

УДК 519.248.23:541.64:539.199

СУЩЕСТВЕННАЯ ВЫБОРКА ПРИ МОДЕЛИРОВАНИИ НЕПРЕРЫВНОГО СПЕКТРА КОНФОРМАЦИЙ МАКРОМОЛЕКУЛ МЕТОДОМ МОНТЕ-КАРЛО

А. Л. Рабинович¹, Д. В. Журкин²

¹ *Институт биологии Карельского научного центра РАН,*

² *Петрозаводский государственный университет*

Описан алгоритм, разработанный для моделирования молекул цепного строения методом Монте-Карло в предположении о непрерывном спектре конформаций, с использованием процедуры существенной выборки. Алгоритм учитывает взаимозависимость каждых трех углов внутреннего вращения вдоль по цепи. Он использован для изучения равновесных свойств углеводородных цепей различного строения.

Ключевые слова: метод Монте-Карло, существенная выборка, цепные молекулы, непрерывный спектр конформаций.

A. L. Rabinovich, D. V. Zhurkin. AN IMPORTANCE SAMPLING FOR MONTE CARLO SIMULATION OF A CONTINUOUS SPECTRUM OF CONFORMATIONS OF MACROMOLECULES

An importance sampling technique for Monte Carlo simulation of chain molecules assuming a continuous spectrum of their conformations is described. The algorithm takes into account the interdependence of each three torsion angles along the chain. It is used to study the equilibrium properties of hydrocarbon chains of different structure.

Key words: Monte Carlo simulation, importance sampling, chain molecules, continuous spectrum of conformations.

ВВЕДЕНИЕ

Для изучения свойств молекулярных систем широко используются методы компьютерного моделирования [10]. Наиболее распространенными являются методы Монте-Карло (МК) и молекулярной динамики (МД). Первый из них позволяет получить средние значения макроскопических наблюдаемых величин по статистическому ансамблю, а второй – средние по времени. Однако используемые в методе МК выборки конечны, а доступные

в методе МД времена моделирования сравнительно невелики, поэтому важно проводить сравнение между собой данных, полученных в рамках различных подходов. Актуальной задачей является также совершенствование методов компьютерного моделирования. В настоящей работе описана специальная вычислительная схема метода МК, разработанная для генерирования конформаций цепных молекул, в предположении о непрерывном

изменении углов внутреннего вращения в полном диапазоне от 0 до 2π . Учтена взаимозависимость каждых трех углов внутреннего вращения вдоль по цепи. При генерировании использована существенная выборка.

СРЕДНИЕ ЗНАЧЕНИЯ

Среднее значение $\langle H \rangle$ любой макроскопической наблюдаемой величины $H(x_1, x_2, \dots, x_n)$ системы частиц (атомов данной молекулы) в каноническом ансамбле можно вычислить следующим образом [1, 11]:

$$\langle H \rangle = \frac{\int \int \dots \int_{\Omega} H(x_1, x_2, \dots, x_n) \cdot \exp[-U(x_1, x_2, \dots, x_n)/k_B T] dx_1 dx_2 \dots dx_n}{\int \int \dots \int_{\Omega} \exp[-U(x_1, x_2, \dots, x_n)/k_B T] dx_1 dx_2 \dots dx_n}. \quad (1)$$

Здесь (x_1, x_2, \dots, x_n) – точка в фазовом пространстве Ω этой системы, k_B – постоянная Больцмана, T – температура, $U(x_1, x_2, \dots, x_n)$ – гамильтониан системы, в котором опущены члены кинетической энергии, поскольку декартовы координаты и импульсы частиц входят в гамильтониан (в потенциальную и кинетическую энергию) по отдельности; переменные разделяются, интеграл по всем импульсам берется тривиальным образом и его можно сократить в числителе и знаменателе. В итоге под фазовым пространством Ω подразумевается конфигурационная его часть, а под $U(x_1, x_2, \dots, x_n)$ – потенциальная энергия, n – размерность пространства Ω . Если количество атомов в системе n_a , то n равно количеству декартовых координат ($3n_a$). Для молекул цепного строения, макромолекул конфигурационное пространство многомерно и функция $U(x_1, x_2, \dots, x_n)$, как правило, очень сложна и в аналитическом виде не представлена. Поэтому провести интегрирование в выражении (1) и выразить $\langle H \rangle$ в конечной аналитической форме в большинстве случаев не удастся. Но можно использовать один из численных методов решения, – например, метод МК; литература о нем и его приложениях весьма обширна (см., напр., [1–3, 5–7, 10–12, 14, 15, 19, 25, 26]).

Метод Монте-Карло

В методе МК задачу сводят к моделированию случайных величин и статистической оценке их характеристик. В простейшей формулировке это означает следующее. Для того чтобы приближенно вычислить некоторую скалярную величину a , предлагается придумать такую случайную величину ξ , у которой существует математическое ожидание $M\xi$, равное a . Тогда для оценки величины a выбирается K независимых реализаций $\xi_1, \xi_2, \dots, \xi_K$ случайной величины ξ и вычисляется среднее арифметическое $\bar{\xi}_K = (1/K) \cdot \sum_{i=1}^K \xi_i$. Последовательность одинаково распределенных независимых случайных величин, у которых существуют математические ожидания, под-

чиняется закону больших чисел (см., напр., [4, 6, 12], теорема Хинчина), поэтому среднее арифметическое этих величин сходится по вероятности к математическому ожиданию, т. е. при $K \rightarrow \infty$ имеет место $\bar{\xi}_K \xrightarrow{P} a$. При больших K величина $\bar{\xi}_K \approx a$, и вместо a можно использовать [6, 12] оценку $\bar{\xi}_K = (1/K) \cdot \sum_{i=1}^K \xi_i$. По определению, последовательность случайных величин $\bar{\xi}_1, \bar{\xi}_2, \dots, \bar{\xi}_K, \dots$ сходится по вероятности [12] к постоянной a , если для любого $h > 0$ вероятность $P\{|\bar{\xi}_K - a| \geq h\} \rightarrow 0$ при $K \rightarrow \infty$ (что равносильно $P\{|\bar{\xi}_K - a| < h\} \rightarrow 1$ при $K \rightarrow \infty$). Здесь $P\{|\bar{\xi}_K - a| \geq h\}$ есть вероятность того, что $\bar{\xi}_K$ удалена от a не менее чем на h . Смысл сходимости по вероятности в том, что вероятность значительного отклонения величины $\bar{\xi}_K$ от величины a при большом K небольшая; тем не менее, сильно отклоняющиеся значения $\bar{\xi}_K$ могут находиться сколь угодно далеко по K , но этих отклонений не очень много, вследствие чего вероятность их появления становится все меньше с ростом K .

Пусть $D\xi$ – (конечная) дисперсия случайной величины ξ . Можно использовать понятие *вероятной ошибки* [12] $r_K = 0,6745 \cdot (D\xi/K)^{1/2}$, для которой $P\{|\bar{\xi}_K - a| < r_K\} \approx P\{|\bar{\xi}_K - a| > r_K\} \approx 1/2$, т. е. одинаково вероятны ошибки, большие чем r_K , и ошибки, меньшие чем r_K , и охарактеризовать *порядок* ошибки метода. Вероятная ошибка $r_K \sim K^{-1/2}$. Критерий вероятной ошибки r_K полезен для сравнения степени "эффективности" [12] (или выгоды, удачности) разных алгоритмов, которые использованы (или можно использовать) для решения одной и той же задачи. Действительно, из выражения для r_K следует, что количества K осредняемых значений случайной величины $K = D\xi \cdot (0,6745/r_K)^2$. Пусть есть два алгоритма. Если задать вероятные ошибки r_K для обоих одинаковыми, то $K' = D\xi' \cdot (0,6745/r_K)^2$, $K'' = D\xi'' \cdot (0,6745/r_K)^2$. Пусть t' и t'' – время, затрачиваемое на расчет одного значения случайных величин ξ' и ξ'' , соответственно. Общее время расчета равно $t' \cdot K' = t' \cdot D\xi' \cdot (0,6745/r_K)^2$

и $t'' \cdot K'' = t'' \cdot D\xi'' \cdot (0,6745/r_K)^2$. Чем меньше время расчета ($t' \cdot K'$ или $t'' \cdot K''$), которое необходимо для достижения заданной вероятной ошибки, при прочих равных условиях, тем более эффективным является алгоритм. А если ввести понятие "трудоемкости" [12] алгоритма, которое используется для сравнения алгоритмов и равно произведению дисперсии осредняемой случайной величины $D\xi$ на время t расчета одного значения ξ (т. е. $t' \cdot D\xi'$ или $t'' \cdot D\xi''$), то можно сказать, что из двух алгоритмов более эффективным является менее трудоемкий.

Случайных величин ξ , – таких, что $M\xi = a$, существует бесконечно много, и теория методов МК отвечает на вопросы о том [12], каким образом выбрать удобную величину ξ для решения данной задачи и как находить значения ξ_1, ξ_2, \dots произвольной случайной величины ξ .

Вычисление интегралов

Представляется целесообразным кратко привести сведения о двух известных схемах применения метода МК для вычисления интегралов; они важны для последующего описания разработанного алгоритма.

Схема 1 ([12], с. 93). Пусть требуется приближенно вычислить интеграл $J = \int_G f(x, y) \cdot p(x, y) dx dy$, где G – произвольная область плоскости x, y , а $p(x, y)$ – некоторая плотность вероятностей: $\int_G p(x, y) dx dy = 1$. Для оценки J методом МК вводится случайная точка Q в области G с плотностью вероятностей $p(x, y)$ и скалярная случайная величина $F = f(x, y)$, – такая, что ее математическое ожидание $MF = J$, т. е. равно искомому значению интеграла. Тогда, если Q_1, \dots, Q_K – независимые реализации случайной точки Q , а $F_1 = f(Q_1), \dots, F_K = f(Q_K)$, то согласно приведенной выше схеме оценкой J будет величина [12] $\bar{F}_K = (1/K) \cdot \sum_{i=1}^K F_i$. При этом $\bar{F}_K \xrightarrow{P} J$.

Важно отметить [12], что если G – ограниченная область, то по этой же схеме можно получить оценку и интеграла вида $\int_G f(x, y) dx dy$. Действительно, пусть площадь области G есть S_G . Тогда плотность вероятностей случайной точки, равномерно распределенной в G , есть $1/S_G$. Если ввести $f_1(x, y) = S_G \cdot f(x, y)$ и обозначить $1/S_G \equiv p_1(x, y)$, то этот интеграл приводится к предыдущему виду: $\int_G f(x, y) dx dy = \int_G f_1(x, y) \cdot p_1(x, y) dx dy$.

