

Internal distances of helical pairs in protein molecules

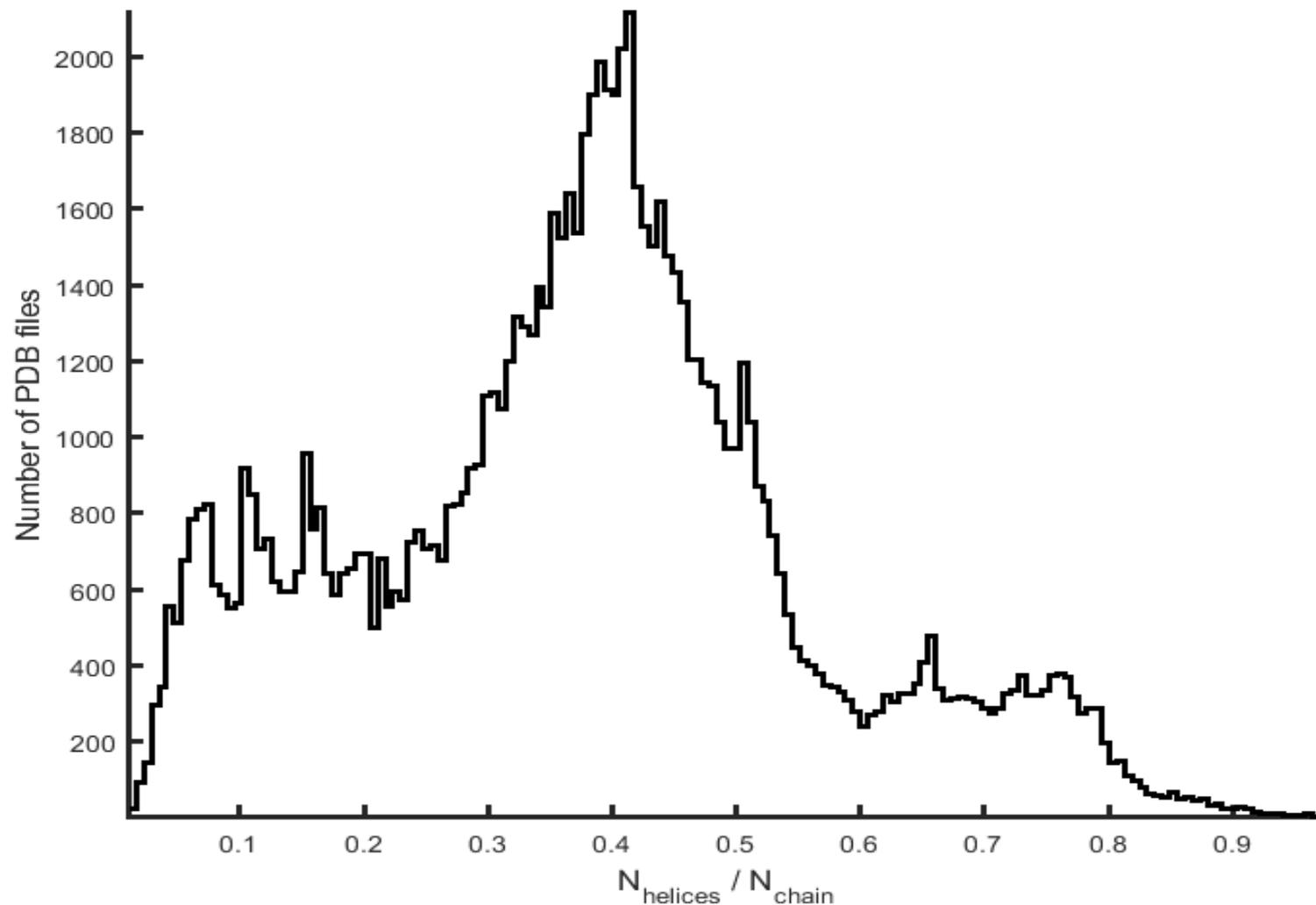
Kulikova L.I.¹, Tikhonov D.A.¹, Efimov A.V.²

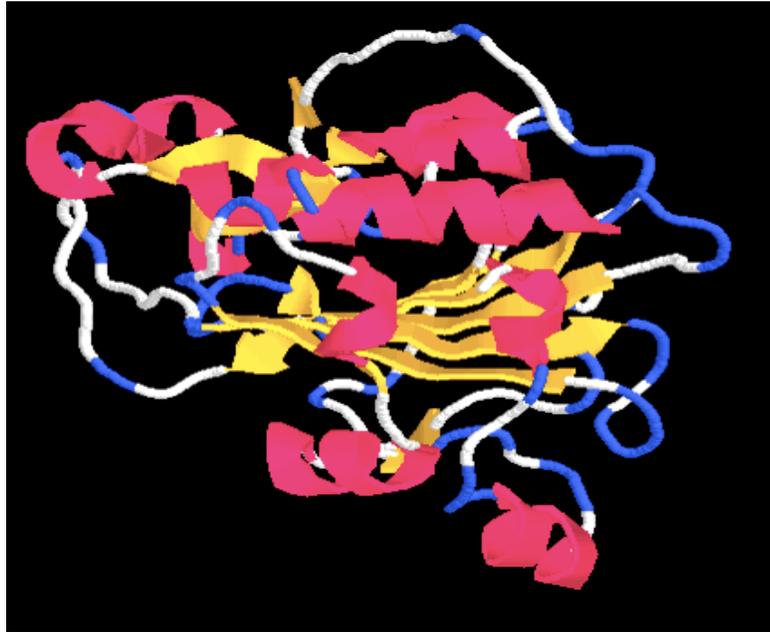
¹Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Pushchino, Russia

