JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Toward the Quantum Chemical Calculation of NMR Chemical Shifts of Proteins. 3. Conformational Sampling and Explicit Solvents Model

Thomas E. Exner,*[,†,‡] Andrea Frank,[†] Ionut Onila,[‡] and Heiko M. Möller*[,†]

[†]Department of Chemistry and Zukunftskolleg, University of Konstanz, 78457 Konstanz, Germany

[‡]Theoretical Medicinal Chemistry and Biophysics, Institute of Pharmacy, Eberhard Karls University Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

Supporting Information

ABSTRACT: Fragment-based quantum chemical calculations are able to accurately calculate NMR chemical shifts even for very large molecules like proteins. But even with systematic optimization of the level of theory and basis sets as well as the use of implicit solvents models, some nuclei like polar protons and nitrogens suffer from poor predictions. Two properties of the real system, strongly influencing the experimental chemical shifts but almost always neglected in the calculations, will be discussed here in great detail: (1) conformational averaging and (2) interactions with first-shell solvent molecules. Classical molecular dynamics simulations in explicit water were carried out for obtaining a representative ensemble including the arrangement of neighboring solvent molecules, which was then subjected to quantum chemical calculations. We could demonstrate with the small test system N-methyl acetamide (NMA) that the calculated chemical shifts show immense variations of up to 6 ppm and 50 ppm for protons and nitrogens, respectively, depending on the snapshot taken from a classical molecular dynamics simulation. Applying the same approach to the HA2 domain of the influenza virus glycoprotein hemagglutinin, a 32-amino-acid-long polypeptide, and comparing averaged values to the experiment, chemical shifts of nonpolar protons and carbon atoms in proteins were calculated with unprecedented accuracy. Additionally, the mean absolute error could be reduced by a factor of 2.43 for polar protons, and reasonable correlations were obtained for nitrogen and carbonyl carbon in contrast to all other studies published so far.

INTRODUCTION

NMR spectroscopy is an irreplaceable tool for investigating biomolecular systems. Besides structure determination of small to medium sized proteins, studies on the dynamics and kinetic processes of even very large systems like protein complex assemblies can be performed using a variety of techniques (for reviews, see refs 1−3). To support such investigations, accurate predictions of NMR chemical shifts are of great benefit since they can be used for structure evaluation and validation and linking NMR chemical shift perturbations to conformational changes. Due to advanced developments in the theoretical treatment but also to the formidable increase in computer power available, quantum mechanical NMR chemical shift calculations can now reach the accuracy needed for these applications, as demonstrated in a number of recent publications,[4−28] including the first two parts of this series of papers from our group.[29,30] Since the small review of methods in the introduction of the second part,[30] important new contributions have been made, which will be shortly highlighted here. Flaig et al.[26] studied the influence of the size of the quantum mechanical region in a QM/MM approach on the calculated chemical shifts. To reach convergence, the use of QM regions up to 1500 atoms was needed only possible due to the use of the density matrix-based linear-scaling coupled perturbed self-consistent field method developed in the group of Ochsenfeld.[31−34] They concluded that 6−10 Å of the environment around the considered nuclei have to be included, and that there is a large benefit of using a QM/MM approach to include additional surroundings. Zhu et al.[25] complemented the automatic fragmentation quantum mechanics/molecular mechanics (AF-QM/MM) approach[16] with a Poisson−Boltzmann model for implicit solvents, resulting in good improvements in the accuracy of ^{1}H and ^{13}C and to a lesser extent ^{15}N chemical shifts. Finally, the fragment molecular orbital method was used by Gao et al. to calculate chemical shifts for α-helix and β-sheet polypeptides[27] as well as for ubiquitin,[28] agreeing well with the experimental values.

In the first two parts of this series of papers, the dependence of the chemical shifts calculated with the fragmentation scheme of the adjustable density matrix assembler (ADMA) approach on the size of the quantum chemical surroundings, i.e., the size of the additional parts explicitly included in the quantum mechanical calculations, the level of theory, and the basis set has been investigated systematically.[29,30] In this third part, we will concentrate on two additional factors that strongly influence chemical shifts and which, thus, cannot be neglected: conformational averaging and solvent effects. Even if solvent effects have been included already in calculations on proteins using implicit solvent models,[24,25,30] the consideration of explicit solvent molecules has proven to be needed for the accurate chemical shift calculations of small molecules,[24,35−38] especially in the case of hydrogen bonding between solvent and solute. Additionally, it was shown that the effects of small conformational changes including zero-point vibration are relevant for small molecules,[18,21,35,39−44] and we expect that

Received: August 8, 2012
Published: October 1, 2012

   dx.doi.org/10.1021/ct300701m | J. Chem. Theory Comput. 2012, 8, 4818−4827

this is even more the case in large biomolecular systems and especially within flexible loops of proteins. To account for these effects in our calculations, we performed classical molecular dynamics (MD) simulations and extracted solute conformations with their correspondent solvent shells from these. Quantum mechanical NMR chemical shift calculations were done using a limited amount of water molecules from these snapshots. Since a representative ensemble of structures is needed for reasonable averages, such calculations are computationally extremely demanding, and we first investigated the relevant trends on a very small test model for a peptide bond: N-methyl acetamide. In order to generalize our observations, we applied our approach to a small polypeptide, the HA2 domain of the influenza virus glycoprotein hemagglutinin

MATERIALS AND METHODS

As already described in the Introduction, we will test here the influence of conformational averaging and explicit solvent molecules on the quantum chemically calculated NMR chemical shifts first for N-methyl acetamide (NMA) as a simple model for peptide bonds and then for the 32-amino-acid-long HA2 domain of the influenza virus glycoprotein hemagglutinin. For generating an ensemble of the molecules in solution, classical molecular dynamics (MD) simulations were used. For representative snapshots of these simulations, the chemical shifts are calculated on the basis of the fragmentation scheme of our adjustable density matrix assembler (ADMA) approach.[45−49] These are then compared to experimental data taken from the BMRB[50] or measured in-house.

Molecular Dynamics Simulations. All simulations were performed with the Amber10 suite of programs.[51] The structure of the HA2 domain was taken from the Protein Data Bank[52] (PDB entry 2KXA[53]). NMA was generated manually twice with a cis and trans peptide bond, respectively. Additionally, dimers for both forms were generated. These were then optimized with density functional theory (DFT) calculations using the B3LYP functional[54] and the 6-31g(d) basis set.[55−62] The MD simulations were performed for the trans isomer. Force-field parameters were taken from the general amber force field (GAFF) with AM1-BCC charges. For the HA2 domain, the modified version of the Cornell et al. force field[63] (parm99SB) was used. The starting structures were placed in a periodic truncated octahedron of TIP3P water molecules,[64] and counterions were added to maintain electroneutrality of the system (2 Na^+ for HA2). The borders of the truncated octahedron were chosen to be at least 15 Å and 10 Å from every solute atom for NMA and HA2, respectively.

The time step for all MD simulations was set to 2 fs with a nonbonded cutoff of 12 Å. The particle mesh Ewald (PME) method[65] was used to treat long-range electrostatic interactions and the SHAKE method[66] to constrain bond lengths of bonds involving hydrogen atoms. Both systems were first minimized with a maximum of 5000 steps. For equilibration, the system was then heated from 100 to 300 K for 100 ps and then relaxed to a density corresponding to 1 bar for 300 ps sampling in the canonical (NVT) and the isothermal isobaric (NPT) ensemble, respectively. To keep the structure close to the experiment during the equilibration, harmonic restraints with force constants of 5 kcal mol^{-1} $Å^{-2}$ were applied to all atoms of the solute in these simulations. These restraints were then gradually reduced to zero over 500 ps of NVT-MD. The production run for NMA had an overall time of 10 ns (NVT ensemble), and snapshots were taken every 2 ps, resulting in 5000 snapshots. For HA2, the corresponding numbers are 5 ns, 10 ps, and 500 snapshots.

