



УДК 577.150.2

## НОВЫЕ СТРУКТУРНЫЕ МОТИВЫ В $\alpha$ -СПИРАЛЬНЫХ БЕЛКАХ

© 1997 г. А. В. Ефимов

Институт белка РАН, 142292, Пущино Московской обл.

Поступила в редакцию 09.09.96 г.

Обнаружены и охарактеризованы четыре новых типа структурных мотивов в  $\alpha$ -спиральных белках. Один из них состоит из трех  $\alpha$ -спиралей и может быть представлен в виде комбинации из  $\alpha$ - $\alpha$ -уголка и L-образной структуры; при этом вторая  $\alpha$ -спираль является одновременно одной из двух спиралей  $\alpha$ - $\alpha$ -уголка и одной из двух спиралей L-структуры. Вторым структурным мотивом, названный ABCD-единицей, состоит из четырех следующих друг за другом по цепи  $\alpha$ -спиралей A, B, C и D, упакованных в пространстве так, что спирали B, C и D образуют левую суперспираль, а спираль A располагается между спиральями B и D приблизительно антипараллельно к ним. Третий структурный мотив состоит из двух  $\alpha$ -спиралей, связанных длинной перетяжкой. В этом мотиве  $\alpha$ -спирали упакованы приблизительно параллельно друг другу и вместе с длинной перетяжкой образуют в пространстве левую  $\alpha$ -L-суперспираль. Четвертым структурным мотивом является четверная  $\alpha$ -спираль, расположенная в пространстве таким образом, что общая укладка цепи напоминает греческую букву  $\phi$ ; поэтому этот мотив был назван  $\phi$ -мотивом. В работе проведен анализ различных разновидностей перечисленных структурных мотивов, приведены многочисленные примеры мотивов и их разновидностей в белках с известными пространственными структурами.

*Ключевые слова:* белки, пространственная структура;  $\alpha$ - $\alpha$ -уголок,  $\alpha$ - $\alpha$ -шпилька, L-структура, правая и левая суперспираль, стереохимический анализ.

В последнее время структурными мотивами принято считать пространственно организованные структурные единицы, образованные двумя, тремя и более соседними по цепи и связанными между собой  $\alpha$ -спиральями и/или  $\beta$ -тяжами, которые часто встречаются как в гомологичных, так и негомологичных белках или многократно повторяются в одном и том же белке. Среди множества возможных и наблюдаемых в белках структурных мотивов только небольшая часть имеет уникальные укладки полипептидной цепи в пространстве и существует лишь в одной из двух “зеркально-симметричных” форм (обзор таких мотивов представлен в работе [1]). Тот факт, что такие структурные мотивы широко распространены в различных белках [2–7], а также то, что многие небольшие белки и домены состоят по существу только из таких мотивов [8], указывают на то, что они являются достаточно устойчивыми образованиями и способны свернуться в свои уникальные структуры сами по себе. Анализ показывает, что если взять структурные мотивы с уникальными укладками цепи в качестве зародышей (или в качестве стартовых структур), то взаимное расположение остальных участков молекулы белка можно получить последовательной пристройкой других  $\alpha$ -спиралей и/или  $\beta$ -тяжей к соответствующему мотиву, принимая во внимание простые правила [4–9]. Таким образом, с одной стороны, структурные мотивы являются “готовыми

структурными блоками” или элементами третичной структуры белков, с другой – их можно рассматривать в качестве зародышей в процессах сворачивания белков или использовать в качестве стартовых структур при моделировании и предсказании пространственной структуры белков.

В настоящей работе рассматриваются несколько новых пространственных структурных мотивов, которые встречаются в  $\alpha$ -спиральных белках. Новые структурные мотивы были обнаружены в результате сравнительного стереохимического анализа известных пространственных структур различных белковых молекул, проведенного с учетом следующих моментов: 1) по определению, только повторяющаяся во многих белках или в пределах одной полипептидной цепи пространственная структурная единица может считаться структурным мотивом [1–10]; 2) структурные мотивы должны быть образованы соседними по цепи элементами вторичной структуры; 3) каждый тип структурного мотива характеризуется определенным числом и типом элементов вторичной структуры, из которых он состоит, а также определенным взаимным расположением этих элементов как вдоль цепи, так и в пространстве; 4) структурные мотивы одного типа, найденные в различных (как в гомологичных, так и в негомологичных) белках, могут иметь различные длины  $\alpha$ -спиралей и(или)  $\beta$ -участков, а также различные длины и конформации соединяющих их

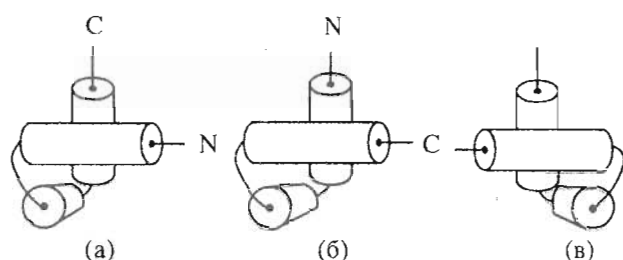


Рис. 1. Схематическое изображение структурных мотивов, образованных  $\alpha$ - $\alpha$ -углом и L-структурой: а) уголок-L-структура; б) L-структура-уголок; в) запрещенная "зеркально-симметричная" форма мотивов.  $\alpha$ -Спирали показаны в виде цилиндров; N и C – обозначения N- и C-концов.

