

ОПИСАНИЕ ПЕРЕХОДА СПИРАЛЬ-КЛУБОК В ГЕТЕРОПОЛИМЕРАХ В РАМКАХ МОДЕЛИ ОМПЦ (GMPC)

А.В. АСАТРЯН^{1,2*}, А.К. КАЛАШЯН³, В.Ф. МОРОЗОВ²

¹Институт химической физики, Ереван, Армения

²Ереванский государственный университет, Ереван, Армения

³Национальный политехнический университет Армении, Ванадзорский филиал

*e-mail: arevik.asatryan@ysu.am

(Поступила в редакцию 9 сентября 2021 г.)

Работа посвящена изучению поведения гетерогенных биополимеров при плавлении вторичной структуры. Генерируется случайный бимодальный биополимер. Каждому компоненту соответствует своя трансфер-матрица ОМПЦ (GMPC), перемножением которых вычисляется свободная энергия системы. Показано, что при длине 5000 повторяющихся единиц приведенная свободная энергия не зависит от реализации последовательности (первой структуры). Соответствующие кривые денатурации показывают, что точка перехода достаточно хорошо определяется, как линейная комбинация соответствующих гомополимеров, однако профиль кривых зависит от реализации. Вычислена поведение числа стыков между спиральным и клубкообразным состояниями. Получен неожиданный результат, что рацемический гетерополимер длиной в 5000 повторяющихся единиц в процессе плавления разбивается на порядка 10 участков. Мы это связываем с поведением корреляции в системе, что и будет исследоваться далее.

1. Введение

Различные процессы в биологических системах часто приводят или сопровождаются переходом спираль-клубок в биополимерах. Несмотря на достигнутые успехи, проблема перехода спираль-клубок до сих пор остается одной из самых популярных проблем в структурной биологии и в настоящее время интерес к проблеме не ослабевает [1–5]. В данной публикации мы основываемся на микроскопической теории перехода спираль–клубок [6–11]. Это стало возможным благодаря подходам и методам современной теоретической физики, позволяющим использовать в статистической физике полимеров большое количество теоретических моделей, более адекватно описывающих предмет изучения, чем это принято на сегодняшний день и не требующих введения дополнительного феноменологического параметра для учета кооперативности системы. На основе Поттс-подобной модели была развита теория перехода спираль–клубок в полипептидах [12–18]. Было также показано, что в пренебрежении петлями большого масштаба характеристическое уравнение для модели ДНК совпадает с таковым для Обобщенной Модели Полипептидной Цепи (ОМПЦ) [15]. Отличие от

полипептидной модели в том, что число повторяющихся единиц, фиксируемых одной водородной связью, входит в ОМПЦ как некий параметр. В итоге получается, что для описания перехода спираль–клубок как в полипептидах, так и в полинуклеотидах, а также для учета влияния факторов различной природы может быть использована одна и та же ОМПЦ, но с разными значениями параметров. Поскольку дальнейшие исследования будут проводиться на основе базовой Обобщенной Модели Полипептидной Цепи, приведем ее основные положения.

2. Модель ОМПЦ

Предположим, имеется биополимер, содержащий N повторяющихся единиц. Предполагается, что каждая повторяющаяся единица может находиться в Q дискретных состояниях, соответствующих различным дискретным значениям конформаций (поворотно-изомерное приближение), энергии этих состояний предполагаются одинаковыми и не зависят от состояний соседних единиц. Опишем состояние i -ой единицы с помощью спиновой переменной γ_i так что $\gamma_i = 1, \dots, Q$. Водородная связь, приводящая к образованию спиральной структуры, образуется, когда Δ соседние повторяющиеся единицы находятся в одной и той же определенной конформации. Пусть это будет конформация номер один. Гамильтониан такой модели ОМПЦ имеет вид

$$-\beta H = J \sum_{i=1}^N \prod_{k=\Delta-1}^0 \delta(\gamma_{i-k}, 1) = J \sum_{i=1}^N \delta_i^{(\Delta)} \quad (1)$$

где $\delta_i^{(\Delta)} = \prod_{k=\Delta-1}^0 \delta(\gamma_{i-k}, 1)$ – произведение символов Кронекера, обеспечивает выделение энергии J только в том случае, когда Δ подряд соседних повторяющихся единиц находятся в спиральной конформации, обозначенной нами номером 1. Трансфер–матрицу ОМПЦ ($\Delta \times \Delta$) можно построить для конкретных значений Δ и Q на базе гамильтониана (1). Она имеет следующие свойства [15–21]:

- элемент $(1, 1)$ равен $W = e^J$;
- все элементы последней строки равны 1;
- элементы первой верхней псевдодиагонали равны 1;
- элементы $(\Delta - 1, \Delta)$ и (Δ, Δ) равны $Q - 1$;
- все остальные элементы равны нулю.

