Published in final edited form as:

J Comput Chem. 2017 November 15; 38(30): 2632–2640. doi:10.1002/jcc.24923.

Efficacy of Independence Sampling in Replica Exchange Simulations of Ordered and Disordered Proteins

Kuo Hao Lee¹ and Jianhan Chen^{1,2,*}

¹Department of Chemistry, University of Massachusetts, Amherst, MA 01003, USA

²Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, MA 01003, USA

Abstract

Recasting temperature replica exchange (T-RE) as a special case of Gibbs sampling has led to a simple and efficient scheme for enhanced mixing (Chodera and Shirts, 2011). To critically examine if T-RE with independence sampling (T-REis) improves conformational sampling, we performed T-RE and T-REis simulations of ordered and disordered proteins using coarse-grained (CG) and atomistic models. The results demonstrate that T-REis effectively increase the replica mobility in temperatures space with minimal computational overhead, especially for folded proteins. However, enhanced mixing does not translate well into improved conformational sampling. The convergences of thermodynamic properties interested are similar, with slight improvements for T-REis of ordered systems. The study re-affirms the efficiency of T-RE does not appear to be limited by temperature diffusion, but by the inherent rates of spontaneous large-scale conformational re-arrangements. Due to its simplicity and efficacy of enhanced mixing, T-REis is expected to be more effective when incorporated with various Hamiltonian-RE protocols.

TOC Graphic

A simple extension of replica exchange (RE) involves multiple exchanges attempted at each exchange cycle with the same Metropolis criteria to enhance mixing. We performed atomistic and coarse-grained simulations of ordered and disordered proteins to critically examine if such temperature RE with independence sampling (T-REis) could improve conformational sampling. The results demonstrate that T-REis effectively increase the replica mobility in temperatures space; yet such enhanced mixing does not translate well into improved conformational sampling. The efficiency of T-RE does not appear to be limited by temperature diffusion, but by the inherent rates of spontaneous large-scale conformational re-arrangements.

Corresponding Author: Phone: (413) 545-3386; jianhanc@umass.edu.

Keywords

Gibbs sampling; enhanced sampling; independence sampling; replica exchange; conformational ensemble; disordered proteins

Introduction

Conformational sampling of large biomolecules and other complex molecular systems is one of the key challenges in molecular dynamics (MD) simulations.^{1–4} Even though conventional explicit solvent simulations have achieved microsecond to millisecond time scales in recent years,^{5–7} enhanced sampling techniques that aim to observe rare dynamic events more frequently remain highly desirable in studies of many biological problems.^{8–15} In particular, temperature-based replica exchange (T-RE) has been widely adopted as a powerful enhanced sampling technique for general biomolecular simulations.¹⁴ The central concept of T-RE is to accelerate energy barriers crossing by random walk in the temperature space. Multiple copies of the system (replicas) are simulated simultaneously and independently at different temperatures using either Monte Carlo (MC) or MD. Periodically, replicas attempt to exchange simulation temperatures according to Metropolis criteria that preserve the detailed balance and ensure canonical distributions at all temperatures.¹⁶ Replicas could sample a broader of conformations as they undergo self-regulated random walk in the temperature space.

In practice, the efficiency of T-RE in sampling biomolecular conformations depends on many parameters including the number of replicas, temperature distributions, and the specific exchange algorithm.^{17–24} The underlying energy barriers of the simulated system have a significant influence on the sampling efficiency. T-RE can be more efficient than conventional MD simulation as long as the activation enthalpies of conformational transitions are positive.^{19,22} On the other hand, protein folding often displays anti-Arrhenius

behavior, where the rate of folding slows down above a certain temperature due to entropic bottlenecks.^{25,26} It has been shown that the number of folding transitions observed at the low temperature of interest depends on the maximum of the harmonic mean of the folding and unfolding transition rates at the high temperature.²⁵ It is therefore important to choose a maximum temperature right above the temperature at which the activation enthalpy of folding vanishes. The number of replicas also influences the efficiency, as sufficient energy overlap is required to achieve optimal exchange acceptance rates.^{21,24} Many approaches have been proposed to reduce the number of replicas required for effective temperature diffusion, such as by reducing the number of particles that participate in temperature exchanges.^{10,11,21,24} Exchange can be attempted as frequent as possible to promote replica diffusion in the temperature space (i.e., mixing) while allowing the system to be sufficient equilibrated.^{27–29} Several advanced RE schemes have also been proposed to directly enhance replica random walk in the temperature space. For example, infinite swapping (INS) considers all possible attempts among replicas,³⁰ whereas partial infinite swapping (PINS) partitions the temperature space into locally symmetrized temperature blocks at each step to reduce the computational overhead.³¹ It has been shown that INS and PINS could provide better mixing and sampling in free energy calculation, but they are computationally more expensive than regular neighbor exchange method.³² Similarly, the Suwa-Todo (ST) algorithm promotes transitions in the replica state space by introducing a weight allocation procedure and minimizing the rejection rate.³³ Metadynamics with bias-exchange replica exchange simulations implemented with INS and ST algorithms show comparable sampling performance for alanine polypeptides,³⁴ and both methods allow enhanced conformational sampling of N-glycans to cross high-energy barrier along ω .³⁵ A replica exchange variant with ST algorithm also demonstrates higher sampling efficiency in explicit solvent simulation of C-peptide.³³

