

УДК: 577.3, 544.35

## Усредненный по молекулярным траекториям метод интегральных уравнений в приближении RISM

Д.А. Тихонов\*, Е.В. Соболев\*\*

*Институт математических проблем биологии, Российская академия наук,  
Пушино, Московская область, 142290, Россия*

**Аннотация.** Предложен новый подход к исследованию влияния растворителя на молекулярные системы, основанный на теории интегральных уравнений теории жидкостей в приближении RISM. Получено новое интегральное уравнение, в котором матрица внутримолекулярных корреляций усредняется по молекулярным траекториям, рассчитываемым методом молекулярной динамики. Такой способ позволяет учитывать подвижность растворенных молекул при оценках эффекта сольватации. Выполнено сравнение результатов полученных в рамках нового подхода со стандартными расчетами средних по траектории термодинамических величин на примере пептида окситоцина. Показано, что все структурные и термодинамические параметры сольватации пептида, полученные в рамках нового метода, лежат в интервале статистической погрешности при усреднении по траектории.

**Ключевые слова:** гидратация макромолекул, стационарная динамическая траектория, термодинамика, интегральные уравнения теории жидкостей, RISM.

### ВВЕДЕНИЕ

Метод интегральных уравнений теории жидкостей в приближении RISM (Reference Interaction Site Model) [1] является одним из теоретических методов для исследования сольватации молекул. В приближении бесконечного разбавления [2], когда плотностью раствора мы можем пренебречь, метод RISM позволяет найти атом-атомные корреляционные функции раствор-растворитель и по ним рассчитать термодинамику сольватации, в частности энергию Гиббса. «Входными данными» теории RISM, помимо потенциалов атом-атомного взаимодействия, является геометрия молекулы раствора, а точнее, матрица межатомных расстояний. Таким образом, корреляционные функции и термодинамика сольватации, найденные как решения уравнений RISM, сопоставляются единственной жесткой геометрии растворенной молекулы. Известно, однако, что большие молекулы при комнатной температуре обладают значительной подвижностью, и даже в равновесном состоянии их геометрия постоянно изменяется, а значит, это необходимо учитывать при анализе термодинамики сольватации.

Можно учесть влияние флуктуаций геометрии молекулы на термодинамику сольватации, если выполнить RISM-расчет для каждой геометрической конфигурации молекулы, а затем вычислить среднее значение энергии Гиббса по всей динамической траектории. Геометрические конфигурации легко получить при помощи молекулярно-динамического моделирования. Такой трудоемкий способ использован в работе [3] для

\* dmitry.tikhonov@gmail.com

\*\* egor@impb.psn.ru

исследования влияния различных способов учета водного окружения на термодинамику гидратации пептида окситоцина. В настоящей работе мы предлагаем другой способ учета конформационной подвижности растворенной молекулы в теории RISM.

В данной работе для случая бесконечного разбавления получена единственная система усредненных по молекулярно-динамической траектории уравнений RISM. Решения этой системы уравнений близки к средним по траектории корреляционным функциям, а рассчитанные по этим функциям термодинамические величины близки к средним по траектории величинам. Предлагаемый способ оценки средней энергии Гиббса по равновесной траектории является полезным и практичным, поскольку исключает многократное решение системы уравнений RISM в каждой точке траектории. Однако априори не ясно, как термодинамические величины, полученные по усредненным уравнениям, будут соответствовать настоящим средним термодинамическим величинам.

Энергия Гиббса и избыточный химический потенциал являются одними из наиболее важных термодинамических величин. В методе интегральных уравнений для простых жидкостей, или теории Орнштейна–Цернике, существует способ вычислить явное выражение для функционала избыточного химического потенциала по формуле Морита–Хироке [4], которые получены аналитически методом термодинамического интегрирования энергии в общем виде. Обобщение этой формулы на случай многоатомных молекул в рамках RISM-формализма выполнено в работе [5]. Однако получить простое выражение функционала избыточного химического потенциала можно только в том случае, если есть простое аналитическое выражение для нелинейной части уравнения RISM, так называемого уравнения замыкания. Иначе необходимо выполнить термодинамическое интегрирование численно. В случае усредненных по траектории уравнений RISM численно интегрировать энергию необходимо только один раз. Это гораздо проще, чем выполнить интегрирование обычного уравнения RISM в каждой точке траектории и затем усреднить результат. Таким образом, метод усредненных уравнений позволяет значительно расширить круг используемых функционалов для оценок энергии Гиббса. Это преимущество метода является дополнительным мотивом для его разработки.

Из-за трудностей вычислительного характера вместо процедуры численного интегрирования используют приближенные оценки по термодинамической теории возмущений. В работе [6] предложено замыкание, которое лучше других учитывает энтропийный вклад в энергию Гиббса. Вводится дополнительная известная функция, так называемая отталкивательная поправка к мостиковому функционалу, которая увеличивает жесткость взаимодействия молекулы с растворителем. Там же предложена формула избыточного химического потенциала с учетом этой поправки, полученная в первом порядке термодинамической теории возмущений. Одна из целей нашей работы — проверка того, насколько оценки энергии Гиббса по теории возмущений близки к результатам численного интегрирования.

## МЕТОДОЛОГИЯ

### 1. Уравнения RISM в пределе бесконечного разбавления

Уравнение для бесконечно разбавленного раствора [2], когда парциальной плотностью растворенного вещества можно пренебречь, связывают прямые  $c_{i\alpha}(r)$  и полные  $h_{i\alpha}(r)$  корреляционные функции, отражающие распределение атома  $\alpha$  молекул растворителя ( $\alpha = 1 \dots M$ ) вокруг атома  $i$  растворенной молекулы ( $i = 1 \dots N$ ):

$$\hat{h}^{uv} = \hat{w}^u \hat{c}^{uv} (\hat{w}^v + \rho \hat{h}^{vv}), \quad (1)$$

где

$$\hat{h}_{i\alpha}^{uv} = \hat{h}_{i\alpha}^{uv}(k) = \frac{4\pi}{k} \int_0^\infty r dr \sin(kr) h_{i\alpha}^{uv}(r),$$

$$\hat{c}_{i\alpha}^{uv} = \hat{c}_{i\alpha}^{uv}(k) = \frac{4\pi}{k} \int_0^\infty r dr \sin(kr) c_{i\alpha}^{uv}(r).$$

Элементы матрицы внутримолекулярных корреляционных функций  $\hat{w}^u$  задают жесткие связи между атомами внутри растворенной молекулы [1]:

$$\hat{w}_{ij}^u = \hat{\omega}_{ij}^u(k) = \delta_{ij} + (1 - \delta_{ij}) \frac{\sin kl_{ij}}{kl_{ij}}, \quad (2)$$

где  $l_{ij}$  — расстояние между атомами  $i$  и  $j$ ,  $\delta_{ij}$  — символ Кронекера.

