# Using molecular dynamics trajectories to predict nuclear spin relaxation behaviour in large spin systems

Ilya Kuprov[a], Laura C. Morris[b], John N. Glushka[b], James H. Prestegard[b*]

[a] School of Chemistry, University of Southampton, Southampton, UK.

[b] Complex Carbohydrate Research Center, University of Georgia, Athens GA 30602, United States

[*] To whom correspondence should be addressed, email: jpresteg@uga.edu

## Highlights

- An implementation of a Redfield integrator for MD trajectories is reported
- Predictions of NOEs on sucrose are compared to experiment
- Glycan force field parameters and advanced water models are tested

## Keywords

spin relaxation, molecular dynamics, water models, glycan force fields

**Abstract**

Molecular dynamics (MD) trajectories provide useful insights into molecular structure and dynamics. However, questions persist about the quantitative accuracy of those insights. Experimental NMR spin relaxation rates can be used as tests, but only if relaxation superoperators can be efficiently computed from MD trajectories – no mean feat for the quantum Liouville space formalism where matrix dimensions quadruple with each added spin 1/2. Here we report a module for the *Spinach* software framework that computes Bloch-Redfield-Wangsness relaxation superoperators (including non-secular terms and cross-correlations) from MD trajectories. Predicted initial slopes of nuclear Overhauser effects for sucrose trajectories using advanced water models and a force field optimised for glycans are within 25% of experimental values.

## 1. Introduction

Molecular dynamics (MD) simulations are useful for studying conformational mobility of small molecules and large biomolecular systems [1,2], but concerns persist about the accuracy of conformer energies and populations predicted by MD force fields. There are encouraging indications that MD predicts reasonable timescales for transitions between conformers [3], but this success must be validated: with few exceptions [4], force field parameters are commonly set to reproduce thermodynamic properties and quantum mechanical energies, not the rates of conformational transitions or residence times. As new water models with correct self-diffusion and viscosity are developed [5], and new approaches to modelling unstructured regions of proteins are explored [6,7], thorough experimental validation is becoming increasingly important [8,9].

Magnetic resonance is unique in its ability to probe conformations and their dynamics, particularly through spin relaxation and cross-relaxation measurements. However, most models used to interpret spin relaxation data make spartan approximations about local mobility [10,11] and use isolated spin pair models. More elaborate treatments of NMR [12-15] and EPR [16,17] relaxation using MD simulations recently emerged, but they still mostly use spin pair approximations; the work from one of our groups is also in this category [18,19]. What is needed is a software framework that would be able to directly convert an MD trajectory into a relaxation superoperator [20] for an arbitrary spin system.

In this communication we report an implementation of this functionality within the *Spinach* library [21]. We tested the proposed methods on aqueous sucrose (GLYCAM06 force field [22]) with OPC [5] and TIP5P [23] water. Conformational mobility in sucrose is well researched, particularly by $^{13}C$ spin relaxation measurements interpreted using isolated spin-pair approximations [24-26]. Here, we will focus on the less straightforward but more informative $^{1}H$-$^{1}H$ nuclear Overhauser effect (NOE) that probes the contacts between the H1 anomeric proton (Figure 1) of glucose and H11,12 protons of fructose positioned across the glycosidic bond connecting glucose and fructose residues. We demonstrate that (cross-)relaxation rates extracted from 1 μs long MD trajectories are in excellent agreement with the experimental data, given that discrepancies in NOEs of 9-25% translate into distance errors of 1.5-4.2%. This is good news for the GLYCAM06 force field, as well as for OPC and TIP5P water models. Nuclear spin (cross-) relaxation rates are easy to measure, and we recommend that such tests are included in standard benchmark sets for novel molecular dynamics force fields.

## 2. Numerical evaluation of Redfield's integral

Standard spin relaxation theory treatment using Bloch-Redfield-Wangsness (BRW) formalism [20,27] yields the following expression for the unthermalised relaxation superoperator:

$$\begin{gathered}\hat{\hat{R}} = -\sum_{km}\int_0^{\infty} g_{km}(\tau)\hat{\hat{Q}}_k e^{-i\hat{\hat{H}}_0\tau}\hat{\hat{Q}}_m^{\dagger} e^{i\hat{\hat{H}}_0\tau}d\tau \\ \hat{\hat{H}}_1(t) = \sum_k q_k(t)\hat{\hat{Q}}_k\,,\; g_{km}(\tau) = \langle q_k(0)q_m^*(\tau)\rangle\end{gathered} \tag{1}$$

where $\hat{\hat{H}}_0$ is the static part of the Hamiltonian commutation superoperator, $\hat{\hat{H}}_1(t)$ is the centred stochastic part, $\left\{\hat{\hat{Q}}_k\right\}$ is a suitable superoperator basis, $q_k(t)$ are the corresponding expansion coefficients, angular brackets denote ensemble average, and the following conditions must hold:

$$\left\|\hat{H}_1(t)\right\| \ll \left\|\hat{H}_0\right\|,\;\; \left\|\hat{\hat{R}}\right\| \ll \frac{1}{\tau_C^{max}} \tag{2}$$

where (||_||) denotes matrix 2-norm (largest singular value) and $\tau_C^{max}$ is the longest characteristic decay time found in the correlation functions $g_{km}(\tau)$.

Numerical calculation of the integral in Eq (1) is best set up by taking a step back in Goldman's derivation [28] and undoing the substitution of $\hat{\hat{H}}_1(t)$ to obtain:

$$\hat{\hat{R}} = -\left\langle\int_0^{\infty}\hat{\hat{H}}_1(0)e^{-i\hat{\hat{H}}_0\tau}\hat{\hat{H}}_1(\tau)e^{i\hat{\hat{H}}_0\tau}d\tau\right\rangle \tag{3}$$

where angular brackets denote ensemble average. Given a sufficiently long molecular dynamics trajectory, this expression may be computed by cutting the trajectory into segments, such that each segment is sufficiently long to converge the integral, and the number of segments is sufficient to approximate the ensemble average.

