Assessing the validity of NMR relaxation rates obtained from coarse-grained simulations of PEG-water mixtures

Simon Gravelle,† David Beyer,† Mariano Brito,† Alexander Schlaich,†‡ and Christian Holm*,†

†Institute for Computational Physics, University of Stuttgart, D-70569 Stuttgart, Germany
‡Stuttgart Center for Simulation Science (SC SimTech), University of Stuttgart, D-70569 Stuttgart, Germany
E-mail: holm@icp.uni-stuttgart.de

Abstract

NMR relaxometry is a powerful and well-established experimental approach to characterize dynamic processes in soft matter systems. All-atom (AA) resolved simulations are typically employed to gain further microscopic insights while reproducing the relaxation rates $R_1$. However, such approaches are limited to time and length-scales that hinder modeling of systems like long polymer chains or hydrogels. Coarse-graining (CG) can overcome this barrier at the cost of loosing atomistic details that impede the calculation of NMR relaxation rates. Here, we address this issue by systematic characterization of dipolar relaxation rates $R_1$ while performing systematic measurements on a PEG-H$_2$O mixture at two different levels of details: AA and CG. Remarkably, we show that NMR relaxation rates $R_1$ obtained at the CG level obey the same trends when compared to AA calculations, but with a systematic offset. This offset is due to, on the one hand, the lack of an intra-monomer component and, on the other hand, the inexact positioning of the spin carriers. We show that the offset can be corrected for quantitatively by reconstructing a posteriori the atomic details for the CG trajectories.

Introduction

Molecular dynamics (MD) simulations are an important tool for the theoretical study of soft matter systems and have been successfully used to model a large variety of systems such as polymers, liquid crystals, amphiphiles, and proteins [1, 2]. When used in combination with experiments, MD simulations help to gain a deeper understanding of complex phenomena by giving access to structural, mechanical and dynamical properties [3, 4]. The all-atom (AA) description of soft matter systems is fairly accurate, but suffers from spatio-temporal limitations. For instance, it remains challenging, if not impossible, to model the characteristic long time scales of certain polymer processes, such as the glass-transition, the dynamics of long chains, or the phase segregation of block copolymers [5–10]. Coarse-grained (CG) simulations offer a good compromise between computational efficiency and accuracy. Generic CG (G-CG) models describe polymer-like macromolecules as beads connected by elastic springs, where each bead represents a group of atoms or monomers, and possibly other interaction potentials. The efficiency of CG simulations comes from a reduced number of degrees of freedom, simpler potential functions, longer simulation time-steps, and the possibility of using an implicit solvent. Despite being comparatively simpler than AA models, G-CG models have been successfully applied to a wide range of polymeric systems, including solutions and melts of neutral polymer chains [11, 12], weak (pH-responsive) polypeptides [13, 14], polyelectrolyte nanogels [15, 16] and biomolecules such as DNA [17] and proteins [18], often yielding quantitative or semi-quantitative agreement with AA simulations and experimental results.

Among the experimental techniques allowing for the characterization of soft matter systems, NMR diffusometry (NMRD) stands out as a noninvasive and relatively fast method to probe molecular dynamics over large frequency ranges [19–21]. By providing direct access to segment diffusion and chain relaxation modes, NMRD allows one to probe the dynamics of soft matter systems like hydrated polymers and hydrogels [22, 23]. NMRD has also been successfully used to investigate the surface reorientational dynamics of fluids nanoconfined in high surface-area microporous materials [21, 24, 25], and is commonly used to study heterogeneous catalytic materials [26] and biomolecules such as proteins [27]. When combined with rheology, NMRD can even provide access to the response of polymer systems under shear [28].