Выражение в скобках в уравнении (1) — это статическая восприимчивость растворителя. Она включает в себя матрицу внутримолекулярных корреляционных функций, задающих жесткие связи между атомами в молекуле растворителя  $\hat{w}^v$ , плотность растворителя  $\rho$  и матрицу полных корреляционных функций чистого растворителя  $\hat{h}^{vv}$ . Диагональные элементы статической восприимчивости равны статическому структурному фактору. Это выражение является константой и может быть либо предварительно рассчитано методом RISM, либо взято из экспериментальных данных. Остается добавить, что размерность матрицы  $\hat{w}^u$  равна числу атомов молекулы раствора  $N$ , и размерность матрицы  $\hat{w}^v$  — числу атомов молекулы растворителя  $M$ . Матрицы корреляционных функций, соответственно, прямоугольные  $N \times M$ .

Вторым уравнением помимо (1), которое связывает корреляционные функции в координатном пространстве и обеспечивает разрешимость системы относительно корреляционных функций, является уравнение замыкания. В качестве уравнения замыкания мы будем использовать гиперцепное замыкание (HNC) [7]:

$$c_{i\alpha} = \exp(-\beta u_{i\alpha} + \gamma_{i\alpha}) - 1 - \gamma_{i\alpha} \quad (3)$$

и его частично линеаризованный вариант (PLHNC) [9]:

$$c_{i\alpha} = \exp(-\beta u_{i\alpha}) F(\gamma_{i\alpha}) - 1 - \gamma_{i\alpha}, \quad (4)$$

где  $\gamma_{i\alpha}(r) = h_{i\alpha}(r) - c_{i\alpha}(r)$  — не прямые корреляционные функции,  $u_{i\alpha}(r)$  — парные потенциалы взаимодействия атома  $i$  молекулы раствора и атома  $\alpha$  молекулы растворителя,  $\beta = 1/kT$  — обратная температура, и

$$F(\gamma_{i\alpha}) = \begin{cases} \exp(\gamma_{i\alpha}), & \text{если } h_{i\alpha}(r) \leq 0, \\ 1 + \gamma_{i\alpha}, & \text{если } h_{i\alpha}(r) > 0. \end{cases}$$

## 2. Молекулярно-динамические траектории и интегральные уравнения

Уравнения RISM позволяют получить корреляционные функции для конкретной геометрической конфигурации растворенной молекулы, точнее, для заданной матрицы межатомных расстояний. Однако в растворе молекула претерпевает существенные конформационные изменения, следовательно, ограничиваться единственным набором межатомных расстояний было бы неверным. Одним из способов учесть подвижность молекулы на равновесном участке траектории является простое усреднение корреляционных функций по геометрическим конфигурациям.

Пусть траектория задана в виде наборов мгновенных межатомных расстояний растворенной молекулы  $\{l_{ij}(t)\}$ , где  $t=1, \dots, N_T$ . Обозначив статическую восприимчивость растворителя буквой  $\mathbf{v}$ , а также  $\mathbf{h} = \text{vec}(\hat{\mathbf{h}}^{uv})$  и  $\mathbf{c} = \text{vec}(\hat{\mathbf{c}}^{uv})$ , где операция  $\text{vec}(\mathbf{x})$  обозначает преобразование матрицы к вектору-столбцу путем последовательной записи столбцов матрицы  $\mathbf{x}$  друг под другом, запишем уравнение (1) для каждой конфигурации  $t$  в виде:

$$\mathbf{h}_t = \mathbf{A}_t \mathbf{c}_t = (\mathbf{v} \otimes \hat{\mathbf{w}}_t^u) \mathbf{c}_t, \quad t=1, \dots, N_T, \quad (5)$$

где  $\otimes$  — прямое произведение матриц, а индексом  $t$  обозначены величины, зависящие от геометрии молекулы. Таким образом, матрица  $\mathbf{A}_t$  в этой записи соответствует набору расстояний  $\{l_{ij}(t)\}$ .

Если получены решения уравнений (5) для всех геометрий  $t$ , тогда среднее значения вектора полных корреляционных функций выражается формулой:

$$\bar{\mathbf{h}} = \frac{1}{N_T} \sum_{t=1}^{N_T} \mathbf{h}_t = \frac{1}{N_T} \sum_{t=1}^{N_T} \mathbf{A}_t \mathbf{c}_t. \quad (6)$$

Выражение (6) имеет тот недостаток, что для вычисления среднего необходимо знать решения уравнений RISM в каждой точке молекулярно-динамической траектории. Введем другое определение среднего вектора полной корреляционной функции:

$$\langle \mathbf{h} \rangle = \left( \frac{1}{N_T} \sum_{t=1}^{N_T} \mathbf{A}_t \right) \left( \frac{1}{N_T} \sum_{t=1}^{N_T} \mathbf{c}_t \right).$$

Если ввести формальное определение среднего вектора прямых корреляционных функций

$$\langle \mathbf{c} \rangle = \frac{1}{N_T} \sum_{t=1}^{N_T} \mathbf{c}_t,$$

то мы приходим к единственному уравнению для таких средних величин:

$$\langle \mathbf{h} \rangle = \bar{\mathbf{A}} \langle \mathbf{c} \rangle, \quad (7)$$

где  $\bar{\mathbf{A}}$  — средняя матрица уравнения RISM:

$$\bar{\mathbf{A}} = \frac{1}{N_T} \sum_{t=1}^{N_T} \mathbf{A}_t. \quad (8)$$

Будем называть решения уравнений (7) псевдосредними корреляционными функциями.