петель, однако общий ход полипептидной цепи в пространстве (т.е. общая укладка полипептидной цепи) должен оставаться неизменным.

#### Комбинации из $\alpha$ - $\alpha$ -углока и L-образной структуры

$\alpha$ - $\alpha$ -Уголок – это структурный мотив, образованный двумя соседними по цепи  $\alpha$ -спиралями, которые располагаются в двух различных плоскостях или слоях и упакованы крестообразно [5]. В белках  $\alpha$ - $\alpha$ -углолки встречаются практически всегда только в одной из двух возможных "зеркально-симметричных" форм, в которой общий ход цепи описывает в пространстве почти виток левой суперспирали. В L-образных структурах две

$\alpha$ -спирали располагаются приблизительно перпендикулярно друг другу в одной плоскости, взаимодействуют только торцами и образуют структуру, напоминающую по форме букву L [10]. В отличие от  $\alpha$ - $\alpha$ -углока L-структуры могут быть правыми и левыми в зависимости от взаимного расположения  $\alpha$ -спиралей относительно гидрофобного ядра. Если смотреть на них со стороны гидрофобного ядра, то в левой L-структуре вторая  $\alpha$ -спираль располагается слева относительно первой, а в правой – справа.

На рис. 1 показаны структурные мотивы, которые представляют собой комбинации из  $\alpha$ - $\alpha$ -углока и L-структуры; при этом вторая спираль входит одновременно и в состав  $\alpha$ - $\alpha$ -углока, и в L-структуру. Если не принимать во внимание направление полипептидной цепи, то оба мотива имеют одинаковую укладку  $\alpha$ -спиралей в пространстве. На самом деле в одном из них первая и вторая  $\alpha$ -спирали образуют  $\alpha$ - $\alpha$ -уголок, а вторая и третья – L-структуру; это уголок-L-структура (рис. 1а). В другом первая и вторая  $\alpha$ -спирали образуют L-структуру, а вторая и третья –  $\alpha$ - $\alpha$ -уголок; это L-структура-уголок (рис. 1б). Поскольку  $\alpha$ - $\alpha$ -уголок существует только в одной из двух форм [5], то "зеркально-симметричная" форма рассматриваемых мотивов (рис. 1в) запрещена и в белках не встречается.

В табл. 1 приведены примеры таких структурных мотивов, которые найдены нами в белках с известными пространственными структурами.

Таблица 1. Структурные мотивы типа уголок-L-структура и L-структура-уголок в белках с известными структурами

Белок	Участки цепи, образующие спирали			Ссылка
	I	II	III	
Мотивы типа уголок-L-структура				
Дифтерийный токсин	98–107	113–119	120–127	[11]
Липаза	265–275	281–292	293–309	[12]
Трансгликозилаза	498–510	516–522	523–540	[13]
Глутамил-тРНК-синтетаза	426–440	446–455	456–468	[14]
Фактор транскрипции LEF-1	7–24	28–43	44–70	[15]
Эндохитиназа	121–133	139–146	146–160	[16]
Гексокиназа	259–270	274–280	281–300	[17]
A-Домен белка HMG1	14–28	38–51	52–74	[18]
Белок RecA	270–280	302–314	315–328	[19]
$\alpha$ -Субъединица РНК-полимеразы	Спираль H1	Спираль H2	Спираль H3	[20]
Белок SIV MA	Спираль H2	Спираль H3	Спираль H4	[21]
РНК-полимераза фага T7	Спираль Q	Спираль R	Спираль S	[22]
Мотивы типа L-структура-уголок				
Цитохром-c-пероксидаза	118–128	129–140	148–168	[23]
Рибонуклеаза Rh	71–80	80–91	97–112	[24]
Лизоцим куриного яйца	4–15	16–20	24–35	[25]
Глобулины	Спираль B	Спираль C	Спираль E	–



