PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2024, **26**, 6386

The accuracy limit of chemical shift predictions for species in aqueous solution†‡

Stefan Maste, [a] Bikramjit Sharma, [b] Tim Pongratz, [a] Bastian Grabe, [a] Wolf Hiller, [a] Markus Beck Erlach, [c] Werner Kremer, [c] Hans Robert Kalbitzer, [c] Dominik Marx*[b] and Stefan M. Kast*[a]

Received 10th November 2023,
Accepted 9th January 2024

DOI: 10.1039/d3cp05471c

rsc.li/pccp

Interpreting NMR experiments benefits from first-principles predictions of chemical shifts. Reaching the accuracy limit of theory is relevant for unambiguous structural analysis and dissecting theoretical approximations. Since accurate chemical shift measurements are based on using internal reference compounds such as trimethylsilylpropanesulfonate (DSS), a detailed comparison of experimental with theoretical data requires simultaneous consideration of both target and reference species ensembles in the same solvent environment. Here we show that *ab initio* molecular dynamics simulations to generate liquid-state ensembles of target and reference compounds, including explicitly their short-range solvation environments and combined with quantum-mechanical solvation models, allows for predicting highly accurate ^{1}H (∼0.1–0.5 ppm) and aliphatic ^{13}C (∼1.5 ppm) chemical shifts for aqueous solutions of the model compounds trimethylamine *N*-oxide (TMAO) and *N*-methylacetamide (NMA), referenced to DSS without any system-specific adjustments. This encompasses the two peptide bond conformations of NMA identified by NMR. The results are used to derive a general-purpose guideline set for predictive NMR chemical shift calculations of NMA in the liquid state and to identify artifacts of force field models. Accurate predictions are only obtained if a sufficient number of explicit water molecules is included in the quantum-mechanical calculations, disproving a purely electrostatic model of the solvent effect on chemical shifts.

1 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is an analytical key technique for structural and dynamical characterization of chemical and biological systems under a variety of conditions. With growing complexity of experimental methods and systems studied, it is increasingly more difficult to clearly assign molecular detail to observed response.[1–3] Computational methods are therefore relevant for interpreting spectral features if they are truly capable of accurately reproducing NMR parameters such as chemical shifts. NMR data sensitively depend not only on the structure and dynamics of the target compound, but also on solvent conditions and composition. In this context, first-principles approaches based on quantum-mechanical (QM) calculations (nowadays dominated by density functional theory, DFT) are the methods of choice for supporting structure determination as they can in principle be applied in situations where empirical methods are problematic due to insufficient training data.[4–6] The computational protocol usually consists of the generation of relevant conformations, geometry optimizations, and NMR calculations. Tools to automatically calculate NMR parameters using such workflows have been developed.[7–10] The variety of methods in terms of structure generation, levels of theory for geometry optimization, and NMR calculations using specific solvent models leads to a wide array of options for the practitioner. However, they are not necessarily transferable between different systems and their quantitative performance, *i.e.* predictive capabilities under new conditions not met before remains unclear. Hence, it is highly desirable to develop a reference framework for predicting chemical shifts, the property focused on in this work, with high accuracy and with as little empirical input as possible, benchmarked for experimentally well-defined model compounds. By gradually increasing the sophistication of the overall method to the quantitative limit we are able to delineate the impact of all individual approximations of the modelling setup

[a] *Fakultät für Chemie und Chemische Biologie, Technische Universität Dortmund, Otto-Hahn-Straße 4a, 44227 Dortmund, Germany. E-mail: stefan.kast@tu-dortmund.de*
[b] *Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany. E-mail: dominik.marx@theochem.ruhr-uni-bochum.de*
[c] *Fakultät für Biologie und Vorklinische Medizin, Universität Regensburg, 93040 Regensburg, Germany*

† Dedicated to Prof. Dr. Klaus Jurkschat on the occasion of his 70th birthday.

‡ Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3cp05471c and https://doi.org/10.17877/RESOLV-2024-lrgkievk

toward achieving predictive power for NMR calculations in the liquid state.

Determination of accurate experimental chemical shifts in solution requires the addition of a reference compound to the solution (internal reference) for cancelling effects of the magnetic susceptibility of the solvent on the magnetic field and thus on the resonance frequency and the measured shifts. The experimental shifts of a target molecule sensitively depend on many precisely controlled factors, the exact composition of the sample, the properties of the solvent, especially the pH and the ionic strength, and external factors such as temperature and pressure. As the ultimate test for the predictive power of a shift calculation is close agreement between measured and predicted values, it is important that the calculations reflect all relevant factors influencing the experimental outcome. In particular, internal referencing has to be properly handled in the calculations.

Consequently, the key factors that determine the accuracy of an NMR shift calculation are: (1) the adequate treatment of the solute's internal degrees of freedom and conformations in solution, (2) the description of the solvent environment, and (3) the QM level of theory for computing the primary quantities, the shielding constants ("chemical shielding") of nuclei of interest. While these issues have frequently been investigated in the past,[11–15] one additional, but rather important ingredient of a successful shift prediction has rarely, if at all, been addressed: (4) the same factors (1–3) also influence the treatment of the NMR "standard" or reference, as the chemical shift is given by the difference between shielding constants of reference and target compound nuclei. However, to the best of our knowledge, the structural and solvation ensemble effects on QM-based calculations of shielding constants of the commonly used reference species tetramethylsilane (TMS) and the water-soluble derivative trimethylsilylpropanesulfonate (DSS) for ^{1}H or ^{13}C have not been studied in the past. Ensemble effects based on snapshots from *ab initio* molecular dynamics (AIMD) simulation[16] of the reference have been reported only for the ^{15}N standard compound nitromethane,[17] whereas AIMD studies of the target species are more common. A recent review[18] lists 28 examples where AIMD simulations in combination with DFT-NMR calculations were performed. This highlights the advantages of generating accurate molecular ensembles including the solvent compared to the use of a limited number of optimized structures only.

The calculation of chemical shifts typically relies on regression with respect to experimental data in the sense of an empirical determination of the intercept representing the reference shielding constant.[7,19] Even if the target species along with a local solvent environment is sampled by AIMD simulations, as has been done for *N*-methylacetamide (NMA) in water[19] the residual error from averaging over QM-based snapshot calculations is quite large when using an empirical reference, on the order of 1 ppm for the amide proton. Alternatively, NMR calculations on geometry optimized structures of TMS or DSS are established and frequently used in the literature,[20–23] allowing for a direct referencing approach mimicking an internal standard. However, fitting the reference to experimental data[7,21] as well as secondary or multi-standard methods[24,25] are usually preferred[26] as they lead to more accurate shift calculations due to improved error compensation.

Generally, this means that only relative chemical shifts are commonly computed in the sense of an additive correction to the directly calculated shielding constant. However, the adequate treatment of the standard is relevant for a robust prediction of absolute shifts in the sense that the methodology should be applicable under conditions where empirical data is scarce or not available. *A priori* it is not known how strongly ensemble and solvation effects influence the experimentally measured data. For instance, we have previously demonstrated that the pressure response of DSS resonances in water is small,[27] which allows for a clear assignment of the pressure dependence of NMR shifts to the target species alone. Similarly, NMR experiments in complex solvent mixtures with direct, *i.e.* internal referencing can be understood only if the solvation and ensemble effects on the standard compound are modelled at a similar level of detail as for the target species.

