ISSN 0026-8933, Molecular Biology, 2018, Vol. 52, No. 6, pp. 929–936. © Pleiades Publishing, Inc., 2018.
Original Russian Text © A.M. Kargatov, E.V. Brazhnikov, A.V. Efimov, 2018, published in Molekulyarnaya Biologiya, 2018, Vol. 52, No. 6, pp. 1074–1081.

STRUCTURAL-FUNCTIONAL ANALYSIS OF BIOPOLYMERS AND THEIR COMPLEXES

UDC 577.150.2

Structure and Features of Amino Acid Sequences of L-Modules in SH3-Like Folds

A. M. Kargatov[a], E. V. Brazhnikov[a], and A. V. Efimov[a, *]

[a] *Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow oblast, 142290 Russia*

**e-mail: efimov@protres.ru*

Received April 17, 2018; in final form, May 29, 2018

Abstract—A novel L-shaped repeat module whose structure can be represented as β-strand–loop–β-strand has been identified in a stereochemical analysis of nonhomologous SH3-like folds. β-Strands of the L-module are positioned at a ~90° angle to each other in different orthogonally packed β-layers. Together with a crossover loop, they form a half-turn of a right-handed superhelix. A database of 60 nonhomologous SH3-like domains has been compiled using the Protein Data Bank to study structural similarities and differences of L-modules. Occurrence frequencies of L-modules have been determined depending on the length of their loops. It has been shown that L-modules with $\beta_m\alpha\alpha\alpha\beta_n$- and $\beta_m\alpha\alpha\alpha\beta\alpha\beta_n$-conformations, where *m* and *n* are numbers of β-residues in the first and second β-strands, occur most often (57 and 8%, respectively). Spatial structures of L-modules of the same type are very similar, demonstrated through superimposing them using computer programs. Structural alignment of the amino acid sequences encoding L-modules has been performed, making it possible to identify key positions for hydrophobic, hydrophilic, and proline residues.

Keywords: α-helix, β-strand, structural motif, structural similarity

DOI: 10.1134/S0026893318060092

INTRODUCTION

SH3- and SH3-like domains are ubiquitous in both homologous and nonhomologous proteins. We have found more than 800 PDB files of SH3-like domains in the Protein Data Bank (PDB). From these data, we selected 60 nonhomologous domains according to the criteria set using software for pairwise amino acid sequence alignment [1]. In general, the structure of SH3-like domains consists of five to six β-strands forming a closed β-cylinder or β-barrel, to which several additional α-helices and/or β-strands can be attached. A distinctive feature of the β-barrels in SH3-like domains (we will refer to them as SH3-folds) is that they have several coexisting structural motifs that often occur in other proteins. For instance, the first three β-strands of the SH3-fold form a 3β-corner [2]. As a rule, the central β-strand of the 3β-corner is very twisted and bent up to 90°. Half of it is located in one β-layer of the β-barrel; the other half, in the other. For this reason, these halves are often formally considered as two β-strands (hence five or six β-strands in SH3-folds). On the other hand, all SH3 folds contain S-shaped β-sheets, which, together with adjacent β-strands, form right-handed β-strand–S-sheet–β-strand superhelices [3, 4]. The first two β-strands in many SH3 folds form strongly twisted and bent β-hairpins, which are very similar to β-hairpins in other proteins [5].

In this study, we identified and characterized another structural element that is common to all SH3-folds. This element is called L-module because its stacking resembles the Latin letter L. Its structure can be represented as β-strand–loop–β-strand. β-Strands of the L-module are positioned at an angle of ~90° in different β-layers of the SH3-fold. Together with a crossover loop, they form a half-turn of a right-handed superhelix. The study demonstrated that structures of nonhomologous L-modules of the same type overlap well when superimposed and that their amino acid sequences have the same key hydrophobic, hydrophilic, and proline residues.

