Computational techniques for efficient conformational sampling of proteins

Adam Liwo[1], **Cezary Czaplewski**[1], **Stanisław Ołdziej**[1,2], and **Harold A Scheraga**[1]

[1]*Baker Laboratory of Chemistry, Cornell University, Ithaca, NY 14853-1301, United States* [2]*Intercollegiate Faculty of Biotechnology, University of Gda sk, Medical University of Gdan sk, ul. Kładki 24, 80-822 Gda sk, Poland*

Abstract

In this review, we summarize the computational methods for sampling the conformational space of biomacromolecules. We discuss the methods applicable to find only lowest energy conformations (global minimization of the potential-energy function) and to generate canonical ensembles (canonical Monte Carlo method and canonical molecular dynamics method and their extensions). Special attention is devoted to the use of coarse-grained models that enable simulations to be enhanced by several orders of magnitude.

Introduction

Searching the conformational space of biological macro-molecules (in particular proteins), also termed *conformational sampling*, is of fundamental interest because of the tight relationship between accessible conformations and function. The primary interest here is its importance for obtaining stable states, that is, structure, kinetic, and thermodynamic properties, and folding pathways of biological macromolecules, for example, proteins. Early approaches to compute protein structure involved global optimization of potential energy [1,2¨], but these were only partially successful in special cases in which entropic effects were less important than enthalpy. At room temperature, the entropic effect can favor a basin that does not contain the global minimum of the potential energy, as illustrated in Figure 1.

The thermodynamically correct approach to conformational sampling should seek the basin with the lowest free energy for the whole system [3¨,4], that is, the protein plus the surrounding medium, usually water. Therefore, this article will first treat global optimization of a target function, with most emphasis on the potential-energy function, and then will consider canonical sampling and generalized-ensemble sampling. The focus of this review is on techniques that facilitate simulations of structure and folding of real-size proteins and we, therefore, emphasize the use of coarse-grained models; for treatment of systems at atomic detail and use of experimental information in sampling, the reader is referred to recently published reviews [5˙,6¨].

Global optimization of the potential-energy function and other target functions

In the protein-folding field, a variety of physics-based procedures for global optimization of potential energy have been proposed. Details of many of them are described in references [1, 2¨]. These procedures fall into four categories:

Deterministic methods—The deterministic methods include (i) a build-up procedure that finds local minima of short fragments by exhaustive energy minimization, and then combines these short fragments into longer ones, with minimization of the energies of the longer fragments [1,2¨]; (ii) a self-consistent electrostatic field (SCEF) method that alters conformation to achieve optimal alignment of each peptide-group dipole with respect to the local electric field at that peptide group [1,2¨]; (iii) a self-consistent multitorsional field (SCMTF) method that is based on the fact that the ground-state solution of the Schrödinger equation provides information about the location of the global minimum of a potential function, that is, the maximum of the square of the ground-state wave function is very often close to the global minimum [1,2¨]; (iv) deformation methods in which the multi-dimensional energy surface is smoothed, either by solving the diffusion equation (DEM) or by scaling the distance variables (DSM) in the potential-energy function, to locate the global minimum [1,2¨].

Stochastic and hybrid methods—The stochastic and hybrid methods include (v) Monte Carlo with minimization (MCM) whereby combinatorial optimization with Monte Carlo is combined with energy minimization to find local minima; related approaches are basin hopping [1,2¨], and αBB, a branch-and-bound method [7¨]; (vi) electrostatically driven Monte Carlo (EDMC), which is a combination of SCEF and MCM, (vii) self-consistent basin-to-deformed basin mapping (SCBDBM) that locates large regions of conformational space containing low-energy minima by coupling them to some of the greatly reduced number of minima on a highly deformed surface.

Genetic algorithms—Conformational space annealing (CSA) [8¨] is a frequently used genetic type algorithm; it combines essential aspects of the build-up procedure and a genetic algorithm and searches the whole conformational space in early stages, and then narrows the search to smaller regions with low energy. CSA has been applied to find the lowest energy structures of proteins [1,2¨] and to protein sequence alignment [9].