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

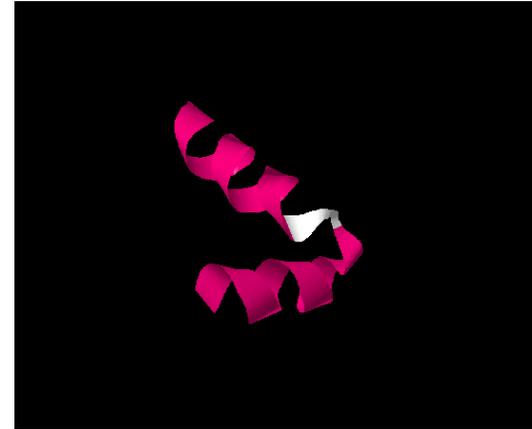
Distribution of the number of proteins in PDB depending on the ratio of amino acids in the helices to the total number of amino acids in the protein

*11th International Conference on Intelligent Data Processing: Theory and Applications (IPD-2016)
Barcelona, Spain, 10-14 October, 2016*

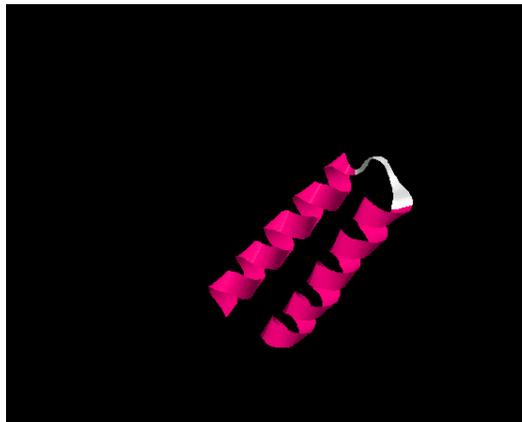




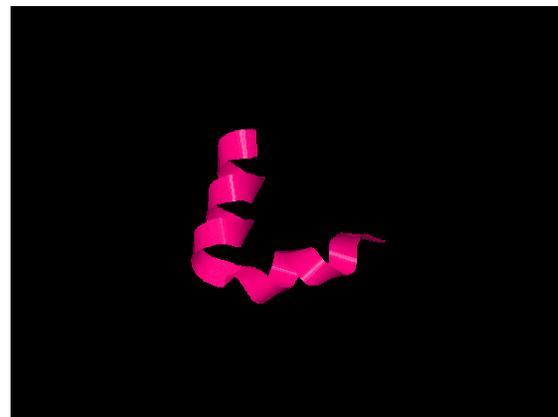
α - α -corner



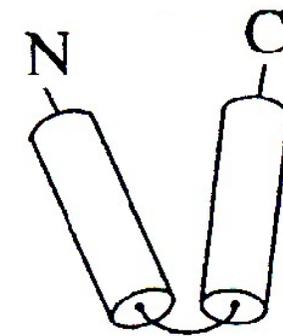
α - α -hairpins



L- shaped structure



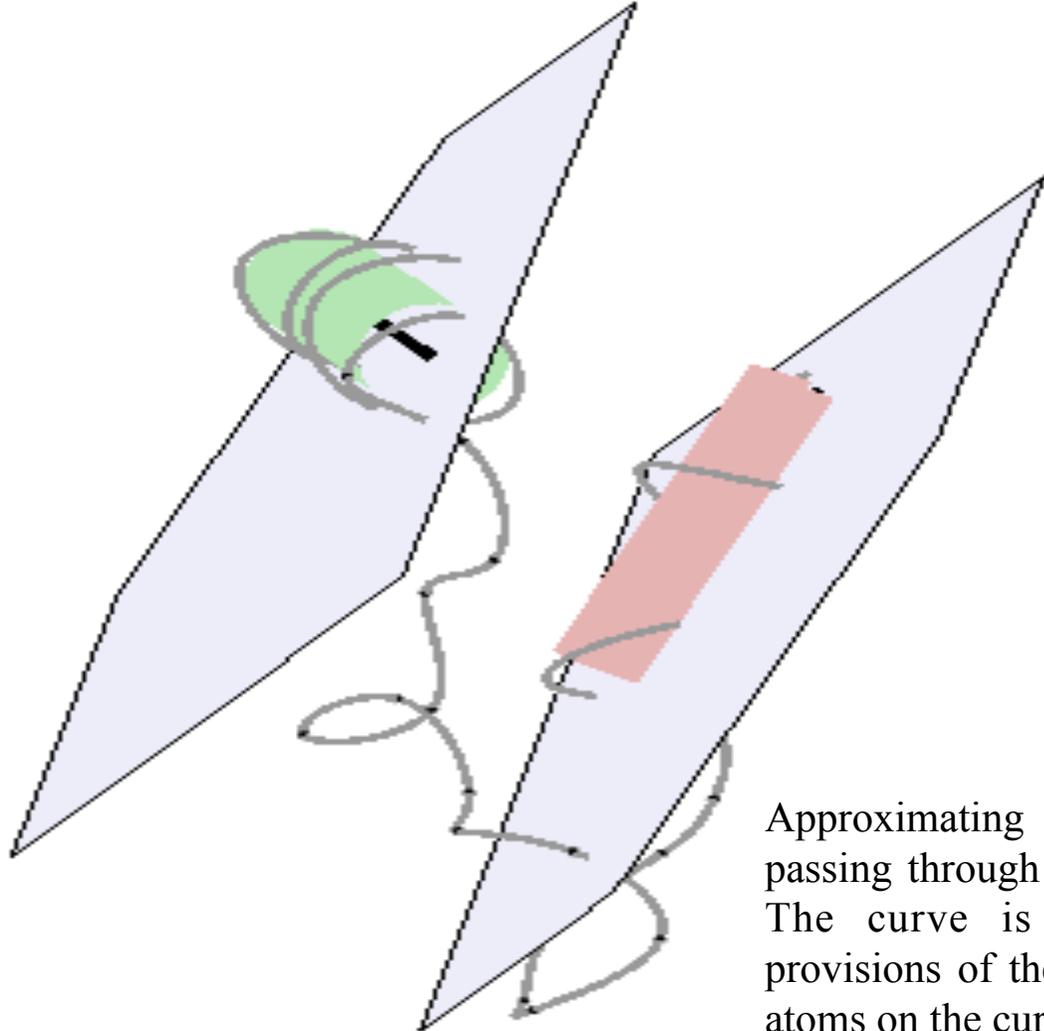
V- shaped structure



