JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Peptide Bond Isomerization in High-Temperature Simulations

Chris Neale,[†,⊥] Régis Pomès,[‡,§] and Angel E. García*[,†,∥,⊥]

[†]Department of Physics, Applied Physics and Astronomy and [∥]Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180-3590, United States

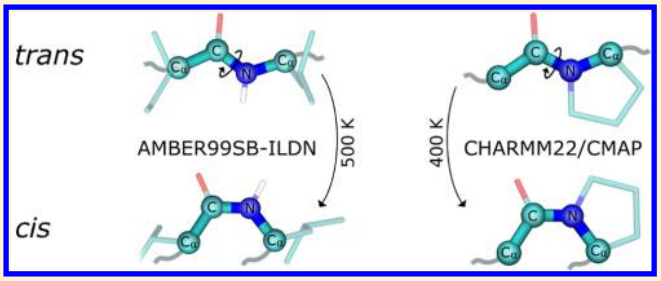
[‡]Molecular Structure and Function, The Hospital for Sick Children, 686 Bay Street, Toronto, Ontario M5G 0A4, Canada

[§]Department of Biochemistry, University of Toronto, 101 College Street, Toronto, Ontario M5G 1L7, Canada

[⊥]Center for NonLinear Studies (CNLS), MS B258, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States

Supporting Information

ABSTRACT: Force fields for molecular simulation are generally optimized to model macromolecules such as proteins at ambient temperature and pressure. Nevertheless, elevated temperatures are frequently used to enhance conformational sampling, either during system setup or as a component of an advanced sampling technique such as temperature replica exchange. Because macromolecular force fields are now put upon to simulate temperatures and time scales that greatly exceed their original design specifications, it is appropriate to re-evaluate whether these force fields are up to the task. Here, we quantify the rates of peptide bond isomerization in high-temperature simulations of three octameric peptides and a small fast-folding protein. We show that peptide octamers with and without proline residues undergo *cis*/*trans* isomerization every 1−5 ns at 800 K with three classical atomistic force fields (AMBER99SB-ILDN, CHARMM22/CMAP, and OPLS-AA/L). On the low microsecond time scale, these force fields permit isomerization of nonprolyl peptide bonds at temperatures ≥500 K, and the CHARMM22/CMAP force field permits isomerization of prolyl peptide bonds ≥400 K. Moreover, the OPLS-AA/L force field allows chiral inversion about the C_α atom at 800 K. Finally, we show that temperature replica exchange permits *cis* peptide bonds developed at 540 K to subsequently migrate back to the 300 K ensemble, where *cis* peptide bonds are present in 2 ± 1% of the population of Trp-cage TC5b, including up to 4% of its folded state. Further work is required to assess the accuracy of *cis*/*trans* isomerization in the current generation of protein force fields.



INTRODUCTION

Proteins are primarily linear polymers in which consecutive amino acid residues are joined by a covalent bond between the backbone nitrogen atom of residue *i* and the backbone carbonyl carbon atom of residue *i*-1.[1,2] Resonance structures with shared π electrons in the $C_{i-1}-N_i$ bond drive coplanarity of the associated sp^2 orbitals such that the peptide bond dihedral angle, ω ($C^{\alpha}_{i-1}-C_{i-1}-N_i-C^{\alpha}_i$), adopts values near (*trans*) 180° or (*cis*) 0°.[3] The apparent requirement for sp^3 hybridization of the N_i atom during isomerization[4,5] and other electronic considerations[6,7] lead to a *cis*/*trans* rotational barrier of 15 to 20 kcal/mol,[8,9] greatly slowing peptide bond isomerization. For example, at pH 5.5 and a temperature, *T*, of ∼298 K, the uncapped Tyr-Ala dipeptide exhibits a *cis* → *trans* isomerization rate of 0.2−0.5 s^{-1}.[8] The reverse (*trans* → *cis*) isomerization rate is 2 orders of magnitude slower, and the resulting population of the *trans* isomer is ∼99.5%.[8]

This abundance of *trans* peptide bonds is overwhelmingly conserved across proteins of known structure because the *cis* conformation of the peptide bond is dramatically disfavored by its predisposition for steric clashes between the side chains of residues *i*-1 and *i*, in addition to other effects.[10] However, when residue *i* is a proline, its unique architecture leads to relative stabilization of the *cis* isomer. As a result, peptide bonds in the protein data bank (PDB) are in *cis* conformations about 5% and 0.03% of the time for Xaa−Pro[11−13] and Xaa−non-Pro[13,14] pairs, respectively, where Xaa is any amino acid. For short peptides in solution at room temperature, *cis* isomers can attain populations of up to 30% and 1% for Xaa−Pro[15] and Xaa−non-Pro[8] pairs, respectively.

Despite their relative rarity, *cis* peptide bonds remain biologically relevant.[16,17] For example, Xaa−Pro peptide bond isomerization can act as a switch in cellular regulatory pathways[18−21] and the activation of ion channels.[22] Furthermore, protein folding kinetics are influenced by Xaa−Pro[23−25] and Xaa−non-Pro[26,27] peptide bond isomerization. However, it is difficult to predict the *cis* population *a priori* because the peptide bond's propensity for the *cis* isomer depends on sequence,[8,15,28] increases with temperature,[8,29] and in some cases decreases with peptide chain length.[8]

Computer simulations have been used to study specific cases where peptide bond isomerization is relevant to protein function.[20,30,31] Nevertheless, peptide bond isomerization is

Received: October 29, 2015
Published: February 11, 2016

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

usually neglected in computer simulations, likely because the potential energy barrier to isomerization is so large. In this article, we assess peptide bond isomerization in computer simulations, with special focus on the enhancement of the rate at which this conformational reorganization occurs at elevated temperatures.[8,29]

Atomistic computer simulations probe macromolecular structure and dynamics with exceptional resolution.[32−37] However, due to computational limitations, the time scales that are routinely accessible to methods such as molecular dynamics (MD) are often substantially less than those of biophysical phenomena of interest. For example, the longest biomolecular simulations reported to date access the 1 to 30 ms time scale,[38,39] whereas lipid flip-flop can take hours,[40] the folding of large RNA molecules can take seconds or minutes,[41,42] and even two-state protein folding times can exceed a second,[43] especially in the presence of rate-limiting proline isomerization.[44]

In some cases, this time scale deficiency can be alleviated by using an enhanced sampling technique.[45,46] One such technique is called temperature replica exchange molecular dynamics (T-REMD) simulation.[47] T-REMD takes advantage of the increased rate of conformational sampling at elevated temperature by employing a set of weakly coupled simulations that undergo reciprocal exchanges along a temperature ladder. In this method, individual replica simulations can migrate to high temperature, undergo relatively rapid conformational decorrelation, and subsequently migrate back to the low-temperature ensemble of interest. Because the exchange acceptance criterion maintains equilibrium sampling at each temperature, replicas sampling states with high potential energies tend to be promoted to higher temperatures whereas those replicas sampling energetically favorable states tend to migrate to lower temperatures.[48]

Efficient and successful T-REMD simulations must satisfy at least the following three criteria. First, in order for the simulations to provide useful models, the states that are highly populated in experiment must also be highly populated in the limit of infinite sampling, and the relative populations should be correct. That is to say the force field must be accurate.[49,50] Second, the highest temperature in the ladder must be high enough to permit rapid conformational decorrelation on time scales that are short in comparison to the simulation.[51,52] Third, the random walk between the highest and lowest temperatures must be sufficiently rapid that many transitions occur between the lower-temperature ensembles of interest and the higher-temperature ensembles in which conformations rapidly decorrelate.[52,53]

Bruccoleri and Karplus noted in 1990 that a simulation temperature of 1500 K (but not 800 K) facilitates the formation of *cis* Xaa−non-Pro peptide bonds in the 235-residue immunoglobulin McPC 603 on the 100 ps time scale using the CHARMM PARAM1 force field.[54] 25 years later, massive advances in hardware speed and software efficiency allow the routine simulation of substantially longer time scales,[55−58] increasing the probability of crossing large potential energy barriers at lower temperatures. This increase in sampling power calls for a reassessment of the temperature dependence of conformational interconversions in simulations of macromolecules such as proteins. An excellent review of peptide bond isomerization, protein/nucleic acid stereochemistry, and related sources of error in simulations of biomolecules is provided by Schreiner et al.[59]

We begin by evaluating peptide bond isomerization as a function of temperature for three octameric peptides using four classical atomistic force fields. We show that temperatures of 500−600 K are sufficient to permit Xaa−non-Pro peptide bond isomerization and that enantiomeric interconversion can occur at 800 K. We then conduct T-REMD simulations of the TC5b Trp-cage miniprotein and show that replicas that develop *cis* peptide bonds at high temperatures can subsequently migrate back to the low-temperature ensemble of interest, where they represent 2 ± 1% of the 300 K ensemble, including 3.6% of the folded state in one of three T-REMD simulations.

