A Method to Derive Structural Information on Molecules from Residual Dipolar Coupling NMR Data

Wilfred F. van Gunsteren[#], Maria Pechlaner[#], Lorna J. Smith[†], Bartosz Stankiewicz[&] and Niels Hansen[&*]

[†]Department of Chemistry, Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford, OX1 3QR, UK,

[#]Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, CH-8093 Zurich, Switzerland,

[&]Institute of Thermodynamics and Thermal Process Engineering, University of Stuttgart, D-70569 Stuttgart, Germany

[*]Corresponding author. Electronic mail: hansen@itt.uni-stuttgart.de

Version/Date: 22-04-2022

For submission to J. Phys. Chem.

Keywords

Structure refinement, nuclear magnetic resonance, residual dipolar couplings, averaging time, restraining force, biased orientation molecular distribution.

Abstract

A method for structure refinement of molecules based on residual dipolar coupling (RDC) data is proposed. It calculates RDC values using rotational and molecule-internal configurational sampling instead of the common refinement procedure that is based on the approximation of the non-uniform rotational distribution of the molecule by a single alignment tensor representing the average non-uniformity of this distribution. By using rotational sampling, as is occurring in the experiment leading to observable RDCs, the method stays close to experiment. It avoids the use of an alignment tensor, and thus the assumption that the overall rotation of the molecule is decoupled from its internal motions and that the molecule be rigid. Two simple molecules, two-united-atomic ethane and a cyclo-octane molecule with eight side chains, containing 24 united atoms, serve as so-called “toy model” test systems. The method demonstrates the influence of molecular flexibility and force-field deficiencies on the outcome of structure refinement based on RDCs. For a molecule of a given size (number of atoms N_{at}), there must be a sufficiently large number N_{RDC} of measured RDC values available in order to allow the restraining forces to bias the overall orientation distribution of the molecule. If the ratio N_{RDC}/N_{at} gets too small, the RDC-restraining forces will either not be strong enough to change the overall rotational direction of the molecule such that the target RDC values are approximated well or will be so strong that they induce a local deformation of the molecule. In the latter case the size or inertia of the molecule hinders a restraining induced overall rotation and the internal structure of the molecule is not strong enough to avoid local deformation due

to the restraining forces.

INTRODUCTION

High-resolution prediction of dominant protein structures from their amino-acid sequences and the refinement of low-resolution model structures using experimental X-ray diffraction or Nuclear Magnetic Resonance (NMR) spectroscopic data in order to obtain accurate atomic-level structural information on dominant protein conformations and folds are long-standing challenges in computational structural biology[1]. This is due to the ruggedness of the atomic-level energy function characterising proteins in aqueous solution or membrane environments, and to the very high number of relevant degrees of freedom to be statistical-mechanically sampled while the free energy differences between the many possible protein conformations are relatively small. Moreover, experimental data alone are not sufficient to uniquely determine dominant protein conformations because of three rather fundamental problems: (i) a too low ratio of number of experimental data over number of relevant degrees of freedom, (ii) the inversion problem of defining a structure in terms of experimental observable quantities depending on it, and (iii) the (conformational or structural) averaging inherent to virtually all experimental measurement techniques.

Structural information on molecules that are biologically relevant, such as proteins, nucleic acids and carbohydrates, can be derived from a variety of observable quantities, such as X-ray diffraction or cryo-electron microscopic (EM) intensities and NMR, Circular Dichroism, Raman or infrared spectra to mention a few[1]. Several quantities that are observable by NMR, such as Nuclear Overhauser enhancements (NOEs), ^{3}J-couplings or chemical shifts, only provide local structural information. Only residual

dipolar couplings (RDCs) do provide longer-range information in terms of the average (relative) directions of bond vectors throughout a molecule. In contrast, X-ray diffraction of crystals yields non-local information. In addition, its information density, i.e. the ratio of the number of independent, measured values of observable quantities for a molecule and the number of independent molecular degrees of freedom, is much higher than for NMR in solution. On the other hand, NMR measurements may provide dynamic information in the form of relaxation data, e.g., expressed as S^2 order parameters. Of all techniques available to obtain information on proteins in solution, NMR shows the highest information density[1].

All techniques to derive structural information from the measurement of observable quantities Q make use of a relation of Q to structure $\vec{r}^N$, a function $Q(\vec{r}^N)$[1]. Here $\vec{r}^N \equiv (\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$ denotes the $3N$ Cartesian coordinates of the N atoms or particles of the molecular system. If an observable quantity Q is dependent on the molecular configuration $\vec{r}^N$, one may try to derive an expression or function $Q(\vec{r}^N)$ that approximates the relation between Q and $\vec{r}^N$. This expression may then be used to derive molecular structures that are compatible with measured (averaged) values of Q, i.e.

$$< Q >= \int Q(\vec{r}^N) \exp(-V(\vec{r}^N)/(k_B T)) \mathrm{d}\vec{r}^N / \int \exp(-V(\vec{r}^N)/(k_B T)) \mathrm{d}\vec{r}^N. \qquad (1)$$

$<Q>$ is a Boltzmann-weighted average of $Q(\vec{r}^N)$ over the $3N$-dimensional molecular configuration space. This means that $<Q>$ constitutes an average over a statistical-mechanical ensemble of molecular configurations. The weights are proportional to $\exp(-V(\vec{r}^N)/(k_B T))$, where $V(\vec{r}^N)$ indicates the energy of a molecular configuration or

structure $\vec{r}^N$, k_B Boltzmann's constant and T the temperature.

Since virtually all experimental techniques measure an average $\langle Q \rangle_{space,time}$ of Q over the molecules (space) in the test tube or crystal and over a time window determined by the type of experiment, the derivation of structural information from a set of $\langle Q \rangle$-values should account for the averaging involved in the measurement. This can be done by applying multi-molecule averaging[2] or time-averaging[3] structure refinement instead of the commonly used single-structure refinement technique. Application of time-averaging structure refinement to proteins based on X-ray data[4,5], NMR NOE[6,7] or ^3J-coupling[8] data showed the protein structural variation to be much larger than that observed using single-structure refinement techniques[1].

For observable quantities Q, such as X-ray reflection intensities I_{hkl}, NOEs, ^3J-couplings or chemical shifts, it is possible to formulate a function $Q(\vec{r}^N)$ relating a Q-value to a particular structure $\vec{r}^N$. RDCs may be directly related to structure by assuming a single alignment tensor representing the anisotropic rotational distribution of the molecule, which is, unfortunately, unknown. For observable quantities such as RDCs or S^2 order parameters, the function relating Q to $\vec{r}^N$ involves some average over the Boltzmann ensemble of structures in solution, $Q(\langle f(\vec{r}^N) \rangle)$, where f denotes the function of $\vec{r}^N$ that is being averaged[1]. This means that structure refinement based on such quantities must involve the averaging $\langle f(\vec{r}^N) \rangle$ in addition to the averaging $\langle Q(\langle f(\vec{r}^N) \rangle) \rangle$. For example, S^2 order parameters can be calculated using the ensemble averaging expression[9]

$$S_{\mathrm{XY}}^2 = \frac{1}{2}\left\{3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}\left\langle\frac{\mu_{\mathrm{XY}\alpha}(t)\mu_{\mathrm{XY}\beta}(t)}{r_{\mathrm{XY}}^3(t)}\right\rangle_{t_{\mathrm{av}}}^2 - \left\langle\frac{1}{r_{\mathrm{XY}}^3(t)}\right\rangle_{t_{\mathrm{av}}}^2\right\}(r_{\mathrm{XY}}^{\mathrm{eff}})^6 \quad , \qquad (2)$$

where t indicates time, and t_{av} the time-averaging window, and

$$\mu_{\mathrm{XY1}} \equiv (x_{\mathrm{X}} - x_{\mathrm{Y}})/r_{\mathrm{XY}}, \qquad \mu_{\mathrm{XY2}} \equiv (y_{\mathrm{X}} - y_{\mathrm{Y}})/r_{\mathrm{XY}}, \qquad \mu_{\mathrm{XY3}} \equiv (z_{\mathrm{X}} - z_{\mathrm{Y}})/r_{\mathrm{XY}}, \qquad (3)$$

are the x-, y-, and z-components of the vector $\vec{r}_{\mathrm{XY}} \equiv \vec{r}_{\mathrm{X}} - \vec{r}_{\mathrm{Y}}$ connecting atoms X and Y, and $r_{\mathrm{XY}} \equiv |\vec{r}_{\mathrm{XY}}|$, its length[10]. To obtain a dimensionless quantity, the term in curly brackets is multiplied with the sixth power of the effective length $r_{\mathrm{XY}}^{\mathrm{eff}}$ of the vector $\vec{r}_{\mathrm{XY}}$.

The dipolar coupling $D_{k_1k_2}$ or in short D_k between two nuclear spins k_1 and k_2 in a homogeneous magnetic field $\vec{H}$ depends on[11] (i) the length $r_{k_1k_2} \equiv |\vec{r}_{k_1k_2}|$ of the inter-nuclear vector $\vec{r}_{k_1k_2} \equiv \vec{r}_{k_1} - \vec{r}_{k_2}$, and (ii) the angle θ_k of this vector with the static magnetic field $\vec{H}$, e.g. $\cos(\theta_k) = (z_{k_1} - z_{k_2})/r_{k_1k_2}$ when $\vec{H}$ lies along the z-axis of the coordinate system,

$$D_k = \frac{-\mu_0 h \gamma_{k_1}\gamma_{k_2}}{8\pi^3}\left(\frac{3\cos^2(\theta_k)-1}{2r_k^3}\right) \quad . \qquad (4)$$

Here μ_0 is the magnetic permeability of vacuum $(\varepsilon_0 c^2)^{-1} = 4\pi\ 10^{-7}$ J s^2 / (C^2 m) or 1.9425913 10^{-8} kJ mol^{-1} ps^2 / (e^2 nm), h is Planck's constant, 6.626176 10^{-34} J s or 0.3990313 kJ mol^{-1} ps,

$$\gamma_i = \frac{e}{2m_{\mathrm{p}}} g_i, \qquad (5)$$

where e is the charge of an electron, 1.6021892 10^{-19} C or 1, if e is used as charge unit, m_{p} is the mass of a proton, 1.672622 10^{-27} kg or 1.007276 u (atomic mass unit), and the g-factor of the proton is

$$g_{\mathrm{p}} = \frac{2\mu_{\mathrm{p}}}{\mu_{\mathrm{N}}} = 5.585695. \qquad (6)$$

Here we have used the proton magnetic moment μ_p = 1.410607 10^{-26} J T^{-1} (m^2 s^{-1} C) or 8.804247 10^{-2} e nm^2 ps^{-1}, and the nuclear magneton

$$\mu_N = \frac{e\hbar}{2m_p} \tag{7}$$

has the value 5.050784 10^{-27} J T^{-1} or 3.152427 10^{-2} e nm^2 ps^{-1}. Thus, the gyromagnetic ratio for the proton becomes γ_p = 2.6752 10^8 rad s^{-1} T^{-1} (rad Hertz/Tesla) or 2.7727 rad e nm^2 ps^{-2} /(kJmol^{-1}). Often the quantity $\gamma_p/(2\pi)$ = 4.2577 10^7 Hz T^{-1} or 0.44129 e nm^2 ps^{-2} /(kJmol^{-1}) is used.

A nucleus placed in an external field $\vec{B}$ (magnetic induction, measured in Tesla), will precess at a frequency f (measured in Hz)

$$f = \frac{\gamma}{2\pi} B. \tag{8}$$

The relation between the magnetic field $\vec{H}$ (measured in C s^{-1} m^{-1} or in e ps^{-1} nm^{-1}) and the magnetic induction $\vec{B}$ (measured in Tesla (J s C^{-1} m^{-2}) or in kJ mol^{-1} ps e^{-1} nm^{-2}), also often called magnetic field $\vec{B}$, is

$$\vec{B} = \mu_0 \mu_r \vec{H} \ , \tag{9}$$

with μ_r being the relative magnetic permeability of the medium.

The constant and variable factors in Eq. (4) can be separated as follows,

$$D_k(t) = D_k^c \left(\frac{r^0}{r_{k_1k_2}(t)}\right)^3 \left(\frac{1}{2}\left(3\cos^2\left(\theta_k(t)\right) - 1\right)\right) = D_k^c R_k(t) P_k(t) \tag{10}$$

with

$$D_k^c \equiv \frac{-\mu_0 h}{2\pi (r^0)^3} \left(\frac{\gamma_{k_1}}{2\pi}\right)\left(\frac{\gamma_{k_2}}{2\pi}\right) \ , \tag{11}$$

$$R_k(t) \equiv \left(\frac{r^0}{r_{k_1k_2}(t)}\right)^3 , \tag{12}$$

$$P_k(t) \equiv \frac{1}{2}\left(3\cos^2\left(\theta_k(t)\right) - 1\right) \ . \tag{13}$$

The second-order Legendre function $P_k(\theta_k(t))$ is shown in Figure 1. It has two zero's, at

the so-called magic angles ($|\cos(\theta)|=3^{-1/2}$) 54.73561° and 125.26439°, two maxima of 1 at 0° and 180°, and a minimum of -1/2 at 90°. It is symmetric around $\theta = 90°$. Since the distance between nuclei k_1 and k_2 is generally of the order of 0.1 nm, one may choose $r^0 = 0.1$ nm, which yields $-\mu_0 h/(2\pi(r^0)^3) = -1.3252\ 10^{-10}$ J^2 s^3 C^{-2} m^{-4} or $-1.2336\ 10^{-6}$ (kJ mol^{-1})^2 ps^3 e^{-2} nm^{-4}. Approximate values of the gyromagnetic ratio for some nuclei are listed[12,13] in Table 1. For an N-H pair of nuclei we find for example $D_k^c(^{14}\text{N-}^1\text{H}) =$ -17.37 kHz and $D_k^c(^{15}\text{N-}^1\text{H}) = +24.36$ kHz. We note, however, that a dipolar coupling is also influenced by external factors, such as neighbouring nuclei[14]. Eq. (4) is thus an approximation.

In a measurement of RDCs the value of D_k in Eq. (4) is averaged over molecules and an experimentally determined time period. Since a dipolar coupling D_k generally involves atoms in a molecule that are covalently bound to each other or at a short distance, while the molecule contains many atoms, one may use the assumption that the fluctuation of the distance $r_{k_1k_2}$ is not coupled to the rotational motion of the vector $\vec{r}_{k_1k_2}$ with respect to the magnetic field direction $\vec{H}$. So, the averaging of D_k may be separated into an average over $(r_{k_1k_2})^{-3}$ and one over the second-order Legendre function of θ_k,

$$\langle D_k \rangle = D_k^c \left\langle \left(\frac{r^0}{r_{k_1k_2}} \right)^3 \right\rangle \left\langle \left(\frac{1}{2}(3\cos^2(\theta_k) - 1) \right) \right\rangle \quad . \tag{14}$$

The average $\langle (r^0/r_{k_1k_2})^3 \rangle$ will be close to one in case $r_{k_1k_2}$ is a fluctuating bond length and r^0 a value close to its average. For an isotropically tumbling molecule the average over the function $P_k(\theta_k)$ will be zero,

$$\left\langle P_k\left(\theta_k\right)\right\rangle \equiv \left\langle \frac{1}{2}\left(3\cos^2\left(\theta_k\right)-1\right)\right\rangle = 0 \quad , \tag{15}$$

leading to $\langle D_k \rangle = 0$.

However, if one would be able to introduce an anisotropy in the rotational tumbling of the molecule, one would have $\langle D_k \rangle \neq 0$, and the values of $\langle D_k \rangle$ would yield values of the average $\langle P_k(\theta_k) \rangle$ and some information about values of the angles θ_k.

Experimentally, such an anisotropy can be induced in different ways[11], (i) using the paramagnetic susceptibility $\chi_r \equiv \mu_r - 1$ of a molecule, (ii) using electrostatic interactions with a molecule, or (iii) by immersing the molecule in a medium that contains some order that will influence the angular distribution $P(\theta)$ of the angle θ of some axis in the molecule with the z-axis (magnetic field). In this way, small values of $\langle D_k \rangle$, called *residual* dipolar couplings, of the order of Hz, that is $10^3 - 10^4$ smaller than the values of the dipolar couplings D_k themselves, which are of the order of kHz, can be obtained. Unfortunately, the size and shape of the experimentally induced anisotropy in the rotational distribution $P(\theta)$ cannot be determined. This anisotropy will be different for different media, leading to rather different $\langle D_k \rangle$ values for different media[15].

The current way to handle this problem is to exclude the molecular rotational degrees of freedom of the molecule from the sampling in a calculation of RDCs by assuming a particular average alignment of the molecule with respect to the z-axis (magnetic field). This involves the assumption that the overall rotation of the molecule is decoupled from its internal motions. In addition, it is often assumed that the average

alignment of the molecule has the same effect on different $<D_k>$ values (RDCs), i.e. that the molecule is rigid.

The unknown average alignment of the molecule can be represented by a so-called 3 x 3 alignment tensor, which depends on 5 parameters, three (Euler) angles plus two non-spherical symmetry parameters for the molecule. The 3 x 3 tensor is symmetric and has trace zero, so only 5 independent tensor elements, which represent the anisotropy of the rotational distribution.

The 5 parameters determining the average molecular alignment can be determined in various ways:

1. Estimation (i) from the shape of the molecule[16], or (ii) from its moments of inertia[17], or (iii) from its gyration tensor[18], or (iv) from its hydrodynamic shape[19], or (v) from its alignment against a hypothetical wall, while considering its charge distribution[20].

2. Use of a given set of (measured) RDC values by optimising the agreement between the given and calculated RDCs of the set, which must contain many more than 5 RDC values in order to reach a sufficiently large ratio of data over parameters fitted. This optimisation can be performed by a linear least-squares fit[14], by singular-value decomposition, which is less efficient[21], or by molecular dynamics (MD), stochastic dynamics (SD) or Monte Carlo (MC) sampling of the domain of the 5 alignment tensor parameters.

In the current article we propose an alternative approach to use measured RDC values for structure determination or refinement of bio-molecules that avoids the use of

an alignment tensor. It thus avoids the assumption that the overall rotation of the molecule is decoupled from its internal motions and that the molecule be rigid, so allows for molecular flexibility. The idea is to perform MD or SD simulations that are sufficiently long to sample the rotational distribution well, i.e., long enough to reduce the $\langle D_k \rangle$ values, which should for infinite sampling become zero, to beyond values that are about 10^{-3} to 10^{-4} of the values of D_k^c. Inclusion of a penalty function that drives the calculated RDCs towards the measured ones in the potential energy function of such a simulation would then generate a conformational ensemble compatible with the given set of (measured) RDC values.

METHODS

RDCs as Restraints in Molecular Simulation

Choice of Restraining Function

A simple RDC restraining function, continuous with continuous derivative, is a quadratic one with a flat bottom that allows for a penalty-free range of deviations $2\Delta D^0$ of the average $\langle D_k \rangle$ value or the instantaneous D_k value from its target (measured) value D_k^0. In analogy to the flat-bottom restraining function for NOEs, ^3J-couplings and S^2 order parameters[1], the corresponding function for RDCs would be

$$V_k^{\mathrm{RDC}}\left(D_k(\vec{r}^N); D_k^0, K_k^{\mathrm{RDC}}, \Delta D^0\right) = \frac{1}{2} K_k^{\mathrm{RDC}} (D_k(\vec{r}^N) - D_k^0 - \Delta D^0)^2$$

for $D_k > D_k^0 + \Delta D^0$, (16a)

$$V_k^{\mathrm{RDC}}\left(D_k(\vec{r}^N); D_k^0, K_k^{\mathrm{RDC}}, \Delta D^0\right) = \frac{1}{2} K_k^{\mathrm{RDC}} (D_k(\vec{r}^N) - D_k^0 + \Delta D^0)^2$$

for $D_k < D_k^0 - \Delta D^0$, (16b)

$$V_k^{\mathrm{RDC}}\left(D_k(\vec{r}^N); D_k^0, K_k^{\mathrm{RDC}}, \Delta D^0\right) = 0$$

otherwise, (16c)

with for a set of N_{RDC} D_k values

$$V^{\mathrm{RDC}}\left(D_k(\vec{r}^N)\right) = \sum_{k=1}^{N_{\mathrm{RDC}}} V_k^{\mathrm{RDC}}\left(D_k(\vec{r}^N); D_k^0, K_k^{\mathrm{RDC}}, \Delta D^0\right). \quad (17)$$

The GROMOS force fields[22-26] treat aliphatic CH_n moieties as united atoms. If C_α-H_α RDCs are to be calculated, the implementation allows for the use of a virtual atom[23, 27, 28].