EXPERIMENTAL

Proteins containing SH3-folds were selected for study from the PDB using the PSBOST structural classification developed in our group (the classification is available at http://strees.protres.ru/ [6]) and directly from the server PDB (http://www.rcsb.org/pdb/) using keywords. Nonhomologous proteins or domains were selected for analysis. Homology testing was performed by the Blast 2 Sequences software for pairwise amino acid sequence alignment (https://blast.ncbi.nlm.nih.gov/Blast.cgi) [1]. Proteins were considered nonhomologous if their E-value, which accounts for both the complete identity of sequence fragments and the conservative substitutions, was higher than 1e-04.

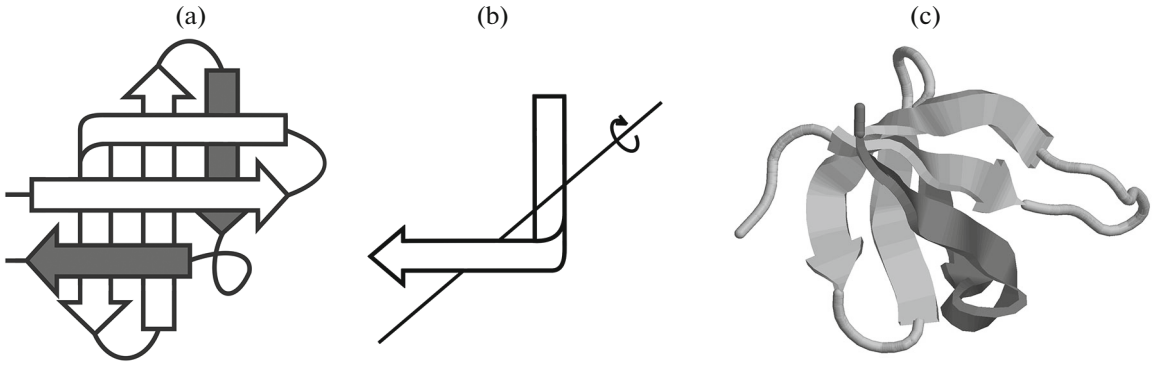

Fig. 1. (a) Schematic representation of the overall stacking of SH3-folds (the L-module is shown in gray). (b) The L-module chain fragment connecting two β-layers forms a half-turn of a right-handed superhelix around the imaginary axis. (c) Image of the SH3-fold structure of protein 1QMC using RasMol [7] (the L-module is shown in a darker color). In all cases, the β-strands are shown as arrows directed from N- to the C-termini.

In the PDB, we found more than 800 PDB files of SH3- and SH3-like domains, from which we selected 60 nonhomologous SH3-folds (see Table 1) according to the criteria of the Blast 2 Sequences software [1].

The main object of the study, the L-module, was identified visually using the RasMol molecular graphics software [7]. MolMol was used to calculate the angles φ and ψ [8]. Conformations of amino acid residues were designated by a set of symbols in accordance with the nomenclature proposed earlier [9]. The DeepView/Swiss-PDBViewer software was used to superimpose spatial structures [10]. Alignment of the amino acid sequences of L-modules based on their structural similarity was performed manually.

RESULTS

The main object of the present study is nonhomologous SH3-like domains, or rather their β-barrels, which are referred to here as SH3-folds. The general packing of the polypeptide chain in the spatial structure of SH3-folds is given in Fig. 1. As noted above, the first three (or four) β-strands form a 3β-corner [2], and the central region is occupied by an S-shaped β-sheet, which, together with its flanking β-strands, forms a turn of a right-handed β-strand–S-sheet–β-strand superhelix [3]. Analysis showed that the C-terminal region of all SH3-folds contains the structural module β-strand–loop–β-strand, which is referred to here as the L-module (Figs. 1a, 1b). β-Strands of the L-module are positioned in different β-sheets of the SH3-fold at an angle of 90°; together with the loop, they form a half-turn of a right-handed superhelix (Fig. 1b).