RESULTS

Standard Simulations of Octapeptides. We begin by assessing the time scales of peptide bond isomerization as a function of temperature in MD simulations using classical atomistic force fields. To this end, we focus on three eight-residue peptides: MTYKLILN, TAEKVFKQ, and DGGPSSGR. Each sequence is simulated 20 times for 100 ns at each of five different temperatures between 400 and 800 K using four different force fields: AMBER99SB-ILDN,[60] AMBER99SB-ILDN with the peptide bond torsion potential modifications of Doshi and Hamelberg,[61] CHARMM22[62] with grid-based energy correction maps (CMAP),[63] and OPLS-AA/L.[64] Examples of the spontaneous transitions from *trans* to *cis* peptide bonds in these simulations are shown in Figure 1.

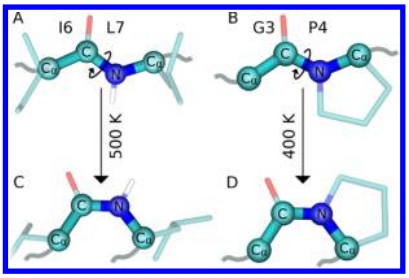


Figure 1. Spontaneous peptide bond isomerization at elevated temperature. Snapshots depict representative (A, B) *trans* and (C, D) *cis* peptide bond configurations in simulations of (A, C) MTYKLILN at 500 K using the AMBER99SB-ILDN force field and (B, D) DGGPSSGR at 400 K using the CHARMM22/CMAP force field. Simulations are initiated from all-*trans* conformations.

All of the evaluated force fields sample *cis* peptide bonds on the simulated time scale (20 independent 100 ns simulations) at $T \geq 700$ K (Figure 2 and Tables S1 and S2). Furthermore, only the Doshi force field fails to sample *cis* peptide bonds in all three peptides at 600 K (Figure 2 and Table S3). The AMBER force field samples *cis* peptide bonds in all three peptides at 500 K (Figure 2 and Table S4), and the CHARMM force field samples Gly−Pro *cis* peptide bonds at temperatures as low as 400 K (Figure 2 and Table S5). Importantly, at $T \geq 700$ K, *cis* peptide bonds arise at all possible locations in all force fields except that of Doshi (Figures S1−S3), which was designed to increase the potential energy barrier between *cis* and *trans* conformations of the peptide bond.[61] Moreover, *cis* conformations of Xaa−non-Pro peptide bonds occur at temperatures as low as 600 K for all force fields and 500 K for AMBER (Figures S1−S3).

The aforementioned simulations were all conducted using an integration time step of 2 fs. To evaluate whether the formation of *cis* peptide bonds at elevated temperatures can be mitigated

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

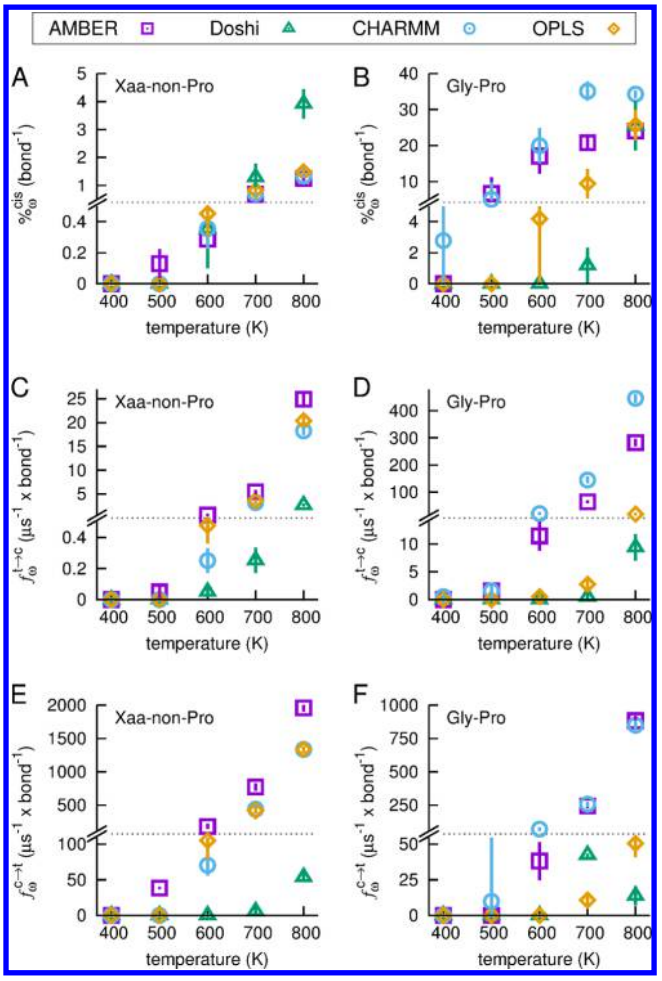


Figure 2. Peptide bond isomerization in octapeptides as a function of temperature. (A, B) Average percentage of peptide bonds in the *cis* isomer, $\%_{\omega}^{cis}$, per bond. (C–F) Average frequencies of peptide bond isomerization from (C, D) *trans* to *cis*, $f_{\omega}^{t\to c}$, or (E, F) *cis* to *trans*, $f_{\omega}^{c\to t}$, per μs per bond. Data shown for (A, C, E) Xaa–non-Pro peptide bonds in all three peptides and (B, D, F) the Gly–Pro peptide bond in DGGPSSGR. Bars estimate the statistical error as the standard deviation from bootstrap resampling. Dotted horizontal line denotes a change in ordinate scale.

by decreasing the integration time step, we repeat the AMBER simulations using a time step of 0.2 fs. Computed kinetic and thermodynamic properties do not depend on the time step in this range (Tables S1–S5 and Figures S1–S3), although the low frequency of peptide bond isomerization at 500 K led one peptide (TAEKVFKQ) to sample *cis* peptide bonds only with the larger time step (Table S4). Furthermore, extending the simulations at 500 K to 2 μs each leads more force fields to access *cis* conformations of peptide bonds but does not dramatically reduce the associated error estimates (Table S6).

While the MTYKLILN and TAEKVFKQ peptides sample *cis* peptide bonds to similar extents, the DGGPSSGR peptide exhibits approximately twice as much *cis* content in Xaa–non-Pro peptide bonds (Figures S4A-C and Tables S1–S6) due to an increase in the rate of *trans* → *cis* isomerization without a compensating increase in the rate of the reverse transformation (Figures S4D-I). However, this increase in Xaa–non-Pro peptide bond *cis* content does not correlate with Gly–Pro peptide bond isomerization (Figures S4J, K) and may instead be related to the presence of glycine residues and/or terminal charges in the DGGPSSGR peptide (Figures S1–S3).

We also note that one OPLS simulation of the TAEKVFKQ peptide at 800 K exhibits chiral inversion at the C_{α} atom of V5 at 70 ns, after which the *D*-form is stable until the simulation is terminated at 100 ns (Figure S5).

T-REMD Simulations of Trp-Cage TC5b. Since peptide bonds undergo isomerization at elevated temperature, T-REMD simulations may permit *cis* peptide bonds developed at high temperature to become temporarily locked-in when a replica migrates to lower temperature. To assess the potential impact of *cis* peptide bond formation on simulations focusing on protein folding, we conduct three T-REMD simulations of the TC5b Trp-cage miniprotein starting from an unfolded conformation using the AMBER99SB-ILDN force field. Here, we reduce the maximum temperature to 540 K in order to make our evaluation more relevant to contemporary T-REMD simulations (32 replicas spanning temperatures from 300 to 540 K).

Starting from an unfolded conformation (Figure S6), all three of our T-REMD simulations attain the folded structure (C_{α} RMSD to folded ≤0.22 nm) within 15–165 ns/replica, and the folded state is sampled 65 ± 9% of the time at 300 K after 250 ns/replica (Figure 3), consistent with our previous T-REMD study of this molecule.[65] In addition, 40 ± 7% of the replicas visit the folded state by 250 ns/replica (Figure S7A), and beyond this time the proportion of replicas that are currently folded is 16 ± 2% (Figure S7B). Further analysis reveal that the protein's conformational ensemble is substantially more expanded at 540 K than at 300 K (Figure S8), overall replica migration in temperature is extensive (Figure S9A), and replicas that remain at low temperatures for lengthy periods sample the folded state often but not always (Figures S9A, B).

