Supporting Information to DOI https://doi.org/10.1039/D3CP05471C

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]

[a] Fakultät für Chemie und Chemische Biologie, Technische Universität Dortmund, Otto-Hahn-Straße 4a, 44227 Dortmund (Germany)

[b] Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum (Germany)

[c] Fakultät für Biologie und Vorklinische Medizin, Universität Regensburg, 93040 Regensburg (Germany)

*Corresponding authors (stefan.kast@tu-dortmund.de, dominik.marx@theochem.ruhr-uni-bochum.de)

Materials and methods

Materials

^{15}N-enriched trimethylamine-N-oxide (TMAO) was synthesized as described earlier.[1] The sample contained 0.4 M ^{15}N-enriched TMAO, 0.1 mM sodium salt of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), 0.1 mM dioxane in 90 % H_2O and 10 % D_2O, pH 7.2 (sample TMAO-1). ^{15}N-enriched and unlabeled N-methylacetamide (NMA) was synthesized by M. Hofmann (sample NMA-1).[2] This sample contained 0.4 mM ^{15}N enriched-NMA, 0.1 mM DSS, 0.1 mM dioxane dissolved in 90 % H_2O and 10 % D_2O, pH = 7.7. A second sample (NMA-2) contained 0.2 M ^{15}N-enriched NMA (Sigma Aldrich) and 0.1 % (m/m) of the sodium salt of DSS in 90 % H_2O and 10 % D_2O. For the study of the concentration and/or pH dependence of chemical shifts, unlabeled NMA (Sigma Aldrich), 0.1 mM DSS dissolved in 90 % H_2O and 10 % D_2O has been used (sample NMA-3).

NMR spectroscopy

The pH of the samples was measured with a Hamilton Spintrode attached to a Beckman Coulter pH meter. Temperature calibration was done via the difference of the resonance lines of the hydroxyl- and methyl-group protons in 100% methanol.[3] The measured pH values have not been corrected for the deuterium isotope effect. The NMR experiments were carried out on a Bruker Avance Neo 600 MHz NMR spectrometer equipped with a TXI (NMA-1) and BBO cold probe (NMA-3), respectively, and Bruker Avance III HD NanoBay 400 MHz NMR spectrometer equipped with a 5 mm room temperature broadband probe (BBFO Smart) (NMA-2), both at 298 K. Spectra were recorded and evaluated with the program TopSpin (Bruker, Karlsruhe). Proton resonances were directly referenced to the methyl resonance of DSS at 0 ppm. ^{13}C and ^{15}N spectra were indirectly referenced to DSS with IndirectRef[4] using Ξ values of 0.251449530 (DSS dissolved in D_2O) and 0.101329118 (liquid NH_3 in a capillary immersed in a DSS solution) in line with the IUPAC-IUB recommendations.[5] Alternatively, ^{13}C shifts were directly referenced to the ^{13}C methyl signal of DSS contained in the solution (see Table S1). ^{1}H chemical shifts were obtained from 1D spectra with presaturation of the HDO signal with spectral resolutions better than 0.0001 ppm/point. Heteronuclear chemical shifts were either obtained from directly detected 1D spectra with proton decoupling or from [^{1}H, ^{15}N] or [^{1}H, ^{13}C] sensitivity enhanced echo-antiecho HSQC-spectra[6] with a resolution of the indirect dimension better than 0.005 ppm/point for also detecting low-intensity *cis*-NMA signals not fully detectable in 1D spectra. In all these cases, the precession of the chemical shift values is limited by the width of the reference signal that is about 1 Hz corresponding to 0.0016 ppm at 600 MHz.

For NMA-2, 1D proton spectra were acquired with presaturation of the HDO signal using the zgpr pulse sequence, a 90° pulse, 16 steady-state scans, and 16 scans. The acquisition time was 2.56 s with a relaxation delay of 2 s using 32768 data points. ^{13}C data was acquired using power-gated proton decoupling with a 30° flip angle, 4 steady-state scans, 256 scans, and an acquisition time of 1.36 s with a relaxation delay of 2 s using 65536 data points. All chemical shifts were directly referenced to the internal DSS signal. The resulting spectra are plotted in Figure S1 (NMA-1) and S2 (NMA-2). While not all *cis*-NMA signals, especially for ^{13}C, could be resolved by 1D spectra, a special focus on the amide proton, also supported by the calculations, confirmed the chemical shifts detailed in Table S1. The data also revealed that additional signals present in the spectrum of the sample NMA-1 originated from impurities.

As the primary goal of the experimental data acquisition in this study was to provide benchmark chemical shifts for calculations where DSS was simulated independently of NMA or TMAO (however at identical environmental conditions), we consistently used only the data obtained from samples TMAO-1 and NMA-1 with very little DSS present for comparison with calculations.

Spectral assignments

Assignments of the TMAO ^{1}H NMR lines from the intensities of the chemical shifts and couplings was trivial. For the assignments of the NMA methyl resonances in the *trans* and *cis* configuration TOCSY and NOESY-spectra were recorded (NMA-1) that allowed the unequivocal assignment of the ^{1}H resonances. The proton assignments allowed the assignment of the corresponding ^{15}N and ^{13}C resonances in

the HSQC spectra recorded from the same samples. NMA is occurring in *trans* and *cis* configuration (see Figures S1 and S2) that are in slow exchange on the NMR time scale at 298 K as proved by exchange cross peaks in NOESY spectra.

Temperature, pH, and concentration dependence of chemical shifts

The precision of the experimental chemical shift values is mainly determined by the line widths and the digital resolution. However, for the comparison of the experimental with the calculated values the equivalence of the external parameters of experiment and calculations are more important that may influence the chemical shifts. Temperature, pressure, solvent composition, pH and possible interactions between target molecules (NMA, TMAO) themselves as well as the reference compound are known to influence the chemical shifts. In the calculations the temperature is 298.15 K, the pressure 0.1 MPa, the solvent is pure H_2O, and no interaction between target molecules or reference molecules is possible. In the experiments the temperature was 298 K ± 0.5 K, the pressure about 0.970 MPa, solvent 90 % H_2O, 10 % D_2O, pH 7.2 ± 0.1 (TMAO) or 7.7 ± 0.1 (NMA), the target molecule concentrations 0.4 ± 0.01 M, and the DSS concentrations 0.1 ± 0.002 mM. DSS may interact with the target molecules, and this interaction may influence the resonance frequencies of DSS itself as well as those of some groups of the target molecules. The first effect is stronger for small molar concentration ratios of DSS to target molecules, the latter effect for large ratios.

For TMAO the temperature effect is -0.0003, 0.003, and 0.0001 ppm/K for ^{1}H, ^{13}C, and ^{15}N, respectively. Increasing the DSS concentration to 10 mM leads to upfield shifts of -0.0003, -0.0007, and -0.015 ppm, for ^{1}H, ^{13}C, and ^{15}N, respectively. An estimate of the total uncertainties would be ±0.001, ±0.004, and ±0.02 ppm for ^{1}H, ^{13}C, and ^{15}N, respectively.

