Ab Initio Studies of Amide-^{15}N Chemical Shifts in Dipeptides: Applications to Protein NMR Spectroscopy†

Hongbiao Le‡ and Eric Oldfield*

Department of Chemistry, University of Illinois at Urbana-Champaign, 505 South Mathews Avenue, Urbana, Illinois 61801

Received: February 29, 1996; In Final Form: July 28, 1996⊗

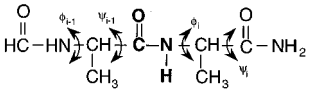
The results of calculations aimed at providing a better understanding of how protein structural parameters affect ^{15}N nuclear magnetic resonance (NMR) chemical shifts, using *ab initio* quantum chemical methods, are reported. The results support previous empirical observations that the two backbone dihedral angles closest to the peptide group (ψ_{i-1} and ϕ_i) have the largest effects on ^{15}N chemical shifts, contributing a range of about 20 ppm. The adjacent torsion angles ϕ_{i-1} and ψ_i have a smaller contribution, up to 8 ppm, but also need to be considered when predicting protein chemical shifts. Different side chain conformations produce chemical shift variations of up to ∼4 ppm. Hydrogen bonding to peptide carbonyl groups can also contribute to ^{15}N shielding, as can longer range electrostatic field effects, but these effects are smaller than those due to torsions. Calculations of ^{15}N chemical shifts of nonhelical alanine residues in a *Staphylococcal* nuclease, dihydrofolate reductase from *Lactobacillus casei*, and ferrocytochrome c_{551} from *Pseudomonas aeruginosa* show a good correlation between experimental observation and *ab initio* prediction, but the shielding of helical residues is overestimated by ∼8 ppm, due most likely to electric field effects from the helix dipole. ^{15}N NMR chemical shifts are very sensitive probes of protein conformation and have potential for structure validation, although at present they are less useful than are ^{13}C shifts for prediction and refinement, because of their more complex dependence on multiple torsional, as well as electrostatic field, effects.

Introduction

Folding a protein into its native conformation causes a large range of chemical shift nonequivalences to be generated,[1,2] a prerequisite for multidimensional nuclear magnetic resonance (NMR) spectroscopic studies of protein structure.[3] However, only in the past four or five years have there been successful attempts to explain the origins of ^{1}H,[4−6] ^{13}C,[7,8] ^{15}N,[8] and ^{19}F shifts[8,9] in proteins. Of all these nuclei, the most studied but least well understood chemical shifts arise from ^{15}N. ^{15}N is readily incorporated into most proteins from $^{15}NH_4^+$, and ^{15}N shifts are routinely obtained in most structure determinations. However, the secondary structural correlations seen with ^{13}C and ^{1}H[7,10] are much weaker with ^{15}N. For example, Wishart et al. observed a 3−4 ppm ^{15}N shift difference between helical and sheet residues,[10] but this is only ∼10−20% of the total ^{15}N shielding range typically seen in proteins. Similarly, we and others have observed a weak correlation between ^{15}N shifts and ϕ_i, ψ_{i-1},[11,12] but clearly additional factors need to be considered in order to achieve better agreement between theory and experiment. As one step in this direction, de Dios et al.[8] used *ab initio* techniques to compute the ^{15}N shifts of valine residues in a nuclease from *Staphylococcus aureus*. Better agreement was observed between theory and experiment than with the empirical methods, which has led us to investigate in more detail the factors which might contribute to ^{15}N shieldings in proteins. *Ab initio* methods have, of course, the advantage that structural parameters can be varied in a highly controlled manner, and the effects of e.g. an isolated ϕ_i change can be investigated in detail. Here, we report the results of calculations in which specific dipeptide torsion angles are varied, and how these torsion angle changes influence ^{15}N chemical shifts. The same *ab initio* methods are then applied to three small proteins, which have high-resolution X-ray structures reported[13−15] as well as solution ^{15}N NMR chemical shifts:[16−18] *Staphylococcal* nuclease (SNase), dihydrofolate reductase (DHFR), and ferrocytochrome c_{551}.

Methods

The model compound chosen for initial study is *N*-formyl-alanyl-alanine amide:

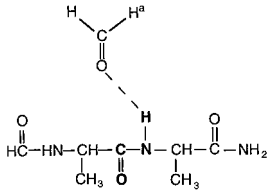


Structural parameters like bond lengths, bond angles, and some dihedral angles were taken from protein structures that were energy minimized using the Discover program with an AMBER force field (Biosym Technologies, Inc., San Diego, CA), a strategy we have shown previously to give good accord with C^{α}, C^{β} shifts in proteins.[8] We used a locally dense[19] basis set: 6-311++G(2d,2p) on the bold-faced atoms above, and 6-31G on the rest of the atoms. NMR shielding calculations were performed primarily by using the gauge including atomic orbitals (GIAO) TEXAS program of Pulay, Wolinski, and Hinton[20,21] on IBM (International Business Machines Corp., Austin, TX) RS/6000 workstations. In the study of dihedral angle effects, a specific torsion angle was changed using a 20° increment, while all other coordinates were fixed. For one set of calculations, each structure was geometry optimized, with the backbone dihedral angles restrained, by using a Hartree−Fock method in Gaussian-94 (Gaussian, Inc., Pittsburgh, PA) employing a 6-31G basis set. Shielding calculations were carried out for several backbone configurations, representative of α-helix, β-sheet, and turn structures. To investigate hydrogen, bonding, we used formaldehyde as a partner molecule, as follows:

† This work was supported by the United States Public Health Service (grant GM-50694).

‡ University Graduate Fellow.

⊗ Abstract published in *Advance ACS Abstracts,* September 1, 1996.



