

Метод количественной оценки хиральности в иерархиях белковых структур

Сидорова А.Э.¹, Твердислов В.А.¹, Левашова Н.Т.¹, Котов А.Р.¹, Устинин М.Н.²

¹Московский государственный университет имени М.В. Ломоносова, Физический факультет, кафедра биофизики, Россия, Москва

²ИМПБ РАН – филиал ИПМ им. М.В. Келдыша РАН, Россия, Московская область, Пуццино
sky314bone@mail.ru, tverdislov@mail.ru, u_m_n@mail.ru

Разработан подход к количественной оценке степени хиральности в иерархических структурах макромолекулярных систем. Предложенный подход позволяет в общем случае количественно оценить знак и степень хиральности вторичных структур белков. Достаточным условием метода для определения знака хиральности вторичной структуры является выбор взаимного расположения α -углеродов, что позволяет на один-два порядка снизить количество обрабатываемой информации по сравнению с полным описанием системы. В результате анализа правых и левых спиральных вторичных структур показано, что угол между вектором направления и суммой векторных произведений выявляет направление закрутки спирали. Наряду с правыми α -спиралями и левыми полипролиновыми спиралями, которые, перевиваясь, образуют правые суперспирали в коллагене, были проанализированы и β -структуры. Данные о пространственной структуре отображаются на специальных «картах хиральности», которые демонстрируют наличие обособленных областей существования для различных типов спиральности. Метод реализован в виде компьютерной программы на языке Python 3.6 с использованием библиотеки SciPy и может быть применен к нуклеиновым кислотам.

Ключевые слова: белки, хиральность, энантиомеры, альфа-углерод, правая альфа-спираль, правая 310 спираль, левая полипролиновая спираль, бета-цепочка, векторное произведение, вектор направления, карта хиральности.

Quantitative approach to the assessment of chirality in hierarchies of protein structures

Sidorova A.E.¹, Tverdislov V.A.¹, Levashova N.T.¹, Kotov A.R.¹, Ustinin M.N.²

¹Department of Biophysics, Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia

²IMPB RAS – Branch of KIAM RAS, Moscow Region, Pushchino, Russia

An approach to the quantitative assessment of the degree of chirality in the hierarchical structures of macromolecular systems is developed. The proposed approach makes it possible in general to quantify the sign and degree of chirality of the secondary protein structures. A sufficient condition for the method for determining the sign of the chirality of the secondary structure is the choice of the relative arrangement of α -carbons, which allows one or two orders to reduce the amount of information processed compared with the complete description of the system. As a result of the analysis of right and left spiral secondary structures, it is shown that the angle between the direction vector and the sum of vector products reveals the direction of twist of the spiral. Along with the right α -helices and the left polyproline helices, which, forming themselves, form the right super-spirals in the collagen, β -structures were analyzed. Data on the spatial structure are displayed on special "chirality maps", which demonstrate the existence of isolated regions of existence for different types of helicity. The method is implemented as a computer program in Python 3.6 using the SciPy library and can be applied to nucleic acids.

Key words: proteins, chirality, enantiomers, alpha-carbon atom, right-handed alpha helix, right-handed 310 helix, left-handed polyproline helix, beta-chain, cross product, direction vector, chirality map.

Введение

Одной из важнейших физических проблем молекулярной биологии является определение механизмов стратификации в иерархиях первичных,

вторичных, третичных и четвертичных структур биомолекул и, как следствие, их функциональных особенностей [1, 2]. В работах [2–4] была выявлена фундаментальная закономерность смены знака хиральности в структурно-

