

СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ БИОПОЛИМЕРОВ И ИХ КОМПЛЕКСОВ

УДК 577.322.23

πDMD-СИМУЛЯЦИЯ КАК СТРАТЕГИЯ ДЛЯ УЛУЧШЕНИЯ СТРУКТУР КОМПЛЕКСОВ НЕУПОРЯДОЧЕННЫХ БЕЛКОВ, СМОДЕЛИРОВАННЫХ С ПОМОЩЬЮ ALPHAFOLD2

© 2025 г. Н. Г. Мурадян^а, А. А. Саргсян^{а, б}, В. Г. Аракелов^а, А. К. Паронян^{а, б},
Г. Г. Аракелов^{а, б, *}, К. Б. Назарян^{а, б}

^аЛаборатория компьютерного моделирования биологических процессов, Институт молекулярной биологии
Национальной академии наук Республики Армения (НАН РА), Ереван, 0014 Армения

^бРоссийско-Армянский университет, Ереван, 0051 Армения

*e-mail: g_arakelov@mb.sci.am

Поступила в редакцию 10.08.2024 г.

После доработки 11.10.2024 г.

Принята к публикации 11.10.2024 г.

Неупорядоченные белки или регионы белков характеризуются отсутствием четко выраженной трехмерной конформации. Их часто называют внутренне неупорядоченными белками или областями, но при этом они играют важную роль в различных клеточных процессах и связаны с развитием ряда патологий. В *in silico* исследовании мы изучили белковые комплексы, состоящие как из упорядоченного белка 14-3-3γ, так и белков с внутренне неупорядоченными областями: белка нуклеокапсида (N) SARS-CoV-2 и p53. Показано, что комплексы, смоделированные с помощью AlphaFold2 и прошедшие этап улучшения с использованием дискретного молекулярно-динамического моделирования, приобретают определенную структуру в неупорядоченных областях. Этот процесс помогает решить одну из ключевых проблем в исследовании неупорядоченных белков – их склонность избегать стабильных конформаций.

Ключевые слова: AlphaFold2, πDMD, внутренне неупорядоченные белки, белок нуклеокапсида, p53, 14-3-3γ

DOI: 10.31857/S0026898425020095, **EDN:** GFYXKD

Внутренне неупорядоченные (неструктурированные) белки (intrinsically disordered proteins, IDP) и регионы (intrinsically disordered regions, IDR) представляют собой широко распространенную группу белков, характеризующихся отсутствием четко определенных трехмерных структур в физиологических условиях. [1, 2]. Термин “нечеткие комплексы”, или “fuzzy complexes”, относится к образованным IDP или IDR и имеет широкий спектр применимости [3–6]. При белок-белковых взаимодействиях (protein-protein interaction, PPI) IDR-содержащие белки могут принимать определенные конформации и подвергаться переходам от неструктурированной к структурированной конформации или остаются неструктурированными. Аномалии в неупорядоченных комплексах могут способствовать развитию различных заболеваний. Например,

нарушение гибкости белковых взаимодействий приводит к нарушениям сигналинга или дезрегуляции клеточных функций [7, 8].

Присутствие IDR в “нечетких комплексах” затрудняет их изучение с использованием традиционных методов структурной биологии. Такие методы, как рентгеновская кристаллография, часто не “улавливают” динамическую природу взаимодействий в этих комплексах [9, 10].

Моделирование комплексов, включающих IDR, требует использования прогрессивных вычислительных подходов, которые учитывают высокую конформационную пластичность и неструктурированную природу исследуемых объектов. Среди доступных инструментов выделяется AlphaFold2 (AF2; <https://deepmind.google/technologies/alphafold/>) — система искусственного интеллекта, специально созданная для

предсказания трехмерных структур белков на основании их аминокислотных последовательностей. AF2 предсказывает структуры с высоким разрешением, что важно для моделирования как упорядоченных, так и неупорядоченных областей белков, хотя и с некоторыми ограничениями для последних. Несмотря на указанные ограничения AF2, мы попытались использовать этот алгоритм для моделирования комплексов, которые содержат упорядоченные и неупорядоченные белки. С этой целью мы выбрали белок 14-3-3 γ в качестве упорядоченного компонента, а нуклеокапсидный (N) белок SARS-CoV-2 и белок p53 в качестве IDR-содержащих компонентов комплекса.

Белок N, взаимодействующий с геномной РНК SARS-CoV-2, образует рибонуклеопротеид, который необходим для репликации вирусной РНК и сборки вириона. Кроме того, N-белок SARS-CoV-2 облегчает репликацию и инфекционность вируса за счет его способности поддерживать разделение фаз жидкость–жидкость (liquid–liquid phase separation, LLPS), что обусловлено присутствием IDR в структуре этого белка. Состоящий из 419 а.о. N-белок SARS-CoV-2 имеет консервативные домены, включая N-концевой (NTD, 49–175 а.о.) и C-концевой (CTD, 248–365 а.о.), а также три IDR: N-концевое плечо (N-плечо, 1–48 а.о.), центральный линкерный участок (LKR, 176–247 а.о.) с Ser/Arg-богатым мотивом и C-концевой хвост (C-хвост, 366–419 а.о.). В инфицированных вирусом клетках белок 14-3-3 γ связывается с N-белком и защищает его от дефосфорилирования, тем самым влияя на его функциональную активность. В этом взаимодействии участвуют аминокислотные остатки S197 и T295 N-белка и K50, R57, K125 и R132 клеточного белка 14-3-3 γ [10–13].

Белок p53 функционирует как фактор транскрипции, участвующий в том числе в регуляции апоптоза, клеточного цикла и его остановки. Этот фактор транскрипции состоит из 393 а.о. и включает 5 доменов: трансаKTиваторный N-концевой (NTD, 1–50 а.о.), Pro-богатый (PR, 63–97 а.о.), ДНК-связывающий (DBD, 102–292 а.о.), тетрамеризующий (TD, 323–356 а.о.) и регуляторный C-концевой (CTD, 363–393 а.о.). Известно, что TD и DBD структурированы, но другие домены и линкерные области между ними неупорядочены. Белки 14-3-3 по-разному влияют на локализацию и функциональность p53 в клетке: изоформы τ и σ повышают концентрацию p53 в цитоплазме, а изоформы γ , β и ϵ усиливают связывание с ДНК (в этом взаимодействии участвуют аминокислотные остатки S366, S378 и T387 белка p53).

Важно отметить, что взаимодействие происходит в тетрамерной структуре, где два мономера 14-3-3 γ , каждый из которых состоит из 9 α -спиралей, связываются с двумя мономерами N-белка или p53. Мишенями 14-3-3 γ служат фосфорилированные формы белков-партнеров [8, 9, 14, 15].