Quantum Chemical NMR Chemical Shift Calculations. For large systems, the use of standard quantum chemical methods is not feasible because of the scaling behavior of such calculations. But fragment-based approaches, where the macromolecule is divided into small parts for which independent calculations can be performed, have proven successful for many molecular properties including NMR chemical shifts. The NMR chemical shift calculations based on the ADMA fragmentation scheme have been described in detail elsewhere.[29,30] Briefly, the macromolecule is first divided into a set of small fragments. Independent calculations using the Gaussian 09 program package with the GIAO (gauge invariant/including atomic orbitals) formalism[67−70] for the NMR isotropic chemical shieldings are then performed. If not otherwise mentioned, the B3LYP[54] hybrid functional with the 6-31g(d) basis set[55] was used. To ensure that influences of the chemical environment are taken into account, not only the fragments themselves but also their surroundings up to a specified distance d are included in the calculation. We will call this combination of a fragment, its surroundings, and the needed capping hydrogen atoms to fill bonds broken in the fragmentation scheme parent molecules in the following. With larger distances d, more of the original macromolecule is covered in the calculation, leading to a higher accuracy. Therefore, this distance criterion can be used to compromise between accuracy and computational speed. Please note that for the calculations with explicit solvents described here, beside parts of the macromolecule, also all water molecules, which are closer to the solute than d, are included in the calculations. From these calculations, the chemical shieldings only of the inner fragment are extracted, subtracted from the corresponding standards calculated at the same level of theory but neglecting solvent effects (tetramethylsilane for ^1H and ^{13}C and ammonia for ^{15}N) to obtain the NMR chemical shifts, and compared to the experiment and calculations based on a single structure (optimized structure and first NMR model in the PDB file for N-methyl acetamide and HA2 domain, respectively). Since each atom is part of exactly one fragment, chemical shifts for all nuclei are obtained in this way. Since we are interested in the chemical shifts for the thermodynamic ensemble of structures and not for minimum structures, the molecular coordinates were taken as provided by the MD simulation without further quantum mechanical optimization.

NMR Spectroscopy. N-methyl acetamide was dissolved in phosphate buffer (H_2O/D_2O = 95:5, 50 mM phosphate, 150 mM NaCl, pH 8) to a final concentration of 50 mM. NMR spectra were acquired at 295 K on a Bruker Avance III 600 MHz spectrometer equipped with a TCI-H/C/N triple resonance cryoprobe. Water suppression was achieved with excitation sculpting.[71] Spectra were processed and analyzed with Bruker's TopSpin software (v3.0). The ^{13}C and ^{15}N chemical shifts were determined from the indirect dimensions of HSQC spectra. Spectra were referenced to internal TSP-d_4 (sodium trimethylsilyl propionate). The ^{13}C and ^{15}N chemical shifts were indirectly referenced as described by Bax and Subramanian[72] and Wishart et al.,[73] respectively, using the conversion factors published for DSS (sodium 4,4-dimethyl-4-silapentane-1-sulfonate).

RESULTS AND DISCUSSION

NMR Chemical Shifts Predictions of N-Methyl Acetamide (NMA) Depending on the Level of Theory and Basis Set Size. We and others have realized in previous studies[4−30] that QM calculations of proteins and peptides yield unsatisfactory results for CO and H^N and N^H most likely due to influences of the solvent, hydrogen bonding, and/or conformational averaging that are not properly reflected in the QM approach. To gain insight in this regard, we performed a detailed investigation with NMA. This molecule was chosen since it is one of the simplest chemical structures containing a peptide bond, and many of the findings discussed in the first two papers of this series[29,30] will be relevant for this molecule. Additionally, its 3D structure/conformational ensemble can be calculated without excessive computational demand, and experimental problems like wrong assignments and structural uncertainties can be avoided. For similar reasons, Moon and Case[43] used the same molecule for their comparison of Hartree−Fock, MP2, and DFT up to the complete basis set (CBS) limit. But since some of the theory/basis set combinations used for the protein calculations in our previous publications[29,30] were not included in the publication, we reperformed some of these calculations and added new ones to be able to directly transfer the results to the studies on solvent effects and conformational averaging described below. The complete results of these new calculations are found in Tables S1−S3, Supporting Information; the experimental values are always provided in column A of these tables for comparison. Our new results, the ones from Moon and Case,[43] and the ones from our earlier publications on oligopeptides and proteins[29,30] converge to a very consistent picture. MP2 with large basis sets gives the best agreement. This is especially important for nitrogen atoms, where all density functionals and small basis sets give values which are too low-field and high field shifted, respectively. For these atoms, some density functionals with small basis sets only give reasonable results because of the very favorable but unreliable cancellation of errors. For carbonyl carbons, the use of large basis sets is much more important than the use of post-HF methods, whereas the results for the remaining carbons and protons are less dependent on both the level of theory and basis set size. For the latter cases, no clear advantage of using a high level of theories and very large basis sets is observable for NMA.

We want to note that, for reasons of statistical significance, calculations of the small molecule NMA should be interpreted with caution; however, very similar trends were found by Moon and Case[43] as well as in our previous publication.[29,30]

Furthermore, our as well as previous results on the influence of the level of theory and basis set size[6,39,40,43,74−76] were obtained without taking any solvent effects into account but compared to experimental spectra obtained in aqueous solution. That solvent effects are important can be seen from other studies on small molecules[24,35−38,42,43,77] and especially from the different experimental values measured in this study and, e.g., used in the study of Moon and Case.[43] For the nitrogen, a chemical shift of 114.24 ppm is measured in aqueous solution while a value of 106.00 ppm is given in the literature,[43] obtained for NMA dissolved in DMSO. Third, the large errors in the chemical shifts of amide protons in all theory/basis set combinations can be explained by direct, first-solvent-shell effects, i.e., the formation of hydrogen bonds with solvent molecules, as will be shown in the next section. A further aspect neglected so far and becoming more important in larger molecules is conformational averaging since the experimental values are not a property of the most stable conformation but of the thermodynamic ensemble at a given temperature and pressure.

NMR Chemical Shifts Predictions of N-Methyl Acetamide (NMA) in Dependence on Conformational Changes and Explicit Solvents. To further investigate the importance of these additional influences, ensembles were generated using classical molecular dynamic simulations. From these simulations, snapshots were extracted in uniform time intervals, and quantum chemical calculations were performed for these sets of snapshots to obtain the desired NMR chemical shifts and an impression of their dependence on the conformation of the macromolecule as well as the specific localization of solvent molecules. To reasonably represent the thermodynamic ensemble, 5000 snapshots over a time period of 10 ns were calculated. Four-Ångstrom surroundings were used to limit the computational demand, and the resulting parent molecules were embedded into an implicit solvent model (IEF-PCM[78−80]) to describe additional solvent effects. The same snapshots were also taken without any explicit solvent molecules but NMA directly embedded into the implicit solvents to quantify the importance of first-solvent-shell effect. Since the amide proton is directly involved in hydrogen-bonding to the solvent, we will start our discussion with this atom. Figure 1 demonstrates clearly that the chemical shift depends very much on the snapshot used in the calculation and, thus, on the relative orientation of the surrounding solvent molecules. Overall, the calculated values span a range of 6 ppm with the minimum at 4.13 ppm and the maximum at 10.12 ppm. The time series also demonstrates that changes in the chemical shift are not caused by different long-lived states representing conformational changes in the molecule or specific solvent structures, which could be artifacts of the simulations, but are due to rapid fluctuations in an equilibrated state. These fluctuations are many orders of magnitude faster than the chemical shift time scale, and thus, only the averaged values are observed in the experiment. In