A hierarchical approach—Finally, a hierarchical approach is based on carrying out an extensive coarse-grained search with the CSA method using a united-residue (UNRES) force field. The set of families of low-energy UNRES conformations obtained with CSA is then converted to an all-atom representation, and the search is continued with the EDMC method [1,2¨,10]. Other coarse-grained approaches have been presented by Levitt [11], Koliński and co-workers [12¨,13¨], Brown and Head-Gordon [14], and Izvekov and Voth [15].

Over the years, the approaches, based on global optimization of potential energy, have facilitated the development and evolution of physics-based methods to compute three-dimensional structures of native proteins and the pathways leading to them. They also serve as components of the more extensive procedures, including introduction of entropic effects, discussed in the remaining sections of this article.

Canonical sampling—The goal of this approach is to obtain a set of low-energy conformations of the system under study distributed according to the Boltzmann law, that is, $P(\mathbf{X}_i) \sim \exp(-E(\mathbf{X}_i)/kT)$ where $P(\mathbf{X}_i)$ is the probability of occurrence of the conformation described by the variables $\mathbf{X}_i$, $E(\mathbf{X}_i)$ is the energy of this conformation, k is the Boltzmann constant and T is the absolute temperature. Two approaches can be considered as prototypes of canonical sampling: the Monte Carlo (MC) method of Metropolis [16¨] and the molecular dynamics (MD) method [17¨]. Both methods provide not only canonical ensembles but also trajectories of conformational changes.

The basic MC algorithm can be summarized as follows: perturb the current conformation at random, compute the energy of the new conformation and, if the energy is lower than the energy of the old conformation, replace the old conformation with the new one; if not, accept the new

conformation with probability $\exp(-\Delta E/kT)$, where ΔE is the difference between the energy of the new and the old conformation. The process is iterated until the ensemble averages converge. The ensemble averages are updated every given number of steps. Thus, the MC approach requires only energy evaluation but its efficiency is hampered by small steps that must be taken to achieve a sufficient acceptance rate.

Several methods have been developed to speed up Monte Carlo search. Koliński and co-workers [12¨,13¨,18] dis-cretized the conformational space of polypeptide chains in a coarse-grained representation to a high-resolution lattice, and designed a set of efficient moves, many of them being collective moves of virtual bonds. Use of a lattice enabled them to pre-compute and store many energy contributions to save time. In connection with knowledge-based information, their model of polypeptide chains and search algorithm were applied successfully to protein-structure prediction [13¨] and to study folding pathways [18]. Recently, Shakhnovich and co-workers [19] designed a knowledge-based move set to carry out all-atom off-lattice simulations of protein folding with their statistical potential. In the high-directional Monte Carlo (HDMC) method [20], the sampling scheme estimates the shape of an energy hypersurface as a covariance tensor to generate variable Monte Carlo steps. The force-bias [21] and hybrid [22] Monte Carlo methods bias a MC step towards the direction of the forces or one determined by an MD step. With these extensions, the efficiency of the MC approach in terms of the convergence of ensemble averages becomes comparable to [23] or higher [24] than that of canonical MD.

In the MD approach, Newton's or Lagrange's equations of motion are solved by numerical integration, which provides much higher efficiency of a single step compared to canonical MC, but involves the expense of computing the energy gradient. Because solving equations of motion in their original form keeps the total energy and not the temperature, or temperature and pressure, constant, thermostat, or thermostat and barostat terms must be included. Details of MC and MD algorithms can be found in the excellent book by Frenkel and Smit [17¨] and in a recent review [25].

Given the potential-energy function, the efficiency of MD depends primarily on the length of the time step in the numerical-integration algorithm, which is 1 fs for all-atom representations with all covalent bonds unconstrained, and 2 fs for those with bonds constrained; this is orders of magnitude smaller than the biological time scale. Nevertheless, *ab initio* folding of the villin headpiece, a 36-residue protein, in explicit water has been studied by Duan and Kollman [26¨]. Extension of the time step can be achieved by reducing the set of variables, for example, to the torsional angles only [27,28˙], although such an approach introduces additional computational cost because of the appearance of a non-diagonal inertia tensor and non-inertial forces [6¨]. Treating the solvent at the mean-field level reduces computational time and accelerates movement along the trajectory; such an approach enables one to study *ab initio* folding of small proteins [29˙]. Use of coarse-grained models of polypep-tide chains provides a 4000-fold increase of efficiency compared to all-atom simulations with explicit solvent and, consequently, *ab initio* studies of protein folding in real time [3¨]. It must be noted, however, that the introduction of coarse-graining or even implicit solvent results in distortion of the time scale of simulations because the fast degrees of freedom are averaged out [3¨].

