



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 232 за 2018 г.



ISSN 2071-2898 (Print)  
ISSN 2071-2901 (Online)

**Лихачев И.В., Лахно В.Д.**

Исследование денатурации  
ДНК на основе модели  
Пейрарда-Бишопа-Доксуа  
методом молекулярной  
динамики

**Рекомендуемая форма библиографической ссылки:** Лихачев И.В., Лахно В.Д. Исследование денатурации ДНК на основе модели Пейрарда-Бишопа-Доксуа методом молекулярной динамики // Препринты ИПМ им. М.В.Келдыша. 2018. № 232. 12 с. doi:[10.20948/prepr-2018-232](https://doi.org/10.20948/prepr-2018-232)  
URL: <http://library.keldysh.ru/preprint.asp?id=2018-232>

**Ордена Ленина  
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ  
имени М.В.Келдыша  
Российской академии наук**

**Лихачев И.В., Лахно В.Д.**

**ИССЛЕДОВАНИЕ ДЕНАТУРАЦИИ  
ДНК НА ОСНОВЕ МОДЕЛИ  
ПЕЙРАРДА-БИШОПА-ДОКСУА  
МЕТОДОМ  
МОЛЕКУЛЯРНОЙ ДИНАМИКИ**

**Москва — 2018**

*Лихачев И.В., Лахно В.Д.*

### **Исследование денатурации ДНК на основе модели Пейрарда-Бишопа-Доксуа методом молекулярной динамики**

Методом прямого молекулярно-динамического моделирования с использованием модели Пейрарда-Бишопа-Доксуа изучен фазовый переход (PolyA/PolyT)<sub>100</sub> дуплекса ДНК в денатурированное состояние. Рассчитаны температурные зависимости полной энергии и теплоёмкости дуплекса. Использованный подход может быть применён для расчёта статистических свойств дуплексов любой длины и состава нуклеотидов.

**Ключевые слова:** одномерное моделирование ДНК, гомогенная нуклеотидная последовательность, столкновительный термостат

*Ilya Viatcheslavovitch Likhachev, Victor Dmitrievitch Lakhno*

### **Investigation of DNA denaturation in Peyrard-Bishop-Dauxois model by molecular dynamics method**

The phase transition of (PolyA/PolyT)<sub>100</sub> duplex into the denaturated state is studied in the Peyrard-Bishop-Dauxois model by the method of direct molecular-dynamical modeling. The temperature dependencies of the total energy and heat capacity of the duplex are calculated. The approach applied can be used to calculate the statistical properties of the duplexes of any length and nucleotide composition.

**Key words:** one-dimensional DNA modeling, homogeneous nucleotide sequence, collisional thermostat

Работа выполнена при поддержке Российского научного фонда, проект РНФ 16-11-10163.

## **Оглавление**

Введение .....	3
Модель Пейрарда-Бишопа-Доксуа .....	3
Термостат .....	4
Численные эксперименты с линейным нагреванием .....	5
Благодарности.....	11
Используемая литература .....	11

## **Введение**

При моделировании молекулярной динамики биомакромолекул используются потенциалы межатомного взаимодействия, которые всегда ограничены по величине при больших расстояниях между атомами (Леннард-Джонс, Кулон, Морз и др. [1]). Это приводит к известным трудностям при расчёте термодинамических характеристик биомакромолекул. Действительно, статистическая сумма для ограниченных потенциалов всегда расходится. Это означает, что макромолекулы должны были бы находиться в денатурированном состоянии при сколь угодно низкой температуре. В действительности, однако, этого не происходит, поскольку в растворе имеется конечная концентрация молекул, что задаёт характерный масштаб обрезания при расчете статсуммы.

Если формально иметь дело не с ансамблем, а с отдельной молекулой в неограниченном растворителе и температурой значительно ниже температуры денатурации, то в расчете всегда имеется четкий выход на стационарное квазиравновесное состояние, которое существует много дольше, чем любой интересующий нас эксперимент. Это равновесное состояние, естественно, будет совпадать с тем, которое мы получили бы, рассматривая ансамбль молекул.

