Molecular Biology, Vol. 39, No. 5, 2005, pp. 769–776. Translated from Molekulyarnaya Biologiya, Vol. 39, No. 5, 2005, pp. 878–886.

STRUCTURAL AND FUNCTIONAL ANALYSIS OF BIOPOLYMERS AND THEIR COMPLEXES

UDC 577.150.2

Analysis of Interactions of Buried Polar Side Chains in β-Proteins

E. V. Brazhnikov and A. V. Efimov

Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow oblast, 142290 Russia
e-mail: efimov@protres.ru

Received February 3, 2005

Abstract—Qualitative and quantitative analysis of polar side chains inaccessible to water molecules, as well as their interactions in 100 globular β-sheet proteins, was performed. It was shown that completely buried polar side chains are widespread in β-proteins, with their vast majority being involved in side chain–side chain or side chain–main chain interactions. An analysis of frequency of occurrence of different side chain–partner pairs demonstrated that these interactions are selective. The results were compared with similar data obtained earlier for α-helical proteins.

Key words: globular proteins, β-structure, accessible surface, hydrogen bond, selectivity of interactions

INTRODUCTION

A characteristic structural feature of globular proteins is the presence in their molecules of hydrophobic cores formed predominantly by hydrophobic side chains and polar envelopes formed by hydrophilic side chains [1]. Hydrophilic side chains may be located in the hydrophobic core if they have partners for the formation of hydrogen, salt, or coordination bonds [2, 3]. Although the presence of several completely buried (inaccessible to water molecules) polar side chains in protein cores was reported long ago [2–5], their structural role has become clearer fairly recently, after the development of model systems based on the so-called leucine zippers, or coiled-coil (CC) proteins. For example, it was shown that the oligomerization state of α-helices and the selection of parallel or antiparallel orientation of α-helices in proteins of this class largely depend on the interaction between buried polar residues [6–8]; i.e., these interactions make α-helical packing specific.

A comparative analysis performed in [9, 10] showed that interhelical polar interactions in CC and globular α-proteins are largely similar, which is suggestive of a similar structural role of these interactions in both classes of proteins. Importantly, interhelical polar interactions are selective; i.e., side chain donors in these proteins predominantly form certain pairs rather than interact with any acceptors [9, 10]. It can be assumed that such selectivity is characteristic only of interhelical interactions and will not be observed in β-sheet proteins. Verification of this assumption was one of the main goals of this study. We found that the interactions between buried side chains in β-proteins are also selective and apparently play a key role in protein folding.

EXPERIMENTAL

Analysis was performed with 100 globular β-proteins with a resolution as high as 2.5 Å. The proteins were selected form the SCOP database (http://scop.mrc-lmb.cam.ac.uk/scop) [11]. This set included proteins with all known types of β-sheet folding. The PDB codes of the proteins selected were as follows:

1ag6, 1aol, 1aq6, 1at3, 1aun,1az3, 1azc, 1b56, 1bd9, 1beb, 1ca9, 1cyn, 1czs, 1d2s, 1d3b, 1d7k, 1dev, 1dg6, 1dlc, 1dmh, 1dnl, 1du5, 1e2w, 1e44, 1e5p, 1edy, 1efi, 1ejf, 1elp, 1epa, 1es6, 1eut, 1exm, 1f35, 1f3u, 1f41, 1g13, 1g31, 1g8f, 1gny, 1goh, 1gtr, 1hbq, 1hd8, 1hug, 1hxr, 1i4u, 1i81, 1iaz, 1ikp, 1j71, 1j8s, 1jac, 1jhj, 1jlx, 1jsg, 1ju3, 1k5j, 1k6w, 1kfu, 1knb, 1lox, 1neu, 1osp, 1p35, 1pdk, 1pgs, 1poo, 1qba, 1qex, 1qfv, 1qhd, 1qou, 1qrw, 1qtf, 1qvc, 1rie, 1rl2, 1sft, 1sgp, 1sgt, 1slu, 1smp, 1spp, 1stm, 1stn, 1tl2, 1vie, 1whi, 1who, 1ycs, 1ygs, 1ytf, 2eng, 2viu, 3cms, 3ezm, 3sil, 4aah, 4tsv.