Recently, Chodera and Shirts recasted RE and expended ensemble simulations as special cases of Gibbs sampling and described several simple approaches for enhancing replicas mixing.³⁶ In particular, they suggested that multiple exchanges could be attempted at each exchange cycle with the same Metropolis criteria to enhance mixing. Traditional nearest neighbor T-RE protocol can be considered as a special case of T-REis with only one set of swapping attempts per exchange cycle. The efficiency of such independence sampling (IS) in replica exchange (T-REis) for enhancing mixing has been illustrated using simple analytical models, and simulation of LJ spheres and alanine dipeptide.³⁶ Various RE schemes based on such multiple exchanges per exchange cycle have been developed.^{37–39} For example, Async REMD was designed for large-scale RE simulations on heterogeneous clusters, and can achieve higher efficiency by optimizing lengths of MD simulations and the number of exchanges attempted per cycle.³⁸ Zhang et al. constructed a Markov state model of RE from long MD simulations to investigate how the number of replica exchange attempts and the number of MD steps per cycle affect sampling efficiency.³⁷ However, how well the observed efficiency in model systems extends to large biomolecular systems with complex conformational space needs to be carefully examined. Specifically, it is yet to be determined how well enhanced mixing achieved with T-REis translates into enhanced conformational sampling and improved convergence in thermodynamic properties of interest. In this work, we carefully assess T-REis with different numbers of exchange

attempts per replica exchange cycle using a set of ordered and disordered proteins at both coarse-grained (CG) and atomistic resolutions. We first evaluate T-REis' ability to enhance replica mixing compared to traditional nearest neighbor T-RE and then rigorously determine its sampling efficiency by directly comparing the rates of local and global conformational transitions and convergence of various thermodynamic properties of interest.

Method

CHARMM⁴⁰ and the Multiscale Modeling Tools for Structural Biology (MMTSB) toolset⁴¹ were used to perform all simulations and facilitate the analysis. The MMTSB toolset has been slightly modified to include multiple independence sampling attempts between MD cycles. Secondary structures of the atomistic structures were analyzed using the CHARMM COOR SEC module. R was used for most statistical analysis.⁴²

Coarse-grained Simulations of protein G B1 domain

Protein G B1 domain consists of 56 residues and contains both α and β secondary structures in its native structure (Figure 1A). It is an ideal model system for benchmarking T-REis' ability to accelerate cooperative folding transitions. We have previously constructed a realistic yet computational treatable CG model of protein G B1,¹⁷ by adding significant nonspecific hydrophobic interactions to the sequence-flavored G⁻-like model.⁴³ The underlying concept of sequence-flavored G⁻-like models is that, while the folded topology governs the essential features of folding, sequence provides different statistical potentials to the populations of structural elements during folding to modulate their prevalence as observable intermediates and affect folding kinetics. Each residue is represented by a single C α bead and the residue-specific native interactions are described by the Miyazawa-Jernigan (MJ) statistical potentials.⁴⁴ The initial G⁻-like model of protein G B1 domain was first generated by MMTSB G⁻-Model Builder (<http://www.mmtsb.org>) based on PDB ID: 3GB1.⁴⁵ Nonspecific van der Waals (vdW) interactions were then introduced between all C α beads with strengths similar to those of the native contacts with a well depth of $\epsilon = -1.1$ kcal/mol. Furthermore, the nonspecific 1–4 vdW interactions as well as those in two short loops (see Figure 1A) were turned off to avoid steric restriction of rotation around C α –C α bonds and facilitate chain diffusion. The model displays more complex energy landscape with multiple intermediate states, but retains a sharp folding transition (see Figure 2). Compared to the original sequence-flavored G⁻-like model, the new model is about 15 times slower in reversible folding and unfolding transitions at the corresponding melting temperatures. All CG simulations of protein G B1 involved 8 replicas distributed exponentially between 300 and 400 K. The MD time step was 10 fs and exchange attempt was made every 5000 MD steps (i.e., every 50 ps). Three IS swapping attempts, 10, 100, or 1000 per exchange cycle were used in T-REis simulations to better evaluate their effects. For each setup, two independent control (from the folded structure) and folding (from an extended structure) runs were performed for convergence diagnosis. The reference T-RE control simulation was taken from the previous study,¹⁷ and a new folding T-RE simulation was performed in this study.

