



Софийски университет „Св. Климент Охридски“
Физически факултет
Катедра „Медицинска физика“

Дипломна работа

за придобиване
на образователно-квалификационна степен „бакалавър“

на
Таня Венелинова Емилова
специалност „Медицинска физика“
Ф.Н. ОРН0500013

**Тема: Взаимодействие на модифицирани циклични пептиди
с неструктурни белтъци на SARS-CoV-2**

Научен ръководител: доц. д-р Пейчо Петков

София, септември 2026 г.

СЪДЪРЖАНИЕ

Въведение	3
1. SARS-CoV-2 – биологична характеристика	5
1.1. Обща характеристика на SARS-CoV-2	5
1.2. Геном и структурна организация	5
1.3. Жизнен цикъл на вируса	8
1.4. Структурни и неструктурни белтъци	10
2. Белтъци на SARS-CoV-2 – значение за вирусната репликация и като терапевтични мишен	11
2.1. Неструктурни белтъци (NSP1–NSP16)	11
2.2. Биологични функции на неструктурните белтъци	12
2.3. Най-изследваните неструктурни белтъци: NSP3, NSP5 (3CLpro), NSP12 (RdRp), NSP13 (Helicase)	12
2.4. Неструктурните белтъци като потенциални терапевтични мишени	16
2.5. Аксесорният белтък ORF6 – структура, функции и значение при SARS-CoV-2	18
3. Модифицирани циклични пептиди	20
3.1. Пептиди	20
3.2. Циклични пептиди	20
3.3. Методи за модифициране на цикличните пептиди	21
3.4. Биологични свойства и терапевтичен потенциал	22
4. Взаимодействие между модифицирани циклични пептиди и белтъците на SARS-CoV-2	23
4.1. Молекулни механизми на взаимодействие	23
4.2. Афинитет и специфичност на свързване	23
4.3. Взаимодействие с NSP13	24
4.4. Взаимодействие с ORF6	24
4.5. Потенциал за инхибиране на вирусната репликация	25
4.6. Съвременни научни изследвания	25
5. Методи за изследване	26
5.1. Молекулен докинг	26
5.2. Молекулна динамика	27
6. Компютърно моделиране на NSP13	32

6.1. Подготовка на структурата на NSP13	32
6.2. Молекулно-динамична симулация на NSP13	33
6.3. Модели на модифицираните циклотиди	36
6.4. Сравнение на двата модифицирани циклотида	41
Заклучение	45
Литература	46

Въведение

Вирусните инфекции са едно от значимите предизвикателства пред съвременната медицина и общественото здраве. Появата на нови вирусни причинители, както и способността им да претърпяват генетични промени, налагат разработването на нови диагностични, профилактични и терапевтични подходи. Въпреки напредъка при създаването на ваксини и антивирусни лекарствени средства, продължава търсенето на по-ефективни и по-специфични терапии [1].

Появата на SARS-CoV-2 и последвалата пандемия от COVID-19 доведоха до засилен интерес към структурата, жизнения цикъл и механизмите на действие на вируса, както и към търсенето на нови средства за неговото лечение [3].

В жизнения цикъл на SARS-CoV-2 участват различни вирусни белтъци. Особен интерес представляват неструктурните белтъци, които имат важна роля за размножаването на вируса и се изследват като възможни мишени за нови противовирусни средства [4].

Наред с изследването на вирусните белтъци се разработват и молекули, които могат да взаимодействат с тях и да повлияят тяхната функция. Сред тях са модифицираните циклични пептиди, които се отличават с по-висока структурна стабилност и устойчивост на ензимно разграждане в сравнение с линейните пептиди. Техните свойства могат да бъдат променяни според целта на конкретното изследване [5].

Съвременните компютърни методи дават възможност да се изследват взаимодействията между пептидите и вирусните белтъци и да се проследява поведението на изследваните молекули. Това ги прави важна част от проучванията, свързани с разработването на нови лекарствени средства [6].

Като част от изследванията, залегнали в научната програма на проекта AVATAR, целта на настоящата дипломна работа е да се разгледат взаимодействията между модифицирани циклични пептиди и белтъци на SARS-CoV-2, като основно внимание е отделено на NSP13 и ORF6. Разгледани са техните структура и функции, свойствата на модифицираните циклични пептиди и методите, използвани за изследване на тези молекули.

В първа глава е направен обзор на основните биологични характеристики, геномната организация и жизнения цикъл на SARS-CoV-2. Втора глава е посветена на основните белтъци на вируса, с по-подробно разглеждане на NSP13 и ORF6 и тяхното значение като възможни терапевтични мишени. В трета глава са представени свойствата и начините за модифициране на цикличните пептиди. В четвърта глава са разгледани взаимодействията между модифицираните циклични пептиди и белтъците на SARS-CoV-2. В пета глава са представени използваните компютърни методи, а в шеста глава е описано компютърното моделиране и молекулно-динамичното изследване на NSP13 и двата модифицирани циклотида.

1. SARS-CoV-2

1.1. Обща характеристика на SARS-CoV-2

В края на 2019 г. в град Ухан, провинция Хубей, Китай, са регистрирани първите случаи на неизвестно до този момент респираторно заболяване. Впоследствие е установено, че причинителят е нов коронавирус, който получава наименованието **Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)**, а причиняваното от него заболяване е наречено **Coronavirus Disease 2019 (COVID-19)**. Поради високата си ефективност на предаване вирусът бързо се разпространява в световен мащаб и през март 2020 г. Световната здравна организация (WHO) официално обявява пандемия. Това води до активни научни изследвания, насочени към изучаване на вируса и към разработването на средства за неговата профилактика и лечение [1,3].

SARS-CoV-2 принадлежи към семейство *Coronaviridae*, подсемейство *Orthocoronavirinae*, род *Betacoronavirus*. Коронавирусите представляват голяма група РНК вируси, които могат да инфектират както хора, така и различни животински видове. При човека част от тях причиняват леки инфекции на горните дихателни пътища, подобни на обикновената настинка, докато други, като SARS-CoV, MERS-CoV и SARS-CoV-2, могат да доведат до тежки респираторни заболявания с висока смъртност при определени групи пациенти [1].

Предаването на SARS-CoV-2 се осъществява главно по въздушно-капков път чрез аерозоли и респираторни капки, отделяни при кашляне, кихане, говорене или дишане. След навлизане в организма вирусът засяга предимно дихателната система, но множество клинични проучвания показват, че инфекцията може да засегне и други органи и системи, включително сърдечно-съдовата, нервната, отделителна и храносмилателна система. Проявите на заболяването варират от липса на симптоми до тежка вирусна пневмония, дихателна недостатъчност и увреждане на няколко органа [1].

Подобно на други РНК вируси, SARS-CoV-2 претърпява генетични промени. По време на пандемията се появяват множество варианти на вируса, някои от които се отличават с по-лесно предаване между хората или със способност частично да избягват имунния отговор. Това налага постоянно проследяване на промените във вируса и редовно обновяване на методите за диагностика, профилактика и лечение [1].

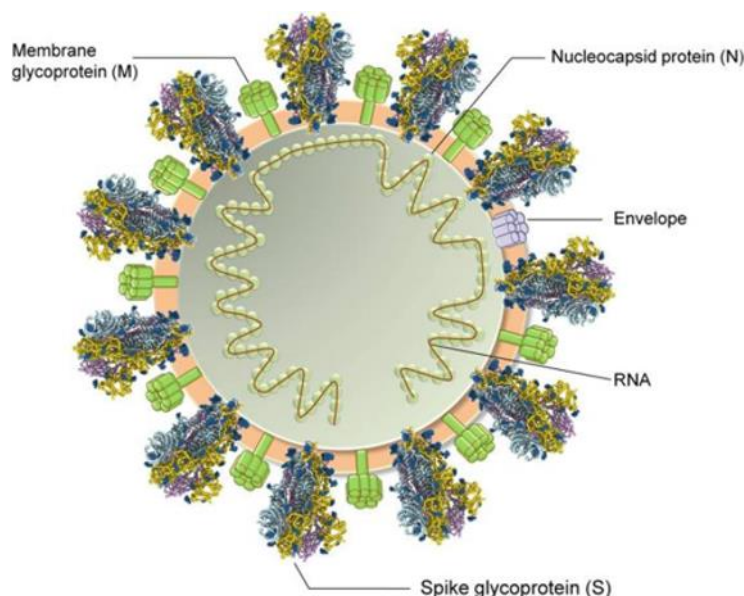
През последните години бяха натрупани значителни знания за SARS-CoV-2, което позволи подробно да бъдат изследвани неговият геном, структурата и начинът, по който вирусът инфектира клетките. Тази информация е важна за разбирането на жизнения цикъл на вируса и за разработването на нови антивирусни лекарствени средства [4].

1.2. Геном и структурна организация

Геномът на SARS-CoV-2 представлява едноверижна РНК с положителна полярност (+ssRNA) и е с дължина приблизително 30 000 нуклеотида. Това е един от най-големите геноми сред РНК вирусите. Вирусната РНК съдържа генетичната информация, необходима

за образуването на вирусните белтъци, които участват в навлизането на вируса в клетката, неговото размножаване и образуването на нови вирусни частици. В двата края на вирусната РНК се намират некодиращи области (5' UTR и 3' UTR), които не участват пряко в образуването на белтъци, но имат важна роля за копирането на вирусната РНК и за регулирането на този процес [4].

SARS-CoV-2 е вирус с липидна обвивка. В нея са разположени структурните белтъци S, М и Е, а във вътрешността на вирусната частица се намира вирусната РНК, свързана с N-белтъка. Структурната организация на вирусната частица е показана на Фиг. 1.2.1 [8].



Фиг. 1.2.1. Схематично представяне на структурата на вирусната частица SARS-CoV-2. Показани са основните структурни белтъци – Spike (S), Membrane (M), Envelope (E) и Nucleocapsid (N), както и вирусната РНК.

Източник: адаптирано по Kumar S., Nyodu R., Maurya V. K., Saxena S. K. *Morphology, Genome Organization, Replication, and Pathogenesis of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)*. In: *Coronavirus Disease 2019 (COVID-19): Epidemiology, Pathogenesis, Diagnosis, and Therapeutics*. Springer; 2020. [8]

Геномът на SARS-CoV-2 има добре организирана структура. В началото на вирусната РНК (5'-края) се намират две отворени рамки за четене (Open Reading Frames, ORF) – участъци, които съдържат информация за образуването на белтъци. Те са означени като ORF1a и ORF1b и заемат приблизително две трети от целия геном. От тях се синтезират полипротеините pp1a и pp1ab. В края на вирусната РНК (3'-края) са разположени гените за структурните белтъци S, Е, М и N, както и за аксесорните белтъци ORF3a, ORF6, ORF7a, ORF7b, ORF8 и ORF9b. Геномната организация на SARS-CoV-2 е показана на Фиг. 1.2.2. [4].



Фиг. 1.2.2. Геномна организация на SARS-CoV-2. Представени са участъците ORF1a и ORF1b, от които се синтезират полипротеините pp1a и pp1ab. След разделянето им се образуват неструктурните белтъци NSP1–NSP16. Показани са също гените за структурните и аксесорните белтъци.

Източник: адаптирано по Brant A. C., Tian W., Majerciak V., Yang W., Zheng Z.-M. SARS-CoV-2: from its discovery to genome structure, transcription, and replication. Cell & Bioscience. 2021 [9].

От полипротеините pp1a и pp1ab се образуват шестнадесетте неструктурни белтъка NSP1–NSP16, които изпълняват различни функции в жизнения цикъл на вируса [4].

Четири основни структурни белтъка изпълняват различни функции. S-белтъкът (Spike) участва в свързването на вируса с рецептора ACE2, разположен върху повърхността на клетките. М-белтъкът има основна роля за оформянето и изграждането на вирусната частица. Е-белтъкът участва в образуването и освобождаването на новите вирусни частици, а N-белтъкът се свързва с вирусната РНК и подпомага нейното подреждане във вътрешността на вирусната частица [7].

Освен структурните и неструктурните белтъци, вирусният геном съдържа информация и за аксесорните белтъци. Те не участват пряко в изграждането на вирусната частица, но влияят върху протичането на инфекцията и върху начина, по който вирусът взаимодейства със заразената клетка. Сред тях особен интерес представлява ORF6, който потиска защитните механизми на клетката срещу вирусната инфекция. Поради значението му за настоящата дипломна работа ORF6 ще бъде разгледан по-подробно в следваща глава [10].

Познаването на геномната и структурната организация на SARS-CoV-2 е важно за разбирането на неговия жизнен цикъл и за установяването на белтъците, към които могат да бъдат насочени нови антивирусни лекарствени средства. Особено важни са неструктурните белтъци и някои аксесорни белтъци, като ORF6, тъй като участват във вирусното размножаване и отслабват защитните механизми на клетката [4].

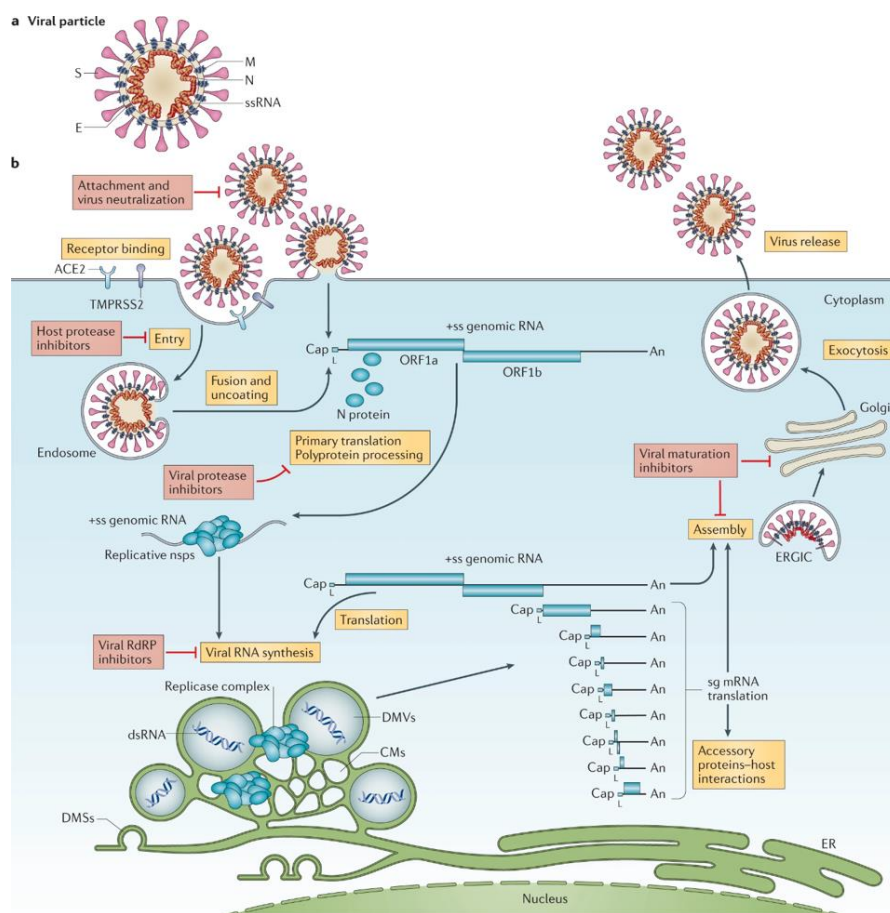
1.3. Жизнен цикъл на вируса

Жизненият цикъл на SARS-CoV-2 представлява последователност от процеси, чрез които вирусът навлиза в клетката, размножава се и образува нови вирусни частици. Тъй като не може да се размножава самостоятелно, той използва механизмите на клетката гостоприемник, за да копира своята РНК и да произвежда необходимите вирусни белтъци [4].

Инфекцията започва, когато S-белтъкът (Spike) се свърже с рецептора ангиотензин конвертиращ ензим 2 (ACE2), разположен върху повърхността на клетката. След това клетъчният ензим TMPRSS2 (Transmembrane Serine Protease 2) активира S-белтъка и улеснява навлизането на вируса в клетката. В зависимост от вида на клетката вирусът може да навлезе чрез сливане на вирусната обвивка с клетъчната мембрана или чрез ендоцитоза – процес, при който клетката обгръща вирусната частица и я въвежда във вътрешността си [7].

