selected in order; a 3D-structure of the molecule is shown. For this 3D-visualization, J-mol is used as a plug-in program for web browsers. The 3DMET is now available as Release 2.1. The Release 2.2 will be opened to access until this presentation. The recent development and changes of 3DMET will be also reported.

2PT009 配列順序を無視したタンパク質構造アライメントの性質

Some characteristic features of non-sequential structural alignment: difference between sequential and non-sequential alignments

Tatsuo Mukai, Kengo Sawada, Shintaro Minami, George Chikenji (*Grad. Sch. of Eng. Univ. of Nagoya*)

Recently, a number of interesting examples that show non-sequential structure similarity have been reported, where non-sequential structural similarity is the structural similarity in which structurally equivalent regions are aligned in different order in the sequence of the compared proteins. For example, Minami et al. demonstrated that the usage of a non-sequential structure alignment program yielded many structure alignments with significant similarities between proteins of known three-dimensional structure and newly determined protein structures that possess a new fold. This finding suggest that many new folds can be modeled by permutating old folds.

In order to develop a new prediction method for new fold targets utilizing the non-sequential structure relationship, we first must clarify characteristic features of non-sequential structural relationship. As a first step towards the goal, we performed a large scale non-sequential alignment between protein pairs deposited in PDB and compare their features with those of sequential alignments. We found several characteristic features found only in non-sequential alignment. For example, protein size and solvent accessibility were largely not conserved in non-sequentially similar pairs. These finding may be usefull for developing a new De Novo prediction method. In the presentation, we will report detailed description of the method and the results.

2PT010 Prediction of the types of carbohydrate-binding proteins

Zhenyi Ge, Wei Cao, Kazuya Sumikoshi, Shugo Nakamura, Kentaro Shimizu (*Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo*)

Carbohydrate-binding proteins or lectins, which are defined as carbohydrate-binding complexes which are not directly modified by carbohydrate, are important complexes that participate in many biological processes as intercellular information transfer and cell recognition. Though there are a variety of methods to classify lectins, classifying lectins based on structural and evolutionary sequence similarities is more popular because of some highly conserved amino acid sequence motifs in CRD (Carbohydrate Recognize Domain). Depending on this way, lectins could be classified into C-type lectins (require calcium for recognition), S-type lectins (also known as Galectins, which require free thiols for stability), L-type lectins (have a feature structure called "jelly roll fold") and so on. We developed a tool to detect and classify lectins by HMM (Hidden Markov Models). Several corresponding profile HMMs are composed and are prepared for scoring entry sequences. Sequences in low score section are regarded as NOT lectins and sequences in the corresponding range of scores are classified into fitting types of lectins.

2PT011 Analysis of DNA-binding sites using novel amino acid index for molecular recognition (MR-index)

Katsutoshi Hibino¹, Yasunori Yokoyama², Shigeki Mitaku² (¹*Dept. of Computational Science and Engineering, Grad. Sch. of Engineering, Nagoya Univ.*, ²*Dept. of Applied Physics, Grad. Sch. of Engineering, Nagoya Univ.*)

Protein-DNA interaction plays an important role in many biological processes. However, the mechanism of protein-DNA interaction is still not clear. Prediction of DNA-binding sites in DNA-binding proteins using their physical properties is very useful for understanding mechanism of protein-DNA interaction. There are many experimentally determined structures in Protein Data Bank database, but only a few protein-DNA complexes are available. In addition, it is very expensive and time-consuming process to solve the structure of a protein-DNA complex through experimental methods. Therefore, a sequence-based computational prediction of DNA-binding sites in DNA-binding proteins is required. In this work, we used a novel amino acid index for molecular recognition (MR-index) to develop a system for predicting DNA-binding sites in DNA-binding proteins. Although MR-index was originally

developed for predicting allergen-unique fragments (AUFs), it is found that most proteins have MR-index peaks at a constant rate. Also, MR-index peaks locate at positions with large structural fluctuations such as loop regions, and those sites tend to interact with other molecules. We took account of physical properties around MR-index peaks and aimed to develop the system for predicting DNA-binding sites in DNA-binding proteins. The discrimination analysis showed that the MR-index peaks around DNA-binding sites have characteristic distribution of physical properties.

2PT101 アミノ酸保存度と相互作用エネルギーを組み合わせたタンパク質-リガンド結合部位予測手法の開発

Protein-Ligand Binding Site Prediction by Using Energy-based Method with Amino Acid Conservation Scores

Hiroto Tsujikawa, Yusuke Umeda, Mizuki Morita, Shugo Nakamura, Kentaro Shimizu (*Grad. Sch. Agric. Life Sci., Univ. Tokyo*)

Protein-ligand interaction is one of the most important steps for expression of protein functions. Identifying ligand binding sites helps protein function analysis and structure-based drug design. Recently, a variety of computational methods for predicting ligand binding sites are proposed. Among these methods, the energy-based method which calculates protein-ligand interaction energy on a three-dimensional protein structure is shown to realize high precision. In this study, we developed a new method which combines the energy-based method with amino acid sequence conservation scores. Our method locates methyl probes around the protein surface and calculates van der Waals interaction energy between each probe and the protein. In addition, we calculate amino acid conservation scores by constructing position-specific scoring matrix (PSSM) and add their weighted scores to the interaction energy. It is considered that highly conserved amino acid residues are expected to be closely related to protein function and appear in functional sites. Probes with energetically favorable for binding sites and within a short distance are clustered and they are ranked according to their total interaction energies. As a result, in 80% proteins, at least one site in top three predicted sites predicts ligand binding sites correctly. We will evaluate the performance of our method and compare it with other existing methods, and show the importance of amino acid sequence conservation in ligand binding sites prediction.

2PT102 蛋白質・蛋白質相互作用ネットワークにおけるドメイン及び非ドメイン間相互作用の解析

Analysis of Domain and Non-Domain Interactions in Protein-Protein Interaction Network

Komori Satoshi, Sarai Akinori, Fujii Satoshi (*Kyusyu Institute of Technology*)

The analysis of protein-protein interactions and their network has been usually based on protein chains and domains. However, non-domain regions are often involved in the interactions and serve function as well. Here, we have considered non-domains as well as domains in the protein-protein interactions. The information of domains within the individual chains was obtained from Pfam, and sequence-based clustering was carried out. Non-domains were defined as the regions outside the domains, within certain length threshold. We have examined physical contacts among all pairs of domains and non-domains within each complex structure of PDB, and identified 15,742 domain-domain interactions, 15,218 domain-non-domain interactions and 12,804 non-domain-non-domain interactions. Then, we constructed the interaction network. These pieces of information on domains and non-domains and their interactions were implemented into PDBnet, which enables users to examine protein interaction network in terms of domain and non-domain interactions. We have compared characteristics of interactions due to domains and non-domains, such as amino acid composition, interaction type and packing. We will also make systematic analysis on the relationship among various properties at molecular structure, interaction and network levels, in order to get insight into the function of proteins. Furthermore, we will apply the present analysis to the systematic predictions of protein-protein interactions based on the catalog of interacting domains and non-domains.

2PT103 進化的に保存されたタンパク質間相互作用の複合体立体構造の予測 3D complex structure assignment for evolutionary conserved protein-protein interactions

Takeshi Kawabata, Haruki Nakamura, Akira Kinjo (*Institute of Protein Research, Osaka University*)