Development of a software tool for the analysis and visualization of protein-protein interactions

Wiltgen M[1], Tilz GP[2], Demel U[2]
[1]*Institute of Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria*
[2]*Clinical Immunology and Jean Dausset Laboratory, Medical University of Graz, Austria*
marco.wiltgen@meduni-graz.at

Introduction The determination and visualization of molecular interfaces between proteins is of special importance for understanding biological processes and diseases at a molecular level. Examples are the analysis of immune complexes, complexes of growing factors and receptors and especially the tumour necrosis factor – receptor complex. Such interactions are of special importance in modern diagnosis and therapy e.g. in the case of infectious diseases, disturbed metabolic situations, incompatibilities in pharmacology etc. The complexity of the protein-protein interface makes it necessary to choose simple representations, allowing the investigator to concentrate on the specific features of interest without being confused by the wealth of information. Hence we concentrate our attention on the development of software for the analysis of macromolecular interfaces of protein complexes.

Material and methods A large number of protein complexes are deposited in the PDB structure database, an international repository for 3D structure files, which can be accessed through Internet [1]. To demonstrate our method we selected protein complexes with two chains (not connected via covalent bonds). The macromolecular interface can be defined by residues on both polypeptide chains which are close enough to form interactions. The considered contact distances, where at least one atom of the adjacent residues must fall in, ranges between 4 and 6 angstroms. First the distances between the residues of the two chains of a complex are calculated. The residues, with distances within the given range, between the two chains are identified and the interface contact matrix, a plot of adjacent residues between the two amino acid chains, is constructed. The residue name and number within each chain is plotted on the respective axis (horizontal and vertical axis) and a corresponding entry is done at the appropriate place in the matrix wherever two residues of each chain come into contact. Then the contact matrix shows the distribution of the residues involved in the macromolecular interface. The elements of the matrix are additionally annotated with several (complementary) physicochemical properties (like: hydrogen bonds, hydrophobicity etc.).

Of course for a realistic and adequate visualization of a macromolecular interface a 3D representation is necessary. Therefore the elements in the interface contact matrix are linked with the 3D structure in that way that mouse manipulations on the matrix elements display the corresponding region in the 3D structure, enabling an interactive visualization and further analysis of the residues involved in the interface. To allow a suitable visualization of the 3D macromolecular interface, the chains are represented by their secondary structure elements whereas the atoms from the residues at the interface are represented as ball-and-stick. Hydrogen bonds between atoms of the two chains are visualized as dotted lines. Additionally the identified residues in the interface contact matrix are used to define the molecular surface at the interface.

The calculations were done by a special developed computer program, where the results are used as input by the Swiss-Pdb Viewer, which is actually one of the most sophisticated visualization tools for macromolecules [2]. The program enables the identification of the residues involved in the interface between the two chains, the determination of hydrogen bonds between appropriate atoms of the interface residues, the calculation and printing of the interface contact matrix, the calculation of the molecular surfaces at the interface out of the elements in the interface contact matrix etc. Most calculations for identifying the residues at the protein-protein interface are done automatically by use of the atomic coordinates (from the PDB file). Interacting windows enables the relationship between the output data of the distance calculations and the corresponding parts of the protein structure. The proprietary program, written in a script language (a kind of PERL derivate), is read in the script window of the Swiss-Pdb Viewer and the different analysis procedures are initiated interactively by the user via mouse clicking. The program runs on a PC with Microsoft Windows. For illustration purposes, calculations were performed on several trypsin (a digestive enzyme of mammals which catalysis the hydrolysis of peptide bonds) complexes.

Results Our approach offers the advantage of connecting the interface contact matrix with a 3D visualization of the complex interfaces. This allows the analysis of the potein-protein interactions at the level of the primary structure (amino acid sequence) and on the level of the quaternary structure. The interface contact matrix enables the investigator an overview about the distribution of the involved residues (annotated with physicochemical properties), an evaluation of interfacial binding "hot spots" and an easy detection of common or similar patterns in different macromolecular interfaces. The analysis of the patterns in the interface contact matrices of slightly different protein complexes allows an easy detection of structural changes. This can be used for studying the effects of mutations. To demonstrate the method, calculations were performed on the enzyme trypsin in complex with several inhibitors. The peptide chains of the inhibitor molecules differ by some changes in amino acids at distinct places. The patterns in the contact matrices are very similar and some of them are showing a typical appearance in all the considered complexes. But in detail there are some changes and effects like for example additional or missing hydrogen bonds can be observed.

The visualization of the selected residues in a 3D view via interacting windows allows a realistic analysis of the macromolecular interface. This connection of the two representations and the visualization reveal from a wealth of information the context and connections without overwhelming the investigator. The representation with molecular surfaces shows complementary shapes. The resulting molecular surface (with projected physiochemical properties of the involved residues) can thereby aid exploration of molecular complementarities at the macromolecular interface.

Discussion A growing section of bioinformatics deals with the computation and visualization of protein 3D structures. Detecting similar protein surfaces provides an important route for discovering unrecognized or novel functional relationships between proteins. This is of special importance for the planning of an individual drug treatment. Better medication can be developed once the structures of binding sites are known. There exist several approaches and methods for studying macromolecular interfaces like: the web resource iPfam, allowing the investigation of protein interactions in the PDB structures at the level of (Pfam) domains and amino acid residues and MolSurfer, which establish a relation between a 2D Map (for navigation) and 3D the molecular surface [3-4].

Our approach offers the advantage of connecting the interface contact matrix with a 3D visualization of the complex interfaces. In the interface contact matrix the involved residues of the macromolecular interface can be determined, (complementary) physicochemical properties be annotated and common pattern of different interfaces detected. The visualization of the selected residues in a 3D view via interacting windows allows a realistic analysis of the macromolecular interface.

References

[1] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE.The Protein Data Bank. Nucleic Acids Research 2000 28:235-242.
[2] Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modelling. Electrophoresis 1997 18: 2714-2723.
[3] Gabdoulline RR, Wade RC, Walther D. MolSurfer: A macromolecular interface navigator. Nucleid Acids Res. 2003 31 (13):3349-51.
[4] Finn RD, Marshall M, Bateman A. iPfam: visualization of protein-protein interactions in PDB at domain and amino acid resolutions. Bioinformatics 20021(3): 410-412.