След навлизането в клетката вирусната РНК се освобождава в цитоплазмата. Там клетъчните рибозоми я използват за синтеза на първите вирусни белтъци. Част от тези белтъци участват в копирането на вирусната РНК. Този процес протича в специални участъци на клетката, наречени двумембранни везикули (double-membrane vesicles, DMVs), които осигуряват подходящи условия за копирането на вирусната РНК [4].

Успоредно с копирането на вирусния геном се образуват и по-къси РНК молекули, които съдържат информация за синтеза на структурните и аксесорните белтъци. По този начин в заразената клетка се създават нови копия на вирусната РНК и белтъците, необходими за изграждането на новите вирусни частици [9].



Фиг. 1.3.1. *Жизнен цикъл на SARS-CoV-2. Представени са основните етапи – свързване с рецептора ACE2, навлизане в клетката, освобождаване на вирусната РНК, синтез на вирусните белтъци, копиране на вирусния геном, сглобяване и освобождаване на новите вирусни частици.*

Източник: V'kovski P., Kratzel A., Steiner S., Stalder H., Thiel V. *Coronavirus biology and replication: implications for SARS-CoV-2*. Nature Reviews Microbiology. 2021;19(3):155–170

След като се образуват нови копия на вирусната РНК и необходимите белтъци, започва изграждането на новите вирусни частици. Този процес протича в областта между ендоплазмения ретикулум и апарата на Голджи (Endoplasmic Reticulum–Golgi Intermediate Compartment или ERGIC). Там вирусната РНК, свързана с N-белтъка, се събира със структурните белтъци S, М и Е и постепенно се оформят новите вирусни частици [4].

След сглобяването си новите вирусни частици се пренасят в малки мембранни мехурчета до повърхността на клетката и се освобождават извън нея чрез екзоцитоза (процес, при който клетката отделя вещества извън себе си). След освобождаването си те могат да инфектират съседни клетки и да започнат нов жизнен цикъл, което води до разпространяване на инфекцията в организма [4].

Разбирането на жизнения цикъл на SARS-CoV-2 е важно, тъй като показва в кои етапи може да бъде ограничено размножаването на вируса. Поради тази причина вирусните белтъци, които участват в различните етапи на жизнения цикъл, се изследват с цел разработване на нови антивирусни лекарствени средства.

1.4. Структурни и неструктурни белтъци

Белтъците на SARS-CoV-2 имат важна роля в жизнения цикъл на вируса. Според своите функции те се разделят на структурни, неструктурни и аксесорни белтъци [7].

Структурните белтъци S, М, Е и N изграждат вирусната частица и участват в нейното навлизане в клетката, оформянето ѝ и свързването с вирусната РНК [7]. Неструктурните белтъци NSP1–NSP16 не участват пряко в изграждането на вирусната частица, но изпълняват важни функции при копирането и обработването на вирусната РНК [9].

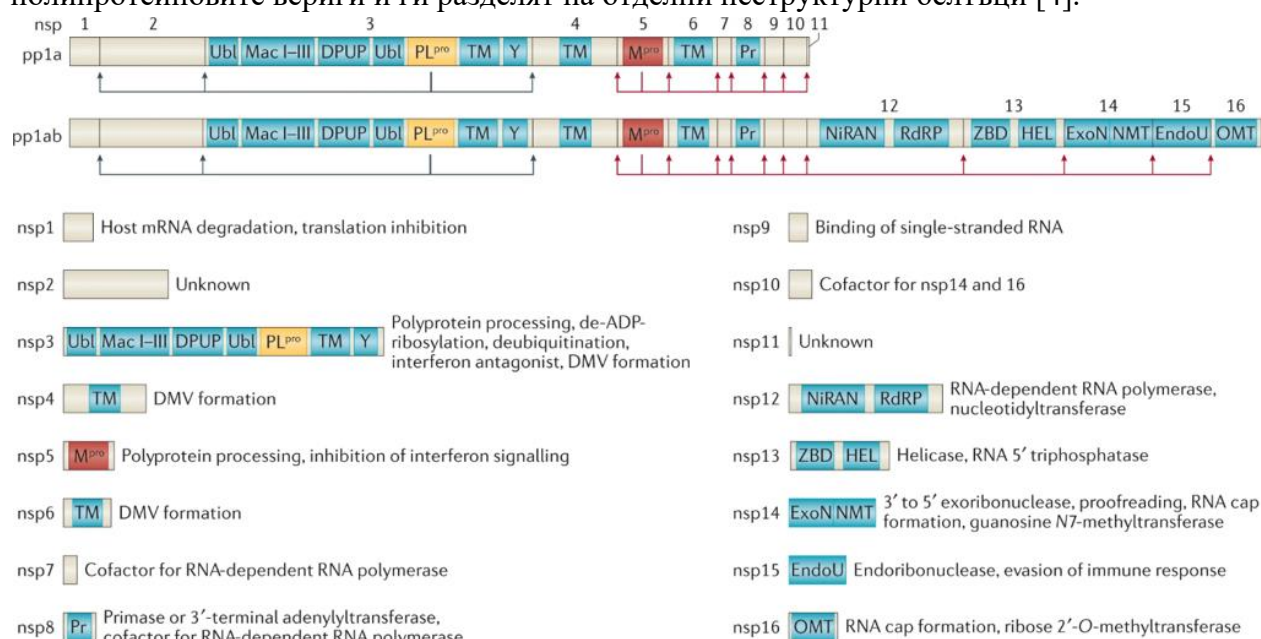
Аксесорните белтъци участват във взаимодействието на вируса със заразената клетка и могат да повлияят нейния защитен отговор. Към тях принадлежи ORF6, чиято структура и функции са разгледани по-подробно в т. 2.5 [10].

2. Белтъци на SARS-CoV-2 – значение за вирусната репликация и като терапевтични мишени

2.1. Неструктурни белтъци (NSP1–NSP16)

Неструктурните белтъци се образуват от полипротеините pp1a и pp1ab. Разделянето им на по-малки белтъци се извършва с помощта на вирусни ензими, наречени протеази. По този начин се образуват шестнадесетте неструктурни белтъка NSP1–NSP16 [4].

Основна роля при разделянето на полипротеините имат две вирусни протеази. Едната е част от NSP3 и се нарича папаиноподобна протеаза (PL_{pro}), а другата е NSP5, известна като главна протеаза (3CL_{pro} или M_{pro}). Те разпознават определени участъци от полипротеиновите вериги и ги разделят на отделни неструктурни белтъци [4].



Фиг. 2.1.1. Неструктурни белтъци на SARS-CoV-2. Представени са полипротеините pp1a и pp1ab, тяхното разделяне от вирусните протеази PL_{pro} и M_{pro} (3CL_{pro}) на неструктурните белтъци NSP1–NSP16, както и основните функции на отделните белтъци.

Източник: V'kovski P., Kratzel A., Steiner S., Stalder H., Thiel V. *Coronavirus biology and replication: implications for SARS-CoV-2*. Nature Reviews Microbiology. 2021;19(3):155–170.

Получените NSP1–NSP16 изпълняват различни функции, свързани с обработването на вирусните полипротеини, копирането на вирусната РНК и взаимодействието със защитните механизми на клетката. Част от тях работят съвместно в репликационно-транскрипционния комплекс. Основните им биологични функции са представени по-подробно в следващите точки [4].

Неструктурните белтъци не действат поотделно, а взаимодействат и изпълняват функциите си съвместно. Повечето от тях участват в образуването на репликационно-транскрипционния комплекс, чрез който се копира вирусната РНК и се образуват РНК молекулите, необходими за синтеза на структурните и аксесорните белтъци [4].

2.2. Биологични функции на неструктурните белтъци

Неструктурните белтъци изпълняват различни биологични функции, които са необходими за размножаването на SARS-CoV-2. Част от тях участват пряко в копирането и обработката на вирусната РНК, други обработват вирусните полипротеини, а трети подпомагат вируса да преодолее защитните механизми на заразената клетка [4].

За разлика от много други РНК вируси, SARS-CoV-2 може да коригира част от грешките, възникнали при копирането на вирусната РНК. Това е особено важно, тъй като коронавирусите имат сравнително голям геном. Така се ограничава натрупването на грешки и се запазва генетичната информация на вируса [4].

Неструктурните белтъци участват и в образуването на двумембранни везикули (double-membrane vesicles, DMVs), които осигуряват защитена среда за протичането на вирусната репликация. В тях се ограничава разпознаването на вирусната РНК от защитните механизми на клетката [4].

2.3. Най-изследваните неструктурни белтъци: NSP3, NSP5 (3CLpro), NSP12 (RdRp) и NSP13 (Helicase)

Сред неструктурните белтъци на SARS-CoV-2 най-подробно изследвани са NSP3, NSP5, NSP12 и NSP13. Всеки от тях изпълнява различна функция, но всички участват в размножаването на вируса. NSP3 и NSP5 участват в обработката на вирусните полипротеини, а NSP12 и NSP13 – в копирането на вирусната РНК [7].

NSP3

NSP3 е най-големият неструктурен белтък на SARS-CoV-2. Той съдържа няколко отделни области, наречени домени, като всяка от тях изпълнява различна функция. Някои от тези домени се свързват с вирусната РНК или с други вирусни белтъци, а други участват в промените в клетъчните мембрани и в образуването на двумембранните мехурчета, в които протича вирусната репликация [4].

Една от най-важните части на NSP3 е папаиноподобната протеаза (PLpro). Тя участва в разделянето на началната част на полипротеините pp1a и pp1ab, при което се отделят NSP1, NSP2 и NSP3. Това е необходима стъпка за образуването на останалите неструктурни белтъци [7].

Освен че участва в обработката на вирусните полипротеини, PLpro отслабва и защитните механизми на клетката. Това улеснява размножаването на вируса в заразената клетка [7].

Друг важен домен на NSP3 е макродоменът Mac1. Смята се, че той също участва в отслабването на защитните механизми на заразената клетка и по този начин подпомага размножаването на вируса [11].

NSP3 взаимодейства с NSP4 и NSP6 при образуването на двумембранните мехурчета. Тези белтъци работят съвместно и подпомагат протичането на вирусната репликация. Така NSP3 участва както в обработката на вирусните полипротеини, така и в създаването на подходяща среда за копирането на вирусната РНК [4].

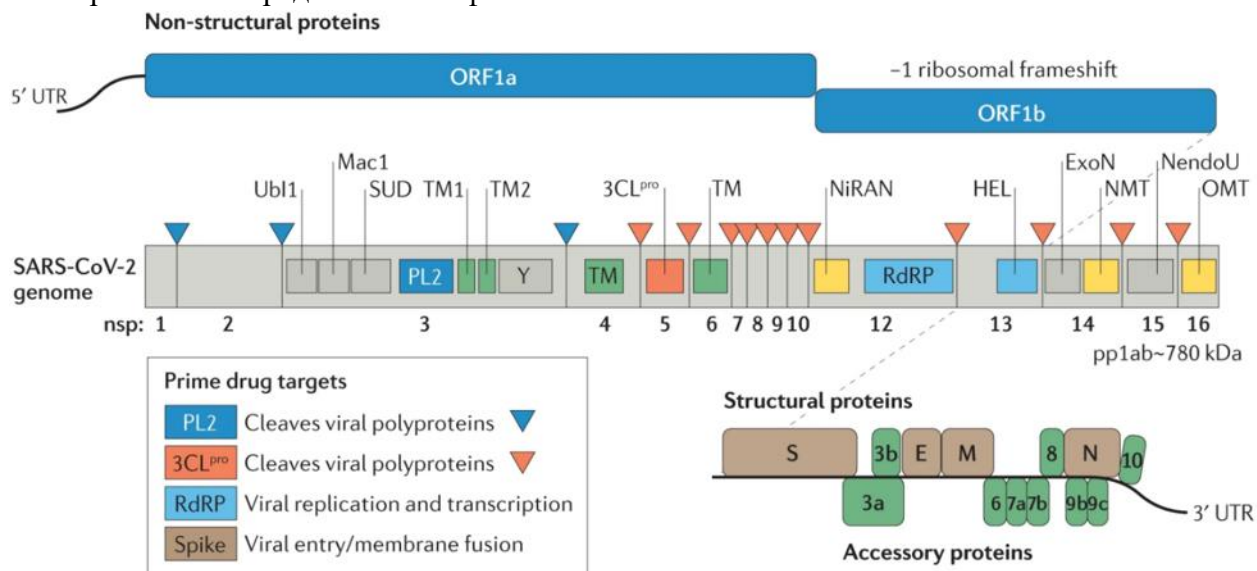
NSP5 – главна протеаза (3CLpro или Mpro)

NSP5 е известен като главна протеаза, 3CLpro или Mpro. Основната му функция е да разделя полипротеините pp1a и pp1ab, при което се образуват повечето неструктурни белтъци – от NSP4 до NSP16. Без действието на NSP5 тези белтъци не могат да изпълняват своите функции и вирусната репликация се нарушава [7].

За да бъде активен, NSP5 се свързва с още една молекула от същия белтък и образува структура, наречена димер. В централната част на белтъка се намира активният център, където протича разделянето на полипротеините [7].

Структурата на NSP5 е подробно проучена, тъй като този белтък има ключово значение за жизнения цикъл на вируса. Дори малки промени в активния му център могат да повлияят върху способността му да обработва полипротеините и съответно върху образуването на останалите неструктурни белтъци [11].

Разположението на неструктурните белтъци и местата на разделяне на вирусните полипротеини са представени на фиг. 2.3.1.



Фиг. 2.3.1. Организация на генома на SARS-CoV-2 и разположение на неструктурните белтъци NSP1–NSP16. Отбелязани са местата, на които вирусните полипротеини се разделят под действието на PLpro (NSP3) и 3CLpro (NSP5), както и някои от основните терапевтични мишени.

Източник: Адаптирано по Yang H., Rao Z. *Structural biology of SARS-CoV-2 and implications for therapeutic development. Nature Reviews Microbiology*. 2021 [7].

NSP12 – РНК-зависима РНК полимераза (RdRp)

NSP12 е основният ензим, който копира вирусната РНК. Той е известен като РНК-зависима РНК полимераза (RdRp). Това е ензим, който използва вирусната РНК като образец, за да синтезира нова РНК верига [7].

NSP12 е основна част от репликационно-транскрипционния комплекс, но не работи достатъчно ефективно самостоятелно. Неговото действие се подпомага от NSP7 и NSP8, които се свързват с него и стабилизируют комплекса по време на копирането на вирусната РНК [4].

Структурата на NSP12 наподобява дясна ръка и включва три основни части – „пръсти“, „длан“ и „палец“. В областта на „дланта“ се намира активният център, където към нарастващата РНК верига последователно се добавят нови нуклеотиди. Останалите части на белтъка подпомагат правилното протичане на синтеза на новата РНК верига [7].

При копирането NSP12 добавя новите нуклеотиди според последователността на РНК веригата, която служи като образец. Така се създават както нови копия на целия вирусен геном, така и по-къси РНК молекули, необходими за синтеза на вирусните белтъци [4].

NSP12 работи съвместно с останалите белтъци от репликационно-транскрипционния комплекс. Ако полимеразата не функционира правилно, вирусът не може да копира своята РНК, което нарушава неговото размножаване [7].

NSP13 – вирусна хеликаза

NSP13 е вирусна хеликаза. Това е ензим, който разделя свързаните или нагънатите участъци на нуклеиновите киселини. При копирането на вирусната РНК могат да се образуват участъци, които затрудняват този процес. NSP13 ги разделя и така позволява копирането на вирусната РНК да продължи [7].