Обобщая на случай произвольных конечных Δ и Q , получим трансфер–матрицу модели вида

$$G = \begin{pmatrix} e^J & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & Q-1 \\ 1 & 1 & 1 & \dots & 1 & 1 & Q-1 \end{pmatrix}. \quad (2)$$

Для гомополимера достаточно большой длины статистическая сумма и характеристическое уравнение для трансфер-матрицы модели имеют вид:

$$Z = \text{Tr}G^N, \quad (3)$$

$$\lambda^{(\Delta-1)}(\lambda - e^J)(\lambda - Q) = (e^J - 1)(Q - 1). \quad (4)$$

Как уже было сказано выше [5], J – энергия водородной связи, приведенная к температуре; Q – число конформаций повторяющихся единиц, спиральному состоянию соответствует конформация номер 1; Δ – число повторяющихся единиц, фиксируемых одной водородной связью; $W = e^J$ – температурный параметр, определяющие энергетические свойства повторяющихся единиц. Последние параметры определяют разные свойства системы. Этому Гамильтониану соответствует трансфер-матрица (2) и вековое уравнение (4).

Для базовой модели, определенной выше температурный параметр W , имеет вид $W = e^{-U/(kT)}$, где U – энергия водородной связи, а энтропийный параметр Q константой. Однако при учете взаимодействия с растворителем и лианами W является сложной функцией температуры, Q тоже становится температурно зависимым [11].

3. Свободная энергия бимодального гетерополимера

Рассмотрим бимодальный гетерополимер состоящий из повторяющихся единиц двух сортов «А» и «В». В таком случае гамильтониан ОМПЦ будет иметь вид

$$-\beta H = \sum_{i=1}^N J_i \delta_i^\Delta, \quad (5)$$

$$\text{где } J_i = \begin{cases} J_A & \text{если } i \text{ - единица типа А} \\ J_B & \text{если } i \text{ - единица типа В} \end{cases}, \quad Q_i = \begin{cases} Q_A & \text{если } i \text{ - единица типа А} \\ Q_B & \text{если } i \text{ - единица типа В} \end{cases}.$$

Поскольку мы будем исследовать зависимость свободной энергии от N , то статистическая сумма такой системы будет иметь вид [18], хотя для больших N можно пользоваться выражением (3).

$$Z = J^* \times \prod_{i=1}^N G_i \times J, \quad (6)$$

где G_i – трансфер-матрица (2) ОМПЦ соответствующая данной модели, $J^* = (0 \ 0 \ \dots \ 1)$ – Δ -компонентная вектор-строка, $J = (0 \ 0 \ \dots \ 0 \ 1)^T$ – вектор-столбец.

Поскольку произведение матриц не обладает коммутативностью, то, вообще говоря, статистическая сумма зависит от последовательности матриц, что соответствует первичной последовательности повторяющихся единиц. Однако, ранее было показано, что для случайной последовательности при достаточно больших N величина $\ln Z/N$ стремится к конечному пределу [22], также как и приведенная свободная энергия, то нам важно определить порядок величины N , когда систему можно считать термодинамической с данной степенью точности.