Accounting for Time-Averaging

As for S^2 order-parameter restraining[10], the use of a biquadratic restraining function, which depends on both the instantaneous value $D_k(t)$ and the average value $\langle D_k \rangle_t$ of an RDC, makes no sense, because an RDC is intrinsically an average quantity, $\langle D_k \rangle$, that cannot be linked to a single configuration. When restraining an averaged quantity, the averaging cannot be over the whole simulated period up till t, because the contribution of the configuration at time t to the average will approach zero for t approaching infinity. This implies that the restraining force, which can only be applied to the current configuration at time t, would also tend to zero for t approaching infinity. For this reason, an exponential memory function is used in the time-average $\overline{D_k}^t \equiv \langle D_k \rangle_t$ that is used in the restraining function[3,6],

$$\overline{D_k(\vec{r}^N(t))}^{t,\exp} = \left(\tau_D^{\mathrm{RDC}}\left(1 - \exp(-t/\tau_D^{\mathrm{RDC}})\right)\right)^{-1} \int_0^t \exp(-(t-t')/\tau_D^{\mathrm{RDC}})\, D_k(\vec{r}^N(t'))\mathrm{d}t', \quad (18)$$

where τ_D^{RDC} is the memory relaxation time, which determines the averaging time period. As discussed before, the averaging over the two time-dependent factors in Eq. (10) may be well approximated, due to the different relevant time scales, by a separation, see Eq. (14), so

$$\overline{D_k(\vec{r}^N(t))}^t = D_k^c \overline{\left(\frac{r^0}{r_{k_1k_2}(t)}\right)^3}^t \overline{\left(\frac{1}{2}(3\cos^2(\theta_k(t)) - 1)\right)}^t = D_k^c\ \overline{R_k(t)}^t\ \overline{P_k(t)}^t. \quad (19)$$

The two time-averaged quantities $\overline{R_k(t)}^t$ and $\overline{P_k(t)}^t$ in Eq. (19) are calculated using the above-mentioned damped-memory manner,

$$\overline{R_k(\vec{r}^N(t))}^{t,\exp} = \left(\tau_r^{\mathrm{RDC}}(1-\exp(-t/\tau_r^{\mathrm{RDC}}))\right)^{-1} \int_0^t \exp(-(t-t')/\tau_r^{\mathrm{RDC}})\, R_k(\vec{r}^N(t'))\mathrm{d}t', \quad (20)$$

and

$$\overline{P_k(\vec{r}^N(t))}^{t,\exp} = \left(\tau_\theta^{\mathrm{RDC}}\left(1-\exp(-t/\tau_\theta^{\mathrm{RDC}})\right)\right)^{-1} \int_0^t \exp(-(t-t')/\tau_\theta^{\mathrm{RDC}})\, P_k(\vec{r}^N(t'))\mathrm{d}t', \quad (21)$$

with τ_r^{RDC} and $\tau_\theta^{\mathrm{RDC}}$ their respective memory relaxation times.

Calculation of the Restraining Force and its Contribution to the Virial

The restraining force on atom *i* with position vector $\vec{r}_i$ is

$$\vec{f}_i(t) = -\frac{\partial V^{\mathrm{RDC}}(t)}{\partial \vec{r}_i(t)} = -\sum_{k=1}^{N_{\mathrm{RDC}}} K_k^{\mathrm{RDC}} \left(\overline{D_k(\vec{r}^N(t))}^t - D_k^0 \mp \Delta D^0\right) \frac{\partial \overline{D_k(\vec{r}^N(t))}^t}{\partial \vec{r}_i(t)}, \quad (22)$$

or in terms of the functions P_k and R_k,

$$\vec{f}_i(t) = -\sum_{k=1}^{N_{\mathrm{RDC}}} K_k^{\mathrm{RDC}} \left(D_k^{\mathrm{c}} \overline{R_k(\vec{r}^N(t))}^t \overline{P_k(\vec{r}^N(t))}^t - D_k^0 \mp \Delta D^0\right) D_k^{\mathrm{c}}$$
$$\left\{\overline{P_k(\vec{r}^N(t))}^t \frac{\partial \overline{R_k(\vec{r}^N(t))}^t}{\partial R_k(\vec{r}^N(t))} \frac{\partial R_k(\vec{r}^N(t))}{\partial \vec{r}_i(t)} + \overline{R_k(\vec{r}^N(t))}^t \frac{\partial \overline{P_k(\vec{r}^N(t))}^t}{\partial P_k(\vec{r}^N(t))} \frac{\partial P_k(\vec{r}^N(t))}{\partial \vec{r}_i(t)}\right\} \quad .(23)$$

In a simulation, the atomic configurations are separated by a time step or interval Δt, so in discretised form we have for the n^{th} time step for the time-averaged quantities in Eq. (23)

$$\overline{R_k(n\Delta t)}^{n\Delta t,\exp} = R_k(n\Delta t)(1-\exp(-\Delta t/\tau_r^{\mathrm{RDC}})) + \exp(-\Delta t/\tau_r^{\mathrm{RDC}})\, \overline{R_k((n-1)\Delta t)}^{(n-1)\Delta t,\exp}, \quad (24)$$

and

$$\overline{P_k(n\Delta t)}^{n\Delta t,\exp} = P_k(n\Delta t)\left(1-\exp(-\Delta t/\tau_\theta^{\mathrm{RDC}})\right) + \exp(-\Delta t/\tau_\theta^{\mathrm{RDC}})\, \overline{P_k((n-1)\Delta t)}^{(n-1)\Delta t,\exp}. \quad (25)$$

Using these equations

$$\frac{\partial \overline{R_k(n\Delta t)}^{n\Delta t,\exp}}{\partial R_k(n\Delta t)} = 1-\exp(-\Delta t/\tau_r^{\mathrm{RDC}}), \quad (26)$$

and

$$\frac{\partial \overline{P_k(n\Delta t)}^{n\Delta t,\exp}}{\partial P_k(n\Delta t)} = 1 - \exp(-\Delta t/\tau_\theta^{\mathrm{RDC}}). \tag{27}$$

Generally, the memory relaxation times τ_r^{RDC} and $\tau_\theta^{\mathrm{RDC}}$ are chosen to be much longer than the MD or SD integration time step Δt,

$$\Delta t \ll \tau_r^{\mathrm{RDC}} \ll \tau_\theta^{\mathrm{RDC}} \ll t^{\mathrm{MD,SD}} \ , \tag{28}$$

where $t^{\mathrm{MD,SD}}$ is the length of the MD or SD simulation. This means that the derivatives Eq. (26) and Eq. (27) can be approximated by $\Delta t/\tau_r^{\mathrm{RDC}}$ and $\Delta t/\tau_\theta^{\mathrm{RDC}}$, respectively.

For the other derivatives we get using Eq. (12)

$$\frac{\partial R_k(\vec{r}^N(t))}{\partial \vec{r}_i(t)} = (\delta_{ik_1} - \delta_{ik_2})(-3(r^0)^3)(r_{k_1k_2}(t))^{-5}\vec{r}_{k_1k_2}(t), \tag{29}$$

and using Eq. (13) and the definition

$$\theta_k \equiv \arccos\left(\frac{z_{k_1k_2}}{r_{k_1k_2}}\right), \tag{30}$$

we find

$$\frac{\partial P_k(\vec{r}^N(t))}{\partial x_i(t)} = -(\delta_{ik_1} - \delta_{ik_2})3\left(\frac{z_{k_1k_2}(t)}{r_{k_1k_2}(t)}\right)^2 (r_{k_1k_2}(t))^{-2}x_{k_1k_2}(t), \tag{31a}$$

$$\frac{\partial P_k(\vec{r}^N(t))}{\partial y_i(t)} = -(\delta_{ik_1} - \delta_{ik_2})3\left(\frac{z_{k_1k_2}(t)}{r_{k_1k_2}(t)}\right)^2 (r_{k_1k_2}(t))^{-2}y_{k_1k_2}(t), \tag{31b}$$

$$\frac{\partial P_k(\vec{r}^N(t))}{\partial z_i(t)} = -(\delta_{ik_1} - \delta_{ik_2})3\left(\frac{z_{k_1k_2}(t)}{r_{k_1k_2}(t)}\right)(r_{k_1k_2}(t))^{-3}(z_{k_1k_2}^2(t) - r_{k_1k_2}^2(t)). \tag{31c}$$

We note that

$$\frac{\partial R_k}{\partial \vec{r}_{k_1}} = -\frac{\partial R_k}{\partial \vec{r}_{k_2}}, \tag{32}$$

and

$$\frac{\partial P_k}{\partial \vec{r}_{k_1}} = -\frac{\partial P_k}{\partial \vec{r}_{k_2}}. \tag{33}$$

In the case that all vectors $\vec{r}_{k_1k_2}$ have a fixed length $r_{k_1k_2}$, i.e., the distance is constrained, the derivative Eq. (29) becomes zero,

$$\frac{\partial R_k(\vec{r}^N(t))}{\partial \vec{r}_i(t)} = 0, \tag{34}$$

and Eqs. (31a-c) become

$$\frac{\partial P_k(\vec{r}^N(t))}{\partial x_i(t)} = \frac{\partial P_k(\vec{r}^N(t))}{\partial y_i(t)} = 0, \tag{35}$$

and

$$\frac{\partial P_k(\vec{r}^N(t))}{\partial z_i(t)} = (\delta_{ik_1} - \delta_{ik_2})3\left(\frac{z_{k_1k_2}(t)}{r_{k_1k_2}}\right)(r_{k_1k_2})^{-1}. \tag{36}$$

The restraining force is an unphysical force, for which reason it is usually not contributing to the virial and pressure in GROMOS[28]. It also should and will be rather small. Only when many (strong) restraints are applied, as in the case of the use of distance restraints to separate fine-grained and coarse-grained water around a solute, the virial contribution of the restraining forces must be accounted for. Otherwise, the computational box becomes unstable[29]. Since the restraining force depends on the angle θ_k of the vector $\vec{r}_{k_1k_2}$ with the z-axis of the coordinate system, the restraining function generates a rotational force on the vector $\vec{r}_{k_1k_2}$, not a translational one. The scalar pressure, i.e., trace of the pressure tensor is thus zero[30,31]. When applying isotropic pressure scaling, the restraining forces do not contribute to the pressure, so their contributions to the virial need not be calculated, as is the case if no pressure scaling, e.g., in vacuo, is applied.

Since the restraining forces are *not* internal in the molecule, they depend on the direction of the z-axis of the coordinate system, their contribution to the virial tensor $W_{\alpha\beta}$ must be calculated in case of anisotropic pressure scaling, not only when an atom-based virial is calculated, but also when a pressure-group based, i.c. molecular, virial is calculated[29],

$$W_{k_{\alpha\beta}} = -\frac{1}{2}\left((\vec{r}_{k_1})_\alpha(\vec{f}_{k_1})_\beta + (\vec{r}_{k_2})_\alpha(\vec{f}_{k_2})_\beta\right) \qquad (\alpha, \beta = x, y, z) \tag{37}$$

with the kinetic energy tensor

$$K_{k_{\alpha\beta}} = \frac{1}{2}\left(m_{k_1}(\vec{v}_{k_1})_\alpha(\vec{v}_{k_1})_\beta + m_{k_2}(\vec{v}_{k_2})_\alpha(\vec{v}_{k_2})_\beta\right) \qquad (\alpha, \beta = x, y, z) \qquad (38)$$

and the pressure tensor

$$P_{\alpha\beta} = \sum_{k=1}^{N_{\mathrm{RDC}}} 2\left(K_{k_{\alpha\beta}} - W_{k_{\alpha\beta}}\right) V^{-1} \qquad (\alpha, \beta = x, y, z) \qquad (39)$$

Using the relation $\vec{f}_{k_2} = -\vec{f}_{k_1}$, the contribution of the k^{th} restraining force to the virial becomes

$$W_{k_{\alpha\beta}} = -\frac{1}{2}\left((\vec{r}_{k_1k_2})_\alpha(\vec{f}_{k_1})_\beta\right). \qquad (\alpha, \beta = x, y, z) \qquad (40)$$

We note that the positions of the particles k_1 and k_2 should satisfy the minimum image condition.

Generating an Anisotropic Rotational Distribution

In order to test the RDC restraining algorithm a simulation generating a non-isotropic rotational distribution should be performed. This can be achieved by introducing a biasing force in a simulation.

Choice of Biasing Function

In an MD or SD simulation the angular distribution $P(\theta_{\mathrm{ab}})$ of an axis 'ab' in the molecule with respect to a hypothetical magnetic field direction $\vec{H}$, e.g. the *z*-direction, can be made (slightly) anisotropic by adding a biasing potential energy function $V^{\mathrm{bias}}\left(\theta_{\mathrm{bias}}; K^{\mathrm{bias}}, \theta^0_{\mathrm{bias}}\right)$ to the potential energy function of the molecular system, that biases the distribution $P(\theta_{\mathrm{bias}})$ towards a value θ^0_{bias}. The extent of the biasing is controlled by a strength parameter or force constant K^{bias}. Such a function V^{bias} should be continuous for all $\theta_{\mathrm{bias}} \in [0,\pi]$ and have a continuous derivative with respect to θ_{bias} on this domain.

One could define the axis 'ab' by specifying two atoms 'a' and 'b' in the molecule. The angle θ_{bias} is then the angle θ_{abz} of the vector $\vec{r}_{\mathrm{ab}} \equiv \vec{r}_{\mathrm{a}} - \vec{r}_{\mathrm{b}}$ with the z-axis,

$$\theta_{\mathrm{abz}} \equiv \arccos\left(\frac{z_{\mathrm{ab}}}{r_{\mathrm{ab}}}\right) \ . \tag{41}$$

The biasing potential energy term could e.g., be chosen as

$$V^{\mathrm{bias}}\left(\theta_{\mathrm{abz}}; K^{\mathrm{bias}}, \theta^0_{\mathrm{bias}}\right) = \frac{1}{2} K^{\mathrm{bias}}\left(\theta_{\mathrm{abz}} - \theta^0_{\mathrm{bias}}\right)^2 \ . \tag{42}$$

The biasing forces $\vec{f}_{\mathrm{a}}(t)$ and $\vec{f}_{\mathrm{b}}(t)$ on the two atoms are then

$$f_{\mathrm{a}x}(t) = -K^{\mathrm{bias}}(\theta_{\mathrm{abz}}(t) - \theta^0_{\mathrm{bias}})(r^2_{\mathrm{ab}}(t) - z^2_{\mathrm{ab}}(t))^{-1/2} r^{-2}_{\mathrm{ab}}(t) z_{\mathrm{ab}}(t) x_{\mathrm{ab}}(t) \ , \tag{43a}$$

$$f_{\mathrm{a}y}(t) = -K^{\mathrm{bias}}(\theta_{\mathrm{abz}}(t) - \theta^0_{\mathrm{bias}})(r^2_{\mathrm{ab}}(t) - z^2_{\mathrm{ab}}(t))^{-1/2} r^{-2}_{\mathrm{ab}}(t) z_{\mathrm{ab}}(t) y_{\mathrm{ab}}(t) \ , \tag{43b}$$

$$f_{\mathrm{a}z}(t) = -K^{\mathrm{bias}}(\theta_{\mathrm{abz}}(t) - \theta^0_{\mathrm{bias}})(r^2_{\mathrm{ab}}(t) - z^2_{\mathrm{ab}}(t))^{-1/2} r^{-2}_{\mathrm{ab}}(t) (z^2_{\mathrm{ab}}(t) - r^2_{\mathrm{ab}}(t)) \ , \tag{43c}$$

and

$$\vec{f}_{\mathrm{b}}(t) = -\vec{f}_{\mathrm{a}}(t) \ . \tag{44}$$

If one would like to avoid the use of the arccosine function, one could use

$$V^{\mathrm{bias}}\left(\theta_{\mathrm{abz}}; K^{\mathrm{bias}}, \theta^0_{\mathrm{bias}}\right) = \frac{1}{2} K^{\mathrm{bias}}(\cos(\theta_{\mathrm{abz}}) - \cos(\theta^0_{\mathrm{bias}}))^2 = \frac{1}{2} K^{\mathrm{bias}}\left(\left(\frac{z_{\mathrm{ab}}}{r_{\mathrm{ab}}}\right) - \cos(\theta^0_{\mathrm{bias}})\right)^2 \tag{45}$$

with

$$f_{\mathrm{a}x}(t) = +K^{\mathrm{bias}}\left(\left(\frac{z_{\mathrm{ab}}(t)}{r_{\mathrm{ab}}(t)}\right) - \cos(\theta^0_{\mathrm{bias}})\right) r^{-3}_{\mathrm{ab}}(t) z_{\mathrm{ab}}(t) x_{\mathrm{ab}}(t) \ , \tag{46a}$$

$$f_{\mathrm{a}y}(t) = +K^{\mathrm{bias}}\left(\left(\frac{z_{\mathrm{ab}}(t)}{r_{\mathrm{ab}}(t)}\right) - \cos(\theta^0_{\mathrm{bias}})\right) r^{-3}_{\mathrm{ab}}(t) z_{\mathrm{ab}}(t) y_{\mathrm{ab}}(t) \ , \tag{46b}$$

$$f_{\mathrm{a}z}(t) = +K^{\mathrm{bias}}\left(\left(\frac{z_{\mathrm{ab}}(t)}{r_{\mathrm{ab}}(t)}\right) - \cos(\theta^0_{\mathrm{bias}})\right) r^{-3}_{\mathrm{ab}}(t) (z^2_{\mathrm{ab}}(t) - r^2_{\mathrm{ab}}(t)), \tag{46c}$$

and

$$\vec{f}_{\mathrm{b}}(t) = -\vec{f}_{\mathrm{a}}(t). \tag{47}$$

If r_{ab} is kept fixed, the biasing forces Eq. (43) reduce to

$$f_{\mathrm{a}x}(t) = f_{\mathrm{a}y}(t) = 0, \tag{48a}$$

$$f_{az}(t) = +K^{\text{bias}}(\theta_{abz}(t) - \theta^0_{\text{bias}})(r^2_{ab} - z^2_{ab}(t))^{-1/2} \ , \tag{48b}$$

and the biasing forces Eq. (46) reduce to

$$f_{ax}(t) = f_{ay}(t) = 0 \ , \tag{49a}$$

$$f_{az}(t) = -K^{\text{bias}}\left(\left(\frac{z_{ab}(t)}{r_{ab}}\right) - \cos(\theta^0_{\text{bias}})\right) r^{-1}_{ab} \ . \tag{49b}$$

We note that the anisotropy introduced in $P(\theta_{abz})$ by a biasing potential energy term $V^{\text{bias}}\left(\theta_{abz}; K^{\text{bias}}, \theta^0_{\text{bias}}\right)$ will unlikely be the same as induced by experiment. However, different atom pairs (a,b) in a molecule may induce different anisotropies, so can be used to investigate different anisotropies and their effects on RDCs.

