Toward Direct Determination of Conformations of Protein Building Units from Multidimensional NMR Experiments I. A Theoretical Case Study of For-Gly-NH_2 and For-L-Ala-NH_2

ANDRÁS PERCZEL,[1] ATTILA G. CSÁSZÁR[2]

[1]*Department of Organic Chemistry, Eötvös University, P.O. Box 32, H-1518 Budapest 112, Hungary*
[2]*Department of Theoretical Chemistry, Eötvös University, P.O. Box 32, H-1518 Budapest 112, Hungary*

Received 29 November 1999; accepted 29 February 2000

ABSTRACT: NMR chemical shielding anisotropy tensors have been computed, employing several basis sets and the GIAO-RHF and GIAO-MP2 formalisms of electronic structure theory, for all the atoms of the five and nine typical backbone conformers of For-Gly-NH_2 and For-L-Ala-NH_2, respectively. Multidimensional chemical shift plots, as a function of the respective backbone fold, have been generated for both peptide models. On the 2D $^{1}H^{NH}$-$^{15}N^{NH}$ and $^{15}N^{NH}$-$^{13}C^{\alpha}$ plots the most notable feature is that at all levels of theory studied the backbone conformers cluster in different regions. Computed chemical shifts, as well as their averages, have been compared to relevant experimental values taken from the BioMagnetic Resonance Bank (BMRB). At the highest levels of theory, for all nuclei but the amide protons, deviations between statistically averaged theoretical and experimental shifts are as low as 1–3%. These results indicate that chemical shift information from selected multiple-pulse NMR experiments (e.g., 2D-HSQC and 3D-HNCA) could directly be employed to extract folding information for polypeptides and proteins.
J Comput Chem 21: 882–900, 2000

Keywords: protein building units; multidimensional NMR experiments; theoretical case study

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

Contract/grant sponsor: Hungarian Scientific Research Foundation; contract/grant numbers: OTKA T024044, T017192, and T017604

Contract/grant sponsor: Hungarian Academy of Sciences; contract/grant number: AKP96/2-427 2.4

Introduction

Traditionally, determination of the three-dimensional (3D) structure of peptides and small proteins from nuclear magnetic resonance (NMR) experiments is based on the assignment of proton (^{1}H) NMR chemical shifts followed by an analysis of the observed nuclear Overhauser effects (NOEs). The observed NOEs related to spatially close pairs of protons are converted into distances, and these proton–proton distances form the basis of structure analysis.[1–4] Backbone and side-chain dihedral angles can be usually determined using NOE-based constraints even if some of the distances are ambiguous.[4a] For a successful 3D NMR structure determination it is mandatory to assign all resonances. To achieve full assignment for unlabeled and singly labeled (^{15}N) proteins, *J*-correlated spectra, and NOESY-type information are essential. In contrast, if a doubly labeled (^{13}C and ^{15}N) protein is available, the full assignment can be achieved without NOEs by using specific 3D experiments, exploring only homo- and heteronuclear coupling constants.[4b–d] Such a strategy could, for example, incorporate the following set of heteronuclear experiments (see ref. 4e for the abbreviations): HNCA, HN(CO)CA, CBCA(CO)NH, HBHA(CO)NH, CC(CO)NH, HCC(CO)NH, and HCCH-TOCSY. At present, however, even if all the assignments are correctly determined from *J*-correlated spectroscopic data, the 3D structure of the molecule cannot be determined without the analysis of the information derived from NOESY-type spectra.

Chemical shielding of a nucleus, located in different proteins or at different sites within the same protein, changes, due either to the individual molecular environment within the macromolecule or to differences in backbone orientations. If the latter factor is dominant, the 3D structure of a protein could be revealed using chemical shift information alone. In the above-mentioned triple-resonance NMR experiments chemical shifts can be resolved since information is spread out, for example, on the ^{1}H-^{15}N, ^{1}H-^{1}H, ^{1}H-^{13}C, and ^{15}N-^{13}C planes. Detailed experimental studies[5–9] have clearly established a few structure-induced $^{13}C^{\alpha}$, $^{15}N^{NH}$, and $^1H^N$ chemical shift changes in peptides and proteins (*vide infra*), providing examples of correlation of backbone folds of peptides and proteins with NMR chemical shifts. These results do provide hope that the direct analysis of NMR chemical shifts from relevant multiple-pulse experiments (e.g., two-dimensional heteronuclear multiple quantum coherence (2D-HMQC),[10] two-dimensional heteronuclear single quantum coherence (2D-HSQC),[11] and 3D-HNCA[12]) may prove to be a plausible alternative to the distance-based (NOE) strategy for elucidation of the dihedral space of protein structures. The major limitation of this approach, at least at present, is that an unambiguous empirical correlation between backbone conformation and NMR chemical shifts has been put forward only for the α-helical and β-sheet regions of the Ramachandran surface.[13, 14] Due to the lack of appropriate experimental data sets, theoretical (*ab initio*) calculation of conformation-dependent NMR shieldings seems to offer the only hope to establish and probe such relations.[13, 15]

The ^{33}S chemical shieldings in inorganic and small organic molecules,[16] the ^{13}C NMR properties[17] of Thr and Tyr amino acids in zwitterions, and the ^{19}F shifts of various fluorinated tryptophan-containing peptides and proteins[18] have been determined by *ab initio* techniques. Nevertheless, the primary goal of NMR computations on peptides and proteins has been the description of chemical shielding anisotropy (CSA) tensors and associated chemical shifts of the ^{13}C and ^{15}N nuclei.[13, 19–24]

Jiao and coworkers[19] were among the first who determined the ^{13}C CSA tensors of For-Gly-NH_2 as a function of the backbone conformation. The C^{α} chemical shifts were calculated over the Ramachandran surface using a grid of 30°. It was found that $^{13}C^{\alpha}$ values in the helical and in the extended regions are shifted by 2.3 ppm downfield and 2.9 ppm upfield, respectively, compared to the random coil value. These shifts should be compared with available experimental results,[19] about +3.2 ppm and −1.2 ppm, respectively. For the alanine residue, Heller and coworkers[23] investigated ways to use C^{α} CSA tensors for the determination of dihedral angles. Using crosspolarization magic-angle-spinning NMR, the C^{α} values of tripeptides of known structure were measured. The obtained experimental shifts were correlated with *ab initio* chemical shifts. Using the gauge-including atomic orbital restricted Hartree–Fock (GIAO-RHF) method, the ^{13}C CSA tensors of the For-L-Ala-NH_2[13c, 20] and For-L-Val-NH_2[13c] model compounds were determined, once again as a function of their main-chain fold. It has been concluded that (a) substitution on C^{α} consistently induces larger shielding shift on C^{β} than on C^{α};[13c] and (b) the $[\phi, \psi]$ and χ_1 conformational parameters have a significant influence on chemical shifts. Focusing on Ala and Val residues, Pearson et al.[21] have found that in a protein such as nuclease the $[\phi, \psi]$ values can be estimated using

chemical shifts. Although the results obtained are less dependable than those derived from NOEs and *J*-coupling constraints, the strategy seemed very promising for the estimation of secondary structures. Analyzing chemical shift nonequivalencies of Ala and Val residues in proteins like calmodulin and nuclease,[13c] the applicability of the direct strategy was further explored.[13c] Prompted by these results, ^{13}C CSA tensor computations were extended[22] to cover additional amino acid residues Val, Ile, Ser, and Thr. For all residues studied, the $^{13}C^{\alpha}$ shift showed the expected ≈5 ppm increase for the β conformation over the helical structure. Furthermore, the diagonal CSA tensor elements were all found to be sensitive to changes in the ϕ, ψ, and χ_1 torsion angles. In summary, all calculations indicated that it is possible to deduce certain backbone and side-chain orientations from $^{13}C^{\alpha}$ shielding tensors.

As well known from different ^{15}N-X-type correlated NMR experiments (e.g., $^{15}N^{NH}$-$^{1}H^{NH}$ HSQC), $^{15}N^{NH}$ chemical shifts are rather sensitive probes of protein main-chain fold, and have considerable potential for structure determination. Nevertheless, due to their fairly complex dependence on several torsion angles and on electrostatic field effects,[24] interpretation of changes in $^{15}N^{NH}$ shifts proved to be somewhat more complicated than that of changes in ^{13}C shifts. In short, while the most important structural factors determining the $^{15}N^{NH}$ NMR chemical shift of amino acid residue *i* appear to be the dihedral angles $\psi^{(i-1)}$ and $\phi^{(i)}$, the angles $\psi^{(i)}$ and $\phi^{(i-1)}$ and the side-chain orientation also have a nonnegligible effect. Consequently, at present, $^{15}N^{NH}$ shifts are less readily useful than ^{13}C shifts.

H-bonds have a crucial role in the formation and stabilization of peptides and proteins. For example, Asakawa and coworkers[25] have found that the C^{α} shielding values are affected not only by the values of the $[\phi, \psi]$ torsional parameters but also by H-bond patterns (standard, bifurcated, etc.) and by H-bond strength. This observation proved to be crucial during determination of the 3D structure of ribonuclease and that of a basic pancreatic trypsin inhibitor.[25] GIAO-RHF calculations also established the influence of H-bonding on the carbonyl carbon of N-methylacetamide interacting with formamide.[26] Further computations,[26] at the 6-31G** GIAO-RHF level, have been performed on the helical and β-sheet structures of For-$(Ala)_5$-NH_2 in order to investigate the effect of intramolecular H-bonds on shieldings of different nuclei, for example, $^{13}C'$. The experimentally determined difference between the helical and the β-sheet values of $^{13}C'$, 4.6 ppm, is close to the calculated value, 4.9 ppm. The most significant perturbation caused by the H-bond was on carbonyl carbons.[27]

When performing any type of *ab initio* computation on any property of a chemical system, issues related to the applied theoretical level (*n*-particle space) and basis set (one-particle space) are to be considered. Laws and coworkers[28] have determined, at the GIAO-RHF level, the effect of basis set extension, employing standard basis sets STO-3G, 3-21G, 4-31G, and 6-311+G(2d), on $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ shifts of two amino acid residues, Ala and Val, in proteins. They found that the chemical shifts of these residues, determined using smaller basis sets, correlate rather well with results obtained with larger basis sets, even when the underlying reference geometries are slightly different. It is encouraging that chemical shifts computed at different levels can be scaled and fine tuned. In a further study of relevance, employing empirical chemical shift surfaces, an attempt was made to predict ^{13}C shift values of valine residues in three well-known proteins (calmoduline, nuclease, and ubiquitine)[29] utilizing their X-ray structures. Most HF results and experimental values showed agreement[29] poorer than expected, but the agreement improved slightly by using density functional theory (DFT)-type calculations. Pearson and coworkers[29] have thus concluded that for accurate computation of chemical shifts geometry optimization and the inclusion of electron correlation in the theoretical treatment appear to be important.

In this report a concerted attempt is made to correlate calculated isotropic NMR shielding and chemical shift values with all characteristic backbone conformations[30] of peptide models, concentrating not only on ^{13}C but also on ^{15}N and ^{1}H. The effect of electron correlation is explicitly incorporated in CSA tensor calculations at the GIAO-MP2 level. To keep the computational problem tractable, systematic calculations are carried out only for conformations of the following amino acid diamides of glycine and alanine: For-Gly-NH_2 and For-L-Ala-NH_2. It is not obvious what is the best use of the large number of quantum chemical data obtained. We felt that apart from investigating traditional issues, such as basis set dependence and the effect of electron correlation on the calculated results, a thorough statistical analysis offers the best insight. Therefore, relevant correlations are investigated in considerable detail, confirming some well-known structure–chemical shift relations and establishing, at the same time, new and potentially useful ones. Although detailed comparison of the computed results with experiment is not yet possible, an attempt

is made to compare direct and statistically averaged chemical shifts from the two sources.

Computational Details

Computation of NMR shielding tensors utilized the Gaussian94,[31] Gaussian98,[32] and AcesII[33] program systems, and were performed at the GIAO-RHF (gauge including atomic orbitals restricted Hartree–Fock)[34] and GIAO-MP2 (GIAO second-order Møller–Plesset perturbation theory)[35] levels of theory employing basis sets of 6-31+G*,[36a] 6-311++G**,[36b] DZ(d),[37] and TZ2P[38] quality.

