

Расчет теплоемкости цепочек Peyrard-Bishop-Dauxois методом Монте-Карло

Теплухин А.В.

Институт математических проблем биологии РАН – филиал ИПМ РАН
им. М.В. Келдыша

tepl@impb.ru

Методом Монте-Карло для канонического *NVT*-ансамбля рассчитана зависимость удельной теплоемкости цепочек *Peyrard-Bishop-Dauxois* от температуры, а также построены соответствующие профили плавления, показывающие, как при этом изменяется среднее по цепочке расстояние между нуклеотидами одного сайта (уотсон-криковской пары). Объем конфигурационного пространства фиксировался путем введения ограничения на величину этого расстояния. Результаты моделирования в таких условиях говорят о том, что температура плавления и ширина переходной области существенно зависят не только от длины цепочки, но и от размера ограничения.

Ключевые слова: ДНК, ПБД-модель, Монте-Карло, моделирование.

Calculation of Heat Capacity of Peyrard-Bishop-Dauxois Chains by Monte Carlo Method

Teplukhin A.V.

Institute of mathematical problems of biology RAS – the Branch of Keldysh Institute of Applied Mathematics RAS

The Monte Carlo method for the canonical *NVT*-ensemble has been used to calculate the dependency of specific heat of *Peyrard-Bishop-Dauxois* chains on temperature as well as to obtain the respective «melting» profiles, showing how the mean distance between nucleotides of the same site (Watson-Crick-pairs) is changed. The volume of the configuration space was limited by imposing restriction on maximal value of the distance. As it follows from the results of simulation of the PDB-chains in such a conditions, the «melting» temperature and width of the transition region are highly depend both on the chain length and the amplitude of allowed variations for the intrasite distance.

Key words: DNA, PBD model, Monte Carlo, simulation.

1. Введение

Цепочки Peyrard-Bishop-Dauxois (ПБД) [1, 2] представляют собой математическую модель процесса денатурации-«плавления» двойной спирали ДНК [3–5]. Небольшое число эмпирических параметров, задающих энергетические и геометрические характеристики взаимодействий азотистых оснований, а также ковалентных связей между нуклеотидами, позволяет добиться высокой степени сходства для кривых плавления, полученных в результате расчетов и экспериментально [6–8].

Следует отметить, что феноменологический характер модели затрудняет интерпретацию получаемых с ее помощью результатов, вызывая серьезные возражения и конструктивную критику со стороны специалистов в области физики ДНК и полимеров [9]. В то же время, благодаря своей

простоте эта модель приобрела высокую популярность среди прикладных задач вычислительной физики и математики [10, 11].

Цель настоящей работы – показать, как температурные профили удельной теплоемкости цепочек ПБД, рассчитываемые в результате вычислительных экспериментов на компьютере, могут зависеть от количества сайтов в цепочке и объема координатного пространства модели.

2. Методика расчетов

В рамках модели ПБД [1, 2] молекула ДНК представляет собой систему N сайтов с потенциальной энергией, вычисляемой по формуле [2]:

$$E = D \sum_{i=1}^N (1 - e^{-aX_i})^2 + \frac{k}{2} \sum_{i=2}^N (1 + \rho e^{-b(X_i + X_{i-1})}) (X_i - X_{i-1})^2.$$

Здесь $D = k = 0.9224$ ккал/моль, $a = 4.45 \text{ \AA}^{-1}$, $\rho = 0.5$, $b = 0.35 \text{ \AA}^{-1}$, X_i – геометрическая характеристика, выражающая удлинение водородных связей между основаниями в уотсон-криковской паре i -го сайта ($X_i = 0$ для «стандартной» геометрии A:T-пары в B-ДНК).

Расчет геометрических (среднее по цепочке значение «координаты» $\langle X \rangle$) и термодинамических (среднее значение потенциальной энергии цепочки $\langle E \rangle$ и ее квадрата $\langle E^2 \rangle$) характеристик проводился методом Монте-Карло для канонического NVT -ансамбля по стандартной процедуре Метрополиса и др. [12]. Объем конфигурационного пространства цепочки из N сайтов определяется параметром L , фиксирующим диапазон значений их X -координат.

