

Analysis Of Torsion Angles Between Helical Axes in Pairs of Helices in Protein Molecules

Tikhonov D.A.*¹, Kulikova L.I.**¹, Efimov A.V.²

¹*Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Pushchino, Russia*

²*Institute of Protein Research, Russian Academy of Sciences, Pushchino, Russia*

Abstract. In this study, an analysis of distribution of the torsion angles Ω between helical axes in pairs of connected helices found in known proteins has been performed. The database for helical pairs was compiled using the Protein Data Bank taking into account the definite rules suggested earlier. The database was analyzed in order to elaborate its classification and find out novel structural features in helix packing. The database was subdivided into three subsets according to criterion of crossing helix projections on the parallel planes passing through the axes of the helices. It was shown that helical pairs not having crossing projections are distributed along the whole range of angles Ω , although there are two maxima at $\Omega = 0^\circ$ and $\Omega = 180^\circ$. Most of helical pairs of this subset are pairs formed by α -helices and 3_{10} -helices. It is shown that the distribution of all the helical pairs having the crossing helix projections has a maximum at $20^\circ < \Omega < 25^\circ$. In this subset, most helical pairs are formed by α -helices. The distribution of only α -helical pairs having crossing axes projections has three maxima, at $-50^\circ < \Omega < -25^\circ$, $20^\circ < \Omega < 25^\circ$, and $70^\circ < \Omega < 110^\circ$.

Key words: structural motifs of proteins, point model, helical pairs in proteins, torsion angle between the axes of helices.

INTRODUCTION

This paper as well as [1, 2] deals with recognition and analysis of various structural motifs formed of two neighboring helices connected by one or more nonhelical amino acid residues. The protein molecules to be analyzed were taken from the Protein Data Bank (PDB) [3].

The rules for selecting helical pairs from the database of protein structures have been described in [1, 2] where helical pairs formed by α -helices, 3_{10} -helices and π -helices in different combinations have been studied. It was shown that in the total number of helical pairs there are 72.16 % of α -helices, 27.73 % of 3_{10} -helices, and 0.1 % of π -helices. Although α -helices prevail (72.16 %), the helical pairs consisting of two α -helices account for only 54.7 % of the total number of helical pairs. The helical pairs formed of 3_{10} -helices account for 45 %, the pairs formed of π -helices account for less than 0.2 %. It was also shown that most of the helical pairs (45 %) not having crossing projections are formed of α -helices and 3_{10} -helices. In the subset of helical pairs having crossing projections, the pairs formed of two α -helices prevail (59 %). In the subset of pairs having crossing axes projections, the helical

*dmitry.tikhonov@gmail.com

**likulikova@mail.ru

pairs formed of two α -helices are predominant (87.5 %). The number of helical pairs not having crossing projections is equal to the number of helical pairs having crossing projections, each contributing 44 % to the total number of helical pairs. The rest 12 % are accounted for by the structures having crossing axes projections. This means that most of helical pairs in proteins (56 %) have crossing projections.

In the above-cited papers, we studied interhelical distances [1], angles between the helical axes, and the number of amino acid residues between the helices [2] in the helical pairs selected in accordance with special rules suggested for selecting helical pairs from the proteins available in the Protein Data Bank. It was shown that the distribution of all the types of helical pairs not having crossing projections covers practically the whole range of the angles with one maximum in the region of the right angle. Helical pairs formed of two α -helices prevail in the subset of pairs having crossing axes projections. Most of these pairs have an acute angle ($20^\circ \leq \varphi \leq 50^\circ$) between the helical axes. The interplane distance for all these structures is equal to 10 Å. The distribution of all the types of helical pairs belonging to the three subsets was analyzed for the length of the connections. It was shown that in all the subsets the structures with a short (from one to seven amino acid residues) connections prevail. We also demonstrated that in the subset of structures formed of two α -helices and having crossing axes projections, the structures with a connection consisting of five amino acid residues are predominant.

In this study, we investigate torsion angles Ω between helical axes in the helical pairs of proteins. The analysis is made for interhelical interactions between neighboring helices connected by connections of different lengths and conformations. In other papers [4–8], investigations were carried out for interhelical interactions in proteins regardless the number of amino acid residues between the helices and without considering their conformations. Our research is performed for a set of helical pairs of proteins registered in the PDB not taking into account the protein classifications. As a rule, such classifications are based on the features of the secondary and tertiary structure of proteins [9]. In the future we are planning to study structural features of helical pairs in each class of proteins.

Compilation of such databases and their investigations are of particular value as it enables to researchers to find out novel structural motifs having unique overall folds of the polypeptide chain and new features of protein structure [10]. The structural motifs formed by two neighboring α -helices connected by relatively short connections are described in [11–13] and are shown to be compact spatial structures. It is also known from literature that α -helices pack in one of three characteristic arrangements, aligned parallel or antiparallel, orthogonal, or slanted. Some examples of such packings in α - α -corners, α - α -hairpins, L-shaped and V-shaped structures have been studied [12], however a comprehensive analysis of these motifs in all the proteins available in the PDB has not been made yet.

RESEARCH SUBJECT

We took from the PDB all the structural motifs formed of two helices of any type arranged one after the other in a polypeptide chain and connected by constrictions of different lengths and having different conformations. The total number of the helical pairs selected was 2206605 [1, 2]. Among them 1207742 structures are formed by two α -helices, 821798 structures by α -helices and 3_{10} -helices, and 172877 structures by two 3_{10} -helices, 3291 structures by α -helices and π -helices, 896 structures by the 3_{10} -helix and the π -helix and 1 helical pair by two π -helices. In total, we have analyzed 100397 protein structures, 66546491 amino acid residues and 384666 polypeptide chains. The large amount of the data processed provides sufficient reliability of the results.

METHODS

In order to investigate the structural motifs of proteins we elaborated special rules for recognition and selection of helical pairs. They are described in [1, 2] and given here for completeness and consistency of presentation.

Analysis of the secondary structure of the polypeptide chains is made by a method developed by the authors of the Dictionary for Secondary Structure of Protein [14] which implies helical sites. Helices of three different types have been analyzed: the first type is α -helix or H in DSSP notation, the second type (β -helix) involves 3_{10} -helices, the third type (π -helix) is π -helix.

