьовуються одночасно обидва ланцюги вірусної ДНК [91]. Їх асоціація формує новостворену дволанцюгову, але лінійну форму геному ВЕБ. Рух реплікативної машинерії епісомою вважається циклічним, що забезпечує продукцію конкатемерної ДНК.

BGLF4 є ще одним продуктом ранньої літичної транскрипції ВЕБ. Ця кіназа є гомологом клітинних циклін-залежних кіназ (CDK від cyclin-dependent kinases (анг.)), здатним фосфорилувати такі клітинні фактори, як UXT, IRF3, TIP60, p27^{KIP}, H2AX, конденсини, топоізомеразу II, MCM4 та ламін A/C [92]. Функціональна мімікрія цього віропротеїну з CDK, модифікована для константного впливу, дозволяє вірусу Епштейна-Барр зупинити життєвий цикл на рівні S-фази як найбільш багатой доступними для синтетичних процесів ресурсами, при цьому унеможливаючи характерну для синтетичної фази реплікацію клітинного геному. Фосфорилування ламінів із їх подальшою деполімеризацією на даному етапі є

підготовкою для успішного формування віріонів на більш пізніх стадіях життєдіяльності ВЕБ і їх транспорту до плазмолемі.

Згаданий вище BARF1 є мембраноасоційованим пептидом, який в процесі дозрівання глікозилується і позбувається свого гідрофобного N-кінця [93]. Це призводить до його екскреції в міжклітинний простір у функціонально активній гексамерній формі, де він шляхом зв'язування інактивує M-CSF — макрофагальний колонієстимулюючий фактор, необхідний для нативної диференції мононуклеарних клітин імунітету.

Фактори BALF0/1 і BHRF1 є прямими гомологами білків родини Bcl-2 [94]. BHRF1 – білок розміром 21,9 кДа, який має у своєму складі три характерних для родини Bcl-2 домени – BH1, BH2 та BH3. Цей віропротеїн здатен, подібно до клітинних антиапоптотичних білків, секвеструвати як проапоптотичні BH3-вмісні регуляторні Bim, Bid та Bcl-2, так і ефекторні Bak та Bax [5]. Це призводить до протидії апоптозу одразу на двох взаємодоповнюваних рівнях: на рівні вивільнення ефекторних Bak та Bax з під впливу антиапоптотичних членів BCL-2 та на рівні збірки порового комплексу, необхідного для ініціації пермеабілізації зовнішньої мембрани мітохондрій. Останнє є важливим, якщо концентрація Bak і Bax є достатньою для подолання першої перепони [95]. Загалом BHRF1 є добре дослідженим фактором, який розглядається на роль потенційної мішені для розробки антивірусних ліків.

Набагато більш суперечливими є дані щодо BALF0/1. Нині існують лише непрямі, хоча і чисельні, підтвердження його сильної антиапоптотичної функції [96, 97].

Як було описано вище, після накопичення допускового рівня білків реплікативної машинерії, починається масовий синтез вірусної конкатемерної дволанцюгової ДНК, де кожна наявна на той момент епісома ВЕБ слугує матрицею. Накопичення ДНК супроводжується нагромадженням білків реплікативної машинерії та допоміжних факторів в зоні розташування вихідної матриці. Таким

чином, формується компартмент реплікації (RC від replication compartments (анг.)) – чітко відділена зона масової вірусної реплікації, транскрипції літичних генів і збірки віріонів [88].

Особливу роль в формуванні RC займає BMRF1. Цей віропротеїн за рахунок спорідненості до дволанцюгової ДНК виконує не тільки функцію фактору процесивності BALF5, а і зв'язує більшість вірусної длДНК [98]. Одночасно BMRF1 володіє здатністю до гомотипових взаємодій, що призводить до утворення так званого BMRF1-ядра — щільного нуклеопротеїнового конгломерату в межах RC, додатково захищеного від дії клітинних ендонуклеаз [99].

Транскрипція пізніх літичних генів ВЕБ не залежить від 5'-енхансерних елементів, необхідних для ініціації транскрипції всіх інших відкритих рамок зчитування (ORF від open reading frame (анг.)). Їх промоторна ділянка містить послідовність TATTWAA замість канонічного ТАТА-боксу [100], яку розпізнає вірусний ТАТА-зв'язуючий білок VcRF1 і прямо рекрутує людську РНК-полімеразу II [101]. Надалі з VcRF1 зв'язується BGLF3, який у свою чергу володіє спорідненістю до BVLF1, BDLF4 та BFRF2 [102]. Таким чином, BGLF3 виконує функцію структурного білка для утворення vPIC. Залежними від vPIC є структурні білки VcLF, BFRF3, BDLF1, BORF1, BGLF1 та BBRF1, білки тегументу, глікопротеїни ліпідної оболонки BALF4 (gB), BXLF2 (gH), BZLF2, BLLF1 та BILF2, а також допоміжні – BVRF2, BFRF0.5, BFLF1, BFRF1 та BXRf1 [103].

Збірка капсиду ВЕБ починається з асоціації між собою BBRF1 (портал), VcLF1, BORF1 та BDLF1. Цей процес відбувається поза BMRF1-ядром і призводить до утворення нестабільних конгломератів, які стабілізуються за рахунок асоціації з VdRF1 і протеазою BVRF2. Для ангуляризації (ущільнення) прокапсиду необхідно, щоб термінальний комплекс асоціювався з порталом і індукував активацію BVRF2. Власне ангуляризована форма стабілізується BFRF3. Ангуляризація і всі наступні етапи дозрівання нуклеокапсиду відбуваються вже в BMRF1-ядрі [98].

Подальший процес пакування геномної ДНК опосередковується гексамерним терміназним комплексом, який за своєю суттю є складним каналом із діаметром близько 39Å. Кожен його мономер є гетеротримером з білків BGRF1/BDRF1, BALF3 і BFRF1A, які функціонують як один структурно-функціональний конгломерат/вузол з АТФазною, нуклеазною та ДНК-зв'язуючою активностями [104]. Первинне розпізнавання ДНК здійснюється в межах одного із *pac* сайтів. Це призводить до гідролізу фосфодієфірного зв'язку в цій ділянці і утворення вільного кінця ДНК, який позиціонується у напрямку каналу термінази. Подальші етапи пакування здійснюються за механізмом «з руки в руку», де взаємодія ДНК з кожним із шести вузлів призводить до активації АТФази, конформаційних перебудов і просування ДНК в напрямку нуклеокапсиду з одночасним зміщенням дуплексу в напрямку наступного вузла. Ці цикли повторюються до досягнення терміназою другого *pac* сайту. Після заповнення капсиду, канал, утворений порталом, закривається кепом, утвореним BVRF1. Слід зазначити, що лише у вершині, зайнятій порталом, цей білок має додаткову функцію, у решті вершин BVRF1 виконує лише структурну функцію.

Після формування зрілого нуклеокапсиду, він покидає ядро за характерним для герпесвірусів механізмом «envelopment-deenvelopment-reenvelopment», який опосередковується компонентами комплексу ядерного виходу (NEC від nuclear egress complex (анг.)) – білками BFRF1 і BFLF2 [105]. BFRF1, розміщуючись у ядерній мембрані, рекрутує компоненти системи ESCRT, що призводить до постійного процесу брунькування, який посилюється за рахунок попереднього ослаблення ядерної ламіни BGLF4. BFLF2 володіє спорідненістю як до нуклеокапсиду, так і до BFRF1 виконуючи таким чином адаптерну функцію, рекрутуючи зрілі частинки до місць локалізації ESCRT. Після утворення первинної вакуолі, яка містить один або кілька нуклеокапсидів ВЕБ і знаходиться в периплазматичному просторі, вона зливається з зовнішньою ядерною мембраною, вивільняючи вміст у цитоплазму.

На цьому етапі доведеною є роль білків BBRF2 і BSRF1 [106]. BSRF1 є тегументним білком, який пальмітується на N-кінці і володіє спорідненістю як до комплексу Гольджі, так і до BBRF2. Останній натомість володіє спорідненістю до BPLF1 і BcLF1. Припускається, що комплекс BSRF1/BBRF2, який до того ж демонструє схильність до олігомеризації, вистеляє цитоплазматичну сторону перетвореного у вірусну фабрику комплексу Гольджі, що дозволяє направити нуклеокапсиди до місця кінцевого дозрівання [106].

Після первинної взаємодії з майбутньою мембраною ініціалізується приєднання білків зовнішнього тегументу, які формують аморфну масу, розташовану в майбутньому віріоні з дистального відносно нуклеокапсиду боку [107]. Далі, за недослідженим нині механізмом, такий тегументизований нуклеокапсид брунькується у везикулярний просвіт формуючи зрілий віріон вірусу Епштейна-Барр з дещо зміщеним відносно центру нуклеокапсидом. Везикули цитоплазматичної фабрики, які містять достатню кількість сформованих віріонів ВЕБ транспортуються до плазмолемі, експлуатуючи для цього клітинні механізми екзоцитозу [108].

Таким чином, даний підрозділ структуровано підсумовує нині відомі експериментальні дані щодо життєвого циклу вірусу Епштейна-Барр. Значно розширена і доповнена версія цієї частини дисертаційної роботи була представлена як окрема самодостатня публікація в одному з фахових міжнародних наукових видань [109].

1.1.3 Хвороби, пов'язані з вірусом Епштейна-Барр

Інфекційний мононуклеоз є найпоширенішим проявом первинної інфекції ВЕБ. Здебільшого, у 35-50% випадків, він реєструється у підлітків, які до того не зустрічались з ВЕБ [110]. Відповідно, є характерним для країн із середнім чи високим рівнем економічного розвитку. Основними симптомами інфекційного мононуклеозу є біль у горлі, лихоманка та збільшення лімфатичних вузлів, що може

супроводжуватися атиповим лімфоцитозом. Причиною цих порушень є масова проліферація CD8⁺ Т-лімфоцитів, активованих значною кількістю уражених В-клітин [111]. Діагностування інфекційного мононуклеозу здійснюється здебільшого на основі морфологічних досліджень крові та детекції анти-ВЕБ антитіл у хворого. Через відсутність дієвих ліків, під час лікування цього захворювання застосовується симптоматична терапія.

Специфічним випадком первинної інфекції ВЕБ є посттрансплантаційне лімфопроліферативне захворювання або ПТЛЗ [8]. Воно виникає у випадку, коли серонегативний реципієнт отримує трансплантат від серопозитивного донора. Супресія імунної системи, необхідна для успішного трансплантування, призводить до набагато більш глибокого ослаблення контролю над пулом уражених В-лімфоцитів, присутніх у трансплантанті, ніж це є можливим у випадку імунокомпетентного господаря [112]. Наслідком відсутності селективного тиску на уражені ВЕБ В-клітини є стабілізація програми транскрипції κ . Це призводить до необмеженого поділу таких псевдотрансформованих В-лімфоцитів – формується картина схожа до класичної В-клітинної лімфоми, хоча справжньої ракової трансформації В-клітин не відбувається [113]. Ефективним способом лікування ПТЛЗ є зниження імуносупресії, хоча це рідко може бути застосовано без ризику відторгнення трансплантанту і є особливо небезпечним для реципієнтів із трансплантатами життєво важливих органів.

Крайнім проявом патології, спричиненої ВЕБ, є ракове переродження. Історично саме під час дослідження раку, якщо бути більш точним, то електронної мікроскопії лімфоми Беркїтта, і був відкритий вірус Епштейна-Барр [114]. Цей тип раку лімфоїдної системи є високоагресивним В-клітинним новоутворенням, яке ділять на три підтипи – ендемічна ЛБ, спорадична ЛБ та ЛБ, асоційована з імунодефіцитом. На молекулярному рівні причиною всіх форм ЛБ є транслокація між хромосомами 8 і 14, яка призводить до посиленої стимуляції гена МУС через наближення із енхансерами важких або легких ланцюгів імуноглобулінів.

Вважається, що транслокація відбуваються в зародковому центрі лімфатичних вузлів внаслідок порушення в процесі дозрівання афінності і, відповідно, соматичної гіпермутації. Підвищена частота саме ВЕБ-позитивної ЛБ є пов'язаною зі спричиненою цим патогеном дерегуляцією клітинного циклу ураженої клітини. Це включає значно знижену чутливість таких клітин до ендогенних сигналів апоптозу і може бути причиною їх покращеного потенціалу до виживання.

Лімфома Ходжкіна (ЛХ) є ще одним раковим захворюванням лімфоїдної системи, пов'язаним із ВЕБ. Морфологічно ЛХ розділяють на класичну і нодулярну [115]. Остання рідко є асоційованою з ВЕБ, натомість класична ЛХ в близько 30%-50% випадків є з цим вірусом пов'язаною. На відміну від лімфоми Беркітта, пухлини ЛХ складають лише на 1-2% безпосередньо з неопластичних клітин. Особливістю ЛХ є відсутність функціонально активного BCR, яка супроводжується конститутивною активацією сигнальних шляхів NF-κB і JAK/STAT [116]. Вважається, що саме на цьому рівні ВЕБ бере участь в канцерогенезі ЛХ [117]. Це пояснюється тим, що активність геному ВЕБ включає білки EBNA1, LMP1 і LMP2a/b (програма транскрипції II), де LMP1 є тим фактором, який прямо стимулює як NF-κB, так і JAK/STAT.

Виходячи з подвійної тропності ВЕБ як до лімфоїдної, так і до епітеліальної тканини, зрозумілим є також ряд карцином асоційованих з цим вірусом. Назофарингеальна карцинома є найбільш поширеним варіантом переродження епітеліальної тканини під дією ВЕБ. За оцінками Всесвітньої організації охорони здоров'я (ВООЗ), у 2018 році в усьому світі було зареєстровано близько 129 000 випадків НК [118]. НК зустрічається нерівномірно серед різних етнічних груп і найчастіше зустрічається кантонців Південного Китаю, а також корінного населення Південно-Східної Азії, Північної Африки та ескімосів в Арктиці. При чому 95% випадків, які реєструються в таких осередках є ВЕБ-позитивними [119]. У складі трансформованих клітин вірусний геном експресується відповідно до модифікованої програми латентності II, де окрім класичних LMPs та EBNA1 додатково

синтезується і повністю літичний BARF1. Таким чином, окрім забезпеченої LMP1 і LMP2s стимуляції сигнальних шляхів NF- κ B, ERK-MAPK, PI3K/AKT, JNK, JAK/STAT і RhoA, додатково забезпечується протидія опосередкованої макрофагами імунної відповіді [120].

1.1.4 Сучасний стан у розробці і пошуку антивірусних препаратів проти вірусу Епштейна-Барр

Відповідно до того, що вірус Епштейна-Барр є герпесвірусом, перші спроби пошуку терапії порушень, спричинених цим вірусом, включали ефективні щодо інших герпесвірусів аналоги нуклеозидів [121]. У ретроспективному дослідженні, проведеному у Франції, ВЕБ-серонегативна дослідна група отримувала валацикловір або валгацикловір протягом 3-6 місяців після трансплантації нирки чи підшлункової залози від ВЕБ-позитивного донора. У цій групі спостерігалось майже двократне зниження (43% проти 84%) частоти ПТЛЗ протягом перших 100 днів після трансплантації, порівняно з групою контролю. Хоча в більш довготривалій перспективі єдиною різницею між групами була відсутність повністю неопластичного переродження у дослідної групи порівняно із 6 випадками в контрольній [122].

Натомість ретроспективне дослідження проведене в Норвегії в період з 2002 по 2015 роки не підтвердило зниження вірусного навантаження після профілактики ганцикловіром реципієнтів печінки [123]. Тут дослідна група внутрішньовенно отримувала препарат протягом 21-38днів. Однак різниці у частці із зниженим вірусним навантаженням між цією групою і контрольною не спостерігалось. Таким чином, попри те, що існує очевидний вплив аналогів нуклеозидів на репродукцію ВЕБ, в тому числі підтверджений *in vitro*, з сучасних досліджень випливає ненадійність отриманого терапевтичного ефекту [121].

Були проведені кілька спроб пошуку і розробки препаратів, безпосередньо націлених на фактори ВЕБ [124]. Серед них найбільшого успіху досяг марибавір

[125]. В первинних дослідженнях його EC_{50} щодо репродукції ВЕБ у клітинах Akata в середньому становить 0,6 мкМ проти 25 мкМ для ацикловіру [126]. Через вплив на активність VMRF1 шляхом порушення фосфорилування цього фактору, було припущено механізм за якого цей препарат прямо інгібує BGLF4. Що, однак, було в подальшому спростовано через відсутність такої активності *in vitro* [127]. В наступному більш глобальному дослідженні було виявлено, що марибавір також знижує експресію gp125 і gp350, а також загалом пригнічує експресію вірусного геному. При цьому пригнічення будь-якого з вірусних факторів транскрипції не спостерігалось [128].

1.2 Вірус SARS-CoV-2

1.2.1 Будова віріону вірусу SARS-CoV-2. Структурні білки

Віріони SARS-CoV-2 мають близьку до сферичної форму із середнім діаметром 108 ± 8 нм [129]. Вони є складними і складаються із зовнішньої ліпідної мембрани, яка містить S-глікопротеїн, М-білок і Е-білок, а також внутрішнього нуклеокапсиду, який складається із N-білка і зв'язаної з ним вірусної РНК.

S-глікопротеїн є значно глікозильованим білком масою близько ~ 600 кДа, гомотримери якого формують славнозвісну «корону», характерну для представників родини Coronaviridae. За рахунок існування подібної до шарніру структури в ділянці поблизу розташування трансмембранної альфа спіралі цей поверхневий білок здатен до порівняно високого ступеня лабільності [130]. Вважається, що подібна особливість дозволяє SARS-CoV-2 більш успішно знаходити hACE2 (основний рецептор для SARS-CoV-2) на поверхні клітини в процесі первинної взаємодії. S-глікопротеїн, як структурний елемент є гетеродимером, який складається з субодиниць S1 розміром близько 700 а.з. і S2 розміром близько 600 а.з.. Субодиниці є результатом протеолітичного розщеплення білка-попередника довжиною ~ 1300 а.з. – S1 є його N-кінцем, а S2 – С-кінцем. В структурі S1 виділяють N-кінцевий домен (NTD від N-terminal domain (анг.)), рецепторзв'язуючий домен і два С-

кінцевих домени. S2, в свою чергу, складається з двох, так званих гептадних повторів – HR1 і HR2, центральної області, контактного домену та згаданої вище трансмембранної спіралі.

Окрім тримеру, рухомою частиною є також RBD. Цей домен здатен приймати дві конформації – «RBD down» та «RBD up». При чому область зв'язування з hACE2 є доступною лише у випадку конформації «RBD up». Вважається, що конформація «RBD down» є необхідною для кращого уникнення дії найбільш ефективних антитіл [131].

Е-білок має масу 8,5 кДа і складається з 75 а.з. Він формує одну трансмембранну спіраль, яка є схильною до гомотипових взаємодій і утворення пентамерної трансмембранної структури катіонного каналу [132]. На основі експериментальних даних відомо, що цей віропротеїн бере участь в брунькуванні та збірці віріонів SARS-CoV-2 [132].

М-глікопротеїн є найбільш поширеним в мембрані віріона. Він є О-глікопротеїном, має вагу ~25–30 кДа і складається з трьох трансмембранних спіралей. Цей білок вважається найбільш важливим в процесі ремоделювання ліпідних мембран під час формування віріону SARS-CoV-2 [133]. Крім того, він є здатним до взаємодії з N-білком, стабілізуючи, таким чином, нуклеокапсид [134].

N-білок є основним компонентом нуклеокапсиду. Він має вагу 46 кДа і складається з 419 амінокислот. Цей віропротеїн можна поділити на п'ять доменів: NTD, РНК-зв'язуючий домен, центральний лінкер, домен димеризації та CTD (останні два інколи об'єднують під назвою CTD). Він забезпечує утворення щільного нуклеопротеїнового комплексу, який в подальшому взаємодіє з М-глікопротеїном, ініціюючи, таким чином, утворення вірусної частинки [135].

1.2.2 Генوم вірусу SARS-CoV-2

Геном SARS-CoV-2 представлений кепованою та поліаденільованою несеgmentованою «+» РНК довжиною 29 903 нуклеотидів [136]. Він містить 12

ORF: ORF1a, ORF1b, S, M, N, E, ORF3a, ORF6, ORF7a, ORF7b, ORF8 і ORF10, які в складі геномної РНК є фланкованим 5'-UTR та 3'-UTR (UTR – untranslated region (анг.)) довжиною 265 п.о. і 239 п.о., відповідно. Дві третини геному займають ORF1a і ORF1b, які перекриваються і кодують поліпротеїни pp1a і pp1ab. Останні після протеолітичного дозрівання перетворюються у 16 неструктурних білків, компонентів здебільшого транскрипційно-реплікативної машинерії. Решту геному займають структурні білки та 6 додаткових білків [137].

Кінцеві UTR формують складні петлеподібні структури необхідні під час реплікації і транскрипції вірусного геному. Зокрема 5'-UTR складається з п'яти висококонсервативних «Stem-Loop» (SL) структур, серед яких SL1 є важливою в трансляції, SL2 – в реплікації, SL3 – у транскрипції, SL4 – у транскрипції і реплікації, а SL5 в упаковці геному [138]. 3'-UTR натомість складається з псевдовузлів BSL і SL1, гіперваріабельної області, висококонсервативного домену S2M і короткого однонуклеотидного мотиву ONM.

1.2.3 Життєвий цикл вірусу SARS-CoV-2

Найпершою подією життєвого циклу SARS-CoV-2 є взаємодія RBD-домену S-глікопротеїну у конформації «RBD up» з мембранною формою hACE2. Після цього, дії можуть розвиватись за двома сценаріями. За першого — поверхнева серинова протеаза TMPRSS2 розщеплює S2 субодиницю в сайті S2'. У випадку другого – асоціація S-глікопротеїну з hACE2 призводить до активації клатрин опосередкованого ендоцитозу [139]. В ендосомі протеолітичне розщеплення по S2' забезпечується неспецифічними ендопептидазами, здебільшого катепсином L.

На молекулярному рівні утворення комплексу RBD/hACE2 призводить до ряду молекулярних перебудов, в наслідок яких збільшується доступність сайту S2' до впливу протеаз. Розщеплення поліпептиду в цій ділянці запускає додатковий етап конформаційних перебудов, які призводять до повної дисоціації S1 від S2. Це звільняє до доступу розчинника гідрофобну поверхню S2, раніше захищену S1. Далі

фузогенний пептид S2 вивільняється у напрямку клітинної мембрани, що призводить до формування містка між ліпідними бішарами, їх наближення і злиття.

Після проникнення в клітину господаря геном SARS-CoV-2 позбавляється від N-білка — вільна кепована і поліаденільована РНК стає доступною до трансляції. Процес починається за класичним для еукаріот сценарієм: фактор ініціації eIF4E розпізнає m⁷-G на 5'-кінці вірусної геномної РНК, далі він залучає комплекс eIF4F (eIF4G + eIF4A), який через eIF4G рекрутує преініціаторний комплекс (PIC від preinitiation complex (анг.)). Після асоціації PIC починає рух у напрямку 3'-кінця скануючи 5'-UTR доки не буде знайденим старт кодон у правильному «сильному» контексті. Після цього залучається 60S-субодиниця рибосоми із формуванням повноцінного 80S рибонуклеопротеїнового конгломерату [140].

Однак на відміну від мРНК господаря, геном SARS-CoV-2 на 5'-UTR має складні тривимірні утвори. Активність eIF4A як гелікази дозволяє PIC все ж просуватись у напрямку 3'-кінця до AUG ORF1a, який знаходиться в основі SL5. Таким чином, процес ініціації трансляції за допомогою класичних механізмів є значно утрудненим для SARS-CoV-2.

Першими продуктами SARS-CoV-2 є поліпротеїни pp1a та pp1ab (pp – polyprotein). Однак в контексті трансляції слід звернути особливу увагу саме на найпершого серед перших – NSP1 (NSP – non-structural protein). Цей білок володіє спорідненістю до ділянки зв'язування мРНК малої субодиниці рибосоми, а також ендонуклеазною активністю. Він забезпечує зниження клітинного білкового синтезу вже з перших етапів життєвого циклу SARS-CoV-2 за рахунок інгібування ініціації трансляції [141].

В цей період ініціація транскрипції вірусної мРНК відбувається за рахунок альтернативних механізмів. Зокрема, згаданий вище AUG є асоційованим з структурними мотивами притаманними так званим «внутрішнім сайтам залучення рибосоми» або IRES (Internal Ribosome Entry Site). Подібні структури є призначеними для специфічних взаємодій типу РНК-РНК, які забезпечують

незалежне від кепу залучення всього преініціаторного комплексу. Особливу роль в цьому процесі має NSP1, N-кінець якого володіє спорідненістю до SL1. Таким чином, рибосоми, які містять цей білок, посилено рекрутуються до 5'-UTR РНК [141].

Після ініціації і збірки функціональної рибосоми розпочинається класичний процес елонгації поліпептидного ланцюга. Це продовжується допоки рибосома не досягне межі між ORF1a і ORF1b. ORF1b не містить класичного страт кодону у відповідному контексті, а також його відкрита рамка зчитування є зміщеною на 1 нуклеотид назад. Для його трансляції SARS-CoV-2 має механізм зсуву рамки зчитування на 1 нуклеотид назад – так званий «запрограмований -1 зсув рамки зчитування рибосоми» або PRF (programmed ribosomal frameshifting). Основним елементом PRF є специфічна ділянка вірусної РНК, яка знаходиться в межах, що відповідають NSP12 і називається «елемент запрограмованого -1 зсуву рамки зчитування» або PFSE (programmed -1 ribosomal frameshifting element). Вона складається з стеблеподібного атенюатора, «слизької» послідовності та масивної псевдовузлової структури стимулятора PRF нижче за течією із стом-кодоном в її межах. Таким чином, під час елонгації, коли рибосома досягає атенюатора вона уповільнюється. Після цього комплекс дестабілізується в межах «слизької» послідовності і майже повністю зупиняється перед стимулятором PRF. В цей момент рибосома має три можливих сценарії розвитку подій. За першого – нестабільний елонгаційний комплекс розпадається і трансляція зупиняється. За другого – використовуючи енергію макроергічних зв'язків рух транскрипційної машинерії все ж продовжується, що призводить до розпізнавання стоп кодону і термінації трансляції. У першому і другому випадках продуктом є pp1a. У третьому ж випадку дестабілізований рибосомальний комплекс може зробити «крок назад» і змістити рамку зчитування. Подальший рух рибосоми призводить до напрацювання pp1ab в ролі продукту.

Напрацювання інших структурних і додаткових білків відбувається з, утворених внаслідок специфічного типу транскрипції, субгеномних мРНК SARS-CoV-2. Цей процес залежить від роботи реплікативно-транскрипційного комплексу (РТК), який складається з неструктурних білків, утворених шляхом обмеженого протеолізу pp1a та pp1ab. В свою чергу протеолітичне розщеплення опосередковується цистеїновими протеазами NSP3 (PLpro) і NSP5 (3CLpro). Перша відповідає за напрацювання вільних форм NSP1, NSP2 та власне NSP3. Друга забезпечує функціональні форми решти факторів, серед яких і компоненти РТК. Власне через це 3CLpro вважається поряд із RBD однією з ключових цілей для розробки антивірусних препаратів. У функціонально активному стані це димер, кожен мономер якого має вагу 33 кДа і складається трьох доменів. Домени I (10–99а.з.) і II (100–182а.з.) мають подібну до протеази пікорнавірусу 3С шестиланцюгову антипаралельну β -складчасту структуру. Власне ця подібність є причиною назви цієї протеази. Активний центр, який є і ділянкою зв'язування субстрату знаходиться між доменами I і II. Він містить класичну каталітичну діаду з Н41 і С145, розміщену в центральній частині кишені. Вважається, що до складу покету також входить і одна молекула води, яка координується Н41 для нуклеофільної атаки. Домен III натомість складається п'яти антипаралельних α -спіралей, є глобулярним і відповідає за формування стабільних гомотипових взаємодій між мономерами.

Після напрацювання достатньої кількості pp1a, pp1ab і їх функціональних продуктів, запускаються активні процеси транскрипції і реплікації. Обидва є опосередкованими згаданим вище РТК, який є складним білковим комплексом і складається із білків NSP7, NSP8, NSP9, NSP10, NSP12, NSP13, NSP14 та NSP16. Серед перелічених власне РНК-залежною РНК-полімеразою (RdRp від RNA-dependent RNA polymerase (анг.)) є NSP12 (103 кДа). Білки NSP7-NSP8 функціонують спільно, формуючи гетерогексамерний комплекс РНК-праймази, NSP13 виконує функцію РНК-гелікази, а NSP9 є РНК-зв'язуючим білком

спорідненим здебільшого до 5'-UTR і 3'-UTR [142].

NSP14 є біфункціональним фактором (60 кДа): він володіє як екзонуклеазною активністю у напрямку 3'→5', так і є гуанін-N7 метилтрансферазою [143]. Його екзонуклеазний N-кінцевий домен функціонує у тандемі з рештою реплікативно-транскрипційної машинерії задля позбавлення новосинтезованого ланцюга РНК від неправильно спарених нуклеоснов. Останнє явище особливо характерне для РНК-вмісних вірусів частота мутацій в яких здебільшого становить одну на 10^3 - 10^6 коректних нуклеотидів [144]. SAM-залежний гуанін-N7 метилтрансферазний домен знаходиться на С-кінці NSP14 і є пов'язаним з екзонуклеазою через довгий шарнір, що дозволяє цим двом функціонувати відносно незалежно. Він є необхідним для метилювання 5'-кінцевого гуаніну новосинтезованого ланцюга РНК по N7, що є одним з етапів утворення кепу. Доповнює процес кепування білок NSP16, який є SAM-залежною 2'-O-метилтрансферазою [145]. При цьому NSP10 є кофактором, необхідним для стабілізації активного центру NSP16 [146].

Після напрацювання всіх необхідних для РТК білків, відбувається його збірка на 3'-UTR геномної РНК. Праймаза NSP7-NSP8 синтезує короткий РНК-праймер, і розпочинається напрацювання вірусної «-» РНК. Характерним для РНК-вмісних вірусів є невіддільність реплікації і транскрипції – РНК є як матрицею для напрацювання білків, так і власне геномом. Однак для SARS-CoV-2 ці процеси, все ж, можна розділити.

5'-кінець кожного ORF SARS-CoV-2 має короткі регуляторні послідовності (transcription regulatory sequence) або TRS. Спільний TRS розташований в районі 5'-UTR називають лідерним або TRS-L, натомість всі інші – TRS-B від «body». Коли машинерія РТК досягає одного з TRS-B, новосинтезований ланцюг містить антисиквенс до неї. Через те, що деякі із компонентів РТК володіють спорідненістю до кепу, 5'-кінець геномної РНК є наближеним до реплісоми. Відповідно, з певною ймовірністю, новосинтезований ланцюг може бути спареним з TRS-L. Це

призводить до зміни матриці синтезу на послідовність поблизу TRS-L. Таким чином, формуються дев'ять вкорочених версій вірусного геному, які називають субгеномними мРНК або sgRNA (subgenomic RNA). Отже, у разі якщо РТК не зазнає заміни матриці зчитування протягом всього процесу синтезу РНК, процес можна умовно назвати реплікацією. У разі ж формування sgRNA, їх «-» ланцюги виступають в ролі матриці для напрацювання відповідних «+» sgRNA. Кожна з них, крім тієї, що містить лише ORF N є полістронною матрицею для трансляції, однак, через наявність стоп-кодону в кінці кожної ORF, далі за першу ORF транскрипційна машинерія не рухається. Процес можна розглядати специфічною формою трансляції.

Загалом, після інфікування, первинного напрацювання білків РТК і sgRNA, більшість подальших синтетичних процесів відбуваються на ремодельованих SARS-CoV-2 мембранних підструктурах, які формують фабрики вірусної репродукції. Часто цей мембранний конгломерат називають органелою реплікації. Для даного патогену він є складним утвором, більше схожим на сітчасту формацію, яка включає двомембранні везикули, поєднані між собою і ендоплазматичною мембраною конекторними перетяжками. Внутрішнє середовище органели реплікації, при цьому, контактує з цитозолем через пори діаметром 2–3 нм, які перетинають обидві мембрани. Хоча детальний механізм утворення цих вірусних фабрик не є дослідженим, все ж відомо, що основну роль в цьому процесі виконують NSP3, NSP4 і NSP6 [147]. Новостворені геномні і sgRNA виходять у цитоплазму через пори, де sgRNA транслуються, а геномна РНК використовується для збірки вірусного потомства.

Процес збірки віріонів починається з накопичення структурних білків S, M, N і E, серед яких найважливішим можна вважати M. Це трансмембранний білок, локалізований у ER. Він може бути представлений у вигляді двох конформерів: компактного M-compact і витягнутого M-long. Він володіє спорідненістю до всіх інших структурних білків та є здатним змінювати поверхневий натяг ліпідних

бішарів. Після обгортання білком N вірусний генوم залучається до ділянок локалізації M і рекрутованих ним S і E. Взаємодія рибонуклеопротеїнового комплексу з насиченою M білком ліпідною мембраною призводить до поступового обгортання нуклеокапсиду останньою і подальшого формування повноцінного віріону [148]. Зараз вважається, що він транспортується у складі везикули до комплексу Гольджі. Звідки за класичним для позаклітинного вантажу шляхом потрапляє до транс-мережі Гольджі і далі в позаклітинний простір.

1.2.4 Сучасний стан у розробці і пошуку антивірусних препаратів проти вірусу SARS-CoV-2

Попри те, що вакцинація як спосіб боротьби з SARS-CoV-2 виявилась достатньо ефективним методом протидії його поширенню, все ж, вона не може бути використаною при вже наявному COVID-19. Лише безпосередньо антивірусні препарати можуть послабити або ж повністю нівелювати вірусну репродукцію. Виходячи з внеску у вірусну репродукцію, наявних суміжних досліджень та наявних структурних знань, перспективними цілями для пошуку і розробки селективних інгібіторів, в майбутньому – етіотропних антивірусних препаратів, вважаються 3CLpro, PLpro, RdRp і RBD S-глікопротеїну [149, 150, 151, 152].

Нірматрелвір (PF-07321332) є розробленим компанією Pfizer ковалентним інгібітором 3CLpro. Структурно він є пептидоміметиком із нітрильною групою в ролі «бойової» [153]. Його експериментально досліджена константа інгібування 3CLpro є рівною 3,11 нМ, при цьому EC_{50} проти SARS-CoV-2 становить 74,5 нМ [154, 155]. Окрім доведеної активності цей інгібітор володіє високою швидкістю метаболізму, оскільки є субстратом для цитохрому CYP3A4 [156]. Цей недолік був вирішений розробником застосуванням ритонавіру в ролі інгібітора системи цитохромів P450. Таким чином, компанією Pfizer був створений комбінований лікарський препарат Паксловід. Він виявився ефективним для протидії ускладненням, спричиненим COVID-19, у пацієнтів старше 65 років, зменшивши

рівень госпіталізації на 73%. При цьому Паксловід ніяким чином не вплинув на цей показник для молодших пацієнтів. Попри це, цей препарат нині є єдиним розробленим етіотропним засобом, рекомендованим ВОЗ [157], хоча і лише для лікування людей похилого віку з високим ризиком розвитку ускладнень.

Порівняно з 3CLpro, інгібітори PLpro є лише в розробці і здебільшого містять акцептори Міхаеля в ролі «бойової» групи. Зокрема Sanders et al. на основі конкурентного інгібітора GRL0617 розробили сполуку 11 із IC_{50} щодо PLpro в 0,094 мкМ [158]. Подальше дослідження на модельній системі інфекції SARS-CoV-2 (Vero E6) дозволили визначити, що при EC_{50} 1,1 мкМ CC_{50} цього похідного є більшим за 30 мкМ [159]. Інша дослідна група, базуючись на вищеописаних даних та на власних розробках, представила сполуку 2, яка характеризується IC_{50} 0,46 мкМ і CC_{50} щодо культури клітин HEK293T ~180 мкМ [160].

Порівняно з інгібіторами протеаз, інгібітори RdRp здебільшого представлені добре відомими ліками, активність яких проти SARS-CoV-2 була підтверджена в експериментах з перепрофілювання ліків [161]. Ремдесивір був першим (22 жовтня 2020 р.) затвердженим FDA лікарським препаратом для лікування важкого COVID-19 [162]. За біохімічною природою він є аналогом аденіну із нітрильною групою на C1' рибози. Ремдесивір відноситься до класу проліків і набуває своєї активної форми вже безпосередньо в клітині. Наслідком цього процесу є монофосфатний міметик до аденіну, який фосфорилується до трифосфату і може бути включений до новосинтезованого ланцюга РНК RdRp, менш розбірливою до субстрату, термінуючи подовження ланцюга РНК [163]. На модельній системі інфекції SARS-CoV-2 (Vero E6) EC_{50} цього лікарського препарату становить 0,77 мкМ, а $CC_{50} = >100$ мкМ. Однак через погану стабільність при пероральному вживанні ремдесивір здебільшого рекомендують застосовувати внутрішньовенно, що значно обмежує його потенціал при лікуванні SARS-CoV-2 та інших вірусних захворювань [164].

Фавіпіравір є ще одним загальновідомим інгібітором RdRp, який був рекомендований до лікування COVID-19 після досліджень з перепрофілювання ліків [165]. Він не є здатним зупиняти реплікативну машинерію RdRp, хоча масове його вбудування у дочірні ланцюги РНК призводить до нежиттєздатного потомства [166]. Такий механізм дії є причиною достатньо великої рекомендованої добової дози цього препарату в 3200 мг при EC_{50} 61.88 мкМ і CC_{50} більше 400 мкМ (Vero E6) [167, 168]. На основі рандомізованого дослідження, яке виявило значне пришвидшення одужання (3 дні проти 5 днів, $p = 0,03$), цей препарат був рекомендований до використання [165]. Хоча слід зазначити, що наступні масштабні дослідження з визначення оптимальної дози виявили, що доза менше 45 мг/кг/добу асоціювалась з подовженим COVID-19 [167]. Це ставить під деякий сумнів отримані раніше результати, а також кидає тінь на основну перевагу фавіпіравіру – можливість зручного перорального застосування хворим в домашніх умовах.

Ще однією перспективною, однак специфічною для SARS-CoV-2, ціллю для пошуку і розробки препаратів є S-глікопротеїн і зокрема його рецепторний домен. Власне реконвалісцентна плазма містить антитіла, здебільшого направлені на RBD і NTD S1 [12]. Поєднуючи сучасні підходи на основі генеративного штучного інтелекту і класичні методи *in vitro*, міжнародна дослідна група, до якої входили в тому числі науковці з України, *de novo* розробила ряд інгібіторів S-глікопротеїну [169]. Серед сконструйованих і синтезованих речовин GEN725 і GEN727 продемонстрували у експериментах як з псевдотипованим лентивірусом, так і з живим SARS-CoV-2 IC_{50} на рівні 18,7 і 2,8 мкМ, відповідно. Подальші дослідження більш перспективної GEN727 показали, що її IC_{50} для по-різному псевдотипованих лентивірусів коливалось від 0,7 до 2,8 мкМ (Вікторія, Альфа, Бета, Гамма, Дельта, Оміврон). При цьому IC_{50} у випадку модельної системи на клітинах Vero-CCL-81 для всіх варіантів, окрім Гамма і Оміврон, було менше 50 мкМ. Спроба молекулярного докінгу з подальшою симуляцією молекулярної динаміки дозволила

ідентифікувати ліпід-зв'язуючу кишеню в основі RBD, як ймовірну ціль дії GEN725 і GEN727. Ця кишеня є присутньою у більшості коронавірусів людини. Зв'язування тут ліпідів плазмолемі призводить до стабілізації RBD в конформації «RBD down» [170].

Інше достатньо масштабне дослідження базувалось на інгібіторів SARS-CoV-1 VE607 [171, 172]. Дослідники показали, що як рецематна суміш, так і кожен із енантіомерів окремо здатен інгібувати SARS-CoV-2 D614G з IC_{50} 2,42 мкМ (Vero E6). При цьому цитотоксичності щодо 293T-ACE2 і Vero-E6 на концентраціях до 100 мкМ не спостерігалось. Подальше розрахункове дослідження, підтверджене експериментально, дозволило визначити механізм дії VE607, який полягає у прямому зв'язуванні з ділянкою взаємодії RBD з hACE2. У дослідженнях *in vivo* було показано зниження вірусного навантаження у тканинах легень в 37 раз. Однак попри те, що дослідна група демонструвала значно меншу втрату ваги, всі тварини померли від ускладнень спричинених SARS-CoV-2.

З даних наведених у цьому підрозділі можна зрозуміти, що здебільшого затвердженими для застосування проти COVID-19 препаратами є добре відомі інгібітори RdRp широкого спектру дії. Кожен з них поряд з цим володіє рядом недоліків, серед яких висока ціна, низька біодоступність, сумнівна ефективність. Єдиним препаратом допущеним до використання і розробленим саме проти SARS-CoV-2 є Паксловід, хоча його ефективність і є доведеною лише для старших людей. При цьому існує достатньо широкий ряд досліджень і розробок у полі антикоронавірусної терапії, які демонструють багатообіцяючі результати у інгібуванні різних ключових факторів життєвого циклу SARS-CoV-2. Це дозволяє більш позитивно поглянути на майбутнє лікування COVID-19, яке, однак, затінюється статистикою класичного процесу розробки ліків, де він займає період в близько 10 років і коштує близько 2 мільярдів доларів США [173]. Таким чином, зараз і на багато років вперед розробка і пошук антикоронавірусних ліків буде однією із нагальних проблем сучасної системи охорони здоров'я.

РОЗДІЛ 2

СТРУКТУРО ОРІЄНТОВАНИЙ ДИЗАЙН ЛІКІВ

Перші ідеї структуро орієнтованого дизайну ліків (structure-based drug design) походять з періоду 1950-1960-х років, коли вперше були отримані дані тривимірної структури білків та проведено спроби зрозуміти зв'язок між «координатами і функціонуванням» [174]. Вже на рівні визначення міоглобіну і гемоглобіну, а також порушень, які супроводжували серповидну анемію, стало зрозуміло, що найбільш перспективним способом відійти від неефективного, наддорогого і повільного сліпого скринінгу природних або синтезованих сполук є розробка і пошук хімічних структур, націлених безпосередньо на білкову ціль.

Починаючи з того часу і донині найпершим етапом у структуро орієнтованому дизайні ліків є відбір мішені. Класично прийнятною ціллю вважається добре описаний і досліджений білок, який володіє ключовою функцією в межах певного патологічного процесу та має надійно визначену за допомогою одного із експериментальних методів тривимірну структуру. До цих методів відносять метод рентгеноструктурного аналізу, ядерного магнітного резонансу і криоелектронної мікроскопії [175]. Здебільшого відповідні дані є представленими у всесвітньо доступних базах даних типу Protein Data Bank, однак частим є явище, коли через складність процедури, фізичні обмеження або ж брак ресурсів даних тривимірної структури цілі немає. В таких випадках застосовуються розрахункові підходи генерації третинної структури мішені.

Серед останніх найбільш поширеним підходом є моделювання на основі гомології [176]. До розробки AlphaFold2 методом, який найкраще відповідав балансу між продуктивністю і надійністю, був підхід на основі шаблонів [177]. Найчастіше він передбачає п'ять етапів: пошук шаблонів, порівняння сиквенсу шаблону і цільового білка для визначення ділянок з найбільшою відповідністю, створення початкової структури цілі та конструювання ділянок, які не відповідали

ні одному із використаних шаблонів на основі загальноприйнятих законів побудови білків. На заключному етапі проводиться узгодження всіх змодельованих ділянок між собою. Поряд зі своєю відносною швидкістю, простотою і надійністю, якщо послідовність найближчого шаблону є відповідною послідовності цільового білка менше ніж на 30%, для пошуку додаткових шаблонів використовуються набагато менш надійні алгоритми, до прикладу, прихована модель Маркова [178].

Після більше ніж двадцяти років домінування, інструменти на основі гомологічного моделювання поступово починають поступатись точністю інструментам, які базуються на глибоко навчених нейромережах [179]. Особливий прорив відбувся після представлення у 2020му році компанією Google DeepMind нового інструменту для прогнозування тривимірної структури білка – AlphaFold2 [180]. В липні того року AlphaFold2 під час участі в конкурсі CASP (Critical Assessment of protein Structure Prediction) досяг середньоквадратичного відхилення (RMSD від root mean square deviation (анг.)) карбонового скелету у 0,9 Å відносно еталонної структури [181, 182]. Для порівняння, наступний найкращий метод мав RMSD в 2,8 Å [183].

В основі AlphaFold2 лежить один із найбільш ефективних у зіставленні даних типів нейромереж – так званий Трансформер [184]. Він є побудованим на основі концепції «уваги» і у випадку AlphaFold2 є навченим співставляти між собою амінокислотну послідовність цілі, споріднених їй білків, відповідні дані тривимірної структури, а також результат їх множинного вирівнювання [177]. У 2023 році у партнерстві між DeepMind і EMBL-EBI (EMBL-European Bioinformatics Institute) було створено базу спрогнозованих на основі послідовності даних тривимірної структури для більше ніж 200млн білків [185].

На основі джерела вихідної інформації для пошуку і розробки ліганду з потенційною біологічною активністю СОДЛ можна поділити на віртуальний скринінг вже існуючих бібліотек хімічних сполук і *de novo* дизайн безпосередньо активного ліганду на основі виключно структури цілі. Віртуальний скринінг, в свою

чергу може бути розділеним на високопродуктивний віртуальний скринінг або HTVS (High-throughput virtual screening) та фармакофорний пошук. Останній володіє рядом недоліків, пов'язаних із складністю оцінки отриманих результатів і не є представленим в цій роботі [186]. Натомість високопродуктивний віртуальний скринінг і молекулярний докінг є широко застосовуваними у сфері відкриття ліків методами [187].

Молекулярний докінг у найпримітивнішому уявленні є розрахунковим процесом стикування молекули ліганду із певною ділянкою рецептора [187]. У ролі рецептора здебільшого виступають білки, зрідка РНК або ДНК, протеїнові чи нуклеопротеїнові конгломерати. У ролі ліганду здебільшого виступають низькомолекулярні сполуки, хоча часто лігандами також можуть бути пептиди і білки [188]. Метою молекулярного докінгу є пошук найбільш енергетично вигідної пози ліганду відносно рецептора. Це є першим етапом у визначенні потенціалу ліганду до зв'язування з рецептором і оцінки його можливостей виступати модулятором активності конкретної цілі [189].

За ступенем рухливості рецептору і ліганду молекулярний докінг поділяють на гнучкий, напівгнучкий і жорсткий. У першому випадку кожен із компонентів може вільно рухатись і змінювати конформації – таким чином алгоритм може найточніше передбачити режим стикування і оцінити його енергетичний вигаш. Однак ступінь затрат обчислювальних ресурсів є непомірно високим – рух кожної з точок системи повинен бути обрахований окремо, а також разом із урахуванням оточуючих змін. Це значно обмежує використання гнучкого докінгу [187]. Натомість, жорсткий молекулярний докінг передбачає, що як ліганд, так і рецептор є повністю жорсткими молекулярними структурами, нездатними до конформаційних змін. Процес стикування тут передбачає виключно пошук пози ліганду із найнижчою умовною енергією утвореного таким чином комплексу. Жорсткий молекулярний докінг є відносно дешевим зі сторони затрат обчислювальних ресурсів, хоча і найменш точним. Найчастіше його застосовують для стикування

великих або дуже великих лігандів і рецепторів.

Компромiсом між точністю і затратами є напівгнучний молекулярний докінг. Цей тип стикування передбачає, що ліганд є повністю рухливою молекулярною конструкцією, здатною до вільної зміни власної конформації. Натомість, рецептор є або повністю жорстким, або ж має лише певні визначені ділянки здатні до руху. Напівгнучний молекулярний докінг є найчастіше застосовуваним для рутинного визначення оптимальної пози і конформації [190].

На рівні розрахунків кожен із описаних вище підходів базується на двох складових: алгоритмі пошуку пози/конформації та функції оцінки. Алгоритми пошуку є необхідними для генерації різних поз ліганду в межах рецептора. У випадку жорсткого молекулярного докінгу ця складова обмежується лише генерацією різних можливих режимів стикування жорсткого ліганду та жорсткого рецептору. При напівгнучному і гнучному молекулярного докінгу окрім цього також генеруються різні конформації ліганду і рецептора, відповідно. Найбільш поширеними є алгоритми на основі симуляції молекулярної динаміки і стохастичного пошуку [191]. Метод симуляції молекулярної динаміки більш детально буде розглянутий нижче. У випадку молекулярного докінгу лише слід зазначити, що використовується загальноприйняте рівняння руху Ньютона, де кожен атом описується як частинка, яка піддається законам класичної механіки. Останні імітують сили, присутні на молекулярному рівні матерії, ігноруючи квантові ефекти. Застосування цього методу дозволяє, шляхом перебору конформацій і поз, знайти положення ліганду з найнижчою умовною енергією комплексоутворення. Найпоширенішим серед методів стохастичного пошуку натомість є також відомий алгоритм Монте-Карло [192]. Він передбачає перебір конформацій і положень ліганду в межах рецептору за рахунок генерації невеликих змін в структурі. Задля зниження конформаційного простору, який є достатньо значним навіть у відносно невеликих лігандів, нині існує велика кількість модифікацій цього методу [193].

Для того, щоб розуміти успішність тієї чи іншої конформації/пози, після кожної ітерації проводиться оцінка потенційної стабільності отриманого комплексу. Функції оцінки здебільшого розділяють на методи прямої оцінки фізики комплексу, емпіричні функції оцінки, «Knowledge-Based» функції та методи на основі машинного навчання [194]. Методи прямої оцінки фізики комплексу в своїй основі мають відкриті нині закони, які описують сили Ван-дер-Ваальса (потенціал Ленард-Джонса) і електростатичні взаємодії (закон Кулона) як основні сили міжмолекулярної взаємодії. Часто до цих двох компонент додають ще внески сольватації явного або неявного розчинника, а також ентропії. Якість подібних обчислень дуже сильно залежить від якості використаного силового поля, яке описує експериментально визначені змінні для кожного із рівнянь в кожному конкретному випадку. Для подолання цього можуть бути використані методи ще більш глибокої фізики, які описують міжмолекулярні ефекти на рівні квантової механіки. Однак якщо застосування силових полів є відносно затратним зі сторони обчислень підходом то обчислення квантової механіки, навіть спрощених її варіацій, є обмеженим випадками, коли точність власне і є метою розрахунків.

Спрощеним варіантом прямої оцінки енергії комплексоутворення є її обчислення на основі емпірично визначених і здебільшого підігнаних під конкретний випадок законів. Для покращення результатів відносна енергія комплексоутворення розбивається на більшу кількість складових, де, окрім енергії Ван-дер-Ваальса, враховується енергія водневих зв'язків, стеричні ефекти і навіть таке абстрактне поняття як гідрофобні зв'язки. Точність емпіричних функцій оцінки є залежною від набору даних на основі якого вони були отримані і способу обробки цих даних, що робить їх загалом найменш надійними, однак не позбавленими переваг – спрощені розрахунки потребують набагато менше часу.

Балансом між емпіричними і фізичними функціями оцінки вважаються «Knowledge-Based» скорингові функції. В своїй основі вони мають статистичний підхід Больцмана за допомогою якого є можливим описати частоту розташування

певних атомів один відносно одного як залежний від відстані енергетичний потенціал. Такий підхід також є значно залежним від набору даних, хоча і може мати в основі набагато більшу кількість даних, порівняно з емпіричними методами і, таким чином, бути більш надійним. Найбільш сучасними і просунутими функціями оцінки є методи, в основі яких лежить машинне навчання. Де найпоширенішими є метод опорних векторів, метод випадкових лісів, штучні нейронні мережі і штучні нейронні мережі з глибоким навчанням [194]. Машинне навчання також базується на великій кількості вхідних даних, однак кореляції, які дозволяють обрахувати певні цільові характеристики, визначаються не безпосередньо дослідником, а безпосередньо сіткою об'єктів (нейронів) об'єднаних між собою зв'язками. Власне конфігурація/вага цих зв'язків і визначається під час процесу навчання. Таким чином, нейронна мережа попри те, що концептуально є результатом програмування, в кінцевому стані може розглядатись як конструкт відмінний від результату роботи дослідника і безпосередньо адаптований набором вхідних даних для виконання певної специфічної функції. Хоча методи на основі нейромереж і нейромереж із глибоким навчанням потребують значного об'єму ресурсів лише в період їх навчання, об'єм необхідних даних, потреба у їх стандартизації і іншій підготовці значно уповільнює процес їх розвитку і впровадження.

Віртуальний скринінг є найбільш поширеним способом застосування молекулярного докінгу [195]. Його основною відмінною рисою є використання бібліотек хімічних структур розміром в тисячі і мільйони молекул. Віртуальний скринінг можна вважати розрахунковим варіантом високопродуктивного скринінгу, де останній має такі обмеження як потреба в синтезі великої кількості сполук, їх очистці і ідентифікації [196]. Віртуальний скринінг дозволяє уникнути пов'язаних ресурсних і часових обмежень, покладаючись задля уточнення результатів на вже отримані експериментальні дані, інтегровані у функції оцінки молекулярного докінгу, як це було описано вище.

Через часто спрощені методи оцінки, як от застосування емпіричних скорингових функцій, невелика кінцева кількість лігандів і їх комплексів додатково оцінюється за допомогою методу симуляції молекулярної динаміки [197]. Метод симуляції молекулярної динаміки в своїй основі має класичну механіку, де кожен з атомів системи є об'єктом, рух якого описується другим законом Ньютона [198]. Сумарна сила, яка діє на кожен із атомів системи обчислюється на основі ковалентних і нековалентних енергетичних компонент. Перша враховується за допомогою класичного рівняння руху пружинного гармонічного маятника, де ковалентний зв'язок виступає в ролі нерозривної пружини з двох сторін від якої знаходяться два умовні атоми. Нековалентні взаємодії натомість описується як потенціали, що по різному змінюються відповідно до відстані. Таким чином, враховуються сили Ван-дер-Ваальса (потенціал Леннарда-Джонса), іонні взаємодії (закон Кулона), водневі взаємодії (складна комбінація Леннарда-Джонса і закону Кулона) та внесок міжатомних і двогранних кутів [199, 200]. Маючи сумарну силу, яка діє на конкретний атом і вектор її дії, можна отримати положення цього атома в наступний момент часу. Безпосередні дані щодо конкретних змінних, які стосуються конкретного атома в конкретному оточенні формують згадані вище силові поля. Вони здебільшого є результатом експериментального визначення певного параметру, до прикладу, довжини ковалентного зв'язку. Однак останнім часом, із розвитком методів квантової механіки і обчислювальної техніки, частим явищем стало уточнення цих даних за допомогою глибоких обрахунків [201].

