Poster I-72

Protein Active Site Comparison and Retrieval System Based on 3D Pattern Matching

Authors:
Dae-Hee Kim (*Electronics and Telecommunications Research Institute*)
Sung-Hee Park (*Electronics and Telecommunications Research Institute*)
Chan-Yong Park (*Electronics and Telecommunications Research Institute*)
Soo-Jun Park (*Electronics and Telecommunications Research Institute*)
Seon-Hee Park (*Electronics and Telecommunications Research Institute*)

Short Abstract: we propose a method which finds and shows some active sites structurally similar to active site of the target protein. The similarities between the active site of target protein and that of other proteins are calculated by our proposed system based on geometric hashing (GH).

Long Abstract:
Active sites of proteins are important targets in drug design. Designed drugs, however, sometimes bind not only the active site of target protein but also that of other proteins, which may cause the side effects of the drugs. Therefore, it is very important to find the other proteins and consider them in drug design in order to prevent the side effects in advance. In this poster, we propose a method to be able to find undesirable proteins structurally similar to the active site of target protein. The method finds and shows some active sites structurally similar to active site of the target protein. In order to represent protein surfaces and recognize the active site, alpha shape method and reorganizing active site information are used. The proposed system has provided the automatic identification of active sites from CASTp, and a functionality to enable users to select active sites manually.

The similarities between the active site of target protein and that of other proteins are calculated by our proposed system based on 3D geometric hashing (GH). Generally, the protein structure comparison by geometric hashing takes much time. In this poster, however, our approach can compare many active sites within a short time by restricting to the active sites.

Experiments have been conducted with 91 protein data and shows that the processing time is average 99 seconds for a single query. From the result of experiments, we confirmed that the proteins with similar active site have the same domain and the similar function, and a similar pocket was found in two apparently non-homologous proteins (1sri-1epa).