

Статистические характеристики гидратных оболочек белков. Компьютерное моделирование

Е. В. Рубцова^{1,a}, А. Б. Соловей^{2,b}, В. И. Лобышев^{3,c}

¹Московский государственный университет имени М. В. Ломоносова, физический факультет, кафедра математики. Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2.

²СУНЦ МГУ. Россия, 121357, Москва, ул. Кременчугская, д. 11.

³Московский государственный университет имени М. В. Ломоносова, физический факультет, кафедра биофизики. Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2.

E-mail: ^aev.rubcova@physics.msu.ru, ^bsoloveybird@gmail.com, ^clobyshev@yandex.ru

Статья поступила 15.01.2015, подписана в печать 01.06.2015.

Молекулы воды и сетка водородных связей являются обязательными структурными компонентами во многих биологических макромолекулярных системах, таких как нуклеиновые кислоты, фибриллярные и глобулярные белки. Известно, что плотность воды в гидратных оболочках белка отличается от плотности объемной воды. В настоящей работе плотность рассмотрена как пространственная функция расстояния от поверхности белка. Аналогично рассмотрены значения функций электростатического потенциала. Полученные профили плотности в слое связанной воды характеризуются наличием ряда четко выраженных максимумов, объясняющихся упорядоченным строением гидратной оболочки на расстоянии, превышающем 7 Å. Наблюдаемые небольшие различия в профилях функций распределения плотности и электростатического потенциала для разных белков обусловлены топологическими изменениями, обнаруженными нами ранее при исследовании функций распределения валентных и торсионных углов связанной воды с этими белками.

Ключевые слова: гидратная оболочка белков, моделирование структуры, плотность.

УДК: 577.3. PACS: 31.15.-p, 87.15.ap.

Введение

Характеристики физических свойств поверхностей «белок — растворитель» являются существенными для понимания структуры, фолдинга белков, их взаимодействия между собой и другими макромолекулами [1]. В отсутствие воды белковые цепи практически неподвижны и ферментативной активности нет или она мала. В работах по дегидратации белков было показано, что белок, как правило, должен быть окружен по крайней мере первой гидратной оболочкой для поддержания своей ферментативной и структурной функции [2].

Молекулы воды в растворе белков можно разделить на четыре группы:

1) сильно и специфически связанные молекулы воды внутри белка, где молекулы воды составляют элемент структуры;

2) молекулы воды первого гидратного слоя, взаимодействующие с поверхностью белка;

3) молекулы воды, находящиеся в промежуточном состоянии между первой гидратной оболочкой и объемной водой;

4) объемная вода.

Сильно связанные молекулы воды занимают внутренние полости белка и могут быть определены методами кристаллографии и ЯМР. Свойства первого гидратного слоя воды, в основном, определяются свойствами поверхности белка, а также дипольным моментом, поляризуемостью и способностью каждой молекулы воды образовывать водородные связи. По-

верхностные молекулы воды, образующие гидратную оболочку, а также молекулы воды, находящиеся в промежуточном состоянии, подвижнее сильно связанных молекул, их водородные связи могут быть искажены, именно этот тип молекул воды рассматривается в настоящей работе.

Внутренние параметры поверхностной воды, т. е. валентные и торсионные углы связанных молекул, значительно отличаются от параметров объемной воды. В работе [3] с помощью комбинированного малоуглового рассеяния было показано, что для различных белков средняя плотность первого гидратного слоя (0–3 Å от поверхности белка) выше плотности объемной воды. Существуют работы, в которых данный факт подтверждается методами рентгеноструктурного анализа и нейтронного рассеяния, а также численными экспериментами с использованием метода молекулярной динамики. Установлено, что плотность водного слоя белка толщиной до 5 Å на 15% превосходит плотность объемной воды. Традиционно гидратация белка обсуждается в терминах гидрофобности/гидрофильности поверхностных групп белка и количестве молекул связанной воды.

