

# Лекция 12: Продвинутые методы молекулярной динамики

## 1 Метод нормальных мод (normal modes analysis)

Значительную роль в изучении динамики белка играет определение и описание его глобальных движений в равновесии, то есть тех движений, которые совершаются одновременно большой совокупностью атомов с одинаковой частотой. Одним из распространенных методов для нахождения таких смещений служит метод нормальных движений (мод) [1, 2]. Главное предположение, на котором основан метод, состоит в том, что рассматриваемая система находится в глобальном или в каком-либо из глубоких локальных минимумов. В таком случае, потенциал полной энергии, зависящий от координат, может быть приближен гармонической функцией. Тогда движение системы можно рассматривать как сумму гармонических осцилляторов с определенной собственной частотой.

Таким образом, нахождение нормальных мод системы сводится к решению задачи диагонализации матрицы гессиана потенциальной энергии:

$$R^T M^{-1/2} H M^{-1/2} R = \text{diag}(\lambda_1 \dots \lambda_1), \quad (1)$$

$$\lambda_i = (2\pi\omega_i)^2 \quad (2)$$

где  $M$  - это матрица масс атомов,  $R$  - матрица собственных векторов системы,  $\lambda_i$  - собственные значения, а  $\omega_i$  - соответствующие им собственные частоты. Матрица  $H$  имеет

размерность  $3N \times 3N$  ( $N$  - количество атомов в системе) и определяется следующим образом:

$$H_{ij} = \frac{\partial^2 V}{\partial q_i^2 \partial q_j^2}, \quad (3)$$

где  $q_i$  и  $q_j$  являются обобщенными, масс-взвешенными координатами атомов. Вообще говоря, для численного решения это уравнение не используется, а элементы матрицы гессиана определяются как:

$$H_{ij} = -\frac{f_i(\mathbf{q} + h\mathbf{e}_j) - f_i(\mathbf{q} - h\mathbf{e}_j)}{2h}, \quad (4)$$

$$f_i = -\frac{\partial V}{\partial q_i}, \quad (5)$$

здесь  $\mathbf{e}_j$  определяет вектор в направлении  $j$ -ого атома, а  $h$  - шаг смещения.

Наибольшую сложность в определении нормальных мод представляет диагонализация матрицы гессиана, так как процесс этот требует затрат времени порядка  $(3N)^2$  и памяти -  $(3N)^3$ , где  $N$  - это число степеней свободы системы. В связи с этим задача в том виде, как она представлена выше, может быть рассмотрена только для небольших систем [3]. Существует немало работ, посвященных адаптации метода для различных систем. Например, засчет диагонализации матрицы меньшего размера [4, 5], использования более упрощенного потенциала функции [6, 7] или изменения способа расчет матрицы [7, 8].

Для симметричных систем, анализ нормальных мод для определения главных движений и спектра частот вирусной капсиды, может быть преобразован, опираясь на алгоритм изложенный в [4]. Пользуясь свойством симметрии, система может быть представлена как совокупность блоков, и движения каждого такого блока можно рассматривать как движения твердого тела. Таким образом, матрица может рассчитываться, например, для  $C_\alpha$ -частиц, на основе модели упругого покрытия (ENM), а затем преобразуется в матрицу меньшей размерности для множества блоков следующим образом:

$$H^b = P^T H P \quad (6)$$

$H$  - матрица размера  $3N \times 3N$ , где  $N$  - число атомов,  $H^b$  - матрица размера  $6n_b \times 6n_b$ , где  $n_b$

- размер блока, матрица  $P$  является проектором, сужающим полноатомное пространство, до пространства рассматриваемых блоков.

Частным случаем этого метода является представление одного блока как центр массы одной аминокислоты. Рассматривается нахождение гессиана для центра массы одного или нескольких аминокислотных остатков [3]. В этом случае матрица-проектор будет описывать преобразования переноса и преобразование поворота. Этот метод с учетом лишь преобразования переноса:

$$P_{Jj} = \sqrt{\frac{m_j}{M_J}}, \quad (7)$$

где  $m_j$  и  $M_J$  соответственно масса атома ( $j$ ) и аминокислотного остатка ( $J$ ). Тогда элемент матрицы определяется по формуле:

$$H_{IJ} = \frac{1}{\sqrt{M_I M_J}} \sum_{i \in I, j \in J} H_{ij}, |J| = |I| = M \quad (8)$$

Элемент матрицы  $H_{ij}$  определяется по формуле (4) из потенциала силового поля.

Для выполнения диагонализации матрицы гессиана, нахождения ее собственных значений и векторов можно использовать стандартные библиотеки линейной алгебры (такие как LAPACK), которые встроены в том числе и в программный пакет GROMACS [9].

## 2 Метод главных компонент (essential dynamics)

Помимо определения равновесных свойств биомолекул, большой интерес представляет изучение динамики белка или его частей. Метод главных компонент (essential dynamics) позволяет исследовать траекторию движения системы полученную с помощью молекулярной динамики [10]. Основная идея состоит в том, чтобы пространство, определяемое  $3N$  степенями свободы, разделить на главное подпространство (essential subspace) и второстепенное. Первое содержит порядка 1% степеней свободы, которые описывают основные движения системы, в то время как оставшиеся степени свободы описывают лишь отдельные флуктуации атомов. Поэтому они могут рассматриваться как неподвижные, а их подпространство

ортогонально к главному подпространству [11].