For these studies, we varied the dihedral angle N−H^N−O−C while the backbone adopted either an α-helical or β-sheet structure. The orientational angles N−H^N−O, H^N−O−C and O−C−H^a were chosen according to the most probable values deduced from X-ray diffraction data.[13−15] An O- - -H^N hydrogen bond length of 2 Å was used, which corresponds to the most stable distance in a hydrogen bond.[22] Shieldings were evaluated without correction for basis set superposition error (BSSE), since in test calculations this effect was typically <0.2 ppm.

The computation of ^{15}N shieldings for protein residues was performed as follow: a dipeptide fragment was clipped out of a Brookhaven Protein Data Bank[23] structure, then capped with N-formyl and C-amide terminating functions, in order to make the model fragment chemically realistic, basically as we have described previously.[8] Then, while keeping all dihedral angles fixed, bond lengths and angles were modified slightly to the values of AMBER Forcefield (Biosym, CA). Finally, the side chain of the first residue was modified to make it an alanine. In some cases, dipeptides containing the actual preceding residues in a protein were investigated, but there were only small effects on shielding. Locally dense basis sets were again used.

Calculations were carried out first at the SCF/HF level of theory, which neglects electron correlation. This can be expected to contribute an error of tens of ppm to the absolute shieldings; however, as noted by others,[24] the shape functions (how shifts vary with torsions) are expected to remain the same at MP2 levels of theory. We should also note that we used planar or nonpyramidalized nitrogen.[25] The reasons for this are that there seems to be no direct experimental evidence for pyramidalization in proteins or peptides, the theoretical ^{15}N shift is quite insensitive to pyramidalization,[25] and it is not known whether a full geometry optimization of a *hydrogen-bonded* fragment would actually retain this effect. Moreover, recent experimental studies[26] support only an extremely small pyramidalization. Given that we do obtain generally good agreement between theory and experiment, and the effects of ϕ_i, ψ_i, ϕ_{i-1}, ψ_{i-1} are all large, we have therefore concentrated on these known structural effects in this investigation. Toward the end of our study, we also investigated shielding in a number of fragments using a sum-over-states density functional (DFT) method,[27] using a uniform IGLO-III basis[28] on all atoms and a Perdew−Wang-91 exchange-correlation functional.[29] The DFT method takes into account the electron exchange-correlation energy using a local density approach.[30] The correlation between the GIAO and DFT results for the alanine residues in the three proteins was very high, $R^2 = 0.98$, and the shape functions for the ϕ_{i-1} and ψ_i rotation plots were essentially identical using both methods. We used an absolute shielding of 244.6 ppm for liquid ammonia at ambient temperature to convert calculated shieldings to the IUPAC chemical shift scale.[31]

Results and Discussion

Backbone Dihedral Angle Effects. Inspecting the *N*-formylalanylalanine amide model, it appears likely that there might be four backbone dihedral angles which could exert an influence on the amide ^{15}N chemical shift, that is, ϕ_{i-1}, ψ_{i-1}, ϕ_i, and ψ_i. Even though ϕ_{i-1} and ψ_i are not directly attached to the amide moiety of interest (shown in Figure 1), they do affect the relative orientations of this amide group with the preceding and following ones (Figure 1), and might reasonably be expected to influence shielding, in large part because of their dipolar nature.

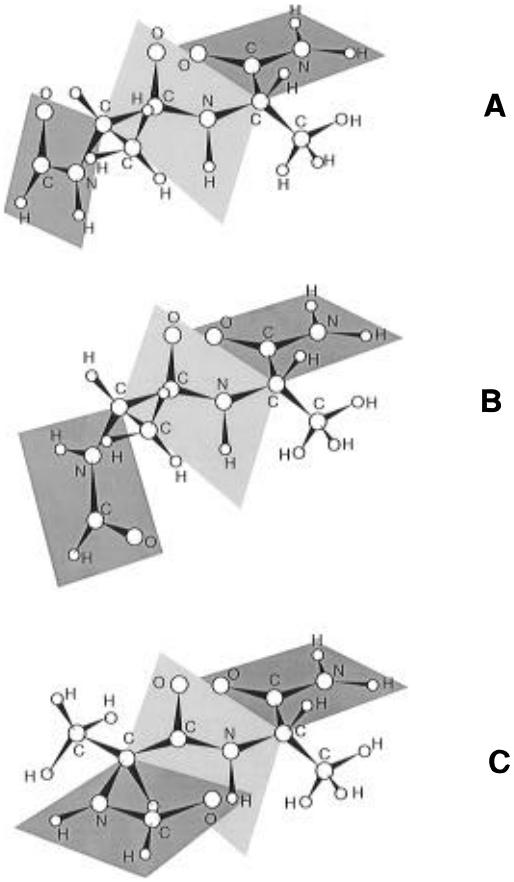


Figure 1. Schematic showing *N*-formyl-alanyl-[^{15}*N*]alanine amide molecules used in ^{15}N shielding calculations. The peptide planes are indicated in boxes. (A) $\phi_{i-1} = -55°$; $\psi_{i-1} = -55°$; $\phi_i = -55°$; ψ_i freely variable (□ in Figure 2A). (B) $\phi_{i-1} = 100°$; $\psi_{i-1} = -55°$; $\phi_i = -55°$; ψ_i freely variable (● in Figure 2A). (C) $\phi_{i-1} = -55°$, $\psi_{i-1} = 120°$; $\phi_i = -55°$; ψ_i freely variable (○ in Figure 2A).