Если искомую оценку получают усреднением по выборке значений функции с использованием плотности вероятностей случайной точки, распределенной в данной области равномерно (как в последнем примере), то такую

выборку в литературе по применению методов МК в физике (статистической, теоретической) принято называть "simple" sampling [15, 25], т. е. "простая" выборка [2] (иногда используется термин "прямая" выборка [14]).

Схема 2 ([12], с. 108). Пусть $J^0 = \int_G f(x, y) dx dy$, где G – ограниченная или неограниченная область, и $\int_G |f(x, y)| dx dy > 0$. В области G вводится плотность вероятности $p(x, y)$. Если $p(x, y) > 0$ в тех точках, в которых $f(x, y) \neq 0$, то $p(x, y)$ называется допустимой по отношению к $f(x, y)$; допустимая плотность $p(x, y)$ может обращаться в нуль только в тех точках, где $f(x, y) = 0$. Если же $p(x, y) > 0$ всюду в G , то она допустима по отношению к любым $f(x, y)$. Пусть G_0 – множество точек, в которых $f(x, y) = 0$, и пусть $G_+ = G - G_0$. Вводится функция $F^0(x, y)$, следующим образом: $F^0(x, y) = f(x, y)/p(x, y)$, если точка (x, y) принадлежит области G_+ , и $F^0(x, y) = 0$, если точка (x, y) принадлежит области G_0 . Здесь $p(x, y)$ – допустимая плотность. Для оценки J^0 методом МК, как и в *схеме 1*, вводится Q – случайная точка, определенная в области G с плотностью вероятностей $p(x, y)$. Математическое ожидание величины $F^0(Q)$ есть $MF^0(Q) = \int_G F^0(x, y) \cdot p(x, y) dx dy = \int_{G_+} f(x, y) dx dy = J^0$ ввиду того, что функция $f(x, y) = 0$ в точках (x, y) , не принадлежащих области G_+ . Итак, $MF^0(Q)$ равно искомому значению интеграла. Дальнейшие действия соответствуют *схеме 1*: если Q_1, \dots, Q_K – независимые реализации случайной точки Q , а $F_1^0 = f(Q_1)/p(Q_1), \dots, F_K^0 = f(Q_K)/p(Q_K)$, то оценкой J^0 будет величина $\bar{F}_K^0 = (1/K) \cdot \sum_{i=1}^K F_i^0$.

Очевидно, плотность $p(x, y)$ желательно выбрать так, чтобы минимизировать дисперсию $DF^0(Q)$. Доказано [12], что минимальная дисперсия $DF^0(Q)$ реализуется в случае, когда плотность $p(x, y)$ пропорциональна $|f(x, y)|$; при этом она равна $|f(x, y)| \cdot [\int_G |f(x, y)| dx dy]^{-1}$. Ясно, однако, что несмотря на привлекательность достижения минимальной дисперсии, строго следовать этой рекомендации нельзя, поскольку задача вычисления $\int_G |f(x, y)| dx dy$ по трудности эквивалентна исходной задаче вычисления искомого интеграла. Но можно сделать и другой вывод: для получения величин F^0 с небольшими дисперсиями *желательно* выбирать плотность $p(x, y)$ *по возможности* пропорциональную $|f(x, y)|$ [12], – тогда в тех частях области G , в которых $|f(x, y)|$ больше и вклад которых в J^0 более существен, будет выбираться больше случайных точек [3, 12]. Иными словами,

чем меньше меняется функция $f(x, y)/p(x, y)$ (чем ближе она к постоянной), тем меньше окажется дисперсия $DF^0(Q)$. В этом состоит основная ценность данного приема; он предложен в работах [22, 23] и широко применяется [1, 2, 7, 12, 14]. Такая выборка называется "importance" sampling [15, 25] или "существенная" выборка [3, 7, 12], выборка "по значимости" [2, 5], выборка "по важности" [1, 6, 7], "предпочтительная" выборка [14].

Общий подход к оценке средних

Обе схемы, описанные выше для двух переменных, обобщаются на случай многих переменных. Возвращаясь к исходному выражению (1) для среднего значения $\langle H \rangle$, перейдем от точек (x, y) плоскости к n -мерным точкам (x_1, x_2, \dots, x_n) конфигурационного пространства Ω . Интеграл в знаменателе выражения (1) есть так называемый

статистический интеграл Z (для дискретного случая – статистическая сумма [9]), $Z = \int \int \dots \int_{\Omega} \exp[-U(x_1, \dots, x_n)/k_B T] dx_1 \dots dx_n$, а величина $(1/Z) \cdot \exp[-U(x_1, \dots, x_n)/k_B T]$ есть плотность вероятности точки (x_1, \dots, x_n) . Для приближенного вычисления интегралов в выражении (1) можно применить как "простую", так и "существенную" выборку.

Реализуем "простую" выборку. n -мерная точка (x_1, \dots, x_n) является аналогом рассмотренной выше 2-мерной точки Q . Независимыми реализациями Q_i случайной точки Q будут точки (x_1^i, \dots, x_n^i) , где $i = 1, 2, \dots, K$. Пусть они распределены равномерно в Ω . Скалярной случайной величиной F (из схемы 1) является выражение $(1/Z) \cdot H(x_1, \dots, x_n) \cdot \exp[-U(x_1, \dots, x_n)/k_B T]$, и его математическое ожидание, согласно (1), равно $\langle H \rangle$. Оценка \bar{H}_K выражения (1) методом МК будет иметь вид [1, 2, 11]

$$\bar{H}_K = \frac{\sum_{i=1}^K H(x_1^i, x_2^i, \dots, x_n^i) \cdot \exp[-U(x_1^i, x_2^i, \dots, x_n^i)/k_B T]}{\sum_{i=1}^K \exp[-U(x_1^i, x_2^i, \dots, x_n^i)/k_B T]} \quad (2)$$

При этом $\bar{H}_K \xrightarrow{P} \langle H \rangle$. Отметим, что выражение (2), в силу специфики исходного выражения (1), имеет вид формулы усреднения в каноническом ансамбле для случая, когда переменные имеют только дискретный спектр. Однако в данном случае суммирование производится не по дискретному спектру (всем состояниям), а по случайным точкам непрерывного спектра, и K является не количеством состояний, а размером выборки. В выражении (2) множитель $\exp[-U(x_1^i, \dots, x_n^i)/k_B T]$ может изменяться на многие порядки и, следовательно, более предпочтительной была бы "существенная" выборка.

Вариант "существенной" выборки, который можно было бы назвать наиболее эффективным, выгодным, т. е. наименее трудоемким (в том смысле, который описан выше), очевидно, был бы достигнут в случае использования функции $(1/Z) \cdot \exp[-U(x_1, x_2, \dots, x_n)/k_B T]$ в качестве плотности распределения случайной точки в области Ω . Пусть точки $(x_1^i, x_2^i, \dots, x_n^i)$ выбираются в Ω не равномерно, как в выражении (2), а именно с такой плотностью. Тогда оценка выражения

(1) сводится к среднему арифметическому $\bar{H}_K = (1/K) \cdot \sum_{i=1}^K H(x_1^i, x_2^i, \dots, x_n^i)$. Несмотря на столь простой вид итоговой оценки, реализовать эту идею невозможно, поскольку по трудности задача предварительного вычисления этой плотности сопоставима с задачей вычисления исходного выражения (1): речь идет о необходимости предварительного вычисления потенциальной энергии системы частиц $U(x_1, x_2, \dots, x_n)$. Тогда нужно подобрать такую плотность $p(x_1, x_2, \dots, x_n)$, которая была бы допустимой по отношению к подынтегральному выражению в (1) и, насколько возможно, пропорциональной ему, но при этом могла быть рассчитана предварительно, до проведения расчета методом МК. В итоге можно добиться дисперсии, значительно меньшей, чем в алгоритме простой выборки, и более эффективного (менее трудоемкого) алгоритма. Пусть некоторую плотность $p(x_1, x_2, \dots, x_n)$, удовлетворяющую этим условиям, найти удалось. Повторяя действия, описанные в схеме 2 выше, получим следующую оценку [1, 2, 11, 14]:

$$\bar{H}_K = \frac{\sum_{i=1}^K H(x_1^i, x_2^i, \dots, x_n^i) \cdot \exp[-U(x_1^i, x_2^i, \dots, x_n^i)/k_B T]/p(x_1^i, x_2^i, \dots, x_n^i)}{\sum_{i=1}^K \exp[-U(x_1^i, x_2^i, \dots, x_n^i)/k_B T]/p(x_1^i, x_2^i, \dots, x_n^i)} \quad (3)$$

Различие между (3) и (2) обусловлено именно способом выбора точек в конфигурационном пространстве.

Если выбрать $p(x_1, x_2, \dots, x_n)$ в (3) пропорционально $\exp[-U(x_1, x_2, \dots, x_n)/k_B T]$, т. е. заранее вычислить какую-то часть энергии $U(x_1, x_2, \dots, x_n)$, то получим искомый результат. Действительно, энергию U можно представить, например, в виде суммы двух слагаемых (назовем их энергией ближних U_{short} и энергией дальних U_{long} взаимодействий), $U = U_{short} + U_{long}$. Тогда $\exp[-U(x_1, x_2, \dots, x_n)/k_B T]$ будет произведением $\exp[-U_{short}(x_1, x_2, \dots, x_n)/k_B T]$ и $\exp[-U_{long}(x_1, x_2, \dots, x_n)/k_B T]$. Выделение U_{short} из общей энергии U следует осуществить таким образом, чтобы была возможность вычислить ее на предварительном этапе,

до осуществления процедуры генерирования конформаций молекул методом МК. Детали реализации этой процедуры будут описаны ниже. Если U_{short} вычислена, то в качестве плотности $p(x_1, x_2, \dots, x_n)$ распределения случайных точек можно использовать $(1/Z_{short}) \cdot \exp[-U_{short}(x_1, x_2, \dots, x_n)/k_B T]$, где $Z_{short} = \int \int \dots \int_{\Omega} \exp[-U_{short}(x_1, \dots, x_n)/k_B T] dx_1 \dots dx_n$. Ясно, что это выражение является допустимой плотностью по отношению к подынтегральной функции: оно ≥ 0 во всей области конфигурационного пространства, а если обращается в нуль при $U_{short} \rightarrow \infty$, то в этих же точках обращается в нуль и экспонента, в которую входит полная потенциальная энергия системы U , поскольку тогда $U \rightarrow \infty$. В итоге вместо (3) получим следующую оценку

$$\bar{H}_K = \frac{\sum_{i=1}^K H(x_1^i, x_2^i, \dots, x_n^i) \cdot \exp[-U_{long}(x_1^i, x_2^i, \dots, x_n^i)/k_B T]}{\sum_{i=1}^K \exp[-U_{long}(x_1^i, x_2^i, \dots, x_n^i)/k_B T]} \quad (4)$$

При компьютерной реализации эту плотность удается воспроизвести лишь приблизительно, и итоговые формулы (см. ниже) отличаются от (4). Тем не менее, выборка оказывается в итоге "существенной".

