

## П-модуль в ОВ-фолде: структура и особенности последовательностей

Бражников Е.В., Ефимов А.В.

*Институт белка РАН, Пущино, Московская область, Россия*

**Аннотация.** В настоящей работе проведен стереохимический анализ комбинаций П-модулей с S-образными  $\beta$ -листами на примере большой выборки негомологичных белков, содержащих ОВ-фолды. Из 16 суперсемейств ОВ-фолдов исследовано 265 представительных белков и доменов, из которых отобрано 70 негомологичных белков с разрешением не менее 2.5 Å. Обнаружено два типа П-модулей: в первом переходная петля содержит  $\alpha$ -спираль, во втором переходная петля имеет нерегулярную структуру. Вход белковой цепи во второй слой происходит через одинаковую дугу с конформацией  $\beta\beta\alpha_L\beta_R$ . Наиболее часто, в 85 % случаев, в конформации  $\alpha_L$  встречаются глицины. Определены частоты встречаемости П-модулей в зависимости от длины их переходных петель. Пространственные структуры П-модулей обоих типов хорошо совмещаются при наложении друг на друга. Структурное выравнивание последовательностей П-модулей позволило определить ключевые позиции, в которых находятся гидрофобные, гидрофильные и глициновые остатки.

**Ключевые слова:** ОВ-фолд,  $\beta$ -тяж,  $\alpha$ -спираль, структурный мотив, фолдинг, хиральность.

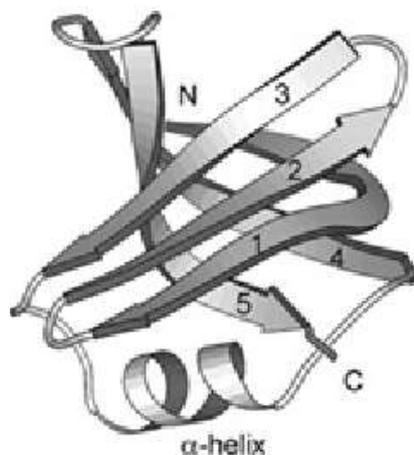
### 1. ВВЕДЕНИЕ

В глобулярных белках полипептидная цепь многократно сложена сама на себя и образует целый ряд часто повторяющихся блоков из двух и более соседних по цепи  $\alpha$ -спиралей и  $\beta$ -тяжей, которые называются супервторичными структурами или структурными мотивами [1–3]. Многие структурные мотивы, имеют компактные уникальные укладки цепей и определенную псевдохиральность. Такие структурные мотивы способны свернуться в свои уникальные структуры сами по себе и могут быть зародышами и/или готовыми структурными блоками при сворачивании белков [2, 3]. В некоторых структурных мотивах полипептидная цепь уложена в одной плоскости, как, например, в S-образных  $\beta$ -листах, или в двух плоскостях, как например, в abcd-единицах [1–3].

В предыдущей статье мы подробно изучили переход цепи из слоя в слой в SH3-подобных доменах и назвали переход  $\beta$ -тяж – петля –  $\beta$ -тяж L-модулем [4]. В настоящей статье аналогичное исследование проведено для ОВ-фолда, в котором тоже есть переход  $\beta$ -тяж – петля –  $\beta$ -тяж, но конформация петли совсем другая.  $\beta$ -Тяжи L-модуля расположены под углом  $\sim 90^\circ$  в разных  $\beta$ -слоях, упакованных ортогонально, и вместе с петлей образуют в пространстве полвитка правой суперспирали. Наиболее часто встречаются L-модули с  $\beta_{m\alpha\alpha\alpha\beta_n}$ - и  $\beta_{m\alpha\alpha\alpha\beta_n}$ -конформациями, где  $m$  и  $n$  - число остатков в первом и втором  $\beta$ -тяжах. L-модули каждого типа хорошо совмещаются при наложении друг на друга. S-образный  $\beta$ -лист состоит из трех  $\beta$ -тяжей, упакованных антипараллельно в одном слое так, что их расположение напоминает букву S. Структура П-модуля может быть представлена также, как и в L-модуле, в виде  $\beta$ -тяж–петля– $\beta$ -тяж, элементы которого образуют в пространстве укладку цепи, напоминающую русскую или греческую заглавную букву “П” или скобу (рис. 1).



модуль [14]. Следует отметить, что первый  $\beta$ -тяж сильно скручен и принадлежит разным  $\beta$ -слоям в  $\beta$ -бочонке. Таким образом, можно считать, что ОВ-фолд состоит из пяти или шести  $\beta$ -тяжей.



**Рис. 2.** Схематическое представление общей укладки полипептидной цепи в ОВ-фолде. Цифрами указаны номера  $\beta$ -тяжей. П-модуль образован тяжами 3 и 4 и перетяжкой между ними.

Переходы полипептидной цепи из слоя в слой в других структурных мотивах изучались нами раньше. Это  $\beta\alpha$ - и  $\alpha\beta$ -дуги в  $\alpha\beta$ Сd-единицах с прямым и обратным ходом полипептидной цепи [15], П-модули в комбинациях с  $\beta\alpha\beta$ -единицами [16],  $\phi$ - и  $\psi$ -мотивами [14, 17], L-модули в SH3-доменах [4]. П-модуль в ОВ-фолде следует за S-образным  $\beta$ -листом, поэтому исследуемая нами структура была названа СП-мотивом [14]. Структурные мотивы с уникальными укладками цепей и определенной хиральностью часто встречаются в глобулярных белках [1, 2, 18]. Было предположено, что они могут быть зародышами при сворачивании белков, или готовыми структурными блоками [2]. Поэтому изучение структур таких мотивов и структурных переходов в них имеет большое значение для теоретических и экспериментальных исследований сворачивания белков [3]. В настоящей работе подробно изучена структура негомологичных СП-мотивов в ОВ-фолдах. Показано, что все эти структуры имеют уникальные конформации цепи при входе во второй слой, а структуры негомологичных СП-модулей хорошо совмещаются при наложении и имеют одинаковое расположение в цепи ключевых гидрофобных, гидрофильных, глициновых и пролиновых остатков.

