Математическая биология и биоинформатика
2017. Т. 12. № 2. С. 398–410. doi: 10.17537/2017.12.398

========== **БИОИНФОРМАТИКА** ==========

УДК: 519

Анализ торсионных углов между осями спиралей в спиральных парах белковых молекул

Тихонов Д.А.[*1,3], Куликова Л.И.[1], Ефимов А.В.[2]**

[1]*Институт математических проблем биологии РАН – филиал ИПМ им. М.В. Келдыша РАН, Пущино, Московская область, Россия*
[2]*Институт белка РАН, Пущино, Московская область, Россия*
[3]*Институт теоретической и экспериментальной биофизики РАН, Пущино, Московская область, Россия*

Аннотация. В данной работе проведен анализ распределения торсионных углов между осями спиралей в парах связанных между собой перетяжками спиралей в пространственных структурах белковых молекул. Исследование проводилось на множестве спиральных пар, отобранных из структур белковых молекул, представленных в PDB. Полученное множество спиральных пар было проанализировано и разбито на три подмножества по критерию пересечения проекций спиралей на параллельные плоскости, проходящие через оси спиралей. Показано, что распределение всех типов спиральных пар, не имеющих пересечений проекций спиралей, в зависимости от торсионного угла охватывает практически весь диапазон углов с двумя пиками в области 0° и 180°. Большинство пар указанного подмножества составляют спиральные пары, состоящие из α- и 3_{10}-спиралей. Спиральные пары, образованные двумя α-спиралями, представляют значительную часть подмножества с пересечением проекций спиралей, и распределение таких структур имеет максимум в области значений торсионного угла между осями спиралей в области 20°–25°. Образованные двумя α-спиралями спиральные пары составляют абсолютное большинство пар подмножества с пересечением проекций и осей спиралей. Для них характерно распределение с тремя максимумами, лежащими в области острых углов: в области отрицательных значений (от –50° до –25°), в области положительных значений углов (от 20° до 25°) и в области прямого угла (от 70° до 110°).

Ключевые слова: *структурные мотивы, точечная модель, спиральные пары в белковых молекулах, торсионный угол между осями спиралей.*

ВВЕДЕНИЕ

Распознаванию и отбору в структурах белковых молекул, зарегистрированных в банке белковых структур PDB [1] всевозможных структурных мотивов, состоящих из двух рядом расположенных по цепи спиралей, соединенных одним или более аминокислотными остатками, конформация которых не является спиральной, и всестороннему исследованию отобранных спиральных пар посвящена данная и опубликованные ранее [2, 3] работы.

В опубликованных работах [2, 3] был описан разработанный нами подход отбора спиральных пар в структурах белковых молекул. В указанных работах изучались

*dmitry.tikhonov@gmail.com
**likulikova@mail.ru

структуры, в образовании которых участвуют две спирали любого типа: α-спирали, спирали 3_{10} и π-спирали. Было показано, что 72.16 % от общего числа всех спиралей приходится на α-спирали; 27.73 % составляют спирали 3_{10}; оставшиеся 0.1 % – π-спирали. Несмотря на то, что α-спирали составляют подавляющее большинство (72.16 %), спиральные пары, состоящие из двух α-спиралей, составляют только 54.7 % к общему числу спиральных пар. Спиральные пары с участием спиралей 3_{10} составляют 45 %, меньше 0.2 % приходится на пары с участием π-спиралей. Также было показано, что большинство спиральных пар (45 %), не имеющих пересечений проекций спиралей, образованы α-спиралью и спиралью 3_{10}. В множестве спиральных пар, проекции спиралей которых имеют пересечение, большинство (59 %) составляют пары, образованные двумя α-спиралями. В множестве пар, включающего структуры, имеющие пересечение проекций и осей спиралей, спиральные пары, образованные двумя α-спиралями, составляют подавляющее большинство (87.5 %). При этом общее число спиральных пар, не имеющих пересечений проекций спиралей равно числу спиральных пар, проекции спиралей которых имеют пересечение, каждое из них дает вклад по 44 % к общему числу спиральных пар, остальные 12 % приходятся на структуры, имеющие пересечение проекций спиралей и осей спиралей. Это означает, что большая часть спиральных пар в аминокислотных цепочках (56 %) имеют пересечения проекций спиралей.

Указанные работы были посвящены исследованию межспиральных расстояний [2], углов между осями спиралей и количества аминокислот между спиралями [3] в спиральных парах, отобранных согласно разработанным правилам отбора спиральных пар в структурах белковых молекул из Банка белковых структур. Было показано, что распределение всех типов спиральных пар, не имеющих пересечения проекций спиралей, охватывает практически весь диапазон углов с одним максимумом в области прямого угла. Образованные двумя α-спиралями спиральные пары составляют абсолютное большинство пар множества с пересечением проекций и осей спиралей. При этом значительная часть указанных пар имеет острый угол $(20° \le \varphi \le 50°)$ между осями спиралей. Межплоскостное расстояние для всех этих структур равно 10 Å. Проведен анализ распределения всех типов спиральных пар, принадлежащих различным множествам, в зависимости от длины перетяжки. Показано, что во всех множествах больше всего исследуемых структур с короткой (от одной до семи аминокислот) перетяжкой, а также что среди всех отобранных по сформулированным критериям структур, состоящих из двух α-спиралей и имеющих пересечение проекций спиралей и осей спиралей, максимальное число структур с длиной перетяжки, равной пяти аминокислотам.

Данная работа посвящена исследованию торсионных углов Ω между осями спиралей в спиральных парах белковых молекул, при этом анализируются межспиральные взаимодействия только тех спиралей, которые расположены в полипептидной цепи друг за другом и соединены между собой перетяжкой, которая может иметь различную длину и конформацию. В работах других исследователей [4–8] изучаются взаимодействия спиралей в белковых молекулах без учета того факта, на сколько аминокислот отстоят друг от друга спирали, и какую вторичную структуру они образуют. Исследование проводилось на множестве спиральных пар из структур белковых молекул, зарегистрированных в PDB, без учета классификации белков. Как правило, такие классификации основаны на особенностях вторичной и третичной структуры белков [9]. В дальнейшем планируется исследование распределения характеристик спиральных пар в каждом отдельно взятом классе белков.

Актуальность создания базы данных указанных структур и их дальнейшее изучение следует из интереса исследователей к структурным мотивам, имеющим уникальные укладки полипептидной цепи в пространстве [10]. Структурные мотивы, образованные

двумя α-спиралями, расположенными в полипептидной цепи друг за другом и соединенными между собой перетяжками описаны в работах [11–13]. Они представляют собой компактные пространственные структуры. Также из литературы известно, что наиболее плотная упаковка двух α-спиралей достигается при антипараллельной, перпендикулярной и, так называемой, скошенной ориентации между спиралями. Примерами таких упаковок являются супервторичные структуры: α-α-уголки, α-α-шпильки, L-образные и V-образные структуры [12], однако всеобъемлющий анализ этих структурных мотивов во всех белках, представленных в PDB, до сих пор не проводился.

