

СТРУКТУРНЫЕ И ФИЗИКО-ХИМИЧЕСКИЕ ХАРАКТЕРИСТИКИ КОНФОРМАЦИОННО-СТАБИЛЬНЫХ ОЛИГОПЕПТИДОВ α -СПИРАЛЬНОГО ТИПА

© 2015 г. А.В. Батяновский, И.Д. Волоотовский, В.А. Намиот*, И.В. Филатов**,
И.А. Галкин***, Н.В. Гнучев***, В.Г. Туманян***, Н.Г. Есипова***

Институт биофизики и клеточной инженерии НАН Беларуси, 220072, Минск, ул. Академическая, 27;

**Институт ядерной физики Московского государственного университета им. М.В. Ломоносова,
119992, Москва, Воробьевы горы, 1;*

***Московский физико-технический институт,
141700, Долгопрудный Московской области, Институтский пер., 9;*

**** Институт молекулярной биологии им. В.А. Энгельгардта РАН, 119991, Москва, ул. Вавилова, 32*

E-mail: nge@immb.ru

Поступила в редакцию 08.10.14 г.

Проведен анализ конформационно-стабильных (конформационно-консервативных) тетрапептидов, отобранных из структур белков, депонированных в банке PDBSelect. Подвыборка содержала 943 разных последовательностей аминокислот тетрапептидов, причем каждая последовательность встречалась в качестве разных сегментов структур белков не менее пяти раз. В результате анализа конформаций на основе разметки DSSP сделан вывод, что в большинстве случаев (900 из 943 последовательностей) реализуется α -спиральная конформация. Для 43 последовательностей имеет место иная конформация, в частности, замечена конформация типа левой спирали полипролина II. Физико-химические свойства конформационно-стабильных пептидов из соответствующей выборки оценивали по средней гидрофобности/гидрофильности тетрапептидов. Из результатов расчетов следует, что для конформационно-стабильных олигопептидов характерна «нейтральность» в аспекте гидрофобности/гидрофильности. Надо отметить, что дисперсия распределения гидрофобности/гидрофильности для конформационно-стабильных пептидов существенно меньше, чем для контрольных выборок. Таким образом, конформационно-стабильные олигопептиды являются выделенной группой локальных структур белка, предельно близких по конформационным и физико-химическим свойствам. В соответствии с разработанной нами ранее теорией специфических дальнедействующих взаимодействий, данные пептиды являются объектами, в наибольшей степени приспособленными для эффективного взаимного молекулярного узнавания.

Ключевые слова: конформационно-стабильные сегменты белка, локальная структура белка, молекулярное узнавание, специфические дальнедействующие взаимодействия.

Проблема связи первичной и третичной структур белка оказалась гораздо сложнее, чем это представлялось полвека назад, когда она впервые была сформулирована. Естественным паллиативом было сведение задачи определения структуры макромолекулы целого белка по последовательности аминокислот к менее масштабным и более частным задачам установления связи структура – последовательность для коротких фрагментов белка. Очевидно, что при таком подходе рассмотрению подлежат только локальные структуры, но зато для коротких участков полипептидной цепи могут быть проанализованы все последовательности (все комбинации символов). Собственно говоря, речь идет о классификации объектов по двум при-

знакам: последовательности и структуре (конформации). Подход, когда за исходные признаки принимается последовательность аминокислотных остатков в варианте, рассмотренном в работах Бэйкера и соав. [1], можно полагать наиболее результативным. Сходные последовательности объединяются в кластеры последовательностей аминокислот, а затем для каждого такого кластера исследуется характерный для него набор структур. Сильно вариабельные по пространственным структурам наборы-классы отбрасываются. Таким образом, каждый из оставшихся классов содержит родственные последовательности, которые принимают только одну или малое число локальных конформаций. При этом для соответствующих последователь-

ностей не составляет труда сделать вывод об их вероятной структуре.

В подходе Хантера [2] систематизация фрагментов ведется на структурном уровне по признаку их структурного сходства, что позволяет определить набор канонических локальных структур. После табулирования аминокислотных последовательностей для каждого из структурных кластеров и проведения расчета позиционно зависимой вероятности аминокислоты появляется возможность устанавливать структуру канонического фрагмента только на основе информации о последовательности. Находят также применение и комбинированные подходы [3].

В развиваемом нами подходе начальной посылкой является последовательность короткого сегмента, но в отличие от исследований Бэйкера требуется значительная структурная определенность (65% конформационных состояний должны быть близки), и степень этой близости также должна быть велика (отличие по двугранным углам не должно превосходить 10°). Мы выделяем, таким образом, с помощью строгого критерия новый класс объектов (конформационно-стабильные или конформационно-консервативные олигопептиды), не имеющих аналогов в работах других авторов [4,5]. В то же время, в силу введенных ограничений, класс подобного рода объектов оказывается достаточно узким. Естественно ожидать от выделенного нами класса специфических свойств, как физико-химических, так и биологических. В частности, конформационно-стабильные олигопептиды удовлетворяют условиям, необходимым для специфического дальнего действия узнавания согласно теории, развитой в работах [6,7].

Подход, заключающийся в анализе структуры олигопептидов по критерию значений двугранных углов в основной цепи, применялся в работах [8–10]. Целью такого анализа было предсказание вторичной структуры белка: α -спиралей и β -структуры в первых двух цитированных работах и вторичной структуры типа полипролин II в третьей работе, выполненной в российском коллективе.

Ранее нами была начата работа по отбору конформационно-стабильных тетрапептидов [4]. Конформационно-стабильный олигопептид – сегмент белка, проявляющий в широкой выборке белковых структур явные конформационные предпочтения, т.е. в большинстве белковых структур встречающийся в каком-либо одном, свойственном ему конформационном состоянии. Изучение таких фрагментов позволило выявить некоторые особенности взаимного рас-

положения этих фрагментов в белковых глобулах, как-то отмечена близость таких фрагментов в глобулярных структурах, что может указывать на координирующую функцию выделенных последовательностей в структурной организации глобулы белка [5].

Цель настоящей работы – исследовать некоторые конформационные и физико-химические характеристики конформационно-стабильных сегментов полипептидных цепей. Тип конформации тетрапептида в данной работе определяли по шкале DSSP; при оценке физико-химических свойств принималась во внимание степень гидрофобности/гидрофильности сегмента.

МАТЕРИАЛЫ И МЕТОДЫ

Список конформационно-стабильных тетрапептидных последовательностей аминокислот был составлен на основе выборки структур белков из банка PDBselect (степень гомологии не выше 35%). Было отобрано более 1400 белковых структур, расшифрованных методом рентгеноструктурного анализа с разрешением не хуже, чем $1,5 \text{ \AA}$. Как указано выше, под конформационно-стабильным тетрапептидом понимается сегмент белка с последовательностью аминокислот, которая в широком наборе белков, низкогомологичных по последовательности аминокислот, встречается преимущественно в определенной конформации (например, последовательность АААК встречается в нашей выборке белков 19 раз, причем 13 раз полипептидная цепь с этой последовательностью находится в конформации α -спирали).

