JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

# Conformational Sampling by Ab Initio Molecular Dynamics Simulations Improves NMR Chemical Shift Predictions

Martin Dračínský,*[,†,‡] Heiko M. Möller,[§] and Thomas E. Exner*[,§,∥]

[†]Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Flemingovo náměstí 2, 166 10 Prague, Czech Republic

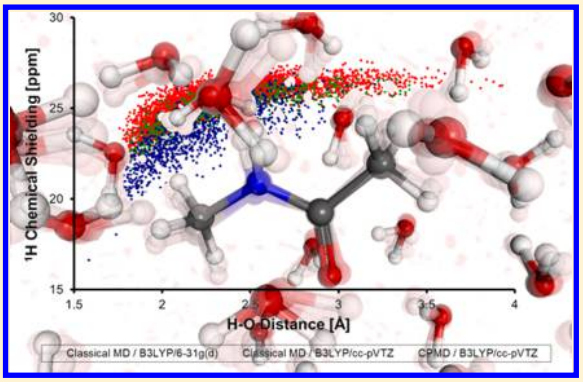
[‡]Department of Chemistry, Durham University, DH1 3LE Durham, United Kingdom

[§]Department of Chemistry, 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:** Car−Parrinello molecular dynamics simulations were performed for *N*-methyl acetamide as a small test system for amide groups in protein backbones, and NMR chemical shifts were calculated based on the generated ensemble. If conformational sampling and explicit solvent molecules are taken into account, excellent agreement between the calculated and experimental chemical shifts is obtained. These results represent a landmark improvement over calculations based on classical molecular dynamics (MD) simulations especially for amide protons, which are predicted too high-field shifted based on the latter ensembles. We were able to show that the better results are caused by the solute−solvents interactions forming shorter hydrogen bonds as well as by the internal degrees of freedom of the solute. Inspired by these results, we propose our approach as a new tool for the validation of force fields due to its power of identifying the structural reasons for discrepancies between the experimental and calculated data.




## INTRODUCTION

The theoretical description of the biochemical world highly relies on biomolecular force fields like AMBER,[1−3] CHARMM,[4,5] GROMACS,[6] OPLS,[7] and others, used e.g. in molecular dynamics simulations, Monte Carlo, and docking applications. Despite all the success of these force fields, however, they are still under continuous development to include such additional effects as polarization or to be able to describe less common secondary- and tertiary-structure elements.[8−12] One specific example is the insufficiency of all current modifications of the AMBER force field to describe unusual DNA/RNA structures like quadruplexes.[11] Further development is hampered by the fact that the beneficial or detrimental effect of a modification is hard to evaluate. The usefulness of the data obtained from NMR spectroscopy in such evaluations has been emphasized in the recent literature. Beauchamp et al.[13] have answered the question of whether protein force fields are becoming better by comparing 524 diverse NMR measurements with back-calculated data from ensembles generated by different state-of-the-art force fields using the empirical program SPARTA+.[14] Unfortunately, it is not clear from the publication how many structures are considered to generate the ensemble average of the NMR chemical shifts and *J* couplings. A similar study has been performed by Aliev and Courtier-Murias with small open-chain peptides as test systems.[15] Yildirim et al.[16] have been able to improve the agreement with the NMR spectra by benchmarking AMBER force fields for RNA and revising $\chi$ torsion parameters. Other groups have used NMR chemical shifts to assess and refine molecular dynamics simulations or a combination of NMR calculations and MD to characterize the dynamic behavior of proteins or nucleic acids.[17−20] All of these studies utilize empirical methods such as SHIFTX2[21] and SPARTA+[14] for the back-calculation of the chemical shifts. Besides their much higher efficiency as compared to quantum-chemical calculations, this is justified by the claim that the quantum mechanical methods available for such large systems are not accurate enough.[17] Nevertheless, recent studies[22−52] have clearly demonstrated that the main errors are not associated with the levels of theory and the basis sets used but arise from the neglect of conformational averaging and first-solvent-shell effects. In our previous paper,[24] we were able to show that NMR chemical shifts of nonpolar protons and carbon atoms in proteins can be calculated with unprecedented accuracy in a purely structure-based approach if multiple structures are taken from molecular dynamics simulations and explicit solvent molecules are considered. The larger errors in

**Received:** April 5, 2013
**Published:** June 18, 2013

polar hydrogens can be attributed to the deficiencies of unpolarized force fields in the description of hydrogen bonding to the solvent. This has only recently been supported by a study of Zhu et al.,[26] who showed that a polarizable charge model can improve the accuracy of the calculated chemical shifts but that water distribution based on a 3D reference interaction site model (3D-RISM) provides even better results.[26]

Taking all these studies together, it seems possible to obtain highly accurate NMR chemical shifts from quantum-chemical calculations based on an ensemble of structures from the best possible molecular dynamics simulation. To further prove this, we will present here calculations on *N*-methyl acetamide (see Figure 1) as a test system for peptide bonds. Classical MD

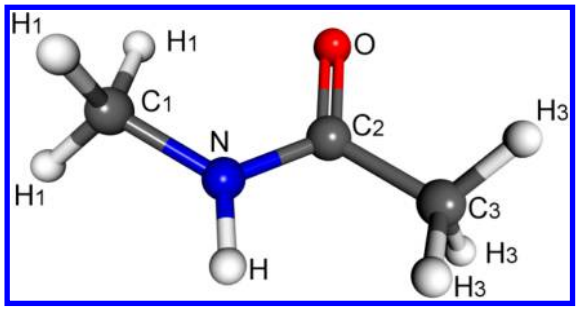



**Figure 1.** The numbering scheme of N-methyl acetamide.

simulations with the General Amber Force Field (GAFF)[53] will be compared to Car–Parrinello MD (CPMD) simulations. Accurate chemical shifts have already been obtained for single amino acids and some other organic molecules using CPMD in our previous papers,[54,55] and the interplay between the CPMD and NMR chemical shifts has been analyzed in detail for liquid water in the group of Sebastiani.[56] It is clear that CPMD cannot be used to simulate large biomolecular systems on a regular basis. However, the clear superiority of the ensemble generated by CPMD over the one generated by classical simulations based on the general AMBER force field (GAFF) demonstrated by the results shown here indicates the improvements obtainable by an NMR chemical shift-guided force-field optimization.

## MATERIALS AND METHODS

**Car–Parrinello Molecular Dynamics Simulations.** A periodic cubic box of 10.062 Å containing one *N*-methyl acetamide and 30 water molecules was created by the HYPERCHEM program.[57] To equilibrate the systems, a classical MD run was performed for 100 ps with 1 fs integration time steps at a temperature of 300 K and constant volume also using HYPERCHEM (the volume had been chosen to have the appropriate density of 1 g/mL). The TIP3P force-field parameters[58] as part of Amber99 were used for water. The geometry was then optimized and transferred to the CPMD software package.[59] The same periodic boundary conditions and 4 au (0.09676 fs) time steps were maintained for all CPMD calculations performed with the BLYP functional[60] and the Vanderbilt ultrasoft pseudopotentials.[61] An energy cutoff of 25 Ry was used. The initial configuration was relaxed by six short CPMD runs comprising 200 steps. After each run, the system was quenched to the Born–Oppenheimer surface by reoptimizing the wave function. A longer 106 ps production run was then performed with a temperature of 300 K maintained with the Nosé-Hoover algorithm,[62] which also kept the system in the canonical (NVT) ensemble. During the long simulations, the trajectory was saved at every 100th step.