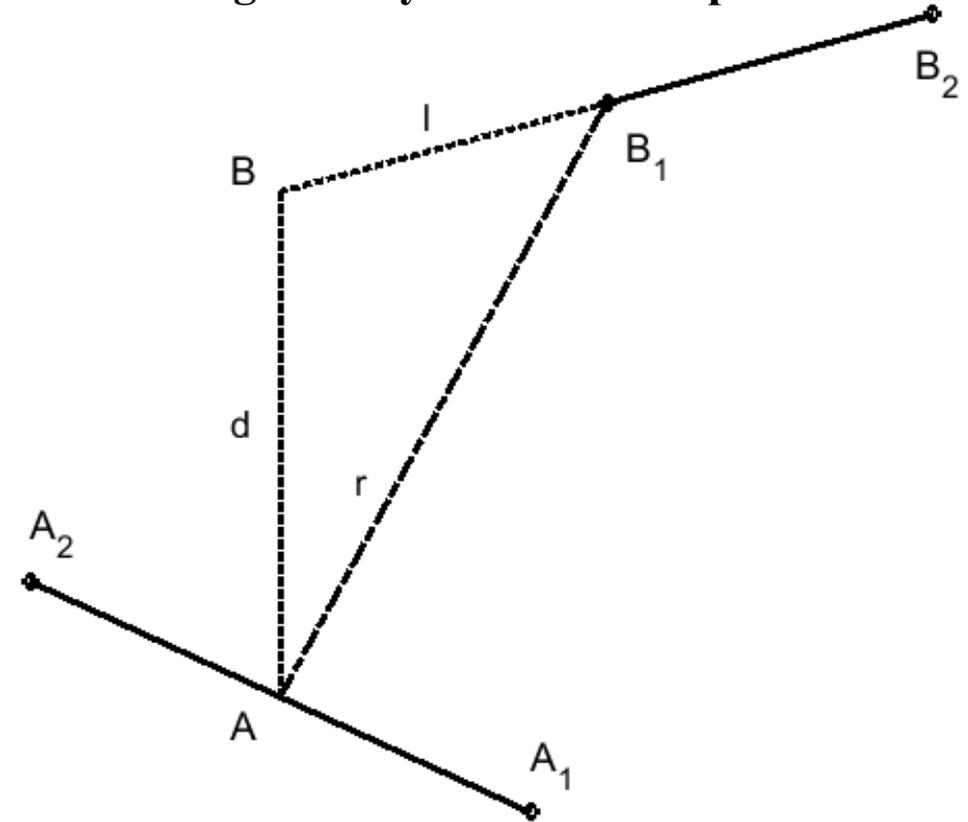
In the present study we delivered two basic tasks:

- select from the protein data bank PDB all structural motifs formed by two helices of any type located in the polypeptide chains one after another and connected by a by constrictions of varying length and conformation, and create a database for the further studies of such two-helical structures;
- an analysis of the distribution between the helical distances in pairs included in the database.

An example of a spiral pair. A fragment of a protein chain from protein data Bank (PDB 3A0B ID, the coordinates of the site: 1000-1037).

