

# ПОВЕРХНОСТЬ ПОТЕНЦИАЛЬНОЙ ЭНЕРГИИ ПОЛИПЕПТИДНЫХ ЦЕПОЧЕК АЛАНИНА

*И. А. Соловьев<sup>a,b\*</sup>, А. В. Якубович<sup>a</sup>, А. В. Соловьев<sup>a,b</sup>, В. Грайнер<sup>b</sup>*

*<sup>a</sup> Физико-технический институт им. А. Ф. Иоффе Российской академии наук  
194021, Санкт-Петербург, Россия*

*<sup>b</sup> Frankfurt Institute for Advanced Studies,  
Johann Wolfgang Goethe University  
60438, Frankfurt am Main, Germany*

Поступила в редакцию 18 мая 2005 г.

Проведено исследование многомерной поверхности потенциальной энергии цепочек, состоящих из трех и шести аланинов, по отношению к связанным с кручением этих молекул относительно полипептидной цепи степеням свободы, которые ответственны за фолдинг подобных молекул и белков. Расчет потенциальных поверхностей выполнен из первых принципов в рамках теории функционала плотности с учетом всех электронов в системе. Проведены оценки вероятностей переходов между различными устойчивыми конформациями молекул. Полученные результаты сравниваются с имеющимися данными молекулярно-динамических расчетов, а также с экспериментальными данными. Исследовано влияние вторичной структуры полипептидной цепи на ее конформационные свойства относительно вращения. Для цепочки из шести аминокислот со структурой спирала и листа продемонстрировано влияние вторичной структуры на устойчивые изомерные состояния молекулы.

PACS: 73.61.Ph

## 1. ВВЕДЕНИЕ

Исследование атомных столкновений является одной из центральных задач атомной физики. Современные методы теории атомных столкновений позволяют рассчитывать сечения многих столкновительных процессов с участием многоатомных молекул. В последнее время все больше внимания уделяется экспериментальным и теоретическим исследованиям столкновений с такими сложными объектами, как атомные кластеры или небольшие биологические молекулы (см., например, [1, 2]). Для теоретического исследования таких процессов необходимо детальное исследование электронной и ионной структуры молекулярных объектов, участвующих в столкновении, которое стало возможным благодаря наличию мощных компьютеров. Важной характеристикой, необходимой для описания подобных процессов, является геометрия ионного остова молекулы и ее поверхности потенциальной энергии.

Поверхность потенциальной энергии системы содержит важную информацию о статических и динамических свойствах молекулы, ее возможных начальных и конечных состояниях. Исследование поверхности потенциальной энергии сложных молекулярных систем представляет самостоятельную и достаточно сложную теоретическую задачу. В настоящей работе приведен пример подобного исследования для поверхности потенциальной энергии небольших полипептидных цепочек.

В последнее время стали возможными экспериментальные исследования небольших фрагментов белков и полипептидов в газообразном состоянии с использованием MALDI масс-спектроскопии [3–6]. С теоретической точки зрения исследования свойств небольших полипептидов представляют заметный интерес, поскольку эти системы можно рассчитывать с достаточно высокой точностью из первых принципов. Соответственно, возможно проводить сравнение результатов теории и эксперимента. Полученная информация может быть использована для построения теоретических моделей более сложных

---

\*E-mail: solovyov@fias.uni-frankfurt.de,  
ilia@th.physik.uni-frankfurt.de

белковых соединений.

Полипептиды отличаются друг от друга первичной и вторичной структурами [7–10]. При одинаковой первичной структуре различия в структуре молекулы соответствуют ее разным конформациям (геометрическим конфигурациям). Вполне естественно, что химические и физические свойства различных конформаций сложных молекул могут существенно различаться. Число различных конформаций (изомерных состояний) быстро возрастает с увеличением размера системы, поэтому поиск наиболее устойчивых конформаций биомолекулы значительно усложняется с увеличением ее размера. С помощью методов ЯМР и дифракции рентгеновских лучей [9] было установлено, что одними из наиболее характерных элементов вторичной структуры белка являются лист и спираль.

Главное отличие листа от спирали связано с различием двугранных углов, образованных атомами полипептидной цепи в этих молекулярных структурах. При повышении температуры должна активироваться степень свободы, связанная с кручением боковых радикалов вокруг полипептидной цепи, при этом меняется конформация молекулы. Большой интерес представляет изучение этого перехода и оценка его характерных времен, поскольку он связан с одним из основных вопросов физики белка — его фолдингом. Для изучения этого перехода необходимо исследовать поверхность потенциальной энергии аминокислотной цепочки относительно кручения боковых радикалов вдоль полипептидной цепи. Кроме вопросов, связанных с фолдингом, такое исследование несет в себе важную информацию о зависимости энергии от угловых переменных, характеризующих цепочку аминокислот, которая очень важна для модельного описания больших белковых структур.

В настоящей работе исследована потенциальная поверхность цепочек аланина. Эти молекулы были выбраны, потому что, во-первых, они часто встречаются в белках и, во-вторых, они являются сравнительно простыми объектами для теоретического расчета ввиду их небольшого размера. Расчет был выполнен в рамках теории функционала плотности, учитывающей все электроны в системе.

Ранее достаточно детально исследовались лишь дипептиды аланина и глицина, а также их аналог (*S*)- $\alpha$ -(формиламино)пропанамид. В работах [11–13] эти молекулы были исследованы в рамках теории Хартри–Фока. В этих работах были рассчитаны потенциальные поверхности систем в зависимости от углов кручения молекулы. Были определены некоторые устойчивые состояния дипептидов, соответству-

ющие различным конформациям молекул. Каждое устойчивое состояние молекулы дополнительно исследовалось на основе теории возмущений, учитывавшей межэлектронные корреляционные взаимодействия. В работах [14–18] были исследованы различные конформационные свойства молекулы, определены устойчивые структуры, рассчитаны их энергии в рамках теории функционала плотности. В работе [19] обсуждалась динамика аналога дипептида аланина, рассчитанная в рамках теории функционала плотности. В этой работе был исследован переход дипептида из одной устойчивой конформации в другую и рассчитано характерное время этого перехода. В работе [20] с использованием методов Монте-Карло было рассчитано характерное время перехода дипептида аланина из одной конформации в другую в водном окружении. В работе [21] был определен эмпирический потенциал парного взаимодействия двух аланинов на основе экспериментальных данных, взятых из базы данных для белков [9]. При этом аланины в белке рассматривались как материальные точки.

В ряде работ исследованы также свойства трипептидов. В работах [22–26] с помощью классической молекулярной механики и полуэмпирических потенциалов (таких как GROMOS, CHARMM и AMBER) была исследована динамика трипептида аланина и глицина. В работе [27] в рамках теории Хартри–Фока были найдены некоторые конформации трипептида аланина и дипептида глицина. В работах [28, 29] с помощью двухмерной колебательной спектроскопии были экспериментально измерены колебательные и вращательные свойства трипептида аланина в водном окружении. В работе [30] были измерены рамановские и ИК-спектры трипептида аланина и глицина в щелочном, нейтральном и кислом окружениях. В этой работе также была разработана феноменологическая модель для определения устойчивых конформаций трипептидов.

Полипептиды исследовались гораздо меньше. Нам известно лишь несколько работ на эту тему. Так, например, в работе [31] обсуждались устойчивые конформации нейтральных и заряженных гексапептидов аланина, рассчитанные на основе эмпирических потенциалов. В работе [32] было проведено экспериментальное исследование различных конформаций цепочки из семи аланинов с помощью ЯМР при различных температурах. В работе [33] методами эмпирической молекулярной динамики, основанной на методах Монте-Карло, был описан полипептид аланина, состоящий из 21 аминокислоты.

В настоящей работе впервые проведено исследо-

вание из первых принципов многомерной поверхности потенциальной энергии цепочек, состоящих из трех и шести аланинов, по отношению к степеням свободы, связанным с кручением этих молекул относительно полипептидной цепи. Этот расчет был выполнен в рамках теории функционала плотности (ТФП). Ранее такого рода вычисления проводились лишь для дипептидов (см., например, [11, 12, 19]). Для больших молекул исследовались лишь отдельные изомерные состояния (см. ссылки, приведенные выше). Следует также отметить, что настоящий расчет поверхностей потенциальной энергии для цепочек из трех и шести аминокислот впервые выполнен с учетом всех электронов в системе. Анализ энергетических поверхностей позволил оценить вероятность перехода между различными устойчивыми конформациями молекул. Было выполнено сравнение полученных результатов с имеющимися результатами молекулярно-динамических расчетов и экспериментальными данными. В настоящей работе также исследовано влияние вторичной структуры полипептидной цепи на ее конформационные свойства относительно вращения. Так, для цепочки из шести аминокислот со структурами спирали и листа продемонстрировано влияние вторичной структуры на устойчивые изомерные состояния молекулы. Некоторые предварительные результаты данной работы были опубликованы на сервере препринтов [34, 35].

