

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. Ломоносова

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ
КАФЕДРА ФИЗИКИ ПОЛИМЕРОВ И КРИСТАЛЛОВ

**Исследование гидратации и
адсорбции боковых цепей
аминокислот методами
компьютерного моделирования**

Дипломная работа
студента 6 курса
Шайтана А.К.

Научный руководитель
д.ф.-м.н, профессор Хохлов А. Р.

Допущен к защите 29 декабря 2006 года

Зав. кафедрой
д.ф.-м.н., проф. А.Р. Хохлов

Москва, 2006

Аннотация

Работа посвящена изучению и применению методов классического компьютерного моделирования для оценки сольватационных и адсорбционных свойств небольших молекул. Проводится обзор теоретических подходов к описанию эффектов сольватации и адсорбции на границе раздела жидкость/жидкость, жидкость/газ. Описываются теоретические основы компьютерного молекулярного моделирования в рамках классических силовых полей, способы расчёта энергии сольватации, способы построения профилей свободной энергии и определения энергии адсорбции в системах с границей раздела фаз методами компьютерного моделирования. Изучаются практические аспекты реализации компьютерного эксперимента методами молекулярной и стохастической динамики.

Практическая часть работы состоит в применении методов молекулярной динамики для оценки энергии гидратации и адсорбции боковых цепей некоторых аминокислот. Проводится сравнение эффективности различных методов для определения энергии гидратации небольших молекул. С высокой точностью определена энергия гидратации 13 R-Н-аналогов (боковых цепей) аминокислот в рамках силового поля OPLS-AA/SPC. Изучена возможность построения профиля свободной энергии молекул и оценки их энергии гидратации в системе с разделом фаз вода/вакуум. Проведены построение и анализ профилей свободной энергии для 5 R-Н-аналогов (боковых цепей) аминокислот в рамках силового поля OPLS-AA/SPC, оценена их энергия адсорбции.

Оглавление

Список сокращений	5
1 Введение	6
1.1 Постановка задачи	8
2 Обзор литературы	9
2.1 Термодинамические характеристики сольватации и адсорбции	9
2.1.1 Энергия гидратации	9
2.1.2 Энергия адсорбции	11
2.2 Элементы статфизики в применении к задаче	13
2.2.1 Энергия сольватации	13
2.2.2 Энергия адсорбции и потенциал средней силы	13
2.3 Классический метод молекулярной динамики	15
2.3.1 Классическая модель молекулярных систем	15
2.3.2 Стандартные силовые поля	15
2.3.3 Моделирование динамики классической системы	16
2.3.4 Теоретические аспекты метода МД	16
2.3.5 Методы задания ансамблей	17
2.3.6 Метод стохастической динамики	18
2.3.7 Технические аспекты моделирования	18
2.4 Элементы неравновесной статистической физики	20
2.4.1 Обобщение II начала термодинамики	21
2.4.2 Равенство Джарзинского	21
2.5 Практические методы расчёта свободной энергии	22
2.5.1 Метод термодинамического интегрирования	22
2.5.2 Метод возмущения свободной энергии	23
2.5.3 Метод Беннета	23
2.6 Расчёт СЭС методами МД	24
2.6.1 Путь превращения системы	25
2.6.2 Особенности реализации метода ТИ	25
2.6.3 Особенности реализации метода Беннета и FEP	25
2.6.4 Источники ошибок	26
2.6.5 Обзор работ	26
2.7 Моделирование границы раздела фаз	27
2.8 Расчёт ПСС методами МД	27
2.8.1 Обзор работ	28

3	Определение свободной энергии гидратации боковых цепей аминокислот	29
3.1	Аннотация	29
3.2	Введение	29
3.3	Силовое поле	29
3.4	Метод расчёта свободной энергии	30
3.5	Получение выборки	32
3.6	Анализ ошибок	34
	3.6.1 Статистическая погрешность	35
	3.6.2 Погрешность выборки	35
3.7	Сравнение ТИ и метода Беннета	37
3.8	Выбор промежуточных состояний	39
3.9	Результаты	40
3.10	Обсуждение и сравнение	40
3.11	Выводы	42
4	Исследование адсорбции боковых цепей аминокислот	44
4.1	Аннотация	44
4.2	Введение	44
4.3	Методы	45
4.4	Результаты	47
4.5	Обсуждение	47
4.6	Энергия адсорбции	52
4.7	Выводы	52
5	Выводы	54
A	Молекулы-объекты исследования	55
B	Вывод формулы для расчёта СЭС в компьютерном моделировании	57
B.1	Введение	57
B.2	Свободная энергия сольватации	57
B.3	Постулаты классической статфизики	58
B.4	Связь статсуммы со свободной энергией	58
	B.4.1 Квантовый подход	58
	B.4.2 Подход Гиббса	59
B.5	Сольватация, рассмотрение статфизики	60

Список сокращений

СЭП - свободная энергия переноса

СЭС - свободная энергия сольватации

СЭГ - свободная энергия гидратации

ПСС - потенциал средней силы

МД - молекулярная динамика, метод молекулярной динамики

Глава 1

Введение

Описание водорастворимых полимеров является одной из основных задач физики и химии полимеров, потому что для большинства природных макромолекул, таких как белки и ДНК, вода является основным растворителем. Более того, за последние десятилетия многие синтетические водорастворимые полимеры нашли широкое применение в биомедицинских целях [1]. Понимание поведения водорастворимых полимеров (как синтетических, так и натуральных) часто является важным для оценки их биологической активности.

Биологическая активность белков сильно завязана на их конформацию и четвертичную структуру. Считается, что конформация нативных белков отвечает глобальному минимуму свободной энергии [2]. Последовательность уровней свободной энергии для разных конформаций может меняться в присутствии других молекул (шаперонов, амилоидов или маленьких молекул, таких как сахара и мочевины) и поверхностей раздела фаз [3, 4]. Конформационные изменения в белках важны для разработки, производства и применения белоксодержащих лекарств [5]. Агрегация белков играет ключевую роль в некоторых заболеваниях, включая болезнь Альцгеймера [6], прионных заболеваниях и губчатых энцефалитах [7].

Для понимания процессов формирования конформации и агрегации белков, а также синтетических полимеров, было предложено большое число моделей. Одной из первых моделей для описания поведения сополимеров в воде была модель Лау и Дилла [8]. Они предложили подразделение мономерных звеньев на два типа: гидрофобные (Н) и полярные (Р). Цепи амфифильных полимеров представлялись состоящими из последовательно соединённых бусинок Н и Р типа (Рис. 1.1). Для водных растворов таких полимеров предполагалось, что гидрофобные бусинки притягиваются друг к другу, в то время как между полярными бусинками преобладают силы отталкивания.

Такая НР модель широко применяется и к белкам, предполагается, что структура глобулярных белков может быть представлена в виде гидрофобного ядра и относительно гидрофильной оболочки. НР модель используется в различных теориях белкового фолдинга.

Однако многие мономерные звенья, как белков, так и синтетических полимеров содержат в своей структуре как гидрофобные, так и полярные группы, то есть могут проявлять определённые амфифильные свойства. В работах [10],[11] была предложена так называемая НА-модель сополимера в водном окружении. В этой модели амфифильные А-звенья состояли из двух элементов: гидрофобные бусинки выполняли роль остова полимерной цепи, а полярные - роль боковых цепей, присоединённых к бусинкам остова. В том числе методами огрублённого компьютерного молекулярного моделирования было показано, что конформационные свойства полимера в такой модели могут серьёзным образом отличаться от свойств предсказываемых НР-моделью сополимера. Так, например, глобулы в НА-модели могут быть не сферическими. За необычные конформационные свойства НА-модели ответственны в основном амфифильные

Рис. 1.1: Схематичное представление НР-модели (а) и НА-модели (b). Н-звенья - гидрофобные, Р-звенья - полярные (гидрофильные), А-звенья - амфифильные. По [9].

звенья, которым энергетически выгодно находиться на поверхности.

Все вышеописанные модели представляют собой достаточно сильное упрощение реальной картины взаимодействий на молекулярном уровне, но в то же время они, по-видимому, отражают наиболее основные особенности поведения таких систем. Важной задачей является соотнесение параметров огрублённой модели со свойствами реальных молекулярных систем.

В работах [9, 12] была предложена одновременная классификация мономеров по их гидрофобности (сродству к воде) и амфифильности (сродству к границе раздела фаз). В качестве численных параметров были использованы свободная энергия переноса молекулы из воды в гексан и свободная энергия адсорбции на границу раздела фаз между водой и гексаном. Эти параметры были определены для набора синтетических мономеров, а также для некоторого количества боковых цепей аминокислот (мономеров белков). Пример, классификационная диаграмма приведён на рис. 1.2.

Рис. 1.2: Двумерная диаграмма для боковых цепей аминокислот по [12]. По осям отложены свободная энергия адсорбции и переноса из фазы в фазу в системе вода/гексан в единицах kT .

Важным методом исследования молекулярных и полимерных систем является компьютерное моделирование. Огрублённое компьютерное моделирование в простых моделях таких, как, например, НА-модель сополимера, позволяет определять важные статистические параметры

полимерной цепи, в частности её конформацию.

На сегодняшний день большое развитие получили и методы атомистического моделирования молекулярных структур, особенно в области биополимеров, в частности, метод молекулярной динамики [13]. Обычно в таких методах используются полноатомные классические модели молекул, в которых атомы представляются в виде “шариков”, взаимодействующих друг с другом и окружением по законам классической механики. Эти методы позволяют с хорошей точностью описывать поведение и детали взаимодействия реальных молекул, однако, из-за сильного увеличения объёма вычислительных процедур по сравнению с огрублёнными моделями, имеют некоторые ограничения в применимости. Эти ограничения, прежде всего, связаны с размерами исследуемых структур и характерными временами молекулярных процессов, которые необходимо исследовать, а также с теми величинами, которые необходимо измерить. Особенно требовательными к вычислительным ресурсам, являются процедуры определения различных термодинамических величин: свободных энергий сольватации, констант связывания, профилей свободной энергии и т.д. Для небольших молекул определение таких термодинамических величин напрямую из атомистического молекулярного моделирования становится на сегодняшний день возможным.

В этой связи представляется интересным применить методы атомистического компьютерного моделирования для описания взаимодействия различных мономерных звеньев с водой и границей раздела фаз между водой и неполярным растворителем. Не следует, однако, забывать, что вычислительный эксперимент всегда идёт в рамках какой-то модели. Поэтому, по результатам моделирования, мы можем делать выводы не только о возможных процессах в реальной молекулярной системе, но и о границах применимости самой модели.

1.1 Постановка задачи

Целью настоящей работы является изучение и применение возможностей полноатомного молекулярного моделирования для анализа взаимодействия небольших молекул с водным окружением и с границей раздела фаз вода/неполярная среда. В качестве небольших молекул для анализа были выбраны молекулы-аналоги боковых цепей незаряженных аминокислот (см. Приложение А). Молекула-аналог (R-H-аналог) представляет собой молекулу аминокислоты, в которой пептидный остов вместе с C- α атомом заменён на атом водорода (Рис. 1.3).

Рис. 1.3: На рисунке слева направо представлены схематичные химические формулы аминокислоты, боковой цепи (радикала) и молекулы-аналога аминокислоты

Глава 2

Обзор литературы

Раздел “Обзор литературы” выделен достаточно условно и больше представляет из себя теоретическое введение и базовое описание некоторых использованных методик. Чтобы сохранить последовательность изложения материала обзор литературных источников также приводится по ходу изложения глав основной части работы.

2.1 Термодинамические характеристики сольватации и адсорбции

Как было сказано в разделе 1.1 “Постановка задачи”, мы стремимся с помощью полноатомного молекулярного моделирования охарактеризовать взаимодействие молекул растворённого вещества с водным окружением и с границей раздела фаз вода/неполярная среда. В химической физике данные взаимодействия обычно описывают двумя параметрами: *свободной энергией сольватации* и *свободной энергией адсорбции*. Рассмотрим подробнее эти понятия.

Предположим, что у нас имеется термодинамическая система, состоящая из двух плохо смешивающихся жидкостей (воды и гексана), которые приведены в соприкосновение по некоторой границе раздела (для простоты предположим, что это плоскость¹, в этой двухфазной системе растворено малое количество вещества S; это вещество некоторым образом перераспределяется в системе. Рис. 2.1.

Систему условно можно разделить на три зоны: вода, гексан и зона перехода между ними.

2.1.1 Энергия гидратации

Применительно к сольватации в воде свободная энергия сольватации обычно называется свободной энергией гидратации (СЭГ) и представляет собой *свободную энергию переноса молекулы из идеального газа в воду*. Для описания переноса молекулы из одной фазы в другую в нашем случае может быть использована теория слабых растворов.

Теория слабых растворов

Предположим, что две жидкости абсолютно несмешиваемые (обладают нулевой взаимной растворимостью), а концентрация растворённого вещества настолько мала, что взаимодействием

¹Хотя с точки зрения концентрационного равновесия форма и площадь границы раздела значения не имеет.

Рис. 2.1: Схематическое представление исследуемой двухфазной системы.

между молекулами этого вещества можно пренебречь, молекулы не агрегируют. *Ещё раз подчеркнём важность этого утверждения, так как все дальнейшие выкладки и способы экспериментального определения искомых параметров относятся лишь к слабым растворам.* Допустим также пока, что суммарный заряд молекулы растворённого вещества равен нулю.²

При таких предположениях наша система является типичным примером двух соприкасающихся слабых растворов. Химический потенциал растворённого вещества в слабом растворе μ_{solv} записывается, как [14]:

$$\mu_{solv} = kT \ln c + \psi(P, T)$$

где c -концентрация растворённого вещества, T -температура, k -постоянная Больцмана, ψ -функция давления и температуры, которую будем называть добавочным химическим потенциалом, и обозначать также μ_{ex} .

Условием равновесия такой системы по растворённому веществу является равенство химических потенциалов растворённого вещества в обеих фазах, откуда следует:

$$\ln \frac{c^1}{c^2} = \frac{\mu_{ex}^2 - \mu_{ex}^1}{kT}$$

Свободной энергией переноса (СЭП) вещества из фазы 1 в фазу 2 назовём величину:

$$F_{trans} = \mu_{ex}^2 - \mu_{ex}^1 = \ln \frac{c^1}{c^2} \quad (2.1)$$

Формула 2.1 указывает на способ экспериментального определения СЭП путём измерения соответствующих концентраций.

²Если молекулы будут заряжены, то рассмотрение усложнится из-за учёта влияния противоионов и дальнедействующего характера кулоновских сил.

Экспериментальное определение СЭГ

Свободные энергии гидратации и переноса из фазы в фазу для R-H аналогов боковых цепей аминокислот с хорошей точностью определяются экспериментально [12, 15, 16, 17, 18]. Для экспериментального определения концентрации растворённого вещества в растворе используют, например, спектрофотометрические методы [9]. Концентрацию вещества в газовой фазе можно измерять по парциальному давлению молекул вещества.

2.1.2 Энергия адсорбции

Теория адсорбции

Для описания процессов адсорбции широкого применения получила теория адсорбции Гиббса [19]. В теории Гиббса предполагается, что поверхность раздела фаз является двумерной плоскостью, а сама граница раздела бесконечно тонкой. Существует и альтернативный подход к описанию явлений адсорбции на границе раздела фаз - это подход Гуггенхайма. Однако, из-за широкого использования подхода Гиббса в экспериментальных работах мы будем придерживаться именно этого подхода.

Предполагается, что поверхность раздела двух несмешивающихся фаз является ещё одной термодинамической фазой, причём объёмные фазы сохраняют свою однородность вплоть до границы раздела [20]. Самой же границе раздела приписывается некоторое количество адсорбированного вещества Γ^S (избыток Гиббса), которое может быть как положительным, так и отрицательным (положительная и отрицательная адсорбция соответственно). Для такой модельной системы справедливо дифференциальное уравнение адсорбции Гиббса, которое связывает активности³ различных компонент системы с поверхностным натяжением границы раздела.

Нас будет интересовать уравнение Гиббса в особом случае. Предположим, что граница раздела представляет собой раздел между растворителем (водой) и вакуумом (или насыщенным паром). Профиль плотности представлен на Рис. 2.2. Далее предположим, что в жидкой фазе у нас растворено очень малое количество вещества S, настолько малое, что взаимодействием между молекулами вещества S можно пренебречь. Исключительно важным является также вопрос о положении поверхности Гиббса. Её положение выбирается так, чтобы избыток Гиббса для растворителя равнялся нулю, то есть:

$$\int^{X_g} (c_1 - c(x))dx + \int_{X_g} (c_2 - c(x))dx = 0$$

где c_1, c_2 -концентрации растворителя в объёмных фазах, $c(x)$ -профиль концентрации растворителя, X_g -положение поверхности Гиббса.

Если выполнены все вышеперечисленные условия, то уравнение Гиббса запишется в виде

$$\Gamma^S = -\frac{1}{RT} \frac{d\gamma}{d \ln c} \quad (2.2)$$

где c -концентрация вещества S в одной из растворённых фаз, γ - поверхностное натяжение границы раздела, R -газовая постоянная, T -температура, Γ^S -избыток Гиббса для растворённого вещества.

³В общем случае уравнение Гиббса записывается для активностей, но в случае малых концентраций его можно записать и для концентраций.

Рис. 2.2: Схематический профиль плотности системы растворитель/насыщенный ($c(x)$) пар и положение поверхности раздела Гиббса. По [20].