Преимущества уравнений (7) при вычислении средних корреляционных функций очевидны. Требуется решить только одно уравнение RISM, но с усредненной по всей траектории матрицей внутримолекулярных корреляций  $\bar{\mathbf{A}}$ . Однако остается неясным, насколько полученные таким способом псевдосредние функции отличаются от истинных средних корреляционных функций. Попытаемся выяснить этот вопрос априори, вычислив разность уравнений (6) и (7). После несложных преобразований выражения для разности получим:

$$\Delta = \bar{\mathbf{h}} - \langle \mathbf{h} \rangle = \frac{1}{N_T^2} \sum_{t=1}^{N_T} \sum_{s=1}^{N_T} (\mathbf{A}_t - \mathbf{A}_s) \mathbf{c}_s. \quad (9)$$

Разницу (9) можно просуммировать:

$$\Delta = \bar{\bar{\mathbf{A}}} \mathbf{c} - \bar{\mathbf{h}}.$$

Таким образом, разница между средней полной корреляционной функцией и полученной из уравнения (7) полной псевдосредней корреляционной функцией

выражается через средние по траектории величины как невязка линейного уравнения со средней матрицей (7), если в качестве пробных функций использованы средние по траектории прямая и полная корреляционные функции. Из общих соображений понятно, что такая невязка может быть нулевой только в том случае, если все матрицы  $A_t$  равны их среднему значению  $\bar{A}$ .

Из выражения (9) следует, что разница  $\Delta$  будет тем меньше, чем меньше разница внутримолекулярных корреляционных матриц для различных точек траектории. Стоит ожидать, что для типичных равновесных траекторий молекулярных систем это правило будет выполняться. Действительно, выделим из уравнения (9) среднее по траектории значение разности внутримолекулярной матрицы для некоторой точки траектории  $s$ :

$$\Delta \bar{A}_s = \frac{1}{N_T} \sum_{t=1}^{N_T} (A_t - A_s).$$

Если молекулярная система испытывает тепловые флуктуации вблизи некоторого равновесного состояния, то следует ожидать, что величина  $\Delta \bar{A}_s$  будет близка к нулю, а, следовательно, будет близко к нулю и среднее значение произведения  $\Delta \bar{A}_s$  на прямую корреляционную функцию

$$\Delta = \frac{1}{N_T} \sum_{s=1}^{N_T} \Delta \bar{A}_s c_s.$$

Анализ линейной части уравнения RISM не позволяет нам дать точных априорных оценок величины погрешности уравнения со средней матрицей (7), кроме самых общих. Тем более трудно учесть влияние на погрешность со стороны нелинейной части уравнений RISM, а именно — уравнений замыкания (3) или (4). Необходимо провести тестовые расчеты на конкретных примерах молекулярных траекторий, чтобы выяснить, насколько малой является величина  $\Delta$  в сравнении, например, с дисперсией средних величин корреляционных функций. В том случае, когда погрешность  $\Delta$  оказывается меньше дисперсии, вполне возможно заменить вычисление средних по траектории на решение уравнения со средней матрицей. В таком случае метод нужно признать практически применимым.

Далее в изложении мы везде полагаем матрицу растворителя  $\nu$  постоянной, и говорим лишь об изменении внутримолекулярной корреляционной матрицы  $\nu^u$ . Тогда средняя матрица уравнения RISM имеет вид:

$$\bar{A} = \nu \otimes \left( \frac{1}{N_T} \sum_{t=1}^{N_T} \hat{w}_t^u \right) = \nu \otimes \bar{w}.$$

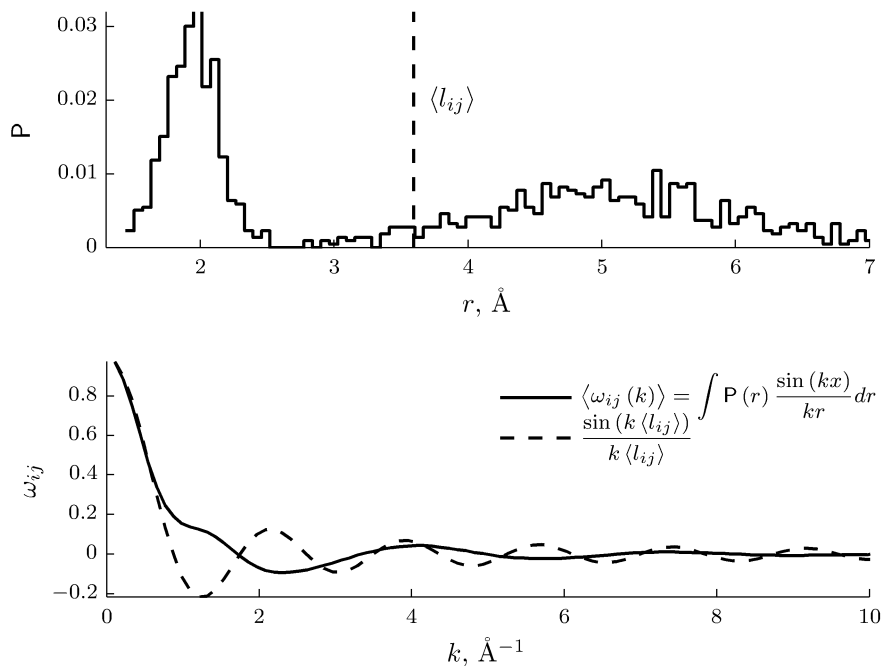
Если же сами функции водного растворителя подвержены значительным флуктуациям, то усреднять нужно матрицу  $A$ . Для водного растворителя при обычных условиях матрица растворителя  $\nu$  является постоянной.

### 3. Детали усреднения матрицы внутримолекулярных корреляций

Остановимся подробнее на вычислительных аспектах получения средней матрицы внутримолекулярных корреляций. Сначала рассмотрим простой случай, когда матрица  $\bar{w}$  выражена через усредненные по траектории межатомные расстояния растворенной молекулы  $\langle l_{ij} \rangle$  по формуле для жестких связей (2):

$$\bar{\omega}_{ij}(\langle l_{ij} \rangle, k) = \delta_{ij} + (1 - \delta_{ij}) \frac{\sin(k \langle l_{ij} \rangle)}{k \langle l_{ij} \rangle}, \quad \langle l_{ij} \rangle = \frac{1}{N_T} \sum_{t=1}^{N_T} l_{ij}(t). \quad (10)$$

Несомненным достоинством определения (10) является параметрическая зависимость функций от скалярной величины среднего межатомного расстояния, что



**Рис. 1.** Внутримолекулярные корреляционные функции, полученные по средним межатомным расстояниям и по гистограммам распределения межатомных расстояний. Сравнение методов усреднения.

