PROTEINS: Structure, Function, and Bioinformatics 66:697–706 (2007)

Enhanced Sampling of Peptide and Protein Conformations Using Replica Exchange Simulations With a Peptide Backbone Biasing-Potential

Srinivasaraghavan Kannan and Martin Zacharias*
School of Engineering and Science, International University Bremen, D-28759 Bremen, Germany

ABSTRACT **During replica exchange molecular dynamics (RexMD) simulations, several replicas of a system are simulated at different temperatures in parallel allowing for exchange between replicas at frequent intervals. This technique allows significantly improved sampling of conformational space and is increasingly being used for structure prediction of peptides and proteins. A drawback of the standard temperature RexMD is the rapid increase of the replica number with increasing system size to cover a desired temperature range. In an effort to limit the number of replicas, a new Hamiltonian-RexMD method has been developed that is specifically designed to enhance the sampling of peptide and protein conformations by applying various levels of a backbone biasing potential for each replica run. The biasing potential lowers the barrier for backbone dihedral transitions and promotes enhanced peptide backbone transitions along the replica coordinate. The application on several peptide cases including in all cases explicit solvent indicates significantly improved conformational sampling when compared with standard MD simulations. This was achieved with a very modest number of 5–7 replicas for each simulation system making it ideally suited for peptide and protein folding simulations as well as refinement of protein model structures in the presence of explicit solvent. Proteins 2007;66:697–706.** © 2006 Wiley-Liss, Inc.

Key words: conformational sampling; molecular dynamics simulation; protein folding; peptide folding; protein structure prediction

INTRODUCTION

The application of classical molecular dynamics (MD) simulations for structure prediction of peptides and proteins is limited by the accuracy of current force fields and the simulation time scale. Peptides and proteins can adopt numerous locally stable conformations separated by large energy barriers. Conformational transitions between stable states can therefore be rare events even on the time scale of tens to hundreds of nanoseconds that have become possible for peptide simulations.[1–12] Various methods have been proposed to overcome the conformational sampling problem during molecular simulations (reviewed in Ref. 12, 13). For example, simulated annealing techniques are frequently used to effectively cross energy barriers at high simulation temperatures followed by slow cooling of the simulation system to select low energy states.[14] However, high initial temperatures used in simulated annealing approaches may interfere with the presence of explicit water molecules during MD simulations. Alternatively, potential scaling methods have been suggested where the original potential is scaled down or replaced by a soft core potential to lower barriers during energy minimization or a MD simulation.[15–22] In the locally enhanced sampling method, multiple conformational copies of a selected region of a molecule are generated and a mean field from the copies is used during the simulation to overcome barriers.[23] The parallel tempering or replica exchange molecular dynamics (RexMD) method is one of the most successful and now most widely used methods to enhance conformational sampling in Monte Carlo (MC)[24–26] and MD simulations.[25,27–34] In RexMD simulations, several copies (replicas) of the system are simulated independently and simultaneously using classical MD or MC methods at different simulation temperatures (or force fields: Hamiltonians). At preset intervals, pairs of replicas (neighboring pairs) are exchanged with a specified transition probability. In its original implementation, temperature is used as a condition to be varied and exchanged among the replicas. The random walk in temperature allows conformations trapped in locally stable states (at a low simulation temperature) to escape by exchanging with replicas at higher simulation temperature. The RexMD method has been successfully applied in folding simulations of several peptides and mini-proteins.[30–34] Unfortunately, efficient exchange between replicas requires sufficient overlap of

Grant sponsor: Pacific Northwest National Laboratories; Grant number: gc11-2002; Grant sponsor: VolkswagenStiftung.

*Correspondence to: Martin Zacharias, School of Engineering and Science, International University Bremen, Campus Ring 1, D-28759 Bremen, Germany. E-mail: m.zacharias@iu-bremen.de

Received 20 May 2006; Revised 6 September 2006; Accepted 15 September 2006

Published online 21 November 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21258

the energy distributions between neighboring replicas. As a consequence, the number of required replicas grows approximately with the square root of the number of particles in the system (to cover a desired temperature range).[35] A larger number of replicas in turn requires also increased simulation times in order to allow efficient "travelling" of replicas in temperature space. One common approach to avoid an excessive increase in the number of replicas in case of studying larger peptides or proteins is to eliminate the solvent degrees of freedom by using an implicit solvent description (e.g. Generalized Born (GB) model).[36] However, it is not clear whether the accuracy of current implicit solvent models is sufficient for a realistic description of the structure and dynamics of peptides and proteins.[31–33] Hybrid explicit/implicit solvent models have been suggested were the simulation of each replica is performed using an explicit solvent description and for each exchange part of the solvent is replaced by a continuum.[37] Another approach employs separate coupling of solute and solvent to different heat baths (target temperatures).[38] Only the solute reference temperatures are varied for each replica. Both methods reduce the effective system size compared at each attempted replica exchange. However, the artificial temperature gradient at the solute–solvent interface may cause artifacts in the latter methods. Instead of using the simulation temperature as a replica coordinate, it is also possible to use the force field or Hamiltonian of the system as a replica-coordinate.[35,39–42] Recently, a promising "Hamiltonian"-RexMD method has been suggested where the solute–solute, solute–solvent, and solvent–solvent interactions are separately (linearly) scaled for each replica.[41] This approach can be used to "effectively" scale only the solute temperature along the replica coordinate. In case of no scaling of the solvent–solvent interactions, the replica exchange probability becomes less dependent on the number of solvent degrees of freedom and hence fewer replicas are required to cover a desired "effective" temperature range compared with standard temperature replica exchange. A similar approach where the nonbonded (Lennard-Jones and electrostatic) interactions within the solute as well as between solute and solvent have been scaled to various degrees has also been suggested.[42]

In the present study, we propose an alternative "Hamiltonian" replica-exchange method that focuses on the protein backbone flexibility and employs a specific biasing potential to promote peptide backbone transitions as a replica coordinate. The purpose of the biasing potential is to reduce the energy barriers associated with peptide backbone dihedral transitions. The level of biasing is gradually changed along the replicas such that frequent transitions are possible at high levels of biasing and the system can escape from getting trapped in local energy minima. Since exchanges between replicas are independent of the number of solvent molecules, the method requires much fewer replicas for efficient sampling compared with standard temperature RexMD. The biasing potential RexMD (BP-RexMD) method has been tested on several examples including alanine and threonine dipeptides, a hexa-alanine (Ala_6) system and one small β-hairpin protein with known structure (all including explicit solvent). In all cases, much better sampling of conformational space compared with standard MD simulations was found. At the same time, the approach required considerably fewer replicas (5–7) than the standard temperature RexMD simulations.