ОБЪЕКТ ИССЛЕДОВАНИЯ

Из банка белковых данных PDB были отобраны все структурные мотивы, образованные двумя спиралями любого типа, расположенными в полипептидной цепи друг за другом и соединенными между собой перетяжками различной длины, имеющими разные конформации. Общее число найденных спиральных пар всех типов равно 2206605 [2, 3]. При этом среди всех отобранных спиральных пар обнаружено 1207742 структур, образованных двумя α-спиралями, 821798 структур, состоящих из α-спирали и 3_{10}-спирали, 172877 – образованных двумя 3_{10}-спиралями структур, 3291 структур, состоящих из α-спирали и π-спирали, 896 структур, образованных 3_{10}- и π-спиралью и одна спиральная пара, образованная двумя π-спиралями. Всего было обработано 100397 белковых структур, количество обработанных аминокислотных остатков составило 66546491, количество обработанных аминокислотных цепочек – 384666. Значительный объём данных, полученный при обработке PDB-структур, обеспечивает достаточную достоверность результатов.

МЕТОДЫ

Для исследования структурных мотивов белковых молекул были разработаны специальные правила для распознавания и отбора спиральных пар. Они были описаны в опубликованных работах [2, 3], здесь приводятся для полноты изложения.

Анализ вторичной структуры аминокислотной цепи осуществлялся с помощью метода, разработанного авторами Dictionary for Secondary Structure of Protein [14]. Находим участки спиралей. Спирали могут быть различных типов. Первый тип – это α-спираль, или H в обозначениях DSSP. Второй тип (G-спираль) включает в себя 3_{10}-спирали. Третий тип (I-спираль) – это π-спираль.

Кандидатом на искомую структуру считается участок белка, который содержит две спирали и любой длины и конфигурации участок белковой нити между спиралями, который называется перетяжкой. Все подобные структуры можно также описать согласно типам входящих в них спиралей.

Для каждой спирали структуры находится ось цилиндра, на который эта спираль намотана. Ось цилиндра определяется по методу наименьших квадратов, при котором минимизируется отклонение параметров спирали от параметров идеальной спирали [15, 16]. Качество оценки осей характеризуется величиной среднеквадратичного отклонения. Отбираются те спирали (и, соответственно, структуры), для которых точность оценки оси удовлетворяет заданному критерию.

Две оси спиралей полностью определяют пространственную структуру. Известно, что через две непересекающиеся прямые в пространстве можно провести две параллельные плоскости так, что первая ось принадлежит первой плоскости, а вторая – второй. Ось, лежащую в одной плоскости, можно спроецировать на другую плоскость. Таким образом, пространственная структура будет полностью описана расстоянием между параллельными плоскостями и проекциями осей спиралей на плоскость.

Все найденные спиральные пары разбиваются на три множества по критерию пересечения проекций спиралей на параллельные плоскости, проходящие через оси спиралей:

– множество $\{A\}$ образуют все спиральные пары, у которых проекции спиралей не имеют пересечений;

– множество $\{B\}$ образуют все спиральные пары с ненулевым пересечением проекций, за исключением тех спиральных пар, где полигон пересечения [17] включает в себя точку пересечения проекций осей спиралей;

– множество $\{C\}$ образуют все спиральные пары, где полигон пересечения включает в себя точку пересечения проекций осей спиралей.

Точечная модель и определения межспиральных расстояний, межспирального угла и торсионного угла между осями в спиральных парах

По известным координатам точек A_1, A_2, B_1 и B_2, являющимися начальными и конечными точками осей двух спиралей, вычисляются координаты векторов $\overrightarrow{A_2A_1}$, $\overrightarrow{B_1B_2}$. Точечная модель спиральной пары приведена на рисунке 1.

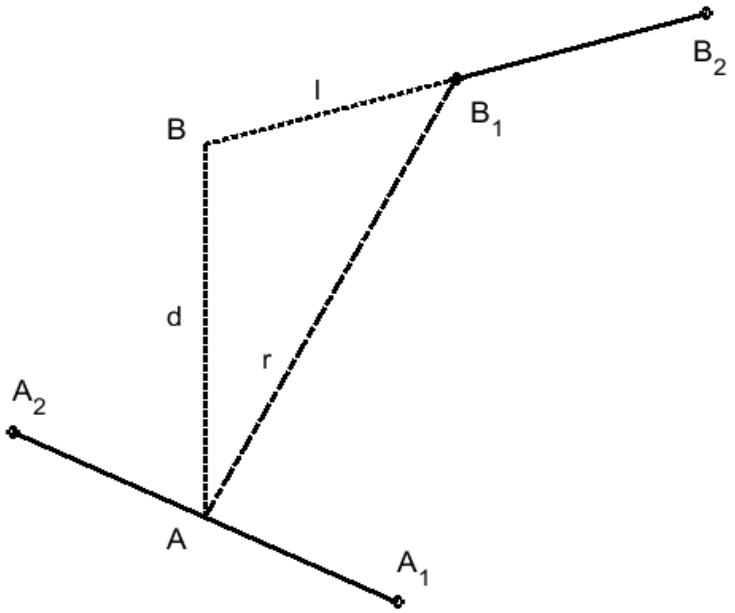

Рис. 1. Точечная модель спиральной пары. Показаны оси спиральной пары. Отрезок [A1, A2] – ось цилиндра первой спирали, [B1, B2] – ось цилиндра второй спирали.

С точки зрения описания взаимного расположения спиралей совершенно естественно возникают три расстояния. Через две непересекающихся прямые в пространстве можно единственным образом построить две параллельные плоскости, кратчайшее расстояние между которыми является расстоянием между прямыми. Расстояние между прямыми, на которых лежат оси цилиндров, будем называть межплоскостным расстоянием d спиральной пары. Второй характеристикой спиральной пары является кратчайшее расстояние r между осями цилиндров. Очевидно, что минимальное значение r всегда будет не меньше межплоскостного расстояния d, оно равно межплоскостному расстоянию d в случае, когда проекции осей спиральных пар пересекаются. Катет l ($l = \sqrt{r^2 - d^2}$) характеризует относительное расположение спиралей в спиральной паре.

Математическая биология и биоинформатика. 2017. Т. 12. № 2. doi: 10.17537/2017.12.398

Межспиральный угол φ – это угол между векторами, лежащими на осях спиралей, где начало первого вектора есть конец первой спирали, конец первого вектора – начало этой спирали, а второй вектор начинается в начале второй спирали, и конец второй спирали есть конец второго вектора. Величина угла φ не имеет знака, определяется без учета хиральности структур.

