J Biomol NMR (2014) 60:169–187
DOI 10.1007/s10858-014-9866-7

ARTICLE

Time-averaged order parameter restraints in molecular dynamics simulations

Niels Hansen · Fabian Heller · Nathan Schmid · Wilfred F. van Gunsteren

Received: 13 May 2014 / Accepted: 25 September 2014 / Published online: 14 October 2014

Abstract A method is described that allows experimental S^2 order parameters to be enforced as a time-averaged quantity in molecular dynamics simulations. The two parameters that characterize time-averaged restraining, the memory relaxation time and the weight of the restraining potential energy term in the potential energy function used in the simulation, are systematically investigated based on two model systems, a vector with one end restrained in space and a pentapeptide. For the latter it is shown that the backbone N–H order parameter of individual residues can be enforced such that the spatial fluctuations of quantities depending on atomic coordinates are not significantly perturbed. The applicability to realistic systems is illustrated for the B3 domain of protein G in aqueous solution.

Keywords Structure refinement · Nuclear magnetic resonance · Force field · Conformational dynamics · Statistical mechanics

Introduction

The interpretation of NMR observables measured for proteins in solution usually requires the consideration of a structurally heterogeneous ensemble of conformers rather than a single static structure (Jardetzky 1980; Braun et al. 1981). This makes molecular dynamics (MD) simulation in explicit solvent a versatile tool to rationalize NMR measurements (van Gunsteren et al. 1994, 2008; Gattin et al. 2009). It can be used to select an ensemble of low-energy structures that are consistent with a given set of experimental measurements. However, the use of MD simulations is limited by finite, often insufficient conformational sampling and force-field inaccuracies (Misura and Baker 2005; Kim et al. 2009; Gniewek et al. 2012). Both limitations are connected because long simulation times might be needed to uncover force-field deficiencies (Raval et al. 2012). In addition to a continuous effort to refine and test atomistic force fields (Beauchamp et al. 2012), the incorporation of information from experiments as restraints contributing to the potential energy function may serve to mitigate both sampling deficiencies and force-field inaccuracies. An important caveat in pursuing this goal is, however, to avoid corrupting the resulting conformational ensemble with spurious and arbitrary biases.

Experiments on biomolecular systems yield values Q_{exp} for an observable $Q(\vec{r}^N, \vec{p}^N)$ that depends on the phase space variables $\vec{r}^N = (\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$, the Cartesian coordinates of the N particles in the system, and $\vec{p}^N = (\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N)$, the conjugate momenta of the N particles. Many measurable quantities Q only depend on the configurations $\vec{r}^N$, and not on $\vec{p}^N$, for which reason we simplify the notation to $Q(\vec{r}^N)$. The great majority of experiments on biomolecular systems yield Q_{exp} values that are averages over molecules in the test tube and over time,

$$Q_{\mathrm{exp}} = \langle \langle Q \rangle_{\mathrm{molecules}} \rangle_{\mathrm{time}}, \tag{1}$$

where the brackets $\langle \ldots \rangle$ denote averaging over configurations $\vec{r}^N$.

MD simulation at constant temperature generally yields a trajectory $\vec{r}^N(t)$ of configurations that are Boltzmann

N. Hansen · F. Heller · N. Schmid · W. F. van Gunsteren
Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zurich, Switzerland
e-mail: wfvgn@igc.phys.chem.ethz.ch

N. Hansen (✉)
Institute of Thermodynamics and Thermal Process Engineering, University of Stuttgart, 70569 Stuttgart, Germany
e-mail: hansen@itt.uni-stuttgart.de

distributed and thus can be averaged straightforwardly to obtain $\langle Q \rangle$ values for comparison to Q_{exp} values. Time-averaging restraints were introduced (Torda et al. 1989) using NOE derived atom-atom (i,j) distance, r_{ij}, information with $Q = r_{ij}^{-3}$ or $Q = r_{ij}^{-6}$, followed by application to chemical shift restraining (Harvey and van Gunsteren 1993), ^3J-coupling restraining (Torda et al. 1993; Scott et al. 1998) and crystallographic structure factor amplitude restraining (Gros and van Gunsteren 1993; Schiffer et al. 1995; Schiffer and van Gunsteren 1999). The time-averaging restraining methods are characterised by two parameters, the force constant or weight of the restraining potential energy term of the interaction function and the memory relaxation time representing the time span over which $Q(\vec{r}^N(t))$ is to be averaged. Their behavior as function of these parameters has been analysed for a variety of systems (Torda et al. 1989; Pearlman and Kollman 1991; Schmitz et al. 1992, 1993; Pearlman 1994a, b; Nanzer et al. 1995, 1997; Dolenc et al. 2010). The force constant should be taken as small as possible in order to avoid a restraining energy bias that destroys the proper Boltzmann weighting in regard to the physical interaction function or force field of the configurations while being large enough to force $\langle Q \rangle$ close to Q_{exp}. The memory relaxation time should be of the order of the experimental averaging time that determines Q_{exp}, but at least an order of magnitude smaller than the length of the MD simulation in order to secure sufficient statistics when averaging $Q(\vec{r}^N(t))$ over t. In addition, the heating of the system due to the non-conservative force resulting from the time-averaging restraining term should be kept small.

The averaging over molecules can be accounted for by simulating in parallel N_m independent systems for which the initial coordinates and momenta $(\vec{r}^N(t_0), \vec{p}^N(t_0))$ are Boltzmann distributed in regard to the physical interaction function that is supposed to approximate the "real" interaction between the atoms of the system, which is easily done for the momenta $\vec{p}^N$ because the kinetic energy term of the Hamiltonian is quadratic in the momenta (in the absence of coordinate constraints), but rather difficult if not impossible for the configurations $\vec{r}^N$ of a biomolecular system, because the potential energy term of the Hamiltonian is generally a complex function of the coordinates. For liquids of small molecules one may choose $\vec{r}^N(t_0)$ more or less arbitrarily, i.e. non-Boltzmann distributed, because their short configurational relaxation times, e.g. of the order of 10–100 ps for liquid water at physiological temperature and pressure, allow each of the N_m different systems to equilibrate to a Boltzmann distributed one within a period about 10 times longer. In addition, the dependence of averages of quantities $Q(\vec{r}^N)$ upon system size is small due to the generally short-ranged nature of the spatial correlations for liquids at physiological thermodynamic conditions. For biomolecules such as proteins in aqueous solution the situation is quite different. Relaxation times of the system generally exceed the MD simulation time which means that if the initial configurations $\vec{r}^N(t_0)$ of the N_m systems are non-Boltzmann distributed, the averages $\langle Q \rangle$ over the N_m systems are very likely to result from non-Boltzmann averaging. This is one of the reasons for which averaging over molecules to obtain $\langle Q \rangle$ should not be applied (Fennen et al. 1995). We note, however, that the long equilibration time needed for each system to reach a Boltzmann distributed trajectory may be shortened by applying replica-exchange techniques (Sugita and Okamoto 1999; Sugita et al. 2000; Fukunishi et al. 2002) that swap configurations between the N_m molecules or systems based on the Boltzmann probability of the respective configurations.

Molecule-averaging restraints were introduced (Scheek et al. 1991; Fennen et al. 1995) using NOE-derived atom-atom distance information. The restraining term of the potential energy function is made dependent on the average $\langle Q \rangle$ of a quantity $Q(\vec{r}^N)$ over the N_m molecules or systems (Huber and van Gunsteren 1998). Molecule-averaging restraining methods are also characterized by two parameters, the force constant or weight of the restraining potential energy term of the interaction function and the number of molecules or systems over which the averaging is performed (Huber and van Gunsteren 1998; Hess and Scheek 2003). However, in case the initial configurations $\vec{r}^N(t_0)$ are not Boltzmann distributed over the N_m systems and the simulation period is not much longer than the system relaxation times, the average $\langle Q \rangle$ over molecules has only be shown to be Boltzmann weighted in the limit of large numbers of molecules and of large restraining force constants (Pitera and Chodera 2012; Roux and Weare 2013; Cavalli et al. 2013) and in the absence of noise in the target data, i.e. the measured values for e.g. NOEs (Olsson et al. 2013).

To evaluate the adequacy of protein force fields, comparisons between calculated and experimentally derived backbone N–H order parameters have been carried out. The backbone N–H order parameter is a measure for the spatial restriction that the N–H vector experiences in a molecular reference frame. It is dominated by the magnitude of the local librations of the peptide plane, i.e. concerted fluctuations of the neighboring ψ_{i-1} and φ_i torsion angles are responsible for the N–H order parameter of the i-th residue (Smith et al. 1995a). Generalized order parameters smaller than one are difficult to interpret in the absence of specific motional models, since they are consistent with a large number of different motional models (Brüschweiler and Wright 1994; Palmer et al. 1966; Luginbühl and Wüthrich

2002; d'Auvergne and Gooley 2003; Johnson et al. 2008). However, a great majority of published NMR backbone relaxation dynamics studies uses a so-called model-free analysis (Lipari and Szabo 1982) of the data (Jarymowycz and Stone 2006). Order parameters calculated from ensembles generated by MD simulations are not subject to a specific motional model but depend on the local flexibility inherent in the force field when solving classical, Newton's equation of motion and on whether the assumption of internal motion being statistically independent of overall tumbling is justified (Peter et al. 2001; Feenstra et al. 2002; Wong and Case 2008; Johnson 2012). Many simulation studies came to the conclusion that protein force fields allow too much local flexibility to reproduce experimental order parameters (Buck et al. 2006; Hornak et al. 2006; Showalter and Brüschweiler 2007; Trbovic et al. 2008), resulting in changes in the energetics of torsional-angle terms for the φ and ψ backbone angles in the CHARMM and Amber force fields (Duan et al. 2003; MacKerell 2004; MacKerell et al. 2004; Hornak et al. 2006) leading to better agreement between calculated and experimentally derived order parameters for proteins.