METHODOLOGY

Test Systems and Simulation Conditions

The initial extended structures for the alanine (Ala) dipeptide (Ace-Ala-Nme), threonine (Thr) dipeptide (Ace-Thr-Nme), hexa-Ala (Ace-Ala_6-Nme) and the chignolin hairpin peptide (sequence: GYDPETGTWG)[43] were generated using the *xleap* module of the Amber8 package.[44] The Ace and Nme groups represent N-terminal acetyl and C-terminal methylamino capping groups, respectively. In all cases, explicit TIP3P water molecules[45] were added (alanine dipeptide: 560 waters; threonine dipeptide: 547 waters; hexa-Ala: 1046 waters; chignolin hairpin peptide: 1121 waters + 2 sodium counter ions) to form truncated octahedral boxes using *xleap*. The parm03 force field[46] was used for all simulations (without modifications). Each simulation system was subjected to energy minimization (1000 steps) using the *Sander* module. During MD simulation, each peptide was initially harmonically restrained (25 kcal mol^{-1} $Å^{-2}$) to the energy minimized start coordinates (extended peptide structure), and the system was heated up to 300 K in steps of 100 K followed by gradual removal of the positional restraints and 0.2 ns unrestrained equilibration of each system at 300 K. During MD, the long range electrostatic interactions were treated with the particle mesh Ewald (PME) method[47] using a real space cutoff distance of $r_{cutoff} = 9$ Å. The RATTLE algorithm[48] was used to constrain bond vibrations involving hydrogen atoms, which allowed a time step of 2 fs.

Biasing Potentials for Peptide φ and ψ Dihedral Angles

A biasing potential for the φ and ψ peptide backbone dihedral angles was constructed by first calculating a potential of mean force (PMF) for each of the two dihedral angles. This was achieved for the alanine dipeptide case using the umbrella sampling method in combination with the weighted histogram analysis (WHAM) method.[49,50] A quadratic umbrella potential ($k = 200$ kcal mol^{-1} rad^{-2}) and a 5° spacing between reference dihedral angles was used. The φ and ψ peptide backbone dihedral angles are usually defined using the peptide backbone atoms C_{i-1}, N_i, $C\alpha_i$ and C_i (in case of φ_i) and N_i, $C\alpha_i$, C_i and N_{i+1} (in case of ψ_i), respectively. For constructing the biasing potential in case of the φ dihedral angle controlling rotation around, the same bond but employing the atoms C_{i-1}, N_i, $C\alpha_i$, and $C\beta_i$ was used. The advantage of this choice is that glycine (which has already relatively low barriers for backbone dihedral transitions in the unmodified force field) is automatically excluded from the biasing potential application (at least for the φ dihedral angle because it

contains no Cβ). The PMF along the dihedral angles was fitted to a Cosinus-Fourier series of the form:

$$V(\alpha)_{\text{torsion}} = \sum_{n}^{N} \frac{V_{\alpha,n}}{2}[1 + \cos(n\alpha - \gamma)]$$

This potential has the same functional form as used in the Amber force field to control dihedral torsion angles. By changing its sign, it can be used as biasing potential to be added to the dihedral angle potential in the Amber (parm03) force field. Addition of the full biasing potential can in principle offset the PMF along the dihedral angle such that barrier less motion is possible. By adding a scaled biasing potential, the free energy barrier along the peptide backbone dihedral angles can be controlled in small steps that can be used as "replica-coordinate" in the biasing potential-replica exchange (BP-RexMD) simulations. The biasing potential was applied during BP-RexMD either in five or seven steps of the full dihedral biasing potential, and the corresponding parameters for each dihedral angle potential are given in Table I. The broader sampling of the Ramachandran plot in case of adding the biasing potential is illustrated in Figure 1. Note, that the calculated backbone dihedral PMF contains several contributions to the free energy change along the dihedral angle (e.g. nonbonded interactions, solute–solvent contributions, etc.) and is not equivalent to just the dihedral angle dependent term in the original force field. Simple removal of the Amber dihedral angle term still results in significant energy barriers for the peptide backbone angles (this is for example nicely illustrated in Straatsma and McCammon).[16]

TABLE I. Dihedral Angle Parameters for Backbone Dihedral Angles φ' and ψ at Different Biasing Levels

Biasing level 5	φ′ (defined by atoms C, N, CA, CB) multiplicity (n)			ψ (N,CA,C,N) multiplicity (n)	
	1	2	3	1	2
0					
k	0.3537	0.8836	0.227	0.6839	1.4537
δ	3.1415	3.1415	3.1415	3.1415	3.1415
1					
k	0.8252	0.8217	0.377	0.7589	1.2787
δ	2.8658	3.1665	3.0755	3.1415	3.2665
2					
k	1.2968	0.7598	0.527	0.8339	1.1037
δ	2.5900	3.1915	3.0095	3.1415	3.3915
3					
k	1.7684	0.6979	0.677	0.9089	0.9287
δ	2.3143	3.2165	2.9435	3.1415	3.5165
4					
k	2.2400	0.6360	0.827	0.9839	0.7537
δ	2.0385	3.2415	2.8775	3.1415	3.6415
Biasing level 7					
0					
k	0.3537	0.8836	0.227	0.6839	1.4537
δ	3.1415	3.1415	3.1415	3.1415	3.1415
1					
k	0.6680	0.8423	0.327	0.7339	1.3370
δ	2.9577	3.1582	3.0975	3.1415	3.2249
2					
k	0.9824	0.8010	0.427	0.7839	1.2203
δ	2.7739	3.1749	3.0535	3.1415	3.3082
3					
k	1.2968	0.7598	0.527	0.8339	1.1037
δ	2.5900	3.1915	3.0095	3.1415	3.3915
4					
k	1.6112	0.7185	0.627	0.8839	0.9870
δ	2.4062	3.2082	2.9655	3.1415	3.4749
5					
k	1.9256	0.6772	0.727	0.9339	0.8703
δ	2.2224	3.2249	2.9215	3.1415	3.5582
6					
k	2.2400	0.6360	0.827	0.9839	0.7537
δ	2.0385	3.2415	2.8775	3.1415	3.6415

Parameters are given according to the dihedral angle force field term of the form: $V(\alpha) = k\cos(n\alpha + \delta)$, k is given in kcal mol^{-1}, δ is in radians, see also Materials and Methods section.

RexMD Using a Backbone Dihedral Angle Biasing Potential

In standard RexMD, copies or replicas of the system are simulated at different temperature (T_0, T_1, $T_2, \ldots, T_N$). Each replica evolves independently and after 500–1000 MD-steps (~1 ps), an exchange of pairs of neighboring replica is attempted according to the Metropolis criterion:

$$w(x_i \to x_j) = 1 \quad \text{for } \Delta \le 0;$$
$$w(x_i \to x_j) = \exp(-\Delta) \quad \text{for } \Delta > 0$$
where
$$\Delta = (\beta_i - \beta_j)[E(r_j) - E(r_i)]$$

with $\beta = 1/RT$ (R: gas constant and T: temperature) and $E(r)$ representing the potential energy of system for a given configuration. It has been recognized that temperature (represented as Boltzmann factor β) and energy (or Hamiltonian of the system) are equivalent in the Metropolis criterion.[35] Hence, instead of modifying the temperature, it is also possible to scale the force field (or part of it) along the replica coordinate. In the present biasing potential replica exchange method, a biasing potential to allow backbone dihedral angle barrier crossing has been added to the force field (last paragraph). Each replica runs at a different level of added biasing potential (the first replica runs with the original force field, see Table I for the parameters for each biasing level). Exchanges at every 250 steps (0.5 ps) or 500 steps (1 ps) between neighboring biasing levels were attempted according to those in Refs. 25 and 35:

$$w(x_i \to x_j) = 1 \quad \text{for } \Delta \le 0;$$
$$w(x_i \to x_j) = \exp(-\Delta) \quad \text{for } \Delta > 0$$
where
$$\Delta = \beta\left[\left(E^j(r_j) - E^j(r_i)\right) - \left(E^i(r_j) - E^i(r_i)\right)\right]$$

