

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/259446446>

The gp41(659–671) HIV-1 Antibody Epitope: A Structurally Challenging Small Peptide

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · DECEMBER 2013

Impact Factor: 3.3 · DOI: 10.1021/jp409355r · Source: PubMed

CITATIONS

4

READS

31

2 AUTHORS:



[Yuan Zhang](#)

North Carolina State University

3 PUBLICATIONS 15 CITATIONS

[SEE PROFILE](#)



[Celeste Sagui](#)

North Carolina State University

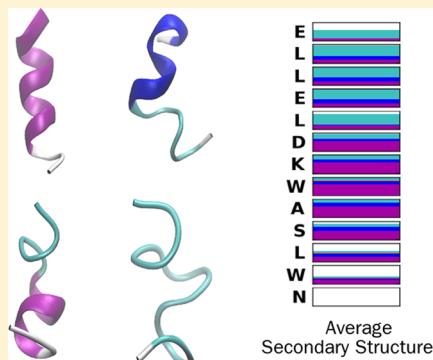
104 PUBLICATIONS 2,172 CITATIONS

[SEE PROFILE](#)

The gp41_{659–671} HIV-1 Antibody Epitope: A Structurally Challenging Small Peptide

Yuan Zhang^{†,‡} and Celeste Sagui^{*,†,‡}[†]Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, United States[‡]Center for High Performance Simulations (CHiPS), North Carolina State University, Raleigh, North Carolina 27695, United States

ABSTRACT: We report on extensive molecular dynamics (MD) simulations of the tridecapeptide corresponding to residues 659–671 of the envelope glycoprotein gp41 of HIV-1, which spans the 2F5 monoclonal antibody epitope ELDKWA. Previously, X-ray crystallography, nuclear magnetic resonance, and circular dichroism experiments have yielded conflicting conformational information, but there is a growing consensus that the monomeric peptide in aqueous solution is disordered. Here, we use the latest, state-of-the-art AMBER force fields to describe the complex conformational landscape of gp41_{659–671}. We have analyzed the conformational ensembles of the peptide in solution both without applied restraints and under successive tensile restraints. In contrast to previous MD simulations, our results are consistent with the bulk of the experimental findings. The amount of helical population is important in aqueous solution, but this structure forms part of a flexible conformational ensemble with a rugged free energy landscape with shallow minima. Under uniaxial tension, the disordered peptide first becomes fully helical before melting into turns, loops, and 3₁₀-helices. The conformational ensemble includes epitope conformations close to an NMR solution structure (PDB ID 1LCX) as well as epitope conformations close to a very different, extended crystal structure (PDB ID 1TJH).



INTRODUCTION

The recent interest in the structural properties of the peptide gp41_{659–671} (ELLELDKWASLWN) is due to two related issues: (i) as possible target for vaccines as it resides in the membrane proximal region of the HIV-1 fusogenic subunit gp41 and (ii) as a short peptide with an elusive structure that challenges both experiments and simulations.

In the HIV-1 virus, gp41 and gp120 are two transmembrane glycoproteins that mediate the fusion of the virus with the host cell. The structure of gp41 consists of a hydrophobic N-terminus fusion domain; two core, highly conserved helical regions that have been solved both by X-ray crystallography^{1–3} and by nuclear magnetic resonance (NMR);⁴ and the C-terminus. Understanding the structure of these domains and the different mechanisms of protein fusion—which eventually result in a six-helix bundle across the host cell membrane with an open channel for the entry of the HIV-1 nucleocapsid—can eventually lead to vaccines. Unfortunately, neutralizing antibodies so far cannot access the helical core structure.⁵ The terminal regions, on the other hand, have proved to be a more susceptible target. For instance, it is known that mutations or deletions in the tryptophan-rich C-terminal region (residues 660–683) can abolish viral fusion.⁶ In addition, this region overlaps part of the peptide T-20 (gp41_{638–673}), which was the first gp41 fusion inhibitor approved by the US Food and Drug Administration in 2003.

A vaccine can also be based on immunogens containing conserved epitopes for monoclonal antibodies in a way that

mimics their structure on the native gp41 and gp120 glycoproteins. In particular, the gp41_{659–671} sequence (which is included in T-20) contains the sequence ELDKWA which is the core epitope for the HIV-1 neutralizing monoclonal antibody 2F5.^{7,8} The 2F5 antibody reacts strongly with peptides that contain the ELDKWA sequence; however, when this sequence is incorporated in various immunogens, these immunogens fail to induce antibodies that react with the native form of the HIV-1 envelope glycoprotein complex. Experimental findings^{8,6} show that all residues in gp41_{659–671} are important to induce antibodies that act like 2F5. Experiments also show that 2F5 only binds to the prefusogenic form of gp41⁹ (before gp120 binds to the T-cell membrane) and suggest that the gp41_{659–671} epitope is solvent exposed in this situation. Thus, it is desirable to know the structure of the gp41_{659–671} epitope in its solvated prefusogenic state.

There have been several experimental and simulation studies to determine the structure of gp41_{659–671} in aqueous solution. An early study based on NMR at pH 7.7 indicated that the conformation of the monomeric peptide in water is an amphiphilic 3₁₀-helix with minor random coil representation.⁶ A second NMR study found no major population of 3₁₀-helical conformers, but a mixture of various conformers.¹⁰ Later, a UV resonance Raman spectroscopy (UVRSS) investigation¹¹

Received: September 18, 2013

Revised: December 11, 2013

Published: December 12, 2013

concluded that there is a rough energy landscape with a wider variety of conformations than found in the NMR studies and previous circular dichroism (CD) studies.⁶ The authors suggested that gp41_{659–671} exhibits a broad distribution of conformations that includes significant population of β -turns as well as 3₁₀-helix and π -helix motifs but little α -helix. Recently, a far-UV CD spectroscopy study¹² also revealed the conformational plasticity of gp41_{659–671}, with no stable α -helical, 3₁₀-helical, or turn motifs. On the other hand, the crystal structure of the peptide bound to the 2F5 antibody shows an extended conformation.^{13,14} Figure 1 shows sample structures obtained

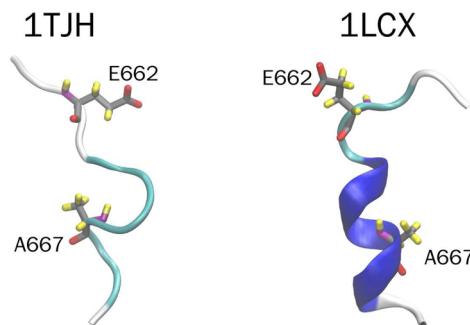


Figure 1. Conformations of the gp41 11-mer peptide (gp41_{659–671} minus the terminal residues, i.e., LLELDKWASLW) from the 1TJH crystal structure¹³ and the gp41_{659–671} 13-mer peptide from the 1LCX NMR solution structure.⁶ The flanking residues E662 and A667 of the epitope ELDKWA are shown. The color scheme for the secondary structure (generated by VMD⁵⁸) is as follows: 3₁₀-helix: blue; turn: cyan; coil: white.

experimentally: PDB ID 1TJH shows the crystal structure of an 11-mer (without the two terminal residues of gp41_{659–671}) bound to 2F5, as found by Ofek et al.,¹³ while PDB ID 1LCX shows the solution NMR 3₁₀-helix, as found by Biron et al.⁶

From the point of view of the simulations, molecular dynamics (MD) results have been no less conflicting. In this case, many of the discrepancies can be attributed to the use of different force fields. A study based on the Gromos96 43a1 force field^{15,16} found no stable structures in aqueous solution, but an equilibrium of interchanging globular structures, as well as non-negligible populations of bends, turns, and α -helices.¹⁷ This force field, however, has been shown to exhibit conformational biases.¹⁸ Another recent MD study used the CHARMM22,¹⁹ AMBER ff03,^{20,21} and AMBER ff99SB²² force fields with the TIP3P water model.²³ In all cases, the authors found far more α -helical content than experimentally observed, with CHARMM22 giving the poorest agreement and ff03 being somewhat closer to experiments. The same group¹² carried out a second study combining simulations based on CHARMM22 and experimental data; unfortunately, the MD results agree poorly with the CD experimental data presented. All three force fields were later shown to have secondary structure biases. In particular, both CHARMM22 and ff03 in general overestimate helices,²⁴ while ff99SB is believed to understabilize helices²⁴ (although the ff99SB results in ref 23 also seem to overestimate the α -helical content with respect to the CD data). Out of the eight force fields employed in the validation study, the best performing fields are CHARMM22* and ff99SB*-ILDN which, with respect to the original CHARMM22 and ff99SB force fields, contain modified parameters for backbone and certain side chains.

Although force fields have come a long way since their initial development, the fact that they can still show disagreement with experimental data for certain biomolecules is an issue of great concern. The previous work showed how a small peptide such as gp41_{659–671} gives rise to rather different results depending on the choice of force field and that these results in general do not agree with experiments. Yet, the same gp41_{659–671} peptide also highlights some of the challenges of validating against experimental data, given that the conclusions of the experimental studies themselves can be quite contradictory, too, since of course all experiments have their own sources of errors and uncontrolled variables.