В ходе итераций этой процедуры выполнялась попытка изменить X -координату случайно выбранного сайта. Номер сайта, а также величина и направление смещения определялись с помощью генератора псевдослучайных чисел [13]. Попытки перехода в состояния с $|X_i| > L$ отвергаются, т. е. используется модифицированный потенциал морзевого типа с бесконечной «стенкой» при $|X_i| > L$. Это обеспечивает сходимость статсуммы и термодинамическое поведение модели.

Чтобы при температуре T отвергалось не более 60% пробных конфигураций цепочек, максимальное смещение по X -координате за одну итерацию не должно превышать $\Delta X = 0.073\sqrt{T} + 0.001 \text{ \AA}$. Для достижения приемлемой статистической погрешности с каждым сайтом выполнялось, в среднем, $5 \cdot 10^9 - 10^{10}$ испытаний, в зависимости от N .

Вычислительные работы проводили на суперкомпьютере «Ломоносов» (МГУ, г. Москва¹) с помощью программы, разработанной автором статьи. Эта программа использует 96 процессорных ядер, что позволяет одновременно (параллельно) производить расчет свойств ПБД-цепочки при 96 значениях температуры (в диапазоне от 300 до 396 К с шагом в один градус).

Значения удельной (на один сайт) теплоемкости ПБД-цепочек в молярном выражении рассчитывались по стандартной флуктуационной формуле: $C_v = \frac{1}{NRT^2} (\langle E^2 \rangle - \langle E \rangle^2) + \frac{R}{2}$. Здесь R – универсальная газовая постоянная, а $R/2$ – вклад кинетической энергии.

3. Результаты расчетов

На рисунке 1 представлены зависимости удельной теплоемкости ПБД-цепочек различной длины ($N = 100, 200, 500$ и 1000 сайтов – «пар оснований») от температуры. В этой группе расчетов максимальное значение X -координаты сайтов было ограничено сверху значением $L = 50 \text{ \AA}$. На этом рисунке хорошо видно, как с ростом длины цепочки растет температура «плавления» и высота

пика теплоемкости, а ширина перехода, наоборот, уменьшается. Такое поведение теплоемкости указывает на высокую кооперативность процесса и вполне соответствует экспериментальным данным о «плавлении» реальной ДНК в физиологическом растворе [3–5].

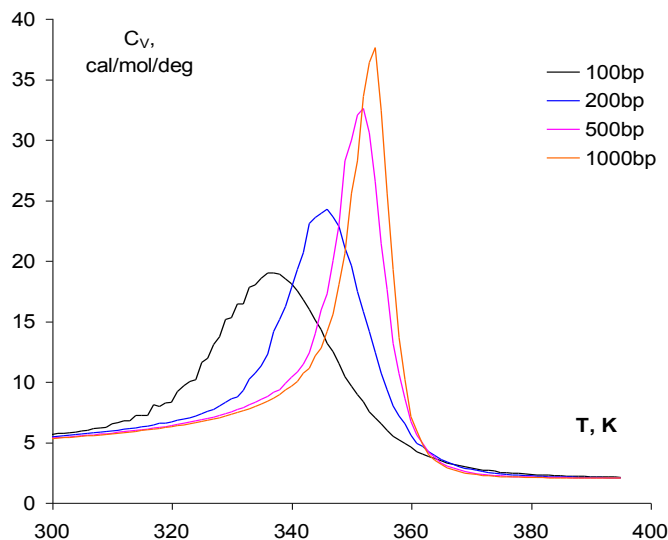


Рис. 1. Зависимость удельной теплоемкости цепочек длиной 100–1000 пар оснований (bp) от температуры. Диапазон изменения X -координаты ограничен 50 \AA .

Как показали наши дальнейшие расчеты, форма кривых $C_v(T)$ существенно зависит от выбора величины параметра L . На рисунке 2 представлено семейство графиков теплоемкости цепочки из 100 сайтов для нескольких значений L . Видно, что увеличение объема конфигурационного пространства модели дает эффект обратный упомянутому выше: с ростом L температура «плавления» и ширина перехода уменьшаются, а высота пика растет.

