Molecular Dynamics Simulation of Folding of a Short Helical Peptide with Many Charged Residues

Chung-Cheng Wei, Ming-Hsun Ho,† Wen-Hung Wang, and Ying-Chieh Sun*

Department of Chemistry, National Taiwan Normal University, 88, TingChow Road Section 4, Taipei 116, Taiwan

Received: May 5, 2005; In Final Form: June 17, 2005

A molecular dynamics simulation of the folding of conantokin-T (con-T), a short helical peptide with 5 helical turns of 21 amino acids with10 charged residues, was carried out to examine folding pathways for this peptide and to predict the folding rate. In the 18 trajectories run at 300 K, 16 trajectories folded, with an averaged folding time of ∼50 ns. Two trajectories did not fold in up to 200 ns simulation. The folded structure in folded trajectories is in good agreement with experimental structure (Skjærbæk; et al. *J. Biol. Chem.* **1997**, *272*, 2291. Lin; et al. *FEBS Lett.* **1997**, *407*, 243). An analysis of the trajectories showed that, at the beginning of a few nanoseconds, helix formation started from residues 5−9 with assistance of a hydrophobic clustering involving Tyr5, Met8, and Leu9. The peptide formed a U-shape mainly due to charge−charge interactions between charged residues at the N- and C-terminus segments. In the next ∼10 ns, several nonnative charge−charge interactions were broken and nonnative Gla10-Lys18 (this denotes a salt bridge between Gal10 and Lys18) and/or Gla10-Lys19 interactions appeared more frequently in this folding step and the peptide became a fishhook J-shape. From this structure, the peptide folded to the folded state in 7 of all 16 folded trajectories in ∼15 ns. Alternatively, in ∼30 ns, the con-T went to a conformation in an L-shape with 4 helical turns and a kink at the Arg13 and Gla14 segment in the other 9 trajectories. Con-T in the L-shape then required another ∼15 ns to fold into the folded state. In addition, in overall folding times, the former 7 trajectories folded faster with the total folding times all shorter than 45 ns, while the latter 9 trajectories folded at a time longer than 45 ns, resulting in an average folding time of ∼50 ns. Two major folding intermediates found in 2 nonfolded trajectories are stabilized by charge clusters of 5 and 6 charged residues, respectively. With inclusion of friction and solvent−solvent interactions, which were ignored in the present GB/SA solvation model, the folding time obtained above should be multiplied by a factor of 1.25−1.7 according to a previous, similar simulation study. This results in a folding time of 65−105 ns, slightly shorter than the folding time of 127 ns for an alanine-based peptide of the same length. This suggests that the energy barrier of folding for this type of peptides with many charged residues is slightly lower than alanine-based helical peptides by less than 1 kcal/mol.

I. Introduction

Recent investigations of protein/peptide folding principles using experimental[1] and theoretical[2−37] methods have broadened and deepened our understanding of the folding kinetics of proteins and peptides. Advances in computational simulation have allowed the folding pathways of protein/peptides to be examined, complementarily to experiments. In addition to an examination of the folding pathways, the folding rates of a number of peptides were found to be in excellent agreement with experimental results, demonstrating the good prediction ability of the simulations.[3,4,7,8,10,11,13] A simulation of the Trp-cage peptide gave a folding time of 4 μs, in excellent agreement with the experimental result of 6 μs.[6] Simulations of alanine-based helical peptides gave folding times of 16−20 and 127 ns for two peptides that were of 16 and 21 amino acids in length,[5,13] respectively, also in excellent agreement with experimental results.[38,39] In the former simulation, it was found that the breaking of nonnative hydrophobic clustering is the rate-determining step in the folding of the alanine-based helical peptide.[5] The energetics of folding pathways was analyzed and was used to aid in the interpretation of folding kinetics.

While recent peptide folding simulation studies have focused more on alanine-based helical peptides, we examined the folding of a helical peptide with many charged residues. To our knowledge, the role of side chain charge−charge interactions in the folding of a helical peptide with many charged residues has not been examined before using the first-principle MD folding simulation. This topic is interesting because, in addition to interacting with other charged side chains, a charged side chain can interact strongly with solvent as well. The magnitude of these two interactions and their effects on peptide folding are open problems in establishing peptide folding principles for peptides with many charged residues. In the present study, we selected a small helical peptide of 21 amino acids of a single-strand helix, conantokin-T, from the protein data bank (pdb code: 1ont).[40] The folding MD was carried out to examine the folding pathways and to predict the folding rate of this peptide.

This peptide was originally isolated from the venom of piscivorous cone snails *Conus Tulipa*, possessing *N*-methyl-D-aspartate receptor antagonist activity, with code sequence

* To whom correspondence should be addressed. E-mail: sun@cc.ntnu.edu.tw.

† Present address: Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104-6323.