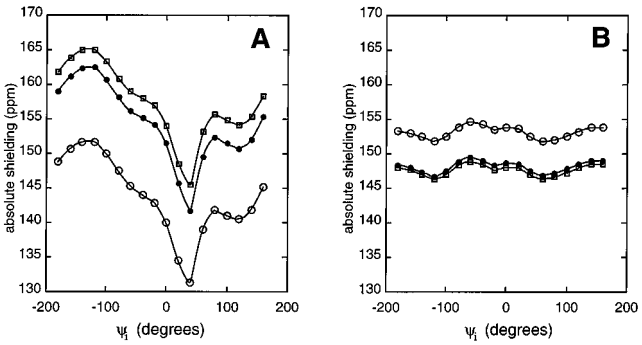


Figure 2. *N*-Formyl-alanyl-[^{15}*N*]alanine amide theoretical ^{15}N shielding ψ_i rotation plots for various backbone dihedral angles ϕ_{i-1}, ψ_{i-1}, and ψ_i. (A) (□) $\phi_{i-1} = -55°$, $\psi_{i-1} = -55°$ $\phi_i = -55°$; (●) $\phi_{i-1} = 100°$, $\psi_{i-1} = -55°$, $\phi_i = -55°$; (○) $\phi_{i-1} = -55°$, $\psi_{i-1} = 120°$, $\phi_i = -55°$. (B) (□) $\phi_{i-1} = -135°$, $\psi_{i-1} = 135°$, $\phi_i = -135°$; (●) $\phi_{i-1} = -80°$, $\psi_{i-1} = 135°$, $\phi_i = -135°$; (○) $\phi_{i-1} = -100°$, $\psi_{i-1} = -30°$, $\phi_i = -135°$.

We show schematically in Figure 1 the three peptide/amide subunits in formyl-alanyl-alanine amide in three conformations. Figure 1A has $\phi_{i-1} = -55°$, $\psi_{i-1} = -55°$, and $\phi_i = -55°$; Figure 1B has $\phi_{i-1} = 100°$, $\psi_{i-1} = -55°$, and $\phi_i = -55°$, while Figure 1C has $\phi_{i-1} = -55°$, $\psi_{i-1} = 120°$, and $\phi_i = -55°$. We evaluated the ^{15}N shieldings of each of these helix or turnlike structures as a function of the (ψ_i) rotation of the C-terminal amide plane, and the results are shown in Figure 2A. The interesting feature of these three curves is that they track each

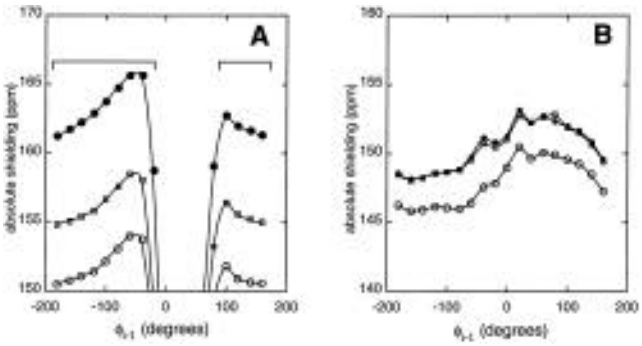


Figure 3. *N*-Formyl-alanyl-[^{15}N]alanine amide theoretical ^{15}N shielding ϕ_{i-1} rotation plots for various backbone dihedral angles (ψ_{i-1}, ϕ_i, ψ_i). (A) (□) $\psi_{i-1} = -55°$, $\phi_i = -55°$, $\psi_i = -55°$; (●) $\psi_{i-1} = -55°$, $\phi_i = -55°$, $\psi_i = 55°$; (○) $\psi_{i-1} = -55°$, $\phi_i = -55°$, $\psi_i = 120°$ (the brackets indicate regions that have favorable steric interactions). (B) (□) $\psi i_{-1} = 135°$, $\phi_i = -135°$, $\psi_i = 135°$; (●) $\psi_{i-1} = 135°$, $\phi_i = -135°$, $\psi_i = -55°$; (○) $\psi_{i-1} = -135°$, $\phi_i = -55°$, $\psi_i = 135°$.

other within about 1 ppm. This implies that changes of the preceding ϕ and ψ torsion angles affect only the absolute position of the curve and not its rotational behavior at a fixed ϕ_i. The biggest change in absolute position of the curve happens when ψ_{i-1} is changed from −55° to 120°, while a change in ϕ_{i-1} only produces about a 2−3 ppm change in absolute shielding. The deshielding of about 16 ppm when ψ_{i-1} goes from −55° to 120° agrees qualitatively with the trend of 6 ppm we reported in our previous empirical study,[12] with the smaller range there being due in part to our fitting of a large data base on a two-dimensional (ϕ_i, ψ_{i-1}) surface. The large range in shielding shown in each curve upon a full rotation cycle also of course reflects energetically unfavorable configurations which are sampled, e.g. for ψ_i from −20° to 60°, which corresponds to a drastic deshielding. Without this section of the curve, the range is reduced from 20 ppm to about 10 ppm, so the effect of ψ_i is smaller than that of ψ_{i-1}. Another set of calculations was then performed at sheet and sheetlike backbone configurations with ϕ_i fixed at −135°, and the results are shown in Figure 2B. Once again, the similarity among the curves is striking. The shielding observed in a full rotation is now only 4 ppm, however, due in this case to the extended conformation the peptide adopts. The biggest absolute shielding change of the curve position of ∼5.5 ppm is caused by changing ψ_{i-1} from 135° to −30°, while changing ϕ_{i-1} only produces a slight change in curve position. The sensitivity of ^{15}N shielding to the more distant dihedral angles, such as ψ_i, is thus strongly influenced by the actual values of the other torsion angles.

