Peptide Models. 18. Hydroxymethyl Side-Chain Induced Backbone Conformational Shifts of L-Serine Amide. All ab Initio Conformers of For-L-Ser-NH_2

András Perczel,*,†,‡ Ödön Farkas,†,§,⊥ and Imre G. Csizmadia†,⊥

Contribution from the Institute of Organic Chemistry, Eötvös University, 112 Budapest P.O.B. 32, H-1117, Hungary, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, England, Department of Chemistry, Wayne State University, Detroit, Michigan 48202, and Department of Chemistry, University of Toronto, Toronto, Ontario M5S 1A1, Canada

Received February 13, 1996. Revised Manuscript Received June 14, 1996[⊗]

Abstract: Using ab initio conformational energy mapping (HF/3-21G) a maximum of nine characteristic backbone conformation clusters (α_L, α_D, β_L, γ_L, γ_D, δ_L, δ_D, ϵ_L, and ϵ_D) were previously established for different amino acid diamides (e.g., For-L-Ala-NH_2, For-L-Val-NH_2, and For-L-Phe-NH_2). Most of the above nine backbone conformers have been located in the [ϕ,ψ] space for various side-chain conformers. The present conformation analysis derives structural parameters of For-L-Ser-NH_2 molecule based on a systematic investigation of the side-chain conformational energy maps $\{E = E(\chi_1,\chi_2)\}$ associated with characteristic backbone structures. The systematic mapping of the $E = E(\phi,\psi,\chi_1,\chi_2)$ four-dimensional Ramachandran-type map has revealed 44 minima. This finding thus established the complete conformational set for For-L-Ser-NH_2. Specific intramolecular hydrogen bonds of the 44 geometry optimized structures were analyzed. These ab initio structures can now be used with greater confidence during force field parameterizations, NMR, and X-ray structure elucidations or even for the characterization of protein backbone structures.

Introduction

To determine the conformational properties of a larger peptide or protein, both diffraction techniques (X-ray or neutron diffraction) and spectroscopical approaches (NMR, CD, IR) require a structural and conformational database of the different building subunits. Thus, peptide and protein chemists established several strategies in the last quarter of the century, to differentiate the conformational units of proteins. Nearly half a century ago Pauling and Corey[1] identified the α-helix and the β-pleated sheet secondary structural elements. Subsequently, additional structural elements such as β-turns[2] or loops[3] became familiar to protein chemists. These elements of secondary structure turned out to be useful in the course of CD, IR, NMR, and X-ray structure investigations. Several computational approaches were developed to predict the above (and additional) subunits in proteins. Some techniques operate on the similarity pattern of characteristic intramolecular hydrogen bonds,[4] while others are based on the classification of the relative spatial arrangement of the successive α carbon atoms.[5] Attention has also been focused on the establishment of different "selection rules" to recognize conformational building units of proteins that form a complete and closed structural set.[6] Needless to say, that the goal to obtain a closed structural set is to eliminate, once and for all, terms like "random structure" or "untypical conformation". For this purpose, the rational subdivision of the Ramachandran-type[7] potential energy surfaces is not a new concept.

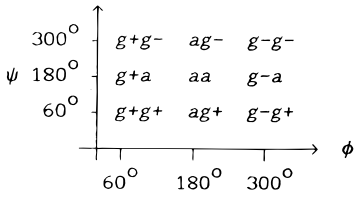
The classical description of the Ramachandran-type $E = E(\phi,\psi)$ surface[7] itself has been improved due to the availability of high performance computers. First, force field computations[8] resulted in valuable structural data on energetically low-laying minima. Recent advances in computational chemistry made it possible to solve the RHF equation even for a molecule as large as an amino acid diamide.[9−19] Moreover, for molecules of this

* Corresponding author. E-mail: perczel@para.chem.elte.hu.
† Eötvös University.
‡ University of Oxford.
§ Wayne State University.
⊥ University of Toronto.
⊗ Abstract published in *Advance ACS Abstracts,* August 1, 1996.

(1) (a) Pauling, L.; Corey, R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 235. (b) Pauling, L.; Corey, R.; Branson, H. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205.
(2) (a) Venkatachalam, C. M. *Biopolymers* **1968**, *6*, 1425.
(3) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1.
(4) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.
(5) (a) Levitt, M. *J. Mol. Biol.* **1976**, *104*, 59. (b) Levitt, M.; Chothia, C. *Nature (London)* **1976**, *261*, 552.
(6) (a) Némethy, G.; Miller, M. H.; Scheraga, H. A. *Macromol.* **1980**, *13*, 914. (b) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95.
(7) It should be noted that in a trans amide bond the ω torsional angle has a single value around 180°. (For a cis peptide bond this value is approximately 0°.) By assuming that both ω_{i-1} and ω_i in the trans amide region can be treated as constants for any typical residue, thus the $E(\omega_{i-1},\phi_i,\psi_i,\omega_i)$ hypersurface may be simplified to $E_{\omega=\text{const}} = E(\phi_i,\psi_i)$.
(8) (a) Vásquez, M.; Scheraga, H. A. *J. Biomol. Struct. Dynamics* **1988**, *5*, 705. (b) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361. (c) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902. (d) Weiner, S. J.; Singh, U. C.; O'Donnell, T. J.; Kollman, P. A. *J. Am. Chem. Soc.* **1984**, *106*, 6243. (e) Brooks, B. B.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187. (f) Seiler, F. J. Res. Lab of the U.S. Air Force Academy, Mopac 5.00, QCPE No. 455, Colorado Springs, CO, 1988.
(9) (a) Sellers, H. L.; Schäfer, L. *J. Am. Chem. Soc.* **1978**, *100*, 7728. (b) Schäfer, L.; Sellers, H. L.; Lovas, F. J.; Suenram, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 6566. (c) Schäfer, R. D.; Van Alsenoy, C.; Scarsdale, J. N. *J. Chem. Phys.* **1982**, *76*, 1439. (d) Klimkowski, V. J.; Schäfer, L.; Momanay, F. A.; Van Alsenoy, C. *J. Mol. Struct. (THEOCHEM)* **1985**, *124*, 143. (e) Scarsdale, J. N.; Van Alsenoy, C.; Klimkowski, V. J.; Schäfer, L.; Momany, F. A. *J. Am. Chem. Soc.* **1983**, *105*, 3438. (f) Schäfer, L.; Klimkowski, V. J.; Momany, F. A.; Chuman, H.; Van Alsenoy, C. *Biopolymers* **1984**, *23*, 2335.
(10) (a) Head Gordon, T.; Head Gordon, M.; Frish, M. J.; Brooks, II, C.; Pople, J. A. *Int. J. Quantum Chem. Quantum Biol. Symp.* **1989**, *16*, 311. (b) Head Gordon, T.; Head Gordon, M.; Frish, M. J.; Brooks, II, C.; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989.

