

УДК: 577.2; 51-76

Моделирование белок-белковых взаимодействий с применением программного комплекса многочастичной броуновской динамики *ProKSim*

С.С.Хрущев^а, А.М.Абатурова, А.Н.Дьяконова, Д.М.Устинин,
Д.В.Зленко, В.А.Федоров, И.Б.Коваленко^б, Г.Ю.Ризниченко, А.Б.Рубин

Московский государственный университет имени М. В. Ломоносова,
Биологический факультет, кафедра биофизики,
Россия, 119992, г. Москва, Ленинские горы, 1-12

E-mail: ^а styx@biophys.msu.ru, ^б ikovalenko78@gmail.com

Получено 5 декабря 2012 г.,
после доработки 20 февраля 2013 г.

Белок-белковые взаимодействия являются основой большинства биологических процессов. Компьютерное моделирование динамики связывания белков дает важную информацию для понимания механизмов их функционирования. Разработана компьютерная программа *ProKSim* (**P**rotein **K**inetics **S**imulator), предназначенная для моделирования взаимодействия макромолекул методом многочастичной броуновской динамики с учетом дальнедействующих электростатических взаимодействий. Проведено исследование диффузионно-столкновительных комплексов для трех пар белков: ферредоксин и ферредоксин:НАДФ⁺-редуктаза, пластоцианин и цитохром *f*, барназа и барстар. Исследована роль электростатических взаимодействий во взаимной ориентации молекул белков при образовании диффузионно-столкновительных комплексов.

Ключевые слова: многочастичная броуновская динамика, белок-белковые взаимодействия, механизмы молекулярного распознавания

Multi-particle Brownian Dynamics software *ProKSim* for protein-protein interactions modeling

S. S. Khrushchev, A. M. Abaturva, A. N. Diakonova, D. M. Ustinin,
D. V. Zlenko, V. A. Fedorov, I. B. Kovalenko, G. Yu. Riznichenko, A. B. Rubin

Lomonosov Moscow State University, School of Biology, Biophysics Department
Leninskie Gory 1-12, 119992, Moscow, Russia

Abstract. — Protein-protein interactions are of central importance for virtually every process in living matter. Modeling the dynamics of protein association is crucial for understanding their functionality. This paper proposes novel simulation software *ProKSim* (**P**rotein **K**inetics **S**imulator) for modeling of protein interactions by means of the multi-particle Brownian Dynamics. Effect of long-range electrostatic interactions on the process of transient encounter complex formation is numerically estimated. Investigation of transient encounter complex formation was performed for three pairs of proteins: ferredoxin and ferredoxin:NADP⁺-reductase, plastocyanin and cytochrome *f*, barnase and barstar.

Keywords: multi-particle Brownian Dynamics, protein-protein recognition, molecular recognition

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 47–64 (Russian).

Работа поддержана грантами РФФИ 11-04-01268-а, 12-07-33036-мол_а_вед и 12-04-31839-мол_а. Работа выполнена с использованием ресурсов суперкомпьютерного комплекса МГУ имени М.В. Ломоносова [Воеводин и др., 2012].

© 2012 Сергей Сергеевич Хрущев, Анна Михайловна Абатурова, Александра Никитична Дьяконова, Дмитрий Михайлович Устинин, Дмитрий Владимирович Зленко, Владимир Андреевич Федоров, Илья Борисович Коваленко, Галина Юрьевна Ризниченко, Андрей Борисович Рубин

1. Введение

Белок–белковые взаимодействия лежат в основе многих физиологических процессов, связанных с ферментативной активностью и ее регуляцией, электронным транспортом и др. Предсказание структуры белковых комплексов и скорости их образования путем компьютерного моделирования представляет собой сложную, до сих пор не решенную до конца проблему. Причиной этого является, в первую очередь, очень высокая вычислительная ресурсоемкость методов моделирования «из первых принципов», поэтому их применение на нынешнем этапе возможно лишь для самых простых систем (десятки атомов). Для более крупных систем приходится использовать те или иные эмпирические или полуэмпирические приближения, имеющие ограниченную область применимости. К числу наиболее часто используемых при исследовании взаимодействия биомакромолекул методов относятся молекулярный докинг и молекулярная динамика.

Процесс образования комплекса двух белковых молекул в растворе можно условно разделить на несколько стадий:

- 1) свободная диффузия молекул в растворе на большом расстоянии от других макромолекул,
- 2) сближение макромолекул и их взаимная ориентация за счет дальнедействующих электростатических взаимодействий с образованием предварительного (диффузионно-столкновительного) комплекса и
- 3) трансформация предварительного комплекса в финальный, т. е. в такую конфигурацию, в которой осуществляется биологическая функция.

Альтернативно диффузионно-столкновительный комплекс может распасться без образования финального комплекса. При трансформации предварительного комплекса в финальный происходят вытеснение молекул растворителя из белок-белкового интерфейса и конформационные изменения самих макромолекул. Важную роль в этом процессе играют гидрофобные взаимодействия и образование водородных связей и солевых мостиков. Различные авторы вкладывают разный смысл в понятие диффузионно-столкновительного (*encounter*) комплекса; обзор нескольких популярных точек зрения приведен в [Gabdoulline, Wade, 1999]. В настоящей работе используется буквальное понимание диффузионно-столкновительного комплекса как ансамбля структур, образующихся в процессе диффузии, в которых макромолекулы сближены друг с другом.

Разработана компьютерная программа *ProKSim* (Protein Kinetics Simulator), предназначенная для моделирования взаимодействия макромолекул в растворе и в клетке методом многочастичной броуновской динамики (БД) с учетом электростатических взаимодействий [Kovalenko et al., 2006; Kovalenko et al., 2011; Коваленко и др., 2009]. В модели молекулы белка рассматриваются как броуновские частицы, совершающие поступательное и вращательное движения в вязкой среде. Конформационная подвижность молекул не учитывается (молекулы рассматриваются как твердые тела), не рассматриваются также гидрофобные взаимодействия белков. В результате компьютерного моделирования может быть получена статистическая информация об образующихся диффузионно-столкновительных комплексах. Для заданного критерия образования белок-белковых комплексов могут быть получены кинетические параметры реакции — зависимость количества (концентрации) комплексов от времени и, следовательно, константа скорости реакции образования комплексов.

Структуры диффузионно-столкновительных комплексов, получаемые с помощью *ProKSim*, могут быть в дальнейшем исследованы другими методами, в частности, с помощью методов молекулярной динамики и квантовой химии, что позволит провести полную реконструкцию процесса взаимодействия белков.

2. Метод броуновской динамики

Метод броуновской динамики является разновидностью метода стохастической динамики и в применении к химической кинетике был развит в работе [Ermaк, McCammon, 1978]. В методе рассматривается решение стохастического уравнения движения (уравнения Ланжевена),

которое для поступательного движения имеет вид

$$m\dot{\mathbf{r}} = -\xi_{\text{tr}}\dot{\mathbf{r}} + \mathbf{F}(\mathbf{r}) + \mathbf{f}(t), \quad (1)$$

где m — масса частицы, \mathbf{r} — положение частицы, t — время, ξ_{tr} — коэффициент вязкого трения частицы в среде, $\mathbf{F}(\mathbf{r})$ — внешняя и $\mathbf{f}(t)$ — случайная силы. Случайная сила $\mathbf{f}(t)$ неявно учитывает тепловое движение молекул среды и их столкновения с макромолекулой и представляет собой дельта-коррелированную случайную величину (белый шум), математическое ожидание которой равно нулю, а автокорреляционная функция равна

$$\langle \mathbf{f}(t) \mathbf{f}(t') \rangle = 2kT\xi_{\text{tr}}\delta\delta(t-t'),$$

где k — постоянная Больцмана, T — температура, δ — символ Кронекера и $\delta(t)$ — дельта-функция Дирака.

Для получения численной схемы мы выполняем проекцию на каждую из координатных осей и почленно усредняем по времени уравнение (1). Для небольших белков в водной среде при нормальных условиях постоянная времени $\frac{m}{\xi} \approx 10^{-12}$ с. Использование шага по времени Δt ,

значительно превышающего $\frac{m}{\xi}$, позволяет пренебречь инерциальным членом ($\langle m\dot{\mathbf{r}} \rangle_{\Delta t} \approx 0$). Для смещений Δx вдоль каждой из осей координат получаем формулу

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

где F_x — проекция внешней силы на соответствующую ось, ξ_{tr}^x — коэффициент вязкого трения для движения вдоль оси, $N(0,1)$ — нормально распределенная случайная величина с нулевым средним и единичной дисперсией.

