J|A|C|S
COMMUNICATIONS

Published on Web 09/11/2004

# α-Helical Peptides Are Not Protonated at the N-Terminus in the Gas Phase

Robert Wieczorek† and J. J. Dannenberg*

*Department of Chemistry, City University of New York−Hunter College and the Graduate School, 695 Park Avenue, New York, New York 10021*

Received April 5, 2004; E-mail: jdannenberg@gc.cuny.edu

Amino acids and small peptides are zwitterionic under most conditions. Consequently, biochemists often draw peptide and protein structures in their respective zwitterionic structures. We have recently shown that formamide chains show a very substantial preference for protonation on the terminal C═O.[1] Protonation of these chains at the $NH_2$ results in rupture of the H-bond between the N-protonated and the other formamides. This result suggested that peptides that contain H-bonding chains that begin at the N-terminus may not be zwitterionic. However, such a conclusion is not completely established, as the nitrogens protonated in the formamide chains are amido, while those in peptides are amino, which are generally more basic. In solution, the probable effect of solvation must also be considered, as the zwitterionic structures of amino acids and small peptides are known to be better solvated than their "normal" counterparts.

Several groups have reported experimental[2] and molecular orbital studies[3] on α-helical peptides that confirm H-bonding cooperativity. We recently reported structural and energetic details of α-helical peptides containing up to 18 amino acids that clearly indicate that the H-bonding chains in these structures are highly cooperative.[4] They behave similarly to the H-bonds in formamide chains, exhibiting much more extensive cooperativity than would be expected for the pairwise, electrostatic interactions employed by most empirical modeling methods.[5] As the proton affinity at the terminal C═O of these formamide chains increases markedly with increasing chain length, similar behavior for α-helices might be anticipated.

Here, we present ONIOM[6] B3LYP/AM1[7] calculations as programmed in Gaussian 98[8] on fully optimized structures of polyalanines, $(ala)_N$ ($N$ = 8, 14, 17), that are protonated at the N-terminus, the COOH-terminus, and the adjacent C═O groups. We have fully described this procedure elsewhere.[4] The specific helices chosen all have three chains containing equal numbers of H-bonds: two, four, and five for $N$ = 8, 14 and 17, respectively. We could not find optimized α-helical structures smaller than $(ala)_8$, or for $(ala)_{11}$ (which would have three chains of three H-bonds). Protonation of $(ala)_8$ at N destroys its helical structure. We include the energetic data for $ala_8$ simply to help define the trends.

The proton affinities (PAs) (see Table 1) at the COOH and C═O groups exceed those at the Ns in each example. Also, the PAs of the COOH and C═O groups increase substantially with increasing peptide size, while those at the Ns decrease. Thus, the differences in the PAs of the COOH and C═O groups with those of the N-termini rapidly become so large that the prospect of the order being reversed by solvation becomes increasingly unlikely as the peptides grow longer. The comparable PA of monomeric alanine at N, which (when protonated) has a $C_5$ H-bond from the $NH_3$ to the adjacent C═O as in Figure 1, is 228.8 kcal/mol. We could not obtain a stable, H-bonding O-protonated alanine.

† On leave from Faculty of Chemistry, University of Wroclaw, F. Joliot-Curie 14, Wroclaw, Poland.

***Table 1.*** Proton Affinities[a] at the Four Positions Described in the Text

| α-helix | $NH_2$ | COOH | C═O | C═O |
|---|---|---|---|---|
| $Ala_8$ | 243.1 | 259.1 | 259.3 | 258.5 |
| $Ala_{14}$ | 212.6 | 272.3 | 274.0 | 274.7 |
| $Ala_{17}$ | 202.5 | 276.5 | 278.0 | 279.0 |

[a] Values given in kcal/mol.

