THE JOURNAL OF PHYSICAL CHEMISTRY B LETTERS

2005, *109,* 20064−20067
Published on Web 10/07/2005

Mechanism of Helix Nucleation and Propagation: Microscopic View from Microsecond Time Scale MD Simulations

Luca Monticelli,*,†,‡ D. Peter Tieleman,‡ and Giorgio Colombo*,†,§

Centre for Biomolecular Interdisciplinary Studies and Industrial Applications, University of Milan, 20131 Milan, Italy, Department of Biological Sciences, University of Calgary, Calgary, Alberta T2N 1N4, Canada, and Istituto di Chimica del Riconoscimento Molecolare, CNR via Mario Bianco, 9, 20131 Milan, Italy

Received: August 22, 2005; In Final Form: September 21, 2005

Microsecond time scale molecular dynamics simulations of the 13-residue peptide RN24 were carried out to investigate the mechanism of helix nucleation and propagation. An extended and an ideal α-helical conformation were used as starting structures. NOE-derived interatomic distances were compared with distances calculated from the simulations, showing good agreement between experimental and simulation results. Based on almost 200 helix nucleation events observed, β-turn and 3_{10}-helix play an important role in helix nucleation; in most cases, helix nucleation is preceded by the formation of a short-lived β-turn (60% probability) or 3_{10}-helix (20% probability), and the conversion from β-turn to α-turn involves bifurcated hydrogen bonds. Helix propagation in RN24 appears to occur preferentially from the N-terminus to the C-terminus, and helix unfolding preferentially in the opposite direction.

Introduction

Understanding protein folding is one of the greatest challenges in biophysics. Using fast time-resolved spectroscopic techniques, it has been shown that α-helix formation occurs on the nanosecond time scale, considerably faster than β-hairpin formation,[1,2] and within reach of molecular simulations.[3] Helix nucleation thermodynamics and kinetics have been studied intensively, using both experimental[1,2,4] and theoretical methods.[3,5−11] Still, the details of the mechanism of helix nucleation and propagation are not yet fully understood at atomic resolution. In this letter, we describe the detailed mechanism of helix nucleation and propagation as observed in microsecond time scale molecular dynamics simulations of the 13-residue α-helical peptide RN24.[12,13]

Methods

Two simulations were carried out at 300 K, starting from an extended and an ideal α-helical conformation, without any structural restraint to aid the folding process. The extended structure was chosen in order to avoid the introduction of any biases toward the helical conformation. The Gromacs software package[14,15] and the GROMOS G43A1 force field[16] were used, with the SPC water model.[17] The PME method was used for the calculation of the electrostatic potential,[18,19] while a twin-range cutoff (0.9/1.4 nm) was used for the Lennard-Jones interactions. All bond lengths were constrained to their equilibrium values using the SETTLE algorithm[20] for water and the LINCS algorithm[21] for the peptide. Dummy atom contructions were used, and the mass of hydrogen atoms was set to 4 au, according to a published procedure,[22] allowing for an integration time step of 5 fs. The introduction of dummy atoms and heavy hydrogen atoms does not influence the thermodynamic properties of the system, and dynamic properties have been shown to be influenced only to a moderate degree.[22] The neighbor list for the calculation of nonbonded interactions was updated every 4 time steps. The peptide and solvent were coupled separately to a temperature bath at 300 K, using the Berendsen algorithm with $\tau_T = 0.1$ ps,[23] and the pressure was kept at 1 bar using weak pressure coupling with $\tau_P = 4.0$ ps.[23] The peptide was placed in a dodecahedral box large enough to contain 1.2 nm of solvent around the peptide in helical conformation (2782 water molecules) and 0.6 nm of solvent for the extended conformation (2828 water molecules); XX, YY, and ZZ components of the dodecahedral box were about 4.9, 4.9, and 3.5 nm, giving a volume of ∼84 nm^3. The length of each trajectory was 1.6 μs, for a total simulation time of 3.2 μs. Secondary structure assignments were based on the DSSP algorithm.[24] The number of helical (N_h), extended (N_β) and turn (N_t) residues was calculated based on DSSP assignments. Hydrogen bonds were determined using a simple geometrical criterion: maximum donor−acceptor distance lower than 0.35 nm, donor−hydrogen−acceptor angle larger than 120°.