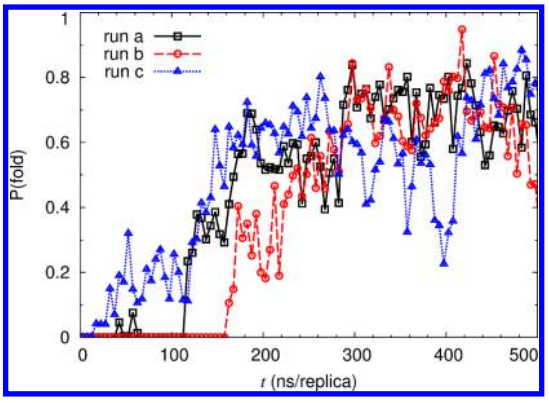


Figure 3. Trp-cage TC5b folding in three T-REMD simulations. Proportion of the folded state (C_{α} RMSD to folded ≤0.22 nm), P(fold), in the 300 K ensemble as a function of simulation time per replica. Data points represent 5 ns block averages[66] and are connected by lines to guide the eye.

Based on the analysis outlined above, these T-REMD simulations of Trp-cage TC5b appear to be a success. The folded state is highly populated (Figure 3), the highest temperature in the ladder is sufficient to unfold the protein (Figure S8), and the random walk carries replicas between temperature extremes (Figure S9), leading most of the replicas to adopt the folded state at some point during the simulation (Figure S7A). Nevertheless, these T-REMD simulations of Trp-

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989–1999

cage TC5b are pervaded by the same type of peptide bond isomerization that we identified in constant-temperature simulations of octapeptides at elevated temperatures. Specifically, 11 and seven of the peptide bonds in this 20-residue protein adopt a *cis* conformation at $T > 362$ K and $T = 300$ K, respectively, during at least one of the three T-REMD simulations (Figure 4).

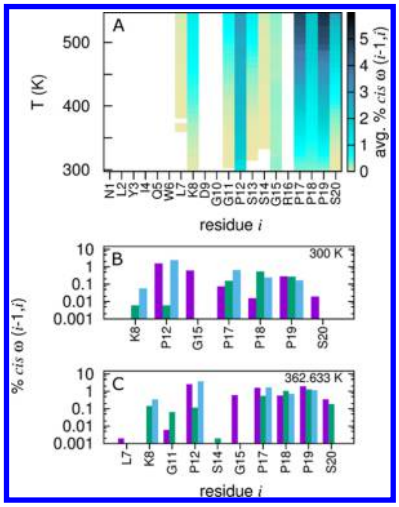


Figure 4. Temperature dependence of *cis* peptide bonds in T-REMD simulations of Trp-cage TC5b. (A) Per-residue % *cis* peptide bond dihedral angle, ω, averaged over three T-REMD simulations, as a function of temperature. (B, C) Histograms depicting, in different colors for each of the three T-REMD simulations, the per residue % *cis* ω at (B) 300 K and (C) 362.633 K. Note the logarithmic ordinate in parts B and C.

All four Xaa−Pro peptide bonds (G11-P12, R16-P17, P17-P18, and P18-P19) sample *cis* conformations at 300 K in all three of our T-REMD simulations, with individual *cis* peptide bond populations ranging from 0.01 to 2.42% (Table S7). Overall, *cis* Xaa−Pro peptide bonds are in 2 ± 1% and 13 ± 6% of the 300 and 540 K ensembles, respectively (Tables S7 and S8), with similar quantities in the last 250 ns/replica (Figure S10A).

Of exceptional relevance to computational studies of protein folding, these *cis* peptide bonds are not excluded from the conformational basin that is ascribed to this protein's folded state based on C_α RMSD to the NMR structure (Figure 5). Two replicas with *cis* peptide bonds, both from T-REMD simulation *c*, visit the folded basin (C_α RMSD $\leq$ 0.22 nm) at 300 K (Figure 5A). In both replicas, peptide bond isomerization occurs at 540 K, while the protein is largely unfolded (C_α RMSD $\geq$ 0.55 nm; Figures 5B, 5C). Both of these peptide bonds (G11-P12 and R16-P17) remain in the *cis* configuration for the remainder of the simulation (Figures 5B, 5C). Together, these *cis* isomers are present in 3.6% of the folded structures at 300 K in simulation *c* (3.2% for G11-P12 and 0.4% for R16-P17). This percentage is 3.3% when considering only the last 250 ns/replica (Figure S10C). The apparent compatibility of these *cis* peptide bonds with the folded state is highlighted in Figures 5D-5F. Figure 5G shows snapshots of the isomerization of the G11-P12 peptide bond from *trans* to *cis* and subsequent folding. There is insufficient data to quantify the statistical significance of the reorientation of the R16 side chain and the loop region around S13 that occurs in the presence of the *cis* R16-P17 peptide bond (Figure 5D).

Surprisingly, all three T-REMD simulations sample *cis* peptide bonds at 300 K for at least one Xaa−non-Pro residue pair (Table S7). Specifically, two of the peptide bonds that sample *cis* isomers at 300 K do not involve a proline residue (L7-K8 and S14-G15; Figure 4B and Table S7), and a third does not have proline as the C-terminal residue in the pair (P19−S20; Figure 4B and Table S7). Overall, *cis* Xaa−non-Pro peptide bonds are present in 0.2 ± 0.3% and 6 ± 1% of the 300 and 540 K ensembles, respectively (Tables S7 and S8), with similar quantities in the last 250 ns/replica (Figure S10A). However, structures containing these *cis* Xaa−non-Pro peptide bonds do not contribute to the folded ensemble at 300 K, with minimum values of C_α RMSD to folded at 300 K of 0.43, 0.55, and 0.65 nm for *cis* L7-K8, S14-G15, and P19−S20, respectively. A C_α RMSD value $\leq$0.22 nm occurs in only three snapshots that contain a *cis* Xaa−non-Pro peptide bond: one snapshot of L7-K8 at 398.694 K and two snapshots of W6-L7 at 430.109 K, all in simulation *a* (a single snapshot represents a population of $\sim 6 \times 10^{-7}$ per simulation across all temperatures and replicas). Therefore, although structures containing *cis* Xaa−non-Pro peptide bonds make up 0.01 to 0.63% of the population at 300 K (Table S7) and 4.27 to 7.09% of the population at 540 K (Table S8), they are almost completely excluded from the folded state ensemble at all temperatures.

Time series of peptide bond isomerization, simulation temperature, and metrics of global conformation are shown in Figure 6 for two replicas that sample Xaa−non-Pro peptide bonds. In simulation *c*, one replica adopts a *cis* L7-K8 peptide bond at time $t = 365$ ns, temperature $T = 540$ K and transiently contributes to the 300 K ensemble, spending 0.3% of its time there, before regaining the *trans* isomer at $t = 475$ ns, $T = 540$ K (Figure 6A). Similarly, one replica in simulation *a* adopts a *cis* S14-G15 peptide bond at $t = 20$ ns, $T = 540$ K and spends 3.1% of its time in the 300 K ensemble before regaining the *trans* isomer at $t = 120$ ns, $T = 500.559$ K (Figure 6B). The later replica eventually finds the folded state (in the absence of any *cis* peptide bonds) at $t = 340$ ns and thereafter spends 11.7% of its time at 300 K (Figure 6B).

The relative populations of peptide bond isomers do not converge on the 500 ns/replica time scale at either 300 K (Table S7) or 540 K (Table S8). When evaluated individually, the statistical error in the estimates of *cis* peptide bond populations in these simulations is generally too large to draw conclusions about equilibrium populations (Tables S7 and S8). Nevertheless, these simulations show, with a high degree of statistical significance, that *cis* peptide bonds exist at 300 K in T-REMD simulations of the TC5b Trp-cage when the maximum temperature is 540 K and the AMBER99SB-ILDN/TIP3P force field combination is used (Table S7).