Side chains of the amino acid residues Glu, Gln, Asp, Asn, Lys, Ser, Thr, Arg, His, Trp, and Tyr were considered polar. Hydrogen bonds were determined and the accessible surface of atoms involved in hydrogen bonding was calculated using the WHAT_IF software (available at http://www.cmbi.kun.nl:1100/WIWWWI/). A side chain was considered completely buried if its accessibility was lower than 2 Å^2. We have taken into account only optimal hydrogen bonds, the number of which is smaller than that of all possible hydrogen bonds. The criteria of formation of optimal and possible hydrogen bonds and their geometric parameters are reported in [12].

0026-8933/05/3905-0769

Table 1. Frequencies of occurrence of buried polar residues in 100 β-proteins

No.	Protein code	E	Q	D	N	K	S	T	R	H	W	Y	Total per protein
1	1ag6				1							1	**2**
2	1aol	1			1			2			2	1	**7**
3	1aq6						3						**3**
4	1at3			1	2		2					1	**6**
5	1aun			1	2		4	1	1		2	2	**13**
6	1az3					1		1				1	**3**
7	1azc				1		1	1		2	1	3	**9**
8	1b56		2				1	2	1		1	1	**8**
9	1bd9			1				3	1	1	2	5	**13**
10	1beb		1				2	1					**4**
11	1ca9						2						**2**
12	1cyn		1				1						**2**
13	1czs	1	2		1		1	2				1	**8**
14	1d2s		1	2			2	2		1		1	**9**
15	1d3b	1						1				1	**3**
16	1d7k			2	1		2	1			1	5	**12**
17	1dev	2		1	1		1					1	**6**
18	1dg6		1				2	2				2	**7**
19	1dlc	1		1	2		4	6		2	2	3	**21**
20	1dmh		1		1		1	2	2	2	1	2	**12**
21	1dnl				1			2	1			1	**5**
22	1du5			1	1		3	1	1		1	2	**10**
23	1e2w		1		1				1				**3**
24	1e44			2	1		1	1				2	**7**
25	1e5p							1				1	**2**
26	1edy			1			1	1				2	**5**
27	1efi	1	1				2	1			1	1	**7**
28	1ejf			2	1	1	2						**6**
29	1elp				1		4				2	3	**10**
30	1epa	1					1	2			1	1	**6**
31	1es6		1				1	2		1		1	**6**
32	1eut		3	1	5		9	19	2	3	1	7	**50**
33	1exm			1	3		3	3		1		2	**13**
34	1f35										1		**1**
35	1f3u						1	2		1			**4**
36	1f41	1		1				1		1		2	**6**
37	1g13						2	1			1		**4**
38	1g31						2	2		1			**5**
39	1g8f		1	1			1	1		2		2	**8**
40	1gny	1	2					1					**4**
41	1goh	2	2	2	4		13	14	1	3	6	5	**52**
42	1gtr	1	2	1	2		1	3	4	1		5	**20**
43	1hbq		1	1	1		3	1			1	1	**9**
44	1hd8			1	1	1	3	2					**8**
45	1hug	1			1			1		4	2	2	**11**
46	1hxr									2	1		**3**
47	1i4u		1	2	1		6	3			1	3	**17**
48	1i81						2					1	**3**
49	1iaz	1	1	1	3		2	1			2	3	**14**
50	1ikp	2	1		2	1	7	4	1	2	2	6	**28**

Table 1. (Contd.)