As just described, the CPMD was performed with the BLYP GGA functional for computational convenience although the hybrid B3LYP is later used for the NMR calculations. The CPMD simulation would have been about 40 times more expensive if performed with the B3LYP functional.[63] Additionally, it is generally accepted that the performance of BLYP and B3LYP is comparable in terms of geometry.[64]

**Quantum-Chemical NMR Chemical-Shift Calculations.** As we are interested in the chemical shifts of the thermodynamic ensemble of structures and not of the minimum structures, the molecular coordinates were taken as provided by the MD simulation without further quantum-mechanical optimization. The NMR chemical-shift calculations were performed using the fragmentation scheme of the adjustable density matrix assembler (ADMA) as described in our previous publications.[22−24] This scheme was designed to cut small fragments out of a large (bio)-macromolecule, on which standard quantum-chemical calculations can be performed. Even if this publication analyzes only the small molecule N-methyl acetamide, the fragmentation can still be used to determine the solvent molecules which should be included in the NMR calculations. This is done by calculating the distance from each solute atom to each atom in a water molecule. If the distance is shorter than the distance cutoff defined for ADMA, set to 4 Å in all of the calculations presented here, the water molecule is included taking also the periodicity of the system into account. The influence of additional solvents was modeled by embedding the supermolecule of the solute and all explicit water molecules, called parent molecules in the following, within an implicit solvent model (IEF-PCM[65−67]). The NMR isotropic chemical shieldings are calculated with the Gaussian09 program package[68] using the B3LYP[60] hybrid functional and the cc-pVTZ basis set[69] within the GIAO (gauge invariant/including atomic orbitals) formalism.[70−72] Most of our discussion will be based on isotropic chemical shieldings. For a direct comparison with the experiment, however, the chemical shieldings have to be converted to chemical shifts by $\boldsymbol{\delta} = \boldsymbol{\sigma}_{standard} - \boldsymbol{\sigma}$ using the corresponding standards. We use the averaged sum of the experimental chemical shift and the calculated chemical shielding as the chemical shielding of the standards ($\boldsymbol{\sigma}_{standard} = ((\sum \boldsymbol{\delta}_{exp} - \boldsymbol{\sigma}_{calc})/n)$ with $n$ as the number of nuclei of a specific type like $^{13}$C or $^{1}$H). Alternatively, the chemical shieldings of the standards (tetramethylsilane for $^{1}$H and $^{13}$C and neat ammonia for $^{15}$N) are calculated at the same level of theory, and the NMR chemical shifts are obtained by subtracting the solute chemical shieldings from these, which is the standard procedure for chemical-shift calculations but also measurements. Nevertheless, please keep in mind that the calculations of the standards used in this publication still neglect solvent effects and flexibility and thus are expected to be error-prone. The third method for calculating chemical shifts is by a linear fit of the experimental chemical shifts versus the calculated shielding values to obtain the shielding values of the corresponding standards, but this is only appropriate if the correlation has a slope of close to 1. Since, as shown below, the results of the classical ensemble differ considerably from this optimal slope, this procedure would artificially favor the ab initio simulation.

## RESULTS AND DISCUSSION

**Ensemble Influences on NMR Chemical Shifts in *N*-Methyl Acetamide.** To analyze the influence of thermodynamic ensembles on the calculated chemical shifts, a representative number of time steps were extracted from the CPMD as well as a previously described classical MD simulation[23] of *N*-methyl acetamide. To limit the computational demand, every 1000th step from the CPMD was taken, resulting in 1104 independent calculations. 1000 chemical-shift calculations were performed with the cc-pVTZ basis set[69] on snapshots taken from the first 2 ns of the classical MD simulation. For comparison, we also include the results (5000 calculations) using the 6-31g(d) basis set[73] taken from our previous publication.[24] Since we could show that 500 and 5000 snapshots taken from 1 and 10 ns simulation time, respectively, resulted in very comparable distributions and almost identical average values,[23] the number of snapshots chosen here should lead to a representative description of conformational averaging and solvent effects. First, we will examine the correlations between the isotropic chemical shieldings and the experimental chemical shifts (see Figure 2 for $^{1}$H chemical shifts and the Supporting Information for $^{13}$C chemical shifts; since the molecule has only one nitrogen atom, no results for $^{15}$N can be shown). Due to the small size of the molecule, the results should not be overinterpreted, and the correlation coefficients are not conclusive. High $r^2$ values are seen for all approaches as expected. Another property of the correlation lines − the slope − can however be analyzed (as long as one keeps in mind the statistical uncertainties). For carbon atoms, these show no clear preference for classical MD or CPMD (see the Supporting Information). To obtain a slope to the optimal value of −1, it is more important to pass from the smaller to the larger basis set than to switch from classical to ab initio MD. This can be justified by the carbonyl carbon chemical shifts, which are predicted as high-field shifted with the small basis set as shown in our previous publications.[22,23] The proton results provide a completely different picture (see Figure 2). The CPMD ensemble has an almost perfect slope of −0.99. For the classical MD, the slope is much smaller, which is caused by the wrong description of the hydrogen bonds and the resulting high-field chemical shielding of the amide proton as already seen in our previous publication.[23]

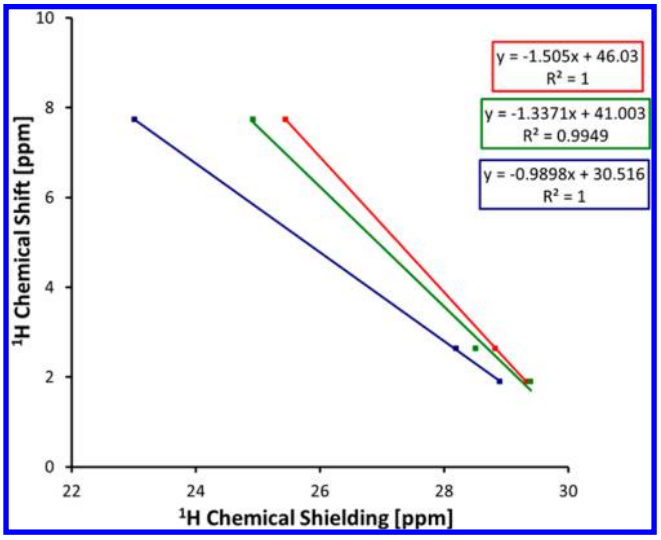



**Figure 2.** The correlation between the experimental $^{1}$H NMR chemical shifts and the calculated isotropic shieldings of the N-methyl acetamide (NMA): red: classical MD/B3LYP/6-31g(d); green: classical MD/B3LYP/cc-pVTZ; and blue: CPMD/B3LYP/cc-pVTZ.