The For-Gly-NH_2 and For-L-Ala-NH_2 models are depicted on Figures 1 and 2, respectively. These figures also contain information concerning average chemical shifts of all nuclei forming the amino acid core, as taken from the BioMagnetic Resonance Bank (BMRB).[39]

The reference geometries employed for the chemical shielding anisotropy (CSA) computations have been determined at the 6-31++G** RHF and 6-311++G** DFT(B3LYP) levels. The latter level was employed extensively for the geometry optimizations, as it proved to be especially reliable during structural studies of the neutral amino acids glycine[40a, 40b] and α-alanine.[40c] To obtain a complete chemical shielding set associated with the five and nine unique backbone orientations of For-Gly-NH_2 and For-L-Ala-NH_2, respectively, additional 6-31++G** RHF constrained geometry optimizations have been performed keeping the characteristic ϕ and ψ torsional angles constant. These fixed ϕ and ψ values were obtained from averaging the appropriate 3-21G RHF backbone conformational parameters of For-L-Ala-L-Ala-NH_2.[44] Description of Levels A–F, resulting from different combinations of full/constrained geometry optimizations and CSA calculations, is given in Table I.

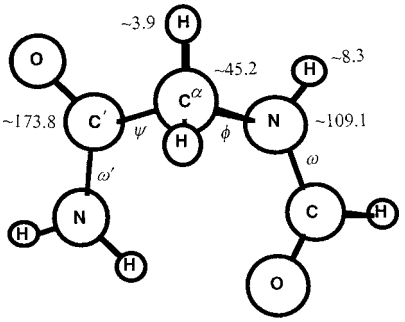

FIGURE 1. The For-Gly-NH_2 model and average chemical shifts (in ppm) taken from BMRB.

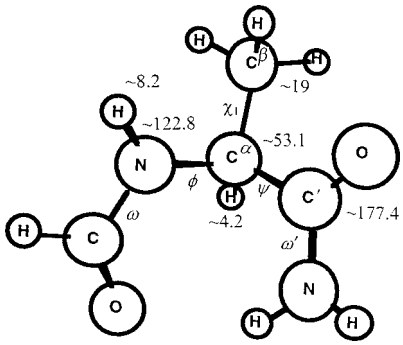

FIGURE 2. The For-L-Ala-NH_2 model and average chemical shifts (in ppm) taken from BMRB.

Selected geometric parameters (including optimized/fixed ϕ and ψ values), energies, and isotropic chemical shielding values (σ-scale) are reported for For-Gly-NH_2 and For-L-Ala-NH_2 in Tables II–VI. When relative chemical shifts (δ-scale) are used, the appropriate isotropic chemical shielding values of ^1H, ^{13}C, and ^{15}N are referenced to ^1H and ^{13}C of tetramethyl-silane (TMS) and to ^{15}N of NH_3. The reference geometry chosen for NH_3 corresponds to the aug-cc-pVTZ CCSD(T) optimized geometry,[42] while the geometry of TMS has been optimized at the 6-311++G** B3LYP level (cf. Tables V and VI).

Average (or random) chemical shift data and associated standard deviations for the amino acid residues Gly and Ala were taken from data deposited in BMRB[39] (cf. Figs. 1 and 2 and Tables V and VI). Experimental chemical shifts associated with each of the 9(5) typical backbone conformers of For-L-Ala-NH_2(For-Gly-NH_2) were determined by a comprehensive analysis of data deposited in BMRB and the Protein Data Bank (PDB).[43]

Structures and Energetics

The usual way to characterize the backbone conformation of peptides (or proteins) is through the $[\phi, \psi]$ torsional angle pairs associated with the amino acid residues.[30] The torsional potentials along both ϕ and ψ should have three minima; therefore, in this simplest picture we expect nine legitimate conformers for each "peptide unit" (...NH-CHR-CO...).[30, 41, 44] These nine conformers are named, according to our established convention,[30c] as follows: α_L, α_D, β_L, γ_L, γ_D, δ_L, δ_D, ε_L, and ε_D (see Figs. 3 and 4 for the Ramachandran surfaces of For-Gly-NH_2 and For-L-Ala-NH_2, respectively). The lack of configurational chirality of

TABLE I.
Designation of Different Combinations of Geometry Optimization and Chemical Shielding Calculations for the For-Gly-NH_2 and For-L-Ala-NH_2 Model Systems.

Model System	Level	NMR Computation	Geometry Optimization	No. of Structures	Comment
For-Gly-NH_2	A	GIAO-RHF/DZ(d)	B3LYP/6-311++G**	5	full opt. plus [ϕ, ψ] constr.
	B	GIAO-MP2/DZ(d)	B3LYP/6-311++G**	5	full opt. plus [ϕ, ψ] constr.
	C	GIAO-RHF/TZ2P	B3LYP/6-311++G**	5	full opt. plus [ϕ, ψ] constr.
For-L-Ala-NH_2	A	GIAO-RHF/6-31+G*	RHF/6-31++G**	9	[ϕ, ψ] constr.
	B	GIAO-RHF/6-31+G*	B3LYP/6-311++G**	6	full opt. only
	C	GIAO-RHF/6-31+G*	RHF/6-31++G**	6	full opt. only
	D	GIAO-RHF/6-311++G**	RHF/6-31++G**	9	[ϕ, ψ] constr.
	E	GIAO-RHF/6-311++G**	B3LYP/6-311++G**	6	full opt. only
	F	GIAO-RHF/TZ2P	B3LYP/6-311++G**	6	full opt. only
	G	GIAO-RHF/TZ2P	B3LYP/6-311++G**	10	full opt. and [ϕ, ψ] constr.

glycine makes it unique among the natural alpha amino acids resulting in a double degeneracy on the Ramachandran PES. In our notation this means that for For-Gly-NH_2 $\alpha_L = \alpha_D = \alpha_{LD}$, $\beta_L = \beta_{LD}$, $\gamma_L = \gamma_D = \gamma_{LD}$, $\varepsilon_L = \varepsilon_D = \varepsilon_{LD}$, and $\delta_L = \delta_D = \delta_{LD}$. Of course, if the glycine residue is embedded in a chiral environment (i.e., in a protein chain), the degeneracy of the conformer pairs of our For-Gly-NH_2 model disappears.

Although classification of hundreds of X-ray determined protein structures revealed the existence of all nine conformer types, systematic *ab initio* calculations[30, 45] carried out on simple amino acid diamides (For-*X*-NH_2, with X = Gly, Ala, Val, Ser,

TABLE II.
Conformational Parameters (Torsion Angles ϕ and ψ in Degrees), Total Energies (E/E_h), Relative Energies [ΔE/(kcal mol^{-1})], and Relative Populations of the Characteristic Conformers of For-Gly-NH_2 and For-L-Ala-NH_2.[a]

Model	Backbone Conformation	ϕ	ψ	E/E_h	ΔE/(kcal mol^{-1})	$p(i)/\sum p(i)$
For-Gly-NH_2	α_{LD}	−68.6[b]	−17.5[b]	−378.027433	3.21	0.002
	β_{LD}	180.00	180.00	−378.032393	0.10	0.444
	γ_{LD}	−81.36	65.78	−378.032551	0.00	0.524
	δ_{LD}	113.18	−18.26	−378.029586	1.86	0.023
	ε_{LD}	−74.7[c]	167.8[c]	−378.028406	2.60	0.007
For-L-Ala-NH_2	α_D	73.4[b]	10.5[b]	−417.352003	5.00	0.000
	α_L	−68.6[b]	−17.5[b]	−417.354165	3.65	0.002
	β_L	−156.59	163.69	−417.358625	0.85	0.185
	γ_D	72.48	−55.06	−417.356591	2.12	0.022
	γ_L	−82.60	73.44	−417.359976	0.00	0.770
	δ_D	−168.06	−42.03	−417.350367	6.03	0.000
	δ_L	−114.02	15.76	−417.356082	2.44	0.013
	ε_D	60[d]	−120[d]	−417.351456	5.35	0.000
	ε_L	−60[d]	120[d]	−417.355659	2.71	0.008

[a] All results reported, unless otherwise noted, correspond to the 6-311++G** B3LYP level. Relative populations, $p(i)/\sum p(i)$, for computed *ab initio* relative energies are determined as $\exp(-\Delta E/RT)/\sum\exp(-\Delta E/RT)$, where $RT = NkT = 0.595371$ kcal mol^{-1}, $T = 300$ K, $k = 1.38 \times 10^{-23}$ J/K, and Avogadro's number (N) is 6.02×10^{23} mol^{-1}.

[b] Constrained backbone parameters of helical conformers found for For-L-Ala-L-Ala-NH_2.[41]

[c] Constrained backbone parameters of polyproline II conformers found for For-L-Ala-L-Ala-NH_2.[41]

[d] For symmetry reasons these idealized backbone values were used, which are characteristic for polyproline II.

TABLE III.

Backbone Conformational Parameters and Isotropic Chemical Shielding Values (σ-Scale), Relative to the γ_{LD} Conformation, of For-Gly-NH_2.[a]

Backbone Conformation	Level[b]	ω_0	ϕ	ψ	ω_1	$^{15}N^{H}$	$^{13}C^{\alpha}$	$^{13}C'$	$^{1}H^{N}$	$^{1}H^{\alpha}$ (Average)
α_{LD}	A	−175.4	−68.6	−17.5	173.1	11.40	3.68	4.86	1.09	0.32
	B	−175.4	−68.6	−17.5	173.1	11.32	4.85	3.13	1.13	0.35
	C	−175.4	−68.6	−17.5	173.1	4.35	1.00	−0.15	0.35	−0.01
β_{LD}	A	180.0	180.0	180.0	180.0	7.97	4.73	0.37	−0.67	0.04
	B	180.0	180.0	180.0	180.0	8.67	6.42	0.87	−0.18	−0.01
	C	180.0	180.0	180.0	180.0	10.03	5.02	0.00	−1.02	−0.08
γ_{LD}	A	−177.6	−81.4	65.8	−178.2	0.00	0.00	0.00	0.00	0.00
	B	−177.6	−81.4	65.8	−178.2	0.00	0.00	0.00	0.00	0.00
	C	−177.6	−81.4	65.8	−178.2	0.00	0.00	0.00	0.00	0.00
δ_{LD}	A	173.5	113.2	−18.3	−172.7	1.67	3.54	−0.12	0.63	−0.27
	B	173.5	113.2	−18.3	−172.7	2.13	5.46	0.45	0.62	−0.19
	C	173.5	113.2	−18.3	−172.7	1.81	3.66	−0.25	0.64	−0.40
ε_{LD}	A	167.1	−74.7	167.8	−175.3	13.28	6.08	4.07	1.00	0.53
	B	167.1	−74.7	167.8	−175.3	13.94	8.01	1.96	1.07	0.51
	C	167.1	−74.7	167.8	−175.3	7.45	3.91	−1.40	0.40	0.09

[a] All relative isotropic shielding values (in ppm) are referenced to the absolute isotropic shieldings of the γ_{LD} conformer, which are as follows: Level A: $^{15}N^{NH} = 164.7$, $^{13}C^{\alpha} = 155.5$, $^{13}C' = 26.9$, $^{1}H^{NH} = 28.5$, $^{1}H^{\alpha} = 29.2$; Level B: $^{15}N^{NH} = 168.3$, $^{13}C^{\alpha} = 157.3$, $^{13}C' = 50.1$, $^{1}H^{NH} = 27.7$, $^{1}H^{\alpha} = 28.6$; Level C: $^{15}N^{NH} = 151.8$, $^{13}C^{\alpha} = 149.2$, $^{13}C' = 14.5$, $^{1}H^{NH} = 27.0$, $^{1}H^{\alpha} = 28.5$.

[b] See Table I for the designation of the *ab initio* levels employed.

TABLE IV.

