Toward Direct Determination of Conformations of Protein Building Units from Multidimensional NMR Experiments. V. NMR Chemical Shielding Analysis of N-formyl-serinamide, a Model for Polar Side-Chain Containing Peptides

ANDRÁS PERCZEL,[1] ANNA K. FÜZÉRY,[2] ATTILA G. CSÁSZÁR[3]
[1]*Department of Organic Chemistry; Eötvös University, Budapest 112, P.O. Box 32, H-1518, Hungary*
[2]*Department of Chemistry, University of Toronto, Toronto, ON Canada M5S 3H6*
[3]*Department of Theoretical Chemistry; Eötvös University, Budapest 112, P.O. Box 32, H-1518, Hungary*

Received 20 December 2002; Accepted 19 February 2003

Abstract: Knowledge of chemical shift–structure relationships could greatly facilitate the NMR chemical shift assignment and structure refinement processes that occur during peptide/protein structure determination via NMR spectroscopy. To determine whether such correlations exist for polar side chain containing amino acid residues the serine dipeptide model, For-L-Ser-NH_2, was studied. Using the GIAO-RHF/6-31+G(d) and GIAO-RHF/TZ2P levels of theory the NMR chemical shifts of all hydrogen ($^1H^N$, $^1H^\alpha$, $^1H^{\beta1}$, $^1H^{\beta2}$), carbon ($^{13}C^\alpha$, $^{13}C^\beta$, $^{13}C'$) and nitrogen (^{15}N) atoms have been computed for all 44 stable conformers of For-L-Ser-NH_2. An attempt was made to establish correlation between chemical shift of each nucleus and the major conformational variables (ω_0, ϕ, ψ, ω_1, χ^1 and χ^2). At both levels of theory a linear correlation can be observed between $^1H^\alpha/\phi$, $^{13}C^\alpha/\phi$, and $^{13}C^\alpha/\psi$. These results indicate that the backbone and side-chain structures of For-L-Ser-NH_2 have a strong influence on its chemical shifts.

 J Comput Chem 24: 1157–1171, 2003

Key words: NMR chemical shift; conformation; serine; *ab initio* calculation

Introduction

In the past 2 decades nuclear magnetic resonance (NMR) spectroscopy has become a powerful and widely used technique for the determination of the structures of peptides and proteins.[1,2] More than 2200 peptide and protein structures have been solved by NMR spectroscopy, and this number is growing rapidly year after year.

The most commonly used structure solution protocol employs[1,2] distance information provided primarily by nuclear Overhauser enhancements (NOEs), because these are related to spatially close pairs of protons, and can be converted into proton–proton distances. The complex relationship between magnetization transfer and the NOE hinders,[3] however, interpretation of the distance information contained in NOEs. This problem is less serious for cases where many NOEs can be collected; however, in other situations it may lead to structures of rather poor quality. In the latter cases use of additional experimental structural information is highly desirable because it allows refinement of NOE-based structures.

Parameters that have been exploited in the past for additional structural information include *J*-couplings,[4,5] residual dipolar couplings,[6,7] and hydrogen exchange rates.[8–10] As many of these parameters are rather complicated to measure,[3] there is a need to find more readily accessable conformationally sensitive parameters. Starting from the 1960s and 1970s, several research projects[11–14] focused on understanding the effects of conformation on the chemical shifts of amino acid residues found in polypeptides. Lack of sufficient chemical shift data hindered the early

Correspondence to: A. Perczel; e-mail: perczel@para.chem.elte.hu

Contract/grant sponsor: Hungarian Scientific Research Fund; contract/grant numbers: OTKA T033074, T030841, and T032486

Contract/grant sponsor: Hungarian Academy of Sciences and Hungarian Ministry of Education (a Domus Hungarica scholarship to A.K.F.)

studies and, as a result, investigations focused on other conformationally sensitive observables. As the amount of chemical shift information on peptides and proteins has grown, interest in the structural information content of chemical shifts has been revived.[14] Numerous empirical studies[15–17] in the early 1990s have ascertained a strong correlation between certain peptide backbone conformations and $^{13}C^{\alpha}$ chemical shifts, and there have been hints[12,13,18] at significant correlations between peptide backbone conformation and $^{1}H^{\alpha}$ as well as $^{1}H^{N}$ and $^{13}C^{\beta}$ chemical shifts. Nevertheless, experimental methods did not seem to provide enough information for detailed chemical shift–structure correlation studies. For example, in the case of side-chain torsion angles significant deviation can occur between the values in solution, where the chemical shifts are measured, and in the crystalline form, where the torsion angles are determined, leading to occasional poor correlation of the two data sets.[19] Furthermore, experimental data for certain structure types is rare, thereby making accurate correlation studies difficult if not impossible.

Over the last decade reasonably accurate, *ab initio* computation of chemical shifts of peptides and proteins[19–35] has become a valuable tool to supplement experimental data for chemical shift–structure correlation studies. Geometry optimizations of small fragments provide a relatively small number of low-energy minima on the related potential energy surface that closely resemble experimentally observed structures.[36,37] *Ab initio* NMR shielding calculations employing such optimized fragments seem to result in reliable chemical shifts. Furthermore, because *ab initio* NMR calculations can be carried out using any reference structure, chemical shift–structure correlations of all kinds can be investigated. Most importantly, *ab initio* calculations allow the structural parameter of interest to be varied in a highly controlled manner, providing more insight than traditional experimental methods.

One of the most important developments in the past decade has been the introduction of the Z-surface technique[24] by the Oldfield group. In a number of thorough investigations (see, e.g., refs. 20, 24, 26, 30, 31, and 35) for selected amino acids they computed, mostly at the GIAO-RHF (Gauge-Including Atomic Orbital-Restricted Hartree–Fock) level, a number of full isotropic chemical shift surfaces as a function of the backbone dihedral angles ϕ and ψ (see Fig. 1). Partial surfaces became also available for the 20 commonly occurring amino acids.[20] In a related series of studies Perczel and Császár[32–34,38] investigated all possible conformational regions of diamide models using geometry-optimized structures and thus an energy criterion for the selection of the structures considered. In all these investigations attempts were made to compare computed results to empirical ones extracted from databases. Perczel and Császár also performed a detailed statistical analysis searching for structure–structure, structure–chemical shift, and chemical shift–chemical shift correlations. Nevertheless, due to the importance of the problem further studies on related and larger models are needed to deepen our understanding of relationships between chemical shifts and diverse structural parameters, especially those of backbone and side-chain dihedral angles.

In this study we have been continuing our investigation of the influence of different side chains and backbone conformations on the NMR properties of nuclei forming the backbone of proteins. We report results on a serine dipeptide, *trans*-For-L-Ser-NH_2, which is the smallest suitable representative of polar, uncharged amino acids. This complements our systematic study of suitable representatives of various amino acid types, whereby results have been reported on glycine,[32] alanine,[32] valine,[33] and phenylalanine[34] dipeptides, For-L-Xxx-NH_2 (where Xxx is one of the previously mentioned amino acids), and on beta-turn models.[38]

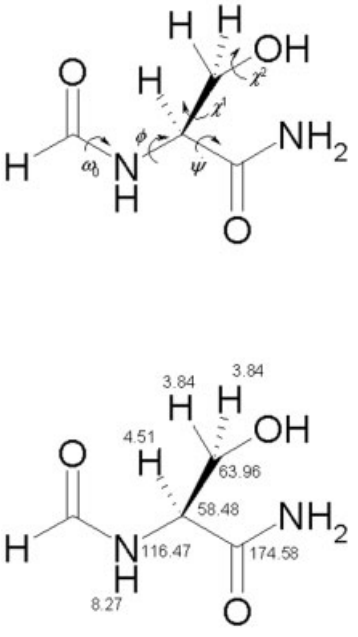

Figure 1. Torsional angles and selected average chemical shifts, taken from the BioMagnetic Resonance Bank (BMRB) and presented next to the respective nuclei, of the For-L-Ser-NH_2 dipeptide model.

Computational Details

Our model system, For-L-Ser-NH_2, is shown in Figure 1. Computation of NMR chemical shielding tensors was performed at the GIAO-RHF (Gauge-Including Atomic Orbital-Hartree–Fock)[39,40] level employing the 6-31+G(d)[41] and TZ2P[42] basis sets. Optimized geometries used for these computations have been determined previously[43] at the RHF/3-21G and RHF/6–311++G(d,p) levels of theory. Four sets of chemical shielding calculations were generated in this study, as detailed in Table 1. The Gaussian98[44] program package was utilized for all electronic structure calculations.

For relative chemical shifts (δ scale) the isotropic chemical shielding values of ^{1}H, ^{13}C, and ^{15}N were referenced to the ^{1}H and ^{13}C values of tetramethylsilane (TMS) and to the ^{15}N value of NH_3. The reference geometry of NH_3 was computed[45] at the CCSD(T)/aug-cc-pVTZ[46,47] level of theory, while that of TMS was obtained at the RHF/3-21G level of theory.

Chemical shift values of Ser associated with a random conformation, the so-called random coil values, were taken from the BioMagnetic Resonance Bank (BMRB)[48] (Fig. 1 and Table 5, *vide infra*). Chemical shift rereferencing, when needed, was done in the group of Dr. Wishart and was generously provided for us for the

Table 1. Designation of the Levels of Theory Employed in This Work for For-L-Ser-NH_2 and the Resulting Number of Structures.[a]

Level	Sublevel	NMR computation	Geometry optimization	No. of structures
A	A1	GIAO-RHF/6-31+G(d)	RHF/3-21G	44 (8)
	A2	GIAO-RHF/6-31+G(d)	RHF/6-311++G(d,p)	36
B	B1	GIAO-RHF/TZ2P	RHF/3-21G	44 (8)
	B2	GIAO-RHF/TZ2P	RHF/6-311++G(d,p)	36

[a]At the RHF/3-21G level 44 minima were found but NMR chemical shifts were only calculated for the eight structures unique to this level of theory.

following 18 proteins: 1BPI, 3LZM, 1LZ1, 2RN2, 2RNT, 1SNC, 1HCB, 1UBQ, 1CEX, 1GZI, 5P21, 1ROP, 1ICM, 193L, 1IGD, 3RN3, 2TRX, and 1A2P. The three-dimensional structures of these proteins were retrieved from the Protein Data Bank (PDB).[49] Data from these two sources were aligned as best as possible. C^{α} chemical shifts were not available for proteins 1LZ1, 2RNT, 1GZI, 1A2P, and 193L. By removing all questionable entries the dihedral parameters and chemical shifts of a total of 79 serine residues were extracted and used as an experimental database on which the theoretically predicted chemical shifts could be tested.

Linear correlation analyses were carried out using Microsoft's Excel program. The quality of each correlation was judged by examining R, the Pearson correlation coefficient.[50] For convenience, Pearson correlation coefficients of chemical shift-structure correlations were designated as R_{cs}. To roughly measure the degree of influence of conformation on chemical shift the average of all R_{cs} values, $\langle R_{cs} \rangle$, was also calculated. Because in most peptides and proteins the ω_0 and ω_1 torsion angles of serine residues (see Fig. 1) take on approximately the same value (close to $\pm 180°$, representing *trans* peptide bonds) correlation studies involving these two torsion angles were of less interest to us. For this reason, these correlations were left out when calculating the average Pearson correlation coefficient, $\langle R_{cs} \rangle$.

For periodic conformational parameters, such as dihedral angles, the definition used for the periodicity strongly influences the outcome of a correlation study.[29,30] To find the optimal definition of the periodicity, that is to determine the maximized R value (R_{max}) for each correlation, each dihedral angle, ϕ, ψ, χ^1, and χ^2, was shifted systematically by increments of 5° (through a total of 360°) and R_{max} was determined. Correlations involving ω_0 and ω_1 were left out for the same reasons described above.