The biasing force is a real force, i.e., also in some way present in experiment, so should contribute to the virial and the pressure. In an experiment to measure RDCs the uniform rotational distribution of the molecule is only slightly perturbed. The RDC values are of the order of Hz. So, the biasing force should only be a rather small perturbation of the forces in the system. Its contribution to the virial tensor $W_{\alpha\beta}$ is

$$W_{ab_{\alpha\beta}} = -\frac{1}{2}\left((\vec{r}_a)_\alpha(\vec{f}_a)_\beta + (\vec{r}_b)_\alpha(\vec{f}_b)_\beta\right) \qquad (\alpha, \beta = x, y, z) \tag{50}$$

Using the relation $\vec{f}_b = -\vec{f}_a$, the contribution of the biasing force to the virial becomes

$$W_{ab_{\alpha\beta}} = -\frac{1}{2}\left((\vec{r}_{ab})_\alpha(\vec{f}_a)_\beta\right). \qquad (\alpha, \beta = x, y, z) \tag{51}$$

We note that the positions of the particles ‘a’ and ‘b’ should satisfy the minimum image condition.

Implementation, Molecular Models, Simulation Set-Up and Analysis

The simulations were performed using the GROMOS simulation software package[32-37]. When treating realistic molecular systems, the use of Standard

International (SI) units is recommended. Apart from restrictions when storing or printing data in non-exponential format, the GROMOS programs are independent of the chosen units. The units are defined by the ones used for the physical constants and atomic and molecular quantities to be specified in the PHYSICALCONSTANTS block[38] in the GROMOS data files[39]. It is recommended to use as basic units nanometer (nm) for length, atomic mass unit (u) for mass, picosecond (ps) for time, Kelvin (K) for temperature, and electronic charge (e) for charge. These basic units then determine the units of other quantities, e.g. kJ/mol for energy, kJ/(mol nm) for force, kJ/(mol nm^3) for pressure, and THz (ps^{-1}) for frequency. If, for example, the non-GROMOS-recommended unit Hz instead of THz is to be used as input unit for dipolar couplings and $kJmol^{-1}Hz^{-2}$ for the RDC-restraining force constant, the scaling factor 10^{12} should be specified in the RDCRESTRAINTS block.

Another implementation issue is the calculation of the factor $(1-\exp(-\Delta t/\tau_D^{\mathrm{RDC}}))$, or rather $(1-\exp(-\Delta t/\tau_r^{\mathrm{RDC}}))$ and $(1-\exp(-\Delta t/\tau_\theta^{\mathrm{RDC}}))$ occurring in Eqs. (18, 20, 21, 24-27) of the algorithm. When $\Delta t/\tau_D^{\mathrm{RDC}}$ (or $\Delta t/\tau_r^{\mathrm{RDC}}$ or $\Delta t/\tau_\theta^{\mathrm{RDC}}$) becomes very small, the factor $(1-\exp(-\Delta t/\tau_D^{\mathrm{RDC}}))$ gets close to zero. Thus, for small values of $x = \Delta t/\tau_D^{\mathrm{RDC}}$ (or $\Delta t/\tau_r^{\mathrm{RDC}}$ or $\Delta t/\tau_\theta^{\mathrm{RDC}}$), the factors $(1-\exp(-\Delta t/\tau_D^{\mathrm{RDC}}))$, (or $(1-\exp(-\Delta t/\tau_r^{\mathrm{RDC}}))$ or $(1-\exp(-\Delta t/\tau_\theta^{\mathrm{RDC}}))$) should be evaluated using the power series expansion[40]

$$(1-\exp(-x)) = x - 1/2x^2 + 1/6x^3 - 1/24x^4 + 1/120x^5 - 1/720x^6 + 1/5040x^7 - O(x^8)\,. \quad (52)$$

For a 48-bit mantissa accuracy the analytical formula (1-exp(-*x*)) is accurate to better

than $1 : 10^6$ when $x = \Delta t/\tau_D^{\mathrm{RDC}} > 0.05$. For $x < 0.05$ this accuracy will be maintained if the expansion given above is used[40].

Use of Simple "Toy" Models to Optimise the Testing of a Method, Model or Algorithm

When a new method is proposed, its performance and range of applicability as function of the values of its parameters are to be evaluated. Use of a real molecular system, for the current RDC-restraining method a protein, has some disadvantages:

1. The measured (RDC) data may be inconsistent with each other due to measurement errors. An example of such a situation for ^3J-couplings, which was detected using a simple "toy" model, can be found in Ref.[41]
2. Only a limited set of measured data or properties (observables) may be available, i.e. can be measured. This restricts the extent of testing in regard to varying the number and type of properties.
3. In contrast to a computational experiment, in a real experiment many quantities cannot be measured, in regard to RDCs for example, the non-uniform rotational distribution of the molecule, which leads to observable RDC values in the Hz range, is not experimentally measurable.
4. In a large molecule, such as a protein, many physical factors will determine the structure and dynamics of a molecule, for example its atomic composition, its covalent structure, its electrostatic interactions, its solvent or crystal environment, etc. The presence of such factors may complicate a straightforward interpretation of an observed correlation between a

parameter from the method or model and the measured data.

So, in order to optimise the testing of a method or model, the (molecular) test system should be chosen as simple as possible without omitting features or degrees of freedom that are fundamental to the method or model. In addition, the smaller the molecular test system, the less computational effort is required to evaluate the many combinations of parameter values of the method or model to be tested.

All models or theories are an abstraction of reality and, as such, have limited applicability. This means that an essential element of any method or model is the conditions or parameter values for which it is valid and knowledge of when it fails to properly describe a given phenomenon. In other words, an attempt to falsify a given model or theory by testing systematically the assumptions and range of parameter values on which the method or model is based, is an essential element of any publication of its characteristics. This generally requires a great many tests, a substantial computational effort.

Ethane

A most simple molecular model to test the proposed algorithm is a molecule of two united atoms, ethane, with a rigid bond. Of the two atoms the masses are 15.035 u (a CH_3 united atom in GROMOS) and the electrostatic partial charges are zero. The length of the bond is 0.153 nm. Taking r^0 = 0.153 nm, $R_k(t)$, Eq. (12), equals 1. The rigid bond vector can be used as RDC, with the gyromagnetic ratio of a ^{13}C atom (see Table 1), so the value of D_k^c, Eq. (11), becomes -4.243 kHz. The rigid bond vector can also be used as vector $\vec{r}_{ab}$, of which the direction is to be influenced by the biasing force.

Cyclo-Octane with Side-Chain CH_2-CH_3 Branches

A cyclo-octane molecule with side-chain CH_2-CH_3 branches at all eight CH_1 united atoms of the ring was used as second simple test system, see Figure 2. It allows for the number of RDC bond vectors to be varied, between 1 and 24 bonds. It also allows for investigating the effect of the degree of intramolecular flexibility on the calculated RDCs, while still being sufficiently small to thoroughly sample the rotational degrees of freedom of the molecule in an SD simulation. In addition, the 24 bonds in the molecule are differently oriented, see Table 2, and thus will generate different $<D_k>$ values, depending on their angle θ_k with the *z*-axis, i.e., the direction of the magnetic field.

The GROMOS 54A7 or 54B7 set of force-field parameters was used[26]. The CHn united atoms carry zero (partial) electrostatic charge, so no electrostatic interaction is present. The molecular topology file contains 16 CHn–CHn torsional dihedral-angle potential energy terms and seven improper dihedral-angle energy terms at the chiral centres of the ring (atoms 3, 7, 10, 11, 15, 16, and 18 in Figure 2, see Supporting Information Table S1). The absence of an improper dihedral angle at the chiral centre atom 4 in the ring allows some flexibility of the ring. The length of the CHn-CHn bonds in this force field is 0.153 nm. Taking r^0 = 0.153 nm, $R_k(t)$, Eq. (12), equals 1. Using the gyromagnetic ratio of a ^{13}C atom (see Table 1), the value of D_k^c, Eq. (11), becomes -4.243 kHz. In order to investigate the effect of enhanced molecular flexibility, an SD simulation was performed, in which the torsional-angle potential-energy terms in the

force field were omitted.

Simulation Set-Up

The molecules were simulated in vacuo using stochastic dynamics. The SD leap-frog algorithm[28,40] was used with a time step Δt = 2 fs, a reference temperature of 300 K, and using different values of the friction coefficients γ_i^{SD}, which were taken the same for all united atoms[42]. The size of the friction coefficients was varied, from 0.0024 ps^{-1} to 24 ps^{-1}, in order to determine the value that would maximise the extent of rotational sampling. All bonds in the molecules were kept fixed using the SHAKE algorithm[43] with a relative geometric precision of 10^{-4}. θ-angle orientation-biasing SD simulations were performed using as biasing force the cosine-dependent form, Eqs. (45-47). The values of the biasing reference angle θ_{bias}^0 and of the biasing force constant K^{bias} were varied in order to investigate their effect upon the orientation distribution of the molecule. The anisotropy induced should be small and the effect of choosing θ_{bias}^0 close to the magic angle is to be investigated.

RDC Restraining

The RDC restraining was performed using the algorithm described above. Since the bonds in the molecules were kept fixed, no averaging of the RDC distances $r_{k_1k_2}$ was required. Only the RDC angles θ_k in Eq. (19) were subject to restraining. The force constants K^{RDC}, the flat-bottom range ΔD^0, the memory relaxation time $\tau_\theta^{\mathrm{RDC}}$ and the number N_{RDC} of RDCs used in the restraining were varied. By using the target RDC values D_k^0 generated using a particular force field, for example the GROMOS force field 54A7 without torsional-angle energy terms, in an RDC restraining SD simulation,

in which these torsional-angle energy terms are present, the effect of a force-field perturbation or a force-field deficiency upon the structures and their variation (flexibility) obtained by RDC restraining can be investigated.

MD Simulations Performed

The following types of SD simulations were performed.

1. Unbiased and unrestrained simulations, aimed at determining the extent and convergence of molecular rotational sampling for various values of the atomic friction coefficients γ_i^{SD} and SD simulation lengths t^{SD}.
2. *z*-axis biased (and unrestrained) simulations using a biasing force dependent on the angle θ_{abz} of a vector $\vec{r}_{\mathrm{ab}} = \vec{r}_{\mathrm{a}} - \vec{r}_{\mathrm{b}}$, defined by two atoms 'a' and 'b' in the molecule, with the *z*-axis. The biasing target angle θ_{bias}^0 and the biasing force constant K^{bias} were varied. The GROMOS force field 54A7 with and without torsional-angle energy terms was used. These simulations are indicated using the notation *bSD* and *bntSD*, respectively.
3. RDC-restraining simulations using various numbers of RDC restraints N_{RDC}, sets of RDC restraint target values D_k^0, restraining force constants K^{RDC}, memory relaxation times $\tau_\theta^{\mathrm{RDC}}$ and (possibly) flat-bottom parameters ΔD^0. These simulations are indicated using the notation *rSD*. In addition, RDC-restraining simulations were performed in which the target values D_k^0 were generated using a force field that differs from the one used in the restraining, this in order to investigate force-field deficiencies of the molecular model used in RDC structure refinement.

Analysis of Atomic Trajectories and RDC Values

Dipolar couplings D_k were calculated using Eq. (10) during the simulations at every time step in order to obtain a high-resolution sampling $\langle D_k \rangle$ value when averaging in Eq. (14), while trajectory atomic coordinates were stored at 2 ps intervals for post-simulation analysis[36]. In the RDC-restraining simulations, the dipolar couplings D_k calculated at every time step were also used to obtain the exponentially damped time averages $\overline{D_k}^{t,\exp}$, or rather $\overline{P_k}^{t,\exp}$ and possibly $\overline{R_k}^{t,\exp}$, of Eqs. (20) and (21), which are used in the RDC-restraining potential-energy term, Eqs. (16) and (17), and force, Eqs. (22) and (23). These exponentially damped RDC values will thus slightly differ from the ones calculated without exponential damping. In order to analyse all trajectories in the same way, when reporting RDC values, Eq. (14) is used for both the unrestrained and the restrained simulations.

RESULTS AND DISCUSSION

Generating Orientation Distributions for Ethane

Figures 3 and 4 show the distributions $P(\theta_{ab})/\sin(\theta_{ab})$ and $P(\varphi_{ab})$ for ethane in unbiased and unrestrained SD simulations using different friction coefficients γ_i^{SD} and SD simulation periods t^{SD}, where φ_{ab} is the angle of the projection of the bond vector onto the *x*-*y*-plane and the *x*-axis, and θ_{ab} is the angle of the bond vector with the *z*-axis of the (right-handed) coordinate system. Figure 5 shows the RDC values $\langle D_{ab} \rangle_t$ as function of time *t*. The $\langle D_{ab} \rangle_t$ values after averaging over a time period t^{SD}, $\langle D_{ab} \rangle_t^{SD}$,

are shown in Table 3.

The lengths t^{SD} of the orientation-biasing and RDC-restraining SD simulations should be chosen such that the $<D_k>_t^{SD}$-values are, due to rotational sampling, well converged to values that are a factor 10^3 or 10^4 smaller than the D_k^c-values, and in addition generate statistically precise $<D_k>_t^{SD}$-values. The data in Table 3 and Figures 3 to 5 suggest for ethane an optimal friction coefficient γ_i^{SD} of 2.4 ps^{-1} and a simulation length t^{SD} of at least 2 μs. The values γ_i^{SD}= 2.4 ps^{-1} and t^{SD} = 4 μs will be used in the simulations of ethane that are performed to investigate the effects of the various parameters of the orientation-biasing force and the RDC-restraining force on the orientation distribution and the calculated RDC-values $<D_k>$.

Generating Orientation Distributions for Cyclo-Octane

Figures 6 and 7 show the distributions $P(\theta_{ab})/\sin(\theta_{ab})$ and $P(\varphi_{ab})$ for the vector $\vec{r}_{ab}$ connecting atoms a = 11 and b = 4 (large arrow in Figure 2) in unbiased and unrestrained SD simulations of cyclo-octane using different friction coefficients γ_i^{SD} and SD simulation periods t^{SD}, where φ_{ab} is the angle of the projection of the bond vector onto the *x*-*y*-plane and the *x*-axis, and θ_{ab} is the angle of the bond vector with the *z*-axis of the (right-handed) coordinate system. Figure 8 shows the 24 RDC values $<D_k>_t$ as function of time *t*. The $<D_k>_t$ values after averaging over a time period t^{SD}, $<D_k>_t^{SD}$, are shown in Table 4.

Using the criteria mentioned before, the data in Table 4 and Figures 6 to 8 suggest an optimal friction coefficient γ_i^{SD} of 0.24 ps^{-1} and a simulation length t^{SD} of at least 5 μs for cyclo-octane. The values γ_i^{SD}= 0.24 ps^{-1} and t^{SD} = 20 μs will be used in the

simulations of the cyclo-octane molecule that are performed to investigate the effects of the various parameters of the orientation-biasing force and the RDC-restraining force on the orientation distribution, the calculated RDC-values $<D_k>$, and the molecular structures resulting from the refinement.

While for the two-(united)atom molecule ethane γ_i^{SD}= 2.4 ps^{-1} appears to optimise the rotational sampling, for the 24-(united)atom molecule cyclo-octane (with eight side chains) a lower friction, γ_i^{SD}= 0.24 ps^{-1}, and thus smaller stochastic forces, seems optimal. For a larger molecule, 24 instead of two atoms, the stochastic forces inducing rotation are distributed over more atoms, which may allow them to be chosen smaller. In addition, the inertial forces inducing rotation may also be larger, allowing smaller stochastic forces too.

Generating RDCs for Ethane

Figure 9 shows the distribution $P(\theta_{ab})/\sin(\theta_{ab})$ for ethane in orientation-biasing *bSD* simulations using different biasing target angles θ_{bias}^0 and biasing force constants K^{bias}. Figure 10 shows the RDC values $<D_{ab}>_t$ as function of time. The $<D_{ab}>_t$ values at the end of the 4 μs simulations, $<D_{ab}>_t^{SD}$, are shown in Table 5.

As can be seen in Figure 1, the second-order Legendre function $P(\theta)$ is symmetric around θ = 90°. Yet, some of the generated θ-angular distributions do not show this symmetry due to the biasing towards only one of the two equivalent biasing reference angles θ_{bias}^0 and 180° - θ_{bias}^0. The symmetry can be obtained by switching the value of the reference biasing reference angle from θ_{bias}^0 to 180° - θ_{bias}^0 half-way through the simulation. Table 5 shows that this does not lead to significantly different $<D_{ab}>_t^{SD}$

values. Figure 11 shows the symmetrised θ-distributions of the *bSD* simulation for the period 1 – 4 μs.

The value of the biasing reference angle θ^0_{bias} influences neither the $<D_{ab}>_t^{SD}$ value (Table 5) nor the θ-distribution (columns in Figure 11) much for the values of the biasing force constant K^{bias} shown. Obviously, for larger biasing force constants K^{bias} the θ-distributions will show peaks at the respective θ^0_{bias} angles. But, the larger the biasing force constant K^{bias} , the larger the $<D_{ab}>_t^{SD}$ values will be. For proteins, the absolute values of measured RDCs lie in the range [0.1 – 30] Hz. In order to obtain such $<D_{ab}>_t^{SD}$ values for ethane, the values K^{bias} = 0.2 kJ/mol and θ^0_{bias} = 0° will be used in further simulations of ethane.

Generating RDCs for Cyclo-Octane

Figure 12 shows the distribution $P(\theta_k)/\sin(\theta_k)$ for the vector $\vec{r}_{\mathrm{ab}}$ connecting atoms a = 11 and b = 4 (large arrow in Figure 2) in cyclo-octane in a biasing *bSD* simulation (dotted lines) with biasing target angle θ^0_{bias}= 0° and biasing force constant K^{bias} = 0.2 kJmol^{-1}. Figure 13 shows the RDC values $<D_k>_t$ as function of time for 10 different bond vectors (indicated by open arrows in Figure 2) for different K^{bias} (line types) and θ^0_{bias}(rows) values. Table 6 shows the generated $<D_k>_t^{SD}$-values.

The value of the biasing reference angle θ^0_{bias} does not influence the $<D_k>_t^{SD}$ values much. Thus, in further simulations the value θ^0_{bias} = 0° will be used. Obviously, the larger the biasing force constant K^{bias} , the larger the $<D_k>_t^{SD}$ values will be. For K^{bias} = 0.1 kJ/mol, their range is [-7,+10] Hz. For K^{bias} = 0.2 kJ/mol, their range is [-12,+19] Hz, and for K^{bias} = 0.5 kJ/mol, their range is [-31,+43] Hz. In order to obtain

$<D_k>_t^{SD}$ values of the order of 10 Hz, the value K^{bias} = 0.2 kJ/mol will be used in further simulations of cyclo-octane.

Recovering RDCs for Ethane

Table 7 shows the <*D*>-values for ethane obtained by RDC-restraining towards the D^0 target RDC-values generated using four different orientation-biasing forces (K^{bias} = 0.05, 0.1, 0.2, 0.5 kJ/mol, θ^0_{bias} = 0°) for eight combinations of values of the restraining-force constant K^{RDC} (K^{RDC} = 0.1 – 100 kJmol^{-1}Hz^{-2}) and the memory relaxation time τ_θ^{RDC} (τ_θ^{RDC} = 0.1 – 10 ns).

Since the memory relaxation time τ_θ^{RDC} is orders of magnitude larger than the simulation time step Δt, the function (1-exp(-$\Delta t/\tau_\theta^{RDC}$)) in Eqs. 26 and 27 is well approximated by $\Delta t /\tau_\theta^{RDC}$ and thus is the restraining force roughly proportional to $K^{RDC} /\tau_\theta^{RDC}$. This is illustrated by the values in Table 7. For τ_θ^{RDC} = 0.1 ns, using K^{RDC} = 0.1 kJmol^{-1}Hz^{-2}, the restraining force is not strong enough to bring the <*D*> value close to the D^0 target value. Increasing K^{RDC} to 1 kJmol^{-1}Hz^{-2} the <*D*> value gets closer, but for K^{RDC} = 10 kJmol^{-1}Hz^{-2}, no further significant improvement is obtained. Using τ_θ^{RDC} = 1 ns and K^{RDC} = 10 or 100 kJmol^{-1}Hz^{-2}, the <*D*> value comes close to the D^0 target value. Lengthening the memory relaxation time to τ_θ^{RDC} = 10 ns or in addition increasing K^{RDC} to 1000 kJmol^{-1}Hz^{-2} worsens the result. In order to approximate the D^0 target RDC-value for ethane well, τ_θ^{RDC} should be at least 1 ns and K^{RDC} as large as 10 or 100 kJmol^{-1}Hz^{-2}.

