Research of residue contacts within the polypeptide chains depending on the density of information in the sequence

S. D. Koshman

Moscow Institute of Physics and Technology, koshman86@mail.ru

A.A.Zinchenko

Shemyakin-Ovchinnikov Institute RAS, alezina@mail.ru

A.N. Nekrasov

Shemyakin-Ovchinnikov Institute RAS, alexei_nekrasov@mail.ru

INTRODUCTION. Globular proteins themselves form compact molecular structures, the packing density of the atoms in which is close to the density of packing of atoms in crystals. We suppose that for the existence of such properties in natural polypeptide chains, there must exist both the areas in which local interactions mainly take place, and the areas in which the interaction between remote chain residues mainly takes place.

THE PROGRAM. To test this prediction, we developed a program for analyzing contacts between residues in the polypeptide chain, depending on the local density of structural information. To estimate the density of the structural information, the earlier proposed in [1] method of analysis of the information structure of protein sequences was used.

The developed program uses a set of files in PDB for input, which initially pass through a correction stage (insertions and deletions are corrected, the non-canonical residues names are restored, etc.) and then information structure of the regarded protein is calculated, and a matrix of contacts for residues belonging to the sections of the polypeptide with a different density of structural information is built.

The program is able to build a matrix of contacts for residues of various information type, i.e. located at different distances from the local extremes in the information structure of the protein under study. This paper discusses in detail the contacts between residues located in the vicinity of the extremes of the information structure, i.e. residues of different information

types. Contacts between the residues of following information types were considered:

ML2 - residue 2 positions to the left from the local minimum;

ML1 - residue in the first position to the left of the local minimum;

M - residue at position of local minimum;

MR1 - residue in the first position to the right of the local minimum;

MR2 - residue in 2 positions to the right of the local minimum;

PL2 - residue 2 positions to the left of the local maximum;

PL1 - residue in the first position to the left of the local maximum;

P - residue at the position of local maximum;

PR1 - residue in the first position to the right of the local maximum;

PR2 - residue 2 positions to the right of the local maximum.

PR3 etc.

RESULTS. The developed program allows exploring the contacts between amino acid residues depending on their information types. It is shown that the information type of residue is an important factor in determining the amino acid residue ability to interact effectively with other amino acid residues of the polypeptide chain. The matrices of contacts were shown for residues of various information types, what confirms the hypothesis of the existence of the regions of the polypeptide chain implementing mainly short- or long-distance interactions within the protein molecule. The results obtained can be widely used in protein computer modeling works and in the design of natural protein molecules.

1. A. N. Nekrasov et al. (2004) Analysis of the information structure of protein sequences: a new method for analyzing the domain organization of proteins, *J. Biomol. Struct. Dyn.*

2. A. N. Nekrasov, A. A. Zinchenko (2010) Structural features of the interfaces in enzyme-inhibitor complexes. *J Biomol Struct Dyn.* 2010 vol. 28(1):85-96