10.1021/jp052349k CCC: $30.25
Published on Web 10/05/2005

GE$\gamma\gamma$Y-QKMLγ-NLRγA-EVKKNA-NH_2. γ denotes the nonnatural Gla residue with a -$CH_2CH(COO^-)_2$ side chain. In a solution at pH = 5.5, at which the experiment was conducted,[40,41] the peptide contains 10 charged residues. CD and ^{1}H NMR spectroscopy experiments indicated that all amino acids in the peptide are involved in a helical conformation, forming a length of 5 helical turns in this peptide in the folded state, as shown in Figure 2. In addition, another conantokin, con-G, was experimentally investigated[40] as well. The effects of divalent counterions, low pH, and solvent in structure were also examined.[40] In the present folding MD simulations, starting from an extended, unfolded structure, a folding simulation for this peptide was carried out. Eighteen trajectories were carried out, 16 folded within 200 ns at 300 K, giving an average folding time of ~50 ns. The MD simulation we used and the analysis methods are described in Methods. The folding pathways and dominant interactions are described and analyzed in section III. The conclusion is given in section IV.

II. Methods

The Amber7 package[42] was used to simulate the folding of this peptide. The N- and C-termini were capped with ACE and NHE, respectively. The charge state of the carboxylates on the side chains of Glu2 and Glu16 were set to −1 because the NMR structure of con-T was obtained in a solution at pH = 5.5. A newly developed force field developed by Duan et al. was employed to describe interactions between the atoms of natural amino acids.[43] For the nonnatural amino acid Gla, the force field parameters were derived using the Antechamber module in Amber7, and a quantum chemistry calculation was carried out to derive the charge of atoms in the Gla residue. The ACE-Gla-NME model molecule was constructed in an α-helix conformation and subjected to a B3LYP/6-31G** level calculation using Gaussian98.[44] In deriving the partial charges of atoms, for simplicity, we took the charges of 6 atoms (N, H, C_α, H_α, C, and O) in the backbone of Glu as the charges of atoms in the backbone of Gla. The charges of atoms in the side chain were adjusted on the basis of the Mulliken charges obtained from the quantum chemistry calculation such that the sum of the charges of the atoms in the Gla residue is −2 due to the presence of two carboxylate groups in the side chain. The adjusted charges differ slightly from the Mulliken charges, as listed in Table S1 in the Supporting Information in which other force field parameters employed are given as well. For the MD simulation, the cutoff distance was set to 100 Å to include all atom−atom nonbonded interactions. For solvation, the GB/SA implicit solvation model developed by the Case group was employed to account for interactions between the peptide and the solution.[45]

Starting from the NMR structure, an energy minimization of 5000 steepest descent and subsequent 10 000 conjugate gradient steps was first carried out. To see whether this structure is stable at 300 K with the present force field in the simulation, an MD simulation of this peptide for 10 ns at 300 K followed starting from the energy-minimized structure. All the bonds associated with hydrogen atoms were fixed at their equilibrium bond lengths. The time step was 2 fs. The temperature coupling was set to 5.0 ps. The computation gave a low root mean square deviation (RMSD) curve of ~1 Å using the NMR structure as the reference structure. This shows that the present computational protocol reproduced the NMR stable structure well. The average energy of the trajectory was −1158 kcal/mol.

For the folding simulation, we generated an unfolded structure of this peptide by running MD at 1000 K for 3 ns. A most extended structure of the RMSD value of 8.1 Å was picked up in the MD trajectory as the unfolded structure of con-T for a later folding simulation. This structure was then subsequently subjected to an energy minimization. Starting from this extended structure, 18 MD simulations at 300 K were carried out with 18 different random number seeds for the velocities of the atoms. The simulations proceeded in lengths of 30−200 ns and were terminated when con-T folded into the helical conformation of a single strand. Sixteen trajectories folded within 200 ns. Two trajectories did not fold in up to 200 ns and were trapped in intermediate states. For those 16 folded trajectories, MD simulations were continued after the time points con-T folded for an additional 15−20 ns in each trajectory. Snapshot structures were saved at 1 ps intervals for 15 trajectories and at 10 ps intervals for those 3 trajectories that were longer than 150 ns. To examine folding pathways in the trajectories, a conformational cluster analysis[46] based on the RMSD values between snapshot structures in a trajectory was employed to cluster conformations of con-T and to analyze conformational transitions in the trajectories. The structures of the RMSD values within a criterion value of 2 Å were clustered as a conformation group in the trajectories. It should be noted that the conformation groups of top 5 populations, which comprise >80% of the total snapshot structures in each trajectory, were found to adequately describe the major folding events of con-T folding. The calculated results and analysis are presented in the next section.

III. Results and Discussion

(a) Folding Pathways: Structure, Conformational Transition, and Interaction. RMSD curves for 18 trajectories are shown in Figure 1. The folding pathways of con-T are summarized in Figure 2. Figure 3 shows the average (i, $i + 4$) H-bond formation for 18 trajectories. The folded structure obtained from the simulation is shown in Figure 2. In the 18 run trajectories, 16 trajectories folded within 200 ns, with an average time of ~50 ns. The folding time was estimated when the entire helix of a single strand formed in a trajectory. The other 2 trajectories did not fold in up to a 200 ns simulation. Their RMSD curves are shown in the last two curves of high RMSD values in Figure 1. Representative peptide structures along the folding pathways are shown in Figure 2 (see below for further discussion). The simulation predicts the folding time of this peptide to be ~50 ns. To compare with an alanine-based helical peptide of the same length,[13] for which the TIP3P explicit water model was employed, the friction and solvent−solvent interactions, which were ignored in the present GB/SA solvation model, need to be considered. A previous study[13] for an alanine-based helical peptide has shown that a factor of 1.25−1.7 should be multiplied to account for the friction effect in folding time. However, to our knowledge, the magnitude of this factor for peptides with many charged residues is not available. A future study using an explicit water model should provide this information. On the basis of current knowledge of this friction effect, multiplication of 1.25−1.7 gave a folding time of 60−105 ns for con-T. This folding time is slightly shorter than the folding time of 127 ns for an alanine-based peptide of 21 amino acids investigated previously.[13] The predicted shorter folding time suggests that the energy barrier for con-T folding is slightly lower than the alanine-based peptide by less than 1 kcal/mol. In contrast to the alanine-based peptide, in which the breaking of nonnative hydrophobic clustering was found to be the rate-determining step in the folding of an alanine-based helical