Backbone Conformational Parameters and Isotropic Chemical Shielding Values (σ-Scale), Relative to the γ_{L} Conformation, of For-L-Ala-NH_2.[a]

Backbone Conformation	Level[b]	ω_0	ϕ	ψ	ω_1	χ_1	$^{15}N^{NH}$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{NH}$	$^{1}H^{\alpha}$
α_{D}	A	171.9	61.8	31.9	−173.9	64.2	8.32	−5.27	−0.28	0.06	−0.12	0.89
	B	171.8	73.4	10.5	−173.6	61.9	8.51	−6.49	0.03	−0.62	−0.03	0.88
	C	169.5	69.9	24.6	−174.7	64.8	8.37	−5.80	0.42	−0.04	−0.02	0.91
	D	171.9	61.8	31.9	−173.9	64.2	9.31	−5.74	−0.36	−0.19	−0.17	0.94
	E	171.8	73.4	10.5	−173.6	61.9	9.56	−6.91	−0.15	−1.02	−0.10	0.90
	F	171.8	73.4	10.5	−173.6	61.9	10.02	−6.87	0.08	−1.34	−0.03	0.97
α_{L}	A	−175.3	−68.6	−17.5	172.9	61.7	3.97	−4.11	−1.62	−1.63	0.17	0.17
	D	−175.3	−68.6	−17.5	172.9	61.7	4.14	−4.14	−1.98	−1.83	0.08	0.17
3_{10}-helix like[c]	F	−176.3	−68.6	−17.5	172.7	−59.9	4.68	−3.74	−2.03	−1.79	0.21	0.11
α-helix-like	F	−175.7	−54.0	−45.0	173.1	61.1	6.25	−5.50	−2.27	−1.83	0.32	0.25
β_{L}	A	−179.6	−167.6	169.9	−179.1	62.4	10.26	−1.53	−3.07	−0.81	−1.45	0.20
	B	177.2	−156.6	163.7	178.7	61.8	9.81	−0.51	−4.42	−1.04	−1.49	0.10
	C	177.9	−155.6	160.4	178.7	61.4	8.81	−1.30	−3.98	−1.00	−1.29	0.11
	D	−179.6	−167.6	169.9	−179.1	62.4	10.91	−1.65	−3.41	−0.87	−1.63	0.17
	E	177.2	−156.6	163.7	178.7	61.8	10.60	−0.54	−4.74	−1.45	−1.65	0.05
	F	177.2	−156.6	163.7	178.7	61.8	10.75	−0.41	−4.55	−1.81	−1.69	0.07

TABLE IV.
(Continued)

Backbone Conformation	Level[b]	ω_0	ϕ	ψ	ω_1	χ_1	$^{15}N^{NH}$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C'$	$^{1}H^{NH}$	$^{1}H^{\alpha}$
γ_D	A	177.6	74.3	−59.5	−179.8	63.4	7.30	−10.59	−0.21	−0.91	−0.30	0.73
	B	175.8	72.5	−55.1	−178.4	62.7	5.94	−10.35	−0.87	−0.83	−0.26	0.73
	C	176.4	75.2	−55.0	−177.8	63.7	6.44	−10.67	−0.14	−1.18	−0.23	0.74
	D	177.6	74.3	−59.5	−179.8	63.4	7.98	−11.25	−0.80	−1.14	−0.37	0.80
	E	175.8	72.5	−55.1	−178.4	62.7	6.68	−11.02	−1.47	−1.30	−0.32	0.78
	F	175.8	72.5	−55.1	−178.4	62.7	6.31	−11.14	−1.35	−1.20	−0.30	0.83
γ_L	A	−178.0	−84.5	68.7	−176.9	62.1	0.00	0.00	0.00	0.00	0.00	0.00
	B	−177.5	−82.6	73.4	−173.8	63.8	0.00	0.00	0.00	0.00	0.00	0.00
	C	−177.8	−86.2	76.5	−173.6	62.1	0.00	0.00	0.00	0.00	0.00	0.00
	D	−178.0	−84.5	68.7	−176.9	62.1	0.00	0.00	0.00	0.00	0.00	0.00
	E	−177.5	−82.6	73.4	−173.8	63.8	0.00	0.00	0.00	0.00	0.00	0.00
	F	−177.5	−82.6	73.4	−173.8	63.8	0.00	0.00	0.00	0.00	0.00	0.00
δ_D	A	173.5	−179.6	−43.7	−170.1	55.4	9.65	−8.08	−1.85	−1.32	0.19	0.12
	B	172.3	−168.1	−42.0	−170.9	58.7	10.39	−6.66	−2.57	−1.05	0.14	−0.01
	C	171.4	−165.9	−41.9	−170.1	58.2	9.70	−7.38	−2.20	−1.33	0.09	0.04
	D	173.5	−179.6	−43.7	−170.1	55.4	10.75	−8.72	−2.55	−1.56	0.18	0.14
	E	172.3	−168.1	−42.0	−170.9	58.7	11.39	−7.32	−3.17	−1.52	0.12	−0.02
	F	172.3	−168.1	−42.0	−170.9	58.7	11.29	−7.30	−3.15	−1.70	0.08	−0.04
δ_L	A	−170.5	−126.2	26.5	171.0	61.0	2.22	−0.30	−2.52	−0.66	0.36	−0.54
	B	−171.5	−114.0	15.8	171.8	59.4	2.56	−0.31	−2.57	−0.54	0.46	−0.50
	C	−171.0	−111.8	13.5	171.2	60.6	2.27	−0.94	−2.45	−1.26	0.43	−0.41
	D	−170.5	−126.2	26.5	171.0	61.0	2.46	−0.41	−2.56	−0.90	0.32	−0.62
	E	−171.5	−114.0	15.8	171.8	59.4	2.73	−0.43	−2.69	−0.69	0.40	−0.60
	F	−171.5	−114.0	15.8	171.8	59.4	2.77	−0.40	−2.70	−0.97	0.49	−0.58
ε_D	A	−163.9	64.7	−178.6	176.2	73.1	12.65	−5.48	−1.44	0.22	0.20	0.93
	D	−163.9	64.7	−178.6	176.2	73.1	13.91	−6.18	−1.73	0.04	0.13	1.00
	F	−173.2	60.0	−120.0	172.3	60.3	8.11	−5.56	−0.70	−3.17	−0.56	1.02
ε_L	A	166.9	−74.7	167.8	−174.0	62.1	6.73	−1.36	−2.31	−2.50	0.04	0.02
	D	166.9	−74.7	167.8	−174.0	62.1	7.35	−1.36	−2.84	−2.81	0.01	0.00
	F	171.2	−60.0	120.0	−173.1	64.0	3.48	−2.34	−0.78	−5.54	0.02	0.48

[a] All relative isotropic shielding values are referenced to the absolute isotropic shielding of the $\gamma_L(g+)$ conformer calculated as: Level A: $^{15}N^{NH} = 158.7$, $^{13}C^{\alpha} = 159.4$, $^{13}C^{\beta} = 185.9$, $^{13}C' = 33.5$, $^{1}H^{NH} = 28.9$, $^{1}H^{\alpha} = 28.8$; Level B: $^{15}N^{NH} = 151.6$, $^{13}C^{\alpha} = 156.4$, $^{13}C^{\beta} = 184.9$, $^{13}C' = 28.3$, $^{1}H^{NH} = 28.3$, $^{1}H^{\alpha} = 28.5$; Level C: $^{15}N^{NH} = 159.2$, $^{13}C^{\alpha} = 159.5$, $^{13}C^{\beta} = 185.9$, $^{13}C' = 33.4$, $^{1}H^{NH} = 28.9$, $^{1}H^{\alpha} = 28.8$; Level D: $^{15}N^{NH} = 141.5$, $^{13}C^{\alpha} = 151.9$, $^{13}C^{\beta} = 179.4$, $^{13}C' = 17.9$, $^{1}H^{NH} = 28.3$, $^{1}H^{\alpha} = 28.5$; Level E: $^{15}N^{NH} = 134.1$, $^{13}C^{\alpha} = 148.8$, $^{13}C^{\beta} = 178.3$, $^{13}C' = 12.8$, $^{1}H^{NH} = 27.7$, $^{1}H^{\alpha} = 28.2$; Level F: $^{15}N^{NH} = 136.1$, $^{13}C^{\alpha} = 148.6$, $^{13}C^{\beta} = 178.1$, $^{13}C' = 12.5$, $^{1}H^{NH} = 27.4$, $^{1}H^{\alpha} = 28.0$.

[b] See Table I for the designation of the *ab initio* levels employed.

and Phe) resulted in fewer minima than the maximum number allowed. For example, at the 3-21G(6-311++G**) RHF level two(three) conformers could not be located for For-L-Ala-NH_2. The results of these systematic calculations[30, 44, 45] show that the α_L, ε_L, and ε_D conformers are found to be absent most often. As far as relative energies are concerned (see Table II), the computational results reveal that the α_L and ε_L conformations have high relative energies compared to the usually most stable γ_L conformation. Nevertheless, existence of the "missing" minima has been verified in other amino acid diamides (e.g., For-L-Ser-NH_2)[44a] as well as in larger peptide models (e.g., For-L-Ala-L-Ala-NH_2).[41]

It is not the principal purpose of this article to discuss structural and energetic features of the two investigated model compounds at considerable length. Nevertheless, it is worth discussing the re-

TABLE V.
Chemical Shifts (δ-Scale) Averaged over All Backbone Conformers of For-Gly-NH_2.[a,b]

Level[c]	Average Shifts (ppm)					Error (%)[d]				
	$^{15}N^{NH}_{avr}$	$^{13}C^{\alpha}_{avr}$	$^{13}C'_{avr}$	$^{1}H^{NH}_{avr}$	$^{1}H^{\alpha}_{avr}$	N	C^{α}	C′	NH	H^{α}
Arithmetic Average										
A	90.97	40.49	170.84	3.88	3.50	17	10	2	53	10
B	105.34	44.06	154.86	3.83	3.35	3	3	11	54	14
C	112.21	42.78	180.50	5.20	3.80	3	5	4	28	3
Boltzmann Average[f]										
A	88.35	39.86	169.91	4.14	3.45	19	12	2	50	11
B	108.53	45.96	155.73	4.42	3.48	1	2	10	47	11
C	112.38	43.15	180.16	5.71	3.76	3	5	4	32	3
Expt.	**109.1**	**45.24**	**173.84**	**8.33**	**3.90**					

[a] Absolute chemical shifts (in ppm) used as references are given as follows: Level A: ^{1}H (TMS) = 32.800, ^{13}C (TMS) = 199.590, ^{15}N (NH_3) = 262.490; Level B: ^{1}H (TMS) = 32.070, ^{13}C (TMS) = 206.265, $^{15}N(NH_3)$ = 280.869; Level C: ^{1}H (TMS) = 32.231, ^{13}C (TMS) = 194.688, ^{15}N (NH_3) = 268.698.

[b] The number of backbone conformers is 5.

[c] See Table I for the designation of the *ab initio* levels used. Expt. = average experimental shift taken form BMRB.[39]

[d] Error = $abs(CSA^{exp} - CSA^{calc})/CSA^{exp} * 100$.

[e] Standard deviation of the average shift.

