JCTC Journal of Chemical Theory and Computation

Improved Replica Exchange Method for Native-State Protein Sampling

Samuel L. C. Moors,* Servaas Michielssens, and Arnout Ceulemans

Department of Chemistry, Katholieke Universiteit Leuven, Celestijnenlaan 200F, 3001 Heverlee, Belgium

Received August 29, 2010

Abstract: We present a new replica exchange method, designed for optimal native state protein sampling in explicit solvent, called replica exchange with flexible tempering (REFT). The method was built upon the recently introduced replica exchange with solute tempering (REST). The potential function is adapted to direct the conformational search toward interdomain movements and the flexible portions of the protein. We demonstrate the improved sampling efficiency of REFT compared to the original REST for the bacteriophage T4 lysozyme.

Introduction

The replica exchange method (REM) is a parallel simulation method aimed to improve the sampling efficiency of complex systems with rugged energy landscapes such as proteins.[1−3] In REM, multiple replicas are simulated in parallel in different thermodynamic states. At regular intervals during the simulation, exchanges between replicas are attempted. The acceptance probability for exchange is based on the detailed balance condition, which ensures that the canonical ensemble is conserved in the limit of an infinitely long simulation time.[4] In its original implementation, the temperature was chosen as the state variable. Exchanges between low and high-temperature states allow the system to cross high barriers much more easily. The temperature REM (T-REM) method proved to be very powerful for small systems such as peptides and small proteins.[5−7] However, as the number of replicas increases with the square root of the system degrees of freedom,[8] its use for medium sized proteins or larger becomes impractical.

To overcome the limitations of the original T-REM, many variations have been suggested in the literature.[9] Hamiltonian REM (H-REM), which uses the potential energy function as the state variable, is particularly promising.[8] In H-REM only parts of the Hamiltonian are scaled between replicas[10] as opposed to the temperature, which is equivalent to a uniform scaling of the potential energy. Reported H-REM implementations include scaled hydrophobicity and hydrophobic aided REM (mimicking the chaperone effect by lowering the hydrophobic interactions),[8,11] phantom chain and soft-core REM (allowing partial atomic overlaps),[8,12] REM with peptide backbone biasing potential (lowering the backbone dihedral angle rotational barriers).[13] Zacharias developed a H-REM method that applies a penalty potential to drive the protein along the soft modes, which were determined from elastic network model calculations.[14]

Berne and co-workers introduced an interesting REM variant, called replica exchange with solute tempering (REST).[15] In REST, the potential energy scales with the temperature in such a way that the solvent appears cold even at high temperatures. The mutual interaction energy between water molecules disappears from the exchange acceptance criterion, giving rise to higher exchange probabilities. Because the bulk of the system degrees of freedom stems from the water molecules, far less replicas are needed.

In this article, we propose to advance the REST method for native-state sampling by using a dual scale approach. Proteins can be subdivided into rigid and flexible regions. Rigid domains usually consist of (single or multiple) secondary structure elements like α-helices and β-sheets. Rigid domains are often connected by flexible hinges or loops, allowing large movements between them. We hypothesize that the sampling rate of native proteins is limited by the slow large-scale movements between rigid domains. In the new method, we selectively heat-up the flexible parts, keeping the rigid domain interactions cold. As such, interdomain motions are facilitated, while the domains themselves maintain a nativelike state. We refer to this method as replica

* Corresponding author phone: (32)16/32.73.84; fax: (32)16/32.79.92; e-mail: sam.moors@chem.kuleuven.be.

10.1021/ct100493v
Published on Web 12/07/2010

exchange with flexible tempering (REFT). The REFT method was applied to the native-state sampling of the bacteriophage T4 lysozyme (T4L) and the sampling efficiency was compared with the REST method.

Materials and Methods

Replica Exchange with Solute Tempering. The REST method is a combination of H-REM and T-REM. In REST, the potential energy function E_0 is decomposed into three terms

$$E_0 = E_{pp} + E_{ww} + E_{pw} \quad (1)$$

where pp, ww, and pw correspond to the internal protein−protein, water−water, and protein−water interactions respectively. The ww and pw energy terms are scaled with the simulation temperature

$$E_i = E_{pp} + (\beta_0/\beta_i)E_{ww} + \sqrt{\beta_0/\beta_i}E_{pw} \quad (2)$$

where β is the reciprocal temperature $1/k_BT$ with k_B the Boltzmann constant. The indices 0 and i refer to the lowest and ith temperature replica. Note that we have slightly adapted the original scaling factor of the pw term $(\beta_0 + \beta_i)/2\beta_i$,[15] to allow defining the rigid and flexible atoms at the force field level.

