

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

УДК 577.322.4

Изучение конформационной подвижности липид-связывающего сайта изоформ $\epsilon 2$, $\epsilon 3$, $\epsilon 4$ аполипопротеина Е методом молекулярной динамики

А.А. Мамчур^{1, 2, *} , В.В. Ерема¹ , Д.А. Каштанова¹ , М.В. Иванов¹, В.С. Юдин¹ ,
А.А. Кескинов¹ , С.А. Краевой¹ , С.М. Юдин¹ 

¹Центр стратегического планирования и управления медико-биологическими рисками здоровью,
Федеральное медико-биологическое агентство, Россия, 119121, г. Москва, ул. Погодинская, д. 10, стр. 1;

²Кафедра биофизики, биологический факультет, Московский государственный университет имени М.В. Ломоносова,
Россия, 119234, г. Москва, Ленинские горы, д. 1, стр. 12

*e-mail: amamchur@cspfmbsa.ru

Нейродегенеративные заболевания, в особенности болезнь Альцгеймера, являются одной из наиболее острых проблем человечества. Несмотря на растущее число исследований в данной области, молекулярные механизмы развития данных заболеваний до сих пор неизвестны. Полиморфизмы в гене *APOE*, кодирующем аполипопротеин Е, достоверно ассоциированы как с развитием нейродегенеративных состояний (rs429358, C112R), так и с механизмами защиты от них (rs7412, R158C). В популяции человека присутствует три изоформы аполипопротеина Е — $\epsilon 2$ (белок с «защитной» заменой), $\epsilon 3$ (белок дикого типа), $\epsilon 4$ (белок с «патогенной» заменой). Данная работа посвящена изучению влияния описанных замен на функциональный участок белка — липид-связывающий сайт — методом молекулярной динамики. В исследовании одновременно рассмотрены все три изоформы аполипопротеина Е. Показано, что как «патогенная», так и «защитная» аминокислотные замены приводят к изменению структуры липид-связывающего сайта по сравнению со структурой дикого типа, но суть этих изменений различна. Расстояние между N-концевой (аминокислоты 88-104) и С-концевой (аминокислоты 251-266) α -спиралями липид-связывающего сайта увеличивается в обеих изоформах. Однако для изоформы $\epsilon 2$ характерно сохранение структуры С-концевой α -спирали, в то время как для изоформы $\epsilon 4$ наблюдается расплетание участка из аминокислот 260-266. Для N-концевой α -спирали наблюдается обратная ситуация. Можно предположить, что именно такие структурные различия в липид-связывающем сайте АпоЕ лежат в основе разного влияния описанных изоформ на клинические показатели носителей соответствующих генотипов. Кластерный анализ позволил выделить структуры, наиболее характерные для изоформ $\epsilon 2$ и $\epsilon 4$, которые будут в дальнейшем использованы как стартовые конформации для последующих молекулярно-динамических исследований.

Ключевые слова: молекулярная динамика, аполипопротеин Е, липид-связывающий сайт, транспорт липидов, нейродегенерация

DOI: 10.55959/MSU0137-0952-16-78-2-3

Последние десятилетия отмечается общемировая тенденция к увеличению продолжительности жизни и демографическому старению населения. Именно пожилой возраст несет с собой бремя возраст-ассоциированных инвалидизирующих состояний, одним из которых является деменция. Причиной деменции часто является нейродегенерация и конкретно болезнь Альцгеймера (БА) [1], наиболее часто развивающаяся у людей старше 65 лет и находящаяся в первой десятке причин смерти по всему миру [2]. На данный момент БА является неизлечимым заболеванием, а имеющиеся лекарственные средства предлагают поддерживающую симптоматическую терапию [1, 2]. Несмотря на

возросшее число исследований в этой области, однозначного ответа на вопросы о молекулярном механизме патогенеза, и, соответственно, патогенетической терапии нет.

Широко известной генетической ассоциацией с болезнью Альцгеймера являются полиморфизмы rs429358 и rs7412 в гене *APOE*, который расположен на 19 хромосоме и кодирует аполипопротеин Е (АпоЕ), участвующий в метаболизме липидов в организме человека [3]. АпоЕ связывается с триглицеридами и холестерином, в основном в виде липопротеинов низкой плотности и липопротеинов очень низкой плотности, обеспечивая их взаимодействие с рецепторами целевых кле-

ток [3]. В периферических тканях АпоЕ в основном синтезируется клетками печени, в центральной нервной системе — астроцитами [3]. Нарушение транспорта и катаболизма липидов предположительно является одним из механизмов развития нейродегенеративных заболеваний, так как липиды необходимы для эффективного формирования миелиновой оболочки. Полиморфизм rs429358 (C112R) формирует вариант АпоЕ ϵ 4, он ассоциирован с развитием БА и меньшей продолжительностью жизни [4]. Полиморфизм rs7412 (R158C) формирует вариант АпоЕ ϵ 2, эффект этого полиморфизма прямо противоположен эффекту rs429358 [4]. Дикий тип белка АпоЕ формирует изоформу ϵ 3.

По своей структуре АпоЕ представляет собой глобулярный белок с молекулярной массой 34 кДа и тремя основными доменами: N-концевым (аминокислоты 1-167), шарнирным (168-205) и С-концевым (206-299) [5]. В пространстве между N- и С-концевыми доменами происходит присоединение липидов [6]. Именно липид-связывающий сайт, включающий в себя аминокислоты 88-104 и 251-266 [6], стал основным объектом нашего исследования. Участок белка, состоящий из аминокислот 140-150, предположительно отвечает за связывание рецептора липопротеинов низкой плотности [6]. В связи с ассоциацией гена *APOE* с болезнью Альцгеймера были проведены исследования *in vitro* и *in silico* по изучению взаимодействий аполипопротеина Е и бета-амилоида. По результатам этих исследований за связывание бета-амилоида могут отвечать аминокислоты 133-135 и 150-161 АпоЕ [7].

