

КОЛИЧЕСТВЕННЫЕ КРИТЕРИИ ХИРАЛЬНОСТИ В ИЕРАРХИЯХ БЕЛКОВЫХ СТРУКТУР

© 2019 г. А.Э. Сидорова, Е.В. Малышко, А.Р. Котов,
В.А. Твердислов, М.Н. Устинин*

*Физический факультет Московского государственного университета имени М.В. Ломоносова,
119991, Москва, Ленинские горы, 1/2*

**Институт математических проблем биологии РАН – филиал ИПМ им. М.В. Келдыша РАН,
142290, Пушкино Московской области, ул. Профессора Виткевича, 1*

E-mail: sky314bone@mail.ru

Поступила в редакцию 13.09.18 г.

После доработки 13.09.18 г.

Принята к публикации 19.11.18 г.

На основе теории формирования знакопеременных иерархических структур в макромолекулярных системах разработан количественный подход к оценке степени хиральности отдельных уровней в иерархиях белковых структур. Оценки необходимы для построения физической модели фолдинга белков и их функционирования как молекулярных машин. Достаточным условием метода для охарактеризования уровня в иерархической структуре и знака хиральности белковых блоков является взаимное расположение α -углеродов аминокислот. Количественной оценкой закрученности спиральной (вторичной) и суперспиральной (третичной) структур является абсолютная величина суммы векторных произведений. Знак скалярного произведения вектора направления на вектор суммы векторных произведений указывает на направление закрутки. Получены карты хиральности для вторичной и третичной структур ряда белков. Достоверность карт подтверждается анализом реальных структур.

Ключевые слова: белки, хиральность, энантиомеры, альфа-углерод, правая альфа-спираль, правая 3_{10} спираль, левая полипролиновая спираль, бета-структура, векторное произведение, вектор направления, карта хиральности.

DOI: 10.1134/S0006302919020029

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]. Для белковых структур эта закономерность выражается следующим образом: в белках левые аминокислоты формируют полипептидную цепь, которая образует правую α -спираль; близко расположенные правые α -спирали, перевиваясь, образуют левую суперспираль; на четвертом надмолекулярном иерархическом уровне белок имеет тенденцию к правой укладке [2–4]. Выявлена последовательность смены знака хиральности в структурно-функциональной иерархии белковых структур L–D–L–D. Аналогично отмечается закономерное чередование знака хиральности D–L–D–L при переходе на более высокий уровень структурно-функциональной организации ДНК. Характерен сдвиг по фазе: белковая иерархия «стартует» с L-аминокис-

лот, а нуклеотидная – с D-углевода дезоксирибозы (рис. 1).

Данные иерархические структуры отражают дискретность в макромолекулярных структурах клеток: от первичной асимметричной до четвертичной. При наличии убедительных качественных оценок, разработанных в данной теории, принципиально необходимо получить количественные оценки доли и знака хиральности структур разных иерархических уровней. Однако, несмотря на многочисленные работы в этой области, проблема на настоящий момент по-прежнему не решена. По существу, в этой проблеме есть два взаимодополняющих вопроса, связанных с введением меры хиральности: как количественно оценить (сравнить) степень хиральности молекулярных конструкций и их знака при одном типе симметрии и как количественно соотнести степень хиральности конструкций разного типа симметрии. Решению первого вопроса посвящена настоящая работа.

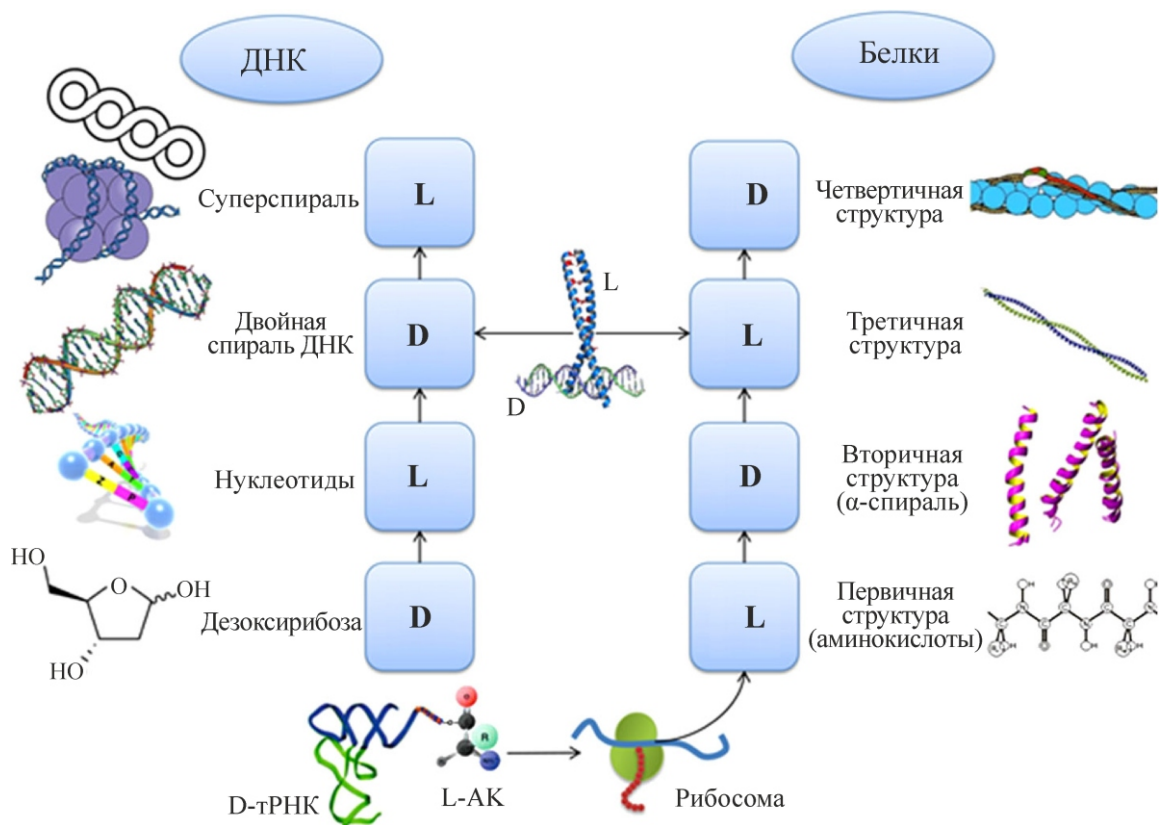


Рис. 1. Знакопеременные хиральные иерархии структурных уровней ДНК и белков: L – левая конфигурация энантиомера или спирали; D – правая конфигурация [2–4].

Обращаясь к истории вопроса, отметим, что один из первых подходов количественной оценки хиральности был основан на оценке асимметричного произведения как непрерывной функции множества точек, абсолютное значение которого инвариантно относительно изометрий [5–7]. Подобные асимметрические произведения были названы «функциями хиральности» [8]. В работах [8,9] была рассмотрена обобщенная функция хиральности для класса тригональных бипирамид, учитывающая некоторую функцию лиганда, расположенного в i -й вершине бипирамиды. Но «функция хиральности» оказалась отличной для различных молекулярных структур, кроме того, анализ сложных молекулярных соединений в рамках этого метода оказался затруднительным, поскольку однозначно было определено правило обхода для разных модельных объектов.

В ряде работ [10–20] хиральность молекулярных объектов рассматривается через «индекс ветвления». Изначально этот индекс был предложен для точек кипения алканов и учитывал количество соседних атомов углерода, с которыми i -й атом углерода образует связь [10]. В дальнейшем были сделаны попытки сопос-

тавления различных вариантов топологического индекса ветвления с физическими свойствами алканов – температурой кипения при нормальном давлении [13], давлением паров при температуре 25°C [15,17], плотностью [18], константой преломления [18,19], критическим давлением [18,20]. Однако результаты работы [12] показали избирательность индекса ветвления относительно определенных физических свойств алканов. Индексы ветвления – топологические индексы, что, несомненно, является достоинством метода. Но метод, к сожалению, применим только для алканов и не позволяет разделять структуры по знаку хиральности.