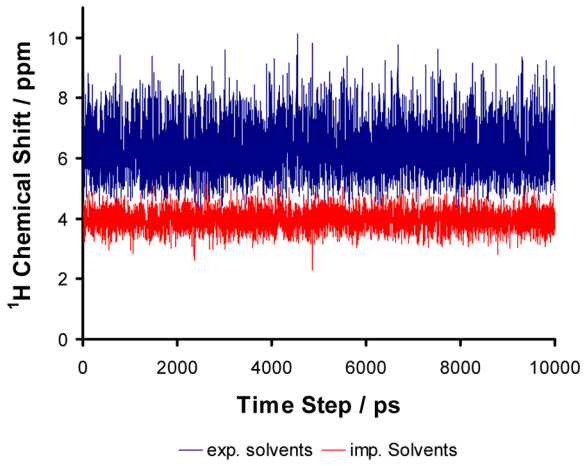

Figure 1. Time series of the ^{1}H chemical shift of H^N of the amide group in N-methyl acetamide (NMA) calculated from snapshots taken every 2 ps of a 10-ns-long MD simulation. Explicit solvents corresponds to a 4 Å solvent shell around the molecule, while the rest of the solvents are approximated by PCM. In the implicit solvent calculation, all water molecules are removed and replaced by PCM.

dx.doi.org/10.1021/ct300701m | J. Chem. Theory Comput. 2012, 8, 4818−4827

Figure 2, the distributions of the chemical shift over the 5000 snapshots using the combined explicit/implicit solvent model

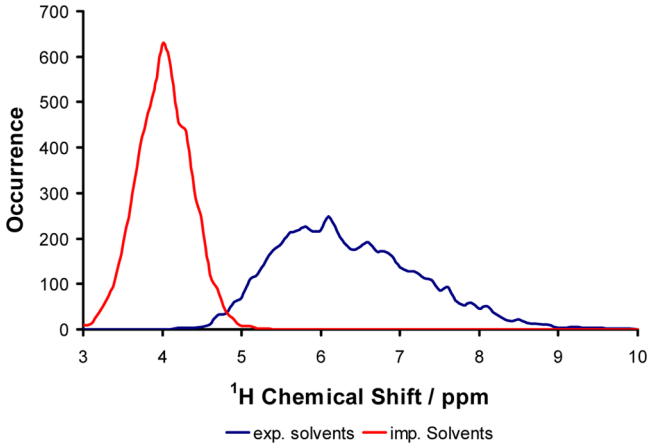

Figure 2. Probability distribution of the ^{1}H chemical shifts of H^N of the amide group in N-methyl acetamide (NMA) calculated from 5000 snapshots along a 10-ns-long MD simulation.

and the implicit-solvent-only model are shown. While for the implicit solvents a relatively narrow distribution of approximate Gaussian shape around an average value of 3.97 ppm (standard deviation STD = 0.34 ppm) is observed, the inclusion of explicit solvent molecules broadens the distribution and pushes it to much higher chemical shifts. The average value in this case is 6.31 ppm with a STD of 0.90 ppm. Only using the 500 snapshots from the first nanosecond of the simulation gave almost exactly the same average value (6.29 ppm with a STD of 0.89 ppm), once more demonstrating the good equilibration and coverage of the thermodynamic ensemble.

Taking solvent and conformational averaging into account lowered the deviation from experimental results by a factor of 2.43; thus, even if this value is still over 1 ppm lower than the experimental value (7.88 ppm), the calculations on the snapshots clearly demonstrate the importance of first-solvent-shell effects (even in the approximate way of the TIP3P[64] water model) especially for polar protons.

To explain these effects, we took a closer look at specific snapshots. In Figure 3, four representative orientations of the solvent molecules are shown. These correspond to the snapshot with the maximum and minimum value of the amide proton chemical shift as well as two arbitrarily chosen snapshots with chemical shifts in the range of 6 and 8 ppm, respectively. The images show that the chemical shift decreases with increasing distance to the closest hydrogen bond acceptor, i.e. the closest oxygen atom of a solvent molecule. While for chemical shifts of 10 ppm and 8 ppm a short hydrogen-bonding distance and also angles close to the optimal values for a hydrogen bond are determined, for the smaller chemical shifts not one but two solvent molecules have almost the same atomic distance to the amide proton. Even if other geometric features are also important, there is a clear correlation between the distance between the proton and the hydrogen-bond acceptor, i.e. the strength of the hydrogen bond, with the chemical shift (see Figure 4). Especially for short distances, the chemical shift decreases almost linearly with a very steep slope with increasing atomic distance. At around 2.2 Å, the hydrogen bond is so weak that its influence on the chemical shift is almost negligible, resulting in chemical shifts distributed around an average value of 5−6 ppm independent of the actual distance. This dependence of the chemical shift on the hydrogen-bond strength can also be observed when calculating dimers of NMA forming intermolecular hydrogen bonds (see Supporting Information). The dimer of the cis isomer forms two very strong hydrogen bonds, and a chemical shift of 9.33 ppm is calculated with the B3LYP/6-31g(d,p). The trans isomer can only maintain one hydrogen bond, and the chemical shift of 7.55 ppm shows that this one is also weaker than the one in the cis isomer. In contrast, the chemical shift in both monomers is around 4 ppm since the formation of a hydrogen bond is not possible. It is surprising that this solvent effect is almost always neglected in theoretical studies since it has been well-known for

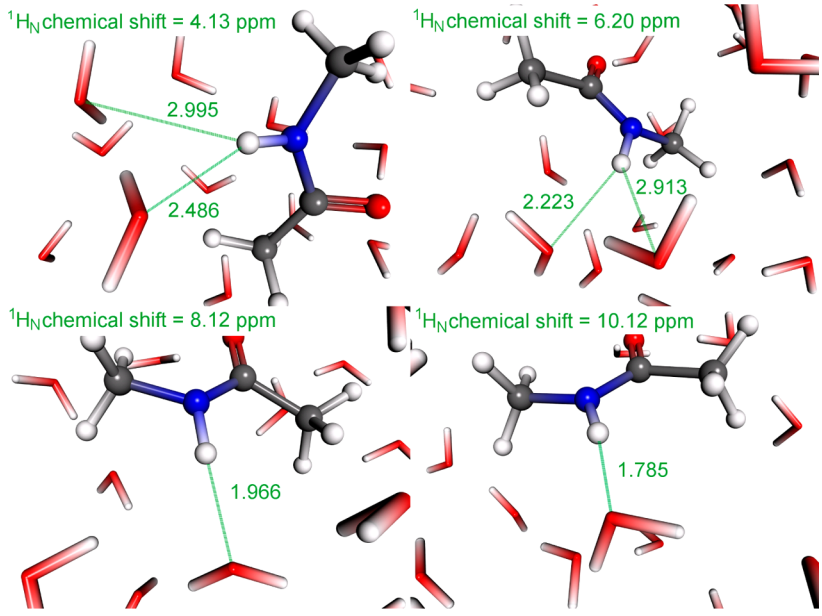

Figure 3. Four representative snapshots from the 10-ns-long MD simulation of N-methyl acetamide (NMA) demonstrating the dependence of the ^{1}H chemical shift of H^N on the distance to the next water oxygen atom.

dx.doi.org/10.1021/ct300701m | J. Chem. Theory Comput. 2012, 8, 4818−4827

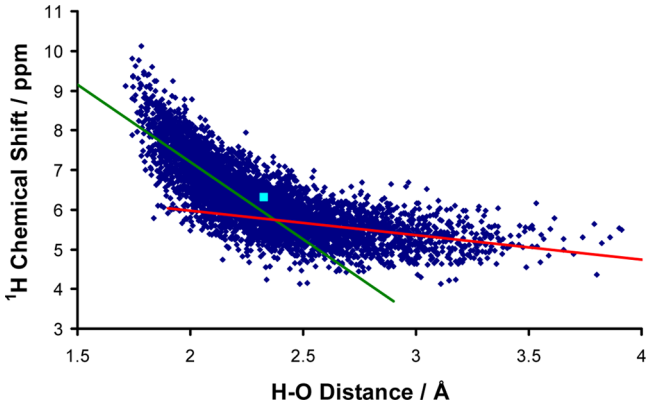

Figure 4. Dependence of the ^{1}H chemical shift of H^N on the distance to the next water oxygen atom as seen in the 10-ns-long MD simulation of N-methyl acetamide (NMA): For short distances, a strong dependence is evident, expressed by the green correlation line for all distances smaller than 2.4 Å. For longer distances (red line), the chemical shifts are almost independent of the actual distance. On average, a distance of 2.328 Å and a chemical shift of 6.31 ppm result as highlighted by the dot in light blue.

a long time in the NMR community (see ref 81 and references therein).