Изучение структуры полноразмерного АпоЕ экспериментальными методами затруднено образованием агрегатов из-за значительной доли гидрофобных аминокислот в С-концевом домене [5]. Внесение определенных мутаций (F257A, W264R, V269A, L279Q, V287E) позволяет избежать агрегации при сохранении биологических функций [8]. Именно при помощи такого метода была получена единственная на данный момент полноразмерная экспериментальная структура АпоЕ — 2L7B в Protein Data Bank [5]. До ее появления представления о структуре АпоЕ базировались на результатах спектроскопии кругового дихроизма [9]. Предполагалось, что нарушение функций АпоЕ в изоформе ϵ 4 связано с формированием солевого мостика R61-E255, который затрудняет связывание липидов [9], однако позже эта гипотеза была опровергнута [5].

Отсутствие однозначной полноразмерной структуры затрудняет анализ молекулярных механизмов, лежащих в основе изменений функций АпоЕ, вызванных заменами C112R и R158C, актуализируя использование теоретических методов молекулярного моделирования, представленных в данной работе.

Материалы и методы

Выбор стартовых конформаций. В качестве стартовой конформации для АпоЕ дикого типа была выбрана структура 2L7B из базы данных Protein Data Bank [5]. Стартовые конформации для изоформ ϵ 2 и ϵ 4 были получены путем внесения замен R158C и C112R соответственно с помощью функции «Мутагенез» программного пакета PyMOL (Schrödinger). Модели, соответствующие изоформам ϵ 2, ϵ 3 и ϵ 4, названы АРОЕ2, АРОЕ3 и АРОЕ4 соответственно.

Молекулярная динамика (МД). Молекулярную динамику (МД) проводили с использованием пакета GROMACS [10] версии 2022.2 и полноатомного силового поля CHARMM27 [11]. Временной шаг интегрирования составил 2 фс. Белок помещался в кубическую ячейку размером 8 нм × 8 нм × 8 нм, применены периодические граничные условия. Температура 300 К и давление 1 атм поддерживались в системе с помощью алгоритмов V-rescale [12] и Parrinello-Rahman [13] соответственно. Растворитель учтен в явном виде (модель воды TIP4P [14]). Для электростатических и Ван-дер-Ваальсовых взаимодействий был установлен радиус отсечки 12 Å. Энергию электростатических взаимодействий рассчитывали с использованием алгоритма PME [15]. Для нейтрализации системы добавлялись ионы Na⁺. Перед МД-моделированием проводилась минимизация энергии (методом сопряженных градиентов, 10000 итераций) с последующим нагревом от 5 до 300 К в течение 5 нс. Для каждой модели были рассчитаны траектории длительностью 500 нс, что суммарно составляет 1,5 мкс динамики аполипопротеина Е.

Анализ данных. Подвижность полноразмерного белка была проанализирована с помощью показателей RMSD (root-mean-square deviation, среднеквадратичное отклонение атомов) и RMSF (root-mean-square-fluctuation, среднеквадратичные колебания альфа-углеродов аминокислот — атомов, которые непосредственно присоединены к карбоксильной и аминогруппам), рассчитанных библиотекой MDAnalysis (Python 3.9.12) [16].

Для прицельного анализа липид-связывающего сайта АпоЕ были выбраны участки белка, состоящие из аминокислот 88-104 и 251-266, описанные в статье Фридена и соавторов [6]. По тексту работы эти участки обозначены как N-концевая и С-концевая α -спираль липид-связывающего сайта соответственно. Для липид-связывающего сайта было дополнительно рассчитано значение RMSD. Анализ вторичных структур осуществлен с использованием функции do_dssp программного пакета GROMACS. Для анализа аминокислотных контактов использовалась библиотека MDAnalysis. Наличие контакта было определено как расположение атомов в составе двух аминокислот на расстоянии меньшем или

равном 4 Å. Программа ENCORE была использована для анализа сходства кластерных ансамблей [17]. Кластерный анализ траекторий проводился с помощью метода DBSCAN. В качестве метрики для оценки сходства кластерных ансамблей выбрана дивергенция Йенсона-Шеннона. Данные визуализированы с помощью библиотек Matplotlib и seaborn языка программирования Python (версия 3.9.12) и программного пакета PyMOL (Schrödinger).

