Utilizing the Charge Field Effect on Amide ^{15}N Chemical Shifts for Protein Structure Validation

Reto Bader*

Department of Physics, Stockholm University, Arrhenius Laboratories, 106 91 Stockholm, Sweden

Received: August 18, 2008; Revised Manuscript Received: October 28, 2008

Of all the nuclei in proteins, the nuclear magnetic resonance (NMR) chemical shifts of nitrogen are the theoretically least well understood. In this study, quantum chemical methods are used in combination with polarizable-continuum models in order to show that consideration of the effective electric field, including charge screening due to solvation, improves considerably the consistencies of statistical relationships between experimental and computed amide ^{15}N shifts between various sets of charged and uncharged oligopeptides and small organic molecules. A single conversion scheme between shielding parameters from first principles using density functional theory (DFT) and experimental shifts is derived that holds for all classes of compounds examined here. This relationship is then used to test the accuracy of such ^{15}N chemical shift predictions in the cyclic decapeptide antibiotic gramicidin S (GS). GS has previously been studied in great detail, both by NMR and X-ray crystallography. It adopts a well-defined backbone conformation, and hence, only a few discrete side chain conformational states need to be considered. Moreover, a charge-relay effect of the two cationic ornithine side chains to the protein backbone has been described earlier by NMR spectroscopy. Here, DFT-derived backbone amide nitrogen chemical shifts were calculated for multiple conformations of GS. Overall, the structural dynamics of GS is revisited in view of chemical shift behavior along with energetic considerations. Together, the study demonstrates proof of concept that ^{15}N chemical shift information is particularly useful in the analysis and validation of protein conformational states in a charged environment.

Introduction

Improving the accuracy of NMR chemical shift calculations has posed a major challenge to theoretical chemists for more than 5 decades,[1] which is somewhat in contrast to experiment, where chemical shifts are relatively easily and highly accurately measured quantities. Hence, there is a large and ever-growing body of experimental data awaiting theoretical understanding, and indeed, magnetic shielding parameters have repeatedly proved useful as benchmarks in tests of new *ab initio* molecular orbital theories.[2–4] After the recognition that isotropic chemical shifts are strongly correlated to protein structure[5]—in particular $^{1}H^{\alpha}$,[6] $^{13}C^{\alpha}$, and $^{13}C^{\beta}$[7]—and the development of powerful computational approaches using hybrid density functionals,[8] the calculation of chemical shifts has recently seen a revival in structural biology.[9] Especially in the case of pharmaceutically relevant structural problems, including peptide and protein amyloid fibrils or membrane-bound peptides and proteins, X-ray crystallography and nuclear-Overhauser-effect (NOE)-based solution NMR methods very often fail due to insolubility, high molecular weight, or noncrystallinity. Systems not amenable by conventional NMR also include structures of short-lived, unstable proteins, or systems where a minor state is in dynamic equilibrium with a more populated state, in which case collection of structural restraints for the minor component can be prohibitively difficult.[10] Acquisition of chemical shift information is nevertheless often still possible in such cases, and data analysis may then rely to a significant extent on the correct interpretation of chemical shifts and shielding tensors.[11]

Overall, chemical shift rather than NOE-based structure determination procedures hold the promise to hasten significantly the rate of high-throughput protein structure generation by NMR, since the assignment of NOE data commonly remains the limiting and most time-consuming step in the whole process.[12] Moreover, rapid interpretation of chemical shift perturbations in a protein NMR spectrum, e.g., due to the binding of a ligand, is highly desirable in structure-based drug design.[13] Such potential applications readily explain the increasing number of NMR structure determination protocols aimed at circumventing the need for collection of NOE data.[10,14–16] More accurate and extensive correlation between chemical shift and local protein conformation is now needed, either through additional empirical data or from recent improvements in the quantum mechanical computation of chemical shifts, in order to push the useful protein size limit of such protocols.[12]

Of all the different nuclei present in proteins, very good agreement between theoretical shift predictions on the basis of small molecular fragments that are excised from published protein structures and the corresponding experimental values has been achieved for ^{13}C.[17] The 4−5 ppm secondary chemical shift separation seen experimentally between the helical and sheet residues for both C^{α} and C^{β}[6,18] was reproduced quantitatively in calculations.[9] In another study, the shielding parameters for C^{α}, C^{β}, $C^{\gamma 1}$, $C^{\gamma 2}$, and C^{δ} of Ile in BPTI and cytochrome *c*, which cover a chemical shift range of about 60 ppm, were predicted with root-mean-square (RMS) deviations of 1.1 and 1.8 ppm, respectively, using the minimal *N*-formyl-L-Ile acid as a model for the computations.[19] These and many similar results therefore suggest that ^{13}C chemical shifts chiefly depend on strictly local conformational parameters, while hydrogen bonding, sequence neighbor residues, solvent, and electrostatic

* E-mail: reto.bader@biologie.uni-r.de. Phone: +49-941-943 4921. Fax: +49-941-943 2479. Present address: Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany.

10.1021/jp807362v CCC: $40.75
Published on Web 12/16/2008

TABLE 1: Correlation Parameters for a Reference Set of Molecules (See Also *Supporting Information Table S2*) between Computed Absolute Nitrogen Shielding Parameters and Experimental Values As Referenced against $\delta(NH_3^{liquid})$

scheme	level of NMR shielding calculation	level of geometry optimization	SCRF-PCM	intercept	slope	r^2	RMSE
1	B3LYP/6-31G**	B3LYP/6-31G**	no	251.47 (±2.32)	−0.9901 (±0.0136)	0.9994	5.17
2	B3LYP/6-31G**	B3LYP/6-31G**	yes	253.62 (±1.74)	−0.9896 (±0.0102)	0.9997	3.88
3	B3LYP/6-31G**	B3LYP/cc-pVQZ	yes	258.56 (±2.97)	−0.9941 (±0.0175)	0.9991	6.63
4	B3LYP/6-311G(2df, 2pd)	B3LYP/6-31G**	yes	224.81 (±1.46)	−0.9261 (±0.0078)	0.9998	3.19
5	B3LYP/6-311G(2df, 2pd)	B3LYP/cc-pVQZ	yes	229.66 (±1.29)	−0.9322 (±0.0070)	0.9998	2.82

effects can safely be ignored. In turn, an approach has recently been presented in which traditional NOE restraints were combined with $^{13}C^{\alpha}$-derived constraints for backbone and side chains to result in refined structural models with better agreement between predicted and observed $^{13}C^{\alpha}$ chemical shifts.[20]

In a comparison of methods for calculating NMR shifts and shielding anisotropies, in which a large range of small molecules with vastly different environments for the nuclei of interest were used, best results were generally obtained using the 3-parameter hybrid functional B3LYP in conjunction with a Pople-style triple split-valence basis set. Again, ^{13}C chemical shift predictions were very accurate with a root-mean-square error (RMSE) of 4.2 ppm over a range of 220 ppm. By contrast, the RMSE of 17.2 ppm for the calculated ^{15}N chemical shifts over a chemical shift range of 160 ppm was more than four times larger.[2] Similar quantitative differences were also reported in comparisons of ^{15}N and ^{13}C chemical shift predictions in proteins.[21] On the basis of various theoretical calculations, it was suggested that, apart from the dependence on torsion angles also seen with ^{13}C chemical shifts, ^{15}N shifts may additionally be highly sensitive to hydrogen bonding[22] and possibly also longer range electrostatic fields. It may be for such potentially complex patterns of contributions to shielding that ^{15}N shifts have been less useful than ^{13}C shifts for protein structure prediction until now.[23]

In previous work concerned with ^{15}N chemical shift predictions for all the different types of nitrogen atoms,[24] the most accurate results were obtained for the nitrogen sp^3 hybrids (amines and hydrazines), thus resembling the good agreement observed for sp^3 carbons mentioned above. By constrast, the higher polarity of sp^2 hybridized amide nitrogens seems to make them more highly sensitive to intermolecular interactions, most notably hydrogen bonding and electrical effects. In experimental studies, amide ^{15}N chemical shifts of the protein backbone have been noted to depend on the hydrogen bonding capacity of the solvent, solvent reaction field effects, the presence of ions, protein self-association or binding of ligands. In fact, the latter has successfully been utilized in a high-throughput approach to discover high-affinity ligands to proteins in target-directed drug research,[25] based on chemical shift perturbations in 2D $^1H-^{15}N$ HSQC NMR titration experiments of amino acids in close proximity to the bound ligand.[26] Despite the repeated recognition of the possibly great potential of ^{15}N chemical shifts in complementing purely conformational information, only little efforts have been taken, however, to incorporate ^{15}N chemical shift information in structural models of peptides and proteins. One previous *ab initio* study of amide ^{15}N chemical shifts in N-Formyl-Ala-Ala dipeptides showed that the backbone dihedral angles have the most dominant effect on ^{15}N chemical shifts, spanning a range of up to 20 ppm for the angles closest to the peptide group and up to 8 ppm for the adjacent (preceding) dihedral angles. By comparison, different side chain conformations in *N*-formyl-Ala-Val produced chemical shift variations of only 4 ppm. Moreover, hydrogen bonding and electrostatic effects were estimated to cause 2−3 ppm fluctuations.[27]

Here, we demonstrate that a significant contribution to peptide amide ^{15}N chemical shifts is arising from charge fields and that the effect can be described reliably from first principles, if the dielectric properties of the surroundings are taken into account. To this particular end, the computationally effective continuum model of the solvent (see ref 28 for a review) is shown to suffice. The method allows one to investigate the consistency between various experimental ^{15}N random-coil chemical shift tables of all 20 proteogenic amino acids that were previously derived using a variety of charged or uncharged single amino acids or di-, tri-, tetra-, and pentapeptide models of random-coil peptides under different solvent conditions.[29–32] Moreover, theoretically predicted charge field effects on ^{15}N chemical shifts for computationally optimized structural models of gramicidin S in different conformational states[33] are compared with an early experimental NMR investigation of this protein.[34] Mimicking the cationic ammonium end groups of the ornithine side chains using point charges or explicit ammonium and each individual peptide bond as *N*-methylacetamide allows one to dissect the gross charge-relay effect on the ^{15}N chemical shift into individual contributions from the close proximity of a partial charge and hydrogen bonding between the peptide backbone and the ammonium group. This provides an example as to how ^{15}N chemical shifts can probe the presence of charges as, e.g., observed in the binding of charged or highly polar ligands to receptors or the formation of intramolecular contacts on folding as examined in the present study.