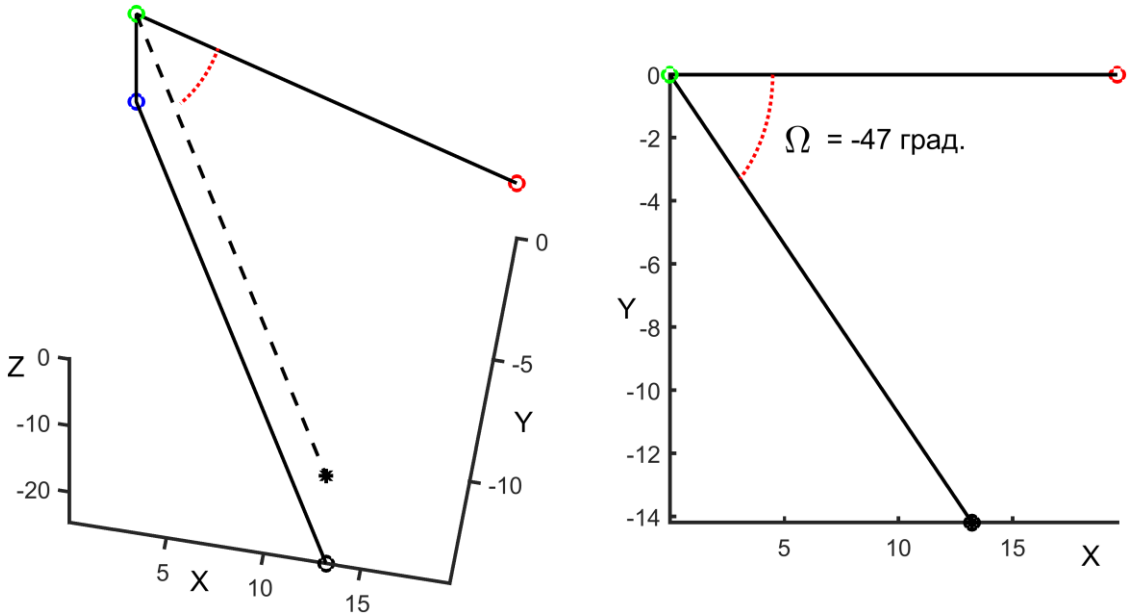

Рис. 2. Определение торсионного угла Ω между осями спиралей.

Пусть есть четыре точки A_1, A_2, B_1, B_2, являющиеся начальными и конечными точками осей двух спиралей, и три вектора $\overrightarrow{A_2A_1}$, $\overrightarrow{A_2B_1}$, и $\overrightarrow{B_1B_2}$. Если поместить точки A_1, A_2 и B_1 на координатную плоскость XY таким образом, что точка A_2 попадает в начало координат, а точка A_1 – на положительное направление оси OX, то получается угол между осью OX+ и проекцией на плоскость XY вектора $\overrightarrow{B_1B_2}$, который называется торсионным углом Ω между осями спиралей. Положительным углом $(0° \le \Omega \le 180°)$ считается направление против часовой стрелки и, соответственно, отрицательным $(-180° \le \Omega \le 0°)$ – направление по часовой стрелке.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Работа посвящена исследованию торсионных углов Ω между осями спиралей в спиральных парах белковых молекул. Все распознанные и отобранные спиральные пары разбивались на три множества по критерию пересечения проекций спиралей на параллельные плоскости, проходящие через оси спиралей [2]. При этом все отобранные спиральные пары разбиты по типу спиралей, образующих исследуемые структуры. Например, спиральная пара типа HH состоит из двух спиралей типа H (α-спиралей), HG-мотивы образованы двумя спиралями, одна из которых спираль типа H (α-спираль), другая – типа G (спираль 3_{10}); спиральная пара типа GG – это мотивы, образованные двумя спиралями типа G.

Рисунок 3 демонстрирует распределение различных типов структур, принадлежащих различным множествам, в зависимости от торсионного угла Ω. Из рисунка 3 хорошо видно существенное различие в распределении структур, принадлежащих различным множествам. Верхний график демонстрирует распределение пар, принадлежащих множеству $\{A\}$, в зависимости от угла Ω, оно достаточно широкое, охватывающее практически весь диапазон углов. Для распределения спиральных пар HH и HG типа

Математическая биология и биоинформатика. 2017. Т. 12. № 2. doi: 10.17537/2017.12.398

характерно наличие двух пиков в области 0° и 180°. Распределение пар типа GG имеет равномерный характер на всем диапазоне углов.