Результаты и обсуждение

Мы рассмотрели поведение всех вариантов белка АпоЕ в растворе методом МД с оценкой среднеквадратичного отклонения атомов RMSD. Интересно отметить, что наиболее стабильным на протяжении всей симуляции оставался вариант АРОЕ3 (максимальное значение RMSD от стартовой конформации 0,7 нм), в то время как АРОЕ4 сильнее всего отклоняется от первоначальной структуры в результате динамики (максимум RMSD 1,02 нс) (рис. 1А, Приложение). Дополнительно была проанализирована подвижность не только полноразмерных белков, но и их отдельных доменов (рис. 1Б и 1В, Приложение), а также среднеквадратичные колебания RMSF альфа-углеродов отдельных аминокислот (рис. 2, Приложение). На основании этих данных мы заключили, что для всех вариантов АпоЕ наибольший вклад в отклонение от первоначальной структуры вносит С-концевой домен (RMSD 0,8–1,1 нм) (рис. 1В, Приложение), тогда как для N-концевого домена RMSD не превышает 0,7 нм (рис. 1Б, Приложение). При этом динамика RMSD для N-концевого домена АРОЕ2 практически совпадает с аналогичной динамикой для АРОЕ3, а N-концевой домен АРОЕ4 отклоняется на 0,15 нм дальше. В то же время при анализе флуктуаций альфа-углеродов аминокислот (рис. 2, Приложение) мы можем заметить, что отличия в подвижности N-концевых доменов незначительны. Тот же анализ RMSF показывает значительное повышение подвижности в шарнирном домене АРОЕ4. Также АРОЕ4 показывает большую подвижность в С-концевом домене, чем дикий тип АРОЕ3. Несмотря на сходство в динамике RMSD С-концевых доменов АРОЕ2 и АРОЕ4, пик подвижности АРОЕ4 приходится на аминокислоты 260–280, тогда как АРОЕ2 наиболее подвижен в областях 230–250 и 290–299. Таким образом, варианты АРОЕ4 и АРОЕ2 являются менее стабильными относительно дикого типа, однако максимальные колебания в этих вариантах локализованы в разных участках аминокислотной последовательности.

Наличие определенных аллелей гена *APOE* (ε2, ε3 или ε4) является одним из факторов, определяющих метаболизм липидов у человека. В связи с этим особое внимание в нашем исследовании

белка АпоЕ было уделено структуре липид-связывающего сайта, описанной группой Фридена и соавторов [6]. В его состав входят две α-спирали, расположенные параллельно в белке дикого типа – N-концевая α-спираль (аминокислоты 88–104) и С-концевая α-спираль (аминокислоты 251–266) [6]. Стабильность липид-связывающих сайтов различных изоформ АпоЕ относительно стартовой конформации также оценивалась с помощью параметра RMSD (рис. 1А). По данным анализа наиболее стабильная структура и, соответственно, наименьшее отклонение от стартовой конформации характерны для модели АРОЕ3 (рис. 1А). RMSD для нее не превышает 0,3 нм. Наибольшее отклонение от первоначальной структуры наблюдается для липид-связывающего сайта модели АРОЕ4 (рис. 1А). Для нее RMSD достигает 0,8 нм, что почти в 3 раза выше, чем для белка дикого типа. Отклонение для модели АРОЕ2 также довольно высоко и составляет в среднем 0,5 нм. При этом для всех моделей наблюдается выход RMSD на плато. Можно предположить, что каждая модель достигает своего оптимального состояния, при этом для АРОЕ4 этот оптимум наиболее удален от стартовой конформации, а значит и от АРОЕ3 (в связи с методикой получения стартовых конформаций). Дальнейшие исследования были направлены на изучение структурных особенностей описанных оптимумов.

Для описания взаимодействия N- и С-концевых α-спиралей липид-связывающего сайта АпоЕ был проведен анализ контактов между их аминокислотами. Наиболее стабильные контакты перечислены в таблице. Карты контактов представлены в приложении (рис. 3. Приложение). Для белка дикого типа характерны 12 стабильных контактов (присутствуют в более чем 70% фреймах траектории) (таблица). Некоторые из них (R92-S263, K95-E255, E96-S263) ранее были описаны исследователями, впервые получившими ЯМР-структуру АпоЕ [5] (рис. 2А). Контакты E88-K262 и T89-E266, также представленные в их статье [5], в нашей траектории присутствовали, но оказались нестабильными (длительность менее 8% времени симуляции) (рис. 2Б). Другие же контакты (T89-K262, Q98-R251), ожидаемые в области липид-связывающего сайта, после этапа минимизации энергии пропадали вовсе (рис. 2Г). Таким образом, проведенный анализ контактов, присутствующих на протяжении 500 нс МД, обновляет наши представления о состоянии белка дикого типа, так как многие из ожидаемых контактов не сохраняются в динамике. В траектории, полученной для модели АРОЕ2, был обнаружен относительно стабильный солевой мостик E96-R264 (рис. 2Г), который почти не присутствовал в траектории АРОЕ3 (рис. 2Д). В то же время в АРОЕ2 нарушены контакты, характерные для белка дикого типа. В траектории АРОЕ4 любые стабильные