Here, the Metropolis criterion involves only a single β or temperature (in the present study 300 K) and the energy

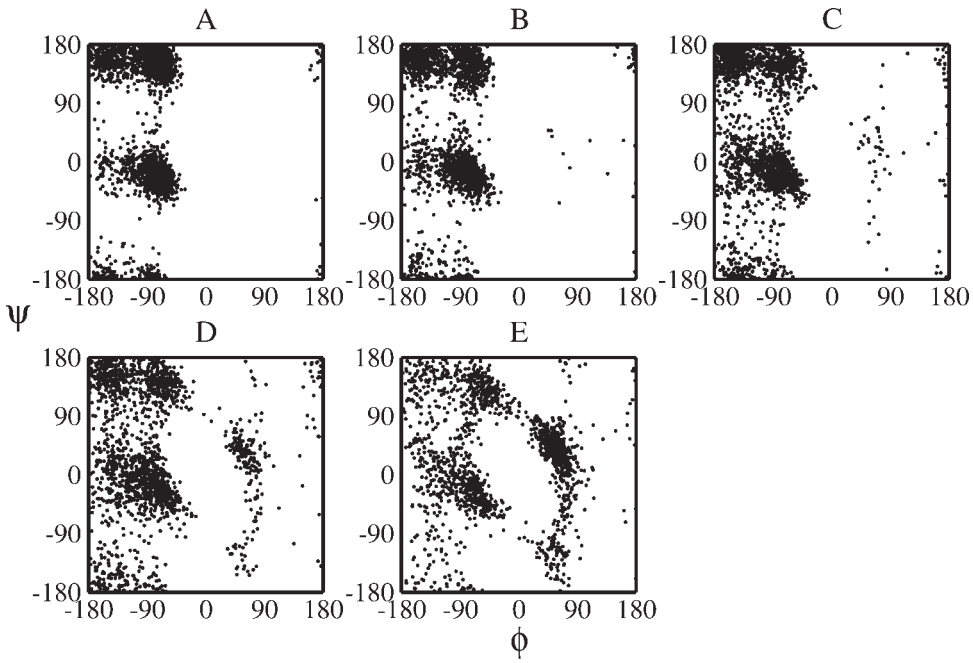

Fig. 1. Comparison of backbone dihedral angle sampling of the alanine dipeptide using 5 dihedral angle biasing potential levels (panels A–E correspond to levels 0–4 in Table I) during 1 ns biasing potential (BP)-RexMD simulation (at 300 K). Each dot in the Ramachandran plots corresponds to a φ–ψ pair of a sampled conformation (conformations were recorded every 1 ps).

difference between neighboring configurations using the force field for replica j (E^j) minus the same difference using force field for replica i (E^i). An advantage compared with temperature RexMD is the fact that the energy differences are only affected by the force field term that change upon going from one replica to another replica run. Hence, the exchange probability is only affected by the backbone dihedral angle terms and not affected by solvent–solvent and solute–solvent (and many other solute–solute) contributions. For the present simulations with 5–7 replicas, the acceptance probability for replica exchanges was in the range of 40–50% for all systems. In the present study, the biasing potential was applied to all φ and ψ peptide backbone dihedral angles of the peptide (except glycine, see above). Note, however, that it is also possible to limit the biasing potential to protein or peptide segments (e.g. loops).

RESULTS

Biasing Potential Replica Exchange Simulations on Dipeptide Test Cases

Potentials of mean force (PMF) for the φ and ψ peptide backbone dihedral angles of the alanine dipeptide in explicit water were obtained using the umbrella sampling approach in combination with the WHAM method. The PMF was used to create a biasing potential (see Methodology section) that was added to the force field description of the peptide in order to lower energy barriers for peptide backbone dihedral angle transitions. During replica exchange simulations, different levels of the biasing potential were added and the corresponding force field parameters to control the backbone dihedral angle potential in the Amber force field are given in Table I. The sampling of φ and ψ peptide backbone dihedral angles at various biasing levels is illustrated in Figure 1, during a 1 ns BP-RexMD simulation on the alanine dipeptide in explicit water. With the original force field [Fig. 1(a)] during 1 ns simulation time (at 300 K), the sampling is dominated by regions in the Ramachandran plot that correspond to α-helical (φ: -160° to -50°; ψ: -60° to $+30^\circ$), β-strand (φ: -180° to -110°; ψ: $+110^\circ$ to $+180^\circ$) as well a P_{II} (polyproline, φ: -110° to -40°; ψ: $+110^\circ$ to $+180^\circ$) states and a few rarely sampled alternative states. With increasing levels of the added biasing potential, the regions in between α-helical and β-strand/P_{II}-regimes are sampled and also other transition regions of the Ramachandran plot [Fig. 1(b–e)]. Note that a uniform sampling of the φ and ψ space was not achieved, presumably, because of inaccuracies of the calculated PMFs (PMFs for φ and ψ were calculated independently) or the fitting to a Cosine-series. However, for the present RexMD method, a uniform sampling of the Ramachandran plot at the full level of the biasing potential is not required. It is not even desirable because parts of the Ramachandran plot correspond to peptide conformations with severe sterical atom overlap. These conformations should be avoided also in the replicas that run in the presence of the biasing potential since such "unphysical" conformations do not correspond to transition regions and have also little chance to be accepted in the run (replica) with the original force field. The main purpose of the biasing potential

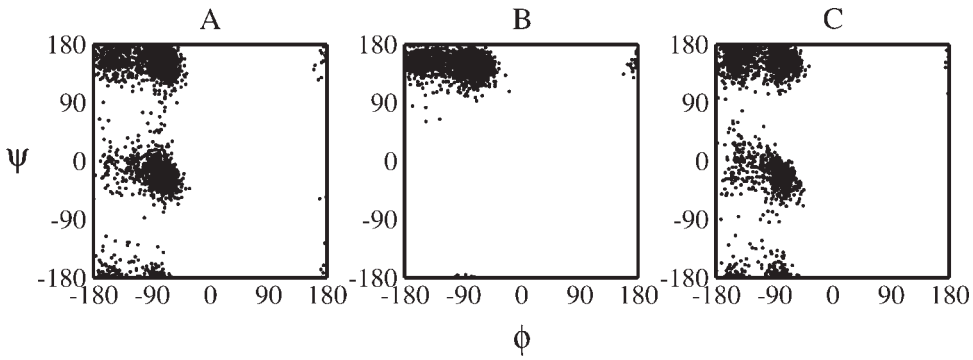

Fig. 2. Comparison of backbone dihedral angle sampling of the alanine dipeptide during 1 ns BP-RexMD (at 300 K) with 5 replicas (5 levels of the dihedral angle biasing potential, see Methods section and Table I) (**A**), during 1 ns conventional MD (**B**) and during 10 ns conventional MD (**C**). For the BP-RexMD only the sampling for the replica at the original force field is shown. Each dot in the Ramachandran plots corresponds to a φ–ψ pair recorded every 1 ps (A, B) or 10 ps (C) to achieve the same total number of dots for each case.