В работе [21] для описания хиральности бензоидов был предложен двоичный код («1» – «0»). Принцип метода заключается в следующем: кодируются только атомы на периферии; внешние вершины бензольных колец (внутренними вершинами пренебрегают) кодируются как «0» (атом является общим для нескольких колец) или «1» (атомы принадлежат определенному кольцу). Молекула считается симметричной, если двоичный код обхода по часовой и против часовой стрелки совпадает, в противном случае – асимметричной. Поэтому для двух противополо-

ложных энантиомеров двоичный код «по часовой стрелке» одного энантиомера совпадает с кодом «против часовой стрелки» противоположного энантиомера [22]. В работе [23] метод двоичного кода был унифицирован за счет введения понятия «частичной суммы атома» k -атома. Преимущества данного метода: начальная точка обхода не имеет значения, а зеркальные молекулы имеют одинаковые по модулю, но противоположные по знаку значения взвешенной атомной суммы. Однако метод применим только к бензоидам в виде 2D-графа, в то время как предметом исследования является трехмерная структура молекул.

Количественные меры хиральности были пересмотрены в 1992 г. и классифицированы по двум категориям [24] – меры отклонения хирального множества от эталонного ахирального множества и меры подобия между множеством и его зеркальным отображением. Автором работы [25] в качестве мер отклонения (подобия) двух множеств было предложено использовать расстояние Хаусдорфа [26], что позволило с использованием теории нечетких множеств применять метод к дискретным и непрерывным множествам точек, например, для облаков электронной плотности [27]. Но, поскольку расстояние Хаусдорфа – это коэффициент, показывающий симметрию молекулы, поэтому метод мер отклонения (подобия) не позволяет оценить «переключение» знака хиральности для противоположных стереоизомеров.

В работе [28] мера хиральности рассмотрена как перекрытие изначального множества с его зеркальным отображением. В качестве функции перекрытия предлагалось использовать объем, поверхность, массу, заряд, волновые функции, электростатический потенциал и взаимодействия Ван-дер-Ваальса. Данный подход нашел свое отражение в разработке метода непрерывных мер симметрии CSM (Continuous Symmetry Measures) и хиральности CCM (Continuous Symmetry Measures) [29–31]. Метод CSM основан на зеркально-поворотных симметриях [31], где мера симметрии – это сумма квадратов расстояния между исходными точками и точками усредненного отображения. Оценка меры хиральности CCM проводится путем сопоставления исходного множества точек со своим зеркальным отображением, при этом выбор отображения основан на принципе минимального расхождения между исходным множеством и зеркальным. Но, как оказалось, у противоположных энантиомеров величины CCM и CSM имеют только положительный знак. Таким образом, по сути, CCM и CSM – это не мера хиральности, а мера симметричности молекул.

Авторами работы [32] был предложен метод функции диссимметрии, который показал себя достаточно универсальным и эффективным инструментом для количественной оценки хиральности и позволил решать ряд задач анализа связи структуры молекул с различными свойствами молекул, обусловленными хиральностью [32,33]. Суть метода сводилась к оценке степени структурного различия молекулы и ее образа, полученного в результате вращения исходного объекта относительно зеркально-поворотных осей (на основе расположения вершин – атомов). Но выбор ориентации зеркально-поворотных осей решается на основе фундаментальных свойств тензора инерции материального тела, а для объектов, у которых имеется два-три момента инерции, сложно определять ориентацию главных осей инерции. Поэтому для хиральных объектов симметрии высоких порядков не может быть корректно оценена мера хиральности [34]. Для адекватного анализа асимметрии необходимо учитывать не только расположение вершин (атомов), но и наличие ребер (связей). Авторы работы [34] предложили модифицированную модель функции диссимметрии, что позволило оценивать асимметрию структур относительно элементов симметрии высоких порядков. Однако метод дает количественную оценку асимметрии, а не хиральности.

Авторами работы [35] для количественной оценки хиральности молекулы предложено использовать отличие молекулы от ее образа, полученного в результате операций симметрии, образующих данный элемент, а именно, поворотов вокруг оси или зеркальных поворотов относительно осей системы координат [36] согласно данному методу, степень симметрии молекулы по координатной оси X_k пропорциональна отношению объема молекулы (V_1) к объему тела (V_2), состоящего из самой молекулы и ее образа, полученного в результате действия операции симметрии. При этом объемы молекул рассчитывают исходя из объемов атомов с помощью численного интегрирования [37], а общую степень симметрии молекулы – через степени симметрии по трем осям. Таким образом, общая диссимметричность (хиральность), согласно данному методу, есть сумма всех асимметричностей, рассчитанных для окружения каждого атома. Разработанный критерий удовлетворяет фундаментальным требованиям, сформулированным в работе [38]: полученные значения диссимметрии молекул отличны от нуля только для хиральных структур различных групп симметрии, а диссимметрии энантиомеров имеют противоположный знак и равное абсолютное значение. Однако сле-

дует отметить техническую сложность использования метода при работе с большими молекулами, поскольку для каждого атома необходимо рассчитать сумму асимметричностей по каждой из трех осей, которые являются осями инерции лишь для определенной молекулы.

Индекс диссимметрии на основе оптической вращательной поляризуемости хиральных молекул можно отнести к физическим методам. В работах [39–42] был предложен индекс диссимметрии на основе анализа моментов вращательной поляризуемости, которая характеризует оптическую активность, где в качестве средней меры хиральности авторы предлагали использовать норму вращательной поляризуемости (χ). Вычисление меры хиральности фактически сводится к анализу хиродиастиального межмолекулярного взаимодействия диссимметричных систем.

В качестве подхода, в котором используется термодинамическая характеристика, можно привести работу [43], где мера хиральности основана на относительной энтропии Кульбака–Лейблера. Расхождение Кульбака–Лейблера – это неотрицательный функционал, являющийся несимметричной мерой удаленности друг от друга двух вероятностных распределений, определенных на общем пространстве элементарных событий. Подход авторов состоит в том, что вместо плотности распределения вероятности $p(x)$ берется функция формы (r) – электронная плотность, приходящаяся на одну частицу вещества энантиомера (R или S), а в качестве нормировочной функции – функция формы рацемата $RS(r)$. Так как метод основан на анализе электронной плотности, индекс связан с экспериментальными измерениями, такими как круговой дихроизм и оптические измерения плоскости вращения плоскополяризованного света, где рацемат часто используется в качестве эталона. В данной работе проанализированы пять хиральных галометанов, содержащие один асимметричный атом углерода и атомы N, F, Cl, Br, I в качестве заместителей, и сравниваются полученные расчеты оптического вращения с мерой непрерывного измерения хиральности Авнира (CCM – Continuous Chirality Measure). К недостаткам данного метода можно отнести следующие:

индекс рассчитывается через электронную плотность, следовательно, для расчетов больших макромолекул предъявляются высокие требования к вычислительной технике;

подход рассчитан только на работу с R/S-энантиомерами («R/S» – абсолютные конфи-

гурации энантиомеров, определяемые правилами Кана–Ингольда–Прелога) [44]);

метод не позволяет выявлять знак хиральности структур биомacroмолекул; индекс всегда имеет положительное числовое значение.

Метод Раоса [45] является геометрическим расчетом индекса хиральности на основе перекрытия правой и левой спиралей на двумерной плоскости. «Степень хиральности» количественно определяется путем вычисления индекса хиральности на основе перекрытия проекций спиральных лент. Данный метод, основанный на сличении объекта с его зеркальным отображением, имеет ряд характерных недостатков:

создание зеркальной копии и анализ перекрытий двух изображений крайне затруднительно проводить с большими глобулярными белками со множеством перекрытий полипептидных цепей;

при сопоставлении объекта с его зеркальной копией индекс хиральности принимает только положительные значения, следовательно, с его помощью можно оценить симметричность, а не знак хиральности;

метод применим только к плоским структурам, в то время как биомacroмолекулы – это 3D-объекты.

