

МЕТОДЫ ОПРЕДЕЛЕНИЯ ЗНАКА ХИРАЛЬНОСТИ СПИРАЛЬНЫХ И СУПЕРСПИРАЛЬНЫХ СТРУКТУР БЕЛКОВ

© 2021 г. А.Э. Сидорова, А.О. Луценко, Д.К. Шпигун, Е.В. Малышко, В.А. Твердислов

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

E-mail: sky314bone@mail.ru

Поступила в редакцию 29.12.2020 г.

После доработки 04.01.2021 г.

Принята к публикации 11.01.2021 г.

Предложен и разработан метод описания спиральных структур разного уровня организации в белках. Метод позволяет охарактеризовать знак хиральности/спиральности и выполнить количественную оценку для α -спиралей, 3_{10} -спиралей и coiled-coil-суперспиралей, что существенно для развития разработанной ранее концепции о знакопеременных иерархиях хиральных структур в белках и нуклеиновых кислотах. Достаточным условием метода является взаимное расположение α -углеродов в спирали, что позволяет на порядок снизить количество обрабатываемой информации и является значительным преимуществом при обработке больших массивов данных.

Ключевые слова: белки, хиральность, спирали, суперспирали, карта хиральности.

DOI: 10.31857/S0006302921030017

Белковые полипептидные цепи сложены исключительно остатками L-аминокислот [1]. В ходе их укладки образуются, в частности, спиральные и суперспиральные структуры. Наиболее распространенной и энергетически выгодной регулярной белковой вторичной структурой является правая α -спираль [2, 3]. α -Спирали возникают лишь на тех участках первичной структуры, где имеются аминокислоты, способные последовательно образовать водородные связи с каждым четвертым аминокислотным остатком.

Другой тип спиральных структур, спирали 3_{10} , составляет 15–20% всех белковых спиралей [4]. Характерная длина спирали 3_{10} — три-пять аминокислотных остатков [5, 6], и полный виток такой спирали образован ровно тремя аминокислотными остатками. Часто данный тип спиралей локализуется на концах регулярных α -спиралей. Существуют и другие типы спиральных вторичных структур, но они встречаются гораздо реже.

Образование несколькими α -спиралями не тривиальных конструкций заинтересовало исследователей в результате исследований Л. Полинга и Ф. Крика [7, 8]: Полинг разработал модель α -спирали, а Крик — модель «coiled coil», суперспирали из α -спиралей. Согласно концепции Крика [8–10], для данного типа суперспирали характерен особый тип упаковки боковых цепей — «knobs into holes». При такой упаковке аминокислотный остаток одной спирали («knob») уклады-

вается в пространство между четырьмя остатками другой спирали («hole»). Последовательности, входящие в состав суперспиралей, составляют до 10% протеома организма [11, 12]. Самой простой конструкцией суперспирали является димер, при этом ориентация α -спиралей может быть как параллельной, так и антипараллельной. Встречаются суперспирали, состоящие из семи и более α -спиралей [13]. Для суперспиралей из двух α -спиралей шаг суперспирали составляет около 140 Å, для суперспиралей из трех-четырех α -спиралей — около 200 Å [14]. Каноническая последовательность суперспирали состоит из гептад — повторяющихся семи аминокислотных остатков [15]. Позиции остатков в гептаде принято обозначать как $a-b-c-d-e-f-g$, где, как правило, в позициях a и d находятся гидрофобные остатки, формирующие гидрофобное ядро, а в позициях e и g — полярные аминокислотные остатки и имеют место ионные взаимодействия [16]. Чередуемость гидрофобных остатков попеременно с тремя и четырьмя остатками дает сдвиг периодичности: идеальная правая α -спираль имеет периодичность около 3.65 остатка на оборот, а левая суперспираль с гептадным повтором — до 3.5 остатка на оборот относительно оси пучка [15].