Even if they do not form a hydrogen bond with the solvents, the protons of the acetyl methyl group are also significantly influenced by the explicit solvent molecules, showing a shift to lower field of 0.8 ppm (implicit solvents, 1.63 ppm; explicit solvents, 2.43 ppm). The experimental value is somewhat in between at 1.99 ppm, and thus, using explicit solvent molecules seems not to be advantageous. But this could be an artifact of the relative small basis set used. For the protons of the N-methyl group, another interesting effect is observed. Even if the average chemical shift does not change significantly (implicit solvents, 2.75 ppm; explicit solvents, 2.93 ppm), the distribution of chemical shifts is very different (see Figure 5). When only using the implicit solvent model, the methyl proton closest to the carbonyl oxygen is low-field shifted compared to the other two protons. This results in two distinct peaks in the distribution. Due to rotation of the methyl group and the resulting fluctuations in the H−O distance and, thus, strength of the influence of the electrostatic field generated by the

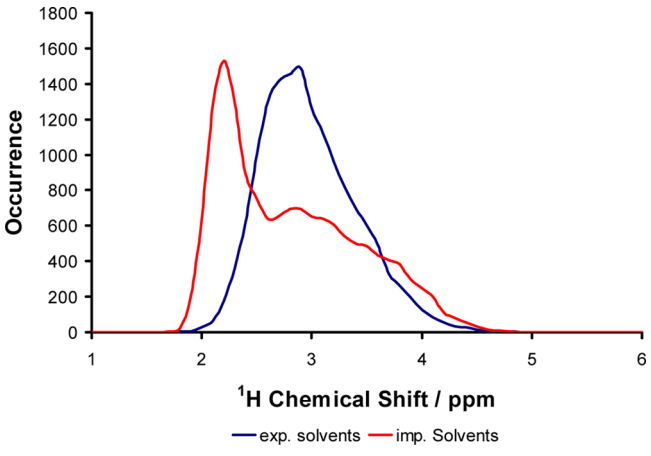

Figure 5. Probability distribution of the ^{1}H chemical shifts of the protons of the N-methyl group in N-methyl acetamide (NMA) calculated from 5000 snapshots along a 10-ns-long MD simulation.

carbonyl group including the electron lone pairs, one of these peaks is very broad. In the explicit-solvents calculation, the two peaks fall almost onto each other since now the other two protons experience an electrostatic field generated by the solvent molecules similar to the intramolecular one.

Regarding the other nuclei, the two methyl carbon atoms show only minor dependence on the conformation and on the solvent model used. The distribution slightly broadens when explicit solvent molecules are included and the average chemical shift is moved to lower field by approximately 3 ppm (see the Supporting Information). In contrast, the carbonyl carbon atom shows a larger dependence of the chemical shift on the conformation of the molecule (STD = 5.4 ppm compared to around 3 ppm for the methyl groups, see Figure 6). It is surprising that even for a small molecule like

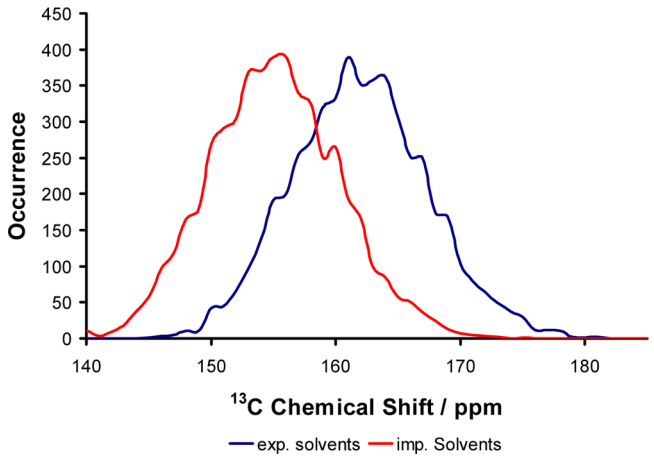

Figure 6. Probability distribution of the ^{13}C chemical shifts of the carbonyl group in N-methyl acetamide (NMA) calculated from 5000 snapshots along a 10-ns-long MD simulation.

NMA the conformational ensemble has a dominating influence on the distribution of carbonyl shift compared to fluctuations of the solvent shell. This is demonstrated by the fact that for the implicit solvent model almost the same STD is obtained (5.2 ppm). This probably documents the increased importance of paramagnetic shielding effects for the carbon chemical shift as compared to protons. But the solvent has a significant influence on the average value. With implicit and explicit solvents, values of 154.60 ppm and 161.43 ppm are calculated, respectively, so that the explicit solvents push it closer to the experimental value of 177.24 ppm. Nevertheless, the larger basis set with its downfield-shifting tendency is probably still needed to obtain a correct prediction.

Finally, the ^{15}N chemical shifts show an even broader distribution. Here, the STD is 7.9 ppm and 6.6 ppm for explicit and implicit solvents, respectively. Thus, both conformational changes and solvent reorganization are important for the broadening. Overall, the chemical shifts of this single atom span a range of more then 40 ppm, which is the same order of magnitude as the experimental range of ^{15}N chemical shifts seen in proteins. Due to this extreme dependence on small variations of geometrical features, it is not surprising that it is impossible to get reliable predictions of ^{15}N chemical shifts based on one single structure. A second observation, which can be made when looking at Figure 7, is the very significantly different average values of 116.90 ppm and 98.70 ppm for the explicit and implicit solvents models, respectively. Including

dx.doi.org/10.1021/ct300701m | J. Chem. Theory Comput. 2012, 8, 4818−4827

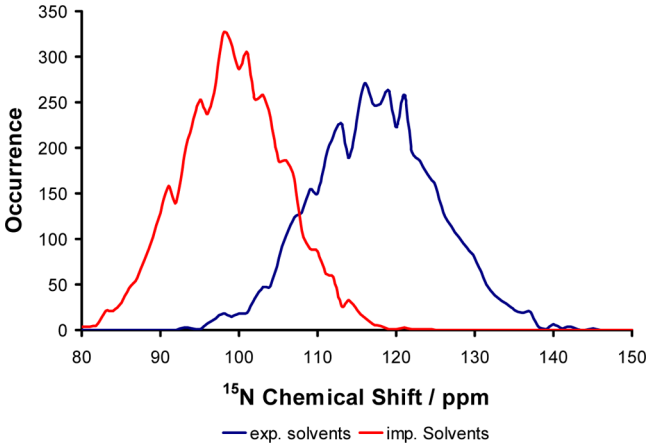

Figure 7. Probability distribution of the ^{15}N chemical shifts of the amide group in N-methyl acetamide (NMA) calculated from 5000 snapshots along a 10-ns-long MD simulation.

conformational averaging in calculations with implicit solvents shows in this example an inferior agreement with experimental results in aqueous solution (exptl. value: 114.24 ppm) than the in vacuo calculation using only one, energy-optimized structure (109.90 ppm). Due to the downfield shift tendency of adding explicit solvent molecules seen for all nuclei, the calculation is getting back close to the experimental value. Similar conclusions were drawn by Cai et al.[24] Their calculations for NMA showed that the augmentation of implicit solvent models with explicit water molecules has a significant influence on isotropic ^{15}N chemical shifts (but not as much as on the chemical shift anisotropy).