NMR order parameters have been used to bias the sampling using molecule or system averaging restraining (Best and Vendruscolo 2004; Richter et al. 2007). This is less straightforward than restraining other quantities obtainable from NMR experiments such as NOEs or ^3J-couplings, because the order parameter S^2 is not a function of a single molecular configuration $\vec{r}^N$, but of the long-time tail of a particular correlation function of a vector along a particular bond or line connecting two atoms of a molecule. This long-time limit or tail can be expressed (Henry and Szabo 1985) as an ensemble- or time-average of a function of the molecular configuration $\vec{r}^N$,

$$S^2(\langle f(\vec{r}^N)\rangle). \tag{2}$$

This average can be taken over a particular, finite time period or over a finite Boltzmann ensemble of configurations $\vec{r}^N$. However, because in the NMR measurement practice the "long-time tail" is determined through a sensitivity time window that depends on the overall rotational tumbling time of the molecule, i.e. of the order of nanoseconds for peptides and proteins, the application of time averaging allows to represent this experimental sensitivity time window in a restraining MD simulation, whereas this is not possible using molecule averaging.

An implementation of molecule averaging is the simultaneous simulation of N_m independent, identical systems, which are coupled together through an S^2 restraining term in the potential energy function that restrains S^2 calculated as an average over the N_m systems. As mentioned before, such a restraining function will allow for a Boltzmann distributed configurational ensemble, in regard to the physical interaction function or force field, to be generated in the simulation in the limits of the number of systems and the force constants of the restraining terms going to infinity (Huber and van Gunsteren 1998; Hess and Scheek 2003), and in the absence of experimental noise (Olsson et al. 2013). Unfortunately, neither condition is met in practice: (1) The number of systems N_m should be kept low, i.e. 10–100, in order to avoid a blow-up of the computational effort; (2) The force constant of the restraining term should be chosen as small as possible, because the restraining term is unphysical in regard to the other force-field terms that represent inter-atomic interactions and it only serves to compensate force-field deficiencies, deficiencies in the function f relating S^2 to configurations, and sampling deficiencies in the absence of which no restraining term would be needed in simulations (Torda et al. 1989; Pearlman and Kollman 1991; Schmitz et al. 1992, 1993; Pearlman 1994a, b; Nanzer et al. 1995, 1997; Dolenc et al. 2010). If these deficiencies happen to be large, a large force constant may force the trajectory to bring $\langle S^2\rangle$ close to S^2_{exp}, but at the cost of misinterpreting the energy of the system; (3) S^2 order parameters are rather sensitive to noise originating from the procedure to derive them from measured data.

A third problematic aspect of the molecule or system averaging restraining is the already mentioned difficulty to choose properly Boltzmann-distributed initial configurations for the N_m systems, i.e. with weights

$$P(\vec{r}^N(t_0)) \propto \exp(-V_{\text{pot}}(\vec{r}^N(t_0))/k_{\text{B}}T) \tag{3}$$

proportional to the Boltzmann factor of the potential energy $V_{\text{pot}}(\vec{r}^N(t_0))$ of each initial configuration, with the Boltzmann constant denoted by k_{B}, and the temperature by T. In practice (Best and Vendruscolo 2004; Richter et al. 2007), the choice

$$P(\vec{r}^N(t_0)) \propto N_m^{-1} \tag{4}$$

is made, i.e. the initial configurations of the N_m systems have equal weights irrespective their potential energies. This means that it is assumed that the equilibration of each of the N_m systems towards a Boltzmann distribution will be sufficiently fast to avoid non-Boltzmann averaging when calculating $\langle S^2\rangle$.

To avoid the problem of inappropriate weighting of configurations i of the N_m molecules in the average

$$\langle Q\rangle = \sum_{i=1}^{N_m} P(\vec{r}^N(i))Q(\vec{r}^N(i)) \tag{5}$$

the weights $P(\vec{r}^N(i))$ can be taken as Boltzmann distributed (Fennen et al. 1995; Huber et al. 1996),

$$P(\vec{r}^N(i)) = \frac{\exp(-V_{\text{pot}}(\vec{r}^N(i))/k_{\text{B}}T)}{\sum_{j=1}^{N_m}\exp(-V_{\text{pot}}(\vec{r}^N(j))/k_{\text{B}}T)}. \quad (6)$$

Because Eq. 6 implies a strong dependence of the weight on the configurational potential energy, which may show large fluctuations in an MD simulation, the instantaneous potential energy value $V_{\text{pot}}(\vec{r}^N(i))$ in Eq. 6 was replaced by an average over parts of the trajectory (Fennen et al. 1995). Summarizing the investigation of Fennen et al. (1995) it was concluded that space- or molecular averaging restraints can be used, but that it is wise to use only time averaging as a first approach and only resort to molecule averaging when justified by evidence of the presence of high barriers separating conformers that contribute significantly to $\langle Q \rangle$, because such barriers might mitigate the sampling in an MD simulation. In the latter case, molecule averaging is only likely to enhance the sampling if the initial configurations of the N_m systems are on different sides of the high barriers (Pepermans et al. 1988).

If the force field used in the simulation would be of infinite accuracy, the sampling of configurations $\vec{r}^N$ would be infinite, and the procedure or function f used in Eq. 2 to link an S^2 order parameter to configurations $\vec{r}^N$ is infinitely accurate, no restraining potential energy term would be required in an MD simulation in order to obtain correct S^2 values from trajectory configurations. Unfortunately, neither of these three conditions is met in practice. If only one of these conditions is wildly violated, restraining the configurational sampling by whichever method is likely to introduce spurious and arbitrary biases. A variety of methods (Best and Vendruscolo 2004; Richter et al. 2007; Pitera and Chodera 2012; Roux and Weare 2013; Cavalli et al. 2013; Olsson et al. 2013, 2014; White and Voth 2014) has been proposed that aim at minimising the bias induced in the configurational distribution $P(\vec{r}^N)$ of the system while using molecule-averaging in the restraining potential energy term that aims at forcing the values of S^2 as obtained by Eq. 2 to be close to a target value S^2_{exp} derived from measured data. When applied to systems of practical interest, all restraining methods suffer from various sources of uncertainty or error (Olsson et al. 2013): (1) errors or noise in the S^2_{exp} values; (2) inconsistency between different S^2_{exp} values used in the restraining; (3) propagation of deficiencies of the procedure or function f into the configurational distribution $P(\vec{r}^N)$ through the particular restraining procedure; (4) impossibility to represent the time-averaging window inherent to the NMR experiment used to determine S^2_{exp} values. These sources of uncertainty cannot be quantified. Their effect may be seen by analysing time series or distributions of values of a set of (observable) quantities of a system and comparing these with their counterparts resulting from unrestrained MD simulation.

In the present work a time-averaging restraining framework for S^2 order parameter restraining is proposed. The mathematical equations are given and the method is tested with respect to the influence of a variation of the values of its parameters upon some characteristic properties of the configurational ensemble generated using two model systems, a vector with one end restrained in space and a pentapeptide. The applicability to proteins in solution is illustrated for the B3 domain of protein G.

Theory

The time-correlation function $\hat{C}(t)$ that describes the relaxation due to dipole-dipole interaction between two nuclei X and Y connected by the internuclear vector $\vec{r}(t)$ of length $r(t)$ and orientation $\theta(t)$, $\varphi(t)$, is given by (Lipari and Szabo 1982)

$$\hat{C}(t) = \frac{4\pi}{5}\sum_{m=-2}^{2}\left\langle \frac{Y_{2m}(\theta^{\text{lab}}(\tau),\varphi^{\text{lab}}(\tau))Y^*_{2m}(\theta^{\text{lab}}(\tau+t),\varphi^{\text{lab}}(\tau+t))}{r(\tau)^3 r(\tau+t)^3}\right\rangle_\tau, \quad (7)$$

where $Y_{2m}(\theta^{\text{lab}},\varphi^{\text{lab}})$ is the second-order spherical harmonic function defined in a laboratory coordinate frame and $\langle \ldots \rangle_\tau$ denotes averaging over the initial times τ. We use the hat-accent to distinguish correlation functions and order parameters that have a dimension of length^{-6} from their dimensionless counterparts. Assuming an isotropically tumbling molecule the overall rotational motion may be decoupled from intramolecular motions, and the correlation function can be factorized,

$$\hat{C}(t) = C_{\text{rot}}(t)\hat{C}_{\text{int}}(t). \quad (8)$$

If we assume that the overall rotational motion correlation decays exponentially (i. e. random or Brownian rotation), we have

$$C_{\text{rot}}(t) = e^{-t/\tau_{\text{rot}}} = e^{-6D_{\text{rot}}t} \qquad (t > 0) \quad (9)$$

where τ_{rot} and D_{rot} are the correlation time and rotational diffusion constant of the macromolecule, respectively. The expression for the intramolecular correlation function $\hat{C}_{\text{int}}(t)$ is now as Eq. 7 but with $\theta(t)$ and $\varphi(t)$ expressed with respect to a molecular coordinate frame, $\theta^{\text{mol}}(t)$ and $\varphi^{\text{mol}}(t)$ (Brüschweiler et al. 1992),

$$\hat{C}_{\text{int}}(t) = \frac{4\pi}{5}\sum_{m=-2}^{2}\left\langle \frac{Y_{2m}(\theta^{\text{mol}}(\tau),\varphi^{\text{mol}}(\tau))Y^*_{2m}(\theta^{\text{mol}}(\tau+t),\varphi^{\text{mol}}(\tau+t))}{r(\tau)^3 r(\tau+t)^3}\right\rangle_\tau. \quad (10)$$