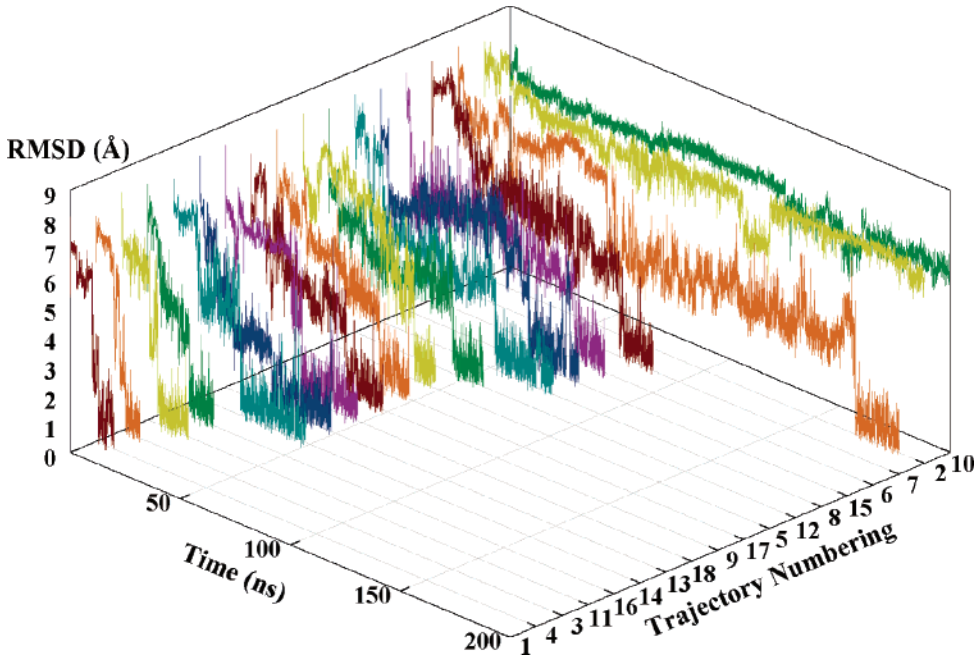

Figure 1. Root-mean-square deviation (RMSD) curves of trajectories shown in a 3D figure. To show them clearly, the trajectories were ordered in increasing order of folding time. The RMSD was calculated for the backbone N, Cα, C atoms using the NMR structure as the reference structure. Sixteen trajectories folded within 200 ns simulation and 2 did not (see text for detailed description). Plateau regions in curves can correspond to the folding steps summarized in Figure 2 based on the calculated results of a conformational cluster analysis described in the methods section (see text for detailed description).

peptide,[5] in con-T, charge−charge interactions are the dominant interactions in determining the folding pathways and the folding rate (see below for further discussion). The folded structure obtained from 16 folded trajectories is a helix of a single strand of 5 helical turns, in excellent agreement with NMR structure.[40] In addition, similar to the NMR structure,[40] the side chains of the nonpolar residues are located at the concave side of the peptide. This is likely due to the hydrophobic effect. However, a slight difference is noted, in which the MD structure is slightly straighter than the NMR structure. In addition to these observations, the MD simulation provides for the time evolution of side chain interactions explicitly in the folding course. In the 16 folded trajectories, Gla4-Lys7, Gla10-Arg13, Gla14-Lys18, and Glu16-Lys19 charge−charge interactions and a hydrophobic cluster involving Tyr5, Met8, Leu9, and Leu12 were observed in the folded state, showing the significance of these interactions in stabilizing the helical conformation of con-T (see further discussion in next paragraphs). The former 4 charge−charge interactions are referred to as native interactions in the folded state. These 4 salt bridges are spread over almost the entire sequence of the con-T peptide. The latter hydrophobic cluster locally stabilizes 1−2 helical turns at the location of residues 5−9 only. As described in the NMR study,[40] the hydrophobic effect was found not to be the major source of stability for this peptide.