Таким образом, задачей моделирования является не формальный расчет статистической суммы на бесконечном времени, соответствующем распавшейся системе, а расчёт, приводящий именно к квазиравновесному состоянию и расчету его термодинамических характеристик.

Проблема, однако, возникает, когда температура оказывается близка к тому значению, при котором два характерных времени сравниваются (т.е. время выхода на стационар и время распада системы).

Такая ситуация возникает, например, при расчете термодинамических характеристик ДНК, температура распада которой ненамного превышает комнатную. Прямой расчет статсуммы в этом случае невозможен, т.к. он уже на конечном времени приводит к развалу молекулы.

Рассмотрению этого вопроса и посвящена данная статья. В ней описывается метод, который может быть эффективно использован для расчета как термодинамических характеристик молекулы вблизи перехода, так и самого такого перехода.

Плавление (денатурация) ДНК [1] происходит из-за увеличения температурных флуктуаций. При этом обе цепочки ДНК отходят друг от друга, разрывая водородные связи. Сами цепочки остаются неразрывными.

## **Модель Пейрарда-Бишопа-Доксуа**

Одна из наиболее используемых динамических моделей ДНК, которая описывает такое поведение молекулы – модель Пейрарда-Бишопа-Доксуа (ПБД) – моделирует поведение как разрывающихся водородных связей (модель

Пейрарда-Бишопа), используя потенциал Морзе, так и неразрывность каждой из цепочек благодаря добавке Пейрарда-Доксуа [2].

В этой модели цепочка нуклеотидов ДНК представляется в виде системы материальных точек в одномерном пространстве, движение которых описывается классическими уравнениями движения:

$$m_i \frac{d^2 x_i}{dt^2} = - \frac{dU}{dx_i},$$

где  $i=1, \dots, N$ ;  $N$  – количество частиц;  $m$  – приведенная масса частицы;  $x$  – координата отклонения от положения равновесия между нуклеотидами;  $U$  – потенциал;  $t$  – время.

Согласно модели ПБД, каждая частица находится в потенциальном поле:

$$U = U_{Morze} + W.$$

В модели ПБД для описания взаимодействующих нуклеотидов в паре используется потенциал Морзе:

$$U_{Morze}(x_i) = D(1 - e^{-\alpha x_i})^2,$$

а взаимодействие соседних пар имеет вид:

$$W(x_i, x_{i-1}) = \frac{k}{2} (1 + \rho e^{-a(x_i + x_{i-1})}) (x_i - x_{i-1})^2.$$

Рассматривается цепочка из 100 пар оснований АТ для значений параметров:  $D=k=0.92$  ккал/моль,  $\alpha=4.45 \text{ \AA}^{-1}$ ,  $a=0.35 \text{ \AA}^{-1}$ ,  $\rho=0.5$ ,  $m=300$  а.е.м., использованных в [2].

Для интегрирования уравнений движения модели ПБД будем использовать скоростной алгоритм Верле [3].

### Термостат

Учет температуры в модели производится, как и в методе полноатомной молекулярной динамики [4], посредством столкновительного термостата, который исходит из представления среды, в которую помещена моделируемая система, в виде виртуальных точечных частиц, имеющих распределение Максвелла по скоростям [5, 6, 7]. Распределение скоростей отвечает заданной температуре  $T_{ref}$ . Виртуальные частицы среды в случайные моменты времени упруго сталкиваются с частицами системы. Уравнения движения имеют вид:

$$m_i \frac{dv_i}{dt} = F_i + \sum_k f_{ik} \cdot \delta(t - t_{ik}),$$

$$F_i = - \frac{dU}{dx_i}.$$

Здесь  $\delta(t)$  – дельта-функция Дирака,  $f_{ik}$  – стохастический источник силы, приводящий к скачку скорости  $i$ -го атома в случайные моменты времени  $t_{ik}$ . Значение скачка скорости вычисляется как результат столкновения двух точечных частиц, имеющих перед столкновением скорости  $v$  (для моделируемой частицы) и  $v_0$  (для виртуальной):

$$\Delta v(t) = \frac{2m_0}{m_0 - m} (v_0(t) - v(t)).$$

Здесь  $m$  – масса моделируемой частицы,  $m_0$  – масса виртуальной частицы. Скорости  $v_0$  выбираются из распределения Максвелла.

