

## ИДЕНТИФИКАЦИЯ ПРОМЕЖУТОЧНЫХ СОСТОЯНИЙ В ПРОЦЕССЕ ДИФФУЗИОННОГО СБЛИЖЕНИЯ ЭЛЕКТРОН-ТРАНСПОРТНЫХ БЕЛКОВ ПЛАСТОЦИАНИНА И ЦИТОХРОМА $f$

© 2015 г. С.С. Хрущев, А.М. Абатурова, В.А. Федоров, И.Б. Коваленко,  
Г.Ю. Ризниченко, А.Б. Рубин

*Биологический факультет Московского государственного университета им. М.В. Ломоносова,  
119991, Москва, Ленинские горы*

*E-mail: styx@biophys.msu.ru*

Поступила в редакцию 29.05.15 г.

С использованием метода броуновской динамики проведено качественное исследование взаимодействия последовательности событий, предшествующих формированию функционально активного комплекса белков пластоцианина и цитохрома  $f$ . С помощью иерархического кластерного анализа идентифицированы промежуточные состояния этого процесса. Предложен сценарий сближения этих белков, в котором ключевую роль играет диффузионный захват молекулы пластоцианина. Охарактеризована подвижность пластоцианина при разных значениях энергии электростатического взаимодействия между молекулами.

*Ключевые слова: взаимодействие белков, броуновская динамика, кластерный анализ, диффузионный захват.*

Короткоживущие (transient) белок-белковые комплексы играют важную роль в различных процессах, протекающих в живой клетке. Формирование короткоживущих комплексов происходит при активации и ингибировании ферментативной активности, в процессе рецепции и передачи информации, при окислительно-восстановительных реакциях и др. В данной работе на основе вычислительного эксперимента производится реконструкция процесса сближения двух молекул белков и проводится идентификация промежуточных состояний комплекса.

Пластоцианин и цитохром  $f$  являются классическими модельными объектами для исследования процесса формирования короткоживущих белок-белковых комплексов. Физиологическая функция пластоцианина состоит в челночном переносе электрона между цитохромным  $b_6f$ -комплексом и фотосистемой I. Передача электрона между реакционными центрами является достаточно быстрым процессом, поэтому общая скорость процесса в значительной мере определяется способностью окислительно-восстановительных реакционных партнеров быстро сформировать функционально активный

комплекс (т.е. такую конфигурацию молекул, в которой происходит перенос электрона) в одном редокс-состоянии и способностью быстро диссоциировать после прохождения реакции [1].

Взаимодействие белков пластоцианина и цитохрома  $f$  высших растений было исследовано методом броуновской динамики в работах [2–4]. Методом многочастичной броуновской динамики было исследовано взаимодействие этих белков в растворе [5] и люмене тилакоида [6], в том числе с учетом электростатических взаимодействий с тилакоидной мембраной [7]. В работе [8] дана количественная оценка вклада электростатических взаимодействий белков в их взаимную ориентацию при контакте. В работе [9] предполагается, что существенную роль в формировании функционально активного комплекса белков пластоцианина и цитохрома  $f$  играет происходящий при сближении макромолекул «диффузионный захват» [10], т.е. формирование такого состояния, при котором электростатические взаимодействия способствуют удержанию белковых молекул в тесном контакте, но в то же время не препятствуют их вращательному движению относительно друг друга.

Сокращения: CD – центральное расстояние, RMSD – среднеквадратичное отклонение положений атомов, RD – расстояние доступности.

## МАТЕРИАЛЫ И МЕТОДЫ

**Объекты исследования.** Пластоцианин является мобильным переносчиком электронов, передающим электрон с цитохромного  $b_6f$ -комплекса на фотосистему I у всех высших растений и некоторых водорослей. Цитохром  $f$  – самая крупная субъединица цитохромного  $b_6f$ -комплекса, содержащая сайт связывания пластоцианина. Окислительно-восстановительная активность пластоцианина обеспечивается наличием в его составе атома меди, меняющего степень окисления между +2 и +1. Цитохром  $f$  содержит гем с атомом железа, меняющим степень окисления между +2 и +3. Комплекс пластоцианина с цитохромом  $f$  находится в быстром равновесии с несвязанными формами белков, для взаимодействия окисленного пластоцианина из шпината *Spinacia oleracea* и восстановленного цитохрома  $f$  из турнепса *Brassica rapa* константа ассоциации комплекса  $K_a = 7 \cdot 10^3 \text{ M}^{-1}$ , константа скорости образования комплекса  $k_a = 2 \cdot 10^8 \text{ M}^{-1}\text{s}^{-1}$  [11].

В настоящей модели используется структура комплекса пластоциана шпината и цитохрома  $f$  турнепса (PDB ID: 2PCF) [1], полученная методом спектроскопии ядерного магнитного резонанса (ЯМР). Рассматривалось взаимодействие восстановленного цитохрома  $f$  и окисленного пластоцианина. Распределение электростатических зарядов на атомах белков соответствует силовому полю CHARMM27 [12,13], на геме — взято из работы [14]. Заряды на атоме меди пластоцианина и связанных с ним аминокислотных остатках аппроксимированы на основе данных работы [15]. Суммарные заряды белков в модели составляли  $-8$  (пластоцианин) и  $-4$  (цитохром  $f$ ) элементарного заряда. Структура нативного  $b_6f$ -комплекса высших растений в настоящее время неизвестна, поэтому для сравнения использовали выравнивание растворимой части цитохрома  $f$  с полной структурой  $b_6f$ -комплекса зеленой водоросли *Chlamydomonas reinhardtii* (PDB ID: 1Q90).