DISCUSSION

We have shown that the peptide bonds of three octapeptides undergo *cis*/*trans* isomerization in MD simulations at elevated temperatures using four atomistic force fields (Figures 1, 2). The increase of the peptide bond isomerization rate and *cis* population with increasing temperature (Figure 2 and Tables S1−S5) is qualitatively consistent with experimental data on model peptides.[8,29] Peptide bond isomerization also occurs in T-REMD simulations of the 20-residue TC5b Trp-cage miniprotein (Figure 4), where high-temperature isomerization contributes structures with *cis* peptide bonds to the 300 K ensemble (Figures 4B, 5, 6). Moreover, these *cis* peptide bonds

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

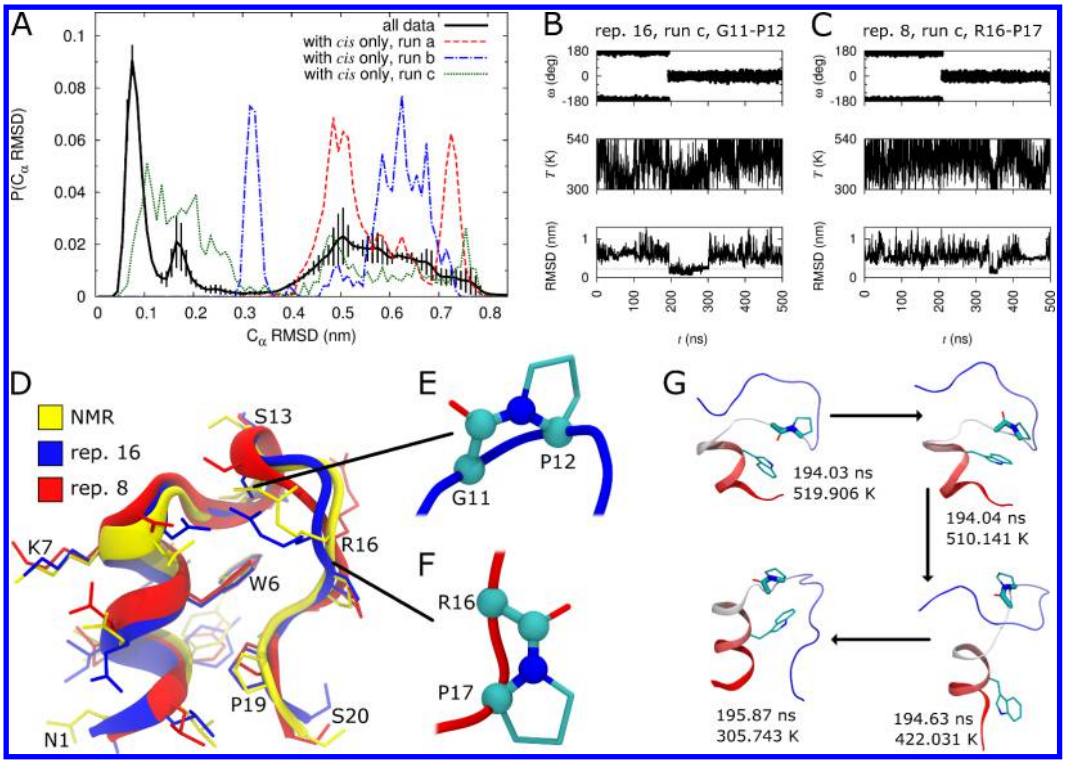


Figure 5. *cis* Xaa−Pro peptide bonds in the folded state ensemble of Trp-cage TC5b at 300 K. (A) Probability histogram of C_α RMSD to folded for (black solid line) all structures and (colored patterned lines) structures with at least one *cis* peptide bond, grouped by T-REMD simulation. For the complete data set, vertical lines depict the standard deviation between histograms based on individual T-REMD simulations. (B and C) Time series of conformational reorganization and replica migration for the two replicas that are folded (C_α RMSD ≤ 0.22 nm) in the presence of a *cis* peptide bond. Time-series depict (top) identified peptide bond dihedral angle, ω (middle), temperature, T, and (bottom) C_α RMSD to folded. The *cis* peptide bond is (part B) G11-P12 and (part C) R16-P17. (D) Superimposed snapshots at 300 K of the replica shown in (blue) part B and (red) part C. The NMR structure (PDB 1L2Y model 1)[67] is in yellow. (E and F) Magnification of the *cis* peptide bonds in the simulation structures depicted in part D. (G) Snapshots showing the development of a *cis* peptide bond between G11-P12 and subsequent folding in replica 16 of simulation *c*. The protein is colored red to blue from N- to C-terminus. The side chain of W6 and the G11-P12 peptide bond are highlighted.

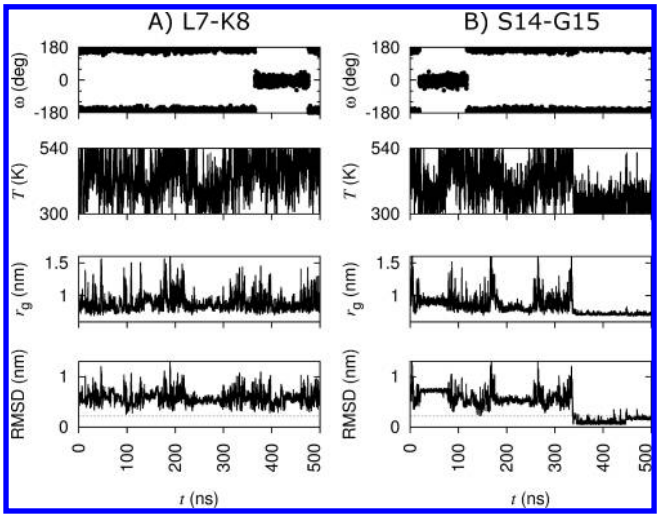


Figure 6. Xaa−non-Pro peptide bond isomerization in T-REMD simulation of Trp-cage TC5b. Representative examples of ω dihedral angle *cis* isomers at 300 K for (A) one replica in simulation *c* (L7-K8 ω) and (B) one replica in simulation *a* (S14-G15 ω). Data show time series of, from top to bottom, ω, temperature T, protein radius of gyration r_g, and C_α RMSD to folded. The dashed gray horizontal line indicates RMSD = 0.22 nm.

are present in 3.6% of the folded state at 300 K in one of three simulations (Figure 5) and are not limited to Xaa−Pro residue pairs (Figure 6). Overall, *cis* peptide bonds are present in 1−3% and 0.01−0.63% of the 300 K ensemble for Xaa−Pro and Xaa−non-Pro, respectively (Table S7).

The highest temperature in our T-REMD simulations, 540 K, is sufficiently high to enable *de novo* folding of the TC5b Trp-cage (Figure 3 and ref 65), while also being sufficiently low so as to remain relevant to similar studies in the literature. For example, higher maximum temperatures, T_{max}, have been used in a variety of T-REMD and standard high-temperature simulations (Table 1).

It is not clear whether the *cis* peptide bonds sampled at low temperatures in our T-REMD simulations of the TC5b Trp-cage accurately represent its conformational ensemble in solution or if they are an artifact of incorrect force field parameters that is only now becoming apparent as longer simulation time scales become accessible. Rovo et al. used temperature-dependent NMR to show that, at pH 3.5, the Trp-cage TC5b folded state is in slow exchange with an unfolded population that has a *cis* G11-P12 peptide bond.[78] This is the same peptide bond that adopts the *cis* isomer in the context of the protein's folded state (C_α RMSD ≤ 0.22 nm) at 300 K in one of our T-REMD simulations (Figures 5D, 5E, 5G). However, Rovo et al. did not detect *cis* peptide bonds in the protein's folded state at acidic pH or in the unfolded state at neutral pH, nor did they detect *cis* isomers of any other Xaa−Pro peptide bonds.[78] Although it seems likely that force field parameters require optimization before they can capture the

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

Table 1. Selected Simulation Studies Employing High Temperatures

T_{max} (K)	protein	simulation method	ref
550	Sac7d	Constant T	68
559	TC10b Trp-cage	T-REMD	69
580	25-residue fragment of amyloid β	T-REMD	70
600	helical transmembrane peptides	T-REMD	71
690	peptides with beta and alpha character	T-REMD	72
700	Met-enkephalin pentapeptide	T-REMD	47
900	peptide derived from protein G	Constant T	73
900	rigin tetrapeptide	Constant T	74
800−1000	synthetic WALP and TMX peptides	T-REMD	75
1000	villin headpiece subdomain	Constant T	76
2000	antimicrobial peptide indolicidin	Constant T	77

nuanced influence of peptide bond isomerization on local and global protein conformation, *cis* peptide bonds may exist in solution at populations that are sufficiently small that they are not resolved by NMR. Moreover, the occurrence of *cis* peptide bonds increases with increasing resolution of crystallographic structures[14] and enhanced conformational processing,[79] suggesting that *cis* peptide bonds may be underrepresented in the PDB. Therefore, we cannot conclusively state that the peptide bond isomerization that we observe in the 300 K ensemble from T-REMD simulations is incorrect. Conversely, the chiral inversion that we observe in one OPLS simulation of the TAEKVFKQ peptide at 800 K (Figure S5) is clearly irrelevant to protein folding at body temperature given that the time scales of biological racemization at 37 °C exceed the human lifetime,[80] indicating that OPLS simulations of ≥100 ns are incompatible with $T \geq 800$ K. Because this type of chiral inversion is bound to become accessible at lower temperatures as simulation time scales increase, we consider the potential for chiral inversion to be a general force field issue that will need to be addressed at some point in the future.