For NMA the largest temperature effect is observed for the amide protons with -0.008 and -0.02 ppm/K in *trans* and *cis* conformation, respectively. The largest ^{13}C temperature coefficient is observed for C-CH_3 (methyl 1) with 0.004 ppm/K. For ^{15}N it is -0.03 ppm/K. In the DSS concentration range between 0.1 and 10.1 mM the largest ^{1}H shift changes are observed for the amide proton. In *trans* and *cis* conformation they are -0.01 ppm and -0.008 ppm. The largest ^{13}C shifts are observed for the carbonyl group with 0.05 in *trans* and *cis* conformation, respectively. For ^{15}N the shifts are even larger with 0.09 ppm. In the pH range between 5.0 and 11 only very small ^{1}H chemical shift changes are observable: They are less than 0.001 ppm for the methyl protons and 0.01 ppm for the amide protons. Dissolution of the highly concentrated NMA sample (0.4 M) to 0.001 M leads only to small upfield shifts of 0.001 ppm and 0.01 ppm of the N-CH_3 (methyl-2) and the C-CH_3 (methyl-1) signals, respectively, and a downfield shift of 0.002 ppm for the NH signals. The direct isotope effect of the D_2O concentration on the chemical shifts by replacing the nitrogen-bound H by D is significant and influences all atoms. However, since separate resonances can be observed, it is of no relevance for this study. An estimate of the total uncertainties would be ±0.01, ±0.04, ±0.03, ±0.06, ±0.12 ppm for the methyl and amide protons, methyl and carbonyl carbons, and ^{15}N, respectively. In summary, for NMA the interaction with DSS appears to cause the largest experimental uncertainties.

Force field molecular dynamics simulations for ensemble generations

All force field molecular dynamics (FFMD) simulations, which also provided starting points for subsequent treatment by *ab initio* MD (AIMD), were carried out using the Gromacs 2016.3 software package.[7] DSS, the *cis* and *trans* conformation of NMA and TMAO were placed in a box of 152 water molecules described by the TIP4P/2005 water model.[8] For DSS, force field parameters were taken from the LigParGen webserver[9] using OPLS/AA[10] parameters with CM1A[11] charges while NMA was parametrized using the ff14SB[12] force field. Parameters for unlike atom pairs were described using Lorentz Berthelot mixing rules for NMA and TMAO and geometric mixing for DSS. The van der Waals interactions were described with a 6-12 Lennard-Jones potential, truncated at 7.7 Å, and the electrostatic interactions were calculated with the smooth particle-mesh Ewald summation[13] with a real-space cutoff of 7.7 Å and a lattice spacing of 1 Å. A time step of 1.5 fs was used during *NpT* simulations, while all bonds including hydrogen were constrained using LINCS.[14] A temperature of 298.15 K was maintained by the stochastic velocity-rescaling thermostat,[15] and a pressure of 1 bar was controlled by the isotropic Parrinello-Rahman barostat[16] using a compressibility of $4.5 \cdot 10^{-5}$ bar^{-1}.

7.5 ns of *NpT* simulations were performed where the last 6 ns were used to determine the density for further *NVT* simulations using FFMD and AIMD. A density of 1016.55 g cm^{-3} was determined for DSS corresponding to a cubic supercell of 16.8601 Å while a density of 998.074 g cm^{-3} was found for NMA corresponding to a cubic supercell of 16.7246 Å. The same density was used for both conformations of NMA. For TMAO a density of 1000.8 g cm^{-3} corresponding to a cubic supercell of 16.7134 Å was obtained. It is noteworthy that the FFMD simulation of TMAO was performed in line with the other FFMD simulations using 152 water molecules, while the previously published AIMD simulation,[17] which was reused here, was performed with 107 water molecules.

For the ensemble generation, *NVT* simulations were performed. Here the constraint of the hydrogen bonds was removed and the timestep was reduced to 1 fs, while otherwise same settings were used. *NVT* simulations were performed for 40 ns. Snapshots for NMR calculations were taken every 0.1 ns resulting in 400 snapshots for FFMD ensembles.

Ab initio molecular dynamics simulations for ensemble generations

From the initial structures obtained from FFMD simulations, AIMD simulations of DSS and both NMA conformations were launched at constant temperature (*NVT*) of 298.15 K using the CP2K package.[18,19] Each of the systems was placed in a cubic box using the dimensions determined from FFMD and was replicated in all three dimensions with periodic cluster boundary conditions. The electronic structure calculations were performed using the QUICKSTEP[20] module within CP2K. Since, QUICKSTEP module dwells on dual basis set approach[20] to describe the wave function with an atom-centered Gaussian basis while the electron density is described with an auxiliary plane-wave basis set. Thus, we employed the atom-centered TZV2P basis set with Goedecker-Teter-Hutter pseudopotential[21,22,23] and plane wave basis with kinetic energy cut-off of 500 Rydberg. The RPBE[24] density functional with Grimme's D3 dispersion correction[25] with zero damping was employed in all the AIMD simulations. During the simulations, each of the cartesian degrees of freedoms was individually thermostated at the given temperature by means of a massive Nose-Hoover chain thermostat.[26]

NMR calculations

For NMR calculations snapshots from AIMD were taken every 0.5 ps for DSS and NMA, while for TMAO every fourth snapshot of previously[17] preselected 1571 snapshots was taken resulting in 400 snapshots for DSS and NMA and 393 snapshots for TMAO. For FFMD, snapshots were taken every 0.1 ns yielding 400 snapshots for all molecules

All NMR calculations were performed at the OLYP/6-311+G(d,p)[27] level of theory with Gaussian16 rev.C01[28] using the Gauge-Independent Atomic Orbital (GIAO) method[29] to compute the nuclear magnetic shielding tensors, and using MP2 with the same basis set for selected calculations (see main text). For the calculation of chemical shifts, the difference between the isotropic shielding constants of DSS and the nucleus of interest was calculated, all averaged over magnetically equivalent nuclei. For calculations with the polarizable continuum model (PCM) the integral equation formalism (IEFPCM) was used with standard settings within Gaussian.[30]

All embedded cluster reference interaction site model (EC-RISM) calculations were performed on a cubic grid with 120^3 points with 0.3 Å spacing in all spatial directions. Convergence criteria were set to 10^{-6} for the maximum direct correlation function deviation in 3D RISM calculations and a maximum free energy difference of 0.01 kcal mol^{-1} between two EC-RISM cycles. Lennard Jones parameters for EC-RISM calculations were taken from TMAO-V3[17] while GAFF[31] parameters were used for NMA and DSS with the addition of silicon parameters from Makrodimitri *et al.*[32] following the previously described procedure for the DSS parametrization.[1] Explicit water molecules as well as EC-RISM water susceptibilities were modeled using modified SPC/E charges and Lennard-Jones parameters in line with previous computations.[33,34] Atom centered point charges were determined with the CHelpG scheme[35] employing default radii and used with the additional constraint to match the quantum chemically obtained dipole moment. Depending on the exact combination of explicit water molecules and background solvation models, a small number of calculations, mostly of DSS, could not be converged and were disregarded in further analysis, however converged shielding constants from at least 358 snapshots were used for every calculated average shielding constant.

All raw data entering the averaging (snapshot and optimized structures with corresponding shielding constants) are provided in machine-readable form at DOI https://doi.org/10.17877/RESOLV-2024-lrgkievk.

Results

NMR spectroscopy

Table S1. Chemical shifts [ppm] for the samples NMA-1 and TMAO-1 using direct and indirect referencing for ^{13}C. The methyl groups bound to C=O and NH are designated as 1 and 2, respectively. For experimental uncertainties see above.