Computational Methods

QM. Small organic reference molecules, amino acids, and short peptide model structures were optimized using density functional theory with B3LYP/6-31G**,[8] unless stated otherwise. An initial MD structure of gramicidin S was further optimized using the double-ζ quality basis set lacvp as implemented in the Jaguar 5.5 package[35] or the double-ζ quality basis set LanLDZ as in Gaussian 03.[36] Default convergence criteria in Gaussian 03 or Jaguar 5.5 were met in most cases.

Shielding parameters of any given compound were usually calculated at the same level of theory used for geometry optimization. Gauge-including atomic orbitals (GIAO) were used to define the vector potential that describes the magnetic field in the Hamiltonian.[2]

For the calculation of charge field effects, the surrounding protein matrix was described as a dielectric continuum through the revised polarizable-continuum-model self-consistent reaction field method (PCM-SCRF) as implemented in Gaussian 03,[37] both during geometry optimizations and shielding calculations. The united atom model was used to build the cavities in combination with the standard electrostatic scaling factor of 1.2 for the sphere radius. The integral-equation formalism version of PCM was applied to solve the Poisson−Boltzmann equation, which defines the classical treatment of electrostatic interactions in solution. Default solvent definitions were used for water,

methanol, and DMSO. It should be noted that the calculation of NMR parameters may be used in conjunction with SCRF methods only as in its most recent implementation in Gaussian 03.

MD. An initial structure of gramicidin S for further refinement by QM methods was generated using GROMACS 3.3.[38] Distance restraints were introduced in the topology file to bring the C and N termini of an initially linear gramicidin S decapeptide into sufficiently close proximity so as to enable subsequent manual cyclization to produce a restraint-fulfilling structure and further geometry optimization by QM methods. The GROMOS96 43b1 vacuum force field was employed to describe the protein. Integration time steps were 1 fs for equilibration and 4 fs during the simulation. All simulations were carried out under constant temperature (300 K) and pressure (1 atm). Equilibration was for 200 ps, before structures were generated over a 10 ns simulation.

Results

Correlating ^{15}N Shielding Parameters with Experimental ^{15}N Shifts. Various previous studies acknowledged the fact that correlated levels of theory are necessary to obtain accurate absolute nitrogen NMR shielding predictions for at least amide nitrogens[27,39] and possibly also other molecular systems in the presence of lone pairs or multiple bonds.[24] Using DFT, however, no particular exchange-correlation functional has been singled out for performing significantly superior over others. While computational NMR methods provide predictions of absolute chemical shielding tensors (and hence also isotropic shielding parameters) which are linearly correlated to the conventional experimental δ shift scales as defined with respect to a given reference compound, the slope of such linear correlation need not be 1 and seem to depend at least as strongly on the choice of basis set as on the density functional method[40] (see also Supporting Information Tables S1 and S2). Hence, linear correlations for conversion of the computed isotropic shielding parameters into chemical shifts need to be established for any given combination of method and basis set. Table 1 presents the parameters of such correlations between shielding predictions by the widely used B3LYP hybrid functional[8] along with the respective experimental values[31] for five small organic reference compounds that cover a wide spectral range of nitrogen shielding between 31 and 510 ppm.

The root-mean-square error of predictions made using the polarized Pople-style split-valence double-ζ basis 6-31G** of 3.88 ppm is larger by only 1 ppm than the value associated with the triple-ζ basis 6-311G(2df,2pd), confirming once again that, overall, B3LYP/6-31G(d,p) geometry optimizations and NMR shielding calculations usually provide reasonable accuracy at acceptable computational costs. In addition, Figure 1a shows the individual residuals between the predicted and experimental chemical shifts for each method listed in Table 1. The compounds used to establish the linear regression parameters are represented by bars in dark gray, while the deviations between experiment and theory for two independent test compounds (gaseous ammonia and nitrogen) are represented by bars in light gray. Using larger basis sets both for geometry optimization and NMR shielding calculations outperforms computations with smaller basis sets, in that the predictions for these two additional compounds are within the error of the linear regression from the sample compounds. The respective coefficients of determination r^2 for the linear relationships between shifts and computed shielding parameters are 0.9994 and 0.9998 for 6-31G** and 6-311G(2df,2pd), respectively.

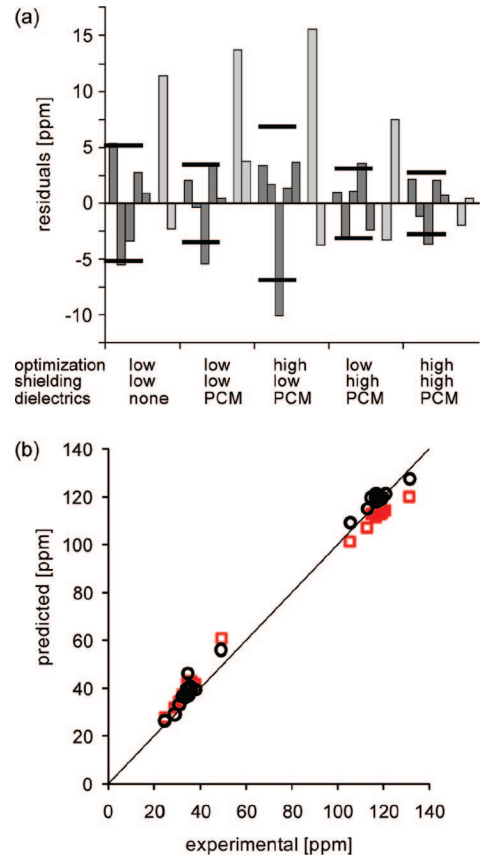

Figure 1. (a) Deviations of computed from experimental nitrogen shifts at different levels of theory. Geometry optimizations were performed using B3LYP/6-31G** or B3LYP/cc-pVQZ for lower or higher quality basis sets, respectively. Shielding parameters were calculated using B3LYP/6-31G** or B3LYP/6-311G(2df,2pd) for lower or higher quality, respectively. PCM was used to include reaction field effects as indicated. Bars in dark gray represent residuals of predictions from the test set (glycine, acetanilide, cyanide, nitrate, and *trans*-azobenzene), while bars in light gray represent residuals of independent predictions for gaseous ammonia and nitrogen. Root-mean-square errors are indicated by horizontal lines. (b) ^{15}N chemical shift predictions for various cationic and acetylated amino acid methylesters[30] (Supporting Information Table S3). Absolute shielding parameters were computed using B3LYP/6-31G(d,p), and the values were converted into the $\delta(NH_3^{liquid})$ chemical shift scale using schemes 1 and 2 of Table 1, dependent on whether shift predictions were made in the gas phase (red squares) or using PCM to include the reaction field (black circles).

Protein nitrogen resonances usually fall into a much narrower spectral window of only about 120 ppm. Hence, the smaller basis set 6-31G(d,p), which is normally used to compute shielding parameters in peptides or fragments of proteins, was further assessed in this spectral range by calculating the NMR shielding parameters of 15 cationic amino acid methylesters as well as 13 acetylated amino acid methylesters with different side chains both in the gas phase and using a PCM to include reaction field effects (see also Supporting Information Table S3). Figure 1b shows a plot of the predicted values for the PCM (in red) and gas phase (in black) versus experimental shifts from ref 30. The upfield resonances associated with the amine forms of the amino acid methylesters are predicted too downfield on average by 4.3 $\pm$ 2.7 and 5.8 $\pm$ 2.3 ppm using the PCM and in the gas phase, respectively. By contrast, the amide ^{15}N shift predictions are insignificantly too downfield by 1.8 $\pm$ 2.2 ppm using PCM to account for the dielectric nature of the environment while too upfield by 4.7 $\pm$ 2.4 ppm in the gas phase.