Atomistic simulations of GB1p, (AAQAA)₃ and KID

GB1p (GEWTY DDA TK TFTVT E) is a β -hairpin derived from protein G B1 domain, and has been estimated to be ~42% folded at 278K based on nuclear magnetic resonance (NMR) chemical shift analysis⁴⁶ (Figure 1B). (AAQAA)₃ is a weakly helical peptide that has been characterized in detail by NMR⁴⁷ (Figure 1C). Both GB1p and (AAQAA)₃ have been extensively utilized in force field optimization and benchmark studies.^{48,49} We note that GB1p resembles larger proteins in many aspects including cooperative folding transitions and microsecond folding time scales. Both control and folding T-REis simulations were performed in the GBSW implicit solvent model⁵⁰ using 8 replicas distributed exponentially between 270 and 450 K. Langevin simulations were performed with a friction coefficient of 0.1 ps⁻¹. SHAKE⁵¹ was used to constrain the lengths of all bonds involving hydrogen atoms, and the dynamic time step was 2 fs. Exchanges attempted every 2 ps with three IS swapping attempts 10, 100, or 1000 per exchange cycle. The same simulation parameters were used in T-REis simulations of (AAQAA)₃. Except that 16 replicas exponentially distributed between 300 and 500 K were used. The reference T-RE simulations of GB1p and (AAQAA)₃ were taken from the previous studies.^{52,53}

The kinase inducible domain (KID) of the transcription factor CREB (residues 119–146: TD SQKRR EILSR RPSYR KILND LSSDA P) is one of the most studied intrinsically disordered proteins (IDPs) (Figure 1D).^{54–56} KID lacks stable tertiary structures in isolation but folds into two helices upon binding to the KIX domain of the CREB binding protein.⁵⁵ NMR chemical shift analysis has revealed that helix α A (residues 120–129) is ~50–60% folded, while helix α B (residues 134–144) is only ~10–15% folded in the unbound state.⁵⁶ Both control and folding T-REis simulations of KID involved 12 replicas exponentially distributed between 270 and 500 K. Otherwise, the same setup was used as described above for GB1p and (AAQAA)₃ simulations. The reference T-RE simulations were taken from the previous study.⁵⁴

Results and Discussion

Effectiveness of T-REis in Sampling CG Cooperative Folding Transitions

We first examined the effectiveness of T-REis using CG simulations of protein G B1 domain. Key statistics of CG T-RE and T-REis simulations are summarized in Table 1. T-REis relies on many cycles of IS swapping attempts to generate uncorrelated and independent samples and enhance mixing.³⁶ Enhanced mixing can be directly measured by N_{TR} , the total number of round trips between the temperature extremes sampled by all replicas. Clearly more IS swapping attempts lead to greater N_{TR} . For example, T-REis can generate about 2-fold N_{TR} compared to T-RE with 10 IS swapping attempts, and >5-fold N_{TR} with 100 swapping attempts. Greater N_{TR} generated in T-REis can be partially attributed to higher replica mobility per change cycle. As shown in Figure 3, with more IS swapping attempts, replicas are more likely to exchange to distal conditions (Figure 3A), and thus have smaller probabilities of remaining at the original conditions (Figure 3B). However, there always exist significant probabilities of replicas remaining at the same condition, and the likelihood of generating moves with condition number difference greater than three is small. Interestingly, replica mobility per exchange cycle quickly reaches its

upper limit with ~100 IS swapping attempts per cycle (Figure 3B). Yet N_{TR} is further increased, to over 10-fold, with higher IS swapping attempts, suggesting that T-REis is highly effective in enhancing mixing.

How well the drastically improved mixing translates into enhanced sampling efficiency is first assessed by examining the number of reversible folding/unfolding transitions sampled by all replicas during the entire simulation (N_{TS}). For this, the folded and unfolded states were identified as Q_c with >0.73 and <0.33 (see Figure 2B), respectively. Disappointingly, the results show that greater N_{TR} in T-REis does not lead to larger N_{TS} . Instead, T-REis sample similar, if not often slightly smaller, numbers of reversible folding transitions. We note that the similar observations have been made previously using adaptive temperature-based RE protocols,¹⁷ where enhanced mixing using either exchange equalization or current maximization was found to have minimal effects in accelerating large-scale conformational transitions. The implication is that the ability to accelerate the large-scale conformational transitions is not limited by temperature diffusion in RE simulations. The fundamental reason is that the free energy barriers of cooperative transitions are dominated by entropic components.¹⁷ Transition states of slow- and fast-folding proteins have been associated with low and high conformational entropies.⁵⁷ We note that similar N_{TS} observed for T-RE and T-REis protocols may also simply that random walk in the temperature space and conformational diffusion are largely decoupled in these simulations, despite the diffusive and multi-step nature of protein conformational transitions. As such, N_{TS} of a replica exchange simulation is largely determined by the temperature distribution itself and not affected by the exchange protocols¹⁸.