In this work, we test how a brand new AMBER force field, mainly ff12SB,²⁵ performs in reproducing the conformations of gp41_{659–671}. We compare it against the more extensively tested ff99SB, whose performance has been deemed satisfactory by a number of studies^{22,26} and for which there are already MD results for gp41_{659–671}. However, it has also been recognized that ff99SB understabilize helical conformations of transiently folded peptides.²⁴ In order to improve the accuracy of secondary structure propensities, the backbone angle ϕ' has been modified in ff12SB. Since residues such as isoleucine, leucine, aspartate, asparagine, and others were shown to sample conformations differently from experiments,^{24,27} the dihedral angles of the side chains in ff12SB have been refined by fitting energy profiles to match ab initio quantum data. Similar side-chain dihedral corrections have been extended to lysine, arginine, glutamate, glutamine, methionine, serine, threonine, valine, tryptophan, cysteine, phenylalanine, tyrosine, and histidine (i.e., all the side chains but Gly, Ala, and Pro), and thus ff12SB is expected to enhance the reproduction of experimentally indicated geometries over ff99SB, although more extensive testing is still needed. In addition to sampling equilibrium configurations, we compare the conformational landscape as the peptide is pulled away from its equilibrium length, and we compare results for both implicit and explicit solvent. Given the dispersion of results in the NMR and CD studies, we hope to motivate single-molecule experiments such as Förster resonance energy transfer (FRET) and single-molecule pulling experiments, which may throw additional information about the conformation of this peptide.

COMPUTATIONAL METHODS

To investigate the conformational phase space of gp41_{659–671}, we have carried out several simulations with the AMBER ff99SB and ff12SB versions of the Cornell et al. force field,²² using both implicit and explicit solvent for each force field. These simulations explore the conformational landscape of the peptide at equilibrium using regular MD. We also use equilibrium umbrella sampling to probe the conformations as the peptide is restrained to different end-to-end distances. Each simulation includes one gp41_{659–671} molecule with sequence Ace-ELLELDKWASLWN-NH₂. The N-acetylated and C-amidated forms are used to avoid excessive interactions between the charged N- and C-terminus in the implicit solvent. These capping groups are generally assumed to make the peptide more similar to the corresponding segment in the native protein. Experimental comparisons of gp41_{659–671} in its zwitterionic and capped forms show essentially the same diffusion constant and the same conformations with a small increase of the helical population in the capped form.⁶ Simulations were carried out at room temperature (300 K) via Langevin dynamics with a coupling parameter $\gamma = 1.0 \text{ ps}^{-1}$.

The peptide has two net negative charges that were neutralized by the addition of two Na^+ ions. Coordinates were sampled every 2 ps for regular MD simulation and 10 ps for the umbrella sampling runs.

Implicit Solvent Simulations. Topology and parameter files along with the coordinates corresponding to the unfolded peptide were generated via the LEAP program of the AMBER v.11²⁸ and v.12²⁵ simulation packages. The implicit water model is based on the generalized Born approximation (GB)²⁹ including the surface area contributions using the LCPO model³⁰ (GB/SA) with the surface tension set to 0.005 kcal/(mol Å²). For the GB model, we used the GB_{OBC} II model,³¹ with a cutoff of 25 Å. With this model, the effective Born radii are rescaled to account for interstitial spaces between atom spheres missed by the GB^{HCT} approximation.

After a short minimization, an initial 10 ns MD simulation was used to generate equilibrium conformations of the peptide. Five different conformations from this run were chosen as initial conformations for five MD runs (with different random number generator seeds), each 40 ns long. One conformation from these sets was then selected as initial conformation for another 100 ns simulation. The 100 ns set and the second half of five 40 ns sets were used for data analysis of equilibrium conformations. Thus, the simulation was run for a cumulative time of 300 ns, and data are presented for 200 ns. Figure 2 shows the five initial conformations for the 40 ns MD simulations for ff99SB. The same procedure was repeated for ff12SB.

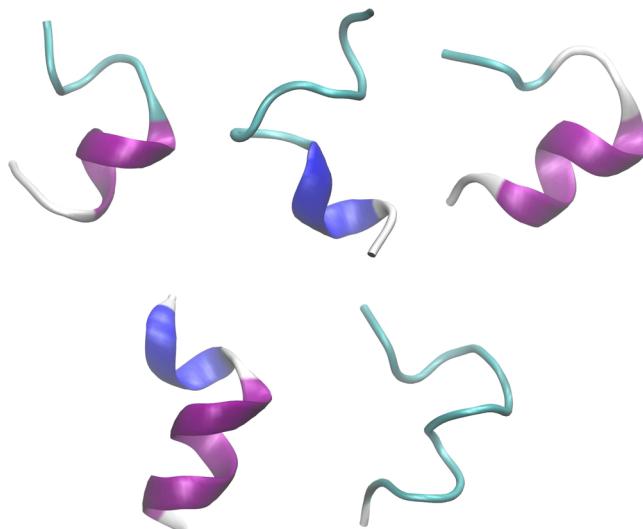


Figure 2. Snapshot of chosen initial conformations for the unrestrained MD simulations under the ff99SB force field. The top of the peptide corresponds to E659 and the bottom to W671. These conformations all evolved from a fully extended initial structure.

We then used equilibrium umbrella sampling to “pull” the peptide through different end-to-end distances by applying a harmonic restraint to the collective variable. The end-to-end distance collective variable is defined as the distance between the peptide-bond nitrogen atoms in E659 and in N671. The end-to-end distance d was restrained with a harmonic potential $V_{\text{rest}} = (1/2)k[d - d_0]^2$ in the range 10–40 Å through 2 Å windows (i.e., $d_0 = 10$ Å, 12 Å, etc.), resulting in 16 windows. 10 ns was used to stretch a given configuration through the 16 windows with a harmonic constant of 100 kcal/(mol Å²). Then,

for each of the configurations centered at 10 Å, 12 Å, etc., we “equilibrated” the system by using a harmonic restraint of 10 kcal/(mol Å²) for 1 ns and then 1 kcal/(mol Å²) for 1 ns. The conformations at this point were taken as initial conformations for 100 ns runs with a harmonic restraint of 1 kcal/(mol Å²), which were finally used for data analysis.

Explicit Solvent Simulations. The TIP3P water model³² was used for the explicit solvent simulations under periodic boundary conditions. Five initial conformations from the implicit solvent simulations for each force field were solvated in a rectangular box, with an average number of waters of approximately 5670. Electrostatics were handled by the PME method,^{33,34} with a direct space cutoff of 9 Å and an average mesh size of approximately 1 Å for the lattice calculations. The equilibration process took place in four steps. First, we applied steepest descent followed by conjugate gradient minimization keeping the peptide atoms fixed at their initial positions. Then we carried out unrestrained steepest descent followed by conjugate gradient minimizations. This was followed by short MD runs under constant volume while the system was heated from 0 to 300 K with weak restraints on the peptide atoms. Finally, the system was kept at 300 K via Langevin dynamics with a collision frequency $\gamma = 1.0 \text{ ps}^{-1}$ and at constant pressure (1 atm) via the Berendsen barostat³⁵ with the isothermal compressibility $\beta = 44.6 \times 10^{-6} \text{ bar}^{-1}$ and the pressure relaxation time $\tau_p = 1.0 \text{ ps}$. These NPT Langevin dynamics simulations were carried out for 2 ns, and the density of the system was found to be stable around 1.0 g/cm³. The two sets of five equilibrated conformations for each force field provided the initial conformations for NPT simulations with Langevin dynamics, each simulation 40 ns long. One configuration for each force field was selected for a further 100 ns run. The second half of each of the 40 ns runs and the 100 ns runs were used for data analysis.

For the umbrella sampling simulations, we also used the NPT ensemble as before, and the collective variable was again selected as the end-to-end distance, with conformations restrained around 14 windows separated by 2 Å, from 10 to 36 Å. The system was pulled through the different end-to-end distances and equilibrated in the same way as in the implicit solvent case. The bond force constant was set at 1.0 kcal/(mol Å²), and the equilibrium bond length was set at 10, 12, 14, ..., 36 Å.

RESULTS

Figure 3 shows the free energy profiles as a function of the end-to-end distance for the two force fields in implicit and explicit solvents. To first order, the two force fields provide similar results, with the equilibrium length at 17.92, 17.75, 16, and 17.18 Å under implicit ff99SB, explicit ff99SB, implicit ff12SB, and explicit ff12SB. The equilibrium end-to-end distances are in good agreement with each other except for implicit ff12SB, which seems to yield a slightly more compact structure. A general feature for these free energies is that for distances greater than the equilibrium length, the explicit solvent lowers the free energy. This can be easily explained, as in the absence of waters, the peptide tends to form hydrogen bonds with itself, making elongated conformations more costly than in the presence of explicit waters, which can replace the internal hydrogen bonds.

Figure 4 shows the Ramachandran plots for each residue under the two force fields and implicit and explicit solvents. For most residues, in addition to α_R , there is an important presence

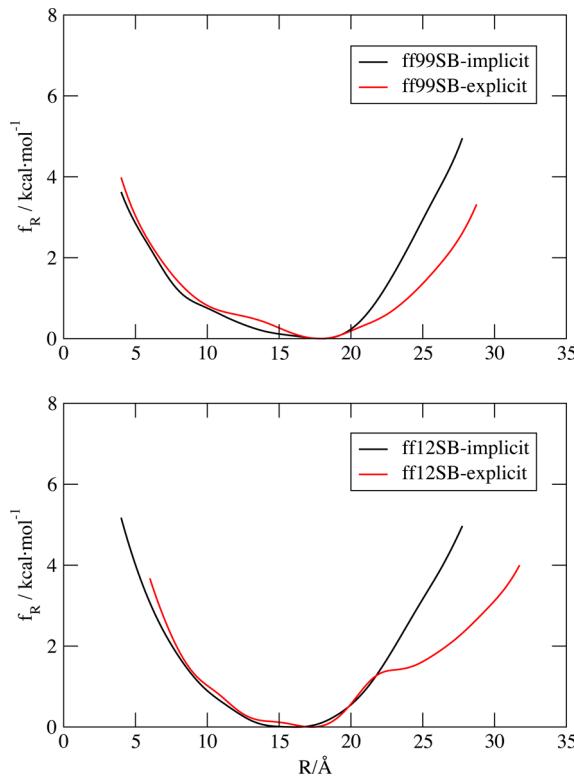


Figure 3. Free energy profiles as a function of the end-to-end distance for gp41_{659–671}. Results for AMBER ff99SB (top) and ff12SB (bottom) in implicit (black) and explicit (red) solvent are shown.

of β and PPII populations. There is better agreement between the implicit and explicit solvent representations for ff12SB than for ff99SB. In the latter, residues Glu-659 to Leu-663 and Leu-669 and Trp-670 show slightly larger β and PPII populations in explicit ff99SB than in implicit ff99SB (indicating less α -helical content in the explicit solvent). On the other hand, for both force fields Trp-666 shows a considerably large PPII population in implicit solvents that is absent in the explicit solvents. Although certain residues in the implicit solvent representation have populations in the α_L region of the Ramachandran plot that are absent in the explicit plots, the α_L region seems to be consistently present in Leu-660 and, especially, Asn-671 (which is systematically characterized as “loop” in Figures 5 and 6).