4. Степень спиральности

Из определения степени спиральности θ , как средняя доля повторяющихся единиц в спиральном состоянии

$$\theta_N = N_h / N, \quad (7)$$

где N_h число повторяющихся единиц в спиральном состоянии. Для бимодального случая положим $J_A = J_0 + \Delta J$, $J_B = J_0 - \Delta J$, $J_0 = (J_A + J_B)/2$, $\Delta J = (J_A - J_B)/2$. Тогда величина стат. суммы будет определяться следующим выражением:

$$Z = \sum_{\{\eta_i\}} e^{\sum_i J_i \delta_i^\Delta}, \quad (8)$$

$J_i = J_0 + \Delta J \varepsilon$, где ε принимает значения ± 1 . Отсюда

$$N_h = \frac{1}{Z} \sum_{\{\eta_i\}} \sum_{i=1}^N \delta_i^\Delta e^{\sum_i J_i \delta_i^\Delta} = \frac{1}{Z} \frac{\partial Z}{\partial J_0} \text{ из (6). Степень спиральности будет иметь вид:}$$

$$\theta_N = \frac{1}{NZ} \tilde{J}(G'_1 G_2 G_3 \dots G_N + G_1 G'_2 G_3 \dots G_N + \dots + G_1 G_2 G'_3 G_4 \dots G_N + \dots + G_1 G_2 G_3 G'_4 \dots G'_N) J, \quad (9)$$

где G'_i – матрица G_i , где все элементы кроме первого равны нулю. Следуя [9,10] выражение в скобке может быть записано как $(E \ O) \prod_{i=1}^N M_i (E \ O)^T$, где M_i –

суперматрица, имеющая вид $\begin{pmatrix} G_i & G'_i \\ O & G_i \end{pmatrix}$, E – единичная матрица размера Δ ,

O – нулевая матрица того же размера и окончательно получаем

$$\theta_N = \frac{1}{NZ} J^* (E \ O) \prod_{i=1}^N M_i \begin{pmatrix} E \\ O \end{pmatrix} J = \frac{1}{N} \frac{J_\theta^* \prod_{i=1}^N M_i J_\theta}{J^* \prod_{i=1}^N G_i J}, \quad (10)$$

Здесь $J_\theta^* = (0 \ 0 \ \dots \ 0 \ 0 \ 1 \ \dots \ 1)$, $J = (0 \ 0 \ \dots \ 1 \ 1)^T$ вектор строка и вектор столбец размером 2Δ .

5. Средняя доля стыков между спиральным и клубкообразным участками

Средняя доля стыков η_N определяется, как доля пар спирального состояния и любого другого состояния.

$$\langle \eta_N \rangle = p(\gamma_i = 1, \gamma_{i+1} = k), k \neq 1. \quad (11)$$

Отсюда $\eta_N = \theta - p(hh)$. Аналогично (8) получим

$$p(hh) = \frac{1}{NZ} J^* (G'_1 G'_2 G_3 G_4 \dots G_N + G_1 G'_2 G'_3 \dots G_N + \dots + G_1 G'_2 \dots G'_{N-1} G'_N) J. \quad (12)$$

Аналогично тому, как было получено выражение для θ , введем суперматрицу L_i

$$L_i = \begin{pmatrix} G_i & G_i' & O \\ O & O & G_i' \\ O & O & G_i \end{pmatrix}. \quad (13)$$

Выражение в скобках в (12) получается через L_i , как

$$(E \ O \ O) \prod_{i=1}^{N-1} L_i (O \ O \ E)^T. \quad (14)$$

Откуда $\eta_N = \frac{1}{ZN} J_L^* \prod_{i=1}^{N-1} L_i J_L$, где J_L^* – вектор-строка, размера $3\Delta(0 \ 0 \ \dots \ 0 \ 0 \ 1 \ \dots \ 1)$, а J_L – вектор-столбец того же размера $(0 \ 0 \ \dots \ 1 \ 1)^T$. Т.о. Средняя доля стыков

$$\eta_N = \frac{1}{ZN} J_\theta^* \prod_{i=1}^N M_i J_\theta - \frac{1}{ZN} J_\eta^* \prod_{i=1}^{N-1} L_i J_\eta, \quad (15)$$

где Z определяется из (6). Отсюда можно перейти к выражению для средний длины спирального участка. Поскольку средняя длина спирального участка $v_N = \theta_N / \eta_N$, то

$$v_N = \left(1 - \frac{J_\eta^* \prod_{i=1}^{N-1} L_i J_\eta}{J_\theta^* \prod_{i=1}^N M_i J_\theta} \right)^{-1}. \quad (16)$$