Given the caveat of N_{TS} to gauge sampling efficiency, we directly examine a set of thermodynamic properties to further assess the potential benefits of enhanced mixing of T-REis. The results suggest that T-REis does not improve the convergence of these properties either. For example, the C_v curves from control and folding T-RE and various T-REis simulations all demonstrate similar limited convergence (Figure 2A). The free energy as a function of Q_c derived from T-RE and T-REis are also similarly converged. We note that the calculated PMFs and T_m curves show substantial variance between T-RE and T-REis with different IS swapping attempts, suggesting a lack of total convergence. This is apparently due to the difficulty of sampling on rugged energy surface of the modified G-like model (see method). The current simulations of 100 μ s per replica only manage sample ~20 reversible transitions. Examination of the autocorrelation function of Q_c does not reveal any evidence that enhanced mixing allows faster equilibration of the structure ensemble sampled at the lowest temperature either (see Figure S1). To confirm the ability of T-REis in recovering the correct thermodynamic properties, we also performed additional T-REis and T-RE simulations using the sequence-flavored G-like model of GB1p. The results demonstrate that T-REis does yield essentially identical ensembles as T-RE (see Text S1 and Figure S2).

Efficiency of T-REis in Sampling Atomistic Folding Transitions

CG simulations of protein G B1 domain reveal enhanced mixing of T-REis failed to translate into more conformational transitions or better convergence in various thermodynamic

properties. Arguably, the G-like models have entropic folding barriers and display anti-Arrhenius behavior in the whole temperature range. This is not fully representative of proteins in more physical and more detailed force fields. Therefore, we have further assessed the efficiency of T-REs for sampling atomistic folding transitions using GB1p, a small β -hairpin peptide with protein-like cooperative folding transitions. The results are summarized in Table 2. Similar observations can be made, that larger IS swapping attempts lead to better mixing but there is no any evidence of enhanced sampling of reversible folding transitions. Compared to CG simulations, replicas in these atomistic simulations appear to have high tendencies to remain at the original conditions with more modest probabilities of exchanging to new conditions 2 steps away (Figure S3 vs. Figure 3A). This is likely a direct consequence of larger energy fluctuations in atomistic simulations. Surprisingly, T-RE generates nearly 4-fold more temperature round trips in the folding verses control run. Most replicas remained in the unfolded states in folding simulation and only three replicas occupied in folded state in the control runs. In contrast, most replicas in the control T-RE simulation also consistently sampled the unfolded state. As a result, swapping between disordered conformers of similar energies in the T-RE folding run much easier than between collapsed states with larger energy differences in the T-RE control run.

The apparent convergence of T-RE simulations is often achieved through mixing of replicas trapped in different conformational states rather than direct sampling of reversible transitions for each replica.^{17,20} Indeed, further examination of key thermodynamic properties suggests potential benefits of enhanced mixing of T-REs, despite its inability to directly accelerate reversible folding/unfolding transitions. We compare the distributions of the number of native hydrogen bonds of GB1p at 270 K derived from T-RE and T-REs in Figure 4. The convergence can be directly quantified by the RMSD of the distributions from the independent folding and control simulations, which is 0.11 for the T-RE simulations but reduced to 0.08 from T-REs with 1000 attempts. However, the RMSD values are higher, at 0.14 and 0.12 respectively, with 10 and 100 IS swapping attempts. Slightly improved convergence is also evident by examining other conformational properties, such as the end-to-end distance (see Figure S4). However, with a lack of true reversible folding transitions, the autocorrelation functions of the number of native hydrogen bond formed at the lowest condition are rugged and do not consistently reach plateaus (see Fig. S5), leading to unreliable single exponential fitting. The calculated decorrelation times does not suggest robust improvement due to enhanced mixing (Table 2). Improvement in convergence due to enhanced mixing appears to be more evident and more consistent when examining multi-dimensional free energy surfaces. For example, the free energy surfaces as a function of the number of native H-bonds and end-to-end distances derived from the folding and control runs have RMSD values ranging from 0.83 to 1.17 kcal/mol from T-REs runs with various swapping attempts per RE cycle, compared to a value of 1.38 for T-RE (Figure 5 and Figure S6).

T-REs for enhanced sampling of disordered protein conformations

CG simulations of protein G B1 domain and atomistic simulations of GB1p clearly demonstrate the challenges of tempering in enhancing cooperative transitions even with T-REs. IDPs are an import class of functional proteins that can partially or as a whole exist as