Results and Discussion

Structure and Energetics

According to multidimensional conformational analysis (MDCA),[37,51] a dihedral angle involving rotation about an sp^3—sp^3 C—C bond should have three stable, minimum-energy conformations. In a simplified picture combination of two such dihedral angles produces $3 \times 3 = 9$ minima, three dihedral angles produce $3 \times 3 \times 3 = 27$ minima, and so on. For example, for the simplest α-amino acids, glycine, and alanine, MDCA predicts nine minimum energy conformers as a result of the two backbone dihedral angles, ϕ and ψ. We have previously established a labeling scheme[52] for these conformers: α_L, β_L, δ_L, ε_L, γ_L, α_D, δ_D, ε_D, and γ_D (see Table 2 and Fig. 2).

For more complicated *trans*-amino acid residues, i.e., those possessing a side chain, the number of theoretically predicted minima increases substantially. Valine, for example, should in theory have 27 stable conformers, while serine, having four variable dihedral angles ϕ, ψ, χ^1, and χ^2, should have 81. However, because MDCA in this form provides an oversimplified approach, certain structural factors, such as perturbations caused by the amino acid side chain, are not considered. This results in deviations between the number of idealized MDCA structures and those obtained by geometry optimization at a given level of electronic structure theory.

Out of the 81 stable structures predicted by MDCA for the serine dipeptide, *trans*-For-L-Ser-NH_2, 44 and 36 were found[36] at the RHF/3-21G and the RHF/6-311++G(d,p) levels of theory, respectively. Representatives of each backbone type, except ε_L, have been found. At the higher level the global minimum, $\gamma_L(g+,g+)$, had (ω_0, ϕ, ψ, χ^1, χ^2) = (−176.8°, −85.4°, 72.2°, 54.4°, 69.8°). Similarly to the other For-L-Xxx-NH_2 amino acid models, examined so far by *ab initio* geometry optimizations,[53–58] all α_L and β_L conformers have relatively high energies with the

Table 2. Two Representations of the Nine Minimum Energy Conformers Predicted by MDCA for the Simplest α-Amino Acid, Glycine.[a]

Present labeling scheme[52]	Alternative labeling scheme	Idealized (ϕ, ψ)
α_L	α_R	(−60, −60)
β_L	C5	(180, 180)
δ_L	β_2	(180, 60)
ε_L	β	(180, −60)
γ_L	$C7_{eq}$	(−60, 60)
α_D	α_L	(60, 60)
δ_D	α'	(−60, −60)
ε_D	α_D	(180, −60)
γ_D	$C7_{ax}$	(60, −60)

[a]Dihedral angles ϕ and ψ (see Fig. 1) are given in degrees.

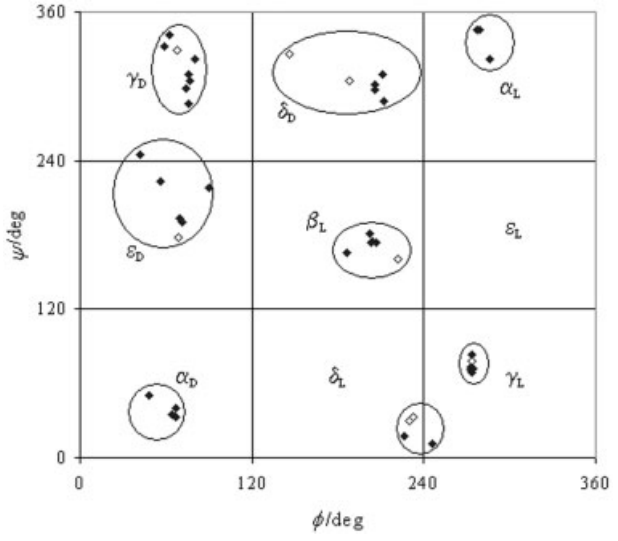

Figure 2. Location of the 44 structures of *trans*-For-L-Ser-NH_2 backbone conformers. Our previously established labeling scheme (α_L, β_L, δ_L, ε_L, γ_L, α_D, δ_D, ε_D, and γ_D) was used. Black diamonds (a total of 36) represent minima found at the RHF/6-311++G(d,p) level of theory, white diamonds (a total of eight) represent minima found only at the RHF/3-21G level of theory.

most stable of these conformers, $\beta_L(a,g+)$, being 2.1 kcal mol^{-1} higher in energy than the global minimum. Selected backbone conformational parameters and relative energies of the 36 minima found at the higher level of theory and the eight minima found only at the lower level of theory can be found in Tables 3 and 4, respectively. Figure 2 illustrates the location of backbone ϕ and ψ angles for the 44 minima. For a more detailed tabulation of results see ref. 36.

Experimentally obtained conformational parameters (Table 8, *vide infra*) reveal that serine residues are most commonly found either in the α_L (α-helical) or β_L (β-pleated sheet) conformations and that other backbone types occur far less frequently. Nevertheless, these other backbone types do occur and sometimes form a significant fraction of the total data (e.g., 15 out of the 79 serine residues in our database take on an ε_L conformation). An extensive search (results not presented) of the Protein Data Bank[49] reveals, that even the eight minima found exclusively at the RHF/3-21G level of theory have natural abundance, albeit at a much smaller percentage then the other calculated conformers. Hence, although the relative energies of the 44 *ab initio trans*-For-L-Ser-NH_2 structures might not be an accurate representation of what is seen in nature, the diversity of backbone types found among the calculated minima certainly is, and thus can be used with confidence during the search for chemical shift–structure correlations.

Chemical Shifts

Relative isotropic chemical shielding values for the 44 unique minima (36 optimized at the higher and 8 at the lower level of theory) can be found in Tables 3 and 4. Chemical shift isotropies (CSI) averaged over individual backbone conformers, using arithmetical averaging, are reported in Table 5A, while CSIs averaged over all backbone conformers, using arithmetical and Boltzmann averaging, are reported in Tables 5B and 5C, respectively. Tables 5B and 5C also contain the experimental average shifts and their standard deviations, obtained from BMRB.

The error of computed values averaged arithmetically over all backbone conformers ranges between 1 and 49% for level A and 1 and 45% for level B. The errors for Boltzmann averaged computed values over the same set ranges from 2% to 38% for level A, and from 0% to 33% for level B. The only large errors are associated with $^1H^N$ and $^1H^\alpha$. Because the large discrepancies are not connected to the computational level employed for the CSI calculations,[38] they must be associated with limitations due to the model chosen (e.g., for $^1H^N$ hydrogen bonding to solvent or to other amino acid residues is neglected). For both types of averaging, level B errors in absolute chemical shift values are generally lower than level A errors, which is in agreement with the relative quality of the basis sets used.

Because it is preferable to have the errors in *ab initio* chemical shifts as small as possible, they were rereferenced so that their averages coincide with the experimental ones. As the arithmetical averaging gave lower errors, these averages were used during the re-referencing process. The resulting chemical shifts were used in the chemical shift–structure correlations (see below).

Tables 6 and 7 contain the Pearson correlation coefficients, R, arising from correlations between selected conformational and isotropic chemical shielding values of all 44 For-L-Ser-NH_2 structures, determined at levels A and B, respectively. Using the 0° to 360° definition of torsional periodicity the average R_{cs} value, $\langle R_{cs}\rangle$, is -0.014 at level B with a standard deviation, σ_{cs}, of 0.268. The Pearson correlation coefficients of the $^1H^\alpha/\phi$, $^{13}C^\alpha/\phi$, and $^{13}C^\alpha/\psi$ correlations fall outside the range determined by $\langle R_{cs}\rangle \pm 2\sigma_{cs}$, i.e., outside the range that 95% of the values fall into, and thus can be considered significant. At level A, the $\langle R_{cs}\rangle$ value is -0.009 with a σ_{cs} of 0.237. Once more the $^1H^\alpha/\phi$, $^{13}C^\alpha/\phi$, and $^{13}C^\alpha/\psi$ correlations all have Pearson coefficients that fall outside the range determined by $\langle R_{cs}\rangle \pm 2\sigma_{cs}$.

Selected R_{max} values can also be found in Tables 6 and 7. Those of the $^1H^\alpha/\phi$ and $^{13}C^\alpha/\phi$ correlations are slightly higher, 0.806 and -0.717 at level A and 0.834 and -0.727 at level B, than their corresponding R_{cs} while that of the $^{13}C^\alpha/\psi$ correlation does not change at either level.

Conformational and CSI parameters for a set of 79 serines, extracted from 18 proteins, are tabulated in Table 8. The results of a correlation analysis involving the experimentally assigned chemical shifts and the major periodic conformational variables, ϕ, ψ, and χ^1, are listed in Table 9. The average R value was 0.066 with a standard deviation (σ) of 0.364.

Chemical Shift–Structure Correlations

The Pearson correlation coefficients arising from crosscorrelations between selected conformational and isotropic chemical shielding values of 44 different *trans*-For-L-Ser-NH_2 structures, determined at levels A and B, are listed in Tables 6 and 7, respectively. Because the results of the two levels are very similar, only those obtained at the higher level of theory (level B) will be discussed here (Table 7).

Table 3. Backbone Conformational Parameters, Relative Energies, and Relative Isotropic Chemical Shielding Values (GIAO-RHF/TZ2P and GIAO-RHF/6-31+G(d), with the Latter Values in Parentheses) of 36 For-L-Ser-NH_2 Conformers Optimized at the RHF/6-311++G(d,p) Level of Theory.