Recovering Orientation Distributions for Ethane

When the orientation-biasing generated RDC target value is approximated well in

the RDC-restraining simulation, this does not necessarily imply that the underlying θ-angle distribution generated in the orientation-biasing *bSD* simulation must be reproduced well too in the RDC-restraining *rSD* simulation. Figure 11 shows both θ-distributions, *bSD* ones (dotted lines) for four biasing force constants K^{bias} (rows) and three biasing reference angles θ^0_{bias}(columns), and the *rSD* ones for two restraining force constants K^{RDC} (dashed and solid lines, 10 and 100 kJmol^{-1}Hz^{-2}, respectively) and memory relaxation time τ_θ^{RDC} = 1 ns. The *bSD* θ-distributions are rather well approximated by the *rSD* ones.

Recovering RDCs for Cyclo-Octane

Table 8 shows the $< D_{k_1k_2} >_{t^{\text{SD}}}$-values obtained for cyclo-octane as calculated from RDC-restraining *rSD* simulations for all 24 bond vectors in cyclo-octane $\vec{r}_{k_1k_2} = \vec{r}_{k_1} - \vec{r}_{k_2}$ (see Figure 2), and for six combinations of the RDC-restraining function parameter values K^{RDC} and τ_θ^{RDC}. The D_k^0 target values were obtained from orientation-biasing *bSD* simulations with K^{bias} = 0.2 kJ/mol and θ^0_{bias} = 0°.

For τ_θ^{RDC} = 1 ns and K^{RDC} = 1 kJmol^{-1}Hz^{-2}, the D_k^0 target values are approximated well, using K^{RDC} = 2 kJmol^{-1}Hz^{-2} even slightly better, whereas for K^{RDC} = 3 kJmol^{-1}Hz^{-2} the simulations failed due to too strong restraining forces that induce local deformations that cannot be corrected by the procedure SHAKE that maintains the bond-length constraints. In SHAKE there is a limit to the maximum displacement, induced by the unconstrained forces, allowed for the atoms involved in each individual constraint. This local convergence criterion means that SHAKE will fail to converge when the forces acting on the specific atoms in a given constraint become very large.

Thus, a SHAKE failure can be used to detect an error in the simulation, specifically the presence and location of unphysically large forces[44]. Reducing the restraining force by lengthening $\tau_\theta^{\mathrm{RDC}}$ from 1 ns to 10 ns the simulations do not fail, but the $<D_k>$ values do not get closer to the D_k^0 target values. In further simulations of cyclo-octane that are performed to investigate the effects of varying the number of RDCs, N_{RDC}, $\tau_\theta^{\mathrm{RDC}}$ = 1 ns will be used and K^{RDC} will be varied, because the size of the restraining forces in comparison to the size of the force-field forces will not only depend on K^{RDC}, but also on N_{RDC}. When investigating the effect of the flexibility of the force field used on the recovery of the D_k^0 target RDC-values by the restraining, $\tau_\theta^{\mathrm{RDC}}$ = 1 ns and K^{RDC} = 2 kJmol^{-1}Hz^{-2} will be used.

Recovering Orientation Distributions for Cyclo-Octane

For cyclo-octane the orientation-biasing generated RDC target values are approximated well in the RDC-restraining *rSD* simulations. The underlying *rSD* θ-angle distributions are shown in Figure 12, together with the *bSD* one (dotted lines), for three restraining force constants K^{RDC} (different line types) and two memory relaxation times $\tau_\theta^{\mathrm{RDC}}$ (1 and 10 ns, columns). The *bSD* θ-distribution is rather well approximated by the *rSD* ones, in particular for $\tau_\theta^{\mathrm{RDC}}$ = 1 ns.

Influence of the Number of RDCs Used in Structure Refinement

When recovering RDCs, all 24 bond vectors in cyclo-octane were used. It would be interesting to investigate how well RDCs are reproduced when only a fraction of these are used as targets in RDC-restraining SD simulations. Table 9 shows RDC $< D_{k_1k_2} >_{t^{\mathrm{SD}}}$ values obtained for cyclo-octane as calculated from RDC-restraining

rSD simulations in which N_{RDC} = 24, 12, 10, 4 or 2 bond-vectors are restrained.

Keeping K^{RDC} constant, lowering N_{RDC} leads to $<D_k>$ values less well approximating the target D_k^0 values. This would suggest to increase the value of the restraining force constant K^{RDC} for lower N_{RDC} values, but that leads to failure of the simulations due to locally too large restraining forces. Keeping N_{RDC} constant, increasing the value of K^{RDC} will bring the $<D_k>$ values closer to the target D_k^0 values, but for too large K^{RDC} values the simulations will fail due to locally too large restraining forces. Thus, for structure refinement based on RDCs of a molecule of a given size (number of atoms N_{at}), there must be a sufficiently large number N_{RDC} of measured RDC values available in order to allow the restraining forces to change the overall orientation of the molecule. If the ratio N_{RDC}/N_{at} gets too small, the restraining forces will not be strong enough to change the overall rotational direction of the molecule (small K^{RDC}) or will be so strong (large K^{RDC}) that they induce a local deformation of the molecule, because the size or inertia of the molecule hinders a restraining induced overall rotation and the internal structure of the molecule is not strong enough to avoid local deformation due to the restraining forces.

Influence of Molecular Flexibility on RDCs

Table 10 shows the effect of enhanced flexibility of cyclo-octane on the orientation-biasing generated RDCs and the RDC-restraining towards these target RDC-values by using the GROMOS 54A7 force field but without any torsional-angle terms (*GRnt*) in the potential energy function in both the orientation-biasing (*bSD*) and the RDC-restraining (*rSD*) simulations.

Introducing flexibility in the force field (compare the D_k^0 target values obtained in the *bSD* simulations with *GRnt* and *GR*), the size of the D_k^0 target values gets slightly smaller for the backbone bonds, slightly larger for the first side-chain bonds. The only sizeable change in these 16 values is observed for bond 5-4, whose value switches from D_k^0 = -4.6 Hz using force field *GR* to D_k^0 = +6.2 Hz using the more flexible *GRnt* force field in the orientation-biasing simulations. This is due to the absence of an improper dihedral-angle force-field term at the chiral atom number 4 in the ring. Upon introducing flexibility into the force field, large changes in D_k^0 target values generated are observed for the second side-chain bonds. RDC values appear to be quite sensitive to the (local) flexibility of the molecule.

The respective $<D_k>$ values (*GR* vs *GRnt*) obtained in the RDC-restraining *rSD* simulations approximate the respective D_k^0 target values (*GR* vs *GRnt*) comparably well for both force fields (compare the combination of *bSD* and *rSD* for *GR* to this combination for *GRnt*).

Table 11 shows the average values of the 16 torsional angles of cyclo-octane and their root-mean-square fluctuations (RMSF) obtained when using the standard GROMOS 54A7 force field (*GR*) and when using this force field without any torsional-angle terms (*GRnt*) in the potential energy function in both, orientation-biasing (*bSD*) and RDC-restraining (*rSD*) simulations. The structures obtained using the standard GROMOS force field (combination of *bSD* and *rSD* for *GR*) show sizeable changes upon introduction of flexibility into the molecule (combination of *bSD* and *rSD* for *GRnt*). Torsional-angle changes in the range [-29°,+57°] are observed. As expected,

without torsional-angle potential energy terms, the simulations generally show larger torsional-angle fluctuations, in particular for the backbone dihedral angles. The relatively large RMSF values for the torsional angles φ(7-3-4-5), φ(5-4-18-16) and φ(3-4-5-6) are due to the absence of an improper dihedral-angle force-field term at the chiral atom with sequence number 4 of the ring.

Influence of a Force-Field Deficiency in RDC Structure Refinement

In structure refinement based on measured RDC data, the orientation distribution underlying these data is unknown. Although bio-molecular force fields are constructed to represent the energy hypersurface of bio-molecules as well as possible, it is likely that such force fields contain deficiencies, which may hinder the restraining towards RDC-values that may be incompatible with the force field used in the structure refinement. This issue can be investigated by generating target RDC D_k^0-values in an orientation-biasing *bSD* simulation using a force field that differs from the one used when restraining RDCs towards these target D_k^0-values in a RDC-restraining *rSD* simulation.

Table 10 shows the effect of using a too flexible force field in RDC structure refinement. Target RDC D_k^0-values were generated in an orientation-biasing *bSD* simulation using the GROMOS 54A7 force field (*GR*), while in the RDC-restraining *rSD* simulation towards these target RDC-values the GROMOS 54A7 force field without any torsional-angle terms in the potential energy function (*GRnt*) was used. Using a too flexible force field (*GRnt* vs *GR*) in the structure refinement leads, as expected, to slightly smaller sizes of the $<D_k>$ values. It also approximates the target

D_k^0-values overall less well.

Table 11 shows the average values of the 16 torsional angles of cyclo-octane and their root-mean-square fluctuations (RMSF) obtained when using the standard GROMOS 54A7 force field (*GR*) in the orientation-biasing *bSD* simulation and using this force field without any torsional-angle terms (*GRnt*) in the potential energy function in the RDC-restraining *rSD* simulations. Using the *GRnt* force field in RDC restraining, sizeable deviations from the values obtained using the standard *GR* force field are observed. These range from –29° to +58°. Using a too flexible force field in structure refinement leads not only to changes in the refined structure, but also to a larger variation in torsional-angle values for the backbone.

Table 10 also shows the effect of using a too rigid force field in RDC structure refinement. Target RDC D_k^0-values were generated in an orientation-biasing *bSD* simulation using the GROMOS 54A7 force field without any torsional-angle terms in the potential energy function (*GRnt*), while in the RDC-restraining *rSD* simulation towards these target RDC-values the complete GROMOS 54A7 force field with all its torsional-angle terms (*GR*) was used. As before, using in structure refinement a force field that is incompatible with the target RDC values, in this case too rigid, leads to slightly smaller sizes of $<D_k>$ values. It also approximates the target D_k^0-values overall less well.

Table 11 shows the average values of the 16 torsional angles of cyclo-octane and their root-mean-square fluctuations (RMSF) obtained when using the GROMOS 54A7 force field without any torsional-angle terms (*GRnt*) in the orientation-biasing *bSD*

simulation and when using the standard GROMOS 54A7 force field (*GR*) in the potential energy function in the RDC-restraining *rSD* simulations. Using the standard *GR* force field in RDC restraining, sizeable deviations from the values obtained using the *GRnt* force field are observed. These range from –58° to +29°. Using a too rigid force field in structure refinement leads not only to changes in the refined structure, but also to a larger variation in torsional-angle values for the backbone.

CONCLUSIONS

A method for structure refinement of bio-molecules based on residual dipolar coupling (RDC) data is proposed, which calculates RDC values using rotational and molecule-internal configurational sampling instead of the standard refinement procedure that is based on the approximation of the non-uniform rotational distribution of the molecule by a single alignment tensor representing the average non-uniformity of the rotational distribution. Thus, by using rotational sampling, as is occurring in the experiment leading to observable RDCs, the algorithm stays close to experiment and avoids the use of an alignment tensor, and thus the assumption that the overall rotation of the molecule is decoupled from its internal motions and that the molecule be rigid. It differs from the methods of RDC-restraining that make use of the alignment-tensor formalism and apply orientational constraints[45] or tensorial constraints[46], in which only a limited part of the rotational space of the molecule is sampled.

The performance of the method is investigated using two simple molecules, two-(united)atomic ethane and a cyclo-octane molecule with eight side chains, containing

24 united atoms. For these small molecules long simulations, up to tens of microseconds, needed to extensively sample their rotational motion, are feasible. Non-uniform rotational distributions for the molecules were generated by adding an orientation-biasing energy term to the potential energy function of the force field used.

For ethane, stochastic dynamics (SD) simulations of at least t^{SD} = 2 μs length using a friction coefficient γ^{SD} of 2.4 ps^{-1} lead to optimal and sufficient sampling of the rotational degrees of freedom of the molecule. For the larger cyclo-octane molecule, a smaller friction, γ^{SD} = 0.24 ps^{-1}, i.e. smaller stochastic forces, is needed to optimise the sampling, but longer simulations of at least t^{SD} = 5 μs length are required to reach sufficient sampling of the rotational degrees of freedom of this larger molecule.

Using an orientation-biasing energy term in the simulations, the size of the generated RDCs depends primarily on the size of the force constant (K^{bias}) of this term, the larger the force constant, the larger the size of the generated RDCs. For proteins, the absolute values of measured RDCs lie in the range [0.1 – 30] Hz. Such values can be reached by using K^{bias} = 0.2 kJ/mol, for both molecules. The value of the biasing reference angle θ^0_{bias} plays a minor role in generating a non-uniform rotational distribution.

Whether the RDC D_k^0 target values can be approximated well in RDC-restraining simulations will primarily depend on the ratio N_{RDC}/N_{at} of the number of target RDC values N_{RDC} and the number of atoms N_{at} in the molecule and on the size of the RDC-restraining forces, which are proportional to $K^{RDC}/\tau_\theta^{RDC}$, the ratio of the force constant K^{RDC} and memory relaxation time τ_θ^{RDC} of the RDC-restraining function. For ethane,

using $\tau_\theta^{\mathrm{RDC}}$ = 1 ns and K^{RDC} = 10 or 100 kJmol^{-1}Hz^{-2}, the <D> value comes close to the D^0 target value. For cyclo-octane, using $\tau_\theta^{\mathrm{RDC}}$ = 1 ns and K^{RDC} = 1 or 2 kJmol^{-1}Hz^{-2}, the D_k^0 target values are approximated well, whereas for K^{RDC} = 3 kJmol^{-1}Hz^{-2} the simulations failed due to too strong restraining forces that induce local deformations. Keeping $K^{\mathrm{RDC}}/\tau_\theta^{\mathrm{RDC}}$ constant, lowering N_{RDC} leads to <D_k> values less well approximating the target D_k^0-values. This would suggest to increase the value of the restraining force constant K^{RDC} for lower N_{RDC} values, but that leads to failure of the simulations due to locally too large restraining forces. Thus, for structure refinement based on RDCs of a molecule of a given size (number of atoms N_{at}), there must be a sufficiently large number N_{RDC} of measured RDC values available in order to allow the restraining forces to change the overall orientation of the molecule. If the ratio $N_{\mathrm{RDC}}/N_{\mathrm{at}}$ gets too small, the restraining forces will not be strong enough to change the overall rotational direction of the molecule (small $K^{\mathrm{RDC}}/\tau_\theta^{\mathrm{RDC}}$) or will be so strong (large $K^{\mathrm{RDC}}/\tau_\theta^{\mathrm{RDC}}$) that they induce a local deformation of the molecule, because the size or inertia of the molecule hinders a restraining induced overall rotation and the internal structure of the molecule is not strong enough to avoid local deformation due to the restraining forces. In a more complex molecule, such as a protein, there are likely to be other factors than just the ratio $N_{\mathrm{RDC}}/N_{\mathrm{at}}$ that determine how well the RDC target values can be approximated. For example, RDCs for backbone NH groups hydrogen-bonded in a helix, backbone NH groups in a flexible loop and NH groups at the end of a long arginine side chain may well behave differently when being restrained.

It appears that the non-uniform rotational distributions generated in the orientation-

biasing simulations can be well reproduced in the RDC-restraining simulations using the RDC target values calculated from the non-uniform rotational distributions.

The effect of varying the molecular flexibility upon the generated RDCs and upon the structures obtained from RDC-restraining simulations was also investigated. Upon introducing flexibility into the force field, large changes in D_k^0 target values generated are observed for the second side-chain bonds in cyclo-octane. RDC values appear to be quite sensitive to the (local) flexibility of the molecule. The structures obtained using the standard GROMOS force field show sizeable changes upon introduction of flexibility into the molecule by omitting the torsional-angle terms in the force field.

Finally, the effect of force-field deficiencies, i.e., using a too flexible or a too rigid force field, in structure refinement based on RDCs was investigated. Using a too flexible force field in structure refinement leads to slightly smaller sizes of the $<D_k>$ values. It also approximates the target D_k^0-values overall less well. Using a too flexible force field in structure refinement leads not only to changes in the refined structure, but also to a larger variation in torsional-angle values. The same holds for using a too rigid force field. It leads to slightly smaller sizes of $<D_k>$ values. It also approximates the target D_k^0-values overall less well, and leads not only to changes in the refined structure, but also to a larger variation in torsional-angle values.

In a next investigation, the method of structure refinement based on RDCs using rotational and molecule-internal sampling will be compared to the commonly applied alignment-tensor approach.

Supporting Information

Table S1. Chirality at the united CH1 atoms of the ring characterised by the definition of the improper dihedral angles and their values. For the definition of a dihedral angle see Refs 39 and 47.

ACKNOWLEDGEMENTS

LJS thanks the Oxford Advanced Research Computing Service (ARC) for access to its facilities, http://dx.doi.org/10.5281/zenodo.22558. NH acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2075 – 390740016 and support by the state of Baden-Württemberg through bwHPC.

References

(1) van Gunsteren, W. F.; Allison, J. A.; Daura, X.; Dolenc, J.; Hansen, N.; Mark, A. E.; Oostenbrink, C.; Rusu, V. H.; Smith, L. J. Deriving Structural Information from Experimentally Measured Data on Biomolecules. *Angew. Chem. Int. Ed.* **2016**, *55*, 15990-16010.

(2) Fennen, J.; Torda, A. E.; van Gunsteren, W. F. Structure Refinement with Molecular Dynamics and a Boltzmann-Weighted Ensemble. *J. Biomol. NMR* **1995**, *6*, 163-170.

(3) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. Time-Dependent Distance Restraints in Molecular Dynamics Simulations. *Chem. Phys. Lett.* **1989**, *157*, 289-294.

(4) Gros, P.; van Gunsteren, W. F.; Hol, W. G. J. Inclusion of Thermal Motion in Crystallographic Structures by Restrained Molecular Dynamics. *Science* **1990**, *249*, 1149-1152.

(5) Schiffer, C. A.; van Gunsteren, W. F. Accessibility and Order of Water Sites in and Around Proteins: A Crystallographic Time-Averaging Study. *Proteins: Struct. Funct. Genet.* **1999**, *36*, 501-511.

(6) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. Time-Averaged Nuclear Overhauser Effect Distance Restraints Applied to Tendamistat. *J. Mol. Biol.* **1990**, *214*, 223-235.

(7) Schiffer, C. A.; Huber, R.; Wüthrich, K.; van Gunsteren, W. F. Simultaneous Refinement of the Structure of BPTI Against NMR Data Measured in Solution and X-Ray Diffraction Data Measured in Single Crystals. *J. Mol. Biol.* **1994**, *241*, 588-599.

(8) Torda, A. E.; Brunne, R. M.; Huber, T.; Kessler, H.; van Gunsteren, W. F. Structure Refinement Using Time-Averaged *J*-Coupling Constant Restraints. *J. Biomol. NMR* **1993**, *3*, 55-66.

(9) Henry, E. R.; Szabo, A. Influence of Vibrational Motion on Solid State Line Shapes and NMR Relaxation. *J. Chem. Phys.* **1985**, *82*, 4753-4761.

(10) Hansen, N., Heller, F., Schmid, N., van Gunsteren, W. F. Time-Averaged Order Parameter Restraints in Molecular Dynamics Simulations. *J. Biomol. NMR* **2014**, *60*, 169–187.

(11) Chen, K.; Tjandra, N. The Use of Residual Dipolar Coupling in Studying Proteins by NMR. *Top. Curr. Chem.* **2012**, *326*, 47-67.

(12) Harris, R. K.; Becker, E. D.; Cabral de Menezes, S. M.; Goodfellow, R.; Granger, P.

NMR Nomenclature: Nuclear Spin Properties and Conventions for Chemical Shifts. IUPAC Recommendations 2001. *Pure Appl. Chem.* **2001**, *73*, 1795-1818.