6. Результаты и их обсуждение

Введем величину x как вероятность повторяющихся единиц «A» в случайной последовательности, а $1-x$ – «B», соответственно. Т.о. Можно генерировать случайную последовательность из повторяющихся единиц «A» и «B» заданной длиной N . Соответственно вводится $C_A = N_A/N$, $C_B = N_B/N$, где N_A и N_B число повторяющихся единиц данного вида в данной последовательности. На рис.1 приведена зависимость величины $\ln Z/N$ от N . Видно, что при достаточно длинной цепи $N > 3000$, приведенная свободная энергия действительно стремится к конечному пределу. Для коротких цепей приведенная свободная энергия сильно зависит от реализации последовательности (первой структуры). На рис. 1 показано поведение приведённой свободной энергии для разных x и двух температур для одной и той же реализации. Следует отметить, что для всех графиков температура указана в приведенных величинах $t = J_A^{-1}$. Такой способ введения температуры позволяет качественно сопоставлять графики зависимости усредненных параметров.

С хорошей степенью точности $N^* \approx 3000$ это подтверждается для широкого набора реализаций, о чем свидетельствует рис. 2. При достаточно больших длин свободная энергия не зависит от реализации. Т.о. На языке приведенной

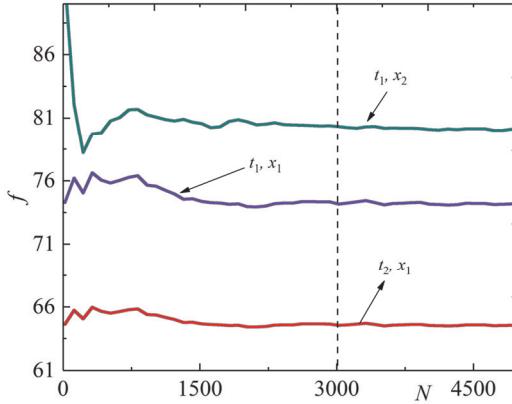


Рис.1. Зависимость приведенной свободной энергии от числа повторяющихся единиц для разных x и температур. $x_1 = 0.4$, $x_2 = 0.5$, $\Delta = 4$, $t_1 = 0.213$, $t_2 = 0.219$, $U_A = 1$, $U_B = 0.8$, $Q_A = 71$, $Q_B = 51$.

свободной энергии при $N > 3000$ можно утверждать, что гетерополимер ведет себя как гомополимер.

На рис.2 показано поведение приведенной свободной энергии при одном и том же значении x и приведенной температуре. Видно, что, однако, если перейти к степени спиральности, то ситуация изменится.

На рисунке 3 показаны кривые денатурации для разных N . Видно, что положение кривых зависит от N , однако даже при $N = 5000$, когда приведенная свободная энергия уже не зависит от первичной структуры (реализации), кривая денатурация зависит.

Из рисунка 4 видно что общие свойства кривых денатурации от реализации не зависят, а вот детали профиля кривых меняются от реализации.

На рисунке 5 приведены кривые денатурации при разных x . Результатом наших расчетов является $t_m(x) = xt_{mA} + (1-x)t_{mB}$, где t_m – точка перехода.

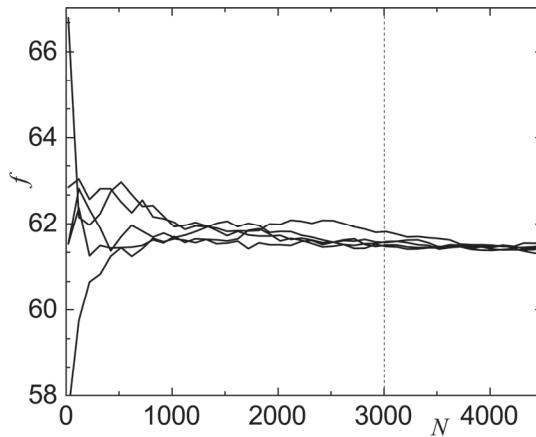


Рис.2. Зависимость приведенной свободной энергии от числа повторяющихся единиц для разных последовательностей. $x_1 = 0.4$, $t = 0.219$, $U_A = 1$, $U_B = 0.8$, $Q_A = 71$, $Q_B = 51$.