Сравнение структур в процессе формирования списка проводили по двугранным углам, оценивая среднее квадратичное отклонение между двугранными углами двух сравниваемых олигопептидов. Мы, как уже указано, считали конформационно-стабильными те олигопептиды, для которых более 65% наблюдаемых конформационных состояний близки друг к другу. Было обнаружено, что только 29 из 8000 комбинаторно возможных последовательностей трипептидов обладают предпочтительными конформациями (более 0,36%). В качестве стабильных тетрапептидов было идентифицировано ~1500 из 160000 комбинаторно возможных последовательностей аминокислотных сегментов размером в четыре аминокислоты (чуть менее 1%). Еще раз напомним, что конформационная определенность растет с увеличением длины сегмента. Особенностью списка конформационно-стабильных структур по сравнению с выборками близких локальных структур в работах других авторов является прежде всего

то, что предпочтительной конформацией для основной массы тетрапептидов из нашего списка является α -спираль [5]. Это неудивительно, так как разброс в углах ϕ и ψ оказывается наименьшим для α -спирали, ввиду особенностей распределения дипольных моментов пептидных групп остова полипептидной цепи в симметрии α -спирали.

Программы, использованные при обработке информации о белковых структурах, в том числе конформационно-стабильных фрагментах, представленных в формате PDB в списках конформационно-стабильных фрагментов, были реализованы нами на языке Python. Статистическую обработку данных вели как с использованием оригинальных программных средств, так и в среде статистических вычислений R.

Гидрофобность фрагментов определяли по суммарной гидрофобности боковых групп аминокислот; оценки значений гидрофобности отдельных боковых групп взяты из работы [11]. В работе [11] определение гидрофобности опирается на три разных литературных источника [12–14], в которых были применены различные методические подходы. Соответствующие значения гидрофобности боковых радикалов аминокислот суммированы в табл. 1. Подчеркнем, что анализ носит сравнительный характер. Оценки были проведены для всех трех наборов значений гидрофобности, приведенных в табл. 1.

Гидрофобности коротких конформационно-стабильных фрагментов (в нашем случае тетрапептидов) сравнивали с гидрофобностями фрагментов с последовательностями аминокислот, сформированными случайным комбинированием аминокислот с частотами, которые были определены двумя способами:

1) в соответствии со встречаемостью аминокислот в α -спиральных фрагментах белковых структур,

2) в соответствии со встречаемостью аминокислот во фрагментах, определенных как конформационно-стабильные олигопептиды из набора исследованных белковых структур.

Значения гидрофобности, определенные по разным шкалам, использовали для сопоставления статистических характеристик физико-химических свойств выборок, содержащих короткие конформационно-стабильные фрагменты, и выборок, полученных случайным комбинированием расположения аминокислот на основе установленных наборов частот встречаемости. Проведен детальный сравнительный анализ статистических свойств пептидных сегментов из выборки конформационно-стабильных тетра-

Таблица 1. Свободная энергия переноса аминокислоты из воды в неполярную среду за вычетом свободной энергии переноса глицина для трех шкал гидрофобности (а), (б) и (в) из работы [11]

Шкала гидрофобности	Свободная энергия переноса, кДж/моль		
	(а) [12]	(б) [13]	(в) [14]
Аминокислота			
Phe F	-12,3	-10,1	-4,79
Ile I	-11,4	-10,2	-1,34
Try W	-11,3	-12,8	-7,81
Leu L	-8,95	-9,64	-2,39
Tyr Y	-8,70	-5,44	-3,99
Met M	-8,23	-6,97	-1,01
Val V	-7,80	-6,91	+0,25
Pro P	-7,15	-4,08	+1,85
Cys C	-4,05	-8,73	-1,05
Glu E	-3,02	+3,63	+8,44
Ala A	-2,67	-1,76	+0,67
Thr T	-2,50	-1,47	+0,55
Gln Q	-1,95	+1,25	+2,39
Asp D	-1,03	+4,36	+5,12
Gly G	0,00	0,00	0,00
Ser S	+0,42	+0,23	+0,50
Asn N	+1,25	+3,40	+1,72
Arg R	+1,75	+5,72	+3,36
Lys K	+3,00	+5,61	+4,12
His H	+4,22	-0,74	+3,99

пептидов и двух выборок реально наблюдаемых тетрапептидов (всех тетрапептидов и только α -спиральных).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Цель нашей работы – проведение анализа структурных и физико-химических характеристик отобранных нами конформационно-стабильных тетрапептидов для установления их роли в процессах молекулярного узнавания в белковых системах. В конкретные задачи входят анализ вклада разных конформаций коротких пептидов и условия их перехода (их стабильности) из одной устойчивой формы – в другую, менее устойчивую.

Чтобы установить природу выделенной конформационно-стабильной группы олигопептидных фрагментов белка, необходимо исследование физико-химических характеристик сегментов полипептидных цепей, выделенных как конформационно-стабильные. Разметка DSSP [15] сама по себе дает возможность ориентации в этом вопросе. Дело в том, что эта разметка косвенно указывает на возможность образования водородных связей, причем это важно для

суждения о степени стабильности фрагмента. В этой связи наше обращение к оценке гидрофобности/гидрофильности как важному свойству, определяющему степень стабильности конформаций, представляется адекватным поставленной задаче.

Чтобы провести корректное исследование физико-химических свойств (в данном случае гидрофобности/гидрофильности) существенное значение имеет однородность выборки в аспекте конформации. Т.е. для объектов выборки должна быть характерна одна определенная конформация или, по крайней мере, речь должна идти о, безусловно, преобладающей конформации. Конечно, современные методы позволяют работать и с объектами, представляющими собой определенное разнообразие, интерпретация данных физико-химического анализа в этом случае, естественно, осложнена.

Анализ структуры конформационно-стабильных пептидов имеет и самостоятельный интерес, поскольку позволяет сделать вывод о том, какие стандартные или нестандартные конформации характерны в этом случае и в какой пропорции.

Анализ в терминах DSSP показал, что в подавляющем числе случаев тетрапептиды принимают α -спиральную конформацию (НННН в номенклатуре DSSP). Исследование конформационно-стабильных тетрапептидов в этой конформации и есть предмет настоящей работы. В табл. 2 приводятся 900 последовательностей аминокислот тетрапептидов, имеющих α -спиральную конформацию. Как видно из таблицы, этим последовательностям соответствуют 5335 структур сегментов полипептидной цепи в глобулярных белках, т.е. в среднем на определенную последовательность аминокислот приходится около шести реальных сегментов белковых структур.

Однако, если говорить обо всей выборке, то для 43 последовательностей либо не все остатки имеют α -спиральную конформацию, либо α -спиральная конформация совсем отсутствует. Соответствующий материал приведен в табл. 3. Число структур тетрапептидов, соответствующих этим 43 последовательностям, равно 335 (см. табл. 3).

Таким образом, для конформационно-стабильных олигопептидов из исследованной выборки характерна одна конформация, а именно α -спиральная конформация. Однако принципиальное значение имеет то, что другие конформации (пусть в малом количестве) также возможны, а среди них левая спираль типа полипролин II.