позволяет нам организовать компактное хранение матрицы  $w$  в памяти компьютера. Хранение матрицы  $w$  становится одной из проблем численной реализации метода, поскольку для молекул с числом атомов порядка тысячи объем необходимой памяти составляет десятки гигабайт. В случае параметрической зависимости (10) сократить требуемый объем памяти можно с помощью двух простых соображений. Во-первых, в большой глобулярной молекуле много приблизительно одинаковых расстояний, так как длины ковалентных связей могут варьироваться лишь в узких пределах. Округлив расстояния до некоторой наперед заданной точности, сведем весь набор межатомных расстояний к небольшому набору неэквивалентных расстояний [9, 10]. Во-вторых, для далеких друг от друга атомов можно ослабить внутримолекулярные корреляции, так как взаимное влияние таких атомов мало. Такое ослабление корреляций позволяет хранить более короткие внутримолекулярные корреляционные функции [11]. В алгоритмическом смысле расчеты со средней геометрией не отличаются от расчетов с жесткой геометрией, и оба эти случая могут быть реализованы в рамках одной программы с входными данными в виде матрицы межатомных расстояний.

Хотя метод (10) является очень простым в реализации, наши расчеты показывают, что он приводит к весьма бедным результатам. Корреляционные функции не совпадают со средними величинами, и часто содержат артефакты в виде паразитных пиков на больших расстояниях. Термодинамика удовлетворительно совпадает со средними величинами только в случае очень жестких и небольших молекул. Для конформационно подвижных молекул этот способ не пригоден.

Второй способ состоит в непосредственном усреднении самих внутримолекулярных корреляционных функций с использованием предварительно полученных из траекторий гистограмм распределения межатомных расстояний  $P_{ij}(r)$ :

$$\langle \omega_{ij}(k) \rangle = \int_0^{\infty} P_{ij}(r) \frac{\sin kr}{kr} dr. \quad (11)$$

Такой способ усреднения является гораздо более адекватным, особенно для конформационно подвижных молекул. Поясним это на модельном примере. На рис. 1 на верхнем графике изображено некоторое условное модельное распределение расстояний между двумя атомами. Из гистограммы видно, что в основном расстояния лежат в пределах  $2 \text{ \AA}$ , однако, в силу конформационной подвижности они могут увеличиваться до  $7 \text{ \AA}$ , а гистограмма имеет несколько пиков. На нижнем графике приведены внутримолекулярные корреляционные функции, соответствующие данной гистограмме. Пунктиром показана функция, усредненная по методу (10), сплошной линией — по методу (11). Хорошо видно, что функции существенно отличаются для всех волновых векторов, исключая окрестность нуля. Следовательно, и результаты решения уравнений с такими матрицами будут сильно отличны.

Также важную практическую ценность имеет тот факт, что функции, усредненные с использованием гистограмм, затухают значительно быстрее. Опыт расчетов показывает, что короткодействующие функции в матрице внутримолекулярных функций обеспечивают лучшую сходимость при итерационном решении уравнений RISM. Время, затраченное на решение усредненного по траектории уравнения, оказывается меньше, чем время, необходимое для получения решения даже для одной жесткой конфигурации.

В способе (11) нельзя выделить эквивалентные функции, чтобы осуществить алгоритм сжатия данных. Однако короткодействующий характер усредненных внутримолекулярных корреляционных функций позволяет осуществить обрезание каждой функции до некоторого максимального значения волнового вектора  $k_{ij}^*$ , удовлетворяющего условию

$$\int_{k_{ij}^*}^{\infty} |\omega_{ij}(k)| dk \Big/ \int_0^{\infty} |\omega_{ij}(k)| dk < \xi, \quad (12)$$

где  $\xi$  — заранее заданная точность обрезания функций. Параметр  $\xi$  влияет как на упаковку матрицы, так и на решения уравнений, потому что элементы точной матрицы  $\omega_{ij}(k)$  заменяются на приближенные выражения  $\omega_{ij}(k)H(k - k_{ij}^*)$ , где  $H(x)$  — функция Хевисайда.  $\xi$  выбирается так, чтобы матрица была достаточно компактной, и в то же время внесенная обрезанием погрешность была меньше точности численного решения уравнений RISM итерационным методом. Дополнительно можно уменьшить влияние обрезания, если граничные значения волновых векторов  $k_{ij}^*$  выбрать так, чтобы совместно с условием (12) выполнялось условие  $\omega_{ij}(k_{ij}^*) = 0$ .

#### 4. Термодинамические функционалы

Для сравнения различных состояний растворенной молекулы по энергии, необходимо найти энергию Гиббса  $\Delta G$ , которая есть сумма избыточного химического потенциала  $\Delta\mu$  и молекулярно-механической энергии  $E_{MM}$ :

$$\Delta G = E_{MM} + \Delta\mu. \quad (13)$$

Избыточный химический потенциал  $\Delta\mu$  представляет собой работу по внедрению одной частицы в среду. Путь внедрения частицы в среду имитируется ослаблением потенциалов взаимодействия с помощью масштабирующего параметра  $\lambda$ :

$$\Delta\mu = 4\pi \frac{\rho}{2\beta} \sum_{i\alpha} \int_0^1 d\lambda \int_0^{\infty} dr r^2 g_{i\alpha}(r, \lambda) \frac{\partial u(r, \lambda)}{\partial \lambda}, \quad (14)$$

где  $g_{i\alpha}(r, \lambda) = h_{i\alpha}(r, \lambda) + 1$  — атом-атомные радиальные функции распределения атома  $\alpha$  молекул растворителя вокруг атома  $i$  растворенной молекулы, которые являются

решениями уравнений RISM для атом-атомных потенциалов взаимодействия  $u(r, \lambda)$ , ослабленных параметром  $\lambda$ . В этом разделе будем говорить о корреляционных функциях, не делая различий между средними, псевдосредними и обычными функциями, и будем подразумевать, что термодинамические функционалы могут быть рассчитаны по любым из них.