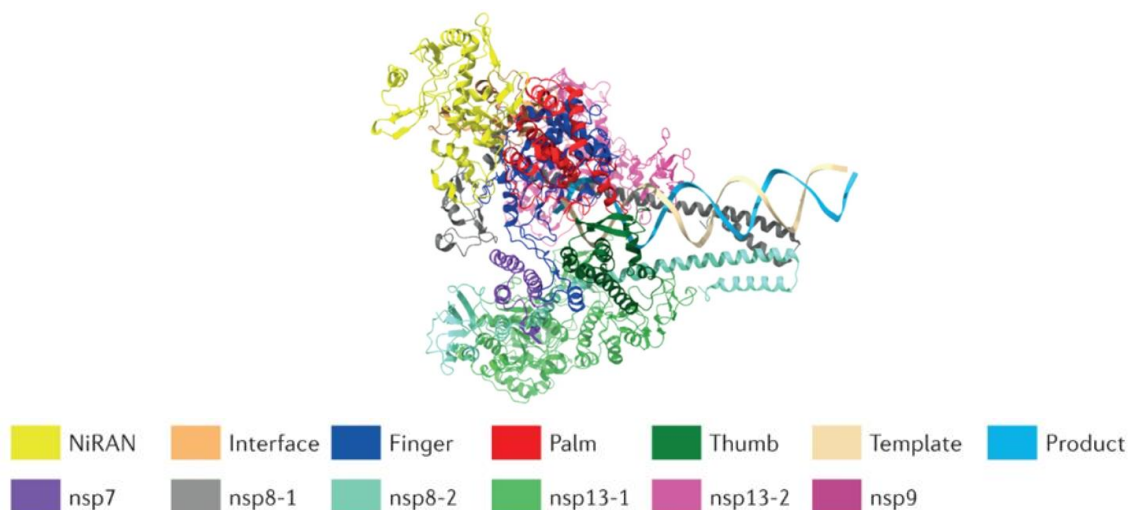
За да извършва тази функция, NSP13 използва енергия, получена от разграждането на аденозинтрифосфат (АТФ). Аденозинтрифосфатът е основният източник на енергия в клетките. Белтъкът се придвижва по РНК веригата и постепенно разделя свързаните участъци. Така NSP13 подпомага работата на NSP12 и на останалите белтъци от репликационно-транскрипционния комплекс [7].

NSP13 се свързва с NSP12, NSP7 и NSP8 и участва в нормалното протичане на вирусната репликация. Освен това има значение и за оформянето на 5'-края на вирусната РНК и подпомага процесите, свързани с коригирането на грешки, които могат да възникнат по време на нейното копиране [7].

NSP13 може да влияе и върху противовирусния отговор на заразената клетка. Установено е, че този белтък потиска част от сигналните процеси, които водят до

образуването на интерферони. По този начин той не само подпомага копирането на вирусната РНК, но и участва в отслабването на клетъчната защита [10].

Взаимодействието между NSP12, NSP7, NSP8 и NSP13 в състава на репликационно-транскрипционния комплекс е представено на фиг. 2.3.2.



Фиг. 2.3.2. Структурен модел на репликационно-транскрипционния комплекс на SARS-CoV-2, включващ РНК-зависимата РНК полимераза NSP12, кофакторите NSP7 и NSP8, хеликазата NSP13 и вирусната РНК.

Източник: Адаптирано по Yang H., Rao Z. *Structural biology of SARS-CoV-2 and implications for therapeutic development. Nature Reviews Microbiology. 2021* [7].

NSP3, NSP5, NSP12 и NSP13 изпълняват различни, но взаимно свързани функции. NSP3 и NSP5 участват в обработването на вирусните полипротеини и образуването на отделните неструктурни белтъци, а NSP12 и NSP13 са необходими за копирането на вирусната РНК. Нарушаването на функцията на който и да е от тези белтъци затруднява размножаването на SARS-CoV-2. Затова те са важни при разработването на нови антивирусни лекарства [7].

2.4. Неструктурните белтъци като потенциални терапевтични мишени

Неструктурните белтъци на SARS-CoV-2 участват в основни процеси от жизнения цикъл на вируса. Те са необходими за обработването на вирусните полипротеини, копирането на вирусната РНК и отслабването на защитните механизми на заразената клетка. Затова блокирането на тяхната функция може да забави или да прекрати размножаването на вируса. Поради това някои от неструктурните белтъци са обект на активни изследвания при разработването на нови антивирусни лекарства [7].

Особен интерес представляват вирусните протеази NSP3 и NSP5. Тези белтъци разделят полипротеините pp1a и pp1ab, като по този начин се образуват отделните неструктурни белтъци. Потискането на тяхната активност нарушава правилното обработване на полипротеините и затруднява образуването на белтъците, необходими за вирусната репликация [7].

Папаиноподобната протеаза (PLpro), която е част от NSP3, представлява интерес при разработването на нови антивирусни вещества. Освен че участва в разделянето на вирусните полипротеини, тя отслабва и защитните механизми на заразената клетка. Затова вещества, които се свързват с активния ѝ център, могат едновременно да нарушат обработването на вирусните полипротеини и да ограничат отслабването на клетъчната защита. Разработването на такива вещества обаче е затруднено, тъй като те трябва да действат избирателно върху вирусната протеаза, без да засягат сходни процеси в човешките клетки [11].

Главната протеаза NSP5 е един от най-добре проучените неструктурни белтъци на SARS-CoV-2. Нейният активен център е добре проучен и към него могат да се свързват вещества, които блокират ензима. Когато активността на NSP5 бъде блокирана, вирусните полипротеини не се разделят правилно и не могат да се образуват повечето неструктурни белтъци. Това прекъсва ранните етапи на вирусната репликация [7].

Пример за лекарство, насочено към главната протеаза, е нирматрелвир (nirmatrelvir). Той се свързва с активния център на NSP5 и потиска разделянето на вирусните полипротеини. Нирматрелвир се прилага заедно с ритонавир (ritonavir), който забавя разграждането му в организма и така поддържа по-висока концентрация на активното вещество. Клинично проучване показва, че тази комбинация може да намали риска от тежко протичане на заболяването при пациенти с повишен риск, когато лечението започне в ранните етапи на инфекцията [12].

Друг важен белтък, към който могат да бъдат насочени антивирусни средства, е NSP12. Тъй като той участва в копирането на вирусната РНК, потискането на неговата активност пречи на образуването на нови вирусни геноми. Някои от веществата, насочени към NSP12, наподобяват естествените нуклеотиди. Те могат да бъдат включени в новообразуваната РНК верига и да нарушат нейното по-нататъшно удължаване [7].

Такъв пример е ремдесивир (remdesivir). След като попадне в клетката, той се превръща в активна форма, която наподобява естествен нуклеотид. NSP12 я включва в синтезираната РНК верига, което затруднява по-нататъшното ѝ копиране. Ремдесивир е сред първите антивирусни средства, използвани за лечение на COVID-19. Клинично проучване показва, че при определени групи хоспитализирани пациенти той съкращава времето за възстановяване [13].

NSP13 е вирусна хеликаза, която също се разглежда като възможна мишена за антивирусни вещества. Нейната хеликазна активност е необходима за разделянето на свързаните участъци на вирусната РНК и за нормалното функциониране на репликационно-транскрипционния комплекс. Блокирането на свързването или хидролизата на АТР може да

ограничи придвижването на NSP13 по РНК веригата и да наруши копирането на вирусния геном. Към момента веществата, насочени към NSP13, все още са предимно в процес на изследване [7].

Неструктурните белтъци имат важна роля в разработването на нови антивирусни средства. Те изпълняват функции, които са пряко свързани с размножаването на вируса, а техните активни центрове позволяват свързването на различни инхибитори (вещества, които потискат активността им). Освен това някои от тези белтъци се променят по-бавно в сравнение с повърхностния S белтък, което може да запази ефективността на лекарствата и при появата на нови вирусни варианти [7].

Разработването на антивирусни средства обаче е свързано и с някои затруднения. Антивирусното вещество трябва да се свързва достатъчно силно с избрания вирусен белтък, да достига до заразените клетки и да не оказва значително влияние върху нормалните процеси в човешките клетки. Възможна е и появата на промени във вирусните белтъци, които да намалят чувствителността им към дадено лекарство. Поради това се изследват не само отделни вещества, но и комбинации от лекарства, които действат върху различни етапи от размножаването на вируса [7].

Неструктурните белтъци NSP3, NSP5, NSP12 и NSP13 са едни от най-важните мишени при разработването на нови антивирусни средства. При NSP3 и NSP5 лечението е насочено към потискане на тяхната протеазна активност, а при NSP12 и NSP13 – към нарушаване на копирането и обработването на вирусната РНК. По-доброто познаване на тяхната структура и функция дава възможност за разработването на нови антивирусни вещества, включително малки молекули и пептиди, които могат да се свързват изборително с определени участъци на вирусните белтъци [7].

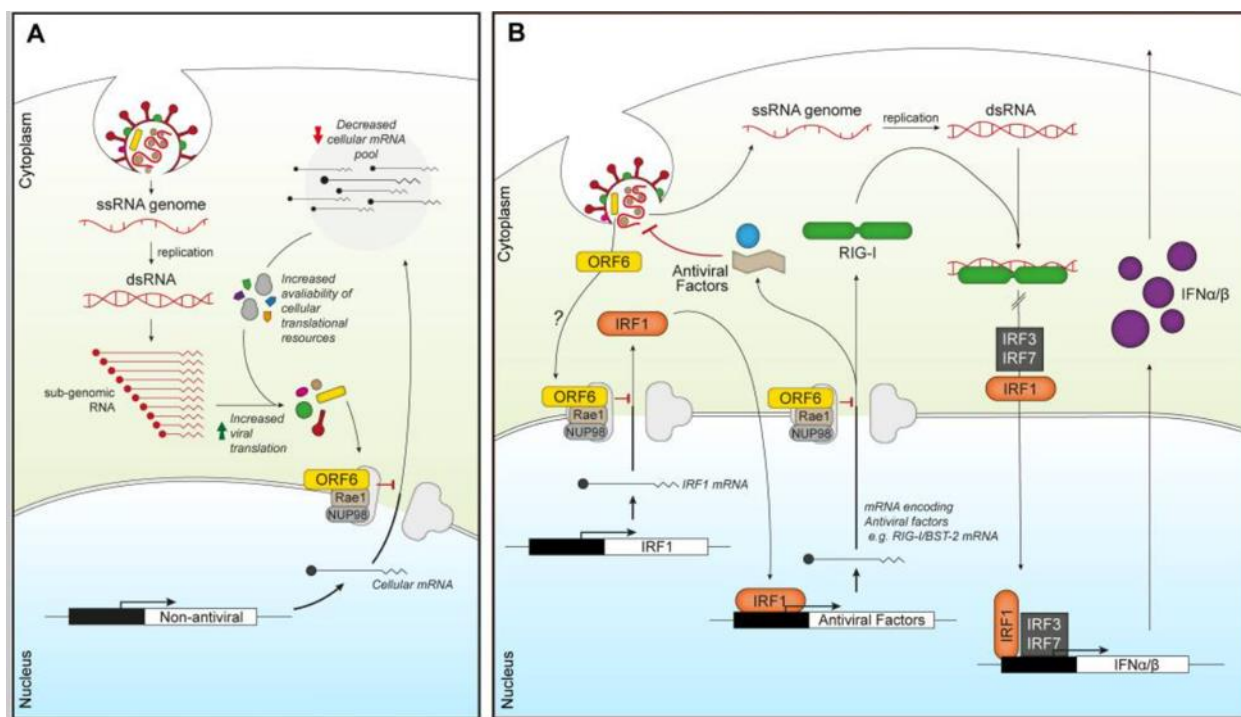
2.5. Аксесорният белтък ORF6 – структура, функции и значение при SARS-CoV-2

Освен структурните и неструктурните белтъци SARS-CoV-2 кодира и няколко аксесорни белтъка, които подпомагат взаимодействието на вируса със заразената клетка. Един от тях е ORF6 – малък белтък, който има важна роля за потискането на противовирусния клетъчен отговор. Той не участва пряко в изграждането на вирусната частица, но помага на вируса да ограничи защитните реакции на клетката и да създаде по-благоприятни условия за своето размножаване [10].

ORF6 е малък белтък, изграден от 61 аминокиселини. В структурата му се различават N-крайна и C-крайна част, които имат различни свойства и функции. N-крайната част е свързана с клетъчните мембрани и съдържа участък с форма на α -спирала. Благодарение на него ORF6 се разполага главно в ендоплазмения ретикулум и апарата на Голджи – клетъчни структури, които участват в синтеза, обработването и транспорта на белтъци. C-крайната част е по-полярна и участва във взаимодействията с клетъчни белтъци, които осъществяват транспорта на молекули между клетъчното ядро и цитоплазмата [14].

ORF6 потиска противовирусния клетъчен отговор чрез взаимодействие с белтъците NUP98 и RAE1, които са част от системата за транспорт през ядрените пори. Ядрените пори представляват белтъчни комплекси в ядрената обвивка, през които се осъществява преносът на молекули между клетъчното ядро и цитоплазмата. Свързването на ORF6 с комплекса RAE1–NUP98 затруднява навлизането в ядрото на клетъчни белтъци, необходими за активирането на противовирусните гени. В резултат защитната реакция на клетката срещу вирусната инфекция се забавя и отслабва, което подпомага размножаването на SARS-CoV-2 [14].

Освен това ORF6 възпрепятства извеждането на клетъчни информационни РНК (mRNA) от ядрото към цитоплазмата. Информационната РНК пренася генетичната информация, необходима за синтеза на белтъци. Когато нейното извеждане от ядрото бъде блокирано, се ограничава образуването на клетъчни белтъци, включително на такива, които участват в противовирусната защита. Вирусните РНК обаче се образуват в цитоплазмата и не зависят от транспорта през ядрените пори. По този начин SARS-CoV-2 може да продължи да използва клетъчния апарат за синтеза на собствените си белтъци [15].



Фиг. 2.5.1. Основни механизми, чрез които аксесорният белтък ORF6 потиска противовирусния клетъчен отговор. (A) ORF6 нарушава транспорта на клетъчните информационни РНК (mRNA) чрез взаимодействие с комплекса RAE1–NUP98. (B) ORF6 потиска интерфероновата сигнализация чрез взаимодействие с комплекса RAE1–NUP98 и нарушаване на транспорта през ядрените пори.

Източник: Hall R., Guedán A., Yap M. W., Young G. R., Harvey R., Stoye J. P., Bishop K. N. SARS-CoV-2 ORF6 disrupts innate immune signalling by inhibiting cellular mRNA export. *PLoS Pathogens*. 2022;18(8):e1010349.

Изследванията показват, че взаимодействието с RAE1 и NUP98 се осъществява чрез С-крайната област на ORF6. По този начин ORF6 нарушава транспорта и в двете посоки през ядрените пори – както навлизането на клетъчни белтъци в ядрото, така и извеждането на информационни РНК към цитоплазмата. Тази функция на С-крайната област е подкрепена от изследвания, според които промени в нея отслабват взаимодействието на ORF6 с комплекса RAE1–NUP98 [10].

Въздействието на ORF6 не се ограничава само до интерфероновата сигнализация. Нарушаването на транспорта между ядрото и цитоплазмата може да засегне движението на различни клетъчни белтъци и РНК молекули, необходими за нормалното функциониране на клетката. По този начин ORF6 допринася за отслабването на клетъчната защита и се разглежда като един от аксесорните белтъци със съществено значение за патогенността на SARS-CoV-2 [14].

Макар да е изграден само от 61 аминокиселини, ORF6 оказва значително влияние върху взаимодействието между SARS-CoV-2 и заразената клетка. Чрез нарушаване на транспорта през ядрените пори той отслабва активирането на противовирусните механизми. По-доброто познаване на неговата структура и взаимодействията му с клетъчните белтъци може да допринесе за разработването на нови антивирусни подходи.

3. Модифицирани циклични пептиди

3.1. Пептиди

Пептидите представляват органични молекули, изградени от две или повече аминокиселини, свързани помежду си чрез пептидни връзки. В зависимост от броя на аминокиселините се различават дипептиди (2 аминокиселини), трипептиди (3 аминокиселини), олигопептиди (обикновено до 20 аминокиселини) и полипептиди (над 20 аминокиселини), които изграждат структурата на много белтъци [5].

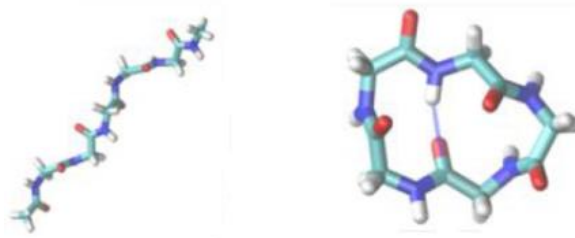
Пептидите изпълняват разнообразни биологични функции и участват в регулирането на различни процеси в организма. Те могат да действат като хормони и невромедиатори (молекули, които предават нервните импулси между клетките), както и да участват в имунната защита, което ги прави важни за нормалното функциониране на организма [5].

Поради своите свойства пептидите все по-често се използват при разработването на нови лекарствени средства. Въпреки това те се разграждат сравнително бързо под действието на ензими и имат ограничена стабилност в организма. Това налага разработването на различни подходи за подобряване на техните свойства чрез химично модифициране [5].