Reference	NMA δ_{NH}		NMA $\delta_{H,met,1}$		NMA $\delta_{H,met,2}$		NMA $\delta_{C,met,1}$		NMA $\delta_{C,met,2}$		NMA $\delta_{Ccarbonyl}$		TMAO δ_H	TMAO δ_C
	trans	*cis*	*trans*	*cis*	*trans*	*cis*	*trans*	*cis*	*trans*	*cis*	*trans*	*cis*		
Direct	7.84	7.10	1.97	2.03	2.71	2.86	24.37	21.45	28.68	31.93	177.20	179.90	3.25	62.18
Indirect	-	-	-	-	-	-	24.44	21.52	28.74	32.00	177.30	180.00	-	62.27

Table S2. ^{15}N chemical shifts [ppm] for the samples NMA-1 and TMAO-1 using indirect referencing.

Reference	NMA δ_N		TMAO δ_N
	trans	*cis*	
Indirect	113.72	111.91	104.55

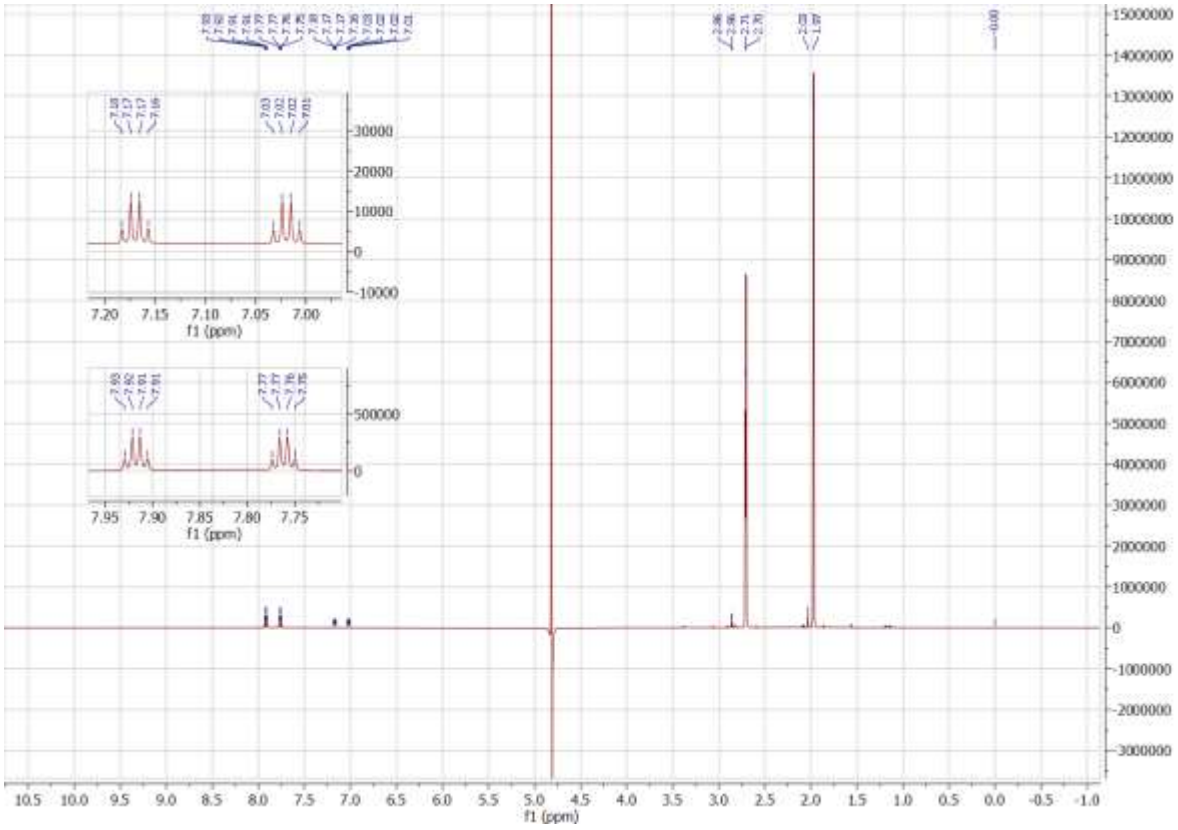

Figure S1. ^1H spectrum of NMA-1 with insets for the amide proton in *cis* (top) and *trans* (bottom) configuration. The spectrum was analyzed and plotted using Mnova.[36]

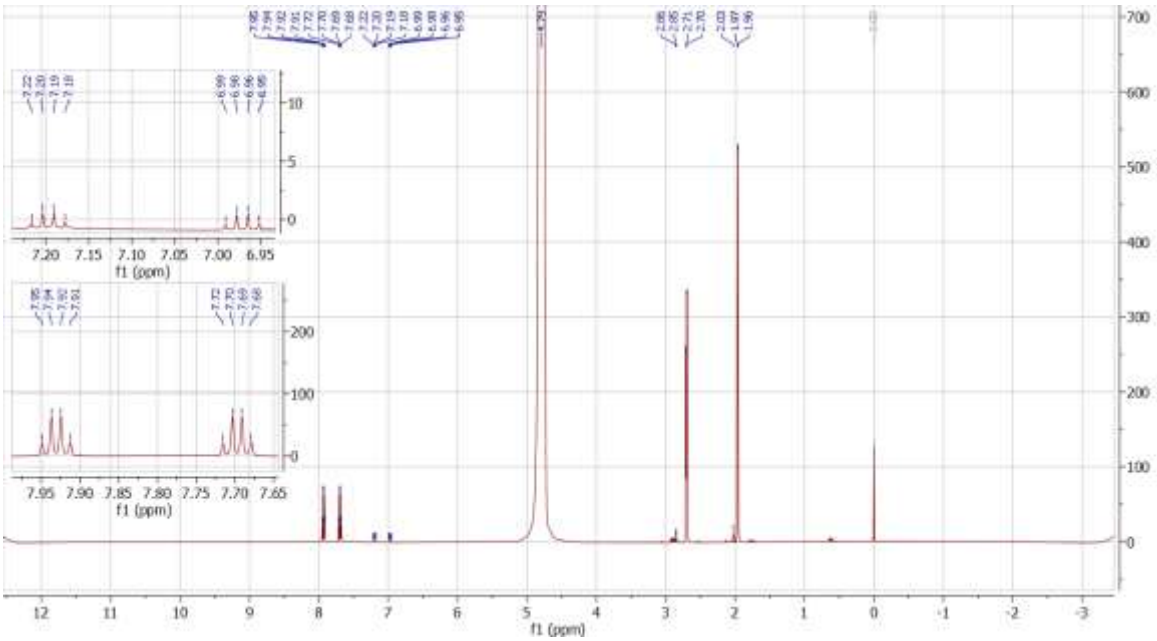

Figure S2. ^{1}H spectrum of NMA-2 with insets for the amide proton in *cis* (top) and *trans* (bottom) configuration. The spectrum was analyzed and plotted using Mnova.[36]