Теперь, после того как мы убедились в том, что структура конформационно-стабильных пептидов практически одинакова, мы можем с тем большим основанием исследовать физико-химические свойства этих объектов. Результаты расчета гидрофобности отображены в табл. 4.

Как мы видим, средние величины значений гидрофобности стабильных олигопептидов близки к средним значениям гидрофобности списка случайных последовательностей, за исключением случая определения гидрофобности по шкале (в), что связано с различием в частотах встречаемости аминокислотных остатков в стабильных фрагментах и всех α -спиральных фрагментах. Если вдуматься, неожиданным является то, что различия в средних величинах гидрофобности для тетрапептидов, полученных на основании частот встречаемости аминокислот в α -спиральных сегментах белковой глобулы и на основании этих частот в конформационно-стабильных фрагментах, проявились лишь в случае со шкалой (в). Проанализировав разности величин произведений гидрофобности и частот встречаемости отдельных аминокислот из α -спиральных сегментов и из стабильных сегментов при определении гидрофобности по шкале (б) (данные представлены в табл. 5) и (а) (данные не приведены) с аналогичными величинами, определенными по шкале (в) (табл. 6), мы обнаружили, что при том, что в шкалах (а) и (б) среднее значение модулей этих величин заметно больше, чем в шкале (в) ((б) : 4,6; (в) : 1,3), суммы их для шкал (а), (б) и (в) по абсолютной величине близки ((б) : -0,57; (в) : 0,52). Это обстоятельство при комбинации аминокислот в тетрапептиды должно приводить к меньшим различиям в средней величине гидрофобности для шкал (а) и (б), поскольку в нашей работе гидрофобности фрагментов определяются суммированием гидрофобностей входящих во фрагмент аминокислот. Поскольку трудно сказать, какой из растворителей, использованных для построения шкал гидрофобности, лучше моделирует среду клетки, остается только констатировать, что для двух шкал гидрофобности различия в средних величинах для разных выборок не наблюдаются, для третьей шкалы имеются достоверные различия.

Что касается важной задачи работы: очевидно значимое отличие по уровню дисперсии гидрофобности в наборе стабильных олигопептидов и в списках случайных последовательностей. Это подтверждает предположение о том, что стабильные фрагменты менее разбросаны относительно среднего значения, а, значит можно сказать, что они проявляют большую независимость от среды, чем случайные фрагменты.

Таблица 2. Конформационно-стабильные тетрапептиды, структура которых является α -спиральной. N – число структур тетрапептидов с указанной последовательностью и с указанной разметкой. Условные обозначения DSSP: H – α -спираль

Последовательность	Разметка DSSP	N	Последовательность	Разметка DSSP	N	Последовательность	Разметка DSSP	N
AAAA	HHHH	18	ADLR	HHHH	5	AIRK	HHHH	4
AAAI	HHHH	8	ADLV	HHHH	6	AKAA	HHHH	7
AAAV	HHHH	10	ADQY	HHHH	4	AKAE	HHHH	4
AAAY	HHHH	4	ADVA	HHHH	4	AKAI	HHHH	5
AADR	HHHH	4	ADVL	HHHH	5	AKAV	HHHH	4
AADV	HHHH	5	ADYL	HHHH	6	AKEL	HHHH	8
AAEF	HHHH	6	AEAF	HHHH	4	AKIL	HHHH	4
AAEI	HHHH	6	AEAI	HHHH	6	AKKL	HHHH	5
AAEL	HHHH	10	AEAL	HHHH	16	AKLA	HHHH	6
AAFL	HHHH	4	AEEA	HHHH	5	AKRL	HHHH	5
AAIA	HHHH	13	AEIA	HHHH	11	ALAE	HHHH	14
AAID	HHHH	7	AEIE	HHHH	5	ALAF	HHHH	4
AAIE	HHHH	11	AEKA	HHHH	9	ALAI	HHHH	8
AAII	HHHH	7	AELA	HHHH	11	ALAK	HHHH	7
AAIN	HHHH	5	AELE	HHHH	6	ALAL	HHHH	8
AAIR	HHHH	6	AELL	HHHH	8	ALAQ	HHHH	8
AAKA	HHHH	9	AELV	HHHH	8	ALDI	HHHH	4
AALA	HHHH	19	AENI	HHHH	5	ALEE	HHHH	8
AALI	HHHH	7	AEQL	HHHH	6	ALEG	HHHH	4
AALK	HHHH	8	AEQV	HHHH	4	ALEK	HHHH	13
AALL	HHHH	8	AERL	HHHH	4	ALEV	HHHH	4
AALQ	HHHH	6	AERQ	HHHH	6	ALHV	HHHH	4
AALR	HHHH	6	AESI	HHHH	4	ALIR	HHHH	5
AAMA	HHHH	5	AESQ	HHHH	4	ALKA	HHHH	7
AAQI	HHHH	6	AETK	HHHH	4	ALKI	HHHH	4
AAQK	HHHH	6	AETL	HHHH	5	ALKK	HHHH	9
AAQL	HHHH	7	AEVL	HHHH	6	ALLA	HHHH	15
AAQQ	HHHH	5	AEVY	HHHH	4	ALLD	HHHH	7
AAQR	HHHH	5	AFEH	HHHH	4	ALLE	HHHH	11
AARA	HHHH	9	AHVE	HHHH	4	ALLK	HHHH	8
AASI	HHHH	5	AIAE	HHHH	6	ALLL	HHHH	8
AATA	HHHH	7	AIAK	HHHH	6	ALLQ	HHHH	6
AAVA	HHHH	13	AIAQ	HHHH	4	ALLR	HHHH	5
AAVR	HHHH	8	AIAR	HHHH	5	ALMD	HHHH	5
AAVS	HHHH	7	AIEA	HHHH	13	ALRA	HHHH	9
AAVA	HHHH	7	AIEM	HHHH	5	ALRD	HHHH	5
ADAL	HHHH	13	AIER	HHHH	7	ALRE	HHHH	6
ADAR	HHHH	5	AIET	HHHH	4	ALRL	HHHH	6
ADAV	HHHH	7	AILE	HHHH	5	ALTR	HHHH	5
ADDI	HHHH	4	AINA	HHHH	4	ALVA	HHHH	8
ADEL	HHHH	6	AIQK	HHHH	5	ALVE	HHHH	7
ADIA	HHHH	6	AIRA	HHHH	5	ALVK	HHHH	8
ADLI	HHHH	9	AIRE	HHHH	6	AMLK	HHHH	4