Столкновения происходят в соответствии с процессом Пуассона, который определяется единственным параметром  $\lambda_0$  – средним числом столкновений атома с частицами среды в единицу времени, т.е. частотой столкновений.

В процессе моделирования температура столкновительного термостата  $T_{\text{текущая}}$  менялась линейно с постоянной скоростью  $V_t=0,1$  К/нс  $T_{\text{текущая}}=T_{\text{начальная}}+V_t*t$ ; что приводит к графику зависимости удельных потенциальной, кинетической и полной энергий от времени (см. рис. 1).

## Численные эксперименты с линейным нагреванием

Начальная температура выбирается заведомо меньшей точки денатурации ДНК. В приведенном примере берется  $T_{\text{начальная}}=250$  К.

Скорость нагрева за единицу времени (1 нс) достаточно мала, так что можно считать, что на любом достаточно малом временном интервале система находится в равновесном состоянии. Как показывают численные эксперименты, релаксация системы – выход потенциальной энергии на плато из температуры 0 К на 250 К – происходит менее чем за 1 нс.

На рис. 1 показана температурная зависимость энергии полинуклеотидной цепочки, содержащей 100 пар оснований в отдельном вычислительном эксперименте (одной реализации). Кинетическая энергия изменяется линейно от температуры в области 300 К до 400 К. Потенциальная энергия и, как следствие, полная энергия, имеет два участка, близких к линейным, и область перехода, где энергия меняется скачком. Средняя по времени кинетическая энергия линейно зависит от времени благодаря столкновительному термостату.

Приведенная на рис. 1 зависимость полной и потенциальной энергии от времени описывает выход из квазигармонического режима в денатурированный. Время  $t_0$  соответствует температуре термостата  $T_{\text{REF}}=250$  К+0.1 К/нс\* $t_0$ . Усреднение этих времен по всем реализациям даёт температуру перехода (см. рис. 2).

