



Научная статья

DOI: 10.15593/RZhBiomeh/2022.4.03

УДК 531/534: [57+61]

ИССЛЕДОВАНИЕ ПРОЦЕССА ДЕНАТУРАЦИИ ДНК НА ОСНОВЕ МОДЕЛИ ПЕЙРАРА – БИШОПА – ДОКСУА И КОЛИЧЕСТВЕННОГО АНАЛИЗА РЕКУРРЕНТНЫХ ДИАГРАММ

А.С. Никитюк, О.С. Бурмистрова, О.Б. Наймарк

Лаборатория физических основ прочности Института механики сплошных сред УрО РАН, Пермь, Россия

О СТАТЬЕ

Получена: 12 августа 2022
Одобрена: 01 декабря 2022
Принята к публикации: 03 декабря 2022

Ключевые слова:

ДНК, открытые состояния, динамика открытых состояний, денатурация, фазовый переход, рекуррентная диаграмма, количественный анализ рекуррентных диаграмм.

АННОТАЦИЯ

Целью работы является исследование фазового перехода при термической денатурации ДНК методом количественного анализа рекуррентных диаграмм, являющегося наиболее современным и эффективным подходом для анализа сложных систем. Для достижения цели реализуется нелинейная динамическая модель процесса денатурации ДНК, основанная на подходе Пейрара – Бишопа – Доксуа, которая позволяет определять состояние пар оснований ДНК в произвольный момент времени. Далее проводится моделирование динамики ДНК при различных температурах, исследуется эволюция растяжения водородных связей при нагревании ДНК. Для анализа полученных результатов используется метод количественного анализа рекуррентных диаграмм, которые позволяют судить о характере протекающих в системе процессов и представляют особый интерес применительно к изучению нелинейной динамики ДНК, в частности фазовых переходов, поскольку на основе их анализа можно выявить состояния повторения и замирания (ламинарности), совершение экстремальных событий, наличие скрытой периодичности и цикличности. С помощью данного метода анализа производится расчет количественных мер в окне, сдвинутом вдоль главной диагонали рекуррентной диаграммы. Данные меры количественно определяют мелкомасштабные структуры рекуррентной диаграммы. При расчетах выявляются наиболее чувствительные к анализу исследуемого процесса меры, с помощью которых определяются предвестники фазового перехода.

© ПНИПУ

Введение

Плавлением ДНК (термической денатурацией) называется инициированный нагревом процесс разрушения комплементарных водородных связей (обусловленных, так называемым стэкинг-взаимодействием) в ДНК при сохранении ковалентных связей. Как показано в ряде экспериментальных и теоретиче-

ских работ [2; 6; 8; 9; 14; 15; 17; 20; 21], плавление ДНК сопровождается структурными превращениями, имеющими природу фазовых переходов первого или второго рода. Исследования возможности существования фазовых переходов в одномерных системах, к которым относится ДНК, является актуальной проблемой теоретической физики [5; 10]. Впервые наличие фазового перехода в растворе гомогенной ДНК (т.е. моле-

© Никитюк Александр Сергеевич – к.ф.-м.н., e-mail: nas@icmm.ru ID: 0000-0001-6777-7525
© Бурмистрова Ольга Сергеевна – инженер, e-mail: ozernykh.o@icmm.ru ID: 0000-0002-9787-6185
© Наймарк Олег Борисович – д.ф.-м.н., зав. лаб., e-mail: naimark@icmm.ru ID: 0000-0001-6537-1177



Эта статья доступна в соответствии с условиями лицензии Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

кулах, состоящих из последовательности одинаковых пар оснований, нуклеотидов) экспериментально было установлено при измерении поглощения света с длиной волны 260–268 нм при медленном нагреве [8]. Дальнейшие исследования нативных ДНК молекул показали, что денатурация происходит в несколько этапов и очень чувствительна к составу последовательности [21]. Теоретические исследования, подтвердившие возможность фазовых переходов в ДНК, были выполнены в работах *Poland et al.* [14], *Scheraga et al.* [15], *Causo et al.* [2], Пейраром и соавт. [3; 4]. При этом наиболее эффективной оказалась динамическая модель Пейрара – Бишопа – Доксуа и ее модификации, уточняющие описание процесса денатурации ДНК, многообразие которых подробно описано в обзоре [18]. Метод количественного анализа рекуррентных диаграмм (РД) применяется для изучения нелинейных процессов, характерных в том числе для биологических систем. Рекуррентные диаграммы позволяют судить о характере протекающих в системе процессов и представляют особый интерес применительно к изучению нелинейной динамики ДНК, в частности фазовых переходов, так как на основе их анализа можно выявить состояния «повторения и замирания» (*recurrence and laminarity states*), совершение экстремальных событий, наличие скрытой периодичности и цикличности. Количественный анализ рекуррентных диаграмм позволяет определить «численные меры», основанные на плотности рекуррентных точек, распределениях длин диагональных, горизонтальных или вертикальных линий, характеризующих эволюцию динамической системы [16; 22]. Целью работы является исследование фазового перехода в ДНК, происходящего при термической денатурации, методом количественного анализа рекуррентных диаграмм. Данная работа состоит из пяти частей. Во введении обсуждаются вопросы актуальности и новизны исследования процесса денатурации ДНК на основе модели Пейрара – Бишопа – Доксуа и количественного анализа рекуррентных диаграмм. Первый раздел посвящен описанию нелинейной динамической модели процесса денатурации ДНК, основанной на подходе Пейрара – Бишопа – Доксуа. Во втором разделе представлены результаты моделирования процесса денатурации ДНК на основе модели Пейрара – Бишопа – Доксуа. Третий раздел посвящен изучению результатов моделирования методом количественного анализа рекуррентных диаграмм. В заключении обсуждаются полученные результаты, делаются выводы и обозначаются перспективы дальнейших исследований.

Нелинейная динамическая модель процесса денатурации ДНК

Оригинальная модель Пейрара – Бишопа – Доксуа была предложена в работах [3; 4] для описания коллективных движений пар оснований, включая флуктуации

большой амплитуды. Модель Пейрара – Бишопа – Доксуа является развитием модели Пейрара – Бишопа, (ПБ-модель), предложенной для описания денатурации ДНК. В ПБ-модели взаимодействие между парами оснований (водородные связи) описывается потенциалом Морзе; стэкинг-взаимодействия вдоль цепи учитываются через гармонический потенциал. Полная энергия системы имеет вид

$$H = E_{kin} + U_{hyd} + U_{stack}, \quad (1)$$

где E_{kin} – кинетическая энергия, U_{hyd} – потенциал Морзе, U_{stack} – потенциал стэкинг-взаимодействия. На рис. 1 приведено схематичное представление динамической модели нелинейной динамики ДНК с указанием водородных связей U_{hyd} и стэкинг-взаимодействий U_{stack} .