Therefore, we set out here to investigate in a proof-of-concept study if essentially quantitative agreement with experiment can be achieved for chemical shift calculations of the model compounds NMA and trimethylamine *N*-oxide (TMAO) in water referenced against DSS and, in addition, which level of sophistication is necessary to meet this goal. The small range of model compounds is a consequence of the enormous computational cost associated with AIMD simulations to generate ensembles. As will be shown below, this methodology is necessary in order to achieve close correspondence with experimental reference data. To this end, it was also important to obtain a reliable NMR dataset for the two compounds under controlled conditions for the present study. As a special problem, NMA contains an NH–CO amide bond representing an analogous structure to a peptide bond that can occur as *cis* and *trans* isomer in solution. The corresponding two isomers were described by NMR for NMA dissolved in CCl_4[28] but not yet in aqueous solution.

All these species have been treated theoretically by us in the past with optimized structures,[20,27,29] indicating the limits of neglecting structural fluctuations and short-range solvation effects by explicitly included solvent species in the context of QM solvation models. As to the latter, the polarizable continuum model[30] (PCM) is frequently employed for NMR calculations, while its performance, especially for protic solvents, can be improved by addition of a few explicit solvent molecules,[31,32] though still leading to large absolute errors for AIMD-sampled NMA with an empirical standard correction.[19]

As an alternative to continuum calculations we have in the past investigated the performance of the embedded cluster reference interaction site model (EC-RISM) as a solvation approach for QM methods in the context of NMR predictions.[27,33] This model combines calculations of the electronic and liquid structure by adding the interaction of a set of polarizing background charges to the electronic Hamiltonian which, as opposed to continuum models, represent the electrostatic potential distribution derived

from the solute–solvent site distribution functions (SDFs) of an atomistically resolved solvent description within the 3D RISM approximation. In this way, the effect of directional hydrogen bonds can be mimicked even in the absence of explicit solvent molecules. While EC-RISM already yielded improved NMR shift results compared to PCM calculations on optimized structures,[27] its performance advantage was particularly evident in the context of calculating isotropic hyperfine coupling constants for electron paramagnetic resonance (EPR) spectroscopy, as demonstrated recently.[34,35] In a similar spirit as in the present work, we worked out the accuracy limit for such calculations on the basis of AIMD-generated ensembles of a spin probe in water. By comparing to fully explicit solvation with a quantum-mechanical/molecular-mechanical (QM/MM) approach, we showed that EC-RISM as a hybrid solvation model is capable of replacing the short-range explicit solvation effects, unlike PCM.

2 Results and discussion

2.1 Experiments and computational approach

On the experimental side, new, more complete reference data for TMAO and NMA than already published by us[20,27] with direct referencing to respective DSS methyl nuclei was produced, for both NMA and TMAO. In addition, the ^{1}H, ^{13}C, and ^{15}N resonances in *cis*-NMA could be identified and characterized for the first time by NMR in aqueous solution (see ESI,‡ Tables S1, S2 and Fig. S1, S2) where the *cis*-isomer occurs in a relative concentration of approximately 1.6%. These data complete the experimental characterization of this important benchmark molecule. The influence of variations of pH, dilution, and temperature on the chemical shifts was checked to account for possible differences in theoretical workflows (see ESI‡).

On the computational side, based on our previous understanding we now turn to the more complicated NMR shift problem by treating the target species, NMA and TMAO, on the same footing as, for the first time, the reference DSS. All species were no longer assumed to be rigid bodies, but represented by AIMD-generated ensembles of fluctuating molecules in a fully explicit aqueous solvent environment. This allows for a one-to-one comparison with experimental data, see ESI,‡ for full details. Dynamics and, therefore, the NMR timescale does not play a role for the statistical analysis as magnetic equivalence of nuclei and distinguishability of conformations is assumed to be known according to the experimental assignments.

We determined the convergence and therefore the accuracy limit by taking *ca.* 400 snapshots (for convergence down to *ca.* 0.05 ppm uncertainty for ^{1}H and 0.2 ppm for ^{13}C with respect to the number of snapshots see Fig. S3, ESI‡) from the trajectories of NMA, TMAO, and DSS and successively enlarged the explicitly treated short-range solvation environment around all nuclei of interest in the spirit of hybrid solvation models. Analyses were performed without further background models and with augmentation by EC-RISM or PCM, and by an MM background (thus providing the standard QM/MM ansatz in addition). This allows us to treat the impact of widely used model approximations on target and reference species separately and, therefore, to study the relevance of error compensation. By comparing with analogous ensembles taken from standard force field-based molecular dynamics (FFMD) simulations we can clearly determine the effect and origin of FF model approximations. We focused on directly referenced ^{1}H and ^{13}C nuclei of the target species with respect to DSS, analyzing amide H as well as methyl H and C in *cis*- and *trans*-NMA and in TMAO in the main text (for experimental NMR data see ESI‡).[15]N data were not included in the computational part of the study since the experimental data (separately shown in ESI,‡ Table S2) could only be referenced indirectly to internal DSS according to the IUPAC-IUB recommendations,[36] which cannot be mimicked by chemical shift calculations as done in this work. For the detailed analysis of all model variants we did not vary the QM level of theory but employed OLYP/6-311+G(d,p)[37–39] GIAO-DFT as this approximation has shown good performance for NMR calculations in a variety of benchmark studies.[40–42] The low computational cost facilitated calculations on hundreds of structures with increasing number of explicit water molecules in the QM zone, bringing standard errors of the calculations down to below 0.1 ppm for ^{1}H and 0.25 ppm for ^{13}C (see Tables S3–S12 in the ESI‡). It turned out that ^{13}C of the carbonyl group in NMA requires special QM treatment, as will be discussed below.

In other studies, varying the QM size for the calculation of isotropic shielding constants was for instance studied for several systems in water based on FFMD but excluding the ensemble of the standard.[43] This latter work recommended a QM zone of 8–10 Å, which could be reduced by a factor of around 2 to 4 depending on the analyzed molecule and nucleus by adding an MM background. In another previous NMR study[19] the authors used a selection of explicit water molecules with a radius of 4 Å around all nuclei and only analyzed the *trans*-conformer of NMA, using empirical referencing. We here call this selection mode ''full'' and examine a second selection protocol, termed ''local'', defined by treating only those solvent molecules in the QM zone whose centers were located within a certain distance from all magnetically equivalent target nuclei (using methyl-H positions as basis for methyl-^{13}C). This way the number of QM solvent molecules could be kept smaller compared to the ''full'' model at the same truncation distance.