heterogeneous ensembles of fluctuating structures under physiological conditions.^{58–62} It is plausible that T-REis may be more effective in enhanced sampling of such disordered conformational ensembles, as the underlying transitions are mostly local and do not involve large entropic barriers. Two disordered proteins, (AAQAA)₃ and unbound KID, are used here to examine the efficacy of T-REis in atomistic simulation of IDPs. Key statistics are summarized in Tables 3 and 4. Interestingly, T-REis does not appear to be able to significantly enhance replica mixing for either (AAQAA)₃ or KID, and often generate fewer temperature round trips per ns of simulation time than T-RE runs. This likely indicates that the conventional neighbor-only exchange protocol is already very efficient in driving replica diffusion in the temperature space for disordered proteins. The apparent reduction of τ_{TR} in T-REis runs is a bit puzzling considering that IS swapping attempts similarly increases the probabilities of replicas exchanging to distal conditions in temperature space (Figure S7). Further examination of the pair-wise exchange acceptance ratios reveals non-trivial consequence of IS swapping attempts. The results shown in Figure S7 includes statistics from all IS swapping attempts throughout the entire T-REis runs. Even though the acceptance ratios are quite high and uniform for T-RE runs, they become increasingly depressed and more non-uniform with more IS swapping attempts. This explains why T-REis may generate smaller number of temperature round trips despite greater individual replica mobility. While depression of apparent exchange acceptance ratios may be expected as T-REis involves more attempts to exchange replicas to distal conditions, the origin of the non-uniformity of the profiles is not directly clear. We have verified that increasing the number of replicas does not help to improve the uniformity of exchange acceptance ratios (Figure S7B). Similar observations can also be made for T-REis simulations of KID (Figure S8). It is likely that the non-uniformity is due to nontrivial distribution of density of states over the temperature range. The slightly reduction of τ_{TR} in T-REis simulations does not appear to have any obvious negative effects on conformational sampling efficiency. Even though the decorrelation times of some residue helical states increase with IS swapping attempts (Figure S9 and S10), similar number of helix-coil transitions were generated during T-RE and T-REis runs for both (AAQAA)₃ and KID (Tables 3 and 4), and the convergence of various thermodynamic properties appears to be slightly improved in T-REis. For example, the RMSD between the residue helicity profiles of (AAQAA)₃ derived from control and folding runs is 0.04 – 0.06 for T-REis with different swapping attempts per RE cycle, consistently lower than that of 0.07 for T-RE simulations even though the later simulations were twice as long. For the larger and more complex KID peptide, T-REis also appears to consistently improve the convergence, capable of generating more consistent residue helicity profiles even with shorter simulations (Figure 7).

Conclusions

We have carefully examined the efficiency of T-REis for enhanced sampling using both CG and atomistic models of ordered and disordered proteins. Clearly, T-REis is highly efficient in enhancing the mixing of replicas in the temperature space without compromising the detail balance or introducing substantial computational overhead. The ability of T-REis to enhance replica mixing appears to be system dependent, where stronger enhancement is achieved for folded proteins with cooperative transitions. However, the observed

enhancement in replica mixing (or lack of it for disordered proteins) does not seem to have any substantial impacts on the efficiency of sampling local and/or global conformational transitions. The current study thus re-affirms the notion that the efficiency of T-RE protocols for proteins does not appear to be limited by temperature diffusion, but by the inherent rates of spontaneous large-scale conformational re-arrangements. In particular, the free energy barriers of cooperative transitions are dominated by entropic components, such that tempering is fundamentally ineffective for driving faster transitions.¹⁷ Nonetheless, there appear to be some benefits of enhanced mixing in achieving slightly better convergence of various thermodynamic properties, despite the lack of direct effects on conformational sampling. In addition, we anticipate that the simple independence sampling scheme could be much more effective when incorporated with various Hamiltonian replica exchange protocols, particularly when the underlying energy landscapes are modified via biasing potentials or selected force field components to reduce the transition energy barriers. For instance, we previously developed a multiscale enhanced sampling (MSES) method, where efficient coarse-grained models are coupled with atomistic ones to accelerate conformational diffusion on atomistic energy landscape.^{52,53} Biased introduced by the coupling is removed via Hamiltonian replica exchange. The sampling efficiency depends critically on how well conformational fluctuations are coupled at coarse-grained and atomistic levels and how effective accelerated dynamics are propagated to the uncoupled limit. Incorporation of infinite sampling is expected to be highly beneficial in our ongoing extension and optimization of MSES to larger and more complex systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by the National Institutes of Health (GM114300). The computing was performed using the Beocat cluster at Kansas State University and Extreme Science and Engineering Discovery Environment (XSEDE) facilities (TG-MCB140210).

References

1. Bernardi RC, Melo MCR, Schulten K. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2015; 1850(5):872–877. [PubMed: 25450171]
2. Christen M, van Gunsteren WF. *Journal of Computational Chemistry*. 2007; 29(2):157–166.
3. Zuckerman DM. *Annual Review of Biophysics*. 2011; 40(1):41–62.
4. Luitz M, Bomblies R, Ostermeir K, Zacharias M. *Journal of physics Condensed matter: an Institute of Physics journal*. 2015; 27(32):323101. [PubMed: 26194626]
5. Freddolino PL, Liu F, Gruebele M, Schulten K. *Biophysical Journal*. 2008; 94(10):L75–77. [PubMed: 18339748]
6. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. *Science*. 2011; 334(6055):517–520. [PubMed: 22034434]
7. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W. *Science*. 2010; 330(6002):341–346. [PubMed: 20947758]
8. Laio A, Parrinello M. *Proceedings of the National Academy of Sciences of the United States of America*. 2002; 99(20):12562–12566. [PubMed: 12271136]