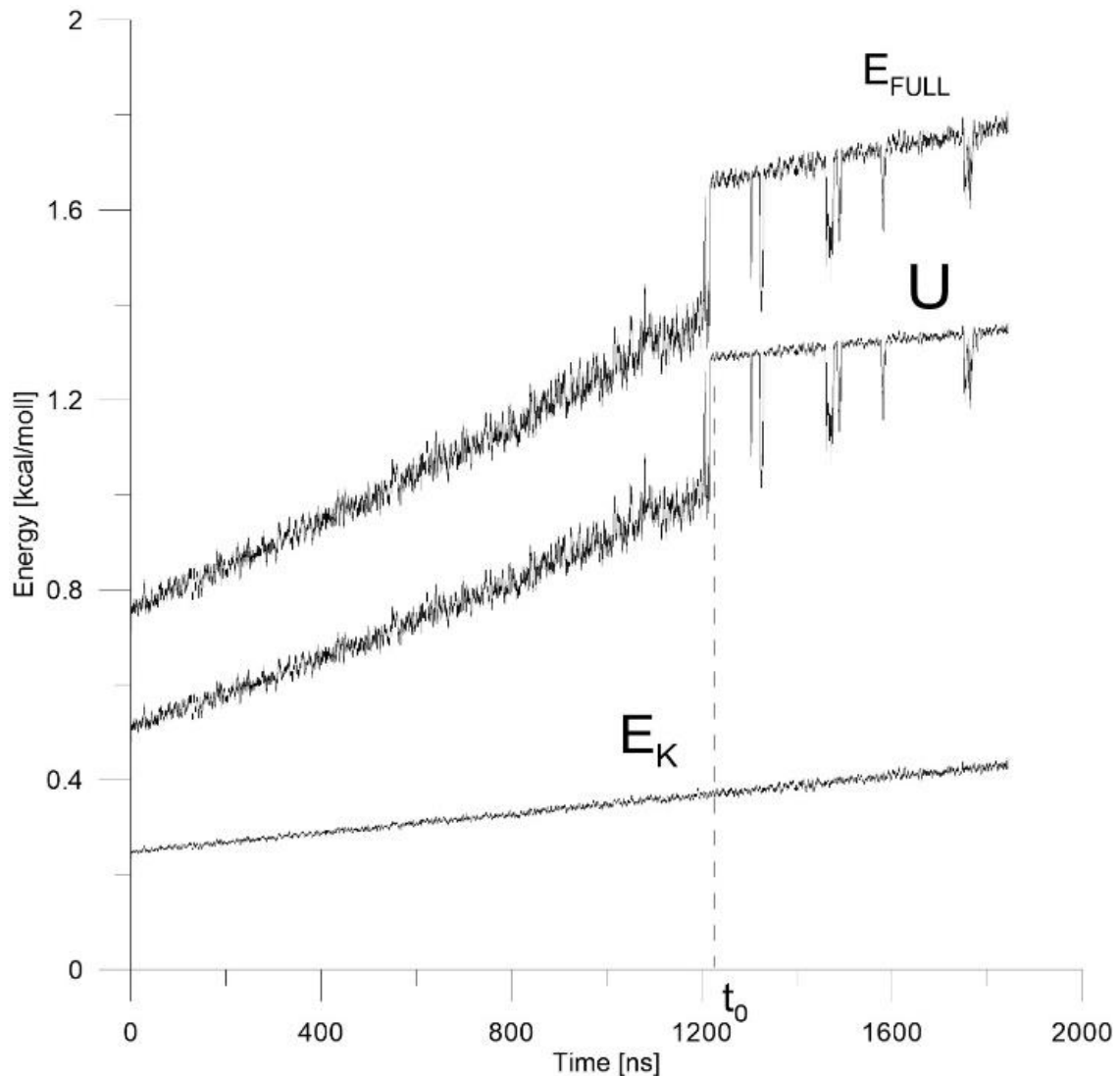


Рис. 1. Зависимость удельной полной ( $E_{FULL}$ ), потенциальной ( $U$ ) и кинетической ( $E_K$ ) энергий от времени для одной реализации в численном эксперименте при линейном повышении температуры термостата  $T_{REF}=250\text{ K}+0.1\text{ K/нс}\cdot t$ .

Фактически рис. 1 описывает энергетическую щель, имеющую порядок величины  $k_B T$ , где  $T$  – температура, соответствующая времени скачка  $t_0$ . Она отделяет денатурированное состояние (справа от скачка) от неденатурированного (слева от скачка). У разных реализаций скачок происходит в разные промежутки времени ввиду случайности действия внешней силы – термостата. Ранний переход иллюстрирует большую энергетическую щель, поздний – меньшую. Наличие такой щели, как это было отмечено выше, обусловлено расходимостью статистической суммы для ограниченных потенциалов, что создает вычислительные трудности для описания процесса плавления молекулы.

Для анализа макроскопических свойств, таких как теплоемкость, необходимо проводить усреднение по большому количеству систем (реализаций).

На рис. 2 показана усредненная по N=512 реализациям температурная зависимость удельной полной энергии  $E(T)$ , приходящаяся на один сайт, цепочки длиной в 100 нуклеотидных пар.

