John von Neumann Institute for Computing

Multiplexed-Replica Exchange Molecular Dynamics with the UNRES Force-Field as an Effective Method for Exploring the Conformational Energy Landscape of Proteins.

Cezary Czaplewski, Sebastian Kalinowski,
Stanislaw Ołdziej, Adam Liwo, Harold A. Scheraga

published in

NIC Workshop 2006,
From Computational Biophysics to Systems Biology,
Jan Meinke, Olav Zimmermann,
Sandipan Mohanty, Ulrich H.E. Hansmann (Editors)
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **34**, ISBN-10: 3-9810843-0-6,
ISBN-13: 978-3-9810843-0-6, pp. 63-66, 2006.

http://www.fz-juelich.de/nic-series/volume34

Multiplexed-Replica Exchange Molecular Dynamics with the UNRES Force-Field as an Effective Method for Exploring the Conformational Energy Landscape of Proteins.

Cezary Czaplewski[1,2], Sebastian Kalinowski[2], Stanisław Ołdziej[1,2], Adam Liwo[1,2], and Harold A. Scheraga[1]

[1] Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853-1301, U.S.A., *E-mail: has5@cornell.edu*

[2] Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland

The theoretical prediction of the three-dimensional structures of proteins purely from an amino-acid sequence is a grand challenge of computational biophysics. Here, we present simulations using an effective search algorithm combined with the united residue force-field. In the replica exchange method, the kinetic trapping problem is overcome by exchanging noninteracting replicas simulated at different temperatures. To enhance sampling, the replicas are multiplexed with a number of independent molecular dynamics runs at each temperature. The simulated free-energy landscape is presented for 1CLB protein. Application of the multiplexed-replica exchange algorithm to large-scale parallel computing is discussed.

1 Introduction

Efficient conformational sampling algorithms are an essential component of methods for studying protein structure and dynamics. Methods such as canonical Molecular Dynamics (MD) or Metropolis Monte Carlo (MC) can be used for estimating thermodynamic properties as well as for a global search, but in practice they easily become trapped, and thus are not effective methods for studying rough free-energy landscapes of proteins.

One of the most effective sampling methods, the Replica Exchange method (also known as Exchange MC[1], or Parallel Tempering[2]), was initially developed to improve sampling in glassy systems in statistical physics. However, following Hansmann's use of the method in Monte Carlo simulations of a simple peptide, Met-enkephalin[2] and Okamoto and Sugita's formulation of an MD version of the algorithm,[3] the Replica Exchange method has been applied extensively in protein-folding simulations.

2 Methods

In the united-residue physics-based UNRES model,[4–6] a polypeptide chain is represented by a sequence of α-carbon (C^α) atoms linked by virtual bonds with attached united side chains (SC) and united peptide groups (p) positioned in the middle between two consecutive C^α atoms. The UNRES force field has been derived as a Restricted Free Energy (RFE) function decomposed into factors arising from interactions within and between a

given number of united interaction sites[5]. The selected parameters of the UNRES force-field were refined by hierarchical optimization[6] of the potential-energy function.

The Replica Exchange MD (REMD) method combines the idea of simulated annealing and Monte Carlo methods, and is one of the generalized-ensemble algorithms that performs a random walk in energy space due to a free random walk in temperature space. In the REMD method, n replica systems, each in the canonical ensemble, and each at a different temperature, are simulated. At given intervals, swaps, or exchanges, of the configurational variables between systems are accepted with the Metropolis criteria. In this paper, we are using multiplex REMD (MREMD) introduced by Pande[7]. In MREMD, to enhance sampling, the replicas are multiplexed with a number of independent molecular dynamics runs at each temperature. Exchanges of configurations only between random replicas with neighboring temperatures are tried.

3 Results and Discussion

The free energy is the most important quantity for the description of equilibrium properties of proteins. An example of the simulated free-energy landscape for the 1CLB protein is shown in Figure 1.