Non-Boltzmann and generalized-ensemble sampling—Because empirical force fields usually produce rough energy landscapes, canonical MC and MD often become trapped at low temperatures and, moreover, sample higher energy regions poorly. A variety of approaches have been developed to overcome energy barriers and to provide sufficient sampling. In the umbrella-sampling method [30˙], a set of restraints is imposed on selected variables or groups of variables [reaction coordinate(s)] to enhance sampling in selected regions of conformational space, and the weighted-histogram analysis method (WHAM) [31]

is applied to remove the contribution from the biasing potential. This approach was applied to construct energy landscapes of proteins [32'] and to generate decoy sets for optimizing protein force fields [4,33]. In the simulated-annealing (SA) technique [34], simulations start at a high temperature, which is gradually decreased until reaching the desired value. The performance of the method depends very much on the cooling schedule.

One of the most effective sampling methods is the replica-exchange method (REM) [35',36'], in which n replica systems, each being a canonical ensemble, and each at a different temperature, are simulated by using MC or MD. At given intervals, exchanges of temperatures are attempted between neighboring replicas; each exchange is accepted or rejected by the Metropolis criterion. Appropriate choice of the temperature distribution is a crucial issue [37']. A recently developed adaptive feedback-optimized algorithm [37'] minimizes the round-trip times between the lowest and highest temperatures. In complex cases, in which there are subtle bottlenecks in the probability of exchange of conformations, the round-trip time is likely to characterize the overall efficiency of parallel tempering better than that obtained with the average acceptance probability.

Instead of, or in addition to temperatures, restraints [38,39'] or Hamiltonians [39',40'] are exchanged between replicas. The multiplexing variant of REM (MREM) [41',42'] enhances sampling by multiplexing the replicas with a number of independent MC or MD runs at each temperature.

Exchanging temperatures between replicas greatly enhances the chance of a trapped conformation to overcome kinetic barriers and, therefore, accelerates the convergence of ensemble averages. However, trajectory information is lost in the process. The REM and MREM methods have been applied extensively in studying protein free-energy landscapes and finding low-energy conformations of proteins both in the all-atom [39',43,44] and coarse-grained approaches [4].

Multicanonical algorithms are another class of methods, in which energy is replaced in the Metropolis criterion by the logarithm of the density of states (the microcanonical entropy) [45"]. A simulation is converged when all energies are sampled with the same frequency; therefore, this method is well suited to overcome energy barriers. Once the density of states is obtained, all ensemble averages can be computed. The multicanonical methods have been applied to peptides and proteins both at the all-atom [46',47'] and coarse-grained [42'] level, both in connection with MC and MD and to study protein folding with lattice models [48]. Two methods that combine replica exchange and a multicanonical algorithm, one by using replica exchange to generate the initial density of states (the REMUCA algorithm) and, the second one, by performing a replica-exchange simulation with a small number of replicas, each covering a different region of the energy space (the MUCAREM algorithm) [47'], respectively, have been introduced recently. We recently [42'] implemented the MUCA, REMUCA, and MUCAREM algorithm to our coarse-grained UNRES force field and found that they do not offer significant advantage over the MREM method, while involving the computation of the gradient of the density of states that has to be evaluated numerically and is unstable particularly in low-energy regions where the density of states is low.

A technique termed locally enhanced sampling (LES) [49,50] was designed to tackle problems, such as ligand binding [49], ligand diffusion or loop optimization in proteins [50], where the conformation of only a small part of the system must be explored thoroughly, while the conformation of the rest of the system stays almost intact. In this method, multiple copies of the smaller part are generated and create a mean field acting on the bulk of the system.

