RAP: Refine a Prediction of Protein Secondary Structure

Ying-Chuan Liu,Chun-Yuan Lin,游坤明,Chuan-Yi Tang
Computer Science & Information Engineering
Computer Science and Informatics
yu@chu.edu.tw

Abstract

RAP aims to refine protein secondary structure prediction from one of famous prediction tools. Protein secondary structure prediction has been extensively discussed for almost 50 years and the machine learning is one of feasible methods for it with more than 70% accuracy. PSIPRED, PHD and PROF are well-known machine learning approaches and based on the three-state prediction: helix, strand, and coil. RAP is a post-processing tool with logical and statistic methods to refine the scores of three-state prediction. Hence, RAP can be easily applied to any protein secondary structure prediction tool if it uses three-state prediction. RAP was tested on the CASP data set with 181 targets and a large-scale data set with 69534 chains of 31402 proteins in PDB database. In the experiment, PSIPRED is used to give the scores of three-state prediction for each test data at first and then RAP predicts the result by refining the scores from PSIPRED. RAP obtains overall prediction accuracy rates 81.3% and 80.7 % for the CASP and PDB data sets, respectively.

Keyword：