Protein Structure Prediction in the 2D HP Model Using Particle Swarm Optimization

A. Băutu, H. Luchian

Modern research in bioinformatics deals with large amounts of data or the simulation of complex biological processes. Many of these problems are too demanding in terms of computational and/or memory requirements for classical algorithms. Nature inspired metaheuristics can be used for finding fast and reasonably accurate solutions for them. Proteins are the most important of all the molecules found in living cells. To carry out its tasks, a protein must fold into a complex three-dimensional structure called native state, which represents the energetic ground state of the protein. Various simplified models for the protein structure exist (e.g. the Toy model, the Functional Model Protein — FMP, the Hydrophobic-Polar model — HP). The high computational complexity of predicting the folded structure of a protein recommends this problem for metaheuristics approaches.

This paper applies the Particle Swarm Optimization (PSO) algorithm to search the ground state of protein foldings. We propose a novel approach, that uses a discrete PSO variant, specially designed for protein folding in the HP model. Despite the simplicity of the model, the protein folding problem in the HP model is $\mathcal{NP}$-hard in both 2D and 3D. Extensive experiments are performed, that cover both artificial data and real protein data, both with relative and absolute 2D folding coordinates. The results indicate that the proposed PSO method is very effective to search for ground states of the proteins structures, with respect to solution accuracy and speed.