Quantum Chemical Study on Oxygen-17 and Nitrogen-14 Nuclear Quadrupole Coupling Parameters of Peptide Bonds in α-Helix and β-Sheet Proteins

Maricel Torrent,[†] Danielle Mansour,[‡] Edmund P. Day,[‡] and Keiji Morokuma*,[†]

Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322, and Department of Physics, Emory University, Atlanta, Georgia 30322

Received: December 7, 2000; In Final Form: February 22, 2001

Theoretical calculations of ^{17}O and ^{14}N nuclear quadrupole coupling (NQC) constants (χ) and asymmetry parameters (η) for small α-helix and β-sheet protein fragments have been carried out using the density functional theory. This computational study is intended to shed light on the differences between the two major structural elements found in the secondary structure of proteins. Specific NQDR spectra are computationally simulated for the ^{17}O and ^{14}N nuclei inherent in protein backbones. The separate signals resulting from α-helices and β-sheet models are predicted to be experimentally distinguishable for ^{17}O but not for ^{14}N. In particular, we predict that the differences in χ (in MHz) between α-helix and β-sheet proteins in solution are $\Delta\chi(^{17}O) = 0.53(15)$ and $\Delta\chi(^{14}N) = 0.14(16)$, with the standard deviations in parentheses. It is found that ^{17}O NQC parameters of proteins are dependent on the particular conformation of the backbone, specifically on the hydrogen bond angle $\theta = \angle H-N\cdots O$ and the backbone dihedral angle $\psi = \angle NC-C(O)N$. Due to this, ^{17}O NQC parameters are observably different in α-helices and β-sheets. Conversely, ^{17}O NQC parameters are not dependent on the length of the hydrogen bond $R_{O\cdots N}$, as had been previously thought, nor are they dependent on either the hydrogen bond dihedral angle $\xi = \angle N-C{=}O\cdots H$ or the backbone dihedral angle $\phi = \angle C-(O)C-NC(O)$. We also conclude that, unlike ^{17}O NQC parameters, ^{14}N NQC parameters of proteins are within the uncertainties identical for both α-helices and β-sheets. Finally, differing residues on protein side chains do not significantly affect the NQC parameters of the backbone C=O and NH groups, and can be modeled computationally by using glycine.

I. Introduction

The folded structures of proteins are composed of simpler structural elements.[1] The most common element found in the secondary structure of proteins is the α helix, comprising on average approximately one-third of all residues in globular proteins. The second major structural element is the β sheet; β-sheet formation plays a critical role in many diseases, including AIDS, Alzheimer's disease, and prion diseases.[2] Helices and sheets are markedly different. However, both types are characterized by having main chain NH and CO groups participating in hydrogen bonds to each other, so that polar atoms are neutralized and the hydrophobic core of the protein is preserved.[3] With the purpose of learning about protein folding (and ultimately developing new drugs, molecular receptors and catalysts), scientists have begun to investigate compounds that fold to mimic these two structural elements. Spectroscopic studies on biomimetic compounds should provide insight into the factors affecting the conformation of the main chain, and help identify fundamental dissimilarities between different structural themes in proteins. The ultimate goal is to study real proteins.

Over the past 15 years, several spectroscopic techniques formerly used to gain information on organic compounds have begun to be successfully employed in biochemical contexts as well, and more specifically in the field of conformational analysis of proteins. For instance, it has been demonstrated that high-resolution solid-state ^{13}C NMR spectroscopy can be used for conformational characterization of polypeptides and proteins in the solid state because the ^{13}C NMR chemical shifts depend on their main-chain conformations.[4] More recent investigations have shown that the ^{15}N isotropic chemical shifts of solid polypeptides are also sensitive to the protein secondary structure (such as right-handed α-helix, antiparallel β-sheet, and other forms), and therefore very useful barometers for the conformational analysis of proteins.[5,6] Besides nitrogen, an equally important atom involved in hydrogen-bonding structure in proteins is oxygen. Unfortunately, the only NMR active isotope of oxygen, ^{17}O ($I = {}^5/_2$), has a low natural abundance (0.037%). This, together with the large quadrupolar interaction experienced by ^{17}O in many solids (including proteins), makes observation and interpretation of ^{17}O NMR spectra difficult, and explains why solid-state ^{17}O NMR studies on proteins are more scarce than ^{13}C and ^{15}N NMR studies.

Another classical spectroscopic technique that has just begun to be employed in biochemical contexts is nuclear quadrupole double resonance (NQDR).[7] One of the most important applications of this experimental method is to study the active sites of metalloproteins, gaining insight into the oxidation states, ligands, coordination geometry, and local magnetic environment of these sites. Interestingly, NQDR also has the potential sensitivity to detect small quantities of quadrupolar nuclei in real proteins. This application, however, is yet to be explored. To correctly interpret the quadrupole spectrum to find the nuclear quadrupole coupling (NQC) parameters, i.e., the NQC constant (NQCC or χ), asymmetry parameter (η), and hyperfine

[†] Cherry L. Emerson Center for Scientific Computation and Department of Chemistry.
[‡] Department of Physics.

10.1021/jp004398x CCC: $20.00
Published on Web 04/06/2001

splitting, at the site of the quadrupolar isotope, first one must be able to distinguish the desired signal from all the background signals. Since proteins are often studied in solution, a very common background signal is that of the ^{17}O contained in water. This signal has been carefully studied both experimentally[8] and theoretically,[9] paving the way for the next step: identifying the signals that result from the ^{17}O and ^{14}N of the peptide backbone.

A few recent experimental studies have investigated the ^{17}O NQC parameters in helices and sheets through both solid state and magic-angle spinning (MAS) NMR spectroscopy.[10,11] It should be noted that these studies did not investigate real proteins, but specially manufactured ^{17}O-labeled poly(L-alanine) chains. The advantage of these chains is that they can be entirely either α-helical or β-sheet. Individual backbone amide NQC values in low molecular weight proteins such as ubiquitin have been measured by solution NMR by LiWang and Bax.[12] They have shown that amide deuteron NQCCs are quite sensitive to hydrogen bonding thereby providing unique spectroscopic probes for studying this interaction in solution. In particular NQCCs calculated from $^2H^N$ T_1 in ubiquitin are found to correlate with the inverse cube of the X-ray crystal-derived hydrogen bond lengths.

A powerful tool that has been proven to be complementary to experimental techniques in many areas is molecular modeling.[13] Recent improvements in quantum chemical methodologies, combined with speed-up in computer hardware, have turned the computer simulation into an ideal partner whenever experiment becomes too difficult, too expensive, too dangerous or simply impractical. A modest contribution from computer simulations to crystallographers and spectroscopists currently working with proteins has been in facilitating refinement of existing solution structures through theoretical/computed estimates. The major contribution from quantum chemistry, however, comes from its predictive character. Computational studies on polypeptides permitted the first successful predictions of the ^{13}C, ^{15}N, and ^{19}F NMR spectra of proteins in solution only a few years ago.[14] Since then, other theoretical studies have provided refinement or additional evidence to support experimental findings, if not accurate predictions in all cases.[15]

By modeling NMR experiments, interesting helix-sheet differences have been recently reported in an ab initio quantum chemical investigation of ^{13}C NMR shielding tensors.[16] It was found that C_α in sheet structures are on average more shielded than in helical structures (by about 4−5 ppm),[16] in good agreement with experimental results.[17] It was also reported that, on average, the overall breadths of the $^{13}C_\alpha$ shielding tensors in sheet residues are about 50% larger than the values computed for helical residues which, unlike the previous finding, was a totally unexpected result.

Also, related theoretical studies on both $\chi(^{17}O)$ and $\chi(^{14}N)$ have been performed to confirm/complement the existing available data from solid state and MAS NMR spectroscopy, but here again, like in the experiments,[10,11] these studies concerned the calculation of the NQC parameters in synthetic polypeptide structures other than α-helix or β-sheet proteins.[18]

To the best of our knowledge, so far there have been no theoretical studies of the NQC parameters of the naturally abundant ^{17}O and ^{14}N in real protein backbones, specifically α-helices and β-sheets. Several important questions remain to be elucidated. For instance, which are the regions of the spectrum where the ^{17}O and ^{14}N signals for both α-helices and β-sheets can be confidently predicted to appear? Are ^{17}O signals for α-helices expected to be experimentally distinguishable from ^{17}O beta-sheet signals? Should the ^{14}N spectra of α/β proteins be expected to show peaks sufficiently isolated to be experimentally resolvable in distinct components (coming from either one conformation or the other)?

With the aim of answering these questions, in this paper we investigate by means of quantum chemical methodologies the NQC parameters of ^{17}O and ^{14}N in the peptide bond of several α-helices and β-sheets extracted from real metalloproteins. In a recent work,[19] we have already shown the suitability of quantum chemical methods to evaluate NQCCs in large systems of biological importance. In particular, the ^{14}N NQCC of the distal nitrogen atom in coenzyme B_{12} was successfully reproduced by our calculations, and a prediction was also made for the (still unavailable) NQCC of the proximal N in the studied coenzyme.[20]

Our goal in the present investigation is, more specifically, to see, first, if model proteins with a high content of alpha secondary structure and model proteins with a high content of beta are distinguishable through their corresponding theoretical ^{17}O and/or ^{14}N quadrupole spectra. Second, it would be desirable to quantify the existing difference, understand its origin, and finally provide some guidelines that may help resolve the signals of complex NQDR spectra in future experiments with proteins (and metalloproteins). It should be noted that real samples of proteins are rarely 100% helical or 100% sheet; very often they contain mixed α/β structures. The advantage of using molecular modeling is that pure alpha polypeptides or pure beta polypeptides can be designed and studied, and their characteristics separately analyzed, avoiding the intrinsic complexities found in experiment while keeping the features of the target system.