Задля відповідності часовим рамкам молекулярних рухів одна ітерація описує рух системи в діапазоні 10^{-12} - 10^{-15} с. Нині вважається, що для того щоб отримати достатню кількість даних для оцінки ліганд/рецепторного комплексу необхідно провести симуляцію молекулярної динаміки цього комплексу у явному розчиннику протягом більше ніж 100 нс [198]. У випадку ж досліджень, які стосуються конформаційних перебудов білків або ж їх фолдингу, цей часовий діапазон може вирости до мілісекунд і довше. Враховуючи, що типові молекулярні системи мають

розмір в 100-500 тис. атомів, кожен з яких описується складною системою математичних рівнянь, лише протягом останнього десятиліття стало можливим проводити симуляції молекулярних систем необхідного розміру протягом достатнього часового проміжку [202].

Існує багато способів і методів аналізу результатів симуляції молекулярної динаміки. В контексті ліганд/рецепторних систем найчастіше використовується обрахунок графіків RMSD, RMSF (root mean square fluctuation), залежності відстані між лігандом і певною амінокислотою чи групою амінокислот і подібні [203]. Однак найбільш сучасним і найбільш інформативним способом відобразити стабільність комплексу є обрахунок вільної енергії зв'язування методами MM/PBSA і MM/GBSA [204].

Ці методи обчислюють енергетичний вигравш системи внаслідок взаємодії ліганду і рецептора [205]. Для цього спочатку обчислюється загальна потенційна енергія комплексу, а потім окремо енергія рецептора і ліганду. Вільна енергія зв'язування ліганду, таким чином, дорівнює різниці потенційної енергії комплексу і суми енергій ліганду та рецептора. Ці підходи потребують лише кінцевого фрейму траєкторії для оцінки, у зв'язку з чим їх часто називають «End-Point»-методами. Хоча в рутинних обрахунках для покращення точності здебільшого використовується не кінцеві координати ліганду і рецептора, а певна вибірка фреймів траєкторії симуляції їх молекулярної динаміки.

З математичної і фізичної точки зору методи MM/PBSA і MM/GBSA базуються на ідентичних до методу молекулярної механіки принципах. Основною відмінністю є спосіб обрахунку внеску сольватації системи. Через велику потребу у ресурсах для обрахунку явного розчинника ці методи передбачають неявний обрахунок внеску розчинника в енергію комплексоутворення. Цей обрахунок відбувається шляхом вирішення рівняння Пуассона-Больцмана (MM/PBSA) або на основі узагальненої моделі Борна (MM/GBSA) для полярної компоненти і площі поверхні, доступної розчиннику для неполярної компоненти [206]. Втрата ж

ступенів свободи ліганду і рецептора, яка супроводжує комплексоутворення, обраховується додатковим членом, який за своєю фізичною суттю є ентропійним штрафом.

Іншим способом розробки нових лікарських сполук на основі структури цілі є *de novo* дизайн ліків [207]. Цей метод передбачає безпосереднє створення ліганду з заданими властивостями: розміщенням, розміром, афінністю, іншими [208]. Найперші алгоритми під час побудови ліганду в межах активного центру цілі оперували атомами як елементарним конструктивним блоком [209]. Однак комбінаторний вибух та складність у оцінці проміжних результатів призвели до зміщення уваги дослідників у напрямку спрощених методів на основі фрагментів [210]. При цьому, спільним для всіх алгоритмів є застосування ітераційного підходу, де основна різниця між ними є зосередженою в межах побудови конкретної ітерації. На цьому рівні нині найбільш поширеними є еволюційні алгоритми, «Growing-Linking-Merging» алгоритм і алгоритми на основі методу Монте-Карло. Останніми роками також активно почали досліджуватись підходи на основі генеративного штучного інтелекту, хоча наразі вони не зазнали широкого практичного застосування [211, 212].

Найбільш поширеним можна вважати «Growing-Linking-Merging» алгоритм. Він передбачає стикування молекулярних фрагментів, джерелом яких можуть бути як бібліотеки наявних нині хімічних сполук, так і бібліотеки створені за допомогою комбінаторної хімії. Найбільшу з останніх надала українська компанія ТОВ «Енамін» у 2020 році [213]. Вона налічує близько 29 мільярдів структур. Початкове «ядро» або «ядра» – молекулярні фрагменти розташовані в ключових частинах активної ділянки цілі – в подальшому виступають в ролі вихідних структур, які, шляхом перебору здатних до кон'югації молекулярних фрагментів, «ростуть» до моменту виконання певної умови поставленої дослідником. До прикладу, до досягнення певного рівня афінності або необхідного ступеня перекриття площі цілі. Окрім необхідних значних затрат обчислювальних потужностей цей підхід, як і всі

підходи *de novo* дизайну ліків, володіє значним недоліком, пов'язаним із низькою синтетичною доступністю отриманих молекулярних конструкцій [214]. Відповідно, дослідник, який контролює процес створення нової молекулярної конструкції, повинен бути залученим не лише на рівні кінцевого результату, а і на рівні власне процесу автоматизованого *de novo* дизайну [215].

РОЗДІЛ 3

МЕТОДОЛОГІЯ РОБОТИ

3.1 Джерела вхідних даних і їх підготовка

3.1.1 Отримання статистичних даних

У дослідженнях пов'язаних з аналізом статистичних даних щодо поширення різних варіантів SARS-CoV-2 як основне джерело статистично значимої інформації були використані ресурси CoVariants, The Johns Hopkins Coronavirus Resource Center та Our World in Data [216, 217, 218]. Кожна вибрана країна аналізувалась за чотирма параметрами: динамікою поширення різних варіантів SARS-CoV-2 та їх взаємодії між собою, динамікою епідемічної ситуації в країні, швидкістю та особливостями вакцинації населення, а також антиковідними заходами. Останнє було оцінено на основі змін в мобільності населення, для чого використовувалися дані Google COVID-19 Community Mobility Reports [219].

3.1.2 Ліганди і їх бібліотеки

У скринінгових дослідженнях як джерело вхідних даних потенційних лігандів використовувались бібліотека затверджених FDA лікарських засобів. Сет даних був отриманий з ресурсу e-Drug3D за червень 2024 року [220]. Для проведення безпосередньо скринінгу на основі вхідних даних, отриманих у форматі SDF за допомогою RDKit 2022.9.5 [221], генерувались відповідні 3D-бібліотеки, які містили всі можливі стереоізомери всіх вхідних сполук. На цьому етапі враховувались всі стереоцентри, а структури, які перекривались були видалені. Далі за допомогою Open Vabel GUI [222] ці 3D-бібліотеки конвертувались у формат PDBQT з одночасним розділенням їх на окремі файли, кожен з яких містив одну тривимірну структуру. Тобто різні ізомери однієї речовини були представлені різними файлами. На цьому етапі, відповідно до умов експерименту, також відбувалась фільтрація вхідних даних за необхідними параметрами, до прикладу за молекулярною масою.

3.1.3 Цілі і їх підготовка

У випадку досліджень, які включали білки в ролі цільових векторів, дані тривимірної структури останніх були отримані з Protein Data Bank (PDB). Експериментальні дані відбирались на основі звіту валідації наданого PDB. Особлива увага зверталась на роздільну здатність ($>3 \text{ \AA}$), кількість амінокислот, що мають неприродне співвідношення кутів ϕ і ψ , а також ступінь повноцінності кінцевої структури. Під час безпосередньої підготовки до подальших етапів з вхідного файду координат видалялись всі дані, окрім координат безпосередньо цільового білка. У випадку наявності амінокислот з подвійними координатами один набір також видалявся.

Дослідження, які включали ДНК в ролі цілі, передбачали генерацію на основі нуклеотидної послідовності класичної Уотсон-Кріківської подвійної спіралі довжиною 50 нуклеотидів за допомогою інструментарію молекулярного редактора Avogadro [223]. Отриманий дуплекс піддавався релаксації за допомогою вбудованого функціоналу Avogadro (силове поле UFF, алгоритм найкрутішого спуску (Steepest Descent)) з подальшим коротким сеансом симуляції молекулярної динамки протягом 5нс (див. підпункт 3.2). Безпосередні дані послідовності ДНК відбирались відповідно до потреб експерименту і були отримані або з баз даних нуклеотидних послідовностей GenBank і RefSeq, або з літературних джерел [224].

3.2 Молекулярний докінг і віртуальний скринінг

Процес віртуального скринінгу проводився за допомогою AutoDock Vina 1.1.2 та bash-скрипта [225]. Останній забезпечував автоматичний запуск докінгу кожного наступного ліганду після закінчення попереднього стикування і збереження результату, а також автоматичне продовження процесу після відновлення електропостачання чи аварійного переривання роботи.

За допомогою MGLTools 1.5.7 зона докування була обмежена попередньо визначеним на основі внеску у життєвий цикл патогену покетом [226]. При цьому ключові амінокислоти, які формували поверхню цієї ділянки, розглядались як повністю рухливі, кількість конформаційних поз була обмежена дев'ятьма з найвищим значенням скорингової функції (*vina*), а вичерпність (*exhaustiveness*) була визначеною як 40.

3.3 Симуляція молекулярної динаміки

Моделювання молекулярної динаміки було виконано за допомогою програмного пакету GROMACS 2019.6 [227]. В усіх дослідженнях застосовувався явний розчинник (модель TIP3P) та фізіологічна концентрація Na^+ і Cl^- зі зміщенням відповідно до потреби нейтралізації системи співвідношенням іонів (0,156 M) [228]. Протонування амінокислотних залишків поліпептидів проводилось за допомогою утиліти *rdb2gmx* відповідно до нейтрального рН та з урахуванням pK_a кожного з них. У випадку гістидину також оцінювалось амінокислотне оточення радикалу на предмет потенційного утворення внутрішньомолекулярних водневих зв'язків. На основі цього визначалось найбільш ймовірне розміщення протону імідазольного кільця. Форма системи була залежною від цільового конгломерату: у випадку молекулярної системи, що містила білок, використовувався ікосаедричний бокс, у випадку ДНК-вмісної – триклінічний. Періодичні граничні умови застосовувались під час побудови всіх молекулярних систем. Цільовий молекулярний конгломерат розміщувався в центральній частині боксу, а найменша відстань від нього до стінок визначалась на рівні 24 Å.

Застосовані силові поля також залежали від молекулярної системи. CHARMM36 (CHARMM27) було використане, як найбільш надійне і точне для опису геометрії білків і поліпептидів, натомість силове поле AMBER99SB було застосоване як підтвержене часом надійне силове поле для опису геометрії нуклеїнових кислот [229, 230]. Додатково надійність AMBER99SB щодо

безпосередньо цільової молекулярної системи була уточнена шляхом рутинного порівняльного симуляційного експерименті із використанням одного із останніх силових полів, призначених для нуклеїнових кислот – OL15 [231].

Параметризація лігандів під час моделювання ліганд/рецепторних комплексів відбувалась відповідно до застосованого силового поля за допомогою онлайн ресурсу SwissParam з використанням методу MMFF [232] для CHARMM36 і інструменту ASPYPE [233] із застосуванням GAFF [234] в ролі силового поля і Gasteiger [235] як методу розподілу заряду для AMBER99SB.

На наступному після побудови етапі всі системи зазнавали мінімізації за алгоритмом найкрутішого спуску до значення енергії системи менше ніж 1000 кДж/моль/нм (максимум 50000 кроків). Після чого було проведено врівноваження системи у дві послідовні фази – класичний NVT і NPT. Кожна фаза була довжиною 100пс для систем в процесі ітеративної розробки лігандів та 400пс для скринінгових досліджень. В останньому випадку також були присутні обмеження руху важких атомів ліганду. В усіх випадках наступним йшов етап молекулярно-динамічного моделювання відповідної тривалості без будь-яких додаткових обмежень, прикладених до складових системи.

Більшість параметрів були стандартно подібними для цих трьох етапів у більшості експериментів. Зокрема, часовий інтервал ітерації дорівнював 2 фс, методом підтримки температури системи був вибраний термостат Буссі-Донадіо-Паррінелло, комплекси білок/ліганд або ДНК/ліганд були виділені як одна група для підтримки температури, а вода і іони як інша. В обох випадках константа зв'язку для термостату Буссі-Донадіо-Паррінелло була рівною 0,1 пс, а цільова температура дорівнювала 300К. Алгоритм LINCS був використаний для визначення і підтримки ковалентних зв'язків між важкими атомами і атомами гідрогену у жорсткому стані [236]. Кількість ітерацій, при цьому, була обмежена однією, а порядок матриці обрахунку зведення дорівнював 4. Створення сітки доступних для нековалентних взаємодій частинок було здійснено за допомогою алгоритму Верле з обмеженням по

відстані в 1,2 нм і частотою оновлення списку «сусідів» кожні 20 фс. При цьому взаємодії Ван-дер-Ваальса розраховувались за допомогою потенціалу Ленард-Джонса із зведенням сили до нуля в інтервалі від 1,0 до 1,2 нм, а електростатичні взаємодії описувались за допомогою згладженої сітки Евальда з кубічним ступенем інтерполяції і інтервалом між вузлами сітки в 0,16 нм. Основна відміна між стадіями була в параметрах баростату системи. Зокрема, для стадії NVT розміри боксу були фіксованими, для NPT був застосований баростат Берендсена, а для стадії MD – баростат Паррінелло-Рахмана. За іншими параметрами обидва баростати були подібними: цільовий тиск 1 бар, щоітераційний обрахунок, ізотропні умови, крок стиснення – $4.5e-5 \text{ бар}^{-1}$. Додатково, на стадії NVT, генерація вихідних швидкостей системи здійснювалась за допомогою генератора псевдовипадкових чисел і відповідно до розподілу Максвела для цільової температури.

Всі стадії були проведені із застосуванням GPU в ролі прискорювача.

3.4 Розрахунок вільної енергії зв'язування

Пакет gmx_MMPBSA 1.6.3 був використаний для обрахунку вільної енергії методами MM/PBSA і MM/GBSA [237]. Джерелом вхідних даних була траєкторія, згенерована GROMACS в процесі симуляції молекулярної динаміки цільових асоціатів. Вхідний файл параметрів обрахунку був створений за допомогою gmx_MMPBSA 1.6.3. Більшість змінних були залишені у їх стандартних значеннях. Однак деякі зміни, все ж, були. Зокрема, перші 150 фреймів вихідної траєкторії не використовувались для аналізу, а з решти був відібраний кожен 10-50тий – таким чином, щоб кінцева кількість взятих до аналізу фреймів знаходилась в діапазоні від 100 до 300. Температура, відповідно до оригінальної симуляції, була визначена на рівні 300К, а побудова топології здійснювалась на основі вхідного файлу топології, згенерованого GROMACS. Власне в залежності від застосованого методу змінювався спосіб обрахунку внеску електростатичної компоненти в енергію сольватації. У випадку MM/PBSA застосовувався метод Пуасона-Больцмана із

явним обчисленням густини розподілу заряду в межах асоціату. Метод MM/GBSA можна розглядати спрощеним варіантом методу MM/PBSA, де задля зменшення обчислювального навантаження використовується узагальнення Борна. Це дозволяє жорстко прив'язати розподіл заряду до конкретних атомів, обмежуючи їх радіусом Борна. Неполарний внесок сольватації розраховується ідентично до обох методів відповідно до площі доступної для розчинника. Остання важлива зміна стосувалась іонної сили, яка була визначена в 0.156 M.

3.5 Аналіз результатів

Результати аналізували за допомогою стандартного програмного забезпечення наданого GROMACS (trjconv, rms, hbond, gyrate, rmsf, distance) і gmx_MMPBSA (gmx_MMPBSA_ana). Візуалізація проводилась за допомогою XMGRACE [238], PyMol 1.8 і BIOVIA Discovery Studio Visualizer [239].

3.6 Ітераційний дизайн малих лігандів

Дослідження, в межах яких проводилась розробка малих органічних сполук *de novo*, на найбільш загальному рівні були побудовані ітераційно. Кожна з ітерацій складалась з послідовних етапів симуляції молекулярної динаміки ліганд рецепторного комплексу протягом 10-50нс, аналізу стабільності комплексу на основі отриманих даних та відповідної модифікації ліганду у напрямку посилення спорідненості до цілі (рис. 3.1). Вихідною точкою в залежності від дослідження могли бути знайдені під час віртуального скринінгу ліганди з умовною відповідністю ділянці націлювання або ж повністю штучні молекулярні зонди, єдиною ціллю яких була ініціація процесу дослідження зони докування.

Визначення прийнятних з точки зору хімічної доступності, адсорбції, токсичності, розподілу, метаболізму і екскреції модифікацій здійснювалось шляхом комбінації аналізу доступних баз хімічних сполук PubChem [240], ChEMBL [241], ZINC [242] та розрахункових методів передбачення потенційної фармакокінетики на

основі хімічної структури малої органічної речовини (див. підпункт 3.7). Власне модифікація ліганду здійснювалась за допомогою Avogadro безпосередньо у ліганд/рецепторному комплексі і передбачала короткий сеанс мінімізації енергії за допомогою вбудованого функціоналу цього молекулярного редактора. Таким чином досягалась первинна адаптація активної ділянки цілі до зміненого ліганду.

Процес ітераційно продовжувався до досягнення стабільності ліганд/рецепторного комплексу, порівнюваної або краще за таку для комплексів успішних етіотропних лікарських сполук [198]. Ключовими параметрами, на основі яких приймалось рішення, були середньоквадратичне відхилення ліганду, яке передбачалось стабільним протягом всього процесу симуляції і таким, що коливалось в межах до 3 Å, та енергетичний виграш на більше ніж 5-10 ккал/моль. Крім того, враховувалась поведінка ліганду і оточуючих його амінокислот в ділянці націлювання. Додаткові параметри типу RMSF, кількості водневих зв'язків і середньої відстані від ключових амінокислот аналізувались за потреби.



Рисунок 3.1 Загальна схема однієї ітерації під час розробки лігандів *de novo*

3.7 Обрахунок фармакокінетичної прийнятності

Для визначення параметрів ADMET (Absorption, Distribution, Metabolism, Excretion, Toxicity) та медичної хімії використовувався безкоштовний онлайн-ресурс ADMETlab 2.0. Кожна із молекул, які викликали інтерес своєю спорідненістю до цільового фактору оцінювались за більше ніж 88 параметрами, значна частина із яких перекривалась за своїм фізичним значенням, але мала інший математичний апарат та/або алгоритм в основі розрахунку прийнятності. Таким чином досягався покращений ступінь надійності передбачення [243].

Задля отримання прогнозованих ADMET під час рутинного ітераційного дизайну структура ліганду безпосередньо будувалась на основі попередньої у вбудованому редакторі JMSE. У випадку ж скринінгових досліджень структури завантажувались у форматі SMILES.

РОЗДІЛ 4

РОЗРОБКА ІНГІБІТОРІВ РЕПРОДУКЦІЇ ВІРУСУ ЕПШТЕЙНА-БАРР ШЛЯХОМ ЗАСТОСУВАННЯ СТРУКТУРО ОРІЄНТОВАНОГО ДИЗАЙНУ ЛІКІВ

4.1 *De novo* розробка ЕВАІ як потенційного інгібітора антиапоптозного фактору вірусу Епштейна-Барр білка VHRF1

Як зазначалось вище, віропротеїн VHRF1 є одним з двох факторів вірусу Епштейна-Барр, які забезпечують потужний антиапоптозний ефект як під час літичної стадії, так і на самому початку інфекції. Його критичний внесок зокрема у процес онкогенезу був підтверджений у ряді масштабних мультидисциплінарних досліджень, в тому числі і з розробки його інгібіторів [244]. Однак наразі більшість з цих досліджень є обмеженими розробкою пептидів або ж повноцінних білків, які імітують ВНЗ-домен факторів родини BCL-2, завдяки чому конкурентно інгібують VHRF1. Хоча загально відомо, що подібні високоімунногенні конструкції можуть бути застосовані лише в крайніх випадках і часто потребують специфічних систем доставки, або ж подальшої модифікації для подолання фізіологічних бар'єрів, головним з яких є плазмолема [245]. Цієї вади позбавлені низькомолекулярні сполуки, тож в межах дисертаційної роботи було проведено дослідження з *de novo* розробки конкурентного інгібітора VHRF1 [246].

Таким чином, після більше ніж 200 ітерацій був отриманий ЕВАІ (Epstein-Barr Antiapoptotic Inhibitor). Його молекулярна маса є рівною 611 Да. Зі структурної сторони він має здебільшого ароматичну природу та, як наслідок, плоску структуру, фланковану, при цьому, індольним та пірольним похідними, які є потенційно більш лабільними у порівнянні з центральною частиною (рис. 4.1). Особливо це стосується індолу, який з'єднаний з рештою молекули метиленовою та карбонільними групами. Окрім наявності кон'югованої електронної системи, ця структура, через наявність трьох внутрішньомолекулярних водневих зв'язків між амідними групами, додатково

знає стабілізації сплющеної форми в центральній частині та набуває саме такої конфігурації розміщення донорів і акцепторів водневого зв'язку, яка забезпечує найефективнішу взаємодію ЕВАІ з ВНЗ-зв'язуючою ділянкою ВНRF1.

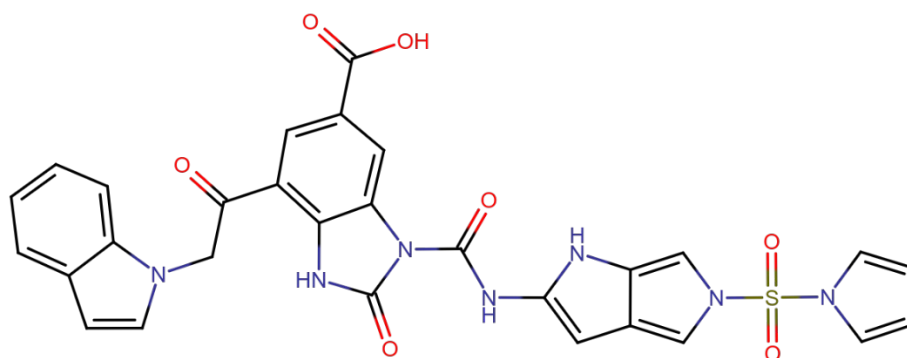


Рисунок 4.1 Структурна формула ЕВАІ.

Задля підтвердження потенціалу ЕВАІ як інгібітора ВНRF1 було проведено два молекулярно-динамічних експеримента. Перший – основний, необхідний для принципового підтвердження придатності ЕВАІ як ліганду ВНRF1. Другий – додатковий як підтвердження результатів першого експерименту. Для обох симуляцій використовувався ідентичний файл координат для основного сеансу симуляції. Під час його генерації, комплекс ЕВАІ-ВНRF1 був навмисно дестабілізований: ліганд був розміщений на відстані більше за суму радіусів Ван-дер-Ваальса сусідніх атомів ЕВАІ і амінокислот ВНЗ-зв'язуючої ділянки ВНRF1. Таким чином, попри те, що ліганд і рецептор були наближеними один до одного, прямих взаємодій, на кшталт водневих, сформовано не було. При цьому далекодіючі електростатичні сили продовжували діяти.

Під час основного моделювання, незважаючи на наявність прямого контакту гідрофобних частин ЕВАІ з розчинником (рис. 4.2), після дестабілізації ліганд

відновив свою конформацію вже до кінця першої наносекунди. Причиною такої швидкої стабілізації були насамперед сильні орієнтуючі взаємодії між центральною частиною інгібітора та BHRF1-R100+N61, де стабільний іонний/водневий зв'язок був характерними для R100, а взаємодія з N61 стабілізувалась шляхом утворення двох водневих зв'язків. Кінцевий індол і сульфонілпірол ліганду в цьому сеансі моделювання також достатньо швидко відновили свою взаємодію з білковим ядром. При цьому, якщо пірольне кільце стабілізувалося в основному за рахунок гідروفобних взаємодій, то індол додатково утворював стійкий T-подібний стекінг з W107.

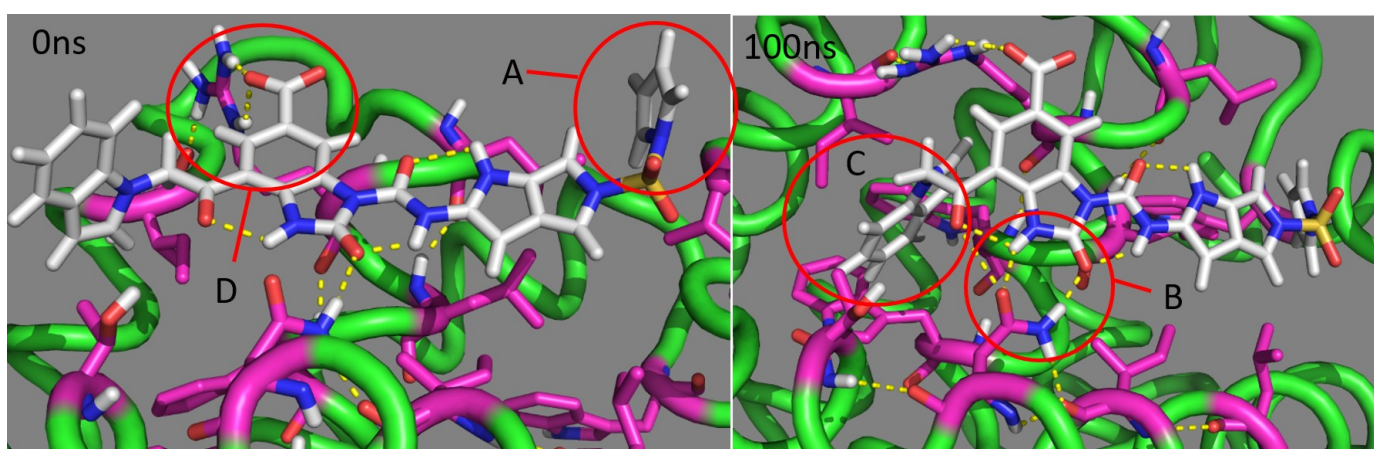


Рисунок 4.2 Поведінка ЕВАІ протягом основного симуляційного експерименту. Сульфонілпірольна (А) та індольна групи не контактують безпосередньо з BHRF1, а центральна частина ліганду стабільно взаємодіє з R100 (D). Сульфонілпірольна та індольна (С) групи відновлюють взаємодію з BHRF1, центральна частина ліганду взаємодіє двома водневими зв'язками з N61 (В)

Попри те, що згадані водневі зв'язки між BHRF1-N61 і центральною частиною ліганду є досить стабільними, за період ~ 70 -85 нс основного моделювання чітко спостерігається втрата одного з них (рис. 4.3; А). А саме того де донором виступає ліганд, а акцептором – карбоніл N61.

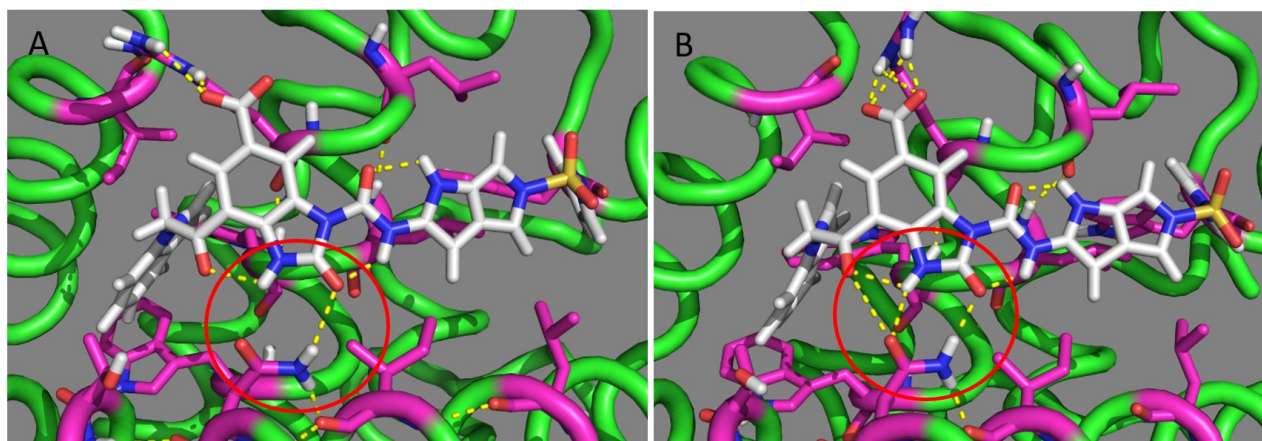


Рисунок 4.3 EBAI у комплексі з VHRF1 на 77ій нс симуляції. А – основне моделювання, один із водневих зв'язків між лігандом та N61 є відсутнім, В – додаткове моделювання, обидва водневих зв'язки між лігандом та N61 є присутніми

Решта взаємодії, особливо іонний зв'язок з R100, спостерігалися протягом усього періоду моделювання. Відповідно, конформація ліганду зазнавала мінімальних змін, які в більшості, відповідали тепловому руху VHRF1.

Додаткове моделювання молекулярної динаміки комплексу VHRF1/EBAI, хоча і було схожим протягом більшої частини симуляції до основного, проте дещо відрізнялося від нього (рис. 4.4). Поряд із порівняною за швидкістю загальною стабілізацією ліганду – відновленням взаємодій, що забезпечуються індолом та центральною частиною інгібітора – кінцевий сульфонілпірол виявив тенденцію до неупорядкованої взаємодії з розчинником протягом перших 5 нс. Однак в подальшому цей пірол вільно відновив свою конформацію та взаємодію з коровою частиною VHRF1. Основною відмінністю на наступних етапах між двома моделюваннями була відносно більша стабільність сформованих орієнтаційних взаємодій EBAI. Зокрема, обидва водневі зв'язки з N61 спостерігалися протягом усього періоду додаткового моделювання (рис. 4.3; В).

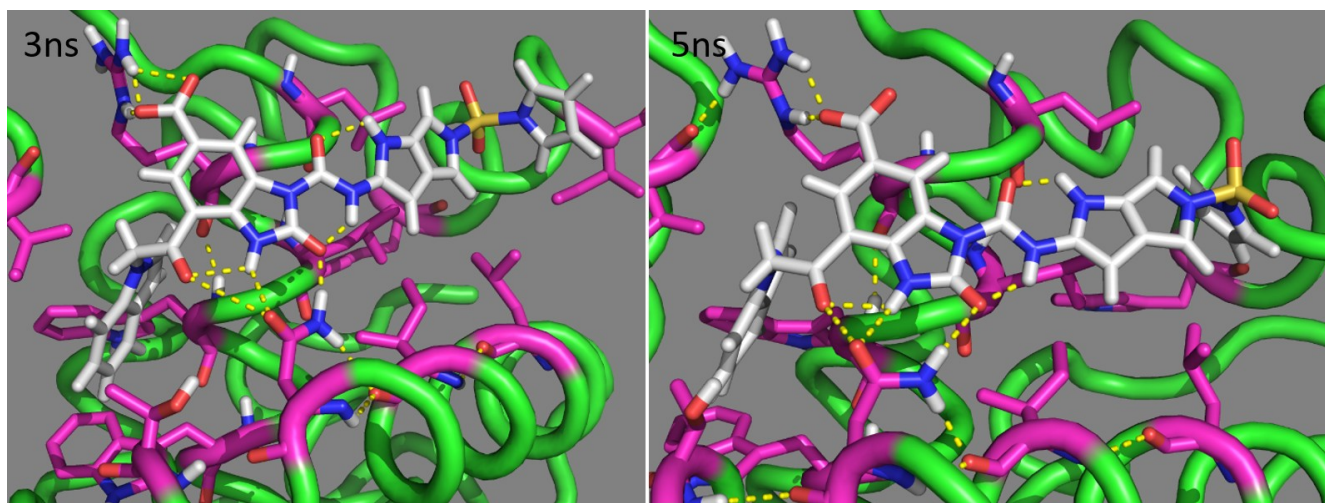


Рисунок 4.4 EBAI протягом додаткового сеансу симуляції. 3 нс – інгібітор відновив всі взаємодії з BHRF1 крім правильного розміщення сульфонілпірролу, 5 нс – сульфонілпіррол також відновив свою взаємодію з BHRF1

Подальший аналіз середньоквадратичного відхилення дозволив підтвердити отримані вище результати. Зокрема, як можна спостерігати на рисунку 4.5, на початку кожного з моделювання ліганд є лабільним – спостерігаються чіткі хвилеподібні рухи графіка залежності RMSD від часу, що закінчується через ~ 30 нс для основного моделювання та ~ 25 нс для додаткового. Згодом RMSD EBAI у разі додаткового моделювання залишається стабільним з діапазоном стандартного відхилення < 1 Å, що чітко узгоджується з траєкторією цього моделювання, описаною вище. При цьому графік RMSD основного моделювання має чіткий період зміни положення ліганду відносно BHRF1, який характеризується зсувом загального стандартного відхилення на $\sim 0,5$ Å і знаходиться в області 70-85 нс моделювання.

Загалом, графіки демонструють хорошу стабільність положення ліганду: після врівноваження щодо ВНЗ-зв'язуючої ділянки BHRF1, його RMSD в обох повторях має чітку тенденцію до стабілізації, а його амплітуда ніколи не перевищує 1 Å.

Особливо чітко це можна побачити на графіку RMSD ліганду під час додаткового моделювання.

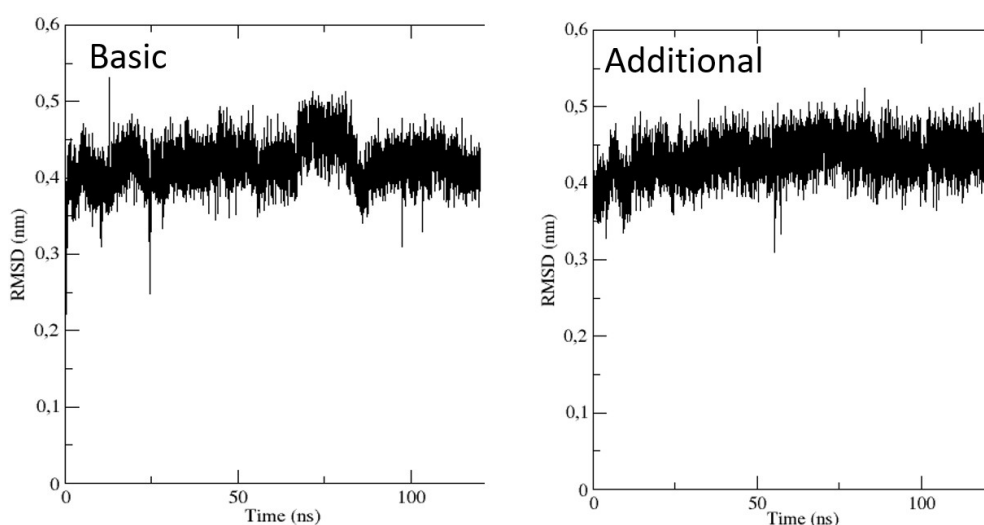


Рисунок 4.5 RMSD ЕВАІ протягом основного і додаткового симуляційних експериментів

Підсумовуючи, в даній частині роботи було показано, що розроблена *de novo* низькомолекулярна структура – ЕВАІ – в імітаційному експерименті, що передбачав значну дестабілізацію її положення, змогла самостійно набути конформації необхідної для стабільної протягом 120 нс взаємодії з ВНЗ-зв’язуючою ділянкою ВНRF1. Відповідно, ЕВАІ можна вважати перспективним кандидатом на звання конкурентного інгібітора цього антиапоптозного фактору вірусу Епштейна-Барр.

4.2 *De novo* розробка селективних ДНК-поліінтеркаляторів з метою протидії вірусу Епштейна-Барр на всіх стадіях життєвого циклу

Вірус Епштейна-Барр, поряд із своєю доведеною здатністю до стимуляції ракового переродження клітин, також, як було описано вище, володіє складним життєвим циклом. Термінальна стадія латентності, характерна для цього патогену,

передбачає майже повну відсутність транскрипційної активності його геному і, відповідно, відсутність віропротеїнів, які могли б розглядатись на роль цілі для розробки специфічного препарату. Таким чином, за умови розробки високоефективного таргетного лікарського засобу, як от ЕВАІ, його застосування дозволить протидіяти ВЕБ у гострій фазі інфекції та у випадку деяких типів раку. Наявність таких препаратів є безперечно важливою, однак той пул патогену, який вже знаходиться у вигляді епісоми в складі В-клітин пам'яті залишиться неушкодженим. Останнє є критичним, зокрема для патологій пов'язаних із супутньою до трансплантації імуносупресією або ж з СНІД. До прикладу, таких як описане вище ПТЛЗ.

Єдиною структурно-функціональною одиницею ВЕБ, наявною у всіх без виключення уражених клітинах, є його геном. Реплікація останнього є синхронізованою і контролюється за допомогою фактора EBNA1, концентрація якого підтримується на достатньому рівні за рахунок функціонування промотора Q_р. Таким чином, ключовою точкою, від якої залежить виживання вірусу Епштейна-Барр навіть у його максимально латентному стані, є промотор Q_р.

Націлювання на ДНК і, особливо, специфічне розпізнавання певних послідовностей є нетривіальним завданням досі невирішеним навіть для протидії такому прямо залежному від генетичних порушень захворюванню як рак. Хоча виходячи з наявності ДНК-зв'язуючих низькомолекулярних агентів, які нині широко використовуються проти лікуванні онкології, принципова можливість розробки подібних речовин існує [247]. Таким чином, в межах дисертаційної роботи було проведено ряд масштабних симуляційних досліджень, направлених на розробку *de novo* селективного ДНК-поліінтеркалятора, здатного диференційно розпізнавати послідовності певних генів.

В ролі структури, аналіз характеристик якої на першому етапі дозволив більш детально зрозуміти механізми взаємодії ДНК/ліганд, був вибраний біс-інтеркалятор XR5944 [248]. XR5944 володіє доведеною біологічною активністю, а також добре

дослідженим механізмом дії. Ця речовина складається з двох трициклічних азотвмісних залишків 9-метилфеназин-1-карбоксаміду в ролі хромофора, з'єднаних між собою дикатіонним лінкером $(\text{CH}_2)_2\text{NH}(\text{CH}_2)_2\text{NH}(\text{CH}_2)_2$ (рис. 4.6). Після проникнення в клітину, XR5944 зв'язується з багатими на $d(\text{ATGCAT})_2$ послідовностями ДНК. Водночас, феназинові хромофори прямо інтеркалюють в ДНК на відстані дві пари основ один від одного. Така відстань є обумовленою довжиною лінкерної ділянки. Власне лінкер при цьому розміщується у великій борозенці ДНК дуплексу. Аміни лінкера, а також амідні хромофору в такому комплексі є здатними до утворення водневих зв'язків з акцепторами нуклеїнових основ, направленими у велику борозенку ДНК. Таким чином, забезпечується деяка вибірковість взаємодії XR5944 з генетичним матеріалом.

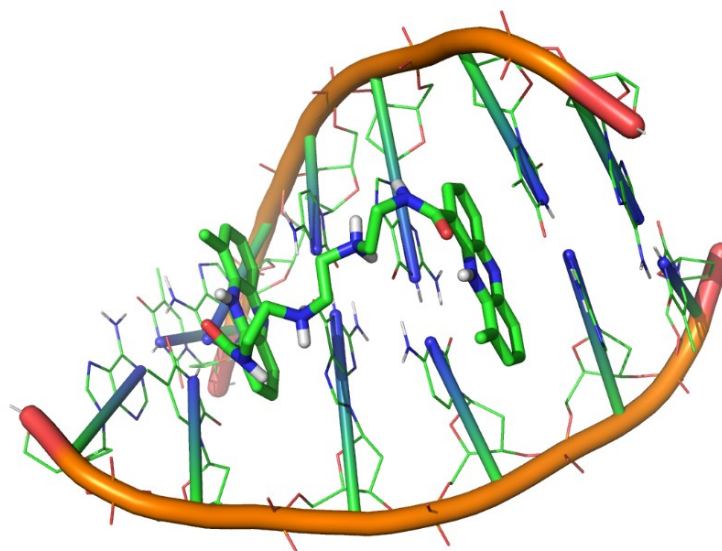


Рисунок 4.6 XR5944 в комплексі з коротким фрагментом ДНК дуплексу

У межах первинного дослідження ітеративно була розроблена тестова молекулярна конструкція, яка складалась з чотирьох феназинових хромофорів, об'єднаних між собою трьома лінкерами таким чином, що останні у комплексі із

тестовою послідовністю фрагмента гена EBNA1 знаходились у великій борозенці над шістьма парами нуклеотидів. Лінкерна ділянка, водночас, була змінена відповідно до концепції формування кожною унікальною послідовністю відповідної унікальної поверхні великої борозенки. Ця концепція базується на наявності направлених у велику борозну ДНК дуплексу трьох точок формування водневих зв'язків від кожної пари нуклеїнових основ. Для пари А-Т це, починаючи з пурину, акцептор (N), донор (NH), акцептор (O). Для пари G-C це, відповідно, акцептор (N), акцептор (O), донор (NH). В залежності від послідовності, паттерн просторового розміщення донорів та акцепторів водневого зв'язку у складі цільового фрагменту ДНК змінюється, формуючи вищезгадану унікальну поверхню великої борозни. У спробі розпізнати карту азотистих основ цільової послідовності (5'-ТТТАСА-3') в ролі скефолдної структури був вибраний транс-декалін модифікований в залежності від відповідної пари нуклеїнових основ у положеннях 1, 4, 6 і 9 (рис. 4.7).

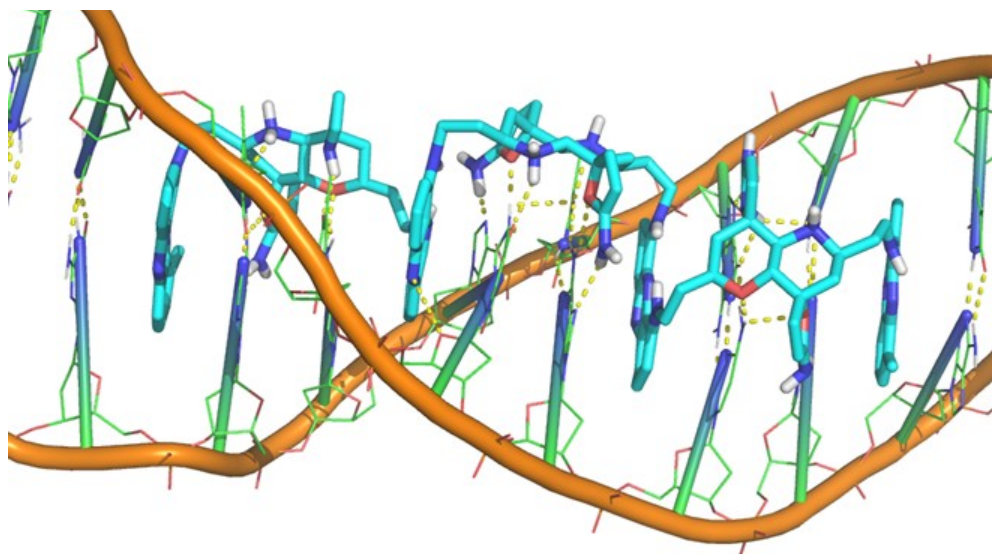


Рисунок 4.7 Тестова конструкція у комплексі з ДНК дуплексом гену EBNA1

Симуляційне дослідження довжиною 50 нс дозволило підтвердити, що, попри модифікації, всі феназинові кільця стабільно зберігають власне положення.

Подібною стабільністю характеризувались під час експерименту і модифіковані декалінові лінкери: всі прогнозовані водневі зв'язки були присутні аж до кінця симуляції.

Однак не скрізь при розпізнаванні були використані всі 3 точки взаємодії на одну пару основ. Зокрема, перша пара основ (Т-А) взаємодіяла з лінкером лише 2ма водневими зв'язками, що є пов'язаним із відносно жорсткою конструкцією декаліну, а також необхідністю розпізнавання двох ідентичних п.о. в межах одного лінкери. Попри це, загально в межах цього експерименту вдалось довести принципову можливість розробки універсального інструменту для контролю рівня транскрипції широкого спектру генів, в тому числі і вірусних [249].

Виходячи з того, що геном людини містить 3 099 734 149 пар основ, а генетична абетка обмежується виключно чотирма літерами, такий інструмент повинен забезпечувати розпізнавання послідовності ДНК певного мінімального розміру. Такого, який, за умови випадкового підбору, буде мати кількість можливих комбінацій нуклеїнових основ фрагменту такого ж розміру, помножених на кількість нуклеотидів у його складі, більшу ніж розмір геному людини. Відповідно до того, що кожна наступна позиція зазнає вибору серед чотирьох можливих нуклеїнових основ – Аденіну, Гуаніну, Тиміну та Цитозину, нерівність буде мати вигляд:

$$4^m * m > 3\,099\,734\,149$$

де m – кількість нуклеотидів в фрагменті ДНК, який розпізнається

Якщо $m = 14$, то кількість варіантів комбінацій становить 268 435 456. Це дає розмір генетичної інформації, серед якої буде зустрічатись в середньому один такий фрагмент, рівною 3 758 096 384 п.о. Відповідно, на всіх наступних етапах для ідентифікації окремої ділянки геному людини або ж ділянки геному патогену мінімальна кількість розпізнаних п.о. має становити 14. В даній роботі, задля

підвищення селективності, послідовність націлювання мала довжину 16 пар основ.

Таким чином, після більше ніж тисячі ітерацій був розроблений наступник першої тестової речовини названий HASDI (**H**igh-**A**ffinity **S**elective **D**N**A** **I**ntercalator) [250]. Задля розпізнавання послідовності ДНК необхідної довжини він складається з восьми сегментів (блоків) кожен з яких є здатним розпізнавати дві пари нуклеїнових основ (рис. 4.8).

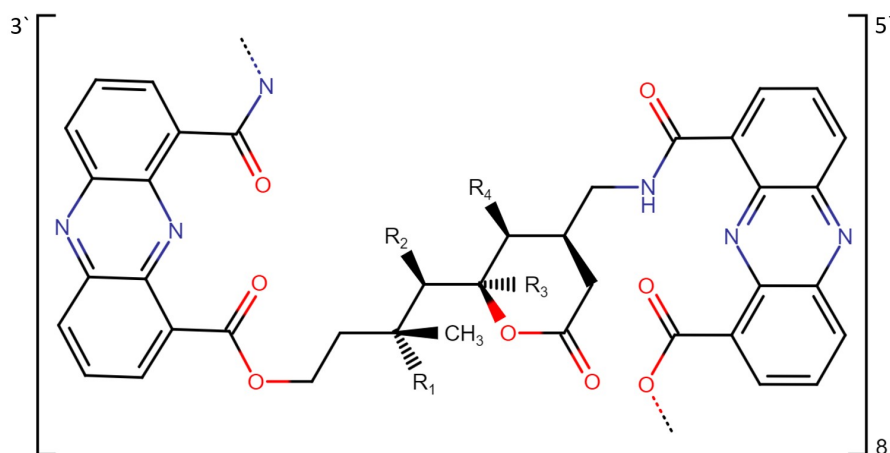


Рисунок 4.8 Загальна структура одного сегменту HASDI, де R1, R2, R3 та R4 є радикалами, що безпосередньо розпізнають послідовність націлювання

Кожен блок, в свою чергу, складається з двох феназинових хромофорів, об'єднаних між собою лінкером і здатних до полімеризації у кінцеву структуру вищого порядку. Скефолдна частина всіх лінкерів є загалом подібною і не залежить від послідовності націлювання – в основі кожного лежать поєднані С-С зв'язком δ -валеріолактон та 2-метилбутил. Для безпосередньої орієнтації на кожну конкретну двійку пар нуклеїнових основ, вони є модифікованими в положеннях 1, 2 аліфатичного метилбутилового ланцюга (R2 та R1 відповідно) і 5, 6 кільця δ -валеріолактону (R4 та R3 відповідно). Конкретні функціональні групи, серед яких під час ітераційної розробки було отримано найкращі результати для розпізнавання

послідовності 5'-TCCSTTTACAACCTCAC-3' (центральна частина дуплексу згенерованого на основі послідовності 5'-TGGAGGTAGTAAGACCTCCSTTTACAACCTCAGGCGAGGAATTGCCCTTG-3' (GenBank: MT164472.1 (28–77); EBNA1-50п.о.)), є представленими у Додатку 1. Молекулярна маса конструкцій, отриманих із застосуванням цього закону, задля націлювання HASDI на певну послідовність довжиною 16 п.о. знаходиться в межах від 7,5 кДа до 7,8 кДа. Назва ж формується на основі аббревіатури HASDI і розміщеної в дужках назви гену в межах якого знаходиться послідовність націлювання. До прикладу, молекулярна конструкція націлена на сиквенс 5'-TCCSTTTACAACCTCAC-3' була названа HASDI(EBNA1).

Задля визначення ступеня її селективності було проведено два симуляційних експерименти. В першому HASDI(EBNA1) був інтеркальований у ДНК дуплекс довжиною 50 п.о., згенерований на основі фрагменту гену EBNA1, який містив власне цей короткий фрагмент. Отже, досліджувався комплекс HASDI з послідовністю націлювання. У другому симуляційному експерименті послідовність, на основі якої було згенеровано ДНК дуплекс, була отримана з випадкової частини гену KCNH2 (5'-AGGCGCTGCCCGAGCCGCGGGCGCTGGAGCGGCTGTCGGCGCGGTGGCAG-3' (RefSeq: NM_000238.4 (28-77); KCNH2-50п.о.). Таким чином, другий комплекс імітував собою інтеркаляцію HASDI у випадкову частину людського геному.

Протягом всього періоду симуляції (150 нс) комплекс EBNA1-50п.о./HASDI(EBNA1) можна було характеризувати як стабільний. Стек-взаємодії всіх феназинових кілець зберігались у незмінному вигляді протягом періоду симуляції: кожна інтеркалююча планарна структура константо знаходилась в місці свого початкового розміщення, тобто на відстані 2 пари нуклеїнових основ один від одного. Однак вибірккові взаємодії у вигляді водневих зв'язків різнились в залежності від конкретної послідовності (Додаток 2).

Відповідно до того, що кожна пара нуклеїнових основ може бути розпізнана максимально трьома водневими зв'язками, кожен модифікований у положеннях R1, R2, R3 та R4 лінкер здатен формувати по 6 водневих зв'язків з двійкою пар основ. Це результується у тому, що сумарно HASDI має змогу розпізнавати паттерн із полярних атомів, який включає 48 точок. Аналіз траєкторії симуляції молекулярної динаміки комплексу HASDI(EBNA1) із послідовністю націлювання свідчить на користь збереження всіх прогнозованих водневих взаємодій протягом всього періоду симуляції. Таким чином, положення всіх лінкерів і, відповідно, пов'язаних з ними феназинів є зафіксованим відносно конкретної пари основ та один відносно одного і змінюється тільки залежно від змін в загальній конформації цільового фрагменту ДНК.

Однак відповідно до згенерованого hbond (GROMACS) графіку залежності кількості водневих зв'язків від часу моделювання, у кожен конкретний проміжок часу кількість водневих зв'язків у комплексі ніколи не перевищувала 43 і в середньому коливалась близько 32 (рис. 4.9). Це не дорівнює 48 можливим водневим зв'язкам, які спостерігались в процесі аналізу траєкторії симуляції та свідчить на користь більш динамічної взаємодії між HASDI(EBNA1) і цільовою ДНК.

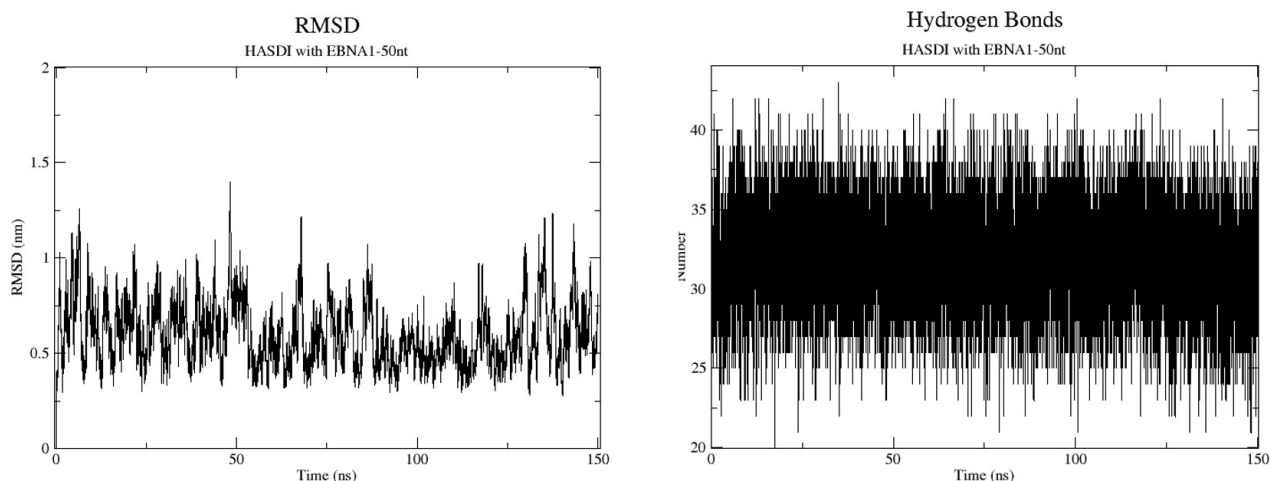


Рисунок 4.9 Середньоквадратичне відхилення HASDI(EBNA1) у комплексі з EBNA1-50п.о. та кількість водневих зв'язків між ним і послідовністю націлювання

Середньоквадратичне відхилення HASDI(EBNA1) в цільовому комплексі коливалось навколо значення в 6.5 \AA . З піками до 10 \AA та мінімумами до 3.5 \AA (рис. 4.9). Це високе значення як для загального RMSD, так і для його розмаху. Однак слід зазначити, що загальна тенденція середньоквадратичного відхилення не зазнає росту, а навпаки є напрочуд стабільною. Крім того, значний розмах коливань є наслідком загальних ундулюючих рухів значення RMSD і у випадку оцінки локального розмаху не перевищує 2 \AA .

Оцінена за допомогою `gmx_MMPBSA` вільна енергія зв'язування HASDI(EBNA1) з EBNA1-50п.о. має глибоко негативне значення ($-235,3 \pm 7,77$ ккал/моль), що також свідчить на користь наявності високого рівня спорідненості між компонентами комплексу.