Един от подходите за подобряване на свойствата на пептидите е циклизацията на пептидната верига. При нея се образува затворена структура, характерна за цикличните пептиди [16].

3.2. Циклични пептиди

При цикличните пептиди пептидната верига е затворена в пръстеновидна структура. За разлика от линейните пептиди, при тях липсват свободни краища на веригата, което определя пространствената им структура и оказва влияние върху техните свойства [17].



Фиг. 3.2.1. Схематично сравнение между линеен и цикличен пептид.

Източник: Адаптирано по: Huang H., Damjanovic J., Miao J., Lin Y.-S. *Cyclic peptides: Backbone rigidification and capability of mimicking motifs at protein–protein interfaces*. *Physical Chemistry Chemical Physics*, 2021, 23(1), 607–616.

Цикличните пептиди могат да бъдат получени чрез различни начини на циклизация. Най-често тя се формира чрез свързване на началото и края на пептидната верига, но е възможно да се образува и чрез връзки между отделни аминокиселини. Различните начини

на циклизация оказват влияние върху структурата и свойствата на получените молекули [17].

Затворената форма намалява подвижността на пептидната верига и спомага за запазването на определена пространствена структура. Това улеснява свързването на цикличните пептиди с определени белтъци, което е важно за изпълнението на техните биологични функции [17].

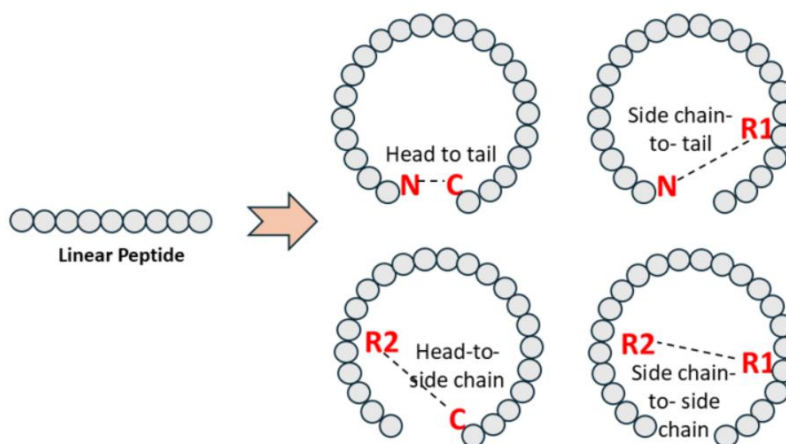
Освен природните циклични пептиди днес се разработват и множество синтетични варианти. Чрез химични промени могат да се изменят техните физикохимични и биологични свойства според целта на конкретното изследване [17].

3.3. Методи за модифициране на цикличните пептиди

За получаване и модифициране на циклични пептиди се използват различни химични подходи. Изборът на метод зависи от структурата на пептида и от свойствата, които се търсят [17].

Един от най-разпространените методи е циклизацията чрез свързване на N-края и C-края на пептидната верига (head-to-tail cyclization). При този подход между двата края на веригата се образува пептидна връзка, с която се затваря пептидната верига. Този метод се използва широко при получаването на различни циклични пептиди [17].

Освен чрез свързване на N-края и C-края, циклизацията може да се осъществи между единия край на пептидната верига и страничната верига на аминокиселинен остатък или между две странични вериги (фиг. 3.3.1)



Фиг. 3.3.1. Основни видове циклизация на пептидите.

Източник: Адаптирано по: INTAVIS Peptide Services. *Closing the Loop: How Cyclic Peptides Are Reshaping Therapeutics*.

При втория вариант могат да се образуват различни химични връзки, включително дисулфидни мостове между два цистеинови остатъка. Това води до получаването на циклични пептиди с различна пространствена структура [16,17].

Освен циклизацията се използват и други методи за химично модифициране на цикличните пептиди. Един от тях е включването на неприродни аминокиселини, които не участват в изграждането на естествените белтъци. Чрез тази модификация могат да се подобрят устойчивостта на пептидите към разграждане и способността им да се свързват с определени белтъци [16].

Друг често използван метод е N-метирирането, при което към азотния атом в пептидната връзка се добавя метилова група ($-\text{CH}_3$). Тази модификация може да подобри проникването на пептидите през клетъчните мембрани и да повиши устойчивостта им към действието на ензимите [18].

Често се комбинират няколко метода за химично модифициране, за да се постигнат желаните свойства на цикличните пептиди. Изборът на подход зависи от структурата на пептида и от целта на неговото приложение [18].

3.4. Биологични свойства и терапевтичен потенциал

Модифицираните циклични пептиди се проучват при бактериални инфекции, онкологични заболявания и вирусни инфекции, поради което интересът към тях в съвременната медицина продължава да нараства [16].

Едно от приложенията им е лечението на бактериални инфекции. Някои циклични пептиди могат да нарушат целостта на бактериалната клетъчна мембрана или да възпрепятстват жизненоважни процеси в бактериалните клетки. Затова те се проучват като възможни антибактериални средства, включително срещу бактерии, устойчиви към съществуващите антибиотици [16].

Модифицираните циклични пептиди се проучват и за лечение на онкологични заболявания. Те могат да повлияят процесите, свързани с растежа и деленето на туморните клетки, както и взаимодействията между различни белтъци, които имат роля в развитието на туморния процес. Това създава възможности за използването им в противотуморната терапия [16].

Друго важно приложение е използването им като антивирусни средства. Някои циклични пептиди могат да взаимодействат с вирусни белтъци и да нарушат тяхната функция, като по този начин възпрепятстват различни етапи от жизнения цикъл на вирусите. Подобни пептиди се разработват и срещу SARS-CoV-2 [16].

Голяма част от модифицираните циклични пептиди все още се намират на етап лабораторни изследвания. Необходими са допълнителни проучвания, за да се оцени тяхната ефективност и безопасност, преди да започнат да се използват в клиничната практика [16].

4. Взаимодействие между модифицирани циклични пептиди и белтъците на SARS-CoV-2

4.1. Молекулни механизми на взаимодействие

Взаимодействието между модифицираните циклични пептиди и белтъците на SARS-CoV-2 зависи от пространственото разположение на аминокиселинните остатъци и от химичните групи на двете молекули. За да се осъществи свързване, е необходимо формата на пептида да съответства на мястото за свързване в белтъка, което позволява образуването на стабилен молекулен комплекс [7].

Свързването се осъществява чрез различни нековалентни взаимодействия, като водородни връзки, електростатични и хидрофобни взаимодействия, както и сили на ван дер Ваалс. Всяко от тях има различен принос за стабилността на комплекса, а съвместното им действие определя силата и устойчивостта на взаимодействието между пептида и белтъка [7].

Пространствената структура на модифицираните циклични пептиди ограничава подвижността на пептидната верига и улеснява свързването с определени участъци от вирусните белтъци. Това повишава вероятността за селективно взаимодействие с избраната молекула мишена и за повлияване на нейната функция [7].

При SARS-CoV-2 особен интерес представляват белтъците NSP13 и ORF6. NSP13 участва във вирусната репликация, а ORF6 потиска клетъчния имуен отговор. Поради важната им роля те се разглеждат като подходящи мишени при разработването на модифицирани циклични пептиди с потенциално противовирусно действие [7].

4.2. Афинитет и специфичност на свързване

При взаимодействието между модифициран цикличен пептид и вирусен белтък важни характеристики са афинитетът и специфичността на свързване. Те показват доколко пептидът може да се свърже с определен белтък и са от значение за неговото потенциално биологично действие [7].

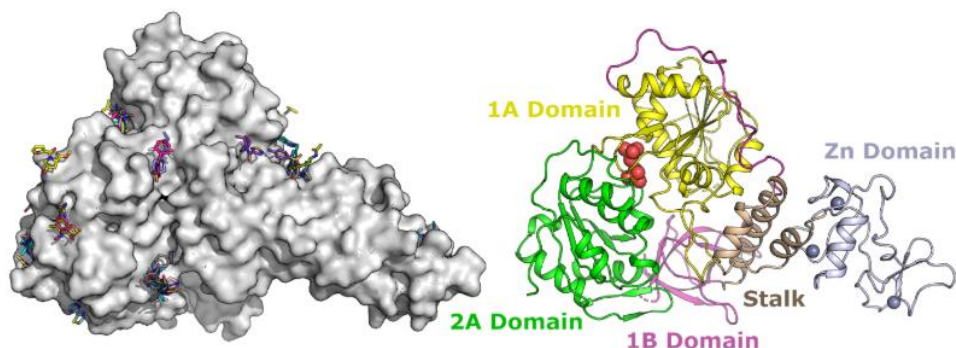
Афинитетът определя силата на свързване между пептида и белтъка. По-високият афинитет означава, че образуваният комплекс е по-устойчив и вероятността за запазване на взаимодействието е по-голяма [7].

Специфичността показва способността на пептида да разпознава определен белтък и да взаимодейства предимно с него. Високата специфичност е важна, тъй като намалява вероятността пептидът да се свързва с други белтъци в клетката и така ограничава възможността от нежелани ефекти [7].

При разработването на модифицирани циклични пептиди се търси съчетание между висок афинитет и добра специфичност. Това увеличава вероятността пептидът да взаимодейства ефективно с избрания вирусен белтък, без да засяга в значителна степен други клетъчни белтъци [7].

4.3. Взаимодействие с NSP13

NSP13 представлява една от възможните мишени при изследването на модифицирани циклични пептиди. Значението му като мишена е свързано с важната му роля във вирусната репликация и участието му в работата на репликационно-транскрипционния комплекс [19].



Фиг. 4.3.1. *Пространствена структура на белтъка NSP13 на SARS-CoV-2 с обозначени основни домени и места на свързване на изследваните химични фрагменти.*

Източник: Diamond Light Source. *NSP13 Helicase crystal structure and XChem fragment screen.* <https://www.diamond.ac.uk/covid-19/for-scientists/NSP13-Helicase-Structure-and-XChem.html>

Структурните изследвания на NSP13 показват наличие на различни области, в които могат да се свързват други молекули. Това дава възможност белтъкът да бъде изследван като мишена при търсенето на нови вещества, които могат да повлияят неговата функция [19].

4.4. Взаимодействие с ORF6

ORF6 представлява друга възможна мишена при изследването на модифицирани циклични пептиди. Белтъкът участва в потискането на противовирусния клетъчен отговор чрез взаимодействия с белтъци, свързани с транспорта между клетъчното ядро и цитоплазмата [10,15].

Поради тази функция се изследва възможността действието на ORF6 да бъде повлияно чрез молекули, които взаимодействат с него. В проекта AVATAR модифицирани циклотиди се изследват като потенциални инхибитори на ORF6, включително чрез компютърно моделиране и последващи експериментални методи [20].

4.5. Потенциал за инхибиране на вирусната репликация

Инхибирането на вирусната репликация представлява ограничаване на процесите, чрез които вирусът образува нови копия в заразената клетка. Това може да се постигне чрез повлияване на вирусни белтъци, които участват в различните етапи на жизнения цикъл на

вируса – навлизане в клетката, копиране на вирусната РНК, синтез на вирусни белтъци и образуване на нови вирусни частици [7].

Поради способността си да взаимодействат с определени вирусни белтъци модифицираните циклични пептиди се разглеждат като перспективни молекули за ограничаване на вирусната репликация. Ако функцията на ключови белтъци бъде нарушена, размножаването на вируса може да бъде затруднено, а разпространението на инфекцията – ограничено [7].

При SARS-CoV-2 особен интерес представляват белтъци като NSP13 и ORF6, тъй като изпълняват важни функции в жизнения цикъл на вируса. Поради това те са сред белтъците, към които са насочени съвременните изследвания за разработване на нови противовирусни средства [7].

4.6. Съвременни научни изследвания

Разработването на лекарства срещу SARS-CoV-2 показва, че вирусните белтъци могат да бъдат използвани като подходящи терапевтични мишени. Нирматрелвир (антивирусно лекарство) потиска действието на главната протеаза M^{pro} (вирусен белтък, необходим за образуването на функционални вирусни белтъци) и така затруднява размножаването на вируса [12]. Ремдесивир (антивирусно лекарство) повлиява процеса на копиране на вирусната РНК чрез действие върху РНК-зависимата РНК полимераза NSP12 (вирусен ензим, който синтезира вирусната РНК) [13].

Успехът на тези лекарства насочва вниманието и към други вирусни белтъци, които също могат да бъдат използвани при разработването на нови терапии срещу SARS-CoV-2. Сред тях са NSP13 и ORF6, към които през последните години се проявява все по-голям научен интерес. Разработват се молекули, които да взаимодействат с тези белтъци и да ограничат развитието на инфекцията [10,19].

Модифицираните циклични пептиди намират приложение в подобни изследвания, тъй като могат да бъдат разработени за взаимодействие с определени вирусни белтъци. След това техните свойства се оценяват чрез компютърни и лабораторни методи [16].

Пример за такова изследване е българският научен проект AVATAR, финансиран от ФНИ. В проекта се разработват модифицирани циклични пептиди, насочени към белтъците ORF6 и NSP13. Оценяват се тяхната стабилност, токсичност, способността им да проникват в клетките и потенциалното им противовирусно действие [20].

Изследванията по проекта продължават, като целта е да се оцени потенциалът на модифицираните циклични пептиди за разработване на нови средства срещу SARS-CoV-2 [20].

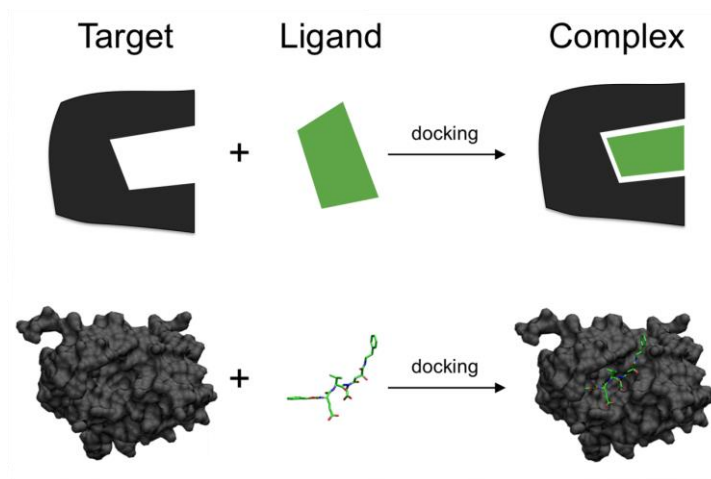
5. Методи за изследване

5.1.

Молекулен

докинг

Молекулният докинг е компютърен метод, който се използва за изследване на взаимодействието между две молекули. Чрез него може да се предвиди дали модифициран цикличен пептид може да се свърже с определен белтък и кои са възможните начини на това свързване [6].



Фиг. 5.1.1. Схематично представяне на процеса на молекулен докинг между лиганд (в случая модифициран цикличен пептид) и белтък и образуването на белтък-лиганден комплекс.

Източник: *Wikimedia Commons. Docking*. Достъпно на: <https://commons.wikimedia.org/wiki/Docking>.

При молекулния докинг компютърната програма разглежда различни положения и ориентации на пептида спрямо белтъка. Получените структури се оценяват, като се вземат предвид формата на молекулите и възможните взаимодействия между тях. По този начин могат да бъдат сравнени различни варианти на свързване и да бъдат избрани най-подходящите за последващ анализ [6].

Едно от основните предимства на молекулния докинг е, че позволява за сравнително кратко време да бъдат изследвани голям брой възможни структури, без всяка от тях първоначално да бъде проверявана експериментално. Получените резултати могат да служат като основа за последващи компютърни и лабораторни изследвания [6].

Основни принципи и функционалности на CABS-dock

За моделиране на взаимодействията между пептиди и белтъци може да се използва CABS-dock. Това е метод за гъвкав пептид-белтъчен докинг, разработен на основата на CABS модела. Една от основните му характеристики е възможността за извършване на

глобален гъвкав докинг, при който не е необходимо предварително да бъде известно точното място на свързване на пептида върху повърхността на белтъчния рецептор [39].

При CABS се използва опростено, или coarse-grained, представяне на белтъчната структура. Всеки аминокиселинен остатък се представя чрез до четири центъра на взаимодействие – C α атом, C β атом, център на страничната верига и псевдоатом, разположен в центъра на виртуалната връзка между два последователни C α атома. Последният се използва при описването на взаимодействията на главната верига и водородните връзки [40]. Това представяне е значително по-опростено от пълноатомното и позволява по-бързо изследване на голям брой възможни конформации.