## 2. МАТЕРИАЛЫ И МЕТОДЫ

Белки для исследования, содержащие ОВ-фолды, отобраны из Банка белковых данных Protein Data Bank с помощью программы SCOP (<https://scop.mrc-lmb.cam.ac.uk>) [19]. Для анализа отбирали негомологичные белки или домены с разрешением не менее 2.5 Å. Проверку на гомологию проводили с помощью программы Blast 2 Sequences (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) [20]. Негомологичными считали белки, для которых показатель E-value, учитывающий как полную идентичность участков последовательностей, так и замены на сходные аминокислоты при выравнивании, был больше  $1e-04$ .

В Банке белковых данных мы нашли 265 PDB-файлов, содержащих ОВ-фолды, с разрешением выше 2.5 Å. Из них отобрали 70 негомологичных (табл. 1, 2) в соответствии с критериями программы Blast 2 Sequences [20]. Основной объект исследования – П-модуль – определяли визуально с помощью программы молекулярной графики RasMol [21]. При вычислении углов  $\phi$  и  $\psi$  использовали

программу MolMol [22]. Конформации аминокислотных остатков обозначали набором символов в соответствии с номенклатурой, предложенной ранее [23]. Чтобы наложить пространственные структуры друг на друга, использовали программу DeepView/Swiss-PDBViewer [24]. Выравнивание аминокислотных последовательностей П-модулей, основанное на их структурном сходстве, проводили вручную.

**Таблица 1.** Список негомологичных ОВ-фолдов с  $\alpha$ -спиралью в петле

№	PDB-код, (цепь)	Границы ОВ-фолдов	Участок П-модуля	Число остатков в петле	Длина $\alpha$ - спирали в петле	Разреше- ние, Å
01	2XSC (A)	1-69	27-55	15	15	2.05
02	1QB5 (D)	20-98	44-83	20	20	1.90
03	1PZK (D)	14-103	47-88	20	20	1.35
04	3SEB (A)	32-99	51-87	10	7	1.48
05	2Z8L (A)	25-91	50-77	15	6	1.65
06	4PQX (A)	33-106	59-86	13	6	2.39
07	5BMT (A)	28-105	72-102	13	11	1.50
08	3H43 (A)	91-149	115-138	12	5	2.10
09	2WFW (A)	141-220	165-182	8	8	1.60
10	2WFW (A)	80-137	99-125	12	5	1.60
11	1B9M (A)	127-260	157-182	13	6	1.75
12	1H9M (A)	2-59	34-59	13	7	1.65
13	1H9M (A)	74-141	106-130	13	7	1.65
14	1GUT (A)	2-68	32-58	13	8	1.50
15	1V43 (A)	247-367	265-289	13	5	2.20
16	1O7I (A)	14-87	49-72	11	3	1.20
17	1QVC (A)	4-111	53-84	15	11	2.20
18	1V1Q (A)	1-96	43-84	15	8	2.10
19	1XJV (A)	19-107	56-82	14	5	1.73
20	1XJV (A)	161-278	200-223	11	11	1.73
21	2B29 (A)	20-103	51-86	16	7	1.60
22	3H15 (A)	284-374	319-345	14	9	1.72
23	1B8A (A)	17-98	44-80	20	9	1.90
24	1CUK (A)	3-64	24-54	12	6	1.80
25	1FL0 (A)	156-232	187-210	14	6	1.50
26	1GD7 (A)	12-64	42-66	12	6	2.00
27	1NTG (A)	11-98	42-65	14	5	2.21
28	3ERS (X)	13-86	42-66	14	6	1.87
29	1HH2 (P)	137-198	157-182	14	3	2.10
30	1IZ6 (A)	71-134	96-124	17	10	2.00
31	1KHI (A)	103-173	130-157	13	13	1.78
32	1SMX (A)	39-118	82-103	9	5	1.80
33	1UEB (A)	64-126	90-115	14	9	1.65
34	2ES2 (A)	1-67	26-53	16	3	1.78
35	2JE6 (I)	66-143	98-130	23	6	1.60
36	3BZC (A)	654-722	679-712	20	6	2.27
37	3L00 (A)	54-127	86-109	13	8	2.35
38	2Q2T (A)	190-293	243-276	17	5	2.30
39	3KF8 (A)	80-181	122-156	18	6	2.40
40	3KF6 (A)	46-127	78-110	18	9	1.65
41	3KF8 (B)	22-99	58-86	17	6	2.40
42	5IHE (A)	263-337	293-323	13	8	2.50