В ряде случаев интегрирование (14) по  $\lambda$  можно выполнить аналитически. Так, в теории гауссовых флуктуаций плотности (GF) было получено выражение [12]:

$$\Delta\mu^{GF} = 4\pi \frac{\rho}{2\beta} \sum_{i\alpha} \int (-2c_{i\alpha}(r) - h_{i\alpha}(r)c_{i\alpha}(r))r^2 dr. \quad (15)$$

Выражение избыточного химического потенциала Сингера–Чандлера (SC) [4, 13, 14] будет иметь различный вид для замыканий HNC (16) и PLHNC (17):

$$\Delta\mu^{SC} = \Delta\mu^{GF} + 4\pi \frac{\rho}{2\beta} \sum_{i\alpha} \int h_{i\alpha}^2(r)r^2 dr, \quad (16)$$

$$\Delta\mu^{SC} = \Delta\mu^{GF} + 4\pi \frac{\rho}{2\beta} \sum_{i\alpha} \int h_{i\alpha}^2(r)H(-h_{i\alpha}(r))r^2 dr. \quad (17)$$

Одно из выражений избыточного химического потенциала, недавно полученное в методе парциальных волн (PW), имеет вид:

$$\Delta\mu^{PW} = \Delta\mu^{GF} + \frac{\rho^2}{2\beta(2\pi)^3} \sum_{i\alpha} \int 4\pi k^2 \hat{h}_{i\alpha}(k) \hat{\zeta}_{i\alpha}(k) dk, \quad (18)$$

где  $\hat{\zeta}_{i\alpha}$  — элементы матрицы  $Z$ :

$$Z = \{ \hat{\zeta}_{i\alpha}(k) \} = \hat{c}^{i\omega} (\hat{w}^v + \rho \hat{h}^{v\omega}) \hat{c}^{v\omega}.$$

Пояснения к этому выражению можно найти в оригинальных работах [15, 16].

Чтобы скорректировать энтропийную составляющую энергии, в работе [6] предложено внести в уравнение замыкания отталкивательную поправку. Так, гиперцепное уравнение с поправкой (HNC/RBC) будет иметь вид:

$$\langle c \rangle_{i\alpha} = \exp(-\beta u_{i\alpha} + b_{i\alpha} + \langle \gamma \rangle_{i\alpha}) - 1 - \langle \gamma \rangle_{i\alpha}, \quad (19)$$

а частично линейаризованное гиперцепное уравнение (PLHNC/RBC), соответственно:

$$\langle c \rangle_{i\alpha} = \exp(-\beta u_{i\alpha} + b_{i\alpha}) F(\langle \gamma \rangle_{i\alpha}) - 1 - \langle \gamma \rangle_{i\alpha}. \quad (20)$$

Поправка  $b_{i\alpha}$  задается выражением:

$$\exp(b_{i\alpha}) = \prod_{\mu \neq \alpha} [\omega_{\mu\alpha} * \exp(-\beta u_{i\mu}^{RBC})],$$

где \* обозначает свертку в координатном пространстве. В качестве отталкивательного потенциала  $u_{i\mu}^{RBC}$  может быть выбран либо отталкивательный член  $r^{-12}$  потенциала Леннарда–Джонса:

$$u_{i\mu}^{r^{-12}}(r) = 4\epsilon_{i\mu} \left( \frac{\sigma_{i\mu}}{r} \right)^{12}, \quad (21)$$

либо отталкивательная часть потенциала Вика–Чандлера–Андерсона (WCA):

$$u_{i\mu}^{WCA}(r) = 4\epsilon_{i\mu} \left[ \left( \frac{\sigma_{i\mu}}{r} \right)^{12} - \left( \frac{\sigma_{i\mu}}{r} \right)^6 + \frac{1}{4} \right] H(2^{1/6} \sigma_{i\mu} - r). \quad (22)$$

Для уравнений RISM с отталкивательной поправкой невозможно получить аналитические выражения для избыточного химического потенциала в явном виде, но можно численно проинтегрировать (14), решив систему уравнений RISM для набора значений  $\lambda$  от 0 до 1. Однако интегрирование из-за его большой вычислительной ресурсоемкости можно выполнить только в случае усредненного по траектории



уравнения (7). В обычном подходе [6] приходится использовать приближенные выражения для избыточного химического потенциала с учетом отталкивательной поправки, полученные в рамках термодинамической теории возмущений (ТРТ) :

$$\Delta\mu^{TPT} = \Delta\mu^{SC} + 4\pi \frac{\rho}{2\beta} \sum_{i\alpha} \int (h_{i\alpha}(r) + 1) (\exp(b_{i\alpha}(r)) - 1) r^2 dr. \quad (23)$$