Figure 2 shows the dependence of the occurrence frequency of L-modules on the length of their loops. It is evident that a vast majority of L-modules have loops of three, five, and six residues. A more detailed analysis shows that the most frequently occurring are L-modules with $\beta_m\alpha\alpha\alpha\beta_n$- and $\beta_m\alpha\alpha\alpha\beta\alpha\beta_n$-conformations, where *m* and *n* are the number of residues in the first and second β-strands, respectively. Taking into account that the first residue to form a hydrogen bond at the N-terminus of the α-helix has, in fact, a β-conformation [9] (the situation is similar in spirals 3_{10}), it can be concluded that a vast majority of L-modules contains a coil of an α-helix or helix 3_{10} in their loops (see Figs. 3 and 4). We note that this first residue of the α-helical coil is simultaneously the last residue of the first β-strand of the L-module. In addition to the turns of the α-helix, the loops of L-modules with $\beta_m\alpha\alpha\alpha\beta\alpha\beta_n$-conformations also contain standard β-half-turns with βαββ conformations (see [9] for their description, and also see Fig. 4).

L-modules from different nonhomologous SH3-folds have very similar spatial structures and overlap well when superimposed. Fig. 3 shows the superposition of 13 L-modules (slightly shorted at their termini) with a $\beta_m\alpha\alpha\alpha\beta_n$-conformation. The representative sample included the following proteins: 1M1G, 1TOV, 1Y71, 2PU9, 2RH2, 2X4X, 3H8Z (2 modules), 3PMT, 4FW1, 4RXJ, and 4YTL (2 modules). Fig. 4 shows the superposition of five L-modules (slightly shorted at their termini) with a $\beta_m\alpha\alpha\alpha\beta\alpha\beta_n$-conformation. These modules came from the following proteins: 1LOJ, 1M5Q, 1OU8, 1U1S, and 2RA2. Note that the structures of the L-modules overlap each other well despite belonging to unrelated, nonhomologous proteins.

The structural similarity of nonhomologous L-modules makes it possible to perform a structural alignment of their amino acid sequences by comparing amino acid residues in structurally equivalent positions. In Fig. 5 shows the structural alignment of amino acid sequences of 34 L-modules with a $\beta_m\alpha\alpha\alpha\beta_n$-conformation. The alignment was performed manually. The sequence fragments of the first β-strands were aligned by the last residue of each β-strand so that the internal (directed towards the hydrophobic core) and the external positions fell into their respective columns. The fragments

Table 1. List of nonhomologous SH3-folds, the resolution of their structures, and the length of loops in their L-modules

№	PDB code, chain	Chain fragment	Resolution, Å	Number of residues in the loop
01	1f39 (A)	178–226	1.9	2
02	1gec (E)	129–215	2.1	16
03	1h5w (A)	61–123	2.1	11
04	1hyo (A)	18–117	1.3	3
05	1i1j (A)	21–90	1.39	3
06	1igq (A)	307–357	1.7	3
07	1jb0 (E)	5–68	2.5	3
08	1loj (A)	24–82	1.9	5
09	1m1g (A)	199–248	2.0	3
10	1m5q (A)	22–76	2.0	5
11	1ne8 (A)	6–95	2.1	3
12	1ou8 (A)	26–95	1.6	5
13	1oz2 (A)	242–292	1.55	5
14	1oz2 (A)	349–399	1.55	5
15	1oz2 (A)	453–503	1.55	5
16	1r77 (A)	284–362	1.75	11
17	1sf9 (A)	84–127	1.71	3
18	1tov (A)	150–219	1.77	3
19	1u1s (A)	21–65	1.6	5
20	1ugp (B)	143–226	1.63	3
21	1v76 (A)	51–99	2.0	3
22	1vq8 (Q)	37–94	2.2	3
23	1vqo (A)	103–157	2.2	2
24	1y71 (A)	9–81	1.95	3
25	1zuk (A)	7–67	1.9	3
26	1zuu (A)	3–57	0.97	3
27	2bjq (A)	32–81	1.75	8
28	2bjq (A)	118–166	1.75	6
29	2bjq (A)	216–263	1.75	6
30	2bjq (A)	295–343	1.75	6
31	2cb5 (A)	321–425	1.85	9
32	2evr (A)	14–79	1.6	6
33	2f5k (A)	15–65	2.2	3
34	2fpe (A)	3–58	1.75	3
35	2ox7 (A)	6–70	1.78	8
36	2pu9 (B)	5–73	1.65	3
37	2ra2 (A)	5–54	1.9	5
38	2rh2 (A)	28–77	0.96	3
39	2v1r (A)	14–76	2.1	3
40	2x4x (A)	1087–1165	1.85	3
41	2xc1 (A)	47–83	1.65	6

Table 1. (Contd.)

