Analysis system for Protein Surface Recognition

Merelli I (1), D'Agostino D (2), Clematis A (2), Milanesi L (1)

(1) Institute of Biomedical Technologies, CNR, Milano (Italy)
(2) Institute for Applied Mathematics and Information Technology, CNR, Genova (Italy)

Motivation

The study of the protein-protein interaction is extremely complex and a method of theoretical analysis to reduce the possibility of interaction to a small dataset of protein would be extremely useful. Discarding the methods based on the analysis of the system from an energetic point of view, a typically engineering approach has been adopted for modelling the protein surface and analyzing its characteristics in order to define correlation between different macromolecules. The analysis of the superficial complementarities of two proteins is the first step to determine which proteins have the possibility to interact between them. The information contained in the morphology of a macromolecular surface is in fact the most important characterization for the definition of the protein-protein interaction. In case of complementariness the docking analysis is directed on the electrostatic characteristics and on the chemical-physical abilities of the proteins. The complementariness represents an important discriminating factor, so a precise model of the superficial characteristics of a protein complex is crucial for the definition of a valid docking.

Methods

The protein modelling always relies on its atomic coordinates, obtained by magnetic nuclear resonance analysis or by crystallography studies. By estimating the occupation volume of each atom it is possible to define the volumetric characteristics of the macromolecule. This three-dimensional grid which describes the protein is therefore analysed through a software which extracts the molecular surface establishing which points of the volume are inside the surface and which instead are outside. To accomplish this task, a high parallelizable algorithm, called Marching Cubes, has been successfully used. The output of this computation is a triangular mesh which describes through its topology the characteristics of the protein surface. But the analysis of the surfaces which has a crucial role to define structural similarities between different macromolecules is a problem that introduces remarkable difficulties. The comparison between different triangular meshes is very challenging because very similar surfaces can be described by different data structure. The basic idea of this work is to use a description method based on a group of images produced by the projection of small pieces of surface on a set of object-oriented coordinate system that rely on a cylindrical vertex-normal reference. The collection of all this images, in coordination with the information about their reference system, describes the whole superficial morphology of a protein. To reduce the mesh noise produced by different descriptions of the same protein surface a bilinear transformation has been chosen to define the images. The projection of each vertex of the surface patch around the reference system falls in a bin of the plane to which four counters are associated. These bins are adjourned in function of the point distance inside the bin. After the projection of all the chosen vertices, each counter will assume a certain value and the ensemble of all the counters represent the image.

Results

Through this local description of the protein surface it is possible to decompose the information about its topology in a set of bi-dimensional images (Fig 1). To enforce the discrimination of the functional sites, the description images can be extracted from each protein in correspondence of the amino acids that have a key role from a function point of view. The comparison of images from different proteins can be performed using both linear correlation or other image-processing algorithms. Nevertheless, this computation is be very time-consuming if a search all against all is used. For this reason a Grid approach will be evaluated in the frame of European Project BioinfoGRID (Bioinformatics Application for Life Science) and the Italian MIUR-FIRB project LITBIO (Laboratory for Interdisciplinary Technologies in Bioinformatics).

Contact email: ivan.merelli@itb.cnr.it

Supplementary informations
BIOINFOGRID web site is available at http://www.itb.cnr.it/bioinfogrid LITBIO web site is available at http://www.litbio.org

Fig.1 Comparison between two images of the L-Chain and of the H-Chain Thrombin.