Numerical evaluation of Eq (3) has two convergence conditions. One is for the ensemble average, where the squared standard deviation of the mean for the matrix-valued content of the angular bracket – abbreviated below as $\langle\ldots\rangle$ – must be much smaller than the mean-squared:

$$\left\|\frac{\langle(\ldots)(\ldots)^{\dagger}\rangle - \langle\ldots\rangle\langle\ldots\rangle^{\dagger}}{N}\right\| \ll \left\|\langle\ldots\rangle\langle\ldots\rangle^{\dagger}\right\| \tag{4}$$

where *N* is the number of samples (in our case, approximately 1000). This condition may be checked directly because the samples are computed explicitly. The standard deviation information

may be used in two alternative ways: either as a convergence condition on the ensemble average and thus on the minimum number of trajectory segments, or (when the number of segments cannot be changed) as a source of statistical uncertainty information on the resulting relaxation and cross-relaxation rates.

The other convergence condition is on the upper limit *T* of the finite integral that approximates the improper integral in Eq (3):

$$\left\| \int_0^T f(\tau) d\tau \right\| \gg \left\| \int_T^\infty f(\tau) d\tau \right\| \tag{5}$$

A practical criterion may be obtained by using the representation given in Eq (1) and noting that the integral may be reduced to a linear combination of Fourier transforms of the correlation functions of the following general form [20,27]:

$$j_{km}(\omega) = \int_0^\infty g_{km}(\tau) e^{-i\omega\tau} d\tau \tag{6}$$

where the frequencies $\omega$ are drawn from the set of all eigenvalue differences of the static Hamiltonian. The convergence condition for this integral is that

$$\left| \int_T^\infty g_{km}(\tau) e^{-i\omega\tau} d\tau \right| \ll \left| \int_0^\infty g_{km}(\tau) e^{-i\omega\tau} d\tau \right| \tag{7}$$

After a variable substitution $\tau \rightarrow \tau + T$ on the left hand side, and an observation that the fixed phase multiplier $e^{-i\omega T}$ has no influence on the modulus of the integral, we obtain:

$$\left| \int_0^\infty g_{km}(\tau + T) e^{-i\omega\tau} d\tau \right| \ll \left| \int_0^\infty g_{km}(\tau) e^{-i\omega\tau} d\tau \right| \tag{8}$$

A sufficient condition here is to have $|g_{km}(\tau + T)| \ll |g_{km}(\tau)|$, for which a posterior check may be performed at run time. We have set the upper integration limit so as to have the integration residual smaller than 5% (corresponding to *T* being approximately 3 times our estimated longest rotational correlation time), but the parameter is of course user-selectable.

In practical calculations, the trajectory was parsed into *Matlab* using *MDToolbox* [29]. Chemical shifts were assumed to be isotropic (CSA has a negligible effect on longitudinal cross-relaxation rates in this system) and *J*-couplings were assumed to be fixed (scalar relaxation of the first and second kind are also negligible). The coordinates from each MD frame were fed into

*Spinach* [21] to obtain spin Hamiltonians corresponding to each frame of the trajectory. The integral in Eq (3) was then computed using the standard trapezium quadrature rule subject to the convergence conditions discussed above. For the purposes of the integration, the molecular dynamics trajectory was down-sampled to the longest time step that still guarantees accurate evaluation of the integral in Eq (6), where two conditions must be satisfied: Nyquist-Shannon sampling condition (a minimum two points per period of the largest frequency) for the complex exponential, and the asymptotic accuracy condition of the *N*-point trapezium quadrature $T\{f(x), a, b, N\}$

$$\left|\int_a^b f(x)dx - T\{f(x), a, b, N\}\right| = \frac{(b-a)^2}{12N^2}\left|f'(b) - f'(a)\right| + O(N^{-3}) \tag{9}$$

In our specific case, $\frac{(b-a)^2}{N^2} \to \Delta t^2$ and the derivative at the upper limit of the integral is zero, meaning that the time step must be such that:

$$\Delta t < \pi/||\widehat{\widehat{H}}_0|| \qquad \frac{\Delta t^2}{12}\|f'(\theta)\| \ll \left\|\int_0^\infty f(\tau)d\tau\right\| \tag{10}$$

where $\theta$ is any point within the integration interval. The former condition is checked before the integration takes place, the latter one afterwards. In our case the time step was 2 ps. The relevant code for Hamiltonian calculations and integration over the segments is available in versions 2.6 and later of the *Spinach* library [21].

### 3. Molecular dynamics trajectories

Initial sucrose coordinates were extracted from crystal structure 3500015 deposited in the Crystallography Open Database [30]. Sucrose atoms were assigned parameters from the GLYCAM06 force field [22] as included in the AMBER18 package [31] and imbedded into a box of either OPC or TIP5P water extending 12 Å beyond the edges of the sucrose bounding box in every direction. After energy minimisation, heating to 300 K, pressure equilibration at 300 K, and a further equilibration run of 100 ns, a production run of 1 μs commenced. Seven such runs were initiated, allowing screening for conformations not adequately sampled in 1 μs.

Figure 1 illustrates the dynamics of the glycosidic dihedral angles: $\phi$ fluctuates rapidly by ±20° around the mean value of –30°, $\varphi$ undergoes smaller fluctuations around +55°, but also makes occasional jumps to –40° and +170°; the –40° conformer is populated about 10% of the time, the 170° conformer only occasionally. This behaviour is generally consistent with previous MD work on sucrose [26]. The intra-ring angle $\chi$ has smaller fluctuations, but in this particular trajectory it exhibits a single transition to another state near +60° where it resides for about 80 ns.