The folding pathways summarized in Figure 2 are based on the calculated results of the conformational cluster analysis described in Methods. It was found that the folding pathways of con-T can be roughly divided into 3 or 4 folding steps from the initial unfolded state to the final folded state with 2 or 3 intermediates, characterized by their shapes, which we refer to as U, J, and L shapes, respectively, as shown in Figure 2. The categorization of structures in trajectories into 3 or 4 main folding steps was found to strongly correlate to the time evolution of the RMSD curves and (i, $i + 4$) H-bond formation shown in Figures 1 and 3, respectively. The first event of con-T folding from the unfolded state in the trajectories is the formation of a helical turn at the segment of residues 5−9, which occurred in the first few nanoseconds in the simulation, and was observed in all trajectories. In all trajectories, a hydrophobic clustering of Tyr5, Met8, Leu9, and Leu12 assisted in the formation of this first helical turn. As this hydrophobic clustering took place, charged side chains interact with each other, making mostly nonnative salt bridges and the peptide resulting in a U-shape. In this folding step, the salt bridge interactions between the negative charged groups at the N-terminus segment, Glu2, Gla3, and Gla4, and the positively charged groups at the C-terminus segment, Lys18 and Lys19, are dominant charge−charge interactions in the 16 folded trajectories. The other 2 nonfolded trajectories are stabilized by charge clusters involving 5 and 6 charged side chains, respectively (see the below subsection of found intermediates below for a detailed description).

As the folding proceeded after the first folding event described above in the 16 folded trajectories, in about 10 ns, the nonnative charge−charge interactions between N- and C-terminus segments began to break. In the meantime, the helix turn segment began to lengthen from the helical turn located at the segment of residues 5−9 to the middle segment around Leu12. In addition, the occasional occurrence of a helical turn at location of Lys18 took place in this time period. These are clearly seen in Figure 3 of (i, $i + 4$) H-bond formation averaged over 18 run trajectories. An examination of 16 folded trajectories showed that the Gla10-Arg13 and Gla14-Lys18 native charge−charge interactions became more stable. Nonnative interactions of Val17 with hydrophobic cluster of Tyr5, Met8, and Leu9 weakened, and nonnative charge−charge interactions of Gla10-Lys18 (Lys19) appeared frequently in this fishhook J-shape folding step.

As seen in Figure 3 for α-helix H-bond formation, in the time period of about 20−50 ns, the helical turn located at residues 15−17 lengthened to both sides. In 7 of the 16 folded trajectories, con-T went into the folded state directly from the J-shape conformer, as indicated in Figure 2. The other 9 trajectories went to L-shape ("titled L-shape") intermediates.