Однако основная причина увеличения плотности — это возмущение сетки водородных связей, состоящее в изменении ее внутренних параметров: уменьшении расстояния O...O и увеличении координационного числа молекулы воды [4–6]. Ранее нами было показано, что структуры гидратной оболочки белков и объемной воды различаются топо-

логически. В структуре гидратной оболочки белков высоко содержание гексациклов и других замкнутых циклов, состоящих из более чем шести молекул воды [7].

Наряду с радиальным распределением плотности интересно также распределение электростатического потенциала в слое гидратной воды, поскольку это именно оно определяет характер межмолекулярного взаимодействия.

Задачей настоящей работы является построение усредненного распределения плотности и электростатического потенциала как функции расстояния от поверхности белка.

Белки разделяются по своей структуре на два больших класса: глобулярные и фибриллярные. Фибриллярные белки представляются собой периодически повторяющиеся спиральные структуры. Примером фибриллярного белка является коллаген. Коллаген — один из самых распространенных и важных структурных белков, он характеризуется специфической последовательностью аминокислотных остатков $[-\text{Gly-X-Y}-]$, в данных обозначениях Gly — глицин, X, Y — чаще всего пролин или гидроксипролин.

Специфическое структурирование воды в проколлагене и коллагеновых фибриллах и связанное с ним формирование нативной структуры коллагена было многократно подтверждено разными методами. С помощью различных экспериментальных техник (например, ЯМР и диэлектрическая релаксация) было показано, что молекулы воды в структурах волокон коллагена менее подвижны, чем в объемной воде [8–11]. Было показано также, что основной вклад в энтальпию тепловой денатурации коллагена дает плавление связанной воды, т. е. связанная вода действительно стабилизирует структуру коллагена [11]. Нами было сделано предположение, что именно в этой структуре топологические паттерны сетки водородных связей могут быть выражены сильнее, чем в водном слое глобулярных белков. По этой причине в данной работе проводилось исследование белков двух типов: глобулярного белка убиквитина, коллагенов и фрагментов коллагенов.

1. Проведение компьютерных экспериментов

Для описания структуры гидратных оболочек белков была использована программа SOLVATE [13]: данная программа в качестве варьируемого параметра использует толщину слоя растворителя, моделирует выпуклую оболочку выбранной толщины вокруг белкового комплекса, который изначально задан структурой, взятой из PDB (международного банка белковых структур высокого разрешения), приближенно вычисляет функцию плотности и на основе полученных данных производит заполнение объема гидратной оболочки молекулами воды с использованием модели воды TIP3P (жесткая трехточечная модель [14]). Таким образом определяется пространственное

расположение связанных молекул воды. Данные PDB-банка представлены для белков в состоянии молекулярных кристаллов, т. е. белков, окруженных минимальным количеством воды, необходимым для поддержания нативной конформации белка. В настоящей работе проводилось изучение структур гидратных оболочек, растворенных в воде фрагментов коллагенов (1CAG.pdb, 1BKV.pdb, 1ITT.pdb), а также убиквитина (1UBQ.pdb).

В качестве первого шага нами были построены гидратные оболочки для вышеперечисленных структур, приведем характеристики структур:

1) 1CAG.pdb — тройная спираль коллагена, определенная с разрешением 1.9 Å, длина каждой цепочки 30 аминокислот, количество молекул растворителя — 10062;

2) 1BKV.pdb — тройная спираль коллагена, определенная с разрешением 2 Å, длина каждой цепочки 30 аминокислот, количество молекул растворителя — 7164;

3) 1ITT.pdb — тройная спираль коллагена, определенная с разрешением 1.9 Å, длина каждой цепочки 7 аминокислот, количество молекул растворителя — 2292;

4) 1UBQ.pdb — белок убиквитин, определенный с разрешением 1.8 Å, количество молекул растворителя — 9630.

В качестве примера структуры гидратных оболочек для фрагмента коллагена 1ITT.pdb в двух проекциях (a — поперечная, b — продольная) изображены на рис. 1.