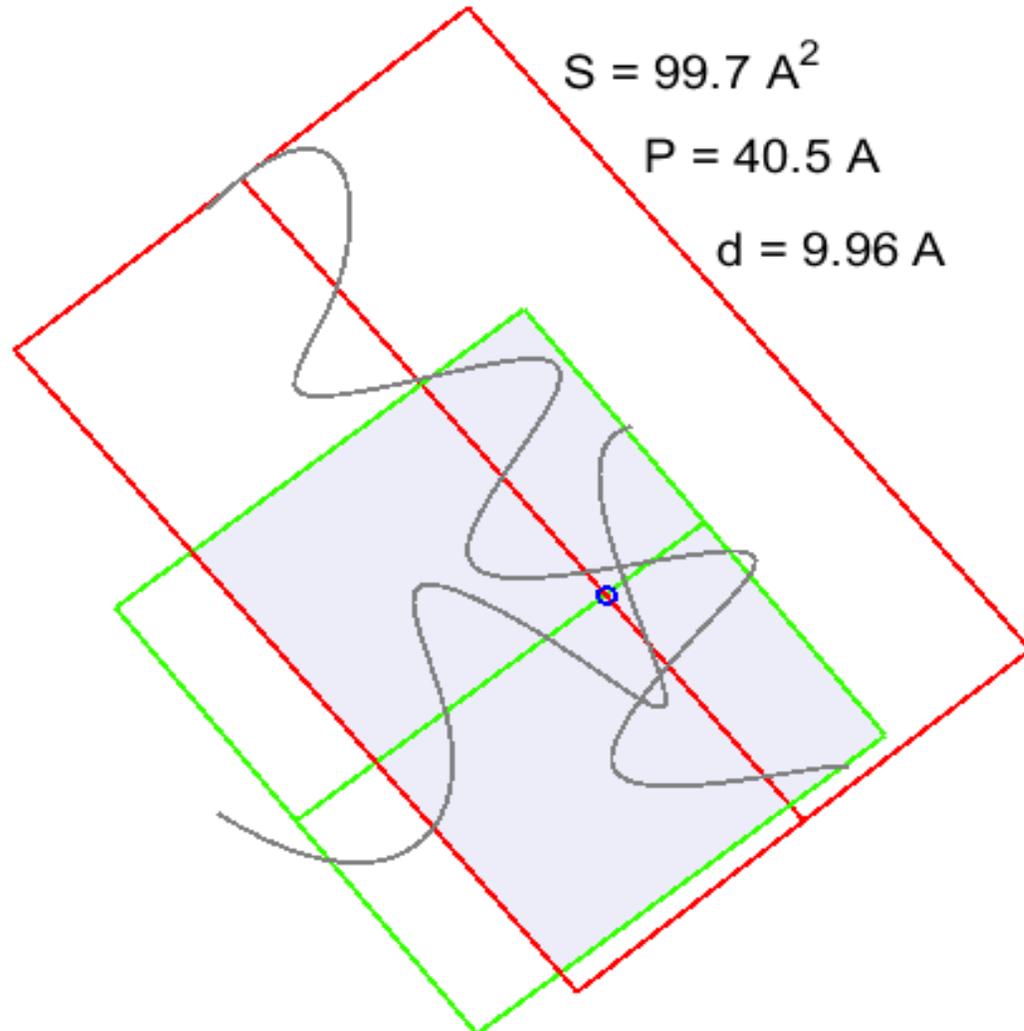


The geometry of the helical pairs

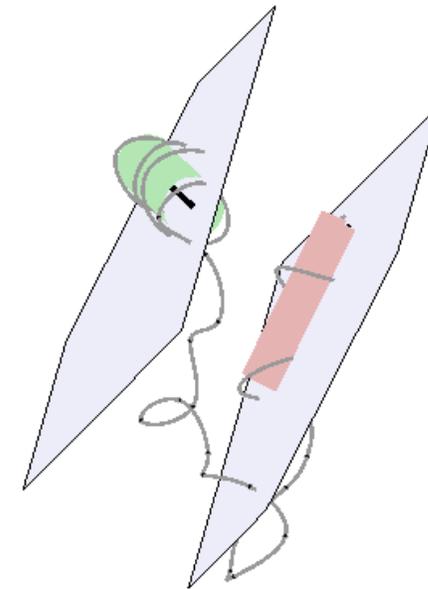


Approximating the helix cylinders and the plane passing through the axis of the cylinders are shown. The curve is approximated according to the provisions of the C α atoms of the protein chain, the atoms on the curve are indicated by dots.

**The intersection of the projections of the cylinder
of helices of the helical pair. Polygon of helices
projections intersection for the helical pair (PDB
ID 3A0B, the coordinates of the site: 1000-1037)**



An example of a spiral pair. A fragment of a protein chain from protein data Bank (PDB 3A0B ID, the coordinates of the site: 1000-1037).