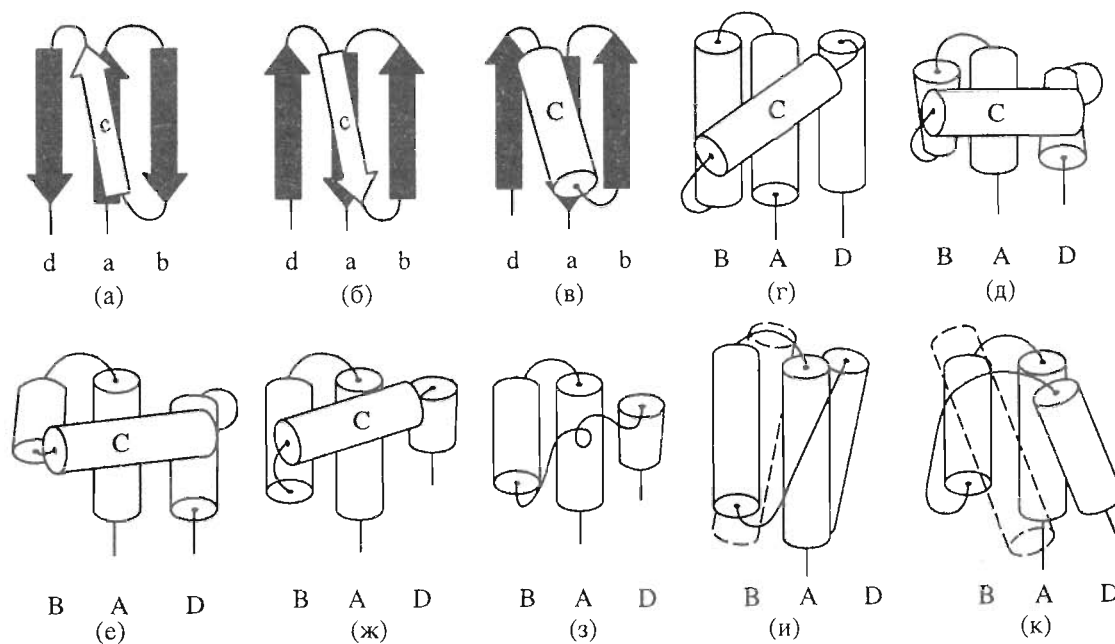


Рис. 2. Укладка цепи в ABCD-единице и ее вариантах (г–к) и сравнение с abcd-единицей (а–в).  $\beta$ -Тяжи показаны стрелками,  $\alpha$ -спирали – в виде цилиндров. Строчными буквами a, b, c и d обозначены  $\beta$ -тяжи, заглавными буквами A, B, C и D –  $\alpha$ -спирали. См. также текст.

Следует отметить, что большинство представленных белков не гомологичны друг другу. В различных белках структурные мотивы могут отличаться друг от друга размерами  $\alpha$ -спиралей и связывающих их петель, однако общая укладка цепи во всех случаях по существу одна и та же. Структуры фактора транскрипции LEF-1 и А-домена белка HMG1 состоят практически полностью из структурных мотивов типа уголок–L-структура и небольших “хвостов”. Это свидетельствует о том, что такие структурные мотивы достаточно стабильны и способны самостоятельно свернуться в свои уникальные структуры. Как установлено ранее [5, 10],  $\alpha$ - $\alpha$ -уголки и L-структуры с короткими перетяжками между  $\alpha$ -спиралями имеют строго определенное чередование вдоль цепи гидрофобных, гидрофильных, глициновых и/или пролиновых остатков, и их локализация в белках может быть предсказана. Это означает, что составленные из них трехспиральные структурные мотивы могут быть также локализованы в белках.

#### ABCD-мотив и его разновидности

На рис. 2 представлена укладка полипептидной цепи, разновидности которой часто встречаются в  $\alpha$ -спиральных белках и доменах.  $\alpha$ -Спиральный вариант укладки (рис. 2г) состоит из четырех последовательно расположенных вдоль цепи  $\alpha$ -спиралей A, B, C и D, упакованных в пространстве так, что спирали B, C и D образуют левую суперспираль, а спираль A располагает

между спиралями B и D приблизительно антипараллельно к ним. Этот структурный мотив назван здесь ABCD-единицей, поскольку его топология напоминает топологию abcd-единицы [4], образованной четырьмя  $\beta$ -участками (см. рис. 2а, 2б). Существенное отличие состоит в том, что  $\beta$ -участки b, c и d в abcd-единице образуют правую суперспираль, а спирали B, C и D – левую суперспираль в ABCD-единице. В целом если не принимать во внимание конформацию участков и хиральность аминокислотных остатков, то общий ход полипептидной цепи в ABCD-единице является зеркальным отображением хода цепи в abcd-единице (ср., например, рис. 2а, 2б и 2г). ABCD-единицы, так же как и abcd-единицы (рис. 2а, 2б), могут иметь противоположные направления полипептидной цепи (т.е. N-конец может находиться как на спирали A, так и на спирали D, рис. 2г), сохраняя при этом общий ход цепи и взаимное расположение спиралей в пространстве. В связи с этим буквы A, B, C и D обозначают положение спиралей в пространственной, а не в первичной структуре.

В белках ABCD-мотив встречается в нескольких вариантах (см. рис. 2г–2к и табл. 2). Семейства гомологичных белков представлены в табл. 2 одним или двумя примерами. Например, гомологичными белками являются утероглобин и ССРВР (сейчас известны пространственные структуры еще трех представителей этого семейства); парвальбумины крысы – один из представителей кальцийсвязывающих белков (известны структуры

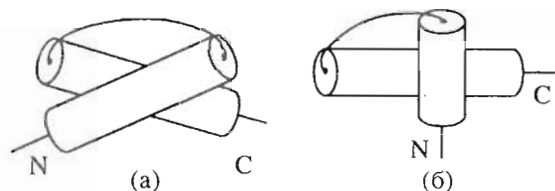


Рис. 3. Схематическое изображение  $\alpha$ -*l*- $\alpha$ -мотива (а) и  $\alpha$ - $\alpha$ -уголка с длинной перетяжкой (б).