II. Computational Details

A. Evaluation of NQCCs. The formulation employed in the evaluation of NQC parameters can be found elsewhere.[21] Briefly, the electric field gradient (EFG) is a traceless, symmetric second-rank tensor whose principal axes are chosen so that its components satisfy[22] $|q_{zz}| \geq |q_{yy}| \geq |q_{xx}|$ ($q_{ij} = \partial^2 V/\partial i \partial j$ where $i, j = x, y,$ and $z,$ and V is the external electrostatic potential). The quantities usually determined experimentally are the NQC constant, $\chi = e^2Qq/h$ (where Q is the nuclear electric quadrupole moment and $q = q_{zz}$), and the asymmetry parameter, $\eta = |(q_{yy} - q_{xx})/q_{zz}|$. Like in many previous studies,[23] here we assume that the nuclear quadrupole moment acts as a simple constant or scaling parameter, and we do not parametrize it (as done in some other works[24]). Among the wide range of $Q(^{17}O)$ and $Q(^{14}N)$ standard values published,[25] we have selected the recent values $Q(^{17}O) = -25.58 \times 10^{-31}$ m^2 and $Q(^{14}N) = 20.44(3) \times 10^{-31}$ m^2 reported, respectively, by Pyykkö et al.[26] and Tokman et al.[27]

To induce transitions between the energy levels of a given nucleus, an oscillating magnetic field is usually applied. If the field oscillates at an appropriate frequency, it interacts with the magnetic dipole moment of the nucleus, causing a time-dependent perturbation. The formulas to calculate the matrix elements of the Hamiltonian operator for pure quadrupole resonance have been given elsewhere. These formulas indicate that there is a mixing of states with $\Delta m = 2$ (m, angular momentum), so that the quadrupole interaction matrix must be diagonalized to find the interaction energy eigenvalues. The resulting secular equations for nuclei of different spins, I, have been worked out by Bersohn,[28] Dean,[29] and Cohen.[30] Here we focus on $I = 1$ (^{14}N) and $I = {}^5/_2$ (^{17}O). For $I = 1$, there are three energy levels corresponding to $m = 1, -1,$ and 0 The frequencies for transition between these levels

are found by evaluating:

$$\nu_o = (E_1 - E_{-1})/h = {}^1/_2\chi\eta \quad (1)$$

$$\nu_+ = (E_1 - E_o)/h = {}^3/_4\chi(1 + \eta/3) \quad (2)$$

$$\nu_- = (E_{-1} - E_o)/h = {}^3/_4\chi(1 - \eta/3) \quad (3)$$

For $I = {}^5/_2$, there are three energy levels corresponding to $m = {}^1/_2$, ${}^3/_2$, and ${}^5/_2$. Unlike the case of $I = 1$, there is no way to obtain exact solutions for these levels in analytical form. Instead, the secular equations have to be solved numerically, and two methods have been used. In the first, the secular equation has been solved numerically with double precision, and the relative energy eigenvalues tabulated for varying η.[31] A second method of solving the secular equation is numerical perturbation. This approximates the eigenvalues in the form of series expansions of η, and can be used when η is sufficiently small or large. For $\eta < 0.5$, numerical perturbation results in the frequencies

$$\nu_1 = {}^3/_{20}\chi(1 + 1.0926\eta^2 - 0.6340\eta^4) \quad (4)$$

$$\nu_2 = {}^6/_{20}\chi(1 - 0.2037\eta^2 + 0.1622\eta^4) \quad (5)$$

B. Choice of Structural Models. When planning the ideal polypeptide model system, certain characteristics have to be kept in mind. The model should be as small as possible, while retaining all the features of the structural element studied (either the α helix or the β sheet), and with the minimum number of interactions necessary for structure formation. On the basis of this criterion, the calculations were performed on polypeptide structures containing six peptide bonds and three H-bond interactions each, and a total of seven C_α atoms for each α-helical fragment and eight C_α atoms for each antiparallel β-sheet fragment. All ends of the structures (two ends for each alpha fragment and four for each beta) terminated with a methyl group.

The geometries employed were divided into two groups: (1) experimental (from the Protein Data Bank, see below) and (2) computer-designed geometries. To simulate α-helical peptides using experimental geometries, a total of five fragments were selected from and cut out of the available X-ray structure of myoglobin (metalloprotein with a 81% contents of alpha). For the antiparallel β-sheet peptides, a total of five fragments were selected from two different proteins: the N-terminal domain of a transcription initiation factor, pftfiibn,[32] and a 178-residue hydrolase, endo-1,4-β-xylanase. The amino acid sequences for each of the 10 selected fragments (five alpha and five beta) are listed in Table 1.[33] All fragments were converted from the original experimental data into polyglycine chains, keeping the original nuclear coordinates but excluding the residues. Representative alpha and beta fragments are shown in Figure 1.

To make our study as general as possible, we also studied several computer-designed structures having the features of either a helix or a pleated sheet, as shown in Table 2. The ϕ and ψ angles for the α-helix models were constrained to the experimental values of −63.0° and −42.0°, respectively,[34] to ensure that the chains would stay in a helical conformation. Concerning β-sheets, the simplest way to bring two antiparallel strands together is a short peptide segment between the C-terminus of one strand and the N-terminus of the other. As noted in previous studies, however, constructing an antiparallel β-sheet structure from a random polypeptide chain by computer simulations is not straightforward.[35] Unlike the case with an α-helix in which only local interactions are important, a β-sheet can arise not only from local interactions but also from nonlocal interactions, so here assembling correctly all intrachain hydrogen bonds and other components becomes critical. All polyglycine structures were first fully optimized (in the case of α-helices, only partially as will be discussed later), and then the EFG was calculated on the optimal geometries. These computer-designed

TABLE 1: Description of Amino Acid Sequences of Selected α-helical and β-sheet Fragments Taken from Experimental Three-Dimensional Structures

fragment	protein[a]	range of residues	sequence
α1	1A6M	14−20	Asp-Ala-Glu-Val-Lys-Ala-Trp
α2	1A6M	20−26	Asp-Val-Ala-Gly-His-Gly-Gln
α3	1A6M	27−33	Asp-Ile-Leu-Ile-Arg-Leu-Phe
α4	1A6M	87−93	Lys-Pro-Leu-Ala-Gln-Ser-His
α5	1A6M	108−114	Ser-Glu-Ala-Ile-Ile-His-Val
β1	1PFT	17−20; 23−26	Ile-Tyr-Asp-Pro; Gly-Glu-Ile-Val
β2	1XYN	7−10; 34−37	Gln-Asn-Tyr-Gln; Phe-Val-Val-Gly
β3	1XYN	17−20; 23−26	Tyr-Ser-Pro-Ser; Gly-Phe-Ser-Val
β4	1XYN	67−70; 73−76	Gly-Trp-Ser-Thr; Leu-Val-Glu-Tyr
β5	1XYN	107−110; 120−123	Asn-Thr-Arg-Val; Thr-Phe-Asn-Gln

[a] Notation as in the PDB files: 1A6M = myoglobin (crystal at 1.0 Å resolution), 1PFT = Transcription Factor IIB−N terminal domain (solution NMR), 1XYN = *endo*-1,4-Xylanase I (crystal at 2.0 Å resolution).

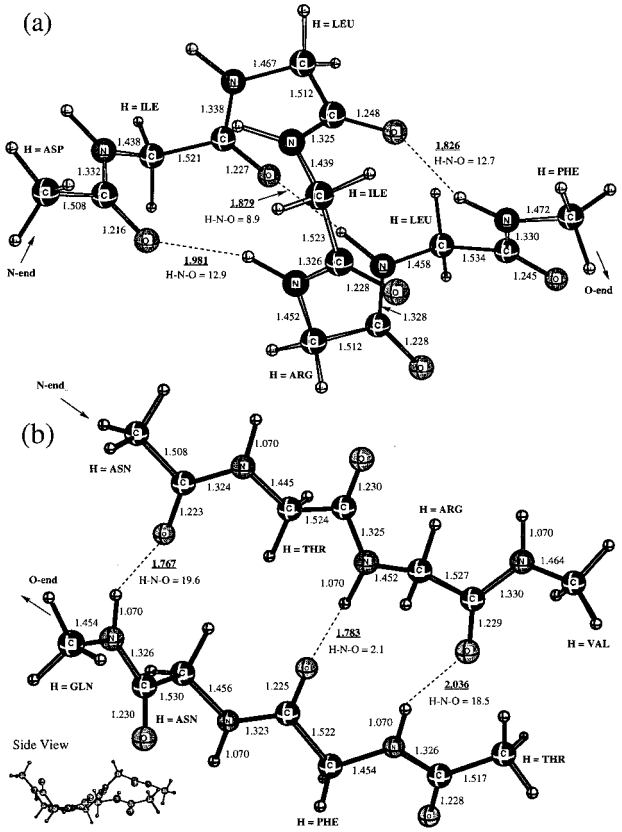

Figure 1. Experimental molecular structure of representative α-helix (a) and β-sheet (b) protein fragments. They correspond, respectively, to fragments α3 and β5 as described in Table 1.

TABLE 2: Fragments Constructed Using Molecular Builder Programs, and Optimization Methods Employed to Obtain Optimal Structures

fragment	optimal method
α6	B3LYP
β6	PM3
β7	AM1

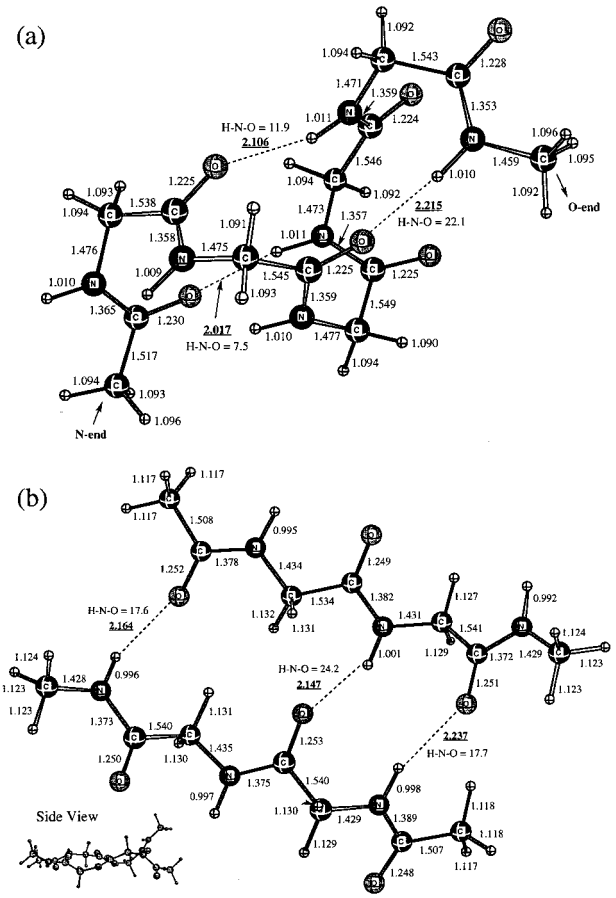

Figure 2. Optimized molecular structure of representative α-helix (a) and *β*-sheet (b) protein fragments. They correspond, respectively, to fragments α6 and *β*7 as described in Table 2.

structures will be referred in the next sections as "generic alpha" and "generic beta". Representative alpha and beta fragments with optimized structures are shown in Figure 2.