Кинетическая энергия вычисляется согласно следующему соотношению

$$E_{kin} = \sum_n \frac{p_n^2}{2m}, \quad (2)$$

где m – масса пары оснований, p_n – импульс пары оснований, n – количество пар оснований.

Потенциальная энергия представляет собой сумму потенциала Морзе, который описывает энергию взаимодействия водородных связей при изменении расстояния между парами оснований и стэкинг-взаимодействий между соседними парами оснований вдоль цепи.

Потенциал Морзе имеет следующий вид

$$U_{hyd} = \sum_n D(e^{-\alpha y_n} - 1)^2, \quad (3)$$

где D – энергия диссоциации пары оснований, α – пространственный масштаб потенциала, y_n – переменная, описывающая динамику оснований внутри пары.

Стэкинг-взаимодействия между соседними основаниями ДНК описываются гармоническим соотношением

$$U_{stack}^h = \sum_n \frac{1}{2} k (y_n - y_{n-1})^2, \quad (4)$$

где k – параметр гармонического взаимодействия между соседними парами оснований, h – индекс, обозначающий гармоническую связь.

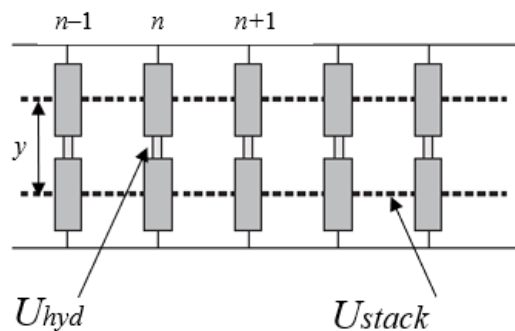


Рис. 1. Динамическая модель нелинейной динамики ДНК [3]

Для описания движения оснований рассматривается производная от функции Гамильтона:

$$y_n = \frac{\partial H}{\partial p_n} = \frac{p_n}{m}. \quad (5)$$

Откуда можно получить соотношение для импульса основания

$$p_n = m\dot{y}_n. \quad (6)$$

Для адекватного моделирования необходимо учитывать термические флуктуации ДНК, поскольку ДНК находится в ядрышке клетки при определенной температуре, определяемой окружением.

Один из способов учесть температуру – это введение термостата Нозе – Гувера, который будет симулировать реальные условия. Так, вводится новая переменная ξ , представляющая собой термодинамический коэффициент трения $\xi = sp_s/M$, где s – дополнительная степень свободы, действующая как внешняя сила на систему из n частиц.

Таким образом

$$\frac{ds}{dt} = \frac{\partial H}{\partial p_s} = \frac{p_s}{M}, \quad (7)$$

где M – масса термостата. Производная от сопряженного импульса имеет вид:

$$\frac{dp_s}{dt} = -\frac{\partial H}{\partial s} = \left(\sum_n \frac{p_n^2}{m \cdot s^2} - n \cdot k_B \cdot T \right) / s. \quad (8)$$

где k_B – константа Больцмана, T – температура термостата. Тогда уравнение движения ДНК будет иметь вид

$$\begin{cases} m\ddot{y}_n = k(y_{n-1} + y_{n+1} - 2y_n) + 2aD(e^{-ay_n} - 1)e^{-ay_n} - \xi m\dot{y}_n \\ \dot{\xi} = \frac{1}{M} \left(\sum_n m\dot{y}_n^2 - n \cdot k_B \cdot T \right). \end{cases} \quad (9)$$

Переменная ξ стабилизирует кинетическую энергию реальной системы до канонического математического ожидания $nk_B T$. Модель Пейрара – Бишопа позволила сравнить результаты моделирования с экспериментом, но численные расчеты давали значения температуры плавления приблизительно на 150 К выше экспериментальных данных. В связи с этим М. Пейраром, А. Бишопом и Т. Доксуа был введен ангармонический потенциал, описывающий нелинейный характер стэкинг-взаимодействия:

$$U_{stack}^a = \frac{k}{2} (1 + \rho e^{-\chi(y_n + y_{n-1})}) (y_n - y_{n-1})^2, \quad (10)$$

где ρ – безразмерный параметр, через который вводится нелинейность процесса стэкинг-взаимодействия, χ – коэффициент затухания, a – индекс, обозначающий ангармоничность потенциала.

Следствием введенного «ангармонизма» потенциала явилось резкое увеличение расстояния между молекулами ДНК при возрастании температуры, что является качественной особенностью модели при описании динамики ДНК. Последнее предполагает учет роли термических флуктуаций на развитие нелинейных процес-

сов в ДНК. Один из способов учета температуры – это введение термостата Нозе – Гувера [19].

С учетом ангармонизма потенциала уравнение движения ДНК принимает следующий вид:

$$\begin{cases} m\ddot{y}_n = k(y_{n-1} + y_{n+1} - 2y_n)(1 + \rho e^{-\chi(y_{n-1} + y_{n+1} + 2y_n)}) + 2aD(e^{-ay_n} - 1)e^{-ay_n} - \xi m\dot{y}_n \\ \dot{\xi} = \frac{1}{M} \left(\sum_n m\dot{y}_n^2 - n \cdot k_B \cdot T \right), \end{cases} \quad (11)$$

где k_B – константа Больцмана, T – температура термостата, M – масса термостата.

Таким образом, модель Пейрара – Бишопа – Доксуа представляет собой нелинейную динамическую модель денатурации ДНК, позволившую привести в соответствие результаты расчетов с экспериментальными данными.

Моделирование процесса денатурации ДНК

Несмотря на то, что существуют модификации модели Пейрара – Бишопа – Доксуа, учитывающие гетерогенность участка ДНК [1; 7; 12], в данной работе рассматривается гомогенный случай.

С использованием модели Пейрара – Бишопа – Доксуа проведено численное моделирование процесса нагрева ДНК в диапазоне от 200 до 540 К для 256 пар оснований. На рис. 2 представлена эволюция растяжения водородных связей при нагревании.

На данной диаграмме изображена динамика сегмента ДНК. Случайный характер флуктуаций связан с видом начальных условий, а именно равномерным случайным распределением открытых и закрытых состояний пар оснований. Белые пятна на рисунках соответствуют областям, где растяжение между парами оснований на 1,2 Å и более превосходит допустимое значение, равное 10 Å, в течение ограниченного времени, когда пара оснований считается открытой.