Проведенное нами исследование было направлено на моделирование комплексов, содержащих IDR, с использованием AF2 и улучшение выбранных моделей с помощью симуляций дискретной молекулярной динамики (π DMD). Основной целью было исследование сборки IDR в комплексах.

МЕТОДЫ

Для предсказания структуры белков использовали версии v2.1.2, v2.2.3, v2.2.4 и v2.3.0, v2.3.1 AlphaFold2 [16, 17].

π DMD – метод молекулярной динамики, разработанный для моделирования макромолекулярных комплексов в биологически значимых масштабах. В отличие от традиционных методов молекулярной динамики, которые используют непрерывные потенциалы, π DMD основан на дискретных ступенчатых потенциалах как для связанных, так и несвязанных взаимодействий. Такой подход упрощает вычисления, решая баллистические уравнения движения для частиц, подвергающихся столкновениям, а не интегрируя уравнения движения Ньютона для каждой частицы в системе. Временные шаги в π DMD являются динамическими и обусловлены событиями. Событие определяется как момент, когда частица приближается к расстоянию взаимодействия с другой частицей, что обычно называют “столкновением”. В этот момент ее скорость мгновенно изменяется в соответствии с принципами сохранения энергии, импульса и углового момента. Этот подход позволяет эффективно симулировать большие биомолекулы в расширенных временных масштабах. Распараллеливание моделирования π DMD – задача сложная из-за той же событийной природы, где каждое событие зависит от предыдущих. Столкновения и события без столкновений, такие как водородные связи, связаны, что означает, что предыдущее столкновение может повлиять на время и место будущих событий. Однако эта связь ограничена скоростью, с которой распространяются эффекты. Многие прогнозы остаются верными, если атомы находятся достаточно далеко друг от друга, что минимизирует влияние ранних столкновений на последующие. Распараллеливание, связанное с событиями, в π DMD делит цикл моделирования на несколько этапов. События столкновений прогнозируются и сохраняются во временной па-

мяти до проверки, после чего действительные события фиксируются в системе. Этот метод позволяет выполнять большинство задач параллельно, в то время как только последний этап фиксации обрабатывается последовательно, что повышает производительность. Силовое поле Medusa, интегрированное в πDMD, основано на неявной модели растворителя EEF1 и методологии оценки водородных связей. Расстояние отсечки/cutoff для всех несвязанных взаимодействий составило 9.0 Å [18, 19].

Симуляции проводили в течение 2×10^6 шагов, что равно 100 нс. Симуляции с шагом 2×10^6 продолжались около 2 недель. Комплексы 14-3-3γ/N содержали 12 988 атомов, а комплексы 14-3-3γ/p53 – 12 398 атомов. Симуляцию проводили с использованием 12 ядер CPU (Central processing unit, центральное процессорное устройство) на процессоре AMD Ryzen Threadripper 3970X 32-Core Processor (“Advanced Micro Devices”, США).

Анализ кластеризации проводили с использованием программы TtClust (<https://github.com/tubiana/TtClust>), предназначенной для кластеризации траекторий молекулярной динамики. Файлы траектории использовали после πDMD. В качестве выходных данных были представлены дендрограмма в виде линейного проекционного отображения и столбчатая диаграмма[20].

Инструменты PyMol (<https://pymol.org/>) и ChmieraX [21] использовали для визуализации, а Grace (<https://plasma-gate.weizmann.ac.il/Grace/>) – для построения графиков.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Моделирование комплексов 14-3-3γ/N и 14-3-3γ/p53 выполнено с использованием AF2-версий v2.1.2, v2.2.3, v2.2.4, v2.3.0 и v2.3.1. Всего было смоделировано 1300 и 1600 структур соответственно для 14-3-3γ/N и 14-3-3γ/p53. Каждая структура прошла анализ вручную, после чего в качестве окончательных моделей были выбраны 2 структуры для 14-3-3γ/N и 2 для 14-3-3γ/p53 – на основе значений RMSD (root mean square deviation) с доступными для рентгеновского анализа доменами: NTD и CTD для N-белка и DBD и TD для p53 – как показано на рис. 1.

Из суперимпозиции выбранных комплексов со структурами доменов, полученных с помощью рентгеноструктурного анализа, и значений RMSD видно, что алгоритм AF2 точно смоделировал комплексы для консервативных областей. В случае IDR мы ожидали низкую оценку прогнозирования. В связи с отсутствием рентгеновских структур для полноразмерных молекул исследованных нами белков, а также для IRD сравнить их с предсказанными AF2 моделями по значениям RMSD не представлялось возможным. Именно поэтому мы использовали

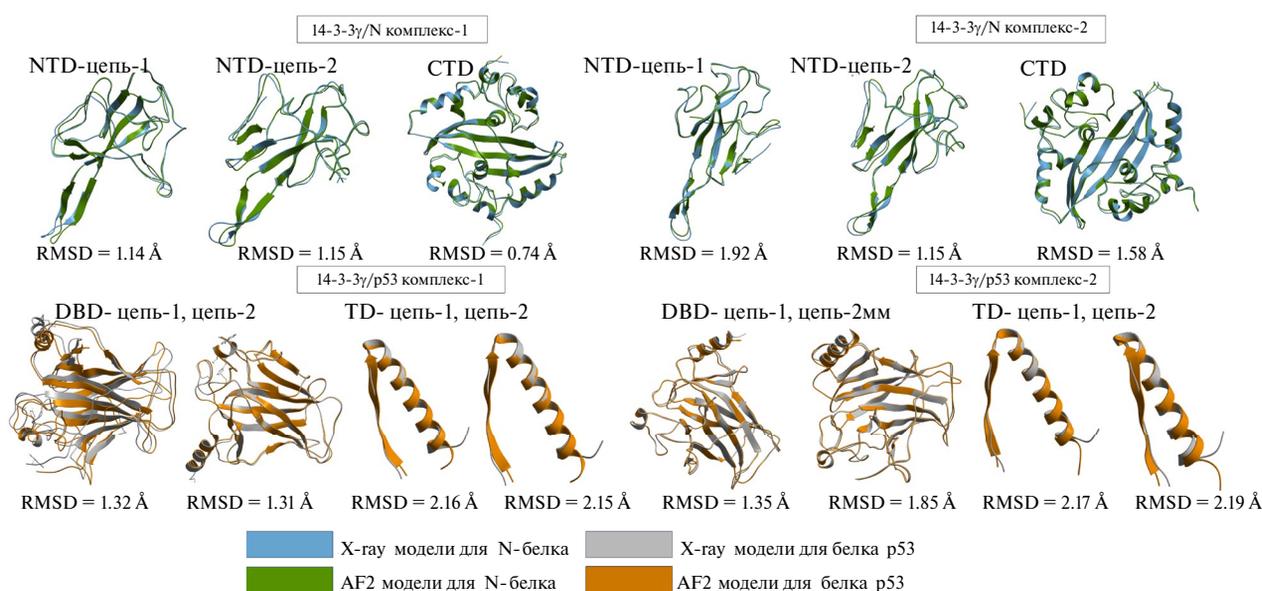


Рис. 1. Смоделированные AlphaFold2 комплексы 14-3-3γ/N и 14-3-3γ/p53. Сравнение с доступными для рентгеноструктурного анализа доменами для N-белка SARS-CoV-2 [PDB ID: 6WZQ (CTD) и 7N0R (NTD)] и белка p53 [PDB ID: 2XWR (DBD) и 1AIE (TD)]. RMSD (root mean square deviation) – среднеквадратичное отклонение; X-ray – рентгеноструктурный анализ.