Results and Discussion

Circular dichroism (CD) measurements indicate the presence of approximately 50% helical conformation for the C-peptide, a peptide closely related to RN24, at the temperature of 1 °C.[25] It was later shown that the helical content above 30 °C is negligible for the C-peptide[26] and decreases sharply as temperature increases for RN24.[12] Moreover, NMR results on RN24

* Corresponding authors. E-mail: lmonti@ucalgary.ca; giorgio.colombo@icrm.cnr.it
† University of Milan.
‡ University of Calgary.
§ Istituto di Chimica del Riconoscimento Molecolare

10.1021/jp054729b CCC: $30.25

TABLE 1: Number of α-Helix and β-Hairpin Formation Events, and Their Mean First Passage Time, Calculated Using Two Different Lengths for Each Secondary Structure

	number of events − short[a]		MFPT[1] (ns)	number of events − long[b]		MFPT[b] (ns)
transition	EXT[c]	HEL[d]	(average)	EXT[c]	HEL[d]	(average)
coil → helix	105	88	4.24	26	26	22.17
helix → coil	105	88	3.71	26	27	10.73
coil → beta	278	263	1.78	43	32	12.17
beta → coil	278	264	1.73	43	32	9.11

[a] α-helix: 4 residues with helical conformation; hairpin: 4 residues hydrogen bonded in β-hairpin or β-sheet. [b] α-helix: 8 residues with helical conformation; hairpin: 6 residues hydrogen bonded in β-hairpin or β-sheet. [c] Simulation starting from the extended conformation. [d] Simulation starting from an ideal α-helical conformation.

have shown long-range NOE peaks that are not typical of helices, suggesting multiple conformations in aqueous solution that include both helical and bent structures.[13] Our simulations show a helical content around 10%, independent from the starting structure chosen, in reasonable agreement with experimental observations and with the Agadir prediction algorithm,[27] both in terms of the total helical content and the helical propensity as a function of the residue number (Figure S1); no violation over 0.05 nm of NOE-derived interproton distances was found (Table S1).

Multiple helix formation events were observed during the simulations, allowing an accurate description of the nucleation and propagation mechanisms in atomic detail. To characterize the mechanism of helix formation, we calculated the hydrogen bonding pattern and the secondary structure content during the simulations. The number of α-helix formation events was calculated using different requirements for the number of helical residues, from a minimum of 4 (here defined as "helix nucleus") to a maximum of 10 (longer helices were never formed during the simulations). Depending on the number of helical residues required, different numbers of α-helix formation events were detected. Table 1 shows the results for two helix lengths (4 and 8 residues), together with their mean first passage time (MFPT); the number of events and MFPT for the formation of beta structures is also shown. In the table, "helix" indicates only α-helical conformations, "beta" indicates both β-hairpin and β-sheet conformations, and "coil" indicates conformations that are neither "helix" nor "beta". All the transitions between α-helix and beta structures passed through a "coil" conformation. We expect qualitative agreement only for the MFPTs due to the limitations of the force field, limited sampling, and slightly altered kinetics due to the use of heavy hydrogens.[22]

It is clear from Table 1 that α-helix nucleation does not always result in the formation of a long α-helix: approximately one out of four helix nuclei transforms into an 8-residue α-helix. Since the transition state ensemble can be defined as the ensemble of structures with 50% probability of folding,[28−30] the helix nucleus as defined here (with 4 consecutive residues in helical conformation) is not representative of the transition state ensemble for the helix folding process as observed in our simulations. The formation and rupture of isolated hydrogen bonds occur on time scales very similar to that of helix propagation (tens/hundreds of picoseconds), while the average lifetime of an α-helix nucleus is approximately 4 ns, significantly shorter than for a long α-helix (∼11 ns). Similar time scales were found by others using MD simulations[30−34] and are compatible with experimental results.[1,2,4] Since the observed kinetics in the simulations is in reasonable agreement with experimental values, we expect that the mechanism of helix formation will also be realistic.