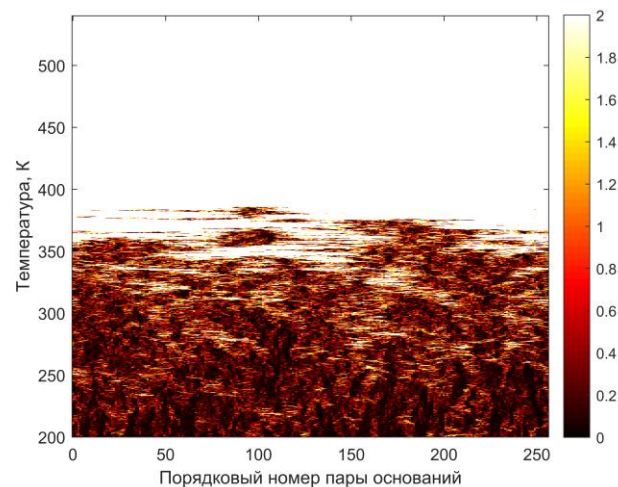


Рис. 2. Эволюция растяжения водородных связей при нагревании ДНК. Черный цвет соответствует случаю, когда пара оснований закрыта, белый – когда открыта

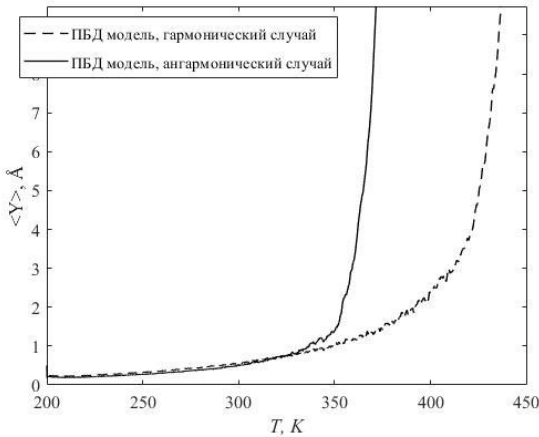


Рис. 3. Диаграмма плавления цепи ДНК

Черные области, в свою очередь, обозначают области, где растяжение меньше допустимого, и пара оснований предполагается закрытой. Рассматривая каждую пару оснований отдельно, можно заметить значительные различия амплитуд растяжения. Темные области разделяются белыми областями, и подобное явление объясняется появлением пузырьков денатурации, обнаруженных экспериментально. При увеличении температуры до максимальной становится всё больше белых областей, это означает, что ДНК полностью денатурирует, то есть связи разрываются.

На рис. 3 приведена осредненная по 1000 реализаций диаграмма плавления ДНК. Сплошная линия соответствует ангармонической модели, а пунктирная – гармонической. Можно заметить, что температуры плавления, полученные на основе данных моделей, различаются.

Данный результат согласуется с результатами эксперимента и оригинальной работой [4], в которой предложена модель.

Количественный анализ рекуррентных диаграмм процесса денатурации ДНК

Рекуррентный анализ является современным и быстро развивающимся подходом к анализу сложных систем, а также является мощным инструментом для исследования повторяемости в динамических системах. С помощью данного метода можно охарактеризовать поведение системы в фазовом пространстве. Среди использующих данный метод исследований, имеется большое количество работ в области анализа нелинейной динамики биологических систем. С помощью рекуррентных диаграмм (РД) появляется возможность анализа динамических процессов при наличии состояний «повторения и замирания», влияния шума. Количественный анализ РД позволяет сопоставить полученной диаграмме некоторые количественные меры, основанные на распределениях длин линий и плотности, так называемых, рекуррентных точек. Главными преимуществами количественного анализа РД являются наглядность полученных результатов, а также возможность извлечь полезную информацию

даже для коротких и нестационарных данных, когда другие методы нелинейного анализа терпят неудачу, например, метод фазовой плоскости, метод марковских процессов или метод кусочно-линейной аппроксимации [12]. Количественный анализ применим к различным видам данных и широко используется при решении инженерных задач, физиологии, геофизики, экономике и медицине [11; 12].

Для количественного анализа рекуррентных диаграмм процесса денатурации ДНК необходимо рассчитать рекуррентные матрицы, являющиеся отображением траектории эволюции ДНК в m -мерном фазовом пространстве (в работе будет использовано одномерное фазовое пространство, соответствующее динамике одной пары оснований) на двумерную двоичную матрицу размером $N \times N$. Единица в ячейке матрицы соответствует проходу траектории в окрестности одной и той же точки фазового пространства при некотором времени i в некоторое другое время j . Обе координатные оси диаграммы являются осями времени. Математически это можно выразить следующим образом

$$R_{i,j}^{m,\epsilon_i} = \Theta\left(\epsilon_i - \|\vec{y}_i - \vec{y}_j\|\right), i, j = 1 \dots N, \quad (12)$$

где y_i – переменная, в случае данного исследования описывающая динамику оснований внутри i -й пары, N – количество рассматриваемых состояний y_i , ϵ_i – размер окрестности точки y_i , Θ – функция Хэвисайда, m – размерность фазового пространства, $\|\cdot\|$ – Евклидова норма.

Размер окрестности ϵ_i определяет радиус в фазовом пространстве с центром в точке y_i . Если точка y_j попадает внутрь данной окрестности, то такое состояние считается подобным состоянию y_i , и, таким образом, на диаграмме устанавливается $R_{i,j} = 1$ ($i=j$). Если значение ϵ будет малым, то малой будет последовательность точек повторения, и следовательно, «рекуррентная структура» исследуемой системы будет неясна. Если же значение будет выбрано слишком большим, тогда почти каждая точка будет являться соседней для любой другой точки, что приведет к ошибочным представлениям о структуре системы [13]. Поэтому для анализа динамики пары оснований были рассмотрены различные ϵ , а именно $\epsilon \geq 5\sigma$, $\epsilon \geq 0,5\sigma$, $\epsilon = \sigma$, где σ – стандартное отклонение сигнала. На основе качественного анализа рекуррентных диаграмм было выявлено, что наиболее оптимальным является $\epsilon = \sigma$.

На рис. 4, б, приведена рекуррентная диаграмма динамики одной пары оснований, где по оси абсцисс указано время (относительные временные единицы (о.в.е.)), а по оси ординат – расстояние внутри пары. Динамика растяжения данной пары отражена на рис. 4, а. Для построения данной диаграммы были приняты следующие параметры: $N = 100$, $\epsilon = 0,31$.

Стоит отметить, на диаграмме присутствует черная диагональная линия, которая называется линией идентичности. Также стоит заметить, рекуррентная диаграмма симметрична относительно линии идентичности, то есть $R_{i,j} \equiv R_{j,i}$.