9. Hamelberg D, Mongan J, McCammon JA. *Journal of Chemical Physics*. 2004; 120(24):11919. [PubMed: 15268227]
10. Wang L, Friesner RA, Berne BJ. *Journal of Physical Chemistry B*. 2011; 115(30):9431–9438.
11. Liu P, Kim B, Friesner RA, Berne BJ. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102(39):13749–13754. [PubMed: 16172406]
12. Darve E, Rodríguez-Gómez D, Pohorille A. *Journal of Chemical Physics*. 2008; 128(14):144120–144120. [PubMed: 18412436]
13. Darve E, Pohorille A. *Journal of Chemical Physics*. 2001; 115(20):9169.
14. Sugita Y, Okamoto Y. *Chemical Physics Letters*. 1999; 314(1):141–151.
15. Hansmann UHE. *Chemical Physics Letters*. 1997; 281(1-3):140–150.
16. Okamoto Y. *Journal of Molecular Graphics and Modelling*. 2004; 22(5):425–439. [PubMed: 15099838]
17. Zhang W, Chen J. *Journal of Chemical Theory and Computation*. 2013; 9(6):2849–2856. [PubMed: 26583873]
18. Zuckerman DM, Lyman E. *Journal of Chemical Theory and Computation*. 2006; 2(4):1200–1202.
19. Predescu C, Predescu M, Ciobanu CV. *Journal of Physical Chemistry B*. 2005; 109(9):4189–4196.
20. Periolo X, Mark AE. *Journal of Chemical Physics*. 2007; 126(1):014903. [PubMed: 17212515]
21. Okur A, Wickstrom L, Layten M, Geney R, Song K, Hornak V, Simmerling C. *Journal of Chemical Theory and Computation*. 2006; 2(2):420–433. [PubMed: 26626529]
22. Nymeyer H. *Journal of Chemical Theory and Computation*. 2008; 4(4):626–636. [PubMed: 26620937]
23. Denschlag R, Lingenheil M, Tavan P. *Chemical Physics Letters*. 2008; 458(1-3):244–248.
24. Cheng X, Cui G, Hornak V, Simmerling C. *Journal of Physical Chemistry B*. 2005; 109(16):8220–8230.
25. Zheng W, Andrec M, Gallicchio E, Levy RM. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(39):15340–15345. [PubMed: 17878309]
26. Zheng W, Andrec M, Gallicchio E, Levy RM. *Journal of Physical Chemistry B*. 2008; 112(19):6083–6093.
27. Abraham MJ, Gready JE. *Journal of Chemical Theory and Computation*. 2008; 4(7):1119–1128. [PubMed: 26636365]
28. Sindhikara D, Meng Y, Roitberg AE. *Journal of Chemical Physics*. 2008; 128(2):024103. [PubMed: 18205439]
29. Sindhikara DJ, Emerson DJ, Roitberg AE. *Journal of Chemical Theory and Computation*. 2010; 6(9):2804–2808. [PubMed: 26616081]
30. Plattner N, Doll JD, Dupuis P, Wang H, Liu Y, Gubernatis JE. *Journal of Chemical Physics*. 2011; 135(13):134111. [PubMed: 21992286]
31. Plattner N, Doll JD, Meuwly M. *Journal of Chemical Theory and Computation*. 2013; 9(9):4215–4224. [PubMed: 26592410]
32. Jo S, Chipot C, Roux B. *Journal of Chemical Theory and Computation*. 2015; 11(5):2234–2244. [PubMed: 26574422]
33. Itoh SG, Okumura H. In *J Chem Theory Comput*. 2013:570–581.
34. Galvelis R, Sugita Y. In *J Comput Chem*. 2015:1446–1455.
35. Galvelis R, Re S, Sugita Y. *J Chem Theory Comput*. 2017:1934–1942. [PubMed: 28399363]
36. Chodera JD, Shirts MR. *Journal of Chemical Physics*. 2011; 135(19):194110. [PubMed: 22112069]
37. Zhang BW, Dai W, Gallicchio E, He P, Xia J, Tan Z, Levy RM. *Journal of Physical Chemistry B*. 2016:2015. acs.jpcc.6b0.
38. Xia J, Flynn WF, Gallicchio E, Zhang BW, He P, Tan Z, Levy RM. *Journal of Computational Chemistry*. 2015; 36(23):1772–1785. [PubMed: 26149645]
39. Gallicchio E, Xia J, Flynn WF, Zhang B, Samlalsingh S, Montes A, Levy RM. *Computer Physics Communications*. 2015; 196:236–246. [PubMed: 27103749]