A candidate for the structure sought-for is a protein site containing two helices of any length and conformation and a protein strand between the helices called a constriction. All such structures can also be described in terms of the helices they contain.

For each helix of the structure, we find the axis of the cylinder around which it is wound. The axis of the cylinder is determined by the least square method which implies minimizing the deviation of the helical parameters from those of an ideal helix [15, 16]. The quality of the axes assessment is characterized by the value of the root mean-square deviation. Only the helices (and accordingly, the structures) for which the accuracy of the assessments satisfies a certain predetermined criterion are selected.

Two helical axes completely determine the three dimensional arrangement of two cylinders of the helical pair. It is known that one can place two parallel planes onto two noncrossing right lines so that the first line would belong to the first plane, while the other one – to the second plane. The axis lying on one plane can be projected onto the other one. Thus, the three dimensional arrangement of the cylinders is fully described by the distance between the parallel planes and the axes projections onto the plane.

All the helical pairs selected are subdivided into three subsets according to criterion of crossing helix projections on the parallel planes passing through the axes of the helices:

- subset $\{A\}$ involves the helical pairs not having crossing projections;
- subset $\{B\}$ involves the helical pairs having crossing projections except for the helical pairs where the overlapping polygon [17] contains the cross point of the helical axes projection;
- subset $\{C\}$ involves the helical pairs for which the overlapping polygon contains the cross point of the helical axes projection.

Point model and determination of the interhelical distances, interhelical angle and torsion angle between the axes in helical pairs

From known coordinates of the points A_1, A_2, B_1 and B_2 , which are the initial and finite points of the axes of two helices we calculate coordinates of the vectors $\overrightarrow{A_2A_1}, \overrightarrow{B_1B_2}$. The point model of the helical pair is shown in Fig. 1.

From the viewpoint of mutual arrangement of helices, three distances naturally come up. The first one is the interplane distance d . As known, it is possible to place in unique way two parallel planes onto two noncrossing right lines so that the shortest distance between the planes is the same as the distance between the lines. The distance between the right lines on which the cylinder axes lie will be called the interplane distance d of a helical pair. The second characteristics of the helical pair is the shortest distance r between the cylinder axes. Obviously, the minimum value of r will always be less than the interplane distance d , and it is equal to the interplane distance d only in the case when the projections of the helical axes of helical pairs intersect. The leg l ($l = \sqrt{r^2 - d^2}$) is the third distance which describes the relative arrangement of helical cylinders in a helical pair.

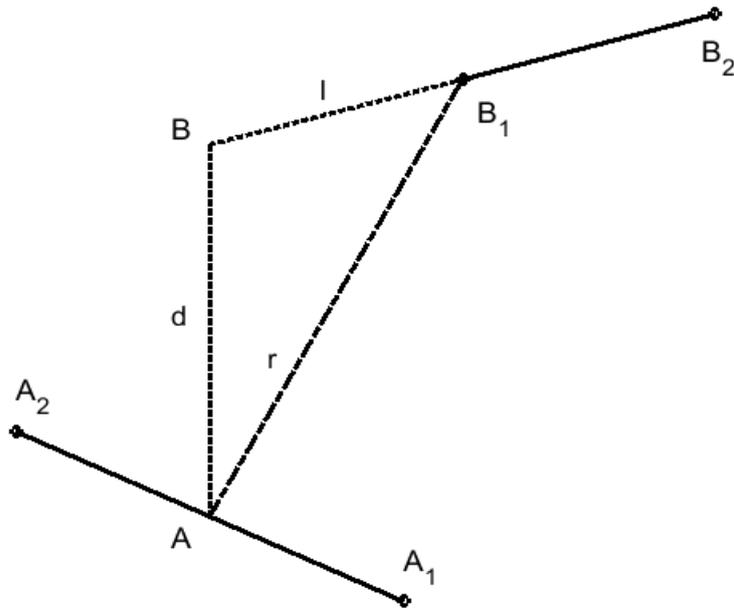


Fig. 1. Point model of a helical pair. The axes of the pair are shown. The interval $[A_1, A_2]$ is the axis of the cylinder of the first helix, $[B_1, B_2]$ is the axis of the cylinder of the second helix.

The interhelical angle φ is an angle between the vectors lying on the helical axes where the origin of the first vector is the end of the first helix and the end of the first vector is the origin of this helix, while the second vector originates at the origin of the second helix and the end of the second helix is the end of the second vector. The value of angle φ has no sign and is determined without regard for the structures chirality.

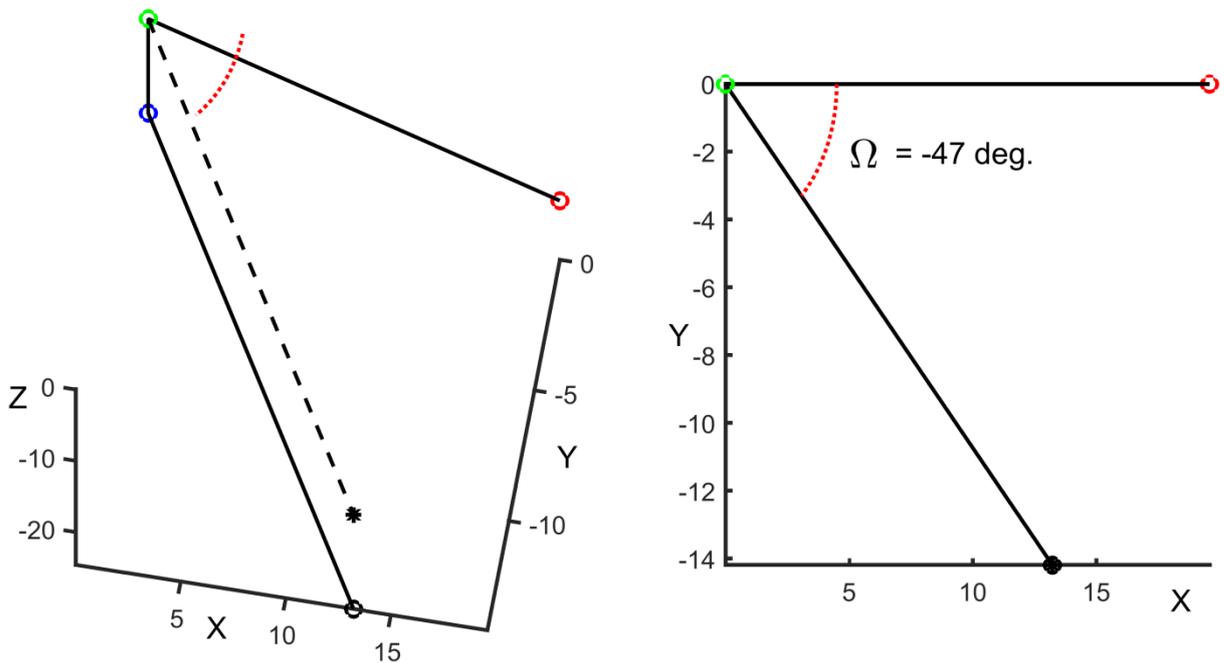


Fig. 2. Determination of the torsion angle Ω between the helical axes.