После «растворения» изучаемых белков в воде обязательным является приведение системы в ненапряженное состояние для сокращения возможного влияния энергетически неоптимального расположения структуры «белок — растворитель». Одной из основных причин появления артефактов является достижение в процессе минимизации методом градиентного спуска (применяемого в нашей работе) точки локального минимума функции потенциальной энергии в зависимости от числа шагов минимизации, но не точки глобального минимума. В случае $|\min - \min_{\text{global}}| \gg \epsilon$ (существенного различия локального и глобального минимумов) после минимизации система будет напряжена (обладать значительным запасом потенциальной энергии), и последующая молекулярная динамика при различных начальных условиях будет приводить к различающимся статистическим характеристикам системы. Для исключения таких неравновесных состояний использовалась следующая последовательность процессов:

1) минимизация потенциальной энергии структуры первичной гидратной оболочки;

2) последующая процедура молекулярной динамики при невысоких (приближенных к нормальным условиям, но с сохранением целостности структуры) температурах, $T \in (50 \div 150)$ К.

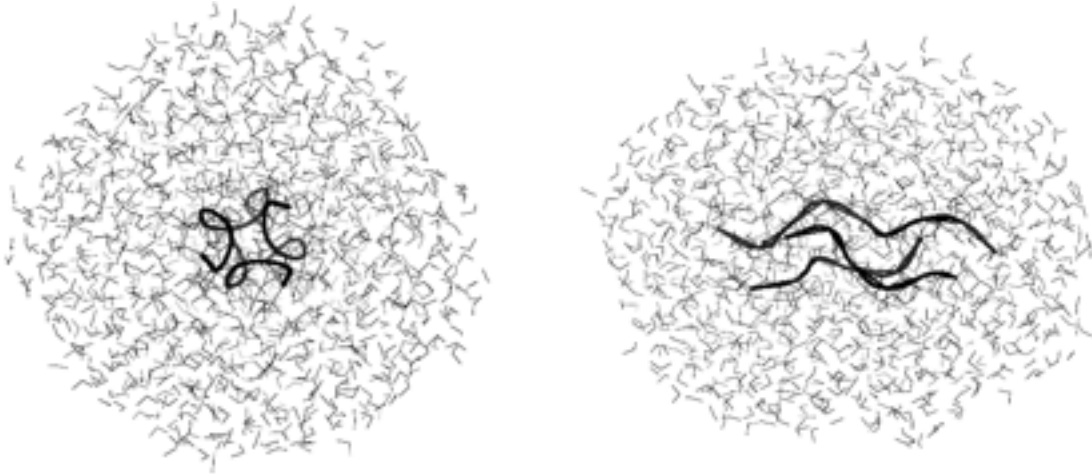


Рис. 1. Гидратная оболочка белка 1TTT.pdb: а — поперечная проекция; б — продольная проекция

Процедура молекулярной динамики не должна осуществляться при больших температурах, так как это может привести к сильно неравновесным структурам, имеющим совсем другие характеристики. Для исключения попадания системы в неглубокий локальный минимум следует чередовать минимизацию энергии с процедурой молекулярной динамики.

Процессы минимизации и молекулярной динамики системы проводили с применением широко используемого программного пакета молекулярной динамики NAMD [15].

После получения квазиравновесного состояния системы «белок — гидратная оболочка» была изучена плотность гидратной воды как функция расстояния от поверхности белка.

Опишем алгоритм нахождения плотности указанной функции.

1. С помощью пакета CGAL [16] формируем внешнюю и внутреннюю оболочки гидратного слоя (на рис. 2 показано, как на невыпуклой оболочке белка находим множество точек, составляющих выпуклую оболочку). Таким образом, получаем слой воды и две поверхности — внутреннюю и внешнюю границы гидратного слоя.

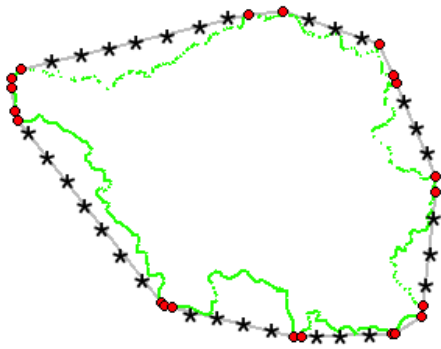


Рис. 2. Приближение замкнутой невыпуклой поверхности тела выпуклой оболочкой (* — выпуклая оболочка исследуемого тела, например белка или гидратного слоя)

2. Находим координаты параллелепипеда, внутри которого можно поместить структуру гидратированного белка.