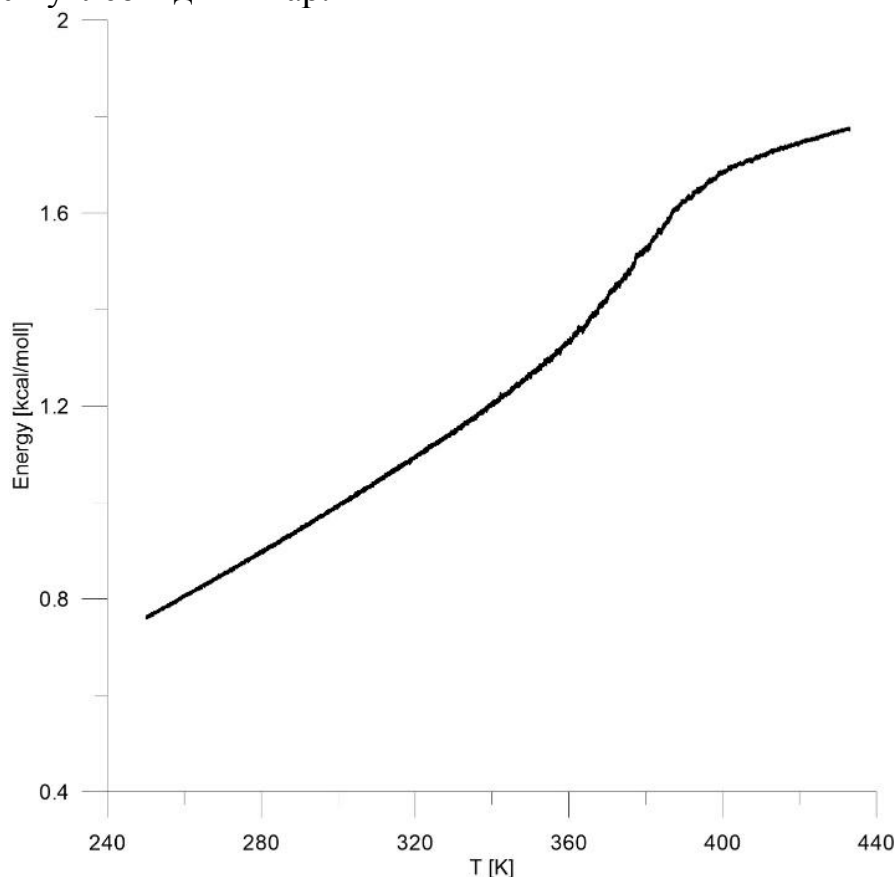


Рис. 2. Зависимость усредненной по 512 реализациям удельной полной энергии от температуры

На рис. 3 приведена температурная зависимость удельной теплоемкости  $C_V = \frac{dE}{dT}$ , для четырех различных систем, усреднённых по 512 реализациям: из 25 частиц ( $\rho=0.5$ ), из 50 частиц ( $\rho=0.5$ ), из 100 частиц ( $\rho=0.5$ ) и из 100 частиц ( $\rho=2$ ). Для поиска производной функция полной энергии была аппроксимирована двумя прямыми (до и после перехода), а также полиномом 5-ой степени в области перехода. После аппроксимации производная находилась аналитически. Границы областей аппроксимации находились при помощи метода минимизации, чтобы полученная кусочно-заданная аналитическая функция имела наименьший квадрат отклонения от данных моделирования.