Replica Exchange with Flexible Tempering. The REFT method further categorizes the protein atoms as either rigid or flexible. The terms E_{pp} and E_{pw} are further subdivided

$$E_0 = E_{ff} + E_{rr} + E_{ww} + E_{fr} + E_{fw} + E_{rw} \quad (3)$$

where ff, rr, ww, fr, fw, and rw represent the flexible−flexible, rigid−rigid, water−water, flexible−rigid, flexible−water, and rigid−water interactions, respectively. The following scaling was used for the potential energy of the higher-temperature replicas

$$E_i = E_{ff} + (\beta_0/\beta_i)E_{ww,rr,wr} + \sqrt{\beta_0/\beta_i}E_{fw,fr} \quad (4)$$

As in eq 2, this scaled potential allows defining the rigid and flexible atoms on the force field level by appropriately scaling the parameters of individual atoms, pairs, triplets and quadruplets. Both REST and REFT satisfy the detailed balance condition, and thus lead to the canonical ensemble in the limit of infinite sampling.

Determination of Rigid Domains. The floppy inclusion and rigid substructure topography (FIRST) program identifies rigid domains in a single protein conformation based on geometrical constraints to model covalent bonds, hydrogen bonds, and hydrophobic tethers.[16] Constraints are defined based on a cutoff energy of interaction. The FIRST analysis was performed on the closed-state X-ray structure (2LZM). Only clusters with more than 15 atoms were included. The constraint cutoff energy value was chosen such that the number of rigid atoms is approximately one-third of all protein atoms. For T4L, a cutoff energy of 0.5 kcal/mol was used, resulting in 12 rigid domains (Figure 1).

Molecular Dynamics Simulations. Molecular dynamics (MD) simulations were performed with the AMBER ff03 force field[17] using the open-source *GROMACS 4.0* software package, which was adapted to account for exchanges between different Hamiltonian states.[18] Long-range electrostatics were computed with the particle mesh Ewald method.[19] The nonbonded cutoff was set to 1.2 nm. A Fourier spacing of 0.19 nm was used. All covalent bonds were constrained using P-LINCS.[20] The integration time step was 2 fs. The neighbor list was updated every 10 steps. The system temperature was controlled by the velocity rescaling thermostat, which preserves the canonical ensemble.[21] Exchanges between replicas were attempted every 1 ps.

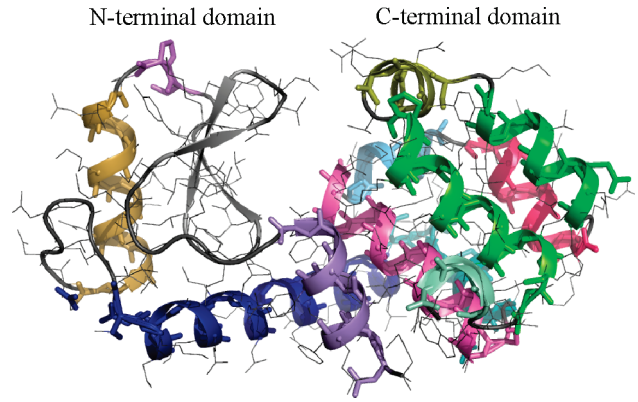

Figure 1. The T4L rigid domains as determined by FIRST. Each colored region represents a different rigid domain. The gray-colored side chains, β-sheet and loops are flexible. The dark blue α-helix connects the N-terminal and C-terminal domains.