Аналогичным образом выводятся уравнения для вращательных степеней свободы:

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

где $\Delta\alpha_x$, M_x и ξ_{rot}^x — угол поворота, момент внешней силы и коэффициент вязкого трения для вращения относительно оси, вокруг которой рассматривается вращение, соответственно.

Шаг по времени является параметром модели и по умолчанию равен 10^{-10} с. В качестве генератора псевдослучайных чисел используется алгоритм, предложенный [Panneton, L'Ecuyer, 2005].

3. Начальные и граничные условия

Модельная сцена представляет собой трехмерный реакционный объем, который может быть разделен на несколько компартментов. Каждый компартмент имеет форму прямоугольного параллелепипеда. Трехмерная модель молекул строится по данным из Protein Data Bank (PDB) [Berman et al., 2007]. Перед началом моделирования заданное количество молекул каждого белка распределяются определенным (по умолчанию случайным) образом в пределах отведенного им компартмента. Границы компартментов в зависимости от параметров модели могут быть либо периодическими (в случае столкновения с границей молекула проходит через нее, появляясь со стороны противоположной границы), либо неупругими (в случае столкновения молекула вплотную подходит к границе компартмента, с которой происходит столкновение).

4. Взаимодействие макромолекул с растворителем и между собой

Метод броуновской динамики учитывает взаимодействие частиц с молекулами растворителя за счет введения случайной силы и силы трения в уравнения движения. Взаимодействие между отдельными броуновскими частицами рассматривается в явном виде. Электростатические взаимодействия являются дальнедействующими и влияют на скорость сближения и взаимную ориентацию макромолекул в процессе диффузии. Контактные взаимодействия проявляются при сближении частиц на расстояние порядка суммы ван-дер-Ваальсовых радиусов входящих в них атомов и препятствуют взаимному проникновению молекул.

4.1. Гидродинамические взаимодействия

В модели растворитель рассматривается неявным образом как непрерывная среда с вязким трением; значение вязкости является параметром модели и по умолчанию принимается равным $10^{-3} \text{ кг}\cdot\text{м}^{-1}\text{с}^{-1}$, что соответствует вязкости воды при температуре 293 К. Гидродинамические взаимодействия между макромолекулами не учитываются.

В гидродинамических расчетах для оценки коэффициентов вязкого трения молекула белка рассматривается как удлинённый эллипсоид вращения с плотностью, однородно распределённой по объёму, и массой, равной массе молекулы. Центр масс эллипсоида совпадает с центром масс молекулы. Ориентация эллипсоида выбирается таким образом, чтобы его ось вращения совпадала по направлению с осью минимальной инерции молекулы; направления двух других осей выбираются произвольно в плоскости, перпендикулярной оси вращения. Определение необходимого направления оси вращения определяется методом последовательного перебора. Длины осей выбираются таким образом, чтобы момент инерции эллипсоида относительно оси вращения был равен минимальному моменту инерции молекулы. Для каждой молекулы вводится собственная ортогональная система координат, связанная с осями соответствующего эллипсоида вращения, в которой рассматривается движение макромолекулы. Размеры эллипсоида используются для вычисления коэффициентов вязкого трения макромолекулы в растворителе [Perrin, 1936]:

$$\xi_{\text{tr}}^a = \frac{8\pi\eta R(p^2 - 1)^{\frac{3}{2}}}{p^{\frac{1}{2}} \left((2p^2 - 1) \ln \left(p + (p^2 - 1)^{\frac{1}{2}} \right) - p(p^2 - 1)^{\frac{1}{2}} \right)},$$

$$\xi_{\text{tr}}^b = \frac{16\pi\eta R(p^2 - 1)^{\frac{3}{2}}}{p^{\frac{1}{2}} \left((2p^2 - 3) \ln \left(p + (p^2 - 1)^{\frac{1}{2}} \right) + p(p^2 - 1)^{\frac{1}{2}} \right)},$$

$$\xi_{\text{rot}}^a = \frac{16\pi\eta R^3 (p^2 - 1)^{\frac{3}{2}}}{3p \left(-\ln \left(p + (p^2 - 1)^{\frac{1}{2}} \right) + p(p^2 - 1)^{\frac{1}{2}} \right)},$$

$$\xi_{\text{rot}}^b = \frac{16\pi\eta R^3 (p^2 - 1)^{\frac{3}{2}} (p^2 + 1)}{3p \left((2p^2 - 1) \ln \left(p + (p^2 - 1)^{\frac{1}{2}} \right) - p(p^2 - 1)^{\frac{1}{2}} \right)},$$

где ξ_{tr}^a и ξ_{tr}^b — коэффициенты вязкого трения при движении вдоль (a) и перпендикулярно (b) осям вращения эллипсоида, ξ_{rot}^a и ξ_{rot}^b — соответствующие коэффициенты вязкого трения для вращательного движения, $R = (ab^2)^{\frac{1}{3}}$ — характерный размер эллипсоида вращения, $p = \frac{a}{b} > 1$ — отношение длин осей эллипсоида вращения (a и b), η — динамическая вязкость среды.

4.2. Электростатические взаимодействия

Для расчета электростатических взаимодействий используется предложенный [Ertnak, McSammon, 1978] приближенный метод. Диэлектрические свойства макромолекул и растворителя, а также экранирующее влияние мобильных зарядов (ионов) учитываются с помощью уравнения Пуассона–Больцмана

$$\varepsilon_0 \nabla [\varepsilon(\mathbf{r}) \nabla \varphi(\mathbf{r})] = -\rho - \sum_i c_i^\infty z_i q_p e^{\left[\frac{z_i q_p \varphi(\mathbf{r})}{kT} \right]}, \quad (2)$$

где ε_0 — электрическая постоянная, $\varepsilon(\mathbf{r})$ — диэлектрическая проницаемость среды, $\varphi(\mathbf{r})$ — электростатический потенциал, ρ — объемная плотность фиксированных электрических зарядов (в молекуле белка), c_i^∞ — концентрация ионов в отсутствие электростатического поля, z_i — заряд иона, q_p — заряд протона, \sum_i обозначает суммирование по всем типам мобильных зарядов (ионов) в растворе [Князева и др., 2010].

Полагая $|z_i q_p \varphi(\mathbf{r})| < kT$, мы можем линеаризовать уравнение (2) подстановкой $e^{\left[\frac{z_i q_p \varphi(\mathbf{r})}{kT} \right]} \approx 1 - \frac{z_i q_p \varphi(\mathbf{r})}{kT}$:

$$\varepsilon_0 \nabla [\varepsilon(\mathbf{r}) \nabla \varphi(\mathbf{r})] = -\rho + \sum_i \frac{c_i^\infty z_i^2 q_p^2 \varphi(\mathbf{r})}{kT} - \sum_i c_i^\infty z_i q_p.$$

$\sum_i c_i^\infty z_i q_p = 0$ из соображений электронейтральности раствора.

Расчет потенциала электростатического поля проводится на кубической сетке по итерационной формуле [Ullmann, Knapp, 1999]

$$\varphi_0 = \frac{q_0 + \sum_j h \varepsilon_0 \varepsilon_j \varphi_j}{h^3 \kappa_0^2 + \sum_j h \varepsilon_0 \varepsilon_j}, \quad (3)$$

где φ_0 , φ_j — электростатический потенциал в ячейке, q_0 — суммарный электрический заряд в ячейке, ε_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, \sum_j обозначает суммирование по ячейкам, контактирующим с данной ячейкой одной гранью. Шаг сетки является параметром модели и по умолчанию принимается равным 10^{-10} м.

Область для расчета потенциала задается как симметрично расположенный относительно начала координат прямоугольный параллелепипед, ребра которого параллельны осям собственной системы координат молекулы. Размеры области определяются как минимальные размеры, при которых любой атом белка расположен не ближе определенного расстояния (радиуса отсечения для электростатических взаимодействий) от границы области. Значение радиуса отсечения является параметром модели и по умолчанию принимается равным $3.5 \cdot 10^{-9}$ м. На границах области расчета потенциала электростатического поля его значение принимается равным нулю.