Conclusions and outlook

Techniques to sample conformational space, both to find the low-energy conformations and to compute canonical ensembles and ensemble averages underwent enormous progress since the seminal work of Metropolis and co-workers [16¨]. The genetic algorithms such as the CSA method [8¨] are currently definitely the most efficient in finding the global minimum of the potential-energy function, while algorithms based on the replica-exchange approach [35˙,36˙] are the most effective to obtain canonical ensembles. All these algorithms are ideal targets for parallelization and have benefited enormously from the advent of distributed computing. Additional acceleration of computations is achieved by introducing the hierarchical (multiscale) approach, in which the bulk of the computations is carried out at the coarse-grained level. Such an approach can soon provide the possibility of reaching the millisecond time scale [3¨]. The crucial factor here is the development of physics-based coarsegrained potentials that could reproduce the structure and thermodynamics of macromolecules [4].

Acknowledgements

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

· of special interest

·· of outstanding interest

1. Scheraga HA, Lee J, Pillardy J, Ye Y-J, Liwo A, Ripoll D. Surmounting the multiple-minima problem in protein folding. J Glob Optimiz 1999;15:235–260.

2. Wales DJ, Scheraga HA. Global optimization of clusters, crystals, and biomolecules. Science 1999;285:1366–1372.of outstanding interestA comprehensive review of global-optimization methods applied to find the lowest energy structures of clusters, molecular crystals, as well as polypeptides, proteins, and other biological molecules.

3. Liwo A, Khalili M, Scheraga HA. Molecular dynamics with the united-residue (UNRES) model of polypeptide chains; test of the approach on model proteins. Proc Natl Acad Sci USA 2005;102:2362–2367. [PubMed: 15677316]of outstanding interestThe paper describes the first mesoscopic ab initio MD of globular proteins using a coarse-grained potential not engineered for a particular sequence. It demonstrates that globular proteins with a size of more than 70 amino-acid residues can be folded from the unfolded conformation in a few hours on a single processor.

4. Liwo A, Khalili M, Czaplewski C, Kalinowski S, Ołdziej S, Wachucik K, Scheraga HA. Modification and optimization of the united-residue (UNRES) potential energy function for canonical simulations. I. Temperature dependence of the effective energy function and tests of the optimization method with single training proteins. J Phys Chem B 2007;111:260–285. [PubMed: 17201450]

5. Lei H, Duan Y. Improved sampling methods for molecular simulation. Curr Opin Struct Biol 2007;17:187–191. [PubMed: 17382533]of special interestA recent comprehensive review of conformational search algorithms.

6. Christen M, van Gunsteren WF. On searching in, sampling of, and dynamically moving through conformational space of biomolecular systems: a review. J Comput Chem 2008;29:157–166. [PubMed: 17570138]of outstanding interestThe paper presents a thorough guide of algorithms for searching conformational space; it classifies the algorithms according to the objectives (obtaining just low-energyconformations or ensembles) and algorithm type.

7. Androulakis IP, Maranas CD, Floudas CA. αBB: a global optimization method for general constrained nonconvex problems. J Glob Optimiz 1995;11:337–363.of special interestThe paper describes the branch-and-bound global optimization method in which the problem of global optimization of a nonconvex function is tackled by replacing nonconvex terms with customized tight convex lower bound functions, successive subdivision of the original region, and the subsequent solution of a series of nonlinear convex minimization problems.

8. Lee J, Scheraga HA, Rackovsky S. New optimization method for conformational energy calculations on polypeptides: conformational space annealing. J Comput Chem 1997;18:1222–1232.of outstanding interestThe paper describes the conformational space annealing method, which combines elements of the build-up method, Monte Carlo with minimization, and a genetic algorithm. The search is started from a sparse set of randomly generated and then energy-minimized conformations, which is subsequently narrowed down to low-energy regions. The method outperforms previous approaches such as the Monte Carlo with minimization and electrostatically driven Monte Carlo methods.

