Comparative Modeling for Protein Structure Prediction

Predicting the 3D structure into which a protein will fold from just its sequence of amino acids remains one of the most central and challenging unsolved problems in Computational Biology. When two proteins are sufficiently close evolutionary relatives, they have both similar sequences and similar folds. As they become more evolutionarily distant, the sequence often diverges faster than the shape. Thus discovering a meaningful weak sequence signal may help predict the folds of proteins that are evolutionary relatives of proteins whose 3D structures are known but whose sequence similarity to these proteins has diverged into "the twilight zone".

We present our work developing novel "wrapping" methods for two specific known structural templates: the beta-helix and beta-trefoil folds. We discuss plans and remaining challenges for making the method high-throughput to solve all beta-structural folds, including the challenge of automatically abstracting structural templates from a set of known 3D structures that share a common core fold. Our recent new multiple structure alignment program that incorporates flexibility into protein structural alignments in order to better align more divergent structures can help construct these templates.

This is joint work with Bonnie Berger, Andrew McDonnell, Jonathan King and Matt Menke.