C. Inclusion of Solvent Effect. Solvent effects were taken into account in selected cases by means of reaction field calculations using the polarizable continuum model (PCM) of Tomasi and co-workers.[36] A dielectric constant of 78.4 (water) was employed. Simulations involving explicit water molecules were not considered because there is no direct solvent in the immediate vicinity of the H-bonds stabilizing the helix or the sheet of a protein. In this sense, the PCM gives a more accurate picture of the real situation in a protein. An additional advantage of using PCM is that not only solvent effects but also other effects of the environment around the studied α-helix or *β*-sheet fragment (side chains or other parts of the protein) can be included implicitly under the long-range averaged interaction modeled by PCM.

D. Method of Calculation. Unless otherwise indicated, geometry optimizations were carried out at the B3LYP/6-31G level for all atoms except N−H and C═O groups where a polarization function was added (6-31G*). Two semiempirical methods (AM1 and PM3) were additionally used for the "generic beta" (*β*6 and *β*7) for geometry optimization for the purpose of comparison. Single-point EFG calculations at PW91P86/6-311G level (6-311G* for N−H and C═O) were performed on (1) the previously optimized structures for the "generic" fragments, and (2) the experimental geometries extracted from selected fragments of the real metalloproteins (some H atoms had to be added to keep all fragments neutral). The use of PW91P86/6-311G(*) to properly evaluate EFGs has been discussed in a previous paper.[19]

For the systematic investigation of the dependence of the NQC parameters on the $R_{O\cdots N}$ bond distance, bond angle, dihedral angle, and phi and psi angles on a given α-helix/*β*-sheet fragment, a semiempirical method (PM3) was used to perform the partial optimizations, accompanied by single-point EFG calculations at PW91P86/6-311G level (6-311G* for N−H and C═O) level.

To assess the effect of the side chains on the EFG of the oxygen atoms in the backbone, a larger model system including all residues in one of the experimental fragment was also tested (*β*5-R). In this case, a two-layer ONIOM calculation was performed on the experimental geometry of the selected fragment, using (PW91P86/6-311G:PM3) with 6-311G* for N−H and C═O in the inner layer. The simple polyglycine version of the fragment (*β*5) was employed as the inner layer, and the C_β atoms served as links between the two layers. A small code was developed to evaluate the effective EFGs from the ONIOM output.[37] All calculations were carried out with the Gaussian 98 package.[38] An energy decomposition analysis (EDA) for both α-helix and *β*-sheet fragments was made using the Amsterdam Density Functional program.[39]

III. Results and Discussion

This section is divided in four subsections. First a comparison of α-helix and *β*-sheet fragments in terms of energy contributions to the total energy is presented. Second, we present the results for ^{17}O, both in the gas phase and in solution, for α-helices and *β*-sheets. Here we evaluate the role of solvent and the effect of protein residues, and we seek to find the main factors that differentiate α-helices and *β*-sheets. Third, we present the results for ^{14}N and compare them with the results for ^{17}O. Finally we predict the NQDR spectra of transition frequencies for both an α-helical and a *β*-sheet protein.

A. Energy Decomposition Analysis. An important aspect of any energy decomposition analysis (EDA) concerns the definition of the two fragments, A and B, in which the system A−B is divided. For a *β*-sheet, the choice of the two fragments is quite obvious. Our *β*-sheet model can be regarded as a supermolecule, the sum of two separate strands, so each strand is a well-defined fragment. For an α-helix, however, the partitioning is not so straightforward because all atoms belong to the same chain. A simple cut on one of the covalent bonds would lead to two radical fragments, which are not a suitable input for the EDA (requiring two saturated entities or molecules). To overcome this difficulty we removed the central peptide unit of the chain (i.e., −N(H)−C(═O)−), and then we saturated the two new ends turning them into methyl groups. Finally we put a constraint to keep the two consecutive helices as if they were still part of a larger one (frozen C−C distance = 3.734 Å and dihedral angles, all as they were in the original model). This truncated α-helix model is the ideal solution that satisfies both the physical and chemical requirements of the system under analysis. Moreover, it is also consistent with the *β*-sheet model to be compared with because the donor (NH) and the acceptor (C═O) groups of a given H-bond interaction are located in separate entities (either a helix or a strand) in both models.

Table 3 shows the individual terms of the energy decomposition (in kcal/mol) for an α-helix and a *β*-sheet. According to the extended transition state method,[40] the total energy of a given system can be written as

$$E_{\text{total}} = E_{\text{int}} + E_{\text{def}} = E^{\circ} + E_{\text{oi}} + E_{\text{def}} = E_{\text{Pauli}} + E_{\text{elstat}} + E_{\text{oi}} + E_{\text{def}}$$

TABLE 3: Terms of the Energy Decomposition (in kcal/mol) for an α-Helix and a *β*-Sheet Fragment

energies	α-helix	*β*-sheet
ΔE_{Pauli}	16.47	5.71
ΔE_{elstat}	−23.00	−15.45
$\Delta E^{\circ a}$	−6.53	−9.74
ΔE_{oi}	−13.90	−7.60
$\Delta E_{\text{int}}{}^{b}$	−20.43	−17.34
ΔE_{def}	22.75	4.72
$\Delta E_{\text{total}}{}^{c}$	2.32	−12.62

[a] Sum of ΔE_{Pauli} and ΔE_{elstat}. [b] Sum of ΔE° and ΔE_{oi}. [c] Sum of ΔE_{int} andΔE_{def}.

Here the deformation energy term, E_{def}, takes into account the geometric distortion of the two fragments in each model (either an α-helix or a *β*-sheet) from their equilibrium geometries to the framework of the final system (i.e., of the entire model). E_{int} represents the net interaction between the two fragments in each model once their geometries are as in the final complex. The E_{int} term is divided into the electrostatic and exchange interaction energy between the two fragments in the combined complex, E°, and the orbital interaction energy, E_{oi}. E° can be further divided into the electrostatic interactions, E_{elstat} and the exchange repulsion, E_{Pauli}. The term E_{oi} originates from stabilizing interactions between occupied and virtual orbitals of the two separate fragments.

As seen from Table 3, the total energy for the α-helix model is positive (+2.32 kcal/mol) whereas the total energy for the *β*-sheet model is negative (−12.62 kcal/mol). This indicates that the model with a helical array is not stable in the gas phase (i.e., without any environmental element stabilizing the tightly coiled structure). The main contribution to this destabilization arises from the deformation energy. The E_{def} for the α-helix (22.75) is almost five times the E_{def} for the *β*-sheet (4.72). Therefore, although the E_{int} is larger in the α-helix (−20.43) than in the *β*-sheet (−17.34), this term cannot compensate for the large deformation energy in the former model. For this reason, in the α-helix models described in the next sections we have kept several dihedral angles frozen (partial optimizations) to preserve the natural rodlike structure.

Also it should be noted that the E_{oi} term is remarkably larger in the α-helix than that in the *β*-sheet. This reveals that the stabilizing interactions between the two fragments (mainly H-bond interactions) are better in the coiled structure than in the pleated sheet, probably as a result of more linear N−H···O bonds. This, in turn, should be reflected in a spatial distribution of electrons around N and O nucleus notably different for α-helices and for *β*-sheets. In the next sections, we investigate theoretically whether ^{17}O and/or ^{14}N NQDR spectroscopies will be able to detect these differences.

B. ^{17}O NQC Parameters. Table 4 shows the calculated average ^{17}O NQC χ and η for each fragment, both in the gas phase (vacuum, $\epsilon = 1.0$) and in solution (water, $\epsilon = 78.4$). All six oxygen atoms have been taken into account in the average of any given single fragment. Even though only half of the oxygen atoms in these models are H-bonded, all have been included in the calculations. Exclusion of the atoms that are not H-bonded does not significantly affect the result. They are equally distributed above and below the average. The effect of introducing these atoms can be regarded as analogous to the effect of having random side chains and other external factors influencing the final value (an effect otherwise neglected by using isolated polyglycine models).

First, it should be noted that the difference between NQC parameters calculated at experimental geometries (α1−α5 and *β*1−*β*5) and parameters obtained at optimized geometries (α6 and *β*6, *β*7) is negligible. Earlier studies[41] showed that the EFG is very sensitive to geometry changes and that the best results are often obtained using the experimental geometry. Our results in Table 4 are consistent with a more recent theoretical study on the nitroethylene molecule[42] where minor differences were found between the values obtained at the experimental and optimized geometries.

TABLE 4: Average Calculated ^{17}O NQC Constants χ (in MHz, Standard Deviation in Parentheses) and Asymmetry Parameters η of Gas-Phase and Solvated α-Helix and *β*-Sheet Protein Fragments Taken for Both Experimental and Optimized Geometries

	gas phase ($\epsilon = 1.0$)		solution ($\epsilon = 78.4$)		difference (solution − gas phase)	
fragment	χ	η	χ	η	$\Delta\chi$	$\Delta\eta$
α1	9.388	0.234	8.983	0.344	−0.405	+0.110
α2	9.203	0.158	8.878	0.252	−0.325	+0.094
α3	9.406	0.142	9.029	0.256	−0.377	+0.114
α4	9.175	0.086	8.793	0.196	−0.382	+0.110
α5	9.370	0.148	8.950	0.273	−0.420	+0.125
α6	9.277	0.072	8.925	0.179	−0.352	+0.107
av α	9.30(9)	0.14(5)	8.93(8)	0.25(5)	−0.37(12)	+0.11(7)
*β*1	8.952	0.213	8.585	0.464	−0.367	+0.251
*β*2	8.795	0.312	8.434	0.680	−0.361	+0.368
*β*3	8.823	0.240	8.461	0.523	−0.362	+0.283
*β*4	8.684	0.240	8.328	0.523	−0.356	+0.283
*β*5	8.651	0.273	8.296	0.595	−0.355	+0.322
*β*6	8.888	0.109	8.524	0.238	−0.364	+0.129
*β*7	8.524	0.319	8.175	0.695	−0.349	+0.376
av *β*	8.76(14)	0.24(7)	8.40 (13)	0.53(14)	−0.36(19)	+0.29(16)
$\Delta(\lvert\alpha-\beta\rvert)$	0.54(17)	0.10(9)	0.53(15)	0.28(15)		

Second, the calculated average gas-phase χ using both the experimental and optimized geometries are $\chi_{\alpha} = 9.30$ (9) MHz and $\chi_{\beta} = 8.76$ (14) MHz. As seen from Table 4, the computed values are remarkably consistent, and the largest deviation from the average is only 2.7% (in the case of the AM1 optimized *β* fragment, *β*7). These results correlate well with those experimentally determined by Ando and co-workers[10,11] ($\chi_{\alpha} = 9.28$ MHz and $\chi_{\beta} = 8.65$ MHz).