9. Joo K, Lee J, Lee S, Seo J-H, Lee SJ, Lee J. High Accuracy Template Based Modeling by Global Optimization. Proteins 2007;69(Suppl 8):83–89. [PubMed: 17894332]

10. Ołdziej S, Czaplewski C, Liwo A, Chinchio M, Nanias M, Vila JA, Khalili M, Arnautova YA, Jagielska A, Makowski M, et al. Physics-based protein-structure prediction using a hierarchical protocol based on the UNRES force field–test with CASP5 and CASP6 targets. Proc Natl Acad Sci USA 2005;102:7547–7552. [PubMed: 15894609]

11. Levitt M. A simplified representation of protein conformations for rapid simulation of protein folding. J Mol Biol 1976;104:59–107. [PubMed: 957439]

12. Koliński A, Godzik A, Skolnick J. A general method for the prediction of the three-dimensional structure and folding pathway of globular proteins: application to designed helical proteins. J Chem Phys 1993;98:7420–7433.of outstanding interestThe paper describes the first successful coarse-grained approach for protein-folding simulations without prior knowledge of the experimental structure. The conformations are discretized to a high-resolution lattice, and a Monte Carlo search is carried out using specially designed virtual-bond moves. The potential-energy function is a sum of components derived from protein crystal structures by using the Boltzmann principle.

13. Skolnick J, Koliński A, Ortiz AR. MONSSTER: a method for folding globular proteins with a small number of distance constraints. J Mol Biol 1997;265:217–241. [PubMed: 9020984]of outstanding interestThe paper describes a protein-structure prediction protocol with the use of a conformational search and potentials described in reference [12··] and restraints from knowledge-based information.

14. Brown S, Head-Gordon T. Intermediates and the folding of proteins L and G. Protein Sci 2004;13:958–970. [PubMed: 15044729]

15. Izvekov S, Voth GA. A multiscale coarse-graining method for biomolecular systems. J Phys Chem B 2005;109:2469–2473. [PubMed: 16851243]

16. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. J Chem Phys 1953;21:1087–1092.of outstanding interestThis is the seminal work that has set the foundations of molecularsimulation methods. It introduces for the first time the Markov-chain scheme to compute ensemble averages in dense systems. The scheme is illustrated with the example of a two-dimensional rigid-sphere system obeying periodic boundary conditions.

17. Frenkel, D.; Smit, B. Understanding Molecular Simulation. From Algorithms to Applications. San Diego: Academic Press; 2002. of outstanding interestAn excellent textbook of molecular dynamics and Monte Carlo algorithms applied in molecular simulations.

18. Kmiecik S, Koliński A. Folding pathway of the B1 domain of protein G explored by a multiscale modeling. Biophys J 2008;94:726–736. [PubMed: 17890394]

19. Chen WW, Yang JS, Shakhnovich EI. A knowledge-based move set for protein folding. Proteins: Struct Funct Bioinform 2007;66:682–688.

20. Shin JK, Yu JY, Jhon MS. Acceleration of convergence-rates by high directional Monte-Carlo sampling procedure. J Chem Phys 1992;97:9283–9290.

21. Kidera A. Smart Monte Carlo simulation of a globular protein. Int J Quantum Chem 1999;75:207–214.

22. Izaquirre JA, Hampton SS. Shadow hybrid Monte Carlo: an efficient propagator in phase space of macromolecules. J Comp Phys 2004;200:581–604.

23. Yamashita H, Endo S, Wako H, Kidera A. Sampling efficiency of molecular dynamics and Monte Carlo method in protein simulation. Chem Phys Lett 2001;342:382–386.

24. Ulmschneider JP, Ulmschneider MB, Di Nola A. Monte Carlo vs molecular dynamics for all-atom polypeptide folding simulations. J Phys Chem B 2007;110:16733–16742. [PubMed: 16913813]

25. Scheraga HA, Khalili M, Liwo A. Protein-folding dynamics: overview of molecular simulation techniques. Annu Rev Phys Chem 2007;58:57–83. [PubMed: 17034338]

26. Duan V, Kollman PA. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. Science 1998;282:740–744. [PubMed: 9784131]of outstanding interestThe paper describes the first large-scale MD simulation of a small protein in explicit water starting from an arbitrary structure, which resulted in the appearance of some native-like structures.

27. Gibson KD, Scheraga HA. Variable step molecular-dynamics: an exploratory technique for peptides with fixed geometry. J Comput Chem 1990;11:468–497.

28. Chen JH, Im W, Brooks CL. Application of torsion angle molecular dynamics for efficient sampling of protein conformations. J Comput Chem 2005;26:1565–1578. [PubMed: 16145655]of special interestThe paper introduces dynamics in torsional angles, which reduces the number of degrees of freedom, and also corrections to the parent Cartesian-coordinate force field necessary to recover the potential surface and incorporate implicit bond and angle flexibility. With this approach, an increase of several fold in conformational sampling efficiency can be achieved reliably compared to Cartesian molecular dynamics.