NMR Chemical Shifts Predictions of the HA2 Domain in Dependence on Conformational Changes and Explicit Solvents. Overall, the distributions of the NMR chemical shifts of the HA2 domain over the snapshots of the MD simulations show very similar behavior as in NMA. Nitrogen atoms and carbonyl carbon atoms have broader distributions in comparison to all other carbon atoms. For the ^{1}H chemical shift, the extreme broadening in the distribution of atoms, which are involved in hydrogen bonding, also manifests in the calculation of the polypeptide. But due to the larger size, two kinds of hydrogen bonds are possible. Intramolecular hydrogen bonds show no difference in the distribution using the implicit-only as well as the explicit solvents model. If the hydrogen bond is formed with a water molecule, the implicit solvent model is again not able to describe this situation, and the chemical shifts are predicted between 4 and 5 ppm. The inclusion of explicit solvent molecules raises the prediction by up to 3 ppm. This is exemplarily shown in Figure 8 with the H^N of amino acids LEU2 and PHE9 located on the solvent accessible surface and being buried deeply inside of the peptide, respectively.

More important than the distributions by themselves is the answer to the question of whether averaging over all different conformations improves the predictions. From Table 1 and Figure 9, it can be clearly seen that this is indeed the case for protons. The errors for all protons as well as for the nonpolar hydrogens are cut by up to 50% compared with a single point calculation with the same level of theory and basis set based on the experimental structure. If one neglects the polar hydrogens the correlation with experimental results is almost perfect, and the mean absolute error (MAE) of 0.26 ppm is much smaller than for any other method neglecting solvents and conformational averaging applied to the same polypeptide so far (single-point calculations with SHIFTX+,[82] the structure-based part of SHIFTX2:[83] MAE = 0.59 ppm[30] and with best quantum chemical calculation mPW1PW91/6-311g(d)/PCM: MAE = 0.45 ppm[30]). This is the case, although the B3LYP functional was used in the calculations performed here, which has been shown to be inferior to mPW1PW91. The results for the explicit and the implicit solvents model are very comparable, showing that the conformational averaging is more important then explicit solvent molecules for nonpolar atoms. Additionally, the large low-field shift of protons involved in hydrogen bonding to water reduced the maximum error from 4.80 ppm

Table 1. Mean Absolute Error (MAE) between Calculated and Experimental ^{1}H NMR Chemical Shifts of the HA2 Domain[a]

basis set	conformation	solvents model	all hydrogens (ppm)	without H^N (ppm)
6-31g(d)	MD	explicit	0.59	0.26
6-31g(d)	MD	implicit	0.68	0.26
6-311g(d)	SP	implicit	0.86	0.45
6-31g(d)	SP	implicit	0.95	0.52

[a]The calculations were performed on the experimental structure (single point, SP) and on representative snapshots of the molecular dynamics simulation (MD).

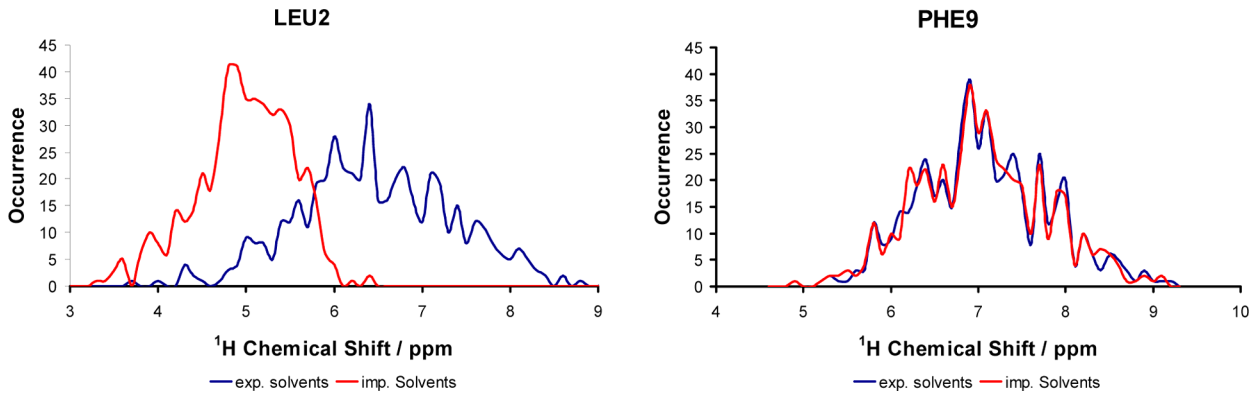

Figure 8. Probability distribution of the ^{1}H chemical shifts of the amide group of residues LEU2 and PHE9 calculated from 500 snapshots along a 10-ns-long MD simulation of the HA2 domain.

dx.doi.org/10.1021/ct300701m | J. Chem. Theory Comput. 2012, 8, 4818−4827

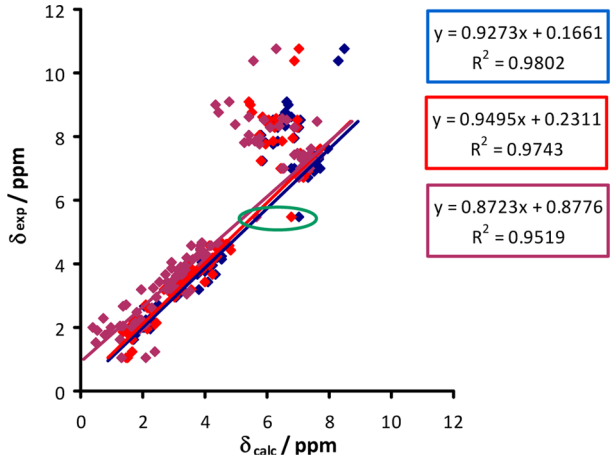

Figure 9. Correlation of ^{1}H NMR chemical shifts of the HA2 domain: blue, B3LYP/6-31g(d), MD, explicit solvents; red, B3LYP/6-31g(d), MD, implicit solvents; magenta, B3LYP/6-31g(d), single point, implicit solvents. Chemical shifts of H^{N} protons are indicated but have not been included in the linear regressions.

for the single-point calculation to 3.73 ppm for the MD snapshots with implicit solvents to finally 2.48 ppm when explicit solvent molecules are considered. Nevertheless, it would still be desirable to decrease these deviations further and identify outliers in order to improve QM-shift-based structure validation and automated assignment procedures. Taking also the results of NMA into account, one explanation for the relatively large errors of polar hydrogens is that the AMBER force field, and perhaps classical force fields in general, predicts hydrogen bonds to be too long and too weak. Studies are on their way using *ab initio* MD simulations to prove this assumption.