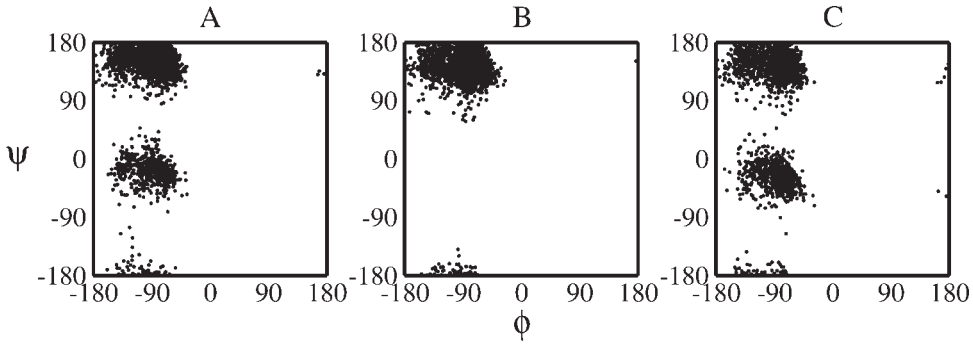

Fig. 3. Comparison of backbone dihedral angle sampling of the threonine dipeptide during 1 ns BP-RexMD with 5 biasing levels (**A**, result for the run with the original force field); during 1 ns standard MD (**B**) and during 10 ns standard MD (**C**). Each dot in the Ramachandran plots corresponds to a φ–ψ pair recorded every 1ps (A, B) or 10 ps (C) to achieve the same total number of dots for each case.

during replica exchange simulations is to lower energy barriers for backbone dihedral transitions and enhance sampling at the transition regions between favorable peptide substrates (see Fig. 1). To demonstrate the efficiency of the biasing potential replica exchange approach, the sampling of the φ and ψ dihedral angles for alanine dipeptide during 1 ns BP-RexMD with 5 replicas [Fig. 2(a)] was compared with a 1 ns MD and 10 ns MD simulation [Fig. 2(b,c)], respectively, using the same start structure (extended conformation, all at 300 K simulation temperature). During the BP-RexMD replica-exchanges were attempted every 500 steps (1 ps) between the biasing potential levels given in Table I (5-replica-level-case). The sampling given in Figure 2 was obtained for the reference replica with the original force field (no biasing potential). During a 1 ns MD simulation, the conventional MD approach sampled only the region of the Ramachandran plot close to the initial structure [Fig. 2(b), e.g. the sampling result depends strongly on the start conditions]. The sampling of the Ramachandran plot during the 1 ns BP-RexMD was very similar to the sampling obtained from a longer MD simulation of 10 ns [compare Fig. 2(a) and (c)]. Note, however, that the total MD simulation time for the replica exchange simulation amounts to 5 ns (5 × 1 ns). Also, even a 10 ns simulation may significantly under sample the available conformational space for alanine dipeptide.[41]

The BP-RexMD approach was also applied to the threonine dipeptide in order to test its efficiency for β-branched amino acids. Also, in this case, a much quicker exploration of the Ramachandran plot during short simulation times was observed when compared with standard MD starting from the same initial conditions (see Fig. 3). It is interesting to note that in the case of the β-branched threonine dipeptide, the P^{II} conformational regime is even more dominantly sampled than in the alanine dipeptide case (see Fig. 2). This observation may relate to the experimental observation of a greater propensity of threonine to be part of β-strands (closer to P^{II} than the α-helical state) when compared with alanine.[51]

BP-RexMD-Application to Hexa-Ala-Peptide

Starting from an extended conformation, a BP-RexMD simulation of Ace-$(Ala)_6$-Nme (termed hexa-Ala) in explicit solvent was performed using 5 levels of the biasing potential. There is experimental evidence that oligo-alanine peptides adopt a polyproline II conformation in

TABLE II. Distribution of Peptide Backbone Conformational States Observed During MD Simulations

	Alanine dipeptide	Hexa-Ala (Ace-$(Ala)_6$-Nme)		
Secondary structure	5 ns standard MD	BP-RexMD	Standard MD	Temp RexMD
Alpha (R) (α^R)	41.2	51.9	53.8	45.6
Beta (β)	15.6	10.5	8.4	9.2
P^{II}	32.2	26.8	29.2	29.6
Alpha(L) (α^L)	2.0	0.2	0.58	2.4

Numbers are given as percentages. The numbers given for the Hexa-Ala case are from 5 ns simulation time. The types of secondary structures are defined as Alpha (R) (α^R-helical; φ: -160° to -50°; ψ: -60° to $+30^\circ$), Beta (β-strand; φ: -180° to -110°; ψ: $+110^\circ$ to $+180^\circ$), P^{II} (polyproline; φ: -110° to -40°; ψ: $+110^\circ$ to $+180^\circ$), and Alpha(L) (α^L-helical; φ: $+20^\circ$ to $+70^\circ$; ψ: -30° to $+70^\circ$).

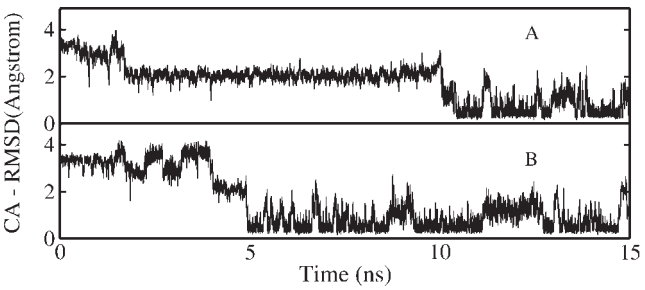

Fig. 4. Root-mean-square deviation (Rmsd of Cα-atoms) of sampled hexa-Ala (Ace-$(Ala)_6$-Nme) conformations from a standard α-helix vs. simulation time. (**A**) Hexa-Ala-Rmsd obtained during 15 ns standard MD at 300 K; (**B**) same for a BP-RexMD (for the replica run with the original force field) without exchanges between replicas but simulation restarts every ps (coordinates and velocities at every restart are taken from stored files). Both simulations started from an extended start conformation.