### ДЕТАЛИ ВЫЧИСЛЕНИЙ

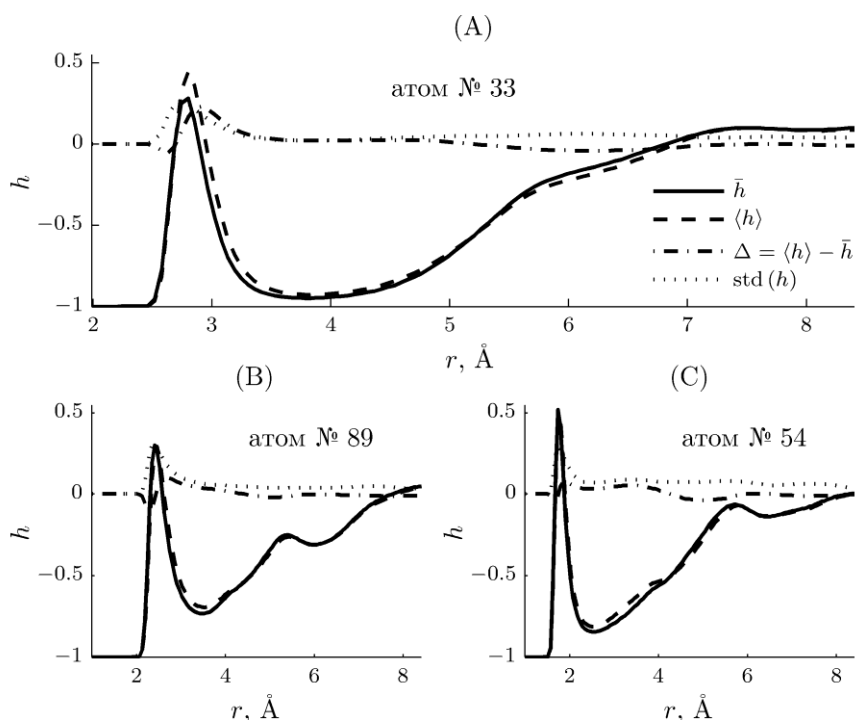
Для сравнения псевдосредних корреляционных функций и термодинамических величин, полученных по псевдосредним функциям (7), с истинными средними (6), мы воспользуемся результатами нашей работы [3], в которой истинные средние были рассчитаны для трех различных траекторий пептида окситоцина. Траектории были получены с помощью трех различных молекулярно-динамических расчетов: пептид в вакууме (А), пептид в явном водном окружении (В), а также расчет, в котором водное окружение учитывалось неявно по обобщенному методу Борна (С). Были вычислены корреляционные функции для 1250 конфигураций пептида в каждой из трех траекторий. По этим корреляционным функциям была рассчитана энергия Гиббса, а затем вычислены средние величины и их дисперсии. Энергия Гиббса была вычислена с использованием разных выражений для избыточного химического потенциала (15)–(18), которые обозначались как GF, SC и PW, соответственно, а также приближенное выражение (23) с учетом отталкивательных поправок (21) и (22), обозначенных как RBC1 и RBC2.

На этом достаточно представительном наборе данных в настоящей работе выполнено сравнение истинных средних с решениями усредненного уравнения. Параметр  $\xi$ , задающий точность обрезания внутримолекулярных корреляционных функций (12), был выбран равным 0.01. Гистограммы распределения межатомных расстояний рассчитывались по 12500 точкам молекулярно-динамической траектории. Решение усредненного уравнения RISM проводилось при тех же параметрах сетки и потенциалов, которые были использованы в работе [3]. Численное интегрирование выражения (14) с учетом отталкивательных поправок (21) и (22) по параметру  $\lambda$  проводилось с адаптивным шагом, точность интегрирования проверялась пересчетом с дробными шагами. Величины, полученные интегрированием, ниже обозначены как INT1 и INT2, соответственно.

### РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Начнем сравнение подходов (6) и (7) с анализа отличий между средними и псевдосредними корреляционными функциями. Для каждой траектории выбрано по одной полной атом-атомной корреляционной функции с самым большим интегралом модуля разности между средней и псевдосредней функцией. На рис. 2 эти функции изображены для случая замыкания HNC, а также приведены их разность и среднеквадратичное отклонение. Там же указаны порядковые номера атомов в молекуле окситоцина. Все функции относятся к корреляциям с атомом водорода воды. Видно, что для разных траекторий функции с максимальным отклонением относятся к разным атомам пептида. Из рисунка видно, что для траектории А разность функций лишь немного выше дисперсии в районе первого пика полной корреляционной функции. Во всей остальной области разность меньше дисперсии. Для траекторий В и С разность всюду меньше дисперсии. Для замыкания PLHNC картина полностью аналогична, вплоть до номеров атомов с максимальным отличием корреляционных функций, поэтому мы не приводим рисунок. Из того факта, что отличия истинных средних от псевдосредних не зависят от выбора уравнения замыкания, можно заключить, что нелинейная часть уравнений RISM не влияет существенно на эти

ОТЛИЧИЯ.



**Рис. 2.** Средние и псевдосредние корреляционные функции, их разность и дисперсия для того атома в каждой из траекторий, интеграл модуля разности функций которого оказался максимальным.

Перейдем к анализу отличий в термодинамике. В табл. 1 приведены средние значения молекулярно-механической энергии, избыточного химического потенциала и энергии Гиббса, полученные в работе [3]. Из таблицы видно, что конформационные состояния в вакууме (А) наименее предпочтительны. Состояния В и С близки и характеризуются энергией Гиббса, отличающейся менее чем на 1 ккал. Состояние С ниже по энергии, чем состояние В.

**Таблица 1.** Средние значения (стандартные отклонения) молекулярно-механической энергии ЕММ, оценки энергии Гиббса сольватации по обобщенному методу Борна плюс поверхностный член  $\Delta\mu$  и энергии Гиббса пептида  $\Delta G$  по методу ММ/GBSA (ккал/моль)

Траектория	$-\langle E_{MM} \rangle$	$-\langle \Delta\mu^{GBSA} \rangle$	$-\langle \Delta G \rangle$
А	-126.6 (11.4)	-62.2 (8.0)	-188.7 (9.4)
В	-5.9 (11.6)	-190.3 (7.9)	-196.2 (8.5)
С	7.5 (12.2)	-204.7 (10.5)	-197.1 (5.6)