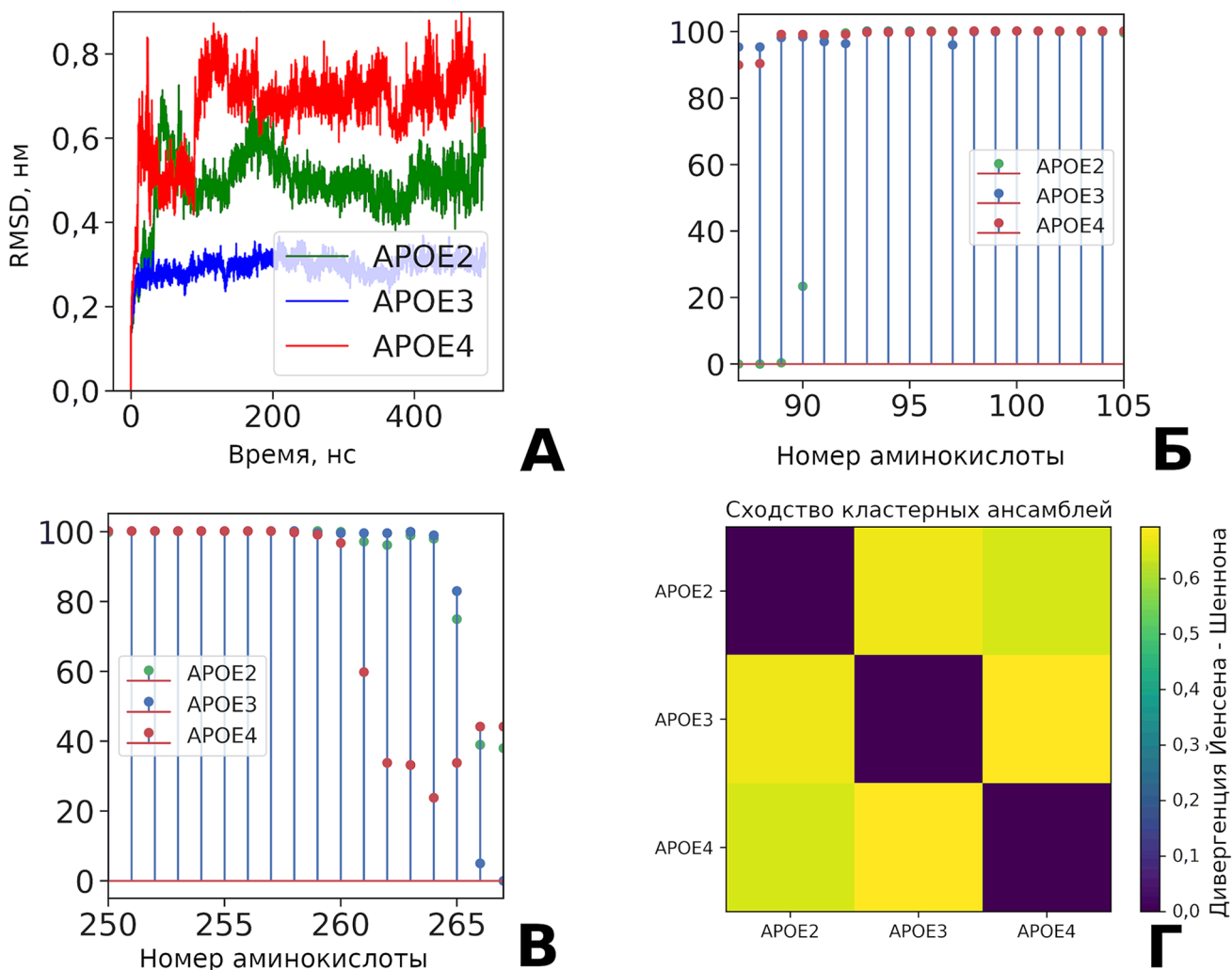


Рис. 1. Результаты анализа конформационной подвижности вариантов АпоЕ. **А** — зависимость RMSD от времени для липид-связывающих сайтов моделей АРОЕ2, АРОЕ3, АРОЕ4. **Б, В** — анализ вторичных структур в составе липид-связывающих сайтов моделей АРОЕ2, АРОЕ3, АРОЕ4. По оси абсцисс расположены номера аминокислот, по оси абсцисс указана доля (в %) времени молекулярно-динамической симуляции, в течение которого данная аминокислота вовлечена в образование α-спирали; **Г** — дивергенция Йенсена-Шеннона, отражающая сходство кластерных ансамблей изоформ АРОЕ2, АРОЕ3, АРОЕ4.

Таблица

Анализ контактов аминокислот С- и N-концевой спиралями липид-связывающего сайта				
Аминокислота N-концевой спирали	Аминокислота С-концевой спирали	Длительность контакта (в % от времени симуляции) в модели		
		АРОЕ2	АРОЕ3	АРОЕ4
R92	A259	2,6%	78,1%	0,2%
R92	K262	2,6%	71,5%	0,4%
R92	S263	7,1%	73,5%	0,7%
K95	E255	4%	96,5%	2,5%
K95	A256	0,5%	99,8%	0%
K95	A259	2,1%	96,4%	0,4%
E96	A256	1,6%	93,9%	0,02%
E96	A259	2,1%	73,1%	0,7%
E96	R260	39,1%	75,5%	3,7%
E96	S263	6,9%	97,1%	1,2%
E96	R264	64,2%	31%	0,6%
R103	L252	3,2%	78,5%	1,3%
R103	Q253	17,9%	98,1%	1,5%

контакты отсутствовали. Среди 40 найденных контактов ни один не длился дольше 4% времени молекулярно-динамической симуляции. Описанное нарушение контактов говорит о том, что для моделей АРОЕ2 и АРОЕ4 характерно большее расстояние между С- и N-концевыми α -спиралями, входящими в состав липид-связывающего сайта, чем для модели АРОЕ3.

Для уточнения характера структурных изменений, происходящих при расширении щели между С- и N-концевыми α -спиралями, входящими в состав липид-связывающего сайта, была проанализирована стабильность вторичной структуры белка. По данным, представленным на рис. 1Б, в моделях АРОЕ3 и АРОЕ4 N-концевая α -спираль (аминокислоты 88-104) остается стабильной на протяжении всей траектории. В модели АРОЕ2 наблюдается расплетение N-концевой α -спирали на участке из аминокислот 88-90 (рис. 1Б). Для изоформы $\epsilon 4$, напротив, характерно нарушение целостности С-концевой α -спирали на участке 261-265, которое отсутствует в моделях АРОЕ2 и АРОЕ3 (рис. 1В).