Shielding constants for all nuclei with increasing number of snapshots

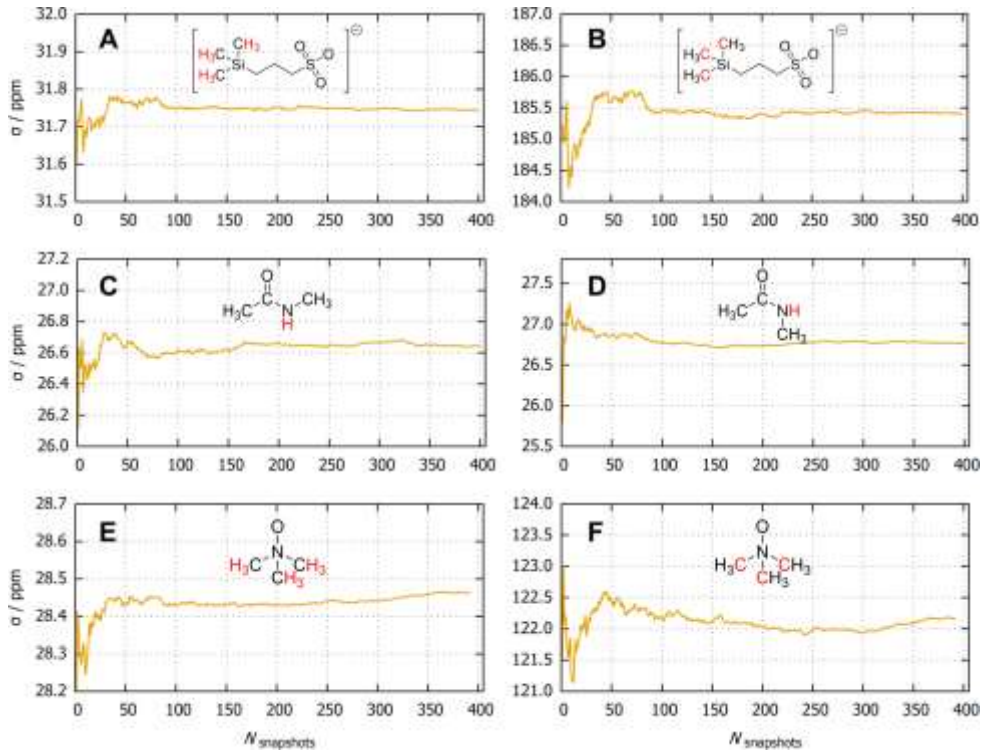

Figure S3. Cumulative average of the Isotropic shielding constants of the highlighted nuclei with increasing number of snapshots considered. The shielding constants were calculated as average over all magnetically equivalent nuclei per snapshot and afterwards averaged up to $N_{snapshots}$. Neither explicit short-range solvation, nor any background solvation model were used for these calculations.

Shielding constants and chemical shifts for all methods and analyzed nuclei

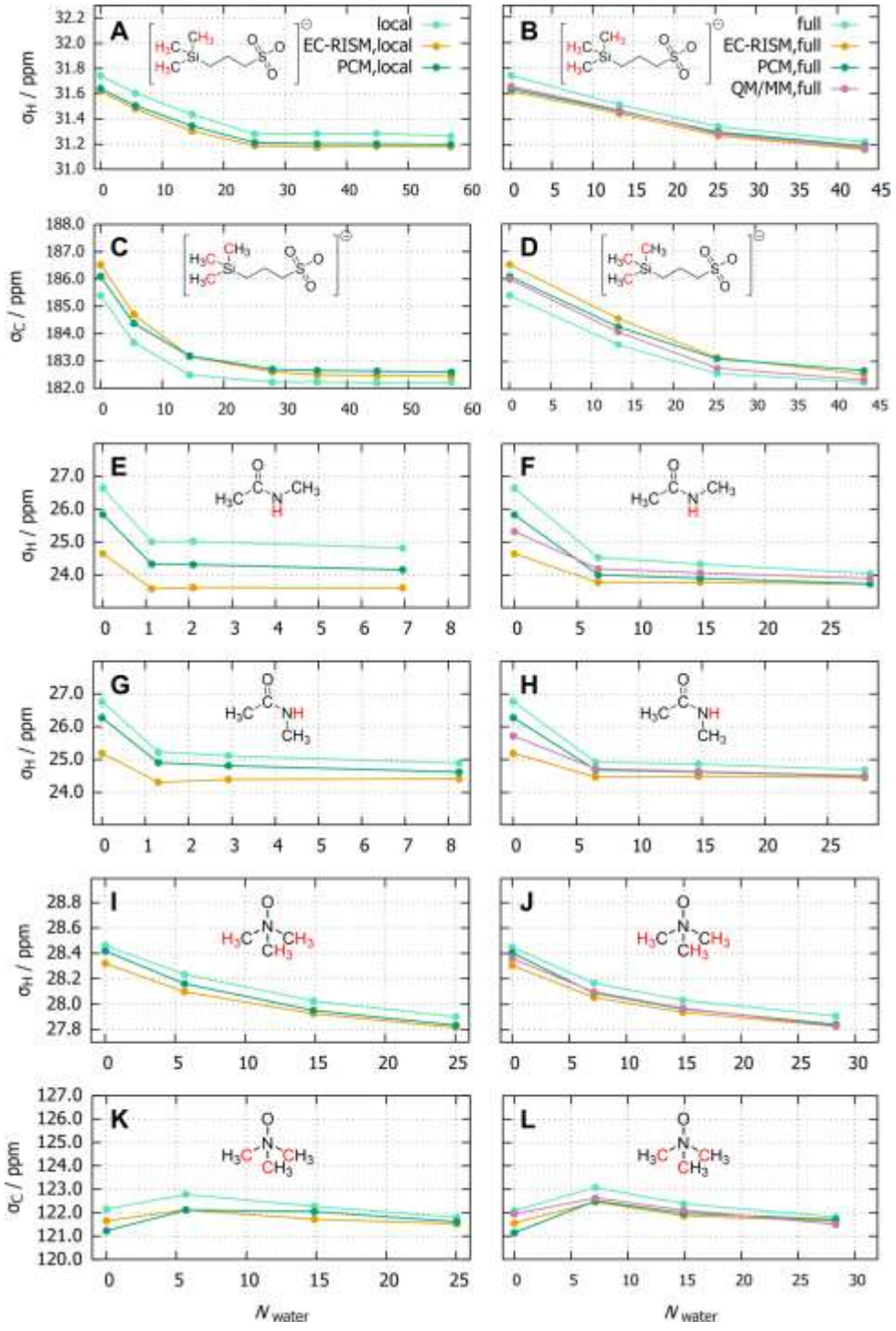

Figure S4. Isotropic shielding constants on the OLYP/6-311+G(d,p) level of theory of the highlighted nuclei with increasing number of explicit water molecules included in the NMR calculations. The "local" solvation (**A**, **C**, **E**, **G**, **I**, **K**) denotes that the water molecules were selected locally around the highlighted nuclei while "full" (**B**, **D**, **F**, **H**, **J**, **K**) denotes that the water molecules were included around the whole molecule. Shielding constants were averaged over snapshots and all magnetically equivalent nuclei. The EC-RISM and PCM subscripts indicate that an additional solvent background modelled with EC-RISM and PCM was used while "QM/MM" denotes that additional TIP3P-point charges [37] were placed on the water molecules outside of the explicit QM zone. The "$ref_{QM/MM}$" value can be identified as the last point of "QM/MM,full" curve.

Table S3. ^{1}H and ^{13}C shielding constants [ppm] of DSS with "local" explicit solvent molecules and additional PCM or EC-RISM background for the AIMD ensemble. The average number of water molecules corresponding to the applied distance criterion is given as well. Errors were calculated as the standard error over all snapshots.

