JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Comparison of Secondary Structure Formation Using 10 Different Force Fields in Microsecond Molecular Dynamics Simulations

Elio A. Cino,[†] Wing-Yiu Choy,*[,†] and Mikko Karttunen*[,‡]

[†]Department of Biochemistry, 1151 Richmond Street North, The University of Western Ontario, London, Ontario N6A 5C1, Canada

[‡]Department of Chemistry, 200 University Avenue West, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

ABSTRACT: We have compared molecular dynamics (MD) simulations of a β-hairpin forming peptide derived from the protein Nrf2 with 10 biomolecular force fields using trajectories of at least 1 μs. The total simulation time was 37.2 μs. Previous studies have shown that different force fields, water models, simulation methods, and parameters can affect simulation outcomes. The MD simulations were done in explicit solvent with a 16-mer Nrf2 β-hairpin forming peptide using Amber ff99SB-ILDN, Amber ff99SB*-ILDN, Amber ff99SB, Amber ff99SB*, Amber ff03, Amber ff03*, GROMOS96 43a1p, GROMOS96 53a6, CHARMM27, and OPLS-AA/L force fields. The effects of charge-groups, terminal capping, and phosphorylation on the peptide folding were also examined. Despite using identical starting structures and simulation parameters, we observed clear differences among the various force fields and even between replicates using the same force field. Our simulations show that the uncapped peptide folds into a native-like β-hairpin structure at 310 K when Amber ff99SB-ILDN, Amber ff99SB*-ILDN, Amber ff99SB, Amber ff99SB*, Amber ff03, Amber ff03*, GROMOS96 43a1p, or GROMOS96 53a6 were used. The CHARMM27 simulations were able to form native hairpins in some of the elevated temperature simulations, while the OPLS-AA/L simulations did not yield native hairpin structures at any temperatures tested. Simulations that used charge-groups or peptide capping groups were not largely different from their uncapped counterparts with single atom charge-groups. On the other hand, phosphorylation of the threonine residue located at the β-turn significantly affected the hairpin formation. To our knowledge, this is the first study comparing such a large set of force fields with respect to β-hairpin folding. Such a comprehensive comparison will offer useful guidance to others conducting similar types of simulations.

INTRODUCTION

Atomistic molecular dynamics (MD) simulations are a versatile tool for studying protein folding and function. They can provide detailed atomistic information, which may be difficult to obtain by experimental techniques. Increases in computational power have allowed for simulations to reach experimentally relevant time scales at the microsecond level: MD simulations have been used to study the folding of peptides and small proteins[1−9] and to model other biological systems. The current record for an atomistic simulation of protein conformational changes, as far as we know, is 1 ms reached by Shaw et al.[7] for the 58-residue protein BPTI.

One of the major challenges in protein folding simulations is choosing an appropriate force field; see, for example, ref 10. This is due to possible biases different force fields have toward certain types of secondary structure.[3,11−14] Ideally, the force field should be fully validated with experimental data, but that is typically not possible as it would involve validation against different structures and other physical properties from a large number of independent and fully validated experiments, mission impossible because experiments have their own error sources due to, for example, instrumentation. While a completely transferable force field does not exist, modifications of existing force fields have led to improvements in agreement with experimental data.[15−22]

In this work, we compared 10 biomolecular force fields with respect to the folding of a peptide derived from Nuclear factor erythroid 2-related factor 2 (Nrf2). Nrf2 is an important transcription factor that regulates the expression of genes responsive to oxidative stress.[23,24] The protein consists of six highly homologous regions (Neh1−6 domains). Structural analysis showed that the N-terminal Neh2 domain is intrinsically disordered, a novel class of proteins that are extremely dynamic in nature.[25−29] Under homeostatic conditions, this domain binds two Kelch units of a Keap1 dimer through two separate motifs: a high affinity "ETGE" motif and a lower affinity "DLG" motif.[30] Crystallographic data have shown that the "ETGE" motif and its surrounding residues (residues 75−83) form a β-hairpin structure upon binding to the Kelch domain of Keap1 (PDB ids: 2FLU and 1X2R).[30,31] NMR-derived ^{1}H, ^{1}H NOEs suggest that residual structure spanning from residues 74−85, likely in the form of a β-hairpin, also exists in the free-state of Neh2.[30,31] Other experimental data have shown that a peptide containing residues 74−87 can compete with the full-length Nrf2 for binding Keap1.[31] Here, we chose to use a 16-mer human Nrf2 peptide with the sequence ^{72}AQLQLDEETGEFLPIQ^{87} for our MD simulations. This peptide contains the "ETGE" motif and should be of sufficient length to form the necessary interactions to stabilize the β-hairpin structure. It is noteworthy that the phosphorylation of Thr-80 has been shown to impair the binding to Keap1.[31] Because Neh2 is largely disordered and lacks a tertiary structure, this β-hairpin likely folds independently, making it a good target for folding simulations.[30]

Received: April 24, 2012
Published: June 19, 2012

   dx.doi.org/10.1021/ct300323g | J. Chem. Theory Comput. 2012, 8, 2725−2740

In addition to Nrf2, several other proteins that contain "ETGE"-like motifs have been shown to interact with the Kelch domain of Keap1. These include PGAM5,[32] FAC1,[33] PTMA,[34] p62,[35] WTX,[36] and PALB2.[37] Some of these Keap1 interacting proteins have only been recently discovered, which suggests that this list of targets may still be growing. Structures of PTMA (Prothymosin alpha) and p62 peptides in complex with Keap1 indicate that their "ETGE"-like motifs bind to the same region as the "ETGE" motif of Nrf2 and form similar hairpin structures in their bound states.[31,34,35] Interestingly, MD simulations from our previous work showed that the binding motifs of Nrf2 and PTMA have a tendency to form hairpin structures that resembled the bound state conformation even in the absence of Keap1.[9] With the list of Keap1 binding partners seemingly expanding and MD simulations becoming an increasingly important and predictive tool, it is important to establish appropriate simulation protocols for these systems, including force field choice.

β-Hairpins are a type of protein supersecondary structure consisting of two hydrogen-bonded antiparallel β-strands connected by a turn. These structural elements are common in globular proteins because they reverse the direction of the protein backbone, allowing the formation of compact structures. β-Hairpin motifs are sometimes involved in protein−protein interactions, and it has been suggested that they can act as nucleation sites for protein folding.[31,38−40] In this study, we compared folding simulations of a β-hairpin peptide derived from the intrinsically disordered Neh2 domain. Starting from an unfolded structure, we performed extensive (1−2 μs each, totaling 37 μs) atomistic molecular dynamics simulations using 10 different force fields (details in next section). We selected these force fields primarily because they are commonly used in biomolecular simulations, including those of β-hairpin folding.[3,9,41,42]

Force field selection is a key factor in the outcome of protein folding simulations. Although force field modifications have led to improved agreements between MD simulations and experimental data, continued testing and comparison with experimental data is required to further these advances. While studies comparing different force fields have been conducted previously, very few of them had included such a large set of force fields with respect to folding of secondary structure elements.[3,14,19,43−45] Small proteins and peptides with folding times on the microsecond time scale are excellent systems to test and compare force fields; such trajectories provide reasonable sampling of conformations and sufficient length to examine the stability of the force field.

In this work, we compare MD simulations of a β-hairpin forming peptide derived from the protein Nrf2, performed with 10 force fields. We assess their agreements with experimental data. The effects of elevated temperatures, charge-groups,[46,47] peptide capping, and phosphorylation of Thr-80 with respect to β-hairpin formation were also examined. Despite using identical starting structures and simulation parameters, we observed clear differences among the various force fields and even between replicate simulations using the same force field. Such a comprehensive comparison will offer useful guidance to others conducting similar types of simulations and for improving force fields.

SIMULATION METHODOLOGY

Starting Structures. The starting structure for our MD simulations was generated on the basis of the amino acid sequence of a 16-mer human Nrf2 peptide (^{72}AQLQLDEETGEFLPIQ^{87}). We used the Crystallography and NMR System (CNS)[48] to generate an extended structure, which subsequently underwent simulated annealing. To avoid any potential bias to the bound-state conformation, a structure from the annealing simulations that did not resemble the bound state structure (PDB id: 2FLU) was chosen as the starting structure.[31] The exact same starting structure was used for all simulations. For the phosphorylated peptide (pThr-80) simulations, a dianionic phosphate group (PO_4^{2-}) was modeled onto residue Thr-80 of the same structure using chimera.[49]

Force Fields. We compared the peptide folding using the following force fields: Amber ff99SB-ILDN,[15,19,20] Amber ff99SB*-ILDN,[15,17,19,20] Amber ff99SB,[15,19] Amber ff99SB*,[15,17,19] Amber ff03,[15,16] Amber ff03*,[15−17] GROMOS96 43a1p,[50,51] GROMOS96 53a6,[21,22] CHARMM27 (version c32b1) with CMAP,[18,52,53] and OPLS-AA/L force fields.[54−56] The "*" designations on the Amber force fields indicate the presence of a modification to the backbone dihedral potentials to improve agreement with experimental data.[17] The "ILDN" designation indicates the presence of a modification to the side-chain torsion potentials of isoleucine, leucine, aspartate, and asparagine to improve agreement with quantum-mechanical calculations.[20] Combination of the "ILDN" and ff99SB* modifications has been demonstrated recently.[44,57] For a recent summary of the evolution of the Amber ff99 and ff03 series of force fields, see the results section of ref 44. The "p" designation on the GROMOS96 43a1 force field indicates the inclusion of phosphorylated amino acid parameters to the otherwise unmodified 43a1 parameters.[50] One major difference between the GROMOS force fields and the others used in this study is that the GROMOS force fields are united atom and do not explicitly have all hydrogen atoms. The "AA" and "/L" designations on the OPLS force field indicate all-atom and the inclusion of updated dihedral parameters from the original distribution.[56]

Simulations of the same peptide with residue Thr-80 phosphorylated (pThr-80) were conducted with several of the above force fields in which phosphothreonine parameters were available. These included Amber ff99SB-ILDN, GROMOS96 43a1p, and CHARMM27. Phosphothreonine parameters from refs 58 and 50 were used for the Amber ff99SB-ILDN and GROMOS96 43a1p force field simulations, respectively. Phosphothreonine parameters included in the CHARMM27 force field distribution were used.[18,53]

Simulation Details. *A. General Parameters.* Simulations were performed using GROMACS (GROningen MAchine for Chemical Simulations) version 4.5.[47] Although GROMACS was used in this work, we expect that our findings will be applicable to other simulation software that utilizes the same force fields.[59] Cubic boxes of linear size 6 nm were used, and periodic boundary conditions were applied in all directions. Sodium (Na^+) and chloride (Cl^-) ions were added to neutralize the system and bring the salt concentration to 0.1 M. Na^+ and Cl^- parameters specific to each force field distribution were used.[60] Protein and nonprotein atoms were coupled to their own temperature baths, which were kept constant at 310 K using the Parrinello−Bussi algorithm.[61] This approach has been shown to perform very well in biomolecular simulations.[46] Pressure was maintained isotropically at 1 bar using the Parrinello−Rahman barostat.[62] A 2 fs time step was employed. Prior to the production runs, the energy of each system was minimized using the steepest descents algorithm. This was

Table 1. Summary of the MD Simulations

force field	water	elevated temp[a]	capped[b]	pThr-80[c]	charge-groups[d]	duplicate[e]
Amber ff99SB-ILDN	TIP3P (7038)					Y
	TIP3P (7030)		Y			Y
	TIP3P (7036)			Y		
Amber ff99SB*-ILDN	TIP3P (7038)	Y				
Amber ff99SB	TIP3P (7038)					Y
Amber ff99SB*[f]	TIP3P (7038)					Y
Amber ff03	TIP3P (7038)					
Amber ff03*	TIP3P (7038)	Y				
GROMOS96 43a1p	SPC (7035)				Y	
	SPC (7030)		Y			
	SPC (7033)			Y	Y	
	SPC (7027)		Y	Y		
GROMOS96 53a6	SPC (7035)	Y			Y	Y
	SPC (7033)	Y	Y			
CHARMM27	TIP3P (7038)	Y				Y
	TIP3P (7030)			Y		
OPLS-AA/L[f]	TIP3P (7037)	Y			Y	
	TIP4P (6969)					

[a]"Y" indicates that elevated temperature simulations were performed at 330, 350, and 370 K from the initial and final (after 1 μs) system configurations. [b]"Y" indicates that the N and C termini of the peptide were capped with acetyl (ACE) and NH_2 groups, respectively. They were otherwise left uncapped (NH_3^+ and COO^-, respectively). [c]"Y" indicates that residue Thr-80 was phosphorylated. [d]"Y" indicates that two simulations were performed: one with default GROMACS charge-groups and one without charge-groups. [e]"Y" indicates that two simulations, each 1 μs, were performed. Duplicates were always performed without charge-groups and were identical to the first simulation except for their initial atom velocities. [f]The trajectory was extended to 2 μs.

followed by 2 ps of position-restrained dynamics with all non-hydrogen atoms restrained with a 1000 kJ mol^{-1} force constant. Initial atom velocities were taken from a Maxwellian distribution at 310 K. All bond lengths were constrained using the LINCS algorithm.[63] A 1.0 nm cutoff was used for Lennard-Jones interactions. Dispersion corrections for energy and pressure were applied. Electrostatic interactions were calculated using the Particle-Mesh Ewald (PME) method[64] with 0.12 nm grid-spacing and a 1.0 nm real-space cutoff. No reaction-field or cutoff methods were tested as they have previously been shown to be inferior to PME.[65,66] System coordinates were written out at 4 ps intervals during the production runs.

B. System-Specific Attributes. The protonation states of all ionizable residues were chosen on the basis of their most probable state at pH 7. Unless specified, simulations were conducted with the amino and carboxyl terminals of the peptide left uncapped (NH_3^+ and COO^-, respectively) for each force field. When studying peptides from the interior of a protein sequence, it is common to add capping groups to the ends. This neutralizes the unphysical charges introduced by the free N- and C-termini, which can potentially disrupt the native structure. To study the effects of peptide capping, several simulations with the N- and C-terminus capped with acetyl (ACE) and NH_2 groups, respectively, were performed (Table 1). The starting structure was solvated in SPC (simple point charge), TIP3P, or TIP4P[67,68] water. The compatibility of these water models with ions has been examined in detail in ref 60. A three-point water model (SPC or TIP3P) was recommended by GROMACS for all of the force fields used in this study, with the exception of OPLS-AA/L, in which the four-point (TIP4P) water model was the recommended choice (Table 1). Simulations with TIP3P and TIP4P were conducted for this force field (Table 1). The nonphosphorylated peptide systems each contained 17 Na^+ and 13 Cl^- ions, while for the pThr-80 systems two extra Na^+ ions were added to neutralize the dianionic phosphate group. For each force field, a simulation was conducted without the use of charge-groups (single atom charge groups); GROMACS uses the concept of charge groups to speed up simulations; see section "Domain Decomposition" in ref 47 for details. It has recently been shown that in some situations charge-groups can lead to pronounced unphysical effects.[46] To examine the effect of charge-groups, additional simulations were conducted with the GROMOS96 and OPLS-AA/L force fields employing the default charge-groups for these force fields. Simulations performed with charge-groups are denoted with brackets around the force field name in the Results. For simulations conducted with the CHARMM27 force field, CMAP correction was applied.[18] A few of the simulations were duplicated to assess reproducibility (Table 1). These systems did not use charge-groups, were prepared in the same manner as stated above, and were assigned different initial atom velocities than their originals. Duplicated simulations are denoted with bracketed sequential numbering beside the force field name in the Results. We also performed elevated temperature simulations at 330, 350, and 370 K with the Amber ff99SB*-ILDN,[15,17,19,20] Amber ff03*,[15−17] GROMOS96 53a6,[21,22] CHARMM27 with CMAP,[18,52,53] and OPLS-AA/L force fields.[54−56] Using the initial and final (after 1 μs) system configurations at 310 K, we reassigned the atom velocities at each higher temperature and performed 0.2 μs MD simulations.

In total, 28 individual 1 μs simulations were conducted, and two of these trajectories were extended to 2 μs (Amber ff99SB* and OPLS-AA/L). An additional 7.2 μs of simulations at elevated temperature was performed. The cumulative simulation time was 37.2 μs. The simulations are summarized in Table 1.

Simulation Analysis. We used either the full 1 μs trajectories or the last 0.1 μs for analysis. By restricting some

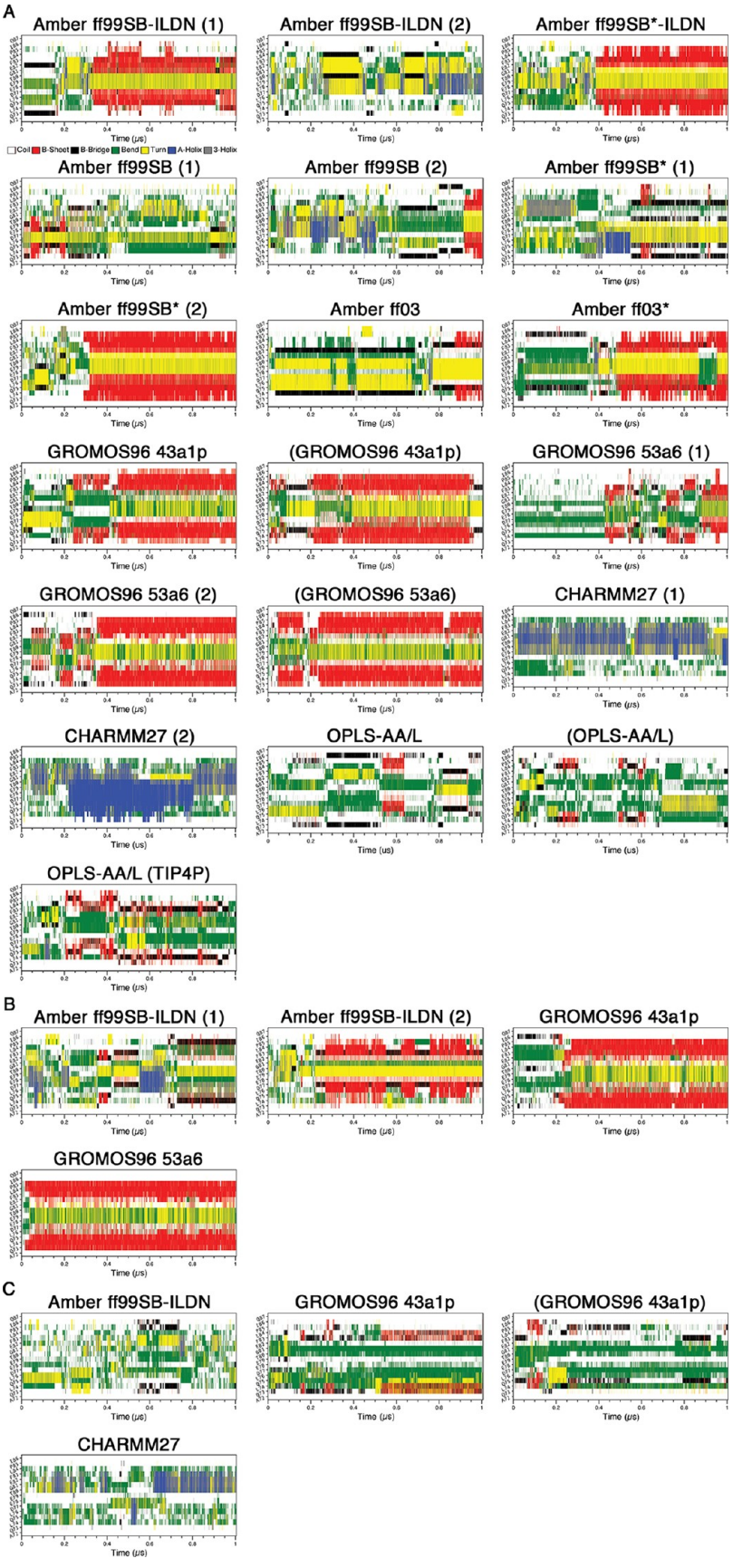

Figure 1. Secondary structure propensity analysis of the trajectories. Secondary structure content was assessed with the DSSP algorithm:[71] coil (white), β-sheet (red), β-bridge (black), bend (green), turn (yellow), α-helix (blue), and 3_{10} helix (gray). (A) Uncapped peptide. (B) Capped peptide. (C) pThr-80 peptide.

analyses to the last 0.1 μs, we allowed as much time as possible for the simulations to converge to a stable conformation. Hydrogen bonds were analyzed as follows: A hydrogen bond between a donor (D−H) and an acceptor (A) was considered to be formed when the DA distance was less than 3.2 Å and the angle between the DA vector and the D−H bond (AD−H angle) was less than 35°. These geometric criteria for defining hydrogen bonds are consistent with those used in prior studies.[69,70] Secondary structure content was assessed with the DSSP algorithm.[71]

RESULTS

We have compared the secondary structures and free- and bound-state contact formations in MD simulations of a β-hairpin forming peptide derived from the intrinsically disordered Neh2 domain of Nrf2 conducted with 10 different biomolecular force fields. The DSSP algorithm was used to monitor the evolution of secondary structures over the entire 1 μs trajectories. β-Hairpin formation was identified by inspection of the cluster center structures and the $C_\alpha-C_\alpha$ atom pair distances during the last 0.1 μs. Resemblance to the native state was gauged via the presence or absence of experimentally determined ^{1}H, ^{1}H NOEs.[30] We have also compared hydrogen bonds, rmsd's, and backbone dihedral angles in the MD structures to the peptide in complex with its binding target, Keap1.[31] The effects of elevated temperatures, terminal capping, charge-groups, and phosphorylation of Thr-80 on hairpin folding were also assessed.

Assessing Secondary Structure Formation at 310 K. To compare the MD trajectories obtained with different force fields, we first assessed the secondary structure content over the course of our simulations at 310 K using the program DSSP.[71] In this analysis, we deemed simulations that had residues from their β-turn regions (^{77}DEET^{80}) in "turn" conformations (yellow) flanked by residues in β-sheet conformations (red) simultaneously, to have formed β-hairpins (Figure 1). For the uncapped peptides, the Amber ff99SB-ILDN (1), Amber ff99SB*-ILDN, Amber ff99SB (2), Amber ff99SB* (2), Amber ff03, Amber ff03*, GROMOS96 43a1p, and GROMOS96 53a6 (1 and 2) simulations, including those which used charge-groups, appeared to adopt β-hairpins at some points in their trajectories (Figure 1A). Cluster center structures from the last 0.1 μs of the simulations, with the potential β-turn region (^{77}DEET^{80}) colored in black, are shown in Figure 2. The result clearly illustrates the 16-mer peptide folds into β-hairpin conformations by using these force fields. Intriguingly, the folding times vary greatly between ~0.05 and >0.9 μs in these simulations (Figure 1A). The CHARMM27 simulations did not form hairpins, and DSSP plots showed that helical content was present in the peptide (Figures 1A and 2A). None of the OPLS-AA/L simulations met our criteria for hairpin formation. Instead, these simulations were enriched in "bend" conformations (green) (Figure 1A). While there was some transient turn and β-sheet content in the expected locations, there were no pronounced β-hairpin signatures (Figures 1A and 2A). Extending of the OPLS-AA/L trajectory (without charge groups and TIP3P water) to 2 μs still did not yield a native-like hairpin (data not shown).

There were clear differences between replicate runs using the same force fields. Specifically, the Amber ff99SB-ILDN (2), Amber ff99SB (1), and Amber ff99SB* (1) simulations did not converge upon β-hairpin conformations (Figures 1A and 2A). To determine if a longer trajectory would lead to hairpin

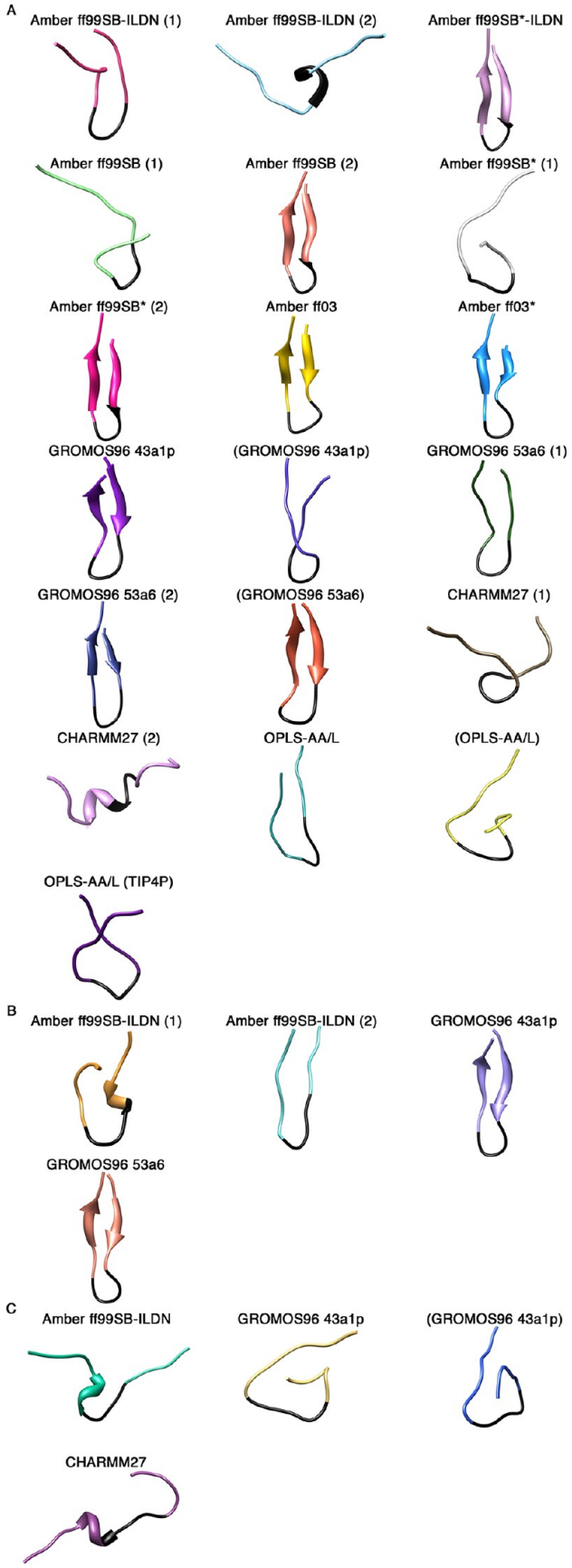

Figure 2. Cluster centroid structures from the last 0.1 μs of the simulations. A single cluster represented all structures in each simulation, and the center structure was extracted. (A) Uncapped peptide. (B) Capped peptide. (C) pThr-80 peptide.

dx.doi.org/10.1021/ct300323g | J. Chem. Theory Comput. 2012, 8, 2725−2740

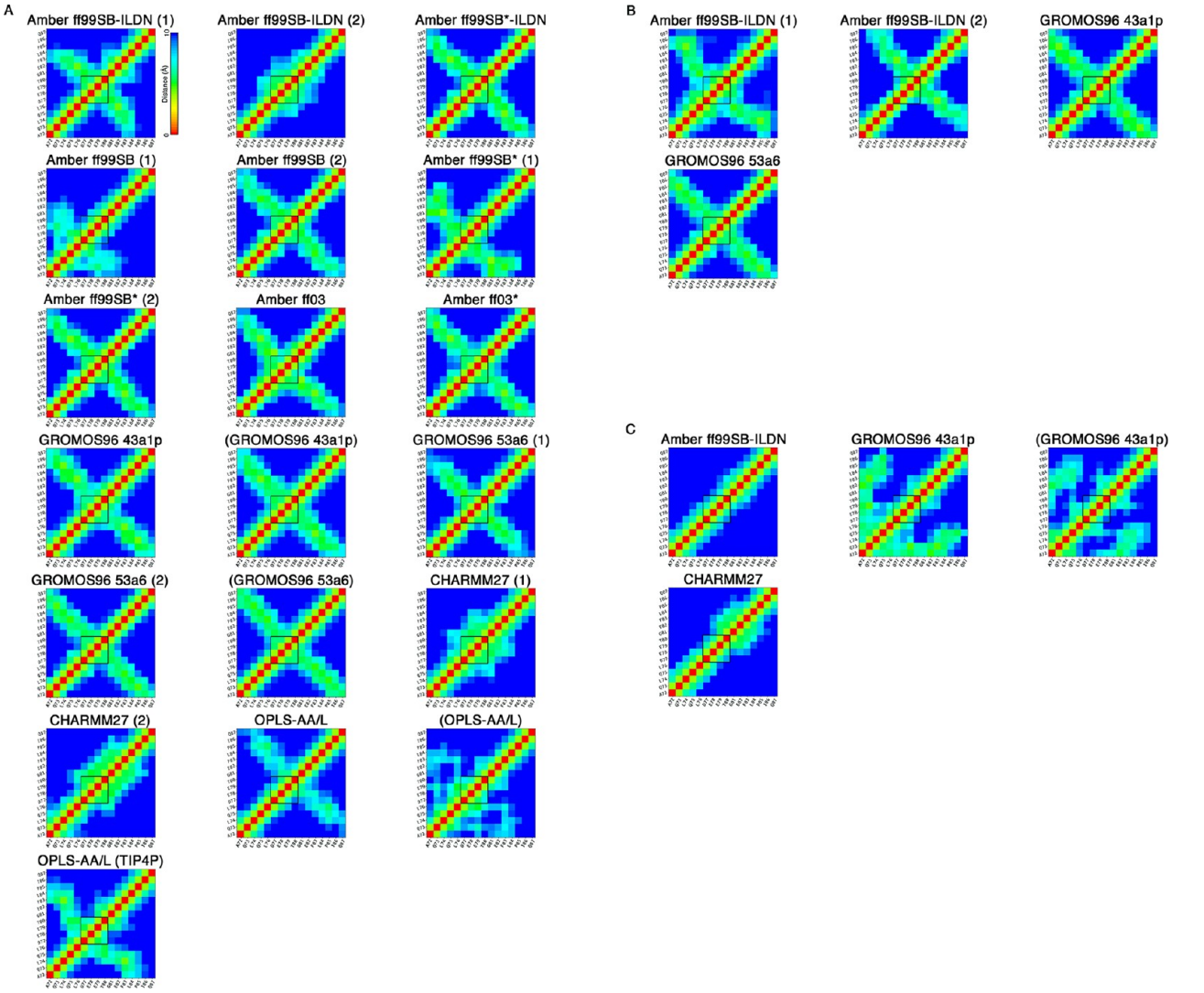

Figure 3. C_α−C_α atom pair distances. Average C_α−C_α distances less than or equal to 10 Å during the last 0.1 μs of the MD simulations. Distances equal to or greater than 10 Å are colored blue. The black square indicates the β-turn (^{77}DEET^{80}) region. (A) Uncapped peptide. (B) Capped peptide. (C) pThr-80 peptide.

formation, we extended the Amber ff99SB* (1) simulation to 2 μs. DSSP analysis, however, still was not indicative of a hairpin structure (data not shown).

For the capped peptides, while the Amber ff99SB-ILDN (2), GROMOS96 43a1p, and GROMOS96 53a6 simulations yielded hairpin signatures throughout large parts of the trajectories (Figure 1B), only the GROMOS96 force fields led to the formation of well-defined β-hairpins in the last 0.1 μs of the simulations (Figure 2B). Again, there were differences between the two Amber ff99SB-ILDN replicates. While the Amber ff99SB-ILDN (1) simulation did have turn and strand contents in the expected region, it appeared to be transient and not as pronounced as the Amber ff99SB-ILDN (2) hairpin signature (Figure 1B). Figure 2B shows that close to the end of the trajectory, Amber ff99SB-ILDN (1) structure adopted a short non-native helix before the turn region, while hairpin structure that is slightly displaced from the expected location was observed in the Amber ff99SB-ILDN (2) trajectory.

The phosphorylation of Thr-80 located at the turn region appears to have significant effects on the peptide folding. β-Hairpin formation was not observed in any of the pThr-80 simulations (Figures 1C and 2C). Interestingly, these simulations all displayed considerable bend content but failed to form a turn in the expected location (Figure 1C).

The averaged C_α−C_α atom pair distances (within 10 Å) were also plotted to identify β-hairpin formation in the simulations (Figure 3). In these plots, the β-turn (^{77}DEET^{80}) region, which the hairpin is approximately centered around, was indicated. For the uncapped peptides, the Amber ff99SB-ILDN (1), Amber ff99SB*-ILDN, Amber ff99SB (2), Amber ff99SB* (2), Amber ff03, Amber ff03*, GROMOS96 43a1p, and GROMOS96 53a6 (1 and 2) simulations, including those which used charge-groups, appeared to form β-hairpins as evidenced by the cross-strand C_α−C_α contacts centered around the β-turn (Figure 3A). Like the DSSP plots, this analysis also revealed clear differences between the replicates of Amber ff99SB-ILDN, Amber ff99SB, and Amber ff99SB* simulations. While Amber ff99SB-ILDN (2) displayed no signature of β-hairpin structure, the hairpins formed in the Amber ff99SB (1) and Amber ff99SB* (1) simulations were found in different regions as compared to the replicas (Figure 3A). The CHARMM27 simulations did not have cross-strand C_α−C_α contacts indicative of β-hairpin structures, but showed regions of compactness in the turn segment (Figure 3A). The OPLS-

AA/L simulations without charge groups had some evident cross-strand contacts, but the β-turn was shifted from the expected location (Figure 3A), while the OPLS-AA/L simulation with default charge-groups did not appear to form a hairpin (Figure 3A). C_α–C_α contacts in the capped peptide simulations were indicative of hairpin structures. However, unlike the conformations observed with the GROMOS force fields, the β-hairpin in the Amber ff99SB-ILDN simulations was shifted from the expected location (Figure 3B). None of the pThr-80 simulations had cross-strand contacts evident of β-hairpin structures (Figure 3C). It is worthwhile to note that in both GROMOS96 43a1p trajectories with Thr-80 phosphorylated there was evidence of close contacts between the positively charged N-terminus and the negatively charged phosphate group (Figure 3C). Because this may represent an unphysical interaction, we performed an additional simulation with a capped version of the peptide (Table 1). In this trajectory, we still did not observe hairpin or turn formation and noticed similar close contacts between the N-terminal region and phosphate group (data not shown).

Comparison to Experimental Data. Next, we compared the results of our simulations to experimental data. To begin, we assessed how many experimentally determined atomic contacts within the Nrf2 β-hairpin were found in our simulations. Even though the free state structure of the 16-mer human Nrf2 peptide used in this study is not currently available, several atomic contacts within the β-hairpin region of mouse Nrf2 have been determined by NMR.[30] The mouse Nrf2 contains the same β-hairpin sequence as the human version used in this study, except with a single conservative amino acid change of L74F (Figure 4A). Given that the human and mouse Nrf2 β-hairpin sequences are nearly identical, they are expected to adopt similar structures.

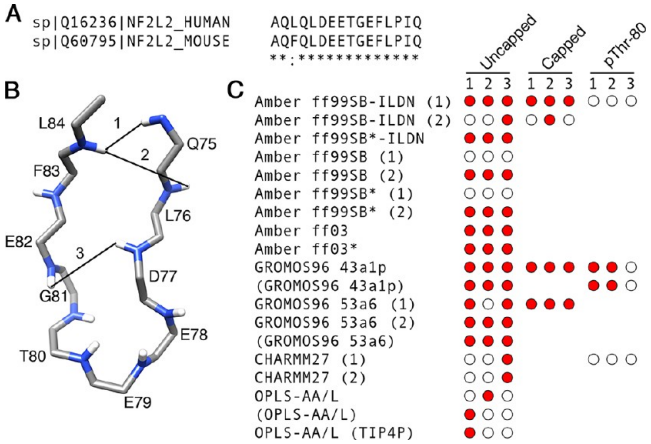

Figure 4. Nrf2 β-hairpin sequence alignment and native contacts. (A) Sequence alignments between the human and mouse Nrf2 β-hairpin segments were generated with ClustalW XXL.[80] The Blosum scoring matrix[81] was used, and gap penalties were set at their default values. Opening and end gap penalties were set to 10. Extending and separation gap penalties were set to 0.05. (B) Three NMR-derived cross-strand ^{1}H, ^{1}H NOEs determined by ref 30 mapped onto a model Nrf2 β-hairpin backbone structure. (C) Presence or absence of each native contact for each force field. Time-averaged distances <6 Å during the last 0.1 μs of the simulations between hydrogen atom pairs matching those observed by Tong et al.[30] were considered to be native contacts.

We compared the NMR-derived cross-strand ^{1}H, ^{1}H NOEs determined in ref 30 to the corresponding time-averaged distances from our MD simulations. Time-averaged distances <6 Å between hydrogen atom pairs matching those observed in[30] were considered to be native contacts. Because the united-atom GROMOS96 force fields used in this study do not explicitly represent every hydrogen atom, we restricted our analysis to backbone amide hydrogens, which were explicitly represented in all force fields. NOEs between adjacent residues and those involving F74 were excluded from the analysis. This reduced the number of experimentally determined native contacts used in this analysis to three (Q75 HN:L84 HN, L76 HN:L84 HN, and D77 HN:E82 HN). They are depicted in Figure 4B.

The presence or absence of each of the three contacts is shown in Figure 4C. For the uncapped peptides, the Amber ff99SB-ILDN (1), Amber ff99SB*-ILDN, Amber ff99SB (2), Amber ff99SB* (2), Amber ff03, Amber ff03*, GROMOS96 43a1p, and GROMOS96 53a6 simulations, including those which used charge-groups, had at least two of the three native contacts (Figure 4C). Once again, there were differences between the Amber replicas (Figure 4C). Notably, in the Amber ff99SB-ILDN (2), Amber ff99SB (1), and Amber ff99SB* (1) simulations, only one or none of the native contacts were present, while their replicas had all three (Figure 4C). The CHARMM27 and OPLS-AA/L simulations had only one out of the three native contacts (Figure 4C). The capped peptides were able to form all three native contacts, but differences between duplicates were also evident. The Amber ff99SB-ILDN (1) simulation had all three contacts, while its duplicate had only one (Figure 4C). Native contacts were reduced in all pThr-80 simulations as compared to their unphosphorylated counterparts (Figure 4C). Interestingly, two of the three native contacts were still present in the GROMOS96 43a1p pThr-80 trajectories (Figure 4C). In these simulations, while the two contacts in the β-sheet region of the hairpin were present, the contact in the turn region was missing (Figure 4C).

In addition to NMR-derived contacts, backbone and side chain hydrogen bonds between Asp-77 and Thr-80 are present when Nrf2 is bound to Keap1 (PDB id: 2FLU).[31] We previously found that hydrogen bonds between these residues may also exist in the free state with high frequencies in simulations conducted with the GROMOS96 53a6 force field.[9] Because hydrogen bonding between Asp-77 and Thr-80 may be correlated with β-hairpin formation, we calculated the frequencies of hydrogen bonding between these residues (Table 2). For the uncapped peptides, we observed high (>0.68) frequencies of Asp-77 to Thr-80 hydrogen bonding in the Amber ff99SB-ILDN (1), Amber ff99SB*-ILDN, Amber ff99SB (2), Amber ff99SB* (2), Amber ff03*, GROMOS96 43a1p, and GROMOS96 53a6 (1 and 2) simulations, including those which used charge-groups (Table 2). Like the aforementioned analyses, clear differences between some replicas were observed (Table 2). Specifically, the Amber ff99SB-ILDN (2), Amber ff99SB (1), and Amber ff99SB* (1) simulations had considerably less Asp-77 to Thr-80 hydrogen bonding as compared to their duplicates (Table 2).

It is noteworthy that in the Amber ff03 simulation, no hydrogen bonding between Asp-77 and Thr-80 was observed (Table 2). Because prior analysis showed that this trajectory formed a hairpin with three native contacts (Figure 4A), the complete lack of hydrogen bonding between these two residues

Table 2. Frequency of Asp-77 to Thr-80 Hydrogen Bonding[a]

force field	uncapped[b]	capped[c]	pThr-80[d]
Amber ff99SB-ILDN (1)	0.95	0.00	0.00
Amber ff99SB-ILDN (2)	0.27	0.00	
Amber ff99SB*-ILDN	0.94		
Amber ff99SB (1)	0.00		
Amber ff99SB (2)	0.97		
Amber ff99SB* (1)	0.00		
Amber ff99SB* (2)	0.97		
Amber ff03	0.00		
Amber ff03*	0.86		
GROMOS96 43a1p	0.69	0.91	0.00
(GROMOS96 43a1p)[e]	0.94		0.00
GROMOS96 53a6 (1)	0.90	0.92	
GROMOS96 53a6 (2)	0.93		
(GROMOS96 53a6)	0.91		
CHARMM27	0.32		0.00
CHARMM36	0.22		
OPLS-AA/L	0.00		
(OPLS-AA/L)	0.00		
OPLS-AA/L (TIP4P)	0.00		

[a]Frequencies of 1 or more hydrogen bonds during the last 0.1 μs of the trajectories. Oxygen and nitrogen atoms were acceptors. Amine groups and the hydroxyl group of Thr-80 were donors. Intraresidue hydrogen bonds were excluded from the analysis. A hydrogen bond between a hydrogen donor (D−H) and a hydrogen acceptor (A) was judged to be formed when the DA distance (r) was less than 3.2 Å, and the angle between the DA vector and the D−H bond (AD−H angle) was less than 35°. [b]Values for the peptides with unmodified N and C termini (NH_3^+ and COO^-, respectively). [c]Values for the peptides with capped N and C termini (ACE and NH_2, respectively). [d]Values for the peptides with residue Thr-80 phosphorylated. [e]Parenthetical values indicate hydrogen-bond frequencies for trajectories with default GROMACS charge groups.

was unexpected. Inspection of the trajectory showed that the side chains of Leu-76 and Asp-77 were on opposite sides of the hairpin as compared to the other simulations (data not shown). Although Asp-77 was not in an appropriate orientation to form hydrogen bonds with Thr-80, we found that the frequency of forming 1 or more hydrogen bonds between Leu-76 and Thr-80 was 0.66 in this trajectory. Low frequencies of Asp-77 to Thr-80 hydrogen bonding were found in the CHARMM27 simulations (between 0.22 and 0.32) and were completely absent in the OPLS-AA/L simulations (Table 2).

For the capped peptides, no hydrogen bonding between Asp-77 and Thr-80 was observed in the Amber ff99SB-ILDN simulations, but was present in over 90% of the structures in the last 0.1 μs of the GROMOS96 trajectories (Table 2). While the capped Amber ff99SB-ILDN (1) simulation was found to have three native contacts (Figure 4C), cluster analysis indicated that there was a short non-native helix before its β-turn region (Figure 1B). Furthermore, the DSSP plot of this trajectory did not have a typical hairpin signature (Figure 2B). These factors likely contributed to the lack of Asp-77 to Thr-80 hydrogen bonding in this trajectory (Table 2). The capped Amber ff99SB-ILDN (2) simulation had only one native contact (Figure 4C), and its β-turn was slightly displaced from the expected location (Figure 3B), factors that likely contributed to the lack of hydrogen bonding (Table 2). Hydrogen bonding between Asp-77 and Thr-80 was not observed in any of the pThr-80 simulations (Table 2).

We also compared the Nrf2 peptide structures from our simulations to that of the Keap1-bound state (PDB id: 2FLU).[31] This comparison is interesting because the ETGE motif of the disordered Nrf2 has been shown to have a tendency to form bound state-like structure even in the absence of the target.[9,30,31] Because the Nrf2 β-hairpin does not adopt a well-defined in the free state,[30] we restricted the rmsd calculations to backbone atoms only.

The rmsd's were calculated separately for the β-turn, ^{77}DEET^{80} and β-hairpin, ^{72}AQLQLDEETGEFL^{84} regions. The rmsd's throughout the trajectories are plotted in Figure 5, and average rmsd values over the last 0.1 μs are shown in Figure 6 and summarized in Tables 3−5. For the uncapped peptides, the Amber ff99SB-ILDN (1), Amber ff99SB*-ILDN, Amber ff99SB (2), Amber ff99SB* (2), Amber ff03, Amber ff03*, GROMOS96 43a1p, and GROMOS96 53a6 (1 and 2) simulations achieved average rmsd's < 1 and <3 Å to the bound state β-turn and hairpin, respectively, including simulations that used charge-groups (Figure 6A). Again, there were some differences between replicas. The Amber ff99SB-ILDN (2) simulation had a β-turn region with an average rmsd <1 Å, but when considering the full β-hairpin, the rmsd was larger than 4.8 Å (Figure 6A). Also, the Amber ff99SB (1) and Amber ff99SB* (1) simulations had substantially higher rmsd's as compared to their duplicates (Figures 5A and 6A). The CHARMM simulations did not lead to a bound state like β-hairpin (rmsd's > 4 Å), but had β-turn rmsd's below 1 Å (Figures 5A and 6A). The OPLS-AA/L simulations had both β-turn and hairpin rmsd's greater than 1 and 3 Å, respectively (Figures 5A and 6A), indicating significant deviations from the bound-state structure.

For the capped peptides, both Amber ff99SB-ILDN simulations had average rmsd's < 3 Å for the hairpin region, but their β-turns had rmsd's > 1 Å (Figures 5B and 6B). In comparison, both GROMOS96 force fields had rmsd's of <1 and <3 Å for the β-turn and hairpin, respectively (Figures 5B and 6B). These rmsd's were similar to their uncapped versions (Figures 5A and 6A). It is worthwhile to note that the capped GROMOS96 53a6 simulation converged to bound state like structure in <0.05 μs, the fastest of all of the simulations (Figure 5B). Among the simulations that had bound state like rmsd's, the amount of time it took to adopt these conformations varied between <0.05 and ∼0.9 μs, even for duplicates using the same force field (Figure 5A and B). However, once a bound state like structure was formed, it tended to remain stable. The β-turn and hairpin rmsd's were higher in all pThr-80 simulations as compared to those of the unphosphorylated peptides (Figures 5C and 6C).

The convergence of the dihedral angles from the trajectories to those from the bound state structure was also assessed (PDB id: 2FLU).[31] The combined ϕ and ψ angles from the simulations and bound state structure are shown in Figure 7, and the average per-residue deviations are shown in Figure 8. For the uncapped peptides, the GROMOS96 43a1p with charge groups and 53a6 force field simulations had the lowest ϕ and ψ deviations from the bound state structure (Figure 8A). These simulations had combined ϕ and ψ deviations of <7° and <17° per residue from the bound state in their β-turn and hairpin regions, respectively (Figure 8A). The Amber ff99SB-ILDN (1) and CHARMM simulations had combined ϕ and ψ deviations of ∼10° in their β-turns, but deviated >20° per residue when considering the entire hairpin (Figure 8A). For the capped peptides, both GROMOS96 force fields had slightly

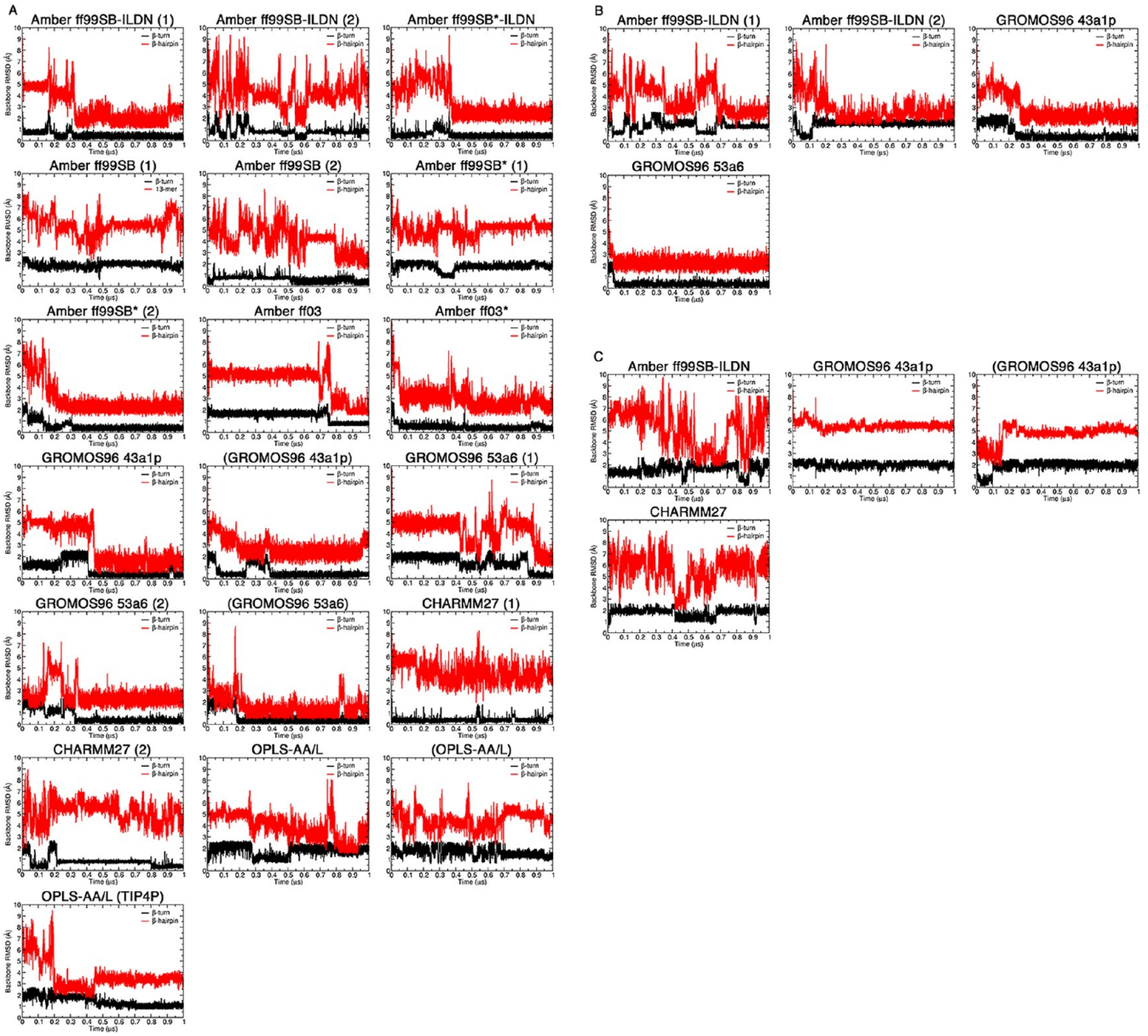

Figure 5. Backbone rmsd's between the bound state and MD structures throughout the trajectories. The rmsd values were calculated for the β-turn 4-mer, ^{77}DEET^{80} (black), and β-hairpin 13-mer, ^{72}AQLQLDEETGEFL^{84} (red), by least-squares fitting the backbone atoms (N, C_α, and C) from each frame to the corresponding atoms of bound state reference structure (PDB id: 2FLU).[31] (A) Uncapped peptide. (B) Capped peptide. (C) pThr-80 peptide.

lower deviations as compared to their uncapped counterparts and had considerably lower deviations than Amber ff99SB-ILDN (Figure 8B). The β-turn and hairpin deviations were higher in all pThr-80 simulations as compared to those of the unphosphorylated peptides (Figure 8B).

Secondary Structure Formation at Elevated Temperatures. Finally, we have performed MD simulations at elevated temperatures using a subset of force fields, Amber ff99SB*-ILDN and ff03*, GROMOS96 53a6, CHARMM27, and OPLS-AA/L, to identify secondary structure formation of the Neh2 peptide under these conditions. Using elevated temperatures provides an additional test to examine possible metastable states. The simulations were performed at 330, 350, and 370 K from both the initial and the final (after 1 μs) system configurations at 310 K. Again, we used DSSP analysis to illustrate the evolution of secondary structures over the trajectories.

In the simulations starting from the initial (unfolded) system coordinates, hairpin formation, at the expected location, was observed in the Amber ff03*, GROMOS96 53a6 (2), and capped GROMOS96 53a6 simulations at 330 K (Figure 9). β-Hairpin structures were also identified in the MD simulations using these force fields at 310 K as mentioned above (Figure 1). Hairpin conformation, which was not observed in CHARMM27 (1) at 310 K, was significantly populated in the trajectory at 330 K (Figure 9). At 350 K, the Amber ff03*, GROMOS96 53a6 (2), and capped GROMOS96 53a6 simulations still had hairpin signatures at some points in their trajectories, but β-hairpin structure was no longer observed in the CHARMM27 (1) simulation (Figure 9). On the other hand, a low population of hairpin conformation was observed

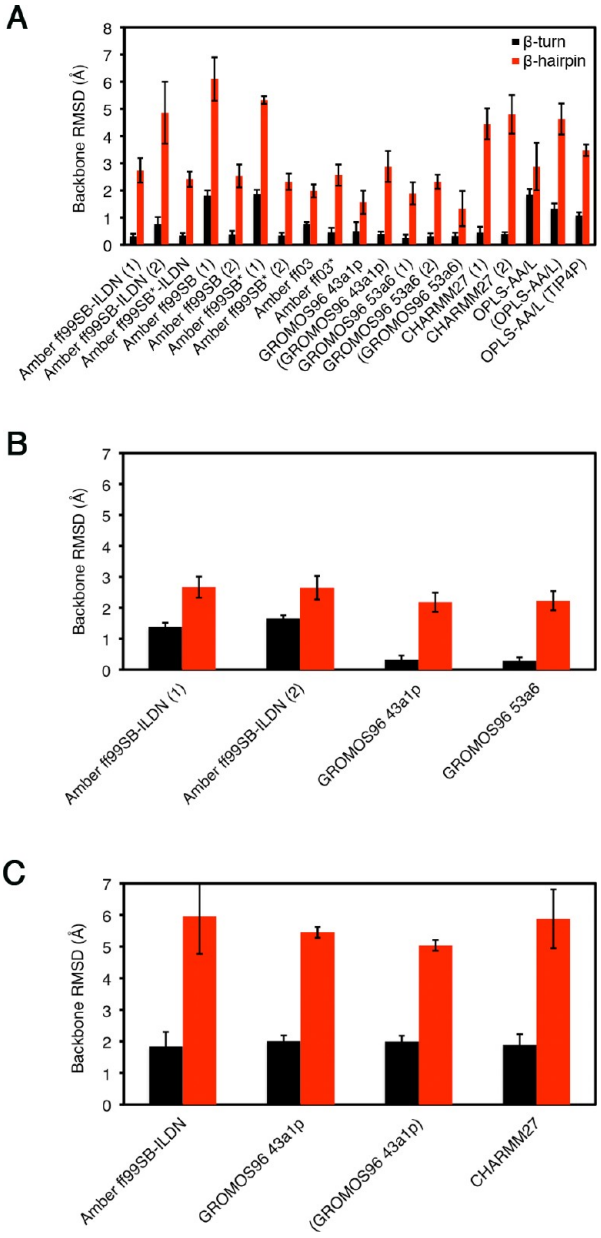

Figure 6. Average backbone rmsd's between the bound state and MD structures. Average rmsd values were calculated over the last 0.1 μs of the simulations for the β-turn 4-mer, ^{77}DEET^{80} (black), and β-hairpin 13-mer, ^{72}AQLQLDEETGEFL^{84} (red), by least-squares fitting the backbone atoms (N, C_α, and C) from each frame to the corresponding atoms of bound state reference structure (PDB id: 2FLU).[31] (A) Uncapped peptide. (B) Capped peptide. (C) pThr-80 peptide.

in the 350 K OPLS-AA/L trajectory (Figure 9). Significant population of β-hairpin structure remained even at 370 K in GROMOS96 53a6 (2) and capped GROMOS96 53a6 simulations (Figure 9), while only transiently formed hairpin was observed in CHARMM27 (1). It is noteworthy that rapid hairpin folding and high thermal stability were observed in the capped GROMOS96 53a6 simulations at all elevated temperatures (Figure 9).

In the simulations starting from the final system coordinates, the Amber ff99SB*-ILDN, Amber ff03*, GROMOS96 53a6 (2), and capped GROMOS96 53a6 trajectories, all of which formed hairpins at 310 K, maintained hairpin signatures at 330 K over 0.2 μs (Figure 10). On the other hand, the

Table 3. Average Backbone rmsd's between the Bound State Structure and MD Structures of the Uncapped Peptides[a]

force field	backbone[b] (Å) ± sdev β-turn[c]	backbone (Å) ± sdev β-hairpin[d]
Amber ff99SB-ILDN (1)	0.30 ± 0.10	2.74 ± 0.45
Amber ff99SB-ILDN (2)	0.77 ± 0.25	4.86 ± 1.14
Amber ff99SB*-ILDN	0.35 ± 0.08	2.41 ± 0.28
Amber ff99SB (1)	1.81 ± 0.19	6.10 ± 0.80
Amber ff99SB (2)	0.38 ± 0.13	2.53 ± 0.42
Amber ff99SB* (1)	1.86 ± 0.16	5.33 ± 0.14
Amber ff99SB* (2)	0.35 ± 0.09	2.32 ± 0.30
Amber ff03	0.77 ± 0.06	1.98 ± 0.24
Amber ff03*	0.46 ± 0.16	2.56 ± 0.39
GROMOS96 43a1p	0.50 ± 0.33	1.56 ± 0.43
(GROMOS96 43a1p)[e]	0.39 ± 0.09	2.88 ± 0.57
GROMOS96 53a6 (1)	0.25 ± 0.12	1.89 ± 0.41
GROMOS96 53a6 (2)	0.30 ± 0.12	2.32 ± 0.26
(GROMOS96 53a6)	0.32 ± 0.12	1.42 ± 0.56
CHARMM27	0.45 ± 0.21	4.45 ± 0.57
CHARMM36	0.39 ± 0.07	4.80 ± 0.71
OPLS-AA/L	1.84 ± 0.21	2.88 ± 0.87
(OPLS-AA/L)	1.33 ± 0.19	4.63 ± 0.57
OPLS-AA/L (TIP4P)	1.08 ± 0.11	3.48 ± 0.21

[a]Average rmsd's were calculated over the last 0.1 μs of the trajectories. [b]Backbone atoms include N, C_α, and C. [c]β-Turn (^{77}DEET^{80}). [d]β-Hairpin (^{72}AQLQLDEETGEFL^{84}). [e]Average rmsd's for trajectory with default GROMACS charge groups.

Table 4. Average rmsd's between the Bound State Conformation and MD Structures of the Capped Peptides[a]

force field	backbone[b] (Å) ± sdev β-turn[c]	backbone (Å) ± sdev β-hairpin[d]
Amber ff99SB-ILDN (1)	1.39 ± 0.13	2.67 ± 0.34
Amber ff99SB-ILDN (2)	1.66 ± 0.10	2.65 ± 0.38
GROMOS96 43a1p	0.32 ± 0.14	2.18 ± 0.31
GROMOS96 53a6	0.29 ± 0.11	2.23 ± 0.31

[a]Average rmsd's were calculated over the last 0.1 μs of the trajectories. [b]Backbone atoms include N, C_α, and C. [c]β-Turn (^{77}DEET^{80}). [d]β-Hairpin (^{72}AQLQLDEETGEFL^{84}).

Table 5. Average rmsd's between the Bound State Conformation and MD Structures of the pThr-80 Peptides[a]

force field	backbone[b] (Å) ± sdev β-turn[c]	backbone (Å) ± sdev β-hairpin[d]
Amber ff99SB-ILDN	1.85 ± 0.45	5.96 ± 1.19
GROMOS96 43a1p	2.01 ± 0.18	5.45 ± 0.17
(GROMOS96 43a1p)[e]	2.00 ± 0.18	5.04 ± 0.17
CHARMM27	1.89 ± 0.34	5.88 ± 0.93

[a]Average rmsd's were calculated over the last 0.1 μs of the trajectories. [b]Backbone atoms include N, C_α, and C. [c]β-Turn (^{77}DEET^{80}). [d]β-Hairpin (^{72}AQLQLDEETGEFL^{84}). [e]Average rmsd's for trajectory with default GROMACS charge groups.

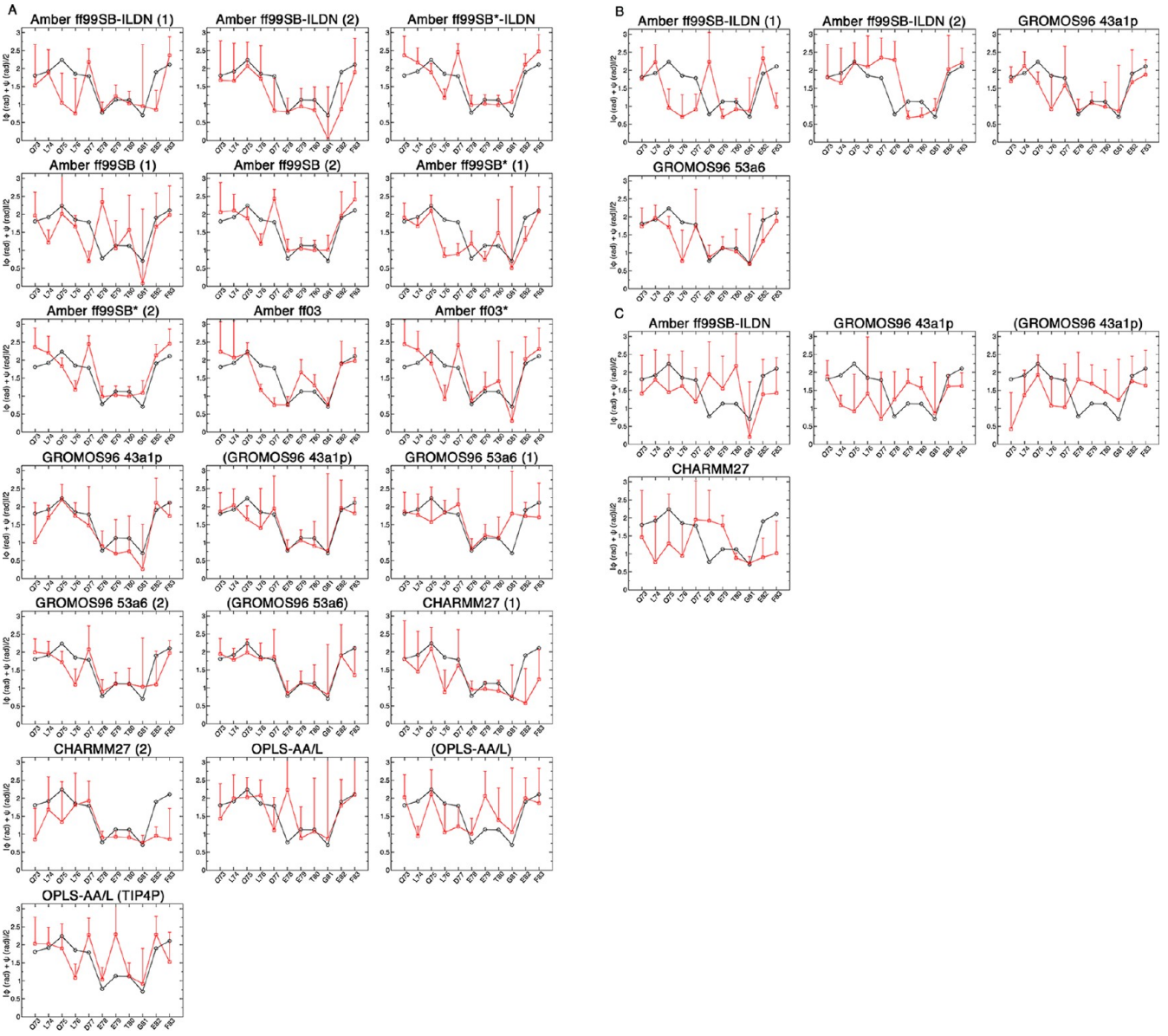

Figure 7. Comparison of the backbone dihedral angles from the bound state structure and MD simulations. The ϕ and ψ angles for residues ^{73}QLQLDEETGEF^{83} were converted to radians, and the absolute values were summed and averaged. Black circles indicate the values from the bound state crystal structure (PDB id: 2FLU).[31] Red squares are the values over the last 0.1 μs of the simulations. (A) Uncapped peptide. (B) Capped peptide. (C) pThr-80 peptide.

CHARMM27 (1) and OPLS-AA/L trajectories at 330 K were heavily biased by α-helical and bend conformations, respectively, similar to what was observed at 310 K (Figure 10). When the temperature was increased to 350 K, the hairpin signature in the Amber ff99SB*-ILDN trajectory disappeared shortly after ~0.1 μs; however, the Amber ff03*, GROMOS96 53a6 (2), and capped GROMOS96 53a6 trajectories maintained their hairpins over the whole 0.2 μs period (Figure 10). The CHARMM27 (1) simulation at 350 K lost its helical properties after about 0.15 μs and appeared to form a hairpin shortly after (Figure 10). At 370 K, both the Amber ff99SB*-ILDN and the Amber ff03* trajectories lost their hairpin signatures after ~0.1 μs, but the GROMOS96 53a6 (2) and capped GROMOS96 53a6 simulations still remained in hairpin conformations throughout almost whole trajectories (Figure 10). On the other hand, the CHARMM27 (1) simulation at 370 K lost its helical property almost immediately and a turn conformation was present in the expected location, but a distinct hairpin signature was not observed (Figure 10). The OPLS-AA/L trajectories did not have any clear hairpin signatures at any of the temperatures (Figure 10). Once again, high thermal stability was observed in the GROMOS96 53a6 simulations (Figure 10).

DISCUSSION AND CONCLUSIONS

We have examined the folding of a 16-mer polypeptide with 10 commonly used biomolecular force fields. The peptide used in this study is derived from the Neh2 domain of Nrf2. Despite that Neh2 has been characterized as being intrinsically disordered, the region encoded by the sequence of this peptide has been shown to contain β-hairpin structure.[30,31,72] Various criteria were used to assess β-hairpin formation of this peptide and compare the results to experimental data. Although the simulations all used the same, non-native, starting structure and

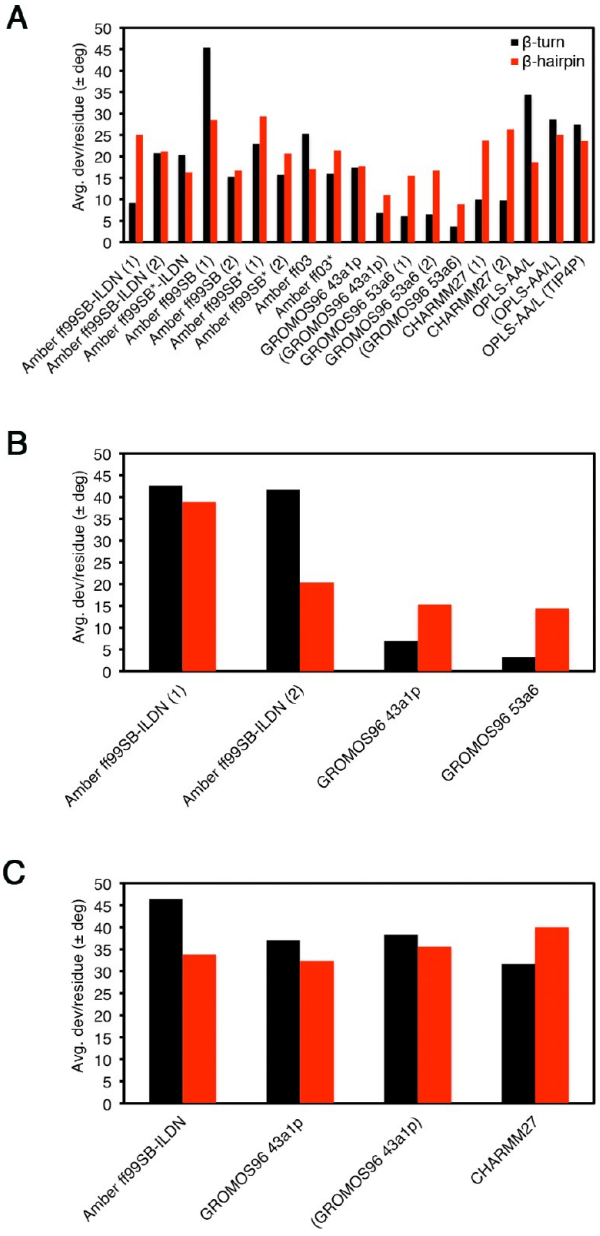

Figure 8. Average combined ϕ and ψ deviations per residue from the bound state crystal structure (PDB id: 2FLU).[31] Black bars are for the β-turn 4-mer, ^{77}DEET^{80}, and red bars are for the β-hairpin 13-mer, ^{72}AQLQLDEETGEFL^{84}. Data were analyzed over the last 0.1 μs of the simulations. (A) Uncapped peptide. (B) Capped peptide. (C) pThr-80 peptide.

were performed with identical parameters, clear differences were observed between different force fields used and even between replicate simulations with the same force field.

While no single type of analysis was sufficient to thoroughly assess and compare β-hairpin formation, the DSSP plots were useful for visualizing potential hairpin formation in this work. In addition, these analyses were also useful in identifying other types of secondary structures. For example, DSSP plots of the CHARMM27 simulations showed that the Nrf2 peptide did not fold into hairpins, but had tendencies to form short α-helices (Figure 2A). This finding was not completely unexpected because CHARMM force fields have been known to have a bias toward helical structures, even when simulating the folding of all β proteins.[3,11−13,73] In addition to CHARMM27, the Amber ff03 force field has also been shown to overstabilize helical structures. Lindorff-Larsen et al.[44] observed that while both CHARMM27 and Amber ff03 could fold the α-helical villin headpiece, proper folding of the β-sheet WW domain could not be achieved even in simulations that were 10 times the experimentally determined folding time in length. On the other hand, they found that the "helix−coil-balanced" Amber ff03* and recently developed CHARMM22* variants could achieve proper folding of both villin and the WW domain.[8,57]

The DSSP plots for the OPLS-AA/L force field simulations also were not indicative of hairpins, but showed considerable amounts of "bend" content. This aligns with the finding of Cao et al.[74] that this force field did not produce the expected β-hairpin structure of the H1 peptide. Interestingly, simulations of the H1 peptide performed with GROMOS96 43a1 yielded a β-hairpin structure consistent with experimental data.[74] It is difficult to determine why the OPLS-AA/L simulations did not form a native-like hairpin structure in our simulations. It is possible that, in general, longer trajectories may be needed for convergence due to the rugged energy landscape and different barriers in these systems.[75,76] The weak hairpin signature observed in the DSSP plot of the OPLS-AA/L trajectory at 350 K supports this notion. Alternatively, there may be incompatibilities between our peptide sequence and OPLS-AA/L, such as high amounts of exposed hydrophobic content.[74]

The $C_\alpha-C_\alpha$ contacts plots also illustrated β-hairpin formation and helped to identify non-native hairpins. For example, these plots showed that the β-turn in the capped peptide Amber ff99SB-ILDN (2) simulation was slightly displaced from its expected location (Figure 2B). This likely explains the lack of Asp-77 to Thr-80 hydrogen bonding in this simulation (Table 2). Together, our findings from the DSSP and $C_\alpha-C_\alpha$ contact analysis suggested that the uncapped Amber ff99SB-ILDN (1), Amber ff99SB*-ILDN, Amber ff99SB (2), Amber ff99SB* (2), Amber ff03, Amber ff03*, GROMOS96 43a1p, GROMOS96 53a6 (1 and 2), and capped GROMOS96 43a1p and 53a6 simulations formed native-like β-hairpins.

Interestingly, the simulations that formed β-hairpins, as judged by DSSP and $C_\alpha-C_\alpha$ contact analysis, also exhibited experimentally determined native contacts present in the free state of Nrf2. Furthermore, we observed that the presence or absence of native contacts was correlated with the frequency of Asp-77 to Thr-80 hydrogen bonding to some extent. Interactions between these residues are thought to be important for the hairpin structure.[9,31] Most of the simulations of the uncapped and unphosphorylated peptides that had two or more native contacts also had high frequencies of Asp-77 to Thr-80 hydrogen bonding. On the other hand, when one or zero native contacts were present, there was usually less hydrogen bonding. One exception was the Amber ff03 simulation, which had all three native contacts, but lacked hydrogen bonding between Asp-77 to Thr-80. Figure 5 shows that in this simulation, Leu-76 and Asp-77 had large backbone dihedral angle deviations from the bound state structure, which could possibly explain the lower hydrogen bonding with Thr-80. It is possible that alternate hydrogen bonds between Leu-76 and Thr-80 may have compensated. The evident positive correlation between native contact formation and a high frequency of Asp-77 to Thr-80 hydrogen bonding in our simulations supports prior suggestions that these interactions are vital for the hairpin structure.[9,31]

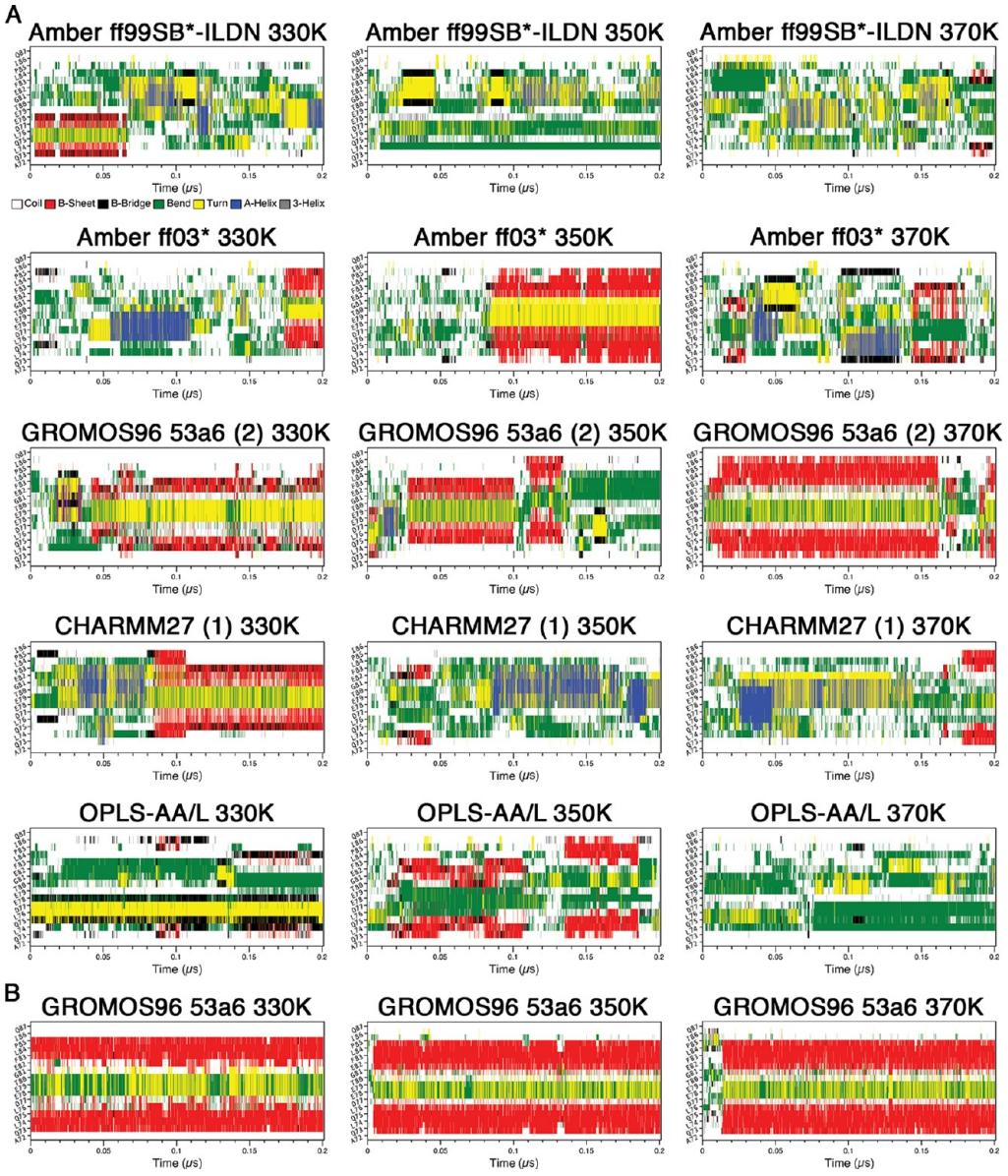

Figure 9. Secondary structure propensity analysis of the elevated temperature simulations from the initial system configurations. Secondary structure content was assessed with the DSSP algorithm:[71] coil (white), β-sheet (red), β-bridge (black), bend (green), turn (yellow), α-helix (blue), and 3_{10} helix (gray). (A) Uncapped peptide. (B) Capped peptide.

We also found that the simulations that formed β-hairpins converged upon conformations that were similar to the structure of Nrf2 bound to Keap1 (PDB id: 2FLU).[31] It is common for disordered proteins, like Nrf2, to contain preformed structural elements in their binding regions.[9,30,77−79] Indeed, NMR data and our prior MD simulations indicated that Nrf2 adopts a hairpin structure in the free-state, which is in high resemblance to its Keap1 bound form (PDB id: 2FLU).[9,31] Therefore, it was expected that simulations with 2−3 free state native contacts also had low rmsd's to the bound state structure. The GROMOS96 simulations clearly had the lowest β-turn and hairpin rmsd's of all of the simulations. These simulations also had very low dihedral angle deviations from the bound state structure.

In general, the simulations that used charge-groups or peptide capping groups were not largely different from their uncapped counterparts with single atom charge-groups. When studying peptides from the interior of a protein sequence, it is common to add capping groups to the ends. This neutralizes the unphysical charges introduced by the free N- and C-termini, which can potentially disrupt the native structure. However, we did not find that the uncapped termini had a detrimental effect on hairpin folding in our current simulations. The GROMOS96 force field simulations employing default charge-groups or peptide capping groups were highly consistent, in all aspects, to their uncapped counterparts. On the other hand, in both capped Amber ff99SB-ILDN replicates, the peptide folded into structures that were moderately different from their uncapped counterparts. It was difficult to determine the cause of this behavior, and it could simply be a convergence issue.

The finding that none of the simulations where Thr-80 was phosphorylated formed β-hairpins was not surprising. Experimental data have shown that phosphorylation of this residue can severely impair binding of Nrf2 to Keap1, likely due to a disruption of β-turn formation.[31] Our pThr-80 simulations

dx.doi.org/10.1021/ct300323g | J. Chem. Theory Comput. 2012, 8, 2725−2740

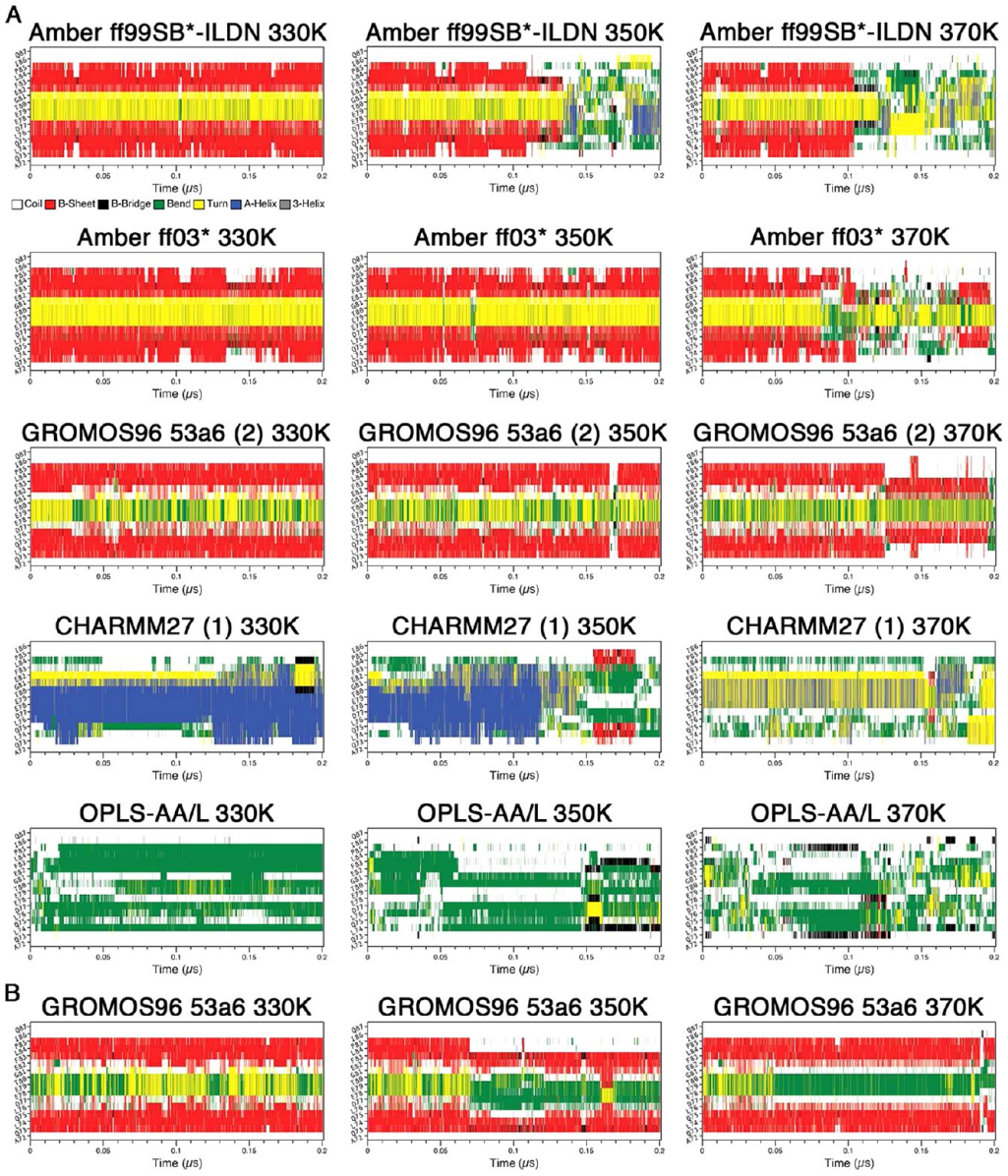

Figure 10. Secondary structure propensity analysis of the elevated temperature simulations from the final (after 1 μs) system configurations. Secondary structure content was assessed with the DSSP algorithm:[71] coil (white), β-sheet (red), β-bridge (black), bend (green), turn (yellow), α-helix (blue), and 3_{10} helix (gray). (A) Uncapped peptide. (B) Capped peptide.

were consistent with this proposition and also suggest that β-turn disruption strongly impairs hairpin formation.

The evident differences between duplicate simulations in this work highlight the importance of replica simulations when performing MD simulations of folding. Even though all duplicate simulations here used identical starting structures and parameters, the assignment of different initial atom velocities led the simulations to follow different pathways. As a result, duplicate simulations did not always converge upon folded structures even with microsecond long trajectories. In this work, we have conducted simulations at elevated temperatures using a subset of force fields to gain insights into the temperature-dependence and metastability of conformational sampling. The results show that with the GROMOS96 53a6 force field, the Neh2 peptide continued to fold into β-hairpin conformation and remained stable even at higher temperatures. This is quite different from what was observed for Amber ff99SB*-ILDN as the hairpin structure becomes less stable under this force field when the temperature increases. Although native-like conformation was not observed in the microsecond long CHARMM27 (1) simulation at 310 K, the peptide quickly folded into a β-hairpin structure at 330 K. Therefore, the lack of conformational convergence at lower temperature may simply be due to insufficient sampling time. However, further increase in temperature (i.e., 350 and 370 K) again led to the disappearance of β-hairpin structure in the CHARMM27 (1) simulations. The results here also show that, although long simulation times are necessary, it is important to have alternative methods of sampling conformations, such as replica-exchange and related methods.[82−84]

Finally, this and other recent comparative studies[44,45,85] show the importance of using different criteria for assessing the properties of different force fields. In addition to more reliable simulations, such studies provide invaluable information about the collective nonadditive properties of amino acids that are helpful in interpreting experiments.

Information Available. Two videos are available: Video 1: The first and last 10 ns of the Amber ff99SB-ILDN (1), Amber ff99SB*-ILDN, Amber ff99SB (2), Amber ff03, Amber ff03* GROMOS96 43a1p, GROMOS96 53a6 (2), CHARMM27 (2), and OPLS-AA/L trajectories (without terminal capping or charge groups). For clarity, water, ions, and hydrogens are not shown, and rotation and translation of the peptide has been removed. Secondary structures were colored as follows in VMD: yellow, β-sheet (arrows indicate chain direction); purple, α helix; blue, 3_{10} helix; white, coil. Video 2: 0−400 ns of the Amber ff99SB* (2) trajectory. For clarity, water, ions, and hydrogens are not shown, and rotation and translation of the peptide has been removed. Secondary structures were colored as follows in VMD: yellow, β-sheet (arrows indicate chain direction); purple, α helix; blue, 3_{10} helix; white, coil. This material is available free of charge via the Internet at http://www.flickr.com/photos/softsimu/.

AUTHOR INFORMATION

Corresponding Author

*E-mail: mikko.karttunen@uwaterloo.ca (M.K.); jchoy4@uwo.ca (W.-Y.C.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank Martin Ulmschneider for critical reading of this manuscript. SharcNet (www.sharcnet.ca) and SciNet HPC Consortium (www.scinethpc.ca) provided the computational resources, and we thank the Natural Sciences and Engineering Research Council of Canada, Canadian Institutes of Health Research (MOP no. 74679), and the Ontario Early Researcher Program for financial support.

REFERENCES

(1) Lei, H.; Wu, C.; Liu, H.; Duan, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4925−4930.
(2) Ensign, D. L.; Pande, V. S. *Biophys. J.* **2009**, *96*, L53−L55.
(3) Best, R. B.; Mittal, J. *Proteins* **2011**, *79*, 1318−1328.
(4) Zhou, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13280−13285.
(5) Monticelli, L.; Sorin, E. J.; Tieleman, D. P.; Pande, V. S.; Colombo, G. *J. Comput. Chem.* **2008**, *29*, 1740−1752.
(6) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. *J. Am. Chem. Soc.* **2010**, *132*, 1526−1528.
(7) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341−346.
(8) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. *Science* **2011**, *334*, 517−520.
(9) Cino, E. A.; Wong-Ekkabut, J.; Karttunen, M.; Choy, W. Y. *PLoS One* **2011**, *6*, e27371.
(10) Ulmschneider, M. B.; Ulmschneider, J. P. *J. Chem. Theory Comput.* **2008**, *4*, 1807−1809.
(11) Best, R. B.; Buchete, N. V.; Hummer, G. *Biophys. J.* **2008**, *95*, L07−L09.
(12) Freddolino, P. L.; Park, S.; Roux, B.; Schulten, K. *Biophys. J.* **2009**, *96*, 3772−3780.
(13) Mittal, J.; Best, R. B. *Biophys. J.* **2010**, *99*, L26−L28.
(14) Patapati, K. K.; Glykos, N. M. *Biophys. J.* **2011**, *101*, 1766−1771.
(15) Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005**, *88*, 2472−2493.
(16) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(17) Best, R. B.; Hummer, G. *J. Phys. Chem. B* **2009**, *113*, 9004−9015.
(18) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(19) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712−725.
(20) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78*, 1950−1958.
(21) Oostenbrink, C.; Soares, T. A.; van der Vegt, N. F.; van Gunsteren, W. F. *Eur. Biophys. J.* **2005**, *34*, 273−284.
(22) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2004**, *25*, 1656−1676.
(23) Itoh, K.; Chiba, T.; Takahashi, S.; Ishii, T.; Igarashi, K.; Katoh, Y.; Oyake, T.; Hayashi, N.; Satoh, K.; Hatayama, I.; Yamamoto, M.; Nabeshima, Y. *Biochem. Biophys. Res. Commun.* **1997**, *236*, 313−322.
(24) Itoh, K.; Wakabayashi, N.; Katoh, Y.; Ishii, T.; Igarashi, K.; Engel, J. D.; Yamamoto, M. *Genes Dev.* **1999**, *13*, 76−86.
(25) Dunker, A. K.; Obradovic, Z.; Romero, P.; Garner, E. C.; Brown, C. J. *Genome Inf. Ser. Workshop Genome Inf.* **2000**, *11*, 161−171.
(26) Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z. *J. Mol. Graphics Modell.* **2001**, *19*, 26−59.
(27) Dunker, A. K.; Obradovic, Z. *Nat. Biotechnol.* **2001**, *19*, 805−806.
(28) Radivojac, P.; Iakoucheva, L. M.; Oldfield, C. J.; Obradovic, Z.; Uversky, V. N.; Dunker, A. K. *Biophys. J.* **2007**, *92*, 1439−1456.
(29) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321−331.
(30) Tong, K. I.; Katoh, Y.; Kusunoki, H.; Itoh, K.; Tanaka, T.; Yamamoto, M. *Mol. Cell. Biol.* **2006**, *26*, 2887−2900.
(31) Lo, S. C.; Li, X.; Henzl, M. T.; Beamer, L. J.; Hannink, M. *EMBO J.* **2006**, *25*, 3605−3617.
(32) Lo, S. C.; Hannink, M. *J. Biol. Chem.* **2006**, *281*, 37893−37903.
(33) Strachan, G. D.; Morgan, K. L.; Otis, L. L.; Caltagarone, J.; Gittis, A.; Bowser, R.; Jordan-Sciutto, K. L. *Biochemistry* **2004**, *43*, 12113−12122.
(34) Padmanabhan, B.; Nakamura, Y.; Yokoyama, S. *Acta. Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2008**, *64*, 233−238.
(35) Komatsu, M.; Kurokawa, H.; Waguri, S.; Taguchi, K.; Kobayashi, A.; Ichimura, Y.; Sou, Y. S.; Ueno, I.; Sakamoto, A.; Tong, K. I.; Kim, M.; Nishito, Y.; Iemura, S.; Natsume, T.; Ueno, T.; Kominami, E.; Motohashi, H.; Tanaka, K.; Yamamoto, M. *Nat. Cell Biol.* **2010**, *12*, 213−223.
(36) Camp, N. D.; James, R. G.; Dawson, D. W.; Yan, F.; Davison, J. M.; Houck, S. A.; Tang, X.; Zheng, N.; Major, M. B.; Moon, R. T. *J. Biol. Chem.* **2012**, *287*, 6539−6550.
(37) Ma, J.; Cai, H.; Wu, T.; Sobhian, B.; Huo, Y.; Alcivar, A.; Mehta, M.; Cheung, K. L.; Ganesan, S.; Kong, A. N.; Zhang, D. D.; Xia, B. *Mol. Cell. Biol.* **2012**, *32*, 1506−1517.
(38) Blanco, F. J.; Rivas, G.; Serrano, L. *Nat. Struct. Biol.* **1994**, *1*, 584−590.
(39) Fasan, R.; Dias, R. L.; Moehle, K.; Zerbe, O.; Obrecht, D.; Mittl, P. R.; Grütter, M. G.; Robinson, J. A. *ChemBioChem* **2006**, *7*, 515−526.
(40) Stefanovic, M.; Markham, N. O.; Parry, E. M.; Garrett-Beal, L. J.; Cline, A. P.; Gallagher, P. G.; Low, P. S.; Bodine, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 13972−13977.
(41) Best, R. B.; Mittal, J. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11087−11092.
(42) van der Spoel, D.; Lindahl, E. *J. Phys. Chem. B* **2003**, *107*, 11178−11187.
(43) Guvench, O.; MacKerell, A. D. *Methods Mol. Biol.* **2008**, *443*, 63−88.
(44) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *PLoS One* **2012**, *7*, e32131.
(45) Ciarkowski, J.; Łuczak, S.; Jagieła, D.; Sikorska, E.; Wójcik, J.; Oleszczuk, M.; Izdebski, J. *J. Mol. Graphics Modell.* **2012**, *32*, 67−74.
(46) Wong-ekkabut, J.; Miettinen, M. S.; Dias, C.; Karttunen, M. *Nat. Nanotechnol.* **2010**, *5*, 555−557.

(47) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(48) Brunger, A. T. *Nat. Protoc.* **2007**, *2*, 2728−2733.
(49) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(50) Hansson, T.; Nordlund, P.; Aqvist, J. *J. Mol. Biol.* **1997**, *265*, 118−127.
(51) Christen, M.; Hünenberger, P. H.; Bakowies, D.; Baron, R.; Bürgi, R.; Geerke, D. P.; Heinz, T. N.; Kastenholz, M. A.; Kräutler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; van Gunsteren, W. F. *J. Comput. Chem.* **2005**, *26*, 1719−1751.
(52) Bjelkmar, P.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E. *J. Chem. Theory Comput.* **2010**, *6*, 459−466.
(53) 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. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(54) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(55) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(56) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(57) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. *Biophys. J.* **2011**, *100*, L47−L49.
(58) Craft, J. W.; Legge, G. B. *J. Biomol. NMR* **2005**, *33*, 15−24.
(59) Sapay, N.; Tieleman, D. P. *J. Comput. Chem.* **2011**, *32*, 1400−1410.
(60) Patra, M.; Karttunen, M. *J. Comput. Chem.* **2004**, *25*, 678−689.
(61) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
(62) Parrinello; Rahman. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(63) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Johannes, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(64) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(65) Patra, M.; Karttunen, M.; Hyvönen, M. T.; Falck, E.; Lindqvist, P.; Vattulainen, I. *Biophys. J.* **2003**, *84*, 3636−3645.
(66) Patra, M.; Hyvonen, M. T.; Falck, E.; Sabouri-Ghomi, M.; Vattulainen, I.; Karttunen, M. *Comput. Phys. Commun.* **2007**, *176*, 14−22.
(67) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullmann, B., Ed.; Reidel: Dordrecht, 1981; p 331.
(68) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(69) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97−179.
(70) Murzyn, K.; Zhao, W.; Karttunen, M.; Kurdziel, M.; Róg, T. *Biointerphases* **2006**, *1*, 98−105.
(71) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577−2637.
(72) Tong, K. I.; Kobayashi, A.; Katsuoka, F.; Yamamoto, M. *Biol. Chem.* **2006**, *387*, 1311−1320.
(73) Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. *Biophys. J.* **2008**, *94*, L75−L77.
(74) Cao, Z.; Liu, L.; Wang, J. *J. Biomol. Struct. Dyn.* **2011**, *29*, 527−539.
(75) Chan, H. S.; Shimizu, S.; Kaya, H. *Methods Enzymol.* **2004**, *380*, 350−379.
(76) Dill, K. A.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10−19.
(77) Fuxreiter, M.; Simon, I.; Friedrich, P.; Tompa, P. *J. Mol. Biol.* **2004**, *338*, 1015−1026.
(78) Gall, C.; Xu, H.; Brickenden, A.; Ai, X.; Choy, W. Y. *Protein Sci.* **2007**, *16*, 2510−2518.
(79) Sivakolundu, S. G.; Bashford, D.; Kriwacki, R. W. *J. Mol. Biol.* **2005**, *353*, 1118−1128.
(80) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23*, 2947−2948.
(81) Henikoff, S.; Henikoff, J. G. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10915−10919.
(82) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058−9067.
(83) Lee, M. S.; Olson, M. A. *J. Chem. Phys.* **2011**, *134*, 244111.
(84) Chaudhury, S.; Olson, M. A.; Tawa, G.; Wallqvist, A.; Lee, M. S. *J. Chem. Theory Comput.* **2012**, *8*, 677−687.
(85) Beauchamp, K. A.; Lin, Y.-S.; Das, R.; Pande, V. S. *J. Chem. Theory Comput.* **2012**, *8*, 1409−1414.