**Моделирование взаимодействия белков методом броуновской динамики.** Метод броуновской динамики является одним из самых простых и в то же время эффективных методов исследования динамики твердых частиц в непрерывной вязкой среде [16,17]. В основе этого метода лежит представление о термодинамическом равновесии между частицей и средой. В результате взаимодействия со средой часть энергии хаотического теплового движения молекул среды передается частице, и та совершает хаотические блуждания. Исследование количественных закономерностей броуновского дви-

жения стало важной вехой в развитии теории случайных процессов [18,19]. Метод броуновской динамики как инструмент молекулярного моделирования нашел широкое применение в биологических исследованиях благодаря работам [20–22]. Развитие метода связано с работами [23,24] (обобщение для цепочки из гидродинамически взаимодействующих сфер) и работами [5,25] (многочастичные методы, в которых рассматривается конкурентное взаимодействие множества белков с учетом интерьера клеточного компартмента). В настоящее время метод броуновской динамики широко применяется для моделирования самых разных объектов, от отдельных ионов до массивных звезд [26]. Подробный обзор метода броуновской динамики применительно к моделированию взаимодействия белков приводится в работе [9].

Для моделирования движения белков мы использовали программное обеспечение ProK-Sim (Protein Kinetics Simulator) [8]. Расчет движения макромолекул проводили с помощью уравнения Ланжевена:

$$m\ddot{r} = -\xi\dot{r} + F(r) + f(t),$$

где  $m$  – масса частицы,  $r$  – ее положение,  $t$  – время,  $\xi$  – коэффициент вязкого трения,  $F(r)$  – внешняя сила, а  $f(t)$  – стохастический член (случайная сила). При исследовании диффузии белков мы пренебрегаем быстрой динамикой системы и используем шаг модели по времени  $\Delta t$ , равный  $10^{-10}$  с, что значительно превышает время релаксации  $m/\xi$  ( $10^{-12}$  с) и позволяет пренебречь инерциальным членом (среднее значение  $(m\ddot{r})_{\Delta t} \approx 0$ ) и перейти к уравнению [24]:

$$\xi\dot{r} = F(r) + f(t).$$

Смещение частицы  $\Delta x$  вдоль координатной оси  $x$ :

$$\Delta x = \frac{F_x \Delta t}{\xi_{tr}^x} + \sqrt{\frac{2kT\Delta t}{\xi_{tr}^x}} N(0, 1),$$

где  $F_x$  – проекция внешней силы на соответствующую ось,  $\xi_{tr}^x$  – коэффициент вязкого трения для движения вдоль оси,  $k$  – постоянная Больцмана,  $T$  – температура,  $N(0, 1)$  – нормально распределенная случайная величина с нулевым средним и единичной дисперсией [24]. Аналогичным образом для вращательных степеней свободы:

$$\Delta a_x = \frac{M_x \Delta t}{\xi_{rot}^x} + \sqrt{\frac{2kT\Delta t}{\xi_{rot}^x}} N(0, 1).$$

где  $\Delta\alpha_x$ ,  $M_x$  и  $\xi_{\text{rot}}^x$  – угол поворота, момент внешней силы и коэффициент вязкого трения для вращения относительно оси, вокруг которой рассматривается вращение, соответственно. Расчет коэффициентов вязкого трения для молекул белка проводили, как описано в работе [8]. Вязкость среды принимали равной  $10^{-3}$  кгм $^{-1}$ с $^{-1}$ , что соответствует вязкости воды при 300 К.

В качестве источника внешней силы рассматривали электростатическое взаимодействие белков с учетом ионной силы раствора. При подготовке модели производили расчет электростатического потенциала вокруг каждой из макромолекул с использованием линеаризованного уравнения Пуассона–Больцмана:

$$\epsilon_0 \nabla[\epsilon(r) \nabla \varphi(r)] = -\rho + \sum_i \frac{c_i^\infty z_i^2 q_p^2 \varphi(r)}{kT} - \sum_i c_i^\infty z_i q_p,$$

где  $\epsilon_0$  – электрическая постоянная,  $\epsilon(r)$  – диэлектрическая проницаемость среды,  $\varphi(r)$  – электростатический потенциал,  $\rho$  – объемная плотность фиксированных электрических зарядов в молекуле белка,  $c_i^\infty$  – концентрация ионов в отсутствие электростатического поля,  $z_i$  – заряд иона,  $q_p$  – заряд протона,  $\Sigma_i$  обозначает суммирование по всем типам мобильных зарядов (ионов) в растворе. Решение находилось на равномерной кубической сетке с шагом 0,1 нм по итерационной формуле [27]:

$$\varphi_0 = \frac{q_0 + \sum_j h \epsilon_0 \epsilon_j \varphi_j}{h^2 \kappa_0^2 + \sum_j h \epsilon_0 \epsilon_j},$$

где  $\varphi_0$ ,  $\varphi_j$  – электростатический потенциал в ячейке,  $q_0$  – суммарный электрический заряд в ячейке (без учета мобильных ионов электролитов),  $\epsilon_j$  – диэлектрическая проницаемость среды в ячейке,  $h$  – размер ячейки,  $\kappa_0 = \sqrt{\frac{2N_A q_p^2 I}{kT}}$  – обратная дебаевская длина для ионной оболочки,  $N_A$  – число Авогадро,  $I = \frac{1}{2} \sum_i c_i^\infty z_i^2$  – ионная

сила раствора,  $\Sigma_j$  обозначает суммирование по ячейкам, контактирующим с данной ячейкой одной гранью. Электростатический потенциал в ячейках, удаленных от любого атома белка более чем на 3,5 нм, принимали равным нулю. Диэлектрическую проницаемость растворителя принимали равной 80, белка – 2. Ионная сила раствора – 100 мМ, температура – 300 К.

Моделирование движения белков проводили в виртуальной кубической ячейке размером  $30 \times 30 \times 30$  нм $^3$ , граничные условия – периодические (тороидальные), т.е. при достижении границы ячейки молекула проходит через нее, появляясь с противоположной стороны ячейки. В каждом вычислительном эксперименте по одной молекуле белков пластоцианина и цитохрома *f* помещали в ячейку в случайных положениях и ориентациях и проводили моделирование их движения до тех пор, пока энергия электростатического взаимодействия молекул не становилась меньше заданного отрицательного значения (т.е. до тех пор, пока электростатические взаимодействия не приводили к достаточно сильному взаимному притяжению молекул белков). Были проведены две серии вычислительных экспериментов, в которых эта энергия составила  $-8,2 \cdot 10^{-21}$  кДж/моль ( $-2kT$ , 20000 экспериментов) и  $-3,3 \cdot 10^{-20}$  кДж/моль ( $-8kT$ , 2200 экспериментов). Для дальнейшего анализа определяли значение модельного времени, прошедшего с начала эксперимента, взаимное расположение белков (т.е. координаты всех атомов) и точное значение величины энергии электростатического взаимодействия между ними.