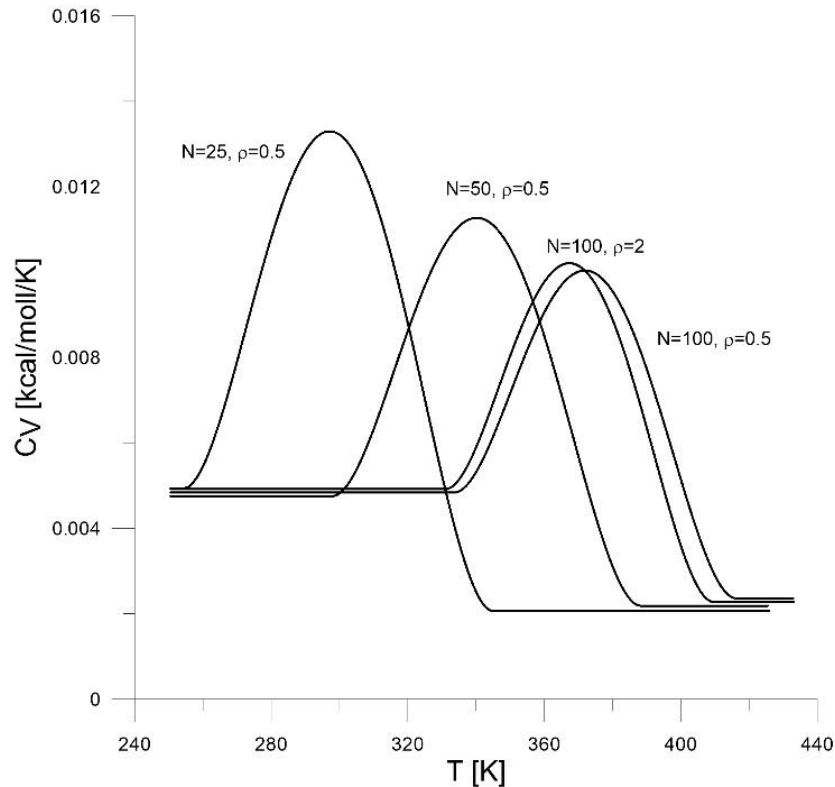


Рис. 3. Удельная теплоемкость как производная полной энергии для четырех систем

Кривые на рис. 3 описывают процесс плавления ДНК, которому соответствует широкий пик теплоёмкости. Величина теплоёмкости до пика больше величины теплоёмкости после пика, поскольку слева от пика подводимое к цепочке тепло расходуется как на увеличение потенциальной энергии между соседними уотсон-криковскими парами ДНК (стэкинг-взаимодействие), так и на увеличение потенциальной энергии каждой из пар, плюс на кинетическую энергию. Справа от пика, когда цепочка полностью денатурирована, ее дальнейший нагрев на потенциальную энергию между нуклеотидами в паре не расходуется. Численно значение слева от пика теплоемкости превосходит две универсальные газовые постоянные ( $2R$ ), а справа от пика стремится к  $R$ . До пика мы имеем два потенциала, близких к параболическим: потенциал Морзе при низких температурах близок к параболическому, а потенциал стэкинг-взаимодействия уже близок к параболическому за счет стремления экспоненты к нулю.

Температура, при которой пик достигает своего максимального значения, соответствует точке плавления ДНК. В этой точке полная энергия однородного дуплекса равна  $ND$ , где  $D$  – константа потенциала Морзе.

В работах [2, 8, 9], в отличие от рис. 3 и от работы [10], были получены скачкообразные кривые теплоемкости. В областях температур, близких к критическим, в модели ПБД кинетическая энергия, полученная от термостата, идёт на увеличение потенциальной энергии межнуклеотидного взаимодействия. Этим объясняется плавный спад теплоемкости после пика. Используемый в [2]

для расчета  $C_V(T)$  метод трансфер-интегралов (ТИ) приводит к  $\delta$ -образному пику, который соответствует результатам моделирования очень длинных цепочек. В отличие от метода ТИ-интегралов, рассмотренный нами подход может быть использован для цепочек любой длины и состава нуклеотидов. Кроме того, метод трансфер-интеграла ограничен моделью Пейрарда-Бишопа. Уже простейшее усложнение модели, например, рассмотрение модели ПБД не допускает аналитических решений соответствующих интегральных уравнений, делает невозможным расчёт реальных систем.

В работе [11] для расчета термодинамических функций был использован прямой метод вычисления статсуммы как многомерного интеграла. Как правило, такой подход оказывается сильно ограниченным длиной цепочки. Самое серьёзное ограничение этого метода, однако, состоит в том, что он позволяет исследовать только статические термодинамические состояния, в то время как использованный нами метод позволяет изучать динамические статистические свойства системы.

Полученный нами пик теплоёмкости зависит от типа нуклеотидной последовательности, что находится в согласии с [12]. При меньшем количестве частиц фазовый переход возникает при меньшей температуре. Данный вывод согласуется также с работой [13], где рассматривалась несколько измененная модель, которая в области малых колебаний сходна с исследуемой.