Conf.[a]	ϕ[b]	ψ[b]	ΔE[c]	^{15}N (c)	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	$^{1}H^{\beta 1}$	$^{1}H^{\beta 2}$
$\beta_L(g-,g+)$	−174.0	165.5	3.86	−4.14 (−4.45)	−8.18 (−7.49)	−8.01 (−7.83)	3.52 (3.69)	−0.43 (−0.26)	−0.04 (−0.12)	−0.30 (−0.24)	0.50 (0.43)
$\beta_L(g+,g-)$	−153.3	174.0	4.51	8.82 (8.55)	−3.29 (−2.87)	−0.91 (−1.10)	3.23 (3.87)	−0.30 (−0.07)	−0.13 (−0.16)	−0.69 (−0.64)	0.73 (0.64)
$\beta_L(g+,a)$	−156.8	173.3	5.78	4.62 (4.37)	−1.66 (−1.29)	−0.85 (−0.68)	4.03 (4.68)	−0.42 (−0.28)	−0.23 (−0.27)	−0.82 (−0.65)	0.59 (0.67)
$\beta_L(a,a)$	−156.0	174.9	2.38	7.36 (6.70)	−0.56 (−0.01)	−0.49 (−0.40)	1.67 (1.97)	−0.15 (0.02)	−0.41 (−0.53)	0.04 (0.12)	−0.10 (−0.06)
$\beta_L(a,g+)$	−157.3	−179.1	2.10	6.50 (5.78)	−2.42 (−1.69)	−0.67 (−0.77)	1.71 (2.05)	−0.15 (0.01)	−0.06 (−0.16)	0.05 (0.01)	0.02 (0.03)
$\delta_L(g-,g-)$	−134.1	17.1	6.36	0.28 (0.31)	−2.30 (−1.62)	−1.47 (−0.81)	2.80 (2.89)	1.93 (1.83)	−0.86 (−0.82)	−1.09 (−0.89)	0.97 (0.75)
$\delta_L(g+,a)$	−114.4	11.8	4.33	5.01 (5.05)	−0.93 (−0.97)	−0.40 (−0.36)	2.43 (2.89)	0.51 (0.49)	−0.64 (−0.62)	−0.11 (−0.02)	−0.48 (−0.37)
$\gamma_L(g-,g-)$	−87.0	73.1	4.89	−3.99 (−3.49)	−2.03 (−1.67)	0.90 (1.31)	3.75 (3.35)	0.99 (0.93)	−0.07 (−0.05)	−0.52 (−0.36)	0.26 (0.08)
$\gamma_L(g-,a)$	−87.00	71.1	4.53	−2.91 (−2.58)	−1.18 (−0.70)	1.22 (1.66)	3.39 (2.96)	1.10 (1.03)	−0.38 (−0.46)	−0.51 (−0.33)	0.2 (0.19)
$\gamma_L(g-,g+)$	−86.7	82.8	5.71	−2.33 (−1.65)	−1.97 (−1.71)	0.37 (0.25)	3.05 (2.63)	1.55 (1.57)	−0.44 (−0.56)	−0.38 (−0.38)	0.41 (0.35)
$\gamma_L(a,g-)$	−86.4	69.2	2.78	−0.57 (−0.13)	−0.08 (0.39)	0.87 (1.06)	0.69 (0.77)	1.62 (1.65)	−0.47 (−0.58)	−0.35 (−0.19)	0.67 (0.51)
$\gamma_L(g+,g+)$	−85.4	72.2	0.00	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
$\alpha_L(a,a)$	−73.9	−38.0	13.06	3.26 (2.79)	−4.64 (−3.75)	−1.32 (−0.67)	3.81 (4.29)	2.05 (2.09)	−0.35 (−0.48)	−0.26 (−0.09)	0.64 (0.65)
$\alpha_L(g-,a)$	−80.8	−14.2	6.41	2.84 (2.43)	−4.89 (−4.51)	0.38 (0.96)	2.37 (2.48)	0.95 (0.83)	−0.37 (−0.44)	−0.80 (−0.59)	0.66 (0.60)
$\alpha_L(g-,g-)$	−82.6	−14.9	6.26	1.09 (0.72)	−6.73 (−6.26)	−0.59 (0.03)	2.92 (3.08)	1.00 (0.91)	−0.13 (−0.12)	−0.80 (−0.6)	0.77 (0.56)
$\varepsilon_D(g-,g+)$	56.0	−137.0	9.52	7.61 (7.43)	−7.98 (−7.50)	2.78 (3.07)	0.84 (1.72)	0.98 (0.96)	0.64 (0.43)	−0.22 (−0.19)	−0.14 (−0.12)
$\varepsilon_D(a,a)$	71.3	−170.2	6.97	4.67 (3.89)	−5.67 (−4.56)	−2.63 (−2.20)	3.14 (3.98)	1.46 (1.36)	0.70 (0.42)	−1.47 (−1.37)	0.23 (0.14)
$\varepsilon_D(a,g+)$	69.7	−167.1	7.82	5.24 (4.56)	−7.85 (−6.74)	−2.45 (−2.12)	2.96 (3.90)	1.46 (1.40)	1.06 (0.85)	−1.49 (−1.52)	0.26 (0.19)
$\varepsilon_D(g+,g-)$	89.9	−142.0	4.90	−3.54 (−3.62)	−6.52 (−5.86)	1.87 (2.25)	0.26 (0.56)	0.52 (0.47)	0.91 (0.83)	−0.55 (−0.45)	0.18 (0.05)
$\varepsilon_D(g+,a)$	41.8	−115.6	12.13	4.30 (3.89)	−5.81 (−5.33)	−3.22 (−3.26)	0.94 (1.13)	1.26 (1.20)	0.89 (0.85)	−0.18 (−0.03)	0.16 (0.14)
$\gamma_D(g-,g-)$	75.3	−50.5	7.85	4.82 (4.82)	−15.78 (−14.49)	2.45 (2.76)	3.34 (3.46)	1.17 (1.15)	0.65 (0.52)	−0.22 (−0.01)	−0.39 (−0.49)
$\gamma_D(g-,a)$	77.0	−55.5	6.88	4.82 (4.66)	−12.34 (−11.09)	3.20 (3.75)	3.96 (3.99)	1.17 (1.12)	0.29 (0.09)	−0.24 (−0.07)	−0.45 (−0.35)
$\gamma_D(g-,g+)$	76.5	−55.4	7.30	2.78 (2.48)	−13.96 (−12.55)	2.83 (3.04)	3.62 (3.58)	1.12 (1.10)	0.34 (0.11)	−0.28 (−0.25)	−0.40 (−0.36)
$\gamma_D(a,a)$	75.3	−74.6	8.64	1.45 (1.53)	−10.11 (−8.89)	−0.58 (−0.06)	3.47 (3.45)	1.00 (0.99)	0.35 (0.13)	−1.37 (−1.13)	0.53 (0.41)
$\gamma_D(a,g+)$	74.2	−61.4	8.78	3.26 (3.22)	−13.05 (−11.6)	0.55 (1.20)	3.26 (3.63)	1.17 (1.19)	0.67 (0.56)	−1.33 (−1.29)	0.57 (0.48)
$\gamma_D(g+,g-)$	80.2	−37.7	6.75	−8.42 (−7.98)	−13.06 (−11.87)	−2.48 (−1.95)	0.56 (0.52)	0.97 (0.96)	0.62 (0.50)	−0.35 (−0.22)	0.02 (−0.09)
$\gamma_D(g+,a)$	58.4	−27.7	12.05	0.90 (0.63)	−11.95 (−11.18)	−5.68 (−5.39)	0.74 (1.00)	1.62 (1.55)	0.66 (0.54)	−0.26 (−0.13)	−0.05 (0.05)
$\gamma_D(g+,g+)$	62.7	−18.6	10.91	1.65 (1.84)	−10.19 (−9.26)	−4.75 (−4.49)	−0.21 (0.14)	1.33 (1.27)	0.63 (0.53)	−0.22 (−0.23)	−0.14 (−0.15)
$\alpha_D(g-,a)$	66.4	33.2	7.09	8.23 (7.12)	−5.78 (−4.84)	2.67 (3.04)	3.32 (3.45)	1.23 (1.06)	0.49 (0.25)	−0.78 (−0.59)	−0.09 (−0.07)
$\alpha_D(g+,g+)$	48.3	50.8	8.57	5.55 (4.39)	−6.02 (−5.04)	−2.20 (−2.72)	1.84 (1.97)	1.76 (1.63)	0.93 (0.82)	0.25 (0.25)	−0.18 (−0.15)
$\alpha_D(a,g-)$	64.3	34.7	6.33	2.97 (1.89)	−5.73 (−4.48)	−0.64 (−0.69)	−0.60 (0.05)	1.30 (1.29)	0.65 (0.40)	−1.21 (−1.08)	1.01 (0.84)
$\alpha_D(a,g+)$	66.5	39.8	14.31	−0.09 (−1.08)	−8.42 (−6.94)	−3.17 (−3.50)	4.50 (4.88)	1.16 (1.18)	0.98 (0.83)	−1.60 (−1.58)	0.89 (0.89)
$\delta_D(g-,g-)$	−148.4	−50.6	10.90	2.87 (2.53)	−7.03 (−5.93)	−3.69 (−3.20)	4.43 (4.58)	1.06 (0.98)	−0.80 (−0.80)	−0.43 (−0.26)	0.91 (0.78)
$\delta_D(a,g+)$	−147.5	−72.0	10.06	4.28 (4.12)	−6.30 (−5.47)	−1.09 (−0.81)	2.70 (2.80)	1.30 (1.27)	−0.60 (−0.62)	0.33 (0.39)	0.01 (−0.04)
$\delta_D(g+,g-)$	−153.9	−62.8	7.35	7.61 (7.38)	−7.00 (−6.47)	1.43 (1.29)	2.34 (2.31)	1.13 (1.14)	−0.25 (−0.29)	−0.81 (−0.63)	0.58 (0.46)
$\delta_D(g+,a)$	−154.0	−58.3	7.20	5.72 (5.44)	−5.14 (−4.59)	0.57 (0.85)	2.36 (2.36)	0.92 (0.88)	−0.51 (−0.59)	−0.78 (−0.55)	0.51 (0.53)

[a]The backbone conformers are labeled according to the set of abbreviations introduced in the past[45]: α_L, α_D, β_L, γ_L, γ_D, δ_L, δ_D, ε_L, and ε_D. For the side-chain orientation the following abbreviations are used: $g+ \cong +60°$, $a \cong \pm 180°$, and $g- \cong -60°$.

[b]Torsional angles ϕ and ψ are given in degrees ($-180° \leq \zeta \leq +180°$, where $\zeta = \phi, \psi$).

[c]Relative energies (in kcal mol^{-1}) are given with respect to $E^{\gamma L(g+,g+)}$, where $E^{\gamma L(g+,g+)} = -489.7936854$ Hartree at the RHF/6-311++G(d,p) level of theory. Absolute isotropic shielding values (in ppm) are referenced to those of the $\gamma_L(g+,g+)$ conformer calculated at GIAO-RHF/TZ2P (GIAO-RHF/6-31+G(d)) levels of theory; ^{15}N: 154.33 (169.19), $^{13}C^{\alpha}$: 147.37 (154.56), $^{13}C^{\beta}$: 137.71 (145.92), $^{13}C'$: 16.47 (32.45), $^{1}H^{N}$: 26.67 (27.62), $^{1}H^{\alpha}$: 28.68 (29.26), $^{1}H^{\beta 1}$: 29.11 (29.43), $^{1}H^{\beta 2}$: 28.39 (29.43).

Table 4. Backbone Conformational Parameters, Relative Energies, and Relative Isotropic Chemical Shielding Values (GIAO-RHF/TZ2P and GIAO-RHF/6-31+G(d), with the Latter Values in Parentheses) of 8 For-L-Ser-NH_2 Conformers Optimized at RHF/3-21G Level of Theory.

Conf.[a]	ϕ^b	ψ^b	ΔE^c	^{15}N (c)	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	$^{1}H^{\beta^1}$	$^{1}H^{\beta^2}$
$\beta_L(g-,a)$	−137.9	160.5	15.39	−0.74 (−0.84)	−0.91 (−0.78)	−7.12 (−6.90)	−2.99 (−2.58)	−0.15 (−0.15)	−1.07 (−1.20)	−0.21 (−0.06)	0.83 (0.79)
$\delta_L(g-,a)$	−130.2	30.0	13.28	−4.95 (−4.71)	−1.25 (−0.45)	−2.36 (−1.8)	−3.66 (−3.12)	1.60 (1.50)	−1.50 (−1.52)	−1.07 (−0.82)	0.79 (0.73)
$\delta_L(a,g-)$	−127.9	33.2	8.32	−4.52 (−4.19)	0.39 (1.25)	−3.55 (−3.57)	−8.31 (−7.30)	1.46 (1.43)	−1.29 (−1.37)	−0.11 (0.04)	0.76 (0.58)
$\gamma_L(a,g+)$	−86.6	77.8	12.51	−7.92 (−6.94)	−2.28 (−1.87)	−0.55 (−0.48)	−0.73 (−0.71)	0.98 (1.02)	−0.29 (−0.25)	−1.11 (−1.04)	0.81 (0.82)
$\varepsilon_D(g-,g-)$	68.9	178.2	20.51	7.29 (6.75)	−9.94 (−9.03)	−1.35 (−0.88)	2.29 (3.18)	1.23 (1.14)	0.89 (0.76)	−0.03 (0.16)	−1.35 (−1.54)
$\gamma_D(a,g-)$	67.5	−31.2	10.59	−5.55 (−4.99)	−10.44 (−9.10)	−1.87 (−1.54)	−8.61 (−7.97)	0.60 (0.62)	0.41 (0.20)	−0.92 (−0.83)	0.77 (0.60)
$\delta_D(g-,g+)$	146.3	−33.9	12.29	−10.38 (−10.43)	−14.73 (−13.47)	−2.50 (−2.18)	−1.93 (−1.56)	1.27 (1.15)	−0.19 (−0.34)	−0.44 (−0.38)	0.41 (0.44)
$\delta_D(a,a)$	−172.4	−55.1	17.23	−2.39 (−2.44)	−7.06 (−5.95)	−3.38 (−2.54)	−0.32 (0.36)	0.89 (0.93)	−0.56 (−0.66)	0.32 (0.49)	−0.93 (−0.92)

[a]See footnotes a, b, and c of Table 3.