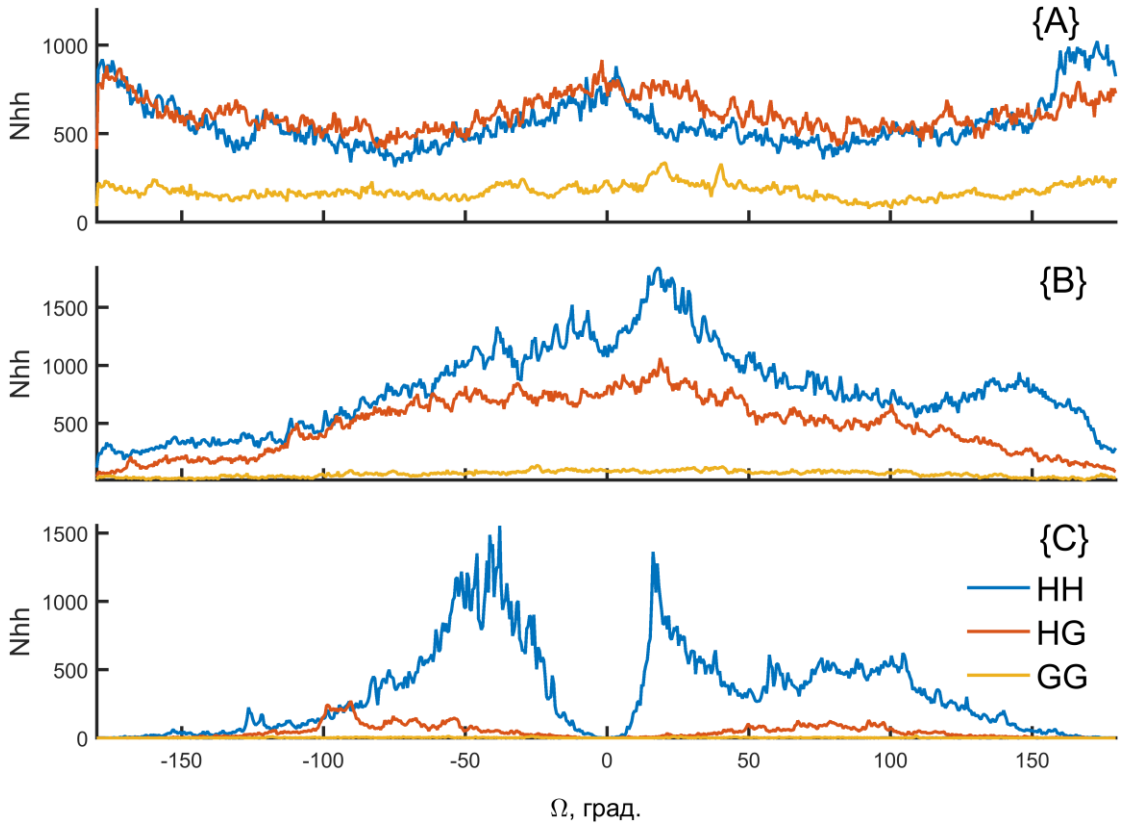

Рис. 3. Распределение структур, принадлежащих различным типам и различным множествам, в зависимости от торсионного угла Ω. (По оси абсцисс – торсионный угол Ω; по оси ординат – фактическое число отобранных структур; Синяя линия HH – структурные мотивы белков, образованные двумя *α*-спиралями; красная линия HG – мотивы, образованные α-спиралями и спиралями 3_{10}; желтая линия GG – мотивы, образованные 3_{10} -спиралями. $\{A\}$, $\{B\}$ и $\{C\}$ – множества спиральных пар, полученные по критерию пересечения проекций спиралей).

На втором графике показано распределение пар, принадлежащих множеству $\{B\}$, в зависимости от угла Ω. Из графика видно, что структур типа GG в этом множестве очень мало, они распределены на всем диапазоне углов. Распределение спиральных пар HG типа имеет максимум в области $(20° \le \Omega \le 25°)$. В этой же области $(20° \le \Omega \le 25°)$ отмечен максимум распределения спиральных пар, состоящих из двух α-спиралей, кроме того, имеется еще три локальных максимума: один – в области $(-15° \le \Omega \le -10°)$, другой – в области $\Omega \approx -30°$ и третий в области $(140° \le \Omega \le 150°)$. Также наблюдается преобладание в данном множестве структур, в образовании которых участвуют две α-спирали.

Нижний график демонстрирует распределение пар, принадлежащих множеству с пересечением проекций и осей спиралей, в зависимости от торсионного угла Ω. Из графика видно, что структур типа GG и HG в этом множестве очень мало, они распределены на всем диапазоне углов. Абсолютное большинство структур этого множества являются спиральными парами HH типа. Распределение спиральных пар, принадлежащих множеству $\{C\}$, совершенно отлично от распределения пар, принадлежащих множествам $\{A\}$ и $\{B\}$. Для этих спиральных пар характерны распределения с двумя максимумами, лежащими в области острых углов: один – в области отрицательных значений $(-50° \le \Omega \le -25°)$, другой – в области положительных

значений углов $(20° \le \Omega \le 25°)$. Провал распределения приходится на угол $\Omega = 0°$, причем в области 0° мы вовсе не наблюдаем наличие спиральных пар. Их нет и при значениях угла $(-180° \le \Omega \le -160°)$ и $(160° \le \Omega \le 180°)$. Необходимо отметить, что в области отрицательных значений Ω распределение пар более локализовано, в то время как при положительных значениях углов отмечается довольно широкое распределение, и наблюдается локальный пик в области $(60° \le \Omega \le 110°)$.

В таблице 1 представлены статистические оценки распределений торсионных углов Ω в области отрицательных и положительных углов для различных типов спиральных пар, принадлежащих множествам $\{B\}$ и $\{C\}$. В данной таблице приведены основные параметры гистограмм, изображенных на рисунке 3.

Таблица 1. Статистические оценки распределений торсионных углов Ω в области отрицательных и положительных углов для различных типов спиральных пар, принадлежащих множествам $\{B\}$ и $\{C\}$

Статистические оценки	Торсионные углы Ω в области отрицательных углов для различных типов спиральных пар			Торсионные углы Ω в области положительных углов для различных типов спиральных пар		
	HH	HG	GG	HH	HG	GG
мода $\{B\}$	–18.3	–30.55	–127.07	22.24	30.99	46.47
медиана $\{B\}$	–53.39	–57.67	–61.3	64.04	59	69.01
среднее $\{B\}$	–63.78	–63.74	–69.89	73.8	66.68	73.41
ср.–кв.откл. $\{B\}$	46.87	43.04	49.4	51.56	46.37	47.38
мода $\{C\}$	–37.69	–94.46	–26.89	16.25	81.13	20.87
медиана $\{C\}$	–49.37	–81.27	–73.18	64.86	77.55	64.59
среднее $\{C\}$	–55.35	–78.53	–77.24	66.02	77.65	69.78
ср.-кв.откл. $\{C\}$	27.26	26.5	37.78	37.94	28.41	37.36

Рисунок 4 демонстрирует распределение спиральных пар, образованных двумя α-спиралями и принадлежащих различным множествам, в зависимости от торсионного угла Ω и длины перетяжки. Верхняя диаграмма демонстрирует распределение спиральных пар, принадлежащих множеству $\{A\}$. На ней наглядно видно, что распределение достаточно широкое, охватывающее практически весь диапазон углов и весь спектр длин перетяжек. Для распределения указанных спиральных пар характерно наличие максимумов в области отрицательных $(-180° \le \Omega \le -165°)$ и положительных $(150° \le \Omega \le 180°)$ значений углов при длине перетяжки, равной 1 (N_P = 1). Наблюдается локальный пик в области 0°, причем длина перетяжки спиральных пар находится в диапазоне $(1 \le N_p \le 10)$.

На средней диаграмме показано распределение состоящих из двух α-спиралей пар, принадлежащих множеству $\{B\}$, в зависимости от угла Ω и длины перетяжки. Распределение имеет два максимума: один в области $(-25° \le \Omega \le 50°)$ при длине перетяжки $(1 \le N_p \le 8)$, другой в области $(125° \le \Omega \le 165°)$, при этом длина перетяжки имеет единственное значение $N_P = 1$. Таких структур много, это видно по шкале соответствия цвета количеству спиральных пар. Нужно отметить, что в распределении имеется еще один локальный пик в области $(125° \le \Omega \le 165°)$ и длине перетяжки $(12 \le N_p \le 18)$. Но таких структур значительно меньше, чем спиральных пар с длиной

Математическая биология и биоинформатика. 2017. Т. 12. № 2. doi: 10.17537/2017.12.398

перетяжки в одну аминокислоту ($N_P = 1$) и торсионным углом между осями спиралей $(125° \le \Omega \le 165°)$.