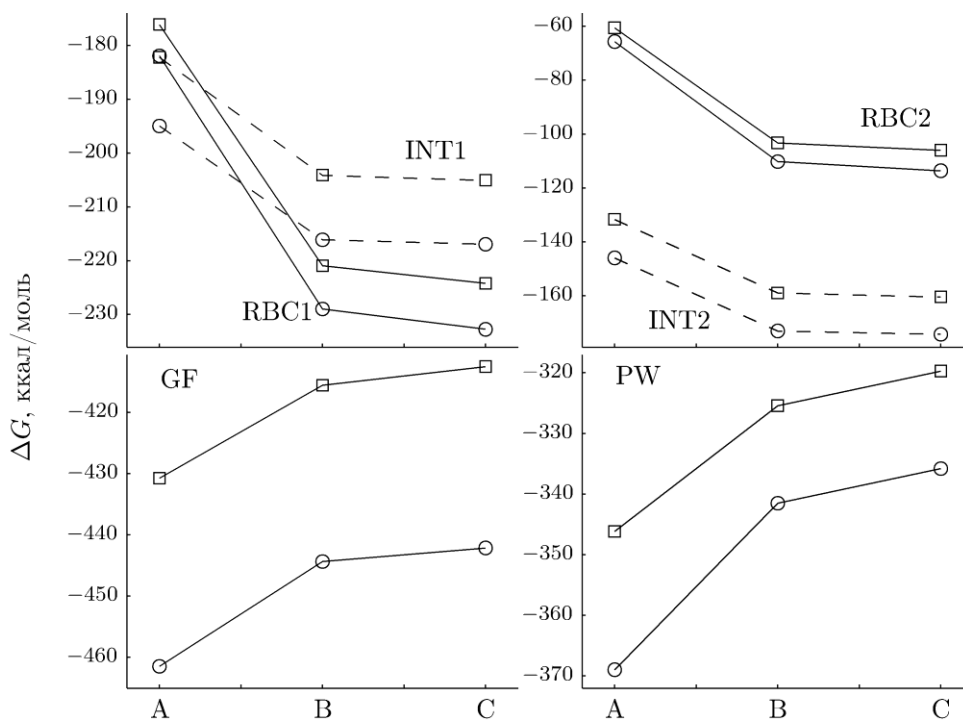
Приведем данные о погрешности усредненного по траектории метода RISM в части термодинамики. В табл. 2 перечислены разности среднего избыточного химического потенциала гидратации и полученного по решениям усредненного уравнения. Для сравнения в скобках приведены значения среднеквадратичного отклонения избыточно химического потенциала, рассчитанного в каждой точке траектории. Данные получены для замыканий HNC и PLHNC. Видно, что результаты группируются. Так, разности химического потенциала, вычисленного по формулам SC, RBC1 и RBC2, отрицательны и примерно равны величинам дисперсии. Для траектории А они слегка превышают

дисперсию, для остальных траекторий — немного меньше дисперсии. Существенной разницы между замыканиями нет. Для формул GF и PW разности положительны и также сравнимы с дисперсией для замыкания HNC. Однако для замыкания PLHNC эти разности, оставаясь положительными, существенно меньше дисперсии. Таким образом, погрешности термодинамики, рассчитанной по решениям усредненного уравнения RISM, лежат в пределах дисперсии мгновенных значений, вычисленных по той же выборке.

**Таблица 2.** Разность истинной средней энергии Гиббса и энергии Гиббса, полученной по решениям усредненного уравнения (ккал/моль). В скобках приведена дисперсия энергии Гиббса, полученной по мгновенным геометрическим конфигурациям

Траектория	Замыкание	SC	GF	PW	RBC1	RBC2
A	HNC	-20.4 (17.0)	18.0 (8.4)	12.6 (7.6)	-20.0 (17.5)	-19.4 (17.0)
	PLHNC	-21.0 (17.0)	6.2 (6.7)	1.7 (6.4)	-21.3 (17.6)	-20.4 (17.0)
B	HNC	-12.2 (10.9)	12.5 (9.5)	9.0 (8.8)	-12.0 (12.5)	-11.7 (11.7)
	PLHNC	-12.5 (10.8)	6.7 (8.5)	3.6 (8.0)	-12.7 (12.4)	-12.2 (11.7)
C	HNC	-12.2 (12.8)	14.7 (14.3)	11.2 (13.2)	-11.3 (14.0)	-11.2 (13.4)
	PLHNC	-12.4 (12.8)	8.3 (12.8)	5.2 (12.0)	-12.0 (14.0)	-11.7 (13.4)

Перейдем к обзору непосредственно энергии Гиббса. На рис. 3 изображены величины энергии Гиббса, вычисленные по псевдосредним функциям для обоих замыканий. Качественная картина изменений энергии Гиббса полностью аналогична той, которая получена в работе [3]. Поэтому мы не будем обсуждать эти изменения в деталях. Остановимся только на отличиях оценок энергии Гиббса, полученных с учетом отталкивательных поправок методами термодинамической теории возмущений (RBC1, RBC2) и термодинамического интегрирования (INT1, INT2). Сравнивая



**Рис. 3.** Энергии Гиббса для окситоцина, полученные по решениям усредненного уравнения RISM. Результаты термодинамического интегрирования отмечены пунктиром.  $\circ$  — HNC,  $\square$  — PLHNC.

полученные этими методами абсолютные величины энергии Гиббса, отметим, что для

разных отталкивательных поправок картина различна. Для поправки (21) термодинамическое интегрирование практически не изменило результат, тогда как для поправки (22) абсолютная величина энергии Гиббса изменилась примерно на 70–100 ккал/моль.

Абсолютная величина энергии Гиббса малоинформативна, следует сравнить изменения энергии Гиббса при переходе из одного состояния в другое. Численные данные о разности энергий, характеризующих разные состояния, приводятся в табл. 3. Данные получены по решениям усредненного уравнения RISM для замыканий HNC и PLHNC. Сравнивая их с теми же разностями истинных средних энергий Гиббса, которые приведены в таблице II в работе [3], можно отметить, что разность энергий, характеризующих сравниваемые состояния, систематически уменьшилась для всех выражений избыточного химического потенциала.

**Таблица 3.** Разности энергий Гиббса, вычисленным по псевдосредним корреляционным функциям, между состояниями A, B и C (ккал/моль)

Разность	Замыкание	SC	GF	PW	RBC1	RBC2	INT1	INT2	GBSA
$\Delta G(B) - \Delta G(A)$	HNC	-43.3	17.1	27.4	-46.9	-44.6	-21.1	-27.0	-7.5
	PLHNC	-41.5	15.2	20.8	-44.8	-42.5	-21.9	-27.3	
$\Delta G(C) - \Delta G(A)$	HNC	-46.5	19.4	33.1	-50.8	-47.9	-21.9	-28.4	-8.4
	PLHNC	-44.4	18.3	26.4	-48.1	-45.3	-22.7	-28.8	
$\Delta G(B) - \Delta G(C)$	HNC	3.2	-2.3	-5.7	3.8	3.3	0.8	1.4	0.9
	PLHNC	2.9	-3.1	-5.6	3.3	2.8	0.9	1.4	

Сравнивая изменение энергии Гиббса при переходе из состояния в вакууме к состоянию в растворе, полученной с учетом отталкивательных поправок (21) и (22) по теории возмущений (RBC1, RBC2) и методом термодинамического интегрирования (INT1, INT2), видим, что в случае численного интегрирования разница уменьшилась примерно в два раза. В работе [3] отмечено, что изменение энергии Гиббса завышено, поэтому уменьшение ее величины при точном учете отталкивательных поправок с помощью интегрирования свидетельствует в пользу самой идеи коррекции энтропийной составляющей с помощью отталкивательных поправок. Однако оценивать в этом случае избыточный химический потенциал методами термодинамической теории возмущений нельзя, поскольку такие оценки не меняют качественно картины перехода пептида из вакуума в раствор.