No.	Protein code	E	Q	D	N	K	S	T	R	H	W	Y	Total per protein
51	1j71		2	5			3	4			1	4	**19**
52	1j8s		1		3		3	2		1	1	1	**12**
53	1jac			1			4	1				1	**7**
54	1jhj		1		1		1	2		1	1	1	**8**
55	1jlx		1	1	3		4	3		3	2	2	**19**
56	1jsg			1			1			1	3	2	**8**
57	1ju3	1	2	6	4		5	3	3		2	4	**30**
58	1k5j	1					1	2		1			**5**
59	1k6w	1	2	1	2		2	2	1	4			**15**
60	1kfu	3	1	1			3	2	2		1	3	**16**
61	1knb			1			1	2			1	1	**6**
62	1lox		1			1	2	1			1		**6**
63	1neu										1		**1**
64	1osp		2				3	1			2	3	**11**
65	1p35		1	1	2		3	1	2			4	**14**
66	1pdk	1	1		2	1	1	2	1			3	**12**
67	1pgs	1			3		5	3	2	2		6	**22**
68	1poo	1	2	2	1	1	5	3			1	7	**23**
69	1qba	2	2	3	2		3	3	2	4	2	6	**29**
70	1qex						2	1		1	1		**5**
71	1qfv	2		2			2	1					**7**
72	1qhd			1				1			1		**3**
73	1qou			1				2		1			**4**
74	1qrw	1	1				7	2		1	2	1	**15**
75	1qtf						1	5		1		2	**9**
76	1qvc	1						2		1		1	**5**
77	1rie			1			1		3			3	**8**
78	1rl2							1				1	**2**
79	1sft	1				1	1	4	1	2		1	**11**
80	1sgp		1		1		4	4	1		1	1	**13**
81	1sgt	1	1				2	2				1	**7**
82	1slu							1					**1**
83	1smp	1		2	4		1	2			1	3	**14**
84	1spp	1						2			4	2	**9**
85	1stm			1	1		1	1				1	**5**
86	1stn	1		2								1	**4**
87	1tl2							1				3	**4**
88	1vie							1				1	**2**
89	1whi												**0**
90	1who				1			1			1		**3**
91	1ycs						2	2				3	**7**
92	1ygs			1			1						**2**
93	1ytf				2			4					**6**
94	2eng		1				2	2					**5**
95	2viu				1	1	5	2		1	1	2	**13**
96	3cms		2	3			4	2			3	3	**17**
97	3ezm						1						**1**
98	3sil	1	1	2	3		11	6	3	2	3	1	**33**
99	4aah		5	2	4	4	9	10	1	3	5	11	**54**
100	4tsv						2			1			**3**
	Total	**38**	**57**	**68**	**82**	**13**	**208**	**196**	**38**	**61**	**77**	**174**	**1012**

