

УДК 577.32; 51.72

В. А. Андрющенко, С. Ф. Чекмарев

*Новосибирский государственный университет
ул. Пирогова, 2, Новосибирск, 630090, Россия*

*Институт теплофизики им. С. С. Кутателадзе СО РАН
пр. Акад. Лаврентьева, 1, Новосибирск, 630090, Россия*

chekmarev@itp.nsc.ru

**ДИНАМИКА УКЛАДКИ БЕЛКА TRP-CAGE
В РАВНОВЕСНЫХ И НЕРАВНОВЕСНЫХ УСЛОВИЯХ:
МОДЕЛИРОВАНИЕ МЕТОДОМ МОЛЕКУЛЯРНОЙ ДИНАМИКИ**

Изучение динамики укладки белка в функциональное (нативное) состояние является одной из актуальных задач молекулярной биологии. Для решения этой задачи широко используется моделирование методом молекулярной динамики. При этом выбор условий, при которых проводится моделирование, играет существенную роль в корректном описании процесса укладки белка. В данной работе мы исследуем динамику укладки одного из «реперных» белков (Trp-cage) в двух условиях – в равновесных, когда белок многократно сворачивается и разворачивается, и в неравновесных, когда генерируется ансамбль траекторий, начинающихся в развернутом состоянии и заканчивающихся в нативном, что соответствует физиологическим условиям. Показано, что поведение белка в этих случаях существенно отличается, в частности в неравновесных условиях наблюдается дополнительное метастабильное состояние, приводящее к новому пути укладки. Моделирование показало также, что кинетика укладки Trp-cage носит одностадийный характер, что соответствует экспериментальным данным.

Ключевые слова: укладка белков, молекулярная динамика, неравновесные условия, тип кинетики, пути укладки.

Введение

После синтеза на рибосоме белок представляет собой протяженную цепочку аминокислотных остатков. Для выполнения своих функций он укладывается в компактную трехмерную структуру, называемую нативной [1]. Наиболее интенсивно исследования механизма укладки белков развивались начиная с 60-х гг. прошлого века, когда было показано, что нативная структура белка в основном определяется последовательностью аминокислотных остатков [2]. За это время был развит широкий арсенал методов

экспериментального исследования [3] и компьютерного моделирования [4], а также достигнуты большие успехи в понимании механизма укладки [1; 5–7]. Тем не менее многие вопросы остаются открытыми. Один из них – это соотношение «равновесной» динамики укладки, когда белок многократно сворачивается до нативного состояния, а потом разворачивается, и «неравновесной», когда белок приходит в нативное состояние и остается в нем [8]. В силу разных причин – по традиции или вследствие более простой реализации – подавляющее число исследований (экспериментов и компьютер-

* Работа выполнена при поддержке Российского научного фонда (проект № 14-14-00325).

ного моделирования) проведено для равновесных условий. В то же время для практики наибольший интерес представляют неравновесные условия, поскольку в физиологических условиях белок стабилен и вероятность его разворачивания очень мала. При компьютерном моделировании обычно выбираются равновесные условия ввиду существенно меньшего объема вычислений – представительную статистику по вероятностям пребывания белка в различных состояниях удастся набрать на одной траектории, которая включает порядка десятка событий сворачивания-разворачивания белка. В неравновесных же условиях, чтобы набрать представительную статистику по путям и времени укладки, требуется рассчитать ансамбль траекторий укладки, состоящий, по крайней мере, из сотни траекторий. Если учесть, что моделирование эволюции короткого белка в несколько десятков аминокислотных остатков методом молекулярной динамики в явном растворителе (в виде молекул воды) на одном процессоре (ядре) дает порядка 1 нс в день, а типичное время укладки – порядка 10 мкс, это выливается в громадный объем вычислительной работы ($100 \times 10\,000$ дней). Поэтому расчеты укладки белков в неравновесных условиях в явном растворителе практически отсутствуют. Более экономичный подход – использование растворителя в неявном виде, путем введения соответствующих добавок в энергию взаимодействия атомов в белке [9; 10]. Отсутствие трения атомов белка о молекулы растворителя повышает скорость конформационных превращений примерно на два порядка, что делает моделирование укладки белков в неравновесных условиях приемлемым по затратам. Существенно, что хотя отсутствие явного растворителя, конечно, искажает «тонкую» динамику укладки белка, наблюдается хорошо известная корреляция по времени формирования элементов вторичной структуры в расчетах с неявным растворителем и в экспериментах: в расчетах альфа-спирали формируются за 1 нс и бета-шпильки за 10 нс [11], а в эксперименте за 0,1 мкс и 1 мкс соответственно [12]. Поэтому реальное время укладки можно оценить, вводя поправочный коэффициент порядка 100.