функциональной иерархии белков и нуклеиновых кислот. Для белковых структур эта закономерность выражается следующим образом: в белках левые аминокислоты образуют полипептидную цепь, которая образует правую α -спираль; правые α -спирали, перевиваясь, образуют левую суперспираль; на четвертом иерархическом уровне белок имеет тенденцию к правой укладке. Последовательность смены знака хиральности в структурно-функциональной иерархии белковых структур L-D-L-D. Первичная структура белка представляет собой последовательность остатков левых аминокислот. Внутри молекулы полипептидная цепь укладывается в α -спираль или в складчатый β -слой. Как известно, α -спираль является правым энантиомером. Непосредственно взаимодействующие правые α -спирали стремятся образовать левую суперспираль. Четвертичная структура белков представлена надмолекулярными структурами, сформированными при правой укладке левых суперспиралей. Наподобие этой белковой иерархии, отмечается закономерное чередование знака хиральности D-L-D-L при переходе на более высокий уровень структурно-функциональной организации ДНК. Очевиден сдвиг по фазе между структурными уровнями этих классов макромолекул, поскольку знакопеременная белковая иерархия «стартовала» с L-аминокислот, а нуклеотидная – с D-углевода дезоксирибозы.

Обсуждаемые иерархии уже не на общепринятом феноменологическом, а с помощью фундаментального бинарного физического критерия хиральности, позволяют однозначно охарактеризовать дискретность в макромолекулярных структурах клеток от первичной (точнее, от «первичной» асимметричной) структуры до четвертичной. Несмотря на убедительность качественных оценок в данном подходе, принципиально необходимо уметь получать количественные оценки доли и знака хиральности структур разных иерархических уровней. Критерии хиральной дискретности принципиально важны для осмысления механизмов функционирования молекулярных машин клетки, поскольку именно эти дискретные уровни структур формируют в них «выделенные механические степени свободы».

На сегодняшний день имеется достаточное количество работ, в которых предлагаются различные методы выработки критерия нарушения симметрии. Однако эта проблема на настоящий момент по-прежнему не решена.

Одним из первых подходов количественной оценки хиральности можно считать асимметрическое произведение, которое ввел Гюи в 1890 году [5, 6]. Однако «функция хиральности» в данном случае специфична для различных классов молекулярных моделей, что неоднозначно определяет правило обхода для разных объектов и затрудняет применение к сложным молекулярным соединениям. Также в качестве меры хиральности

использовались отклонения двух множеств, позволяющие применять данный метод, как к дискретным, так и непрерывным множествам точек [7]. Концепция Рассата была расширена Мези [8] для облаков электронной плотности с использованием теории нечетких множеств. К недостаткам данного метода можно отнести следующее. Вычисление расстояния Хаусдорфа подразумевает наличие двух множеств для сравнения: первое множество – исследуемый энантиомер, а в качестве второго надо выбрать или противоположный стереоизомер, или эталонный рацемат. Но в молекулярной биологии далеко не на каждом иерархическом уровне можно выделить такие множества для сравнения. Кроме того, данный метод не фиксирует «переключение» знака хиральности для противоположных стереоизомеров. Непрерывные меры симметрии CSM (Continuous Symmetry Measures) и хиральности CCM (Continuous Symmetry Measures) были введены Забродским, Пелегом и Авниром [9–11], однако, по сути, CCM – это не мера хиральности, а лишь мера сходства исходного множества точек со своим зеркальным отображением.

В качестве работы, в которой исследовалась степень хиральности спиральных структур, можно отметить работу [12], в которой «степень хиральности» количественно определялась путем вычисления индекса хиральности на основе перекрытия проекций спиральных лент методом сличения объекта с его зеркальным отображением. В этом случае индекс хиральности принимает только положительные значения, что позволяет оценить симметричность, но не знак хиральности. Метод применим лишь к плоским спиральным лентам, коими являются далеко не все структуры в биомолекулах. К типичным исследованиям количественных аспектов дисимметрии относятся работы [13–17], основанные на геометрических построениях.

Наиболее широко известен метод, основанный на двугранных углах для аминокислотных остатков цепи – карты Рамачандрана [18–20]. Но карты, основанные на этих углах, определяют только преобладающую конформацию аминокислот в белке [21], не позволяют определять знак хиральности и не применимы к нуклеиновым кислотам. Одним из таких направлений является применение среднеквадратичной количественной оценки хиральности как меры симметрии [22] или рассмотрение симметрии как непрерывного свойства, но не обязательно как двоичного свойства [23]. Известным методом оценки хиральности является метод, основанный на анализе моментов вращательной поляризуемости, характеризующей оптическую активность [24]. Наш подход оценки хиральных структур носит более общий характер и позволяет более полно определять вторичную структуру: не только ее тип, но и знак хиральности с информацией о пространственной структуре.