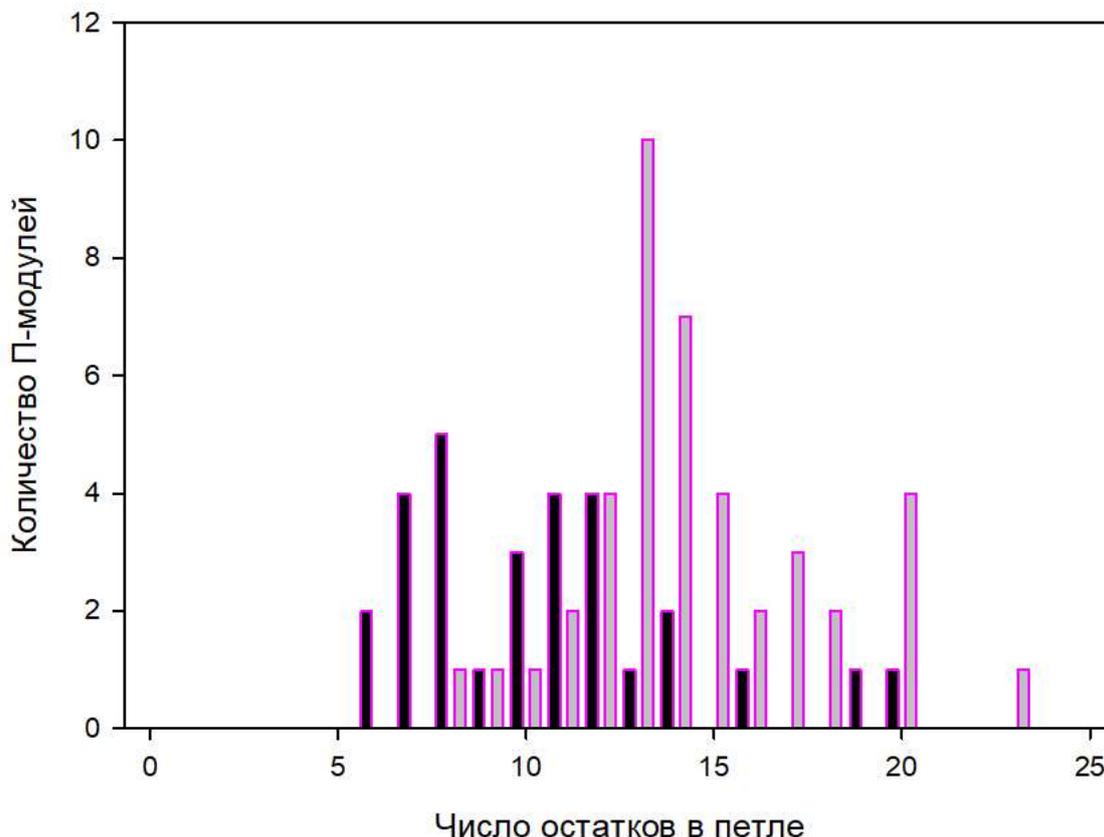
**Таблица 2.** Список негомологичных ОВ-фолдов с петлями, имеющими нерегулярную структуру

№	PDB-код, (цепь)	Границы ОВ-фолдов	Участок II-модулей	Число остатков в петле	Разрешение, Å
01	2PRD (A)	14–157	70–89	8	2.00
02	2QW7 (A)	1–96	40–59	8	2.40
03	1JB3 (A)	16–111	62–90	12	1.60
04	2IJ0 (B)	18–91	63–86	11	2.25
05	3QWN (H)	28–100	57–76	10	2.42
06	3K6O (A)	34–105	60–87	14	2.00
07	2Z1C (A)	2–53	24–45	11	1.80
08	2WG5 (A)	62–120	86–110	12	2.10
09	2AWN (B)	238–357	265–285	8	2.30
10	1QZG (A)	18–140	84–104	11	1.90
11	2PI2 (E)	9–88	50–73	8	2.00
12	1JB7 (A)	221–299	252–274	11	1.86
13	1UWV (A)	5–74	37–53	6	1.95
14	1N9W (B)	15–86	41–62	10	2.30
15	2JA9 (A)	70–150	95–123	19	2.20
16	2R7D (A)	404–461	428–447	6	1.80
17	1UEB (A)	127–184	158–174	7	1.65
18	2EIF (A)	74–132	99–122	10	1.80
19	1T9H (A)	2–64	25–51	16	1.60
20	3KBG (A)	116–162	140–157	7	1.75
21	1CKM (A)	241–296	264–284	7	2.50
22	1GVP (A)	1–87	43–63	9	1.60
23	1K3R (A)	104–163	127–144	7	2.30
24	2I5H (A)	16–78	53–71	8	1.74
25	1U5K (A)	4–79	32–61	20	2.00
26	4DA2 (B)	8–72	30–56	12	1.80
27	4ABN (A)	316–411	365–387	12	2.05
28	3KF6 (B)	520–584	542–569	14	1.65
29	3FLO (E)	293–367	329–353	13	2.50