В иерархичных структурах белков переходы с нижнего на высший уровень сопровождаются чередованием знака хиральности структурных уровней L-D-L-D для белков и D-L-D-L для

ДНК, что связано не только с изменением пространственных характеристик, но и с изменением механизмов их формирования, физических свойств и молекулярно-биологические функций [17, 18]. В развитии данной концепции существенным представляется получение количественных оценок хиральности белковых структур разных иерархических уровней. Решению вопроса оценки хиральности спиральных и суперспиральных структур белков посвящена настоящая работа.

МЕТОД ОЦЕНКИ ЗНАКА ХИРАЛЬНОСТИ СПИРАЛЬНЫХ СТРУКТУР БЕЛКОВ

Плоский характер и фиксированная длина пептидных связей определяют практически единственную возможность для сворачивания пептидной цепи путем поворота плоскостей пептидных связей относительно их связей с α -углеродными атомами C_α . Такое вращение было описано двугранными углами, значения которых лежат в основе оценки хиральности белков методом Рамачандрана, позволяющим увидеть преобладающую конформацию элементов вторичной структуры биомолекул [19]. Однако оказалось, что возможно существование конформаций полипептидной цепи и вне разрешенных областей карты Рамачандрана. (Подробный анализ различных методов оценки хиральности приведен в работе [20].)

В ранее представленном нами методе оценки степени хиральности в иерархических структурах белков [20, 21] для определения знака хиральности вторичной структуры достаточным условием анализа является только взаимное расположение C_α . Для оценки знака хиральности было предложено использовать значение косинуса угла между вектором направления (суммой векторов между соседними атомами C_α) и вектором произведения (суммой всех векторных произведений). Этот метод был нами усовершенствован по следующим причинам. Во-первых, в качестве меры хиральности бралась длина вектора произведений, что приводило к неправильной оценке хиральности. Например, для двух последовательных правых спиралей, соединенных разными способами, получаются кардинально разные результаты. Во-вторых, для оценки знака хиральности использовался знак косинуса угла между вектором направления и вектором произведения. Это могло привести к неверному определению знака хиральности.

В качестве экспериментальных объектов были использованы структуры белков, представленные в международной базе данных PDB [22].

Сущность предлагаемого метода оценки хиральности спиральных структур белков рассмот-

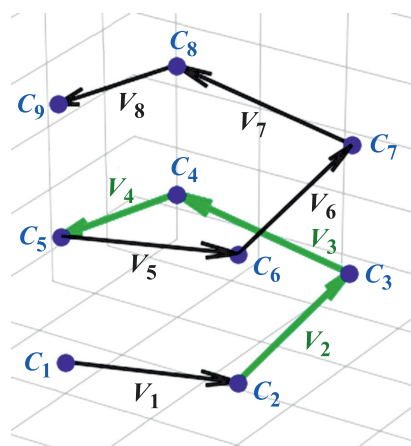


Рис. 1. Модельное изображение полипептидной цепи правой спирали.

рим на модели полипептидной цепи, состоящей из n аминокислотных остатков и имеющей соответственно n атомов C_α в качестве опорных точек. Между каждыми двумя соседними опорными точками строится вектор v_i . Для n опорных точек получаем $(n-1)$ векторов (рис. 1).

Далее для каждых трех последовательных векторов рассчитываем их смешанное произведение (с учетом их координат):

$$([v_1, v_2], v_3) = (v_1 z_2 - v_2 z_1) x_3 + (z_1 x_2 - z_2 x_1) y_3 + (x_1 y_2 - x_2 y_1) z_3. \quad (1)$$

Сумма всех смешанных произведений позволяет оценивать знак хиральности

$$\chi_{\text{total}} = \sum_{i=1}^{n-3} ([v_i, v_{i+1}], v_{i+2}). \quad (2)$$

Результат смешанного произведения — число, знак которого определяет взаимное расположение векторов: для правой тройки векторов знак положительный, для левой — отрицательный. Смешанные произведения не зависят друг от друга, любое изменение положения каждого атома влияет на несколько смешанных произведений (от одного до четырех, в зависимости от порядкового номера атома в цепи).