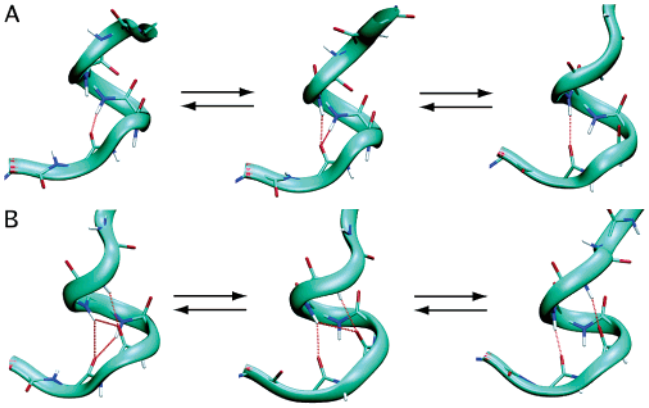

Figure 1. Most frequently observed mechanism of helix nucleation (A) and propagation (B). Red dotted lines indicate hydrogen bonds. Side chains are not shown for clarity.

Focusing on the peptide-water interaction, we observed a significant correlation between helix formation and backbone desolvation; the rupture of backbone−water hydrogen bonds assists the formation of backbone−backbone hydrogen bonds. This supports the suggestion that the use of explicit solvent may help avoid the unphysical trapping of peptide structures that may occur in implicit solvent simulations.[35]

The probability of nucleation was calculated for each residue, assigning nucleation events to the N-terminal residue in the helix nucleus (Figure S2). Residues Glu2 to Ala5 have significantly higher probabilities of being part of the helix nucleus compared to C-terminal residues. Based on an analysis of hydrogen bonding and secondary structure, we found two common mechanisms for helix nucleation. In the first mechanism (∼60% probability), a β-turn is quickly transformed into an α-turn; this shift usually occurs through a bifurcated hydrogen bond involving residues (i+3) and (i+4) at the same time (Figure 1A). The second mechanism (∼20% probability) involves the conversion of a β-turn into a short 3_{10}-helix, characterized by two consecutive (i, i+3) hydrogen bonds, followed by a hydrogen bond shift from the amide group of residue (i+3) to the one of residue (i+4). Modifying the criteria for the calculation of hydrogen bonds results in minor changes in the number and lifetime of the hydrogen bonds detected, but the sequence of events registered in helix nucleation is found to be completely independent of the geometrical criteria, as expected. Based on backbone dihedral angle analysis, the most commonly found β-turn nucleating helix formation is type II′. Although the population of isolated β-turns (not including β-hairpins) and 3_{10}-helices is low and their lifetime is short compared to α-helices, both secondary structure elements appear to play an important role in the initial stages of helix folding. While this has previously been predicted by others,[36−40] the atomic-detail description provided here is based on almost 200 nucleation events registered in unbiased MD simulations and allows us to predict the relative probability of each mechanism described.