Third, the calculated average χ in solution, which should provide a more accurate prediction than the gas-phase results for the signals in real proteins to be determined in future experiments, are $\chi_{\alpha} = 8.93$ (8) MHz and $\chi_{\beta} = 8.40$ (13) MHz. These values are obtained averaging both the experimental and optimized geometries. It should be noted that, regardless of the medium being vacuum or water, the calculated difference in χ between α-helices and *β*-sheets is approximately half a MHz: 0.53 (15) MHz in solution, 0.54 (17) MHz in the gas phase. Absolute values are slightly shifted from vacuum to water; relative differences stay the same. With line widths typical of frozen samples at 2K, NQDR spectrometers can resolve differences well below 100 kHz. Therefore, our results show that separate signals resulting from α-helices and *β*-sheet models will be experimentally distinguishable.

Fourth, the average change in the NQCCs between the gas-phase and solvated protein fragments in Table 4 is $\Delta\chi = -4.1\%$, with the negative sign indicating that χ decreases from the gas phase to solution. The effect of solvation on EFG's has been theoretically investigated in previous studies with a supermolecule approach for a few neutral molecules (H_2O, HCOOH, H_2CO, CH_3OH, and $HCONH_2$).[43] Those studies showed that hydrogen-bonding interactions cause a 10−30% decrease in EFG with respect to the isolated molecule. Here we do not use explicit molecules but a continuum model. A more suitable reference is our previous study on coenzyme B_{12},[19] where the

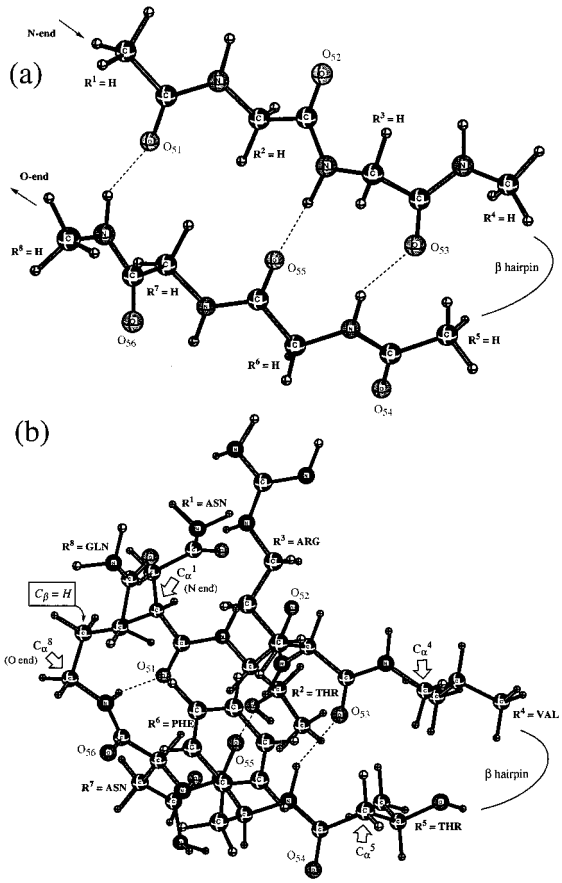

Figure 3. Comparison of a beta-sheet fragment, β5, without (a) and with (b) residues. The relevant oxygen atoms have been labeled as O_{xy}, where the first digit of the subscript, x, indicates the fragment (5), and the second digit, *y*, indicates the corresponding O atom in the fragment (starting from the N-terminus).

TABLE 5: Comparison of χ (in MHz) and η for Fragment β5 with Residues (β5-R) and without (β5-H; i.e., R = H)

O atom	$\chi(\beta5\text{-H})$	$\chi(\beta5\text{-R})$	$\Delta\chi^a$	$\eta(\beta5\text{-H})$	$\eta(\beta5\text{-R})$	$\Delta\eta^b$
O_{51}	8.333	8.538	+0.205	0.423	0.383	−0.040
O_{52}	8.312	8.452	+0.140	0.252	0.212	−0.040
O_{53}	8.690	8.634	−0.056	0.283	0.323	+0.040
O_{54}	9.064	9.077	−0.013	0.130	0.129	−0.001
O_{55}	8.537	8.730	+0.193	0.425	0.357	−0.068
O_{56}	8.972	8.691	−0.281	0.123	0.190	+0.067

[a] Average deviation $\Delta\chi = 0.003$ MHz [b] Average deviation $\Delta\eta = 0.007$

decrease in EFGs from the gas phase to PCM-simulated solution was found to be in the range of 2−9%.

Role of Protein Residues. The role played by the side chains has been investigated by comparing the NQCCs in a selected fragment (β5) with and without residues, as shown in Figure 3. The results are shown in Table 5. As can be seen, substitution of the real side chains by hydrogen atoms results in an average $\Delta\chi$ of only ±0.003 MHz (less than 1% of the $\Delta\chi_{\alpha,\beta}$, difference at averaged values for α and β, 0.54 MHz). The maximum deviation is approximately ±0.2 MHz, and occurs at an oxygen surrounded by two amide ligands (O_{56}). The average $\Delta\eta$ is also very small (±0.07). It follows from these results that the protein residues do not significantly affect the EFG components of the main-chain atoms. For the purpose of the present study, the protein residues can be safely replaced by glycines. A polyglycine model might not be accurate enough for a more sophisticated study of a particular system where subtle effects from the side chains may be important.

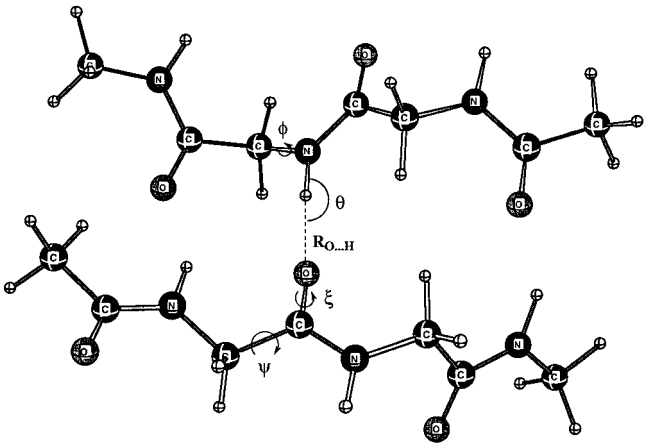

Figure 4. Structural model used to study the effects of altering the parameters $R_{O\cdots N}$, θ, ξ, ϕ, ψ (see also text for definition) on χ and η.

Factors Differentiating α*-Helices and* β*-Sheets.* To obtain a deeper insight into the difference between the two secondary structural elements, additional theoretical calculations have been carried out using the model compound shown in Figure 4. Single-point EFG calculations have been performed on the PM3-optimized geometry of this model. The dependencies on the hydrogen bond length (defined by $R_{O\cdots N}$), hydrogen bond angle (defined by $\theta = \angle H{-}N\cdots O$), and hydrogen bond dihedral angle (defined by $\xi = \angle N{-}C{=}O\cdots H$) of the calculated EFG have been investigated.

As evidenced by the graphical representations shown in Figures 5 and 6, the above-mentioned difference in χ of 0.54 MHz between the averaged values of the α and β fragments is found to be a direct result of the different conformations of the hydrogen-bonded systems. Figure 5a displays χ as a function of the bond length $R_{O\cdots N}$ and clearly shows that, in the relevant range of hydrogen bonds there is no dependence of χ on this bond distance. Figure 5c indicates that there is also minimal, if any, systematic dependence of χ on the dihedral angle $\angle N{-}C{=}O\cdots H$. Figure 5b, however, illustrates that χ varies dramatically with changes in the hydrogen bond angle $\theta = \angle H{-}N\cdots O$. These results explain the distinction between α and β. The difference in θ between these two systems is about 10°. H-bonds in α-helices are in general more linear ($\theta = \angle H{-}N\cdots O$ ca. 10−12°) than in β-sheets ($\theta = \angle H{-}N\cdots O$ ca. 20−22°).[44] According to our results in Figure 5b, such a 10° difference corresponds to an approximate variation of 0.34 MHz in the NQC constant. This accounts for ca. 63% of the difference we found between the α-helix and β-sheet conformations ($\Delta\chi(|\alpha{-}\beta|) = 0.54$ MHz).

It should be noted that χ and η do not necessarily follow the same trends. For instance, η shows a strong dependence on the bond length $R_{O\cdots N}$, Figure 6a, whereas χ remains almost constant, Figure 5a. As the O···H distance increases, the lone pair electrons of O become less polarized in the direction of the H atom, i.e., the electron distribution around the O atom becomes more spherical. Consequently, the *x*, *y*, and *z* components of the EFG (and in particular q_x and q_y) become more alike, yielding a smaller η at large $R_{O\cdots N}$ values. Unlike χ, which depends only on the q_z component, η is clearly a function of $R_{O\cdots N}$. Also notice that, at infinite distance, one would expect $q_x \approx q_y$ and therefore $\eta \approx 0$. As can be extrapolated from Figure 6(a), η tends asymptotically to zero with increasing $R_{O\cdots N}$. It should be noted, however, that around 2.8−3.0 Å, i.e., in the typical range of $R_{O\cdots N}$ for both α-helices and β-sheets, η is expected to be ca. 0.4.