Using the 0° to 360° definition of the torsional periodicity, approximately 90% of the chemical shift–structure Pearson coefficients, R_{cs}, fall in the −0.3 to +0.3 range and their average value, $\langle R_{cs} \rangle$, is −0.014. When the optimal definition is used for each dihedral angle, ϕ, ψ, χ^1, and χ^2, the results are approximately the same; around 90% of the R_{max} values fall in the −0.5 to +0.5 range and their average value, $\langle R_{max} \rangle$, is −0.010. Because $R \approx 0$ indicates no correlation, these results imply that most chemical shifts show only insignificant correlation with conformational (ϕ, ψ, χ^1, and χ^2) parameters. These results support previous findings that chemical shifts typically do not correlate with torsional variables.

Three correlations, however, are much larger than the others. The Pearson coefficients of the $^1H^{\alpha}/\phi$, $^{13}C^{\alpha}/\phi$, and $^{13}C^{\alpha}/\psi$ correlations, −0.816, 0.707, and −0.675, respectively, are significant. The corresponding R_{max} values for $^1H^{\alpha}/\phi$ and $^{13}C^{\alpha}/\phi$ are even larger, 0.834 and −0.727, respectively. These linear correlation values indicate that conformation exerts a significant influence over the $^1H^{\alpha}$ and $^{13}C^{\alpha}$ chemical shifts of For-L-Ser-NH_2. A closer look at the $^1H^{\alpha}/\phi$, $^{13}C^{\alpha}/\phi$, and $^{13}C^{\alpha}/\psi$ plots (Fig. 3A–C, respectively) reveal several important features of these chemical shift–backbone structure relationships: (1) regardless of the side-chain orientation "normal" γ-turn backbone conformers, γ_D, have the largest relative, with respect to the random coil value, $^{13}C^{\alpha}$ chemical shift (≈62–68 ppm); (2) β-pleated sheet (β_L), inverse γ-turn (γ_L), and δ_L conformers have the smallest relative $^{13}C^{\alpha}$ chemical shifts (≈51–55 ppm); (3) both types of helices (α_L and α_D), δ_D, and ε_D conformers have $^{13}C^{\alpha}$ chemical shifts that range in between those of (1) and (2); and (4) D-type conformers, with the exception of δ_D, have smaller (≈3.5–4.3 ppm) while L-type conformers have larger (≈4.8–6.0 ppm) relative $^1H^{\alpha}$ chemical shifts.

As the R_{cs} and R_{max} values indicate, however, the $^1H^{\alpha}/\phi$, $^{13}C^{\alpha}/\phi$, and $^{13}C^{\alpha}/\psi$ correlations are not perfectly linear (i.e., neither R_{cs} nor R_{max} is ±1.0). Hence, chemical shifts can be assigned only with some uncertainty to a specific type of backbone conformer. In these cases a number of acceptable possibilities exist and all of them must be considered during a structure solution process. Nevertheless, this information is highly useful because it limits the (ϕ,ψ) dihedral space of the residue that has to be explored during the structure solution process.

Chemical Shift–Chemical Shift Correlation Maps

Typical two-dimensional (2D) and three-dimensional (3D) heteronuclear experiments, based on single-quantum coherence (HSQC) and multiple-quantum coherence (HMQC) transfer, are an indirect source of structural information. If each typical backbone conformer (i.e., α_L, β_L, etc.) had a unique set of hetero- and homonuclear chemical shifts (or at least part of the set was unique to the backbone conformer type) then nonoverlapping regions of chemical shifts would arise on appropriately selected 2D and 3D spectra. Identification of these regions to specific backbone conformer types would yield important structural information.

Two distinctly different types of 2D-HSQC-type spectra, ^{15}N-$^1H^N$ and $^{13}C^{\alpha}$-$^1H^{\alpha}$, calculated at level B, are shown for For-L-Ser-NH_2 in Figures 4 and 5, respectively. Note that during common spectral analysis of doubly labeled (^{13}C and ^{15}N) proteins both types of heterocorrelations are typically recorded within more

Table 5. Selected Parameters Associated with Backbone Conformational Averages of For-L-Ser-NH_2 Determined at Two Levels of Theory (GIAO-RHF/6-31+G(d) and GIAO-RHF/TZ2P).

A—Averaged over the individual backbone conformers

bb averages[a]	Level[b]	^{15}N (c)	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	$^{1}H^{\beta1}$	$^{1}H^{\beta2}$	ϕ	ψ	ΔE[d]	$p(i)/\Sigma p(i)$[e]
α_L (3)[f]	A	95.77	51.52	55.21	165.51	3.90	3.89	3.80	3.32	−79.09	−22.38	9.29	0.00
	B	111.97	52.74	57.49	175.18	4.23	3.83	3.74	3.15			8.61	0.00
α_D (4)	A	94.67	52.00	56.29	166.20	3.89	2.97	4.12	3.54	61.38	39.64	9.71	0.00
	B	110.20	53.81	57.81	175.95	4.20	2.79	3.96	3.43			9.10	0.00
β_L (6)	A	94.39	49.03	58.27	166.51	5.30	3.95	3.61	3.50	−155.87	171.53	5.37	0.03
	B	110.63	50.15	59.99	176.36	5.83	3.87	3.44	3.41			5.14	0.04
γ_L (6)	A	100.21	47.61	54.69	167.29	4.15	3.86	3.75	3.60	−86.51	74.39	5.07	0.97
	B	117.32	48.57	56.51	176.53	4.52	3.83	3.60	3.45			4.95	0.96
γ_D (9)	A	97.06	57.79	55.62	167.48	4.07	3.19	3.83	3.91	71.91	−45.84	9.42	0.00
	B	113.73	59.64	57.68	177.09	4.43	3.04	3.70	3.79			9.13	0.00
δ_L (4)	A	98.63	47.13	56.95	169.95	3.87	4.62	3.79	3.50	−126.63	23.01	7.29	0.00
	B	115.41	48.34	58.92	179.90	4.19	4.62	3.72	3.33			7.76	0.00
δ_D (6)	A	96.65	53.66	56.42	166.98	4.12	4.09	3.53	3.71	−165.00	−55.47	10.80	0.00
	B	113.08	55.19	58.42	176.62	4.47	4.04	3.42	3.59			10.80	0.00
ε_D (6)	A	93.93	53.18	55.84	166.38	4.09	2.85	3.94	4.11	66.28	−152.26	10.29	0.00
	B	110.10	54.61	57.81	176.48	4.41	2.70	3.78	3.95			10.01	0.00

B—Averaged over all backbone conformers using arithmetical averaging

bb Averages[a]	Level[b]	^{15}N (c)	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	$^{1}H^{\beta1}$	$^{1}H^{\beta2}$
Average	A	96.48	52.09	56.15	167.08	4.21	3.62	3.78	3.70
Standard deviation	A	4.54	4.14	2.56	2.80	0.57	0.62	0.50	0.50
Average	B	112.93	53.52	58.07	176.79	4.58	3.53	3.65	3.57
Standard deviation	B	4.84	4.42	2.53	2.91	0.62	0.66	0.50	0.52
Experimental average shift[h]		**116.47**	**58.48**	**63.96**	**174.58**	**8.27**	**4.51**	**3.84**	**3.84**
Exp. standard deviation		**4.03**	**2.23**	**5.25**	**1.67**	**0.83**	**0.44**	**0.32**	**0.33**
Diff = exp. − calc.	A	19.99	6.39	7.81	7.50	4.06	0.89	0.06	0.14
% error[g]	A	17	11	12	4	49	20	1	4
Diff = exp. − calc.	B	3.54	4.96	5.89	−2.21	3.69	0.98	0.19	0.27
% error	B	3	8	9	1	45	22	5	7

C—Averaged over all backbone conformers using Boltzmann averaging

bb Averages[a]	Level[b]	^{15}N (c)	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	$^{1}H^{\beta1}$	$^{1}H^{\beta2}$
Average	A	97.44	46.72	55.33	168.62	5.16	3.56	3.37	3.91
	B	114.15	47.40	57.00	178.14	5.55	3.56	3.12	3.83
Experimental average shift		**116.47**	**58.48**	**63.96**	**174.58**	**8.27**	**4.51**	**3.84**	**3.84**
Exp. standard deviation		**4.03**	**2.23**	**5.25**	**1.67**	**0.83**	**0.44**	**0.32**	**0.33**
Diff = exp. − calc.	B	19.00	11.74	8.61	5.90	3.11	0.95	0.47	−0.07
% error[h]	B	16	20	13	3	38	21	12	2
Diff = exp. − calc.	A	2.32	11.08	6.96	−3.56	2.72	0.95	0.72	0.01
% error	A	2	19	11	2	33	21	19	0

[a]44 fully optimized conformers (36 RHF/6-311++G(d,p) and 8 RHF/3-21G) were used for averaging according to the backbone parameters.

[b]Level A: GIAO-RHF/6-31+G(d)//RHF/6-311++G(d,p) plus GIAO-RHF/6-31+G(d)//RHF/3-21G. Level B: GIAO-RHF/TZ2P//RHF/6-311++G(d,p) plus GIAO-RHF/TZ2P//RHF/3-21G.

[c]Averages of absolute chemical shielding values relative to TMS or NH_3

Method	Used for	H (TMS)	C (TMS)	N (NH_3)
GIAO-RHF/6-31+G(d)	A	32.80	201.24	266.94
GIAO-RHF/TZ2P	B	32.23	194.69	268.70

TMS and NH_3 geometry were calculated at RHF/3-21G and CCSD(T)/aug-cc-pVTZ levels of theory, respectively. Chemical shifts have not been rereferenced prior to averaging.

[d]Averages of relative energies determined at GIAO-RHF/631+G(d) (Levels A) and GIAO-RHF/TZ2P (Level B) levels of theory.

[e]Relative populations are calculated as $\exp(-\Delta E/kT)/\Sigma \exp(-\Delta E/kT)$, where $kT = 0.595371$ (kcal/mol) at $T = 300$ K (if $k = 1.38*10^{-23}$ [J/K] and Avogadro's number is $6.02*10^{23}$ (mol^{-1})).

[f]Number of side chain rotamers optimized previously.[36]

[g]Experimental average shift taken from the BMRB on 02-02-1999.

[h]% error = $[|(CSI^{exp.} - CSI^{calc.})|/CSI^{exp.}] * 100\%$.