The secondary structure motifs were characterized by using the program STRIDE,³⁶ since it is the code used in a couple of other simulations of gp41_{659–671}, and therefore comparisons are more straightforward. Figures 5 and 6 show the secondary structure population for each residue (y-axis) as a function of “sampling window”: strictly speaking, it is not completely as a function of time because the first 100 ns are obtained as aggregates of the last 20 ns in each of the five independent 40 ns runs; the period between 100 and 200 ns corresponds to the time evolution of another independent run (each simulation was therefore “equilibrated” for over 20 ns, and this data is not shown). We see that the second half of the plot constructed in this way is qualitatively similar to the first half, suggesting that the system has undergone ergodic sampling. The average secondary structure for each residue can be appreciated at a glance in Figure 7. The major differences between ff99SB and ff12SB occur in the N-terminal ELLEL sequence. This is due to both corrections in residues L and E, which was found to be very helical in ff99SB.²⁷ Thus, the α -helical content of ELLEL is

51% in implicit ff99SB while only 11% in implicit ff12SB, and it is 31% in explicit ff99SB while only 19% in ff12SB. For the individual E’s in ELLEL, the differences in α -helical content of E659 and E662 amounts to 43% for implicit ff99SB while only 8% for implicit ff12SB and 24% for explicit ff99SB while 18% for explicit ff12SB.

For each force field, the maps between the implicit and the explicit solvent are in qualitative agreement. If one considers the implicit and explicit solvent populations for a given type of secondary structure for *each residue*, and then sums the absolute value of their differences, then it is possible to obtain some sort of measure of how much the implicit and explicit solvent representations in Figures 5 and 6 differ (notice that there is emphasis on absolute differences of individual residues; in the averages presented in Table 1 positive differences at a given residue can cancel out negative differences at another residue). Overall, the differences between implicit and explicit ff99SB are larger than the differences between implicit and explicit ff12SB. In other words, there seems to be a better agreement between both solvents in ff12SB than in ff99SB, a result that is more pronounced if DSSP^{37,38} is used for the secondary structure analysis. The major differences between implicit and explicit ff99SB occur in the population of α -helices, while the major differences between implicit and explicit ff12SB occur in the population of turns. Although π structures have been reported before,¹¹ we find that the population of π -helix is negligible, in agreement with more recent measurements.^{12,23} The amount of 3_{10} -helices and turn helices is also important. In particular, 3_{10} -helices, which have been reported in a previous NMR study,⁶ vary between 5% for implicit ff99SB and 13% for explicit ff12SB (the analysis was repeated with DSSP, which yields 13% and 14% for 3_{10} -helices in implicit and explicit ff99SB and 19% and 21% in implicit and explicit ff12SB). In other words, for each force field 3_{10} -helices tend to show up more in explicit than implicit solvents, and explicit ff12SB exhibits more of them than explicit ff99SB. Turns comprise 28% of the total populations of both implicit and explicit ff99SB and 41% and 31% of the total populations of implicit and explicit ff12SB. The third most common structural “motif” is “loops”, i.e., disordered structures which comprise 24% and 30% of the total populations of implicit and explicit ff99SB and 20% and 22% of the total populations of implicit and explicit ff12SB. The average populations for the whole peptide and for the epitope are shown in Table 1.

Figure 8 shows the distribution of conformations of the ELDKWA epitope based on the root-mean-square deviation (RMSD) values of the backbone atoms N, O, C', and C_α with respect to a solution NMR structure (1LCX)⁶ of gp41_{659–671} and with respect to the crystal structure of a heptamer ELDKWAS in complex with 2F5 (1TJG).¹³ The 1LCX solution NMR structure is the one reported by Biron et al.,⁶ which is a 3_{10} -helix from E662 to W670 according to DSSP and from D664 to L669 according to STRIDE (as shown in Figure 1). The crystal structure of the heptamer shows the same structure as 1TJH for the 11-mer (Figure 1) and 1TJI for the 17-mer, and all correspond to the conformations found by Ofek et al.¹³ When the structure of gp41_{659–671} is analyzed as part of the complex in 1TJH, then DSSP and STRIDE predict a β -strand on residues E662 and L663, followed by a turn, followed by disorder. If we try the isolated gp41_{659–671}, then the secondary structure is PPII for E662 and L663 and a type I β -turn for residues D664–A667. The simulations results in Figure 8 for both explicit ff99SB and explicit ff12SB show an important

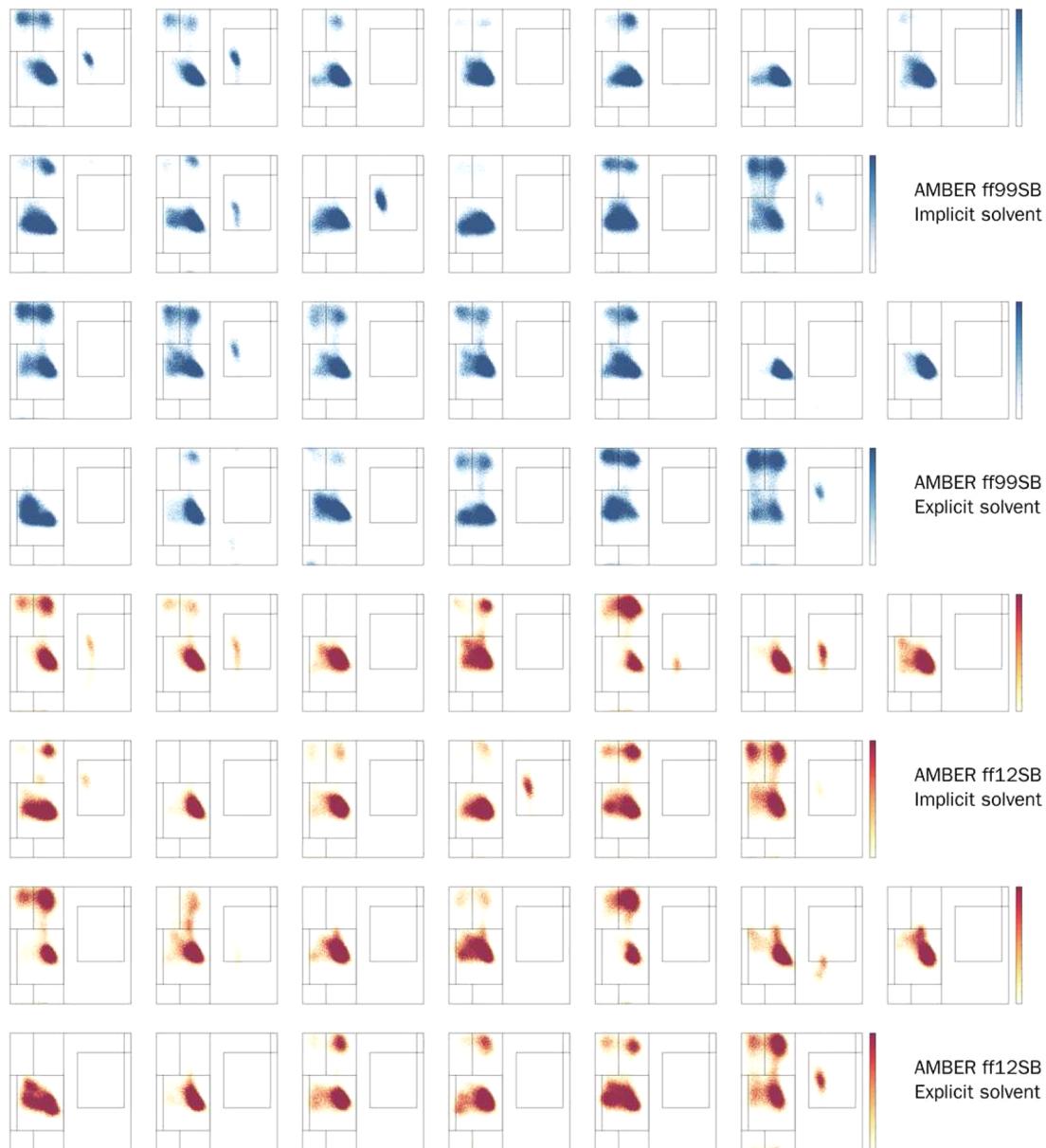


Figure 4. Ramachandran plots for each residue in ELLELDKWASLWN for the unrestrained MD simulations. Results for the AMBER ff99SB are presented in blue and for ff12SB in red/yellow. For each force field, the top two rows show the plots for implicit solvent, while the bottom two rows show the plots for the explicit solvent.

population presence in the bottom quadrants, which partly correspond to the solution NMR structure, with RMSDs between 1 Å (even less for ff99SB) and 2 Å. There is another important contribution in the upper-right quadrant, for conformations that are similar neither to the crystal structure nor to the NMR solution structure and that are mainly disordered. In addition, both force fields show a contribution to the lower-left quadrant, closer to the crystal structure. The implicit solvent representations (not shown) tend to increase the population in the upper-right quadrant and to decrease the rare, relatively extended conformations in the lower-left quadrant that have smaller RMSDs with respect to 1TJG.