As described in the method sections, there are multiple options to arrive from the isotropic chemical shieldings at the chemical shifts. The most common approach is to calculate the isotropic chemical shielding of the standards (tetramethylsilane for $^{1}$H and $^{13}$C and ammonia for $^{15}$N) with the same level of theory and basis set and subtract the values of the molecule of interest from these. The experimental chemical shifts and the corresponding calculated values averaged over the thermodynamic ensembles taken from the classical as well as CPMD are summarized in Table 1. At first glance, it seems that going from classical MD to ab initio MD worsens the agreement because of the shift to a lower field seen in the CPMD ensemble. This is in contrast to the just-discussed improvement in the correlation. All deviations from the experimental values are larger for the CPMD ensemble than for the classical ensemble using the same larger basis set. This is also true for the amide proton even if the deviations are comparable and in the opposite direction. However, one has to keep in mind that, for N-methyl acetamide, the ensemble averages with explicit solvent molecules are taken, but the chemical shielding value of the standard is still based on a single-point calculation. Therefore, the largest error is expected to be associated with the calculation of the standard. In principle, the standards should also be determined while taking conformational sampling and solvent effects into account. Even with such a treatment, however, we might expect large errors, arising from the different quality of the description of such standard systems as molecules/proteins in water and very special systems like liquid ammonia.

The dependence on an accurate value for the standards can be circumvented by looking at relative values. Using the offset of the correlation line as the chemical shielding of the corresponding standard is only appropriate for slopes close to −1 and would thus strongly disfavor the classical simulations. Defining the chemical shielding of the standards as the average sum of the experimental chemical shifts and the calculated chemical shieldings of the target molecule is, in our opinion, a method not artificially favoring one of the ensembles even if the information on systematic errors is lost. The resulting values are given in Table 2. With a standard for $^{1}$H of 30.79 ppm (which

**Table 1. Experimental and Averaged Calculated NMR Chemical Shifts of N-Methyl Acetamide (NMA)[a]**

| | $C_1$ | $H_1$ | N | H | CO | $C_3$ | $H_3$ |
|---|---|---|---|---|---|---|---|
| experiment | 25.80 | 2.63 | 113.80 | 7.73 | 174.50 | 21.50 | 1.90 |
| CPMD/cc-pVTZ | 35.84 | 3.55 | 159.99 | 8.72 | 191.80 | 30.50 | 2.85 |
| classical MD/cc-pVTZ | 33.04 | 3.23 | 139.58 | 6.82 | 177.36 | 26.76 | 2.35 |
| classical MD/6-31g(d) | 30.17 | 2.93 | 116.90 | 6.31 | 161.43 | 23.78 | 2.43 |

[a]The calculated chemical shifts are obtained by building the difference of the isotropic-chemical shieldings of the standard and the target nucleus.

Table 2. Experimental and Averaged Calculated NMR Chemical Shifts of N-Methyl Acetamide (NMA)[a]

| | $C_1$ | $H_1$ | N | H | CO | $C_3$ | $H_3$ |
|---|---|---|---|---|---|---|---|
| experiment | 25.80 | 2.63 | 113.80 | 7.73 | 174.50 | 21.50 | 1.90 |
| CPMD/cc-pVTZ | 23.73 | 2.60 | n.a. | 7.77 | 179.69 | 18.39 | 1.90 |
| classical MD/cc-pVTZ | 27.92 | 3.18 | n.a. | 6.77 | 172.24 | 21.64 | 2.30 |
| classical MD/6-31g(d) | 32.31 | 3.13 | n.a. | 6.51 | 163.57 | 25.92 | 2.63 |

[a]The calculated chemical shifts are obtained using the averaged sum of the experimental chemical shift and the isotropic chemical shielding of the target nucleus as the standard. This approach is applicable only if a larger number of nuclei can be averaged; therefore, the values for 15N are not given here.

corrects the value of the standard calculated with B3LYP/cc-pVTZ by −0.95 ppm), almost perfect agreement between the CPMD ensemble calculation and the experiment can be obtained. Additional calculations on a larger set of molecules are needed to show whether this offset is of general nature. For the two aliphatic protons in NMA, the low-field shift caused by the conformational sampling and explicit solvent effects, when comparing the single-point vacuum calculation and the CPMD-ensemble averages, is very systematic and of the same order of magnitude as the correction needed for the standard.

**Structural Reasons for Ensemble Influences.** The results presented above show a clear dependence on the ensemble used and a clear preference for the ab initio CPMD simulation. To identify the reasons for these different behaviors, the distributions of the chemical shieldings depending on the conformations adopted in the MD simulations will be analyzed. The main motivation for the use of the ab initio molecular dynamics simulation were the extremely low values of the chemical shift of the amide proton in classical MD-generated ensembles. Even if explicit water molecules improved the results considerably, there was still a disagreement of almost 1 ppm. We argued in our previous paper[24] that this high-field shift of the classical MD-derived data was caused by an inaccurate description of the hydrogen bonds to the solvent molecules, which had been predicted to be too long and too weak by classical MD. CPMD simulations were already able to correct for this phenomenon.[54,55] Our new calculations also show the desired effect. It can be clearly seen in Figure 3 that the distribution of amide-proton shieldings is only slightly shifted to the lower field (which translates to higher chemical shifts) if the larger basis set is used but that with the structures from the CPMD an almost 2 ppm shift is obtained. As a short note, we would like to mention that while the distribution of the other nuclei can be reasonably modeled by Gaussian distributions, a LogNormal model gives much better agreement for the amide proton due to the long tail toward small chemical shieldings with values reaching 16 ppm.

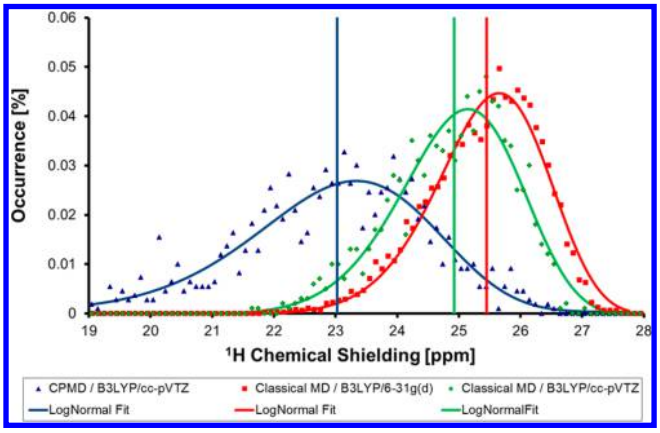



Figure 3. The probability distribution of the $^1$H chemical shielding of the amide proton in N-methyl acetamide (NMA). The vertical lines indicate the average shielding values.