[f] Boltzmann averaging over all conformers was achieved with the use of the corresponding Level A–F energies of the GIAO calculations.

sults obtained and compare them with previously determined values.[46]

The ϕ and ψ torsional angles do not change much among the model compounds For-Gly-NH_2, For-L-Ala-NH_2, and For-L-Ala-L-Ala-NH_2. For example, the constrained ϕ and ψ torsion angles taken[41] from For-L-Ala-L-Ala-NH_2 never deviate more than 15° from the optimized ones (cf. Table II). The structural characteristics, when appropriate, are similar to those observed for the parent amino acids, glycine[40a] and α-alanine.[40c]

The relative energies of the conformers presented in Table II reveal that the energy order of the L-type conformers (cf. Figs. 3 and 4) is identical for both models, regardless of whether electron correlation is incorporated into the computation or not (relevant energy data, not detailed here, can be obtained from the authors upon request). The observed energy order of the L-type conformers is as follows: $E(\gamma_L) < E(\beta_L) < E(\delta_L) < E(\varepsilon_L) < E(\alpha_L)$. Typically, the D-type backbone conformers of L-amino acid diamides have a lower stability. The most stable D-type conformer, γ_D of For-L-Ala-NH_2, has a low relative energy due to a strong intramolecular H-bond, which makes its relative energy similar to that of δ_L. The δ_D conformer, though a minimum on the PES of For-L-Ala-NH_2, has a very high relative energy. The α_D conformer of For-L-Ala-NH_2, a minimum at the 6-311++G** RHF level, is not a stationary point at the 6-311++G** DFT(B3LYP) level. For For-Gly-NH_2, the RHF energy difference between the γ_{LD} and β_{LD} conformers is typically 1 kcal mol^{-1}, favoring γ_{LD}. However, electron correlation stabilizes β_{LD} over γ_{LD}. As a result, while the RHF energy order, $E(\gamma_L) < E(\beta_L)$, is preserved, using the DFT(B3LYP) or MP2 methods, the difference between the total energies of the two conformers becomes almost negligible.

Accuracy of Computed Chemical Shifts

Average NMR chemical shifts (and their standard deviations) of all nuclei of glycine and alanine residues are available from BMRB.[39] These data are reported in Tables V and VI. Therefore, one can attempt to answer how well the computed chemical shifts compare with their experimental counterparts. The experimental shifts of the different nuclei ($^{15}N^{NH}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, $^{1}H^{NH}$) retrieved from the database refer to a hypothetical average backbone conformation of the relevant amino acid residue. For purposes of comparison either a simple arithmetical average or a Boltzmann-type (energy-weighted) average of the shifts of the nuclei of individual conformers can be determined. These

TABLE VI.
Chemical Shifts (δ-Scale) Averaged over All Backbone Conformers of For-L-Ala-NH_2.[a,b]

Level[c]	No. Conf.	Average Chemical Shifts (ppm)						Error (%)[d]					
		$^{15}N^{NH}_{avr}$	$^{13}C^{\alpha}_{avr}$	$^{13}C^{\beta}_{avr}$	$^{13}C'_{avr}$	$^{1}H^{NH}_{avr}$	$^{1}H^{\alpha}_{avr}$	N	C^{α}	C^{β}	C′	NH	H^{α}
Arithmetical Average													
A	9	101.46 (4.06)[e]	45.91 (3.65)	16.81 (1.10)	168.58 (0.89)	3.97 (0.54)	3.69 (0.48)	17	13	11	5	51	13
B	6	109.09 (4.19)	48.89 (4.37)	18.05 (1.76)	173.64 (0.39)	4.73 (0.68)	4.09 (0.52)	11	8	5	2	42	4
C	6	101.80 (3.93)	46.06 (4.27)	16.74 (1.75)	168.64 (0.62)	4.09 (0.59)	3.75 (0.50)	17	13	11	5	50	12
D	9	113.86 (4.47)	48.04 (3.92)	17.97 (1.18)	178.69 (0.94)	4.25 (0.59)	3.62 (0.53)	7	9	5	1	48	15
E	6	121.82 (4.60)	51.09 (4.66)	19.26 (1.85)	183.78 (0.58)	5.02 (0.72)	4.04 (0.56)	1	4	2	4	39	5
F	6	125.76 (4.67)	50.49 (4.72)	18.92 (2.16)	183.34 (0.65)	5.12 (0.75)	4.04 (0.59)	2	5	0	3	37	5
Boltzmann Average[f]													
A	9	104.77	42.46	16.38	168.06	4.33	3.91	15	20	13	5	47	8
B	6	110.45	45.16	18.51	173.48	5.25	4.24	10	15	2	2	36	0
C	6	104.01	42.34	17.03	168.28	4.45	3.93	15	20	10	5	46	7
D	9	117.11	44.41	17.50	178.06	4.67	3.85	5	16	7	0	43	9
E	6	122.97	47.08	19.77	183.56	5.63	4.20	0	11	5	3	31	1
F	6	126.50	46.43	19.26	183.20	5.83	4.21	3	12	2	3	29	1
Expt.		**122.82 (5.88)**	**53.05 (2.17)**	**18.91 (1.78)**	**177.40 (8.32)**	**8.17 (0.62)**	**4.24 (0.47)**						

[a] Absolute chemical shifts (in ppm) used as references are given as follows: Levels A, B, and C (6-31+G* GIAO-RHF): ^{1}H (TMS) = 32.000, ^{13}C (TMS) = 201.239, ^{15}N (NH_3) = 266.937; Levels D and E (6-311++G** GIAO-RHF): ^{1}H (TMS) = 32.415, ^{13}C (TMS) = 195.559, ^{15}N (NH_3) = 262.784; Level F (TZ2P GIAO-RHF): ^{1}H (TMS) = 32.231, ^{13}C (TMS) = 194.688, ^{15}N (NH_3) = 268.698.

[b] The number of backbone conformers is 9 when constrained ϕ, ψ and 6 when optimized ϕ, ψ values are used.

[c] See Table I for the designation of the *ab initio* levels used. Expt. = average experimental shift taken from BMRB.[39]

[d] Error = $\mathrm{abs}(\delta^{exp} - \delta^{calc})/\delta^{exp} * 100$.

[e] Standard deviation of the average shift is given in parentheses.

[f] Boltzmann averaging over all conformers was achieved with the use of the corresponding Level A–F energies of the GIAO calculations.

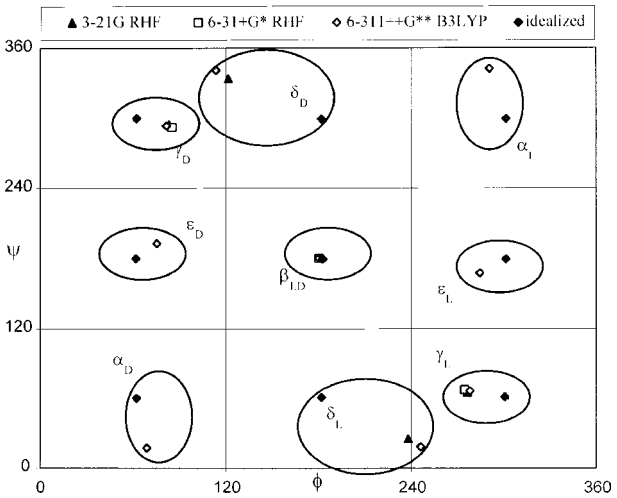

FIGURE 3. The 3-21G RHF, 6-31+G* RHF, 6-311++G** B3LYP and the idealized locations of the different backbone conformers of For-Gly-NH_2 on a Ramachandran surface. Note that the lack of configuration chirality results in $\alpha_L = \alpha_D = \alpha_{LD}$, $\beta_L = \beta_{LD}$, $\gamma_L = \gamma_D = \gamma_{LD}$, $\varepsilon_L = \varepsilon_D = \varepsilon_{LD}$, and $\delta_L = \delta_D = \delta_{LD}$. In the case of the conformers $\alpha_L = \alpha_D = \alpha_{LD}$ and $\varepsilon_L = \varepsilon_D = \varepsilon_{LD}$ the constrained geometry optimizations, performed at the 6-311++G** B3LYP level, utilized characteristic ϕ and ψ torsional angles (see text for details).

data are presented in Tables V and VI for For-Gly-NH_2 and For-L-Ala-NH_2, respectively. Overall, the agreement between the average theoretical and experimental shifts is rather impressive, especially for those data obtained using Levels E and F. It is noteworthy that as the size of the basis set employed increases, the deviation between theoretically and experimentally determined average shifts decreases for all the nuclei but the amide proton. Disregarding the shifts of the amide protons ($^1H^{NH}$), the error for all nuclei, both for the Gly and Ala model systems, is down to a few percent. This means, on the absolute (ppm) scale, that the difference between the best calculated and experimentally determined shifts for ^{13}C and $^{15}N^{NH}$ is typically in the range of only a few ppm, while for the alpha protons the same difference falls to the tenths of ppm range. At this stage it is not clear to us why the difference between the calculated and observed $^1H^{NH}$ values is so much larger than for any other nucleus. It is perhaps of interest to note that the largest error is associated with the most acidic (i.e., deshielded) protons, H^{NH}, and not with the less acidic H^{α} protons. Though deviation of the calculated $^1H^{NH}$ shift from experiment decreases when higher levels of theory are applied, it remains substantial. While the $^1H^{NH}$ shift is off on the absolute (ppm) scale, it is sensitive to conformational changes (*vide infra*). Until now, theoretical studies were concerned with ^{13}C and $^{15}N^{NH}$ shifts only, and with their relation to α-helical and β-sheet conformers of the peptide models. It is demonstrated here that for all nuclei but one, in both peptide models studied and considering all typical backbone orientations, the average shifts can be determined with considerable accuracy.

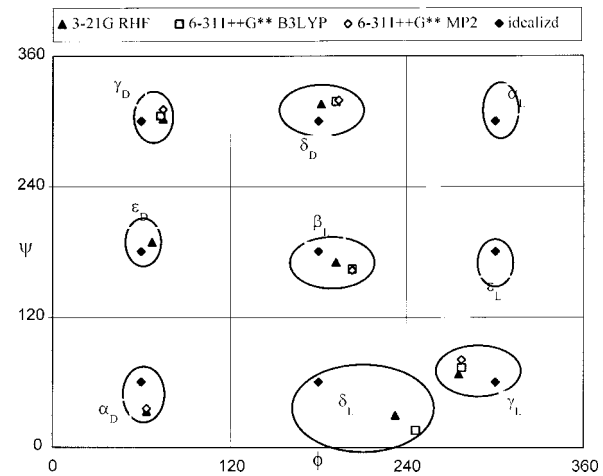

FIGURE 4. The 3-21G RHF, 6-311++G** B3LYP, 6-311++G** MP2, and the idealized locations of the different backbone conformers of For-L-Ala-NH_2 on a Ramachandran surface. Only fully optimized geometries are reported.

Chemical Shift—Chemical Shift Correlations for Main-Chain Fold Determination

Typical two-dimensional (2D) single-quantum coherence (2D-HSQC)[11] or multiple-quantum coherence (2D-HMQC)[10] experiments correlate chemical shifts of selected hetero nuclei (e.g., ^{13}C and ^{15}N) with those of their attached protons. It is widely known that not only the side chains modify the observed chemical shifts (e.g., there is a characteristic shift of $^{15}N^{NH}$ in Gly, Ser, and Thr residues)[22] but for the same residue a significant up- or downfield shift can be detected, depending upon the local molecular environment and/or the backbone conformation. These and similar chemical shift–chemical shift correlations are worthy exploring in detail.