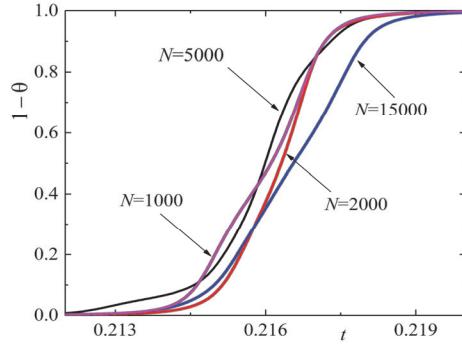


Рис.3. Зависимость $1 - \theta$ от температуры для последовательностей с разными N . $x_1 = 0.4$, $\Delta = 4$, $U_A = 1$, $U_A = 0.8$, $Q_A = 71$, $Q_B = 51$. (Температура приведена к величине).

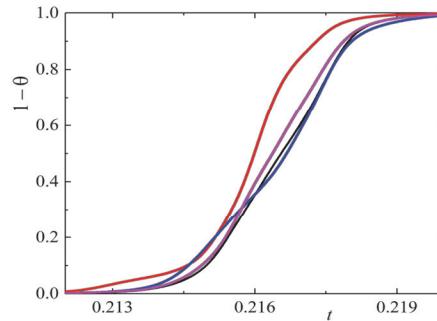


Рис.4. Зависимость $1 - \theta$ от температуры для разных реализаций при $x_1 = 0.4$, $\Delta = 4$, $U_A = 1$, $U_A = 0.8$, $Q_A = 71$, $Q_B = 51$, $N = 5000$. (Температура приведена к величине $t = J_A^{-1}$).

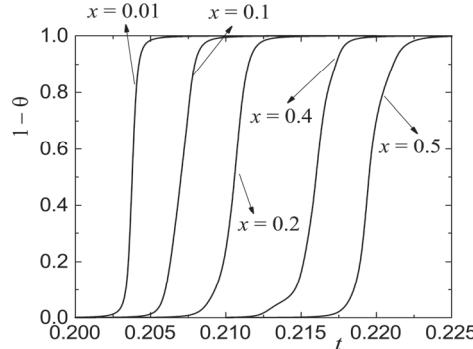


Рис.5. Зависимость $1 - \theta$ от температуры для последовательностей с разными x в термодинамическом придел $\Delta=4$, $U_A = 1$, $U_A = 0.8$, $Q_A = 71$, $Q_B = 51$, $N = 5000$. (Температура приведена к величине $t = J_A^{-1}$).

Рассмотрим детали плавления спиральной структуры. При плавлении макромолекула разбивается на спиральные фрагменты разделенные клубкообразными участками. Доля таких фрагментов определяется долей стыков (h, \bar{h}) , т.е. на языке ОМПЦ $p(\gamma_i = 1, \gamma_{i+1} = k)$, $k \neq 1$, что выражается формулой (15).

На рис. 6 изображена зависимость доли стыков η от степени спиральности

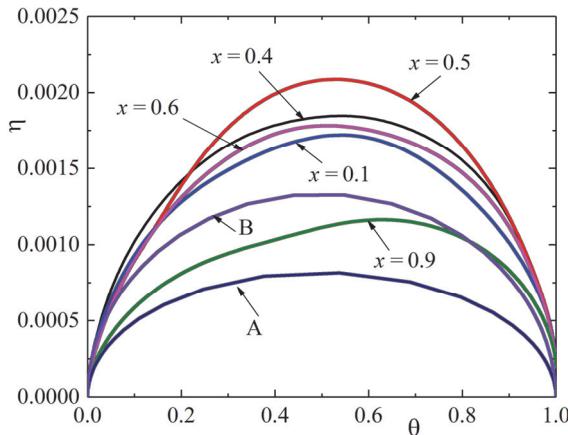


Рис.6. Зависимость $\eta(\theta)$ для разных x -ов, $\Delta=4$, $U_A = 1$, $U_A = 0.8$, $Q_A = 71$, $Q_B = 51$, $N = 5000$.

при разных x , включая гомополимер «А» и «В». Такая зависимость выбрана для того, чтобы сравнивать поведение при различных x , поскольку $\eta(t)$ для разных x будут находиться в разных температурных участках. Поэтому мы выбрали зависимость $\eta(\theta)$.