Table 6. Pearson Correlation Coefficients, R, for the Correlations of Selected Conformational and Isotropic Chemical Shielding Values of For-L-Ser-NH_2, Calculated at Level A (GIAO-RHF/6-31+G(d)).[a,b]

	ω_0	ϕ	ψ	ω_1	χ^1	χ^2	^{15}N	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	$^{1}H^{\beta1}$	$^{1}H^{\beta2}$
ω_0	1	0.338	−0.127	−0.448	0.220	0.140	−0.160	0.311	0.092	0.002	0.160	−0.178	−0.224	0.102
							−0.197	**0.692**	0.032	−0.035	−0.165	**−0.779**	0.186	0.349
ϕ		1	−0.249	−0.208	0.132	0.100	(0.374)	**(−0.717)**	(−0.330)	(0.294)	(−0.474)	**(0.806)**	(0.400)	(−0.402)
							0.078	**−0.689**	0.028	0.085	0.003	0.204	0.131	−0.159
ψ			1	−0.008	−0.072	0.030	(−0.531)	**(−0.689)**	(−0.324)	(0.402)	(0.518)	(−0.423)	(0.229)	(0.408)
ω_1				1	−0.095	−0.111	−0.238	−0.273	−0.043	−0.041	0.251	0.220	−0.012	−0.108
							−0.158	−0.079	0.201	0.081	0.160	−0.248	−0.091	0.023
χ^1					1	0.055	(0.392)	(−0.271)	(−0.370)	(−0.307)	(−0.210)	(−0.248)	(0.476)	(−0.353)
							−0.022	0.090	0.141	−0.172	0.052	−0.146	−0.008	0.096
χ^2						1	(0.332)	(−0.222)	(0.344)	(0.343)	(−0.296)	(0.282)	(0.305)	(−0.332)
^{15}N							1	0.024	0.316	0.540	−0.074	0.216	0.082	−0.278
$^{13}C^{\alpha}$								1	−0.072	−0.149	−0.191	−0.579	0.062	0.337
$^{13}C^{\beta}$									1	0.282	0.133	0.106	−0.052	−0.222
$^{13}C'$										1	0.007	0.300	−0.129	−0.172
$^{1}H^{N}$											1	0.085	−0.136	−0.007
$^{1}H^{\alpha}$												1	−0.230	−0.355
$^{1}H^{\beta1}$													1	−0.563
$^{1}H^{\beta2}$														1

[a]A total of 44 conformers were used to generate data vectors to be correlated. All conformational variables (ω_0, ϕ, ψ, ω_1, χ^1, and χ^2) are scaled between 0° and 360°. Autocorrelation values (e.g., $R[\phi/\phi]$) are all equal to 1. The Pearson correlation coefficient of chemical shift–structure correlations has been designated as R_{cs}. The average of all R_{cs} values, $\langle R_{cs}\rangle$, is −0.010 with a standard deviation (σ_{cs}) of 0.263. For reasons mentioned in the text correlations involving the ω_0 and ω_1 torsion angles have been left out of the $\langle R_{cs}\rangle$ calculation. All values outside the range of $\langle R_{cs}\rangle \pm 2\sigma_{cs}$ (−0.010 ± 0.526) are marked in bold.

[b]Selected maximized R values, R_{max}, are given in parentheses. These were obtained by shifting systematically (with an increment of 5°) the periodic unit of the conformational variables ϕ, ψ, χ^1, and χ^2.

Table 7. Pearson Correlation Coefficients, R, for the Correlations of Selected Conformational and Isotropic Chemical Shielding Values of For-L-Ser-NH_2, Calculated at Level B (GIAO-RHF/TZ2P).[a,b]

	ω_0	ϕ	ψ	ω_1	χ^1	χ^2	^{15}N	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	$^{1}H^{\beta 1}$	$^{1}H^{\beta 2}$
ω_0	1	0.338	−0.127	−0.448	0.220	0.140	−0.173	0.322	0.045	−0.015	0.184	−0.195	−0.259	0.148
							−0.225	**0.707**	0.026	0.004	−0.182	**−0.816**	0.152	0.347
ϕ		1	−0.249	−0.208	0.132	0.100	(0.381)	**(−0.727)**	(−0.332)	(0.264)	(−0.486)	**(0.834)**	(0.394)	(−0.412)
							0.071	**−0.675**	−0.022	0.082	0.004	0.214	0.114	−0.161
ψ			1	−0.008	−0.072	0.030	(−0.530)	**(−0.675)**	(−0.300)	(0.391)	(0.530)	(−0.438)	(0.237)	(0.370)
ω_1				1	−0.095	−0.111	−0.228	−0.284	−0.023	0.001	0.267	0.226	0.020	−0.105
							−0.153	−0.090	0.156	0.090	0.183	−0.209	−0.130	0.067
χ^1					1	0.055	(0.380)	(−0.280)	(−0.354)	(−0.320)	(0.227)	(−0.279)	(0.462)	(−0.380)
							−0.033	0.084	0.114	−0.173	0.059	−0.154	−0.113	0.200
χ^2						1	(0.329)	(−0.226)	(0.332)	(0.339)	(−0.315)	(0.265)	(0.396)	(−0.399)
^{15}N							1	0.018	0.316	0.503	−0.080	0.249	0.074	−0.292
$^{13}C^{\alpha}$								1	−0.037	−0.132	−0.225	−0.600	0.073	0.315
$^{13}C^{\beta}$									1	0.303	0.118	0.133	−0.066	−0.203
$^{13}C'$										1	0.004	0.256	−0.109	−0.181
$^{1}H^{N}$											1	0.111	−0.163	0.006
$^{1}H^{\alpha}$												1	−0.187	−0.348
$^{1}H^{\beta 1}$													1	−0.593
$^{1}H^{\beta 2}$														1

[a]A total of 44 conformers were used to generate data vectors to be correlated. All conformational variables (ω_0, ϕ, ψ, ω_1, χ^1, and χ^2) are scaled between 0° and 360°. Autocorrelation values (e.g., $R[\phi/\phi]$) are all equal to 1. The Pearson correlation coefficient of chemical shift–structure correlations has been designated as R_{cs}. The average of all R_{cs} values, $\langle R_{cs} \rangle$, is −0.014 with a standard deviation (σ_{cs}) of 0.268. For reasons mentioned in the text correlations involving the ω_0 and ω_1 torsion angles have been left out of the $\langle R_{cs} \rangle$ calculation. All values outside the range of $\langle R_{cs} \rangle \pm 2\sigma_{cs}$ (−0.014 ± 0.537) are marked in bold.

[b]Selected maximized R values, R_{max}, are given in parentheses. These were obtained by shifting systematically (with an increment of 5°) the periodic unit of the conformational variables ϕ, ψ, χ^1, and χ^2.

Table 8. Experimental Chemical Shift Isotropy Values (δ Scale) of 79 Serines with Conformational Parameters ϕ, ψ, and χ^1 Retrieved from 18 Carefully Selected and Scaled Proteins.[a]

^{15}N	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	*bb*[b]	*sc*[b]	ϕ	ψ	χ^1	ϕ^c	ψ^c	diff[d]	$\chi^{1\,c}$	diff
113.7	61.3	61.6	175.3	8.0	4.3	α_L	*a*	−69.2	−44.5	178.7	290.8	315.5	15	178.7	1
114.1	60.9	62.8	****	7.5	4.4	α_L	*a*	−65.7	−40.4	−179.6	294.3	319.6	10	180.4	0
114.4	61.2	63.5	****	8.0	4.2	α_L	*a*	−64.3	−44.0	−179.9	295.7	316.0	14	180.1	0
****	59.3	****	****	8.6	4.3	α_L	*a*	−90.0	−22.2	−176.3	270.0	337.8	26	183.7	4
117.7	63.4	62.6	173.4	8.5	4.2	α_L	*g*−	−62.2	−41.3	−61.9	297.8	318.7	12	298.1	2
111.7	61.7	62.9	176.8	8.3	3.7	α_L	*g*−	−63.0	−43.1	−70.6	297.0	317.0	13	289.4	11
118.0	63.2	62.4	175.4	7.6	4.3	α_L	*g*−	−63.0	−38.5	−77.8	297.0	321.5	9	282.2	18
118.1	59.1	64.0	173.6	9.1	4.6	α_L	*g*−	−105.7	−24.7	−70.4	254.4	335.3	41	289.6	10
118.7	62.5	64.0	****	8.6	4.6	α_L	*g*−	−78.4	−36.7	−57.6	281.6	323.3	15	302.4	2
116.6	56.4	66.0	174.2	9.4	4.8	α_L	*g*−	−49.9	−47.1	−84.4	310.1	312.9	23	275.6	24
111.7	58.9	63.8	175.4	7.1	4.4	α_L	*g*−	−84.8	−17.0	−71.5	275.2	343.0	24	288.5	12
113.5	58.6	66.1	174.2	9.1	4.8	α_L	*g*−	−77.6	−25.0	−81.7	282.4	335.0	14	278.3	22
113.4	61.4	62.8	175.7	7.1	4.3	α_L	*g*−	−64.2	−45.6	−70.1	295.8	314.4	16	289.9	10
118.1	61.6	62.2	177.3	9.2	4.2	α_L	*g*−	−61.3	−36.9	−67.2	298.7	323.1	8	292.8	7
119.1	61.2	****	****	9.3	4.3	α_L	*g*−	−68.1	−43.9	−70.5	292.0	316.1	14	289.5	10
116.2	62.0	62.8	****	8.3	4.4	α_L	*g*−	−63.2	−45.1	−76.2	296.8	314.9	15	283.8	16
****	59.7	****	****	9.0	4.4	α_L	*g*−	−74.8	−11.1	−66.6	285.2	349.0	21	293.4	7
112.0	59.4	64.4	171.5	7.3	4.4	α_L	*g*+	−79.2	−16.1	73.1	280.8	343.9	20	73.1	13
113.2	60.2	62.0	174.9	9.0	4.3	α_L	*g*+	−65.3	−41.9	68.7	294.7	318.1	12	68.7	9
116.9	59.9	62.8	175.2	8.8	4.4	α_L	*g*+	−68.2	−18.9	72.2	291.8	341.1	12	72.2	12
117.4	59.4	63.7	175.3	8.7	4.4	α_L	*g*+	−79.9	−18.2	83.2	280.1	341.8	19	83.2	23
114.9	61.1	63.4	174.4	8.8	4.5	α_L	*g*+	−73.6	−26.4	64.4	286.4	333.6	9	64.4	4
120.4	57.9	62.5	175.5	7.4	4.7	α_L	*g*+	−72.1	−21.3	54.0	287.9	338.7	11	54.0	6
112.7	59.2	64.6	174.0	9.0	4.6	α_L	*g*+	−87.7	−17.8	62.3	272.3	342.2	26	62.3	2
116.9	61.8	****	177.2	9.4	4.1	α_L	*g*+	−58.8	−17.7	77.1	301.2	342.3	14	77.1	17
112.4	60.8	63.1	174.8	8.2	4.1	α_L	*g*+	−64.8	−37.1	28.6	295.2	323.0	7	28.6	31
105.0	57.2	63.4	175.0	7.1	4.4	α_L	*g*+	−79.8	−8.1	35.7	280.2	351.9	26	35.7	24
115.0	60.9	62.3	178.6	8.5	4.3	α_L	*g*+	−63.9	−29.6	59.4	296.1	330.4	1	59.4	1
119.3	61.0	63.4	175.3	9.1	4.8	α_L	*g*+	−73.4	−13.7	67.9	286.6	346.3	18	67.9	8
122.2	62.1	62.8	175.0	8.9	4.0	α_L	*g*+	−56.7	−35.3	74.6	303.3	324.7	10	74.6	15
112.9	61.7	****	****	8.1	4.2	α_L	*g*+	−67.9	−32.0	63.4	292.1	328.0	4	63.4	3
116.2	61.5	****	****	7.9	4.3	α_L	*g*+	−68.0	−25.3	80.5	292.0	334.7	6	80.5	20
****	59.3	****	****	8.1	4.4	α_L	*g*+	−85.1	−9.3	90.3	274.9	350.7	29	90.3	30
****	58.2	****	****	8.0	4.3	α_L	*g*+	−99.4	−14.8	23.2	260.6	345.2	38	23.2	37
****	60.4	****	****	9.0	4.4	α_L	*g*+	−57.9	−24.3	62.5	302.1	335.7	9	62.5	3
****	61.6	****	****	8.6	4.2	α_L	*g*+	−64.0	−28.7	67.6	296.0	331.3	2	67.6	8
117.4	58.9	64.2	****	8.1	4.3	α_L	*g*+	−89.1	−6.8	80.1	270.9	353.2	33	80.1	20
113.5	57.4	65.1	173.3	8.8	4.8	β_L	*a*	−145.4	153.3	−156.5	214.6	153.3	6	203.5	23
113.9	57.8	65.1	171.9	8.4	4.6	β_L	*a*	−165.2	138.3	−173.7	194.9	138.3	28	186.3	6
120.8	56.1	64.4	174.0	8.7	5.7	β_L	*a*	−113.1	127.0	173.1	246.9	127.0	35	173.1	7
118.9	56.5	64.4	173.0	8.5	4.6	β_L	*a*	−129.9	119.3	165.5	230.2	119.3	32	165.5	14
110.5	56.3	62.8	174.2	6.7	4.3	β_L	*a*	−168.4	−175	−160.7	191.6	184.7	45	199.3	19
107.8	56.9	****	****	7.5	4.6	β_L	*a*	−152.1	144.4	155.0	207.9	144.4	13	155.0	25
111.1	57.1	****	****	8.8	4.9	β_L	*a*	−136.4	115.8	173.5	223.6	115.8	34	173.5	7
121.6	56.8	****	172.8	9.0	5.5	β_L	*a*	−135.0	127.6	175.1	225.0	127.6	23	175.1	5
117.5	57.0	63.2	172.3	8.1	4.8	β_L	*g*−	−126.6	160.9	−59.1	233.4	160.9	17	301.0	1
117.0	55.9	65.6	173.6	9.9	6.1	β_L	*g*−	−130.3	114.1	−51.2	229.7	114.1	37	308.8	9
120.3	55.9	****	174.5	7.5	5.0	β_L	*g*−	−112.4	142.9	−60.7	247.6	142.9	28	299.4	1
****	56.7	****	****	8.9	4.9	β_L	*g*−	−104.5	156.5	−92.1	255.5	156.5	36	267.9	32
103.0	55.8	66.8	171.9	7.5	4.5	β_L	*g*+	−141.9	153.7	72.2	218.1	153.7	4	72.2	12
112.4	57.7	66.7	172.2	7.6	4.7	β_L	*g*+	−151.3	157.8	52.3	208.7	157.8	14	52.3	8
113.8	56.8	65.5	171.9	7.2	3.9	β_L	*g*+	−155.2	167.9	65.3	204.9	167.9	23	65.3	5
113.4	56.3	64.2	175.3	7.4	4.6	β_L	*g*+	−113.4	144.8	57.8	246.6	144.8	27	57.8	2
118.9	56.7	70.8	175.9	8.5	5.9	β_L	*g*+	−138.9	167.6	79.8	221.1	167.6	18	79.8	20
112.7	55.8	****	****	8.4	5.6	β_L	*g*+	−115.7	146.0	55.5	244.3	146.0	25	55.5	5
119.6	57.2	****	174.5	9.0	4.9	β_L	*g*+	−151.5	−1790	68.4	208.5	181.0	33	68.4	8