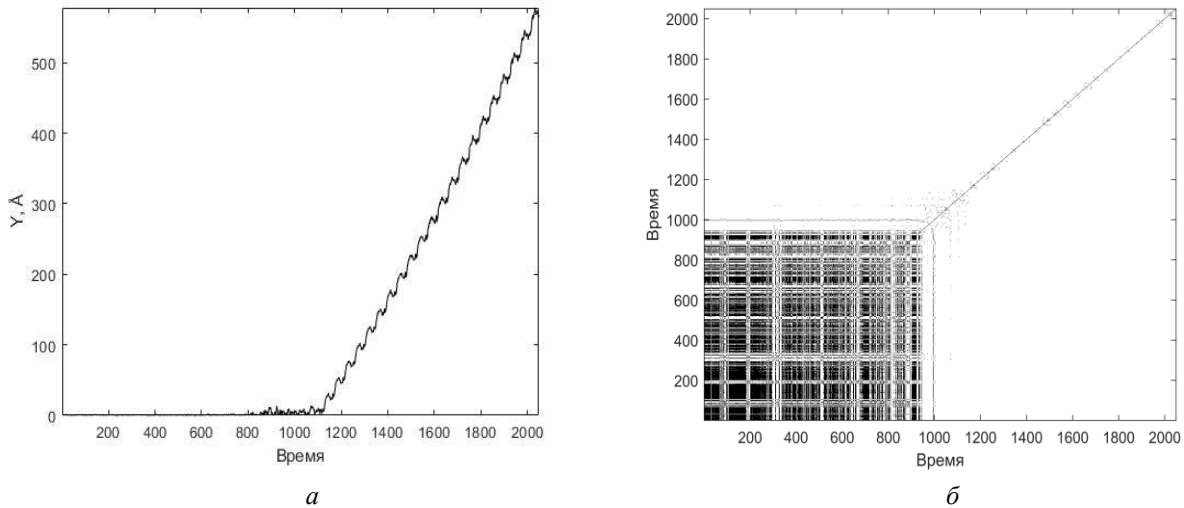


Рис. 4. Анализируемая пара оснований: *a* – динамика растяжения пары оснований; *б* – рекуррентная диаграмма динамики, по оси абсцисс и ординат – время, в течение которого происходит нагрев (о.в.е.)

С точки зрения топологии и текстуры можно выделить несколько видов рекуррентных диаграмм. Если на диаграмме видны белые полосы, это означает нестационарность данных, некоторые состояния редки или далеки от нормы. Если прослеживаются единичные точки, значит в системе происходят сильные колебания. Возникновение отдельных изолированных точек означает, что процесс может быть некоррелированным, случайным или антикоррелированным. Наличие вертикальных/горизонтальных полос показывает, что некоторые состояния не меняются или изменяются медленно в течение некоторого времени; индикация ламинарных состояний [13].

Проанализируем рекуррентную диаграмму (РД), представленную на рис. 4, б. В окрестности 900 относительных временных единиц (о.в.е.) происходит смена режима.

Проанализируем часть диаграммы до данной точки с точки зрения топологии и текстуры. На диаграмме четко видны белые и черные вертикальные полосы. Их наличие означает нестационарный характер динамики пары оснований ДНК, система не находится в состоянии равновесия. При этом стоит отметить, что на коротких промежутках времени динамика пары оснований практически не меняет свой характер. Чем ближе к отметке 900 о.в.е., тем заметнее становятся единичные изолированные точки, это означает, что в системе происходят сильные колебания. На основе анализа топологии и текстуры рекуррентной диаграммы можно предположить, что наиболее чувствительными к анализу плавления ДНК являются меры, основанные на вертикальных/горизонтальных структурах, по причине явного выделения на диаграмме вертикальных линий.

Для перехода к количественному анализу рекуррентных диаграмм необходимо ввести меры, которые количественно определяют мелкокомасштабные структуры рекуррентных диаграмм. Данные меры основаны на плотности

точек рекуррентности или структурах линий на диаграммах. Меры могут быть сгруппированы по структурам, на основе которых они рассчитаны. К мерам, основанным на плотности точек, относится мера рекуррентности (*RR*). Меры, рассчитанные на основе диагональных структур: мера дивергенции (*DIV*), мера детерминизма (*DET*), мера отношения (*RATIO*), мера средней длины диагональной линии (*L*) и мера энтропии (*ENTR*). Меры, основанные на вертикальных/горизонтальных линиях: мера замирания (*LAM*), мера времени запаздывания (*TT*) и мера максимальной длины вертикальной линии (*V_{max}*).

Мера рекуррентности определяется как

$$RR(\epsilon) = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j}(\epsilon). \quad (13)$$

На рис. 5 представлен результат расчета меры рекуррентности на основе РД (см. рис. 4, б). Данная мера показывает плотность рекуррентных точек, определяемая их количеством.