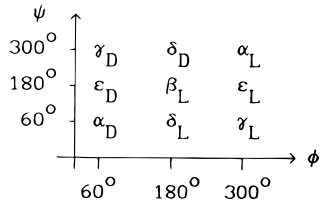
S0002-7863(96)00464-7 CCC: $12.00

Scheme 1



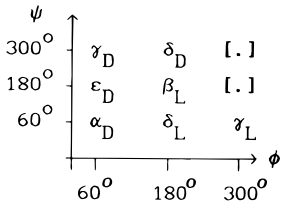
MDCA[11] *predicted 9 legitimate conformers (standard conformational symbolism)*

A



MDCA[11] *predicted 9 legitimate conformers using the conformational symbolism suggested for peptide building units.*

B



Optimized ab initio minima of For-Xxx-NH$_2$ *(Xxx=Gly,Ala,Val,Phe)*

C

size it is now possible to perform rather expensive grid-searches[10,14,20−22] to compute quantum mechanically the actual shape of a Ramachandran-type potential energy surfaces.

Previous ab initio calculations of the $E = E(\phi,\psi)$-type surfaces produced a total of nine basic backbone-conformation clusters.[11−13] The rational for such an observation is due to the fact that three minima ($g+$, a, and $g-$) are expected to occur along both ϕ and ψ variables. In an idealized case, based on *m*ulti*d*imensional *c*onformational *a*nalysis (MDCA) the nine legitimate minima are expected to be located on a 2D-potential energy surface as shown on Scheme 1A. For the time being, there is no uniform IUPAC-IUB[23] recommendation for the labeling of diamide conformers. We have proposed previously[11−13,24] a suitable notation based on these nine conformers (Scheme 1A). This notation uses labels which are familiar to peptide chemists such as γ-turn, β-pleated sheet, α-helix, etc. We are using Greek letters subscripted by L or D to denote the chirality of the conformer, i.e., α_L, α_D, β_L, γ_L, γ_D, δ_L, δ_D, ϵ_L, and ϵ_D (Scheme 1B). These labels are used to discriminate the nine basic conformers that form a closed conformational set like the one shown in Scheme 1A. Some of these diamide conformers are the conformational building units of well-known secondary structural elements. It was possible to preserve practically all previously applied well-known symbols with some additional specification by indexing the conformational codes.[25] Most of these conformational codes are associated with the conformers of well-known secondary structures, such as helices, sheets, etc. Only three among the nine, the δ_L-, δ_D-, and ϵ_D-type backbone structures were not well-known until recently but were found useful to describe

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

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

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

(14) McAllister, M. A.; Perczel, A.; Császár, P.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1993**, *288*, 181.

(15) Liegener, C.-M.; Endredi, G.; McAllister, M. A.; Perczel, A.; Ladik, J.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1993**, *115*, 8275.

(16) Perczel, A.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1993**, *286*, 75.

(17) (a) Perczel, A.; McAllister, M. A.; Csaszar, P.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1993**, *115*, 4849. (b) Perczel, A.; Mcallister, M. A.; Császár, P.; Csizmadia, I. G. *Can. J. Chem.* **1994**, *72*, 2050.

(18) Cheung, M.; McGovern, M. E.; Danchman, T. J.; Jin, T.; Zhao, D. C.; McAllister, M. A.; Farkas, Ö.; Perczel, A.; Császár, P.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1994**, *309*, 151.

(19) Perczel, A.; Daudel, R.; Ángyán, J. G.; Csizmadia, I. G. *Can. J. Chem.* **1990**, *68*, 1182.

(20) Perczel, A.; Farkas, Ö.; Csizmadia, I. G. *J. Comput. Chem.* **1996**, *17*, 000.

(21) (a) Farkas, Ö.; Perczel, A.; Marcoccia, J. F.; Hollósi, M.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1995**, *331*, 27. (b) Perczel, A.; Farkas, Ö.; Marcoccia, J. F.; Csizmadia, I. G. *J. Quant. Chem.* **1996**,

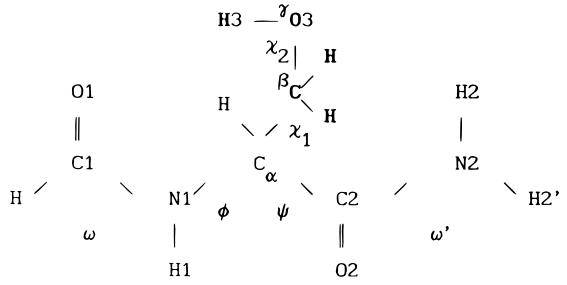
(22) Perczel, A.; Farkas, Ö.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1995**, *117*, 1653.

(23) IUPAC-IUB Commission on Biochemical Nomenclature. *Biochemistry* **1970**, *9*, 3471.

(24) (a) Perczel, A.; Csizmadia, I. G. *Int. Rev. Phys. Chem.* **1995**, *14*, 127. (b) Farkas, Ö.; McAllister, M. A.; Ma, J. H.; Perczel, A.; Hollósi, M.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1996**, in press. (c) Perczel, A.; Farkas, Ö.; Csizmadia, I. G. *Can. J. Chem.* in press.

(25) For example the normal-γ turn (the $[g+g-]$ minimum in Scheme 1/A) and the inverse-γ turn (the $[g-g+]$ minimum in Scheme 1A) are called as γ_D or γ_L, respectively. Also the building unit of the β-pleated sheets, which has an $[aa]$-type diamide conformation (Scheme 1A) is now called β_L. Similarly, for the periodic building unit of poly prolin II secondary structural element ($\phi \approx -60°$, $\psi \approx 140°$), the $[g-a]$-type minimum (Scheme 1A), is called ϵ_L. The conformational building unit of the right-handed ($\approx[g-g-]$ or α_L) or that of the left-handed ($\approx[g+g+]$ or α_D) α-helix are also familiar for peptide and protein chemists.

Scheme 2



protein backbone structures. It is interesting to note that these diamide conformations were first obtained from quantum mechanical calculations of amino acid diamide systems[24] and sequentially applied for describing protein main chain conformation.

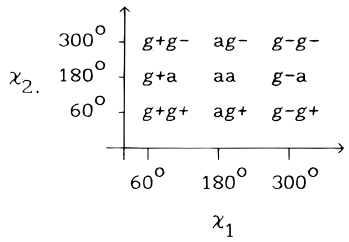
During the last several years, numerous *N*-acetylamino acid−*N*′-methylamides ($CH_3CONH-CHR-CONHCH_3$) and *N*-formylamino acid amides ($HCONH-CHR-CONH_2$) have been the target of *ab initio* calculations.[9−22] However, two incompatibilities were noticed when the results of the *ab initio* calculations were compared with those of general conformational expectations based on force field calculations. During *ab initio* geometry optimization of glycine-,[10−12] alanine-,[10−12] valine-,[12,13] and phenylalanine[24b,c] containing model compounds, the expected ($g-g-$ or αL) minimum, around the [$\phi \approx 300° \pm 30°$, $\psi \approx 300° \pm 30°$] region of the $E = E(\phi,\psi)$ map was not found (Scheme 1C). Such a diamide backbone torsional angle pair corresponds to the right-handed α-helix conformation. Recent ab initio investigations[22] revealed that side-chain backbone interactions operative in For-L-Ser-NH_2 can stabilize this "missing" α-helix-like backbone conformation even in diamides. On the other hand the minimum at [$\phi \approx 300° \pm 30°$, $\psi \approx 180° \pm 30°$], called ϵ_L has not yet been found in any diamide system. The [$g-a$]- or ϵ_L-type backbone orientation was ab initio computed only as a subconformer of a larger peptides (e.g. For-L-Ala-L-Ala-NH_2).