40. Brooks BR, Brooks CL III, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, JZ Pu, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. *Journal of Computational Chemistry*. 2009; 30(10):1545–1614. [PubMed: 19444816]
41. Feig M, Karanicolas J, Brooks CL III. *Journal of Molecular Graphics and Modelling*. 2004; 22(5): 377–395. [PubMed: 15099834]
42. Team, R. D. C. R Foundation for Statistical Computing. Vienna Austria: 2008.
43. Karanicolas J, Brooks CL III. *Protein Science*. 2002; 11(10):2351–2361. [PubMed: 12237457]
44. Miyazawa S, Jernigan RL. *Journal of Molecular Biology*. 1996; 256(3):623–644. [PubMed: 8604144]
45. Kuszewski J, Gronenborn AM, Clore GM. *Journal of the American Chemical Society*. 1999; 121(10):2337–2338.
46. Fesinmeyer RM, Hudson FM, Andersen NH. *Journal of the American Chemical Society*. 2004; 126(23):7238–7243. [PubMed: 15186161]
47. Shalongo W, Dugad L, Stellwagen E. *Journal of the American Chemical Society*. 1994; 116(18): 8288–8293.
48. Best RB, Mittal J, Feig M, Mackerell AD Jr. *Biophysical Journal*. 2012; 103(5):1045–1051. [PubMed: 23009854]
49. Huang J, Mackerell AD Jr. *Biophysical Journal*. 2014; 107(4):991–997. [PubMed: 25140435]
50. Im W, Lee MS, Brooks CL III. In *J Comput Chem*. 2003:1691–1702.
51. Ryckaert J-P, Ciccotti G, Berendsen HJC. *Journal of Computational Physics*. 1977; 23:327–341.
52. Zhang W, Chen J. *Journal of Chemical Theory and Computation*. 2014; 10(3):918–923. [PubMed: 26580171]
53. Lee KH, Chen J. *Journal of Computational Chemistry*. 2016; 37(6):550–557. [PubMed: 26052838]
54. Ganguly D, Chen J. *Journal of the American Chemical Society*. 2009; 131(14):5214–5223. [PubMed: 19278259]
55. Radhakrishnan I, Pérez-Alvarado GC, Parker D, Dyson HJ, Montminy MR, Wright PE. *Cell*. 1997; 91(6):741–752. [PubMed: 9413984]
56. Radhakrishnan I, Pérez-Alvarado GC, Dyson HJ, Wright PE. *FEBS letters*. 1998; 430(3):317–322. [PubMed: 9688563]
57. Wallin S, Chan HS. In *J Phys Condens Matter*. 2006:S307–S328.
58. Uversky VN, Oldfield CJ, Dunker AK. *J Mol Recognit*. 2005; 18(5):343–384. [PubMed: 16094605]
59. Dyson HJ, Wright PE. *Nat Rev Mol Cell Biol*. 2005; 6(3):197–208. [PubMed: 15738986]
60. Smock RG, Gierasch LM. *Science*. 2009; 324(5924):198–203. [PubMed: 19359576]
61. Hilser VJ, Thompson EB. *Proc Natl Acad Sci U S A*. 2007; 104(20):8311–8315. [PubMed: 17494761]
62. Tsai CJ, Ma B, Sham YY, Kumar S, Nussinov R. *Proteins*. 2001; 44(4):418–427. [PubMed: 11484219]

Figure 1. CG and atomistic model protein systems. (A) CG model of protein G B1 domain, and atomistic model of (B) GB1p, (C) (AAQAA)₃, and (D) KID. All presentations are colored based on secondary structures.

Figure 2. (A) The heat capacity (C_v) as a function of temperature and (B) free energy as a function of the native contact fraction (Q_c) at $T_m = 354$ K, calculated from T-RE and T-REis simulations.

Figure 3.
(A) Distributions of the difference between the original and new condition numbers and (B) the ratios of replicas remaining at the same conditions.

Figure 4.

Distributions of the number of native hydrogen bonds formed at 270 K calculated from (A) T-RE and T-REis simulations with (B) 10, (C) 100, (D) 1000 IS swapping attempts per cycle. The last 80% of data was used for calculating all distributions. Results from both control and folding runs are shown for examining the level of convergence.

Figure 5.

Free energy surfaces as a function of the number of native H-bonds and end-to-end distances at 270 K, calculated from control (A) and folding (B) T-RE simulations as well as control (C) and folding (D) T-REis simulations with 10 IS swapping attempts per RE cycle. Data from the last 80 ns sampled at all temperatures were used in the WHAM analysis. The RMSD of 2D PMFs between control and folding runs is 1.38 kcal/mol for T-RE and 0.83 kcal/mol for T-REis. Note that regions with free energy >8 kcal/mol were not included in the RMSD calculation.

Figure 6. Convergence of residue helicity profiles of (AAQAA)₃ at 270K, calculated using (A) T-RE and T-REis with (B) 10, (C) 100, (D) 1000 IS swapping attempts per RE cycle. The RMSDs between helicity profiles derived from control and folding runs are 0.07 for T-RE (using 120-200 ns data), and 0.06, 0.04 and 0.06 for results from control and folding T-REis runs with 10, 100 and 1000 IS swapping attempts (using 60-100 ns data), respectively.

Figure 7.

Residue helicity profile of KID. Convergence of residue helicity profiles of KID using (A) T-RE and T-REis with (B) 10, (C) 100, (D) 1000 IS swapping attempts per cycle. The RMSDs between control and folding helicity profiles are 0.20 for T-RE, and 0.23, 0.06 and 0.13 for T-REis runs with 10, 100 and 1000 IS swapping attempts, respectively.

Table 1

Key statistics of CG T-RE and T-REis simulations of protein G B1 domain.