(13) Harris, R. K.; Becker, E. D.; Cabral de Menezes, S. M.; Goodfellow, R.; Granger, P. NMR Nomenclature: Nuclear Spin Properties and Conventions for Chemical Shifts. IUPAC Recommendations 2001. *Magn. Reson. Chem.* **2002**, *40*, 489-505.

(14) Wirz, L. N. Graph Theoretic and Electronic Properties of Fullerenes and Biasing Molecular Modelling Simulations with Experimental Residual Dipolar Couplings. Ph.D. Dissertation, Massey University, Albany, New Zealand, 2015.

(15) Higman, V. A.; Boyd, J.; Smith, L. J.; Redfield, C. Residual Dipolar Couplings – Are Multiple Independent Alignments Always Possible? *J. Biomol. NMR* **2011**, *49*, 53-60.

(16) Zweckstetter, M.; Bax, A. Prediction of Sterically Induced Alignment in a Dilute Liquid Crystalline Phase: Aid to Protein Structure Determination by NMR. *J. Am. Chem. Soc.* **2000**, *122*, 3791-3792.

(17) Azurmendi, H. F.; Bush, C. A. Tracking Alignment from the Moment of Inertia Tensor (TRAMITE) of Biomolecules in Neutral Dilute Liquid Crystal Solutions. *J. Am. Chem. Soc*. **2002**, *124*, 2426-2427.

(18) Wu, B.; Petersen, M.; Girard, F.; Tessari, M.; Wymenga, S. S. Prediction of Molecular Alignment of Nucleic Acids in Aligned Media. *J. Biomol. NMR* **2006**, *35*, 103-115.

(19) Almond, A.; Axelsen, J. B. Physical Interpretation of Residual Dipolar Couplings in Neutral Aligned Media. *J. Am. Chem. Soc.* **2002**, *124*, 9986- 9987.

(20) Zweckstetter, M. NMR: Prediction of Molecular Alignment from Structure using the PALES Software. *Nat. Protoc.* **2008**, *3*, 679-690.

(21) Wirz, L. N.; Allison, J. A. Fitting Alignment Tensor Components to Experimental RDCs, CSAs and RQCs. *J. Biomol. NMR* **2015**, *62*, 25-29.

(22) van Gunsteren, W. F.; Berendsen, H. J. C., Groningen Molecular Simulation (GROMOS) Library Manual; BIOMOS b.v., University of Groningen, Groningen, The Netherlands, 1987.

(23) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. Biomolecular Simulation: The GROMOS96 Manual and User Guide, Vdf Hochschulverlag AG an der ETH Zürich: Zürich, Switzerland, 1996.

(24) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(25) van Gunsteren, W. F. et al. The GROMOS Software for (Bio)Molecular Simulation. Volume 3: Force Field and Topology Data Set. 2011, http://www.gromos.net/, (accessed 4 July 2019).

(26) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and Testing of the GROMOS Force-Field Versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843−856.

(27) van Gunsteren, W. F.; Boelens, R.; Kaptein, R.; Scheek, R. M.; Zuiderweg, E. R. P. In *Molecular Dynamics and Protein Structure*; Hermans, J., Ed.; Polycrystal Book Service: Western Springs, Ill., USA, 1996, pp 92-99.

(28) van Gunsteren, W. F. et al. The GROMOS Software for (Bio)Molecular Simulation.

Volume 2 : Algorithms and Formulae for Modelling of Molecular Systems, 2011, http://www.gromos.net/, (accessed 4 July 2019).

(29) Riniker, S.; Eichenberger, A. P.; van Gunsteren, W. F. Structural Effects of an Atomic-Level Layer of Water Molecules Around Proteins Solvated in Supra-Molecular Coarse-Grained Water. *J. Phys. Chem. B* **2012**, *116*, 8873-8879.

(30) Bekker, H; Ahlstrom, P. The Virial of Angle Dependent Potentials in Molecular Dynamics Simulations. *Mol. Simul.* **1994**, *13*, 367-374.

(31) Bekker, H.; Berendsen, H. J. C.; van Gunsteren, W. F. Force and Virial of Torsional-Angle Dependent Potentials. *J. Comput. Chem.* **1995**, *16*, 527-533.

(32) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P., van Gunsteren, W. F. Architecture, Implementation and Parallelization of the GROMOS Software for Biomolecular Simulation. *Comput. Phys. Commun.* **2012**, *183*, 890-903.

(33) Kunz, A. P. E.; Allison, J. R.; Geerke, D. P.; Horta, B. A. C.; Hünenberger, P. H.; Riniker, S.; Schmid, N.; van Gunsteren, W. F. New Functionalities in the GROMOS Biomolecular Simulation Software. *J. Comput. Chem.* **2012**, *33*, 340-353.

(34) Schmid, N.; Allison, J. R.; Dolenc, J.; Eichenberger, A. P.; Kunz, A. P. E.; van Gunsteren, W. F. Biomolecular Structure Refinement Using the GROMOS Simulation Software. *J. Biomol. NMR* **2011**, *51*, 265-281.

(35) Riniker, S.; Christ, C. D.; Hansen, H. S.; Hünenberger, P. H.; Oostenbrink, C.; Steiner, D.; van Gunsteren, W. F. Calculation of Relative Free Energies for Ligand-Protein Binding, Solvation and Conformational Transitions Using the GROMOS Software. *J. Phys. Chem. B* **2011**, *115*, 13570-13577.

(36) Eichenberger, A. P.; Allison, J. R.; Dolenc, J.; Geerke, D. P.; Horta, B. A. C.; Meier, K.; Oostenbrink, C.; Schmid, N.; Steiner, D.; Wang, D.; van Gunsteren, W. F. GROMOS++ Software for the Analysis of Biomolecular Simulation Trajectories. *J. Chem. Theory. Comp.* **2011**, *7*, 3379-3390.

(37) van Gunsteren, W. F. et al. The GROMOS Software for (Bio)Molecular Simulation. Volumes 1 – 9, 2011, http://www.gromos.net/, (accessed 4 July 2019).

(38) van Gunsteren, W. F. et al. The GROMOS Software for (Bio)Molecular Simulation Volume 4: Data Structures and Formats, 2011, http://www.gromos.net/, (accessed 4 July 2019).

(39) van Gunsteren, W. F. et al. The GROMOS Software for (Bio)Molecular Simulation. Volume 6: Technical Details, 2011, http://www.gromos.net/, (accessed 4 July 2019).

(40) van Gunsteren, W. F.; Berendsen, H. J. C. A Leap-Frog Algorithm for Stochastic Dynamics. *Mol. Simul.* **1988**, *1*, 173-185.

(41) Allison, J. R.; van Gunsteren, W. F. A Method to Explore Protein Side Chain Conformational Variability Using Experimental Data. *ChemPhysChem* **2009,** *10*, 3213-3228.

(42) Shi, Y. Y.; Wang, L.; van Gunsteren, W. F. On the Approximation of Solvent Effects on the Conformation and Dynamics of Cyclosporin A by Stochastic Dynamics Simulation Techniques. *Mol. Simul.* **1988**, *1*, 369-383.

(43) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(44) van Gunsteren, W. F.; Daura, X; Fuchs, P. F. J.; Hansen, N.; Horta, B. A. C.; Hünenberger, P. H.; Mark, A. E.; Pechlaner, M.; Riniker, S.; Oostenbrink, C. On the Effect of the Various Assumptions and Approximations used in Molecular Simulation on the Properties of Bio-Molecular Systems: Overview and Perspective on Issues. *ChemPhysChem* **2021**, *22*, 264–282.

(45) Sternberg, U.; Witter, R.; Ulrich, A. S. All-Atom Molecular Dynamics Simulations Using Orientational Constraints from Anisotropic NMR Samples. *J. Biomol. NMR* **2007**, *38*, 23-39.

(46) Tzvetkova, P.; Sternberg, U.; Gloge, T.; Navarro-Vàzquez, A.; Luy, B. Configuration Determination by Residual Dipolar Couplings: Accessing the full Conformational Space by Molecular Dynamics with Tensorial Constraints. *Chem. Sci.* **2019**, *10*, 8774-8791.

(47) IUPAC-IUB commission on biochemical nomenclature. Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains. Tentative Rules (1969). *Biochemistry* **1970**, *9*, 3471–3479.

Figure Captions

Figure 1. Graph of the second-order Legendre function $P(\theta) = (3\cos(\theta) - 1)/2$. The two "magic" angles at which the function $P(\theta)$ vanishes, are indicated by two vertical lines.

Figure 2. Molecular topology of the cyclo-octane molecule with eight side chains and with atom sequence numbers of the CH_1, CH_2, and CH_3 united atoms. The 24 small grey, open and solid, arrows indicate the 24 bond vectors, for which RDC values are calculated. The open arrows are bond vectors of which the motion is more extensively analysed. The large grey arrow indicates the atom-atom vector that is used in the biasing force, Eqs. (45-47), to generate an anisotropic orientation distribution of the molecule. For the chirality at the united CH1 atoms of the ring, see Table S1 of Supporting Information.

Figure 3. Angular distribution $P(\theta_{ab})/\sin(\theta_{ab})$ for ethane in unbiased and unrestrained SD simulations using five different friction coefficients γ_i^{SD} (0.0024, 0.024, 0.24, 2.4, 24 ps^{-1}, rows from top to bottom) and for seven different SD simulation periods t^{SD} (0.01, 0.1, 0.5, 1, 5, 10, 20 μs, columns from left to right), where θ_{ab} is the angle of the bond vector with the *z*-axis of the (right-handed) coordinate system.

Figure 4. Angular distribution $P(\varphi_{ab})$ for ethane in unbiased and unrestrained SD simulations using five different friction coefficients γ_i^{SD} (0.0024, 0.024, 0.24, 2.4, 24

ps^{-1}, rows from top to bottom) and for seven different SD simulation periods t^{SD} (0.01, 0.1, 0.5, 1, 5, 10, 20 μs, columns from left to right), where φ_{ab} is the angle of the projection of the bond vector onto the *x*-*y*-plane and the *x*-axis of the (right-handed) coordinate system.

Figure 5. RDC value $\langle D_{ab} \rangle_t$ as function of time *t* (up till 20 μs) for ethane in unbiased and unrestrained SD simulations using five different friction coefficients γ_i^{SD} (0.0024, 0.024, 0.24, 2.4, 24 ps^{-1}, rows from top to bottom). The $\langle D_{ab} \rangle_t$ values after averaging over a time period t^{SD} (0.01 μs – 20 μs), $\langle D_{ab} \rangle_t^{SD}$, are given in Table 3.

Figure 6. Angular distribution $P(\theta_{ab})/\sin(\theta_{ab})$ for the vector $\vec{r}_{ab}$ connecting atoms a = 11 and b = 4 (large arrow in Figure 2) in cyclo-octane in unbiased and unrestrained SD simulations using four different friction coefficients γ_i^{SD} (0.024, 0.24, 2.4, 24 ps^{-1}, rows from top to bottom) and for seven different SD simulation periods t^{SD} (0.01, 0.1, 0.5, 1, 5, 10, 20 μs, columns from left to right), where θ_{ab} is the angle of the bond vector with the *z*-axis of the (right-handed) coordinate system.

Figure 7. Angular distribution $P(\varphi_{ab})$ for the vector $\vec{r}_{ab}$ connecting atoms a = 11 and b = 4 (large arrow in Figure 2) in cyclo-octane in unbiased and unrestrained SD simulations using four different friction coefficients γ_i^{SD} (0.024, 0.24, 2.4, 24 ps^{-1}, rows from top to bottom) and for seven different SD simulation periods t^{SD} (0.01, 0.1, 0.5, 1, 5, 10, 20 μs, columns from left to right), where φ_{ab} is the angle of the

projection of the bond vector onto the *x*-*y*-plane and the *x*-axis of the (right-handed) coordinate system.

Figure 8. RDC values $\langle D_k \rangle_t$ (in Hz) ($k = k_1 k_2$) as function of time t (up till 20 μs) for cyclo-octane in unbiased and unrestrained SD simulations using four different friction coefficients γ_i^{SD} (0.024, 0.24, 2.4, 24 ps^{-1}, rows from top to bottom), for the 24 bond vectors (Figure 2, Table 4). The $\langle D_k \rangle_t$ values after averaging over a time period t^{SD} (0.1 μs – 20 μs), $\langle D_k \rangle_t^{\mathrm{SD}}$, are given in Table 4.

Figure 9. Angular distribution $P(\theta_{\mathrm{ab}})/\sin(\theta_{\mathrm{ab}})$ for ethane in (t^{SD} = 4 μs) orientation-biasing *bSD* simulations with γ^{SD} = 2.4 ps^{-1} using three different biasing reference angles θ_{bias}^0 (0°, 55°, 90°, columns from left to right) and four different biasing force constants K^{bias} (dotted lines: 0.05 kJmol^{-1}; dashed lines: 0.1 kJmol^{-1}; solid lines: 0.2 kJmol^{-1}; dot-dashed lines: 0.5 kJmol^{-1}), where θ_{ab} is the angle of the bond vector with the *z*-axis of the (right-handed) coordinate system.

Figure 10. RDC value $\langle D_{\mathrm{ab}} \rangle_t$ as function of time t for ethane in (t^{SD} = 4 μs) orientation-biasing *bSD* simulations with γ^{SD} = 2.4 ps^{-1} using three different biasing reference angles θ_{bias}^0 (0°, 55°, 90°, columns from left to right) and four different biasing force constants K^{bias} (dotted lines: 0.05 kJmol^{-1}; dashed lines: 0.1 kJmol^{-1}; solid lines: 0.2 kJmol^{-1}; dot-dashed lines: 0.5 kJmol^{-1}), where θ_{ab} is the angle of the bond vector with the *z*-axis of the (right-handed) coordinate system. The $\langle D_{\mathrm{ab}} \rangle_t$

values after averaging over a time period t^{SD}, $\langle D_{ab} \rangle_t^{SD}$, are given in Table 5.

Figure 11. Angular distribution $P(\theta_{ab})/\sin(\theta_{ab})$ for ethane as obtained from orientation-biasing *bSD* simulations and from RDC-restraining *rSD* simulations both with γ^{SD} = 2.4 ps^{-1} and t^{SD} = 4 μs. In the orientation-biasing *bSD* simulations the biasing reference angles are θ_{bias}^0 = 0°, 55°, and 90° (columns from left to right) and the biasing force constants K^{bias} = 0.05, 0.1, 0.2, and 0.5 kJmol^{-1}, where θ_{ab} is the angle of the bond vector with the *z*-axis of the (right-handed) coordinate system. The *bSD* θ-distributions (dotted lines) were symmetrised with respect to θ = 90°. The parameters of the *rSD* simulations were K^{RDC} = 10 kJmol^{-1}Hz^{-2} (dashed lines) and K^{RDC} = 100 kJmol^{-1}Hz^{-2} (solid lines), τ_θ^{RDC}= 1 ns, ΔD^0 = 1 Hz, and the D^0 target values generated in the 12 corresponding *bSD* simulations were used. The first microsecond of these *rSD* simulations were not used in the calculation of the θ-distributions.

Figure 12. Angular distribution $P(\theta_{ab})/\sin(\theta_{ab})$ for the vector $\vec{r}_{ab}$ connecting atoms a = 11 and b = 4 (large arrow in Figure 2) in cyclo-octane as obtained from an orientation-biasing *bSD* simulation and from RDC-restraining *rSD* simulations both with γ^{SD} = 0.24 ps^{-1} and t^{SD} = 20 μs. In the orientation-biasing *bSD* simulation the biasing reference angle θ_{bias}^0= 0° and the biasing force constant K^{bias} = 0.2 kJmol^{-1} were used, where θ_{ab} is the angle of the bond vector with the *z*-axis of the (right-handed) coordinate system. The *bSD* θ-distributions (dotted lines) were symmetrised with respect to θ = 90°. The parameters of the *rSD* simulations were K^{RDC} = 0.5

kJmol^{-1}Hz^{-2} (dashed lines), 1 kJmol^{-1}Hz^{-2} (solid lines) and K^{RDC} = 2 kJmol^{-1}Hz^{-2} (dot-dashed lines), $\tau_\theta^{\mathrm{RDC}}$= 1 ns (left panel) and 10 ns (right panel), ΔD^0 = 1 Hz, and the D^0 target values generated in the *bSD* simulations were used.

Figure 13. RDC values $\langle D_k \rangle_t$ (in Hz) ($k = k_1 k_2$) as function of time t for 10 different bond vectors in cyclo-octane (indicated by the open arrows in Figure 2) in (t^{SD} = 20 μs) orientation-biasing *bSD* simulations with γ^{SD} = 0.24 ps^{-1} using three different biasing reference angles θ_{bias}^0 (0°, 55°, 90°, rows from top to bottom) and three different biasing force constants K^{bias} (dashed lines: 0.1 kJmol^{-1}, solid lines: 0.2 kJmol^{-1}, dot-dashed lines: 0.5 kJmol^{-1}), where θ_k is the angle of the bond vector with the z-axis of the (right-handed) coordinate system. The $\langle D_k \rangle_t$ values after averaging over a time period t^{SD}, $\langle D_k \rangle_t^{\mathrm{SD}}$, are given in Table 6.

Tables

Table 1. Approximate values of the gyromagnetic ratio for some nuclei[12,13].

	γ	γ/(2π)	γ	γ/(2π)
Nucleus	10^6 rad s^{-1} T^{-1}	10^6 Hz T^{-1}	rad e $nm^2/(kJmol^{-1}$ $ps^2)$	e $nm^2/(kJmol^{-1}$ $ps^2)$
^{1}H	267.5221	42.5775	2.77269	0.441288
^{12}C	0	0	0	0
^{13}C	67.2828	10.7084	0.697343	0.110986
^{14}N	19.3378	3.0777	0.200424	0.0318984
^{15}N	-27.1262	-4.3173	-0.281146	-0.0447457
^{16}O	0	0	0	0
^{17}O	-36.2808	-5.7743	-0.376027	-0.0598466
^{32}S	0	0	0	0

Table 2. Relative directions, i.e., values of the angle θ with the z-axis for the 24 bond vectors $\vec{r}_{k_1k_2} = \vec{r}_{k_1} - \vec{r}_{k_2}$, in an energy minimised structure of the cyclo-octane molecule of Figure 2 for an arbitrary orientation of the molecule with respect to the z-axis.

Backbone bonds		First side-chain bonds		Second side-chain bonds	
Bond vector $\vec{r}_{k_1k_2}$	Angle θ (degree)	Bond vector $\vec{r}_{k_1k_2}$	Angle θ (degree)	Bond vector $\vec{r}_{k_1k_2}$	Angle θ (degree)
3-4	87	**2-3**	20	**1-2**	84
4-18	103	**5-4**	145	**6-5**	88
18-16	71	**19-18**	13	**20-19**	75
16-15	88	**17-16**	147	**23-17**	102
15-11	69	**21-15**	30	**22-21**	92
11-10	100	**13-11**	165	**14-13**	105
10-7	88	**12-10**	38	**24-12**	97
7-3	114	**8-7**	159	**9-8**	93

Table 3. RDC $<D_{ab}>_t$ values (Hz) after averaging over a time period t^{SD} in unbiased and unrestrained SD simulations of ethane, $<D_{ab}>_t^{SD}$, using five different friction coefficients γ_i^{SD} (0.0024 ps^{-1} – 24 ps^{-1}) and for eight different SD simulation periods t^{SD} (0.01 μs – 20 μs).