Определение энергии адсорбции

Для определения численного значения свободной энергии адсорбции обычно используют следующий приём. Предполагают, что переходный слой между фазами имеет всё же некоторую толщину τ , которая обычно полагается равной 0.6 нм. Избыточной поверхностной концентрации Гиббса сопоставляется объёмная концентрация [21]. И энергия адсорбции может быть записана в форме предложенной Лэнгмюром [22].

$$\Delta f_{ads} = kT \ln \left(1 + \frac{\alpha}{RT\tau} \right) \quad (2.3)$$

где α -производная поверхностного давления⁴ по концентрации вещества S в одной из объёмных фаз, когда концентрация стремится к нулю.

$$\alpha = \left(\frac{\gamma_0 - \gamma}{c} \right)_{c \rightarrow 0}$$

Важно отметить, что в случае уравнения 2.3 энергия адсорбции берётся относительно энергии сольватации вещества в той фазе, относительно которой ведётся измерение α .

2.2 Элементы статфизики в применении к задаче

В предыдущих разделах было дано термодинамическое описание интересующих нас величин. В этом разделе будет дано далее детальное рассмотрение с помощью методов статистической физики.

2.2.1 Энергия сольватации

Для получения более детального понимания процесса сольватации его необходимо рассмотреть с точки зрения статистической физики.

⁴Поверхностное давление - двумерный аналог давления монослоя на поверхности, численно равен разности поверхностного натяжения чистого растворителя и поверхностного натяжения системы $\gamma_0 - \gamma$.

Можно показать, что в слабых растворах СЭС, которую мы измеряем как логарифм отношения концентраций, фактически является свободной энергией “включения взаимодействий” между молекулой и окружающим раствором (free energy of coupling a molecule into the rest of solution) [23].

Выводу и обсуждению этого факта посвящено приложение В.

Результат этого рассмотрения следующий. Гамильтониан нашей системы растворитель + молекула вещества представим как:

$$H = H_{sol} + H_{mol} + H_{int}$$

Нам нужно найти разницу свободной энергии между системами с гамильтонианами $H_1 = H_{sol} + H_{mol} + H_{int}$ и $H_2 = H_{sol} + H_{mol}$. Удобно записать это в виде зависимости гамильтониана от параметра λ , тем самым заодно определив некоторую плавную связь между двумя гамильтонианами по параметру λ

$$H(\lambda = 0) = H_{sol} + H_{mol} \quad (2.4)$$

$$H(\lambda = 1) = H_{sol} + H_{mol} + H_{int} \quad (2.5)$$

тогда

$$\Delta G = G(\lambda = 1) - G(\lambda = 0)$$

или, используя выражение статсуммы через интеграл по фазовому пространству:

$$\Delta G = -kT \ln \frac{\int dV d\vec{p} d\vec{q} \exp \left\{ -\frac{H(\vec{p}, \vec{q}, \lambda=1) + PV}{kT} \right\}}{\int dV d\vec{p} d\vec{q} \exp \left\{ -\frac{H(\vec{p}, \vec{q}, \lambda=0) + PV}{kT} \right\}} \quad (2.6)$$

Более точное выражение для свободной энергии сольватации (см. Приложение В, формула В.33) включает ещё поправку, связанную с изменением объёма системы, однако в большинстве случаев она оказывается пренебрежимо мала.

2.2.2 Энергия адсорбции и потенциал средней силы

Для энергии адсорбции (2.3) вывести аналитическую формулу так же, как это было сделано для энергии сольватации несколько затруднительно. Однако, описывать адсорбцию на молекулярном уровне можно в терминах профиля концентрации или профиля свободной энергии.

Потенциал средней силы

Профиль концентрации вещества в системе может быть связан с потенциалом средней силы (ПСС) (см. [24]). Рассмотрим теорию потенциала средней силы применительно к молекуле в двухфазной системе растворитель/газ (Рис. 2.2). Ось X направим в направлении перпендикулярном границе раздела фаз. Нас интересует зависимость концентрации растворённого вещества S , от координаты X . Предположим, что концентрация растворённого вещества мала, тогда рассмотрение концентрационного профиля можно свести к определению вероятности обнаружить одну молекулу вещества S с той или иной координатой центра масс. Пусть система обладает некоторым гамильтонианом H . Выделим в Гамильтониане координату

центра масс молекулы вещества S вдоль оси X:

$$H = H(\vec{p}, \vec{q}, X) \quad (2.7)$$

Для простоты рассмотрение будем проводить в каноническом (NVT) ансамбле. Плотность вероятности для центра масс молекулы иметь то или иное значение X, запишется в виде 2.8

$$P(X_0) = \frac{\int e^{-\frac{H}{kT}} \delta(X - X_0) d\vec{q} d\vec{p} dX}{Z} = \langle \delta(X - X_0) \rangle \quad (2.8)$$

Функция $P(X_0)$ является и нормированным профилем концентрации вещества S (при малых концентрациях S). Тогда профиль свободной энергии (он же профиль добавочного химического потенциала см. 2.1) запишется в виде:

$$\mu_{ex}^s(X_0) = -kT \ln\left(\frac{P(X_0)}{P_{gas}}\right) \quad (2.9)$$

где P_{gas} -плотность вероятности для молекулы находиться в газовой фазе (вакууме). Нетрудно показать [24], что

$$-\frac{\partial \mu_{ex}^s(X_0)}{\partial X_0} = \left\langle -\frac{\partial H}{\partial X} \right\rangle_{X=X_0} \quad (2.10)$$

где в правой части стоит средняя сила, действующая на центр масс молекулы S, в случае если сам центр масс закреплён в точке X_0 . В левой же части уравнения стоит производная от *потенциала средней силы*.

Связь с энергией адсорбции

Итак, измеряя среднюю силу, действующую на молекулу в разных положениях, можно, посчитав потенциал средней силы, перейти к профилю концентрации растворённого вещества S. Соответствующий данному концентрационному профилю избыток Гиббса запишется в виде:

$$\Gamma^s = \int^{X_g} (c_{sol} - c(x)) dx + \int_{X_g} (c_{gas} - c(x)) dx$$

где c_{sol} , c_{gas} -концентрации вещества S в жидкой и газовой фазах соответственно, X_g -координата поверхности Гиббса, рассчитанная по растворителю. Собственно, избыток Гиббса и является мерой способности вещества к адсорбции. Его можно сопоставить свободной энергии адсорбции в понимании Лэнгмюра 2.3.

О неполярной среде: газ или вакуум

Если мы хотим рассмотреть адсорбцию молекул на поверхность воды, то в качестве смежной с водой средой в простейшем случае можно взять пустоту (вакуум). Однако, не следует забывать, что в реальности вода будет испаряться, пока не установится равновесие между насыщенным паром воды (который образуется на месте вакуумной фазы) и жидкой водой.

Но в компьютерном моделировании мы будем рассматривать весьма небольшие системы (с размерами порядка 3 нм) и на небольших временах (около 2 нс). Принимая во внимание плотность насыщенного пара воды и кинетику испарения, можно ожидать, что в такой системе мы едва ли сможем наблюдать процесс испарения. Даже если молекула воды оторвётся от поверхности воды, объём “газового” слоя не позволит ей отойти далеко, и она

в скором времени опять сольётся с основной массой воды. Кроме того, даже если в системе будет присутствовать насыщенный пар, из-за его разреженности его можно приближённо считать идеальным газом, и особого влияния на измеряемые параметры он оказывать не будет. Поэтому в дальнейшем под словами “вакуумная фаза” и “газовая фаза” понимается одно и то же.

2.3 Классический метод молекулярной динамики

Одним из основных методов исследования статистических характеристик молекулярных систем на основе их классических моделей является метод молекулярной динамики. В данном разделе будут подробнее рассмотрены основы метода.

2.3.1 Классическая модель молекулярных систем

Для исследования молекулярных систем мы будем сопоставлять им некоторые модели классической механики. Описание молекулы в рамках классического приближения с заданием соответствующих констант взаимодействия называется *силовым полем*. На сегодняшний день разработано большое количество различных классических и полуклассических (включающих квантовые поправки) силовых полей.

2.3.2 Стандартные силовые поля

В данном исследовании использовалось силовое поле OPLS-AA ([25]) и тип воды SPC ([26]) с замороженными внутренними степенями свободы. Силовое поле OPLS-AA широко используется при моделировании биологических систем, обеспечивая достаточную быстроту и точность расчёта, о чём свидетельствуют результаты работы [27].

В поле OPLS-AA, как и почти во всех других классических силовых полях, атомы представляются в виде материальных точек, взаимодействующих по классическим законам, потенциал взаимодействия является функцией от координат всех атомов.

Межмолекулярные взаимодействия представляются в виде парных взаимодействий между атомами с потенциалами Леннарда-Джонса и Кулона:

$$U_{pair} = \sum 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] + \sum \frac{q_i q_j}{r} \quad (2.11)$$

Взаимодействия внутри молекул описываются как парными взаимодействиями 2.11, так и внутримолекулярными потенциалами:

$$U_{intmol} = U_{bond} + U_{angle} + U_{torsion} + U_{improp} \quad (2.12)$$

где U_{bond} -потенциал связи (обычно гармонический по длине связи), U_{angle} -потенциал для угла между двумя смежными связями (также обычно гармонический), $U_{torsion}$ -потенциал на торсионные (двухгранные) углы в молекуле (обычно берётся в виде суммы косинусов), U_{improp} -потенциал на ложноторсионные углы, используется, чтобы некоторые атомы оставались в одной плоскости.

Кроме того, зачастую вместо задания гармонического потенциала для атомной связи используют моделирование с постоянной длиной связи (замороженной связью), то есть накладывают голономную связь на расстояния между соседними атомами. С технической

точки зрения при моделировании таких связей используются специальные алгоритмы, такие как SHAKE, RATTLE, LINCS, SETTLE [28] и т.д.

2.3.3 Моделирование динамики классической системы

В основе метода молекулярной динамики лежит достаточно простая идея численного решения уравнений движения для классической механической системы, являющейся моделью молекулярного объекта [29].

Однако эта “простая” идея связана с достаточно важными проблемами в статистической физике. Методы молекулярной динамики прокладывают мост между представлениями об устройстве и движении систем на микроскопическом уровне и описанием макроскопических явлений методами статфизики и термодинамики. На данном этапе развития науки в этой области существуют нерешённые проблемы как теоретического, так и прикладного характера. Зачастую в МД эксперименте нас интересует не конкретное движение атомов и молекул в течение небольшого промежутка времени, а некоторые усреднённые, статистические характеристики системы. *Возможность получения термодинамических характеристик системы из расчётов молекулярной динамики остаётся, вообще говоря, теоретически не до конца обоснованной [29].*

2.3.4 Теоретические аспекты метода МД

Обратимость

Одним из проблемных моментов в понимании результатов метода МД является парадокс обратимости впервые сформулированный Лошмидтом (см. [30]). Для понимания того, каким образом из обратимых по времени уравнений движения получаются необратимые следствия макроскопической термодинамики, используются методы теории хаоса [31], а также методы компьютерного моделирования [30]. Следует отметить, что в большинстве реализаций метода МД достичь обратимости на больших временах нельзя из-за численной ошибки при решении дифференциальных уравнений.

Гипотеза эргодичности

Важным элементом в теории статистической физики является гипотеза эргодичности [32]. В полной мере это утверждение справедливо и для МД экспериментов. Следует обращать внимание, как на возможность достижения системой всех возможных микросостояний, так и на необходимые для этого времена моделирования.

Численные алгоритмы

Предполагается, что система дифференциальных уравнений движения для динамической системы крайне неустойчива по Ляпунову относительно малых возмущений ([30]). Поэтому из-за конечной точности расчёта компьютерный эксперимент в принципе не может воспроизводить точные аналитические траектории системы в фазовом пространстве. Становится невозможным, численными методами получить траекторию движения системы на достаточных временах сколько-нибудь близкую к реальной траектории системы. Однако, значительный накопленный опыт применения МД к различным системам свидетельствует о том, что использование таких “неправильных” траекторий всё же допустимо, и не сказывается на средних статистических

характеристиках системы. Предполагается наличие, так называемых, теневых орбит (shadow orbits) - настоящих траекторий механической системы в фазовом пространстве, которые находятся вблизи от “численных” траекторий на временах больших, чем времена развития неустойчивости Ляпунова в системе [33].

Среди численных алгоритмов предпочтение зачастую отдаётся алгоритму Верле (см. [29]). Алгоритм Верле обладает следующими особенностями:

1. удовлетворяет законам сохранения импульса и энергии
2. сохраняет объём фазового пространства (удовлетворяет теореме Лиувилля)
3. может быть получен из дискретного принципа наименьшего действия. Алгоритм Верле генерирует траекторию, которая соответствует экстремуму действия для некоторой “теневого” траектории (см. [29] §4.3.5)

2.3.5 Методы задания ансамблей

Важным аспектом технологии молекулярной динамики является моделирование в заданных ансамблях. В силу некоторых сложностей широко используются лишь ансамбли NVT (канонический ансамбль) и NPT (изобароизотермический ансамбль). Методы позволяющие задавать нужный ансамбль условно можно разделить на три типа:

- (a) использование уравнений термомеханики
- (b) добавление некоторого случайного шума в решение уравнений движения (термостат и баростат Андерсена, метод столкновительной динамики)
- (c) метод стохастической динамики

Метод уравнений термомеханики (см. [30]) заключается в модификации уравнений движения и добавления в них дополнительных членов линейных по импульсу. Используются такие модификации уравнений как термостат Берендсена [34], термостат Нозе, термостат Нозе-Гувера, цепи Нозе-Гувера [35], баростат Берендсена, баростат Нозе-Гувера. Убедительные теоретические и практические данные корректности реализации канонического ансамбля получены лишь для семейства термостатов Нозе и Нозе-Гувера. Если для уравнений гамильтоновой механики стационарным является любое распределение вида $f(H)$, то, например, для уравнений термомеханики Нозе-Гувера стационарным является лишь каноническое распределение $\exp(-H/kT)$ ([30]).

Все стохастические методы, упомянутые в пункте (b), правильным образом формируют соответствующие распределения по состояниям системы в фазовом пространстве, однако, могут влиять на динамические характеристики системы (коэффициенты диффузии и т.д.). Различие между методами состоит лишь в способах внесения случайного шума в систему (см. [36]).

2.3.6 Метод стохастической динамики

Для термостатирования системы также может применяться метод стохастической (броуновской) динамики с маленьким коэффициентом трения. Метод стохастической динамики позволяет правильным образом задавать распределение системы в соответствии с заданной температурой.

Суть метода состоит в численном решении уравнения Ланжевена 2.13 для молекулярной системы.

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = -m_i \xi_i \frac{d\vec{r}_i}{dt} + \vec{F}_i(\vec{r}) + \tilde{\vec{r}}_i \quad (2.13)$$

где ξ_i -коэффициент трения, $\tilde{\vec{r}}_i$ - случайный шумовой процесс такой, что

$$\langle \tilde{\vec{r}}_i(t) \tilde{\vec{r}}_j(t+s) \rangle = 2m_i \xi_i kT \delta(s) \delta_{ij}$$

Для численного интегрирования уравнения 2.13 используется, в частности, лип-фрэг алгоритм третьего порядка (см. [37]).

2.3.7 Технические аспекты моделирования

Данный раздел будет посвящён обзору некоторых технических методов применяемых при моделировании. С той или иной степенью подробности описание методик может быть найдено в книге [29]. Ниже перечислены основные методики, использованные в данной работе, и краткий к ним комментарий.

Периодические граничные условия

для уменьшения влияния граничных условий обычно при моделировании используют периодические граничные условия.

Методы ускорения расчёта

Для ускорения расчётов обычно прибегают к учёту взаимодействий лишь между ближайшими атомами, вводя так называемые радиусы обрезания. Однако, эта процедура приводит к появлению неточностей при моделировании, которые затем пытаются подправить. Для учёта межмолекулярных взаимодействий большое значение имеет не только задание самого силового поля, но и способов расчёта этих межмолекулярных взаимодействий. При параметризации большинства известных на сегодняшний день силовых полей предполагалось, что межмолекулярные взаимодействия “обрезаются” на некотором расстоянии, то есть молекулы, находящиеся на расстояниях больше некоторого, не взаимодействуют. Однако, в современных работах по молекулярному моделированию исследователи стараются использовать методики учитывающие межмолекулярные взаимодействия в полной мере. Причин этому две, во-первых, считается, что “обрезать” кулоновские взаимодействия при использовании периодических граничных условий в принципе не корректно из-за медленно убывающего характера взаимодействия. И, во-вторых, использование методологий, которые не зависят от значений радиусов обрезания, позволяет сделать результаты моделирования более универсальными, переносимыми, так как это устраняет зависимость наблюдаемых от произвольных нефизических параметров, которые могут сильно влиять на результат. При использовании компьютерного моделирования для получения количественных результатов последнее замечание становится особенно важным.

