John von Neumann Institute for Computing

Optimization of the United-Residue UNRES Force Field for Langevin Dynamics Simulations

Adam Liwo, Mey Khalili, Cezary Czaplewski,
Sebastian Kalinowski, Stanisław Ołdziej,
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. 107-110 , 2006.

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

Optimization of the United-Residue UNRES Force Field for Langevin Dynamics Simulations

Adam Liwo[1,2], Mey Khalili[1], Cezary Czaplewski[1,2], Sebastian Kalinowski[2], Stanisław Ołdziej[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

Recently, with success, we applied the united-residue UNRES force field developed in our laboratory to carry out molecular dynamics simulations. This communication is a preliminary report of the development of a procedure for optimizing the UNRES force field for canonical simulations.

1 Introduction

In the last decade we have been developing a united-residue physics-based force field termed UNRES[1–3] for energy-based prediction of protein structures from amino-acid sequences. Recently[4], we extended the application of UNRES to mesoscopic molecular dynamics simulations and found that, with this approach, we can simulate protein folding pathways in real time. This new application requires reparameterization of the force field, which is the subject of our current work. We describe preliminary results in this communication.

2 Methods

In the UNRES model, a polypeptide chain is represented as a sequence of α-carbon atoms (C^α) with attached united side chains (SC) and united peptide groups (p), each of which is positioned in the middle between two consecutive C^α atoms. The effective energy function is a sum of different terms corresponding to interactions between the SC, SC and p, and p sites, as well as local and correlation terms, each of which is multiplied by an appropriate weight[3], w. The expressions for these terms had been derived[2] based on a Kubo cluster cumulant[5] expansion of a polypeptide chain in water, where the degrees of freedom not present in the model had been integrated out. In the present work we introduced explicit dependence of the cumulant-based terms on temperature.

The hierarchical method of force field optimization developed in our laboratory[3] aims at obtaining energy landscapes of selected training proteins such that the free energy of each of the training proteins decreases with increasing native likeness. In the present study, we computed the free energies below, at, and above the folding-transition temperatures and extended the approach by the requirements that the free-energy relations be inverted above the folding-transition temperature. We used the multiplexing replica-exchange molecular dynamics (MREMD)[6] to generate decoy sets with given energy-function parameters

and the weighted-histogram analysis method (WHAM)[7] to process the results in order to compute the free energies of the ensembles.

3 Results and Discussion

The structures of the three training proteins used in this work and the representatives of the most probable conformations obtained with optimized force fields (a separate optimization was carried out for each protein) are shown in Figure 1.

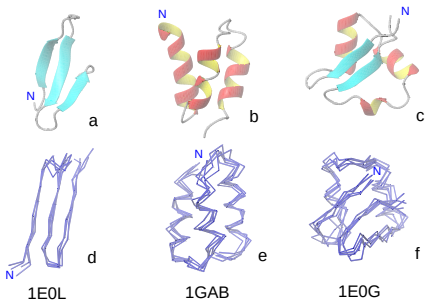

Figure 1. Experimental structures of 1E0L (a), 1GAB (b), and 1E0G (c) and some of the most probable conformations at temperatures below the folding-transition temperature of these proteins (d–f). The experimental structures are shown in the ribbon representation and the simulated structure are shown as C^{α}-traces. The N-termini are marked for tracing purposes.

The results of test *ab initio* prediction runs (using the multiplexing replica-exchange molecular dynamics and the force fields parameterized using 1GAB on a number of α-helical proteins are summarized in Table 1. It can be seen that the force field predicts the structures of proteins both with simple three- or four-helix bundle folds (1BDD, 1CLB, 1LQ7, 1E68, 1P68) and those with more complex topology (1POU, 1KOY, 1PRU), although its performance on the former is better. Therefore the force field parameterized using 1GAB appears transferable. We are currently working on including proteins with more complex α-helix topology as well as more complex $\alpha + \beta$-proteins in parameterization.

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, the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdańsk, and the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) at the University of Warsaw is gratefully acknowledged.

PDB ID	No of residues	lowest RMSD	index(es) of native cluster(s)	proba-bility
1BDD	46	2.2	1–4	0.84
1LQ7	67	1.6	5	0.10
1E68	70	3.9	4	0.08
1CLB	75	3.9	3	0.15
1P68	102	2.7	2	0.23
1POU	71	5.3	7	0.05
1PRU	56	5.3	–[a]	–[a]
1KOY	62	4.4	9	0.003

Table 1. Results of tests on α-helical proteisn of the force field parameterized on 1GAB.

[a] No clear cluster of native-like structures was located.

References

1. 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.
2. 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.
3. S. Ołdziej, A. Liwo, C. Czaplewski, J. Pillardy, and H. A. Scheraga. Optimization of the unres force field by hierarchical design of the potential-energy landscape: 2. Off-lattice tests of the method with single proteins. *J. Phys. Chem. B*, 108:16934–16949, 2004.
4. A. Liwo, M. Khalili, and H. A. Scheraga. Ab initio simulations of protein-folding pathways by molecular dynamics with the united-residue model of polypeptide chains. *Proc. Natl. Acad. Sci. U.S.A.*, 102:2362–2367, 2005.
5. R. Kubo. Generalized cumulant expansion method. *J. Phys. Soc. Japan*, 17:1100–1120, 1962.
6. U. H. E. Hansmann and Y. Okamoto. Comparative study of multicanonical and simulated annealing algorithms in the protein folding problem. *Physica A*, 212:415–437, 1994.
7. S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.*, 13:1011–1021, 1992.