Focusing on the direction of helix propagation, we found that N-terminus to C-terminus propagation has ∼60% probability, while the opposite direction is observed only in ∼15% of the registered events, and simultaneous propagation in ∼25%. When propagation occurs from the C-terminus to the N-terminus, helices are significantly shorter, reaching only 5.5 residues on average instead of 7.3 residues. The unfolding process appears to occur preferentially from the C-terminus to the N-terminus (see example in Figure 2B), which is consistent with results from previous MD simulations.[3,11,41−43] The propagation mechanism was also determined as a function of the helix length,

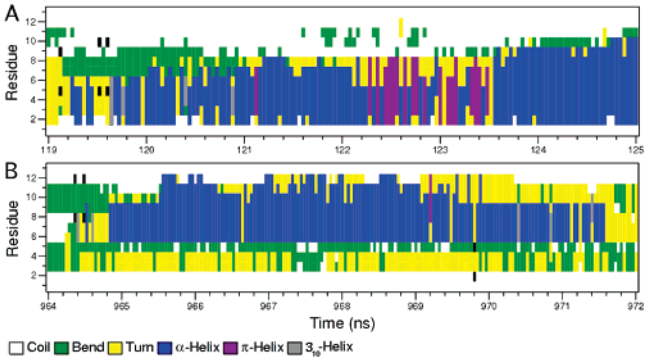

Figure 2. Secondary structure as a function of simulation time during two distinct helix formation events. (A) The formation of β-turns or 3_{10}-helices usually precedes the formation of a longer α-helix. (B) Helix propagation takes place preferentially from the N-terminus to the C-terminus, and helix unfolding in the opposite direction.

and was found to be completely independent of the number of residues in the helix: the same sequence of events is observed in the formation of α-helices of all possible lengths, as expected. Although the preference for a specific direction for helix propagation is clear from our simulations, we cannot conclude that this is a general feature of helix folding; the observed preference could be due to a sequence effect, because the N-terminal region of the peptide includes residues with higher helical propensity, according to the Agadir helical content prediction algorithm[27] (see Figure S1 and S2). Investigations on the physical basis for this preference are currently underway.

We found a relatively high population for the π-helical conformation, which may be an overestimate related to force field issues.[44] Despite their relatively high population, π-helices are generally short in the simulations (~5.7 residues on average) and their formation rarely precedes the formation of long α-helices. Also, bifurcated hydrogen bonds involving residues i and (i+4) or (i+5) are rarely observed in the initial stages of helix formation. Therefore, we conclude that the formation of π-helix conformations is important in reducing the conformational entropy of the peptide in the unfolded state,[45] but does not have a direct involvement in the nucleation of the α-helix.

Bifurcated hydrogen bonds are frequently observed also in the propagation mechanism. Two kinds of bifurcated hydrogen bond occur, involving either residues (i+3) and (i+4) or residues (i+4) and (i+5). The first type has a probability approximately 2.5 times higher than the second, and its formation normally precedes the formation of a longer α-helix (~75% probability; see example in Figure 1B), playing therefore an important role in helix propagation. Interestingly, the formation of bifurcated hydrogen bonds involving residue (i+5) usually precedes the formation of a π-helix (~70% probability) instead of an α-helix. Therefore, we conclude that bifurcated hydrogen bonds of the second type are not important in the α-helix propagation mechanism.

Conclusions

The present study provides an unprecedented view of the early stages of helix folding. Based on almost 200 helix nucleation events observed, we conclude that β-turn and 3_{10}-helix play an important role in helix nucleation; the conversion from β-turn to α-turn is the most commonly observed nucleation mechanism, with ~60% probability, and involves bifurcated hydrogen bonds. Helix propagation in RN24 occurs from the N-terminus to the C-terminus with ~60% probability, while the opposite direction was observed in only ~15% and simultaneous propagation in ~25% of the events. Helix unfolding takes places preferentially in the opposite direction. The formation of bifurcated hydrogen bonds between residue i and (i+3) or (i+4) normally precedes the formation of longer α-helices (~75% probability). Although in this paper we only studied the RN24 peptide, we expect that similar helix formation mechanisms apply in other natural peptides and possibly in proteins.

Acknowledgment. D.P.T. is a Senior Scholar of the Alberta Heritage Foundation for Medical Research (AHFMR). L.M. was supported by a postdoctoral fellowship by the CISI—University of Milan and the AHFMR.