In both ''full'' and ''local'' cases, the snapshots of the respective solute species together with their explicit short-range solvation environments were treated as taken from AIMD without further background field (no prefix to ''full''/''local'', in Fig. 1 only), and by EC-RISM, PCM, and QM/MM (annotated by prefixes ''ECR'', ''PCM'', ''QM/MM'' to ''full/local'') backgrounds, in the QM/MM case representing water point charges by the TIP3P model.[44] Trends can be illustrated in two variants, by plotting NMR results as a function of distance or of the number of surrounding water molecules (corresponding to a certain distance). While convergence of shieldings is better represented in the latter way when starting from an origin of 0, for chemical shifts the distance representation allows for

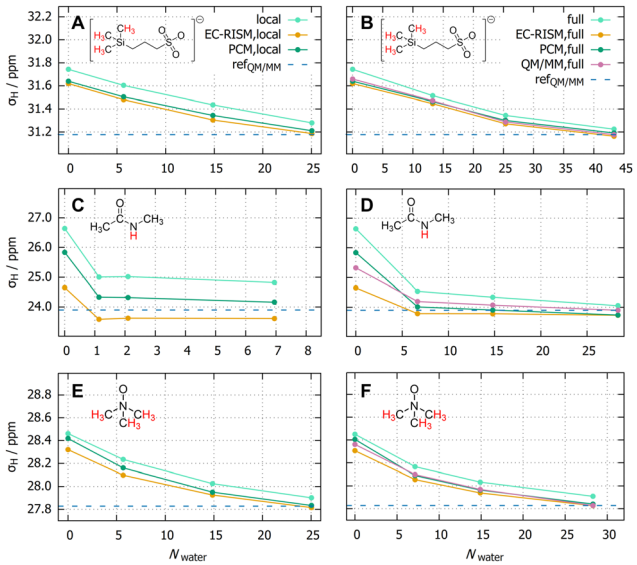

Fig. 1 Isotropic shielding constants of the highlighted nuclei (in red), averaged over magnetically equivalent atoms, for increasing number of explicit water molecules included in the NMR calculation with "local" (A), (C) and (E). and "full" (B), (D) and (F) explicit water selection modes and different background models annotated by subscripts. "ref$_{QM/MM}$" corresponds to the "full" selection scheme at 4 Å and is also shown as a dashed blue line for easier comparison. The complete depiction including all highlighted nuclei is shown in Fig. S4 in the ESI.‡ Raw data is found in Tables S3–S12 (ESI‡).

different numbers of solvent molecules for target and reference at the same truncation distance. In order to test the electrostatic hypothesis as the dominant factor for solvent effect on chemical shifts, we evaluated the models QM/MM, EC-RISM, and PCM without a single explicit water in the QM zone, annotated by the additional suffix "0". This allows for a comparison with the respective "full" models in order to evaluate the limits of a purely electrostatic solvation treatment. We furthermore add analyses based on FFMD and on a single optimized structure, designated by superscripts "FFMD" and "opt" to shieldings and shifts instead of "AIMD".

2.2 Shielding constants

Table 1 shows the shielding constants of a representative set of nuclei in the reference "AIMD-QM/MM,full" solvation framework ($\sigma^{AIMD}_{QM/MM,full}$) and deviations of this quantity for other solvation schemes with respect to this reference (see Table S12 for other nuclei, ESI‡). We first note that the "full" scheme at the largest number of water molecules (corresponding to a 4 Å shell) consistently yields a stable value with very little deviation between different background models (up to 0.17 ppm for amide ^{1}H, 0.02 ppm for methyl ^{1}H, and 0.34 ppm for ^{13}C), *i.e.* close to the statistical uncertainty. We can therefore assume that these values are truly converged and represent the accuracy limit under the remaining approximations.

In the following the abbreviation "ref$_{QM/MM}$" is used for the theoretical reference, obtained from a "full" 4 Å explicit water shell based on the AIMD ensemble with MM background charges.

Fig. 1 shows the calculated absolute values of the shielding constants for the representative set of nuclei (as in Table 1) with all the solvation schemes. For completeness, pure AIMD snapshot calculations without any background are also shown in Fig. 1, indicating the necessity to add a background model to reach convergence. While the difference between "local" and "full" schemes is quite small for DSS and TMAO, and even no inclusion of explicit water molecules induces an error with respect to the limit of *ca.* 0.5 ppm only (with a small advantage of EC-RISM over PCM), for NMA the difference is larger. 0.4 ppm remain for PCM and EC-RISM even at 4 Å "local" explicit solvent corresponding to 7 water molecules, which can be attributed to the missing explicit solvation of the carbonyl group, see Fig. 2B. Similarly, missing explicit solvation of the sulfonate group of DSS (Fig. 2A) correlates with the 0.1 ppm difference in Fig. 1A that remains using local explicit solvent when EC-RISM or PCM are added compared to pure local explicit solvation. EC-RISM shows consistent advantages over PCM for the less expensive "local" scheme, being closer to the limit for all truncations including 0 water molecules (where the error is *ca.* 0.8 ppm). Notably, just one explicit water near the amide H in NMA reduces the error to less than 0.3 ppm compared to reference that requires *ca.* 29 water molecules.

2.3 Chemical shifts

Turning now to the benchmarking of solvation schemes for representative chemical shifts as shown in Table 2 and Fig. 3, we can confirm the accuracy limit for the "full" scheme at *ca.* 4 Å ("AIMD-QM/MM,full"), corresponding to "ref$_{QM/MM}$" in the figures. Comparison of the deviations between calculated and experimental data, $\Delta\delta$, for AIMD with QM/MM, EC-RISM, and PCM backgrounds in the "full" scheme among each other reveals close agreement (−0.43/−0.56, −0.39/−0.40, −0.40/−0.38 ppm for *cis*/*trans*-NMA$_{NH}$; 0.10, 0.09, 0.10 ppm for TMAO$_H$; −1.35, −1.42, −1.34 ppm for TMAO$_C$). The variation among the background models (*ca.* 0.1 ppm for amide ^{1}H, 0.01 ppm for methyl ^{1}H, 0.1 ppm for methyl ^{13}C) does not exceed much the experimental uncertainties (Table S2 and ESI‡). Hence, the "full" approach represents the converged limit within our chosen models.

Next, we can assign the remaining difference to the experimental reference as the realistic accuracy limit of chemical shift calculations under the given approximations. Consequently, we are only now in the position to systematically test the QM model performance for chemical shift predictions in the future, by excluding most other sources of error. In our case, the remaining theoretical uncertainties are *ca.* 0.5 ppm for amide-^{1}H, 0.1 ppm for methyl-^{1}H, and 1.5 ppm for methyl-^{13}C, as determined from the "full"-scheme AIMD ensemble data and the three background models (see columns $\Delta\delta^{AIMD}_{QM/MM,full}$, $\Delta\delta^{AIMD}_{ECR,full}$, $\Delta\sigma^{AIMD}_{PCM,full}$ in Table 2). This is the key result of the present work. Moreover, the consistency of the deviation between predicted *cis* and *trans* NMA amide-^{1}H shifts from the experimentally assigned peaks removes any remaining uncertainty about the conformational *cis* origin of the signal at 7.10 ppm.