№	PDB code, chain	Chain fragment	Resolution, Å	Number of residues in the loop
42	2zgw (A)	190–235	1.5	3
43	2zpl (A)	129–218	1.7	18
44	3bz7 (A)	33–80	2.0	3
45	3fdr (A)	32–79	1.75	3
46	3g1j (A)	18–89	1.7	16
47	3g7z (A)	3–83	2.35	3
48	3h8z (A)	15–60	1.92	3
49	3h8z (A)	74–126	1.92	3
50	3pmt (A)	560–608	1.8	3
51	3qir (A)	31–115	2.45	3
52	4edj (A)	14–105	1.9	10
53	4edj (A)	112–196	1.9	10
54	4fw1 (A)	222–268	1.7	3
55	4kxt (A)	263–345	2.29	3
56	4rxj (A)	962–1019	2.1	3
57	4ytl (A)	538–583	1.6	3
58	4ytl (A)	588–632	1.6	3
59	5jqc (A)	256–329	2.15	3
60	5jqc (A)	335–411	2.15	3

of α-helical sequences were aligned by the first residues of the α-helices (these residues are also the last residues of the first β-strands). Thus, each column of the multiple alignment contains structurally equivalent residues. Their conformation is indicated by the letters α and β in the upper line. A similar alignment of amino acid sequences of the L-modules having a $\beta_m\alpha\alpha\alpha\beta\alpha\beta_n$-conformation is shown in Fig. 6.

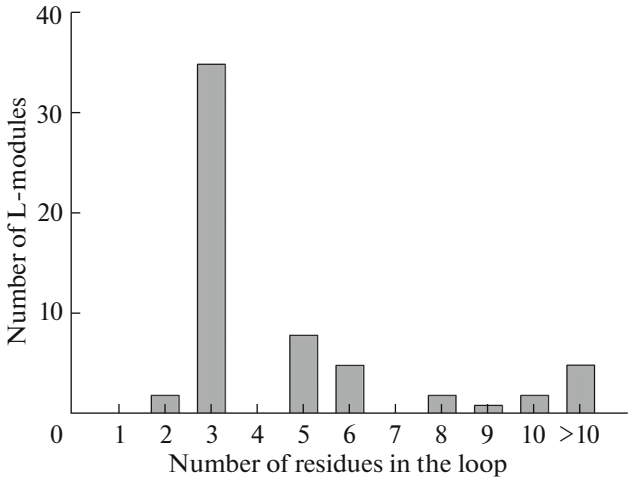

Fig. 2. Occurrence frequencies of L-modules depending on the length of their loops.

In accordance with the principles of the structural organization of proteins, side chains of amino acid residues immersed in the hydrophobic core or local hydrophobic environment should be hydrophobic, whereas those located at the surface are mostly hydrophilic. In ideal cases, there should be a strict alternation of hydrophobic and hydrophilic residues in β-strands of two-layer structures [11]. However, this rule does not always holds in β-strands located at the edges of β-sheets: their side chains, which are directed towards the hydrophobic core, can be accessible to water molecules and, consequently, may be hydrophilic. As can be seen in Fig. 1a, β-strands forming the L-module are located on the edges of β-layers. For this reason, in these β-strands, many side chains directed toward the hydrophobic core (their letter designations are underlined in Figs. 5 and 6) are hydrophilic. Nevertheless, many of them have long hydrophobic legs (Lys, Arg, Glu, and Gln). In the L-modules with $\beta_m\alpha\alpha\alpha\beta_n$-conformations, the rule is most strictly implemented in two positions: these are the second to last position of the first β-strand and the first position of the second β-strand. In most cases, these positions are occupied by hydrophobic residues (shown in gray in Fig. 5). In the L-modules with