Let we have four points A_1, A_2, B_1, B_2 , which are the original and terminal points of the axes of two helices and three vectors $\overrightarrow{A_2A_1}$, $\overrightarrow{A_2B_1}$, and $\overrightarrow{B_1B_2}$. If we place points A_1, A_2 and B_1 onto the coordinate plane XY in such a way that the point A_2 be at the coordinate origin, while the point A_1 – on the positive direction of the axis $OX+$, then we have an angle between the axis $OX+$ and the projection of the vector $\overrightarrow{B_1B_2}$ (which is called a torsion angle Ω between the

helical axes) onto the plane XY . The anticlockwise direction is considered to be a positive angle ($0^\circ \leq \Omega \leq 180^\circ$) while the clockwise direction is believed to be a negative angle ($-180^\circ \leq \Omega \leq 0^\circ$).

RESULTS AND DISCUSSION

We investigated torsion angles Ω between the helical axes in the helical pairs of proteins. All recognized and selected helical pairs were subdivided into three subsets according to the criterion of crossing helix projections on the parallel planes passing through the axes of the helices [1]. Besides, all the helices selected were classified by the type of helices in the structures analyzed. For example, the helical pair HH consists of two H type helices (α -helices), HG-motifs are formed of two helices one of which is an H type helix (α -helix), while the other is a G type helix (3_{10} -helix); the helical pair of GG type is formed of two G type helices.

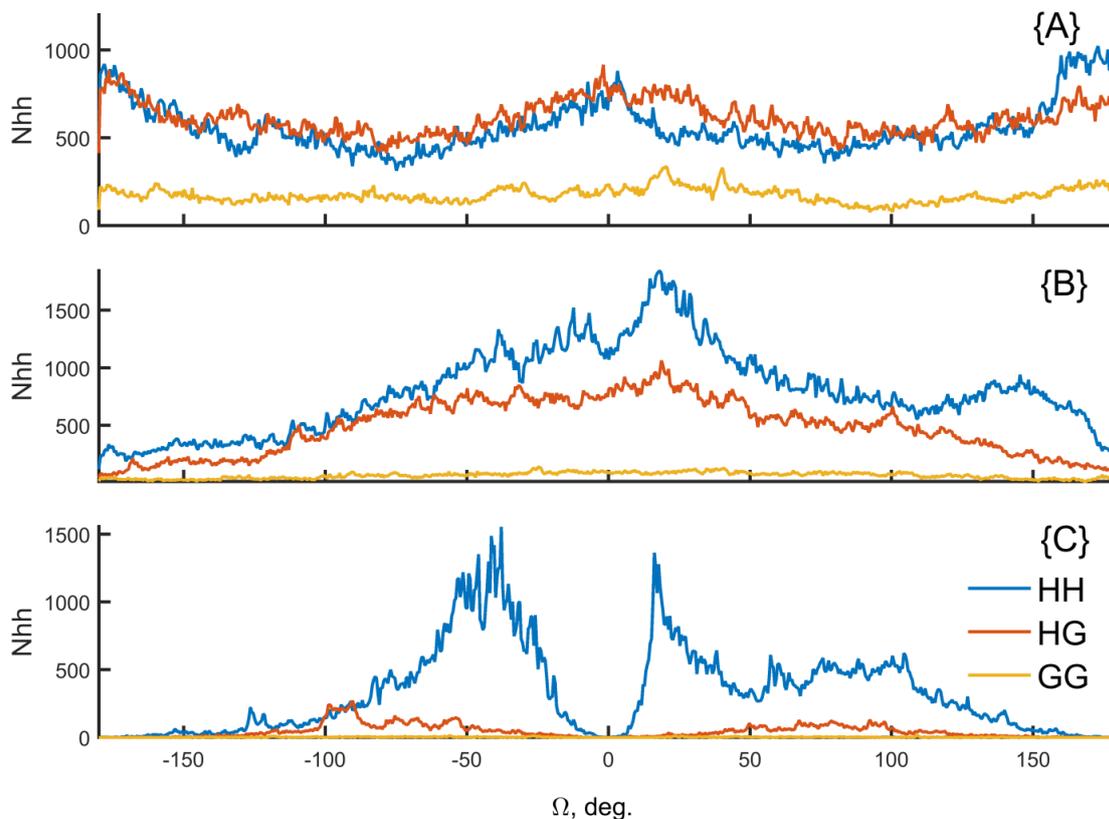


Fig. 3. Distribution of different-type structures belonging to different subsets depending on the torsion angle Ω . (along the x-axis – the torsion angle Ω s; along the y-axis – the actual number the structures selected; the blue line HH indicates the structural motifs of the proteins formed of two α -helices; the red line HG – motifs formed of α -helices and 3_{10} -helices; the yellow line GG – motifs formed of 3_{10} -helices. {A}, {B} and {C} are subsets of helical pairs selected according to the criterion of crossing helix projections).

Figure 3 illustrates the distribution of different-type structures belonging to different subsets depending on the torsion angle Ω . It is obvious from Figure 3 that the structures belonging to different subsets differ considerably. The upper graph demonstrates the distribution of the pairs belonging to subset {A} depending on the angle Ω . It is rather wide and covers nearly the whole range of Ω values. The distribution of the helical pairs of HH and HG types has two maxima at 0° and 180° . The distribution of the helical pairs of GG type is uniform over the whole range of Ω values.

The central graph shows the distribution of the pairs belonging to subset {B} depending on the angle Ω . It is seen from the graph that the structures of the GG type are few and

distributed over the whole range of Ω values. The distribution of the helical pairs of HG type has a maximum for $(20^\circ \leq \Omega \leq 25^\circ)$. In the same region $(20^\circ \leq \Omega \leq 25^\circ)$ there is a maximum of the distribution of the helical pairs consisting of two α -helices. Besides, there are three local maxima: for $(-15^\circ \leq \Omega \leq -10^\circ)$, for $\Omega \approx -30^\circ$ and for $(140^\circ \leq \Omega \leq 150^\circ)$. Prevailing in this subset are the structures formed of two α -helices.

The lower graph demonstrates the distribution of the pairs belonging to the subset having crossing axes projections depending on the torsion angle Ω . It is seen that the numbers of pairs of GG and HG type are few and they are distributed over the whole range of Ω values. The vast majority of the structures from this subset are helical pairs of HH type. The distribution of the helical pairs belonging to subset $\{C\}$ differs considerably from the distributions of the pairs belonging to subsets $\{A\}$ and $\{B\}$. These helical pairs demonstrate distribution with two maxima lying in the region of acute angles: one – in the region of negative Ω values $(-50^\circ \leq \Omega \leq -25^\circ)$, the other – in the region of positive Ω values $(20^\circ \leq \Omega \leq 25^\circ)$. The dip of the distribution corresponds to the angle $\Omega = 0^\circ$, notice that for $\Omega = 0^\circ$ there are no helical pairs at all. They are also lacking for $\Omega (-180^\circ \leq \Omega \leq -160^\circ)$ and $(160^\circ \leq \Omega \leq 180^\circ)$. It should be noted that in the region of negative Ω values the distribution of the pairs is more localized, while in the region of positive Ω values the scattering is wide and a local maximum for $(60^\circ \leq \Omega \leq 110^\circ)$ is observed.

Table 1. Statistical estimates of the distribution of torsion angles Ω in the region of negative and positive Ω values for helical pairs of different types belonging to subsets $\{B\}$ and $\{C\}$

Statistical estimates	Torsion angles Ω in the region of negative Ω values for helical pairs of different types			Torsion angles Ω in the region of positive Ω values for helical pairs of different types		
	HH	HG	GG	HH	HG	GG
mode $\{B\}$	-18.3	-30.55	-127.07	22.24	30.99	46.47
median $\{B\}$	-53.39	-57.67	-61.3	64.04	59	69.01
mean $\{B\}$	-63.78	-63.74	-69.89	73.8	66.68	73.41
std $\{B\}$	46.87	43.04	49.4	51.56	46.37	47.38
mode $\{C\}$	-37.69	-94.46	-26.89	16.25	81.13	20.87
median $\{C\}$	-49.37	-81.27	-73.18	64.86	77.55	64.59
mean $\{C\}$	-55.35	-78.53	-77.24	66.02	77.65	69.78
std $\{C\}$	27.26	26.5	37.78	37.94	28.41	37.36

Table 1 lists statistical estimates of the distribution of torsion angles Ω in the region of negative and positive Ω values for the helical pairs of different types belonging to subsets $\{B\}$ and $\{C\}$. The table presents the main parameters of the histograms shown in Figure 3.

Figure 4 illustrates the distribution of the helical pairs formed of two α -helices belonging to different subsets depending on the torsion angle Ω and the constriction length. The upper diagram demonstrates the distribution of the helical pairs belonging to subset $\{A\}$. It is clearly seen that the distribution is rather wide and covers nearly the whole range of Ω values and the whole spectrum of constrictions. This distribution demonstrates maxima in the region of negative $(-180^\circ \leq \Omega \leq -165^\circ)$ and positive $(150^\circ \leq \Omega \leq 180^\circ)$ Ω values for the constriction length equal to 1 ($N_p = 1$). There is a local maximum at 0° , while the constriction length falls on the interval $(1 \leq N_p \leq 10)$.

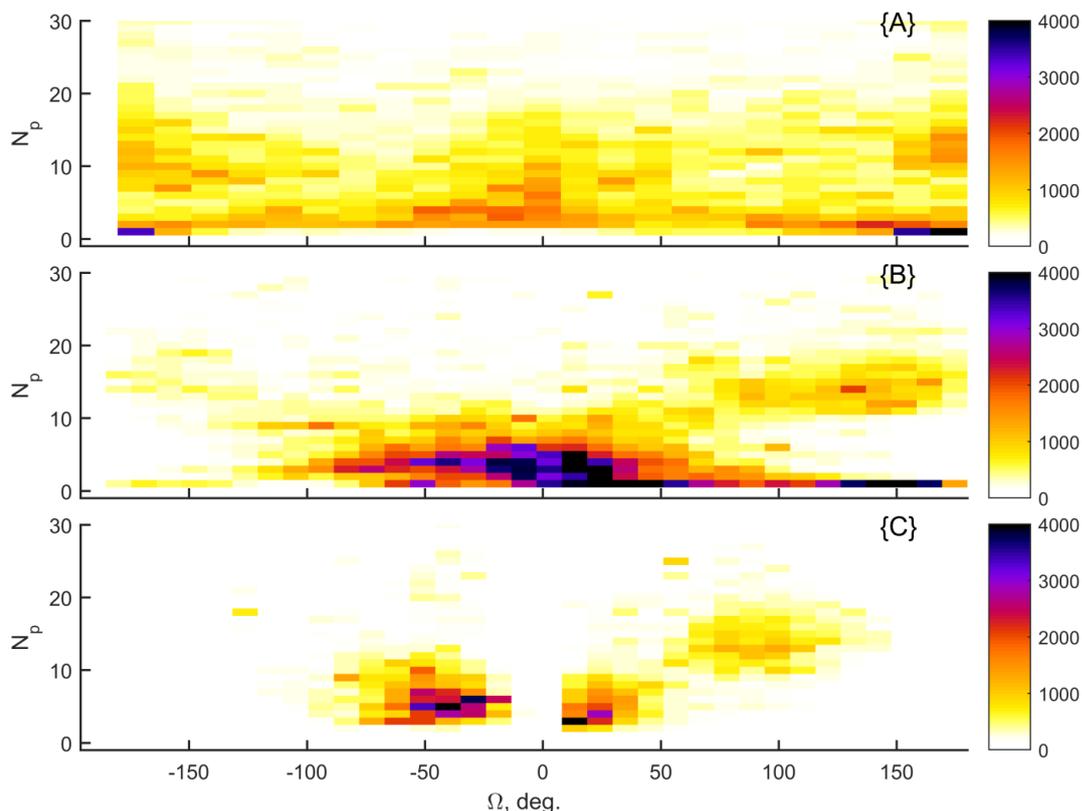


Fig. 4. Distribution of helical pairs of HH type belonging to different subsets depending on the torsion angle Ω and the constriction length N_p . Along the x-axis – the torsion angle Ω is, along the y-axis – the number of aminoacids between the helices. {A}, {B} and {C} are subsets of helical pairs selected according to the criterion of crossing helix projections. On the right-hand side there is a scale for correspondence between the color (from white to black) and the number of helical pairs.