- 1) - {A}**
- 2) - {B}**
- 3) - {C}**

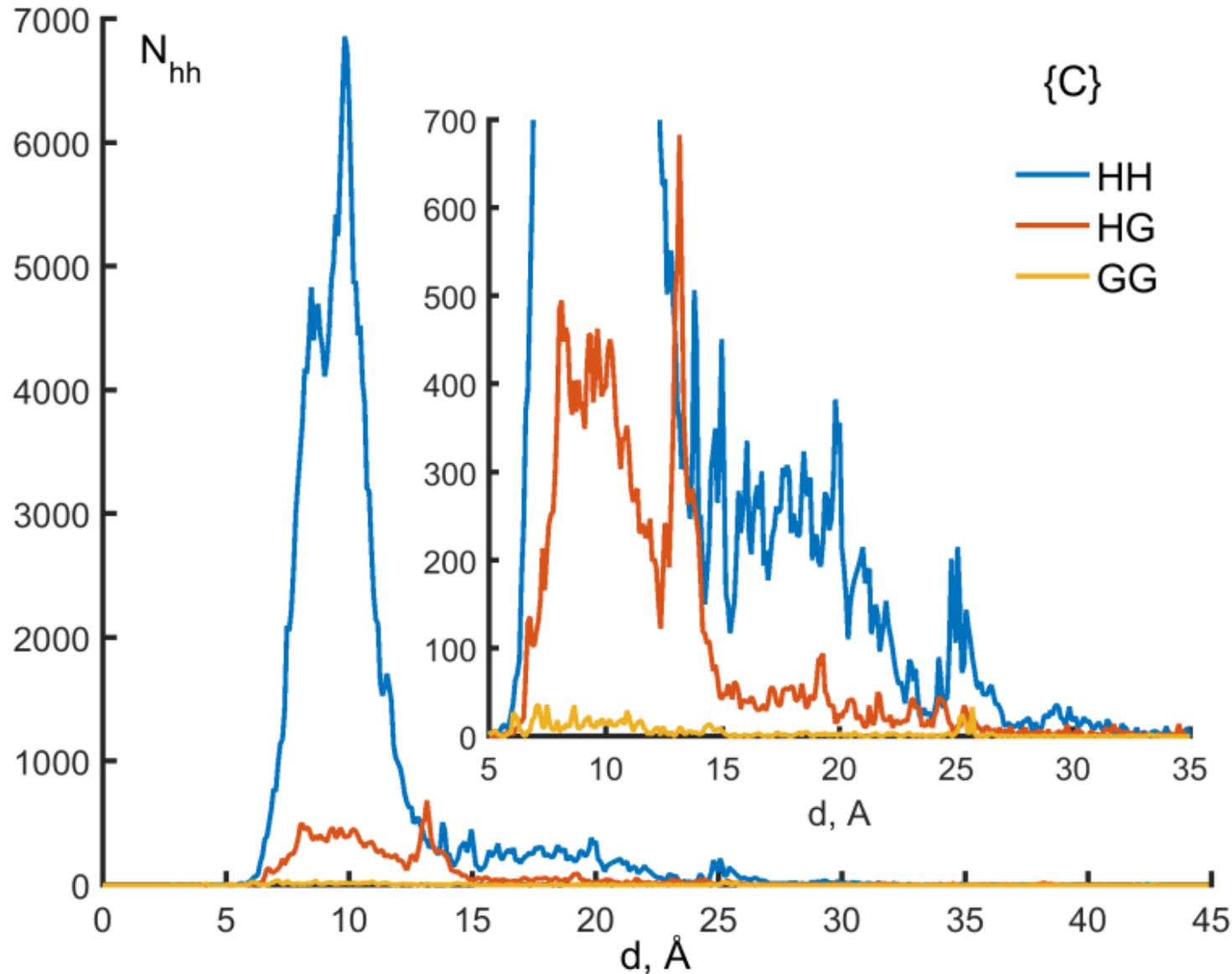
Total information

*11th International Conference on Intelligent Data Processing: Theory and Applications (IPD-2016)
Barcelona, Spain, 10-14 October, 2016*

The number of processed protein structures PDB	The number of the processed of amino acid residues	The number of processed amino acid chains	Number of H-type helices	Number of G-type helices	Number of I-type helices
100397	66546491	384666	1952658 <i>(72.16%)</i>	750605 <i>(27.73%)</i>	2908 <i>(0.1%)</i>