Положенията на C α атомите са ограничени върху дискретна пространствена решетка с висока разделителна способност, докато останалите центрове се определят спрямо C α гръбнака. По този начин се намалява броят на необходимите изчисления, като в същото време се запазват основните структурни характеристики на молекулата [40].

Изследването на различните конформации се извършва чрез метода Монте Карло. При симулацията се извършват последователни малки промени в структурата. Новите конформации се приемат или отхвърлят според критерия на Метрополис в зависимост от промяната в енергията на системата. По този начин се изследват множество възможни пространствени състояния на пептида и белтъчния рецептор [40].

При глобалния CABS-dock пептидът започва от различни случайни положения около рецептора и може да изследва различни области от неговата повърхност. По време на симулацията пептидът може да променя положението, ориентацията и конформацията си. Белтъчният рецептор също не се разглежда като напълно неподвижен, а може да проявява определена гъвкавост. За да се запази общата му структура, могат да бъдат използвани ограничения по разстояние между определени C α атоми [39].

Оценъчна функция на CABS

За оценяване на генерираните конформации CABS използва статистически потенциал, базиран на знания (knowledge-based statistical potential). Неговите параметри са получени от статистически зависимости, наблюдавани в експериментално определени белтъчни структури от Protein Data Bank. По това CABS се различава от класическите физични силови полета, използвани при пълноатомната молекулна динамика [40].

Оценъчната функция включва няколко вида взаимодействия. Отчитат се локалните конформационни предпочитания на главната верига, взаимодействията между страничните вериги на аминокиселинните остатъци, изключеният обем и образуването на водородни връзки в главната верига. Взаимодействията между страничните вериги зависят от вида на аминокиселините и тяхното пространствено разположение. Ефектът на водната среда се отчита неявно (implicit solvent) чрез статистическите потенциали, без отделните водни молекули да бъдат включвани в модела [40].

Когато има предварителна информация за контакт между определени аминокиселинни остатъци, към CABS модела може да бъде добавен допълнителен контактен потенциал. Неговият вид е [43].

$$E(D) = \begin{cases} 0, & D \leq D_0 \\ s(D - D_0), & D > D_0 \end{cases}$$

където:

$E(D)$ – допълнителен енергиен бариер;

D – текущото разстояние между псевдоатомите, представящи избраните странични вериги;

D_0 – граничното разстояние, под което потенциалът е равен на нула;

s – коефициент, определящ нарастването на допълнителния енергиен бариер [43].

Когато разстоянието D е по-малко или равно на D_0 , не се добавя допълнителен енергиен бариер. Когато разстоянието стане по-голямо от зададената стойност, допълнителният енергиен бариер нараства линейно. Този допълнителен потенциал може да се използва за насочване на пептида към предварително известна област на свързване [43]. Важно е, че този член служи за насочване на докинга и не участва във финалната конформационна енергия на комплекса.

Генериране и подбор на моделите

Работата на CABS-dock преминава през няколко основни етапа. Първо се задават триизмерната структура на белтъчния рецептор и аминокиселинната последователност или структурата на пептида. След това чрез Монте Карло симулации се генерират множество възможни конформации на пептид-белтъчния комплекс [39].

Получените модели се оценяват чрез CABS потенциала и сходните структури се групират чрез структурно клъстериране. Клъстерирането позволява големият брой генерирани конформации да бъде разделен на групи от структурно подобни модели, от които се избират представителни структури [39].

Тъй като основната симулация се извършва с опростеното CABS представяне, избраните модели след това могат да бъдат възстановени до пълноатомно представяне. Това позволява получените структури да бъдат използвани за по-подробен анализ и за следващи етапи на компютърното моделиране [39].

Приложение при циклични пептиди

CABS-dock може да бъде използван и при докинг на циклични пептиди. При тях по време на симулацията е необходимо да бъде запазена затворената структура на пептида. За тази цел могат да бъдат зададени подходящи ограничения по разстояние. Методът може да

се използва както при пептиди с циклизирана основна верига, така и при пептиди, стабилизиращи чрез дисулфидни връзки [41].

При изследване на докинга на циклични пептиди с CABS-dock се генерира голям брой възможни модели, които след това се оценяват и клъстерират. Избраните модели могат да бъдат възстановени до пълноатомни структури и допълнително анализирани. Това позволява методът да бъде използван при търсенето на възможни начини на взаимодействие между циклични пептиди и белтъчни рецептори [41].

CABS-flex Standalone 3 обединява възможности за моделиране на структурната гъвкавост на белтъци, предсказване на пептидни структури и пептид-белтъчен докинг. Програмата е реализирана на Python 3 и позволява работа както с линейни, така и с циклични пептиди [29,42].

В настоящата дипломна работа молекулярният докинг е предвиден като последващ етап след молекулно-динамичните симулации на NSP13 и модифицираните циклотиди. За двата циклотиди lib9-L6-D1 и lib9-L6-D2 са избрани представителни структури от най-големите конформационни клъстери. В практическата част на работата избраната структура на lib9-L6-D1 е използвана за пептид-белтъчен докинг с NSP13, чието провеждане и получени резултати са представени в Глава 6. Изследването е част от научната програма на проекта AVATAR [20].

5.2. Молекулярна динамика

Молекулярната динамика е компютърен метод, чрез който може да се проследи как се движат атомите в една молекулярна система с течение на времето. За разлика от молекулярния докинг, чрез който се предвиждат възможни положения на свързване между пептида и белтъка, молекулярната динамика позволява да се проследи как се променя полученият комплекс с течение на времето и дали той остава стабилен [6].

Основата на молекулярната динамика са уравненията на Нютон. За всеки атом се разглежда силата, която действа върху него, неговата маса и начина, по който се променят координатите му във времето. Това може да бъде представено чрез уравнението:

$$m_i \frac{d^2 r_i}{dt^2} = F_i = - \frac{\partial U}{\partial r_i},$$

където m_i е масата на атома, r_i представя неговите координати, F_i е силата, която действа върху него, а U е потенциалната енергия на системата [23].

При система с много атоми се получават голям брой уравнения, които обикновено не могат да бъдат решени директно. Затова се използва числено интегриране. При него времето се разделя на много малки стъпки и за всяка от тях се изчисляват новите координати и скорости на атомите. Така постепенно се получава описание на движението на цялата система във времето [24].

За да се извърши това числено интегриране, могат да се използват различни алгоритми. Един от тях е алгоритъмът на Верле (Verlet). При него новите координати на

частицата се изчисляват, като се използват координатите ѝ в настоящия и в предходния момент, както и силата, която действа върху нея [30]:

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + \frac{F(t)}{m}(\Delta t)^2$$

Тук $r(t)$ представя координатите на частицата в момент t , Δt е малката времева стъпка, $F(t)$ е силата, която действа върху частицата, а m е нейната маса. Така чрез последователно повтаряне на изчисленията може да се проследи как се променят координатите на частицата във времето [30].

Друг вариант е алгоритъмът на Верле със скорости (Velocity Verlet). При него се изчисляват както новите координати, така и новата скорост на частицата. Първо се определят координатите ѝ за следващата времева стъпка [30]:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{F(t)}{2m}(\Delta t)^2$$

След това се определя и новата скорост:

$$v(t + \Delta t) = v(t) + \frac{F(t) + F(t + \Delta t)}{2m}\Delta t$$

Тук $v(t)$ е скоростта на частицата в момент t . При този алгоритъм координатите и скоростта се получават за един и същи момент от времето, което е удобно при проследяване на движението на частиците [30].

Друг често използван алгоритъм е Leapfrog („прескочи кобила“). При него координатите и скоростите не се изчисляват за един и същи момент. Скоростта се определя на половин времева стъпка спрямо координатите [30]:

$$v\left(t + \frac{\Delta t}{2}\right) = v\left(t - \frac{\Delta t}{2}\right) + \frac{F(t)}{m}\Delta t$$

След това чрез получената скорост се изчисляват новите координати:

$$r(t + \Delta t) = r(t) + v\left(t + \frac{\Delta t}{2}\right)\Delta t$$

По този начин изчисляването на скоростите и координатите се редува във времето, сякаш те се „прескачат“ едно друго. Оттук идва и наименованието Leapfrog. И трите алгоритъма позволяват движението на частиците да бъде проследявано чрез последователни малки времеви стъпки [30].

При молекулно-динамичните симулации алгоритмите от фамилията на Верле и Leapfrog често се предпочитат пред класическите методи като Runge–Kutta. Причината е, че те са симплектични и времево обратими. Това е важно при продължителни симулации, защото позволява по-добре да се запази поведението на системата във времето [30].

Методът Runge–Kutta от четвърти ред (RK4) има висока локална точност от четвърти ред $O(\Delta t^4)$. При дълги симулации на консервативни системи обаче малките грешки могат постепенно да се натрупват, което да доведе до изкуствен дрейф на енергията. При алгоритмите на Верле енергията също може да се изменя в малки граници, но остава около определена средна стойност, без да се наблюдава постоянен дрейф при продължителна симулация [30].

Друга важна причина за използването на алгоритмите на Верле е тяхната изчислителна ефективност. При симулации на системи с много частици голяма част от изчислителния ресурс се използва за определяне на силите между тях. Алгоритъмът на Верле изисква едно ново изчисляване на силите за една времева стъпка, докато при RK4 са необходими четири междинни изчисления. Така се спестява изчислителен ресурс и става възможно да се проследяват по-дълги процеси [30].

В заключение, за затворени консервативни системи алгоритмите от фамилията на Верле съчетават стабилност при продължителни симулации с малък изчислителен разход. Това ги прави подходящ избор и за молекулно-динамичните симулации, предвидени в дипломната работа. В сравнение с тях несимплектични методи като RK4 са по-подходящи при системи с дисипация или при задачи, при които е необходима висока точност за по-кратък период от време [30].

Преди началото на симулацията трябва да бъдат зададени начални условия. Те включват началните координати и началните скорости на атомите. Координатите могат да бъдат взети от експериментално определена структура или от база данни като Protein Data Bank (PDB). Началните скорости обикновено се задават така, че да отговарят на избраната температура на системата [25].

За да се изчислят силите между атомите, се използват математически модели, наречени силови полета (force fields). Те описват различните взаимодействия между атомите и позволяват да се изчисли потенциалната енергия на системата. От тази енергия след това се определят силите, които действат върху отделните атоми [23].

Силовото поле отчита както взаимодействията между атомите, които са свързани помежду си, така и взаимодействията между атоми, които не са свързани с химична връзка. Към първата група спадат разтягането на връзките, промяната на ъглите и торзионните движения. Към втората група се отнасят основно ван дер Ваалсовите и електростатичните взаимодействия [23].

В опростен вид потенциалната енергия може да бъде представена като:

$$U = U_{\text{връзки}} + U_{\text{ъгли}} + U_{\text{торзии}} + U_{\text{van der Waals}} + U_{\text{електростатични}}.$$

След като се изчисли потенциалната енергия, от нея се определят силите върху атомите. От силите се получават ускоренията, а чрез численото интегриране се намират

новите скорости и координати на атомите. По този начин се проследява развитието на молекулната система с течение на времето [23].

Съществуват различни видове силови полета според начина, по който е представена молекулната система. При all-atom моделите се разглеждат всички атоми. При united-atom моделите някои атоми се обединяват, а при coarse-grained моделите цели групи от атоми се представят чрез по-опростени частици. Сред често използваните силови полета за биомолекули са CHARMM, AMBER, OPLS и GROMOS [26].

При молекулно-динамичните симулации е важно да се вземат предвид и условията, при които се намира системата. Затова могат да се използват различни статистически ансамбли. При NVE ансамбъла броят на частиците, обемът и общата енергия са постоянни. При NVT ансамбъла се поддържат постоянни броят на частиците, обемът и температурата, а при NPT ансамбъла постоянни са броят на частиците, температурата и налягането [24].

За поддържане на температурата се използва термостат, който влияе върху скоростите на атомите. За поддържане на налягането се използва баростат, който позволява да се променя обемът на системата. По този начин могат да се създадат условия, които са по-близки до тези, при които молекулите се намират в реална биологична среда [24].

Преди основната част на молекулно-динамичната симулация обикновено се извършва уравновесяване на системата. При него първо може да се стабилизира температурата, а след това и налягането. Целта е системата да достигне по-стабилно състояние преди да започне същинското събиране на данни от симулацията [24].

Крайният резултат от молекулно-динамичната симулация е траектория, която показва как се променят координатите на атомите във времето. Тази траектория позволява да се проследяват измененията в структурата на молекулите и начина, по който те взаимодействат една с друга [23].

При анализа не се разглежда само една структура в определен момент. Проследяват се различни показатели за определен период от симулацията и често се използват техните средни стойности. По този начин може по-надеждно да се оцени дали структурата е стабилна и как се променя във времето [23].

При изследване на пептид-белтъчни комплекси молекулната динамика може да се използва за проследяване на структурните промени на комплекса във времето и за оценка на неговото поведение при зададените условия на симулацията [6].

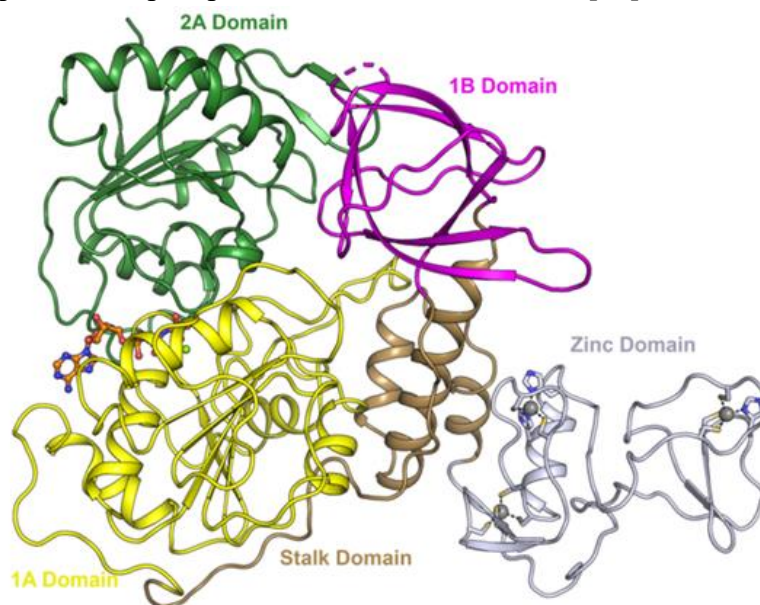
6. Компютърно моделиране на NSP13

6.1. Подготовка на структурата на NSP13

За провеждане на молекулно-динамичната симулация е използвана триизмерната структура на хеликазата NSP13 на SARS-CoV-2 с идентификационен код 7NIO от базата данни Protein Data Bank (PDB) [25]. Структурата съдържа две белтъчни вериги – А и Е, като за симулацията е избрана верига А. Всяка от двете вериги съдържа по три структурни Zn^{2+} йона, които имат значение за запазването на структурата на N-терминалния цинк-свързващ домен.

Наличието на тези три цинкови йона е характерна особеност на структурата на NSP13. Белтъкът съдържа пет основни домена – цинк-свързващ домен (ZBD), Stalk домен, 1B домен и два RecA-подобни домена – 1A и 2A. Последните два участват в свързването и разграждането на нуклеотиди [19].

Структурните изследвания на NSP13 показват също, че между отделните домени могат да настъпват промени в пространственото им положение [19].



Фиг. 6.1.1.. Пространствена структура на NSP13 на SARS-CoV-2 с означени основни домени.

Източник: *Newman J. A. et al. [19]*.

За подготовката на системата е отделена веригата А заедно с трите Zn^{2+} йона. Отстранени са кристалографските водни молекули и другите хетероатоми, които не са необходими за изграждането на модела. Особено внимание е отделено на цинковите центрове, тъй като Zn^{2+} йоните са координирани от определени цистеинови и хистидинови остатъци и тази организация трябва да се запази по време на симулацията.

За описване на взаимодействията между атомите на белтъка е използвано силовото поле CHARMM36/CHARMM36m [31–34]. Силовото поле представлява набор от математически зависимости и параметри, чрез които се изчисляват взаимодействията между атомите по време на молекулно-динамичната симулация. Водната среда е представена чрез модела TIP3P [36].