**Анализ взаимного расположения макромолекул.** Анализ взаимного расположения макромолекул в ансамбле финальных конфигураций, полученных методом броуновской динамики, представляет собой нетривиальную задачу, так как даже для представления молекул в виде твердых тел каждая конфигурация описывается шестью параметрами (шесть степеней свободы). В настоящее время не существует метода наглядного представления таких ансамблей. Часто используется редукция параметрического пространства до трех измерений [28–30]. В этом случае положение центра масс одной из молекул (рассматриваемой как «лиганд») отображается в виде облака точек вокруг другой молекулы («рецептора»). Такой подход позволяет достаточно точно определить сайты связывания «лиганда» на «рецепторе», однако информация об ориентации «лиганда» полностью игнорируется.

Полноценный анализ многомерных данных возможен с помощью методов факторного анализа и кластерного анализа. В данной работе был применен однопараметрический иерархический метод классификации структур, адаптированный нами для обработки результатов молекулярного моделирования [8]. Этот метод основан на алгоритмах кластеризации по плотности (в смысле частоты встречаемости в некотором смысле «похожих» структур, density-based clustering) [31–33]. Задача иерархической

классификации в общем случае может быть сведена к некоторому упорядочению элементов множества, в данном случае – набору полученных методом броуновской динамики взаимных ориентаций двух белков, после чего в упорядоченном списке могут быть определены диапазоны номеров элементов, принадлежащих к каждому классу. Нашей целью было определить классы (ансамбли) энергетически выгодных конфигураций, в которых отдельные состояния могут переходить друг в друга. Точное решение этой задачи требует детального исследования энергетического ландшафта взаимодействия двух макромолекул, однако приближенное решение может быть получено с помощью анализа «плотности» ансамбля.

В используемом методе под кластером понимается плотная группа, отделенная от смежных групп менее плотными областями. В качестве критерия сходства двух конфигураций мы использовали среднеквадратичное отклонение координат атомов пластоцианина в одной конфигурации по сравнению с их положением в другой при «выравнивании» конфигураций совмещением атомов цитохрома *f*. Так как исследуемые конфигурации были получены методом броуновской динамики, в котором отсутствует внутримолекулярная конформационная подвижность, такое выравнивание всегда может быть проведено с идеальной точностью. В дальнейшем мы по традиции будем называть рассчитываемое значение данного параметра сходства RMSD (однако важно отметить его отличие от минимального среднеквадратичного отклонения атомов при наилучшем совмещении структур [34], которое также принято обозначать этой аббревиатурой).

Обнаружение нескольких плотных групп конфигураций, разделенных незаселенными регионами, свидетельствует о существовании нескольких кластеров (ансамблей) энергетически выгодных взаимных расположений молекул. Мы предполагаем, что переход между такими ансамблями требует преодоления энергетического либо энтропийного барьера.

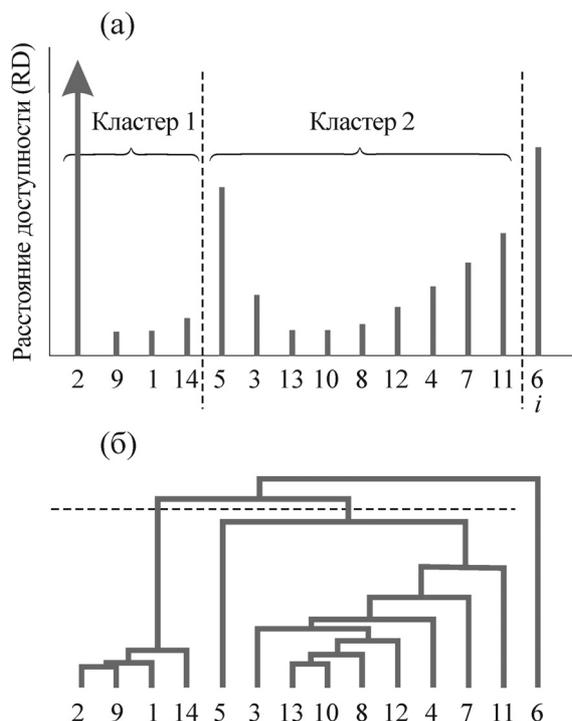
Для уменьшения влияния случайного появления похожих конфигураций в разреженных областях в качестве показателя плотности для каждой конфигурации из исходного множества вычисляется так называемое *центральное расстояние* CD (core distance). Это расстояние тем больше, чем меньше плотность конфигураций, похожих на данную. Для вычисления CD для одного взаимного расположения молекул рассчитываются значения среднеквадратичного отклонения положений атомов (RMSD) каждой конфигурации в ансамбле по сравнению с дан-

ной конфигурацией. Эти значения упорядочиваются в порядке возрастания, и в качестве CD берется *M*-е значение, где *M* – параметр метода (минимальное число конфигураций в кластере). Такие вычисления повторяются для всех взаимных расположений молекул в ансамбле. Параметр *M* в данной работе принимали равным 100.

После определения значений CD проводится упорядочивание конфигураций. На первом шаге в список добавляется произвольно выбранная конфигурация; для определенности мы берем конфигурацию с наименьшим центральным расстоянием. На каждом последующем шаге работы в список алгоритма добавляется конфигурация, которая наиболее похожа на одну из уже добавленных ранее, т.е. имеет минимальное из всех еще не добавленных конфигураций *расстояние доступности* RD (reachability distance). RD принимается равным минимальному значению RMSD между данной конфигурацией и любой конфигурацией, уже добавленной в список, если это значение больше CD данной структуры, либо равным соответствующему значению CD – в противном случае. На каждом шаге работы алгоритма RD пересчитывается для всех конфигураций, еще не добавленных в список. При добавлении конфигурации в список значение RD сохраняется для дальнейшего анализа; для первой добавленной конфигурации оно принимается равным  $+\infty$ .