оценку pLDDT (predicted Local Distance Difference Test), которую широко применяют для отображения достоверности структур, смоделированных AF2. pLDDT (<https://www.ebi.ac.uk/training/online/courses/alphafold/inputs-and-outputs/evaluating-alphafolds-predicted-structures-using-confidence-scores/plddt-understanding-local-confidence/>) оценивает уровни достоверности отдельных аминокислотных остатков в прогнозируемых структурах белка. Оценка варьирует от 0 до 100, при этом значения выше 90 указывают на высокую достоверность, а ниже 50 – на низкую. pLDDT оценивает, совпадает ли предсказанное расположение остатка с расстояниями между его атомом C α и соседними атомами C α (в пределах 15 Å) в фактической структуре белка. Регионы с pLDDT-оценками в пределах от 50 до 70 указывают на IDR в структуре белка [17]. На рис. 2 все 4 комплекса показаны с использованием оценки достоверности pLDDT.

Также мы рассчитали расстояния между сайтами связывания в белках. Наименьшие расстояния между сайтами связывания комплекса-1 14-3-3 γ /N составляют: между мономером-1 [T205 и K50] 1.57 Å, между мономером-2 [T205 и R57] 1.62 Å. В комплексе-2 расстояние между мономером-1 [T205 и K125] составляет 1.78 Å, между мономером-2 [T205 и K125] – 5.75 Å.

В комплексе-1 14-3-3 γ /p53 расстояние между мономером-1 [S378 и R57] составляет 4.41 Å, между мономером-2 [S378 и K50] – 9.03 Å, в комплексе-2 расстояние между мономером-1 [S378 и K50] равно 8.68 Å, между мономером-2 [S378 и K50] – 7.21 Å.

Симуляцию π DMD использовали для улучшения выбранных белковых моделей. Метод π DMD эффективен при изучении сборки белковых комплексов, прогнозировании конформаций в белок-белковых взаимодействиях и анализе IDR [22–24]. Учитывая эти возможности, мы использовали π DMD для исследования сборки IDR-содержащих белков в комплексах с 14-3-3 γ .

π DMD – основанный на физике метод симуляции, который использует дискретные энергетические потенциалы вместо непрерывных. Этот подход позволяет проводить симуляции биомолекулярных систем в масштабе микросекунд на стандартных персональных компьютерах, без привлечения суперкомпьютеров или специализированного оборудования. Вместо того, чтобы непрерывно вычислять силы на каждом временном шаге для обновления скоростей и положений, π DMD использует баллистическое движение и назначает временной шаг как время до следующего происходящего взаи-

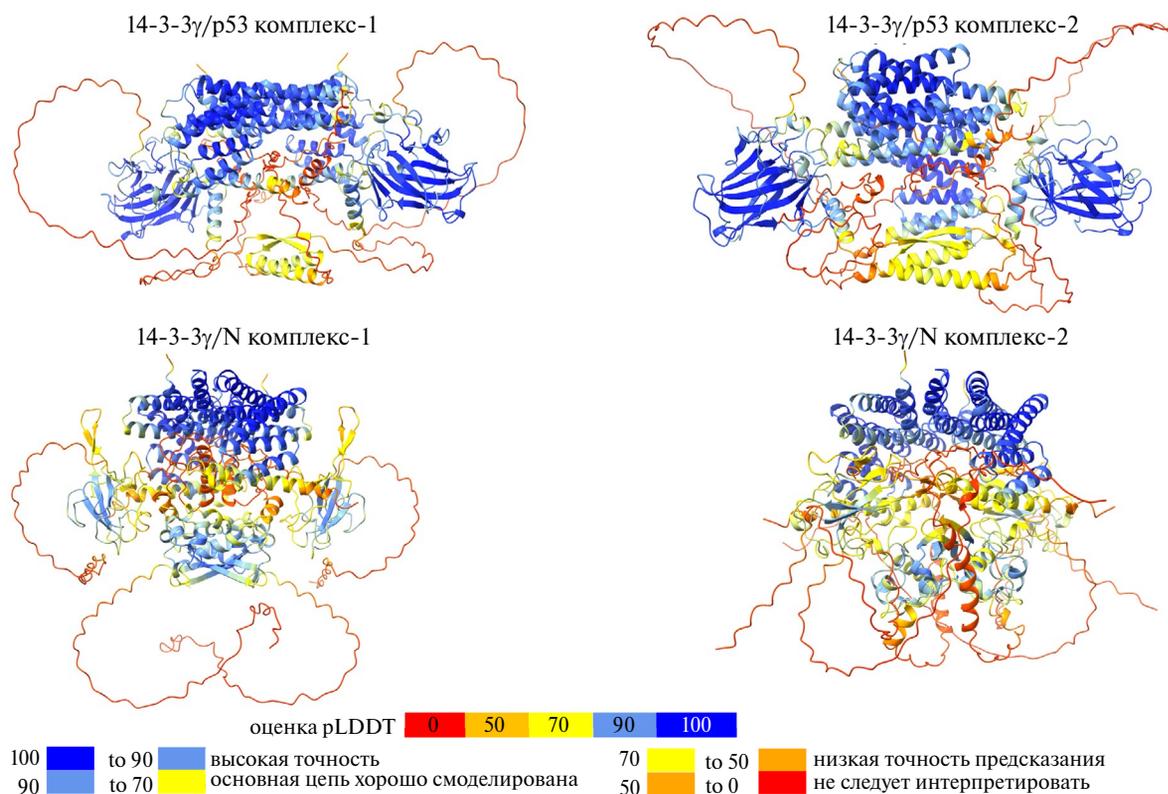


Рис. 2. Смоделированные алгоритмом AF2 окончательные структуры, окрашенные по шкале pLDDT.