У другому симуляційному експерименті, серед всього вибраного в межах KCNH2 фрагменту, HASDI(EBNA1) був інтеркальований в ділянку послідовності 5'-CGGGCGCTGGAGCGGC-3', яка значно відрізняється від послідовності націлювання. Попри невідповідність, на рівні неспецифічної компоненти (стек і гідрофобні взаємодії) молекулярна динаміка HASDI була подібною до такої для

EBNA1-50п.о./HASDI(EBNA1). Зокрема, кожне планарне кільце феназину незмінно знаходилось в місці своєї вихідної інтеркаляції, не зазнаючи значних змін положення. При цьому спостерігались значні зміни у взаємодіях характерних для специфічної компоненти. Серед всієї послідовності 5'-CGGGCGCTGGAGCGGC-3' було лише два однонуклеотидних перекриття з послідовністю 5'-TCCCTTACAACCTCA-3' – A55 та C57, які взаємодіють з R4/ R3 6го лінкера та R4/ R3 7го, відповідно (Додаток 3). В обох випадках спостерігалось стабільне збереження 3 водневих зв'язків із кожною із пар нуклеотидів. Решта пар основ нецільового фрагменту обмежувались в середньому 1-2 стабільними зв'язками на один не призначений для їх розпізнавання лінкер. Цікаві аберації спостерігались у випадку взаємодії HASDI(EBNA1) із ДНК в ділянці 4-го лінкера (R4/ R3): з початку симуляції ця ділянка була в рівному ступені нестабільною, як і інші ділянки взаємодії ліганду із нецільовими нуклеїновими основами, і формувала в середньому 2-3 водневих зв'язки. Однак починаючи з 93-ї нс моделювання C51 покинув площину дуплексу, що призвело до стабілізації лінкеру і формування 5-ти водневих зв'язків. Серед іншого слід зазначити, що гідроксигрупа R4 першого лінкера також виступала донором у формуванні стабільного внутрішньомолекулярного зв'язку з киснем амідної групи.

Таким чином, в процесі аналізу траєкторії симуляції молекулярної динаміки спостерігались в середньому близько 17-20 водневих зв'язків між специфічною компонентою HASDI(EBNA1) та KCNH2-50п.о. Ці спостереження підтверджуються згенерованим GROMACS більш детальним оглядом кількості водневих взаємодій в часі (рис. 4.10). Їх кількість між HASDI(EBNA1) та KCNH2-50п.о. з часом моделювання динамічно змінювалась. Від початку симуляції і до 110 нс їх число зростало з 16ти до 19ти, після чого знижувалось до 17ти в кінці моделювання. Піково HASDI(EBNA1) взаємодіяв з KCNH2-50п.о. через близько 25 водневих зв'язків (108 нс). Що, попри все, є значно нижчим значенням навіть за середню кількість цих взаємодій характерних для комплексу EBNA1-50п.о./HASDI(EBNA1).

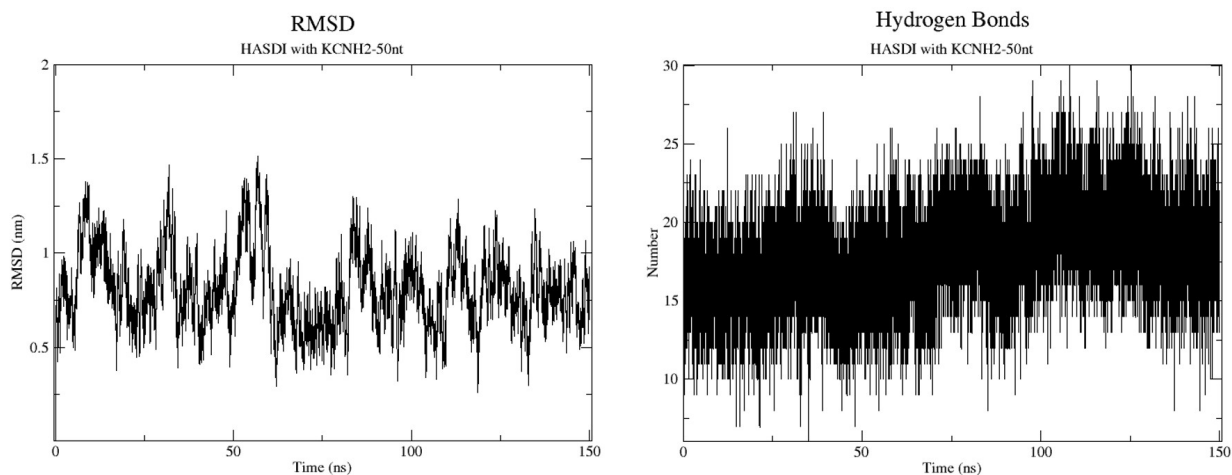


Рисунок 4.10 Середньоквадратичне відхилення HASDI(EBNA1) у комплексі з KCNH2-50п.о. та кількість водневих зв'язків між ним і нецільовою послідовністю

RMSD HASDI(EBNA1) в комплексі із KCNH2-50п.о. значно коливалось протягом симуляції молекулярної динаміки (рис. 4.10). Від початку можна спостерігати коливання середньоквадратичного відхилення від 4,5 Å аж до 14,5 Å з відповідним розмахом близько 10 Å. Це є більшим значенням, порівняно з HASDI(EBNA1) в комплексі з EBNA1-50п.о., та є наслідком виражених ундулюючих рухів значення RMSD, а також локального розмаху середньоквадратичного відхилення в 4 Å. Однак перелічене є більш характерним для першої частини моделювання. Починаючи з 85 нс спостерігається збільшення ступеня періодичності ундулюючих рухів середньоквадратичного відхилення, а також зменшення їх розмаху до 6 Å. При цьому локальний розмах зменшився до 3 Å.

Розрахована вільна енергія комплексоутворення HASDI(EBNA1) з KCNH2-50п.о., схоже до HASDI/EBNA1-50п.о., знаходиться глибоко в негативних значеннях і становить $-193,47 \pm 14,09$ ккал/моль. Це свідчить на користь збереження в HASDI(EBNA1) деякого рівня спорідненості і до неспецифічних ділянок ДНК.

Таким чином, можна стверджувати, що HASDI, попри всі зміни і модифікації, поряд із здатністю підтримувати стабільну взаємодію з ДНК, володіє також достатньо високим рівнем селективності. Останнє виражається у падінні середньої кількості водневих взаємодій між лігандом і ДНК з 32х до 17-19ти, рості і дестабілізації RMSD, а також зниженні падіння вільної енергії зв'язування з $-235,3 \pm 7,77$ ккал/моль до $-193,47 \pm 14,09$ ккал/моль у випадку інтеркаляції чітко націленої молекулярної конструкції до ДНК дуплексу із невідповідною послідовністю. Показово, що початок руйнування комплексу HASDI(EBNA1)/KCNH2-50п.о., виражений у локальному плавленні ДНК дуплексу, спостерігався вже в межах 100 нс від початку симуляції. Враховуючи набагато більший період, характерний для конформаційних перебудов макромолекул, це говорить про близьку до нуля ймовірність утворення комплексу HASDI з настільки неспорідненою ділянкою ДНК. Хоча слід зазначити, що, опираючись на збереження положення феназинових хромофорів, а також значно негативні значення вільної енергії зв'язування, в тому числі і у комплексі з HASDI(EBNA1)/KCNH2-50п.о., HASDI володіє надлишковою спорідненістю до ДНК дуплексу як такого. Враховуючи значне падіння кількості водневих зв'язків, а також фізику процесу, це явище з високою ймовірністю забезпечується феназиновими хромофорами і їх потужною спряженою π -електронною системою [251].

Таким чином, наступним кроком у розвитку HASDI як інструменту/агенту, здатного до селективного розпізнавання послідовності ДНК, є дизайн його варіанту із зниженою неселективною компонентою комплексоутворення. Його наступник був розроблений після більше ніж 200 ітерацій під час кожної з яких ретельно контролювалось збереження кількості міжмолекулярних водневих взаємодій. Він отримав назву HASDI-G2 від **H**igh-Affinity Selective **D**N_A Intercalator – **G**eneration 2 [252] (рис. 4.11).

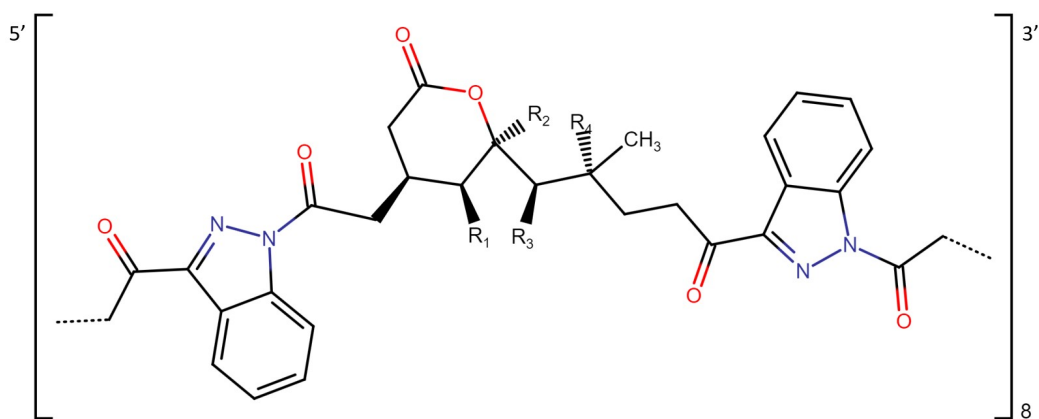


Рисунок 4.11 Загальна структура одного сегменту HASDI-G2, де R1, R2, R3 та R4 є радикалами, що безпосередньо розпізнають послідовність націлювання

Розпізнавання ДНК послідовності HASDI-G2 є реалізованим на основі тих же особливостей, що і у попередника. Власне скефолдна ділянка лінкерної частини кожного блоку ліганду взагалі не зазнала змін. Поєднані ковалентним С-С зв'язком, δ -валеріолактон та 2-метилбутил, в залежності від послідовності, модифікувались в положеннях 1, 2 аліфатичного ланцюга метилбутилу (R3 та R4 відповідно) і 5, 6 кільця δ -валеріолактону (R1 та R2 відповідно), що забезпечує розпізнавання двох пар основ на основі їх донорів і акцепторів водневого зв'язку, направлених у напрямку великої борозенки ДНК дуплексу. Однак безпосередньо перелік замісників для R1, R2, R3 та R4 застосованих тут є дещо іншим і більш повноцінним. Замісники, а також закон їх комбінації, є представленим у Додатку 4.

Порівняно з лінкерною частиною, хромофори зазнали набагато більш значних змін. Зокрема, феназинові кільця були замінені на залишки індазолу, карбоксиловані за першим і третім атомами. Через низький ступінь поляризації та наявність кон'югованої π -електронної системи планарні індазольні кільця легко беруть участь у гідрофобних та Ван-дер-Ваальсових взаємодіях, що забезпечує їх неспецифічну спорідненість до ДНК як дволанцюгової спіралі за механізмом класичним для ДНК-інтеркаляторів. При цьому, порівняно з феназиновими хромофорами HASDI,

індазольні залишки є значно меншими за розміром, що відповідно знижує загальну гідрофобність HASDI-G2, зміщує баланс енергетичних складових комплексоутворення в сторону їх селективної частки, а також призводить до відчутного зниження його молекулярної маси, яка, для здатних розпізнавати 16 п.о. похідних, коливається в межах від 5,8 кДа до 6 кДа.

Для визначення ступеня впливу змін внесених при розробці HASDI-G2 найперше було повторено експеримент описаний раніше для HASDI. Зокрема було створено HASDI-G2(EBNA1), замісники якого для націлювання на послідовність 5'-TCCCTTTACAACCTCA-3' в положеннях R1, R2, R3 та R4 лінкерних ділянок були вибрані відповідно до описаного в Додатку 4 закону. Після цього HASDI-G2(EBNA1) був інтеркальований в дуплекс, згенерований на основі гена EBNA1, довжиною 50 п.о., що містив послідовність націлювання, а також в дуплекс такої ж довжини, згенерований на основі послідовності гену KCNH2. Отримані два комплекси HASDI-G2(EBNA1)/ДНК були досліджені методом симуляції молекулярної динаміки протягом 150 нс.

Протягом всього періоду симуляції комплекс HASDI-G2(EBNA1)/EBNA1-50п.о. можна, подібно до такого у випадку HASDI, описати як напрочуд стабільний. Стек-взаємодії всіх індазольних кілець зберігались у незмінному вигляді протягом періоду симуляції – кожна інтеркалююча планарна структура константо знаходилась в місці свого початкового розміщення, тобто на відстані 2 пар нуклеїнових основ один від одного (рис. 4.12).

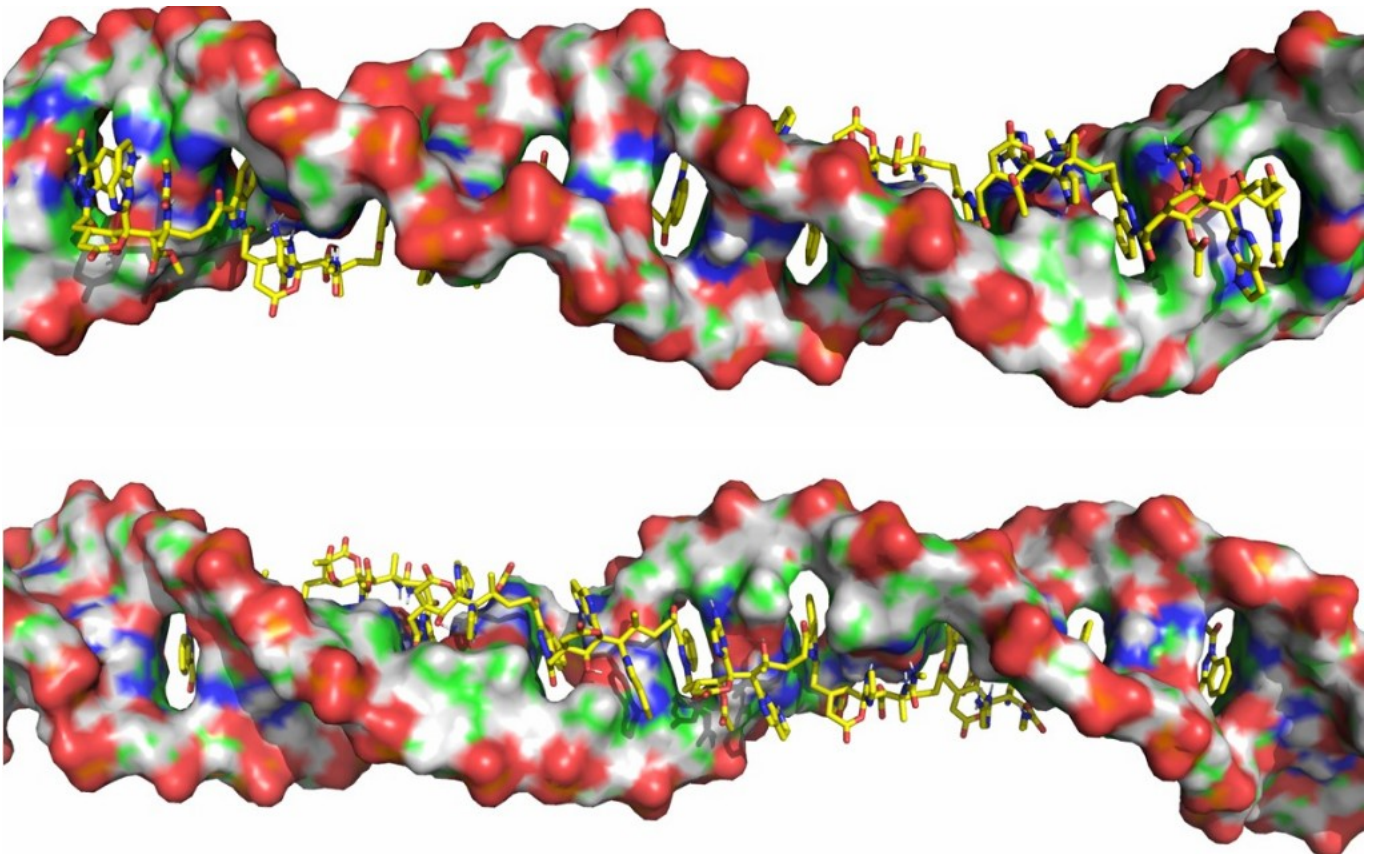


Рисунок 4.12 Загальний вигляд комплексу HASDI-G2(EBNA1) з EBNA1-50п.о. під час симуляції молекулярної динаміки

Водневі зв'язки як селективна компонента взаємодії коливались в межах від 30ти до 42х (рис. 4.13), із середнім значенням в 36 контактів на комплекс. Що є кращим значенням порівняно із попередником у подібному комплексі – середнє 32 водневих зв'язки. Відповідно, в середньому один лінкер, який теоретично здатен розпізнавати 2 пари нуклеїнових основ шляхом формування бти водневих зв'язків, в межах даної симуляції взаємодіє з ДНК в 4-5 точках. Без тенденції до дестабілізації взаємодії.

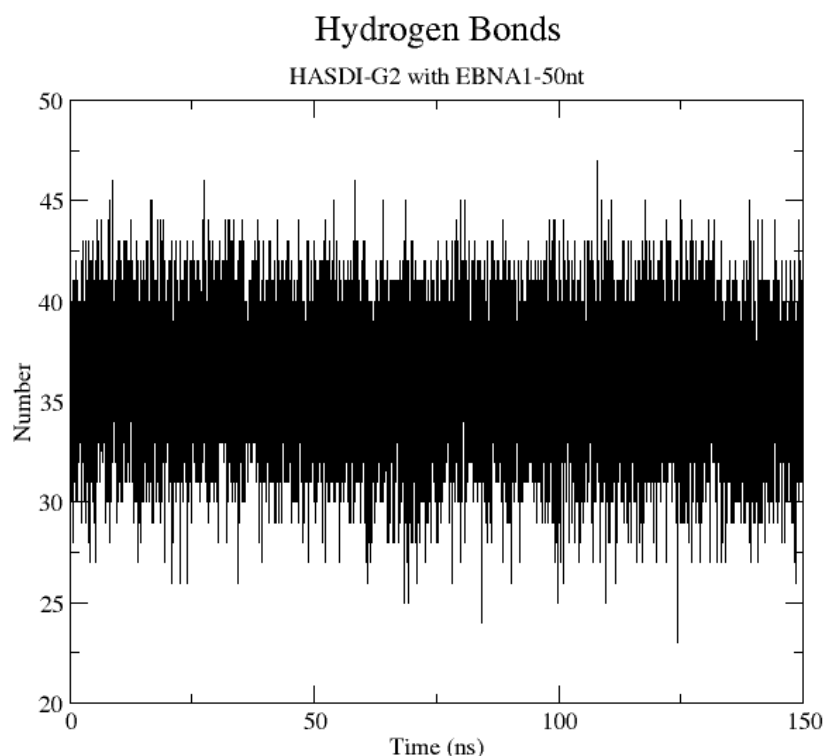


Рисунок 4.13 Залежність кількості водневих зв'язків у комплексі HASDI-G2(EBNA1)/EBNA1-50п.о. від часу симуляції молекулярної динаміки

Високий рівень стабільності отриманого конгломерату підтверджує і енергія комплексоутворення (ΔG), розрахована на основі траєкторії симуляції методом ММ/РBSA, яка дорівнює $-157.72 \pm 6,57$ ккал/моль (таблиця 4.1). Водночас, ентальпія (ΔH) цієї взаємодії є рівною $-177.31 \pm 6,57$ ккал/моль, а ентропійний штраф ($-T\Delta S$) $19,59 \pm 0,05$ ккал/моль. Серед ентальпійних складових стабілізуючим внеском володів вклад сил Ван-дер-Ваальса (E_{vdw}), електростатичних взаємодій (E_{el}) та неполярного компонента енергії сольватації (E_{npol}) – $-395,43 \pm 6,72$ ккал/моль, $-76,33 \pm 3,55$ ккал/моль та $-198,88 \pm 2,62$ ккал/моль відповідно. Натомість, дестабілізуючими були внески полярної енергії сольватації (E_{pb}) і енергії дисперсійних сил (E_{disp}) – $140,79 \pm 3,59$ ккал/моль та $352,54 \pm 3,52$ ккал/моль,

відповідно. Значний внесок саме енергії Ван-дер-Ваальса і дисперсійних сил пов'язаний, очевидно, з порівняно значною площею саме індазольних кілець у складі HASDI-G2(EBNA1). Натомість, внесок електростатичних взаємодій першочергово є наслідком специфічних взаємодій лінкерних ділянок ліганду з нуклеїновими основами.

Таблиця 4.1

Загальна енергія зв'язування та її компоненти, розраховані для комплексу HASDI-G2(EBNA1)/EBNA1-50п.о у ккал/моль

	Evdw	Eel	Epb	Enpol	Edisp	ΔH	$-T\Delta S$	ΔG
Average	-395,43	-76,33	140,79	-198,88	352,54	-177,31	19,59	-157,72
SD	6,72	3,55	3,59	2,62	3,52	6,57	0,05	6,57

Симуляція молекулярної динаміки HASDI-G2(EBNA1) у комплексі з випадковою послідовністю ДНК, якою в даному випадку виступив також фрагмент гена KCNН2, натомість, значно відрізнялась від такої у випадку «правильного» комплексу. Вже через 6 нс симуляції швидке наростання дестабілізації 5`-кінця ліганду призвело до виходу крайнього індазольного кільця за межі дуплексу і, відповідно, до втрати відповідним лінкером будь-яких орієнтуючих взаємодій з нуклеосоновами. Протягом наступних 20 нс відбулась стабілізація вивільненої частини HASDI-G2(EBNA1) у проміжній формі, за якої індазольне кільце, позбавлене стек взаємодій з нуклеосоновами, все ж розміщується у великій борозенці поряд із лінкером другого модуля ліганду. Ця конформація проіснувала до 80ї наносекунди симуляції. Після чого спостерігалась повторна зміна конформації, яка цього разу стосувалась ДНК дуплексу в зоні розміщення першого і другого лінкера HASDI-G2(EBNA1). Це виражалось у локальному плавленні ДНК з

подальшим виходом С30 антисенсового ланцюга за межі дуплексу і стабілізацією його взаємодії з індазольним кільцем ліганду, яке також розміщувалось поза межами своєї вихідної локалізації (рис. 4.14). Ця конформація додатково стабілізувалась двома рівноцінними внутрішньомолекулярними водневими зв'язками між гідроксилом R1, призначеним для розпізнавання тиміну сенсового ланцюга, та киснем кетогрупи з однієї сторони і киснем замісника R3, призначеного для розпізнавання цитозину, з іншої. Подібне розміщення ліганду і ДНК в цій ділянці зберігалось до кінця симуляції.

Окрім значної дестабілізації ДНК дуплексу в ділянці 5'-кінця ліганду, в районі взаємодії 5го лінкеру HASDI-G2(EBNA1) подібно спостерігалось локальне плавлення ДНК. Після 60ти наносекунд стабільної Уотсон-Кріківської взаємодії, G26 сенсового ланцюга майже повністю покинув дуплекс. Єдиний водневий зв'язок, який залишався, був представлений взаємодією O6 гуаніну і 4NH2 цитозину. Така конформація G26 була присутньою до кінця симуляції і додатково стабілізувалась неканонічним водневим зв'язком між 2NH2 гуаніну і киснем фосфатної кислоти у складі негативно зарядженого кістяка ДНК. При цьому індазольне кільце, яке фланкує цей лінкер з 5'-кінця, було значно зміщеним у напрямку від G26. Ця конформація стабільно спостерігалась до кінця симуляції.

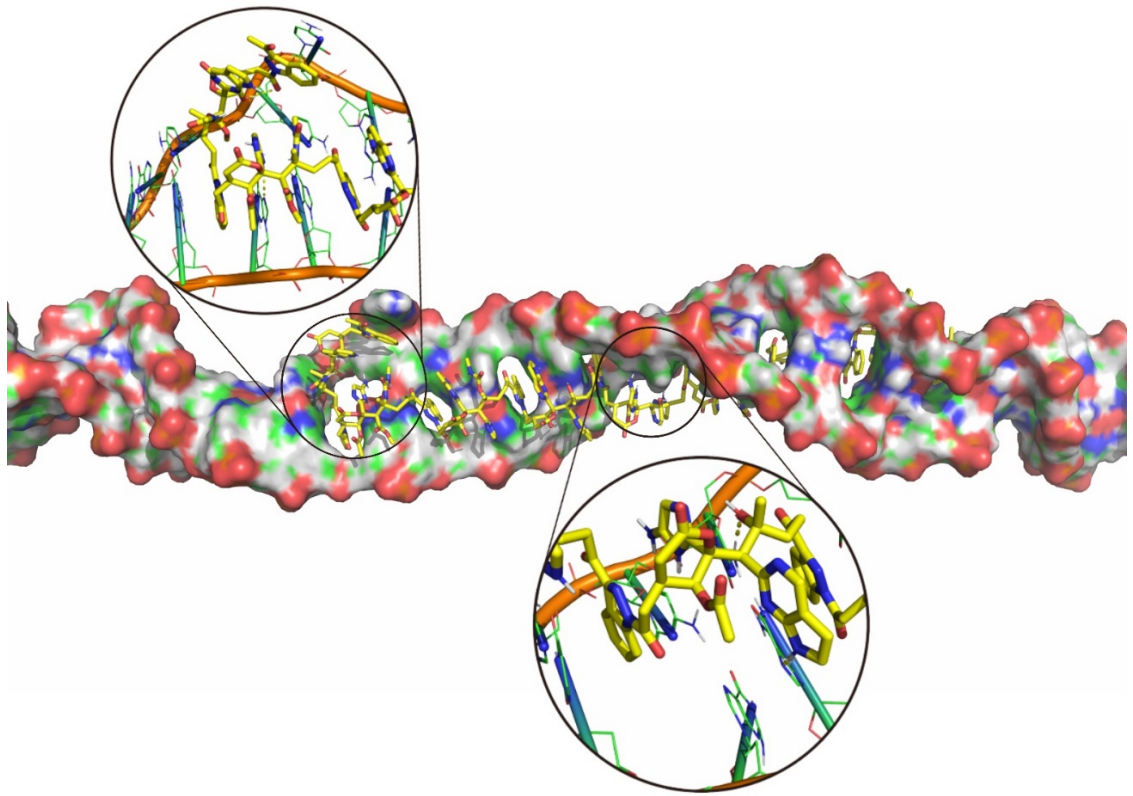


Рисунок 4.14 Загальний вигляд комплексу HASDI-G2(EBNA1) з KCNH2-50п.о. під час симуляції молекулярної динаміки

Відповідно до прогресуючих конформаційних змін, паттерн водневих зв'язків як основи специфічної взаємодії HASDI-G2 з певною послідовністю ДНК характеризувався значним зниженням їх кількості (рис. 4.15). Зокрема, на початку симуляції середня кількість водневих зв'язків між HASDI-G2(EBNA1) і KCNH2-50п.о. становила 24. Починаючи з 8 нс і до 140 нс їх кількість знизилась до 20ти. В кінці симуляції спостерігається ще більше падіння – до 19ти міжмолекулярних водневих зв'язків. Відповідно, на рівні кожного окремого лінкера із 6ти можливих є наявними в середньому 2,5 водневих зв'язки. Це підтверджувалось і аналізом траєкторії симуляції молекулярної динаміки. Лише у випадку випадкових співпадінь між послідовностями EBNA1-50п.о. та KCNH2-50п.о. (A28 і C30) спостерігалась наявність всіх водневих зв'язків. У решті випадків були присутні лише випадкові

донорно-акцепторні взаємодії – від 1єї до 3х на один лінкер, хоча перший лінкер після виходу 5`-індазолу на мав ніяких орієнтуючих взаємодії з ДНК дуплексом.

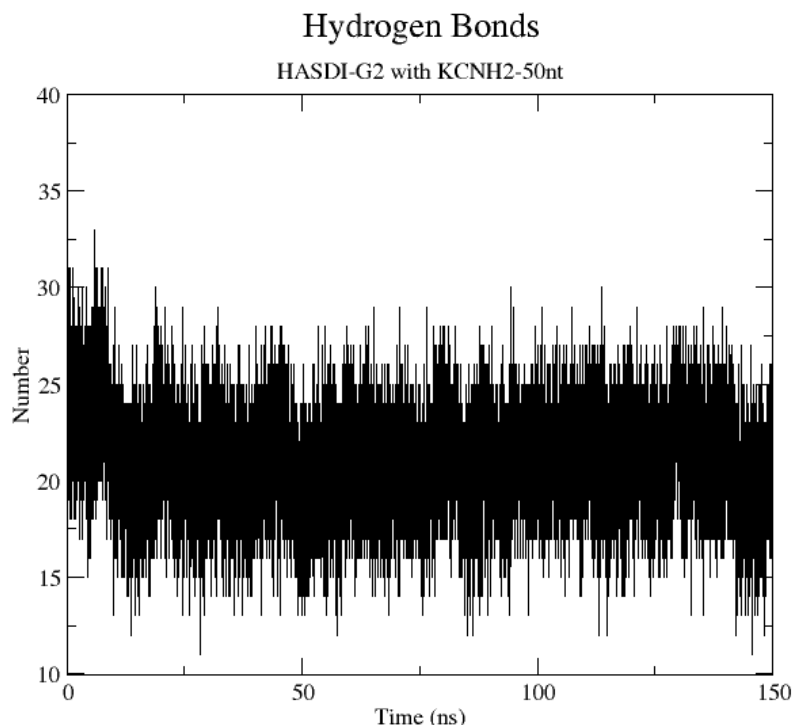


Рисунок 4.15 Залежність кількості водневих зв'язків у комплексі HASDI-G2(EBNA1)/KCNH2-50п.о. від часу симуляції молекулярної динаміки

Порівняно низький рівень стабільності отриманого асоціату підтверджується і на основі розрахунку вільної енергії зв'язування, яка у даному випадку є рівною $-97,25 \pm 9,18$ ккал/моль (таблиця 4.2). При цьому ентальпія комплексоутворення становить $-139,65 \pm 9,18$ ккал/моль, а ентропійний штраф $42,39 \pm 0,05$ ккал/моль. Серед ентальпійних складових, як і у випадку комплексу HASDI-G2(EBNA1)/EBNA1-50п.о., стабілізуючим внеском володіли сили Ван-дер-Ваальса, електростатичні взаємодії та неполярний компонент енергії сольватації – $-356,74 \pm 11,89$ ккал/моль, $-53,89 \pm 5,17$ ккал/моль та $-180,53 \pm 5,76$ ккал/моль,

відповідно. Натомість, дестабілізуючими були внески полярної енергії сольватації і енергії дисперсійних сил – $123,08 \pm 5,88$ ккал/моль та $328,43 \pm 8,74$ ккал/моль, відповідно. Найбільша різниця в енергетичних складових порівняно з оригінальним комплексом спостерігалась у випадку електростатичних взаємодій (-29,4 %), які найкраще відображають внесок специфічної компоненти взаємодії HASDI-G2(EBNA1)/ДНК.

Таблиця 4.2

Загальна енергія зв'язування та її компоненти, розраховані для комплексу HASDI-G2(EBNA1)/KCNH2-50п.о у ккал/моль

	Evdw	Eel	Epb	Enpol	Edisp	ΔH	$-T\Delta S$	ΔG
Average	-356,74	-53,89	123,08	-180,53	328,43	-139,65	42,39	-97,25
SD	11,89	5,17	5,88	5,76	8,74	9,18	0,05	9,18

Враховуючи отримані результати, було вирішено поглибити дослідження рядом додаткових симуляційних експериментів, в яких HASDI-G2 націлювався на цілком реальні мутовані ділянки, характерні для перероджених клітин. Як найбільш яскравий приклад був розглянутий зрощений ген BCR_ABL1, який є молекулярною основою так званої філадельфійської хромосоми – основної причини хронічного мієлоїдного лейкозу [253]. Замісники HASDI-G2 в положеннях R1, R2, R3 та R4 підбирались відповідно до закону описаному в Додатку 4 задля націлювання на ділянку зрощення BCR з ABL1 – 5`-CTGTGGAGTGGGTTTT-3` (центральна частина дуплексу згенерованого на основі послідовності 5`-ACACSTTTGACCCTGGCCGCTGTGGAGTGGGTTTTATCAGCTTCCATACC-3` (GenBank: JQ425253.1 (88–137); BCR_ABL1-50п.о.)). Таким чином, ця частина

дослідження включала три симуляційні експерименти, в яких було визначено спорідненість HASDI-G2(BCR_ABL1) до ділянки націлювання, а також незмінених генів BCR і ABL1 окремо. Де BCR-50п.о. був згенерований на основі послідовності 5'-ACACSTTTGACCCTGGCCGCTGTGGAGTGTTTGTGCTGGTTGATGCCTTC-3' (GenBank: LC775148.1 (109–158)), а ABL1-50п.о. на основі 5'-CTTTGTATTTCCATATACATTTTAGAGTGGGTTTTATCAGCTTCCATACC-3' (RefSeq: NG_012034.1 (22856–22905)). Таким чином, безпосередня зона інтеркаляції кожного разу знаходилась в точці зрощення BCR з ABL1. Тобто HASDI-G2(BCR_ABL1) в будь-якому випадку хоча б частково відповідав послідовності ділянки інтеркаляції.

Симуляція молекулярної динаміки комплексу HASDI-G2(BCR_ABL1)/BCR_ABL1-50п.о. дозволила підтвердити рівень спорідненості ліганду, подібний до такого для комплексу HASDI-G2(EBNA1)/EBNA1-50п.о. (рис. 4.16). Зокрема, під час симуляції всі індазольні кільця знаходились безпосередньо в зоні інтеркаляції HASDI-G2(BCR_ABL1), тобто через кожні дві пари нуклеїнових основ. Кожен лінкер формував численні водневі зв'язки із донорами і акцепторами водневого зв'язку, що знаходяться у великій борозенці дуплексу. А кожна пара нуклеїнових основ ДНК дуплексу незмінно знаходилась в межах подвійної спіралі і була стабілізована класичними Уотсон-Кріківськими взаємодіями.

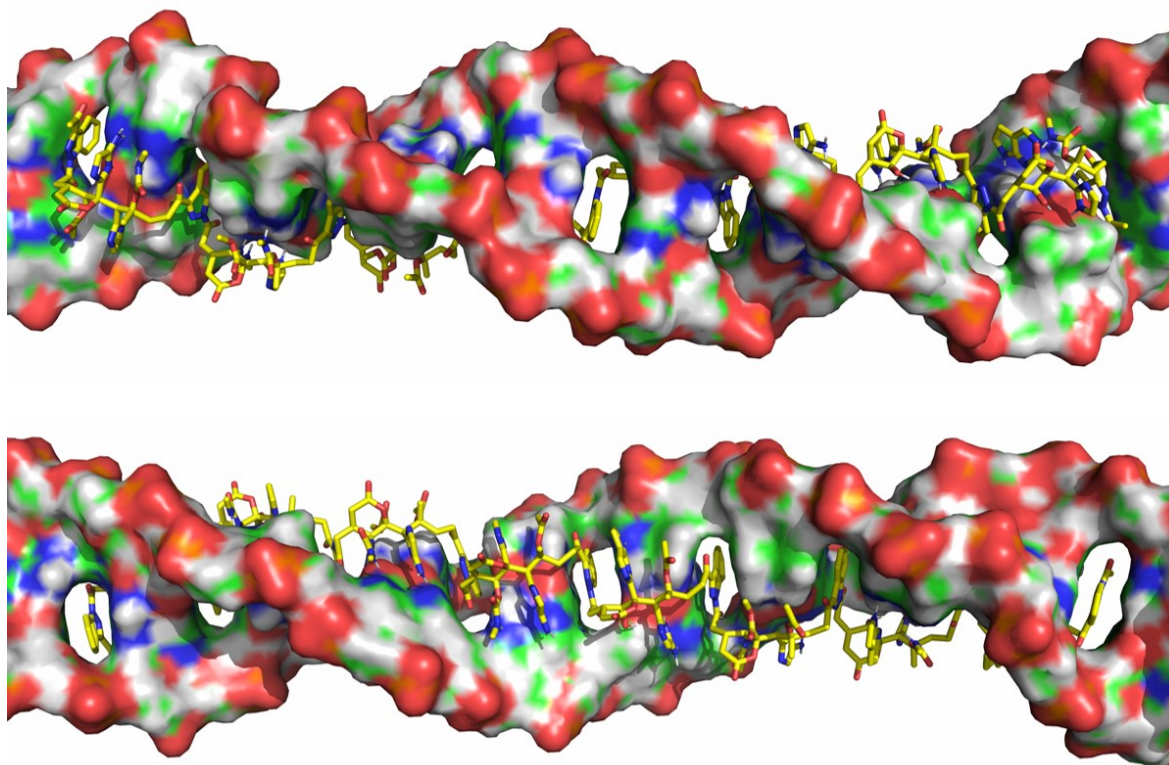


Рисунок 4.16 Загальний вигляд комплексу HASDI-G2(BCR_ABL1) з BCR_ABL1-50п.о. під час симуляції молекулярної динаміки

Кількість водневих зв'язків при націлюванні HASDI-G2 на BCR_ABL1-50п.о. була дещо нижчою, ніж у випадку його націлювання на EBNA1-50п.о., і в середньому становила 34 водневі зв'язки (рис. 4.17). Що відповідає 4,25 селективній взаємодії на один лінкер і є вищим значенням, порівняно із попередньою версією HASDI. Слід зазначити, що з 1ї по 7му наносекунду симуляції кількість водневих зв'язків була нижчою і коливалась близько 31го на комплекс, однак решта симуляції характеризувались саме 34 взаємодіями.

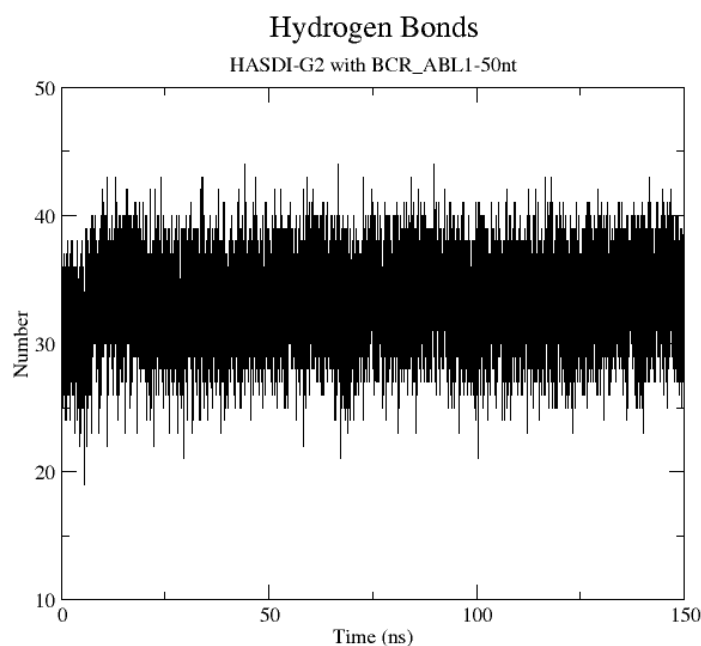


Рисунок 4.17 Залежність кількості водневих зв'язків у комплексі HASDI-G2(BCR_ABL1)/BCR_ABL1-50п.о. від часу симуляції молекулярної динаміки

Ступінь стабільності системи підтверджується і розрахунком її вільної енергії зв'язування, яка в даному випадку є рівною $-154,95 \pm 6,93$ ккал/моль, що всього на $2,77$ ккал/моль є меншим, порівняно із еталонним комплексом HASDI-G2(EBNA1)/EBNA1-50п.о.. Подібна ситуація спостерігається і у випадку інших енергетичних складових (таблиця 4.3). При цьому, всі стабілізуючі сили (зв'язки Ван-дер-Ваальса, електростатичні взаємодії, неполярний компонент енергії сольватації) мають більш негативні значення, порівняно із такими для еталонного комплексу. Зворотна ситуація спостерігається у випадку дестабілізуючих сил – полярної енергії сольватації і енергії дисперсійних взаємодій. Обидва цих компоненти ентальпії володіли більш позитивним значенням. При чому E_{rb} відрізнялось на $13,52$ ккал/моль, що сумарно і призвело до дещо меншої ентальпії комплексоутворення і, відповідно, меншої енергії зв'язування.

Таблиця 4.3

Загальна енергія зв'язування та її компоненти, розраховані для комплексу HASDI-G2(BCR_ABL1)/BCR_ABL1-50п.о. у ккал/моль

	Evdw	Eel	Epb	Enpol	Edisp	ΔH	$-\Delta S$	ΔG
Average	-399,6	-80,36	154,31	-202,17	354,68	-173,15	18,2	-154,95
SD	6,91	3,44	3,68	2,77	3,58	6,93	0,05	6,93

Порівняно з стабільною взаємодією, якою характеризувався комплекс HASDI-G2(BCR_ABL1)/BCR_ABL1-50п.о., комплекс HASDI-G2(BCR_ABL1)/BCR-50п.о. виявляв ознаки дестабілізації вже з самого початку симуляції (рис. 4.18). Очікувано, що в зоні нижче ділянки зрощення, міжмолекулярні водневі зв'язки були характерні лише під час взаємодії з T32 і T34 – двох пар нуклеїнових основ, які перекривались з такими у відповідній частині нативного BCR. Однак з плином симуляції спостерігались значні конформаційні перебудови як HASDI-G2(BCR_ABL1), так і ДНК дуплексу. Вже через 4 нс моделювання третій лінкер HASDI-G2(BCR_ABL1) втратив майже всі водневі зв'язки з парами основ, на розпізнавання яких він власне був направлений. В подальшому ця ділянка характеризувалась низькою стабільністю взаємодії ДНК дуплексу із селективною частиною ліганду із поступовою повною їх втратою. На 146 нс симуляції G24 одна з основ, яку повинен був розпізнавати HASDI-G2(BCR_ABL1) в цій ділянці, покинула подвійну спіраль.

Останнє відбулось після того як на 119 нс симуляції T30, перша основа після точки зрощення, також покинула дуплекс, що свідчить на користь значного внеску кооперативного ефекту у стабільність взаємодії HASDI-G2(BCR_ABL1)/ціль. Також слід зазначити, що 3'-кінець сенсового ланцюга та відповідна йому частина

антисенсового ланцюга у складі дуплексу до кінця симуляції формували стабільну, дещо зігнуту відносно уявної прямої решти подвійної спіралі, конформацію.

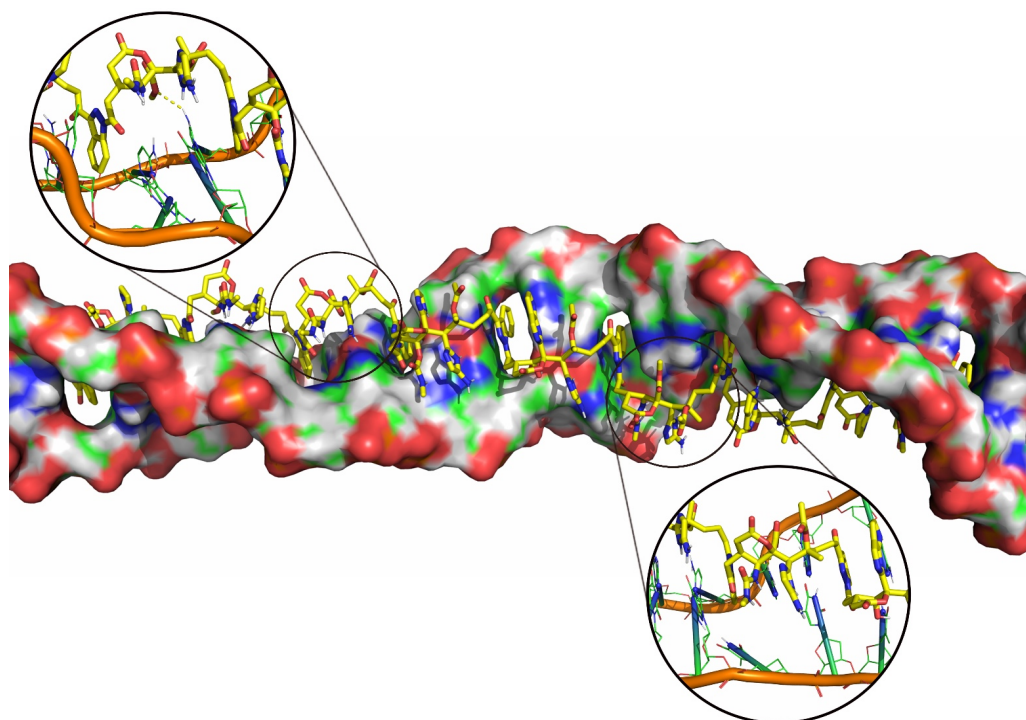


Рисунок 4.18 Загальний вигляд комплексу HASDI-G2(BCR_ABL1) з BCR-50п.о. під час симуляції молекулярної динаміки

Кількість водневих зв'язків між HASDI-G2(BCR_ABL1) і BCR-50п.о. змінювалась з плином симуляції (рис. 4.19). Зокрема, з 1ї по 3тю наносекунду траєкторії їх кількість дорівнювала 29. Після чого спостерігалось їх різке зменшення до близько 26ти. Водночас, до 33ї наносекунди їх кількість відновилась до вихідної. Однак решта симуляції характеризувалась поступовим і невпинним їх зменшенням. Таким чином, на 150й нс моделювання їх кількість дорівнювала 25 без будь-якої тенденції до стабілізації чи росту. Відповідно, в кінці симуляції один лінкер в середньому взаємодіяв з ціллю шляхом формування 3,125 водневих зв'язків із 6ти можливих.

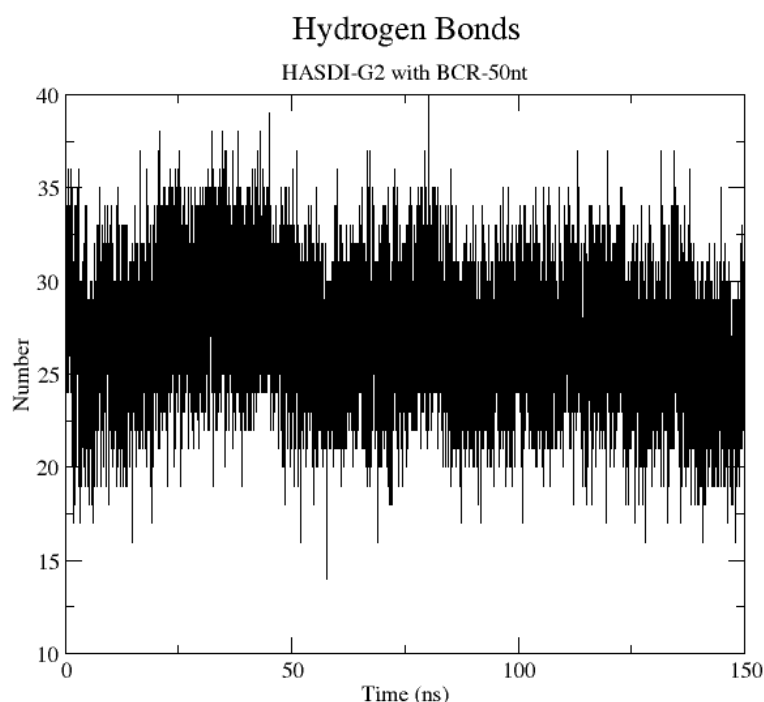


Рисунок 4.19 Залежність кількості водневих зв'язків у комплексі HASDI-G2(BCR_ABL1)/BCR-50п.о. від часу симуляції молекулярної динаміки

Розрахунок вільної енергії зв'язування загалом також підтверджує низький рівень стабільності комплексу HASDI-G2(BCR_ABL1)/ДНК у випадку інтеркаляції ліганду у невідповідну ділянку (таблиця 4.4). Навіть попри те, що 12 із 16 нуклеоснов, які здатен розпізнавати HASDI-G2(BCR_ABL1), відповідають оптимальній послідовності. Загалом, ΔG комплексу HASDI-G2(BCR_ABL1)/BCR-50п.о. була рівною $-146,19 \pm 7,34$ ккал/моль, що на 8,76 ккал/моль є меншим, порівняно з вільною енергією зв'язування HASDI-G2(BCR_ABL1) з BCR_ABL1-50п.о. І є наслідком загально дещо нижчих значень кожного із розрахованих складових ентальпії взаємодії. Водночас, найбільших змін зазнали електростатичний компонент і полярна енергія сольватації – кожен змінився

на 5%. Що, очевидно, є пов'язаним зі зменшенням кількості водневих зв'язків між компонентами комплексу та дестабілізацією власне ДНК дуплексу із виходом деяких із нуклеоснов за межі подвійної спіралі.

Таблиця 4.4

Загальна енергія зв'язування та її компоненти, розраховані для комплексу HASDI-G2(BCR_ABL1)/BCR-50п.о. у ккал/моль

	Evdw	Eel	Epb	Enpol	Edisp	ΔH	$-T\Delta S$	ΔG
Average	-387,29	-76,19	146,54	-196,27	348,27	-164,94	18,75	-146,19
SD	8,14	3,78	4,24	3,08	4,41	7,79	0,88	7,34

Подібно до симуляції молекулярної динаміки комплексу HASDI-G2(BCR_ABL1) з нативним BCR-50п.о., інтеркаляція HASDI-G2(BCR_ABL1) до нативного варіанту ABL1 призвела до відсутності більшості орієнтуючих взаємодій між лігандом та рецептором до ділянки зрощення вже на початку моделювання (рис. 4.20). Лише у випадку одноклеотидного перекриття (T21 та T23) спостерігались класичні взаємодії у повному обсязі.

Однак з плином симуляції спостерігалось наростання конформаційних змін як ліганду, так і ДНК в зоні інтеркаляції, що призводило до швидкого формування неканонічних водневих зв'язків і їх подальшого розриву. Починаючи з 62 нс моделювання третій лінкер повністю втратив усі направлені взаємодії з нуклеосоновами в зоні контакту. Це відбулось попри те, що G25 є нуклеосоною, яку HASDI-G2(BCR_ABL1) є здатним розпізнати в межах «правильного» комплексу. В подальшому взаємодії лінкеру і G25 все ж спостерігались, однак новосформовані водневі зв'язки були вкрай нестабільними та не стосувались цитидину цієї нуклеотидної пари.

Починаючи з 102ї наносекунди симуляції спостерігалась різка втрата всіх орієнтуючих взаємодій між шостим лінкером HASDI-G2(BCR_ABL1) і відповідною йому частиною ДНК дуплексу. Яка, в свою чергу, є розташованою після точки злиття генів BCR і ABL1. Відповідно, кооперативний ефект від часткової невідповідності послідовності HASDI-G2 чітко проявляється і в даному випадку також. Заради справедливості слід зазначити, що до кінця симуляції спостерігався, крім того, короткий період повного відновлення всіх взаємодій між шостим лінкером HASDI-G2(BCR_ABL1) і відповідною частиною ДНК. Однак під час закінчення симуляції взаємодія була відсутньою.

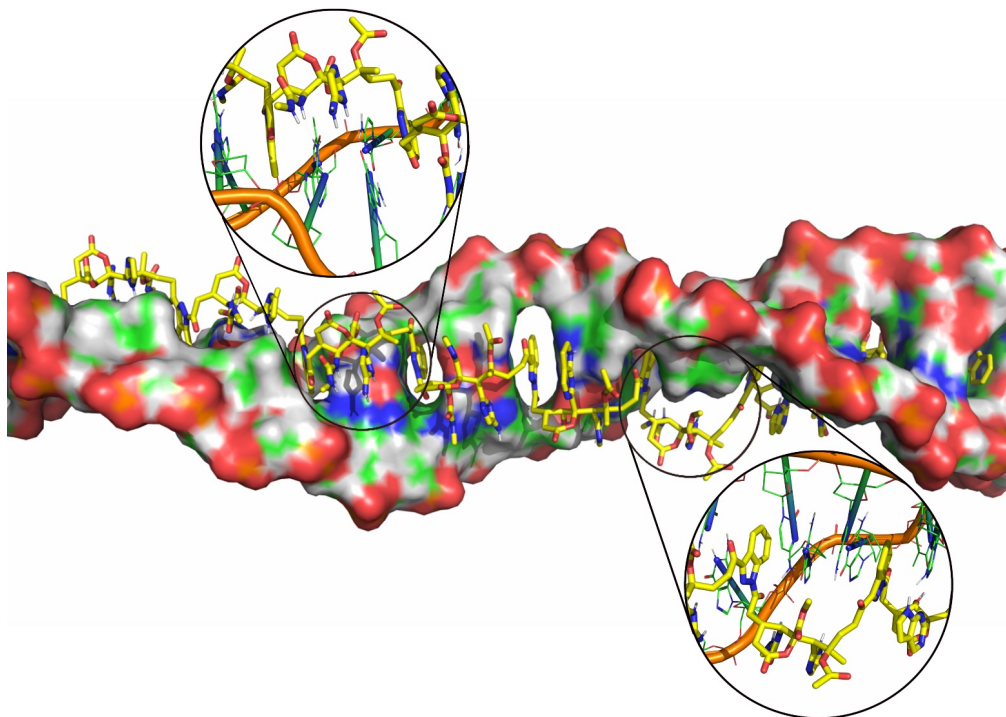


Рисунок 4.20 Загальний вигляд комплексу HASDI-G2(BCR_ABL1) з ABL1-50п.о. під час симуляції молекулярної динаміки

Як можна помітити з рисунку 4.21, загальна зміна кількості водневих зв'язків між ABL1-50п.о. і HASDI-G2(BCR_ABL1) значно коливається з плином симуляції, що підтверджує особливості траєкторії ліганд-рецепторного комплексу описані вище. З початку симуляції і до 112 нс середня кількість водневих взаємодій стрімко падає з 30-31 зв'язку до 25 на комплекс. Після чого спостерігається короточасний ріст їх кількості до 28. Однак починаючи з 143 нс моделювання і до його кінця кількість водневих зв'язків стрімко знижується до 22 на увесь комплекс. Що дорівнює 2,75 зв'язки на один лінкер із 6ти можливих, і є близьким значенням до такого у випадку майже повністю невідповідного комплексу HASDI-G2(EBNA1)/KCNH2-50п.о..