В работах [46,47] была разработана целостная концепция количественной оценки хиральности. Согласно этой теории, индекс хиральности не должен зависеть от способа выбора зеркального отображения и должен быть нечувствительным к любому изометрическому преобразованию распределения, а мера хиральности должна быть непрерывной характеристикой. Согласно данной концепции:

хиральность чувствительна к размерности пространства: треугольник с разными вершинами является хиральным в 2D-пространстве, а в трехмерном пространстве – ахиральным. Мера хиральности должна определяться для пространства, имеющего любую размерность;

хиральность – это не только геометрия. Поэтому хиральный индекс не должен зависеть от выбора способа зеркального отображения и не должен быть чувствительным к любому изометрическому преобразованию распределения;

мера хиральности должна быть непрерывной характеристикой;

любое распределение массы или заряда можно обрабатывать через плотность распределения – дискретного или непрерывного, конечного или бесконечного.

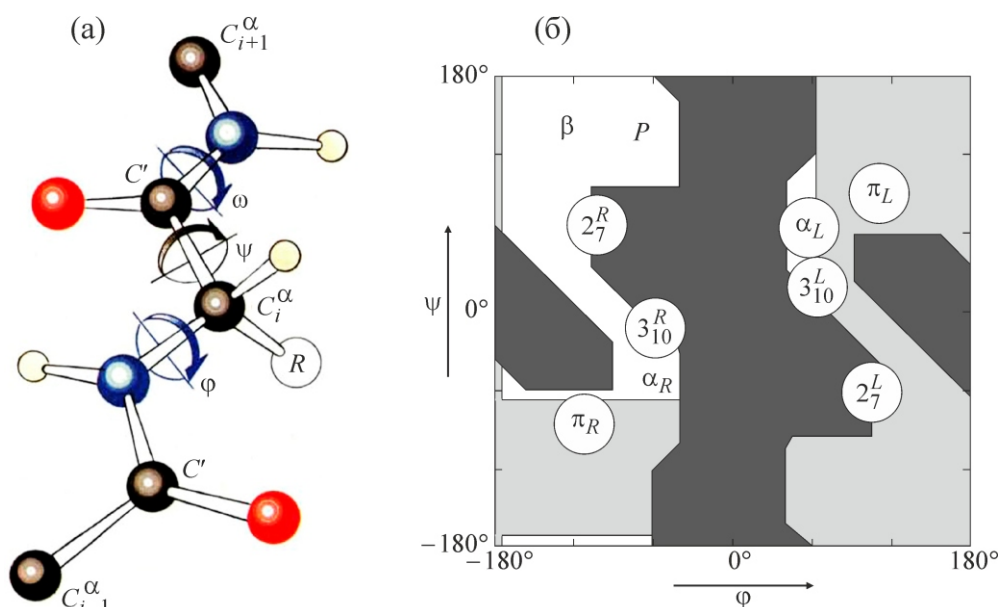


Рис. 2. Карта Рамачандрана: (а) – двугранные углы, характеризующие структуру белка [52], (б) – карта Рамачандрана [51].

Наиболее широко известен метод разрешенных конформаций вторичных структур, основанный на двугранных углах для аминокислотных остатков цепи (в полипептидной цепи) – карты Рамачандрана [48–51]. На этих картах, согласно статистическим данным, конформация около 1% неглициновых остатков считается «запрещенной» [51], а к наиболее перспективным участкам (при поиске аминокислот-кандидатов для замены на остаток глицина) относятся изгибы полипептидной цепи. Но двугранные углы позволяют увидеть только преобладающую конформацию аминокислот в белке [52], а карты, основанные на этих углах, не позволяют определять знак хиральности и не применимы к нуклеиновым кислотам (рис. 2).

МЕТОД ВЕКТОРНЫХ ПРОИЗВЕДЕНИЙ ОЦЕНКИ КОЛИЧЕСТВЕННЫХ КРИТЕРИЕВ ХИРАЛЬНОСТИ В ИЕРАРХИЯХ БЕЛКОВЫХ СТРУКТУР

Разработанный авторами подход количественной оценки хиральных структур имеет общий характер и позволяет достаточно полно характеризовать вторичную и третичную структуры белков: не только их тип, но и знак хиральности с информацией о пространственной структуре [53]. Для определения уровня и знака хиральности блоков или фрагментов вторичной структуры достаточным условием анализа

является взаимное расположение α -углеродов C . Этот подход позволяет на порядок снизить количество обрабатываемой информации (по сравнению с анализом информации обо всех атомах молекулы), что является явным преимуществом при обработке больших массивов данных. Несмотря на то что в PDB-формате существует классификация из 10 различных спиралей [54], наиболее часто встречающимся типом спирали является правая α -спираль. Левая спираль представлена в виде спирали типа poly(Pro)II, в отличие от α -спирали не скрепленная коаксиальными водородными связями (как в коллагене). В нашей работе вторичные спиральные структуры белков рассмотрены как два крайних примера нарушения симметрии на втором уровне структурной иерархии белков.

Предлагаемый подход основан на использовании векторных произведений. Рассмотрим модель витка правой α -спирали: остов из α -углеродов и ось спирали направлена от C-конца к N-концу (рис. 3). Начальные условия: считаем α -углерод (C-конец) первым атомом остова C_1 , а α -углерод (N-конец) – последним атомом остова C_n . Построив векторы между соседними атомами (от предыдущего атома к последующему), получаем векторы $C_1C_2, C_2C_3, \dots, C_{(n-1)}C_n$. Для n атомов получается $(n-1)$ таких векторов, сумма которых – вектор, направленный от первого атома к последнему – «вектор направления» \mathbf{d} :