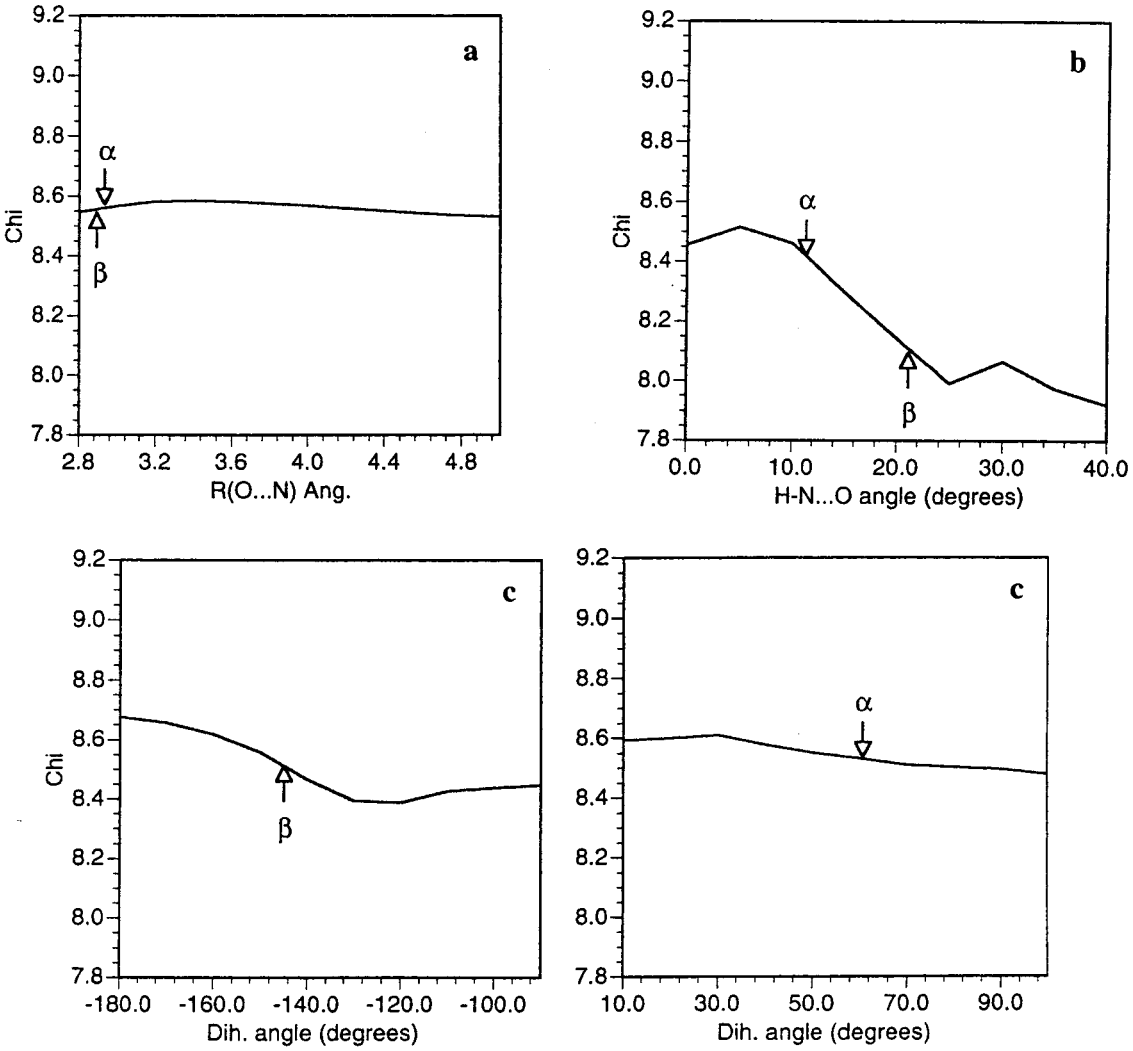

Figure 5. The calculated $\chi(^{17}O)$ values (in MHz) for the model system in Figure 4, as functions of (a) $R_{O\cdots N}$, (b) the hydrogen bond angle $\theta = \angle H{-}N\cdots O$ and (c) the hydrogen bond dihedral angle $\xi = \angle N{-}C{=}O\cdots H$.

Comparison with Previous Results. The results of this study are different from those reported by Kuroki et al.,[18] which indicated that χ is primarily a function of $R_{O\cdots N}$, with χ increasing with increasing length. These authors suggested that the distinction in χ between α and β was due to the slight difference in the $R_{O\cdots N}$ distance in these two structural elements.[18] X-ray diffraction has determined these lengths to be 2.87 and 2.83 Å for the α and β conformations, respectively.[45] Using the data of Kuroki et al.,[18] a difference of 0.04 Å would only cause $\chi_\alpha - \chi_\beta \approx 0.025$ MHz which cannot account for the large 0.54 MHz shift in χ that has been found both theoretically and experimentally. It follows from both our study and that of Kuroki et al.[18] that χ is strongly dependent on the ∠HNO bond angle, not $R_{O\cdots N}$. We found that the phenomenon of χ increasing with increasing $R_{O\cdots N}$ only occurs at larger distances, and has no effect over the small range of valid hydrogen bond lengths in real proteins.

Influence of the Backbone Conformation. To complete our study on the relationship between structure and NQC parameters, we have also investigated the effect of the backbone dihedral angles ϕ (phi) = ∠C(O)C−NC(O) and ψ (psi) = ∠NC−C(O)N on χ and η, as shown in Figures 7 and 8. It is well-known that the conformation of a given main chain is completely defined when ϕ and ψ are specified for each residue in the chain. Typical average values of ϕ and ψ for α-helices and antiparallel β-sheets are, respectively, $\phi_\alpha = -63.0°$ and $\psi_\alpha = -42.0°$, and $\phi_\beta = -109.0°$ and $\psi_\beta = 130.0°$. From Figure 7a, it can be seen that χ depends only minimally on ϕ: for $-150° \le \phi \le -50°$, 8.4 MHz $\le \chi \le$ 8.6 MHz. Consequently, the difference between α (ca. 8.40 MHz) and β (ca. 8.54 MHz) is only 0.14 MHz. In contrast, Figure 7b shows a clear trend that as ψ increases from 90° to 160°(β region), χ increases steadily, and as ψ increases from −70° to −30° (α region), χ increases steadily again. In particular, the difference between α (ca. 9.10 MHz) and β (ca. 8.66 MHz) is 0.44 MHz. This factor alone accounts for most (ca. 81%) of the difference we found between the α-helix and β-sheet conformations (0.54 MHz). This factor combined with the effect of a different H-bond angle, $\theta = \angle H{-}N\cdots O$, for α-helix and β-sheets (discussed above) explains the calculated difference of half a MHz. The two effects are probably cooperative.

Mulliken population analysis on these fragments reveal that the change in χ as a function of ψ (but not as a function of ϕ) can be mainly ascribed to a change in the electron distribution around the target O atom as ψ varies (the electron distribution remains almost constant with variations of ϕ). Geometry considerations play a minor role. The connection between electron density and EFG was already established in pioneering studies, where field gradients were regarded as holes of σ and/or π electrons.[46] A similar, more sophisticated concept has been applied in the present study. Thus, a change of 70° in ψ (from 100° to 170°) causes a variation of 16% on the total gross atomic

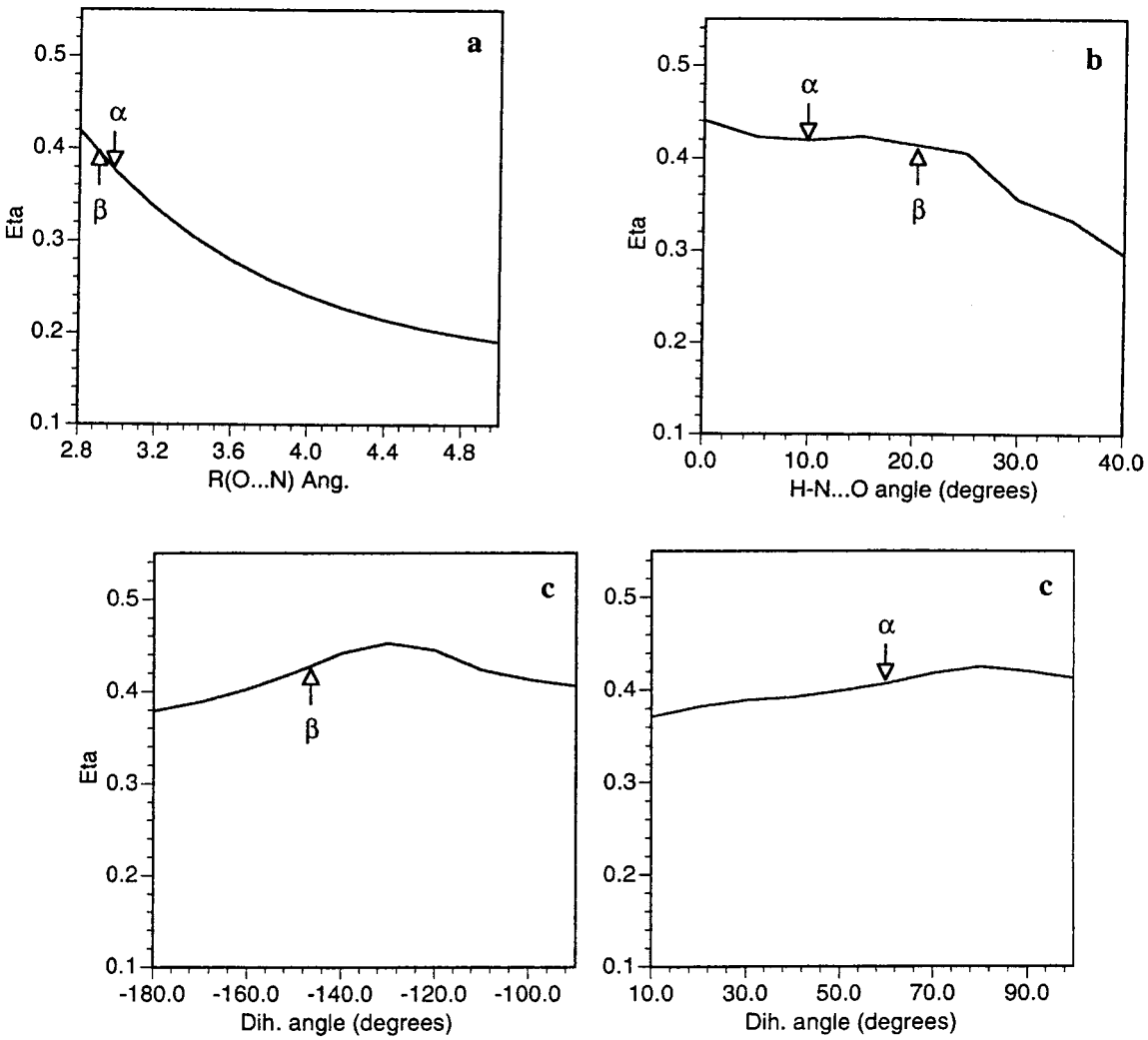

Figure 6. The calculated $\eta(^{17}O)$ values for the model system in Figure 4, as functions of (a) $R_{O\cdots N}$, (a) the hydrogen bond angle $\theta = \angle H-N\cdots O$ and (b) the hydrogen bond dihedral angle $\xi = \angle N-C{=}O\cdots H$.

population on the target O atom. A change of 70° in ϕ (from −150° to −80°) causes only a variation of 2% in the total gross atomic population on the target atom.

We conclude that for ^{17}O, NQC parameters of proteins depend on the particular conformation of the backbone and the hydrogen-bond system, specifically on the backbone dihedral angle ψ and the hydrogen bond angle θ. These are the two main factors making α-helices have observably higher NQC parameters than β-sheets. Conversely, ^{17}O NQC parameters are not dependent on $R_{O\cdots N}$, as had been previously thought,[18] nor are they dependent on either ξ or ϕ.