Friction coefficient	SD simulation period t^{SD} (μs)							
γ_i^{SD} (ps^{-1})	0.01	0.1	1.0	2.0	5.0	10.0	15.0	20.0
0.0024	107.8	27.0	-29.5	-23.1	-1.3	2.7	4.0	5.6
0.024	60.8	10.3	-0.4	-1.2	4.0	2.6	2.7	1.2
0.24	0.3	-2.2	-2.5	-1.0	0.5	-0.0	-0.0	0.2
2.4	13.5	1.0	-0.9	-0.1	-0.0	0.0	-0.2	-0.1
24.0	-14.2	-10.3	-1.5	-1.9	-0.6	0.0	-0.4	-0.4

Table 4. RDC $< D_{k_1 k_2} >_t$ values (Hz) after averaging over a time period t^{SD} in unbiased and unrestrained SD simulations of cyclo-octane, $< D_{k_1 k_2} >_{t^{SD}}$, using four different friction coefficients γ_i^{SD} (0.024, 0.24, 2.4, 24 ps^{-1}), for all 24 bond vectors $\vec{r}_{k_1 k_2} = \vec{r}_{k_1} - \vec{r}_{k_2}$, see Figure 2, and for four different SD simulation periods t^{SD} (0.1, 1, 5, 20 μs).[a]

Friction coefficient γ_i^{SD} (ps^{-1})	Backbone bonds					First side-chain bonds					Second side-chain bonds				
	Bond vector $\vec{r}_{k_1 k_2}$	$< D_{k_1 k_2} >_{t^{SD}}$ (Hz)				Bond vector $\vec{r}_{k_1 k_2}$	$< D_{k_1 k_2} >_{t^{SD}}$ (Hz)				Bond vector $\vec{r}_{k_1 k_2}$	$< D_{k_1 k_2} >_{t^{SD}}$ (Hz)			
		t^{SD} (μs)					t^{SD} (μs)					t^{SD} (μs)			
		0.1	1	5	20		0.1	1	5	20		0.1	1	5	20
0.024	**3-4**	14.8	-4.2	-2.9	-0.8	**2-3**	-20.9	-1.2	3.4	0.2	**1-2**	22.0	5.5	-1.2	0.4
	4-18	12.7	-1.6	-1.2	-0.3	**5-4**	-29.6	1.0	3.8	0.6	**6-5**	15.8	5.7	-0.2	0.5
	18-16	21.9	7.1	-0.3	0.5	**19-18**	-29.6	-4.3	1.6	-0.2	**20-19**	20.3	1.8	-2.5	0.3
	16-15	14.7	3.8	-2.1	0.2	**17-16**	-18.8	-1.2	3.2	0.2	**23-17**	11.0	-4.9	-3.6	-0.9
	15-11	11.1	-6.1	-3.9	-1.0	**21-15**	-31.2	-1.0	3.2	0.5	**22-21**	17.4	-1.0	-1.8	-0.8
	11-10	20.8	-0.6	-0.9	-0.3	**13-11**	-22.1	-1.7	2.4	-0.0	**14-13**	15.7	3.0	-2.8	-0.2
	10-7	16.6	7.6	0.1	0.8	**12-10**	-24.6	-0.4	3.4	0.5	**24-12**	14.4	-5.6	-3.7	-0.9
	7-3	15.2	2.1	-3.0	-0.2	**8-7**	-31.8	-4.4	2.0	-0.1	**9-8**	17.2	1.5	-0.6	-0.0
0.24	**3-4**	14.8	2.2	-0.9	0.4	**2-3**	-4.6	-0.8	0.0	-0.3	**1-2**	-8.6	-4.3	0.7	-0.3
	4-18	22.6	3.2	-0.4	-0.4	**5-4**	-2.4	3.3	0.9	-0.8	**6-5**	-3.0	-0.9	1.2	-0.5
	18-16	-10.0	-2.2	1.1	-0.3	**19-18**	-10.8	-3.4	-1.1	0.4	**20-19**	-6.5	-4.2	-0.1	0.5
	16-15	-12.4	-1.4	0.7	0.6	**17-16**	-1.6	1.6	0.3	-0.5	**23-17**	10.2	-0.0	-1.3	1.1
	15-11	11.8	-0.3	-1.7	0.9	**21-15**	1.8	2.4	0.3	-0.6	**22-21**	19.4	1.9	0.1	-0.5
	11-10	20.0	2.9	-0.1	-0.5	**13-11**	-14.7	-3.5	-0.5	0.1	**14-13**	-14.0	-2.7	-0.2	1.1
	10-7	-4.8	0.2	1.6	-0.5	**12-10**	-1.0	3.2	0.8	-0.5	**24-12**	12.2	1.2	-1.6	0.4

	7-3	-17.9	-4.8	-0.2	1.0	**8-7**	-1.1	-0.4	-0.8	0.1	**9-8**	14.5	2.4	0.2	-0.7
2.4	**3-4**	20.1	9.3	1.5	-0.6	**2-3**	-9.9	-5.6	-0.0	0.3	**1-2**	-11.0	-5.8	-1.5	0.7
	4-18	-36.4	-1.0	-4.6	-0.7	**5-4**	17.7	0.4	1.0	0.4	**6-5**	-8.4	-3.2	-2.0	0.2
	18-16	-21.0	-9.4	-1.2	0.9	**19-18**	-12.2	-0.8	0.8	-0.1	**20-19**	28.5	2.7	2.0	0.4
	16-15	43.4	4.7	4.1	0.5	**17-16**	9.3	-1.5	0.2	0.0	**23-17**	16.0	7.9	1.2	-0.5
	15-11	2.3	9.0	1.4	-0.8	**21-15**	3.8	0.3	-1.9	-0.4	**22-21**	-27.2	-1.8	-3.6	-0.6
	11-10	-34.4	-3.3	-3.8	-0.3	**13-11**	-14.3	-4.8	1.0	0.3	**14-13**	27.0	2.5	4.3	0.8
	10-7	-6.9	-6.1	-1.8	0.7	**12-10**	29.8	1.7	0.7	-0.1	**24-12**	3.4	7.2	0.6	-1.1
	7-3	23.3	2.7	4.4	0.6	**8-7**	-11.9	0.5	-1.0	-0.7	**9-8**	-32.9	-2.3	-4.2	-0.2
24	**3-4**	40.2	19.6	-7.1	-0.7	**2-3**	38.1	-22.2	7.4	7.4	**1-2**	-32.6	-1.4	-8.4	-2.1
	4-18	-30.5	11.9	-5.6	-0.2	**5-4**	-16.7	-11.2	9.1	-5.3	**6-5**	-31.3	2.5	-3.4	-2.8
	18-16	-23.1	-9.9	-9.4	-1.7	**19-18**	-2.1	-11.9	12.8	7.7	**20-19**	-14.0	9.1	-6.3	-2.9
	16-15	0.4	9.1	-5.1	-7.6	**17-16**	48.5	-15.7	6.2	3.9	**23-17**	25.8	20.4	-6.0	-0.6
	15-11	13.6	22.6	-3.8	2.0	**21-15**	-36.3	-5.7	13.5	-1.7	**22-21**	-3.3	8.5	-7.7	0.1
	11-10	10.3	4.4	-10.6	1.8	**13-11**	19.3	-20.9	9.0	8.6	**14-13**	2.2	10.7	-6.2	-3.6
	10-7	-49.2	-3.2	-7.2	-7.6	**12-10**	27.1	-9.0	9.8	-1.8	**24-12**	29.1	21.4	-5.2	1.0
	7-3	-0.3	5.8	-4.7	-1.7	**8-7**	-21.2	-4.0	15.2	4.7	**9-8**	-18.7	7.6	-8.6	-0.7

[a]The sequence numbers of the atoms are given in Figure 2.

Table 5. RDC-values $\langle D \rangle_t$ obtained for ethane (single RDC: $N_{RDC} = 1$) as calculated from $t^{SD} = 4$ μs orientation-biasing *bSD* simulations with $\gamma^{SD} = 2.4$ ps^{-1}.[a]

	θ^0_{bias} = 0°		θ^0_{bias} = 55°		θ^0_{bias} = 90°
K^{bias} (kJ/mol)	$\langle D \rangle$ (Hz)	$\langle D \rangle_{sym}$ (Hz)	$\langle D \rangle$ (Hz)	$\langle D \rangle_{sym}$ (Hz)	$\langle D \rangle$ (Hz)
0.05	4.9	4.6	5.3	5.4	5.4
0.1	11.2	10.6	11.1	10.5	11.3
0.2	20.5	20.0	21.9	21.8	22.4
0.5	45.0	44.7	52.1	51.8	55.5

[a]Different orientation-biasing function parameter values (K^{bias} = 0.05, 0.1, 0.2, 0.5 kJmol^{-1}, θ^0_{bias} = 0°, 55°, 90°) were used to generate $\langle D \rangle_t^{SD}$ values.

Table 6. RDC $< D_{k_1k_2} >_{t^{SD}}$ values (Hz) obtained for cyclo-octane as calculated from t^{SD} = 20 μs orientation-biasing *bSD* simulations with γ^{SD} = 0.24 ps^{-1}, for all 24 bond vectors $\vec{r}_{k_1k_2} = \vec{r}_{k_1} - \vec{r}_{k_2}$, see Figure 2, and for different orientation-biasing function parameter values (K^{bias}, θ^0_{bias}), K^{bias} = 0.1, 0.2 and 0.5 kJmol^{-1}, θ^0_{bias} = 0°, 55°, 90°.[a]

θ^0_{bias} (degree)	Backbone bonds				First side-chain bonds				Second side-chain bonds			
	Bond vector $\vec{r}_{k_1k_2}$	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)			Bond vector $\vec{r}_{k_1k_2}$	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)			Bond vector $\vec{r}_{k_1k_2}$	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)		
		K^{bias} (kJmol^{-1})				K^{bias} (kJmol^{-1})				K^{bias} (kJmol^{-1})		
		0.1	0.2	0.5		0.1	0.2	0.5		0.1	0.2	0.5
0°	**3-4**	-1.5	-6.7	-11.7	**2-3**	-4.5	-7.5	-18.0	**1-2**	6.0	13.8	29.7
	4-18	-2.6	-7.5	-13.6	**5-4**	-2.1	-4.6	-10.5	**6-5**	4.9	11.6	24.4
	18-16	6.4	15.7	32.5	**19-18**	-4.5	-7.4	-17.6	**20-19**	7.1	15.6	33.1
	16-15	7.8	16.2	34.3	**17-16**	-5.0	-10.1	-22.7	**23-17**	0.2	-2.9	-4.5
	15-11	-1.9	-6.7	-11.8	**21-15**	-3.7	-7.2	-16.3	**22-21**	-1.8	-5.8	-10.0
	11-10	-2.4	-7.0	-12.7	**13-11**	-3.2	-3.9	-10.7	**14-13**	6.3	13.8	28.3
	10-7	6.8	15.6	32.8	**12-10**	-3.1	-6.8	-15.5	**24-12**	-1.6	-6.2	-11.0
	7-3	7.1	16.1	33.6	**8-7**	-5.5	-10.6	-23.8	**9-8**	-1.3	-3.5	-5.6
55°	**3-4**	-3.6	-7.7	-15.2	**2-3**	-3.4	-7.1	-20.1	**1-2**	6.8	15.8	35.2
	4-18	-4.3	-7.2	-16.4	**5-4**	-1.5	-4.6	-11.5	**6-5**	5.0	12.4	29.2
	18-16	7.6	17.6	39.2	**19-18**	-3.2	-7.9	-20.5	**20-19**	7.9	15.6	37.9
	16-15	8.2	15.8	39.8	**17-16**	-5.0	-10.3	-25.8	**23-17**	-1.4	-4.1	-6.8
	15-11	-3.4	-7.9	-15.7	**21-15**	-3.4	-7.9	-18.7	**22-21**	-3.1	-5.3	-12.2
	11-10	-4.1	-6.1	-15.0	**13-11**	-1.4	-3.5	-11.7	**14-13**	6.8	13.3	33.0
	10-7	7.4	16.9	39.4	**12-10**	-3.2	-7.5	-17.5	**24-12**	-3.4	-6.9	-14.6
	7-3	8.3	16.0	38.9	**8-7**	-4.9	-11.6	-27.9	**9-8**	-2.3	-2.9	-6.9

90°	**3-4**	-2.7	-6.3	-15.2	**2-3**	-4.5	-8.1	-22.8	**1-2**	8.1	15.0	38.2
	4-18	-4.4	-9.1	-16.0	**5-4**	-2.4	-5.1	-12.6	**6-5**	5.9	11.5	31.5
	18-16	8.3	17.0	42.0	**19-18**	-4.7	-8.2	-23.4	**20-19**	8.9	17.6	40.5
	16-15	9.7	18.5	42.6	**17-16**	-5.6	-11.0	-28.3	**23-17**	-0.9	-2.6	-6.9
	15-11	-3.2	-6.5	-16.1	**21-15**	-4.6	-9.2	-20.5	**22-21**	-2.8	-6.2	-11.6
	11-10	-3.9	-7.9	-14.6	**13-11**	-2.7	-4.0	-14.3	**14-13**	7.9	15.7	35.3
	10-7	8.3	16.3	42.7	**12-10**	-3.6	-7.6	-19.4	**24-12**	-2.6	-6.4	-14.4
	7-3	9.3	18.7	41.0	**8-7**	-6.7	-12.5	-30.7	**9-8**	-2.1	-4.7	-5.9

[a]The sequence numbers of the atoms are given in Figure 2.

Table 7. RDC-values $<D>_t$ obtained for ethane (single RDC: N_{RDC}=1) as calculated from (t^{SD} = 4 μs) orientation-biasing *bSD* simulations (second column) and from (t^{SD} = 4 μs) *rSD* simulations (both with γ^{SD} = 2.4 ps^{-1}) using the RDC-restraining algorithm and the orientation-biasing generated RDC-values <*D*> as D^0 target values (right columns).[a]

Orientation-biasing *bSD* simulation parameters and <*D*> values		RDC restraining *rSD* simulations <*D*> (Hz)								
K^{bias} (kJ/mol)	<*D*> = D^0 (Hz)	K^{RDC} = 0.1 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 0.1 (ns)	K^{RDC} = 1 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 0.1 (ns)	K^{RDC} = 10 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 0.1 (ns)	K^{RDC} = 0.1 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 1 (ns)	K^{RDC} = 1 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 1 (ns)	K^{RDC} = 10 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 1 (ns)	K^{RDC} = 100 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 1 (ns)	K^{RDC} = 100 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 10 (ns)	K^{RDC} = 1000 (kJmol^{-1}Hz^{-2}) τ_θ^{RDC}= 10 (ns)
0.05	4.9	2.2	2.8	2.5	0.8	3.1	4.2	4.3	1.1	1.3
0.1	11.2	7.0	8.8	8.7	2.1	7.5	10.2	10.6	7.0	7.4[b]
0.2	20.5	13.8	17.7	18.0	4.2	14.3	19.1	19.8	15.8	16.6[b]
0.5	45.0	31.7	41.2	42.3	9.5	32.3	42.6	44.0[b]	37.8[b]	40.2[c]

[a]Reference orientation angle θ_{bias}^0 = 0°. Different orientation-biasing function force-constant values K^{bias} (0.05, 0.1, 0.2, 0.5 kJmol^{-1}) were used to generate <*D*> values, and different restraining function force-constant K^{RDC} values (0.1, 1, 10, 100, 1000 kJmol^{-1}Hz^{-2}) and τ_θ^{RDC} values (0.1, 1, 10 ns) were used to reproduce these. N_{RDC} = 1, D^0 = <*D*>, ΔD^0 = 1 Hz. [b] These simulations were carried out with Δt = 1 fs. [c] Simulation was carried out with Δt = 0.5 fs.

Table 8. RDC $< D_{k_1k_2} >_{t^{SD}}$ values (Hz) obtained for cyclo-octane as calculated from 24 bond-vector RDC-restraining *rSD* simulations with γ^{SD} = 0.24 ps^{-1} (t^{SD} = 20 μs), for all 24 bond vectors $\vec{r}_{k_1k_2} = \vec{r}_{k_1} - \vec{r}_{k_2}$, see Figure 2, and for different restraining function parameter values (K^{RDC}, τ_θ^{RDC}), K^{RDC} = 1, 2, or 3 kJmol^{-1}Hz^{-2} and τ_θ^{RDC} = 1 ns or 10 ns.[a]

K^{RDC} (kJmol^{-1} Hz^{-2})	Backbone bonds				First side-chain bonds				Second side-chain bonds			
	Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)		Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)		Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)	
			τ_θ^{RDC} (ns)				τ_θ^{RDC} (ns)				τ_θ^{RDC} (ns)	
			1	10			1	10			1	10
1	**3-4**	-6.7	-5.7	-3.1	**2-3**	-7.5	-6.2	-4.8	**1-2**	13.8	12.3	8.4
	4-18	-7.5	-6.3	-2.6	**5-4**	-4.6	-3.9	-3.9	**6-5**	11.6	10.4	7.4
	18-16	15.7	12.9	8.9	**19-18**	-7.4	-6.2	-4.5	**20-19**	15.6	13.8	8.9
	16-15	16.2	13.3	8.1	**17-16**	-10.1	-8.3	-6.4	**23-17**	-2.9	-3.1	-1.2
	15-11	-6.7	-5.8	-2.7	**21-15**	-7.2	-5.9	-4.7	**22-21**	-5.8	-5.1	-2.3
	11-10	-7.0	-5.9	-2.4	**13-11**	-3.9	-3.3	-2.7	**14-13**	13.8	11.7	7.7
	10-7	15.6	12.9	8.6	**12-10**	-6.8	-5.5	-5.1	**24-12**	-6.2	-5.7	-2.6
	7-3	16.1	13.1	8.5	**8-7**	-10.6	-8.9	-6.2	**9-8**	-3.5	-3.2	-0.7
2	**3-4**	-6.7	-5.9	-3.8	**2-3**	-7.5	-6.3	-5.8	**1-2**	13.8	12.3	9.9
	4-18	-7.5	-6.4	-4.2	**5-4**	-4.6	-3.9	-3.5	**6-5**	11.6	10.4	8.7
	18-16	15.7	13.1	10.3	**19-18**	-7.4	-6.4	-5.9	**20-19**	15.6	13.7	10.7
	16-15	16.2	13.5	10.8	**17-16**	-10.1	-8.5	-7.0	**23-17**	-2.9	-3.1	-1.8
	15-11	-6.7	-6.1	-4.0	**21-15**	-7.2	-6.0	-5.1	**22-21**	-5.8	-5.0	-3.3
	11-10	-7.0	-5.9	-3.9	**13-11**	-3.9	-3.4	-3.6	**14-13**	13.8	11.6	9.2

	10-7	15.6	13.2	10.6	**12-10**	-6.8	-5.5	-4.6	**24-12**	-6.2	-5.9	-3.9
	7-3	16.1	13.2	10.3	**8-7**	-10.6	-9.2	-7.6	**9-8**	-3.5	-3.0	-1.8
3	**3-4**	-6.7	-	-4.5	**2-3**	-7.5	-	-6.2	**1-2**	13.8	-	11.1
	4-18	-7.5	-	-4.7	**5-4**	-4.6	-	-3.9	**6-5**	11.6	-	9.1
	18-16	15.7	-	11.8	**19-18**	-7.4	-	-6.3	**20-19**	15.6	-	11.9
	16-15	16.2	-	11.9	**17-16**	-10.1	-	-7.9	**23-17**	-2.9	-	-2.1
	15-11	-6.7	-	-4.5	**21-15**	-7.2	-	-5.7	**22-21**	-5.8	-	-3.7
	11-10	-7.0	-	-4.4	**13-11**	-3.9	-	-3.7	**14-13**	13.8	-	10.3
	10-7	15.6	-	11.8	**12-10**	-6.8	-	-5.4	**24-12**	-6.2	-	-4.3
	7-3	16.1	-	11.6	**8-7**	-10.6	-	-8.4	**9-8**	-3.5	-	-2.0

[a]ΔD^0 = 1 Hz. The sequence numbers of the atoms are given in Figure 2. The $D^0_{k_1k_2}$ target values for the 24 bond vectors were obtained from orientation-biasing *bSD* simulations with γ^{SD} = 0.24 ps^{-1} (t^{SD} = 20 μs), using the orientation-biasing function parameter values K^{bias} = 0.2 kJmol^{-1} and θ^0_{bias} = 0°.

