

## Development of a Scoring Method for Predicting Protein Complex Structures

Y. Tsuchiya<sup>1</sup>, E. Kanamori<sup>3</sup>, D. M. Standley<sup>2</sup>, H. Nakamura<sup>2</sup>,  
K. Kinoshita<sup>1</sup>

<sup>1</sup>*Institute of Medical Science, the University of Tokyo, Tokyo, Japan*

<sup>2</sup>*Institute for Protein Research, Osaka University, Osaka, Japan*

<sup>3</sup>*Biomedical Information Research Center, Tokyo, Japan*

The information about protein-protein interactions increases much more rapidly than the increase of the number of the tertiary structures of those protein complexes. Therefore, precise prediction of protein complex structures by protein-protein docking simulations is required. When the protein complex is rebuilt from its component protomers which derive from experimentally determined complex structure (native structure) by docking, the complex models with rmsd < 10 Å from the native structure (near-native model) could be obtained, along with a great number of false positives (decoy). The separation of near-native models from many decoys is therefore needed in the prediction of complex structures by docking. In this study, we developed the method for scoring docking models so that the near-native models were higher in rank than decoys, based on the assumption that the interfaces of near-native models are more complementary in terms of surface properties and shapes compared to those of decoys.

We used 125 non-redundant hetero-dimers (native structures) as targets. For each target, maximum 500 complex models were generated by our docking method. We also observed these targets in terms of the shape of the interfaces of their native structures. As a result, we found that these targets could be classified into two groups according to their interface shapes, and moreover, that these classes correlated with the groups which was based on the number of models with high docking score, namely, the difficulty in the separation of near-native models. We therefore only focused on 75 targets classified as difficult targets which need the separation. So far our method could separate the near-native models from the decoys in 70% of these targets.

---

Proceedings of The Joint 2<sup>nd</sup> Pacific Rim International  
Conference on Protein Science and 4<sup>th</sup> Asian-Oceania  
Human Proteome Organization, Cairns- Australia, 22-26  
June 2008