The same analysis that revealed the existence of the α_L-type backbone conformation in For-L-Ser-NH_2 focused our attention to the importance of side-chain potential energy mapping. The serine residue represents the simplest species with a polar side chain $-CH_2OH$ (a hydroxymethyl group), where a total of nine different side chain conformers are expected to be associated with each and every backbone structure. Three (3) minima are presumed to occur for the rotation about the $C^{\alpha}-C^{\beta}$ covalent bond (χ_1 torsional angle according to the IUPAC-IUB convention) to be combined with an additional three (3) minima which is expected to exist along the rotation about the $C^{\beta}-O^{\gamma}$ covalent bond (χ_2 torsional angle) as shown in Scheme 2. One may use again the common *gauche*+ ($g+$), *anti* (a), and *gauche*-($g-$) notation for the typical side-chain conformations along both χ_1 and χ_2. In an idealized case, the locations of the above nine minima are expected at the following positions: 60°,60°; ($g+g+$), 60°,180°; ($g+a$),..., 300°,300°; ($g-g-$). Thus, the nine minima are the result of the three orientations ($g+$, a, and $g-$) as shown in Scheme 3.

Unfavorable interactions between side-chain/side-chain and/or side-chain/backbone atoms can annihilate certain side-chain conformers. Which of the side chain conformers are eliminated for a given backbone conformation may well depend on the backbone conformer type.

Although serine has an average frequency (6.9%) in proteins,[26a] it is an important conformation inducing factor in protein folding. Although, it can rarely be assigned in α-helices and β-strands, serine plays an important role in the formation of

Scheme 3



β-turns.[26b] Analyzing X-ray structures of different globular proteins the result suggests that the most important conformer among the hairpin structures is the type I β-turn, where at position ($i+2$) serine is among the most frequently assigned amino acid residue type.[26b] In spite of the efforts,[27] up to the present no "ideal" type I β-turn model has been synthesized, but among the best candidates several have Ser at their ($i+2$) position (e.g., Piv-Pro-Ser-$NHCH_3$,[27e] Boc-Pro-Ser-$NHCH_3$,[27i] Boc-Val-Ser-$NHCH_3$,[27j] and *cyclo*[(δ)Ava-Gly-Pro-Ser(O^tBu)-Gly][27k,l]). In order to interpret the recorded NMR, CD, and IR spectra of the different model compounds it is essential to determine all the conformers involved in the equilibrium system with their conformational properties (such as conformational weight, torsional angles, etc.). Most frequently the variation of specific torsional angles (ϕ_i, ψ_i, ω_i, χ_i^1, χ_i^2) is of primary interest in practice. Therefore, the ab initio determination of N- and C-protected serine residue is an important step to understand more complicated conformational features of peptides and proteins incorporating serine residue.

The *first aim* of the present analysis was to establish all possible conformers of For-L-Ser-NH_2 by a systematic side-chain mapping. The *second aim* was to compare the location of the ab initio computed conformers on the four dimensional $E = E(\phi,\psi,\chi_1,\chi_2)$ potential energy surface (PES) with the predicted 9*9 = 81 minima. The *third aim* was the analysis of the specific side-chain backbone hydrogen bonds which is presumed to play a major role in side-chain/backbone interaction and could be the basis of backbone conformational shifts.

Computational Methods

Initial backbone conformers were determined by molecular mechanics (ECEPP/2)[8a] and/or semiempirical (AM1/MOPAC)[8f] type calculations. Fully relaxed ab initio structures were obtained subsequentially by optimized [ϕ,ψ] backbone structures (Table 1A) associated with different side-chain orientations [χ_1,χ_2]. Structural parameters of For-L-Ser-NH_2 were computed using natural internal coordinates[28a,b] in full geometry optimization. Minimization was performed at the RHF level of theory by a modified GDIIS[28c] gradient method using a standard 3-21G

(26) (a) McCaldon, P.; Argos, P. *Proteins* **1988**, *4*, 99. (b) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221.

(27) (a) Aubry, A.; Protas, J.; Boussard, G.; Marraud, M. *Acta Crystallogr.* **1980**, *B36*, 2825. (b) Aubry, A.; Lecomte, C.; Boussard, G.; Marraud, M. *J. Chim. Phys. Chim. Biol.* **1983**, *80*, 609. (c) Aubry, A.; Marraud, M. *Acta Crystallogr.* **1985**, *C41*, 65. (d) Aubry, A.; Marraud, M.; Protas, J.; Neel, J. *Comput. Rend. Acad. Sci. Paris* **1974**, *287c*, 163. (e) Aubry, A.; Protas, J.; Marraud, M.; Neel, J. *Acta. Crystallogr.* **1976**, *B32*, 2749. (f) Bruch, M. D.; Noggle, J. H.; Gierasch, L. M. *J. Am. Chem. Soc.* **1985**, *107*, 1400. (g) Kessler, H.; Anders, U.; Schudok, M. *J. Am. Chem. Soc.* **1990**, *112*, 5908. (h) Gierasch, L. M.; Deber, C. M.; Madison, V.; Niu, C.-H.; Blout, E. R. *Biochemistry* **1981**, *20*, 4730. (i) Perczel, A.; Hollósi, M.; Fülöp, V.; Kálmán, A.; Sándor, P.; és Fasman, G. D. *Biopolymers* **1990**, *30*, 763. (j) Perczel, A.; Foxman, B. M.; Fasman, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8210. (k) Hollósi, M.; Kövér, K. E.; Holly, S.; Radics, L.; Fasman, G. D. *Biopolymers* **1987**, *26*, 1555. (l) Perczel, A.; Hollósi, M.; Foxman, B. M.; Fasman, G. D. *J. Am. Chem. Soc.* **1991**, *113*, 9772.

(28) (a) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550. (b) Fogarasi, G.; Zhou, X.; Taylor, P. W.; Pulay, P. *J. Am. Chem. Soc.* **1992**, *114*, 8191. (c) Császár, P.; Pulay, P. *J. Mol. Struct. (THEOCHEM)* **1984**, *114*, 31.

Table 1.