The influence of the hydrogen-bond distance on the amide-hydrogen chemical shielding can be seen in Figure 4, where the

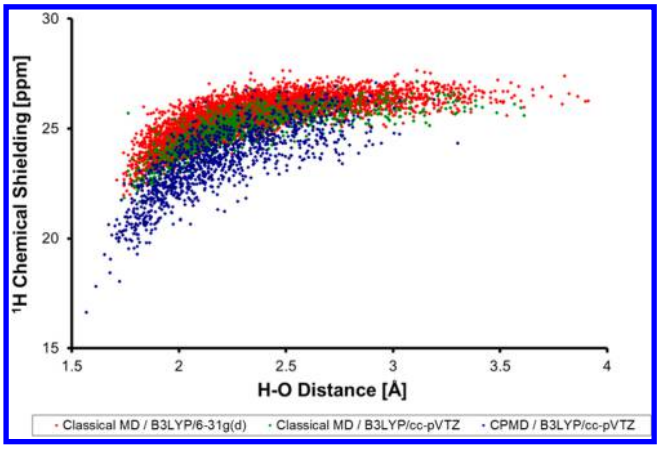



Figure 4. The dependence of the $^1$H chemical shielding of the amide proton in NMA on the distance to the closest water-oxygen atom.

chemical shielding is plotted against the distance of the amide proton to the closest water oxygen atom. First, while the shortest distance in the classical MD is around 1.8 Å, it is around 1.5 Å in the CPMD. Second, there are many snapshot geometries with distances between 3 and 4 Å in the classical MD and almost none in the CPMD. These findings are also evident if the radial distribution functions of the water oxygen (Figure 5) atoms around the amide proton are compared. The shorter distances (the CPMD maximum is shifted by 0.1 Å) and the much higher peak of the first solvent shell are indicative of a stronger and more ordered hydrogen bond. The higher ordering of the hydrogen-bonding network also manifests itself in the prominent peak for the second solvent shell in the CPMD, which is absent in the classical simulation. One could argue that the radial-distribution functions differ so strongly between the ab initio and classical MD due to the suboptimal force-field parameters of NMA. Nevertheless, very similar distributions without a peak for the second solvent shell are obtained for the simulation of a small polypeptide[23] using the highly optimized AMBER force field (see also Figure 5). Since deviations from the experimental values of more than 1 ppm are also seen for the solvent-exposed amide protons in this peptide, the results presented for NMA can, in our opinion, be directly applied to polypeptide calculations, and the claim of our last paper, namely that the hydrogen bonds predicted by classical force fields are too long and too weak, is supported. The amide proton forms a strong hydrogen bond to water


dx.doi.org/10.1021/ct400282h | J. Chem. Theory Comput. 2013, 9, 3806−3815

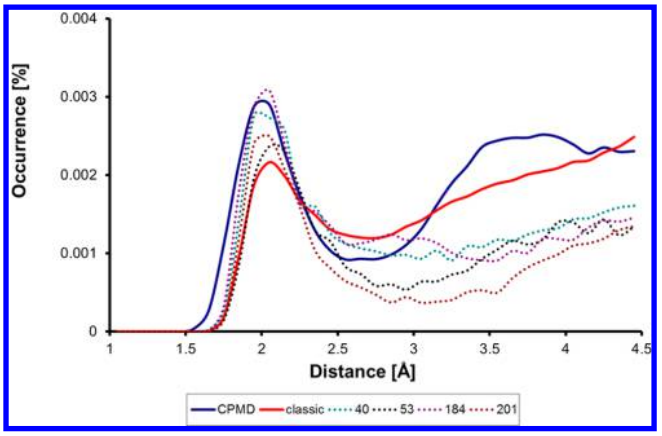



**Figure 5.** The radial distribution functions of the water-oxygen atoms around the amide proton of NMA from the CPMD and classical simulations and of the HA2 domain of the influenza virus glycoprotein hemagglutinin presented in ref 23. The numbers in the legend correspond to the atom numbers in the Protein Data Bank (PDB)[74] file of the HA2 domain (pdb entry 2KXA[75]).

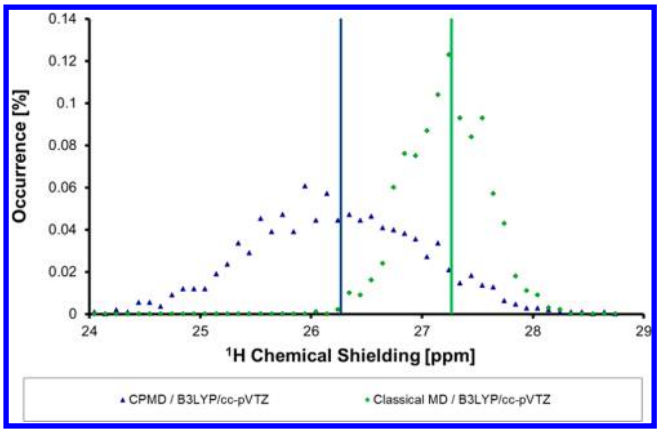



**Figure 6.** The probability distribution of the $^1$H chemical shielding of the amide proton in N-methyl acetamide (NMA) using the implicit solvent model alone.

molecules most of the time of the CPMD simulation. For some of the amide protons in the peptide (atom 40 and 184 of HA2), the first solvent shell peak is as high as in the CPMD simulation of NMA, but even for such strong hydrogen bonds the shortest distance is the same as for the weaker ones (atoms 53 and 201 of HA2 as well as NMA). Additionally, in the classical MD this hydrogen bond is often replaced by weaker interactions with two water molecules.

Besides this shift caused by shorter hydrogen-bond distances, however, the chemical shieldings are on average also low-field shifted in the CPMD as compared to the classical MD snapshots if one considers values corresponding to similar hydrogen-bond lengths. The reason for this could be different orientations of the coordinating water molecule (e.g., the donor-hydrogen-acceptor angle) or could result from the different conformations in NMA. To test the second assumption, we calculated again the chemical shieldings for each snapshot with the larger basis set, but this time using only the implicit solvent model (IEF-PCM[65−67]) without including any explicit water molecules. In comparison with the calculations with explicit water molecules, the amide-hydrogen shielding values of the isolated NMA are shifted by ca. 3.2 ppm and 2.3 ppm upfield (see Table 3) for the CPMD and classical MD, respectively, and the long tail toward small chemical shieldings caused by short hydrogen bonds is missing (see Figure 6). This corresponds to the influence of the solvents just described. Although the average amide-hydrogen chemical shieldings move closer to each other when explicit water molecules are neglected, there is still a 1 ppm difference between the two ensembles. Therefore, there have to be also intramolecular reasons for the distinct features of the two distributions clearly visible in Figure 4.