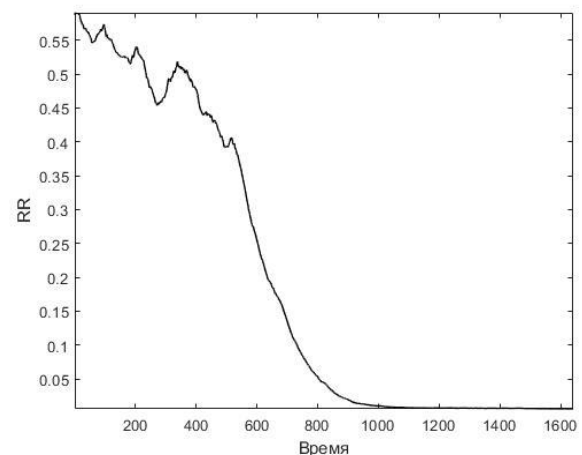


Рис. 5. Мера рекуррентности, основанная на РД растяжения пары оснований

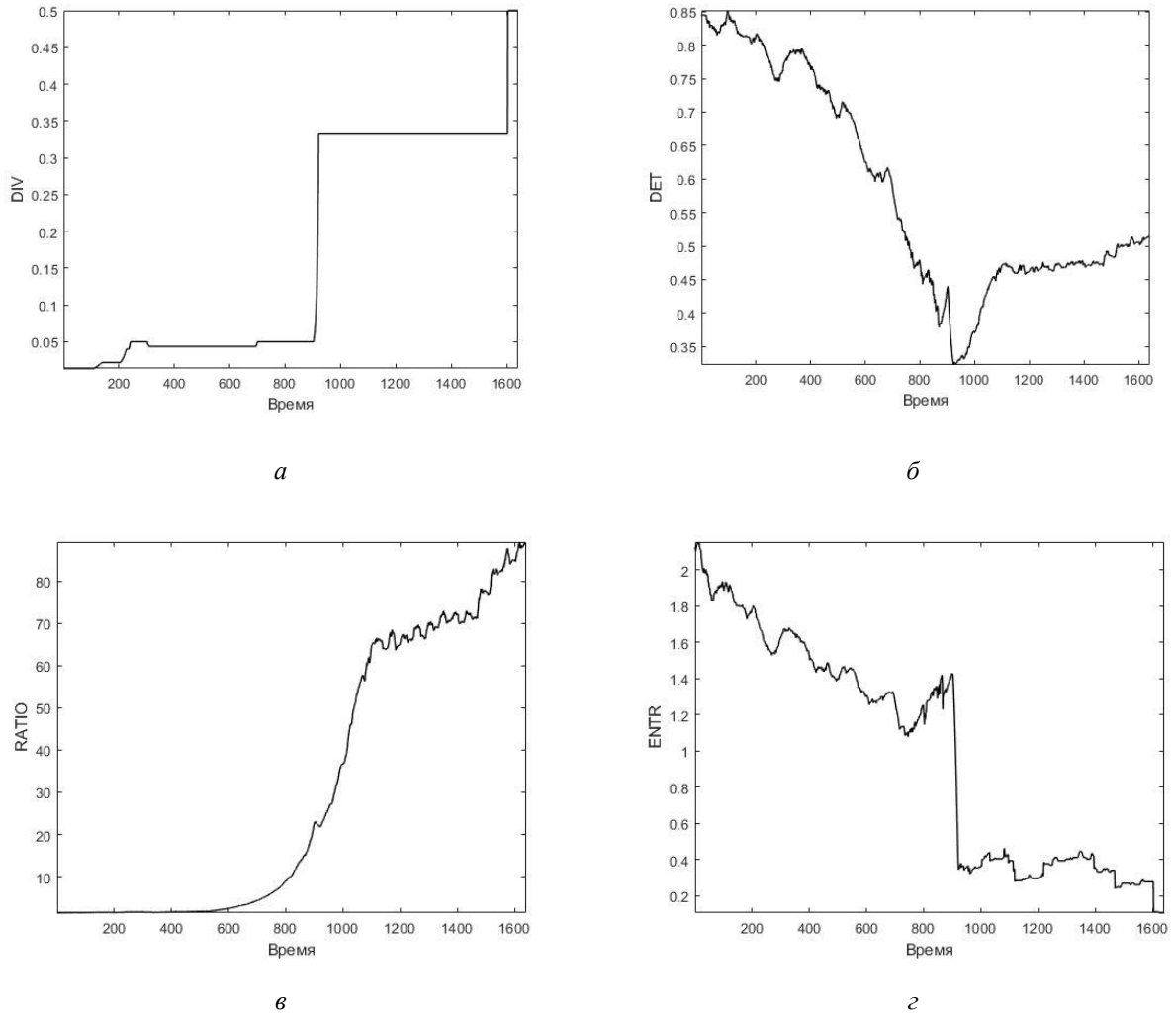


Рис. 6. Результаты расчета мер, основанные на РД растяжения пары оснований: *a* – мера дивергенции; *б* – мера детерминизма; *в* – мера отношения; *г* – мера энтропии

Рассмотрим меры, основанные на диагональных структурах. Мера дивергенции задается следующим соотношением

$$DIV = \frac{1}{L_{max}}, \quad (14)$$

где L_{max} – максимальный размер диагональной линии.

Мера детерминизма имеет вид

$$DET = \frac{\sum_{l=l_{min}}^N IP(l)}{\sum_{l=1}^N IP(l)}, \quad (15)$$

где l – длина диагональной линии, l_{min} – минимальная длина диагональной линии, $P(l)$ – количество диагональных линий размера l . Меры отношения и энтропии могут быть определены согласно

$$RATIO = N^2 \frac{\sum_{l=l_{min}}^N IP(l)}{(\sum_{l=1}^N IP(l))^2}, \quad (16)$$

$$ENTR = - \sum_{l=l_{min}}^N p(l) \ln p(l), \quad (17)$$

$$p(l) = \frac{P(l)}{\sum_{l=l_{min}}^N P(l)},$$

$$L = \frac{\sum_{l=l_{min}}^N lP(l)}{\sum_{l=l_{min}}^N P(l)}. \quad (18)$$

На рис. 6 представлены результаты расчета мер, основанных на диагональных структурах, по РД динамики пары оснований ДНК (см. рис. 4, б). Мера дивергенции основана на определении размера самой длинной диагональной линии рекуррентной диаграммы. Данная мера связана с расходимостью траектории фазового пространства. Чем больше значение данной меры, тем быстрее расходятся сегменты траектории.

Мера детерминизма отражает предсказуемость поведения ДНК. Диагональные структуры показывают время, в течение которого участок фазовой траектории подходит близко к другому участку фазовой траектории. Мера

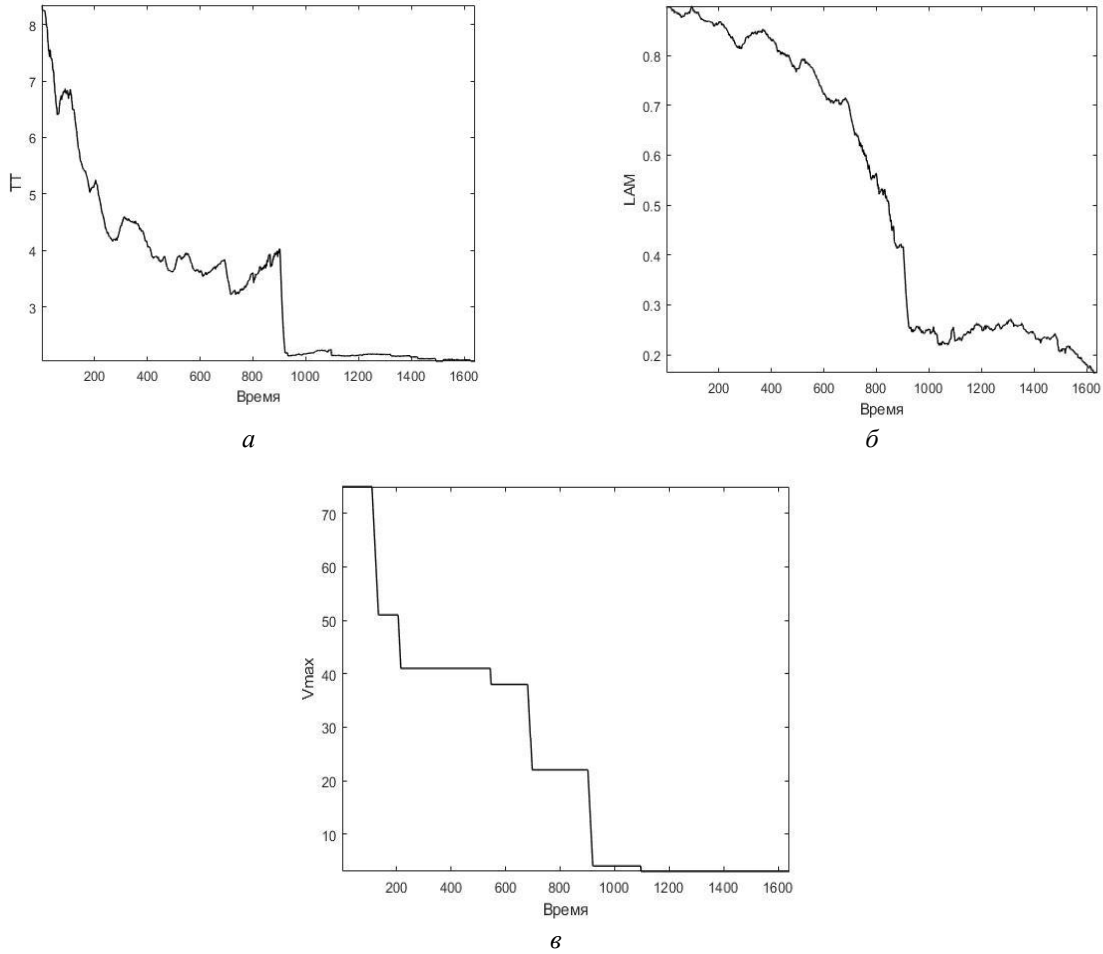


Рис. 7. Результаты расчета мер, основанные на РД растяжения пары оснований: *а* – мера времени запаздывания; *б* – мера замирания; *в* – мера максимальной длины вертикальной линии

отношения помогает выявлять переходы в динамике исследуемых молекул.