След подготовката на структурата NSP13 е поставен в периодична ромбично-додекаедрична симулационна кутия. Минималното разстояние между белтъка и границата на кутията е 1.2 nm. Системата е запълнена с 39 699 водни молекули, представени чрез модела TIP3P [36]. Добавени са Na^+ и Cl^- йони до крайна концентрация 150 mM.

Преди започване на основната молекулно-динамична симулация е извършена енергийна минимизация. При нея се намаляват неблагоприятните взаимодействия и прекалено близките контакти между атомите в началната структура. След това системата е уравновесена последователно в NVT и NPT ансамбъл. Проведени са 500 ps NVT и 1 ns NPT уравновесяване при температура 310 K и налягане 1 bar.

Температурата е контролирана чрез velocity-rescale термостат [27], а налягането – чрез баростат на Parrinello–Rahman [28]. Избраната температура от 310 K съответства приблизително на 37 °C и позволява симулацията да се провежда при температура, близка до физиологичната.

За електростатичните взаимодействия е използван методът Particle-Mesh Ewald (PME) [37]. Връзките, включващи водородни атоми, са ограничени чрез алгоритъма LINCS [38], а използваната времева стъпка при симулацията е 2 fs.

За подготовката на структурата, провеждането на молекулната динамика и анализа на получената траектория е използвана програмата GROMACS 2023.5 [35]. GROMACS е специализиран софтуер за провеждане и анализ на молекулно-динамични симулации. Използването му е предвидено и в проекта AVATAR като един от основните програмни пакети за компютърно моделиране [20].

6.2. Молекулно-динамична симулация на NSP13

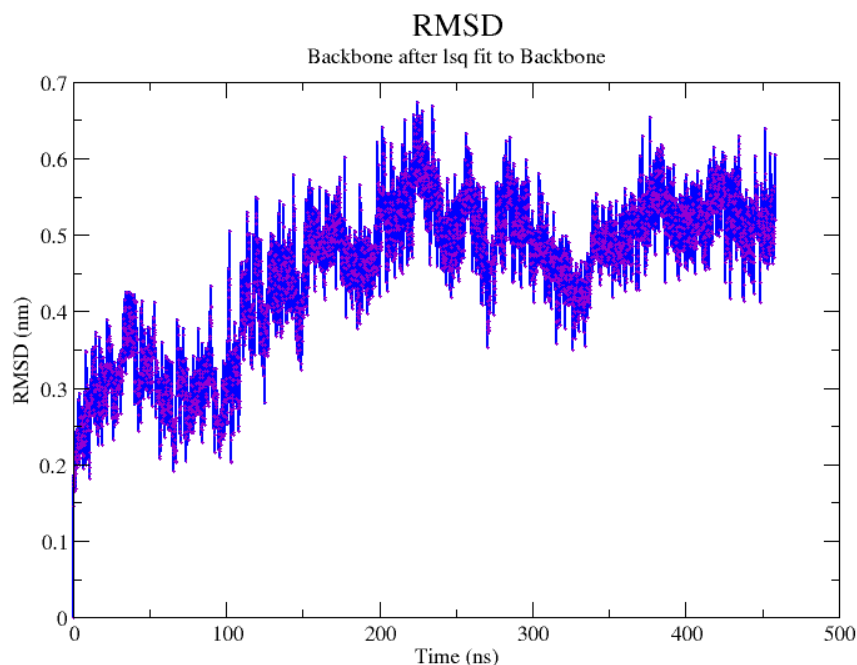
Основната молекулно-динамична симулация на NSP13 е проведена в NPT ансамбъл при температура 310 K и налягане 1 bar. По време на симулацията периодично са записвани координатите на атомите, което позволява да бъде проследено движението на белтъка във времето. Кадрите от траекторията са записвани през 10 ps, а данните за енергията – през 2 ps.

По време на симулацията е проследявано и състоянието на трите структурни цинкови центъра. Проверявани са разстоянията между Zn^{2+} йоните и координиращите ги атоми, както и пространствената геометрия около всеки цинков йон. Целта е да се установи дали тези области запазват своята структура по време на симулацията.

За оценка на поведението на NSP13 в получената траектория са използвани два основни показателя – RMSD (Root Mean Square Deviation) и RMSF (Root Mean Square

Fluctuation). Те позволяват да се проследят както общите промени в структурата на белтъка, така и подвижността на отделните му аминокиселинни остатъци [30].

RMSD представлява средноквадратично отклонение на атомните позиции спрямо избрана референтна структура. Чрез този показател може да се проследи как се променя цялостната структура на NSP13 по време на молекулно-динамичната симулация [30].

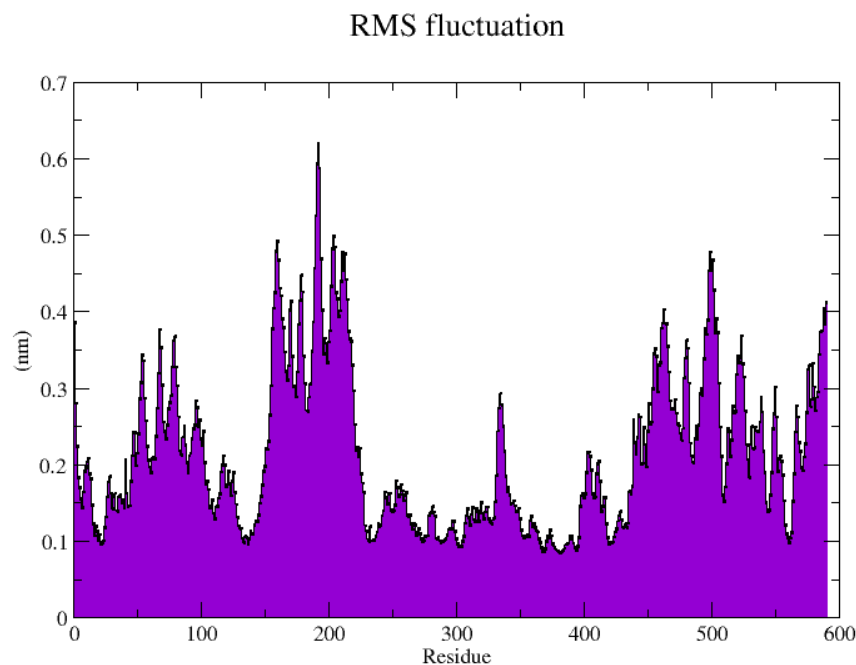


Фиг. 6.2.1 Изменение на RMSD на белтъчния гръбнак на NSP13 по време на молекулно-динамичната симулация.

По хоризонталната ос е представено времето в наносекунди (ns), а по вертикалната – RMSD в нанометри (nm). В началната част на симулацията се наблюдава постепенно изменение на RMSD, свързано с приспособяването на белтъчната структура към условията на симулационната среда. След приблизително 150–200 ns стойностите се изменят основно в сравнително ограничен диапазон. Наблюдават се отделни повишения и понижения, но липсва постоянно нарастване на RMSD до края на проследения период.

Полученият резултат показва, че NSP13 претърпява структурни промени по време на симулацията, но не се наблюдава непрекъснато увеличаване на отклонението от началната структура. Подобна подвижност е характерна за NSP13, тъй като експерименталните структурни изследвания показват промени в относителното положение на отделните му домени [19]. В публикуваните структури например са установени движения на цинк-свързващия домен и на домовете 1B, 1A и 2A [19].

За разлика от RMSD, който показва общото изменение на структурата във времето, RMSF се използва за оценка на подвижността на отделните аминокиселинни остатъци. Повисоките RMSF стойности показват области с по-голяма подвижност, а по-ниските стойности съответстват на по-слабо подвижни участъци [30].



Фиг. 6.2.2 RMSF на аминокиселинните остатъци на NSP13 по време на молекуло-динамичната симулация.

По хоризонталната ос са представени аминокиселинните остатъци на NSP13, а по вертикалната – стойността на RMSF в nm. От графиката се вижда, че подвижността не е еднаква по цялата белтъчна верига. Значителна част от аминокиселинните остатъци показват сравнително малки отклонения, докато в определени области се наблюдават по-високи стойности.

По-изразена подвижност се наблюдава в областта приблизително между остатъци 150 и 220, където се намира и най-високият пик на графиката. Повишени стойности се наблюдават и в някои участъци от втората половина на белтъчната верига. Това показва, че NSP13 съдържа както по-стабилни, така и по-гъвкави области.

Полученото поведение е в съответствие с известната конформационна гъвкавост на NSP13 [19].

Структурните изследвания показват, че цинк-свързващият домен може да променя положението си с около 10–15°, а при 1B домена са наблюдавани промени от около 25° [19].

Съвместното разглеждане на RMSD и RMSF позволява да се получи по-пълна представа за поведението на NSP13 по време на молекулно-динамичната симулация. RMSD показва промените на цялата структура във времето, докато RMSF показва кои отделни участъци на белтъка са по-подвижни.

След началния период на еквилибриране на структурата, RMSD се изменя в ограничен диапазон от стойности, без да показва непрекъснато нарастване. RMSF от своя страна показва, че по-големите движения са съсредоточени в определени области на белтъчната верига. Тези резултати са в съответствие с известната структурна гъвкавост на NSP13 и с наличието на различни пространствени състояния на белтъка [19].

Newman и съавт. описват отворени и затворени конформации на NSP13, които са свързани с различни етапи от каталитичния му цикъл [19]. Това показва, че определена степен на относително движение на домените е естествена част от функционирането на белтъка.

Получената след молекулно-динамичната симулация структура може да бъде използвана в следващия етап на изследването – молекулен докинг с модифицирани циклични пептиди.

6.3. Модели на модифицираните циклотиди

За последващото изследване на взаимодействието с NSP13 са подготвени два модифицирани циклотиди – 1ib9-L6-D1 и 1ib9-L6-D2. И двата са създадени на основата на циклотиди MCoTI-II (*Momordica cochinchinensis* trypsin inhibitor II), чиято структура е налична в Protein Data Bank с идентификационен код 1IB9 [25]. MCoTI-II съдържа 34 аминокиселинни остатъка и има циклична структура, стабилизирана от три дисулфидни връзки – Cys7–Cys25, Cys8–Cys21 и Cys15–Cys33. Те образуват характерния за циклотидите цистинов възел.

Структурата на MCoTI-II съдържа шест участъка, разположени между цистеиновите остатъци. За модифицирането е избран участък 6 (loop 6), който е най-дългият и е разположен на повърхността на молекулата. Този участък позволява включването на допълнителни аминокиселинни последователности, като основната структура на циклотиди се запазва.

При двата варианта естествената последователност SGSDGG в участък 6 е заменена с модифицирани последователности, включващи части от d1 и d2 фрагментите на човешкия интерферон- γ (IFN- γ). Така са получени циклотидите 1ib9-L6-D1 и 1ib9-L6-D2, при които са запазени цикличната пептидна верига и цистиновият възел на MCoTI-II.

Модифициран циклотид 1ib9-L6-D1

При 1ib9-L6-D1 естествената последователност SGSDGG в участък 6 на MCoTI-II е заменена с последователността ALAKTGRR, която включва основни аминокиселинни остатъци, свързани с d1 фрагмента на човешкия IFN- γ . Полученият циклотид съдържа 36

аминокиселинни остатъка и има общ заряд +9 при неутрално рН. Запазени са цистиновият възел и трипсин-инхибиторният участък PKILKK.

Аминокиселинната последователност на lib9-L6-D1 е:

ALAKTGRRVCPKILKKCRRDSDCPGACICRGNGYCG.

Модифициран циклотид lib9-L6-D2

При lib9-L6-D2 естествената последователност SGSDGG в участък 6 е заменена с последователността RGRRDGG, която включва основни аминокиселинни остатъци, свързани с d2 фрагмента на човешкия IFN- γ . Полученият циклотид съдържа 35 аминокиселинни остатъка и има общ заряд +6 при неутрално рН. При него също са запазени цистиновият възел и трипсин-инхибиторният участък PKILKK.

Аминокиселинната последователност на lib9-L6-D2 е:

RGRRDGGVCPKILKKCRRDSDCPGAICRGNGYCG.

6.3.1 Симулационен протокол

Двата модифицирани циклотида са подготвени по еднакъв начин за молекулно-динамичните симулации. Всеки от тях е поставен в додекаедрична симулационна кутия, като минималното разстояние между пептида и границата на кутията е 1.0 nm. Системите са запълнени с водни молекули, представени чрез модела TIP3P [36], и са добавени Na⁺ и Cl⁻ йони до концентрация 0.15 M NaCl. За lib9-L6-D1 са добавени 14 Na⁺ и 23 Cl⁻ йона, а за lib9-L6-D2 – 10 Na⁺ и 16 Cl⁻ йона.

За описване на взаимодействията между атомите е използвано силовото поле CHARMM36 [31–34]. При подготовката са запазени дисулфидните връзки и цикличната структура на пептидите. Цикличната връзка при lib9-L6-D1 е между Gly36 и Ala1, а при lib9-L6-D2 – между Cys35 и Arg1.

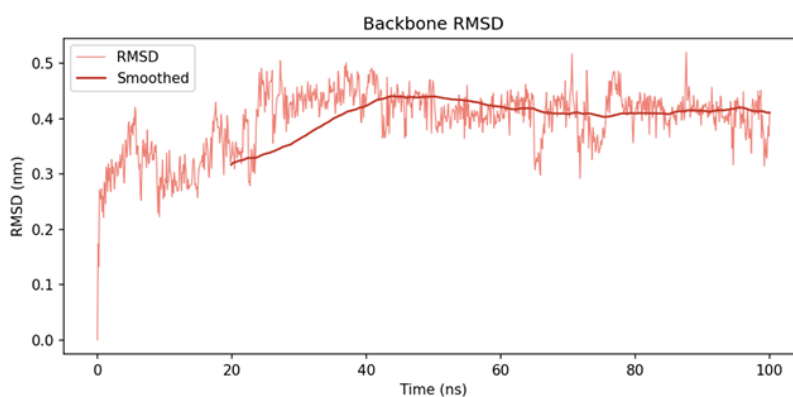
Симулациите са проведени с програмата GROMACS [35] при времева стъпка 2 fs. Първо е извършена енергийна минимизация, след което системите са нагрети и уравновесени при температура 310 K. След NVT уравновесяването е проведено NPT уравновесяване при температура 310 K и налягане 1 bar, последвано от основна молекулно-динамична симулация с продължителност 100 ns. Температурата е контролирана чрез V-rescale термостат [27], а налягането – чрез баростат на Parrinello–Rahman [28]. Електростатичните взаимодействия са изчислявани чрез метода Particle-Mesh Ewald [37], а връзките, включващи водородни атоми, са ограничени чрез алгоритъма LINCS [38].

За оценка на поведението на двата циклотида са анализирани RMSD, RMSF, радиусът на инерция, вторичната структура, вътрешните водородни връзки и разпределението на получените конформации в клъстери. За всеки циклотид е определена представителна структура от най-големия клъстер, която може да бъде използвана при последващ молекулен докинг.

6.3.2 Резултати

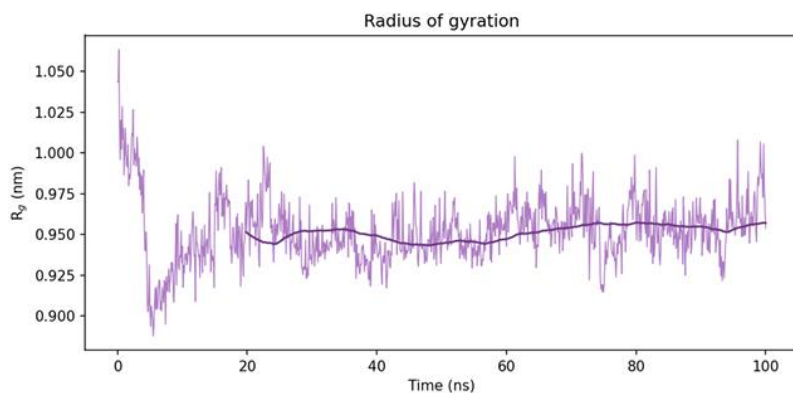
Циклотид *lib9-L6-D1*

RMSD на пептидния гръбнак на *lib9-L6-D1* постепенно нараства през началната част на симулацията и след приблизително 40 ns достига плато около 0.41 nm. Средната стойност на RMSD за цялата траектория е 0.396 nm, като стойностите се изменят в интервала 0.00–0.52 nm. След достигането на платото се наблюдават сравнително малки колебания, което показва, че циклотидът запазва оформена пространствена структура през останалата част от симулацията.