C. ^{14}N NQC Parameters. The calculated average χ and η for each protein fragment, both in the gas phase and in solution, are shown in Table 6. As can be seen, the χ values are remarkably consistent, and the largest deviation is only 6.1% (which corresponds to the PM3-optimized β fragment, β6). The overall averages in the gas phase are $\chi_\alpha = 4.46$ (9) MHz for α, and $\chi_\beta = 4.30$ (12) MHz for β, with an approximate difference, $\Delta\chi(|\alpha-\beta|)$, of 0.16 (15) MHz. This is within the uncertainties of zero, indicating that conformation has no effect on $\chi(^{14}N)$. In light of these results, we predict that ^{14}N NQDR spectroscopy will not be able to distinguish separate signals coming from two different secondary structures (α-helices and β-sheets).

Like in the case of ^{17}O, the difference between α-helices and β-sheets, $\Delta(|\alpha-\beta|)$, for ^{14}N does not change significantly from the gas phase, 0.16 (15) MHz, to solution, 0.14 (16) MHz, although the absolute values of χ decrease by ca. 0.15 MHz. The average change in the NQCCs between the gas phase and solvated fragments is $\Delta\chi = -9.1\%$.

Interpretation of the ^{17}O and ^{14}N Results. To understand why the ^{17}O parameters for α-helices and for β-sheets are distinguishable while the ^{14}N parameters are not, one should keep in mind the physical origin of the NQC constants. NQC constants result from the interaction between quadrupolar atoms and their immediate neighbors. Such an interaction creates a nonspherical charge distribution at the nucleus of the quadrupolar atom. In proteins, two related factors affect such interaction: the hydrogen-bonding network and the conformation of the protein. Let us analyze each atom, ^{17}O and ^{14}N, separately.

In protein backbones, the ^{17}O are all hydrogen bonded, so the electron distribution in oxygen atoms should be in principle dependent on the length, bond angle, and dihedral angle of this bond (only the bond angle seems to play a role according to our study). Second, both the hydrogen-bonding frame and the directionality of the bonds of the backbone atoms depend on the particular conformation of the protein. Thus, the hydrogen-bonding picture in an α-helix, composed of one twisted chain, is quite different from hydrogen bonding in a β-sheet conformation, with two approximately straight and parallel chains. Similarly, the directionality of the bonds involving O and N are quite different in a rodlike structure (α-helix, linear H-bond) and in a pleated sheet (β-sheet, bent H-bond). Therefore, one

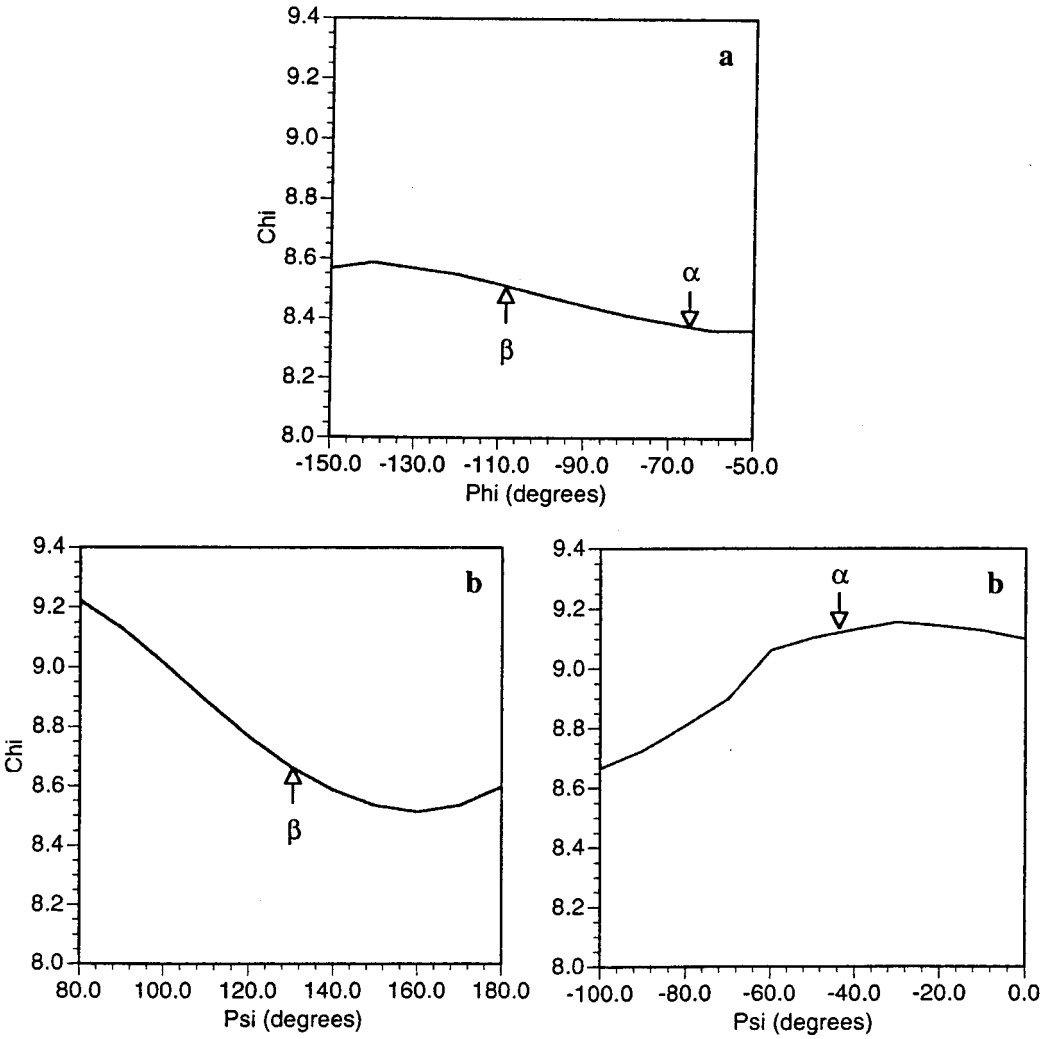

Figure 7. The calculated $\chi(^{17}O)$ values (in MHz) for the model system in Figure 4, as functions of the backbone dihedral angle (a) $\phi = \angle C(O)C-NC(O)$ and (b) $\psi = \angle NC-C(O)N$.

could conclude that, in principle, both the particular conformation of the backbone and the hydrogen-bonding picture play a key role in the resulting NQC parameters for ^{17}O.

On the other hand, the relationship between backbone ^{14}N NQC parameters and protein structure is less direct than that of the backbone ^{17}O. In the case of ^{17}O, it is clear that variations in the secondary protein structure have a direct effect on the hydrogen bond angle between the O and H. The nitrogen nuclei are also each bonded to a hydrogen, but here the bonds are covalent, and a change in the overall structure does not affect this bond.

We conclude that for ^{14}N, NQC parameters of proteins do not depend on the particular conformation of the backbone. Accordingly, ^{17}O NQDR spectroscopy will be suitable to evaluate the α/β contents of a given protein sample while ^{14}N NQDR spectroscopy will not.

D. NQDR Spectra and Frequency Predictions. The NQDR frequencies were calculated for all transitions of oxygen and nitrogen as described in section II.B, using eqs 1−5. Frequencies were originally evaluated using the exact η calculated. These results were not very conclusive. With the currently available methods, theoretical calculations of η are still very inaccurate and show unpredictably large deviations. As already pointed out by Torrent et al.[19] in a previous work, the computation of η utilizes the three diagonal components of the EFG. Each of the components represents a very small difference between two large numbers (electronic and nuclear contributions). Minor variations in either of these large numbers affect η drastically. This problem is particularly severe when the x and y components are very close; a minute fractional shift in one of them automatically turns into a large variation in η. The deviation in η here was also expected to be wide. To correct for this, the transition frequencies were recalculated with η restricted to its average values. This removed the extraneous outlying data, and limited the calculations to a more reasonable range of η. The results for both ^{14}N and ^{17}O frequencies are shown in Tables 7 and 8, respectively.

For ^{14}N, since it has an integral spin $I = 1$, eqs 1−3 were utilized. The corresponding frequencies (ν_o, ν_+, and ν_-), along with their sum, are tabulated in Table 7 to obtain the predicted NQDR spectra for ^{14}N.

This relatively simple method cannot be used for ^{17}O. Unlike ^{14}N, ^{17}O has a half-integer spin, and hence no exact solutions for the frequencies may be found in analytical form. The two methods described in the method section to obtain the eigenvalues of the secular equation were tested for comparison. The second method, numerical perturbation, was found to be more accurate and is the one we used below. The average calculated transition frequencies to obtain the predicted NQDR spectra for ^{17}O are summarized in Table 8.

From the results in Tables 7 and 8, the corresponding NQDR spectra can be partially predicted. What these tables do not show

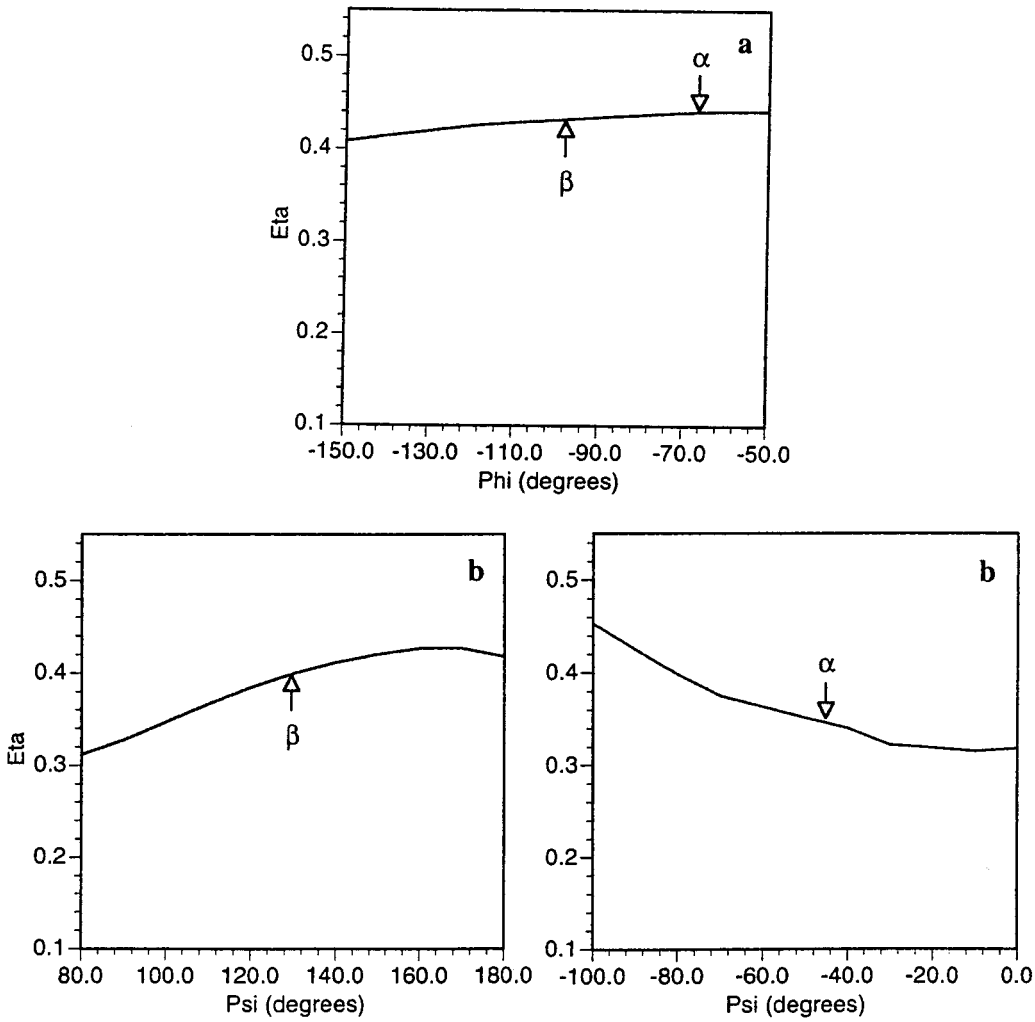

Figure 8. The calculated $\eta(^{17}O)$ values for the model system in Figure 4, as functions of the backbone dihedral angle (a) $\phi = \angle C(O)C-NC(O)$ and (b) $\psi = \angle NC-C(O)N$.