(*continued*)

Table 8. (*Continued*)

^{15}N	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$	bb^{b}	sc^{b}	ϕ	ψ	χ^{1}	ϕ^{c}	ψ^{c}	diff[d]	$\chi^{1\,c}$	diff
****	57.1	****	****	7.7	4.7	β_L	$g+$	−160.3	166.4	75.0	199.7	166.4	26	75.0	15
****	56.9	****	****	8.1	5.3	β_L	$g+$	−125.0	140.9	55.8	235.0	140.9	18	55.8	4
****	57.4	****	****	7.7	4.5	δ_L	$g-$	−93.2	1.8	−43.2	266.8	1.8	34	316.8	17
104.2	56.5	63.3	174.9	6.9	3.7	δ_L	$g+$	−109.2	−1.0	68.6	250.8	−1.0	23	68.6	9
114.3	56.0	66.9	173.4	7.7	4.9	δ_L	$g+$	−126.7	5.0	59.3	233.3	5.0	10	59.3	1
113.7	56.0	63.3	173.7	8.1	4.8	δ_L	$g+$	−130.5	46.9	47.3	229.6	46.9	32	47.3	13
****	57.0	****	****	7.4	4.6	δ_L	$g+$	−125.7	23.6	75.3	234.3	23.6	9	75.3	15
118.8	56.3	63.9	****	8.8	4.7	ε_L	a	−71.8	143.4	142.5	288.2	143.4	8	142.5	38
115.4	56.1	63.0	171.9	8.3	4.7	ε_L	a	−87.7	124.3	163.4	272.3	124.3	28	163.4	17
115.5	58.4	62.8	176.1	8.3	4.5	ε_L	a	−63.5	120.2	166.4	296.5	120.2	20	166.4	14
****	56.9	****	****	8.1	4.9	ε_L	a	−84.6	122.2	165.0	275.4	122.2	26	165.0	15
****	58.8	****	****	7.0	4.0	ε_L	a	−67.6	125.0	176.0	292.4	125.0	15	176.0	4
116.4	60.7	65.2	172.4	7.7	4.7	ε_L	$g-$	−71.2	159.5	−73.2	288.9	159.5	20	286.8	13
113.8	57.7	64.5	172.7	7.7	4.8	ε_L	$g+$	−84.0	155.5	66.5	276.0	155.5	25	66.5	6
110.7	58.2	63.7	171.8	7.4	4.1	ε_L	$g+$	−72.2	165.8	78.8	287.8	165.8	27	78.8	19
117.4	56.9	66.3	174.9	8.1	5.0	ε_L	$g+$	−90.8	171.3	81.6	269.3	171.3	41	81.6	22
116.6	57.6	65.9	172.6	8.6	5.4	ε_L	$g+$	−90.1	171.6	72.5	269.9	171.6	40	72.5	12
120.4	56.4	66.1	176.7	8.7	5.2	ε_L	$g+$	−93.5	164.1	77.8	266.5	164.1	37	77.8	18
115.6	56.7	65.8	175.1	8.4	4.7	ε_L	$g+$	−84.9	155.9	75.0	275.1	155.9	25	75.0	15
****	57.7	****	****	9.2	4.3	ε_L	$g+$	−73.4	148.9	75.1	286.6	148.9	12	75.1	15
****	57.3	****	****	9.1	4.7	ε_L	$g+$	−71.2	143.8	64.9	288.8	143.8	7	64.9	5
120.7	55.9	65.8	****	8.4	4.6	ε_L	$g+$	−85.2	162.9	76.2	274.8	162.9	30	76.2	16
118.3	56.9	****	****	7.9	4.4	ε_D	$g+$	53.1	−134	60.4	53.1	226.3	16	60.4	0

[a]Structures have been obtained from the Protein Data Bank (PDB). Chemical shifts for the corresponding structures were taken from the BioMagnetic Resonance Bank (BMRB). Chemical shifts not submitted to the BMRB are indicated by ****.

[b]Symbols for relevant conformational types: α_L, β_L, ε_L, etc., for backbone (bb), and $g+$, a or $g-$ for the side-chain (sc) orientation.

[c]$0° \leq \phi \leq 360°$ and $0° \leq \psi \leq 360°$.

[d]Difference between $[\phi, \psi]_i$ value of residue i and the appropriate reference value associated with the given type of backbone conformer.

complex spectral data. Clearly, the ^{15}N-$^{1}H^{N}$ spectrum cannot be used to extract structural information from chemical shifts alone. Although chemical shifts associated with a β_L-type backbone orientation fall into a well-separated region, the resonances of other conformers are mixed together and can hardly be distinguished from each other. The $^{13}C^{\alpha}$-$^{1}H^{\alpha}$ spectrum appears to be considerably more useful for structure analysis than the ^{15}N-$^{1}H^{N}$ plot. In the latter type of heterocorrelation spectrum the clustering of the C^{α}/H^{α} chemical shift values as a function of the different backbone conformations is clear. Regardless of the side chain orientation most of the serine backbone structures can be distinguished, as done previously for Gly, Ala and Phe. The δ_D, α_L, and γ_D chemical shifts fall into separate regions, the ε_D and α_D regions overlap only slightly, just like the β_L and δ_L ones, and it is only the

Table 9. Pearson Correlation Coefficient, R, Values Determined Between Major Periodic Conformational Variables (ϕ, ψ, and χ^1) and Chemical Shifts Assigned to Serines in 18 Proteins.[a,b]

	^{15}N	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$
ϕ	0.091 (−0.316)[c]	0.580 (−0.704)	−0.496 (0.538)	0.482 (−0.499)	0.202 (−0.216)	−0.307 (0.506)
ψ	−0.043 (0.218)	0.719 (0.745)	−0.482 (0.584)	0.460 (0.460)	0.132 (0.290)	−0.509 (−0.530)
χ^1	0.190 (−0.240)	0.220 (−0.354)	−0.188 (0.280)	−0.015 (−0.165)	0.128 (−0.168)	0.027 (−0.153)

[a]Structures have been obtained from the Protein Data Bank (PDB).[42] Chemical shifts for the corresponding structures were taken from the BioMagnetic Resonance Bank (BMRB).[41]

[b]The average R value, $\langle R \rangle$, is 0.066 with a standard deviation (σ) of 0.364. None of the values are larger in magnitude than $\langle R \rangle \pm 2\sigma$, but those of $^{13}C^{\alpha}/\phi$ and $^{13}C^{\alpha}/\psi$ are close enough.

[c]Maximized R values, R_{max}, are given in parentheses, obtained by shifting systematically (with an increment of 5°) the periodic unit of the conformational variables ϕ, ψ, and χ^1.

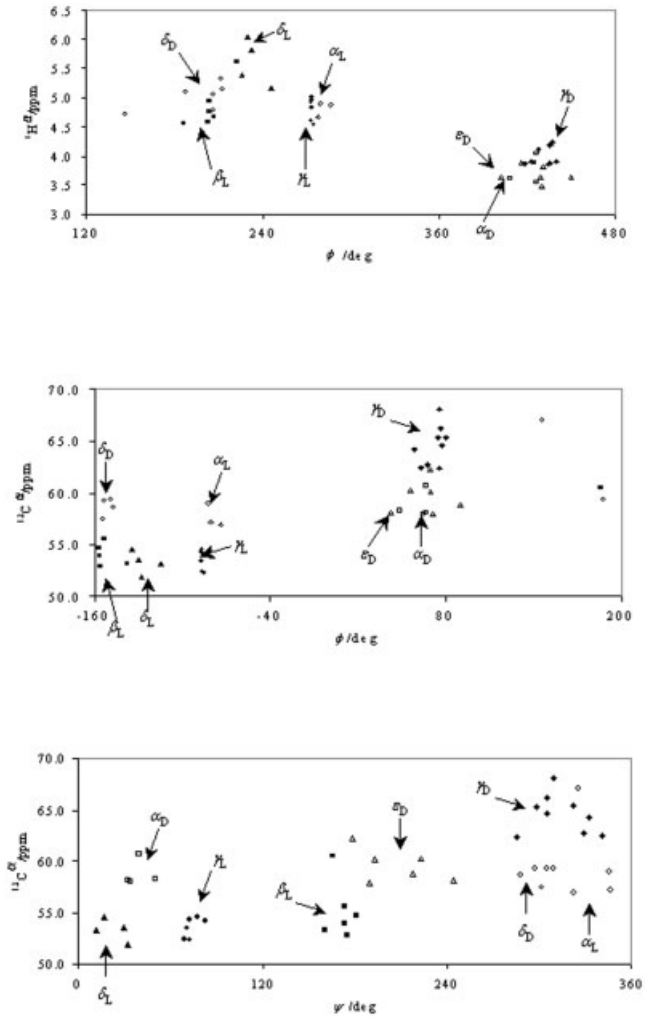

Figure 3. Illustration of chemical shift–structure correlations for the 44 optimized conformers of *trans*-For-L-Ser-NH_2. All chemical shifts have been calculated at the GIAO-RHF/TZ2P level of theory. (A) $^1H^\alpha/\phi$, (B) $^{13}C^\alpha/\phi$, and (C) $^{13}C^\alpha/\psi$ where different symbols represent different backbone conformations [■ = β_L (β-pleated sheet), ◇ = α_L (α-helix), ◆ = γ_D (γ-turn), ● = γ_L (inverse γ-turn), ○ = δ_D, ▲ = δ_L, △ = ε_D, and □ = α_D]. Correction factors for $^1H^\alpha$ and $^{13}C^\alpha$ are 0.98 and 4.96 ppm, respectively (see text for details). In these images ϕ and ψ are not illustrated from −180° to 180°, as is customary, rather their ranges correspond to those used in achieving R_{max} (see text for details).