The initial T4L structure was obtained from the X-ray structure (2LZM). Protonation states were determined using the H++ web server.[22] The charged protein was neutralized with Cl^- ions and solvated in a dodecahedral box using 9093 TIP3P water molecules. The neutral solvated system was relaxed, heated to 300 K using position restraints and briefly equilibrated for 50 ps. Next, one REST (s1) and one REFT (f1) simulation were started using identical starting structures for each replica. After 2 ns of conventional MD, again one REST (s2) and one REFT (f2) simulation were started. The REM simulations were carried out for 120 ns, the last 60 ns of which were used for analysis. The f1 simulation was continued until 360 ns for further analysis of the equilibrated ensemble. Replica temperatures were 300.0, 308.4, 317.4, 326.8, 336.6, 346.8, 357.3, 368.2, 379.5, 391.2, 403.3, 415.8, 428.7, 442.1, 455.9, 470.2, 484.9, 500.1, 515.9, 532.2, 549.0, 566.3, 584.3, 602.8 K for REFT, and 300.0, 307.8, 315.9, 324.1, 332.6, 341.2, 350.2, 359.3, 368.7, 378.3, 388.2, 398.3, 408.7, 419.4, 430.3, 441.5, 453.1, 464.9, 477.0, 489.5, 502.2, 515.4, 528.8, 542.6, 556.8, 571.3, 586.2, 601.5 K for REST. The replica temperatures were adapted from an exponentially distributed temperature set, based on preliminary 2 ns replica exchange simulations. Special care was taken to obtain pairwise exchange probabilities between all neighboring replicas close to 20%.

The cleft volume was calculated by generating a spherical grid of oxygen atoms with a radius of 1.1 nm around the center of the C^α atoms of residues Glu11, Gly30, and Phe104. These three residues are located at the active site core inside the cleft. The grid spacing was 0.1 nm. Grid atoms overlapping with protein atoms were removed and the

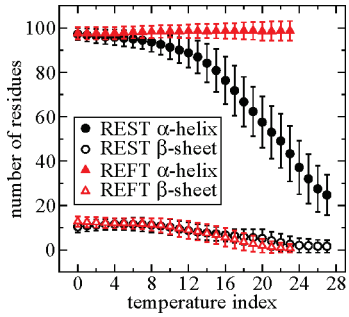

Figure 2. Number of residues that are part of an α-helix or β-sheet as a function of temperature index, averaged over the last 60 ns of the 120 ns simulation. The temperature indices denote the serial numbers of the consecutive thermodynamic states, starting with zero for the 300 K state. Error bars represent the standard deviations of the averages.

number of remaining grid atoms was counted. The principal component analysis (PCA) was carried out on the backbone atoms of a series of 38 X-ray structures as described by de Groot et al.[23] Snapshots of the MD simulations were projected onto the calculated eigenvectors.

Results

T4L is a 164-residue protein which is known to display large conformational changes upon substrate binding through hinge bending motion.[24,25] The protein consists of two domains connected by a central α-helix. The active site is located in the cleft between the N-terminal and C-terminal domains. The hinge bending motion mediates a closing and opening of the active-site cleft. Starting from the closed state, the equilibration rates of the open and closed states and their relative populations were used to compare the sampling efficiency of the REST and REFT methods.

Four explicit solvent 120 ns simulations were carried out: two independent REST (s1, s2) and two independent REFT (f1, f2) simulations. To span a 300−600 K range with 20% acceptance probability between successive replicas, using REST and REFT 28 and 24 replicas were needed respectively. For comparison, with T-REM we estimate that about 124 replicas would have been needed to span the same exchange rates and temperature range.

We used the FIRST method to identify the rigid domains.[16] The effect of increased rigid domain interactions on the secondary structure is readily seen in Figure 2. In the REST simulation, at high temperatures both the α-helix and β-sheet content decays. Using REFT, α-helix content stays intact even at 600 K. The β-sheet content however decreases with temperature. The β-strands were not identified by FIRST as part of a rigid domain, reflecting the more flexible nature of β-sheets compared to helices.

We define the crossing rate as the number of replicas that cross temperature space between 300 K and a given temperature >300 K. High crossing rates are essential for fast equilibration at low temperature. Compared to the REST simulation, on account of the preservation of secondary structure (Figure 2), the REFT crossing rates have drastically increased (part A of Figure 3). The crossing rate differences become more pronounced as the simulation progresses.