A. Optimized[a] ab Initio (HF/3-21G) Geometries and Relative Energies for For-L-Ser-NH_2 and For-L-Ala-NH_2.

backbone [side chain]	ϕ[b]	ψ	χ_1	χ_2	ΔE[c] (kcal/mol)
		For-L-Ser-NH_2			
$\gamma_L[g+g+]$	−83.6	+71.5	+51.9	+69.8	0.0
$\gamma_L[a\ g+]$	−86.5	+77.8	−169.2	+74.8	12.5
$\gamma_L[a\ g-]$	−83.4	+62.7	+179.8	−68.7	4.8
$\gamma_L[g-g+]$	−85.4	+67.4	−65.3	+55.2	10.5
$\gamma_L[g-a\ ]$	−77.1	+61.4	−44.0	−178.5	7.5
$\gamma_L[g-g-]$	−77.4	+63.4	−41.2	−75.5	7.8
$\beta_L[g+a\ ]$	−170.6	+174.9	+68.0	−172.9	11.2
$\beta_L[g+g-]$	−166.7	+174.8	+67.4	−60.5	9.1
$\beta_L[a\ g+]$	−170.3	−171.5	−171.8	+90.1	3.3
$\beta_L[a\ a\ ]$	−171.8	−173.4	−173.0	+155.1	3.8
$\beta_L[g-g+]$	−179.0	+172.9	−89.2	+55.5	10.5
$\beta_L[g-a\ ]$	−137.3	+160.0	−64.5	+171.5	15.4
$\delta_L[g+a\ ]$	−118.1	+20.2	+51.4	+159.8	14.0
$\delta_L[a\ g-]$	−128.5	+32.9	−172.3	−60.5	8.3
$\delta_L[g-a\ ]$	−129.9	+29.8	−53.1	−167.5	13.3
$\delta_L[g-g-]$	−151.9	+35.6	−46.8	−54.7	11.0
$\alpha_L[g-a\ ]$	−70.5	−24.9	−47.5	−174.5	16.9
$\alpha_L[g-g-]$	−72.0	−23.7	−45.6	−77.5	12.5
$\alpha_L[a\ a\ ]$	−62.4	−42.8	+179.9	−168.7	20.6
$\alpha_D[g+g+]$	+46.4	+53.6	+56.2	+62.1	12.1
$\alpha_D[a\ g+]$	+60.1	+43.8	−156.9	+79.1	20.5
$\alpha_D[a\ g-]$	+62.3	+34.1	−168.1	−65.4	9.1
$\alpha_D[g-a\ ]$	+60.3	+37.6	−58.0	−179.0	12.9
$\delta_D[g+a\ ]$	−163.7	−63.3	+55.1	−169.8	11.2
$\delta_D[g+g-]$	−159.3	−67.6	+52.3	−76.1	10.5
$\delta_D[a\ g+]$	−173.3	−49.4	+163.9	+68.2	15.7
$\delta_D[a\ a]$	−172.4	−55.1	+168.6	+165.9	17.2
$\delta_D[g-g+]$	+146.3	−33.9	−75.2	+74.4	12.3
$\delta_D[g-g-]$	−157.9	−51.8	−47.5	−32.2	15.7
$\epsilon_D[g+a\ ]$	+43.0	−105.5	+92.8	+171.6	18.5
$\epsilon_D[g+g-]$	+99.8	−116.9	+76.3	−68.7	4.9
$\epsilon_D[a\ a\ ]$	+68.4	−172.4	−162.5	−179.3	9.4
$\epsilon_D[a\ g+]$	+66.9	−169.0	−166.4	+82.4	10.1
$\epsilon_D[g-g+]$	+64.5	+177.9	−63.5	+67.4	16.3
$\epsilon_D[g-g-]$	+68.9	+178.2	−58.1	−67.3	20.5
$\gamma_D[g+g+]$	+62.9	−40.3	+41.7	+48.7	14.0
$\delta_D[g+a]$	+51.9	−28.7	+65.8	+173.3	17.1
$\gamma_D[g+g-]$	+78.0	−45.2	+81.9	−62.2	9.4
$\gamma_D[a\ g+]$	+71.3	−52.2	+170.8	+49.0	12.0
$\gamma_D[a\ a\ ]$	+74.0	−65.0	−177.5	−158.0	12.7
$\gamma_D[a\ g-]$	+67.5	−31.2	−163.8	−39.9	10.6
$\gamma_D[g-g+]$	+72.2	−57.5	−61.1	+79.1	12.5
$\gamma_D[g-a\ ]$	+75.4	−56.1	−58.6	+176.0	12.0
$\gamma_D[g-g-]$	+74.7	−55.2	−57.6	−78.2	12.9
		For-L-Ala-NH_2[d]			
α_D	+63.8	+32.7	+60.6		6.0
β_L	−168.3	+170.5	+50.9		1.3
γ_L	−84.5	+67.3	+64.2		0.0
γ_D	+74.0	−57.4	+60.4		2.5
δ_L	−128.1	+29.8	+58.5		3.8
δ_D	−178.6	−44.1	+58.6		7.3
ϵ_D	+67.2	−171.9	+65.6		8.2

B. Comparison of the ϕ,ψ Torsional Angles (deg), Total Energies, E [hartree], and Relative Energies, ΔE [kcal/mol], Associated with the β_L and γ_L Conformations of For-L-Ala-NH_2 Computed at HF/3-21G and MP2/6-311++(d,p) Levels of Theory

	β_L			γ_L			
method	ϕ	ψ	E	ϕ	ψ	E	ΔE
RHF/3-21G	−168.3	+170.6	−412.472 783	−84.5	+67.3	−412.474 780	1.25
MP2/6-311++(d,p)	−157.1	+163.2	−416.382 776	−82.8	+80.6	−416.384 696	1.21

[a] Remaining gradients $<1E$-6 in au. [b] Torsional angles in degrees according to the IUPAC-IUB convention. [c] ΔE in kcal/mol relative to the $E(\gamma_L[g+g+]) = -486.919\,567$ hartree for For-L-Ser-NH_2 and relative to the $E(\gamma_L) = -412.474\,780$ hartree for For-L-Ala-NH_2. [d] Data taken from reference.

basis set.[29] Ab initio energy values were calculated by the TX90 program.[30] Each of the different side-chain geometries with common backbone conformation (e.g., extended-like [β_L] backbone structures) resulted in characteristic [ϕ,ψ] values. At such [ϕ,ψ] points of the Ramachandran potential energy surface, side-chain maps ($f[\chi_1,\chi_2]$) were computed using 30° increments along both χ_1 and χ_2. Thus, side-chain maps ($f[\chi_1,\chi_2]$) associated

(29) Baker, J. *J. Comput. Chem.* **1986**, *7*, 385.

(30) (a) Pulay, P. and co-workers, TX90; Fayetteville, AR, 1990. (b) Pulay, P. *Theor. Chim. Acta*. **1979**, *50*, 229. (c) Frank, J. Selier Research Laboratory, U.S. Air Force Academy, Mopac 5.00 QCPE No. 455, Colorado Springs, CO, 1988. (d) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