Продолжение

Последовательность	Разметка DSSP	<i>N</i>	Последовательность	Разметка DSSP	<i>N</i>	Последовательность	Разметка DSSP	<i>N</i>
ANAV	HHHH	8	AVEK	HHHH	4	DVEA	HHHH	5
ANLL	HHHH	4	AVEQ	HHHH	5	DVLE	HHHH	6
ANLR	HHHH	4	AVLD	HHHH	5	DVLT	HHHH	5
AQAA	HHHH	8	AVNK	HHHH	5	DYIK	HHHH	6
AQAV	HHHH	5	AVRQ	HHHH	5	DYLA	HHHH	5
AQEA	HHHH	4	AVTL	HHHH	4	EAAA	HHHH	10
AQEI	HHHH	5	AWLA	HHHH	5	EAAE	HHHH	5
AQEL	HHHH	7	AYAA	HHHH	8	EAAI	HHHH	6
AQIA	HHHH	7	CALL	HHHH	4	EAAK	HHHH	6
AQII	HHHH	4	DAAI	HHHH	6	EAAI	HHHH	9
AQIL	HHHH	5	DAFL	HHHH	4	EAAQ	HHHH	8
AQLA	HHHH	9	DAIA	HHHH	7	EAAAR	HHHH	8
AQLL	HHHH	5	DAIE	HHHH	5	EAAV	HHHH	5
AQLR	HHHH	5	DALA	HHHH	10	EAAAY	HHHH	4
AQQE	HHHH	4	DALK	HHHH	8	EAEA	HHHH	6
AQQL	HHHH	5	DALL	HHHH	6	EAFK	HHHH	7
AQSL	HHHH	4	DALR	HHHH	8	EAFI	HHHH	5
AQVL	HHHH	7	DALT	HHHH	5	EAFR	HHHH	5
ARAA	HHHH	6	DAVA	HHHH	9	EAGL	HHHH	4
ARAE	HHHH	4	DAVR	HHHH	8	EAI	HHHH	6
ARAL	HHHH	9	DDIL	HHHH	4	EAI	HHHH	4
AREL	HHHH	7	DDIR	HHHH	5	EAI	HHHH	5
ARKL	HHHH	6	DEAA	HHHH	6	EAKA	HHHH	5
ARLK	HHHH	4	DEAI	HHHH	5	EALA	HHHH	24
ARLQ	HHHH	6	DEEI	HHHH	4	EALD	HHHH	6
ARQA	HHHH	5	DEIA	HHHH	4	EALE	HHHH	18
ARQF	HHHH	4	DFEK	HHHH	4	EALI	HHHH	7
ARTL	HHHH	5	DFLA	HHHH	4	EALL	HHHH	9
ASAL	HHHH	9	DIAA	HHHH	6	EALQ	HHHH	6
ASAV	HHHH	4	DIAE	HHHH	4	EALR	HHHH	9
ASIE	HHHH	6	DIAL	HHHH	12	EAQK	HHHH	5
ASLI	HHHH	9	DIAT	HHHH	4	EARA	HHHH	5
ASLL	HHHH	5	DKIR	HHHH	4	EARK	HHHH	5
ASRL	HHHH	5	DLAK	HHHH	4	EASA	HHHH	4
ATEL	HHHH	6	DLIK	HHHH	5	EATL	HHHH	7
ATLA	HHHH	5	DLLE	HHHH	9	EAVA	HHHH	9
ATLR	HHHH	7	DLLK	HHHH	8	EAVE	HHHH	6
ATRI	HHHH	4	DLLL	HHHH	5	EAVK	HHHH	8
AVAA	HHHH	13	DLLR	HHHH	6	EAVQ	HHHH	4
AVAK	HHHH	5	DLLV	HHHH	6	EAVR	HHHH	12
AVAR	HHHH	5	DLVD	HHHH	4	EAVS	HHHH	4
AVAT	HHHH	5	DQLA	HHHH	4	EAVV	HHHH	7
AVDA	HHHH	7	DRIL	HHHH	5	EDFI	HHHH	4
AVDY	HHHH	4	DSVN	HHHH	4	EDIA	HHHH	6
AVEE	HHHH	5	DTLI	HHHH	4	EDLK	HHHH	4

Продолжение

Последовательность	Разметка DSSP	N	Последовательность	Разметка DSSP	N	Последовательность	Разметка DSSP	N
EDLR	HHHH	5	EIIS	HHHH	4	EQIQ	HHHH	6
EDVE	HHHH	5	EILQ	HHHH	5	EQLA	HHHH	8
EDVL	HHHH	4	EILR	HHHH	6	EQLL	HHHH	9
EEAA	HHHH	5	EIRE	HHHH	5	EQLR	HHHH	7
EEAF	HHHH	4	EIRR	HHHH	6	EQRL	HHHH	4
EEAK	HHHH	7	EIVD	HHHH	4	ERAA	HHHH	7
EEAL	HHHH	10	EIVE	HHHH	5	ERAL	HHHH	4
EEAR	HHHH	8	EIVR	HHHH	6	ERIA	HHHH	5
EEAV	HHHH	10	EKAA	HHHH	8	ERIL	HHHH	5
EEDI	HHHH	5	EKAL	HHHH	6	ERLA	HHHH	9
EEDL	HHHH	7	EKAQ	HHHH	5	ERLL	HHHH	5
EEEI	HHHH	7	EKAR	HHHH	5	ERLR	HHHH	9
EEEL	HHHH	7	EKIA	HHHH	6	ERRI	HHHH	4
EEFL	HHHH	7	EKIR	HHHH	4	ESAI	HHHH	4
EEFV	HHHH	4	EKLE	HHHH	6	ESAR	HHHH	4
EEIE	HHHH	7	EKLL	HHHH	8	ESIR	HHHH	5
EEII	HHHH	6	EKLR	HHHH	7	ESLR	HHHH	5
EEIK	HHHH	4	EKRL	HHHH	8	ETAK	HHHH	4
EEIL	HHHH	8	EKVK	HHHH	5	ETFL	HHHH	5
EEIQ	HHHH	5	ELAA	HHHH	11	ETLL	HHHH	8
EEIR	HHHH	6	ELAE	HHHH	6	ETLR	HHHH	5
EELA	HHHH	9	ELAQ	HHHH	9	ETVA	HHHH	6
EELE	HHHH	9	ELAR	HHHH	8	EVAA	HHHH	6
EELK	HHHH	8	ELAV	HHHH	4	EVAE	HHHH	8
EELL	HHHH	8	ELEA	HHHH	7	EVAR	HHHH	4
EELR	HHHH	9	ELIE	HHHH	5	EVAT	HHHH	4
EERL	HHHH	4	ELKA	HHHH	8	EVIA	HHHH	5
EETL	HHHH	8	ELKK	HHHH	6	EVKA	HHHH	5
EEVK	HHHH	7	ELKR	HHHH	6	EVKR	HHHH	4
EEVR	HHHH	5	ELLA	HHHH	10	EVLK	HHHH	7
EEYL	HHHH	5	ELLK	HHHH	9	EVRA	HHHH	5
EFAK	HHHH	9	ELLL	HHHH	5	EVTR	HHHH	4
EFEK	HHHH	5	ELLN	HHHH	5	EVVK	HHHH	7
EFLA	HHHH	5	ELLQ	HHHH	7	EYAA	HHHH	4
EFLE	HHHH	6	ELLR	HHHH	7	FADY	HHHH	4
EFLK	HHHH	7	ELQE	HHHH	5	FAKA	HHHH	4
EFLQ	HHHH	4	ELRE	HHHH	6	FAKI	HHHH	6
EFTR	HHHH	4	ELRK	HHHH	4	FAKL	HHHH	5
EFVK	HHHH	6	ELVE	HHHH	7	FALA	HHHH	4
EGVR	HHHH	6	ELVR	HHHH	6	FAQL	HHHH	4
EHLA	HHHH	5	ELYR	HHHH	4	FAVA	HHHH	4
EIAA	HHHH	4	EMIR	HHHH	5	FEEA	HHHH	6
EIAR	HHHH	11	EQAA	HHHH	5	FEKA	HHHH	4
EIER	HHHH	4	EQAV	HHHH	4	FIRK	HHHH	4
EIIR	HHHH	5	EQIA	HHHH	7	FKEA	HHHH	4