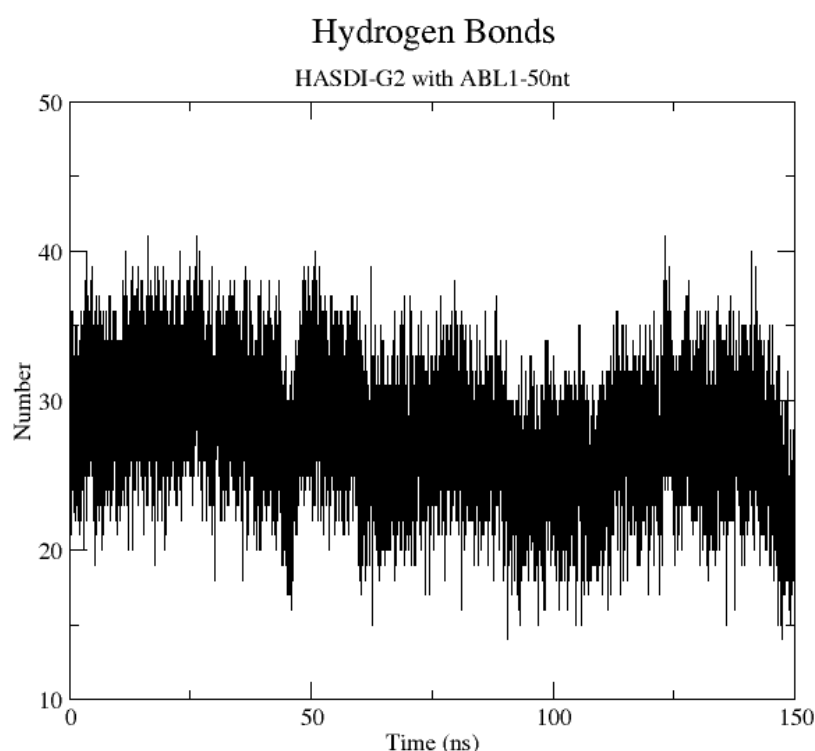


Рисунок 4.21 Залежність кількості водневих зв'язків у комплексі HASDI-G2(BCR_ABL1)/ABL1-50п.о. від часу симуляції молекулярної динаміки

Попри відповідність 13 з 16 пар основ цільовій для HASDI-G2(BCR_ABL1) послідовності, розрахована методом ММ/РBSA енергія зв'язування з ABL1-50п.о. є нижчою за таку у випадку взаємодії ліганду з BCR-50п.о. (таблиця 4.5), де відповідними були 12 із 16 нуклеоснов. В даному випадку ΔG становить $-138,99 \pm 8,36$ ккал/моль, що на 15,96 ккал/моль менше, порівняно з вільною енергією зв'язування еталонного комплексу, або на 7,2 ккал/моль – порівняно із комплексом HASDI-G2(BCR_ABL1)/BCR-50п.о. Останнє є порівнюваним із різницею вільної енергії зв'язування HASDI-G2(BCR_ABL1) з BCR_ABL1-50п.о. і BCR-50п.о.

Енергетичні компоненти ентальпії взаємодії по різному співвідносились як з такими для еталонного комплексу, так і для комплексу HASDI-G2(BCR_ABL1) з ABL1-50п.о. Зокрема, енергія взаємодій Ван-дер-Ваальса мала середнє значення порівняно з цими двома комплексами як крайніми точками. Натомість, неполярна енергія сольватації і внесок дисперсійних сил були наближеними для таких у випадку еталонного комплексу. Водночас, полярна енергія сольватації була близькою до такої для комплексу HASDI-G2(BCR_ABL1) з BCR-50п.о., а електростатичний компонент – основне відображення селективних взаємодій – був нижчим. Однак основний внесок у таку значну зміну вільної енергії зв'язування був отриманий через ріст ентропійної складової більше ніж на 69%, порівняно еталонним комплексом.

Таблиця 4.5

Загальна енергія зв'язування та її компоненти, розраховані для комплексу HASDI-G2(BCR_ABL1)/ABL1-50п.о. у ккал/моль

	Evdw	Eel	Epb	Enpol	Edisp	ΔH	$-T\Delta S$	ΔG
Average	-392,86	-75,12	146,72	-201,61	353,0	-169,88	30,89	-138,99
SD	9,86	4,56	5,0	3,64	5,5	8,28	1,12	8,36

Для розуміння можливостей HASDI-G2 щодо дискримінації цілі серед справді споріднених послідовностей, було проведено ще одне віртуальне дослідження. Ліганд був націлений на мутантний варіант гену KRAS – KRAS_G12S і, зокрема, послідовність 5`-AGTTGGAGCTAGTGGC-3` (центральна частина використаної для генерації дуплексу послідовності 5`-ACTGAATATAAACTTGTGGTAGTTGGAGCTAGTGGCGTAGGCAAGAGTGC-3` [254]). ДНК дуплекс, згенерований на основі фрагменту нативного KRAS (5`-ACTGAATATAAACTTGTGGTAGTTGGAGCTGGTGGCGTAGGCAAGAGTGC-3` (10623–10672), RefSeq: NG_007524.2), натомість, був використаний в ролі контрольної послідовності, яка буде конкурувати за зв'язування HASDI-G2(KRAS_G12S). Ці генетичні варіанти відрізняються єдиною нуклеотидною парою (G→A), яка є причиною амінокислотної заміни і, як наслідок, значних порушень у функціонуванні KRAS [254]. Таким чином, дискримінація однієї з двох послідовностей, які відрізняються між собою лише однією нуклеотидною парою із шістнадцяти, може розглядатись як найскладніше завдання, яке може бути поставлене перед HASDI-G2 у майбутньому.

Симуляція молекулярної динаміки комплексу HASDI-G2(KRAS_G12S) з KRAS_G12S-50п.о. підтвердила сформоване на основі двох попередніх експериментів твердження щодо відсутності значної різниці у стабільності асоціату HASDI-G2/ціль за умови націлювання на конкретну послідовність (рис. 4.22). Зокрема, більшість водневих зв'язків між лігандом і рецептором зберігались протягом всього періоду симуляції. Кожна пара нуклеооснов розпізнавалась шляхом формування здебільшого шести водневих зв'язків. Індазольні кільця стабільно зберігають своє положення в межах початкової інтеркаляції, тобто кожні дві пари нуклеїнових основ. Відповідно, комплекс HASDI-G2(KRAS_G12S)/KRAS_G12S-50п.о. можна вважати, як класично дуже стабільний.

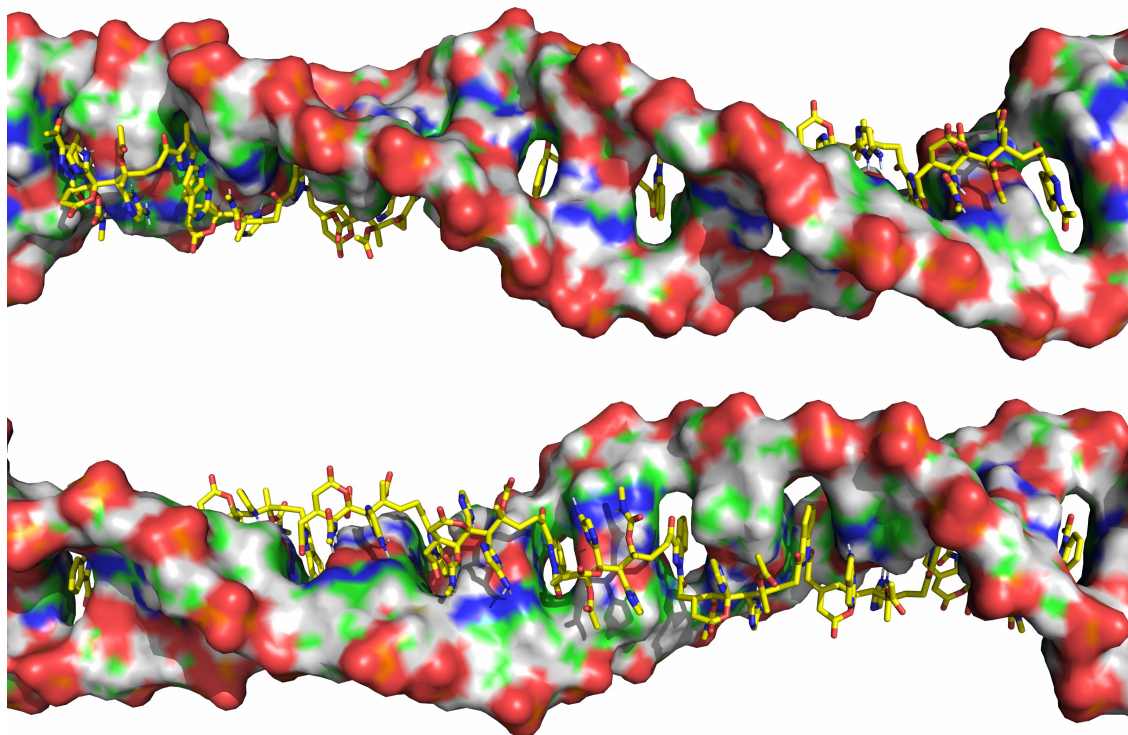


Рисунок 4.22 Загальний вигляд комплексу HASDI-G2(KRAS_G12S) з KRAS_G12S-50п.о. під час симуляції молекулярної динаміки

Кількість водневих зв'язків, розрахована за допомогою внутрішнього функціоналу GROMACS, при цьому, в середньому дорівнювала 35-36 взаємодій (рис. 4.23). Це значення є актуальним для більшої частини симуляції. Лише починаючи з 138 наносекунди спостерігається повільне їх зниження до ~34, яке, однак, не вплинуло на картину, що спостерігалась під час аналізу траєкторії симуляції цього комплексу. Відповідно, на один лінкер, що розпізнає дві пари нуклеїнових основ припадає 4,25-4,5 водневих зв'язки. Це вище значення, порівняно із попередньою версією HASDI, однак в даному дослідженні знаходиться між HASDI-G2(BCR_ABL1) і HASDI-G2(EBNA1) у випадку їх комплексів із послідовностями націлювання.

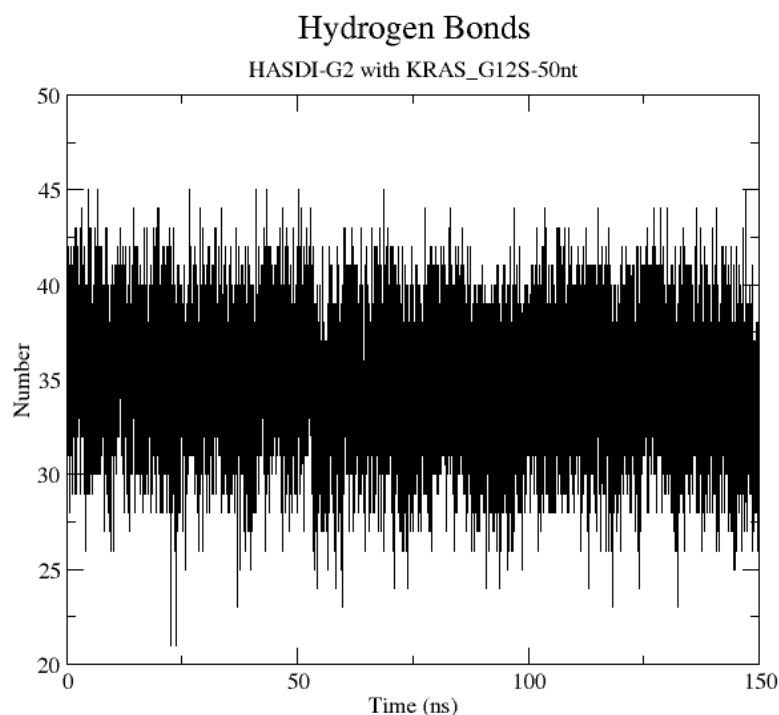


Рисунок 4.23 Залежність кількості водневих зв'язків у комплексі HASDI-G2(KRAS_G12S)/KRAS_G12S-50п.о. від часу симуляції молекулярної динаміки

Розрахована на основі траєкторії симуляції молекулярної динаміки енергія зв'язування загалом свідчить на користь навіть більшої стабільності комплексу HASDI-G2(KRAS_G12S)/KRAS_G12S-50п.о., порівняно із подібними асоціатами і дорівнює $-159,6 \pm 6,61$ ккал/моль (таблиця 4.6). Що є на 4,65 ккал/моль більше, порівняно з HASDI-G2(BCR_ABL1)/BCR_ABL1-50п.о. і на 1,88 ккал/моль – за HASDI-G2(EBNA1)/EBNA1-50п.о. Однак слід зазначити, що більшість енергетичних складових ентальпії є порівнюваними із такими для комплексу HASDI-G2(BCR_ABL1)/BCR_ABL1-50п.о. Ентропійний штраф при цьому має проміжне значення між згаданими вище асоціатами. Єдина значна різниця виражена у енергії електростатичних взаємодій, яка в даному випадку становить $85,44 \pm 3,56$ ккал/моль. Відповідно, саме електростатичний внесок тут є ключовим.

Таблиця 4.6

Загальна енергія зв'язування та її компоненти, розраховані для комплексу HASDI-G2(KRAS_G12S)/KRAS_G12S-50п.о. у ккал/моль

	Evdw	Eel	Epb	Enpol	Edisp	ΔH	$-T\Delta S$	ΔG
Average	-398,67	-85,44	154,23	-202,33	353,66	-178,56	18,95	-159,6
SD	7,04	3,56	3,74	2,64	3,94	6,61	0,04	6,61

Поряд із загальною очікуваною поведінкою HASDI-G2(KRAS_G12S) у комплексі із власне послідовністю націлювання, результати, отримані внаслідок симуляції молекулярної динаміки HASDI-G2(KRAS_G12S) з послідовністю нативного гена KRAS, свідчать на користь неочікувано виняткової селективності HASDI-G2. Зокрема, попри те, що HASDI-G2(KRAS_G12S) був здатен розпізнавати 15 із 16 нуклеотидних пар послідовності 5'-AGTTGGAGCTGGTGGC-3', вже з самого початку симуляції, окрім очікуваної відсутності орієнтуючих взаємодій в районі положення 31ї пари нуклеїнових основ ДНК дуплексу, спостерігалась стрімка дестабілізація взаємодій 4го лінкера ліганду з відповідними йому парами основ (рис. 4.24). Водночас, A27 змінив власне положення відносно площини подвійної спіралі частково покинувши її. Фактично в цій ділянці спостерігалось локальне плавлення з самого початку симуляції.

В значній мірі подібна ситуація спостерігалась у випадку сьомого лінкера HASDI-G2(KRAS_G12S). Починаючи з 7ї наносекунди моделювання T33 втратив класичні Уотсон-Кріківські взаємодії з відповідним йому аденіном і також вийшов за межі площини дуплексу. В подальшому в цій ділянці спостерігались короткотривалі періоди значного збільшення лабільності ліганду із розривом

більшості орієнтуючих взаємодій. Однак всі вони завершувались стабілізацією в конформації, де три основи з чотирьох розпізнаються за допомогою класичного патерну водневих зв'язків, а T33 формує екзотичний водневий зв'язок між 3NH гетероциклу і киснем гідроксилу ліганду. Також слід зазначити, що 32га пара нуклеїнових основ, яка безпосередньо контактує із місцем одонуклеотидної заміни, також втратила більшість із можливих взаємодій з відповідним лінкером і його замісниками. Лише зрідка шостий лінкер формував слабкі водневі зв'язки між R2/R3 і G31. При цьому, значно змінювалась власне конформація лінкерної ділянки.

У випадку решти п'яти лінкерів HASDI-G2(KRAS_G12S) спостерігалось збереження всіх класичних орієнтуючих взаємодій із відповідними їм парами нуклеїнових основ.

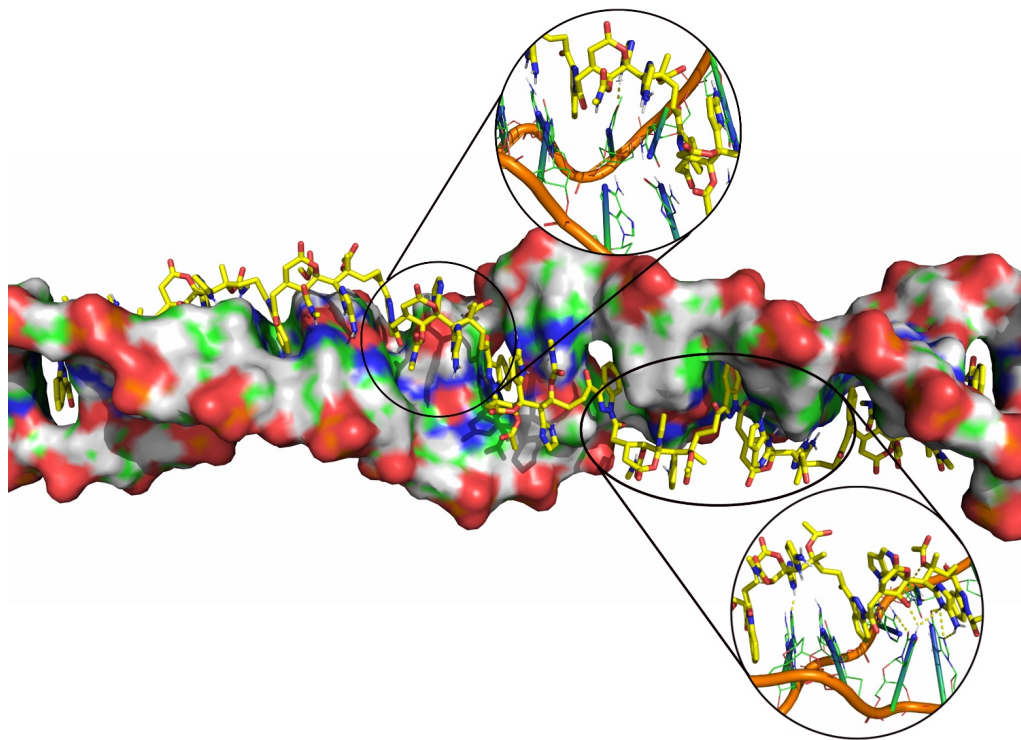


Рисунок 4.24 Загальний вигляд комплексу HASDI-G2(KRAS_G12S) з KRAS-50п.о. під час симуляції молекулярної динаміки

Розглядаючи загальну кількість водневих зв'язків у комплексі HASDI-G2(KRAS_G12S)/KRAS-50п.о. в період з 1ї по 14ту наносекунду, вона дорівнювала 30-31 (рис. 4.25). Однак далі спостерігалось стрімке їх зниження до 26-28, яке зберігалось порівняно рівномірно до кінця симуляції з деякою тенденцією до їх зниження в кінці. Відповідно, в даному випадку кожен окремий лінкер в середньому формував 3,375 водневих зв'язки із відповідною парою основ, що є відносно високим значенням, порівняно з іншими невідповідними комплексами, і може свідчити на користь наявності деякого рівня додаткової стабілізації орієнтуючих взаємодій решти лінкерних ділянок ліганду.

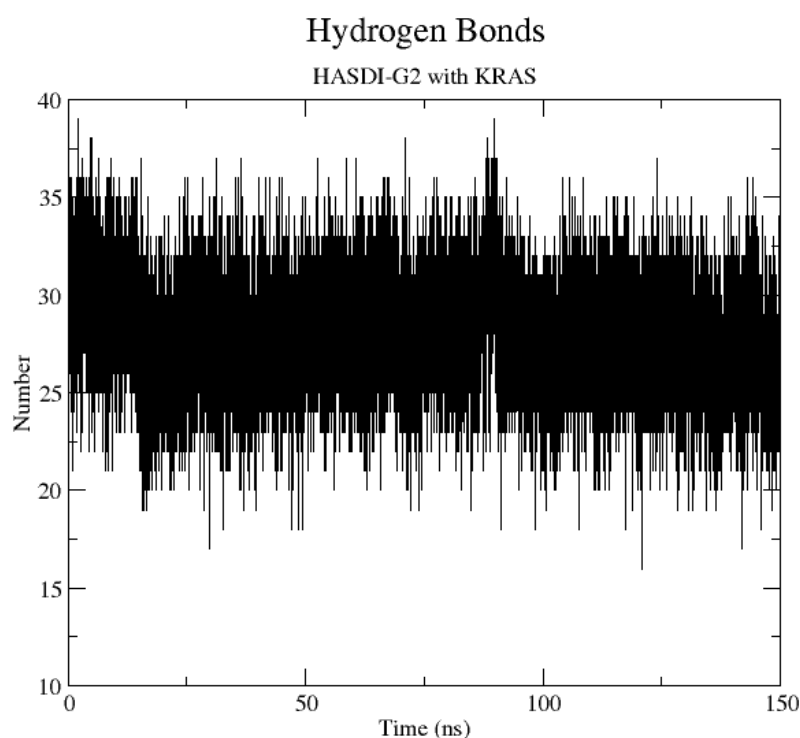


Рисунок 4.25 Залежність кількості водневих зв'язків у комплексі HASDI-G2(KRAS_G12S)/KRAS-50п.о. від часу симуляції молекулярної динаміки

Розрахована методом MM/PBSA вільна енергія зв'язування асоціату HASDI-G2(KRAS_G12S)/KRAS-50п.о. дорівнювала $-146,13 \pm 7,45$ ккал/моль (таблиця 4.7). Що є на 13,46 ккал/моль менше, порівняно з нативним комплексом, однак більше порівняно з комплексом HASDI-G2(BCR_ABL1)/BCR-50п.о., де відповідними були лише 12 із 16 нуклеоснов. Всі стабілізуючі енергетичні компоненти ентальпії, при цьому, очікувано і навіть класично були нижчими порівняно з нативним комплексом. Водночас, внесок полярної сольватації характеризувався меншим енергетичним штрафом, а внесок дисперсійних сил був майже ідентичним. Ентропійний штраф одночасно був значно зниженим.

Таблиця 4.7

Загальна енергія зв'язування та її компоненти, розраховані для комплексу HASDI-G2(KRAS_G12S)/KRAS-50п.о. у ккал/моль

	Evdw	Eel	Epb	Enpol	Edisp	ΔH	$-T\Delta S$	ΔG
Average	-390,37	-75,82	150,44	-198,28	353,58	-160,46	14,33	-146,13
SD	8,54	5,22	6,17	3,64	5,45	7,45	0,05	7,45

Таким чином, HASDI-G2 у трьох віртуальних експериментах, які передбачали сумарно 7 симуляцій (кожна довжиною 150 нс), продемонстрував значний рівень спорідненості саме до послідовностей, на які був націлений. Це виявлялось у високій стабільності комплексів HASDI-G2(EBNA1)/EBNA1-50п.о., HASDI-G2(BCR_ABL1)/BCR_ABL1-50п.о. та HASDI-G2(KRAS_G12S)/KRAS_G12S-50п.о. В кожному із них HASDI-G2 був змінений, відповідно до закону описаного в Додатку 4, для націлювання на EBNA1-50п.о., BCR_ABL1-50п.о. та

KRAS_G12S-50п.о., відповідно. Ці асоціати характеризувались напрочуд високим рівнем однорідності траєкторії симуляції, відсутністю деструктивних змін, високим рівнем орієнтуючих взаємодій, а також схожою і високою розрахованою енергією комплексоутворення – $-157,72 \pm 6,57$ ккал/моль, $-154,95 \pm 6,93$ ккал/моль та $-159,6 \pm 6,61$ ккал/моль, відповідно. Натомість комплекси HASDI-G2(EBNA1)/KCNH2-50п.о., HASDI-G2(BCR_ABL1)/BCR-50п.о., HASDI-G2(BCR_ABL1)/ABL1-50п.о. та HASDI-G2(KRAS_G12S)/KRAS-50п.о. характеризувались відносно низьким рівнем стабільності ліганд-рецепторної взаємодії, значно меншою кількістю водневих взаємодій, частими конформаційними перебудовами та порівняно низькою енергією комплексоутворення – $-97,25 \pm 9,18$ ккал/моль, $-146,19 \pm 7,34$ ккал/моль, $-138,99 \pm 8,36$ ккал/моль, $-146,13 \pm 7,45$ ккал/моль, відповідно. Кожен з цих асоціатів мав в ролі рецептора послідовність ДНК, що відрізнялась від послідовності націлювання HASDI-G2, однак різною мірою: KCNH2-50п.о. мав перекриття в 2 п.о., BCR-50п.о. – 12 п.о., ABL1-50п.о. – 13 п.о., а KRAS-50п.о. – 15 п.о. І хоча стабільність HASDI-G2(EBNA1)/KCNH2-50п.о. очевидно була найнижчою, решта комплексів, попри значну різницю в ступені подібності до послідовності націлювання, загалом характеризувались схожими конформаційними перебудовами та іншими параметрами.

Під час аналізу отриманих даних вдалось відслідкувати схожі патерни в молекулярній динаміці цих конгломератів. У всіх випадках вони є наслідком прояву кооперативного ефекту, який починає виявлятися вже за умови одонуклеотидної різниці між послідовністю націлювання та конкретною ділянкою інтеркаляції і реалізується у вигляді деструктивних змін поза межами зони безпосередньої невідповідності. Більш детально ці зміни проявляються у втраті всіх орієнтуючих взаємодій між певним лінкером HASDI-G2 і нуклеосоновами, навіть за умови їх повної відповідності послідовності націлювання, локальним плавленням ДНК дуплексу і навіть деінтеркаляцією деяких індазольних кілець. Останнього не спостерігалось у випадку HASDI із його феназиновими хромофорами навіть за

умови максимального рівня невідповідності послідовності націлювання ділянці взаємодії поліінтеркалятора. Відповідно, HASDI-G2 може характеризуватись як більш селективна структура, порівняно зі своїм попередником – вершина розвитку концепції селективного ДНК-поліінтеркалюючого агенту на даний момент.

Підсумовуючи, в цій частині роботи було розроблено два покоління селективних ДНК-інтеркаляторів: HASDI і HASDI-G2. Обидва вони були досліджені методом симуляції молекулярної динаміки у порівняльних експериментах, де параметри стабільності ліганд-рецепторного комплексу, утвореного на основі послідовності націлювання, порівнювались із такими для комплексу з по-різному відмінною послідовністю ДНК дуплексу. Високий рівень спорідненості до дволанцюгової спіралі, згенерованої на основі послідовності націлювання, був підтверджений як для HASDI, так і для HASDI-G2, однак селективність взаємодії була різною. Зокрема, попри чіткі ознаки зниження стабільності комплексу при взаємодії HASDI з випадковою частиною геному (різке падіння кількості водневих зв'язків, зниження енергетичного виграшу), ознак деструктивних змін на рівні феназинових кілець не спостерігалось. Врахування цих даних під час розробки HASDI-G2 дозволило знизити рівень неселективної спорідненості. Таким чином, друге покоління HASDI виявилось здатним дискримінувати послідовності, які відрізняються від послідовності націлювання починаючи вже з однієї нуклеотидної пари. Це було підтверджено в численних порівняльних експериментах, в яких в ролі послідовностей націлювання та тестування були використані послідовності нативних та мутованих генів пов'язаних із злоякісним переродженням клітин.

РОЗДІЛ 5

АНАЛІЗ МУТАЦІЙНОГО ПОТЕНЦІАЛУ SARS-COV-2 І РОЗРОБКА/ПОШУК ІНГІБІТОРІВ ЙОГО ПРОНИКНЕННЯ ШЛЯХОМ ЗАСТОСУВАННЯ СТРУКТУРО ОРІЄНТОВАНОГО ДИЗАЙНУ ЛІКІВ

5.1 Аналіз статистично значимих змін у поширенні різних варіантів SARS-CoV-2, пов'язаних з поствакцинальним імунітетом в період пандемії

В умовах роботи з емерджентним патогеном дуже важливо розуміти швидкість його пристосувальних змін. Це, поряд із структурними і функціональними даними, дозволить більш точно вибрати безпосередню ціль/цілі, інгібування якої дозволить успішно протидіяти життєдіяльності патогену, а також точно детермінувати зону, прийнятну для взаємодії з лігандом [255].

З цією метою був проведений аналіз даних, отриманих з відкритих джерел інформації, щодо поширення різних варіантів SARS-CoV-2, ступеня мобільності населення, рівня вакцинованості, захворюваності і смертності. В даному контексті особлива увага зверталась на найбільш активно мутуючий і поширений на той момент штам Дельта. Таким чином, в даній частині роботи проводиться вивчення закономірностей впливу вакцинації на динаміку поширення різних варіантів вірусу SARS-CoV-2 шляхом порівняльного аналізу інформації про зміни епідемічної ситуації та особливості динаміки виявлення варіанту Дельта поряд з іншими штамми від початку пандемії до листопада 2021 року [256].

Як об'єкт дослідження були вибрані три країни на основі їх соціально-економічних особливостей: Індія, Німеччина та Україна. Німеччина є представником економічно розвинених країн. Її економіка є найпотужнішою серед країн Європейського Союзу [257]. Україна та Індія є країнами, що розвиваються, однак сильно відрізняються одна від одної за демографічними, культурними та соціальними параметрами. Крім того, однією з найперших країн, де був виявлений

штам Дельта, була Індія [258].

Аналіз динаміки поширення штамів SARS-CoV-2 в Індії, показав, що вперше варіант Дельта в цій країні був зареєстрований ще у серпні 2020р. Однак його поява в той період не спричинила витіснення інших штамів. Навпаки був витіснений саме він. Далі, в жовтні-листопаді спостерігався невеликий підйом частки сиквенсів даного варіанту (рис. 5.1).

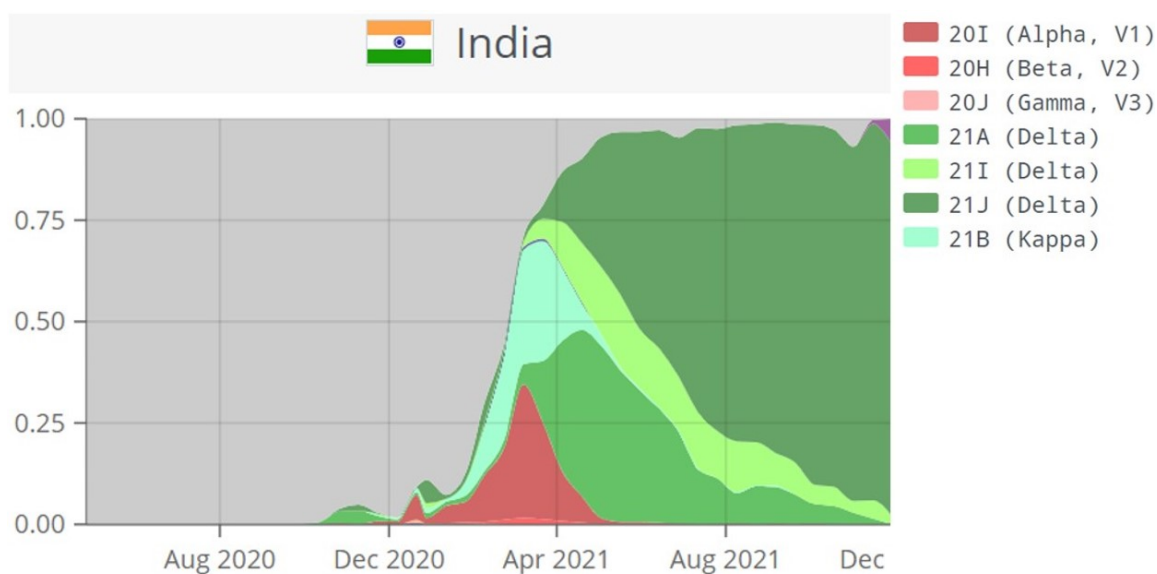


Рисунок 5.1 Зміна частки сиквенсів різних варіантів SARS-CoV-2 у Індії

Цей підйом тривав достатньо короткий період часу і швидко змінився іншим підйомом у грудні 2020р – січні 2021р. Тут, поряд із Дельта варіантом почав поширюватись і варіант Альфа. Можна помітити як обидва ці варіанти співіснують в один проміжок часу. Схожа ситуація спостерігається і в наступному підйомі, який тривав до травня 2021р, різниця лише в масштабі: домінуючі штами майже повністю витіснили інші варіанти, а також варіант Каппа, який деякий період часу був більше представлений у популяції SARS-CoV-2, ніж варіант Дельта. Однак вже з середини березня останній почав витісняти інші штами – спочатку Альфа, а потім і Каппа.

Як видно з рисунку 5.2, на момент кінця 2021р Індія пережила дві хвилі загострення епідемічної ситуації. Перша тривала з січня по червень 2020р та була спричинена вихідним варіантом SARS-CoV-2. Друга характеризувалась набагато більшим рівнем захворюваності і смертності та тривала з кінці березня по червень 2021р. Саме в цей період штам Дельта почав витісняти інші варіанти і його поширення призвело до такого загострення епідемічної ситуації у цій країні.

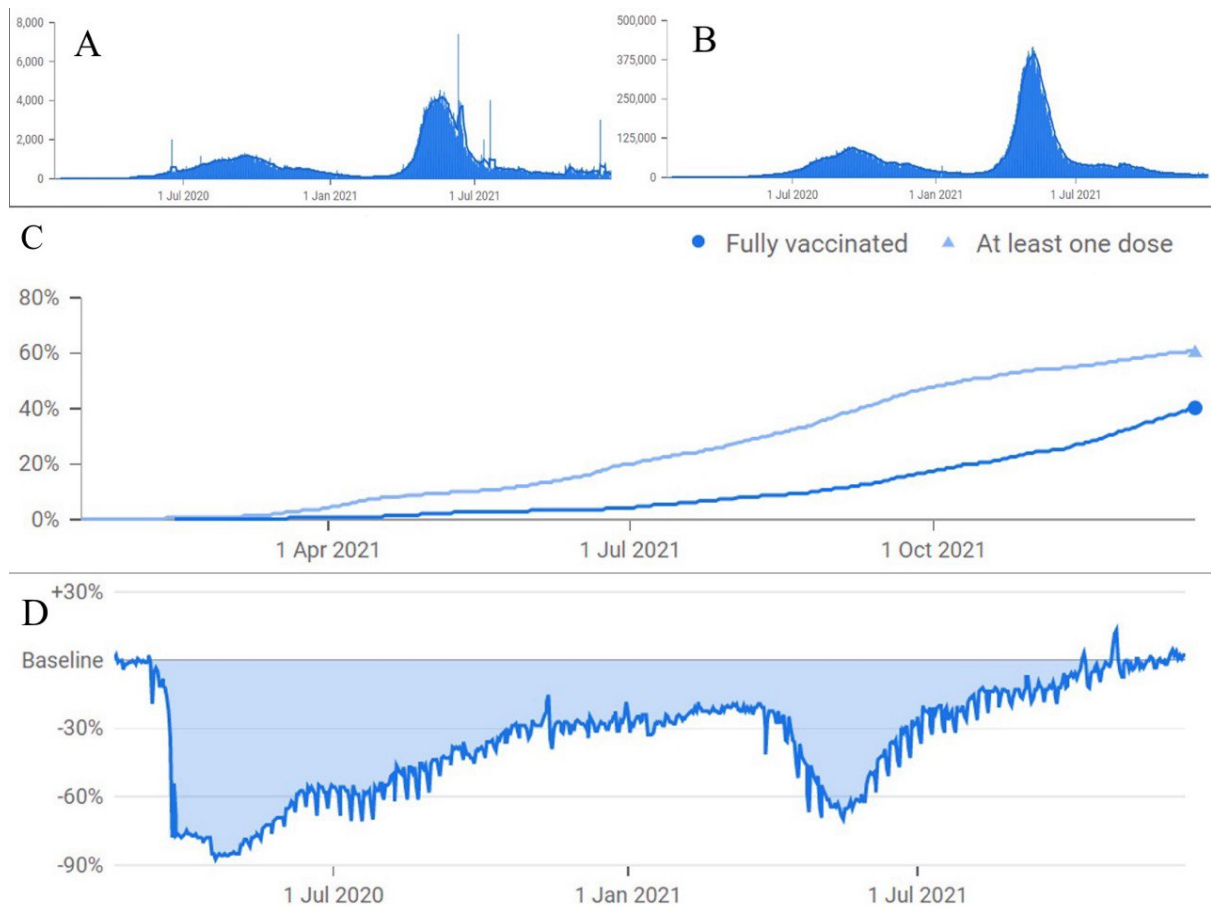


Рисунок 5.2 Залежність смертності (A), захворюваності (B), вакцинованості населення (C) та його мобільності (D) від часу в Індії

Таким чином, дві хвилі загострення епідемічної ситуації, характерні для Індії, були спричинені різними варіантами SARS-CoV-2. Крім того, вони відрізнялись і за

ступенем небезпеки – перша характеризувалась відносно повільним наростанням захворюваності та смертності і таким же повільним затуханням, друга ж швидко прогресувала і була в кілька разів більш потужною, хоча також і більш короткотривалою (рис. 5.2; А, В). Динаміку першої хвилі можна пояснити різким зменшенням ефективності поширення патогену в популяції господаря, що було спричинено потужними антиковідними заходами, запровадженими ще до початку масового поширення SARS-CoV-2 серед населення країни [259]. В дечому інша ситуація спостерігається у випадку другої хвилі. За кілька місяців до її початку були значно послаблені протиепідемічні заходи, що призвело до швидкого поширення серед населення варіантів Альфа, Каппа та Дельта, що однак не спричинило приросту захворюваності. Лише з березня 2021р. можна спостерігати різкий підйом захворюваності. Саме в цей період штам Дельта почав активно витіснити інші варіанти. Однак цікавим є інше – за кілька тижнів до загострення ситуації населення Індії почало вакцинуватись. Після початку другої хвилі, в країні були повернуті більш жорсткі антиковідні заходи, а також пришвидшена вакцинація населення, що разом із поступовим наростанням резистентності населення внаслідок стрімкого природного перенесення інфекції результувалось в швидкому затухання спалаху [260].

Розглядаючи графік поширення різних штамів SARS-CoV-2 в Німеччині (рис. 5.3), можна помітити, що вперше варіант Дельта на території цієї країни був сиквенований ще в кінці квітня 2020р. Подібно до Індії, цей штам не витіснив усі інші варіанти одразу після інвазії. А також до певної міри був витіснений іншими, очевидно більш успішними в конкретних умовах, варіантами. При чому в даному випадку штам Дельта майже не спостерігався в популяції господаря достатньо довгий період часу. В період з жовтня 2020р. по травень 2021р. панували штами Альфа, EU1 та EU2, а також варіанти 98F і 439K. Після чого, починаючи з червня 2021р. варіант Дельта дуже швидко, буквально одномоментно, витіснив основний на той період штам Альфа і став домінуючим.

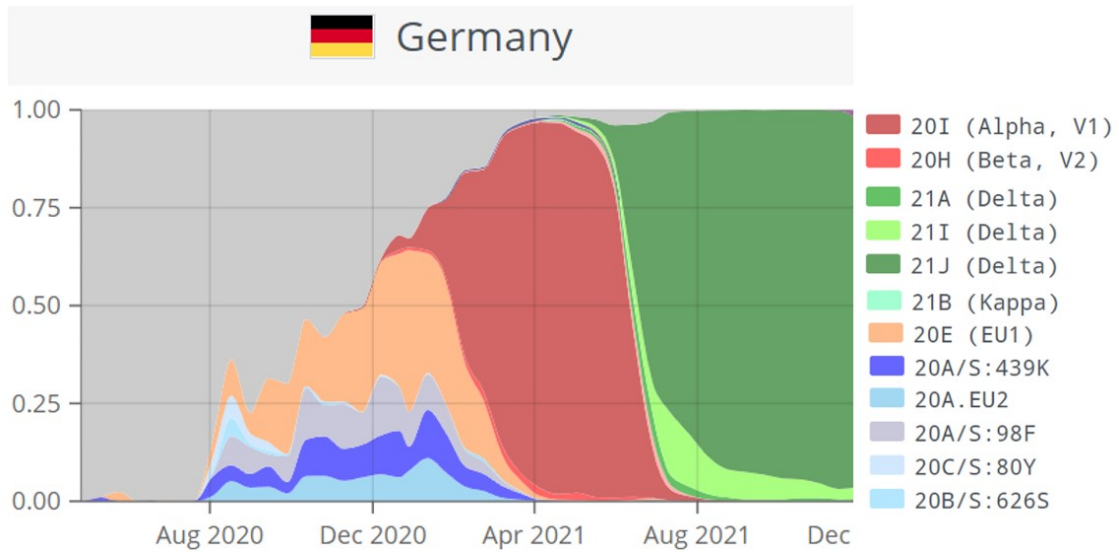


Рисунок 5.3 Зміна частки сиквенсів різних варіантів SARS-CoV-2 у Німеччині

Порівнюючи рисунок 5.4 (А, В) з рисунком 5.3, можна помітити, що у Німеччині, подібно до Індії, усі хвилі захворюваності були спричинені різними варіантами SARS-CoV-2. Перша, коротка, спостерігалась в період первинного проникнення даного вірусу в популяцію людей та була відносно незначною через потужні та ефективні антиковідні заходи, запроваджені в країні з початку пандемії. Друга хвиля була набагато більш масштабною і одночасно гетерогенною. Перший потужний та небезпечний підйом був спричинений варіантами EU1, EU2, 98F і 439K. Після чого спостерігалось деяке зниження захворюваності з подальшим підйомом. В цей період свого поширення зазнав варіант Альфа. Саме на той час припадає початок активної вакцинаційної програми Німеччини, тому зрозумілою є значна різниця в тривалості та небезпечності кінця другої хвилі, у порівнянні з її початком. Масове наростання резистентності населення дозволило пом'якшити наслідки інфікування кожної окремої особи та призвело до припинення поширення патогену. В процесі затухання цього спалаху, дуже швидко варіант Дельта витіснив варіант Альфа і залишився єдиним штамом, який циркулював на території

Німеччини. Через кілька місяців, після значного послаблення протиепідемічних заходів, даний варіант спричинив третю, найбільш потужну хвилю захворюваності, яка однак характеризується порівняно низьким рівнем смертності. Що, по аналогії з варіантом Альфа, можна пояснити наявністю значного рівня індивідуальної резистентності, отриманого внаслідок вакцинації.

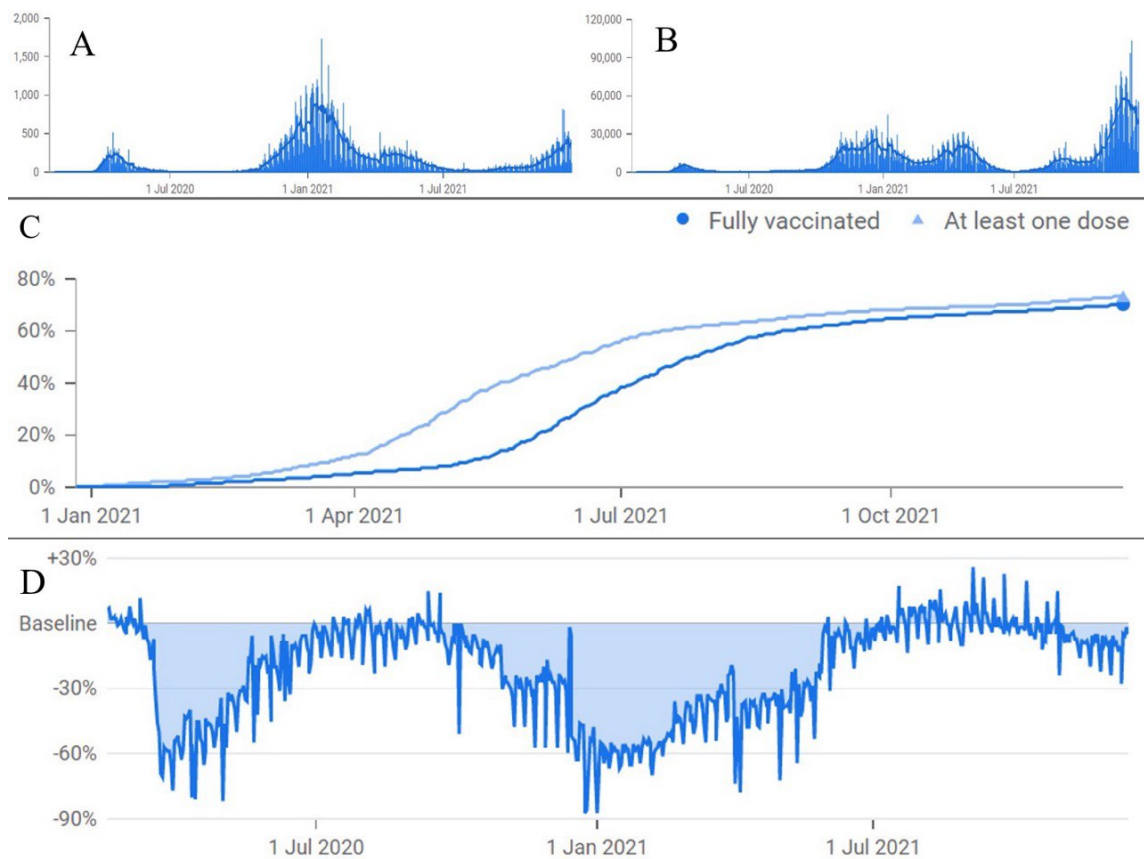


Рисунок 5.4 Залежність смертності (A), захворюваності (B), вакцинованості населення (C) та його мобільності (D) від часу в Німеччині

У даному випадку варіант Дельта, подібно до Індії, також не став панувати в популяції одразу після появи. Ба більше, він був повністю витіснений на довгий період аж до початку наростання групового імунітету, пришвидшеного використанням вакцин. Висока імунність популяції попри поширення частки

варіанту Дельта призвела до пригнічення поширення SARS-CoV-2 в цілому. Однак з часом, послабленням постінфекційного імунітету, збільшенням мобільності населення та наближенням зимового періоду, за якого міняється паттерн контактів людей, штам Дельта все ж спричинив загострення епідемічної ситуації [261].

Відповідно рисунку 5.5, до грудня 2020р. Україна не зустрічалась ні з одним із варіантів, які потребують особливого нагляду. З початку зими, тоді коли в Європі і, зокрема, в Німеччині зазнали поширення варіанти EU1, 98F і 439K, у нашій країні вони були представлені короткотривало і зовсім маленькою часткою. Починаючи з січня 2021р., варіант Альфа стрімко витіснив інші штами та став домінуючим аж до червня, коли, в свою чергу, був витіснений варіантом Дельта. Водночас, в літні місяці 2021р., після повного витіснення варіанту Альфа, спостерігається короткотривале збільшення його частки у пулі сиквенсів українського походження в GISAID [262], що можна пов'язати із сезонним посиленням трудової та туристичної міграції населення нашої країни.

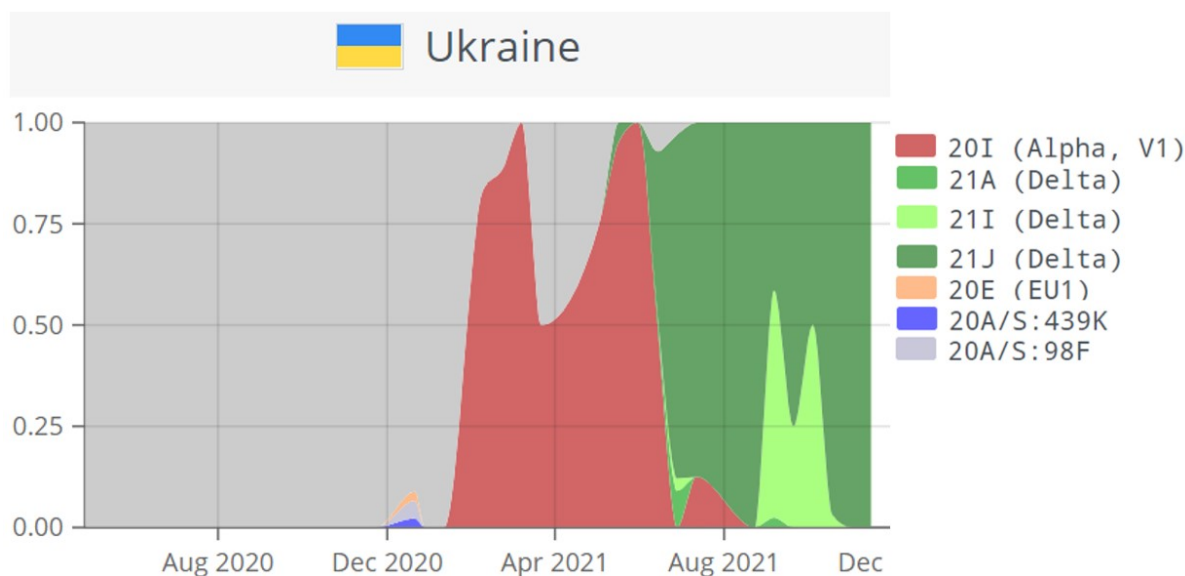


Рисунок 5.5 Зміна частки сиквенсів різних варіантів SARS-CoV-2 в Україні

Порівнюючи захворюваність і смертність з активністю населення (рис. 5.6) можна помітити, що завдяки вчасним та потужним протиепідемічним заходам запровадженими в Україні на початку пандемії, вдалось уникнути значного підвищення рівня захворюваності, спричиненого SARS-CoV-2, подібного до того, що спостерігався в Німеччині. Однак варіант Альфа спричинив спалах дещо пізніше – осінню 2020р. Йому передував довгий період відсутності контролю над переміщеннями населення, що, ймовірно, і було причиною високих показників як захворюваності, так і смертності. Без значного покращення епідемічної ситуації перша хвиля перейшла в другу. Друга хвиля, однак, була спричинена вже іншим штамом – штамом Альфа, який дуже швидко витіснив вихідний варіант. Запроваджені ще на початку першої хвилі антиковідні заходи, наростання колективного імунітету та наближення літнього періоду призвело до значного зниження рівня захворюваності. Паралельно із покращенням епідемічної ситуації, штам Дельта почав витісняти варіант Альфа. Також слід зазначити, що, хоча даний варіант (Дельта) і не реєструвався на території України задовго до появи інших штамів (як у випадку Індії та Німеччини), він вперше був сиквенований в травні 2021р., тобто за місяць до початку активного витіснення інших варіантів. Це узгоджується з початком посилення колективного імунітету внаслідок вакцинації. Третя хвиля почалась стрімко, після кількох літніх місяців високої мобільності населення країни. Подібних до запроваджених в минулому значних обмежень пересування громадян країни не було, що може розглядатись однією з причин більшого рівня захворюваності, який ми можемо спостерігати на момент дослідження. Крім того, виглядає ймовірним і внесок відносно низького рівня вакцинованості населення.

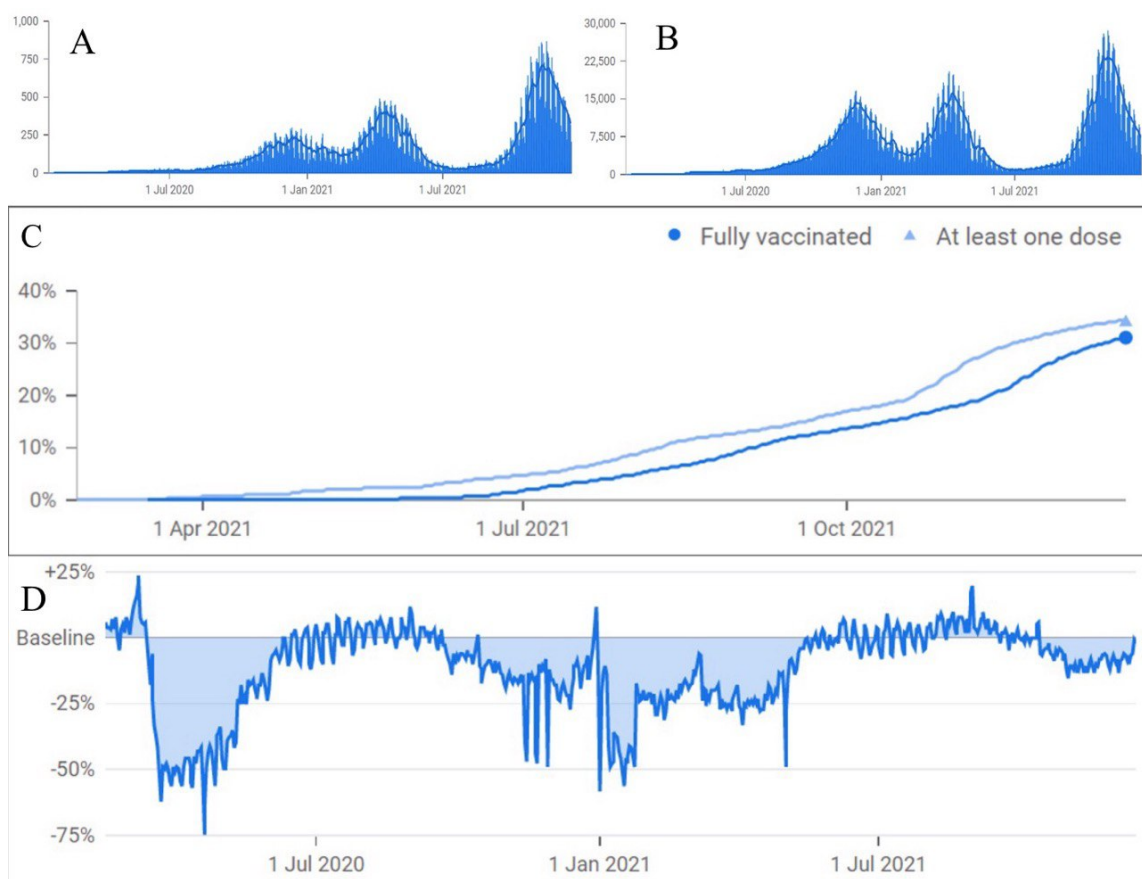


Рисунок 5.6 Залежність смертності (A), захворюваності (B), вакцинованості населення (C) та його мобільності (D) від часу в Німеччині

Підсумовуючи результати цієї частини роботи, можна помітити, що у випадку кожної із розглянутих країн штаб Дельта був наявний на їх території до початку активного його домінування. Він не спричинив витіснення інших варіантів одразу ж. Водночас, з наростанням резистентності населення внаслідок вакцинації у кожному випадку можна було спостерігати незмінно стрімке та впевнене витіснення варіантом Дельта всіх інших штабів SARS-CoV-2. Така його поведінка може свідчити про появу значної переваги у даного варіанту порівняно з іншими у випадку наявності деякої частки вакцинованого населення в країні. З високою ймовірністю це можна пояснити меншим супресуючим впливом поствакцинального імунітету на конкретно цей штаб у порівнянні з іншими, який насамперед

проявляється в більшій кількості так званих проривних (breakthrough) інфекцій [263]. Це підтверджується також характерним для Німеччини та України збільшенням частки Дельта варіанту під час загального зменшення захворюваності – Дельта варіант, завдяки меншій вразливості до дії поствакцинального імунітету, здатен довший час циркулювати в популяції господаря. У випадку ж Індії, де витіснення всіх інших штамів варіантом Дельта супроводжувалось підйомом захворюваності, можна пояснити особливостями культури та високою густотою населення [264]. Крім того, низький рівень захворюваності під час першої хвилі (і, відповідно, постінфекційного імунітету) та вакцинованості, помножений на збільшення мобільності населення, з високою ймовірністю мали вирішальне значення.