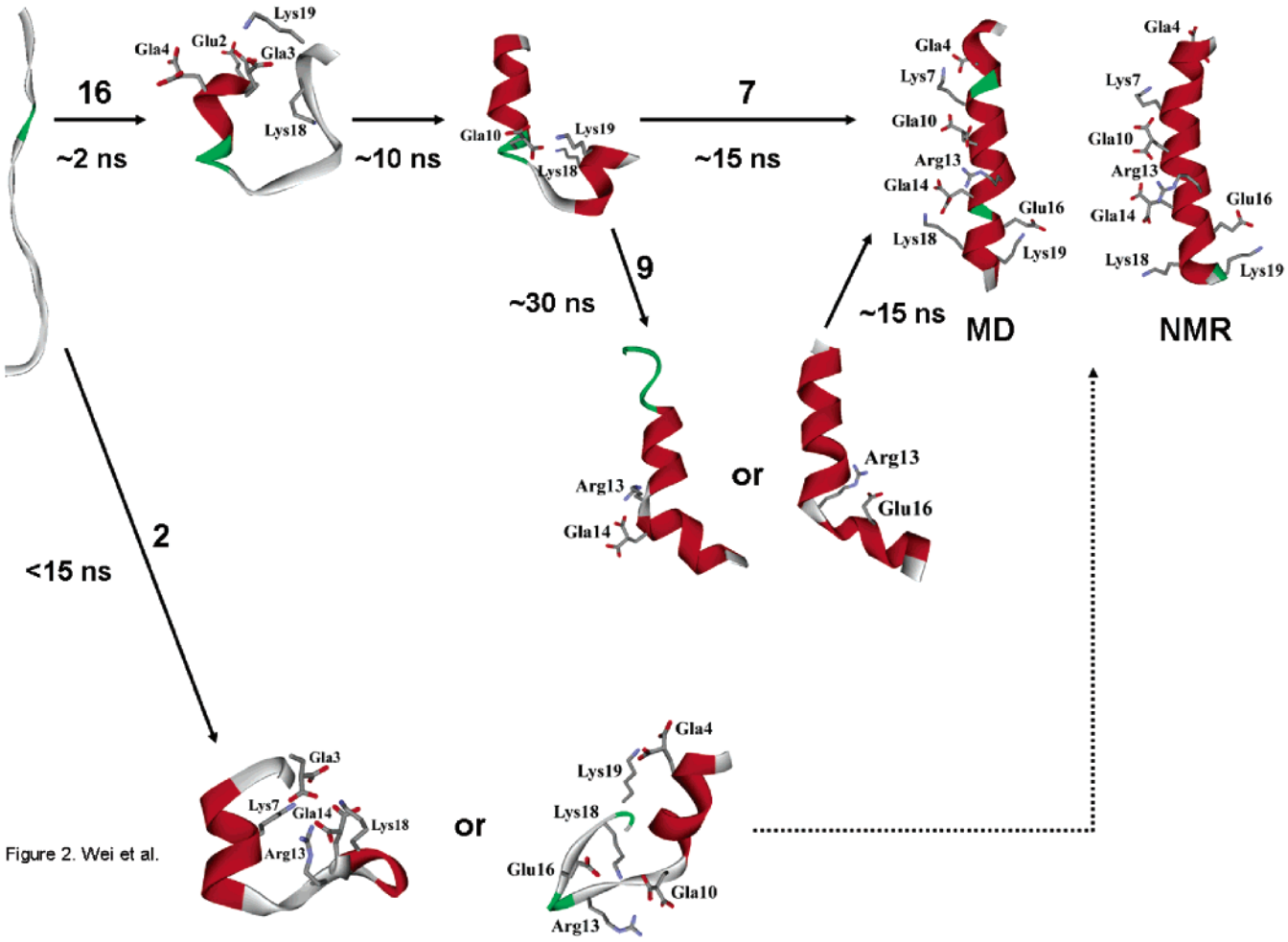

Figure 2. Summary of folding pathways. Peptide structures and interactions of interest in each folding step are shown. Numbers on arrow signs are the number of trajectories going through that conformational transition. The NMR structure (pdb code: 1ont) is shown beside the folded structure obtained from the MD simulation. The time below each arrow sign is the averaged, approximate time for that conformational transition in the folding pathways. The side chains of nonnative charge−charge interactions are shown for the intermediates in the folding pathways. Native charge−charge interactions are shown in the MD folded structure (see text for detailed description). A structure from the NMR structure ensemble is shown beside the MD folded structure. Charged side chains in the NMR structure, in which 7 salt bridges were speculated to be main contributors in stabilizing the con-T folded structure based on CD and ^{1}H NMR spectroscopy, are shown as well (see text for further discussion). Oxygen atoms on the side chains are shown in red; nitrogen atoms on the side chains, in blue.

The folding times of the former 7 trajectories are all shorter than 45 ns, while the latter 9 trajectories have folding times longer than 45 ns, with an average folding time of ∼50 ns. In the latter case, breaking of Gla10-Lys18 (Lys19) nonnative charge−charge interactions and the formation of Gla14-Lys18 and Glu16-Lys19 native salt bridges turned the J-shape peptide into an L-shape before folding into the folded state, as illustrated in Figure 2. This folding step required ∼30 ns. In this L-shape structure, con-T almost became a single-strand helix of 5 complete helical turns except for a "kink" located around residue 15. Native charge−charge Gla10-Arg13, Gla14-Lys18, and Glu16-Lys19 interactions and the interaction of Arg13-Gla14 formed, but the backbone around residues 13 and 14 was not in the native conformation, with con-T resulting in an L-shape. Four of these 9 trajectories had additional Arg13-Glu16 nonnative charge−charge interactions, and one had an additional Gla10-Lys18 nonnative charge−charge interaction. The breaking of these two nonnative interactions led con-T to the folded state in these 5 trajectories. In another 4 trajectories, it was observed that, although native charge−charge interaction formed, the charge−charge "interaction chain" of Gla10-Arg13-Gla14-Lys18 was in a nonnative conformation. A slide motion of this interaction chain, mainly the motion of the pair of two oxygen atoms on the Gla14 side chain over the pair of nitrogen atoms on the Arg13 side chain, completed the formation of a single-strand helix of 5 turns and led the con-T to the folded state. This step from the L-shape intermediate to the folded state required, on average, ∼15 ns in these 9 folded trajectories. The folding of con-T through this L-shape intermediate pathway required a much longer time than the pathways for the 7 trajectories in which the L-shape intermediate was not involved. About half of all the folded trajectories proceed by folding through the L-shape intermediate and weight significantly in the estimated folding time of ∼50 ns for con-T folding.

(b) Found Intermediates. Among 18 run trajectories, 2 trajectories were not folded in up to a 200 ns simulation. The RMSD curves of these 2 trajectories are shown in the last two curves for high RMSD values in Figure 1. As described above, a helix turn at the segment of residues 5−9 formed in early in the time scale of a few nanosceconds in con-T was observed as well in these 2 trajectories. In one of these 2 trajectories, the con-T, in the first 30 ns, is in a conformation, in which a side chain charge cluster involving Glu2, Gla3, Lys7, Arg13, Gal14, Lys18, and Lys19 was observed. Other charged side chains, Gla4, Gal10, and Glu16 were somewhat solvated and pointed outward in this conformation. The RMSD value of this