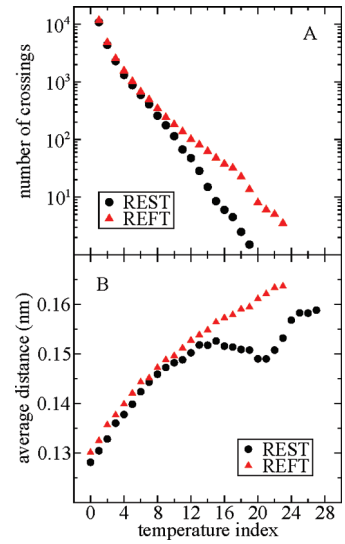

Figure 3. Mobility in temperature and conformational space. (A) Number of temperature crossings between 300 K and the temperature index given in the abscissa. (B) Average distance covered per ps along the first eigenvector of the PCA analysis. Averages were taken over the last 60 ns of the two REST and REFT simulations, respectively.

Above 500 K no single crossing was observed for any of the 28 REST replicas during the last 60 ns. Thus, an effective decoupling occurs between the high and low temperature replicas.

Compared to REST, a larger part of the REFT system is kept cold, which might affect the dynamics in the high-temperature simulations. It is therefore important to evaluate the mobility of the protein in conformational space. The average distance covered every ps along the first PCA eigenvector (Materials and Methods) as a function of temperature is plotted in part B of Figure 3. This collective coordinate corresponds to the hinge bending opening-closing motion. At low temperatures, the mobility along the hinge bending coordinate is comparable. As the temperature increases, part of the kinetic energy is channeled toward the slow hinge bending mode. Above 420 K however, the REST hinge bending mobility starts to degrade. This is easily explained by the loss of native structure, which leads to a distribution of kinetic energy into many local non-native high-frequency motions. By contrast, in the REFT simulation, a steady increase of the hinge bending mobility is seen along the full temperature range.

To better understand how the potential function affects conformational sampling at high temperature, we performed two MD simulations of the REST and REFT states at temperature index 17, but without replica exchanges (Figure 4). The thermodynamic state at temperature index 17 induces strongly elevated protein dynamics, while the crossing rates are still considerable for both methods (parts A and B of Figure 3). Initially good open−closed transition rates are seen for REST-state MD. However, after 4.3 ns the protein starts to unfold, as can be seen from the backbone root-mean-square deviation (bRMSD) plot (part C of Figure 4). As a reference, at 300 K the bRMSD from the X-ray structure fluctuates between 0.07 and 0.47 nm. The unfolding coincides with a sharp decline of the open−closed transitions to

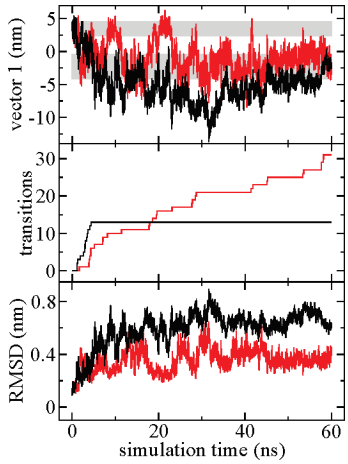

Figure 4. Time evolution of the first PCA eigenvector (A), the cumulative number of transitions between the open and closed states (B), and the bRMSD from the X-ray structure (C) of two high-temperature MD simulations using a potential function corresponding to temperature index 17, which is 464.9 K for REST (black) and 500.1 K for REFT (red). The upper and lower gray bars in (A) indicate the closed and open states, respectively.

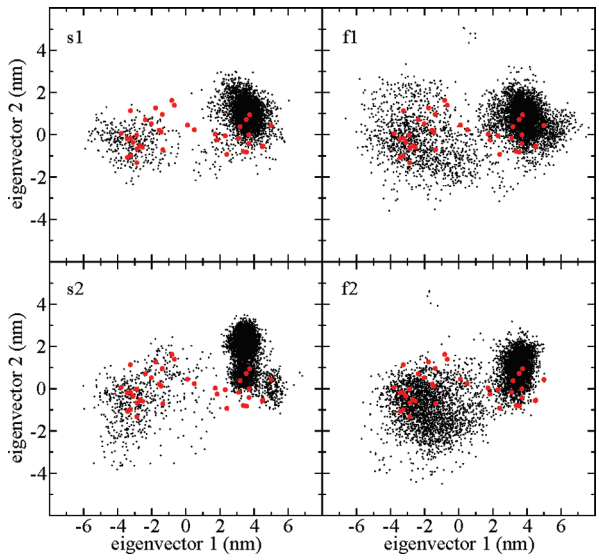

Figure 5. 2D projection of the 300 K trajectories onto the first two eigenvalues of a PCA on the backbone of a series of X-ray structures. REST (s1, s2), REFT (f1, f2). Black dots indicate samples taken every 10 ps. Red circles indicate 38 X-ray reference structures.