Мера энтропии отражает вероятность $p(l)$ нахождения диагональной линии длины l в рекуррентной диаграмме. Данная мера количественно определяет информативность случайных процессов.

Также мера энтропии показывает сложность рекуррентной диаграммы по отношению к диагональным линиям в системе.

Для диаграмм, в которых преобладают линии одинаковой длины, значение меры энтропии будет низкое, в ином случае, когда линии разных длин появляются одинаково часто, $ENTR$ будет иметь довольно высокое значение.

Мера времени запаздывания, основанная на вертикальных/горизонтальных структурах имеет вид

$$TT = \frac{\sum_{v=v_{min}}^N vP(v)}{\sum_{v=v_{min}}^N P(v)}, \quad (19)$$

где v – длина вертикальной линии, v_{min} – минимальная длина вертикальной линии, $P(v)$ – количество диагональных линий длины v . Мера ламинарности определяется как

$$LAM = \frac{\sum_{v=v_{min}}^N vP(v)}{\sum_{i,j}^N RP_{i,j}^e}, \quad (20)$$

где знаменатель определяет общее количество точек в рекуррентной диаграмме.

Максимальная длина вертикальной линии на РД задается

$$v_{max} = \max(\{v_l\}_{l=1}^{N_v}), \quad (21)$$

где N_v – абсолютное количество вертикальных линий.

На рис. 7 представлены результаты расчета мер, основанных на вертикальных структурах, по РД динамики пары оснований ДНК (см. рис. 4, б). С помощью меры времени запаздывания определяется средняя длина вертикальных структур. Также данная мера позволяет оценить среднее время, в течение которого ДНК будет находиться в определенном состоянии, или как долго это состояние будет заблокировано.

Мера замирания определяется отношением количества рекуррентных точек, образующих вертикальные линии, к общему количеству рекуррентных точек. Величина LAM характеризует наличие состояний замирания исследуемой молекулы. С помощью меры максимальной

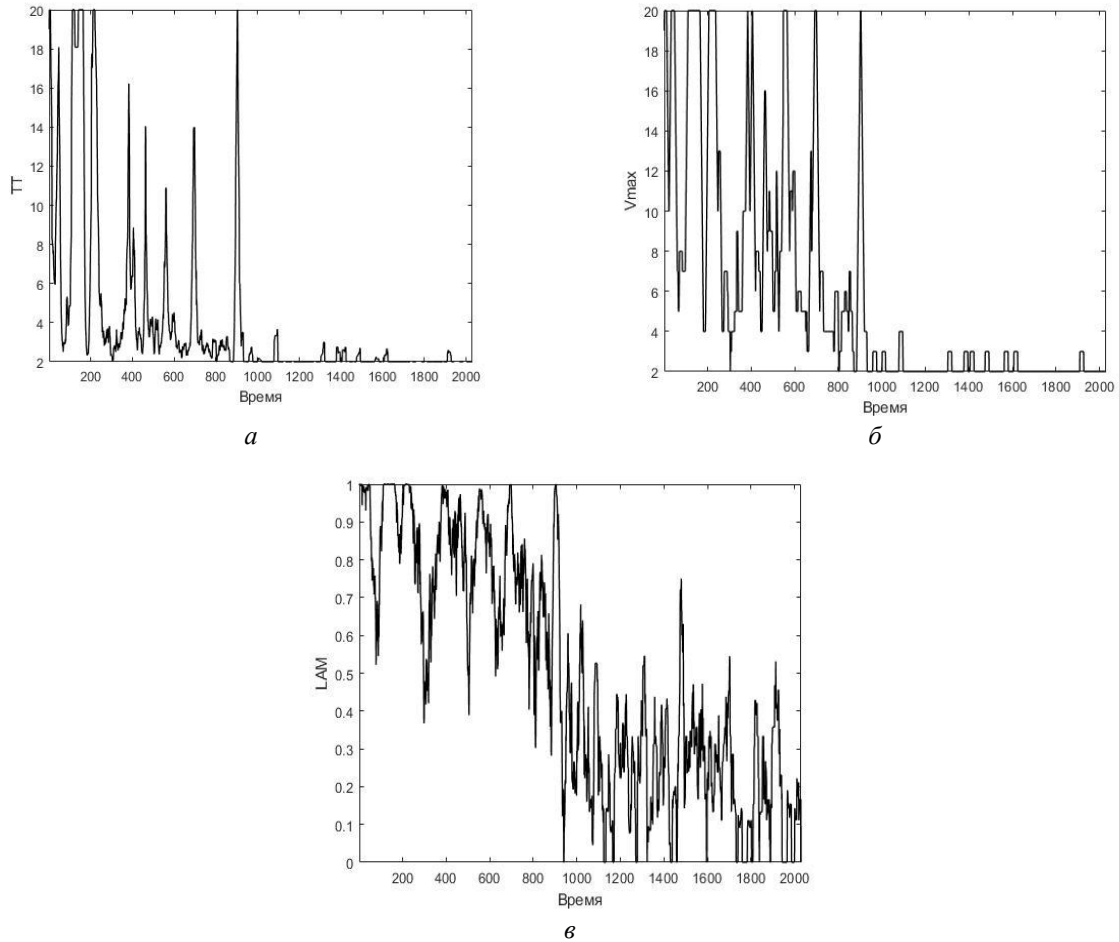


Рис. 8. Графики эволюции мер, основанные на РД растяжения первой пары оснований с использованием окна размером $S = 20$: *a* – мера времени запаздывания; *б* – мера максимальной длины вертикальной линии; *в* – мера замирания

длины вертикальной структуры можно определить максимальную длину вертикальной линии.

Фазовый переход исследуется при различных значениях температуры, которая варьируется в диапазоне от 200 до 540 К. Расчёт рекуррентных свойств ДНК выполняется для центральной пары оснований. По рис. 4 проведен расчет мер в окне размером $S = 400$, выбор которого установлен экспериментально и вычислен по формуле $S \approx N/5$. Важным моментом является выбор размера окна. Слишком малый размер окна может повлечь за собой излишний шум и флуктуации, а слишком большой – даст более гладкие графики мер ввиду незаметного влияния событий во всем объеме данных. Вычисление мер в окне аналогично вычислению мер для всей рекуррентной диаграммы и позволяет проследить динамику изменения мер в зависимости от времени (эпохи).