около 20 этих белков); А-цепь рибина, шигатоксин и момордин – также гомологи. Как видно, общая укладка цепи во всех случаях по существу одинакова, хотя длины  $\alpha$ -спиралей, длины петель и взаимная ориентация спиралей (т.е. углы  $\Omega$  между осями спиралей) изменяются в довольно широких пределах, а вместо спирали С может быть длинный нерегулярный участок (рис. 2з, 2и, 2к). Отметим, что в *abcd*-единице участок “с” также может иметь нерегулярную или  $\alpha$ -спиральную (рис. 2в) конформацию [4, 8, 9].

Общая укладка цепи в *ABCD*-единице и ее хиральность определяется, по-видимому, тем, что в состав этого структурного мотива входят один или два  $\alpha$ - $\alpha$ -уголка. Как отмечалось выше, полипептидная цепь в  $\alpha$ - $\alpha$ -уголке описывает почти виток левой суперспирали, что и определяет образование левой суперспирали из  $\alpha$ -спиралей В, С и

Д. Для этого достаточно, чтобы пара спиралей В и С образовала  $\alpha$ - $\alpha$ -уголок, а пара С и D – V- или L-образную структуру или, наоборот, пара В и С образовала V- или L-образную структуру, а пара С и D –  $\alpha$ - $\alpha$ -уголок. Понятно, что если обе пары спиралей ВС и CD образуют  $\alpha$ - $\alpha$ -уголки, то это также приведет к формированию левой суперспирали *VCD*. Почему суперспираль *VcD* (где *c* – нерегулярный или вытянутый участок цепи) также левая, до конца неясно. Можно предположить, что в процессе сворачивания участок “с” был сначала в виде  $\alpha$ -спирали, которая затем преобразовалась в нерегулярную структуру. Однако более вероятным представляется то, что такого типа левые суперспирали являются наиболее стабильными формами упаковки  $\alpha$ -спиралей, связанных длинными петлями (см. следующий раздел статьи).

### $\alpha$ -*l*- $\alpha$ -Мотивы

Две соседние по цепи  $\alpha$ -спирали, связанные короткими или средними по длине перетяжками, могут образовывать  $\alpha$ - $\alpha$ -шпильки,  $\alpha$ - $\alpha$ -уголки, L-структуры и V-образные структуры. Если же длина перетяжки сопоставима с длинами связанных ею  $\alpha$ -спиралей или приблизительно равна длине одной из них, то возможно образование суперспирали типа  $\alpha$ -спираль–петля– $\alpha$ -спираль (рис. 3а). Это не означает, что связанные такими

Таблица 2. *ABCD*-мотивы в белках с известными структурами

Белок	Участки цепи, образующие спирали				Аналог на рис. 2	Ссылка
	А	В	С	Д		
Утероглобин	3–16	17–29	31–47	49–66	г	[26]
ССРВР	3–16	17–29	31–47	49–66	г	[27]
Парвальбумин крысы	7–18	25–34	39–51	59–71	г	[28]
Трансглюкозилаза	581–603	571–579	555–566	544–553	г	[13]
Термолизин	301–314	280–297	260–274	233–246	г	[29]
Эластаза (РАЕ)	282–294	264–280	243–259	221–239	г	[30]
Glu-тРНК-синтетаза	391–404	208–424	426–440	455–468	г	[14]
Цитратсинтаза	167–193	136–160	103–119	88–99	г	[31]
Редуктаза R1	76–95	54–69	37–49	19–31	г	[32]
Taq-полимераза	688–700	675–683	656–670	638–648	д	[33]
Пероксидаза (АРР)	96–111	78–88	44–60	20–39	е	[34]
Рецептор RAR- $\gamma$	266–277	296–301	308–320	325–336	е	[35]
Белок р17	73–88	55–69	33–45	22–26	ж	[36]
Кислая фосфатаза	197–215	170–186	153–163	143–149	ж	[37]
Рибин, А-цепь	161–178	142–156	132–136	122–130	з	[38]
Шигатоксин	152–170	132–144	–	114–121	з	[39]
Момордин	145–163	128–140	–	111–119	з	[40]
Интерлейкин-4	40–59	69–94	–	108–129	и	[41]
	69–94	40–59	–	4–20	к	
Фактор LIF	75–105	108–138	–	154–180	и	[42]
	108–138	75–105	–	21–49	к	
Ферритин	9–38	44–71	–	91–120	и	[43]

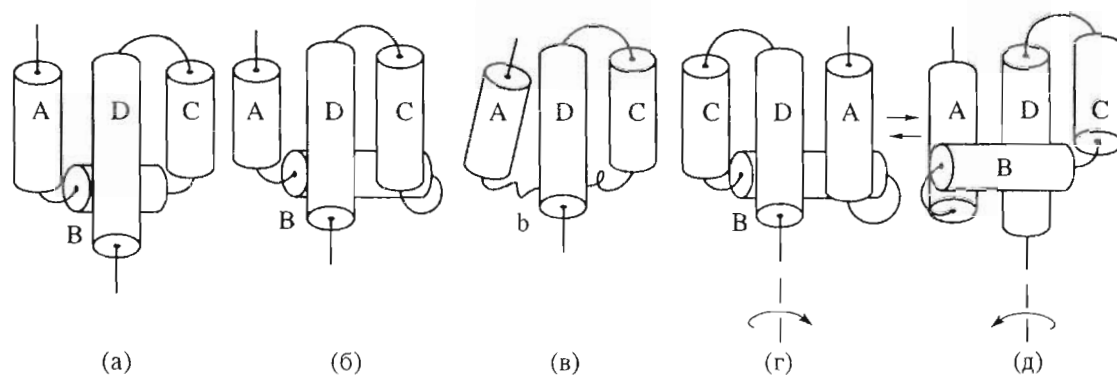


Рис. 4. Схематическое изображение правых (а-в) и левых (г-д)  $\phi$ -образных мотивов. В виде цилиндров показаны  $\alpha$ -спирали. Детальные пояснения даны в тексте.