Table 1 Isotropic shielding constants [ppm] for all analyzed nuclei, averaged over magnetically equivalent atoms, for a 4 Å "full" (AIMD ensemble) explicit solvation shell augmented by an EC-RISM ("ECR"), PCM, and QM/MM (corresponding to "$\mathrm{ref}_{\mathrm{QM/MM}}$" in the figures, relative to which all other numbers are reported) background and corresponding 2.5 Å "local" (AIMD and FFMD ensembles) data; EC-RISM and PCM results for AIMD ensembles with 0 explicit water molecules (subscript "0") and for optimized structures from OLYP/6-311+G(d,p)/PCM (superscript "opt", subscript "0"). Additional data for all other ^{1}H and ^{13}C nuclei of NMA (restricted to the "0" and "full" solvation schemes) is provided in Table S12 (ESI‡)

Nucleus	$\sigma^{\mathrm{AIMD}}_{\mathrm{QM/MM,full}}$	$\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{QM/MM,0}}$	$\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{ECR,full}}$	$\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{ECR,loc}}$	$\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{ECR,0}}$	$\Delta\sigma^{\mathrm{opt}}_{\mathrm{ECR,0}}$	$\Delta\sigma^{\mathrm{FFMD}}_{\mathrm{ECR,full}}$	$\Delta\sigma^{\mathrm{FFMD}}_{\mathrm{ECR,loc}}$	$\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{PCM,full}}$	$\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{PCM,loc}}$	$\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{PCM,0}}$	$\Delta\sigma^{\mathrm{opt}}_{\mathrm{PCM,0}}$	$\Delta\sigma^{\mathrm{FFMD}}_{\mathrm{PCM,full}}$	$\Delta\sigma^{\mathrm{FFMD}}_{\mathrm{PCM,loc}}$
$\mathrm{DSS_H}$	31.18	0.48	−0.02	0.30	0.44	0.56	0.42	0.75	0.01	0.32	0.46	0.59	0.45	0.77
$\mathrm{DSS_C}$	182.33	3.66	0.21	2.38	4.20	5.09	3.32	5.63	0.34	2.04	3.77	4.80	3.42	5.17
cis-$\mathrm{NMA_{NH}}$	24.51	1.21	−0.06	−0.20	0.68	1.16	1.07	0.74	−0.02	0.39	1.76	2.22	1.15	1.41
trans-$\mathrm{NMA_{NH}}$	23.90	1.42	−0.17	−0.32	0.75	1.14	1.10	1.04	−0.17	0.43	1.94	2.25	1.15	1.82
$\mathrm{TMAO_H}$	27.83	0.53	0.00	0.27	0.49	0.72	0.51	0.75	0.01	0.32	0.59	0.82	0.55	0.83
$\mathrm{TMAO_C}$	121.50	0.44	0.14	0.62	0.16	1.28	5.54	5.55	0.20	0.61	−0.27	0.89	5.30	5.54

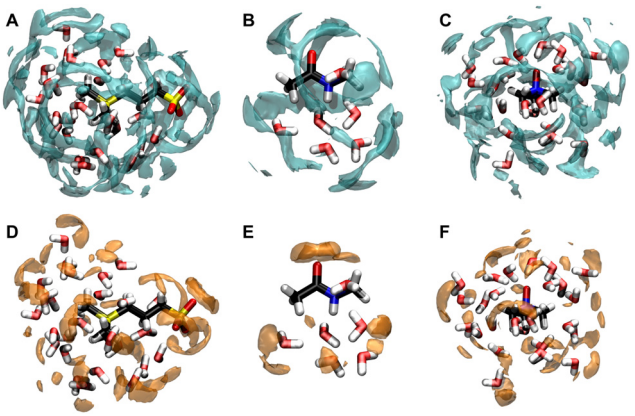

Fig. 2 Snapshots of DSS (A) and (D), NMA (B) and (E), and TMAO (C) and (F) with 4 Å of explicit water locally selected around the three methyl groups of DSS and TMAO and the amide proton of NMA, combined with water–O (A–C) and water–H (D–F) SDF isosurfaces (at 4-fold relative bulk density) from EC-RISM calculations. The figures were created by VMD.[45]

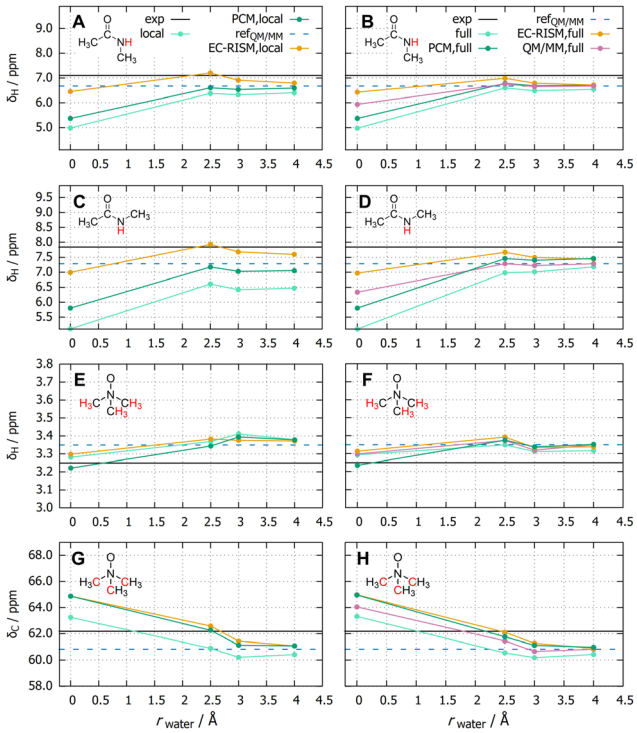

Fig. 3 Chemical shifts of the highlighted nuclei (in red), averaged over magnetically equivalent atoms, with increasing radius of the explicit water shell with "local" (A), (C), (E) and (G) and "full" (B), (D), (F) and (H) explicit water selection schemes and different background models annotated by subscripts, using same shell distances for target and reference. For ^{13}C shifts of TMAO (G) and (H) the distance criterion was applied to the hydrogen atoms of the methyl groups. The experimental chemical shifts are shown by a horizontal black line. "$\mathrm{ref}_{\mathrm{QM/MM}}$" corresponds to the "full" selection scheme at 4 Å and is also shown as dashed blue line for easier comparison. Experimental chemical shifts of TMAO and NMA are shown in Table S1 (ESI‡). Data of the corresponding nuclear shielding constants is shown in Tables S3–S12 (ESI‡).

While similar small uncertainties remain for the methyl groups of *trans*-NMA presented in Table S13 (ESI‡) ("full" approach only), the carbonyl-^{13}C is obviously more problematic (Table S13, ESI‡). Results for the shielding constant show again convergence of the sampling approach within the "full" model, whereas the absolute deviation from experiment remains large with 6.8 ppm for the QM/MM reference (slightly better with corresponding EC-RISM and PCM "full" setups). This points at the QM level of theory as the source of error for this particular nucleus. As a calculation of an ensemble average including all relevant explicit water molecules is computationally very demanding when a more reliable level such as MP2 is used (which is expected to be relevant particularly for carbonyl-^{13}C[46]), we estimated a correction to the DFT level by an extrapolation scheme. The difference between MP2/6-311+G(d,p) and OLYP results for AIMD ensembles with zero water and EC-RISM/PCM backgrounds provides an extrapolation approximation to be added to the respective "full" results. As can be seen from Table S14 (ESI‡), the deviation from experiment was found to be only 1.1 ppm, while the deviations for all other nuclei within this extrapolation scheme remain essentially similar to original OLYP values. We have tested this approximation on a few snapshots of *trans*-NMA with 2.5 Å "full" solvation for which explicitly solvated MP2 calculations could be performed, and found on average *ca.* 0.5 ppm deviation between extrapolated and exact MP2 results for carbonyl-^{13}C shielding constants. This clearly shows that our methodology is able to dissect and isolate the

