3DinSight: An Integrated Database and Search Tool for Structure, Function and Property of Biomolecules

Jianghong An [1] ajh@rtc.riken.go.jp
Takao Nakama [2] nakama@rtc.riken.go.jp
Yasushi Kubota [2] kubota@rtc.riken.go.jp
Akinori Sarai [1] sarai@rtc.riken.go.jp

[1] Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki, 305 Japan
[2] Advanced Technology Institute Inc., 3-23-15 Jinbo, Kanda, Tokyo, 101 Japan

1 Introduction

In order to understand many biological phenomena, it is critical to get insight into the relationship among structure, function and property of biomolecules. However, it is usually difficult to infer the relation from individual data, since we usually need to examine several databases and literatures to obtain the necessary information. It would be useful to have an integrated database where one can examine the relationship among structure, function and property. There are some services available in the Internet to link various databases, but the relational information of biomolecular structure, function and property is rather scarce.

3DinSight is an integrated database, search and visualization tool for structure, function and property of biomolecules, developed to help researchers to get insight into their relationship [1]. Various kinds of searches can be carried out though WWW interfaces. The locations of motif sequences and mutations are automatically mapped on the structure, and visualized by interactive viewers, VRML (Virtual Reality Modeling Language) and RasMol, where the mapped 3D objects are hyper-linked to the corresponding document data. Also, the thermodynamic data of proteins and mutants are integrated into 3DinSight. The amino-acid properties of a molecule, together with structural and functional information, can be displayed as a graph plot. 3DinSight is freely accessible through the Internet (http://www.rtc.riken.go.jp/3DinSight.html).

2 Relational database

We have implemented structural, functional and property data of biomolecules into SYBASE relational database by creating various tables. The structural data are taken from PDB and daily updated. Those data are also classified into several subsets according to sequence homology and functions.

As to the functional data, we have implemented PROSITE, which contains motifs of proteins. The functional sites are mapped on amino-acid sequence and structure, and stored in a relational table. We have also implemented the Protein Mutant Database (PMD), which contains a collection of information about protein mutants. The mutant positions are mapped on the amino-acid sequences of PDB by using the PDB-PIR cross-reference table. We created a table to relate the protein chains with locations of mutations to make one-to-one correspondence to the location on structure.

We are developing Thermodynamic Database for Proteins and Mutants (ProTherm), which is a collection of various thermodynamic data of proteins and their mutants. Some of these data have direct link to structural and functional information. Various properties of amino acids based on their physico-chemical properties such as hydrophobicity and propensities for secondary structures are included as a table, by which one can examine the relationship among amino-acid property, structural and functional information.

The present relational database also contains links to amino-acid databases, SWISS-PROT and PIR, where link information is updated according to the sequence homology. We are also developing Protein-Nucleic Acid Recognition Database, which include structural database for protein-nucleic acid complex, base-amino acid interactions and prediction tools for transcription factor targets.

3 WWW interface

3DinSight can be accessed through the Internet by a WWW interface to the database. The address of the main page is: http://www.rtc.riken.go.jp/3DinSight.html. One can combine various searches such as entry, keywords, author, structure resolution and sequence pattern in the secondary structure, for the specified dataset and display option linking functional information. 3DinSight also contains various interfaces for searching PROSITE, Protein Mutant Database, PIR, SWISS-PROT, ProTherm and Protein-Nucleic Acid Recognition Database. There are other convenient interfaces such as finding motifs and sequence homologues in PDB for user's sequence. 3DinSight can also be searched by SQL directly.

The screened structure and associated functional or mutation sites can be visualized by using visualization tools, VRML and RasMol [2]. In either case, structures can be displayed automatically by selecting from the menu in the entry display, and the functional or mutation sites in the 3D structure are clickable objects linked to the PROSITE or mutant documents, so that one can obtain the corresponding functional or mutational information such as description of the site and associated literature. One can also do amino-acid analysis, where properties of the molecule can be examined; mathematical operations such as sum and average for each protein chain can be performed and displayed; the property profiles can be plotted and displayed as a graph, where the actual locations of secondary structures, functional and mutation sites are automatically mapped.

References

[1] An, J., Nakama, T., Kubota, Y., and Sarai, A., 3DinSight: An Integrated Relational Database and Search Tool for Structure, Function and Property of Biomolecules, *Bioinformatics*, 14:188–195, 1998.

[2] Nakama, T., An, J., Kubota, Y., and Sarai, A. Visualization of Functional Sites on Protein Structures by Virtual Reality Modeling Language, *Bioimages*, 5:59–64, 1997.