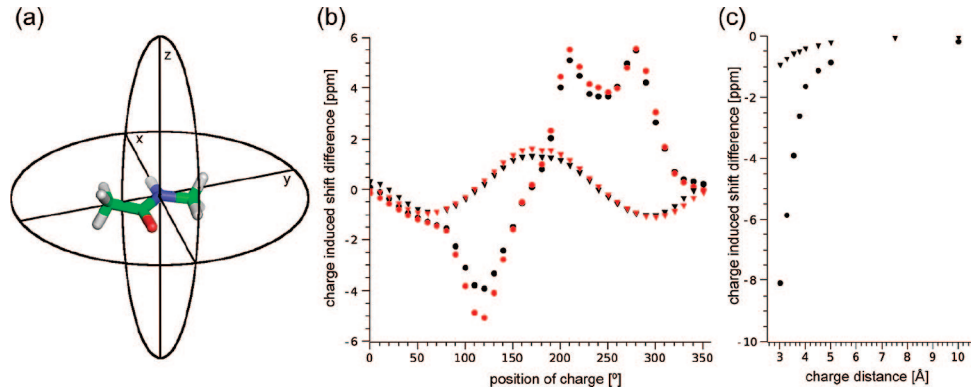

Figure 2. Predicted effect of a partial positive charge of +0.25 proton units on the amide ^{15}N chemical shift of *N*-methyl-acetamide in a cavity surrounded by a dielectric environment of $\varepsilon = 4.0$. (a) The charge was positioned on circles around the ^{15}N nucleus with a radius of 3.5 Å. (b) Circles denote positions of the charge in the *xy* plane, where *x* is at an angle of 0° and *y* at 90°, while triangles indicate positions of the charge in the *xz* plane, where *x* is again at 0°. Shift predictions were made using basis set 6-31G(d,p) (black symbols) in combination with scheme 2 of Table 1 or 6-311G(2df,2pd) (red symbols) in combination with scheme 5 of Table 1 to convert the theoretical shielding values into the experimental shift scale. (c) Charge-induced effect on ^{15}N chemical shift as predicted for varying distances of a single partial charge of +0.25 proton units along two different directions from the nitrogen nucleus: triangles denote charge positions in the *xz* plane at 60° from *x*, and circles represent charge positions in the *xy* plane at 120° from *x*.

Overall, the PCM including computations (in black) result in a closer agreement between theory and experiment, in particular at the amide ^{15}N nuclei.

Charge-Induced Effects. The effect of charges on the amide ^{15}N chemical shift tensors has been the topic of previous investigations that were mainly concerned with solid-state effects of ionic compounds (see, e.g., refs 1 and 41). A systematic approach to utilize charge field effects on the more widely abundant isotropic chemical shifts in solution for protein structure validation and refinement by introducing additional geometric constraints has only very recently been proposed by Vila and Scheraga for the $^{13}C^{\alpha}$ chemical shift.[42] It has been noted earlier, however, that, in computing protein chemical shifts in the gas phase, the use of neutral rather than charged side chains resulted in better agreement with experimental values obtained in aqueous solutions.[43] This can be rationalized by the concept of induced surface charge,[44] where a charge of opposite polarity to the electrostatic potential is induced at the molecular surface through the solvent, an effect which is usually called charge screening. In effect, ionizable residues on the surface of a protein are usually only partially charged and, hence, shielding/deshielding induced by changes in the protonation/deprotonation of ionizable groups in water is expected to be significantly reduced as compared to the gas phase. Here, we will first examine the sensitivity of the amide ^{15}N chemical shift to the presence of charges and then move on to compute amide ^{15}N chemical shifts of charged and uncharged peptides both in the gas phase and in water using the polarizable-continuum model to take dielectric effects of the solvent into account.

Figure 2 shows how the computed amide ^{15}N resonance of a model for the peptide bond moiety shifts if a partial positive charge is placed at a distance of 3.5 Å and at different positions with respect to the orientation of the peptide bond. Here, the NMA−peptide bond was in a cavity surrounded by a dielectric environment of $\varepsilon = 4.0$ which has been proposed to resemble the average dielectric properties of the protein interior. For example, at a distance of 3.5 Å along the direction of the N−C$^{\alpha}$ bond of a peptide group, an arbitrary apparent charge of +0.25 proton units will induce a shift of about −5 ppm (Figure 2b). The polarizabilites of the bonds around the nitrogen nucleus seem at maximum if the charge is located in the plane spanned by the atoms defining the peptide bond. The effect is most negative in the direction of the N−C$^{\alpha}$ bond and most positive at about the direction of the N−C′ bond (reaching a maximum at angles of ±50° from the latter bond). By contrast, chemical shift changes tend to be much smaller as such charges are moved outside the peptide bond plane, adopting maximal values of ±1.8 ppm at a distance of 3.5 Å and perpendicular in either direction of the plane. This result agrees very well with the experimental observations in short, extended peptides that C-terminal protonation and deprotonation reactions result in large chemical shift changes at the preceding amide ^{15}N nucleus, while backbone nuclei are not sensitive to proton transfer effects at the side chains of aspartate, glutamate, lysine, or arginine in such random-coil peptides.[29,31]

Computations of the charge-induced effect on shielding using basis sets of different sizes result in basis-set-dependent theoretical values. Upon conversion of these theoretical values from the shielding scale into the experimental shift scale (using the correlation parameters derived in Table 1), however, the predicted effects are virtually identical (compare the red and black data points in Figure 2b).

Figure 2c shows how the charge-induced effect on the amide ^{15}N chemical shift decreases as the distance between the charge and ^{15}N nucleus is increased on a straight line through the center of the sphere. The data were computed for two different directions of such lines (within and outside the peptide bond plane), on which the partial charge of +0.25 proton units was moved. The computed values were then parametrized using an exponential function. For any given direction of the lines in an environment of $\varepsilon = 4$, the charge-induced effect on the amide ^{15}N shift doubles as the charge is moved closer to the nucleus by 0.64 Å. As expected, the effect is independent of the particular direction of the lines through the center of the sphere. Overall, the above theoretical calculations show that charge-induced chemical shift effects on the ^{15}N nucleus possess a number of features that may prove useful for restraints for protein structure validation and refinement: the magnitude and sign of the effect are shown to be strongly dependent on the relative position of the charge with respect to the orientation of the peptide bond. Indeed, the effect is also inversion symmetrical with respect to the sign of the added charge.

Charged vs Uncharged Oligopeptides. The charge-induced chemical shift effect on ^{15}N resonances has been observed experimentally in oligopeptides by either change of pH or chemical modification (acetylation/amidation) of ionizable

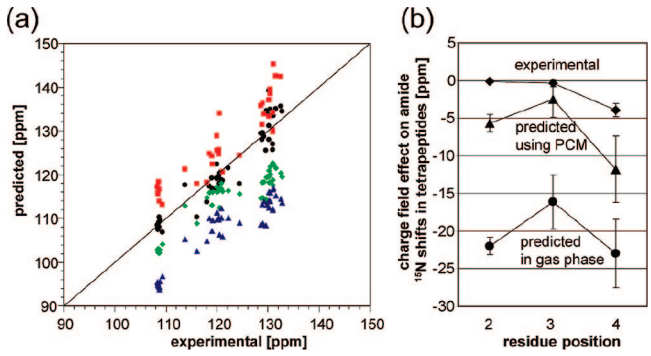

Figure 3. ^{15}N chemical shift predictions for zwitterionic tetrapeptides,[29] modeled as extended structures (Supporting Information Table S5). (a) Absolute shieldings were computed using B3LYP/6-31G(d,p), and the values were converted into the $\delta(NH_3^{liquid})$ chemical shift scale using the relationship established in Table 1. Shift predictions were made in the gas phase and using PCM to include the reaction field, as denoted by red squares and black circles, respectively. Shift predictions were additionally made using uncharged model compounds in the gas phase (blue triangles) or using PCM (green diamonds). (b) Average of charge-induced shift differences for residue positions 2−4 in zwitterionic tetrapeptides as compared to their uncharged states calculated for the gas phase (circles), using PCM (triangles), and extrapolated from experimental values measured in water at different pHs (diamonds).[29]

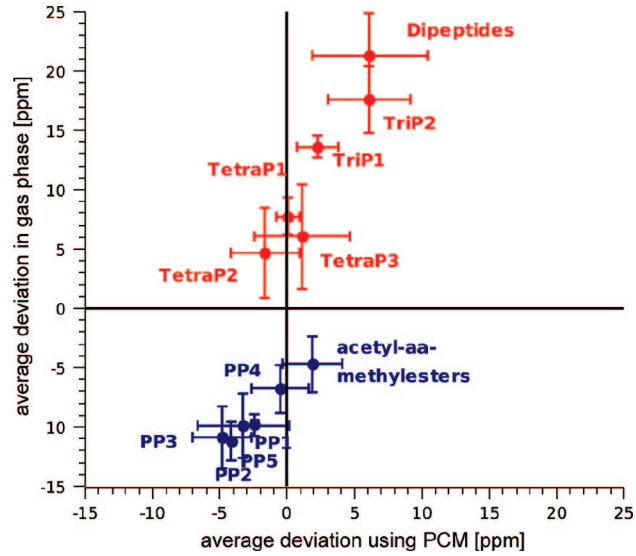

Figure 4. Average deviations of amide ^{15}N chemical shift predictions in charged (red) and uncharged (blue) peptides using PCM to include the reaction field (*x*-axis) as compared to predictions made in the gas phase (*y*-axis). Absolute shieldings were computed using B3LYP/6-31G(d,p), and the values were converted into the $\delta(NH_3^{liquid})$ chemical shift scale using the relationship established in Table 1. Values for uncharged, N-terminally acetylated and C-terminally amidated pentapeptides measured in water[32,45] are shown in blue and are labeled PP1−5, respectively, according to the residue position along the peptide sequence; acetylated amino acid methylesters measured in DMSO[30] are denoted in blue and labeled acetyl-aa-methylesters; zwitterionic di-,[31] tri-,[31] and tetrapeptides[29] measured in water are indicated in red and labeled Dipeptides, TriP1, TriP2, and TetraP1−3, respectively, according to the amide bond position along the peptide sequence. See Supporting Information Tables S3−S6 for individual data points.

groups in attempts to define reference ^{15}N chemical shifts for all 20 proteinogenic amino acids[29,31,32,45] (see also Supporting Information Tables S4−S6). In Figure 3a, experimental shifts for tetrapeptides of the type G-G-X-A as measured by Wüthrich et al. are compared to those computed for different charge states and dielectric environments. Under physiological conditions in aqueous solution, the peptides are zwitterionic with a positively charged ammonium group at the N terminus and a negatively charged carboxyl group at the C terminus, but suitable changes of pH can render either end neutral.