Исследована возможность нормировки хиральности спиральных белковых структур с соблюдением следующего принципа: величина хиральности должна стремиться к некоторому значению при увеличении числа опорных точек, расположенных на том же объекте (т.е. при увеличении плотности точек). Поэтому каждое смешанное произведение нормировалось на определенную степень к средней длине используемых в нем векторов:

$$C_i = \left(\frac{1}{3} \sum_{j=0}^2 |v_{i+j}| \right)^k, \quad (3)$$

где C_i — нормировочный множитель.

Подбор степени k осуществлялся исходя из следующего: поскольку основными объектами исследования являются белковые структуры с четко определенным расположением атомов углерода, в качестве эталона была взята цилиндрическая винтовая линия с равномерно расположенными точками. Исследование показало, что для сохранения конечной ненулевой хиральности при неограниченном увеличении числа точек следует взять степень, равную пяти.

Окончательное выражение для нормированной хиральности имеет вид:

$$\chi_{\text{norm}} = \sum_{i=1}^{n-3} \frac{([v_i, v_{i+1}], v_{i+2})}{C_i}. \quad (4)$$

Метод реализован в виде компьютерной программы на языке Python 3.7.

Посредством разработанного метода получены карты хиральности для белков-ферментов, представленных в базе данных PDB (рис. 2).

Анализ представленных данных свидетельствует о том, что, во-первых, α -спирали и 3_{10} -спирали представлены в основном правыми структурами (левые α -спирали составляют 0.06% от общего количества α -спиралей, а левые спирали 3_{10} — 4.38% от общего количества спиралей 3_{10}), что адекватно отражает данные, представленные в научной литературе. Нормированное значение хиральности линейно зависит от числа атомов в спирали.

МЕТОД ОЦЕНКИ ЗНАКА ХИРАЛЬНОСТИ СУПЕРСПИРАЛЬНЫХ СТРУКТУР БЕЛКОВ

Со времени открытия структур типа «coiled coil» предпринимались попытки их математического описания. Согласно работе [23], шаг суперспирали является ключевым фактором, определяющим внутри- и межмолекулярные взаимодействия. В соответствии с экспериментальными данными для суперспиральных структур, состоящих из двух спиралей, среднее значение длины шага составляет около 140 Å, что согласуется с наблюдаемым количеством остатков на виток в α -спиралях глобулярных белков. В работе [24] была предложена математическая модель определения знака хиральности суперспиральной структуры с учетом взаимосвязи между шагом и радиусом. В работе [23] изучается хиральность структур типа «coiled coil» и предлагается определять хиральность (τ) «поверхностной кривой» отдельных α -спиралей. Под поверхностной кривой автором понимается некоторая линия, находящаяся

в середине между кривой, соединяющей гептадные остатки в двух позициях. Хиральность рассчитывается в зависимости от периодичности спирали (число остатков в повторе), угла между соседними остатками и сдвигом вдоль оси в расчете на аминокислоту. Знак хиральности определяется по знаку τ — для правых структур τ положительно, для левых — отрицательно. При последующем расчете поверхностная кривая получилась левой. Следует отметить, что в представленных методах не учитывалась возможность того, что суперспираль может быть правозакрученной. Таким образом, данные подходы не дают адекватной оценки знака хиральности суперспиральных структур.

В предлагаемом нами методе для оценки знака хиральности суперспиральных белковых структур типа «coiled coil» в качестве начальных условий используется информация о взаимном расположении α -углеродов аминокислотных остатков цепи. В основе метода лежит определение направления закрутки каждой отдельной α -спирали относительно оси всей суперспирали. В качестве критерия для определения этого направления выбран угол между осью суперспирали и осями образующих ее спиралей. Ориентация этого угла относительно оси суперспирали (по часовой или против часовой стрелки) служит показателем для определения знака хиральности суперспирали. Для определения ориентации угла используется свойство векторного произведения образовывать правые тройки векторов.