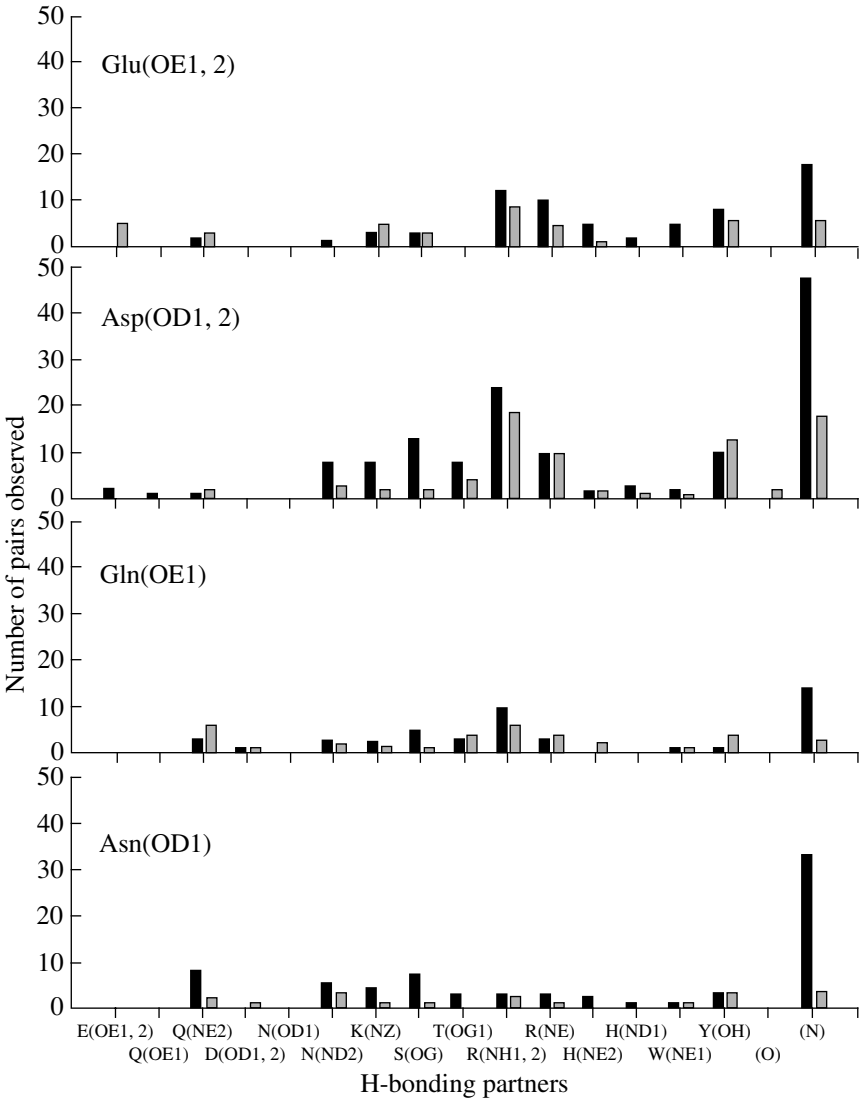

Fig. 1. Frequencies of occurrence of the pairs "buried acceptor–partner." The ordinate shows the number of respective buried acceptors; the abscissa shows their H-bonding partners. Black and gray rectangles indicate β-sheet and α-helical proteins, respectively.

RESULTS AND DISCUSSION

The results of quantitative analysis of buried polar side chains for 100 β-proteins are summarized in Table 1. In total, these proteins contain 8136 amino acid residues present in the β-structure. They include 3790 polar residues, of which 1012 residues are completely inaccessible to water molecules (Table 1). Most often, the side chains of serine (208 side chains), threonine (196 side chains) and tyrosine (174 side chains) are inaccessible to water molecules. The side chains of these residues occur two to three times more frequently than other buried side chains. The number of buried polar side chains varies significantly from protein to protein; however, the average number of buried polar residues in β-sheet proteins (1012 residues per 100 proteins) is almost twice as great as in α-helical proteins (557 residues per 80 proteins [9]). Thus, buried polar side chains are widespread in globular proteins, and their presence in proteins is a rule rather than an exception. Analysis showed that the majority of buried polar side chains have partners for the formation of hydrogen or salt bonds.

In the absence of such partners, burying a polar side chain into a hydrophobic environment results in breaking hydrogen bonds with water without formation of new bonds. Therefore, hydrophilic potential of a polar side chain becomes uncompensated, which considerably decreases the stability of the structure. Such a loss of partners may take place, for example, in the case of protein misfolding, which should destabilize misfolded structures and results in their subsequent unfolding. This means that buried polar side chains may play a pivotal role in protein folding, forcing a protein to "search for" a proper relative position