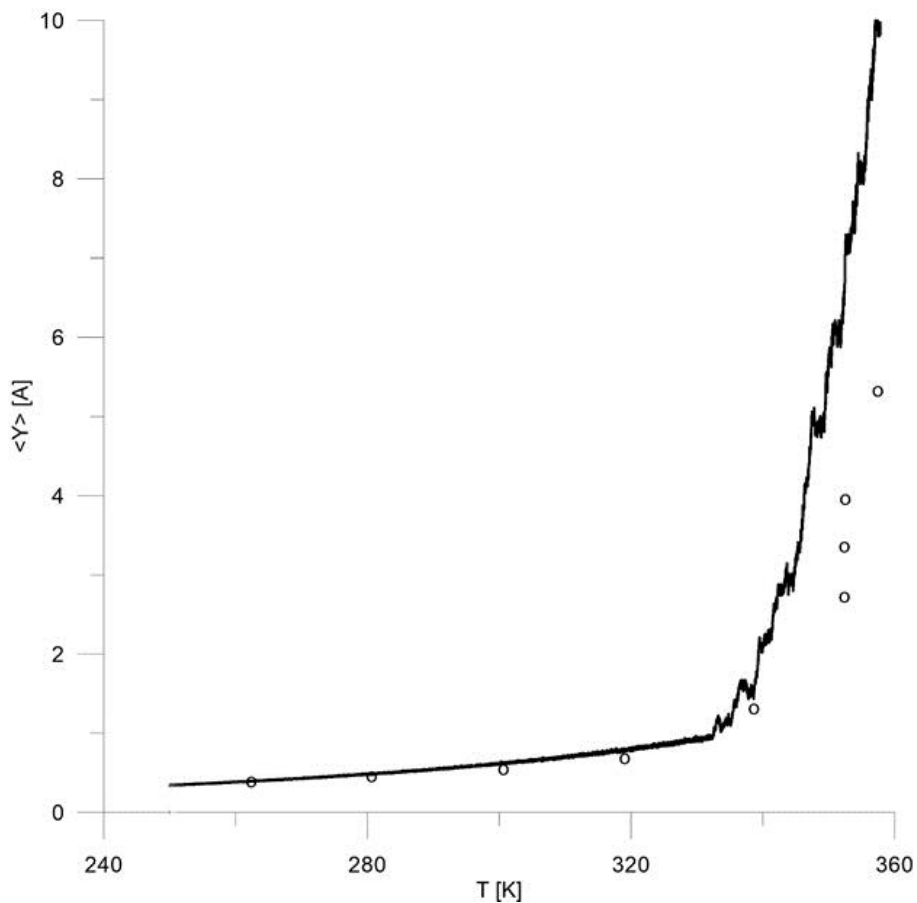
Сравнение результатов, полученных данным методом и ТИ-методом статьи [2], приведено на рис. 4, где изображена зависимость среднего значения координаты смещений нуклеотидной пары из равновесного положения по всем реализациям от температуры. Степень совпадения достаточно велика, что говорит о корректности проведенных расчетов. Некоторые отличия можно объяснить тем, что в статье [2] использовался термостат Нозе-Нувера, а в данной работе – столкновительный.

Термостаты Берендсена [14], Нозе и Нозе-Нувера [15] могут приносить некоторые артефакты в исследования. Все эти термостаты хорошо устанавливают среднюю температуру. Однако при появлении одной сильно горячей частицы (равно как и сильно холодной) эти термостаты для поддержания средней температуры охлаждают (разогревают) остальную систему, которая находится при температуре, близкой к заданной. Спустя некоторое время температуры всех частиц системы уравниваются благодаря межчастичному взаимодействию, но не благодаря термостату. Если говорить о потенциале Морзе, возникновение холодных частиц возможно при попытке выхода из потенциальной ямы. Если частица не скатывается обратно в потенциальную яму, то она имеет скорость ниже, нежели в квазигармоническом режиме.

В отличие от термостата Берендсена и Нозе-Нувера столкновительный термостат воздействует на каждую частицу индивидуально, независимо от температуры остальных частиц системы, что более соответствует реальной ситуации.

Обращаясь к работе [2], мы обращаем внимание на то, что у нас практически совпадают средние значения координат, а теплоемкость – нет. Это объясняется тем, что в работе [2] координаты посчитаны методом моделирования молекулярной динамики только до момента денатурации, а теплоемкость приводится только по методу трансфер-интегралов. В нашей работе был использован метод моделирования молекулярной динамики как до денатурации, так и после.