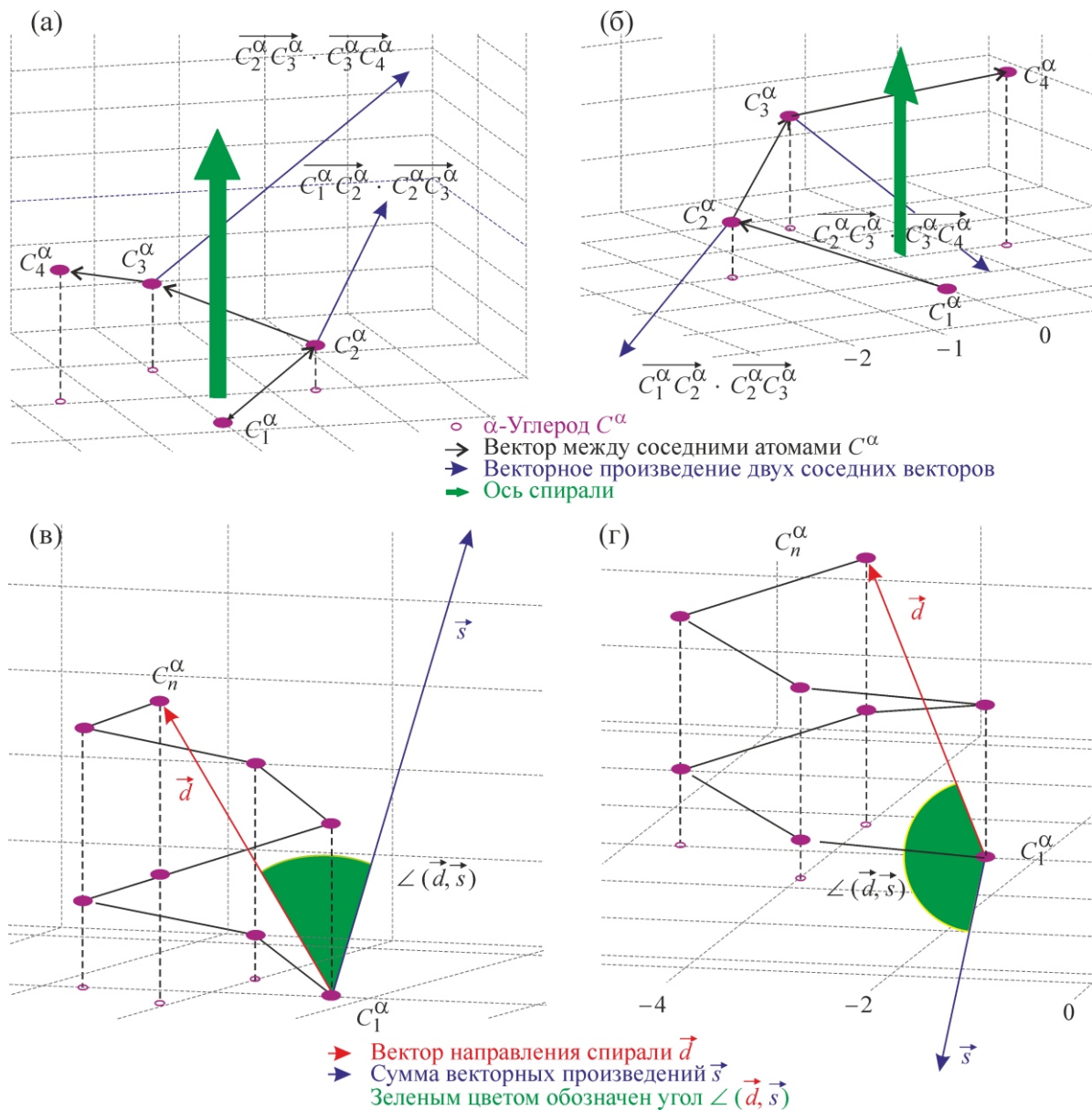


Рис. 3. Модели витков спирали: виток правой спирали (а) и виток левой спирали (б); угол между вектором направления \vec{d} и суммой векторных произведений \vec{s} : для правых спиралей (в) и для левых спиралей (г).

$$C_1 C_2 + C_2 C_3 + \dots + C_{(n-1)} C_n =$$

$$\sum_{i=1}^{(n-1)} C_i C_{(i+1)} = C_1 C_n = \vec{d}. \quad (1)$$

Использование векторного произведения позволяет заложить в метод чувствительность к направлению закрутки спирали: поднимаясь в правой спирали по оси «вверх», векторные произведения направлены также «вверх» относительно направления оси спирали (рис. 3а), а поднимаясь «вверх» в левой спирали – вектор-

ные произведения, по правилу правой руки, будут направлены вниз (рис. 3б).

Для $(n - 1)$ векторов между соседними атомами получаем $(n - 2)$ векторных произведений, сумма которых – вектор \vec{s} :

$$C_1 C_2 + C_2 C_3 + C_3 C_4 + \dots = (2)$$

$$\sum_{i=1}^{(n-2)} C_i C_{(i+1)} + C_{(i+1)} C_{(i+2)} = \vec{s}.$$

Скалярное произведение векторов \vec{d} и \vec{s} позволяет определить косинус угла между ними:

Таблица 1. Количественный состав исследованных молекул белков

PDB	Тип вторичной структуры	Кол-во, шт.	PDB	Тип вторичной структуры	Кол-во, шт.
1L2P	Правые -спирали	1	1BDG	Правые -спирали	15
1QL2	Правые -спирали	3		Правые спирали 3_{10}	2
1TJC	Правые -спирали	10		-Структуры	12
2DGC	Правые -спирали	1	1DZI	Правые -спирали	7
2V5F	Правые -спирали	5		Правые спирали 3_{10}	1
5U9R	Правые -спирали	1		Левые полипролиновые спирали	3
5UA7	Правые -спирали	1		-Структуры	5
5WST	Правые -спирали	2	1PWC	Правые -спирали	12
2YQ8	Правые -спирали	24		Правые спирали 3_{10}	1
1CAG	Левые полипролиновые спирали	3		-Структуры	7
1BKV	Левые полипролиновые спирали	3	1OUT	Правые -спирали	11
3U29	Левые полипролиновые спирали	6		Правые спирали 3_{10}	4
2CUO	Левые полипролиновые спирали	6	2A7U	Правые -спирали	7
2D3H	Левые полипролиновые спирали	6		Правые спирали 3_{10}	1
2KLW	Левые полипролиновые спирали	3	2N8R	Правые -спирали	3
3A1H	Левые полипролиновые спирали	6		Правые спирали 3_{10}	1
4OY5	Левые полипролиновые спирали	1		Левые полипролиновые спирали	3
3DMW	Левые полипролиновые спирали	3		-Структуры	4
4Z1R	Левые полипролиновые спирали	3	3POB	Левые полипролиновые спирали	3
4DMT	Левые полипролиновые спирали	3		-Цепочки	7
3A0A	Левые полипролиновые спирали	6	5CJB	Правые спирали 3_{10}	1
7DFR	Правые -спирали	4		Левые полипролиновые спирали	3
	-Структуры	7		-Структуры	11

$$\cos(\mathbf{d}; \mathbf{s}) = \frac{\mathbf{d} \cdot \mathbf{s}}{|\mathbf{d}| |\mathbf{s}|} \quad (3)$$

Согласно методу, для правых спиралей угол $(\mathbf{d}; \mathbf{s})$ должен быть менее 90° , т.е. $\cos(\mathbf{d}; \mathbf{s}) > 0$ (рис. 3в), а для левой спирали угол – больше 0° , т.е. $\cos(\mathbf{d}, \mathbf{s}) < 0$ (рис. 3г).

Проведенные исследования (табл. 1) подтвердили, что для каждого типа спиралей характерен свой узкий диапазон углов. Для правых -спиралей: $0^\circ < (\mathbf{d}, \mathbf{s}) < 90^\circ$, $0 < \cos(\mathbf{d}, \mathbf{s}) < 1$; для левых полипролиновых спиралей: $90^\circ < (\mathbf{d}, \mathbf{s}) < 180^\circ$, $-1 < \cos(\mathbf{d}, \mathbf{s}) < 0$. На примерах простых моделей показано, что введенные параметры спиральных структур позволяют характеризовать знак хиральности белковых структур (рис. 4).

Ограничение метода: для определения направления закрутки спирали необходимо мини-

мум 4 опорные точки (атома). Связано это с тем, что для реализации метода необходимо построить минимум два векторных произведения \mathbf{c}_i , то есть минимум три исходных вектора между соседними атомами. При наличии только трех точек-атомов A_i получаем два исходных вектора и одно векторное произведение \mathbf{c}_i , и в этом случае угол (\mathbf{d}, \mathbf{s}) всегда будет равен 90° , что не позволяет определить направление закрутки, поскольку любые три точки образуют плоскость, которая в трехмерном пространстве ахиральна.

У распространенных -спиральных вторичных структур в белковых молекулах на один виток приходится целых 3,6–3,7 остатка (3 -углерода) [51]. Следовательно, вторичную закрутку в третичной структуре можно оценить, анализируя каждый первый -углерод во всех витках, т.е. в качестве опорных точек для