Последовательность	Разметка DSSP	<i>N</i>	Последовательность	Разметка DSSP	<i>N</i>	Последовательность	Разметка DSSP	<i>N</i>
FKKL	HHHH	5	IKAV	HHHH	7	KEEV	HHHH	5
FKTA	HHHH	4	IKEK	HHHH	5	KEIR	HHHH	6
FLEA	HHHH	5	IKQL	HHHH	5	KELA	HHHH	7
FLKD	HHHH	4	IKRL	HHHH	4	KELE	HHHH	7
FQEA	HHHH	4	ILAQ	HHHH	4	KELI	HHHH	6
FRQA	HHHH	5	ILEA	HHHH	4	KELR	HHHH	5
FSAV	HHHH	4	ILKE	HHHH	5	KELS	HHHH	6
FTRL	HHHH	5	ILNT	HHHH	4	KEVA	HHHH	8
FYQA	HHHH	4	ILQA	HHHH	4	KEVE	HHHH	4
GAAI	HHHH	5	INAI	HHHH	6	KFVE	HHHH	7
GDAF	HHHH	5	INAV	HHHH	5	KIAK	HHHH	7
GDFL	HHHH	4	INKL	HHHH	5	KIAQ	HHHH	4
GEAL	HHHH	5	IQEV	HHHH	4	KKAI	HHHH	4
HAAA	HHHH	4	IRAQ	HHHH	4	KKAL	HHHH	5
HLLA	HHHH	4	IRDE	HHHH	4	KKIA	HHHH	6
HRLR	HHHH	4	IREI	HHHH	6	KKLG	HHHH	5
IAAA	HHHH	9	IREL	HHHH	5	KKLQ	HHHH	5
IAAF	HHHH	6	IRET	HHHH	4	KLAE	HHHH	9
IAAI	HHHH	4	IRKA	HHHH	6	KLAQ	HHHH	5
IAAQ	HHHH	6	IRLA	HHHH	4	KLIE	HHHH	7
IAAW	HHHH	4	IRRI	HHHH	4	KLIK	HHHH	5
IAEE	HHHH	4	ISAL	HHHH	4	KLIQ	HHHH	4
IAEK	HHHH	4	ISLA	HHHH	4	KLKE	HHHH	7
IAEL	HHHH	4	IVEA	HHHH	5	KLLA	HHHH	6
IAET	HHHH	4	IVQQ	HHHH	5	KLLK	HHHH	7
IAGL	HHHH	7	KAAA	HHHH	6	KLLL	HHHH	6
IAKA	HHHH	8	KAAE	HHHH	5	KLQD	HHHH	4
IAKE	HHHH	5	KAAI	HHHH	6	KLRA	HHHH	6
IAKF	HHHH	4	KAAC	HHHH	5	KLVE	HHHH	11
IAKS	HHHH	4	KAAL	HHHH	11	KLYK	HHHH	5
IALA	HHHH	5	KAAQ	HHHH	6	KQLQ	HHHH	5
IALV	HHHH	7	KAFL	HHHH	6	KQRL	HHHH	5
IAQA	HHHH	9	KAIA	HHHH	4	KRAL	HHHH	5
IARK	HHHH	5	KAID	HHHH	5	KRAV	HHHH	6
IDAL	HHHH	7	KAIE	HHHH	9	KRLE	HHHH	5
IDLL	HHHH	4	KAKE	HHHH	5	KTAL	HHHH	6
IDTL	HHHH	5	KALA	HHHH	10	KVAE	HHHH	4
IEAA	HHHH	6	KAVD	HHHH	4	KVLA	HHHH	7
IEAL	HHHH	10	KAYE	HHHH	5	KVLE	HHHH	5
IEEL	HHHH	5	KDAL	HHHH	4	KYLA	HHHH	5
IEEV	HHHH	4	KDIA	HHHH	4	LAAE	HHHH	7
IEIA	HHHH	5	KDIE	HHHH	5	LAAI	HHHH	8
IERL	HHHH	9	KDLE	HHHH	5	LAAL	HHHH	11
IIDD	HHHH	4	KEAL	HHHH	7	LAAQ	HHHH	6
IIRL	HHHH	6	KEAQ	HHHH	4	LAAR	HHHH	13

Продолжение

Последовательность	Разметка DSSP	<i>N</i>	Последовательность	Разметка DSSP	<i>N</i>	Последовательность	Разметка DSSP	<i>N</i>
LAAY	HHHH	4	LEIA	HHHH	4	LLQL	HHHH	5
LADL	HHHH	9	LEII	HHHH	4	LLRA	HHHH	6
LADV	HHHH	5	LEKI	HHHH	9	LNAL	HHHH	5
LAEL	HHHH	17	LEKL	HHHH	9	LNRL	HHHH	6
LAER	HHHH	8	LEKR	HHHH	7	LNSL	HHHH	6
LAFL	HHHH	5	LELL	HHHH	6	LQAL	HHHH	9
LAIA	HHHH	9	LEQL	HHHH	13	LQDV	HHHH	4
LAIE	HHHH	6	LESA	HHHH	7	LQEY	HHHH	5
LAII	HHHH	5	LEVA	HHHH	5	LQLT	HHHH	5
LAKA	HHHH	12	LEYL	HHHH	5	LQRL	HHHH	5
LAKE	HHHH	8	LFEE	HHHH	4	LQVL	HHHH	4
LAKI	HHHH	4	LFES	HHHH	4	LQVQ	HHHH	4
LAKR	HHHH	6	LGEL	HHHH	5	LRAA	HHHH	7
LALA	HHHH	7	LIAA	HHHH	6	LRAI	HHHH	8
LALS	HHHH	5	LIAK	HHHH	5	LRAL	HHHH	10
LAQA	HHHH	5	LIAL	HHHH	5	LREK	HHHH	4
LAQE	HHHH	6	LIEE	HHHH	5	LREL	HHHH	10
LAQI	HHHH	4	LIEK	HHHH	7	LRGI	HHHH	5
LAQL	HHHH	8	LIKE	HHHH	6	LRKA	HHHH	4
LARA	HHHH	4	LKAA	HHHH	12	LRKL	HHHH	6
LARL	HHHH	6	LKEA	HHHH	10	LRKY	HHHH	4
LASL	HHHH	9	LKEV	HHHH	8	LRLI	HHHH	7
LAVE	HHHH	9	LKFL	HHHH	4	LRNV	HHHH	4
LDAI	HHHH	5	LKKA	HHHH	7	LRQL	HHHH	6
LDAL	HHHH	7	LKKI	HHHH	4	LRTL	HHHH	7
LDDI	HHHH	4	LKKK	HHHH	5	LTAA	HHHH	6
LDDV	HHHH	4	LKLA	HHHH	4	LTAF	HHHH	4
LDEI	HHHH	9	LKQL	HHHH	4	LTEL	HHHH	6
LDEV	HHHH	4	LKRI	HHHH	6	LTRL	HHHH	4
LDIA	HHHH	5	LLAA	HHHH	11	LTVW	HHHH	4
LDKA	HHHH	6	LLAI	HHHH	5	LVEF	HHHH	6
LDKI	HHHH	5	LLAL	HHHH	7	LVEK	HHHH	6
LDVL	HHHH	5	LLAR	HHHH	6	LVER	HHHH	4
LDYL	HHHH	6	LLDA	HHHH	5	LVKK	HHHH	6
LEAA	HHHH	13	LLED	HHHH	10	LVRA	HHHH	4
LEAF	HHHH	5	LLEE	HHHH	8	LYRE	HHHH	4
LEAI	HHHH	6	LLEQ	HHHH	5	MARL	HHHH	4
LEAL	HHHH	12	LLIE	HHHH	4	MEAL	HHHH	5
LEAT	HHHH	7	LLKA	HHHH	7	NAAI	HHHH	6
LEAV	HHHH	7	LLKE	HHHH	12	NAFN	HHHH	4
LEDF	HHHH	4	LLKK	HHHH	8	NAIV	HHHH	5
LEEA	HHHH	10	LLKQ	HHHH	7	NAVA	HHHH	6
LEEI	HHHH	6	LLLA	HHHH	8	NAVK	HHHH	7
LEET	HHHH	5	LLLE	HHHH	10	NEAL	HHHH	6
LEEV	HHHH	7	LLNE	HHHH	4	NEVL	HHHH	4

