Monte Carlo Sampling of Protein Conformations

Peter Freddolino, John Stone, Klaus Schulten
Theoretical and Computational Biophysics, Beckman, UIUC

Introduction

Protein structural and thermodynamic properties are often probed through the use of Monte Carlo simulations. In such simulations, random perturbations are applied to the structure according to a predefined set of rules, and the resulting moves are accepted or rejected according to a probability based on the energy of the states before and after the moves. Long Monte Carlo simulations allow significant sampling of conformational space, and can be used in protein structure prediction or to aid in building structures for other types of simulation.

Description

MCS is a small program allowing Monte Carlo sampling of the structure of a protein. A series of Monte Carlo steps are attempted; at each step, some portion of the protein backbone is perturbed, and the orientation of one or more side chains of the protein are changed. The energy of the newly generated conformation is calculated, and the step is then accepted or rejected according to standard Monte Carlo criteria. If the step is not accepted, a new conformation is attempted instead. At the end of the simulation, both a final conformation and statistics about the run are output. Each Monte Carlo step may be made more or less "severe" based on how many degrees of freedom are changed during one step; allowing more changes simultaneously will allow larger overall structural changes in fewer steps, but reduce the acceptance rate of each individual step.

Objective

The objective of the project is to implement the Monte Carlo sampling portion of MCS on a GPU using CUDA. Complete implementation of a finely tuned Monte Carlo sampling algorithm for protein structure should be quite possible in the time allotted for the project. The sampling implementation will be benchmarked using a set of target proteins ranging in size from 10 to 8000 residues. Analysis of the limitations of the hardware for this problem is required, including characterization of differing limitations that may occur with different protein sizes and step severities.

Background

No specialized knowledge is necessary to work on the application, since the algorithm to be implemented is straightforward and a reference implementation is provided. Background in basic biology will make the target application easier to understand, but can easily be made up for by some brief reading on protein structure. Programming expertise is assumed.

Resources

Algorithm references, papers, and sample code will be made available online:
http://www.ks.uiuc.edu/Research/vmd/projects/ece498/

Contact Information

Peter Freddolino: petefred@ks.uiuc.edu
John Stone: johns@ks.uiuc.edu