При учёте взаимодействий Ван-дер-Ваальса иногда применяют модифицированный с помощью свитч-функции потенциал Леннард-Джонса [38]. Тогда сила взаимодействия может быть

записана следующим образом:

$$F_s(r) = F(r) \quad r < r_1 \quad (2.14)$$

$$F_s(r) = F(r) + S(r) \quad r_1 \leq r \leq r_c \quad (2.15)$$

$$F_s(r) = 0 \quad r_c < r \quad (2.16)$$

$$S(r_1) = 0 \quad (2.17)$$

$$S'(r_1) = 0 \quad (2.18)$$

$$S(r_c) = -F(r_c) \quad (2.19)$$

$$S'(r_c) = -F'(r_c) \quad (2.20)$$

$$S(r) = A(r - r_1)^2 + B(r - r_1)^3 \quad (2.21)$$

$$(2.22)$$

где $F(r)$ -немодифицированная сила, $F_s(r)$ -модифицированная сила, коэффициенты A и B находятся из соответствующих условий на значение и производную функции $S(r)$ в точках r_1 и r_c . r_1 и r_c - два параметра характеризующие начало и конец обрезания потенциала. При высоких требованиях к точности расчёта используется также аналитическая поправка для “хвоста” дисперсионного взаимодействия ($\frac{1}{r^6}$). Предполагается, что на расстояниях больших, чем радиус обрезания r_c , вещество можно считать размазанным по пространству с некоторой плотностью [39]. Тогда поправка к энергии может быть записана через радиальную корреляционную функцию $g(r)$:

$$V_{cor} = \frac{1}{2}N\rho \int_0^\infty 4\pi r^2 g(r)(V(r) - V_c(r))dr \quad (2.23)$$

где $V_c(r)$ -”обрезанный” потенциал. Обычно полагают $g(r) = 1$ на расстояниях больших радиуса обрезания.

Следует отметить, что такая аналитическая поправка применима лишь для однородных сред. На Рис. 2.3 представлена аналитически вычисленная поправка к энергии гидратации молекулы метана в зависимости от радиуса обрезания взаимодействий с потенциалом Леннард-Джонса.

Из поправки к энергии системы можно вывести поправку к вириалу сил и, следовательно, давлению в системе. Для простейшего алгоритма обрезания взаимодействий на расстоянии r_c после интегрирования можно получить следующую поправку к давлению:

$$P_{cor} = -\frac{4}{3}\pi\rho^2(4\epsilon\sigma^6)r_c^{-3} \quad (2.24)$$

где ρ -плотность среды, ϵ и σ - параметры потенциала Леннард-Джонса, r_c - радиус обрезания.

Таким образом, например, при радиусе обрезания в 0.9 нм поправка к энергии составит -0.75 кДж/моль на молекулу, а поправка к давлению 280 атмосфер [40].

Учёт электростатических взаимодействий

Энергия электростатических взаимодействий согласно закону кулона убывает как $\frac{1}{r}$. Такой закон убывания взаимодействия, вообще говоря, не позволяет использовать методику обрезания взаимодействия на некотором конечном расстоянии (в случае использования периодических граничных условий). Поэтому для вычисления кулоновских взаимодействий в молекулярном моделировании широко применяются методы, в основе которых лежит

Рис. 2.3: Зависимость не учитываемой энергии взаимодействия метана с водой от радиуса обрезания потенциала Леннард-Джонса

метод сумм Эвальда [29]. Практическая реализация оригинального метода сумм Эвальда характеризуется уже не одним, а тремя параметрами: радиусами обрезания в прямом и Фурье-пространствах, а также относительным весом между прямой и обратной суммами. Однако, время вычисления суммы в обратном пространстве для метода Эвальда растёт как N^2 или $N^{3/2}$, где N -число частиц в системе.

На базе метода Эвальда существуют модифицированные методы расчёта электростатических взаимодействий такие как, например, PME (Particle Mesh Ewald) [41], PPPM (Particle-Particle Particle-Mesh), SPME (Smooth Particle Mesh Ewald). Метод PME в настоящее время реализован во многих программных пакетах. Вычислительные затраты для метода PME растут как $N \ln N$.

2.4 Элементы неравновесной статистической физики

Для расчёта изменения свободной энергии в различных состояниях молекулярных систем методами компьютерного моделирования большое значение имеют некоторые элементы неравновесной статистической физики.

Системы рассматриваемые при моделировании хотя и подчиняются законам термодинамики, но настолько малы по сравнению с макроскопическими масштабами, что для них большое значение имеют флуктуации.

2.4.1 Обобщение II начала термодинамики

Обобщение II закона термодинамики на небольшие системы (*Флуктуационная теорема*) было теоретически разработано Эвансом в начале 1990-ых годов [42] и получило экспериментальное

подтверждение в начале 21 века [43]. Одно из следствий флуктуационной теоремы может быть сформулировано следующим образом: “существует вероятность того, что энтропия неравновесной изолированной системы может спонтанно уменьшиться”. На практике такие флуктуации энтропии могут наблюдаться лишь на небольших временах в небольших системах, которые как раз и исследуются при моделировании.

2.4.2 Равенство Джарзинского

Тесно связанным, по сути, с флуктуационной теоремой является равенство Джарзинского [44], которое имеет широкое практическое применение при моделировании неравновесных процессов в небольших системах.

Допустим, что у нас имеются два состояния системы (А и В) с некоторыми значениями свободной энергии. Тогда разница свободной энергии будет равна:

$$\Delta F = F_B - F_A$$

Предположим, что для перевода системы из состояния А в состояние В мы затрачиваем работу W , тогда, основываясь на втором законе термодинамики, можно утверждать, что:

$$\Delta F \leq W \quad (2.25)$$

причём равенство выполняется только в случае бесконечно медленного квазиравновесного процесса.

Уравнение Джарзинского обобщает это неравенство на случай произвольных неравновесных процессов:

$$\exp\left(-\frac{\Delta F}{kT}\right) = \overline{\exp\left(-\frac{W}{kT}\right)} \quad (2.26)$$

Усреднение в правой части уравнения 2.26 проводится по равновесному ансамблю А-состояний системы. При этом ансамбль А-состояний может быть как микроканоническим, так и каноническим. Обобщение понятия работы на микроскопическом уровне вводится следующим образом. Допустим, что $\vec{z}(t)$ -траектория точки фазового пространства. Гамильтониан системы зависит от параметра λ : $H(\vec{z}, \lambda)$. У нас есть некоторый закон изменения λ во времени: $\lambda(t)$, который и определяет наше воздействие на систему (работу). Тогда работа внешних условий над системой может быть записана как:

$$W = \int_0^{t_s} dt \dot{\lambda} \frac{\partial H}{\partial \lambda}(\vec{z}(t)) \quad (2.27)$$

Из 2.26 напрямую следует 2.25, только под работой уже понимается средняя по ансамблю:

$$\Delta F \leq \bar{W}$$

Равенство Джарзинского может быть обобщено на случай двух процессов: перевода системы из состояния А в состояние В (прямой процесс) и обратно (обратный процесс). Этот результат может быть получен из следствий более общей *теоремы Крукса* [45]. Одним из следствий этой теоремы является следующее утверждение 2.28:

$$\ln\left(\frac{P_F(W)}{P_R(-W)}\right) = kT(W - \Delta F) \quad (2.28)$$

где $P_F(W)$ и $P_R(W)$ - функции плотности вероятности для величины работы W , затрачиваемой при прямом и обратном процессах соответственно.

2.5 Практические методы расчёта свободной энергии

Итак, предположим, что молекулярная система характеризуется некоторым Гамильтонианом, зависящим от параметра: $H(\lambda)$. Будем говорить, что каждому значению λ соответствует некоторое λ -состояние системы. Каким образом возможно рассчитать разницу в свободной энергии между двумя λ -состояниями?

Начиная с 1960-ых годов для практической реализации в численных экспериментах широкое применение получили такие методы, как *метод термодинамического интегрирования*, *метод возмущения свободной энергии*, *метод Беннета*. С появлением работ Джарзинского, Крукса, Shirts et al. [44, 45, 46] выявилась более глубокая связь между этими методами. А соотношения 2.26 и 2.28 предлагают более общий подход к решению задачи отыскания свободной энергии через неравновесные процессы. При этом, упомянутые ранее методы являются частными случаями в этом общем подходе, но могут также быть получены независимыми способами. В данной работе нам будут интересны эти три метода, так как они имеют широкое практическое применение.

2.5.1 Метод термодинамического интегрирования

Метод термодинамического интегрирования (ТИ) может быть получен из 2.26-2.27 при бесконечно медленном изменении параметра λ со временем. Кроме того метод ТИ может быть получен и из более простых соображений. Свободная энергия Гиббса системы даётся в общем случае выражением 2.29.

$$G = -kT \ln \int \exp \left(-\frac{H(V, \vec{p}, \vec{q}, \lambda)}{kT} \right) d\vec{p} d\vec{q} dV \quad (2.29)$$

Для метода ТИ основным является равенство 2.30.

$$\frac{\partial G}{\partial \lambda} = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle \quad (2.30)$$

И разница свободной энергии между разными λ -состояниями даётся выражением 2.31.

$$\Delta G = \int \frac{\partial G}{\partial \lambda} d\lambda \quad (2.31)$$

Метод достаточно эффективен и часто применяется. Однако, необходимость интегрирования функции $\frac{\partial G}{\partial \lambda}$ требует достаточно большого количества вычислений и вносит дополнительную ошибку численного интегрирования, которую трудно оценить при ограниченном наборе известных значений функции.

2.5.2 Метод возмущения свободной энергии

Метод возмущения свободной энергии (Free energy perturbation - FEP) является другим предельным случаем соотношений 2.26-2.28, когда $d\lambda/dt = \infty$, то есть работа над системой

совершается бесконечно быстро, адиабатически. В этом случае разница свободной энергии между разными λ -состояниями может быть записана в виде 2.32.

$$\Delta G = -kT \ln \frac{\int dV \vec{p} d\vec{q} \exp \left\{ -\frac{H(V, \vec{p}, \vec{q}, \lambda_1)}{kT} \right\}}{\int dV \vec{p} d\vec{q} \exp \left\{ -\frac{H(V, \vec{p}, \vec{q}, \lambda_0)}{kT} \right\}} = -kT \ln \langle \exp \frac{H_0 - H_1}{kT} \rangle_0 \quad (2.32)$$

где $H_0 = H(V, \vec{p}, \vec{q}, \lambda_0)$, $H_1 = H(V, \vec{p}, \vec{q}, \lambda_1)$, усреднение ведётся по NPT-ансамблю соответствующему состоянию системы с $\lambda = \lambda_0$.

2.5.3 Метод Беннета

Метод Беннета (Bennet acceptance ratio method) позволяет наиболее эффективным образом комбинировать данные, получаемые при прямом и обратном процессе. Иными словами, мы переводим систему из состояния с $\lambda = \lambda_0$ в состояние с $\lambda = \lambda_1$ и при этом получаем некоторое значение работы W_F . Однако возможен и обратный процесс перевода системы из состояния λ_1 в состояние λ_0 с работой W_R . В 1976 году [47] метод был предложен Беннетом как обобщение метода FEP. Однако в настоящее время стало ясно, что таким образом можно обобщить практически любые методы расчёта свободной энергии. Этот *обобщённый метод Беннета* основывается на соотношениях Крукса 2.28 и позволяет построить наилучшую статистику (оценочную функцию максимального правдоподобия) для оценки значения свободной энергии по наборам параметров W_F и W_R (см. [46]).

В простейшем случае как обобщение метода FEP метод Беннета (оригинальная работа [47], 1976 год) имеет следующий вид. Итак, допустим, что у нас есть набор работ по бесконечно быстрой трансформации системы из состояния λ_0 в состояние λ_1 - W_F , и обратно - W_R . Метод Беннета задаёт оценочную функцию, которая определяет значение разницы свободной энергии с минимальной дисперсией среди всех оценочных функций. Оценочная функция задаётся неявно, как решение уравнения 2.33 по параметру C .

$$\sum_m^{N_R} f(W_R^m + C) = \sum_n^{N_F} f(W_F^n - C) \quad (2.33)$$

В уравнении (2.33) f -функция Ферми-Дирака:

$$f(x) \equiv \frac{1}{1 + \exp \beta x}$$

N_F и N_R - соответственно количество полученных замеров в прямом и обратном направлениях, $\beta = 1/kT$. Можно показать, что наилучший результат всегда достигается при $N_F \approx N_R$. Более того метод Беннета имеет преимущество перед простым методом FEP. Так, более эффективным является осуществление выборки значений работы в прямом и обратном направлениях, и затем применение метода Беннета, чем осуществление в два раза большей выборки в одном направлении и применении метода FEP.

2.6 Расчёт СЭС методами МД

Итак, расчёт свободной энергии сольватации вещества S в жидкости A (формула 2.6) может быть сведён к расчёту разницы свободной энергии между двумя состояниями ($\lambda = 0$ и $\lambda = 1$) системы, которая представляет из себя некоторое количество жидкости A и растворённую в ней молекулу S . При этом состояние $\lambda = 1$ соответствует системе, в которой молекула S взаимодействует с окружающей её жидкостью, а состояние $\lambda = 0$ - в котором молекула не взаимодействует с молекулами жидкости.

Однако, способы реализации расчёта могут быть весьма различными, что отражается на их скорости и точности. Для организации расчёта необходимо выбрать некоторую методику, которая состоит из многих элементов. Основные элементы приведены в таблице 2.1.

Элемент метода определения СЭС	Возможные варианты
Силовые поля	Большое множество (AMBER, OPLS-AA, CHARMM и т.д.) включая различные модели воды (TIP3P, SPC, SPC-E, TIP4P, TIP5P и т.д.)
Метод расчёта разницы свободной энергии между состояниями	FEP, ТИ, метод Беннета и т.д.
Путь превращения системы $H(\lambda)$	Различные варианты см. пункт 2.6.1
Количество этапов в превращении системы (в случае метода FEP и Беннета)	
Метод генерации выборки по заданному ансамблю	МД, Стохастическая динамика, Монте-Карло, гибридный Монте-Карло и т.д.
Особенности учёта взаимодействий при генерации выборки	Радиусы обрезания, учёт электростатических взаимодействий и т.д.

Таблица 2.1: Таблица-схема элементов методики по определению свободной энергии сольватации в компьютерном эксперименте

2.6.1 Путь превращения системы

Важным элементом при расчёте СЭС является зависимость Гамильтониана от λ . Так обычно, отключение взаимодействия происходит постепенно. Сначала снимается кулоновское взаимодействие, а затем уже взаимодействие Леннард-Джонса. Для отключения кулоновского взаимодействия обычно используют линейный закон зависимости потенциальной энергии от λ :

$$U(\lambda_C, \lambda_{LJ}) = U_{solv} + U_{mol} + \lambda_C U_{int}^{coulomb} + U_{int}^{VdW}(\lambda_{LJ})$$

Для отключения взаимодействия Леннард-Джонса в литературе предлагаются различные подходы. Общее мнение состоит в том, что линейная параметризация крайне неэффективна из-за сингулярности потенциала Леннард-Джонса в нуле. Поэтому используют потенциалы

с “мягким ядром” вида 2.35:

$$U^{VdW}(\lambda_{LJ}) = \lambda_{LJ}^t \sum_{ij} 4\epsilon_{ij} \left(\left(\frac{\sigma}{r_S} \right)^{12} - \left(\frac{\sigma}{r_S} \right)^6 \right) \quad (2.34)$$

$$r_S = (\alpha\sigma^6\lambda_{LJ}^p + r_{ij}^6)^{1/6} \quad (2.35)$$

Здесь α, p и t -параметры этого потенциала. В литературе нет пока чётких указаний по наилучшему выбору этих параметров. По данным работы [48] значения $t = 1, p = 1$ являются более оптимальными.

2.6.2 Особенности реализации метода ТИ

При использовании метода термодинамического интегрирования необходимо стараться задать такой путь превращения системы $H(\lambda)$, чтобы получающаяся в результате зависимость $\frac{dG}{d\lambda}(\lambda)$ была как можно более гладкой и легко интегрируемой.

2.6.3 Особенности реализации метода Беннета и FEP

При использовании метода Беннета и метода FEP следует руководствоваться следующими соображениями. Обратимся к уравнению 2.32. Здесь процедура усреднения будет работать эффективно (с достаточной точностью) если разница энергии между λ -состояниями 0 и 1 для генерируемых конформаций из равновесного ансамбля состояний 0 будет не слишком велика. То есть конформации с достаточной вероятностью встречающиеся в системе 0 будут встречаться и в системе 1. Таким образом, распределения плотности вероятности по конформациям в фазовом пространстве для систем 1 и 0 должны перекрываться. И чем больше они перекрываются, тем выше эффективность метода.

Принимая во внимание вышесказанное, обычно при расчёте энергии сольватации рассчитывают разницу энергии не сразу между взаимодействующим и не взаимодействующим состоянием, а разбивают процесс на некоторое количество этапов. Пользуясь формализмом пути превращения эти этапы можно задавать как λ -состояния системы с некоторыми значениями λ .

2.6.4 Источники ошибок

При определении СЭС в компьютерном эксперименте присутствуют как статистические, так и систематические ошибки. В таблице 2.2 кратко представлены типы ошибок и их происхождение.