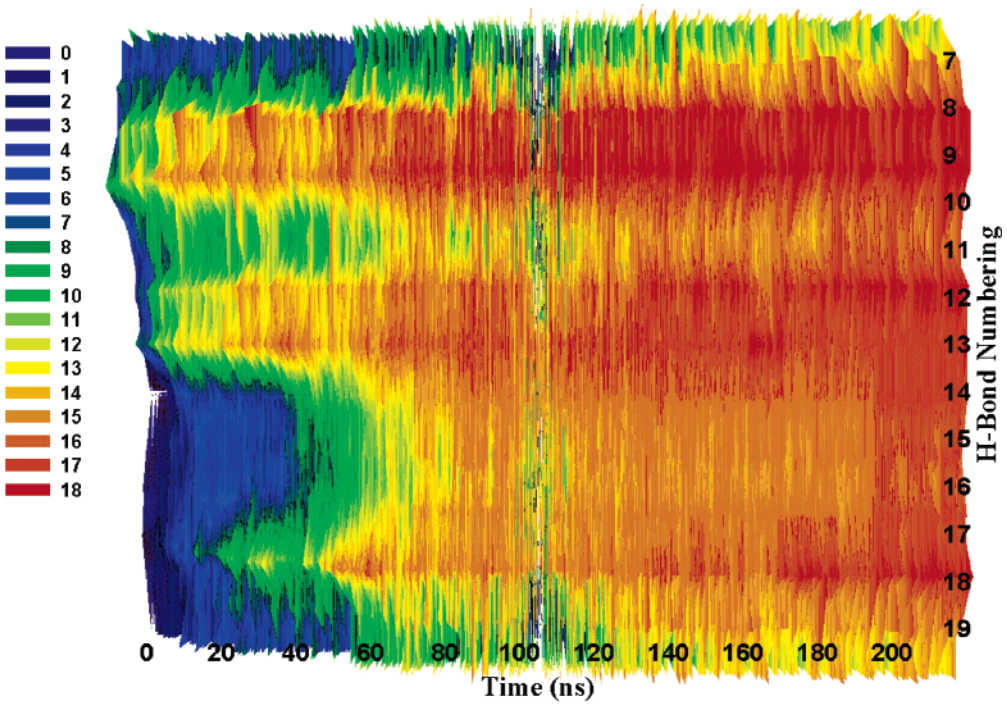

Figure 3. (i, $i + 4$) backbone hydrogen bond formation averaged over a run of 18 trajectories. The numbering in the y-axis refers to the ($i + 4$)th residue. For example, 7 stands for the H-bond between NH of residue 7 and CO of the residue 3. Color codes the number of trajectories in which a specific (i, $i + 4$) bond occurs in the 18 run trajectories at each time point. Hydrogen bond definition: the N of the NH group and the O of the CO group are within 3.5 Å, and the angle of N−H−O is larger than 120°. For those 16 folded trajectories, trajectories of 15−20 ns after con-T folded were copied to fill the time period between the ends of the folded trajectories and 200 ns in each trajectory, mimicking an equilibrium run after con-T folded. The figure is the top view of a 3D figure in which x,y,z-axes represent simulation time, H-bond numbering, and the number of trajectories out of 18 trajectories in which H-bonds formed, respectively. The value in the z-axis is shown in color ranging from 0 (in blue) to 18 (red), representing the number of trajectories having a H-bond at a specific time in the simulation trajectories and the location in the peptide sequence. It is clearly seen that con-T completed folding at ∼50 ns. The small blank region in the middle of the figure is because of a visual result of the 3D image drawn by the software, Sigmaplot, which does not relate to H-bond formation.

conformation was ∼7 Å. As the folding proceeded, the RMSD value decreased to 6−6.5 Å, fluctuating around these values until the end of the 200 ns simulation. At ∼30 ns, at which time the RMSD curve decreased from 7 to 6−6.5 Å; the charge cluster described above reshuffled and adjusted. Glu2 swung outward and was solvated, leaving a cluster of Gla3, Lys7, Arg13, Gal14, and Lys18 (or Lys19), as shown in Figure 2. An examination of this trajectory indicated that Gla3, Lys7, Arg13, and Gla14 remained in the cluster in parts of the trajectory of the period of the RMSD value of 6−6.5 Å. The Lys18 and Lys19 switched in participating in the charge cluster, while the con-T underwent structural fluctuation during this period. These findings suggest that the side chains of Gla3, Lys7, Arg13, and Gla14 are core participants in this cluster, which stabilize this intermediate structure. Note that, although the net charge of this cluster is −2, solvation of these charged side chains should serve to screen down the charge−charge interactions significantly. The occasional involvement of Lys18 and Lys19 should neutralize this charge cluster more. In addition to these interactions in the trajectory, at the C-terminus site, a helix turn was observed at the Glu16 to Lys19 segment. An examination of the structure showed that the native Glu16-Lys19 side chain charge−charge interaction assisted in the formation of this helix turn. This helix turn occurred at ∼30 ns after the stable nonnative charge cluster in the intermediate state had formed in this trajectory and remained for the remainder of this trajectory. To compare the energy of this intermediate with the folded-state structure, a calculation of the average energy value in 130−200 ns of this trajectory gave a value of −1149 kcal/mol, 9 kcal/mol higher than the energy of −1158 kcal/mol for the folded state.

In another nonfolded trajectory, two major intermediates occurred at about 40 and 160 ns and existed in the trajectory for about 110 and 40 ns, respectively. The structure of the former is shown in the second structure of the found intermediates in the 2 nonfolded trajectories in Figure 2. The structure of the latter is not shown. In the former, Gla4-Lys19 and Glu16-Lys18 nonnative charge−charge interactions along with the Gla10-Arg13 native interaction but in a nonnative orientation comprise a charge cluster, stabilizing this intermediate. The Leu12 and Val17 were distant from the hydrophobic cluster of Tyr5, Met8, and Leu9. In the latter intermediate, Val17 and Leu12 were close to Leu9, forming a hydrophobic cluster, while Tyr5 and Met8 were in different orientations and rather distant from the Leu9 side chain. It should be noted here that, in addition to the 5 nonpolar residues discussed above, the other 2 nonpolar residues, Ala15 and Ala21, were found not to have a significant involvement in the hydrophobic clustering in con-T folding. Glu2-Lsy18 and Gla3-Lys19 interactions are nonlocal, nonnative charge−charge interactions, stabilizing this intermediate, while Arg13-Glu16 is a minor, local interaction involved in the charge cluster composed of these charged residues in this intermediate state. These results suggest that to stabilize these intermediates in nonfolded trajectories, 2−3 nonnative, nonlocal interactions between residues at the N- and Cterminus segments along with the assistance of some other minor, local charge−charge interactions are needed. The former nonnative interactions can be charge−charge interactions close to the ends of the N-and C-termini between negatively charged residues Glu2, Gla3, Gla4, Gla10 and positively charged residues Arg13, Lys18, and Lys19, or the Lys7-Gal14 charge−charge interaction close to the middle segment, or the nonnative hydrophobic cluster of