Последовательность	Разметка DSSP	N	Последовательность	Разметка DSSP	N	Последовательность	Разметка DSSP	N
NLLR	HHHH	4	QNAI	HHHH	4	RILN	HHHH	5
NRLV	HHHH	4	QNLL	HHHH	5	RKAL	HHHH	5
PAAL	HHHH	5	QNNL	HHHH	4	RKEL	HHHH	4
PAYL	HHHH	7	QQAA	HHHH	6	RKLL	HHHH	6
PDVI	HHHH	4	QQLA	HHHH	6	RLAE	HHHH	5
PEEA	HHHH	4	QQLI	HHHH	4	RLAK	HHHH	6
PEEI	HHHH	4	QQLL	HHHH	7	RLAQ	HHHH	6
PEEL	HHHH	11	QRAK	HHHH	4	RLEE	HHHH	6
PLLL	HHHH	6	QREL	HHHH	6	RLER	HHHH	4
PLVQ	HHHH	4	QRLL	HHHH	6	RLFЕ	HHHH	4
QAAE	HHHH	7	QVAA	HHHH	5	RLGE	HHHH	4
QAAI	HHHH	5	QVAD	HHHH	5	RLKQ	HHHH	4
QAAT	HHHH	5	RAAA	HHHH	8	RLLE	HHHH	5
QAER	HHHH	6	RAAE	HHHH	7	RLLK	HHHH	5
QAIA	HHHH	4	RAAI	HHHH	5	RLRA	HHHH	7
QAIE	HHHH	4	RAAL	HHHH	5	RLRE	HHHH	5
QAIK	HHHH	5	RAAV	HHHH	6	RLRR	HHHH	5
QAKE	HHHH	4	RAEL	HHHH	6	RQAL	HHHH	5
QALA	HHHH	6	RAER	HHHH	6	RRAA	HHHH	4
QALL	HHHH	8	RAIE	HHHH	10	RTLL	HHHH	8
QARA	HHHH	5	RAKA	HHHH	6	RVAQ	HHHH	5
QAVS	HHHH	4	RALA	HHHH	8	SAAL	HHHH	8
QDLL	HHHH	5	RALE	HHHH	5	SDAL	HHHH	5
QEAL	HHHH	7	RALK	HHHH	6	SDAV	HHHH	4
QELA	HHHH	4	RALM	HHHH	4	SDVL	HHHH	5
QELL	HHHH	6	RARA	HHHH	5	SEFL	HHHH	4
QELV	HHHH	5	RAVE	HHHH	7	SELK	HHHH	5
QETL	HHHH	4	RDAV	HHHH	5	SELL	HHHH	5
QEVЛ	HHHH	5	RDEL	HHHH	4	SEQL	HHHH	4
QIAA	HHHH	7	RDFL	HHHH	4	SIAK	HHHH	4
QIAE	HHHH	5	REAA	HHHH	4	SIEK	HHHH	5
QKAA	HHHH	6	REAL	HHHH	6	SLAI	HHHH	5
QKAL	HHHH	5	REAV	HHHH	5	SRAV	HHHH	5
QKEI	HHHH	4	REEA	HHHH	8	SRLA	HHHH	6
QKLI	HHHH	4	REEL	HHHH	5	STAI	HHHH	4
QLAL	HHHH	4	REKL	HHHH	5	STLQ	HHHH	4
QLAQ	HHHH	7	RELA	HHHH	12	SVLE	HHHH	6
QLID	HHHH	4	RELE	HHHH	5	SYAL	HHHH	4
QLLA	HHHH	6	RELI	HHHH	5	TAAQ	HHHH	4
QLLE	HHHH	8	RELK	HHHH	5	TAAV	HHHH	5
QLLL	HHHH	5	RELL	HHHH	6	TAFЛ	HHHH	6
QLQA	HHHH	6	RELS	HHHH	4	TAIA	HHHH	7
QLRA	HHHH	5	REQI	HHHH	5	TALA	HHHH	8
QLRE	HHHH	5	RERI	HHHH	4	TALE	HHHH	4
QLVK	HHHH	4	RGIA	HHHH	6	TALK	HHHH	5