Protocol	Run	Length (μ s)	T_m (K)	N_{TR}	N_{TS}	τ_{TS} (μ s)	τ_{icf} (μ s)
T-RE	Control	100	352	1411	21	38.1	0.31
	Folding	100	351	1937	25	32.0	0.78
T-REis (10)	Control	100	352	2588	21	38.1	0.30
	Folding	100	346	4756	18	44.4	1.81
T-REis (100)	Control	100	343	9746	13	61.5	0.32
	Folding	100	346	11183	15	53.3	3.57
T-REis (1000)	Control	100	343	19817	21	38.1	0.50
	Folding	100	342	23930	22	36.4	1.38

T_m values were estimated based on the CV curves. N_{TR} is the total number of round trips that all replicas travel between the temperature extremes. N_{TS} is the number of reversible folding transitions sampled by all replicas during the entire course of the simulation. The effective reversible transition rate, τ_{TS} , is calculated as N_{TS} divided by the total simulation time. τ_{icf} is the decay half time derived from the autocorrelation function of Q_C at the lowest condition. The correlation functions are shown in Fig. S1.

Table 2

Summary of atomistic T-RE and T-REis simulations of GB1p.

Protocol	Run	Length (ns)	N_{TR}	N_{TS}	τ_{TS} (ns ⁻¹)	τ_{def} (ns)
T-RE	Control	100	283	1	0.0013	9.18
	Folding	100	1247	0	undef	8.43
T-REis (10)	Control	100	1483	0	undef	10.26
	Folding	100	1432	0	undef	2.39
T-REis (100)	Control	100	1745	0	undef	9.82
	Folding	100	1572	0	undef	4.97
T-REis (1000)	Control	100	2527	0	undef	8.62
	Folding	100	2371	0	undef	10.49

See Table 1 for the definition of N_{TR} , N_{TS} , and τ_{TS} . A folded state is defined by 5 native hydrogen bonds formed, and an unfolded state has no native hydrogen bond formed. All simulations involve eight replicas. The effective reversible transition rate τ_{TS} is calculated as N_{TS} divided by total simulation time (800 ns in all cases). τ_{def} is the decay half time derived from the autocorrelation function of the number of native hydrogen bond formed at the lowest condition. The correlation functions are shown in Fig. S5.

Table 3

Summary of atomistic T-RE and T-REis simulations of (AAQAA)₃.

Protocol	Run	Length (ns)	Temperature τ_R (ns ⁻¹)	Residue 4 (Å)		Residue 8 (Q)		Residue 12 (Å)	
				τ_{TS} (ns ⁻¹)	τ_{ref} (ns)	τ_{TS} (ns ⁻¹)	τ_{ref} (ns)	τ_{TS} (ns ⁻¹)	τ_{ref} (ns)
T-RE	Control	200	0.39	12.26	1.67	23.57	1.66	18.11	1.15
	Folding	200	0.39	12.67	0.68	23.62	2.95	16.77	1.49
T-REis (10)	Control	100	0.25	12.43	1.18	23.91	1.85	17.47	4.09
	Folding	100	0.26	13.07	0.29	23.38	1.39	16.64	1.96
T-REis (100)	Control	100	0.28	11.78	0.67	23.53	1.75	15.32	2.04
	Folding	100	0.26	12.34	0.58	22.58	2.57	17.44	0.83
T-REis (1000)	Control	100	0.31	11.05	1.45	20.97	4.50	17.34	3.54
	Folding	100	0.28	11.78	0.50	22.26	2.43	15.72	0.60

See Table 1 for the definition of τ_{TS} $\tau_{TR} = NTR/(\text{simulation})$ Length. Three residues are selected for reporting the helix-coil transition rates (τ_{TS}), which locate in the middle, the first and the third quarter of the peptide. τ_{ref} is the decay half time of residue helical state calculated from the autocorrelation function at the lowest condition.

Table 4

Summary of atomistic T-RE and T-REis simulations of KID.

Protocol	Run	Length (ns)	Temperature τ_{TR} (ns ⁻¹)	Residue 124 (R)		Residue 138 (L)	
				τ_{TS} (ns ⁻¹)	τ_{act} (ns)	τ_{TS} (ns ⁻¹)	τ_{act} (ns)
T-RE	Control	200	0.18	25.77	23.45	11.41	9.91
	Folding	200	0.19	22.37	16.93	5.33	1.50
T-REis (10)	Control	100	0.08	26.32	5.22	11.34	1.38
	Folding	100	0.10	15.24	4.19	10.38	3.34
T-REis (100)	Control	100	0.12	26.05	9.96	8.12	1.32
	Folding	100	0.15	20.09	4.86	6.88	0.68
T-REis (1000)	Control	100	0.19	23.62	5.92	6.95	0.44
	Folding	100	0.21	13.80	8.17	5.61	1.06

See Table 3 for the definition of τ_{TS} and τ_{TR} . Residues 124 and 138 are selected for reporting the helix-coil transition rates (τ_{TS}), which are located near the middle of helices αA (residues 120-129) and αB (residues 134-144), respectively. τ_{act} is the decay half time of residue helical state calculated from the autocorrelation function at the lowest condition.