3. Рассматриваем случайную точку A_{rand} в структуре выделенного параллелепипеда (эта точка не обязательно имеет координаты, совпадающие с каким-либо атомом кислорода).

4. Определяем, находится ли эта точка внутри слоя связанной воды: для этого проверяем, что точка лежит внутри оболочки W , но вне оболочки P .

5. Если точка внутренняя (лежит в гидратном слое), то окружаем A_{rand} сферой радиуса R — в нашем случае в качестве параметра R было выбрано значение $1.5 \text{ \AA}$. Предварительно проверяем, что данная сфера целиком лежит в водном слое.

6. Рассчитываем плотность молекул воды в шаре: $\rho = N_{\text{mol}}/V$, где N_{mol} — количество атомов молекул воды, V — объем шара радиуса R .

7. Под расстоянием от выделенной точки от белка будем понимать $\text{dist} = \min(r(A_{\text{rand}}, C_{\text{protein}}), C_{\text{protein}}) \in \Omega$, где Ω — множество точек белка.

8. Повторяя итерации 1–6 M раз ($M \gg 1$) и усредняя значения ρ для каждого dist , получаем зависимость средней плотности гидратной оболочки как функции расстояния от белка.

9. Погрешность считаем следующим образом. Пусть $g(r)$ — функции распределения плотности.

Рассмотрим значения функций $g_1 = g(r_m)$ и $g_2 = g(r_{m+M})$ — на шаге m и $(m+M)$ соответственно, где M — количество молекул воды в структуре. M выбирается равным количеству молекул воды, чтобы в среднем каждая молекула «оказала влияние» на функцию $g(r_{m+M})$.

Затем рассмотрим невязку по норме пространства $h_{[a,b]}$: $\Delta g = 2/(\|g_1\| + \|g_2\|) \cdot \sqrt{\int (g_1 - g_2)^2 dr}$. Критерием прекращения итерационного процесса является достижение $|\Delta| < \varepsilon$, где $\varepsilon \ll 1$, в нашем случае $\varepsilon = 0.001$. Физический смысл этого критерия: функция распределения плотности выходит на стабильный уровень, и с увеличением числа итераций распределение значительно не меняется.

Для нахождения электрического потенциала как функции расстояния от белка использовался следующий алгоритм: процедура нахождения электрического потенциала как функции расстояния от белка аналогична процедуре нахождения плотности гидратной оболочки, но в каждой точке A_{rand} вычисляем не плотность $\rho = N_{\text{mol}}/V$ из вышеуказанного алгоритма, а потенциал электрического поля $\sum_i \frac{q_i}{r_i}$, где суммирование берется по всем атомам воды для данной точки. Приведем значения зарядов для нашей модели: $q_{\text{oxygen}} = -0.64$, $q_{\text{hydrogen}} = 0.32$.

2. Результаты и обсуждения

На рис. 3 представлены усредненные радиальные распределения плотности гидратных оболочек двух белков с разным количеством чередования процедур минимизации и нагрева ($N = 2$, три раза). Видно качественное различие радиального распределения связанной воды глобулярного и фибриллярного белков. У коллагена наблюдается первый пик на расстоянии $2.9 \text{ \AA}$, а у глобулярного белка положение максимума такое же, но этот пик уширен, наблюдается некоторое количество молекул воды на более близком расстоянии. Скорее всего это обусловлено неоднородностью поверхности глобулярного белка. Во всех радиальных распределениях массовой плотности в гидратных оболочках наблюдаются четко выраженные максимумы, принадлежащие первой, второй, третьей и четвертой гидратным оболочкам

(что соответствует следующим значениям расстояния от поверхности белка: $2.9 \text{ \AA}$, $4.2 \text{ \AA}$, $5.5 \text{ \AA}$, $6.8\text{--}7.1 \text{ \AA}$). Дальнейшие пики, возможно, являются результатом граничных эффектов, поэтому мы их не рассматриваем. Заметим, что второй, третий и четвертый максимумы в гидратной оболочке коллагена слегка смещены на большие расстояния, чем у глобулярного белка. А третий и четвертый максимумы более четко выражены в случае коллагена. Возможно, это связано с различиями в топологии структуры гидратной оболочки, что обнаруживается при сравнении валентных и торсионных углов [7].