При расчете потенциала электростатического поля молекула белка рассматривается как набор ячеек кубической сетки с отличной от растворителя диэлектрической проницаемостью и пространственно-распределенными фиксированными зарядами. Принадлежность ячейки к области внутри белка определяется по ее доступности для молекул растворителя. В эту область входят

- (а) все ячейки, в центре которых нельзя разместить сферу заданного радиуса (по умолчанию – $1.4 \cdot 10^{-10}$ м, что соответствует радиусу молекулы воды [Connolly, 1983; Richards, 1977]) так, чтобы внутри нее не оказалось ни одной точки, являющейся координатой центра атома в файле структуры молекулы (PDB), а также
- (б) те ячейки, для которых условие (а) не выполнено, но которые принадлежат к замкнутой области, со всех сторон окруженной ячейками, для которых условие (а) выполнено.

Все остальные ячейки считаются заполненными растворителем с мобильными зарядами (ионами). Величина ионной силы раствора является параметром модели. Диэлектрическая проницаемость растворителя принимается равной 80; в ячейках, принадлежащих белку и контактирующих с растворителем, $\epsilon = 40$; в ячейках, принадлежащих белку и не контактирующих с растворителем, $\epsilon = 2$. Для ячеек растворителя учитывается экранирующее влияние мобильных зарядов (ионов).

Значения электрических зарядов на атомах белков могут быть прочитаны из файла структуры белка или дополнительного файла (поддерживается чтение файлов в формате PDB и PQR, а также файлов топологии программного пакета GROMACS [Hess et al., 2008]), либо рассчитаны автоматически по уравнению Хендерсона–Хассельбальха со стандартными значениями константы диссоциации pK_a для аминокислотных остатков [Dawson et al., 1986; Hasselbalch, 1917; Henderson, 1908] на основании заданного в конфигурационном файле значения pH раствора.

Расчет электростатического поля и его градиента проводится однократно для каждого типа молекул. Электростатическая сила и ее момент, действующие на каждую макромолекулу со стороны окружающих ее макромолекул, вычисляются на каждом шаге модели, исходя из заранее рассчитанных значений напряженности электростатического поля; при этом влияние диэлектрической проницаемости самой молекулы не учитывается. Это позволяет значительно увеличить скорость расчета, так как используются предварительно рассчитанные значения напряженности поля и не требуется повторный расчет напряженности электростатического поля на каждом временном шаге модели. Расчет электростатических взаимодействий для пар макромолекул, расстояние между всеми атомами которых заведомо превышает радиус отсечения электростатического взаимодействия, не производится.

4.3. Контактные взаимодействия

Общепринятым способом описания контактных взаимодействий в молекулярном моделировании является использование эмпирического потенциала Леннард–Джонса («потенциала 6–12»). В то же время этот способ расчета контактных взаимодействий является весьма ресурсоемким и требует достаточно малого шага модели по времени (порядка $10^{-15} - 10^{-12}$ с), что плохо совместимо с методом броуновской динамики, в котором шаг по времени ограничен снизу (см. раздел 2). В *ProKSim* реализована более простая модель контактных взаимодействий как неупругих взаимодействий твердых тел. Для каждого белка определяется запрещенный объем, то есть область пространства, в которой не может находиться другая макромолекула [Northrup et al., 1987]. В том случае, если перемещение молекулы приводит к столкновению с другими молекулами или границами компартмента (если не используются периодические граничные условия), величина перемещения пропорционально уменьшается методом двоичного поиска по всем поступательным и вращательным степеням свободы. Расчет новых координат молекул и проверка их взаимопересечения осуществляется на каждом шаге модели последовательно для каждой молекулы.

При старте программы определяются радиусы сфер, вмещающих молекулы белков. Сначала проверка возможности взаимопересечения белковых молекул производится путем сравнения расстояния между их центрами масс с суммой радиусов вмещающих сфер. Для дальнейшей

более точной проверки столкновения макромолекул между собой в программе реализовано несколько алгоритмов:

- (а) Наиболее простой и очевидный (но и самый требовательный к вычислительным ресурсам) алгоритм использует представление каждого атома в виде шаров. Радиус шара равен ван-дер-ваальсовому радиусу для соответствующего химического элемента. Если два шара, относящиеся к разным макромолекулам, пересекаются, то такое взаимное расположение молекул считается невозможным и должно быть исправлено путем уменьшения величины перемещения молекулы в процессе работы модели, либо путем размещения молекулы в другом месте реакционного объема, если проверка производится на этапе начальной подготовки модельной сцены.
- (б) В другом варианте для проверки взаимопересечения молекул проводится проверка факта пересечения наборов ячеек сетки, используемой при расчете потенциала электростатического поля (см. рис. 1). Молекулы считаются столкнувшимися в том случае, если координаты центра поверхностной ячейки одного белка, то есть принадлежащей белку, но контактирующей с растворителем, либо, в зависимости от параметра модели, приповерхностной, то есть принадлежащей растворителю, но контактирующей с белком, оказываются внутри принадлежащей другому белку области. Этот вариант дает менее точные результаты, чем предыдущий, однако такая проверка требует значительно меньших вычислительных ресурсов.
- (в) Молекула белка представляется в виде набора шаров различных диаметров, приближенно описывающего форму макромолекулы [Kovalenko et al., 2006].

Для ускорения вычислений без ущерба для точности полученных результатов указанные алгоритмы могут использоваться совместно, например, вначале проводится грубая проверка взаимопересечения молекул по алгоритму (б), а более точная проверка с использованием координат атомов и их ван-дер-ваальсовых радиусов проводится только в том случае, если менее точный алгоритм предполагает возможность пересечения. Какие именно алгоритмы используются и в каком порядке, задается в конфигурационном файле программы.

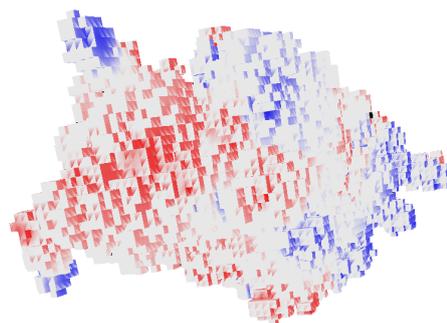


Рис. 1. Представление молекул белка в виде набора ячеек с ребром 10^{-10} м (алгоритм 4.3б) на примере цитохрома *f* (см. далее). Цветом обозначено значение электростатического потенциала в ячейках от -0.3 В (красный) до $+0.3$ В (синий). Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

5. Анализ статистики взаимодействия макромолекул и кинетики образования белок-белковых комплексов

5.1. Статистика межмолекулярных взаимодействий

В программе *ProKSim* предусмотрена возможность сохранения взаимного расположения макромолекул в одном из стандартных форматов файлов, предназначенных для визуализации и анализа молекулярных структур. Возможна запись структур комплексов в файлы формата PDB (RCSB Protein Data Bank) или в платформенно-независимом формате файла для траекторий GROMACS XTC [Green et al., 1995]. Предусмотрено несколько вариантов отбора структур для последующего анализа:

- (а) Отбираются конфигурации, в которых расстояние между любыми двумя атомами, относящимися к разным молекулам, не превышает пороговое значение. Пороговое значение является параметром модели и по умолчанию принимается равным $5 \cdot 10^{-10}$ м.

- (б) Отбираются конфигурации, для которых электростатические взаимодействия приводят к взаимному притяжению молекул с энергией, превышающей (по модулю) пороговое значение. Пороговое значение является параметром модели и по умолчанию принимается равным $kT \approx 4.1 \cdot 10^{-21}$ Дж.
- (в) В конфигурационном файле программы задается набор расстояний между заранее выбранными точками в собственной системе координат молекулы. В том случае, если для проверяемых молекул заданное число расстояний (по умолчанию – все расстояния, указанные в файле конфигурации) меньше или равны указанным в конфигурационном файле значениям, считается, что взаимное расположение молекул удовлетворяет критерию.
- Проверка критериев производится на каждом шаге модели для каждой пары молекул.

5.2. Кинетика межмолекулярных взаимодействий

В программе предусмотрена возможность прямого моделирования протекания биохимических реакций. Для каждого из перечисленных в разделе 5.1 критериев образования предварительного (диффузионно-столкновительного) комплекса может быть задана вероятность образования финального комплекса. Реакция образования комплексов рассматривается как необратимая, прореагировавшие молекулы исключаются из дальнейшего цикла работы модели (их положение в реакционном объеме фиксируется; они участвуют в контактных и электростатических взаимодействиях с другими молекулами, но не могут участвовать в других реакциях).