In principle, T-REMD should be able to sort out *cis*/*trans* peptide bond populations and provide reliable estimates of these populations at each sampled temperature. However, two issues arise. First, the force field must generate the correct populations, which is not always the case.[61] Second, even if everything is correct in the limit of infinite sampling, peptide bond isomerization may still introduce a source of sampling inefficiency because the simulations may not be long enough to reach equilibrium distributions of peptide bond torsion angles. The stringency of this later requirement is highlighted by the fact that, on the 500 ns/replica time scale, our T-REMD simulations attain converged estimates of the folded state population at 300 K (Figure 3) but not the proportion of *cis* petide bonds at 300 K (Table S7) or 540 K (Table S8) or in the folded state (Figure S10C). Moreover, if the highest temperature permits peptide bond isomerization, but only rarely, then replicas may occasionally develop *cis* peptide bonds and become temporarily trapped at elevated temperatures, possibly influencing sampling efficiency along the entire temperature ladder.

In essence, our finding is that the increased duration of accessible simulation time scales has led to an era in which peptide bond isomerization is now a source of sampling inefficiency and/or statistical error in MD simulations. Historically, the potential energy barrier to peptide bond isomerization has been sufficiently high to generally disallow conformational transitions about this dihedral angle (or it simply has not been systematically checked). In a *trans*-locked regime, there may be inaccuracy (e.g., when a peptide bond should really adopt the *cis* isomer), but there is no related issue with statistical convergence because only *trans* states are sampled. Conversely, future simulation time scales may be sufficiently long to equilibrate *cis*/*trans* isomerization of all peptide bonds. In the meantime, statistical sampling errors must be considered as simulation time scales are sufficiently long to permit some isomerization but remain insufficient to sample the many back-and-forth transitions required to attain converged populations. One stop-gap strategy is to use modified potentials to abolish peptide bond isomerization in cases where the protein backbone is assumed to exist in an all-*trans* form. To this end, the AMBER simulation software provides the makeCHIR_RST script,[81] and the CHARMM simulation software provides flat-bottom harmonic potentials for torsions,[82] as does the NAMD simulation software via the cispeptide plugin[59] to VMD.[83] To our knowledge, the GROMACS simulation software is not presently capable of disallowing *cis* peptide bonds without affecting fluctuations within the *trans* basin. Conversely, simulations aimed at equilibrating *cis*/*trans* peptide bond ratios may consider lowering the height of the potential energy barriers opposing isomerization.

We note in passing that our three T-REMD simulations, which are identical except for random variation during the runs, require 140 to 240 ns/replica to attain converged values of the proportion of the conformational ensemble that is folded at 300 K (Figure 3). This range of equilibration times may be relevant when trying to optimize T-REMD simulations.

CONCLUSION

High-temperature simulations of three octapeptides show that three commonly used atomistic protein force fields allow *cis*/*trans* isomerization of Xaa−non-Pro peptide bonds within 100 ns at $T \geq 600$ K (500 K for the AMBER99SB-ILDN force field). Furthermore, Xaa−Pro peptide bonds isomerize even more readily (400 K for the CHARMM22 force field with CMAP) and the OPLS-AA/L force field permits C_α chiral inversion at 800 K. In comparison to constant-temperature simulations, T-REMD simulations provide a new, faster pathway for peptide bond isomerization at ambient temperature via heating, isomerization, and subsequent cooling. In this way, even Xaa−non-Pro *cis* peptide bonds introduced at high temperatures in T-REMD simulations are integrated into the 300 K ensemble within a few hundred nanoseconds per replica.

Our AMBER99SB-ILDN T-REMD simulations indicate that *cis* G11-P12 and R16-P17 peptide bonds may be present in the folded state of Trp-cage TC5b, with a cumulative population of 0 to 3.6%. It remains to be shown whether these *cis* peptide bonds are accurately represented by the current generation of force fields and therefore whether these structural interconversions are merely an inconvenient sampling problem or whether they represent an underappreciated source of simulation error. As simulation time scales continue to increase, peptide bond isomerization and chiral inversion will become accessible at ever lower temperatures, making the treatment of these hitherto seemingly irrelevant degrees of freedom increasingly important. Moreover, the peptide bond isomerization and chiral inversions that we identify in standard high-temperature and T-REMD simulations are equally relevant to other high-

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

temperature methods such as simulated annealing and simulated tempering and may complicate the construction of Markov state models for protein folding.

We conclude with the following recommendations. First, all high-temperature simulations should be assessed for chiral inversion and peptide bond isomerization. Second, to avoid chiral inversion the OPLS-AA/L force field should not be used at $T > 700$ K. Third, for simulation systems that are unlikely to attain converged estimates of *cis*/*trans* ratios on accessible time scales, one may choose to explicitly limit sampling to the all-*trans* regime by using modified potentials that specifically inhibit peptide bond isomerization. Alternatively, the chance of peptide bond isomerization can be reduced by limiting maximum temperatures to <400 K, < 500 K, < 600 K, and <600 K for the CHARMM22/CMAP, AMBER99SB-ILDN, Doshi, and OPLS-AA/L force fields, respectively. Nevertheless, peptide bond isomerization is real,[8,9,11−14,16−27] and future work should be directed toward evaluating the consistency between estimates of *cis*/*trans* populations and isomerization rates in experiment and simulation. One fruitful avenue of future research involves the evaluation of different simulation force fields for their ability to reproduce the sequence and temperature dependence of *cis* peptide bonds in the experimental work of Scherer et al.[8]

METHODS

Molecular dynamics simulations are conducted with a single-precision compilation of versions 4.6.7 and 4.6.3 of the GROMACS simulation package[84] for constant-temperature and replica exchange simulations, respectively, as outlined below. The integration time step is 2 fs, except as noted. The nonbonded pairlist is updated every 10 timesteps. Water molecules are rigidified with SETTLE,[85] and bond lengths in peptides are constrained with P-LINCS[86] using sixth-order coupling and a single iteration. Electrostatics are computed with the smooth particle-mesh Ewald (PME) method[87,88] with a Fourier grid spacing of 0.12 nm. Temperature is controlled using velocity Langevin dynamics[89] with a coupling constant of 1 ps. Equilibration in the *NpT* ensemble is achieved by isotropic coupling to a Berendsen barostat[90] at 1 bar with a coupling constant of 4 ps.

For Amber simulations, peptides are modeled by the AMBER99SB-ILDN protein force field.[60] The water model is TIP3P.[91] Lennard-Jones (LJ) interactions are evaluated using a group-based cutoff, truncating interactions at 1 nm without a smoothing function. Dispersion corrections[92] are applied. Simulations of the Amber force field with the peptide bond torsion potential modifications of Doshi and Hamelberg[61] are conducted similarly.

For CHARMM simulations, peptides are modeled by the CHARMM 22/27 protein force field[62] with grid-based energy correction maps.[63] The water model is TIP3P[91] with CHARMM modifications.[62] LJ interactions are evaluated using an atom-based cutoff and a long-range neighbor-list constructed to 1.3 nm, gradually switching off the potential energy of interactions between 0.8 and 1.2 nm.

For OPLS simulations, peptides are modeled by the OPLS-AA/L protein force field.[64] The water model is TIP4P.[91] LJ interactions are evaluated using a group-based cutoff, truncating interactions at 1 nm without a smoothing function.

Octapeptides. We simulate three eight-residue peptides with zwitterionic termini: MTYKLILN, TAEKVFKQ, and DGGPSSGR. The first two of these sequences are based on fragments of *Streptococcus sp.* protein G that have been simulated by Ho and Dill.[93] Specifically, MTYKLILN represents the first 8 residues of a construct of protein *G*'s immunoglobulin binding domain, which is structurally characterized in PDB 2GB1.[94] The sequence of this octapeptide corresponds to protein G residues 227−234 in UniProt ID P06654, except that protein G residue 227 is glutamic acid, not methionine. We use the sequence simulated by Ho and Dill,[93] with an N-terminal methionine, because we are interested in evaluating *cis*−*trans* isomerization in peptides that have already been simulated using T-REMD and we are not in this instance interested in the folding properties of protein G, *per se*. The second protein G-derived peptide that we simulate is TAEKVFKQ, representing protein G residues 251−258. Finally, we simulate an eight residue fragment of the TC5b Trp-cage miniprotein, DGGPSSGR, selected because it contains a single proline residue.

An all-trans conformation of each peptide is generated with the Molefacture plugin to VMD 1.9.2.[83] Each peptide is embedded in a rhombic dodecahedral unit cell of volume 46.3 nm^3, hydrated with ∼1480 water molecules, neutralized with a single Cl^- ion, and energy minimized. Each peptide system is then simulated 20 times for 2 ns in the *NpT* ensemble. In each case, the snapshot from the last 1 ns whose system volume most closely matches the average volume over the 1 ns of simulation is selected for further simulation. In each case, systems are simulated for 100 ns at 400, 500, 600, 700, or 800 K. In this way, we conduct a total of one hundred 100-ns simulations of each peptide, with 20 independent repeats at each of five temperatures.

