Peptide Models. 17. The Role of the Water Molecule in Peptide Folding. An ab Initio Study on the Right-Handed Helical Conformations of *N*-Formylglycinamide and *N*-Formyl-L-alaninamide Monohydrates [H(CONH−CHR−CONH)H·H_2O; R = H or CH_3]

András Perczel, Ödön Farkas, and Imre G. Csizmadia*

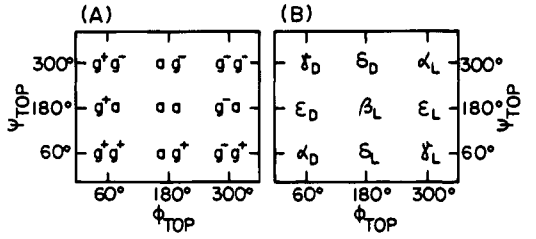
Department of Chemistry, University of Toronto Toronto, Ontario, Canada M5S 1A1
Department of Organic Chemistry, Institute of Chemistry Eötvös University, Budapest, Hungary

Received July 7, 1994

Building up the electron density of larger molecules like peptides by using the electron density of molecular fragments has been the subject of several recent studies.[1,2] From a synthetic point of view, peptides are built up from 20 natural amino acids as a sequential assembly of amino acid residues: −NH−CHR−CO−, but structurally the backbone conformations of peptides may also be regarded as built up from discrete conformations of the appropriate amino acid diamide fragments. The latter diamide model compounds,[3−11] incorporating different types of α-amino acids (all possessing a chiral center with the exception of glycine), have one common conformational feature: the adoption of a maximum number of nine discrete backbone conformations. The existence and the geometry of the 3 × 3 = 9 discrete diamide orientations have been predicted by multidimensional conformation analyses (MDCA).[12] Accordingly, we have found changes along both torsional modes ϕ and ψ leading to three discrete minima (g^+, a, and g^-), which in turn yield a total of nine minimum energy conformations (g^+g^+, ..., g^-g^-), as shown in Chart 1A.

For the sake of simplicity, these conformations are abbreviated (cf. Chart 1B) in terms of subscripted Greek letters (α_L, α_D, β_L, γ_L, δ_L, δ_D, ϵ_L, and ϵ_D).[7,9,10] The validity of the MDCA prediction has also been proven by *ab initio* MO calculations. The pseudo-three-dimensional Ramachandran map,[13] i.e., $E = E(\phi,\psi)$, for H−CONH−CHMe−CONH−H is shown in Figure 1 of the supplementary material. The approximate locations of the nine conformations denoted by subscripted Greek letters on a Ramachandran map, ($E = E[\phi,\psi]$), are shown by solid dots on Figure 2 of the supplementary material. These conformational prototypes could be rather useful in deciphering the main chain or backbone conformation of proteins, since these nine conformational centers are located in the same region of the $E = E[\phi,\psi]$ 2D Ramachandran map, where the "ϕ,ψ plots" of the protein's X-ray structure show a high population (Figure 7 in ref 23b). Although these nine minima are typical for large varieties of amino acids, the building block of the right-handed helix (the α_L minimum) was not found during the thorough *ab initio* investigations[3−10] carried out for an appreciable period of time on model peptides such as For-Gly-NH_2, For-L-Ala-NH_2, and For-L-Val-NH_2. In the case of For-L-Ala-NH_2, even an elaborated grid search[9] failed to reveal the expected α_L minimum in the [$\phi = -60 \pm 30°$, ψ $-30 \pm 30°$] region, suggesting the annihilation of this minimum. Yet, alanine is often described as an α-helix former amino acid residue on the basis of two secondary structure prediction algorithms. In order to resolve this contradiction, three different attempts have been made in our laboratories to investigate the conformational properties of amino acid diamides in the [$\phi = 60 \pm 30°$, $\psi = -30 \pm 30°$] region of the 2D Ramachandran map. Varying both the molecular constitution and the size of the model compounds, efforts were made to stabilize the annihilated α_L conformation through three different types of interactions: (1) backbone/backbone [BB/BB], (2) backbone/side chain [BB/SC], and (3) intermolecular [IM].