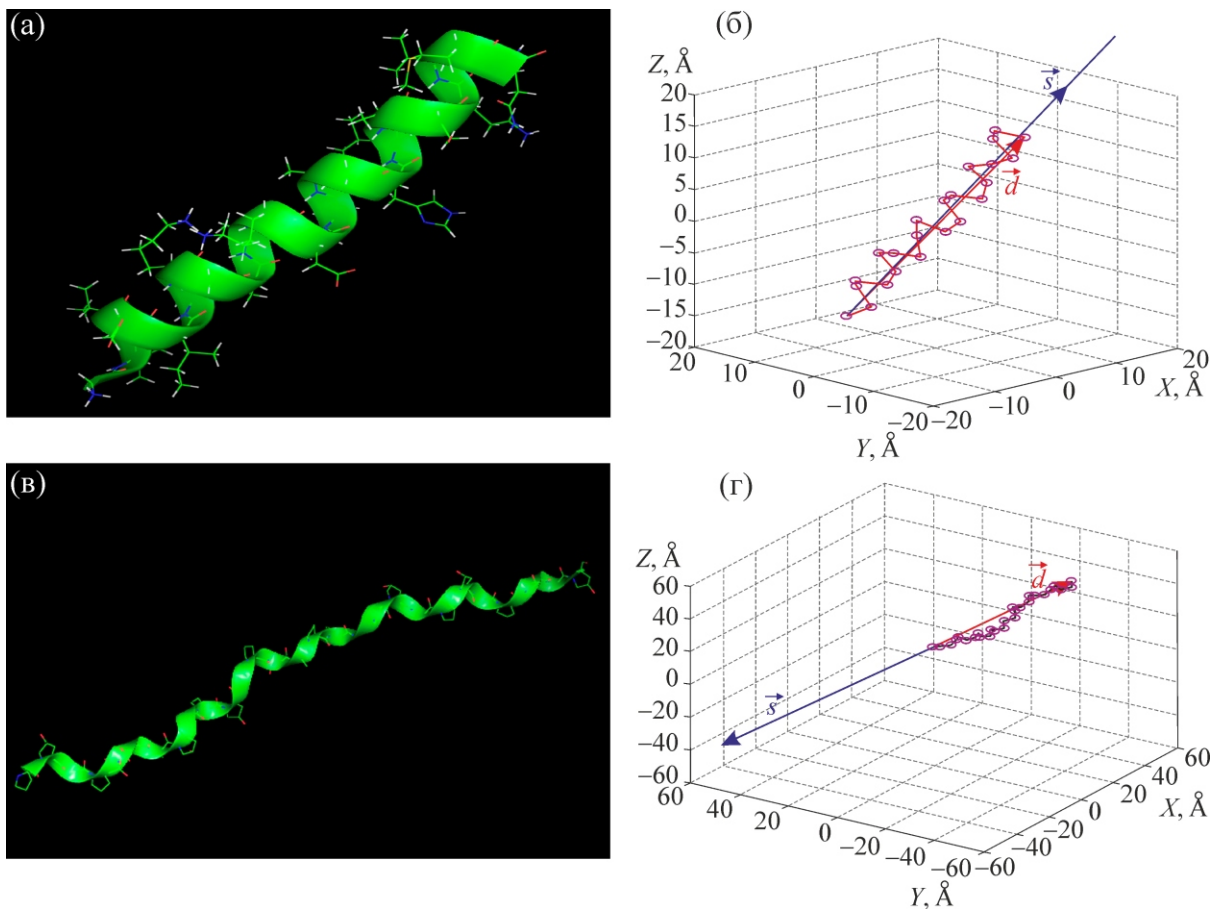


Рис. 4. Правая α -спираль (PDB ID: 5UA7 [54]): (а) – исходное изображение; (б) – изображение после обработки. Левая полипролиновая спираль (сегмент «chain A», PDB ID: 1CAG [54]): (в) – исходное изображение; (г) – изображение после обработки.

оценки вторичной закрутки, начиная с первого атома α -углерода, брать каждый четвертый атом: $C_1, C_4, C_7, \dots, C_{i+3}$ и т.д. Альтернативный подход, который реализован в данной работе, заключается в том, чтобы в качестве опорной точки брать среднюю по трем исходным атомам α -углерода, что также позволяет оценить смещение витков – вторичную закрутку. Так, для α -структур, у которых на один виток приходится лишь два остатка (два α -углерода) [51], в качестве опорных точек следует брать либо каждый третий атом, либо точку, среднюю по двум исходным атомам. На рис. 5 и 6 представлены примеры оценки вторичной закрутки правой α -спирали и левой полипролиновой спирали. Анализ приведенных примеров показывает, что вторичная закрутка третичной структуры имеет противоположный знак: правая α -спираль закручивается левым образом, а левая полипролиновая – правым образом.

МЕРА КОЛИЧЕСТВЕННОЙ ОЦЕНКИ ХИРАЛЬНОСТИ

Абсолютная величина суммы векторных произведений $|\mathbf{s}|$ фактически является количественной оценкой, «мерой» хиральности закрученных спиральных структур. Знак косинуса угла между векторами \mathbf{d} и \mathbf{s} (4) позволяет определить направление закрутки: при $\cos(\mathbf{d}, \mathbf{s}) < 0$ – закрутка левая, при $\cos(\mathbf{d}, \mathbf{s}) > 0$ – закрутка правая. Как следует из формулы (3), знак $\cos(\mathbf{d}, \mathbf{s})$ совпадает со знаком скалярного произведения (\mathbf{d}, \mathbf{s}) . Таким образом, в качестве меры хиральности предлагается использовать величину:

$$\text{total} = \text{sign}[(\mathbf{d}, \mathbf{s})]|\mathbf{s}|. \quad (4)$$

Величина хиральности определяется модулем суммы векторных произведений $|\mathbf{s}|$, а направление закрутки – знаком скалярного произведения вектора суммы векторных произведений \mathbf{s} на вектор направления \mathbf{d} . При увеличении количества правых витков величина total также увеличи-

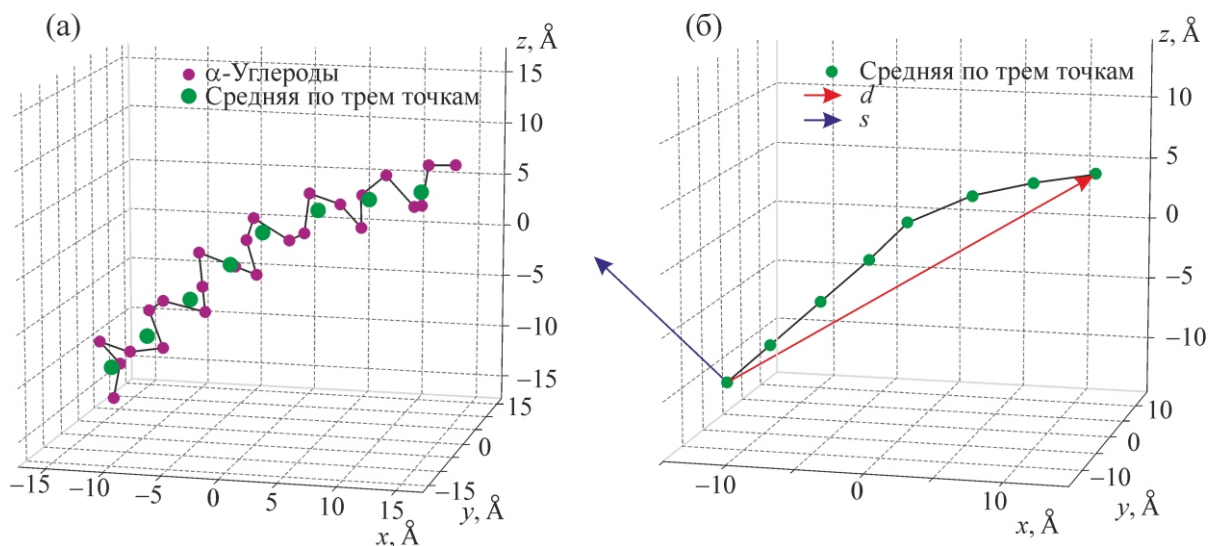


Рис. 5. Вторичная закрутка правой α -спирали (PDB: 5UA8 [54]): (а) – переопределение опорных точек; (б) – результат расчета.

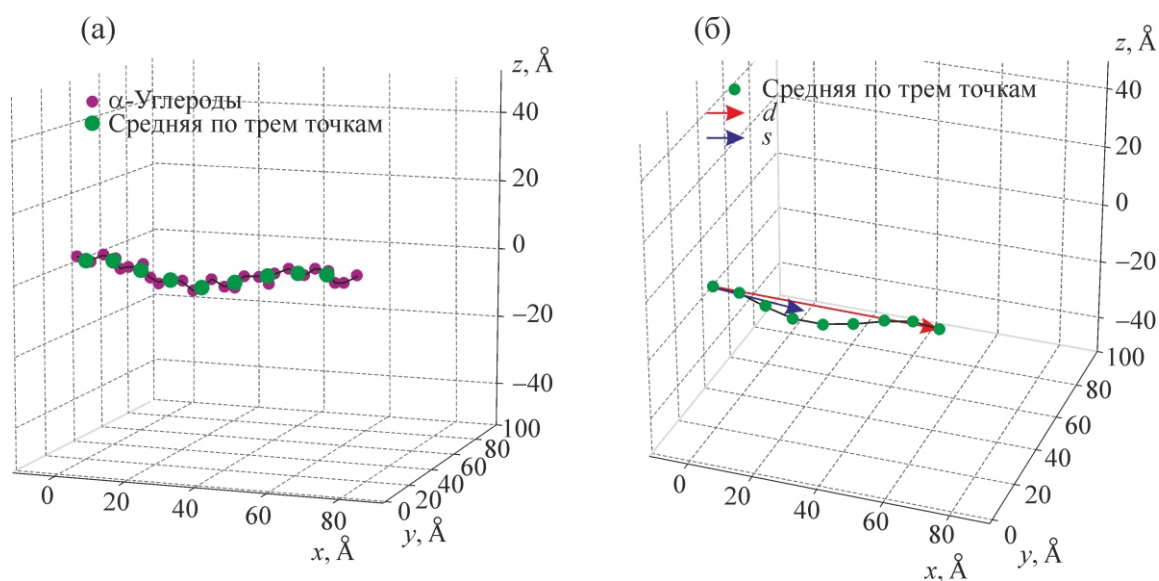


Рис. 6. Вторичная закрутка левой полипролиновой спирали (PDB: 1CAG.chain A [54]): (а) – переопределение опорных точек; (б) – результат расчета.

вается, а при увеличении количества левых витков total уменьшается. При наличии равного количества правых и левых витков $\text{total} = 0$ – это ахиральный объект. Выборка исследованных молекул представлена в табл. 2.

На рис. 7 показана зависимость изменения величины total от типа вторичной структуры и количества опорных α -углеродов (т.е. от длины цепочки). Анализ показывает, что спиральные структуры с правой закруткой располагаются выше условной желтой линии, а структуры с левой закруткой – ниже желтой линии. Прос-

транственную структуру α -структур можно определить как «почти плоские зигзаги» с незначительной левой закруткой (обладают слабо выраженной хиральностью) (рис. 7б). Несколько зигзагообразных полипептидных цепей, в которых водородные связи внутри каждой цепи отсутствуют, образуют β -структуру. В состав β -листов входят 20–28% всех аминокислотных остатков глобулярных белков [55]. Из-за некоторой скрученности отдельных β -тяжей β -листы всегда несколько скручены как целое [56], и направленность водородных связей меняется

Таблица 2. Количественный состав исследованных молекул белков

PDB**	Тип вторичной структуры	Количество, шт.
1BKV*, 1CAG*, 1G9W*, 1ITG*, 1K6F*, 1Q7D*, 1QSU*, 1WZB*, 2CUO*, 2D3F*, 2D3H*, 2DRT*, 2G66*, 2KLW*, 3A08*, 3A0A*, 3A0M*, 3A19*, 3A1H*, 3ADM*, 3AH9*, 3B0S*, 3DMW*, 3IPN*, 3P46*, 3POD*, 3PON*, 3T4F*, 3U29*, 3WN8*, 4DMT*, 4GYX*, 4OY5*, 4Z1R*	Левая полипролиновая спираль	34
1E79	Правая -спираль	97
	Правая спираль 3_{10} -структура	37 94
1BMF	Правая -спираль	87
	Правая спираль 3_{10} -структура	6 78
1HV4	Правая -спираль	56
	Правая спираль 3_{10}	16

по ходу тяжа. Отдельный -тяж имеет левое скручивание. Таким образом, скрученность -листа левая, если смотреть с края этого листа (и правая, если смотреть на поворот линии водородных связей вдоль -тяжей) [51]. Принято смотреть вдоль хода -тяжей, и поэтому считается, что у -листа правопреллерная скрученность [56]. Здесь проявляется общее свойство двоякоизогнутой поверхности с центральной осью симметрии, когда суммарная хиральность объекта отсутствует, и объект – ахирален (рис. 8). Можно полагать, что -лист – единственная внутримолекулярная структура, являющаяся регулярной, но ахиральной. При этом существенно, что такая структура не включается в рассматриваемые нами хиральные иерархии, а принадлежность ее ко вторичным структурам основывается исключительно на участии водородных связей в ее формировании. Эти особенности их пространственной структуры объясняют близкое расположение -структур к условной горизонтальной линии на рис. 7а. Существенно, что часть точек, принадлежащие, согласно данным PDB, к правым спиральям 3_{10} (рис. 7б), на нашей карте хиральности расположены в области левых структур (располагаются ниже условной горизонтальной линии). Мы полагаем, что, согласно [57], в результате отсутствия специфичного идентификатора в PDB-формате для этой структуры, короткие левые спирали 3_{10} были зашифрованы как правые.

Данные по вторичным (-спирали и -листы) и третичным структурам как функция меры хиральности $total$ представлены в виде оригинальных «карт хиральности» на рис. 9.

По горизонтальной оси (рис. 9) отложена величина $norm$:

$$norm = \frac{total}{(n-2)|v|_{mean}^2} \frac{\text{sign}[\cos(\mathbf{d}, \mathbf{s})]|\mathbf{s}|}{(n-2)|v|_{mean}^2}, \quad (5)$$

где нормировка – произведение количества векторных произведений $(n-2)$ на квадрат среднего расстояния между каждыми двумя соседними опорными точками $|v|_{mean}^2$. По вертикальной оси – степень растянутости структуры – отношение $|\mathbf{d}|/L$ (L – длина цепочки). На карте хиральности первичной закрутки (рис. 9а) спиральные структуры D (правая закрутка) располагаются справа от условной желтой линии, а структуры L (левая закрутка) – слева от условной вертикальной линии. Правые -спирали располагаются ниже, что объясняется плотной упаковкой витков. Левые полипролиновые спирали и -цепочки располагаются выше, так как это растянутые и плоские структуры. На графике видно, что для всех правых спиралей $norm > 0$ (так как $\cos(\mathbf{d}, \mathbf{s}) > 0$), а для левых полипролиновых $norm < 0$ (так как $\cos(\mathbf{d}, \mathbf{s}) < 0$). Таким образом, можно констатировать, что метод векторных произведений адекватно отражает направление закрутки. Отмечается соответствие полученных данных картам Рамачандрана: правые -спирали распола-

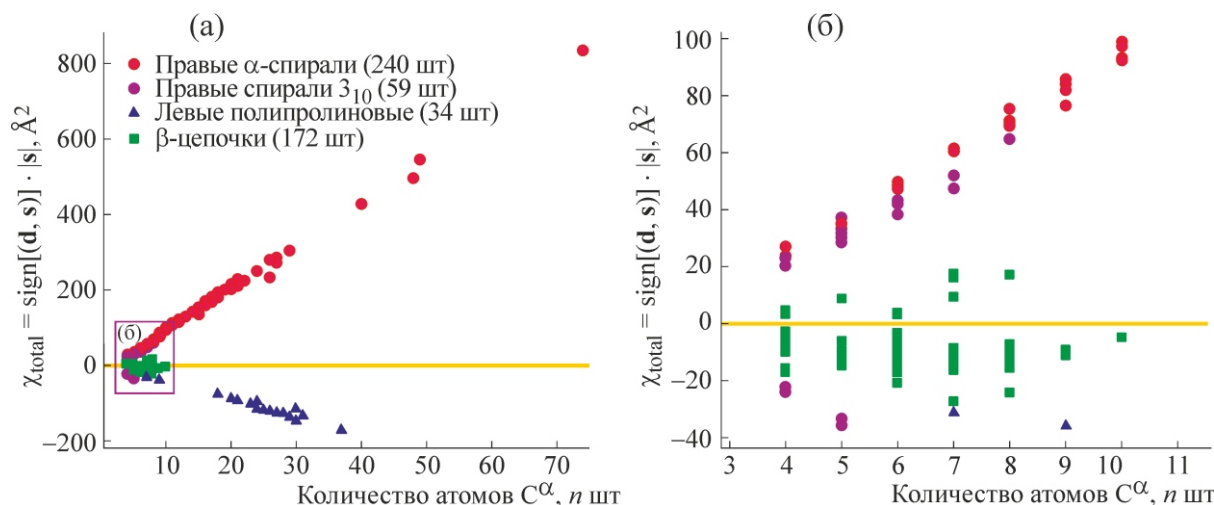


Рис. 7. Зависимость величины χ_{total} от типа вторичной структуры и количества опорных точек: (а) – исходный масштаб; (б) – увеличенный масштаб.