Как было ясно и как видно на рис. 6 кривые $\eta(\theta)$ являются кривыми с максимумом в районе $\theta = 0.5$. Кривая при $x = 0$ (менее стабильный гомополимер polyB) расположена выше кривой при $x=1$ (более стабильный polyA). Однако гетерополимеры располагаются выше обоих гомополимеров. Это ясно, поскольку в гетерополимерах больше вариантов стыков. Из рисунка видно, что более высокий максимум получается при $x = 0.5$, т.е. при наивысшей гетерогенности. Тем не менее случай с $x = 0.5$ проявляет и большой разброс для различных реализаций рис. 7. Разброс при $x = 0.4$ меньше (рис.8), а при $x = 0.1$ и $x = 0.9$ разброс уже не виден и не представлен на рисунках. Что же касается величины максимума, то она варьируется от 2.2×10^{-3} ($x = 0.5$) до 6×10^{-4} ($x=1$), что соответствует от 11 до 3 стыков.

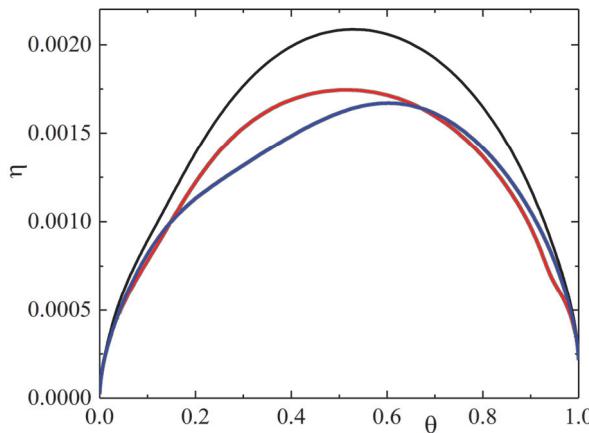


Рис.7. Зависимость $\eta(\theta)$ для 3 реализаций, $x = 0.5$, $\Delta = 4$, $U_A = 1$, $U_A = 0.8$, $Q_A = 71$, $Q_B = 51$, $N = 5000$.

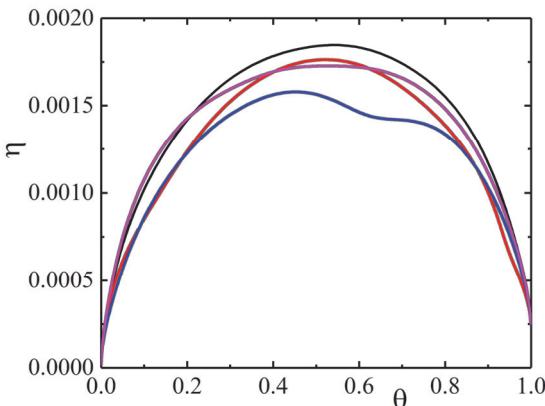


Рис.8. Зависимость $\eta(\theta)$ для четырех реализаций, $x = 0.5$, $\Delta = 4$, $U_A = 1$, $U_A = 0.8$, $Q_A = 71$, $Q_B = 51$, $N = 5000$.

Таким образом, даже при таком поверхностном рассмотрении картина перехода получается весьма парадоксальной. При длине 5000 повторяющихся единиц мы наблюдаем 11 стыков для гетерополимера и порядка единичных стыков для гомополимеров. А значит, анализируя поведения доли стыков мы приходим к заключению, что длина в 5000 повторяющихся единиц хотя и достаточно для того, чтобы свободная энергия гетерополимера (а гомополимер и подавно) выходило бы на насыщение, а разбиение на участки к таким результатам не приводит. В недавно вышедшей работе А. Бадасяна [25] на примере модели Зимма-Брэгга было показано, что даже в гомополимерном случае существует промежуточный по длине режим связанный с корреляционной длиной. В связи с этим, следующие наши исследования будут связаны с исследованием корреляций конформаций в гетерополимерах.

Исследование выполнено при финансовой поддержке ГКН МОН РА в рамках научного проекта № 19YR-1F057.