In this article, we evaluate the relevance of generic CG models for the determination of the NMR relaxation rate $R_1$ for protons of spin quantum number $I = 1/2$, for which the primary relaxation mechanism is dipole-dipole interaction. To do so, a model polymer system made of short PEG chains in water is used. Two types of simulations are employed, AA and G-CG, and their respective values of $R_1$ are compared. In addition, in order to differentiate between the effects of using a model with reduced number of degrees of freedom (i.e., a CG
model) on the NMR relaxation rate measurement, and the known discrepancies between AA and G-CG models, \( R_1 \) was also measured from monomers center-of-mass (CoM) in the AA simulation. The latter yields the same reduced level of detail as the G-CG simulations, but with exactly the same molecular dynamics as the AA. We discuss the differences in \( R_1 \) between the AA and the CG/CoM models by splitting the signal into intramonomer, intermonomer, and intermolecule components. Finally, we propose a method to correct the values of \( R_1 \) obtained from CG models, allowing us to directly compare \( R_1 \) results from CG simulations with results from NMR experiments.

\[ \text{PEG-H}_2\text{O} \]

Figure 1: (a-c) Snapshots from the AA simulations of PEG-H\(_2\)O mixtures with PEG concentration \( c_{\text{PEG}} = 0.1 \) (a), 1 (b), and 10 (c) [29]. (d) AA, CoM-CG, and G-CG representations of a PEG molecule with 5 monomers (molar mass \( M_{\text{PEG}} = 200 \text{ g/mol} \)). The double arrows \( r_{ij} \) indicate the distance between two spin carrier locations for the respective models. Oxygen, hydrogen, and carbon atoms of the PEG molecule are in red, white, and gray, respectively. Water molecules (in blue) are explicitly considered only for AA NMR calculations.

\[ \text{Method} \]

\[ \text{Simulation models} \]

In total, two types of simulations of the model polymer system were performed: AA simulations with explicit water, and G-CG simulations with implicit water.

\[ \text{AA} \]

All-atom (AA) molecular dynamics (MD) simulations were used to study a bulk polyethylene glycol-water (PEG-H\(_2\)O) system with a large range of PEG concentrations \( c_{\text{PEG}} \in [0.01, 100] \) (Fig. 1abc), where \( c_{\text{PEG}} \) is unitless and defined as the ratio between the total number of PEG’s oxygen atoms and the total number of \( \text{H}_2\text{O} \) molecules. A PEG concentration \( c_{\text{PEG}} = 0.01 \) corresponds to 2 PEG molecules and 1000 water molecules, and \( c_{\text{PEG}} = 100 \) corresponds to 160 PEG molecules and 8 water molecules. AA MD simulations were performed with the open-source GROMACS simulation package [30]. The number of atoms in each simulation was chosen approximately equal to 5000, corresponding to a typical box size of \((3.4 \text{ nm})^3\). For most of the calculations, PEG200 with molar mass equal to 200 g/mol was used. PEG200 is composed of 3 monomer units, O(CH\(_2\))\(_2\), and two ends, CH\(_2\)OH, for a total of 5 oxygen atoms per PEG molecule (Fig 1d). Simulations with PEG molecules with molar mass ranging from \( M = 100 \text{ g/mol} \) (3 monomers) to 600 g/mol (10 monomers) were also performed. The CSVR thermostat [31] was used with a time constant of 0.5 ps to impose a temperature \( T = 298.15 \text{ K} \). An isotropic pressure coupling was applied using the Berendsen barostat [32] with a time constant of 1 ps. A cutoff of 1.4 nm was used for the Lennard-Jones interaction with long range dispersion correction. Long range electrostatic interactions were handled using the smooth particle mesh Ewald method (SPME) [33]. The TIP4P/\( \epsilon \) force field was used for water [34] together with the CHARMM36 force field for the PEG molecules [35]. The simulation scripts are available on the DaRUS repository [36].

The simulations were carried out according to the following protocol: after an energy minimisation, two successive equilibration steps were performed, respectively, a NVT step of 5 ps with a timestep of 0.1 fs, a first NPT step of 1 ns with imposed pressure of 1000 bar and timestep of 1 fs, and a second NPT step of 1 ns with imposed pressure of 1 bar and timestep of 1 fs. Then, production runs of total duration \( t_{\text{tot}} \) were performed in the NPT ensemble with a timestep of 1 fs. Trajectories were recorded with a dumping period \( \Delta t \). For the systems with the lowest viscosity, corresponding to the lowest PEG concentration, \( c_{\text{PEG}} < 1 \), \( t_{\text{tot}} \) was chosen equal to 50 ns, and \( \Delta t \) was chosen equal to 0.2 ps. For the systems with the highest viscosity, \( c_{\text{PEG}} \geq 1 \), \( t_{\text{tot}} \) was chosen equal to 500 ns, and \( \Delta t \) was chosen equal to 2 ps. We ensured that the dumping period \( \Delta t \) had no significant impact on the result by performing convergence tests (Fig S1). For the lowest PEG concentrations \( c_{\text{PEG}} = 0.01 \) and 0.03, up to six independent production runs were performed in order to improve the statistics.