В данной работе, используя моделирование методом молекулярной динамики в неявном растворителе, мы сравниваем укладку

одного из «реперных» альфа-бета белков (Trp-cage) в неравновесных и равновесных условиях. В равновесных условиях процесс укладки этого белка исследовался во многих работах [13–21], в частности экспериментально [13; 16; 17] и методом молекулярной динамики [14; 15; 18–21]. Основной обсуждаемый вопрос – это кинетика укладки, т. е. либо кинетика является одностадийной (экспоненциальной), когда переход осуществляется между двумя состояниями – от развернутых конформаций белка (реагента) к нативному (продукту), либо она носит более сложный характер. Единственной известной нам работой, где динамика укладки Trp-cage моделировалась в неравновесных условиях, является недавно опубликованная работа [21], в которой было рассчитано 25 траекторий укладки Trp-cage в явном растворителе и сделан вывод, что кинетика укладки сложнее, чем одностадийная, что противоречит эксперименту [13]. Проведенное нами исследование с улучшенной статистикой (100 траекторий) показывает, однако, что кинетика одностадийная, а это соответствует экспериментальным данным [13].

Белок и метод исследования

Белок Trp-cage представляет собой последовательность аминокислотных остатков Asn1-Leu2-Tyr3-Ile4-Gln5-Trp6-Leu7-Lys8-Asp9-Gly10-Gly11-Pro12-Ser13-Ser14-Gly15-Arg16-Pro17-Pro18-Pro19-Ser20 (1L2Y.pdb [22]). Для моделирования динамики укладки белка использовался подход, подобный тому, что применялся ранее для beta3s мини-протеина [8; 23; 24]. Конкретно, моделирование проводилось в среде пакета CHARMM [25]. Все тяжелые атомы и атомы водорода, связанные с атомами азота и кислорода, рассматривались явно. Для моделирования взаимодействия атомов внутри белка использовалось поле сил PARAM19 [9] и порог 7,5 Å для несвязывающих взаимодействий, т. е. между аминокислотными остатками, которые не являются соседями по белковой цепи. Водное окружение учитывалось неявно, используя метод поверхности, доступной для растворителя (SAS) [10]. Процесс укладки белка моделировался методом молекулярной динамики. Для поддержания температуры использовался термостат Берендсена [26] как в равновесных,

так и в неравновесных условиях. Длины ковалентных связей, вовлекающих атомы водорода, ограничивались с помощью алгоритма SHAKE [27], что позволяло проводить расчеты с шагом 2 фс.