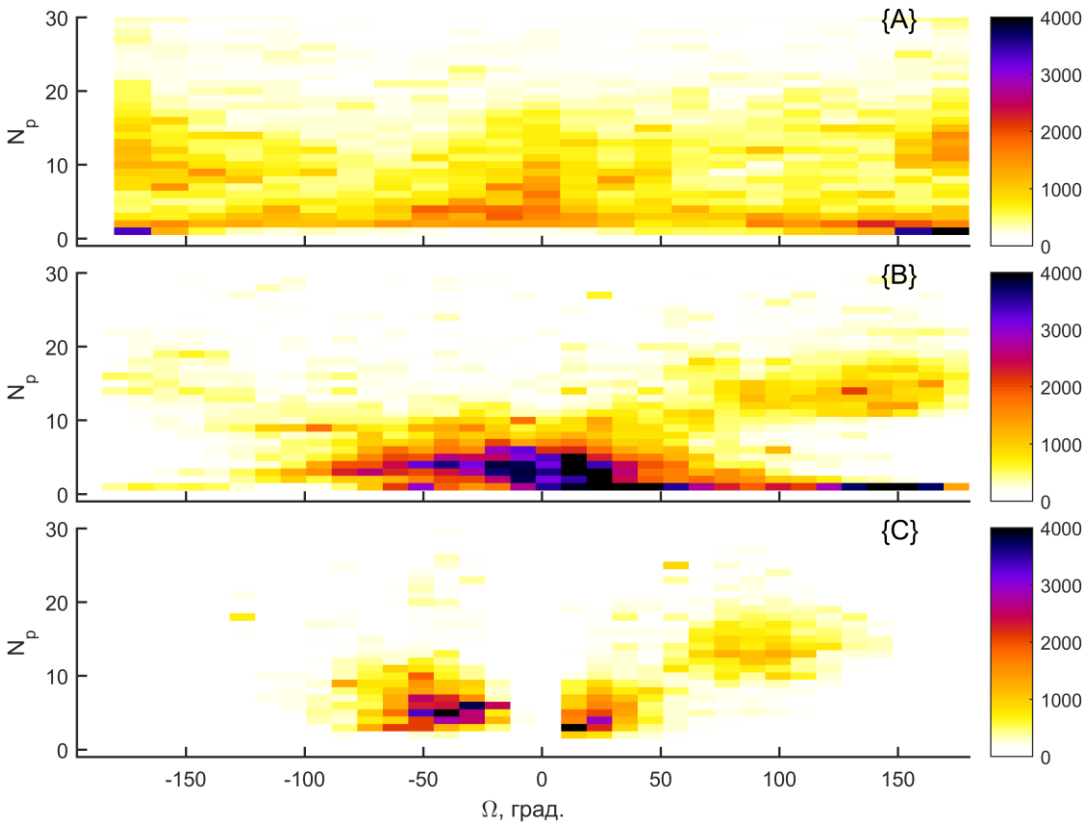

Рис. 4. Распределение спиральных пар типа НН, принадлежащих различным множествам, в зависимости от торсионного угла Ω и длины перетяжки N_p. По оси абсцисс – торсионный угол Ω, по оси ординат – количество аминокислот между спиралями. $\{A\}$, $\{B\}$ и $\{C\}$ – множества спиральных пар, полученные по критерию пересечения проекций спиралей. Справа дана шкала соответствия цвета (от белого до черного) количеству спиральных пар.

На нижней диаграмме приведено распределение пар, принадлежащих множеству с пересечением проекций и осей спиралей, в зависимости от угла Ω и длины перетяжки. Видно, что спиральных пар, состоящих из двух α-спиралей, нет совсем в областях $(-180° \le \Omega \le -100°)$, $(-15° \le \Omega \le 10°)$ и $(150° \le \Omega \le 180°)$ при любых значениях длин перетяжек. Необходимо отметить, что распределение очень локализовано. Значительное количество рассматриваемых пар имеют угол Ω в области $(-100° \le \Omega \le -20°)$ и длину перетяжки N_P в диапазоне $(3 \le N_p \le 7)$. При этом максимум распределения находится в области $(-55° \le \Omega \le -25°)$. Достаточно большое количество исследуемых структур имеют угол Ω в области $(10° \le \Omega \le 50°)$ и длину перетяжки $(2 \le N_p \le 5)$. Максимум распределения приходится на $(10° \le \Omega \le 20°)$. В распределении пар также наблюдается еще один локальный пик в области $(70° \le \Omega \le 110°)$ и длине перетяжки $(10 \le N_p \le 18)$.

Изучая распределение спиральных пар типа НН, принадлежащих различным множествам, в зависимости от торсионного угла Ω и длины перетяжки N_p можно отметить, что большинство образованных двумя α-спиралями структур имеют короткую длину перетяжки.

На рисунке 5 представлено распределение спиральных пар типа НН, принадлежащих различным множествам, в зависимости от торсионного угла Ω и

межплоскостного расстояния *d*. Верхняя диаграмма демонстрирует распределение образованных двумя α-спиралями спиральных пар, принадлежащих множеству спиральных пар, проекции спиралей которых не имеют пересечений. Распределение исследуемых структур достаточно широкое, охватывающее практически весь диапазон Ω и *d*. Для распределения спиральных пар НН типа характерно наличие двух пиков в областях углов: 0° и 180° и межплоскостного расстояния от одного до 3 Å. Причем нужно отметить, пар в указанных областях углов и, имеющих межплоскостное расстояние $d > 5$ Å, не обнаружено.