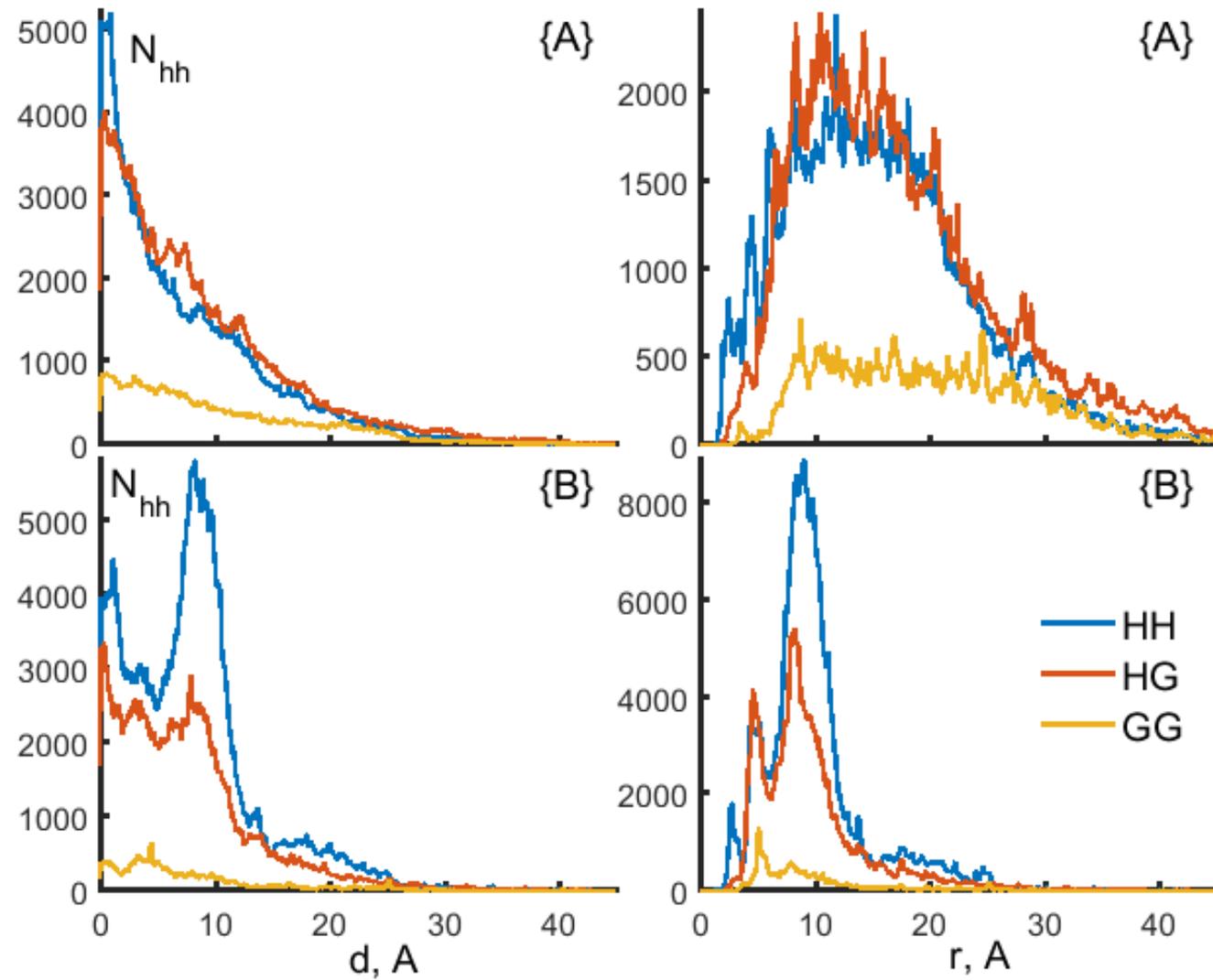
The sets of helical pairs	Types of helical pairs						Number of sets of elements
	HH	HG	GG	HI	GI	II	
{A}	402912	441055	125766	1588	643	0	971964
{B}	570830	349024	45513	1677	244	1	967289
{C}	234000	31719	1598	26	9	0	267352
The total number of helical pairs by types	1207742 <i>(54.7%)</i>	821798	172877	3291	896	1	2206605

Distribution of helical pairs belonging to the set {C}, depending on the interplanar distance d

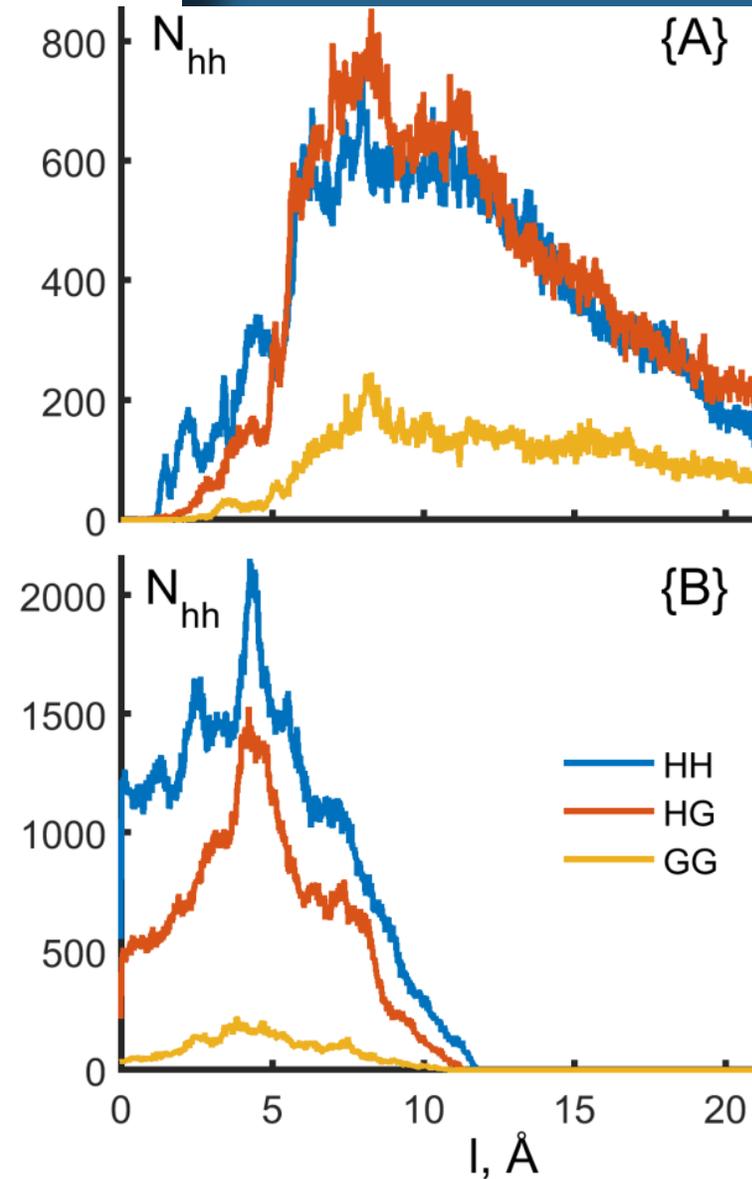
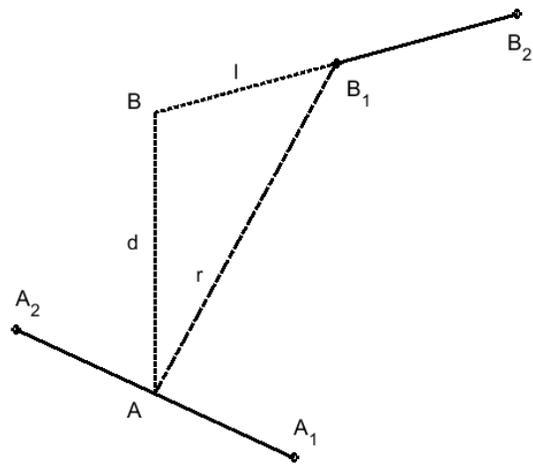