With the linear-regression-solver-node and backward-feature elimination within the KNIME workflow managing system,[76] it was possible to identify the important structural features (the distances between the atom pairs provided the input) describing the amide-proton shielding distributions in the isolated NMA. As the main feature, the distance between the amide proton and nitrogen was identified, and an already reasonable correlation ($r^2$ = 0.6985) was obtained with this criterion alone (see Figure 7). It also explains why the chemical

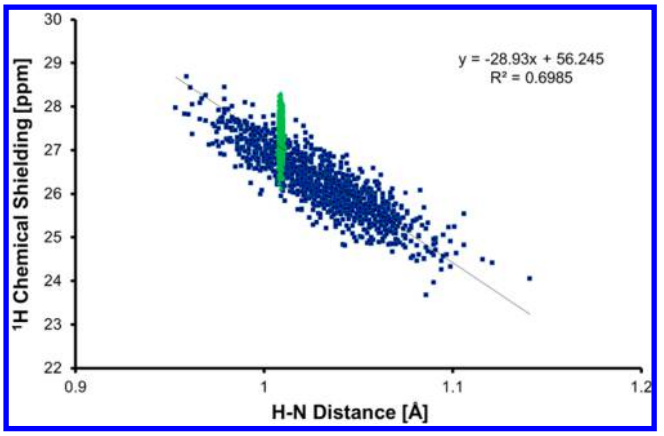



**Figure 7.** The correlation between the N−H distance and the calculated amide proton shieldings using implicit solvents. The values based on the CPMD and the classical ensembles are color-coded in blue and green, respectively.

shieldings of all protons show a much narrower distribution in the classical MD as compared to CPMD. The SHAKE algorithm[77] was used to constrain the distances of all bonds with attached protons to allow for a larger integration time step. In the case of the amide group, this limits the conformational space to arrangements with N−H distances between 1.008 and

**Table 3. Average NMR Chemical Shieldings of N-Methyl Acetamide (NMA) from Classical MD and CPMD Calculated with Explicit Hydration**[a]

| | $C_1$ | $H_1$ | N | H | CO | $C_3$ | $H_3$ |
|---|---|---|---|---|---|---|---|
| CPMD/cc-pVTZ | 148.64 (2.83) | 28.20 (0.26) | 106.04 (27.56) | 23.02 (3.25) | −7.32 (7.23) | 153.98 (4.75) | 28.90 (0.58) |
| classical MD/cc-pVTZ | 151.44 (3.45) | 28.51 (0.26) | 126.45 (21.43) | 24.92 (2.34) | 7.12 (6.81) | 157.72 (4.41) | 29.33 (0.53) |

[a]The differences between implicit and explicit hydration are in parentheses.

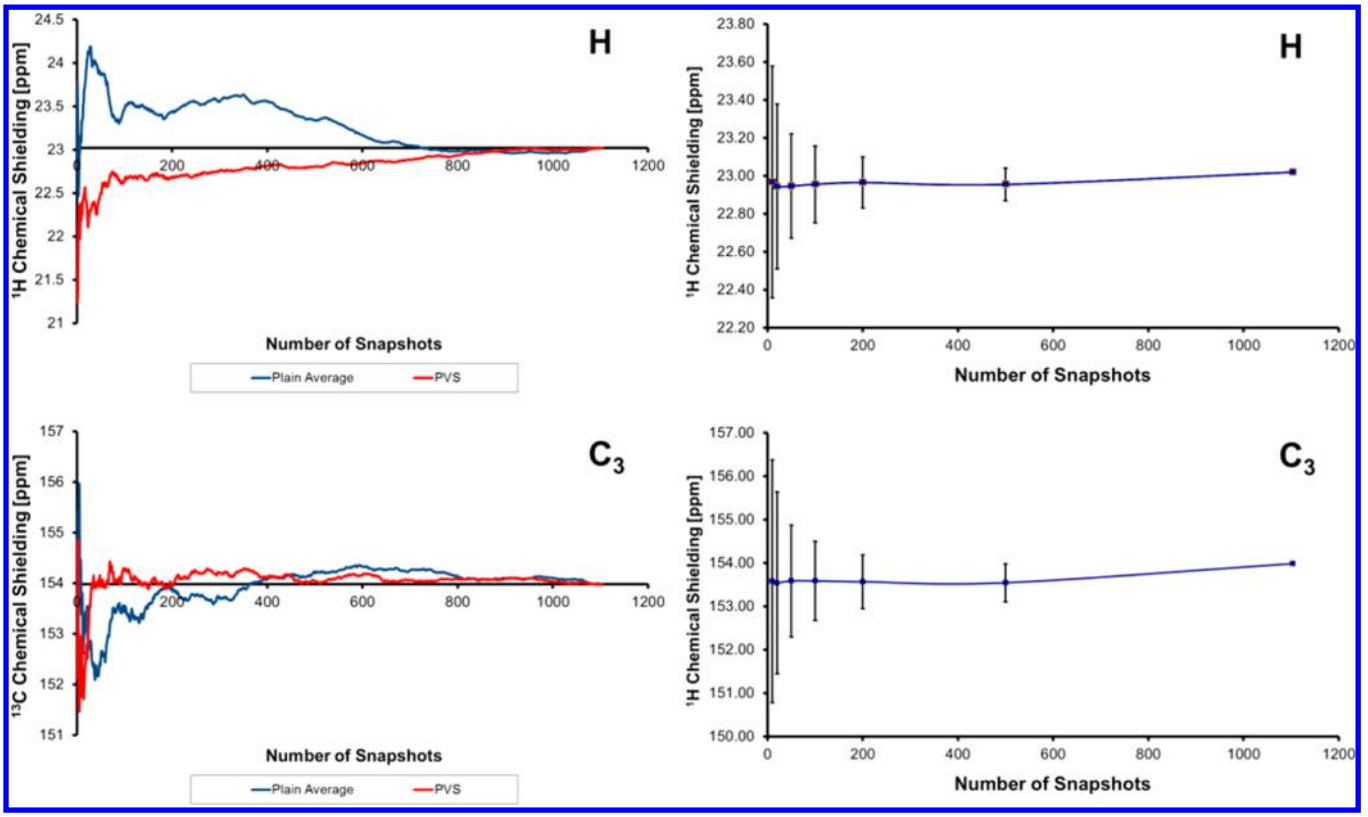



**Figure 8.** Convergence behavior of the NMR chemical shifts of the amide proton and $C_3$ with the number of used structures using three different analyses: (left) convergence when the snapshots are added in sequential order or if they are chosen according to the Parallel Variable Selection (PVS) method, (right) averages and standard deviations of 10000 independent random samples.

1.010 Å (the variance is probably caused by the limited precision of the atomic coordinates in the pdb files used to convert the AMBER output to the GAUSSIAN input), and the extreme chemical shielding of 28 and 25 ppm caused by bond lengths of 0.95 and 1.10 Å can only be seen in the CPMD ensemble. If three distances, namely N−H, $C_2$−O, and $C_2$−H (which could also be modeled by the $C_2$−N−H angle), are correlated with the calculated chemical shielding, the prediction can be improved to an $r^2$ = 0.8735 (see the Supporting Information). The inclusion of the $C_2$−O and $C_2$−H distances now makes it possible to describe the spread in the chemical shieldings based on the classical ensemble. One practical result from this analysis is that SHAKE should probably be deactivated when generating ensembles for NMR chemical-shift calculations. In this paper and our previous paper,[23] we tried to adhere to the standard MD protocols as closely as possible, and that covalent bonds to hydrogens are constrained (like when using SHAKE) can be considered standard nowadays. Still, a perfect agreement between CPMD and non-SHAKE classical MD is not expected, because the second feature of the correlation is the $C_2$−O distance whose distributions show large differences between the two ensembles (see the Supporting Information).