Table 2 Chemical shifts [ppm] (calculated by the difference of shielding constants of corresponding DSS and target nuclei within a given model) for all analyzed nuclei, averaged over magnetically equivalent atoms, for a 4 Å "full" (AIMD ensemble) explicit solvation shell augmented by an EC-RISM ("ECR"), PCM, and QM/MM (corresponding to "$ref_{QM/MM}$" in the figures) background and corresponding 2.5 Å "local" (AIMD and FFMD ensembles) data; EC-RISM and PCM results for AIMD ensembles with 0 explicit water molecules (subscript "0") and for optimized structures from OLYP/6-311+G(d,p)/PCM (superscript "opt", subscript "0"), referenced to DSS with the corresponding solvation scheme. Experimental shifts of TMAO and NMA are reported in Table S1 (ESI‡). Including small effects of pH, temperature, and sample concentration on the proton shifts of TMAO and NMA, the experimental uncertainty is 0.01 ppm or less for all protons. For ^{13}C (and ^{15}N, not calculated in this work) it is smaller than 0.06 ppm (see ESI‡). Indirect referencing of ^{13}C to ^{1}H of DSS produces essentially identical shift values (see Table S1, ESI‡). Additional data for all other ^{1}H and ^{13}C nuclei of NMA (restricted to the "0" and "full" solvation schemes) is provided in Table S13 (ESI‡)

Nucleus	δ_{exp}	$\Delta\delta^{AIMD}_{QM/MM,full}$	$\Delta\delta^{AIMD}_{QM/MM,0}$	$\Delta\delta^{AIMD}_{ECR,full}$	$\Delta\delta^{AIMD}_{ECR,loc}$	$\Delta\delta^{AIMD}_{ECR,0}$	$\Delta\delta^{opt}_{ECR,0}$	$\Delta\delta^{FFMD}_{ECR,full}$	$\Delta\delta^{FFMD}_{ECR,loc}$	$\Delta\delta^{AIMD}_{PCM,full}$	$\Delta\delta^{AIMD}_{PCM,loc}$	$\Delta\delta^{AIMD}_{PCM,0}$	$\Delta\delta^{opt}_{PCM,0}$	$\Delta\delta^{FFMD}_{PCM,full}$	$\Delta\delta^{FFMD}_{PCM,loc}$
cis-NMA_{NH}	7.10	−0.43	−1.16	−0.39	0.07	−0.67	−1.04	−1.08	−0.42	−0.40	−0.52	−1.73	−2.06	−1.13	−1.09
trans-NMA_{NH}	7.84	−0.56	−1.51	−0.40	0.06	−0.87	−1.14	−1.24	−0.85	−0.38	−0.69	−2.04	−2.22	−1.25	−1.63
$TMAO_H$	3.25	0.10	0.05	0.09	0.14	0.05	−0.06	0.01	0.09	0.10	0.08	−0.03	−0.13	0.00	0.02
$TMAO_C$	62.18	−1.35	1.87	−1.42	0.40	2.55	2.32	−3.71	−1.41	−1.34	0.08	2.69	2.56	−3.37	−1.86

various approximation sources to gradually approach the accuracy limit.

An interesting phenomenon arises from combining two shielding calculations to yield the chemical shift: trends with increasingly large short-range explicit solvation shells are no longer monotonic and even allow for the identification of "sweet spots" in terms of the adequate explicit shell size. Apparently, 2.5–3 Å are enough for our systems, leading to very good results for all TMAO nuclei, regardless of background solvation, where the error compensation between target and reference nuclei is especially pronounced for ^{1}H. However, we again find substantial differences between TMAO and NMA regarding model performance at such small distances. At 2.5 Å (see Table 2) ^{1}H deviations for all background models are *ca.* 0.1 ppm on average for methyl groups irrespective of "full" or "local" schemes (ranging from 0.01 to 0.14 ppm for all AIMD-based "full" and "local" calculations shown in Table 2) and similarly *ca.* 1 ppm on average for methyl-^{13}C (range −1.42 to +0.40 ppm). Not unexpectedly, "full" or "local" schemes differ more strongly for amide-^{1}H and various background models using the AIMD ensemble by *ca.* 0.6 ppm (*cis*) and 0.8 ppm (*trans*), respectively. However, EC-RISM apparently outperforms PCM also for the experimental benchmark, yielding less than 0.1 ppm difference at 2.5 Å and being consistently the best model even in the absence of explicit water. For 0 water molecules, EC-RISM improves the error compared to PCM (*ca.* 2 ppm, in line with previous results[19]) for NMA to *ca.* 1 ppm, while a single water molecule (corresponding to 2.5 Å) reduces it further to less than 0.1 ppm in the "local" approach (see also ref. 47 for a detailed analysis of the NMA-water complex). Our combination of explicit treatment of the reference and an EC-RISM background model therefore yields considerable improvement over earlier approaches.

From these results, we are now in the position to answer the long-standing question as to the ability of a purely electrostatic model to explain the water-induced chemical shift modulation. With 0 explicit water molecules, both EC-RISM and PCM correspond to such purely electrostatically modeled environments. A third pure electrostatic variant is provided by using QM/MM with 0 water models, *i.e.* we can compare "QM/MM,full" as the reference with "QM/MM,0", as also shown in Table 2. Again, deviations for the amide proton are on the order of 1 ppm (−1.16 − (−0.43) = −0.73 ppm for *cis*-NMA and −1.51 − (−0.56) = −0.95 ppm for *trans*-NMA), in line with EC-RISM and PCM. Hence, explicit water molecules and their associated direct electronic interaction with the solute are essential for quantitative agreement, ruling out a purely electrostatic model at least for the amide-H.

Further insight into the relevance of individual contributions to the limiting values for shielding and shift can be gained from comparing the 0-water AIMD ensembles with the corresponding data for PCM-optimized single structures for all species (superscript "opt" in tables). In the spirit of our previous works this corresponds to "vertically de-solvating" the solute from the AIMD ensemble members and "re-solvating" them with PCM and EC-RISM background models. The difference between ensemble-averaged and optimized data then yields a measure for the relative importance of structural fluctuations of the target molecules to be compared with the reference values for explicitly solvated species. The results are also collected in Table 1 (difference of columns $\Delta\sigma^{AIMD}_{ECR,0} - \Delta\sigma^{opt}_{ECR,0}$, $\Delta\sigma^{AIMD}_{PCM,0} - \Delta\sigma^{opt}_{PCM,0}$), and Table 2 (difference of columns $\Delta\delta^{AIMD}_{ECR,0} - \Delta\delta^{opt}_{ECR,0}$, $\Delta\delta^{AIMD}_{PCM,0} - \Delta\delta^{opt}_{PCM,0}$) for (EC-RISM and PCM respectively). For ^{1}H shielding constants and chemical shifts the difference between re-solvated flexible ensemble averages and rigid structures is *ca.* 0.5 ppm at most, closer to 0.2 ppm for methyl-^{1}H, more or less independent of the background models PCM or EC-RISM. For ^{13}C the difference is *ca.* 1.1 ppm for shielding constants, and *ca.* 0.2 ppm for shifts. This clearly shows that taking rigid structures from geometry optimization is in principle a reasonable approach to NMR calculations. However, the effect of ignoring even a few explicit water molecules is considerably larger for all nuclei except for the ^{13}C shielding constant and the ^{1}H shift of TMAO, which are, probably coincidentally, rather small. Very strikingly, explicit solvation is particularly important for ^{13}C of DSS (*ca.* 5.1 ppm relative to the de-/re-solvated state) which directly translates into a *ca.* 2.5 ppm error for the TMAO ^{13}C shift. This can explain

an observation made earlier[20] that ^{13}C spectra and their pressure trends cannot be reproduced by EC-RISM model calculations in the absence of explicit water, and emphasizes again the relevance of the present attempt to fully take the reference compound into consideration.