Another comparison to experiment can be carried out through NMR chemical shifts. Figure 9 gives the distribution function of conformations versus the RMSD of experimental ($\delta_{j,\text{exp}}$) and predicted ($\delta_{j,\text{pred}}$) chemical shifts computed via SHIFTX2³⁹ for HN and HA, according to

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{j=1}^N (\delta_{j,\text{pred}} - \delta_{j,\text{exp}})^2} \quad (1)$$

where N is the number of residues of each conformation. The experimental data are provided in ref 6 and correspond to the structure 1LCX. We notice that the RMSD for the algorithm itself as implemented in SHIFTX2³⁹ is 0.2351 ppm for HN and 0.1081 ppm for HA. This means that in addition to discrepancies between simulations and experiment, there are extra sources of errors due to the accuracy of the chemical shift predicting algorithm. Thus, if we run the 1LCX structure itself through SHIFTX2, the computed RMSDs are 0.3425 ppm for HN and 0.2073 ppm for HA (these are indicated as a red bar in the graphic). Again, although both explicit force fields are similar, ff12SB seems to perform slightly better than ff99SB.

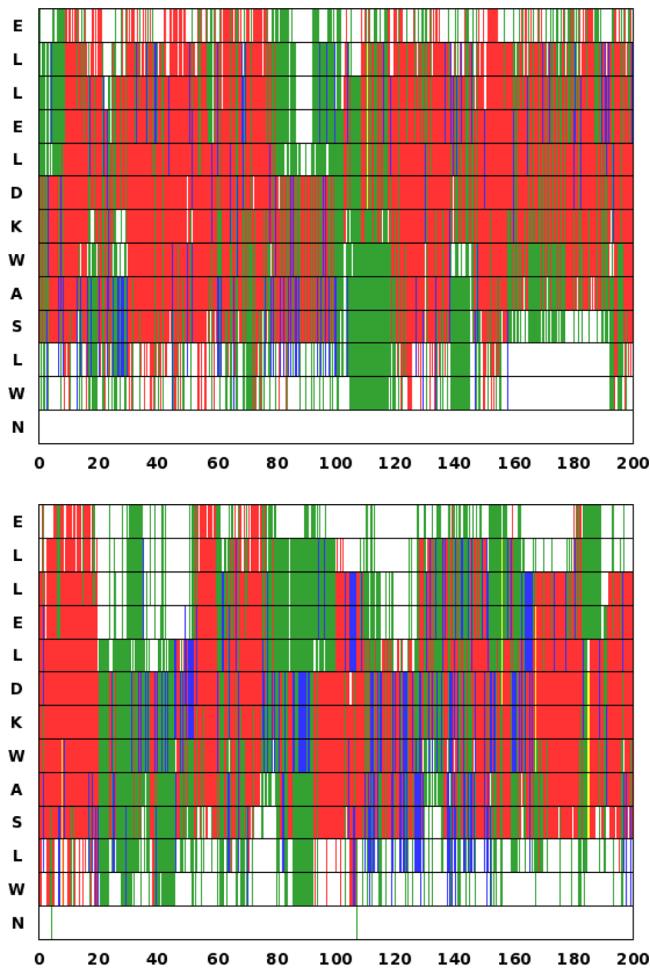


Figure 5. Secondary structure map, as obtained with STRIDE, of the unrestrained MD simulations under implicit ff99SB (top) and explicit ff99SB (bottom). The red, blue, yellow, green, and white colors correspond to α -helix, 3_{10} -helix, π -helix, turn, and “other” secondary structural motifs, respectively.

Figures 10 and 11 show the average secondary structure population for each residue (y-axis) as a function of the end-to-end distance (in Å) used in the restraint for the umbrella sampling runs. Each window represents an average over 100 ns. For ff99SB, the implicit solvent structures show more α -helical content than the corresponding explicit solvent ones, a trend that was already present in the unrestrained simulations. There is more consistent agreement between implicit and explicit solvents under the ff12SB force field. If one chooses a middle residue in the implicit graphs, and follows it along the row as the peptide is stretched, one can see a relatively smooth behavior, where the α -helical content (red) increases from its initial value at 10 Å until it reaches a maximum at approximately 20 Å and then it decreases accompanied by an increase of 3_{10} -helices, turns, and disordered extended structures. This behavior is considerably less “regular” in the explicit solvents (especially in ff99SB, where no clear trend is discernible). In particular, the number of α -helical structures decreases dramatically after 22 Å, giving rise to many more 3_{10} -helices, turns, and disordered extended structures, in agreement with the lowering in free energies in Figure 3 for the explicit solvents with respect to the implicit solvents.

Finally, Figure 12 shows a similar plot to Figure 8, i.e., the distribution of conformations of the ELDKWA epitope based

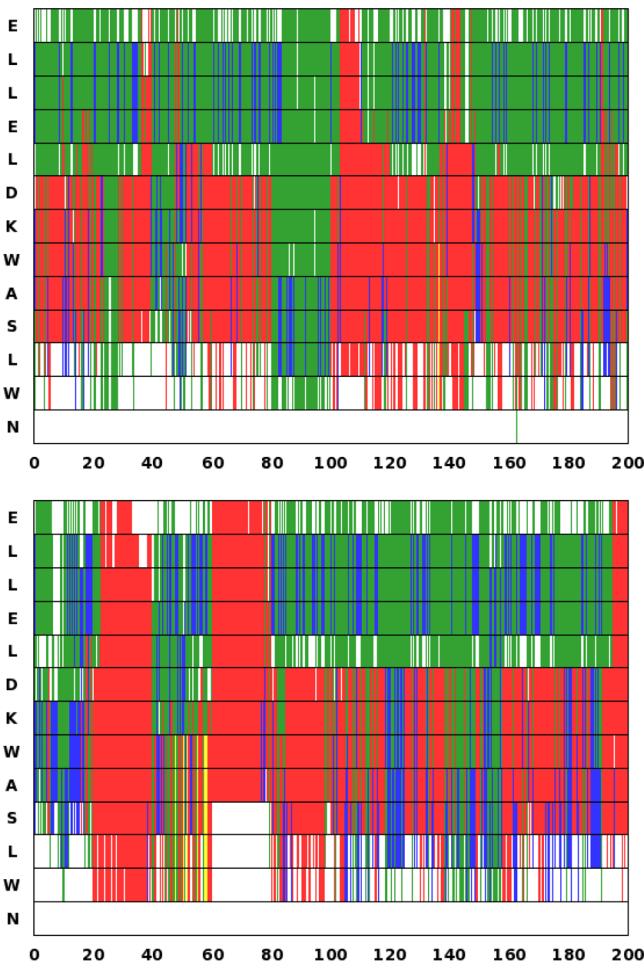


Figure 6. Secondary structure map, as obtained with STRIDE, of the unrestrained MD simulations under implicit ff12SB (top) and explicit ff12SB (bottom). The red, blue, yellow, green, and white colors correspond to α -helix, 3_{10} -helix, π -helix, turn, and “other” secondary structural motifs, respectively.

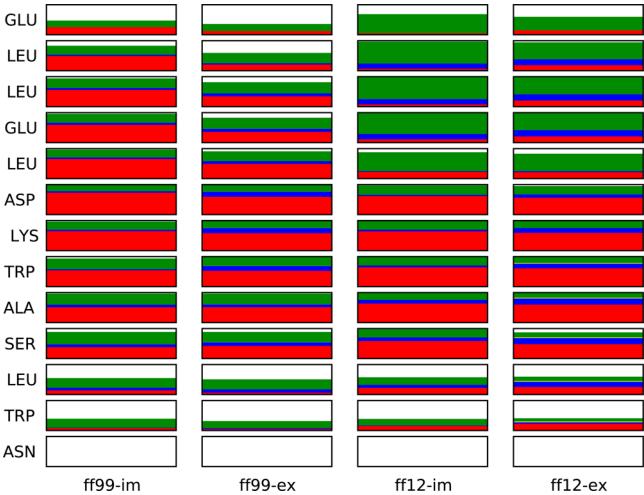


Figure 7. Average secondary structure map, as obtained with STRIDE, of the unrestrained MD simulation under implicit ff99SB, explicit ff99SB, implicit ff12SB, and explicit ff12SB (from left to right). Secondary structural motifs are α -helix (red), 3_{10} -helix (blue), π -helix (yellow), turn (green), and “other” (white).