Трансформация предварительного комплекса в финальный рассматривается как случайный процесс. Реакционное событие считается произошедшим в случае, если полученное от генератора равномерно распределенных в диапазоне $[0,1)$ псевдослучайных чисел значение оказывается меньшим или равным заданной в конфигурационном файле вероятности образования комплекса за 1 шаг модели. На основе полученных данных проводится оценка константы скорости бимолекулярной реакции взаимодействия белков из предположения равенства начальных концентраций реагентов и отсутствия конкурирующих реакций по формуле

$$k = \frac{c(t)}{c_0(c_0 - c(t))t}, \quad (4)$$

где c_0 — начальная концентрация каждого из реагентов и $c(t)$ — концентрация комплекса в момент времени t .

6. Получение структур диффузионно-столкновительных комплексов и анализ их конфигурации

Проведено исследование диффузионно-столкновительных комплексов, получаемых в результате диффузии с учетом дальнедействующих электростатических взаимодействий, для трех пар белков: ферредоксин и ферредоксин:НАДФ⁺-редуктаза, пластоцианин и цитохром *f*, барназа и барстар.

В кубический реакционный объем с периодическими граничными условиями и размером грани $1.6 \cdot 10^{-7}$ м помещали по 100 молекул каждого белка, что соответствует концентрации $4 \cdot 10^{-5}$ М. В качестве параметров модели использовали приведенные выше значения по умолчанию. Контактные взаимодействия учитывали по алгоритму 4.3а. Для каждой пары белков проводили по два вычислительных эксперимента: в одном электростатические взаимодействия учитывались, в другом — не учитывались. Электростатические заряды на ионизированных аминокислотных остатках рассчитывали по уравнению Хендерсона–Хассельбальха для pH 7; ионную силу раствора брали равной 10^{-1} М. В каждом эксперименте по критериям 5.1а и 5.1б

отбирали взаимные положения белков для дальнейшего анализа. Таким образом, для каждой пары белков было получено три набора структур:

- (а) диффузионно-столкновительные комплексы, отобранные по критерию 5.1а при диффузии белков без учета электростатических взаимодействий,
- (б) диффузионно-столкновительные комплексы, отобранные по критерию 5.1а при диффузии белков с учетом электростатических взаимодействий,
- (в) структуры, отобранные по критерию 5.1б при диффузии белков с учетом электростатических взаимодействий.

Для каждой пары белков анализировали выборку из 20 000 случайно отобранных конфигураций комплексов из набора 6б. В качестве меры сходства между комплексами было использовано среднеквадратичное отклонение положений α -углеродных атомов белков (RMSD) [Kabsch, 1976]. Были вычислены значения RMSD для всех пар конфигураций в каждой выборке и получена треугольная матрица расстояний 20000 x 20000. Для каждой выборки была определена типичная структура, то есть такая конфигурация, среднее значение RMSD которой по сравнению с остальными структурами в выборке минимально. Эта типичная структура в дальнейшем использовалась как эталон при расчете среднеквадратичного отклонения атомов в конфигурациях, отобранных по критерию 5.1а для вычислительных экспериментов с учетом и без учета электростатических взаимодействий и 5.1б для вычислительного эксперимента с учетом электростатических взаимодействий.

6.1. Неиерархический кластерный анализ структур комплексов белков

Для анализа представительности выбранной типичной структуры проводили кластерный анализ по методу, описанному в [Daura et al., 1999] и реализованный в утилите `g_cluster` программного пакета GROMACS. Процедура кластеризации состоит в том, что для всех отобранных комплексов подсчитывается количество сходных с ним. Два комплекса считаются схожими между собой в случае, если величина RMSD для них меньше заданного порогового значения (*расстояния кластеризации*). Кластер структур включает комплекс, имеющий наибольшее количество схожих с ним структур комплексов, и все эти комплексы. После нахождения кластера все комплексы, входящие в него, исключаются из дальнейшего рассмотрения, и процедура кластеризации повторяется до тех пор, пока все комплексы не будут отнесены к какому-либо кластеру. Таким образом, в результате кластеризации каждая структура оказывается отнесенной к одному и только одному кластеру.

Для каждой выборки может быть определено такое минимальное значение расстояния кластеризации, начиная с которого все структуры относятся к одному кластеру. Если выборка состоит из случайно ориентированных друг относительно друга пар белков, отобранных по критерию 5.1а, то при расстоянии кластеризации, лишь незначительно меньшем этого минимального значения, выборка будет разделена на два кластера, первый из которых будет включать почти все структуры комплексов, а второй — лишь одну или несколько структур. При дальнейшем уменьшении расстояния кластеризации доля структур, входящих во второй кластер, будет возрастать, пока практически не сравняется с долей структур в первом кластере; дальнейшее уменьшение расстояния кластеризации приведет к быстрому росту числа кластеров. В том случае, если взаимная ориентация белков в парах белков не полностью случайна, и имеется какая-то предпочтительная ориентация, то зависимости числа кластеров и соотношения доли структур, относящихся ко второму и к первому кластеру, от расстояния кластеризации, будут иными: на периферии первого кластера появится большое количество микрокластеров, каждый из которых включает лишь небольшую долю всех структур. Соотношение доли структур во втором и в первом кластере приблизится к единице лишь тогда, когда общее число кластеров будет достаточно большим. Таким образом, анализируя характер зависимости этих параметров от расстояния кластеризации, можно оценить, насколько адекватно описание выборки с помощью единственной типичной структуры комплекса.

6.2. Иерархический кластерный анализ структур комплексов белков

Для иерархической группировки структур по частоте встречаемости использовали однопараметрический метод классификации структур с использованием алгоритмов, описанных в [Ankerst et al., 1999; Elke et al., 2006; Sander et al., 2003]. Алгоритм был адаптирован с учетом специфики анализируемых данных. В этом методе под кластером понимается плотная (в смысле частоты встречаемости похожих структур) группа, отделенная от смежных групп менее плотными областями. Для уменьшения влияния случайного появления похожих структур в разреженных областях для каждой структуры вычисляется так называемое *центральное расстояние CD (core distance)*. Это расстояние зависит от заданного исследователем параметра — минимального количества структур в кластере M — и определяется как значение RMSD данной структуры по сравнению с M -той (по порядку увеличения RMSD) структурой из выборки. В данной работе принимали $M = 300$.

Задача кластеризации решается в два этапа. Вначале проводится упорядочивание структур: начиная с произвольно выбранной структуры (в нашей работе мы начинали со структуры с наименьшим центральным расстоянием), составляется список, в который на каждом шаге работы алгоритма добавляется та из отсутствующих в списке структур, которая наиболее похожа на любую из добавленных ранее. В качестве критерия добавления используется расстояние доступности RD (*reachability distance*). RD принимается равным минимальному значению RMSD между данной структурой и любой структурой, 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 структур, еще не добавленных в список. При добавлении структуры в список значение RD сохраняется для дальнейшего анализа; для первой структуры оно принимается равным $+\infty$.

На втором этапе упорядоченный список используется для выделения кластеров структур; при этом кластеры являются вложенными друг в друга, и каждая структура относится как минимум к одному (корневому) кластеру. Проводится поиск локальных максимумов зависимости расстояния доступности структуры от ее номера в списке. Предшествующие локальному максимуму и следующие за ним подряд группы структур, имеющие расстояние доступности не более расстояния доступности локального максимума, выделяются в качестве отдельных подкластеров в том случае, если количество входящих в каждую из них структур не менее M . Структура комплекса, соответствующая локальному максимуму RD , считается принадлежащей к «правому» кластеру (рис. 2).

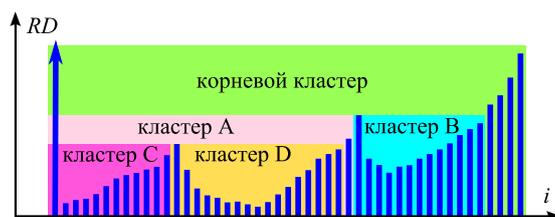


Рис. 2. Схематическое представление результата работы алгоритма поиска кластеров с помощью диаграммы зависимости расстояния доступности RD от номера структуры комплекса в списке i . Кластеры А и В вложены в корневой кластер, кластеры С и D — в кластер А. Цветная версия рисунка доступна в электронной версии журнала на сайте <http://cfm.ics.rg.ru/>

В результате применения этого алгоритма структуры комплексов белков оказываются расклассифицированными по иерархическому дереву, каждая ветвь которого соответствует более или менее обособленной группе схожих структур. Для каждой группы может быть определена своя типичная структура; можно сравнить группы по частоте (плотности) встречаемости похожих структур в группе и оценить степень обособленности двух близких групп. Группы структур (кластеры) могут быть вложены друг в друга, поэтому одна структура может принадлежать одновременно к нескольким кластерам.