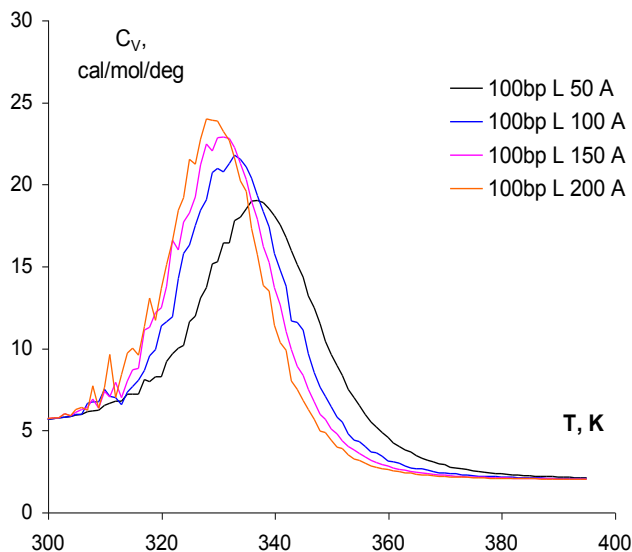


Рис. 2. Зависимость удельной теплоемкости цепочек длиной 100 пар оснований (bp) от температуры для четырех значений L .

¹ <http://parallel.ru/cluster>

Аналогичный эффект влияния величины параметра L был выявлен и в расчетах для цепочки из 1000 сайтов (см. рис. 3). Следует отметить, что в отличие от 100-сайтовых цепочек, уменьшение температуры «плавления» и ширины перехода с ростом L стало значительно слабее, напротив, высота пика теплоемкости стала меняться резко.

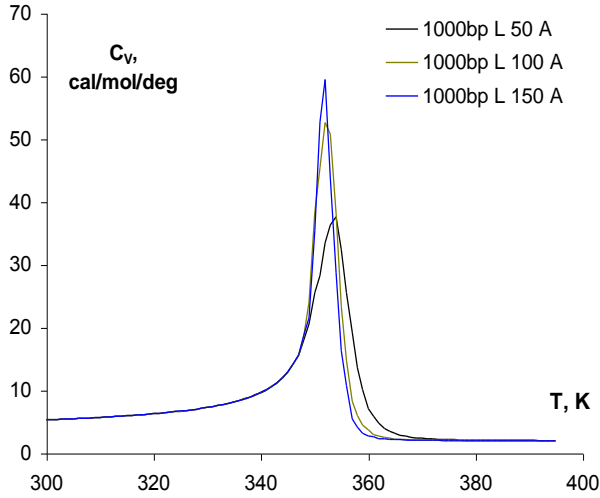


Рис. 3. Зависимость удельной теплоемкости цепочек длиной 1000 пар оснований (bp) от температуры для трех значений L .

Профили «плавления» ПБД-цепочек (см. рис. 4 и 5) показывают изменение среднего по ансамблю конфигураций и сайтам цепочки значения X -координат с ростом температуры. В ПБД-модели X -координата характеризует степень «открытия» отдельного сайта цепочки. Проводя аналогию с реальной ДНК, можно сказать, что эта величина задает в некотором обобщенном смысле расстояние между нуклеотидами в уотсон-криковской паре.

На рисунке 4 видно, что эффект кооперативности заметно повышает стабильность длинных цепочек. Тем не менее, у цепочек, состоящих более чем из 500 сайтов, он становится слабее, чем у более коротких.

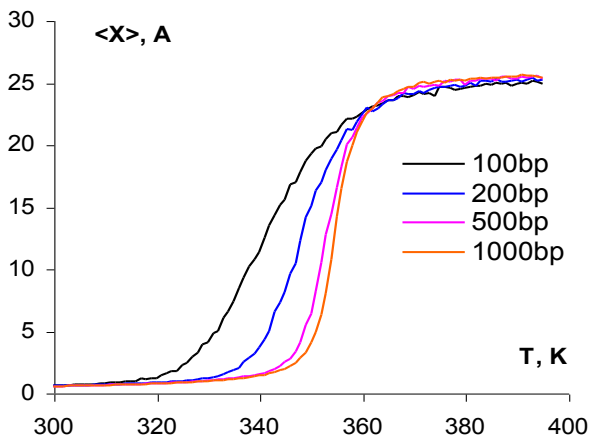


Рис. 4. Профили «плавления» ПБД-цепочек разной длины для $L = 50 \text{ \AA}$.