На рис. 4, 5 приведены функции электростатических потенциалов как для структуры гидратной оболочки, включающей атомы кислорода и протоны, так и для оболочек, учитывающих только атомы кислорода.

Как видно из полученных результатов, функции электростатического потенциала в гидратных оболочках белков, как и функции плотности, имеют упорядоченную квазипериодическую структуру на расстояниях, включающих более трех молекулярных слоев воды. Полученные функции от структур с атомами кислорода и протонами являются более детальными по сравнению со структурами с учетом только атомов кислорода. Профили потенциала для белка 1TTT.pdb качественно отличаются от профилей для остальных белков. Это отличие объясняется различием топологии, что видно на особенностях

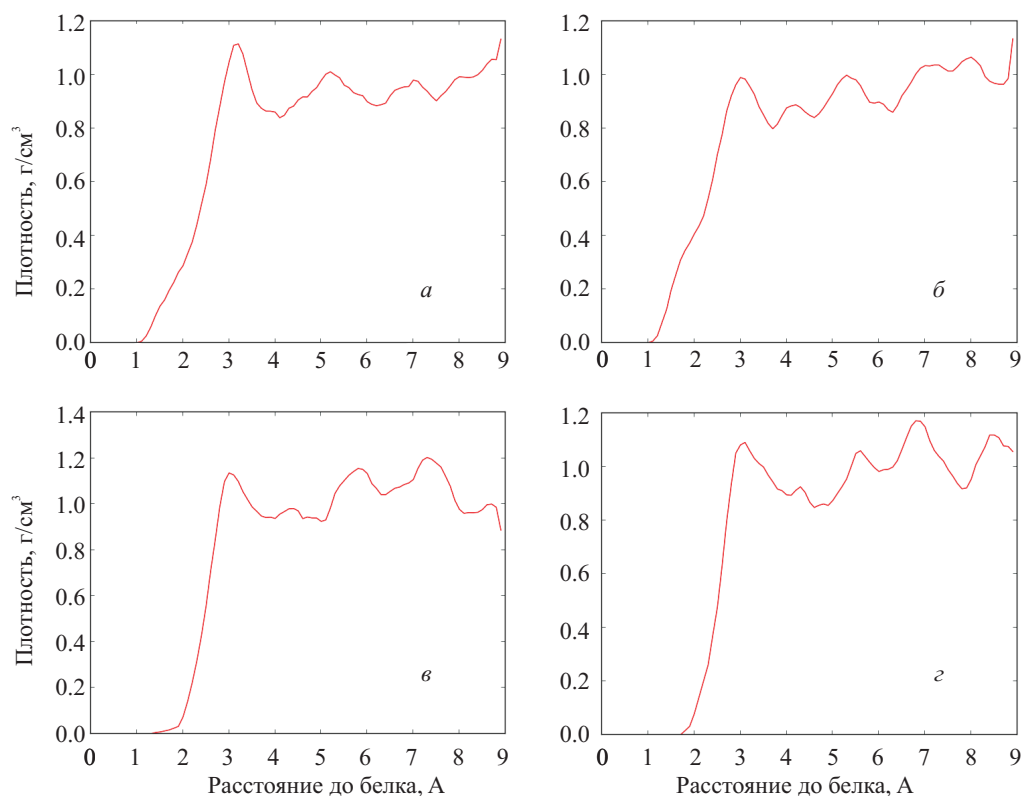


Рис. 3. Функции распределения плотности в гидратных оболочках, N — число чередований процессов минимизации и нагрева: а — 1UBQ.pdb, $N = 3$; б — 1UBQ.pdb, $N = 2$; в — 1CAG.pdb, $N = 3$; г — 1CAG.pdb, $N = 2$