We also investigated the effects of changing ϕ_{i-1} in a systematic manner, and these results are shown in Figure 3A,B. Like the calculations performed for ψ_i, where ϕ_i was fixed at either −55° or 135°, in this case ψ_{i-1} was fixed at −55° (Figure 3A) or 135° (Figure 3B), and ϕ_{i-1} was then subsequently varied for each curve. From inspection of the results shown in Figure 3, our general conclusion is the same, that is, the shape of the curve is *only* determined by ψ_{i-1} (ϕ_i for Figure 2A,B), while ϕ_i, ψ_i (ϕ_{i-1}, ψ_{i-1} for Figure 2A,B) determine the curve position, with ϕ_i (ψ_{i-1} for Figure 2A,B) being the dominant factor, in accord with previous empirical findings.[12] The dramatic deshielding around $\phi_{i-1} = 0°$ in Figure 3A is caused by the close proximity of the oxygen atom of the preceding amide group to the H^N atom of the nitrogen atom being investigated and is unphysical. By excluding this part of the curve, the predicted range is about 5−6 ppm, again in general accord with experiment.[12] Notice in Figure 3B that the two curves which have the same ϕ_i angle but different ψ_i angles actually collapse onto each other, while in Figure 2A,B this effect is not seen. This reflects the diminished effect of ϕ_{i-1} and ψ_i dihedral angles in sheetlike configurations, the likely reason being that when ψ_{i-1} or ϕ_i is at −55° (the helical conformation), the two adjacent amide groups are closer to each other than when ψ_{i-1} or ϕ_i is at 135° or −135°. The conclusion we draw from the results of the calculations shown in Figures 2 and 3 is, therefore, that all four torsion angles need to be considered when evaluating ^{15}N chemical shieldings (or shifts)—a much more complex situation than that found for $^{13}C^{\alpha}$ or $^{13}C^{\beta}$ shifts, where typically ϕ_i and ψ_i dominate shielding. However, we can also see that it is ϕ_i and ψ_{i-1} which have the largest effects on shielding. This is in good agreement with previous results, which showed a correlation between experimental ^{15}N shifts and ϕ_i, ψ_{i-1}.[11,12] The large ϕ_i, ψ_{i-1} effect can be readily seen for example in valine residues 66 and 99 of SNase. It is observed experimentally that residue 66 is 28 ppm more deshielded than residue 99.[16] A large part of this deshielding, ∼15 ppm, can be explained by the ψ_{i-1} differences: −50° in residue 66 and 118° in residue 99, as estimated from Figure 2. Another 7 ppm in the shift difference can be attributed to the χ^1 effect, which is discussed below.

In all the above calculations, relaxation of structural parameters other than the backbone dihedral angles was not considered. To test the validity of this approach, we carried out a series of constrained geometry optimizations prior to the shielding calculations for the open box symbols in Figure 3A. The constrained geometry optimizations were performed in Gaussian-94 using a 6-31G basis set on all atoms. Structures were allowed to relax under the condition that all backbone dihedral angles were fixed. Geometry optimization reduced the range of the ϕ_{i-1} rotation curve (open boxes in Figure 3A) from 50 ppm to less than 15 ppm, and also the effect of geometry optimization was different for different parts of the rotation curve. For the part of the curve relevant to protein structures (indicated by the bracket in the figure), structural relaxation almost invariably increased the shielding by about 7 ppm, with an rmsd of about 0.4 ppm. For one section of the curve (ϕ_{i-1} around 0), the change of shielding upon optimization is large. However, due to the presence of unphysical steric interactions, this region is not of great relevance for protein structural studies. Therefore, omitting geometry optimization appears appropriate for shielding behavior at relevant geometries.

There are, in addition, at least four other factors which might, in principle, be important in determining ^{15}N shielding, although, based on the results shown in Figures 2 and 3, and the empirical correlations,[11,12] it appears that ψ_i, ϕ_{i-1} have the largest overall effects. These additional factors are the effects of alkyl substituents on the immediately preceding residue (both their nature, and conformation); the effects of side-chain conformation in the residue under consideration; the effects of hydrogen bonding to the peptide ^{15}N; and the effects of longer range electrostatic fields on shielding. To begin with, we consider the effects of incorporating a preceding valine residue, which can in principle adopt different side-chain conformations.

Side-Chain Dihedral Angle Effects. *N*-formyl-valyl-[^{15}N]-alanine amide calculations were carried out at two different backbone conformations: α-helical ($\phi_{i-1} = -55°$, $\psi_{i-1} = -55°$, $\phi_i = -55°$, $\psi_i = -55°$) and β-sheet ($\phi_{i-1} = -135°$, $\psi_{i-1} = 135°$, $\phi_i = -135°$, $\psi_i = 135°$), with a full valine $\chi^1{}_{i-1}$ rotation performed at 20° intervals. Figure 4, A and B, shows the results of these alanine ^{15}N shielding calculations for sheet and helical conformations, respectively, in which the valine χ^1 angle (of the preceding residue) is varied. The main shielding−deshielding effect of the rotation appears to correlate with the degree of overlap between one of the CH_3 groups of the valine residue and the amide group. Figure 4C,D shows how the distances of $C_\gamma{}^1$ and $C_\gamma{}^2$ to the N atom vary with χ^1, for the two backbone