Using the addition theorem of spherical harmonics

$$P_l(\cos(\theta_{12})) = \frac{4\pi}{2l+1}\sum_{m=-l}^{+l} Y_{lm}(\theta_1,\varphi_1)Y_{lm}^*(\theta_2,\varphi_2) \tag{11}$$

in which θ_{12} is the angle between the radius vectors (θ_1,φ_1) and (θ_2,φ_2) and

$$P_2(\cos\theta_{12}) = \frac{1}{2}\left[3\cos^2\theta_{12}-1\right], \tag{12}$$

we get

$$\begin{aligned}\hat{C}_{\text{int}}(t) &= \left\langle \frac{P_2(\cos\theta_{12}(\tau,t))}{r(t)^3 r(\tau+t)^3}\right\rangle_\tau \\ &= \frac{1}{2}\left\langle \frac{3\cos^2\theta_{12}(\tau,t)-1}{r(t)^3 r(\tau+t)^3}\right\rangle_\tau .\end{aligned} \tag{13}$$

We may also write this expression in Cartesian tensor notation defining (Henry and Szabo 1985)

$$\mu_1 = \frac{x}{r},\quad \mu_2 = \frac{y}{r},\quad \mu_3 = \frac{z}{r}\quad \text{or}\quad \vec{\mu} = \frac{\vec{r}}{r} \tag{14}$$

and

$$\Phi_{\alpha\beta}(t) = \frac{\mu_\alpha(t)\mu_\beta(t)}{r^3(t)} \tag{15}$$

which leads to

$$\hat{C}_{\text{int}}(t) = \frac{1}{2}\{3\,\text{tr}\langle\Phi(0)\Phi(t)\rangle - \langle\text{tr}\Phi(0)\text{tr}\Phi(t)\rangle\} \tag{16}$$

with

$$\text{tr}\Phi = \sum_{i=1}^{3}\Phi_{ii}. \tag{17}$$

In the limit of fast internal motion compared to overall rotation we get using the property of correlation functions that $\lim_{t\to\infty}\langle A(0)A(t)\rangle = \langle A\rangle^2$,

$$\begin{aligned}\lim_{t\to\infty}\hat{C}_{\text{int}}(t) &= \frac{1}{2}\left\{3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}\langle\Phi_{\alpha\beta}\rangle^2 - \left\langle\sum_{\alpha=1}^{3}\Phi_{\alpha\alpha}\right\rangle^2\right\} \\ &= \frac{1}{2}\left\{3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}\left\langle\frac{\mu_\alpha(\tau)\mu_\beta(\tau)}{r^3(\tau)}\right\rangle_\tau^2 - \left\langle\frac{1}{r^3(\tau)}\right\rangle_\tau^2\right\} \\ &= \hat{S}^2 = \frac{S^2}{(r^{\text{eff}})^6}\end{aligned} \tag{18}$$

where $\hat{S}^2$ is a generalized order parameter (Lipari and Szabo 1982), S^2 is a dimensionless order parameter and r^{eff} is an effective internuclear distance between atoms X and Y, evaluated for example as $\langle 1/r^6(\tau)\rangle_\tau^{-1/6}$ (Brüschweiler et al. 1992).

Note that the internal molecular reference frame, although commonly used, is far from being defined unambiguously. The reason of this is that due to the dynamics of the molecule there is no "fix part" which could be objectively chosen for reference (Gáspári and Perczel 2010; Gapsys and de Groot 2013). In practical applications, measurement of geometric parameters relative to the reference frame is often implemented by using least-squares fitting of backbone atoms in regular secondary structure before calculating the values.

Order parameters as restraints

To use order parameters $S^2_{\text{XY}}(\text{exp})$ derived from experiment to restrain the motion of atoms X and Y in an MD simulation we use the following restraining function

$$V^{\text{restr}}(\vec{r}^N(t)) = \frac{1}{2}K^{\text{sr}}\left[\overline{S^2_{\text{XY}}(\vec{r}^N(t))} - S^2_{\text{XY}}(\text{exp})\right]^2 \tag{19}$$

or a flat bottom alternative (Christen et al. 2007), allowing for some uncertainty in the reference $S^2_{\text{XY}}(\text{exp})$ values. With K^{sr} we denote the force constant and $\overline{S^2_{\text{XY}}(\vec{r}^N(t))}$ is the time-averaged order parameter calculated from

$$\overline{S^2_{\text{XY}}(t)} = \frac{1}{2}\left\{3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}\left[\overline{Q_{\alpha\beta}(t)}\right]^2 - \left[\overline{D(t)}\right]^2\right\}\cdot(r^{\text{eff}}_{\text{XY}})^6 \tag{20}$$

The time averaged quantities $\overline{Q_{\alpha\beta}(t)}$ and $\overline{D(t)}$ are calculated in the usual damped memory manner (Torda et al. 1989) with the memory relaxation time τ_{sr},

$$\overline{Q_{\alpha\beta}(t)} = \frac{1}{\tau_{\text{sr}}[1-e^{-t/\tau_{\text{sr}}}]}\int_0^t e^{-(t-t')/\tau_{\text{sr}}}Q_{\alpha\beta}(t')dt' \tag{21}$$

and

$$\overline{D(t)} = \frac{1}{\tau_{\text{sr}}[1-e^{-t/\tau_{\text{sr}}}]}\int_0^t e^{-(t-t')/\tau_{\text{sr}}}D(t')dt' \tag{22}$$

and

$$Q_{\alpha\beta}(t') = \frac{(r_{\text{X}\alpha}(t')-r_{\text{Y}\alpha}(t'))(r_{\text{X}\beta}(t')-r_{\text{Y}\beta}(t'))}{(r_{\text{XY}}(t'))^5} \tag{23}$$

and

$$D(t') = \frac{1}{r_{\text{XY}}(t')^3} \tag{24}$$

with

$$\vec{r}_{XY} = \vec{r}_X - \vec{r}_Y \quad \text{and} \quad r_{XY} = [(\vec{r}_X - \vec{r}_Y)\cdot(\vec{r}_X - \vec{r}_Y)]^{1/2} \tag{25}$$

and

$$\begin{aligned} r_{X1} &= \text{x-component of vector } \vec{r}_X \\ r_{X2} &= \text{y-component of vector } \vec{r}_X \\ r_{X3} &= \text{z-component of vector } \vec{r}_X \end{aligned} \tag{26}$$

and likewise for $\vec{r}_Y$. The discretized form, applicable to atomic trajectories, in which configurations are separated by a time interval Δt, is

$$Q_{\alpha\beta}(m\Delta t) = \frac{(r_{X\alpha}(m\Delta t) - r_{Y\alpha}(m\Delta t))(r_{X\beta}(m\Delta t) - r_{Y\beta}(m\Delta t))}{(r_{XY}(m\Delta t))^5} \tag{27}$$

and

$$\overline{Q_{\alpha\beta}(n\Delta t)} = Q_{\alpha\beta}(n\Delta t)\left\{1 - e^{-\Delta t/\tau_{sr}}\right\} + e^{-\Delta t/\tau_{sr}}\overline{Q_{\alpha\beta}((n-1)\Delta t)} \tag{28}$$

and

$$\overline{S^2_{XY}(n\Delta t)} = \frac{1}{2}\left\{3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}\left[\overline{Q_{\alpha\beta}(n\Delta t)}\right]^2 - \left[\overline{D(n\Delta t)}\right]^2\right\}\cdot(r^{\text{eff}}_{XY})^6 \tag{29}$$

with

$$D(m\Delta t) = \frac{1}{r_{XY}(m\Delta t)^3} \tag{30}$$

and

$$\overline{D(n\Delta t)} = D(n\Delta t)\left\{1 - e^{-\Delta t/\tau_{sr}}\right\} + e^{-\Delta t/\tau_{sr}}\overline{D((n-1)\Delta t)}. \tag{31}$$