Цікавою була ситуація у випадку Німеччини. Перша і друга частина другої хвилі, поряд із тим, що були спричинені різними варіантами SARS-CoV-2 та мали відносно схожий рівень захворюваності, значно відрізнялись за смертністю. Порівняно низький рівень смертності спостерігається в цій країні і за третьою хвилі. Таку особливість динаміки епідемічної ситуації можна пояснити різким нарощенням кількості вакцинованих осіб. Таким чином, попри можливо не такий значний вплив на захворюваність в цілому, вакцинація здатна в кілька разів знизити частоту смертельних випадків – потужно послабити прояв небезпечних для життя наслідків COVID19.

Однак такий ефект спостерігається лише за високого рівня вакцинованості населення. У випадку низького його значення, як, наприклад, у випадку України та Індії, частка смертельних випадків не зазнала значних змін. Таким чином, можна стверджувати, що в залежності від культурних, економічних, демографічних та інших особливостей конкретної країни, існує певний стан популяції, коли вакцинованість є достатньою для надання переваги варіанту Дельта поряд з іншими варіантами і при цьому недостатньою щоб ефективно знизити рівень смертності від даного штаму.

5.2 Ідентифікація консервативної кишені в межах ділянки взаємодії RBD з hACE2, придатної для зв'язування низькомолекулярної сполуки

Ця частина роботи, задля ідентифікації найбільш придатної для розробки і пошуку нових антивірусних сполук ділянки RBD, складалась з двох етапів. В межах першого був проведений аналіз паттерну міжмолекулярних взаємодій у комплексі RBD/hACE2 для варіантів Ухань, Р.1 і Кластер-5 [226]. Де останні два через заміни N501Y і Y453F/N501Y, відповідно, розглядались як представники, які розширюють розуміння інтерфейсу зв'язування. Другий етап базувався на отриманих даних щодо hACE-зв'язуючої ділянки RBD різних варіантів SARS-CoV-2 і включав симуляцію їх молекулярної динаміки задля аналізу цієї ділянки білкової поверхні на предмет наявності перспективного покету для взаємодії з малим лігандом.

Розглядаючи цю частину роботи більш детально, симуляційне дослідження молекулярної динаміки комплексу RBD варіанту Ухань з hACE2 дозволило ідентифікувати амінокислоти Q24, D30, H34, E35, E37, Y41, Y83, K353, D355 і R393 – для hACE2 та K417, Y453, F486, N487, Y489, Q493, Q498, T500, N501, Y505 – для RBD, як безпосередній інтерфейс взаємодії між цими факторами (рис. 5.7).

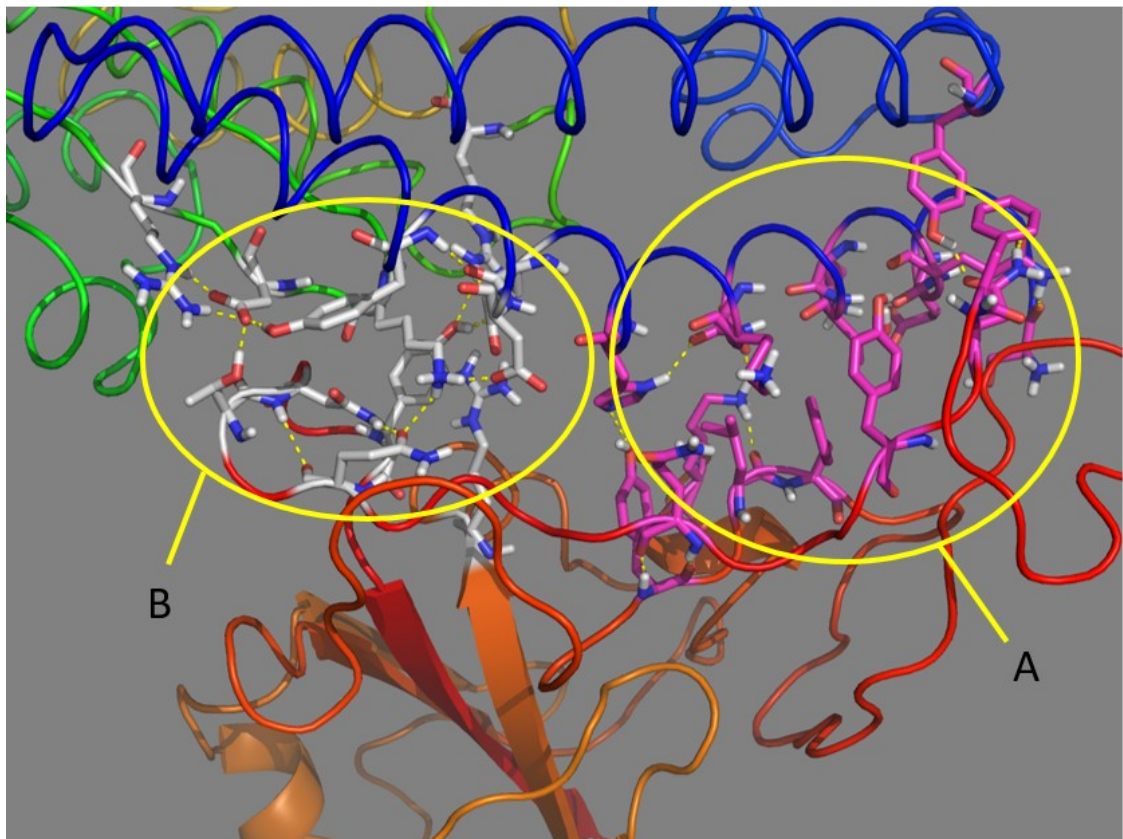


Рисунок 5.7 Ділянка взаємодії немутантного RBD з hACE2. А – перший субінтерфейс взаємодії, В – другий субінтерфейс взаємодії

При чому взаємодія RBD з hACE2 виявилась неоднорідною. В межах одного великого описаного вище інтерфейсу можна чітко виділити два субінтерфейси.

Перший охоплює N-кінець Н1 (Н - helix) hACE2 (тирозин-83 входить до Н2) та відповідну частину RBD (рис. 5.7; А). Основними стабілізуючими взаємодіями тут є водневі зв'язки hACE2 Q24/Y83 + RBD N487, hACE2 H34 + RBD Y453, hACE2 E35 + RBD Q493, а також іонний зв'язок hACE2 D30 + RBD K417. Додаткову стабілізацію забезпечує RBD F486, який екранує водневий зв'язок hACE2 Y83 + RBD N487 від впливу розчинника, а також формує міжплощинний стекінг із цими ж амінокислотами.

Другий субінтерфейс розташовується на С-кінці відносно Н1 hACE2

(амінокислоти K353, D355 та R393 входять до центральної бета-складчастої структури) та охоплює відповідну частину RBD (рис. 5.7; В). Основними орієнтуючими взаємодіями тут є водневі зв'язки hACE2 E37/R393 + RBD Y505, hACE2 K353 + RBD Q498/G502 та hACE2 D355 + RBD T500. Цікаво, що водневий зв'язок hACE2 K353 + RBD G502 спостерігається не між радикалом однієї з них (K353) і пептидною групою іншої, а саме між їх пептидними групами (карбонілом K353 та амідом G502). Це може свідчити про високу стабільність положення частин білків в яких вони знаходяться один відносно одного.

Молекулярна динаміка комплексу мутанта N501Y відрізняється від вищеписаної відповідно до заміни аспарагіну на тирозин (рис. 5.8). Амінокислота hACE2 K353, поряд із збереженим зв'язку із RBD G502, втрачає взаємодію з RBD Q498. Тирозин 501 рекрутує Q498 шляхом утворення внутрішньо ланцюгового водневого зв'язку. Паралельно амід даного глутаміну формує стабільний водневий зв'язок із hACE2 Q42. Виступаючи, таким чином, в ролі посередника. Додатково стабілізація Y501 здійснюється за рахунок формування Т-подібного стекінгу з Y41. В центральній частині ділянки взаємодії амінокислота RBD Q493, яка раніше формувала лише один водневий зв'язок із hACE2 E35 взаємодіє також з hACE2 K31, чим додатково фіксує своє положення. Більш впорядкований радикал лізину 31 формує, крім того, стабільний, не характерний для немутантного RBD, зв'язок з карбонілом F490. Натомість, взаємодія hACE2 H34 з RBD Y453 проявлялась лише в Т-подібному стекінгові без утворення водневого зв'язку. Решта взаємодій комплексу N501Y RBD/hACE2 були подібними до таких для немутантного RBD.

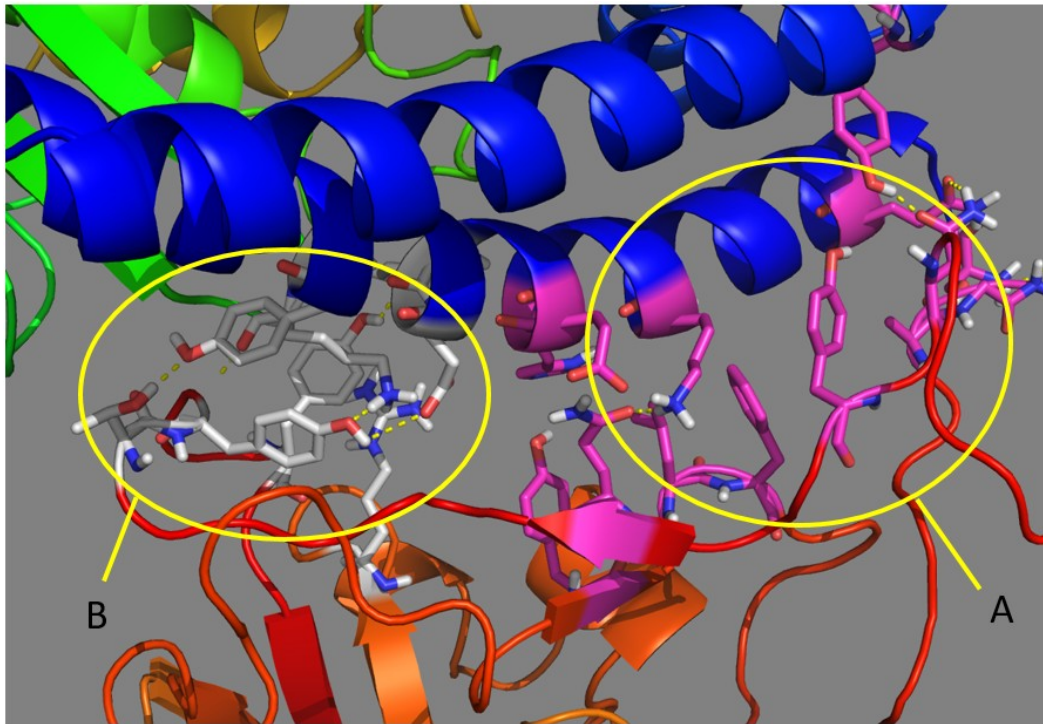


Рисунок 5.8 Ділянка взаємодії RBD із заміною N501Y з hACE2. А – перший субінтерфейс взаємодії, В – другий субінтерфейс взаємодії

Схожа ситуація, але з деякими змінами, спостерігається і у випадку N501Y + Y453F. Через заміну тирозину на фенілаланін в положенні 453, будь-які орієнтуючі взаємодії не могли бути утворені в цій ділянці. Крім того, згаданий додатковий водневий зв'язок hACE2 K31 + RBD F490, характерний для мутанта N501Y, в даному випадку також не спостерігався.

Динаміка загальної кількості утворених розглянутими варіантами RBD водневих зв'язків вказує на відносну стабільність їх валової кількості з плином симуляції (рис. 5.9). У всіх трьох випадках їх кількість коливалась в районі 9-10. Однак загальний рівень взаємодій між RBD та hACE2, на що вказує динаміка попарних контактів $\sim 3,5$ А, є дещо вищим для комплексу немутантного S1 рецептор розпізнаючого домену. Водночас, мутант N501Y відзначається більшою стабільністю цього пулу взаємодій.

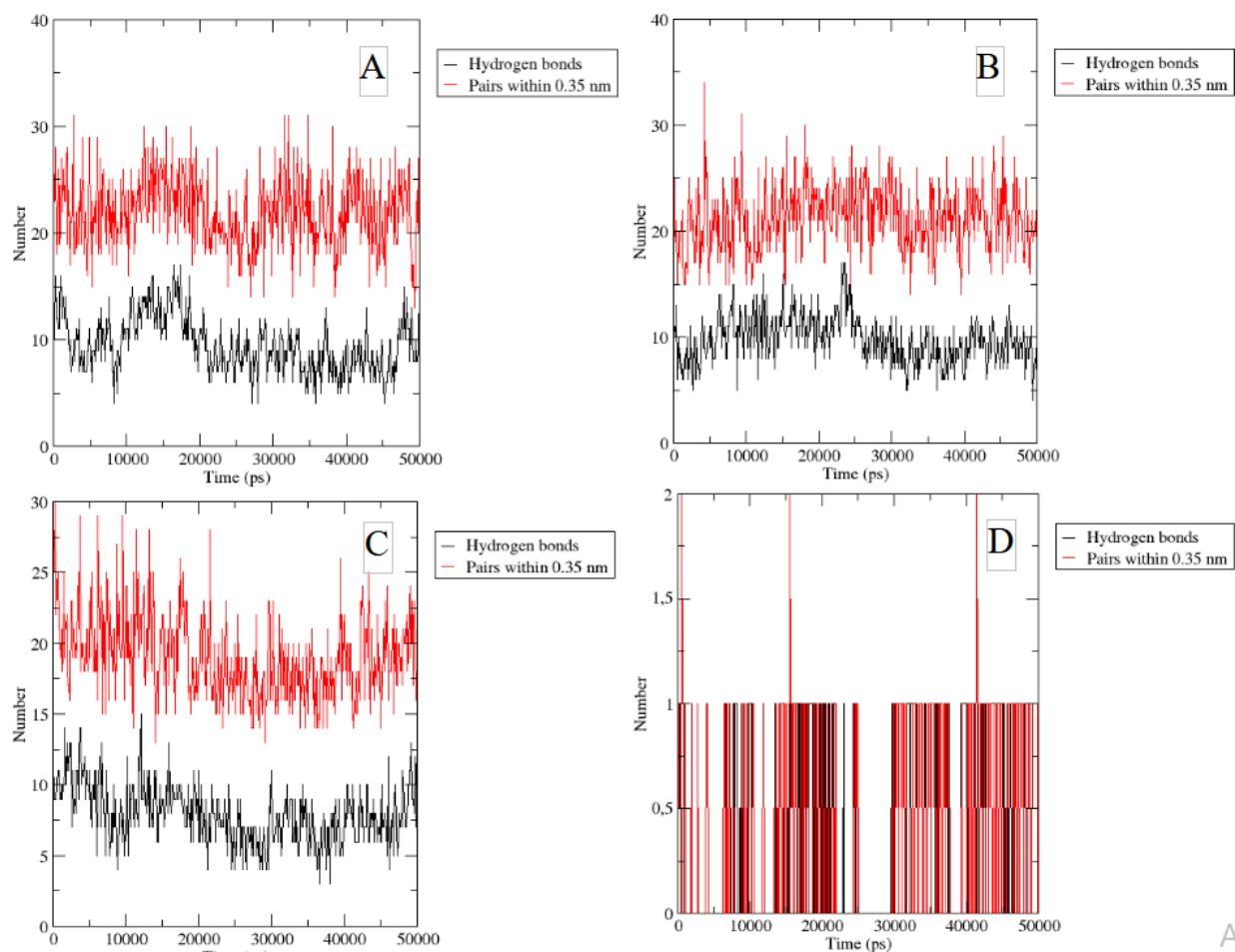


Рисунок 5.9 Залежність кількості утворених водневих зв'язків від часу симуляції. А – між диким варіантом RBD та hACE2, В – між N501Y варіантом RBD та hACE2, С – між N501Y + Y453F варіантом RBD та hACE2, D – між RBD 501 та карбонілом G496/Y495

Таким чином, молекулярна динаміка розглянутих RBD, попри заміни, в загальному була подібною: валова кількість водневих зв'язків та загальна кількість попарних контактів коливалась схожим чином для всіх варіантів. Лише деяка більша стабільність та кількість взаємодій 0,35 нм, яка спостерігалась протягом симуляції молекулярної динаміки немутантного RBD, може, з деякою ймовірністю, говорити про його більший потенціал до утворення комплексу з hACE2. Попри це, ретельний

аналіз траєкторії молекулярної динаміки всіх трьох варіантів говорить про наявність значних змін в конфігурації інтерфейсу взаємодії RBD/hACE2.

Зокрема, заміна N501Y призвела до неможливості утворення водневого зв'язку hACE2 K353 + RBD Q498. При цьому, вакантна амінокислота Q498 стабільно взаємодіяла в ролі донора водневого зв'язку з гідроксилом тирозину 501 та в ролі акцептора/донора із hACE2 Q42. Таким чином, у випадку згаданої заміни в ділянці виділеного тут С-кінцевого субінтерфейсу, взаємодії формуються на один водневий зв'язок більше. Крім того, тирозин, в порівнянні із аспарагіном, містить набагато більшу за площею спряжену електронну хмару, тому Т-подібний стекінг з Y41 також буде впливати на стабільність комплексу розглянутого варіанту з hACE2. Також, поряд із формуванням додаткових взаємодій між hACE2 K31 та RBD Q493/F490, втрачається контакт hACE2 H34 з RBD Y453, що також в результаті додає один теоретичний водневий зв'язок.

Молекулярна динаміка комплексу варіанта N501Y + Y453F в загальному була дуже схожою до такої для N501Y. Однак додаткова заміна тирозину на фенілаланін в 453 положенні призвела до збільшення гідрофобності цієї ділянки. Крім того, відсутність гідроксилу в пара-положенні зменшило розмір самого радикалу, чим потенціувало hACE2 H34 до утворення міжплощинного стекінгу з RBD Q493. Що, в свою чергу, стабілізувало водневі зв'язки Q493 з hACE2 K31/E35 та зменшило ймовірність взаємодії лізину 31 з RBD F490.

Подальший аналіз топологічних особливостей розглянутих варіантів RBD дозволив ідентифікувати виражену каверну між описаними вище субінтерфейсами взаємодії (рис. 5.10). Вона формується залишками R403, Y453, Q493, S494, Y495, G496, F497, Q498, N501 та Y505 варіанту Ухань (рис. 5.10; В). Відповідно, варіанти із замінами N501Y та Y453F в даних положеннях мають інші амінокислоти. Випадково склалось, що ці мутації знаходяться в протилежних кінцях виділеної кишені (рис. 5.10; А). Y453F слабо впливає на загальну топологію цієї ділянки, однак відсутність високополярної гідроксильної групи робить цю частину кишені

набагато більш гідрофобною і, як наслідок, не здатною до утворення водневих зв'язків.

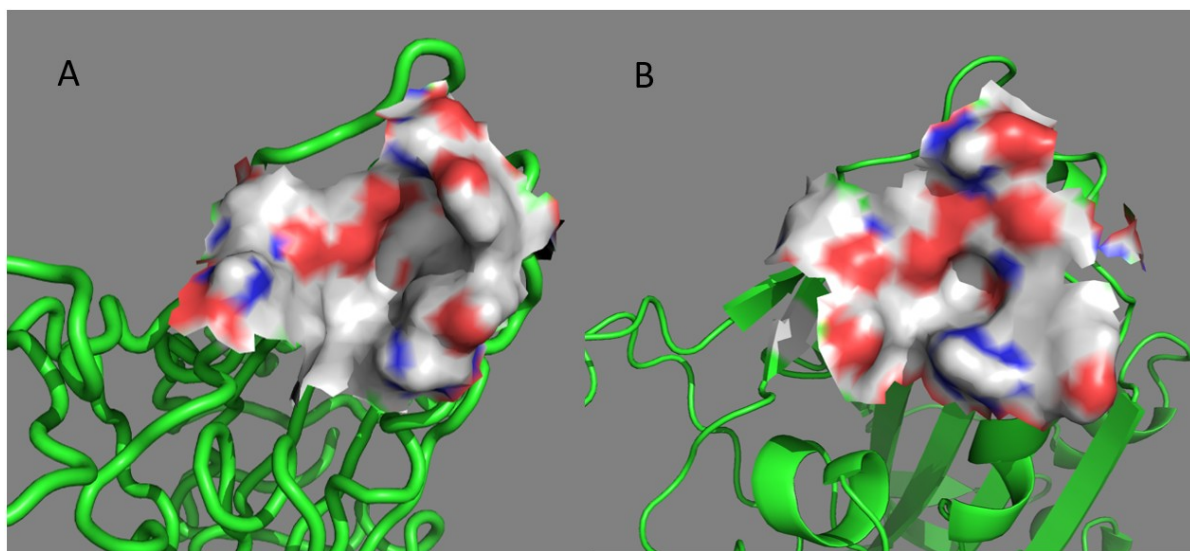


Рисунок 5.10 Кишеня RBD. А – мутант N501Y + Y453F, В – незмінений RBD

Натомість, N501Y не є настільки синонімічною заміною. Через більший розмір радикалу та наявність парагідроксилу дана амінокислота формує два рівнозначних водневих зв'язки з карбонілами пептидних груп G496 та Y495, чим фіксує своє положення та частково перекриває ідентифікований покет (рис. 5.11). Додаткова стабілізація Y501 забезпечується міжплощинним стекінгом із Y505. Однак ця амінокислота може знаходитись і в іншій, не фіксованій водневими зв'язками, формі (рис. 5.11; А).

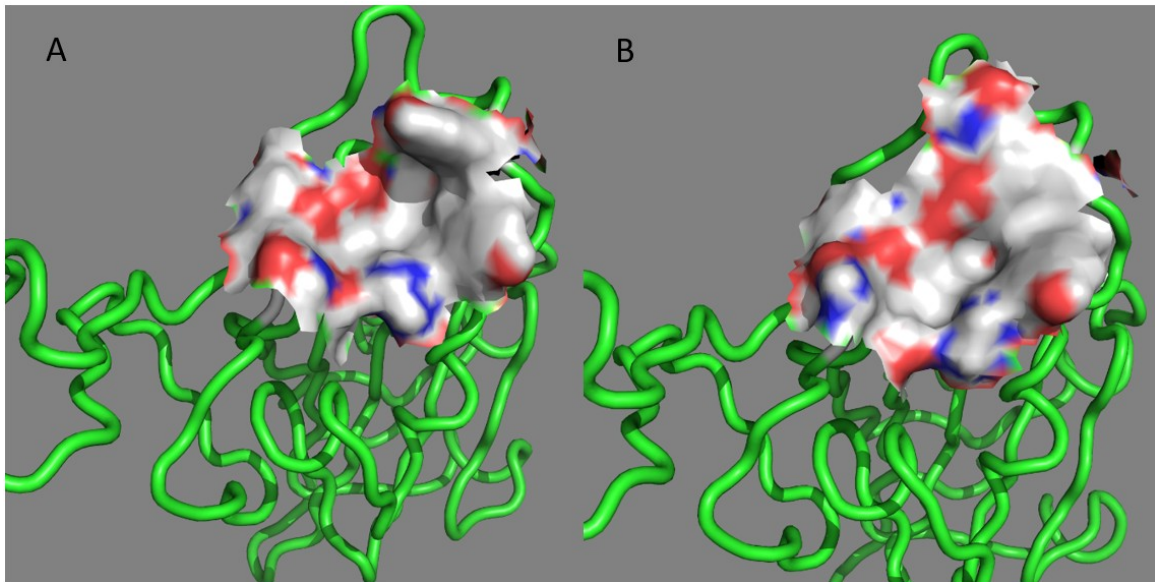


Рисунок 5.11 Кишеня RBD N501Y. А – невпорядкована форма Y501, В – фіксована форма Y501

У такому випадку її радикал прямо контактує з розчинником і не формує впорядкованих взаємодій з рештою RBD. Співвідношення цих двох форм в цій симуляції було $\sim 3:1$ (рис. 5.9; D).

Узагальнюючи, розглянуті в цій частині роботи, характерні для різних варіантів SARS-CoV-2, заміни в межах функціонально активної ділянки RBD знаходяться безпосередньо в ідентифікованій кишені та можуть достатньо сильно модифікувати її топологічні особливості. Зокрема, N501Y призводить до зменшення лінійних розмірів кавеоли, оскільки тирозин в цьому положенні формує два стабільних взаємовиключних водневих зв'язки із G496 та Y495, чим потужно стабілізує своє положення та потенційно перешкоджає взаємодії кандидата на ліки із RBD в цій ділянці. Хоча слід зазначити, що з плином моделювання було виявлено і більш рідкісну конформацію цієї амінокислоти. В близько 25% симуляції залишок тирозину 501 не формував орієнтуючих взаємодій з рештою білка, чим відкривав близько 20% кишені для потенційного інгібітора. Тому, можна стверджувати, що у

випадку пошуку кандидатів на лікарські засоби, направлених на цю ділянку, є сенс використовувати в ролі рецептора як мінімум дві конформації RBD.

Заміна Y453F не впливала настільки сильно на топологію кишені. Тирозин та фенілаланін є подібними за своїми характеристиками амінокислотами. Однак відсутність пара-гідроксилу, як було сказано вище, призвела до зменшення полярності відповідної частини каверни і, як наслідок, до відсутності орієнтуючих взаємодій. Таким чином, потенційно активна відносно дикого варіанту RBD сполука, яка в тому числі орієнтована на утворення водневих зв'язків в цій ділянці, може достатньо сильно змінювати свою активність у випадку варіанта Y453F.

5.3 *De novo* розробка інгібітора RBD S-глікопротеїну, потенційно активного проти чотирьох штамів SARS-CoV-2

На основі даних щодо чіткої наявності в межах hACE2-зв'язуючої ділянки RBD покету, придатного для взаємодії з низькомолекулярним інгібітором, в цій частині роботи було розроблено сполуку, здатну до стабільної взаємодії з цією частиною поверхні RBD [265]. Для цього був застосований ітераційний підхід, який базувався на методі симуляції молекулярної динаміки. Відповідно до також отриманих попередньо даних про топологічні зміни в тій частині цільового білка, які супроводжують еволюцію SARS-CoV-2 під час його адаптації до людини як господаря, процес розробки включав тестування отриманого інгібітора додатково на трьох варіантах патогену – Омикрон, Дельта та Кластер 5.

Представлена на рисунку 5.12 структура є наслідком більше ніж 700 ітерацій. Вона має молекулярну масу 337 Да, є похідним тіазолу із замісниками в другому і четвертому положенні та за своїми топологічними особливостями відповідає кишені RBD, розташованій в ділянці взаємодії RBD/hACE2, що була ідентифікована в попередній частині роботи. Замісник в другому положенні тіазолу є орієнтованим на взаємодію з E406 та Y495, з якими, завдяки наявності карбонільної та карбамідної груп, утворює стабільні водневі зв'язки. Циклопропіл в даному випадку слугує і як

структурний елемент, що забезпечує оптимальне взаємне розташування полярних груп, і як ліпофільний замісник. Замісник в четвертому положенні тiazолу містить дві карбонільні групи та розроблявся направленим на взаємодію з R403 та G496. В даному випадку розміщений між полярними угрупованнями атомів ізопропіл занурюється в невелику ліпофільну кишеню між цими двома амінокислотами. Власне тiazольне кільце є орієнтованим на міжплощинний стекінг з R403 та Y453.

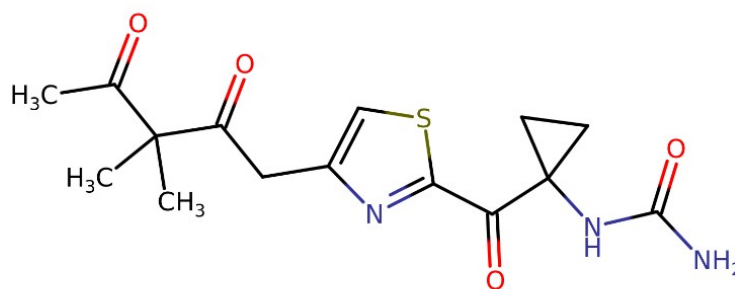


Рисунок 5.12 Структура розробленого інгібітора RBD

Протягом всіх 100 нс моделювання розроблена молекула стабільно взаємодіяла з RBD вихідного варіанту SARS-CoV-2 в ділянці зв'язування. Усі передбачені на стадії розробки водневі зв'язки були збережені. Особливої уваги заслуговує конфігурація R403 та E406 один відносно одного та відносно інгібітора (рис. 5.13; синє та помаранчеве кола, відповідно). Ці дві амінокислоти, окрім потужних електростатичних взаємодій, додатково стабілізують та детермінують своє взаємне положення шляхом утворення двох водневих зв'язків між гуанідиною та карбонільною групами. В такій витягнутій конфігурації формується ділянка з донорами водневого зв'язку з однієї сторони та акцептором з іншої. Таким чином, утворюється сприятливе оточення для вискоєфективної взаємодії інгібітора з RBD, за якої розроблена молекула утворює один водневий зв'язок між карбонільною групою та R403 і два водневі зв'язки між карбамідною

групою та E406. Додаткову стабілізацію забезпечують водневі зв'язки з Y495 і G496 (рис. 5.13; червоне коло). Крім того, тiazольне кільце з обох сторін від його площини контактує з гуанідиною групою R403 та фенільною групою Y453 (рис. 5.13; чорне коло), формуючи таким чином стабільний стекінг з цими амінокислотами.

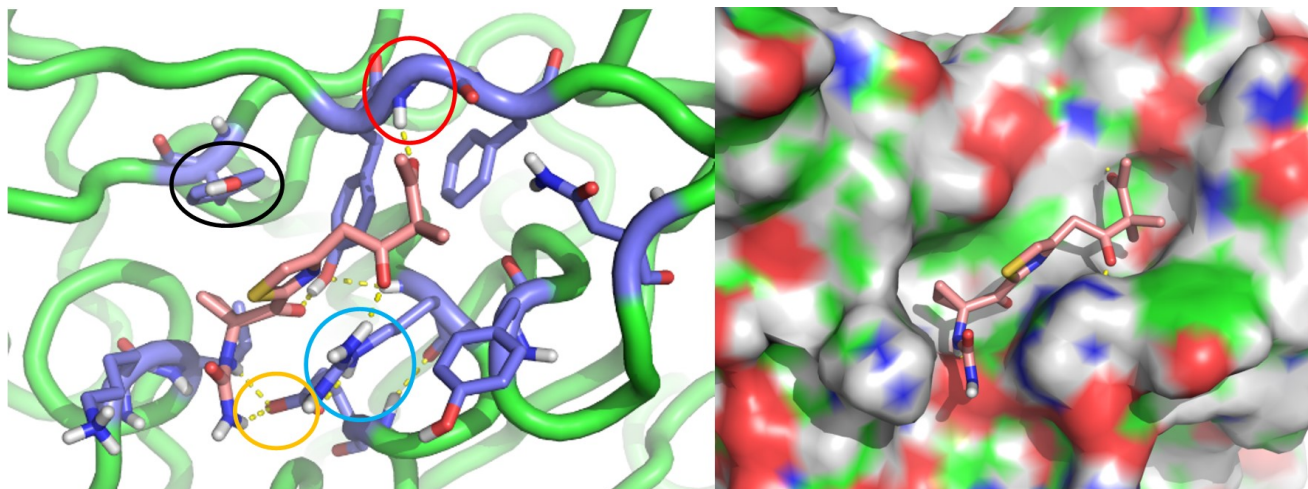


Рисунок 5.13 Розроблений інгібітор в комплексі з RBD вихідного варіанту SARS-CoV-2. R403, E406, Y453 і G496 позначені кольоровими колами – синім, помаранчевим, чорним та червоним, відповідно

Середньоквадратичне відхилення розробленого інгібітора в комплексі з RBD коливається в межах 1,5 Å без різких аберацій, що свідчить про високу стабільність положення розробленої речовини відносно RBD (рис. 5.14). А це, в свою чергу, свідчить про загально високу спорідненість ліганду до рецепторної молекули. Додатковим підтвердженням цього є розподіл лабільності між різними амінокислотами RBD. Амінокислоти R403, E406, Y453, Y495 та G496, тобто ті, які безпосередньо взаємодіють з розробленою молекулою, незмінно мають мінімальне RMSF порівняно з іншими поверхневими амінокислотами. Що свідчить про більшу стабільність оточення розробленого інгібітора.

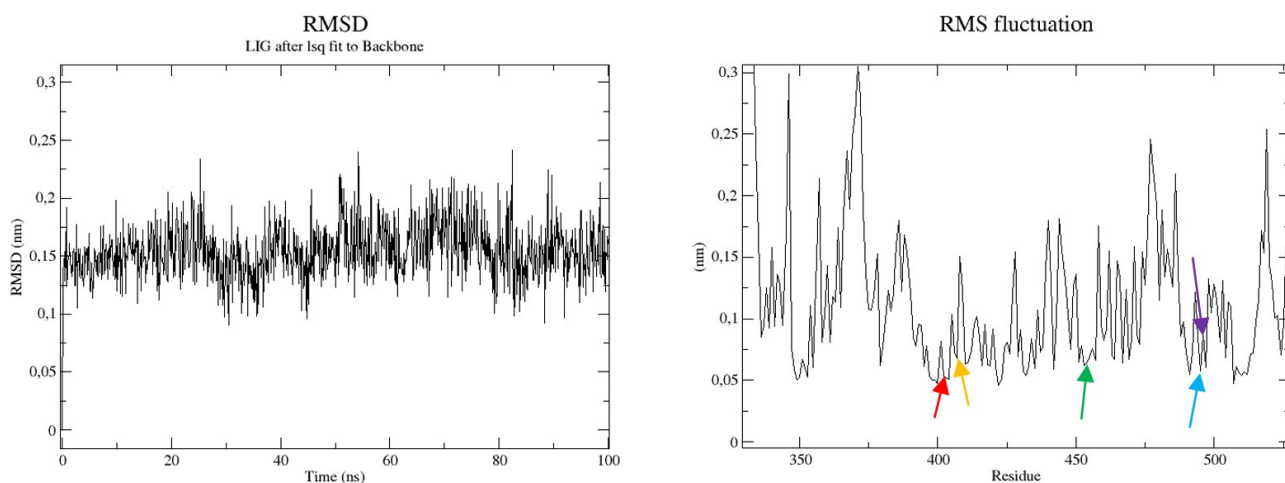


Рисунок 5.14 Середньоквадратичне відхилення ліганд-рецепторного комплексу та RMSF RBD вихідного варіанту SARS-CoV-2. Амінокислоти R403, E406, Y453, Y495 та G496 позначені червоною, жовтою, зеленою синьою і фіолетовою стрілками, відповідно

Вільна енергія взаємодії розробленої речовини з RBD варіанту Ухань становить $-18,27$ ккал/моль, що повністю узгоджується з описаними вище результатами та втретє підтверджує високий потенціал даного низькомолекулярного інгібітора до утворення стабільного комплексу із рецепторзв'язуючим доменом вихідного варіанту SARS-CoV-2.

В комплексі з RBD інших трьох варіантів SARS-CoV-2 розроблена речовина, загалом, продемонструвала подібну поведінку як і у випадку варіанту Ухань (рис. 5.15). Зокрема, всі згадані вище водневі зв'язки спостерігаються для всіх трьох моделювань та протягом всього періоду симуляції. Однак, в залежності від варіанту і, відповідно, конкретного амінокислотного оточення, виникали деякі зміни в загальному паттерні взаємодій. Зокрема, розроблена речовина в комплексі з RBD Кластер 5 стабілізувалась стекинговою взаємодією не з Y453, а з F453, що, попри наявність деякої різниці в полярності радикалу, не вплинуло ні на положення ліганду, ні на інші взаємодії. У випадку варіанту Оміврон, через набагато більшу

кількість несинонімічних замін, різниця є більш помітною. Зокрема, спостерігається періодична короткочасна поява слабких водневих зв'язків з H505 та N417. Крім того, наявність більш масивної амінокислоти в положенні 501 – тирозину, замість аспарагіну, призводить до зміщення кінцевого карбонілу. Поряд з цим канонічні взаємодії з R403, E406, Y453 та Y495 не зазнають ніяких значних змін.

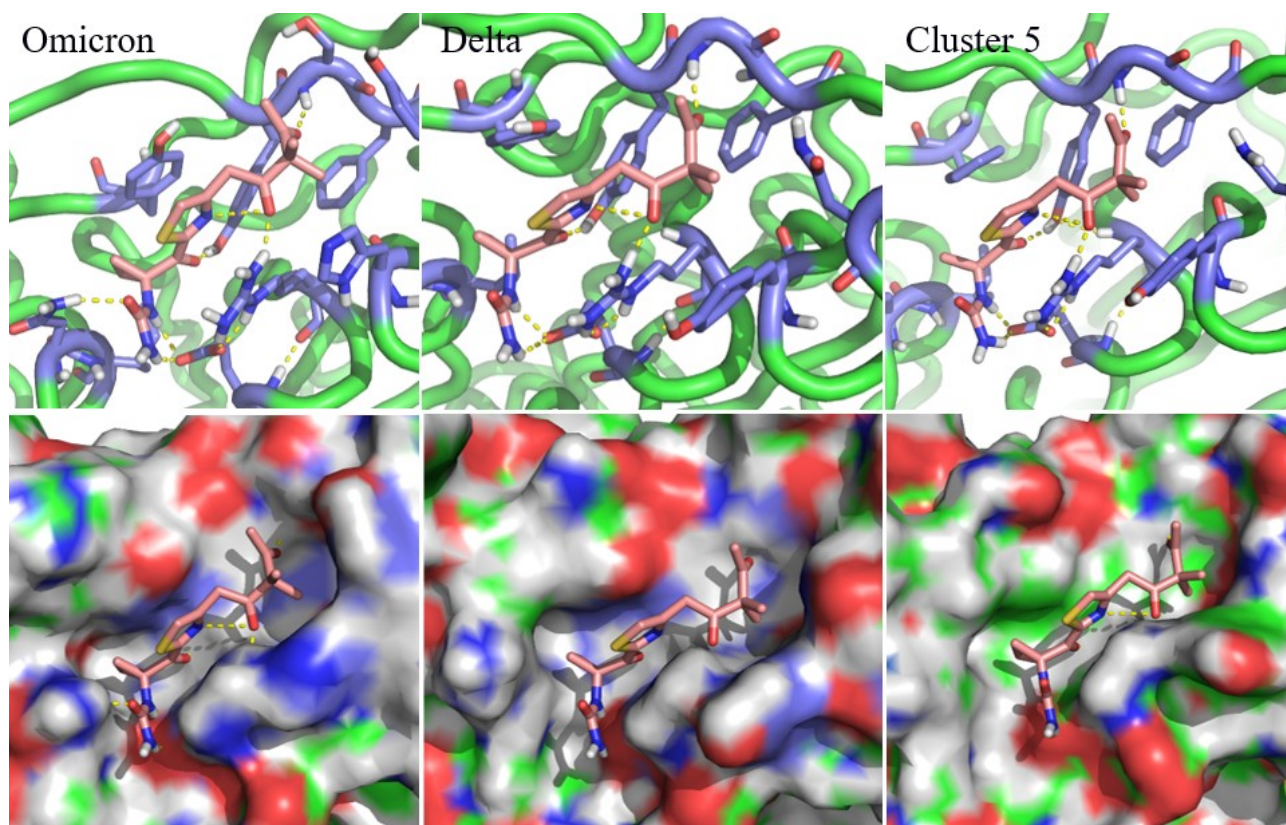


Рисунок 5.15 Розроблений інгібітор в комплексі з RBD різних варіантів SARS-CoV-2. Загальне його положення майже не змінюється в залежності від варіанту

Середньоквадратичне відхилення розробленого інгібітора в комплексі з різними варіантами RBD також дещо відрізняється в залежності від варіанту (рис. 5.16). Зокрема, розроблена структура в комплексі з RBD Кластер 5 показала RMSD навіть нижче ніж у комплексі з RBD Ухань – близько 1,3 Å. Що, поряд із

збереженням цього значення протягом всього періоду моделювання, говорить про ще кращу стабільність утвореного комплексу. Дещо більша лабільність розробленої молекули спостерігається у комплексах з RBD Оміврон та Дельта – близько 2 Å. Що може свідчити про деяку меншу спорідненість розробленого інгібітора до RBD цих варіантів. Однак загальний розмах цих графіків знаходиться в межах 0,5 Å, що свідчить про достатньо константне положення ліганду відносно рецептора.

Додатковим підтвердженням стабільності комплексів розробленої речовини з розглянутими варіантами RBD є RMSF амінокислот, з якими взаємодіє безпосередньо ліганд. Як і у випадку вихідного варіанту, амінокислоти R403, E406, Y453, Y495 та G496 всіх трьох варіантів RBD мають порівняно низьке RMSF (рис. 5.16). При цьому, в більшості їх радикали знаходяться на поверхні білкової глобули та безпосередньо зазнають дії розчинника. Таким чином, низькі флуктуації цих амінокислот дозволяють припустити значний стабілізуючий вплив ефектора, який з ними взаємодіє.

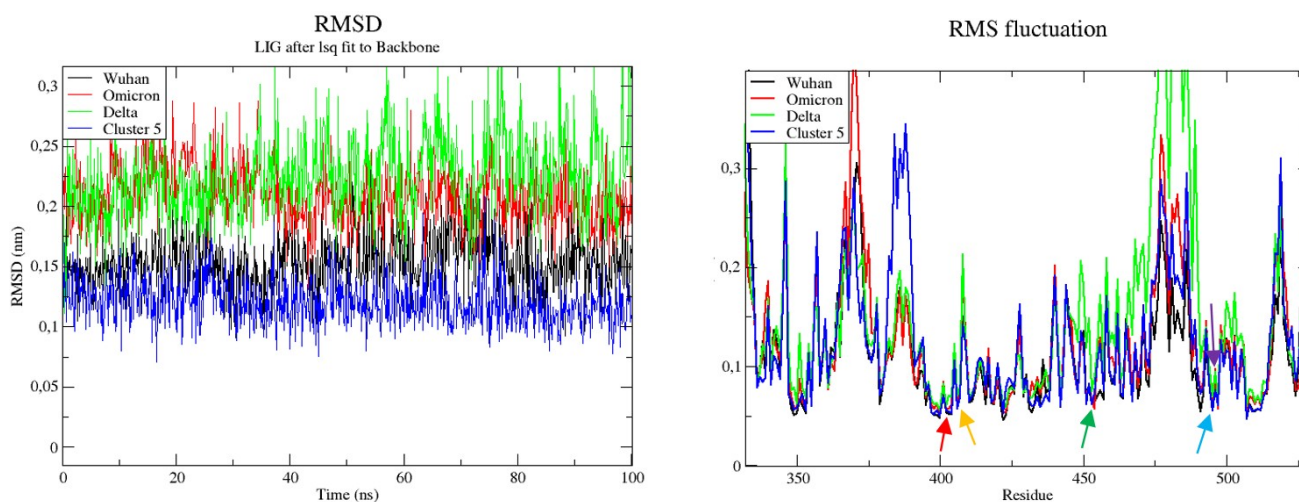


Рисунок 5.16 RMSD ліганд-рецепторного комплексу та RMSF RBD варіантів Ухань, Оміврон, Дельта та Кластер 5. Амінокислоти R403, E406, Y453, Y495 та G496 позначені червоною, жовтою, зеленою синьою і фіолетовою стрілками, відповідно

За значенням вільної енергії комплексоутворення розроблена молекула характеризується не меншою, а навіть більшою спорідненістю до RBD варіантів Оміврон, Дельта та Кластер 5, порівняно із вихідним варіантом Ухань: -21,61, -20,89 та -20,49 ккал/моль проти -18,27 ккал/моль, відповідно. Що, попри дещо зворотну ситуацію у випадку їх RMSD, може свідчити про вищу стабільність комплексів розробленого інгібітора з RBD більш пізніх варіантів SARS-CoV-2. Одночасно різниця між їх розрахованими ΔG знаходиться близько 15%. Тож спорідненість розробленого ліганду до RBD різних варіантів SARS-CoV-2 може розглядатись як подібно висока.

Відповідно до того, що потенційна лікарська структура повинна володіти також рядом фармакокінетичних властивостей і параметрів медичної хімії для подальшого успішного виробництва і застосування, розроблена тут речовина також була оцінена за допомогою ADMETlab 2.0.

Відповідно до рисунку 5.17, її обраховані фізико-хімічні властивості можна вважати в загальному прийнятними. Вона не порушує жодного із правил Ліпінського [266]. Єдиний параметр, розрахункове значення якого виходить за межі запропонованого розробниками ADMETlab 2.0 оптимуму, є $\log D$, який знаходиться для розробленої речовини дещо нижче, що потенційно говорить про невелику надлишкову гідрофільність.

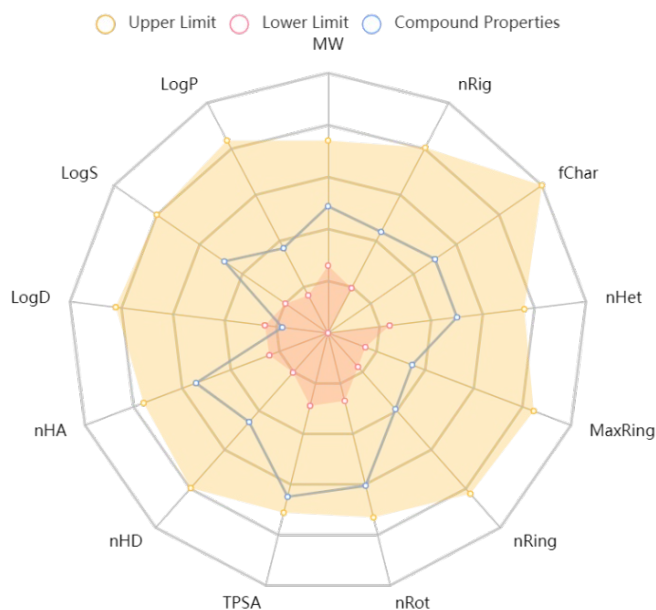


Рисунок 5.17 Фізико-хімічні властивості розробленої структури розраховані ADMETlab 2.0. MW – молекулярна маса, nRig – кількість жорстких зв'язків, fChar – формальний заряд, nHet – кількість гетероатомів, MaxRong – кількість атомів у найбільшому кільці, nRing – кількість кілець, nRot – кількість ротаційних зв'язків, TPSA – площа топологічної полярної поверхні, nHD – кількість донорів водневих зв'язків, nHA – кількість акцепторів водневих зв'язків, LogD – логарифм коефіцієнтів розподілу n-октанол/вода при pH=7,4, LogS – логарифм значення розчинності у воді, LogP – логарифм коефіцієнта розподілу n-октанол/вода

Серед інших показників особливої уваги заслуговують параметри абсорбції, розподілу, канцерогенності та токсичності. Абсорбційна здатність розробленої речовини потенційно знаходиться на високому рівні, оскільки проникність для культури клітин Caco-2 (Caco-2 Permeability) (-5,076) [267], MDCK (MDCK Permeability) (3e-05) [268], загальна абсорбційність (HIA), F20% та F30% знаходяться в прийнятних межах. Єдиною ознакою, що класично вважається негативною є ймовірна придатність розробленого інгібітора на роль субстрату для

Р-глікопротеїну. Однак слід пам'ятати, що рецептор для розробленої молекули – RBD – є функціональним і важливим лише у випадку розміщення віріону SARS-CoV-2 поза клітиною. Таким чином, зменшена концентрація інгібітора в клітині відповідно зменшує ймовірність прояву негативних побічних наслідків, а підвищена у міжклітинному просторі призводить до потенційного підвищення ефективності.

Розраховані параметри розподілу для розробленої молекули можна вважати близькими до ідеальних: вільна і зв'язана з білками плазми фракції приблизно співвідносяться як 1:1, а об'єм розподілу дорівнює 0,376 л/кг. Одночасно дана потенційна лікарська речовина володіє низькою розрахунковою здатністю до проникнення через гематоенцефалічний бар'єр. Останнє не є надзвичайно важливим показником, однак може зменшити потенційні ризики нецільової нейротоксичності.

Розроблена речовина також має низький потенціал до стимулювання канцерогенезу та модуляції активності надважливого калієвого каналу hERG. Одночасно потенційно є достатньо гепатотоксичною, про що свідчать розраховані параметри Н-НТ (human hepatotoxicity) та DILI (drug-induced liver injury).

Розраховані параметри медичної хімії, зокрема SAscore (синтетична доступність), QED (подібність до ліків) та Fsp3 (частка sp3 гібридизованих атомів карбону), також коливаються навколо прийнятних значень для речовини подібної до ліків [269]. SAscore, розрахована ADMETlab 2.0, дорівнює 3,315, що свідчить про відносну синтетичну доступність розробленої сполуки. При цьому QED, який є більш складним та широким параметром, мав значення 0,573, що, хоча і не є привабливим значенням, все ж не може розглядатись неприйнятним для новоствореної сполуки та цілком узгоджується із SAscore. Розрахований параметр Fsp3, який здебільшого пов'язують із кращим прогнозом для речовин подібних на ліки, має значення 0,533, що свідчить про збалансоване співвідношення між розгалуженою аліфатичною та планарною ароматичною компонентами розробленого ліганду.

Підсумовуючи цю частину роботи, молекулярна динаміка розробленого інгібітора у комплексі з RBD всіх чотирьох варіантів SARS-CoV-2 була дуже подібною: у всіх випадках розроблений інгібітор показав високу стабільність утворених комплексів. Базові взаємодії, якими можна вважати характерні для асоціату вихідного варіанту RBD водневі та стек-взаємодії, зберігались в повній мірі протягом молекулярних симуляцій кожного із трьох додаткових комплексів – Оміврон, Дельта та Кластер 5. Два з них мають амінокислотні заміни безпосередньо в ділянці взаємодії з розробленою молекулою, або ж поряд з цією ділянкою (Кластер 5 – Y453F, Оміврон – K417N та Y505H). Що одночасно призвело і до наявності деяких невеликих змін в загальному паттерні взаємодій розробленого інгібітора та відповідного RBD. Зокрема, характерна для RBD Кластер 5 заміна Y453F, призводить до утворення стекінгу з менш полярним порівняно з тирозином радикалом фенілаланіну. Така заміна мінімально впливає на стабільність досліджуваного комплексу, оскільки різниця для утворення стек-взаємодій є. Це підтверджується RMSD ліганду в цьому асоціаті, яке є навіть нижчим ніж у вихідного комплексу. А також розрахованою вільною енергією взаємодії -20,49 ккал/моль проти -18,27 ккал/моль, характерних для комплексу ліганду з RBD вихідного варіанту SARS-CoV-2. Порівняно низьке RMSF амінокислот ділянки зв'язування RBD Кластер 5, асоційованого з розробленою речовиною, також, хоча і опосередковано, свідчить про зниження вільної енергії комплексу. RBD Оміврон, відповідно до згаданих замін, також має дещо змінений паттерн взаємодій з розробленим лігандом. Зокрема, формуються короткочасні водневі зв'язки з H505 та N417, що пояснює характерне, найвище серед інших варіантів SARS-CoV-2, падіння вільної енергії взаємодії (-21,61 ккал/моль). Водночас, середньоквадратичне відхилення ліганду в складі цього комплексу коливається близько 2 Å. Це вище, ніж у випадку комплексів варіантів Кластер 5 та Ухань. Подібною є ситуація і у випадку варіанту Дельта: поряд із високим падінням вільної енергії взаємодії (-20,89 ккал/моль) та високою стабільністю положення амінокислот ділянки зв'язування,

спостерігається дещо вище RMSD. Однак, слід зазначити, що значення середньоквадратичного відхилення є сенс розглядати в певних діапазонах, а не по абсолютному значенню (до 3 Å – положення ліганду достатньо стабільне, більше 3 Å – взаємодія ліганду з рецептором є малоймовірною) [270, 271]. Таким чином, всі комплекси RBD розглянутих варіантів можна вважати подібно і достатньо стабільними. При цьому, якщо звернути увагу саме на загальну зміну вільної енергії комплексоутворення для всіх розглянутих варіантів, то виявиться, що її різниця між термінальними значеннями не перевищує 15%. Це також може вважатись ознакою подібної спорідненості розробленого потенційного інгібітора до всіх чотирьох розглянутих варіантів RBD.

Одночасно, виходячи з розрахованих за допомогою ADMETlab 2.0 даних, розроблений потенційний інгібітор має щонайменше прийнятні фармакокінетичні властивості та параметри медичної хімії. Таким чином, можна стверджувати, що розроблена в цій частині роботи *de novo* сполука має високий потенціал до подальшого використання її або ж її похідних в ролі етіотропного препарату, направленою на подолання наслідків ураження SARS-CoV-2.

5.4 Пошук і ідентифікація потенційного інгібітора RBD S-глікопротеїну SARS-CoV-2 серед затверджених FDA лікарських сполук

Ця частина роботи також є присвяченою розробці інгібітора взаємодії hACE2/RBD і опирається на дані топології функціонально-активної поверхні RBD, отримані попередньо. Однак тут був використаний більш класичний протокол пошуку нових сполук з потенціалом до антивірусної дії [272]. Він передбачав класичний віртуальний скринінг, де в ролі джерела даних лігандів була використана бібліотека затверджених FDA лікарських сполук, а також стандартний у таких випадках багатоступінчатий процес оцінки і відсіву потенційних кандидатів методом симуляції молекулярної динаміки з фінальною оцінкою вільної енергії зв'язування найбільш перспективної речовини.

На першому етапі було проведене фільтрування всієї вхідної бібліотеки за розміром. Виходячи з того, що покет, придатний для безпосередньої взаємодії з лігандом, складається з R403, Y453, Q493, S494, Y495, G496, F497, Q498, N501 та Y505, тобто є достатньо значним за площею, а також опираючись на загальноприйняті правила пошуку ліків, діапазон прийнятних за розміром лігандів був обмежений структурами з масою від 100 Да до 500 Да [273]. Таким чином, кількість потенційних кандидатів була зменшена з 2118 до 1749.

На другому етапі було проведено віртуальний скринінг, який базувався на автоматизованому напівгнучкому молекулярному докінгові. На основі розміщення в межах ділянки взаємодії RBD/hACE2, сформованої у випадку варіанту Оміврон SARS-CoV-2 з амінокислотами функціонально активної поверхні цілі, а також паттерну водневих і інших взаємодій, серед відфільтрованих на основі розміру кандидатів було вибрано пропранолол, циноксацин, метіпронолол, нефазодон, кладрибіне, бікалутамід, метокарбамол, метаксалон і кетодаролутамід як найбільш перспективні (рис. 5.18). Вони характеризувались безпосереднім розміщенням в межах описаної вище каверни і взаємодіяли з хоча б однією з її полярних амінокислот за допомогою водневого зв'язку або іншої направленої взаємодії.

каверни, та безпосередньо центром маси ліганду (рис. 5.19).