$\beta_m\alpha\alpha\alpha\beta\alpha\beta_n$-conformations, there are three such positions: the second to last position of the first β-strand, the β-position in the loop, and the first internal position of the second β-strand. These positions are occupied only by hydrophobic residues (shown in gray in Fig. 6). Of particular interest are the first positions of α-helixes, which are also the last positions of the first β-strands in both types of L-modules (also shown in grey in Figs. 5 and 6). In these cases, the so-called "perpendicular entrance" to the α-helix was discovered. It occurs very often at the N-termini of α-helices in proteins and has a $\beta\beta\alpha_m$-conformation, where *m* is the number of residues in the α-helix [9, 12]. Statistical analysis showed that the first (β) positions of α-helices in proteins are mostly occupied by hydrophilic (~80%, the majority of which are Ser, Thr, Asp, and Asn) and small (~ 17%, Gly, Ala, and Pro) residues. In general, a similar distribution of residues was found in the first (β) positions of α-helices of L-modules, although they contain an anomalously large number of prolines. In the L-modules with a $\beta_m\alpha\alpha\alpha\beta_n$-conformation, 50% of these positions are occupied by the proline. In addition, prolines are found in the second and third positions of α-helices (Fig. 5).

DISCUSSION

In globular proteins, the polypeptide chain repeatedly folds into itself and forms a series of frequently repeating blocks of two or more adjacent α-helices and/or β-strands, which are called supersecondary structures or structural motifs [13–16]. Of particular interest are structural motifs with unique stacking patterns and a certain pseudochirality [15, 16]. Each type of structural motif can be characterized by the same number of α-helices and/or β-strands and their identical arrangement along the chain and in space. It is believed that these structural motifs are stable, capable of folding into their unique structures on their own, and can act as nucleating centers and/or structural blocks in protein folding [15, 17].

In previous papers [18, 19], we considered in detail different variants of Π-modules whose structure resembles the Greek (or Russian) letter Π. Apparently, the Π-modules are unstable on their own. However, in combination with β-hairpins, S- and Z-shaped β-sheets, and βαβ-units, they form compact, closed, and stable structural motifs, which are very common in proteins. As can be seen in Fig. 1, L-modules are noncompact structures. Their β-strands hardly interact with each other but their spatial position is always the same and, together with the loop, they form a half-turn of a right-handed superhelix. The interaction of the L-module with the N-terminal part of the SH3-fold leads to the formation of a closed β-barrel structure, which has enhanced cooperativity and greater stability (see, for example, [20]). Apparently, this is the main structural role of the L-modules.

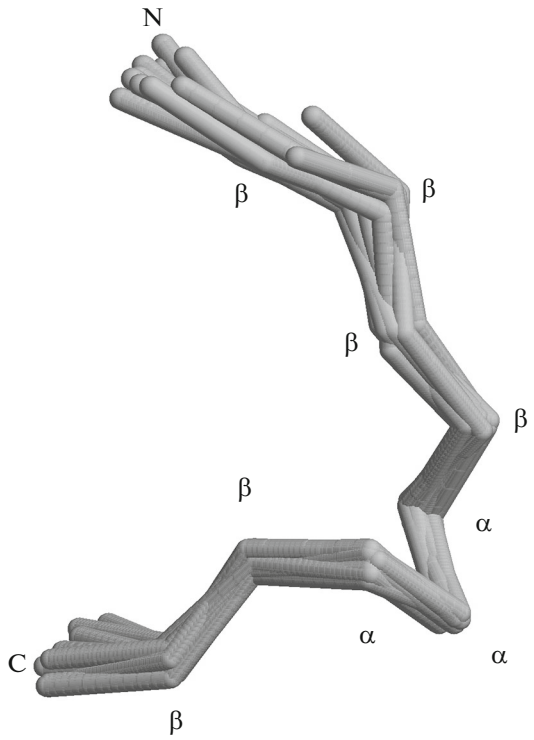

Fig. 3. Superposition of the structures of 13 L-modules with the $\beta_m\alpha\alpha\alpha\beta_n$-conformation performed in DeepView/Swiss-PDBViewer [10]. The image came from RasMol [7]. The letters α and β denote conformations of the corresponding residues. The letters N and C denote the ends of the chain segments.