Отметим, что при изучении свойств макромолекул идея вычисления или аппроксимации части энергии $U(x_1, x_2, \dots, x_n)$ в методе МК использовалась многократно, предложено много вариантов для разных молекулярных систем, и в каждом из них учтена специфика этих систем и характер задачи; ряд вариантов упомянут, например, в обзоре [11]. В настоящей работе эта идея использована для разработки алгоритма, который можно применить для исследования свойств молекул цепного строения определенных классов, и в нем учтена специфика этих цепей. В частности, ниже алгоритм расчета реализован для углеводородных цепей вида $CH_3 - (CH_2)_f - (CH = CH - CH_2)_d - (CH_2)_g - CH_3$, где f, d, g – целые.

ОПИСАНИЕ ЦЕПНЫХ МОЛЕКУЛ

Приведем основные сведения о принятом описании цепных молекул. Конформации мак-

ромолекулы могут быть охарактеризованы либо декартовыми координатами всех атомов, либо внутренними переменными: совокупностями валентных связей $\{l\}$, валентных углов $\{\theta\}$, углов внутреннего вращения $\{\varphi\}$ (т. е. торсионных углов). В рамках классической статистической механики существуют две схемы приближений [20], – "классическая гибкая модель" (согласно которой макромолекула – это система материальных точек, взаимодействие которых описывается некоторыми потенциалами) и "классическая жесткая модель" (согласно которой потенциалы действуют лишь между определенными участками макромолекулы, остальные же участки скреплены между собой геометрическими связями).

"Гибкая" модель при изучении равновесных свойств считается более предпочтительной [18, 20] (несмотря на ряд нерешенных вопросов [18, 20]), и она позволяет для среднего значения $\langle H \rangle$ вместо выражения (1) в декартовых координатах записать выражение, в котором все величины зависят от внутренних переменных:

$$\langle H \rangle = \frac{\int \int \dots \int H(\{\varphi\}, \{\theta\}, \{l\}) \cdot \exp[-U(\{\varphi\}, \{\theta\}, \{l\})/k_B T] d\{\varphi\} d\{\theta\} d\{l\}}{\int \int \dots \int \exp[-U(\{\varphi\}, \{\theta\}, \{l\})/k_B T] d\{\varphi\} d\{\theta\} d\{l\}} \quad (5)$$

Здесь, как и в (1), H – некоторая макроскопическая наблюдаемая величина изучаемой системы частиц (молекулы), U – потенциальная энергия; $\{\varphi\}$, $\{\theta\}$ и $\{l\}$ – совокупности углов

внутреннего вращения, валентных углов и валентных связей, соответственно, которые образуют конфигурационное пространство.

Для получения оценки величины (5), аналогичной (3), методом МК можно использовать существенную выборку. Пусть случайные точки в конфигурационном пространстве есть $\{\varphi\}^\nu$, $\{\theta\}^\nu$ и $\{l\}^\nu$, а количество конформаций (состояний ν) молекулы в выборке есть

$$\bar{H}_\omega = \frac{\sum_{\nu=1}^{\omega} H(\{\varphi\}^\nu, \{\theta\}^\nu, \{l\}^\nu) \cdot \exp[-U(\{\varphi\}^\nu, \{\theta\}^\nu, \{l\}^\nu)/k_B T]/p(\{\varphi\}^\nu, \{\theta\}^\nu, \{l\}^\nu)}{\sum_{\nu=1}^{\omega} \exp[-U(\{\varphi\}^\nu, \{\theta\}^\nu, \{l\}^\nu)/k_B T]/p(\{\varphi\}^\nu, \{\theta\}^\nu, \{l\}^\nu)}. \quad (6)$$

Выражение (6) можно сделать более компактным, если убрать явное перечисление совокупностей случайных значений внутренних пере-

ω (для внутренних переменных обозначение ν использовано вместо i , а ω – вместо объема выборки K). Тогда, выбирая случайные точки с некоторой плотностью вероятностей $p(\{\varphi\}, \{\theta\}, \{l\})$, получим следующую оценку \bar{H}_ω :

$$\bar{H}_\omega = \frac{\sum_{\nu=1}^{\omega} H^\nu \cdot \exp[-U^\nu/k_B T]/p^\nu}{\sum_{\nu=1}^{\omega} \exp[-U^\nu/k_B T]/p^\nu}. \quad (7)$$

Силовое поле

Рассмотрим вопрос об общем подходе к расчету энергии цепных молекул. Как правило, в расчетах используется метод атом-атомных потенциальных функций. В литературе существуют наборы таких функций и апробированные параметры (силовые поля) для цепных молекул разных типов. В частности, для молекул липидов и их компонентов – углеводородных цепей рассматриваемого типа предложено поле CHARMM27 [16, 17, 24, 28]; в настоящей работе использован модифицированный вариант этого поля [21]. Энергия U системы частиц вычисляется в виде суммы энергий валентных связей $U_l = \sum_{n=1}^{N_l} K_{l_n} \cdot (l_n - l_{0_n})^2$ и уг-

лов $U_\theta = \sum_{n=1}^{N_\theta} K_{\theta_n} \cdot (\theta_n - \theta_{0_n})^2$, энергии Юри-

Брэдли $U_u = \sum_{n=1}^{N_u} K_{u_n} \cdot (u_n - u_{0_n})^2$, энергии

неплоских отклонений атомов $U_\eta = \sum_{n=1}^{N_\eta} K_{\eta_n} \cdot$

$(\eta_n - \eta_{0_n})^2$, торсионной энергии $U_\varphi = \sum_{n=1}^{N_\varphi} K_{\varphi_n} \cdot$

$[1 + \cos(p_n \cdot \varphi_n - \delta_n)]$, энергии невалентных взаимодействий $U_{nb} = \sum_{i,j} \varepsilon_{ij} \cdot [(R_{min_{ij}}/r_{ij})^{12} -$

$2 \cdot (R_{min_{ij}}/r_{ij})^6]$ и электростатической энергии $U_{el} = \sum_{i,j} [(q_i \cdot q_j)/(\varepsilon_{el} \cdot r_{ij})]$.

Здесь N_l , N_θ , N_u , N_η , N_φ – соответственно, количество валентных связей; валентных углов; различных пар атомов, разделенных двумя валентными связями; углов неплоских отклонений в молекуле; торсионных углов. K_{l_n} , K_{θ_n} , K_{u_n} , K_{η_n} , K_{φ_n} – силовые постоянные для

расчета перечисленных компонентов энергии, соответственно. Среди параметров слагаемых энергии используются текущие и равновесные значения: длины n -й валентной связи (l_n и l_{0_n}), n -го валентного угла (θ_n и θ_{0_n}), n -го расстояния между атомами, разделенными двумя валентными связями (u_n и u_{0_n}), n -го угла неплоских отклонений (η_n и η_{0_n}). φ_n – это значение n -го торсионного угла, а p_n , δ_n – множитель и сдвиг фаз; r_{ij} – расстояние между атомами с номерами i и j , разделенными более чем двумя валентными связями; ε_{ij} и $R_{min_{ij}}$ – глубина и положение соответствующего минимума энергии взаимодействия; q_i , q_j – парциальные заряды на атомах, ε_{el} – диэлектрическая постоянная.

Сравнение внутренних переменных

Переменные $\{\varphi\}$, $\{\theta\}$, $\{l\}$ не являются, в определенном смысле, равноправными между собой. Действительно, при данной температуре, в условиях, близких к состоянию равновесия молекулы, каждый угол внутреннего вращения из совокупности $\{\varphi\}$ может с высокой вероятностью изменяться во всем диапазоне $[0, 2\pi]$, тогда как изменения валентных связей $\{l\}$ и валентных углов $\{\theta\}$ происходят лишь в малых областях вокруг их равновесных значений (это так называемые "малые колебания"). Данное различие использовано при разработке настоящей вычислительной схемы метода МК: генерирование каждой конформации макромолекулы осуществляли в 2 этапа.

На 1-м этапе генерирования конформации ν все величины $\{l\}$ и величины $\{\theta\}$ (или их большинство) фиксировали при равновесных значениях, т. е. соответствующие компонен-

ты энергии обращали в нуль. Конформацию номер ν генерировали только по торсионным углам $\{\varphi\}$. Последние изменялись в диапазонах $[0, 2\pi]$, но если рассматриваемая молекула содержала двойные связи *cis* (вращения вокруг которых сопряжены с преодолением большого потенциального барьера), то, аналогично валентным углам, соответствующие углы вращения φ фиксировали в положении равновесия. При генерировании конформации ν по совокупности углов $\{\varphi\}$ был предложен вариант существенной выборки, детали которого будут описаны ниже. В отличие от углов внутреннего вращения, независимое варьирование в методе МК всех валентных углов осуществить невозможно, – например, углов, имеющих вершину на одном и том же атоме. Поэтому на 2-м этапе осуществляли переход в полученной конформации от внутренних переменных (валентных связей, валентных углов, торсионных углов) к декартовым координатам атомов и проводили варьирование всех координат, – независимо для каждого атома. При этом было учтено, что, поскольку валентные связи $\{l\}$ и валентные углы $\{\theta\}$ претерпевают лишь малые колебания вокруг положений равновесия, изменения декартовых координат атомов также

могут быть лишь малыми. Оценка величины диапазона варьирования декартовых координат проведена из условия, чтобы максимальные изменения значений энергии валентных углов U_θ и валентных связей U_l , отвечающих этому диапазону, в сумме не превышали изменений торсионной энергии U_φ цепи и энергии невалентных взаимодействий U_{nb} . Расчеты показали, что этому условию отвечает диапазон $\pm 0,02 \text{ \AA}$ вокруг равновесных положений всех атомов. Ввиду малости этого диапазона полагали, что в его пределах значения координат распределены равномерно; процедуру генерирования координат осуществляли с использованием генератора (датчика) псевдослучайных чисел [27], равномерно распределенных в интервале $[0,1)$.