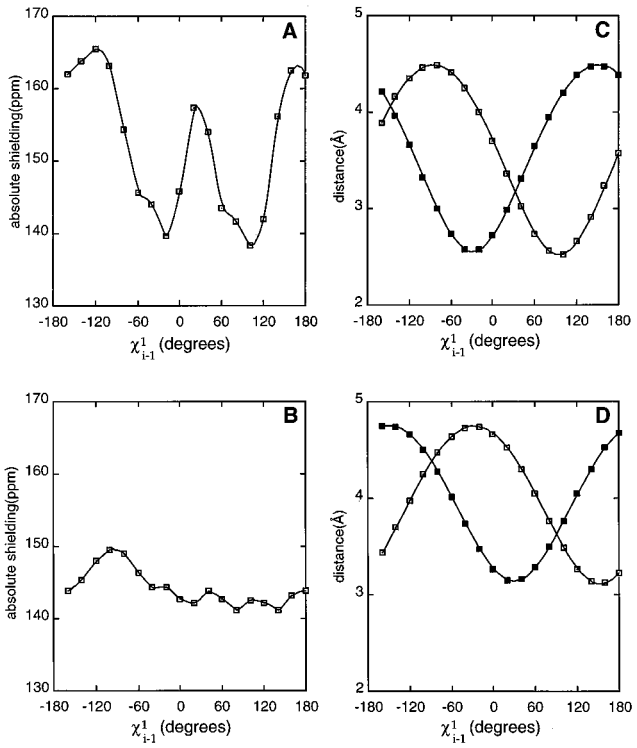


Figure 4. Effect of $\chi^1{}_{i-1}$ on *N*-formyl-valyl-[^{15}N]-alanine amide theoretical ^{15}N shielding: (A) rotation plot with backbone dihedral analges $\phi_{i-1} = -55°$, $\psi_{i-1} = -55°$, $\phi_i = -55°$, $\psi_i = -55°$; (B) rotation plot with backbone dihedral angles $\phi_{i-1} = -135°$, $\psi_{i-1} = 135°$, $\phi_i = -135°$, $\psi_i = 135°$; (C) N to $C_\gamma{}^1$ and $C_\gamma{}^2$ distance rotation plots with the same backbone conformation as in A (□, N to $C_\gamma{}^1$ distance; ■, N to $C_\gamma{}^2$ distance); (D) N to $C_\gamma{}^1$ and $C_\gamma{}^2$ distance rotation plots with the same backbone conformation as in B (□, N to $C_\gamma{}^1$ distance; ■, N to $C_\gamma{}^2$ distance).

conformations. As can be seen in Figure 4, the nitrogen is more highly shielded when $C_\gamma{}^1$ and $C_\gamma{}^2$ are furthest from the N atom. This large deshielding effect seen in the helical conformation calculation is not seen experimentally, however, most probably because it is energetically very unfavorable. For example, X-ray structures of a mostly helical protein, *Drosophila melanogaster* calmodulin,[32] show that five helical valine residues have χ^1 either close to 60° or −60°, which according to our calculations, Figure 4A, would produce a deshielding of 15−20 ppm for ^{15}N of the next residue, which is not seen experimentally. In fact, based on *J*-couplings[33] and solution C^α, C^β shifts,[34,35] these valines must actually have $\chi^1 = 180°$ in solution, with no large deshielding consequences for the following nitrogens. In the sheet conformation calculations, Figure 4B, the overall ^{15}N shift range is down to 8 ppm, and for the three most populated rotamers, the range is only ∼3 ppm. Therefore, based on this "worst case" scenario, we feel it is reasonable to use the dipeptide Ala-Ala, instead of X-Ala, where X is the actual preceding residue, in our ^{15}N shift calculations.

In summary, the effects of alkyl group conformation of the preceding residue are small, compared with the effects of the backbone torsion angles we have already considered. While certain $i - 1$ residue side-chain conformations *can* have large effects on shielding, these conformations are energetically unfavorable, and indeed given the solution ^{13}C shift and $^3J_{\alpha\beta}$ measurements on calmodulin, may be quite infrequent. Correlation of our calculational results with the effects of $i - 1$ variation in random-coil peptides[36,37] is not possible at present, since their dynamic structures are unknown.

Another potentially important factor to consider in ^{15}N shielding is the side-chain conformation of the residue in question. We therefore carried out ^{15}N shielding calculations

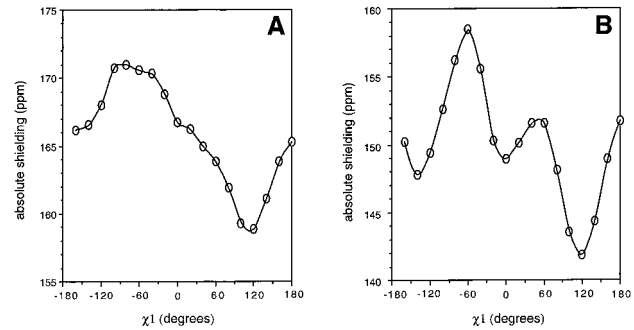


Figure 5. Effect of valine χ^1 on *N*-formyl-alanyl-[^{15}N]-valine ^{15}N shielding for (A) helical and (B) sheetlike fragments.

for an *N*-formyl-alanyl-[^{15}N]-valine amide peptide, as a function of the valine side-chain torsion angle χ^1, and the results for helical and sheetlike conformations are presented in Figure 5. As can be seen from Figure 5, both helical (Figure 5A) and sheetlike (Figure 5B) fragments have similar χ^1-shielding rotations, with the $\chi^1 = 60°$, 180° conformations having very similar shieldings, while the second most popular $\chi^1 = -60°$ conformation is ∼7 ppm more shielded. Based on the solution NMR results for calmodulin discussed above, and upon the generally low occurrence of the $\chi^1 = -60°$ conformer in X-ray structures,[34] these results indicate that in most cases χ^1 effects are unlikely to influence ^{15}N shieldings significantly. The overwhelming effect is due to the backbone conformation, with the $\chi^1 = 60°$, 180° helical fragments being more shielded than a typical sheetlike fragment, Figure 5.

One clear example of such a χ^1 effect can be seen in residues 66 and 104 of SNase, as alluded to above. The four backbone dihedral angles of these two valine residues are very similar with $\phi_{i-1} = -54.2°$, $\psi_{i-1} = -50.0°$, $\phi_i = -80.1°$, and $\psi_i = -28.2°$ for residue 66 and $\phi_{i-1} = -60.9°$, $\psi_{i-1} = -49.1°$, $\phi_i = -61.9°$, and $\psi_i = -49.4°$ for residue 104. The X-ray χ^1 conformations though are different. The χ^1 of residue 66 is −60° while that of residue 104 is 180°. The experimentally observed 8.6 ppm deshielding from residue 66 to residue 104 is thus quite well accounted for by a primarily χ^1 effect—about 7 ppm, as deduced from Figure 5A.