модействия (“события”). Этот метод позволяет экономить время и вычислительные ресурсы и поэтому востребован при изучении сложных биомолекулярных систем [22].

Важно отметить, что πDMD использует силовое поле Medusa, которое применяется к стандартным аминокислотам и не учитывает их модифицированные формы – такие как фосфорилирование [25]. Это стало проблемой и ограничением для анализа выбранных нами комплексов, так как фосфорилирование партнера играет решающую роль во взаимодействиях белка 14-3-3γ. Учитывая этот факт, мы продолжили эксперименты по исследованию сборки IDR. Было особенно интересно выяснить, изменится ли и, если да, то как расстояние в сайтах связывания по сравнению с первоначальной структурой.

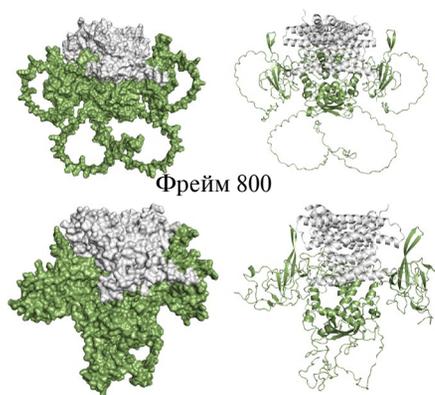
Мы провели πDMD-симуляции со всеми 4 выбранными комплексами. Система комплексов 14-3-3γ/N состояла из 12988 ато-

мов, и πDMD-симуляции были запущены для 2×10^6 шагов. Траектории записывали через каждые 100 шагов и получили 20 000 фреймов.

Во время симуляций происходили некоторые изменения в структурах белков. Визуализация траектории динамики показала, что для комплекса-1 сборка IDR началась около фреймов 750–800, тогда как для комплекса-2 – в районе фреймов 1 200–1 500. На рис. 3а, б проиллюстрирована начальная структура каждого комплекса после AF2-моделирования и выделены фреймы, где наблюдалась сборка IDR.

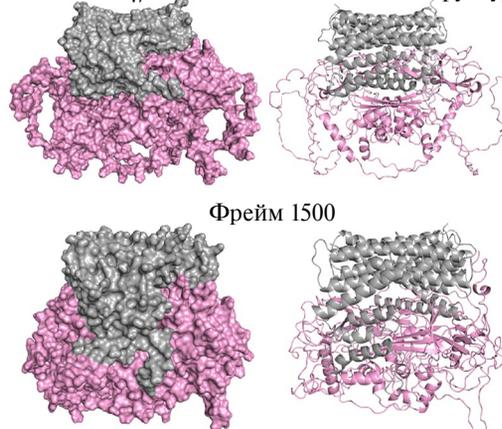
Для двух комплексов 14-3-3γ/p53 были выполнены πDMD-симуляции по 2×10^6 шагов, где каждый комплекс состоял из 12 398 атомов. Траектории симуляции записывали с интервалом в 100 шагов, в результате чего было сгенерировано 20 000 фреймов. Визуализация показала, что сборка IDR для комплекса-1 началась примерно с 200 фрейма, а для комплекса-2 – примерно с 300 фрейма. Интересно, что комп-

а 14-3-3γ/N комплекс-1 начальная структура



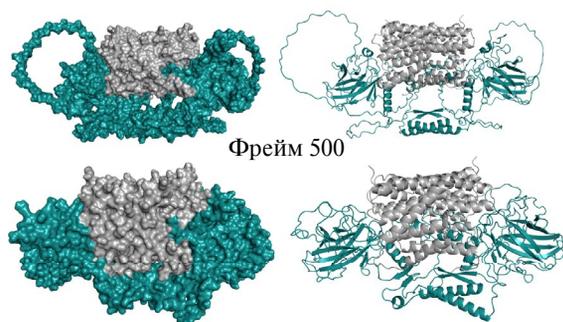
Фрейм 800

б 14-3-3γ/N комплекс-2 начальная структура



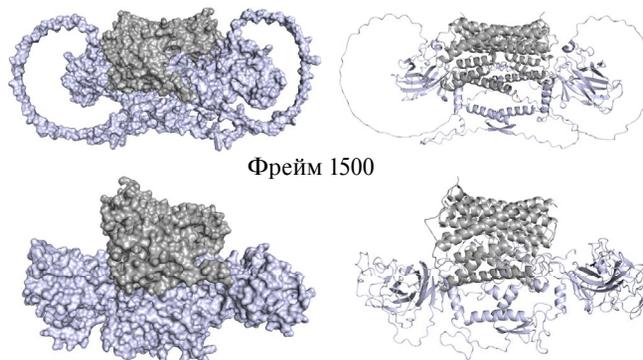
Фрейм 1500

в 14-3-3γ/p53 комплекс-1 начальная структура



Фрейм 500

г 14-3-3γ/p53 комплекс-2 начальная структура



Фрейм 1500

Рис. 3. Результаты πDMD-симуляции. а – Комплекс-1 14-3-3γ/N: начальная структура и фрейм 800; б – Комплекс-2 14-3-3γ/N: начальная структура и фрейм 1500; в – Комплекс-1 14-3-3γ/p53: начальная структура и фрейм 500; г – Комплекс-2 14-3-3γ/p53: начальная структура и фрейм 1500. Все структуры представлены сбоку.

лекс-1 был полностью собран после 500 фрейма (рис. 3*в*), а комплекс-2 — после 1500 (рис. 3*з*).

После π DMD-симуляции был проведен кластерный анализ с использованием программы TtClust [20]. Обнаружено 4 кластера для комплекса-1 14-3-3 γ /N с почти одинаковым числом конформаций, тогда как комплекс-2 образовывал 2 кластера со значительно различающимся числом конформаций (рис. 4*а*, *б*).

Из каждого кластера были выбраны репрезентативные центроиды: 4 для комплекса-1 (рис. 4*а*) и 2 для комплекса-2 (рис. 4*б*). Сборки IDR выражены в каждом центроиде, что доказывает, что комплексы оставались собранными в течение всего эксперимента. Это подтверждено при визуализации траектории.