For all other nuclei, the distributions of the calculated isotropic chemical shieldings using the three different theoretical approaches (CPMD/B3LYP/cc-pVTZ, classical MD/B3LYP/cc-pVTZ and classical MD/B3LYP/6-31g(d)) show comparable behaviors if explicit solvent molecules are considered (see the Supporting Information). The values obtained with the smaller basis set based on the structures from the classical MD simulation are shifted to the higher field, the classical MD and triple-ζ basis set combination is in the middle, and the CPMD results are in the lowest field. Additionally, the distribution of the chemical shielding is broader for CPMD than for the classical MD structures especially for protons, which can probably again be attributed to the usage of the SHAKE algorithm.

To separate the influence of the solvent from the intraligand differences between the two ensembles (CPMD and classical MD) with the larger basis set, the average chemical shieldings and the changes when moving from the mixed explicit/implicit solvent model (the closest water molecules are explicitly included) to the entirely implicit solvent model are summarized in Table 3. When the explicit water molecules are removed, all nuclei show a significant high-field shift, which is of the same order of magnitude in the CPMD and classical MD ensembles. The already-discussed amide hydrogen as well as the amide nitrogen show the most significant dependence on the simulation method. The extremely large shift for the nitrogen atom, specifically of more than 20 ppm, and the difference of 6 ppm between ab initio and classic MD can be attributed to hydrogen bonding and the differences in the intermolecular H−O distances. When looking at the calculations with the implicit solvent model alone, the relevant intramolecular features can be analyzed as already done for the amide hydrogen above. The two nuclei showing the most distinct distributions and chemical shieldings differing by more than 14 ppm in the classical and CPMD ensembles are the carbonyl carbon and the amide nitrogen (see Table 3 and the Supporting Information). Since we have removed the solvent for these calculations, these differences are caused by internal degrees of freedom and not by solute−solvent interactions. We would like to emphasize

here again that the differences in the solute−solvent interaction only account for 6 and less than 0.5 ppm for nitrogen and carbonyl carbon, respectively, when comparing the two ensembles (the difference of the differences between the explicit and implicit solvents, i.e. the difference of the numbers in parentheses in Table 3). KNIME[76] was again used to correlate the distances with the calculated chemical shieldings using the implicit solvent model. For the carbonyl carbon atom, KNIME identifies the $C_2$−O distance as the most important factor, and a correlation with $r^2 = 0.8703$ can be obtained with this distance alone. The multilinear fit including also the $C_2$−N and $C_2$−H distances as criteria provides an almost perfect fit (see also the Supporting Information). The chemical shift of the amide nitrogen is much harder to reproduce using correlations to atomic distances. The best two distance criteria are the $C_2$−O and the N−H distances, but they only give correlations with $r^2 = 0.5010$ and $r^2 = 0.3143$. Five distance criteria are needed to obtain a reasonable correlation coefficient ($r^2 = 0.8152$). It should also be highlighted that, besides the local surroundings ($C_2$−O, N−H, and $C_2$−N distances), nonbonding distances ($C_1$−O and N−$H_3$) are used in the correlation. On the one hand, this shows that the differences between the two ensembles can clearly be described by the two distances showing the largest variations in the two ensembles ($C_2$−O and N−H, see the Supporting Information); on the other hand, however, it indicates that $^{15}$N chemical shieldings are influenced by larger parts of the molecular system, and changes in distant regions also have to be sampled for an accurate prediction of these nuclei.

**Convergence Behavior of NMR Chemical Shifts.** Performing thousands of NMR chemical shift calculations is very time-consuming and can only be performed for a small number of molecules (even if one neglects the demand for the MD simulation). Therefore, it is of interest to analyze the convergence behavior of the chemical shifts to find the optimal number of calculations as a compromise of accuracy and efficiency. This was done using three different approaches. First, the snapshots are added one after the other in sequential order, and the averages values are built using all selected snapshots. This is called plain average in the following. Second, we picked 10000 samples each consisting of 10 random snapshots from the 1104 total snapshots of the CPMD simulation. For each of these random samples an average value of the NMR chemical shift is calculated. To see the convergence, the average of these averages and the standard deviation is then calculated. This procedure is then repeated for 20, 50, 100, and 200 as well as 500 snapshots in the random samples. Third, we used the Parallel Variable Selection (PVS) method recently introduced for a faster convergence of a MD-based calculation of spectroscopic properties by a preselection of clusters.[78] The results of these analysis are shown in Figure 8 for the amid proton and $C_3$ (similar figures for the amid nitrogen and the carbonyl carbon are given in the Supporting Information, averages of averages and standard deviations of all nuclei can also be found in the Supporting Information). The plain average and the standard deviation clearly show that a large number of snapshots has to be taken into account. For the best performing example - $C_3$, the plain average is reasonably converged after the first 200 snapshots although some significant fluctuations of the calculated chemical shift can be seen after this point. For the amide proton, the predicted chemical shift converges only after 600 snapshots. For the 10000 random samples, significant standard deviations are obtained for all nuclei up to at least 200 snapshots. The PVS method, which uses water density distribution as a secondary variable to preselect the most probable cluster geometries, performs better for most nuclei, and good predictions can be obtained with 100 snapshots. Only the amide proton with very special chemical shift distribution (see Figure 3) has to be described by many more snapshots (>600).

## CONCLUSION AND DISCUSSION

Taking NMA as a model system, we have shown that the combination of ab initio molecular dynamics with explicit solvent molecules and quantum-chemical calculations of NMR parameters based on the generated ensembles results in an almost perfect agreement with experiment. Significant differences have been found between the structures of the NMA solvation shell modeled by ab initio and classical MD. Furthermore, the distribution of bond distances differs considerably between the two MD approaches. It should be noted that ab initio MD allow for anharmonic molecular vibrations, whereas current force fields rely almost exclusively on the harmonic approximation.