Explicit solvent				
r_{water}	N_{water}	Solvent background	σ_H	σ_C
-	0	-	31.74 ± 0.01	185.40 ± 0.13
2.5 Å	5.4	-	31.60 ± 0.02	183.68 ± 0.14
3.0 Å	14.5	-	31.44 ± 0.02	182.49 ± 0.18
4.0 Å	27.9	-	31.28 ± 0.02	182.22 ± 0.13
4.5 Å	35.2		31.28 ± 0.03	182.23 ± 0.14
5.0 Å	44.9		31.28 ± 0.02	182.20 ± 0.13
5.5 Å	56.9		31.26 ± 0.01	182.22 ± 0.13
-	0	PCM	31.64 ± 0.01	186.10 ± 0.12
2.5 Å	5.4	PCM	31.50 ± 0.01	184.37 ± 0.13
3.0 Å	14.5	PCM	31.34 ± 0.01	183.18 ± 0.13
4.0 Å	27.9	PCM	31.21 ± 0.01	182.70 ± 0.13
4.5 Å	35.2	PCM	31.20 ± 0.01	182.65 ± 0.13
5.0 Å	44.9	PCM	31.20 ± 0.01	182.63 ± 0.13
5.5 Å	56.9	PCM	31.20 ± 0.01	182.60 ± 0.13
-	0	EC-RISM	31.62 ± 0.01	186.53 ± 0.14
2.5 Å	5.4	EC-RISM	31.48 ± 0.01	184.71 ± 0.14
3.0 Å	14.5	EC-RISM	31.30 ± 0.01	183.17 ± 0.14
4.0 Å	27.9	EC-RISM	31.19 ± 0.01	182.60 ± 0.13
4.5 Å	35.2	EC-RISM	31.18 ± 0.01	182.49 ± 0.13
5.0 Å	44.9	EC-RISM	31.18 ± 0.01	182.46 ± 0.13
5.5 Å	56.9	EC-RISM	31.18 ± 0.01	182.45 ± 0.01

Table S4. ^{1}H and ^{13}C shielding constants [ppm] of DSS with "full" explicit solvent molecules and additional PCM or EC-RISM background for the AIMD ensembles. The average number of water molecules corresponding to the applied distance criterion is given as well. Errors were calculated as the standard error over all snapshots.

Explicit solvent				
r_{water}	N_{water}	Solvent background	σ_H	σ_C
-	0	-	31.74 ± 0.01	185.40 ± 0.13
2.5 Å	13.3	-	31.51 ± 0.01	183.62 ± 0.10
3.0 Å	25.4	-	31.34 ± 0.01	182.57 ± 0.10
4.0 Å	43.3	-	31.22 ± 0.01	182.24 ± 0.09
-	0	PCM	31.64 ± 0.01	186.10 ± 0.12
2.5 Å	13.3	PCM	31.46 ± 0.01	184.26 ± 0.13
3.0 Å	25.4	PCM	31.30 ± 0.01	183.10 ± 0.13
4.0 Å	43.3	PCM	31.19 ± 0.01	182.67 ± 0.13
-	0	EC-RISM	31.62 ± 0.01	186.53 ± 0.12
2.5 Å	13.3	EC-RISM	31.45 ± 0.01	184.57 ± 0.13
3.0 Å	25.4	EC-RISM	31.27 ± 0.01	183.15 ± 0.13
4.0 Å	43.3	EC-RISM	31.16 ± 0.01	182.54 ± 0.13
-	0	QM/MM	31.66 ± 0.01	185.99 ± 0.13
2.5 Å	13.3	QM/MM	31.47 ± 0.01	184.08 ± 0.13
3.0 Å	25.4	QM/MM	31.29 ± 0.01	182.77 ± 0.13
4.0 Å	43.3	QM/MM	31.18 ± 0.01	182.33 ± 0.13

Table S5. Amide ^{1}H shielding constants [ppm] of NMA in the *trans* conformation with "local" explicit solvent molecules and additional PCM or EC-RISM background for the AIMD ensembles. The average number of water molecules corresponding to the applied distance criterion is given as well. Errors were calculated as the standard error over all snapshots.

Explicit solvent			
r_{water}	N_{water}	Solvent background	σ_H
-	0	-	26.64 ± 0.04
2.5 Å	1.1	-	25.01 ± 0.05
3.0 Å	2.1	-	25.02 ± 0.06
4.0 Å	7.0	-	24.82 ± 0.07
-	0	PCM	25.84 ± 0.04
2.5 Å	1.1	PCM	24.33 ± 0.07
3.0 Å	2.1	PCM	24.32 ± 0.07
4.0 Å	7.0	PCM	24.16 ± 0.07
-	0	EC-RISM	24.65 ± 0.04
2.5 Å	1.1	EC-RISM	23.58 ± 0.07
3.0 Å	2.1	EC-RISM	23.62 ± 0.07
4.0 Å	7.0	EC-RISM	23.61 ± 0.07

Table S6. Amide ^{1}H shielding constants [ppm] of NMA in the *trans* conformation with "full" explicit solvent molecules and additional PCM, EC-RISM or QM/MM background for the AIMD ensembles. The average number of water molecules corresponding to the applied distance criterion is given as well. Errors were calculated as the standard error over all snapshots.

Explicit solvent			
r_{water}	N_{water}	Solvent background	σ_H
-	0	-	26.64 ± 0.04
2.5 Å	6.7	-	24.53 ± 0.07
3.0 Å	14.8	-	24.34 ± 0.07
4.0 Å	28.3	-	24.05 ± 0.07
-	0	PCM	25.84 ± 0.04
2.5 Å	6.7	PCM	24.01 ± 0.07
3.0 Å	14.8	PCM	23.90 ± 0.07
4.0 Å	28.3	PCM	23.73 ± 0.07
-	0	EC-RISM	24.65 ± 0.04
2.5 Å	6.7	EC-RISM	23.78 ± 0.07
3.0 Å	14.8	EC-RISM	23.78 ± 0.07
4.0 Å	28.3	EC-RISM	23.73 ± 0.07
-	0	QM/MM	25.32 ± 0.05
2.5 Å	6.7	QM/MM	24.19 ± 0.07
3.0 Å	14.8	QM/MM	24.07 ± 0.07
4.0 Å	28.3	QM/MM	23.90 ± 0.07

Table S7. Amide ^{1}H shielding constants [ppm] of NMA in the *cis* conformation with "local" explicit solvent molecules and additional PCM or EC-RISM background for the AIMD ensembles. The average number of water molecules corresponding to the applied distance criterion is given as well. Errors were calculated as the standard error over all snapshots.

Explicit solvent			
r_{water}	N_{water}	Solvent background	σ_H
-	0	-	26.77 ± 0.03
2.5 Å	1.3	-	25.23 ± 0.05
3.0 Å	2.9	-	25.11 ± 0.05
4.0 Å	8.2	-	24.89 ± 0.05
-	0	PCM	26.27 ± 0.04
2.5 Å	1.3	PCM	24.90 ± 0.07
3.0 Å	2.9	PCM	24.81 ± 0.07
4.0 Å	8.2	PCM	24.62 ± 0.07
-	0	EC-RISM	25.19 ± 0.04
2.5 Å	1.3	EC-RISM	24.31 ± 0.07
3.0 Å	2.9	EC-RISM	24.39 ± 0.07
4.0 Å	8.2	EC-RISM	24.41 ± 0.07

Table S8. Amide ^{1}H shielding constants [ppm] of NMA in the *cis* conformation with "full" explicit solvent molecules and additional PCM, EC-RISM or QM/MM background for the AIMD ensembles. The average number of water molecules corresponding to the applied distance criterion is given as well. Errors were calculated as the standard error over all snapshots.