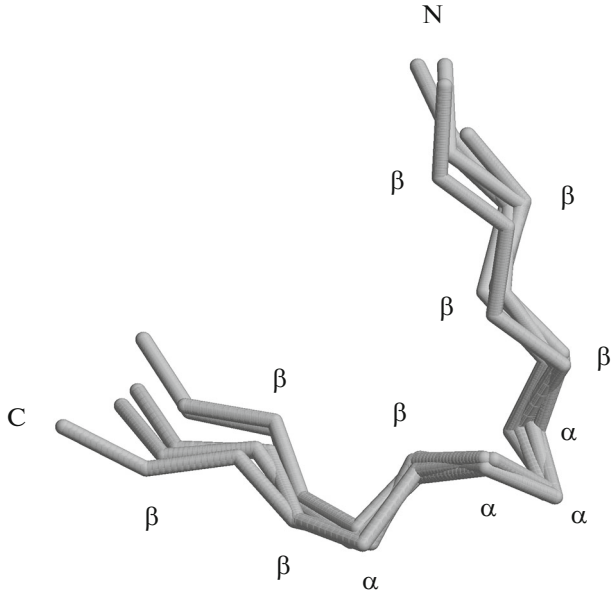

Fig. 4. Superposition of the structures of five L-modules with the $\beta_m\alpha\alpha\alpha\beta\alpha\beta_n$-conformation as described in Fig. 3.

It should be noted that structures similar to L-modules were first found in five proteins with orthogonal packing of β-layers and have been described in [21]. It has also been shown that β-strands, which go

No.	PDB code	ββββββββ	ααα	ββββββ	Chain fragment
01	1HYO : A	fts	QAS	atmh	107–116
02	1I1J : A	gyfp	SSI	vre	77–86
03	1IGQ : A	efean	LAD	vkl	342–352
04	1JB0 : E	tnnfa	LHE	vqe	57–67
05	1M1G : A	rmtpveld	FDQ	vek	234–247
06	1NE8 : A	irtld	KQR	lt	80–89
07	1SF9 : A	ealp	ISV	lqe	117–126
08	1TOV : A	ggfvr	PVD	vkv	208–218
09	1UGP : B	svyydcw	EPY	iel	213–225
10	1V76 : A	iwkvp	KDV	sifefe	85–98
11	1VQ8 : Q	kektiivt	AAH	lrr	80–93
12	1Y71 : A	qtnip	EQM	vkk	70–80
13	1ZUK : A	viegifp	KSF	vav	54–66
14	1ZUU : A	etglvp	TTY	ir	46–56
15	2F5K : A	ewvp	ESR	vlk	55–64
16	2FPE : A	rgvfp	AYY	aie	47–57
17	2PU9 : B	kahfr	PDE	vtl	62–72
18	2RH2 : A	vqiyp	VAA	ler	66–76
19	2V1R : A	igyip	YNY	iei	65–75
20	2X4X : A	wqwlp	RTK	lvp	1154–1164
21	3BZ7 : A	drqvk	KDD	anq	69–79
22	3FDR : A	ngdcp	LKD	lra	68–78
23	3G7Z : A	asvp	VSV	ig	69–77
24	3H8Z : A	rqip	FGD	vr	51–59
25	3H8Z : A	eivt	LER	lrp	116–125
26	3PMT : A	yeevl	LSN	ikp	597–607
27	3QIR : A	illl	PEL	sf	106–114
28	4FW1 : A	viwvp	SRK	vkp	257–267
29	4KXT : A	tylp	LEV	cni	335–344
30	4RXJ : A	dyywvh	QGR	vfp	1007–1018
31	4YTL : A	lefp	IST	lrk	573–582
32	4YTL : A	evtit	ANN	ls	622–631
33	5JQC : A	wmd	KRA	fvy	320–328
34	5JQC : A	wld	ARA	fhv	402–410

Fig. 5. Structural alignment of amino acid sequences of L-modules with the $\beta_m\alpha\alpha\alpha\beta_n$-conformation. Conformations of residues in the columns are indicated in the upper line by the letters α and β. α-Spiral residues are shown in uppercase, and β-structural residues are shown in lower case letters. β-Residues directed toward the hydrophobic core are underlined. In each line, the PDB code of the protein is given on the left, and the chain section forming the L-module is shown on the right. See text for additional information.