Работа построена следующим образом. В разд. 2 дано краткое описание используемых теоретических методов, в разд. 3 приведены результаты выполненных расчетов и проведено их детальное обсуждение, в разд. 4 подведены итоги работы. В работе использована атомная система единиц  $\hbar=m=|e|=1$ , если не указаны другие единицы.

## 2. ТЕОРЕТИЧЕСКИЕ МЕТОДЫ

В данной работе исследованы свойства многомерных поверхностей потенциальной энергии полипептидов аланина на основе теории функционала плотности. Поверхности потенциальной энергии являются многомерной функцией координат ядер атомов, образующих полипептид. Мы исследовали поведение поверхности потенциальной энергии как функции углов, образованных атомами полипептидной цепи. Для исследования поверхностей потенциальной энергии было использовано приближение Борна–Оппенгеймера, в рамках которого можно разделить движение электронной и ионной подсистем.

Теория функционала плотности является распространенным средством для расчета квантовых многоэлектронных систем, в которых многоэлектронные корреляции играют существенную роль. Формализм ТФП хорошо известен и изложен во множестве учебников [1, 2]. Поэтому в настоящей работе мы привели лишь основные уравнения и изложили главные идеи этого метода.

Основное уравнение ТФП, уравнение Кона–Шама, имеет следующий вид [1, 2]:

$$\left( \frac{\hat{p}^2}{2} + U_{ions} + V_H + V_{xc} \right) \psi_i = \varepsilon_i \psi_i. \quad (1)$$

Здесь первое слагаемое описывает кинетическую энергию  $i$ -го электрона с волновой функцией  $\psi_i$  и энергией  $\varepsilon_i$ ;  $U_{ions}$  описывает притяжение электрона к ионным центрам,  $V_H$  — хартриевскую часть межэлектронного взаимодействия [36], а  $V_{xc}$  называется локальным обменно-корреляционным потенциалом, который определяется как функциональная производная обменно-корреляционного функционала энергии:

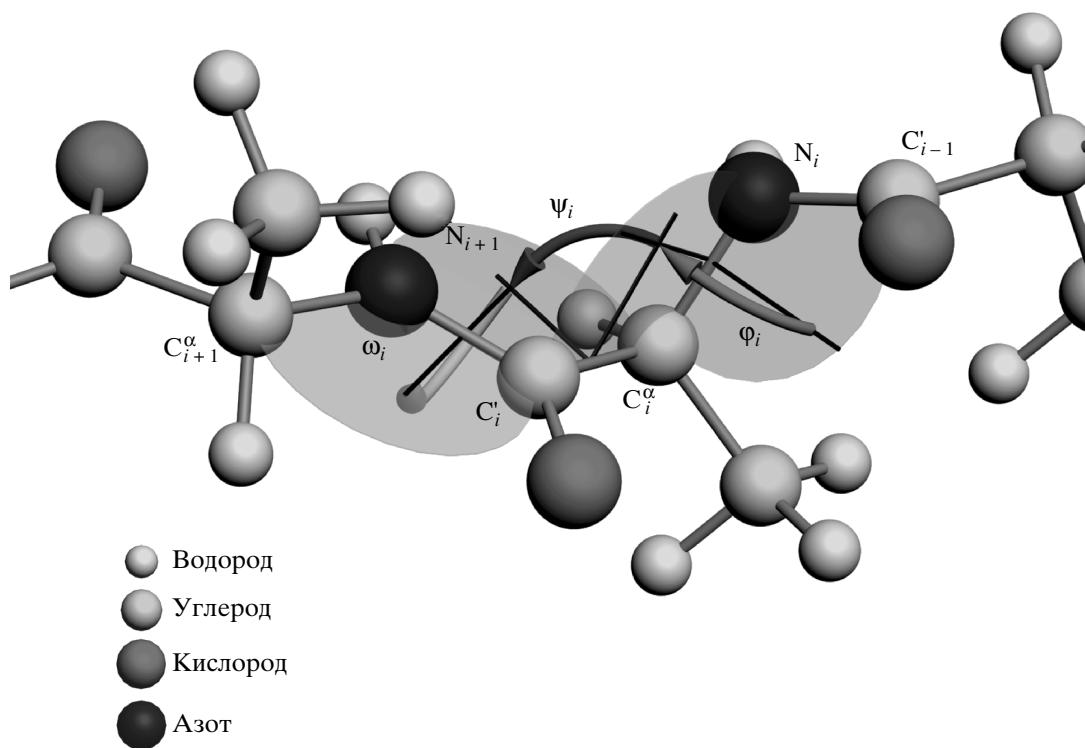
$$V_{xc} = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}. \quad (2)$$

Соотношение (2) является точным. Оно получено в рамках теории Хонберга [37]. Однако установить точный вид функционала  $E_{xc}$ , который был бы применим к различным системам и условиям, оказывается невозможным. Как следствие этого существует множество различных функционалов (см., например, [1]), применимых к различным случаям.

Приближенные обменно-корреляционные функционалы, которые используются в рамках ТФП, делятся на две части, обменную и корреляционную:

$$E_{xc}[\rho] = E_x(\rho) + E_c(\rho). \quad (3)$$

Обе части являются функционалами электронной плотности и могут быть двух типов: локальные, зависящие только от плотности  $\rho$ , и с градиентным поправками, зависящими от плотности  $\rho$  и ее градиента  $\nabla \rho$ . Существует множество различных обменно-корреляционных функционалов. В данной работе использован трехпараметрический обменный функционал Бекке в сочетании с корреляционным функционалом Ли, Янга и Пара, которые учитывают градиентные поправки [38–40] (B3LYP).



**Рис. 1.** Двугранные углы  $\varphi$ ,  $\psi$  и  $\omega$ , используемые для описания поверхности потенциальной энергии полипептидной цепи

### 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

#### 3.1. Определение крутильных степеней свободы полипептида

Приведем результаты расчета поверхностей потенциальной энергии цепочек аланина. Мы исследовали потенциальные энергии цепочек, состоящих из трех и шести аминокислот, в зависимости от двугранных углов  $\varphi$  и  $\psi$ , определенных на рис. 1.

Оба угла определяются четырьмя соседними атомами в полипептидной цепочке. Угол  $\varphi_i$  определяется как двугранный угол между плоскостями, образованными атомами ( $C'_{i-1}-N_i-C_i^{\alpha}$ ) и ( $N_i-C_i^{\alpha}-C'_i$ ). Угол  $\psi_i$  определяется как двугранный угол между плоскостями, образованными атомами ( $N_i-C_i^{\alpha}-C'_i$ ) и ( $C_i^{\alpha}-C'_i-N_{i+1}$ ). Кроме углов  $\varphi_i$  и  $\psi_i$  существует третий угол, образованный атомами полипептидной цепи — угол  $\omega_i$ , который определяется как двугранный угол между плоскостями, образованными атомами ( $C_i^{\alpha}-C'_i-N_{i+1}$ ) и ( $C'_i-N_{i+1}-C_{i+1}^{\alpha}$ ). Нумерация атомов полипептидной цепи ведется, начиная с  $\text{NH}_2$ -конца полипептида. Углы  $\varphi_i$ ,  $\psi_i$  и  $\omega_i$  могут принимать все возможные значения в интервале  $[-180^\circ; 180^\circ]$ . Для однозначного определения углов необходимо задать

их направление отсчета. В данной работе использовано общепринятое определение [10], согласно которому отсчет углов  $\varphi_i$ ,  $\psi_i$  и  $\omega_i$  проводится по часовой стрелке, если смотреть на молекулу с ее  $\text{NH}_2$ -конца (см. рис. 1).

Углы  $\varphi_i$  и  $\psi_i$  могут быть определены для любой аминокислоты в цепочке, кроме крайних. В настоящей работе были рассмотрены углы  $\varphi$  и  $\psi$ , соответствующие средней аминокислоте полипептида. Для простоты индексы в обозначении углов опущены.

Для изучения кручения цепочек аминокислот вдоль полипептидной цепи необходимо задать начальную структуру молекулы. С увеличением размера системы возрастает количество ее стабильных изомерных состояний, различающихся лишь своей вторичной структурой. В настоящей работе исследованы цепочки аминокислот с двумя наиболее распространенными типами вторичной структуры, имеемыми конформациями спирали и листа.

При изменении углов  $\varphi$  и  $\psi$  в центральной аминокислоте полипептид может принимать состояния, заметно отличающиеся от структуры листа и спирали. В этих случаях для удобства обсуждения полипептид относится к типу листа или спирали, если

структуря листа или спирали возникает при определенных значениях углов  $\varphi$  и  $\psi$  в центральной аминокислоте.