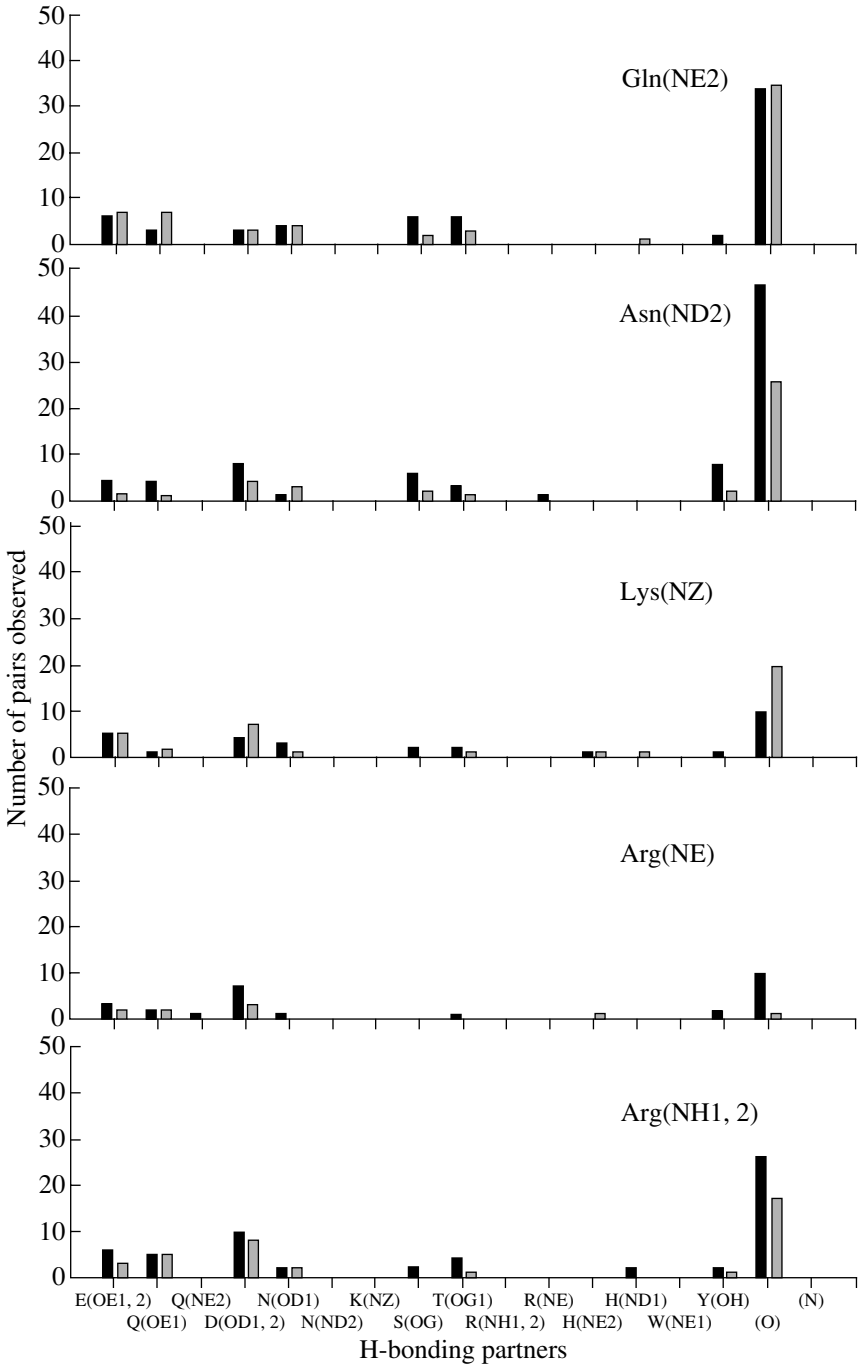

Fig. 2. Frequencies of occurrence of the pairs "buried donor–partner." The ordinate shows the number of respective buried donors; the abscissa shows their H-bonding partners. Black and gray rectangles indicate β-sheet and α-helical proteins, respectively.

of structural elements, when partners are located close to one another, thereby implementing a peculiar folding correction. On the other hand, the presence of buried polar side chains without partners in native protein molecules may play a functional role, because this may change the stability of proteins and mobility of certain parts of molecules relative to others.

To reveal possible selectivity of interactions of buried polar side chains, we studied the frequencies of occurrence of the pairs "buried side chain–partner," which form hydrogen bonds with one another. We proceeded from the fact that a high frequency of occurrence of some buried pairs "side chain–partner" compared to the number of other pairs formed by the same side chain is indicative of selectivity of interaction of the given side chain. Table 2 summarizes the data on the partners and the frequency of hydrogen bonding by different donors and acceptors of buried side

Table 2. Frequencies of occurrence of different buried donor–acceptor pairs involved in the formation of optimal hydrogen bonds in β-proteins