Chart 1



(1) The elongations of the For-L-Ala-NH_2 molecule by the insertion of another alanine residue resulted in the For-L-Ala-L-Ala-NH_2 model compound. This made the formation of a 10-member intramolecular hydrogen bond (1 ← 4 H-bond) possible in a special $\phi_1 = -68.6°$, $\psi_1 = -17.5°$, $\phi_2 = -113.1°$, and $\psi_2 = +21.3°$ backbone conformation. In this structure, the first diamide unit adopts a backbone torsional angle combination close to the targeted α_L [$\phi \simeq -60°$, $\psi \simeq -30°$] substructure.[14]

(2) The polar hydroxymethyl group of serine diamide [H−CONH−CH(CH_2OH)−CONH−H], or simply For-L-Ala-NH_2, can also provide a stabilizing effect, resulting in three α_L-type backbone structures associated with three different side chain orientations.[11b]

Points 1 and 2 have already been reported on. Now, in this paper we report some solid evidence for the existence of the third type of molecular stabilization, where the key role is played by intermolecular interaction. The water molecule in *N*-formyl-L-alaninamide monohydrate, i.e., [H(CONH−$CHCH_3$−CONH)H]·H_2O, or simply For-L-Ala-NH_2·H_2O, can stabilize the otherwise annihilated α_L backbone conformation. This observation also underlines the importance of the H_2O molecule in the peptide folding, to be discussed later. Recently, Shang and Head-Gordon[15] proposed *a reaction field model of solvent* to apply for "restabilizing" the annihilated minimum at the ϕ =

(1) (a) Bader, R. F. W. *Acc. Chem. Res.* **1985**, *9*, 18. (b) Bader, R. F. W.; Caroll, M. T.; Chessman, J.; Chang, C. *J. Am. Chem. Soc.* **1987**, *109*, 7068. (c) Bader, R. F. W.; Popelier, P. L. A.; Keith, T. A. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 620.

(2) Walker, P. D.; Mezey, P. G. *J. Am. Chem. Soc.* **1993**, *115*, 12423.

(3) (a) Schafer, L.; Van Alsenoy, C.; Scarsdale, J. N. *J. Chem. Phys.* **1982**, *76*, 1439. (b) Schafer, L.; Klimkowski, V. J.; Momany, F. A.; Chuman, H.; Van Alsenoy, C. *Biopolymers* **1984**, *23*, 2335.

(4) (a) Klimkowski, V. J.; Schafer, L.; Momanay, F. A.; Van Alsenoy, C. *J. Mol. Struct.* **1985**, *124*, 143. (b) Scarsdale, J. N.; Van Alsenoy, C.; Klimkowski, V. J.; Schafer, L.; Momany, F. A. *J. Am. Chem. Soc.* **1983**, *105*, 3438.

(5) Weiner, S. J.; Singh, V. C.; O'Donnel, T. J.; Kollman, P. *J. Am. Chem. Soc.* **1984**, *106*, 6243.

(6) (a) Head-Gordon, T.; Head-Gordon, M.; Frish, M. J.; Brooks, C., II; Pople, J. A. *Int. J. Quantum Chem. Quantum Biol. Symp.* **1989**, *16*, 311−319. (b) Head-Gordon, T.; Head-Gordon, M.; Frish, M. J.; Brooks, C., II; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989.

(7) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J.-L.; Marcoccia, J.-F.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1991**, *113*, 6256.

(8) Bohm, H.-J.; Brode, S. *J. Am. Chem. Soc.* **1991**, *113*, 7129.

(9) McAllister, M. A.; Perczel, A.; Császár, P.; Viviani, W.; Rivail, J.-L.; Csizmadia, I. G. *J. Mol. Struct.* **1993**, *288*, 161.

(10) Viviani, W.; Rivail, J.-L.; Perczel, A.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1993**, *115*, 8321.

(11) (a) Perczel, A.; Daudel, R.; Ángyán, J.; Csizmadia, I. G. *Can. J. Chem.* **1990**, *68*, 1182. (b) Perczel, A.; Farkas, Ö; Marcoccia, J. F.; Csizmadia, I. G. *J. Am. Chem. Soc.*, submitted. (c) Perczel, A.; Farkas, Ö.; Csizmadia, I. G., manuscript in preparation.