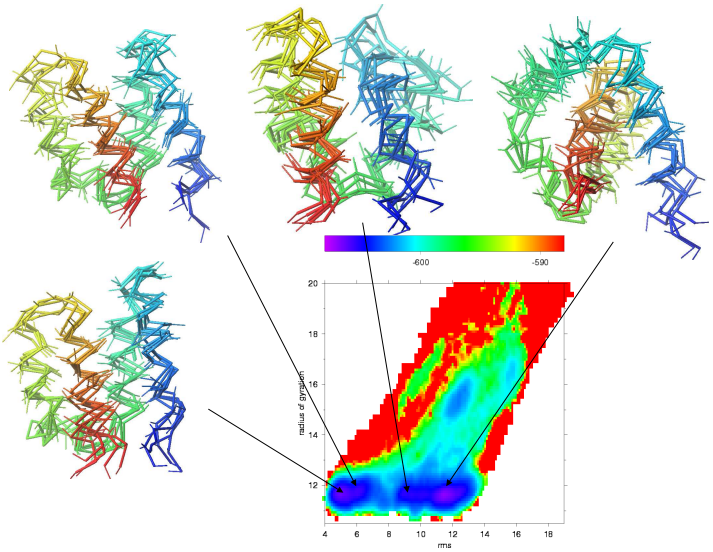

Figure 1. Free energy landscape of 1CLB protein as a function of rmsd and radius of gyration, and representative structures from low free-energy regions.

To calculate free energy profiles, we used the densities of states obtained from the multi-histogram analysis of MREMD trajectories.[8] From the densities of states, we calculated the microcanonical entropy, $S(E_i) = k_B ln\Big[n(E_i)\Big]$, for all conformations collected from the simulations, and used it to compute the microcanonical free energies with the

following expression: $F(E_i, T) = E_i - TS(E_i)$. To plot the free energy as a function of rmsd (r) and radius of gyration (ρ), we calculated the free energy by evaluating the following expression for each grid point:

$$F(r, \rho, T) = -k_B T ln \sum_{E_i \in N(r,\rho)} exp\left(\frac{-F(E_i, T)}{k_B T}\right) \quad (1)$$

where the index i enumerates conformations within the histogram bins, $N(r, \rho)$, for given ranges of rmsd and radius of gyration.

Parallelization of the MREMD method has been enhanced by removing the synchronization step. The exchanges of configurations between random replicas with neighboring temperatures are tried, not at the same number of MD steps for each replica but are forced by the replica with the lowest temperature independent of the number of steps performed by the other replicas. The improved algorithm scales almost linearly up to 1,000 processors with over 80% average performance.

Acknowledgments

Support from the National Institutes of Health (grant GM-14312), the National Science Foundation (grant MCB00-03722), the NIH Fogarty International Center (grant TW7193), the Polish Ministry of Education and Science (grant 3 T09A 032 26) as well as use of supercomputing resources at the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputer Center is gratefully acknowledged.

References

1. K. Hukushima and K. Nemoto. Exchange monte carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.*, 65:1604–1608, 1996.
2. U. H. E. Hansmann. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.*, 281:140–150, 1997.
3. Y. Sugita and Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, 314:141–151, 1999.
4. A. Liwo, S. Ołdziej, M. R. Pincus, R. J. Wawak, S. Rackovsky, and H. A. Scheraga. A united-residue force field for off-lattice protein-structure simulations. I. functional forms and parameters of long-range side-chain interaction potentials from protein crystal data. *J. Comput. Chem.*, 18:849–873, 1997.
5. A. Liwo, C. Czaplewski, J. Pillardy, and H. A. Scheraga. Cumulant-based expressions for the multibody terms for the correlation between local and electrostatic interactions in the united-residue force field. *J. Chem. Phys.*, 115:2323–2347, 2001.
6. S. Ołdziej, J. Łągiewka, A. Liwo, C. Czaplewski, M. Chinchio, M. Nanias, and H. A. Scheraga. Optimization of the unres force field by hierarchical design of the potential-energy landscape. 3. use of many proteins in optimization. *J. Phys. Chem. B*, 108:16950–16959, 2004.
7. Y. M. Rhee and V. S. Pande. Multiplexed-replica exchange molecular dynamics method for protein folding simulation. *Biophys. J.*, 84:775–786, 2003.

8. M. Nanias, C. Czaplewski, and H. A. Scheraga. Replica exchange and multicanonical algorithms with the coarse-grained united-residue (UNRES) force field. *J. Chem. Theor. Comput.*, 2:513–528, 2006.