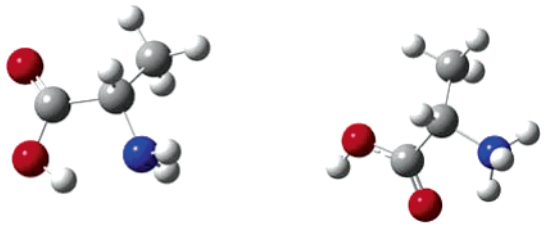


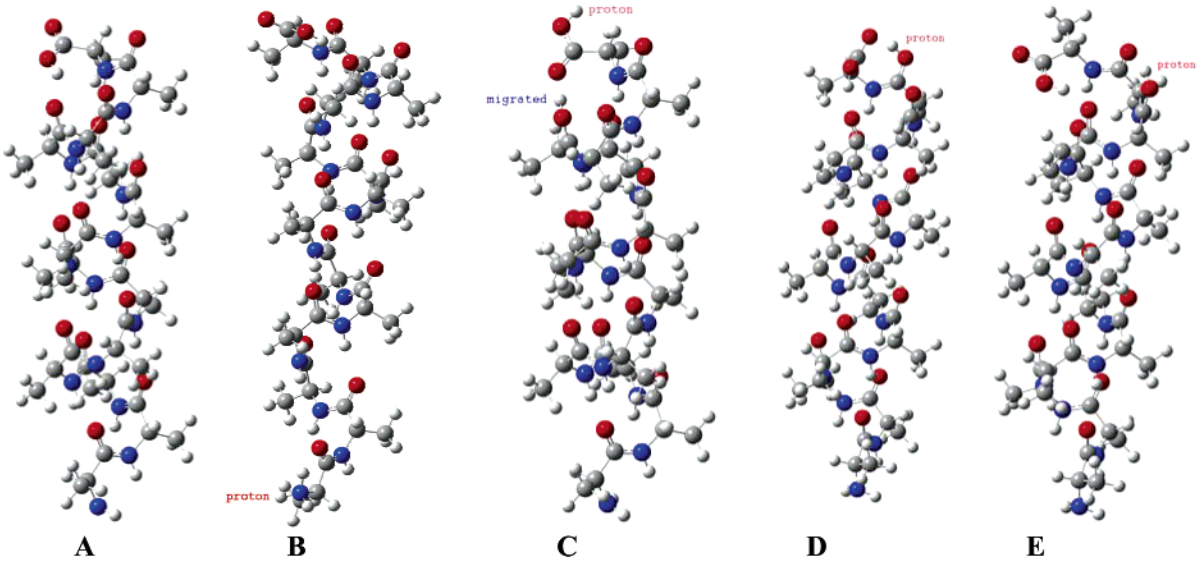
***Figure 1.*** Alanine (left) and N-protonated alanine (right).

We illustrate the geometries of those structures that maintain an α-helix upon protonation in Figure 2 using $(ala)_{14}$ as an example. The neutral structure, **A**, contains three chains of four H-bonds apiece. One has an O−H···O H-bond between the carboxyl OH and a carbonyl, one has an N−H···O H-bond between the C═O of the N-terminal alanine and an N−H; all the others are typical amide N−H···O H-bonds. **B** depicts the peptide protonated at the N-terminus. A $C_5$ H-bond from the C═O to the $-NH_3^+$ (analogous to those found in planar $\beta$-strands) replaces the broken H-bond from the C═O of the N-terminal alanine that was part of a cooperative chain. The terminal helical dihedral NCCN angle ($\psi$) becomes 180° versus −45.5° in **A**.

Protonation of the carboxyl occurs on its C═O, as in **C**. The proton of the COOH migrates to the C═O to which it was H-bonded, analogous to migrations of N−Hs reported for chains of five or more formamides.[1] The protonated carboxyl is more acidic than the protonated amidic C═O to which it migrates. Protonation of the other two proximate C═O groups of **A** (both unsatisfied H-bond acceptors) produce **D** and **E**, each containing a new H-bond to an adjacent C═O (that of the carboxyl for **D**, and that of the next C═O for **E**). As seen from the table, the PAs for these three structures do not differ substantially. In each of these cases, a strong H-bond is formed from the protonated C═O (including the one to which the carboxyl proton was transferred in **C**) to an adjacent C═O. These participate in the H-bonding chains whose conjugated $\pi$-systems traverse the H-bonds.

The PAs at the COOH and C═O groups of the α-helices exceed those reported for H-bonding formamide chains since (1) these lead to formation of the additional H-bonds shown for **D** and **E** and to the migrated OH proton in **C**; and (2) enthalpies (reported for the formamide chains) are smaller than energies (reported here) by vibrational corrections. On the other hand, the PAs at the N-termini of $ala_{14}$ and $ala_{17}$ are less than for alanine, itself, as α-helical


10.1021/ja048033s CCC: $27.50 




H-bonds are broken to the C═O of N-terminal ala in the two N-protonated helices. As the broken H-bonds are stronger for $N = 17$ than 14, the PA at N decreases as N increases. The N-terminus PA of $ala_8$ exceeds that of ala due to H-bonding in the (no longer helical, but globular) peptide (likely one of several local minima). Gas-phase studies of protonated $(gly)_n$ and $(ala)_n$ suggest they are globular for $n$ values up to 20. However, (ala) with a lys at the C-terminus is helical. MD simulations suggest the protonated lys H-bonds to a C═O.[9] **C**, **D**, and **E** resemble the structure of ac-$(ala)_{19}lysH^+$ in ref 9b with $H^+$ instead of $lysH^+$. Other energy minima might exist for $(ala)_n$, which may be globular like that of $(ala)_8$.