Метод исследования

В предлагаемом нами методе не обязательно обрабатывать всю информацию обо всех атомах молекулы, которая записана в PDB-файлах [10]. Для определения знака хиральности вторичной структуры достаточным условием анализа является лишь взаимное расположение α -углеродов C_α . Это позволяет на порядок снизить количество обрабатываемой информации, что является явным преимуществом при обработке больших массивов данных, на что и направлен наш метод. Несмотря на то, что в PDB-формате существует классификация из 10 различных спиралей [10], наиболее часто встречающимся типом правой спирали является α -спираль. Левая же спираль представлена в виде спирали типа poly(Pro)II (как в коллагене). В данной работе рассмотрены вторичные спиральные структуры белков как две крайности нарушений симметрий на втором уровне структурной иерархии белков. На рисунке 1,А показана исходная молекула (PDB ID: 1L2P [11]), состоящая из 548 атомов, на рисунке 1,Б – остов этой молекулы из 61-го α -углерода.

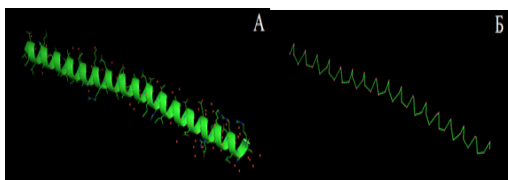


Рис. 1. Изображения субъединицы «б» АТФсинтазы *E. coli* (PDB ID: 1L2P) [11]: А – исходное изображение; Б – изображение остова только из α -углеродов.

Предлагаемый подход основан на векторных произведениях. Рассмотрим модель витка правой спирали: остов из α -углеродов и ось спирали, направленную от N-конца к С-концу (рис. 2). Считаем альфа-углерод, принадлежащий N-концу, первым атомом остова C_1^α , альфа-углерод, принадлежащий С-концу, – последним атомом остова C_n^α . Построив векторы между соседними атомами: от предыдущего атома к следующему, получаем векторы $C_1^\alpha C_2^\alpha$, $C_2^\alpha C_3^\alpha$, ..., $C_{(n-1)}^\alpha C_n^\alpha$ (рис. 2). Таким образом, для n атомов получается $(n-1)$ таких векторов, сумма которых является вектором $C_1^\alpha C_n^\alpha$. Этот вектор – «вектор направления» – обозначим \mathbf{d} :

$$\begin{aligned} C_1^\alpha C_2^\alpha + C_2^\alpha C_3^\alpha + \dots + C_{(n-1)}^\alpha C_n^\alpha = \\ = \sum_{i=1}^{(n-1)} C_i^\alpha C_{(i+1)}^\alpha = C_1^\alpha C_n^\alpha = \mathbf{d} \end{aligned} \quad (1)$$

Для каждой пары соседних векторов строим векторное произведение (рис. 2): $C_1^\alpha C_2^\alpha \times C_2^\alpha C_3^\alpha$, $C_2^\alpha C_3^\alpha \times C_3^\alpha C_4^\alpha$ и т.д., которое даст вектор, перпендикулярный двум исходным, и образующий с ними «правую тройку» [9]. Использование векторного произведения позволяет заложить в модель чувствительность к направлению закрутки спирали: поднимаясь в правой спирали по оси «вверх», векторные произведения направлены также «вверх» относительно направления оси спирали (рис. 2,А), а поднимаясь «вверх» в левой спирали – векторные произведения, по правилу правой руки, будут направлены уже «вниз» (рис. 2,Б). Для $(n-1)$ векторов получаем $(n-2)$ векторных произведений, сумма которых – вектор \mathbf{s} :

$$\begin{aligned} C_1^\alpha C_2^\alpha \times C_2^\alpha C_3^\alpha + C_2^\alpha C_3^\alpha \times C_3^\alpha C_4^\alpha + \dots = \\ = \sum_{i=1}^{(n-2)} C_i^\alpha C_{(i+1)}^\alpha \times C_{(i+1)}^\alpha C_{(i+2)}^\alpha = \mathbf{s} \end{aligned} \quad (2)$$