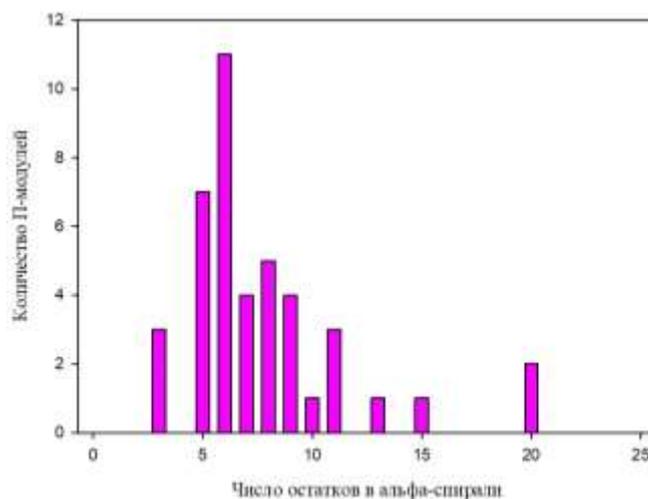
### 3. РЕЗУЛЬТАТЫ

В общем виде структура II-модуля может быть представлена как  $\beta$ -тяж–переходная петля– $\beta$ -тяж. В отличие от  $\beta$ -шпильки,  $\beta$ -тяжи II-модуля водородных связей между собой не образуют, но образуют их с дополнительными  $\beta$ -тяжами, которые расположены между  $\beta$ -тяжами в общем  $\beta$ -слое [14]. В ОВ-фолде переходная петля может состоять из  $\alpha$ -спирали, из неупорядоченной структуры (которую мы назвали дугой) или их комбинаций. Мы проанализировали все эти три случая. На рисунке 3 представлены зависимости частоты встречаемости II-модулей от длины их петель для двух типов.

Переходные петли с  $\alpha$ -спиралью встречаются чаще. Мы сравнили длины переходных петель II-модулей с  $\alpha$ -спиралью и без неё. В случае переходной петли со спиралью наибольшее число ОВ-фолдов имеют длину перетяжки 13–14 аминокислотных остатков. В ОВ-фолдах без  $\alpha$ -спирали длина петли сильно варьирует примерно от 6 до 14 аминокислотных остатков (рис. 3). В целом, петли со спиралью существенно длиннее петель в II-модулях без спиралей. Это закономерно, так как скрученная в спираль полипептидная цепь всегда компактнее и короче, чем развернутая. Исследование длин  $\alpha$ -спиралей показало, что чаще всего встречаются спирали длиной 6 и 5 аминокислотных остатков (рис. 4).



**Рис. 3.** Частоты встречаемости П-модулей с  $\alpha$ -спиралью в петле (светлые столбики) и без неё (темные столбики) в зависимости от длины петель.



**Рис. 4.** Зависимость количества  $\alpha$ -спиралей от их длины в П-модулях в ОВ-фолдах.

П-модули из различных негомологичных ОВ-фолдов обладают весьма схожими пространственными структурами и хорошо совмещаются, если их петли имеют одинаковую длину. На рисунке 5 представлено стерео-изображение наложения друг на друга дуг из укороченных П-модулей, содержащих в петле  $\alpha$ -спирали. Наложены шесть дуг с использованием программы DeepView/Swiss-PDB Viewer [20]. Выбраны следующие участки в структуре белков: 1B9M:A (171–178), 1GUT:A (48–55), 1H9M:A (120–127), 3L0O:A (97–104), 1UEB:A (103–110), 1QVC (70–77). Видно, что структуры хорошо совмещаются, хотя эти белки не являются гомологами.

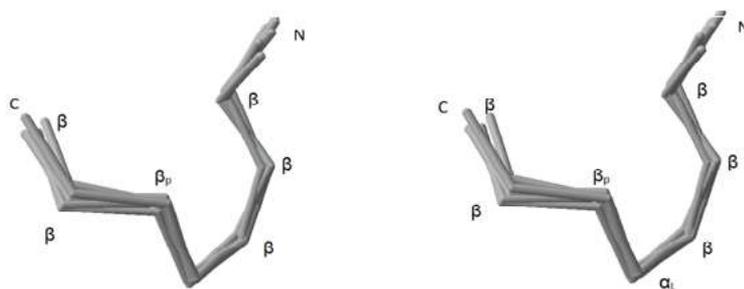


Рис. 5. Стереорезервирование дугообразных структур из Π-модулей. Конформации аминокислотных остатков показаны символами α, β, α<sub>L</sub>, β<sub>P</sub> [23].

