A New Type of Helix Pattern in Polyalanine Peptide

5 **AUTHORS**, INCLUDING:

Kuk-Hwan Kim
University of Michigan

255 PUBLICATIONS   **4,472** CITATIONS

A New Type of Helix Pattern in Polyalanine Peptide

Hyeon S. Son, Byung Hee Hong, Chi-Wan Lee,
Sunggoo Yun, and Kwang S. Kim*

*National Creative Research Initiative Center for
Superfunctional Materials, Department of Chemistry
Division of Molecular and Life Sciences
Pohang University of Science and Technology
Hyojadong, Namgu, Pohang 790-784, Korea*

Protein folding is a fundamental problem in life sciences. It is generally known that nonlocal interactions determine the folding conformation to the context of the folding process.[1,2] As the most common regular secondary conformation in proteins, the helix has been an important ingredient of the protein folding problem.[3] In particular alanine-based polypeptides are widely studied to identify the helix-folding process in that the alanine amino acid is known to have one of the highest helix propensities. In principle, intrinsic helix propensities can be obtained from gas-phase measurements where the solvent effect is absent. Hudgins et al. studied alanine-based peptides in vacuo using the high-resolution ion mobility measurement technique.[4] It was reported that introduction of a single lysine at the C terminus (to form Ac−A_n−LysH$^+$) resulted in the formation of very stable, monomeric polyalanine helices.[5] Here, we have investigated (using ab initio calculation and simulaton approach) the helix formation in vacuo in different terminal charge conditions from those of Hudgins et al.[5] Then we have found a new type of helix motif. To the best of our knowledge, this type of helix conformation (to be named λ-helix) has not been characterized before.

It has been reported that ab initio calculations and simulations can produce helical structures with use of alanine-based sequences.[6] Further, a recent computation study showed that proteins with other amino acid sequences can be designed to adopt the desired structures.[7] We have carried out ab initio calculations (Gaussian 98)[8] and molecular dynamics (MD) simulations (CHARMm[9] V26 with the CHARMM22 parameter set[10]) of alanine-based polypeptides. Two terminal conditions are investigated: H_3N^+−RH−CO−[NH−RH−CO]$_{n-2}$−NH−RH−COO$^-$ (NH$_3^+$−A_n−COO$^-$ in short) and H_3CCO−[NH−RH−CO]$_n$−NH$_2$ (CH$_3$−A_n−NH$_2$ in short). While the α-helices are observed in the latter case, the unusually stable λ-helix formation is observed in the former case. This left-handed λ-helix (λ_L-helix or 6_{19}-helix) is shown together with a right-handed α-helix (α_R helix) in Figure
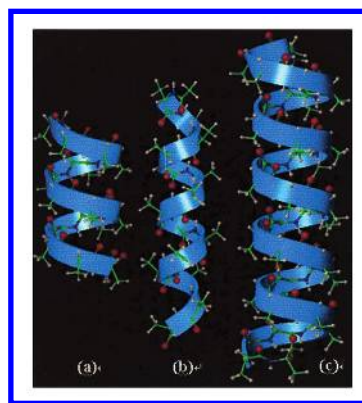
Figure 1. The idealized λ_L–helix pattern for NH$_3^+$−A_{15}−COO$^-$ (a), the idealized α_R-helix pattern for CH$_3$−A_{15}−NH$_2$ (b), and the λ_L–helix pattern from the NH$_3^+$−A_{30}−COO$^-$ MD simulation (c). The idealized λ_L- and α_R-helices have (ϕ, ψ) angles of (−100°, −80°) and (−57°, −47°), respectively.
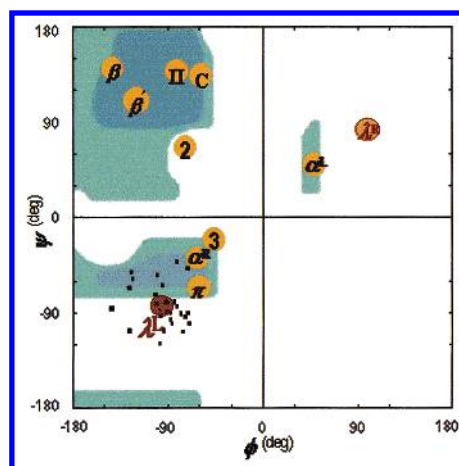