В качестве примера рассмотрен белок 1BB1 [https://www.rcsb.org/structure/1BB1], состоящий из трех α -спиралей (на рис. 3а показано его отображение в базе данных PDB). Поскольку в α -спирали на один виток в среднем приходится 3.6 аминокислотных остатка, в качестве условного центра витка спирали взят геометрический центр четырех последовательных атомов (для первого витка — точки C_1, C_2, C_3, C_4 и т.д.) — точки осей спиралей $A_1, A_2, \dots, A_n$, через которые проводится ось α -спирали (рис. 3б). Далее строятся «средняя спираль», где геометрический центр первых атомов каждой α -спирали ($C_{1,1}, C_{2,1}, C_{3,1}$) является первой точкой «средней спирали» и т.д., и ось «средней спирали» (рис. 3в). В общем случае как ось средней спирали, так и оси α -спиралей являются ломаными линиями в трехмерном пространстве. Но для данного метода достаточно информации о взаимных направлениях осей только в узком поперечном сечении суперспирали (два полных витка α -спиралей), на таком интервале оси можно рассматривать как прямые линии, что позволяет непосредственно строить для них векторы направления. Поскольку метод оценивает лишь знак хиральности, такое приближение вполне уместно.

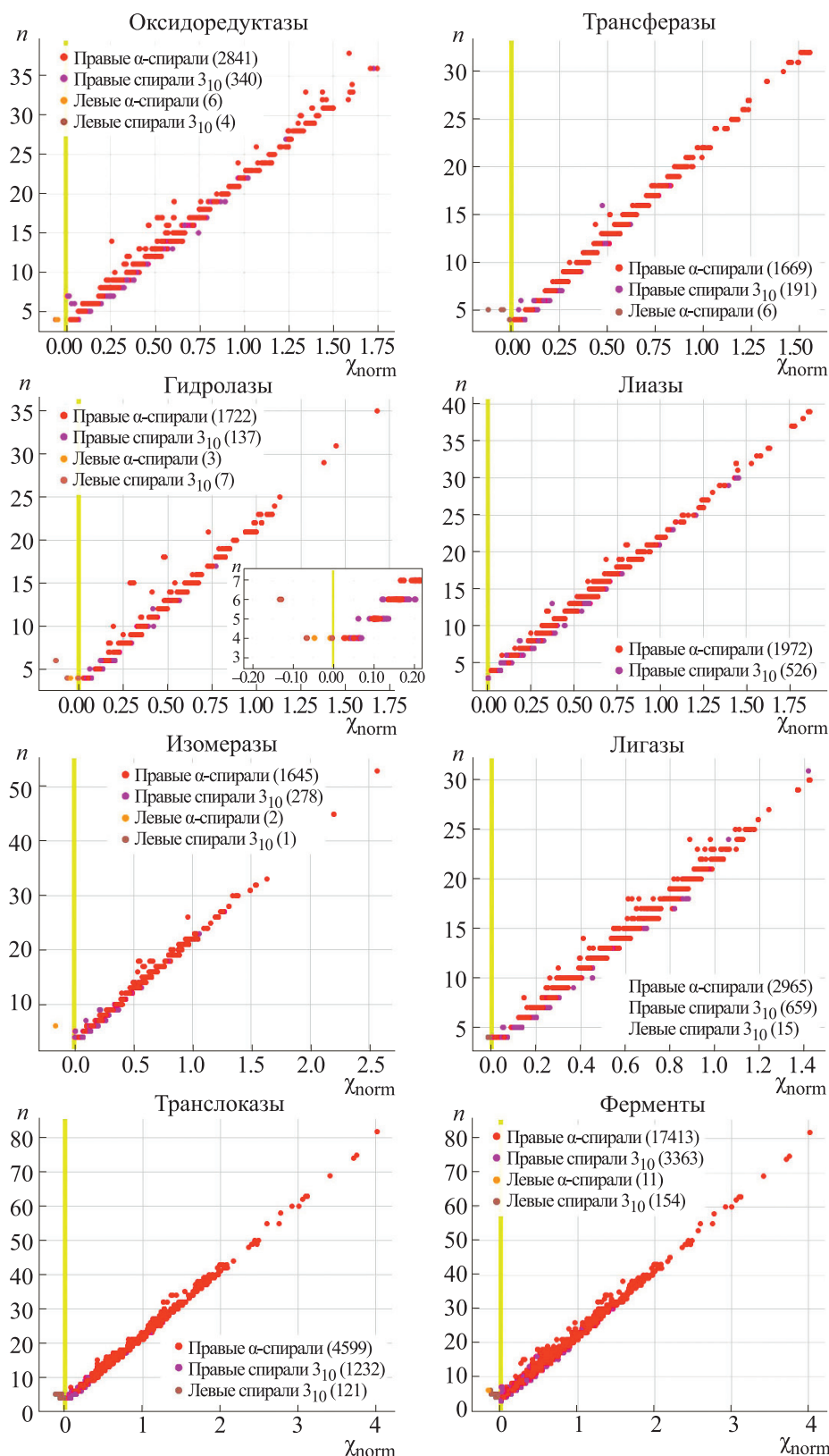


Рис. 2. Карты хиральности белков-ферментов. По горизонтальной оси — нормированное значение хиральности, по вертикальной оси — длина вторичной структуры в аминокислотных остатках.

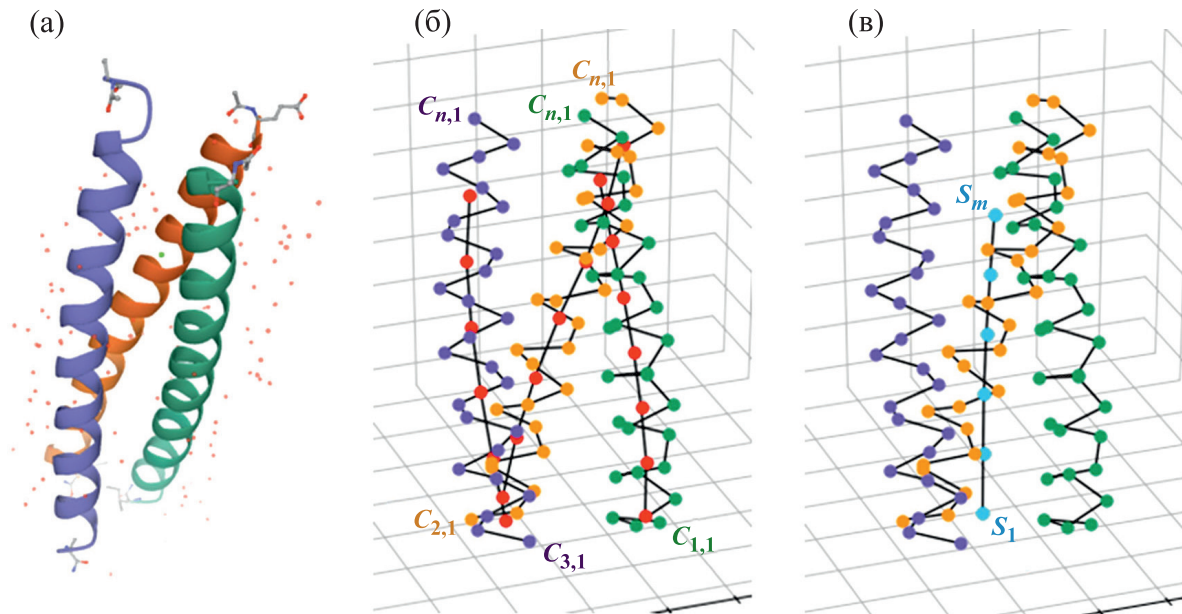


Рис. 3. Схематическое изображение работы метода по определению знака хиральности структуры типа «coiled coil» на примере белка 1BB1.