Only one nucleus (HD1 from TRP21) can be clearly identified as an outlier, for which the prediction deteriorates if MD snapshots are used. This is clearly a consequence of the MD trajectory evaluated here. In the experimental structure, this nucleus forms a T-shaped proton−π interaction with TYR22 in the experimental structure (Figure 10). But TYR22 turns out to be rather flexible in the simulation and moves away from TRP21 in a very early stage of our simulation. Thus, for most of the MD, the TRP21-HD proton is not positioned centrally in the anisotropy field of TYR22's phenyl ring, causing

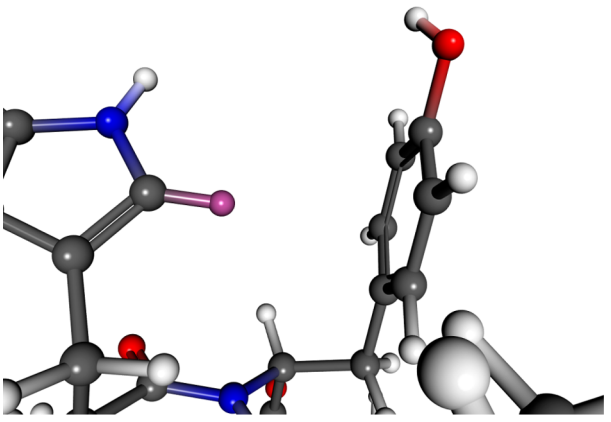

Figure 10. T-shaped interaction of TRP21 and TYR22 as seen in the NMR structure of the HA2 domain.

the QM calculation to yield chemical shifts that approach those of a largely solvent exposed Trp residue.

For carbon atoms, the inclusion of conformational averaging and explicit solvents also leads to an improved performance compared to the single-point calculations (see Table 2). But,

Table 2. Mean Absolute Error (MAE) between Calculated and Experimental ^{13}C NMR Chemical Shifts of the HA2 Domain[a]

basis set	conformation	solvents model	carbons (ppm)
6-31g(d)	MD	explicit	5.15
6-31g(d)	MD	implicit	5.51
6-311g(d)	SP	implicit	1.85
6-31g(d)	SP	implicit	7.11

[a]The calculations were performed on the experimental structure (single point, SP) and on the snapshots of the molecular dynamic simulation (MD).

the systematic errors caused by the small double-ζ basis set already seen in our previous publications[30] are still present. Therefore, the mean absolute error is still almost 3 times larger than the one using the larger 6-311g(d) basis set with the experimental structure even if the correlation is slightly better when the solvents are considered explicitly (see Supporting Information). This improved correlation becomes even clearer if the CO groups are analyzed separately from the rest of the carbon atoms (see Supporting Information). For the aliphatic nuclei, an almost perfect correlation (r^2 = 0.9921) is obtained (see also Supporting Information), which is even more remarkable when considering the much narrower chemical shift distribution of this set of nuclei compared to all carbons including carbonyl groups.

The correlation of the CO chemical shifts (r^2 = 0.5788, see Supporting Information) is significantly improved in our ensemble calculations when compared to the single-point calculations using the 6-31g(d) and even the 6-311g(d) basis set, r^2 = 0.4638 and r^2 = 0.4604, respectively. However, this correlation coefficient is still far from a situation that would allow reliable structure validation. To summarize these findings, we are confident that the combination of a triple-ζ basis set like 6-311g(d) with conformational averaging and explicit solvent molecules would lead to excellent mean absolute errors as well as correlation coefficients.

^{15}N chemical shifts still show the largest errors of all nuclei considered here. They even increase if conformational averaging and explicit solvent molecules are considered. Thus, besides the favorable compensation of errors, when using DFT in combination with small basis sets discussed in our previous publication,[30] also the neglect of conformational sampling and explicit solvents contributes to the apparent accuracy of the single-point/small basis set calculations (see Table 3). Having stated this, there are also some very promising aspects of the new explicit solvent calculations. The square of the correlation coefficient is improving from r^2 = 0.5829 for the single-point calculation to 0.7277 and 0.8117 when using conformational averaging and conformational averaging/explicit solvents, respectively (see Figure 11). This demonstrates that the stochastic error can be significantly decreased by the approaches followed here. Since the systematic error could be a result of problems of the NH_3 used as standard and it could be circumvented by an internal one,[30] highly accurate quantum chemical ^{15}N chemical shift calculations are in reach if large

Table 3. Mean Absolute Error (MAE) between Calculated and Experimental ^{15}N NMR Chemical Shifts of the HA2 Domain[a]

basis set	conformation	solvents model	nitrogens (ppm)
6-31g(d)	MD	explicit	14.59
6-31g(d)	MD	implicit	11.55
6-311g(d)	SP	implicit	43.89
6-31g(d)	SP	implicit	7.84

[a]The calculations were performed on the experimental structure (single point, SP) and on representative snapshots of the molecular dynamic simulation (MD).

basis sets are combined with conformational averaging, explicit solvents, and the use of standards more similar to the groups present in proteins. At the moment, such calculations are restricted by the extreme amount of computer resources needed for the averaging of hundreds of calculations for each fragment.

CONCLUSION

The results presented here demonstrate that both conformational changes as well as the location of explicit solvent molecules have a massive influence on the NMR chemical shifts. If both effects are included, the calculated values for the amide proton in N-methyl acetamide (NMA) span a range of 6 ppm depending on the actual conformation and solvent shell arrangement taken from individual snapshots of a classical MD simulation, and the average value over all snapshots is shifted by more than 2 ppm to lower field reducing the error compared to the experiment by a factor of 2.43. Even if the other protons are less affected, differences of more than 2 ppm are still seen between snapshots. For ^{13}C and ^{15}N, the values span ranges of 30 and 50 ppm, respectively, which is especially remarkable if one considers that the calculated distribution of the single nitrogen nucleus in NMA has a comparable magnitude to that of the full range of all experimentally determined amide nitrogen atoms in proteins.

The effects of the better approximation of the real system when including conformational averaging and explicit solvents (as seen in the small test model) manifest themselves in a higher accuracy of the predicted chemical shift of the HA2 domain. The mean absolute error for nonpolar protons almost halved compared to the single-point calculation using the same level of theory/basis set combination, and the overall accuracy surpasses any other method applied to the same polypeptide so far, even if these used larger basis sets and specially selected density functionals. An even larger relative improvement is seen for the polar protons forming hydrogen bridges to the solvents, demonstrating the importance of explicit solvent molecules. The still relative large errors of these latter protons can probably be attributed to the approximations of the TIP3P water model and other parameters of the AMBER force field, which seem to result in too long and too weak hydrogen bonds. This is supported by the work of Dracinsky et al.,[35−37] in which the benefits of *ab initio* molecular dynamics simulations to obtain starting conformations was shown. For carbon atoms, the correlation and the mean absolute errors compared to the experiment were excellent already in the single-point calculations. From conformational averaging, very comparable values result as from a single experimental structure. Thus, large basis sets are more important than ensemble calculations and solvent effects especially for the carbonyl group despite the large variation of values seen between the snapshots. Performing averaging with the large basis set on the same level as described in this publication is out of reach at the moment due to the computational demand, but we expect that this would result in a further improvement of the correlation for the carbonyl carbon atoms.