Полученные нами результаты, по-видимому, являются достаточно общими и могут быть распространены на другие системы. В частности, широкий пик теплоемкости экспериментально наблюдался в белках [16].



*Рис. 4.* Зависимость среднего значения координаты от температуры (черная кривая). Длина цепочки 100 нуклеотидов, количество реализаций 100. Кружками отмечены данные молекулярного моделирования с использованием термостата Нозе-Гувера из источника [2]

## Благодарности

Авторы выражают благодарность Балабаеву Николаю Кирилловичу за предложения по методике проведения экспериментов, а также Фиалко Надежде Сергеевне за критические замечания.

Расчеты проводились на суперкомпьютерном комплексе Ломоносов Московского государственного университета [17], а также на суперкомпьютерах k100 и k60 ИПМ им. М.В.Келдыша РАН.

## Используемая литература

1. Wartell R.M., Benight A.S. Thermal denaturation of DNA molecules: A comparison of theory with experiment, *Physics Reports* 126 (1985) 67-107.
2. Dauxois T., Peyrard M., Entropy-driven DNA Denaturation, *J. Phys. Rev. E.* 47 (1993) 44-47.
3. Swope W.G., Andersen H.C., Berens P.H., Wilson K.R., A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters, *J. Chem. Phys.* 76, 637 (1982) 637-649.
4. Hoover W.G., Ladd A.J.C., Moran B., Analytic and numerical surface dynamics of the triangular lattice, *Phys. Rev. Lett.* 48 (1982) 1818-1820.
5. Балабаев Н.К., Лемак А.С., Молекулярная динамика линейного полимера в гидродинамическом потоке, *Журн. Физ. Химии.* 69 (1995) 28-32.
6. Lemak A.S., Balabaev N.K., Comparison between collisional dynamics and Brownian dynamics *Molecular Simulation*, 15 (1995) 223-231.
7. Lemak A.S., Balabaev N.K., Molecular dynamics simulation of polymer chain in solution by collisional dynamics method, *J. Comput. Chem.* 17 (1996) 1685-1695.
8. Dauxois T., Payrard M. // *J. Phys. Rev. E.* 1995, Vol. 51 N 5. P 4027-4040.
9. Zoli M., Path integral method for DNA denaturation. *J. Phys. Rev. E.*, 79 (2009) 041927-7.
10. Vaitiekunas, Paulius, Colyn Crane-Robinson, Privalov P.L. The Energetic Basis of the DNA Double Helix: A Combined Microcalorimetric Approach, *Nucleic Acids Research* 43.17 (2015) 8577–8589.
11. Srivastava S., Singh N. The probability analysis of opening of DNA. *J Chem Phys.* 21;134(2011):115102.
12. Marmur J., Doty P. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *Journal of Molecular Biology.* 1962. V. 5. Iss. 1. P. 109–118.
13. Joyeux M., Buyukdagli S. A dynamical model based on finite stacking enthalpies for homogeneous and inhomogeneous DNA thermal denaturation, *Phys. Rev. E. Stat. Nonlin. Soft. Matter Phys.* 2005 Nov;72(5 Pt 1):051902. Epub 2005 Nov 1.

14. Berendsen H. J. C., Postma J. P. M., Gunsteren W. F. van, Di Nola A., Haak J. R. J. *Chem. Phys.* 1984. Vol. 81. P. 3684–3690.
15. Hoover W. G. *Phys. Rev. A.* 1985. Vol. 31. P. 1695–1697.
16. Privalov P.L., Khechinashvili N.N., A thermodynamics approach to the problem of stabilization of globular protein structure: a calorimetry study, *J. Mol. Biol.* 86 (1974) 665-684.
17. Sadovnichy V., Tikhonravov A., Voevodin V.I., and Opanasenko V. "Lomonosov": Supercomputing at Moscow State University. In *Contemporary High Performance Computing: From Petascale toward Exascale* (Chapman & Hall/CRC Computational Science), pp.283-307, Boca Raton, USA, CRC Press, 2013.