Finally, as expensive AIMD simulations cannot routinely be performed for new systems, let alone large ones, it is advisable to investigate the performance of the explicit short-range solvation approaches using FFMD ensembles from commonly employed force fields. Results for the ''local'' (2.5 Å) and ''full'' (4 Å) schemes for shielding constants and shifts are also reported in Tables 1 and 2. For the ''local'' approach we find slightly larger deviations for shielding constants compared to the corresponding AIMD ensemble values for all nuclei, most notably for DSS/TMAO-^{13}C, where the discrepancy is large with more than 5 ppm (Table 1, column $\Delta\sigma^{\mathrm{FFMD}}_{\mathrm{ECR,local}}$). The ''full'' approach, which can faithfully be assumed to exhibit the converged limit also for FFMD, and for which EC-RISM and PCM background models again yield very similar results, particularly improves the error of the ^{13}C shielding constant of DSS to *ca.* 3 ppm (Table 1, column $\Delta\sigma^{\mathrm{FFMD}}_{\mathrm{ECR,full}}$), though not measurably for TMAO (5.54 *vs.* 5.55 ppm). Error cancellation then yields a satisfactory ^{13}C shift for TMAO only with the ''local'' solvation scheme (*ca.* 1.4 and 1.9 ppm difference for EC-RISM and PCM, respectively, see Table 2, columns $\Delta\delta^{\mathrm{FFMD}}_{\mathrm{ECR,local}}$ and $\Delta\delta^{\mathrm{FFMD}}_{\mathrm{PCM,local}}$), while the ^{1}H shift is practically quantitatively predicted for the FFMD ensemble. Also similar to AIMD, EC-RISM yields overall better agreement than PCM for amide ^{1}H in the ''local'' case, which hints at the capability of EC-RISM to mimic the water solvation structure at more distant nuclei better than PCM.

For hydrogens, similar to what has been found earlier[19] (*trans* only), the relative deviation between AIMD and FFMD is considerably stronger for NMA than for TMAO, which can be traced back to two main factors regarding the intra- and intermolecular description of NMA by a force field. As dominant intramolecular factor, a slightly (0.01 Å) shortened bond length of the amide proton was found from FFMD that correlates with an increased shielding constant of 0.76 ppm as demonstrated in Fig. S5 and S6 (ESI‡). This increase is significant as, despite the large variance shown in Fig. S6 (ESI‡), the standard errors are on the order of 0.1 ppm only throughout (Tables S3–S11, ESI‡). A slightly larger displacement (1.1 deg) for the amide dihedral angle was also observed for AIMD (Fig. S5 and S6, ESI‡). However, a correlation with shielding constants cannot be detected. On the intermolecular side, a less accurate description of the H-bonding situation around the amide-H was found for FFMD which yields a further 0.6 ppm difference in shielding constants even with just one explicit water molecule in the ''local'' solvation scheme. This gives rise to a total difference between AIMD and FFMD of 1.36 ppm for the amide proton of *trans*-NMA with the ''local'' solvation and EC-RISM background as shown in Table 1, columns $\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{ECR,local}} - \Delta\sigma^{\mathrm{FFMD}}_{\mathrm{ECR,local}}$, which stays about the same at 1.27 ppm when switching to the ''full'' solvation scheme (Table 1, columns $\Delta\sigma^{\mathrm{AIMD}}_{\mathrm{ECR,full}} - \Delta\sigma^{\mathrm{FFMD}}_{\mathrm{ECR,full}}$). This again shows the importance of the accurate description of the nearest solvent molecules. As can be seen from the radial distribution functions in Fig. S7 (ESI‡), the force field yields a less pronounced H-bond to the water–O whose distance is slightly larger than for AIMD, while the correspondence is very good for other species. As an explanation, a larger H-bond acceptor distance was attributed to increased shielding as a consequence of Pauli repulsion,[48] in line with our observation. As this relative ensemble error for ^{1}H is smaller for DSS, this effect directly translates into a stronger FFMD error for the shift, which is yet about half as large for EC-RISM than for a PCM background.

Still, the error can be directly attributed to shortcomings of the force field of the amide group coupled to a water model. Notably, the shift errors for AIMD and FFMD ensembles of the TMAO ^{1}H are almost identical (for both EC-RISM and PCM), which is the result of error compensation on the level of shielding constants. This compensation also partly comes into play for ^{13}C, though to a lesser extent, effectively reducing the error of *ca.* 5.5 ppm for FFMD shielding constants. From the discrepancy between EC-RISM and PCM in the ''full'' solvation scheme we can estimate the magnitude of the force field artifact. Compared to AIMD, shift predictions for TMAO-^{13}C deviate on the order of 2 ppm (see Table 2, columns $\Delta\delta^{\mathrm{AIMD}}_{\mathrm{ECR,full}} - \Delta\delta^{\mathrm{FFMD}}_{\mathrm{ECR,full}} = (-1.42 - (-3.71))$ ppm and columns $\Delta\delta^{\mathrm{AIMD}}_{\mathrm{PCM,full}} - \Delta\delta^{\mathrm{FFMD}}_{\mathrm{PCM,full}} = (-1.34 - (-3.37))$ ppm) more from the experimental reference, forming an additional layer of uncertainty over the accuracy limit identified in this work.

3 Conclusions

In summary, our results lead to the following conclusions:

(1) The accuracy limit for NMR calculations in aqueous solution at the selected QM level of theory can be reached by using a ''full'' solvation scheme with a 4 Å water shell sampled from explicit AIMD simulations of the solution together with a QM solvation model background.

(2) For a meaningful comparison of chemical shift predictions with experimental NMR data it is required to calculate simultaneously the chemical shifts of the reference compound used as internal standard in the experiments with the same methods, including solvation effects. For our model compounds representing polar and apolar solute–water interactions this yields on the order of at most 5% deviation from experiment throughout, corresponding to mean absolute errors of 0.18 ppm for the seven ^{1}H and 2.64 ppm for the seven ^{13}C nuclei (0.90 ppm without carbonyl) at the DFT level with the ''QM/MM,full'' solvation model (Table 2 and Tables S13, S14, ESI‡).