The restraining force on atom X then becomes

$$\begin{aligned} \vec{f}_X(t) = -\frac{\partial V^{\text{restr}}(\vec{r}^N(t))}{\partial \vec{r}_X(t)} &= -K^{sr}\left[\overline{S^2_{XY}(\vec{r}^N(t))} - S^2_{XY}(\text{exp})\right] \\ &\cdot\frac{1}{2}\left\{3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}2\overline{Q_{\alpha\beta}(t)}\frac{\partial\overline{Q_{\alpha\beta}(t)}}{\partial Q_{\alpha\beta}(t)}\frac{\partial Q_{\alpha\beta}(t)}{\partial \vec{r}_X(t)}\right. \\ &\left. -2\overline{D(t)}\frac{\partial\overline{D(t)}}{\partial D(t)}\frac{\partial D(t)}{\partial \vec{r}_X(t)}\right\}\cdot(r^{\text{eff}}_{XY})^6 \end{aligned} \tag{32}$$

and the restraining force on atom Y becomes

$$\begin{aligned} \vec{f}_Y(t) = -\frac{\partial V^{\text{restr}}(\vec{r}^N(t))}{\partial \vec{r}_Y(t)} &= -K^{sr}\left[\overline{S^2_{XY}(\vec{r}^N(t))} - S^2_{XY}(\text{exp})\right] \\ &\cdot\frac{1}{2}\left\{3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}2\overline{Q_{\alpha\beta}(t)}\frac{\partial\overline{Q_{\alpha\beta}(t)}}{\partial Q_{\alpha\beta}(t)}\frac{\partial Q_{\alpha\beta}(t)}{\partial \vec{r}_Y(t)}\right. \\ &\left. -2\overline{D(t)}\frac{\partial\overline{D(t)}}{\partial D(t)}\frac{\partial D(t)}{\partial \vec{r}_Y(t)}\right\}\cdot(r^{\text{eff}}_{XY})^6 \end{aligned} \tag{33}$$

although using Eq. 28 we have

$$\frac{\partial\overline{Q_{\alpha\beta}(t)}}{\partial Q_{\alpha\beta}(t)} = \frac{\partial\overline{Q_{\alpha\beta}(n\Delta t)}}{\partial Q_{\alpha\beta}(n\Delta t)} = \left[1 - e^{-\Delta t/\tau_{sr}}\right] \tag{34}$$

the approximation (Scott et al. 1998)

$$\frac{\partial\overline{Q_{\alpha\beta}(t)}}{\partial Q_{\alpha\beta}(t)} = 1 \tag{35}$$

is often used, which only leads to a rescaling of K^{sr} in practice, and likewise

$$\frac{\partial\overline{D(t)}}{\partial D(t)} = 1. \tag{36}$$

For the derivatives $\frac{\partial Q_{\alpha\beta}(t)}{\partial \vec{r}_X(t)}$ we find using Eq. 23, where we omit the variable t and denote the three components of the position vector $\vec{r}_X$ of atom X by $r_{X\gamma}$ with $\gamma = 1, 2, 3,$

$$\begin{aligned} \frac{\partial Q_{\alpha\beta}}{\partial r_{X_\gamma}} &= \frac{(r_{XY})^5\{\delta_{\gamma\alpha}(r_{X\beta} - r_{Y\beta}) + \delta_{\gamma\beta}(r_{X\alpha} - r_{Y\alpha})\}}{(r_{XY})^{10}} \\ &\quad - \frac{(r_{X\alpha} - r_{Y\alpha})(r_{X\beta} - r_{Y\beta})\cdot 5(r_{XY})^4(r_{X\gamma} - r_{Y\gamma})(r_{XY})^{-1}}{(r_{XY})^{10}} \\ &= \frac{(r_{XY})^2\{\delta_{\gamma\alpha}(r_{X\beta} - r_{Y\beta}) + \delta_{\gamma\beta}(r_{X\alpha} - r_{Y\alpha})\} - 5(r_{X\alpha} - r_{Y\alpha})(r_{X\beta} - r_{Y\beta})(r_{X\gamma} - r_{Y\gamma})}{(r_{XY})^7} \end{aligned} \tag{37}$$

where δ_{ij} is the Kronecker delta. For the derivatives $\frac{\partial Q_{\alpha\beta}(t)}{\partial \vec{r}_Y(t)}$ we find likewise

$$\frac{\partial Q_{\alpha\beta}}{\partial r_{Y\gamma}} = -\frac{\partial Q_{\alpha\beta}}{\partial r_{X\gamma}}. \tag{38}$$

Thus the restraining force on atom Y is the negative of the restraining force on atom X for this restraining function. For the derivatives $\frac{\partial D(t)}{\partial \vec{r}_X(t)}$ we find using Eq. 24 likewise

$$\frac{\partial D}{\partial r_{X\gamma}} = \frac{-3(r_{X\gamma} - r_{Y\gamma})}{(r_{XY})^5} \tag{39}$$

and

$$\frac{\partial D}{\partial r_{X\gamma}} = -\frac{\partial D}{\partial r_{Y\gamma}}. \tag{40}$$

Note that the use of order parameters to bias MD simulations is conceptually different from the use of other properties such as nuclear Overhauser enhancement (NOE) atom-atom distance bounds or ^3J-coupling constants, because order parameters are not instantaneous observables, i.e. the order parameter for a single configuration is by definition equal to unity. Therefore, τ_{sr} represents not only the memory relaxation but also the experimentally determined averaging period. Thus it should be chosen larger than the decay time of the internal autocorrelation function of the vector connecting the two atoms, but not larger than the sensitivity time window of the NMR experiment.

Computational details

Simulated systems

The method was first tested on two systems, a single vector with one end restrained to the origin of the coordinate system and a freely rotating α-pentapeptide. In both cases the solvent was treated as external field acting on the solute by stochastic and frictional forces. The reason for using stochastic dynamics (SD) (van Gunsteren et al. 1981; Yun-Yu et al. 1988) instead of explicit solvent is that this allows a much longer simulation period. But, biomolecular force fields, such as the GROMOS one used here, are generally developed, i.e. their parameters calibrated, for use in the condensed phase, i.c. using water as solvent. Since water has a dielectric permittivity of 78 at room temperature and pressure, it has a large damping effect on Coulomb interactions within the solute. Thus the force field used in the pentapeptide simulation may not perform well under vacuum boundary conditions, which may induce conformational transitions that are an artefact of the toy character of the model. The restraining force may also induce conformational transitions in the solute, if the restraining force or energy becomes large, i.e. in the case the restraining energy is not a small perturbation of the energy of the system. Thus we report in Table 2 energies of unrestrained and restrained simulations in order to illustrate that the restraining perturbation is small. Additionally, in a simulation of the peptide without explicit solvent, the peptide is more easily trapped in local energy minima (Daura et al. 1999), which allows focusing the order parameter analysis on local fluctuations.

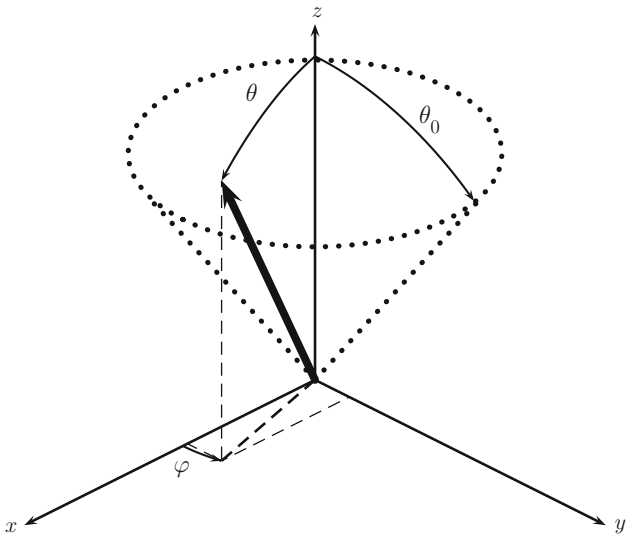

Fig. 1 Illustration of a bond vector moving in the angular region $0 \le \theta \le \theta_0 < \pi$, $-\pi \le \varphi \le \pi$

A vector of constant length, with orientation $\Omega = (\theta, \varphi)$, moving randomly in the angular region $0 \le \theta \le \theta_0 < \pi$, $-\pi \le \varphi \le \pi$, where θ and φ are the polar and azimuth angles, respectively (see Fig. 1) was simulated using two non-interacting atoms of mass 5 and 5,000 amu, respectively, the latter one harmonically positionally restrained to the origin with a force constant of 10,000 kJ mol^{-1} nm^{-2}, connected through a constrained bond of length 0.153 nm.

The α-pentapeptide had the amino acid sequence Val^1–Tyr^2–Arg^3–Lys^4–Gln^5. Initial coordinates were taken from the tutorial files of the GROMOS11 program package (http://www.gromos.net).

The method was subsequently applied to the B3 domain of protein G, further called GB3, in aqueous solution. Fifty backbone N–H order parameters derived from NMR experiments using an anisotropic model of the overall tumbling were taken from Hall and Fushman (Hall and Fushman 2003). Initial coordinates were taken from the NMR model structure deposited in the PDB as entry 2OED.