ЛИТЕРАТУРА

1. V.A. Bloomfield. American J. of Phys., **67**, 1212, (1999).
2. J.S. Schreck, J.-M. Yuan. Phys. Rev. E, **81**, 061919 (2010).
3. K. Ghosh, K.A. Dill. J. Am. Chem. Soc., **131**, 2306 (2009).
4. A.C. Hausrath. Prot. Science, **15**, 2051 (2006).
5. D.J. Jacobs, G.G. Wood. Biopolymers, **75**, 1 (2004).
6. G. Rayan, A.D. Tsamaloukas, R.B. Macgregor, H. Heerklotz. J. of Phys. Chem. B, **113**, 1738 (2009).
7. S.V. Kornilova, P. Yasem, D.N. Grigorev, et al. Biofizika, **43**(1), 46 (1998).
8. E.V. Petrovskaya, V.V. Vasilevskaya, A.R. Khokhlov. Moscow University Physics Bulletin, **62**(2), 99 (2007).
9. C.J. Benham. Phys. Rev. E, **53**, 2984 (1996).
10. T. Einert, R. Netz. Biophys. J., **99**, 578 (2010).
11. T. Einert, P. Naeger, H. Orland, R. Netz. Phys. Rev. Lett., **101**(1-4), 048103 (2008).
12. N.S. Ananikyan, Sh.A. Hayryan, E.Sh. Mamasakhlisov, V.F. Morozov. Biopolymers,

- 30**, 357 (1990).
13. **Sh.A. Hayryan, E.Sh. Mamasakhlisov, V.F. Morozov.** Biopolymers, **35**, 75 (1995).
 14. **N.S. Ananikyan, E.Sh. Mamasakhlisov, V.F. Morozov.** Phys. Chem., **27**(3), 603 (1993).
 15. **V.F. Morozov, E.Sh. Mamasakhlisov, Sh.A. Hayryan, Hu. Chin-Kun.** Physica A, **281**(1–4), 51 (2000).
 16. **A.V. Badasyan, A.V. Grigoryan, A.Yu. Chukhadzhyan, Y.Sh. Mamasakhlisov, V.F. Morozov.** Izvestiya Natsional'noi Akademii Nauk Armenii, Fizika, **37**, 320 (2002).
 17. **V.F. Morozov, A.V. Badasyan, A.V. Grigoryan, M.A. Sahakyan, E.Sh. Mamasakhlisov.** Biopolymers, **75**(5), 434 (2004).
 18. **A.V. Grigoryan, A.V. Badasyan, V.F. Morozov, E.S. Mamasakhlisov.** Izvestiya Natsional'noi Akademii Nauk Armenii, Fizika, **39**, 265 (2004).
 19. **Sh.A. Tonoyan.** PhD thesis, The theory of helix-coil transition in presence of low molecular compounds, (2011), Yerevan, Armenia.
 20. **A.V. Tsarukyan.** PhD thesis, Effects of intra- and intermolecular interactions in helix-coil transitions in frame of Generalized Model of Polypeptide Chain, (2009), Yerevan, Armenia.
 21. **A.V. Asatryan.** PhD thesis, Influence of aqueous solutions of low-molecular ligands on helix-coil transition in heterogeneous biopolymers, (2020), Yerevan, Armenia.
 22. **V.I. Oseledets, Tr. Mosk. Mat. Obs.**, **19**, 179 (1968).
 23. **M. Serva, G. Paladin,** Phys. Rev. Lett., **70**, 105 (1993).
 24. **A. Badasyan.** PhD thesis, Переход Спираль-Клубок в Биополимерах. Влияние неоднородностей, (2005), Yerevan, Armenia.
 25. **A. Badasyan.** Polymers, **13**(12), 1985 2021.

DESCRIPTION OF HELIX-COIL TRANSITION IN HETEROPOLYMERS IN FRAME OF GMPC

A.V. ASATRYAN, A.K. KALASHYAN, V.F. MOROZOV

The work is devoted to the study of the behavior of heterogeneous biopolymers during the melting of the secondary structure. A random bimodal biopolymer is generated. Each component corresponds to its own transfer-matrix of GMPC, by multiplying these transfer-matrices the free energy of the system is calculated. It is shown that for a length of 5000 repeating units, the reduced free energy does not depend on the realization of the system (primary structure). The corresponding denaturation curves show that the transition point is well defined as a linear combination of the corresponding homopolymers, but the profile of the curves depends on the implementation. The behavior of the number of junctions between the helical and coiled states is calculated. An unexpected result is obtained that the racemic heteropolymer with a length of 5000 repeated units is divided into about 10 sections during the melting process. We associate this with the behavior of correlation in the system, which will be investigated further.