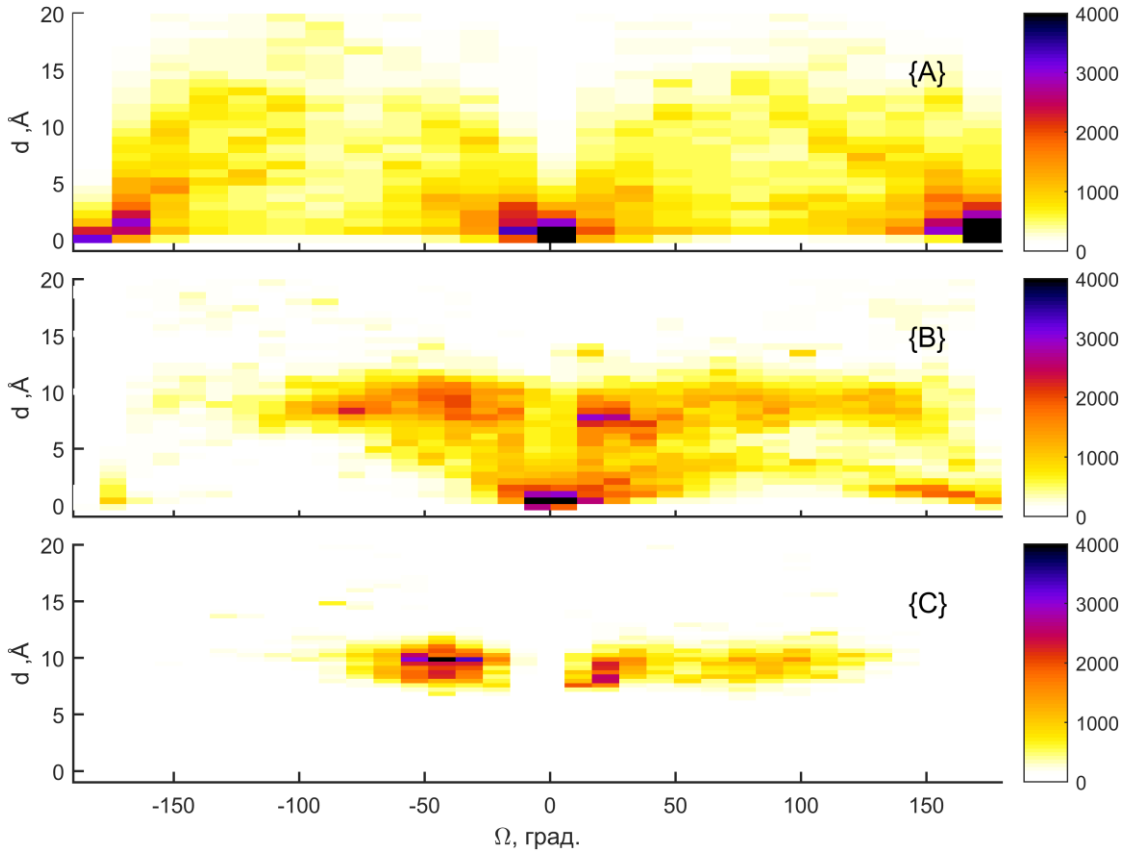

Рис. 5. Распределение спиральных пар типа НН, принадлежащих различным множествам, в зависимости от торсионного угла Ω и межплоскостного расстояния *d*. По оси абсцисс – торсионный угол Ω, по оси ординат – межплоскостное расстояние между спиралями. {*A*}, {*B*} и {*C*} – множества спиральных пар, полученные по критерию пересечения проекций спиралей. Справа дана шкала соответствия цвета (от белого до черного) количеству спиральных пар.

На второй диаграмме представлено распределение состоящих из двух α-спиралей пар, принадлежащих множеству спиральных пар, проекции спиралей которых имеют пересечение, в зависимости от торсионного угла Ω и межплоскостного расстояния *d*. Показано, что распределение данных структур имеет максимум при значениях Ω в области $(-10° \le \Omega \le 10°)$ и межплоскостного расстояния от одного до 2 Å. Имеется также три локальных максимума: $(-90° \le \Omega \le -40°)$, $(10° \le \Omega \le 40°)$ при значениях межплоскостного расстояния от 7 Å до 10 Å, и $(140° \le \Omega \le 170°)$ при $d \approx 3$ Å.

На нижней диаграмме показано распределение пар, принадлежащих множеству с пересечением проекций и осей спиралей, в зависимости от угла Ω и межплоскостного расстояния *d*. На диаграмме явно видно, что распределение сильно локализовано. Среди исследованных пар структур с пересечением проекций и осей спиралей, имеющих межплоскостное расстояние $d < 7$ Å и $d > 12$ Å и торсионный угол между осями спиралей, лежащий в областях $(-180° \le \Omega \le -80°)$, $(-10° \le \Omega \le 10°)$ и

$(140° \le \Omega \le 180°)$, не обнаружено. Есть два максимума распределения: в области отрицательных значений углов $(-60° \le \Omega \le -30°)$ и в области положительных значений торсионных углов ($\Omega \approx 20°$). Еще один пик наблюдается в области $(80° \le \Omega \le 100°)$, но количество спиральных пар с углом в области 90° между осями спиралей гораздо меньше. Межплоскостное расстояние для всех указанных локализаций спиральных пар, принадлежащих множеству $\{C\}$, находится в диапазоне от 7 Å до 12 Å, причем максимум приходится на $d \approx 10$ Å, что связано со стерическими ограничениями для спиральных пар из этого множества [18–22].

В данной работе также проведен анализ распределения спиральных пар типа НН, принадлежащих различным множествам, в зависимости от угла φ и торсионного угла Ω между осями спиралей. Результаты исследования межспиральных углов φ в структурных мотивах, образованных двумя спиралями, в том числе и в спиральных парах типа НН, были опубликованы ранее [3]. Межспиральным углом φ или углом между осями спиралей считается угол между векторами, лежащими на осях спиралей, где начало первого вектора есть конец первой спирали, конец первого вектора – начало этой спирали, а второй вектор начинается в начале второй спирали, и конец второй спирали есть конец второго вектора. Величина угла φ не имеет знака, определяется без учета хиральности структур.