Simulation parameters

All simulations were carried out using a modified version of the GROMOS11 program package (Schmid et al. 2011a, 2012). For the SD simulations the 54B7 force field (Schmid et al. 2011b) was used which is the one corresponding to the 54A7 force field, but adapted in order to be used for simulations of biomolecules in vacuum. These simulations were performed under vacuum boundary conditions. The Langevin equations of motion were integrated using the leapfrog scheme (van Gunsteren and Berendsen 1988) with a timestep of 0.5 fs for the randomly moving vector and 2 fs for the peptide. Bond lengths were constrained by application of the SHAKE procedure (Ryckaert et al. 1977) with a relative geometric tolerance of 10^{-4}. The reference temperature for all simulations was set to 298 K. A single friction coefficient γ of 91 ps^{-1} was used for all atoms.

In case of the peptide, the non-bonded van der Waals and electrostatic interactions were calculated using a twin-range cutoff scheme (Berendsen 1985), with short- and long-range cutoff distances set to 0.8 and 1.4 nm, respectively. The short-range interactions were calculated every timestep using a group-based pairlist updated every fifth timestep. The intermediate-range interactions were re-evaluated at each pairlist update and assumed constant in between. A reaction-field correction (Barker and Watts 1973; Tironi et al. 1995) was applied to account for the mean effect of electrostatic interactions beyond the long-range cutoff distance, using a relative dielectric permittivity of one. The simulations were carried out for 110 ns of which the latter 100 ns were used for analysis. Atom coordinates and energies were saved for analysis every 10 ps.

For protein GB3, MD simulations were carried out using the 54A7 force field (Schmid et al. 2011b). The simulations were performed under minimum image periodic boundary conditions based on cubic boxes, using 7,486 simple point charge (SPC) water molecules (Berendsen et al. 1981). The equations of motion were integrated using the leapfrog algorithm (Hockney 1970) with a timestep of 2 fs. Bond lengths and the bond angle of water molecules were constrained by applying the SHAKE algorithm (Ryckaert et al. 1977) with a relative geometric tolerance of 10^{-4}. The center of mass motion of the computational box was removed every 2 ps. All simulations were performed at constant pressure and temperature. The temperature was maintained at 298 K by weak coupling to an external bath (Berendsen et al. 1984) with a relaxation time of 0.1 ps. Solute and solvent were coupled to separate heat baths. The pressure was calculated using a group-based virial and held constant at 1 atm using the weak coupling method with a relaxation time of 0.5 ps (Berendsen et al. 1984) and an isothermal compressibility of 4.575×10^{-4} $(\mathrm{kJ\,mol^{-1}\,nm^{-3}})^{-1}$.

The nonbonded van der Waals and electrostatic interactions were calculated as described above, but with a relative dielectric permittivity ε_{RF} of 61, appropriate for SPC water (Heinz et al. 2001). The reaction-field self-term and excluded-atom-term contributions to the energy, forces, and virial were included as described previously (Christen et al. 2005). The simulations were carried out for 25.5 ns of which the latter 25 ns were used for analysis. Atom coordinates and energies were saved for analysis every 2 ps.

Trajectory analysis

All analyses were carried out using the GROMOS++ suite of programs (Eichenberger et al. 2011). The moving vector was analyzed in terms of the (θ, φ)-space sampled, the distribution of the θ-angle and the plateau value of the internal autocorrelation function, Eq. 41.

The pentapeptide was analyzed in terms of the time series and fluctuations of the φ-ψ angles and N–H order parameters of residues 2–4 after superimposing all structures according to the backbone atoms of these residues.

The protein was analyzed in terms of the N–H order parameters of all residues after superimposing all structures according to the backbone atoms of residues 3–56.

In the present study only atom pairs with constrained internuclear distance were considered. Because the effect of constraining bond lengths on simulated order parameters is negligible (Pfeiffer et al. 2001), the expression for the internal autocorrelation function can be simplified to

$$C_{\mathrm{int}}(t) = \frac{1}{2}\left\langle 3\cos^2\theta(\tau, t) - 1\right\rangle_{\tau} \tag{41}$$

and the order parameter S^2 is obtained from

$$\lim_{t\to\infty} C_{\mathrm{int}}(t) = \frac{1}{2}\left\{3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}\left\langle \mu_{\alpha}(\tau)\mu_{\beta}(\tau)\right\rangle_{\tau}^{2} - 1\right\} = S^2. \tag{42}$$

Results and discussion

While the interpretation of trajectories of the vector moving in the (θ, φ)-space in terms of an order parameter is straightforward, there are subtleties present in the interpretation of the peptide and protein trajectories. First, order parameters are defined in a molecule-fixed reference frame while the equations of motion are integrated in a simulation box-fixed coordinate frame. This may cause an underestimation of the time-averaged order parameter if the memory relaxation time is not significantly smaller than the characteristic time for overall tumbling of the molecule. In

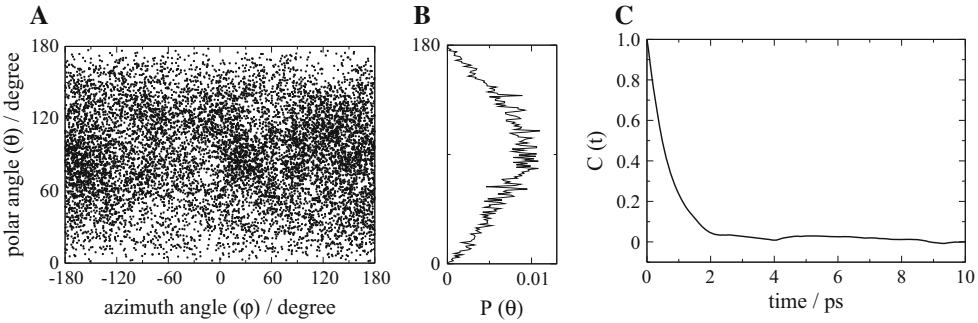

Fig. 2 Analysis of the unrestrained simulation of 500 ps of a vector moving randomly in the angular region. **a** Distribution in (θ, φ)-space, **b** distribution of the polar angle θ, **c** internal autocorrelation function (Eq. 41)

post-simulation analysis of proteins, order parameters are usually calculated after a least-squares fit of the backbone atom positions of trajectory structures in order to remove overall rotation of the molecule. However, in case of major structural transitions in the backbone this fit may introduce spurious rotations leading to a decrease of the order parameter. If such transitions are observed, Eq. 42 should be applied separately to those parts of the trajectories which are in the same conformational state.

Vector moving in (θ, φ)-space

In an unrestrained SD simulation the motion of the bond vector is isotropic, shown in Fig. 2a, leading to an order parameter of $S^2 = 0$. In the setup studied here the motion is in addition azimuthally symmetric about the z-axis, making the order parameter dependent on the θ-angle only. The latter is distributed symmetrically with respect to the x-y plane, see Fig. 2b. In that case, also referred to as diffusion in a cone model, the order parameter vanishes not only for $\theta = \pi$ but also for $\theta = \pi/2$ (Brainard and Szabo 1981; Lipari and Szabo 1982). Therefore, we restrict the analysis of the restrained simulations to one of the two hemispheres. Figure 2c shows the internal autocorrelation function of the bond vector, which decreases to zero quickly.

The restrained simulations using time-averaging with an exponentially decaying memory are characterized by two parameters, i.e. the force constant K^{sr} and the memory relaxation time τ_{sr}. The force constant controls the relative weight of the artificial restraining energy with respect to the force-field energy while τ_{sr} determines the length of the exponential decay in Eqs. 28 and 31 and the system's sensitivity to its past. Figure 3 summarises a series of simulations in which the two parameters were systematically varied. Note that for the vector randomly moving in the (θ, φ)-space no contribution from the physical force field is present, i.e. the restraining forces only act against the forces from the stochastic thermostat. Agreement with the imposed restraints was judged by comparing the imposed order parameter with the plateau value of the internal autocorrelation function (Eq. 41) of the vector, while the influence of the restraining parameters on the configurational ensemble is evaluated from the distributions of the polar angles sampled, and the dynamics of the vector is also analyzed (Fig. 4). The rows of Fig. 3 correspond to the imposed S^2 values of $S_0^2 = 1.0, 0.9, 0.7$ and 0.5 while the columns correspond to different force constants. The different curves in each panel correspond to different values of the memory relaxation time τ_{sr}. The plateau values of the internal autocorrelation function are given in the legends of each panel. Note that in the simulations a flat bottom restraining potential energy function

$$V^{\mathrm{restr}}(S^2) = 0 \quad \text{if} \quad S_0^2 - \Delta S^2 < S^2 < S_0^2 + \Delta S^2, \tag{43}$$

was used allowing for differences of up to $\Delta S^2 = 0.01$ between internal order parameters and imposed ones without resulting in a restraining force. Only those results are reported in Fig. 3 for which the maximum of the θ-angle distribution was clearly less than $\pi/2$. At constant relaxation times, increased force constants lead to more narrow distributions of the polar angle. At constant force constant, increasing relaxation times lead to more narrow distributions of the polar angle. Both trends are to be expected because a larger force constant penalizes deviation from the imposed order parameter more strongly, while long relaxation times usually decrease the internal order parameter leading to larger deviation from the imposed one and thus to a larger force according to Eq. 19. Restraining the motion to a restricted angular region is not possible for all combinations of K^{sr} and τ_{sr}. The larger the force constant the smaller is the minimal value of τ_{sr} for which restraining is still observed. However, good agreement with the imposed value is only achieved for the largest force constants and relaxation times. To further study the role of the relaxation time, the dynamic behavior of the interatomic vector was investigated. Figure 4 shows the polar coordinates of the positionally unrestrained atom as function of time for an imposed order parameter of $S_0^2 = 0.9$ and a force constant of 1,000 kJ mol^{-1}. The different rows demonstrate the effect of different values of τ_{sr} on the time series of the φ and θ angles, i.e. the right panels in rows B, C, and D correspond to the black, green and orange