Structures of sucrose with ($\phi$, $\varphi$) of (–30°,55°) and (–30°,–40°) are shown on the centre right and top right of Figure 1, respectively. The other ring conformer is shown at the bottom right; it is an alternate chair form, $^{1}C_{4}$, as opposed to the dominant $^{4}C_{1}$ form. It was occasionally seen for both OPC and TIP5P waters, also by others [32,33] with time-domain population of 1-15% and free energy estimates that suggest a 2% population [33]. Our own estimates based on $^{1}J_{H1,C1}$ coupling constants are consistent with this estimate. A combination of 2% population and the occasional jump within one microsecond trajectories makes this conformer insignificant for longitudinal cross-relaxation in NMR – for relaxation analysis, we have therefore chosen a 1 μs trajectory in which this conformer does not make an appearance.

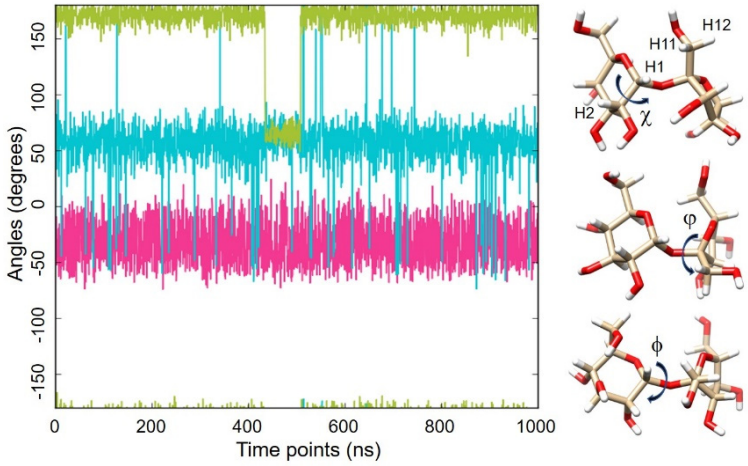


***Figure 1.*** *Dynamics of the indicated dihedral angles within sucrose (1 μs MD trajectory, GLYCAM06 sucrose, OPC water, 300 K). $\chi$ (H2-C2-C1-O2') is shown in green, $\varphi$ (C1-O2'-C2-C1) is shown in cyan, and $\phi$ (H1-C1-O2'-C2) is shown in red.*

Figure 2 shows distance dynamics for the relevant pairs of protons from the OPC water trajectory, showing no transitions to a $^{1}C_{4}$ chair. Inter-residue distances show deep modulation, primarily as a result of rotation about the exocyclic C1-C2 bond in fructose, with occasional jumps as far as 5 Å; these correspond to the above mentioned $\varphi$ jumps to +170°. These fluctuations are not rotational – analytical relaxation theory treatment using purely rotational correlation functions is expected to fail here. Inside the glucose ring, the distance fluctuations are smaller: ±0.3 Å between H1 and H2, with predominantly global rotational dynamics contributing to spin relaxation.

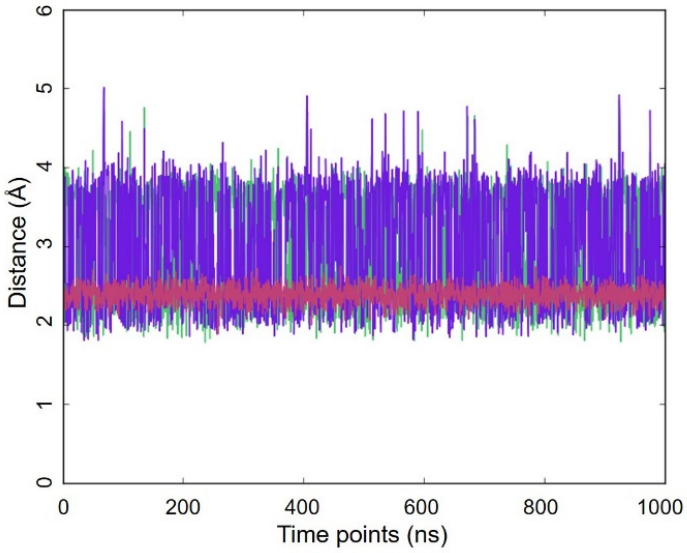


*Figure 2. Dynamics of the distances between the atoms indicated in Figure 1 (1 μs MD trajectory, GLYCAM06 sucrose, OPC water, 300 K): glucose H1 to fructose H11 (green), glucose H1 to fructose H12 (purple), glucose H1 to glucose H2 (red).*

**4. Breakdown of simple rotational diffusion models**

The first stage of the data analysis was a test of standard analytical results from textbook spin relaxation theory [28] that assumes purely rotational modulation of dipolar couplings. An inspection of the MD trajectories (Figure 2) suggests that the modulation of inter-nuclear dipolar couplings inside the glucose ring is predominantly due to global rotation, but the modulation of the dipolar couplings between protons belonging to the different rings is not.