(12) (a) General and Theoretical aspects of the thiol group. In *The chemistry of the thiol group*; Csizmadia, I. G., Patai, S., Eds.; The chemistry of functional groups; Wiley and Sons: New York, 1974; pp 1−109. (b) Peterson, M. R.; Csizmadia, I. G. Analytic Equations for Conformational Energy Surfaces. In *Progress of Theoretical Organic Chemistry*; Elsevier: Amsterdam, 1982; Vol. 3, pp 190. (c) Multidimensional Theoretical Stereochemistry and Conformational Potential Energy Surface Topology in *New Theoretical Concept for Understanding Organic Reactions*; Csizmadia, I. G., Bertran, J. D., Eds.; Reidel Publishing Co.: Dordrecht, 1989; pp 1−31.

(13) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95.

(14) (a) Pauling, L.; Corey, R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 729−740. (b) Pauling, L.; Corey, R.; Branson, H. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205−211. (c) Levitt, M.; Chothia, C. *Nature (London)* **1976**, *261*, 552−558.

(15) Shang, H. S.; Head-Gordon, T. *J. Am. Chem. Soc.* **1994**, *116*, 1528.

0002-7863/95/1517-1653$09.00/0

Table 1. Selected Structural Data of the For-Gly-NH_2·HOH and For-L-Ala-NH_2·HOH Supermolecules Optimized Using 3-21G and DZP Basis Sets

	For-Gly-NH_2		For-L-Ala-NH_2	
	3-21 G	DZP[a]	3-21 G	DZP
O1–H10	2.10 Å	2.39 Å	2.09 Å	2.36 Å
O1–O11	2.97 Å	3.23 Å	2.96 Å	3.20 Å
O1–H10–O11	147.8°	147.1°	148.2°	148.6°
O7–H9	2.07 Å	2.34 Å	2.06 Å	2.33 Å
O7–O11	2.89 Å	3.15 Å	2.88	3.15 Å
O7–H9–O11	141.1°	143.3°	141.7°	144.2°
$\omega 1$	−175.55°	−169.73°	−175.25°	−169.84°
ϕ	−70.65°	−74.90°	−69.82°	−74.09°
ψ	−35.16°	−29.70°	−36.83°	−31.73°
$\omega 2$	+179.71°	+172.78°	−178.92°	+173.63°
E(total)[b]	−449.254 011 0	−451.756 210 9	−488.078 445 7	−490.792 461 8
F(max)[c]	4.0×10^{-5}	3.6×10^{-5}	4.0×10^{-5}	3.9×10^{-5}
F(av)[c]	1.3×10^{-5}	1.1×10^{-5}	1.3×10^{-5}	1.2×10^{-5}

[a] The following Huzinaga (DZP) basis set (*J. Chem. Phys.* **1965**, *42*, 1293) has been used for C, N and O, (8s,4p,1d) → [5111,31,1], and H, (4s,1p) → [31,1]. [b] Hartree. [c] In mdyn.

Table 2. Comparison of the [ϕ,ψ] Torsional Angles (deg) for the α_L Conformation of the Various Model Compounds with Backbone/Backbone, Backbone/Side Chain, and Intermolecular Interactions

		model compound							
			BB/SC				IM		
		BB/BB	For-L-Ser-NH_2[a]				For-Gly-NH_2·H_2O		For-L-Ala-NH_2·H_2O	
theoretical method		For(L-Ala)$_2NH_2$ 3-21G	3-21G	4-21G	4-21G*	6-31G*	3-21G	DZP	3-21G	DZP
ab initio	ϕ	−86.6[b]	−62.4	−67.5	−63.0	−69.1	−70.7	−74.9	−69.8	−74.1
	ψ	−17.5[b]	−42.8	−40.1	−45.5	−39.9	−35.2	−29.7	−36.8	−31.7
reaction field	ϕ						−75.0	−84.5[c]	−74.5	−81.3[c]
	ψ						−27.7	−21.3[c]	−28.3	−24.6[c]

[a] Side chain conformation is [*a*,*a*]. [b] In the case of For(L-Ala)$_4NH_2$ in its $(\alpha_L)_3\delta_L$ conformation as computed at HF/3-21G level of theory, the helical portion of the 3_{10} helix had the following three pairs of torsional angles: $\phi = 61.9°$, $\psi = -27.6°$; $\phi = -61.9°$, $\psi = -21.1°$; and $\phi = -72.6$, $\psi = -3.9$. [c] Computed at HF/6-31+G* level of theory.