Для неравновесных условий процесс укладки стартовал с развернутого состояния белка, которое готовилось путем выдерживания белка, сгенерированного по аминокислотной последовательности с помощью CHARMM [25], при повышенной температуре ($T = 500$ К) в течение 5×10^3 временных шагов, и заканчивался при достижении нативного состояния. Достижение нативного состояния определялось по числу нативных контактов, т. е. контактов, которые присутствуют в нативной структуре. Контакт считался сформированным, если расстояние между C_α -атомами в аминокислотных остатках, не являющихся соседями по цепи, не превышало $6,5 \text{ \AA}$, как в работе [14]. В качестве нативного состояния была принята ЯМР структура Trp-cage под номером 16, для которой, из общего числа 38 ЯМР структур [22], число нативных контактов максимально (37 контактов). Нативное состояние считалось достигнутым, если число нативных контактов в текущей конформации становилось равным числу контактов в нативной структуре. Соответственно время укладки белка определялось как время первого достижения нативного состояния при старте белка из начального, развернутого состояния. Число нативных контактов в начальных состояниях не превышало 7.

Для получения обозримой картины динамики укладки многомерное конформационное пространство белка сводилось к сокращенному пространству коллективных переменных. В качестве таких переменных использовались доля нативных контактов (q), т. е. отношение текущего числа нативных контактов к числу контактов в нативном состоянии, и среднеквадратичное отклонение от нативного состояния (rmsd), также рассчитанное по C_α -атомам.

Результаты и обсуждение

Первым этапом изучения динамики укладки было исследование зависимости времени укладки белка от температуры. Для этого было проведено моделирование укладки белка в неравновесных условиях при различных температурах. Для каждой темпера-

туры было смоделировано 100 траекторий. На рис. 1 показано среднеквадратичное отклонение (rmsd) конечных структур, отвечающих достижению нативного состояния, от ЯМР структуры № 16 [22], которая была принята в качестве нативной структуры. Видно, что нативное состояние достигается с хорошей точностью – в среднем rmsd составляет $\approx 2 \text{ \AA}$, со стандартным отклонением от него $\approx 0,6 \text{ \AA}$, что соизмеримо с соответствующими величинами для ансамбля ЯМР структур Trp-cage [22] – среднее rmsd от структуры № 16 равно $\approx 1 \text{ \AA}$, а стандартное отклонение от этого среднего значения равно $\approx 0,3 \text{ \AA}$.

Установлено, что зависимость среднего времени укладки от температуры имеет типичный U-образный характер, в данном случае с минимумом в районе 300–305 К, что соответствует физиологическим условиям (рис. 2). Минимум формируется за счет того, что с понижением температуры уменьшаются скорости переходов, а с повышением температуры белок большее время находится в полукompактном состоянии [28]. Время укладки белка при $T = 300$ К составляет 35 нс, что находится в хорошем соответствии с экспериментальными данными (4,1 мкс) [13], если учесть отмеченное выше увеличение скорости конформационных превращений в неявном растворителе по сравнению с экспериментом на два порядка [11; 12].

При температуре 300 К было проведено моделирование укладки белка в равновесных условиях. Моделирование стартовало из нативного состояния, а длина траектории составляла 6 мкс, в течение которой белок

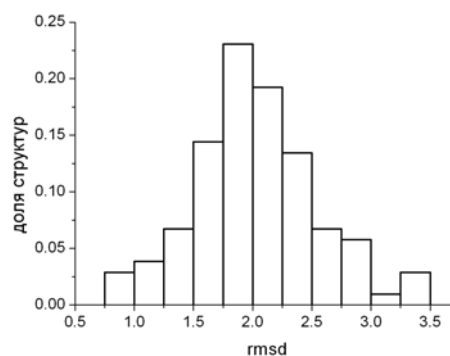


Рис. 1. Распределение достигнутых структур по среднеквадратичному отклонению (в ангстремах) от ЯМР нативной структуры

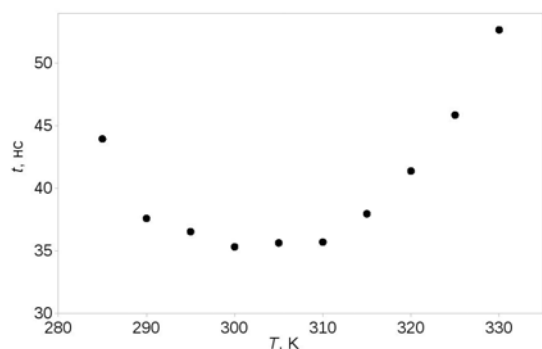


Рис. 2. Зависимость среднего времени укладки от температуры

сворачивался и разворачивался около 200 раз. На рис. 3 показаны поверхности свободной энергии для равновесных и неравновесных условий в зависимости от доли нативных контактов q и среднеквадратичного

отклонения от нативной структуры rmsd. Свободная энергия вычислялась как

$$F(q, \text{rmsd}) = -k_B T \ln P(q, \text{rmsd}),$$

где $P(q, \text{rmsd})$ – найденная вероятность пребывания системы в точке (q, rmsd) , а k_B – постоянная Больцмана. Буквами А, В и С отмечены бассейны притяжения, соответствующие характерным состояниям белка. Бассейн А содержит нативноподобные состояния белка. На поверхности свободной энергии для неравновесных условий он выражен менее отчетливо (более мелкий), потому что траектории обрываются при достижении нативного состояния. Бассейны, отмеченные буквами В и С, соответствуют метастабильным состояниям. Видно, что поверхность свободной энергии для неравновесных условий содержит дополнительный бассейн С. Наличие двух областей с различным числом нативных контактов,

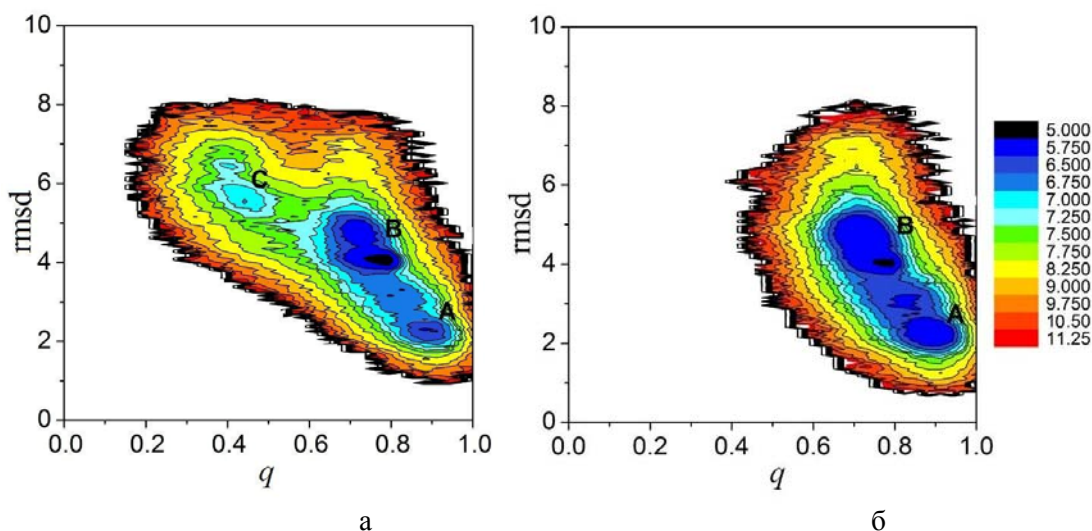


Рис. 3. Поверхности свободной энергии при $T = 300$ К: а – неравновесные условия; б – равновесные условия