## ЗАКЛЮЧЕНИЕ

Из приведенных сравнений можно сделать вывод, что усредненное уравнение (7) и метод усреднения (11) дают весьма неплохие результаты не только для малых жестких молекул, но и для конформационно подвижных пептидов. Псевдосредние корреляционные функции совпадают с истинными средними функциями с точностью до статистической ошибки усреднения. Также и термодинамика, полученная по псевдосредним функциям, отличается от средних величин в пределах погрешности. При этом затраты на вычисление псевдосредних меньше минимум на два порядка в сравнении с затратами на вычисление средних по траектории величин.

В том случае, когда молекулярная траектория нестационарная, например, молекула претерпевает существенные конформационные изменения, метод усредненного уравнения также применим с некоторыми поправками. Нужно выделить несколько квазистационарных участков, и на них решать задачу о псевдосредних корреляционных функциях. В целом такой подход оказывается более эффективным, чем решение полной задачи RISM в каждой точке нестационарной траектории.

Использование усредненного уравнения позволяет решать более сложные вычислительные задачи, чем применение обычных методов решения уравнений RISM с последующим усреднением результатов. Например, для псевдосредних возможно выполнить термодинамическое интегрирование или исследовать температурные зависимости сольватации с использованием больших молекулярно-динамических массивов данных. Однако следует помнить, что для каждой конкретной системы отклонение средних величин от псевдосредних может быть значительно больше дисперсии. Вопрос в том, будут ли зависимости псевдосредних от изучаемых параметров симпатны тем же зависимостям средних величин. Этот вопрос должен решаться применимо к каждому конкретному случаю.

### СПИСОК ЛИТЕРАТУРЫ

1. Chandler D., Andersen H.C. Optimized cluster expansions for classical fluids. II. Theory of molecular liquids. *Journal of Chemical Physics*. 1972. V. 57. P. 1930–1937.
2. Hirata F., Rossky P.J., Pettitt B.M. The interionic potential of mean force in a molecular polar solvent from an extended RISM equation. *Journal of Chemical Physics*. 1983. V. 78. № 6. P. 4133–4144.
3. Тихонов Д.А., Соболев Е.В. Оценки энергии Гиббса гидратации по молекулярно-динамическим траекториям методом интегральных уравнений теории жидкостей в приближении RISM. *Журнал физической химии*. 2011. Т. 85. № 3 (в печати).
4. Morita T., Hiroike K. A new approach to the theory of classical fluids. I. *Progress of Theoretical Physics*. 1960. V. 23. P. 1003–1027.
5. Lue L., Blankschtein D. Liquid-state theory of hydrocarbon-water systems: application to methane, ethane, and propane. *The Journal of Physical Chemistry*. 1992. V. 96. № 21. P. 8582–8594.
6. Kovalenko A., Hirata F. Hydration free energy of hydrophobic solutes studied by a reference interaction site model with a repulsive bridge correction and a thermodynamic perturbation method. *Journal of Chemical Physics*. 2000. V. 113. P. 2793–2805.
7. van Leeuwen J.M.J., Groeneveld J., de Boer J. New method for the calculation of the pair correlation function. I. *Physica*. 1959. V. 25. № 7–12. P. 792–808.
8. Kovalenko A., Hirata F. Self-consistent description of a metal-water interface by the Kohn–Sham density functional theory and the three-dimensional reference interaction site model. *Journal of Chemical Physics*. 1999. V. 110. № 20. P. 10095–10112.
9. Tikhonov D.A., Polozov R.V., Timoshenko E.G. et al. Hydration of a B-DNA fragment in the method of atom-atom correlation functions with the reference interaction site model approximation. *Journal of Chemical Physics*. 1998. V. 109. P. 1528–1539.
10. Тихонов Д.А. Метод интегральных уравнений теории жидкостей для изучения гидратации макромолекул. В: *Компьютеры и суперкомпьютеры в биологии*. Под ред. Лахно В.Д., Устинина М.Н. Москва–Ижевск: Институт компьютерных исследований, 2002. С. 209–233.
11. Соболев Е.В., Тихонов Д.А. Применение метода RISM для оценки свободной энергии связывания 4',6-диамидино-2-фенилиндола в малом желобе ДНК по молекулярно-динамической траектории. *Математическая биология и биоинформатика*. 2010. Т. 5. № 2. С. 98–113. URL: [http://www.matbio.org/downloads/Sobolev2010\(5\\_98\).pdf](http://www.matbio.org/downloads/Sobolev2010(5_98).pdf) (дата обращения 15.12.2010).
12. Chandler D., Singh Y., Richardson D.M. Excess electrons in simple fluids. I. General equilibrium theory for classical hard sphere solvents. *The Journal of Chemical Physics*. 1984. V. 81. № 4. P. 1975–1982.
13. Singer S.J., Chandler D. Free energy functions in the extended RISM approximation. *Molecular Physics*. 1985. V. 55. P. 621–625.

14. Zichi D.A., Rossky P.J. Molecular conformational equilibria in liquids. *The Journal of Chemical Physics*. 1986. V. 84. № 3. P. 1712–1723.
15. Ten-no S. Free energy of solvation for the reference interaction site model: Critical comparison of expressions. *Journal of Chemical Physics*. 2001. V. 115. P. 3724–3731.
16. Ten-no S., Iwata S. On the connection between the reference interaction site model integral equation theory and the partial wave expansion of the molecular Ornstein–Zernike equation. *Journal of Chemical Physics*. 1999. V. 111. P. 4865–4868.

Материал поступил в редакцию 01.12.2010, опубликован 17.12.2010