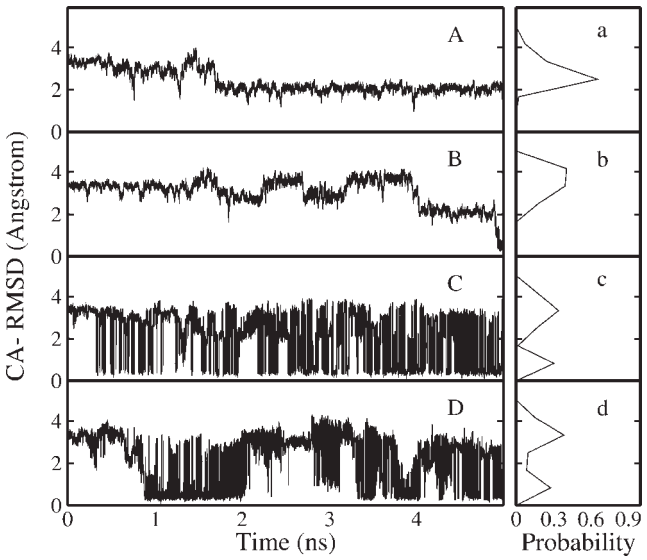

Fig. 5. Rmsd of Cα-atoms of sampled hexa-Ala (Ace-$(Ala)_6$-Nme) conformations from a standard α-helix vs. simulation time. (**A**) Hexa-Ala-Rmsd obtained during 5 ns standard MD at 300 K [first 5 ns of the plot shown in Fig. 4(a)]; (**B**) same for a BP-RexMD without exchanges [original force field, first 5 ns of the Rmsd curve shown in Fig. 4(b)]; (**C**) same for a temperature RexMD (Rmsd curve for the replica run at 300 K); (**D**) same for BP-RexMD (for the replica simulation with the original force field). All simulations were started from the same fully extended start structure. The panels on the right of each Rmsd plot (a–d) correspond to the Rmsd-distribution (with respect to the α-helical state) during the second half of each simulation.

solution.[52] To check whether the parm03 force field parameters employed in the present study favor a P^{II} conformation, a standard temperature replica exchange simulation with 16 replicas was performed (~6 ns, simulation temperatures: 300, 303, 306.6, 310.8, 316.2, 321, 327.6, 334.2, 341.4, 349.2, 357.6, 366.6, 376.6, 386.4, 397.2, 408.6 K). The above temperature spacing resulted in an exchange acceptance ratio of ~30%. As a second control, a long MD reference simulation (300 K, 15 ns) starting from the same extended hexa-Ala structure was performed. In all the simulations, the P^{II} state was sampled extensively (Table II) with a contribution of ~29% similar to the P^{II} probability found for the alanine dipeptide simulations (Table II). However, in all cases, the α-helical state had considerably higher probability than the P^{II} state (Table II). The bias of the Amber force fields toward "over-stabilization" of the α-helical state has been noted in previous studies and several attempts have been made to correct for it.[30,33,37] Consequently, in both the standard MD as well as in the temperature replica exchange simulations conformations with small root mean square deviation (Rmsd) with respect to an α-helical reference for the hexa-Ala appeared as the most frequently sampled state during the final part of the simulations (Figs. 4 and 5). This state was reached after ~8 ns during the conventional MD simulation (see Fig. 4) and was the most sampled state for the rest of the simulation (~15 ns). It is important to note that for the purpose of testing the BP-RexMD approach, the artificial stabilization of the α-helical conformation of the hexa-Ala by the current force field is even desirable since it defines a stable target structure for the hex-Ala simulations. For further comparison, a BP-RexMD simulation without exchanges between replicas but following otherwise exactly the same BP-RexMD protocol (see Methodology section) was run. The resulting trajectories were compared with an energy-minimized standard α-helical conformation of the hexa-Ala sequence. Ultimately, all simulations lead to a significant proportion of α-helical states at the final stages of the simulations. However, low-free energy α-helical conformations were much more rapidly sampled (in less than 1 ns) in case of the standard RexMD (with 16 temperature replicas, see above) and in the BP-RexMD compared with the standard MD or the BP-RexMD without exchanges (Figs. 4 and 5). The distribution of α-helical vs. non-α-helical structures during the last 2 ns of the simulation is similar for the temperature RexMD and the BP-RexMD simula-

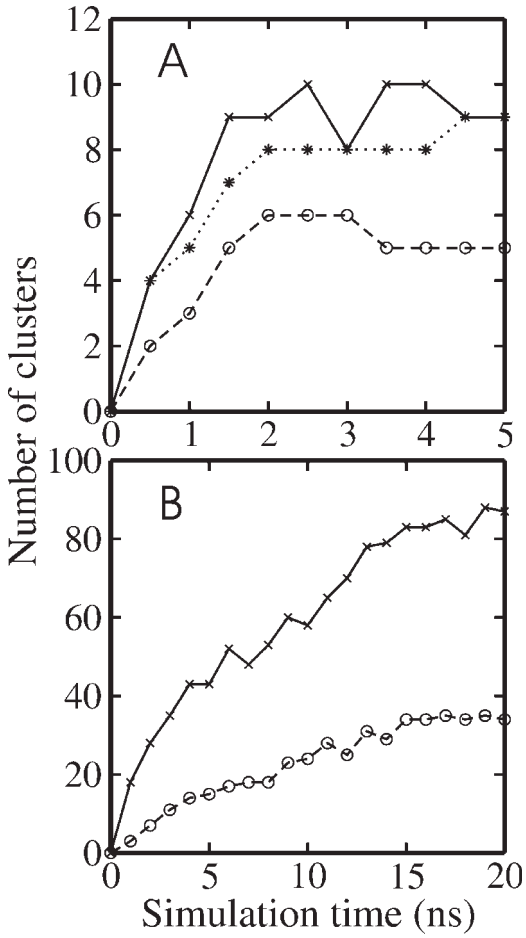

Fig. 6. Accumulation of conformational clusters during standard MD simulations (dashed lines) and BP-RexMD simulations (continuous lines) of hexa-Ala (**A**) and chignolin (**B**) peptides. In case of hexa-Ala (A), the result of temperature RexMD is also shown (dotted lines). Cluster analysis was performed on all recorded structures up to the simulation time given on the *x*-axis using the program kclust of the MMTSB tools[53] and a 2 Å Rmsd(Cα) exclusion cutoff for the distance of conformations relative to each cluster center. The number of accumulated distinct clusters is plotted vs. simulation time.

tions (~40% α-helix if one counts all conformations within and Rmsd of 2 Å from the helical reference as α-helix, Fig. 5). The helix probability translates to a free energy of helix formation close to zero. However, this can only be considered as an estimate since accurate converged sampling of conformational probability distributions may require significantly longer simulation times even with the present replica exchange method. Besides the rapid sampling of α-helical states, a cluster analysis of the sampled peptide conformations clearly demonstrates a much more efficient sampling of the hexa-Ala conformations using the BP-RexMD (or temperature RexMD) compared with standard continuous MD simulations (see Fig. 6). Cluster analysis was based on the pair-wise Cartesian (backbone) Rmsd between conformations with an Rmsd cutoff of 2 Å and using the kclust program in the MMTSB-tools.[53] Both the BP-RexMD and the temperature RexMD covered approximately twice the number of distinct conformational clusters after a few nanoseconds of simulations time [Fig. 6(a)]. Besides a dominant cluster representing the α-helical state, two alternative states that were also significantly populated correspond to β-hairpin type structures with the turn located at different positions along the sequence (not shown).