All calculations were performed both in the gas phase and employing the PCM for water, and all peptide models were geometry optimized in extended conformation with the side chain adopting the *trans* configuration. Recent calculations of potential energy surfaces of dipeptides indicate that the most stable backbone geometry is indeed extended,[46] and hence, it can be assumed that this conformation provides a reasonable first approximation to model the dielectric effects on the ^{15}N chemical shifts. In doing so, the shifts were computed for both the zwitterionic and the uncharged forms of the peptides. The results reveal clearly that, on average, most unbiased agreement between theoretical and experimental values for the zwitterionic state is obtained if both the peptide model includes terminal charges and reaction field effects are taken into account by the PCM method, resulting in a RMSE of the predicted from the experimental values of 3.0 ppm (black circles in Figure 3a).

It has been noted previously that deprotonation of the C-terminal carboxyl group gives rise to a low-field shift of 4−6 ppm at the amide ^{15}N resonance of the C-terminal residue,[29,31] while deprotonation of the N-terminal ammonium group has virtually no effect on any downstream amide ^{15}N shift. Figure 3b compares the average residue position-dependent experimental shift differences on deprotonation of the peptide end groups (diamonds) with the shift differences between charged and uncharged peptide models as predicted both in the gas phase (circles) and using the PCM approach (triangles). It is evident that the effect of the terminal charges on amide ^{15}N chemical shifts in tetrapeptides is much better reproduced using the PCM approach than through computations in the gas phase. Typically, the PCM for water included spheres on the terminal atoms which account for polarization charges of about −0.5 to −0.6 proton units at the N terminus and +0.5 to +0.6 proton units at the C terminus. Such charge-screening results in a dampening of the chemical shift effects observed in the gas phase of about 17, 13, and 10 ppm at the amide ^{15}N nitrogen of residues 2, 3, and 4, respectively. Nevertheless, the charge-induced shift effects are overestimated also with the PCM approach, by about 6, 2, and 8 ppm on average at residue positions 2, 3, and 4, respectively (Figure 3b). Interestingly, a partial charge of −0.25 protons at the relative position from an amide ^{15}N nucleus in the direction and distance that corresponds to the position of a C-terminal CO_2^- group of zwitterionic tetrapeptides in water is predicted to induce a chemical shift effect of about +4 ppm (Figure 2b with reverse signs for the charge-induced shift difference), which is in very good agreement with the experimentally observed shift difference. Moreover, the strong distance dependence of the charge field effect (Figure 2b) is also in agreement with the fact that protonation and deprotonation at either terminal end of zwitterionic tetrapeptides do not perturb the experimental amide nitrogen shifts at neither residue positions 2 nor 3. Overall, these considerations therefore suggest that significant charge-screening effects can only be observed with an upper limit of about 5 Å for the distance between a given nitrogen nucleus and any surrounding charges.

In order to examine this problem in more detail, the ^{15}N chemical shifts were calculated for a variety of both charged and uncharged peptides for which experimental data were available in the literature. Figure 4 summarizes the results comparing the statistical deviation between the experimental values and theoretical prediction made either in the gas phase

(*y*-axis) or using the PCM (*x*-axis) for individual residue positions in different peptides. Overall, the inclusion of the PCM in the calculations reduces the span of average deviations between theory and experiment from about 32 ppm in the gas phase to 11 ppm. Gas phase predictions of uncharged peptides underestimate the experimental values by 4−10 ppm, while they overestimate the measured shifts of zwitterionic di-, tri-, and tetrapeptides by 5−22 ppm. Using the PCM, there is no such obvious segregation, although biases toward slightly positive or negative deviations for charged or uncharged peptides, respectively, are observed.

Moreover, the magnitude of the deviation seems to be dependent on the residue position along the peptide chain. For instance, the predictions using the PCM for uncharged acetylated pentapeptide amides Ac−GGXGG−NH_2 are correct for glycine following immediately after the variable central amino acid, but they are significantly too upfield for the other four residue positions. A similar observation is made for zwitterionic di- and tripeptides, in which the amide ^{15}N shift of the C-terminal amino acid is predicted at too downfield a resonance, while the predictions for the central residue in the tripeptide are only slightly overestimated on average. It is likely that there are multiple contributions to these systematic biases including errors from different shift referencing procedures employed in different experimental studies and specific conformational and hydration preferences at different positions within the peptides or from approximations made in building the cavity to produce the PCM reaction field.

Underestimation of solvent effects on ^{15}N NMR shielding parameters has been reported recently also in the case of three diazines in aqueous solution.[47] In this study, the PCM reaction field was found to be very sensitive to the dimension of the cavity. To bypass this dependence, the authors modeled explicitly the water molecules that were closest to the charge and thereby reached better agreement with the experimental solvent-induced shifts. A similar approach, with a few explicit water molecules placed in hydrogen bonding distance around the C-terminal CO_2^- group of the tetrapeptides NH_3^+−GGAA−CO_2^- and NH3−GGAA−COOH, however, did not significantly reduce the predicted shift difference at any amide nitrogen nucleus of the peptides examined here.

Despite these limitations, the PCM approach produces a reasonably quantitative prediction of the charge- and reaction-field-induced effects on the amide ^{15}N chemical shift. The magnitude of these effects is shown to be comparable to those associated with conformational changes and hence needs to be taken into account in the prediction of protein ^{15}N chemical shifts. In turn, the effect will also prove useful for the validation of local features of protein models in the surroundings of nearby charges. Typical situations may be encountered in the binding of charged ligands and ions or protonation and deprotonation at the protein surface. A practical example shall be discussed in the following.

^{15}N Chemical Shifts of Gramicidin S. The cyclic decapeptide gramicidin S [cyclo-(Val-Orn-Leu-D-Phe-Pro)$_2$] (GS) is an antibiotic which acts on virtually all types of cellular membranes by making them ion-permeable.[48] The conformational properties of GS are commonly expected to play a role in such activity, and much effort has been directed toward a better understanding of the structure and dynamics of wild-type and mutant GS in different environments and their relationships with activity.[49] Several NMR studies indicate that GS possesses a relatively rigid backbone, which was found to be the same within experimental error in both the aprotic solvent DMSO and the protic solvent methanol.[50–52] The studies are also in agreement with early crystallographic work on a GS−urea complex, which revealed an antiparallel β-pleated sheet structure with two type II′ β-turns.[53,54] The conformational rigidity arises from four trans-annular symmetric hydrogen bonds between Leu and Val residues. An additional hydrogen bond was observed in the crystal structure between a N_δ of one of the two Orn and the CO group of the following D-Phe. By contrast, the NMR studies were less clear about the side chain conformations and there was evidence for solvent-dependent behavior, in particular for the ornithine side chains. Moreover, the NMR data are differing from the X-ray results in that the time-averaged NMR parameters in solution are only consistent with a C_2 symmetric structure, which had been suggested in early studies of GS[55] based on the primary sequence of the peptide. In principle, a C_2 symmetry consistent NMR data set could be realized either in individual molecules or, alternatively, as a time-averaged quantity in a dynamical system.[56]

Several attempts to resolve the tertiary structure of Orn have been hampered by the apparent lack of long-range NOEs between the aminopropyl side chain and the protein backbone. The absence of such NOEs is somewhat surprising considering that a distinct preference of more than 70% for one side chain $C^\alpha-C^\beta$ rotamer of the ornithine residues was predicted on the basis of $^3J_{\alpha\beta}$ coupling constants.[33] Whether the latter was corresponding to $\chi^1 = 180°$ or $\chi^1 = -60°$, however, could not be clarified. A preferred conformation was also postulated for the distal portion of Orn, for which ^{13}C T_1 relaxation times in both methanol and DMSO indicated a decrease in the apparent rotational correlation time between C^β and C^γ.[57,58] Moreover, early semiempirical calculations[59,60] and a subsequent crystalline urea complex of GS[53] suggested one or two hydrogen bonds, involving Orn $N^\delta H_3^+$ and the backbone carboxyl groups of D-Phe. While the former predicted two such hydrogen bonds in the i → i − 3 sense,[59] the latter revealed a single i → i + 2 Orn $N^\delta H_3^+$ −O═C bond per molecule.[53] In a subsequent study, evidence by NMR was reported which suggested that the two Orn side chains fluctuate quickly (on the NMR time scale) between a fully unbound state and an i → i + 2 Orn $N^\delta H_3^+$ −O═C D-Phe intramolecular hydrogen bond.[34]

With regard to chemical shift analyses, at least three studies made extensive use of the ^{15}N chemical shift information in order to evaluate secondary and tertiary structural features of gramicidin S in various solvents. Khaled et al.[61] and Live et al.[62] used this information to examine solvent-induced perturbation of the four trans-annular hydrogen bonds. Krauss and Chan[34] studied the conformation and intramolecular hydrogen bonding of the ornithine side chains through a comparison of ^{15}N chemical shifts of GS and a chemically modified GS variant with permethylated Orn side chains to prevent them from forming intramolecular hydrogen bonds.