№	PDB-код	ββββββββ	ααααααα-----	βββα <sub>L</sub> β <sub>P</sub>	βββββββ---	Цепь
01	2XSC (A)	-KELFTNR	WNLQSLLLSA-----	QITGM	TVT <sub>L</sub> IK----	27-53
02	1QB5 (D)	RISRAKDY	PDNVMTAEMRKIAMAA----	VLSGM	RVN <sub>L</sub> MC <sub>A</sub> ----	48-82
03	1PZK (D)	EVPGSQHI	DSQKKAIERMKD <sub>L</sub> TLR-----	YLTEA	KVEKLCVW--	51-88
04	3SEB (A)	NVRVEFKN	KDLAD-----	KYKDK	YVDVFGANY	63-89
05	2Z8L (A)	KFTRIQIF	GKDIERLKTR-----	KNPGL	DIFV-----	51-77
06	4PQX (A)	-----PSN	SQAWGGEK-----	FENGQ	RAFVIFN--	63-85
07	5BMT (A)	-----LYV	ATIANDPT-----	YRAGD	CVLANFTV--	56-79
08	3H43 (A)	-SFLVNVS	HFVNPDD-----	LAPGK	RVCL-----	115-137
09	2WFW (A)	-ERIVWLA	APL-----	AAVFA	DPE-----	165-182
10	2WFW (A)	-KMRLTCS	PNIDTDT-----	LALGQ	TVRL-----	103-125
11	1B9M (A)	-RLKVAIT	AQSGARLG-----	LDEG	EVLILL----	157-182
12	1H9M (A)	-KLA <sub>L</sub> AVVT	LESARSLQ-----	LAAGK	EVVAVV----	34-59
13	1H9M (A)	-EITSMVT	KEAVAELG-----	LKPGA	SASAVI----	106-131
14	1GUT (A)	-KITSIIS	LDSVEELG-----	VKEGA	ELTAV----	34-58
15	1V43 (A)	--FRIELP	Q MDLLK-----	DYVVK	TVLFGI----	265-289
16	1O7I (A)	---KLT <sub>L</sub> W	GKHAGS-----	IKEGQ	VVKIE-----	52-72
17	1QVC (A)	EWHRVVL <sub>F</sub>	GKLAEVASEY-----	LRKGS	QVYIEGQLR	53-84
18	1V1Q (A)	COMPVIVS	GHENQAI <sub>L</sub> THS-----	ITVGS	RITVQGFIS	48-79
19	1XJV (A)	KLTCLLFS	GNYEALPII-----	YKNGD	IVRFH----	56-82
20	1XJV (A)	--QDLVLE	GDLSHIHR-----	LQN	LTIDILVY--	199-223
21	2B29 (A)	TLSSFMLA	TQLNPLVEEQ--- --+---	LSSNC	VCQIHRFIV	52-84
22	3H15 (A)	--ISLFLF	GDVHKEHWK-----	TDQGT	VIGLL----	321-345
23	1B8A (A)	-GIVQITA	PKKKVDPELFK <sub>L</sub> IPK---+---	LRSED	VVAVEGVVN	44-79
24	1CUK (A)	VGYEVHMP	MTCFYEL-----	PEAGQ	EAI <sub>L</sub> VTFHFV	24-52
25	1FL0 (A)	---RTVVS	GLVNHVPLE-----	QMQR	MVILL----	187-210
26	1GD7 (A)	VKQSSAQI	TELYRPE-----	DLVGR	LVVCA----	42-66
27	1NTG (A)	---RTVVS	GLVQFVPKE-----	ELQDR	LVVV-----	42-65
28	3ERS (X)	--TLQTVT	SLVPYYSEE-----	ELM <sub>L</sub> GK	TVVV-----	42-66
29	1HH2 (P)	--LETRLP	KKEWIPGEE-----	IKAGD	LVKVYI----	157-182
30	1IZ6 (A)	---TFEVP	IDTGVAD <sub>L</sub> EIRDQ-----	LKEGI	NVEYWE <sub>T</sub> --	96-124
31	1KHI (A)	-PVIDQSS	LWNRLQKA-----	FESGR	GSRV <sub>L</sub> LVV--	130-157
32	1SMX (A)	YSAHGRPN	IKDV-----	LREGQ	EVIVQ----	82-103
33	1UEB (A)	---QFAVP	RSRVVGA <sub>L</sub> EF-----	FKEGM	TALGDMY--	90-115
34	2ES2 (A)	----VFVH	FSAIQGE <sub>L</sub> GFKT-----	LEEGQ	AVSFEIVE--	26-53
35	2JE6 (I)	----AYLP	ASNLLGRSINVGED <sub>L</sub> RRY--	LDVGD	YVIARI----	98-130
36	3BZC (A)	---DGLVH	ISALSEKFVKDPYEV-----	VKAGD	IVKVKMEV	679-712
37	3L00 (A)	---IYIS	PSQIRKFN-----	LNTGD	IISGVIR--	86-109
38	2Q2T (A)	FSIGTGFD	ADQRRDFWQ <sub>L</sub> NKE-----	SYIGK	MVKFKY----	245-275
39	3KF8 (A)	-KIRVRLS	QEQFKEVGLTLDK-----	KNYGK	IVELEGEI--	122-154
40	3KF6 (A)	--LRVFI	IQDDFSMSKRAIS-----	MSPGN	VVCVFGKI--	78-109
41	3KF8 (B)	----INIC	IFNLL <sub>L</sub> EACSSRV-----	LVPGT	LVNIDAFY--	58-86
42	5INE (A)	KVFLPKDS	EDYREAFK-----	VLPDA	VVAFKGF--	295-322

Рис. 6. Структурное выравнивание последовательностей Π-модулей в ОВ-фолдах, содержащих в петле α-спираль. Конформации остатков в столбцах показаны символами α, β, α<sub>L</sub>, β<sub>P</sub> в верхней строке. Подчеркнуты остатки, направленные в сторону гидрофобного ядра. В каждой строке слева показан PDB-код белка, а справа – участок цепи, образующий Π-модуль. Серым цветом выделены остатки, находящиеся в α<sub>L</sub> – конформации.

На рисунке 6 представлено структурное выравнивание аминокислотных последовательностей 42 Π-модулей, содержащих в петле α-спираль. Выравнивание

проводили вручную. Участки последовательностей первых  $\beta$ -тяжей выравнивали по последнему остатку каждого  $\beta$ -тяжа так, чтобы внутренние (направленные в гидрофобное ядро) и внешние (направленные наружу) позиции попадали в одни и те же столбцы соответственно. Участки  $\alpha$ -спиральных последовательностей выравнивали по первым остаткам  $\alpha$ -спиралей, которые одновременно являются и последними остатками первых  $\beta$ -тяжей. Таким образом, в каждом столбце множественного выравнивания расположены структурно-эквивалентные остатки, конформация которых показана буквами в верхней строке. Последний остаток в первом  $\beta$ -тяже, который одновременно и первый остаток  $\alpha$ -спирали, является гидрофильным, небольшого размера или пролином. Аналогичное выравнивание проведено для аминокислотных последовательностей П-модулей, не имеющих в петле  $\alpha$ -спираль. На рисунке 7 представлена укладка аминокислотных последовательностей в таких ОВ-фолдах. Структурное выравнивание было выполнено так же, как и на предыдущем рисунке. Отметим, что закономерности укладки полипептидной цепи аналогичны П-модулям, содержащих в петле  $\alpha$ -спираль, за исключением того, что здесь не требуется наличия гидрофобных остатков, как это требовалось для образования  $\alpha$ -спирали. Как и в случае П-модуля с  $\alpha$ -спиралью последний остаток первого тяжа в основном гидрофильный, а первый остаток петли преимущественно находится в  $\alpha$ -спиральной конформации.