No.		**E**	**Q** OE1	**Q** NE2	**D**	**N** OD1	**N** ND2	**K** NZ	**S** OG	**T** OG1	**R** NHi	**R** NE	**H** NE2	**H** ND1	**W** NE1	**Y** OH	**O**	**N**	Do not form H-bonds
1	OE1 **E** OE2	–	–	2	–	–	1	3	3	–	12	10	5	2	5	8	–	**18**	**2**
2	OE1 **Q** NE2	–	–	3	1	–	3	2	5	3	10	3	–	–	1	1	–	**14**	**4**
		6	3	–	3	4	–	–	6	6	–	–	–	–	–	2	**34**	–	
3	OD1 **D** OD2	2	1	1	–	–	8	8	13	8	24	10	2	3	2	10	–	**48**	**4**
4	OD1 **N** ND2	–	–	8	–	–	5	4	7	3	3	3	2	1	1	3	–	**33**	**2**
		4	4	–	8	1	–	–	6	3	–	1	–	–	–	8	**47**	–	
5	**K** NZ	5	1	–	4	3	–	–	2	2	–	–	1	–	–	1	**10**	–	–
6	**S** OG	13	4	8	18	10	5	6	8	15	5	4	7	5	1	9	**82**	**80**	**19**
7	**T** OG1	10	5	10	10	4	4	6	15	8	6	2	4	3	5	9	**79**	**68**	**16**
8	NH1,2 **R** NE	6	5	–	10	2	–	–	2	4	–	–	–	2	–	2	**26**	–	**2**
		3	2	1	7	1	–	–	–	1	–	–	–	–	–	2	**10**	–	
9	NE2 **H** ND1	5	1	–	5	1	2	1	3	5	1	–	–	–	–	5	**11**	**2**	**6**
		2	1	–	6	1	1	–	4	5	1	1	1	2	1	5	**8**	**7**	
10	**W** NE1	1	2	–	3	–	–	–	3	7	–	–	1	1	–	3	**32**	–	**20**
11	**Y** OH	18	4	7	16	5	9	7	9	13	15	7	3	6	1	4	**45**	**20**	**27**