β_L and γ_L chemical shifts that are inseparable from each other. This suggests that the location of chemical shifts in such spectra could be used to identify certain backbone conformers (mainly γ_D, α_L, and δ_D) of serine residues. The degree of chemical shift separation in other 2D-HSQC-type spectra (not shown) ranges between those of the ^{15}N-$^1H^N$ and $^{13}C^\alpha$-$^1H^\alpha$ plots.

Comparing the results of these analyses with the 1D analogues it is clear that while $^{13}C^\alpha$-$^1H^\alpha$ plots allow conformer types to be identified, the 1D $^{13}C^\alpha$ or $^1H^\alpha$ plots do not. On the other hand, the ^{15}N-$^1H^N$ plot does not provide significant advantages over its 1D constituents.

Of all the 3D-HSQC-type spectra examined the most promising one, involving the $^{13}C^\alpha$, $^1H^\alpha$, and $^1H^N$ nuclei, is illustrated in Figure 6. In this spectrum the γ_D, γ_L, δ_L, β_L and α_L backbone conformers can be clearly separated, and only the ε_D and α_D backbone conformers are interwoven in the same region of space. This is an improvement over the results of the $^{13}C^\alpha$-$^1H^\alpha$ 2D-HSQC-type spectrum where the chemical shifts of only three out of eight backbone conformer types were completely separable from each other. However, care should be taken with this observation. Because it was noted earlier that the average calculated and average experimental amide proton chemical shifts disagree significantly, the results obtained using experimental data might differ slightly from those obtained for the *ab initio* $^{13}C^\alpha$-$^1H^\alpha$-$^1H^N$ 3D-HSQC-type plot. We hope to explore this topic in the future by the use of *ab initio* calculations including solvent simulation.

Other *ab initio* 3D-HSQC-type spectra (not shown) showed slightly to moderately worse chemical shift separation than that found in the $^{13}C^\alpha$-$^1H^\sigma$-$^1H^N$ plot.

Chemical Shifts of Serine as a Function of Its Backbone Conformation in Proteins

The experimental data support our *ab initio* findings. Analyzing resonances associated with serines in a set of selected proteins the average magnitude of the Pearson correlation coefficients, arising from correlations involving the ^{15}N, $^{13}C^\alpha$, $^{13}C^\beta$, $^{13}C'$, $^1H^N$, and $^1H^\alpha$ chemical shifts and ϕ, ψ, and χ^1 dihedral angles, is 0.293, while that of the R_{max} values is 0.544. Clearly, a moderate structure–chemical shift correlation is observed for the experimental data. This agrees well with the weak chemical shift-structure correlation found previously in our *ab initio* data.

The $^{13}C^\alpha/\phi$ and $^{13}C^\alpha/\psi$ correlations are somewhat stronger than the other chemical shift–structure correlations: their maximized Pearson coefficients are −0.704 and 0.745, respectively (Table 9). This observation is analogous with what we found in the *ab initio* data set, indicating that the $^{13}C^\alpha$ chemical shift of the Ser residue is significantly affected by its conformation. The $^{13}C^\alpha/\phi$ and $^{13}C^\alpha/\psi$ plots, Figure 7A and B, respectively, clearly illustrate this structural influence: β_L, δ_L, ε_L, and ε_D conformers have smaller (≈55–58 ppm), while the α_L conformers have larger (≈58–64 ppm) $^{13}C^\alpha$ chemical shifts than the random coil value of 58.48 ppm. Although the *ab initio* calculations indicated that the $^1H^\alpha/\phi$ correlation is significantly larger than the others, the experimental data support only a less significant correlation.

Most of the *ab initio* chemical shift–structure correlations are retrieved when the analysis is performed using our experimental data set. Some of the Pearson correlation coefficient values are different but the qualitative conclusions are in good agreement. The small discrepancies between the *ab initio* and experimental observations could result from a number of factors: (1) the relatively large error in the *ab initio* $^1H^\alpha$ chemical shifts; and (b) the incompleteness of the experimental structure database, namely that the experimental set of 79 serine residues, retrieved from 18 proteins, contains only five backbone types with a predominance of α_L, β_L, and ε_L, while the *ab initio* conformational library includes an almost complete representation of all backbone con-

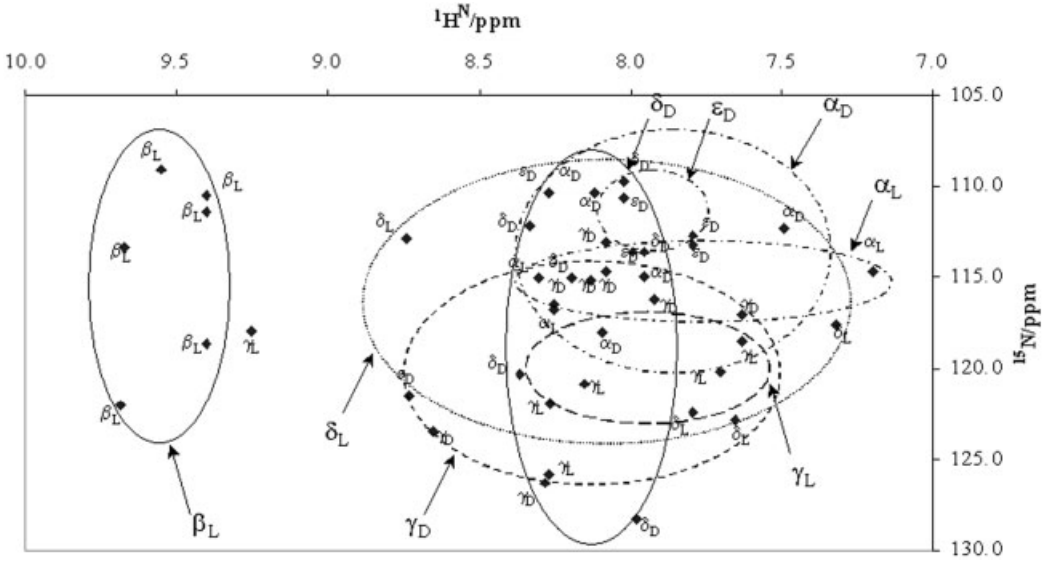

Figure 4. A heterocorrelation plot (e.g., 2D-HSQC) of calculated ^{15}N and $^1H^N$ chemical shifts (δ scale), of the 44 conformers of *trans*-For-L-Ser-NH_2, in which no obvious chemical shift–structure correlation can be seen. All chemical shifts have been determined at the GIAO-RHF/TZ2P level of theory. Chemical shift isotropies are referenced (δ scale) and corrected with the difference found between calculated (*ab initio*) and experimentally observed (BMRB) conformational shift values (correction factors for $^1H^N/^{15}N$: 3.69/3.54 ppm; see text for details).

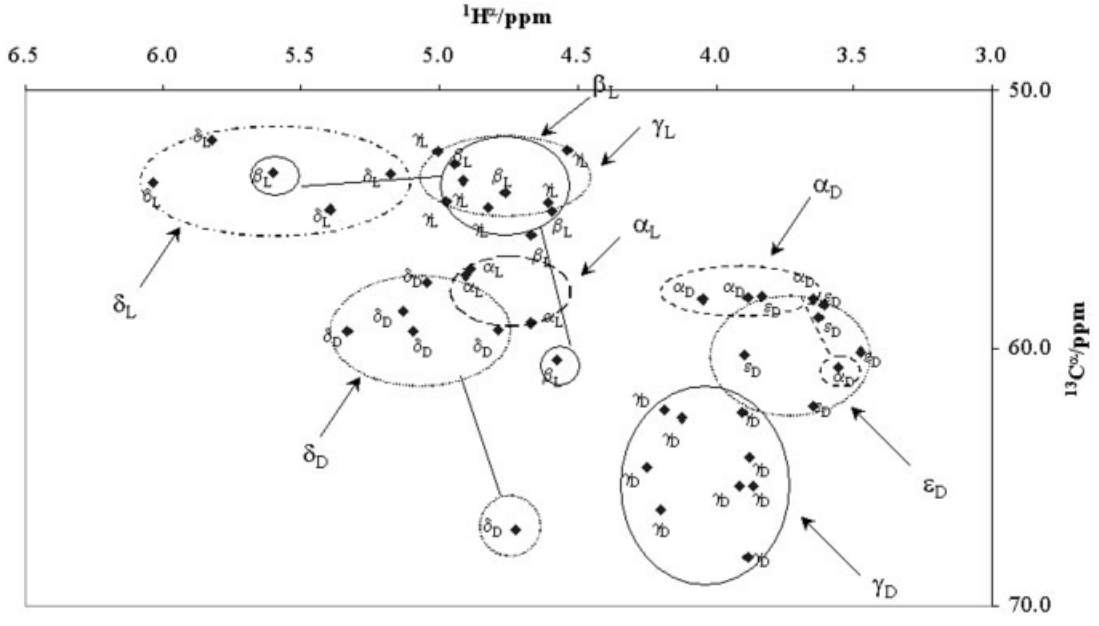

Figure 5. A heterocorrelation plot (e.g., 2D-HSQC) of calculated $^{13}C^\alpha$ and $^1H^\alpha$ chemical shifts (δ scale), of the 44 conformers of *trans*-For-L-Ser-NH_2, in which a significant chemical shift–structure correlation is observed. All chemical shifts have been determined at the GIAO-RHF/TZ2P level of theory. Chemical shift isotropies are referenced (δ scale) and corrected with the difference found between calculated (*ab initio*) and experimentally observed (BMRB) conformational shift values (correction factors for $^1H^\alpha/^{13}C^\alpha$: 0.98/4.96 ppm; see text for details).

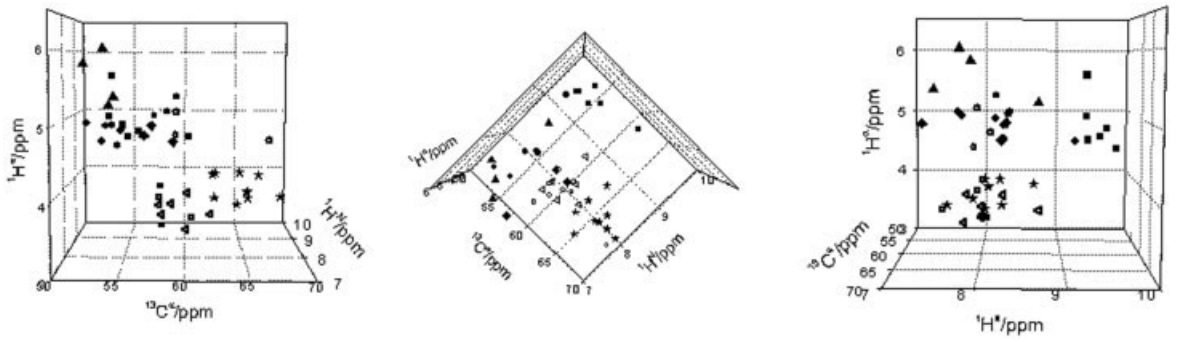

Figure 6. Three views of a H—C heterocorrelation plot of calculated $^{13}C^{\alpha}$, $^{1}H^{\alpha}$, and $^{1}H^{N}$ chemical shifts (δ scale) determined at the GIAO-RHF/TZ2P level for the 44 conformers of *trans*-For-L-Ser-NH_2. Using a suitable projection each backbone type can be distinguished [■ = β_L (β-pleated sheet), ◆ = α_L (α-helix), ★ = γ_D (γ-turn), ○ = γ_L (inverse γ-turn), ● = δ_D, □ = δ_L, ◁ = ε_D, and 7▲ = α_D]. Chemical shift isotropies are referenced (δ scale) and corrected with the difference found between calculated (*ab initio*) and experimentally observed (BMRB) conformational shift values (correction factors for $^{13}C^{\alpha}$/$^{1}H^{\alpha}$/$^{1}H^{N}$: 4.96/0.98/3.96 ppm; see text for details).

