

Статистические характеристики гидратных оболочек белков.  
Компьютерное моделирование.  
И.О.Ф.

Аннотация.

Молекулы воды и сетка водородных связей являются необходимым структурным компонентом во многих биологических макромолекулярных системах, например, таких как нуклеиновые кислоты, фибриллярные и глобулярные белки. Известно, что плотность воды в гидратных оболочках белка отличается от плотности объёмной воды [?]. В данной работе мы рассматривали плотность как пространственную функцию удаления от белка. В ходе исследования были найдены характеристические особенности полученной функции плотности, а также дано объяснение указанным особенностям.

Ключевые слова: гидратная оболочка белков, молекулярная динамика, плотность

**1. Введение.** Характеристики физических свойств поверхностей “белок - растворитель” являются существенными для понимания структуры и фолдинга белков. Биологические молекулы растворены в воде, и взаимодействие с белками, аминокислотами и мембранами формирует их структуру и свойства [?].

В отсутствие окружающего водного слоя белки практически неподвижны и ферментативная активность пренебрежимо мала. В работах по дегидратации белков показано, что белок должен быть окружён по крайней мере первой гидратной оболочкой для поддержания своей функциональности [?].

Молекулы воды в растворе белков можно разделить на 3 группы:

- сильно и специфически связанная молекулы воды внутри белка (молекулы воды составляют элемент структуры)
- молекулы воды, взаимодействующие с поверхностью белка
- объёмная вода

Сильно связанные молекулы воды занимают внутренние полости белка и могут быть определены методами кристаллографии. Поверхностные молекулы воды, образующие гидратную оболочку, подвижнее сильно связанных молекул, их водородные связи искажены, именно этот тип молекул исследуется в данной работе.

Свойства слоя воды регулируются дипольным моментом, поляризацией и способностью каждой молекулы образовывать водородные связи. Водные молекулы взаимодействуют с макромолекулами на малых расстояниях ( $\sim \text{\AA}$ ).

В экспериментальной работе [?] с помощью рентгеновского и нейтронного рассеяния установлено, что на поверхности белка плотность воды значительно выше плотности объёмной воды. Также в ряде численных работ по моделированию установлено, что плотность слоя толщины  $N \text{\AA}$  на 15

Показано, что различия в плотности первой гидратной оболочки по сравнению с объёмной водой определяются топологическими и электростатическими свойствами поверхностей белков [?]. В среднем, более плотная

вода находится в сжатом состоянии, при котором водные диполи стремятся выровняться параллельно друг другу электростатическим полем, генерируемым атомами белка.

Характеристики гидратации белков существенны для понимания их структуры и функций. Эти характеристики требуют выяснения взаимного влияния атомов белка и молекул воды. Для термодинамического описания гидратации белка – необходимо получить более глобальную картину, в которой растворитель описан в терминах вероятностных распределений. Известный метод для получения радиальных функций распределения в этом направлении – это молекулярное моделирование.

Внутренние параметры поверхностной воды значительно отличаются от объёмной. В работе [?] было с помощью комбинированного малоуглового рассеяния показано, что для различных белков средняя плотность первого гидратного слоя (0–3 Å от поверхности белка) выше плотности объёмной воды. Существуют работы, в которых данный факт подтверждается численными экспериментами (с помощью метода молекулярной динамики), а также методами кристаллографии.

В работе [?] объясняются физические причины этого эффекта с помощью численных экспериментов. Традиционно гидратация белка обсуждается в терминах гидрофобности/гидрофильности поверхностных групп. Но в работе [?] указывается, что здесь имеет место изменения в поверхностной топологии. Изменение топологии сетки водородных связей, а так же электростатическое поле, генерируемое атомами белка, определяют повышенную плотность гидратной оболочки. Наиболее интересными для нас являются: функция распределения плотности и плотности электрического заряда в гидратных оболочках различных белков, а также сравнение полученных функций с аналогичными для объёмной воды.