Taken together, the wealth of available ^{15}N shift studies and the comparably few degrees of dynamical freedom of GS in methanol and DMSO suggest that this decapeptide antibiotic is an attractive target to test the potential of ^{15}N chemical shift information in the validation of protein structures. Here, a model for the overall secondary structure of GS was obtained through an MD simulation, in which an initially extended and linearized conformation of the decapeptide was cyclized by applying a minimal distance constraint between atoms of the N and C terminus. The simulation converged quickly toward the correct hydrogen bonding scheme between the two strands and remained stable throughout the 10 ns simulation, which is consistent with earlier MM simulations[56] and several experimental NMR

TABLE 2: Computed Geometries (ψ/φ and χ^1) and Relative Total Energies Using B3LYP/lacvp with PCM for Methanol of Three Conformers of GS with Three Different Combinations of Orn Side Chain Hydrogen Bonding States Each (The Geometries Are Compared to Experimental NMR and X-ray Data[54])

	conformer 1			conformer 2			conformer 3			X-ray	NMR
H-bonded Orn sc.	2	1	0	2	1	0	2	1	0	1	n.d.
rel. energy (kcal/mol)	21	3	8	20	10	4	31	12	0		
Val 1	−118/157	−128/162	−128/139	−100/130	−101/135	−103/142	−102/133	−104/139	−103/139	−120/157	−120/120
Val 6	−119/160	−118/152	−126/132	−102/134	−102/141	−96/134	−105/135	−92/130	−92/129	−125/153	−120/120
side chains	gauche+/gauche−			gauche−/trans			gauche−/trans			g+/g−	
Orn 2	−111/131	−105/131	−123/90	−120/130	−117/130	−123/124	−115/126	−115/126	−114/118	−108/131	−110/110
Orn 7	−113/133	−106/125	−127/87	−123/133	−119/131	−122/116	−116/128	−106/123	−108/118	−105/136	−110/110
Leu 3	−124/105	−123/102	−149/123	−127/107	−127/107	−138/116	−125/103	−126/98	−133/101	−121/93	−120/110
Leu 8	−125/106	−141/115	−152/130	−129/108	−132/110	−138/116	−127/105	−141/106	−145/111	−139/121	−120/110
side chains	gauche−			gauche−			trans			g−	
Phe 4	60/−120	59/−118	56/−117	62/−121	62/−122	60/−123	59/−122	61/−122	61/−124	61/−126	55/−110
Phe 9	58/−120	69/−119	59/−116	62/−121	60/−122	60/−123	59/−123	60/−124	60/−125	56/−136	55/−110
side chains	trans			trans			gauche+			trans	
Pro 5	−88/1	−87/2	−83/−2	−91/3	−91/3	−96/−1	−87/3	−86/−2	−87/−3	−81/−2	−60/−40
Pro 10	−87/−1	−90/1	−84/−4	−90/2	−90/4	−89/−1	−87/2	−88/3	−88/3	−93/11	−60/−40

analyses of GS in solution.[33,34,50,63] An arbitrary snapshot at the end of a 10 ns trajectory was then used as a starting conformation for a high-level geometry optimization using B3LYP and the lacvp basis set of the Jaguar 5.5 program package. An initial optimization of GS in the gas phase resulted in a structure in which the two ornithine side chains formed two hydrogen bonds i → i + 2/i + 3 Orn $N^{\delta}H_3^+$—O═C D-Phe/O═C Pro each. After the optimization had converged, the side chain dihedral angles were manually adjusted to produce three different conformers of GS which were inspired by the findings of an early $C^{\alpha}-C^{\beta}$ conformational analysis by NMR.[33] Particular attention was paid to the conformation of the Orn side chain. The latter is either in *trans* orientation if its N^{δ} forms an intramolecular hydrogen bond to the CO of Phe or gauche (−60°), if unbound.[54] Of each conformer, triplicates were designed in which both, one, or none of the two Orn side chains formed an intramolecular hydrogen bond. All conformers were then subjected to another cycle of geometry optimization using B3LYP/lacvp and the PCM for methanol.

Table 2 compares the total energies and backbone geometries of the energy minimized conformers 1−3 along with the geometries derived from X-ray and NMR analyses.[54] Overall, the data suggest that the three conformers, which differ in the side chain χ^1 torsion angles of Val1, Leu3, and D-Phe4, are associated with similar electronic energies for identical states with regard to Orn $N^{\delta}H_3^+$—O═C D-Phe4 hydrogen bonds. By contrast, in all conformers 1−3, the energies are significantly larger by 10−20 kcal/mol for structures in which both Orn side chains are hydrogen bonded to the peptide backbone as compared to structures with only one or no intramolecular hydrogen bond. The electronic energy does not account for entropic effects. Intramolecular hydrogen bond formation will, however, reduce the number of degrees of freedom and hence be energetically even less favorable. Taken together, these calculations strongly suggest that C_2 symmetrical structures with two intramolecular Orn $N^{\delta}H_3^+$—O═C D-Phe4 hydrogen bonds are hardly populated. The formation of just one such hydrogen bond is energetically well accessible and even favorable in the case of conformer 1. Indeed, both backbone and side chain conformations of conformer 1 with a single Orn side chain being hydrogen bonded to the peptide backbone are in excellent agreement with the geometry of the GS crystal structure.[53,54] As the latter could only be obtained in the presence of urea, it has never been ruled out that the symmetry-broken structure was energetically more favorable only as a complex with urea in the crystal structure. The present results, however, suggest that the crystal structure presents a true low-energy conformation in solution. The C_2 symmetry that was apparent in the time-averaged NMR parameters would then arise by rapid fluctuations between a fully unbound state and a single i → i + 2 Orn $N^{\delta}H_3^+$—O═C D-Phe intramolecular hydrogen bond, as suggested by Krauss et al.[34]

It is now particularly interesting to see how ^{15}N shift predictions made using conformers 1−3 from Table 2 relate to experimentally measured NMR ^{15}N shifts of GS in DMSO[61,62] (Supporting Information Table S8). Comparing the three models with non-hydrogen-bonded Orn side chains from Table 2, the experimental ^{15}N shifts are best reproduced theoretically using conformer 1 (Figure 5), which is known to exist in the crystalline phase. The predictions seem to be systematically biased by about +3 ppm with an overall root-mean-square error of 4 ppm. Such a value is within the range of potential experimental reference offsets.[64] By contrast, the ^{15}N chemical shifts are significantly less well reproduced using conformers 2 and 3 with RMSEs of 8.8 and 8.7 ppm, respectively. Extremely large deviations of about 15 ppm are observed at Val in both conformers 2 and 3. This provides strong evidence that the χ^1 rotamer of Val1 and 6 is well-defined in solution and adopts a gauche+/gauche− configuration (as in conformer 1 and the X-ray structure) rather than gauche+/trans as in conformers 2 and 3. Conformer 3 differs from the X-ray structure additionally in the side chain rotamer configurations of Leu3 and D-Phe4. Indeed, using this model, the residuals of the shift predictions at Leu 3 are

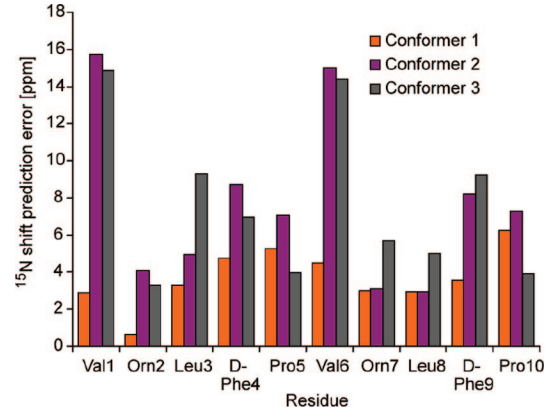

Figure 5. Comparison of amide ^{15}N chemical shift predictions for three different conformers of gramicidin S. The bars indicate the error of prediction as the difference between the predicted shifts (after conversion from computed shieldings using the relationship established in Table 1) and experimental shifts measured in DMSO.[62] The individual geometries of conformers 1−3 are as specified in Table 2.