2.6.5 Обзор работ

В литературе представлено достаточно большое количество работ по определению свободной энергии гидратации различными методами в компьютерном моделировании. Однако среди более ранних работ не всегда могло быть приведено однозначное сравнение результатов из-за различных методик по ускорению вычислений. Так многие работы проводились с разными значениями радиусов обрезания без использования соответствующих поправок. Наиболее масштабное исследование по определению свободной энергии гидратации для различных силовых полей и различных моделей воды было проведено в работах [27, 48]. По результатам работ было произведено сравнение качества силовых полей для описания

Ошибка	Примечание
Систематическая ошибка силового поля	Мы измеряем не столько реальную энергию, сколько энергию сольватации для данной модели
Ошибка, связанная с вычислительными процедурами	Прежде всего её происхождение связано с различными приближениями и упрощениями, которые используются, чтобы ускорить расчёт. Например, введение всевозможных радиусов обрезания взаимодействий.
Стандартная статистическая ошибка	Так как мы измеряем некоторую статистическую величину, которая сама по себе обладает дисперсией, то ограниченная выборка порождает некоторую статистическую ошибку
Статистическая ошибка, связанная с плохим качеством выборки (Погрешность выборки)	Ошибка возникает, если время моделирования было недостаточным, чтобы выборка покрыла все значимые состояния системы. Такое может происходить, если на поверхности свободной энергии присутствуют некоторые барьеры между локальными минимумами. Один из простейших примеров, это анти-гош конформации молекул.
Конечный размер моделируемой системы	Небольшой по макроскопическим меркам размер моделируемой системы может приводить к различным артефактам

Таблица 2.2: Таблица-классификация различных типов ошибок в определении СЭС компьютерными методами

эффекта гидратации. В общем же, для классических силовых полей систематическая среднеквадратичная ошибка в определении энергии гидратации составляет порядка 1 ккал/моль. Оказалось, что поле OPLS-AA описывает энергии гидратации несколько лучше остальных. Для полей AMBER, CHARMM, OPLS-AA соответствующие среднеквадратичные ошибки составляют 1.35, 1.31, 0.85 ккал/моль.

2.7 Моделирование границы раздела фаз

Моделирование границы раздела жидкой фазы представляет достаточно большой интерес с точки зрения молекулярного моделирования. Однако, основной проблемой для молекулярного моделирования является неспособность предсказать толщину границы раздела фаз [49]. Общепринятой является теория капиллярных волн. Она предсказывает следующее значение

для толщины раздела фаз:

$$\sigma_{cw}^2 = \frac{kT}{4\pi^2\gamma} \int_{q_{min}}^{q_{max}} \frac{1}{q^2 + k_g^2} d^2q$$

где ρ -плотность среды, γ -поверхностное натяжение, $k_g = \rho g/\gamma$, $q_{max} = 2\pi/\xi$, ξ -корреляционная длина, которая обычно соответствует размеру молекулы, q_{min} -наименьший волновой вектор. В случае компьютерного моделирования $q_{min} = 2\pi/L$, где L -размер элементарной ячейки. Тогда после некоторых приближений и уточнений (см. [49]) можно получить следующее выражение:

$$\sigma_{cw}^2 = \frac{kT}{2\pi\gamma} \ln \frac{L}{\xi}$$

Компьютерное моделирование подтверждает такую логарифмическую зависимость толщины интерфейса от размера элементарной ячейки. Профили плотности жидкость/пар обычно хорошо аппроксимируется зависимостью

$$\psi(z) = \frac{1}{2} \left((\rho^{liq} + \rho^{vap}) - (\rho^{liq} - \rho^{vap}) \operatorname{erf} \left[\frac{(z - z_0)}{\sigma\sqrt{2}} \right] \right)$$

где σ -толщина границы раздела, а z_0 - положение поверхности Гиббса.

2.8 Расчёт ПСС методами МД

Наиболее используемые методы расчёта потенциала средней силы в компьютерном моделировании это [40]:

- Метод АСМ тяги (AFM pulling): к группе атомов условно привязывается один конец пружинки, а другой конец пружины начинают медленно смещать. При этом вычисляется работа силы. Для анализа работы силы можно применить равенство Джарзинского 2.26.
- Метод сдерживающих сил (constraint forces): расстояние между центрами масс необходимых атомных групп фиксируется в процессе динамики с помощью какого-либо удерживающего алгоритма (SHAKE, LINCS и т.д.), то есть налагается некоторая голономная связь. При этом на каждом шаге вычисляется сила реакции связи, которая необходима для поддержания удерживающей связи.
- Метод зонтичной выборки (Umbrella sampling): к Гамильтониану системы добавляется дополнительный гармонический потенциал (пружинка) на расстояние между некоторыми исследуемыми атомными группами. При этом мы получаем смещённую выборку, однако эта выборка затрагивает области фазового пространства представляющие для нас интерес. С помощью простых статфизических рассмотрений можно учесть смещение выборки и получить необходимый результат.
- Возможны другие варианты, в частности, можно использовать подход с параметризацией Гамильтониана по λ , где от параметра зависит расстояние между группами атомов.

Для задач этой работы наиболее простым и понятным является применение метода сдерживающей силы.

2.8.1 Обзор работ

В работе [50] определялись профили свободной энергии для молекул углекислого газа и азота методами молекулярной динамики. Аналогичных работ по R-H-аналогам аминокислот нами в литературе найдено не было.

Глава 3

Определение свободной энергии гидратации боковых цепей аминокислот

3.1 Аннотация

Произведена численная оценка свободной энергии гидратации соединений соответствующих боковым цепям незаряженных аминокислотных остатков (кроме пролина) методами классического компьютерного моделирования. Для генерации выборки по состояниям использовался метод стохастической динамики с силовым полем OPLS-AA и моделью воды SPC. Для вычисления свободной энергии сольватации применялся метод Беннета. Достигнута статистическая точность в определении энергии порядка 0.1 кДж/моль. Обсуждается сравнение метода Беннета с методом термодинамического интегрирования. Анализируется происхождение ошибок в определении энергии.

3.2 Введение

Определение свободной энергии сольватации с хорошей точностью в атомистическом компьютерном эксперименте является достаточно трудоёмкой задачей. С практической точки зрения, например, в фармакологии, обычно необходимо определение энергетических параметров с погрешностью не более 2 кДж/моль. Средние систематические ошибки, которые уже заложены в классических силовых полях, составляют 3-6 кДж/моль. Однако, наличие методов, которые позволяют быстро оценивать термодинамические параметры для разных силовых полей, открывает возможности для дальнейшего улучшения этих полей. Возможность осуществлять необходимые по вычислительному объёму расчёты появилась сравнительно недавно.

3.3 Силовое поле

В работе [27] было произведено вычисление энергий гидратации для боковых цепей аминокислотных остатков с использованием трёх разных силовых полей (OPLS-AA, CHARMM22 и AMBER99) и моделью воды TIP3P. Было показано, что среди этих силовых полей поле OPLS-AA даёт результат наиболее близкий к экспериментальному. В работе [48] были проанализированы значения энергии гидратации для поля OPLS-AA и различных моделей воды (TIP3P, SPC, SPC/E, TIP4P, TIP5P, TIP3P-MOD и др.). Наиболее используемыми из

Рис. 3.1: Схематичное представление парциальных зарядов в соответствии с полем OPLS-AA для аминокислотного остатка серина и скорректированные заряды для его R-Н-аналога (справа). По [27].

них на сегодняшний день являются модели TIP3P и SPC. По результатам работы [48] для этих двух моделей наблюдается одинаковое среднеквадратичное отклонение вычисленных значений энергии гидратации от экспериментальных. Лучший результат показала лишь модель TIP3P-MOD, однако, следует учесть, что она широко не используется, и специально создавалась путём модификации параметров модели TIP3P для лучшего описания взаимодействия молекулы метана с водой.

Исходя из вышеприведённых результатов, нами было решено использовать силовое поле OPLS-AA [25] и модель воды SPC [26].

На основе силового поля OPLS-AA для аминокислотных остатков были созданы топологии для веществ-аналогов боковых цепей. Вещество-аналог боковой цепи получалось отделением от аминокислотного остатка остова и заменой C- α -атома на атом водорода. Заряд на C- β -атоме корректировался соответствующим образом, чтобы система была в целом электронейтральна (см. Рис. 3.1). Параметры взаимодействия Леннарда-Джонса для C- β -атома соответствовали стандартному алифатическому типу углеродных атомов поля OPLS-AA. Для атомов серы брались несколько модифицированные параметры из поля OPLS-AA/L [51]. Исследованные вещества приведены в приложении А.

Взаимодействие между несвязанными атомами сводилось к парным взаимодействиям с потенциалами Леннарда-Джонса и Кулона (см. 2.11).

3.4 Метод расчёта свободной энергии

Как было показано, определение свободной энергии сольватации фактически сводится к расчёту свободной энергии взаимодействия молекулы с водным окружением. То есть расчёту разницы свободной энергии между системами с полным Гамильтонианом:

$$H_0 = H_{water} + H_S + H_{inter}$$

и системой с Гамильтонианом без последнего члена, описывающего энергию взаимодействия молекулы вещества S с растворителем:

$$H_1 = H_{water} + H_S$$

Однако, в силу технических сложностей по выделению Гамильтониана взаимодействия H_{inter} при моделировании иногда используют другой путь. Учитывая, что свободная энергия является функцией состояния системы и следовательно не зависит от способа достижения состояния, предлагается следующий термодинамический цикл Рис. 3.2. Фактически удобно бывает превратить взаимодействующую молекулу вещества S в воде в некоторое невзаимодействующее состояние N. А потом в вакууме посчитать энергию превращения молекулы N обратно в молекулу S. Третья часть цикла на рисунке 3.2, а именно ΔF_2 (помещение невзаимодействующей молекулы N в растворитель из вакуума) всегда равняется 0. Дополнительное состояние

Рис. 3.2: Термодинамический цикл по определению свободной энергии сольватации через дополнительные “ненатуральные” состояния.

N как раз выбирается из удобства представления Гамильтониана силовым полем. Удобно обычно выключать не взаимодействие молекулы с растворителем, а занулять на молекуле S заряды и убирать взаимодействие с потенциалом Леннард-Джонса.

$$H_{inter} = U_{inter}^{LJ} + U_{inter}^{qq}$$

Поэтому разница между молекулой S и молекулой N обычно состоит в отсутствии электростатических и Ван-дер-ваальсовых взаимодействий внутри самой молекулы. Далее везде в работе под H_{inter} будет пониматься энергия взаимодействия атомов молекулы S как с окружением, так и друг с другом. А при вычислении свободной энергии будет учитываться поправка на взаимодействие атомов друг с другом.

Путь превращения системы задавался следующим образом. Сначала по линейному закону отключалось кулоновское взаимодействие:

$$U_{inter}^{qq} = \lambda_{qq} \sum_{ij} \frac{q_i q_j}{r_{ij}}$$

Затем отключалось взаимодействие Леннард-Джонса, для этого использовалась “мягкоядерная” интерполяция потенциала следующего вида:

$$U_{inter}^{LJ}(\lambda_{LJ}) = \lambda_{LJ} \sum_{ij} 4\epsilon_{ij} \left(\frac{1}{(0.5(0.3)^6 \lambda_{LJ} + r_{ij}^6)^2} - \frac{1}{(0.5(0.3)^6 \lambda_{LJ} + r_{ij}^6)} \right)$$

где r_{ij} -задаётся в нанометрах.

Таким образом, для вычисления энергии гидратации вычислялись энергии 4 процессов: энергия разрядки молекулы в воде (ΔF_C^{wat}), энергия выключения взаимодействий Леннард-Джонса у уже разряженной молекулы в воде (ΔF_{VdW}^{wat}), и энергии тех же процессов в вакууме ΔF_C^{vac} и ΔF_{VdW}^{vac} . Искомая энергия гидратации вычислялась по формуле:

$$\Delta F_{hydr} = \Delta F_C^{vac} + \Delta F_{VdW}^{vac} - \Delta F_C^{wat} - \Delta F_{VdW}^{wat}$$

Для вычисления разницы свободной энергии в основном использовался метод Беннета. При этом весь путь превращения разбивался на некоторое количество дополнительных состояний. Последовательно вычислялась разница свободной энергии между ближайшими состояниями, которая затем суммировалась для получения конечного значения. Промежуточные состояния, которые использовались при вычислении свободной энергии с использованием метода Беннета, приведены в таблице 3.1.

Номер состояния п/п	λ_{qq}	λ_{LJ}
1(полное взаимодействие)	1	1
2	0.5	1
3	0	1
4	0	0.8
5	0	0.7
6	0	0.6
7	0	0.5
8	0	0.4
9	0	0.3
10	0	0.2
11(отсутствие взаимодействия)	0	0

Таблица 3.1: Таблица с параметрами промежуточных состояний, использованных при расчёте энергии гидратации методом Беннета

Для молекулы метана расчёты также были выполнены с применением метода термодинамического интегрирования. В этом случае производная Гамильтониана по параметру вычислялась в 21 точке с интервалом 0.05 (предполагалось, что $\lambda = \lambda_{qq} = \lambda_{LJ}$).

3.5 Получение выборки

Для создания выборки состояний и вычисления параметров различных конфигураций молекулярной системы использовался программный пакет GROMACS 3.3 [52]. Применялся метод стохастической динамики с коэффициентом трения 1 ps^{-1} и соответствующим случайным процессом при температуре 300K. Кроме того, для создания NPT ансамбля использовался

алгоритм баростатирования Паринелло-Рамана [53] с временем релаксации 2 пс, заданным давлением 1 атм и заданной сжимаемостью воды $4.5 * 10^{-5}$ бар⁻¹. Для интегрирования использовался лип-форг алгоритм третьего порядка [37] с шагом 2 фс. Вычисления проводились с двойной точностью. Кроме того моделирование проводилось с “замороженными” связями между атомами. Для молекул воды использовался алгоритм SETTLES [54], а для боковых цепей аминокислот алгоритм LINCS [55] четвёртого порядка с двумя итерациями на шаг.

Моделируемая система состояла из одной молекулы R-H-аналога аминокислотного остатка и около 900 молекул воды (см. Рис. 3.3). Использовались периодические граничные условия, размер элементарной ячейки составлял порядка $3 \times 3 \times 3$ нм³. Для учёта взаимодействий Леннард-Джонса использовалась следующая методика. Сила взаимодействия плавно обрезаалась на расстоянии от 1 до 1.2 нм по закону 2.22. При этом потенциал непрерывен вместе со своей производной, однако вторая производная потенциала уже не непрерывная функция, это должно приводить к небольшому несохранению энергии, однако, для практических применений, как было показано в работе [48], это не существенно. Для учёта “хвостов” взаимодействия применялась корректировка дисперсионного взаимодействия вида 2.23. Соответствующая поправка добавлялась и к вириалу сил при расчёте давления.

Рис. 3.3: Графическое представление моделируемой системы для определения энергии гидратации молекулы бутана.

Для вычисления электростатических взаимодействий использовался алгоритм PME (Particle Mesh Ewald [41]) с радиусом обрезания в прямом пространстве 1.3 нм, порядком интерполяции 6, параметром относительной точности между реальной и фурье частью энергии 10^{-6} . Расстояние между узлами решётки, определяющее точность в обратном пространстве, равнялось 0.12 нм.

Для построения списка соседей использовался ячеечный метод, список соседей обновлялся каждые 10 шагов. Радиус создания списка соседей был выбран равным 1.3 нм.

Перед каждым этапом сбора данных производилась релаксация системы в течение 1 нс, затем производился сбор данных в течении 5 нс. Конфигурации записывались каждые

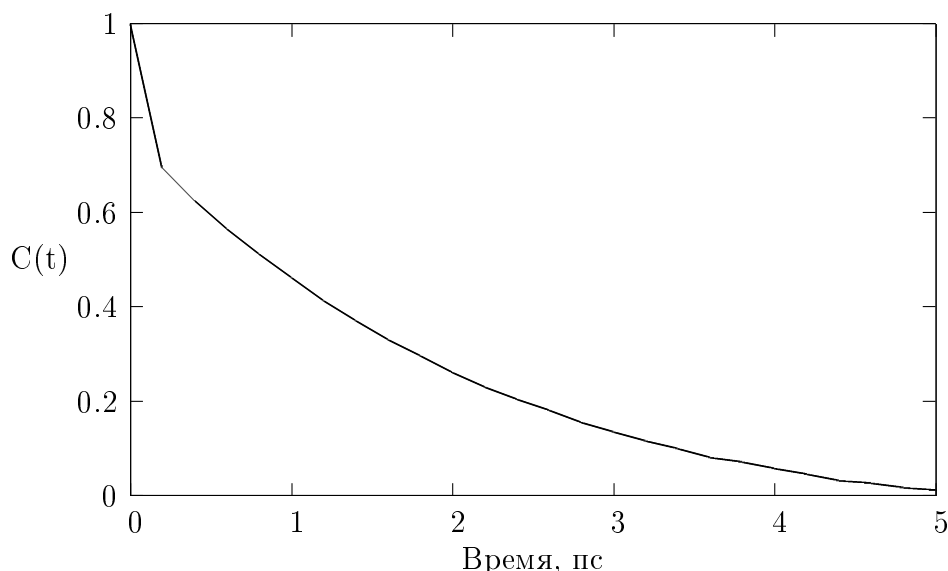


Рис. 3.4: Автокорреляционная функция потенциальной энергии одной из систем

1000 шагов, то есть раз в 2 пс. На рисунке 3.4 приведена автокорреляционная функция энергии одной из систем. Из графика видно, что запись данных раз в 2 пс обеспечивает достаточно некоррелированный набор данных. Всего для каждого значения λ усреднение производилось по 2500 конформациям.