Other studies,[54−56,79−83] applying similar approaches to include conformational and solvent effects on different systems, found also the importance of including the dynamics of the system and its surroundings correctly to get a good agreement with experiment. By taking conformational sampling and explicit solvent effects into account, the accuracy of the quantum chemical calculations of NMR chemical shifts for proteins can be pushed to a limit where they become very interesting for a number of applications. Therefore the statement that quantum chemical calculations are too unreliable[17] has probably to be relativized and to be limited to studies with single structures without ensemble averaging. Concentrating on the amide group as a very important fragment in biomolecular system, we have shown here that the remaining errors especially for polar protons can be attributed to the inaccuracies in the ensembles generated by state-of-the-art force fields, particularly regarding hydrogen bonds, which are predicted as too long and too short-lived, as well as to the slightly inaccurate distributions of internal coordinates demonstrated by multilinear regressions. Especially the probability distribution of the $C_2$−O and N−H distances shows the largest deviations between the classical and ab initio MD; it is hence not surprising that they are important to describe the spread in the distributions as well as the differences of the chemical shieldings in the two ensembles. We have shown that the SHAKE algorithm or similar algorithms should probably be avoided in MD simulations when geometry snapshots are used for the calculations of NMR parameters as these algorithms limit the distributions of the distances of covalently bound hydrogen atoms. A proper sampling of all bond distances is necessary to obtain chemical shifts comparable with the experiment.

Since ab initio simulations are not possible for a large biochemical system on a routine basis, improvements of the force fields are highly desirable especially for noncanonical structures. As described in the Introduction, these new developments are already guided by a comparison of the calculated (at the moment almost exclusively using empirical methods) and experimental chemical shifts. Even though empirical methods have the clear advantage of speed, there are at least three reasons why we think that quantum-chemical

NMR chemical-shift calculations are beneficial for force-field development:

(1) Quantum-chemical calculations of NMR parameters can be applied to structures even if the available experimental data are not sufficient for the parametrization of empirical methods. This is especially the case for RNA/DNA as well as protein−ligand complexes; studies to evaluate the accuracy obtainable for such systems are being performed.

(2) The quantum-chemical calculations are heavily dependent on the instantaneous structure which was demonstrated by the very broad distributions of calculated chemical shifts in our previous publication[24] as well as the study presented here. When experimenting with empirical methods like SHIFTX2[21] and SPARTA+,[14] we saw a much weaker influence of slight variations in the structure (data not shown). This is not unexpected, because the empirical methods are parametrized based on experimental NMR parameters, which intrinsically reflect the averaging of fast molecular motions. Therefore, changes in the thermodynamic ensembles caused by modified force-field parameters will be much more pronounced with QM than with empirical NMR chemical-shift calculations. Additionally, also interactions between the solute and solvent can be evaluated, which is not possible with empirical methods since these interactions are implicitly included in the parameters.

(3) The empirical methods are parametrized on experimental structures. Nevertheless, there are issues with structures determined with both of the two important structure-determination methods: NMR and X-ray. On the one hand, the refinement steps of protein-structure determination by NMR use, in addition to the experimental data, force fields in simulated-annealing optimizations. Therefore, the force field puts a bias onto the structures, which has an impact on the parameters of the empirical NMR calculations and, in turn, influences the validation of the improved force field. On the other hand, protons are usually not resolved in X-ray structures and have to be added most often using the standard placement and force-field-based optimization. One practical consequence is the problem in the determination of hydrogen-bonding distances: A comparison to the CPMD ensemble shows that the intermolecular H−O distance is too long in the classical MD ensemble. In contrast, the C═O distance and the N−H distances are too short. In combination, the acceptor−donor distance, the only distance seen in the crystal, is approximately the same in both ensembles. As a consequence, the parameters of the empirical NMR prediction methods will be based on the wrong proton positions, and the inaccuracies herein cannot be identified.

Since the results presented here and elsewhere[54−56,79−83] prove the fragment-based approach as well as the theory/basis set combination is valid and identified the sampling as the main reason for success or failure, we have started investigations of biologically more relevant systems like peptides and DNA/RNA fragments using different versions of the AMBER force fields to see the influence of the modified parameter sets as well as polarization effects on the generated ensembles and the predicted NMR chemical shifts. Simultaneously, SHAKE will be investigated in more detail.

## ■ ASSOCIATED CONTENT

### Ⓢ Supporting Information

The correlation between the calculated and experimental $^{13}C$ NMR chemical shifts of the N-methyl acetamide, the distributions of the chemical shieldings for specific atoms for the two ensembles, the correlation of chemical shieldings with internal degrees of freedom, the distributions of specific distances as well as the averages and standard deviation over 10000 random samples and the visualization of the convergence of the chemical shifts with the number of snapshots for two additional nuclei. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION

### Corresponding Author

*Phone: +49-7071-2976969. E-mail: thomas.exner@uni-tuebingen.de. Corresponding author address: Institute of Pharmacy, Eberhard Karls University Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany (T.E.E.). Phone: +420-220-183-327. E-mail: dracinsky@uochb.cas.cz. Corresponding author address: Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Flemingovo náměstí 2, 166 10 Prague, Czech Republic (M.D.).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Dr. Bour for providing the software for the PVS method. The work has been 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 Juniorprofessoren-Programm of the state Baden-Württemberg, and the Czech Science Foundation (grant no. 13-24880S). Additionally, we thank the Baden-Württemberg grid (bwGRiD), which is part of the D-Grid system, for providing the computer resources making the computations possible.

## ■ REFERENCES

(1) 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.

(2) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(3) Best, R. B.; Hummer, G. *J. Phys. Chem. B* **2009**, *113*, 9004−9015.

(4) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(5) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400−1415.

(6) van Gunsteren, W. F. *GROMOS. Groningen Molecular Simulation Program Package*; University of Groningen: Groningen, 1987.

(7) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.

(8) Mackerell, A. D. *J. Comput. Chem.* **2004**, *25*, 1584−1604.

(9) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236−242.

(10) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712−725.

(11) Krepl, M.; Zgarbova, M.; Stadlbauer, P.; Otyepka, M.; Banas, P.; Koca, J.; Cheatham, T. E.; Jurecka, P.; Sponer, J. *J. Chem. Theory Comput.* **2012**, *8*, 2506−2520.

(12) Best, R. B.; Buchete, N. V.; Hummer, G. *Biophys. J.* **2008**, *95*, L07−L09.

(13) Beauchamp, K. A.; Lin, Y. S.; Das, R.; Pande, V. S. *J. Chem. Theory Comput.* **2012**, *8*, 1409−1414.

(14) Shen, Y.; Bax, A. *J. Biomol. NMR* **2010**, *48*, 13−22.

(15) Aliev, A. E.; Courtier-Murias, D. *J. Phys. Chem. B* **2010**, *114*, 12358−12375.