zero (part B of Figure 4). For a large part of the simulation time, the system samples a region in conformational space that is non-native (PCA eigenvector 1 < −5, part A of Figure 4). In contrast, the REFT-state MD stays mainly within the native bRMSD region, and the number of open−closed transitions increases steadily throughout the trajectory. For comparison, a 184 ns conventional MD simulation started from the closed state did not yield a single transition to the open state.[26]

In Figure 5, the sampling in the space of the two largest PCA eigenvectors, as well as a series of 38 X-ray structures[23] is shown. The second eigenvector describes a twisting mode. Using REFT, the area of open conformations (eigenvector 1 < 1.25 nm) is populated more densely. At this stage the systems are not yet equilibrated, which explains the differences between the four simulations in Figure 5. All X-ray conformations are sampled by both methods.

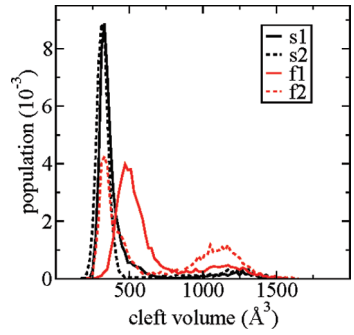

Figure 6. Cleft volume distribution of the two independent REST (s1, s2) and REFT (f1, f2) simulations, from the last 60 ns of the 120 ns simulations at 300 K.

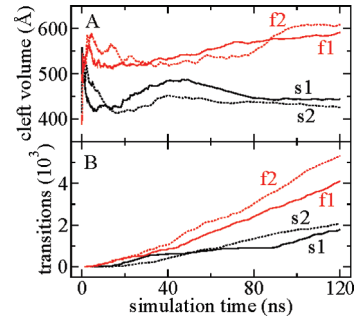

Figure 7. Time evolution of the cumulative average cleft volume (A) and the total number of open−closed transitions (B) in the 300 K state. Starting cleft volumes were 489 Å^3 (s1, f1) and 406 Å^3 (s2, f2).

Whereas the principal eigenvectors of a PCA analysis represent the collective coordinate with the largest spatial displacement, the cleft volume (Materials and Methods) provides information about the local dynamics at the active site. Again, the REFT simulations display a higher population for the open state (>750 Å^3) compared to the REST simulation (Figure 6). Moreover, the closed state distribution is much more dispersed, indicating better local sampling of the active site cleft.

Part A of Figure 7 shows the time evolution of the average cleft volume. Shortly after an initial increase, the REST simulations' cleft volumes decline sharply. A possible explanation for this decline might be that open structures unfold faster than closed structures, and hence are trapped inside the high-temperature states more easily. A similar but much smaller cleft volume decline is also visible in the REFT simulations. After about 10−30 ns, the REFT cleft volumes increase steadily, whereas the REST cleft volumes remain at a much lower level.

The rate of open−closed transitions in the 300 K state, which reflects exchanges between replicas with open and closed conformations, provides an additional measure of the sampling efficiency. As shown in part B of Figure 7, the transition rate is relatively constant for all four simulations, indicating good local replica exchange frequencies. During the last 60 ns of the simulation, the average REFT transition