Зная длину и направление векторов \mathbf{d} и \mathbf{s} , можно через скалярное произведение найти угол между этими векторами:

$$\angle(\mathbf{d}, \mathbf{s}) = \arccos \left(\frac{\mathbf{d} \cdot \mathbf{s}}{|\mathbf{d}| |\mathbf{s}|} \right). \quad (3)$$

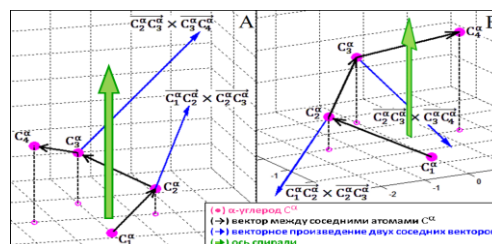


Рис. 2. Модели витков спирали: А – виток правой спирали; Б – виток левой спирали.

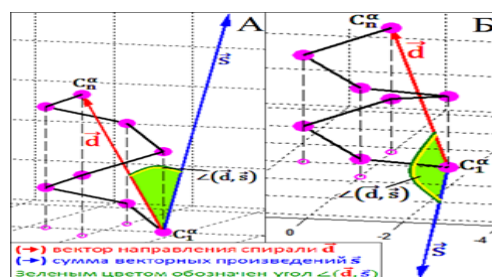


Рис. 3. Угол между вектором направления \mathbf{d} и суммой векторных произведений \mathbf{s} : А – правая спираль; Б – левая спираль.

Таблица. Перечень молекул из [11], использованные для построения «карты хиральности» вторичных структур белков

| № точки | PDB ID | № точки | PDB ID | № точки | PDB ID | № точки | PDB ID | № точки | PDB ID |
|---------|--------|---------|--------|---------|--------|---------|--------|---------|--------|
| 1–3 | 1CAG | 24 | 5U9R | 43–48 | 2D3H | 128–143 | 1OUT | 186–188 | 4DMT |
| 4 | 1L2P | 25 | 5UA7 | 49–51 | 2KLW | 144–152 | 2A7U | 189–194 | 3A0A |
| 5–7 | 1QL2 | 26–27 | 5WST | 52–57 | 3A1H | 153–178 | 2YQ8 | 195–205 | 2N8R |
| 8–17 | 1TJC | 28–30 | 1BKV | 58–88 | 1BDG | 179 | 4OY5 | 206–216 | 3POB |
| 18 | 2DGC | 31–36 | 3U29 | 89–106 | 1DZI | 180–182 | 3DMW | 217–232 | 5CJB |
| 19–23 | 2V5F | 37–42 | 2CUO | 107–127 | 1PWC | 183–185 | 4Z1R | 233–244 | 7DFR |

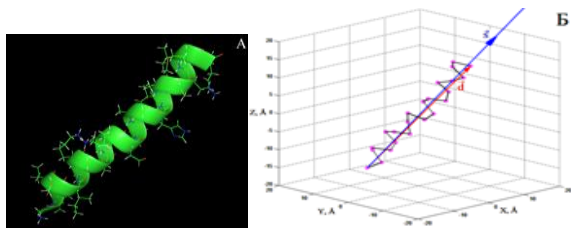


Рис. 4. Изображение правой α -спирали (PDB ID: 5UA7 [11]): А – до обработки; Б – после обработки.

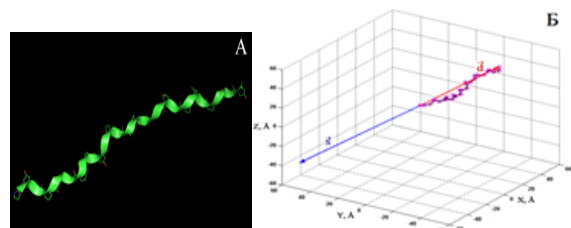


Рис. 5. Изображение левой полипролиновой спирали (сегмент молекулы PDB ID: 1CAG chain A [11]): А – до обработки; Б – после обработки.