As can be seen from data in Tables III and IV, simple reorientation of the amide planes can have a remarkable effect on the chemical shifts of the nuclei

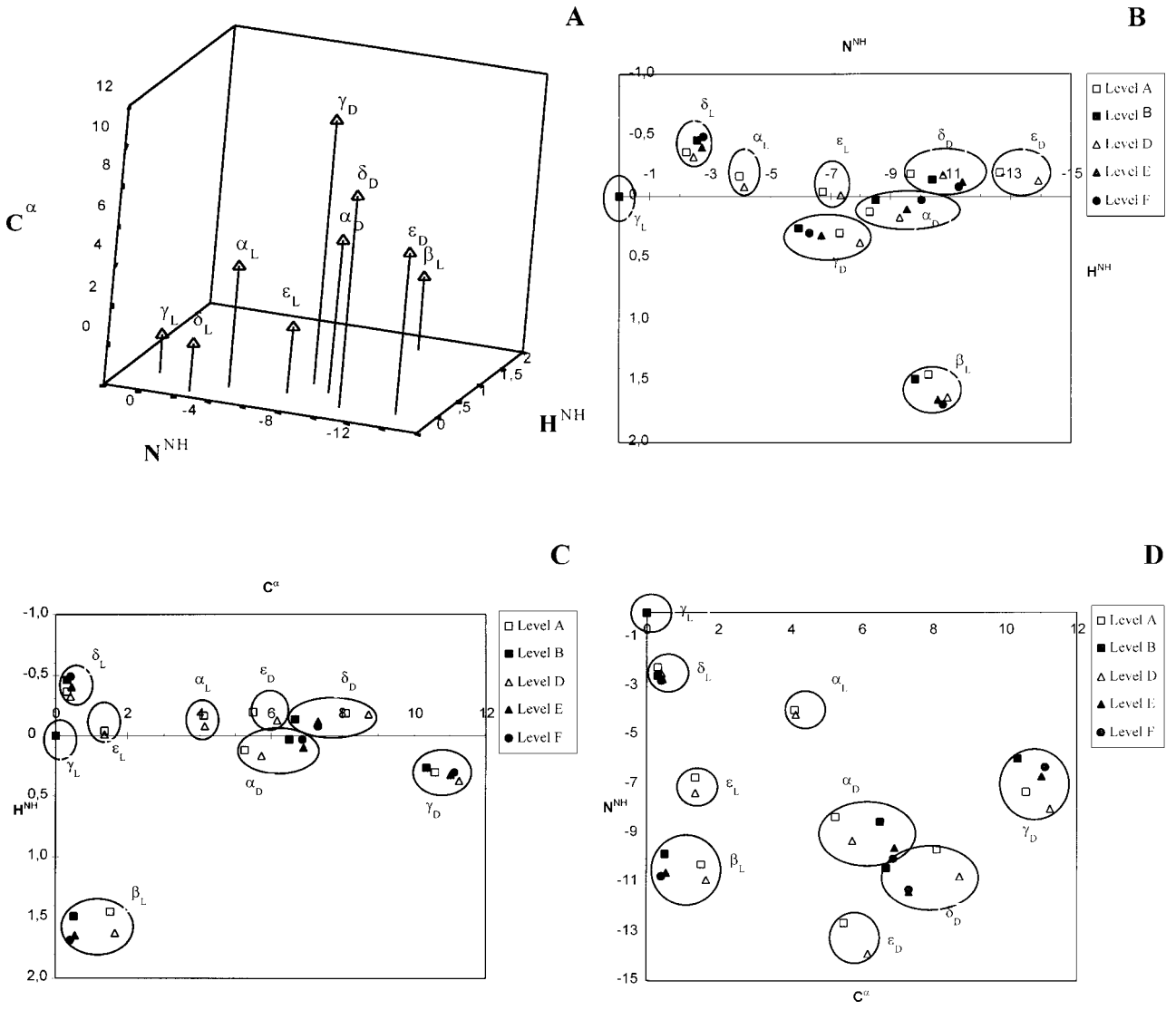

FIGURE 5. $^{15}N^{NH}$-$^{1}H^{NH}$-$^{13}C^{\alpha}$ correlated 3D plot obtained at Level D [**A**] and its three 2D projections ($^{15}N^{NH}$-$^{1}H^{N}$ [**B**], $^{13}C^{\alpha}$-$^{1}H^{NH}$ [**C**], and $^{13}C^{\alpha}$-$^{15}N^{NH}$ [**D**]) for For-L-Ala-NH_2. All shift values of the nine backbone conformers of For-L-Ala-NH_2 are relative to the energetically most stable conformer, γ_L. NMR computations (6-31+G* GIAO-RHF [square], 6-311++G** GIAO-RHF [triangle], and TZ2P GIAO-RHF [circle]) were performed by using two sets of optimized geometries, 6-31++G** RHF [empty symbols] and 6-311++G** B3LYP [filled symbols].

of our model compounds. For example, for For-L-Ala-NH_2 the $\gamma_L \rightarrow \varepsilon_D$ (Level A: $\phi_{\gamma L} = -84.5°$, $\psi_{\gamma L} = 68.7° \rightarrow \phi_{\varepsilon D} = 64.7°$, $\psi_{\varepsilon D} = -178.6°$) conformational shift is followed by a 5.5 ppm change in the $^{13}C^{\alpha}$ and a 12.7 ppm change in the $^{15}N^{NH}$ shifts. Furthermore, interconversion of two conformational neighbors, for example that of $\gamma_L \rightarrow \beta_L$, can easily be monitored through changes of selected shifts (e.g., at Level B the appropriate changes are $\Delta^{15}N^{NH} = 9.8$ ppm and $\Delta^{13}C^{\beta} = -4.4$ ppm). These and related effects can clearly be observed on the $^{15}N^{NH}$-$^{1}H^{NH}$-$^{13}C^{\alpha}$ correlated 3D plot [Fig. 5(A)]. Figure 5(A), as well as the the related plots referring to the $^{15}N^{NH}$-$^{1}H^{N}$, $^{13}C^{\alpha}$-$^{15}N^{NH}$, and $^{13}C^{\alpha}$-$^{1}H^{HN}$ planes [Fig. 5(B–D)], reveal distinct locations for the nine backbone conformers of For-L-Ala-NH_2. Structure-induced changes in the experimentally easily amenable $^{1}H^{NH}$ shifts are small [$\Delta\delta(^{1}H^{NH}) \approx \pm 0.5$ ppm]. Nevertheless, if the amide proton ($^{1}H^{NH}$) shifts are correlated with the amide nitrogen ($^{15}N^{NH}$) shifts, a seemingly useful 2D plot emerges [see Fig. 5(B)]: on this ^{1}H-^{15}N HSQC-like plane the characteristic backbone orientations separate clearly. Note, furthermore, that

all L-type conformers but β_L occupy the left part of the 2D plot. Because D-type conformers are rare in globular proteins, the following useful rule emerges for the relative chemical shifts (δ-scale in ppm) of ^{15}N of amides: $N^{NH}(\gamma_L) > N^{NH}(\delta_L) > N^{NH}(\alpha_L) > N^{NH}(\varepsilon_L) > N^{NH}(\beta_L)$. On the $^1H^{NH}$-$^{13}C^{\alpha}$ plane [Fig. 5(C)], the L-type backbone conformations are once again close to each other. In contrast, all D-type conformers are shifted upfield on the ^{13}C scale. Undoubtedly, each characteristic backbone orientation has its distinct position even on the least commonly used subplane of the 3D correlation plot, $^{13}C^{\alpha}$-$^{15}N^{NH}$ [Fig. 5(D)]. A similarly useful 2D plot of correlated $^{13}C^{\alpha}$-$^1H^{\alpha}$ shifts is presented on Figure 6. Overall, in the case of the two model peptides studied, positions of the correlated shifts of the various backbone conformers differ from each other to such an extent that, if this described the real case in peptides and proteins, it would be straightforward to recognize and to assign the backbone conformations of the amino acid residues from their relative chemical shifts. Considering the resolution of modern spectrometers, the chemical shift differences, which might be observed, seem to be sufficient to distinguish and assign the different backbone conformer types of the same residue in a protein.

There are several factors determining the position of the backbone conformers on these correlated plots. At least three of them are related to computations. These include the effects of one-particle basis set deficiency, the extent of electron correlation, and the effect of geometry optimization. Results reported in Tables III and IV as well as data shown on Figures 5 and 6, reveal that the conformers of the two model peptides are clearly distinguishable on the correlated plots, independent of the basis set employed [6-31+G*, 6-311++G**, DZ(d) or TZ2P] for the calculation of the chemical shieldings. Although the choice whether the GIAO-RHF calculations were performed at constrained (Levels A and D) or fully optimized (Levels B, C, E, and F) reference geometries of For-L-Ala-NH_2 has an effect on the calculated shieldings (Table IV), separation of the backbone conformers is clearly not affected (Figs. 5–7). In the case of For-Gly-NH_2, the effect of electron correlation on CSA was determined directly by performing GIAO-RHF and GIAO-MP2 calculations (Table III). The changes in the CSA values of all nuclei (of all conformers) following inclusion of electron correlation (SCF $\leftrightarrow$ MP2) and extension of basis set [DZ(d) $\leftrightarrow$ TZ2P] (Levels A and B) have a similar magnitude. More specifically, for the α_{LD} and ε_{LD} conformers of For-Gly-NH_2 the effect of basis set extension slightly overwhelms the effect of electron correlation, while for β_{LD} and δ_{LD} the opposite holds. In the future, the effect of electron correlation on the calculated chemical shieldings should be further investigated on chiral peptide models by using much more demanding GIAO-CCSD[47] calculations. Although MP2 calculations have induced significant relative chemical shift changes, it is not expected that electron correlation will alter profoundly the qualitative picture: the basic peptide conformers appear at different regions on chemical shift–chemical shift plots (cf. Figs. 5 and 6). Therefore, the usefulness of the approach of direct determination of conformations of protein building units from multidimensional NMR experiments depends on what effect the side chains, solvation, anisotropic factors, and inter- and intramolecular H-bonding might have on the relative chemical shifts of the selected nuclei. Detailed theoretical investigations of more model compounds and more correlated chemical shielding plots are needed to establish the magnitude of these effects, while it is also hoped that the present theoretical results will encourage experimental work in this direction.

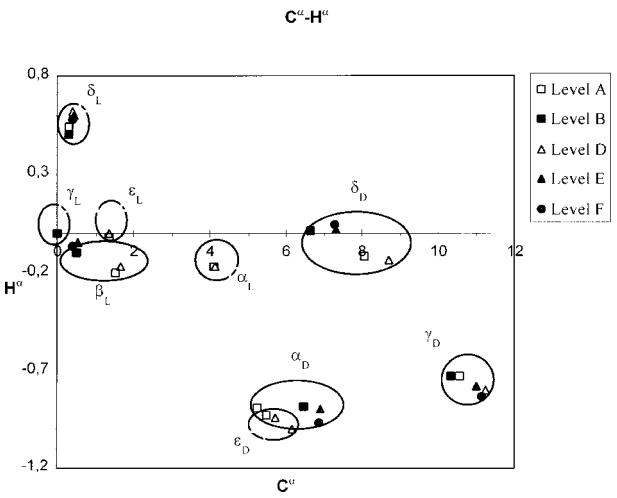

FIGURE 6. $^{13}C^{\alpha}$-$^1H^{\alpha}$ correlated 2D plot. All shifts of the nine characteristic backbone conformers of For-L-Ala-NH_2 are relative to γ_L. NMR computations (6-31+G* GIAO-RHF [square], 6-311++G** GIAO-RHF [triangle], and TZ2P GIAO-RHF [circle]) were performed by using two sets of geometries: 6-31++G** RHF [empty symbols] and 6-311++G** B3LYP [filled symbols].

Chemical Shift–Structure Correlations

One of the primary goals of the research reported in this article has been the investigation of

the use of chemical shifts for predicting the main chain fold of peptides and proteins. So far, we have demonstrated that (a) there is convergence between theoretically determined and experimentally based conformational-average shifts (see above); and (b) chemical shifts cluster as a function of main-chain fold (see previous section). Prompted by these promising results, we performed a systematic linear correlation analysis of all important conformational parameters (ϕ, ψ, ω_0, ω_1, and χ_1) and all chemical shifts ($^{15}N^{NH}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, $^{1}H^{NH}$, and $^{1}H^{\alpha}$). We repeated these linear correlation studies at all 3(6) different levels of theory (see Table I) available for For-Gly-NH_2(For-L-Ala-NH_2). This analysis utilizes the linear correlation (Pearson) coefficient, R, and the related standard error, $S_{Y.X}$, defined as follows:

$$R = \frac{\left[N\left(\sum XY\right) - \left(\sum X\right)\left(\sum Y\right)\right]}{\sqrt{\left[N\sum X^2 - \left(\sum X\right)^2\right]\left[N\sum Y^2 - \left(\sum Y\right)^2\right]}} \quad (1)$$

and

$$S_{Y.X} = \left(\left[\frac{1}{N(N-2)}\right]\left[N\sum Y^2 - \left(\sum Y\right)^2 - \frac{\left[N\sum XY - \left(\sum X\right)\left(\sum Y\right)\right]^2}{N\sum X^2 - \left(\sum X\right)^2}\right]\right)^{1/2} \quad (2)$$

where the vectors **X** and **Y** of length N contain the results to be correlated.