Trp-Cage TC5b. The Trp-cage simulation system consists of terminally capped TC5b (acetyl-NLYIQWLKDGGPSSG-RPPPS-methylamide), a single Cl^- ion, and 2487 water molecules in a rhombic dodecahedal unit cell. The initial structure is model 1 from PDB 1L2Y,[67] to which terminal caps are added with PyMOL.[95] This folded structure is hydrated, energy minimized, and simulated for 200 ps in the *NpT* ensemble at 300 K. To generate an unfolded protein structure, the final frame from this *NpT* simulation is simulated for 10 ns in the *NVT* ensemble at 700 K. Although *cis* peptide bonds occur transiently in this 10 ns simulation (P13-S14, R17-P18, and P18-P19), none are present in the structure that we use to initiate T-REMD simulations, whose peptide bonds are all within 25° of perfectly *trans* dihedral angles. There are no chiral inversions in this high temperature *NVT* simulation. The protein's starting structure for T-REMD simulation is depicted in Figure S6.

Three independent T-REMD simulations of TC5b are conducted in the *NVT* ensemble using the AMBER99SB-ILDN and TIP3P parameters. Simulations use 32 replicas exponentially spaced between 300 and 540 K (300, 305.743, 311.595, 317.559, 323.638, 329.833, 336.147, 342.581, 349.139, 355.822, 362.633, 369.574, 376.649, 383.858, 391.206, 398.694, 406.326, 414.104, 422.031, 430.109, 438.342, 446.733, 455.284, 463.999, 472.881, 481.932, 491.157, 500.559, 510.141, 519.906, 529.858, and 540 K), selected using the approach of Okamoto et al.[96]

$$T_i = T_0 \times (T_{max}/T_0)^{(i-1)/(N-1)} \quad (1)$$

where the initial temperature of the i^{th} of $N = 32$ replicas, T_i, varies from $T_0 = 300$ K to $T_{max} = 540$ K, for i from 1 to N. Replica exchanges are attempted every 0.4 ps, and snapshots are

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

saved every 10 ps. Each simulation is conducted for 500 ns/replica.

Data Analysis. Peptide bond dihedral angles, ω, between backbone $C_{\alpha i-1}$, C_{i-1}, N_i, and $C_{\alpha i}$ atoms of residues $i-1$ and i are computed with the GROMACS 4.6.7 analysis tool g_chi. The peptide bond's *cis* and *trans* conformational basins are taken as $-60° < \omega < 60°$ and $120° < \omega < -120°$, respectively (with periodicity of 360°). To avoid spurious definition of peptide bond isomerization, conformations with $60° \geq \omega \geq 120°$ and $-120° \geq \omega \geq -60°$ are defined as *cis* or *trans* based on the previous value in the time trajectory. For example, a time series of ω = 0°, 90°, 180°, 90° is defined as *cis*, *cis*, *trans*, *trans*.

For constant-temperature simulations, data from all three peptide sequences are grouped together such that, for each combination of force field and temperature, B peptide bonds are sampled in R repeat simulations (B = 20 for Xaa−non-Pro and B = 1 for Xaa−Pro; R = 20 for all cases). The average frequency for *trans* → *cis* isomerization of ω, $f_\omega^{t\to c}$ is computed according to

$$f_\omega^{t\to c} = \frac{\sum_{b=1}^{B}\sum_{r=1}^{R} N(\omega_{trans\to cis}^{b,r})}{\sum_{b=1}^{B}\sum_{r=1}^{R} N^{\dagger}(\omega_{trans}^{b,r})} \quad (2)$$

where $N(\omega_{trans\to cis}^{b,r})$ and $N(\omega_{trans}^{b,r})$ are the number of *trans* → *cis* isomerization events and the number of snapshots in which the peptide bond is defined as *trans*, respectively, for the b^{th} peptide bond in the r^{th} repeat simulation and $N^{\dagger}(\omega_{trans}^{b,r})$ is the value of $N(\omega_{trans}^{b,r})$ excluding the final frame of the trajectory, which cannot contribute to the numerator. Bootstrap resampling[97] is used to estimate the error in $f_\omega^{t\to c}$ by resampling the trajectories, with replacement, to obtain a value of $^*f_\omega^{t\to c}$ analogously to eq 2 except that each paired value of b,r is replaced by a randomly selected pair of b^*,r^* whose values are the same in the numerator and the denominator. The estimated error is then the standard deviation among 10,000 evaluations of $^*f_\omega^{t\to c}$ conducted with different random seeds. The numerator and denominator are summed separately to avoid introducing large error estimates when some trajectories exhibit rapid, stable isomerization and hence small values of $N^{\dagger}(\omega_{trans}^{b,r})$. The average frequency for *cis* → *trans* isomerization of ω, $f_\omega^{c\to t}$, is computed analogously to eq 2, and the *cis* peptide bond percentage, $\%_\omega^{cis}$, is computed according to

$$\%_\omega^{cis} = \frac{\sum_{b=1}^{B}\sum_{r=1}^{R} N(\omega_{cis}^{b,r})}{\sum_{b=1}^{B}\sum_{r=1}^{R} N(\omega_{cis}^{b,r}) + N(\omega_{trans}^{b,r})} \times 100\% \quad (3)$$

Chiral inversion is assessed with the chirality plugin[59] to VMD 1.9.2,[83] and chirality is quantified for Figure S5 with the GROMACS 4.6.7 analysis tool g_angle.

The folded state of the TC5b Trp-cage miniprotein is defined based on an NMR structure (PDB 1L2Y model 1).[67] In simulations, the folded basin is defined as C_α RMSD to folded ≤0.22 nm, as in our previous work with Trp-cage proteins.[65,69,98]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b01022.

Tables S1−S6, peptide bond isomerization in standard simulations; Tables S7 and S8, *cis* peptide bonds in T-REMD simulations; Figures S1−S3, pre-residue percent *cis* peptide bonds in standard simulations; Figure S4, comparison of Xaa−non-Pro *cis*/*trans* peptide bond isomerization; Figure S5, chiral inversion at the C_α atom of V5; Figure S6, unfolded initial structure for T-REMD simulations of Trp-cage TC5b; Figure S7, protein folding in three T-REMD simulations of Trp-cage TC5b; Figure S8, compaction of Trp-cage TC5b; Figure S9, replica mobility and conformational exchange in three T-REMD simulations of Trp-cage TC5b; Figure S10, temperature dependence of *cis* peptide bonds in T-REMD simulations of Trp-cage TC5b (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 1-505-667-3883. E-mail: agarcia@lanl.gov.

Author Contributions

C.N., R.P., and A.E.G. designed the research, C.N. performed the research, C.N. analyzed the data, and C.N. and A.E.G. wrote the paper.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Computations are performed at (i) Colosse at the CLUMEQ high performance computing (HPC) consortium of Calcul Québec (www.calculquebec.ca), (ii) GPC at the SciNet HPC consortium,[99] (iii) the Center for Computational Innovations at Rensselaer Polytechnic Institute, and (iv) Stampede at the Texas Advanced Computing Center at the University of Texas at Austin (www.tacc.utexas.edu), to which access is provided by the Extreme Science and Engineering Discovery Environment (XSEDE grant TG-MCB130178). CLUMEQ is funded by the Canada Foundation for Innovation (CFI), the Natural Sciences and Engineering Research Council of Canada (NSERC), and Fonds de Recherche Nature et Technologies Québec. SciNet is a resource of Compute Canada (www.computecanada.ca) and is funded by the CFI under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto. XSEDE is supported by National Science Foundation (NSF) grant number ACI-1053575. C.N. is funded by a postdoctoral fellowship from the Canadian Institutes of Health Research (CIHR). R.P. is funded by CIHR Operating Grant MOP-43998 and NSERC Discovery grant 418679. This work is funded in part by NSF grant MCB-1050966.

■ REFERENCES

(1) Fischer, E. Üeber die Hydrolyse der Proteinstoffe. *Chem.-Ztg.* **1902**, *26*, 939−940.

(2) Hofmeister, F. Über Bau und Gruppierung der Eiweisskörper. *Rev. Physiol., Biochem. Pharmacol.* **1902**, *1*, 759−802.

(3) Pauling, L.; Corey, R. B.; Branson, H. R. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 205−211.

(4) Yonezawa, Y.; Nakata, K.; Sakakura, K.; Takada, T.; Nakamura, H. Intra- and intermolecular interaction inducing pyramidalization on both sides of a proline dipeptide during isomerization: an ab initio QM/MM molecular dynamics simulation study in explicit water. *J. Am. Chem. Soc.* **2009**, *131*, 4535−4540.