\[ \text{G-CG} \]

G-CG simulations were performed with the open-source software package ESPResSo [37]. Short polymer chains consisting of 5 beads each were initially placed in a cubic box, and water was treated implicitly. The number of PEG molecules and the box dimensions were chosen to mimic the AA simulations. A generic bead-spring model derived from the Kremer-Grest polymer model [11] was used to model the PEG chains. Each monomer...
was represented by a bead of mass \( m \approx 39 \text{u} \) interacting via the purely repulsive Weeks-Chandler-Andersen potential [38], with an energy scale \( \epsilon = k_B T \) and a bead diameter \( \sigma = 0.355 \text{nm} \). The chemical bonds between adjacent pairs of monomers were modeled using the finite extensibility nonlinear elastic (FENE) potential [11] with magnitude \( k = 30 \epsilon / \sigma^2 \), maximal stretching length \( \Delta r_{\text{max}} = 1.5 \sigma \), and equilibrium length \( r_0 = 0 \). Additionally, an angular potential

\[
V_{\text{ba}}(\phi) = \frac{k_{\text{ba}}}{2} (\phi - \phi_0)^2 ,
\]

was imposed to adjacent triplets of monomers, where \( k_{\text{ba}} \) is the bond-angle bending constant, \( \phi \) the bond angle, and \( \phi_0 = \pi \) the equilibrium bond angle. The value of \( k_{\text{ba}} = 1 \text{e} \) was used as a fitting parameter to match the radius of gyration of the AA PEG molecules (Fig. S2). The system was thermalised to a temperature of \( T = 298.15 \text{K} \) by coupling the particles to a heat bath using the Langevin thermostat [2] with friction coefficient

\[
\gamma = 3 r_0 \pi \eta \sigma ,
\]

where \( r_0 \eta \) is a numerical prefactor. The value of \( \eta = 1 \text{mPa \cdot s} \) was chosen to mimic the interaction with an implicit solvent corresponding to bulk water, and the coefficient \( r_0 \) was chosen so that \( R_1(f) \) spectra show the same dependence in the frequency \( f \) as the one extracted from CoM calculations (Fig. S3): for low PEG concentration \( c_{\text{PEG}} < 1 \), \( r_0 \) was taken equal to 1.5, and for \( c_{\text{PEG}} \geq 1 \), \( r_0 \) was taken equal to 5. Finally the stochastic equations of motion were integrated using the Velocity-Verlet method [1]. The scripts are available in the DaRUS repository [36].

The simulations were performed according to the following simulation protocol: after an energy minimisation using the method of steepest descent, two successive equilibration steps were performed (both in the NVT ensemble): first, an integration of 10 ns with a timestep of 5 fs was performed, where a force capping that gradually increased over time was applied. The next step consisted of 50 ns (again with a timestep of 5 fs) without any force capping. Then, production runs of total duration \( \tau_{\text{tot}} = 50 \text{ns} \) were performed in the NVT ensemble with a timestep of 10 fs. For \( r < 1 \), the positions were written out with a period of \( \Delta t = 0.2 \text{ps} \), while for \( r \geq 1 \), \( \Delta t \) was chosen equal to 2 ps (i.e., the same \( \Delta t \) values as for AA calculations). We made sure that this choice provided us with a high accuracy on the calculation of \( \tau_T \) by performing similar convergence tests as the ones presented in Fig. S1. For each value of \( r \), 3 independent production runs were performed in order to improve the statistics.