A typical set of results (R^2 values), obtained at Level D, is shown in Table VII. For structure–shift correlation only those R^2 values are relevant, which relate structural parameters with shift values, as given in the box of broken lines in Table VII. (All other R^2 data, reported in Table VII in italics, are irrelevant for this part of the study.) When correlating a set of chemical shifts [$^{13}C^{\alpha}(\alpha_L)$, $^{13}C^{\alpha}(\beta_L)$, etc.] with conformational variables [$\phi(\alpha_L)$, $\phi(\beta_L)$, etc.], a problem arises due to the fact that torsion angles are variables of periodic nature. In conventional definitions torsions sample regions of [−180°, +180°] or [0°, +360°]. Obviously, the correlation is dependent on which definition is actually applied. To obtain reproducible R^2 values, unless otherwise noted, we selected the period [0°, +360°] for the definition of the torsion angles. It is further noted that the linear correlation can be optimized by a (moving) shift in the period. The optimum correlation, for each pair, will be described, hereafter, by R^2_{max}.

The arithmetical average of the calculated R^2 values is very low; for example, it is 0.14 for the data in Table VII. This indicates that on the average there is no meaningful linear correlation between conformational and shift parameters. However, for certain pairs the R^2 values can be three to four times higher than the average value (e.g., $R^2[\phi/C^{\alpha}] = 0.488$, $R^2[\phi/H^{\alpha}] = 0.678$, $R^2[\psi/C^{\alpha}] = 0.38$, and $R^2[\chi_1/H^{\alpha}] = 0.371$; all printed in bold in Table VII), revealing a significant correlation. The "optimized" correlations, R^2_{max} (Table VIII), become even more significant. The interesting structure-shift pairs, for example, ϕ/C^{α}, are collected into Table IX for both model systems. The most important conclusions that can be drawn from data in Tables VII–IX are as follows: (a) there is significant correlation be-

TABLE VII.
R^2 Values between Selected Conformational Parameters[a] and Chemical Shifts of For-L-Ala-NH_2.[b]

	ω_0	ϕ	ψ	ω_1	χ_1	$^{13}C^{\alpha}$	$^{15}N^{NH}$	$^{13}C'$	$^{1}H^{NH}$	$^{13}C^{\beta}$	$^{1}H^{\alpha}$
ω_0	*1*	*0.013*	*0.007*	*0.636*	*0.312*	0.016	0.001	0.296	0.021	0.000	0.000
ϕ		*1*	*0.001*	*0.019*	*0.254*	**0.488**	**0.462**	0.278	0.016	0.111	**0.678**
ψ			*1*	*0.002*	*0.046*	**0.380**	0.086	0.247	0.001	0.059	0.040
ω_1				*1*	*0.139*	0.060	0.089	0.022	0.016	0.015	0.052
χ_1					*1*	0.002	0.138	0.250	0.000	0.061	**0.371**
$^{13}C^{\alpha}$						*1*	*0.296*	*0.000*	*0.006*	*0.052*	*0.452*
$^{15}N^{NH}$							*1*	*0.003*	*0.100*	*0.099*	*0.423*
$^{13}C'$								*1*	*0.007*	*0.325*	*0.144*
$^{1}H^{NH}$									*1*	*0.085*	*0.019*
$^{13}C^{\beta}$										*1*	*0.211*
$^{1}H^{\alpha}$											*1*

[a] All conformational variables are defined in the range [0°, 360°].
[b] Autocorrelation values (e.g., $R^2[\phi/\phi]$) are all equal to 1, the average of R^2 values is 0.144, with a standard deviation of 0.176. Isotropic shielding and (constrained) geometry optimizations have been performed at the 6-311++G** GIAO-RHF and 6-31++G** RHF levels, respectively (Level D of Table I).

TABLE VIII.

Maximized R^2 Values (R^2_{max}) between Selected Conformational Parameters and Chemical Shifts of For-L-Ala-NH_2.[a]

	$^{13}C^{\alpha}$	$^{15}N^{NH}$	$^{13}C'$	$^{1}H^{NH}$	$^{13}C^{\beta}$	$^{1}H^{\alpha}$
ϕ	0.727	0.626	0.340	0.340	0.412	0.678
ψ	0.421	0.475	0.247	0.243	0.411	0.264

[a] See footnotes of Table VII. See text for definition of R^2_{max}.

TABLE IX.

Correlation of Selected Conformational Parameters ([0°, 360°] Representation) with Chemical Shifts for For-Gly-NH_2 and For-L-Ala-NH_2.[a]

Model	Level	ϕ-$^{1}H^{\alpha 1}$	ϕ-$^{1}H^{\alpha 2}$	ϕ-$^{13}C^{\alpha}$	ψ-$^{13}C^{\alpha}$	ψ-$^{15}N^{NH}$	χ_1-$^{13}C'$
For-Gly-NH_2	A	0.93 (4.08)	0.49 (2.14)	0.01 (0.06)	0.12 (0.55)	0.21 (0.92)	
	B	0.98 (4.24)	−0.67 (2.00)	−0.22 (0.22)	0.41 (0.75)	0.42 (0.78)	
	C	0.96 (3.54)	−0.98 (3.67)	−0.59 (1.32)	0.17 (0.11)	−0.02 (0.00)	
For-L-Ala-NH_2	A	−0.83 (4.72)		0.68 (3.16)	−0.64 (2.80)	−0.67 (3.07)	0.52 (1.86)
	B	−0.90 (3.57)		0.86 (3.30)	−0.61 (1.64)	−0.51 (1.16)	0.53 (1.25)
	C	−0.93 (3.85)		0.81 (2.94)	−0.71 (2.27)	−0.59 (1.57)	0.60 (1.60)
	D	−0.82 (4.72)		0.70 (3.39)	−0.62 (2.65)	−0.68 (3.21)	0.50 (1.74)
	E	−0.89 (3.39)		0.86 (3.15)	−0.61 (1.63)	−0.54 (1.24)	0.46 (0.93)
	F	−0.70 (3.56)		0.63 (3.23)	−0.53 (1.68)	−0.55 (1.25)	0.03 (1.30)

[a] For each conformational parameter–chemical shift correlation first R, the linear correlation coefficient [see eq. (1) of the text] is reported followed by R^2/ρ, where ρ is the standard deviation of all R^2 values of Table VII. For For-Gly-NH_2 the ρ values are 0.23, 0.23, and 0.26 for Levels A, B, and C, respectively. For For-L-Ala-NH_2 the ρ values are {0.14, 0.23, 0.22, 0.14, 0.23, 0.23} for Levels {A, B, C, D, E, F}, respectively.

TABLE X.

Chemical Shifts (in ppm), as a Function of the Applied Level of Theory, for the Nuclei C^{α} and H^{α} of For-L-Ala-NH_2.

Nucleus	Conformer	Level A	Level B	Level C	Level D	Level E	Level F
C^{α}	α_D	47.10	51.33	47.51	49.40	53.63	53.01
	α_L	45.94			47.80		
	β_L	43.36	45.35	43.01	45.31	47.26	46.55
	γ_D	52.42	55.19	52.38	54.91	57.74	57.28
	γ_L	41.83	44.84	41.71	43.66	46.72	46.14
	δ_D	49.91	51.50	49.09	52.38	54.04	53.44
	δ_L	42.13	45.15	42.65	44.07	47.15	46.54
	ε_D	47.32			49.84		
	ε_L	43.19			45.02		
H^{α}	α_D	3.08	3.41	3.07	2.97	3.33	3.28
	α_L	3.80			3.74		
	β_L	3.77	4.19	3.87	3.74	4.18	4.18
	γ_D	3.24	3.56	3.24	3.11	3.45	3.42
	γ_L	3.97	4.29	3.98	3.91	4.23	4.25
	δ_D	3.85	4.30	3.94	3.77	4.25	4.29
	δ_L	4.51	4.79	4.39	4.53	4.83	4.83
	ε_D	3.04			2.91		
	ε_L	3.95			3.91		

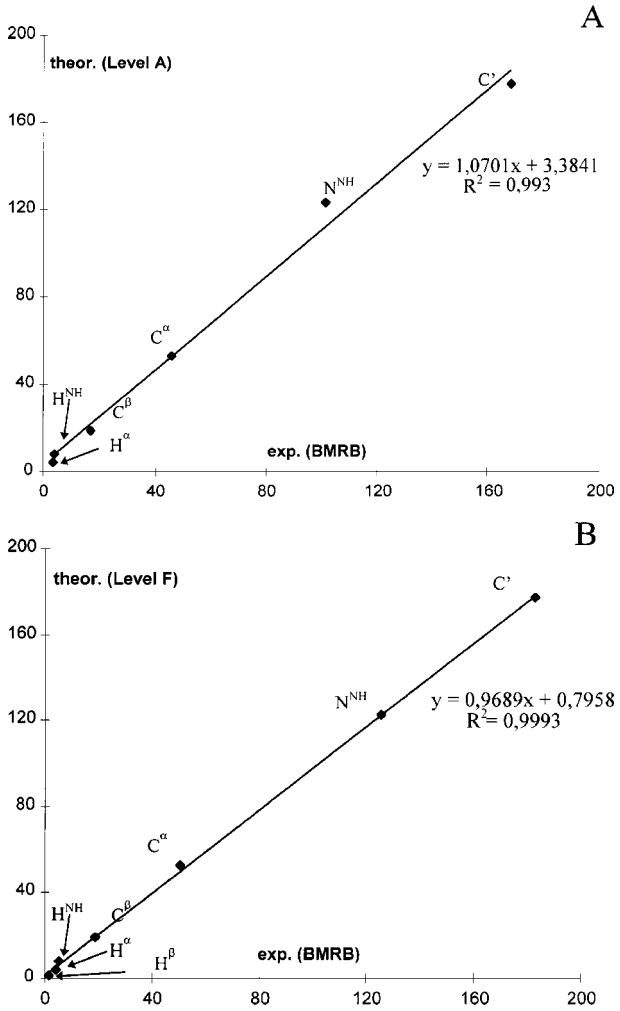

FIGURE 7. Correlation between *ab initio* determined and experimentally found average (conformation independent) chemical shifts plotted for all relevant nuclei in For-L-Ala-NH_2 at two levels of theory: [**A**] = GIAO-RHF/6-31+G*//RHF/6-31++G** (Level A) and [**B**] = GIAO-RHF/TZ2P//B3LYP/6-311++G** (Level F).

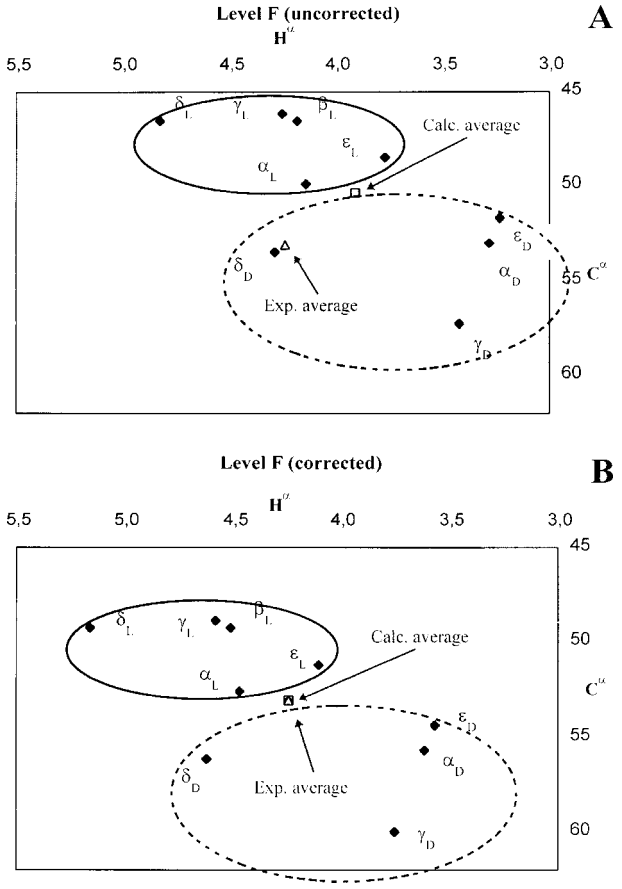

FIGURE 8. The C^{α}/H^{α} correlation plot of the nine typical backbone conformers associated with For-L-Ala-NH_2: (**A**) purely theoretical (Level F), (**B**) corrected with the difference found between calculated and experimental average chemical shifts.

tween ϕ and the $^1H^{\alpha}$ shift, with a typical $R[\phi/H^{\alpha}]$ value larger than 0.8; (b) the $^{13}C^{\alpha}$ shift of For-L-Ala-NH_2 correlates with ϕ just as well as with ψ; (c) somewhat disturbingly, the ϕ-$^{13}C^{\alpha}$ and ψ-$^{13}C^{\alpha}$ correlations become, at certain levels of theory, almost insignificant for For-Gly-NH_2, but perhaps this only indicates that achiral For-Gly-NH_2 is not an ideal model for peptides; and (d) none of the correlations seem to become more significant when higher levels of theory are applied for the computation of the underlying quantities.