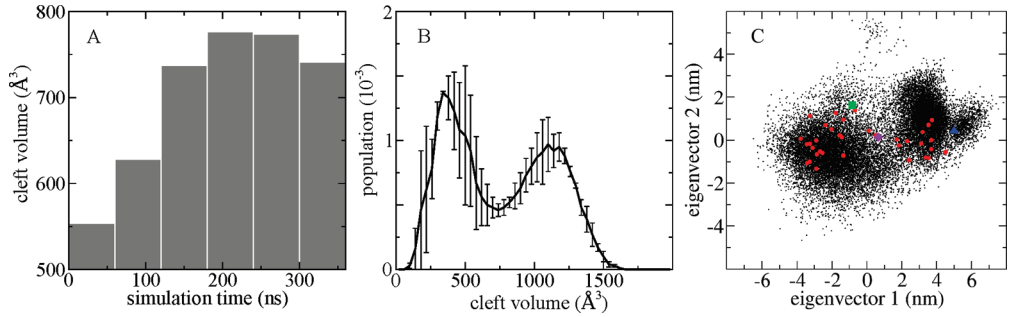

Figure 8. Native state equilibration of T4L in the 300 K state. (A) Evolution of the cleft volume shown as block averages over 60 ns. (B) Cleft volume distribution taken over the last 240 ns of a 360 ns simulation. Error bars represent the standard deviations from the average cleft populations of two 120 ns blocks. The open state has a population of 49 ± 5%. (C) 2D projection of a PCA analysis of the last 240 ns of a 360 ns REFT simulation. The purple diamond indicates the average position in 2D space. The green square and blue triangle represent 150L(c) and 152L, respectively.

rate is 47 transitions/ns, 2.5 times the REST transition rate of 18 transitions/ns. Considering acceptance probabilities ∼20% with one replica exchange attempt every picosecond, the fraction of REFT and REST replica exchanges accompanied by an open−closed transition are ∼23% and ∼9%, respectively.

To estimate the simulation time necessary to obtain a fully equilibrated ensemble, and to get an estimate of the equilibrium distribution, the f1 REFT simulation was extended to 360 ns. Judged by the evolution of the cleft volume (part A of Figure 8), about 120 ns of REFT simulation time seems necessary to reach an equilibrated state. Note that the same starting conformation was used for each replica within a given simulation to avoid any bias in the comparison of both methods. If a diverse set of starting conformations was used across all replicas, a much shorter equilibration time would likely be obtained.

As shown in part B of Figure 8, T4L forms a distinct two-state open−closed equilibrium in solution. Both states occupy approximately half of the population. This two-state model is confirmed by the PCA (part C of Figure 8). In addition, a small third superclosed state is visible in part C of Figure 8, which has a population of ∼5% (eigenvector 1 > 4.9 nm). Here the N- and C-terminal domains are tightly held together by two side chain H-bonds (Asp20 - Gln141) and (Glu22 - Arg137). The latter H-bond is much less prominent, and the former H-bond is virtually nonexistent in the normal closed state.

Discussion

The efficiency of replica exchange methods depends critically upon (i) the mobility in temperature space and (ii) the mobility in conformational space in the high-temperature range. As was observed by the Berne group, the REST method did not perform well for larger systems.[27] Despite the high acceptance ratios, the rate at which replicas traveled temperature space between the highest and lowest temperature state remained low, effectively decoupling the high and low temperature replicas. Replicas at high temperatures were mainly unfolded, replicas at low temperatures were mostly in the native state, but native-unfolded state transitions within a single replica occurred rarely. Better results were obtained by excluding a randomly chosen subset of water molecules from the solvent interaction scaling,[27] partially negating the benefits of REST compared to the original T-REM.

With REST, the high-temperature replicas are mostly unfolded and thus sample a region in conformational space that does not significantly contribute to the equilibration of the native state. Because of the high entropy of the unfolded state, the probability of refolding to the native state is very low. As shown in Figure 4, partial unfolding also hinders open−closed interconversion rate of T4L. At room temperature, however, the native state is much more stable than the unfolded state. Thus, in REST the only meaningful way for an unfolded protein at high temperature to contribute to the native state at room temperature is by refolding into the native state while traveling down the temperature ladder, which is a very slow process.

The REFT method is based on (i) the notion that only about 0.00001% to 0.01% of polypeptide chains are unfolded at physiological temperatures,[28] and (ii) the assumption that the equilibration rate of proteins is usually limited by slow interdomain dynamics. The REFT potential function tends to keep the rigid domains intact over the full temperature range, while stimulating domain−domain movements at higher temperatures. As such, more simulation time is spent in the native state, which, as is demonstrated in this article, allows the native state to reach an equilibrated ensemble much faster.