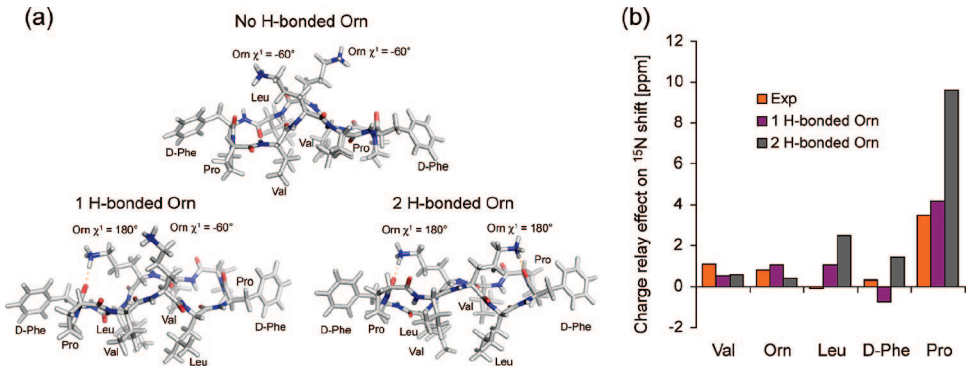

Figure 6. Charge-relay effect on the ^{15}N chemical shifts of gramicidin S. (a) Side-views of geometry optimized models of GS conformer 1 with no, one, or two ornithine side chains intramolecularly hydrogen bonded to the carboxyl group of D-Phe. Orn $\chi^1 = 180°$ if the side chain Orn $N^{\delta}H_3$ is hydrogen bonded and $\chi^1 = -60°$ otherwise[33] (see Supporting Information Table S7 for coordinates). (b) Average charge-relay effect on GS amide ^{15}N chemical shifts for the different types of residues in the three conformational states shown on the left-hand side. The average is taken to account for the fact that the observed time-averaged NMR parameters are only consistent with a C_2 symmetrical structure. The experimental values are from Krauss et al.[34] and are calculated as the amide ^{15}N shift differences between GS in methanol, in which the ornithine side chains are intramolecularly hydrogen bonded to the carboxyl groups of D-Phe, and [2,2′-N^{δ}-trimethylornithyl]GS in methanol, in which such intramolecular hydrogen bonds cannot form.

significantly larger than using conformers 1 or 2. By contrast, the ^{15}N shift prediction error at D-Phe 4 is larger both in conformers 2 and 3 than conformer 1, although the side chain configurations of conformers 1 and 2 are identical. This suggests that the ^{15}N shift of D-Phe is more strongly determined by the backbone (ψ/φ) rather than side chain (χ^1) conformation.

While the focus of this work is on the computation of amide ^{15}N chemical shifts, it is clear that the promise of using chemical shifts in protein conformational analysis is that there may be corroborating pieces of information from different types of nuclei. For example, the various carbon chemical shifts are known to be strongly dependent on local backbone and side chain conformation.[65] The isotropic carbon shieldings for GS in the three Orn rotamer states of conformers 1−3 are listed in Supporting Information Table S9 along with experimental data measured in DMSO.[66] The linear correlations between the theoretical and experimental values are equally good in all cases ($r^2 > 0.96$). Moreover, for Leu, the ^{13}C shift predictions (C^{α}, C^{β}, C^{γ}, and C^{δ}) are similar, despite the fact that they differ in side chain torsion angles. For Val, however, the C^{α}, C^{β}, and C^{γ} resonance predictions are in better agreement with the experimental values, if $\chi^1 = 60°$ (as in conformers 2 and 3) rather than 180° (as in conformer 1). This is in contrast to both the crystal structure[53,54] and the respective ^{15}N chemical shifts, which are significantly better reproduced by DFT shielding calculations using conformer 1 (Figure 5). Since the Val rotamer state of conformer 1 ($\chi^1 = 180°$) is energetically more favorable, it seems likely that this state is also most populated in solution, although it may nevertheless exist in conformational equilibrium with other Val side chain conformations. To conclude, the ^{13}C chemical shift predictions for the side chain nuclei prove less useful in the present case, presumably because the carbon chemical shifts are more strongly dependent on backbone rather than side chain torsion angles and the former are largely kept fixed in all three conformers. Moreover, the empirical dependence of the C^{γ} resonances on amino acid side chain torsion angles has only recently been investigated and seems to cover a rather narrow range of shifts on average (1−2 ppm),[67] which makes specific predictions less reliable.

Taken together, the above energy and NMR shift calculations on various conformational models of GS provide a considerable number of constraints in describing the structure and dynamics of this peptide antibiotic in solution. Indeed, the view emerges that GS in solution predominantly exists in a conformational state that is highly similar to the crystal structure. Whether this also holds for the rotamer and hydrogen bonding states of the two Orn side chains shall now be discussed again using ^{15}N chemical shift information.

The side chain hydrogen bonding of Orn has been a prominent focus of debate.[34,53,56,60] As mentioned above, the main questions are whether it exists at all and, if so, if both or just one Orn side chain is involved and which carboxyl groups provide the hydrogen-accepting oxygen. To shed light on some of those questions, Krauss et al.[34] designed a GS variant with permethylated Orn side chains. Quaternization of Orn eliminates the possibility of intramoleculear hydrogen bonding while keeping the backbone conformation of GS intact.[34] The ^{15}N chemical shift differences observed at backbone amide nitrogens with and without permethylation of the Orn side chains are then merely resulting from the combined effects of intramolecular hydrogen bonds and charge field effects due to varying distances to the charged ornithine side chain N^{δ}. The experimental shift differences as measured in methanol are shown in Figure 6b. The release of the intramolecular hydrogen bond between Orn $N^{\delta}H_3^+$ and the carboxyl group of D-Phe by permethylation of the ornithine side chain results in a specific charge-relay effect of 3.5 ppm at the nitrogen nucleus of the amide bond between D-Phe_i and Pro_{i+1}. Minor shifts of 1.1 and 0.8 ppm are observed at amide nitrogens of Val_{i+2} and Orn_{i+3}.

Along with the experimental estimate for the charge-relay effect on backbone amide ^{15}N chemical shifts as a result of intramolecular hydrogen bonding to the Orn side chain, the bar graph of Figure 6 also shows the calculated average effect on the backbone ^{15}N shift of the residue types on formation of intramolecular hydrogen bonds to a single or both Orn side chain $N^{\delta}H_3^+$. In good agreement with the experimental observations, the ^{15}N shift effects of Orn $N^{\delta}H_3$ hydrogen bond formation to the carboxyl group of D-Phe are limited to this particular amide linkage with a strong downfield secondary shift on Pro ^{15}N. Moreover, the magnitude of the average effect on the two Pro is predicted to be 4.2 ppm, if just one of the Pro residues is involved in an intramolecular hydrogen bond to an Orn side chain at any time, and 9.6 ppm, if both at the same time. The former value compares clearly much better with the experimental estimate of 3.5 ppm, and hence, the shift calculations strongly support the view that GS in methanol solution exists in a

TABLE 3: Charge Field and Hydrogen Bonding Effects on ^{15}N Chemical Shifts of Gramicidin S Studied Using Minimal *N*-Methyl-acetamide Models for the Individual Peptide Bonds in Combination with NH_4^+ or NH_3 at the Position of the Charged Amino Groups from the Two Ornithine Side Chains in the Three Different Conformational States of GS Model 1 (see Table 2 for Geometries and Figure 6 for Graphical Representations)

	2 Orn-$N^\delta H_3$—O═C D-Phe H-bonds			1 Orn-$N^\delta H_3$—O═C D-Phe H-bond[b]			no Orn-$N^\delta H_3$—O═C D-Phe H-bonds			
amide bond	NH_4^+	NH_3	estimated charge field effect	NH_4^+	NH_3	estimated charge field effect	NH_4^+	NH_3	estimated charge field effect	exp. effect[a]
Val	0.3	0.1	0.2	−0.2/0	−0.1/−0.3	−0.1/0.3	−0.3	0.1	−0.2	1.1
				−0.1	**−0.2**	**0.1**				
Orn	−1	−0.7	−0.3	−0.5/−1.1	−0.0/−1.1	−0.5/0.0	−0.7	−0.2	−0.4	0.8
				−0.8	**−0.5**	**−0.3**				
Leu	0.3	0	0.3	0.2/0.5	−0.1/0.1	0.2/0.4	0.4	0.1	0.3	−0.1
				0.3	**0**	**0.3**				
D-Phe	−0.9	−0.1	−0.8	−1.0/0.1	−0.1/−0.0	−1.0/0.1	0	−0.1	0	0.3
				−0.5	**0**	**−0.5**				
Pro	6.7	2.3	4.5	6.5/0.1	2.2/−0.1	4.4/0.1	−0.3	−0.6	0.3	3.5
				3.3	**1.1**	**2.2**				

[a] Taken from ref 34. [b] Charge field and hydrogen bonding effects on the ^{15}N shift of a given residue type differ in the two strands of the β-pleated sheet and are listed separately. In addition, average values for each residue type are shown in bold.

conformation that is highly similar to the X-ray structure, with just one intramolecular hydrogen bond between Orn $N^\delta H_3^+$ and the peptide backbone at any one time.

The charge field effect was shown above to be dependent on the distance and orientation of the charge with respect to a given amide bond. This implies that it should be possible to reproduce at least qualitatively the relative order of magnitudes of the experimental shift differences (Figure 6) by just using reduced model systems for an amide bond such as N-acetyl amide in combination with NH_4^+ or the even simpler model of point charges at positions which correspond to the center of the Orn $N^\delta H_3^+$. The results of such computations are presented in Table 3. In order to quantify the relative contributions to the overall shift difference resulting from hydrogen bonding and charge field effects, respectively, the shift difference calculations were performed both modeling $N^\delta H_3^+$ as NH_4^+ and as uncharged NH_3, while the molecular conformations and possible hydrogen bonds were left unaltered with respect to the complete GS template conformer 1.