### 3.2. Зависимость энергии полипептида от двугранного угла $\omega$

Для каждой аминокислоты существуют лишь три двугранных угла, образованных атомами полипептидной цепи и описывающих кручение полипептида. Угол  $\omega$  (см. рис. 1) образован на стыке двух аминокислот и тем самым отличается от углов  $\varphi$  и  $\psi$ . Между атомами  $C'_i$  и  $N_{i+1}$  образуется квазидвойная связь, которая препятствует кручению полипептида. Поэтому можно говорить об угле  $\omega$  как о жесткой степени свободы, слабо зависящей от типа полипептида и значений остальных характерных углов.

Методом B3LYP/6-31G(d) была рассчитана зависимость энергии трипептида аланина от угла  $\omega$  при различных значениях углов  $\varphi$  и  $\psi$ . Расчет показал, что в системе существуют два минимума при  $\omega = 0$  и  $180^\circ$ , положение которых практически не зависит от углов  $\varphi$  и  $\psi$ . Состояние системы с  $\omega = 180^\circ$  энергетически более выгодно. Разница в энергии зависит от значений углов  $\varphi$  и  $\psi$  и составляет 0.12–0.30 эВ. Высота барьера между двумя минимумами меняется незначительно с изменением углов  $\varphi$  и  $\psi$  и составляет около 0.81 эВ = 18.68 ккал/моль.

Расчет показывает, что при температуре, близкой к комнатной, значение угла  $\omega$  меняется незначительно. Поэтому его изменение можно не учитывать при дальнейшем изучении поверхности потенциальной энергии, характеризующей кручение полипептидов.

Зависимость потенциальной энергии полипептидов от углов  $\varphi$  и  $\psi$  при кручении молекулы при температурах, близких к комнатной, оказывается значительно более сложной и поэтому представляет значительный интерес.

### 3.3. Поверхность потенциальной энергии трипептида аланина

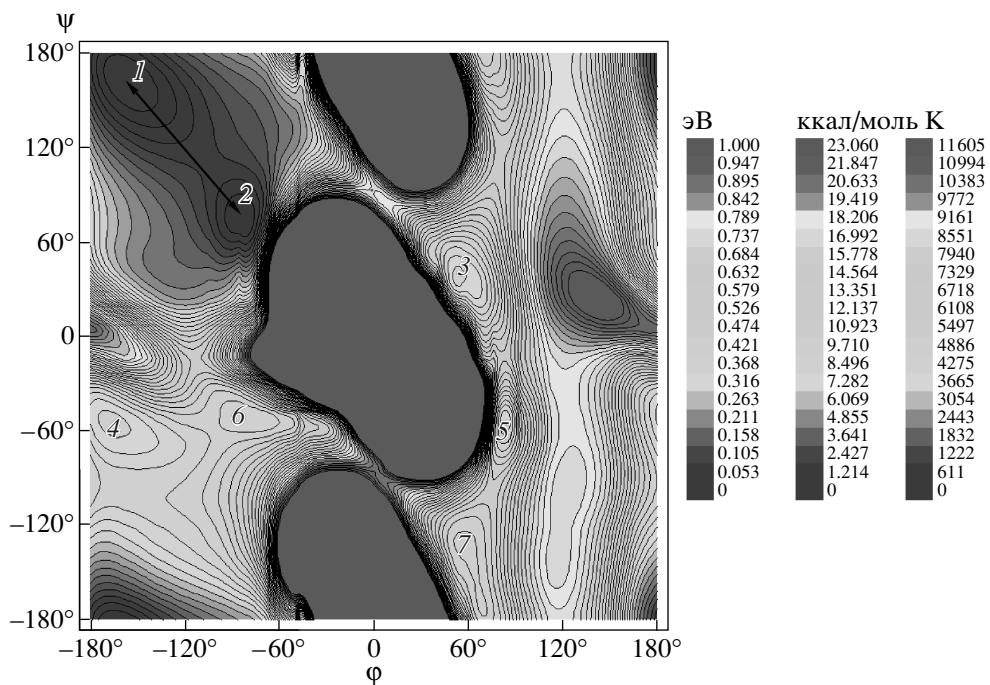
На рис. 2 показана поверхность потенциальной энергии трипептида аланина, рассчитанная методом B3LYP/6-31G(2d,p). Шкала энергии приведена в эВ, ккал/моль и К и отсчитана от глобального минимума данной поверхности.

Как видно на рис. 2, на энергетической поверхности наблюдается несколько минимумов, они обозначены цифрами в порядке возрастания их энергии. Каждому минимуму соответствует устойчивое

изомерное состояние молекулы. На рисунке видно два выраженных минимума, обозначенных цифрами 1 и 2, а также несколько менее глубоких, обладающих большей энергией (3–7). Конформации молекулы, соответствующие этим минимумам, приведены на рис. 3. Штриховыми линиями на рисунке показаны основные водородные связи в системе, которые определены как связи между атомами кислорода и атомами водорода, связанными с атомами азота и расположенными на расстоянии, не превышающем 2.9 Å.

Как видно на этом рисунке, геометрии, соответствующие различным устойчивым состояниям, заметно отличаются друг от друга. Для исследования потенциальных поверхностей использовалась следующая процедура. После определения устойчивой структуры молекулы фиксировались все степени свободы системы и менялись лишь углы  $\varphi$  и  $\psi$  в центральной аминокислоте. Таким методом были рассчитаны все потенциальные поверхности, представленные в этом разделе. Данный метод позволяет эффективно получать информацию об основных конформациях, соответствующих минимумам на поверхности потенциальной энергии. Таким образом были определены соответствующие им углы  $\varphi$  и  $\psi$ . Заметим, что абсолютные значения энергий, соответствующих различным конформациям, определены не совсем точно, поскольку данный метод расчета не учитывает релаксацию всех атомов при кручении молекулы. Вычисления потенциальной поверхности цепочки с учетом ее релаксации требуют в 20–30 раз больше компьютерного времени. Систематическое исследование роли релаксации при формировании поверхности потенциальной энергии с учетом всех электронов системы в работе не проводилось. Однако была выполнена полная оптимизация структур полипептидов, соответствующих различным минимумам на поверхности потенциальной энергии с учетом всех степеней свободы молекулы.

На рис. 3 сравниваются стабильные состояния, соответствующие трипептиду аланина, с учетом и без учета релаксации молекулы. Как видно из этого сравнения, характерные двугранные углы  $\varphi$  и  $\psi$  меняются в пределах погрешности в 10 %, что может быть объяснено незначительным влиянием остальных степеней свободы на значение двугранных углов в полипептиде вблизи минимума энергии. В результате релаксации по остальным степеням свободы происходит небольшая перестройка боковых атомов (или радикалов в случае больших аминокислот), что приводит к понижению энергии системы. В связи с этим относительная энергия минимумов может



**Рис. 2.** Поверхность потенциальной энергии трипептида аланина, рассчитанная методом B3LYP/6-31G(2d,p). Энергии приведены в эВ, ккал/моль и К. Цифрами помечены минимумы на потенциальной поверхности. Стрелками показаны рассмотренные переходы между различными конформациями молекулы

измениться при учете полной релаксации молекулы. В данной работе рассчитаны поверхности потенциальной энергии полипептидов на сетке с шагом  $18^\circ$ . Выполнение расчета с более мелким шагом не имело бы смысла, потому что учет всех степеней свободы приводил бы к большим изменениям углов  $\varphi$  и  $\psi$ , чем шаг сканирования.