To highlight the breakdown of the rotational diffusion approximation within BRW theory, in Figure 3 we compare analytical and numerical calculations on a standalone three-spin subsystem (H1, H2, H4) inside the glucose ring and a six-spin system (H1, H2, H3, H5 of glucose; H11, H12 of fructose) that spans both rings. These systems were selected by setting a distance cut-off of 3 Å from the central H2 spin in the 3-spin system and 4 Å from the central H1 spin in the 6-spin system using a representative frame from the trajectory. All calculations were run at 14.1 Tesla (600 MHz proton frequency). For the three-spin subsystem of the glucose ring, the agreement with the analytical model using pure rotational diffusion is good (Figure 3, left panel) – rotational correlation time in the analytical model was the only fitting parameter. Within the experimental

error, the fitted value of 78 ps agrees well with ~100 ps reported from $^{13}C$ relaxation analysis of data from a 100 mM sucrose sample in $D_2O$ [24], particularly if corrections for differences in $H_2O$ vs $D_2O$ viscosity are made (a reduction of ~10%). This suggests that isotropic rotational diffusion approximates the dynamics of the relatively rigid glucose ring well. However, the six-spin system that spans both rings shows both an incorrect slope and a greater scatter around the straight line (Figure 3, right panel). This is predictable – Figure 2 clearly shows the presence of significant motion that modulates distances between protons positioned on different sugar residues. That the simple rotational diffusion model with fixed distances – a cornerstone of early NMR structure analysis [34] – should be so fragile is lamentable but not unexpected.

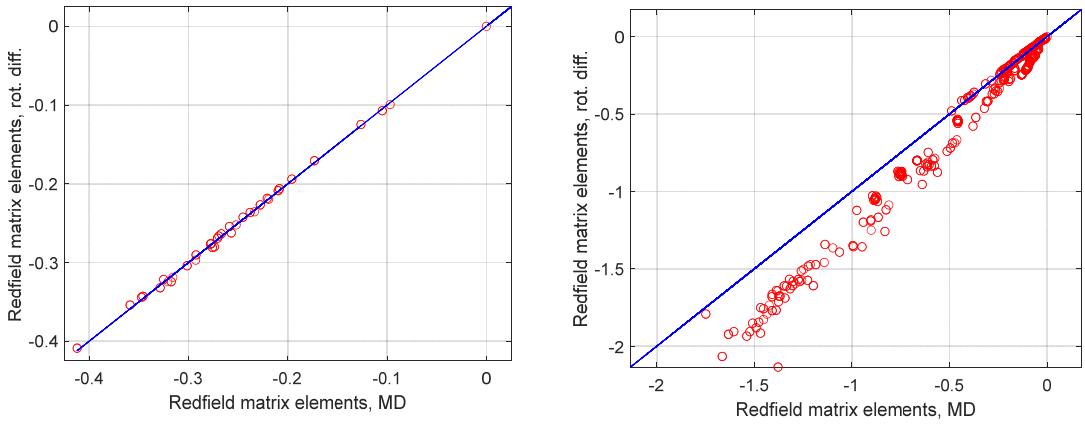


***Figure 3.*** *An illustration of the divergence in agreement between the non-zeroes of the relaxation superoperator for the analytical treatment (using pure rotational diffusion approximation) and the numerical treatment (using MD trajectory integration, OPC water) as the contribution to spin relaxation from internal motion increases.* ***Left panel:*** *3-spin subsystem (H1, H2, H4) in the relatively rigid glucose ring where the dipolar coupling modulation is almost entirely rotational.* ***(Right panel)*** *6-spin subsystem (H1, H2, H3, H5 of glucose; H11, H12 of fructose) where significant internal motion about the glycosidic and exocyclic C1-C2 fructose bond is present. The blue line in both panels corresponds to the rotational correlation time of 78 ps, which was obtained by fitting the left panel data.*

## 5. Rotational diffusion in TIP3P, TIP5P, and OPC water

The well-documented viscosity problem [23] of TIP3P water is apparent in Figure 4, where points representing Redfield matrix elements for our three spin (H1, H2, H4) system calculated numerically for different water models and analytically with a 100 ps isotropic rotational correlation time fall on distinctly different lines. Correlation times predicted by the different models can be extracted by dividing the correlation time for the diagonal (100 ps) by the slopes of line fitted to the points for a given model. Relative to the correlation time extracted from a simulation in OPC

water (78 ps), which is close to experimental expectations, the correlation time extracted from a simulation in TIP3P water is a factor of 1.9 lower. This is in agreement with the factor of 2.2 underestimate of viscosity reported in the literature [23]. It is clear that TIP3P, while computationally efficient, should not be used when modelling rotational diffusion.

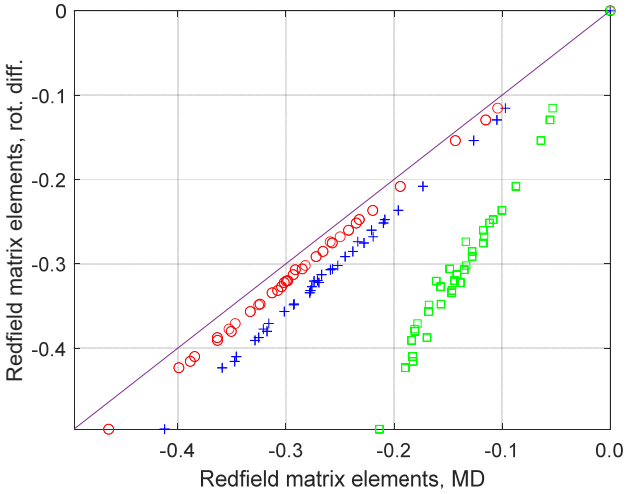


***Figure 4.*** *Non-zeroes of the Redfield relaxation superoperator (X axis) obtained from molecular dynamics trajectories in TIP3P (green squares), TIP5P (red circles), and OPC (blue crosses) water, plotted against the corresponding elements of the analytical relaxation superoperator obtained using the pure rotational diffusion approximation with the rotational correlation time of 100 ps.*

For TIP5P, the fitted rotational correlation time is 88 ps – even closer than OPC to the experimental value of 100 ps [24]. This is somewhat unexpected, because TIP5P is known to overestimate water self-diffusion constant and, hence, underestimate viscosity by 9-14% [5,23]. This may suggest that solute-water interactions are significantly different from water-water interactions in the two models.