№	PDB-код	$\beta\beta\beta\beta\beta\beta\beta\beta$	$\alpha$ -----	$\beta\beta\beta\alpha\beta\beta$	$\beta\beta\beta\beta\beta$ ----	Цепь
01	2PRD (A)	----DGLVLS	TYP-----	LLPGV	VVEVRV----	70-89
02	2QW7 (A)	----YEVAGD	VVG-----	AGLNE	WVLVAR----	40-59
03	1JB3 (A)	-KVVIGGFGD	PLICDNQ-----	VSTGD	TRIFFVNP--	62-90
04	2IJ0 (B)	----DLNTRK	TKKSQH-----	TSEGT	YIHFIQIS---	63-86
05	3QWN (H)	----LTVSKD	RTGSA-----	ISPNT	SRRV-----	57-76
06	3K60 (A)	-----KLWV	VQHIVPYRD-----	LKAGE	RIFGNYSFLE	60-87
07	2Z1C (A)	----VKREVR	LDLMPD-----	TKPGD	WVIVH-----	24-45
08	2WG5 (A)	---KFVVNTS	QYINEEE-----	LKPGA	RVALNQ----	86-110
09	2AWN (B)	---QWLPVE	SRD-----	VQVGA	NMSLGI----	265-285
10	1QZG (A)	-QIHLFSKQG	NDLPVI-----	KQVGQ	P-----	84-104
11	2PI2 (E)	-NGTIELMEP	LDE-----	EISGI	VEVVGRV---	50-73
12	1JB7 (A)	---VFYTLSL	KLKFPH-----	VRTGE	VVRIR-----	252-274
13	1UWV (A)	---KTLFIP	G-----	LLPQE	NAEVT-----	37-53
14	1N9W (B)	---GVVQVVT	GGLKL-----	PLPES	ALRVR-----	41-62
15	2JA9 (A)	-----VSL	YMAFPNASKNRPT--	LQVGD	LVYARV----	95-123
16	2R7D (A)	---FDVQVNT	P-----	AAPGT	ALQVQFA---	428-447
17	1UEB (A)	-----VVQVP	LF-----	VEPGE	VIKVD-----	158-174
18	2EIF (A)	---TLELPIP	EGIEG-----	LEPGG	EVEYIEA---	99-122
19	1T9H (A)	---KVIQCR	GRGIFRKNKIT----	PLVGD	YVVYQ-----	25-51
20	3KBG (A)	---TFITDD	KS-----	IKVGD	VLAVS-----	140-157
21	1CKM (A)	--NVPVGKLD	GY-----	YNKGS	IVECGF----	264-284
22	1GVP (A)	-VLVKITLDE	GQPA-----	YAPGL	YTV-----	43-63
23	1K3R (A)	-----LAL	CR-----	LTVNR	IMSFRV----	127-144
24	2I5H (A)	--LLEVSIRK	GKQ-----	PLVMD	RVY-----	53-71
25	1U5K (A)	--GKLKAIAR	GGVKGPLSSSLNLF--	HHVGV	QVY-----	32-61
26	4DA2 (B)	-RIERALITN	TGRLNEF-----	MIPGR	IGYCTP----	30-56
27	4ABN (A)	---YAVMVY	NVVQSWG-----	VLIGD	SVAIP-----	365-387
28	3KF6 (B)	----SSVTCD	VTVVLPDVS-----	IQKHE	WLNIVGRK--	542-569
29	3FLO (E)	-----RVRLD	LSQVNELS-----	FFLGQ	IVAFKGGK---	329-353

Рис. 7. Структурное выравнивание последовательностей П-модулей в ОВ-фолдах, не содержащих в петле  $\alpha$ -спираль. Обозначения такие же, как на рисунке 5.

Из рисунков 6 и 7 видно, что в обоих случаях вход во второй слой происходит через одинаковую дугу, имеющую конформацию  $\beta\beta\beta\alpha\beta\beta$ . Похожие дуги ранее наблюдались при переходе цепи из одного слоя в другой слой в 2-х слойных  $\beta$ -белках [25]. Отметим, что первый  $\beta$ -остаток в дуге погружен в гидрофобное ядро и должен быть

гидрофобным. Боковая цепь пятого остатка дуги (который является одновременно первым остатком второго  $\beta$ -тяжа) спрятана частично и может быть как гидрофобным, так и гидрофильным. Четвертый  $\alpha_L$ -остаток должен быть глицином или остатком с гибкой боковой цепью. Пролины во втором или третьем остатках облегчают образование дуги, но их присутствие не является обязательным. Детальный анализ  $\beta\beta\alpha_L\beta$ -дуги в П-мотивах (рис. 6, 7) показывает, что пятое положение (оно же первое положение второго тяжа в П-модуле) предпочитает гидрофильные остатки. Боковые цепи во втором и третьем положениях дуги в П-модулях доступны молекулам воды и поэтому преимущественно заняты гидрофильными остатками. Что касается  $\beta$ -тяжей в П-модуле, то внутренние положения предпочтительно заняты гидрофобными остатками, а внешние – гидрофильными, что хорошо согласуется с теорией [26].