In this case, interplanar distance d equal to the minimal r

Distribution of all types of helical pairs belonging to the set {A} and {B} depending on the interplanar and minimal distances

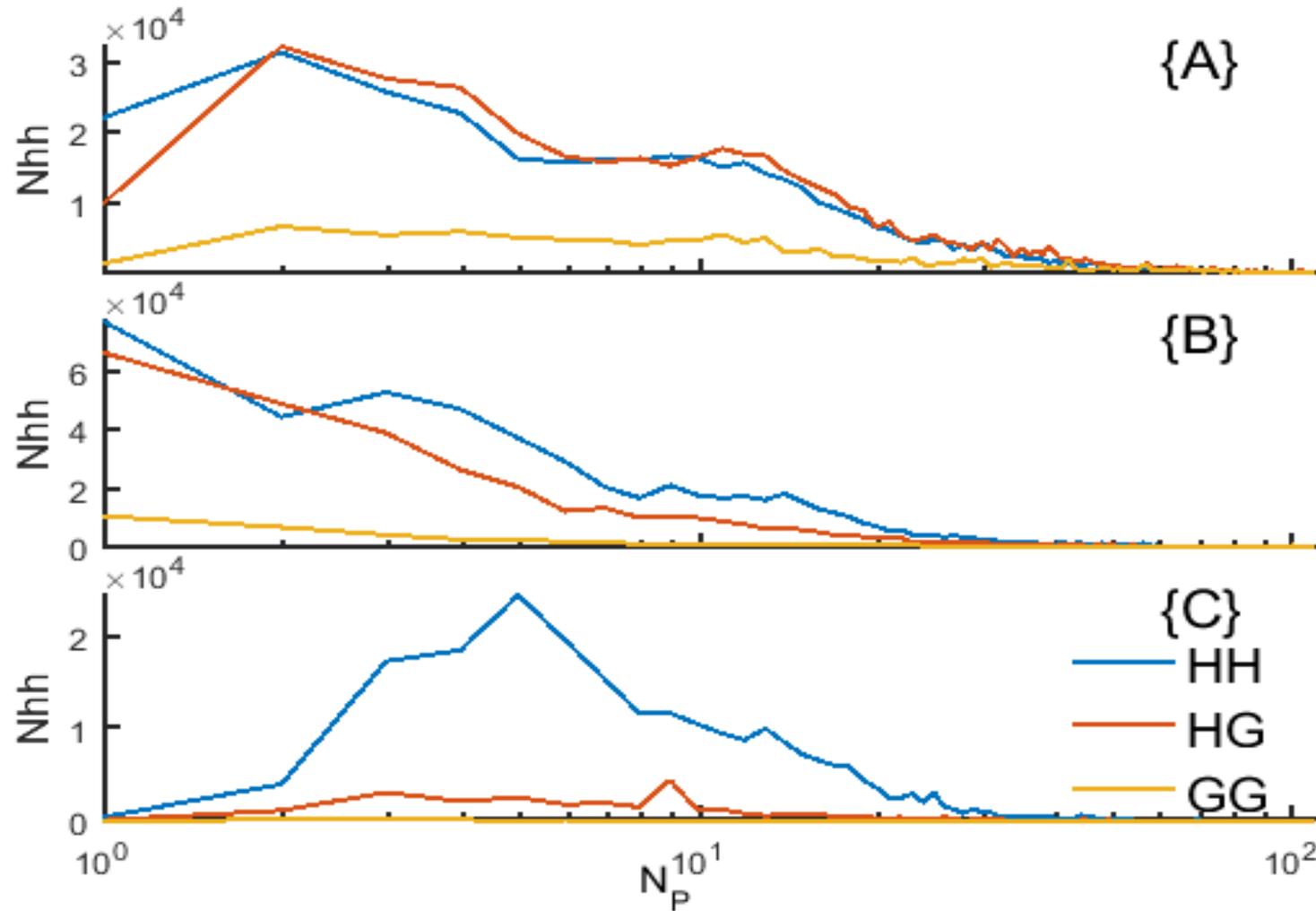


Distribution of all types of helical pairs depending on the value of leg of a triangle l



Distribution of selected structures belonging to different types and different sets, depending on the number of amino acids in the connection

*11th International Conference on Intelligent Data Processing: Theory and Applications (IPD-2016)
Barcelona, Spain, 10-14 October, 2016*



By using a point model of helical pairs, we have selected the set of researched structures of protein molecules, registered in the PDB.