Although the linear correlation of selected chemical shifts with conformational parameters is not as high as one would hope, pairs with large R^2 values certainly warrant further investigation. All available results suggest that relative changes in the chemical shifts of $^{13}C^{\alpha}$ and $^1H^{\alpha}$ with the conformational characteristics are important. Therefore, we decided to investigate the dependence of the calculated shifts of these two nuclei as a function of the applied level of theory (Tables X and XI). All chemical shifts calculated at different levels of theory (Levels A–F) correlate well with themselves (e.g., $R(^{13}C^{\alpha}) > 0.978$, $R(^1H^{\alpha}) > 0.987$). This means that extension of the basis or inclusion of electron correlation into the shielding calculation do not increase significantly the reliability of predicting $[\phi, \psi]$ values from calculated $^{13}C^{\alpha}$ and $^1H^{\alpha}$ values. (These findings are based also on results obtained for For-Gly-NH_2 but not presented here.) Figure 8(A) reports an *ab initio* (Level F) C^{α}/H^{α} correlation plot with respect to all typical backbone conformers of For-L-Ala-NH_2. Figure 8(B) shows how these *ab initio* determined shift values would change if a constant, correcting for the

TABLE XI.
Chemical Shift Correlations (R and $S_{Y,X}$), as a Function of the Applied Level of Theory (A–F), for the Nuclei C^{α} and H^{α} of For-L-Ala-NH_2.[a]

Nucleus	Parameter		A	B	C	D	E	F
C^{α}	R	A	1.000	0.978	0.994	0.999	0.980	0.980
		B	0.978	1.000	0.992	0.978	1.000	1.000
		C	0.994	0.992	1.000	0.994	0.993	0.993
		D	0.999	0.978	0.994	1.000	0.981	0.980
		E	0.980	1.000	0.993	0.981	1.000	1.000
		F	0.980	1.000	0.993	0.980	1.000	1.000
	$S_{Y,X}$	A	—	1.029	0.539	0.156	0.977	0.988
		B	1.020	—	0.628	1.010	0.098	0.106
		C	0.523	0.613	—	0.518	0.569	0.572
		D	0.168	1.084	0.569	—	1.024	1.039
		E	1.034	0.105	0.621	1.018	—	0.066
		F	1.058	0.115	0.632	1.045	0.067	—
H^{α}	R	A	1.000	0.992	0.989	0.999	0.991	0.989
		B	0.992	1.000	0.998	0.991	0.999	1.000
		C	0.989	0.998	1.000	0.987	0.996	0.998
		D	0.999	0.991	0.987	1.000	0.992	0.988
		E	0.991	0.999	0.996	0.992	1.000	0.999
		F	0.989	1.000	0.998	0.988	0.999	1.000
	$S_{Y,X}$	A	—	0.074	0.088	0.019	0.076	0.087
		B	0.074	—	0.032	0.079	0.029	0.018
		C	0.084	0.031	—	0.089	0.051	0.037
		D	0.021	0.087	0.102	—	0.079	0.097
		E	0.082	0.032	0.057	0.078	—	0.025
		F	0.098	0.020	0.044	0.100	0.026	—

[a] For primary entries see Table X. See Table I for the description of the theoretical levels.

deviation between experimental and calculated average shift values, were added to them. Figure 8(B) is perhaps the most meaningful way to present the relevant theoretical results to experimentalists.

Because the two predominant secondary structural elements of proteins are the α-helix and the β-sheet, accurate description of the NMR characteristics of these two conformational building units has been of primary interest for experimentalists and theoreticians alike.[5, 9, 13, 15, 48] The most important results are as follows: (a) there is a characteristic downfield shift of the C^{α} and C′ chemical shifts associated with the right-handed helix conformation; and (b) there is a significant upfield shift when the peptide is in a β-sheet structure. Nevertheless, there is still a lot to be learned as important atoms, such as protons, are typically excluded from these analysis. Consequently, we collected previous theoretical[15c, 19, 22, 49] and experimental chemical shifts in Table XII and compare them with the present chemical shift data. Note that previous shifts might refer to slightly different model compounds with slightly different torsion angles and the calculations have been performed at somewhat different levels of theory. Nevertheless, for peptide models containing alanine (Table XIIA), all calculated chemical shifts for carbon nuclei (C^{α}, C^{β}, and C′) are in agreement with literature data. The positive value of $\Delta\delta$ calculated for C^{α} agrees with the expected downfield shift of a helical backbone and the upfield shift characteristic of an extended conformation. The smallest difference between calculated ($\Delta\delta^{\alpha-\beta}_{\mathrm{calc}} = 3.3$ ppm) and experimentally determined $^{13}C^{\alpha}$ shifts ($\Delta\delta^{\alpha-\beta}_{\mathrm{exp}} \approx 3.0$ ppm) is obtained when Level G is compared to a 3_{10}-helix/β-sheet conversion (data not shown). In the case of For-Gly-NH_2, the best agreement between theory and experiment for $^{13}C^{\alpha}$ is obtained at the MP2 level ($\Delta\delta^{\alpha-\beta}_{\mathrm{exp}} \approx$ 2.0 ppm and $\Delta\delta^{\alpha-\beta}_{\mathrm{calc}} = 1.6$ ppm). The sign of the $\Delta\delta$ quantity for a helix-sheet conversion should be opposite for C^{β} than for C^{α}, which is indeed the case. Furthermore, the computed difference, typi-

TABLE XII.

Comparison of Theoretical and Experimental NMR Relative Chemical Shift Differences ($\delta^{\alpha\text{-helix}} - \delta^{\beta\text{-sheet}}$) between α-Helix- and β-Sheet-Like Conformers of the For-Gly-NH_2 and For-L-Ala-NH_2 Model Systems.

For-Gly-NH_2	N^{NH}	C^{α}	C'	H^{NH}	H^{α}	H^{α}	H^{α}
IGLO/DZ//SCF/3-21G[e] (ref. 19)	—	5.2	—	−4.6			—
RHF/6-311++G(2p2d) (ref. 22)	—	4.3	—	—			—
GIAO-RHF/DZ(d)// B3LYP/6-311++G** (Level A)	−3.4	1.0	−4.5	−1.8	−0.7	0.1	−0.3
GIAO-MP2/DZ(d)//B3LYP/6-311++G** (Level B)	−2.6	1.6	−2.3	−1.3	−0.7	0	−0.4
GIAO-RHF/TZ2P//B3LYP/6-311++G** (Level C)	5.7	4.0	0.2	−1.4	−0.5	0.4	−0.1
Experimental data from Wishart et al. (ref. 5a)		2.8					
Experimental BMRB[f] (Gly[in β-sheet]-Gly[in α-helix])		0.9					−0.2

For-L-Ala-NH_2	N^{NH}	C^{α}	C^{β}	C'	H^{NH}	H^{α}
deMon-NMR/IGLOIII[a] (ref. 15c)	−1.7	6.8	0	0	—	—
GIAO-SCF/TZP[a] (ref. 15c)	−2.2	4.7	−1.1	−0.6	—	—
GIAO-SCF/IGLOIII[a] (ref. 15c)	−1.9	5.0	−1.2	−0.4	—	—
RHF/6-311++G(2p2d) (ref. 22)	—	5.0	—	—	—	—
DFT deMon (ref. 43)	—	—	—	—		−0.4
GIAO-RHF/6-31+G*//SCF/6-31++G**[b] (Level A)	6.3	2.6	−1.4	0.8	−1.6	0.0
GIAO-RHF/6-311++G**//SCF/6-31++G**[b] (Level D)	6.8	2.5	−1.4	1.0	−1.7	0.0
RHF/TZ2P//B3LYP/6-311++G**[b] (Level G) {3_{10}-helix or type I β-turn}	6.1	3.3	−2.5	0	−1.9	0.0
RHF/TZ2P//B3LYP/6-311++G**[c] (Level G)	4.5	5.1	−2.3	0	−2.0	−0.2
Experimental data from Kricheldorf and Müller (ref. 48)	—	3.5	−4.8	4.6	—	—
Experimental data from Wishart et al. (ref. 5a)	0	3.5	—	2.1	—	—
Experimental data from Spera and Bax (ref. 9d)	—	4.6	−2.5	—	—	—
Experimental BMRB[d] (Ala[in β-sheet] − Ala[in α-helix])		3.0				−0.8

[a] Computed values are for *N*-acetyl-*N'*-methylalaninamide; see Table I of ref. 15c for results for the α-helix ($\phi = -60°, \psi = -60°$) and the extended conformer ($\phi = -120°, \psi = 120°$).

[b] Values calculated for *N*-formyl-alaninamide are for 3_{10}-helices and/or for residue ($i+1$) in a type I β-turn ($\phi = -68.6°, \psi = -17.5°$) and extended ($\phi = -167.6°, \psi = 169.9°$) conformers.

[c] Values calculated for *N*-formyl-alaninamide are for the α-helix ($\phi = -54°, \psi = -45°$) and for the extended ($\phi = -167.6°, \psi = 169.9°$) conformer.

[d] Some 50 alanine residues, all having C^{α} and H^{α} assignments, have been included.

[e] Values calculated for *N*-acetyl-*N'*-methylglycinamide are for the α-helix ($\phi = -48°, \psi = -57°$) and for the extended ($\phi = -139°, \psi = 135°$) conformers taken from Table I of ref. 19.

[f] Some 40 glycine residues, all having C^{α} and H^{α} assignments, have been included.

cally around -2.5 ppm $\leq \Delta\delta^{\alpha-\beta}_{calc} \leq -1.1$ ppm, agrees well with the experimental value of $\Delta\delta^{\alpha-\beta}_{exp} \approx -2.5$ ppm.[9c] Carbonyl carbons are less frequently used for monitoring conformational shifts. Once again, our calculations confirm that ^{13}C σ values contain conformational information[50, 51] and chemical shifts have information about ϕ, ψ values. The result obtained at Level D (Table XIIA) is the closest to the experimental value of Wishart and Sykes,[5a] deviating by only about 1 ppm ($\Delta\delta^{\alpha-\beta}_{calc} = 1.0$ ppm, $\Delta\delta^{\alpha-\beta}_{expt} = 2.1$ ppm). Although $^{15}N^{NH}$ shifts are also affected by dihedral parameters,[51] they are expected to be more complicated to use for a conformational analysis. For example, the experimental value for the amide nitrogen, $\Delta\delta^{\alpha-\beta}_{expt} \approx 0$ ppm, has an uncertainty of ± 4 ppm.[5a] As reported in Table XII, values computed for the $^{15}N^{NH}$ shift in different conformational states can be rather different.

