MEGADOCK: a High-speed Protein-protein Interaction Prediction System by All-to-all Physical Docking

Masahito Ohue[1,2,†] Yuri Matsuzaki[1] Nobuyuki Uchikoga[3]
Takashi Ishida[1] Yutaka Akiyama[1,‡]

1. Graduate School of Information Science and Engineering, Tokyo Institute of Technology, Tokyo, Japan.
2. Research Fellow of the Japan Society for the Promotion of Science.
3. Department of Physics, Chuo University, Tokyo, Japan.

[†]ohue@bi.cs.titech.ac.jp [‡]akiyama@cs.titech.ac.jp

The elucidation of protein-protein interaction (PPI) networks is important for understanding cellular systems and structure-based drug designs. However, the development of an effective method to conduct exhaustive PPI screening represents a computational challenge.

We have been investigating a protein-docking approach based on shape complementarity and physico-chemical properties. To realize the procedures required to sample a huge number of protein dockings, we have developed "MEGADOCK", a high-speed protein-protein docking software package. MEGADOCK reduces the calculation time required for docking by using several techniques such as a novel scoring function called the real-Pairwise Shape Complementarity (rPSC) score. We demonstrate that MEGADOCK is capable of exhaustive PPI screening by completing docking calculations 8.8 times faster than the conventional docking software, ZDOCK, while maintaining an acceptable level of accuracy. When our PPI prediction system was applied to a subset of a general benchmark dataset to predict 120 relevant interacting pairs from 120x120=14,400 combinations of proteins, an F-measure value of 0.231 was obtained.

MEGADOCK showed comparable docking accuracy to other FFT-based software programs, such as ZDOCK, while employing a much simpler and thus computationally less expensive score function. Additionally, our software was shown to be applicable to a large scale protein-protein interaction screening problem with accuracy better than random. With our approach combined with parallel high-performance computing systems, searching and analyzing protein-protein interactions with consideration to three-dimensional structures at the interactome scale is now a feasible problem.