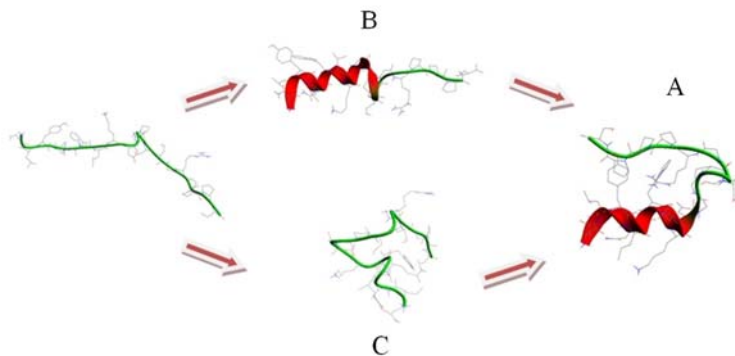
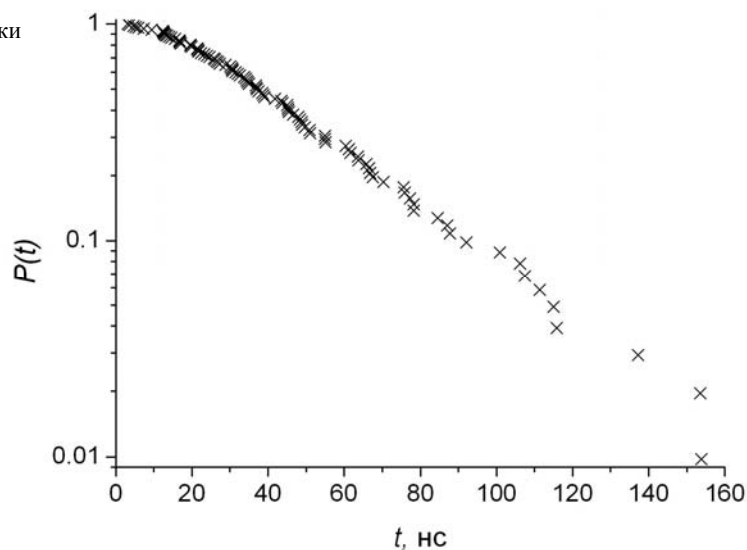


Рис. 4. Иллюстрация путей перехода белка из развернутого в нативное состояние

Рис. 5. Распределение времени укладки при $T = 300$ К



но близкими значениями $rmsd$ на этой поверхности, т. е. бассейнов В и С, свидетельствует о существовании различных траекторий укладки (рис. 4), что согласуется с результатами, полученными в работах [15; 21], где моделирование проводилось с применением явного растворителя (в [15] моделировался лишь начальный участок неравновесной траектории в 36 нс от полного среднего времени укладки в 3,73 мкс, полученного в [21]).

Чтобы определить тип кинетики укладки, мы рассчитали распределение времен укладки. На рис. 5 показано распределение вероятности нахождения белка в развернутом состоянии для $T = 300$ К, определяемое как

$$P(t) = \int_t^{\infty} p(t) dt,$$

где $p(t)$ – плотность вероятности времени укладки. Экспоненциальный характер этой зависимости свидетельствует о том, что существуют только два существенных состояния – несвернутый и свернутый белок, а внепутевые метастабильные состояния (интермедиаты) отсутствуют [28]. Это означает, что в отличие от выводов работы [21], кинетика укладки носит одностадийный характер в соответствии с экспериментальными данными [13].

Заключение

Показано, что динамика укладки Trp-cage в неравновесных условиях, т. е. когда в соответствии с поведением белков в физиоло-

гических условиях траектории укладки заканчиваются в нативном состоянии белка, отличается от динамики белка в равновесных условиях. Это выражается, в частности, в том, что на поверхности свободной энергии появляется бассейн притяжения (метастабильное состояние), который отсутствует в равновесных условиях. Присутствие этого метастабильного состояния приводит к дополнительному пути реакции. С учетом известной корреляции между скоростями конформационных превращений при моделировании укладки белка в приближении неявного растворителя и в эксперименте среднее время укладки хорошо согласуется с экспериментальными данными. Расширение ансамбля траекторий укладки по сравнению с работой [21], где ранее проводилось моделирование укладки Trp-cage в неравновесных условиях, позволило также вычислить распределение времени укладки. Оно показало, что в отличие от выводов, сделанных в [21], кинетика укладки носит одностадийный характер, что соответствует экспериментальным данным [13].