3.6 Анализ ошибок

Первым шагом на пути определения свободной энергии является задание силового поля. Теоретически, задав силовое поле, мы тем самым определим и значение энергии гидратации. Однако, поскольку любое силовое поле является лишь попыткой представить реальную систему в виде некоторой модели классических тел, то, задавая эту модель, мы уже вносим некоторое отклонение от реальности. Другим источником неточности в определении энергии является уже погрешность определения величины в процессе расчёта. Эту погрешность условно можно разделить на два вклада: статистическая погрешность самого расчёта из-за термодинамических флуктуаций и возможная ошибка, вызванная тем, что система за время расчёта не побывала достаточное количество раз во всех статистически значимых конформациях. Назовём эту погрешность *погрешностью выборки*. Другими словами, чтобы избежать последней ошибки, необходимо, чтобы была произведена выборка состояний по всему доступному объёму фазового пространства, а не только в окрестности какого-либо локального минимума свободной энергии.

Возможны в принципе ещё ошибки связанные с огрублённым вычислением взаимодействий и конечным размером системы, однако, основываясь на результатах работ [27, 48], можно утверждать, что при выбранном наборе параметров эти ошибки являются незначительными.

3.6.1 Статистическая погрешность

Статистическая погрешность оценивалась с помощью блочного усреднения (block averaging). Набор из 2500 независимых измерений разбивался на 5 блоков по 500 измерений. Для каждого блока вычислялись среднее и дисперсия, которая оценивалась по следующей

стандартной формуле:

$$\sigma^2 = \sum \frac{(\langle X \rangle - X)^2}{(n - 1)}$$

Конечная величина свободной энергии являлась суммой изменений энергии в результате нескольких последовательных трансформаций. Дисперсия конечной величины вычислялась как сумма дисперсий свободной энергии на каждом этапе. Применяв распределение Стьюдента, можно полагать, что точное значение энергии лежит с 90% вероятностью в интервале от $\langle X \rangle - \sigma$ до $\langle X \rangle + \sigma$.

В случае метода термодинамического интегрирования статистическая ошибка оценивалась по временному ряду стандартным образом, однако, к статистической ошибке ещё добавлялась погрешность интегрирования, которая оценивалась как разница между значениями интеграла, вычисленного методом трапеций по всем точкам, и интеграла, вычисленного по набору с в два раза меньшим количеством точек. Следует отметить, что при получающемся типе функций, такая оценка является достаточно разумной. Вычислялся также и верхний предел оценки погрешности интегрирования. При получающемся достаточно гладком типе функций верхняя оценка может быть произведена следующим образом. Весь интервал разбивается на отрезки, содержащие три точки данных. На каждом отрезке из трёх точек считается интеграл методом трапеций по трём точкам и по двум крайним. Разница между этими значениями является максимальной оценкой погрешности на отрезке. Сумма модулей этих погрешностей даёт максимальную оценку погрешности интеграла на полном отрезке интегрирования. Следует, однако, отметить, что при достаточно гладком виде функции верхняя граница обычно сильно превосходит реальную погрешность.

3.6.2 Погрешность выборки

Под погрешностью выборки мы понимаем погрешность, связанную с тем, что в системе существуют некоторые степени свободы, времена корреляции которых больше или сопоставимы со временем наблюдения за системой. И при этом покрытие выборкой данных степеней свободы важно для конечного усреднения.

Более подробно поясним сказанное. Одними из важнейших степеней свободы являются степени, связанные с перераспределением молекул воды вокруг растворённой молекулы. Чтобы убедиться, что степени свободы связанные с положением молекул воды относительно исследуемой молекулы в достаточной степени покрываются при выборке, можно оценить среднее время, проводимое молекулой воды в первой координационной сфере вокруг молекулы воды. Такие оценки проводились в работе [27] и было получено значение порядка 0.5 пс, что сильно меньше времени моделирования. Аналогичный вывод можно было бы сделать и исходя из коэффициента диффузии воды. Для воды SPC он равен примерно $2.3 \cdot 10^{-5}$ см²/с, а практически для всех компьютерных моделей воды он обычно оказывается на 100-150% больше. Тогда не сложно получить оценку, что средний квадрат смещения молекулы воды станет порядка квадрата длины ячейки (3 нм) за 0.5 пс.

Более сложным является вопрос о покрытии выборкой внутренних степеней свободы самой растворённой молекулы. Большинство исследованных молекул являются небольшими, практически жёсткими, и для них такой вопрос не стоит. Однако, более крупные молекулы с “медленными” внутренними степенями свободы вызывают больше вопросов.

Среди исследованного набора молекул медленными степенями свободы обладают молекулы с торсионной подвижностью. Это в основном относится к R-H-аналогам метионина и изолейцина. На Рис. 3.5 приведена зависимость торсионного угла вдоль C-C связи для аналога метионина. Видно, что за время вычисления совершается достаточно много переходов между анти- и гош-ротамерами. Следовательно, можно заключить, что используемого времени моделирования достаточно для генерации хорошей выборки по этой степени свободы. Разумно предположить, что при всех других значениях λ , когда заряды и объёмные взаимодействия атомов играют меньшую роль, переходы между разными конформациями будут лишь облегчаться.

Рис. 3.5: График зависимости расстояния между крайними C-атомами в молекуле аналоге метионина от времени. На графике хорошо видны анти-гош-переходы.

Сложнее обстоит дело с молекулой аналогом изолейцина – бутаном. На Рис. 3.7 приведена зависимость торсионного угла вдоль C-C связи для молекулы бутана. Анти- состояние бутана (Рис. 3.6), как известно, более стабильно, чем гош-состояние. Выигрыш в анти-состоянии по сравнению с гош-состоянием составляет около 0.9 ккал/моль, то есть 1.5 кТ. В зависимости на Рис. 3.7 молекула изначально была помещена в гош-конформацию, и затем перешла в стабильную анти-конформацию.

Следовательно, нельзя сделать вывода о том, что времени моделирования достаточно для покрытия этой степени свободы. Однако, в работе [27] было показано, что разница в энергии гидратации для анти- и гош-ротамеров невелика, порядка 0.2 ± 0.17 кДж/моль. То есть погрешность выборки при определении энергии сольватации бутана будет невелика, однако этот аспект нужно иметь в виду.

3.7 Сравнение ТИ и метода Беннета

Для молекулы метана вычисление свободной энергии гидратации было выполнено методами термодинамического интегрирования и методом Беннета, описанными выше способами. На Рис. 3.8 приведена полученная зависимость усреднённой производной гамильтониана по λ от λ при вычислении энергии гидратации метана методом ТИ. После интегрирования зависимости методом трапеций было получено значение энергии гидратации. В таблице

Рис. 3.6: Молекула бутана в анти (слева) и гаус-конформациях

Рис. 3.7: График зависимости расстояния между крайними С-атомами в молекуле бутана.

3.2 приведены данные для сравнения двух методов. Следует отметить, что в методе ТИ интегрирования использовалось одновременное отключение кулоновского и Леннард-Джонсовского взаимодействий, тогда, как в работе проведённой методом Беннета, эти взаимодействия отключались по очереди. Однако, это не должно отразиться на результатах, так как свободная энергия отключения кулоновского взаимодействия молекулы метана с водой равняется нулю в пределах статистической погрешности.

Время в таблице соответствует процессорному времени при расчёте в пакте GROMACS 3.3. Для тестирования использовался процессор Pentium 4 Xeon с тактовой частотой 3.2 ГГц. В таблице приведены лишь результаты для двух конкретных измерений, однако, из их сравнения уже можно сделать некоторые выводы. В проведённых измерениях вычисление энергии методом термодинамического интегрирования заняло примерно в два раза больше времени. Это связано с тем, что вычисления проводились для в два раза большего количества состояний, а также с тем, что для метода ТИ в процессе расчёта приходится считать не только энергию системы, но и производную энергии по параметру λ . Результирующая статистическая погрешность при этом всё равно оказалась больше, чем в методе Беннета. Однако, в методе ТИ присутствует также погрешность интегрирования, которую обычно

Рис. 3.8: График зависимости усреднённой производной гамильтониана по λ от λ при вычислении энергии гидратации метана методом ТИ.

	ΔF_{hydr} , кДж/моль	Станд. отклон.	Погр. интегр./Верхний предел
ТИ	9.195	0.053	0.002/0.46
МБ	9.199	0.041	нет
	Количество вычислений	Время вычисления	Полное время
ТИ	21	46 часов	40 дней
МБ	11	44 часа	20 дней

Таблица 3.2: Таблица с данными по сравнению метода термодинамического интегрирования и метода Беннета на примере определения энергии гидратации метана.

не так легко оценить. В таблице приведена верхняя оценка (0.46 кДж/моль), но она явно сильно завышена. С другой стороны интеграл взятый по в два раза меньшему набору данных отличается всего на 0.002 кДж/моль, то есть фактически не отличается. Это означает, что в методе ТИ можно было бы сократить время вычисления в два раза. Однако статистическая погрешность при этом увеличилась бы до 0.07 кДж/моль.

Можно сделать следующий вывод: несмотря на то, что параметры метода ТИ интегрирования были выбраны далеко не оптимальным образом, результаты свидетельствуют о том, что метод Беннета является более эффективным. Он за то же время даёт меньшую статистическую погрешность (0.04 vs 0.07) и свободен от проблем с оценкой ошибки интегрирования.

3.8 Выбор промежуточных состояний

Руководящей идеей при выборе промежуточных состояний для вычисления свободной энергии методом Беннета является минимизация погрешности. Наиболее выгодным вариантом является случай, когда погрешности на всех этапах “мутации” молекулы являются величинами одного порядка. В таблице 3.3 приведены погрешности, получающиеся при отключении взаимодействий в системе: молекула изобутана (R-H-аналог LEU) в воде. Из таблицы видно, что при заданных зависимостях потенциалов от λ и выбранных значениях λ , свободные энергии трансформации на каждом этапе приблизительно одного порядка, что

Обозначение $\lambda_{qq}, \lambda_{LJ}$	состояния	ΔF , кДж/моль	Погрешность, кДж/моль
1.0 1.0	(полное взаимодействие)		
		-2.007	0.00067
0.5 1.0			
		-1.700	0.0043
0.0 1.0			
		0.924	0.0097
0.0 0.8			
		0.727	0.0095
0.0 0.7			
		0.602	0.0093
0.0 0.6			
		0.109	0.0089
0.0 0.5			
		-0.889	0.021
0.0 0.4			
		-3.003	0.040
0.0 0.3			
		-6.016	0.0058
0.0 0.2			
		-6.032	0.059
0.0 0.0	(отсутствие взаимодействия)		

Таблица 3.3: Данные по разностям свободных энергий между промежуточными состояниями в системе: молекула изобутана в воде.

обеспечивает минимальность конечной погрешности.

Для молекулы бутана (R-Н-аналог ILE), для сравнения, было произведено отключение кулоновского взаимодействия с параметрами $\lambda_{qq} = 1; 0.25; 0$, что привело к большой погрешности в конечном значении энергии гидратации. В таблице 3.4 представлены погрешности на этапах отключения кулоновского взаимодействия для этой системы. Видно, что на первом этапе погрешность сильно возросла.

3.9 Результаты

Используя вышеописанную технологию определения свободной энергии гидратации, были определены энергии гидратации для 13 R-Н-аналогов аминокислот. В таблице 3.5 приведены полученные значения и соответствующие статистические погрешности, соответствующие 90% доверительному интервалу. Экспериментальные данные по работе [15]. Следует отметить, что экспериментальные данные разных исследователей отличаются, что может быть вызвано условиями эксперимента. Средний разброс данных между разными экспериментальными работами составляет порядка 0.2 кДж/моль.

Обозначение состояния λ_{qq}	ΔF , кДж/моль	Погрешность, кДж/моль
1.0 (полное взаимодействие)		
	-5.674	0.14
0.25		
	-2.046	0.0045
0.0 (отсутствие кулоновского взаимодействия)		

Таблица 3.4: Данные по разностям свободных энергий при нестандартном отключении кулоновского взаимодействия в системе: молекула бутана в воде.

R-H-аналог	ΔF_{hydr} , кДж/моль	$\Delta F_{hydr}^{experiment}$, кДж/моль	Отклонение, кДж/моль
ALA	9.20 ± 0.04	8.12	1.0
VAL	10.1 ± 0.1	8.33	1.8
LEU	10.57 ± 0.08	9.54	1.0
ILE	10.5 ± 0.3	9.0	1.5
CYS	-2.24 ± 0.06	-5.19	2.95
MET	-1.74 ± 0.18	-6.2	4.45
PHE	-3.26 ± 0.11	-3.18	0.08
TRP	-18.7 ± 0.3	-24.6	5.9
SER	-19.2 ± 0.1	-21.18	2.0
THR	-18.65 ± 0.13	-20.4	1.8
ASN	-35.69 ± 0.10	-40.5	4.8
GLN	-36.45 ± 0.16	-39.26	2.8
TYR	-22.33 ± 0.18	-25.6	3.3

Таблица 3.5: Рассчитанные значения энергии гидратации и экспериментальные значения.

3.10 Обсуждение и сравнение

Как видно из таблицы 3.5 в настоящей работе были получены значения энергий гидратации в модели OPLS-AA/SPC с высокой точностью, соответствующей приблизительно точности экспериментального определения. Однако сравнение полученных значений и экспериментальных свидетельствует о том, что систематическая ошибка, вносимая моделью достаточно велика. Она составляет от 0 до 6 кДж/моль. Причём во всех случаях наша модель завышает реальное значение энергии гидратации. Такой тип поведения свойственен почти всем классическим моделям широко используемым на сегодняшний день. Следует отметить, что при измерении относительных энергий гидратации эта систематическая ошибка может частично гаситься.

Точность такого порядка, который был получен в работе, обычно не требуется для практических целей, так как систематическая ошибка модели намного превосходит её. Однако, способ получения надёжных и точных результатов в рамках конкретной модели является весьма полезным для дальнейшего развития и уточнения параметров силовых полей, параметризации моделей по термодинамическим характеристикам.

Наконец, такие данные важны для дальнейшего более детального изучения профилей и ландшафтов свободной энергии методами компьютерного моделирования, что обычно не возможно с помощью экспериментальных методик. В частности, полученные энергии гидратации молекул в данной модели будут использованы в дальнейшем в главе 4 при описании адсорбции молекул и построении профилей свободной энергии для них.

Проведём теперь сравнение полученных результатов с ранее проводившимися работами по компьютерному моделированию. К сожалению, сравнение результатов работы со многими работами проводившимся до этого может быть достаточно сложным из-за различных подходов к вычислительным процедурам. Основной проблемой является отказ многих исследователей от использования “корректных” процедур учёта дальнедействующих взаимодействий. Ещё одной проблемой может быть использование не совсем корректных процедур по созданию ансамблей постоянной температуры и давления.

Наиболее достоверной работой, с которой можно сравнить результаты проведённого исследования, является работа [48]. В этой работе применялся во многом похожий набор алгоритмов и методик. В работе использовалось также ядро комплекса GROMACS версии 3.1.4, что во многом облегчает сравнение. Однако, были и некоторые различия. Сравнение результатов приведено в таблице 3.6.

R-H-аналог	Данная работа ΔF_{hydr} , кДж/моль	Работа [48] ΔF_{hydr} , кДж/моль	Совпадение в пределах статистической погрешности
ALA	9.20 ± 0.04	9.08 ± 0.04	-
VAL	10.1 ± 0.1	10.08 ± 0.13	+
LEU	10.57 ± 0.08	10.46 ± 0.13	+
ILE	10.5 ± 0.3	10.38 ± 0.13	+
CYS	-2.24 ± 0.06	-2.09 ± 0.08	-
MET	-1.74 ± 0.18	-1.76 ± 0.13	++
PHE	-3.26 ± 0.11	-3.31 ± 0.17	+
TRP	-18.7 ± 0.3	-19.7 ± 0.2	-
SER	-19.2 ± 0.1	-19.33 ± 0.13	+
THR	-18.65 ± 0.13	-18.54 ± 0.17	+
ASN	-35.69 ± 0.10	-35.70 ± 0.17	++
GLN	-36.45 ± 0.16	-35.95 ± 0.17	-
TYR	-22.33 ± 0.18	-22.56 ± 0.17	+

Таблица 3.6: Сравнение результатов данной работы и работы [48]

Как следует из сравнения, не все результаты совпадают в пределах статистической погрешности. Погрешность выборки также маловероятна, так как несовпадения идут в основном на молекулах не обладающих “медленными” степенями свободы. Особенно важным является различие результатов для самой простой молекулы - метана.

Различия следует искать в вычислительных процедурах. Электростатическим вкладом в энергию гидратации метана можно пренебречь. Вклад взаимодействий Ван-дер-Ваальса

составляет 9.21 ± 0.03 кДж/моль по нашим результатам и 9.081 ± 0.04 по результатам работы Shirts et al.[48].

В этих двух работах использовались несколько различные способы учёта взаимодействий Леннарда-Джонса. Радиусы обрезания этого взаимодействия в работе Shirts et al. составляли 0.7-0.75 нм с использованием свитч функции. Во время моделирования не применялся алгоритм коррекции дисперсионного взаимодействия. А эта поправка вычислялась аналитически уже после определения энергии гидратации и добавлялась к ней. Так, поправка для молекулы метана составляла -1.64 кДж/моль. Тем не менее, такая процедура давала результаты, хорошо согласующиеся с результатами, полученными при использовании больших радиусов обрезания. В нашей работе использовались радиусы обрезания 1-1.2 нм, однако, поправка (коррекция дисперсионного взаимодействия) применялась не только к энергии системы, что не оказывает никакого эффекта на динамику системы, но и к вириалу сил, используемому для вычисления давления в системе.