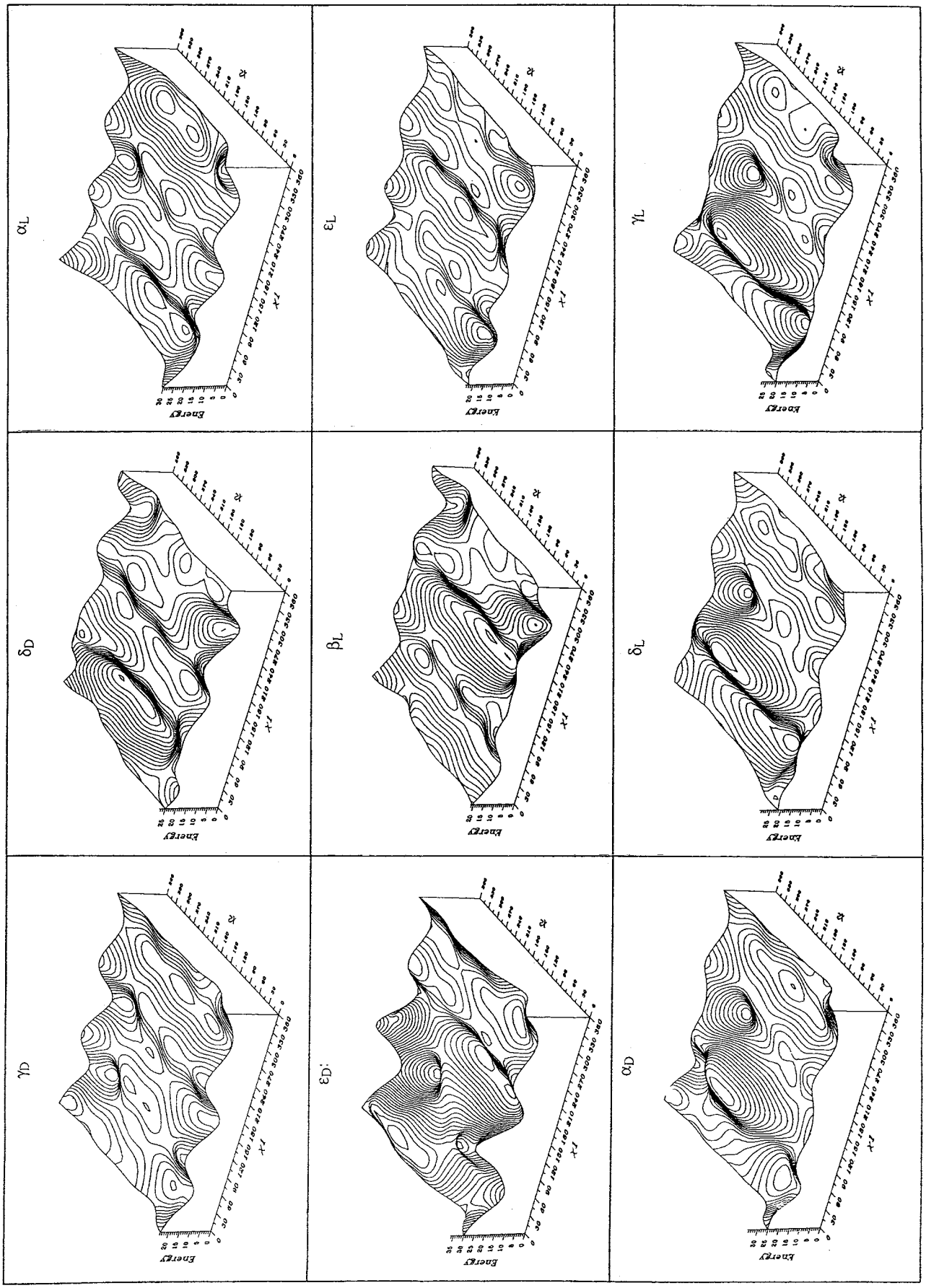


Figure 1. Side chain-maps ($f[\chi_1,\chi_2]$) associated with the α_L, α_D, β_L, γ_L, γ_D, δ_L, δ_D, ϵ_L, and ϵ_D backbone conformations. (Each and every $f[\chi_1,\chi_2]$ map consisted of 144 grid points.)

Table 2. Ab Initio Backbone Conformation and Relative Energy Shifts of For-L-Ser-NH_2 with Respect to For-L-Ala-NH_2[a]

For-L-Ser-NH_2 bb[sc.] conformers	conf shift[b]		rel conf shift[c]		rel energy shift[d]	
	$\Delta\phi$	$\Delta\psi$	$\Delta\Delta\phi$	$\Delta\Delta\psi$	$\Delta\Delta r$[e]	$\Delta\Delta E$
$\gamma_L[g+g+]$	0.9	4.2	0.0	0.0	0.0	0.0
$\gamma_L[a\ g+]$	−2.0	10.5	−2.9	6.3	6.9	12.5
$\gamma_L[a\ g-]$	1.1	−4.6	0.2	−8.8	8.8	4.8
$\gamma_L[g-g+]$	−0.9	0.1	−1.8	−4.1	4.5	10.5
$\gamma_L[g-a\ ]$	7.4	−5.9	6.5	−10.1	12.0	7.5
$\gamma_L[g-g-]$	7.1	−3.9	6.2	−8.1	10.2	7.8
$\beta_L[g+a\ ]$	−2.3	4.4	−3.2	0.2	3.2	9.9
$\beta_L[g+g-]$	1.6	4.3	0.7	0.1	0.7	7.8
$\beta_L[a\ g+]$	−2.0	18.5	−2.9	14.3	14.9	2.0
$\beta_L[a\ a\ ]$	−3.5	16.1	−4.3	11.9	12.7	2.5
$\beta_L[g-g+]$	−10.7	2.4	−11.6	−1.8	11.7	7.2
$\beta_L[g-a\ ]$[f]	31.0	−10.5	30.1	−14.7	33.5	14.1
$\delta_L[g+a\ ]$	10.0	−9.6	9.1	−13.8	16.5	10.2
$\delta_L[a\ g-]$	−0.4	3.1	−1.3	−1.1	1.7	4.5
$\delta_L[g-a\ ]$	−1.8	0.0	−2.7	−4.2	5.0	9.5
$\delta_L[g-g-]$[f]	−23.8	5.8	−24.7	1.6	24.8	7.2
$\alpha_L[g-a\ ]$						
$\alpha_L[g-g-]$						
$\alpha_L[a\ a\ ]$						
$\alpha_D[g+g+]$[f]	−17.4	20.9	−18.3	16.7	24.8	6.1
$\alpha_D[a\ g+]$	−3.7	11.1	−4.6	6.9	8.3	14.5
$\alpha_D[a\ g-]$	−1.5	1.4	−2.4	−2.8	3.7	3.1
$\alpha_D[g-a\ ]$	−3.5	4.9	−4.4	0.7	4.5	6.9
$\delta_D[g+a\ ]$[f]	14.9	−19.2	14.0	−23.4	27.3	3.9
$\delta_D[g+g-]$[f]	19.3	−23.5	18.4	−27.7	33.3	3.2
$\delta_D[a\ g+]$	5.3	−5.3	4.4	−9.5	10.5	8.4
$\delta_D[a\ a\ ]$	6.2	−11.0	5.3	−15.2	16.1	9.9
$\delta_D[g-g+]$[f]	−35.1	10.2	−36.0	6.0	36.5	5.0
$\delta_D[g-g-]$[f]	20.7	−7.7	−19.8	−11.9	23.1	8.4
$\epsilon_D[g+a\ ]$[f]	−24.2	66.4	−25.1	62.2	67.1	10.3
$\epsilon_D[g+g-]$[f]	32.6	55.0	31.7	50.8	59.9	3.3
$\epsilon_D[a\ a\ ]$	1.2	−1.0	0.3	−5.2	5.2	1.2
$\epsilon_D[a\ g+]$	−1.7	2.9	−2.6	−1.3	2.1	1.9
$\epsilon_D[g-g+]$	−2.7	−10.2	−3.6	−14.6	15.0	8.1
$\epsilon_D[g-g-]$	1.7	−9.9	1.8	−14.1	14.2	12.3
$\gamma_D[g+g+]$	−11.1	17.1	−12.0	12.9	17.6	11.5
$\gamma_D[g+a\ ]$[f]	−22.1	28.7	−23.0	24.5	33.6	14.6
$\gamma_D[g+g-]$	4.0	12.2	3.1	8.0	8.6	6.9
$\gamma_D[a\ g+]$	−2.7	3.2	−3.6	−1.0	3.7	9.5
$\gamma_D[a\ a\ ]$	0.0	−7.6	−0.9	−11.8	11.8	10.2
$\gamma_D[a\ g-]$[f]	−6.5	26.2	−7.4	22.0	23.2	8.1
$\gamma_D[g-g+]$	−1.8	−0.1	−2.7	−4.3	5.1	10.0
$\gamma_D[g-a\ ]$	1.4	1.3	0.5	−2.9	2.9	9.5
$\gamma_D[g-g-]$	0.7	2.2	−0.2	−2.0	2.0	10.4