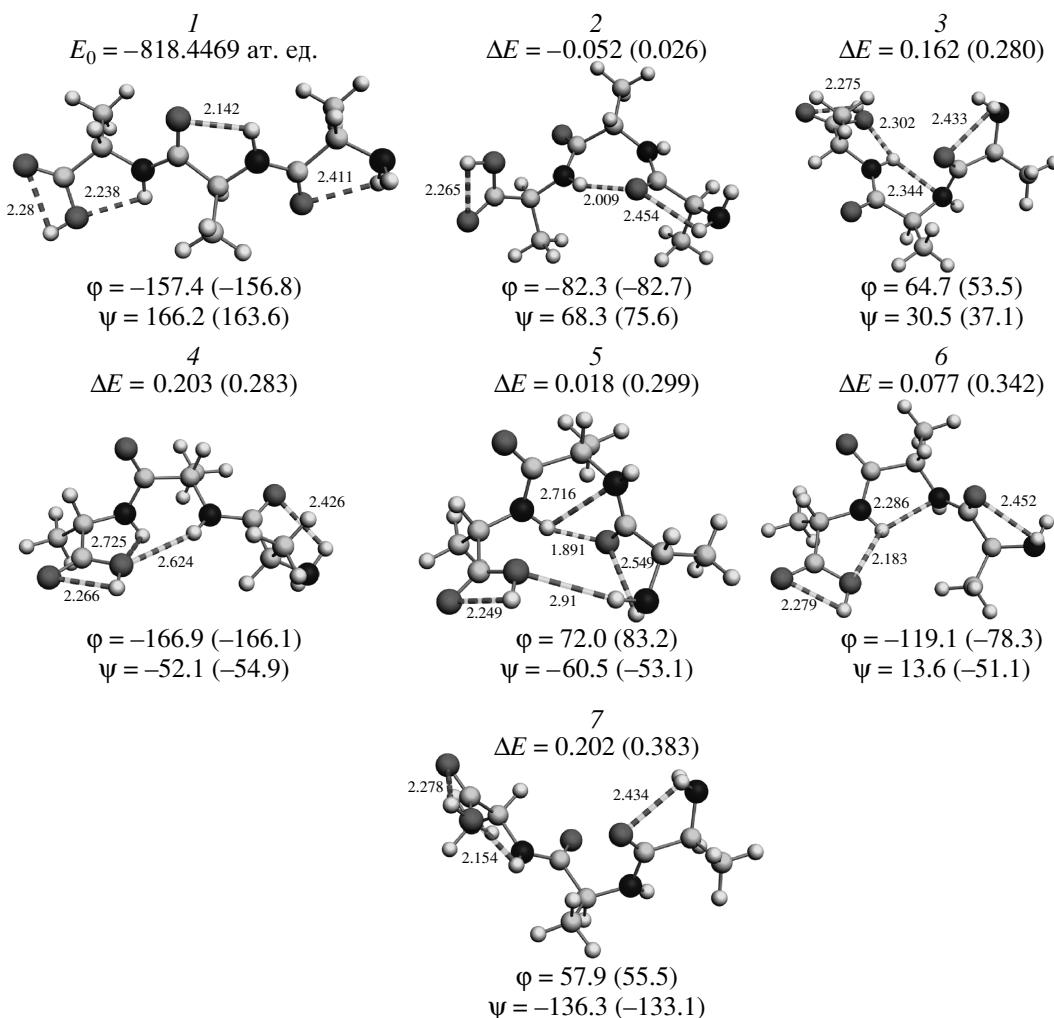
На поверхности потенциальной энергии трипептида аланина, в отличие от поверхности потенциальной энергии трипептида глицина [34], существует дополнительный максимум при  $\varphi = 120^\circ \pm 50^\circ$ ,  $\psi = 30^\circ \pm 30^\circ$ . Он возникает при перекрытии боковых  $\text{CH}_3$ -радикалов аланина, которые отсутствуют у глицина.

Отметим, что не все минимумы на поверхности потенциальной энергии соответствуют тем или иным устойчивым конформациям полипептида. Например, минимума 6 на реальной многомерной энергетической гиперповерхности не существует. Это связано с тем, что при построении энергетической поверхности учитывались только две степени свободы. Поэтому для достоверного определения устойчивых конформаций молекулы необходимо проводить оптимизацию соответствующих минимумам структур, показанных на рис. 3, с учетом всех степеней

свободы.

На рис. 2 видно, что некоторые области на потенциальной поверхности являются запрещенными. В этих областях потенциальная энергия молекулы заметно возрастает из-за сближения отдельных атомов в процессе кручения полипептида. Учет релаксации молекул, хотя и уменьшает значения энергии системы, мало влияет на положение запрещенных зон. На рис. 2 видны две выраженные области вблизи значений углов  $\varphi$  и  $\psi$  ( $0, 0$ ) и ( $0, 180^\circ$ ), соответственно. В первой области происходит сближение атомов водорода и кислорода на расстояния, значительно меньшие характерной длины связи в молекуле. При этом возникает сильное межатомное отталкивание, обусловленное перекрытием электронных оболочек атомов и обменным взаимодействием электронов. Второй максимум соответствует сближению двух атомов кислорода на расстояние порядка длины связи, что также приводит к увеличению энергии системы за счет кулоновского отталкивания этих атомов.

Рассмотрим подробнее энергетический барьер для перехода между наиболее глубокими минимумами  $1 \leftrightarrow 2$ , отмеченный стрелкой на рис. 2. Для более точного определения высоты барьера для дан-



**Рис. 3.** Оптимизированные геометрии трипептида аланина. Различные конформации молекулы, соответствующие минимумам на поверхности потенциальной энергии, представленной на рис. 2, отмечены соответствующими цифрами. Под каждой структурой приведен набор углов  $\varphi$  и  $\psi$ , полученных с учетом релаксации всех степеней свободы в системе. В скобках указаны величины этих углов, найденные без учета релаксации. Энергии молекул приведены над каждой из них, при этом энергии даны в эВ и отсчитаны от энергии состояния 1 (энергия этого состояния приведена в ат. ед.). В скобках указаны энергии, вычисленные без учета релаксации всех степеней свободы в системе.

Штриховыми линиями показаны основные водородные связи в системе. Их длины приведены в ангстремах

ногого перехода была проведена дополнительная оптимизация структуры полипептида по всем степеням свободы, кроме двугранных углов  $\varphi$  и  $\psi$ , отвечающих переменной сканирования. В результате оптимизации высота барьера уменьшилась приблизительно на 0.03 эВ.

На основе анализа энергетических барьеров можно сделать оценку характерного времени перехода из одного состояния в другое. Для этого необходимо воспользоваться формулой Аррениуса, которая

имеет вид

$$\frac{1}{\tau} = \Omega e^{-\Delta E/kT}, \quad (4)$$

где  $\tau$  — характерное время перехода системы из одного состояния в другое,  $\Omega$  — предэкспоненциальный фактор, определяющий частоту подхода системы к барьеру,  $\Delta E$  — высота барьера,  $T$  — температура системы,  $k$  — постоянная Больцмана.

Чтобы проверить применимость этой модели для анализа динамики цепочек аланина, был проведен тестовый расчет и оценено характерное время пе-

рехода из одного устойчивого состояния в другое для аналога дипептида аланина ((S)- $\alpha$ -(формиламино)пропанамид), поскольку эта система исследовалась ранее. Методом B3LYP/6-31+G(2d,p) были рассчитаны высоты барьеров для переходов между наиболее глубокими минимумами  $1 \rightarrow 2$  и  $2 \rightarrow 1$  с учетом релаксации всех степеней свободы в системе. Для перехода  $1 \rightarrow 2$  высота барьера составила

$$\Delta E_{1 \rightarrow 2} = 0.047 \text{ эВ},$$

а для перехода  $2 \rightarrow 1$

$$\Delta E_{2 \rightarrow 1} = 0.079 \text{ эВ.}$$

Частота  $\Omega$ , соответствующая колебаниям аналога дипептида аланина вблизи минимума 1 вдоль координаты, соединяющей минимумы 1 и 2, рассчитанная методом B3LYP/6-31+G(2d,p) составила  $42.87 \text{ см}^{-1}$ . Поэтому получаем

$$\tau_{2 \times Ala}^{1 \rightarrow 2} \approx 5 \text{ пс}, \quad \tau_{2 \times Ala}^{2 \rightarrow 1} \approx 17 \text{ пс.}$$

Этот результат находится в прекрасном согласии с результатом [19], полученным на основе молекулярной динамики, предсказывающим

$$\tau_{2 \times Ala}^{1 \rightarrow 2} \approx 7 \text{ пс}, \quad \tau_{2 \times Ala}^{2 \rightarrow 1} \approx 19 \text{ пс.}$$

Данное сравнение показывает, что формула (4) может быть использована для оценки времен перехода между различными устойчивыми состояниями молекулы.

Используя метод B3LYP/6-31G(2d,p), мы рассчитали характерные частоты для цепочек аланина. Для трипептида аланина частота равна  $32.04 \text{ см}^{-1}$ . Из проведенного расчета следует, что для перехода  $1 \rightarrow 2$  высота барьера составила

$$\Delta E_{1 \rightarrow 2} = 0.066 \text{ эВ,}$$

а для перехода  $2 \rightarrow 1$  высота барьера составила

$$\Delta E_{2 \rightarrow 1} = 0.114 \text{ эВ.}$$

В результате получаем

$$\tau_{3 \times Ala}^{1 \rightarrow 2} \approx 13 \text{ пс}, \quad \tau_{3 \times Ala}^{2 \rightarrow 1} \approx 86 \text{ пс.}$$

Заметим, что эти времена могут быть измерены методами ЯМР [10, 41].