На следующем этапе упорядоченный список используется для выделения кластеров конфигураций; при этом кластеры являются вложенными друг в друга и каждая конфигурация относится как минимум к одному кластеру, включающему все конфигурации в выборке (корневому кластеру). Проводится поиск локальных максимумов зависимости RD структуры от ее номера в списке. Предшествующие локальному максимуму и следующие за ним подряд группы структур, имеющие RD не более RD локального максимума, выделяются в качестве отдельных подкластеров в том случае, если количество входящих в каждую из них структур не менее *M*. Конфигурация, соответствующая локальному максимуму RD, считается принадлежащей к «правому» кластеру.

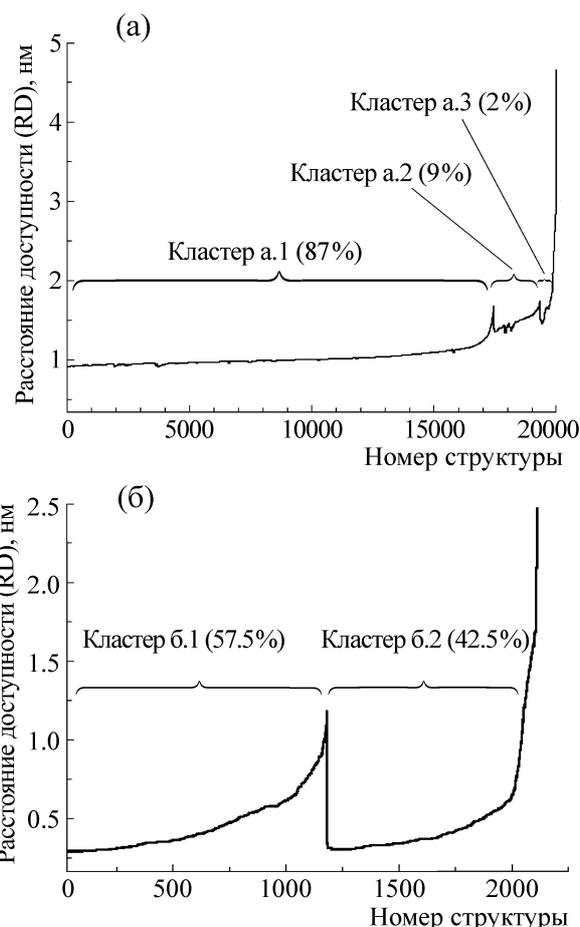
На рис. 1 показано, каким образом результат работы алгоритма может быть сопоставлен с иерархическим деревом кластеров. В представленной выборке можно выделить два кластера, первый из которых включает объекты #1, #2, #9 и #14, а второй – #3, #4, #5, #7, #8, #10, #11, #12 и #13. Объект #6 значительно отличается как от объектов первого, так и второго кластера.



**Рис. 1.** Схематическое представление результата работы алгоритма поиска кластеров. (а) – Зависимость расстояния доступности RD от номера объекта в списке  $i$ , (б) – эквивалентное представление в виде иерархического дерева.

В результате применения этого алгоритма взаимные ориентации макромолекул оказываются расклассифицированными по иерархическому дереву, каждая ветвь которого соответствует более или менее обособленной группе (ансамблю) схожих конфигураций. По графику зависимости RD от номера структуры в упорядоченном списке можно сравнить группы по частоте встречаемости (плотности) похожих конфигураций в группе и качественно оценить степень обособленности двух близких групп. Для всех отобранных ансамблей были определены их типичные конфигурации, т.е. такие взаимные расположения молекул, относительно которых среднее значение RMSD для всех входящих в кластер конфигураций минимально.

Для того чтобы оценить характер подвижности молекул белков в ансамбле, для каждого кластера по всем конфигурациям были рассчитаны значения температурного фактора Дебая–Валлера для каждого атома. Этот параметр обычно используется в рентгеноструктурном анализе и тому подобных методах в качестве характеристики тепловой подвижности атома в кристалле. Для расчета фактора Дебая–Валлера по результатам молекулярного моделирования используется формула  $B_i = 8\pi^2 \langle u_i^2 \rangle$ , где



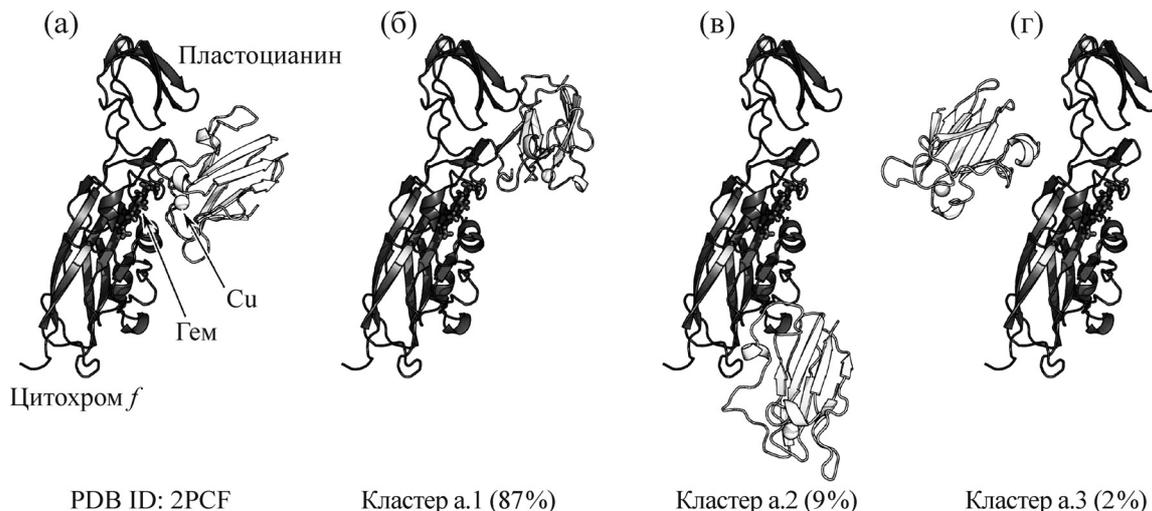
**Рис. 2.** Зависимость расстояния доступности RD от номера конфигурации в упорядоченном списке (см. пояснения в разделе «Методы»): (а) – для 20000 конфигураций с энергией электростатического взаимодействия между молекулами пластоцианина и цитохрома  $f$  менее  $-8,2 \cdot 10^{-21}$  кДж/моль ( $-2kT$ ), (б) – для 2200 конфигураций с энергией взаимодействия менее  $-3,3 \cdot 10^{-20}$  кДж/моль ( $-8kT$ ).