## 6. Comparison with experiment

Figure 5 shows NOE enhancements predicted and observed following inversion of the glucose H1 proton resonance in an 8-spin system (H1, H2, H3, H4, H5 of glucose and H11, H12 H3 of fructose) in which the H4 of glucose and H3 of fructose were added explicitly based on their close proximity to one or more of the original six spins. The left panel shows NOEs for the sum of fructose H11 and H12 protons (those resonances overlap); the right panel shows NOEs for the

glucose H2 proton. Given the number of approximations involved in a molecular dynamics simulation followed by the spin Hamiltonian generation, followed by the calculation of the Redfield integral over a finite trajectory and a finite ensemble, the agreement between experimental and predicted NOE accumulation curves is better than expected. Initial slopes are within experimental error for the H11-H12 pair of fructose and within 25% for the H2 proton of glucose. Simulations using analytical relaxation superoperators calculated from static (single frame) structures, including one close to the crystal structure and one representing to the most frequently sampled MD conformer, show larger deviations, particularly for H11+H12 NOEs. This underlines the importance of using MD trajectories in these simulations.

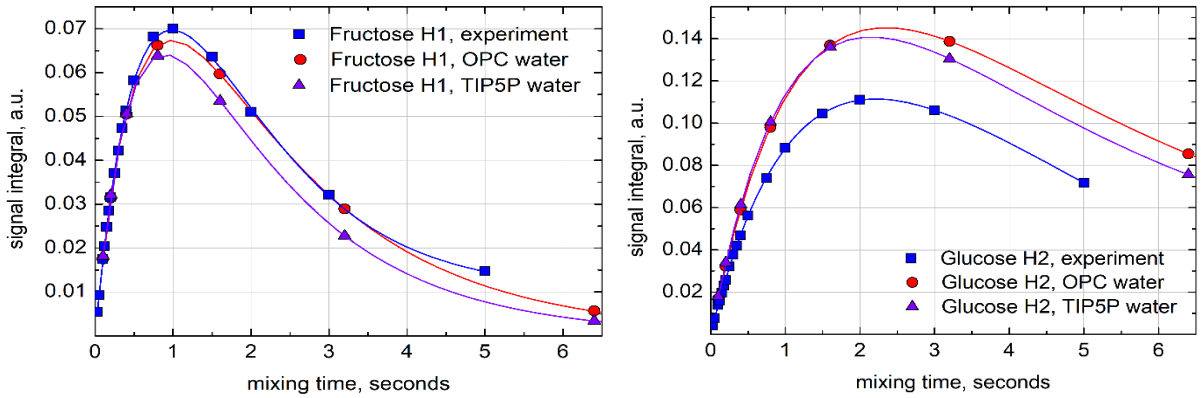


***Figure 5.*** *Comparison of experimental (blue) and simulated (red for OPC water, violet for TIP5P water) NOE accumulation curves following the inversion of glucose H1 proton spin (the curves are biexponential fits). Simulations were performed using an 8-spin system centred on the H1 of glucose (Glc H1-H5; Fruc H11, H12, H3). Signal integrals are fractions of the inverted resonance intensity. Experimental data were collected at 600 MHz and 308 K; the higher temperature was chosen to compensate for the increased viscosity of $D_2O$. No attempt was made to compensate for the effect of sucrose concentration on the viscosity – that effect is expected to be ~0.5%. The sample was degassed prior to NMR experiments.*

Even with the level of agreement shown between experimental data and superoperator simulations, it is tempting to speculate on the origin of such deviations as do exist. The mismatch is largest for the glucose H2 proton. The NOEs are here dependent primarily on global rotational motion of the relatively rigid glucose ring. The deviations for both the initial slopes and NOE maxima are in the direction expected for the shorter correlation times predicted by the two water models. Going back to a spin pair model, one expects the positive NOEs to stem from dominance of double quantum transitions over zero quantum transitions for this small molecule. The contributions do cancel to some extent, and this extent increases with longer correlation times leading to lower NOE magnitudes for the experimental observation. The deviations from the experiment

for the fructose H1 protons are minimal, especially for the initial slopes. In this case, the spectral power density giving rise to NOEs may have larger contributions from modulation of dipolar interactions by internal motions about the glycosidic and the exocyclic bonds. The relative translational motion of proton pairs spanning the two rings may be better reproduced by the more reliably parameterised dihedral angle and direct repulsion terms of the force field. The global rotational modulation responsible for the less accurately reproduced intra-ring NOE would rely on the same indirect effects that are responsible for the overall viscosity of the water in both models, and therefore be harder to match.

**7. Full spin system *versus* spin-pair simulations**

Three-spin and higher effects are plentiful, if sparsely researched, in liquid state magnetic resonance systems – it is of some interest to compare our simulation results, which do include multi-spin effects and DD-DD cross-correlations, to previous models based on spin-pair approximations where those effects were neglected.

The cross-correlation contribution [35] to the relaxation rate for the central spin of a 3-spin system depends on the distances $r_{1,2}$ and the angle $\theta$ between the distance vectors as $r_1^{-3} r_2^{-3}(3\cos^2\theta - 1)/2$. This function has a directional extremum at $r_1 = r_2$, which ranges from +100% of the dipolar rate at $\theta = 180^\circ$ to –50% at $\theta = 90^\circ$ – a large contribution that one cannot reasonably dismiss, particularly in multi-spin systems. Figure 6 illustrates this by comparing the results of our full spin system calculation to the MD-based spin-pair approach similar to one we reported a few years ago [18], where correlation functions were computed from MD trajectories, and then used to assemble a spin-pair approximation to the relaxation superoperator.