Leu9, Leu12, and Val17. These interactions caused the con-T to become a U-shape with a deeper energy well than the U-shape conformers in the folded trajectories. Some other minor local charge−charge interactions may contribute to the stabilization of these intermediates to some extent.

(c) Simulation at 360 K. In addition to the simulations at 300 K, simulations at 360 K were also carried out to examine the thermal effect in the folding of this peptide. Ten MD trajectories were run and all folded within a 30 ns simulation time, giving an average folding time of ∼10 ns, faster than ∼50 ns at 300 K. An examination of the trajectories indicated that the folding pathways at 360 K are not significantly different from 300 K. This suggests that the shape of the free energy surface of folding at 360 K is similar to that at 300 K except that the magnitude of the energy barriers changes slightly. The folded structures in the ends of the trajectories showed a high helical content of ∼80%, which is higher than the NMR experimental value of 28% at 348 K,[41] for which backbone (ϕ, φ) angles for α-helix formation criterion[5] was used. It should be noted here that, using a tighter criterion in the (ϕ, φ) angle criterion or the (i, $i + 4$) H-bond in terms of the distance and angle between NH and CO can result in a lower helical content. The calculated helical content at 360 K is consistently lower than that at 300 K, although the difference between 300 K and the higher temperature is smaller than the experimental result.[40] This can be attributed to the overstabilization of the helical conformation in the force field at this high temperature. It should be noted that most of the current force fields were developed to fit to the experimental and/or high-level theoretical results at room temperature. Nevertheless, weak temperature dependence of the free energy surface and folding pathways of structured peptides were reported,[21,37] supporting that high-temperature simulation is useful in examination of folding pathways of structured peptides. The interactions stabilizing the intermediates, described in the previous paragraph, are much more easily broken at 360 K. Finally, the found folding rate result suggests that, at 360 K, the reduction of free energy barrier of folding was mainly due to the entropy contribution. An MD simulation at this temperature can speed up folding of this peptide.

IV. Conclusion

In summary, an MD simulation of the folding of the con-T peptide was carried out to examine its folding pathways and to predict folding rate. The folding time with the GB/SA solvation model in computation was predicted to be ∼50 ns at 300 K. Structures and interactions in the folding pathways were described and analyzed. In contrast to alanine-based peptides, hydrophobic interactions in con-T are not the major interaction affecting the peptide folding pathways although a clustering of 3 nonpolar side chains assists on the formation of a helical turn locally in the early folding of con-T. Charge−charge interactions are major factors affecting the folding pathways of con-T. Time evolution for the breaking of nonnative charge−charge interactions and the formation of native interactions in con-T folding gave a picture of the folding process in 3 or 4 folding steps, and the folding intermediates are characterized by U, J, and L shapes, respectively, between the initial unfolded and final folded states in the folded trajectories. The pathways going through the L-shape were found to have a longer folding time than those that did not proceed through this state and account for half of all folded trajectories, resulting in a significant weight of these pathways in the estimated folding time for con-T. Interactions of significance were examined and analyzed in each folding step. In addition, the conformation characteristics of 3 intermediates in 2 nonfolded trajectories and interactions that stabilize these intermediates were described and analyzed. These calculated results can be useful in interpreting the structural stability principle of this peptide and folding kinetics in future experiments. The prediction of a merely slightly shorter folding time of this 21-mer α-helix peptide with 10 charged residues than an alanine-based peptide of the same length should stimulate for further investigations of the role of charge−charge interactions in peptide folding dynamics in either experimental or theoretical investigations. The computed results for the folding pathways, folding rate, and intermediate structures in this peptide folding can then be verified in future experiments.

Acknowledgment. We thank Prof. Yong Duan for helpful discussions and sending us the 2003 force field prior to its availability in the Amber package. We also thank Hwung-Wen Chen for useful discussion and computer technical support. NSC is acknowledged for providing funding. NCHC and the computer center at Science College of NTNU are acknowledged for providing partial CPU time. Our gratitude also goes to the Academic Paper Editing Clinic, NTNU.

Supporting Information Available: Force field parameters for the nonnatural amino acid, Gla. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) See, for example, (a) Eaton, W. A.; Munoz, V.; Hagen S. J.; Jas, G. S.; Lapidus, L. J.; Henry, E. R.; Hofrichter, J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327. (b) Chen, R. P.; Huang, J. J.; Chen, H. L.; Jan H.; Velusamy M.; Lee, C. T.; Fann, W.; Larsen, R. W.; Chan S. I. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7305. (c) Chen, P. Y.; Gopalacushina, B. G.; Yang, C. C.; Chan, S. I.; Evans, P. A. *Protein Sci.* **2001**, *10*, 2063, and references therein.