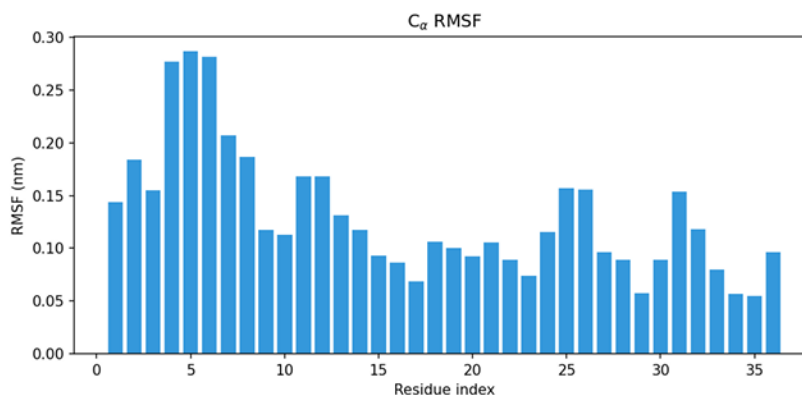
Фиг. 6.3.2.1 Изменение на RMSD на пептидния гръбнак на *lib9-L6-D1* по време на 100 ns молекулно-динамична симулация.

Инерчният радиус (radius of gyration, R_g) остава сравнително постоянен през цялата симулация със средна стойност 0.953 nm и диапазон 0.89–1.06 nm. Това показва, че *lib9-L6-D1* запазва компактната си форма по време на симулацията.



Фиг. 6.3.2.2 Изменение на инерчния радиус на *lib9-L6-D1* по време на молекулно-динамичната симулация.

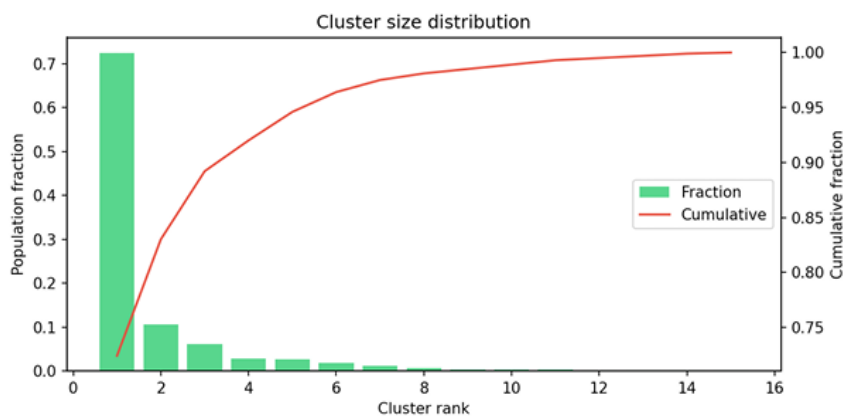
RMSF анализът показва, че подвижността не е еднаква във всички части на 1ib9-L6-D1. Най-големи колебания се наблюдават в модифицирания участък, включващ остатъци 1–8. Най-високите стойности са отчетени при Thr5 – 0.29 nm, и Gly6 – 0.28 nm. Областите на цистиновия възел показват по-ниски RMSF стойности, което показва по-ограничена подвижност на тази част от молекулата.



Фиг. 6.3.2.3 RMSF на аминокиселинните остатъци на 1ib9-L6-D1 по време на молекулно-динамичната симулация.

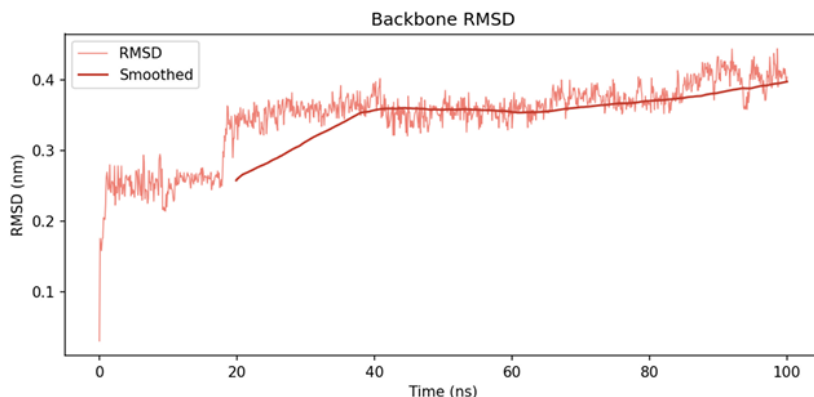
При конформационния анализ на 1ib9-L6-D1 са определени 15 различни клъстера. Най-големият от тях включва 725 от 1001 анализирани структури, което представлява приблизително 72% от траекторията. Наличието на един преобладаващ клъстер показва, че циклотидът прекарва значителна част от симулацията в близки пространствени състояния.

Структурата центроид на най-големия клъстер е избрана като представителна за 1ib9-L6-D1 и може да бъде използвана при последващ молекулен докинг.



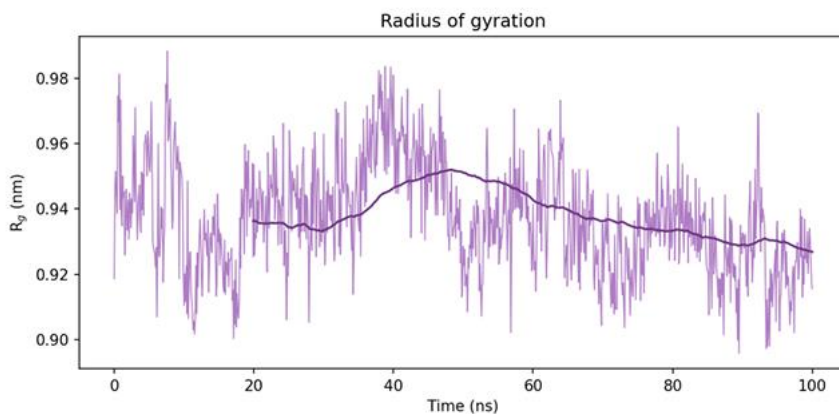
Фиг. 6.3.2.4 Разпределение на конформационните клъстери на 1ib9-L6-D1.
Циклотид 1ib9-L6-D2

При 1ib9-L6-D2 RMSD на пептидния гръбнак постепенно нараства през началната част на симулацията и след приблизително 50 ns достига плато около 0.35 nm. Средната стойност за цялата траектория е 0.347 nm, като стойностите се изменят между 0.03 и 0.44 nm. След началното пренареждане стойностите се изменят около сравнително постоянно ниво.



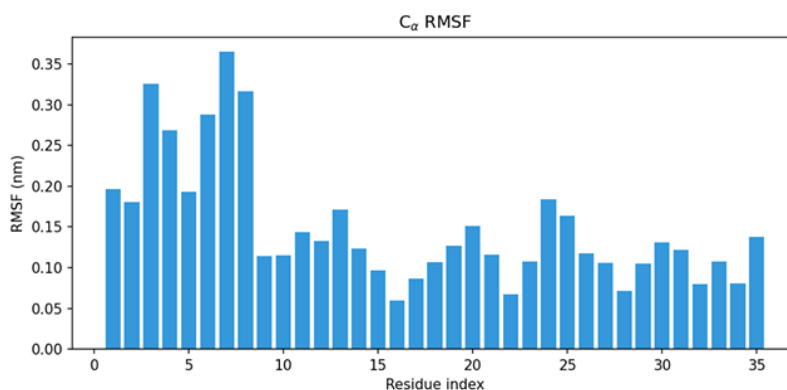
Фиг. 6.3.2.5 Изменение на RMSD на пептидния гръбнак на 1ib9-L6-D2 по време на 100 ns молекулно-динамична симулация.

Инерчният радиус остава сравнително постоянен със средна стойност 0.937 nm и диапазон 0.85–1.02 nm. Това показва, че и 1ib9-L6-D2 запазва компактната си форма по време на симулацията.



Фиг. 6.3.2.6. Изменение на инерчният радиус на 1ib9-L6-D2 по време на молекулно-динамичната симулация.

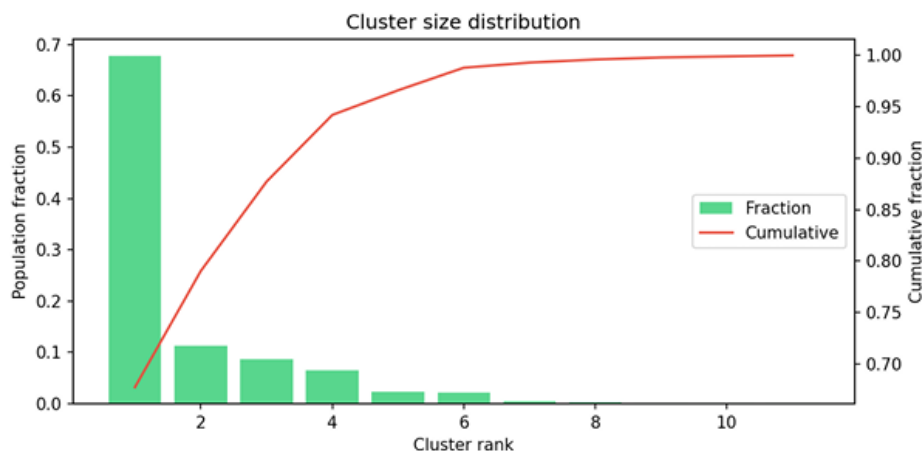
RMSF анализът показва по-висока подвижност основно в модифицирания участък на 1ib9-L6-D2. Най-високите стойности са 0.37 nm при Cys7 и 0.33 nm при Arg3. В областта на цистиновия възел RMSF стойностите са по-ниски, което показва по-ограничена подвижност на основната структура на циклотида.



Фиг. 6.3.2.7 RMSF на аминокиселинните остатъци на *lib9-L6-D2* по време на молекулно-динамичната симулация.

При *lib9-L6-D2* са определени 11 различни конформационни клъстера. Най-големият включва 678 от 1001 анализирани структури, или приблизително 68% от траекторията. И при този циклотид се наблюдава едно ясно преобладаващо конформационно състояние.

Представителната структура може да бъде използвана като начална структура при последващ молекулен докинг.



Фиг. 6.3.2.8 Разпределение на конформационните клъстери на *lib9-L6-D2*.

6.4. Сравнение на двата модифицирани циклотида

Резултатите от 100 ns молекулно-динамични симулации показват, че и двата модифицирани циклотида запазват компактна пространствена структура. Средният RMSD е 0.396 nm за *lib9-L6-D1* и 0.347 nm за *lib9-L6-D2*, а средният инерчен радиус е съответно 0.953 nm и 0.937 nm. Средната RMSF стойност е 0.130 nm за *lib9-L6-D1* и 0.150 nm за *lib9-L6-D2*.

Пептид-Протеин докинг

След молекулно-динамичните симулации на NSP13 и двата модифицирани циклотида са избрани представителни структури за последващ молекулен докинг. За 1ib9-L6-D1 и 1ib9-L6-D2 са използвани центроидите на най-големите клъстери, определени при анализа на техните молекулно-динамични траектории. Като представителна структура на NSP13 е избран последният кадър от молекулно-динамичната симулация на белтъка. Представителните структури на двата циклотида са получени от проведените 100 ns симулации и клъстерния анализ. Докладът за симулациите определя центроидите на доминиращите клъстери като подходящи начални структури за последващ пептид-белтъчен докинг.

За провеждането на докинга е използван CABSflex подход, при който се отчита гъвкавостта както на пептида, така и на белтъчния рецептор. Методът използва опростено (coarse-grained) представяне на молекулите, което позволява да бъдат изследвани голям брой възможни положения и конформации на пептида спрямо белтъка [39,42]. Подробните принципи на CABS-dock и използваната оценъчна функция са описани в т. 5.1.

Докинг на NSP13 с 1ib9-L6-D1

За изследване на взаимодействието между NSP13 и модифицирания циклотид 1ib9-L6-D1 са проведени 25 независими CABSflex симулации при еднакви параметри. При всяка симулация са изследвани различни положения и конформации на циклотида спрямо NSP13.

След оценяване и клъстериране на получените конформации от всяка симулация са избрани по десет от най-добре класираните представители. Те са възстановени до пълноатомно представяне чрез cg2all и Modeller. Така от 25-те независими симулации са получени общо 250 пълноатомни модела на комплекса NSP13–1ib9-L6-D1.

При моделирането е проследено запазването на цикличната структура на 1ib9-L6-D1 и на трите дисулфидни връзки – Cys10–Cys27, Cys17–Cys29 и Cys23–Cys35. Запазването на цикличната структура и дисулфидните връзки е важно при моделирането на циклични пептиди [41].

Оценяване и подбор на моделите

Получените пълноатомни модели са оценени чрез DOPE (Discrete Optimized Protein Energy) и MolPDF (Modeller objective function). DOPE е използван като основен показател при подреждането на моделите, като по-ниските и по-отрицателни стойности съответстват на по-добре оценени структури. MolPDF е използван като допълнителен показател при сравняването на моделите.

Крайният подбор не е направен само според една стойност. Вzeti са предвид DOPE и MolPDF, появата на сходни решения при различните независими симулации, запазването на цикличната структура и дисулфидните връзки и структурното сходство между най-добре оценените модели.

За оценка на структурното сходство е използван RMSD на Ca атомите след пространствено напасване на структурите. По този начин може да се установи дали различни добре оценени модели показват сходно положение и структура на комплекса.

Проверка на геометрията на циклотида

При получените модели е проверено дали цикличната структура на 1ib9-L6-D1 е запазена. За тази цел е проследено разстоянието между N1 и C36. При стойности под 2.2 Å структурата е приемана за затворена, а стойности между 2.2 и 4.0 Å са разглеждани като близки до затваряне.

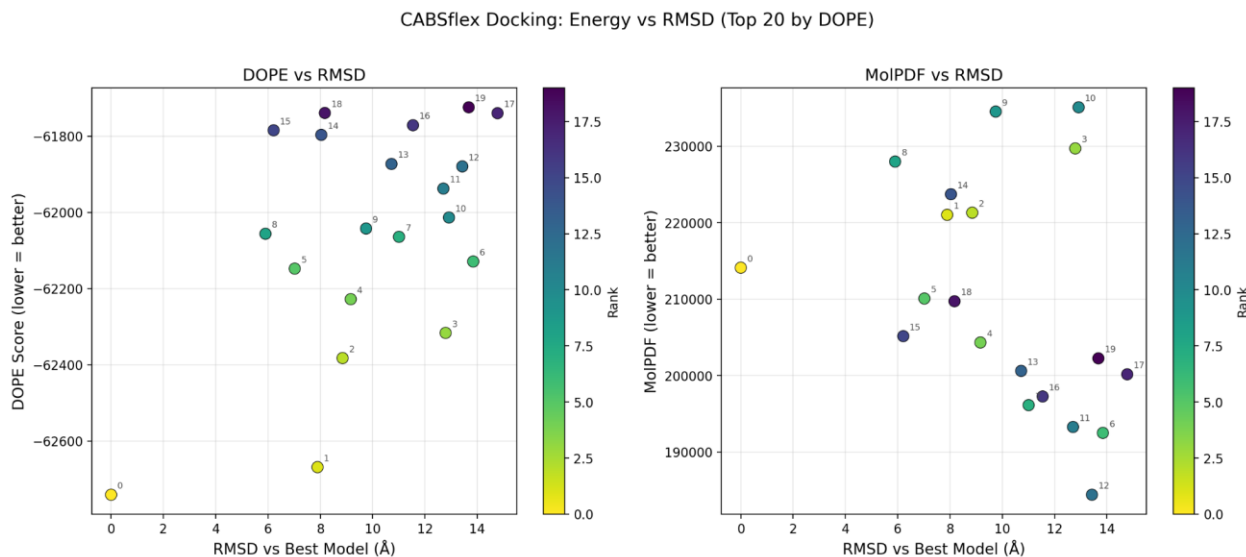
Проверена е и геометрията на трите дисулфидни връзки чрез разстоянията между серните атоми (SG–SG) на съответните цистеинови остатъци. Разстояния под 2.8 Å са приемани като съответстващи на образувана дисулфидна връзка, между 2.8 и 3.6 Å – като близки до необходимата геометрия, а над 3.6 Å – като нарушена връзка. Моделите с неподходяща геометрия не са предпочитани при крайния подбор.

