March 2016

Master's Thesis Abstract

Protein ligand-binding pocket comparison by a reduced vector representation derived from multidimensional scaling of generalized description of binding pockets

(多次元尺度構成法を用いた蛋白質ポケット部位の
縮約ベクトル表現によるポケット構造比較)

Department of Computational Biology and Medical Sciences
Graduate School of Frontier Sciences
The University of Tokyo

中村　司

47-146908　**Tsukasa Nakamura**

Adviser: Visiting Associate Professor Kentaro Tomii

Abstract

Proteins serve various functions in living cells. When they exert their functions, physical contact with other molecules occurs. A close connection therefore exists between their functions and structures. Therefore, comparison and classification about known and predicted protein structures provides important insight into the structural features of proteins, elucidating their functions and structures. Analyzing the mutual interactions between proteins and small molecules is important to predict the ligands which bind to parts of putative ligand binding pockets. Such analysis demands a fast and efficient method for comparing ligand binding pockets because of the recent increase of protein structure information. A method has been developed for representing a ligand binding pocket with one reduced vector for binding pocket comparison. Using our new method, one can calculate the similarity between ligand binding pockets merely by calculating the inner product of about 140-dimensional vectors. The new method exhibits higher detection performance of similar binding pockets than metrics used in existing alignment-free methods and another existing accurate alignment-dependent method: APoc. We also investigated the effects of modifications in expansion and revision of edge classes with two datasets for improving the ability to detect similar binding pockets. Results show that the effects of the modifications depend on the difficulty of the dataset, implying some avenues for methods of improvement.

Keywords

ligand-binding site, alignment-free comparison, multidimensional scaling, vector representation, large-scale comparison