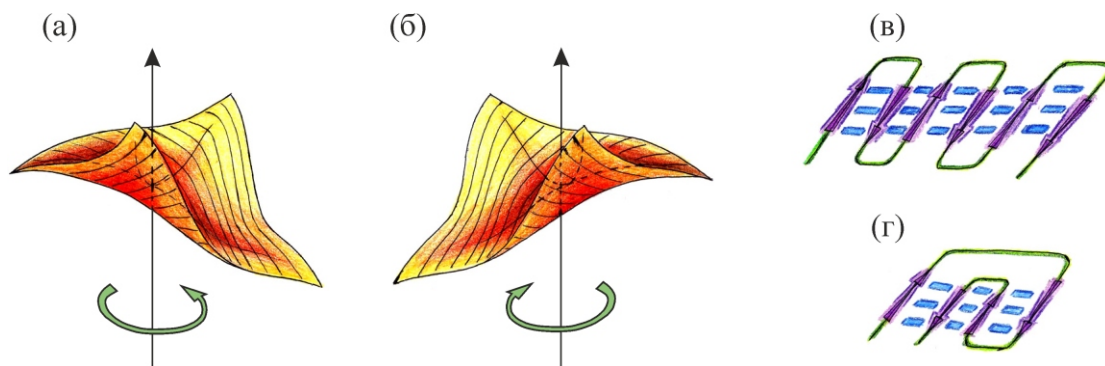


Рис. 8. Двоикоизогнутая поверхность модели β -листа с центральной осью симметрии. При наблюдении с одного торца структура модели характеризуется левой хиральностью (а), при наблюдении с другого торца – правой хиральностью (б). Типы β -структур: (в) – «меандр», (г) – «греческий ключ».

гаются рядом с правыми спиралями 3_{10} , а левые полипролиновые – рядом с β -цепочками.

На рис. 9б представлена карта хиральности третичных структур, характеризующихся вторичной закруткой (количество точек на этом графике уменьшилось, по сравнению с графиком 7а, поскольку отсеялись короткие структуры). Согласно методу для оценки вторичной закрутки требуются более длинные структуры, что дает возможность в качестве опорных точек для β -спиралей брать среднюю по трем точкам (для β -структур – среднюю по двум точкам). Если для определения первичной закрутки β -спиралей необходимо минимальное количество атомов – четыре атома С (условие метода), то для определения вторичной закрутки спиральных структур минимум составляет $4 - 3 = 12$ атомов С ($4 - 2 = 8$ атомов С для β -структур).

Из правых β -спиралей вторичная правая закрутка отмечена для 16 молекул из исследованных 147. Также только одна из 12 длинных β -структур не меняет знак хиральности (имеет вторичную левую закрутку). Таким образом, сопоставление карт хиральности для вторичной (рис. 7а) и третичной структур (рис. 7б) выявляет смену направления закрутки (знака хиральности) при переходе на следующий иерархический уровень, что полностью укладывается в теорию формирования иерархичных структур белков [2–4].

Разработанный метод позволяет выявлять качественные нестыковки в классификации вторичных структур в PDB-файлах. Небольшая часть D-спиралей 3_{10} попала в область левых структур (слева от условной разграничительной линии) – это L-спирали 3_{10} , зашифрованные в PDB-файле как правые. Кроме того, метод смог

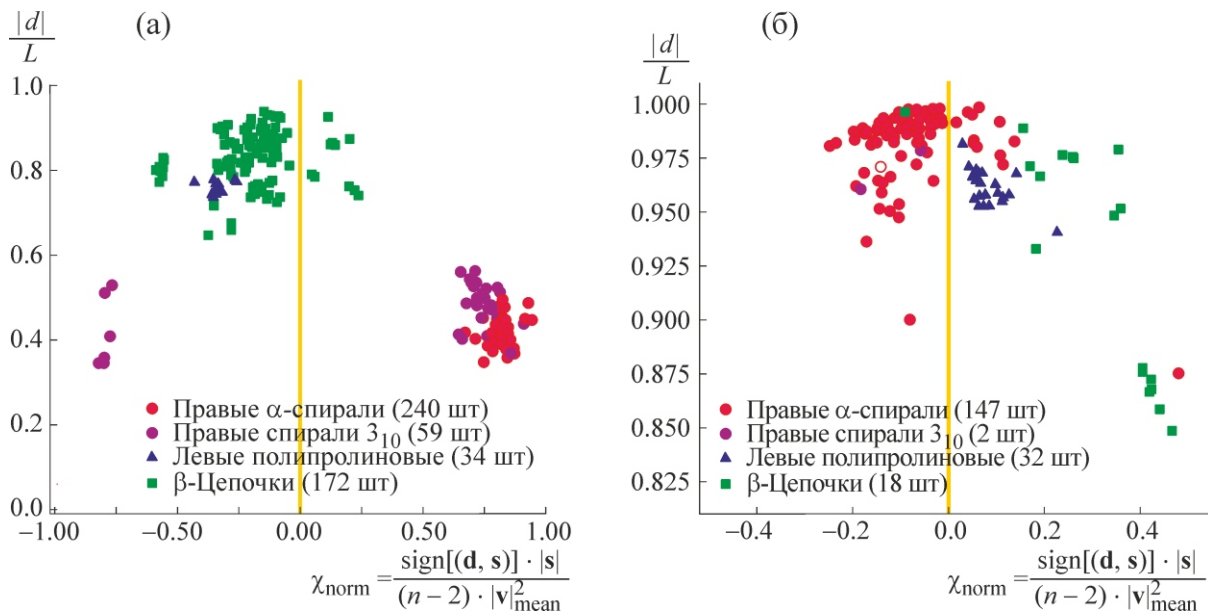


Рис. 9. Карты хиральности вторичных (а) и третичной (б) структур.

выявить неочевидную при визуальном анализе тенденцию к левой закрутке β -структур (рис. 7а), о чем было сообщено еще в 1982 г. [58], и количественно ее оценить.

ЗАКЛЮЧЕНИЕ

Предложен и разработан новый метод, характеризующий вторичную и третичную структуры белков (знак и меру хиральности фрагмента цепи) и дающий информацию о пространственной структуре белков и ДНК. Установлено, что знак косинуса угла между векторами \mathbf{d} и \mathbf{s} (4) позволяет определить направление закрутки: при $\cos(\mathbf{d}, \mathbf{s}) < 0$ – закрутка левая, при $\cos(\mathbf{d}, \mathbf{s}) > 0$ – закрутка правая. Получена количественная мера хиральности спиральных структур – абсолютная величина суммы векторных произведений $|\mathbf{s}|$. Знак косинуса угла между векторами \mathbf{d} и \mathbf{s} позволяет определить направление закрутки: при $\cos(\mathbf{d}, \mathbf{s}) < 0$ – закрутка левая, при $\cos(\mathbf{d}, \mathbf{s}) > 0$ – закрутка правая. На основании этой меры хиральности представлены адекватные реальным структурам карты хиральности вторичной и третичной структур, анализ которых демонстрирует смену направления закрутки (знака хиральности) при переходе на следующий иерархический уровень от спиральных к суперспиральным структурам. Метод наиболее эффективен для количественного сравнения знака и степени хиральности структур одного типа симметрии. Разрабо-

танный метод – геометрическое решение данной биофизической проблемы – может служить существенным дополнением к известным подходам к моделированию процессов фолдинга макромолекул.