Explicit solvent			
r_{water}	N_{water}	Solvent background	σ_H
-	0	-	26.77 ± 0.04
2.5 Å	6.5	-	24.92 ± 0.08
3.0 Å	14.7	-	24.86 ± 0.07
4.0 Å	27.9	-	24.69 ± 0.08
-	0	PCM	26.27 ± 0.02
2.5 Å	6.5	PCM	24.68 ± 0.07
3.0 Å	14.7	PCM	24.62 ± 0.07
4.0 Å	27.9	PCM	24.49 ± 0.07
-	0	EC-RISM	25.19 ± 0.04
2.5 Å	6.5	EC-RISM	24.46 ± 0.07
3.0 Å	14.7	EC-RISM	24.48 ± 0.07
4.0 Å	27.9	EC-RISM	24.45 ± 0.08
-	0	QM/MM	25.72 ± 0.05
2.5 Å	6.5	QM/MM	24.72 ± 0.08
3.0 Å	14.7	QM/MM	24.64 ± 0.07
4.0 Å	27.9	QM/MM	24.51 ± 0.08

Table S9. ^{1}H and ^{13}C shielding constants [ppm] of TMAO with "local" explicit solvent molecules and additional PCM or EC-RISM background. The average number of water molecules corresponding to the applied distance criterion is given as well. Errors were calculated as the standard error over all snapshots.

Explicit solvent				
r_{water}	N_{water}	Solvent background	σ_H	σ_C
-	0	-	28.46 ± 0.01	122.15 ± 0.16
2.5 Å	5.7	-	28.24 ± 0.02	122.78 ± 0.19
3.0 Å	14.9	-	28.02 ± 0.02	122.28 ± 0.17
4.0 Å	25.0	-	27.90 ± 0.01	121.80 ± 0.16
-	0	PCM	28.42 ± 0.01	121.23 ± 0.17
2.5 Å	5.7	PCM	28.16 ± 0.02	122.12 ± 0.17
3.0 Å	14.9	PCM	27.95 ± 0.02	122.05 ± 0.22
4.0 Å	25.0	PCM	27.83 ± 0.01	121.61 ± 0.16
-	0	EC-RISM	28.32 ± 0.01	121.66 ± 0.17
2.5 Å	5.7	EC-RISM	28.10 ± 0.01	122.12 ± 0.17
3.0 Å	14.9	EC-RISM	27.93 ± 0.01	121.75 ± 0.17
4.0 Å	25.0	EC-RISM	27.82 ± 0.01	121.53 ± 0.16

Table S10. ^{1}H and ^{13}C shielding constants [ppm] of TMAO with "full" explicit solvent molecules and additional PCM or EC-RISM background. The average number of water molecules corresponding to the applied distance criterion is given as well. Errors were calculated as the standard error over all snapshots.

Explicit solvent				
r_{water}	N_{water}	Solvent background	σ_H	σ_C
-	-	-	28.45 ± 0.01	122.09 ± 0.16
2.5 Å	7.1 H_2O	-	28.17 ± 0.02	123.07 ± 0.17
3.0 Å	14.9 H_2O	-	28.03 ± 0.01	122.38 ± 0.17
4.0 Å	28.3 H_2O	-	27.91 ± 0.01	121.81 ± 0.16
-	-	PCM	28.40 ± 0.01	121.14 ± 0.17
2.5 Å	7.1 H_2O	PCM	28.09 ± 0.01	122.50 ± 0.17
3.0 Å	14.9 H_2O	PCM	27.96 ± 0.01	121.98 ± 0.16
4.0 Å	28.3 H_2O	PCM	27.84 ± 0.01	121.70 ± 0.16
-	-	EC-RISM	28.31 ± 0.01	121.56 ± 0.17
2.5 Å	7.1 H_2O	EC-RISM	28.05 ± 0.01	122.46 ± 0.17
3.0 Å	14.9 H_2O	EC-RISM	27.94 ± 0.01	121.87 ± 0.17
4.0 Å	28.3 H_2O	EC-RISM	27.83 ± 0.01	121.64 ± 0.16
-	-	QM/MM	28.36 ± 0.01	121.94 ± 0.17
2.5 Å	7.1 H_2O	QM/MM	28.10 ± 0.02	122.62 ± 0.17
3.0 Å	14.9 H_2O	QM/MM	27.97 ± 0.01	122.11 ± 0.16
4.0 Å	28.3 H_2O	QM/MM	27.83 ± 0.02	121.50 ± 0.17

Table S11. Comparison of shielding constants [ppm] for the nuclei featured in the main text derived from AIMD- and FFMD-ensembles with "local" and "full" explicit solvation augmented by an EC-RISM- or PCM-background.

Nucleus	$\sigma^{AIMD}_{ECR,local}$	$\sigma^{FFMD}_{ECR,local}$	$\sigma^{AIMD}_{PCM,local}$	$\sigma^{FFMD}_{PCM,local}$	$\sigma^{AIMD}_{ECR,full}$	$\sigma^{FFMD}_{ECR,full}$	$\sigma^{AIMD}_{PCM,full}$	$\sigma^{FFMD}_{PCM,full}$
DSS_H	31.48 ± 0.01	31.93 ± 0.01	31.50 ± 0.01	31.95 ± 0.01	31.16 ± 0.01	31.60 ± 0.01	31.19 ± 0.01	31.63 ± 0.01
DSS_C	184.71 ± 0.13	187.96 ± 0.12	184.37 ± 0.13	187.50 ± 0.13	182.54 ± 0.13	185.65 ± 0.13	182.67 ± 0.13	185.75 ± 0.12
trans-NMA_{NH}	23.58 ± 0.07	24.94 ± 0.06	24.33 ± 0.07	25.72 ± 0.06	23.73 ± 0.07	25.00 ± 0.06	23.73 ± 0.07	25.05 ± 0.01
cis-NMA_{NH}	24.31 ± 0.07	25.25 ± 0.06	24.90 ± 0.07	25.92 ± 0.06	24.45 ± 0.08	25.58 ± 0.07	24.49 ± 0.07	25.66 ± 0.06
$TMAO_H$	28.10 ± 0.01	28.58 ± 0.02	28.16 ± 0.01	28.66 ± 0.02	27.83 ± 0.02	28.34 ± 0.02	27.84 ± 0.01	28.38 ± 0.02
$TMAO_C$	122.12 ± 0.17	127.05 ± 0.25	122.12 ± 0.17	127.04 ± 0.22	121.64 ± 0.16	127.04 ± 0.24	121.70 ± 0.16	126.80 ± 0.53

Table S12. Shielding constants [ppm] for the further nuclei of NMA.

Nucleus	$\sigma^{AIMD}_{QM/MM,full}$	$\Delta\sigma^{AIMD}_{QM/MM,0}$	$\Delta\sigma^{AIMD}_{ECR,full}$	$\Delta\sigma^{AIMD}_{ECR,0}$	$\Delta\sigma^{opt}_{ECR,0}$	$\Delta\sigma^{FFMD}_{ECR,full}$	$\Delta\sigma^{AIMD}_{PCM,full}$	$\Delta\sigma^{AIMD}_{PCM,0}$	$\Delta\sigma^{opt}_{PCM,0}$	$\Delta\sigma^{FFMD}_{PCM,full}$
trans-$NMA_{Hmet,1}$	29.14	0.45	-0.02	0.33	0.67	0.42	-0.01	0.51	0.74	0.45
trans-$NMA_{Hmet,2}$	28.42	0.31	-0.01	0.26	0.58	0.21	0.00	0.32	0.55	0.24
trans-$NMA_{Cmet,1}$	157.62	3.17	-0.16	2.32	4.22	1.33	-0.07	3.37	4.35	1.52
trans-$NMA_{Cmet,2}$	152.07	2.84	0.00	2.49	4.85	2.24	0.07	2.91	4.26	2.29
trans-$NMA_{Ccarbonyl}$	11.84	0.16	-0.68	-3.16	1.44	6.25	-0.57	2.03	5.32	6.66
cis-$NMA_{Hmet,1}$	29.14	0.42	-0.02	0.27	0.47	0.38	-0.01	0.45	0.64	0.41
cis-$NMA_{Hmet,2}$	28.27	0.34	-0.06	0.21	0.46	0.23	-0.04	0.33	0.58	0.26
cis-$NMA_{Cmet,1}$	160.96	3.61	-0.25	2.64	2.03	0.40	-0.17	3.76	3.06	0.58
cis-$NMA_{Cmet,2}$	149.23	2.15	-0.24	1.27	3.90	2.04	-0.14	1.94	4.12	2.11
cis-$NMA_{Ccarbonyl}$	9.68	0.62	-1.06	-3.42	-0.91	6.88	-0.72	2.20	5.44	7.48

Table S13. Chemical shift differences [ppm] compared to experimental values for the further nuclei of NMA.