Hydration of the α-helices will alter the relative energies calculated for the gas phase. Although we did not explicitly consider solvation in this work, we can approach it in several ways. Explicit water molecules should H-bond most strongly to the protonated sites. For **B**, one expects the $-NH_3^+$ to be surrounded by waters, which would likely further unravel the proximate helical structure. Thus, hydration would destroy the local helical character, as it stabilizes the positive charge. Use of a continuum solvation model would favor the N-protonated form as the dipole moment of this structure increases from that of the neutral helix upon protonation, while those of the others decrease. However, since the methyl groups that form the cylindrical "walls" of the helix are hydrophobic, the water molecules that contact these walls should preferentially assemble to best preserve the liquid water structure rather than counteract the dipole of the protonated helix. Furthermore, α-helical regions of proteins often form "bundles" that are held together by hydrophobic interactions. Thus, they may not be very well hydrated. The C═O groups that do not participate in the helical H-bonds could be readily protonated in these "bundles". Clearly, further detailed investigation of solvation is indicated.

We conclude from these studies that (1) terminal α-helical regions of proteins should not be N-protonated. If the N-termini of such helices become protonated, they lose their local helical character; (2) α-helical bundles might provide facile sites for protonation on C═O groups; and (3) the α-helical structures of artificially synthesized peptides may be protonated near the COOH terminus, especially in those solvent mixtures that have been reported to promote α-helix formation (such as water/trifluoroethanol) rather than in pure water.[2f,10]

**Acknowledgment.** This work was supported by a grant from the donors of the Petroleum Research Foundation, administered by the American Chemical Society.

**Supporting Information Available:** Cartesian coordinates of the relevant structures (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

## References

(1) Moisan, S.; Dannenberg, J. J. *J. Phys. Chem. B* **2003**, *107*, 12842.
(2) (a) Kennedy, R. J.; Tsang, K.-Y.; Kemp, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 934. (b) Smythe, M. L.; Huston, S. E.; Marshall, G. R. *J. Am. Chem. Soc.* **1995**, *117*, 5445. (c) Hanson, P.; Martinez, G.; Millhauser, G.; Formaggio, F.; Crisma, M.; Toniolo, C.; Vita, C. *J. Am. Chem. Soc.* **1996**, *118*, 271. (d) Wolf, W. M.; Stasiak, M.; Leplawy, M. T.; Bianco, A.; Formaggio, F.; Crisma, M.; Toniolo, C. *J. Am. Chem. Soc.* **1998**, *120*, 11558. (e) Karle, I. L. *J. Mol. Struct.* **1999**, *474*, 103. (f) Jaravine, V. A.; Alexandrescu, A. T.; Grzesiek, S. *Protein Sci.* **2001**, *10*, 943. (g) Silva, R. A. G. D.; Yasui, S. C.; Kubelka, J.; Formaggio, F.; Crisma, M.; Toniolo, C.; Keiderling, T. A. *Biopolymers* **2002**, *65*, 229.
(3) (a) Mohle, K.; Hofmann, H.-J. *J. Pept. Res.* **1998**, *51*, 19. (b) Topol, I. A.; Burt, S. K.; Deretey, E.; Tang, T.-H.; Perczel, A.; Rashin, A.; Csizmadia, I. G. *J. Am. Chem. Soc.* **2001**, *123*, 6054−6060. (c) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2000**, *256*, 15. (d) Wu, Y.-D.; Zhao, Y.-L. *J. Am. Chem. Soc.* **2001**, *123*, 5313. (e) Wu, X.; Wang, S. *J. Phys. Chem. B* **2001**, *105*, 2227. (f) Kubelka, J.; Gangani, R. A.; Silva, D.; Keiderling, T. A. *J. Am. Chem. Soc.* **2002**, *124*, 5325. (g) Bour, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2002**, *65*, 45. (h) Ireta, J.; Neugebauer, J.; Scheffler, M.; Rojo, A.; Galvan, M. *J. Phys. Chem. B* **2003**, *107*, 1432.
(4) (a) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc* **2003**, *125*, 8124. (b) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2004**, *126*, in press.
(5) Kobko, N.; Dannenberg, J. J. *J. Phys. Chem. A* **2003**, *107*, 10389.
(6) Dapprich, S.; Komiromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *THEOCHEM* **1999**, 0166.
(7) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(8) 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.; 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. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11; A.11 ed.; Gaussian, Inc.: Pittsburgh, PA, 1998−2001.
(9) (a) Hudgins, R. R.; Mao, Y.; Ratner, M. A.; Jarrold, M. F. *Biophys. J.* **1999**, *76*, 1591. (b) Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1999**, *121*, 3494.
(10) (a) Goodman, M.; Listowsky, I.; Masuda, Y.; Boardman, F. *Biopolymers* **1963**, *1*, 33. (b) Wallimann, P.; Kennedy, R. J.; Miller, J. S.; Shalongo, W.; Kemp, D. S. *J. Am. Chem. Soc.* **2003**, *125*, 1203.

JA048033S