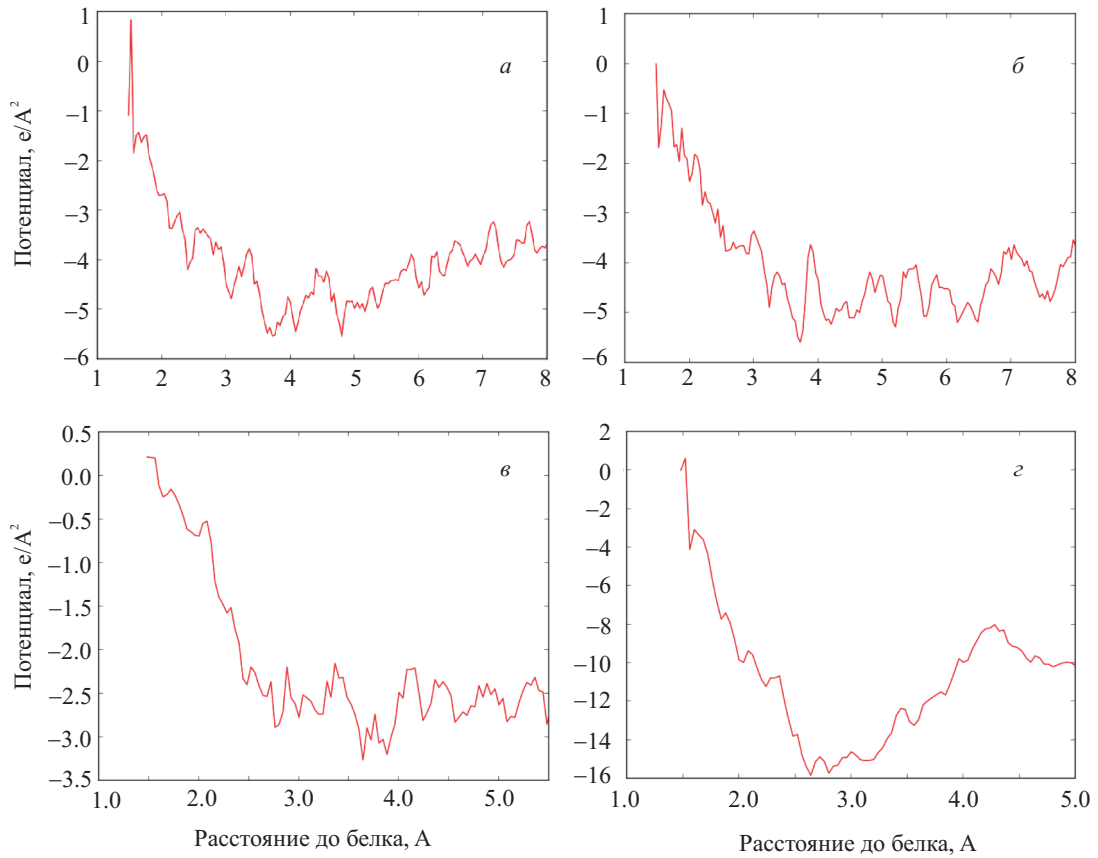


Рис. 4. Функция распределения потенциала, включая как атомы кислорода, так и протоны, для оболочек:
a — 1UBQ.pdb, *б* — 1CAG.pdb, *в* — 1BKV.pdb, *г* — 1ITT.pdb