**2. Проведение компьютерных экспериментов.** В данной работе проводилось изучение структур:

- 1CAG.pdb – тройная спираль коллагена, определённая с разрешением 1.9 Å
- 1UBQ.pdb – белок убиквитин, определённый с разрешением 1.8 Å
- water ball - шарообразное скопление молекул воды (3186 молекул), данная структура является аналогом объёмной воды

Для создания структуры water ball и гидратных оболочек белков была использована программа SOLVATE [?].

После растворения изучаемых белков в воде необходимо привести систему в ненапряжённое состояние, сократить возможные влияния энергетически неоптимального расположения структуры “белок-растворитель”. Для наших целей подходит такая последовательность процессов:

1. минимизация потенциальной энергии структуры
2. дальнейшая молекулярная динамика при невысоких (приближенных к нормальным условиям, но с сохранением целостности структуры) температурах
3. для наших целей необходимо было взять несколько структур и сравнить их характеристики. Это необходимо для понимания: какие из

особенностей статистических характеристик являются значимыми и воспроизводимыми, и какие – артефактами

Одной из основных причин появления артефактов является достижение в процессе минимизации методом градиентного спуска (применяемого в нашей работе) точки локального минимума  $MIN_1$  (рис. ??, точка А), но не точки глобального минимума  $MIN_{ABS}$ . Таким образом, в случае  $|MIN_1 - MIN_{ABS}| \gg \epsilon$  (существенного различия локального и глобального минимумов) после минимизации система будет напряжена (обладать существенным запасом потенциальной энергии), и последующая молекулярная динамика при различных начальных условиях будет приводить к различным статистическим характеристикам системы. Для уменьшения влияния этого факта необходимо чередовать минимизацию для преодоления барьера АВ.

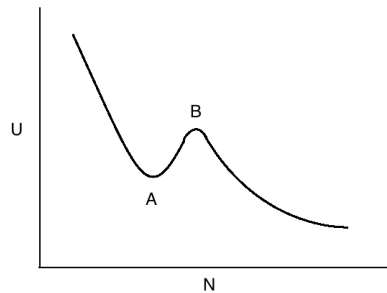


Рис. 1: Зависимость потенциальной энергии системы, т. А - локальный минимум

Процессы минимизации и молекулярной динамики системы проводились с использованием программного пакет NAMD [?]

**2. 1 Проведение компьютерных экспериментов.** Была изучена плотность не только как усреднённая величина, относящаяся к системе множества молекул, но и как функция расстояния до поверхности белка.

Алгоритм нахождения плотности гидратной оболочки как функции минимального расстояния до молекулы белка:

1. С помощью пакета `pyhull.py` [?] приближаем оболочку водного слоя симплексами, полученную оболочку обозначим как S
2. Находим  $x_i^{min}$  и  $x_i^{max}$ ,  $i = 1..3$  (т.е. находим внешних параллелепипед, внутри которого лежит структура «белок + вода»)
3. Определяем случайную точку  $A_{RAND}$  в структуре (эта точка не обязательно имеет координаты, совпадающие с каким-либо атомом кислорода).
4. Определяем, является ли точка внутренней, т.е. лежит ли она внутри оболочки S. Для этой задачи используем алгоритм Möller-Trumbore [?].