This set was subdivided into three subsets according to criterion of crossing helix projections on the parallel plane passing through the axis of the helix.

Investigated a variety distances between helices in helical pairs in protein molecules.

It was shown that the distribution of distances between the pairs of helices whose projections are not crossed has a more long-range nature than those whose projections are overlapped.

The vast majority of helical pairs of HH type, belonging to the sets {C} and {B}, have the interplanar distance 10\AA .

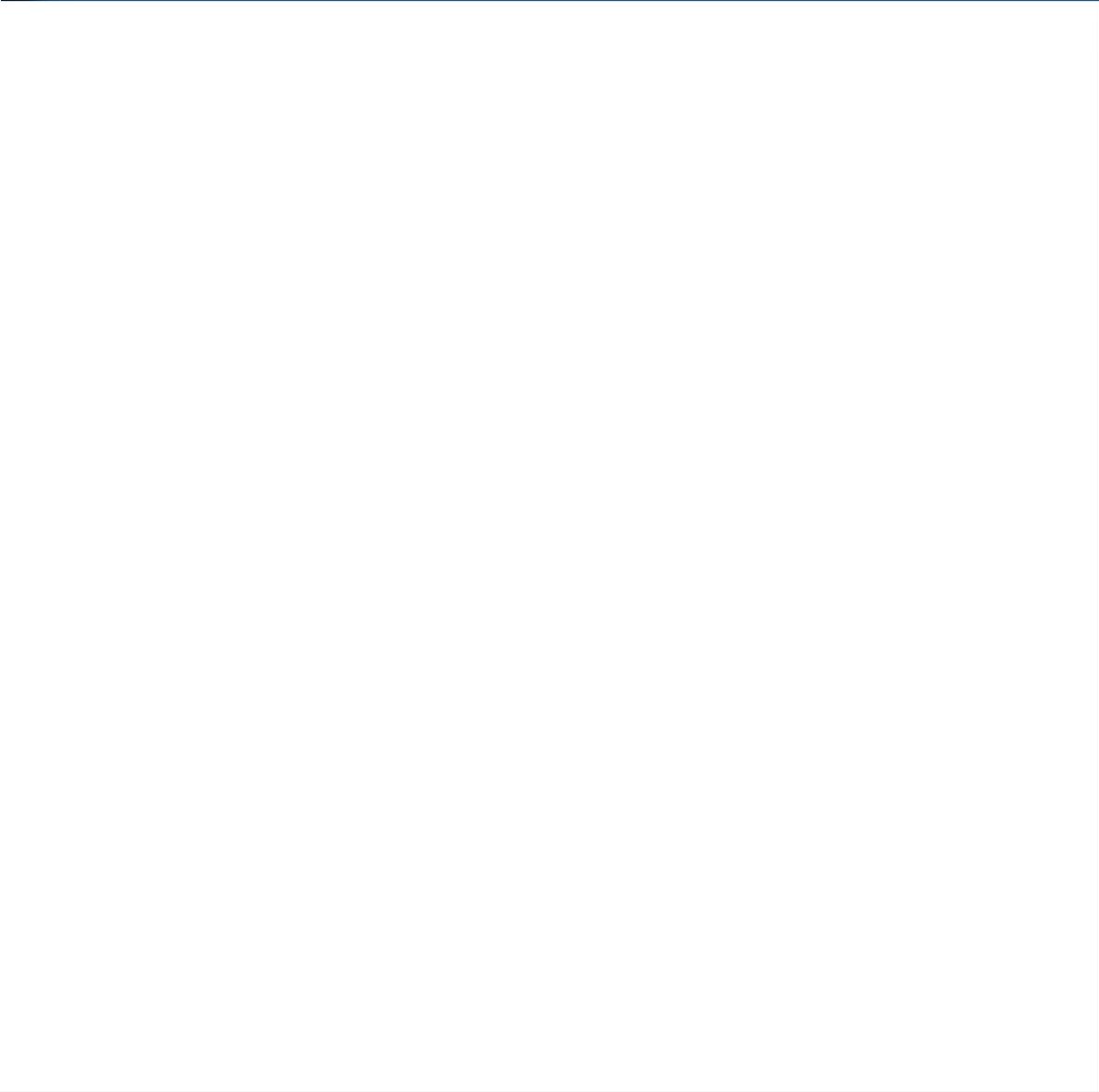
It is shown that the vast majority of all pairs of set {C} is helical pair HH (87.5%).

Amongst all selected according to formulated criteria structures, maximum number of structures with the length of the connection of 5.

It was also shown that in all sets structural motifs with short connection form the majority.

Thank you for your attention!

likulikova@mail.ru



The geometry of the helical pairs