Table 9. RDC $< D_{k_1k_2} >_{t^{SD}}$ values (Hz) obtained for cyclo-octane as calculated from N_{RDC} (2, 4, 10, 12, 24) bond-vector RDC-restraining *rSD* simulations with γ^{SD} = 0.24 ps^{-1} (t^{SD} = 20 μs), for all 24 bond vectors $\vec{r}_{k_1k_2} = \vec{r}_{k_1} - \vec{r}_{k_2}$, see Figure 2, and for different restraining function parameter values (K^{RDC}, τ_θ^{RDC}), K^{RDC} = 1, 0.5, 0.2, or 0.1 kJmol^{-1}Hz^{-2} and τ_θ^{RDC} = 1 ns.[a]

K^{RDC} (kJ mol^{-1} Hz^{-2})	Backbone bonds							First side-chain bonds							Second side-chain bonds						
	Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)					Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)					Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{SD}}$ (Hz)				
			N_{RDC}							N_{RDC}							N_{RDC}				
			24	12	10	4	2			24	12	10	4	2			24	12	10	4	2
1	**3-4**	-6.7	*-5.7*	*-5.1*	-	-	-	**2-3**	-7.5	*-6.2*	*-6.5*	-	-	-	**1-2**	13.8	*12.3*	*11.9*	-	-	-
	4-18	-7.5	*-6.3*	-5.1	-	-	-	**5-4**	-4.6	*-3.9*	-3.1	-	-	-	**6-5**	11.6	*10.4*	9.0	-	-	-
	18-16	15.7	*12.9*	*11.9*	-	-	-	**19-18**	-7.4	*-6.2*	-6.1	-	-	-	**20-19**	15.6	*13.8*	*12.9*	-	-	-
	16-15	16.2	*13.3*	*12.5*	-	-	-	**17-16**	-10.1	*-8.3*	*-8.4*	-	-	-	**23-17**	-2.9	*-3.1*	*-2.7*	-	-	-
	15-11	-6.7	*-5.8*	-5.0	-	-	-	**21-15**	-7.2	*-5.9*	-5.2	-	-	-	**22-21**	-5.8	*-5.1*	-3.9	-	-	-
	11-10	-7.0	*-5.9*	*-5.2*	-	-	-	**13-11**	-3.9	*-3.3*	*-3.8*	-	-	-	**14-13**	13.8	*11.7*	*11.1*	-	-	-
	10-7	15.6	*12.9*	12.3	-	-	-	**12-10**	-6.8	*-5.5*	-5.2	-	-	-	**24-12**	-6.2	*-5.7*	-4.8	-	-	-
	7-3	16.1	*13.1*	11.9	-	-	-	**8-7**	-10.6	*-8.9*	*-8.3*	-	-	-	**9-8**	-3.5	*-3.2*	-2.4	-	-	-
0.5	**3-4**	-6.7	*-5.0*	*-4.4*	*-4.8*	-	*-3.7*	**2-3**	-7.5	*-6.0*	*-6.4*	*-6.3*	-	0.8	**1-2**	13.8	*11.6*	*10.8*	*10.4*	-	1.1
	4-18	-7.5	*-5.5*	-4.5	-3.8	-	-2.2	**5-4**	-4.6	*-4.0*	-2.6	-2.9	-	1.8	**6-5**	11.6	*9.8*	8.4	8.1	-	1.2
	18-16	15.7	*11.9*	*10.6*	*10.4*	-	1.4	**19-18**	-7.4	*-5.8*	-6.2	-4.9	-	1.9	**20-19**	15.6	*12.7*	*11.7*	10.5	-	1.4
	16-15	16.2	*12.2*	*11.7*	*10.6*	-	1.1	**17-16**	-10.1	*-7.9*	*-7.7*	*-8.2*	-	-0.6	**23-17**	-2.9	*-2.4*	*-2.3*	*-2.3*	-	-2.8
	15-11	-6.7	*-5.0*	-4.7	-3.8	-	-2.6	**21-15**	-7.2	*-5.6*	-4.4	-4.2	-	1.9	**22-21**	-5.8	*-4.4*	-3.4	-3.0	-	-2.8
	11-10	-7.0	*-5.1*	*-4.8*	*-4.6*	-	*-3.3*	**13-11**	-3.9	*-3.2*	*-4.1*	*-3.5*	-	1.9	**14-13**	13.8	*10.5*	*10.2*	*9.5*	-	0.2
	10-7	15.6	*11.9*	11.5	11.0	-	1.7	**12-10**	-6.8	*-5.3*	-4.3	-5.3	-	0.5	**24-12**	-6.2	*-5.0*	-4.4	-4.3	-	-3.5
	7-3	16.1	*12.1*	10.7	10.5	-	1.4	**8-7**	-10.6	*-8.3*	*-7.7*	-6.4	-	1.6	**9-8**	-3.5	*-2.6*	-1.9	-1.8	-	-1.9

0.2	**3-4**	-6.7	*-4.1*	*-2.5*	*-3.7*	*-1.3*	*-2.0*	**2-3**	-7.5	*-5.4*	*-5.5*	*-5.3*	-3.2	0.5	**1-2**	13.8	*9.6*	*9.0*	*7.8*	2.6	-0.2
	4-18	-7.5	*-4.0*	-2.8	-2.4	-1.0	-1.1	**5-4**	-4.6	*-3.7*	-4.1	-2.6	-2.3	0.8	**6-5**	11.6	*8.3*	7.0	6.3	2.0	0.2
	18-16	15.7	*10.3*	*8.6*	*7.8*	2.6	-0.2	**19-18**	-7.4	*-5.2*	-5.2	-3.2	-0.0	2.0	**20-19**	15.6	*10.6*	*9.9*	*8.0*	4.4	0.2
	16-15	16.2	*10.1*	*9.1*	*7.8*	3.8	-0.0	**17-16**	-10.1	*-7.1*	*-6.7*	*-7.1*	*-4.8*	-0.3	**23-17**	-2.9	*-1.7*	*-0.8*	*-1.5*	0.5	-1.1
	15-11	-6.7	*-4.0*	-2.3	-2.2	1.1	-0.8	**21-15**	-7.2	*-4.9*	-4.9	-3.2	-1.7	1.5	**22-21**	-5.8	*-3.6*	-1.4	-2.1	-1.0	-1.3
	11-10	-7.0	*-3.8*	*-2.7*	*-3.6*	*-2.5*	*-2.1*	**13-11**	-3.9	*-3.0*	*-3.7*	*-2.5*	-0.9	1.4	**14-13**	13.8	*8.9*	*8.1*	*7.8*	*5.3*	-0.2
	10-7	15.6	*10.3*	8.7	8.2	2.6	0.1	**12-10**	-6.8	*-5.1*	-4.7	-4.9	-3.7	0.3	**24-12**	-6.2	*-3.9*	-2.2	-3.0	-0.2	-1.5
	7-3	16.1	*10.1*	8.9	8.3	5.2	0.6	**8-7**	-10.6	*-7.1*	*-6.6*	-4.3	-0.7	2.0	**9-8**	-3.5	*-1.5*	-0.9	-1.4	-1.2	-1.3
0.1	**3-4**	-6.7	*-3.0*	*-1.3*	*-2.6*	*-0.3*	*-1.6*	**2-3**	-7.5	*-5.0*	*-4.6*	*-4.1*	-2.4	-0.1	**1-2**	13.8	*8.2*	*6.3*	*6.1*	1.5	0.3
	4-18	-7.5	*-3.1*	-2.6	-1.3	-0.9	-0.8	**5-4**	-4.6	*-3.0*	-2.9	-2.3	-1.3	1.0	**6-5**	11.6	*7.1*	5.0	4.9	2.0	0.8
	18-16	15.7	*8.1*	*6.2*	*5.6*	1.3	0.6	**19-18**	-7.4	*-4.6*	-4.3	-2.7	-0.6	0.4	**20-19**	15.6	*9.0*	*7.9*	6.2	2.8	0.5
	16-15	16.2	*8.7*	*7.7*	*5.7*	3.0	0.8	**17-16**	-10.1	*-6.0*	*-5.2*	*-5.2*	*-2.8*	-0.5	**23-17**	-2.9	*-1.3*	*-0.1*	*-1.1*	0.7	-1.0
	15-11	-6.7	*-2.9*	-1.3	-1.6	0.7	-1.2	**21-15**	-7.2	*-3.9*	-4.0	-2.1	-0.9	1.0	**22-21**	-5.8	*-2.4*	-1.7	-1.5	-0.8	-1.1
	11-10	-7.0	*-3.2*	*-2.6*	*-2.3*	*-1.8*	*-1.5*	**13-11**	-3.9	*-3.2*	*-3.1*	*-2.3*	-1.1	0.3	**14-13**	13.8	*7.6*	*6.9*	*5.2*	*3.8*	0.4
	10-7	15.6	*8.5*	6.3	6.2	1.7	1.1	**12-10**	-6.8	*-3.9*	-3.4	-3.5	-1.8	0.4	**24-12**	-6.2	*-2.9*	-1.3	-1.8	0.2	-1.3
	7-3	16.1	*8.3*	7.4	5.7	3.3	0.6	**8-7**	-10.6	*-5.9*	*-5.3*	-3.2	-0.8	0.6	**9-8**	-3.5	*-1.4*	-1.1	-0.4	-0.9	-0.5

[a]ΔD^0 = 1 Hz. The sequence numbers of the atoms are given in Figure 2. The $D^0_{k_1k_2}$ target values for the 24 bond vectors were obtained from orientation-biasing *bSD* simulations with γ^{SD} = 0.24 ps^{-1} (t^{SD} = 20 μs), using the orientation-biasing function parameter values K^{bias} = 0.2 kJmol^{-1} and θ^0_{bias} = 0°. The results are from RDC-restraining to N_{RDC} = 2 bond vectors (3-4, 11-10), to N_{RDC} = 4 bond vectors (in addition 17-16, 14-13), to N_{RDC} = 10 bond vectors (in addition 18-16, 16-15, 2-3, 13-11, 1-2, 23-17, 14-13), to N_{RDC} = 12 bond vectors (in addition 8-7, 20-19), and to all 24 bond vectors. The values for the bonds that are part of the set of N_{RDC} RDC restraints applied, are in italics.

Table 10. RDC $< D_{k_1k_2} >_{t^{\mathrm{SD}}}$ values (Hz) obtained for cyclo-octane as calculated from 24 bond-vector RDC-restraining *rSD* simulations with γ^{SD} = 0.24 ps^{-1} (t^{SD} = 20 μs), for all 24 bond vectors $\vec{r}_{k_1k_2} = \vec{r}_{k_1} - \vec{r}_{k_2}$, see Figure 2, and using the restraining function parameter values K^{RDC} = 2 kJmol^{-1}Hz^{-2} and $\tau_\theta^{\mathrm{RDC}}$ = 1 ns, ΔD^0 = 1 Hz, and either the GROMOS force field (*GR*) or the GROMOS force field without torsional-angle terms (*GRnt*).[a]

Force field used in *bSD* simulation to generate D^0 values	Backbone bonds				First side-chain bonds				Second side-chain bonds			
	Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{\mathrm{SD}}}$ (Hz) Force field used in *rSD* simulation		Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{\mathrm{SD}}}$ (Hz) Force field used in *rSD* simulation		Bond vector $\vec{r}_{k_1k_2}$	$D^0_{k_1k_2}$ (Hz)	$< D_{k_1k_2} >_{t^{\mathrm{SD}}}$ (Hz) Force field used in *rSD* simulation	
			GR	*GRnt*			*GR*	*GRnt*			*GR*	*GRnt*
GR	**3-4**	-6.7	-5.9	-4.5	**2-3**	-7.5	-6.3	-6.4	**1-2**	13.8	12.3	8.9
	4-18	-7.5	-6.4	-4.9	**5-4**	-4.6	-3.9	-1.3	**6-5**	11.6	10.4	10.1
	18-16	15.7	13.1	12.6	**19-18**	-7.4	-6.4	-6.7	**20-19**	15.6	13.7	10.0
	16-15	16.2	13.5	12.7	**17-16**	-10.1	-8.5	-8.4	**23-17**	-2.9	-3.1	-4.3
	15-11	-6.7	-6.1	-5.4	**21-15**	-7.2	-6.0	-6.0	**22-21**	-5.8	-5.0	-1.9
	11-10	-7.0	-5.9	-5.3	**13-11**	-3.9	-3.4	-1.9	**14-13**	13.8	11.6	12.9
	10-7	15.6	13.2	12.4	**12-10**	-6.8	-5.5	-5.5	**24-12**	-6.2	-5.9	-2.3
	7-3	16.1	13.2	12.6	**8-7**	-10.6	-9.2	-8.9	**9-8**	-3.5	-3.0	-4.0
GRnt	**3-4**	-5.4	-4.1	-5.0	**2-3**	-8.3	-6.7	-6.8	**1-2**	2.0	4.7	1.9
	4-18	-6.3	-5.3	-5.6	**5-4**	6.2	0.6	5.4	**6-5**	4.7	5.7	4.4
	18-16	15.1	10.2	12.4	**19-18**	-8.2	-6.7	-6.9	**20-19**	3.4	6.5	2.7
	16-15	16.5	12.9	13.5	**17-16**	-10.6	-7.9	-8.8	**23-17**	-7.5	-4.8	-6.8
	15-11	-5.1	-4.5	-4.9	**21-15**	-7.1	-4.9	-5.9	**22-21**	4.1	1.0	3.7
	11-10	-5.8	-5.2	-5.1	**13-11**	-4.6	-5.5	-3.9	**14-13**	17.6	13.5	15.2
	10-7	15.6	11.9	13.0	**12-10**	-6.8	-4.4	-5.5	**24-12**	5.1	0.8	4.3

	7-3	15.9	11.0	12.8	**8-7**	-10.9	-8.4	-9.3	**9-8**	-7.2	-4.8	-6.1

[a]The sequence numbers of the atoms are given in Figure 2. The $D^0_{k_1k_2}$ target values for the 24 bond vectors were obtained from orientation-biasing *bSD* simulations with γ^{SD} = 0.24 ps^{-1} (t^{SD} = 20 μs), using either the GROMOS force field (*GR*) or the GROMOS force field without torsional-angle terms (*GRnt*) and the orientation-biasing function parameter values K^{bias} = 0.2 kJmol^{-1} and θ^0_{bias} = 0°.

Table 11. Average torsional-angle values <φ> and torsional-angle root-mean-square fluctuations RMSF*(φ)* obtained for cyclo-octane as calculated from 24 bond-vector RDC-restraining *rSD* simulations (*rSD*) with γ^{SD} = 0.24 ps^{-1} (t^{SD} = 20 μs), for all 24 bond vectors $\vec{r}_{k_1k_2} = \vec{r}_{k_1} - \vec{r}_{k_2}$, see Figure 2, and using the restraining function parameter values K^{RDC} = 2 kJmol^{-1}Hz^{-2} and τ_θ^{RDC} = 1 ns, ΔD^0 = 1 Hz, and using either the GROMOS force field (*GR*) or the GROMOS force field without torsional-angle terms (*GRnt*).[a]

Force field used in orientation-biasing *bSD* simulation	Backbone torsional angles					Side-chain torsional angles				
	Torsional angle *(k1-k2-k3-k4)*	Force field used in restraining *rSD* simulation				Torsional angle *(k1-k2-k3-k4)*	Force field used in restraining *rSD* simulation			
		GR		*GRnt*			*GR*		*GRnt*	
		<φ> (°)	RMSF*(φ)* (°)	<φ> (°)	RMSF*(φ)* (°)		<φ> (°)	RMSF*(φ)* (°)	<φ> (°)	RMSF*(φ)* (°)
GR	**7-3-4-5**	-88.8	17.1	-116.5	18.0	**1-2-3-4**	-83.2	28.5	-112.6	23.3
	4-3-7-10	-50.8	6.3	-42.7	10.7	**3-4-5-6**	-70.8	44.2	-12.2	97.3
	5-4-18-16	89.8	16.7	116.5	18.0	**3-7-8-9**	77.4	21.0	105.6	24.2
	3-7-10-11	48.0	6.5	40.5	10.4	**7-10-12-24**	-77.7	22.1	-103.9	24.9
	7-10-11-15	-48.7	6.5	-39.8	10.8	**10-11-13-14**	79.8	25.5	104.8	25.3
	10-11-15-16	48.6	6.7	39.8	10.8	**16-15-21-22**	80.0	26.5	103.9	24.9
	11-15-16-18	-47.4	6.7	-40.5	10.4	**18-16-17-23**	-77.4	21.8	-105.6	24.2
	15-16-18-4	49.7	6.5	42.7	10.7	**4-18-19-20**	80.1	26.3	112.6	23.1
GRnt	**7-3-4-5**	-89.0	17.5	-116.5	18.0	**1-2-3-4**	-83.3	28.9	-112.6	23.4
	4-3-7-10	-50.8	6.3	-42.7	10.7	**3-4-5-6**	-70.1	45.1	-12.2	97.3
	5-4-18-16	90.1	17.1	116.5	18.0	**3-7-8-9**	77.3	21.3	105.6	24.3
	3-7-10-11	48.0	6.5	40.5	10.4	**7-10-12-24**	-77.7	22.1	-103.9	24.7
	7-10-11-15	-48.6	6.5	-39.7	10.8	**10-11-13-14**	79.7	25.7	104.7	25.6
	10-11-15-16	48.6	6.7	39.7	10.8	**16-15-21-22**	80.1	26.2	103.8	25.1
	11-15-16-18	-47.4	6.7	-40.5	10.4	**18-16-17-23**	-77.4	21.9	-105.6	24.2

	15-16-18-4	49.7	6.5	42.7	10.7	**4-18-19-20**	80.3	26.7	112.5	23.4

[a]The $D^0_{k_1k_2}$ target values for the 24 bond vectors were obtained from orientation-biasing *bSD* simulations with $\gamma^{SD} = 0.24$ ps^{-1} ($t^{SD} = 20$ μs), using either the GROMOS force field (*GR*) or the GROMOS force field without torsional-angle terms (*GRnt*) and the orientation-biasing function parameter values $K^{bias} = 0.2$ kJmol^{-1} and $\theta^0_{bias} = 0°$. The atom numbers *(k1-k2-k3-k4)* defining a torsional angle are given in Figure 2.