Для нахождения главных степеней свободы строится ковариационная матрица атомных координат:

$$C_{ij} = \frac{1}{T} \sum_t (q_i(t) - \langle q_i \rangle)(q_j(t) - \langle q_j \rangle) \quad (9)$$

где  $T$  - количество конфигураций системы:  $t = 1, 2 \dots T$ ,  $q_i(t)$  - обобщенные координаты  $i$ -ого атома в момент  $t$ ,  $i = 1, 2 \dots 3N$ ,  $\langle q_i \rangle$  - среднее значение координаты для всех конфигураций. Матрица  $C$  диагонализуется, полученные собственные вектора задают новое пространство. Проецируя траекторию на эти вектора, можно найти представление движения в базисе нового пространства:

$$\mathbf{p}_i(t) = \mathbf{q}(t) \cdot \mathbf{R}_i, \quad (10)$$

полученные компоненты  $\mathbf{p}_i$  называются главными компонентами, а собственное значение  $\lambda_i$  задает их среднее квадратичное отклонение для  $i$ -ой компоненты. Первые несколько главных мод и определяют основные движения системы. Теперь динамику исследуемой системы можно представить одной или несколькими главными модами [12]:

$$\mathbf{q}'_i(t) = \langle \mathbf{q} \rangle + \mathbf{p}_i(t) \cdot \mathbf{R}_i. \quad (11)$$

Основное отличие существенной динамики от нормальных мод состоит в том, что нормальные моды описывают гармонические движения, происходящие в системе вокруг какого-либо локального минимума и не выводящие ее из равновесия, в то время как метод главных компонент выделяет негармоничность движения и характеризует систему в неравновесном состоянии. Если же рассматривается динамика на небольшой временной шкале, которое не выводит систему из ее локального минимума, то флуктуации атомов будут преобладать над главными модами движения. В этом случае анализ главных компонент почти эквивалентен определению нормальных мод и называется *квазигармоническим анализом* [2].

Метод главных компонент может также быть выполнен с использованием пакета GROMACS [9] для центров масс одного или двух аминокислотных остатков рассматриваемой системы. Как было показано [12], при таком приближении не происходит потери информации, но значительно уменьшается количество атомарных флуктуаций, позволяя более четко определить основные типы движения системы.

## Список литературы

- [1] L. D. Landau and E. M. Lifshiz, *Theoretical physics. Mechanics*. Moscow: Nauka, 2nd ed., 1965.
- [2] B. R. Brooks, D. Janežič, and M. Karplus, “Harmonic analysis of large systems. I. Methodology,” *J.Comp.Chem*, vol. 16, no. 12, pp. 1522–1542, 1995.
- [3] S. Jääskeläinen, C. S. Verma<sup>1</sup>, R. E. Hubbard, P. Lestko, and L. S. Caves, “Conformational change in the activation of lipase: An analysis in terms of low-frequency normal modes,” *Protein Sci.*, vol. 345, pp. 299–314, 2005.
- [4] F. Tama, F. X. Gadea, O. Marques, and Y.-H. Sanejouand, “Building-block approach for determining low-frequency normal modes of macromolecules,” *Proteins*, vol. 41, pp. 1–7, 2000.
- [5] G. Lia and Q. Cui, “A coarse-grained normal mode approach for macromolecules:an efficient implementation and application to  $ca^{2+}$ -ATPase,” *Biophys.J.*, vol. 83, pp. 2457–2474, 2002.
- [6] M. M. Tirion, “Large amplitude elastic motions in proteins from a single-parameter, atomic analysis,” *Phys.Rev.Lett.*, vol. 77, pp. 1905–1908, 1996.
- [7] I. Bahar, A. R. Atilgan, and B. Erman, “Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential,” *Fold.Des.*, vol. 2, pp. 173–181, 1997.
- [8] A. Atilgan, S. Durell, R. Jernigan, M. Demirel, O. Keskin, and I. Bahar, “Anisotropy of fluctuation dynamics of proteins with an elastic network model,” *Biophys.J.*, vol. 80, pp. 505–515, 2001.
- [9] E. Lindahl, B. Hess, and D. van der Spoel, “GROMACS 3.0: a package for molecular simulation and trajectory analysis,” *J.Mol.Model*, vol. 7, pp. 306–317, 2001.
- [10] G. H. Peters, D. M. van Aalten, A. Svendsen, and R. Bywater, “Essential dynamics of lipase binding sites: the effect of inhibitors of different chain length,” *Protein Eng.*, vol. 10, no. 2, pp. 149–158, 1997.
- [11] A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, “Essential dynamics of proteins,” *Proteins*, vol. 17, pp. 412–425, 1993.

- [12] D. M. F. V. Aalten, B. L. D. Groot, J. B. C. Findlay, H. J. C. Berendsen, and A. Amadei, “A comparison of techniques for calculating protein essential dynamics,” *J.Comp.Chem*, vol. 18, no. 2, pp. 169–181, 1997.