Определение знака хиральности суперспирали сводится к определению направления угла между осью суперспирали и осью составляющей спирали. Геометрический анализ суперспиралей показывает, что для левой суперспирали составляющие спирали будут отклонены в правую сторону относительно оси суперспирали и угол между векторами $\mathbf{v}_s$ и $\mathbf{v}_{ci}$ будет отсчитываться против часовой стрелки (рис. 4а). Воспользуемся свойством векторного произведения образовывать правые тройки векторов. Обозначим вектор от оси суперспирали к оси i -й спирали через $\mathbf{v}_{ki}$, а векторное произведение векторов $\mathbf{v}_{ci}$ и $\mathbf{v}_s$ — через $\mathbf{v}_{pi}$. Если вектор $\mathbf{v}_{ci}$ отклонен от вектора $\mathbf{v}_s$ вправо,

т.е. по часовой стрелке (суперспираль в этом случае является левой), то векторы $\mathbf{v}_{ki}$ и $\mathbf{v}_{pi}$ будут противонаправлены (рис. 4б). Если вектор $\mathbf{v}_{ci}$ отклонен от вектора $\mathbf{v}_s$ влево (суперспираль правая), то векторы $\mathbf{v}_{ki}$ и $\mathbf{v}_{pi}$ будут сонаправлены. Величина угла между $\mathbf{v}_{ki}$ и $\mathbf{v}_{pi}$ определяется с помощью скалярного произведения, которое для тупых углов отрицательно, для острых — положительно.

Оценка знака хиральности суперспирали рассчитывается с помощью усреднения значения косинуса соответствующего угла для всех спиралей, образующих суперспираль:

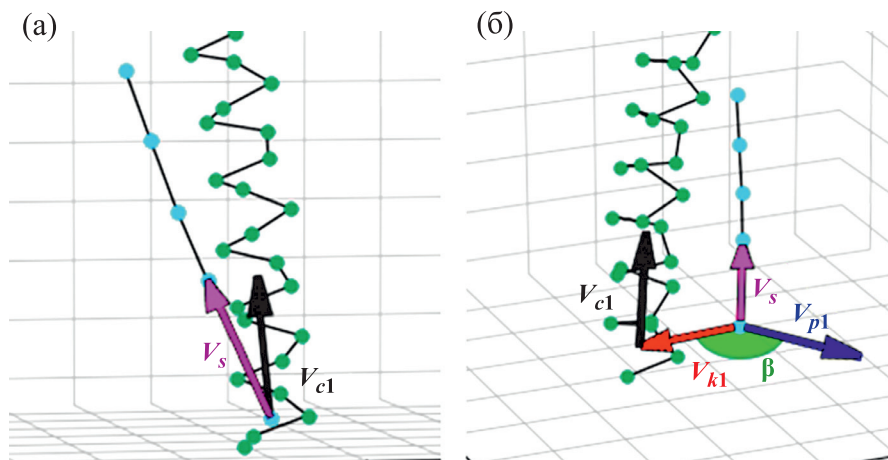


Рис. 4. Определение знака хиральности структуры типа «coiled coil» по углу между векторами на примере белка 1BB1.

Расчетные параметры структур типа «coiled coil»