(3) In the spirit of hybrid solvation approaches, a simplified short-range explicit solvation scheme can only be employed if the background model retains molecular detail as provided by EC-RISM.

(4) The benchmark limit achieved allows for a clear characterization of force field artifacts. These have been identified not only for the amide hydration structure influencing the ^{1}H shielding, but also for ^{13}C of DSS.

(5) The drastic influence of tiny structural changes of only a few picometers on amide ^{1}H shifts should not only be considered for force field model development. This observation also has consequences for the interpretation of NMR responses as a function of environmental variables such as solvent composition, temperature, and pressure, which can all alter structures and, simultaneously, electronic polarization. Accurate calculations will support disentangling these effects.

(6) Short-range explicit solvation is more important than accounting for flexibility within a conformational well.

(7) The strong and unexpected impact of explicit water for treating DSS suggests that this reference compounds needs special consideration in the context of the chosen solvent conditions in order to yield a predictive theoretical model.

(8) A purely electrostatic interpretation of water-solvation effects on chemical shifts is clearly insufficient.

(9) More expensive levels of theory than DFT are necessary for adequately capturing the chemical shift of carbonyl-C.

Taken together, these insights provide the fundamentals for designing predictive first-principles NMR prediction workflows in future work which are computationally efficient and highly accurate at the same time. From a target compound perspective, extension to more and challenging systems and problems, such as the discrimination between diastereomers for which much smaller differences of corresponding chemical shifts can be expected, is an important goal. Clearly, given the computational cost required for reaching most accurate predictions, we are not yet at the point to deliver a recipe for structure determination, yet our results provide knowledge about quantitative limitations of available tools and approaches. From the reference standard perspective, it will be interesting to study the impact of co-solvents to water as well as non-ambient conditions such as extreme pressures on conformational DSS ensembles. The methodology presented is robust in the sense that it is directly applicable and scalable to more complex systems as long as costly AIMD can be replaced by empirical potential models for ensemble generation. Given the limitations of FFMD found in this work, absolutely accurate numbers are difficult to obtain. Hence, further force field optimization (for which the present methodology can provide suitable reference data), or switching to modern machine learning-based interaction potentials can represent a route to further progress.

Author contributions

S. M.: data curation: lead; formal analysis: lead; investigation: lead; methodology: lead; software: equal; writing – original draft: equal. B. S.: data curation: equal; investigation: equal; writing – review & editing: supporting. T. P.: formal analysis: supporting; methodology: equal. B. G.: data curation: equal; formal analysis: supporting; investigation: equal. W. H.: data curation: equal; formal analysis: equal; investigation: supporting. M. B. E.: data curation: equal; investigation: supporting. W. K.: data curation: equal; formal analysis: equal; investigation: equal. H. R. K.: data curation: lead; formal analysis: equal; supervision: lead; writing – review & editing: supporting. D. M.: conceptualization: equal; project administration: equal; supervision: lead; writing – review & editing: equal. S. M. K.: conceptualization: lead; methodology: equal; project administration: lead; supervision: lead; writing – original draft: equal; writing – review & editing: lead.

Data availability

Detailed experimental and computational methods as well as the full set of calculated results is available at {publication DOI}. Raw data (snapshot and optimized structures with corresponding shielding constants) is provided in machine-readable format under **https://doi.org/10.17877/RESOLV-2024-lrgkievk**. The authors have cited additional references in the ESI.‡ [49–82]

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2033 – Projektnummer 390677874 (to D. M. and S. M. K.), under the Research Unit FOR 1979 (to W. K., H. R. K., D. M. and S. M. K.), and under the Chemistry Consortium of the National Research Data Infrastructure, NFDI4Chem (to S. M. K.). We also thank the IT and Media Center (ITMC) of the TU Dortmund for computational support. The AIMD simulations have been carried out using computational resources provided by HPC@ZEMOS, HPC-RESOLV, and BoViLab@RUB. M. Hofmann and O. Reiser are acknowledged for providing us with ^{15}N-labeled NMA, as well as W. Kauter for supporting NMR measurements.

Notes and references

1 F. L. P. Costa, A. C. F. de Albuquerque, R. G. Fiorot, L. M. Lião, L. H. Martorano, G. V. S. Mota, A. L. Valverde, J. W. M. Carneiro and F. M. dos Santos Junior, *Org. Chem. Front.*, 2021, **8**, 2019.
2 J. Xiong, P.-J. Zhou, H.-W. Jiang, T. Huang, Y.-H. He, Z.-Y. Zhao, Y. Zang, Y.-M. Choo, X. Wang, A. G. Chittiboyina, P. Pandey, M. T. Hamann, J. Li and J.-F. Hu, *Angew. Chem., Int. Ed.*, 2021, **60**, 22270.
3 A. G. Kutateladze and D. S. Reddy, *J. Org. Chem.*, 2017, **82**, 3368.
4 E. Jonas, S. Kuhn and N. Schlörer, *Magn. Reson. Chem.*, 2022, **60**, 1021.
5 L. B. Krivdin, *Magn. Reson. Chem.*, 2019, **57**, 897.
6 A. Navarro-Vázquez, *Magn. Reson. Chem.*, 2017, **55**, 29.
7 P. H. Willoughby, M. J. Jansma and T. R. Hoye, *Nat. Protoc.*, 2014, **9**, 643.
8 P. H. Willoughby, M. J. Jansma and T. R. Hoye, *Nat. Protoc.*, 2020, **15**, 2277.