7. Объекты исследования

7.1. Ферредоксин и ферредоксин:НАДФ⁺-редуктаза

Ферредоксин служит конечным акцептором электронов фотосистемы I. Он передает электроны в разные пути метаболизма, в том числе в путь фиксации углерода через ферредоксин:НАДФ⁺-редуктазу (ФНР), восстанавливающую никотинамиддинуклеотидфосфат (НАДФ⁺) до НАДФН, необходимого в цикле Кальвина. Ферредоксин содержит железо-серный кластер [2Fe-2S], в котором один из атомов железа меняет степень окисления между +2 и +3. В ФНР электроны распределяются по сопряженной системе ароматических связей флавинадениндинуклеотида (ФАД). Константа скорости образования комплекса ферредоксина и ФНР из цианобактерии *Anabaena* составляет $k_a = 6.9 \cdot 10^8 \text{ M}^{-1} \text{ c}^{-1}$, константа диссоциации $K_d = 2.2 \cdot 10^{-6} \text{ M}$ [Hurley et al., 1999].

Взаимодействие белков ферредоксина и ФНР было исследовано [Kovalenko et al., 2010] методом многочастичной броуновской динамики. Были вычислены зависимости константы скорости образования комплексов белков от ионной силы для дикого типа и мутантов ФНР.

В настоящей модели рассматривается полученная методом рентгеноструктурного анализа (РСА) структура комплекса ферредоксина и ФНР из *Anabaena* (PDB ID: 1EWY) [Morales et al., 2000]. В качестве исходных данных для построения модели использовали цепь С (ферредоксин) и цепь А (ФНР) указанного комплекса. Недостающие в структуре легкие атомы были добавлены с помощью утилиты `pdb2gmx` программного комплекса GROMACS. Рассматривалось взаимодействие восстановленного ферредоксина и окисленной ФНР. Суммарные заряды белков в модели составляли -14.9 (ферредоксин) и -2.2 (ферредоксин:НАДФ⁺-редуктаза) элементарного заряда.

7.2. Пластоцианин и цитохром *f*

Пластоцианин является мобильным переносчиком электронов, передающим электрон с цитохромного *b₆f* комплекса на фотосистему I у всех высших растений и некоторых водорослей. Цитохром *f* — самая крупная субъединица цитохромного *b₆f* 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 $k_a = 2 \cdot 10^8 \text{ M}^{-1} \text{ c}^{-1}$ [Kannt et al., 1996].

Взаимодействие белков пластоцианина и цитохрома *f* высших растений было исследовано методом БД в работах [Pearson, Gross, 1998; Rienzo et al., 2001; Ullmann et al., 1997]. Методом многочастичной БД было исследовано взаимодействие пластоцианина и цитохрома *f* в растворе [Kovalenko et al., 2006] и люмене тилакоида [Коваленко и др., 2008], в том числе с учетом электростатических взаимодействий с тилакоидной мембраной [Князева и др., 2010].

В настоящей модели рассматривается структура комплекса пластоцианина шпината и цитохрома *f* турнепса (PDB ID: 2PCF) [Ubbink et al., 1998], полученная методом спектроскопии ядерного магнитного резонанса (ЯМР). Рассматривалось взаимодействие восстановленного цитохрома *f* и окисленного пластоцианина. Распределение электростатических зарядов на геме взято из [Autenrieth et al., 2004]. Суммарные заряды белков в модели составляли -8.00 (пластоцианин) и -3.95 (цитохром *f*) элементарного заряда.

7.3. Барназа и барстар

Бактериальная РНК-аза (барназа) и ее ингибитор барстар являются одним из классических объектов при моделировании белок-белковых взаимодействий [Hoefling, Gottschalk, 2010;

Kieslich et al., 2010; Spaar et al., 2006; Wang et al., 2004; Ермакова, 2006], так как формируют (в среде с водородным показателем рН от 7 до 9) очень прочно связанный комплекс. Период полураспада этого комплекса составляет около двух дней (константа диссоциации комплекса $K_d = 10^{-14}$ М, константа скорости образования комплекса $k_a = 4 \cdot 10^8$ М⁻¹с⁻¹). В кислой среде (рН ниже 4.5) сродство белков друг к другу полностью исчезает [Schreiber, 2001]. Барназа является ядом, секретлируемым бактерией *Bacillus amyloliquefaciens* во внешнюю среду. Формирование устойчивого комплекса белков барназы и барстара внутри клетки бактерии является непременным условием выживания организма, так как белок чрезвычайно цитотоксичен из-за крайне высокой РНК-азной активности.

В модели рассматривали структуру комплекса барназы и барстара (PDB ID: 1BGS) [Guillet et al., 1993], полученную методом рентгеноструктурного анализа. В качестве исходных данных для построения модели использовали цепь А (барназа) и цепь Е (барстар) указанного комплекса. Недостающие в структуре легкие атомы были добавлены с помощью утилиты `pdb2gmx` программного комплекса GROMACS. Суммарные заряды белков в модели составляли +2.23 (барназа) и -5.89 (барстар) элементарного заряда.

8. Типичные структуры диффузионно-столкновительных комплексов

Проведено исследование диффузионно-столкновительных комплексов для трех пар белков: ферредоксин и ферредоксин:НАДФ⁺-редуктаза, пластоцианин и цитохром *f*, барназа и барстар. На рисунке 3 показаны результаты неиерархической кластеризации структур диффузионно-столкновительных комплексов, отобранных по критерию 5.1а, с помощью метода, описанного в разделе 6.1. Вид приведенных графиков свидетельствует о том, что распределение конфигураций диффузионно-столкновительных комплексов ферредоксина и ферредоксин:НАДФ⁺-редуктазы и пластоцианина и цитохрома *f* значительно отличается от случайного и для них можно выбрать представительные типичные структуры. Эти структуры визуализированы на рис. 4в и 5б. Распределение структур диффузионно-столкновительных комплексов барназы и барстара соответствует случайному, и, следовательно, полученная типичная структура (рисунок 6б) не может быть признана представительной.

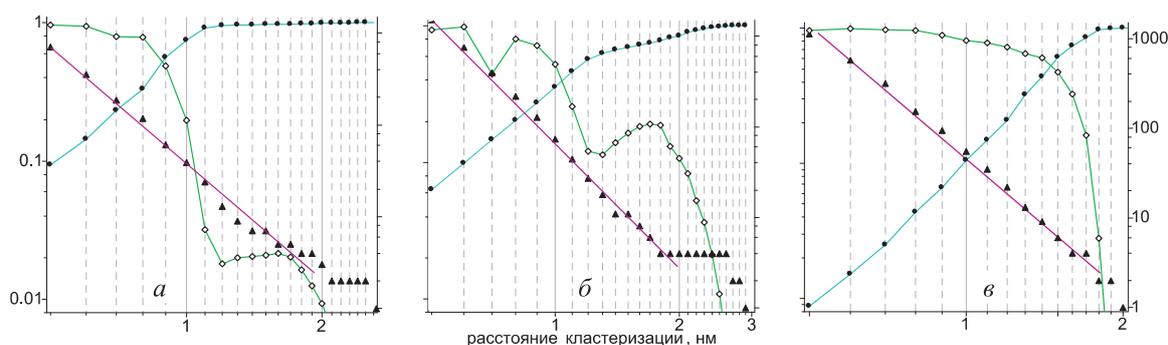


Рис. 3. Зависимости числа кластеров (\blacktriangle , по правой оси ординат), доли самого крупного кластера от всей выборки (\bullet , по левой оси ординат) и соотношения размеров второго и первого кластера (\diamond , по левой оси ординат) от расстояния кластеризации. Графики приведены в логарифмическом масштабе. (а) – ферредоксин и ферредоксин:НАДФ⁺-редуктаза, (б) – пластоцианин и цитохром *f*, (в) – барназа и барстар.