29. Jang S, Kim E, Shin S, Pak Y. Ab initio folding of helix bundle proteins using molecular dynamics simulations. J Am Chem Soc 2003;125:14841–14846. [PubMed: 14640661]of special interestThe paper describes large-scale ab initio folding simulations of two small proteins: the villin headpiece (36 amino-acid residues) and the N-terminal fragment of the B-domain of staphylococcal protein A (46 residues) using AMBER with the GB implicit-solvent model. For both proteins, stable native structures were obtained, starting from arbitrary conformations.

30. Torrie GM, Valleau JP. Non-physical sampling distribution in Monte Carlo free-energy estimation–umbrella sampling. J Comput Phys 1977;23:187–199.of special interestThe paper introduces the umbrella-sampling method by using weights that bias Monte Carlo simulations to sample all conformational space regions of interest with nearly uniform probability.

31. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J Comput Chem 1992;13:1011–1021.

32. Boczko EM, Brooks CL III. First-principles calculation of the folding free energy of a 3-helix bundle protein. Science 1995;269:393–396. [PubMed: 7618103]of special interestThe paper describes a study of the free-energy landscape of the Nterminal fragment of the B-domain of staphylococcal protein A by umbrella-sampling MD simulations with explicit solvent.

33. Koretke KK, Luthey-Schulten Z, Wolynes PG. Self-consistently optimized energy functions for protein structure prediction by molecular dynamics. Proc Natl Acad Sci USA 1998;95:2932–2937. [PubMed: 9501193]

34. Chou KC, Carlacci L. Simulated annealing approach to the study of protein structures. Protein Eng 1991;4:661–667. [PubMed: 1946323]

35. Hansmann UHE. Parallel tempering algorithm for conformational studies of biological molecules. Chem Phys Lett 1997;281:140–150.of special interestThe paper describes the application of the replica-exchange (parallel tempering) method to sample the conformational space of peptides and proteins by the Monte Carlo method, with the pentapeptide [Met5]enkephalin as an example. The replica-exchange method has performed much better than the canonical Monte Carlo method, because it enables the system to overcome barriers in the rough energy landscape.

36. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 1999;314:141–151.of special interestThe paper describes the replica-exchange molecular dynamics algorithm and a test of its effectiveness on the pentapeptide [Met5]enkephalin.

37. Trebst S, Troyer M, Hansmann UHE. Optimized parallel tempering simulations of proteins. J Chem Phys 2006;124:174903. [PubMed: 16689600]of special interestThe paper describes a scheme for choosing the optimal distribution of temperatures in replica-exchange simulations based on maximizing the rate of round-trips between the lowest and the highest temperature replicas.

38. Gront D, Koliński A, Hansmann UHE. Protein structure prediction by tempering spatial constraints. J Comput Aided Mol Des 2005;19:603–608. [PubMed: 16267688]

39. Sugita Y, Kitao A, Okamoto Y. Multidimensional replicaexchange method for free-energy calculations. J Chem Phys 2000;113:6042–6051.of special interestThe paper describes the extension of the umbrella-sampling version of the replica-exchange method, in which not only temperatures but also restraints are exchanged between replicas.

40. Liu P, Huang XH, Zhou RH, Berne BJ. Hydrophobic aided replica exchange: an efficient algorithm for protein folding in explicit solvent. J Phys Chem B 2006;110:19018–19022. [PubMed: 16986898] of special interestThe paper describes a modification of the replica-exchange method for MD simulations with explicit solvent to incorporate replicas with attenuated protein–water interactions, which mimics the interaction with chaperone.

41. Rhee YM, Pande VS. Multiplexed-replica exchange molecular dynamics method for protein folding simulation. Biophys J 2003;84:775–786. [PubMed: 12547762]of special interestThe paper reports the extension of the replica-exchange method by running a number of replicas instead of one at each temperature, which enhances the performance of the algorithm on massively parallel computers.