9 M. O. Marcarino, M. A. M. Zanardi, S. Cicetti and A. M. Sarotti, *Acc. Chem. Res.*, 2020, **53**, 1922.
10 Y. Yesiltepe, J. R. Nuñez, S. M. Colby, D. G. Thomas, M. I. Borkum, P. N. Reardon, N. M. Washton, T. O. Metz, J. G. Teeguarden, N. Govind and R. S. Renslow, *J. Cheminform.*, 2018, **10**, 52.
11 G. K. Pierens, *J. Comput. Chem.*, 2014, **35**, 1388.
12 D. Flaig, M. Maurer, M. Hanni, K. Braunger, L. Kick, M. Thubauville and C. Ochsenfeld, *J. Chem. Theory Comput.*, 2014, **10**, 572.
13 G. K. Pierens, T. K. Venkatachalam and D. C. Reutens, *Sci. Rep.*, 2017, **7**, 5605.
14 C. Benzi, O. Crescenzi, M. Pavone and V. Barone, *Magn. Reson. Chem.*, 2004, **42**, S57–S67.
15 L. B. Casabianca, *Magn. Reson. Chem.*, 2020, **58**, 611.
16 D. Marx and J. Hutter, *Ab Initio Molecular Dynamics. Basic Theory and Advanced Methods*, Cambridge University Press, 2009.
17 I. C. Gerber and F. Jolibois, *Phys. Chem. Chem. Phys.*, 2015, **17**, 12222.
18 A. H. Mazurek, Ł. Szeleszczuk and D. M. Pisklak, *Int. J. Mol. Sci.*, 2021, **22**, 4378.
19 M. Dračínský, H. M. Möller and T. E. Exner, *J. Chem. Theory Comput.*, 2013, **9**, 3806.
20 T. Pongratz, P. Kibies, L. Eberlein, N. Tielker, C. Hölzl, S. Imoto, M. Beck Erlach, S. Kurrmann, P. H. Schummel, M. Hofmann, O. Reiser, R. Winter, W. Kremer, H. R. Kalbitzer, D. Marx, D. Horinek and S. M. Kast, *Biophys. Chem.*, 2020, **257**, 106258.
21 T. Mineva, Y. Tsoneva, R. Kevorkyants and A. Goursot, *Can. J. Chem.*, 2013, **91**, 529.
22 J. M. Fonville, M. Swart, Z. Vokáčová, V. Sychrovský, J. E. Šponer, J. Šponer, C. W. Hilbers, F. M. Bickelhaupt and S. S. Wijmenga, *Chem. – Eur. J.*, 2012, **18**, 12372.
23 T. Stadelmann, C. Balmer, S. Riniker and M.-O. Ebert, *Phys. Chem. Chem. Phys.*, 2022, **24**, 23551.
24 A. M. Sarotti and S. C. Pellegrinet, *J. Org. Chem.*, 2009, **74**, 7254.
25 A. M. Sarotti and S. C. Pellegrinet, *J. Org. Chem.*, 2012, **77**, 6059.
26 H. D. Watts, M. N. A. Mohamed and J. D. Kubicki, *J. Phys. Chem. B*, 2011, **115**, 1958.
27 R. Frach, P. Kibies, S. Böttcher, T. Pongratz, S. Strohfeldt, S. Kurrmann, J. Koehler, M. Hofmann, W. Kremer, H. R. Kalbitzer, O. Reiser, D. Horinek and S. M. Kast, *Angew. Chem., Int. Ed.*, 2016, **55**, 8757.
28 M. Akiyama and H. Torii, *Spectrochim. Acta, Part A*, 2000, **56A**, 137.
29 R. Frach and S. M. Kast, *J. Phys. Chem. A*, 2014, **118**, 11620.
30 S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.
31 H. C. Da Silva and W. B. de Almeida, *Chem. Phys.*, 2020, **528**, 110479.
32 V. A. Semenov, D. O. Samultsev and L. B. Krivdin, *Magn. Reson. Chem.*, 2014, **52**, 686.
33 (*a*) M. Reimann and M. Kaupp, *J. Phys. Chem. A*, 2020, **124**, 7439; (*b*) K. Imamura, D. Yokogawa, M. Higashi and H. Sato, *J. Chem. Phys.*, 2022, **157**, 204105.
34 B. Sharma, V. A. Tran, T. Pongratz, L. Galazzo, I. Zhurko, E. Bordignon, S. M. Kast, F. Neese and D. Marx, *J. Chem. Theory Comput.*, 2021, **17**, 6366.
35 V. A. Tran, M. Teucher, L. Galazzo, B. Sharma, T. Pongratz, S. M. Kast, D. Marx, E. Bordignon, A. Schnegg and F. Neese, *J. Phys. Chem. A*, 2023, **31**, 6447.
36 R. K. Harris, E. D. Becker, S. M. Cabral de Menezes, P. Granger, R. E. Hoffman and K. W. Zilm, *Pure Appl. Chem.*, 2008, **80**, 59.
37 N. C. Handy and A. J. Cohen, *Mol. Phys.*, 2001, **99**, 403.
38 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B Condens. Matter Mater. Phys.*, 1988, **37**, 785.
39 W.-M. Hoe, A. J. Cohen and N. C. Handy, *Chem. Phys. Lett.*, 2001, **341**, 319.
40 Y. Y. Rusakov, I. L. Rusakova, V. A. Semenov, D. O. Samultsev, S. V. Fedorov and L. B. Krivdin, *J. Phys. Chem. A*, 2018, **122**, 6746.
41 Y. Zhang, A. Wu, X. Xu and Y. Yan, *Chem. Phys. Lett.*, 2006, **421**, 383.
42 E. Toomsalu and P. Burk, *J. Mol. Model.*, 2015, **21**, 244.
43 D. Flaig, M. Beer and C. Ochsenfeld, *J. Chem. Theory Comput.*, 2012, **8**, 2260.
44 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
45 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33.
46 S. Moon and D. A. Case, *J. Comput. Chem.*, 2006, **27**, 825.
47 D. Sitkoff and D. A. Case, *J. Am. Chem. Soc.*, 1997, **119**, 12262.
48 M. N. C. Zarycz and C. Fonseca Guerra, *J. Phys. Chem. Lett.*, 2018, **9**, 3720.
49 M. Hofmann, *PhD thesis*, University of Regensburg, Regensburg (Germany), 2018.
50 D. S. Raiford, C. L. Fisk and E. D. Becker, *Anal. Chem.*, 1979, **51**, 2050.
51 **https://auremol.de** (last access 2023/10/26).
52 J. L. Markley, A. Bax, Y. Arata, C. W. Hilbers, R. Kaptein, B. D. Sykes, P. E. Wright and K. Wüthrich, *Pure Appl. Chem.*, 1998, **70**, 117.
53 G. Bodenhausen and D. J. Ruben, *Chem. Phys. Lett.*, 1980, **69**, 185.
54 M. R. Willcott, *J. Am. Chem. Soc.*, 2009, **131**, 13180.
55 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19.
56 J. L. F. Abascal and C. Vega, *J. Chem. Phys.*, 2005, **123**, 234505.
57 L. S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives and W. L. Jorgensen, *Nucleic Acids Res.*, 2017, **45**, W331–W336.
58 M. J. Robertson, J. Tirado-Rives and W. L. Jorgensen, *J. Chem. Theory Comput.*, 2015, **11**, 3499.
59 W. L. Jorgensen and J. Tirado-Rives, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6665.
60 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696.
61 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577.

62 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463.
63 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
64 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
65 C. Hölzl, P. Kibies, S. Imoto, R. Frach, S. Suladze, R. Winter, D. Marx, D. Horinek and S. M. Kast, *J. Chem. Phys.*, 2016, **144**, 144104.
66 J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 15.
67 T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack and J. Hutter, *J. Chem. Phys.*, 2020, **152**, 194103.
68 J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comput. Phys. Commun.*, 2005, **167**, 103.
69 S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B Condens. Matter Mater. Phys.*, 1996, **54**, 1703.
70 C. Hartwigsen, S. Goedecker and J. Hutter, *Phys. Rev. B Condens. Matter Mater. Phys.*, 1998, **58**, 3641.
71 M. Krack, *Theor. Chem. Acc.*, 2005, **114**, 145.
72 B. Hammer, L. B. Hansen and J. K. Nørskov, *Phys. Rev. B Condens. Matter Mater. Phys.*, 1999, **59**, 7413.
73 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
74 G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635.
75 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.
76 R. Ditchfield, *Mol. Phys.*, 1974, **27**, 789.
77 S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.
78 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157.
79 Z. A. Makrodimitri, V. E. Raptis and I. G. Economou, *J. Phys. Chem. B*, 2006, **110**, 16047.
80 T. Kloss, J. Heil and S. M. Kast, *J. Phys. Chem. B*, 2008, **112**, 4337.
81 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269.
82 L. E. Chirlian and M. M. Francl, *J. Comput. Chem.*, 1987, **8**, 894.