8.1. Ферредоксин и ферредоксин:НАДФ⁺-редуктаза

На рисунке 4 видно, что ориентация белков ферредоксина и ФНР в диффузионно-столкновительном комплексе в общем сходна с их ориентацией в кристалле (PDB ID: 1EWY). Следует отметить, что в кристаллической структуре на каждую молекулу ферредоксина приходится по две молекулы ФНР, и она оказывается зажатой между этими молекулами. Одна из этих

молекул (цепь А) сближена с железно-серным кластером ферредоксина (цепь С) флавиновым трициклом (расстояние $7.4 \cdot 10^{-10}$ м), другая (цепь В) — адениновым остатком (расстояние $1.1 \cdot 10^{-9}$ м). Расстояние между флавиновым трициклом цепи В и железно-серным кластером ферредоксина составляет $1.5 \cdot 10^{-9}$ м. В полученной типичной структуре диффузионно-столкновительного комплекса расстояние между флавиновым трициклом ФНР и железно-серным кластером ферредоксина составляет $2.3 \cdot 10^{-9}$ м. Среднеквадратичное отклонение положений α -углеродных атомов составляет $2.4 \cdot 10^{-9}$ м (как при сравнении с цепями А и С, так и при сравнении с цепями В и С).

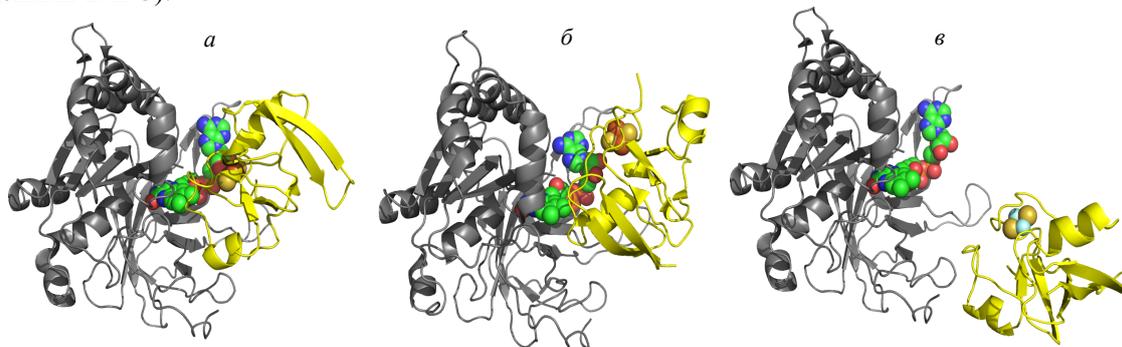


Рис. 4. Структура комплекса ферредоксина (желтый) и ФНР (серый), полученного методом РСА (PDB ID: 1EWY) (*a* — цепи А и С, *б* — цепи В и С), и типичная структура, полученная при анализе траекторий броуновской динамики (*в*). В виде сфер обозначены атомы, входящие в состав ФАД и железно-серного кластера [2Fe-2S]. Для визуализации структур использован PyMol [Schrödinger LLC, 2012]. Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

8.2. Пластоцианин и цитохром *f*

В типичной структуре диффузионно-столкновительного комплекса пластоцианина и цитохрома *f* (рис. 5) пластоцианин расположен вблизи экспериментально обнаруженного сайта связывания, однако ориентация его значительно отличается от экспериментально определенной. Расстояние между гемовым железом цитохрома *f* и ионом меди пластоцианина в экспериментально изученной структуре комплекса (PDB ID: 2PCF) составляет $1.1 \cdot 10^{-9}$ м, в то время как в полученном методом броуновской динамики диффузионно-столкновительном комплексе — $4.1 \cdot 10^{-9}$ м. Среднеквадратичное отклонение положения α -углеродных атомов от положения в комплексе, полученном методом ЯМР (PDB ID: 2PCF, модель 9), составляет $1.3 \cdot 10^{-9}$ м.

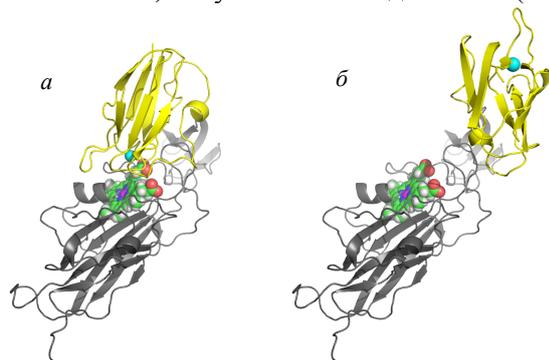


Рис. 5. Структура комплекса пластоцианина (желтый) и цитохрома *f* (серый), полученного методом ЯМР (PDB ID: 2PCF, модель 9) (*a*), и типичная структура, полученная при анализе траекторий броуновской динамики (*б*). В виде сфер обозначены ион меди и атомы, входящие в состав гема. Для визуализации структур использован PyMol [Schrödinger LLC, 2012]. Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

8.3. Барназа и барстар

На рисунке 6 показаны экспериментально полученные варианты ориентации белка барстара относительно барназы в кристалле (PDB ID: 1BGS) и типичная модельная структура диффузионно-столкновительного комплекса. Можно видеть, что эта структура не соответствует ни одному из вариантов расположения этих белков в кристалле. Более подробный анализ модель-

ной структуры не проводился, так как методами кластерного анализа было показано, что данная структура не может быть признана представительной для исследованной выборки. Использование иерархической классификации структур диффузионно-столкновительных комплексов, отобранных по критерию 5.1а, с помощью метода, описанного в разделе 6.2, также не позволило выявить какие-либо конфигурации, частота встречаемости которых в выборке значительно отличалась от средней. В то же время анализ структур, отобранных по критерию 5.1б, то есть таких конфигураций, в которых энергия притяжения между двумя молекулами превышает пороговое значение $kT \approx 4.1 \cdot 10^{-21}$ Дж, позволил выявить две обособленные группы энергетически выгодных взаимных расположений белков (рис. 7). Среди энергетически выгодных ориентаций наиболее часто встречаются структуры, схожие с приведенной на рисунке 7г. Расположение барстара относительно барназы в этой структуре соответствует одному из физиологически незначимых положений в кристалле, хотя его ориентация отличается от экспериментально полученной (рис. 7б; среднеквадратичное отклонение положений α -углеродных атомов — $1.1 \cdot 10^{-9}$ м). Вторая по встречаемости группа показана на рисунке 7в и близка к физиологически значимому комплексу (рис. 7а; среднеквадратичное отклонение положений α -углеродных атомов — $1.3 \cdot 10^{-9}$ м). Полученные энергетически выгодные состояния находятся в согласии с результатами [Spraag et al., 2006]; использованный нами метод иерархической классификации структур позволил более наглядно представить эти состояния, так как использованный в [Spraag et al., 2006] метод позволяет только определить расположение центра масс одного белка относительно другого.

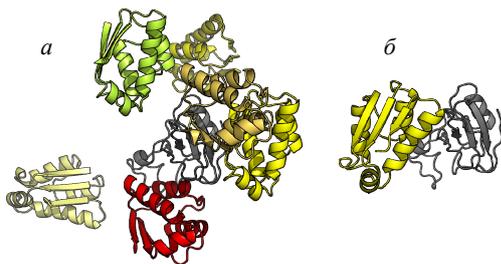


Рис. 6. а — варианты расположения барстара относительно барназы (серый) в комплексе, исследованном методом PCA (PDB ID: 1BGS). Красным показано физиологически значимое положение, различными оттенками желтого — альтернативные положения с учетом симметрии кристаллической решетки. б — неrepresentative типичная структура, полученная при анализе траекторий броуновской динамики. Для визуализации структур использован PyMol [Schrödinger LLC, 2012]. Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

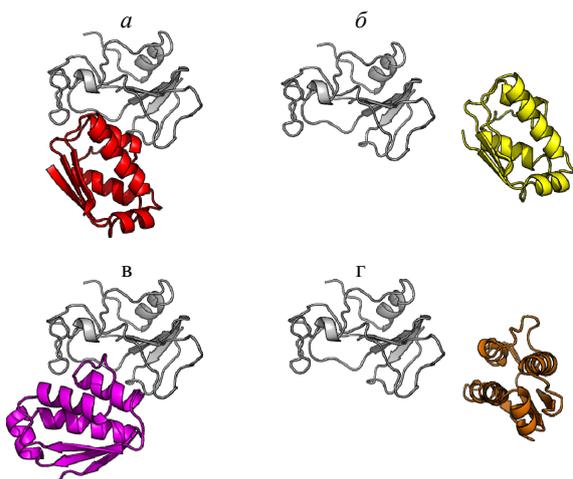


Рис. 7. а, б — варианты расположения барстара относительно барназы (серый) в комплексе, исследованном методом PCA (PDB ID: 1BGS). Красным показано физиологически значимое положение (а). в, г — соответствующие (а) и (б) группы энергетически выгодных ориентаций, полученных методом броуновской динамики. Для визуализации структур использован PyMol [Schrödinger LLC, 2012]. Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