Полученные меры реагируют на смену структуры диаграммы. Действительно, в окрестности координаты 900 о.в.с. на диаграмме видна резкая смена структуры, что говорит о смене режима. Меры *RR*, *DET*, *RATIO*, *ENTR* и *TT* отреагировали на это возникновением локальных экстремумов. Вычисление данных мер позволя-

ет установить факт смены режима. Характер перехода лучше всего демонстрируют меры *TT*, *LAM*, *Vmax*, поскольку расчёт данных мер основан на вертикальных структурах. Поскольку при анализе топологии и структуры рекуррентной диаграммы результатов моделирования денатурации ДНК четко видны именно вертикальные структуры, значит именно меры, основанные на них, будут чувствительны к анализу исследуемого процесса.

Рассмотрим процесс более детализировано, уменьшив размер окна при расчёте мер, а также рассчитав меры для сигнала других частиц. На рис. 8 приведены графики эволюции мер *TT*, *LAM*, *Vmax*, рассчитанные для динамики одной пары оснований с использованием окна размером $S = 20$. Затем приведем графики эволюции мер *TT*, *LAM*, *Vmax*, рассчитанные для динамики другой пары оснований с использованием окна с таким же размером, $S = 20$. На рис. 8, 9 можно заметить явные предвестники фазового перехода. В частности, по графикам эволюции меры *TT* можно увидеть непосредственно перед фазовым переходом снижение значения амплитуды, а затем резкий всплеск. По графикам эволюции меры *Vmax* – увеличение промежутков времени между локаль-

ными экстремумами, а по графикам эволюции меры LAM – резкое снижение амплитуды колебаний.

Заключение

Целью работы являлось исследование фазового перехода в ДНК, происходящего при термической денатурации, методом количественного анализа рекуррентных диаграмм. Для достижения цели реализована модель Пейрара – Бишопа – Доксуа, которая позволяет определять состояние пар оснований ДНК в произвольный момент времени, проведено моделирование динамики ДНК при различных температурах.

Для анализа полученных результатов использован метод количественного анализа рекуррентных диаграмм, благодаря которому появилась возможность судить о характере протекающих в системе процессов. С помощью данного метода анализа произведен расчет мер и представлены их эволюции, исходя из которых был найден момент фазового перехода в системе.

По результатам исследования выявлено, что наиболее чувствительными к анализу исследуемого процесса оказались меры, основанные на вертикальных структурах, а именно TT , LAM , V_{max} . Данные меры количе-

ственно определяют мелкомасштабные структуры рекуррентных диаграмм, изменения в топологии и «восприимчивость» системы к флуктуациям различной природы в окрестности фазового перехода.

В отличие от классической теории фазовых переходов, процессы в ДНК демонстрируют качественно новые закономерности, связанные с термодинамикой фазовых переходов, в так называемых одномерных системах. Применительно к термодинамике ДНК это является следствием сочетания в молекуле ДНК свойств, характерных для кристаллической решетки – дальний порядок взаимодействия оснований вдоль цепей ДНК, и «слабых» взаимодействий между цепями (водородные связи), характерные для жидкостей. Это предопределяет низкотемпературный барьер структурных трансформаций в ДНК (локальное плавление), обусловленных термоактивированным разрывом водородных связей, которые также определяются напряженным состоянием (деформированием) ДНК.

Энергетический вклад обоих типов взаимодействий отражен в модели Пейрара – Бишопа – Доксуа, которая является базовым представлением при исследовании динамики ДНК и дает уникальную возможность изучения роли коллективных мод дефектов, так называемых

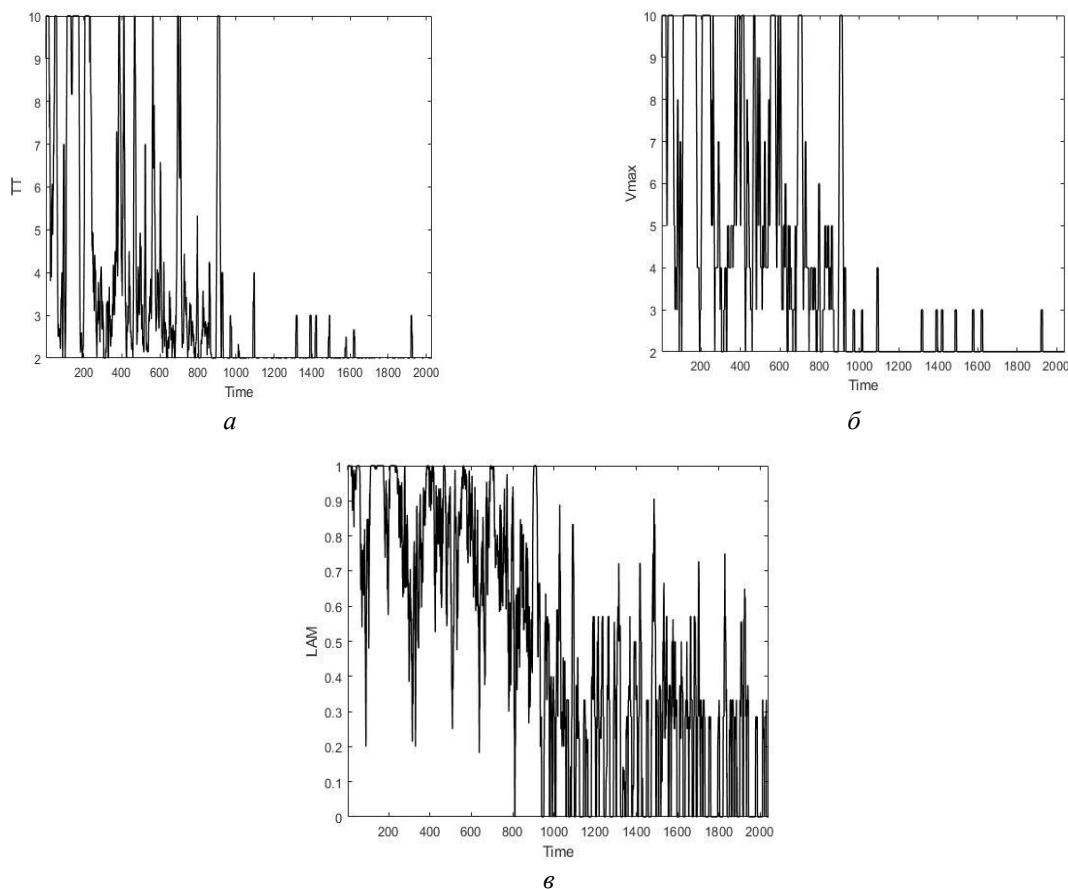


Рис. 9. Графики эволюции мер, основанные на РД растяжения второй пары оснований с использованием окна размером $S = 20$: a – мера времени запаздывания; b – мера максимальной длины вертикальной линии; c – мера замирания

открытых комплексов, на термодинамику структурных переходов. Последние определяют фундаментальные закономерности функционирования ДНК (транскрипцию, репликацию) и являются основой при построении термодинамики трансформаций в клетке с учетом роли коллективных степеней свободы, представляющих различные нелинейные моды открытых комплексов.