Hydrogen Bonding. The next factor to consider in more detail is the effect of hydrogen bonding on ^{15}N chemical shifts. In earlier work, we noted that N−O hydrogen bond lengths in proteins typically vary between 3 and 4 Å, and this can produce a shift range of up to about 3 ppm for residues in regular secondary structure. However, the effects of the *orientation* of the CO−HN hydrogen bond on shielding have not yet been explored in detail and in principle could be an important contributor to shielding nonequivalence. We have therefore investigated the change of ^{15}N chemical shielding upon varying the C−N−H^N−O dihedral angle. For residues in β-sheet structure, about 70% of residues in proteins have C−N−H^N−O dihedral angles from −75° to 60°, while in helical structures, about 90% cluster around ±100°.

We show in Figure 6A a plot of shielding vs the C−N−H^N−O dihedral angle for a backbone sheet conformation ($\phi_{i-1} = 140°$, $\psi_{i-1} = -120°$, $\phi_i = 140°$, $\psi_i = -120°$), using formaldehyde as the hydrogen-bond partner. Other parameters specifying the hydrogen-bond partner are $d(H^N{-}O) = 2$ Å, bond angles N−H^N−O and H^N−O−C of 160° and 150°, respectively, and dihedral angles N−H^N−O−C = 175° and H^N−O−C−H^a = 50°. The overall shielding range is 3 ppm, or only 2 ppm from −75° to 60°. Figure 6B shows the results of calculations for a helical backbone geometry ($\phi_{i-1} = -55°$, $\psi_{i-1} = -55°$, $\phi_i = -55°$, $\psi_i = -55°$), and bond angles N−H^N−O = 155°, H^N−O−C = 150°, with the other dihedral angles being the same as those given above. The shielding variation is more compli-

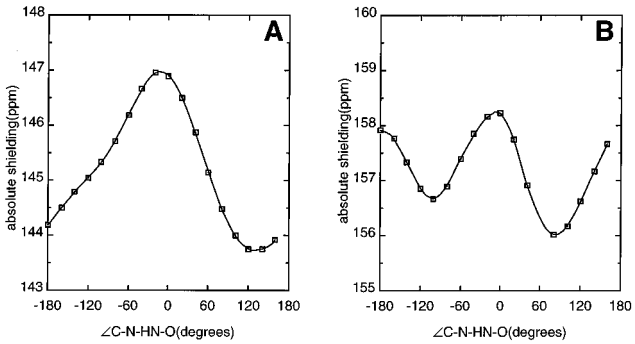


Figure 6. Effect of the $C^{\alpha}-N-H^{N}-O$ dihedral angle on ^{15}N shielding in the *N*-formyl-alanyl-[^{15}N]-alanine amide/formaldehyde dimer, for sheet- and helixlike peptides: (A) $\phi_{i-1} = -135°$, $\psi_{i-1} = 135°$, $\phi_i = -135°$, $\psi_i = 135°$, hydrogen-bond length $d_{HO} = 2$ Å, $\angle O-H^{N}-N = 160°$ and $\angle C-O-H^{N} = 150°$, dihedral angles $C-O-H^{N}-O = 175°$ and $H'-C-O-H^{N} = 50°$; (B) $\phi_{i-1} = -55°$, $\psi_{i-1} = -55°$, $\phi_i = -55°$, $\psi_i = -55°$, hydrogen-bond length $d_{HO} = 2$ Å, $\angle O-H^{N}-N = 155°$ and $\angle C-O-H^{N} = 150°$, dihedral angles $C-O-H^{N}-O = 175°$ and $H'-C-O-H^{N} = 50°$.

cated than found for the sheet geometry, but the total range is only 2 ppm, and for the two most preferred orientations, less than 1 ppm. BSSE was not corrected for in the reported shieldings, due to the our finding tht the biggest BSSE in Figure 6A was less than 0.2 ppm.

From the above model computations, we conclude that angular and hydrogen bond length distributions may cause about a 2−3 ppm fluctuation in ^{15}N chemical shielding, quite a small contribution to the total ∼20 ppm ^{15}N chemical shielding range observed experimentally. In the future, it will be of interest to investigate the effects of geometry optimization of hydrogen-bonded dimers on ^{15}N shielding. However, geometry optimization applications to proteins are much more remote, since the accuracy of torsion angles of protein structures is only ≈10−20°. This makes a more detailed comparison with experiment very difficult, since ϕ_{i-1}, ψ_{i-1}, ϕ_i, and ψ_i are all known to have very large effects on shielding, as we have shown above, and uncertainties in these parameters will tend to mask other smaller shielding contributions.

^{15}N Chemical Shifts in Proteins. In order to test the accuracy of some of the ideas we have presented above, we have investigated the ^{15}N chemical shielding of 38 alanine residues in proteins. The proteins chosen for investigation were a nuclease from *Staphylococcus aureus*, dihydrofolate reductase from *Lactobacillus casei*, and cytochrome c_{551} from *Pseudomonas aeruginosa*. Each protein has a large number of alanine residues whose ^{15}N chemical shifts and assignments have been reported, plus each protein has a high-resolution (∼1.5−1.6 Å) X-ray structure. As described in Methods, the model fragment we used is *N*-formyl-ala-ala amide. The first residue is always alanine, irrespective of the actual residue which precedes the alanine residue under investigation in the protein. The change of chemical shieldings caused by this is small (<2 ppm, data not shown). The backbone dihedral angles of the model dipeptide (ϕ_{i-1}, ψ_{i-1}, ϕ_i, ψ_i) were obtained from the reported X-ray structure of the respective protein, while the other structural parameters were again from the AMBER forcefield. On the basis of the results described above, we again did not include hydrogen-bond partners in this series of calculations.