TABLE 6: Average Calculated ^{14}N NQC Constants χ (in MHz, Standard Deviation in Parentheses) and Asymmetry Parameters η of Gas-Phase and Solvated α-Helix and β-Sheet Protein Fragments Taken for Both Experimental and Optimized Geometries

	gas phase ($\epsilon = 1.0$)		solution ($\epsilon = 78.4$)		difference (gas phase − solution)	
fragment	χ	η	χ	η	$\Delta\chi$	$\Delta\eta$
α1	4.298	0.229	3.870	0.360	−0.428	+0.131
α2	4.533	0.189	4.215	0.245	−0.318	+0.056
α3	4.449	0.227	4.031	0.305	−0.418	+0.078
α4	4.583	0.199	4.142	0.296	−0.441	+0.097
α5	4.465	0.213	4.015	0.326	−0.450	+0.113
α6	4.415	0.236	4.060	0.315	−0.355	+0.079
av α	4.46(9)	0.21(2)	4.05(11)	0.31(3)	−0.41(14)	+0.10(4)
β1	4.287	0.324	3.897	0.486	−0.390	+0.162
β2	4.297	0.260	3.906	0.390	−0.391	+0.130
β3	4.423	0.255	4.020	0.383	−0.403	+0.128
β4	4.365	0.241	3.968	0.362	−0.397	+0.121
β5	4.306	0.286	3.914	0.429	−0.392	+0.143
β6	4.040	0.317	3.672	0.476	−0.368	+0.159
β7	4.405	0.266	4.004	0.399	−0.401	+0.133
av β	4.30(12)	0.28(3)	3.91(11)	0.42(4)	−0.39(16)	+0.14(5)
$\Delta(\lvert\alpha-\beta\rvert)$	0.16(15)	0.07(4)	0.14(16)	0.11(5)		

TABLE 7: Calculated ^{14}N Transition Frequencies (in MHz, Standard Deviation in Parentheses), ν_0, ν_-, and ν_+, of α-Helical and β-Sheet Proteins, Averaging Experimental and Optimized Geometries

	ν_0	ν_-	ν_+
α-helix	0.48(2)	3.10(13)	3.58(15)
β-sheet	0.60(3)	2.93(14)	3.53(18)
$\Delta(\lvert\alpha-\beta\rvert)$	0.12(4)	0.17(19)	0.05(23)

TABLE 8: Calculated ^{17}O Transition Frequencies (in MHz, Standard Deviation in Parentheses), ν_1, ν_2, and $\nu_1+\nu_2$, of α-Helical and β-Sheet Proteins, Averaging Experimental and Optimized Geometries

	ν_1	ν_2	$\nu_1+\nu_2$
α-helix	1.67(3)	2.69(6)	4.36(9)
β-sheet	1.51(5)	2.56(8)	4.07(13)
$\Delta(\lvert\alpha-\beta\rvert)$	0.16(6)	0.13(10)	0.29(16)

is how isolated a certain peak will be or how crowded a given region of the spectra will become due to overlap of numerous adjacent peaks. According to Table 8, in the ^{17}O spectra, there should be a distinct peak at 1.668 MHz for α-helices and at 1.513 MHz for β-sheets. These peaks would represent, respectively, the predicted transition frequency ν_1 for α-helices and for β-sheets. Provided they do not overlap much (here we have information only for the intensity), they should be experimentally distinguishable. The transition ν_2 appears at higher frequencies. Here again, the α and β peaks should be resolvable unless they overlap significantly. The clearest indication of whether a sample structure is α-helical or β-sheet, however, comes from $\nu_1+\nu_2$, the frequency for transitions from the first energy level to the third. Since $\nu_1+\nu_2$ is basically an addition of the first two frequencies, the peak separation between α and

β is also additive. Hence, $\nu_1+\nu_2$ should have a larger peak separation than either ν_1 or ν_2, regardless of the overlap between α and β. The gap between α and β shown in Table 8, 0.29(16) MHz, indicates that the experimental signals for α-helices and β-sheets in ^{17}O spectra should be distinct.

It follows from Table 7 that ^{14}N signals from α and β conformations might be harder to identify. Of the three frequencies, two (ν_- and ν_+) are almost completely indistinguishable because, for both of them, α and β values are within much less than a standard deviation of each other. The only distinguishable transition frequency for ^{14}N is ν_o. Here, like ν_1 (^{17}O) and $\nu_1+\nu_2$ (^{17}O), the peaks should be more separated. From these strictly theoretical considerations, we predict that this experimental signal for α-helices and β-sheets in ^{14}N spectra are likely to be distinct. However, this may not be the case after taking into account experimental factors. Due to its low frequency, this signal for ν_o is often difficult to detect. The current procedure of NQDR may make low-frequency detection unviable. On the other hand, Rabbani et al.[47] have been able to detect such low frequency signals in other structures using NQDR. Therefore, we should conclude that, only given the right experimental setup, the $\nu_o(^{14}N)$ signal might be distinguishable for α-helices and β-sheets.

To sum up, it follows from our predictions that it will be experimentally feasible to identify α-helices and β-sheets in the NDQR spectra of protein backbone ^{17}O, by correctly interpreting the spectra according to the above guidelines. In the case of ^{14}N NQDR spectra, a more sophisticated experimental setup is required.

IV. Conclusions

According to the above presented results and discussions, we may draw the following conclusions.

(1) Specific NQDR spectra are computationally predicted for the ^{14}N and ^{17}O inherent in protein backbones. The separate signals resulting from α-helices and β-sheet models are predicted to be experimentally distinguishable only for ^{17}O NQDR spectroscopy: $\Delta(|\alpha-\beta|)$ for $^{17}O = 0.54(17)$ MHz in gas phase and 0.53(15) MHz in solution, whereas $\Delta(|\alpha-\beta|)$ for $^{14}N = 0.16(15)$ MHz in the gas phase and 0.14(16) MHz in solution, which are within the uncertainties of zero. Hence, protein conformation may be identified through analysis of the ^{17}O NQDR spectra, but not ^{14}N NQDR spectra.

(2) The theoretically calculated ^{17}O NQC parameters for α-helices and β-sheets in gas phase are $\chi_\alpha = 9.30(9)$ and $\chi_\beta = 8.76(14)$, and $\eta_\alpha = 0.14(5)$ and $\eta_\beta = 0.24(7)$, with the standard deviation in parentheses. These parameters correlate well with available NMR data for α-helices and β-sheets. The average change in the NQCCs between the gas-phase and solvated protein fragments is $\Delta\chi = -4.1\%$.

(3) The theoretically calculated ^{14}N NQC parameters (with the standard deviation in parentheses) for α-helices and β-sheets in the gas phase are $\chi_\alpha = 4.46(9)$ and $\chi_\beta = 4.30(12)$, and $\eta_\alpha = 0.21(2)$ and $\eta_\beta = 0.28(3)$. Neither $\chi(^{14}N)$ nor $\eta(^{14}N)$ has been experimentally determined specifically for α-helices and β-sheets yet. However, the computed values here reported are similar to experimental results for ^{14}N nuclei in related nitrogen-containing molecules.

(4) Different amino acid residues, R, on a given protein do not significantly affect the NQC parameters of the backbone C═O and NH groups. Certainly the specific residues play a determining role in establishing the backbone conformation. However, for the purpose of calculating the EFG, these residues can be safely replaced by hydrogen atoms (R = H, i.e polyglycines) without a significant loss of accuracy, and with the advantage of minimizing the computational cost.

(5) ^{17}O NQC parameters of proteins are dependent on the particular conformation of the backbone, specifically on θ and ψ. Due to this, α-helices and β-sheets have observably different ^{17}O NQC parameters.

(6) Conversely, ^{17}O NQC parameters are not dependent on $R_{O\cdots N}$, as had been previously thought, nor are they dependent on either ξ or ϕ.

(7) Unlike ^{17}O NQC parameters, ^{14}N NQC parameters of proteins do not depend on the particular conformation of the backbone.

Overall, these results strongly support the idea that DFT methods give excellent representations of local electrostatic properties such as EFG for proteins with a different secondary structure. It encourages the future use of these methods in situations where experimental results are lacking or difficult to obtain. Thus, computationally predicted spectra like the ones reported here can be confidently used (1) to help elucidate the components of a complex NQDR spectra of a protein sample, (2) to indicate which signals can safely be disregarded when examining other features of proteins, such as enzyme active sites, and (3) to investigate proteins with unknown secondary structures.

Experimental work is currently in progress in our lab using the above predicted spectra as a guide for identifying the α-helix and β-sheet signals of oxy-myoglobin and human transcription factor IIB.

Acknowledgment. The present research is in part supported by grants CHE96-27775 from the National Science Foundation (K.M.) and R21 RR13365 from the National Institute of Health (E.P.D.). M.T. acknowledges a Postdoctoral Fellowship from the Spanish Ministerio de Educación, Cultura y Deportes. Acknowledgment is made to the Cherry L. Emerson Center of Emory University for the use of its resources, which is in part supported by a National Science Foundation grant (CHE-0079627) and an IBM Shared University Research Award. Computer time allocated at the Center for Supercomputer Applications (NCSA) is also acknowledged.