СПИСОК ЛИТЕРАТУРЫ

1. Б. Албертс, Д. Брей, Дж. Льюиси др., *Молекулярная биология клетки*, в 3-х т., 2-е изд., перераб. и доп. (Мир, М., 1994), т. 1.
2. V. A. Tverdislov, A. E. Sidorova, and L. V. Iakovenko, *Biophysics* **57** (1), 146 (2012).
3. В. А. Твердислов, *Биофизика* **58** (1), 159 (2013).
4. V. A. Tverdislov, E. V. Malyshko, S. A. Il'chenko, et al., *Biophysics* **62** (3), 331 (2017).
5. P. A. Guye, *Compt. Rendus* **110**, 714 (1890).
6. P. A. Guye, *Compt. Rendus* **116**, 1451 (1893).
7. P. A. Guye and L. Chavanne, *Compt. Rendus* **116**, 1454 (1893).
8. M. Petitjean, *Entropy* **5** (3), 271 (2003).
9. E. Ruch and A. Schönhofer, *Theor. Chim. Acta* **10** (2), 91 (1968).
10. M. Randić, *J. Am. Chem. Soc.* **97** (23), 6609 (1975).
11. M. Randić, *Chemometrics and Intelligent Lab. Systems* **10**, 213 (1991).
12. X. L. Peng, K.-T. Fang, Q.-N. Hu, Y.-Z. Liang, *Molecules* **9** (12), 1089 (2004).
13. G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci.* **39** (5), 788 (1999).
14. Y. Du, Y. Liang, B. Li, and Ch. Xu, *J. Chem. Inf. Comput. Sci.* **42** (5), 993 (2002).

15. D. Yaffe and Y. Cohen, *J. Chem. Inf. Comput. Sci.* **41** (2), 463 (2001).
16. E. S. Goll and P. C. Jurs, *J. Chem. Inf. Comput. Sci.* **39** (6), 1081 (1999).
17. H. E. McClelland and P. C. Jurs, *J. Chem. Inf. Comput. Sci.* **40** (4), 967 (2000).
18. J. A. Dean, *Lange's Handbook of chemistry* (McGraw-Hill, New York, 1998).
19. A. R. Katritzky, S. Sild, and M. Karelson, *J. Chem. Inf. Comput. Sci.* **38** (5), 840 (1998).
20. B. E. Turner, C. L. Costello, and P. C. Jurs, *J. Chem. Inf. Comput. Sci.* **38** (4), 639 (1998).
21. M. Randić and M. Razinger, *J. Chem. Inf. Comput. Sci.* **36** (3), 429 (1996).
22. T. Zhao, Q. Zhang, H. Long, and L. Xu, *PLoS One* **9** (7) (2014). DOI: 10.1371/journal.pone.0102043.
23. M. Randić, *J. Chem. Inf. Comput. Sci.* **41** (3), 639 (2001).
24. A. B. Buda, T. A. der Heyde, and K. Mislow, *Angewandte Chemie Int. Ed.* **31** (8), 989 (1992).
25. A. Rassat, *Comptes-rendus des séances de l'Académie des sciences. Ser. 2, Mécanique-physique, chimie, sciences de l'univers, sciences de la terre* **299** (2), 53 (1984).
26. Ф. Хаусдорф, *Теория множеств* (ОНТИ, М.–Л., 1937).
27. P. G. Mezey, *J. Mol. Struct.: THEOCHEM* **455** (2–3), 183 (1998).
28. G. Gilat and L. S. Schulman, *Chem. Phys. Lett.* **121** (1–2), 13 (1985).
29. H. Zabrodsky, S. Peleg, and D. Avnir, *J. Am. Chem. Soc.* **114** (20), 7843 (1992).
30. H. Zabrodsky, S. Peleg, and D. Avnir, *J. Am. Chem. Soc.* **115** (18), 8278 (1993).
31. S. Peleg and D. Avnir, *J. Am. Chem. Soc.* **115** (18), 8278 (1993).
32. M. Pinsky, C. Dryzun, D. Casanova, et al., *J. Comput. Chem.* **29** (16), 2712 (2008). DOI: 10.1002/jcc.20990.
33. V. E. Kuz'min, I. B. Stel'makh, M. B. Bekker, and D. V. Pozigun, *J. Phys. Org. Chem.* **5** (6), 295 (1992).
34. L. A. Kutulya, V. E. Kuz'min, I. B. Stel'mach, et al., *J. Phys. Org. Chem.* **5** (6), 308 (1992).
35. С. Э. Алиханиди и В. Е. Кузьмин, *Журн. структур. химии* **41** (4), 795 (2000).
36. В. М. Марков, В. А. Потемкин и А. В. Белик, *Журн. структур. химии* **42** (1), 91 (2001).
37. П. М. Зоркий и Н. Н. Афолина, *Симметрия молекул и кристаллов* (Изд-во МГУ, М., 1979).
38. А. В. Белик и В. А. Потемкин, *Журн. физ. химии* **66** (1), 140 (1992).
39. P. W. Fowler, *Symmetry: Culture and Science* **16** (4), 321 (2005).
40. А. В. Лузанов, В. В. Иванов и Р. М. Миняев, *Журн. структур. химии* **39** (2), 319 (1998).
40. A. V. Luzanov and D. Nerukh, *Functional Materials* **12** (1), 55 (2005).
41. A. V. Luzanov and D. Nerukh, *J. Math. Chem.* **41** (4), 417 (2007).
42. A. V. Luzanov *Functional Materials* **22** (3), 355 (2015).
43. S. Janssens, P. Bultinck, A. Borgoo, et al., *J. Phys. Chem. A* **114** (1), 640 (2010). DOI: 10.1021/jp9081883.
44. Ш. Бакстон и С. Робертс, *Введение в стереохимию органических соединений* (Мир, М., 2015).
45. G. Raos, *Macromolec. Theory and Simulations* **11** (7), 739 (2002).
46. M. Petitjean, *The Mathematical Theory of Chirality*. URL: <http://petitjeanmichel.free.fr/itoweb.petitjean.html> (дата обращения: 10.04.2018).
47. M. Petitjean, *J. Math. Phys.* **43** (8), 4147 (2002).
48. G. N. Ramachandran, C. Ramakrishnan, and V. Sasi-sekharan, *J. Mol. Biol.* **7** (1), 95 (1963).
49. Ч. Кантор и П. Шиммель, *Биофизическая химия* (Мир, М., 1984), т. 2.
50. D. A. Brant and P. R. Schimmel, *Proc. Natl. Acad. Sci. USA* **58**, 428 (1967).
51. А. В. Финкельштейн и О. Б. Птицын, *Физика белка* (Изд-во КДУ, М., 2012).
52. Ю. А. Овчинников, *Биоорганическая химия* (Просвещение, М., 1987).
53. А. Р. Котов, А. Э. Сидорова, В. А. Твердислов и М. Н. Устинин, *Ученые записки физич. ф-та МГУ* **1830701** (3), 1 (2018).
54. The Protein Data Bank. URL: <http://www.rcsb.org> (дата обращения: 10.04.2018).
55. W. Kabsch and C. Sander, *Biopolymers* **22** (12), 2577 (1983).
56. C. Chothia, *J. Mol. Biol.* **75**, 295 (1973).
57. *Документация формата PDB-файла*. URL: <http://www.wwpdb.org/documentation/file-format>.
58. Г. Шульц и Р. Ширмер, *Принципы структурной организации белков* (Мир, М., 1982).

Quantitative Criteria of Chirality in Hierarchical Structure of Proteins

A.E. Sidorova*, E.V. Malyshko*, A.R. Kotov*, V.A. Tverdislov*, and M.N. Ustinin**

*Faculty of Physics, Lomonosov Moscow State University, Leninskiye Gory 1/2, Moscow, 119991 Russia

**Institute of Mathematical Problems of Biology – Branch of Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, ul. Professora Vitkevicha 1, Pushchino, Moscow Region, 142290 Russia

Based on the theory of formation of sign alternating hierarchical structures in macromolecular systems, a quantitative approach has been developed to assess the chirality sign of individual levels in hierarchical structures of protein. The quantitative estimates are needed to model protein folding used as molecular machines. Mutual attraction between α -carbon atoms of amino acids is a sufficient condition for the method in order to characterize the level in the hierarchical structure and determine the chirality sign of protein blocks. A quantitative estimate of the twist of the helical (secondary) and superhelical (tertiary) structures is the absolute value of the sum of cross products. The sign of the scalar product of the direction vector to the vector of the sum of vector products indicates the direction of the twist. Chiral maps for the secondary and tertiary structures of a number of proteins have been obtained. The reliability of the maps is confirmed by the analysis of real structures.

Keywords: proteins, chirality, enantiomers, alpha-carbon atom, right-handed alpha helix, right-handed 3_{10} helix, left-handed polyproline helix, beta-chain, cross product, direction vector, chiral map