5. Если точка внутренняя, то окружаем  $A_{RAND}$  кубом  $B_1..B_8$  с длиной ребра 1 таким, что длина ребра  $B_1B_2 = 1 \text{ \AA}$ ,  $A_{RAND}$  - центр куба  $B_1..B_8$ . Предварительно проверяем, что данный куб целиком лежит в водном слое (является внутренней областью)
6. Плотность молекул воды в кубе:  $\rho = N_{MOL}/V$ , где  $N_{MOL}$  - количество атомов молекул воды (считаем, что центр масс молекулы воды совпадает с координатами атома кислорода)
7. Под расстоянием по оси абсцисс будем понимать  $dist = \min(r(A_{RAND}, C_{protein}), C_{protein} \in \Omega)$ , где  $\Omega$  - множество точек белка. Таким образом получаем пары  $(x, y) = (dist, \rho)$
8. Повторяя итерации  $N$  раз ( $N \gg 1$ ), и усредняя значения  $\rho$  для каждого  $dist$  - получаем зависимость "средней плотности гидратной оболочки как функции.
9. Ошибку считаем следующим образом: пусть  $g(r)$  - функции распределения плотности.
  - (a) Рассмотрим значения функции  $g_1 = g(r_n)$  и  $g_2 = g(r_{n+N})$  - функции на шаге  $n$  и  $(n+N)$  соответственно, где  $N$  - количество молекул воды в структуре.  $N$  выбирается равным количеству молекул воды, чтобы в среднем каждая молекула «оказала влияние» на функцию  $g(r_{n+N})$
  - (b) Затем рассмотрим невязку по норме пространства  $h_{[a,b]} : \Delta g = 2/(\|g_1\| + \|g_2\|) * \sqrt{\int (g_1 - g_2)^2 dr}$ . Критерием прекращения итерационного процесса является достижение  $|\Delta| < \epsilon$ , где  $\epsilon \ll 1$ , в нашем случае  $\epsilon = 0.001$ . Физический смысл этого критерия: функция распределения плотности вышла на стабильный уровень, и с увеличением числа итераций - распределение значительно не меняется.

Алгоритм нахождения плотности объёмной воды как функции координаты  $r$ :

- данный алгоритм полностью повторяет описанный выше алгоритм для гидратной оболочки за исключением физического смысла оси абсцисс.
- определим ось абсцисс: шар воды вписан в прямой параллелепипед. Найдём центр этого параллелепипеда  $C_{CENTER}$  (точка пересечения диагоналей), под координатой  $r$  будем понимать расстояние от  $A_{RAND}$  до  $C_{CENTER}$ .

Функция распределения плотности для объёмной воды (рис. 1):

Распределение отрицательного электрического заряда (плотность атомов кислорода) свидетельствует о неслучайном распределении молекул. Учитывая тот факт, что они находятся достаточно далеко от поверхности белка (10Å - третья - четвертая (вторая-третья)? гидратная оболочка) можно говорить об электрическом и структурном упорядочивании...? Силовое поле гидратной оболочки определено структурой белка. Малые молекулы

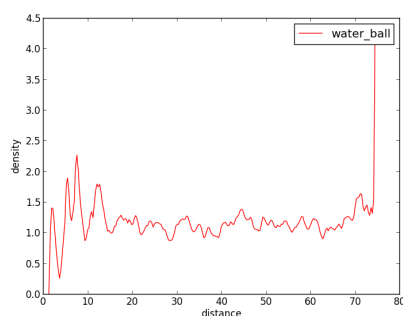


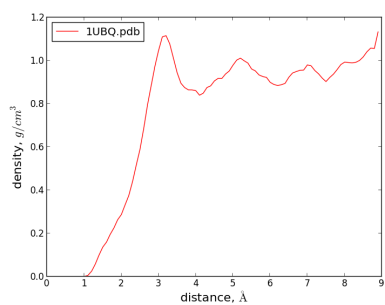
Рис. 2: Функция распределения плотности для объёмной воды

(для которых этот белок является ферментом), а также макромолекулярные структуры, присутствующие в клетке в случае структурного белка реагируют именно на силовое поле. Возможно, заданное силовое поле гидратной оболочки является структурным паттерном как для структурных и регуляторных белков, так и для ферментов. Кулоновское дальнеедействие гидратной оболочки, выстроенной благодаря водородным связям молекул воды и молекул белка, а также молекул воды между собой может обеспечивать ферментативное соответствие.