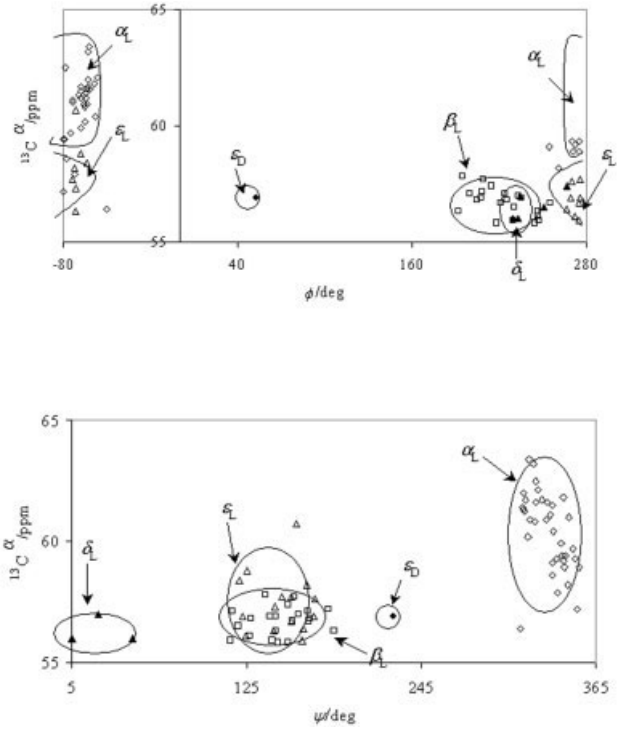

Figure 7. Illustration of chemical shift–structure correlations for 79 serines with chemical shift and conformational parameters extracted from the BMRB and PDB, respectively. In these images ϕ and ψ are not illustrated from −180° to 180°, as is customary, rather their ranges correspond to those used in achieving R_{max} (see text for details). (A) $^{13}C^{\alpha}/\phi$ and (B) $^{13}C^{\alpha}/\psi$ where □ = β_L (β-pleated sheet), ◇ = α_L (α-helix), ● = ε_D, ▲ = δ_L, △ = ε_L.

former types, only the backbone conformation type associated with the polyproline II (ε_L) form is missing. In the future we hope to assemble a significantly larger experimental database to test our results.

CONCLUSIONS

Currently the most common method of solving solution–phase structures of peptides and proteins by NMR spectroscopy involves generating distance information from NOEs. Due to the complex relationship between magnetization transfer and the NOE, precise interpretation of the distance information contained in NOEs may be difficult.

Previous experimental and theoretical chemical shift–structure studies suggest that the chemical shift is a conformationally sensitive parameter and could provide some structural information of high demand. A detailed search for chemical shift-structure relationships in the For-L-Ser-NH_2 dipeptide, employing the GIAO-RHF method, yielded promising results. A statistical analysis of $^{1}H^{\alpha}$ and $^{13}C^{\alpha}$ chemical shifts indicates that they are significantly influenced by the torsion angles ϕ and ψ and 2D and 3D HSQC-type plots appear useful in extracting conformational information from selected chemical shifts. Correlation studies using experimentally determined chemical shifts and conformational parameters of 79 serine residues support the computed results.

Although the α-helix and β-sheet information content of $^{1}H^{\alpha}$ and $^{13}C^{\alpha}$ chemical shifts has been known for some time and have now been incorporated in commercial software such as TALOS[59] and Dyana,[60,61] our results suggest that these chemical shifts can also be used to identify serine residues in other secondary structural elements (e.g., β-turns, loops) as well as in most typical backbone conformations. With these data we are convinced that easily obtainable chemical shift values (e.g., $^{1}H^{\alpha}$, $^{13}C^{\alpha}$) could be of great value for structure refinement of peptides and proteins.

Acknowledgments

This research was partially supported by grants from the Hungarian Scientific Research Fund (OTKA T033074, T030841, and T032486). A.K.F. thanks the Hungarian Academy of Sciences and the Hungarian Ministry of Education for a Domus Hungarica scholarship. The kind help of Zoltán Gáspári in the search of the Protein Data Bank is appreciated.

References

1. Wüthrich, K. NMR of Proteins and Nucleic Acids; John Wiley & Sons: New York, 1986.
2. Roberts, G. C. K. NMR of Macromolecules—A Practical Approach; Oxford University Press: New York, 1993.
3. Cavanagh, J.; Fairbrother, W. J. ; Palmer, A. G. III; Skelton, N. J. Protein NMR Spectroscopy—Principles and Practice; Academic Press: San Diego, 1996.
4. Karplus, M. J Phys Chem 1959, 30, 11.
5. Pardi, A.; Billeter, M.; Wüthrich, K. J Mol Biol 1984, 180, 741.
6. Prestegard, J. H.; Kishore, A. I. Curr Opin Chem Biol 2001, 5, 584.
7. Tian, F.; Valafar, H.; Prestegard, J. H. J Am Chem Soc 2001, 123, 11791.
8. Eglander, S. W.; Sosnick, T. R.; Englander, J. J.; Mayne, L. Curr Opin Struct Biol 1996, 6, 18.
9. Blandl, T.; Warder, S. E.; Prorok, M.; Castellino, F. J. FEBS Lett 2000, 470. 139.
10. Kobasyashi, T.; Ikeguchi, M.; Sugai, S. J Mol Biol, 2000, 299, 757.
11. Sternlicht, H.; Wilson, D. Biochemistry 1967, 6, 2881.
12. Markley, J. L.; Meadows, D. H.; Jardetzky, O. J Mol Biol 1967, 27, 25.
13. Tigelaar, H. L.; Flygare, W. H. J Am Chem Soc 1972, 94, 343.
14. Szliágyi, L.; Iardetzky, O. J Mag Reson 1989, 83, 441.
15. Spera, S.; Bax, A. J Am Chem Soc 1991, 113, 5490.
16. Wishart, D. S.; Sykes, B. D.; Richards, F. M. J Mol Biol 1991, 222, 311.
17. Reily, M. D.; Thanabal, V.; Omecinsky, D. O. J Am Chem Soc 1992, 114, 6251.
18. Wishart, D. S.; Sykes, B. D.; Richards, F. M. Biochemistry 1992, 31, 1647.
19. Pearson, J. G.; Le, H.; Sanders, L. K.; Godbout, N.; Havlin, R. H.; Oldfield, E. J Am Chem Soc 1997, 119, 11941.
20. Sun, H.; Sanders, L. K.; Oldfield, E. J Am Chem Soc 2002, 124, 5486.
21. Jiao, D.; Barfield, M.; Hruby, V. J. J Am Chem Soc 1993, 115, 10883.
22. Bühl, M.; Schleyer, P. V. R. J Am Chem Soc 1992, 114, 477.
23. de Dios, A. C.; Oldfield, E. J Am Chem Soc 1994, 116, 5307.
24. Le, H.; Pearson, J. G.; de Dios, A. C.; Oldfield, E. J Am Chem Soc 1995, 117, 3800.
25. Liu, F.; Phung, C. G.; Alderman, D. W.; Grant, D. M. J Am Chem Soc 1996, 118, 10629.
26. de Dios, A. C.; Pearson, J. G.; Oldfield, E. J Am Chem Soc 1993, 115, 9768.
27. Laws, D. D.; Le, H.; de Dios, A. C.; Havlin, R. H.; Oldfield, E. J Am Chem Soc 1995, 117, 9542.
28. Havlin, R. H.; Le, H.; Laws, D. D.; de Dios, A. C.; Oldfield, E. J Am Chem Soc 1997, 119, 11951.
29. Heller, J.; Laws, D. D.; Tomaselli, M.; King, D. S.; Wemmer, D. E.; Pines, A.; Havlin, R. H.; Oldfield, E. J Am Chem Soc 1997, 199, 7827.
30. Le, H.; Oldfield, E. J Phys Chem 1996, 100, 16423.
31. de Dios, A. C.; Pearson, J. G.; Oldfield, E. Science 1993, 260, 1491.
32. Perczel, A.; Császár, A. G. J Comp Chem 2000, 21, 882.
33. Perczel, A.; Császár, A. G. Chem Eur J 2001, 7, 1069.
34. Perczel, A.; Császár, A. G. J Eur Phys D 2002, 20, 513.
35. Oldfield, E. Annu Rev Phys Chem 2002, 53, 349.
36. Baldoni, H. A.; Rodriguez, A. M.; Zamora, M. A.; Zamarbide, G. N.; Enriz, R. D. J Mol Struct (Theochem) 1999, 465, 79.
37. Császár, A. G. Perczel, A. Prog Biophys Mol Biol 1999, 71, 243.
38. Czinki, E.; Császár, A. G. Perczel, A. Chem Eur J 2003, to appear.
39. Wolinski, K.; Hinton, J. F.; Pulay, P. J Am Chem Soc 1990, 112, 8251.
40. Ditchfield, R. Mol Phys 1974, 27, 789.
41. Hariharan, P. C.; Pople, J. A. Theor Chim Acta 1973, 28, 213.
42. Schafer, A.; Huber, R.; Ahlrichs, R. J Chem Phys 1994, 100, 5829.
43. Jákli, I.; Perczel, A.; Farkas, Ö.; Császár, A. G.; Sosa, C.; Csizmadia, I. G. J Comp Chem 2000, 21, 626.
44. 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.; Oritz, J. V.; Baboul, A. G.; Stefano, 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. Gaussian98; Revision A.5; Gaussian, Inc.: Pittsburgh, PA, 1998.
45. Császár, A. G.; Allen, W. D.; Schaefer, H. F, III. J Chem Phys 1998, 108, 9751.
46. Woon, D. E.; Dunning, T. H., Jr. J Chem Phys 1993, 98, 1358.
47. Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. J Chem Phys 1992, 96, 6796.
48. (a) Seavey, B. R.; Farr, E. A.; Westler, W. M.; Markley, J. L. J Biomol NMR 1991, 1, 217; (b) Markley, J. L.; Farr, E. A.; Seavey, B. R. J Biomol NMR 1991 1, 231; (c) The Biomolecular NMR Data Bank (BioMagResBank or BMRB) can be accessed at http://www.bmrb.wisc.edu/.
49. (a) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. Nucleic Acids Res 2000, 28, 235; (b) http://www.rcsb.org/pdb/.
50. Wonnacott, T. M.; Wonnacott, R. J. Introductory Statistics; Wiley: New York, 1990; 5th ed.
51. Csizmadia, I. Basic Principles for Introductory Organic Chemistry; Quirk Press: Toronto, 1997.
52. Perczel, A.; Ángyán, J. G.; Kajtár, M.; Viviani, W.; Rivail, J.-L; Csizmadia, I. G. J Am Chem Soc 1991, 113, 6256.
53. Perczel, A.; Farkas, Ö Császár, A. G.; Csizmadia, I. G. Can J Chem 1997, 75, 1120.
54. 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.
55. Hudáky, P.; Jákli, I.; Császár, A. G.; Perczel, A. J Comp Chem 2001, 22, 732.
56. Zamora, M. A.; Baldoni, H. A.; Bombasaro, J. A.; Mak, M. L.; Perczel, A.; Farkas, O.; Enriz, R. D. J Mol Struct (Theochem) 2001, 540, 271.
57. Baldoni, H. A.; Rodriguez, A. M.; Zamora, M. A.; Zamarbide, G. N.; Enriz, R. D.; Farkas, O.; Csaszar, P.; Torday, L. L.; Sosa, C. P.; Jakli, I.; Perzel, A.; Papp, J. G.; Hollosi, M.; Csizmadia, I. G. J Mol Struct (Theochem) 1999, 465, 79.
58. Aleman, C.; Puiggali, J. J Phys Chem B 1997, 101, 3441.
59. Cornilescu, G.; Delaglio, F.; Bax, A. J Biomol NMR 1999, 13, 289.
60. Güntert, P.; Mumenthaler, C.; Wührich, K. J Mol Biol 1997, 273, 283.
61. Gütert, P. Q Rev Biophys 1998, 31, 145.