It is important to note that the REFT method in its current form cannot be applied a priori to an amino acid sequence of unknown fold. The method is however not meant to study protein folding, but to achieve faster equilibration of the protein native state. Note that protein unfolding is not prevented at any temperature. At higher temperatures however, the fraction of time spent in the unfolded state will generally be small, just as is the case at low temperature with conventional MD.

The choice of the rigid domains will affect the efficiency of the REFT method. If all atoms are considered rigid, the system is equivalent to multiple conventional MD simulations at 300 K. The REST method corresponds to a REFT system with only flexible protein atoms. In this work, we have used the FIRST algorithm to determine the rigid and flexible domains, which derives geometrical constraints from a static protein structure to identify the rigid domains. In the case

of T4L, while large interdomain motions occur, the domains themselves remain intact. This was verified by FIRST calculations on both open and closed X-ray structures, which gave very similar results. For more flexible proteins, a FIRST analysis of multiple protein structures, experimental or from simulation, could be used to distinguish the truly rigid domains from the more flexible ones, as was proposed in the original FIRST article.[16] Other methods could be used to obtain rigidity information, including simple secondary structure identification algorithms, tCONCOORD,[26] which is similarly based on distance constraints, normal mode analysis, and elastic network models.[29]

Independent evidence for the equilibrated state is provided by a study of Goto et al. who obtained approximately equal contributions from open and closed conformers by fitting experimental dipolar couplings to a two-state open−closed model.[30] The superclosed state in part C of Figure 8 is similar to the solution conformation of the T21C/T142C mutant, which contains a disulfide bond between the N- and C-terminal domains.[30] The cleft of this mutant is even more closed than the most closed X-ray T4L conformation (152L, blue triangle in part C of Figure 8), a mutant with three disulfide bonds. Earlier tCONCOORD and multiple conventional MD simulations started at different conformational states did not reveal this compacted state.[26]

The equilibrated average position on the 2D PCA plot (violet diamond, part C of Figure 8) agrees reasonably well with the position of the M6I mutant X-ray structure (150L(c), green square, part C of Figure 8), which resembles the average solution conformation of the T4L homologue C54T/C97A calculated from $^{1}HN-^{15}N$ dipolar couplings.[30] Possible explanations for the difference include: unsatisfactory equilibration, inaccurate force field parameters (the AMBER ff03 force field has been shown to favor α-helices, which might lead to an overly rigid protein,[31] differences between the 150L(c) X-ray structure and the solution structure which were not resolved by the NMR data, and the presence of the two mutations (C54T/C97A) in the NMR structure.

Further improvements of the method are possible. To optimize the balance between rigidity and flexibility, one could refine the scaled potential in eq 4 by explicitly enumerating the atom pairs whose interactions need to be scaled. More accurate rigidity information could be extracted from pregenerated conformational ensembles. Alternatively, using a fast algorithm such as FIRST, rigid domains can be calculated dynamically during the simulation, which would facilitate the formation and destruction of transient domains. Moreover, kinetic information could be extracted from the REM simulations[32] as a way to validate the assumption that the large-scale motion between rigid domains is the slow rate-limiting step.

Acknowledgment. We thank B. de Groot for kindly providing the file with reference X-ray structures. S.L.C.M. is a research mandate holder of the Flemish agency for Innovation by Science and Technology (IWT). S.M. is the recipient of a doctoral grant from the Flemish Science Foundation (FWO). This research was conducted utilizing high-performance computational resources provided by the University of Leuven (http://ludit.kuleuven.be/hpc).

References

(1) Swendsen, R. H.; Wang, J. S. Replica Monte-Carlo simulation of spin-glasses. *Phys. Rev. Lett.* **1986**, *57* (21), 2607–2609.

(2) Hukushima, K.; Nemoto, K. Exchange Monte Carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.* **1996**, *65* (6), 1604–1608.

(3) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314* (1−2), 141–151.

(4) Frenkel, D.; Smit, B. *Understanding Molecular Simulation. From Algorithms to Applications*; Academic Press: London, U.K., 2002.

(5) Zhou, R. Trp-cage: folding free energy landscape in explicit water. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (23), 13280–13285.

(6) Garcia, A. E.; Onuchic, J. N. Folding a protein in a computer: An atomic description of the folding/unfolding of protein A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (24), 13898–13903.