$B_i$  – температурный фактор для  $i$ -го атома, а  $\langle u_i^2 \rangle$  – усредненный по ансамблю квадрат отклонения положение атома от его среднего положения. Для расчета значений температурного фактора Дебая–Валлера использовали программу gmx rmsf пакета GROMACS [35].

Для визуализации структур макромолекул использовали пакет PyMol [36].

## РЕЗУЛЬТАТЫ

Методом кластерного анализа (см. рис. 2) показано, что большая часть (87%) из 20000 энергетически выгодных взаимных ориентаций молекул пластоцианина и цитохрома  $f$  с энергией электростатического притяжения между молекулами более  $2kT$  (т.е. с энергией взаимодействия меньше  $-8,2 \cdot 10^{-21}$  кДж/моль) относится



**Рис. 3.** Относительное расположение молекул белков пластоцианина и цитохрома *f*: (а) – полученный методом ЯМР функционально активный комплекс [1], (б–г) – центральные структуры кластеров (а.1), (а.2) и (а.3) (см. рис. 2а) соответственно. Структуры соориентированы по молекуле цитохрома *f*. Показаны элементы вторичной структуры белков и реакционные центры – атом меди пластоцианина и гем цитохрома *f*.

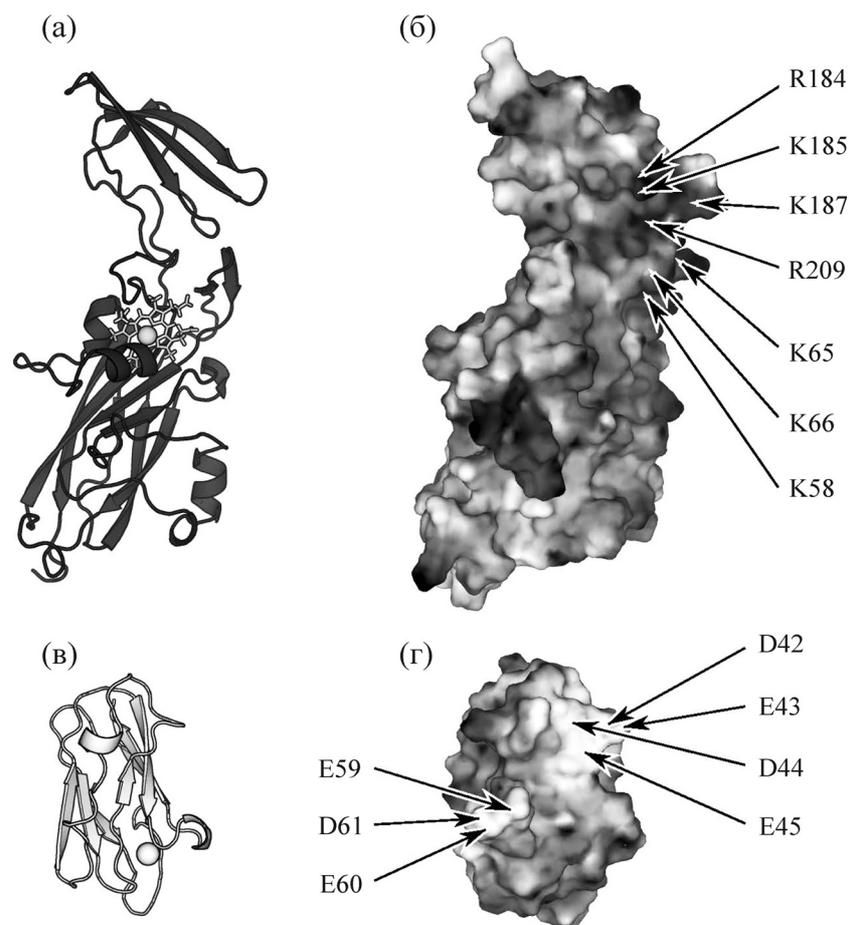
к одному равномерно-плотному ансамблю (а.1). Расположение пластоцианина относительно цитохрома *f* в целом соответствует таковому в функционально активном комплексе [1] (ср. рис. 3а и 3б), однако его ориентация может значительно варьировать: среднее среднеквадратичное отклонение положений атомов, входящих в ансамбль структур по сравнению с его центральной структурой составляет более 2,6 нм.

Мы предполагаем, что этот ансамбль структур соответствует состоянию диффузионного захвата молекулы пластоцианина цитохромом *f*. В этом состоянии электростатические взаимодействия препятствуют увеличению расстояния между поверхностями молекул, однако позволяют пластоцианину практически свободно вращаться вокруг собственного центра масс. Возможность такого вращения обеспечивается значительным общим отрицательным зарядом молекулы пластоцианина, благодаря которому его энергетически выгодное электростатическое взаимодействие с положительно заряженной областью цитохрома *f* возможно в широком диапазоне ориентаций. В то же время положительно заряженная область цитохрома *f* (аминокислоты К58, К65, К66, R184, К185, К187 и R209) окружена отрицательными зарядами (рис. 4б), что препятствует уходу пластоцианина из данной области. Это подтверждается результатами расчета температурного фактора Дебая–Валлера по всему ансамблю структур (рис. 5а): в то время как для атомов пластоцианина температурный фактор меняется в диапазоне от 1160 до 14500 Å<sup>2</sup> (среднее значение 5100 Å<sup>2</sup>), для

цитохрома *f* этот диапазон составляет от 120 до 3500 Å<sup>2</sup> (среднее значение 890 Å<sup>2</sup>).