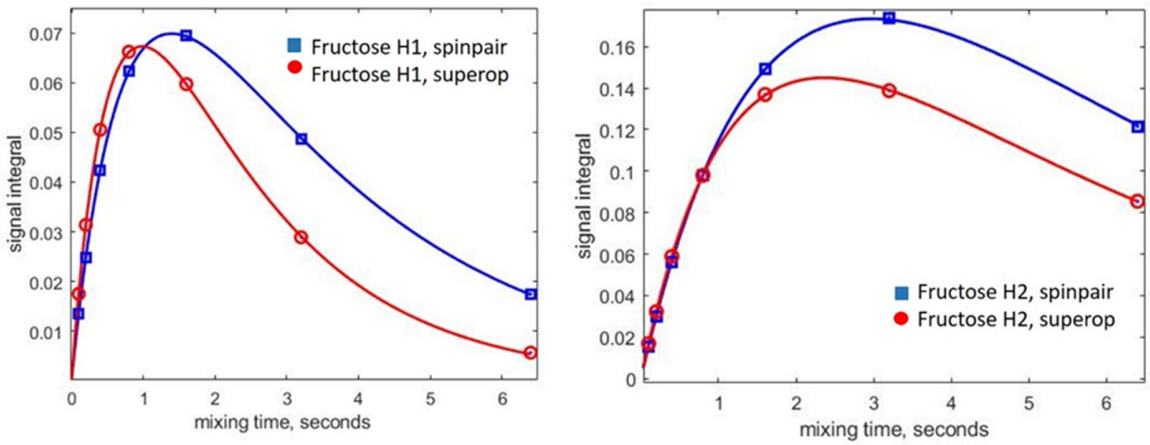


***Figure 6.*** *Comparison of simulated NOE build-up curves (on inversion of the glucose H1 resonance) using spin-pair (blue) and full spin system (red) approaches. Solid lines are biexponential fits. Left panel: fructose H11+H12 protons. Right panel: glucose H2 proton.*

Figure 6 shows simulated NOE accumulation curves for the same 8-spin system using the full spin system approach (red curves) and an updated version of the spin pair approximation (blue curves) which separates the overall rotation from the internal motion, as described by Gu *et al.* [12]. The spin pair model requires an independent estimate of the overall rotational correlation time; we set this to the effective correlation time we had extracted above from the 3-spin glucose cluster in OPC water (78 ps).

The difference between the full spin system treatment and the spin pair approximation is tolerable for cross-relaxation effects used in structure analysis (18% reduction of the NOE maximum for H2, 6% reduction for H11+H12, and 26% increase in the initial slope of the build-up curve for H11+H12). However, differences are comparable to the differences seen between spin-pair simulations and experiment, suggesting that full spin system methods that include cross-correlations and strong coupling effects are essential when using magnetic resonance data to fine-tune MD force field parameters.

## 8. Conclusions and outlook

We have presented a new tool for comparing experimental spin relaxation data to simulations based on molecular dynamics trajectories. It directly calculates spin relaxation superoperators by numerical evaluation of Redfield's integral. The tool is integrated into the *Spinach* package, and enables simulations of a variety of spin relaxation experiments, including those on spin systems with prominent strong coupling and relaxation interference effects.

Calculations of this type have only very recently become feasible, with five principal factors making a contribution: (a) ready availability of graphics cards with the processing power (over 10 TFLOPS for NVidia Ampere architecture) roughly equal to the world's largest supercomputer in 2002; (b) development and implementation of large-scale methods for relaxation superoperators and sparse matrix exponentials in magnetic resonance [36,37]; (c) continued improvement in the performance and syntactic flexibility of Matlab, where program code is often shorter and more readable than its description by human language and standard mathematical notation; (d) emergence of liquid state spin dynamics simulation methods with linear complexity scaling, with the result that protein-sized systems are no longer out of reach [38]; (e) availability of sufficiently long molecular dynamics trajectories to achieve simultaneous convergence on both the upper limit and the ensemble average in Eq (3).

We have tested this tool by simulating NOE experiments on sucrose, using 1 μs MD trajectories under the GLYCAM06 force field with OPC, TIP5P, and TIP3P water. The agreement with experiment is very good, given that discrepancies in NOE initial slopes of 0-25% translate to distance errors in rigid structures of 0-4.2%. This speaks well for the GLYCAM06 parameterisation of carbohydrates, as well as OPC and TIP5P water models; TIP3P fell short.

The CPU time requirements for the simulations reported above do not exceed those required to run the MD trajectory. Once the relaxation superoperator is calculated, the CPU time consumed by the calculation of NOEs at multiple mixing times, or by a simulation of any other relaxation experiment, is insignificant. The 8-spin relaxation superoperator calculation using IK-1(3,3) basis set and a 1 μs trajectory sampled every 2 ps took approximately 30 hours in a system with 24 contemporary CPU cores; this time is reduced significantly when GPUs are used. Larger spin systems are expected to be straightforward: once the number of spins exceeds the size of the strongly interacting cluster within the restricted state space approximation [39] (typically fewer than five spins in liquid state NMR simulations), the computational complexity scaling with respect to the size of the spin system is approximately linear [38].

In summary, nuclear magnetic relaxation offers subtle and detailed measures of rotational and conformational dynamics, with much potential for molecular dynamics force field refinement. The simulation tools presented here will be valuable in this pursuit because they are highly accurate at the level of the spin physics. This eliminates the concerns about, for example, the effects of strong scalar coupling and relaxation interference. Additional relaxation mechanisms can easily be added and, with the ongoing improvement in computer hardware, the wall clock times will continue to decrease.