Итак, валентные связи l и валентные углы θ сначала фиксировали при равновесных значениях. В выражении (5) можно явно указать поэтому только торсионные углы. Пусть N – количество атомов основной цепи молекулы, $\varphi_1, \varphi_2, \dots, \varphi_{N-1}$ – углы внутреннего вращения вокруг связей основной цепи, $U(\varphi_1, \varphi_2, \dots, \varphi_{N-1})$ – потенциальная (конформационная) энергия. Тогда вместо (5) будем иметь

$$\langle H \rangle = \frac{\int_0^{2\pi} \int_0^{2\pi} \dots \int_0^{2\pi} H(\varphi_1, \varphi_2, \dots, \varphi_{N-1}) \cdot \exp[-U(\varphi_1, \varphi_2, \dots, \varphi_{N-1})/k_B T] d\varphi_1 d\varphi_2 \dots d\varphi_{N-1}}{\int_0^{2\pi} \int_0^{2\pi} \dots \int_0^{2\pi} \exp[-U(\varphi_1, \varphi_2, \dots, \varphi_{N-1})/k_B T] d\varphi_1 d\varphi_2 \dots d\varphi_{N-1}}. \quad (8)$$

Энергия $U(\varphi_1, \varphi_2, \dots, \varphi_{N-1})$ при рассмотрении системы макромолекула – растворитель является энергией системы, состоящей из макромолекулы и окружающих молекул растворителя, усредненной (при любом данном состоянии $\varphi_1, \varphi_2, \dots, \varphi_{N-1}$ макромолекулы) по всем поло-

жениям и ориентациям молекул растворителя [18].

При использовании существенной выборки, выбирая случайные точки с некоторой допустимой плотностью $p(\varphi_1, \varphi_2, \dots, \varphi_{N-1})$, получим следующую оценку \bar{H}_ω методом МК:

$$\bar{H}_\omega = \frac{\sum_{\nu=1}^{\omega} H(\varphi_1^\nu, \varphi_2^\nu, \dots, \varphi_{N-1}^\nu) \cdot \exp[-U(\varphi_1^\nu, \varphi_2^\nu, \dots, \varphi_{N-1}^\nu)/k_B T] / p(\varphi_1^\nu, \varphi_2^\nu, \dots, \varphi_{N-1}^\nu)}{\sum_{\nu=1}^{\omega} \exp[-U(\varphi_1^\nu, \varphi_2^\nu, \dots, \varphi_{N-1}^\nu)/k_B T] / p(\varphi_1^\nu, \varphi_2^\nu, \dots, \varphi_{N-1}^\nu)}. \quad (9)$$

РАСЧЕТ ЭНЕРГИИ ЦЕПНОЙ МОЛЕКУЛЫ

Опишем более подробно процедуру расчета энергии U , а затем выбора допустимой плотности $p(\varphi_1, \varphi_2, \dots, \varphi_{N-1})$. Как упоминалось в разделе «Общий подход к оценке средних», потенциальная энергия всей молекулы U была

представлена суммой энергий ближних U_{short} и дальних U_{long} взаимодействий, $U = U_{short} + U_{long}$. Выделение U_{short} проведено так, чтобы ее можно было вычислить на предварительном этапе, до осуществления процедуры генерирования конформаций методом МК.

Исследуемая цепная молекула была условно разделена на небольшие фрагменты, – такие, что каждый из них содержал 3 последовательных угла внутреннего вращения вокруг связей основной цепи, $\varphi_\gamma, \varphi_{\gamma+1}, \varphi_{\gamma+2}$, варьирование которых приводило к изменению взаимных положений атомов. При конструировании цепи принято следующее правило: фрагменты соединяются не "встык", а с наложением друг на друга, – так, чтобы последовательные (соседние, смежные) фрагменты имели по два общих варьлируемых торсионных угла из трех. Этим учтена непрерывная взаимозависимость каждых трех углов вдоль по цепи.

Энергия U_{short} является суммой энергий U_{m_γ} этих фрагментов цепи, а энергия взаимодействий более удаленных вдоль по цепи атомов входит в U_{long} . В итоге $U_{short} = \sum_{\gamma=1}^{N-3} U_{m_\gamma}(\varphi_\gamma, \varphi_{\gamma+1}, \varphi_{\gamma+2})$. Здесь m_γ – индекс фрагмента: m – его идентификационный номер, который характеризует химическое строение фрагмента ($m = 1, 2, \dots, m_f$, где m_f – общее количество разных фрагментов, необходимых для конструирования избранных молекул по указанному правилу), а γ – его номер в последовательности фрагментов вдоль по цепи. Номера m для некоторых γ в индексах m_γ в сумме U_{short} могут быть одинаковыми. Если молекула – это полимерная цепь, то номера m будут циклически повторяться для одинаковых мономерных единиц.

Как упоминалось, алгоритм расчета реализован для углеводородных цепей вида $CH_3 - (CH_2)_f - (CH = CH - CH_2)_d - (CH_2)_g - CH_3$, где f, d, g – целые. Анализ показал, что любая из цепей такого вида может быть сконструирована должным выбором нескольких фрагментов из следующей совокупности (три связи, углы вращения вокруг которых варьируются, ниже выделены дефисом):

$CH_3 - CH_2 - CH_2 - CH_2CH_2$;
 $CH_3CH_2 - CH_2 - CH_2 - CH_2CH_2$;
 $CH_2CH_2 - CH_2 - CH_2 - CH_2CH_2$;
 $CH_2CH_2 - CH_2 - CH_2 - CH_2CH$;
 $CH_2CH_2 - CH_2 - CH_2 - CHCHCH_2$;
 $CH_2CH_2 - CH_2 - CHCH - CH_2CH_2$;
 $CH_3CH_2 - CH_2 - CH_2 - CH_2CH$;
 $CH_2CH_2 - CH_2 - CHCH - CH_2CH$;
 $CH_2CH_2 - CHCH - CH_2 - CHCHCH_2$;
 $CH_3 - CH_2 - CHCH - CH_2CH$;
 $CH_3CH_2 - CHCH - CH_2 - CHCHCH_2$;
 $CH_2CHCH - CH_2 - CHCH - CH_2CH$;
 $CH_2CHCH - CH_2 - CH_2 - CH_2CH_3$;
 $CHCH_2 - CH_2 - CH_2 - CH_3$;
 $CHCH_2 - CHCH - CH_2 - CH_2CH_3$;
 $CH_2CHCH - CH_2 - CH_2 - CH_3$.

Один из этих фрагментов представлен в виде схемы на рис. 1.

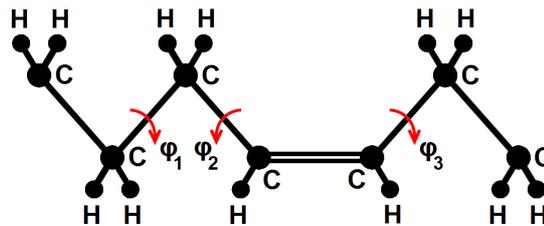


Рис. 1. Схема фрагмента $CH_2CH_2 - CH_2 - CHCH - CH_2CH_2$. Три варьлируемых торсионных угла ($\varphi_1, \varphi_2, \varphi_3$) обозначены стрелками

Весовые множители

Энергию $U_m(\varphi_1, \varphi_2, \varphi_3)$ каждого из 16 представленных выше фрагментов вычисляли (с учетом только компонентов U_φ, U_{nb} и U_{el}) таким образом, чтобы сумма энергий любой совокупности фрагментов оказалась равной энергии ближних взаимодействий U_{short} той цепи, которая составлена из этих фрагментов. Для этого при вычислении каждой U_m вводили для трех перечисленных компонентов энергии весовые множители (1/3, 1/2 или 1). Множители были согласованы между смежными фрагментами и исключали возможность того, что одни и те же слагаемые окажутся дважды (или более раз) просуммированы в итоговой энергии U_{short} всей цепи. Избран принцип равноправного относительно смежных фрагментов учета каждого компонента энергии. А именно, слагаемые энергии, зависящие строго от одного торсионного угла, учитывали с множителем 1/3 в трех фрагментах (вдоль по цепи), в которых этот угол варьируется. В первом фрагменте из этих трех углов варьируется как φ_3 , во втором смежном фрагменте – как φ_2 , в третьем смежном фрагменте – как φ_1 . Слагаемые, зависящие строго от двух торсионных углов, учитывали с множителем 1/2 в каждом из двух смежных фрагментов, в которых эти углы варьируются (в первом фрагменте – как φ_2, φ_3 , а во втором смежном фрагменте – как φ_1, φ_2). Слагаемые, зависящие строго от трех торсионных углов, учитывали с множителем 1 только в том фрагменте, в котором эти углы варьируются. В энергии U_m каждого фрагмента из рассмотрения были исключены такие слагаемые, которые формально зависят от трех (или двух) углов вращения, но в данном фрагменте m один из них не варьируется. Эти слагаемые включали в энергию U_m того смежного фрагмента,

в котором варьируются все три (или, соответственно, два) торсионных угла.

В концевых фрагментах весовые множители некоторых компонентов энергии отличались от таковых для "срединных" фрагментов (для корректного расчета U_{short} всей цепи). Например, в концевых фрагментах все слагаемые энергии, которые зависели только от одного торсионного угла, учитывали полностью (с множителем 1, а не 1/3), если этот угол был первым (φ_1), поскольку ни в один из последующих фрагментов эти слагаемые больше не входят. Все слагаемые энергии, которые в концевых фрагментах зависели от второго угла (φ_2), учитывали не с множителем 1/3, а с 1/2, поскольку эти слагаемые входят не в три, а в два фрагмента: в концевой и смежный с ним фрагмент. Множители 1/3 для слагаемых энергии в концевых фрагментах были использованы только для третьего угла (φ_3). Табулирование значений энергии $U_m(\varphi_1, \varphi_2, \varphi_3)$ по трем углам вращения осуществлено с интервалом 1° .