60 ± 30°, $\psi = -30 \pm 30°$ region of the Ramachandran map.[13] Even though no molecular water had been included to represent the first solvation shell, the Onsager model[16] was sufficient for them to recreate the annihilated minimum energy α_L conformation. Using a set of natural internal coordinates,[17,18] we optimized the For-Gly-NH_2·H_2O and For-L-Ala-NH_2·H_2O supermolecules using both a smaller (3-21G) and a larger (DZP) basis set at the HF level of theory, using the TEXAS 90 program system.[19] It should be emphasized that while the computations were carried out within the HF formalism, the results, due to the modest basis set size, were well above the HF limit. Selected structural data of these supermolecules are reported in Table 1. The atom labeling and space filling models of the peptide–monohydrate complex show two H-bonds (cf. Figure 3 of the supplementary material).

Table 2 compares the ϕ and ψ values of the three model compound types 1, 2, and 3 used to restabilize the annihilated α_L helix. In the footnote of Table 2, the three pairs of torsional angles of the helical positions of the 3_{10} helix, i.e, the $(\alpha_L)_3\delta_L$ conformation of For(L-Ala)$_4NH_2$,[20] are also shown. The corresponding ϕ,ψ values of the 3_{10} helix of For(L-Ala)$_5NH_2$ reported by Schäfer and co-workers[21] are quite similar. Comparing the backbone torsional angle values of *ab initio* results with the adjacent torsional angles obtained from the *reaction field* model (cf. Table 2), a remarkable similarity can be observed. The deviation is always smaller than 10°. These resemblances confirm the relevance of such a *reaction field* model calculation in certain cases, and it could save valuable CPU time, too. However, the *ab initio* calculations of the herein reported supermolecule not only provide an α_L-type backbone orientation but also give precious data on a possible hydration pattern of the diamide molecule, even in a conformation which is intrinsically unstable. The two H-bonds that make possible the adaptation of a stable α_L-type backbone conformation for these diamide systems are the first and unique example of how backbone folding may be influenced by solvent molecules.

Moreover, the type I β-turn is in itself a stable structure for various model triamide systems, such as in the case for the For-Gly-Gly-NH_2, For-L-Ala-Gly-NH_2, For-Gly-L-Ala-NH_2, and For-L-Ala-L-Ala-NH_2 molecules.[22,23] A hairpin structure, which is about one-half of a turn in an α-helix, can serve as a template for the additional amide units, entering the formation of a 3_{10}- or an α-helical segment. Using the X-ray diffraction data of an N- and a C-protected -L-Val-L-Ser- model compound, it had been shown previously how β-turn structures can associate in the solid state, forming a helix, like molecular packing.[24] On the other hand, the importance of a single water molecule in the transition of β-turn α-helices and vice versa has already been demonstrated[25] by analyzing X-ray structures of protein molecules.

Supplementary Material Available: Figures 1–3 (4 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA942190H

(16) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 11486.
(17) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.
(18) Fogarasi, G.; Zhou, X.; Taylor, P. W.; Pulay, P. *J. Am. Chem. Soc.* **1992**, *114*, 8191–8201.
(19) (a) Pulay, P.; et al. *TX90*; Fayetteville, AR, 1990. (b) Pulay, P. *Theor. Chim. Acta* **1979**, *50*, 229.
(20) Perczel, A.; Endredi, G.; McAllister, M. A.; Farkas, Ö.; Csaszar, P.; Ladik, J.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)*, in press.
(21) Schäfer, L.; Newton, S. Q.; Ming, C.; Peeters, A.; VanAlsenoy, C.; Wolinski, K.; Momany, F. A. *J. Am. Chem. Soc.* **1993**, *115*, 272.
(22) Perczel, A.; McAllister, M. A.; Csizmadia, I. G., to be published.
(23) (a) Perczel, A.; McAllister, M. A.; Csaszar, P.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1993**, *115*, 4849. (b) Perczel, A.; McAllister, M. A.; Csaszar, P.; Csizmadia, I. G. *Can. J. Chem.* **1994**, *72*, 2050.
(24) Perczel, A.; Foxman, B. M.; Fasman, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8210.
(25) Sundaralingham, M.; Sukharudu, Y. C. *Science* **1989**, *244*, 1333.