### 3.4. Поверхность потенциальной энергии гексапептида аланина со вторичной структурой листа и спиралы

На рис. 4 и 5 приведены поверхности потенциальной энергии гексапептида аланина со вторичной

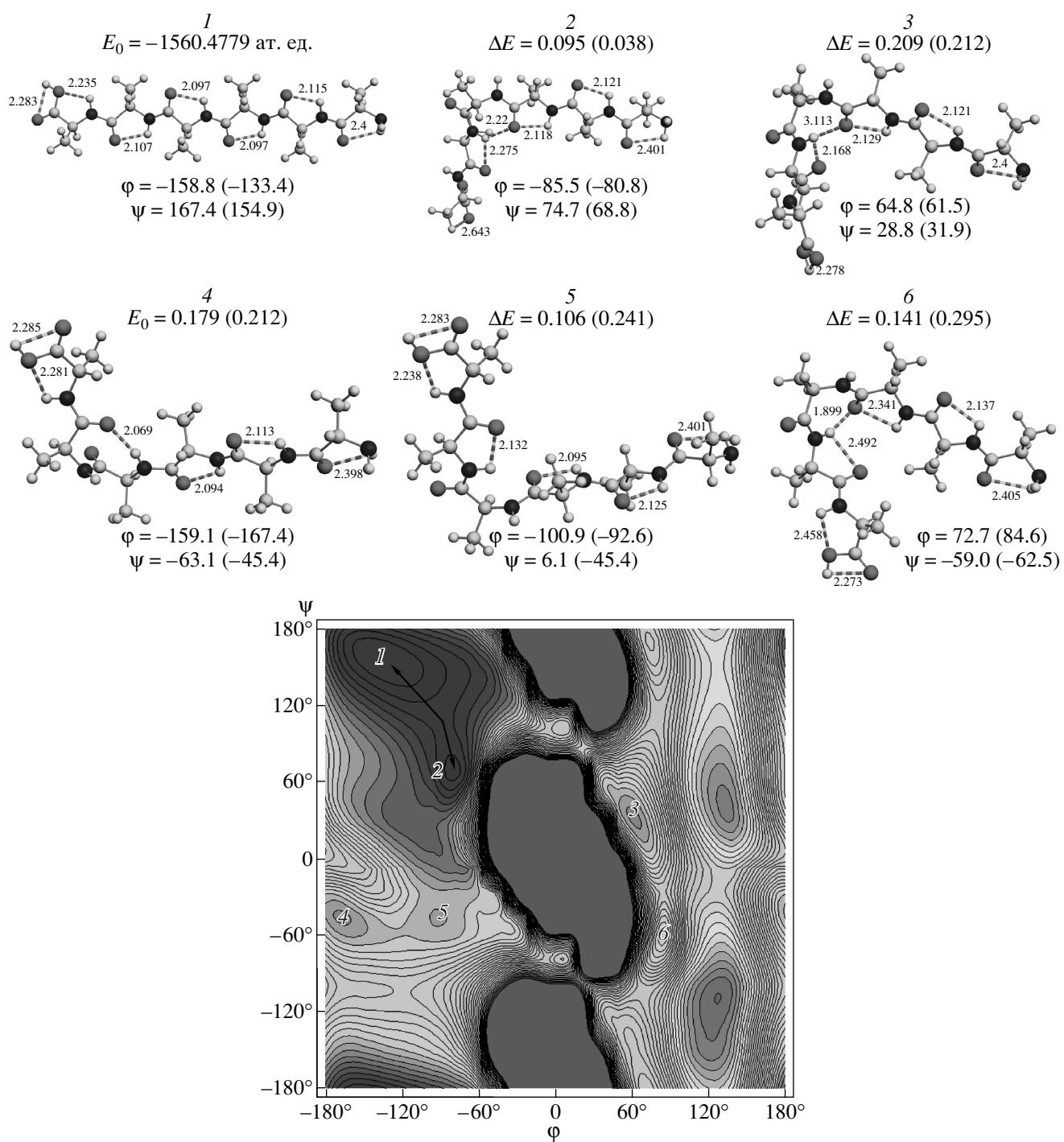
структурой листа и спирали соответственно в зависимости от двугранных углов  $\varphi$  и  $\psi$ . Конформации, соответствующие основным минимумам на энергетических поверхностях, приведены рядом с соответствующими потенциальными поверхностями.

Энергетические поверхности трипептида и гексапептида аланина со вторичной структурой листа очень похожи, что говорит о том, что конформации листа энергетическая поверхность формируется в основном за счет взаимодействий соседних аминокислот. Как и в случае трипептида, на поверхности потенциальной энергии гексапептида аланина со вторичной структурой листа наблюдаются два основных минимума. Барьер перехода между ними показан стрелкой на рис. 4. Для более точного определения высоты барьера для данного перехода была проведена дополнительная оптимизация структуры полипептида по всем степеням свободы, кроме двугранных углов  $\varphi$  и  $\psi$ , отвечающих переменной сканирования. Расчет, выполненный методом B3LYP/6-31G(2d,p), показал, что для перехода  $1 \rightarrow 2$  высота барьера составляет 0.095 эВ, а для перехода  $2 \rightarrow 1$  она равна 0.023 эВ. Частота колебаний гексапептида для рассматриваемой координаты, рассчитанная методом B3LYP/STO-3G, составила  $6.24 \text{ см}^{-1}$ . В результате при комнатной температуре времена переходов составляют

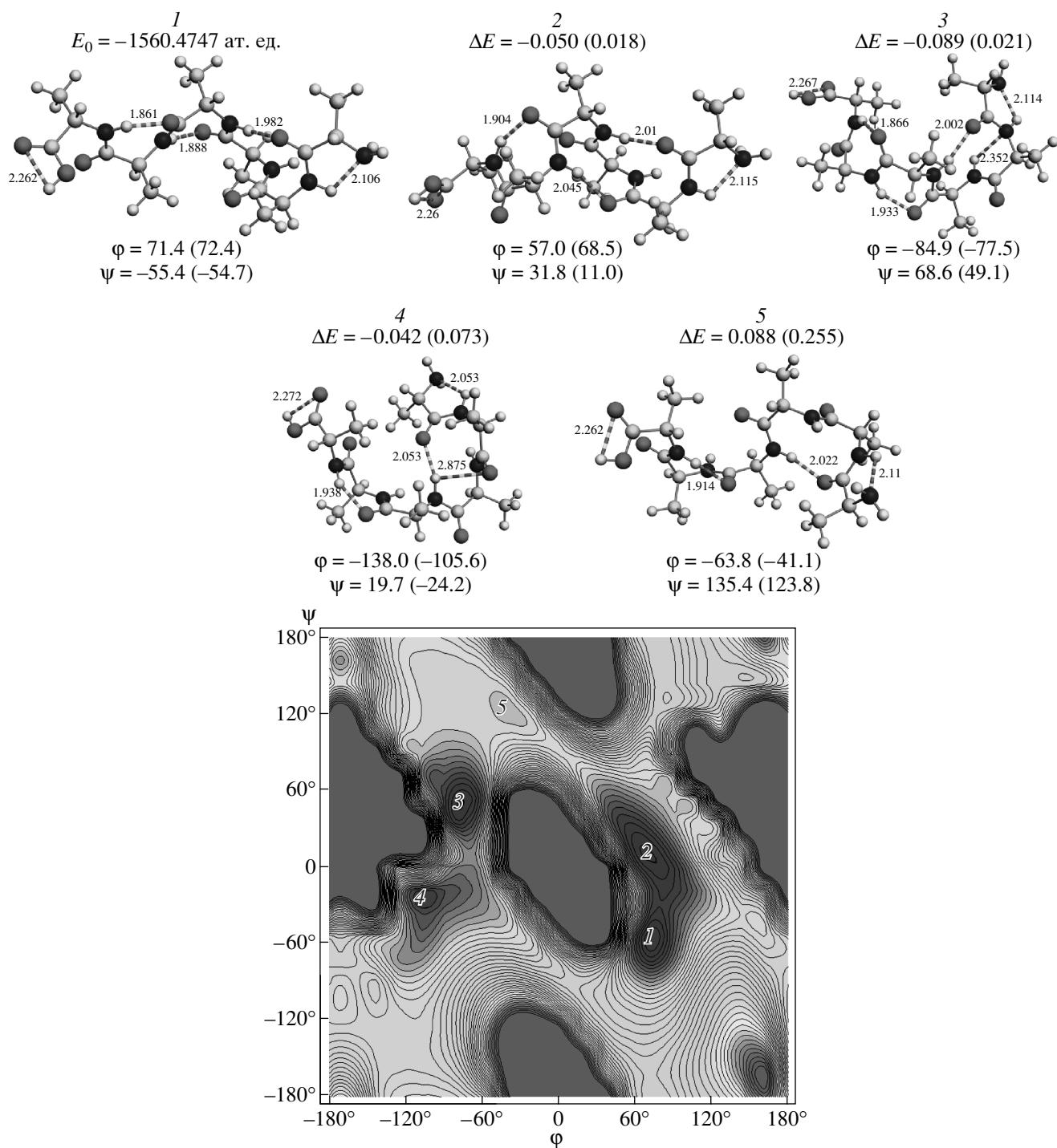
$$\tau_{6 \times Ala}^{1 \rightarrow 2} \approx 211 \text{ пс}, \quad \tau_{6 \times Ala}^{2 \rightarrow 1} \approx 13 \text{ пс.}$$

Перейдем к рассмотрению гексапептида аланина со вторичной структурой спирали. Поверхность потенциальной энергии для этого полипептида приведена на рис. 5. На ней видны две центральные запрещенные области примерно при тех же значениях углов  $\varphi$  и  $\psi$ , что и на остальных энергетических поверхностях. В дополнение к двум центральным максимумам на поверхности гексапептида аланина со вторичной структурой спирали наблюдается максимум при  $\varphi \approx 180^\circ$  и  $\psi \approx 40^\circ$ . Возникновение этого максимума обусловлено сильным сближением боковых радикалов крайних аминокислот гексапептида.