СУЩЕСТВЕННАЯ ВЫБОРКА

Если энергия U_{short} исследуемой цепной молекулы рассчитана (по фрагментам), то алгоритм существенной выборки при генерировании конформаций этой молекулы методом МК будет эффективным, если организовать выборку совокупности случайных значений торсионных углов с плотностью вероятностей $p = (1/Z_{short}) \cdot \exp[-U_{short}(\varphi_1, \varphi_2, \dots, \varphi_{N-1})/k_B T]$.

$$\text{Здесь } Z_{short} = \int_0^{2\pi} \dots \int_0^{2\pi} \exp[-U_{short}(\varphi_1, \dots, \varphi_{N-1})/k_B T] \cdot d\varphi_1 \dots d\varphi_{N-1}.$$

$$\text{Поскольку } U_{short} = \sum_{\gamma=1}^{N-3} U_{m_\gamma}(\varphi_\gamma, \varphi_{\gamma+1}, \varphi_{\gamma+2}),$$

то числитель

$$\begin{aligned} & \exp[-U_{short}(\varphi_1, \dots, \varphi_{N-1})/k_B T] / \\ & = \prod_{\gamma=1}^{N-3} \exp[-U_{m_\gamma}(\varphi_\gamma, \varphi_{\gamma+1}, \varphi_{\gamma+2})/k_B T], \end{aligned}$$

а

$$\begin{aligned} Z_{short} & = \int_0^{2\pi} \dots \int_0^{2\pi} \prod_{\gamma=1}^{N-3} \exp[-U_{m_\gamma}(\varphi_\gamma, \varphi_{\gamma+1}, \varphi_{\gamma+2}) / \\ & k_B T] \cdot d\varphi_1 \dots d\varphi_{N-1}. \end{aligned}$$

Методика реализации этой идеи такова. Используя табулированные значения энергии U_m каждого фрагмента, при данной температуре

T рассчитывали, с интервалом 1° , функции $\exp[-U_m(\varphi_1, \varphi_2, \varphi_3)/k_B T]$ и численными методами вычисляли интегралы

$$\int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \exp[-U_m(\varphi_1, \varphi_2, \varphi_3)/k_B T] d\varphi_1 d\varphi_2 d\varphi_3.$$

Пусть величина такого интеграла для фрагмента m равна I_m . Для каждого m области от 0 до 2π по каждому из трех углов ($\varphi_1, \varphi_2, \varphi_3$), т. е. "куб", разбивали на $100^3 = 1000000$ параллелепипедов (состояний) так, чтобы вероятности осуществления каждого из состояний были равны друг другу. Указанное разбиение осуществлено последовательно, численно, – с использованием метода прямоугольников [8]. Сначала были рассчитаны такие значения углов по оси φ_1 (положения плоскостей), которые разбивают трехмерное пространство ($\varphi_1, \varphi_2, \varphi_3$) на 100 "слоев", интегралы по которым от функции $\exp[-U_m(\varphi_1, \varphi_2, \varphi_3)/k_B T]$ равны $I_m/100$. Каждый из слоев затем разделили по углу φ_2 на 100 равновероятных "подслоев" из условия, чтобы интегралы от функции $\exp[-U_m(\varphi_1, \varphi_2, \varphi_3)/k_B T]$ по каждому подслою были равны $I_m/10000$, т. е. рассчитали значения углов по оси φ_2 , которые разделяют между собой подслои в каждом слое. Наконец, каждый подслой разделили по углу φ_3 на 100 состояний (параллелепипедов) так, чтобы интегралы от функции $\exp[-U_m(\varphi_1, \varphi_2, \varphi_3)/k_B T]$ по каждому параллелепипеду были равны $I_m/1000000$ (рассчитали соответствующие граничные значения углов между параллелепипедами по оси φ_3 в каждом подслое). Очевидно, что при таком способе разбиения границы параллелепипедов в областях минимумов энергии каждого фрагмента m сгущаются, т. е. количество состояний – параллелепипедов в областях вблизи минимумов энергии оказывается больше, чем в областях вблизи максимумов. Массив чисел – границ между 1000000 параллелепипедов по всем трем углам $\varphi_1, \varphi_2, \varphi_3$ данного фрагмента используется далее для генерирования состояний с плотностью вероятностей, пропорциональной величине $\exp[-U_m(\varphi_1, \varphi_2, \varphi_3)/k_B T]$, которая входит в качестве сомножителя в плотность p .

Например, на рис. 2 схематично приведено несколько плоскостей, которые делят область изменения трех варьируемых углов вращения ("куб") некоторого фрагмента на слои по оси φ_1 , – это слои вблизи середины (180°) диапазона изменения угла φ_1 .

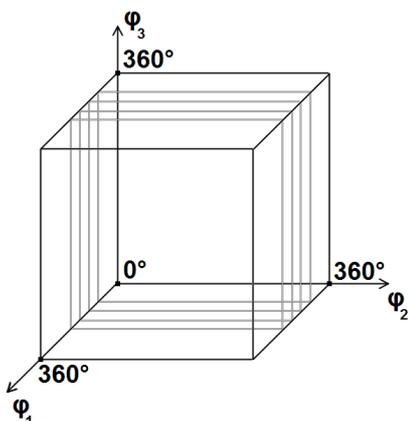


Рис. 2. Разбиение части области изменения трех варьируемых углов любого фрагмента на равновероятные слои по оси φ_1

Зависимость энергии фрагмента, изображенного на рис. 1, от углов φ_2 и φ_3 при $\varphi_1 = 180^\circ$, т. е. в одном из таких слоев, представлена в виде поверхности на рис. 3:

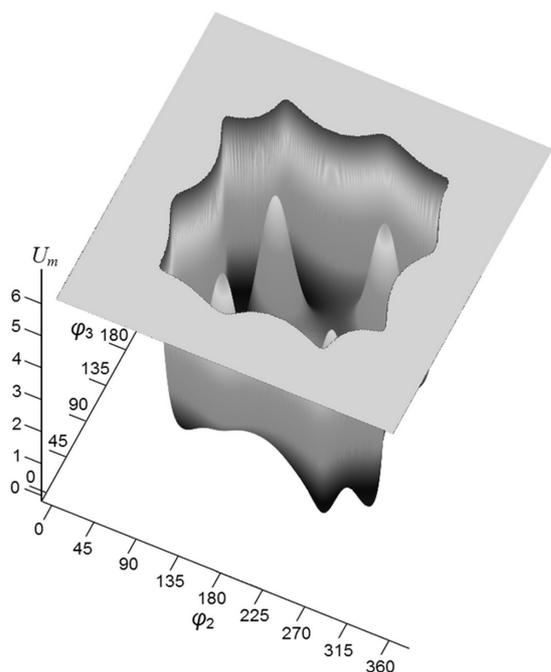


Рис. 3. Зависимость энергии (ккал/моль) фрагмента рис. 1 от углов φ_2 и φ_3 при $\varphi_1 = 180^\circ$. Отсчет энергии – от глобального минимума энергии данного фрагмента

Эту поверхность можно представить в виде эквиэнергетических линий, т. е. конформационной карты, рис. 4:

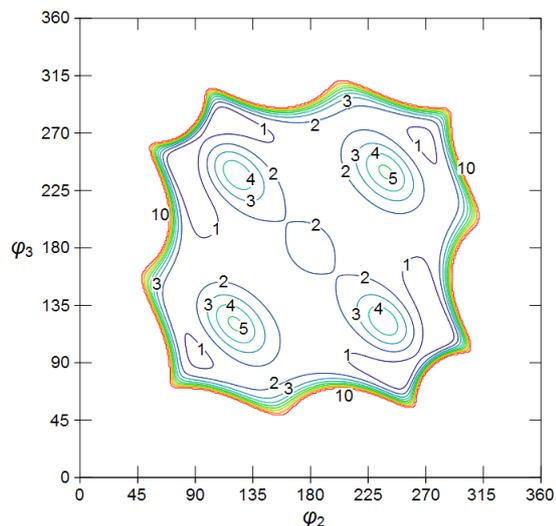


Рис. 4. Конформационная карта фрагмента рис. 1 при $\varphi_1 = 180^\circ$. Числа у кривых – энергия в ккал/моль, отсчет энергии – от глобального минимума

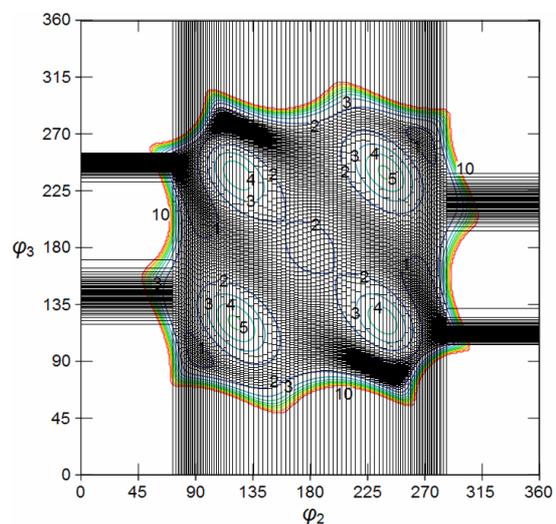


Рис. 5. Проекция границ параллелепипедов на плоскость (φ_2, φ_3) при $\varphi_1 = 180^\circ$ фрагмента рис. 1, совмещенная с соответствующей конформационной картой. Числа у кривых – энергия, ккал/моль

Слой, в который попадает значение $\varphi_1 = 180^\circ$, как и другие слои, разделен на подслои по оси φ_2 ('столбики'), а каждый подслой – на параллелепипеды по оси φ_3 . Проекция границ итоговых параллелепипедов на плоскость φ_2, φ_3 , в которой изображена также конформационная карта, имеет вид, представленный на рис. 5.

Проекция границ параллелепипедов на плоскость φ_2, φ_3 представлена на рис. 6:

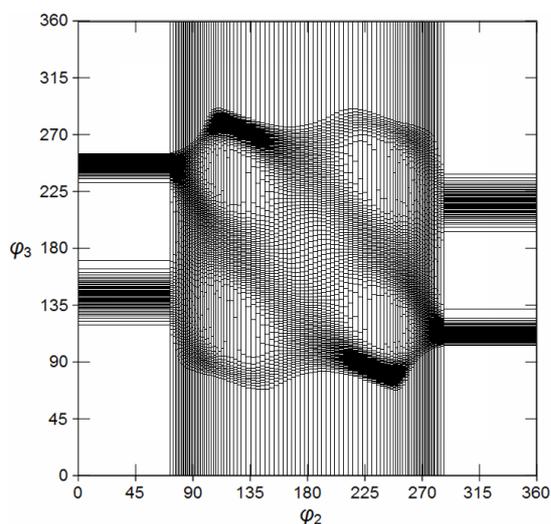


Рис. 6. Проекция границ параллелепипедов на плоскость (φ_2, φ_3) при $\varphi_1 = 180^\circ$ для фрагмента, изображенного на рис. 1

Легко видеть, что совокупность проекций границ параллелепипедов на рис. 6 воспроизводит очертания конформационной карты. Эти границы делят ось φ_2 фрагмента на полосы неравной ширины (это проекции "подслоев"), а каждую из полос по оси φ_3 — на прямоугольники (это проекции параллелепипедов) неравного размера. По количеству прямоугольников на единицу площади можно определить положения областей, в которых расположены минимумы энергии, — в них границы прямоугольников, действительно, сгущаются.