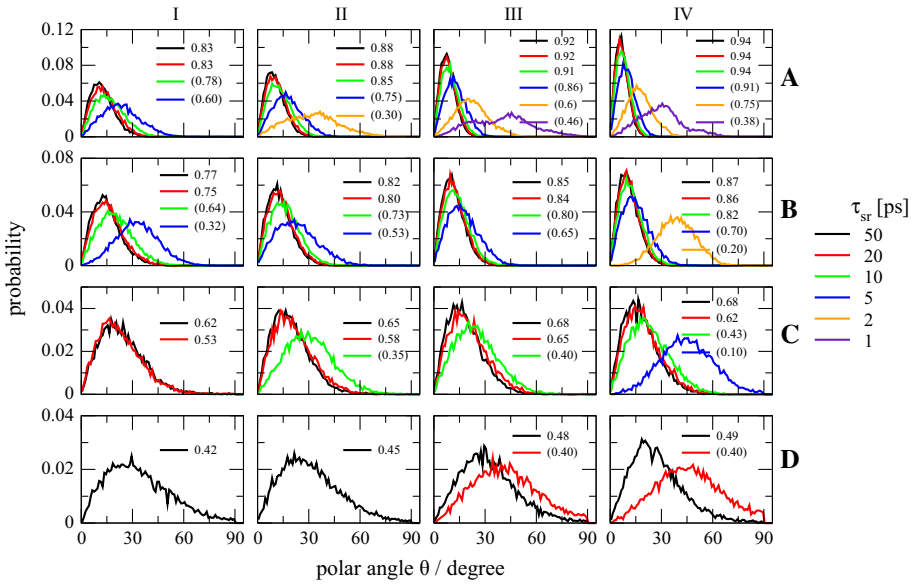

Fig. 3 Effect of restraining parameters on the distribution of the polar angle of a vector moving in the angular region $0 \leq \theta \leq \pi/2$. Each *panel* shows the distributions of the θ-angle for different values of τ_{sr}. The plateau values of the internal autocorrelation function are reported in the legends of each *panel*. *Numbers in parentheses* indicate that no distinct plateau value could be extracted. The force constants were constant for *each column* and set to *I* 100 kJ mol^{-1}, *II* 200 kJ mol^{-1}, *III* 500 kJ mol^{-1}, *IV* 1,000 kJ mol^{-1}. The imposed order parameters S_0^2 were constant for each *row* and set to **a** 1.0, **b** 0.9, **c** 0.7, **d** 0.5. A flat bottom restraining function (Eqs. 19, 43) with a flat range of $\Delta S^2 = 0.01$ was used

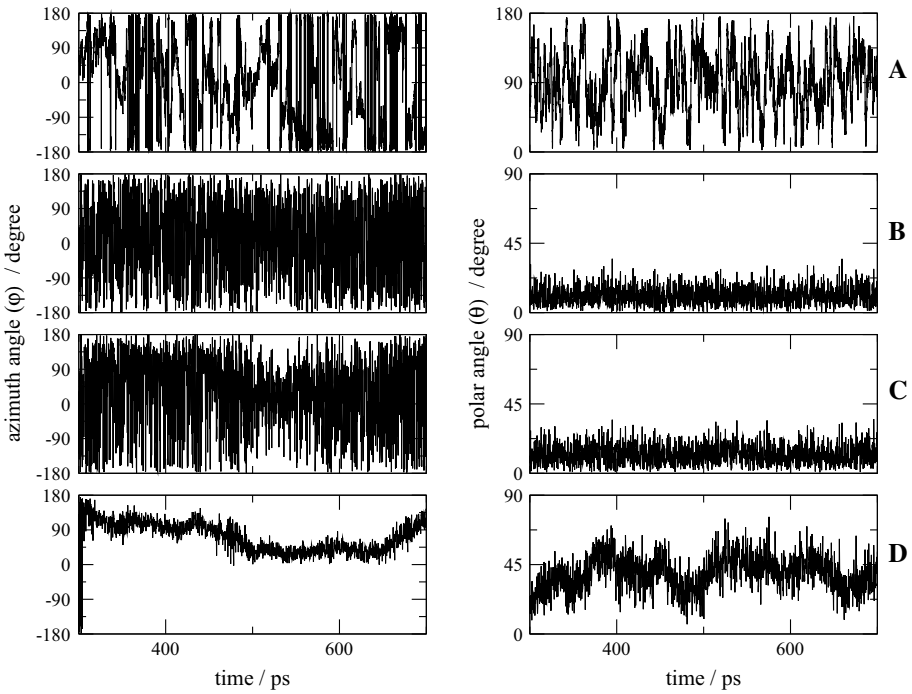

Fig. 4 Effect of the memory relaxation time τ_{sr} on the dynamic behavior of a vector moving in the angular region $0 \leq \theta \leq \pi/2$. The *columns* show the φ and θ coordinates (from *left to right*) of the positionally unrestrained end of the vector. The *topmost panels* **a** show the results for the unrestrained simulation. For the other *rows* τ_{sr} was set to the following values: **b** 50 ps, **c** 10 ps, **d** 2 ps. The force constant K^{sr} was set to 1,000 kJ mol^{-1} in all restrained simulations. The imposed order parameter S_0^2 was 0.9 and $\Delta S^2 = 0.01$

distributions in Fig. 3, panel IV-B. For the smallest relaxation time of 2 ps, corresponding to the panels in the bottom row, the vector is drifting in the coordinate space, a consequence of the decreasing influence of formerly visited configurations on the restraining forces. The panels in the first row of Fig. 4 show the time series of the polar coordinates of the unrestrained simulation. In the latter, the configuration space is sampled in an irregular manner as expected due to the stochastic nature of the driving forces. The individual coordinates show small amplitude

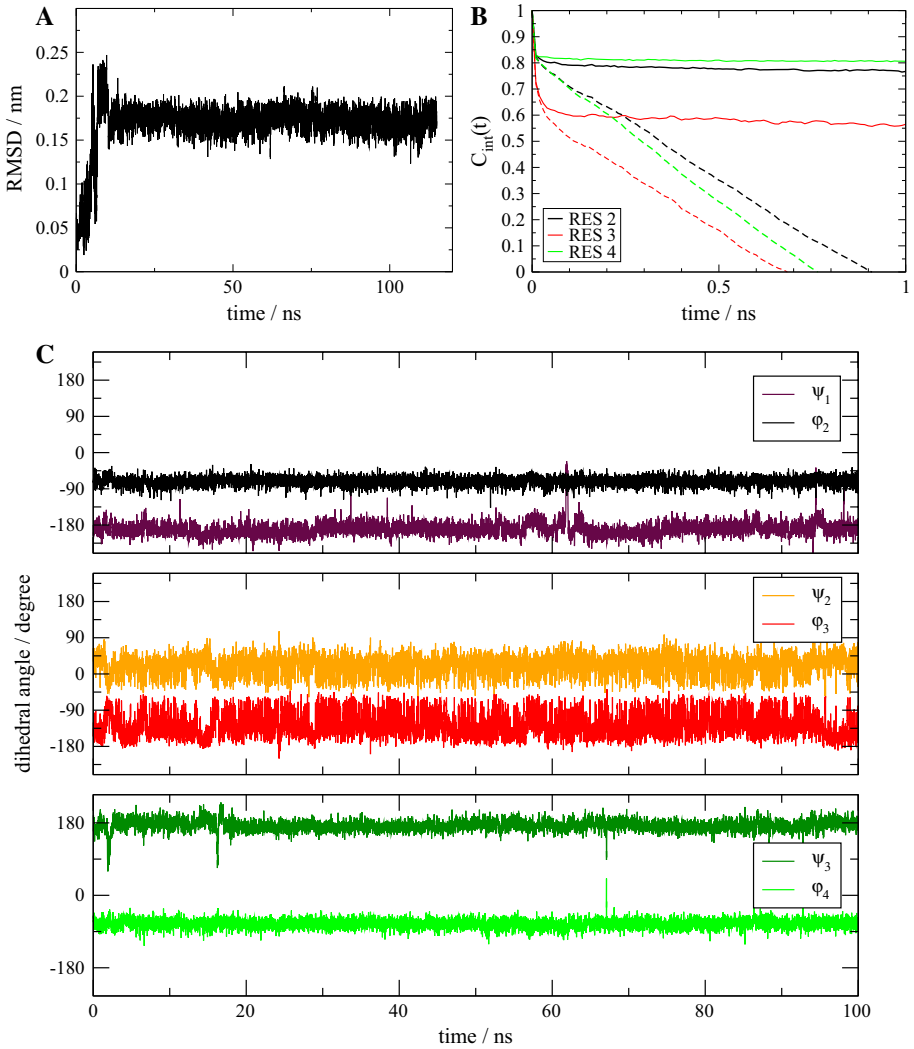

Fig. 5 **a** Backbone atom-positional root-mean-square deviation (RMSD) from the starting structure for the pentapeptide in an unrestrained SD simulation at 298 K. **b** Internal (Eq. 10, *solid lines*) and non-internal (Eq. 7, *dashed lines*, i.e. without removal of overall rotations) autocorrelation function for N–H vectors of residues 2–4 in the stable RMSD regime. The correlation functions are reported in dimensionless form, i.e. after multiplication with r_{NH}^6. **c** Time series of ψ_{i-1} and φ_i angles for residues 2–4 of the last 100 ns of the unrestrained simulation

fluctuations at high frequency and have slower periods of motion of up to 10 ps. The restrained simulations show the same type of small amplitude fluctuations at high frequency. The two rows in the middle demonstrate how the configurational space of the θ-angle is restricted by the restraining force.