[a] Values of this table are based on the data presented in Table 1. [b] Torsional angle shifts (ϕ,ψ) of For-L-Ser-NH_2 are relative to those of For-L-Ala-NH_2. [c] Relative to the $\Delta\phi$ and $\Delta\psi$ values of the $\gamma_L[g+g+]$ conformer of For-L-Ser-NH_2. [d] $\Delta\Delta E = \Delta E\{\text{For-L-Ser-NH}_2\} - \Delta E\{\text{For-L-Ala-NH}_2\}$. Values are in kcal/mol. [e] $\Delta\Delta r = \{(\Delta\Delta\phi)^2 + (\Delta\Delta\psi)^2\}^{1/2}$; the Pythagorean distances of the relative conformational shifts. [f] "ghost" conformers with significantly shifted ϕ and ψ values.

with the α_L, α_D, β_L, γ_L, γ_D, δ_L, δ_D, ϵ_L, and ϵ_D backbone conformations were determined, where each and every $f[\chi_1,\chi_2]$ map consisted of 144 grid points (Figure 1). Since these maps were associated with a particularly fixed backbone conformation, they are called fixed side-chain maps ($f_{bb=fixed}[\chi_1,\chi_2]$), even though $[(3n-6)-4]$ internal coordinates were fully relaxed.

The usual concern is the role of basis set size and the omission or inclusion of the correlation energy in determining the accuracy of the computed results. The β_L and γ_L conformations of the For-L-Ala-NH_2 have been reoptimized[17b] at the MP2/6311++G[d,p] level of theory, and the results are summarized in Table 1B. Since, along the torsional mode of motion the conformational regions are separated from each other by several tens of degrees (on the idealized surface the separation is 120°) the variations in torsional angles are not considered to be very dramatic on going from the HF/3-21G to the MP2/6-311++(d,p) levels of theory. In addition, the relative energies are practically the same at the two levels of theory.

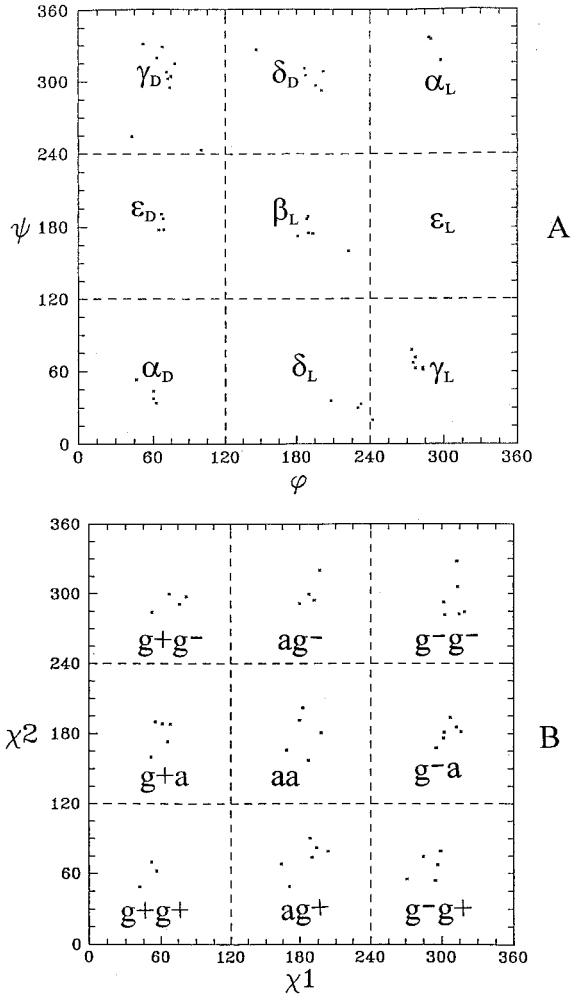


Figure 2. The distribution of the 44 molecular structures of For-L-Ser-NH_2 on a $f(\phi,\psi)$ (A) or on a $f(\chi_1,\chi_2)$ (B) type map.

Results and Discussion

Serine can be derived from alanine by replacing a β proton in the later molecule with a hydroxy group (cf. Scheme 2). Such a modification has significant impact on the overall conformational feature of the molecule. First of all, the simpler 2D Ramachandran-type potential energy surface, $E = E(\phi,\psi)$, relevant for the For-L-Ala-NH_2 conformations should be replaced by a 4D potential energy hypersurface of the following form: $E = E(\phi,\psi,\chi_1,\chi_2)$. Furthermore, the polar −OH functional group can be a proton acceptor as well as a proton donor in hydrogen bonds. The formation of intramolecular hydrogen bonds give a "chance" to the molecule to stabilize specific backbone conformers. One important consequence of this has been discussed already, namely that the stability of the conformational building unit of the right-handed α-helix (α_L) is at least partially due to the formation of specific side-chain backbone interactions.[20,21b]

Our expectation seems to be justified by the present ab initio results since a maximum of nine different side-chain conformers ($[g+g+]$, $[g+a]$,...,$[g-g-]$) are possible involving the χ_1 and χ_2 torsional angles (Scheme 3). As reported in Table 1 and shown in Figure 2B the "normal" γ-turn ($\phi \approx 60°$, $\psi \approx 300°$) denoted as γ_D-type backbone orientation makes the adoption of all nine different side-chain structures possible. If this were