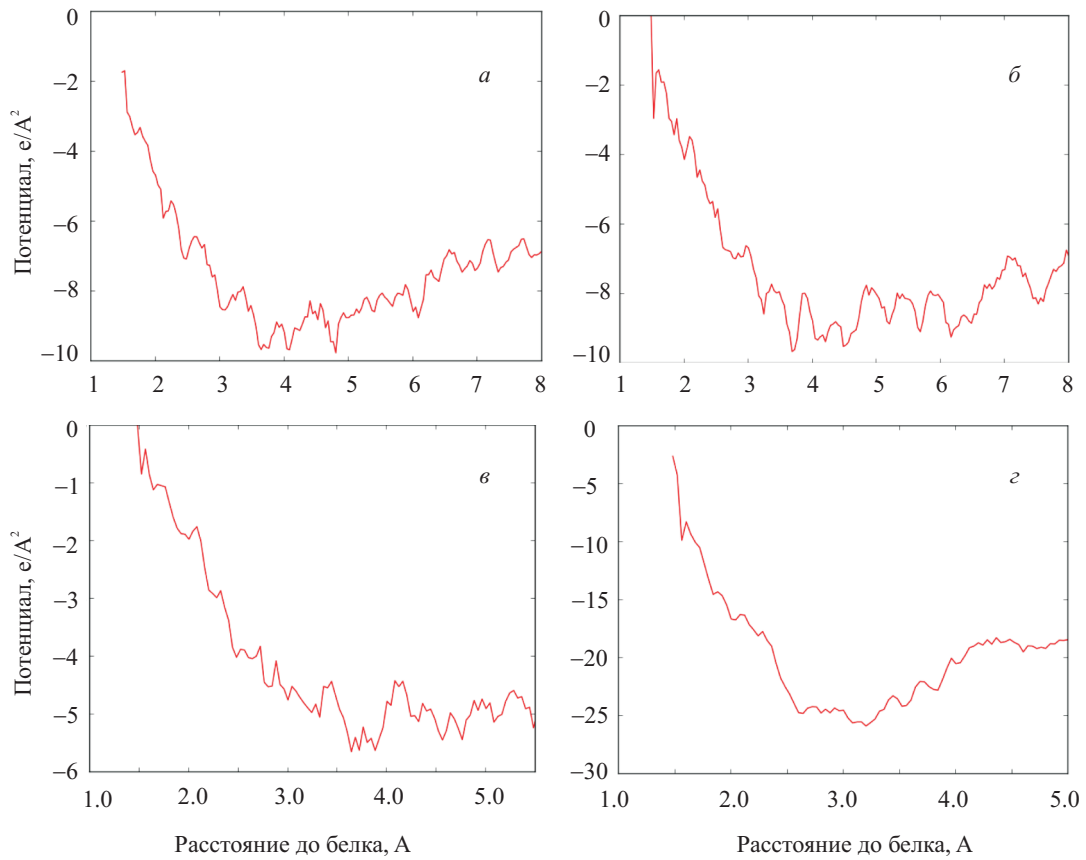


Рис. 5. Функция распределения потенциала, включая только атомы кислорода, для оболочек:
a — 1UBQ.pdb, *б* — 1CAG.pdb, *в* — 1BKV.pdb, *г* — 1ITT.pdb

распределений валентных и торсионных углов по сравнению с другими белками [7].

Заключение

Показано, что распределение плотности в объемном слое гидратной воды указывает на наличие многослойной организации гидратной оболочки рассматриваемых белков. Учитывая выраженную периодичность, можно говорить о структурном упорядочивании на расстояниях от поверхности белка вплоть до 7 Å. Распределения плотности гидратной оболочки глобулярного и фибриллярного белков качественно похожи, однако наблюдаемые небольшие различия функций плотности в гидратных оболочках разных белков объясняются топологическими изменениями, коррелирующими с результатами, полученными нами ранее [7]. Аналогичные выводы делаются на основании анализа функций распределения электростатического потенциала.

Развитием этой работы следует рассматривать более детальное изучение топологических особенностей гидратной оболочки.

