Construction of Biomolecular Network in Structurome and Analysis of Cooperativity

Mitsunobu Sekitou  
f673048m@bio.kyutech.ac.jp

Akinori Sarai  
sarai@bio.kyutech.ac.jp

Dept. Bioscience and Bioinformatics, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, 820-8502 Japan

Keywords: biomolecular network, structurome, database, cooperativity

1 Introduction

Most of biological functions are achieved by cooperative action of huge number of molecules in the network such as signal transduction and metabolic networks. In order to understand the molecular mechanism of the cooperative action in the molecular network and to develop drugs, structural information of molecules and their interactions are essential. There are more than 39,000 known structures of biomolecules registered in Protein Data Bank (PDB) [1]. The size of biomolecular structural space (structurome) is expanding very rapidly, which is further accelerated by structural genomics projects. Furthermore, the structurome is becoming more complex, as it is populated by many molecular complexes in recent years. Therefore, it is also becoming more difficult to make a systematic analysis of the structurome and to extract useful information. Thus, we have developed a database/tool of biomolecular network, PDBnet, based on the structural information of PDB. PDBnet enables users to extract various kinds of information on the network of molecular clusters and to visualize the network. By using this tool, we have analyzed the cooperativity in molecular interactions.

2 Method and Result

2.1 Clustering

We have decomposed the PDB structures into chains (e.g., peptide and nucleotide chains). The original number of chains we used is 90,806. There is a redundancy in PDB such as the structures of the same protein and mutant proteins. Thus, we first removed the redundancy by clustering chains based on sequence similarity. For this purpose, we used the cluster list based on BLAST (with an identity threshold of 90%) obtained from PDB. The number of the resulting clusters is 13,621. We also made clustering of nucleic acids based on the similarity of their sequences, and obtained 569 clusters.

2.2 Cluster-Cluster Link

We have examined all the pairs of chains whether they share the same PDB code. If two chains share the same PDB code we created a link between the clusters containing these chains. The total number of the links created is 9,808. These links indicate that there are some physical contacts and therefore some functional relations between the clusters.

2.3 Network

By putting the linked clusters together, we obtained 804 separate networks. These networks represent segments of molecular networks in structurome, and should corresponds to some biological functions.

2.4 Database

Various kinds of information on molecules, clusters and networks were implemented into a relational database using MySQL. These data were also integrated with sequence, structure, property and functional
information through a backbone database, 3DinSight, developed in our laboratory. Also, the information on the domain of proteins was added to the chain information.

2.5 Search and Visualization Interfaces
All the data in PDBnet can be searched through a search interface, and the interaction of clusters and the network are visualized by using NEATO [2]. The network viewer enables users to overlook the molecular networks and visualize each molecular structure and browse through cluster information by clicking on each node on the screen. PDBnet can be accessed through the Internet [3].

![Example of network](image)

White nodes are protein clusters and orange nodes are DNA clusters or molecules. This figure shows a network associated with the signal transduction in T cell.

00944: Nuclear factor of activated T-CELLS
02015: C-JUN, Transcription factor AP-1
04953: C-FOS, P55-C-FOS proto-oncogene protein
08014: Interferon regulatory factor 3
12025: Cyclic-AMP-dependent transcription factor ATF

2.6 Analysis of Cooperativity
We are analyzing the cooperativity of molecular interactions at a structural level by using PDBnet. Currently, we are analyzing the cooperativity of protein-DNA interaction in a systematic way. We compare the complex structures of monomer and heterodimer of transcription factors bound to DNA (see Fig. 1 for example), and examine the effect of cooperative recognition of DNA sequences by transcription factors. We have observed that the cooperative interactions of proteins to DNA enhance the specificity of protein-DNA recognition.

Discussion

It is inconvenient to analyze the interaction of molecules systematically by using the present PDB. PDBnet we constructed provides a bird’s-eye view of molecular organization in structurome. This kind of tool enables us to analyze the structural mechanism of cooperative action of molecules in a systematic manner. At present, the domain structure of proteins is not considered for the construction of the network. Moreover, PDB contains artificial chimera proteins. The domain and chimera information is described for each cluster in the database. We are currently constructing the network by considering the information of domain and chimera. We are also combining the networks of PDBnet with the functional biological networks.

References