Окончание

Последовательность	Разметка DSSP	<i>N</i>	Последовательность	Разметка DSSP	<i>N</i>	Последовательность	Разметка DSSP	<i>N</i>
TALS	HHHH	6	VDEA	HHHH	5	VLKK	HHHH	4
TARE	HHHH	7	VDLL	HHHH	7	VLKL	HHHH	4
TEAA	HHHH	6	VDNI	HHHH	5	VLLA	HHHH	7
TELA	HHHH	5	VDQF	HHHH	4	VLRK	HHHH	5
TELL	HHHH	6	VEAA	HHHH	6	VLVR	HHHH	4
TKAF	HHHH	4	VEAF	HHHH	5	VNAA	HHHH	6
TLAL	HHHH	4	VEAI	HHHH	4	VNEA	HHHH	6
TLKA	HHHH	7	VEAL	HHHH	8	VQDL	HHHH	6
TLKE	HHHH	6	VEEA	HHHH	5	VQEL	HHHH	9
TLLA	HHHH	6	VEEI	HHHH	7	VQQQ	HHHH	6
TLLE	HHHH	6	VEEL	HHHH	10	VRAA	HHHH	5
TLRE	HHHH	4	VEEV	HHHH	6	VRAL	HHHH	6
TLRK	HHHH	7	VEFA	HHHH	5	VRQL	HHHH	5
TSLI	HHHH	5	VEKA	HHHH	7	VSIL	HHHH	4
TVRE	HHHH	4	VELL	HHHH	5	VSLL	HHHH	6
TVWG	HHHH	4	VERA	HHHH	7	VTLA	HHHH	7
VAAL	HHHH	10	VERV	HHHH	4	VVQA	HHHH	5
VAAM	HHHH	4	VESL	HHHH	6	VVRA	HHHH	4
VAAQ	HHHH	5	VEVL	HHHH	6	VVSL	HHHH	4
VAAV	HHHH	5	VGEA	HHHH	4	VYRL	HHHH	4
VADA	HHHH	7	VIDF	HHHH	4	YAAA	HHHH	4
VADV	HHHH	6	VKDA	HHHH	5	YADI	HHHH	6
VAEA	HHHH	7	VKEA	HHHH	12	YALA	HHHH	5
VAEE	HHHH	5	VKKA	HHHH	9	YALL	HHHH	5
VAEL	HHHH	7	VKLI	HHHH	5	YEAA	HHHH	5
VAIQ	HHHH	4	VKLV	HHHH	9	YEAL	HHHH	5
VAKA	HHHH	6	VKQE	HHHH	4	YETL	HHHH	5
VAQL	HHHH	7	VLAK	HHHH	7	YLAA	HHHH	5
VASA	HHHH	4	VLAQ	HHHH	5	YLAD	HHHH	5
VATA	HHHH	5	VLEA	HHHH	9	YLEA	HHHH	6
VDAI	HHHH	5	VLEK	HHHH	5	YNQL	HHHH	4
VDAV	HHHH	6	VLER	HHHH	5	YVAA	HHHH	4

Однако, возможно, корректнее сопоставлять гидрофобные свойства конформационно-стабильных тетрапептидов не с такими же характеристиками тетрапептидов, скомбинированных за счет случайным образом расположенных в них радикалов аминокислот, а с гидрофобными свойствами реально наблюдаемых тетрапептидов из исследованного набора белков. Были рассмотрены выборки, состоящие из наблюдаемых α -спиральных тетрапептидов, а также всех возможных тетрапептидов. Результаты представлены в табл. 7. Видно, что уровень дисперсии гидрофобности реальных тетрапеп-

тидов, как только α -спиральных, так и вообще всех наблюдаемых, несколько меньше случайных, но все же превышает таковые величины у конформационно-стабильных фрагментов.

По результатам этой части исследования можно говорить, что конформационно-стабильные олигопептиды, полученные нами, чаще склонны проявлять отсутствие ярко выраженной гидрофобности/гидрофильности, а точнее «нейтральность» в этом плане, в сравнении с набором всех наблюдаемых последовательностей и с выборками случайных последовательностей, генерируемых согласно наблюдаемым

Таблица 3. Конформационно-стабильные тетрапептиды, хотя бы часть конформаций которых не является α -спиральной. N – число структур тетрапептидов с указанной последовательностью; M – число структур тетрапептидов с указанной разметкой. Условные обозначения DSSP: H – α -спираль, T – поворот, G – спираль 3_{10} , S – изгиб, E – β -тяж

Последовательность	N	Разметка DSSP	M	Последовательность	N	Разметка DSSP	M
AAAE	10	HHHT	1	ETLA	6	HHHG	1
		HHHH	9			HHHH	5
AAAK	7	HHHT	1	LEFL	10	HHHG	1
		HHHH	6			HHHH	9
AAAL	18	HHHT	1	LVEA	6	HHHG	1
		HHHH	17			HHHH	5
AIAA	17	HHHT	3	NQLA	5	HHHG	1
		HHHH	14			HHHH	4
ALEA	12	HHHT	1	QVKA	5	HHHG	1
		HHHH	11			HHHH	4
ALKL	7	HHHT	1	ALAA	23	HHHT	1
		HHHH	6			HHHG	1
EALK	17	HHHT	1	QYPD	4	HHHH	21
		HHHH	16			H_TT	2
ELAK	12	HHHT	1			H_GG	1
		HHHH	1			HSTT	1
EVLA	4	HHHT	1	SEEE	5	_HHH	4
		HHHH	3			SHHH	1
GVLA	10	HHHT	1	SPAQ	4	_HHH	3
		HHHH	9			SHHH	1
KALK	9	HHHT	1	TAEE	6	_HHH	6
		HHHH	8			_HHH	11
LAAA	15	HHHT	1	SRED	4	_HHH	4
		HHHH	14			_HHH	5
LAEA	7	HHHT	1	SDAE	4	_HHH	4
		HHHH	6			_HHH	5
LDAA	10	HHHT	1	NNQI	4	SS_	4
		HHHH	9			PGPP	5
LKDL	8	GGGT	1	PPGP	5	_____	5
		HHHH	7			EVNG	4
LLEA	11	HHHT	1	GPPG	5	_____	5
		HHHH	10			ADGS	8
LLES	6	HHHT	1	DEKG	5	TTS_	8
		HHHH	5			_TTS	4
QLAA		HHHT	1	LPDG	8	ETT_	1
		HHHH	6			_TTS	8
AKSA	5	HHHG	1	EKYP	6	TT_H	1
		HHHH	4			HH_S	1
						HH_T	4

частотам встречаемости аминокислот, в том числе, и частотам встречаемости в самих списках стабильных фрагментов.

ЗАКЛЮЧЕНИЕ

Таким образом, анализ конформационно-стабильных тетрапептидов, встречающихся дос-

Таблица 4. Средние значения и стандартные отклонения гидрофобности/гидрофильности в случае Списков конформационно-стабильных (I), случайных в соответствии с частотами в α -спиральных областях (II) и случайных в соответствии с частотами всех наблюдаемых тетрапептидных последовательностей (III)

Шкала гидрофобности	(а)			(б)			(в)		
	I	II	III	I	II	III	I	II	III
Выборка									
Среднее	-16,7	-16,1	-16,7	-8,5	-8,7	-8,5	6,3	4,4	6,3
Стандартное отклонение	6,4	9,4	8,3	8,5	11,4	11,4	6,0	7,0	7,2

Таблица 5. Величины произведений значений гидрофобности/гидрофильности, определенных по шкале (б) гидрофобности аминокислотных остатков, на частоты встречаемости в списке конформационно-стабильных сегментов (Стаб.) и в α -спиральных сегментах белков (Альфа) и различия этих величин для двух шкал