Два минорных ансамбля энергетически выгодных взаимных ориентаций молекул (а.2 и а.3) включают в себя 9 и 2% структур. Плотность структур в этих ансамблях существенно меньше, чем в доминантном, и их разделение с ним можно четко проследить на графике расстояния доступности RD (рис. 2а). При выравнивании использованной нами структуры цитохрома *f* турнепса с полной структурой *b<sub>6</sub>f*-комплекса зеленой водоросли *Chlamydomonas reinhardtii* (PDB ID: 1Q90) видно, что участки молекулы цитохрома *f*, к которым приближается молекула пластоцианина в этих ансамблях, участвуют в связывании с другими субъединицами мультиферментного комплекса (в частности, с железосерным белком Риске), поэтому мы расцениваем эти ансамбли конфигураций как характерные только для выделенного цитохрома *f* и, следовательно, не имеющие физиологического значения.

Для более детального анализа доминантного ансамбля энергетически выгодных взаимных ориентаций молекул была исследована также выборка из 2200 структур с энергией электростатического притяжения между молекулами более  $8kT$  (т.е. с энергией взаимодействия меньше  $-3,3 \cdot 10^{-20}$  кДж/моль). На рис. 2б видно, что выборка состоит из двух близких по численности и плотности, но сильно разделенных между собой ансамблей структур. Рассчитанные значения температурного фактора Дебая–Валлера для атомов цитохрома *f* достаточно малы

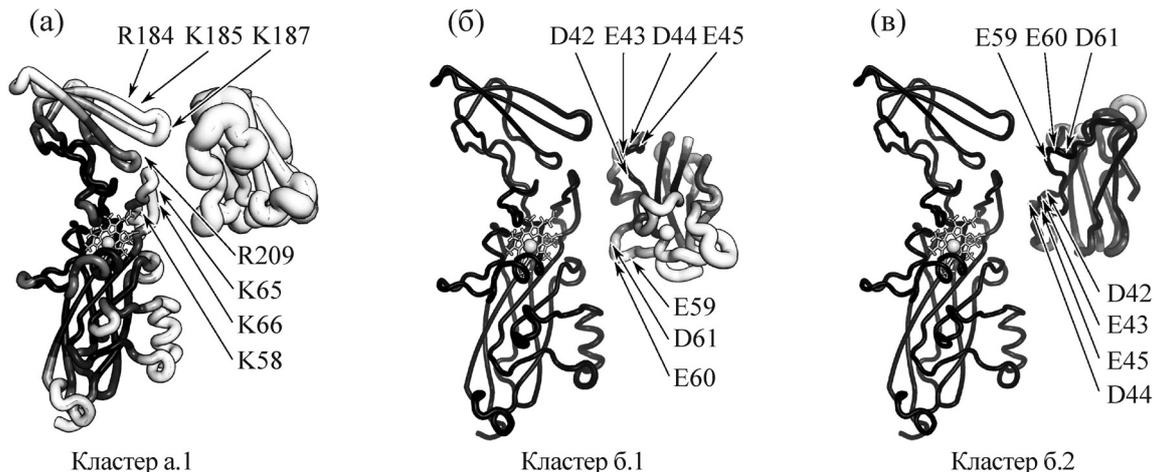


**Рис. 4.** Схематическое изображения третичной структуры белков пластоцианина (а) и цитохрома *f* (в). Показаны элементы вторичной структуры, (а) – выделен гем, в виде сферы показан атом железа, (в) – в виде сферы показан атом меди. Карта распределения электростатического заряда для этих белков (б и г соответственно) на доступной растворителю поверхности [37]. Интенсивность заполнения пропорциональна электрическому потенциалу, белый цвет соответствует значению  $-0,2$  В, черный –  $+0,2$  В.

(хотя, конечно, значительно превышают обычные для кристаллографии значения – что неудивительно, так как в данном случае оценка температурного фактора проводится для ансамбля молекул в растворе) и изменяются в диапазоне от 20 до  $700 \text{ \AA}^2$  (среднее значение  $180 \text{ \AA}^2$ ) для кластера (б.1) (обозначение по рис. 2б) и от 12 до  $480 \text{ \AA}^2$  (среднее значение  $120 \text{ \AA}^2$ ) для кластера (б.2). Однако эти два ансамбля значительно отличаются по ориентации пластоцианина относительно цитохрома *f* и по характеру его подвижности.

В первом из них, обозначенном как (б.1) (57,5% структур, рис. 5б), относительная подвижность атомов пластоцианина значительно варьирует: значения температурного фактора Дебая–Валлера находятся в диапазоне от 180 до  $3200 \text{ \AA}^2$  (среднее значение  $870 \text{ \AA}^2$ ). Наименее подвижный участок включает петли D42E43D44E45, отрицательно заряженные аминокислоты которой сближены с петлями

R184K185E186K187 цитохрома *f*. Соответственно, наибольшая подвижность наблюдается на противоположном от этой петли краю молекулы пластоцианина. Мы предполагаем, что это состояние можно охарактеризовать как следующую фазу диффузионного захвата пластоцианина цитохромом *f*: в отличие от ансамбля (а.1), где пластоцианин практически свободно вращается вокруг своего центра масс, здесь наблюдается тесная стыковка конкретных областей двух макромолекул, однако все еще сохраняется возможность их вращения вокруг области контакта. На этом основании можно предположить, что основной вклад в энергию взаимодействия вносят именно перечисленные выше аминокислоты, хотя в некоторых конфигурациях аминокислоты E59, E60 и D61 пластоцианина сближаются с аминокислотами K58, K65 и K66 цитохрома *f*. Мы предположили, что более тесное сближение этих областей должно соответствовать дальнейшему продвижению



**Рис. 5.** Относительное расположение молекул белков пластоцианина и цитохрома *f* в центральных структурах кластеров (а.1), (б.1) и (б.2) (см. рис. 2) – (а–в) соответственно. Структуры ориентированы по молекуле цитохрома *f*. Показана основная полипептидная цепь белков (толщина линии пропорциональна значению температурного фактора Дебая–Валлера в диапазоне от 600 до 15000 Å<sup>2</sup>) и реакционные центры – атом меди пластоцианина и гем цитохрома *f*. Показано расположение положительно заряженных аминокислот цитохрома *f*, участвующих в контакте с пластоцианином (а) и соответствующих им отрицательно заряженных аминокислот пластоцианина (б,в).