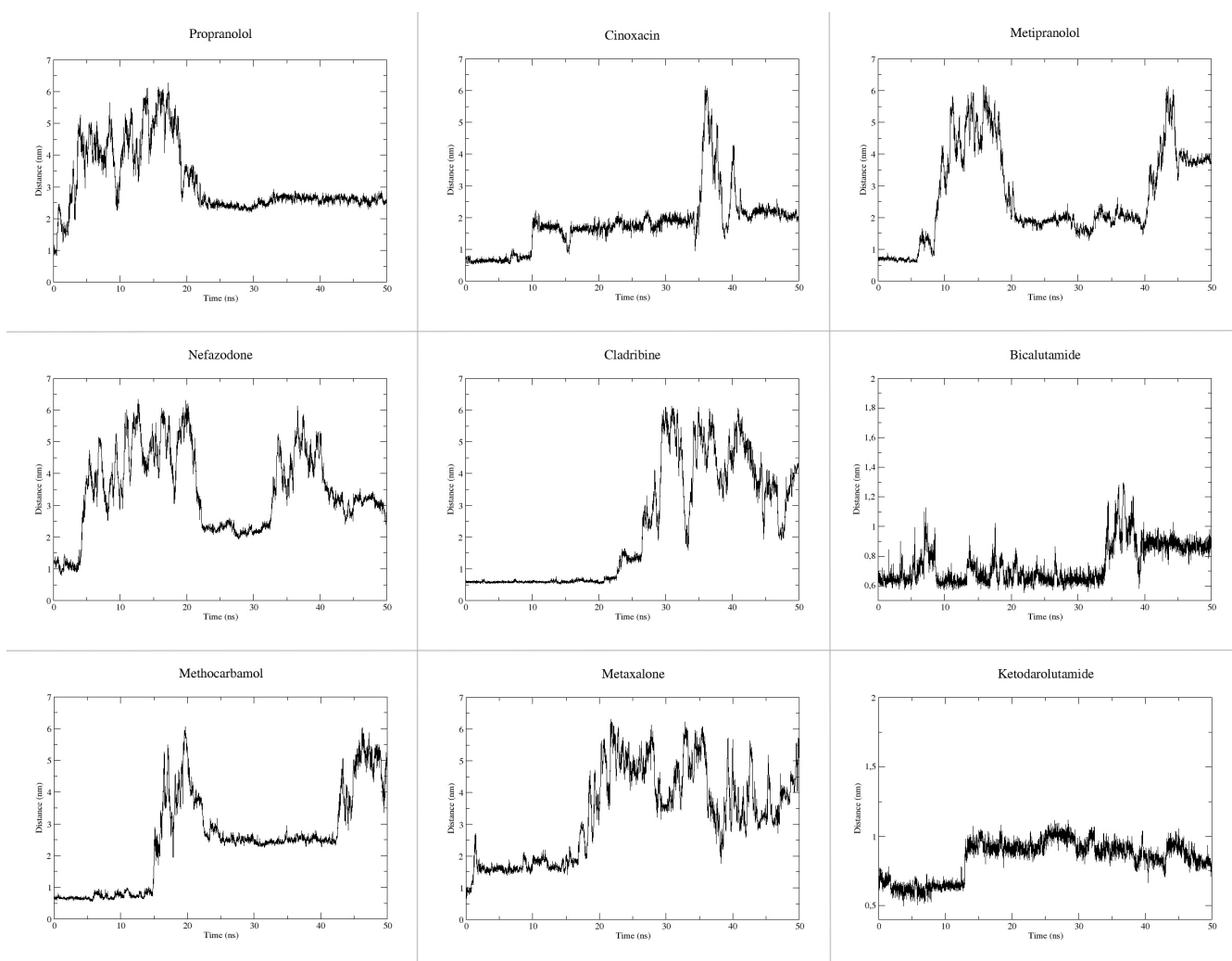


Рисунок 5.19 Динаміки зміни середньої відстані між центрами мас вибраних лігандів і R403 RBD

Як можна зрозуміти з рисунку 5.19, лише бікалутамід і кетодаролутамід під час симуляційного дослідження зміщувались відносно початкової ділянки докування менше ніж на 1 нм. Решта сполук демонстрували набагато більш хаотичну поведінку, яка супроводжувалась різким збільшенням відстані між ділянкою взаємодії RBD/hACE2 і лігандом до 2-6 нм. Таким чином, саме ці структури були відібрані для подальшого більш глибокого дослідження.

Бікалутамід під час подовженої до 200 нс симуляції дисоціював з зони докування вже через 17 нс від початку дослідження. До кінця експерименту цей ліганд не відновив свою взаємодію ні з hACE2-зв'язуючою ділянкою, ні з RBD загалом зазнаючи Броунівського руху у розчиннику (рис. 5.20).

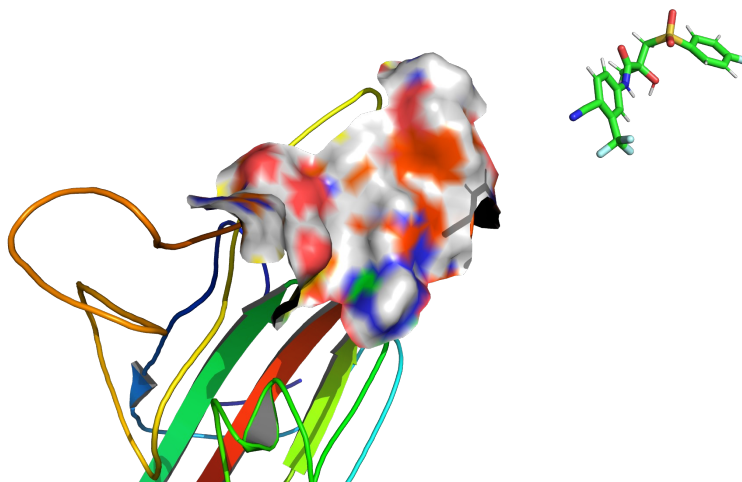


Рисунок 5.20 Положення бікалутаміду після 17ї нс симуляції молекулярної динаміки

Підтвердженням низької стабільності комплексу бікалутамід/RBD є також наведені на рисунку 5.21 зміни у відстані між центрами мас R403 і лігандом та середньоквадратичне відхилення бікалутаміду у цьому асоціаті. Після короткого періоду стабільності, на обох графіках можна зафіксувати вибухоподібний ріст значень, що говорить про значну лабільність ліганду, яка супроводжується віддаленням від ділянки докування.

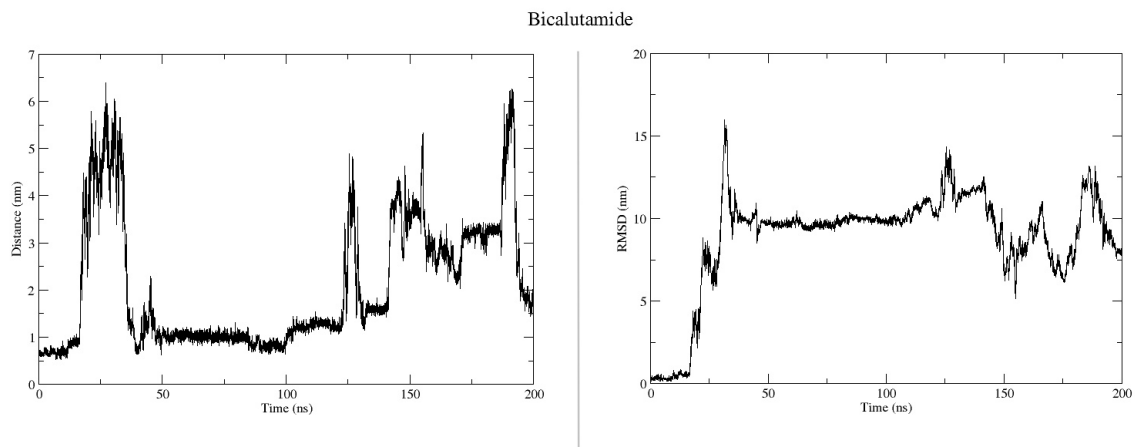


Рисунок 5.21 Середня відстань між центрами мас R403 RBD і бікалутаміду (А), а також RMSD бікалутаміду у комплексі з цим білком (В)

Натомість, комплекс кетодаролутамід/RBD протягом всього періоду симуляції молекулярної динаміки знаходився в зоні докування. Лише в першу секунду симуляції відбулась єдина конформаційна перебудова за якої ацетилпіразолкарбоксамід як частина кетодаролутаміду втратив стек-взаємодію з Н505 на користь ідентичної взаємодії з Y449. Таким чином, кетодаролутамід протягом переважної частини симуляційного дослідження мав конформацію, за якої він перпендикулярно перетинав сайт взаємодії RBD/hACE2 (рис. 5.22).

Також слід зазначити, що окрім очікуваних взаємодій із амінокислотами, які формують цікавий для нас покет, кетодаролутамід додатково формував два водневих зв'язки з Q409 і N417 – амінокислотами поза цим сайтом. При чому, у випадку N417 зв'язок стосувався аміду його пептидної групи та був присутній протягом всього періоду симуляції.

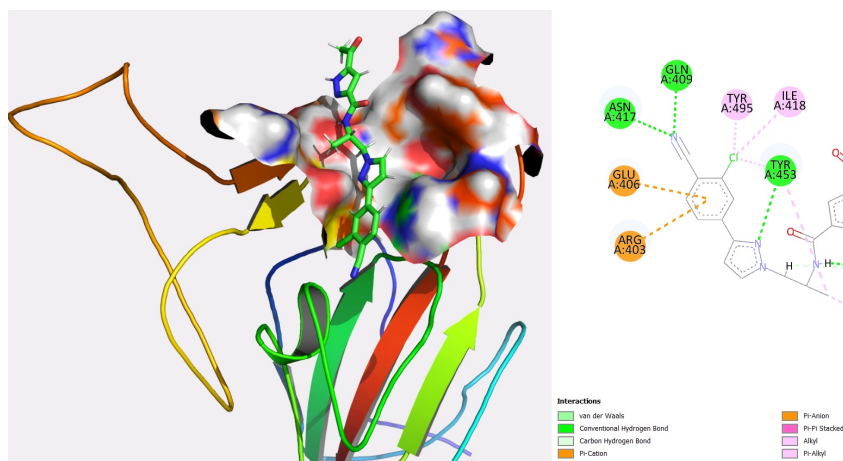


Рисунок 5.22 Репрезентативна поза кетодаролутаміду у комплексі з RBD та паттерн його взаємодій з оточуючими амінокислотами

Висока стабільність асоціату кетодаролутамід/RBD відображається, крім того, у стабільній відстані до R403 RBD, яка в середньому становить 8Å. RMSD, водночас, коливається в межах 5Å і не зазнає відчутних тенденцій до зміни (рис. 5.23).

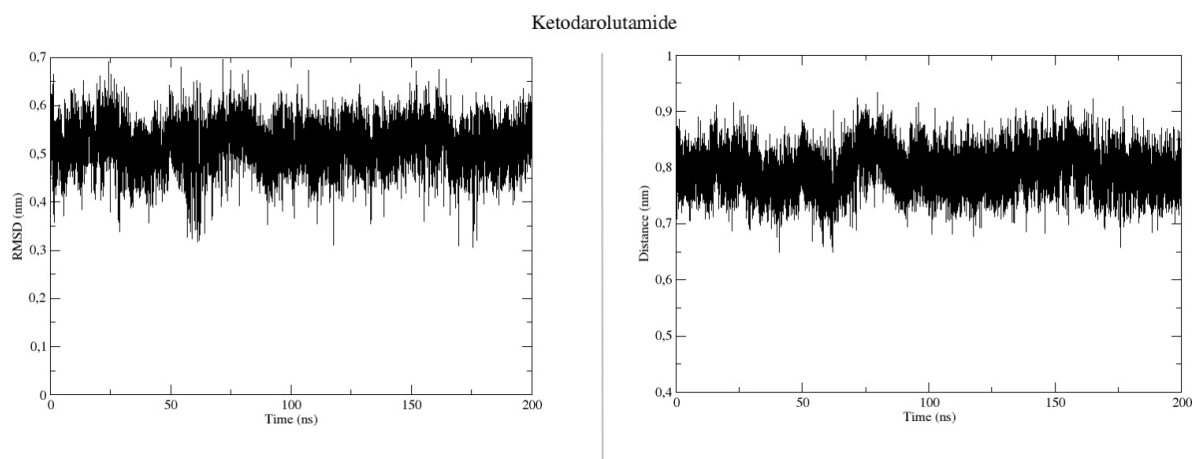


Рисунок 5.23 Середня відстань між центрами мас R403 RBD і кетодаролутаміду (A), а також RMSD кетодаролутаміду у комплексі з цим білком (B)

Задля кращого розуміння потенціалу кетодаролутаміду на фінальній стадії роботи було проведено обрахунок ММ/PBSA, який базувався на отриманій вище траєкторії поглибленого симуляційного дослідження. Як зрозуміло з таблиці 5.1, розрахована вільна енергія зв'язування цього препарату до RBD у зоні його взаємодії з hACE2 становить $-11,33 \pm 3,41$ ккал/моль. Водночас, ентальпія системи становить $-21,29 \pm 3,37$ ккал/моль, а ентропійний штраф дорівнює $9,96 \pm 0,48$ ккал/моль. Найбільш значний внесок в загальну ентальпію, а отже у стабільність взаємодії, вносять неполярні енергетичні компоненти. Хоча слід зазначити, що внесок електростатичної компоненти у $-19,96 \pm 4,81$ ккал/моль є також достатньо значним.

Таблиця 5.1

Розрахована методом ММ/PBSA енергія зв'язування кетодаролутаміду з RBD, виражена у ккал/моль

	Evdw	Eel	Epb	Enpol	ΔH	$-T\Delta S$	ΔG
Average	-32,39	-19,96	34,29	-3,23	-21,29	9,96	-11,33
SD	2,15	4,81	3,8	0,11	3,37	0,48	3,41

Підсумовуючи, у цій частині роботи було проведено скринінгове дослідження щодо потенційних інгібіторів RBD S-глікопротеїну SARS-CoV-2. Воно базувалось на бібліотеці затверджених FDA лікарських сполук. Серед більше ніж двох тисяч структур послідовно у три етапи було ідентифіковано кетодаролутамід, як речовину із виразним потенціалом до взаємодії з hACE2-зв'язуючою ділянкою RBD у зоні попередньо ідентифікованої кишені. Було виявлено, що протягом всього періоду симуляції ліганд стабільно розміщувався в межах перспективної каверни і формував

серед іншого близько чотирьох водневих зв'язків. При чому два з них стосувались Q409 і N417. Ці амінокислоти не входять до переліку тих, які формують попередньо визначену кишеню, однак служать як додаткова компонента, яка підвищує стабільність комплексу і селективність взаємодії. Цікаво, що взаємодія з N417 відбувається за рахунок аміду пептидної групи. Сумарно це говорить про відчутну селективність взаємодії – водневі зв'язки є дуже чутливими до кута розміщення донор-акцепторної пари [274]. А також про те, що мутації, які будуть стосуватись N417 не будуть впливати на взаємодію кетодаролутаміду з RBD у цій ділянці. Стабільне значення RMSD підтверджує стабільність утвореного конгломерату.

На основі отриманої траєкторії задля більш точного розуміння потенціалу кетодаролутаміду щодо інгібування взаємодії RBD/hACE2 було обраховано ΔG комплексоутворення. Вона становила $-11,33 \pm 3,41$ ккал/моль, що підтверджує наведені вище висновки.

РОЗДІЛ 6

ОБГОВОРЕННЯ РЕЗУЛЬТАТІВ РОБОТИ

Об'єктами цієї дисертаційної роботи були два достатньо різних за свою природою патогени – вірус Епштейна-Барр і SARS-CoV-2. Перший є представником родини Herpesviridae і відповідно характеризується ДНК-геномом, дуже складним життєвим циклом і потенціалом до стимуляції канцерогенезу. SARS-CoV-2, натомість, є представником родини Coronaviridae, який має відносно великий РНК-геном і, відповідно, високий потенціал до мутаційних змін. Однак обидва цих патогени мають спільну рису – їх вплив на світову систему охорони здоров'я важко переоцінити.

Вірус Епштейна-Барр у латентному стані є присутнім у близько 90% населення планети, що говорить про довгий і тісний шлях цього патогену та людини як виду [275]. Його здатність до надзвичайного контролю життєвого циклу клітини під час інфекції, окрім згаданої стимуляції канцерогенезу, особливо є небезпечною для людей із пригніченою імунною системою. Лише ПТЛЗ за останніми даними виникає більше ніж в 20% реципієнтів трансплантантів [276]. При чому близько 70% ПТЛЗ є ВЕБ-позитивним, а смертність може становити більше ніж 60% у п'ятирічному інтервалі.

SARS-CoV-2, натомість, є емерджентним патогеном, який спричинив найбільш потужну в новітній історії людства пандемію. За період з початку його глобального поширення і до офіційного завершення пандемії 5 травня 2023 року близько 2млн смертельних випадків COVID-19 було зареєстровано лише у європейському регіоні [277]. Тільки зараз як світова система охорони здоров'я, так і загалом економіка починають оговтуватись від потужного удару, спричиненого цим патогеном та значними зусиллями, що були застосованими задля уповільнення його поширення [278]. Наразі SARS-CoV-2 можна розглядати на роль нової сезонної

загрози, яка нині успішно пройшла адаптацію до людського виду.

Попри глобальну небезпеку, ефективна терапія проти обох цих патогенів не є розробленою досі. Зокрема, у випадку ВЕБ класичні підходи до протидії репродукції патогену шляхом застосування інгібіторів ДНК-полімерази не мають належного ефекту попри загалом високу ефективність *in vitro* [121]. Це пов'язано із складністю життєвого циклу цього вірусу і, особливо, здатністю до майже повного маскування власної присутності у інфікованих клітинах під час максимального рівня латентності. На цій стадії ВЕБ існує у вигляді епісоми у ядрі В-клітини, яка подвоюється за допомогою клітинних ДНК-полімераз, при чому лише під час клітинного поділу.

SARS-CoV-2, при цьому, хоча і може вважатись потенційно «простішим» зі сторони розробки таргетної терапії, є новим патогеном, який після вже більше ніж п'яти років від початку поширення, продовжує активно змінюватись під дією зовнішніх факторів [279]. Серед них вакцинацію і накопичення реконвалісцентного популяційного імунітету можна вважати найбільш потужними. Таким чином, попри значні зусилля, найбільш значуща частина з яких є описаною в одному з попередніх розділів, лише обмежені успіхи були досягнутими. При чому, навіть паксловід – комплексний препарат, розробка якого почалась ще в період виникнення SARS-CoV-1, хоча і володіє доведеним антивірусним впливом, досі не показав статистично значимого ефекту на всі вікові групи населення, окрім старших людей [157].

Таким чином, обидва об'єкти даної дисертаційної роботи, попри свій негативний вплив на світову систему охорони здоров'я, досі не мають розробленої ефективної таргетної терапії. Дослідження проведені тут є присвяченими пошуку і розробці *de novo* молекулярних структур, придатних для отримання на їх основі лікарських засобів здатних прицільно інгібувати певні фактори життєвого циклу цих патогенів, що в перспективі дозволить значно знизити негативний вплив вірусу Епштейна-Барр і SARS-CoV-2 на наш соціум.

Опираючись на важливість білка BHRF1 в життєвому циклі ВЕБ і особливо

його ролі в іморталізації уражених патогеном В-лімфоцитів, цей антиапоптичний фактор був вибраний як перша ціль для розробки селективного інгібітора. BHRF1 характеризується наявністю чітко вираженої і функціонально-важливої кишені, доступної для націлювання, що дозволяє розглядати BHRF1, крім того, як компромісну між простотою і важливістю для патогену ціль [280].

Шляхом застосування ітераційного підходу, який опирався на симуляцію молекулярної динаміки ліганд-рецепторного комплексу як джерело даних щодо спорідненості конкретного варіанту ліганду з цілью, була розроблена молекулярна конструкція здатна до стабільного комплексоутворення з BHRF1 [246]. Ця структура була названа EBAI від Epstein-Barr Antiapoptotic Inhibitor. Задля підтвердження високого рівня спорідненості було проведено поглиблене подвійне симуляційне дослідження довжиною в 120 нс. Воно складалось з двох окремих симуляцій, де власне ліганд-рецепторний комплекс був навмисне дестабілізований шляхом віддалення EBAI від вихідної ділянки на відстань достатню для втрати всіх прямих взаємодій з амінокислотами білка. За результатами експерименту було підтверджено здатність розробленої структури до самостійного відновлення конформації, необхідної для успішної взаємодії з BHRF1. При чому, найшвидше відновлюваними взаємодіями були водневі та іонні зв'язки з R100 та N61, де аргінін і його позитивний заряд вважається вирішальним для виконання BHRF1 своєї антиапоптотичної функції [6].

Потенціал BHRF1 як цілі для розробки терапії проти, зокрема, раку спричиненого ВЕБ був помічений і іншими дослідниками. У 2014 році дослідною групою з США була проведена спроба розробки високомолекулярного інгібітора цього білка [280]. Тут дослідники шляхом еволютивного як автоматичного, так і ручного *de novo* дизайну розробили штучний білок, здатний до специфічного інгібування цього антиапоптичного фактору. В подальшому цей поліпептид був названий BINDI від BHRF1-INhibiting Design acting Intracellularly. Дослідження стабільності комплексу BINDI/BHRF1 дозволили визначити його KD на рівні $220 \pm$

50 пМ. Водночас, спорідненість до людських білків родини BCL-2 виявилась значно нижчою — KD 40 ± 10 нМ, Bcl-2 з KD $2,1 \pm 0,1$ мкМ, Bcl-w з KD 870 ± 40 нМ і Bcl-XL з KD 810 ± 80 нМ. Для підтвердження потенціалу BINDI дослідники також провели *in vivo* експеримент з лікування ксенотрансплантантів Ramos-AW у голих мишей BALB/c, де для покращення фармакокінетики BINDI був ковалентно з'єднаний через С-кінцевий цистеїн до Pol950/CD19. Лікування мишей з ксенотрансплантантами дозволило підтвердити, що BINDI володіє здатністю до протидії раку, індукованому трансформованими ВЕБ клітинами: на 11ий день розмір пухлин у таких мишей становив 140 ± 60 мм³ порівняно з 330 ± 140 мм³ у випадку контрольної групи, якій вводили лише Pol950/CD19. При цьому, об'єм пухлин в групі, яка не отримувала лікування взагалі становив 1080 ± 500 мм³.

Однак попри весь потенціал BHRF1 як цілі для розробки таргетних терапевтичних засобів, цей фактор, як і інші віропротеїни, є відсутнім у випадку становлення максимального рівня латентності, характерного для нормального життєвого циклу ВЕБ. Це робить потенціал ЕВАІ малозастосовуваним у випадку потреби повного знищення патогену. Щоб вирішити цю проблему, ціллю була вибрана єдина структура патогену наявна у всіх без виключення уражених В-клітинах — його ДНК-геном.

Задля націлювання на ДНК дуплекс за основу був вибраний клас ДНК-поліінтеркаляторів. Ці сполуки є здатними здебільшого невивірковано зв'язуватись з дволанцюговою спіраллю шляхом інтеркаляції їх планарних ароматичних частин між нуклеїновими основами та індукувати, таким чином, зупинку транскрипційної і реплікативної машинерії [281]. Лінкери, які забезпечують об'єднання хромофорів, залежно від структурних особливостей конкретної структури, розташовуються в малій або ж великій борозенці дуплексу. Опіраючись на ці теоретичні дані, була розроблена тестова конструкція, де в ролі хромофорів були використані феназинові кільця, а лінкери були модифіковані таким чином, щоб розпізнавати карту донорів і акцепторів водневого зв'язку, розташованих у великій борозенці. Спроектowana

конструкція була неоптимізованою і потенційно здатною розпізнавати послідовність лише із бти пар нуклеотидів. Попри це, під час симуляційного експерименту вона характеризувалась стабільним збереженням положення феназинових хромофорів в місцях їх інтеркаляції та довготривалою взаємодією лінкерних ділянок із нуклеїновими основами відповідної послідовності [249]. Це дозволило підтвердити потенціал подібного підходу.

Застосовуючи ітераційний підхід і симуляцію молекулярної динаміки в ролі джерела даних про стабільність ліганд-рецепторного комплексу, вихідна конструкція була послідовно покращена і набула здатності бути націленою на послідовність будь-якого розміру і складності [250]. Цей перший повноцінний варіант селективного поліінтеркалятора був названий HASDI від High-Affinity Selective DNA Intercalator. Достатньою для повноцінного дослідження було вибрано зону взаємодії довжиною в 16 пар нуклеотидів, оскільки саме така послідовність за розрахунками є більше ніж унікальною в межах людського геному за умови випадкового розміщення нуклеотидів.

Задля підтвердження селективності було проведено два симуляційних дослідження в яких HASDI був націлений на коротку послідовність в межах гена EBNA1 однак інтеркальований не лише в ДНК дуплекс, згенерований на основі цільової послідовності, а і на основі випадкової послідовності гену KCNH2. Очікувано, що стабільність конгломерату HASDI(EBNA1) та ДНК з послідовністю націлювання була подібною і навіть вищою ніж у випадку короткої тестової конструкції. Ознак дестабілізації не спостерігалось протягом всього періоду симуляції, а водневі зв'язки, які власне забезпечували розпізнавання послідовності націлювання, були незмінно присутні і стабільні. Зворотне спостерігалось у випадку інтеркаляції HASDI(EBNA1) до випадкового фрагменту геному. Зокрема, кількість водневих зв'язків була знижена з 32-х до 17-19-ти, а ДНК дуплекс зазнав локального плавлення в місцях максимального напруження. Фінальний обрахунок вільної енергії зв'язування також підтвердив різний рівень спорідненості HASDI(EBNA1)

до цільової і нецільової послідовності: $-235,3 \pm 7,77$ ккал/моль проти $-193,47 \pm 14,09$ ккал/моль, відповідно. При цьому, слід зазначити, що попри всі перелічені зміни, спричинені невідповідністю послідовності до HASDI(EBNA1), не спостерігалось очікуваних процесів деінтеркаляції феназинових хромофорів. Це свідчить на користь наявності у даної конфігурації деякого надлишкового рівня невибіркової спорідненості до ДНК як такої. З метою зниження внеску неселективної компоненти у процес комплексоутворення, було проведено іще один пул досліджень, результатом якого стала розробка наступного покоління HASDI — HASDI-G2 [252].

HASDI-G2, поряд із принциповим збереженням конструкції лінкерної частини, містив замість феназинових індазольні кільця в ролі хромофорів. За рахунок фізичного зменшення площі, придатної для інтеркаляції, прогнозувалось, що неселективна компонента буде відповідно послабленою. Задля підтвердження цього припущення було проведено масштабне порівняльне дослідження, яке включало три незалежних експерименти в кожному з яких HASDI-G2 був націлений на певну послідовність. Подібно до попереднього експерименту, його інтеркаляція проводилась в ДНК дуплекси, згенеровані як на основі послідовності націлювання, так і на основі сиквенсів, в різній ступені відмінних від еталонного.

Перший експеримент був клоном проведеного раніше дослідження з визначення потенціалу HASDI. Однак, на відміну від нього, наслідком невідповідності нуклеотидної послідовності між EBNA1 і KСNH2 стали набагато більш драматичні конформаційні зміни. Зокрема, окрім локального плавлення ДНК дуплексу, спостерігалась швидка деінтеркаляція одразу кількох індазольних кілець. При цьому, кількість водневих зв'язків зменшилась із в середньому 36 до 20 з тенденцією до їх поступового зменшення, а ΔG становила $-97,25 \pm 9,18$ ккал/моль проти $-157,72 \pm 6,57$ ккал/моль для HASDI-G2(EBNA1)/EBNA1-50п.о. Це загалом підтверджує припущення про зниження впливу неселективної компоненти на комплексоутворення HASDI-G2, порівняно з HASDI. Однак, опираючись на ріст кількості водневих зв'язків, отримані результати також свідчать на користь

збільшення спорідненості селективної компоненти даного варіанту поліінтеркалятора. Це цікавий ефект, що не був передбачений на рівні розробки HASDI-G2, але який може розглядатись як підтвердження набагато більш оптимального співвідношення ступеня неселективної спорідненості до ДНК дуплексу і селективного зчитування поверхні великої борозенки.

У другому експерименті HASDI-G2 був націлений на зрощений ген BCR_ABL1, який є молекулярною основою так званої філадельфійської хромосоми – основної причини хронічного мієлоїдного лейкозу [253]. Як послідовності, частково схожі до послідовності націлювання, були використані нативні сиквенси BCR і ABL1 в точці потенційного зрощення. Таким чином, різниця між цими двома послідовностями і послідовністю націлювання була набагато меншою в порівнянні з EBNA1 і KCNH2. Як і очікувалось, взаємодія лінкерів HASDI-G2 з парами нуклеотидів, які не відповідали послідовності націлювання, була майже відсутньою, а у випадку комплексу HASDI-G2(BCR_ABL1)/BCR-50п.о. в цій ділянці також спостерігалось локальне плавлення ДНК дуплексу. Однак поряд з цим в ділянках, які повністю були відповідними послідовності націлювання, також спостерігалась подібна втрата водневих зв'язків, а у випадку комплексу HASDI-G2(BCR_ABL1)/BCR-50п.о. і локальне плавлення ДНК дуплексу. Подібна поведінка також не була передбаченою на рівні розробки, однак може пояснюватись значним внеском так званого кооперативного ефекту в процес взаємодії HASDI-G2 з ДНК [282]. Обрахунок вільної енергії зв'язування загалом підтвердив значну різницю у спорідненості HASDI-G2(BCR_ABL1) до послідовності націлювання, порівняно з нативними BCR-50п.о./ABL1-50п.о. Зокрема, у першому випадку ΔG була рівною $-154,95 \pm 6,93$ ккал/моль, натомість у випадку взаємодії з BCR-50п.о. дорівнювала $-146,19 \pm 7,34$ ккал/моль, а ABL1-50п.о. – $-138,99 \pm 8,36$ ккал/моль. Таким чином, на цілком практичному прикладі було підтверджено можливість нового покоління селективного поліінтеркалятора до розпізнавання послідовності націлювання в умовах наявності споріднених фрагментів ДНК. Слід зазначити, що

потреба в кооперативній взаємодії всього HASDI-G2 задля формування стабільного конгломерату з ДНК дуплексом потенційно значно збільшує його вибірковість, порівняно з очікуваною.

Однак крайнім випадком і найскладнішим можливим завданням для конструкції, призначеної для розпізнавання певного фрагменту геному, є його дискримінація серед послідовностей, відмінних лише однією парою нуклеїнових основ. Третій і фінальний експеримент мав на меті визначити чи має HASDI-G2 такий потенціал. В ролі цільової послідовності був вибраний фрагмент мутантного гену KRAS – KRAS_G12S. Від нативної послідовності він відрізняється єдиною нуклеотидною парою (G→A), яка є причиною амінокислотної заміни і, як наслідок, значних порушень у функціонуванні даного фактору [254]. Під час відповідного симуляційного дослідження було визначено, що, попри відповідність 15 з 16 пар нуклеотидів, ступінь стабільності комплексу HASDI-G2(KRAS_G12S)/KRAS-50п.о. був подібним для такого у випадку HASDI-G2(BCR_ABL1)/BCR-50п.о. та HASDI-G2(BCR_ABL1)/ABL1-50п.о. Зокрема, поряд із очевидною відсутністю орієнтуючих взаємодій з невідповідною парою нуклеїнових основ, були присутніми деструктивні зміни у цілком відповідних частинах конгломерату. Водночас, ці деструктивні зміни проявлялись в тому числі і плавленням подвійної спіралі. Обрахунок енергетики взаємодії відповідав даним, отриманим під час аналізу траєкторій симуляцій. Зокрема, ΔG системи на основі HASDI-G2(KRAS_G12S)/KRAS_G12S-50п.о. була оцінена у $-159,6 \pm 6,61$ ккал/моль, тоді як вільна енергія зв'язування HASDI-G2(KRAS_G12S)/KRAS-50п.о. становила лише $-146,13 \pm 7,45$ ккал/моль.

Таким чином, ймовірно завдяки описаному вище явищу кооперативності, HASDI-G2 здатен чітко розрізняти навіть дуже схожі послідовності, починаючи з різниці у одну пару основ. При чому в ступені, який значно перевищує очікуваний і є подібним для такого у випадку невідповідності у 3 і більше нуклеотидних пар.

В даному контексті слід додатково окреслити особливості інтерпретації даних, отриманих за допомогою використаного в роботі методу ММ/PBSA. Цей підхід, як

один з «End-Point» методів оцінки вільної енергії комплексоутворення, використовує для обрахунку абсолютних значень як рецептор, так і ліганд у конформації присутній у їх комплексі [283]. Таким чином, у випадку ДНК-інтеркаляторів не враховується, що процес власне інтеркаляції передбачає утворення кавеоли між окремими нуклеїновими основами, де останнє потребує енергетичних затрат. Ця особливість MM/PBSA саме в контексті ДНК-інтеркаляції є однією з причин дослідження HASDI і HASDI-G2 шляхом порівняння їх комплексів з цільовими і нецільовими фрагментами ДНК та передбачає саме порівняльний їх аналіз, а не оцінку на основі абсолютних значень.

Підсумовуючи вищесказане та враховуючи, зокрема, і методологічні обмеження, розроблені тут модульні поліінтеркалюючі молекулярні конструкції можна розглядати як такі, що потенційно здатні до винятково селективної і стабільної взаємодії з послідовністю націлювання, яка може бути довжиною щонайменше в 16 пар нуклеїнових основ.

Слушним є порівняти HASDI і HASDI-G2 з відомими молекулярними конструкціями подібного призначення. Безсумнівно ними можна вважати відкриті в 1996 році Дерван та ін. пірол-імідазол поліаміди (ППП) [284]. Цей клас сполук в своїй основі має залишки N-метилпіролу і N-метилімідазолу, сполучені між собою класичними амідними зв'язками. Після певної граничної кількості таких ланок, достатніх здебільшого максимум для взаємодії з шістьма парами нуклеїнових основ, PPP містять лінкер із γ -аміномасляної кислоти. Взаємодія PPP з цільовою послідовністю ДНК передбачає розміщення цих структур в малій борозенці таким чином, що лінкерна ділянка формує так званий γ -поворот, а пірол-імідазольні поліамідні фрагменти спарюються між собою шляхом Ван-дер-Ваальсових взаємодій у стабільну структуру, де донори і акцептори водневого зв'язку лежать у напрямку нуклеосонов. Таким чином, формується подібна до ДНК антипаралельна молекулярна конструкція, де нековалентна пара пірол-пірол розпізнає пари А/Т або Т/А, а пара імідазол-пірол G/C.

За останні кілька десятиліть було проведено ряд досліджень, що підтвердили значний потенціал пірол-імідазол поліамідів в ролі агентів, здатних до модуляції транскрипції певних частин геному. Зокрема, група під керівництвом Пітера Дерван, продовжуючи власні дослідження, розробила поліамід направлений на послідовності ДНК, пов'язані із реалізацією стимулюючого впливу фактору NF-κB (вважається пов'язаним із багатьма видами раку та старінням) [285]. Для цього дослідники синтезували структуру [ImImPyPy-(R)α-aminoγ-PyPyPyPy] (~1370 Да), здатну до теоретичного розпізнавання послідовності 5'-WGGWWW-3'. Додавання даного поліаміду до коротких фрагментів генів IL6 та IL8 (регулюються NF-κB) призвело до підвищення температури плавлення на більше ніж 10°C у обох випадках. Подальші дослідження на A549 дозволили ідентифікувати зниження кількості мРНК IL6 та IL8 у 6,5 та 4 рази, відповідно (інкубація 10 мкМ конструкту з клітинами протягом 48 год з подальшою стимуляцією TNF-α протягом 12 годин). У кінцевому рахунку Дерван та команда досліджували розроблений ними пірол-імідазолний поліамід на предмет його спрямованої активності проти росту ксенотрансплантантів раку простати (LNCaP) *in vivo* на моделі самців мишей NSG. Вони виявили виразну активність розробленої ними конструкції застосованої у концентрації ~1 мг/кг протягом 9 днів: розмір пухлин у дослідній групі був на 64% меншим, порівняно з контролем. Крім того, особливості гістологічного профілю і біохімічних показників свідчили на користь стрімкого некрозу саме ксенотрансплантантів.

Описані вище дослідження є лише невеликою, хоча і репрезентативною, частиною широкого різноманіття напрацювань у напрямку розробки антиракових ліків на основі пірол-імідазолних поліамідів.

Однак, попри загалом високу цінність підтвердження самої можливості селективного інгібування певних частин геному, вже сам підхід, застосований під час створення ППП, вносить конструктивні обмеження щодо подальшого прогресу у цьому напрямку. Зокрема, більшість із наявних нині розробок володіють здатністю

розпізнавати не більше ніж 6 нуклеотидів нерегулярної послідовності ДНК [286]. Це пов'язано із більшим кроком ППП, порівняно із ДНК дуплексом, що з ростом розміру ліганду призводить до стрімкого наростання напруження у кон'югаті і, відповідно, зниження його стабільності [287, 288]. Були спроби виправити цю проблему, які проводили, зокрема, як Дерван з командою (шляхом включення β -аланіну в структуру ППП), так і інші дослідники [289, 290, 291]. Це, однак, також є дієвим лише до певної межі, оскільки передбачає значний ріст внеску ентропійного штрафу у комплексоутворення, що пов'язано із доповненням ліганду довгим аліфатичним лінкером/лінкерами [292].

Для повноти розуміння подальшого потенціалу HASDI і HASDI-G2 як представників нового, до того невідомого, підходу до націлювання на ДНК слід зазначити, що на відміну від ППП їх хімічний синтез ймовірно буде достатньо складним завданням. Попри присутність широкого спектру похідних δ -валеріолактону, 2-метилбутилу, а також всіх замісників з Додатку 1 і Додатку 4 в, до прикладу, PubChem [240], широке застосування як феназинових, так і індазольних хромофорів [293, 294] та добре відпрацьовану методологію хімічної кон'югації різних функціональних груп за допомогою C-C зв'язків [295], HASDI і HASDI-G2 характеризуються таким ускладнюючим фактором як оптична активність. Кожен лінкер кожного окремого модуля обох розроблених структур має п'ять хіральних центрів. Кожен із останніх під час хімічного синтезу окремого лінкера буде призводити до утворення рацемічної суміші, яка повинна бути очищена від одного з енантіомерів. Разом це призводить не тільки до збільшених затрат реагентів щонайменше вдвічі, а і до потреби у застосуванні складних вартісних методів очистки кожного окремого енантіомеру. З іншої сторони, об'єднання окремих модулів між собою може бути процесом, опосередкованим відносно простою і ефективною реакцією конденсації. Останнє дозволяє припустити, що відпрацювання методики синтезу кожного окремого модуля буде найбільш затратним на початку досліджень HASDI і HASDI-G2. З часом, напрацювання методологічної бази і

початок масового синтезу окремих модулів дозволить будувати молекулярну конструкцію необхідної довжини, націлену на конкретний ген в умовах, що не передбачають ні залучення висококваліфікованого персоналу, ні дороговартісних приладів і методів. Власне подібний процес стрімкого розвитку методології синтезу пройшли і пірол-імідазол поліаміди, які на початку свого розвитку були обмеженими класичними на той момент методами синтезу у рідкій фазі. Проте вийшли на рутинний рівень доступності після розробки протоколів їх твердофазного синтезу [287].

Однак попри можливі майбутні покращення методології отримання та такі переваги, як здатність чітко ідентифікувати кожен окрему пару нуклеїнових основ, а також відсутність обмежень у розмірі молекули і, відповідно, довжині ділянки націлювання, слід враховувати, що розроблені тут конструкції можуть мати ускладнення зі сторони їх хімічного синтезу. Тож на даному етапі HASDI і HASDI-G2 слід вважати, хоча і перспективними, однак лише першими спробами розробки агентів, здатних до націлювання безпосередньо на геном патогену.

Повертаючись до SARS-CoV-2, попри адаптаційне зниження патогенності, цей коронавірус досі є причиною відчутного рівня захворюваності з відповідною кількістю смертельних випадків. Водночас, через свою природу як емерджентного РНК-вмісного вірусу, SARS-CoV-2 володіє високою здатністю до адаптації задля протидії таким факторам еволюційного відбору як, до прикладу, наростання популяційної резистентності господаря. Власне процес пристосувальних змін реалізується на рівні змін первинної і, відповідно, вторинної та третинної структури певних, особливо поверхневих, білкових факторів патогенів. Таким чином, задля розуміння особливостей цього процесу у випадку безпосередньо SARS-CoV-2 на початковому етапі частини роботи, яка стосується цього патогену, було проведено статистичне дослідження взаємозалежності ступеня вакцинованості населення, тобто рівня резистентності популяції, і поширеності різних варіантів SARS-CoV-2 [256]. Для більш широкого розуміння взаємозв'язків, це дослідження охоплювало

три країни з різним соціоекономічним профілем – Німеччину як представника економічно розвинених країн, Індію як країну, що розвиваються і має велику густину населення, і Україну як країну, що розвивається, але має середню густину населення [257].

Було визначено, що в кожній з трьох країн один з варіантів SARS-CoV-2, зокрема, варіант Дельта був присутній у популяції господаря задовго до того як стати домінуючим. Його поява не спричинила витіснення інших варіантів, що свідчить на користь відсутності в нього відчутних еволюційних переваг порівняно з іншими варіантами SARS-CoV-2 в той момент. Однак в подальшому в різний період часу варіант Дельта незмінно витісняв інші варіанти. При чому спостерігалась чітка асоціація його поширення і наростання відсотка вакцинованого населення у випадку всіх трьох країн. Таким чином, можна стверджувати, що варіант Дельта з високою ймовірністю набував еволюційної переваги саме за умов підвищеного рівня резистентності популяції господаря.

Відповідно, опираючись на отримані дані, з точки зору розробки антивірусної терапії COVID-19 слід враховувати, що, окрім характерного і відомого високого потенціалу SARS-CoV-2 як РНК-вмісного патогену до швидкого пристосування, цей емерджентний вірус володіє значною гетерогенністю популяції. Це дозволяє йому надзвичайно швидко пристосовуватись до впливу можливого позитивного відбору, спричиненого як вакцинацією, так і впливом потенційного етіотропного лікарського засобу. Таким чином, під час подальших досліджень у напрямку розробки і пошуку нових селективних інгібіторів, наприклад білків прикріплення, слід враховувати потребу у додатковій стадії аналізу потенційної цілі не тільки на предмет наявності придатної для взаємодії з низькомолекулярним лігандом кишені, а і на предмет стабільності асоціації потенційного ліганду за умови змін в межах ділянки взаємодії.

Відповідно до визначеного внеску S-глікопротеїну як структурно функціональній одиниці SARS-CoV-2, яка забезпечує прикріплення та проникнення патогену до чутливої клітини, його рецепторзв'язуючий домен був вибраний в ролі

цілі для подальших досліджень з розробки/пошуку селективного інгібітора. Відповідно до загальноприйнятого процесу дизайну лікарських сполук, спочатку було проведено дослідження hACE2-зв'язуючої ділянки RBD на предмет особливостей інтерфейсу взаємодії і, особливо, наявності кишені в його межах, потенційно придатної для взаємодії з низькомолекулярним лігандом. При чому, опираючись на підвищений потенціал SARS-CoV-2 до адаптаційних змін, спричинених серед іншого, як згадувалось вище, значною гетерогенністю популяції патогену, ця частина роботи включала одразу три варіанти – Ухань, Р.1 і Кластер 5 [226].

Було визначено, що інтерфейс взаємодії RBD з hACE2 не є суцільним і його умовно можна поділити на два субінтерфейси: N-кінцевий та С-кінцевий. Кожен з субінтерфейсів характеризується своїм специфічним паттерном сформованих зв'язків, які змінюються відповідно до замін RBD N501Y та Y453F. При цьому загальна кількість водневих зв'язків є відносно однаковою для всіх мутантів, що свідчить на користь орієнтовно подібного ступеня спорідненості різних варіантів SARS-CoV-2 до людського hACE2 і відповідну виродженість переліку можливих мутацій в цій ділянці.

Подальші симуляційні дослідження кожного з RBD дозволили ідентифікувати розміщену між згаданими вище субінтерфейсами відносно консервативну кишеню, потенційно придатну для взаємодії з низькомолекулярним лігандом. В процесі аналізу зміни її топологічних особливостей відповідно до замін N501Y та Y453F було підтверджено відносно значні перебудови, які включали в тому числі і зміни в лінійних розмірах покету. Таким чином, було підтверджено хоча і обмежену, але все ж варіативність топології поверхні взаємодії RBD з hACE2, залежну від конкретного варіанту SARS-CoV-2. Опираючись на отримані дані, в наступній частині роботи було розроблено низькомолекулярний інгібітор, здатний до взаємодії з ідентифікованою кишенею попри наявні зміни в її топології [266].

Для цього був застосований згаданий раніше ітераційний підхід, який

опирався на симуляцію молекулярної динаміки ліганд-рецепторного комплексу як джерело даних щодо спорідненості конкретного варіанту ліганду з ціллю. Задля збільшення ймовірності довготривалої активності потенційного інгібітора в умовах гетерогенності популяції патогену кількість варіантів SARS-CoV-2 була збільшена до чотирьох. Зокрема було враховано зміни топології ACE2-зв'язуючої ділянки RBD варіантів Ухань, Омикрон, Дельта та Кластер 5.

Розроблений, таким чином, потенційний інгібітор утворював співставні за своєю стабільністю комплекси з кожним із досліджуваних варіантів RBD, попри наявні у випадку Кластер 5 і Омикрон заміни Y453F і K417N/Y505H. Слід зазначити, що вплив мутацій все ж був присутній, що особливо чітко спостерігалось у випадку RBD Омикрон, де під час симуляції молекулярної динаміки ліганд-рецепторного комплексу формувались додаткові водневі зв'язки між інгібітором і амінокислотами H505 та N417. Однак за рештою параметрів, зокрема середньоквадратичним відхиленням інгібітора і вільною енергією зв'язування, всі комплекси були орієнтовно однаково високостабільними.

Для розуміння більш віддаленого потенціалу цієї структури додатково було розрахованої її можливі фармакокінетичні параметри. Серед більше ніж 84 розрахованих ADMETlab 2.0 параметрів лише кілька виходили за прийнятні межі для класичної лікарської сполуки. Серед них LogD, H-НТ та DILI, що свідчить на користь наявності у сполуки деякої надлишкової гідрофільності і гепатотоксичності. Це, однак, у порівнянні з переважаюче прийнятним прогнозом, може розглядатись як ознака хоча не ідеального, але достатнього потенціалу цієї речовини щонайменше у площині можливої подальшої розробки на її основі структур з покращеним ADMET профілем. Що насправді є класичним підходом в сфері структуро орієнтованого дизайну ліків.

Враховуючи наявність відкритої і дослідженої в межах цієї роботи кишені, розташованої у складі ACE2-зв'язуючої ділянки RBD, було проведено ще одне дослідження у напрямку пошуку нових сполук із потенціалом до протидії SARS-

CoV-2 у розпізнаванні клітини господаря [296]. Ця частина роботи була зосереджена на більш класичному і поширеному віртуальному скринінгові. При чому, враховуючи, що джерелом вихідних даних була бібліотека затверджених FDA лікарських речовин, це дослідження можна розглядати і з точки зору перепрофілювання ліків.

Після трьох послідовних етапів фільтрації, молекулярного докінгу і симуляції молекулярної динаміки, кетодаролутамід був ідентифікований як молекулярна структура, здатна до взаємодії з RBD SARS-CoV-2 у ділянці його контакту з людським ACE2. Подальше поглиблене симуляційне дослідження підтвердило отримані дані. Кетодаролутамід характеризувався стабільним середньоквадратичним відхиленням протягом 200нс симуляції, не виявляв тенденції до відділення від ділянки взаємодії, а його вільна енергія зв'язування становила $-11,33 \pm 3,41$ ккал/моль. Що для неоптимізованої під специфічне амінокислотне оточення сполуки є достатньо прийнятним показником [297].

Цікаво, що під час поглибленого вивчення літературних даних щодо кетодаролутаміду, було виявлено звіти, де стверджується нижчий рівень захворюваності і послаблення симптоматики в хворих раком передміхурової залози, які приймали антиандрогени [15]. Серед останніх у протоколах лікування часто зустрічається і кетодаролутамід [298]. В цих дослідженнях науковці пояснюють такий ефект здебільшого за рахунок залежності кількості hACE2 від концентрації андрогенів. Однак, виходячи з отриманих результатів, можна припустити етіотропний вплив щонайменше одного із застосованих препаратів.

Підсумовуючи, в цій частині дисертаційної роботи було проведено ряд досліджень із визначення еволюційних і пов'язаних з цим структурних особливостей SARS-CoV-2 як емерджентного патогену. Ці дані були використані для розробки низькомолекулярної речовини, потенційно здатної до стабільної взаємодії з основною структурою SARS-CoV-2, призначеною для розпізнавання сприйнятливої клітини – рецепторзв'язуючого домену S-глікопротеїну. Це дозволяє

вважати цю сполуку перспективним кандидатом для подальшої розробки на її основі або на основі її похідних етіотропного антиковідного лікарського засобу. Хоча слід зазначити, що враховані вище варіанти Ухань, Омикрон, Дельта та Кластер 5 були актуальними на період проведення досліджень (2020-2021pp). Наразі варіант Омикрон як домінуючий має щонайменше 49 підштамів і продовжує мутувати, пристосовуючись до наростання популяційного імунітету. Мутації, характерні для цих підваріантів залучають в тому числі амінокислоти, які формують згадану вище кишеню [299]. І хоча на прикладі збереження спорідненості розробленої структури до щонайменше чотирьох варіантів можна припустити відносно низьку її чутливість до навіть тих амінокислотних замінів, які безпосередньо стосуються відкритої кишені, таке різноманіття нових варіацій RBD породжує відповідне різноманіття мутацій ACE2-зв'язуючої ділянки цього вірусного фактору, відповідно знижуючи ймовірність прояву достатнього рівня активності. Таким чином, розроблена речовина повинна розглядатись як така, що має значний потенціал до прояву антикоронавірусної активності відносно досліджених тут варіантів SARS-CoV-2, а також володіє певним потенціалом до його збереження у випадку сучасних ліній цього емерджентного патогену. Хоча останнє припущення потребує подальших послідовних розрахункових і *in vitro* досліджень.

Заключна частина цієї дисертаційної роботи також опиралась на структурні дані, отримані під час вивчення RBD і, в даному випадку, включала класичне дослідження з віртуального скринінгу. Пошук був проведений на основі бібліотеки затверджених FDA лікарських сполук і включав кілька послідовних етапів оцінки та фільтрації. У підсумку це дозволило ідентифікувати кетодаролутамід як затверджений препарат, потенційно здатний впливати на функціонування RBD SARS-CoV-2. Отримані дані, окрім цілком прикладного інтересу, можуть мати і фундаментальне значення оскільки проливають світло на потенційний механізм занотованого раніше іншими дослідниками впливу даного антиандрогену на перебіг COVID-19.

ВИСНОВКИ

В дисертаційній роботі розроблені та охарактеризовані ряд молекулярних структур, потенційно здатних інгібувати репродукцію двох визначних за своїм соціоекономічним впливом патогенів – вірусу Епштейна-Барр та SARS-CoV-2. За результатами роботи зроблені наступні висновки:

1. *De novo* розроблено EBAl – перший в своєму роді потенційний інгібітор BCL-2-подібного віропротеїну, білка ВЕБ ВHRF1.
2. За допомогою підходів структуро орієнтованого дизайну ліків розроблено HASDI і його покращений варіант HASDI-G2 – унікальні високомолекулярні конструкції, які в численних симуляційних експериментах довели свою здатність високоселективно розпізнавати фрагмент нуклеїнової кислоти певної послідовності. Їх націлювання на ген EBNA1 вірусу Епштейна-Барр – єдиний періодично активний за максимального рівня латентності ген цього патогену – потенційно дозволяє протидіяти ВЕБ, навіть в умовах максимального рівня аутопригнічення його транскрипційної активності.
3. На основі даних порівняльного аналізу визначено, що SARS-CoV-2 володіє значною гетерогенністю популяції. Вперше на період публікації відповідних даних прямо акцентовано увагу на зв'язку між наростанням популяційного імунітету проти SARS-CoV-2 і його варіантним складом як явища критичного для успішної розробки антиковідних ліків.
4. В межах hACE2-зв'язуючого інтерфейсу RBD S-глікопротеїну SARS-CoV-2 ідентифіковано до того невідому топологічно спільну серед варіантів Ухань, P.1 і Кластер 5 кишеню придатну для розробки і пошуку низькомолекулярних інгібіторів взаємодії цього патогену з hACE2.
5. Опираючись на напрацьовані як біологічні, так і структурні дані *de novo* розроблено перший у своєму роді малий ліганд, потенційно здатний до стабільної взаємодії з RBD варіантів Ухань, Омікрон, Дельта та Кластер 5. Це,

окрім прикладного значення, на фундаментальному рівні підтверджує можливість розробки універсального інгібітора прикріплення SARS-CoV-2.

6. Серед затверджених FDA лікарських сполук ідентифіковано кетодаролутамід як ліганд, потенційно здатний до зв'язування з RBD у його hACE2-зв'язуючій ділянці. Попри те, що через потужний антиандрогенний ефект цей препарат не може бути використаний безпосередньо в ролі антикоронавірусного засобу, він може стати структурною основою для розробки сполук як з кращою спорідненістю до цілі, так і нижчим або відсутнім небажаним антадрогенним потенціалом. Побічним, але важливим ефектом отриманих даних, є розкриття можливого механізму впливу кетодаролутаміду на перебіг COVID-19 під час лікування андроген залежного раку передміхурової залози.