Понять, как именно проявляются описанные выше закономерности в трехмерной структуре, помогло сравнение кластерных ансамблей, образуемых моделями в ходе молекулярно-динамической симуляции. Дивергенция Йенсена-Шеннона, близкая к нулю, описывает абсолютную идентичность кластерных ансамблей структур. Соответственно, рис. 1Г демонстрирует, что сходство между ансамблями моделей АРОЕ3 и АРОЕ4 практически отсутствует. При этом присутствуют похожие паттерны в кластеризации моделей АРОЕ2 и АРОЕ3, а также АРОЕ2 и АРОЕ4. Для более глубокого понимания структурных особенностей моделей АРОЕ2, АРОЕ3, АРОЕ4 был проанализирован состав полученных кластерных ансамблей.

Суммарно вся система из трех траекторий содержит в себе 10 кластеров, образованных структурами, принадлежащими как одной, так и нескольким моделям. Кластерный анализ подтверждает стабильность липид-связывающего сайта изоформы $\epsilon 3$ — вся траектория модели АРОЕ3 вошла в кластер 0. Этот же кластер содер-

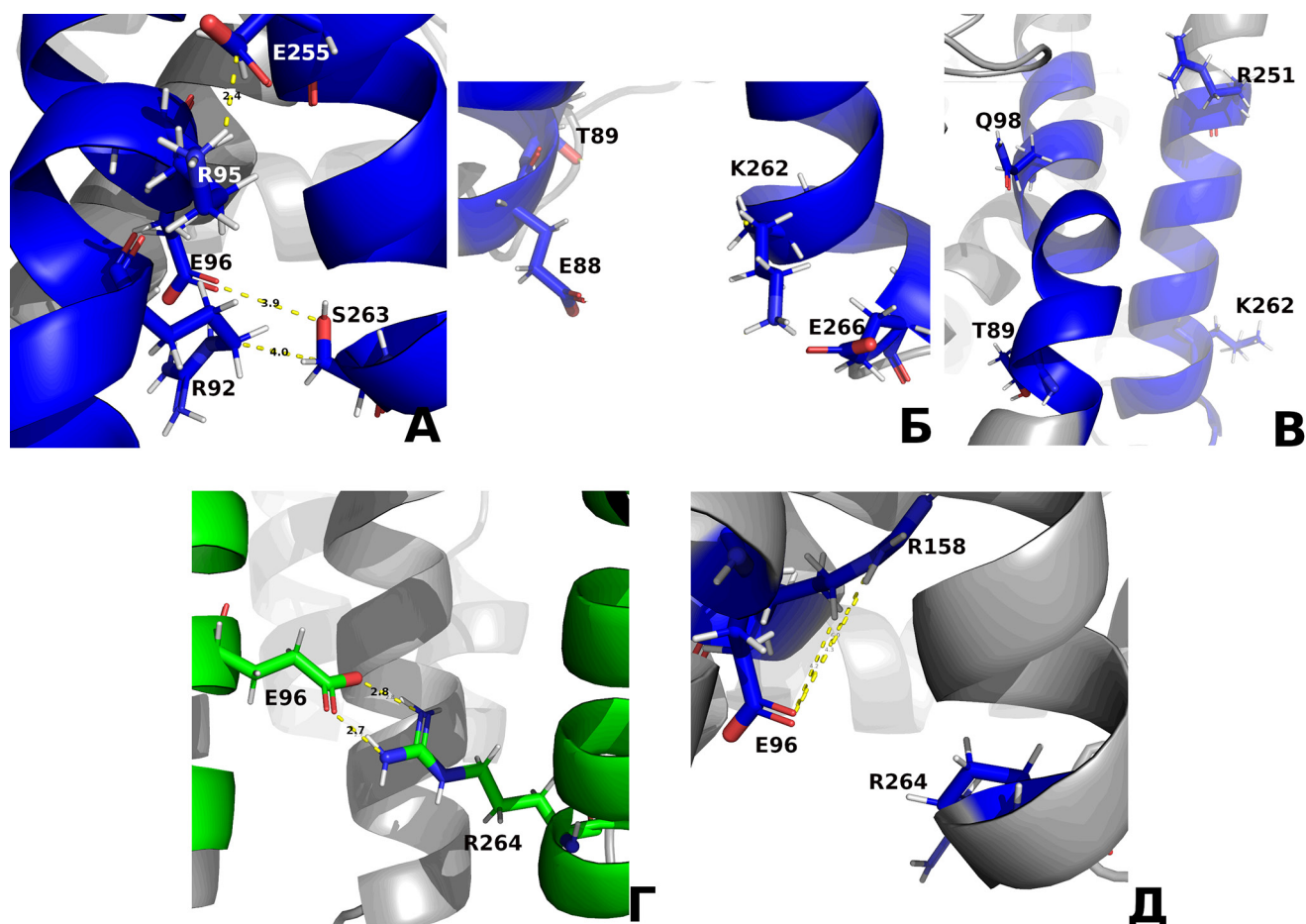


Рис. 2. Некоторые из анализируемых контактов. А — стабильные контакты в структуре липид-связывающего сайта изоформы АРОЕ3, ранее описанные авторами структуры 2L7B. Б — нестабильные контакты в структуре липид-связывающего сайта изоформы АРОЕ3, ранее описанные авторами структуры 2L7B. В — полностью отсутствующие контакты в структуре липид-связывающего сайта изоформы АРОЕ3, ранее описанные авторами структуры 2L7B. Г — присутствие солевого мостика E96-R264 в структуре липид-связывающего сайта изоформы АРОЕ2. Д — отсутствие солевого мостика E96-R264 в структуре липид-связывающего сайта изоформы АРОЕ3.