(16) Yildirim, I.; Stern, H. A.; Tubbs, J. D.; Kennedy, S. D.; Turner, D. H. *J. Phys. Chem. B* **2011**, *115*, 9261−9270.
(17) Allison, J. *Biophys. Rev.* **2012**, *4*, 189−203.
(18) Lehtivarjo, J.; Tuppurainen, K.; Hassinen, T.; Laatikainen, R.; Peräkylä, M. *J. Biomol. NMR* **2012**, *52*, 257−267.
(19) Camilloni, C.; Cavalli, A.; Vendruscolo, M. *J. Phys. Chem. B* **2013**, *117*, 1838−1843.
(20) Shapiro, Y. E.; Meirovitch, E. *J. Phys. Chem. B* **2012**, *116*, 4056−4068.
(21) Han, B.; Liu, Y.; Ginzinger, S.; Wishart, D. *J. Biomol. NMR* **2011**, *50*, 43−57.
(22) Frank, A.; Onila, I.; Möller, H. M.; Exner, T. E. *Proteins* **2011**, *79*, 2189−2202.
(23) Frank, A.; Möller, H. M.; Exner, T. E. *J. Chem. Theory Comput.* **2012**, *8*, 1480−1492.
(24) Exner, T. E.; Frank, A.; Onila, I.; Möller, H. M. *J. Chem. Theory Comput.* **2012**, *8*, 4818−4827.
(25) Zhu, T.; He, X.; Zhang, J. Z. H. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7837−7845.
(26) Zhu, T.; Zhang, J. Z. H.; He, X. *J. Chem. Theory Comput.* **2013**, *9*, 2104−2114.
(27) Elgabarty, H.; Schmieder, P.; Sebastiani, D. *Chem. Sci.* **2013**, *4*, 755−763.
(28) Mulder, F. A. A.; Filatov, M. *Chem. Soc. Rev.* **2010**, *39*, 578−590.
(29) Casabianca, L. B.; de Dios, A. C. *J. Chem. Phys.* **2008**, *128*, 052201−052210.
(30) Vila, J. A.; Arnautova, Y. A.; Martin, O. A.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16972−16977.
(31) Sun, H.; Oldfield, E. *J. Am. Chem. Soc.* **2004**, *126*, 4726−4734.
(32) 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.
(33) Vila, J. A.; Scheraga, H. A. *Proteins* **2008**, *71*, 641−654.
(34) Vila, J. A.; Arnautova, Y. A.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1891−1896.
(35) Vila, J. A.; Ripoll, D. R.; Scheraga, H. A. *J. Phys. Chem. B* **2007**, *111*, 6577−6585.
(36) Jacob, C. R.; Visscher, L. *J. Chem. Phys.* **2006**, *125*, 194104.
(37) Lee, A. M.; Bettens, R. P. A. *J. Phys. Chem. A* **2007**, *111*, 5111−5115.
(38) Johnson, E. R.; DiLabio, G. A. *J. Mol. Struct. THEOCHEM* **2009**, *898*, 56−61.
(39) He, X.; Wang, B.; Merz, K. M., Jr. *J. Phys. Chem. B* **2009**, *113*, 10380−10388.
(40) Hori, S.; Yamauchi, K.; Kuroki, S.; Ando, I. *Int. J. Mol. Sci.* **2002**, *3*, 907−913.
(41) Tang, S.; Case, D. *J. Biomol. NMR* **2007**, *38*, 255−266.
(42) Tang, S.; Case, D. *J. Biomol. NMR* **2011**, *51*, 303−312.
(43) Xu, X. P.; Case, D. A. *Biopolymers* **2002**, *65*, 408−423.
(44) Manalo, M. N.; de Dios, A. C. *J. Mol. Struct. THEOCHEM* **2004**, *675*, 1−8.
(45) Cai, L.; Fushman, D.; Kosov, D. *J. Biomol. NMR* **2009**, *45*, 245−253.
(46) Cai, L.; Kosov, D.; Fushman, D. *J. Biomol. NMR* **2011**, *50*, 19−33.
(47) Cai, L.; Fushman, D.; Kosov, D. *J. Biomol. NMR* **2008**, *41*, 77−88.
(48) Flaig, D.; Beer, M.; Ochsenfeld, C. *J. Chem. Theory Comput.* **2012**, *8*, 2260−2271.
(49) Gao, Q.; Yokojima, S.; Fedorov, D. G.; Kitaura, K.; Sakurai, M.; Nakamura, S. *J. Chem. Theory Comput.* **2010**, *6*, 1428−1444.
(50) Gao, Q.; Yokojima, S.; Kohno, T.; Ishida, T.; Fedorov, D. G.; Kitaura, K.; Fujihira, M.; Nakamura, S. *Chem. Phys. Lett.* **2007**, *445*, 331−339.
(51) Pandey, M. K.; Ramamoorthy, A. *J. Phys. Chem. B* **2012**, *117*, 859−867.
(52) Precechtelova, J.; Munzarova, M. L.; Vaara, J.; Novotny, J.; Dracinsky, M.; Sklenar, V. *J. Chem. Theory Comput.* **2013**, *9*, 1641−1656.
(53) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(54) Dracinsky, M.; Kaminsky, J.; Bour, P. *J. Phys. Chem. B* **2009**, *113*, 14698−14707.
(55) Dracinsky, M.; Bour, P. *J. Chem. Theory Comput.* **2009**, *6*, 288−299.
(56) Banyai, D. R.; Murakhtina, T.; Sebastiani, D. *Magn. Reson. Chem.* **2010**, *48*, S56−S60.
(57) *HyperChem* 8.0.3; Hypercube, Inc.: Gainesville, 2007.
(58) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(59) *CPMD*; IBM Corp. and MPI für Festkörperforschung: Stuttgart, 2008.
(60) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(61) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892−7895.
(62) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(63) Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I. F.; Mundy, C. J. *J. Phys. Chem. B* **2005**, *110*, 3685−3691.
(64) Dracinsky, M.; Castano, O.; Kotora, M.; Bour, P. *J. Org. Chem.* **2010**, *75*, 576−581.
(65) Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032−3041.
(66) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151−5158.
(67) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 253−260.
(68) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; 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.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision B.01*; Gaussian, Inc.: Wallingford, CT, 2009.
(69) Dunning, J. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(70) McWeeny, R. *Phys. Rev.* **1962**, *126*, 1028−1034.
(71) Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789−807.
(72) Wolinski, K.; Hilton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.
(73) Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209−214.
(74) 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.
(75) Lorieau, J. L.; Louis, J. M.; Bax, A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 11341−11346.
(76) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B. KNIME: The Konstanz Information Miner. In *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)*; Springer: 2007; pp 319−326.
(77) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(78) Kessler, J.; Dracinsky, M.; Bour, P. *J. Comput. Chem.* **2013**, *34*, 366−371.
(79) Pauwels, E.; Claeys, D.; Martins, J. C.; Waroquier, M.; Bifulco, G.; Speybroeck, V. V.; Madder, A. *RSC Adv.* **2013**, *3*, 3925−3938.
(80) Sternberg, U.; Birtalan, E.; Jakovkin, I.; Luy, B.; Schepers, U.; Brase, S.; Muhle-Goll, C. *Org. Biomol. Chem.* **2013**, *11*, 640−647.
(81) Eriksen, J. J.; Olsen, J. M.; Aidas, K.; Agren, H.; Mikkelsen, K. V.; Kongsted, J. *J. Comput. Chem.* **2011**, *32*, 2853−2864.
(82) Rohrig, U. F.; Sebastiani, D. *J. Phys. Chem. B* **2008**, *112*, 1267−1274.

(83) Komin, S.; Gossens, C.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *J. Phys. Chem. B* **2007**, *111*, 5225−5232.