The central diagram shows the distribution of the pairs formed of two α -helices belonging to subset {B} depending on the angle Ω and the connection length. The distribution has two maxima: on the interval $(-25^\circ \leq \Omega \leq 50^\circ)$ for the connection length $(1 \leq N_p \leq 8)$ and on the interval $(125^\circ \leq \Omega \leq 165^\circ)$ for the connection length $N_p = 1$. As seen, there are many these structures. It should be noted that the distribution demonstrates one local maximum on the interval $(125^\circ \leq \Omega \leq 165^\circ)$ for the connection length $(12 \leq N_p \leq 18)$. But these constructions are significantly fewer than the constructions whose connection consists of one amino acid residue ($N_p = 1$) and the torsion angle between the helical axes is in the region of $(125^\circ \leq \Omega \leq 165^\circ)$.

The lower diagram presents the distribution of the pairs belonging to subset having crossing axes projections depending on the angle Ω and the constriction length. It is seen that the helical pairs consisting of two α -helices are lacking in the regions of $(-180^\circ \leq \Omega \leq -100^\circ)$, $(-15^\circ \leq \Omega \leq 10^\circ)$ and $(150^\circ \leq \Omega \leq 180^\circ)$ for any connection length. It should be noted that the distribution is rather localized. A good deal of the pairs have a torsion angle Ω in the region of $(-100^\circ \leq \Omega \leq -20^\circ)$ and the connection length N_p in the range of $(3 \leq N_p \leq 7)$. A maximum is observed on the interval $(-55^\circ \leq \Omega \leq -25^\circ)$. A fairly large number of the structures have angle Ω in the region of $(10^\circ \leq \Omega \leq 50^\circ)$ and the connection length $(2 \leq N_p \leq 5)$. Maximum of the distribution is observed for $(10^\circ \leq \Omega \leq 20^\circ)$. There is also a local maximum for $(70^\circ \leq \Omega \leq 110^\circ)$ and the connection length $(10 \leq N_p \leq 18)$.

Investigating the distribution of helical pairs of HH type belonging to different subsets depending on the torsion angle Ω and the connection length N_p it may be noted that most of the pairs formed of two α -helices have a short constriction.

Figure 5 shows the distribution of the helical pairs of HH type that belong to different subsets depending of the torsion angle Ω and the interplane distance d . The upper diagram presents the distribution of the helical pairs formed of two α -helices belonging to the subset of helical pairs not having crossing projections. The distribution of these structures is rather wide and covers nearly the whole range of Ω and d values. The distribution of helical pairs of HH type demonstrates two maxima for Ω equal to 0° and 180° and the interplane distance in the range from 1 Å to 3 Å. It should be noted that for $d > 5$ Å, there are no helical pairs in the range of the above indicated Ω values.

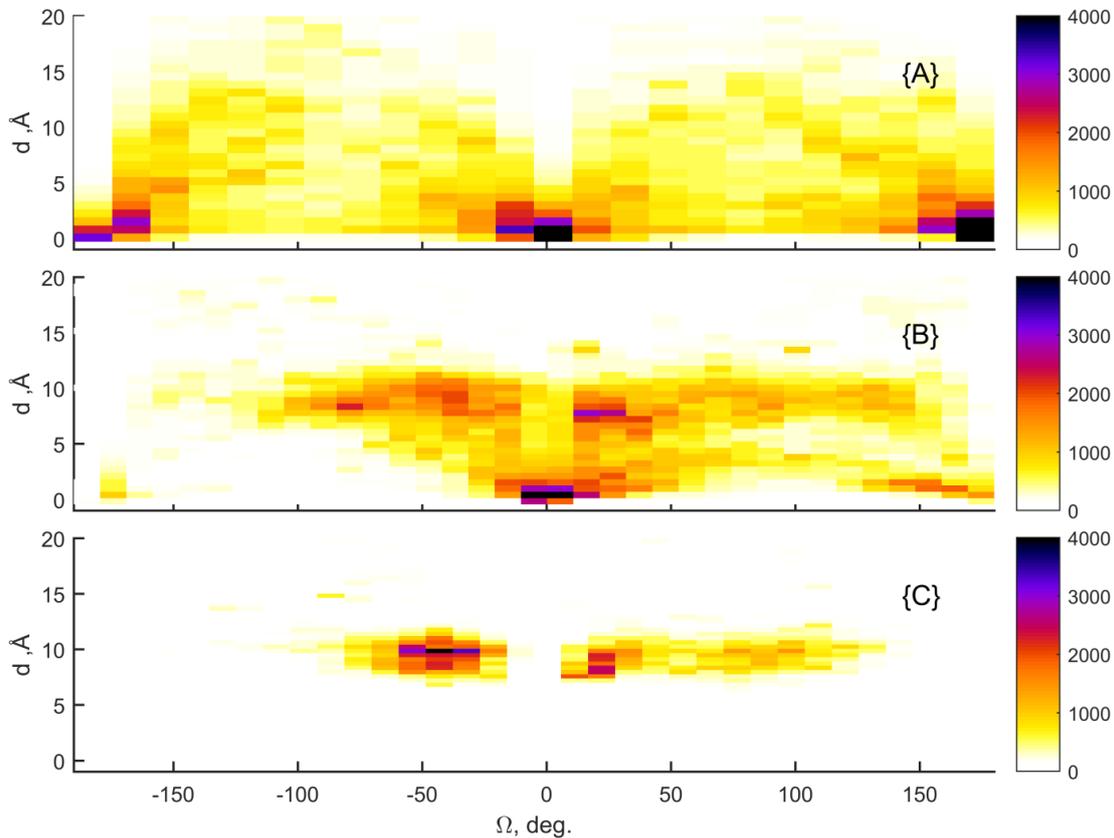


Fig. 5. Distribution of helical pairs of HH type belonging to different subsets depending on the torsion angle Ω and the interplane distance d . Along the x-axis – the torsion angle Ω , along the y-axis – the interplane distance between the helices. {A}, {B} and {C} are subsets of helical pairs selected according to the criterion of crossing helix projections. On the right-hand side there is a scale for correspondence between the color (from white to black) and the number of helical pairs.