Nucleus	δ_{exp}	$\Delta\delta^{AIMD}_{QM/MM,full}$	$\Delta\delta^{AIMD}_{QM/MM,0}$	$\Delta\delta^{AIMD}_{ECR,full}$	$\Delta\delta^{AIMD}_{ECR,0}$	$\Delta\delta^{opt}_{ECR,0}$	$\Delta\delta^{FFMD}_{ECR,full}$	$\Delta\delta^{AIMD}_{PCM,full}$	$\Delta\delta^{AIMD}_{PCM,0}$	$\Delta\delta^{opt}_{PCM,0}$	$\Delta\delta^{FFMD}_{PCM,full}$
trans-$NMA_{Hmet,1}$	1.97	0.08	0.10	0.07	0.19	-0.03	0.08	0.09	0.03	-0.07	0.08
trans-$NMA_{Hmet,2}$	2.71	0.05	0.22	0.04	0.24	0.03	0.26	0.06	0.19	0.09	0.26
trans-$NMA_{Cmet,1}$	24.37	0.34	0.83	0.70	2.22	1.21	2.33	0.75	0.74	0.79	2.24
trans-$NMA_{Cmet,2}$	28.68	1.58	2.40	1.79	3.29	1.82	2.66	1.85	2.44	2.11	2.71
trans-$NMA_{Ccarbonyl}$	177.20	-6.71	-3.21	-5.79	0.63	-3.06	-9.64	-5.79	-4.97	-7.23	-9.94
cis-$NMA_{Hmet,1}$	2.03	0.01	0.06	0.01	0.17	0.09	0.04	0.03	0.02	-0.04	0.05
cis-$NMA_{Hmet,2}$	2.86	0.05	0.19	0.09	0.28	0.15	0.23	0.10	0.18	0.06	0.24
cis-$NMA_{Cmet,1}$	21.45	-0.08	-0.03	0.38	1.48	2.98	2.84	0.43	-0.07	1.66	2.76
cis-$NMA_{Cmet,2}$	31.93	1.17	2.68	1.62	4.10	2.36	2.45	1.65	3.00	1.85	2.48
cis-$NMA_{Ccarbonyl}$	179.90	-7.25	-4.21	-5.98	0.36	-1.26	-10.82	-6.19	-5.68	-7.89	-11.31

Table S14. Shielding constants and chemical shift differences [ppm] compared to experimental values for all analyzed nuclei derived from AIMD-ensembles using the MP2/6-311+G(d,p) level of theory in the NMR calculation with the "PCM,0" solvation method and extrapolated (additional subscript "extrap.") to a hypothetical "PCM,full" solvation. The extrapolation took the difference between "PCM,full" and "PCM,0" on the OLYP level of theory in addition to MP2 "PCM,0" shielding constants.

Nucleus	δ_{exp}	$\Delta\delta^{AIMD}_{PCM,0,MP2}$	$\Delta\delta^{AIMD}_{PCM,MP2,full\ extrap.}$	$\sigma^{AIMD}_{PCM,0,MP2}$	$\sigma^{AIMD}_{PCM,MP2,full\ extrap.}$
DSS_H	-	-	-	31.68	31.23
DSS_C	-	-	-	198.80	195.37
cis-NMA_{NH}	7.10	-1.74	-0.41	26.33	24.55
trans-NMA_{NH}	7.84	-2.27	-0.61	26.11	24.00
$TMAO_H$	3.25	-0.15	-0.02	28.58	28.00
$TMAO_C$	62.18	5.15	1.25	131.47	131.94
trans-$NMA_{Hmet,1}$	1.97	-0.04	0.03	29.75	29.24
trans-$NMA_{Hmet,2}$	2.71	0.16	0.03	28.81	28.49
trans-$NMA_{Cmet,1}$	24.37	2.41	2.41	172.02	168.59
trans-$NMA_{Cmet,2}$	28.68	3.10	2.50	167.02	164.19
trans-$NMA_{Ccarbonyl}$	177.20	-0.15	-0.97	21.75	19.14
cis-$NMA_{Hmet,1}$	2.03	-0.02	-0.01	29.68	29.22
cis-$NMA_{Hmet,2}$	2.86	0.11	0.03	28.71	28.34
cis-$NMA_{Cmet,1}$	21.45	1.43	1.94	175.92	171.98
cis-$NMA_{Cmet,2}$	31.93	3.30	1.95	163.57	161.49
cis-$NMA_{Ccarbonyl}$	179.90	-0.24	-0.76	19.14	16.23

Structural analysis

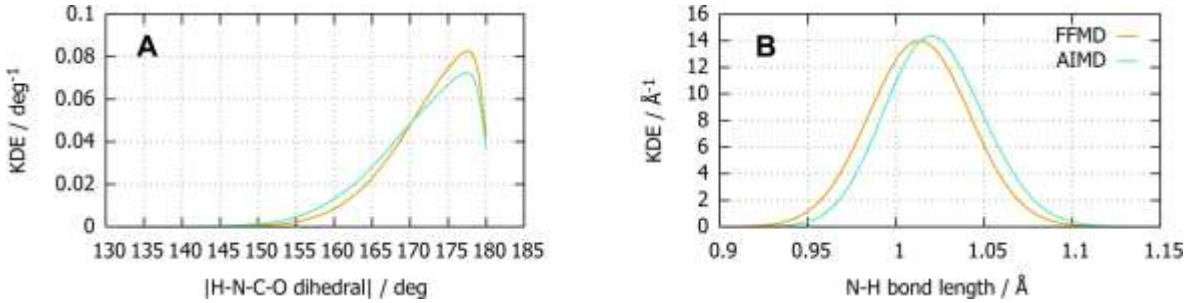

Figure S5. Kernel density estimate (KDE) of the absolute H-N-C-O dihedral angle (**A**) and the N-H bond length (**B**) for NMA in the *trans* conformation simulated by FFMD (orange) and AIMD (blue). For the analysis, 300000 snapshots of either simulation were considered. For the KDE a bandwidth of 1 deg and 0.01 Å was used, yielding average values of 171.52 deg and 1.022 Å for AIMD and 172.64 deg and 1.013 Å for FFMD.