9. Роль электростатических взаимодействий во взаимной ориентации белков в процессе диффузии

Для двух исследованных пар фотосинтетических электрон-транспортных белков ферредоксина и ферредоксин:НАДФ⁺-редуктазы и пластоцианина и цитохрома *f* получены предста-

вительные типичные структуры диффузионно-столкновительных комплексов. Показано, что в процессе диффузии с учетом электростатических взаимодействий мобильные переносчики электронов (ферредоксин и пластоцианин) подходят к соответствующим сайтам своих партнеров по взаимодействию, однако их ориентация может отличаться от ориентации в финальном комплексе (то есть конфигурации, в которой возможен перенос электрона). Для количественного определения вклада электростатических взаимодействий в формирование ансамбля структур диффузионно-столкновительного комплекса было проведено сравнение (по параметру среднеквадратичного отклонения положения α -углеродных атомов белков) всех отобранных в вычислительных экспериментах конфигураций диффузионно-столкновительных комплексов с соответствующими типичными структурами. График функции плотности полученных распределений показан на рис. 8. Видно, что распределение, полученное по критерию $6б$, может быть представлено в виде взвешенной суперпозиции распределений $6а$ и $6в$. Полученные по методу наименьших квадратов веса соответствующих распределений приведены в таблице 1. Полученные веса дают количественную оценку роли далекодействующих электростатических взаимодействий для процесса сближения и взаимной ориентации белков в процессе диффузии.

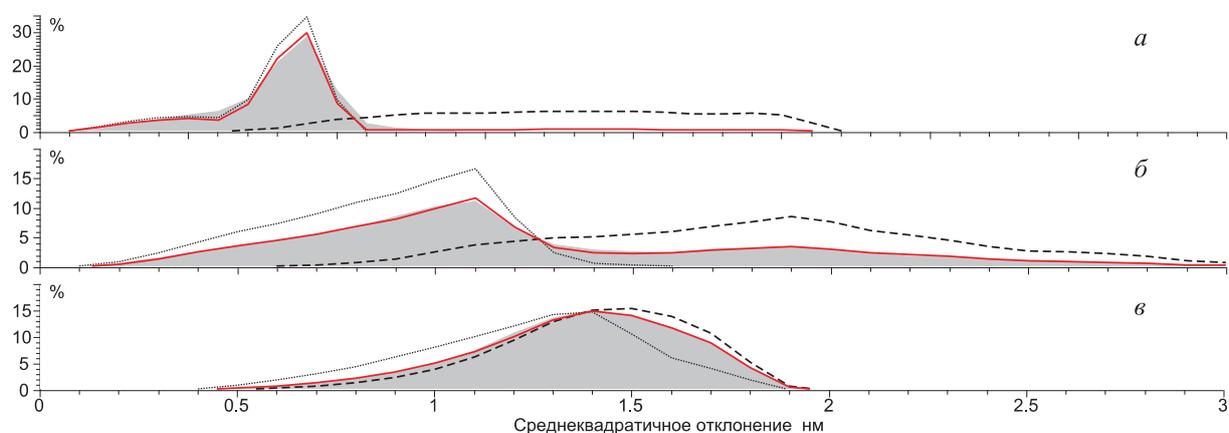


Рис. 8. Функция плотности распределения среднеквадратичного отклонения положения α -углеродных атомов диффузионно-столкновительных комплексов белков относительно типичной структуры. Пунктирная линия — диффузия без учета электростатических взаимодействий, отбор комплексов при расстоянии между атомами двух белков меньше $5 \cdot 10^{-10}$ м ($6а$); точечная линия — диффузия с учетом электростатических взаимодействий, отбор комплексов при энергии электростатического притяжения $\geq kT$ ($6в$). Полутоновая заливка — диффузия с учетом электростатических взаимодействий, отбор комплексов при расстоянии между атомами двух белков меньше $5 \cdot 10^{-10}$ м ($6б$); красная сплошная линия — результат фитирования распределения $6б$ в виде взвешенной суммы распределений $6а$ и $6в$, соответствующие веса приведены в таблице 1. $а$ — ферредоксин и ферредоксин:НАДФ⁺-редуктаза, $б$ — пластоцианин и цитохром f , $в$ — барназа и барстар

Таблица 1. Коэффициенты фитирования по методу наименьших квадратов функции плотности распределения $6б$ в виде взвешенной суммы распределений $6а$ и $6в$

| Белки | Вес функции плотности распределения в суперпозиции | |
|--|--|---|
| | $6а$ (электростатические взаимодействия) | $6в$ (без учета электростатических взаимодействий) |
| Ферредоксин и ферредоксин:НАДФ ⁺ -редуктаза | 0.85 | 0.15 |
| Пластоцианин и цитохром f | 0.61 | 0.39 |
| Барназа и барстар | 0.30 | 0.70 |

В представленной работе показана возможность применения метода броуновской динамики для определения роли далекодействующих электростатических взаимодействий во взаимной

ориентации макромолекул в процессе диффузии и получения типичных структур диффузионно-столкновительных комплексов. Этот метод также широко применяется для изучения кинетики диффузионно-контролируемых реакций [Kovalenko et al., 2006 и др. работы]. Однако при сближении молекул в их взаимодействии существенную роль начинают играть те факторы, которыми мы пренебрегаем в методе броуновской динамики. К числу таких факторов относятся внутримолекулярная подвижность, формирование водородных связей и солевых мостиков, а также гидрофобные взаимодействия. Все эти факторы участвуют в трансформации предварительных диффузионно-столкновительных комплексов в финальные (обладающие физиологической активностью, например, способствующие переносу электрона между активными центрами электрон-транспортных белков). Трансформация диффузионно-столкновительных комплексов, получаемых с помощью программного обеспечения **ProKSim**, может быть исследована с помощью методов молекулярной динамики и квантовой химии, что выходит за рамки представленной работы.

Программное обеспечение **ProKSim** может быть загружено по адресу http://www.biophys.msu.ru/rus/science/complex_systems/ProKSim/. Исходный код может быть предоставлен по запросу.

10. Список литературы

- Воеводин В. В., Жуматий С. А., Соболев С. И., Антонов А. С., Брызгалов П. А., Никитенко Д. А., Стефанов К. С., Воеводин В. В. Практика суперкомпьютера «Ломоносов» // Открытые системы. — 2012. — № 7.
- Ермакова Е. А. Сравнительное изучение реакции взаимодействия *Bacillus umyloliquefaciens* рибонуклеазы (барназы) и *Bacillus intermedius* рибонуклеазы (биназы) с барстаром методом броуновской динамики // Биофизика. — 2006. — Т. 51, № 2. — С. 242–249.
- Князева О. С., Коваленко И. Б., Абатурова А. М., Ризниченко Г. Ю., Грачев Е. А., Рубин А. Б. Многочастичная модель диффузии и взаимодействия пластоцианина с цитохромом *f* в электростатическом поле фотосинтетической мембраны // Биофизика. — 2010. — Т. 55, № 2. — С. 259–268.
- Коваленко И. Б., Абатурова А. М., Громов П. А., Устинин Д. М., Грачев Е. А., Ризниченко Г. Ю., Рубин А. Б. Компьютерное моделирование образования комплекса между пластоцианином и цитохромом *f* в люмене тилакоида // Биофизика. — 2008. — Т. 53, № 2. — С. 261–270.
- Коваленко И. Б., Абатурова А. М., Ризниченко Г. Ю., Рубин А. Б. Новый метод компьютерного моделирования образования белок-белковых комплексов // Доклады Академии наук. — 2009. — Т. 427, № 5. — С. 696–698.
- Ankerst M., Breunig M. M., Kriegel H.-P., Sander J. OPTICS: Ordering Points To Identify the Clustering Structure // Proceedings of ACM SIGMOD, Philadelphia, PA., 1999. — P. 49–60.
- Autenrieth F., Tajkhorshid E., Baudry J., Luthey-Schulten Z. Classical force field parameters for the heme prosthetic group of cytochrome *c*. // Journal of computational chemistry. — 2004. — V. 25, No. 13. — P. 1613–22.
- Berman H., Henrick K., Nakamura H., Markley J.L. The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data // Nucleic Acids Research. — 2007. — V. 35. — Database issue. P. D301–D303.
- Connolly M. L. Analytical molecular surface calculation // J. Appl. Crystallogr. — 1983. — V. 16. — P. 548–558.
- Daura X., Gademann K., Jaun B., Seebach D., Gunsteren W.F. van, Mark A.E. Peptide Folding: When Simulation Meets Experiment // Angewandte Chemie International Edition. — 1999. — V. 38, No. 1–2. — P. 236–240.
- Dawson R. M. C., Elliott D. C., Elliott W. H., Jones K. M. Data for Biochemical Research (third edition). Oxford: Clarendon Press, 1986.