The central diagram presents the distribution of the helical pairs consisting of two α -helices belonging to the subset of helical pairs having crossing projections depending on the torsion angle Ω and the interplane distance d . It is shown that the distribution of these structures has a maximum for Ω values in the range of $(-10^\circ \leq \Omega \leq 10^\circ)$ and the interplane distance in the range from 1 Å to 2 Å. There are also three local maxima: on the intervals $(-90^\circ \leq \Omega \leq -40^\circ)$ and $(10^\circ \leq \Omega \leq 40^\circ)$ for d varying from 7 Å to 10 Å, and on the interval $(140^\circ \leq \Omega \leq 170^\circ)$ for $d \approx 3$ Å.

The lower diagram shows the distribution of the pairs belonging to the subset having crossing axes projections depending on the angle Ω and the interplane distance d . It is clearly seen that the distribution is strongly localized. Among the pairs analyzed there are no structures having the interplane distance $d < 7$ Å and $d > 12$ Å and the torsion angle Ω in the range of $(-180^\circ \leq \Omega \leq -80^\circ)$, $(-10^\circ \leq \Omega \leq 10^\circ)$ and $(140^\circ \leq \Omega \leq 180^\circ)$. There are two maxima: in the range of negative Ω values $(-60^\circ \leq \Omega \leq -30^\circ)$ and in the range of positive Ω values ($\Omega \approx 20^\circ$). One more peak is observed in the region of $(80^\circ \leq \Omega \leq 100^\circ)$, however the

number of helical pairs with $\Omega \approx 90^\circ$ is much fewer. For all the indicated localizations of the helical pairs belonging to subset $\{C\}$, the interplane distance is in the range from 7 Å to 12 Å, the maximum corresponding to $d \approx 10$ Å, which is associated with steric constraints for the helical pairs from this subset [18–22].

In this paper we also analyzed the distribution of the helical pairs of HH type belonging to different subsets depending on the angle φ and the torsion angle Ω between the helical axes. The results of investigation of interhelical angles φ in structural motifs formed of two helices such as helical pairs of HH type were published in [3]. The interhelical angle φ is an angle between the vectors lying on the helical axes where the origin of the first vector is the end of the first helix and the end of the first vector is the origin of this helix, while the second vector originates at the origin of the second helix and the end of the second helix is the end of the second vector. The value of angle φ has no sign and is determined regardless of the structure chirality.

Figure 6 demonstrates the distribution of the helical pairs of HH type belonging to different subsets depending on the angle φ between the helical axes and the torsion angle Ω . As is clearly seen, the distributions of the structures belonging to different subsets are identical. The only difference is that the distribution of the pairs belonging to subset $\{A\}$ is wider than those of the structures belonging to subsets $\{B\}$ and $\{C\}$.

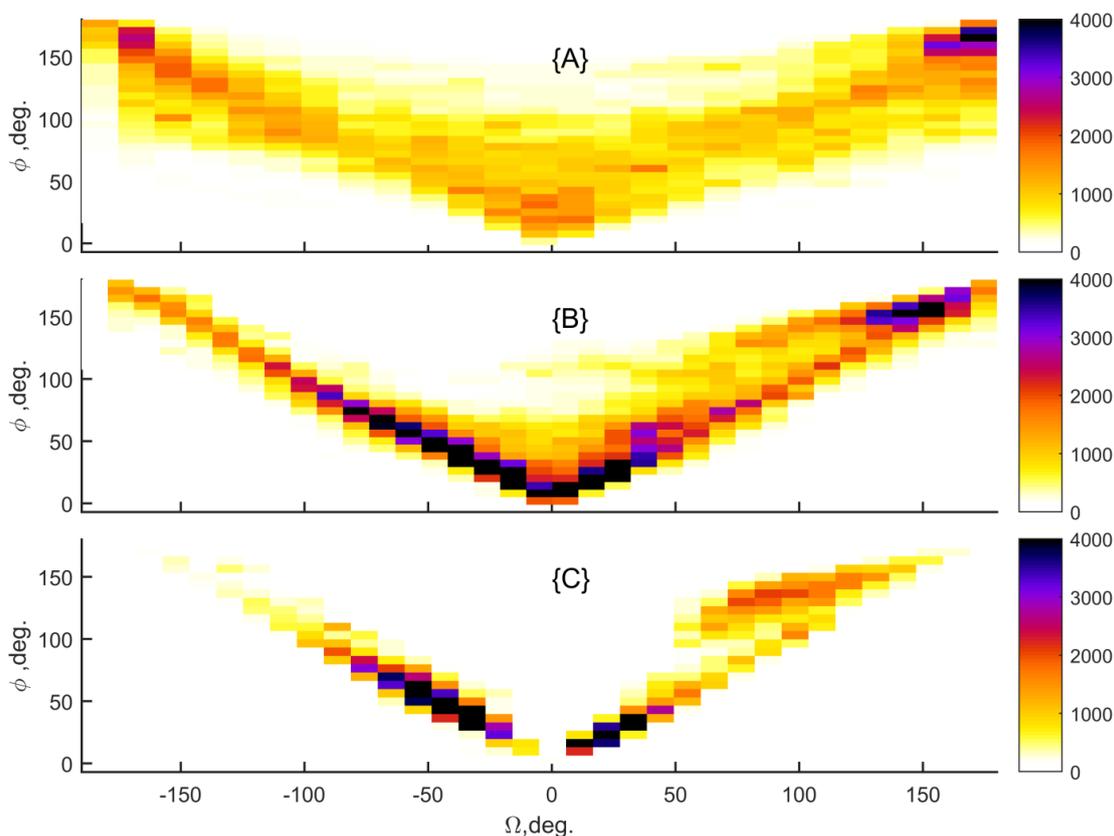


Fig. 6. Distribution of helical pairs of HH type belonging to different subsets depending on the angle φ and the torsion angle Ω between the helical axes. Along the x-axis – the torsion angle Ω , along the y-axis – the angle φ between the helical axes. $\{A\}$, $\{B\}$ and $\{C\}$ are subsets of the helical pairs selected according to the criterion of crossing helix projections. On the right-hand side there is a scale for correspondence between the color (from white to black) and the number of the helical pairs.