Список литературы

1. Финкельштейн А. В., Птицын О. Б. Физика белка. М.: Университет, 2002. 376 с.
2. Anfinsen C. B., Haber E., Sela M. et al. The Kinetics of Formation of Native Ribonuclease During Oxidation of the Reduced Polypeptide Chain // Proc. Natl. Acad. Sci. USA. 1961. Vol. 47. P. 1309–1314.

3. Bartlett A. I., Radford S. E. An Expanding Arsenal of Experimental Methods Yields an Explosion of Insights into Protein Folding Mechanisms // *Nature Struct. Mol. Biol.* 2009. Vol. 16. P. 582–588.
4. Lippert R. A., Predescu C., Ierardi D. J. et al. Accurate and efficient integration for molecular dynamics simulations at constant temperature and pressure // *J. Chem. Phys.* 2013. Vol. 139. P. 164106.
5. Onuchic J. N., Luthey-Schulten Z., Wolynes P. G. Theory of Protein Folding: The Energy Landscape Perspective // *Annu. Rev. Phys. Chem.* 1997. Vol. 48. P. 545–600.
6. Dinner A. R., Šali A., Smith L. J. et al. Understanding protein folding via free energy surfaces from theory and experiment // *Trends Biochem. Sci.* 2000. Vol. 25. P. 331–339.
7. Dill K. A., Ozkan S. B., Shell M. S. et al. The protein folding problem // *Annu. Rev. Biophys.* 2008. Vol. 37. P. 289–316.
8. Kalgin I. V., Chekmarev S. F., Karplus M. First Passage Analysis of the Folding of a Beta-Sheet Miniprotein: Is it More Realistic Than the Standard Equilibrium Approach? // *J. Phys. Chem. B.* 2014. Vol. 118. P. 4287–4299.
9. Neria E., Fischer S., Karplus M. Simulation of activation free energies in molecular systems // *J. Chem. Phys.* 1996. Vol. 105. P. 1902–1921.
10. Ferrara P., Apostolakis J., Caflisch A. Evaluation of a fast implicit solvent model for molecular dynamics simulations // *Proteins: Structure, Function, and Bioinformatics.* 2002. Vol. 46. P. 24–33.
11. Ferrara P., Apostolakis J., Caflisch A. Thermodynamics and Kinetics of Folding of Two Model Peptides Investigated by Molecular Dynamics Simulations // *J. Phys. Chem. B.* 2000. Vol. 104. P. 5000–5010.
12. Eaton W. A., Muñoz V., Hagen S. J. et al. Fast Kinetics and Mechanisms in Protein Folding. *Annu. Rev. Biophys. Biomol. Struct.* 2000. Vol. 29. P. 327–359.
13. Qiu L., Pabit S. A., Roitberg A. E., Hagen S. J. Smaller and Faster: The 20-Residue Trp-cage Protein Folds in 4 μ s // *J. Am. Chem. Soc.* 2002. Vol. 124. P. 12953.
14. Zhou R. Trp-cage: Folding free energy explicit water // *PNAS.* 2003. Vol. 100. P. 13280–13285.
15. Juraszek J., Bolhuis P. G. Sampling the multiple folding mechanisms of Trp-cage in explicit solvent // *PNAS.* 2006. Vol. 103. P. 15859–15864.
16. Byrne A., Williams D. V., Barua B. et al. Folding Dynamics and Pathways of the Trp-Cage Miniproteins // *Biochemistry* 2014. Vol. 53. P. 6011–6021.
17. Meuzelaar H., Marino K. A., Huerta-Viga A. et al. Folding Dynamics of the Trp-Cage Miniprotein: Evidence for a Native-Like Intermediate from Combined Time-Resolved Vibrational Spectroscopy and Molecular Dynamics Simulations // *J. Phys. Chem. B.* 2013. Vol. 117. P. 11490–11501.
18. Hatch H. W., Stillinger F. H., Debenedetti P. G. Computational Study of the Stability of the Miniprotein Trp-cage, the GB1 β -Hairpin, and the AK16 Peptide, under Negative Pressure // *J. Phys. Chem. B.* 2014. Vol. 118. P. 7761–7769.
19. Orsi M., Ding W., Palaiokostas M. Direct Mixing of Atomistic Solutes and Coarse-Grained Water // *J. Chem. Theory Comput.* 2014. Vol. 10. P. 4684–4693.
20. Doshi U., Hamelberg D. Towards fast, rigorous and efficient conformational sampling of biomolecules: Advances in accelerated molecular dynamics // *Biochimica et Biophysica Acta.* 2015. Vol. 1850. P. 878–888.
21. Kim S. B., Dsilva C. J., Kevrekidis I. G., Debenedetti P. G. Systematic characterization of protein folding pathways using diffusion maps: Application to Trp-cage miniprotein // *J. Chem. Phys.* 2015. Vol. 142. P. 085101.
22. Neidigh J. W., Fesinmeyer R. M., Andersen N. H. Designing a 20-residue protein // *Nat. Struct. Biol.* 2002. Vol. 9. P. 425–430.
23. Калгин И. В., Чекмарев С. Ф. Самоподобие потоков укладки белков // *Вестн. Новосиб. гос. ун-та. Серия: Физика.* 2012. Т. 7, вып. 4. С. 136–141.
24. Kalgin I. V., Chekmarev S. F. Folding of a Beta-Sheet Miniprotein: Probability Fluxes, Streamlines, and the Potential for the Driving Force // *J. Phys. Chem. B* 2015. Vol. 119. P. 1380–1387.
25. Brooks B. R., Bruccoleri R. E., Olafson B. D. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations // *J. Comp. Chem.* 1983. Vol. 4. P. 187–217.
26. Berendsen H. J. C., Postma J. P. M., DiNola A. et al. Molecular Dynamics with Coupling to an External Bath // *J. Chem. Phys.* 1984. Vol. 81. P. 3684–3690.