Pentapeptide

The results of the unrestrained simulation of the pentapeptide are presented in Fig. 5a in terms of atom-positional root-mean-square deviation (RMSD) from the starting structure as function of time. After some rearrangements in the initial phase, the RMSD adopts a stable regime for the rest of the simulation period of 110 ns, showing only local fluctuations but no major structural transitions. Figure 5b presents the internal autocorrelation function of the N–H bond vectors of residues 2–4 in the regime where the RMSD is stable. The decay time of the internal autocorrelation function is about 10 ps for residues 2 and 4 and about 40 ps for residue 3. The dashed lines in Fig. 5b represent the autocorrelation functions calculated without prior superposition of the backbone atoms, showing the influence of overall tumbling of the molecule. The rotational correlation times for the peptide, calculated from single-exponential fits to the autocorrelation function of the three normalized vectors defined by the C_α atoms of residues 2 and 3, 3 and 4, and by the cross product of the latter (Feenstra et al. 2002) were in the range of 600–760 ps. This defines an upper bound for the memory relaxation time used in Eqs. 28 and 31 which should be considerably smaller than the characteristic time for overall tumbling. Figure 5c displays the time series of the ψ_{i-1} and φ_i angles of residues 2–4 in the stable RMSD regime. As expected, no major conformational transitions take place. In some occasions the ψ_1-angle adopts a different conformation reflecting an enhanced flexibility of the outer residues compared to the inner ones. Consistent with a lower

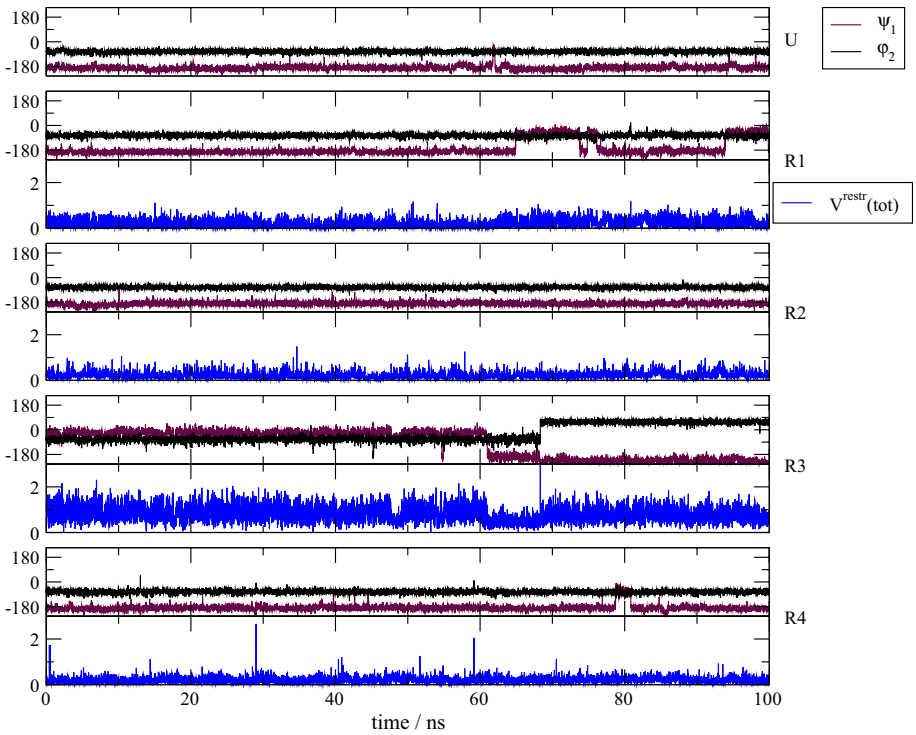

Fig. 6 Time series of the ψ_1 and φ_2 angles of the pentapeptide for four different restraining cases, compared to the unrestrained simulation. The *upper panel* *U* shows the results for the unrestrained simulation. For the other *panels* the parameters were as follows: *R1* $K^{sr} = 100\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, S_0^2 for residues 2–4 taken from the unrestrained simulation, *R2* $K^{sr} = 200\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, $S_0^2(\mathrm{residues2{-}4}) = 0.9$, *R3* $K^{sr} = 200\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, S_0^2 (residues 2−4) = 0.5, and *R4* $K^{sr} = 100\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, $S_0^2(\mathrm{residues\ 2{-}4}) = 0.7$. For each restraining case the time series of the total restraining energy is shown in *blue*. Angles are presented in degree and energies in $\mathrm{kJ\,mol^{-1}}$

plateau value of the internal autocorrelation function of residue 3, i.e. a lower order parameter, the ψ_2 and φ_3-angles show larger fluctuations than the other ψ and φ-angles.

In a series of simulations the force constants and memory relaxation times were systematically varied for four different restraining cases. These represent situations in which (1) the imposed S_0^2 values for the N–H vectors of residues 2–4 are compatible with the force field used, i.e. $S_0^2 = 0.8$, 0.6, and 0.8, taken from the unrestrained simulation (case R1), (2) the imposed S_0^2 values are with a value of $S_0^2 = 0.9$ for all three residues larger than the ones compatible with the force field (case R2), (3) the imposed values $S_0^2 = 0.5$ for all three residues are smaller than the force-field compatible ones (case R3), and (4) the imposed values $S_0^2 = 0.7$ for all three residues are larger for residue 3 and smaller for residues 2 and 4 than the force-field compatible ones (case R4). The restrained simulations were in each case carried out with force constants of 50, 100, 200, 500, 1,000 $\mathrm{kJ\,mol^{-1}}$ and memory relaxation times of 10, 20, 50, 100, 500, 1,000 ps. For each restraining case the results showing the best agreement with the imposed order parameters are presented in Figs. 6, 7 and 8 in terms of the times series of the ψ_{i-1} and φ_i angles of residues 2–4. Table 1 reports the average fluctuations σ_α of the ψ_{i-1} and φ_i angles and the N–H order parameters evaluated according to Eq. 42. The average fluctuations were evaluated as averages of the root-mean-square fluctuations of simulation periods of 1 ns and as the root-mean-square fluctuations of the entire 100 ns trajectory. Likewise, order parameters were evaluated using averaging time windows of 1 and 100 ns. Disagreement between the two numbers points to structural transitions in the backbone during the simulations. An additional evaluation using an averaging time window of 0.5 ns gives very similar results compared to the one using an averaging time window of 1 ns, as expected from the fast decay of the internal autocorrelation functions to their plateau values. For practical applications such as protein simulations in explicit solvent an averaging time window of 1 ns would be realistic, corresponding to the observation time in an NMR experiment (Chandrasekhar et al. 1992; Smith et al. 1995b; Evenäs et al. 1999; Stocker and van Gunsteren 2000; Sapienza and Lee 2010). Therefore, the agreement with the imposed restraints was judged by comparing the imposed order parameters to the