The upper diagram presents the distribution of the helical pairs belonging to the subset of the structures having crossing projections and formed of two α -helices depending on the angle φ between the helical axes and the torsion angle Ω . It is seen that for all the pairs, the value of the angle φ is approximate equal to the value of the torsion angle Ω . It should be noted that this is true for all the structures from to all the subsets. The distribution of the pairs belonging

to subset $\{A\}$ demonstrates two maxima: in the range of negative Ω values ($-170^\circ \leq \Omega \leq -160^\circ$) and in the range of positive Ω values ($150^\circ \leq \Omega \leq 180^\circ$), while the angle φ between the helical axes in both the cases has the values in the range of ($150^\circ \leq \varphi \leq 160^\circ$). There are also many helical pairs whose angle φ and the torsion angle Ω are approximately equal to 0° .

The central diagram shows the distribution of the helical pairs belonging to subset $\{B\}$ depending on the angle φ and the torsion angle Ω between the helical axes. This distribution is more localized and can be presented as $|\Omega| \approx \varphi$. Notice, that most of the structures are localized in the regions ($-90^\circ \leq \Omega \leq 30^\circ$) and ($30^\circ \leq \varphi \leq 90^\circ$). There is one more maximum for $\varphi \approx 150^\circ$ и $\Omega \approx 150^\circ$. It can also be seen that a fairly large number of the pairs have the angles $\varphi \approx 150^\circ$ and $\Omega \approx 90^\circ$.

The lower diagram presents the distribution of the helical pairs consisting of two α -helices belonging to the subset of helical pairs having crossing axes projections depending on the angle Ω and the angle φ between the helical axes. It should be noted that among the helical pairs analyzed there are few or no structures with Ω values in the range of ($-180^\circ \leq \Omega \leq -120^\circ$) and φ values in the range of ($120^\circ \leq \varphi \leq 180^\circ$). The structures in which the torsion angle Ω and the angle φ between the helical axes are in the vicinity of 0° ($\varphi \approx \Omega \approx 0^\circ$) are also lacking. There are two maxima in the region of acute angles: one in the regions of ($-60^\circ \leq \Omega \leq -30^\circ$) and ($30^\circ \leq \varphi \leq 60^\circ$) and the other – in the regions of ($5^\circ \leq \Omega \leq 40^\circ$) and ($10^\circ \leq \varphi \leq 45^\circ$), and one local maximum in the regions of ($70^\circ \leq \Omega \leq 110^\circ$) and ($120^\circ \leq \varphi \leq 140^\circ$).

CONCLUSION

We analyzed the distribution of the torsion angles between the helical axes in the pairs of connected helices. This paper continues our earlier work on overall investigation of helical pairs. Earlier, we collected helical pairs from the structures of proteins available in the Protein Data Bank using special rules suggested for identifying helical pairs. We suggested a point model of helical pairs. The resulting set of helical pairs was analyzed in order to elaborate its classification and find out novel structural features in helix packing. The database was subdivided into three subsets according to criterion of crossing helix projections on the parallel planes passing through the axes of the helices. In our earlier work we investigated the interhelical distances, the number of amino acid residues between the helices and the interhelical angles in the structures selected. In the future we are planning to analyze the distribution of the areas and perimeters of polygons of the helices projections intersection in pairs of connected helices found in known proteins. Each work devoted to investigation of interhelical interactions in helical pairs provides new insights into structural motifs having unique packing of the polypeptide chain. So far there are no investigations in which all the motifs available in the PDB have been analyzed.

In this study, we have analyzed the distribution of the helical pairs of different types belonging to different subsets depending on the torsion angle Ω between the helical axes. It was found that the distribution of all the helical pairs not having crossing projections covers nearly the whole range of Ω values and demonstrates two peaks at 0° and 180° . Most of these structures are the pairs consisting of α -helices and 3_{10} -helices. In the subset of helices having crossing projections, the pairs formed of two α -helices prevail. We also demonstrated that the distribution of the structures belonging to the subset of helices having crossing projections has a maximum at ($20^\circ \leq \Omega \leq 25^\circ$). Besides, we have found that the helical pairs formed by two α -helices prevail in the subset of helical pairs having crossing axes projections. These helical pairs demonstrate distribution with two maxima lying in the region of acute angles: one – in the region of negative Ω values ($-50^\circ \leq \Omega \leq -25^\circ$), the other – in the region of positive Ω values ($20^\circ \leq \Omega \leq 25^\circ$) and a local maximum for ($60^\circ \leq \Omega \leq 110^\circ$).

We also analyzed the distribution of the helical pairs formed of α -helices belonging to different subsets depending of the torsion angle Ω and the interplane distance d . It was shown that the distribution of the structures formed of two α -helices not having crossing projections demonstrates a maximum for small values of d : from 1 Å to 3 Å and the torsion angle $\Omega \approx 0^\circ$. The distribution of the pairs having crossing projections has one more maximum at 180° . Among the structures having crossing projections, there are a lot of pairs whose torsion angles lie in the range of $(-90^\circ \leq \Omega \leq -40^\circ)$ and $(10^\circ \leq \Omega \leq 40^\circ)$ for d varying from 7 Å to 10 Å. We found that for all the pairs having crossing axes projections, the interplane distance d varies from 7 Å to 12 Å, reaching maximum at $d \approx 10$ Å. The torsion angles of these structures have three localizations: $(-60^\circ \leq \Omega \leq -30^\circ)$, $\Omega \approx 20^\circ$ and $\Omega \approx 90^\circ$.