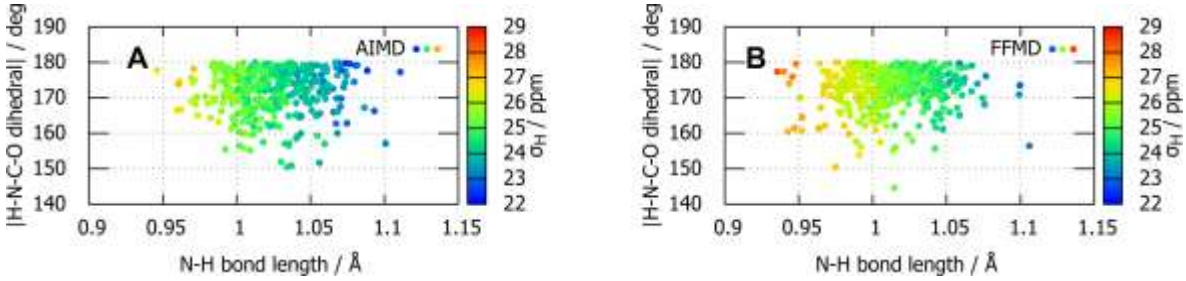

Figure S6. Heatmap of the amide shielding constants of *trans*-NMA in dependence of the absolute H-N-C-O dihedral angle and the N-H bond length for the snapshots from AIMD (**A**) and FFMD (**B**) used in the NMR calculations. Here the "EC-RISM,0" solvation method was employed, yielding average shielding constants of 24.65 ppm (AIMD) and 25.41 ppm (FFMD).

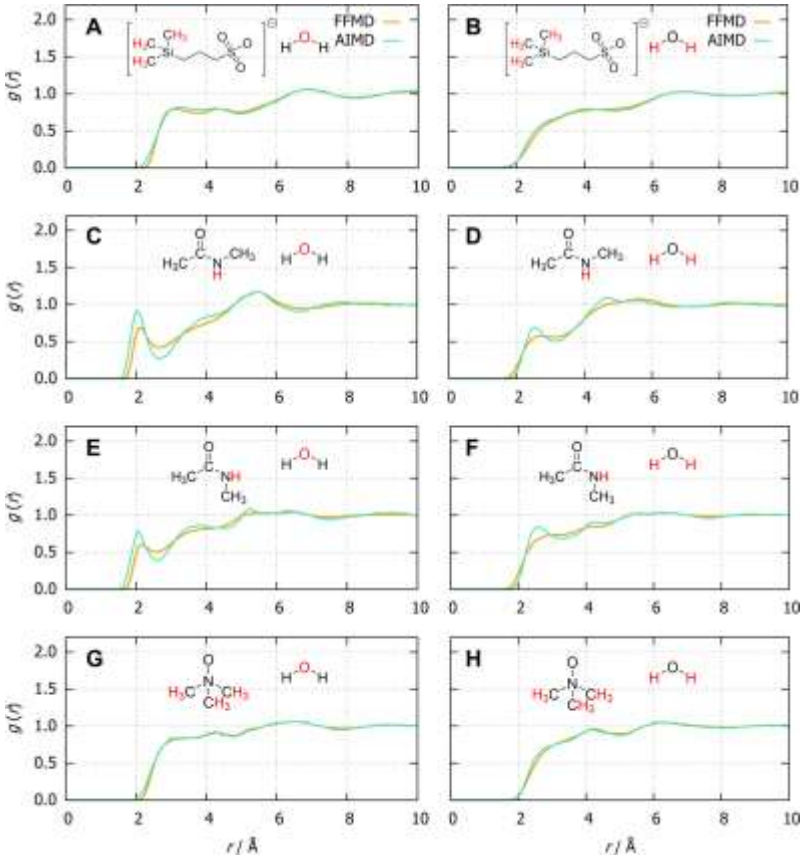

Figure S7. Radial distribution functions of the highlighted water nuclei surrounding the highlighted solute nuclei for which shielding constants have been calculated. RDFs from FFMD (orange) and AIMD (blue) are compared.

Supplementary references

[1] 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, S. M. Kast, *Biophys. Chem.* 2020, **257**, 106258.
[2] M. Hofmann, PhD thesis, University of Regensburg, Regensburg (Germany), 2018.
[3] D. S. Raiford, C. L. Fisk, E. D. Becker, *Anal. Chem.* 1979, **51**, 2050.
[4] https://auremol.de (last access 2023/10/26).
[5] J. L. Markley, A. Bax, Y. Arata, C. W. Hilbers, R. Kaptein, B. D. Sykes, P. E. Wright, K. Wüthrich, *Pure Appl. Chem.* 1998, **70**, 117.
[6] G. Bodenhausen, D. J. Ruben, *Chem. Phys. Lett.* 1980, **69**, 185.
[7] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* 2015, **1-2**, 19.
[8] J. L. F. Abascal, C. Vega, *J. Chem. Phys.* 2005, **123**, 234505.
[9] L. S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives, W. L. Jorgensen, *Nucleic Acids Res.* 2017, **45**, W331-W336.
[10] M. J. Robertson, J. Tirado-Rives, W. L. Jorgensen, *J. Chem. Theory Comput.* 2015, **11**, 3499.
[11] W. L. Jorgensen, J. Tirado-Rives, *Proc. Natl. Acad. Sci. USA* 2005, **102**, 6665.
[12] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* 2015, **11**, 3696.
[13] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* 1995, **103**, 8577.
[14] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* 1997, **18**, 1463.
[15] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* 2007, **126**, 014101.
[16] M. Parrinello, A. Rahman, *J. Appl. Phys.* 1981, **52**, 7182-7190.
[17] C. Hölzl, P. Kibies, S. Imoto, R. Frach, S. Suladze, R. Winter, D. Marx, D. Horinek, S. M. Kast, *J. Chem. Phys.* 2016, **144**, 144104.
[18] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, *WIREs Comput. Mol. Sci.* 2014, **4**, 15.
[19] 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, J. Hutter, *J. Chem. Phys.* 2020, **152**, 194103.
[20] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, *Computer Physics Commun.* 2005, **167**, 103.
[21] S. Goedecker, M. Teter, J. Hutter, *Phys. Rev. B Condens. Matter* 1996, **54**, 1703.
[22] C. Hartwigsen, S. Goedecker, J. Hutter, *Phys. Rev. B Condens. Matter* 1998, **58**, 3641.
[23] M. Krack, *Theor. Chem. Acc.* 2005, **114**, 145.
[24] B. Hammer, L. B. Hansen, J. K. Nørskov, *Phys. Rev. B Condens. Matter* 1999, **59**, 7413.
[25] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* 2010, **132**, 154104.
[26] G. J. Martyna, M. L. Klein, M. Tuckerman, *J. Chem. Phys.* 1992, **97**, 2635.
[27] N. C. Handy, A. J. Cohen, *Mol. Phys.* 2001, **99**, 403.
[28] 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, Williams, 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, D. J. Fox, *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.
[29] R. Ditchfield, *Mol. Phys.* 1974, **27**, 789.
[30] S. Miertuš, E. Scrocco, J. Tomasi, *Chem. Phys.* 1981, **55**, 117.
[31] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* 2004, **25**, 1157.
[32] Z. A. Makrodimitri, V. E. Raptis, I. G. Economou, *J. Phys. Chem. B* 2006, **110**, 16047.
[33] T. Kloss, J. Heil, S. M. Kast, *J. Phys. Chem. B* 2008, **112**, 4337.
[34] H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.* 1987, **91**, 6269.
[35] L. E. Chirlian, M. M. Francl, *J. Comput. Chem.* 1987, **8**, 894.
[36] M. R. Willcott, *J. Am. Chem. Soc.* 2009, **131**, 13180.
[37] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* 1983, **79**, 926.