Таким образом, необходимые условия образования П-модулей, содержащих  $\alpha$ -спираль в петле, состоят в том, что в определенных ключевых позициях их аминокислотных последовательностей должны находиться гидрофобные, гидрофильные остатки, а также глицины и пролины. По-видимому, сам по себе П-модуль нестабилен, однако в комбинациях с S-образными  $\beta$ -листами в ОВ-фолдах он образует замкнутые компактные структуры с повышенной кооперативностью и стабильностью и может быть зародышем или готовыми структурным блоком в процессе сворачивания белков.

#### 4. ЗАКЛЮЧЕНИЕ

В статье проведен стереохимический анализ П-модуля в ОВ-фолде в негомологичных белках. Мы показали, что существуют два типа модулей. Первый тип содержит в петле  $\alpha$ -спираль, второй тип её не содержит. П-модули разных типов имеют разные длины петель, соединяющих два слоя. В случае первого типа со спиралью они короче с геометрической точки зрения. Конформация остатков при входе и выходе из петли не зависят от того, содержится в петле спираль или нет. Вход в петлю в обоих случаях происходит без напряжений в цепи и не содержит глицинов, в основном это  $\beta\alpha$ -переход. Наоборот, выход из петли происходит через напряженную структуру, имеющую конформацию  $\beta\beta\alpha_L\beta$ . В этом случае в  $\alpha_L$ -конформации находятся, в основном, глицины. Распределение гидрофобных, гидрофильных и глициновых остатков в обоих типах П-модулей одинаково. Таким образом, можно предположить, что в белках суперсемейства ОВ-фолдов архитектура упаковки цепи в большой степени определяется П-модулями.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 20-04-00453).

Авторы не имеют финансовой заинтересованности в представленных материалах или методах. Конфликт интересов отсутствует.

#### СПИСОК ЛИТЕРАТУРЫ

1. Efimov A.V. Standard structures in proteins. *Prog. Biophys. Molec. Biol.* 1993. V. 60. P. 201–239. doi: [10.1016/0079-6107\(93\)90015-c](https://doi.org/10.1016/0079-6107(93)90015-c)
2. Efimov A.V. Favoured structural motifs in globular proteins. *Structure.* 1994. V. 2. P. 999–1002. doi: [10.1016/s0969-2126\(94\)00102-2](https://doi.org/10.1016/s0969-2126(94)00102-2)
3. Efimov A.V. Super-secondary structures and modeling of protein folds. *Methods Molecular Biology.* 2013. V. 932. P. 177–189. doi: [10.1007/978-1-62703-065-6\\_11](https://doi.org/10.1007/978-1-62703-065-6_11)
4. Каргатов А.М., Бражников Е.В., Ефимов А.В. L-модули в SH3-подобных фолдах: структура и особенности последовательностей. *Молекуляр. Биология.* 2018. V. 52. P. 1074–1081. doi: [10.1134/S0026898418060095](https://doi.org/10.1134/S0026898418060095)