На рисунке 5 показано, как будут меняться профили плавления цепочки из 100 сайтов при увеличении параметра L . Видно, что с ростом L массовое «открытие» сайтов начинается при более низких температурах. Можно предположить, что этот эффект имеет энтропийную природу. Действительно, увеличение объема конфигурационного пространства повышает долю конфигураций в которых X -координата сайта удовлетворяет простейшему геометрическому критерию «открытия» сайта, например $|X_i| > 5 \text{ \AA}$, даже при низких температурах, что неизбежно ведет к росту статистического веса таких состояний.

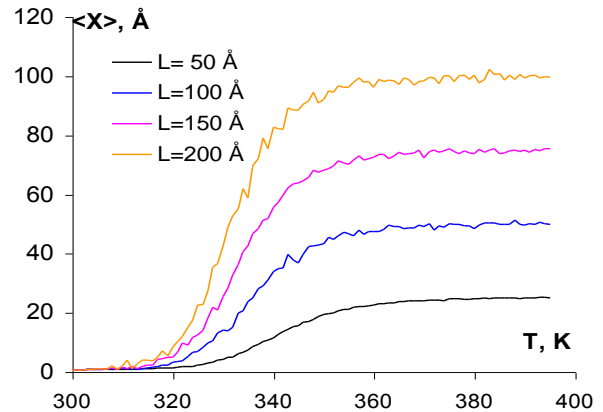


Рис. 5. Профили «плавления» ПБД-цепочек из 100 сайтов при разных L .

4. Заключение

Метод Монте-Карло для NVT-ансамбля позволяет быстро и с высокой точностью рассчитывать термодинамические характеристики ПБД-цепочек. В перспективе этот подход может успешно применяться для получения профилей «плавления» моделей ДНК с заданной последовательностью нуклеотидов. Тем не менее, как показали наши расчеты, при планировании вычислительных экспериментов следует учитывать влияние размера области изменения X -координат сайтов на получаемые результаты.

5. Список литературы

1. Peyrard M., Bishop A.R. *Phys. Rev. Lett.* 1989. V. 62. P. 2755.
2. Dauxois T., Peyrard M., Bishop A.R. *Phys. Rev. E.* 1993. V. 47. P. R44.
3. Zimm B.H. *J. Chem. Phys.* 1960. V. 33. P. 1349.
4. Wartell R.M., Benight A.S. *Phys. Rep.* 1985. V. 126. P. 67.
5. Vologodskii A., Frank-Kamenetskii M.D. *Phys. Life Rev.* 2018. V. 25. P. 1.
6. Ares S., Voulgarakis N.K., Rasmussen K.Ø., Bishop A.R. *Phys. Rev. Lett.* 2005. V. 94. P. 035504.
7. van Erp T.S., Cuesta-López S., Peyrard M. *Eur. Phys. J. E.* 2006. V. 20. P. 421.

8. Peyrard M., Cuesta-López S., Angelov D. *J. Phys.: Condens. Matter.* 2009. V. 21. P. 034103.
9. Frank-Kamenetskii M.D., Prakash S. *Phys. Life Rev.* 2014. V. 11. P. 153.
10. Peyrard M., Cuesta-López S., James G. *Nonlinearity.* 2008. V. 21. P. T91.
11. Likhachev I.V., Lakhno V.D. *Chem. Phys. Lett.* 2019. V. 727. P. 55.
12. Metropolis N.A., Rosenbluth A.W., Rosenbluth M.N., Teller A.H. *J. Chem. Phys.* 1953. V. 21. P. 1087.
13. Wichmann B.A., Hill I.D. *Comput. Statist. Data Anal.* 2006. V. 51. P. 1614.