Аминокислота	A	C	E	D	G	F	I	H	K	M
Стаб.	-0,4	0	0,52	0,17	0	-0,19	-0,7	0	0,39	-0,02
Альфа	-0,21	-0,08	0,32	0,23	0	-0,40	-0,62	-0,02	0,37	-0,16
Разность	-0,19	0,08	0,2	-0,06	0	0,21	-0,08	0,02	0,02	0,14
Аминокислота	L	N	Q	P	S	R	T	W	V	Y
Стаб.	-1,76	0,04	0,06	-0,02	0	0,37	-0,03	-0,01	-0,4	-0,06
Альфа	0,11	0,06	-0,08	0,01	0,33	-0,07	-0,2	-0,45	-0,45	-0,16
Разность	-1,87	-0,02	0,14	-0,03	-0,33	0,44	0,17	0,44	0,05	0,1

Таблица 6. Величины произведений значений гидрофобности/гидрофильности, определенных по шкале (в) гидрофобности аминокислотных остатков, на частоты встречаемости в списке конформационно-стабильных сегментов (Стаб.) и в α -спиральных сегментах белков (Альфа) и различия этих величин для двух шкал

Аминокислота	A	C	E	D	G	F	I	H	K	M
Стаб.	0,15	0	1,21	0,2	0	-0,09	-0,09	0,01	0,29	0
Альфа	0,08	-0,01	0,75	0,27	0	-0,19	-0,08	0,08	0,27	-0,02
Разность	0,07	0,01	0,46	-0,07	0	0,1	-0,01	-0,07	0,02	0,02
Аминокислота	L	N	Q	P	S	R	T	W	V	Y
Стаб.	-0,44	0,02	0,12	0,01	0,01	0,22	0,01	-0,01	0,02	-0,04
Альфа	-0,28	0,06	0,12	0,04	0,02	0,19	0,02	-0,12	0,02	-0,14
Разность	-0,16	-0,04	0	-0,03	-0,01	0,03	-0,01	0,11	0	0,1

Таблица 7. Средние значения и стандартные отклонения гидрофобности/гидрофильности в случае конформационно-стабильных (I), случайных в соответствии с частотами в α -спиральных областях (II) и случайных в соответствии с частотами всех наблюдаемых тетрапептидных последовательностей (III)

Шкала гидрофобности	(а)			(б)			(в)		
	I	II	III	I	II	III	I	II	III
Выборка									
Среднее	-16,7	-15,7	15,1	-8,5	-8,6	-8,5	6,3	4,2	3,7
Стандартное отклонение	6,4	8,4	8,9	8,5	10,0	10,2	6,1	6,6	6,3

таточно часто в банке PDBSelect, показал, что в большинстве случаев в них реализуется α -спиральная конформация. Для ~4% выборки характерны иные конформации. Оценка физико-химические свойства конформационно-стабильных пептидов из выборки по средней гидрофобности/гидрофильности тетрапептидов по-

зволяет сделать вывод о том, что для них свойственны нейтральные значения гидрофобности/гидрофильности при характерной малой величине дисперсии этой величины. Таким образом, выделенные ранее по признаку конформационной стабильности пептиды являются особой группой весьма близких по структуре

и физико-химическим свойствам объектов. Эти объекты удовлетворяют требованиям теории, разработанной нами ранее, к молекулярным единицам, осуществляющим эффективное специфическое узнавание на далеких (по молекулярным масштабам) расстояниях.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты № 14-04-90034-Бел_а и 15-04-99605).

СПИСОК ЛИТЕРАТУРЫ

1. C. Bystroff, K. T. Simons, K. F. Han, and D. Baker, *Curr. Opin. Biotechnol.* **7**, 417 (1996).
2. C. Hunter and S. Subramaniam, *Proteins* **50**, 572 (2003).
3. O. Sunder, I. Sommer, and T. Lengauer, *BMC Bioinformatics* **7**, 14 (2006). DOI:10.1186/1471-2105-7-14.
4. А. В. Батяновский и П. К. Власов, *Биофизика* **53**, 556 (2008).
5. А. В. Батяновский, Н. Г. Есипова и С. Э. Шноль, *Биофизика* **54**, 1137 (2009).
6. В. А. Намиот, А. В. Батяновский, И. В. Филатов и др., *Биофизика* **56**, 594 (2011).
7. V. A. Namiot, A. V. Batyanovskii, I. V. Filatov, et al., *Phys. Lett. A* **375**, 2911 (2011).
8. E. A. Kabat and T. T. Wu, *Proc. Natl. Acad. Sci. USA* **71**, 4217 (1974).
9. A. Figureau, M. A. Soto, and J. Toha, *Protein* **16**, 103 (2003).
10. P. K. Vlasov, A. V. Vlasova, V. G. Tumanyan, and N. G. Esipova, *Proteins* **61**, 763 (2005).
11. H. S. Chan, *Amino Acid Side-chain Hydrophobicity. eLS.* (2002). DOI: 10.1038/npg.els.0003005.
12. D. R. DeVido, J. G. Dorsey, H. S. Chan, and K. A. Dill, *J. Phys. Chem. B* **102**, 7272 (1998).
13. J.-L. Fauchère and V. Plika, *Eur. J. Med. Chem. – Chimie Thérapeutique* **18**, 369 (1983).
14. W. C. Wimley and S. H. White, *Nature Struct. Biol.* **3**, 842 (1996).
15. C. A. Andersen, A. G. Palmer, S. Brunak, and B. Rost, *Structure* **10**, 175 (2002).

Structural and Physicochemical Characteristics of Conformationally Stable α -Helical Oligopeptides

A.V. Batyanovskii*, **I.D. Volotovskiy***, **V.A. Namiot****, **I.V. Filatov*****, **I.A. Galkin******, **N.V. Gnuchev******, **V.G. Tumanyan******, and **N.G. Esipova******

**Institute of Biophysics and Cell Technology, National Academy of Sciences of Belarus, ul. Akademicheskaya 27, Minsk, 220072 Belarus*

***Institute of Nuclear Physics, Lomonosov Moscow State University, Leninskie Gory 1, Moscow, 119992 Russia*

****Moscow Institute of Physics and Technology, Institutskiy per. 9, Dolgoprudny, Moscow Region, 141700 Russia*

*****Engelhard Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, Moscow, 119991 Russia*

The analysis of conformationally stable (conformational conservative) tetrapeptides selected from protein structures deposited in PDBSelect data bank has been fulfilled. The subset contained 943 tetrapeptide amino acid sequences and there were merely five 3D protein segment representatives for each sequence. As a result, the conclusion has been drawn on the basis of DSSP annotation analysis that in the majority of cases (900 of 943) α -helical conformation is obvious. Different than α -helix, in particular, the left-handed polyproline II helical conformation was observed in 43 sequences. The physical and chemical properties of conformationally stable peptides taken from the appropriate sample were estimated by the average hydrophobicity/hydrophilicity of tetrapeptides. The results of calculations show that the "neutrality" towards hydrophobicity/hydrophilicity is representative of conformationally stable oligopeptides. It should be noted, that dispersion of hydrophobicity/hydrophilicity distribution is sufficiently lower than for the test subsets. Thus, the conformationally stable oligopeptides present a distinct group of local protein structures which are very close with respect to conformational and physicochemical properties. In accordance with our developed theory of specific long range interactions these peptides are the objects being quite useful for effective mutual molecular recognition.

Key words: conformationally stable protein segments, local protein structure, molecular recognition, specific long-range interactions