Следует отметить, что с учетом релаксации всех степеней свободы молекулы значение углов  $\varphi$  и  $\psi$ , соответствующих некоторым минимумам на поверхностях потенциальной энергии для гексапептидов аланина, значительно меняется (см., например, состояния 1, 4, 5 на рис. 4 и состояния 2, 4 на рис. 5). Это фактически означает, что вблизи этих минимумов поверхность потенциальной энергии гексапептида аланина с учетом релаксации всех степеней свободы молекулы сильно отличается от поверхности потенциальной энергии, рассчитанной без уч-



**Рис. 4.** Поверхность потенциальной энергии гексапептида аланина со вторичной структурой листа, рассчитанная методом B3LYP/6-31G(2d,p). Шкала энергии приведена на рис. 2. Цифрами помечены минимумы на потенциальной поверхности. Изображения оптимизированных конформаций гексапептида аланина приведены рядом с потенциальной поверхностью. Значения углов  $\varphi$  и  $\psi$ , а также относительные энергии состояний приведены аналогично рис. 3



**Рис. 5.** Поверхность потенциальной энергии гексапептида аланина со вторичной структурой спирали, рассчитанная методом B3LYP/6-31G(2d,p). Шкала энергии приведена на рис. 2. Цифрами помечены минимумы на потенциальной поверхности. Изображения оптимизированных конформаций гексапептида аланина приведены рядом с потенциальной поверхностью. Значения углов  $\varphi$  и  $\psi$ , а также относительные энергии состояний приведены аналогично рис. 3

та релаксации. Расчет поверхности потенциальной энергии полипептида без учета релаксации является трудоемкой задачей в рамках теории функционала плотности. Так, например, для расчета поверхности потенциальной энергии гексапептида, рассмотренной в настоящей работе, потребовалось около 2000 ч компьютерного времени компьютера класса 2.4 ГГц, Pentium Xeon. Расчет потенциальной поверхности с учетом релаксации всех степеней свободы молекулы потребовал бы примерно в 20-30 раз больше компьютерного времени, что составляет приблизительно 5 лет работы компьютера упомянутого класса. Подобный расчет был бы неоправданно дорогим, поскольку поверхность потенциальной энергии, рассчитанная без учета релаксации, уже содержит весьма большой объем полезной информации. Так, с помощью этой поверхности можно эффективно определять различные состояния полипептида, которые могут в дальнейшем быть использованы как начальные конфигурации для оптимизации системы по всем степеням свободы.

### 3.5. Сравнение результатов расчета с экспериментальными данными

В настоящее время стало возможным экспериментальное определение пространственной структуры многих белков. Зная структуру белка, можно определить характеризующие ее углы  $\varphi$  и  $\psi$ .

На рис. 6а представлена карта запрещенных и разрешенных конформаций аланина (стерическая диаграмма Рамачандрана), взятая из работы [42]. Она получена из чисто геометрических соображений. При этом структура полипептида предполагается фиксированной, определяемой межатомными радиусами ван-дер-ваальсовского взаимодействия. В зависимости от величины расстояний между атомами на стерической диаграмме различают три области: полностью разрешенную, условно разрешенную и запрещенную. Конформация считается полностью разрешенной, если расстояние между атомами различных аминокислот превышает некоторое критическое значение  $r_{ij} \geq r_{max}$ . Условно разрешенными считаются области, в которых расстояния между отдельными атомами различных аминокислот находятся в интервале

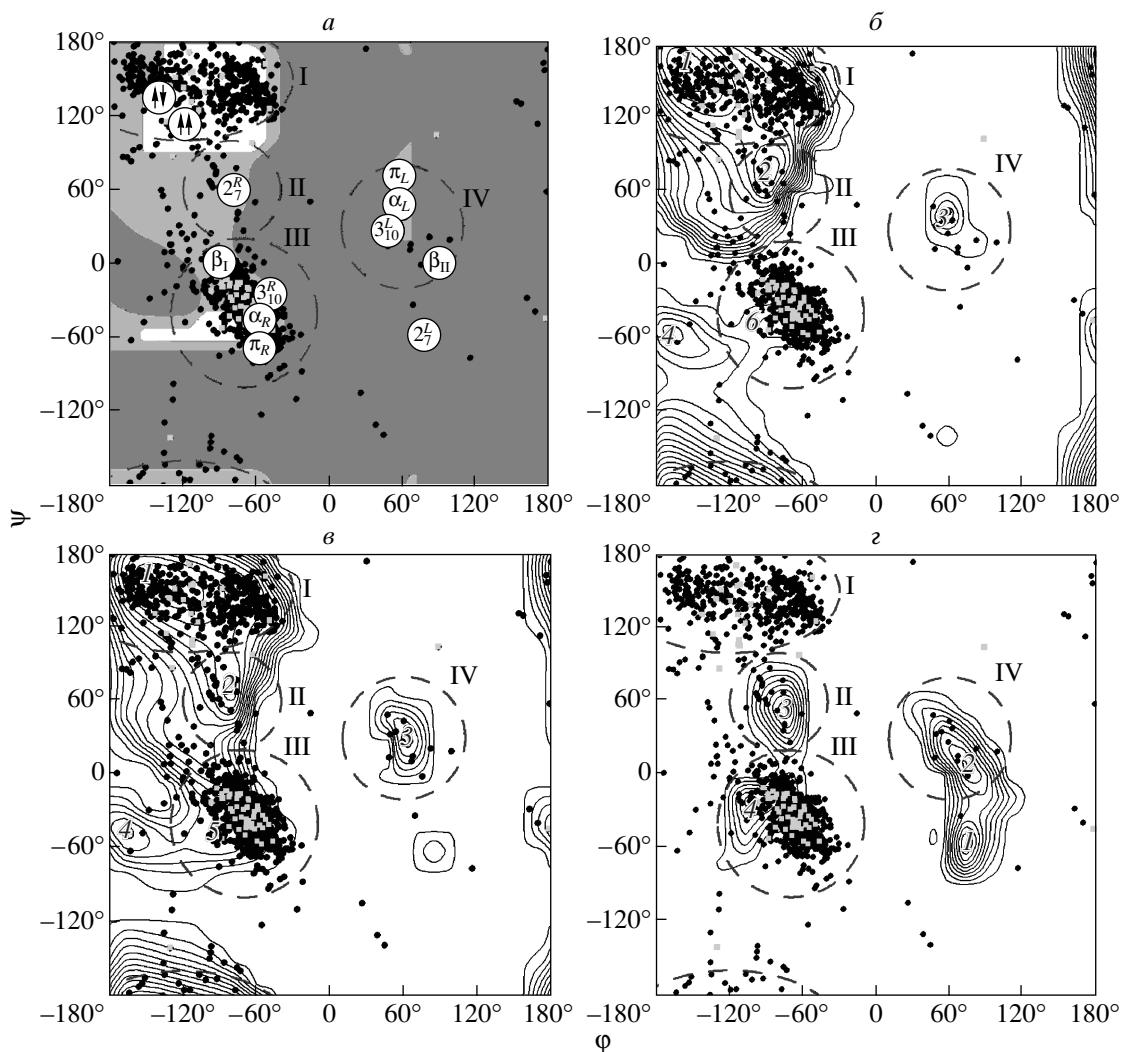
$$r_{min} \leq r_{ij} < r_{max}.$$

Все остальные состояния считаются запрещенными. Значения  $r_{min}$  и  $r_{max}$  определяются типами пар атомов и могут быть найдены во множестве учебников (см., например, [42]). На рис. 6а белым

цветом показаны полностью разрешенные области, светло-серым — условно разрешенные области, а темно-серым — запрещенные области. На этом рисунке точками отмечены углы, соответствующие основным геометриям аланина, периодическое повторение которых приводит к возникновению цепочек с определенной вторичной структурой. В случае  $\beta$ -изгиба эти углы соответствуют углам в третьем аланине, если считать от начала  $\beta$ -изгиба. Так,  $2_7^R, 2_7^L$  — правая и левая спираль  $2_7$ ;  $3_{10}^R, 3_{10}^L$  — правая и левая спираль  $3_{10}$ ;  $\alpha_R, \alpha_L$  — правая и левая  $\alpha$ -спираль ( $4_{13}$ );  $\pi_R, \pi_L$  — правая и левая  $\pi$ -спираль ( $5_{16}$ );  $\uparrow\uparrow, \uparrow\downarrow$  — параллельный и антипараллельный  $\beta$ -лист;  $\beta_I, \beta_{II}$  соответствуют  $\beta$ -изгибам первого и второго рода.