(2) Gnanakaran, S.; Nymeyer, H.; Portman; J.; Sanbonmatsu, K. Y.; Garcia, A. E. *Curr. Opin. Struct. Biol.* **2003**, *13*, 168.

(3) Chowdhury, S.; Lee, M. C.; Xiong, G.; Duan, Y. *J. Mol. Biol.* **2003**, *327*, 711.

(4) Chowdhury, S.; Lee, M. C.; Xiong, G.; Duan, Y. *J. Phys. Chem. B* **2003**, *108*, 13855.

(5) Chowdhury, S.; Zhang, W.; Wu, C.; Xiong, G.; Duan, Y. *Biopolymers* **2003**, *68*, 63.

(6) Simmerling, C.; Strockbine, B.; Roitberg, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11258.

(7) Garcia, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13898.

(8) Vila, J. A.; Ripoll, D. R.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14812.

(9) Zhou, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13280.

(10) Chowdhury, S.; Lee, M. C.; Duan, Y. *J. Phys. Chem. B* **2004**, *108*, 13855.

(11) Liwo, A.; Khalili, M.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2362.

(12) Duan, Y. Kollman, P. A. *Science* **1998**, *282*, 740.

(13) Pande, V. S.; Baker, I.; Chapman, J.; Elmer, S. P.; Khaliq, S.; Larson, S. M.; Rhee, Y. M.; Shirts, M. R.; Snow, C. D.; Sorin, E. J.; Zagrovic, B. *Biopolymers* **2003**, *68*, 91.

(14) Ferrara, P.; Apostolakis, J.; Caflisch, A. *Proteins* **2002**, *39*, 252.

(15) Nymeyer, H.; Garcia, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13934.

(16) Pak, Y.; Jang, S.; Shin, S. *J. Am. Chem. Soc.* **2002**, *124*, 4976.

(17) Pak, Y.; Kim, E.; Jang, S. *J. Chem. Phys.* **2004**, *121*, 9184.

(18) Snow, C. D.; Zagrovic, B.; Pande, V. S. *J. Am. Chem. Soc.* **2002**, *124*, 14548.

(19) Sung, S. S. *Biophys. J.* **1999**, *76*, 164.

(20) Wang, H.; Sung, S. S. *J. Am. Chem. Soc.* **2000**, *122*, 1999.

(21) Snow, C. D.; Sorin, E. J.; Rhee, Y. M.; Pande, V. S. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 43.

(22) Islam, S. A.; Karplus, M.; Weaver, D. L. *Structure* **2004**, *12*, 1833.

(23) Cheng, X.; Cui, G.; Hornak, V.; Simmerling, C. *J. Phys. Chem. B* **2005**, *109*, 8220.

(24) Elber, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 151.

(25) Herges, T.; Wenzel, W. *Structure* **2005**, *13*, 661.

(26) Sorin, E. J.; Pande, V. S. *J. Comput. Chem.* **2005**, *26*, 682.

(27) Liwo, A.; Khalili, M.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2362.

(28) Ulmschneider, J. P.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2004**, *126*, 1849.

(29) Shea, J. E.; Onuchic, J. N.; Brooks, C. L., III. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16064.

(30) Li, F. Y.; Chen, Y.; You, Z. Q.; Mou, C. Y. *J. Chin. Chem. Soc.* **2002**, *49*, 783.

(31) Lei, H.; Duan, Y. *J. Chem. Phys.* **2004**, *121*, 12104.

(32) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2782.

(33) DeMarco, M. L.; Alonso, D. O. V.; Daggett, V. *J. Mol. Biol.* **2004**, *341*, 1109.

(34) Chowdhury, S.; Lei, H.; Duan, Y. *J. Phys. Chem. B*, in press.

(35) Tsai, Y. L.; Sun, Y. C. Submitted to *Biophys. J.*

(36) Hansmann, U. H. E.; Onuchic, J. N. *J. Chem. Phys.* **2001**, *115*, 1601.

(37) Cavalli, A.; Ferrara, P.; Caflisch, A. *Proteins* **2002**, *47*, 305.

(38) Hummer, G.; Garcia, A. E.; Garde, S. *Proteins* **2001**, *42*, 77.

(39) (a) Williams, S.; Causgrove, T. P., Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. *Biochemistry* **1996**, *35*, 691. (b) Thompson, P.; Eaton, W.; Hofrichter, J. *Biochemistry* **1997**, *36*, 9200.

(40) Skjærbæk, N.; Nielsen, K. J.; Lewis, R. J.; Alewood, P.; Craik, D. J. *J. Biol. Chem.* **1997**, *272*, 2291.

(41) Lin, C. H.; Chen, C. S.; Hsu, K. S.; King, D. S.; Lyu, P. C. *FEBS Lett.* **1997**, *407*, 243.

(42) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C.; Darden, T. Merz, K. M.;. Stanton, R. V.; Cheng, A.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G.. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. *Amber7 Users' Manual*; University of California: San Francisco, 2002.

(43) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Ang R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999.

(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople. J. A. *Gaussian 98*; Gaussian Inc.: Pittsburgh, PA, 1998.

(45) Tsui, V.; Case, D. A. *Biopolymers* **2001**, *56*, 275.

(46) Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins* **1999**, *34*, 269.