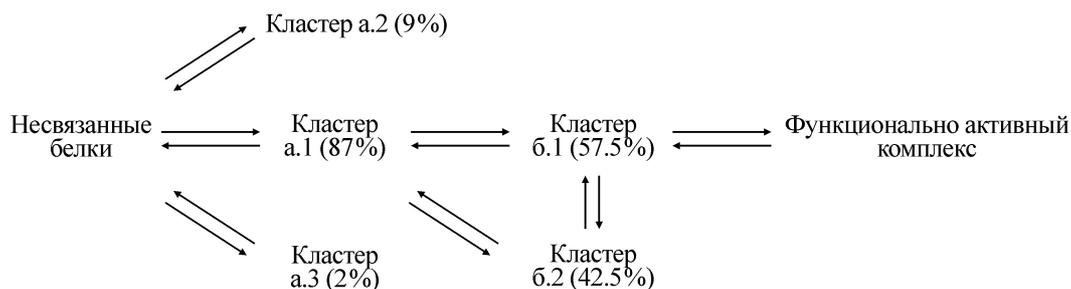
на пути к формированию функционально активного комплекса, так как при этом сближаются и активные центры макромолекул. Мы добавили конфигурации нативного комплекса (PDB ID: 2PCF) к полученной методом броуновской динамики выборке и повторили процедуру кластеризации. Структуры нативного комплекса были идентифицированы как входящие в кластер (б.1), что подтверждает предположение о том, что это состояние является промежуточным в процессе формирования функционально активного комплекса.

В другом ансамбле (б.2, 42,5% структур, рис. 5в) относительная подвижность атомов пластоцианина значительно ниже: значения температурного фактора Дебая–Валлера находятся в диапазоне от 12 до 470 Å<sup>2</sup> (среднее значение 120 Å<sup>2</sup>). Наименьшая подвижность наблюдается в двух областях: в петле E59E60D61, которая сближена с петлей R184K185E186K187 цитохрома *f* и в петле D42E43D44E45, сближенной с областью K58, K65 и K66 цитохрома *f*. Это препятствует свободному вращению белков. При этом молекула пластоцианина развернута на 180° относительно ориентации в функционально активном комплексе, и на основании полученных результатов мы полагаем, что переход конфигураций ансамбля (б.2) в состояние функционально активного комплекса возможен только после преодоления энергетического (либо энтропийного) барьера, отделяющего этот ансамбль от ансамбля (б.1).

## ОБСУЖДЕНИЕ

С помощью методов броуновской динамики и иерархического кластерного анализа были проанализированы энергетически выгодные состояния, возникающие в процессе диффузионного сближения фотосинтетических электрон-транспортных белков пластоцианина и цитохрома *f*. Полученные результаты подтверждают гипотезу, что формирование функционально активного комплекса этих белков происходит по механизму диффузионного захвата [10]. Нами было идентифицировано несколько ансамблей взаимных расположений молекул, которые, как мы предполагаем, характеризуют разные стадии диффузионного захвата.

На начальной стадии молекула пластоцианина за счет теплового движения случайно оказывается вблизи от положительно заряженной области цитохрома *f*. Вне зависимости от ориентации молекулы пластоцианина, большой суммарный отрицательный заряд молекулы приводит к ее удержанию в этой области. При этом молекула пластоцианина может практически свободно вращаться вокруг своего центра масс за счет теплового движения (ансамбль а.1). В результате этого вращения происходит сближение противоположно заряженных областей двух белков, что приводит к усилению их электростатического притяжения друг к другу и ограничению взаимной подвижности. Нами показано, что на этой стадии может произойти образование двух типов взаимных ориентаций белков.



**Рис. 6.** Гипотетическая схема переходов между состояниями на пути образования функционально активного комплекса пластоцианина и цитохрома *f* из несвязанных белков. См. пояснения в тексте.

В одном из ансамблей (б.1) петля D42E43D44E45 пластоцианина сближена с петлей R184K185E186K187 цитохрома *f*. Взаимодействие этих петель формирует гибкий шарнир, который прочно связывает две молекулы, но не препятствует их вращению вокруг области сочленения. Этот ансамбль включает в себя большое разнообразие взаимных ориентаций макромолекул, некоторые из которых близки к конфигурации функционально активного комплекса. На этом основании мы предполагаем, что путь формирования функционально активного комплекса пластоцианина и цитохрома *f* проходит через состояние (б.1).

В ансамбле (б.2) относительная подвижность макромолекул значительно ограничена и исключает возможность непосредственной трансформации этого состояния в функционально активный комплекс. Мы полагаем, что за счет теплового движения это состояние может либо медленно трансформироваться в первое сильно связанное состояние (б.1), либо вернуться в подвижное состояние (а.1).

На рис. 6 показана гипотетическая схема переходов между состояниями на пути образования функционально активного комплекса пластоцианина и цитохрома *f* из несвязанных белков.

## ЗАКЛЮЧЕНИЕ

Молекулярное моделирование событий, происходящих в процессе функционирования живых систем, позволяет глубже понять физические основы процессов жизнедеятельности. В данной работе предложен оригинальный подход к реконструкции последовательности событий, происходящих при диффузионном сближении белков, с использованием методов броуновской динамики и иерархического кластерного анализа. С использованием этого метода проведено исследование взаимодействия мобильного переносчика электрона фотосинтетической электрон-транспортной цепи пластоциа-

нина с его реакционным партнером – цитохромом *f*, экспонированной в люмен тилакоида субъединицей цитохромного *b<sub>6</sub>f*-комплекса. Применение предложенного метода позволило качественно описать последовательность событий, предшествующих формированию функционально активного комплекса этих белков. Этот же метод может быть использован и для количественной характеристики происходящих при сближении белков процессов, что будет продемонстрировано в будущих публикациях, а объединение его с методом молекулярной динамики в рамках единого комплексного подхода позволит провести полную реконструкцию процесса образования функционально активного белок-белкового комплекса.

Вычислительные эксперименты проводили с использованием суперкомпьютера «Ломоносов» в Суперкомпьютерном центре МГУ им. М.В. Ломоносова [38].