Было сделано предположение, что разница в результатах происходит именно из-за дополнительного учёта поправки к вириалу сил. Было произведено определение энергии гидратации метана без учёта поправки к вириалу сил, но с учётом поправки к энергии системы. В результате было получено значение Ван-дер-ваальсовой составляющей энергии гидратации 9.086 ± 0.03 , что с точностью до третьего знака совпадает со значением полученным Shirts et al. Учёт “дисперсионной поправки” к вириалу сил увеличивает давление в системе на величину порядка 100 атмосфер.

Можно утверждать, что “дисперсионная поправка” увеличивающая давление в системе влияет на величину энергии гидратации и изменяет её для молекулы метана где-то на 0.12 кДж/моль.

Возможно, несостыковки для молекул аналогов цистеина, триптофана и глутамина также связаны с этой поправкой. Однако, для того, чтобы утверждать это наверняка, необходимы дополнительные исследования.

3.11 Выводы

В данной работе была изучена и применена методика измерения свободной энергии гидратации небольших молекул с помощью компьютерного эксперимента. Методика предполагает измерение свободной энергии в рамках классической модели молекул (некоторого силового поля) с высокой точностью. Особое значение придавалось эффективности (скорости расчёта) методик, однако, при этом акцент делался на использование методик, не вносящих систематические ошибки в результат измерения. Были проанализированы различные методы оценки свободной энергии; показано, что для оценки свободной энергии гидратации наиболее эффективным является метод Беннета. В работе обсуждались алгоритмы учёта межмолекулярных взаимодействий и их вклад в систематическую ошибку измерения. Показано, что использование радиуса обрезания при учёте взаимодействий Ван-дер-Ваальса даёт систематическую ошибку, которую можно разбить на два вклада. Первый появляется из-за недоучёта энергии взаимодействия молекул, второй из-за недооценки давления в системе. Неучёт второй поправки в большинстве исследований предыдущих авторов может отвечать за небольшое расхождение результатов с данными нашей работы для некоторых молекул.

В работе была произведена оценка энергии гидратации для веществ-аналогов боковых

цепей аминокислот в модели силовых полей OPLS-AA/SPC. Подтверждено наличие систематической ошибки данной модели, приводящей к переоценке энергии гидратации по сравнению с реальной. То есть модель завышает вероятность нахождения молекул в воде. Ошибка в определении энергии гидратации составляет от 0 до 4.5 кДж/моль.

Глава 4

Исследование адсорбции боковых цепей аминокислот

4.1 Аннотация

В работе проведена оценка адсорбционных свойств пяти R-Н-аналогов аминокислотных остатков (TYR, THR, ALA, SER, TRP) методами компьютерного моделирования. Использовался классический метод стохастической динамики с силовым полем OPLS-AA/SPC. Для всех молекул был построен профиль свободной энергии в системе с разделом фаз вода/вакуум. Обсуждаются методы оценки свободной энергии адсорбции. Проводится сравнение результатов для различных молекул.

4.2 Введение

Наряду со сродством к объёмным фазам, сродство молекул к поверхности является зачастую важнейшим элементом в функционировании молекулярных систем. Как было показано в разделе “Обзор литературы”, экспериментальным параметром, отражающим сродство молекул к поверхности раздела является свободная энергия адсорбции. Как и в случае гидратации, интересным представляется попытка использования компьютерного моделирования для определения энергии адсорбции. Кроме того, с помощью компьютерного моделирования можно получить и намного более детальную характеристику - профиль свободной энергии, или потенциал средней силы (ПСС). Анализ профиля свободной энергии позволяет судить, как о термодинамических, так и о кинетических характеристиках адсорбции.

Построение потенциалов средней силы получило широкое применение в компьютерном моделировании. Существуют работы по определению ПСС для молекул в двухфазных системах. В литературе пока нами не было найдено работ, где проводилось бы рассмотрение профилей свободной энергии в двухфазных системах для некоторого набора R-Н-аналогов аминокислот. В этой связи попытка такого исследования является достаточно интересной.

4.3 Методы

Способы определения ПСС были изложены в разделе 2.8. В данной работе использовался метод “сдерживающих сил”. Моделируемая система представлена на Рис. 4.1. Толщина

водного слоя составляла приблизительно 4 нм, однако, молекулы находящиеся у левого края в слое толщиной около 0.9 нм были заморожены. То есть их координаты не обновлялись в ходе расчёта. Это было сделано для того, чтобы создать своего рода ледяную подложку, непроницаемую для молекул воды, находящихся в жидкой фазе. Для создания выборки использовался метод стохастической динамики с параметрами аналогичными тем, которые приведены в разделе 3. Единственным отличием было отсутствие баростатирования, то есть моделирование проводилось в NVT ансамбле.

Следует отметить, что поправка к энергии, призванная учесть “хвост” взаимодействия

Рис. 4.1: Вид моделируемой системы для вычисления профиля свободной энергии молекулы в системе вода/вакуум. Стрелкой показано направление оси для построение профиля.

Леннард-Джонса, в данном случае никакого эффекта не оказывала. Так как она влияет лишь на измеряемое значение энергии, но не на силу взаимодействия.

Для фиксации молекулы налагалось ограничение на абсолютную X-координату центра масс молекулы. Для реализации ограничения использовался алгоритм SHAKE. На каждом шаге измерялась проекция силы на ось X, действующая на центр масс молекулы.

После подготовки системы и её релаксации сбор данных проводился в течение 1-2 нс. Характерная зависимость автокорреляционной функции случайной силы от времени представлена на Рис. 4.2. Время релаксации для силы составляет порядка 0.01 пс. То есть за время расчёта получалось порядка 100 000 независимых значений силы. График распределения для силы представлен на Рис. 4.3. Вид распределения соответствует общепринятым представлениям о гауссовом распределении случайной силы. Стандартное отклонение составляет порядка нескольких сотен, а средняя сила всего лишь от единиц до десятков кДж/моль/нм. Поэтому для достижения удовлетворительной точности в определении среднего необходимо проводить усреднение по большой выборке. Оценить размер выборки (и следовательно время моделирования) можно исходя из стандартной формулы для погрешности среднего (см. [39]):

$$\delta F = \sqrt{\langle F^2 \rangle} \sqrt{2\tau/T}$$

где δF -статистическая погрешность в определении среднего, $\langle F^2 \rangle$ -дисперсия случайной величины (силы), τ -время релаксации случайной величины, T -время моделирования (снятия выборки).

Исходя из этой формулы видно, что погрешность порядка единиц кДж/моль/нм, достигается, когда время моделирования на 4-5 порядков больше времени релаксации силы.

Рис. 4.2: Автокорреляционная функция силы, действующей на центр масс молекулы

Рис. 4.3: Распределение для силы, действующей на центр масс молекулы

Для каждой молекулы проводилось измерение силы в разных точках оси X. Сила интегрировалась для получения профиля свободной энергии. По профилю свободной энергии определялся избыток Гиббса. А избытку Гиббса сопоставлялась свободная энергия адсорбции в соответствии с соображениями, описанными в пункте 2.2.2.

4.4 Результаты

На рисунках 4.4, 4.5, 4.7, 4.6, 4.8 представлены полученные зависимости. Пунктирная синяя кривая соединяет полученные значения средней силы в различных точках. В каждой точке, где проводилось измерение, отложена погрешность. Нормальная красная линия - потенциал средней силы (ПСС) (он же профиль свободной энергии), получаемый путём интегрирования средней силы. Зелёным пунктиром обозначен профиль плотности воды, а чёрная пунктирная линия - поверхность Гиббса, определённая по профилю плотности воды. Также представлен рассчитанный по профилю свободной энергии профиль концентрации растворённого вещества в системе, нормированный на единичную концентрацию в газовой фазе. Слева чёрной горизонтальной линией отмечается значение энергии гидратации рассчитанное в разделе 3.

Рис. 4.4: Графики зависимости средней силы, потенциала средней силы, профиля концентрации растворённого вещества, профиля плотности от X для молекулы R-H-аналога **ALA**. Также чёрными линиями на рисунке отмечены поверхность Гиббса и значение энергии гидратации полученное в разделе 3.

4.5 Обсуждение

Рассмотрим подробнее полученные зависимости. Все вещества, рассмотренные в данной работе, кроме молекулы метана (R-H аналог аланина) обладают отрицательной энергией гидратации.

На Рис. 4.4 представлены зависимости средней силы и потенциала средней силы для молекулы метана. Следует отметить, что молекула метана обладает положительной свободной энергией гидратации и отрицательной энтальпией гидратации, так называемый, эффект гидрофобной гидратации. При этом энтальпия гидратации составляет около -11 кДж/моль

Рис. 4.5: Графики зависимости средней силы, потенциала средней силы, профиля концентрации растворённого вещества, профиля плотности от X для молекулы R-H-аналога **SER**. Также чёрными линиями на рисунке отмечены поверхность Гиббса и значение энергии гидратации полученное в разделе 3.

([23] Таблица 8). Отрицательная энтальпия гидратации возникает главным образом из-за Ван-дер-ваальсова притяжения молекулы метана и молекул воды. Это взаимодействие ответственно и за притяжение молекулы к границе раздела, что хорошо видно на графике. Более того, по форме графики средней силы и ПСС напоминают в правой части потенциал взаимодействия Леннард-Джонса. Сила взаимодействия молекулы метана и воды увеличивается по мере приближения молекулы к водному слою до тех пор, пока молекула не начинает сталкиваться с молекулами воды, что хорошо видно на Рис. 4.4. Как только профиль плотности системы становится отличным от нуля, сила притяжения начинает уменьшаться. По мере продвижения в водный слой возрастает сила отталкивания, которая, в частности, обусловлена энтропийными потерями воды. Это приводит к формированию адсорбционного минимума свободной энергии, который расположен в двух ангстремах от поверхности Гиббса. Можно предположить, что небольшой адсорбционный минимум на границе раздела для молекулы метана обусловлен балансом между энтропийными и энтальпийными составляющими свободной энергии. Когда молекула входит в приповерхностный водный слой, сила отталкивания сильно возрастает, возрастает и ошибка в определении значения средней силы. Зависимость силы от координаты становится намного менее плавной, появляются “зигзагообразные” колебания силы. Возможно, такие колебания связаны с дискретностью структуры воды в приповерхностном слое и конечными размерами системы, однако, не исключено, что они возникают из-за погрешности выборки. Затем в “объёмной” фазе сила уменьшается и в идеале должна выходить на ноль. Однако, в районе значений оси X 1.5-3 нм, средняя сила ведёт себя непредсказуемым образом. В районе 2.5 нм появляется значительная сила притяжения. С чем может быть связано такое поведение средней силы? Если посмотреть на графики для других веществ, то мы увидим, что те или иные отклонения средней силы от нулевого значения в области значений оси X 2-3 нанометра присутствуют для всех веществ. При этом плотность воды в этой области уже соответствует объёмной. Однако

Рис. 4.6: Графики зависимости средней силы, потенциала средней силы, профиля концентрации растворённого вещества, профиля плотности от X для молекулы R-H-аналога **THR**. Также чёрными линиями на рисунке отмечены поверхность Гиббса и значение энергии гидратации полученное в разделе 3.

это не означает, что эта область может считаться “объёмной” водой, так как размеры системы невелики. Кроме того, своё влияние начинает оказывать и “замороженный” слой воды, который заканчивается на значении оси X в 1 нм.

Для прояснения этого вопроса в дальнейших исследованиях предлагается отказаться от использования слоя “замороженных” молекул воды. Расчёт профиля вдоль всей толщины слоя, благодаря ожидаемой симметрии профиля, позволит сделать более аргументированные выводы о поведении профиля средней силы и ПСС на расстоянии в 1-2 нанометра от границы раздела.

В пункте 2.7 говорилось о том, что при моделировании границ раздела, существенное влияние на толщину границы оказывает размер системы. И в компьютерном моделировании толщина границы раздела оказывается заниженной. Поэтому к профилям свободной энергии стоит относиться не как к реальным профилям, а как к некоторым модельным кривым, отражающим взаимодействия молекулы со средой в различных положениях.

Рассмотрим теперь полученные зависимости для веществ с отрицательной свободной энергией гидратации. Общий вид зависимости средней силы от координаты оказывается приблизительно одинаков для всех исследованных молекул (как гидрофобных, так и гидрофильных). Справа от поверхности Гиббса на графиках молекулы притягиваются к границе раздела. Максимальная сила притяжения действует на молекулу приблизительно на расстоянии 0.5 нм от поверхности Гиббса. По мере продвижения в объёмную фазу сила притяжения сменяется силой отталкивания, что проявляется в виде появления минимума на профиле свободной энергии. Максимум силы отталкивания находится где-то на расстоянии 0.3 нм от поверхности Гиббса в “водной” фазе. Далее по мере углубления в объёмную фазу сила

Рис. 4.7: Графики зависимости средней силы, потенциала средней силы, профиля концентрации растворённого вещества, профиля плотности от X для молекулы R-Н-аналога **TYR**. Также чёрными линиями на рисунке отмечены поверхность Гиббса и значение энергии гидратации полученное в разделе 3.

отталкивания обычно значительно уменьшается и сходит на нет в объёмной фазе. Однако, характер зависимости средней силы, относительные размеры разных частей зависимости приводят к достаточно различным профилям свободной энергии.

Все 4 гидрофильных соединения (метанол/SER, этанол/THR, п-крезол/TYR и 3-метилиндола/TRP) обладают по сравнению с молекулой метана значительно более глубоким и широким минимумом средней силы вблизи интерфейса в вакуумной фазе. Такое притяжение, является проявлением сильного и дальнodelствующего характера электростатического взаимодействия. Можно предположить, что весомый вклад в притяжение даёт взаимодействие с приповерхностным слоем воды, который, как известно, обладает дополнительным дипольным моментом. Когда молекулы пересекают поверхность Гиббса, они начинают испытывать на себе отталкивание. Сила отталкивания и ширина слоя, в котором молекулы испытывают отталкивание, варьируются и коррелируют с размером молекулы. Для относительно больших ароматических молекул сила отталкивания достигает 30 кДж/моль/нм, а ширина слоя - около 6 ангстрем. У этанола ширина слоя остаётся приблизительно такой же, но максимальная сила уменьшается до 20 кДж/моль/нм. У метанола максимальная сила уменьшается до 10 кДж/моль/нм, что приводит к тому, что на профиле средней силы метанола нет выраженного адсорбционного минимума.

На основании анализа профилей свободной энергии могут быть сделаны некоторые выводы о кинетике поглощения вещества из газовой фазы в водную и обратно. Так для молекулы аналога триптофана (TRP Рис. 4.8) виден максимум профиля вблизи поверхности в объёмной фазе. Величина максимума по сравнению с энергией гидратации - порядка 0.5 кТ. Это максимум называется в литературе [50] барьером десольватации.

Как видно из приведённых графиков, данная модель на качественном уровне неплохо

Рис. 4.8: Графики зависимости средней силы, потенциала средней силы, профиля концентрации растворённого вещества, профиля плотности от X для молекулы R-Н-аналога **TRP**. Также чёрными линиями на рисунке отмечены поверхность Гиббса и значение энергии гидратации полученное в разделе 3.

описывает профиль плотности свободной энергии. В вакууме и вблизи границы раздела ошибка в определении средней силы невелика, однако, когда молекула заходит в объёмную фазу воды ошибка сильно возрастает, и, кроме того, зависимость средней силы от координаты становится менее плавной. Весьма вероятно, что в этой области сказываются конечные размеры системы и влияние границ на внутреннюю структуру воды. Следует отметить, что, строго говоря, слой толщиной в 3 нм не может считаться объёмной фазой. Тем не менее, по порядку величины предсказания энергии гидратации совпадают со значениями, измеренными в разделе 3. Можно сделать вывод, что методика, использовавшаяся в предыдущем исследовании, определяет энергию гидратации намного более эффективно и точно.

В то же время для получения адсорбционных характеристик в случае сильно адсорбирующихся веществ для нас важен лишь минимум свободной энергии вблизи границы раздела. Поскольку этот минимум обычно глубже значений энергии в вакууме и в объёмной воде, то график зависимости концентрации имеет ярко выраженный максимум. А для определения избытка Гиббса нам важно точно знать лишь форму и положение этого максимума. Поэтому неточности в определении профиля энергии в объёмной фазе мало сказываются на определении избытка Гиббса и энергии адсорбции. Качественный максимум концентрации отсутствует лишь для молекулы метанола (Рис. 4.5), что делает количественное определение энергии адсорбции несколько сложным. Однако, в случае метанола энергия гидратации получаемая из потенциала средней силы оказывается несколько заниженной. То есть энергия молекулы вблизи интерфейса, оцененная по профилю потенциала средней силы, оказывается на несколько кТ меньше реальной энергии гидратации в данной модели.

4.6 Энергия адсорбции

Для определения энергии адсорбции относительно вакуумной фазы сначала определялся избыток Гиббса, и ему сопоставлялось значение энергии адсорбции по формуле Лэнгмюра 2.3. Далее вычислялась энергия адсорбции относительно фазы максимального сродства, для этого при необходимости использовались уже полученные в разделе 3 значения энергии гидратации. В таблице 4.1 приведены значения полученной энергии для пяти молекул. Также приведены экспериментальные данные работы [12] или работы [56] по оценке энергии адсорбции на **границу раздела вода/гексан**. Ещё раз отметим, что результат для SER является крайне неточным из-за слабовыраженного минимума свободной энергии вблизи интерфейса. В таблице приведены именно экспериментальные результаты для адсорбции

Молекула	F_{ads}^{vacuum} , кДж/моль	$F_{ads}^{max-affin}$, кДж/моль	Эксперимент вода/гексан, F_{ads} , кДж/моль
ALA	-1.5	1.5	-0.25
SER	-21.3	2	—
TYR	-29.5	7.2	5.5
TRP	-28.0	9.3	14.5
THR	-21.9	3.3	—

Таблица 4.1: Энергии адсорбции, рассчитанные методом компьютерного моделирования, и их экспериментальные значения

на границу раздела вода/гексан, тогда как в моделировании измерялась величина адсорбции на границу раздела вода/вакуум (вода/газ). Качественно результаты совпадают. Экспериментальных данных по адсорбции исследуемых веществ на границу вода/газ в литературе нами найдено не было.