Важный параметр, который планировали оценить, — расстояние между сайтами взаимодействия. Мы изучили расстояние между всеми потенциальными взаимодействиями с участием аминокислотных остатков в сайтах связывания для 4 комплексов, как показано на рис. 5 и 6. Как видно из представленных графических данных, расстояние между аминокислотами в сайтах связывания значительно не уменьшалось на протяжении всего эксперимента — симуляции π DMD. Хотя в некоторых фреймах расстояние между сайтами связывания иногда достигало 5 Å, среднее значение составляло ~ 20 Å для комплексов 14-3-3 γ /N и ~ 7.6 Å для комплексов 14-3-3 γ /p53.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В этом исследовании мы продемонстрировали, что алгоритм AF2 можно использовать для моделирования тетрамерных комплексов, содержащих IDR, а дальнейшее улучшение моделей проводить с помощью π DMD-симуляции. Одна из проблем, с которой мы столкнулись в ходе эксперимента, — это моделирование IDR, что связано с отсутствием определенной трехмерной структуры. Традиционные методы часто не справляются с IDR-содержащими белками из-за подвижности их структур. Таким образом, указанный подход может быть полезным и важным для изучения комплексов белков с IDR [1, 2, 26–29].

Конечно, моделирование неупорядоченных регионов было одним из ограничений для AF2, поскольку эти области обычно не имеют гомологов и имеют низкие значения rLDDT после моделирования. Вот почему было выбрано улучшение моделей с помощью дальнейшей симуляции π DMD, с основной целью — увидеть сборку IDR. Моделирование π DMD оказалось довольно удобным подходом для того, чтобы увидеть сборку неупорядоченных участков в тетрамерных комплексах белков за короткий период моделирования. Это стало возможным потому, что метод моделирования π DMD использует физику баллистического движения для генерации траекторий частиц в пространстве с течением времени в соответствии с дискретными энергетическими потенциалами на основе расстояния [22].

Выбранные нами для исследования IDR-содержащие белки: N SARS-CoV-2 и p53 — были использованы в качестве мишеней клеточного белка 14-3-3 γ , который связывается с фосфо-

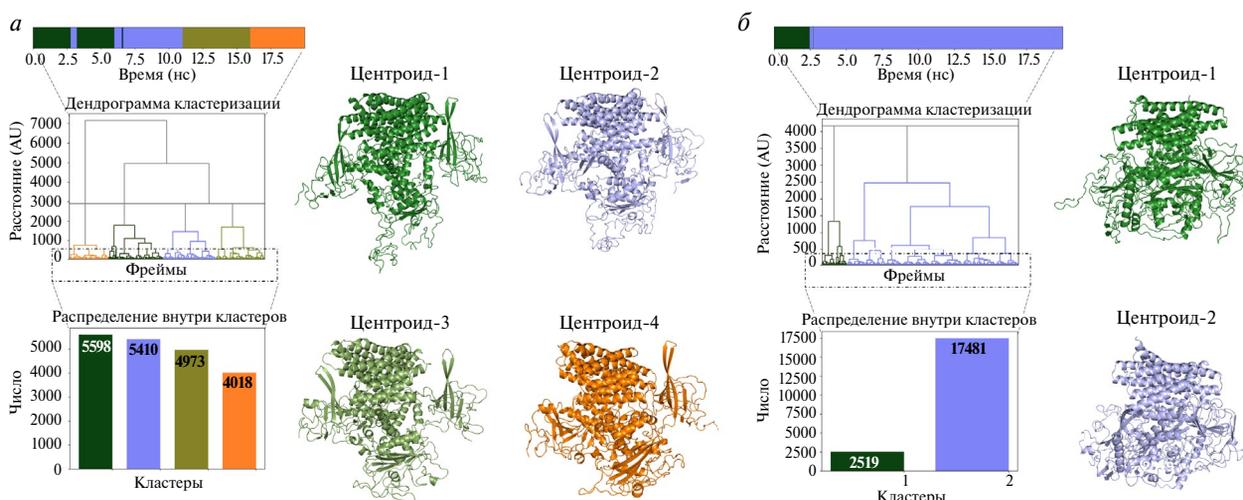


Рис. 4. Результаты кластерного анализа комплексов 14-3-3 γ /N (программа TtClust). *а* — Кластеры комплекса-1 14-3-3 γ /N с 4 репрезентативными центроидами; *б* — Кластеры комплекса-2 14-3-3 γ /N с двумя репрезентативными центроидами.