Необходимо также отметить следующее. Для расчета свойств цепных молекул часто применяется поворотнo-изомерная модель [13], в которой используются только дискретные состояния, отвечающие положениям минимумов энергии. Из рис. 3, 4 видно, что потенциальная поверхность данного молекулярного фрагмента сложна и имеет пологие, асимметричные минимумы. Такая ситуация характерна и для потенциальных поверхностей при других значениях угла φ_1 данного фрагмента, а также для потенциальных поверхностей других фрагментов молекул данного типа. Если ярко выраженных, симметричных минимумов энергии нет, то использование поворотнo-изомерного приближения затруднено, поскольку открытым остается весьма существенный вопрос о выборе значений параметров [13]. При использовании в настоящей работе непрерывного спектра конформаций учитывается асимметрия заселенности состояний. Это позволяет более адекватно моделировать макромолекулы заданной структуры, обеспечивает более широкие возмож-

ности в описании их характеристик в разных условиях, а также предсказании свойств в тех случаях, когда экспериментальные данные отсутствуют. Это создает основу для более корректного рассмотрения, например, эффектов, в которых важную роль играет "накопление" вдоль по цепи небольших отклонений углов внутреннего вращения от поворотнo-изомерных состояний. В этом состоят преимущества использования непрерывного спектра конформаций цепных молекул в теоретических моделях и, в частности, целесообразность разработки настоящего алгоритма. В последнем учитываются и другие важные факторы: воспроизводится точное химическое строение цепи и учитывается взаимозависимость каждого трех последовательных углов внутреннего вращения.

ГЕНЕРИРОВАНИЕ КОНФОРМАЦИЙ

При генерировании конформаций цепи с углами $\varphi_1, \varphi_2, \varphi_3, \dots, \varphi_{N-1}$ составляли надлежащую последовательность фрагментов, воспроизводящих строение данной цепи, согласно принятому правилу: любые два соседних (смежных) фрагмента должны иметь два общих варьируемых торсионных угла из трех. Номерам каждой тройки углов во фрагментах присваивали последовательные номера $(\varphi_1, \varphi_2, \varphi_3)$, $(\varphi_2, \varphi_3, \varphi_4)$, $(\varphi_3, \varphi_4, \varphi_5)$ и т. д. вдоль по цепи. Для выбора значений первой тройки углов $(\varphi_1, \varphi_2, \varphi_3)$ в данной цепи методом МК выбирали один из 1000000 параллелепипедов из надлежащего концевго фрагмента, т. е. номера трех сторон параллелепипеда: номер $k^{(1)}$ слоя (от 1 до 100), номер $k^{(2)}$ подслоя (от 1 до 100) и номер $k^{(3)}$ состояния в подслое (от 1 до 100). Каждый номер выбирался с помощью генератора псевдослучайных чисел [27], равномерно распределенных в интервале $[0,1)$. Пусть $\eta^{(i)} \in [0,1)$ — такие числа ($i=1, 2, 3$), тогда $k^{(i)} = \text{INT}(100 \cdot \eta^{(i)}) + 1$, где INT — целая часть числа.

В пределах избранного параллелепипеда выбирали фиксированные значения углов φ_1, φ_2 и φ_3 , полагая, что ввиду большого количества (1000000) параллелепипедов, на которые разделено конфигурационное пространство каждого трех углов, изменениями плотности вероятности внутри каждого параллелепипеда можно пренебречь, т. е. распределение значений углов в нем считали равномерным (отметим, что при этом также вычисляли и реализуемую плотность вероятностей генерирования углов; см. ниже). При выборе этих значений использовали информацию о положе-

нии границ по осям φ_1 , φ_2 и φ_3 уже избранного параллелепипеда и три очередных псевдослучайных числа $\eta^{(i)} \in [0, 1)$; $i=4, 5, 6$. Например, если $\varphi_{1k^{(1)}}$, $\varphi_{1(k^{(1)}+1)}$ – значения границ $k^{(1)}$ -го слоя по углу φ_1 , то фиксированное значение $\varphi_{1k^{(1)}}^{(fix)}$ угла φ_1 внутри этого слоя вычисляли согласно $\varphi_{1k^{(1)}}^{(fix)} = \varphi_{1k^{(1)}} + (\varphi_{1(k^{(1)}+1)} - \varphi_{1k^{(1)}}) \cdot \eta^{(4)}$. Аналогично, для расчета фиксированного значения $\varphi_{2k^{(2)}}^{(fix)}$ угла φ_2 внутри $k^{(2)}$ -го слоя использовали положения соответствующих границ и число $\eta^{(5)}$, а для $\varphi_{3k^{(3)}}^{(fix)}$ угла φ_3 внутри $k^{(3)}$ -го слоя – положения границ и число $\eta^{(6)}$.

Очередной угол φ_4 выбирали из разбиения, отвечающего смежному фрагменту с тройкой углов $\varphi_2, \varphi_3, \varphi_4$. Но при этом рассматривали лишь те 100 параллелепипедов смежного фрагмента, в которые попали уже избранные на предыдущем шаге фиксированные значения двух предыдущих углов, φ_2 и φ_3 . Аналогично вышеприведенной процедуре выбирали некоторый номер $k^{(4)}$ одного из 100 этих параллелепипедов, для чего генерировали очередное (7-е) псевдослучайное число $\eta^{(7)} \in [0, 1)$ и вычисляли $k^{(4)} = \text{INT}(100 \cdot \eta^{(7)} + 1)$. Затем внутри избранного параллелепипеда выбирали фиксированное значение угла φ_4 . По-

следняя процедура, как и в приведенном выше примере, выполнялась с использованием известных значений границ по оси φ_4 данного параллелепипеда и еще одного генерированного псевдослучайного числа $\eta^{(8)} \in [0, 1)$.

Процедуру выбора фиксированных значений всех последующих углов φ_γ , где $\gamma = 5, 6, \dots, N-1$, проводили аналогично процедуре выбора значения угла φ_4 (выбирали номер одного параллелепипеда из 100 и затем – фиксированное значение угла), с использованием двух очередных псевдослучайных чисел, полученных с помощью датчика. Таким образом, при переходе к очередному смежному фрагменту выбирали только один (следующий) торсионный угол. Отметим, что в описанном алгоритме номера торсионных углов могут следовать не подряд вдоль по цепи, в общем случае используется последовательность только варьируемых углов.

Если бы при генерировании конформации молекулы в рамках описанной процедуры удалось воспроизвести требуемую плотность $(1/Z_{short}) \cdot \prod_{\gamma=1}^{N-3} \exp[-U_{m_\gamma}(\varphi_\gamma, \varphi_{\gamma+1}, \varphi_{\gamma+2})/k_B T]$ строго, то искомая оценка \bar{H}_ω , полученная методом МК, имела бы вид, аналогичный выражению (4), т. е.

$$\bar{H}_\omega = \frac{\sum_{\nu=1}^{\omega} H(\varphi_1^\nu, \varphi_2^\nu, \dots, \varphi_{N-1}^\nu) \cdot \exp[-U_{long}(\varphi_1^\nu, \varphi_2^\nu, \dots, \varphi_{N-1}^\nu)/k_B T]}{\sum_{\nu=1}^{\omega} \exp[-U_{long}(\varphi_1^\nu, \varphi_2^\nu, \dots, \varphi_{N-1}^\nu)/k_B T]} \quad (10)$$

Но состояния в "кубах" вероятностей осуществления углов внутреннего вращения фрагментов задаются лишь с точностью до размеров параллелепипедов. Следовательно, необходимо применять выражение (9), или, в общем виде, выражения (6), (7), в которых следует вычислить реально использованную плотность вероятностей генерирования конформаций p . Это возможно (см. следующий раздел), поскольку размеры параллелепипедов, выбранных при генерировании очередной конформации цепной молекулы, известны. При этом соблюдается и общая рекомендация [12], согласно которой желательно выбирать плотность p по возможности пропорциональной подынтегральной функции, поэтому в итоге можно ожидать небольших дисперсий для оценок \bar{H}_ω .

Плотность вероятностей

Размеры каждого параллелепипеда, который выбирается в процессе генерирования конформации цепи, можно идентифицировать

индексом фрагмента m_γ и тремя случайными номерами $\lambda_{1,\gamma}^\nu$, $\lambda_{2,\gamma}^\nu$, $\lambda_{3,\gamma}^\nu$ сторон 1, 2 и 3 конкретного параллелепипеда для каждого номера фрагмента γ (т. е. торсионного угла) в каждой конформации ν данной цепи, где $\lambda_{1,\gamma}^\nu = 1, 2, \dots, 100$; $\lambda_{2,\gamma}^\nu = 1, 2, \dots, 100$; $\lambda_{3,\gamma}^\nu = 1, 2, \dots, 100$. Пусть $(L_1)_{m_\gamma, \lambda_{1,\gamma}^\nu}$, $(L_2)_{m_\gamma, \lambda_{2,\gamma}^\nu}$, $(L_3)_{m_\gamma, \lambda_{3,\gamma}^\nu}$ – длины сторон, в угловых единицах, параллелепипеда, выбранного случайно в конформации номер ν во фрагменте номер γ типа m . Согласно предложенному алгоритму, фиксированные значения углов $\varphi_1, \varphi_2, \varphi_3$ в концевом (т. е. первом, $\gamma = 1$) фрагменте выбраны равновероятно в объеме $(L_1)_{m_1, \lambda_{1,1}^\nu} \cdot (L_2)_{m_1, \lambda_{2,1}^\nu} \cdot (L_3)_{m_1, \lambda_{3,1}^\nu}$, поэтому плотность вероятностей генерирования точки с фиксированными значениями углов $(\varphi_1, \varphi_2, \varphi_3)$ равна

$$100^{-3} \cdot [(L_1)_{m_1, \lambda_{1,1}^\nu} \cdot (L_2)_{m_1, \lambda_{2,1}^\nu} \cdot (L_3)_{m_1, \lambda_{3,1}^\nu}]^{-1}.$$