No.	PDB code	ββββββββ	αααβα	ββββ	Chain fragment
01	1LOJ : A	vlir	GDNiV	yisr	69–81
02	1M5Q : A	rvfim	YRYiV	hid	62–74
03	1OU8 : A	vsrelyip	MGAaL	aiya	78–94
04	1U1S : A	qmvy	KHAiS	tvvp	52–64
05	2RA2 : A	kqqin	RTDvK	emva	40–53

Fig. 6. Structural alignment of amino acid sequences encoding L-modules with the $\beta_m\alpha\alpha\alpha\beta\alpha\beta_n$-conformation. The notation is the same as in Fig. 5.

from one β-layer to another, are strongly twisted and bent by ~ 90° forming a half-turn of a right-handed superhelix just as in the L-module. Two adjacent β-strands in ββ-corners form two right-handed superhelices (i.e., two L-modules) at the junction between two β-layers [22]. However, L-modules are most commonly found in 3β-corners and, hence, in proteins containing them. The structure of a 3β-corner is a Z-shaped β-sheet folded into itself so that the central β-strand forms a half-turn of a right-handed superhelix, i.e., the L-module, at the junction between two β-layers [23]. At the junction points, the central β-strands can acquire β-bend conformation [20] or form small standard structures having βγββ, ββα_Lβ-, βαγβ, βγεβ-, or βεβ-conformations [22–24]. While constructing an updated structural tree, we found 720 proteins containing 3β-corners in the PDB, of which 224 are nonhomologous proteins (the tree is available at http://strees.protres.ru/2_1.xml; see also [6, 24]). This database included most of the 60 SH3 folds (not all of them as the list of these domains was shorter in 2014 [24]) because the first three (or four) β-strands in each of these folds form a 3β-corner and, respectively, the second or second and third β-strands (depending on the conformation and the number of residues in the transition) form a right L-module. Thus, each SH3-fold contains two right L-modules, one of which is located in the N-terminal region and belongs to the 3β-corner, whereas the other is isolated at the C-terminus. As follows from the foregoing, the present paper is mainly devoted to these isolated L-modules.

Comparative analysis shows that the majority of L-modules in the 3β-corners of SH3 folds have shorter junctions between their layers than isolated L-modules. As a rule, the former L-modules contain from zero to two residues in α-, α_L-, γ-, or ε-conformations (see also [24]), whereas most of the latter (as evident from Fig. 2) consist of three, five, or more residues. A key role in the formation of many L-module junctions belonging to the 3β-corners is played by residues with an α_L- or ε-conformation, which must be glycines or residues with flexible side chains. The junctions (or loops) of most of the isolated L-modules contain no residues with α_L- and ε-conformations, and glycines are relatively rare in these positions (see Figs. 5 and 6). Prolines, on the contrary, are found quite often and, apparently, play an essential role in the formation of junctions in isolated L-modules.

In conclusion, we should note that isolated L-modules exist in proteins of other structural classes and superfamilies other than SH3 folds. For example, L-modules with $\beta_m\alpha\alpha\alpha\beta_n$-conformation are found in the following proteins in the PDB: 1B9M, 1FR3, 1PL6, 1Q12, 2J7N, 2ZB3, 2ZB3, 5A4G, and others. Despite the fact that these L-modules are found in different structural classes of proteins, they also form right-handed superhelices and overlap well with each other and other L-modules from SH3-folds when superimposed. The positions of their key hydrophobic, hydrophilic, and proline residues are also similar. All this indicates that this structural similarity is caused not by homology but by common physical and chemical principles.