Supporting Information Available: NOE-derived distances and violations (Table S1), helical propensity according to the Agadir algorithm (Figure S1), and nucleation probability for each residue (Figure S2) in PDF format. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) 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.
(2) Eaton, W. A.; Munoz, V.; Thompson, P. A.; Henry, E. R.; Hofrichter, J. *Acc. Chem. Res.* **1998**, *31*, 745.
(3) Hummer, G.; Garcia, A. E.; Garde, S. *Proteins* **2001**, *42*, 77.
(4) Werner, J. H.; Dyer, R. B.; Fesinmeyer, R. M.; Andersen, N. H. *J. Phys. Chem. B* **2002**, *106*, 487.
(5) Brooks, C. L. *J. Phys. Chem.* **1996**, *100*, 2546.
(6) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2782.
(7) Hiltpold, A.; Ferrara, P.; Gsponer, J.; Caflisch, A. *J. Phys. Chem. B* **2000**, *104*, 10080.
(8) Hong, Q.; Schellman, J. A. *J. Phys. Chem.* **1992**, *96*, 3987.
(9) Schaefer, M.; Bartels, C.; Karplus, M. *J. Mol. Biol.* **1998**, *284*, 835.
(10) Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1991**, *30*, 3864.
(11) Young, W. S.; Brooks, C. L., III *J. Mol. Biol.* **1996**, *259*, 560.
(12) Shoemaker, K. R.; Kim, P. S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Nature* **1987**, *326*, 563.
(13) Osterhout, J. J.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Dyson, H. J.; Wright, P. E. *Biochemistry* **1989**, *28*, 7059.
(14) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Comm.* **1995**, *91*, 43.
(15) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306.
(16) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535.
(17) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(18) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(19) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(20) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952.
(21) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463.
(22) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. *J. Comput. Chem.* **1999**, *20*, 786.
(23) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(24) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.
(25) Brown, J. E.; Klee, W. A. *Biochemistry* **1971**, *10*, 470.
(26) Bierzynski, A.; Kim, P. S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2470.
(27) Munoz, V.; Serrano, L. *Biopolymers* **1997**, *41*, 495.
(28) Gsponer, J.; Caflisch, A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6719.
(29) Mirny, L.; Shakhnovich, E. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 361.
(30) Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005**, *88*, 2472.
(31) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 236.
(32) Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins* **1999**, *34*, 269.
(33) Simmerling, C.; Strockbine, B.; Roitberg, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11258.
(34) van Gunsteren, W. F.; Burgi, P.; Peter, C.; Daura, X. *Angew. Chem., Int. Ed.* **2001**, *40*, 351.

(35) Nymeyer, H.; Garcia, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13934.
(36) Breaux, G. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **2003**, *125*, 10740.
(37) Dasgupta, B.; Pal, L.; Basu, G.; Chakrabarti, P. *Proteins* **2004**, *55*, 305.
(38) Kentsis, A.; Mezei, M.; Gindin, T.; Osman, R. *Proteins* **2004**, *55*, 493.
(39) Mu, Y. G.; Nguyen, P. H.; Stock, G. *Proteins* **2005**, *58*, 45.
(40) Pal, U.; Chakrabarti, P.; Basu, G. *J. Mol. Biol.* **2003**, *326*, 273.
(41) Daggett, V.; Levitt, M. *J. Mol. Biol.* **1992**, *223*, 1121.
(42) Hummer, G.; Garcia, A. E.; Garde, S. *Phys. Rev. Lett.* **2000**, *85*, 2637.
(43) Soman, K. V.; Karimi, A.; Case, D. A. *Biopolymers* **1991**, *31*, 1351.
(44) Feig, M.; MacKerell, A. D.; Brooks, C. L. *J. Phys. Chem. B* **2003**, *107*, 2831.
(45) Shi, Z. S.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190.