Множитель 100^{-3} обусловлен тем, что в данной процедуре выбирается также один из 100

номеров каждой из 3-х сторон параллелепипеда. Генерирование значения угла φ_4 производится по фрагменту $\gamma = 2$, в котором есть три торсионных угла: $\varphi_2, \varphi_3, \varphi_4$, но первые два уже избраны на предыдущем шаге. Плотность вероятностей выбора φ_4 есть $100^{-1} \cdot [(L_3)_{m_2, \lambda_{3,2}^\nu}]^{-1}$. Аналогичный сомножи-

тель появляется для всех последующих углов во всех фрагментах вдоль по цепи в конформации с номером ν , т. е. для очередного угла номер γ это $100^{-1} \cdot [(L_3)_{m_\gamma, \lambda_{3,\gamma}^\nu}]^{-1}$. Итоговая плотность вероятностей p , с которой реально осуществляется генерирование данной конформации, равна

$$p = 100^{-(N-1)} \cdot [(L_1)_{m_1, \lambda_{1,1}^\nu} \cdot (L_2)_{m_1, \lambda_{2,1}^\nu} \cdot (L_3)_{m_1, \lambda_{3,1}^\nu}]^{-1} \cdot \prod_{\gamma=2}^{N-3} [(L_3)_{m_\gamma, \lambda_{3,\gamma}^\nu}]^{-1}. \quad (11)$$

В общем случае, если варьируются не все углы, в этой формуле в произведении вместо $N-3$ должно быть число $N_{angle} \leq N-3$. Выражение p в итоге является простым по сути: в него входят лишь произведения размеров тех параллелепипедов, внутри которых были избраны значения углов. Очевидно, что p из (11) является плотностью, допустимой по отношению к любым функциям, поскольку, в силу конечности размеров всех параллелепипедов, $p > 0$.

Оценка средних

Для того чтобы осуществить варьирование не только торсионных углов, но и валентных углов и валентных связей (варьирование обобщенное и независимое для каждого валентного угла и связи), после генерирования совокупности торсионных углов проводили, как упоминалось выше, варьирование декартовых коор-

динат x_1, x_2, \dots, x_n всех атомов молекулы. При решении ряда специальных задач эту процедуру не осуществляли. Оценка величины диапазона варьирования декартовых координат атомов проведена с использованием параметров поля CHARMM27 в варианте [21].

После завершения этапа варьирования декартовых координат всех атомов данной конформации ν и получения измененных наборов значений $\{l\}^\nu, \{\theta\}^\nu, \{\varphi\}^\nu$ проводили расчет полной энергии молекулы с использованием всех компонентов энергии поля CHARMM27 [21] и соответствующих Больцмановских множителей.

Формула (11) дает плотность вероятностей p генерирования конформации, полученной перед проведением этапа варьирования декартовых координат атомов. Именно это выражение следует подставить в (9) или (6), (7) вместо p для получения оценки \bar{H}_ω . Например, используя выражение (7), получим

$$\bar{H}_\omega = \frac{\sum_{\nu=1}^{\omega} H^\nu \cdot \exp[-U^\nu/k_B T] \cdot [(L_1)_{m_1, \lambda_{1,1}^\nu} (L_2)_{m_1, \lambda_{2,1}^\nu} (L_3)_{m_1, \lambda_{3,1}^\nu}] \cdot \prod_{\gamma=2}^{N-3} (L_3)_{m_\gamma, \lambda_{3,\gamma}^\nu}}{\sum_{\nu=1}^{\omega} \exp[-U^\nu/k_B T] \cdot [(L_1)_{m_1, \lambda_{1,1}^\nu} (L_2)_{m_1, \lambda_{2,1}^\nu} (L_3)_{m_1, \lambda_{3,1}^\nu}] \cdot \prod_{\gamma=2}^{N-3} (L_3)_{m_\gamma, \lambda_{3,\gamma}^\nu}}. \quad (12)$$

Если варьирование декартовых координат атомов при генерировании конформаций не проводится, то в оценке (12) использовано точное выражение для плотности вероятностей p генерирования конформаций. Эта плотность, хотя и не равна строго требуемой плотности

$$(1/Z_{short}) \cdot \prod_{\gamma=1}^{N-3} \exp[-U_{m_\gamma}(\varphi_\gamma, \varphi_{\gamma+1}, \varphi_{\gamma+2})/k_B T],$$

но близка к ней. В итоге выражение (12) дает состоятельную оценку, $\bar{H}_\omega \xrightarrow{P} \langle H \rangle$ при $\omega \rightarrow \infty$.

Можно продемонстрировать степень эффективности (выгодности) разработанного алгоритма существенной выборки строго по торсионным углам по сравнению с алгоритмом простой выборки, осуществив 2 варианта ((i) и (ii)) компьютерного эксперимента, в ко-

торых варьирование декартовых координат атомов не проводится, – валентные связи и валентные углы фиксированы в равновесных положениях.

Ниже приведены оценки ряда средних характеристик для неразветвленной углеводородной цепи без двойных связей, содержащей 10 углеродных атомов (условное обозначение цепи 10:0). Расчеты проведены при температуре $T = 293$ К в так называемых Θ – условиях [13] молекулы. Доверительные интервалы во всех приведенных ниже результатах отвечают 95 % надежности согласно t -распределению Стьюдента.

Вариант (i). Цепь 10:0, простая выборка (без варьирования декартовых координат), $\omega = 9,6 \cdot 10^{10}$ конформаций. Оценки средних значений получены из выражения (9), в ко-

тором плотность p избрана постоянной, случайные значения торсионных углов распределены равномерно: расстояние между концевыми атомами углерода $9,345 \pm 0,003 \text{ \AA}$, квадрат расстояния между концевыми атомами углерода $88,48 \pm 0,06 \text{ \AA}^2$, радиус инерции $3,3434 \pm 0,0006 \text{ \AA}$.

Вариант (ii). Цепь 10:0, существенная выборка (без варьирования декартовых координат), $\omega = 7,2 \cdot 10^8$ конформаций. Оценки средних значений получены из выражения (9) с плотностью p распределения торсионных углов в генерируемой выборке согласно (11): расстояние между концевыми атомами углерода $9,344 \pm 0,001 \text{ \AA}$, квадрат расстояния между концевыми атомами углерода $88,45 \pm 0,02 \text{ \AA}^2$, радиус инерции $3,3433 \pm 0,0002 \text{ \AA}$.

Легко видеть, что оценки всех соответственных характеристик в (i) и (ii) в пределах доверительных интервалов совпадают, но в варианте (ii) использован примерно в 130 раз меньший размер выборки ω , чем в варианте (i), и при этом в 3 раза меньшим оказался доверительный интервал. Если учесть, что вероятная ошибка $\sim \omega^{-1/2}$ [3, 12], то для достижения одинаковых ошибок в (i) и (ii) можно было бы уменьшить ω в варианте (ii) еще примерно в 9 раз. В итоге ω для существенной выборки можно избрать примерно в 1000 раз меньше, чем для выборки простой.

Далее, расчеты показали, что суммарные времена t генерирования случайных точек и вычисления всех заданных случайных величин в обоих алгоритмах соизмеримы: на такой расчет для одной конформации цепи по алгоритму (ii) потребовалось всего в 1,25 раза больше времени, чем по алгоритму (i). В итоге произведение времени t на ω в варианте существенной выборки (ii) оказалось в несколько раз меньше, чем в варианте простой выборки (i) и, таким образом, разработанный алгоритм существенной выборки является намного более эффективным, выгодным.

Варианты, в которых осуществляется варьирование декартовых координат всех атомов, требуют дополнительного анализа: такое варьирование приводит к изменению не только валентных связей и валентных углов (по отношению к своим равновесным значениям), но и к дополнительному изменению торсионных углов (по отношению к значениям, которые уже были получены в процедуре существенной выборки). Это изменение в формуле (11) не учтено, и для таких условий формула (11) дает лишь приблизительное выражение для плотности вероятностей p генерирования данной конформации.

В этой связи отметим, что проводимая в настоящей работе последовательность действий может быть рассмотрена и как приложение иных известных способов уменьшения дисперсии [12]. В частности, вычисление интегралов при фиксированных ("замороженных") значениях всех величин $\{l\}$ и $\{\theta\}$, осуществленное на 1-м этапе, фактически означает "интегрирование по части переменных (понижение порядка интеграла)" [12]. Доказано, что если аналитически взять интеграл по некоторым из переменных, а по остальным переменным использовать тот же метод МК, то дисперсия уменьшится. В данном случае этот этап проведен, и можно ожидать, что "размораживание" величин $\{l\}$ и $\{\theta\}$ на 2-м этапе приведет к некоторому увеличению дисперсии.

Далее, варьирование декартовых координат атомов (которое эквивалентно "размораживанию" $\{l\}$ и $\{\theta\}$), проводится лишь в малом диапазоне ($\pm 0,02 \text{ \AA}$) вокруг равновесных положений атомов; это отвечает малым колебаниям последних. Ограничение диапазона варьирования декартовых координат до конечных значений означает применение еще одного известного способа уменьшения дисперсии: "интегрирования по части области" [12]. Идея способа такова: если возможно (аналитически) вычислить интеграл по некоторой части области (допустим, что результат интегрирования равен W), то всегда выгодно вычислять простым методом МК только интеграл по оставшейся области и добавить W к итоговой оценке в качестве слагаемого. Этот способ для данного случая варьирования валентных углов и связей дает $W \rightarrow 0$, поскольку интегрирование проводится по той части области Ω , в которой энергия валентных углов и связей $\rightarrow \infty$, т. е. Больцмановский множитель в интеграле $\rightarrow 0$. Это есть вся область с большими отклонениями углов и связей. В итоге вычислению методом МК подлежит только интеграл по оставшейся области – вблизи положений равновесия атомов.

Для вычисления интеграла по этой оставшейся области можно рассчитать плотность p , – она будет иметь достаточно сложный вид. Более целесообразно, однако, использовать выражение (11) для генерирования торсионных углов, и рекомендуемый простой метод МК (равномерное распределение) для плотности декартовых координат атомов в узких "потенциальных ямах" ($\pm 0,02 \text{ \AA}$) вокруг равновесных положений. Типичные значения равновесных длин разных валентных связей в поле CHARMM27 [21] –

это $\sim 1,11 - 1,53 \text{ \AA}$ (что отвечает данным физического эксперимента), и амплитуда варьирования $\sim 0,02 \text{ \AA}$ соответствует лишь нескольким процентам от этих величин. Согласно исходным условиям, описанным выше, энергия таких колебаний должна быть сравнительно малой. Естественно ожидать, что влияние таких колебаний и на другие средние характеристики молекул также будет малым.