42. Nanias M, Czaplewski C, Scheraga HA. Replica exchange and multicanonical algorithms with the coarse-grained unitedresidue (UNRES) force field. J Chem Theor Comput 2006;2:513–528.of special interestThe paper describes an extensive study of replica-exchange and multicanonical MC and MD algorithms applied to the coarse-grained UNRES force field using three proteins: decaalanine, the N-terminal part of the Bdomain of staphylococcal protein A (46 residues), and 1E0G (48 residues) as examples. The replica-exchange method appears to be most stable for 1E0G that possesses the roughest energy landscape of the three proteins.

43. Paschek D, Garcia AE. Reversible temperature and pressure denaturation of a protein fragment: A replica exchange molecular dynamics simulation study. Phys Rev Lett 2004;93:238105. [PubMed: 15601210]

44. Yang JS, Chen WW, Skolnick J, Shakhnovich EI. All-atom ab initio folding of a diverse set of proteins. Structure 2007;15:53–63. [PubMed: 17223532]

45. Berg BA, Neuhaus T. Multicanonical ensemble–a new approach to simulate 1st order phase-transitions. Phys Rev Lett 1992;68:9–12. [PubMed: 10045099]of outstanding interestThe paper describes the application of a multicanonical algorithm, in which the acceptance criterion is related to the density of states instead of the Boltzmann factor, to a 2D 10-state Potts model on a lattice. The algorithm has an approximately 100-fold faster convergence rate compared to an isothermal run.

46. Mitsutake A, Sugita Y, Okamoto Y. Generalized-ensemble algorithms for molecular simulations of biopolymers. Biopolymers 2001;60:96–123. [PubMed: 11455545]of special interestThe paper describes the application of the multicanonical algorithm as well as replica-exchange sampling and simulated tempering to peptide systems.

47. Mitsutake A, Sugita Y, Okamoto Y. Replica-exchange multicanonical and multicanonical replica-exchange Monte Carlo simulations of peptides. I. Formulation and benchmark test. J Chem Phys 2003;118:6664–6675.of special interestThe paper introduces two new algorithms for sampling the conformational space of peptides and proteins, which combine multicanonical sampling with the replica-exchange method.

48. Hao M-H, Scheraga HA. Theory of two-state cooperative folding of proteins. Acc Chem Res 1998;31:433–440.

49. Czerminski R, Elber R. Computational studies of ligand diffusion in globins. 1. Leghemoglobin. Proteins: Struct Funct Genet 1991;10:70–80. [PubMed: 2062829]

50. Hornak V, Simmerling C. Generation of accurate protein loop conformations through low-barrier molecular dynamics. Proteins: Struct Funct Genet 2003;51:577–590. [PubMed: 12784217]
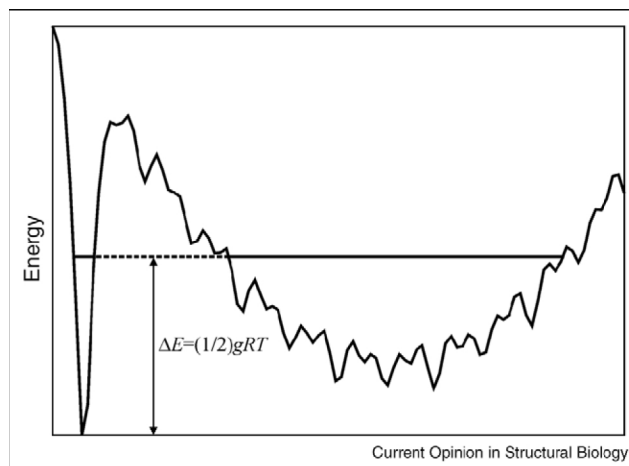
Figure 1.
Illustration of the importance of the entropic effect for the probability of occurrence of a particular family of structures. The energy surface shown in the picture contains a narrow minimum (left), which is the global minimum, and a broad basin of minima (right). The horizontal line is the approximate boundary of the accessibility of the regions of conformational space belonging to these two basins, DE being the thermal energy at absolute temperature T, g the number of degrees of freedom, and R the universal gas constant. The area closed by the energy curve and the corresponding horizontal line is related to the entropy of each basin. Consequently, at sufficiently high temperature, the free energy of conformations that belong to the right basin will be significantly lower than that of the global-minimum region and these conformations will, therefore, be more probable than those corresponding to the global minimum of the potential energy.