The results are in very good agreement with the calculations done using the complete GS model. Indeed, the two sets of data correlate with coefficients of correlation of 0.97 and 0.91 for the structures with 2 or 1 intramolecular Orn-$N^\delta H_3$—O═C D-Phe hydrogen bonds, respectively, while no significant charge-relay effect is predicted for the conformation without any intramolecular hydrogen bond formed between Orn $N^\delta H_3$ and the peptide backbone. Once more, these calculations suggest that on average only one intramolecular Orn-$N^\delta H_3$—O═C D-Phe is formed by GS in solution at any one time. In addition, using charged NH_4^+ and uncharged NH_3 as a model for the Orn $N^\delta H_3$ enabled estimation of the individual contributions of hydrogen bonding and the charge field effect to the total ^{15}N chemical shift change on hydrogen bond formation to the carboxyl group of D-Phe. From these calculations, the charge field effect seems to account for 2.2 ppm (=2/3) of the total shift change, while a hydrogen bond from the Orn $N^\delta H_3$ to the carboxyl oxygen of the amide linkage results in a downfield shift of 1.1 ppm. In a separate calculation, the charge field effect of 2.2 ppm at the nitrogen atom was found to correspond to the effect of a simple positive point charge of 0.15 proton units, placed at the position of the Orn N^δ (instead of NH_4^+). Overall, the calculations with minimal model systems suggest therefore that charge-relay effects are very local in nature. While they seem to be strongly dependent on the hybridization state of the concerned atom, the particular conformation of the bonds to the neighboring atoms proves to be less important. To conclude, simple models for the immediate chemical environment of the nucleus under investigation should usually suffice to account for charge field effects in a quantitative manner.

Discussion

In this paper, we have demonstrated that the amide ^{15}N chemical shift is sensitive to charge fields in an orientation- and distance-dependent manner with respect to both the magnitude and sign of the effect. Hence, a notable change in charge density on an amide nitrogen nucleus is observed due to polarization effects from the solvent at the adjoining peptide bond.

Furthermore, the magnitude of the effect is significantly dependent on the dielectric properties of the environment, which can be taken into account quantitatively by the reaction field contributions from a polarizable-continuum model.[37] In a previous study that did not take into account such contributions,[24] it has been noticed that most accurate predictions of chemical shifts in different types of nitrogen nuclei are generally obtained for the nitrogen sp^3 hybrids, e.g., in amines or hydrazines. The fact that these nitrogen types are hardly polar and therefore little if any susceptible to charge, dipole, and reaction field effects may readily explain this observation.

Buckingham has shown[68] that chemical shift changes at any nucleus can be expressed in terms of a power series in the electric field at the nucleus, of which only linear and second-order electric field effects prove significant. In the case of a uniform field, the linear term is simply a result of polarization of charge along the bonds on the atom leading to a change in charge density at its nucleus, while the second-order term is due to orbital distortion.[69] By contrast, if the field source is a molecular point charge or dipole, it falls off over distance and the polarization of the bond depends now also on the field averaged by integration along the bond.[69] In fact, the linear electric field shift was shown to consist primarily of two terms, a uniform field and a field gradient term. From a semiempirical analysis of ^{13}C resonances in various small molecules, Batchelor confirmed that the maximum field gradient shift coefficients for unsaturated carbons are larger than those of saturated carbons as are the maximum uniform field shift coefficients, because saturated bonds have much smaller polarizabilities than unsatur-

ated bonds. It is therefore obvious that the higher polarity of amides due to their partial double-bond character makes the charge, dipole, and reaction field effects on amide ^{15}N chemical shifts due to the solvent much more significant and renders reliable computations difficult as long as such effects are not taken into account.[24]

In our work, the polarizable-continuum model handles this effect reasonably well and allows interpretation of the electric field shifts of amide nitrogens close to a field source. Since in particular the field gradient term of an unsaturated bond will show a very strong orientational dependence,[69] the linear electric field shift might therefore be expected to provide considerable structural information. In the case of the amide nitrogen, it is shown that the directions of maximum electric field shifts are close to the N−C′ and N−C$^{\alpha}$ bonds, while charges outside the plane defined by the peptide bond have significantly less effect on the chemical shift. Moreover, the directions of field gradient shifts are determined solely by the sign of the field gradient; i.e., positive and negative charges will produce downfield and upfield linear electric field effects, respectively.[69]

As an example, we have studied linear electric field effects in zwitterionic di-, tri-, and tetrapeptides. The values for the apparent N- and C-terminal (Mulliken) charges in water are predicted by the PCM method to be +0.1 and −0.2, respectively, while they are +0.6 and −1.2, respectively, in the gas phase. NMR shielding calculations of zwitterionic tetrapeptides confirm that the dielectric properties of water need to be taken into account in order to obtain chemical shift predictions that are in agreement with the experimental values measured in water.[29] Although the problem is virtually identical in zwitterionic di- and tripeptides, the computed shifts appear somewhat too downfield by about 5 ppm on average, as compared to the experimental shifts.[31] This points toward a systematic error in the reporting of the ^{15}N chemical shifts in the latter study, a problem that has frequently been encountered when different studies made use of different reference compounds.[70] Deviation plots between experimental and computed shifts such as the one presented in Figure 4 could provide a convenient way to detect such systematic biases that may have resulted from the use of different methods of referencing. Overall, the PCM approach results in ^{15}N shift predictions in zwitterionic tetrapeptides which are on average much closer to the experimentally observed values than the predictions made in the gas phase. Nevertheless, a systematic overestimation of the shift effects seen on protonation/deprotonation of the terminal amine or carboxyl groups of some 6−8 ppm, which has recently also been reported in a computational study on the L-Ala−L-Ala dipeptide,[46] will require clarification.

The utility of the electric field effect on amide ^{15}N chemical shifts has been further assessed by comparing ^{15}N chemical shift predictions for various conformational models of the cyclic decapeptide gramicidin S (GS) with experimental values. Due to the large size of this peptide, a discrete cluster model would be computationally much too demanding and dielectric solvation effects are preferably taken into account by the PCM approach. GS is particularly suitable to study electric field effects because it is strongly amphiphilic with two positively charged ornithine side chains on one side, while the backbone is relatively rigid as a result of four trans-annular hydrogen bonds. Indeed, our *de novo* computation of a GS model by MD, followed by geometry optimization from first principles (using DFT), resulted in very good agreement of the backbone torsion angles with the crystal structure of GS. Different side chain rotameric states were then imposed and the structures again energy optimized. In our analysis, the ^{13}C chemical shifts are not sensitive enough to side chain conformational effects to enable distinction and identification of different rotameric states; all of the proposed models give rise to ^{13}C shift predictions which agree reasonably well with the experimental data. In contrast, the charge field effect on the amide ^{15}N chemical shift arising from positive charges of two ornithine side chains was successfully utilized in the present study to distinguish different Orn rotameric states of GS in methanol. The experimentally observed shift differences of GS upon quaternization of Orn can be explained quantitatively on the basis of charge field and hydrogen bonding effects from the cationic ornithine side chains alone. Interestingly, the magnitude of the experimentally observed ^{15}N shift effect is much closer to the theoretically predicted charge-relay effect resulting from intramolecular hydrogen bonding of just one rather than both Orn side chains to the peptide backbone, as it is known from the crystal structure.[53,54] Moreover, tertiary structural states with two such intramolecular hydrogen bonds are calculated to be 10−20 kcal/mol higher in energy than the respective states with just one intramolecular hydrogen bond, which clearly corroborates our conclusion.

Together, this provides strong evidence that the nonsymmetric state with one intramolecular hydrogen bond, as observed in the crystal structure, is also prevailing in methanol. Moreover, using simple *N*-methylacetamide (NMA) models to mimic individual amide bonds, it could be shown that the overall shift effects on backbone amide nitrogen resonances due to close proximity of the cationic Orn side chain can be delineated into additive contributions from the charge field and the hydrogen bond donating capacity of the Orn side chain. The former contribution is reproduced simply using point charges as a substitute for the charged ammonium groups of ornithine at the respective positions, while the latter is computed by substituting the ammonium group with uncharged ammonia. In effect, excellent agreement between the values derived from the complete GS model and values computed using such simple NMA models was achieved. The present work demonstrates that charge field effects can be studied by DFT using simple model systems for the peptide bond in the presence of point charges. Such calculations should prove helpful in the validation of protein models in the surroundings of charges. In turn, the fractional size of an atomic charge can easily be probed if one or more amide ^{15}N nuclei are in sufficient proximity to pick up the charge field effect and provided that all the other structural parameters possibly contributing to the ^{15}N shift (conformation, hydrogen bonding) are known and quantified. Typical situations, in which charge field effects on amide ^{15}N shifts could be relevant, include charged side chains that are involved in long-range contacts to another site on a protein, charges introduced through binding of a ligand, pH effects at the protein termini, or specific binding of ions. Moreover, the results may contribute to the refinement of empirical or semiempirical ^{15}N chemical shift prediction tools[17,43,71] and reveal significant relationships between protein structure and chemical shifts, including medium- or long-range effects, that could possibly be used to further improve shift-based NMR structure determination.[10,14,23] Such chemical-shift-based structure determination protocols use as input experimental parameters the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, ^{15}N, $^{1}H^{\alpha}$, and $^{1}H^{N}$ NMR chemical shifts in conjunction with a fragment-based assembly method.[10,14] Recent advances have made it clear that accurate and extensive correlation between chemical shift and local conformation is crucial to the success of this approach.[10] For example, the most recently developed empirical prediction tool SPARTA considers only contributions from amino acid

types, backbone ϕ and ψ torsion angles and side chain χ^1 rotameric states for the prediction of ^{15}N chemical shifts, resulting in an overall RMSE between predicted and experimental values of 2.52 ppm.[71] The benefit of additionally considering electric field contributions may be twofold: (1) Conformational effects on ^{15}N chemical shifts should become better defined if corrections for nonconformational effects are taken into account; (2) in globular proteins, charge field effects can provide additional medium- or long-range spatial restraints for structure determination.