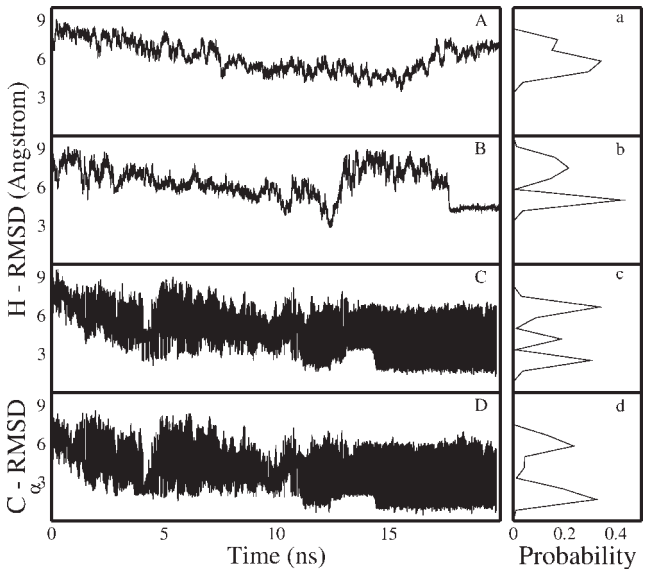

Fig. 7. Rmsd of sampled chignolin peptide conformations from the experimental NMR structure (first entry of pdb1ua0) vs. simulation time starting from a fully extended peptide structure. (**A**) Heavy atom-Rmsd obtained during 20 ns standard MD at 300 K; (**B**) same for a BP-RexMD without exchanges (original force field); (**C**) same for a BP-RexMD (Rmsd curve for the replica with the original force field); (**D**) same as (C) but showing the Cα-Rmsd. The panels on the right of each Rmsd plot (a–d) correspond to the Rmsd-distribution (with respect to experimental structure) during the second half of each simulation.

Folding Simulations on a β-Hairpin Forming Peptide

The efficiency of the BP-RexMD approach was further evaluated on the chignolin peptide, one of the smallest β-hairpin peptides known to be stable in solution.[43] The structure of this protein was recently determined by NMR experiments.[43] It has also been demonstrated that extensive conventional temperature RexMD simulations (using 16 replicas) of more than 100 ns including ~890 water explicit molecules can lead to a folded structure very similar to the experimental NMR structure.[10]

A BP-RexMD with 7 biasing potential levels was used to study this peptide in explicit solvent simulations with more than 1100 water molecules starting from an extended conformation. Within ~5–10 ns transitions to a conformation with a backbone Rmsd of ~1.9 Å (heavy atom Rmsd: 2.8 Å) with respect to the experimental NMR structure[43] (first structure of NMR ensemble in Protein Data Bank entry 1UA0) appeared (see Fig. 7). After ~15 ns, increased sampling of conformations with a backbone Rmsd < 1 Å (heavy atom Rmsd: ~1.5 Å) was observed (see Fig. 7). The superposition of a snapshot from the last part of the BP-RexMD simulation onto the experimental structure indicates very close agreement [Fig. 8(b)] and a large

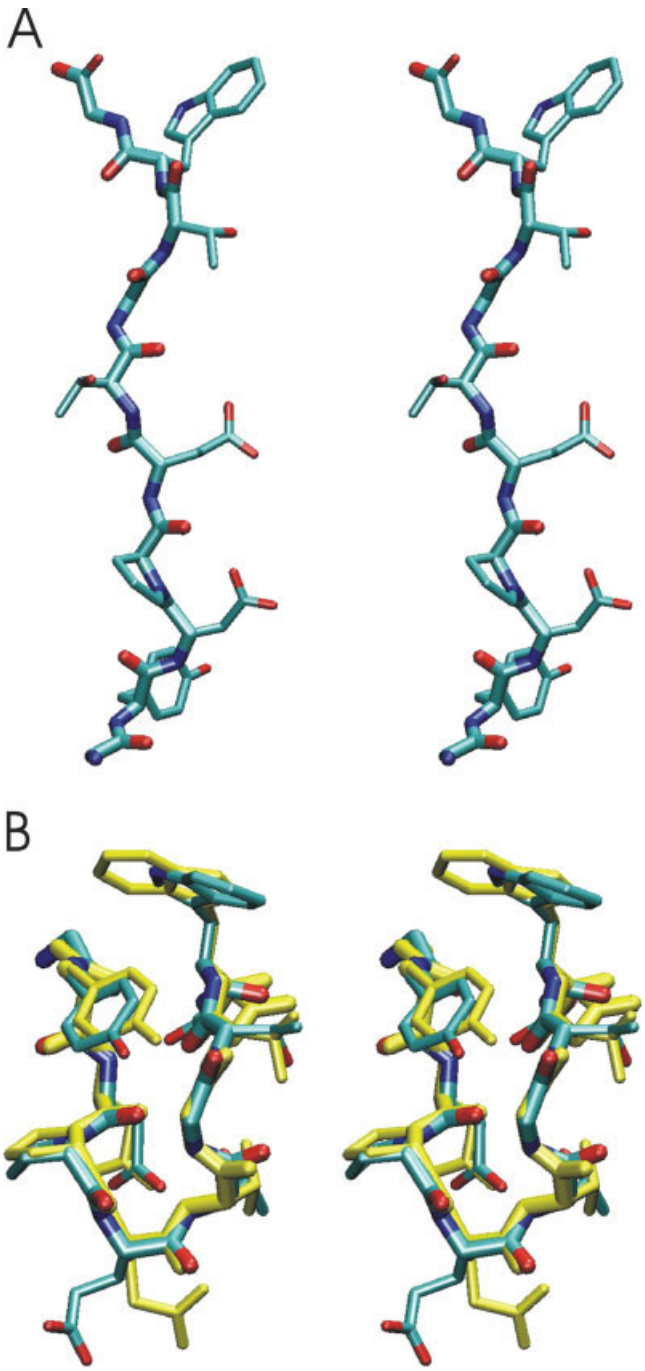

Fig. 8. (**A**) Stereo view of the extended chignolin peptide start structure (atom color code, only heavy atoms are shown). The Rmsd (heavy atoms) of the start structure from the experimental structure (pdb1ua0) was 6.5 Å. (**B**) Stereo view of a folded chignolin peptide structure (atom color code) obtained after ~15 ns PB-RexMD with a heavy atom Rmsd of 1.7 Å from experiment superimposed on the experimental structure (yellow).

scale conformational transition with respect to the extended start structure [Fig. 8(a)]. Note that the average pair-wise Rmsd between the NMR models is in the same order of ~1 Å (pdb1UA0). This indicates that the level of agreement with experiment as observed during the simulation is within the uncertainties of the NMR structure determination. Around 40% of the conformations sampled during the last 5 ns of the BP-RexMD are within a backbone Rmsd of 1.5 Å from the reference structure. This fraction is slowly increasing over time indicating that for obtaining a converged probability distribution longer simulations might be necessary. In none of the control simulations (standard MD at 300 K or RexMD without exchanges), peptide conformations within 1.5 Å from the reference structure were observed indicating that time scales beyond 20 ns are required to reach a folded state for this peptide in standard MD approaches (control simulations beyond 40 ns still did not lead to folded structures, not shown).

Similar to the previous example (hexa-Ala), a cluster analysis of the simulations [Fig. 6(b)] indicates rapid increase in the number of sampled distinct conformational states (clusters) already during the early phase of the BP-RexMD simulation whereas very limited coverage and much slower appearance of new distinct clusters during the entire simulation time (20 ns) was observed for the conventional MD simulation (see Fig. 6).

