

Модель конформационной подвижности белка на основе принципа переноса масс

Г. С. Тамазян

*Центр геномной биоинформатики им. Ф. Г. Добржанского**
Санкт-Петербургский государственный университет
gaik.tamazian@gmail.com

С. Н. Князев

Лаборатория биоинформатики
Санкт-Петербургский национальный исследовательский университет ИТМО
sergey.n.knyazev@gmail.com

Е. О. Степанов

Кафедра математической физики, математико-механический факультет
Санкт-Петербургский государственный университет,
Санкт-Петербургское отделение Математического института им. В. А. Стеклова РАН
e.stepanov@sns.it

Ю. Б. Порозов

Лаборатория биоинформатики
Санкт-Петербургский национальный исследовательский университет ИТМО
porozov@ifc.cnr.it

Аннотация

Предлагается новый способ моделирования конформационных движений белковых молекул, основанный на принципе переноса массы и приближенном (coarse-grained) представлении белка. В отличие от методов молекулярной динамики, рассмотренный метод предназначен для моделирования движений, которые происходят на сравнительно протяженных (порядка миллисекунд) временных интервалах. Предложена математическая модель движения белка, поставлена задача получения конформационного движения и дан алгоритм ее решения. Приводятся результаты численных экспериментов, показывающие адекватность предложенной модели.

1. Введение

Белковые молекулы играют основную роль в биохимических процессах, протекающих в живых клетках. Относясь к макромолекулам, белки характеризуются не только атомарным составом, но и пространственной (третичной) структурой. При этом существует зависимость между структурной организацией белка и его функциональной активностью, которая проявляется в моменты перехода белка из одной конформации в другую [1].

Ввиду связи между структурой, подвижностью и функцией белка возникает необходимость в моделировании конформационной подвижности белковых молекул. В настоящее время существует множество методов для решения подобных задач, однако большинство из них (например, методы молекулярной динамики [2] или эластичных сетей [3]) основываются на анализе внутренней энергии рассматриваемой молекулы и не применимы к моделированию протяженных (порядка нескольких миллисекунд) движений, в ходе которых молекула будет подвергаться воздействию внешних сил. В данной рабо-

*Работа поддержана мегагрантом министерства образования и науки номер 11.G34.31.0068, руководитель доктор С. Дж. О'Брайен.

те предлагается модель конформационного движения, основывающаяся на приближенном представлении белковой молекулы и процесса ее перемещения между двумя пространственными конфигурациями.

2. Описание модели

Белковая молекула представляет собой полимер, основной единицей которого являются аминокислоты. Поскольку конформационная подвижность белка обеспечивается изменением структуры главной цепи, в дальнейшем мы будем рассматривать только принадлежащие ей атомы. Подобная модель конформации, в которой игнорируются атомы боковых цепей белка, относится к приближенным (coarse-grained) [4] моделям.

2.1. Представление конформации

Рассмотрим конформацию $x = (x_1, \dots, x_n)$, заданную декартовыми координатами атомов главной цепи белка. Перейдем к представлению конформации x в виде набора трех векторов: вектора длин связей r_i , планарных углов α_i и торсионных углов ψ_i :

$$x = \{ (r_i)_{i=1}^{n-1}, (\alpha_i)_{i=1}^{n-2}, (\psi_i)_{i=1}^{n-3} \}, \quad (1)$$

где величины r_i , α_i и ψ_i задаются следующим образом:

$$r_i = \|x_{i+1} - x_i\|, \quad (2)$$

$$\alpha_i = \arccos \left(\frac{\Delta x_{i+1} \cdot \Delta x_i}{\|\Delta x_{i+1}\| \|\Delta x_i\|} \right), \quad (3)$$

$$\psi_i = \text{sign}((N_i \times N_{i+1}) \cdot \Delta x_{i+1}) \arccos \left(\frac{\Delta N_{i+1} \cdot \Delta N_i}{\|\Delta N_{i+1}\| \|\Delta N_i\|} \right), \quad (4)$$

где $\Delta x_i = x_{i+1} - x_i$, $N_i = \Delta x_i \times \Delta x_{i+1}$, $\|\cdot\|$ обозначает евклидову норму вектора. Данное представление конформации удобно тем, что позволяет явно оперировать торсионными углами ψ_i , оставляя длины связей r_i и планарные углы α_i неизменными, что согласуется с биохимическими законами, по которым подвижность главной цепи белка обуславливается изменением торсионных углов.

2.2. Представление конформационного движения

Пусть имеются две конформации x^0 и x^1 одного и того же белка, заданные в приведенном выше виде:

$$x^0 = \{ (r_i^0)_{i=1}^{n-1}, (\alpha_i^0)_{i=1}^{n-2}, (\psi_i^0)_{i=1}^{n-3} \}, \quad (5)$$

$$x^1 = \{ (r_i^1)_{i=1}^{n-1}, (\alpha_i^1)_{i=1}^{n-2}, (\psi_i^1)_{i=1}^{n-3} \}. \quad (6)$$

Будем называть любую конформацию $x = \{ (r_i)_{i=1}^{n-1}, (\alpha_i)_{i=1}^{n-2}, (\psi_i)_{i=1}^{n-3} \}$ *допустимой относительно x^0* , если $r_i = r_i^0$ и $\alpha_i = \alpha_i^0$. Пусть x^0 и x^1

допустимы относительно друг друга. Будем называть *допустимым движением* молекулы белка между конформациями x^0 и x^1 липшицеву функцию

$$\gamma: [0, 1] \rightarrow \mathbb{R}^{3n} \quad (7)$$

такую, что $\gamma(t)$ является допустимой относительно x^0 для любого $t \in [0, 1]$. Иначе говоря, функция γ в момент времени t характеризует положение всех n атомов главной цепи рассматриваемого белка. Допустимость движения γ означает постоянство длин связей и планарных углов в ходе такого движения.