Scheme 4

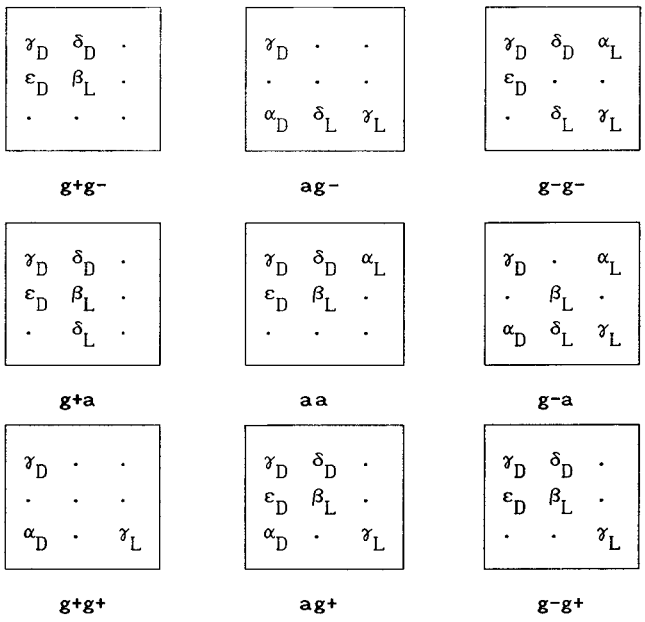


Table 3. Ab Initio Backbone Conformations of For-L-Ser-NH_2 and For-L-Ala-NH_2

backbone conformation type	For-L-Ser-NH_2[a]		For-L-Ala-NH_2		av backbone conformational shift[b]	
	ϕ	ψ	ϕ	ψ	$\Delta\phi$	$\Delta\psi$
γ_L	−82.2	67.4	−84.5	67.3	2.3	−0.1
β_L	−166.0	176.3	−168.3	170.5	2.3	−5.8
δ_L	−132.1	30.0	−128.1	29.8	−4.0	0.2
α_L	−68.3	−30.5				
α_D	57.3	42.3	63.8	32.7	−6.5	9.6
δ_D	−173.4	−53.5	−178.6	−44.1	5.2	−9.4
ϵ_D	68.6	−154.7	67.2	−171.9	1.4	17.2
γ_D	69.8	−47.9	74.0	−57.4	−4.2	9.5

[a] Averaged for all stable side-chain conformations. [b] The conformational shift calculated between the appropriate backbone conformers of For-L-Ser-NH_2 and For-L-Ala-NH_2 molecules.

Table 4. Ab Initio Backbone Conformations of For-L-Ser-NH_2 and For-L-Ala-NH_2

H-bond[a]	type[b] atom	acceptor[c] atom	donor atom	dist and angle ranges			obsd in
				X... Y[d]	X... HY[d]	X−H−Y[e]	
A	bb/bb	O1	N2	1.95 ± 0.25	1.95 ± 0.25	145 ± 10	all γ_L and γ_D
B	sc/bb	O1	O3	2.10 ± 0.30	2.80 ± 0.15	130 ± 20	$\beta_L(g-g+)$, $\delta_L(g-g-)$ $\gamma_D(g+g-)$, $\gamma_D(g+g+)$ $\delta_D(g-g-)$, $\delta_D(g-g+)$ $\epsilon_D(g+g-)$, $\epsilon_D(g+a)$
C	bb/bb	O2	N1	2.20 ± 0.15	2.60 ± 0.05	105 ± 6	all β_L
D	sc/bb	O2	O3	2.20 ± 0.10	2.65 ± 0.10	107 ± 3	$\gamma_L(g+g+)$, $\gamma_L(ag-)$, $\delta_L(ag-)$, $\gamma_D(ag-)$, $\alpha_D(ag-)$
E	sc/bb	O3	N1	1.90 ± 0.15	2.65 ± 0.10	137 ± 10	$\gamma_L(g-g-)$, $\gamma_L(g-a)$, $\delta_L(g+a)$, $\alpha_L(g-g-)$, $\alpha_L(g-a)$
F	sc/bb	O3	N2	1.95 ± 0.25	1.95 ± 0.25	145 ± 10	$\beta_L(aa)$, $\delta_D(g+g-)$ $\delta_D(g+a)$, $\epsilon_D(g+g-)$

[a] For typical structure see Figure 3A−F. [b] The hydrogen bond can be observed between backbone/backbone (bb/bb) or side-chain/backbone (sc/bb) atoms of For-L-Ser-NH_2. [c] For the numbering of the atoms see Scheme 2. [d] Distances in Å. [e] Angles in deg.

the case with each and every backbone structures (γ_L, γ_D, β_L, etc.) of For-L-Ser-NH_2, then 9*9 = 81 distinguishable molecular conformers could be computed. In reality, due to unfavorable interactions only 44 out of the above 81 possible molecular conformers exist at the HF/3-21G level of theory (Scheme 4 and Figure 2A,B). Since the ϵ_L-type structure is still not among the stable backbone conformers, the 44 geometries are shared among the remaining eight backbone structures. A total of nine γ_D-, six γ_L-, β_L-, ϵ_D-, δ_D-, four α_D-, δ_L-, and three α_L-type For-L-Ser-NH_2 backbone conformers were determined with various side-chain conformers. Computed conformational and energetic parameters of the stable structures are reported in Table 1. The relative stability of these structures should be considered with some caution due to the inherent limitations of the applied level of theory (HF/3-21G). The 3-21G basis set is regarded as a small but relevant approach to build up the HF wave function of the molecule since our previous finding for For-L-Ala-NH_2 revealed semiqualitative similarities between the results of HF/3-21G and MP2/6-311++G(d,p) computations (c.f. Table 2 of ref 17b). Furthermore, it should be mentioned that for For-L-Ser-NH_2 the global minimum is the $\gamma_L[g+g+]$ and that all L-type structures had relatively low energy contents (<4.0 kcal/mol) compared to $\gamma_L[g+g+]$. This is in good agreement with general expectations and with previous ab initio computa-

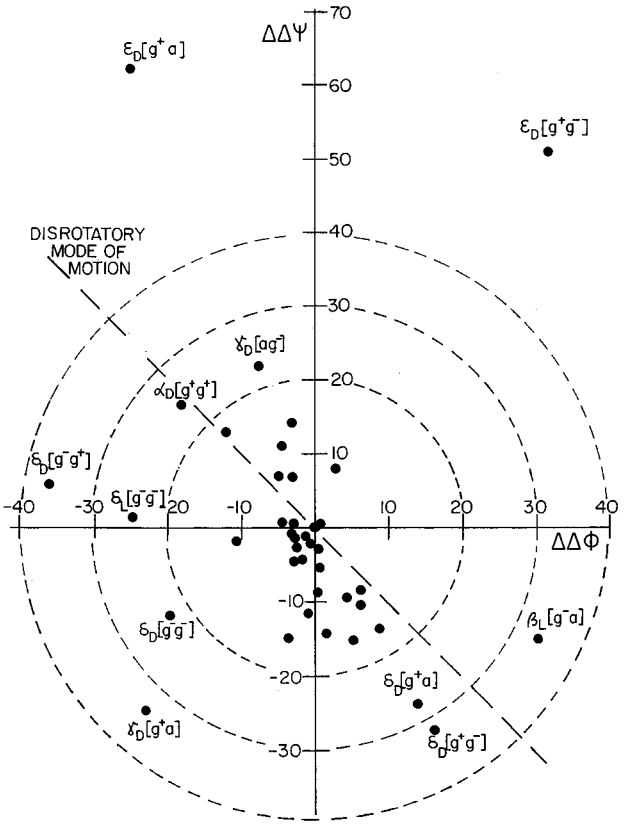


Figure 3. Distribution of computed serine diamide conformations on a double difference ($\Delta\Delta\phi$ and $\Delta\Delta\psi$) plot.

tions carried out on N- and C-protected Gly,[10−12] Ala,[10−12] Val,[12,13] and Phe.[24c]