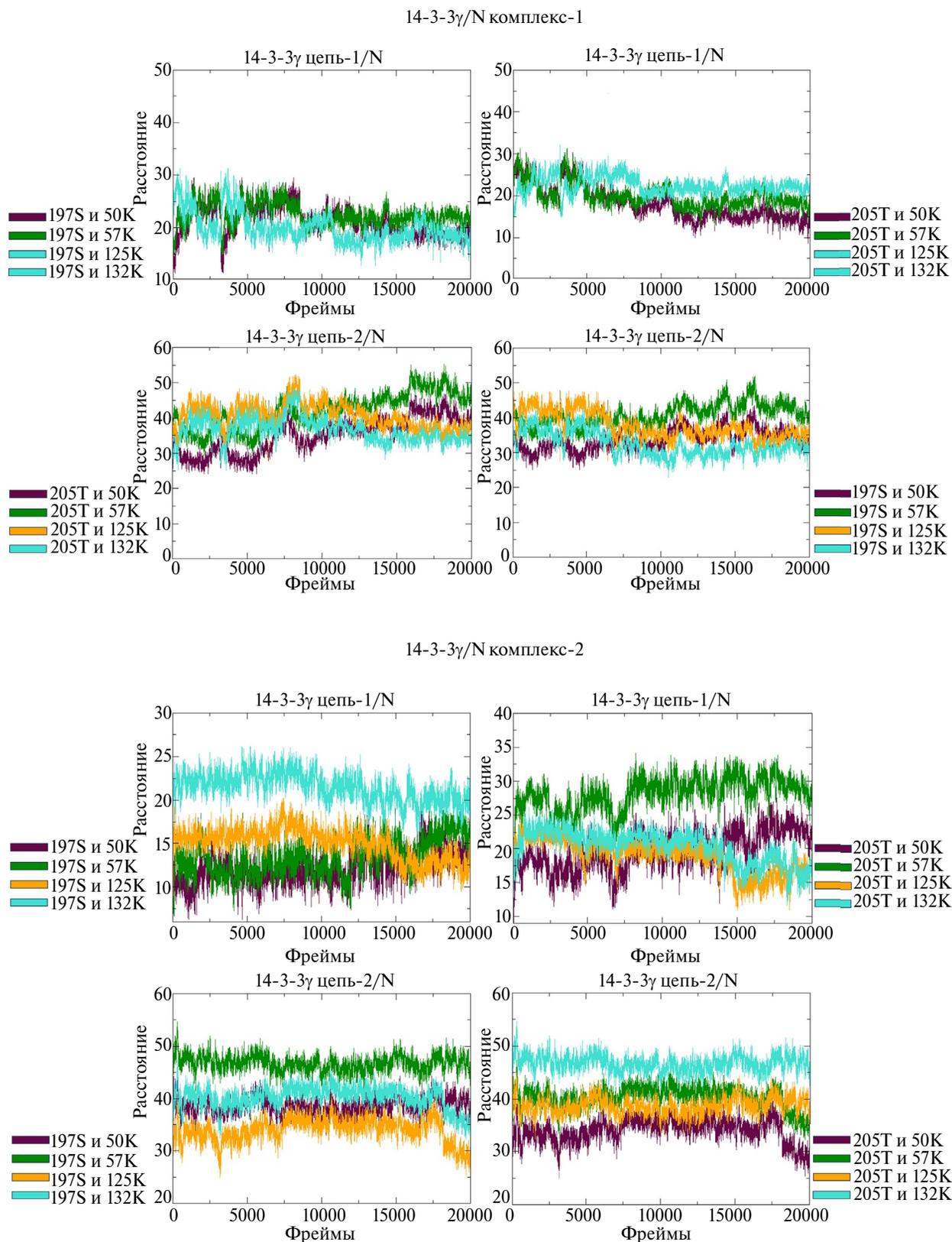
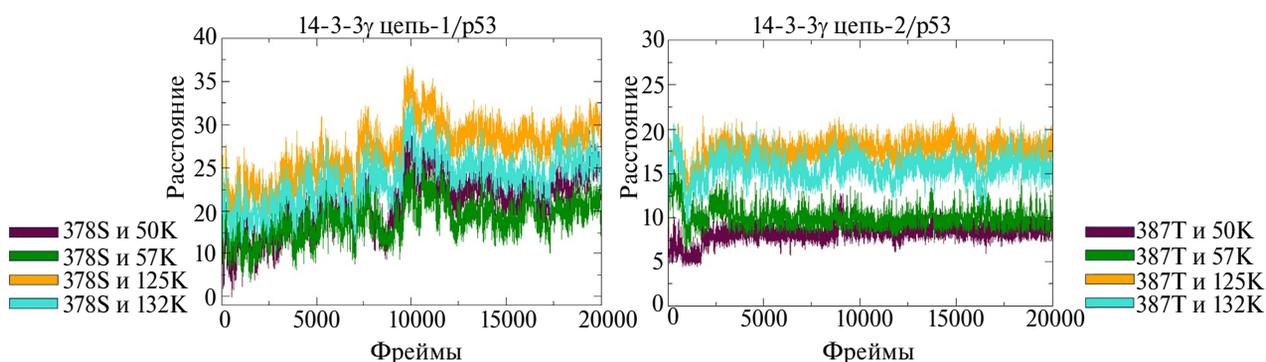
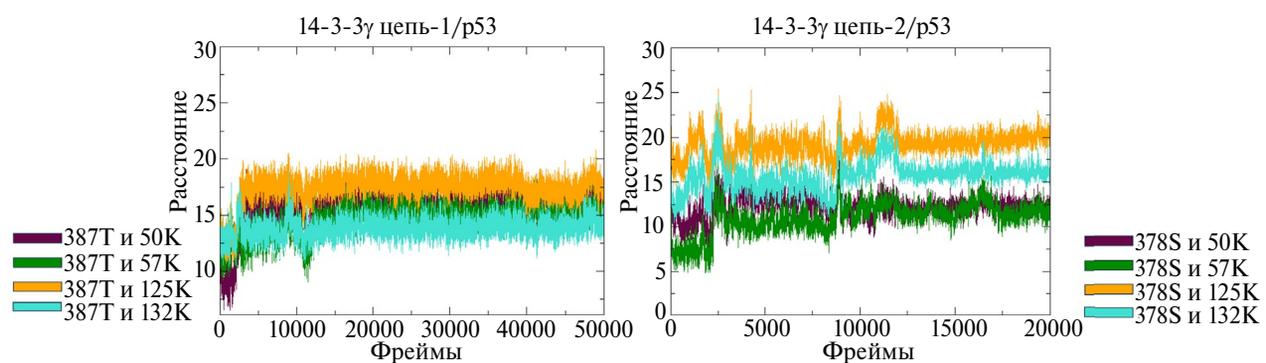


Рис. 5. Расстояния сайтов связывания для комплексов 14-3-3γ/N. Во взаимодействии участвуют S197 и T205 N-белка N и K50, R57, K125 и R132 белка 14-3-3γ. Расстояния указаны в ангстремах (Å).

14-3-3γ/p53 комплекс-1



14-3-3γ/p53 комплекс-2

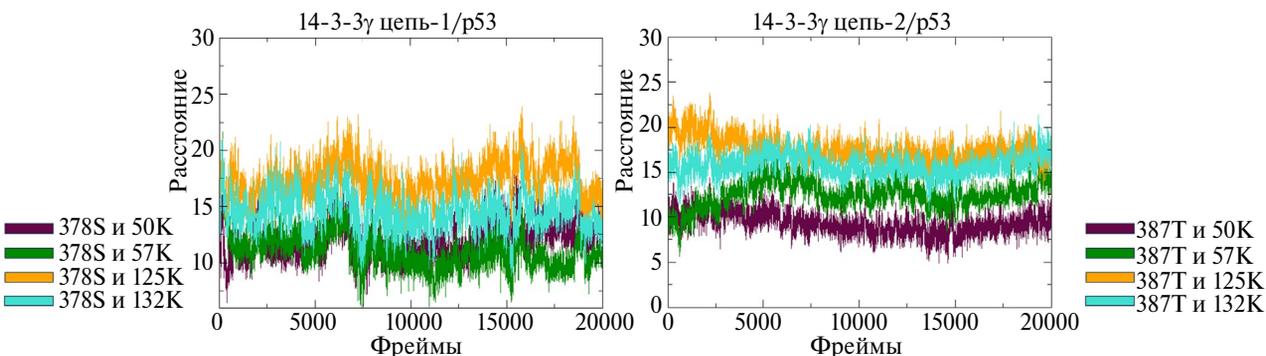
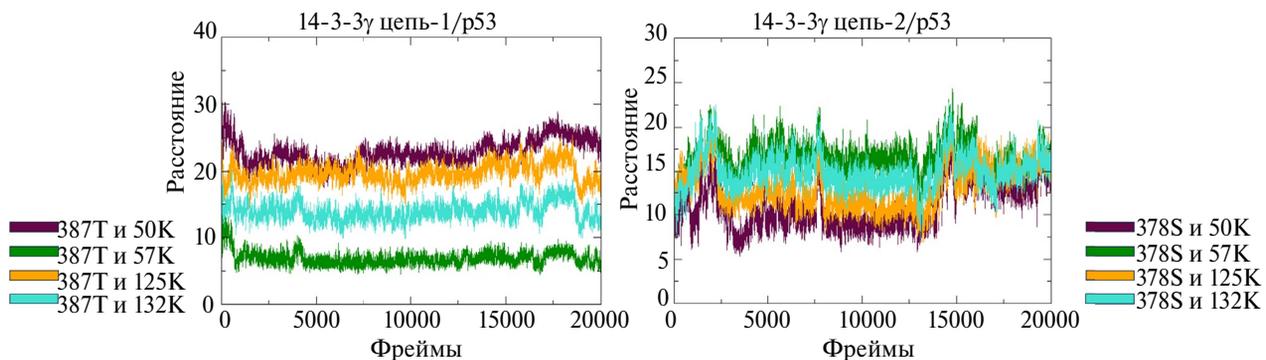


Рис. 6. Расстояния сайтов связывания для комплексов 14-3-3γ/p53. Во взаимодействии участвуют S366, S378 и T387 белка p53 и K50, R57, K125 и R132 белка 14-3-3γ. Расстояния указаны в ангстремах (Å).