Finally, we were able for the first time to get a reasonable correlation of ^{15}N chemical shifts in a protein (r^2 = 0.8117) when using explicit solvent molecules and averaging. Cai et al.[24] also investigated ^{15}N chemical shifts and showed the importance of augmenting implicit solvent models with explicit first-solvent-shell water molecules for NMA and that the implicit solvent model could serve as a reasonable approximation for the effect of the protein environment for selected fragments of the protein GB3. Thus, we can conclude that explicit solvent molecules are more important than a QM/MM treatment of additional parts of the biomolecule as long as an implicit solvent model replaces the bulk water as well as the neglected parts of the protein. Nevertheless, the mean absolute error is still high. Increasing the basis set size leads into the wrong direction, as shown in our previous publication,[30] and

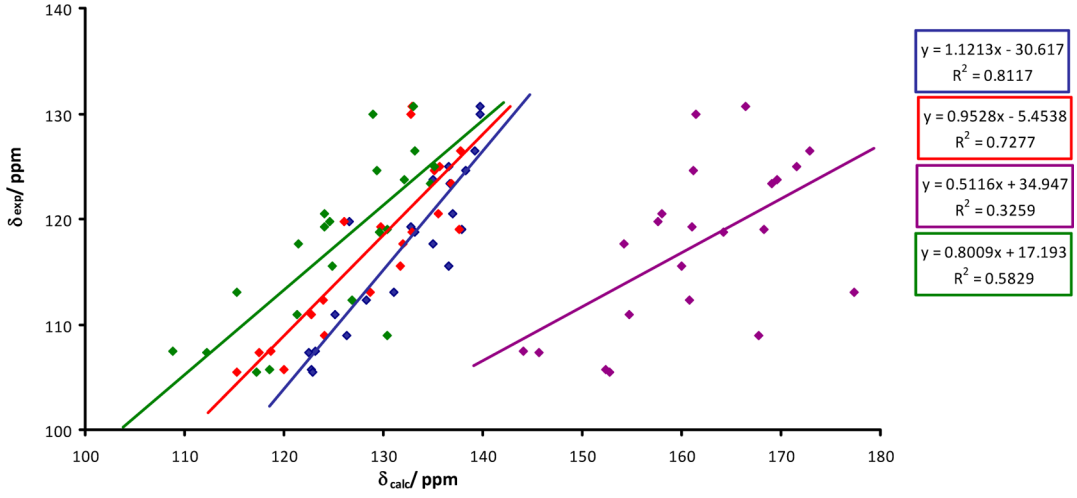

Figure 11. Correlation of ^{15}N NMR chemical shifts of the HA2 domain: blue, B3LYP/6-31g(d), MD, explicit solvents; red, B3LYP/6-31g(d), MD, implicit solvents; magenta, B3LYP/6-311g(d), single point, implicit solvents; green, B3LYP/6-31g(d), single point, implicit solvents.

thus, higher levels of theory are probably needed for quantitative results. Perhaps, the description of hydrogen bonds by the classical molecular dynamics simulation (inadequate as basis for the NMR calculations performed here) also contributes to the errors for nitrogen atoms (and carbonyl carbon atoms). Additionally, it should be kept in mind that the calculations for the standard are still performed *in vacuo* here. This definitely adds to the systematic error, especially for ^{15}N. Using an internal standard is probably the best way to decrease the influence of the standard since it is complicated to include solvent effects accurately and cannot profit from error cancellation due to very different experimental conditions (liquid ammonium compared to protein/peptide in aqueous solution).

In summary, chemical shifts of nonpolar protons and carbon atoms in proteins can be calculated with unseen accuracy in a purely structure-based approach using density functional theory and medium-size basis sets if conformational averaging and solvent effects (first solvent shell by explicit solvent molecules and bulk with a implicit solvent model) are accounted for. Probably the combination of an empirical structure-based and sequence-based approach like SHIFTX2[82] would be able to yield even better results. But this could not be tested fairly since the used peptide is part of the sequence database used for this empirical method. The database-supported prediction of SHIFTX2 is furthermore not applicable to complexes of proteins with nonprotein ligands. Our QM-based approach is not limited to proteins and protein−protein complexes but can handle all other types of molecular complexes without the need for modification. ^{15}N and ^{13}CO shifts are more difficult to handle. One additional reason becoming evident through the use of the new calculations (beside level-of-theory and basis set deficiency) is the high dependency of these nuclei on the actual conformation. On the one hand, the obtained values are more or less random numbers if a single structure is used since only slight deviations can result in very different chemical shifts. Conformational sampling, on the other hand, will be more demanding since a larger ensemble is needed to obtain converged averages. But if the goal is to achieve the same level of accuracy for these nuclei as for aliphatic carbons and nonpolar hydrogens, there is probably no alternative to investing more computer resources to combine all important aspects in one calculation. We hope that this will become possible by linear-scaling methods like the density matrix-based linear-scaling coupled perturbed self-consistent field method[26] and approaches to describe correlation effects at the post-Hartree−Fock level at least in an approximate way as done in concept of intermediate references,[84,85] in which the difference between the absolute shielding between the compound and the reference is calculated in a two step process. First, the difference between the reference and an intermediate molecular system, corresponding to small surroundings in ADMA, is obtained using a high level of theory and a large basis set. The influences of additional parts (larger surroundings) as well as explicit solvent molecules can then be added on a lower level of theory by comparing the intermediate with the full system corresponding to a parent molecule with the anticipated size of surroundings complemented with partial charges for additional parts as well as explicit and implicit solvents.

Last but not least, additional work is on the way to analyze the benefit of the generation of more accurate conformations using *ab initio* MD simulations. But such calculations will limit the method to a relatively small system size, and we hope that new parametrizations of classical force fields with improved hydrogen-bonding geometries will become available as a replacement. Perhaps our chemical shift calculation could even be used in the development of such force fields due to the possibility of verifying conformations very sensitively based on the agreement between calculated and experimental NMR chemical shifts.

■ ASSOCIATED CONTENT

Supporting Information

Calculated and experimental NMR chemical shifts of N-methyl acetamide using different levels of theory and basis sets, distribution of ^{13}C chemical shifts calculated for snapshots taken from the MD simulation of N-methyl acetamide, and correlations of ^{13}C NMR chemical shifts of the HA2 domain are available. This material can be downloaded free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +49-7071-2976969 (T.E.E.), +49-7531-88-5174 (H.M.M.). E-mail: thomas.exner@uni-tuebingen.de (T.E.E.), heiko.moeller@uni-konstanz.de (H.M.M.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work was supported by the German Research Foundation (grant no. EX15/17-1), the Konstanz Research School Chemical Biology (KoRS-CB), the Zukunftskolleg of the Universität Konstanz and the Juniorprofessoren-Programm of the state Baden-Württemberg. I.O. acknowledges the support by the ERASMUS program of the European Union. Additionally, we thank the Common Ulm Stuttgart Server (CUSS) and the Baden-Württemberg grid (bwGRiD), which is part of the D-Grid system, for providing the computer resources making the computations possible.

■ REFERENCES

(1) Tzakos, A. G.; Grace, C. R. R.; Lukavsky, P. J.; Riek, R. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 319−342.
(2) Kay, L. E. *J. Magn. Reson.* **2005**, *173*, 193−207.
(3) Wider, G.; Wüthrich, K. *Curr. Opin. Struct. Biol.* **1999**, *9*, 594−601.
(4) Oldfield, E. *Annu. Rev. Phys. Chem.* **2002**, *53*, 349−378.
(5) Mulder, F. A. A.; Filatov, M. *Chem. Soc. Rev.* **2010**, 39.
(6) Casabianca, L. B.; de Dios, A. C. *J. Chem. Phys.* **2008**, *128*, 052201−052210.
(7) Vila, J. A.; Arnautova, Y. A.; Martin, O. A.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 16972−16977.
(8) Sun, H.; Oldfield, E. *J. Am. Chem. Soc.* **2004**, *126*, 4726−4734.
(9) Vila, J. A.; Aramini, J. M.; Rossi, P.; Kuzin, A.; Su, M.; Seetharaman, J.; Xiao, R.; Tong, L.; Montelino, G. T.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 14389−14394.
(10) Vila, J. A.; Scheraga, H. A. *Proteins* **2008**, *71*, 641−654.
(11) Vila, J. A.; Arnautova, Y. A.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 1891−1896.
(12) Vila, J. A.; Ripoll, D. R.; Scheraga, H. A. *J. Phys. Chem. B* **2007**, *111*, 6577−6585.
(13) Jacob, C. R.; Visscher, L. *J. Chem. Phys.* **2006**, *125*, 194104.
(14) Lee, A. M.; Bettens, R. P. A. *J. Phys. Chem. A* **2007**, *111*, 5111−5115.
(15) Johnson, E. R.; DiLabio, G. A. *J. Mol. Struct.: THEOCHEM* **2009**, *898*, 56−61.