Резултати от докинга

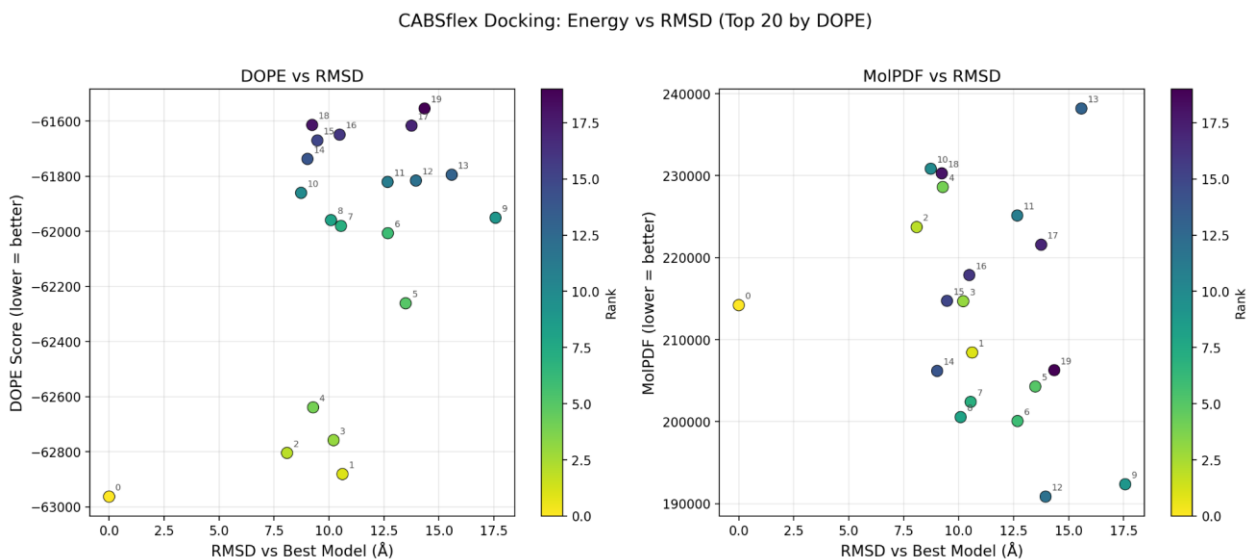
Всички 25 независими симулации са завършили успешно и са получени общо 250 структури на атомно ниво на комплексите NSP13–1ib9-L6-D1 и NSP13–1ib9-L6-D2.

DOPE стойностите за получените модели са в интервала от –62740.9 до –56463.5, със средна стойност –60476.6. Най-ниската DOPE стойност е получена за model_4.pdb от run_22, при който DOPE = –62740.9, а MolPDF = 214103.9 (виж фиг. 6.4.1 и фиг. 6.4.2).

При най-добре оценените модели е направено и RMSD сравнение. Целта е да се установи дали модели с близки енергийни оценки показват сходни пространствени структури. Появата на сходни решения в независими симулации дава допълнителна подкрепа за определено положение на циклотида спрямо NSP13.



Фиг. 6.4.1 Графично представяне на DOPE, MolPDF и RMSD на най-добрите 20 комплекса NSP13–1ib9-L6-D1. Цветовата скала (и числото до маркерите) представя ранга на подреждане в DOPE-RMSD пространството като 0 най-високия ранг.



Фиг. 6.4.2 Графично представяне на DOPE, MolPDF и RMSD на най-добрите 20 комплекса NSP13–1ib9-L6-D2. Цветовата скала (и числото до маркерите) представя ранга на подреждане в DOPE-RMSD пространството като 0 най-високия ранг.

Избор на представителни модели на комплексите

След сравняване на получените резултати модела с ранг 6 е определен като най-добър за комплекса NSP13–1ib9-L6-D1 при използваните условия и модела с ранг 8 за комплекса NSP13–1ib9-L6-D2 (виж фиг. 6.4.1 и фиг. 6.4.2). Тези модели имат относително ниска DOPE стойност сред получените модели, като при избора са отчетени също MolPDF, геометрията на циклотиди и структурното сходство с останалите добре оценени модели.

След внимателно изследване на структурата на циклотидите в комплексите се оказва, че въпреки наложените ограничения за цикличност на пептидните им вериги и запазването на всеки от дисулфидните мостове поотделно, в геометрията на цистиновите възли има значително отклонения от нативните структури. Последното няма как да се случи спонтанно поради здравината на цистиновите възли. Ето защо е необходимо оценъчната функция да бъде модифицирана така, че да запазва геометрията на цистиновия възел или да бъде съставена система от ограничения в разстоянията между специфични групи от атоми, осигуряващи това запазване по време на докинг процедурата. За реализирането на първото е необходимо да се напише допълнителен модул към програмата CABSflex, докато за второто е нужно допълнително изследване за определяне на подходящи ограничения на разстоянията между атомите от цистиновия възел, така че да се имобилизират само нужните части на молекулата без да бъде нарушена присъщата гъвкавост на страничните вериги.

Заклучение

В настоящата дипломна работа е изследвано взаимодействието между модифицирани циклични пептиди и белтъци на SARS-CoV-2, като основно внимание е отделено на NSP13 и ORF6. От направения литературен обзор се вижда, че тези белтъци имат важна роля в развитието на вирусната инфекция и могат да бъдат използвани като мишени при търсенето на нови противовирусни средства.

В работата са представени основните свойства на цикличните пептиди и начините, по които тяхната структура може да бъде променяна. По-подробно са изследвани два модифицирани циклотида – lib9-L6-D1 и lib9-L6-D2, получени в ранките на изследванията, част от научната протема на проекта AVATAR, на основата на нативния цикличен пептид MCoTI-II. При тях са направени промени в участък 6, като основната циклична структура и цистиновият възел са запазени.

В практическата част е извършена молекулно-динамична симулация на NSP13 като верификация на параметризирането на структурата със силово поле CHARMM36m. Това не е тривиална задача, поради нетривиалното параметризиране на взаимодействията на цинковите йони с околните аминокиселинни остатъци в цинксвързания домейн на молекулата на NSP13. Резултатите показват структура на протеина, която е в добро съгласие с експериментално получената, както и подвижност, свързана с характерната за този белтък структурна гъвкавост, докладвана в литературата [19].

Молекулно-динамични симулации с продължителност 100 ns са проведени и за двата модифицирани циклотида. При lib9-L6-D1 и lib9-L6-D2 се запазва компактната структура на молекулите, като по-голяма подвижност се наблюдава главно в модифицираните участъци. От получените резултати за всеки циклотид е избрана представителна структура, която може да бъде използвана при следващия етап на изследването – молекулен докинг с NSP13.

За циклотидите lib9-L6-D1 и lib9-L6-D2 е проведен пептид-протеинов докинг с NSP13. Получените резултати позволяват да бъдат сравнени различните модели на комплексите и да бъде избран най-походящия модел на взаимодействие въз основа на използваните критерии за оценка.

Получените резултати, натрупаният опит и установеният протокол за пептид-протеинов докинг в рамките на дипломната работа имат пилотен характер и са важни за по-нататъшните изследвания от научната програма на проекта AVATAR. Установеният докинг протокол и анализът на получените структури показват, че CABSflex е подходящ за целите на проекта и може да подпомогне моделирането на модифицирани циклични пептиди с потенциален терапевтичен ефект след като бъде подобрена оценъчната процедура или добавени допълнителни ограничения върху гъвкавостта на цистиновия възел.

Литература:

- [1] Hu B., Guo H., Zhou P., Shi Z.-L. Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*. 2021;19(3):141–154. doi:10.1038/s41579-020-00459-7.
- [2] Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*. 2020;5(4):536–544. doi:10.1038/s41564-020-0695-z.
- [3] World Health Organization (WHO). WHO Director-General's opening remarks at the media briefing on COVID-19 – 11 March 2020. Geneva: World Health Organization; 2020.
Available from:
<https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
Accessed: 10 July 2026.
- [4] V'kovski P., Kratzel A., Steiner S., Stalder H., Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology*. 2021;19(3):155–170. doi:10.1038/s41579-020-00468-6.
- [5] Fosgerau K., Hoffmann T. Peptide therapeutics: current status and future directions. *Drug Discovery Today*. 2015;20(1):122–128. doi:10.1016/j.drudis.2014.10.003.
- [6] Frenkel D., Smit B. Understanding Molecular Simulation: From Algorithms to Applications. 3rd ed. London: Academic Press; 2023.
- [7] Yang H., Rao Z. Structural biology of SARS-CoV-2 and implications for therapeutic development. *Nature Reviews Microbiology*. 2021;19(11):685–700. doi:10.1038/s41579-021-00630-8.
- [8] Kumar S., Nyodu R., Maurya V. K., Saxena S. K. Morphology, Genome Organization, Replication, and Pathogenesis of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). In: Saxena S. K., editor. *Coronavirus Disease 2019 (COVID-19): Epidemiology, Pathogenesis, Diagnosis, and Therapeutics*. Singapore: Springer; 2020. p. 23–31. doi:10.1007/978-981-15-4814-7_3.

-
- [9] Brant A. C., Tian W., Majerciak V., Yang W., Zheng Z.-M. SARS-CoV-2: from its discovery to genome structure, transcription, and replication. *Cell & Bioscience*. 2021;11:136. doi:10.1186/s13578-021-00643-z.
- [10] Yuen C. K., Lam J. Y., Wong W. M., Mak L. F., Wang X., Chu H., Cai J. P., Jin D. Y., To K. K. W., Chan J. F. W., Yuen K. Y., Kok K. H. SARS-CoV-2 nsp13, nsp14, nsp15 and ORF6 function as potent interferon antagonists. *Emerging Microbes & Infections*. 2020;9(1):1418–1428. doi:10.1080/22221751.2020.1780953.
- [11] Yan W., Zheng Y., Zeng X., He B., Cheng W. Structural biology of SARS-CoV-2: open the door for novel therapies. *Signal Transduction and Targeted Therapy*. 2022;7:26. doi:10.1038/s41392-022-00884-5.
- [12] Hammond J., Leister-Tebbe H., Gardner A. et al. Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. *The New England Journal of Medicine*. 2022;386:1397–1408.
- [13] Beigel J. H., Tomashek K. M., Dodd L. E. et al. Remdesivir for the Treatment of Covid-19 – Final Report. *The New England Journal of Medicine*. 2020;383:1813–1826.
- [14] Lee J. G., Huang W., Lee H., van de Leemput J., Kane M. A., Han Z. Characterization of SARS-CoV-2 proteins reveals Orf6 pathogenicity, subcellular localization, host interactions and attenuation by Selinexor. *Cell & Bioscience*. 2021;11:58. doi:10.1186/s13578-021-00568-7.
- [15] Hall, R., Guedán, A., Yap, M. W., Young, G. R., Harvey, R., Stoye, J. P., Bishop, K. N. SARS-CoV-2 ORF6 disrupts innate immune signalling by inhibiting cellular mRNA export. *PLoS Pathogens*, 2022, 18(8), e1010349. DOI: 10.1371/journal.ppat.1010349.
- [16] Muttenthaler M., King G. F., Adams D. J., Alewood P. F. *Trends in peptide drug discovery*. *Nature Reviews Drug Discovery*. 2021.
- [17] Zorzi A., Deyle K., Heinis C. *Cyclic peptide therapeutics: past, present and future*. *Current Opinion in Chemical Biology*. 2017.
- [18] Dougherty P. G., Sahni A., Pei D. *Understanding cell penetration of cyclic peptides*. *Chemical Reviews*. 2019.
- [19] Newman J. A., Douangamath A., Yadzani S., et al. Structure, mechanism and crystallographic fragment screening of the SARS-CoV-2 NSP13 helicase. *Nature Communications*. 2021;12:4848. doi:10.1038/s41467-021-25166-6.

-
- [20] Начева Г. и съавт. Разработване и изследване на пептидни аптамери с противовирусни свойства (AVATAR). Научно описание на проект. Конкурс за финансиране на фундаментални научни изследвания – 2023 г., Фонд „Научни изследвания“, Институт по молекулярна биология „Акад. Румен Цанев“ – БАН, Софийски университет „Св. Климент Охридски“, Институт по информационни и комуникационни технологии – БАН, 2023.
- [21] Lau J. L., Dunn M. K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & Medicinal Chemistry*. 2018;26(10):2700–2707.
- [22] Kolinski A. Protein modeling and structure prediction with a reduced representation. *Acta Biochimica Polonica*. 2004;51(2):349–371. doi:10.18388/abp.2004_3575.
- [23] González M. A. Force fields and molecular dynamics simulations. *Ecole thématique de la Société Française de la Neutronique*. 2011;12:169–200.
- [24] GROMACS Development Team. GROMACS 2024.1 Manual: Molecular Dynamics Algorithms. 2024.
- [25] Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E. The Protein Data Bank. *Nucleic Acids Research*. 2000;28(1):235–242.
- [26] Leach A. R. *Molecular Modelling: Principles and Applications*. 2nd ed. Harlow: Prentice Hall; 2001.
- [27] Bussi G., Donadio D., Parrinello M. Canonical sampling through velocity rescaling. *Journal of Chemical Physics*. 2007;126(1):014101.
- [28] Parrinello M., Rahman A. Crystal structure and pair potentials: A molecular-dynamics study. *Physical Review Letters*. 1980;45(14):1196.
- [29] CABS-flex Standalone 3. Peptide–Protein Docking. Documentation. Available from: CABS-flex Standalone 3 – Peptide–Protein Docking Accessed: 7 September 2026.
- [30] Frenkel D., Smit B. *Understanding Molecular Simulation: From Algorithms to Applications*. 2nd ed. San Diego: Academic Press; 2002.
- [31] Best R. B., Zhu X., Shim J., Lopes P. E., Mittal J., Feig M., MacKerell A. D. Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the

backbone phi, psi and side-chain chi1 and chi2 torsional angles. *Journal of Chemical Theory and Computation*. 2012;8:3257–3273.

[32] Huang J., Rauscher S., Nawrocki G., Ran T., Feig M., de Groot B. L., Grubmüller H., MacKerell A. D. Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*. 2017;14:71–73.

[33] MacKerell A. D. Jr. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*. 1998;102:3586–3616.

[34] MacKerell A. D. Jr., Feig M., Brooks C. L. III. Improved treatment of the protein backbone in empirical force fields. *Journal of the American Chemical Society*. 2004;126:698–699.

[35] Abraham M. J., Murtola T., Schulz R., Páll S., Smith J. C., Hess B., Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1–2:19–25.

[36] Jorgensen W. L., Chandrasekhar J., Madura J. D., Impey R. W., Klein M. L. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*. 1983;79:926–935.

[37] Essmann U., Perera L., Berkowitz M. L., Darden T., Lee H., Pedersen L. G. A smooth particle mesh Ewald method. *Journal of Chemical Physics*. 1995;103:8577–8593.

[38] Hess B. P-LINCS: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation*. 2008;4:116–122.

[39] Kurcinski M., Jamroz M., Blaszczyk M., Kolinski A., Kmiecik S. CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. *Nucleic Acids Research*. 2015;43(W1):W419–W424. doi:10.1093/nar/gkv456.

[40] Kmiecik S., Gront D., Kolinski M., Wieteska L., Dawid A. E., Kolinski A. Coarse-Grained Protein Models and Their Applications. *Chemical Reviews*. 2016;116(14):7898–7936. doi:10.1021/acs.chemrev.6b00163.

[41] Zalewski M., Badaczewska-Dawid A., Kmiecik S. Flexible Docking of Cyclic Peptides to Proteins Using CABS-dock. *Journal of Chemical Theory and Computation*. 2025;21(17):8249–8254. doi:10.1021/acs.jctc.5c00995.

[42] Nithin C., Wróblewski K., Szukało P., Fasemire A., Kuriata A., Kurcinski M., Kolinski A., Kmiecik S. CABS-flex standalone 3: an open command-line platform for protein flexibility simulation, peptide structure modeling, and protein-peptide docking. arXiv. 2026;arXiv:2606.24487.

[43] Kurcinski M., Ciemny M. P., Oleniecki T., Kuriata A., Badaczewska-Dawid A. E., Kolinski A., Kmiecik S. Protein–peptide docking using CABS-dock and contact information. *Briefings in Bioinformatics*. 2019;20(6):2299–2305. doi:10.1093/bib/bby080.