5. Murzin A.G. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* 1993. V. 12. P. 861–867. doi: [10.1002/j.1460-2075.1993.tb05726.x](https://doi.org/10.1002/j.1460-2075.1993.tb05726.x)
6. Agrawal V., Kishan K.V. OB-fold: growing bigger with functional consistency. *Curr. Protein. Pept. Sci.* 2003. V. 4. P. 195–206. doi: [10.2174/1389203033487207](https://doi.org/10.2174/1389203033487207)
7. Agrawal V., Kishan K.V. Functional evolution of two subtly different (similar) folds. *BMC Struct. Biol.* 2001. V. 1. P. 5. doi: [10.1186/1472-6807-1-5](https://doi.org/10.1186/1472-6807-1-5)
8. Arcus V. OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr. Opin. Struct. Biol.* 2002 V. 12. P. 794–801. doi: [10.1016/s0959-440x\(02\)00392-5](https://doi.org/10.1016/s0959-440x(02)00392-5)
9. Bianco P.R. OB-fold Families of Genome Guardians: A Universal Theme Constructed From the Small  $\beta$ -barrel Building Block. *Front. Mol. Biosci.* 2022. V. 11. P. 784451. doi: [10.3389/fmolb.2022.784451](https://doi.org/10.3389/fmolb.2022.784451)
10. Nguyen D.D., Kim E.Y., Sang P.B., Chai W. Roles of OB-Fold Proteins in Replication Stress. *Front. Cell Dev. Biol.* 2020. V. 11. doi: [10.3389/fcell.2020.574466](https://doi.org/10.3389/fcell.2020.574466)
11. Bianco P.R. The mechanism of action of the SSB interactome reveals it is the first OB-fold family of genome guardians in prokaryotes. *Protein Sci.* 2021 V. 30. P. 1757–1775. doi: [10.1002/pro.4140](https://doi.org/10.1002/pro.4140)
12. Efimov A.V. Structure of coiled  $\beta$ - $\beta$ -hairpins and  $\beta$ - $\beta$ -corners. *FEBS Lett.* 1991. V. 284. P. 288–292. doi: [10.1016/0014-5793\(91\)80706-9](https://doi.org/10.1016/0014-5793(91)80706-9)
13. Boshkova E.A., Brazhnikov E.V., Efimov A.V. Relationship between structure and amino acid sequence of strongly twisted and coiled  $\beta$ -hairpins in globular proteins. *Mol. Biol. (Mosk.)*. 2016. V. 50. P. 777–782. doi: [10.7868/S0026898416050049](https://doi.org/10.7868/S0026898416050049)
14. Efimov A.V. Structural motifs in which  $\beta$ -strands are clipped together with the  $\Pi$ -like module. *Proteins*. 2017. V. 85 P. 1925–1930. doi: [10.1002/prot.25346](https://doi.org/10.1002/prot.25346)
15. Бражников Е.В., Каргатов А.М., Ефимов А.В. Структура петлевых участков в  $\beta$ - $\alpha$ - и  $\alpha$ - $\beta$ -дугах abCd-единиц в глобулярных белках. *Матем. биол. и биоинф.* 2016. V. 11. P. 159–169. doi: [10.17537/2016.11.159](https://doi.org/10.17537/2016.11.159)
16. Kargatov A.M., Efimov A.V. Unique combinations of  $\beta\alpha\beta$ -units and  $\Pi$ -like modules in proteins and features of their amino acid sequences. *Mol. Biol. (Mosk.)*. 2018. V. 1. P. 36–41. doi: [10.7868/S002689841801007X](https://doi.org/10.7868/S002689841801007X)
17. Efimov A.V. Structural trees for proteins containing  $\varphi$ -motifs. *Biochemistry (Mosc.)*. 2008. V. 3. P. 23–28. doi: [10.1134/s0006297908010033](https://doi.org/10.1134/s0006297908010033)
18. Sternberg M.J., Thornton J.M. On the conformation of proteins: the handedness of the beta-strand-alpha-helix-beta-strand unit. *J. Mol. Biol.* 1976. V. 105. P. 367–382. doi: [10.1016/0022-2836\(76\)90099-1](https://doi.org/10.1016/0022-2836(76)90099-1)
19. Murzin A.G., Brenner S.E., Hubbard T, Chothia C. SCOP: A Structural Classification of Proteins Database for the Investigation of Sequences and Structures. *J. Mol. Biol.* 1995. V. 247. P. 536–540. doi: [10.1006/jmbi.1995.0159](https://doi.org/10.1006/jmbi.1995.0159)
20. Tatusova T.A., Madden T.L. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.* 1999. V. 174. P. 247–250. doi: [10.1111/j.1574-6968.1999.tb13575.x](https://doi.org/10.1111/j.1574-6968.1999.tb13575.x)
21. Sayle R.A., Milner-White E.J. RASMOL: biomolecular graphics for all. *Trends Biochem. Sci.* 1995. V. 20. P. 374–376. doi: [10.1016/S0968-0004\(00\)89080-5](https://doi.org/10.1016/S0968-0004(00)89080-5)
22. Koradi R., Billeter M., Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 1996. V. 14. P. 51–55. doi: [10.1016/0263-7855\(96\)00009-4](https://doi.org/10.1016/0263-7855(96)00009-4)
23. Efimov A.V. Standard conformations of polypeptide chain in irregular regions of proteins. *Mol. Biol. (Mosk.)*. 1986. V. 20. P. 208–216.

24. Guex N., Peitsch M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 1997. V. 18. P. 2714–2723. doi: [10.1002/elps.1150181505](https://doi.org/10.1002/elps.1150181505)
25. Efimov AV. Pseudo-homology of protein standard structures formed by two consecutive  $\beta$ -strands. *FEBS Lett*. 1987. V. 224. P. 372–376. doi: [10.1016/0014-5793\(87\)80487-8](https://doi.org/10.1016/0014-5793(87)80487-8)
26. Lim V.I. Structural principles of the globular organization of protein chains. A stereochemical theory of globular protein secondary structure. *J. Mol. Biol.* 1974. V. 88. P. 857–872. doi: [10.1016/0022-2836\(74\)90404-5](https://doi.org/10.1016/0022-2836(74)90404-5)

Рукопись поступила в редакцию 29.09.2022, переработанный вариант поступил 24.11.2022.  
Дата опубликования 06.12.2022.

===== BIOINFORMATICS =====

## Structure and features of amino acid sequences of Π-modules in OB-folds

**Brazhnikov E.V., Efimov A.V.**

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

**Abstract.** Stereochemical analysis has been performed for Π-modules from the large set of non-homologous protein structures containing the OB-fold. That module consists of two  $\beta$ -strands connected by a loop and placed in different sheets in such a way which looks as Greek letter Π. Total 70 non-homologous proteins at resolution not less than 2.5Å have been selected for the analysis from 265 suitable structures belonging to sixteen OB-fold super families. We have disclosed two types of Π-modules: the first with the connecting loop containing  $\alpha$ -helix, and second one without helix. Entrance of protein chain into second  $\beta$ -sheet is carried out by the same arch with conformation  $\beta\beta\beta\alpha_L\beta_P$ . In most cases, 85 % of total,  $\alpha$ -positions are occupied by glycine residue, while at entrance in the loop such residues are absent. Occupancy frequency of Π-modules has been obtained in dependence on the loop length. Spatial pathway of structures of all modules are superimposed very well. Structural alignment of amino acid for Π-module sequences allows us to determine the key positions of the hydrophobic, hydrophilic, and glycine residues.

**Key words:** OB-fold,  $\beta$ -strand,  $\alpha$ -helix, structural motif, folding, handedness.