1P057**Properties of Residue Contact Occurrence in Denatured Proteins Calculated with the Potential from the Interresidue Average Distances**

○Takeshi Kikuchi

Dept. Biosci. and Bioinf., Col. Inf. Sci. and Eng., Ritsumeikan Univ.

We have analyzed contact maps derived from interresidue average distance statistics for several proteins and compared with their folding units. The results have suggested that local appearances of ADMs reflect positions of folding initiation sites of proteins, i.e., ADMs reflect folding process of proteins. In particular, ADMs for proteins which clearly show two-state folding predicts positions of sub-domains containing residues with high ϕ values. This fact indicates that the interresidue average distance statistics effectively contain the information on protein folding processes. In this study, we try to make simulations of structure sampling of proteins in denatured states with interresidue potential derived from the interresidue average distance statistics from a totally randomized conformation and calculate residue contact frequencies. The contact frequency $g(i, j)$ is defined as a frequency of a pair of residues i and j makes a contact, i.e., their $C\alpha$ distance $R_{ij} < R_{cut}$. In this study, we take $10 \text{ \AA}$ as R_{cut} . Furthermore, we calculate $Q(i, j)$ which denotes the deviation from the average contact frequency $g(m)$ where $m=i-j$ ($i > j$). Thus, $Q(i, j)$ implies a pair of residue i and j with high frequency of contacts. Furthermore, we calculate $p(i) = \sum_j Q(i, j)$. We investigate the properties of $Q(i, j)$ and $p(i)$ with folding properties of proteins exhibiting two-state folding. We also discuss with the predictability of protein folding processes with our analyses.

1P059**Predicting ligand binding sites of uncharacterized protein structure**

○Mizuki Morita, Shugo Nakamura, Kentaro Shimizu

Graduate School of Agricultural and Life Sciences, The University of Tokyo

We have developed a computational algorithm for the accurate identification of ligand (small-molecule) binding sites on three-dimensional protein structures. A variety of computational methods for identifying and characterizing binding sites have been proposed over the past two decades. Because the ultimate goal of the binding site prediction methods is to find active sites on uncharacterized structures, it is important to test and validate the algorithm on data sets of the unbound (apo) structures. Most of these methods, however, only tested ligand-bound (holo) structures instead of performing a realistic test using unbound (apo) structures. Our method is simple, accurate and fast enough to predict protein-ligand binding site of uncharacterized protein structures. In this method, called StickyFingers, the protein surface is coated with a layer of probes to calculate interaction energies with the protein. Probes with favorable interaction energies are retained and clusters of these probes are ranked according to the number of probes in a cluster or their total interaction energies. The largest cluster or energetically most favorable cluster is then ranked first. The success rate in the first predicted site was 89% for the ligand-bound state and 83% for the unbound state. The percentages of proteins with at least one success in the first two sites were 100% for the ligand-bound state and 94% for the unbound state. The average speed of overall process is 0.5 minutes on a Dual 2.5 GHz Apple PowerPC G5.

1P058**Development and evaluation of a system for scoring the local structural segment configuration**

○Yoshihide Makino, Nobuya Itoh

Biotechnology Research Center, Toyama Prefectural University

The data deposited in the protein data bank (PDB) provides increasingly valuable information for the structural study of the protein. Here we report the development of a simple structural quality check system using the coordinate data extracted from the PDB. The coordinates of specific amino acid pairs that are not adjacent to the primary structure were collected. Then a score similar to a dipole interaction function was calculated for each pair using their C-alpha and pseudo C-beta atom coordinates. The effectiveness of this scoring was tested with barnase as the test structure. A partial structural segment, such as the N-terminal helix, was selected, and the relative translation position of the partial segment was scanned. The distinctiveness of the native position was evaluated according to whether the maximum score could be obtained at the native location of the segment relative to the remaining protein structure. Simple distance-based scoring was confirmed effective when the rough (such as three or five angstroms) scanning step was employed. This dipole-like scoring system also worked well to pick up the native position. When scanning using a smaller step (one angstrom or so) was examined, native positions were picked up for most cases with dipole-like scoring, whereas less accurate positions were predicted with distance-based scoring. The simple dipole-like scoring system would be applicable for a rapid check of structural quality in a structure determination or prediction study.

1P060**Development of a protein tertiary structure prediction server**

○Shugo Nakamura, Mizuki Morita, Masanori Kakuta, Kentaro Shimizu

Dept. of Biotechnology, The Univ. of Tokyo

We developed an automated protein structure prediction server named ENABLE. This server can generate model structures for a given target sequence even if the target does not have any structural templates. Overview of the prediction procedure of our server is as follows: 1) Search templates for the target sequence using PSI-BLAST and fold recognition tools such as FUGUE. 2) Secondary structure prediction using PSIPRED, disorder prediction using disABLE developed in our laboratory, and searching fragments as candidate sub-structures of each part of the target are executed. 3) Build a number of tertiary structure models according to template structures and alignments found in step 1. 4) Assess qualities of generated models in step 3 and determine where to improve and whether de novo prediction is needed for this target or not. 5) Generate models from extended structure (de novo prediction) or improve parts of model structures generated in step 3. 6) Pick up several models as prediction results according to clustering and assessment of qualities of the models. For model generation in step 4, we used IDDD/ABLE system based on fragment assembly method developed in our laboratory. For model quality assessment in step 5, we developed a model quality predictor based on support vector regression (SVR). We test our server system on CASP7 (Seventh Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction).