

# **Exploiting a Mechanical Perturbation of Titin Domain to Identify How Force Field Parameterization Affects Protein Refolding Pathways**

David Wang\* and Piotr E. Marszalek\*

*Department of Mechanical Engineering and Materials Science, Duke University, Durham,  
NC 27708, USA*

E-mail: david.wang3@duke.edu; pemar@duke.edu

## Abstract

Molecular mechanics force fields have been shown to differ in their predictions of processes such as protein folding. To test how force field differences affect predicted protein behavior, we created a mechanically perturbed model of the beta-stranded I91 titin domain based on atomic force spectroscopy data and examined its refolding behavior using six different force fields. To examine the transferability of the force field discrepancies identified by this model, we compared the results to equilibrium simulations of the weakly helical peptide Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub>. The total simulation time was 80  $\mu$ s. From these simulations we found significant differences in I91 perturbation refolding ability between force fields. Concurrently, Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> equilibration experiments indicated that although force fields have similar overall helical frequencies, they can differ in helical lifetimes. The combination of these results suggests that differences in force field parameterization may allow a more direct transition between the beta and alpha regions of the Ramachandran plot thereby affecting both beta-strand refolding ability and helical lifetimes. Furthermore, the combination of results suggests that using mechanically perturbed models can provide a controlled method to gain more insight into how force fields affect protein behavior.

## Introduction

Molecular dynamics simulations describe macromolecular motion in all-atom detail, a feat that is extremely difficult to achieve experimentally.<sup>1-5</sup> From these simulations, a wide variety of thermodynamic and kinetic observables can be inferred thereby providing a better picture of the nature of various biological processes such as protein folding.<sup>6-12</sup> The utility of molecular dynamics simulations however, depends on the accuracy of its force field - the set of potentials used to describe the energetics behind atomic movements. Due to the complexity of the parameterization process, various force fields are not guaranteed to exhibit the same effect on biomolecular behavior.<sup>13-26</sup> This has been seen previously when modifica-

tions to torsion angle parameters have been shown to impede folding and generate incorrect structures.<sup>27-29</sup> Thus force field variations can affect sensitive biomolecular processes such as protein folding pathways.<sup>2,7,12,15,30,31</sup>

Although current validation schemes of molecular mechanics force fields use a wide variety of experimental data to ensure transferability,<sup>32-37</sup> these validations often occur over small peptides. Other studies involving the folding of complete proteins are often limited by simulation time scales<sup>19,20,33</sup> or involve the use of enhanced sampling techniques like replica exchange.<sup>2,7,9,32</sup> Thus, examining more detailed pictures of processes like protein folding that could also be simulated in a small time would be helpful in identifying force field discrepancies as these processes are more sensitive to force field inaccuracies than validation and comparison studies on native state structures.<sup>9,12,30</sup>

One potential way to experimentally observe the behavior of macromolecular structures is through atomic force spectroscopy.<sup>38-40</sup> With atomic force microscopes, various proteins can be unfolded using force. Indeed, certain portions of proteins have been observed to unfold under force and refold after force relaxation.<sup>38,41,42</sup> Thus, a force perturbation on a small portion of protein could serve as an ideal model to examine how various force fields comparatively affect biomolecular motion. Since the perturbations refold after force relaxation, we expect that a reconstruction of the structural perturbation will refold if the molecular mechanics force field can accurately reproduce the perturbation.

In particular, the structural perturbations involved in many atomic force microscopy refolding experiments are relatively tiny compared to the overall structure.<sup>38,42</sup> Therefore, by the thermodynamic hypothesis it is expected that structural perturbations will robustly refold into the native structure.<sup>5,29,43</sup> Furthermore the small size of the perturbation allows for faster computation because of the reduction of the size of the water box. Along with an increase in speed of computation, the perturbations should also be more sensitive to force field discrepancies as the perturbations should refold from the unfolded state to the native structure.<sup>9,12,30,44</sup>

From these considerations, we created a model of a mechanically perturbed protein from titin I91 (formerly I27), a domain of the titin protein that has been extensively studied using force spectroscopy.<sup>26,45–49</sup> Under force, the first ten residues of I91 comprising beta strand A have been shown to unfold and refold upon force extension.<sup>38</sup> Thus, I91 provides an opportunity to test the refolding of beta proteins with various force fields. In particular, this mechanical perturbation provides a more controlled method to examine the refolding of a single beta strand as compared to the refolding of beta-hairpins. Previous literature has indicated that beta-hairpin refolding occurs under multiple stages and the particulars of the folding process are still contested.<sup>13,19</sup> Furthermore, beta-strand refolding may require timescales greater than the simulation timescale of 1  $\mu s$ .<sup>7,9,13,19</sup> Thus, by examining the movement of a single strand, this mechanical perturbation provides a more controlled and faster method to examine the dynamics of beta strands.

The Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> helical peptide was chosen to model the folding and unfolding of helix structures. This helical peptide has been used as a benchmark to determine empirical corrections to various force fields<sup>30,32,34</sup> and previous NMR data also provides an empirical observable with which helical frequencies can be validated.<sup>50</sup> The model provides a way to test whether force field discrepancies observed with the I91 perturbation are transferable and whether such observations can provide new insight into the folding process.

We tested five different force field/water model combinations using Amber simulation software: Amber ff14SB/Tip3p,<sup>33</sup> Amber ff14SB without backbone dihedral correction/Tip3p (abbreviated as Amber ff14SBonlySC),<sup>33</sup> Amber FB15/Tip3p,<sup>36</sup> Amber ff99SB\*-ILDN/Tip3p,<sup>32</sup> and Amber ff99SB-ILDN/Tip3p;<sup>1</sup> and two different force field/water model combinations using Desmond: Charmm 22\*/Tip3p-Charmm<sup>30</sup> and Amber ff99SB-ILDN/Tip3p<sup>1</sup> for time spans amounting to at least 1  $\mu s$  per simulation. The 1  $\mu s$  timescale allowed for the observation of folding events while being short enough to run repeated trials to ensure robust sampling of the unfolded state. From these simulations, we observed differences in the I91 perturbation refolding pathway which were then supported by Ac-(AAQAA)<sub>3</sub>NH<sub>2</sub> equilibra-

tion simulations. This suggests that mechanical perturbations may serve as good models to examine the effect of force field discrepancies.

## Simulation Methodology

In order to generate the structure perturbation, steered molecular dynamics simulations were conducted with NAMD.<sup>51</sup> The N-terminus of the I91 domain (using chain A of pdb structure 1waa, residues 1-89<sup>45</sup>) was pulled in the upwards z-direction away from the rest of the protein using a previously established protocol.<sup>46</sup> A structure was selected when residues 1-10 had lost contact with the rest of the I91 domain. The upward z-direction was chosen in order to examine the behavior of the strand in the unfolded state as z-directional pulling leaves more space for the strand to sample various configurations. Although the direction of pulling is not the same as that done during experiment, we expect that given the small size of the perturbation and the downhill nature of the folding process<sup>29,43</sup> that the perturbation will still exhibit robust refolding. The selection of these structures was then used as the starting point of equilibration simulations.

Equilibration simulations of the I91 perturbation were performed using two different software packages. A water box was constructed around the initial structure with a buffer of 10 Å on each side. In order to test the Charmm 22\* and Amber ff99SB-ILDN force fields, the Desmond software package was used alongside GPU acceleration.<sup>52</sup> Using Desmond, the initial perturbed structure was equilibrated at 310K with a Nose-Hoover thermostat in an NPT ensemble for an initial 1 $\mu$ s of simulation time. If the structure did not reach the folded state by the end of the simulation, the simulation was continued for a second 1 $\mu$ s of simulation time. Five trials were performed for both the Charmm 22\* and Amber ff99SB-ILDN force field with each trial changing the randomization seed to provide independence. In order to test the Amber ff14SB, Amber ff14SBonlySC, Amber FB15, Amber ff99SB\*-ILDN, and Amber ff99SB-ILDN force fields the Amber simulation package was used alongside

GPU acceleration<sup>53</sup> Using Amber, the perturbed structure was equilibrated at 310K with a Langevin thermostat (friction of  $2 \text{ ps}^{-1}$ ) in an NPT ensemble for an initial  $1 \mu\text{s}$  of simulation time. Similarly, if the structure did not reach the folded state by the end of the initial  $1 \mu\text{s}$  of simulation time, the simulation would be continued for another  $1 \mu\text{s}$ . Five trials were performed for each of the Amber ff14SB, Amber ff14SBonlySC, Amber ff99SB\*-ILDN, Amber ff99SB-ILDN and Amber FB15 force fields.

Equilibration simulations of Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> were performed at 303K to test whether differences in force fields captured by the mechanical perturbation refolding simulations could also be observed in Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> helical distributions. For simulations using Amber software, the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> structure was created as a linear peptide using tleap. The structure was placed in a water box with a  $20 \text{ \AA}$  buffer and equilibrated using an NPT ensemble with a Langevin thermostat (friction of  $1 \text{ ps}^{-1}$ ) for  $2 \mu\text{s}$  for the following force fields: Amber ff14SB,<sup>33</sup> Amber ff14SBonlySC,<sup>33</sup> Amber ff99SB\*-ILDN,<sup>32</sup> Amber FB15,<sup>36</sup> and Amber 99SB-ILDN.<sup>54</sup> For simulations using Desmond software, the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> structure was again created as a linear peptide using tleap and run in an NPT ensemble with a Nose-Hoover thermostat for  $2 \mu\text{s}$  for the following force fields: Charmm 22\*<sup>30</sup> and Amber ff99SB-ILDN.<sup>1</sup> Unlike Amber software, Desmond software does not allow the automatic addition of NH<sub>2</sub> caps therefore the c-terminus of Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptide was manually adjusted to fit a residue template in the Charmm 22\* force field (see Supplemental Information), while in the Amber ff99SB-ILDN force field an NMA cap was used instead.

To determine how force fields affected the I91 perturbation refolding process, molecular dynamics trajectories were divided according to whether the structure was in the folded or unfolded states. To locate the folded and unfolded states, RMSD and Q-value plots for the 10 perturbed residues of the I91 perturbation were created with cutoffs of 0.4 and 0.85 being used for the folded and unfolded states respectively (see Supplementary Information).<sup>29</sup> Upon identification of the time of folding, the structures in the unfolded state were grouped from each of the five trials and analyzed to see whether force fields affected the unfolded

state. Analysis algorithms were used to produce the potential of mean force distribution,<sup>55</sup> secondary structure percentage,<sup>56</sup> and clustered structures.<sup>57</sup> Secondary structure percentage was analyzed via DSSP<sup>56</sup> using MDTraj.<sup>58</sup> Clustering of the unfolded structures was done using the KMedoids and KCenters algorithms as implemented in MSMBuiler<sup>59</sup> and MDTraj.<sup>58</sup> Arrangement of clusters for visualization was made using the Multialign viewer<sup>60</sup> in UCSF Chimera.<sup>61</sup>

To examine Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> helicity we used previously established criteria to define the a<sub>h</sub> region:  $\phi \in [-100^\circ, -30^\circ]$  and  $\psi \in [-67^\circ, -7^\circ]$  and counted a residue as helical if three consecutive residues were found in the a<sub>h</sub> region.<sup>32,34</sup> We computed the frequency with which each individual residue was deemed helical in order to match previously published results and NMR data.<sup>30,32,34,50</sup> The helicity of each Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> residue as a function of time was computed using DSSP criteria<sup>56</sup> with an implementation found in MDTraj.<sup>58</sup> The lifetime of each helix was defined as the length of time which a set of helical residues and their adjacent partners could propagate in the secondary structure vs time plot generated by DSSP (see Supplemental Information). These counts were then used to compute the mean helical lifetimes and max helical lifetimes of the simulation. The helical lifetime per residue was defined as the length of time with which a specific residue could remain helical (see Supplemental Information). From this set of times, the mean and max helical lifetimes per residue were also computed. Finally, dihedral potential of mean force plots were created in order to examine the distribution of  $\phi$   $\psi$  values found in the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> equilibrium simulation distributions.<sup>55</sup>

## Results

### I91 refolds robustly under most force fields

From the simulations it can be seen that the perturbed structure of I91 domain refolds robustly under most force fields (see Table 1). For most force fields, the refolding times

range from tens of nanoseconds to over 1  $\mu s$ . This distribution suggests that while the given 2  $\mu s$  of simulation time is feasible to capture some refolding events, the true range of refolding times is somewhat greater than the simulation time of 2  $\mu s$  which was the maximum possible simulation time given the computational power at our disposal. There were noticeable differences in the ability of various force fields to capture refolding events. In particular, the Amber ff14SB and Amber ff99SB\*-ILDN force fields displayed difficulty refolding the perturbed intermediate in the allotted time (see Table 1). This difference in refolding capacity suggests that differences in force field parameterization may affect the refolding pathways of I91. Since the unfolded state is more sensitive to force field discrepancies, we then examined the unfolded state to see if different force fields alter the unfolded state of the I91 mechanically perturbed intermediate.

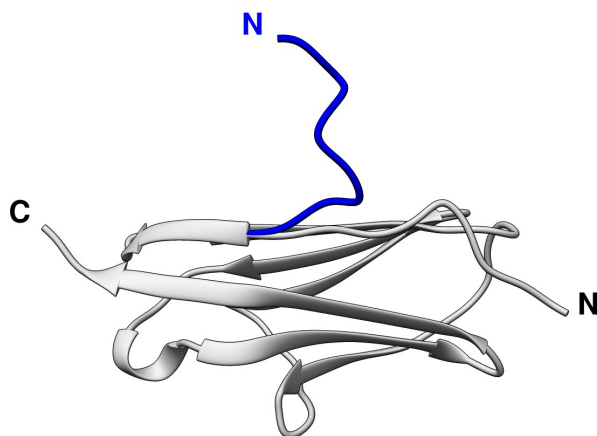


Figure 1: I91 Mechanical Perturbation. The native structure of I91 (PDB ID: 1waa<sup>45</sup>) is shown in gray while the mechanical perturbation is shown in blue.

After separating the molecular dynamics trajectories into folded and unfolded states, we examined the distribution of dihedral angles in the unfolded state. Potentials of mean force were constructed in order to examine the magnitude with which the dihedral angles were found in various states (see Figure 2). From these potential of mean force diagrams it can be seen that there are significant minima in the helical regions of the Charmm 22\*, Amber



**Table 1: I91 Perturbation Refolding Simulation Summary**

Force Field	Trials Refolded	Refolding Times (ns)
Charmm 22*	5/5	498, 749, 1089, 346, 1136
Amber ff99SB-ILDN (Desmond)	2/5	1372, 1270
Amber ff99SB-ILDN (Amber)	3/5	411, 833, 455
Amber FB15	4/5	1384, 1208, 890, 57
Amber ff14SB	1/5	600
Amber ff14SBonlySC	3/5	852, 384, 51
Amber ff99SB*-ILDN	2/5	439, 1602

ff99SB-ILDN, Amber ff99SB\*-ILDN, and Amber ff14SB force fields. The difference between the Amber ff14SB and Amber ff14SBonlySC force fields indicates that the empirical dihedral correction of Amber ff14SB force field sharply increases the amount of helical torsion angles found within the unfolded region by almost  $2 k_B T$ . Additionally, it is interesting that while the Charmm 22\* force field appears to refold I91 robustly, the Amber ff99SB\*-ILDN and Amber ff14SB force fields struggle to refold the perturbation. Since the Charmm 22\* force field also contains a significant population of dihedral angles in the alpha helical region, this suggests that simply possessing residues with helical dihedral angles may not actually affect refolding.

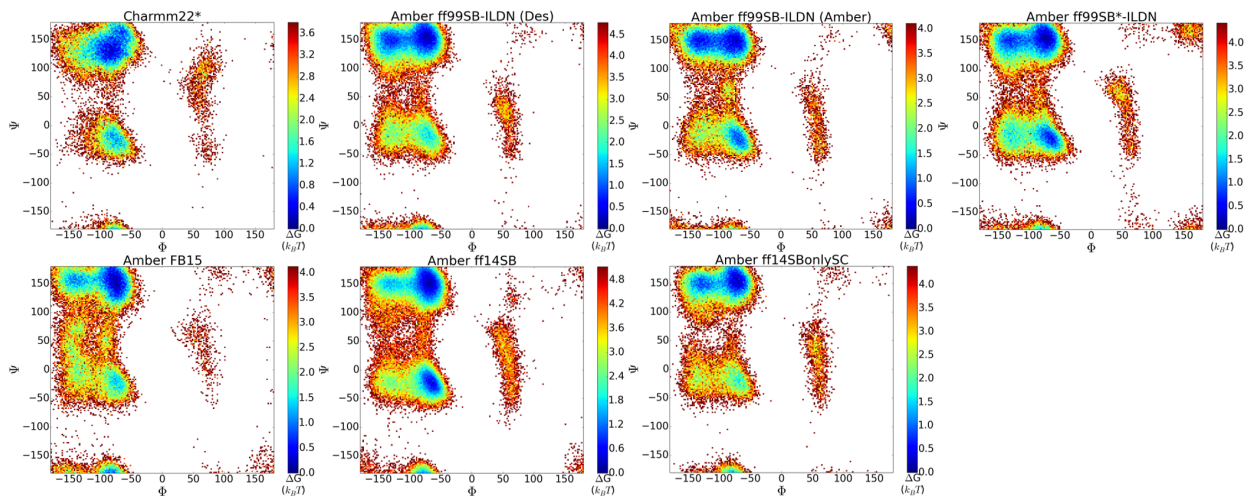


Figure 2: Potentials of mean force for the unfolded state of the I91 perturbation for each force field.

Another interesting point of comparison lies between the Amber FB15 force field and all the other force fields. Dihedral sampling with the Amber FB15 force field for the I91 unfolded state is much more prevalent in the region between the alpha and beta regions of the Ramachandran plot ( $\psi \in [25^\circ, 75^\circ]$ ). The potential of mean force plot for the unfolded state of the Amber FB15 force field then samples dihedral pairs which are not found in any of the other force fields (see Figure 2). This may be related to the parameterization method of the Amber FB15 force field which reparameterized all the bonded terms of the Amber ff99SB force field based on an automated methodology,<sup>35</sup> notably relaxing bond force constants from their equilibrium values by up to an extra 10 percent.<sup>36</sup> The effects of this methodology are visible in the I91 potential of mean force map as the Amber FB15 force field samples more regions of the Ramachandran plot than any other force field.<sup>13,36</sup>

Finally, a point of comparison arises between the Amber ff99SB-ILDN force field simulations conducted with Desmond software and the Amber ff99SB-ILDN force field simulations conducted with Amber software. In comparison to the rest of the potential of mean force plots, the Amber ff99SB-ILDN potential of mean force plot is similar in shape to both the Amber ff14SB and Amber ff14SBonlySC potential of mean force plots. This is because the same backbone dihedrals are used for all three force fields - each being based on the Amber ff99SB backbone dihedrals.<sup>1,33,62</sup> By comparing the two Amber ff99SB-ILDN force fields together it becomes apparent that the Amber ff99SB-ILDN force field simulations conducted with Amber possess a greater amount of helical torsion angles than the Amber ff99SB-ILDN force field using Desmond. However, given the higher amount of refolding events produced by the Amber ff99SB-ILDN force field simulations conducted with Amber, we believe that this is more likely to be the result of differences due to sampling than it is of any innate differences between the software packages.

Even though dihedral maps give a good idea of which residues possess helical or beta-sheet characteristics the actual formation of helices and beta-sheets does not occur until multiple residues have adopted the same repeating dihedral angles. Thus, in order to examine the

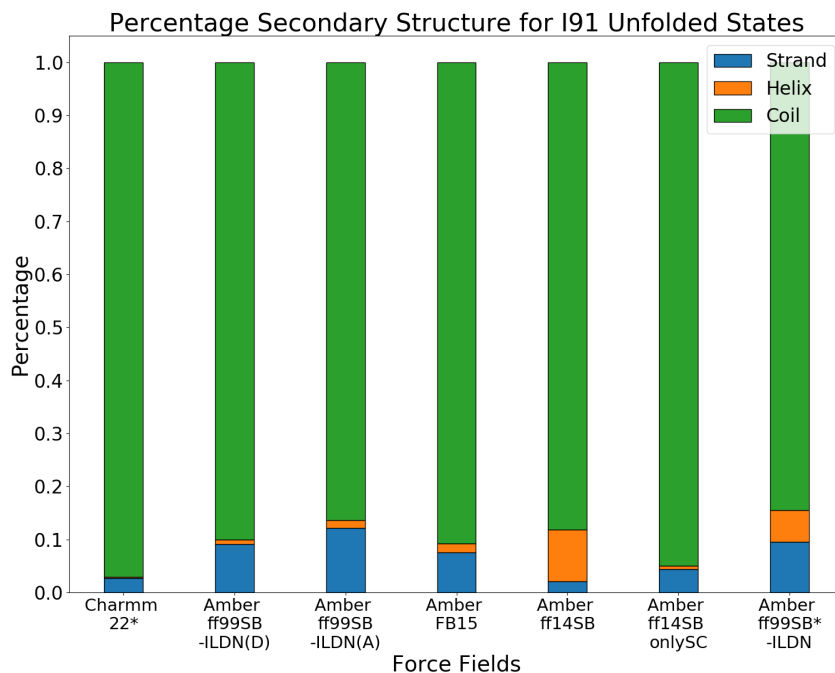


Figure 3: Secondary structure percentages for the unfolded states of the I91 perturbation. A reduced characterization is shown where a helix is either an alpha, 3/10, or pi helix; a strand is either a beta bridge or extended strand; and a coil is either a turn, bend, or irregular element.

actual secondary structure content of the unfolded state. We used DSSP to assign the secondary structure of each residue and then compiled the secondary structure frequencies together from all the unfolded states of all the trials (see Figure 3). By comparing the various secondary structures for the unfolded state of the I91 perturbation we can see that the Amber ff14SB and Amber ff99SB\*-ILDN force fields possess a much higher frequency of helical residues than the rest of the force fields which produce almost no helical residues. This supports the potential of mean force plots (see Figure 2) which suggest large minima in the helical regions of both Amber ff14SB and Amber ff99SB\*-ILDN. However, the absence of helices notably in Charmm 22\* suggests that residues simulated using the Charmm 22\* force field could easily leave the alpha helical minima region and did not remain in the helical configuration long enough for the formation of stable helices.

Further comparisons can be made between the Amber FB15 force field and the rest of

the structures. Overall, the Amber FB15 force field displays roughly similar amounts of secondary structure as the other force fields. This suggests that although the dihedral angle distributions are different, the FB15 force field will end up producing roughly equivalent amounts of secondary structure as the rest of the force fields. It is interesting to note that although most Amber force fields have a significant amount of extended conformation, that the Amber ff14SBonlySC and Charmm 22\* force fields both possess low amounts of extended conformation. This suggests that the force fields do not form strands in the unfolded state and only adopt the secondary structure when folded into the native state.

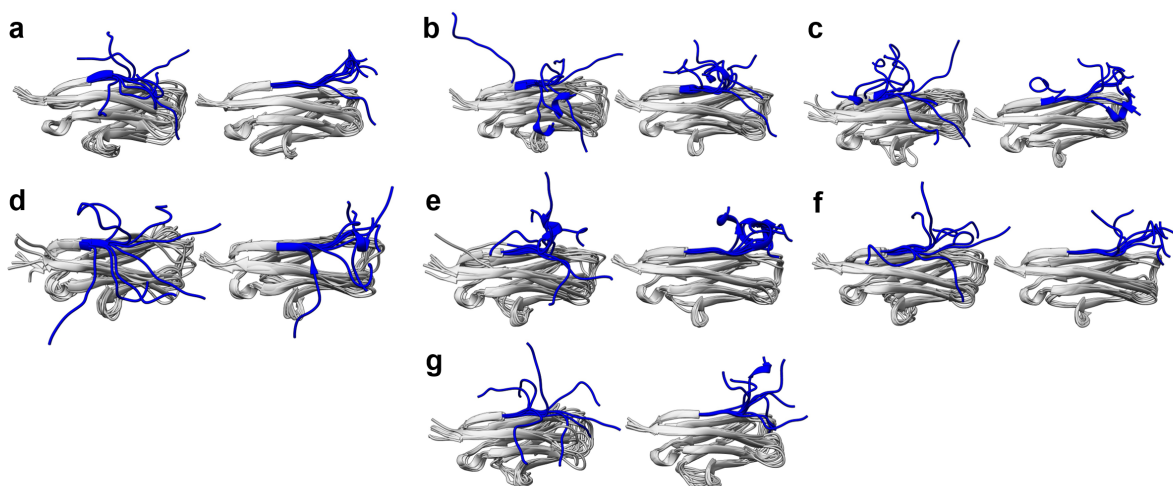


Figure 4: Clusters of the unfolded states of the I91 perturbations for each force field. K-Centers clusters are found to the left and K-Medoids clusters are found to the right. The force fields used are as follows: a) Amber ff99SB-ILDN (in Desmond) b) Amber ff99SB-ILDN (in Amber) c) Amber ff99SB\*-ILDN d) Amber FB15 e) Amber ff14SB f) Amber ff14SBonlySC g) Charmm 22\*

After identifying common secondary structures, we wished to examine the most common structures sampled by each of the force fields. To that end we used the KMedoids and KCenters clustering algorithms to determine the most common structures of the unfolded state. Two different clustering algorithms were used because the KCenters clustering algorithm commonly picks up outliers in the structures and was thus used to identify the range of structures that were sampled in the unfolded state.<sup>57</sup> Meanwhile, the KMedoids algorithm

was used to find the most common unfolded state structures and thus the structures that were most likely to be found in local minima.

Through this method we observed that the Amber ff14SB force field, while sampling much of the same space as the other force fields (see KCenters structures for Amber ff14SB in Figure 4e), possesses many structures found in helices as seen with the KMedoids algorithm (see Figure 4e). This corroborates the potential of mean force plots and secondary structure analysis using DSSP. The heavy amount of helical structures suggests that the sampling of this minima with the Amber ff14SB force field might temporarily impeded the creation of beta structures thereby altering the refolding pathway of I91. The presence of helical segments in the Amber ff99SB\*-ILDN force field also supports the refolding times of this force field (see Table 1). The various helical residues of the Amber ff99SB\*-ILDN force field may also impede refolding and cause the formation of unnecessary local minima.

The observation that the Amber FB15 force field samples a wider space on the Ramachandran map is corroborated by the KCenters algorithm clusters for the Amber FB15 force field (see Figure 4d). The KCenters algorithm for Amber FB15 identifies clusters which exist in a wide variety of different configurations that the other force fields do not seem to sample. This suggests that the Amber FB15 force field does increase the space which the perturbations sample. Overall, it can be seen that most structures sample close to the native structure as seen through the KMedoids algorithm. This suggests that the pathways for the folding of the I91 perturbation are limited. Thus, the I91 perturbation is a useful tool for examining the refolding of peptides because force fields must go through one pathway in order to find the correct native structure.

## **Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> Mean Helical Lifetime Can Prove Useful In Identifying Force Field Discrepancies**

To examine whether the results of I91 perturbed intermediate refolding could be replicated in other peptides and linked to experimental observables, we also ran Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub>

equilibration simulations. Previously, the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptide was found to form low amounts of helical structure that sharply decreased as temperature increased.<sup>32,50</sup> By running equilibration simulations under previously described conditions<sup>30,32</sup> we hoped to to examine how our simulation methods compared against previously published results as well as add data for newly tested force fields. Furthermore, we hoped that examining Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> refolding patterns might provide further insight into possible force field discrepancies.

We first analyzed the helical frequency of each residue in the same manner as performed in previous literature.<sup>30,32,32,44</sup> We found that the Amber ff99SB\*-ILDN and Amber FB15 force fields accurately reproduce the 19% helical frequency per residue as observed by NMR experiments (see Figure 5).<sup>34,50</sup> Furthermore, the Amber ff99SB-ILDN force fields for both Desmond and Amber do not produce many helices, hovering at less than 10% for each residue. This supports the accuracy of our simulation methods as it mirrors previously published results.<sup>32,34</sup> The Amber ff14SBonlySC force field also produces this low amount of helical content which is understandable since the dihedral terms were not modified for this force field and should be identical to that of Amber ff99SB-ILDN.

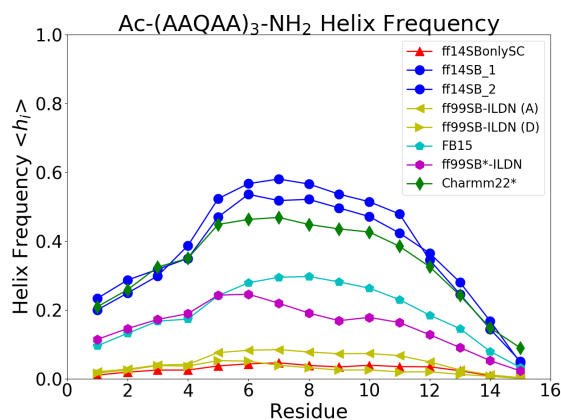


Figure 5: Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> helicity per residue. The helicities were determined with previously established criteria<sup>32</sup>

In contrast to the low helicities of the Amber ff99SB-ILDN and Amber ff14SBonlySC force fields, the Amber ff14SB force field displays a sharp increase in the amount of helical

content. This is corroborated by the I91 perturbed intermediate model which suggested that the empirical dihedral correction for Amber ff14SB created an unexpectedly deep minima in the alpha helical region of the Ramachandran plot. More surprisingly however, the Charmm 22\* force field possesses a helical frequency very similar to that of Amber ff14SB. Such a result is interesting because the Charmm 22\* and Amber ff14SB force fields display drastically different refolding pathways in the I91 perturbation. Specifically, the Amber ff14SB force field displayed large amounts of helices during I91 perturbation refolding while the Charmm 22\* force field did not (see Figure 3). Why then, would the Charmm 22\* force field not form helices during I91 perturbation refolding when it does form a large amount of helices in the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptide?

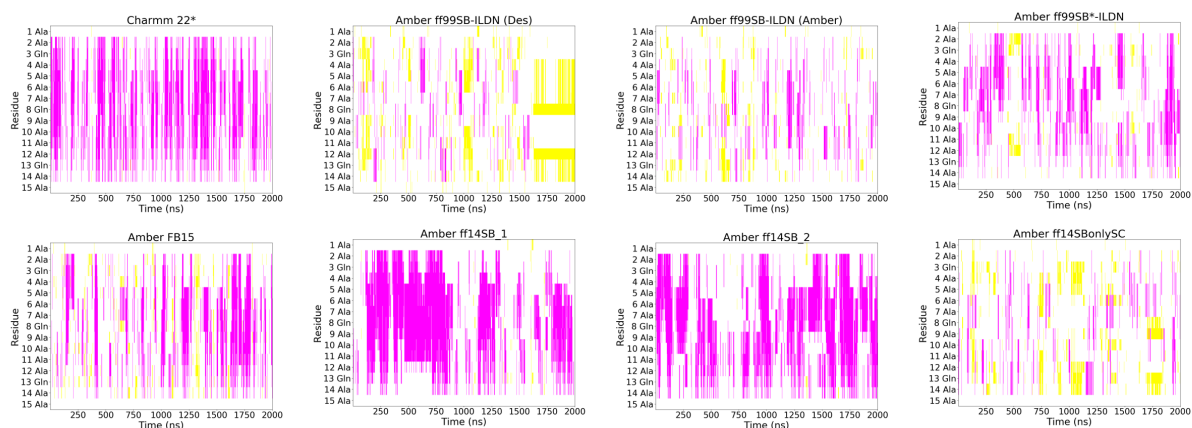


Figure 6: Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> secondary structure formation with respect to time for each force field. Magenta represents helical structures, yellow represents helical structures, and white represents coiled structures. Secondary structures were generated using DSSP with abbreviated criteria<sup>58</sup>

Although different force fields may have similar overall helical frequencies, how those aggregate helical frequencies are created might differ. A force field with two stretches of very stable helix and a force field with lots of small stretches of unstable helix may end up with the same helical frequency over a long period of time. Thus we examined the change in helicity with respect to time to provide a more detailed analysis of the helical frequencies found during Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> equilibration (see Figure 6).



From the graphs of secondary structure with respect to time (see Figure 6) it can be seen that while the Amber ff14SB force field possesses larger amounts of stable helices, the Charmm 22\* force field possesses almost as many helices except that they form and break more quickly. This suggests that the primary difference between the Amber ff14SB and Charmm 22\* force fields is that the Charmm 22\* force field does not stabilize formed helices nearly as well as the Amber ff14SB force field does. It also suggests that the Charmm 22\* force field is more likely to initiate the formation of helices than the Amber ff14SB force field.

Additionally, the Amber ff99SB\*-ILDN force field also displays large intervals of stable helical content interspersed with smaller amounts of helices. Similar behavior can be found in the Amber FB15 force field (see Figure 6). From these graphs it might be suggested that the lack of helical initiation might be what separates the Amber FB15 and Amber ff99SB\*-ILDN force fields from the Charmm 22\* force field rather than the stability of individual helices.

Another comment must be made on the Amber ff99SB-ILDN force fields for Desmond and Amber. In a large portion of the Desmond simulation, the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptide forms a beta hairpin which does not break. This is not expected since the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptide should form alpha helices. However, we believe that the discrepancy is the result of either sampling or the use of the NMA cap instead of the NH<sub>2</sub> cap in the peptide simulated using Desmond rather than any differences between the simulation packages. This is because Amber ff99SB-ILDN has already been shown to possess extremely low amounts of helical content. Therefore, it could be suggested that this might translate to an observation of a beta structure. Furthermore, although less likely, the presence of an NMA cap could also create differences in folding behavior.<sup>19</sup>

To quantify differences in helical content over time we counted the helices formed in the secondary structure vs time graphs created using DSSP (see Figure 7). After identifying each helix that was formed, the mean helical lifetime ]was calculated for each force field. Further-



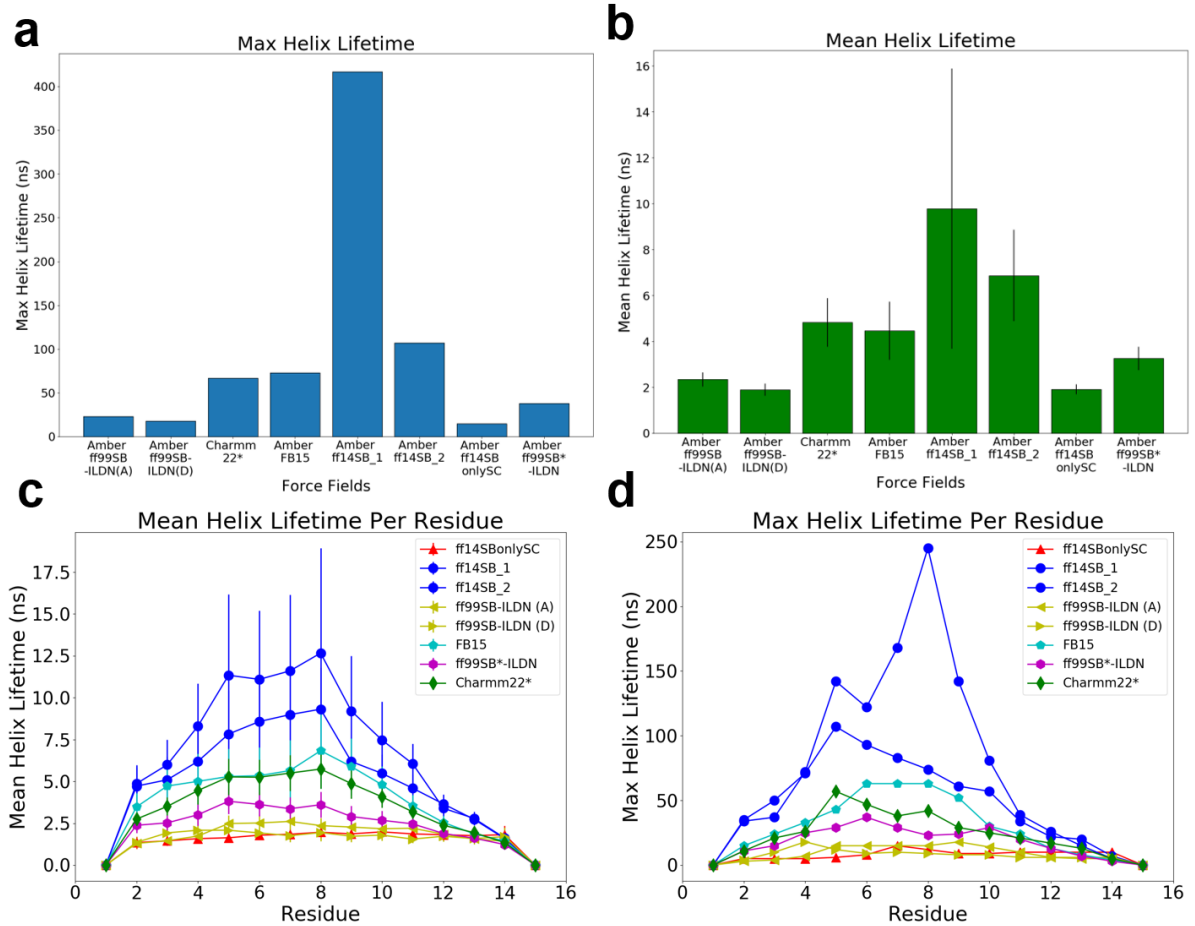


Figure 7: Examination of differences in Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> helical lifetimes. A) The mean helix lifetime for each force field B) The maximum helix lifetime for each force field C) The mean helix lifetime for each residue in each force field D) The maximum helix lifetime for each residue in each force field

more, the mean helical lifetime of each residue of Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> was also determined. We performed this extra analysis because one helical region may have individual residues breaking apart and reforming. In such cases, examining the mean helical lifetime of each residue would give a better picture of the dynamics of individual amino acids in a helix for a given force field. The max helical lifetimes overall and for each residue were also counted because the distribution of helix lifetimes is skewed toward very short helical lifetimes and the maximum helical lifetimes gives an idea of the range of helical lifetimes found for each force field.

Calculating helical lifetimes confirms several observations gained qualitatively from observing the secondary structure of Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> over time (see Figure 6). By examining helical lifetime rather than overall helicity, the helicity of Charmm 22\* lowers considerably (see Figure 7). While examining helical frequency, Charmm 22\* would have similar frequencies to Amber ff14SB. When examining mean helical frequency however, the Charmm 22\* equilibration simulation of Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> reveals helical lifetimes equal to or lower than Amber FB15, a force field which had helical frequency about equal to the experimentally parameterized value. This suggests that Charmm 22\* is increasing the initiation of helix formation rather than maintaining helical stability. Although the error bars of the mean helix lifetimes between Charmm 22\* and Amber ff14SB overlap, a t-test combining both Amber ff14SB simulations gives a p-value of 0.057 which we believe is likely to be significant.

The difference in mean helical lifetimes does not explain the difference in I91 perturbation refolding of the Amber ff99SB\*-ILDN force field compared to the Amber FB15 or Charmm 22\* force fields. This can be seen in the analysis of differences in helical lifetimes. While the Charmm 22\* and Amber FB15 force fields have similar lifetimes to that of Amber ff99SB\*-ILDN, the Amber ff99SB\*-ILDN force field forms far more helices than Charmm 22\* or Amber FB15 in the I91 perturbed unfolded state. This is likely due to the difference in sequences between the two simulated structures. While the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptide does not contain any lysine or glutamic acid residues, the I91 perturbation contains two glutamic acid residues and one lysine. These residues have been shown to be overly helical in the Amber ff99SB\* force field.<sup>44</sup> Since the Amber ff99SB\*-ILDN force field does not reparameterize the lysine or glutamic acid residues we would expect this result to be transferable. Thus, the change in systems may explain why helical constructs are observed in the I91 perturbation but not during Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> equilibration simulations.

To examine force field parameterization differences which might lead to differing Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> and I91 perturbation folding results, we constructed potentials of mean force for each of the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> simulations. Through these potential of mean force

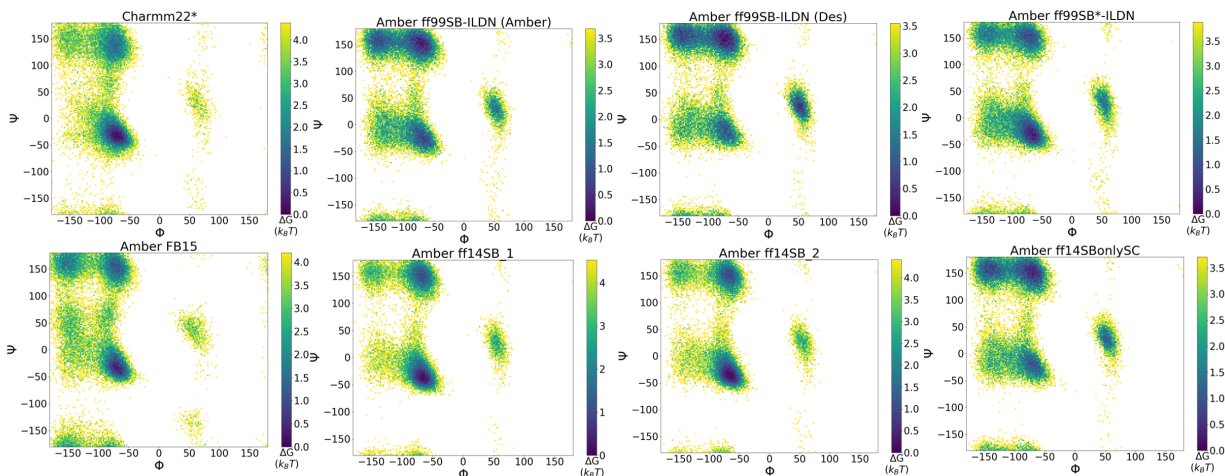


Figure 8: Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> potentials of mean force created from the dihedral angles found in equilibration simulations in each of the observed force fields

plots we hoped to identify regions of the Ramachandran plot which a force field might sample that might explain the discrepancies between the I91 Perturbation and Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> results (see Figure 8).

One notable difference occurs between the helical region of the Charmm 22\* force field and the Amber force fields. While the Charmm 22\* force field possesses one minimum for the alpha helical region of the Ramachandran plot, all Amber force fields possess at least two minima for the alpha helical region. All of the force fields tested have a minimum at the  $a_h$  region:  $\phi \in [-100^\circ, -30^\circ]$  and  $\psi \in [-67^\circ, -7^\circ]$ . Amber force fields however, possess an additional region of low energy found between  $\phi \in [-160^\circ, -100^\circ]$  and  $\psi \in [-50^\circ, 50^\circ]$  which is around  $2 k_B T$  greater than the minimum found in the  $a_h$  region (see Figure 2 and Figure 8).

Additionally, in the area between the beta and alpha regions ( $\psi \in [25^\circ, 75^\circ]$ ), the Charmm 22\* force field potential of mean force plot possesses a gradient that is much more widespread and evenly distributed than in any of the other potential of mean force plots (see Figure 8). We further note that the area between the alpha and beta region varies between each of the Amber force fields. The Amber FB15 force field appears to have the most widespread sampling of the dihedral angles in this region while the Amber ff14SB force field samples the region the least. Indeed, the Amber ff14SB force field barely samples the region  $\phi \in$

$[-160^\circ, -100^\circ]$  and  $\psi \in [-50^\circ, 50^\circ]$  likely as a result of the increased barrier provided by the Amber ff14SB empirical backbone dihedral correction.<sup>33</sup>

The differences in sampling in the area between the beta and alpha region and in the extra minimum found in Amber forcefields suggests a possible mechanism for the discrepancy in helix formation for the Charmm 22\* force field. Whereas other force fields possessed much greater barriers, the lack of phi barriers to the alpha region of the Charmm 22\* force field allows for an easy transition between beta and alpha regions thereby increasing the rate of helix formation by the Charmm 22\* force field. In contrast, the Amber ff14SB force field possesses nearly no points in half of the transition region (see Figure 8), illustrating the difficulties of moving from the alpha region to the beta region in the Amber ff14SB force field as compared to the Charmm 22\* force field.

## Discussion and Conclusion

In this paper we examined how various force fields could affect the refolding of the I91 mechanically perturbed intermediate and the helical peptide Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub>. By examining the distribution of dihedral angles, the frequency of secondary structures, and the arrangement of clustered structures we observed the increased formation of helices in the unfolded state for the Amber ff14SB and Amber ff99SB\*-ILDN force fields. Additionally we observed that even though the Charmm 22\* force field had similar overall helical frequencies to the Amber ff14SB force field, that the Charmm 22\* force field possesses a much smaller mean helical lifetime.

The Amber ff99SB\*-ILDN and Amber ff14SB force fields both displayed difficulty refolding the I91 mechanically perturbed intermediate. The difference is particularly interesting since these force fields form helices in the unfolded state of the I91 perturbation (see Figure 3). In the native state however, the I91 protein perturbation forms a beta strand and thus forms native contacts with other beta strands. In contrast, the presence of helical interme-

diates for the small perturbation results in the creation of a significant amount of non-native contacts. Since the I91 perturbation is known to refold experimentally,<sup>38</sup> and given the idea of downhill folding,<sup>29,43</sup> it is unlikely that the helical constructs of the unfolded state contribute to the folding pathway and rather represent a local minima which the perturbation gets trapped in.

Although the Amber ff99SB\*-ILDN and Amber ff14SB force fields both display difficulties refolding the I91 perturbation, the reasons for these difficulties likely differ. It has been shown that the Amber ff99SB\* force field is overly helical for glutamic acid and lysine residues.<sup>44</sup> This can be seen in how both the Amber ff14SBonlySC and Amber FB15 force fields, which do reparameterize the glutamic acid and lysine side chain torsion angles, can more robustly refold the I91 perturbation than the Amber ff99SB\*-ILDN force field (see Table 1).

The Amber ff14SB force field also displays difficulties refolding the I91 perturbation. However, the force field displays higher amounts of helical content than all of the other tested Amber force fields. Furthermore, the Amber ff14SB force field likely does not contain incorrect helicities for glutamic acid and lysine as it is based on the Amber ff14SBonlySC force field which reparameterized all side chains and robustly refolds the I91 perturbation (see Table 1). Thus, the only difference that could cause much greater helicities for the Amber ff14SB force field is the Amber ff14SB empirical backbone dihedral correction. Although this correction has been validated for other small peptides,<sup>33</sup> the failure of this correction to match experimentally observed helicities in the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptide as well as its difficulty refolding the I91 perturbation suggest that the correction overstabilizes helices. These observations have been replicated in implicit solvent,<sup>14</sup> although it should be noted that such observations may not be reliable as the dihedral correction for ff14SB was intended for explicit solvent simulations.<sup>33</sup> Furthermore, such an increase in helicities was not replicated in an analysis of A $\beta$ <sub>16-22</sub> dimer. We note that these discrepancies may be explained however, by the aforementioned study examining dimerization instead of refolding and also the more limited time scales per simulation (200 ns vs 2  $\mu$ s).<sup>25</sup>

The ease of Amber ff14SBonlySC and Amber FB15 in refolding the I91 perturbation also raises potential questions about how parameterization affects protein behavior. From the dihedral maps of the I91 mechanical perturbation unfolding state it can be observed that there are noticeable gaps in the beta and alpha region for the Amber ff14SBonlySC force field while the AMBER FB15 force field samples the entire region very widely. This presents two different mechanisms for the efficacy of these force fields in refolding the I91 protein perturbation. The Amber ff14SBonlySC potential of mean force plots along with its low helical frequency during Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> equilibration simulations suggest that the force field is biased toward the beta strand region of the Ramachandran plot. This and the lack of the additional minima in the region suggest that the ff14SBonlySC force field easily forms beta stranded proteins. In contrast, the wide sampling of the Amber FB15 force field suggests that the Amber FB15 force field does not get trapped in any single minima. Instead, the Amber FB15 force field will eventually make its way to the true minima after broadly sampling a larger configuration space.

Furthermore, the difference in helices present in the unfolded state of the I91 perturbation unfolded state suggests that even if two force fields have similar overall helicities as observed with the Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> peptide, the mean helical lifetime can affect whether those helices are actually observed during the refolding pathway. This can be seen when comparing Charmm 22\* to Amber ff14SB. Although Charmm 22\* Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> equilibration simulations resulted in the same overall helical frequencies as the Amber ff14SB equilibration simulations, the mean helical lifetimes differed. The difference was noticeable in the unfolded state of the I91 perturbation where the Charmm 22\* force field did not form many helices while the Amber ff14SB force field did. These observations can be correlated with observations of the potential of mean force plots between the Charmm 22\* and Amber force fields. The singular alpha region of the Charmm 22\* force field, without the extra minima found in the region ( $\phi \in [-160^\circ, -100^\circ]$  and  $\psi \in [-50^\circ, 50^\circ]$ ) may result in a more direct transition between the beta and alpha regions of the Ramachandran plot which may

result in more frequent but less stable helices as compared to other Amber force fields with similar overall helical frequencies.

The extra data from the I91 perturbed intermediate allowed for the controlled examination of beta strand refolding. The differences in this refolding pathway then led to observations about force field discrepancies that might have been otherwise unnoticed by examining bulk measurements like Ac-(AAQAA)<sub>3</sub>-NH<sub>2</sub> helical frequency. Thus, investigating force field accuracy through protein structure perturbations provides a way of testing the transferability of parameterization schemes and actual protein folding simulations while still keeping a system small enough to reduce the time and simulation costs of folding. Thus the creation of mechanically perturbed intermediates can prove to be a useful tool in identifying various force field discrepancies.

## Acknowledgement

The authors thank the following funding sources: NSF Grant MCB-1517245 and NSF Grant MCB-1817556, Sigma Xi Grant G2018031596204454, Duke Deans' Summer Research Fellowship, and NCICU Undergraduate Research Fellowship.

We thank R.B. Best, E. Josephs, and J. Pajak for helpful discussions regarding the nature of this publication. We thank J. Wu for help with the creation and set up of our computational equipment.

Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

## References

- (1) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How fast-folding proteins fold. *Science* **2011**, *334*, 517–520.



- (2) Best, R. B.; Mittal, J. Balance between  $\alpha$  and  $\beta$  Structures in Ab Initio Protein Folding. *J. Phys. Chem. B* **2010**, *114*, 8790–8798.
- (3) Lovas, S.; Zhang, Y.; Yu, J.; Lyubchenko, Y. L. Molecular Mechanism of Misfolding and Aggregation of A $\beta$ (13–23). *J. Phys. Chem. B* **2013**, *117*, 6175–6186.
- (4) Echeverria, I.; Makarov, D. E.; Papoian, G. A. Concerted Dihedral Rotations Give Rise to Internal Friction in Unfolded Proteins. *J. Am. Chem. Soc.* **2014**, *136*, 8708–8713.
- (5) Dill, K. A.; MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **2012**, *338*, 1042–1046.
- (6) 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. et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010**, *330*, 341–346.
- (7) Best, R. B.; Mittal, J. Protein Simulations with an Optimized Water Model: Cooperative Helix Formation and Temperature-Induced Unfolded State Collapse. *J. Phys. Chem. B* **2010**, *114*, 14916–14923.
- (8) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. Atomic-level description of ubiquitin folding. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 5915–5920.
- (9) Best, R. B.; Mittal, J. Free-energy landscape of the GB1 hairpin in all-atom explicit solvent simulations with different force fields: Similarities and differences. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 1318–1328.
- (10) Zhang, Y.; Hashemi, M.; Lv, Z.; Lyubchenko, Y. L. Self-assembly of the full-length amyloid A $\beta$ 42 protein in dimers. *Nanoscale* **2016**, *8*, 18928–18937.
- (11) Satija, R.; Das, A.; Makarov, D. E. Transition path times reveal memory effects and anomalous diffusion in the dynamics of protein folding. *J. Chem. Phys.* **2017**, *147*, 152707.



- (12) Piana, S.; Klepeis, J. L.; Shaw, D. E. Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2014**, *24*, 98–105.
- (13) McKiernan, K. A.; Husic, B. E.; Pande, V. S. Modeling the mechanism of CLN025 beta-hairpin formation. *J. Chem. Phys.* **2017**, *147*, 104107.
- (14) Shao, Q.; Zhu, W. Assessing AMBER force fields for protein folding in an implicit solvent. *Phys. Chem. Chem. Phys.* **2018**, *20*, 7206–7216.
- (15) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Systematic Validation of Protein Force Fields against Experimental Data. *PLoS ONE* **2012**, *7*, e32131.
- (16) Man, V. H.; Nguyen, P. H.; Derreumaux, P. High-Resolution Structures of the Amyloid- $\beta$  1–42 Dimers from the Comparison of Four Atomistic Force Fields. *J. Phys. Chem. B* **2017**, *121*, 5977–5987.
- (17) Martín-García, F.; Papaleo, E.; Gomez-Puertas, P.; Boomsma, W.; Lindorff-Larsen, K. Comparing Molecular Dynamics Force Fields in the Essential Subspace. *PLoS ONE* **2015**, *10*, e0121114.
- (18) Chen, W.; Shi, C.; MacKerell, A. D.; Shen, J. Conformational Dynamics of Two Natively Unfolded Fragment Peptides: Comparison of the AMBER and CHARMM Force Fields. *J. Phys. Chem. B* **2015**, *119*, 7902–7910.
- (19) Cino, E. A.; Choy, W.-Y.; Karttunen, M. Comparison of Secondary Structure Formation Using 10 Different Force Fields in Microsecond Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 2725–2740.
- (20) Beauchamp, K. A.; Lin, Y.-S.; Das, R.; Pande, V. S. Are Protein Force Fields Get-

- ting Better? A Systematic Benchmark on 524 Diverse NMR Measurements. *J. Chem. Theory Comput.* **2012**, *8*, 1409–1414.
- (21) Palazzesi, F.; Prakash, M. K.; Bonomi, M.; Barducci, A. Accuracy of Current All-Atom Force-Fields in Modeling Protein Disordered States. *J. Chem. Theory Comput.* **2015**, *11*, 2–7.
- (22) Rauscher, S.; Gapsys, V.; Gajda, M. J.; Zweckstetter, M.; de Groot, B. L.; Grubmüller, H. Structural Ensembles of Intrinsically Disordered Proteins Depend Strongly on Force Field: A Comparison to Experiment. *J. Chem. Theory Comput.* **2015**, *11*, 5513–5524.
- (23) Watts, C. R.; Gregory, A.; Frisbie, C.; Lovas, S. Effects of force fields on the conformational and dynamic properties of amyloid  $\beta(1-40)$  dimer explored by replica exchange molecular dynamics simulations. *Proteins: Struct., Funct., Bioinf.* **2018**, *86*, 279–300.
- (24) Carballo-Pacheco, M.; Strodel, B. Comparison of force fields for Alzheimer’s A $\beta$ 42: A case study for intrinsically disordered proteins: Comparison of Force Fields for Alzheimer’s A  $\beta$ 42. *Protein Sci.* **2017**, *26*, 174–185.
- (25) Man, V. H.; He, X.; Derreumaux, P.; Ji, B.; Xie, X.-Q.; Nguyen, P. H.; Wang, J. Effects of All-Atom Molecular Mechanics Force Fields on Amyloid Peptide Assembly: The Case of A $\beta$ <sub>16–22</sub> Dimer. *J. Chem. Theory Comput.* **2019**, *15*, 1440–1452.
- (26) Zhang, Y.; Lyubchenko, Y. The Structure of Misfolded Amyloidogenic Dimers: Computational Analysis of Force Spectroscopy Data. *Biophys. J.* **2014**, *107*, 2903–2910.
- (27) Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. Ten-Microsecond Molecular Dynamics Simulation of a Fast-Folding WW Domain. *Biophys. J.* **2008**, *94*, L75–L77.
- (28) Freddolino, P. L.; Park, S.; Roux, B.; Schulten, K. Force Field Bias in Protein Folding Simulations. *Biophys. J.* **2009**, *96*, 3772–3780.

- (29) Best, R. B.; Hummer, G.; Eaton, W. A. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 17874–17879.
- (30) Piana, S.; Lindorff-Larsen, K.; Shaw, D. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophys. J.* **2011**, *100*, L47–L49.
- (31) Best, R. B.; Buchete, N.-V.; Hummer, G. Are Current Molecular Dynamics Force Fields too Helical? *Biophys. J.* **2008**, *95*, L07–L09.
- (32) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.
- (33) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (34) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone  $\phi$ ,  $\psi$  and Side-Chain  $\chi_1$  and  $\chi_2$  Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (35) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.
- (36) Wang, L.-P.; McKiernan, K. A.; Gomes, J.; Beauchamp, K. A.; Head-Gordon, T.; Rice, J. E.; Swope, W. C.; Martínez, T. J.; Pande, V. S. Building a More Predictive Protein Force Field: A Systematic and Reproducible Route to AMBER-FB15. *J. Phys. Chem. B* **2017**, *121*, 4023–4039.
- (37) Robustelli, P.; Piana, S.; Shaw, D. E. Developing a molecular dynamics force field for

- both folded and disordered protein states. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E4758–E4766.
- (38) Marszalek, P. E.; Lu, H.; Li, H.; Carrion-Vazquez, M.; Oberhauser, A. F.; Schulten, K.; Fernandez, J. M. Mechanical unfolding intermediates in titin modules. *Nature* **1999**, *402*, 100.
- (39) Dudko, O. K.; Graham, T. G. W.; Best, R. B. Locating the Barrier for Folding of Single Molecules under an External Force. *Phys. Rev. Lett.* **2011**, *107*, 208301.
- (40) Makarov, D. E. Shapes of dominant transition paths from single-molecule force spectroscopy. *J. Chem. Phys.* **2015**, *143*, 194103.
- (41) Hoskins, J. R.; Yanagihara, K.; Mizuuchi, K.; Wickner, S. ClpAP and ClpXP degrade proteins with tags located in the interior of the primary sequence. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11037–11042.
- (42) Ganim, Z.; Rief, M. Mechanically switching single-molecule fluorescence of GFP by unfolding and refolding. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 11052–11056.
- (43) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545–600.
- (44) Best, R.; de Sancho, D.; Mittal, J. Residue-Specific  $\alpha$ -Helix Propensities from Molecular Simulation. *Biophys. J.* **2012**, *102*, 1462–1467.
- (45) Stacklies, W.; Vega, M. C.; Wilmanns, M.; Gräter, F. Mechanical Network in Titin Immunoglobulin from Force Distribution Analysis. *PLoS Comput. Biol.* **2009**, *5*, e1000306.
- (46) Gao, M.; Lu, H.; Schulten, K. Simulated refolding of stretched titin immunoglobulin domains. *Biophysical Journal* **2001**, *81*, 2268–2277.

- (47) Li, P.-C.; Makarov, D. E. Theoretical studies of the mechanical unfolding of the muscle protein titin: Bridging the time-scale gap between simulation and experiment. *J. Chem. Phys.* **2003**, *119*, 9260–9268.
- (48) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3694–3699.
- (49) Rief, M. Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *Science* **1997**, *276*, 1109–1112.
- (50) Shalongo, W.; Dugad, L.; Stellwagen, E. Distribution of Helicity within the Model Peptide Acetyl(AAQAA)<sub>3</sub>amide. *J. Am. Chem. Soc.* **1994**, *116*, 8288–8293.
- (51) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (52) Bowers, K. J.; Chow, D. E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D. et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. ACM/IEEE SC 2006 Conference (SC'06). Tampa, FL, 2006; pp 43–43.
- (53) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *9*, 3878–3888.
- (54) 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., Bioinf.* **2010**, NA–NA.

- (55) Sippl, M. J. Calculation of Conformational Ensembles from Potentials of Mean Force. *J. Mol. Biol.* **1990**, *213*, 859–883.
- (56) Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637.
- (57) Peng, J.-h.; Wang, W.; Yu, Y.-q.; Gu, H.-l.; Huang, X. Clustering algorithms to analyze molecular dynamics simulation trajectories for complex chemical and biological systems. *Chin. J. Chem. Phys.* **2018**, *31*, 404–420.
- (58) McGibbon, R.; Beauchamp, K.; Harrigan, M.; Klein, C.; Swails, J.; Hernández, C.; Schwantes, C.; Wang, L.-P.; Lane, T.; Pande, V. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109*, 1528–1532.
- (59) Beauchamp, K. A.; Bowman, G. R.; Lane, T. J.; Maibaum, L.; Haque, I. S.; Pande, V. S. MSMBuilder2: Modeling Conformational Dynamics on the Picosecond to Millisecond Scale. *J. Chem. Theory Comput.* **2011**, *7*, 3412–3419.
- (60) Meng, E. C.; Pettersen, E. F.; Couch, G. S.; Huang, C. C.; Ferrin, T. E. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinf.* **2006**, *7*, 339.
- (61) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
- (62) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.

## Graphical TOC Entry