жит в себе начальные фреймы траекторий АРОЕ2 (0,9% от общего числа фреймов) и АРОЕ4 (0,2%), что ожидаемо с учетом методики формирования стартовых конформаций. Остальные фреймы траектории АРОЕ2 были распределены между кластерами 1, 2, 3, 4, 5 и 6 (2,2%, 7,4%, 2,2%, 4,8%, 81,2% и 1,4% соответственно), АРОЕ4 — между кластерами 2, 7, 8 и 9 (5,6%, 1,6%, 11,8% и 80,8% соответственно). Центроиды самых заселенных кластеров (№№ 0, 5, 9), соответствующие наиболее характерным для каждой из трех моделей структурам, представлены на рис. 3.

В течение траектории молекулярной динамики изоформы $\epsilon 2$ происходит поворот α -спиралей, входящих в состав липид-связывающего домена, в параллельных плоскостях, а затем эти плоскости удаляются друг от друга. Вероятно, этот поворот происходит вокруг солевого мостика E96-R264, описанного выше (таблица, рис. 2Д). При этом структура С-концевой α -спирали сохраняется (рис. 1В, 3А), тогда как участок N-концевой α -спирали теряет вторичную структуру (рис. 1Б, 3А). Сопоставляя информацию о структуре центроидов и стабильности вторичных структур, можно предположить, что в моделях АРОЕ3 и АРОЕ4 N-концевая α -спираль стабилизирована солевым мостиком E96-R158 (рис. 2Г). Однако при внесении замены R158C данный контакт разрушается, и E96 притягивается к R264 (рис. 2Д), тем самым дестабилизируя α -спираль.

Для изоформы $\epsilon 4$ также характерно удаление N-концевой и С-концевой α -спиралей друг от друга, однако оно происходит по другому механизму. В данном случае С-концевая α -спираль как бы «оттягивается», увеличивая размер щели и постепенно расплетая участок, состоящий из аминокислот 260-266 (рис. 1В и 3В). Такая

ситуация объяснима стабильностью N-концевой α -спирали при одновременно более мобильной С-концевой, что согласуется с ранее полученными экспериментальными данными о высокой подвижности изоформы $\epsilon 4$ [6].

О нарушении прочности взаимодействий между N- и С-концевой α -спиралями липид-связывающего сайта изоформ $\epsilon 2$ и $\epsilon 4$ свидетельствуют также энергетические характеристики (рис. 6 и 7, Приложение). Средняя энергия электростатических взаимодействий для $\epsilon 2$ и $\epsilon 4$ составляет -74,17 кДж/моль и -10,66 кДж/моль соответственно, тогда как для $\epsilon 3$ — -259,16 кДж/моль. Аналогичная ситуация наблюдается для Ван-дер-Ваальсовых взаимодействий: -15,11 кДж/моль для $\epsilon 2$, -83,23 кДж/моль для $\epsilon 3$ и -1,69 кДж/моль для $\epsilon 4$. Уменьшение модуля энергии в изоформах $\epsilon 2$ и $\epsilon 4$ вероятнее всего является следствием увеличения расстояния между описываемыми спиралями.

Необходимо отметить, что структура 2I7b из базы данных Protein Data Bank, используемая в данном исследовании как основа для построения стартовых конформаций, была получена с использованием мутантного белка АПОЕ (A257R264A269Q279E287). Вводимые замены (F257A, W264R, V269A, L279Q, V287E) делают С-концевой домен менее гидрофобным. Некоторые из них (F257A, W264R) находятся в исследуемой области белка и могут вносить существенный вклад в конформационную подвижность липид-связывающего сайта. Более того, именно аминокислота R264 образует солевой мостик с E96 N-концевой α -спирали липид-связывающего сайта АРОЕ2, который сильно влияет на динамику этой изоформы. При удалении этой аминокислоты и возвращении триптофана в положение 264 обнаруженный контакт может быть нарушен. Не-