Список литературы

1. Kauzmann W. // Adv. Protein Chem. 1959. **14**. P. 1.
2. Mattos C. // Curr. Opin. Struct. Biol. 2001. **11**, N 6. P. 761.
3. Svergun D.I., Richard S., Koch M.H.J. et al. // Proc. Natl. Acad. Sci. 1998. **95**, N 5. P. 2267.
4. Lee S.H., Rossky P.J. // J. Chem. Phys. 1994. **100**, N 4. P. 3334.
5. Friesen A.D., Matyushov D.V. // J. Chem. Phys. 2011. **135**, N 10. P. 104501.
6. Sarupria S., Garde S. // Phys. Rev. Lett.. 2009. **103**, N 3. P. 037803.
7. Rubtsova E.V., Solovei A.B., Lobyshev V.I. // Biofizika. 2014. **59**, N 6. P. 1071.
8. Cheng Y.K., Sheu W.S., Rossky P.J. // Biophys. J. 1999. **76**, N 4. P. 1734.
9. Chapman G.E., Danyluk S.S., McLauchlan K.A. // Proc. of the Royal Soc. of London. Ser. B. Biological Sciences. 1971. **178**, N 1053. P. 465.
10. Bella J., Brodsky B., Berman H. // Structure, 1995.
11. Лобышев В.И., Калинин Л.П. Изотопные эффекты D₂O в биологических системах. М., 1978.
12. Есипова Н.Г., Андреева Н.С., Гатовская Т.В. // Биофизика, 1958. **3**, № 5. С. 529.
13. [Max Planck Institute for Biophysical Chemistry | Research | Research Groups | Theoretical and Computational Biophysics | Research | Methods | Solvate.](#)
14. Mark P., Nilsson L. // J. Chem. Phys. 2001. **105**, N 43. P. 9954.
15. [NAMD — Scalable Molecular Dynamics.](#)
16. [CGAL — Computational Geometry Algorithms Library.](#)
17. Burling F.T., Weis W.I., Flaherty K.M., Brunger A.T. // Science. 1996. **271**, N 5245. P. 72.
18. Levitt M., Sharon R. // Proc. Natl. Acad. Sci. U.S.A. 1988. **85**, N 20. P. 7557.
19. [RCSB Protein Data Bank — RCSB PDB — ICAG Structure Summary.](#)
20. [RCSB Protein Data Bank — RCSB PDB - IUBQ Structure Summary.](#)
21. Moeller T., Trumbore B. // J. Graph. Tools. 1997. **2**, N 1. P. 21.
22. Levy Y., Onuchic J. N. // Annu. Rev. Biophys. Biomol. Struct. 2006. **35**. P. 389.
23. Vitagliano L., Berisio R., De Simone A. // Biophys J. 2011. **100**, N 9. P. 2253.

The statistical characteristics of protein hydration shells: A computer simulation

E. V. Rubtsova^{1,a}, A. B. Solovey^{2,b}, V. I. Lobyshev^{3,c}

¹Department of Mathematics, Faculty of Physics, Lomonosov Moscow State University, Moscow 119991, Russia.

²AESC, Moscow State University, 11 Kremenchugskaya str., Moscow 121357, Russia.

³Department of Biophysics, Faculty of Physics, Lomonosov Moscow State University, Moscow 119991, Russia.

E-mail: ^aev.rubtsova@physics.msu.ru, ^bsoloveybird@gmail.com, ^clobyshev@yandex.ru.

Water molecules and the hydrogen-bond network are essential structural components of many biological macromolecular systems, such as nucleic acids, as well as fibrillar and globular proteins. The water density in the protein hydration shell is known to differ from that of bulk water. In this paper, spatial density as function of the distance from the protein surface is discussed. The electric potential function is considered in the same way. The resulting profiles of the bound water density are characterized by the presence of several distinct maxima due to the existence of the regular structure of bound water at distances over 7 Å. The minor discrepancies that are observed in both radial functions for different proteins are explained by topological variations that were revealed earlier by studying the valence and dihedral angle-distribution functions for water bound to the same proteins.

Keywords: protein hydration shell, structure modeling, density.

PACS: 31.15.-p, 87.15.ap.

Received 15 January 2015.

English version: *Moscow University Physics Bulletin* 5(2015).

Сведения об авторах

1. Рубцова Екатерина Владимировна — аспирантка, e-mail: ev.rubtsova@physics.msu.ru.

2. Соловей Алексей Борисович — кандидат физ.-мат. наук, ст. преподаватель; e-mail: soloveybird@gmail.com.

3. Лобышев Валентин Иванович — доктор физ.-мат. наук, профессор; e-mail: lobyshev@yandex.ru.