DISCUSSION

Replica exchange MD simulations are frequently used to enhance the conformational sampling during MD simulations.[7–12,27,29–34] A drawback of the conventional temperature RexMD is the rapid increase of the number of replicas with increasing system size to cover a desired temperature range.[35] The ratio of the standard deviation of the system potential energy (a measure of the energy fluctuation) vs. average energy decreases with the square-root of the system size. Hence, to achieve sufficient overlap of the energy distributions between replicas run at different temperatures (required to achieve a reasonable exchange acceptance ratio), the temperature "spacing" between neighboring replicas is required to decrease with system size. Another drawback of large numbers of replicas is the need to run longer simulations (or more exchanges) to allow sufficient "travelling" or exchanges between high and low temperature replicas when compared with a small number of replicas. Especially in case of simulations that include a large number of explicit water molecules, the rapid increase of the number of replicas in temperature RexMD simulations limits the applicability to peptide or small protein systems. To decrease the number of replicas during RexMD simulations, separate temperature coupling of solute and solvent degrees of freedom has been used with the solvent temperature kept at the reference temperature.[38] A hybrid explicit/implicit solvent RexMD approach has been developed where all replicas are run in the presence of explicit solvent but for the evaluation of the exchange probability part of the water is replaced by a continuum description.[37] Another alternative is to scale the potential energy function (Hamiltonian) along the replicas.[35,39–42] Promising Hamilitonian replica exchange approaches have been developed that scale nonbonded interactions (Lennard-Jones and electrostatic interactions) within the solute and between solute and sol-

vent to various degrees.[41,42] A critical issue in the application of such Hamiltonian replica exchange methods is the choice and magnitude of force field energy terms to be scaled along the replicas. In the present study, a new type of Hamiltonian-RexMD method has been presented that employs various levels of biasing potential for φ and ψ peptide backbone dihedral angles along the system replicas. The biasing potential lowers the barrier for backbone dihedral transitions and promotes an increased tendency for peptide backbone transitions along the replica coordinate. Such backbone biasing potentials have been used successfully in previous simulation studies on oligo-Ala peptides[16] and on cyclic peptides[20] and proteins[19] (in single conventional simulations). In these applications, the peptide backbone dihedral potential was scaled or a biasing potential was applied during an early stage of a simulation and gradually transformed to the original potential during a single MD simulation. This is similar to a simulated annealing simulation where typically one starts from a high simulation temperature and gradually reduce the temperature to a low value with the drawback that the result of the simulation depends strongly on the system and speed of temperature decrease during the annealing run. The φ and ψ dihedral angles are the main conformational determinants of the main chain conformation of peptides and proteins and correspond to soft degrees of freedom for the peptide structure with relatively small transition energy barriers. In contrast to temperature RexMD and other Hamiltonian-RexMD approaches, the present BP-RexMD method employs structural knowledge on peptides and proteins to design a replica coordinate most appropriate for enhanced peptide/protein conformational sampling because it focuses on a soft degree of freedom of a peptide or protein. Only between 5 and 7 replicas were necessary for the hexa-Ala-peptide and the hairpin forming peptide (chignolin), respectively, including between 600 (hexa Ala) and 1100 (chignolin) explicit water molecules to achieve folding of these peptides in explicit solvent to conformations in close agreement with experimental structures (or a stable structure for the Amber force field in case of the hexa-Ala peptide). The relatively high replica-exchange acceptance probability of 40–50% (compared with ~30% in case of most temperature replica simulations) for these systems indicates that an optimization of the approach with respect to the number of replicas may result in even smaller number of replicas for typical peptide simulation systems. Temperature replica exchange simulations on the same systems required 16 replicas in case of the hexa-Ala system (present study) and 16 replicas in a case of a published study on the chignolin peptide[10] including less water molecules than the present study. In contrast to the standard temperature RexMD, for the present BP-RexMD, the number of required replicas is independent of the number of water molecules included during the simulation. Only the biasing potential energy term enters into the exchange probability meaning that the number of required replicas is expected to scale approximately linearly with the number of included backbone dihedral angles. A drawback compared with standard RexMD methods, however, is the fact that the present BP-RexMD is restricted to peptide or protein simulations and per se not generally applicable to any organic or biomolecule of interest. Also, for each peptide force field, a specific biasing potential needs to be constructed to apply the current method. However, it is in principle also possible for other types of biomolecules to identify the most important variables that control the biomolecule structure and to construct an appropriate biasing potential. Another advantage is that the replica exchange sampling can be easily focused to parts of a protein (e.g. a loop region) keeping the "unbiased" (original) backbone dihedral angle potential for the rest of the protein in all replicas. For example, comparative protein modeling based on sequence similarity of a protein to a protein with known structure often relies on a nonuniform similarity along the sequence alignment. Parts of the protein can be modeled with high confidence and other regions (e.g. loop regions with low target-template similarity) have to be modeled by ab initio methods. In this case, enhanced sampling of parts of proteins under realistic conditions (in the presence of explicit water molecules) is desired during refinement steps of comparative protein modeling or at protein–protein and protein–ligand interfaces. Instead of independent biasing potentials for peptide backbone dihedral angles, it is also possible to apply a similar approach to a collective degree of freedom (e.g. a combination of dihedral angles). The present BP-RexMD method is ideally suited to refine model proteins by focusing the biasing potential during BP-RexMD to critical protein segments. This further limits the number of required replicas. Other applications of the present BP-RexMD could include improved sampling of peptide backbone conformations during ab initio protein or peptide folding simulations in explicit solvent. It is also straight forward to extend the method to include enhanced conformational transition of side chain conformations by constructing an appropriate biasing potential for amino acid side chain transitions.

ACKNOWLEDGMENTS

This work was performed using the computational resources of the CLAMV (Computational Laboratories for Analysis, Modeling and Visualization) at IUB and supercomputer resources of the EMSL (Environmental Molecular Science Laboratories) at the PNNL (Pacific Northwest National Laboratories). S.K. is supported by the BIOlogical RECognition graduate program at IUB and Volkswagen Stiftung.

REFERENCES

1. Daura X, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Reversible peptide folding in solution by molecular dynamics simulation. J Mol Biol 1998;280:925–932.
2. Duan Y, Kollman PA. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. Science 1998;282:740–744.
3. Roccatano D, Amadei A, Di Nola A, Berendsen HJ. A molecular dynamics study of the 41–56 β-hairpin from b1 domain of protein G. Protein Sci 1999;10:2130–2143.
4. Pande VS, Roshkar DS. Molecular dynamics simulations of unfolding and refolding of a β-hairpin fragment of protein G. Proc Natl Acad Sci USA 1999;96:9062–9067.