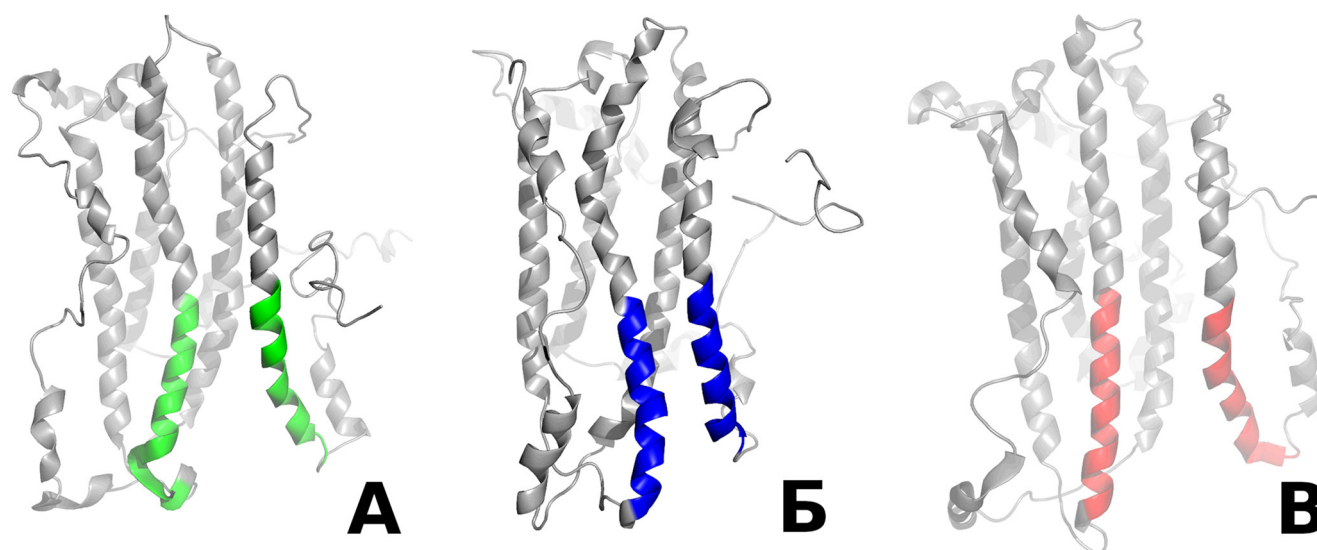


Рис. 3. Структуры липид-связывающих сайтов АРОЕ2 (А), АРОЕ3 (Б), АРОЕ4 (В), представляющие центроиды наиболее заселенных кластеров в каждой модели.

смотря на имеющиеся данные о сохранении биологической активности изоформы АРОЕ3 при введении указанных замен [8], экспериментов по их влиянию на изоформы АРОЕ2 и АРОЕ4 не проводилось. Данный факт накладывает ограничения на проведенное исследование, из-за чего мы вынуждены утверждать, что описанные выше результаты справедливы только для мутантной (A257R264A269Q279E287) формы белка.

Таким образом, изоформы $\epsilon 2$ и $\epsilon 4$ аполипопротеина Е содержат в себе измененный, по сравнению с $\epsilon 3$, липид-связывающий сайт, однако механизм этих изменений различен. Несмотря на то, что замены, определяющие изоформу белка, расположены вне липид-связывающего сайта, их внесение влияет на его структуру и динамику через сеть аминокислотных контактов. Для изоформы $\epsilon 2$ характерно сохранение структуры С-концевой α -спирали, в то время как для изоформы $\epsilon 4$ наблюдается расплетание участка из аминокислот 260–266. Расстояние между N-концевой (аминокислоты 88–104) и С-концевой (аминокислоты 251–266) α -спиралями липид-связывающего сайта увеличивается в обеих изоформах. Для N-концевой α -спирали наблюдается обратная ситуация. Можно предположить, что именно такие структурные различия в липид-связывающем сайте АпоЕ лежат в основе разного влияния описанных изоформ на клинические показатели людей-носителей. Проведенный кластерный анализ позволил выделить структуры, наиболее характерные для изоформ $\epsilon 2$ и $\epsilon 4$, которые будут в дальнейшем использованы как стартовые конформации для последующих молекулярно-динамических исследований, в том числе моделирования непосредственно комплексов АпоЕ и липидов.

Работа выполнена без привлечения финансирования. Исследование проводили без использования животных и без привлечения людей в качестве испытуемых. Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Ballard C., Gauthier S., Corbett A., Brayne C., Aarsland D., Jones E. Alzheimer's disease. *Lancet*. 2011;377(9770):1019–1031.
2. 2021 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2021;17(2):327–406.
3. Huang Y, Mahley R.W. Apolipoprotein E: structure and function in lipid metabolism, neurobiology, and Alzheimer's diseases. *Neurobiol. Dis.* 2014;72(Pt. A):3–12.
4. Deelen J., Evans D.S., Arking D.E., Tesi N., Nygaard M., Liu X., et al. A meta-analysis of genome-wide association studies identifies multiple longevity genes. *Nat. Commun.* 2019;10:3669.
5. Chen J., Li Q., Wang J. Topology of human apolipoprotein E3 uniquely regulates its diverse biological functions. *Proc. Natl. Acad. Sci. U.S.A.* 2011;108(36):14813–14818.
6. Frieden C., Wang H., Ho C.M.W. A mechanism for lipid binding to apoE and the role of intrinsically disordered regions coupled to domain–domain interactions. *Proc. Natl. Acad. Sci. U.S.A.* 2017;114(24):6292–6297.
7. Tsiolaki P.L., Katsafana A.D., Baltoumas F.A., Louros N.N., Iconomidou V.A. Hidden aggregation hotspots on human apolipoprotein E: A structural study. *Int. J. Mol. Sci.* 2019;20(9):2274.
8. Zhang Y., Vasudevan S., Sojitrawala R., Zhao W., Cui C., Xu C., Fan D., Newhouse Y., Balestra R., Jerome W.G., Weisgraber K., Li Q., Wang J. A monomeric, biologically active, full-length human apolipoprotein E. *Biochemistry*. 2007;46(37):10722–10732.
9. Hatters D.M., Peters-Libeu C.A., Weisgraber K.H. Apolipoprotein E structure: insights into function. *Trends Biochem. Sci.* 2006;31(8):445–454.
10. Abraham M.J., Murtola T., Schulz R., Pall 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.
11. Klauda J.B., Venable R.M., Freites J.A., O'Connor J.W., Tobias D.J., Mondragon-Ramirez C., Vorobyov I., MacKerell A.D., Pastor R.W. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B*. 2010;114(23):7830–7843.
12. Bussi G., Donadio D., Parrinello M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 2007;126(1):014101.
13. Parrinello M., Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* 1981;52(12):7182–7190.
14. Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R.W., Klein M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 1983;79(2):926–935.
15. Essmann U., Perera L., Berkowitz M.L., Darden T., Lee H., Pedersen L.G. A smooth particle mesh Ewald method. *J. Chem. Phys.* 1995;103(19):8577–8593.
16. Michaud-Agrawal N., Denning E.J., Woolf T.B., Beckstein O. MDAAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* 2011;32(10):2319–2327.
17. Tiberti M., Papaleo E., Bengtsen T., Boomsma W., Lindorff-Larsen K. ENCORE: Software for quantitative ensemble comparison. *PLoS Comput. Biol.* 2015;11(10):e1004415.