The correlation between theory and experiment for the nonhelical alanine residues of each protein is as follows: slope = −1.1, $R^2 = 0.96$ for cytochrome c_{551}; slope = −1.2, $R^2 = 0.68$ for DHFR and slope = −0.42, $R^2 = 0.46$ for SNase. For nonhelical alanines of cytochrome c_{551} and DHFR, the correlation is slope = −1.04 and $R^2 = 0.72$ without any experimental referencing correction. After subtracting 1.4 ppm from the

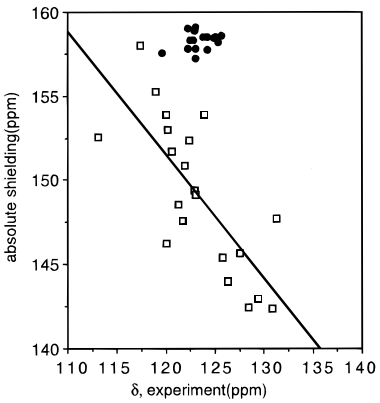


Figure 7. Plot of experimental ^{15}N chemical shifts of alanine residues in *Staphylococcal* nuclease, *L. casei* DHFR, and *P. aeruginosa* cytochrome c_{551}, vs computed absolute shieldings. Helical residues, ●; nonhelical residues, □.

experimental values of cytochrome c_{551} (based on the average helical ^{15}N shifts), the correlation became slope = −1.13 and $R^2 = 0.80$. With SNase, the slope is −0.73 and R^2 is 0.56. The highly scattered points in SNase are Ala^{90}, Ala^{109} and Ala^{112} with errors of −5.3, −4.1, and 4.5 ppm from the correlation line. These errors grow to −6.9, −8.4, and 7.3 ppm for each residue if they are omitted from the correlation line which, however, improves to slope = −1.12 and $R^2 = 0.81$. These results strongly suggest that there is a nonnegligible referencing difference between cytochrome c_{551} and DHFR, and there appears to be a structural differences between the crystal and solution structures of SNase in the region of residues 90, 109, and 112. For Ala^{109}, we found similar apparent differences between crystal and solution structure based on C^{α} and C^{β} shifts.[40] In order to more fully understand the origins of these discrepancies, it will be necessary in the future to investigate both ^{15}N and ^{13}C crystal chemical shifts.[41,42]

We find that computed helical Ala residues are more shielded than expected, by about 8 ppm (Figure 7), an effect which may be due to the presence of the helix dipole (electric field). Incorporation of point charges does bring the helix cluster into closer register with the sheet correlation, but at the expense of an increase in overall scatter. This suggests that while the basic idea of a long-range electrostatic field contribution to α-helical ^{15}N shielding is probably correct, our representation of the E-field contributions to shielding, based on a static charge field, is inadequate. In the future, it may be possible to improve this by using molecular dynamics based methods, as we have reported for ^{19}F,[8,9] but this is not practical at present since it would necessitate evaluation of $N \times M$ shieldings as a function of ϕ_{i-1}, ψ_{i-1}, ϕ_i, and ψ_i, where N is the number of points in the dynamics trajectory and M is the number of residues! Ring current effects could also in principle influence ^{15}N shielding, but there is no significant improvement in the results shown in Figure 7 when ring currents are included, presumably due to the "buried" nature of the peptide backbone nitrogen atoms. Also, as noted by Grant et al., a more detailed description of the hydrogen-bond network may bring closer accord with experiment, but the basically low resolution of protein structures precludes us from attempting this at present, unlike the situation with benzamide.[43]

Finally, we compared results from a different approach, density functional theory, which incorporates the effects of electron correlation, using the deMon program.[44] The multiple bond character of the amide moiety makes electron correlation a potentially important contributor to shielding, a factor which is neglected in the Hartree−Fock method. It has been noted in previous studies using the Møller−Plesset method that the inclusion of electron correlation produces an overall shift of all

TABLE 1: Experimental Alanine ^{15}N Chemical Shifts of Cytochrome c_{551} and the Theoretical Shieldings from GIAO and deMon-DFT Theories

	exptl shift (ppm)	GIAO shielding (ppm)	DFT shielding (ppm)
ala^{14}	123.6	150.89	121.50
ala^{17}	121.8	152.96	124.77
ala^{26}	124.7	149.12	120.86
ala^{31}	122.5	158.36	128.17
ala^{32}	119.1	158.03	128.47
ala^{35}	125.0	158.53	131.98
ala^{38}	132.6	142.35	114.30
ala^{40}	124.0	152.36	121.14
ala^{42}	122.2	159.06	129.00
ala^{45}	122.8	158.34	127.92
ala^{65}	129.2	145.63	116.35
ala^{71}	123.0	159.13	129.18
ala^{75}	124.2	158.53	128.21

shielding components;[24] however, direct comparison with experiment appears not to have been made. Here, we calculate alanine ^{15}N shieldings for cytochrome c_{551} using both SCF-HF and DFT methods. Both approaches give a good correlation between experimental and theoretical shifts for nonhelical residues. DFT shieldings have a correlation line of slope -1.04 and $R^2 = 0.97$, while GIAO-SCF shieldings have a slope of -1.11 and $R^2 = 0.96$. The intercepts of the lines are different, 288.5 ppm for GIAO-SCF and 251.3 ppm for DFT, with DFT giving a closer absolute shielding to the experimental value of 244.6 ppm.[30] The experimental values and the GIAO-SCF and DFT theoretical shieldings are listed in Table 1. Our results support the conclusion of a previous study by Sulzbach et al.[25] on the effects of the electron correlation on the amide ^{15}N chemical shift. Moreover, it appears that inclusion of electron correlation enables quite good absolute shielding predictions to be made, even for planar fragments.