27. Ryckaert J.-P., Ciccotti G., Berendsen H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-alkanes // J. Comput. Phys. 1977. Vol. 23. P. 327–341.

28. Chekmarev S. F., Krivov S. V., Karp-lus M. Folding Time Distributions as an Approach to Protein Folding Kinetics // J. Phys. Chem. B. 2005. Vol. 109. P. 5312–5330.

Материал поступил в редколлегию 22.04.2015

V. A. Andryushchenko, S. F. Chekmarev

*Novosibirsk State University
2 Pirogov Str., Novosibirsk, 630090, Russian Federation*

*Institute of Thermophysics of SB RAS
1 Lavrentiev Ave., Novosibirsk, 630090, Russian Federation*

chekmarev@itp.nsc.ru

**FOLDING OF TRP-CAGE PROTEIN
IN EQUILIBRIUM AND NONEQUILIBRIUM CONDITIONS:
MOLECULAR DYNAMICS SIMULATIONS**

The study of the dynamics of protein folding into its functional (native) state is one of the actual problems of molecular biology. For this, molecular dynamics simulations are widely used. The conditions under which the simulations are performed are important for the correct description of the folding process. In the present paper, we study the folding dynamics of one of the benchmark proteins (Trp-cage) under two conditions – the equilibrium conditions, when the protein repeatedly folds and unfolds, and under nonequilibrium conditions, when an ensemble of trajectories is generated that start in an unfolded state of the protein and are terminated in the native state, which corresponds to the physiological conditions. It is shown that the behavior of the protein under these conditions is essentially different; in particular, in the case of nonequilibrium conditions an additional metastable state is formed, which leads to a separate folding pathway. The simulations have also shown that the kinetics of Trp-cage are two-state, which corresponds to the experimental results.

Keywords: protein folding, molecular dynamics, nonequilibrium conditions, kinetics type, folding pathway.