Table 1. Secondary Structure Population Based on STRIDE for gp41_{659–671} and the Epitope ELDKWA

| peptide | simulation model | α -helix | 3_{10} -helix | π -helix | turn | other |
|-------------------------|------------------|-----------------|-----------------|--------------|------|-------|
| gp41 _{659–671} | ff99SB-implicit | 43 | 5 | 0 | 28 | 24 |
| | ff99SB-explicit | 33 | 9 | 0 | 28 | 30 |
| | ff12SB-implicit | 31 | 8 | 0 | 41 | 20 |
| | ff12SB-explicit | 33 | 13 | 1 | 31 | 22 |
| ELDKWA | ff99SB-implicit | 61 | 6 | 0 | 29 | 4 |
| | ff99SB-explicit | 51 | 13 | 0 | 30 | 5 |
| | ff12SB-implicit | 48 | 8 | 0 | 41 | 3 |
| | ff12SB-explicit | 46 | 15 | 1 | 34 | 4 |

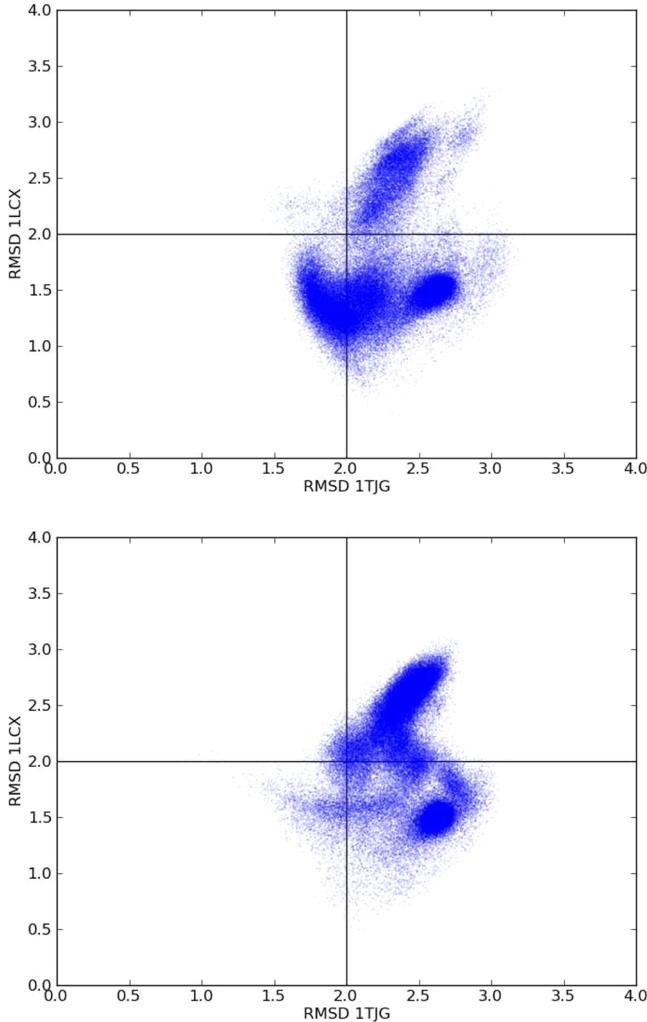


Figure 8. Scatter plot of the RMSD values of the epitope ELDKWA with respect to the corresponding sequence motifs of the peptide in the 1TJG crystal structure (*x*-axis) and in the 1LCX NMR structure (*y*-axis). The coordinate data are from the unrestrained MD simulation under explicit ff99SB (top) and explicit ff12SB (bottom).

on the root-mean-square deviations (RMSD) values of the backbone atoms N, O, C', and C_α with respect to the solution NMR structure (1LCX) of gp41_{659–671} and with respect to the crystal structure 1TJG, except that the simulation coordinates belong to the umbrella sampling runs with end-to-end distance of 28 Å. The net effect of stretching the polymer in explicit ff99SB (top panel) results in approximately the same spread in RMSDs with respect to 1TJG but augments the population with conformations closer to 1LCX. On the other hand, the

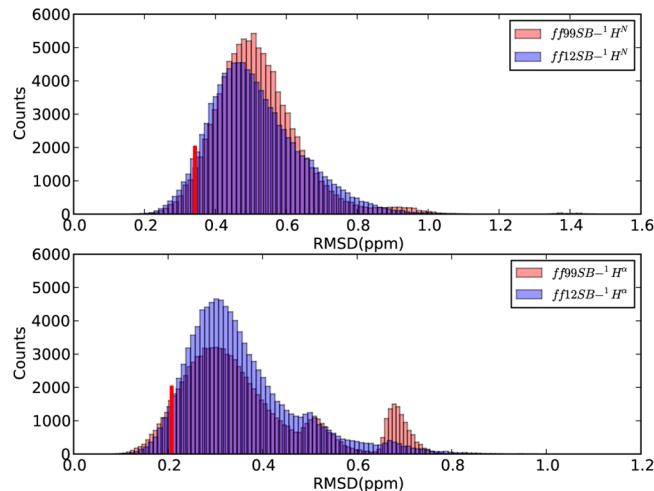


Figure 9. RMSD of predicted chemical shifts with respect to experimental chemical shifts⁶ for HN (top) and HA (bottom). The experimental data correspond to chemical shifts reported for 1LCX, and the predicted data are from the simulations under explicit ff99SB (red) and explicit ff12SB (blue). The red bar shows the RMSD of the predicted chemical shift for the 1LCX NMR structure with respect to the experimentally provided value.⁶

effect of stretching the polymer in explicit ff12SB is slightly opposite: it gives approximately the same spread in RMSDs with respect to 1LCX but increases the spread of RMSDs with respect to 1TJG that extend from 0.5 Å to more than 3.5 Å. Remarkably, there is an important population contribution to the left quadrants, i.e., close to the extended crystal structure in 1TJG, including a small population of RMSDs of approximately 0.5 Å. Implicit solvent representations (not shown) suppress the extended conformations and are closer to 1LCX.

■ DISCUSSION

Comparison to Previously Published Experimental Data. The NMR study by Biron et al.⁶ of the peptide in aqueous solution at pH 7.7 found a monomeric, amphiphilic 3_{10} -helix with minor random coil representation⁶ (Figure 1), a finding that seemed to be confirmed by CD measurements carried out by the same group. In particular, the exposed surface of the helix was found to consist of conserved, hydrophobic HIV-1 residues. The reconstructed, energy-minimized structure found a 3_{10} -helix conformation with 3.2 residues per turn. This analysis was carried out for the zwitterionic form of the peptide, but the authors observed additional helical stabilization with acetylation and amidation of the peptide. However, a second NMR study and a CD spectrum analysis both carried out by Barbato et al.¹⁰ found no major population of 3_{10} -helical conformers, but a mixture of various conformers. These authors showed that all the 13 residues in gp41_{659–671} were important for binding the human monoclonal antibody 2F5, as the flanking residues induce local structure (rather than providing additional contacts to 2F5). The authors also proposed that gp41 exhibits a “conformational plasticity” where the final configuration depends on the microenvironment. Thus, the peptide could undergo a conformational transition between disordered and helical structures as it approaches the lipid membrane, a fact that was later further explored and confirmed by Gregor et al.²³ Close to the membrane, in membrane-mimetic environments, or in conformations that involve dimers or trimers, the X-ray

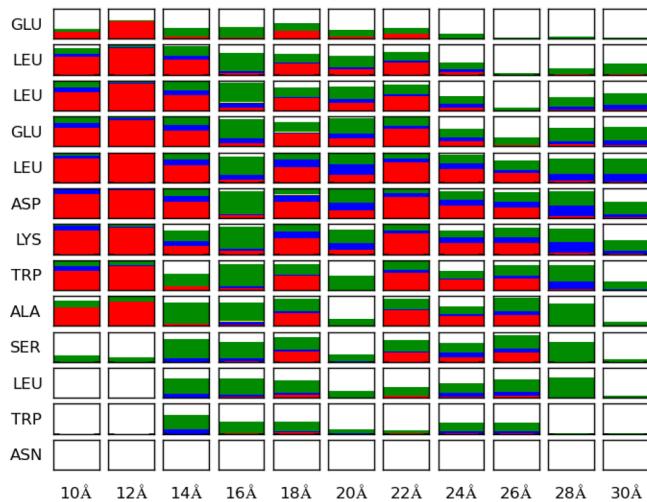
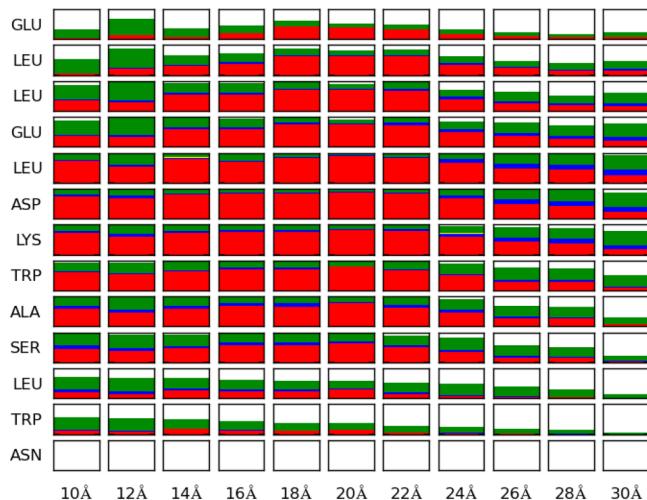


Figure 10. Secondary structure calculated with STRIDE for the ff99SB umbrella sampling runs. The *y*-axis shows the residue, and the *x*-axis shows the end-to-end distance used as restraint in the umbrella sampling runs. Each small window represents an average for the given residue over 100 ns. Top graph corresponds to implicit solvent and bottom graph to explicit solvent. Red, blue, yellow, green, and white correspond to α -helix, 3_{10} -helix, π -helix, turn, and “other” secondary structural motifs, respectively.

and NMR structures of gp41_{659–671} are found to be mainly helical (see for instance PDB IDs 1ENV,² 2PV6,⁴⁰ 1JAV,³ 3G9R,⁴² 3H01,⁴³ and 2X7R⁴⁴).