рированными аминокислотными остатками в полипептидной цепи. Однако, как упоминалось выше, силовое поле Medusa, используемое πDMD, не учитывает нестандартные или модифицированные аминокислоты [25]. В качестве одного из возможных сценариев мы рассматривали замену фосфорилированных аминокислотных остатков на мутанты-фосфомиметики и предполагали, что в этом случае не будет проблем с силовым полем Medusa. Однако недавно А. Kozeleková и др. [30] показали, что фосфомиметики не заменяют фосфорилированные формы белков, поэтому мы исключили этот вариант из рассмотрения.

Тем не менее основная цель проводимого нами исследования заключалась в понимании общей динамики взаимодействия IDR-содержащих белков – в частности определить, происходит ли сборка IDR во время молекулярной симуляции. И это было подтверждено с помощью πDMD. При относительно короткой πDMD-симуляции мы наблюдали сборку комплексов 14-3-3γ/N в диапазоне фреймов 800–1500, а комплексов 14-3-3γ/p53, начиная с фрейма 500 и заканчивая 1500 (все 20 000 конформаций после 2×10^6 шагов πDMD-моделирования).

Для взаимодействия с 14-3-3γ белки N и p53 должны быть фосфорилированы в специфических сайтах связывания, а именно по S197 и T205 в N-белке и по S366, S378 и T387 в p53. Мы проанализировали расстояния между этими сайтами связывания (нефосфорилированными) по всей траектории и обнаружили, что расстояния между ними увеличивались со временем, а не уменьшались (на рис. 5 и 6 показана флуктуация расстояний между всеми возможными сайтами связывания для комплексов 14-3-3γ/N и 14-3-3γ/p53), подтверждая, что фосфорилирование аминокислот необходимо для взаимодействия с белком 14-3-3γ.

Таким образом, в результате проведенного исследования показано, что алгоритм AF2 можно использовать для моделирования комплексов, содержащих IDR, а система πDMD – для изучения сборки IDR на ранней стадии симуляции. Кроме того, получено подтверждение, что для взаимодействия 14-3-3γ с мишенями: N-белком SARS-CoV-2 и p53 – необходимо их фосфорилирование.

Ограничения исследования. Структуры, смоделированные и описанные в этом исследовании для комплексов 14-3-3γ/N и 14-3-3γ/p53, являются теоретическими. В связи с отсутствием рентгеноструктурных данных по полноразмерным структурам этих комплексов для подтверждения точности полученных моделей необходима их экспериментальная проверка.

Работа выполнена при поддержке Комитета высшего образования и науки Республики Армения в рамках научных проектов №№ 21AG-1F057, 22AA-1F026 и 22AA-1F025.

Статья не содержит исследований с использованием людей или животных.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Lotthammer J.M., Ginell G.M., Griffith D., Eme-necker R.J., Holehouse A.S. (2024) Direct prediction of intrinsically disordered protein conformational properties from sequences. *Nat. Methods*. **21**(3), 465–476. doi:10.1038/s41592-023-02159-5
2. Shrestha U.R., Smith J.C., Petridis L. (2021) Full structural ensembles of intrinsically disordered proteins from unbiased molecular dynamics simulations. *Commun. Biol.* **4**(1), 243.
3. Gong X., Zhang Y., Chen J. (2021) Advanced sampling methods for multiscale simulation of disordered proteins and dynamic interactions. *Biomolecules*. **11**, 1416.
4. Hartman, A.M., Elgaher W.A.M., Hertrich N., Andrei S.A., Ottmann C., Hirsch A.K.H. (2020) Discovery of small-molecule stabilizers of 14-3-3γ protein–protein interactions via dynamic combinatorial chemistry. *ACS Med. Chem. Lett.* **11**, 1041–1046.
5. Somsen B.A., Cossar P.J., Arkin M.R., Brunsveld L., Ottmann C. (2024) 14-3-3 protein–protein interactions: from mechanistic understanding to their small-molecule stabilization. *ChemBiochem*. **25**(14), e202400214.
6. Liu J., Cao S., Ding G., Wang B., Li Y., Zhao Y., Shao Q., Feng J., Liu S., Qin L., Xiao Y. (2021) The role of 14-3-3 proteins in cell signalling pathways and virus infection. *J. Cell Mol. Med.* **25**, 4173–4182.
7. Yang X., Lee W.H., Sobott F., Papagrigoriou E., Robinson C.V., Grossmann J.G., Sundström M., Doyle D.A., Elkins J.M. (2006) Structural basis for protein–protein interactions in the 14-3-3γ protein family. *Proc. Natl. Acad. Sci. USA*. **103**, 17237–17242.
8. Pitasse-Santos P., Hewitt-Richards I., Abeywickrama Wijewardana Sooriyaarachchi M.D., Doveston R.G. (2024) Harnessing the 14-3-3γ protein–protein interaction network. *Curr. Opin. Struct. Biol.* **86**, 102822.
9. Falcicchio M., Ward J.A., Macip S., Doveston R.G. (2020) Regulation of p53 by the 14-3-3γ protein interaction network: new opportunities for drug discovery in cancer. *Cell Death Discov.* **6**(1), 126.
10. Muradyan N., Arakelov V., Sargsyan A., Paronyan A., Arakelov G., Nazaryan K. (2024) Impact of mutations on the stability of SARS-CoV-2 nucleocapsid protein structure. *Sci. Rep.* **14**(1), 5870.
11. Cubuk J., Alston J.J., Incicco J.J., Singh S., Stuchell-Breteron M.D., Ward M.D., Zimmerman M.I.,