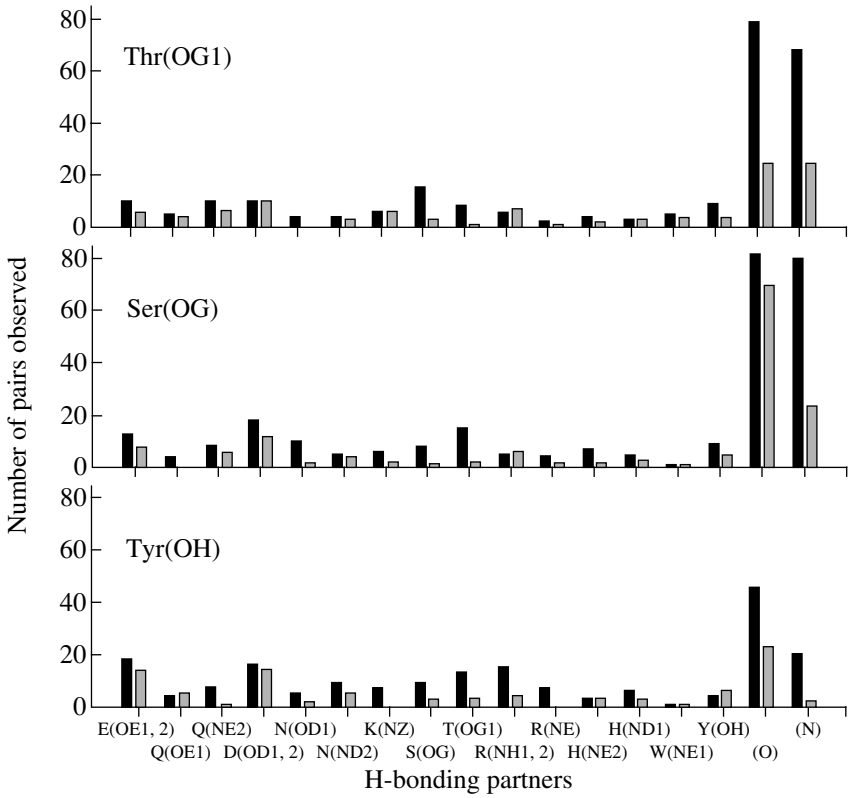

Fig. 3. Frequencies of occurrence of the pairs "buried donor/acceptor–partner" for buried side chains of threonine, serine, and tyrosine residues. The ordinate shows the number of buried oxygen atoms; the abscissa shows their H-bonding partners. Black and gray rectangles indicate β-sheet and α-helical proteins, respectively.

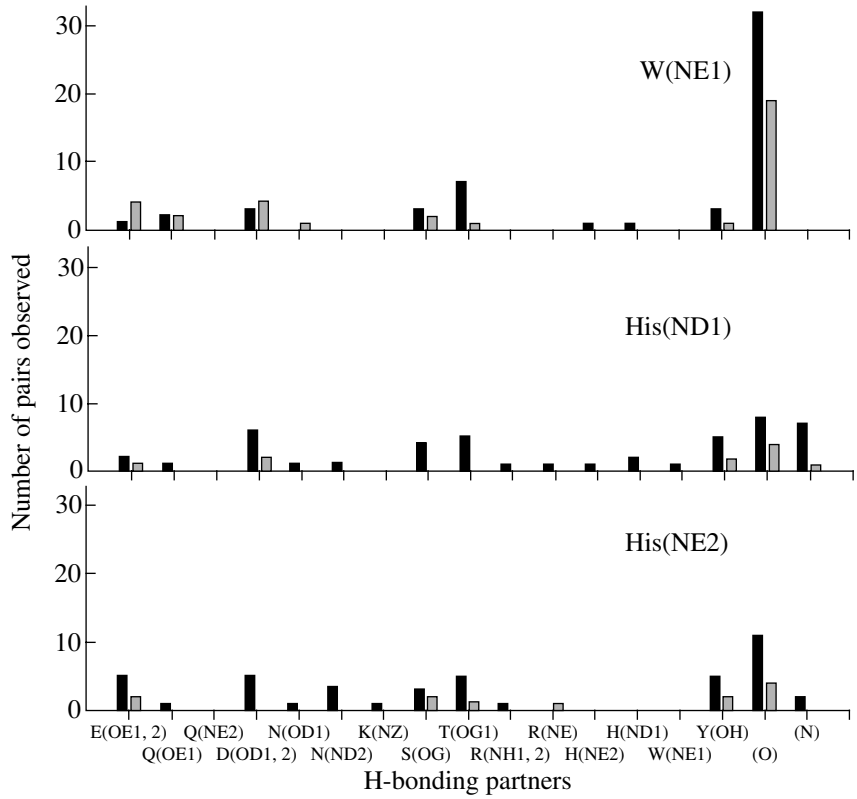

Fig. 4. Frequencies of occurrence of buried hydrogen bonds formed by the side chains of tryptophane and histidine residues. The ordinate shows the number of buried atoms; the abscissa shows their H-bonding partners. Black and gray rectangles indicate β-sheet and α-helical proteins, respectively.

chains. Lines in this table show the donors and acceptors of buried side chains; in columns, there are the partners for hydrogen bonding, which may be completely or partly buried. To make these data more illustrative, they are presented as a histogram in Figs. 1–4. The frequencies of occurrence of analogous pairs in α-helical proteins are also shown for comparison.

It can be seen that the "side chain–main chain" interactions predominate in all buried donors and acceptors and that the side chains of threonine, serine, and tyrosine residues make the greatest contribution to these interactions. The high frequency of occurrence of the "side chain–main chain" pairs is caused precisely by an excess of donors and acceptors of the main chain in proteins.