длинными перетяжками  $\alpha$ -спирали всегда образуют  $\alpha$ - $l$ - $\alpha$ -мотивы –  $\alpha$ - $\alpha$ -шпильки и  $\alpha$ - $\alpha$ -уголки также могут иметь длинные перетяжки (см., например, рис. 3б). Закономерность состоит в том, что если  $\alpha$ - $l$ - $\alpha$ -мотив образуется, то образуется преимущественно левая суперспираль (рис. 3а). Как отмечалось выше,  $\alpha$ - $\alpha$ -уголок встречается в белках практически всегда в одной форме, в которой общий ход цепи описывает в пространстве почти виток левой суперспирали. Таким образом,  $\alpha$ - $l$ - $\alpha$ -мотив и  $\alpha$ - $\alpha$ -уголок с длинной перетяжкой очень похожи друг на друга и различаются в основном взаимной ориентацией  $\alpha$ -спиралей (ср. рис. 3а и 3б). Можно предположить, что  $\alpha$ - $\alpha$ -уголок является промежуточной структурой при образовании  $\alpha$ - $l$ - $\alpha$ -мотива, и это определяет отбор левой формы  $\alpha$ - $l$ - $\alpha$ -суперспирали. Важно отметить существенное отличие  $\alpha$ - $l$ - $\alpha$ -мотивов от  $\alpha$ - $\beta$ - $\alpha$ -единиц, входящих в состав  $\beta$ - $\alpha$ - $\beta$ - $\alpha$ -мотивов (или укладок по Россманну) в  $\alpha$ / $\beta$ -белках.  $\alpha$ - $\beta$ - $\alpha$ -Единицы, более подробную структуру которых можно представить как  $\alpha$ -спираль-петля- $\beta$ -тяж-петля- $\alpha$ -спираль, в  $\alpha$ / $\beta$ -белках всегда образуют правые суперспирали в отличие от левых суперспиралей, образуемых мотивами типа  $\alpha$ -спираль-петля- $\alpha$ -спираль.

В табл. 3 приведены примеры  $\alpha$ - $l$ - $\alpha$ -мотивов, обнаруженных в белках с известными структура-

ми. Во всех представленных случаях суперспирали – левые. К этому следует добавить левые суперспирали BcD в ABcD-единицах (см. рис. 2з, 2и, 2к и табл. 2), которые встречаются в А-цепи рибцина и его гомологах, ферритине и во всех цитохинах.

Необходимо отметить, что длинные перетяжки в  $\alpha$ - $l$ - $\alpha$ -мотивах и с-участки в ABcD-единицах могут быть отнесены к так называемым переходным перетяжкам. Согласно определению [48], “переходные перетяжки” (crossover connections or overhand connections) – это длинные перетяжки, которые переходят с одного торца  $\alpha$ -спирального пучка на другой по поверхности пучка. Однако важно подчеркнуть, что термины “левый” и “правый” пучок означают, что четырехспиральные пучки (включая пучки с “переходными перетяжками”) имеют различное взаимное расположение  $\alpha$ -спиралей, а не различные формы суперспиралей, которые остаются левыми.

#### $\phi$ -Образные мотивы

Во многих белках четыре следующие друг за другом по цепи  $\alpha$ -спирали А, В, С и D укладываются в структуру, которая по форме напоминает греческую букву  $\phi$  (рис. 4). Будем называть этот структурный мотив  $\phi$ -образным мотивом или

Таблица 3.  $\alpha$ - $l$ - $\alpha$ -Мотивы в белках с известными структурами

Белок	Участки цепи, образующие спирали		Ссылка
	$\alpha_1$	$\alpha_2$	
Глутамил-тРНК-синтаза	262–275	285–293	[14]
Дифтерийный токсин	98–107	113–120	[11]
Taq-полимераза	453–477	487–496	[33]
Барстар	12–25	33–43	[44]
Резолваза RuvC	117–129	139–152	[45]
Zn <sup>2+</sup> -G-пептидаза	16–30	43–55	[46]
Лизил-тРНК-синтаза	347–458	467–478	[47]

просто  $\phi$ -мотивом.  $\phi$ -Мотивы могут иметь прямое и обратное направление полипептидной цепи, т.е. N-конец цепи может находиться как на А-, так и на D-спирали. В этой связи буквы А, В, С и D обозначают положение  $\alpha$ -спиралей в пространственной структуре, а не в аминокислотной последовательности. Спираль в  $\phi$ -мотивах обозначены в той же последовательности, в которой перо ручки описывает траекторию при написании буквы  $\phi$  (рис. 4).  $\phi$ -Мотивы встречаются в белках в двух “зеркально-симметричных” формах. В правых  $\phi$ -мотивах спираль В (рис. 4а, 4б) или нерегулярный участок b (рис. 4в) располагаются снизу от спирали D, а общий ход цепи может быть представлен в виде “сплющенной” правой суперспирали А-ВС-D. В левом  $\phi$ -мотиве (рис. 4д) спираль В располагается сверху относительно спирали D, и общий ход цепи может быть представлен в виде левой “сплющенной” суперспирали А-ВС-D. Для наглядности на рис. 4г представлено “зеркальное” отображение правых  $\phi$ -мотивов (рис. 4а, 4б), которое является левым  $\phi$ -мотивом (рис. 4д), что легко видеть при повороте этой структуры на  $180^\circ$  (как это показано стрелкой).