Для наглядности на рисунке 3 векторы \mathbf{d} и \mathbf{s} построены из одной точки – атома $C\alpha_1$. Согласно нашему методу, для правых спиралей угол $\angle(\mathbf{d}, \mathbf{s})$ должен быть менее 90° (рис. 3,А), а для левой спирали – более 90° (рис. 3,Б). Исследования показали, что для каждого типа спирали (правой α -спирали и левой полипролиновой) этот угол лежит в своем узком диапазоне. На рисунках 4–5 показаны соответственно правая и левая спирали «до» и «после» обработки. Предлагается представлять данные расчета в виде «карт хиральности», где по горизонтальной оси откладывается угол $\angle(\mathbf{d}, \mathbf{s})$, а по вертикальной оси – отношение $|\mathbf{s}|/|\mathbf{d}|$ (рис. 6, таблица).

На представленной «карте» правые α -спирали располагаются выше, что объясняется плотной упаковкой витков. Левые полипролиновые спирали и бета-цепочки располагаются внизу, т.к. это растянутые и плоские структуры соответственно. Как можно видеть, все правые спирали имеют угол менее 90° , а левые полипролиновые – более 90° , таким образом, наш метод верно определяет направление закрутки. Также надо отметить, что наш численный метод смог выявить неочевидное для визуального анализа: β -цепочки имеют тенденцию к левой закрутке.

Для некоторых вторичных структур угол $\angle(\mathbf{d}, \mathbf{s})$ равен 90° – это следствие наличия у данных структур 3 остатков и, следовательно, 3 атомов α -углерода, между которыми можно построить лишь 2 вектора, для которых можно построить лишь 1 векторное произведение перпендикулярное исходным векторам. Таким образом, для определения направления закрутки спирали необходимо иметь минимум 4 атома, по которым строятся векторы. Это условие ограничения метода.

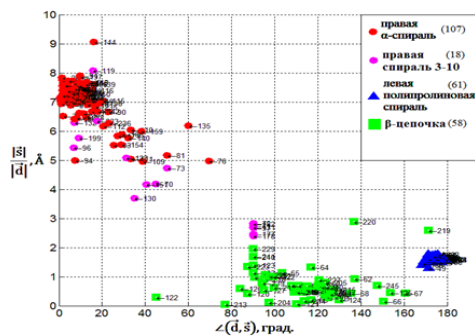


Рис. 6. «Карта хиральности» вторичных структур: α -спиралей – 107, правых спиралей $\langle 3-10 \rangle$ – 18, левых полипролиновых – 61, β -цепочек – 58.

Таким образом, предложен и отработан метод, характеризующий вторичную структуру белков – тип и знак хиральности.

Об автоволновой природе фолдинга

Неоднородные активные среды вследствие наличия порогового эффекта могут гасить автоволновые процессы или изменять тип симметрии. Имеется тенденция трансформации тривиальных структур, формирующихся в однородных активных средах, в хиральные паттерны в неоднородных средах [1]. Подобная эволюция носит необратимый характер. Активная среда с локальными неоднородностями может выступать в роли эффективного однонаправленного хирального фильтра [1–3].

Хиральность – дуалистическое свойство, определяющее не только статическую конструкцию макромолекул, но и траекторию процессов их пространственной самоорганизации – фолдинга. Гомохиральность первичной структуры белков, составленной левыми аминокислотными остатками, придает ей свойства одномерной активной среды и создает предпосылки автоволновой самоорганизации – формирования спиральных, суперспиральных и складчатых регулярных структур [3].

Предложена пространственная математическая модель фолдинга белков как процесса автоволновой самоорганизации в активных средах, позволяющая независимо от длины цепи учитывать знак хиральности структуры и иерархический уровень. В ходе компьютерной симуляции, согласно модели, формируется спиральная структура u , определяемая переменными x , y , с заданным направлением закрутки. Модель представляет собой следующее двумерное автоволновое уравнение:

$$D\Delta u = u \left(u - \frac{z_v(x, y)}{2} \cdot \frac{x^2 + y^2}{x_v^2 + y_v^2} \right) (u - z_v(x, y)), \quad (4)$$

где $\mathbf{v}(x_v, y_v, z_v)$ – радиус-вектор элемента спирали. Каждый второй элемент в витке спирали рассчитывается умножением первого элемента на матрицу C_1 – поворота в плоскости витка на угол $(\pi - \alpha)$. Каждый последний элемент витка рассчитывается посредством преобразования

поворотов и сдвига в начало следующего витка. Автоволновой фронт будет иметь форму спирали, проекция которой на плоскости OXY будет описываться уравнением (рис. 7).