## 9. Materials and Methods

**Molecular dynamics simulation:** simulations were done using the AMBER 18 package [31]. Initial sucrose coordinates were taken from an X-ray structure deposited in the Crystallography Open Database (3500015), and the force field parameters were provided by GLYCAM_06j-1 [22]. The initial structure was embedded in rectangular boxes of either OPC, TIP3P, or TIP5P water that extended 12 Å in each direction from the outermost sucrose atoms. In each case, the initial energy minimisation was done using the SANDER module of AMBER and the system was then heated to 300 K over 50 ps using 2 fs steps under constant pressure. Production runs were

carried out using the PMEMD module of AMBER 18, saving frames every 2 ps. Trajectories were analysed and structures in selected frames visualised using tools in UCSF Chimera v1.13.1 [40].

**NMR simulations:** calculations were done using *Matlab R2020a*, *Spinach 2.5* [21], and the *MDToolbox* [29]. All relevant scripts are enclosed with the example set provided with versions 2.6 and later of *Spinach*.

**NMR experiments:** a 7.8 mM solution of sucrose in 99.99% $D_2O$ was degassed by nitrogen bubbling. Spectra were acquired on a Bruker Avance III 600 MHz spectrometer at 35°C using a one-dimensional NOE pulse sequence (selnogpzs.2) that uses selective refocusing [41] for excitation, and incorporates a simultaneous frequency shifted pulse and pulsed field gradient during the mixing time to suppress zero-quantum signals [42].[42]. Additional pulsed field gradients served to suppress radiation damping effects. Spectra were acquired at mixing times from 30 to 5000 ms with a sufficient recycling delay to avoid scan interference. The spectra were processed and integrated in *MestReNova* using 1 Hz exponential line broadening and zero-filling to 65,536 points.

## Author Contributions

**Ilya Kuprov:** conceptualisation, software, writing; **Laura C. Morris:** MD simulation, review and editing; **John N. Glushka:** data acquisition, review and editing; **James H. Prestegard:** conceptualisation, software, writing, funding.

## Funding

This work was supported by the US National Institutes of Health, grant number R01 GM134335-01. Manuscript content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References:

[1] S.A. Hollingsworth, R.O. Dror, Molecular Dynamics Simulation for All, Neuron, 99 (2018) 1129-1143.

[2] M. Karplus, J.A. McCammon, Molecular dynamics simulations of biomolecules, Nature Structural Biology, 9 (2002) 646-652.

[3] N. Smolin, R. Biehl, G.R. Kneller, D. Richter, J.C. Smith, Functional Domain Motions in Proteins on the similar to 1-100 ns Timescale: Comparison of Neutron Spin-Echo Spectroscopy

of Phosphoglycerate Kinase with Molecular- Dynamics Simulation, Biophysical Journal, 102 (2012) 1108-1117.

[4] F. Hoffmann, F.A.A. Mulder, L.V. Schafer, Predicting NMR relaxation of proteins from molecular dynamics simulations with accurate methyl rotation barriers, Journal of Chemical Physics, 152 (2020).

[5] S. Izadi, R. Anandakrishnan, A.V. Onufriev, Building Water Models: A Different Approach, Journal of Physical Chemistry Letters, 5 (2014) 3863-3871.

[6] P.C. Chen, M. Hologne, O. Walker, J. Hennig, Ab Initio Prediction of NMR Spin Relaxation Parameters from Molecular Dynamics Simulations, Journal of Chemical Theory and Computation, 14 (2018) 1009-1019.

[7] Z.F. Jing, C.W. Liu, S.Y. Cheng, R. Qi, B.D. Walker, J.P. Piquemal, P.Y. Ren, Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications, in: K.A. Dill (Ed.) Annual Review of Biophysics, Vol 48, 2019, pp. 371-394.

[8] A. Srivastava, S.P. Tiwari, O. Miyashita, F. Tama, Integrative/Hybrid Modeling Approaches for Studying Biomolecules, Journal of Molecular Biology, 432 (2020) 2846-2860.

[9] S. Bottaro, K. Lindorff-Larsen, Biophysical experiments and biomolecular simulations: A perfect match?, Science, 361 (2018) 355-+.

[10] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, Journal of the American Chemical Society, 104 (1982) 4546-4559.

[11] E. Meirovitch, Y.E. Shapiro, A. Polimeno, J.H. Freed, Structural dynamics of bio-macromolecules by NMR: The slowly relaxing local structure approach, Progress in nuclear magnetic resonance spectroscopy, 56 (2010) 360.

[12] Y. Gu, D.W. Li, R. Bruschweiler, NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment, Journal of Chemical Theory and Computation, 10 (2014) 2599-2607.

[13] C.A. Smith, A. Mazur, A.K. Rout, S. Becker, D. Lee, B.L. de Groot, C. Griesinger, Enhancing NMR derived ensembles with kinetics on multiple timescales, Journal of Biomolecular Nmr, 74 (2020) 27-43.

[14] P.J. Nichols, A. Born, M.A. Henen, D. Strotz, C.N. Celestine, P. Guntert, B. Vogeli, Extending the Applicability of Exact Nuclear Overhauser Enhancements to Large Proteins and RNA, Chembiochem, 19 (2018) 1695-1701.

[15] B. Vogeli, The nuclear Overhauser effect from a quantitative perspective, Progress in Nuclear Magnetic Resonance Spectroscopy, 78 (2014) 1-46.

[16] V.S. Oganesyan, A general approach for prediction of motional EPR spectra from Molecular Dynamics (MD) simulations: application to spin labelled protein, Physical Chemistry Chemical Physics, 13 (2011) 4724-4737.