The analysis of the distribution of the pairs formed of two α -helices belonging to different subsets depending on the torsion angle Ω and the connection length revealed that the vast majority of the helical pairs having crossing projections have a short connection (N_P). The distribution of the structures formed of two α -helices having crossing axes projections demonstrates two maxima: in the region of negative Ω values $(-55^\circ \leq \Omega \leq -25^\circ)$ and the connection length varying from 3 to 7 amino acid residues and the other – in the region of positive Ω values $(10^\circ \leq \Omega \leq 20^\circ)$ and the connection length varying from 2 to 5 amino acid residues and one local peak in the region of $(70^\circ \leq \Omega \leq 110^\circ)$, the interhelical distance varying from 10 to 18 amino acid residues.

The study was supported by the RFBR (projects № 16-01-00692 and № 18-07-01031).

REFERENCES

1. Tikhonov D.A., Kulikova L.I., Efimov A.V. Statistical analysis of the internal distances of helical pairs in protein molecules. *Mathematical Biology and Bioinformatics*. 2016. V. 11. No. 2. P. 170–190 (in Russ.). doi: [10.17537/2016.11.170](https://doi.org/10.17537/2016.11.170)
2. Tikhonov D.A., Kulikova L.I., Efimov A.V. The study of interhelical angles in the structural motifs formed by two helices. *Mathematical Biology and Bioinformatics*. 2017. V. 12. No. 1. P. 83–101 (in Russ.). doi: [10.17537/2017.12.83](https://doi.org/10.17537/2017.12.83)
3. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. The Protein Data Bank. *Nucleic Acids Research*. 2000. V. 28. P. 235–242. doi: [10.1093/nar/28.1.235](https://doi.org/10.1093/nar/28.1.235)
4. Crick F.H.C. The Packing of α -helices: simple coiled-coils. *Acta Crystallographica*. 1953. V. 6. P. 689–697. doi: [10.1107/S0365110X53001964](https://doi.org/10.1107/S0365110X53001964)
5. Lee H.S., Choi J., Yoon S. QHELIX: A Computational tool for the improved measurement of inter-helical angles in proteins. *Proteins*. 2007. V. 26. P. 556–561. doi: [10.1007/s10930-007-9097-9](https://doi.org/10.1007/s10930-007-9097-9)
6. Walther D., Eisenhaber F., Argos P. Principles of helix-helix packing in proteins: the helical lattice superposition model. *Journal of Molecular Biology*. 1996. V. 255. P. 536–553. doi: [10.1006/jmbi.1996.0044](https://doi.org/10.1006/jmbi.1996.0044)
7. Chothia C., Levitt M., Richardson D. Structure of proteins: packing of α -helices and pleated sheets. *Proc. Natl. Acad. Sci. USA*. 1977. V. 74. P. 4130–4134. doi: [10.1073/pnas.74.10.4130](https://doi.org/10.1073/pnas.74.10.4130)
8. Chothia C., Levitt M., Richardson D. Helix to helix packing in proteins. *Journal of Molecular Biology*. 1981. V. 145. P. 215–250. doi: [10.1016/0022-2836\(81\)90341-7](https://doi.org/10.1016/0022-2836(81)90341-7)
9. Levitt M., Chothia C. Structural patterns in globular proteins. *Nature*. 1976. V. 261. P. 552–558. doi: [10.1038/261552a0](https://doi.org/10.1038/261552a0)
10. 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)

11. Gordeev A.B., Kargatov A.M., Efimov A.V. PCBOST: Protein classification based on structural trees. *Biochemical and Biophysical Research Communications*. 2010. V. 397. P. 470–471. doi: [10.1016/j.bbrc.2010.05.136](https://doi.org/10.1016/j.bbrc.2010.05.136)
12. Efimov A.V. Super-secondary structures and modeling of protein folds. In: *Methods in Molecular Biology*. Ed. Kister A.E. Clifton: Humana Press, 2013. Vol. 932. P. 177–189.
13. Brazhnikov E.V., Efimov A.V. Structure of α - α -hairpins with short connections in globular proteins. *Molecular Biology*. 2001. V. 35. No. 1. P. 89–97. doi: [10.1023/A:1004859003221](https://doi.org/10.1023/A:1004859003221)
14. Kabsch W., Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 1983. V. 22. No. 12. P. 2577–2637. doi: [10.1002/bip.360221211](https://doi.org/10.1002/bip.360221211)
15. Kabsch W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica*. 1976. V. 32. P. 922–923. doi: [10.1107/S0567739476001873](https://doi.org/10.1107/S0567739476001873)
16. Kabsch W. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica*. 1978. V. 34. P. 827–828. doi: [10.1107/S0567739478001680](https://doi.org/10.1107/S0567739478001680)
17. Legland D. *MatGeom: Matlab geometry toolbox for 2D/3D geometric computing*. URL: <http://github.com/dlegland/matGeom> (accessed 11 May 2017).
18. Calhoun J.R., Kono H., Lahr S., Wang W., DeGrado W.F., Saven J.G. Computational design and characterization of a monomeric helical dinuclear metalloprotein. *Journal of Molecular Biology*. 2003. V. 334. No. 5. P. 1101–1115. doi: [10.1016/j.jmb.2003.10.004](https://doi.org/10.1016/j.jmb.2003.10.004)
19. Calhoun J.R., Natri F., Maglio O., Pavone V., Lombardi A., DeGrado W.F. Artificial diiron proteins: From structure to function. *Peptide Science*. 2005. V. 80. No. 2–3. P. 264–278. doi: [10.1002/bip.20230](https://doi.org/10.1002/bip.20230)
20. Chino M., Maglio O., Natri F., Pavone V., DeGrado W.F., Lombardi A. Artificial diiron enzymes with a de novo designed four-helix bundle structure. *European Journal of Inorganic Chemistry*. 2015. V. P. 3371–3390. doi: [10.1002/ejic.201500470](https://doi.org/10.1002/ejic.201500470)
21. Chino M., Leone L., Maglio O., Lombardi A. Designing Covalently Linked Heterodimeric Four-Helix Bundles. *Methods in enzymology*. 2016. V. 580. P. 471–499. doi: [10.1016/bs.mie.2016.05.036](https://doi.org/10.1016/bs.mie.2016.05.036)
22. Trovato A., Seno F. A new perspective on analysis of helix-helix packing preferences in globular proteins. *Proteins: structure, function, bioinformatics*. 2004. V. 55. P. 1014–1022. doi: [10.1002/prot.20083](https://doi.org/10.1002/prot.20083)

Received 24 April 2018.

Published 30 April 2018.