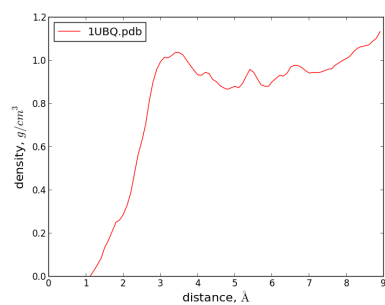
## Список литературы

- [1] D. I. Svergun, S. Richard, M. H. J. Koch, Z. Sayers, S. Kuprin, and G. Zaccai. Protein hydration in solution: Experimental observation by x-ray and neutron scattering // Proc. Natl. Acad. Sci. 1998. T. 95. № 5. С. 2267–2272.
- [2] Kauzmann W. Some factors in the interpretation of protein denaturation. // Adv. Protein Chem. 1959. T. 14. С. 1–63.
- [3] Mattos C. Proteins in organic solvents // Curr. Opin. Struct. Biol. 2001. T. 11. № 6. С. 761–764.
- [4] Burling F.T. и др. Direct Observation of Protein Solvation and Discrete Disorder with Experimental Crystallographic Phases // Science (80-. ). 1996. T. 271. № 5245. С. 72–77.
- [5] Lee S.H., Rossky P.J. A comparison of the structure and dynamics of liquid water at hydrophobic and hydrophilic surfaces—a molecular dynamics simulation study // J. Chem. Phys. 1994. T. 100. № 4. С. 3334.
- [6] Levitt M., Sharon R. Accurate simulation of protein dynamics in solution. // Proc. Natl. Acad. Sci. U. S. A. 1988. T. 85. № 20. С. 7557–61.
- [7] Cheng Y.K., Sheu W.S., Rossky P.J. Hydrophobic hydration of amphipathic peptides. // Biophys. J. 1999. T. 76. № 4. С. 1734–43.
- [8] RCSB Protein Data Bank - RCSB PDB - 1CAG Structure Summary. URL: <http://www.rcsb.org/pdb/explore.do?structureId=1CAG>

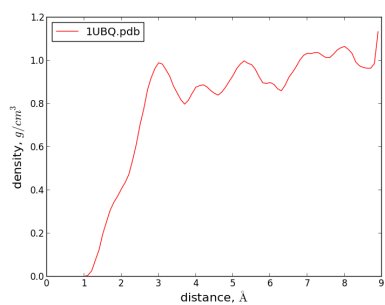
- [9] RCSB Protein Data Bank - RCSB PDB - 1UBQ Structure Summary. URL:  
<http://www.rcsb.org/pdb/explore/explore.do?structureId=1UBQ>
- [10] Max Planck Institute for Biophysical Chemistry | Research | Research  
 Groups | Theoretical and Computational Biophysics | Research | Methods  
 | Solvate URL: <http://www.mpibpc.mpg.de/grubmueller/solvate>
- [11] NAMD - Scalable Molecular Dynamics URL:  
<http://www.ks.uiuc.edu/Research/namd/>
- [12] pyhull Package — pyhull 1.4.3 URL:  
<https://pythonhosted.org/pyhull/pyhull.html>
- [13] Moeller T., Trumbore B. Fast, Minimum Storage Ray-Triangle Intersection  
 // J. Graph. Tools. 1997. T. 2. № 1. C. 21–28.



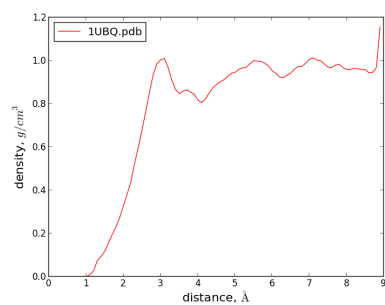
a)



b)

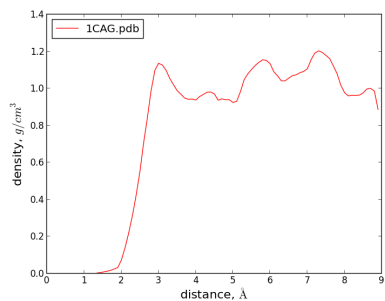


c)