A UVRSS investigation¹¹ revealed a rough energy landscape with a wide variety of conformations. In particular, an examination of the conformation-sensitive AmIII₃ region suggested that gp41_{659–671} exhibits a broad distribution of conformations that includes significant population of β -turns as well as 3_{10} -helix and π -helix motifs but little α -helix. The authors of this work also found a lack of temperature-induced change (between 1 and 30 °C) in the AmIII₃ band intensities, a fact that they interpreted as indicative of an energy landscape where folded and unfolded conformations occupy broad, degenerate free energy minima. Approximately half of the peptides were found in folded conformations (predominantly 3_{10} - and π -helices) and the other half in unfolded (β -turn/PPII) conformations. However, the assignation of secondary structure to this type of spectra is not straightforward and not unequivocal.^{11,12} Recently, a far-UV CD spectroscopy study

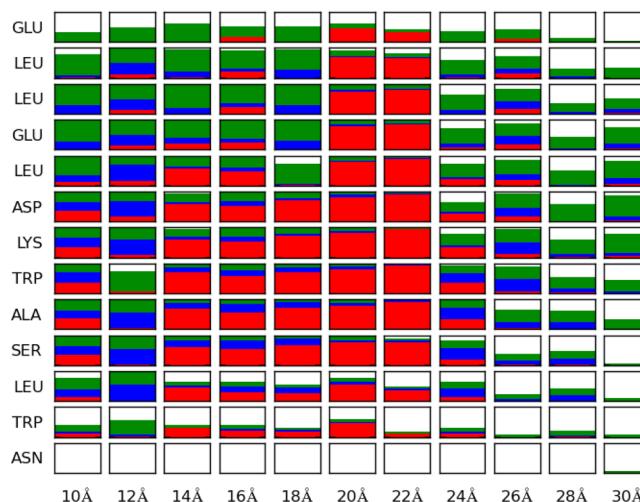
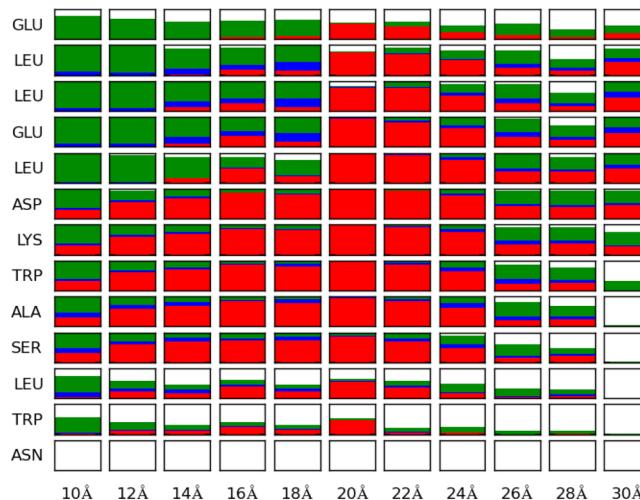


Figure 11. Secondary structure calculated with STRIDE for the ff12SB umbrella sampling runs. The *y*-axis shows the residue, and the *x*-axis shows the end-to-end distance used as restraint in the umbrella sampling runs. Each small window represents an average for the given residue over 100 ns. Top graph corresponds to implicit solvent and bottom graph to explicit solvent. Red, blue, yellow, green, and white correspond to α -helix, 3_{10} -helix, π -helix, turn, and “other” secondary structural motifs, respectively.

carried out by Tulip et al.¹² also revealed the conformational plasticity of gp41_{659–671}, which exhibited different competing folding propensities and no stable α -helical, 3_{10} -helical, or turn motifs. The spectrum of gp41_{659–671} at 5 °C in various aqueous conditions showed a strong negative band just before 200 nm and a negative shoulder at approximately 225 nm, which are indicative of disordered polypeptides with residues in the α_R and β regions of the Ramachandran plot.

Our simulations are consistent with the bulk of these experimental findings. First, the amount of helical population is important in aqueous solution, but this structure forms part of a flexible conformational ensemble with a rugged free energy, with shallow minima, as reported by Barbato et al.¹⁰ and Tulip et al.¹² This is strongly supported by the Ramachandran plots in Figure 4. The analysis of secondary structure, which depends partially on the criteria used, shows that the 3_{10} -helix is present but in small amounts: for explicit ff12SB its population reaches 13% under STRIDE or 21% under DSSP. For all cases, we find negligible π conformation, which is in agreement with the

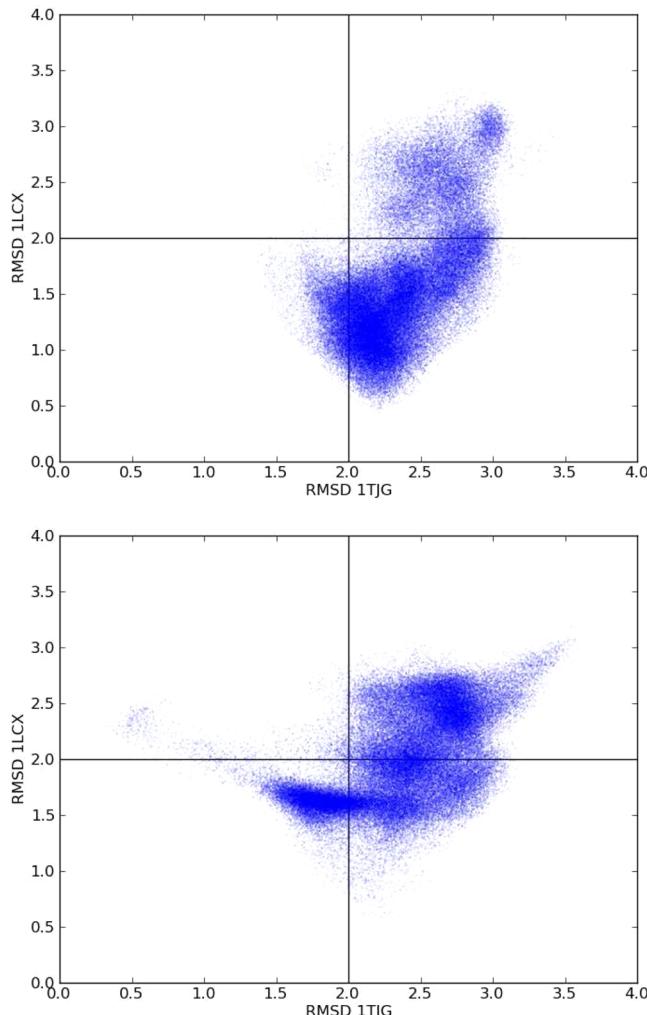


Figure 12. Scatter plot of the RMSD values of the epitope ELDKWA with respect to the corresponding sequence motifs of the peptide in the 1TJG crystal structure (*x*-axis) and in the 1LCX NMR structure (*y*-axis). The coordinate data are from the umbrella simulation under explicit ff99SB (top) and explicit ff12SB (bottom) for an end-to-end distance of 28 Å.

knowledge that π -helices are seldom found to be stable experimentally. Mainly, the simulations results exhibit a large presence of α -helical populations, turns, and loops, as clearly shown in Figures 5 and 6. Still, the explicit solvent conformations show important populations whose epitope is relatively close to the NMR structure 1LCX of Biron et al. as well as an ensemble of conformations comprising other structures (Figure 8). The relatively extended conformation of Ofek et al.¹³ seems to be under-represented in the regular MD simulations in Figure 8 (left quadrants). However, in the crystal structure gp41_{659–671} is bound to 2F5, and this clearly imposes restraints in the conformational space. For instance, the binding stretches the conformation¹³ such that the peptide spans roughly 25 Å from the C_α of residue Glu₆₅₉ to the C_α of residue Trp₆₇₀. In our restrained umbrella runs, the end-to-end distance collective variable is defined as the distance between the peptide-bond nitrogen atoms in Glu₆₅₉ and Asn₆₇₁, and the restraining harmonic potential has a weak constant of 1 kcal/(mol Å²). Thus, the conformations at an end-to-end distance of 28 Å can (weakly) mimic some of the effects of binding, and a comparison between these structures and 1TJG is more fair.

This is done in Figure 12, which for explicit ff12SB shows a non-negligible population of conformations close to 1TJG (left quadrants), including some conformations whose RMSDs are about 0.5 Å.

Comparison to Previous Simulations. As stated in the Introduction, there have been a few simulation studies of this peptide. Martins Do Canto et al.¹⁷ used the Gromos96 43a1 force field to examine different peptides (T-1249 and T-20) based on the C-region of HIV gp41. In particular, they also examined two shorter homologous peptides, one of which was gp41_{659–671} (called 3f5 in their work). The authors found no stable structures in aqueous solution, but an equilibrium of interchanging coil structures, as well as non-negligible populations of turns, bends, and (less) α -helices. This force field, however, has been shown to exhibit conformational biases,¹⁸ tending to underestimate α -helices and overestimate π -helices. The authors used a cutoff of 10 Å for the electrostatic interactions, which—as the authors stated—could affect the results. The authors concluded that the peptides have no stable structure in solution in the studied time scale (100 ns) and also supported the idea of conformational plasticity. This conformational plasticity was also previously found in a study by Lapejosa et al.⁴⁵ for various peptides containing the epitope ELDKWA. The authors used the OPLS force field^{46,47} in implicit solvent and the program STRIDE for secondary structure calculation. The authors found that the heptamer ELDKWAS is highly flexible and its conformational propensities in solution can include both helical structures and type I β -turns, which would be of help for intra- and intermolecular hydrophobic interactions. The heptamer in solution, however, does not necessarily exhibit the same properties of gp41_{659–671} in solution as it lacks the interactions between the flanking residues.