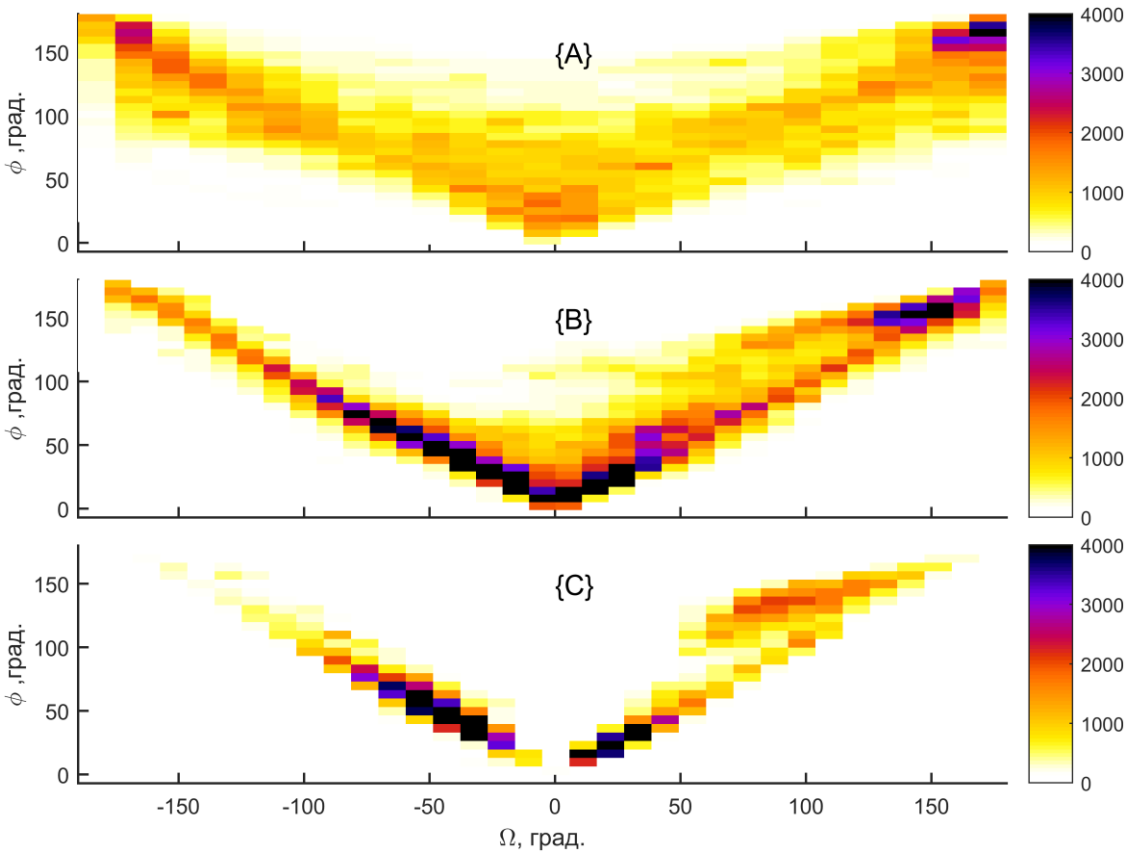

Рис. 6. Распределение спиральных пар типа НН, принадлежащих различным множествам, в зависимости от угла φ и торсионного угла Ω между осями спиралей. По оси абсцисс – торсионный угол Ω, по оси ординат – угол φ между осями спиралей. $\{A\}$, $\{B\}$ и $\{C\}$ – множества спиральных пар, полученные по критерию пересечения проекций спиралей. Справа дана шкала соответствия цвета (от белого до черного) количеству спиральных пар.

Рисунок 6 демонстрирует распределение спиральных пар типа НН, принадлежащих различным множествам, в зависимости от угла φ между осями спиралей и торсионного угла Ω. Хорошо видна идентичность распределения исследуемых структур, принадлежащих различным множествам. Разница лишь в том, что распределение пар,

принадлежащих множеству $\{A\}$, более широко, чем распределение структур, принадлежащих множествам $\{B\}$ и $\{C\}$.

На верхней диаграмме представлено распределение спиральных пар, принадлежащих множеству пар, проекции спиралей которых не имеют пересечений и в образовании которых участвуют две α-спирали, в зависимости от угла φ между осями спиралей и торсионного угла Ω. На диаграмме видно, что для всех пар указанного множества величина угла φ между осями спиралей примерна равна величине торсионного угла Ω. Нужно отметить, что это правило верно для исследуемых структур, принадлежащим всем множествам. Распределение пар множества $\{A\}$ имеет два максимума: в области отрицательных углов Ω $(-170° \le \Omega \le -160°)$ и в области положительных углов Ω $(150° \le \Omega \le 180°)$, а угол между осями спиралей φ в обоих случаях имеет значения $(150° \le \varphi \le 160°)$. Также много спиральных пар с углом φ и торсионным углом Ω между осями спиралей в области 0°.

Средняя диаграмма демонстрирует распределение спиральных пар, принадлежащих множеству $\{B\}$, в зависимости от угла φ и торсионного угла Ω между осями спиралей. Распределение более локализовано, его можно представить в виде $|\Omega| \approx \varphi$. При этом основная часть структур локализовано в области $(-90° \le \Omega \le 30°)$ и $(30° \le \varphi \le 90°)$. Имеется еще один максимум в области $\varphi \approx 150°$ и $\Omega \approx 150°$. Также из рисунка видно, что достаточно много пар указанного множества имеют углы $\varphi \approx 150°$ и $\Omega \approx 90°$.

На нижней диаграмме представлено распределение состоящих из двух α-спиралей пар, принадлежащих множеству спиральных пар, проекции спиралей и оси спиралей которых имеют пересечение, в зависимости от угла Ω и угла φ между осями спиралей. Нужно отметить, что среди исследуемых спиральных пар почти нет структур с торсионным углом, лежащим в области $(-180° \le \Omega \le -120°)$, и углом φ, лежащим в области $(120° \le \varphi \le 180°)$, также не найдено структур, у которых торсионный угол Ω и угол φ между осями спиралей имеют величину в области 0° ($\varphi \approx \Omega \approx 0°$). Имеются два максимума, лежащие в области острых углов: один – в области $(-60° \le \Omega \le -30°)$ и $(30° \le \varphi \le 60°)$, другой – в области $(5° \le \Omega \le 40°)$ и $(10° \le \varphi \le 45°)$, и еще одним локальным пиком в области: $(70° \le \Omega \le 110°)$ и $(120° \le \varphi \le 140°)$.

ЗАКЛЮЧЕНИЕ

В данной работе проведен анализ распределения торсионных углов между осями спиралей в парах связанных между собой перетяжками спиралей в пространственных структурах белковых молекул. В соответствии с разработанными правилами были отобраны спиральные пары в структурах всех белковых молекул, зарегистрированных в банке белковых структур PDB. Полученное множество спиральных пар было проанализировано с целью его классификации и установления закономерностей структурной организации. По критерию пересечения проекций спиралей на параллельные плоскости, проходящие через оси спиралей, оно было разбито на три подмножества. Были исследованы межспиральные расстояния спиральных пар, характерное количество аминокислот между спиралями и значения межспиральных углов в отобранных структурах. Каждое исследование, посвященное изучению межспиральных взаимодействий в спиральных парах, расширяет понимание и дополняет знания о структурных мотивах, имеющих уникальные укладки полипептидной цепи в пространстве. До сих пор исследований на базе структурных мотивов, отобранных из всех белковых структур, представленных в PDB, не проводилось.

В данной работе проведен анализ распределения всех типов спиральных пар, принадлежащих различным множествам, в зависимости от торсионных углов Ω между

Математическая биология и биоинформатика. 2017. Т. 12. № 2. doi: 10.17537/2017.12.398

осями спиралей отобранных структур. Показано, что распределение всех типов спиральных пар, не имеющих пересечения проекций спиралей, охватывает практически весь диапазон углов с двумя пиками в области 0° и 180°. Большинство пар указанного множества составляют спиральные пары, состоящие из α- и 3_{10}-спиралей, а множества с пересечением только проекций спиралей – спиральные пары, образованные двумя α-спиралями. Показано, распределение структур множества с пересечением проекций спиралей имеет максимум в области от 20° до 25°. Также показано, что образованные двумя α-спиралями спиральные пары составляют абсолютное большинство пар множества с пересечением проекций и осей спиралей. Для этих спиральных пар характерно распределение с двумя максимумами, лежащими в области острых углов: один – в области отрицательных значений от – 50° до – 25°, другой – в области положительных значений углов от 20° до 25° и еще одним локальным пиком в области от 70° до 110°.