**NMR measurement of \( R_1 \)**

The measurements of the NMR relaxation rate \( R_1 \) were performed at two different levels of details: AA and CG. In addition, two types of CG calculations were made, CoM-CG from the CoM of the AA model, and G-CG from the generic bead-spin model. The Python script used for the calculation of \( R_1 \) is openly available [39].

**AA**

The NMR relaxation rate \( R_1 \) for interacting dipolar nuclei can be calculated from AA molecular dynamics without adjustable parameters, allowing for direct comparisons between simulations and experiments [4, 21, 40]. For isotropic systems, \( R_1 \) is calculated from the autocorrelation function of fluctuating magnetic dipole-dipole interactions [41],

\[
G(t) = \frac{1}{N_H} \sum_{i \neq j} \langle F_{ij}(t) F_{ij}(0) \rangle ,
\]

where \( N_H \) is the total number of pairs of hydrogen atoms (or pairs of spins) where \( i \neq j \). Here, when both \( \text{H}_2\text{O} \) and PEG are considered, all possible pairs respecting \( i \neq j \) are included in Eq. (2): \( \text{H}_2\text{O}-\text{H}_2\text{O} \), PEG-PEG, and PEG-PEG. When only PEG molecules are considered, which eases the comparison with CG models, only the PEG-PEG pairs that are included in Eq. (2). Finally

\[
F_{ij}(t) = \sqrt{\frac{16 \pi \gamma J (\theta_{ij}(t))}{5 \pi r_{ij}^3(t)}} ,
\]

where \( J(0) \) is the normalized spherical harmonic with \( \ell = 2 \) and \( m = 0 \), \( \theta_{ij}(t) \) is the nuclear spin separation, and \( \theta_{ij}(t) \) is the polar angle of the direction \( r_{ij} \) with respect to laboratory axes, assuming that the applied static magnetic field \( B_0 \) is parallel to \( e_z \). Then, the NMR relaxation rate \( R_1 \) is calculated using

\[
R_1 = K [ J(\omega_0) + 4 J(2\omega_0) ] ,
\]

where \( J \) is the spectral density obtained as the Fourier transform of \( G \), \( \omega_0 = \gamma B_0 \) is the Larmor frequency with \( \gamma \) the gyro-magnetic ratio for \(^1\text{H} \) with spin \( I = 1/2 \), and

\[
K = \frac{1}{4} \left( \frac{\mu_0}{4 \pi} \right)^2 \hbar^2 \gamma^2 I (I + 1) ,
\]

where \( \mu_0 \) is the vacuum permeability. Here, the hydrogen atoms from both PEG and \( \text{H}_2\text{O} \) molecules were assumed to carry one spin \( I \) each. More information about the method and the theory can be found in Refs. [4, 42–46].

**CoM-CG**

The position \( \textbf{R} \) of the CoM of each monomer/bead was first calculated from the AA simulations using

\[
\sum_i m_i (\textbf{r}_i - \textbf{R}) = 0 ,
\]

where the sum runs over all the atoms of a monomer, and \( m_i \) is the mass of the atom of position \( r_i \) (Fig. 1d). Contributions from \( \text{H}_2\text{O} \) were neglected, and the NMR relaxation time \( R_1 \) was calculated using Eq. (6) (see the next subsection) together with Eqs. (3-5).

**G-CG**

G-CG calculations of \( R_1 \) were done assuming that each bead carries an effective number of spins \( n_{\text{eff}} \), where \( n_{\text{eff}} \) was chosen such that the total number of spins per PEG
molecule is the same as for the AA model. For PEG molecules with molar mass $M_{\text{PEG}} = 200\,\text{g/mol}$, each end contains 3 hydrogen atoms, and each central monomer contains 4 hydrogen atoms (Fig. 1d), so $n_{\text{eff}} = (2 \times 3 + 3 \times 4)/5 = 3.6$. For G-CG models, the autocorrelation function (Eq. (2)) was modified as follows:

$$G(t) = \frac{n_{\text{eff}}}{N_B} \sum_{i \neq j} \langle F_{ij}(t + \tau)F_{ij}(\tau) \rangle,$$

where $N_B$ is the total number of bead pairs where $i \neq j$. Finally, $R_1$ was calculated using Eq. (6) together with Eqs. (3-5).