References and Notes

(1) (a) Branden, C.; Tooze, J. *Introduction to Protein Structure*; Garland Publishing: New York, 1991. (b) Stryer, L. *Biochemistry*, 4th ed.; Freeman: New York, 1995; Chapters 2 and 3.

(2) Nowick, J. S. *Acc. Chem. Res*. **1999**, *32*, 287.

(3) Kellis, J. T. *Nature* **1988**, *333*, 784.

(4) (a) Saito, H.; Ando, I. *Annu. Rep. NMR Spectrosc*. **1989**, *21*, 209. (b) Ando, I.; Yamanobe, T.; Asakura, T. *Prog. NMR Spectrosc*. **1990**, *22*, 349. (c) Asakawa, N.; Kurosu, H.; Ando, I.; Shoji, A.; Ozaki, T. *J. Mol. Struct*. **1994**, *317*, 119.

(5) Shoji, A.; Ozaki, T.; Fujito, T.; Deguchi, K.; Ando, I.; Magoshi, J. *J. Mol. Struct*. **1998**, *441*, 251.

(6) Ashikawa, M.; Shoji, A.; Ozaki, T.; Ando, I. *Macromolecules* **1999**, *32*, 2288.

(7) Poplett, I. J. F. *Adv. Nucl. Quadrupole. Reson*. **1980**, *4*, 115.

(8) (a) Verhoeven, J.; Dymanus, A.; Bluyssen, H. *J. Chem. Phys*. **1969**, *50*, 3330. (b) Brosnan, S. G. P.; Edmonds, D. T. *J. Mol. Struct*. **1980**, *58*, 23. (c) Poplett, I. J. F. *J. Magn. Reson*. **1982**, *50*, 397.

(9) (a) Butler, L. G.; Brown, T. L. *J. Am. Chem. Soc*. **1981**, *103*, 6541. (b) Cummins, P. L.; Bacskay, G. B.; Hush, N. S. *J. Chem. Phys*. **1985**, *82*, 2002. (c) Keshari, V.; Karna, S. P.; Chandra, P. *J. Mol. Struct*. **1989**, *192*, 271. (d) Eggenberger, R.; Gerber, S.; Huber, H.; Searles, D.; Welker, M. *J. Mol. Spectrosc*. **1992**, *151*, 474. (e) Webster, B. C.; Buttar, D. *J. Chem. Soc., Faraday Trans*. **1992**, *88*, 1087. (f) Bagno, A.; Scorrano, G. *J. Phys. Chem*. **1996**, *100*, 1545.

(10) Takahashi, A.; Kuroki, S.; Ando, I.; Ozaki, T.; Shoji, A. *J. Mol. Struct*. **1998**, *442*, 195.

(11) Yamauchi, K.; Kuroki, S.; Ando, I.; Ozaki, T.; Shoji, A. *Chem. Phys. Lett*. **1999**, *302*, 331.

(12) LiWang, A. C.; Bax, A. *J. Magn. Res*. **1997**, *127*, 54.

(13) (a) Cundari, T. R.; Deng, J.; Fu, W.; Klinckman, T. R.; Yoshikawa, A. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 941.

(14) (a) Oldfield, E. *J. Biomol. NMR* **1995**, *5*, 217. (b) de Dios, A. C.; Oldfield, E. *Solid State NMR* **1996**, *6*, 101.

(15) (a) Perczel, A.; Farkas, Ö.; Csizmadia, I. G. *J. Am. Chem. Soc*. **1996**, *118*, 7809. (b) Gresh, N.; Guo, H.; Salahub, D. R.; Roques, B. P.; Kafafi, S. A. *J. Am. Chem. Soc*. **1999**, *121*, 7885. (c) Bursulaya, B. D.; Brooks III, C. L. *J. Am. Chem. Soc*. **1999**, *121*, 9947. (d) Wu, Y.-D.; Wang, D.-P. *J. Am. Chem. Soc*. **1999**, *121*, 9352.

(16) Havlin, R. H.; Le, H.; Laws, D. D.; deDios, A. C.; Oldfield, E. *J. Am. Chem. Soc*. **1997**, *119*, 11951.

(17) (a) Spera, S.; Bax, A. *J. Am. Chem. Soc*. **1991**, *113*, 5490. (b) Kricheldorf, H. R.; Müller, D. *Macromolecules* **1983**, *16*, 615. (c) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647.

(18) Kuroki, S.; Ando, S.; Ando, I. *Chem. Phys*. **1995**, *195*, 107.

(19) Torrent, M.; Musaev, D. G.; Morokuma, K.; Ke, S.-C.; Warncke, K. *J. Phys. Chem. B* **1999**, *103*, 8618.

(20) Ke, S.-C.; Torrent, M.; Musaev, D. G.; Morokuma, K.; Warncke, K. *Biochemistry* **1999**, *38*, 12681.

(21) (a) Kern, C. W.; Karplus, M. *J. Chem. Phys*. **1965**, *42*, 1062. (b) Prasad, G.; Keshari, V.; Chandra, P. *J. Mol. Struct*. **1989**, *192*, 253.

(22) Lucken, E. A. C. *Nuclear Quadrupole Coupling Constants*; Academic: New York, 1969.

(23) (a) Prasad, G.; Lal, A.; Chandra, P. *Theor. Chim. Acta* **1989**, *75*, 475. (b) Palmer, M. H. *Chem. Phys*. **1988**, *127*, 335. (c) Pati, R.; Srinivas, S.; Briere, T.; Das, T. P. *J. Phys. Chem*. **1995**, *99*, 9051.

(24) (a) Cummins, P. L.; Bacskay, G. B.; Hush, N. S.; Ahlrichs, R. *J. Chem. Phys*. **1987**, *86*, 6908. (b) Brown, R. D.; Head-Gordon, M. P. *Mol. Phys*. **1987**, *61*, 1183.

(25) (a) Harvey, J. S. M. *Proc. R. Soc*. **1965**, *285A*, 581. (b) Schaefer, H. F., III.; Klemm, R. A.; Harris, F. E. *Phys. Rev*. **1969**, *181*, 137. (c) Fuller, G. H. *J. Phys. Chem*. **1976**, *5*, 835. (d) Mason, J., Ed. *Multinuclear NMR*; Plenum Press: London, 1987. (e) Raghavan, P. *At. Nucl. Data Tables* **1989**, *42*, 203. (f) Boykin, D. W. *^{17}O NMR Spectroscopy in Organic Chemistry*; Chemical Rubber Co.: Boca Raton, FL, 1990. (g) Pyykkö, P. *Z. Naturforsch*. **1992**, *47a*, 189.

(26) Pyykkö, P.; Li, J. Report HUKI 1-92; Department of Chemistry, University of Helsinki: Helsinki, 1992.

(27) Tokman, M.; Sundholm, D.; Pyykkö, P.; Olsen, J. *Chem. Phys. Lett*. **1997**, *265*, 60.

(28) Bersohn, R. *J. Chem. Phys*. **1952**, *20*, 1505.

(29) Dean, C. *Phys. Rev*. **1954**, *96*, 1053.

(30) Cohen, M. H. *Phys. Rev*. **1954**, *96*, 1278.

(31) Chihara, H.; Nakamura, N. *Numerical Data and Functional Relationships in Science and Technology*; Hellwege, K.-H., Hellwege, A. M., Eds.; Springer-Verlag: Berlin, 1988; Volume 20, Subvol. a.

(32) This metalloprotein was selected because it is currently being studied in our laboratory at the Department of Physics; future comparisons once experiments are completed will be useful.

(33) It should be noted that pftfiibn has only 50 residues and a low content of beta (16%); only one fragment containing six peptide bonds in an antiparallel pleated sheet could be extracted from such a protein. Endo-1,4-β-xylanase has an exceptionally high content of beta (62%); from this protein four eight-residue β-sheet fragments were obtained.

(34) (a) Richardson, J. *Adv. Prot. Chem*. **1981**, *34*, 174. (b) Kohn, W. D.; Hodges, R. S. *Trends Biotechnol.* **1998**, *16*, 379.

(35) Okamoto, Y.; Masuya, M.; Nabeshima, M.; Nakazawa, T. *Chem. Phys. Lett*. **1999**, *299*, 17.

(36) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys*. **1981**, *55*, 117. (b) Miertus, S.; Tomasi, J. *Chem. Phys*. **1982**, *65*, 239. (c) Cossi, M.; Barone, R.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett*. **1996**, *255*, 327.

(37) Torrent, M. *ONIOM2EFG*; Emory University: Atlanta, 2000.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A. Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.1; Gaussian, Inc.: Pittsburgh, PA, 1998.

(39) (a) *ADF2000*; Vrije Universiteit Amsterdam: Amsterdam, The Netherlands, 2000. (b) Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem. Phys*. **1973**, *2*, 41. (c) te Velde, G.; Baerends, E. J. *J. Comput. Phys*. **1992**, *99*, 84. (d) Fonseca Guerra, C.; et al. METECC-95 **1995**, 305.

(40) Ziegler, T.; Rauk, A. *Theor. Chim. Acta* **1977**, *46*, 1.

(41) O'Konski, C. T.; Ha, T.-K. *J. Chem. Phys*. **1968**, *49*, 5354.

(42) Åstrand, P.-O.; Ruud, K.; Mikkelsen, K. V.; Helgaker, T. *Mol. Phys*. **1997**, *92*, 89.

(43) (a) Cummins, P. L.; Bacskay, G. B.; Hush, N. S. *Mol. Phys*. **1987**, *61*, 795. (b) Butler, L. G.; Brown, T. L. *J. Am. Chem. Soc*. **1981**, *103*, 6541. (c) Gready, J. E. *Chem. Phys*. **1981**, *55*, 1. (d) Bagno, A.; Lovato, G.; Scorrano, G.; Wijnen, J. W. *J. Phys. Chem*. **1993**, *97*, 4601.

(44) Values obtained by averaging these parameters in all α-helices and all β-sheet fragments under study.

(45) (a) Arnott, S.; Wonacott, A. L. *J. Mol. Biol*. **1966**, *21*, 371. (b) Arnott, S.; Dover: S. D.; Elliot, A. *J. Mol. Biol*. **1967**, *30*, 201.

(46) (a) Bersohn, R. J. *J. Chem. Phys*. **1954**, *22*, 2078. (b) Morokuma, K.; Fukui, K.; Yonezawa, T.; Kato, H. *Bull. Chem. Japan* **1963**, *36*, 47.

(47) (a) Rabbani, S. R.; Edmonds, D. T.; Gosling, P. *J. Magn. Res*. **1987**, *72*, 230. (b) Rabbani, S. R.; Edmonds, D. T.; Gosling, P. *J. Magn. Res*. **1987**, *72*, 422.