

METHODS FOR PROTEIN STRUCTURE PREDICTION

Zhiquan He

Dr. Dong Xu, Dissertation Supervisor

ABSTRACT

The lack of tertiary structures is a main obstacle to fully understanding the functions of these proteins. Computational structure prediction from amino acid sequence is a viable solution for this.

The widely used method for protein structure prediction is comparative protein modeling, which heavily relies on fold recognition performance and alignment accuracy. Another step in protein structure prediction is the structural assessment for predicted protein structures, which obviously plays a critical role.

We implemented a protein sequence alignment tool which is based on Conditional Random Field and improved its alignment quality by incorporating more complex scoring models. After deeper study of fold recognition and alignment problem, we proposed a new protocol to improve the quality of sequence profiles, which intrinsically affects the performance of fold recognition and alignment accuracy.

Besides this, several machine learning methods, such as Hidden Markov Model and Conditional Random Field, have been proposed to combine knowledge scoring functions and consensus methods from different perspectives for structural quality assessment purpose.