Clustering the 44 conformers in the 4D space ($f = f[\phi,\psi,\chi_1,\chi_2]$) only 11 conformers have significantly shifted $[\phi,\psi]$ values. Conformations that have ϕ and ψ values which are significantly different from the "norm" are sometimes referred to as "ghost" conformers. The above "ghost" conformers are marked by *f* in Table 2. By averaging the ϕ,ψ torsional angles, given in Table 1, of qualitatively similar structures of For-L-Ser-NH_2 [e.g., γL(sc), βL(sc)] we have obtained characteristic backbone conformational "locations" on the Ramachandran surface (Table 3 and Figure 2A). Comparing the deviations of the averages presented in Table 3, with the individual deviations, shown in Table 2, it is clear that some of the differences have averaged out. Nevertheless, the data presented in Table 3 are quite useful. Comparing the averaged backbone structures of For-L-Ser-NH_2 with similar properties of For-L-Ala-NH_2 a clear resemblance was obtained (Table 3). Not considering the ψ value of the ϵD conformer, influenced by two ϵD "ghosts" the structural shift induced by the modification of the side chain (Ala → Ser) is always smaller than 10°. These similarities support our previous assumption[24a] that the existence and approximate location of the nine backbone structures is a rather intrinsic feature of diamide systems, regardless of the side-chain type of the involved amino acid residue. (Similar resemblance was also observed when the characteristic $[\phi,\psi]$-values of the phenyl substituted alanine diamide, For-L-Phe-NH_2, were compared with that of the For-L-Ala-NH_2.[24c])

Although, the conformational properties of diamide systems are generally preserved in For-L-Ser-NH_2, the possibility of specific side-chain backbone interactions should be mentioned and discussed in details. The lone-pairs of all three oxygen atoms (O1, O2 of the amide groups and O3 of the side chain) have a proton-acceptor character (Scheme 2). Both the amide protons (NHs) and the "acidic" hydrogen atom of the side-chain oxygen are potential candidates for the formation of intramolecular hydrogen bonds with the above acceptor atoms. A maximum of six different hydrogen-bond types (A−F) can be presumed (Figure 4 and Table 4). Two among the six H-bond types, A and C, are formed between backbone/backbone atoms, and, therefore, they cannot be amino acid specific. Of these two (A and C) hydrogen bonded structures the A-type H-bond is characteristic for both forms of γ-turns (γL and γD), while the C-type is associated with the extended-like (βL) backbone structure, which has a five-membered ring hydrogen bond. In all four remaining hydrogen bonded structures (B, D, E, and F) a side-chain atom (O or H) is interacting with a backbone atom (H or O), respectively. These side-chain backbone type hydrogen bonds are expected to be operative not only in serine but also in threonine and cysteine containing diamides. Since in B- and D-type interactions (Figure 4B,D) the OH proton of the side chain is donated to one of the amide oxygen atoms, not less than three torsional angles $\{\phi,\chi_1,\chi_2$ (in B) or ψ,χ_1,χ_2 (in D)$\}$ should have the appropriate orientation. By contrast, in E- and F-type intramolecular H-bond (Figure 4E,F) the side-

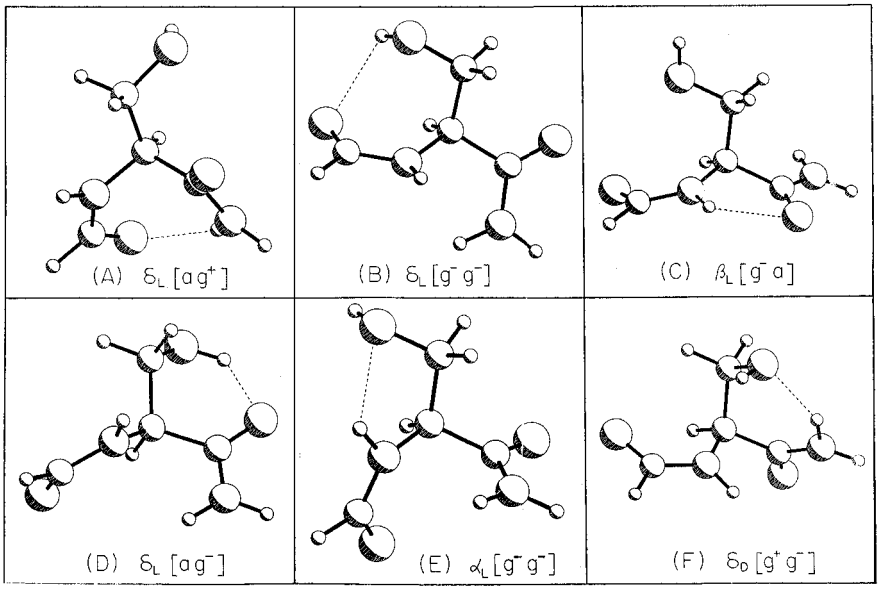


Figure 4. The six possible hydrogen bond-types (A−F) between side-chain backbone or backbone backbone atoms in For-L-Ser-NH_2. Only one example is shown for each type even though there are several conformers listed in Table 4.

chain oxygen is the proton acceptor, thus only the relative orientation of two torsional angles count. These are ϕ,χ_1 in E- and ψ,χ_1 in F-type side-chain backbone interactions. As shown in Table 4 all the latter four side-chain/backbone interactions (B, D, E, and F) could be associated with at least three (or more) different backbone orientations in For-L-Ser-NH_2. Therefore, all these intramolecular interactions should be considered as "not backbone-conformation specific" interactions. Yet all these conformers (A−F) may well occur in protein structures.

Conclusion

A systematic side-chain conformational energy mapping has been performed to determine all possible molecular conformations of For-L-Ser-NH_2. It was shown that the $f = f(\chi_1,\chi_2)$ side-chain potential energy surface may have nine minima as predicted MDCA. In the case of the γD-type backbone orientation all the possible nine side-chain minima exist. However, side-chain surfaces associated with other backbone conformation types have less than nine side-chain conformers.

The only backbone conformation type that has not been found in For-L-Ser-NH_2 is the ϵL. Both the MDCA predicted locations and the $f = f_{\epsilon L}^{fixed}(\chi_1,\chi_2)$ full geometry optimization failed to provide a single ϵL-type diamide orientation. Thus, the existence of the conformational building unit of poly proline II (ϵL) in diamide systems remains an open question.

On the other hand, three different αL-type backbone conformations were obtained in the case of For-L-Ser-NH_2 (αL[$g-g-$], αL[$g-a$], and α_L[a_1a]).

Particular attention was focused on the four different side-chain backbone type intramolecular hydrogen bonds (Figure 3B,D−F). All these hydrogen bonds were previously assigned in X-ray determined structures of peptides. Ab initio results confirmed the existence of these interactions in vacuum and demonstrated that for several (more than three) backbone conformations the same hydrogen bond pattern can be present.

Acknowledgment. This research was supported by grants from the Hungarian Scientific Research Foundation (OTKA FO13799, FO17188, and TO17192). Thanks are also due to FEFA (CFA/III-254) for providing computational facilities. The continued financial support of the National Sciences and Engineering Research Council of Canada is gratefully acknowledged.

JA960464Q