Acknowledgment. I thank Prof. Per Siegbahn for granting access to the computational resources at Stockholm University and KTH Stockholm. Financial support through a Fellowship by the Swiss National Science Foundation (PA002-109015) is gratefully acknowledged.

Supporting Information Available: Molecular modeling coordinates of three geometry optimized conformational states of gramicidin S conformer 1, ^{15}N and ^{13}C chemical shift predictions for the three gramicidin S models included in Table 2, and experimental and computed NMR shifts of linear oligopeptides and small organic molecules used to establish the linear relationship between shifts and computed shielding parameters for different basis sets and with or without polarizable-continuum models. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Di Fiori, N.; Orendt, A. M.; Caputo, M. C.; Ferraro, M. B.; Facelli, J. C. *Magn. Reson. Chem.* **2004**, *42*, S41−47.
(2) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996**, *104*, 5497–5509.
(3) Karplus, M. *Rev. Mod. Phys.* **1960**, *32*, 455–460.
(4) Pople, J. A *Proc. R. Soc. London, Ser. A* **1957**, *239*, 550–556.
(5) Markley, J. L.; Meadows, D. H.; Jardetzky, O. *J. Mol. Biol.* **1967**, *27*, 25–40.
(6) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311–333.
(7) Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490–5492.
(8) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
(9) de Dios, A. C.; Pearson, J. G.; Oldfield, E. *Science* **1993**, *260*, 1491–1496.
(10) Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K. K.; Lemak, A.; Ignatchenko, A.; Arrowsmith, C. H.; Szyperski, T.; Montelione, G. T.; Baker, D.; Bax, A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4685–4690.
(11) Tycko, R. *Q. Rev. Biophys.* **2006**, *39*, 1–55.
(12) Gryk, M. R.; Hoch, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4533–4534.
(13) McCoy, M. A.; Wyss, D. F. *J. Am. Chem. Soc.* **2002**, *124*, 11758–11763.
(14) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9615–9620.
(15) Delaglio, F.; Kontaxis, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 2142–2143.
(16) Gong, H.; Shen, Y.; Rose, G. D. *Protein Sci.* **2007**, *16*, 1515–1521.
(17) Xu, X. P.; Case, D. A. *J. Biomol. NMR* **2001**, *21*, 321–333.
(18) Wang, Y.; Jardetzky, O. *J. Am. Chem. Soc.* **2002**, *124*, 14075–14084.
(19) Sun, H.; Sanders, L. K.; Oldfield, E. *J. Am. Chem. Soc.* **2002**, *124*, 5486–5495.
(20) Vila, J. A.; Ripoll, D. R.; Scheraga, H. A. *J. Phys. Chem. B* **2007**, *111*, 6577–6585.
(21) Oldfield, E. *J. Biomol. NMR* **1995**, *5*, 217–225.
(22) de Dios, A. C.; Pearson, J. G.; Oldfield, E. *J. Am. Chem. Soc.* **1993**, *115*, 9768–9773.
(23) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289–302.
(24) Fadda, E.; Casida, M. E.; Salahub, D. R. *J. Phys. Chem. A* **2003**, *107*, 9924–9930.
(25) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531–1534.
(26) Stark, J.; Powers, R. *J. Am. Chem. Soc.* **2008**, *130*, 535−545.
(27) Le, H.; Oldfield, E. *J. Phys. Chem.* **1996**, *100*, 16423–16428.
(28) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3093.
(29) Braun, D.; Wider, G.; Wuethrich, K. *J. Am. Chem. Soc.* **1994**, *116*, 8466–8469.
(30) Kricheldorf, H. R. *Org. Magn. Reson.* **1981**, *15*, 162–177.
(31) Markowski, V.; Posner, T. B.; Loftus, P.; Roberts, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 1308–1309.
(32) Schwarzinger, S.; Kroon, G. J.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. *J. Am. Chem. Soc.* **2001**, *123*, 2970–2978.
(33) Jones, C. R.; Kuo, M.; Gibbons, W. A. *J. Biol. Chem.* **1979**, *254*, 10307–10312.
(34) Krauss, E. M.; Chan, S. I. *J. Am. Chem. Soc.* **1982**, *104*, 6953–6961.
(35) *Jaguar 5.5*; Schrödinger: Portland, OR, 1991−2003.
(36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.05; Gaussian, Inc.: Wallingford, CT, 2004.
(37) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43–54.
(38) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
(39) Sulzbach, H. M.; Schleyer, P. v. R.; Schaefer, H. F *J. Am. Chem. Soc.* **1994**, *116*, 3967–3972.
(40) Baldridge, K. K.; Siegel, J. S. *J. Phys. Chem. A* **1999**, *103*, 4038–4042.
(41) Scheurer, C.; Skrynnikov, N. R.; Lienin, S. F.; Straus, S. K.; Bruschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 4242–4251.
(42) Vila, J. A.; Scheraga, H. A. *Proteins* **2008**, *71*, 641−654.
(43) Xu, X. P.; Case, D. A. *Biopolymers* **2002**, *65*, 408–423.
(44) Honig, B.; Nicholls, B. *Science* **1995**, *268*, 1144–1149.
(45) Schwarzinger, S.; Kroon, G. J.; Foss, T. R.; Wright, P. E.; Dyson, H. J. *J. Biomol. NMR* **2000**, *18*, 43–48.
(46) Sychrovsky, V.; Budesinsky, M.; Benda, L.; Spirko, V.; Vokacova, Z.; Sebestik, J.; Bour, P. *J. Phys. Chem. B* **2008**, *112*, 1796–1805.
(47) Kongsted, J.; Mennucci, B. *J. Phys. Chem. A* **2007**, *111*, 9890–9900.
(48) Katsu, T.; Ninomiya, C.; Kuroko, M.; Kobayashi, H.; Hirota, T.; Fujita, Y. *Biochim. Biophys. Acta* **1988**, *939*, 57–63.
(49) Lee, D. L.; Hodges, R. S. *Biopolymers* **2003**, *71*, 28–48.
(50) Jones, C. R; Sikakana, C. T; Hehir, S; Kuo, M. C; Gibbons, W. A. *Biophys. J.* **1978**, *24*, 815–832.
(51) Stern, A.; Gibbons, W. A.; Craig, L. C. *Proc. Natl. Acad. Sci. U.S.A.* **1968**, *61*, 734–741.
(52) Xu, Y.; Sugar, I. P.; Krishna, N. R. *J. Biomol. NMR* **1995**, *5*, 37–48.
(53) Hull, S. E.; Karlsson, R.; Main, P.; Woolfson, M. M.; Dodson, E. J. *Nature* **1978**, *275*, 206–207.
(54) Tishchenko, G. N.; Andrianov, V. I.; Vainstein, B. K.; Woolfson, M. M.; Dodson, E. *Acta Crystallogr., Sect. D* **1997**, *53*, 151–159.
(55) Hodgkin, D. C.; Oughton, B. M. *Biochem. J.* **1957**, *65*, 752–756.
(56) Mihailescu, D.; Smith, J. C. *J. Phys. Chem. B* **1999**, *103*, 1586–1594.
(57) Allerhand, A.; Komoroski, R. A. *J. Am. Chem. Soc.* **1973**, *95*, 8228–8231.
(58) Komoroski, R. A.; Peat, I. R.; Levy, G. C. *Biochem. Biophys. Res. Commun.* **1975**, *65*, 272–279.
(59) Dygert, M.; Go, N.; Scheraga, H. A. *Macromolecules* **1975**, *8*, 750–761.
(60) Momany, F. A.; Vanderkooi, G.; Tuttle, R. W.; Scheraga, H. A. *Biochemistry (Moscow)* **1969**, *8*, 744–746.
(61) Khaled, M. A.; Urry, D. W.; Sugano, H.; Miyoshi, M.; Izumiya, N. *Biochemistry (Moscow)* **1978**, *17*, 2490–2494.
(62) Live, D. H.; Davis, D. G.; Agosta, W. C.; Cowburn, D. *J. Am. Chem. Soc.* **1984**, *106*, 1939–1941.
(63) Rae, I. D.; Scheraga, H. A. *Biochem. Biophys. Res. Commun.* **1978**, *81*, 481–485.
(64) Wang, Y.; Wishart, D. S. *J. Biomol. NMR* **2005**, *31*, 143–148.
(65) Pearson, J. G.; Le, H.; Sanders, L. K.; Godbout, N.; Havlin, R. H.; Oldfield, E. *J. Am. Chem. Soc.* **1997**, *119*, 11941–11950.

(66) Gibbons, W. A.; Sogn, J. A.; Stern, A.; Craig, L. C.; Johnson, L. F. *Nature* **1970**, *227*, 840–842.
(67) London, R. E.; Wingad, B. D.; Mueller, G. A. *J. Am. Chem. Soc.* **2008**, *130*, 11097–11105.
(68) Buckingham, A. D *Can. J. Chem.* **1960**, *38*, 300–307.
(69) Batchelor, J. G. *J. Am. Chem. Soc.* **1975**, *97*, 3410–3415.
(70) Wishart, D. S.; Nip, A. M. *Biochem. Cell Biol.* **1998**, *76*, 153–163.
(71) Shen, Y.; Bax, A. *J. Biomol. NMR* **2007**, *38*, 289–302.

JP807362V