[17] D. Sezer, J.H. Freed, B. Roux, Simulating electron spin resonance spectra of nitroxide spin labels from molecular dynamics and stochastic trajectories, The Journal of chemical physics, 128 (2008) 04B625.

[18] G. Chalmers, J.N. Glushka, B.L. Foley, R.J. Woods, J.H. Prestegard, Direct NOE simulation from long MD trajectories, Journal of Magnetic Resonance, 265 (2016) 1-9.

[19] G.R. Chalmers, A. Eletsky, L.C. Morris, J.Y. Yang, F. Tian, R.J. Woods, K.W. Moremen, J.H. Prestegard, NMR Resonance Assignment Methodology: Characterizing Large Sparsely Labeled Glycoproteins, J Mol Biol, 431 (2019) 2369-2382.

[20] A.G. Redfield, On the theory of relaxation processes, Ibm Journal of Research and Development, 1 (1957) 19-31.

[21] H.J. Hogben, M. Krzystyniak, G.T.P. Charnock, P.J. Hore, I. Kuprov, Spinach - A software library for simulation of spin dynamics in large spin systems, Journal of Magnetic Resonance, 208 (2011) 179-194.

[22] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. Gonzalez-Outeirino, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: A generalizable Biomolecular force field. Carbohydrates, J. Comput. Chem., 29 (2008) 622-655.

[23] M.W. Mahoney, W.L. Jorgensen, Diffusion constant of the TIP5P model of liquid water, Journal of Chemical Physics, 114 (2001) 363-366.

[24] C. Baraguey, D. Mertens, A. Dolle, Anisotropic reorientation and intermolecular interactions of sucrose molecules in aqueous solution. A temperature and concentration-dependent C-13 NMR relaxation study, Journal of Physical Chemistry B, 106 (2002) 6331-6337.

[25] M. Effemey, J. Lang, J. Kowalewski, Multiple-field carbon-13 and proton relaxation in sucrose in viscous solution, Magnetic Resonance in Chemistry, 38 (2000) 1012-1018.

[26] J.C. Xia, D.A. Case, Sucrose in aqueous solution revisited, Part 2: Adaptively biased molecular dynamics simulations and computational analysis of NMR relaxation, Biopolymers, 97 (2012) 289-302.

[27] R.K. Wangsness, F. Bloch, The dynamical theory of nuclear induction, Physical Review, 89 (1953) 728.

[28] M. Goldman, Formal theory of spin–lattice relaxation, Journal of Magnetic Resonance, 149 (2001) 160-187.

[29] Y. Matsunaga, Y. Sugita, Refining Markov state models for conformational dynamics using ensemble-averaged data and time-series trajectories, Journal of Chemical Physics, 148 (2018) 241731-241737.

[30] S. Grazulis, A. Daskevic, A. Merkys, D. Chateigner, L. Lutterotti, M. Quiros, N.R. Serebryanaya, P. Moeck, R.T. Downs, A. Le Bail, Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration, Nucleic Acids Research, 40 (2012) D420-D427.

[31] D.A. Case, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, V.W.D.I. Cruzeiro, T.A. Darden, R.E. Duke, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R. Harris, N. Homeyer, S. Izadi, A.T.K. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York, P.A. Kollman, AMBER 2018, University of California, San Francisco, (2018).

[32] M. Delbianco, A. Kononov, A. Poveda, Y. Yu, T. Diercks, J. Jimenez-Barbero, P.H. Seeberger, Well-Defined Oligo- and Polysaccharides as Ideal Probes for Structural Studies, Journal of the American Chemical Society, 140 (2018) 5421-5426.

[33] B.M. Sattelle, A. Almond, Microsecond kinetics in model single-and double-stranded amylose polymers, Physical Chemistry Chemical Physics, 16 (2014) 8119-8126.

[34] P. Guntert, Structure calculation of biological macromolecules from NMR data, Quarterly Reviews of Biophysics, 31 (1998) 145-237.

[35] A. Kumar, R.C.R. Grace, P.K. Madhu, Cross-correlations in NMR, Progress in Nuclear Magnetic Resonance Spectroscopy, 37 (2000) 191-319.

[36] I. Kuprov, Diagonalization-free implementation of spin relaxation theory for large spin systems, Journal of Magnetic Resonance, 209 (2011) 31-38.

[37] D. Goodwin, I. Kuprov, Auxiliary matrix formalism for interaction representation transformations, optimal control, and spin relaxation theories, The Journal of chemical physics, 143 (2015) 084113.

[38] L.J. Edwards, D. Savostyanov, Z. Welderufael, D. Lee, I. Kuprov, Quantum mechanical NMR simulation algorithm for protein-size spin systems, Journal of Magnetic Resonance, 243 (2014) 107-113.

[39] A. Karabanov, I. Kuprov, G. Charnock, A. van der Drift, L.J. Edwards, W. Köckenberger, On the accuracy of the state space restriction approximation for spin dynamics simulations, The Journal of chemical physics, 135 (2011) 084106.

[40] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF chimera - A visualization system for exploratory research and analysis, J. Comput. Chem., 25 (2004) 1605-1612.

[41] K. Stott, J. Stonehouse, J. Keeler, T.L. Hwang, A.J. Shaka, Excitation sculpting in hogh-resolution nuclear magnetic resonance spectroscopy - application to selective NOE experiments, Journal of the American Chemical Society, 117 (1995) 4199-4200.

[42] M.J. Thrippleton, J. Keeler, Elimination of zero-quantum interference in two-dimensional NMR spectra, Angewandte Chemie-International Edition, 42 (2003) 3938-3941.