3. Получение конформационного движения

3.1. Постановка задачи

Поставим задачу получения конформационного движения белка между заданными конформациями x^0 и x^1 следующим образом: искомое движение γ находится как решение задачи минимизации функционала

$$F(\gamma) = \sum_{j=1}^n m_j l_j^p, \quad (8)$$

где l_j — длина пути, пройденного j -ым атомом главной цепи при движении γ от x^0 к $T(x^1)$, m_j — масса этого атома, $p \geq 1$ — параметр модели. Минимизация производится по всевозможным допустимым относительно x_0 движениям γ и твердотельным перемещениям T .

3.2. Решение задачи

Задачу (8) предлагается решать численно путем дискретизации модели движения следующим образом. Искомое допустимое движение γ представляется в виде последовательности из M конформаций, из которых начальная и конечная определяются заданными конформациями x^0 и x^1 . Функционал (8) приобретает следующий вид:

$$\tilde{F}(\tilde{\gamma}) = \sum_{j=1}^n m_j \left(\sum_{k=2}^M \|x_j^k - x_j^{k-1}\| \right)^p, \quad (9)$$

Движение $\tilde{\gamma} = \{ (x_j^k)_{j=1}^n \}_{k=1}^M$, являющееся решением этой задачи, представляет собой приближение к искомому движению γ . Задача минимизации функционала (9) может быть решена методами локальной оптимизации [5].

4. Результаты

Был проведен сравнительный анализ описанного метода со следующими методами моделирования конформационной подвижности белков [6]

Таблица 1. Трансформации белков, использованные для сравнения методов моделирования конформационной подвижности. N соответствует количеству аминокислот в главной цепи белка. Значения RMSD приведены в ангстремах. Начальные и конечные конформации заданы номерами доступа в Protein Data Bank и номерами моделей.

Белок	N	Начало	Конец	RMSD
Eel Calcitonin	33	1BYV:1	1BYV:7	7.77
Eel Calcitonin	33	1BZB:1	1BZB:10	10.7
Суановирин-N	101	1J4V:1	1L5B:1	6.41
Calmodulin	149	1PRW:1	1OSA:1	16.25
$\gamma\delta$ Resolvase	183	1GDT:1	2RSL:1	2.11

- модель эластичных сетей [7] (сервер PATH-ENM: <http://enm.lobos.nih.gov>),
- метод линейной интерполяции декартовых координат [8] (сервер MovieMaker: <http://wishart.biology.ualberta.ca/moviemaker>),
- метод комбинированной интерполяции декартовых координат и торсионных углов с минимизацией энергий промежуточных конформаций [9] (сервер Morph: <http://molmovdb.org>).

Для сравнения методов были выбраны трансформации между парами конформаций пяти белков (см. таблицу 1). RMSD в таблице даны для атомов альфа-углеродов главной цепи белка.

Трансформации, полученные для указанных белков методами PATH-ENM и MovieMaker, отличаются искаженной геометрией молекулы (нарушениями длин связей и планарных углов). Результаты, полученные методом molmovdb и предложенным методом, свободны от данного недостатка и в целом близки друг к другу.

Список литературы

[1] Bahar I., Lezon T. R., Yang L.-W., Eyal E. Global dynamics of proteins: bridging between structure and function // Annual review of biophysics. 2010. Vol. 39. P. 23.

[2] Karplus M., Petsko G. A. Molecular dynamics simulations in biology // Nature. 1990. Vol. 347, no. 6294. P. 631–639.

[3] Atilgan A., Durell S., Jernigan R. et al. Anisotropy of fluctuation dynamics of proteins with an elastic network model // Biophysical Journal. 2001. Vol. 80, no. 1. P. 505–515.

[4] Levitt M. A simplified representation of protein conformations for rapid simulation of protein folding // Journal of molecular biology. 1976. Vol. 104, no. 1. P. 59–107.

[5] Nocedal J., Wright S. Numerical optimization, series in operations research and financial engineering // Springer, New York. 2006.

[6] Кошевой А. А. Метод предсказания и оптимизации конформационной подвижности белков путем решения транспортной задачи переноса массы: Диссертация на соискание ученой степени магистра / Санкт-Петербургский национальный университет ИТМО. Санкт-Петербург, 2011.

[7] Zheng W., Doniach S. A comparative study of motor-protein motions by using a simple elastic-network model // Proceedings of the National Academy of Sciences. 2003. Vol. 100, no. 23. P. 13253–13258.

[8] Maiti R., Van Domselaar G. H., Wishart D. S. MovieMaker: a web server for rapid rendering of protein motions and interactions // Nucleic acids research. 2005. Vol. 33, no. suppl 2. P. W358–W362.

[9] Flores S., Echols N., Milburn D. et al. The Database of Macromolecular Motions: new features added at the decade mark // Nucleic acids research. 2006. Vol. 34, no. suppl 1. P. D296–D301.