Методом МК получены оценки степени влияния малых колебаний на ряд характеристик. Для этого осуществлены компьютерные эксперименты (*iii*) и (*iv*) с варьированием декартовых координат всех атомов для молекулы 10:0 при $T = 293 \text{ K}$ в Θ – условиях [13].

Вариант (*iii*). Цепь 10:0, простая выборка (с варьированием декартовых координат), $\omega = 9,6 \cdot 10^{10}$ конформаций. Оценки средних значений получены из выражения (6) (или (7)), в котором положено $p = \text{Const}$, т. е. точки выбирались равномерно в пространстве торсионных углов и в малых областях вокруг равновесных значений атомов: расстояние между концевыми атомами углерода $9,347 \pm 0,007 \text{ \AA}$, квадрат расстояния между концевыми атомами углерода $88,51 \pm 0,10 \text{ \AA}^2$, радиус инерции $3,3441 \pm 0,0012 \text{ \AA}$. Эти оценки $\overline{H}_\omega \xrightarrow{P} \langle H \rangle$ при $\omega \rightarrow \infty$, поскольку все состояния выбирали равномерно, а дополнительное варьирование торсионных углов при этом не нарушает равномерности распределения по торсионным углам.

Вариант (*iv*). Цепь 10:0, существенная выборка (с варьированием декартовых координат), $\omega = 28,8 \cdot 10^8$ конформаций. Оценки средних значений получены из выражения (12), т. е. торсионные углы выбирали неравномерно, с плотностью p , но декартовые координаты в пределах областей $\pm 0,02 \text{ \AA}$ вокруг равновесных значений атомов – равномерно: расстояние между концевыми атомами углерода $9,348 \pm 0,001 \text{ \AA}$, квадрат расстояния между концевыми атомами углерода $88,54 \pm 0,02 \text{ \AA}^2$, радиус инерции $3,3443 \pm 0,0002 \text{ \AA}$. Размер выборки ω в варианте (*iv*) был увеличен в 4 раза по сравнению с таковым в варианте (*ii*) для того, чтобы добиться совпадения доверительных интервалов в (*ii*) и (*iv*).

Как и ожидалось, вследствие учета малых колебаний оценки геометрических характеристик в варианте (*iii*) увеличились по сравнению с (*i*) (и, соответственно, в варианте (*iv*) – увеличились по сравнению с (*ii*)). Вместе с тем относительный рост этих характеристик составил всего 0,02–0,04 %; т. е. влияние обсуждаемого варьирования декартовых координат

атомов данной цепной молекулы на ее геометрические характеристики, действительно, оказалось малым.

Отметим, что оценки всех соответственных характеристик вариантов (*iii*) и (*iv*) в пределах доверительных интервалов совпадают (аналогично тому, как совпадают оценки характеристик в вариантах (*i*) и (*ii*)). Размер выборки ω в варианте (*iii*) примерно в 30 раз больше, чем в варианте (*iv*), а доверительные интервалы – больше примерно в 5–7 раз, т. е. для получения одинаковых интервалов в (*iii*) и (*iv*) следует увеличить ω в варианте (*iii*) еще примерно в 25–50 раз, и получим в ~ 750 – 1500 раз больше, чем в варианте (*iv*). Поскольку времена расчета случайных величин соизмеримы, вариант существенной выборки с варьированием координат (*iv*) оказался также в сотни раз эффективнее, чем соответствующий вариант простой выборки (*iii*).

Разработанный алгоритм компьютерной имитации цепных молекул позволяет воспроизвести точное химическое строение цепи, учитывает взаимозависимость каждого трех последовательных углов внутреннего вращения и непрерывный спектр конформаций цепей, а при генерировании этих конформаций использует метод существенной выборки. В итоге он позволяет передать основные особенности молекул цепного строения и, обладая общностью, может быть адаптирован для конструирования и исследования свойств цепных молекул разных классов.

Работа выполнена при поддержке РФФИ (проект 10-03-00201а) и программы Президента РФ – Ведущие научные школы (НШ-1642.2012.4).

ЛИТЕРАТУРА

1. Биндер К. Методы Монте-Карло в статистической физике / Ред. К. Биндер. М.: Мир, 1982. С. 9–57.
2. Биндер К., Хеерман Д. В. Моделирование методом Монте-Карло в статистической физике: Введение. М.: Наука; Физматлит, 1995. 144 с.
3. Бусленко Н. П., Голенко Д. И., Соболев И. М., и др. Метод статистических испытаний (метод Монте-Карло). М.: ГИФМЛ, 1962. 332 с.
4. Гнеденко Б. В. Курс теории вероятностей. М.: Гос. изд. физ.-мат. лит., 1961. 408 с.
5. Гулд Х., Тобочник Я. Компьютерное моделирование в физике. М.: Мир, 1990. Ч. 2. 400 с.
6. Ермаков С. М. Метод Монте-Карло и смежные вопросы. М.: Наука; Гл. ред. физ.-мат. лит., 1971. 328 с.

7. *Ермаков С. М., Михайлов Г. А.* Статистическое моделирование. М.: Наука; Гл. ред. физ.-мат. лит., 1982. 296 с.
8. *Зализняк В. Е.* Основы научных вычислений. Введение в численные методы для физиков. М.: Едиториал УРРС, 2002. 296 с.
9. *Ландау Л. Д., Лифшиц Е. М.* Статистическая физика. Ч. 1. М.: Наука; Гл. ред. физ.-мат. лит., 1976. 584 с.
10. Методы компьютерного моделирования для исследования полимеров и биополимеров / Отв. ред. В. А. Иванов, А. Л. Рабинович, А. Р. Хохлов. М.: Книжный дом ЛИБРОКОМ, 2009. 696 с.
11. *Рабинович А. Л., Иванов В. А.* Методы компьютерного моделирования для исследования полимеров и биополимеров / Отв. ред. В. А. Иванов, А. Л. Рабинович, А. Р. Хохлов. М.: Книжный дом ЛИБРОКОМ, 2009. С. 63–119.
12. *Соболь И. М.* Численные методы Монте-Карло. М.: Наука; Гл. ред. физ.-мат. лит., 1973. 312 с.
13. *Флори П.* Статистическая механика цепных молекул. М.: Мир, 1971. 440 с.
14. *Хеерман Д. В.* Методы компьютерного эксперимента в теоретической физике. М.: Наука; Гл. ред. физ.-мат. лит., 1990. 176 с.
15. *Allen M. P., Tildesley D. J.* Computer Simulation of Liquids. Oxford: Clarendon Press, 1987. 385 p.
16. *Feller S. E., MacKerell, Jr. A. D.* An improved empirical potential energy function for molecular simulations of phospholipids // *J. Phys. Chem. B.* 2000. Vol. 104. P. 7510–7515.
17. *Feller S. E., Yin D., Pastor R. W., MacKerell, Jr. A. D.* Molecular dynamics simulation of unsaturated lipids at low hydration: parametrization and comparison with diffraction studies // *Biophys. J.* 1997. Vol. 73. P. 2269–2279.
18. *Flory P. J.* Foundations of rotational isomeric state theory and general methods for generating configurational averages // *Macromolecules.* 1974. Vol. 7, N 3. P. 381–392.
19. *Frenkel D., Smit B.* Understanding Molecular Simulation. From Algorithms to Applications. 2nd edition, San Diego: Academic Press, 2002. 638 p.
20. *Gö N., Scheraga H. A.* On the use of classical statistical mechanics in the treatment of polymer chain conformations // *Macromolecules.* 1976. Vol. 9, N 4. P. 535–542.
21. *Högberg C.-J., Nikitin A. M., Lyubartsev A. P.* Modification of the CHARMM force field for DMPC lipid bilayer // *J. Comput. Chem.* 2008. Vol. 29. P. 2359–2369.
22. *Kahn H.* Random sampling (Monte Carlo) techniques in neutron attenuation problems. 1 // *Nucleonics.* 1950. Vol. 6, N 5. P. 27–33.
23. *Kahn H.* Random sampling (Monte Carlo) techniques in neutron attenuation problems. 2 // *Nucleonics.* 1950. Vol. 6, N 6. P. 60–65.
24. *Klada J. B., Brooks B. R., MacKerell, Jr. A. D. et al.* An ab initio study on the torsional surface of alkanes and its effect on molecular simulations of alkanes and DPPC bilayer // *J. Phys. Chem. B.* 2005. Vol. 109. P. 5300–5311.
25. *Landau D. P., Binder K.* A Guide to Monte Carlo Simulations in Statistical Physics. 2nd edition. Cambridge University Press, 2005. 432 с.
26. Simulation Methods for Polymers / Eds. Kotelyanskii M., Theodorou D.N. N.Y.: Marcel Dekker, Inc., 2004. 602 p.
27. *Matsumoto M., Nishimura T.* Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator // *ACM Transactions on Modeling and Computer Simulations: Special Issue on Uniform Random Number Generation.* 1998. Vol. 8. P. 1–25.
28. *Schlenkrich M., Brickmann J., MacKerell, Jr. A. D., Karplus M.* In: Biological membranes: a molecular perspective from computation and experiment / K. M. Merz, B. Roux, Eds. Birkhauser, Boston. 1996. P. 31–81.

СВЕДЕНИЯ ОБ АВТОРАХ:

Рабинович Александр Львович

гл. науч. сотр., д. ф.-м. н.
 Институт биологии Карельского научного центра РАН
 ул. Пушкинская, 11, Петрозаводск, Республика Карелия, Россия, 185910
 эл. почта: rabinov@krc.karelia.ru
 тел.: (8142) 571879

Журкин Дмитрий Викторович

аспирант
 Петрозаводский государственный университет,
 физико-технический факультет
 ул. Университетская, 10 А, Петрозаводск, Республика Карелия, Россия, 185910
 эл. почта: zhurkin_dm@mail.ru
 тел.: (8142) 719671

Rabinovich, Alexandr

Institute of Biology, Karelian Research Centre, Russian Academy of Sciences
 11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia
 e-mail: rabinov@krc.karelia.ru
 tel.: (8142) 571879

Zhurkin, Dmitrii

Petrozavodsk State University, Faculty of Physical Engineering
 10 A University St., 185910 Petrozavodsk, Karelia, Russia
 e-mail: zhurkin_dm@mail.ru
 tel.: (8142) 719671