**Viscosity measurements**

The shear viscosity $\eta$ of the PEG-water mixtures was evaluated using a non-equilibrium method [47]. In short, a static, cosinusoidal acceleration $A \cos(2\pi t/L_z)$ was applied along the $x$-direction, where $A = 0.03\,\text{nm/ps}^2$ ($A = 0.1\,\text{nm/ps}^2$ for G-CG) is the amplitude and $L_z$ is the box height. The viscosity was then calculated from the generated velocity profile, see Ref. 47 for details. For the AA simulations, the GROMACS implementation was used [30, 47], and for the G-CG simulations, a homemade implementation on the Python level in ESPResSo was used.

For both, AA and G-CG models, a number $N_\eta$ of production runs of 10 ns each were performed. For the AA model, $N_\eta = 6$ for the least viscous systems ($\zeta_{\text{PEG}} \rightarrow 0$), and $N_\eta = 20$ for the most viscous systems ($\zeta_{\text{PEG}} \rightarrow 100$). For the G-CG model, a larger number of runs were needed to achieve sufficient statistics due to reduced degrees of freedom as compared to the AA model. We used $N_\eta = 2000 \forall \zeta_{\text{PEG}}$. For both AA and G-CG calculations, the mean values and standard deviations of the viscosities were calculated from the different runs.

**Reverse coarse graining procedure**

Reverse CG (RCG) methods allow to reconstruct atomistic details from their corresponding CG representation [48]. RCG was performed using a numerical procedure similar to the one developed by Rzepiela et al. [48], here implemented using LAMMPS [49]. First, an AA simulation containing a single PEG molecule was performed. Harmonic potentials between the CoM of each monomer and the positions of the beads extracted from the G-CG simulation were added. Then, the energy of the system was minimized using steepest descent. This minimization step was conducted for every frame until a full trajectory was obtained for each PEG molecule, and then all the molecules were merged and wrapped into a single simulation trajectory. The final trajectory, called G-CG-H where the ‘-H’ refers to the reconstructed hydrogen atoms, was then used for calculating NMR relaxation rate $R_1$ using Eqs. (2-5).

**Results**

**Validation of the AA model**

First, in order to validate the AA model as well as the measurement methods, both the dynamic viscosity $\eta$ and the NMR relaxation rate $R_1$ of PEG-H$_2$O mixtures were measured and compared to experimental results by Jora et al. [50].

The dynamic viscosity $\eta$ of the PEG-H$_2$O mixtures was determined for increasing PEG concentration $c_{\text{PEG}}$ using the non-equilibrium method of Hess et al. [47]. Our results show an increase of $\eta$ by a factor of about 30 when increasing $c_{\text{PEG}}$ from 0.01 to 100, in good agreement with experimental results from Jora et al. (Fig. 2a). The steepest increase occurs for intermediate PEG concentration ($c_{\text{PEG}} \approx 1$), corresponding to the semi-dilute regime where the interaction between PEG molecules are non negligible [50]. In the diluted regime $c_{\text{PEG}} \rightarrow 0$, one recovers the viscosity of pure water, $\eta \rightarrow 1\,\text{mPa}\cdot\text{s}$. In the concentrated regime $c_{\text{PEG}} > 10$, the viscosity of the AA simulation is $40\pm10\,\text{mPa}\cdot\text{s}$, which is only slightly lower than the experimental viscosity of $50\pm5\,\text{mPa}\cdot\text{s}$.