Авторы выражают благодарность доценту Т.Ю. Плюсниной за конструктивные замечания по содержанию статьи.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проекты 15-07-08927, 14-04-00302 и 15-04-08681.

## СПИСОК ЛИТЕРАТУРЫ

1. M. Ubbink, M. Ejdebeck, B.G. Karlsson, et al., *Structure* **6**, 323 (1998).
2. D. C. Pearson Jr. and E. L. Gross, *Biophys. J.* **75**, 2698 (1998).
3. F. De Rienzo, R. R. Gabdoulline, M. C. Menziani, et al., *Biophys. J.* **81** (3), 3090 (2001).
4. G. M. Ullmann, E.-W. Knapp, and N. M. Kostic, *J. Am. Chem. Soc.* **119**, 42 (1997).
5. I. B. Kovalenko, A. M. Abaturova, P. A. Gromov, et al., *Phys. Biol.* **3**, 121 (2006).
6. И. Б. Коваленко, А. М. Абатурова, П. А. Громов и др., *Биофизика* **53** (2), 261 (2008).
7. О. С. Князева, И. Б. Коваленко, А. М. Абатурова и др., *Биофизика* **55** (2), 259 (2010).

8. С. С. Хрущев, А. М. Абатурова, А. Н. Дьяконова и др., Компьютерные исследования и моделирование **5** (1), 47 (2013).
9. С. С. Хрущев, А. М. Абатурова, А. Н. Дьяконова и др., Биофизика **60** (2), 270 (2015).
10. S. H. Northrup and H. P. Erickson, Proc. Natl. Acad. Sci. USA **89** (8), 3338 (1992).
11. A. Kannt, S. Young, and D. S. Bendall, Biochim. Biophys. Acta **1277**, 115 (1996).
12. A. D. MacKerell, D. Bashford, M. Bellott, et al., J. Phys. Chem. B **102** (18), 3586 (1998).
13. A. D. Mackerell, M. Feig, and C. L. Brooks, J. Comput. Chem. **25** (11), 1400 (2004).
14. F. Autenrieth, E. Tajkhorshid, J. Baudry, et al., J. Comput. Chem. **25** (13), 1613 (2004).
15. N. V. Zolotareva, N. M. Alykov, and I. B. Kovalenko, Int. J. Appl. Eng. Res. **10** (5), 12419 (2015).
16. R. F. Fox, Phys. Rep. **48** (3), 179 (1978).
17. Ю. Л. Климонтович, Успехи физ. наук **164** (8), 811 (1994).
18. А. Эйнштейн и М. Смолуховский, *Брауновское движение* (Л.: ОНТИ, 1934).
19. N. Wiener, *Differential-space* (1923).
20. M. E. Davis, J. D. Madura, B. A. Luty, et al., Comput. Phys. Commun. **62** (2), 187 (1991).
21. J. Madura, J. Briggs, R. Wade, et al., Comput. Phys. Commun. **91** (1–3), 57 (1995).
22. S. H. Northrup, *MacroDox v.2.0.2: Software for the Prediction of Macromolecular Interaction* (Cookeville, TN: Tennessee Technological University, 1995).
23. M. Fixman, Macromolecules **19** (4), 1204 (1986).
24. D. L. Ermak and J. A. McCammon, J. Chem. Phys. **69** (4), 1353 (1978).
25. I. B. Kovalenko, A. M. Abaturova, D.M. Ustinin, et al., Biophysics (Oxf.), **52** (5), 492 (2007).
26. D. Merritt, P. Berczik, and F. Laun, Astron. J. **133** (2), 553 (2007).
27. G. M. Ullmann and E.-W. Knapp, Eur. Biophys. J. **28** (7), 533 (1999).
28. C. Tang, J. Iwahara, and G. M. Clore, Nature **444**, 7117, 383 (2006).
29. A. Spaar, C. Dammer, R. R. Gabdouliline, et al., Biophys. J. **90** (6), 1913 (2006).
30. S. Scanu, J. M. Foerster, G. M. Ullmann, et al., J. Am. Chem. Soc. **135** (20), 7681 (2013).
31. M. Ankerst, M. M. Breunig, H.-P. Kriegel, et al., Proc. ACM SIGMOD (Philadelphia, PA, 1999), pp. 49–60.
32. A. Elke, C. Böhm, and P. Kröger, in *Proc. 10th Pacific-Asian Conf. Adv. Knowl. Discov. Data Mining*. (2006), pp. 119–128.
33. J. Sander, X. Qin, Z. Lu, et al., in *Proc. 7th Pacific-Asia Conf. Knowl. Discov. Data Mining (PAKDD)* (Seoul, Korea, 2003), pp. 75–87.
34. W. Kabsch, Acta Crystallogr. Sect. A **32** (5), 922 (1976).
35. D. Van Der Spoel and B. Hess, Wiley Interdiscip. Rev. Comput. Mol. Sci. **1** (5), 710 (2011).
36. Schrödinger LLC (2012).
37. M. L. Connolly, J. Appl. Crystallogr. **16**, 548 (1983).
38. В. В. Воеводин, С. А. Жуматий, С. И. Соболев и др., Открытые системы **7** (2012).

## Identification of Intermediate States of Electron Transfer Proteins Plastocyanin and Cytochrome *f* in the Process of Diffusional Encounter

S.S. Khruschev, A.M. Abaturova, V.A. Fedorov, I.B. Kovalenko,  
G.Yu. Riznichenko, and A.B. Rubin

*Department of Biology, Lomonosov Moscow State University, Leninskie Gory 1, Moscow, 119991 Russia*

The Brownian dynamics method is used for qualitative analysis of events leading to formation of a functionally active plastocyanin–cytochrome *f* complex. Intermediate states of this process are identified by density-based hierarchical clustering. Diffusive entrapment of plastocyanin by cytochrome *f* is a key point of the suggested putative scenario of protein–protein approaching. Mobility of plastocyanin is characterized for different values of protein–protein electrostatic interaction energy.

*Key words: protein-protein interactions, brownian dynamics, cluster analysis, diffusive entrapment*