In view of this, we were interested predominantly in the "side chain–side chain" pairs. Similarly to α-proteins, β-proteins are characterized by sufficiently high frequencies of occurrence of the pairs Asp–Arg, Asp–Lys, Glu–Arg, Glu–Lys, His–Glu, and His–Asp. This may by accounted for by the fact that the side chains of these residues are charged and form ionic pairs with a greater gain in energy than the hydrogen bonds between uncharged donors and acceptors. Tyr, Ser, and Thr preferentially formed hydrogen bonds with one another, as well as with Asp and Glu. Note that like-charged side chains do not form hydrogen bonds with one another, which can be expected from general physicochemical considerations. A small number of hydrogen bonds observed in the pairs Asp–Glu and Glu–Glu is accounted for by the fact that one or both side chains may be uncharged and function as a donor.

A comparison of the frequencies of occurrence and the selectivity of formation of different pairs in α-helical and β-sheet proteins (Figs. 1–4) shows that the behavioral pattern of buried polar residues in both classes of proteins are similar, which is indicative of their similar roles in these proteins.

On the one hand, interactions between buried polar residues considerably contribute to protein stability. Recent studies showed that buried hydrogen bonds are more favorable in terms of energy than the bonds accessible to water molecules [13]. This can be accounted for by two facts. First, dielectric "constant" in the protein interior is lower than on its surface or in aquatic medium, which enhances electrostatic interactions between buried polar groups. Second, water molecules do not compete for the formation of hydrogen bonds with buried donors and acceptors of a protein.

On the other hand, buried polar groups can decrease the stability of proteins if they have no partners for the formation of hydrogen, coordination, or salt bonds. First, this may have a functional significance (e.g., lead to an increase in mobility of some parts of a molecule relative to other parts). Second, as mentioned above, in the case of misfolding of a protein, burial of polar groups into the hydrophobic environment without partners may sharply decrease protein stability, thereby ensuring correction of protein folding. At last, buried polar residues may determine the specificity of packing of secondary-structure elements [6–8, 10]. Taken together, all these facts show that buried polar residues play a key role in the formation and maintenance of native structure of proteins and emphasize the necessity of its further study.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 04-04-49343a) and the Scientific School Support grant (NSh-1968.2003.4).

REFERENCES

1. Perutz M.F., Kendrew J.C., Watson H.C. 1965. Structure and function of haemoglobin: 2. Some relations between polypeptide chain configuration and amino acid sequence. *J. Mol. Biol*. **13**, 669–678.
2. Lee, B., Richards F.M. 1971. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol*. **55**, 379–400.
3. Chothia C. 1976. The nature of the accessible and buried surfaces in proteins. *J. Mol. Biol*. **105**, 1–14.
4. Barlow D.J., Thornton J.M. 1983. Ion-pairs in proteins. *J. Mol. Biol*. **168**, 867–885.
5. Rashin A.A., Honig B. 1984. On the environment of ionizable groups in globular proteins. *J. Mol. Biol*. **173**, 515–521.
6. Gonzales L., Jr., Woolfson D.N., Alber T. 1996. Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nature Struct. Biol*. **3**, 1011–1018.
7. Oakley M.G., Kim P.S. 1998. A buried polar interaction can direct the relative orientation of helices in a coiled coil. *Biochemistry.* **37**, 12603–12610.
8. Lumb K.J., Kim P.S. 1995. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry.* **34**, 8642–8648.
9. Kondratova M.S., Efimov A.V. 2002. A systematic analysis of buried polar side chains and their interactions in α-helical proteins. *Mol. Biol*. **36**, 144–151.
10. Efimov A.V., Kondratova M.S. 2003. Comparative analysis of interhelical polar interactions in proteins with differently packed α-helices. *Mol. Biol*. **37**, 515–521.
11. Murzin A.G., Brenner S.E., Hubbard T., Chothia C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol*. **247**, 536–540.
12. Hooft R.W.W., Sander C., Vriend G. 1996. Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins*. **26**, 363–376.
13. Efimov A.V., Brazhnikov E.V. 2003. Relationship between intramolecular hydrogen bonding and solvent accessibility of side-chain donors and acceptors in proteins. *FEBS Lett*. **554**, 389–393.