![Figure 2: a) Dynamic viscosity $\eta$ of the PEG-H$_2$O mixtures as a function of the PEG concentration $c_{\text{PEG}}$. PEG molecules have 5 monomers. Full disks are all-atom (AA) molecular dynamics simulations, and full squares are experimental data from Ref. [50]. Error bars correspond to the standard error. b) NMR relaxation rate $R_1$ for the PEG-water mixtures as a function of $c_{\text{PEG}}$. Open disks denote CoM-CG calculations performed from AA simulations and including the hydrogen atoms of both water and PEG molecules.](image)
One specificity of our CoM model, and of CG models in general, is the absence of an explicit solvent. To access the error introduced on the NMR $R_1$ when ignoring the spins of the H2O molecules, $R_1$ was re-calculated from the AA PEG-H2O mixtures using the same procedure as before [i.e. using Eqs. (2-5)], but while disregarding all the contributions involving the hydrogen atoms of the H2O molecules in the double summation in Eq. (2). Our results show that ignoring the H2O molecules only matters at low PEG concentration $c_{PEG} \leq 0.1$, where it leads to an overestimation of $R_1$ by a factor up to 2 (Fig. S5a). Such overestimation of $R_1$ is likely due to the comparatively slower translational and rotational motions of PEG as compared to H2O (Fig. S5b), since a slower molecular dynamics means faster spin relaxation, i.e. a larger value of $R_1$ [4, 46].

CoM analysis

To evaluate the validity of deducing $R_1$ from CG models, we performed NMR $R_1$ measurements from the CoM of the PEG monomers as extracted from the AA models, assuming that an effective number of spins $n_{eff}$ was located at each CoM position. This a-posteriori coarse-graining procedure allows us to separate the known intrinsic discrepancies between AA and G-CG models [53–56] from the potential error on the values of $R_1$ introduced by the coarse-graining itself.

Our results show that CoM-CG calculations underestimate the value of $R_1$ by a multiplicative factor of about 6 (Fig. 2b) as compared to AA calculations. Interestingly, this factor of $\approx 6$ on $R_1$, which we refer to in the following as an offset in the log-log scale, is found to be quasi-independent of the PEG concentration $c_{PEG}$, PEG molar mass $M_{PEG}$, and simulation temperature $T$ (Fig. S4). Our results show that even though NMR calculations applied at a coarse-grained level do not provide a quantitative estimate for $R_1$, they can remarkably still provide the right trends.

In order to better identify the origin(s) of the mismatch between AA and CoM-CG calculations, the NMR rate $R_1$ was broken down into 3 components: inter-molecule ($R^T_1$), inter-monomer ($R^{R\text{-chain}}_1$), and intra-monomer ($R^{R\text{-bead}}_1$) (Fig. 3ab). The total NMR relaxation rate follows as

$$R_1 = R^T_1 + R^{R\text{-chain}}_1 + R^{R\text{-bead}}_1. \quad (7)$$

The inter-molecule component $R^T_1$ was calculated by performing the double summation in Eq. (2) only between atom pairs belonging to different molecules. $R^T_1$ typically provides information on the translational (T) motion of molecules [4]. The inter-monomer component $R^{R\text{-chain}}_1$ was calculated from atoms belonging to different monomers from the same molecule, and the intra-monomer component $R^{R\text{-bead}}_1$ was calculated from atoms belonging to the same monomer. The letter R in R-chain and R-bead stands for rotational [4]. While T and R-chain contributions are present within both AA and CoM-CG calculations, R-bead can only be calculated from an AA representation (Fig. 3ab). Therefore, $R^{R\text{-bead}}_1 = 0$ in Eq. (7) in the case of CoM-CG calculations.

All-three contributions $R^T_1$, $R^{R\text{-chain}}_1$, and $R^{R\text{-bead}}_1$ were extracted from the AA calculations. For the sake of comparing AA to CoM-CG, contributions involving the water molecules were excluded from all calculations here. Our results show that, in the low PEG concentration limit, $c_{PEG} \to 0$, the intra-monomer contribution $R^{R\text{-bead}}_1$ dominates over the two others, while in the high PEG concentration limit, $R^{R\text{-bead}}_1$ and $R^T_1$ are of similar magnitude and dominate over $R^{R\text{-chain}}_1$ (Fig. 3c). As expected, the relative importance of $R^T_1$ with respect to $R^{R\text{-bead}}_1$ and $R^{R\text{-chain}}_1$ significantly increases as $c_{PEG}$ increases, due to PEG-PEG interactions becoming more and more prominent at larger PEG concentration. Overall, our results obtained with the AA model show that $R^{R\text{-bead}}_1$ is the dominant contribution to $R_1$ for all $c_{PEG}$. Therefore the absence of the $R^{R\text{-bead}}_1$ contribution from a CoM-CG representation explains, at least in part, the offset measured on the value of $R_1$ when performing CoM-CG NMR calculations (Fig. 2b).