Майбутні перспективи. Важливим є окреслити обмеження представлених оригінальних досліджень. Попри весь спектр застосованих тут загальноприйнятих і відпрацьованих десятиліттями методів та підходів, їх розрахункова природа робить отриманий результат хоча і науковосмісним, але також і ймовірнісним за своєю суттю. Тож біологічний потенціал кожної із розроблених молекулярних конструкцій і кожної із *de novo* отриманих або ідентифікованих малих молекул потребує подальшої *in vitro* оцінки, як наступного кроку класичного процесу структуро орієнтованого дизайну ліків.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Favaretto, F., Matsumura, E.E., Ferriol, I., Chitarra, W., and Nerva, L. (2025). The four Ws of viruses: Where, Which, What and Why - A deep dive into viral evolution. *Virology* 606, 110476. <https://doi.org/10.1016/j.virol.2025.110476>.
2. Elkhatib, W.F., Abdelkareem, S.S., Khalaf, W.S., Shahin, M.I., Elfadil, D., Alhazmi, A., El-Batal, A.I., and El-Sayyad, G.S. (2022). Narrative review on century of respiratory pandemics from Spanish flu to COVID-19 and impact of nanotechnology on COVID-19 diagnosis and immune system boosting. *Virol J* 19, 167. <https://doi.org/10.1186/s12985-022-01902-2>.
3. Murata, T., Sugimoto, A., Inagaki, T., Yanagi, Y., Watanabe, T., Sato, Y., and Kimura, H. (2021). Molecular Basis of Epstein-Barr Virus Latency Establishment and Lytic Reactivation. *Viruses* 13, 2344. <https://doi.org/10.3390/v13122344>.
4. Kaján, G.L., Doszpoly, A., Tarján, Z.L., Vidovszky, M.Z., and Papp, T. (2020). Virus-Host Coevolution with a Focus on Animal and Human DNA Viruses. *J Mol Evol* 88, 41–56. <https://doi.org/10.1007/s00239-019-09913-4>.
5. Foight, G.W., and Keating, A.E. (2015). Locating Herpesvirus Bcl-2 Homologs in the Specificity Landscape of Anti-Apoptotic Bcl-2 Proteins. *J Mol Biol* 427, 2468–2490. <https://doi.org/10.1016/j.jmb.2015.05.015>.
6. Fitzsimmons, L., Cartlidge, R., Chang, C., Sejic, N., Galbraith, L.C.A., Suraweera, C.D., Croom-Carter, D., Dewson, G., Tierney, R.J., Bell, A.I., et al. (2020). EBV BCL-2 homologue BHRF1 drives chemoresistance and lymphomagenesis by inhibiting multiple cellular pro-apoptotic proteins. *Cell Death Differ* 27, 1554–1568. <https://doi.org/10.1038/s41418-019-0435-1>.
7. Song, S., Jiang, Z., Spezia-Lindner, D.E., Liang, T., Xu, C., Wang, H., Tian, Y., and Bai, Y. (2020). BHRF1 Enhances EBV Mediated Nasopharyngeal Carcinoma Tumorigenesis through Modulating Mitophagy Associated with Mitochondrial Membrane Permeabilization Transition. *Cells* 9, 1158. <https://doi.org/10.3390/cells9051158>.

8. Abbas, F., El Kossi, M., Shaheen, I.S., Sharma, A., and Halawa, A. (2020). Post-transplantation lymphoproliferative disorders: Current concepts and future therapeutic approaches. *World J Transplant* 10, 29–46. <https://doi.org/10.5500/wjt.v10.i2.29>.
9. Sarker, R., Roknuzzaman, A.S.M., Nazmunnahar, null, Shahriar, M., Hossain, M.J., and Islam, M.R. (2023). The WHO has declared the end of pandemic phase of COVID-19: Way to come back in the normal life. *Health Sci Rep* 6, e1544. <https://doi.org/10.1002/hsr2.1544>.
10. Sarkar, M., and Madabhavi, I. (2024). COVID-19 mutations: An overview. *World J Methodol* 14, 89761. <https://doi.org/10.5662/wjm.v14.i3.89761>.
11. Contes, K.M., Liu, B.M., Contes, K.M., and Liu, B.M. (2025). Epidemiology, Clinical Significance, and Diagnosis of Respiratory Viruses and Their Co-Infections in the Post-COVID Era. *Pathogens* 14. <https://doi.org/10.3390/pathogens14030262>.
12. Ju, B., Zhang, Q., Ge, J., Wang, R., Sun, J., Ge, X., Yu, J., Shan, S., Zhou, B., Song, S., et al. (2020). Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* 584, 115–119. <https://doi.org/10.1038/s41586-020-2380-z>.
13. Saini, M., Mehra, N., Kumar, G., Paul, R., and Kovács, B. (2025). Chapter Seven - Molecular and structure-based drug design: From theory to practice. In *Advances in Pharmacology*, C. V. Narayan, S. Verma, A. S. Grewal, N. Singh, and H. Nimesh, eds. (Academic Press), pp. 121–138. <https://doi.org/10.1016/bs.apha.2025.02.004>.
14. Lyu, J., Irwin, J.J., and Shoichet, B.K. (2023). Modeling the expansion of virtual screening libraries. *Nat Chem Biol* 19, 712–718. <https://doi.org/10.1038/s41589-022-01234-w>.
15. Giotis, E.S., Cil, E., and Brooke, G.N. (2022). Use of Antiandrogens as Therapeutic Agents in COVID-19 Patients. *Viruses* 14, 2728. <https://doi.org/10.3390/v14122728>.
16. Möhl, B.S., Chen, J., Sathiyamoorthy, K., Jardetzky, T.S., and Longnecker, R. (2016). Structural and Mechanistic Insights into the Tropism of Epstein-Barr Virus. *Mol Cells* 39, 286–291. <https://doi.org/10.14348/molcells.2016.0066>.
17. Draganova, E.B., Valentin, J., and Heldwein, E.E. (2021). The Ins and Outs of

- Herpesviral Capsids: Divergent Structures and Assembly Mechanisms across the Three Subfamilies. *Viruses* 13, 1913. <https://doi.org/10.3390/v13101913>.
18. Li, Z., Zhang, X., Dong, L., Pang, J., Xu, M., Zhong, Q., Zeng, M.-S., and Yu, X. (2020). CryoEM structure of the tegumented capsid of Epstein-Barr virus. *Cell Res* 30, 873–884. <https://doi.org/10.1038/s41422-020-0363-0>.
19. Židovec Lepej, S., Matulić, M., Gršković, P., Pavlica, M., Radmanić, L., and Korać, P. (2020). miRNAs: EBV Mechanism for Escaping Host's Immune Response and Supporting Tumorigenesis. *Pathogens* 9, 353. <https://doi.org/10.3390/pathogens9050353>.
20. Smatti, M.K., Al-Sadeq, D.W., Ali, N.H., Pintus, G., Abou-Saleh, H., and Nasrallah, G.K. (2018). Epstein-Barr Virus Epidemiology, Serology, and Genetic Variability of LMP-1 Oncogene Among Healthy Population: An Update. *Front Oncol* 8, 211. <https://doi.org/10.3389/fonc.2018.00211>.
21. Ba Abdullah, M.M., Palermo, R.D., Palser, A.L., Grayson, N.E., Kellam, P., Correia, S., Szymula, A., and White, R.E. (2017). Heterogeneity of the Epstein-Barr Virus (EBV) Major Internal Repeat Reveals Evolutionary Mechanisms of EBV and a Functional Defect in the Prototype EBV Strain B95-8. *J Virol* 91, e00920-17. <https://doi.org/10.1128/JVI.00920-17>.
22. Wongwiwat, W., Fournier, B., Bassano, I., Bayoumy, A., Elgueta Karstegl, C., Styles, C., Bridges, R., Lenoir, C., BoutBoul, D., Moshous, D., et al. (2022). Epstein-Barr Virus Genome Deletions in Epstein-Barr Virus-Positive T/NK Cell Lymphoproliferative Diseases. *J Virol* 96, e0039422. <https://doi.org/10.1128/jvi.00394-22>.
23. Borozan, I., Zapatka, M., Frappier, L., and Ferretti, V. (2018). Analysis of Epstein-Barr Virus Genomes and Expression Profiles in Gastric Adenocarcinoma. *J Virol* 92, e01239-17. <https://doi.org/10.1128/JVI.01239-17>.
24. Yu, H., and Robertson, E.S. (2023). Epstein-Barr Virus History and Pathogenesis. *Viruses* 15, 714. <https://doi.org/10.3390/v15030714>.
25. Dunmire, S.K., Grimm, J.M., Schmeling, D.O., Balfour, H.H., and Hogquist, K.A. (2015). The Incubation Period of Primary Epstein-Barr Virus Infection: Viral Dynamics

- and Immunologic Events. *PLoS Pathog* 11, e1005286. <https://doi.org/10.1371/journal.ppat.1005286>.
26. Hadinoto, V., Shapiro, M., Greenough, T.C., Sullivan, J.L., Luzuriaga, K., and Thorley-Lawson, D.A. (2008). On the dynamics of acute EBV infection and the pathogenesis of infectious mononucleosis. *Blood* 111, 1420–1427. <https://doi.org/10.1182/blood-2007-06-093278>.
27. Jean-Pierre, V., Lupo, J., Buisson, M., Morand, P., and Germe, R. (2021). Main Targets of Interest for the Development of a Prophylactic or Therapeutic Epstein-Barr Virus Vaccine. *Front Microbiol* 12, 701611. <https://doi.org/10.3389/fmicb.2021.701611>.
28. Anderson, L.R., Owens, T.W., and Naylor, M.J. (2014). Structural and mechanical functions of integrins. *Biophys Rev* 6, 203–213. <https://doi.org/10.1007/s12551-013-0124-0>.
29. Chesnokova, L.S., and Hutt-Fletcher, L.M. (2011). Fusion of Epstein-Barr virus with epithelial cells can be triggered by $\alpha\beta 5$ in addition to $\alpha\beta 6$ and $\alpha\beta 8$, and integrin binding triggers a conformational change in glycoproteins gHgL. *J Virol* 85, 13214–13223. <https://doi.org/10.1128/JVI.05580-11>.
30. Backovic, M., Longnecker, R., and Jardetzky, T.S. (2009). Structure of a trimeric variant of the Epstein-Barr virus glycoprotein B. *Proc Natl Acad Sci U S A* 106, 2880–2885. <https://doi.org/10.1073/pnas.0810530106>.
31. Valencia, S.M., and Hutt-Fletcher, L.M. (2012). Important but differential roles for actin in trafficking of Epstein-Barr virus in B cells and epithelial cells. *J Virol* 86, 2–10. <https://doi.org/10.1128/JVI.05883-11>.
32. Möhl, B.S., Chen, J., and Longnecker, R. (2019). Gammaherpesvirus entry and fusion: A tale how two human pathogenic viruses enter their host cells. *Adv Virus Res* 104, 313–343. <https://doi.org/10.1016/bs.aivir.2019.05.006>.
33. Sathiyamoorthy, K., Hu, Y.X., Möhl, B.S., Chen, J., Longnecker, R., and Jardetzky, T.S. (2016). Structural basis for Epstein-Barr virus host cell tropism mediated by gp42 and gHgL entry glycoproteins. *Nat Commun* 7, 13557. <https://doi.org/10.1038/ncomms13557>.

34. Döhner, K., Cornelius, A., Serrero, M.C., and Sodeik, B. (2021). The journey of herpesvirus capsids and genomes to the host cell nucleus. *Curr Opin Virol* 50, 147–158. <https://doi.org/10.1016/j.coviro.2021.08.005>.
35. Liashkovich, I., Hafezi, W., Kühn, J.M., Oberleithner, H., and Shahin, V. (2011). Nuclear delivery mechanism of herpes simplex virus type 1 genome. *J Mol Recognit* 24, 414–421. <https://doi.org/10.1002/jmr.1120>.
36. Brandariz-Nuñez, A., Liu, T., Du, T., and Evilevitch, A. (2019). Pressure-driven release of viral genome into a host nucleus is a mechanism leading to herpes infection. *Elife* 8, e47212. <https://doi.org/10.7554/eLife.47212>.
37. López-Muñoz, A.D., Rastrojo, A., Kropp, K.A., Viejo-Borbolla, A., and Alcamí, A. (2021). Combination of long- and short-read sequencing fully resolves complex repeats of herpes simplex virus 2 strain MS complete genome. *Microb Genom* 7, 000586. <https://doi.org/10.1099/mgen.0.000586>.
38. Clatterbuck Soper, S.F., and Meltzer, P.S. (2023). ATRX/DAXX: Guarding the Genome against the Hazards of ALT. *Genes (Basel)* 14, 790. <https://doi.org/10.3390/genes14040790>.
39. Inagaki, T., Sato, Y., Ito, J., Takaki, M., Okuno, Y., Yaguchi, M., Masud, H.M.A.A., Watanabe, T., Sato, K., Iwami, S., et al. (2020). Direct Evidence of Abortive Lytic Infection-Mediated Establishment of Epstein-Barr Virus Latency During B-Cell Infection. *Front Microbiol* 11, 575255. <https://doi.org/10.3389/fmicb.2020.575255>.
40. Quinn, L.L., Williams, L.R., White, C., Forrest, C., Zuo, J., and Rowe, M. (2016). The Missing Link in Epstein-Barr Virus Immune Evasion: the BDLF3 Gene Induces Ubiquitination and Downregulation of Major Histocompatibility Complex Class I (MHC-I) and MHC-II. *J Virol* 90, 356–367. <https://doi.org/10.1128/JVI.02183-15>.
41. Elliott, J., Goodhew, E.B., Krug, L.T., Shakhnovsky, N., Yoo, L., and Speck, S.H. (2004). Variable methylation of the Epstein-Barr virus Wp EBNA gene promoter in B-lymphoblastoid cell lines. *J Virol* 78, 14062–14065. <https://doi.org/10.1128/JVI.78.24.14062-14065.2004>.

42. Kang, M.-S., and Kieff, E. (2015). Epstein-Barr virus latent genes. *Exp Mol Med* 47, e131. <https://doi.org/10.1038/emm.2014.84>.
43. Friedrich, T., Ferrante, F., Pioger, L., Nist, A., Stiewe, T., Andrau, J.-C., Bartkuhn, M., Giaimo, B.D., and Borggrefe, T. (2022). Notch-dependent and -independent functions of transcription factor RBPJ. *Nucleic Acids Res* 50, 7925–7937. <https://doi.org/10.1093/nar/gkac601>.
44. Portal, D., Zhou, H., Zhao, B., Kharchenko, P.V., Lowry, E., Wong, L., Quackenbush, J., Holloway, D., Jiang, S., Lu, Y., et al. (2013). Epstein-Barr virus nuclear antigen leader protein localizes to promoters and enhancers with cell transcription factors and EBNA2. *Proc Natl Acad Sci U S A* 110, 18537–18542. <https://doi.org/10.1073/pnas.1317608110>.
45. Zhao, B., Zou, J., Wang, H., Johannsen, E., Peng, C., Quackenbush, J., Mar, J.C., Morton, C.C., Freedman, M.L., Blacklow, S.C., et al. (2011). Epstein-Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. *Proc Natl Acad Sci U S A* 108, 14902–14907. <https://doi.org/10.1073/pnas.1108892108>.
46. Pomerantz, M.M., Ahmadiyeh, N., Jia, L., Herman, P., Verzi, M.P., Doddapaneni, H., Beckwith, C.A., Chan, J.A., Hills, A., Davis, M., et al. (2009). The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat Genet* 41, 882–884. <https://doi.org/10.1038/ng.403>.
47. Dhanasekaran, R., Deutzmann, A., Mahauad-Fernandez, W.D., Hansen, A.S., Gouw, A.M., and Felsher, D.W. (2022). The MYC oncogene - the grand orchestrator of cancer growth and immune evasion. *Nat Rev Clin Oncol* 19, 23–36. <https://doi.org/10.1038/s41571-021-00549-2>.
48. Yin, H., Qu, J., Peng, Q., and Gan, R. (2019). Molecular mechanisms of EBV-driven cell cycle progression and oncogenesis. *Med Microbiol Immunol* 208, 573–583. <https://doi.org/10.1007/s00430-018-0570-1>.
49. Saha, A., and Robertson, E.S. (2013). Impact of EBV essential nuclear protein EBNA-3C on B-cell proliferation and apoptosis. *Future Microbiol* 8, 323–352. <https://doi.org/10.2217/fmb.12.147>.

50. Kalchschmidt, J.S., Gillman, A.C.T., Paschos, K., Bazot, Q., Kempkes, B., and Allday, M.J. (2016). EBNA3C Directs Recruitment of RBPJ (CBF1) to Chromatin during the Process of Gene Repression in EBV Infected B Cells. *PLoS Pathog* 12, e1005383. <https://doi.org/10.1371/journal.ppat.1005383>.
51. Hoffman, B., and Liebermann, D.A. (2008). Apoptotic signaling by c-MYC. *Oncogene* 27, 6462–6472. <https://doi.org/10.1038/onc.2008.312>.
52. Styles, C.T., Paschos, K., White, R.E., and Farrell, P.J. (2018). The Cooperative Functions of the EBNA3 Proteins Are Central to EBV Persistence and Latency. *Pathogens* 7, 31. <https://doi.org/10.3390/pathogens7010031>.
53. Maruo, S., Zhao, B., Johannsen, E., Kieff, E., Zou, J., and Takada, K. (2011). Epstein-Barr virus nuclear antigens 3C and 3A maintain lymphoblastoid cell growth by repressing p16INK4A and p14ARF expression. *Proc Natl Acad Sci U S A* 108, 1919–1924. <https://doi.org/10.1073/pnas.1019599108>.
54. Anderton, E., Yee, J., Smith, P., Crook, T., White, R.E., and Allday, M.J. (2008). Two Epstein-Barr virus (EBV) oncoproteins cooperate to repress expression of the proapoptotic tumour-suppressor Bim: clues to the pathogenesis of Burkitt's lymphoma. *Oncogene* 27, 421–433. <https://doi.org/10.1038/sj.onc.1210668>.
55. Lin, J., Johannsen, E., Robertson, E., and Kieff, E. (2002). Epstein-Barr virus nuclear antigen 3C putative repression domain mediates coactivation of the LMP1 promoter with EBNA-2. *J Virol* 76, 232–242. <https://doi.org/10.1128/jvi.76.1.232-242.2002>.
56. Zhao, B., Mar, J.C., Maruo, S., Lee, S., Gewurz, B.E., Johannsen, E., Holton, K., Rubio, R., Takada, K., Quackenbush, J., et al. (2011). Epstein-Barr virus nuclear antigen 3C regulated genes in lymphoblastoid cell lines. *Proc Natl Acad Sci U S A* 108, 337–342. <https://doi.org/10.1073/pnas.1017419108>.
57. White, R.E., Groves, I.J., Turro, E., Yee, J., Kremmer, E., and Allday, M.J. (2010). Extensive co-operation between the Epstein-Barr virus EBNA3 proteins in the manipulation of host gene expression and epigenetic chromatin modification. *PLoS One* 5, e13979. <https://doi.org/10.1371/journal.pone.0013979>.

58. Tao, Q., Robertson, K.D., Manns, A., Hildesheim, A., and Ambinder, R.F. (1998). The Epstein-Barr virus major latent promoter Qp is constitutively active, hypomethylated, and methylation sensitive. *J Virol* 72, 7075–7083. <https://doi.org/10.1128/JVI.72.9.7075-7083.1998>.
59. Schaefer, B.C., Strominger, J.L., and Speck, S.H. (1997). Host-cell-determined methylation of specific Epstein-Barr virus promoters regulates the choice between distinct viral latency programs. *Mol Cell Biol* 17, 364–377. <https://doi.org/10.1128/MCB.17.1.364>.
60. Reisman, D., Yates, J., and Sugden, B. (1985). A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two cis-acting components. *Mol Cell Biol* 5, 1822–1832. <https://doi.org/10.1128/mcb.5.8.1822-1832.1985>.
61. Moriyama, K., Yoshizawa-Sugata, N., Obuse, C., Tsurimoto, T., and Masai, H. (2012). Epstein-Barr nuclear antigen 1 (EBNA1)-dependent recruitment of origin recognition complex (Orc) on oriP of Epstein-Barr virus with purified proteins: stimulation by Cdc6 through its direct interaction with EBNA1. *J Biol Chem* 287, 23977–23994. <https://doi.org/10.1074/jbc.M112.368456>.
62. Mrozek-Gorska, P., Buschle, A., Pich, D., Schwarzmayer, T., Fechtner, R., Scialdone, A., and Hammerschmidt, W. (2019). Epstein-Barr virus reprograms human B lymphocytes immediately in the prelatent phase of infection. *Proc Natl Acad Sci U S A* 116, 16046–16055. <https://doi.org/10.1073/pnas.1901314116>.
63. Forte, E., Salinas, R.E., Chang, C., Zhou, T., Linnstaedt, S.D., Gottwein, E., Jacobs, C., Jima, D., Li, Q.-J., Dave, S.S., et al. (2012). The Epstein-Barr virus (EBV)-induced tumor suppressor microRNA MiR-34a is growth promoting in EBV-infected B cells. *J Virol* 86, 6889–6898. <https://doi.org/10.1128/JVI.07056-11>.
64. Soni, V., Cahir-McFarland, E., and Kieff, E. (2007). LMP1 TRAFficking activates growth and survival pathways. *Adv Exp Med Biol* 597, 173–187. https://doi.org/10.1007/978-0-387-70630-6_14.
65. Wang, L.W., Jiang, S., and Gewurz, B.E. (2017). Epstein-Barr Virus LMP1-Mediated

- Oncogenicity. *J Virol* 91, e01718-16. <https://doi.org/10.1128/JVI.01718-16>.
66. Capece, D., Verzella, D., Flati, I., Arboretto, P., Cornice, J., and Franzoso, G. (2022). NF- κ B: blending metabolism, immunity, and inflammation. *Trends Immunol* 43, 757–775. <https://doi.org/10.1016/j.it.2022.07.004>.
67. Kober-Hasslacher, M., Oh-Strauß, H., Kumar, D., Soberon, V., Diehl, C., Lech, M., Engleitner, T., Katab, E., Fernández-Sáiz, V., Piontek, G., et al. (2020). c-Rel gain in B cells drives germinal center reactions and autoantibody production. *J Clin Invest* 130, 3270–3286. <https://doi.org/10.1172/JCI124382>.
68. Klapproth, K., Sander, S., Marinkovic, D., Baumann, B., and Wirth, T. (2009). The IKK2/NF- κ B pathway suppresses MYC-induced lymphomagenesis. *Blood* 114, 2448–2458. <https://doi.org/10.1182/blood-2008-09-181008>.
69. Panagopoulos, D., Victoratos, P., Alexiou, M., Kollias, G., and Mosialos, G. (2004). Comparative analysis of signal transduction by CD40 and the Epstein-Barr virus oncoprotein LMP1 in vivo. *J Virol* 78, 13253–13261. <https://doi.org/10.1128/JVI.78.23.13253-13261.2004>.
70. Vrazo, A.C., Chauchard, M., Raab-Traub, N., and Longnecker, R. (2012). Epstein-Barr virus LMP2A reduces hyperactivation induced by LMP1 to restore normal B cell phenotype in transgenic mice. *PLoS Pathog* 8, e1002662. <https://doi.org/10.1371/journal.ppat.1002662>.
71. Yang, J., and Reth, M. (2016). Receptor Dissociation and B-Cell Activation. *Curr Top Microbiol Immunol* 393, 27–43. https://doi.org/10.1007/82_2015_482.
72. Rovedo, M., and Longnecker, R. (2007). Epstein-barr virus latent membrane protein 2B (LMP2B) modulates LMP2A activity. *J Virol* 81, 84–94. <https://doi.org/10.1128/JVI.01302-06>.
73. Price, A.M., and Luftig, M.A. (2014). Dynamic Epstein-Barr virus gene expression on the path to B-cell transformation. *Adv Virus Res* 88, 279–313. <https://doi.org/10.1016/B978-0-12-800098-4.00006-4>.
74. Price, A.M., and Luftig, M.A. (2015). To be or not IIb: a multi-step process for

- Epstein-Barr virus latency establishment and consequences for B cell tumorigenesis. *PLoS Pathog* 11, e1004656. <https://doi.org/10.1371/journal.ppat.1004656>.
75. Delgoffe, G.M., and Vignali, D.A.A. (2013). STAT heterodimers in immunity: A mixed message or a unique signal? *JAKSTAT* 2, e23060. <https://doi.org/10.4161/jkst.23060>.
76. Thorley-Lawson, D.A. (2015). EBV Persistence--Introducing the Virus. *Curr Top Microbiol Immunol* 390, 151–209. https://doi.org/10.1007/978-3-319-22822-8_8.
77. Roughan, J.E., and Thorley-Lawson, D.A. (2009). The intersection of Epstein-Barr virus with the germinal center. *J Virol* 83, 3968–3976. <https://doi.org/10.1128/JVI.02609-08>.
78. Hislop, A.D., Kuo, M., Drake-Lee, A.B., Akbar, A.N., Bergler, W., Hammerschmitt, N., Khan, N., Palendira, U., Leese, A.M., Timms, J.M., et al. (2005). Tonsillar homing of Epstein-Barr virus-specific CD8⁺ T cells and the virus-host balance. *J Clin Invest* 115, 2546–2555. <https://doi.org/10.1172/JCI24810>.
79. Palm, A.-K.E., and Henry, C. (2019). Remembrance of Things Past: Long-Term B Cell Memory After Infection and Vaccination. *Front Immunol* 10, 1787. <https://doi.org/10.3389/fimmu.2019.01787>.
80. Murata, T. (2023). Epstein-Barr virus: the molecular virology and the associated diseases. *Fujita Med J* 9, 65–72. <https://doi.org/10.20407/fmj.2022-018>.
81. Kenney, S.C., and Mertz, J.E. (2014). Regulation of the latent-lytic switch in Epstein-Barr virus. *Semin Cancer Biol* 26, 60–68. <https://doi.org/10.1016/j.semcancer.2014.01.002>.
82. Robinson, A.R., Kwek, S.S., and Kenney, S.C. (2012). The B-cell specific transcription factor, Oct-2, promotes Epstein-Barr virus latency by inhibiting the viral immediate-early protein, BZLF1. *PLoS Pathog* 8, e1002516. <https://doi.org/10.1371/journal.ppat.1002516>.
83. Raver, R.M., Panfil, A.R., Hagemeyer, S.R., and Kenney, S.C. (2013). The B-cell-specific transcription factor and master regulator Pax5 promotes Epstein-Barr virus latency

- by negatively regulating the viral immediate early protein BZLF1. *J Virol* 87, 8053–8063. <https://doi.org/10.1128/JVI.00546-13>.
84. Bryant, H., and Farrell, P.J. (2002). Signal Transduction and Transcription Factor Modification during Reactivation of Epstein-Barr Virus from Latency. *J Virol* 76, 10290–10298. <https://doi.org/10.1128/jvi.76.20.10290-10298.2002>.
85. Deng, Z., Fan, T., Xiao, C., Tian, H., Zheng, Y., Li, C., and He, J. (2024). TGF- β signaling in health, disease and therapeutics. *Sig Transduct Target Ther* 9, 61. <https://doi.org/10.1038/s41392-024-01764-w>.
86. Li, H., Liu, S., Hu, J., Luo, X., Li, N., M Bode, A., and Cao, Y. (2016). Epstein-Barr virus lytic reactivation regulation and its pathogenic role in carcinogenesis. *Int J Biol Sci* 12, 1309–1318. <https://doi.org/10.7150/ijbs.16564>.
87. Aubry, V., Mure, F., Mariamé, B., Deschamps, T., Wyrwicz, L.S., Manet, E., and Gruffat, H. (2014). Epstein-Barr virus late gene transcription depends on the assembly of a virus-specific preinitiation complex. *J Virol* 88, 12825–12838. <https://doi.org/10.1128/JVI.02139-14>.
88. Sugimoto, A. (2022). Replication Compartments-The Great Survival Strategy for Epstein-Barr Virus Lytic Replication. *Microorganisms* 10, 896. <https://doi.org/10.3390/microorganisms10050896>.
89. Liao, G., Wu, F.Y., and Hayward, S.D. (2001). Interaction with the Epstein-Barr virus helicase targets Zta to DNA replication compartments. *J Virol* 75, 8792–8802. <https://doi.org/10.1128/jvi.75.18.8792-8802.2001>.
90. Baumann, M., Feederle, R., Kremmer, E., and Hammerschmidt, W. (1999). Cellular transcription factors recruit viral replication proteins to activate the Epstein-Barr virus origin of lytic DNA replication, oriLyt. *EMBO J* 18, 6095–6105. <https://doi.org/10.1093/emboj/18.21.6095>.
91. Weller, S.K., and Coen, D.M. (2012). Herpes simplex viruses: mechanisms of DNA replication. *Cold Spring Harb Perspect Biol* 4, a013011. <https://doi.org/10.1101/cshperspect.a013011>.

92. Massacci, G., Perfetto, L., and Sacco, F. (2023). The Cyclin-dependent kinase 1: more than a cell cycle regulator. *Br J Cancer* 129, 1707–1716. <https://doi.org/10.1038/s41416-023-02468-8>.
93. Lo, A.K.-F., Dawson, C.W., Lung, H.L., Wong, K.-L., and Young, L.S. (2020). The Therapeutic Potential of Targeting BARF1 in EBV-Associated Malignancies. *Cancers (Basel)* 12, 1940. <https://doi.org/10.3390/cancers12071940>.
94. Fu, Q., He, C., and Mao, Z. (2013). Epstein-Barr virus interactions with the Bcl-2 protein family and apoptosis in human tumor cells. *J Zhejiang Univ Sci B* 14, 8–24. <https://doi.org/10.1631/jzus.B1200189>.
95. Caria, S., Chugh, S., Nhu, D., Lessene, G., and Kvansakul, M. (2012). Crystallization and preliminary X-ray characterization of Epstein-Barr virus BHRF1 in complex with a benzoylurea peptidomimetic. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 68, 1521–1524. <https://doi.org/10.1107/S1744309112043333>.
96. Cabras, G., Decaussin, G., Zeng, Y., Djennaoui, D., Melouli, H., Brouilly, P., Bouguermouh, A.M., and Ooka, T. (2005). Epstein-Barr virus encoded BALF1 gene is transcribed in Burkitt's lymphoma cell lines and in nasopharyngeal carcinoma's biopsies. *J Clin Virol* 34, 26–34. <https://doi.org/10.1016/j.jcv.2004.12.016>.
97. Bellows, D.S., Howell, M., Pearson, C., Hazlewood, S.A., and Hardwick, J.M. (2002). Epstein-Barr virus BALF1 is a BCL-2-like antagonist of the herpesvirus antiapoptotic BCL-2 proteins. *J Virol* 76, 2469–2479. <https://doi.org/10.1128/jvi.76.5.2469-2479.2002>.
98. Sugimoto, A., Yamashita, Y., Kanda, T., Murata, T., and Tsurumi, T. (2019). Epstein-Barr virus genome packaging factors accumulate in BMRF1-cores within viral replication compartments. *PLOS ONE* 14, e0222519. <https://doi.org/10.1371/journal.pone.0222519>.
99. Murayama, K., Nakayama, S., Kato-Murayama, M., Akasaka, R., Ohbayashi, N., Kamewari-Hayami, Y., Terada, T., Shirouzu, M., Tsurumi, T., and Yokoyama, S. (2009). Crystal structure of epstein-barr virus DNA polymerase processivity factor BMRF1. *J Biol Chem* 284, 35896–35905. <https://doi.org/10.1074/jbc.M109.051581>.
100. Dremel, S.E., and Didychuk, A.L. (2023). Better late than never: A unique strategy

- for late gene transcription in the beta- and gammaherpesviruses. *Semin Cell Dev Biol* 146, 57–69. <https://doi.org/10.1016/j.semcdb.2022.12.001>.
101. Gruffat, H., Kadjouf, F., Mariamé, B., and Manet, E. (2012). The Epstein-Barr virus BcRF1 gene product is a TBP-like protein with an essential role in late gene expression. *J Virol* 86, 6023–6032. <https://doi.org/10.1128/JVI.00159-12>.
102. Nishimura, M., Watanabe, T., Yagi, S., Yamanaka, T., and Fujimuro, M. (2017). Kaposi's sarcoma-associated herpesvirus ORF34 is essential for late gene expression and virus production. *Sci Rep* 7, 329. <https://doi.org/10.1038/s41598-017-00401-7>.
103. Dremel, S.E., and Didychuk, A.L. (2023). Better late than never: A unique strategy for late gene transcription in the beta- and gammaherpesviruses. *Semin Cell Dev Biol* 146, 57–69. <https://doi.org/10.1016/j.semcdb.2022.12.001>.
104. Yang, K., Poon, A.P.W., Roizman, B., and Baines, J.D. (2008). Temperature-sensitive mutations in the putative herpes simplex virus type 1 terminase subunits pUL15 and pUL33 preclude viral DNA cleavage/packaging and interaction with pUL28 at the nonpermissive temperature. *J Virol* 82, 487–494. <https://doi.org/10.1128/JVI.01875-07>.
105. Lv, Y., Zhou, S., Gao, S., and Deng, H. (2019). Remodeling of host membranes during herpesvirus assembly and egress. *Protein Cell* 10, 315–326. <https://doi.org/10.1007/s13238-018-0577-9>.
106. He, H.-P., Luo, M., Cao, Y.-L., Lin, Y.-X., Zhang, H., Zhang, X., Ou, J.-Y., Yu, B., Chen, X., Xu, M., et al. (2020). Structure of Epstein-Barr virus tegument protein complex BBRF2-BSRF1 reveals its potential role in viral envelopment. *Nat Commun* 11, 5405. <https://doi.org/10.1038/s41467-020-19259-x>.
107. Tsai, M.-H., Raykova, A., Klinke, O., Bernhardt, K., Gärtner, K., Leung, C.S., Geletneky, K., Sertel, S., Münz, C., Feederle, R., et al. (2013). Spontaneous lytic replication and epitheliotropism define an Epstein-Barr virus strain found in carcinomas. *Cell Rep* 5, 458–470. <https://doi.org/10.1016/j.celrep.2013.09.012>.
108. Nanbo, A. (2020). Epstein-Barr Virus Exploits the Secretory Pathway to Release Virions. *Microorganisms* 8, 729. <https://doi.org/10.3390/microorganisms8050729>.

109. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2025). A thorough insight into the life cycle of the Epstein-Barr virus. From the molecular to the organismal level. *Curr Res Microb Sci* 9, 100505. <https://doi.org/10.1016/j.crmicr.2025.100505>.
110. Huang, W., Bai, L., and Tang, H. (2023). Epstein-Barr virus infection: the micro and macro worlds. *Virol J* 20, 220. <https://doi.org/10.1186/s12985-023-02187-9>.
111. Rickinson, A.B., Long, H.M., Palendira, U., Münz, C., and Hislop, A.D. (2014). Cellular immune controls over Epstein-Barr virus infection: new lessons from the clinic and the laboratory. *Trends Immunol* 35, 159–169. <https://doi.org/10.1016/j.it.2014.01.003>.
112. Xuan, L., Jiang, X., Sun, J., Zhang, Y., Huang, F., Fan, Z., Guo, X., Dai, M., Liu, C., Yu, G., et al. (2013). Spectrum of Epstein-Barr virus-associated diseases in recipients of allogeneic hematopoietic stem cell transplantation. *Transplantation* 96, 560–566. <https://doi.org/10.1097/TP.0b013e31829d38af>.
113. Fujimoto, A., and Suzuki, R. (2020). Epstein-Barr Virus-Associated Post-Transplant Lymphoproliferative Disorders after Hematopoietic Stem Cell Transplantation: Pathogenesis, Risk Factors and Clinical Outcomes. *Cancers (Basel)* 12, 328. <https://doi.org/10.3390/cancers12020328>.
114. Patel, P.D., Alghareeb, R., Hussain, A., Maheshwari, M.V., and Khalid, N. (2022). The Association of Epstein-Barr Virus With Cancer. *Cureus* 14, e26314. <https://doi.org/10.7759/cureus.26314>.
115. Swerdlow, S.H., Campo, E., Pileri, S.A., Harris, N.L., Stein, H., Siebert, R., Advani, R., Ghielmini, M., Salles, G.A., Zelenetz, A.D., et al. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* 127, 2375–2390. <https://doi.org/10.1182/blood-2016-01-643569>.
116. Aldinucci, D., Celegato, M., and Casagrande, N. (2016). Microenvironmental interactions in classical Hodgkin lymphoma and their role in promoting tumor growth, immune escape and drug resistance. *Cancer Lett* 380, 243–252. <https://doi.org/10.1016/j.canlet.2015.10.007>.
117. Shannon-Lowe, C., Rickinson, A.B., and Bell, A.I. (2017). Epstein-Barr virus-

associated lymphomas. *Philos Trans R Soc Lond B Biol Sci* 372, 20160271. <https://doi.org/10.1098/rstb.2016.0271>.

118. de Martel, C., Georges, D., Bray, F., Ferlay, J., and Clifford, G.M. (2020). Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. *Lancet Glob Health* 8, e180–e190. [https://doi.org/10.1016/S2214-109X\(19\)30488-7](https://doi.org/10.1016/S2214-109X(19)30488-7).

119. Lu, T., Guo, Q., Lin, K., Chen, H., Chen, Y., Xu, Y., Lin, C., Su, Y., Chen, Y., Chen, M., et al. (2020). Circulating Epstein-Barr virus microRNAs BART7-3p and BART13-3p as novel biomarkers in nasopharyngeal carcinoma. *Cancer Sci* 111, 1711–1723. <https://doi.org/10.1111/cas.14381>.

120. Lin, D.-C., Meng, X., Hazawa, M., Nagata, Y., Varela, A.M., Xu, L., Sato, Y., Liu, L.-Z., Ding, L.-W., Sharma, A., et al. (2014). The genomic landscape of nasopharyngeal carcinoma. *Nat Genet* 46, 866–871. <https://doi.org/10.1038/ng.3006>.

121. Keith, K.A., Hartline, C.B., Bowlin, T.L., and Prichard, M.N. (2018). A standardized approach to the evaluation of antivirals against DNA viruses: Polyomaviruses and lymphotropic herpesviruses. *Antiviral Res* 159, 122–129. <https://doi.org/10.1016/j.antiviral.2018.09.016>.

122. Ville, S., Imbert-Marcille, B.-M., Coste-Burel, M., Garandeau, C., Meurette, A., Cantarovitch, D., Giral, M., Hourmant, M., Blancho, G., and Dantal, J. (2018). Impact of antiviral prophylaxis in adults Epstein-Barr Virus-seronegative kidney recipients on early and late post-transplantation lymphoproliferative disorder onset: a retrospective cohort study. *Transpl Int* 31, 484–494. <https://doi.org/10.1111/tri.13085>.

123. Østensen, A.B., Sanengen, T., Holter, E., Line, P.-D., and Almaas, R. (2017). No effect of treatment with intravenous ganciclovir on Epstein-Barr virus viremia demonstrated after pediatric liver transplantation. *Pediatr Transplant* 21. <https://doi.org/10.1111/petr.13010>.

124. Andrei, G., Trompet, E., and Snoeck, R. (2019). Novel Therapeutics for Epstein-Barr Virus. *Molecules* 24, 997. <https://doi.org/10.3390/molecules24050997>.

125. Biron, K.K., Harvey, R.J., Chamberlain, S.C., Good, S.S., Smith, A.A., Davis, M.G.,

- Talarico, C.L., Miller, W.H., Ferris, R., Dornsife, R.E., et al. (2002). Potent and selective inhibition of human cytomegalovirus replication by 1263W94, a benzimidazole L-riboside with a unique mode of action. *Antimicrob Agents Chemother* 46, 2365–2372. <https://doi.org/10.1128/AAC.46.8.2365-2372.2002>.
126. Zacny, V.L., Gershburg, E., Davis, M.G., Biron, K.K., and Pagano, J.S. (1999). Inhibition of Epstein-Barr virus replication by a benzimidazole L-riboside: novel antiviral mechanism of 5, 6-dichloro-2-(isopropylamino)-1-beta-L-ribofuranosyl-1H-benzimidazole. *J Virol* 73, 7271–7277. <https://doi.org/10.1128/JVI.73.9.7271-7277.1999>.
127. Gershburg, E., Hong, K., and Pagano, J.S. (2004). Effects of maribavir and selected indolocarbazoles on Epstein-Barr virus protein kinase BGLF4 and on viral lytic replication. *Antimicrob Agents Chemother* 48, 1900–1903. <https://doi.org/10.1128/AAC.48.5.1900-1903.2004>.
128. Wang, F.-Z., Roy, D., Gershburg, E., Whitehurst, C.B., Dittmer, D.P., and Pagano, J.S. (2009). Maribavir inhibits epstein-barr virus transcription in addition to viral DNA replication. *J Virol* 83, 12108–12117. <https://doi.org/10.1128/JVI.01575-09>.
129. Hardenbrook, N.J., and Zhang, P. (2022). A structural view of the SARS-CoV-2 virus and its assembly. *Curr Opin Virol* 52, 123–134. <https://doi.org/10.1016/j.coviro.2021.11.011>.
130. Yao, H., Song, Y., Chen, Y., Wu, N., Xu, J., Sun, C., Zhang, J., Weng, T., Zhang, Z., Wu, Z., et al. (2020). Molecular Architecture of the SARS-CoV-2 Virus. *Cell* 183, 730–738.e13. <https://doi.org/10.1016/j.cell.2020.09.018>.
131. Zhu, X., Zhang, X., Sun, Z., Zheng, Y., Su, S., Chen, F., Zhu, X., Zhang, X., Sun, Z., Zheng, Y., et al. (2021). Identification of Oil Tea (*Camellia oleifera* C.Abel) Cultivars Using EfficientNet-B4 CNN Model with Attention Mechanism. *Forests* 13. <https://doi.org/10.3390/f13010001>.
132. Mandala, V.S., McKay, M.J., Shcherbakov, A.A., Dregni, A.J., Kolocouris, A., and Hong, M. (2020). Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat Struct Mol Biol* 27, 1202–1208.

<https://doi.org/10.1038/s41594-020-00536-8>.

133. Yadav, R., Chaudhary, J.K., Jain, N., Chaudhary, P.K., Khanra, S., Dhamija, P., Sharma, A., Kumar, A., and Handu, S. (2021). Role of Structural and Non-Structural Proteins and Therapeutic Targets of SARS-CoV-2 for COVID-19. *Cells* 10, 821. <https://doi.org/10.3390/cells10040821>.

134. Ravi, V., Saxena, S., and Panda, P.S. (2022). Basic virology of SARS-CoV 2. *Indian J Med Microbiol* 40, 182–186. <https://doi.org/10.1016/j.ijmmb.2022.02.005>.

135. Sheikh, A., Al-Taher, A., Al-Nazawi, M., Al-Mubarak, A.I., and Kandeel, M. (2020). Analysis of preferred codon usage in the coronavirus N genes and their implications for genome evolution and vaccine design. *J Virol Methods* 277, 113806. <https://doi.org/10.1016/j.jviromet.2019.113806>.

136. Wu, S.-N., Xiao, T., Chen, H., and Li, X.-H. (2024). Decoding the genome of SARS-CoV-2: a pathway to drug development through translation inhibition. *RNA Biol* 21, 1–18. <https://doi.org/10.1080/15476286.2024.2433830>.

137. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O’Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 583, 459–468. <https://doi.org/10.1038/s41586-020-2286-9>.

138. Bassett, M., Salemi, M., and Rife Magalis, B. (2022). Lessons Learned and Yet-to-Be Learned on the Importance of RNA Structure in SARS-CoV-2 Replication. *Microbiol Mol Biol Rev* 86, e0005721. <https://doi.org/10.1128/mnbr.00057-21>.

139. Bayati, A., Kumar, R., Francis, V., and McPherson, P.S. (2021). SARS-CoV-2 infects cells after viral entry via clathrin-mediated endocytosis. *J Biol Chem* 296, 100306. <https://doi.org/10.1016/j.jbc.2021.100306>.

140. Borden, K.L.B., and Volpon, L. (2020). The diversity, plasticity, and adaptability of cap-dependent translation initiation and the associated machinery. *RNA Biol* 17, 1239–1251. <https://doi.org/10.1080/15476286.2020.1766179>.

141. Simeoni, M., Cavinato, T., Rodriguez, D., and Gatfield, D. (2021). I(nsp1)ecting

- SARS-CoV-2-ribosome interactions. *Commun Biol* 4, 715. <https://doi.org/10.1038/s42003-021-02265-0>.
142. Tam, D., Lorenzo-Leal, A.C., Hernández, L.R., and Bach, H. (2023). Targeting SARS-CoV-2 Non-Structural Proteins. *Int J Mol Sci* 24, 13002. <https://doi.org/10.3390/ijms241613002>.
143. von Brunn, A., Teepe, C., Simpson, J.C., Pepperkok, R., Friedel, C.C., Zimmer, R., Roberts, R., Baric, R., and Haas, J. (2007). Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome. *PLoS One* 2, e459. <https://doi.org/10.1371/journal.pone.0000459>.
144. Domingo, E., García-Crespo, C., Lobo-Vega, R., and Perales, C. (2021). Mutation Rates, Mutation Frequencies, and Proofreading-Repair Activities in RNA Virus Genetics. *Viruses* 13, 1882. <https://doi.org/10.3390/v13091882>.
145. Decroly, E., Ferron, F., Lescar, J., and Canard, B. (2011). Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* 10, 51–65. <https://doi.org/10.1038/nrmicro2675>.
146. Rosas-Lemus, M., Minasov, G., Shuvalova, L., Inniss, N.L., Kiryukhina, O., Wiersum, G., Kim, Y., Jedrzejczak, R., Maltseva, N.I., Endres, M., et al. (2020). The crystal structure of nsp10-nsp16 heterodimer from SARS-CoV-2 in complex with S-adenosylmethionine. *bioRxiv*, 2020.04.17.047498. <https://doi.org/10.1101/2020.04.17.047498>.
147. Sergio, M.C., Ricciardi, S., Guarino, A.M., Giaquinto, L., and De Matteis, M.A. (2024). Membrane remodeling and trafficking piloted by SARS-CoV-2. *Trends Cell Biol* 34, 785–800. <https://doi.org/10.1016/j.tcb.2023.12.006>.
148. Neuman, B.W., Kiss, G., Kunding, A.H., Bhella, D., Baksh, M.F., Connelly, S., Droese, B., Klaus, J.P., Makino, S., Sawicki, S.G., et al. (2011). A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol* 174, 11–22. <https://doi.org/10.1016/j.jsb.2010.11.021>.
149. Diogo, M.A., Cabral, A.G.T., and de Oliveira, R.B. (2024). Advances in the Search

- for SARS-CoV-2 Mpro and PLpro Inhibitors. *Pathogens* 13, 825. <https://doi.org/10.3390/pathogens13100825>.
150. Huang, J., Ma, Q., Su, Z., and Cheng, X. (2024). Advancements in the Development of Anti-SARS-CoV-2 Therapeutics. *Int J Mol Sci* 25, 10820. <https://doi.org/10.3390/ijms251910820>.
151. Dong, M., Galvan Achi, J.M., Du, R., Rong, L., and Cui, Q. (2024). Development of SARS-CoV-2 entry antivirals. *Cell Insight* 3, 100144. <https://doi.org/10.1016/j.cellin.2023.100144>.
152. Zhou, B., Chen, D., Zhang, T., Song, C., Zhang, X., Lin, L., Huang, J., Peng, X., Liu, Y., Wu, G., et al. (2024). Recent advancements in the discovery of small-molecule non-nucleoside inhibitors targeting SARS-CoV-2 RdRp. *Biomedicine & Pharmacotherapy* 171, 116180. <https://doi.org/10.1016/j.biopha.2024.116180>.
153. Singh, R.S.P., Toussi, S.S., Hackman, F., Chan, P.L., Rao, R., Allen, R., Van Eyck, L., Pawlak, S., Kadar, E.P., Clark, F., et al. (2022). Innovative Randomized Phase I Study and Dosing Regimen Selection to Accelerate and Inform Pivotal COVID-19 Trial of Nirmatrelvir. *Clin Pharmacol Ther* 112, 101–111. <https://doi.org/10.1002/cpt.2603>.
154. Citarella, A., Dimasi, A., Moi, D., Passarella, D., Scala, A., Piperno, A., and Micale, N. (2023). Recent Advances in SARS-CoV-2 Main Protease Inhibitors: From Nirmatrelvir to Future Perspectives. *Biomolecules* 13, 1339. <https://doi.org/10.3390/biom13091339>.
155. Joyce, R.P., Hu, V.W., and Wang, J. (2022). The history, mechanism, and perspectives of nirmatrelvir (PF-07321332): an orally bioavailable main protease inhibitor used in combination with ritonavir to reduce COVID-19-related hospitalizations. *Med Chem Res* 31, 1637–1646. <https://doi.org/10.1007/s00044-022-02951-6>.
156. Hashemian, S.M.R., Sheida, A., Taghizadieh, M., Memar, M.Y., Hamblin, M.R., Bannazadeh Baghi, H., Sadri Nahand, J., Asemi, Z., and Mirzaei, H. (2023). Paxlovid (Nirmatrelvir/Ritonavir): A new approach to Covid-19 therapy? *Biomed Pharmacother* 162, 114367. <https://doi.org/10.1016/j.biopha.2023.114367>.
157. Hillary, V.E., and Ceasar, S.A. (2023). An update on COVID-19: SARS-CoV-2

- variants, antiviral drugs, and vaccines. *Heliyon* 9, e13952. <https://doi.org/10.1016/j.heliyon.2023.e13952>.
158. Brian Chia, C.S., and Pheng Lim, S. (2023). A Patent Review on SARS Coronavirus Papain-Like Protease (PLpro) Inhibitors. *ChemMedChem* 18, e202300216. <https://doi.org/10.1002/cmdc.202300216>.
159. Sanders, B.C., Pokhrel, S., Labbe, A.D., Mathews, I.I., Cooper, C.J., Davidson, R.B., Phillips, G., Weiss, K.L., Zhang, Q., O'Neill, H., et al. (2023). Potent and selective covalent inhibition of the papain-like protease from SARS-CoV-2. *Nat Commun* 14, 1733. <https://doi.org/10.1038/s41467-023-37254-w>.
160. Wang, Q., Chen, G., He, J., Li, J., Xiong, M., Su, H., Li, M., Hu, H., and Xu, Y. (2023). Structure-Based Design of Potent Peptidomimetic Inhibitors Covalently Targeting SARS-CoV-2 Papain-like Protease. *Int J Mol Sci* 24, 8633. <https://doi.org/10.3390/ijms24108633>.
161. Bekheit, M.S., Panda, S.S., and Girgis, A.S. (2023). Potential RNA-dependent RNA polymerase (RdRp) inhibitors as prospective drug candidates for SARS-CoV-2. *Eur J Med Chem* 252, 115292. <https://doi.org/10.1016/j.ejmech.2023.115292>.
162. Lamb, Y.N. (2020). Remdesivir: First Approval. *Drugs* 80, 1355–1363. <https://doi.org/10.1007/s40265-020-01378-w>.
163. Kokic, G., Hillen, H.S., Tegunov, D., Dienemann, C., Seitz, F., Schmitzova, J., Farnung, L., Siewert, A., Höbartner, C., and Cramer, P. (2021). Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nat Commun* 12, 279. <https://doi.org/10.1038/s41467-020-20542-0>.
164. Humeniuk, R., Mathias, A., Kirby, B.J., Lutz, J.D., Cao, H., Osinusi, A., Babusis, D., Porter, D., Wei, X., Ling, J., et al. (2021). Pharmacokinetic, Pharmacodynamic, and Drug-Interaction Profile of Remdesivir, a SARS-CoV-2 Replication Inhibitor. *Clin Pharmacokinet* 60, 569–583. <https://doi.org/10.1007/s40262-021-00984-5>.
165. Udwardia, Z.F., Singh, P., Barkate, H., Patil, S., Rangwala, S., Pendse, A., Kadam, J., Wu, W., Caracta, C.F., and Tandon, M. (2021). Efficacy and safety of favipiravir, an oral

RNA-dependent RNA polymerase inhibitor, in mild-to-moderate COVID-19: A randomized, comparative, open-label, multicenter, phase 3 clinical trial. *Int J Infect Dis* 103, 62–71. <https://doi.org/10.1016/j.ijid.2020.11.142>.

166. Zhao, L., and Zhong, W. (2021). Mechanism of action of favipiravir against SARS-CoV-2: Mutagenesis or chain termination? *Innovation (Camb)* 2, 100165. <https://doi.org/10.1016/j.xinn.2021.100165>.

167. Rattanaumpawan, P., Jirajariyavej, S., Lerdlamyong, K., Palavutitotai, N., and Saiyarin, J. (2022). Real-World Effectiveness and Optimal Dosage of Favipiravir for Treatment of COVID-19: Results from a Multicenter Observational Study in Thailand. *Antibiotics (Basel)* 11, 805. <https://doi.org/10.3390/antibiotics11060805>.

168. Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., Shi, Z., Hu, Z., Zhong, W., and Xiao, G. (2020). Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res* 30, 269–271. <https://doi.org/10.1038/s41422-020-0282-0>.

169. Chenthamarakshan, V., Hoffman, S.C., Owen, C.D., Lukacik, P., Strain-Damerell, C., Fearon, D., Malla, T.R., Tumber, A., Schofield, C.J., Duyvesteyn, H.M.E., et al. (2023). Accelerating drug target inhibitor discovery with a deep generative foundation model. *Sci Adv* 9, eadg7865. <https://doi.org/10.1126/sciadv.adg7865>.

170. Toelzer, C., Gupta, K., Yadav, S.K.N., Borucu, U., Davidson, A.D., Kavanagh Williamson, M., Shoemark, D.K., Garzoni, F., Staufer, O., Milligan, R., et al. (2020). Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. *Science* 370, 725–730. <https://doi.org/10.1126/science.abd3255>.

171. Kao, R.Y., Tsui, W.H.W., Lee, T.S.W., Tanner, J.A., Watt, R.M., Huang, J.-D., Hu, L., Chen, G., Chen, Z., Zhang, L., et al. (2004). Identification of novel small-molecule inhibitors of severe acute respiratory syndrome-associated coronavirus by chemical genetics. *Chem Biol* 11, 1293–1299. <https://doi.org/10.1016/j.chembiol.2004.07.013>.

172. Ding, S., Ullah, I., Gong, S.Y., Grover, J.R., Mohammadi, M., Chen, Y., Vézina, D., Beaudoin-Bussi eres, G., Verma, V.T., Goyette, G., et al. (2022). VE607 stabilizes SARS-

CoV-2 Spike in the “RBD-up” conformation and inhibits viral entry. *iScience* 25, 104528. <https://doi.org/10.1016/j.isci.2022.104528>.

173. DiMasi, J.A., Hansen, R.W., and Grabowski, H.G. (2003). The price of innovation: new estimates of drug development costs. *Journal of Health Economics* 22, 151–185. [https://doi.org/10.1016/S0167-6296\(02\)00126-1](https://doi.org/10.1016/S0167-6296(02)00126-1).

174. van Montfort, R.L.M., and Workman, P. (2017). Structure-based drug design: aiming for a perfect fit. *Essays Biochem* 61, 431–437. <https://doi.org/10.1042/EBC20170052>.

175. Wang, K., Huang, Y., Wang, Y., You, Q., and Wang, L. (2024). Recent advances from computer-aided drug design to artificial intelligence drug design. *RSC Med. Chem.* 15, 3978–4000. <https://doi.org/10.1039/D4MD00522H>.

176. Saharkhiz, S., Mostafavi, M., Birashk, A., Karimian, S., Khalilollah, S., Jaferian, S., Yazdani, Y., Alipourfard, I., Huh, Y.S., Farani, M.R., et al. (2024). The State-of-the-Art Overview to Application of Deep Learning in Accurate Protein Design and Structure Prediction. *Top Curr Chem (Cham)* 382, 23. <https://doi.org/10.1007/s41061-024-00469-6>.