Conclusions

Establishing correlation between peptide main-chain folds and chemical shifts provides a continu-

ous challenge for experimentalists and theoreticians alike. For theoreticians, uncertainties arise, for example, from the problem of ideal size and type of a peptide model to be used for the computations, the required minimum level of *ab initio* theory, and the incorporation of important structural factors (e.g., nonplanarity of the amide groups). There are just as severe experimental difficulties in establishing such correlations. Consequently, unambiguous correlations have been put forward only for the α-helical and β-sheet regions of the Ramachandran surface. One of the principal aims of this computational study has been the confirmation of existing correlations and derivation of new ones. Some more important findings of this study, relevant or related to this issue, are as follows:

The energy order of the L-type conformers of the two peptide models, For-Gly-NH_2 and For-L-Ala-NH_2, is identical, regardless whether electron correlation is incorporated into the computation or not, and it is as follows: $E(\gamma_L) < E(\beta_L) < E(\delta_L) < E(\varepsilon_L) < E(\alpha_L)$. Typically, the D-type backbone conformers have higher relative energy or lower stability. The α_L, ε_L, and ε_D conformers are found not to correspond to local minima on the potential energy hypersurfaces of For-Gly-NH_2 and For-L-Ala-NH_2.

Agreement between theoretical and experimental (BMRB) chemical shifts averaged over all conformers is impressive. As the size of the basis set employed for the shielding calculation increases, the deviation between the shifts decreases to a few percent for all nuclei but the amide proton.

If $^1H^{NH}$ shifts are correlated with $^{15}N^{NH}$ shifts (the $^{15}N^H$-$^1H^N$ planes could result from HSQC ($^{15}N^{NH}$, $^1H^{NH}$)-type experiments) or if $^1H^\alpha$ shifts are plotted against $^{13}C^\alpha$ shifts, useful 2D plots emerge, independently of the basis set employed and whether electron correlation has been included in the theoretical treatment. In the case of the two model peptides studied, positions of the various backbone conformers on the chemical shift–chemical shift plots differ from each other to such an extent that, if this described the real case in peptides, it would be straightforward to recognize and to assign the main-chain fold from the relative chemical shifts.

The empirical structure–chemical shift correlations observed by Wishart,[5] Oldfield,[13] Bax,[4d] and others have been successfully applied in structural determinations of helical and extended-like secondary structures. The data presented in this paper give further justification to these empirical findings and reveals the intrinsic correlation of all nine characteristic backbone conformations of peptides and their chemical shift information.

In summary, *ab initio* isotropic NMR shielding results presented in this article for the model systems For-Gly-NH_2 and For-L-Ala-NH_2 facilitate and encourage the application of correlated relative chemical shift information from {^{1}H-^{15}N}HSQC, {^{1}H-^{13}C}HSQC, HNCA, HNCB, and other multiple-pulse NMR experiments to extract structural information directly from these measurements, thus opening an alternative route to NOE's to derive structures of proteins from their NMR spectra. Detailed theoretical investigations of more model compounds (e.g., For-Ser-NH_2, For-Val-NH_2, and For-Phe-NH_2) are underway in our laboratories. The preliminary results support all the conclusions presented in this article for the simpler model systems For-Gly-NH_2 and For-L-Ala-NH_2.

Acknowledgments

The authors thank I. Jákli and Dr. G. D. Show for helpful discussions.

References

1. Neuhaus, D.; Williamson, M. The Nuclear Overhauser Effect in Structural and Conformational Analysis; VCH: New York, 1989.
2. (a) Wüthrich, K. In NMR of Proteins and Nucleic Acids; Wiley: New York, 1986; (b) Wüthrich, K. Science 1989, 243, 45.
3. Smith, J. A.; Pease, L. G. CRC Crit Rev Biochem 1980, 8, 315.
4. (a) Nilges, M. J Mol Biol 1995, 245, 645; (b) Fesik, S. W.; Zuiderweg, E. R. P. Q Rev Biophys 1990, 23, 97; (c) Clore, G. M.; Gronenborn, A. M. Prog NMR Spectrosc 1991, 23, 43; (d) Bax, A.; Grzesiek, S. Acc Chem Res 1993, 26, 131; (e) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Skelton, N. J. Protein NMR Spectroscopy. Principles and Practice; Academic: San Diego, 1996.
5. (a) Wishart, D. S.; Sykes, B. D.; Richards, F. M. J Mol Biol 1991, 222, 311; (b) Wishart, D. S.; Sykes, B. D. Methods in Enzimology; James, T. L.; Oppenheimer, N. J., eds.; Academic: New York, 1992.
6. Szilágyi, L. Prog NMR Spectrosc 1995, 27, 325.
7. Saito, H. Magn Reson Chem 1986, 24, 835, and references therein.
8. Johnson, C. E.; Bovey, F. A. J Chem Phys 1958, 29, 1012.
9. (a) Perkins, S. J.; Wüthrich, K. Biochim Biophys Acta 1979, 576, 409; (b) Haigh, C. W.; Mallion, R. B. Prog NMR Spectrosc 1980, 13, 303; (c) Williamson, M. P. Biopolymers 1990, 29, 1423; (d) Spera, S.; Bax, A. J Am Chem Soc 1991, 113, 5490.
10. (a) Müller, L. J Am Chem Soc 1979, 101, 4481; (b) Bax, A.; Griffey, R. H.; Hawkins, B. L. Magn Res 1983, 55, 301; (c) Redfield, A. G. Chem Phys Lett 1987, 96, 537.

11. Bodenhausen, G.; Ruben, D. J. Chem Phys Lett 1980, 69, 185.
12. Ikura, M.; Kay, L. E.; Bax, A. Biochemistry 1990, 29, 4659.
13. (a) de Dios, A. C.; Pearson, J. G.; Oldfield, E. Science 1993, 260, 1491; (b) de Dios, A. C.; Pearson, J. G.; Oldfield, E. J Am Chem Soc 1993, 115, 9768; (c) de Dios, A. C.; Oldfield, E. J Am Chem Soc 1994, 116, 5307; (d) Le, H.-B.; Pearson, J. G.; de Dios, A. C.; Oldfield, E. J Am Chem Soc 1995, 117, 3800.
14. (a) Reily, M. D.; Thanabal, V.; Omecinsky, D. O. J Am Chem Soc 1992, 114, 6251; (b) Pastore, A.; Saudek, V. J Magn Reson 1990, 90, 165.
15. (a) Sulzbach, H. M.; Schleyer, P. v. R.; Schaefer, H. F., III J Am Chem Soc 1994, 116, 3967; (b) Sulzbach, H. M.; Schleyer, P. v. R.; Schaefer, H. F., III J Am Chem Soc 1995, 117, 2632; (c) Sulzbach, H. M.; Vacek, G.; Schreiner, P. R.; Galbraith, J. M.; Schleyer, P. v. R.; Schaefer, H. F., III J Comp Chem 1997, 18, 126.
16. Bagno, A. J Mol Struct 1997, 418, 243.
17. de Dios, A. C.; Laws, D. D.; Oldfield, E. J Am Chem Soc 1994, 116, 7784.
18. Pearson, J. G.; Oldfield, E.; Lee, F. S.; Warshel, A. J Am Chem Soc 1993, 115, 6851.
19. Jiao, D.; Barfield, M.; Hruby, V. J. J Am Chem Soc 1993, 115, 10883.
20. Asakawa, N.; Kurosu, H.; Ando, I. J Mol Struct 1994, 323, 279.
21. Pearson, J. G.; Wang, J. F.; Markley, J. L.; Le, H.-B.; Oldfield, E. J Am Chem Soc 1995, 117, 8823.
22. Havlin, R. H.; Le, H.-B.; Laws, D. D.; deDios, A. C.; Oldfield, E. J Am Chem Soc 1997, 119, 11951.
23. 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, 119, 7827.
24. Le, H.-B.; Oldfield, E. J Phys Chem 1996, 100, 16423.
25. Asakawa, N.; Kurosu, H.; Ando, I.; Shoji, A.; Ozaki, T. J Mol Struct (Theochem) 1994, 317, 119.
26. de Dios, A. C.; Oldfield, E. J Am Chem Soc 1994, 116, 11485.
27. Zheng, G.; Wang, L. M.; Hu, J. Z.; Zhang, X. D.; Shen, L. F.; Ye, C. H.; Webb, G. A. Magn Res Chem 1997, 35, 606.
28. Laws, D. D.; Le, H.-B.; de Dios, A. C.; Havlin, R. H.; Oldfield, E. J Am Chem Soc 1995, 117, 9542.
29. Pearson, J. G.; Le, H. B.; Sanders, L. K.; Godbout, N.; Havlin, R. H.; Oldfield, E. J Am Chem Soc 1997, 119, 11941.
30. (a) Ramakrishnan, C.; Ramachandran, G. N. Biophys J 1965, 5, 909; (b) IUPAC-IUB Commission on Biochemical Nomenclature Biochemistry 1970, 9, 3471; (c) Perczel, A.; Ángyán, J. G.; Kajtár, M.; Viviani, W.; Rivail, J.-L.; Marcoccia, J.-F.; Csizmadia, I. G. J Am Chem Soc 1991, 113, 6256; (d) McAllister, M. A.; Perczel, A.; Császár, P.; Viviani, W.; Rivail, J.-L.; Csizmadia, I. G. J Mol Struct 1993, 290, 161; (e) Vladia, W.; Rivail, J.-L.; Perczel, A.; Csizmadia, I. G. J Am Chem Soc 1993, 115, 8321; (f) Perczel, A.; Farkas, Ö.; Császár, A. G.; Csizmadia, I. G. Can J Chem 1997, 75, 1120; (g) Császár, A. G.; Perczel, A. Prog Biophys Mol Biol 1999, 71, 243.
31. Frisch, M. J., et al. Gaussian94, Revision B.2, Gaussian Inc.: Pittsburgh, PA (1995).
32. Frisch, M. J., et al. Gaussian98, Revision A.5, Gaussian Inc.: Pittsburgh, PA (1998).
33. Stanton, J. F.; Gauss, J.; Watts, J. D.; Lauderdale, W. J.; Bartlett, R. J. Int J Quant Chem, Symp 1992, 26, 879.
34. (a) Ditchfield, R. Mol Phys 1974, 27, 789; (b) Wolinski, K.; Hinton, J. F.; Pulay, P. J Am Chem Soc 1990, 111, 8251.
35. Gauss, J. J Chem Phys 1993, 99, 3629.
36. (a) Hariharan, P. C.; Pople, J. A. Theor Chim Acta 1973, 28, 213; (b) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. J Chem Phys 1980, 72, 650.
37. Schäfer, A.; Horn, H.; Ahlrichs, R. J Chem Phys 1992, 97, 2571.
38. Schäfer, A.; Huber, C.; Ahlrichs, R. J Chem Phys 1994, 100, 5829.
39. Seavey, X. J Biomol NMR 1991, 1, 217.
40. (a) Császár, A. G. J Am Chem Soc 1992, 114, 9568; (b) Barone, V.; Adamo, C.; Lelj, F. J Chem Phys 1995, 102, 364; (c) Császár, A. G. J Phys Chem 1996, 100, 3541.
41. Perczel, A.; McAllister, M. A.; Császár, P.; Csizmadia, I. G. J Am Chem Soc 1993, 115, 4849.
42. Császár, A. G.; Allen, W. D.; Schaefer, H. F., III J Chem Phys 1998, 108, 9751.
43. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. E.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. J Mol Biol 1977, 112, 535.
44. (a) Perczel, A.; Farkas, Ö.; Csizmadia, I. G. J Am Chem Soc 1996, 118, 7809; (b) Perczel, A.; McAllister, M. A.; Császár, P.; Csizmadia, I. G. Can J Chem 1994, 72, 2050.
45. Head–Gordon, T.; Head–Gordon, M.; Frisch, M. J.; Brooks, C., III; Pople, J. A. J Am Chem Soc 1991, 113, 5989.
46. Endrédi, G.; Perczel, A.; Farkas, Ö.; McAllister, M. A.; Csonka, G. I.; Ladik, J.; Csizmadia, I. G. J Mol Struct (Theochem) 1997, 391, 15.
47. Gauss, J.; Stanton, J. F. J Chem Phys 1995, 102, 251.
48. Kricheldorf, H. R.; Müller, D. Macromolecules 1983, 16, 615.
49. Sitkoff, D.; Case, D. A. J Am Chem Soc 1997, 119, 12262.
50. deDios, A. C. Prog NMR Spectrosc 1996, 29, 229.
51. Oldfield, E. J Biomol NMR 1995, 5, 217.