Figure 2. Ramachandran plot for protein secondary structures composed of alanine amino acids. The dots indicate the (ϕ, ψ) angles of the λ_L–helix from the NH$_3^+$−A_{30}−COO$^-$ MD simulation. The circles indicate the standard angles of the secondary structures. (−57°, −47°): right-handed α-helix (α). (−119°, 113°): parallel β-pleated sheet (β). (−139°, 135°): antiparallel β-pleated sheet (β'). (−49°, −26°): right-handed 3_{10}-helix (**3**). (−57°, −70°): right-handed π-helix (π). (−78°, 59°): 2.2_7 ribbon (**2**). (−79°,150°): left-handed polyglycine II and poly-L-proline II helices (**II**). (−51°, 153°): collagen (**C**). (57°, 47°): left-handed α-helix (α$_L$). (−100°, −80°): left-handed λ-helix (λ_L). (100 °, 80 °): right-handed λ-helix (λ_R).

1.[11] A Ramachandran plot to represent various protein secondary structures is given in Figure 2. The λ_L-helix has five residues per turn with H-bonds between NH of residue n and C′=O of residue $n + 4$. A right-handed λ_L-helix is, however, not observed in our simulations due to the close approach of the side chains and the C′O group of L amino acids (as the left-handed α-helix is not observed in the experiments).

Our ab initio and simulation results show that the λ_L-helix is lower in energy than the α_R-helix when both ends are charged (Table 1). The Becke-3 parameters with Lee−Yang−Parr functionals (B3LYP) using the 6-31G* basis set show that the λ_L-helix with charged ends (NH$_3^+$−A_n−COO$^-$) is more stable than the neutral α_R-helix (NH$_2$−A_n−COOH) by ∼2 kcal/mol for $n =$ 8, which implies that the zwitterionic λ_L-helix can exist in the gas phase as a global or at least a local minimum energy structure. The existence of λ_L-helix is ensured from the all-positive

(1) Minor, D. L.; Kim, P. S. Nature **1996,** *380,* 730.
(2) Kim, K. S. *Bull. Korean Chem. Soc.* **1993,** *14,* 18. (b) Kim, K. S. *Chem. Phys. Lett.* **1989,** *156,* 261.
(3) Darby, N. J.; Creighton, T. E. *Protein Structure*; IRL Press at Oxford University Press: Oxford, 1993; pp 9−16.
(4) Hudgins, R. R.; Woenckhaus, J.; Jarrold, M. F. *Int. J. Mass. Spectrom. Ion Processes* **1997,** *165/166,* 497.
(5) Hudgins, R. R.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1998,** *120,* 12974.
(6) (a) Park, C.; Carlson, M. J.; Goddard, W. A., III *J. Phys. Chem. A* **2000,** *104,* 7784. (b) Park, C.; Goddard, W. A., III *J. Phys. Chem. B* **2000,** *104,* 2498. (c) Wu, Y.-D.; Wang, D.-P.; Chan, K. W. K.; Yang, D. *J. Am. Chem. Soc.* **1999,** *121,* 11189. (d) Takano, M.; Yamato, T.; Higo, J.; Suyama, A.; Nagayama, K. *J. Am. Chem. Soc.* **1999,** *121,* 605. (e) Klein, C. Th.; Mayer, B.; Kohler, G.; Wolschann, P. *J. Mol. Struct.* **1996,** *370,* 33. (f) Sung, S.-S. *Biophys. J.* **1995,** *68,* 826
(7) Harbury, P. B.; Plecs, J. J.; Tidor, B.; Alber, T.; Kim, P. S. *Science* **1998,** *282,* 1462.
(8) Frisch, M. J. et al. Gaussian 98, Gaussian, Inc.: Pittsburgh, PA, 1998.
(9) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *Comput. Chem.* **1983,** *4,* 187.
(10) MacKerell, A. D. et al. *J. Phys. Chem. B* **1998,** *102,* 3586.