(5) Fischer, S.; Dunbrack, R. L.; Karplus, M. Cis-trans imide isomerization of the proline dipeptide. *J. Am. Chem. Soc.* **1994**, *116*, 11931−11937.

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

(6) Lauvergnat, D.; Hiberty, P. C. Role of conjugation in the stabilities and rotational barriers of formamide and thioformamide. An ab initio valence-bond study. *J. Am. Chem. Soc.* **1997**, *119*, 9478−9482.
(7) Hinderaker, M. P.; Raines, R. T. An electronic effect on protein structure. *Protein Sci.* **2003**, *12*, 1188−1194.
(8) Scherer, G.; Kramer, M. L.; Schutkowski, M.; Reimer, U.; Fischer, G. Barriers to rotation of secondary amide peptide bonds. *J. Am. Chem. Soc.* **1998**, *120*, 5568−5574.
(9) Grathwohl, C.; Wüthrich, K. NMR studies of the rates of proline cis−trans isomerization in oligopeptides. *Biopolymers* **1981**, *20*, 2623−2633.
(10) Mathieu, S.; Poteau, R.; Trinquier, G. Estimating the "steric clash" at cis peptide bonds. *J. Phys. Chem. B* **2008**, *112*, 7894−7902.
(11) Lorenzen, S.; Peters, B.; Goede, A.; Preissner, R.; Frömmel, C. Conservation of cis prolyl bonds in proteins during evolution. *Proteins: Struct., Funct., Genet.* **2005**, *58*, 589−595.
(12) Pal, D.; Chakrabarti, P. Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations. *J. Mol. Biol.* **1999**, *294*, 271−288.
(13) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. Non-proline Cis peptide bonds in proteins. *J. Mol. Biol.* **1999**, *286*, 291−304.
(14) Stewart, D. E.; Sarkar, A.; Wampler, J. E. Occurrence and role of cis peptide bonds in protein structures. *J. Mol. Biol.* **1990**, *214*, 253−260.
(15) Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J. Mol. Biol.* **1998**, *279*, 449−460.
(16) Craveur, P.; Joseph, A.; Poulain, P.; de Brevern, A.; Rebehmed, J. Cis−trans isomerization of omega dihedrals in proteins. *Amino Acids* **2013**, *45*, 279−289.
(17) Joseph, A. P.; Srinivasan, N.; De Brevern, A. G. Cis-trans peptide variations in structurally similar proteins. *Amino Acids* **2012**, *43*, 1369−1381.
(18) Yaffe, M. B.; Schutkowski, M.; Shen, M.; Zhou, X. Z.; Stukenberg, P. T.; Rahfeld, J.-U.; Xu, J.; Kuang, J.; Kirschner, M. W.; Fischer, G.; Cantley, L. C.; Lu, K. P. Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. *Science* **1997**, *278*, 1957−1960.
(19) Follis, A. V.; Llambi, F.; Merritt, P.; Chipuk, J. E.; Green, D. R.; Kriwacki, R. W. Pin1-induced proline isomerization in cytosolic p53 mediates BAX activation and apoptosis. *Mol. Cell* **2015**, *59*, 677−684.
(20) Hamelberg, D.; Shen, T.; McCammon, J. A. Phosphorylation effects on cis/trans isomerization and the backbone conformation of serine−proline motifs: accelerated molecular dynamics analysis. *J. Am. Chem. Soc.* **2005**, *127*, 1969−1974.
(21) Lu, K. P.; Finn, G.; Lee, T. H.; Nicholson, L. K. Prolyl cis-trans isomerization as a molecular timer. *Nat. Chem. Biol.* **2007**, *3*, 619−629.
(22) Lummis, S. C. R.; Beene, D. L.; Lee, L. W.; Lester, H. A.; Broadhurst, R. W.; Dougherty, D. A. Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature* **2005**, *438*, 248−252.
(23) Bächinger, H. P. The influence of peptidyl-prolyl cis-trans isomerase on the in vitro folding of type III collagen. *J. Biol. Chem.* **1987**, *262*, 17144−17148.
(24) Lang, K.; Schmid, F. X.; Fischer, G. Catalysis of protein folding by prolyl isomerase. *Nature* **1987**, *329*, 268−270.
(25) Wedemeyer, W. J.; Welker, E.; Scheraga, H. A. Proline cis−trans iomerization and protein folding. *Biochemistry* **2002**, *41*, 14637−14644.
(26) Pappenberger, G.; Aygun, H.; Engels, J. W.; Reimer, U.; Fischer, G.; Kiefhaber, T. Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics. *Nat. Struct. Biol.* **2001**, *8*, 452−458.
(27) Schmid, F. X.; Grafl, R.; Wrba, A.; Beintema, J. J. Role of proline peptide bond isomerization in unfolding and refolding of ribonuclease. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 872−876.
(28) Guan, R.-J.; Xiang, Y.; He, X.-L.; Wang, C.-G.; Wang, M.; Zhang, Y.; Sundberg, E. J.; Wang, D.-C. Structural mechanism governing cis and trans isomeric states and an intramolecular switch for cis/trans isomerization of a non-proline peptide bond observed in crystal structures of scorpion toxins. *J. Mol. Biol.* **2004**, *341*, 1189−1204.
(29) Li, P.; Chen, X. G.; Shulin, E.; Asher, S. A. UV resonance Raman ground and excited state studies of amide and peptide isomerization dynamics. *J. Am. Chem. Soc.* **1997**, *119*, 1116−1120.
(30) Xia, J.; Levy, R. M. Molecular dynamics of the proline switch and its role in Crk signaling. *J. Phys. Chem. B* **2014**, *118*, 4535−4545.
(31) Hamelberg, D.; McCammon, J. A. Fast peptidyl cis−trans isomerization within the flexible gly-rich flaps of HIV-1 protease. *J. Am. Chem. Soc.* **2005**, *127*, 13778−13779.
(32) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646−652.
(33) Baker, C. M.; Best, R. B. Insights into the binding of intrinsically disordered proteins from molecular dynamics simulation. *WIREs Comput. Mol. Sci.* **2014**, *4*, 182−198.
(34) Perilla, J. R.; Goh, B. C.; Cassidy, C. K.; Liu, B.; Bernardi, R. C.; Rudack, T.; Yu, H.; Wu, Z.; Schulten, K. Molecular dynamics simulations of large macromolecular complexes. *Curr. Opin. Struct. Biol.* **2015**, *31*, 64−74.
(35) Šponer, J.; Banáš, P.; Jurečka, P.; Zgarbová, M.; Kührová, P.; Havrila, M.; Krepl, M.; Stadlbauer, P.; Otyepka, M. Molecular dynamics simulations of nucleic acids. from tetranucleotides to the ribosome. *J. Phys. Chem. Lett.* **2014**, *5*, 1771−1782.
(36) Roberts, C. J.; Debenedetti, P. G. Structure and dynamics in concentrated, amorphous carbohydrate−water systems by molecular dynamics simulation. *J. Phys. Chem. B* **1999**, *103*, 7308−7318.
(37) Huynh, L.; Neale, C.; Pomès, R.; Allen, C. Computational approaches to the rational design of nanoemulsions, polymeric micelles, and dendrimers for drug delivery. *Nanomedicine* **2012**, *8*, 20−36.
(38) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How fast-folding proteins fold. *Science* **2011**, *334*, 517−520.
(39) Lane, T. J.; Shukla, D.; Beauchamp, K. A.; Pande, V. S. To milliseconds and beyond: challenges in the simulation of protein folding. *Curr. Opin. Struct. Biol.* **2013**, *23*, 58−65.
(40) Contreras, F. X.; Sánchez-Magraner, L.; Alonso, A.; Goñi, F. M. Transbilayer (flip-flop) lipid motion and lipid scrambling in membranes. *FEBS Lett.* **2010**, *584*, 1779−1786.
(41) Zarrinkar, P. P.; Williamson, J. R. Williamson Kinetic intermediates in RNA folding. *Science* **1994**, *265*, 918−924.
(42) Thirumalai, D.; Woodson, S. A. Maximizing RNA folding rates: a balancing act. *RNA* **2000**, *6*, 790−794.
(43) Jackson, S. E. How do small single-domain proteins fold? *Folding Des.* **1998**, *3*, R81−R91.
(44) Schmid, F. X. Prolyl Isomerase: Enzymatic Catalysis of Slow Protein-Folding Reactions. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 123−143.
(45) Zwier, M. C.; Chong, L. T. Reaching biological timescales with all-atom molecular dynamics simulations. *Curr. Opin. Pharmacol.* **2010**, *10*, 745−752.
(46) Schlick, T. Molecular dynamics-based approaches for enhanced sampling of long-time, large-scale conformational changes in biomolecules. *F1000 Biol. Rep.* **2009**, *1*, 51.
(47) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(48) Periole, X.; Mark, A. E. Convergence and sampling efficiency in replica exchange simulations of peptide folding in explicit solvent. *J. Chem. Phys.* **2007**, *126*, 014903.
(49) Buck, M.; Bouguet-Bonnet, S.; Pastor, R. W.; MacKerell, A. D., Jr Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophys. J.* **2006**, *90*, L36−L38.
(50) Ponder, J. W.; Case, D. A.; Valerie, D. Force Fields for Protein Simulations. *Adv. Protein Chem.* **2003**, *66*, 27−85.
(51) Ruscio, J. Z.; Fawzi, N. L.; Head-Gordon, T. How hot? Systematic convergence of the replica exchange method using multiple reservoirs. *J. Comput. Chem.* **2010**, *31*, 620−627.
(52) Rosta, E.; Hummer, G. Error and efficiency of replica exchange molecular dynamics simulations. *J. Chem. Phys.* **2009**, *131*, 165102.

