An approach for constructing a database of manually curated contacts in proteins

Pedro M Martins[1], Diego C B Mariano[1], Isabela Pastorini[1], Naiara Pantuza[1], Marcos F M Silva[1], Raquel C de Melo-Minardi[1]

[1] *LBS - Laboratory of Bioinformatics and Systems. Department of Computer Science. Federal University of Minas Gerais. Belo Horizonte, Brazil*

The comprehension of created patterns through the interaction of atoms and residues, either in proteins or other biomolecules, has been used to solve a range of problems in bioinformatics. For instance: protein folding inferences, structure prediction, functional likeness, structural alignments, thermodynamic stability, protein-ligand and protein-protein interactions, and so on. As a pdb file shows the three-dimensional structure of a protein as x, y and z coordinates of the amino acids, a pair of close residues are considered to be in contact if the distance between their specific atoms is less than a distance threshold. There are a couple of databases where contacts can be found. For instance, Piccolo is a database of structurally-characterized protein-protein interactions described at atomic level. The established contacts are calculated as follows: first, a radial cutoff search is used to identify atoms within 6.05 angstrom. After that, atom pairs are annotated with a specific type of bond depending on the atoms types, distance and the angle between them. However, it is not guaranteed that only the first layer of neighbor atoms are connected by edges and occlusions may occur, making the problem of inferring residue-residue contacts in proteins still unsolved. In this work, we propose an interface to build a manually curated database of protein-protein contacts. The tool counts on a friendly interface that allows the specialists to analyze a contact by visualizing the atoms pairs of residues in separated chains of proteins. Four distinct contact types are considered in the present study: hydrogen bond, hydrophobic interaction, ionic interaction and aromatic stacking. Hydrogen bonds occur between atoms with different electronegativities and have an important role in the enzymatic catalysis. The preference of nonpolar atoms for nonaqueous environments is known as hydrophobic interactions; in globular proteins, hydrophobic effect is important to keep the atoms in an arrangement such that, atoms with higher polarity remain on the surface of the protein and may interact with other molecules, whereas the hydrophobic atoms tend to remain within the protein. Ionic interactions happen between anions (negatively charged atoms) and cations (positively charged atoms). Finally, the aromatic stacks are attractive, noncovalent interactions between aromatic rings and play an important part in the protein folding. The interface also permits a specialist to explain the reasons why (or why not) a contact is truly established and these data may be used afterwards to predict protein-protein contacts through data mining methods.

Supported by: FAPEMIG, CNPq, and CAPES (51/2013 - 23038.004007/2014-82).