Another recent MD study by Gregor et al.²³ employed the CHARMM22, AMBER ff03, and AMBER ff99SB force fields with the TIP3P water model. In all cases, the results show considerably far more α -helical content than experimentally observed, with CHARMM22 giving the worst agreement and ff03 being closer to experiments. CHARMM22 was also used in a second study by the same group,¹² and the MD results were in poor agreement with the accompanying experimental data. All three force fields were later shown to have secondary structure biases. For instance, in a validation study of different force fields against experimental data, CHARMM22 fails to maintain the structure of protein GB3 (which unfolds during the simulation), and both CHARMM22 and ff03 in general overstabilize helices.²⁴ On the other hand, ff99SB is believed to destabilize helices.²⁴

The simulation results by Gregor et al. disagree considerably with the experimental data, especially in the case of ff99SB in explicit water, which greatly overestimates the α -helical content with respect to the CD data. As we are particularly interested on how well the AMBER force fields perform on disordered peptides and intrinsically disordered proteins, we carried out our simulations specifically comparing the two most recent force fields with the TIP3P water model. In particular, we found that the ff99SB force field with the TIP3P water model produces different results than those reported by Gregor et al. This can be appreciated by comparing the bottom panel in Figure 5 and the bottom panel of Figure 4 in ref 23 (the secondary structures in both figures have been computed with STRIDE). Except for the terminal residues, the results in ref 23 show a much higher α -helical content than our results. One

obvious difference in the simulations is that the previous authors employed the zwitterionic form of the peptide, while we are using capped ends. This, however, favors the opposite trend: the capped peptide favors the α -helical conformation more than its zwitterionic version. Given that we are using the same force field, the same water model, the same treatment of electrostatics, and the same program for the analysis of the secondary structure, we speculate that the main sources of disagreement can be attributed to the size of the simulation box and/or to the amount of sampling. With respect to the system size, we notice that Gregor et al. used only 33% of the numbers of waters that we used. Given that the peptide is disordered and in principle it can adopt a fully unfolded conformation, it is important to pad the box with a large number of waters to avoid periodic artifacts which could indeed stabilize a helical conformation (especially in a zwitterionic form): we chose the minimum number of waters that seemed to be "safe". A definite answer to what the minimum number of waters for this system would probably require a careful finite effect (i.e., periodic artifact) study. With respect to the amount of sampling, we notice that the simulations in ref 23 only extend to 15 ns. Although we show only 200 ns of data, we actually carried out each simulation for 300 ns (see computational details). Thus, the extensive α -helical conformations in ref 23 could simply be due to lack of sampling. For example, the first 20 ns window for explicit ff99SB in our Figure 5 (which corresponds to the last 20 ns in a 40 ns run) corresponds mostly to α -helical conformations and qualitatively resemble Figure 4 in ref 23. However, later there are also extended periods where "turns" are more stable (these were found unstable in ref 23). All in all, our simulation data seem to be more consistent with the experimental results and with the concept of a rugged free energy with multiple conformations.

Comparison between the Force Fields. We are interested on how well the AMBER force fields can describe disordered peptides and intrinsically disordered proteins, whose satisfactory description is more challenging. Thus, we compared a brand new AMBER force field, mainly ff12SB²⁵ against the more extensively tested ff99SB, whose performance has been deemed satisfactory by a number of studies^{22,26} but which has also been found to suffer a number of inaccuracies in backbone and side-chain dihedral angles.^{24,27} Basically, ff12SB builds upon ff99SB with modified parameters for the backbone ϕ' torsion angle and for all side-chain torsions, except Gly, Ala, and Pro.

To first order, the force fields are in good agreement. The implicit and explicit simulations for each force field yield essentially the same free energy for conformations near equilibrium and for the compact structures. However, as the end-to-end distance grows, the polymer in explicit solvent tends to make more hydrogen bonds with waters, and therefore it is easier to stretch than the polymer in implicit solvent, which has more intramolecular hydrogen bonds. Other general trends for gp41_{659–671}, valid independently of whether one computes the secondary structure with STRIDE or with DSSP, are the following. The implicit ff99SB gives the largest number of α -helical conformations, while explicit ff99SB and both implicit and explicit ff12SB are comparable in this respect. The percentage of 3₁₀-helices is larger for explicit ff12SB, and the percentage of loop structures is larger for explicit ff99SB. The major differences between the two force fields are in N-terminal ELLEL sequence, where refinements in the side-chain torsions

of Glu and Leu decrease dramatically the α -helical content of ELLEL in ff12SB.

Force fields are parametrized near equilibrium conformations, and comparisons far from equilibrium present a more stringent test of their performance. Here, we stretched the polymer away from its equilibrium end-to-end distance. The differences between the two force fields as reflected in Figures 10 and 11 are in general larger than the differences near equilibrium as shown in Figure 7. Implicit ff99SB always shows more helical content than the explicit version, while ff12SB shows better agreement between implicit and explicit solvents. Unfortunately, in spite of the measurable differences that we have described, we cannot unequivocally discern from the simulations which force field is better. The stretch simulations perhaps favor ff12SB as the better field, as there is least disagreement between the implicit and explicit solvent representations and as ff12SB encompasses in its conformational ensemble a larger population of turns, including those structures—when the polymer is stretched—that approach the crystallographic structure.

Conformational Equilibrium in Stretched Conformations. To study the mechanical properties of polypeptides, it is customary to carry out single-molecule stretching experiments using, for instance, atomic force microscopy or optical tweezers. The results are generally shown as a plot of the tensile force as a function of the end-to-end distance of the polymer. Computationally, this can be easily achieved, e.g., by using steered molecular dynamics (SMD). In this work, we have used umbrella sampling, so instead of pulling the peptide continuously, we take pictures of the equilibrium conformations under different (rather weak) end-to-end distance restraints.

First, the equilibrium, unbiased free energy is similar to that obtained in many other simulation approaches, such as in the case of alanine-rich peptides explored with adaptively biasing techniques.⁴⁸ The stretched conformations in the implicit simulations (top panels in Figures 10 and 11) reveal a smooth behavior with stretching: for each residue, the α -helix population along a row (as the conformations are stretched) increases from its initial value until it reaches a maximum in the window centered at 20 Å. After that it decreases with a concomitant increase in 3₁₀-helices, turns, and disordered extended structures, which is qualitatively similar to the results in ref 48. An analogous behavior can also be observed in explicit ff12SB, although the α -helical content decreases abruptly after 22 Å, giving rise to many more 3₁₀-helices, turns, and disordered extended structures than the implicit solvents, in agreement with the free energy maps in Figure 3.

However, one main difference between our work and most work on stretching helical peptides is precisely given by the fact that at the free energy minimum gp41_{659–671} is not a helical but a disordered peptide (with some helical content), and its α -helical content increases to its maximum on stretching, before it starts decreasing again. This agrees with the theoretical prediction that a partially helical chain should increase its helicity upon tensile stretching by an applied force due to the loss of the conformational entropy of the disordered state under uniaxial tension.^{49–51} This situation however has not been verified experimentally at the single molecular level due to the difficulties associated with determining the secondary structure of a single polypeptide chain,⁵² although it has been confirmed at the macroscopic level.^{53,54}

The differences in configurations and resulting free energies between the implicit and explicit solvents beyond the maximum

helical content provides an indirect confirmation of the theoretical argument proposed by Chakrabarti and Levine^{55,56} that the tensile “melting” of a partially helical chain will be determined by the stretching mode of the locally denatured parts of the polymer. This local denaturation can take place spontaneously due to local energy fluctuations. In the explicit solvent the carbonyl of the backbones can form hydrogen bonds with the waters,⁵⁷ providing an additional source of fluctuations over the implicit solvent and the consequent easier melting of the helical chain in explicit waters.

CONCLUSION

In this work, we have studied the conformational ensembles of gp41_{659–671} under the most recent AMBER force fields ff99SB and ff12SB in both implicit and explicit waters via MD for a cumulative time longer than 7.2 μ s. Intrinsic disordered peptides are hard to properly characterize, but our results are consistent with the bulk of the experimental findings. The amount of helical population is important in aqueous solution, but this structure forms part of a flexible conformational ensemble with a rugged free energy landscape with shallow minima. The newest force field ff12SB seems to perform slightly better than ff99SB, especially under restraints that take the peptide away from its equilibrium end-to-end distance. When the originally disordered peptide at the equilibrium end-to-end distance corresponding to the free energy minimum is subject to uniaxial stretching, it first becomes almost fully helical due to the loss of the conformational entropy of the disordered state under uniaxial tension. Upon further stretching, the helix melts into turns, loops, and 3₁₀-helices. The explicit waters provide additional hydrogen bonds that facilitate the “melting” of the helical structure and the lowering of the free energy when compared with the implicit solvent. Important populations in the conformational ensemble show epitope conformations that are close to the NMR aqueous solution structure (1LCX) of Biron et al.⁶ as well as epitope conformations close to the very different crystal structure of the peptide bound to 2F5 (1TJH) of Ofek et al.¹³ The latter agreement is obtained under a weak end-to-end distance restraint that mildly mimics one of the several effects of the binding to 2F5.

AUTHOR INFORMATION

Corresponding Author

*E-mail sagui@ncsu.edu, Ph (919)515-3111 (C.S.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the National Science Foundation (Grants 1021883 and 1148144). We also thank the NC State HPC Center for extensive computational support.