Поступила в редакцию 07.03.2023

После доработки 19.04.2023

Принята в печать 28.04.2023

RESEARCH ARTICLE

Molecular dynamics simulation of the conformational mobility of the lipid-binding site in the Apolipoprotein E isoforms $\epsilon 2$, $\epsilon 3$, and $\epsilon 4$

A.A. Mamchur^{1, 2, *} , V.V. Erema¹ , D.A. Kashtanova¹ , M.V. Ivanov¹, V.S. Yudin¹ ,
A.A. Keskinov¹ , S.A. Kraevoy¹ , S.M. Yudin¹ 

¹Centre for Strategic Planning and Management of Biomedical Health Risks, Federal Medical Biological Agency (Centre for Strategic Planning of FMBA of Russia), 10-1 Pogodinskaya str., Moscow, 119121, Russia;

²Department of Biophysics, School of Biology, Lomonosov Moscow State University, 1-12 Leninskie gory, Moscow, 119234, Russia

*e-mail: amamchur@cspfmba.ru

Neurodegenerative disorders, particularly Alzheimer's disease, have become a major healthcare issue all over the world. Despite extensive research, the molecular mechanisms underlying these conditions are yet to be identified. APOE is the apolipoprotein E-coding gene. Its polymorphisms have been found to be significantly associated with both neurodegenerative disorders (rs429358, C112R) and protective effects against these disorders (rs7412, R158C). Humans carry three apolipoprotein E isoforms: $\epsilon 2$ (protein with a protective mutation), $\epsilon 3$ (wild-type protein), and $\epsilon 4$ (protein with a pathogenic mutation). The study sought to investigate how these substitutions affect the lipid-binding site, which is the functional region of the protein. Molecular dynamics simulation was used to analyze all three isoforms. We found that, unlike the wild-type isoform, both the pathogenic and protective mutations caused changes in the lipid-binding site. The changes, however, were different. Both $\epsilon 2$ and $\epsilon 4$ lead to an increased distance between the N-terminal (amino acids 88-104) and the C-terminal (amino acids 251-266) helices. However, in $\epsilon 2$, the C-terminal helix retained its structure, whereas in $\epsilon 4$, it unwound between amino acids 260 and 266. The opposite was true for the N-terminal helix. It is safe to assume that it is these structural differences in the lipid-binding site that account for the different effects of these two isoforms and clinical characteristics of their carriers. The clustering analysis allowed for the identification of the structures, most typical of $\epsilon 2$ and $\epsilon 4$, which could be used as the foundation for further molecular dynamics studies.

Keywords: *molecular dynamics, Apolipoprotein E, lipid-binding site, lipid transport, neurodegeneration*

Funding: The study was performed without financial support.

Сведения об авторах

Мамчур Александра Александровна — аналитик отдела медицинской геномики ФГБУ «ЦСП» ФМБА России; аспирант кафедры биофизики биологического факультета МГУ. Тел.: 8-495-540-61-75 (доб. 4141); e-mail: amamchur@cspfmba.ru; ORCID: <https://orcid.org/0000-0002-6025-7663>

Ерема Вероника Вячеславовна — канд. мед. наук, аналитик отдела медицинской геномики ФГБУ «ЦСП» ФМБА России. Тел.: 8-495-540-61-75; e-mail: verema@cspfmba.ru; ORCID: <https://orcid.org/0000-0003-0547-3280>

Каштанова Дарья Андреевна — канд. мед. наук, вед. аналитик отдела медицинской геномики ФГБУ «ЦСП» ФМБА России. Тел.: 8-495-540-61-75; e-mail: dkashtanova@cspfmba.ru; ORCID: <https://orcid.org/0000-0001-8977-4384>

Иванов Михаил Вячеславович — аналитик 2-й категории отдела медицинской геномики ФГБУ «ЦСП» ФМБА России. Тел.: 8-495-540-61-75; e-mail: mivanov@cspfmba.ru

Юдин Владимир Сергеевич — канд. биол. наук, начальник отдела медицинской геномики ФГБУ «ЦСП» ФМБА России. Тел.: 8-495-540-61-75; e-mail: vyudin@cspfmba.ru; ORCID: <https://orcid.org/0000-0002-9199-6258>

Кескинов Антон Артурович — канд. мед. наук, канд. эконом. наук, начальник управления организации проведения научных исследований ФГБУ «ЦСП» ФМБА России. Тел.: 8-495-540-61-75; e-mail: keskinov@cspfmba.ru; ORCID: <https://orcid.org/0000-0001-7378-983X>

Краевой Сергей Александрович — первый заместитель генерального директора ФГБУ «ЦСП» ФМБА России. Тел.: 8-495-540-61-75; e-mail: skraevoy@cspfmba.ru; ORCID: <https://orcid.org/0000-0003-1775-9235>

Юдин Сергей Михайлович — докт. мед. наук, проф., ген. директор ФГБУ «ЦСП» ФМБА России. Тел.: 8-495-540-61-75; e-mail: yudin@cspfmba.ru; ORCID: <https://orcid.org/0000-0002-7942-8004>