(16) He, X.; Wang, B.; Merz, K. M., Jr. *J. Phys. Chem. B* **2009**, *113*, 10380−10388.
(17) Hori, S.; Yamauchi, K.; Kuroki, S.; Ando, I. *Int. J. Mol. Sci.* **2002**, *3*, 907−913.
(18) Tang, S.; Case, D. *J. Biomol. NMR* **2007**, *38*, 255−266.
(19) Tang, S.; Case, D. *J. Biomol. NMR* **2011**, *51*, 303−312.
(20) Xu, X. P.; Case, D. A. *Biopolymers* **2002**, *65*, 408−423.
(21) Manalo, M. N.; de Dios, A. C. *J. Mol. Struct.: THEOCHEM* **2004**, *675*, 1−8.
(22) Cai, L.; Fushman, D.; Kosov, D. *J. Biomol. NMR* **2009**, *45*, 245−253.
(23) Cai, L.; Kosov, D.; Fushman, D. *J. Biomol. NMR* **2011**, *50*, 19−33.
(24) Cai, L.; Fushman, D.; Kosov, D. *J. Biomol. NMR* **2008**, *41*, 77−88.
(25) Zhu, T.; He, X.; Zhang, J. Z. H. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7837−7845.
(26) Flaig, D.; Beer, M.; Ochsenfeld, C. *J. Chem. Theory Comput.* **2012**, *8*, 2260−2271.
(27) Gao, Q.; Yokojima, S.; Fedorov, D. G.; Kitaura, K.; Sakurai, M.; Nakamura, S. *J. Chem. Theory Comput.* **2010**, *6*, 1428−1444.
(28) Gao, Q.; Yokojima, S.; Kohno, T.; Ishida, T.; Fedorov, D. G.; Kitaura, K.; Fujihira, M.; Nakamura, S. *Chem. Phys. Lett.* **2007**, *445*, 331−339.
(29) Frank, A.; Onila, I.; Möller, H. M.; Exner, T. E. *Proteins* **2011**, *79*, 2189−2202.
(30) Frank, A.; Möller, H. M.; Exner, T. E. *J. Chem. Theory Comput.* **2012**, *8*, 1480−1492.
(31) Ochsenfeld, C.; Kussmann, J.; Koziol, F. *Angew. Chem., Int. Ed.* **2004**, *43*, 4485−4489.
(32) Improta, R.; Barone, V.; Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2006**, *125*, 054103−054109.
(33) Beer, M.; Ochsenfeld, C. *J. Chem. Phys.* **2008**, *128*, 221102−221104.
(34) Beer, M.; Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2011**, *134*, 074102−074115.
(35) Dracinsky, M.; Kaminsky, J.; Bour, P. *J. Phys. Chem. B* **2009**, *113*, 14698−14707.
(36) Dracinsky, M.; Bour, P. *J. Chem. Theory Comput.* **2009**, *6*, 288−299.
(37) Dracinsky, M.; Budesinsky, M.; Warzajtis, B.; Rychlewska, U. *J. Phys. Chem. A* **2011**, *116*, 680−688.
(38) Aidas, K.; Mogelhoj, A.; Kjaer, H.; Nielsen, C. B.; Mikkelsen, K. V.; Ruud, K.; Christiansen, O.; Kongsted, J. *J. Phys. Chem. A* **2007**, *111*, 4199−4210.
(39) Auer, A.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2003**, *118*, 10407.
(40) Prochnow, E.; Auer, A. A. *J. Chem. Phys.* **2010**, *132*, 064109−7.
(41) Woodford, J. N.; Harbison, G. S. *J. Chem. Theory Comput.* **2006**, *2*, 1464−1475.
(42) Eriksen, J. J.; Olsen, J. M.; Aidas, K.; Agren, H.; Mikkelsen, K. V.; Kongsted, J. *J. Comput. Chem.* **2011**, *32*, 2853−2864.
(43) Moon, S.; Case, D. A. *J. Comput. Chem.* **2006**, *27*, 825−836.
(44) Dumez, J. N.; Pickard, C. J. *J. Chem. Phys.* **2009**, *130*, 104701−104708.
(45) Mezey, P. G. *Adv. Quantum Chem.* **1996**, *27*, 163−222.
(46) Mezey, P. G. *Internat. Rev. Phys. Chem.* **1997**, *16*, 361−388.
(47) Mezey, P. G. *J. Math. Chem.* **1995**, *18*, 141−168.
(48) Exner, T. E.; Mezey, P. G. *J. Phys. Chem. A* **2002**, *106*, 11791−11800.
(49) Exner, T. E.; Mezey, P. G. *J. Phys. Chem. A* **2004**, *108*, 4301−4309.
(50) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Wenger, R. K.; Yao, H.; Markley, J. L. *Nucleic Acids Res.* **2007**, *36*, D402−D408.
(51) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*, University of California: San Francisco, CA, 2008.
(52) Berman, H. M.; Westbrook, J.; Feng, Z.; Filliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(53) Lorieau, J. L.; Louis, J. M.; Bax, A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 11341−11346.
(54) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(55) Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209−214.
(56) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213−222.
(57) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654−3665.
(58) Blaudeau, J. P.; McGrath, M. P.; Curtiss, L. A.; Radom, L. *J. Chem. Phys.* **1997**, *107*, 5016−5021.
(59) Binning, R. C., Jr.; Curtiss, L. A. *J. Comput. Chem.* **1990**, *11*, 1206−1216.
(60) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. *J. Comput. Chem.* **2001**, *22*, 976−984.
(61) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. *J. Chem. Phys.* **1998**, *109*, 1223−1229.
(62) Gordon, M. S. *Chem. Phys. Lett.* **1980**, *76*, 163−168.
(63) Cornell, W. D.; Cieplak, P.; Bayby, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(64) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(65) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(66) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(67) London, F. *J. Phys. Radium* **1937**, *8*, 397−409.
(68) McWeeny, R. *Phys. Rev.* **1962**, *126*, 1028−1034.
(69) Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789−807.
(70) Wolinski, K.; Hilton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.
(71) Hwang, T. L.; Shaka, A. J. *J. Magn. Reson. A* **1995**, *112*, 275−279.
(72) Bax, A.; Subramanian, S. *J. Magn. Reson.* **1986**, *67*, 565−569.
(73) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135−140.
(74) Sefzik, T. H.; Turco, D.; Iuliucci, R. J.; Facelli, J. C. *J. Phys. Chem. A* **2005**, *109*, 1180−1187.
(75) Kupka, T.; Stachow, M.; Nieradka, M.; Kaminsky, J.; Pluta, T. *J. Chem. Theory Comput.* **2010**, *6*, 1580−1589.
(76) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996**, *104*, 5497−5509.
(77) van Mourik, T. *J. Chem. Phys.* **2006**, *125*, 191101−191104.
(78) Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032−3041.
(79) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151−5158.
(80) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 253−260.
(81) Grzesiek, S.; Cordier, F.; Jaravine, V.; Barfield, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2004**, *45*, 275−300.
(82) Han, B.; Liu, Y.; Ginzinger, S.; Wishart, D. *J. Biomol. NMR* **2011**, *50*, 43−57.
(83) Neal, S.; Nip, A. M.; Zhang, H.; Wishart, D. S. *J. Biomol. NMR* **2003**, *26*, 215−240.
(84) Ochsenfeld, C.; Koziol, F.; Brown, S. P.; Schaller, T.; Seelbach, U. P.; Klärner, F. G. *Solid State Nucl. Magn. Reson.* **2002**, *22*, 128−153.
(85) Zienau, J.; Kussmann, J.; Ochsenfeld, C. *Mol. Phys.* **2010**, *108*, 333−342.