REFERENCES

- (1) Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. Core Structure of gp41 from the HIV Envelope Glycoprotein. *Cell* **1997**, *89*, 263–273.
- (2) Weissenhorn, W.; Dessen, A.; Harrison, S. C.; Skehel, J. J.; Wiley, D. Atomic Structure of the Ectodomain from HIV-1 gp41. *Nature* **1997**, *387*, 426–430.
- (3) Tan, K. M.; Liu, J. H.; Wang, J. H.; Shen, S.; Lu, M. Atomic Structure of a Thermostable Subdomain of HIV-1 gp41. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12303–12308.
- (4) Caffrey, M.; Cai, M. L.; Kaufman, J.; Stahl, S. J.; Wingfield, P. T.; Covell, D. G.; Gronenborn, A. M.; Clore, G. M. Three-dimensional Solution Structure of the 44 kDa Ectodomain of SIV gp41. *EMBO J.* **1998**, *17*, 4572–4584.
- (5) Nabel, G. J. Challenges and Opportunities for Development of an AIDS Vaccine. *Nature* **2001**, *410*, 1002–1007.
- (6) Biron, Z.; Khare, S.; Samson, A.; Hayek, Y.; Naider, F.; Anglister, J. A Monomeric 3(10)-Helix is Formed in Water by a 13-residue Peptide Representing the Neutralizing Determinant of HIV-1 on gp41. *Biochemistry* **2002**, *41*, 12687–12696.
- (7) Purtscher, M.; Trkola, A.; Gruber, G.; Buchacher, A.; Predl, R.; Steindl, F.; Tauer, C.; Berger, R.; Barrett, N.; Jungbauer, A.; Katinger, H. A Broadly Neutralizing Human Monoclonal antibody Against GP41 of Human-Immunodeficiency-Virus Type-1. *AIDS Res. Hum. Retroviruses* **1994**, *10*, 1651–1658.
- (8) Parker, C. E.; Deterding, L. J.; Hager-Braun, C.; Binley, J. M.; Schulke, N.; Katinger, H.; Moore, J. P.; Tomer, K. B. Fine Definition of the Epitope on the gp41 Glycoprotein of Human Immunodeficiency Virus Type 1 for the Neutralizing Monoclonal Antibody 2F5. *J. Virol.* **2001**, *75*, 10906–10911.
- (9) Sattentau, Q. J.; Zollapazner, S.; Poignard, P. Epitope Exposure on Functional, Oligomeric HIV-1 GP41 Molecules. *Virology* **1995**, *206*, 713–717.
- (10) Barbato, G.; Bianchi, E.; Ingallinella, P.; Hurni, W.; Miller, M.; Cilibertol, G.; Cortese, R.; Bazzo, R.; Shiver, J.; Pessi, A. Structural Analysis of the Epitope of the Anti-HIV Antibody 2F5 Sheds Light into its Mechanism of Neutralization and HIV Fusion. *J. Mol. Biol.* **2003**, *330*, 1101–1115.
- (11) Ahmed, Z.; Asher, S. A. UV Resonance Raman Investigation of a 3(10)-helical Peptide Reveals a Rough Energy Landscape. *Biochemistry* **2006**, *45*, 9068–9073.
- (12) Tulip, P. R.; Gregor, C. R.; Troitzsch, R. Z.; Martyna, G. J.; Cerasoli, E.; Tranter, G.; Crain, J. Conformational Plasticity in an HIV-1 Antibody Epitope. *J. Phys. Chem. B* **2010**, *114*, 7942–7950.
- (13) Ofek, G.; Tang, M.; Sambor, A.; Katinger, H.; Mascola, J. R.; Wyatt, R.; Kwong, P. D. Structure and Mechanistic Analysis of the Anti-human Immunodeficiency Virus Type 1 Antibody 2F5 in Complex with its gp41 Epitope. *J. Virol.* **2004**, *78*, 10724–10737.
- (14) Ofek, G.; Guenaga, F. J.; Schief, W. R.; Skinner, J.; Baker, D.; Wyatt, R.; Kwong, P. D. Elicitation of Structure-specific Antibodies by Epitope Scaffolds. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 17880–17887.
- (15) van Gunsteren, W. F.; Daura, X.; Mark, A. E. *Encyclopedia of Computational Chemistry*; John Wiley and Sons: New York, 2002; Chapter The GROMOS force field.
- (16) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596–3607.
- (17) Martins Do Canto, A. M. T.; Carvalho, A. J. P.; Ramalho, J. P. P.; Loura, L. M. S. T-20 and T-1249 HIV Fusion Inhibitors’ Structure and Conformation in Solution: a Molecular Dynamics Study. *J. Pept. Sci.* **2008**, *14*, 442–447.
- (18) Matthes, D.; de Groot, B. L. Secondary Structure Propensities in Peptide Folding Simulations: A Systematic Comparison of Molecular Mechanics Interaction Schemes. *Biophys. J.* **2009**, *97*, 599–608.
- (19) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (20) 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. A Point-charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
- (21) Lee, M. C.; Duan, Y. Distinguish Protein Decoys by Using a Scoring Function Based on a New Amber Force Field, Short Molecular Dynamics Simulations, and the Generalized Born Solvent Model. *Proteins* **2004**, *55*, 620–634.

- (22) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712–725.
- (23) Gregor, C. R.; Cerasoli, E.; Tulip, P. R.; Ryadnov, M. G.; Martyna, G. J.; Crain, J. Autonomous Folding in the Membrane Proximal HIV Peptide gp41(659–671): pH Tuneability at Micelle Interfaces. *Phys. Chem. Chem. Phys.* **2011**, *13*, 127–135.
- (24) 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* **2010**, *78*, 1950–1958.
- (25) Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; et al. AMBER 12; University of California: San Francisco, CA, 2012.
- (26) 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.
- (27) Best, R. B.; de Sancho, D.; Mittal, J. Residue-Specific alpha-Helix Propensities from Molecular Simulation. *Biophys. J.* **2012**, *102*, 1462–1467.
- (28) Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; et al. AMBER 11; University of California: San Francisco, 2010.
- (29) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.
- (30) Weiser, J.; Shenkin, P. S.; Still, W. C. Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). *J. Comput. Chem.* **1999**, *20*, 217–230.
- (31) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins* **2004**, *55*, 383–394.
- (32) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (33) Darden, T. A.; York, D. M.; Pedersen, L. G. Particle Mesh Ewald: An N log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (34) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (35) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (36) Frishman, D.; Argos, P. Knowledge-Based Secondary Structure Assignment. *Proteins* **1995**, *23*, 566–579.
- (37) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22*, 2577–2637.
- (38) Joosten, R. P.; Te Beek, T. A. H.; Krieger, E.; Hekkelman, M. L.; Hooft, R. W. W.; Schneider, R.; Sander, C.; Vriend, G. A Series of PDB Related Databases for Everyday Needs. *Nucleic Acids Res.* **2011**, *39*, D411–D419.
- (39) Han, B.; Liu, Y.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: Significantly Improved Protein Chemical Shift Prediction. *J. Biomol. NMR* **2011**, *50*, 43–57.
- (40) Sun, Z.-Y. J.; Oh, K. J.; Kim, M.; Yu, J.; Brusic, V.; Song, L.; Qiao, Z.; huai Wang, J.; Wagner, G.; Reinherz, E. L. HIV-1 Broadly Neutralizing Antibody Extracts Its Epitope from a Kinked gp41 Ectodomain Region on the Viral Membrane. *Immunity* **2008**, *28*, 52–63.
- (41) Schibli, D. J.; Montelara, R. C.; Vogel, H. J. The Membrane-proximal Tryptophan-Rich Region of the HIV Glycoprotein, gp41, Forms a Well-Defined Helix in Dodecylphosphocholine Micelles. *Biochemistry* **2001**, *40*, 9570–9578.
- (42) Liu, J.; Deng, Y.; Dey, A. K.; Moore, J. P.; Lu, M. Structure of the HIV-1 gp41 Membrane-Proximal Ectodomain Region in a Putative Prefusion Conformation. *Biochemistry* **2009**, *48*, 2915–2923.
- (43) Liu, J.; Deng, Y.; Li, Q.; Dey, A. K.; Moore, J. P.; Lu, M. Role of a Putative gp41 Dimerization Domain in Human Immunodeficiency Virus Type 1 Membrane Fusion. *J. Virol.* **2010**, *84*, 201–209.
- (44) Buzon, V.; Natrajan, G.; Schibli, D.; Campelo, F.; Kozlov, M. M.; Weissenhorn, W. Crystal Structure of HIV-1 gp41 Including Both Fusion Peptide and Membrane Proximal External Regions. *PLoS Pathog.* **2010**, *6*, 880.
- (45) Lapelosa, M.; Gallicchio, E.; Arnold, G. F.; Arnold, E.; Levy, R. M. In Silico Vaccine Design Based on Molecular Simulations of Rhinovirus Chimeras Presenting HIV-1 gp41 Epitopes. *J. Mol. Biol.* **2009**, *385*, 675–691.
- (46) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- (47) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (48) Henin, J.; Schulten, K.; Chipot, C. Conformational Equilibrium in Alanine-Rich Peptides Probed by Reversible Stretching Simulations. *J. Phys. Chem. B* **2006**, *110*, 16718–16723.
- (49) Tamashiro, M. N.; Pincus, P. Helix-coil Transition in Homopolypeptides under Stretching. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2001**, *63*, 021909–1–012909–8.
- (50) Saito, N.; Go, M. Helix-Coil Transition of Polypeptides under an External Force. *J. Phys. Soc. Jpn.* **1968**, *24*, 376–379.
- (51) Doi, M. Helix-Coil Transition under External Forces. *Polym. J. (Tokyo, Jpn.)* **1974**, *6*, 222–229.
- (52) Ikai, A. Tensile Mechanics of α -Helical Coil Springs. *Adv. Polym. Sci.* **2010**, *232*, 65–96.
- (53) County, S.; Gornall, J. L.; Terentjev, E. M. Induced Helicity in Biopolymer Networks under Stress. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13457–13460.
- (54) County, S.; Gornall, J. L.; Terentjev, E. M. Mechanically Induced Helix-Coil Transition in Biopolymer Networks. *Biophys. J.* **2006**, *90*, 1019–1027.
- (55) Chakrabarti, B.; Levine, A. J. Nonlinear Elasticity of an Alpha-helical Polypeptide: Monte Carlo Studies. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2006**, *74*, 031903–1–031903–11.
- (56) Afrin, R.; Takahashi, I.; Shiga, K.; Ikai, A. Tensile Mechanics of Alanine-Based Helical Polypeptide: Force Spectroscopy versus Computer Simulations. *Biophys. J.* **2009**, *96*, 1105–1114.
- (57) Gnanakaran, S.; Hochstrasser, R. M.; Garcia, A. E. Nature of Structural Inhomogeneities on Folding a Helix and their Influence on Spectral Measurements. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9229–9234.
- (58) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.