Возможны три варианта правых  $\phi$ -мотивов. В одном из них обе пары спиралей АВ и ВС образуют L-образные структуры (рис. 4а); в другом – пара АВ образует L-образную структуру, а пара ВС –  $\alpha$ - $\alpha$ -уголок (рис. 4б); в третьем варианте вместо спирали В представлен нерегулярный или

вытянутый участок b (рис. 4в). В левом  $\phi$ -мотиве (рис. 4д) пара АВ образует  $\alpha$ - $\alpha$ -уголок (а не пара ВС, как в варианте, показанном на рис. 4б), а пара ВС – L-образную структуру. По-видимому, левый  $\phi$ -мотив может быть только в том случае, если пара спиралей АВ образует  $\alpha$ - $\alpha$ -уголок. В других случаях образуются правые  $\phi$ -мотивы. Примеры рассмотренных выше вариантов  $\phi$ -мотивов, которые обнаружены в белках с известными пространственными структурами, представлены в табл. 4.

К настоящему времени стало очевидным, что, для того чтобы решить проблему структуры белков и понять, “как” предсказывать их пространственную структуру, сначала нужно выяснить, “что” предсказывать, т.е. необходимо знать все возможные и стереохимически разрешенные структуры, которые полипептидная цепь способна образовать в принципе. Это можно видеть на следующем примере. Допустим, одним из известных методов предсказано, что в белке имеются четыре  $\alpha$ -спирали, следующие друг за другом по цепи. До сих пор наиболее вероятной структурой, в которую эти  $\alpha$ -спирали могут быть упакованы, считалась бы одна из разновидностей четырехспирального пучка (это наиболее изученный и часто встречающийся в белках структурный мотив из четырех  $\alpha$ -спиралей). Теперь мы знаем, что четыре  $\alpha$ -спирали могут образовать также ABCD-единицу или

Таблица 4.  $\phi$ -Образные мотивы в белках с известными структурами

Белок	Участки цепи, образующие спирали				Аналог на рис. 4	Ссылка
	А	В	С	Д		
Редуктаза	33–47	56–65	66–96	101–130	а	[49]
	152–170	132–142	101–130	66–96	б	
	268–291	257–267	222–255	185–221	а	
Интерферон $\beta$	112–131	102–111	76–98	47–66	а	[50]
Гидроксилаза $\alpha$ -субъединица	180–193	Спираль 2	131–161	97–127	а	[51]
	63–89	–	97–127	131–161	в	
$\beta$ -субъединица	219–233	Спираль 3	167–199	132–164	б	[52]
	101–125	–	132–164	167–199	в	
Фосфолипаза С	44–56	66–70	85–104	105–124	а	[53]
	141–151	–	171–191	206–242	в	
Глутатионтрансфераза	184–192	172–182	164–171	147–163	а	[54]
Рибонуклеаза Rh	81–91	97–112	113–117	128–144	д	[24]
Лизин	80–96	52–76	44–51	13–38	д	[54]
Реакционный центр субъединица L	116–140	149–164	170–200	225–250	д	[55]
	142–167	176–190	197–223	259–284	д	
Лизоцим фага T4	2–12	59–80	81–91	92–106	д	[56]



$\phi$ -мотив, и должны анализировать все возможные укладки  $\alpha$ -спиралей.

Другой важный вопрос, который широко обсуждается в литературе: является ли наблюдаемое сходство между пространственными структурами различных (в том числе и неродственных) белков следствием эволюционной дивергенции, функциональной конвергенции, или это результат действия общих физико-химических законов при сворачивании полипептидных цепей (см., например, работу [57])? Многочисленные примеры одинаковой укладки полипептидных цепей белков, не имеющих никакой гомологии первичной структуры, а также белков с совершенно различными функциями [4–8, 57] говорят в пользу последнего утверждения. Более того, двухслойные  $\alpha/\beta$ - и  $\beta$ -белки, принадлежащие к пространственно различным структурным классам, также имеют очень похожие укладки полипептидных цепей, а возможные структуры белков этих классов могут быть рассчитаны на основании одинаковых правил и запретов [9]. Еще одним аргументом в пользу того, что физико-химические принципы лежат в основе структурного сходства белков, является тот факт, что полипептидные цепи разных белков способны свернуться по существу в одну и ту же пространственную укладку, независимо от направления цепи. В частности, ABCD-единицы и  $\phi$ -мотивы могут иметь направление цепи от А- к D-спирали или от D- к А-спирали. То же самое наблюдается в других структурных мотивах, а также в структурах более высокого порядка [4–9]. Очевидно, что, какими бы далекими родственниками белки не были и какой бы низкий уровень гомологии они не имели, они не могут изменить направление полипептидной цепи в ходе своего эволюционного развития. Следовательно, с точки зрения эволюционного развития невозможно объяснить одинаковую пространственную укладку полипептидных цепей, имеющих разные направления.

Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (проект № 95-04-11851а) и Международного научного фонда (ISF, гранты NKK000 и NKK300).

### СПИСОК ЛИТЕРАТУРЫ

1. Efimov A.V. // Structure. 1994. V. 2. P. 999–1002.
2. Rao S.T., Rossmann M.G. // J. Mol. Biol. 1973. V. 76. P. 241–256.
3. Sternberg M.J.E., Thornton J.M. // J. Mol. Biol. 1976. V. 105. P. 367–382.
4. Ефимов А.В. // Молекуляр. биология. 1982. Т. 16. С. 799–806.
5. Ефимов А.В. // Молекуляр. биология. 1984. Т. 18. С. 1524–1537.
6. Efimov A.V. // FEBS Lett. 1992. V. 298. P. 261–265.
7. Efimov A.V. // FEBS Lett. 1993. V. 334. P. 253–256.
8. Efimov A.V. // FEBS Lett. 1994. V. 355. P. 213–219.
9. Efimov A.V. // J. Mol. Biol. 1995. V. 245. P. 402–415.
10. Ефимов А.В. // Молекуляр. биология. 1992. Т. 26. С. 1370–1376.
11. Bennett M.J., Eisenberg D. // Protein Sci. 1994. V. 3. P. 1464–1475.
12. Schrag J.D., Cygler M. // J. Mol. Biol. 1993. V. 230. P. 575–591.
13. Thunnissen A.-M.W.H., Dijkstra A.J., Kalk K.H., Rozeboom H.J., Engel H., Keck W., Dijkstra B.W. // Nature. 1994. V. 367. P. 750–753.
14. Nureki O., Vassilyev D.G., Katayanagi K., Shimizu T., Sekine S., Kigawa T., Miyazawa T., Yokoyama S., Morikawa K. // Science. 1995. V. 267. P. 1958–1965.
15. Love J.J., Li X., Case D.A., Giese K., Grosschedl R., Wright P.E. // Nature. 1995. V. 376. P. 791–795.
16. Hari P.J., Pflugger H., Monzingo A.F., Hollis T., Robertus J.D. // J. Mol. Biol. 1995. V. 248. P. 402–413.
17. Bennett W.S., Jr., Steitz T.A. // Proc. Natl. Acad. Sci. USA. 1978. V. 75. P. 4848–4852.
18. Hardman C.H., Broadhurst R.W., Raine A.R.C., Grasser K.D., Thomas J.O., Laue E.D. // Biochemistry. 1995. V. 34. P. 16596–16607.
19. Story R.M., Bishop D.K., Kleckner N., Steitz T.A. // Science. 1993. V. 259. P. 1892–1896.
20. Jeon Y.H., Negishi T., Shirakawa M., Yamazaki T., Fujita N., Ishihama A., Kyogoku Y. // Science. 1995. V. 270. P. 1495–1497.
21. Rao Z., Belyaev A.S., Fry E., Roy P., Jones I.M., Stuart D.I. // Nature. 1995. V. 378. P. 743–747.
22. Sousa R., Chung Y.J., Rose J.P., Wang B.-C. // Nature. 1993. V. 364. P. 593–599.
23. Fülöp V., Ridout C.J., Greenwood C., Hajdu J. // Structure. 1995. V. 3. P. 1225–1233.
24. Kurihara H., Mitsui Y., Ohgi K., Irie M., Mizuno H., Nakamura K.T. // FEBS Lett. 1992. V. 306. P. 189–192.
25. Strynadka N.C.J., James M.N.G. // J. Mol. Biol. 1991. V. 220. P. 401–424.
26. Morize I., Surcouf E., Vaney M.C., Epelboin Y., Buehner M., Fridlansky F., Milgrom E., Mornon J.P. // J. Mol. Biol. 1987. V. 194. P. 725–739.
27. Umland T.C., Swaminathan S., Singh G., Warty V., Furey W., Pletcher J., Sax M. // Nature Struct. Biol. 1994. V. 1. P. 538–545.
28. McPhalen C.A., Sielecki A.R., Santarsiero B.D., James M.N.G. // J. Mol. Biol. 1994. V. 235. P. 718–732.
29. Holmes M., Matthews B.W. // J. Mol. Biol. 1982. V. 160. P. 623–639.
30. Thayer M.M., Flaherty K.M., McKay D.B. // J. Biol. Chem. 1991. V. 266. P. 2864–2871.
31. Remington S., Wiegand G., Huber R. // J. Mol. Biol. 1982. V. 158. P. 111–152.
32. Uhlin U., Eklund H. // Nature. 1994. V. 370. P. 533–539.
33. Kim Y., Eom S.H., Wang J., Lee D.-S., Suh S.W., Steitz T.A. // Nature. 1995. V. 376. P. 612–616.
34. Kunishima N., Fukuyama K., Matsubara H., Hatana-ka H., Shibano Y., Amachi T. // J. Mol. Biol. 1994. V. 235. P. 331–344.