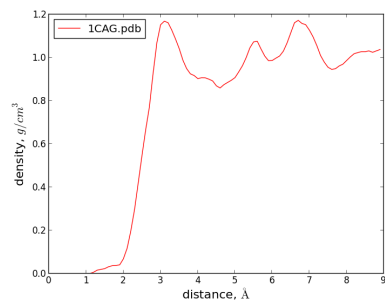


d)

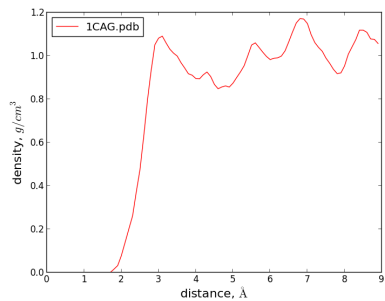
Рис. 3: Функции распределения плотности в гидратных оболочках 1UBQ



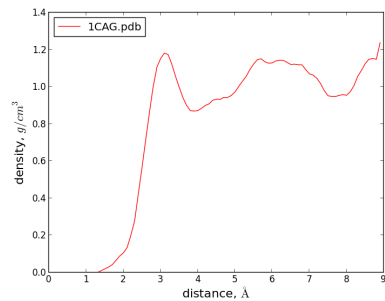
a)



b)

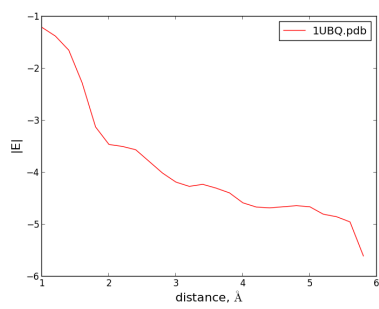


c)

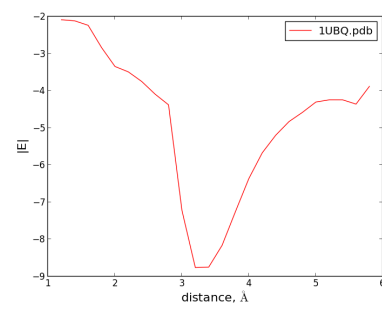


d)

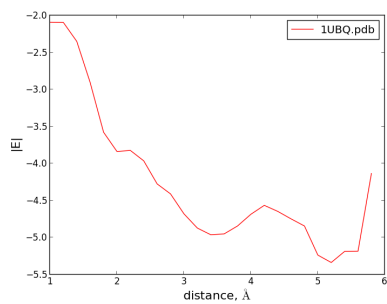
Рис. 4: Функции распределения плотности в гидратных оболочках 1CAG



a)



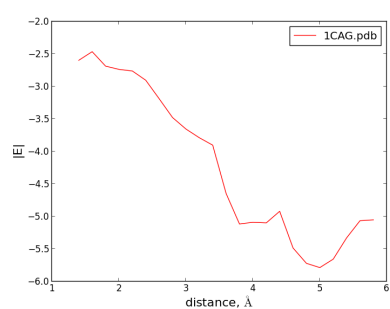
b)



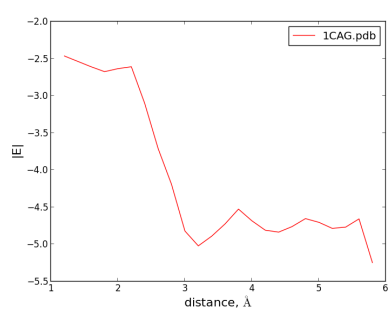
c)

Рис. 5: Функции распределения плотности отрицательных зарядов в гидратных оболочках 1UBQ

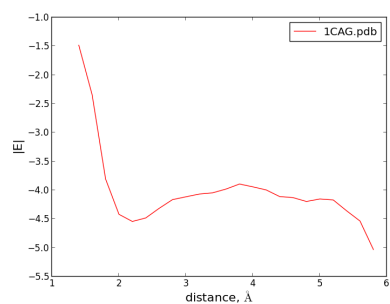




a)



b)



c)

Рис. 6: Функции распределения плотности отрицательных зарядов в гидратных оболочках 1CAG