Conclusions

From the results of our *ab initio* calculations, it appears that protein backbone amide ^{15}N chemical shifts are mainly determined by the backbone dihedral angles ϕ_{i-1}, ψ_{i-1}, ϕ_i, and ψ_i. ψ_{i-1} and ϕ_i dominate, agreeing with earlier empirical studies.[11,12] Our results also show that the side chain of the immediately preceding residue has a relatively small influence on shielding, presumably because conformations which could influence shielding substantially are energetically unfavorable. For valine residues, χ^1 rotations of the residue of interest can also affect shielding, but the $\chi^1 = 60°$, $180°$ conformers of valine have essentially the same shielding. Investigation of the effects of hydrogen-bond orientation also reveals a very minor influence on shielding. Our calculations of ^{15}N-alanine shielding in three proteins show a good correlation with experimental shifts for sheet residues, considerably better than that seen using empirical methods,[12] making ^{15}N shifts a potentially useful tool in structure validation, although future progress is likely to depend on the availability of crystal, rather than solution, chemical shifts.

Acknowledgment. We thank Professors P. Pulay and J. F. Hinton and Dr. K. Wolinski for providing us with a copy of their TEXAS program, and Professor D. Salahub for providing his DEMON/MASTER-CS program.

References and Notes

(1) McDonald, C. C.; Phillips, W. D. *J. Am. Chem. Soc.* **1967**, *89*, 6332.
(2) Allerhand, A.; Childers, R. F.; Oldfield, E. *Biochemistry* **1973**, *12*, 1335.
(3) Wüthrich, K. *Acta Crystallogr.* **1995**, *D51*, 249.
(4) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647.
(5) Ösapay, K.; Case, D. A. *J. Biomol. NMR* **1994**, *4*, 215.
(6) Blanco, F. J.; Herranz, J.; González, C.; Jiménez, M. A.; Rico, M.; Santoro, J.; Nieto, J. L. *J. Am. Chem. Soc.* **1992**, *114*, 9676.
(7) Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490.
(8) de Dios, A. C.; Pearson, J. G.; Oldfield, E. *Science* **1993**, *260*, 1491.
(9) Pearson, J. G.; Oldfield, E.; Lee, F. S.; Warshel, A. *J. Am. Chem. Soc.* **1993**, *115*, 6851.
(10) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311.
(11) Glushka, J.; Lee, M.; Coffin, S.; Cowburn, D. *J. Am. Chem. Soc.* **1989**, *111*, 7716.
(12) Le, H.; Oldfield, E. *J. Biomol. NMR* **1994**, *4*, 341.
(13) Loll, P. J.; Lattman, E. E. *Proteins: Struct. Funct. Genet.* **1989**, *5*, 183.
(14) Bolin, J. T.; Filman, D. J.; Matthews, D. A.; Hamlin, R. C. *J. Biol. Chem.* **1982**, *257*, 13650.
(15) Matsuura, Y.; Takano, T.; Dickerson, R. E. *J. Mol. Biol.* **1982**, *156*, 389.
(16) Yamazaki, T.; Torchia, D. Private communication.
(17) Carr, M. D.; Birdsall, B.; Frenkiel, T. A.; Bauer, C. J.; Jimenez-Barbero, J.; Polshakov, V. I.; McCormick, J. E.; Roberts, G. C. K.; Feeney, J. *Biochemistry* **1991**, *30*, 6330.
(18) Timkovich, R. *Biochemistry* **1990**, *29*, 7773.
(19) Chesnut, D. B.; Moore, K. D. *J. Comput. Chem.* **1989**, *10*, 648.
(20) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251.
(21) Pulay, P.; Wolinski, K.; Hinton, J. F. *The Texas Program*; University of Arkansas: Fayetteville, AR, 1991.
(22) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1994.
(23) Abola, E. E.; Bernstein, F. C.; Bryant, S. H.; Koetzle, T. F.; Weng, J. *Protein Data Bank in Crystallographic Databases—Information Content, Software Systems, Scientific Applications*; Data Commission of the International Union of Crystallography: Bonn, 1987.
(24) Sulzbach, H. M.; Schleyer, P. v. R.; Schaefer III, H. F. *J. Am. Chem. Soc.* **1994**, *116*, 3967.
(25) Sulzbach, H. M.; Schleyer, P. v. R.; Schaefer III, H. F. *J. Am. Chem. Soc.* **1995**, *117*, 2632.
(26) Tjandra, N.; Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6264.
(27) Malkin, V. G.; Malkina, O. L.; Salahub, D. R. *J. Am. Chem. Soc.* **1994**, *116*, 5898.
(28) Kutzelnigg, W.; Fleischer, U.; Schindler, M. In *NMR-Basic Principles and Progress*: Springer-Verlag; Heidelberg, 1990; Vol. 23, pp 165.
(29) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
(30) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, U.K., 1989.
(31) Jameson, C. J.; Jameson, A. K.; Cohen, S. M.; Parker, H.; Opposunggu, D.; Burrell, P. M.; Wille, W. *J. Chem. Phys.* **1981**, *74*, 1608.
(32) Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quiocho, F. A. *J. Mol. Biol.* **1992**, *228*, 1177.
(33) Ikura, M. Private communication.
(34) Ikura, M.; Spera, S.; Barbato, G.; Kay, L. E.; Krinks, M.; Bax, A. *Biochemistry* **1991**, *30*, 9216.
(35) de Dios, A. C.; Oldfield, E. *J. Am. Chem. Soc.* **1994**, *116*, 5307.
(36) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67.
(37) Braun, D.; Wider, G.; Wüthrich, K. *J. Am. Chem. Soc.* **1994**, *116*, 8466.
(38) Ponder, J. W.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 775.
(39) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135.
(40) Pearson, J. G.; Wang, J.-F.; Markley, J. L.; Le, H.; Oldfield, E. *J. Am. Chem. Soc.* **1995**, *117*, 8823.
(41) Sparks, S. W.; Cole, H. B. R.; Torchia, D.; Young, P. E. *Chem. Scr.* **1989**, *29A*, 31.
(42) Cole, H. B. R.; Sparks, S. W.; Torchia, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6362.
(43) Facelli, J. C.; Pugmire, R. J.; Grant, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 5488.
(44) St-Amant, A.; Salahub, D. R. *Chem. Phys. Lett.* **1990**, *169*, 387.

JP9606164