PDB ID белка	Количество спиралей	Цепи	Номера структур	Средний косинус угла	Средний угол	Направление закрутки	Происхождение
1BB1	3	A,B,C	1,2,3	−0,86544	151,61627	L	Искусственный
1COI	3	A	1,2,3	−0,95287	164,27641	L	Искусственный
1D7M	2	A,B	1,2	−0,89340	153,30395	L	Природный
1L8D	2	A	1,2	−0,83381	146,49241	L	Природный
1L8D	2	B	3,4	−0,79023	142,20691	L	Природный
1NYH	2	A	1,2	−0,70155	134,55180	L	Природный
1P9I	2	A	1,2	−0,79532	142,68591	L	Неизвестно
1RH4	4	A	1,2,3,4	0,97138	12,60585	D	Искусственный
1T6F	2	A,B	1,2	−0,94825	161,48665	L	Природный
1WT6	3	A,B,D	1,2,3	−0,82796	146,74953	L	Природный
2HY6	7	A,B,C,D,E,F,G	1,2,3,4,5,6,7	−0,93137	160,34975	L	Природный
2V7I	2	A,B	1,2	−0,50522	120,34623	L	Природный
2XU6	2	A,B	1,2	−0,92664	157,91675	L	Природный
3QH9	2	A	1,2	−0,99355	173,48756	L	Природный
4BRY	2	A,B	1,2	−0,89581	153,61220	L	Природный
4BWD	2	A,B	1,2	−0,97126	166,23026	L	Природный
4FXX	2	A	1,2	−0,80737	143,83989	L	Природный
4JBZ	3	A,B,C	19,37,55	−0,93823	162,33884	L	Природный
4YTO	2	A,B	1,4	−0,91952	156,85593	L	Природный
5AI6	4	A	1,2,3,4	−0,96225	166,96110	L	Природный
5D3A	2	A,B	1,2	−0,85570	148,83745	L	Природный
5JXC	3	A,B,C	1,2,3	−0,78597	142,28240	L	Природный
5JXC	3	D,E,F	4,5,6	−0,82593	147,97427	L	Природный
5LXN	2	A,B	1,2	−0,77531	140,83288	L	Природный
5LXO	2	A,B	1,2	−0,69884	134,33381	L	Природный
5M48	3	A	1,2,3	−0,87375	151,17191	L	Природный
5NFD	2	A	1,2	−0,95346	162,45090	L	Природный
5T1Y	2	A	1,2	−0,95321	162,40315	L	Природный
5UXT	3	A,B,C	1,2,3	−0,92353	160,90101	L	Неизвестно
1FB1	5	A,B,C,D,E	9,17,26,34,42	−0,32922	110,45772	L	Природный
1H12	6	A	4,7,10,12,14,17	−0,70685	141,34958	L	Природный

$$\chi_{\text{norm}} = \frac{1}{q} \sum_{i=1}^q \cos(\beta_i), \quad (5)$$

где β_i — угол между направлением оси i -й α -спирали и направлением оси суперспирали, т.е. между векторами $\mathbf{v}_{ki}$ и $\mathbf{v}_{pi}$; q — число спиралей в суперспирали.

Для расчетов использовали координаты атомов из файлов PDB. Метод опробован на 30 суперспиралях, из которых 29 были идентифицированы как левые, а суперспираль из искусственного белка 1RH4 определилась как правая. Данные расчетных параметров оценки хиральности структур типа «coiled coil» представлены в таблице. Анализ 30 различных суперспиралей показал,

что такие структуры являются в абсолютном большинстве левыми, что полностью согласуется с известными из научной литературы данными, а также с концепцией хиральной иерархии белковых структур [17, 18].

Представленные нами методы оценки спиральных и суперспиральных структур отличаются небольшим количеством необходимой для его использования информации, а также простотой вычислений. На основе методов реализованы компьютерные программы на языке Python 3.7. Графический интерфейс реализован с помощью библиотеки tkinter. Среднее время расчета структуры — 50 мс.

ЗАКЛЮЧЕНИЕ

Представлены методы идентификации знака хиральности спиральных структур белков и суперспиралей «coiled coil». Достаточным условием методов является взаимное расположение α -углеродов в спирали, что позволяет на порядок снизить количество обрабатываемой информации и является явным преимуществом при обработке больших массивов данных. На основе методов реализованы компьютерные программы на языке Python 3.7.

Анализ представленных карт хиральности α - и 3_{10} -спиралей свидетельствует о том, что α -спирали и 3_{10} -спирали представлены в основном правыми структурами, а суперспирали «coiled coil» — левыми. Полученные результаты соответствуют теории знакопеременных хиральных структур и адекватно отражает данные, представленные в научной литературе. Нормированное значение хиральности линейно зависит от числа атомов в спирали.