In the case of CoM-CG calculations, our results show that $R_1$ is dominated by the inter-monomer component ($R^{R\text{-chain}}_1$) in the low PEG concentration limit (Fig. 3d). In the high PEG concentration limit, $R^{R\text{-chain}}_1$ and $R^T_1$ are of similar magnitude. Our results also indicate that
there is no exact agreement between AA and CoM-CG for the calculations of $R_{\text{chain}}^1$ and $R_{\text{chain}}^2$; both contributions calculated from CoM-CG are shifted toward lower values as compared to AA (Fig. 4a-b). Such offsets on $R_{\text{chain}}^1$ and $R_{\text{chain}}^2$ can be explained by the comparatively larger typical distances between spin carriers under the CoM-CG representation as compared with the AA representation, as indicated by radial distribution functions $g(r)$ (Fig. 4c-d). Such increased typical distance between spin carriers leads to a reduction in the absolute value of the correlation function $G$, as $G$ depends on the inter-species distances as $G \sim r^{-6}$ [Eqs. (2-3)], which in turn translates into a reduction of the relaxation rate $R_1$ [Eq. (4)].

G-CG

Finally, the relaxation rate $R_1$ was measured from a G-CG simulation performed with ESPResSo [37]. Short polymer chains consisting of 5 beads each were placed in a cubic box. The number of polymer molecules and the box dimensions were chosen to mimic the AA simulations. Results obtained from the G-CG simulations show that the value of $R_1$ is underestimated by a factor of $\approx 5 - 7$ by the G-CG model compared to AA (Fig. 5a), in good agreement with the results obtained from CoM-CG (Fig. 2b).

Reverse coarse graining

In order to overcome the limitation of CG models and compare directly CG simulations with experiments, we propose here to re-introduce hydrogen atoms based on the positions of the beads during post-processing. The latter is achieved via RCG, and has been successfully used previously to reconstruct molecules with different levels of complexity [48], as well as to generate X-ray scattering data from CG lipid simulations [57]. The NMR relaxation rate $R_1$ was then extracted from the G-CG simulation trajectories with re-introduced hydrogens, which we refer to as G-CG-H (Fig. 5a). Our results show that the inter-molecular term, $R_{\text{chain}}^1$, was measured from a G-CG-H trajectory, differs from the AA prediction calculated from the inter-molecule (T) distances. Overall, our results show that NMR measurements performed from a CG model lead to an offset in the value of $R_1$. In the case of the PEG molecules considered here, there are two major differences between AA and CG calculations; the first difference is the absence of intramonomer contribution $R_{\text{head}}^1$ in CG calculations, and between AA and G-CG-H trajectories for $R_{\text{head}}^1$ and $R_{\text{chain}}^1$ also show good agreement (Fig. S6).

It should be noted that the excellent prediction of $R_1$ by G-CG-H does not imply that G-CG-H trajectories reproduce perfectly the molecular dynamics, or that G-CG and AA systems share the same properties. For instance, the viscosity of the G-CG simulations, as calculated using the same protocol as for the AA simulations, is systematically larger than the experimental viscosity, by a factor up to 3 for the largest values of $c_{\text{PEG}}$ (Fig. S7). Such discrepancies in viscosity between AA and G-CG models are well documented [53-56], and are typically attributed to the missing degrees of freedom in the G-CG models, and to differences between the use of an explicit (AA) and implicit (G-CG) solvents. In addition, the molecular motion of the PEG molecules, as predicted by the G-CG-H trajectory, differs from the AA simulation, as evidenced by comparing the correlation functions (Figs. S8) and the full $R_1$ spectrum (Figs. S9).