(11) α$_R$/λ$_L$ means α$_P$/λ$_M$ in IUPAC notation.

Table 1. Relative Stabilities (kcal/mol) of α, λ, and SB Structures in Charged ($NH_3^+-A_n-COO^-$) and Uncharged ($CH_3-A_n-NH_2$) States[a]

Ala$_n$	$n = 8$		$n = 15$		$n = 30$	
	ΔE^c	ΔE^u	ΔE^c	ΔE^u	ΔE^c	ΔE^u
MD (0 K): λ: α	23.1	16.9	157.9	34.6	150.7	42.6
MD 0 K): λ: SB	2.9		17.3			
MD (300 K)	6.3		48.5		83.0	
B3LYP/6-31G*//HF/3-21G	3.4	16.8	0.3	35.8		
B3LYP/6-31G*	1.7	15.2				

[a] ΔE^c is E(charged α_R/SB) $- E$(charged λ_L); ΔE^u is E(uncharged λ_L) $- E$(uncharged α_R). MD simulations (300 K) were carried out for 5 ns, while the last 1 ns was used for sampling. Then, the final structures as well as other well-equilibrated low-energy structures were energy-minimized (0 K). The lowest energy structures were found to be λ_L-helices and SB structures. These relative energies listed in the 2nd row of MD (0 K) are quite different from those listed in the first row of MD (0 K) whose structures were optimized starting from their idealized (ϕ, ψ) angles in Figure 1. Thus, in the first and second rows of MD (0 K), λ_L-helices were compared with α_R-helices and SB structures, respectively. This results in large variations in ΔE^c. However, ΔE^u shows little variation because uncharged α- and λ-helices are little changed from their idealized structures. With the terminal charges, α_R-helices are much less stable than SB structures, while λ_L-helices are quite stable. B3LYP/6-31G*// HF/3-21G values denote the single-point energies (B3LYP/6-31G*) at the HF/3-21G optimized geometries, while the B3LYP/6-31G* energies were obtained at the B3LYP/6-31G* optimized geometries. In these cases, for the charged ends, the energies of λ_L-helices were compared with those of SB structures, since α_R-helices with charged ends were not stable. With charged ends, the energies of SB structures are considered to be the lowest next to those of λ_L-helices (or could be the lowest in the case of long chains). The average H-bond distance and angle of the λ_L-helix for $n = 8$ are 1.93 Å and 59° (B3LYP/6-31G*).

frequencies at the Hartree−Fock (HF/3-21G) level for $n = 8$. For long, charged chains, the coiled salt-bridge (SB) structures can be much lower in energy than α_R-helices, and could be lower than λ_L-helices. When the tendency to form SB is suppressed by the presence of surrounding molecular environments, λ_L-helices would be the lowest energy structures even for long chains. Thus, the λ_L-helix motif could play an important role in protein folding.

The dipole of an α_R-helix (~3.5 D per amino acid residue) generates an electrostatic field along the helix axis, producing an effective positive charge at the amino end and an effective negative charge at the carboxyl end.[6b] This dipole alignment enables the helix motif to be extended up to a long distance in the same direction. However, when both terminals of the helix are positively and negatively charged, respectively, the orientation of helix dipoles is restricted to the same direction of electric field between the end charges. In the case of α_R-helix with charged ends, the directions of the helix dipole and the electric field are opposite, and so it is unstable. In the case of the λ_L-helix, the direction of the helix dipole is the same as the electric field, which stabilizes the structure through the charge−dipole interaction.[12] The orientation of the λ_L-helix is not observed in any other helix conformations. A π-helix shows a similarity in shape, however, the peptide unit dipole direction of the π-helix is opposite to that of the λ-helix.