Однако не все из перечисленных выше структур в равной мере присутствуют в аланиновых звеньях белков. На рис. 6а показано распределение значений углов  $\varphi$  и  $\psi$  аланиновых звеньев 50 белков, взятых случайным образом из базы данных для белков [9, 43]. В данной работе были взяты белки со следующими PDB-кодами: 1ALD, 1BRP, 1BTC, 1C1L, 1C1M, 1CCA, 1CCR, 1CMS, 1CTY, 1EST, 1GBT, 1HBG, 1LEC, 1LEH, 1LH1, 1LTE, 1MBA, 1MBS, 1MPP, 1MUP, 1POC, 1RCB, 1SHG, 1YEB, 2AAA, 2APR, 2C2C, 2CNA, 2IMM, 2LHB, 2MHR, 2RHE, 2STV, 351C, 3APP, 3REP, 1GKU, 1AA2, 1G8X, 1AFA, 2SCQ, 1CX2, 1CFB, 1A00, 1MT5, 1IL8, 1ICW, 1GOS, 1PRH, 1C1G. Видно, что можно выделить четыре основные области, в пределах которых распределены почти все экспериментальные точки. На рис. 6 эти области схематически обозначены штриховыми эллипсами. Следует подчеркнуть, что подобное разбиение носит исключительно иллюстративный характер и служит для облегчения понимания экспериментально наблюдаемых данных и упрощения дальнейшего обсуждения. Каждая из областей соответствует различным типам вторичной структуры поли-аланина. Так, область I соответствует параллельным и антипараллельным  $\beta$ -листам. Область II — правой спирали  $2_7^R$ . Область III — правой  $\alpha_R$ -спирали, правой спирали  $3_{10}^R$ , правой  $\pi_R$ -спирали и  $\beta$ -изгибу первого рода. Область IV соответствует левой  $\alpha_L$ -спирали, левой спирали  $3_{10}^L$ , левой  $\pi_L$ -спирали и  $\beta$ -изгибу второго рода. Видно, что в некоторых случаях внутри одной области присутствуют несколько вторичных структур, и однозначно сопоставить точки какому-либо типу вторичной структуры на основе имеющихся экспериментальных данных нельзя.

Перейдем теперь к сравнению экспериментально наблюденных значений углов  $\varphi$  и  $\psi$  в белках с полу-



**Рис. 6.** Сравнение значений углов  $\varphi$  и  $\psi$ , соответствующих аланиновым звеньям 50 белков, взятых случайным образом из базы данных для белков [9, 43], со стерической диаграммой поли-аланина [42] (*a*). Сравнение значений углов  $\varphi$  и  $\psi$  аланиновых звеньев ряда белков, взятых из базы данных для белков [9, 43], с топологией рассчитанных поверхностей потенциальной энергии для трипептида аланина (*б*), гексапептида аланина в конформации листа (*в*), гексапептида аланина в конформации спирали (*г*). Квадраты соответствуют аланинам в окружении аланинов, кружки соответствуют аланинам в окружении других аминокислот. Штриховыми эллипсами отмечены области повышенной концентрации наблюдаемых углов

жением минимумов на поверхностях потенциальной энергии, рассчитанных для полипептидов. Отметим, что, насколько нам известно, подобного рода сравнение ранее не проводилось. Целью этого сравнения является установление соответствия рассчитанных углов  $\varphi$  и  $\psi$ , соответствующих различным устойчивым состояниям полипептидных цепочек, углам  $\varphi$  и  $\psi$ , наблюдаемым в эксперименте, а также сопоставление вторичной структуры рассчитанных конформаций полипептидов вторичным структурам Рамачандрана.

Область I можно сопоставить минимумам 1 на поверхностях потенциальной энергии трипептида аланина (рис. 6*б*) и гексапептида аланина со вторичной структурой листа (рис. 6*в*). Эти состояния в точности соответствуют цепочкам аланина в состоянии  $\beta$ -листа (см. рис. 3 и 4). На поверхности потенциальной энергии для гексапептида аланина со вторичной структурой спирали (рис. 6*г*) минимум в этой области потенциальной поверхности не возникает.

Область II соответствует минимумам 2 на поверхностях на рис. 6б, 6в, а также минимуму 3 на поверхности потенциальной энергии на рис. 6г. На стерической диаграмме поли-аланина область II соответствует правой спирали  $2_7^R$ . Структура состояний 2 на поверхностях на рис. 6б, в отличается от структуры этой спирали. Фрагменту спирали  $2_7^R$  соответствуют лишь центральные аланины, которым соответствуют углы  $\varphi$  и  $\psi$  на рис. 6б, в. Поэтому можно судить о состояниях 2 как о смешанных состояниях, в которых центральная часть имеет конформацию спирали, а крайние — конформацию листа. В реальных белках полноценная спираль  $2_7^R$  образуется крайне редко, поскольку соответствующая ей область на стерической диаграмме лежит на самом краю запрещенной области, а также угол схождения групп N–H и C=O практически прямой, что невыгодно для образования водородной связи. Состояние 3 на поверхности на рис. 6г тоже является смешанным состоянием. Здесь можно выделить один виток спирали  $3_{10}^R$  и два витка спирали  $2_7^R$  с характерными водородными связями между атомами (см. рис. 5).

Области III соответствуют структура правой  $\alpha_R$ -спирале, правой спирали  $3_{10}^R$ , правой  $\pi_R$ -спирале и  $\beta$ -изгиба I типа. Область III можно сопоставить минимумам 6, 5 и 4 на поверхностях соответственно на рис. 6б, в, г. Состояние 6 на поверхности потенциальной энергии трипептида аланина не может соответствовать ни одной вторичной структуре, характерной для этой области на потенциальной поверхности, поскольку полипептид слишком короткий. В этой области потенциальной поверхности наиболее вероятными для аланина являются структуры правой  $\alpha_R$ -спирали и  $\beta$ -изгиба. Однако для формирования одного витка такой спирали (или для формирования изгиба) необходимо минимум четыре аминокислоты. Состояние 5 на потенциальной поверхности гексапептида аланина на рис. 6в можно охарактеризовать как несформировавшийся  $\beta$ -изгиб, поскольку аланин, в котором меняются двугранные углы, принимает состояние  $\beta$ -изгиба, а соседний с ним аланин находится в состоянии листа (см. рис. 4). Состояние 4 на потенциальной поверхности гексапептида аланина на рис. 6г заметно изменяется при учете релаксации молекулы по всем степеням свободы и попадает за рамки области III. По характерным водородным связям в состоянии 4 можно выделить фрагмент спирали  $2_7^R$  и фрагмент спирали  $3_{10}^R$ . Тот факт, что точка, соответствующая минимуму 4, лежит за пределами областей II и III, объясняется тем, что углы  $\varphi$  и  $\psi$ , представленные на рис. 6г, соответствуют аминокислоте на стыке фрагментов двух

спиралей. Это приводит к значительным искажениям структуры аминокислоты.