Discussion and Conclusion

Overall, our results show that NMR measurements performed from a CG model lead to an offset in the value of $R_1$. In the case of the PEG molecules considered here, there are two major differences between AA and CG calculations, the first difference is the absence of intramonomer contribution $R_{\text{head}}^1$ in CG calculations, and
the second difference is the mismatch of both $R_{1}^{\text{R-chain}}$ and $R_{1}^{\text{T}}$ induced by the modified distances between the spin carriers.

The good agreement on the values of $R_{1}$ obtained between experimental results or AA simulations and G-CG model with reconstructed hydrogens, G-CG-H, attests that only the low-frequency motions of the system matter for the calculation of NMR relaxation rates. Higher frequency motions, such as bond vibrations, are not present in the G-CG-H trajectories, with no apparent consequences on the calculation of $R_{1}$. Therefore, as long as the CG model reproduces the rotational and translational dynamics of the molecules, and as long as the average distance between the re-introduced hydrogen atoms is correct, NMR relaxation rates can be extracted from a CG simulation. As alternatives to the numerical procedure used here to perform RCG, several other RCG methods have proven successful, including the use of virtual sites [58], or the backmapping of atoms using machine learning approaches [59, 60].

In the present contribution, only the longitudinal relaxation rate $R_{1}$ induced by dipole-dipole interaction was considered because it is the relevant relaxation mechanism for the PEG-water system. A priori, the reconstruction of atomic details employed here for the calculation of $R_{1}$ could be used as well to calculate the transversal dipolar relaxation rate $R_{2}$ from a G-CG model since a similar level of atomic details is required. However, in the case of nuclei with spin quantum number $I \geq 1$, the NMR relaxation tends to be dominated by quadrupolar interaction [61]. Its calculation requires the accurate estimate of the autocorrelation function of the electric field gradient at the nucleus, a quantity that depends on the intra-atomic electronic charge distribution and is typically computed using ab initio molecular dynamics methods or classical MD with a polarizable model [61–63]. Therefore, the determination of quadrupolar interaction from a G-CG model would require the reconstruction of a higher level of details than the one required for dipolar interaction, such as Hartee-Fock [64] or QM/MM [65] level. The possibility of using a G-CG model to measure such quadrupolar interaction should be subject of future investigation.

In summary, we have evaluated the relevance of generic CG models for the characterization of the dipolar NMR relaxation rate $R_{1}$. Our results obtained from a PEG-H$_2$O mixture show that $R_{1}$ calculated from a G-CG model follow the same trend as $R_{1}$ from a reference AA model, but with a systematic offset. For PEG, the main origin of the offset on $R_{1}$ is the absence of intra-monomer contributions, which is an intrinsic feature of G-CG models. The second reason for the offset are the differences in the typical distances between spin carriers between G-CG and AA models, which affect both inter-monomer and inter-molecule contributions. We have shown that the offset in $R_{1}$ between G-CG and AA simulations can be corrected exactly by re-introducing hydrogen atoms at their equilibrium positions from the G-CG trajectory using a RCG procedure. Our results obtained from the generated G-CG-H trajectories compare well with the experimental data. Therefore our study is the first step toward the development of joined experimental and numerical NMR studies of complex systems that are not accessible on the AA level, such as long polymer chains and hydrogels.

**Supporting Information**

Supporting figures S1 to S9.

**Acknowledgements**

The authors acknowledge funding by the German Research Foundation (DFG) under grant 268449726. Further funding was provided by the DFG under the grants No. 429529433, 327154368 – SFB 1313, 397384169 – FOR2811, and 390740016 – EXC 2075 under Germany’s Excellence Strategy, as well as by the Stuttgart Center for Simulation Science (SimTech). The authors acknowledge support by the state of Baden-Württemberg through bwHPC.
Availability of Data

All simulation input files and post-processing scripts are available in the DaRUS data repository [36].

References


(39) Gravelle, S. NMRforMD, version 0.0.9, doi: 10.5281/zenodo.6821215.


(50) Jora, M. Z.; Cardoso, M. V.; Sabadini, E. Dynamic aspects of water-poly (ethylene glycol) solutions studied by 1H NMR. *Journal of Molecular Liquids* 2016, 222, 94–100.


TOC Graphic