(7) Moors, S. L. C.; Jonckheer, A.; De Maeyer, M.; Engelborghs, Y.; Ceulemans, A. Tryptophan conformations associated with partial unfolding in ribonuclease T1. *Biophys. J.* **2009**, *97* (6), 1778–1786.

(8) 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* (20), 9058–9067.

(9) Earl, D. J.; Deem, M. W. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.* **2005**, *7* (23), 3910–3916.

(10) Moors, S. L. C.; Michielssens, S.; Flors, C.; Dedecker, P.; Hofkens, J.; Ceulemans, A. How is cis-trans isomerization controlled in Dronpa mutants? A replica exchange molecular dynamics study. *J. Chem. Theory Comput.* **2008**, *4* (6), 1012–1020.

(11) Liu, P.; Huang, X. H.; Zhou, R. H.; Berne, B. J. Hydrophobic aided replica exchange: An efficient algorithm for protein folding in explicit solvent. *J. Phys. Chem. B* **2006**, *110* (38), 19018–19022.

(12) Hritz, J.; Oostenbrink, C. Hamiltonian replica exchange molecular dynamics using soft-core interactions. *J. Chem. Phys.* **2008**, *128* (14), 144121.

(13) Kannan, S.; Zacharias, M. Enhanced sampling of peptide and protein conformations using replica exchange simulations with a peptide backbone biasing-potential. *Proteins* **2007**, *66* (3), 697–706.

(14) Zacharias, M. Combining elastic network analysis and molecular dynamics simulations by hamiltonian replica exchange. *J. Chem. Theory Comput.* **2008**, *4* (3), 477–487.

(15) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (39), 13749–13754.

(16) Jacobs, D. J.; Rader, A. J.; Kuhn, L. A.; Thorpe, M. F. Protein flexibility predictions using graph theory. *Proteins* **2001**, *44* (2), 150–165.

(17) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; 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* (16), 1999–2012.

(18) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447.

(19) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.

(20) Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **2008**, *4* (1), 116–122.

(21) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.

(22) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: a server for estimating pK(a)s and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **2005**, *33* (Web Server issue), W368–W371.

(23) de Groot, B. L.; Hayward, S.; van Aalten, D. M.; Amadei, A.; Berendsen, H. J. Domain motions in bacteriophage T4 lysozyme: A comparison between molecular dynamics and crystallographic data. *Proteins* **1998**, *31* (2), 116–127.

(24) Zhang, X. J.; Wozniak, J. A.; Matthews, B. W. Protein flexibility and adaptability seen in 25 crystalforms of T4 lysozyme. *J. Mol. Biol.* **1995**, *250* (4), 527–552.

(25) Mchaourab, H. S.; Oh, K. J.; Fang, C. J.; Hubbell, W. L. Conformation of T4 lysozyme in solution. Hinge-bending motion and the substrate-induced conformational transition studied by site-directed spin labeling. *Biochemistry* **1997**, *36* (2), 307–316.

(26) Seeliger, D.; Haas, J.; de Groot, B. L. Geometry-based sampling of conformational transitions in proteins. *Structure* **2007**, *15* (11), 1482–1492.

(27) Huang, X. H.; Hagen, M.; Kim, B.; Friesner, R. A.; Zhou, R. H.; Berne, B. J. Replica exchange with solute tempering: Efficiency in large scale systems. *J. Phys. Chem. B* **2007**, *111* (19), 5405–5410.

(28) Murphy, R. M.; Tsai, A. M. *Misbehaving Proteins. Protein (Mis)folding, Aggregation and Stability*; Springer Science+ Business Media: New York, N.Y., 2006.

(29) Cui, Q.; Bahar, I. *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems*; Chapman & Hall/CRC: Boca Raton, FL, 2006.

(30) Goto, N. K.; Skrynnikov, N. R.; Dahlquist, F. W.; Kay, L. E. What is the average conformation of bacteriophage T4 lysozyme in solution? A domain orientation study using dipolar couplings measured by solution NMR. *J. Mol. Biol.* **2001**, *308* (4), 745–764.

(31) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65* (3), 712–725.

(32) Yang, S.; Onuchic, J. N.; Garcia, A. E.; Levine, H. Folding time predictions from all-atom replica exchange simulations. *J. Mol. Biol.* **2007**, *372* (3), 756–763.

CT100493V