4.7 Выводы

В данной работе была предложена методика оценки свободной энергии адсорбции на поверхность воды для некоторых небольших молекул. Для молекул с хорошей адсорбцией метод представляет собой весьма точный способ определения энергии адсорбции в рамках заданной модели. Комбинация данного метода и метода определения энергии гидратации позволяет получить для молекулы два важных параметра - энергию гидратации и энергию адсорбции. Получаемый профиль свободной энергии даёт более детальную информацию о распределении молекул вблизи границы раздела, позволяет оценить положение и ширину адсорбционной зоны.

В данной работе были построены профили свободной энергии для пяти молекул, R-Н-аналогов аминокислот ALA, SER, TYR, TRP, THR. Для этих молекул проведена оценка энергии адсорбции, проанализированы профили свободной энергии. Полученные значения коррелируют с экспериментальными данными по адсорбции молекул на границу раздела вода/гексан.

Глава 5

Выводы

По результатам работы можно сделать следующие основные выводы.

Методами компьютерного моделирования:

- Реализована методика оценки свободной энергии небольших молекул гидратации с высокой статистической точностью
- Предложен метод оценки энергий адсорбции небольших молекул.
- Произведена оценка энергии гидратации для 13 боковых цепей аминокислот с высокой точностью
- Для 5 веществ построены профили свободной энергии на границе раздела с водой.
- Произведена оценка энергий адсорбции 5 молекул.
- Предложенные модели и методы хорошо описывают адсорбционные и гидратационные свойства молекул на качественном уровне.
- Количественные оценки значений энергий гидратации для некоторых веществ с хорошей точностью соответствуют экспериментальным значениям.

Заключение

В заключении хотелось бы выразить благодарность моим научным руководителям за постановку интересной задачи, обучение и руководство в течение трёх лет, сотрудникам кафедры биоинженерии Биологического факультета за советы и предоставленные вычислительные возможности, Николаю Кирилловичу Балабаеву за лекции, советы и любезно предоставленный программный пакет ПУМА, позволившие освоить метод молекулярной динамики, семье за поддержку, друзьям за многочисленные обсуждения и создание творческой атмосферы.

Приложение А

Молекулы-объекты исследования

В таблице А.1 приведены вещества исследованные в данной работе и соответствующие им аминокислоты. Из 20 встречающихся в природе аминокислот исследовались лишь 13 аминокислот. Не исследовались аминокислоты, которые при стандартном рН обладают заряженными группами из-за трудности учёта электростатических взаимодействий при моделировании в этом случае. Не рассмотрена также аминокислота пролин из-за её нестандартного строения и невозможности описанным образом разделить её на остов и боковую цепь.

Название	Обозначение	Формула	R-Н-аналог	Название R-Н-аналога
Аланин	ALA			Метан
Валин	VAL			Пропан
Лейцин	LEU			изобутан
Изолейцин	ILE			н-бутан
Цистеин	CYS			Этилмеркаптан
Метионин	MET			Этилметилсульфид
Фенилаланин	PHE			Толуол
Триптофан	TRP			3-метилиндол
Серин	SER			Метанол
Треонин	THR			Этанол
Аспарагин	ASN			Ацетамид
Глутамин	GLN			Пропионамид
Тирозин	TYR			пара-метилфенол (п-крезол)

Таблица А.1: Аминокислоты и их R-Н-аналоги, исследованные в данной работе

Приложение В

Вывод формулы для расчёта СЭС в компьютерном моделировании

В.1 Введение

Классическая теория жидкостей и растворов является достаточно хорошо разработанной областью статистической физики (см. например, [14]). Развёрнутое описание процессов сольватации различных молекул с точки зрения термодинамики приведено в работе [23]. В этой же работе приводится соотношение между свободной энергией сольватации и энергией взаимодействия молекулы с окружающим растворителем. Вывод формул пригодных для вычисления энергии сольватации в компьютерном эксперименте (при использовании изобаро-изотермического ансамбля) приводится в приложении к работе [27].

В данном разделе приводится развёрнутый вывод конечной расчётной формулы В.33. При этом, отличие от работы [27] состоит в использовании при выводе постулатов только классической статистической физики, так как, по мнению автора, использование “квантовых” подходов несколько усложняет понимание вывода, и вводит неискущённого читателя в заблуждение относительно границ применимости квантовых и классических моделей.

В.2 Свободная энергия сольватации

Поясним ещё раз, что же такое свободная энергия сольватации. В большинстве экспериментальных работ свободную энергию сольватации молекулярного вещества S в жидкости В измеряют следующим образом. Допустим, у нас есть двухфазная система, состоящая из жидкой фазы (вещество В) и газообразной фазы. Кроме того, в системе присутствует вещество S в малой концентрации. Тогда, измеряя концентрации вещества S в жидкой и газообразной фазах (или парциальное давление вещества S в газообразной фазе), говорят, что величина (В.1) является свободной энергией сольватации вещества S в жидкости В [27].

$$\Delta G_{solv} = kT \ln \frac{P_S}{c_s^B kT} \quad (\text{В.1})$$

В литературе [23] величина ΔG_{solv} обычно интерпретируется как свободная энергия, связанная с взаимодействием (coupling) молекулы растворённого вещества с остальной частью раствора. Из следующих параграфов станет ясно, почему используется такая интерпретация.

В.3 Постулаты классической статфизики

Проблема, излагаемая в данной работе, напрямую связана с определением термодинамического параметра (свободной энергии сольватации) из микроскопических параметров наших молекул. Получением соотношений между микроскопическими параметрами системы и её термодинамическими характеристиками занимается статистическая физика. Мы будем широко использовать связь между статистической суммой и свободной энергией (В.2-В.3).

$$F = -kT \ln \frac{Z_V}{N!} \quad (\text{В.2})$$

$$Z_V = \int \exp \left\{ -\frac{H(\vec{q}, \vec{p})}{kT} \right\} d\vec{q} d\vec{p} \quad (\text{В.3})$$

Для нашего дальнейшего рассмотрения необходимо понимать в какой момент и откуда берётся множитель $N!$ в выражении (В.2).

В.4 Связь статсуммы со свободной энергией

В.4.1 Квантовый подход

Изложение статфизики можно изначально строить из квантовых предположений, вводя энтропию через логарифм вырожденности энергетического уровня с энергией E . Тогда статсумма запишется как (В.4).

$$Z_V^Q = \text{Tr} \exp(-\frac{\hat{H}}{kT}) \quad (\text{В.4})$$

Тогда при переходе к классическому пределу по квазиклассическим формулам с некоторыми приближениями (см. например, [29]) можно получить (В.5), где d -размерность пространства.

$$Z_Q^V \approx \frac{1}{h^{dN}} \int d\vec{p} d\vec{q} \exp \left\{ -\frac{H(\vec{p}, \vec{q})}{kT} \right\} \quad (\text{В.5})$$

Однако, в формуле (В.5) пока нет множителя $1/N!$. Появление этого множителя объясняется в литературе из принципа неразличимости частиц в квантовой механике. Так для ферми- или бозе- газа правильное выражение для квантового состояния должно быть записано как полная симметризация или антисимметризация по всем перестановкам, например, для бозе-газа (В.6).

$$\begin{aligned} |p_1 \dots p_N \rangle_s &= \frac{1}{N!} \sum_{\sigma} |p_{\sigma_1} \dots p_{\sigma_N} \rangle \\ |q_1 \dots q_N \rangle_s &= \frac{1}{N!} \sum_{\sigma} |q_{\sigma_1} \dots q_{\sigma_N} \rangle \end{aligned} \quad (\text{В.6})$$

Можно показать, что множитель $1/N!$ появившийся в формуле (В.6) перейдёт в квазиклассическом приближении в выражение для статсуммы, и мы получим (В.7)

$$Z_Q^V \approx \frac{1}{h^{dN} N!} \int d\vec{p} d\vec{q} \exp \left\{ -\frac{H(\vec{p}, \vec{q})}{kT} \right\} \quad (\text{В.7})$$

Итак, смысл множителя $1/N!$ связан с учётом неразличимости частиц и отражает количество способов, которыми можно переставить одинаковые частицы. Обобщением формулы (В.7) на случай, когда в системе присутствуют два сорта частиц будет формула, где $N!$ заменено на соответствующее количество перестановок, то есть $N_1! * N_2!$.

В.4.2 Подход Гиббса

Приведённые выше рассуждения, как считается в литературе, позволяют объяснить, так называемый, *парадокс Гиббса* [57]. Суть парадокса состоит в том, что энтропия смешения двух “идеальных” газов не зависит от природы этих газов, однако меняется скачком, если мы считаем газы неразличимыми. В формуле (В.7) это будет соответствовать замене $N_1! * N_2!$ на $(N_1 + N_2)!$.

В ходе дальнейшего рассмотрения методики компьютерного эксперимента мы столкнёмся с необходимостью чётко представлять себе, в каком случае нужно использовать множитель $N_1! * N_2!$, а в каком $-(N_1 + N_2)!$. Для целей данной работы, по мнению автора, квантовый подход к разрешению парадокса Гиббса не совсем прозрачен по двум причинам:

- Во-первых, мы изначально оговорили, что будем рассматривать систему как классическую. И с этой точки зрения объяснение парадокса Гиббса через квантовую механику представляет собой некоторый трюк.
- И во-вторых, в дальнейшем мы столкнёмся с необходимостью плавно изменять Гамильтониан взаимодействия одной частицы со всей остальной системой. И, разумеется, в классических вычислениях не будет никакой разницы между динамикой системы с изначальным Гамильтонианом и Гамильтонианом, отличающимся от исходного на бесконечно малую величину. Поэтому нам потребуется ясное понимание происходящих процессов, или хотя бы универсальных алгоритмов, в соответствии с которыми нужно действовать.

Замечательное рассмотрение вопросов связанных с парадоксом Гиббса приведено в работе [58]. Объяснение, излагаемое в этой работе, по-видимому, соответствует оригинальному взгляду Гиббса. Для нашего дальнейшего рассмотрения мы рассмотрим выводы работы [58].

Итак, прежде чем заниматься описанием системы с точки зрения статфизики мы должны определиться с набором макропараметров и не менять этот набор по ходу рассмотрения. Набор макропараметров зависит от тех экспериментальных фактов и возможностей, которые нам необходимо описать. Допустим, у нас есть смесь двух различных газов, однако, наши экспериментальные методики не позволяют нам их различить. Тогда на некоторых примерах можно показать, что экспериментальные результаты будут согласовываться с теорией в двух случаях:

- если мы честно записываем свободную энергию системы как сумму энергий двух различных газов

$$F/kT = N_1 \ln \frac{N_1}{V} + N_2 \ln \frac{N_2}{V}$$

- или если мы считаем газы неразличимыми и пишем

$$F/kT = (N_1 + N_2) \ln \frac{(N_1 + N_2)}{V}$$

Однако, если в нашем эксперименте существуют и используются методики, которые могут различить два газа, то описать экспериментальные результаты можно будет только с помощью первого выражения.

Хотя в [58] это показывается только при конкретном рассмотрении одного из возможных экспериментов, авторы данной работы предполагают, что может быть сделано следующее общее утверждение:

Если в рамках модели не существенно различие между частицами, одним макропараметром можно описать характеристики сразу нескольких типов частиц, то можно использовать выражение для статсуммы одинаковых частиц, если же мы ожидаем, что различие между частицами проявится в эксперименте, мы соотносим им разные макропараметры, то при теоретическом рассмотрении необходимо делить статсумму на произведение факториалов.

В.5 Сольватация, рассмотрение статфизики

Рассмотрим теперь эксперимент по определению свободной энергии сольватации, описанный в пункте В.2, с точки зрения классической статфизики. Рассмотрим газ вещества S с парциальным давлением P_s . Свободная энергия Гиббса системы задаётся выражением (В.8)

$$G(N_s^g, P_s, T) = -kT \ln Z^P(N_s^g, P_s, T) \quad (\text{В.8})$$

где статсумма Z^P определяется как

$$Z^P(N_s^g, P_s, T) = \int_0^\infty dV e^{(-\frac{P_s V}{kT})} Z^V(N_s^g, V, T) \quad (\text{В.9})$$

Наша задача состоит в том, чтобы найти равновесные концентрации вещества S в двух фазах. Равновесие будет достигнуто, когда химпотенциалы вещества S в обеих фазах сравняются.

$$\mu = \frac{\partial G}{\partial N} \quad (\text{В.10})$$

$$\mu_s^{gas} = \mu_s^{liquid} \quad (\text{В.11})$$

Займемся нахождением химпотенциалов.

Газовая фаза

Предположим, что в газовой фазе можно пренебречь взаимодействием молекул между собой, тогда Гамильтониан системы примет вид

$$H(\vec{p}, \vec{q}) = \sum_i^{N_s} H_s$$

, где H_s -гамильтониан каждой молекулы, может быть представлен в виде движение центра масс молекулы и внутреннего Гамильтониана взаимодействий:

$$H_s(p_1, \dots, p_n, q_1, \dots, q_n) = \frac{p_{mc}^2}{2m} + H_{int}(\tilde{p}_1, \dots, \tilde{p}_{n-1}, \tilde{q}_1, \dots, \tilde{q}_{n-1})$$

Тогда

$$Z^V(N_s^g, V, T) = \frac{1}{h^{dN_s^g} N_s^g!} (\Lambda * V * q_s(T))^{N_s^g}$$

где

$$q_s(T) = \int \exp\left(-\frac{H_{int}(\tilde{p}_1 \dots \tilde{p}_{n-1}, \tilde{q}_1 \dots \tilde{q}_{n-1})}{kT}\right) d\tilde{p}_1 \dots d\tilde{p}_{n-1} d\tilde{q}_1 \dots d\tilde{q}_{n-1}$$

$$\Lambda = (2\pi M k T)^{3/2}$$

Подставляя эти выражения в (B.8) получим

$$G(N_s^g, P_s, T) = -kT N_s^g \ln \left\{ \frac{\Lambda q_s(T) kT}{h^{dN_s^g} P_s} \right\} \quad (\text{B.12})$$

Найдём также выражение для химпотенциала

$$\mu(N_s^g, P_s, T) = G(N_s^g, P_s, T) - G(N_s^g - 1, P_s, T)$$

$$\mu_s^g(N_s^g, P_s, T) = -kT \ln \left\{ \frac{\Lambda q_s(T) kT}{h^d * P_s} \right\} \quad (\text{B.13})$$

Жидкая фаза

Найдём теперь химпотенциал вещества S в жидкой фазе B.

$$G(N_s^l, N_b, P, T) = -kT \ln Z^P(N_s^l, N_b, P, T)$$

$$\mu_a^l(N_s^l, N_b, P, T) = G(N_s^l, N_b, P, T) - G(N_s^l - 1, N_b, P, T) \quad (\text{B.14})$$

теперь привлечём соображения пункта B.4.2. Поскольку в нашем случае молекулы вещества S и B разные даже в экспериментальном смысле, веществу S соответствует отдельный макропараметр - концентрация вещества S, то для теоретического рассмотрения мы должны использовать следующее выражение для статсуммы (B.15)

$$Z^P(N_s^l, N_b, P, T) = \frac{1}{h^{d(N_s^l + N_b)} N_s^l! N_b!} \int d\vec{p} d\vec{q} dV \exp \left\{ -\frac{H(\vec{p}, \vec{q}) + PV}{kT} \right\} \quad (\text{B.15})$$

Тогда обозначив интеграл в выражении (B.15) за $\Delta(N_s^l, N_b, P, T)$ получим (B.16)

$$Q(N_s^l, N_b, V, T) = \int d\vec{p} d\vec{q} \exp \left\{ -\frac{H(\vec{p}, \vec{q})}{kT} \right\}$$

$$\Delta(N_s^l, N_b, P, T) = \int dV \exp \left\{ -\frac{PV}{kT} \right\} * Q(N_s^l, N_b, V, T)$$

$$\mu_s^l(N_s^l, N_b, P, T) = -kT \ln \left\{ \frac{\Delta(N_s^l, N_b, P, T)}{N_s^l * h^d * \Delta(N_s^l - 1, N_b, P, T)} \right\} \quad (\text{B.16})$$

Расчёт (B.16) свёлся у нас лишь к расчёту интегралов типа $\Delta(N_s^l, N_b, P, T)$. Предположим, что Гамильтониан нашей системы (жидкость B + вещество S) можно записать в следующем виде (B.17) [27].

$$H(N_s^l, N_b) = T_s(N_s^l) + T_b(N_b) + U_s(N_s^l) + U_b(N_b) + U_{b-b}(N_b) + U_{s-s}(N_s^l) + U_{s-b}(N_s^l, N_b) \quad (\text{B.17})$$