177. Wuyun, Q., Chen, Y., Shen, Y., Cao, Y., Hu, G., Cui, W., Gao, J., and Zheng, W. (2024). Recent Progress of Protein Tertiary Structure Prediction. *Molecules* 29, 832. <https://doi.org/10.3390/molecules29040832>.

178. Wu, S., and Zhang, Y. (2008). MUSTER: Improving protein sequence profile-profile alignments by using multiple sources of structure information. *Proteins* 72, 547–556. <https://doi.org/10.1002/prot.21945>.

179. Chen, L., Li, Q., Nasif, K.F.A., Xie, Y., Deng, B., Niu, S., Pouriye, S., Dai, Z., Chen, J., and Xie, C.Y. (2024). AI-Driven Deep Learning Techniques in Protein Structure Prediction. *Int J Mol Sci* 25, 8426. <https://doi.org/10.3390/ijms25158426>.

180. Callaway, E. (2020). “It will change everything”: DeepMind’s AI makes gigantic leap in solving protein structures. *Nature* 588, 203–204. <https://doi.org/10.1038/d41586-020-03348-4>.

181. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Applying and

improving AlphaFold at CASP14. *Proteins* 89, 1711–1721. <https://doi.org/10.1002/prot.26257>.

182. Moult, J., Pedersen, J.T., Judson, R., and Fidelis, K. (1995). A large-scale experiment to assess protein structure prediction methods. *Proteins* 23, ii–v. <https://doi.org/10.1002/prot.340230303>.

183. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.

184. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. ukasz, and Polosukhin, I. (2017). Attention is All you Need. In *Advances in Neural Information Processing Systems* (Curran Associates, Inc.). Available at: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

185. UniProt Consortium. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 49, D480–D489. <https://doi.org/10.1093/nar/gkaa1100>.

186. Kaserer, T., Beck, K.R., Akram, M., Odermatt, A., and Schuster, D. (2015). Pharmacophore Models and Pharmacophore-Based Virtual Screening: Concepts and Applications Exemplified on Hydroxysteroid Dehydrogenases. *Molecules* 20, 22799–22832. <https://doi.org/10.3390/molecules201219880>.

187. Blanes-Mira, C., Fernández-Aguado, P., de Andrés-López, J., Fernández-Carvajal, A., Ferrer-Montiel, A., and Fernández-Ballester, G. (2022). Comprehensive Survey of Consensus Docking for High-Throughput Virtual Screening. *Molecules* 28, 175. <https://doi.org/10.3390/molecules28010175>.

188. Lee, A.C.-L., Harris, J.L., Khanna, K.K., and Hong, J.-H. (2019). A Comprehensive Review on Current Advances in Peptide Drug Development and Design. *Int J Mol Sci* 20, 2383. <https://doi.org/10.3390/ijms20102383>.

189. Pinzi, L., and Rastelli, G. (2019). Molecular Docking: Shifting Paradigms in Drug

- Discovery. *Int J Mol Sci* 20, 4331. <https://doi.org/10.3390/ijms20184331>.
190. Wang, J., and Dokholyan, N.V. (2019). MedusaDock 2.0: Efficient and Accurate Protein-Ligand Docking With Constraints. *J Chem Inf Model* 59, 2509–2515. <https://doi.org/10.1021/acs.jcim.8b00905>.
191. Yadava, U. (2018). Search algorithms and scoring methods in protein-ligand docking. *Endocrinology&Metabolism International Journal* 6. <https://doi.org/10.15406/emij.2018.06.00212>.
192. Huang, S.-Y., and Zou, X. (2010). Advances and challenges in protein-ligand docking. *Int J Mol Sci* 11, 3016–3034. <https://doi.org/10.3390/ijms11083016>.
193. Lorenzen, S., and Zhang, Y. (2007). Monte Carlo refinement of rigid-body protein docking structures with backbone displacement and side-chain optimization. *Protein Sci* 16, 2716–2725. <https://doi.org/10.1110/ps.072847207>.
194. Li, J., Fu, A., and Zhang, L. (2019). An Overview of Scoring Functions Used for Protein-Ligand Interactions in Molecular Docking. *Interdiscip Sci* 11, 320–328. <https://doi.org/10.1007/s12539-019-00327-w>.
195. Maia, E.H.B., Assis, L.C., de Oliveira, T.A., da Silva, A.M., and Taranto, A.G. (2020). Structure-Based Virtual Screening: From Classical to Artificial Intelligence. *Front Chem* 8, 343. <https://doi.org/10.3389/fchem.2020.00343>.
196. Blay, V., Tolani, B., Ho, S.P., and Arkin, M.R. (2020). High-Throughput Screening: today's biochemical and cell-based approaches. *Drug Discov Today* 25, 1807–1821. <https://doi.org/10.1016/j.drudis.2020.07.024>.
197. Talarico, C., Gervasoni, S., Manelfi, C., Pedretti, A., Vistoli, G., and Beccari, A.R. (2020). Combining Molecular Dynamics and Docking Simulations to Develop Targeted Protocols for Performing Optimized Virtual Screening Campaigns on The hTRPM8 Channel. *Int J Mol Sci* 21, 2265. <https://doi.org/10.3390/ijms21072265>.
198. Hollingsworth, S.A., and Dror, R.O. (2018). Molecular Dynamics Simulation for All. *Neuron* 99, 1129–1143. <https://doi.org/10.1016/j.neuron.2018.08.011>.
199. Chen, I.J., Yin, D., and MacKerell Jr., A.D. (2002). Combined ab initio/empirical

- approach for optimization of Lennard-Jones parameters for polar-neutral compounds. *Journal of Computational Chemistry* 23, 199–213. <https://doi.org/10.1002/jcc.1166>.
200. Fabiola, F., Bertram, R., Korostelev, A., and Chapman, M.S. (2002). An improved hydrogen bond potential: impact on medium resolution protein structures. *Protein Sci* 11, 1415–1423. <https://doi.org/10.1110/ps.4890102>.
201. Sami, S., Menger, M.F.S.J., Faraji, S., Broer, R., and Havenith, R.W.A. (2021). Q-Force: Quantum Mechanically Augmented Molecular Force Fields. *J Chem Theory Comput* 17, 4946–4960. <https://doi.org/10.1021/acs.jctc.1c00195>.
202. Dos Santos Nascimento, I.J., Gomes, J.N.S., de Oliveira Viana, J., de Medeiros E Silva, Y.M.S., Barbosa, E.G., and de Moura, R.O. (2024). The Power of Molecular Dynamics Simulations and Their Applications to Discover Cysteine Protease Inhibitors. *Mini Rev Med Chem* 24, 1125–1146. <https://doi.org/10.2174/1389557523666230901152257>.
203. Ghahremanian, S., Rashidi, M.M., Raeisi, K., and Toghraie, D. (2022). Molecular dynamics simulation approach for discovering potential inhibitors against SARS-CoV-2: A structural review. *J Mol Liq* 354, 118901. <https://doi.org/10.1016/j.molliq.2022.118901>.
204. Poli, G., Granchi, C., Rizzolio, F., and Tuccinardi, T. (2020). Application of MM-PBSA Methods in Virtual Screening. *Molecules* 25, 1971. <https://doi.org/10.3390/molecules25081971>.
205. King, E., Aitchison, E., Li, H., and Luo, R. (2021). Recent Developments in Free Energy Calculations for Drug Discovery. *Front Mol Biosci* 8, 712085. <https://doi.org/10.3389/fmolb.2021.712085>.
206. Genheden, S., and Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* 10, 449–461. <https://doi.org/10.1517/17460441.2015.1032936>.
207. Lin, X., Li, X., and Lin, X. (2020). A Review on Applications of Computational Methods in Drug Screening and Design. *Molecules* 25, 1375. <https://doi.org/10.3390/molecules25061375>.

208. Allen, W.J., Fochtman, B.C., Balius, T.E., and Rizzo, R.C. (2017). Customizable de novo design strategies for DOCK: Application to HIVgp41 and other therapeutic targets. *J Comput Chem* 38, 2641–2663. <https://doi.org/10.1002/jcc.25052>.
209. Nishibata, Y., and Itai, A. (1991). Automatic creation of drug candidate structures based on receptor structure. Starting point for artificial lead generation. *Tetrahedron* 47, 8985–8990. [https://doi.org/10.1016/S0040-4020\(01\)86503-0](https://doi.org/10.1016/S0040-4020(01)86503-0).
210. de Souza Neto, L.R., Moreira-Filho, J.T., Neves, B.J., Maidana, R.L.B.R., Guimarães, A.C.R., Furnham, N., Andrade, C.H., and Silva, F.P. (2020). In silico Strategies to Support Fragment-to-Lead Optimization in Drug Discovery. *Front Chem* 8, 93. <https://doi.org/10.3389/fchem.2020.00093>.
211. Wang, Y., Zhao, H., Sciabola, S., and Wang, W. (2023). cMolGPT: A Conditional Generative Pre-Trained Transformer for Target-Specific De Novo Molecular Generation. *Molecules* 28, 4430. <https://doi.org/10.3390/molecules28114430>.
212. Martinelli, D.D. (2022). Generative machine learning for de novo drug discovery: A systematic review. *Comput Biol Med* 145, 105403. <https://doi.org/10.1016/j.combiomed.2022.105403>.
213. Grygorenko, O.O., Radchenko, D.S., Dziuba, I., Chuprina, A., Gubina, K.E., and Moroz, Y.S. (2020). Generating Multibillion Chemical Space of Readily Accessible Screening Compounds. *iScience* 23, 101681. <https://doi.org/10.1016/j.isci.2020.101681>.
214. Tang, Y., Moretti, R., and Meiler, J. (2024). Recent Advances in Automated Structure-Based De Novo Drug Design. *J Chem Inf Model* 64, 1794–1805. <https://doi.org/10.1021/acs.jcim.4c00247>.
215. Sundin, I., Voronov, A., Xiao, H., Papadopoulos, K., Bjerrum, E.J., Heinonen, M., Patronov, A., Kaski, S., and Engkvist, O. (2022). Human-in-the-loop assisted de novo molecular design. *J Cheminform* 14, 86. <https://doi.org/10.1186/s13321-022-00667-8>.
216. Hodcroft, E.B. (2021). "CoVariants: SARS-CoV-2 Mutations and Variants of Interest." Available at: <https://covariants.org/>
217. Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to

track COVID-19 in real time. *Lancet Infect Dis* 20, 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1).

218. Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Gavrilov, D., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., et al. (2020). Coronavirus (COVID-19) Vaccinations. Our World in Data. Available at: <https://ourworldindata.org/>

219. Google LLC "Google COVID-19 Community Mobility Reports". Available at: <https://www.google.com/covid19/mobility/>

220. Pihan, E., Colliandre, L., Guichou, J.-F., and Douguet, D. (2012). e-Drug3D: 3D structure collections dedicated to drug repurposing and fragment-based drug design. *Bioinformatics* 28, 1540–1541. <https://doi.org/10.1093/bioinformatics/bts186>.

221. RDKit: Open-source cheminformatics. Available at: <https://www.rdkit.org>

222. O’Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., and Hutchison, G.R. (2011). Open Babel: An open chemical toolbox. *J Cheminform* 3, 33. <https://doi.org/10.1186/1758-2946-3-33>.

223. Hanwell, M.D., Curtis, D.E., Lonie, D.C., Vandermeersch, T., Zurek, E., and Hutchison, G.R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminform* 4, 17. <https://doi.org/10.1186/1758-2946-4-17>.

224. Sayers, E.W., Beck, J., Bolton, E.E., Brister, J.R., Chan, J., Connor, R., Feldgarden, M., Fine, A.M., Funk, K., Hoffman, J., et al. (2025). Database resources of the National Center for Biotechnology Information in 2025. *Nucleic Acids Res* 53, D20–D29. <https://doi.org/10.1093/nar/gkae979>.

225. Eberhardt, J., Santos-Martins, D., Tillack, A.F., and Forli, S. (2021). AutoDock Vina 1.2.0: new docking methods, expanded force field, and Python bindings. *J Chem Inf Model* 61, 3891–3898. <https://doi.org/10.1021/acs.jcim.1c00203>.

226. Zaremba, A., Zaremba, P., Muchnyk, F.V., Baranova, G., and Zahorodnia, S. (2021). In silico Identification of a Viral Surface Glycoprotein Site Suitable for the Development of Low Molecular Weight Inhibitors for Various Variants of the SARS-CoV-2. *Mikrobiolohichni Zhurnal* 84, 34–43. <https://doi.org/10.15407/microbiolj84.01.034>.

227. Abraham, M., Murtola, T., Schulz, R., Páll, S., Smith, J., Hess, B., and Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1. <https://doi.org/10.1016/j.softx.2015.06.001>.
228. Mark, P., and Nilsson, L. (2001). Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* 105, 9954–9960. <https://doi.org/10.1021/jp003020w>.
229. Huang, J., and MacKerell Jr, A.D. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry* 34, 2135–2145. <https://doi.org/10.1002/jcc.23354>.
230. Tian, C., Kasavajhala, K., Belfon, K.A.A., Raguetta, L., Huang, H., Migués, A.N., Bickel, J., Wang, Y., Pincay, J., Wu, Q., et al. (2020). ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* 16, 528–552. <https://doi.org/10.1021/acs.jctc.9b00591>.
231. Zgarbová, M., Šponer, J., Otyepka, M., Cheatham, T.E.I., Galindo-Murillo, R., and Jurečka, P. (2015). Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *J. Chem. Theory Comput.* 11, 5723–5736. <https://doi.org/10.1021/acs.jctc.5b00716>.
232. Bugnon, M., Goullieux, M., Röhrig, U.F., Perez, M.A.S., Daina, A., Michielin, O., and Zoete, V. (2023). SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization. *J Chem Inf Model* 63, 6469–6475. <https://doi.org/10.1021/acs.jcim.3c01053>.
233. Kagami, L., Wilter, A., Diaz, A., and Vranken, W. (2023). The ACPYPE web server for small-molecule MD topology generation. *Bioinformatics* 39, btad350. <https://doi.org/10.1093/bioinformatics/btad350>.
234. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., and Case, D.A. (2004). Development and testing of a general amber force field. *J Comput Chem* 25, 1157–1174. <https://doi.org/10.1002/jcc.20035>.

235. Zhao, S., Schaub, A.J., Tsai, S.-C., and Luo, R. (2021). Development of a Pantetheine Force Field Library for Molecular Modeling. *J Chem Inf Model* 61, 856–868. <https://doi.org/10.1021/acs.jcim.0c01384>.
236. Hess, B., Bekker, H., Berendsen, H.J.C., and Fraaije, J.G.E.M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* 18, 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12%253C1463::AID-JCC4%253E3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12%253C1463::AID-JCC4%253E3.0.CO;2-H).
237. Valdés-Tresanco, M.S., Valdés-Tresanco, M.E., Valiente, P.A., and Moreno, E. (2021). gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J Chem Theory Comput* 17, 6281–6291. <https://doi.org/10.1021/acs.jctc.1c00645>.
238. Turner, P. J. (2005). XMGRACE, Version 5.1.19. Beaverton, OR: Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology
239. Dassault Systèmes. Available at: <https://discover.3ds.com/discovery-studio-visualizer-download>
240. Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B.A., Thiessen, P.A., Yu, B., et al. (2025). PubChem 2025 update. *Nucleic Acids Res* 53, D1516–D1525. <https://doi.org/10.1093/nar/gkae1059>.
241. Zdrzil, B., Felix, E., Hunter, F., Manners, E.J., Blackshaw, J., Corbett, S., de Veij, M., Ioannidis, H., Lopez, D.M., Mosquera, J.F., et al. (2024). The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res* 52, D1180–D1192. <https://doi.org/10.1093/nar/gkad1004>.
242. Irwin, J.J., Tang, K.G., Young, J., Dandarchuluun, C., Wong, B.R., Khurelbaatar, M., Moroz, Y.S., Mayfield, J., and Sayle, R.A. (2020). ZINC20-A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *J Chem Inf Model* 60, 6065–6073. <https://doi.org/10.1021/acs.jcim.0c00675>.
243. Xiong, G., Wu, Z., Yi, J., Fu, L., Yang, Z., Hsieh, C., Yin, M., Zeng, X., Wu, C., Lu, A., et al. (2021). ADMETlab 2.0: an integrated online platform for accurate and

comprehensive predictions of ADMET properties. *Nucleic Acids Res* 49, W5–W14. <https://doi.org/10.1093/nar/gkab255>.

244. Reddy, C.N., and Sankararamakrishnan, R. (2021). Designing BH3-Mimetic Peptide Inhibitors for the Viral Bcl-2 Homologues A179L and BHRF1: Importance of Long-Range Electrostatic Interactions. *ACS Omega* 6, 26976–26989. <https://doi.org/10.1021/acsomega.1c03385>.

245. Hemmer, B., Kondo, T., Gran, B., Pinilla, C., Cortese, I., Pascal, J., Tzou, A., McFarland, H.F., Houghten, R., and Martin, R. (2000). Minimal peptide length requirements for CD4(+) T cell clones--implications for molecular mimicry and T cell survival. *Int Immunol* 12, 375–383. <https://doi.org/10.1093/intimm/12.3.375>.

246. Zaremba, A.A., Zaremba, P.Y., and Platonov, M.O. (2023). De novo designed EBAI as a potential inhibitor of the viral protein BHRF1. Research in silico. *J Biomol Struct Dyn* 41, 3680–3685. <https://doi.org/10.1080/07391102.2022.2053746>.

247. Rescifina, A., Zagni, C., Varrica, M.G., Pistarà, V., and Corsaro, A. (2014). Recent advances in small organic molecules as DNA intercalating agents: synthesis, activity, and modeling. *Eur J Med Chem* 74, 95–115. <https://doi.org/10.1016/j.ejmech.2013.11.029>.

248. Buric, A.J., Dickerhoff, J., and Yang, D. (2021). Novel DNA Bis-Intercalator XR5944 as a Potent Anticancer Drug-Design and Mechanism of Action. *Molecules* 26, 4132. <https://doi.org/10.3390/molecules26144132>.

249. Zaremba, A., Zaremba, P., and Zagorodnya, S. Selective DNA intercalation of massive molecules as a new method of highly specific inhibition of transcription. 6th International Electronic Conference on Medicinal Chemistry, 1–30 November 2020, Basel, Switzerland. <https://doi.org/10.3390/ECMC2020-07364>.

250. Zaremba, A.A., Zaremba, P.Y., and Zahorodnia, S.D. (2023). In silico study of HASDI (high-affinity selective DNA intercalator) as a new agent capable of highly selective recognition of the DNA sequence. *Sci Rep* 13, 5395. <https://doi.org/10.1038/s41598-023-32595-4>.

251. Sasaki, K., Irie, T., Nozaki, M., Kawawaki, T., Das, S., and Negishi, Y. (2025). A

- phenazine-linked π -conjugated covalent organic framework for conjugation-driven drug loading. *Nanoscale Horizons* 10, 3309–3318. <https://doi.org/10.1039/D5NH00470E>.
252. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2025). In silico development of HASDI-G2 as a novel agent for selective recognition of the DNA sequence. *Sci Rep* 15, 8577. <https://doi.org/10.1038/s41598-025-89967-1>.
253. Osman, A.E.G., and Deininger, M.W. (2021). Chronic Myeloid Leukemia: Modern therapies, current challenges and future directions. *Blood Rev* 49, 100825. <https://doi.org/10.1016/j.blre.2021.100825>.
254. Gao, Q., Ouyang, W., Kang, B., Han, X., Xiong, Y., Ding, R., Li, Y., Wang, F., Huang, L., Chen, L., et al. (2020). Selective targeting of the oncogenic KRAS G12S mutant allele by CRISPR/Cas9 induces efficient tumor regression. *Theranostics* 10, 5137–5153. <https://doi.org/10.7150/thno.42325>.
255. Khan, M.Z.I., Nazli, A., Al-Furas, H., Asad, M.I., Ajmal, I., Khan, D., Shah, J., Farooq, M.A., and Jiang, W. (2022). An overview of viral mutagenesis and the impact on pathogenesis of SARS-CoV-2 variants. *Front Immunol* 13, 1034444. <https://doi.org/10.3389/fimmu.2022.1034444>.
256. Zaremba, A., Zaremba, P., Budzanivska, I., and Zahorodnia, S. (2022). Patterns of the influence of vaccination on the dynamics of different SARS-CoV-2 variants spread. Two-year analysis. *Bulletin of Taras Shevchenko National University of Kyiv Series Biology* 89, 39–45. DOI: <https://doi.org/10.17721/1728.2748.2022.89.39-45>.
257. Boytchev, H. (2020). How Germany retains one of the world’s strongest research reputations. *Nature* 587, S104–S105. <https://doi.org/10.1038/d41586-020-03318-w>.
258. Tao, K., Tzou, P.L., Nouhin, J., Gupta, R.K., de Oliveira, T., Kosakovsky Pond, S.L., Fera, D., and Shafer, R.W. (2021). The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat Rev Genet* 22, 757–773. <https://doi.org/10.1038/s41576-021-00408-x>.
259. Zhang, Y., Olufadewa, I.I., Adesina, M.A., Ekpo, M.D., Akinloye, S.J., Iyanda, T.O., Nwachukwu, P., and Kodzo, L.D. (2021). Lessons from the coronavirus disease 2019

- (COVID-19) pandemic response in China, Italy, and the U.S.: a guide for Africa and low- and middle-income countries. *Glob Health J* 5, 56–61. <https://doi.org/10.1016/j.glohj.2021.02.003>.
260. Machado, B.A.S., Hodel, K.V.S., Fonseca, L.M.D.S., Pires, V.C., Mascarenhas, L.A.B., da Silva Andrade, L.P.C., Moret, M.A., and Badaró, R. (2022). The Importance of Vaccination in the Context of the COVID-19 Pandemic: A Brief Update Regarding the Use of Vaccines. *Vaccines (Basel)* 10, 591. <https://doi.org/10.3390/vaccines10040591>.
261. Icard, P., Simula, L., Rei, J., Fournel, L., De Pauw, V., and Alifano, M. (2021). On the footsteps of Hippocrates, Sanctorius and Harvey to better understand the influence of cold on the occurrence of COVID-19 in European countries in 2020. *Biochimie* 191, 164–171. <https://doi.org/10.1016/j.biochi.2021.09.009>.
262. Khare, S., Gurry, C., Freitas, L., Schultz, M.B., Bach, G., Diallo, A., Akite, N., Ho, J., Lee, R.T., Yeo, W., et al. (2021). GISAID's Role in Pandemic Response. *China CDC Wkly* 3, 1049–1051. <https://doi.org/10.46234/ccdcw2021.255>.
263. Ahmad, L. (2021). Implication of SARS-CoV-2 Immune Escape Spike Variants on Secondary and Vaccine Breakthrough Infections. *Front Immunol* 12, 742167. <https://doi.org/10.3389/fimmu.2021.742167>.
264. Mourya, D.T., Yadav, P.D., Ullas, P.T., Bhardwaj, S.D., Sahay, R.R., Chadha, M.S., Shete, A.M., Jadhav, S., Gupta, N., Gangakhedkar, R.R., et al. (2019). Emerging/re-emerging viral diseases & new viruses on the Indian horizon. *Indian J Med Res* 149, 447–467. https://doi.org/10.4103/ijmr.IJMR_1239_18.
265. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2023). De novo designed inhibitor has high affinity to four variants of the RBD of S-glycoprotein of SARS-CoV-2 - an in silico study. *J Biomol Struct Dyn* 41, 9389–9397. <https://doi.org/10.1080/07391102.2022.2141886>.
266. Lipinski, C.A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods* 44, 235–249. [https://doi.org/10.1016/s1056-8719\(00\)00107-6](https://doi.org/10.1016/s1056-8719(00)00107-6).

267. Markopoulos, C., Thoenen, F., Preisig, D., Symillides, M., Vertzoni, M., Parrott, N., Reppas, C., and Imanidis, G. (2014). Biorelevant media for transport experiments in the Caco-2 model to evaluate drug absorption in the fasted and the fed state and their usefulness. *Eur J Pharm Biopharm* 86, 438–448. <https://doi.org/10.1016/j.ejpb.2013.10.017>.
268. Bittermann, K., and Goss, K.-U. (2017). Predicting apparent passive permeability of Caco-2 and MDCK cell-monolayers: A mechanistic model. *PLoS One* 12, e0190319. <https://doi.org/10.1371/journal.pone.0190319>.
269. Ertl, P., and Schuffenhauer, A. (2009). Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J Cheminform* 1, 8. <https://doi.org/10.1186/1758-2946-1-8>.
270. Ramírez, D., and Caballero, J. (2018). Is It Reliable to Take the Molecular Docking Top Scoring Position as the Best Solution without Considering Available Structural Data? *Molecules* 23, 1038. <https://doi.org/10.3390/molecules23051038>.
271. Seeliger, D., and de Groot, B.L. (2010). Conformational transitions upon ligand binding: holo-structure prediction from apo conformations. *PLoS Comput Biol* 6, e1000634. <https://doi.org/10.1371/journal.pcbi.1000634>.
272. Tuz-Zohura, F.-, Shawon, A.R.M., Hasan, M.M., Aeyas, A., Chowdhury, F.I., and Khandaker, M.U. (2023). In-silico approach to designing effective antiviral drugs against SARS-CoV-2 and SARS-CoV-1 from reported phytochemicals: a quality improvement study. *Ann Med Surg (Lond)* 85, 3446–3460. <https://doi.org/10.1097/MS9.0000000000000839>.
273. Lipinski, C.A. (2004). Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies* 1, 337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007>.
274. Herschlag, D., and Pinney, M.M. (2018). Hydrogen Bonds: Simple after All? *Biochemistry* 57, 3338–3352. <https://doi.org/10.1021/acs.biochem.8b00217>.
275. Cirac, A., Poirey, R., Dieckmeyer, M., Witter, K., Delecluse, H.-J., Behrends, U., and

- Mautner, J. (2021). Immunoinformatic Analysis Reveals Antigenic Heterogeneity of Epstein-Barr Virus Is Immune-Driven. *Front Immunol* 12, 796379. <https://doi.org/10.3389/fimmu.2021.796379>.
276. Morscio, J., and Tousseyn, T. (2016). Recent insights in the pathogenesis of post-transplantation lymphoproliferative disorders. *World J Transplant* 6, 505–516. <https://doi.org/10.5500/wjt.v6.i3.505>.
277. World Health Organization. Available at: <https://www.who.int/europe/emergencies/situations/covid-19#>
278. Dofuor, A.K., Quartey, N.K.-A., Osabutey, A.F., Boateng, B.O., Lutuf, H., Osei, J.H.N., Ayivi-Tosuh, S.M., Aiduenu, A.F., Ekloh, W., Loh, S.K., et al. (2023). The Global Impact of COVID-19: Historical Development, Molecular Characterization, Drug Discovery and Future Directions. *Clin Pathol* 16, 2632010X231218075. <https://doi.org/10.1177/2632010X231218075>.
279. Andersson, N.W., Thiesson, E.M., and Hviid, A. (2025). Safety of JN.1-Updated mRNA COVID-19 Vaccines. *JAMA Netw Open* 8, e2523557. <https://doi.org/10.1001/jamanetworkopen.2025.23557>.
280. Procko, E., Berguig, G.Y., Shen, B.W., Song, Y., Frayo, S., Convertine, A.J., Margineantu, D., Booth, G., Correia, B.E., Cheng, Y., et al. (2014). A computationally designed inhibitor of an Epstein-Barr viral Bcl-2 protein induces apoptosis in infected cells. *Cell* 157, 1644–1656. <https://doi.org/10.1016/j.cell.2014.04.034>.
281. Rhoden Smith, A., and Iverson, B.L. (2013). Threading polyintercalators with extremely slow dissociation rates and extended DNA binding sites. *J Am Chem Soc* 135, 12783–12789. <https://doi.org/10.1021/ja4057344>.
282. Hilser, V.J., Dowdy, D., Oas, T.G., and Freire, E. (1998). The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble. *Proc Natl Acad Sci U S A* 95, 9903–9908. <https://doi.org/10.1073/pnas.95.17.9903>.
283. Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J.Z.H., and Hou, T. (2019). End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies

and Applications in Drug Design. *Chem. Rev.* 119, 9478–9508. <https://doi.org/10.1021/acs.chemrev.9b00055>.

284. Trauger, J.W., Baird, E.E., and Dervan, P.B. (1996). Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* 382, 559–561. <https://doi.org/10.1038/382559a0>.

285. Raskatov, J.A., Meier, J.L., Puckett, J.W., Yang, F., Ramakrishnan, P., and Dervan, P.B. (2012). Modulation of NF- κ B-dependent gene transcription using programmable DNA minor groove binders. *Proc Natl Acad Sci U S A* 109, 1023–1028. <https://doi.org/10.1073/pnas.1118506109>.

286. Bhaduri, S., Ranjan, N., and Arya, D.P. (2018). An overview of recent advances in duplex DNA recognition by small molecules. *Beilstein J Org Chem* 14, 1051–1086. <https://doi.org/10.3762/bjoc.14.93>.

287. Blackledge, M.S., and Melander, C. (2013). Programmable DNA-binding small molecules. *Bioorg Med Chem* 21, 6101–6114. <https://doi.org/10.1016/j.bmc.2013.04.023>.

288. Lin, J., and Nagase, H. (2020). The Road Not Taken with Pyrrole-Imidazole Polyamides: Off-Target Effects and Genomic Binding. *Biomolecules* 10, 544. <https://doi.org/10.3390/biom10040544>.

289. Wang, C.C.C., Ellervik, U., and Dervan, P.B. (2001). Expanding the recognition of the minor groove of DNA by incorporation of β -alanine in hairpin polyamides. *Bioorganic & Medicinal Chemistry* 9, 653–657. [https://doi.org/10.1016/S0968-0896\(00\)00282-0](https://doi.org/10.1016/S0968-0896(00)00282-0).

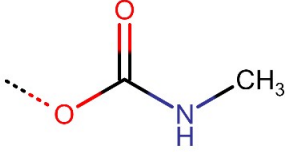
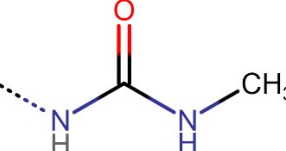
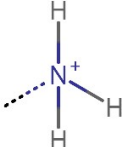
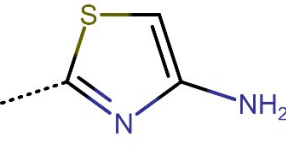
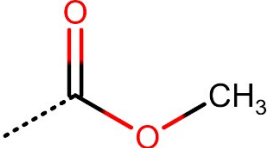
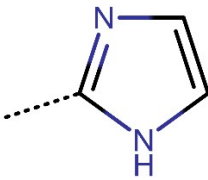
290. Kawamoto, Y., Sasaki, A., Hashiya, K., Ide, S., Bando, T., Maeshima, K., and Sugiyama, H. (2015). Tandem trimer pyrrole–imidazole polyamide probes targeting 18 base pairs in human telomere sequences. *Chem. Sci.* 6, 2307–2312. <https://doi.org/10.1039/C4SC03755C>.

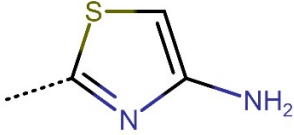
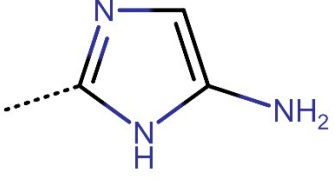
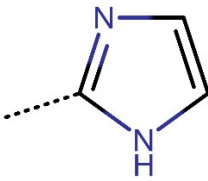
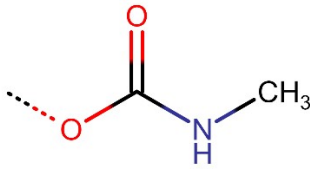
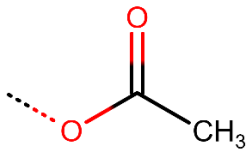
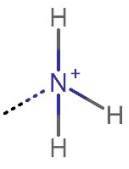
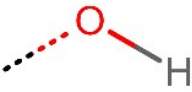
291. Kawamoto, Y., Sasaki, A., Chandran, A., Hashiya, K., Ide, S., Bando, T., Maeshima, K., and Sugiyama, H. (2016). Targeting 24 bp within Telomere Repeat Sequences with Tandem Tetramer Pyrrole-Imidazole Polyamide Probes. *J Am Chem Soc* 138, 14100–14107. <https://doi.org/10.1021/jacs.6b09023>.

292. Siebert, X., and Amzel, L.M. (2004). Loss of translational entropy in molecular associations. *Proteins* 54, 104–115. <https://doi.org/10.1002/prot.10472>.
293. Yan, J., Liu, W., Cai, J., Wang, Y., Li, D., Hua, H., and Cao, H. (2021). Advances in Phenazines over the Past Decade: Review of Their Pharmacological Activities, Mechanisms of Action, Biosynthetic Pathways and Synthetic Strategies. *Mar Drugs* 19, 610. <https://doi.org/10.3390/md19110610>.
294. Uppulapu, S.K., Alam, M.J., Kumar, S., and Banerjee, S.K. (2022). Indazole and its Derivatives in Cardiovascular Diseases: Overview, Current Scenario, and Future Perspectives. *Curr Top Med Chem* 22, 1177–1188. <https://doi.org/10.2174/1568026621666211214151534>.
295. Lam, N.Y.S., Wu, K., and Yu, J.-Q. (2021). Advancing the Logic of Chemical Synthesis: C-H Activation as Strategic and Tactical Disconnections for C-C Bond Construction. *Angew Chem Int Ed Engl* 60, 15767–15790. <https://doi.org/10.1002/anie.202011901>.
296. Zaremba A., and Zahorodnia S. In silico study of ketodarolutamide as a potential inhibitor of the SARS-CoV-2 RBD interaction with human ACE2. The V Scientific Conference “Youth and Modern Problems of Microbiology and Virology”, 19-20 November 2024, Kyiv, Ukraine, P.48.
297. Hata, H., Phuoc Tran, D., Marzouk Sobeh, M., and Kitao, A. (2021). Binding free energy of protein/ligand complexes calculated using dissociation Parallel Cascade Selection Molecular Dynamics and Markov state model. *Biophys Physicobiol* 18, 305–316. <https://doi.org/10.2142/biophysico.bppb-v18.037>.
298. Lee, K.M., Heberer, K., Gao, A., Becker, D.J., Loeb, S., Makarov, D.V., Gulanski, B., DuVall, S.L., Aslan, M., Lee, J., et al. (2022). A Population-Level Analysis of the Protective Effects of Androgen Deprivation Therapy Against COVID-19 Disease Incidence and Severity. *Front Med (Lausanne)* 9, 774773. <https://doi.org/10.3389/fmed.2022.774773>.
299. Gong, T., Zhang, X., Lin, H., Li, J., Tao, J., Zeng, T., Ren, X., Xie, Z., Lei, X.,

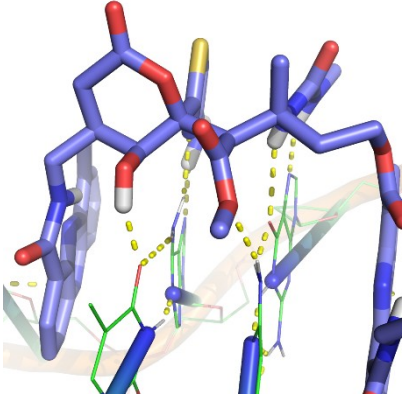
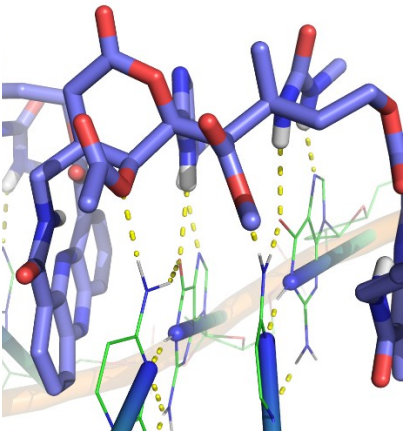
Zhang, S., et al. (2025). Mutation profiling, evolution analysis, molecular dynamics simulation, and functional characterization of Omicron sub-strains. *Virus Res* 360, 199626. <https://doi.org/10.1016/j.virusres.2025.199626>.

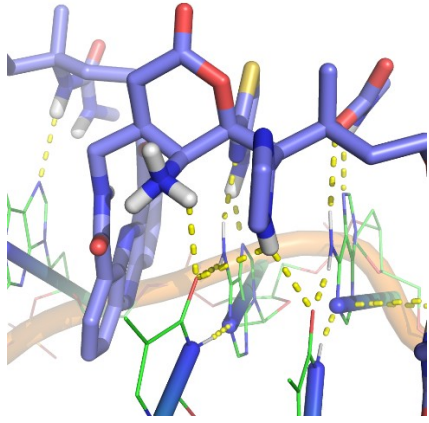
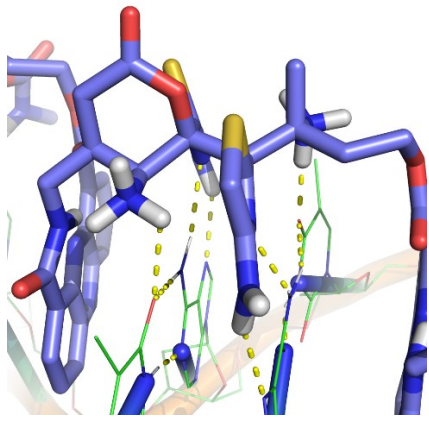
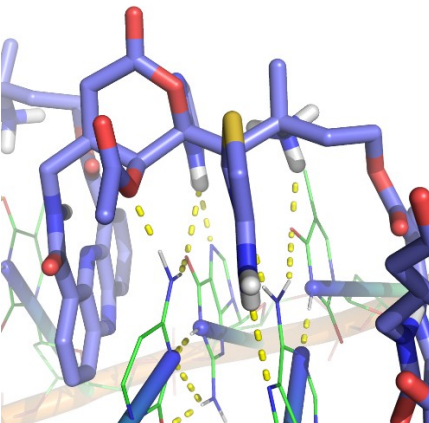
Модифікації лінкеру HASDI направлені на розпізнавання різних нуклеїнових основ
у складі ДНК-дуплексу

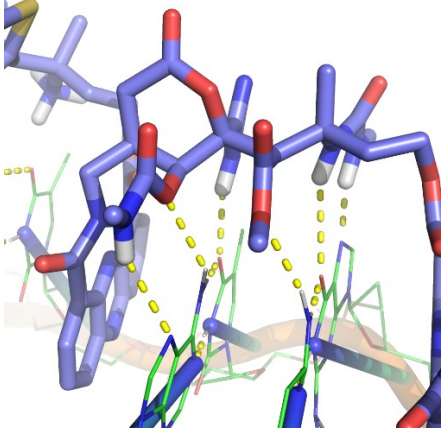
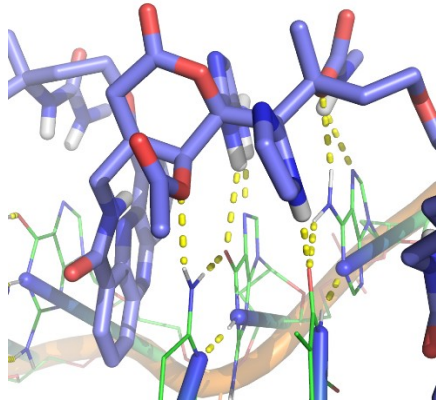
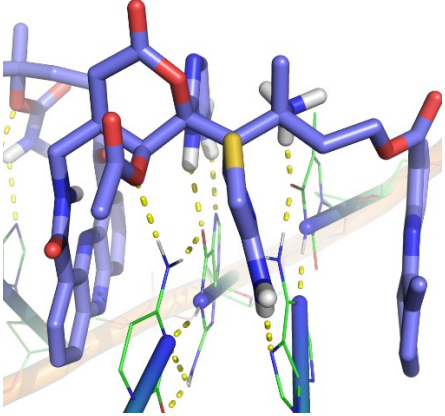
Радикал	Нуклеосонова, що розпізнається	Замісник
R ₁	Adenine (A)	
	Guanine (G)	
	Thymine (T)	
R ₂	Adenine (A)	
	Cytosine (C)	
	Thymine (T)	

R ₃	Adenine (A)	
	Guanine (G)	
	Thymine (T)	
R ₄	Adenine (A)	
	Cytosine (C)	
	Thymine (T)	
		

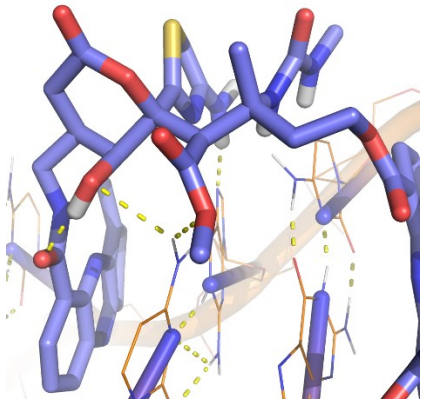
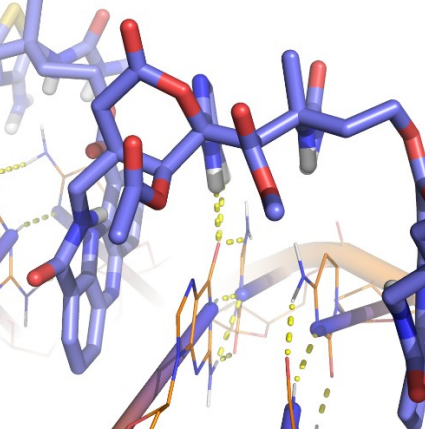
Особливості взаємодії HASDI із послідовністю 5`-TCCCTTTACAACCTCA-3`
(EBNA1-50nt).

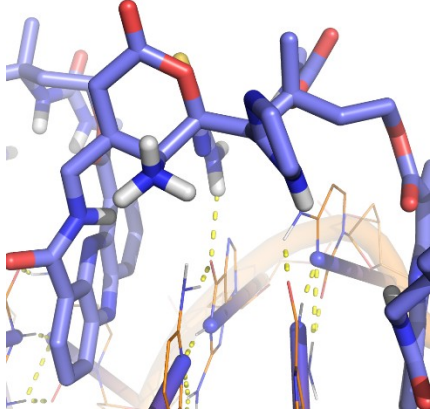
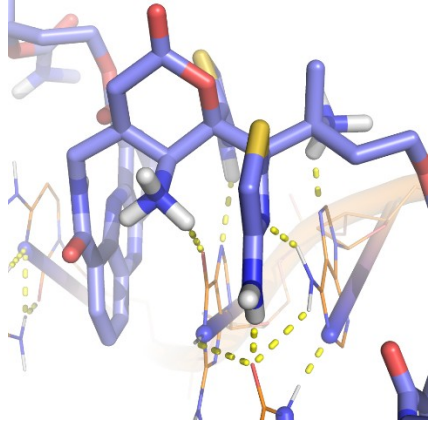
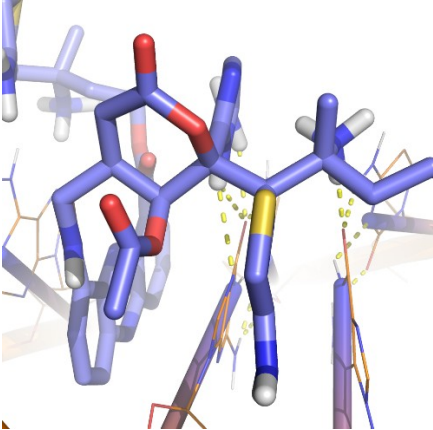
№ (5` to 3`)	Пари нуклеїнових основ	Розпізнавання HASDI фрагменту EBNA1-50nt (5`-TCCCTTTACAACCTCA-3`)	Середня кількість водневих зв'язків
1	T-A C-G		6
2	C-G C-G		6

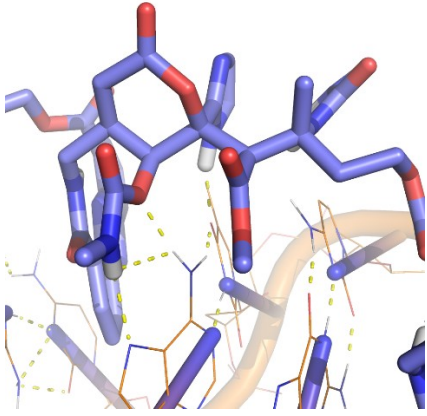
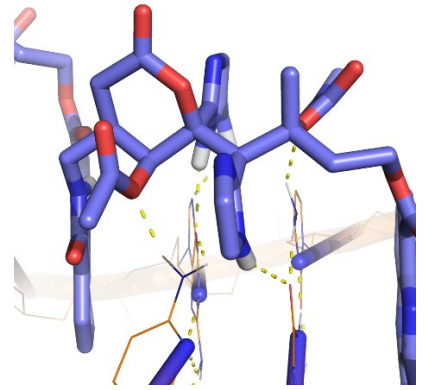
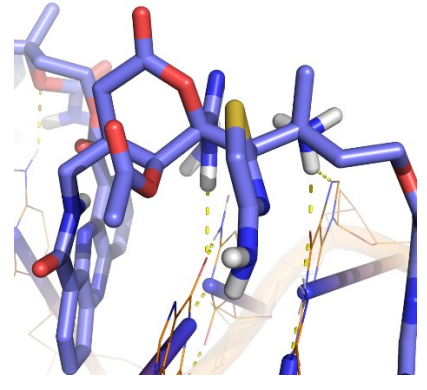
3	T-A T-A		6
4	T-A A-T		6
5	C-G A-T		6

6	A-T C-G		6
7	C-G T-A		6
8	C-G A-T		6

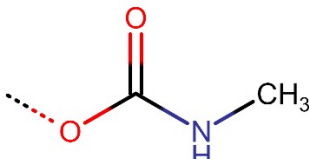
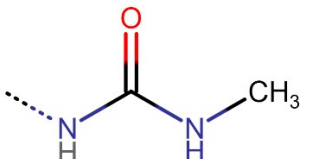
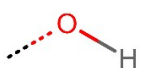
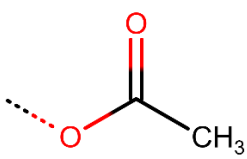
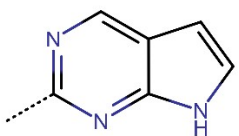
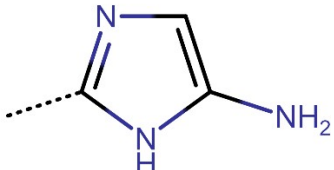
Особливості взаємодії HASDI з послідовністю 5'-CGGGCGCTGGAGCGGC-3'
(KCNH2-50bp).

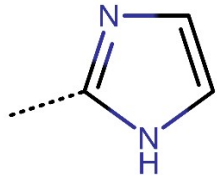
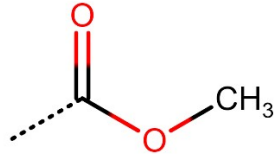
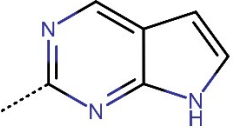
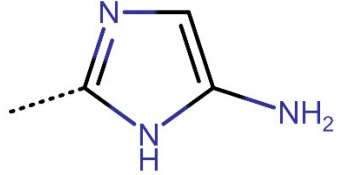
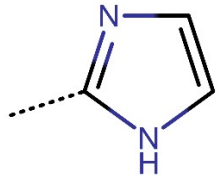
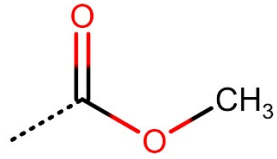
№ (5' to 3')	Пари нуклеїнових основ	Взаємодія HASDI з фрагментом KCN2-50nt (5'-CGGGCGCTGGAGCGGC-3')	Середня кількість водневих зв'язків	Примітки
1	C-G G-C		2	Є наявним внутрішньо молекулярний водневий зв'язок
2	G-C G-C		1	-

3	C-G G-C		1	-
4	C-G T-A		(5)	Cytosine покидає площину ДНК після 96нс моделювання
5	G-C G-C		2	-

6	A-T G-C		3	A-T пара є першим перекриттям між EBNA1-50bp і KCNH2-50bp
7	C-G G-C		4	C-G пара є другим перекриттям між EBNA1-50bp і KCNH2-50bp
8	G-C C-G		2	-

R1, R2, R3 та R4 лінкерів HASDI-G2 та закон їх комбінування для розпізнавання ДНК-послідовності націлювання

Радикал	Нуклеосонова, що розпізнається	Замісник
R ₁	Adenine (A)	
	Guanine (G)	
	Thymine (T)	
	Cytosine (C)	
R ₂	Adenine (A)	
	Guanine (G)	

	Thymine (T)	
	Cytosine (C)	
R ₃	Adenine (A)	
	Guanine (G)	
	Thymine (T)	
	Cytosine (C)	

R ₄	Adenine (A)	
	Guanine (G)	
	Thymine (T)	
	Cytosine (C)	

Наукові статті за основною темою наукової роботи

(* – особистий внесок здобувача)

1. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2025). A thorough insight into the life cycle of the Epstein-Barr virus. From the molecular to the organismal level. *Current Research in Microbial Sciences* 9, 100505. <https://doi.org/10.1016/j.crmicr.2025.100505>.

(*проаналізовано весь спектр наукових робіт щодо особливостей життєвого циклу вірусу Ештейна-Барр депонованих у PubMed, сформовано і письмово викладено повноцінний механістичний погляд на життєвий цикл цього патогену)

2. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2025). In silico development of HASDI-G2 as a novel agent for selective recognition of the DNA sequence. *Sci Rep* 15, 8577. <https://doi.org/10.1038/s41598-025-89967-1>. (*розроблено HASDI-G2, проведено ряд симуляційних експериментів де підтверджено покращений ступінь його селективності щодо ДНК-послідовності)

3. Zaremba, A.A., Zaremba, P.Y., and Zahorodnia, S.D. (2023). In silico study of HASDI (high-affinity selective DNA intercalator) as a new agent capable of highly selective recognition of the DNA sequence. *Sci Rep* 13, 5395. <https://doi.org/10.1038/s41598-023-32595-4>. (*розроблено HASDI, проведено ряд симуляційних експериментів де підтверджено наявність селективності щодо ДНК-послідовності)

4. Zaremba, A., Zaremba, P., and Zahorodnia, S. (2023). De novo designed inhibitor has high affinity to four variants of the RBD of S-glycoprotein of SARS-CoV-2 - an in silico study. *Journal of Biomolecular Structure and Dynamics* 41, 9389–9397. <https://doi.org/10.1080/07391102.2022.2141886>. (*розроблено малу молекулу здатну зв'язуватись з RBD SARS-CoV-2, проведено ряд симуляційних експериментів де підтверджено можливість її зв'язування з чотирма варіантами SARS-CoV-2)

5. Zaremba, A.A., Zaremba, P.Y., and Platonov, M.O. (2023). De novo designed EBAI as a potential inhibitor of the viral protein BHRF1. Research in silico. Journal of Biomolecular Structure and Dynamics 41, 3680–3685. <https://doi.org/10.1080/07391102.2022.2053746>. (*розроблено EBAI, проведено два паралельних симуляційних експерименти де підтверджено здатність до комплексоутворення з BHRF1 вірусу Епштейна-Барр)
6. Zaremba, A., Zaremba, P., Muchnyk, F.V., Baranova, G., and Zahorodnia, S. (2021). In silico Identification of a Viral Surface Glycoprotein Site Suitable for the Development of Low Molecular Weight Inhibitors for Various Variants of the SARS-CoV-2. Mikrobiolohichnyi Zhurnal 84, 34–43. <https://doi.org/10.15407/microbiolj84.01.034>. (*проведено ряд симуляційних експериментів, визначено амінокислотний склад інтерфейсу взаємодії RBD/hACE2 і ідентифіковано кишеню придатну для СОДЛ)
7. Zaremba, A., Zaremba, P., Budzanivska, I., and Zahorodnia, S. (2022). Patterns of the influence of vaccination on the dynamics of different SARS-CoV-2 variants spread. Two-year analysis. Bulletin of Taras Shevchenko National University of Kyiv Series Biology 89, 39–45. <https://doi.org/10.17721/1728.2748.2022.89.39-45>. (*проаналізовано дані з відкритих джерел щодо різних епідемічних і соціальних аспектів поширення SARS-CoV-2 у Індії, Німеччині та Україні, визначено залежність між ступенем вакцинованості населення і часткою варіанту Дельта)

Наукові праці, які засвідчують апробацію матеріалів дисертації

8. Zaremba, A., Zaremba, P., and Zahorodnia, S. In silico identification of a new potential drug-binding pocket on the surface of the receptor-binding domain of the SARS-CoV-2 S-glycoprotein. The 1st International Electronic Conference on Medicinal Chemistry and Pharmaceutics, 1–30 November 2025, Basel, Switzerland, online. (форма участі – очна)

9. Zaremba A., and Zahorodnia S. In silico study of ketodarolutamide as a potential inhibitor of the SARS-CoV-2 RBD interaction with human ACE2. The V Scientific Conference “Youth and Modern Problems of Microbiology and Virology”, 19-20 November 2024, Kyiv, Ukraine, P. 48. (форма участі – очна)

10. Zaremba A., and Zahorodnia S. 3D structure data validation of the SARS-CoV-2 protein E transmembrane domain pentamer form as a potential target for drug development. Modern aspects of microbiology, virology and biotechnology in wartime and post-war period, 15-16 November 2023, Kyiv, Ukraine, P. 277-278. (форма участі – очна)

11. Zaremba, A., Shalimov, O., Onys'ko, P., and Zahorodnia, S. In silico identification of a potential inhibitor of the SARS-CoV-2 S-glycoprotein receptor-binding domain interaction with human ACE2. 9th International Electronic Conference on Medicinal Chemistry, 1–30 November 2023, Basel, Switzerland, online. (форма участі – очна)

12. Zaremba, A., Zaremba, P., and Zagorodnya, S. Selective DNA intercalation of massive molecules as a new method of highly specific inhibition of transcription. 6th International Electronic Conference on Medicinal Chemistry, 1–30 November 2020, Basel, Switzerland, online. (форма участі – очна)

Документ підписано у сервісі Вчасно (продовження)
Дисертація_фін_Заремба.pdf

Документ відправлено: 16:59 26.03.2026
Документ отримано (05417087): 16:59 26.03.2026

Відправник документу

Електронний підпис

16:59 26.03.2026

Ідентифікаційний код: 3604600776

Заремба Андрій Анатолійович

Власник ключа: Заремба Андрій Анатолійович

Час перевірки КЕП/ЕЦП: 16:59 26.03.2026

Статус перевірки сертифікату: Сертифікат діє

Серійний номер: 514B5C86A1E5DA1104000000CE4A160125423205

Тип підпису: кваліфікований

Тип сертифікату: кваліфікований