Благодаря возможности выявления фазового перехода, количественный анализ рекуррентных диаграмм является одним из перспективных методов. Данный подход применим не только для исследования денатурации, но и для большинства биологических процессов, например канцерогенез, влияние геомагнитного поля на человека и других.

Список литературы

- Alexandrov B.S., Gelev V., Monisova Y., Alexandrov L.B., Bishop A.R., Rasmussen K.O., Usheva A. A nonlinear dynamic model of DNA with a sequence-dependent stacking term // *Nucleic Acids Research*. – 2009. – Vol. 37. – P. 2405–2410.
- Causo M.S., Coluzzi B., Grassberger P. Simple model for the DNA denaturation transition // *Physical Review E*. – 2000. – Vol. 62. – P. 3958–3973.
- Dauxois T., Peyrard, M., Bishop, A.R. Dynamics and thermodynamics of a nonlinear model for DNA denaturation // *Physical Review E*. – 1993. – Vol. 47(1). – P. 684.
- Dauxois T., Peyrard M., Bishop A.R. Entropy-driven DNA denaturation // *Physical Review E*. – 1993. – Vol. 47(1). – P. R44.
- Fitzpatrick M., Sundaresan N.M., Li A.C., Koch J., Houck A.A. Observation of a dissipative phase transition in a one-dimensional circuit QED lattice // *Physical Review X*. – 2017. – Vol. 7(1). – Art. 011016.
- Hall A.C., Ostrowski L.A., Mekhail K. Phase separation as a melting pot for DNA repeats // *Trends in Genetics*. – 2019. – Vol. 35(8). – P. 589–600.
- Hillebrand M., Kalosakas G., Schweltnus A., Skokos C. Heterogeneity and chaos in the Peyrard-Bishop-Dauxois DNA model // *Physical Review E*. – 2019. – Vol. 99(2). – Art. 022213.
- Inman R.B., Baldwin R.L. Helix–random coil transitions in DNA homopolymer pairs // *J. Mol. Biol.* – 1964. – Vol. 8. – P. 452–469.
- Likhachev I.V., Lakhno V.D. The direct investigation of DNA denaturation in Peyrard-Bishop-Dauxois model by molecular dynamics method // *Chemical Physics Letters*. – 2019. – Vol. 727. – P. 55–58.
- Longhi S. Topological phase transition in non-Hermitian quasicrystals // *Physical review letters*. – 2019. – Vol. 122(23). – Art. 237601.
- Lopes M.A. Recurrence quantification analysis of dynamic brain networks // *European Journal of Neuroscience*. – 2021. – Vol. 53. – no. 4. – P. 1040–1059.
- Cadoni M., de Leo R., Demelio S., Gaeta G. Propagation of twist solitons in fully inhomogeneous DNA chains // *Journal of Nonlinear Mathematical Physics*. – 2010. – Vol. 17(4). – P. 557–569.
- Marwan N. Recurrence plots for the analysis of complex systems // *Physics reports*. – 2007. – Vol. 438. – no. 5-6. – P. 237–329.
- Poland D., Scheraga H.A. Occurrence of a phase transition in nucleic acid models // *The Journal of Chemical Physics*. – 1966. – Vol. 45. – P. 1464–1469.
- Poland D., Scheraga H.A. Phase transitions in one dimension and the helix-coil transition in polyamino acids // *The Journal of Chemical Physics*. – 1966. – Vol. 45. – P. 1456–1463.
- Ramdani S., Boyer A., Caron S., Bonnetblanc F., Bouchara F., Duffau H., Lesne A. Parametric recurrence quantification analysis of autoregressive processes for pattern recognition in multichannel electroencephalographic data // *Pattern Recognition*. – 2021. – Vol. 109. – Art. 107572.
- Ritort F. Single-molecule experiments in biological physics: methods and applications // *Journal of Physics: Condensed Matter*. – 2006. – Vol. 18(32). – P. R531.
- Shigaev A.S., Ponomarev O.A., Lakhno V.D. Theoretical and experimental investigations of DNA open states // *arXiv preprint arXiv*. – 2014. – Art. 1402.6533.
- Nose S. A unified formulation of the constant temperature molecular dynamics methods // *The Journal of Chemical Physics*. – 1984. – Vol. 81, no. 1. – P. 511–519.
- Theodorakopoulos N., Dauxois T., Peyrard M. Order of the phase transition in models of DNA thermal denaturation // *Physical Review Letters*. – 2000. – Vol. 85(1). – P. 6–11.
- Wartell R.M., Benight A.S. Thermal denaturation of DNA molecules: a comparison of theory with experiments // *Phys. Rep.* – 1985. – Vol. 126. – P. 67–107.
- Zou Y., Donner R.V., Marwan N., Donges J.F., Kurths J. Complex network approaches to nonlinear time series analysis // *Physics Reports*. – 2019. – Vol. 787. – P. 1–97.

Финансирование. Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (Государственный контракт № АААА-А19-119013090021-5), а также поддержана Грантом Президента Российской Федерации для государственной поддержки молодых российских ученых № МК-44.2022.1.1 (Соглашение № 075-15-2022-362 от 05 мая 2022 года).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

STUDY OF THE DNA DENATURATION BASED ON THE PEYRARD-BISHOP-DAUXOIS MODEL AND RECURRENCE QUANTIFICATION ANALYSIS

A.S. Nikitiuk, O.S. Burmistrova, O.B. Naimark

Laboratory of Physical Foundations of Strength
Institute of continuous media mechanics UB RAS, Perm, Russia

ARTICLE INFO

Received: 12 August 2022
Approved: 01 December 2022
Accepted for publication: 03 December 2022

Key words:

DNA, open states, dynamics of open states, denaturation, phase transition, recurrence plot, quantitative analysis of recurrence plots.

ABSTRACT

The aim of this work was to study of the phase transition during thermal denaturation of DNA by the method of recurrence quantification analysis, which is the most modern and efficient approach for the analysis of complex systems. To achieve the goal, the nonlinear dynamics model of the DNA denaturation is implemented, based on the Peyrard-Bishop-Dauxois approach, which allows determining the state of DNA base pairs at an arbitrary point in time. Next, the dynamics of DNA is simulated at different temperatures, and the evolution of hydrogen bond stretching during DNA heating is studied. To analyze the results obtained, the method of recurrence quantification analysis is used, which make it possible to judge the nature of the processes occurring in the system and are of particular interest in relation to the study of nonlinear DNA dynamics, in particular phase transitions, since on the basis of their analysis it is possible to identify states of repetition and fading (laminarity), the occurrence of extreme events, the presence of hidden periodicity and cyclicity. Using this method of analysis, quantitative measures are calculated in a window shifted along the main diagonal of the recurrent plot (RP). These measures quantify fine-scale RP structures. The calculations reveal the most sensitive measures to the analysis of the process under study, with the help of which the precursors of the phase transition are determined.