Evolving Local Minima in the Protein Energy Surface

Brian Olson

Department of Computer Science
George Mason University
Fairfax, VA, 22030
bolson3@gmu.edu

Abstract

Proteins are the molecular tools of the living cell and the path to unraveling their function is through modeling and understanding their structure. Many diseases occur when a protein loses its intended function due to its inability to form the appropriate structure with which it binds to other molecules in the cell. A holistic approach to protein modeling would be to characterize all possible structural states accessible by a protein under native, physiologic conditions. However, this task is infeasible. The question then becomes, how can we model the subset of these structural states most relevant to the function or disfunction of a protein?

This thesis proposes a novel computational framework to obtain an expansive view of the protein conformational space relevant for function while controlling computational cost. The framework complements experimental and high-resolution computational methods which limit their focus to a single region of the conformational space. The framework employs the knowledge that functionally-relevant conformations are those low in energy and the framework incorporates the latest understanding of protein structure and energy from biophysics. Specifically, this thesis proposes a novel stochastic search framework for exploring a diverse ensemble of conformations which capture low-energy basins in the protein energy surface.

The proposed search framework employs a hybrid or memetic approach for explicit sampling of local minima in the protein energy surface. This hybrid search framework combines a global evolutionary search approach with a local search component to take advantage of the latest advances from the computational biology community. Specifically, the following questions are addressed to effectively model the protein conformational space: (1) How to balance limited computational resources between exploration of the conformational space in global search with exploitation of local minima in local search? The hybrid search framework combines a global evolutionary search to explore the breadth of the conformational space with a local search for efficiently exploiting local minima in the underlying energy surface. (2) How to sample new conformations at the global level? Two complementary approaches are investigated. One approach proposes an enhanced fragment selection method for sampling a new conformation based on an existing structure. The other approach employs a genetic algorithm to combine features from multiple existing structures to sample a new conformation. (3) How to employ energy to better discriminate between interesting conformations and noise in the conformational search space? A multi-objective decomposition of the energy function is employed to guide the search towards more biologically relevant, low-energy conformations by focusing on the energy terms with the most discriminatory power.

Work in this thesis shows that, by combining advanced algorithmic components with the latest understanding of protein biophysics, the proposed search framework is able to more effectively model functionally-relevant conformational states. A direct comparison between the proposed framework and a state-of-the-art coarse-grained sampling algorithm shows that the enhanced sampling strategies lead to a more comprehensive picture of the underlying protein energy surface. By taking this more comprehensive view, the framework is able to capture the protein native state as well as or better than methods relying primarily on protein-specific sampling strategies.