ФИНАНСИРОВАНИЕ РАБОТЫ

Исследование выполнено при поддержке Междисциплинарной научно-образовательной школы Московского университета «Фундаментальные и прикладные исследования космоса» и частичной финансовой поддержке Российского научного фонда (проект № 19-74-00082).

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая работа не содержит описания исследований с использованием людей и животных в качестве объектов.

СПИСОК ЛИТЕРАТУРЫ

1. А. В. Финкельштейн и О. Б. Птицын, *Физика белка: курс лекций с цветными и стереоскопическими иллюстрациями и задачами*, 3е изд. (КДУ, М., 2012).
2. L. Pauling and R. B. Corey, Proc. Natl. Acad. Sci. USA **37**, 256 (1951).
3. L. Pauling and R. B. Corey, Proc. Natl. Acad. Sci. USA **37**, 729 (1951).
4. D. J. Barlow and J. M. Thornton, J. Mol. Biol. **201**, 601 (1988).
5. M. N. Fodje and S. Al-Karadaghi, Protein Eng. **15**, 353 (2002).
6. J. S. Richardson and D. C. Richardson, Science **240**, 1648 (1988).
7. L. Pauling, R. B. Corey, and H. R. Branson, Proc. Natl. Acad. Sci. USA **37**, 205 (1951).
8. F. H. C. Crick, Nature **170**, 882 (1952).
9. F. H. C. Crick, Acta Crystallogr. **6**, 685 (1953a).
10. F. H. C. Crick, Acta Crystallogr. **6**, 689 (1953b).
11. A. Rose, S. J. Schraegle, E. A. Stahlberg, and I. Meier, BMC Evol Biol. **5**, 66 (2005).
12. O. J. Rackham, M. Madera, C. T. Armstrong, et al., J. Mol. Biol. **403**, 480 (2010).
13. E. Moutevelis and D. N. Woolfson, J. Mol. Biol. **385**, 726 (2009).
14. J. Seo and C. Cohen, Proteins **15**, 223 (1993).
15. A. Lupas and M. Gruber, Adv. Prot. Chem. **70**, 37 (2005).
16. Y. B. Yu. Adv. Drug Deliv. Rev. **54** (8), 1113 (2002).
17. В. А. Твердислов, Биофизика **58** (1), 159 (2013).
18. В. А. Твердислов и Е. В. Малышко, Успехи физ. наук **189**, 375 (2019).
19. G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan, J. Mol Biol. **7**, 95 (1963).
20. A. E. Sidorova, E. V. Malysheko, A. R. Kotov, et al., Biophysics **64** (2), 155 (2019).
21. A. E. Sidorova, N. T. Levashova, E. V. Malysheko, and V. A. Tverdislov, Moscow Univ. Phys. Bull. **74** (3), 213 (2019).
22. RCSB PDB. Available online: <http://www.rcsb.org> (accessed on 15 December 2020).
23. G. N. Phillips, Proteins **14**, 425 (1992).
24. R. D. B. Fraser and T. P. MacRae, *Conformation in Fibrous Proteins and Related Synthetic Polypeptides* (Academic Press, Lond., 1973).

Methods for Determining the Chirality Sign of Helical and Superhelical Protein Structures

A.E. Sidorova, A.O. Lutsenko, D.K. Shpigun, E.V. Malyshko, and V.A. Tverdislov

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

A method for describing helical structures of different levels of organization in proteins has been developed. The proposed method can be used to characterize the sign of chirality/helicity and quantify α -helices, 3_{10} -helices and coiled coil structures; it is important for the development of the previously developed concept of sign-alternating chiral hierarchies of DNA and protein structures. A sufficient condition for this method is the relative location of α -carbons in the helix. Satisfying this condition ensures that the amount of processed information can be reduced by an order of magnitude thereby offering an evident benefit for processing large data sets.

Keywords: proteins, chirality, helices, superhelices, chirality map