ACKNOWLEDGMENTS

This study was performed within the framework of state task no. 0115-2018-0008 and supported in part by the Russian Foundation for Basic Research, project no. 17-04-00242.

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interest. This article does not contain any studies involving animals or human participants performed by any of the authors.

REFERENCES

1. Tatusova T.A., Madden T.L. 1999. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.* **177**, 187–188.
2. Efimov A.V. 1997. A structural tree for proteins containing 3β-corners. *FEBS Lett.* **407**, 37–41.
3. Efimov A.V. 1998. A structural tree for proteins containing S-like β-sheets. *FEBS Lett.* **437**, 246–250.
4. Efimov A.V. 2010. Structural motifs are closed into cycles in proteins. *Biochem. Biophys. Res. Commun.* **399**, 412–415.
5. Boshkova E.A., Brazhnikov E.V., Efimov A.V. 2016. Relationship between structure and amino acid sequence of strongly twisted and coiled β-hairpins in globular proteins. *Mol. Biol.* (Moscow). **50** (5), 777–782.
6. Gordeev A.B., Kargatov A.M., Efimov A.V. 2010. PCBOST: Protein classification based on structural trees. *Biochem. Biophys. Res. Commun.* **397**, 470–471.
7. Sayle R.A., Milner-White E.J. 1995. RASMOL: Biomolecular graphics for all. *Trends Biochem. Sci.* **20**, 374–376.
8. Koradi R., Billeter M., Wüthrich K. 1996. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.
9. Efimov A.V. 1986. Standard polypeptide chain conformations in irregular protein regions. *Mol. Biol.* (Moscow). **20**, 250–260.
10. Guex N., Peitsch M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis.* **18**, 2714–2723.
11. Lim V.I. 1974. Structural principles of the globular organization of protein chains. A stereochemical theory of globular protein secondary structure. *J. Mol. Biol.* **88**, 857–872.
12. Efimov A.V. 1984. A novel supersecondary structural motif in proteins: αα-corner. *Mol. Biol.* (Moscow). **18**, 1524–1537.

13. Rao S.T., Rossmann M.G. 1973. Comparison of supersecondary structures in proteins. *J. Mol. Biol.* **76**, 241–256.
14. Levitt M., Chothia C. 1976. Structural patterns in globular proteins. *Nature*. **261**, 552–558.
15. Efimov A.V. 1994. Favoured structural motifs in globular proteins. *Structure*. **2**, 999–1002.
16. Efimov A.V. 2018. Chirality and handedness of protein structures. *Biochemistry* (Moscow). **83 (**Suppl. 1), 103–110.
17. Efimov A.V. 1997. Structural trees for protein superfamilies. *Proteins*. **28**, 241–261.
18. Efimov A.V. 2017. Structural motifs in which β-strands are clipped together with the Π-like module. *Proteins*. **85**, 1925–1930.
19. Kargatov A.M., Efimov A.V. 2018. Unique combinations of βαβ-units and Π-like modules in proteins and specific features of their amino acid sequences. *Mol. Biol.* (Moscow). **52** (1), 36–41.
20. Kutyshenko V.P., Gushchina L.V., Khristoforov V.S., Prokhorov D.A., Timchenko M.A., Kudrevatykh Yu.A., Fedyukina D.V., Filimonov V.V. 2010. NMR structure and dynamics of the chimeric protein SH3-F2. *Mol. Biol.* (Moscow). **44** (6), 948–957.
21. Chothia C., Janin J. 1982. Orthogonal packing of beta-pleated sheets in proteins. *Biochemistry*. **21**, 3955–3965.
22. Efimov A.V. 1991. Structure of coiled β-β-hairpins and β-β-corners. *FEBS Lett.* **284**, 288–292.
23. Efimov A.V. 1992. A novel super-secondary structure of β-proteins: A triple-strand corner. *FEBS Lett.* **298**, 261–265.
24. Efimov A.V., Boshkova E.A. 2014. Two mechanisms of protein folding: A theoretical analysis (review article). *Russ. J. Bioorg. Chem*. **40** (6), 612–619.

Translated by Yu. Modestova