Области IV соответствуют структуры левой  $\alpha_L$ -спирали, левой спирали  $3_{10}^L$ , левой  $\pi_L$ -спирали и  $\beta$ -изгиба типа II. Фрагменты с такими типами вторичной структуры возникают в аланиновых звеньях белков крайне редко. Отметим тот факт, что для формирования всех этих структур необходимо как минимум четыре аминокислоты, поэтому возникающий на потенциальной поверхности для трипептида аланина минимум 3 не может быть отнесен к той или иной вторичной структуре. На поверхностях потенциальной энергии для гексапептидов области IV можно сопоставить состояния 3 и 2 на поверхностях на рис. 6в и г, соответственно. Состояние 3 на поверхности на рис. 6в также соответствует несформировавшемуся  $\beta$ -изгибу, поскольку аланин, в котором меняются двугранные углы, принимает состояние  $\beta$ -изгиба, а соседний с ним аланин находится в состоянии листа (см. рис. 4). Состояние 2 на поверхности на рис. 6г лежит за пределами области IV, однако учет релаксации молекулы по всем степеням свободы заметно сдвигает минимум по поверхности, и он попадает в разрешенную зону для левой  $\alpha_L$ -спирали и левой спирали  $3_{10}^L$  (см. рис. 5). Состояние 2 соответствует структуре деформированной левой спирали  $3_{10}^L$  (см. рис. 5). Здесь видно два витка этой спирали. Основные отличия вызваны недостаточной длиной полипептида для формирования правильной спиралевидной структуры.

Также следует отметить, что состояние 1 на поверхности на рис. 6г совпадает со структурой левой спирали  $2_7^L$  на стерической диаграмме на рис. 6а, хотя и является смешанным (здесь присутствуют два витка левой спирали  $2_7^L$  и один виток левой спирали  $3_{10}^L$ ). Это совпадение вызвано тем, что углы  $\varphi$  и  $\psi$  на рис. 6г определены для аланина из спирали  $2_7^L$ .

#### 4. ЗАКЛЮЧЕНИЕ

В данной работе исследованы многомерные поверхности потенциальной энергии цепочек аминокислот, состоящих из трех и шести аланинов. Работа посвящена изучению конформационных свойств этих систем по отношению к степеням свободы, связанным с кручением молекул относительно полипептидной цепи. Расчет был выполнен из первых принципов в рамках теории функционала плотности с учетом всех электронов в системе. Насколько известно авторам, для рассматриваемых молекул такого рода исследование ранее не проводилось.

Анализ энергетических поверхностей позволил провести оценки времен переходов между различными устойчивыми конформациями полипептидов. Проведено сравнение полученных результатов с имеющимися результатами молекулярно-динамических расчетов. Продемонстрирована сильная асимметрия переходов между соседними устойчивыми состояниями на поверхности потенциальной энергии. Установлено, что времена переходов лежат в пикосекундном диапазоне.

Впервые было проведено сравнение рассчитанных минимумов на поверхностях потенциальной энергии с экспериментально наблюдаемыми в реальных белках распределениями углов  $\varphi$  и  $\psi$ . Это сравнение показало, что все основные минимумы на поверхностях потенциальной энергии, которые обсуждаются в данной работе, могут быть поставлены в соответствие наиболее часто встречающимся в белках значениям углов  $\varphi$  и  $\psi$ , полученным экспериментально. В работе также исследовано влияние вторичной структуры полипептидной цепи на ее конформационные свойства относительно вращения. Так, для цепочки из шести аминокислот со структурой спираль и листа продемонстрировано влияние вторичной структуры на формирование устойчивых изомерных состояний молекулы.

Результаты данной работы могут служить основой для изучения более сложных систем. Например, используя предложенную модель для оценки времени перехода, можно провести аналогичные вычисления для полипептидов, состоящих из различных аминокислот, и для белков и тем самым оценить время фолдинга таких систем. Полученные результаты можно также использовать для установления области применимости различных модельных расчетов, требующих значительно меньших затрат компьютерного времени и мощности, чем расчет из первых принципов.

Авторы выражают благодарность Studienstiftung des deutschen Volkes, INTAS (грант № 03-51-6170) и РФФИ (грант № 06-02-17227-а) за поддержку данной работы, а также Frankfurt Center for Scientific Computing за предоставление необходимого количества компьютерного времени.

## ЛИТЕРАТУРА

1. *Atomic Clusters and Nanoparticles, NATO Advanced Study Institute, Les Houches Session LXXIII*, Les Houches (2000), ed. by C. Guet, P. Hobza, F. Spiegelman, and F. David, EDP Sciences and Springer-Verlag, Berlin (2001).

2. *Latest Advances in Atomic Cluster Collisions Fission, Fusion, Electron, Ion and Photon Impact*, ed. by A. Solo'yov and J.-P. Connerade, World Scientific Press (2004).
3. M. Karas and F. Hillenkamp, *Anal. Chem.* **60**, 2299 (1988).
4. F. Hillenkamp and M. Karas, *Int. J. Mass Spectr.* **200**, 71 (2000).
5. M. Karas, U. Bahr, I. Fournier, M. Gluckmann, and A. Pfenninger, *J. Mass Spectr.* **226**, 239 (2003).
6. M. Wind and W. Lehmann, *J. Anal. At. Spectr.* **19**, 20 (2004).
7. A. Finkelstein and O. Ptizin, *Physics of Proteins*, Moscow University Press, Moscow (2002).
8. A. Mülberg, *Protein Folding*, St. Petersburg University Press (2004).
9. H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, and P. Bourne, *Nucleic Acids Res.* **28**, 235 (2000).
10. A. Rubin, *Biophysics: Theoretical Biophysics*, Moscow University Press, Moscow (2004).
11. T. Head-Gordon, M. Head-Gordon, M. Frisch, C. B. III, and J. Pople, *J. Amer. Chem. Soc.* **113**, 5989 (1991).
12. I. Gould, W. Cornell, and I. Hillier, *J. Amer. Chem. Soc.* **116**, 9250 (1994).
13. Z. Wang and Y. Duan, *J. Comp. Chem.* **25**, 1699 (2004).
14. A. Percel, O. Farkas, I. Jákli, I. Topol, and I. Csizmadia, *J. Comp. Chem.* **24**, 1026 (2003).
15. I. Húdaky, P. Húdaky, and A. Percel, *J. Comp. Chem.* **25**, 1522 (2004).
16. R. Improta and V. Barone, *J. Comp. Chem.* **25**, 1533 (2004).
17. R. Vargas, J. Garza, B. Hay, and D. Dixon, *J. Phys. Chem. A* **106**, 3213 (2002).
18. R. Kashner and D. Hohl, *J. Phys. Chem. A* **102**, 5111 (1998).
19. D. Wei, H. Guo, and D. Salahub, *Phys. Rev. E* **64**, (2001).
20. X. Wu and S. Sung, *Proteins: Structure, Function and Genetics* **34**, 295 (1999).

21. P. Pliego-Pastrana and M. D. Carbajal-Tinoco, Phys. Rev. E **68**, 011903 (2003).
22. S. Woutersen, Y. Mu, G. Stock, and P. Hamm, Chem. Phys. **266**, (2001).
23. S. Woutersen, R. Pfister, Y. Mu, D. Kosov, and G. Stock, J. Chem. Phys. **117**, 6833 (2002).
24. Y. Mu and G. Stock, J. Chem. Phys. B **106**, 5294 (2002).
25. Y. Mu, D. Kosov, and G. Stock, J. Chem. Phys. B **107**, 5064 (2003).
26. P. Nguyen and G. Stock, J. Chem. Phys. **119**, 11350 (2003).
27. H. Torii and M. Tasumi, J. Raman. Spectr. **29**, 81 (1998).
28. S. Woutersen and P. Hamm, J. Chem. Phys. B **104**, 11316 (2000).
29. S. Woutersen and P. Hamm, J. Chem. Phys. **114**, 2727 (2001).
30. R. Schweitzer-Stenner, F. Eker, Q. Huang, and K. Griebenow, J. Amer. Chem. Soc. **123**, (2000).
31. Y. Levy and O. Becker, J. Chem. Phys. **114**, 993 (2001).
32. Z. Shi, C. Olson, G. Rose, R. Baldwin, and N. Kallenbach, PNAS **99**, 9190 (2002).
33. A. Garsia, Polymer **45**, 669 (2004).
34. A. Yakubovitch, I. Solov'yov, A. Solov'yov, and W. Greiner, E-print archives, physics/0406093.
35. A. Yakubovitch, I. Solov'yov, A. Solov'yov, and W. Greiner, E-print archives, physics/0406094.
36. L. Lindgren and J. Morrison, *Atomic Many-Body Theory*, Springer-Verlag, New York, Heidelberg, Berlin (1986).
37. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
38. A. Becke, Phys. Rev. **38**, 3098 (1988).
39. C. Lee, W. Yang, and R. Parr, Phys. Rev. B **37**, 785 (1988).
40. R. Parr and W. Yang, *Density-Functional of Atoms and Molecules*, Oxford University Press, Oxford, New York (1989).
41. A. Bax, Prot. Sci. **12**, 1 (2003).
42. D. Voet and J. Voet, *Biochemistry*, John Wiley and Sons, Inc., USA (2004).
43. S. Sheik, P. Sundararajan, A. Hussain, and K. Sekar, Bioinformatics **18**, 1548 (2002).