- Elke A., Böhm C., Kröger P.* DeLiClu: Boosting Robustness, Completeness, Usability, and Efficiency of Hierarchical Clustering by a Closest Pair Ranking // Proceedings of the 10th Pacific-Asian Conference on Advances in Knowledge Discovery and Data Mining., 2006. — P. 119–128.
- Ermak D. L., McCammon J. A.* Brownian dynamics with hydrodynamic interactions // J. Chem. Phys. — 1978. — V. 69, No. 4. — P. 1353–1360.
- Gabdoulline R. R., Wade R. C.* On the protein-protein diffusional encounter complex. // Journal of molecular recognition JMR. — 1999. — V. 12, No. 4. — P. 226–234.
- Green D., Meacham K., Surridge M., Hoesel F. van, Berendsen H.* Methods and Techniques in Computational Chemistry: METECC-95 // Methods and Techniques in Computational Chemistry: METECC-95. : STEF, Cagliari, 1995. — P. 435.
- Guillet V., Laphorn A., Hartley R.W., Mauguen Y.* Recognition between a bacterial ribonuclease, barnase, and its natural inhibitor, barstar // Structure (London, England : 1993). — 1993. — V. 1, No. 3. — P. 165–76.
- Hasselbalch K. A.* Die Berechnung der Wasserstoffzahl des Blutes aus der freien und gebundenen Kohlensäure desselben, und die Sauerstoffbindung des Blutes als Funktion der Wasserstoffzahl // Biochemische Zeitschrift. — 1917. — V. 78. — P. 112–144.
- Henderson L. J.* Concerning the relationship between the strength of acids and their capacity to preserve // Am. J. Physiol. — 1908. — V. 21. — P. 173–179.
- Hess B., Kutzner C., Van Der Spoel D., Lindahl E.* GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation // Journal of Chemical Theory and Computation. — 2008. — V. 4, No. 3. — P. 435–447.
- Hoefling M., Gottschalk K. E.* Barnase-Barstar: From first encounter to final complex. // Journal of Structural Biology. — 2010. — V. 171, No. 1. — P. 52–63.
- Hurley J. K., Hazzard J. T., Martinez-Julvez M., Medina M., Gomez-Moreno C., Tollin G.* Electrostatic forces involved in orienting Anabaena ferredoxin during binding to Anabaena ferredoxin:NADP⁺ reductase: site-specific mutagenesis, transient kinetic measurements, and electrostatic surface potentials // Protein Sci. 1999. V. 8. P. 1614–1622.
- Kabsch W.* A solution for the best rotation to relate two sets of vectors // Acta Crystallographica Section A. — 1976. — V. 32, No. 5. — P. 922–923.
- Kannt A., Young S., Bendall D.S.* The role of acidic residues of plastocyanin in its interaction with cytochrome f // Biochim. Biophys. Acta. — 1996. — V. 1277. — P. 115–126.
- Kieslich C. A., Gorham R. D., Morikis D.* Is the rigid-body assumption reasonable?. Insights into the effects of dynamics on the electrostatic analysis of barnase-barstar // Journal of Non-Crystalline Solids. — 2010.
- Kovalenko I. B., Abaturova A. M., Diakonova A. N., Knyazeva O. S., Ustinin D. M., Khruschev S. S., Riznichenko G. Y., Rubin A. B.* Computer simulation of protein-protein association in photosynthesis // Math. Model. Nat. Phenom. — 2011. — V. 6, No. 7. — P. 39–54.
- Kovalenko I. B., Abaturova A. M., Gromov P. A., Ustinin D. M., Grachev E. A., Riznichenko G. Y., Rubin A.B.* Direct simulation of plastocyanin and cytochrome f interactions in solution // Phys. Biol. — 2006. — V. 3. — P. 121–129.
- Kovalenko I. B., Diakonova A. N., Abaturova A. M., Riznichenko G. Y., Rubin A. B.* Direct computer simulation of ferredoxin and FNR complex formation in solution // Phys. Biol. — 2010. — V. 7, No. 2. — P. 26001.
- Morales R., Kachalova G., Vellieux F., Charon M.H., Frey M.* Crystallographic studies of the interaction between the ferredoxin-NADP⁺ reductase and ferredoxin from the cyanobacterium Anabaena: looking for the elusive ferredoxin molecule. // Acta crystallographica. Section D, Biological crystallography. — 2000. — V. 56, No. Pt 11. — P. 1408–12.
- Northrup S., Boles J., Reynolds J.* Electrostatic effects in the brownian dynamics of association and orientation of heme proteins // J Phys Chem. — 1987. — V. 91. — P. 5991–5998.
- Panneton F., L'Ecuyer P.* On the xorshift random number generators // ACM T. Model. Comput. S. — 2005. — V. 15, No. 4. — P. 346–361.

- Pearson Jr. D. C., Gross E. L.* Brownian Dynamics Study of the interaction between plastocyanin and cytochrome f // *Biophysical Journal*. — 1998. — V. 75. — P. 2698–2711.
- Perrin F.* Mouvement Brownien d'un ellipsoïde (II). Rotation libre et dépolariisation des fluorescences. Translation et diffusion de molécules ellipsoïdales // *J. Phys. Radium*. — 1936. — V. 7, No. 1. — P. 1–11.
- Richards F. M.* Areas, volumes, packing and protein structure. // *Annual review of biophysics and bioengineering*. — 1977. — V. 6. — P. 151–76.
- De Rienzo F., Gabdoulline R. R., Menziani M. C., De Benedetti P. G., Wade R. C., Rienzo F., Benedetti P.* Electrostatic analysis and brownian dynamics simulation of the association of plastocyanin and cytochrome f // *Biophys. J.* — 2001. — V. 81, No. 3. — P. 3090–3104.
- Sander J., Qin X., Lu Z., Niu N., Kovarsky A.* Automatic Extraction of Clusters from Hierarchical Clustering Representations // *Proceedings of the 7th Pacific-Asia Conference on Knowledge Discovery and DataMining (PAKDD)*, Seoul, Korea.: Springer-Verlag, 2003. — P. 75–87.
- Schreiber G.* Methods for Studying the Interaction of Barnase with Its Inhibitor Barstar // *Nuclease Methods and Protocols (Methods in Molecular Biology)* / под ред. *C.H. Schein*.: Humana Press, 2001. — P. 213–226.
- Schrödinger LLC.* The PyMOL Molecular Graphics System. 2003–2012.
- Spaar A., Dammer C., Gabdoulline R.R., Wade R.C., Helms V.* Diffusional Encounter of Barnase and Barstar // *Biophys. J.* — 2006. — V. 90, No. 6. — P. 1913–1924.
- Ubbink M., Ejdebeck M., Karlsson B. G., Bendall D. S.* The structure of the complex of plastocyanin and cytochrome f, determined by paramagnetic NMR and restrained rigid-body molecular dynamics // *Structure*. — 1998. — V. 6. — P. 323–335.
- Ullmann G. M., Knapp E.-W.* Electrostatic models for computing protonation and redox equilibria in proteins // *Eur. Biophys. J.* — 1999. — V. 28, No. 7. — P. 533–551.
- Ullmann G. M., Knapp E.-W., Kostic N.M.* Computational simulation and analysis of dynamic association between plastocyanin and cytochrome f. Consequences for the electron-transfer reaction // *J. Am. Chem. Soc.* — 1997. — V. 119. — P. 42–52.
- Wang T., Tomic S., Gabdoulline R. R., Wade R.C.* How Optimal Are the Binding Energetics of Barnase and Barstar? // *Biophysical Journal*. — 2004. — V. 87, No. 3. — P. 1618–1630.