Для простоты рассмотрим случай парных взаимодействий:

$$H(N_s^l, N_b) = \sum^{N_s^l} H_s + \sum^{N_b} H_b + \sum^{N_s^l} \sum^{N_s^l} U_{s-s} + \sum^{N_b} \sum^{N_b} U_{b-b} + \sum^{N_b} \sum^{N_s^l} U_{s-b} \quad (\text{B.18})$$

Выделим в Гамильтониане одну из молекул растворённого вещества:

$$H(N_a^l, N_b) = \sum^{N_a^l-1} H_a + \sum^{N_b} H_b + \sum^{N_a^l-1} \sum^{N_a^l-1} U_{a-a} + \sum^{N_b} \sum^{N_b} U_{b-b} + \sum^{N_b} \sum^{N_a^l-1} U_{a-b} + \left\{ H_a + \sum^{N_a^l-1} U_{a-a} + \sum^{N_b} U_{a-b} \right\} \quad (\text{B.19})$$

Далее сделаем следующий трюк: пусть у нас есть $H(N_s^l, N_b, \lambda)$, обладающий следующим свойством:

$$\begin{aligned} H(N_s^l, N_b, \lambda = 0) &= \sum^{N_s^l-1} H_s + \sum^{N_b} H_b + \sum^{N_s^l-1} \sum^{N_s^l-1} U_{s-s} + \sum^{N_b} \sum^{N_b} U_{b-b} + \sum^{N_b} \sum^{N_s^l-1} U_{s-b} + \{H_s\} \\ H(N_s^l, N_b, \lambda = 1) &= H(N_s^l, N_b) \end{aligned} \quad (\text{B.20})$$

Простейший вариант такого Гамильтониана:

$$H(N_s^l, N_b) = \sum^{N_s^l-1} H_s + \sum^{N_b} H_b + \sum^{N_s^l-1} \sum^{N_s^l-1} U_{s-s} + \sum^{N_b} \sum^{N_b} U_{b-b} + \sum^{N_b} \sum^{N_s^l-1} U_{s-b} + H_s + \lambda \left\{ \sum^{N_s^l-1} U_{s-s} + \sum^{N_b} U_{s-b} \right\} \quad (\text{B.21})$$

Сопоставим каждому Гамильтониану соответствующий интеграл Q . Заметим, что из-за разделения переменных при интегрировании:

$$\begin{aligned} Q(N_s^l, N_b, V, T, \lambda = 1) &= Q(N_s^l, N_b, V, T) \\ Q(N_s^l, N_b, V, T, \lambda = 0) &= Q(N_s^l - 1, N_b, V, T) * Q(N_s^l = 1, N_b = 0, V, T) \end{aligned} \quad (\text{B.22})$$

Теперь сопоставим им интегралы типа Δ

$$\begin{aligned} \Delta(N_s^l, N_b, P, T, \lambda = 1) &= \Delta(N_s^l, N_b, P, T) \\ \Delta(N_s^l, N_b, P, T, \lambda = 0) &= \int dV \exp(-\frac{PV}{kT}) Q(N_s^l - 1, N_b, V, T) * Q(N_s^l = 1, N_b = 0, V, T) \end{aligned} \quad (\text{B.23})$$

Рассмотрим подробнее второе выражение в (B.24), его можно преобразовать к следующему

виду (B.27)

$$\begin{aligned}\Delta(N_s^l, N_b, P, T, \lambda = 0) &= \int_0^\infty dV \exp(-\frac{PV}{kT}) Q(N_s^l - 1, N_b, V, T) * Q(N_s^l = 1, N_b = 0, V, T) \\ &= \int_0^\infty dV \exp(-\frac{PV}{kT}) Q(N_s^l - 1, N_b, V, T) * (\Lambda * V * q_s(T))\end{aligned}\quad (B.24)$$

$$= \Lambda q_s(T) \Delta(N_s^l - 1, N_b, P, T) \frac{\int_0^\infty dV V \exp(-\frac{PV}{kT}) Q(N_s^l - 1, N_b, V, T)}{\Delta(N_s^l - 1, N_b, P, T)} \quad (B.25)$$

$$= \Delta(N_s^l - 1, N_b, P, T) * \Lambda q_s(T) V^* \quad (B.26)$$

$$= \Delta(N_s^l - 1, N_b, P, T) * Q(N_s = 1, V^*, T) \quad (B.27)$$

где V^* обозначен последний множитель, по сути это средний объём системы состоящей из $N_s^l - 1$ молекул вещества А и N_b вещества В при давлении Р.

Итак, проведя данные преобразования, мы можем записать химпотенциал в следующем виде (B.30)

$$\mu_s^l(N_s^l, N_b, P, T) = -kT \ln \left\{ \frac{\Delta(N_s^l, N_b, P, T)}{N_s^l * h^d * \Delta(N_s^l - 1, N_b, P, T)} \right\} \quad (B.28)$$

$$= -kT \ln \left\{ \frac{1}{N_s^l * h^d} \frac{\Delta(N_s^l, N_b, P, T, \lambda = 1)}{\Delta(N_s^l, N_b, P, T, \lambda = 0)} \frac{\Delta(N_s^l, N_b, P, T, \lambda = 0)}{\Delta(N_s^l - 1, N_b, P, T)} \right\} \quad (B.29)$$

$$= -kT \ln \frac{1}{N_s^l * h^d} + \Delta G_{sim} - kT \ln \Lambda q_s(T) V^* \quad (B.30)$$

где ΔG_{sim} определено (B.31)

$$\Delta G_{sim} = -kT \ln \frac{\Delta(N_s^l, N_b, P, T, \lambda = 1)}{\Delta(N_s^l, N_b, P, T, \lambda = 0)} \quad (B.31)$$

Равновесие

Итак, в равновесии у нас должны быть равны химпотенциалы вещества А в обоих фазах, приравнявая (B.13) и (B.30) получаем:

$$-kT \ln \left\{ \frac{\Lambda q_s(T) kT}{h^d * P_s} \right\} = -kT \ln \frac{1}{N_s^l * h^d} + \Delta G_{sim} - kT \ln \Lambda q_s(T) V^* \quad (B.32)$$

Произведём упрощение выражения:

$$-kT \ln \left\{ \frac{kT}{P_s} \right\} = -kT \ln \frac{1}{N_s^l} + \Delta G_{sim} - kT \ln V^*$$

Пусть V_l -объём системы, состоящей из N_s^l молекул вещества А и N_b вещества В, тогда

$$-kT \ln \left\{ \frac{kT N_s^l}{P_s V_l} \right\} = \Delta G_{sim} - kT \ln \frac{V^*}{V_l}$$

И, вспоминая (B.1), получаем

$$-kT \ln \left\{ \frac{kT c_s^B}{P_s} \right\} = \Delta G_{sim} - kT \ln \frac{V^*}{V_l} = \Delta G_{solv} \quad (B.33)$$

Итак, мы свели задачу расчёта свободной энергии сольватации к расчёту следующего отношения интегралов

$$\Delta G_{sim} = -kT \ln \frac{\int dV d\vec{p} d\vec{q} \exp \left\{ -\frac{H(\vec{p}, \vec{q}, \lambda=1) + PV}{kT} \right\}}{\int dV d\vec{p} d\vec{q} \exp \left\{ -\frac{H(\vec{p}, \vec{q}, \lambda=0) + PV}{kT} \right\}} \quad (B.34)$$

По сути выражение B.34 определяет разницу свободной энергии двух систем с *одинаковым числом частиц*, но разными параметрами взаимодействия. Это изменение свободной энергии уже может быть рассчитано в компьютерном эксперименте.

Литература

- [1] H. Vihola, A. Laukkanen, J. Hirvonen, and H. Tenhu. Binding and release of drugs into and from thermosensitive poly(n-vinyl caprolactam) nanoparticles. *Eur J Pharm Sci*, 16(1):69, 2002.
- [2] P.M. Harrison, H.S. Chan, S.B. Prusiner, and F.E. Cohen. Thermodynamics of model prions and its implications for the problem of prion protein folding. *J. Mol. Biol.*, 289:593, 1999.
- [3] S. Srebnik, A.K. Chakraborty, and D. Bratko. Random heteropolymer adsorption on disordered multifunctional surfaces: Effect of specific intersegment interactions. *J. Chem. Phys*, 109:6415, 1998.
- [4] R.E. Anderson, V.S. Pande, and C.J. Radke. Dynamic lattice monte carlo simulation of a model protein at an oil/water interface. *J. Chem. Phys*, 112:9167, 2000.
- [5] H.R. Costantino, R. Langer, and A.M. Klibanov. Aggregation of lyophilized pharmaceutical protein, recombinant human albumin: Effect of moisture and stabilization by experiments. *Biotechnology*, 13:493, 1995.
- [6] R. Wetzel. Mutations and off-pathway aggregation of proteins. *Trends Biotechnol*, 12:193, 1994.
- [7] S.B. Prusiner and S. J. DeArmond. Prion diseases and neurodegeneration. *Annu. Rev. Neurosci*, 17:311, 1994.
- [8] K.F. Lau and K.A. Dill. A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules*, 22:3986, 1989.
- [9] I.M. Okhapkin, E.E. Makhaeva, and A.R. Khokhlov. Two-dimensional classification of amphiphilic monomers based on interfacial and partitioning properties. 1. monomers of synthetic water-soluble polymers. *Colloid Polym Sci*, 284(2):117, 2005.
- [10] V.V. Vasilevskaya, P.G. Khalatur, and A.R. Khokhlov. Conformational polymorphism of amphiphilic polymers in a poor solvent. *Macromolecules*, 36:10103, 2003.
- [11] V.V. Vasilevskaya, A.A. Klochkov, A.A. Lazutin, P.G. Khalatur, and A.R. Khokhlov. H/a (hydrophobic/amphiphilic) copolymer model: Coil-globule transition versus aggregation. *Macromolecules*, 37:5444, 2004.
- [12] I.M. Okhapkin, A.A. Askadskii, V.A. Markov, E.E. Makhaeva, and A.R. Khokhlov. Two-dimensional classification of amphiphilic monomers based on interfacial and partitioning properties. 2. amino acids and amino acid residues. *Colloid Polym Sci*, 284(6):575, 2006.

- [13] К.Б. Терёшкина и К.В. Шайтан. *Молекулярная динамика белков и пептидов*. Издательство Московского Университета, 2005.
- [14] Л.Д. Ландау и Е.М. Лившиц. *Теоретическая физика, т.5 Статистическая физика, §87-91, Теория слабых растворов*. ФИЗМАТЛИТ, 1995.
- [15] R. Wolfenden, L. Andersson, P. M. Cullis, and C. C. B. Southgate. Affinities of amino acid side chains for solvent water. *Biochemistry*, 20:849, 1981.
- [16] M.H. Abraham, G.S. Whiting, R. Fuchs, and E.J. Chambers. Thermodynamics of solute transfer from water to hexadecane. *J. Chem. Soc., Perkin Trans. 2*, 2:291, 1990.
- [17] J. Hine and P. K. Mookerjee. The intrinsic hydrophilic character of organic compounds. correlations in terms of structural contributions. *J. Org. Chem.*, 40:292, 1975.
- [18] A.V. Plyasunov and E.L. Shock. Infinite dilution partial molar properties of aqueous solutions of nonelectrolytes. ii. equations for the standard thermodynamic functions of hydration of volatile nonelectrolytes over wide ranges of conditions including subcritical temperatures. *Geochim. Cosmochim. Acta*, 64:439, 2000.
- [19] Л.Д. Ландау и Е.М. Лившиц. *Теоретическая физика, т.5 Статистическая физика, §157, Поверхностное натяжение растворов*. ФИЗМАТЛИТ, 1995.
- [20] Kh. Graf, Michael Kappl, and Hans-Jurgen Butt. *Physics and Chemistry of Interfaces*. Wiley-VCH, 2003.
- [21] C. Kempball and E.K. Rideal. The adsorption of vapours on mercury. i. non-polar substances. *Proc Roy Soc A*, 187:53, 1946.
- [22] I. Langmuir. The constitution and fundamental properties of solids and liquids. ii. liquids. *J Am Chem Soc*, 39:1848, 1917.
- [23] A. Ben-Naim and Y. Marcus. Solvation thermodynamics of nonionic solutes. *J Chem Phys*, 81:2016, 1984.
- [24] T.L. Hill. *Statistical Mechanics*, chapter 31. Potential of Mean Force and Superposition Approximation. Courier Dover Publications, 1987.
- [25] W. L. Jorgensen and J. Tirado-Rives. The opls force field for proteins. energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110:1657, 1988.
- [26] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. In B. Pullman, editor, *Intermolecular Forces*, page 331. Reidel, Dordrecht, 1981.
- [27] M.R. Shirts, J.W. Pietra, W.C. Swope, and V.S. Pande. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J Chem Phys*, 119:5740, 2003.
- [28] S. Miyamoto and P.A. Kollman. Settle: an analytical version of the shake and rattle algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8):952, 1992.
- [29] D. Frenkel and B. Smit. *Understanding molecular simulation*. Academic Press, 2002.

- [30] W.G. Hoover. *Time Reversibility, Computer Simulation, and Chaos*. World Scientific, 2001.
- [31] А.Ю. Лоскутов и А.С. Михайлов. *Введение в синергетику*. УРСС, 1990.
- [32] http://en.wikipedia.org/wiki/ergodic_hypothesis.
- [33] R.E. Gillilan and K.R. Wilson. Shadowing, rare events, and rubber bands - a variational verlet algorithm for molecular dynamics. *J. Chem. Phys.*, 97:1757, 1992.
- [34] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81:3684, 1984.
- [35] G.J. Martyna, M.L. Klein, and M. Tuckerman. Nose-hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97:2635, 1992.
- [36] A.S. Lemak and N.K. Balabaev. A comparison between collisional dynamics and brownian dynamics. *Molecular Simulation*, 15:223, 1995.
- [37] W. F. van Gunsteren and H. J. C. Berendsen. A leap-frog algorithm for stochastic dynamics. *Molecular Simulation*, 1:173, 1998.
- [38] D. van der Spoel and P. J. van Maaren. The origin of layer structure artifacts in simulations of liquid water. *J. Chem. Theor. Comp.*, 2:1, 2006.
- [39] M. P. Allen and D. J. Tildesley. *Computer Simulations of Liquids*. Oxford Science Publications, 1987.
- [40] David van der Spoel, Erik Lindahl, Berk Hess, Aldert R. van Buuren, Emile Apol, Pieter J. Meulenhoff, D. Peter Tieleman, Alfons L.T.M. Sijbers, K. Anton Feenstra, Rudi van Drunen, and Herman J.C. Berendsen. Gromacs user manual, version 3.3, 2006. <ftp://ftp.gromacs.org/pub/manual/manual-3.3.pdf>.
- [41] T. Darden, D. York, and L. Pedersen. Particle mesh ewald: An n-log(n) method for ewald sums in large systems. *J. Chem. Phys.*, 98:10089, 1993.
- [42] Denis J. Evans and Debra J. Searles. Equilibrium microstates which generate second law violating steady states. *Physical Review E*, 50:1645, 1994.
- [43] Wang GM, Sevick EM, Mittag E, Searles DJ, and Evans DJ. Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales. *Phys Rev Lett*, 89:50601, 2002.
- [44] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys Rev Lett*, 78:2690, 1997.
- [45] Gavin E. Crooks. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E*, 61:2361, 2000.
- [46] Michael R. Shirts, Eric Bair, Giles Hooker, and Vijay S. Pande. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys. Rev. Lett.*, 91:140601, 2003.

- [47] CH Bennet. Efficient estimation of free energy differences from monte carlo data. *J. Comp. Phys.*, 22:245, 1976.
- [48] Michael R. Shirts and Vijay S. Pande. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J Chem Phys*, 122:134508, 2005.
- [49] J. P. Nicolas and N. R. de Souza. Molecular dynamics study of the n-hexane–water interface: Towards a better understanding of the liquid–liquid interfacial broadening. *J. Chem. Phys.*, 120:2564, 2003.
- [50] T. Somasundaram, R. M. Lynden-Bell, and C. H. Patterson. The passage of gases through the liquid water/vapour interface: a simulation study. *Phys. Chem. Chem. Phys.*, 1:143, 1999.
- [51] G.A. Kaminski, R.A. Friesner, J. Rives, and W.L. Jorgensen. Evaluation and reparametrization of the opls-aa force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B*, 105:6474, 2001.
- [52] E. Lindahl, B. Hess, and D. van der Spoel. Gromacs 3.0: a package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling*, 7:306, 2001.
- [53] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, 52:7182, 1981.
- [54] S. Miyamoto and P.A. Kollman. Settle: An analytical version of the shake and rattle algorithms for rigid water models. *J. Comp. Chem.*, 13:952, 1992.
- [55] B. Hess, H. Bekker, H.J.C. Berendsen, and J.G.E.M. Fraaije. Lincs: A linear constraint solver for molecular simulations. *J. Comp. Chem.*, 18:1463, 1997.
- [56] E.B. Leodidis and T.A. Hutton. Amino acids in aot reversed micelles. 2. the hydrophobic effect and hydrogen bonding as driving forces for interfacial solubilization. *J. Phys. Chem.*, 94:6411, 1990.
- [57] J.W. Gibbs. *Heterogeneous Equilibrium*. 1875-78.
- [58] E.T. Jaynes. The gibbs paradox. In Smith C.R., Erickson G.J., and Neudorfer P.O., editors, *Maximum Entropy and Bayesian Methods*, pages 1–22. Kluwert Academic Publishers, Dordrecht, Holland, 1992.