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

(53) Abraham, M. J.; Gready, J. E. Ensuring mixing efficiency of replica-exchange molecular dynamics simulations. *J. Chem. Theory Comput.* **2008**, *4*, 1119−1128.
(54) Bruccoleri, R. E.; Karplus, M. Conformational sampling using high-temperature molecular dynamics. *Biopolymers* **1990**, *29*, 1847−1862.
(55) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542−1555.
(56) Kutzner, C.; Páll, S.; Fechner, M.; Esztermann, A.; de Groot, B. L.; Grubmüller, H. Best bang for your buck: GPU nodes for GROMACS biomolecular simulations. *J. Comput. Chem.* **2015**, *36*, 1990−2008.
(57) Shaw, D. E.; Grossman, J. P.; Bank, J. A.; Batson, B.; Butts, J. A.; Chao, J. C.; Deneroff, M. M.; Dror, R. O.; Even, A.; Fenton, C. H.; Forte, A.; Gagliardo, J.; Gill, G.; Greskamp, B.; Ho, C. R.; Ierardi, D. J.; Iserovich, L.; Kuskin, J. S.; Larson, R. H.; Layman, T.; Lee, L.-S.; Lerer, A. K.; Li, C.; Killebrew, D.; Mackenzie, K. M.; Mok, S.Y.-H.; Moraes, M. A.; Mueller, R.; Nociolo, L. J.; Peticolas, J. L.; Quan, T.; Ramot, D.; Salmon, J. K.; Scarpazza, D. P.; Schafer, U. B.; Siddique, N.; Snyder, C. W.; Spengler, J.; Tang, P. T. P.; Theobald, M.; Toma, H.; Towles, B.; Vitale, B.; Wang, S. C.; Young, C. Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*; 2014; pp 41−53.
(58) Shirts, M.; Pande, V. S. Screen savers of the world unite! *Science* **2000**, *290*, 1903−1904.
(59) Schreiner, E.; Trabuco, L. G.; Freddolino, P. L.; Schulten, K. Stereochemical errors and their implications for molecular dynamics simulations. *BMC Bioinf.* **2011**, *12*, 190.
(60) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.
(61) Doshi, U.; Hamelberg, D. Reoptimization of the AMBER force field parameters for peptide bond (omega) torsions using accelerated molecular dynamics. *J. Phys. Chem. B* **2009**, *113*, 16590−16595.
(62) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(63) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(64) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(65) English, C. A.; García, A. E. Charged termini on the Trp-cage roughen the folding energy landscape. *J. Phys. Chem. B* **2015**, *119*, 7874−7881.
(66) Flyvbjerg, H.; Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **1989**, *91*, 461−466.
(67) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Designing a 20-residue protein. *Nat. Struct. Biol.* **2002**, *9*, 425−430.
(68) de Bakker, P. I. W.; Hünenberger, P. H.; McCammon, J. A. Molecular dynamics simulations of the hyperthermophilic protein sac7d from Sulfolobus acidocaldarius: contribution of salt bridges to thermostability. *J. Mol. Biol.* **1999**, *285*, 1811−1830.
(69) English, C. A.; Garcia, A. E. Folding and unfolding thermodynamics of the TC10b Trp-cage miniprotein. *Phys. Chem. Chem. Phys.* **2014**, *16*, 2748−2757.
(70) Baumketner, A.; Shea, J.-E. The structure of the Alzheimer amyloid β 10−35 peptide probed through replica-exchange molecular dynamics simulations in explicit solvent. *J. Mol. Biol.* **2007**, *366*, 275−285.
(71) Bu, L.; Im, W.; Brooks, C. L., III Membrane assembly of simple helix homo-oligomers studied via molecular dynamics simulations. *Biophys. J.* **2007**, *92*, 854−863.
(72) Felts, A. K.; Harano, Y.; Gallicchio, E.; Levy, R. M. Free energy surfaces of β-hairpin and α-helical peptides generated by replica exchange molecular dynamics with the AGBNP implicit solvent model. *Proteins: Struct., Funct., Genet.* **2004**, *56*, 310−321.
(73) Pande, V. S.; Rokhsar, D. S. Molecular dynamics simulations of unfolding and refolding of a β-hairpin fragment of protein G. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9062−9067.
(74) Kumar, N.; Kishore, R. Determination of an unusual secondary structural element in the immunostimulating tetrapeptide rigin in aqueous environments: insights via MD simulations, 1H NMR and CD spectroscopic studies. *J. Pept. Sci.* **2010**, *16*, 456−464.
(75) Im, W.; Brooks, C. L. Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6771−6776.
(76) Duan, Y.; Kollman, P. A. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science.* **1998**, *282*, 740−744.
(77) Leitgeb, B. Characteristic structural features of indolicidin: effects of the cis-trans isomerism on its conformation. *Chem. Biol. Drug Des.* **2014**, *83*, 132−140.
(78) Rovó, P.; Stráner, P.; Láng, A.; Bartha, I.; Huszár, K.; Nyitray, L.; Perczel, A. Structural insights into the Trp-cage folding intermediate formation. *Chem. - Eur. J.* **2013**, *19*, 2628−2640.
(79) Touw, W. G.; Joosten, R. P.; Vriend, G. Detection of trans-cis flips and peptide-plane flips in protein structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *71*, 1604−1614.
(80) Helfman, P. M.; Bada, J. L. Aspartic acid racemization in tooth enamel from living humans. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 2891−2894.
(81) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J. *AMBER 9*; University of California: San Francisco, 2006.
(82) Okur, A.; Miller, B. T.; Joo, K.; Lee, J.; Brooks, B. R. Generating reservoir conformations for replica exchange through the use of the conformational space annealing method. *J. Chem. Theory Comput.* **2013**, *9*, 1115−1124.
(83) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(84) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(85) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952−962.
(86) Hess, B. P-LINCS: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116−122.
(87) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N\bullet\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(88) 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*, 8577−8593.
(89) van Gunsteren, W. F.; Berendsen, H. J. C. A leap-frog algorithm for stochastic dynamics. *Mol. Simul.* **1988**, *1*, 173−185.
(90) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999

(91) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(92) Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Pande, V. S. Accurate and efficient corrections for missing dispersion interactions in molecular simulations. *J. Phys. Chem. B* **2007**, *111*, 13052−13063.
(93) Ho, B. K.; Dill, K. A. Folding very short peptides using molecular dynamics. *PLoS Comput. Biol.* **2006**, *2*, e27.
(94) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* **1991**, *253*, 657−661.
(95) DeLano, W. L. *The PyMOL molecular graphics system*; DeLano Scientific: San Carlos, CA, 2002.
(96) Okamoto, Y.; Fukugita, M.; Nakazawa, T.; Kawai, H. Alpha-helix folding by Monte Carlo simulated annealing in isolated C-peptide of ribonuclease A. *Protein Eng., Des. Sel.* **1991**, *4*, 639−647.
(97) Efron, B. Bootstrap Methods: Another Look at the Jackknife. *Ann. Statist.* **1979**, *7*, 1−26.
(98) Day, R.; Paschek, D.; Garcia, A. E. Microsecond simulations of the folding/unfolding thermodynamics of the Trp-cage miniprotein. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1889−1899.
(99) Loken, C.; Gruner, D.; Groer, L.; Peltier, R.; Bunn, N.; Craig, M.; Henriques, T.; Dempsey, J.; Yu, C.-H.; Chen, J.; Dursi, L. J.; Chong, J.; Northrup, S.; Pinto, J.; Knecht, N.; Van Zon, R. SciNet: lessons learned from building a power-efficient top-20 system and data centre. *J. Phys. Conf. Ser.* **2010**, *256*, 012026.

DOI: 10.1021/acs.jctc.5b01022
J. Chem. Theory Comput. 2016, 12, 1989−1999