5. Garcia AE, Sanbonmatsu KY. Exploring the energy landscape of a β-hairpin in explicit solvent. Proteins 2001;42:345–354.
6. Simmerling C, Strockbine B, Roitberg AE. All-atom structure prediction and folding simulations of a stable protein. J Am Chem Soc 2002;124:11258–11259.
7. Zhou R, Berne BJ, Germain R. The free energy landscape for β-hairpin folding in explicit water. Proc Natl Acad Sci USA 2001;98: 14931–14936.
8. Rao F, Caflisch A. Replica exchange molecular dynamics simulations of reversible folding. J Chem Phys 2003;119:4035–4042.
9. Roccatano D, Nau WM, Zacharias M. Structural and dynamic properties of the CAGQW peptide in water: a molecular dynamics simulation study using different force fields. J Phys Chem 2004;108: 18734–18742.
10. Seibert MM, Patriksson A, Hess B, van der Spoel D. Reproducible polypeptide folding and structure prediction using molecular dynamics simulations. J Mol Biol 2005;354:173–183.
11. Nguyen P, Stock G, Mittag E, Hu C-K, Li MS. Free energy landscape and folding mechanism of a β-hairpin in explicit water: a replica exchange molecular dynamics study. Proteins 2006;61:795–808.
12. Gnanakaran S, Nymeyer H, Portman J, Sanbonmatsu KY, Garcia AE. Peptide folding simulations. Curr Opin Struct Biol 2003;15: 168–174.
13. Kaihsu T. Conformational sampling for the impatient. Biophys Chem 2004;107:213–220.
14. Brunger AT, Adams PD, Rice LM. New applications of simulated annealing in X-ray crystallography and solution NMR. Structure 1997;5:325–336.
15. Kostrowicki J, Scheraga HA. Application of the diffusion equation method for global optimization to oligopeptides. J Chem Phys 1992;96:7442–7449.
16. Straatsma TP, McCammon JA. Treatment of rotational isomers. III. The use of biasing potentials J Chem Phys 1994;101:5032–5039.
17. Huber T, Torda AE, van Gunsteren WF. Structure optimization combining soft-core interaction functions, the diffusion equation method and molecular dynamics. J Phys Chem A 1997;10:5926–5930.
18. Tappura K, Lahtela-Kakkonen M, Teleman O. A new soft-core potential function for molecular dynamics applied to the prediction of protein loop conformations. J Comput Chem 2000;21:388–397.
19. Tappura K. Influence of rotational energy barriers to the conformational search of protein loops in molecular dynamics and ranking the conformations. Proteins Struct Funct Genet 2001; 44:167–179.
20. Riemann RN, Zacharias M. Reversible scaling of dihedral angle barriers during molecular dynamics to improve structure prediction of cyclic peptides. J Pept Res 2004;63:354–364.
21. Riemann RN, Zacharias M. Refinement of protein cores and protein-peptide interfaces using a potential scaling approach. Prot Eng Des Select 2005;18:465–476.
22. Hornak V, Simmerling C. Generation of accurate protein loop conformations through low-barrier molecular dynamics. Proteins 2003;51:577–590.
23. Simmerling C, Miller JL, Kollman PA. Combined locally enhanced sampling and partical mesh Ewald as a strategy to locate the experimental structure of a nonhelical nucleic acid. J Am Chem Soc 1998;120:7149–4.
24. Swendsen RH, Wang JS. Replica Monte Carlo simulations of spin glasses. Phys Rev Lett 1986;57:2607–2609.
25. Okamoto Y. Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. J Mol Graph Model 2004;22:425–439.
26. Predescu C, Predescu M, Ciobanu CVJ. On the efficiency of exchange in parallel tempering Monte Carlo simulations J Phys Chem B 2005;109:4189–4196.
27. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 1999;314:141–151.
28. Okabe T, Kawata M, Okamoto Y, Mikami M. Replica-exchange Monte Carlo method for the isobaric–isothermal ensemble. Chem Phys Lett 2001;335:435–439.
29. Mitsutake A, Sugita Y, Okamoto Y. Generalized-ensemble algorithms for molecular simulations of biopolymers. Biopolymers 2001;60:96–123.
30. Sanbonmatsu KY, Garcia AE. Structure of Met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. Proteins 2002;46:225–234.
31. Zhou R, Berne BJ. Can a continuum solvent model reproduce the free energy landscape of a β-hairpin folding in water? Proc Natl Acad Sci USA 2002;99:12777–12782.
32. Zhou R. Free energy landscape of protein folding in water: explicit vs. implicit solvent. Proteins 2003;53:148–161.
33. Nymeyer H, Garcia AE. Simulation of the folding equilibrium of α-helical peptides: a comparison of the generalized born approximation with explicit solvent. Proc Natl Acad Sci USA 2003;100: 13934–13939.
34. Yoshida K, Yamaguchi T, Okamoto Y. Replica-exchange molecular dynamics simulation of small peptide in water and in ethanol. Chem Phys Lett 2005;41:2280–2284.
35. Fukunishi H, Watanabe O, Takada S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction. J Chem Phys 2002;116:9058–9067.
36. Bashford D, Case DA. Generalized born models of macromolecular solvation effects. Annu Rev Phys Chem 2000;51:129–152.
37. Okur A, Wickstrom L, Layten M, Geney R, Song K, Hornak V, Simmerling CJ. Improved efficiency of replica exchange simulations through use of a hybrid explicit/implicit solvation model. Chem Theory and Comput 2006;2:420–433.
38. Cheng X, Cui G, Hornak V, Simmerling C. Modified replica exchange simulation methods for local structure refinement. J Phys Chem B 2005;109:8220–8230.
39. Jang S, Shin S, Pak Y. Replica-exchange method using the generalized effective potential. Phys Rev Lett 2003;91:58305–58309.
40. Zhu Z, Tuckerman ME, Samuelson SO, Martyna GJ. Using novel variable transformations to enhance conformational sampling in molecular dynamics. Phys Rev Lett 2002;88:100201.
41. Liu P, Kim B, Friesner RA, Berne BA. Replica exchange with solute tempering: a method for sampling biological systems in explicit water. Proc Natl Acad Sci USA 2005;102:13749–13754.
42. Affentranger R, Tavernelli I, Di Iorio EE. A novel Hamiltonian replica exchange MD protocol to enhance protein conformational space sampling J Chem Theory Comput 2006;2:217–228.
43. Honda S, Yamasaki K, Sawada Y, Morii H. 10 Residue folded peptide designed by segment statistics. Struct Fold Des 2004;12:1507–1518.
44. Case D, Pearlman DA, Caldwell JW, Cheatham TE, III, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Tsui V, Radmer RJ, Duan Y, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, Kollman PA. Amber 8. University of California: San Francisco; 2003.
45. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
46. Duan Y, Wu A, Chowdhury CS, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 2003;24:1999–2012.
47. Darden T, York D, Pedersen L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. J Chem Phys 1993;98: 10089–10092.
48. Miyamoto S, Kollman PA. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. J Comput Chem 1992;13:952–962.
49. Kumar SD, Bouzida R, Swendsen H, 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.
50. Grossfield A. 2003. Weighted histogram analysis method. http://dasher.wustl.edu/alan.
51. Chou PY. Prediction of protein structural class from amino acid composition. In: Fasman GD, editor. Prediction of protein structure and the principles of protein conformation. New York: Plenum Press; 1989. pp 549–586.
52. Shi Z, Olson CA, Rose GD, Baldwin RL, Kallenbach NR. Polyproline II structure in a sequence of seven alanine residues. Proc Natl Acad Sci USA 2002;99:9190–9195.
53. Feig M, Karanicolas J, Brooks CL. MMTSB tool set: enhanced sampling and multiscale modeling methods for applications in structural biology. J Mol Graph Model 2004;22:377–395.