- Vithani N., Griffith D., Wagoner J.A., Bowman G.R., Hall K.B., Soranno A., Holehouse A.S. (2021) The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nat. Commun.* **12**(1), 1936.
12. Ni X., Han Y., Zhou R., Zhou Y., Lei J. (2023) Structural insights into ribonucleoprotein dissociation by nucleocapsid protein interacting with non-structural protein 3 in SARS-CoV-2. *Commun. Biol.* **6**(1), 193.
 13. Tugaeva K.V., Hawkins D.E.D.P., Smith J.L.R., Bayfield O.W., Ker D.S., Sysoev A.A., Klychnikov O.I., Antson A.A., Sluchanko N.N. (2021) The mechanism of SARS-CoV-2 nucleocapsid protein recognition by the human 14-3-3 γ proteins. *J. Mol. Biol.* **433**, 166875.
 14. Joerger A.C., Fersht A.R. (2008) Structural biology of the tumor suppressor p53. *Annu. Rev. Biochem.* **77**, 557–582.
 15. Rajagopalan S., Sade R.S., Townsley F.M., Fersht A.R. (2009) Mechanistic differences in the transcriptional activation of p53 by 14-3-3 γ isoforms. *Nucleic Acids Res.* **38**, 893–906.
 16. Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., Tunyasuvunakool K., Bates R., Žídek A., Potapenko A., Bridgland A., Meyer C., Kohl S.A.A., Ballard A.J., Cowie A., Romera-Paredes B., Nikolov S., Jain R., Adler J., Back T., Petersen S., Reiman D., Clancy E., Zielinski M., Steinegger M., Pacholska M., Berghammer T., Bodenstein S., Silver D., Vinyals O., Senior A.W., Kavukcuoglu K., Kohli P., Hassabis D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature.* **596**, 583–589.
 17. Evans R., O'Neill M., Pritzel A., Antropova N., Senior A., Green T., Žídek A., Bates R., Blackwell S., Yim J., Ronneberger O., Bodenstein S., Zielinski M., Bridgland A., Potapenko A., Cowie A., Tunyasuvunakool K., Jain R., Clancy E., Kohli P., Jumper J., Hassabis D. (2021) Protein complex prediction with AlphaFold-Multimer. *bioRxiv*. doi: 10.1101/2021.10.04.463034
 18. Dokholyan N.V., Buldyrev S.V., Stanley H.E., Shakhnovich E.I. (1998) Discrete molecular dynamics studies of the folding of a protein-like model. *Fold. Des.* **3**, 577–587.
 19. Proctor E.A., Ding F., Dokholyan N.V. (2011) Discrete molecular dynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 80–92.
 20. Tubiana T., Carvaille J.-C., Boulard Y., Bressanelli S. (2018) TTClust: a versatile molecular simulation trajectory clustering program with graphical summaries. *J. Chem. Inf. Model.* **58**, 2178–2182.
 21. Pettersen E.F., Goddard T.D., Huang C.C., Meng E.C., Couch G.S., Croll T.I., Morris J.H., Ferrin T.E. (2020) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82.
 22. Proctor E.A., Dokholyan N.V. (2016) Applications of discrete molecular dynamics in biology and medicine. *Curr. Opin. Struct. Biol.* **37**, 9–13.
 23. Szöllösi D., Horváth T., Han K.H., Dokholyan N.V., Tompa P., Kalmár L., Hegedűs T. (2014) Discrete molecular dynamics can predict helical prestructured motifs in disordered proteins. *PLoS One.* **9**, e95795
 24. Zamel J., Chen J., Zaer S., Harris P.D., Drori P., Lebendiker M., Kalisman N., Dokholyan N.V., Lerner E. (2023) Structural and dynamic insights into α -synuclein dimer conformations. *Structure.* **31**, 411–423.e6.
 25. Ding F., Dokholyan N.V. (2006) Emergence of protein fold families through rational design. *PLoS Comput. Biol.* **2**, e85.
 26. Kasahara K., Terazawa H., Takahashi T., Higo J. (2019) Studies on molecular dynamics of intrinsically disordered proteins and their fuzzy complexes: a mini-review. *Comput. Struct. Biotechnol. J.* **17**, 712–720.
 27. Fatafta H., Samantray S., Sayyed-Ahmad A., Coskuner-Weber O., Strodel B. (2021) Molecular simulations of IDPs: from ensemble generation to IDP interactions leading to disorder-to-order transitions. *Prog. Mol. Biol. Transl. Sci.* **183**, 135–185. doi: 10.1016/bs.pmbts.2021.06.003
 28. Tesei G., Trolle A.I., Jonsson N., Betz J., Knudsen F.E., Pesce F., Johansson K.E., Lindorff-Larsen K. (2024) Conformational ensembles of the human intrinsically disordered proteome. *Nature.* **626**, 897–904.
 29. Shrestha U.R., Smith J.C., Petridis L. (2021) Full structural ensembles of intrinsically disordered proteins from unbiased molecular dynamics simulations. *Commun. Biol.* **4**(1), 243.
 30. Kozeleková A., Náplavová A., Brom T., Gašparik N., Šimek J., Houser J., Hritz J. (2022) Phosphorylated and phosphomimicking variants may differ – a case study of 14-3-3 γ protein. *Front. Chem.* **10**, 835733.

πDMD Simulation as a Strategy for Refinement of AlphaFold2 Modeled Fuzzy Protein Complex Structures

© 2025 N. G. Muradyan¹, A. A. Sargsyan^{1, 2}, V. G. Arakelov¹, A. K. Paronyan^{1, 2},
G. G. Arakelov^{1, 2, *}, K. B. Nazaryan^{1, 2}

¹Laboratory of Computational Modeling of Biological Processes, Institute of Molecular Biology, National Academy of Sciences of the Republic of Armenia (NAS RA), Yerevan, 0014 Armenia

²Russian–Armenian University, Yerevan, 0051 Armenia

*e-mail: g_arakelov@mb.sci.am

Disordered proteins are of great interest due to their structural features, as they do not have well-defined three-dimensional structures. These proteins, often called intrinsically disordered proteins or regions, play critical roles in various cellular processes and are associated with the development of a number of diseases. Our *in silico* research focused on the investigation of protein complexes that include both the ordered protein, such as 14-3-3γ, and proteins with intrinsically disordered regions, such as nucleocapsid (N) of SARS-CoV-2 and p53. Our findings demonstrate that complexes modeled by AlphaFold2 and refined using discrete molecular dynamics simulations acquire assembled structures in disordered regions. After refinement, the modeled complexes exhibit a degree of structural assembly that addresses a key challenge in studying disordered proteins – their propensity to evade stable conformations.

Keywords: AlphaFold2, πDMD, intrinsically disorder regions, nucleocapsid protein, p53, 14-3-3γ