In many previous MD simulations, the right-handed α-helix patterns are found in the neutral terminal charge conditions of $CH_3-(A)_n-NH_2$.[6d] However, in charged terminal conditions, we find that the λ_L-helix is formed. Starting with a fully extended linear structure, a complete α_R-helix is formed at ~1.8 ns in the $CH_3-A_{15}-NH_2$ simulation and a complete λ_L-helix is formed at

(12) Besides, in the charged λ_L-helix, the electric field is stronger than in the charged α_R-helix because of the shortened distance between two opposite charges (by ~25%), which makes the charge−dipole interaction stronger.
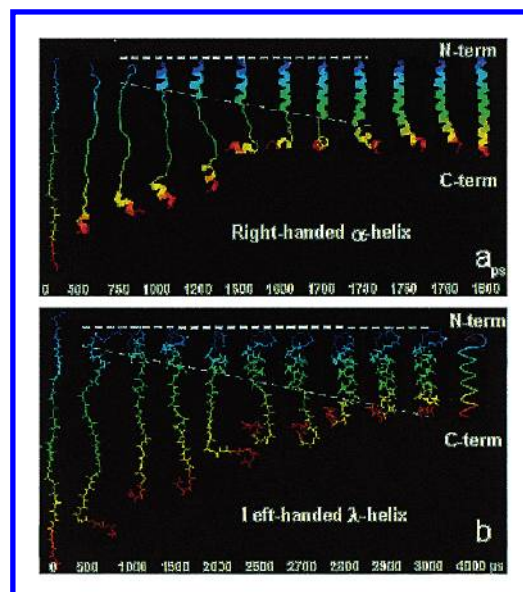
Figure 3. Selected snapshots from MD trajectories. An α_R-helix (a) and a λ_L-helix (b) are produced by the $CH_3-A_{30}-NH_2$ and the $NH_3^+-A_{30}-COO^-$ simulations, respectively.

~3.0 ns in the $NH_3^+-A_{30}-COO^-$ simulation. During the helix-folding process, helix-nucleation first occurs at the terminal sites (in particular, near the N-terminus in the case of the λ_L-helix) and it promotes the propagation of the helix pattern along the segment (Figure 3). In both types of simulations only the terminal composition is different. Therefore, the handedness of the final conformation is related to the terminal charge conditions. The propagation of the helix pattern along the segment clearly shows that sequential local interactions determine the nascent folding patterns of the protein. The initial folding in the λ_L-helix arises from the electrostatic interactions of the positively charged NH_3^+ group with the adjacent carbonyl dipole moiety, followed by the dipole−dipole interactions between two adjacent carbonyl moieties. Our study using the semiempirical approach also indicates the fold-induced folding that is propagated by the initial folding arisen from the charge (NH_3^+)−dipole (CO) interactions, followed by two adjacent carbonyl moiety interactions.[13]

In conclusion, we have found that, in the presence of charged terminal ends, an unusual helical motif, a left-handed λ-helix, is much more stable than a right-handed α-helix. This result could be useful for protein folding study, in particular, in the gas phase and in the nonpolar solvent environments,[13] in consideration of the recent development in gas-phase experiments.[5] The terminal end patching approach (using different charge conditions) could be utilized to control the handedness of protein folding as well as unfolding in nonpolar protein terminal segments.

(13) We have also conducted sets of MD simulations in water. The MD simulations of ~2 ns did not show significant conformational changes in both λ_L- and α-helices. Although a charged α-helix is unstable in the gas phase, it is still intact in water; the water molecules around the terminal ends prevent the effects of excessive charges introduced by the terminal ends. Therefore the terminal charge-induced folding/unfolding mechanism might not be effective in the aqueous solvent environment. However, it could be effective in nonpolar solvent environments, because the property in this case would be more similar to that of the gas phase. In the aqueous simulation, both terminals are almost capped by the charged ion−water interactions, and no water is found inside the λ-helix.