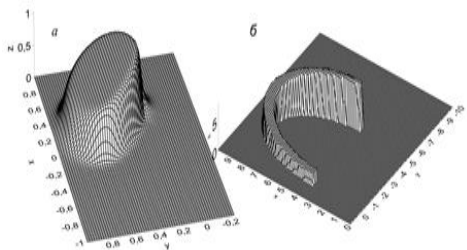


Рис. 7. Результаты математического моделирования автоволнового процесса фолдинга: а) виток из 4 аминокислот, б) спираль из 12 аминокислот.

На рисунке 7,а представлена начальная стадия формирования правой спирали движением влево от условно принятого первого элемента, на рисунке 7,б – правая спираль из 12 элементов (4 витка). Активатором хиральной системы является первичный носитель знака хиральности. Учет знака хиральности заложен в модель и позволяет описывать стратификацию в структурно-функциональной иерархии белковых молекул, что обуславливает процесс фолдинга.

Заключение

Предложен и отработан метод, определяющий вторичную структуру белков – тип и знак хиральности (направление закрутки) и дающий информацию о пространственной структуре, которая позволяет перейти на следующий иерархический уровень: суперспирали (например, левые полипролиновые спирали в коллагене, перевиваясь, образуют правую супер спираль), параллельные и антипараллельные β -слои, третичную структуру белков.

Данная работа – это шаг к главной цели наших исследований: связать термодинамику и симметрию, что есть ключ к пониманию фолдинга.

Список литературы

1. Твердислов В.А., Сидорова А.Э., Яковенко Л.В. От симметрий – к законам эволюции. I. Хиральность как инструмент стратификации активных сред. *Биофизика*. М.: Изд-во Наука, 2012. Т. 57. № 1. С. 146–154.
2. Твердислов В.А. Хиральность как первичный переключатель иерархических уровней в молекулярно-биологических системах. *Биофизика*. М.: Изд-во Наука, 2013. Т. 58. № 1. С. 159–164.
3. Твердислов В.А., Мальшко Е.В., Ильченко С.А. От волновых механизмов самоорганизации к молекулярным машинам.

Известия РАН. Серия физическая. 2015. Т. 79. № 3. С. 1729–1733.

4. Petitjean M. On the root mean square quantitative chirality and quantitative symmetry measures. *Journal of Mathematical Physics*. 1999. V. 40. № 9. P. 4587–4595.
5. Avnir D., Hel-Or H., Mezey P. Symmetry and chirality: Continuous measures. *The Encyclopedia of Computational Chemistry*. 1998. V. 4. P. 2890–2901.
6. Лузанов А.В., Иванов В.В., Миняев Р.М. Индексы дисимметрии на основе оптической вращательной поляризуемости хиральных молекул. *Журнал структурной химии*. 1998. Т. 39. № 2. С. 319–327.
7. Марценюк М.А., Марценюк Н.М., Фуфачев М.А. Тензорные параметры как «меры» киральности молекул. *Вестник Пермского университета (Химия)*. 2011. № 4. С. 61–70.
8. Овчинников Ю.А. *Биоорганическая химия*. М.: Просвещение, 1987. 815 с.
9. Аленицын А.Г., Бутиков Е.И., Кондратьев А.С. *Краткий физико-математический справочник*. М.: Наука, 1990. 368 с.
10. *Worldwide Protein Data Bank*. URL: <http://www wwpdb.org/documentation/file-format> (дата обращения: 03.09.2018).
11. *RCSB Protein Data Bank*. URL: <http://www.rcsb.org/> (дата обращения: 03.09.2018).