35. Renaud J.-P., Rochel N., Ruff M., Vivat V., Chambon P., Gronemeyer H., Moras D. // *Nature*. 1995. V. 378. P. 681–689.
36. Matthews S., Barlow P., Boyd J., Barton G., Russell R., Mills H., Cunningham M., Meyers N., Burns N., Clark N., Kingsman S., Kingsman A., Campbell I. // *Nature*. 1994. V. 370. P. 666–668.
37. Schneider G., Lindqvist Y., Vihko P. // *EMBO J*. 1993. V. 12. P. 2609–2615.
38. Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J., Pauptit R.A. // *J. Mol. Biol.* 1994. V. 244. P. 410–422.
39. Fraser M.E., Chernai M.M., Kozlov Y.V., James M.N.G. // *Nature Struct. Biol.* 1994. V. 1. P. 59–64.
40. Husain J., Tickle I.J., Wood S.P. // *FEBS Lett.* 1994. V. 342. P. 154–158.
41. Wlodaver A., Pavlovsky A., Gustchina A. // *FEBS Lett.* 1992. V. 309. P. 59–64.
42. Robinson R.C., Grey L.M., Staunton D., Vankelecom H., Vernallis A.B., Moreau J.-F., Stuart D.I., Heath J.K., Jones E.Y. // *Cell*. 1994. V. 77. P. 1101–1116.
43. Trikha J., Theil E.C., Allewell N.M. // *J. Mol. Biol.* 1995. V. 248. P. 949–967.
44. Guillet V., Laphorn A., Hartley R.W., Mauguen Y. // *Structure*. 1993. V. 1. P. 165–176.
45. Ariyoshi M., Vassilyev D.G., Iwasaki H., Nakamura H., Shinagawa H., Morikawa K. // *Cell*. 1994. V. 78. P. 1063–1072.
46. Dideberg O., Charlier P., Dive G., Joris B., Frere J.M., Ghuysen J.M. // *Nature*. 1982. V. 299. P. 469–470.
47. Onesti S., Miller A.D., Brick P. // *Structure*. 1995. V. 3. P. 163–176.
48. Presnell S.R., Cohen F.E. // *Proc. Natl. Acad. Sci. USA*. 1989. V. 86. P. 6592–6596.
49. Nordlund P., Eklund H. // *J. Mol. Biol.* 1993. V. 232. P. 123–164.
50. Senda T., Saitoh S., Mitsui Y. // *J. Mol. Biol.* 1995. V. 253. P. 187–207.
51. Rosenzweig A.C., Frederick C.A., Lippard S.J., Nordlund P. // *Nature*. 1993. V. 366. P. 537–543.
52. Hough E., Hansen L.K., Birknes B., Jynge K., Hansen S., Hordvik A., Little C., Dodson E., Derewenda Z. // *Nature*. 1989. V. 338. P. 357–360.
53. Dirr H., Reinemer P., Huber R. // *J. Mol. Biol.* 1994. V. 243. P. 72–92.
54. Shaw A., McRee D.E., Vacquier V.D., Stout C.D. // *Science*. 1993. V. 262. P. 1864–1867.
55. Deisenhofer J., Epp O., Miki K., Huber R., Michel H. // *Nature*. 1985. V. 318. P. 618–624.
56. Remington S.J., Anderson W.F., Owen J., Ten Eyck L.F., Grainger C.T., Matthews B.W. // *J. Mol. Biol.* 1978. V. 118. P. 81–98.
57. Ptitsyn O.B., Finkelstein A.V. // *Quart. Rev. Biophys.* 1980. V. 13. P. 339–386.

## Novel Structural Motifs in $\alpha$ -Helical Proteins

A. V. Efimov

*Institute of Protein Research, Russian Academy of Sciences, Pushchino, 142292, Moscow oblast, Russia*

**Abstract**—Four novel structural motifs were found and characterized in  $\alpha$ -helical proteins. One of them consists of three  $\alpha$  helices and can be represented as a combination of an  $\alpha$ - $\alpha$  corner and an L-shaped structure. Its second  $\alpha$  helix is a part of both one of the two  $\alpha$  helices of the  $\alpha$ - $\alpha$  corner and one of the two L-structure helices. The second structural motif, named the ABCD unit, consists of four  $\alpha$  helices A, B, C, and D. These are consecutive in sequence and arranged in space in such a manner that the helices B, C, and D form a left-handed superhelix and the helix A is located between the helices B and D and is approximately antiparallel to them. The third motif,  $\alpha$  helix-loop- $\alpha$  helix, consists of two  $\alpha$  helices and a long connection. Its  $\alpha$  helices are arranged in an approximately parallel manner and together with the long connection, form a left-handed  $\alpha$ -l- $\alpha$  superhelix in space. The fourth structural motif considered is formed by four consecutive  $\alpha$  helices. It is named the  $\phi$  motif, because its overall shape is reminiscent of the Greek letter  $\phi$ . Various variants of these motifs are analyzed on numerous examples of proteins with the known spatial structures.

**Key words:** *proteins, spacial structure;  $\alpha$ - $\alpha$  corner;  $\alpha$ - $\alpha$  hairpin; L-structure; left- and right-handed superhelix; stereochemical analysis.*