Learning from Evolution to Predict Protein Structure

Burkhard Rost
European Molecular Biology Laboratory
Heidelberg, Germany
email: rost@embl-heidelberg.de

Abstract. In the wake of the genome data flow, we need - more urgently than ever - accurate tools to predict protein structure. The problem of predicting protein structure from sequence remains fundamentally unsolved despite more than three decades of intensive research effort. However, the wealth of evolutionary information deposited in current databases enabled a significant improvement for methods predicting protein structure in 1D: secondary structure, transmembrane helices, and solvent accessibility. In particular, the combination of evolutionary information with neural networks proved extremely successful. The new generation of prediction methods proved to be accurate and reliable enough to be useful in genome analysis, and in experimental structure determination. Moreover, the new generation of theoretical methods is increasingly influencing experiments in molecular biology.

[A more extended version is given as a postscript file.]