Figure 1

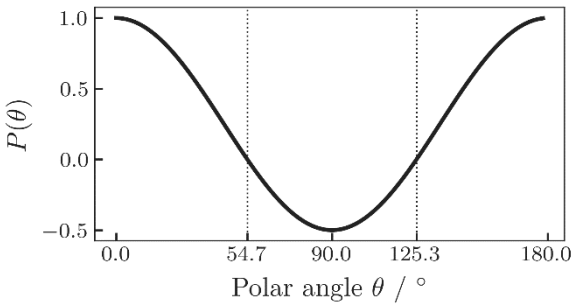


Figure 2

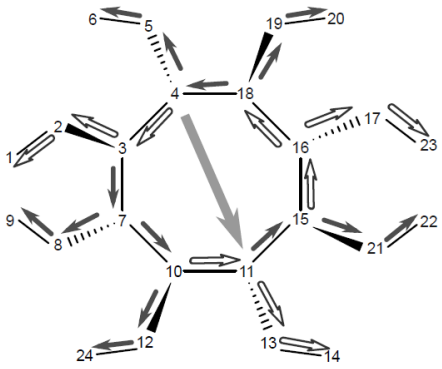


Figure 3

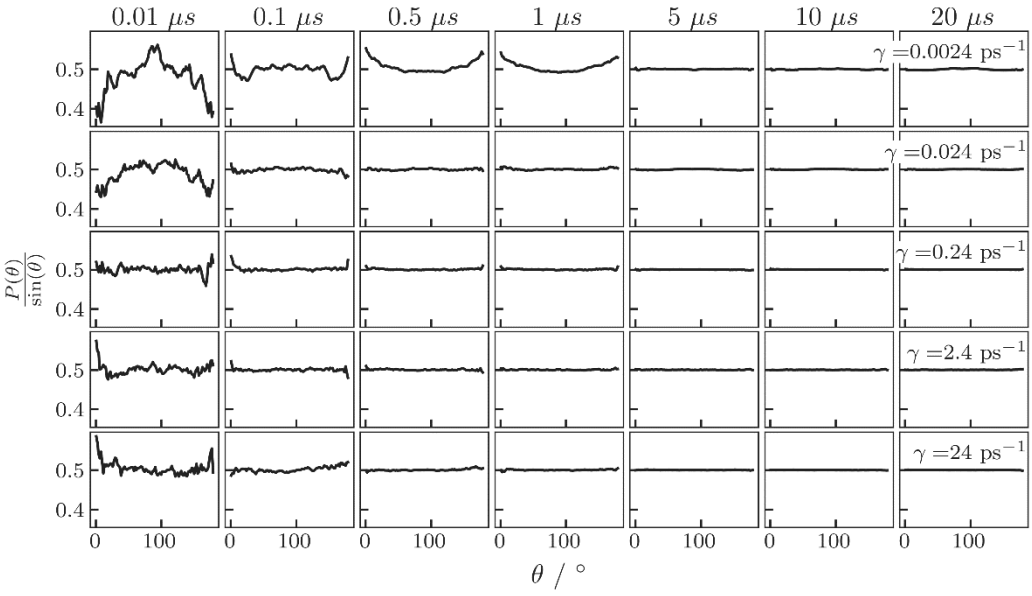

0.01 μs
0.1 μs
0.5 μs
1 μs
5 μs
10 μs
20 μs
γ =0.0024 ps^{-1}
γ =0.024 ps^{-1}
γ =0.24 ps^{-1}
γ =2.4 ps^{-1}
γ =24 ps^{-1}
$\frac{P(\theta)}{\sin(\theta)}$
0.5
0.4
0
100
θ / °

Figure 4

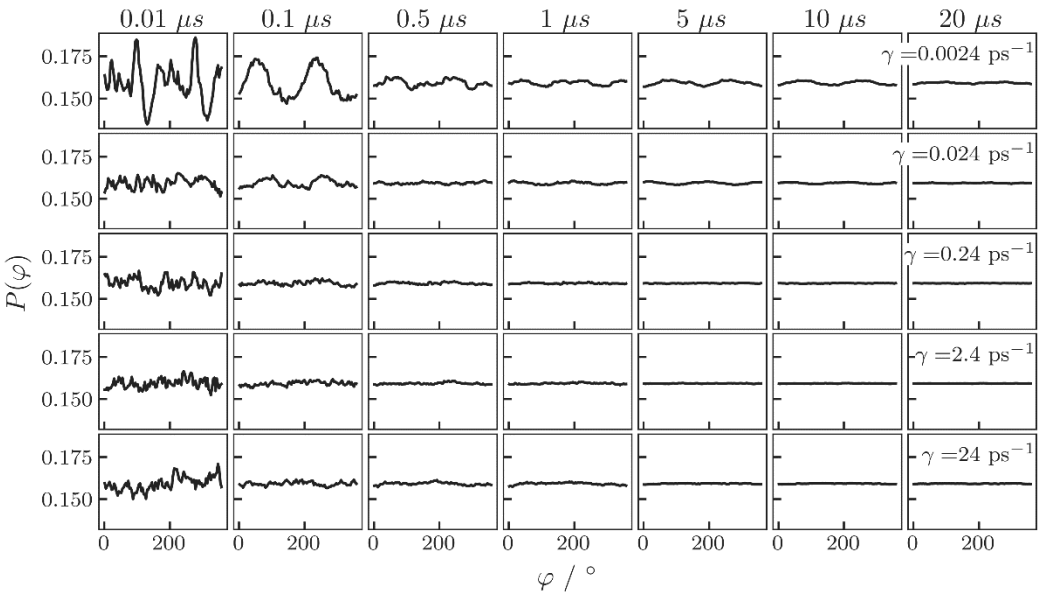

0.01 μs
0.1 μs
0.5 μs
1 μs
5 μs
10 μs
20 μs
γ =0.0024 ps^{-1}
γ =0.024 ps^{-1}
γ =0.24 ps^{-1}
γ =2.4 ps^{-1}
γ =24 ps^{-1}
$P(\varphi)$
0.175
0.150
0
200
φ / °

Figure 5

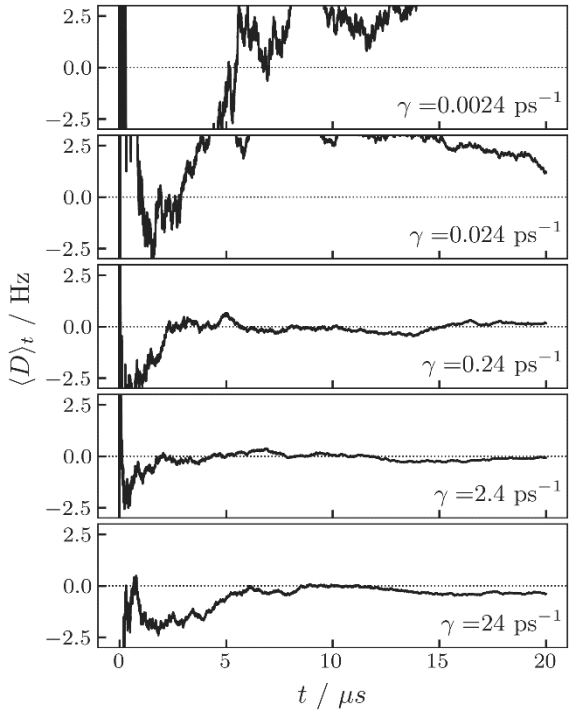

2.5
0.0
−2.5
γ =0.0024 ps^{-1}
γ =0.024 ps^{-1}
γ =0.24 ps^{-1}
γ =2.4 ps^{-1}
γ =24 ps^{-1}
$\langle D \rangle_t$ / Hz
0
5
10
15
20
t / μs

Figure 6

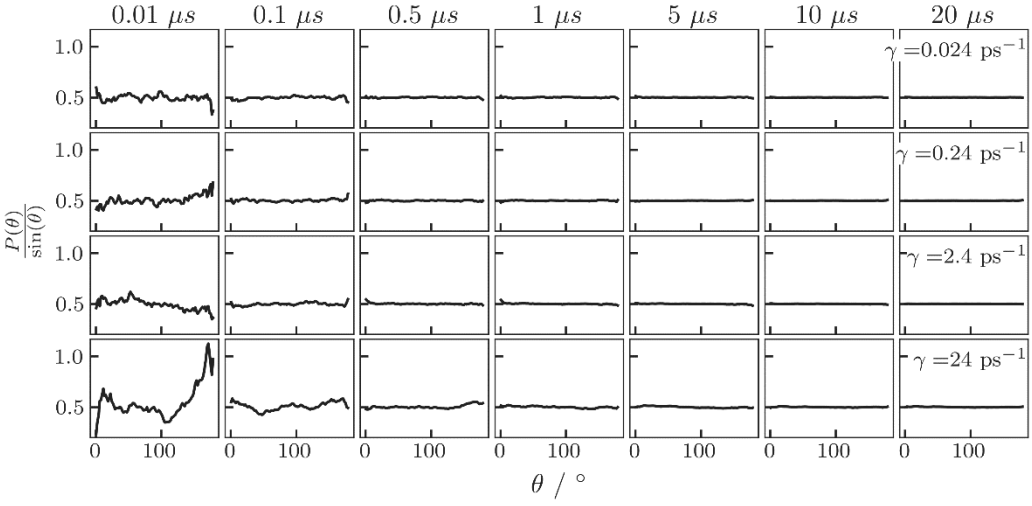

0.01 μs
0.1 μs
0.5 μs
1 μs
5 μs
10 μs
20 μs
γ =0.024 ps^{-1}
γ =0.24 ps^{-1}
γ =2.4 ps^{-1}
γ =24 ps^{-1}
$\frac{P(\theta)}{\sin(\theta)}$
1.0
0.5
0
100
θ / °

Figure 7

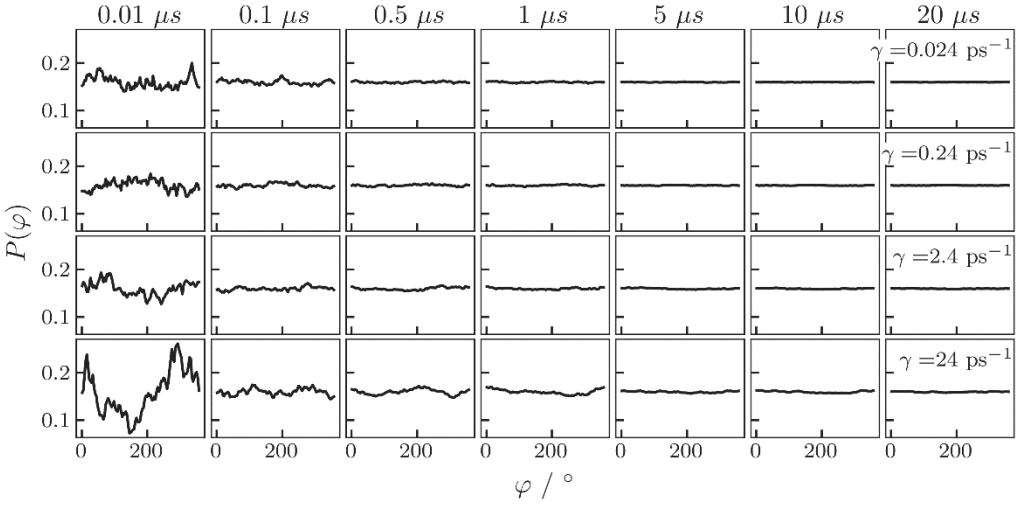

0.01 μs
0.1 μs
0.5 μs
1 μs
5 μs
10 μs
20 μs
γ =0.024 ps^{-1}
γ =0.24 ps^{-1}
γ =2.4 ps^{-1}
γ =24 ps^{-1}
$P(\varphi)$
0.2
0.1
0
200
φ / °

Figure 8

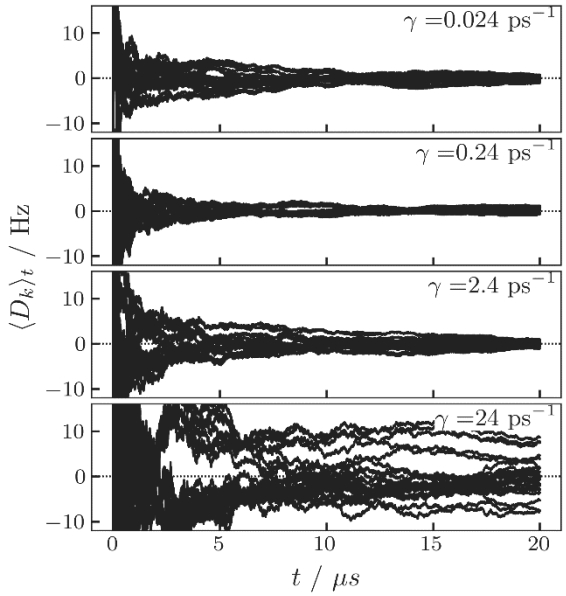


Figure 9

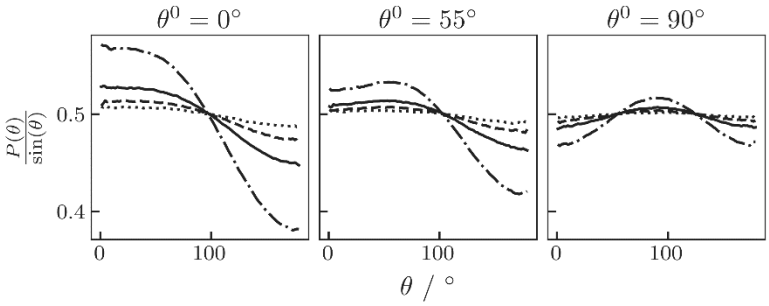

$\theta^0 = 0°$
$\theta^0 = 55°$
$\theta^0 = 90°$
$\frac{P(\theta)}{\sin(\theta)}$
0.5
0.4
0
100
0
100
0
100
θ / °

Figure 10

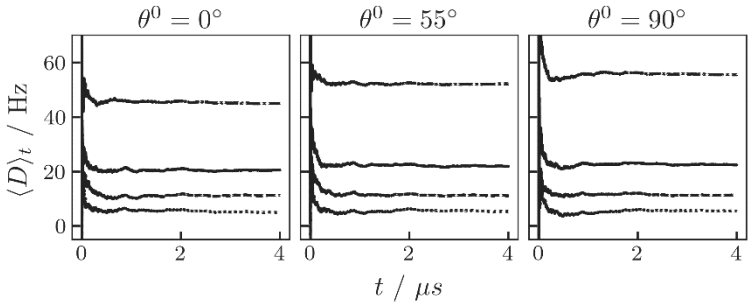

$\theta^0 = 0°$
$\theta^0 = 55°$
$\theta^0 = 90°$
$\langle D \rangle_t$ / Hz
60
40
20
0
0
2
4
t / μs

Figure 11

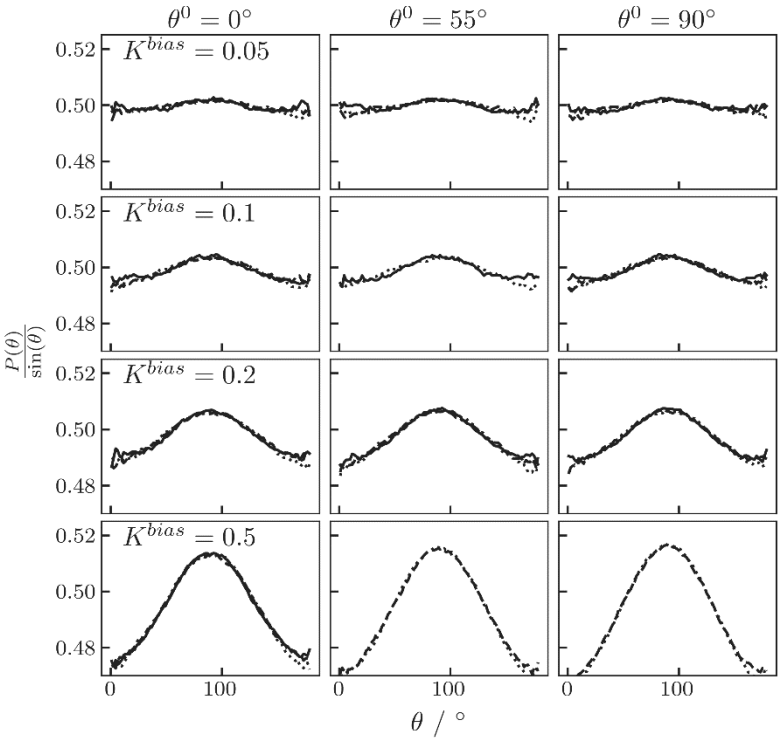

$\theta^0 = 0°$
$\theta^0 = 55°$
$\theta^0 = 90°$
$K^{bias} = 0.05$
$K^{bias} = 0.1$
$K^{bias} = 0.2$
$K^{bias} = 0.5$
0.52
0.50
0.48
$\frac{P(\theta)}{\sin(\theta)}$
0
100
θ / °

Figure 12

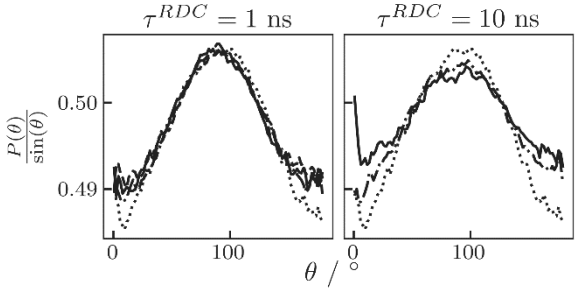
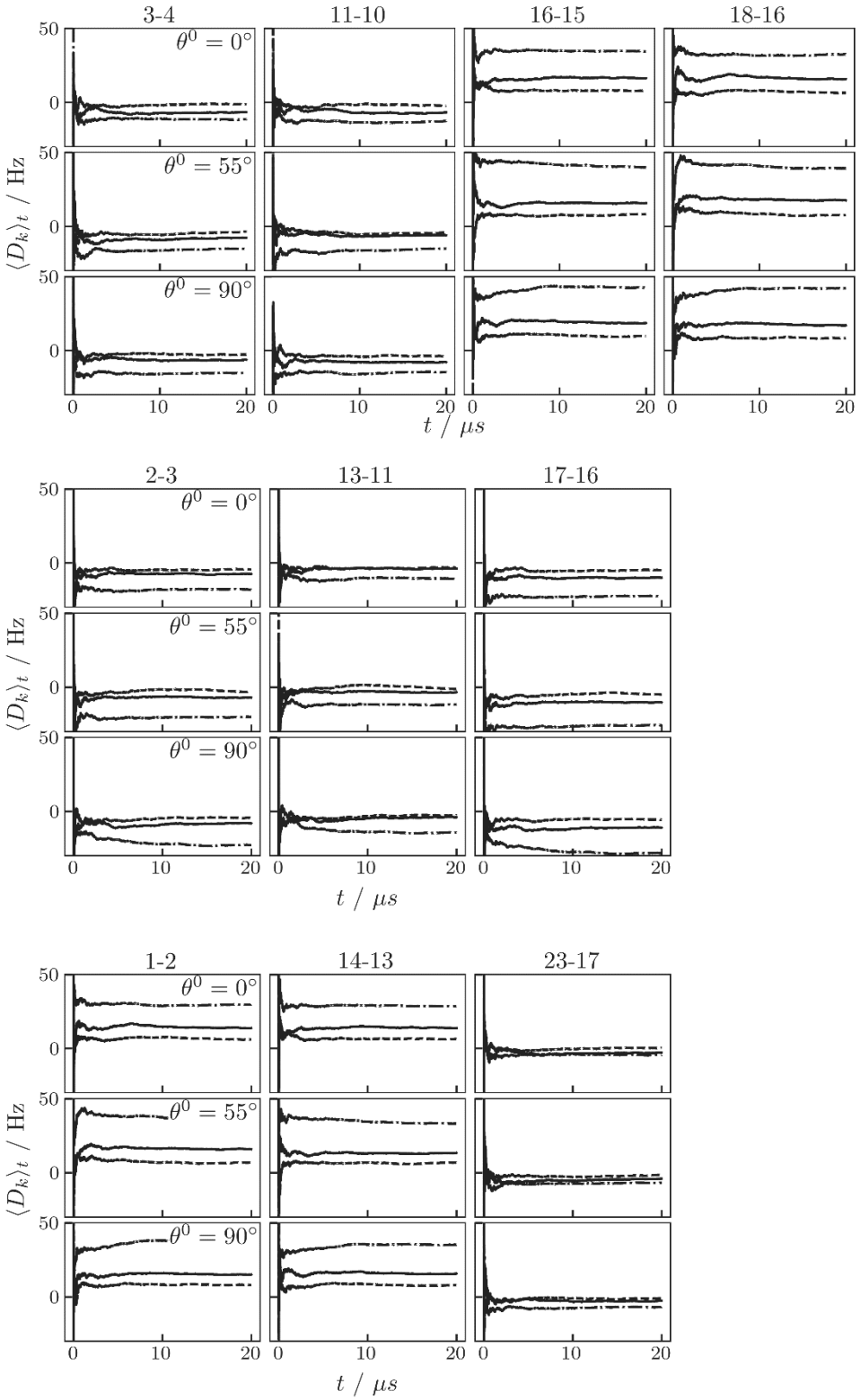

$\tau^{RDC} = 1$ ns
$\tau^{RDC} = 10$ ns
$\frac{P(\theta)}{\sin(\theta)}$
0.50
0.49
0
100
0
100
θ / °

Figure 13


3-4
11-10
16-15
18-16
$\theta^0 = 0°$
$\theta^0 = 55°$
$\theta^0 = 90°$
$\langle D_k \rangle_t$ / Hz
t / μs
2-3
13-11
17-16
1-2
14-13
23-17

TOC Graphic

$$D(t) = D^c R(t)\left(\frac{1}{2}\left(3\cos^2\left(\theta(t)\right) - 1\right)\right)$$