Проведён также анализ распределения состоящих из двух α-спиралей спиральных пар, принадлежащих различным множествам, в зависимости от торсионного угла Ω и межплоскостного расстояния *d*. В ходе исследования показано, что распределение образованных двумя α-спиралями структур, не имеющих пересечения осей спиралей, имеет максимум при малых значениях межплоскостного расстояния от 1 Å до 3 Å и значениях торсионного угла Ω, лежащего в области 0°, а в распределении пар, не имеющих пересечения и проекций спиралей, наблюдается еще один максимум в области 180°. Среди структур, имеющих пересечение только проекций спиралей очень много пар, торсионные углы которых находятся в диапазоне от –90° до –40° и от 10° до 40° при значениях межплоскостного расстояния от 7 Å до 10 Å. Показано, что межплоскостное расстояние *d* для всех спиральных пар, имеющих пересечение осей спиралей, находится в диапазоне от 7 Å до 12 Å, причем максимум приходится на $d \approx 10$Å. При этом торсионные углы указанных структур имеют три локализации: $(-60° \le \Omega \le -30°)$, в области 20° и 90°.

Проведённый анализ распределения образованных двумя α-спиралями спиральных пар, принадлежащих различным множествам, в зависимости от торсионного угла Ω и длины перетяжки показал, что значительное количество исследуемых пар с пересечением проекций спиралей имеют короткую длину перетяжки N_P. В распределении образованных двумя α-спиралями структур, имеющих пересечения осей спиралей, наблюдаются два максимума: один в области отрицательных углов – в области $(-55° \le \Omega \le -25°)$ и длине перетяжки от трех до семи аминокислот, другой – в области положительных значений углов $(10° \le \Omega \le 20°)$ и длине перетяжки от двух до пяти аминокислот и еще одним локальным пиком в области $(70° \le \Omega \le 110°)$, а между спиралями в данных структурах находится от десяти до восемнадцати аминокислот.

Исследование выполнено при поддержке Российского фонда фундаментальных исследований (проект № 16-01-00692).

СПИСОК ЛИТЕРАТУРЫ

1. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. The Protein Data Bank. *Nucleic Acids Research*. 2000. Т. 28. С. 235–242.
2. Тихонов Д.А., Куликова Л.И., Ефимов А.В. Статистический анализ внутренних расстояний спиральных пар в белковых молекулах. *Математическая биология и биоинформатика.* 2016. Т. 11. № 2. С. 170–190. doi: 10.17537/2016.11.170
3. Тихонов Д.А., Куликова Л.И., Ефимов А.В. Исследование межспиральных углов в структурных мотивах, образованных двумя спиралями. *Математическая биология и биоинформатика.* 2017. Т. 12. № 1. С. 83–101. doi: 10.17537/2017.12.83

4. Crick F.H.C. The Packing of a-helices: simple coiled-coils. *Acta Crystallographica*. 1953. V. 6. P. 689–697.
5. Lee H.S., Choi J., Yoon S. QHELIX: A Computational tool for the improved measurement of inter-helical angles in proteins. *Protein.* 2007. V. 26. P. 556–561. doi: 10.1007/s10930-007-9097-9
6. Walther D., Eisenhaber F. and Argos P. Principles of helix-helix packing in proteins: the helical lattice superposition model. *Molecular Biology.* 1996. V. 255. P. 536–553.
7. Chothia C., Levitt M., Richardson D. Structure of proteins: packing of α-helices and pleated sheets. *Proc. Natl. Acad. Sci.* 1977. V. 74. P. 4130–4134.
8. Chothia C., Levitt M., Richardson D. Helix to helix packing in proteins. *Molecular Biology*. 1981. V. 145. P. 215–250.
9. Levitt M., Chothia C. Structural patterns in globular proteins. *Nature*. 1976. V. 261. P. 552–558.
10. Efimov A.V. Standard structures in proteins. *Prog. Biophys. Molec. Biol*. 1993. V. 60. P. 201–239.
11. Gordeev A.B., Kargatov A.M., Efimov A.V. PCBOST: Protein classification based on structural trees. *Biochemical and Biophysical Research Communications.* 2010. V. 397. P. 470–471.
12. Efimov A.V. Super-secondary structures and modeling of protein folds. In: *Methods in Molecular Biology.* Ed. Kister A.E. Clifton: Humana Press, 2013. V. 932. P. 177–189.
13. Brazhnikov E.V., Efimov A.V. Structure of α-α-hairpins with short connections in globular proteins. *Molecular Biology*. 2001. V. 35. № 1. P. 89–97.
14. Kabsch W., Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 1983. V. 22. № 12. P. 2577–2637.
15. Kabsch W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica*. 1976. V. 32. P. 922–923.
16. Kabsch W. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica*. 1978. V. 34. P. 827–828.
17. Legland D. *MatGeom: Matlab geometry toolbox for 2D/3D geometric computing.* URL: http://github.com/dlegland/matGeom (дата обращения: 11.05.2017).
18. Calhoun J.R., Kono H., Lahr S., Wang W., DeGrado W.F., Saven J.G. Computational design and characterization of a monomeric helical dinuclear metalloprotein. *Journal of Molecular Biology*. 2003. V. 334. No. 5. P. 1101–1115.
19. Calhoun J.R., Nastri F., Maglio O., Pavone V., Lombardi A., DeGrado W.F. Artificial diiron proteins: From structure to function. *Peptide Science*. 2005. V. 80. No. 2–3. P. 264–278.
20. Chino M., Maglio O., Nastri F., Pavone V., DeGrado W.F., Lombardi A. Artificial diiron enzymes with a de novo designed four-helix bundle structure. *European Journal of Inorganic Chemistry*. 2015. P. 3371–3390. doi: 10.1002/ejic.201500470
21. Chino M., Leone L., Maglio O., Lombardi A. *Designing Covalently Linked Heterodimeric Four-Helix Bundles*. 2016. P. 471–499. (Methods in Enzymology; V. 580). doi: 10.1016/bs.mie.2016.05.036.
22. Trovato A., Seno F. A new perspective on analysis of helix-helix packing preferences in globular proteins. *Proteins: Structure, Function, Bioinformatics.* 2004. V. 55. P. 1014–1022.

Рукопись поступила в редакцию 19.09.2017.
Дата опубликования 27.11.2017.