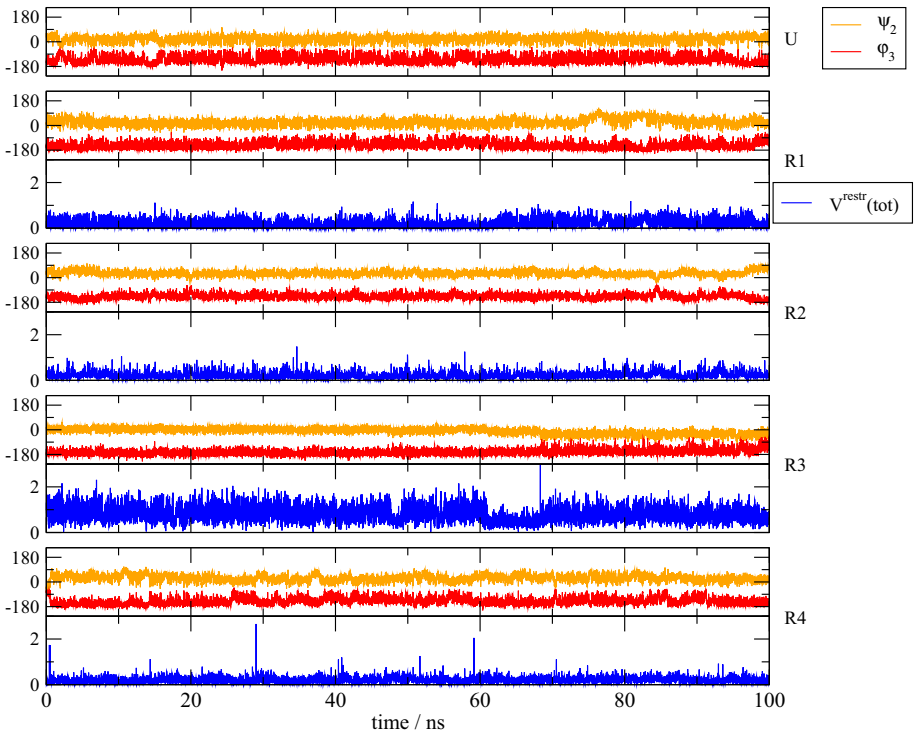

Fig. 7 Time series of ψ_2 and φ_3 angles of the pentapeptide for four different restraining cases, compared to the unrestrained simulation. The *upper panel U* shows the results for the unrestrained simulation. For the other *panels* the parameters were as follows: *R1* $K^{sr} = 100\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, S_0^2 for residues 2–4 taken from the unrestrained simulation, *R2* $K^{sr} = 200\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, $S_0^2(\text{residues } 2{-}4) = 0.9$, *R3* $K^{sr} = 200\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, $S_0^2(\text{residues } 2{-}4) = 0.5$, and *R4* $K^{sr} = 100\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, $S_0^2(\text{residues } 2{-}4) = 0.7$. For each restraining case the time series of the total restraining energy is shown in *blue*. Angles are presented in degree and energies in $\mathrm{kJ\,mol^{-1}}$

ones evaluated with Eq. 42 using an averaging time window of 1 ns.

Although the N–H order parameters result from a combination of molecular motions that cannot directly be interpreted in terms of backbone torsional-angle fluctuations (Marchand and Roux 1998), the change in order parameters upon using restraints is reflected in the average root-mean-square fluctuations of backbone torsional angles. The panels in Figs. 6, 7 and 8 illustrate the relationship between fluctuations of the neighboring ψ_{i-1} and φ_i torsion angles and the N–H order parameter of the i-th residue (Smith et al. 1995a). For residue 3 (Fig. 7), having the smallest order parameter in the unrestrained simulation, the fluctuations in the ψ_2 and φ_3 angles are 23.1° and 27.1°, respectively. For residue 2 (Fig. 6), having the second largest order parameter the fluctuations are $\sigma_{\psi_1} = 13.4°$ and $\sigma_{\varphi_2} = 10.5°$ while for residue 4 (Fig. 8), having the largest order parameter they evaluate to $\sigma_{\psi_3} = 11.7°$ and $\sigma_{\varphi_4} = 10.6°$.

The purpose of restraining case R1 was to identify values for K^{sr} and τ_{sr} that resemble the unrestrained simulation. The best agreement within the investigated parameter space was obtained for a force constant of $100\,\mathrm{kJ\,mol^{-1}}$ and a memory relaxation time of 20 ps. Increasing the relaxation time led to mixing in contributions from overall tumbling, resulting in restraining forces that are too large. Decreasing the relaxation time led to structural transitions in the backbone similar to those seen in case R3 (see below). This behavior is to be expected because for memory relaxation times smaller than the decay time of the internal autocorrelation function, the internal order parameter is overestimated. Increasing the force constants to 200 or 500 $\mathrm{kJ\,mol^{-1}}$ had little effect on the observed order parameters, but slightly reduced the fluctuations in parts of the trajectory. One might wonder why the time series of the dihedral angles of restraining case R1 slightly differ from the unrestrained case. This difference reflects the difficulty of exactly imposing the order parameters of the unrestrained simulations as these are slightly dependent on the averaging window used in the analysis (see Table 1) and thus subject to some uncertainty. As a result the restraining energy, shown in blue in Figs. 6, 7 and 8 is not zero in this case.

For case R2, good agreement with the imposed order parameters of $S_0^2 = 0.9$ was found for a force constant of

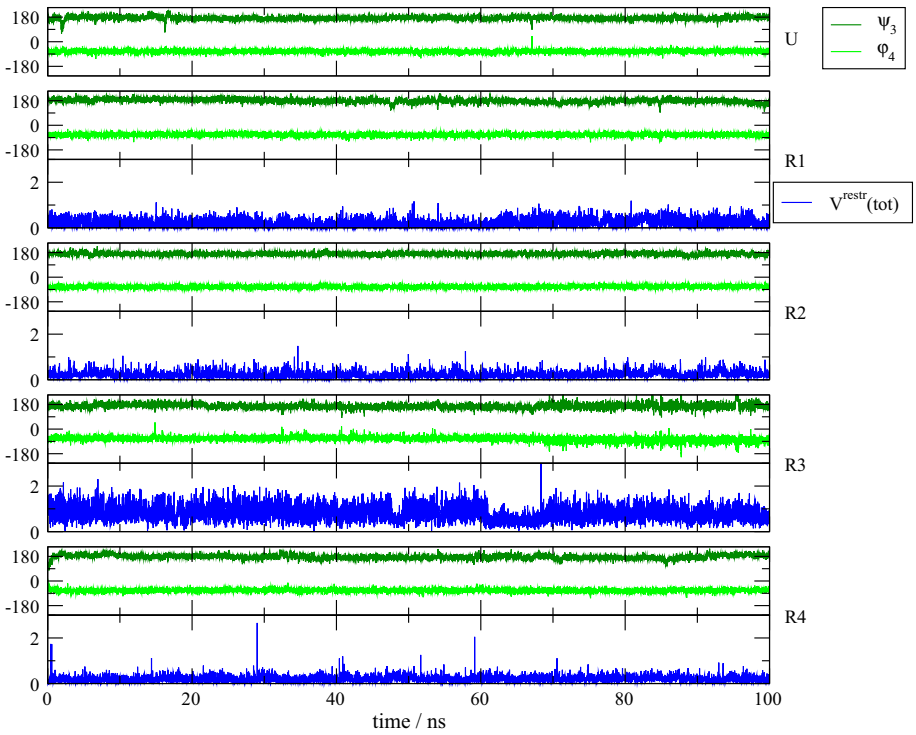

Fig. 8 Time series of ψ_3 and φ_4 angles of the pentapeptide for four different restraining cases, compared to the unrestrained simulation. The *upper panel (U)* shows the results for the unrestrained simulation. For the other *panels* the parameters were as follows: *R1* $K^{sr} = 100\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, S_0^2 for residues 2 to 4 taken from the unrestrained simulation, *R2* $K^{sr} = 200\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, $S_0^2(\text{residues } 2{-}4) = 0.9$, *R3* $K^{sr} = 200\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, $S_0^2(\text{residues } 2{-}4) = 0.5$, and *R4* $K^{sr} = 100\,\mathrm{kJ\,mol^{-1}}$, $\tau_{sr} = 20\,\mathrm{ps}$, $S_0^2(\text{residues } 2{-}4) = 0.7$. For each restraining case the time series of the total restraining energy is shown in *blue*. Angles are presented in degree and energies in $\mathrm{kJ\,mol^{-1}}$

$200\,\mathrm{kJ\,mol^{-1}}$ and a relaxation time of 20 ps. Increasing the force constant did not improve the agreement with the imposed values substantially, but rather led to an oscillation of the φ-angle of residue 3 between two states with rather long residence times of about 20 ns. The fluctuations of the ψ_2 and φ_3 angles decrease significantly compared to the unrestrained simulation as expected because residue 3 has the lowest order parameter in the unrestrained simulation. Case R3 assumes a scenario in which an unrestrained simulation shows too large order parameters. For a force constant of $K^{sr} = 200\,\mathrm{kJ\,mol^{-1}}$ the obtained order parameters are somewhat too large, while $K^{sr} = 500\,\mathrm{kJ\,mol^{-1}}$ leads to good agreement with the imposed order parameters for residues 2 and 4 while for residue 3 the order parameter was underestimated. Note that for case R3 the backbone adopts a different conformation regarding the ψ_1 and φ_2 angles compared to all other cases. Finally, case R4 attempts to lower the order parameters for residues 2 and 4 while increasing the one of residue 3. Using a force constant of $K^{sr} = 100\,\mathrm{kJ\,mol^{-1}}$ leads to the desired values for residues 2 and 4 while the one for residue 3 is still too small. By allowing for individual force constants for each residue the agreement with the imposed values could be improved using $K^{sr} = 200\,\mathrm{kJ\,mol^{-1}}$ for residue 3 and $K^{sr} = 100\,\mathrm{kJ\,mol^{-1}}$ for residues 2 and 4, respectively. The resulting order parameters were 0.75, 0.71 and 0.72 for residues 2, 3 and 4, while the torsional angle fluctuations were 15.0°, 16.0° and 14.9° for $\sigma_{\psi_{i-1}}$. and 9.7°, 20.6° and 15.4° for σ_{φ_i}.

On the basis of the present results a force constant between 100 and $200\,\mathrm{kJ\,mol^{-1}}$ combined with a relaxation time of 20 ps appears to be an appropriate choice. The transferability of these parameters to proteins in solution is discussed below.

Table 2 gives an overview of the most important energetic terms of the simulations. The averaged values for the total potential energy of all five simulations are of comparable size. The fluctuation of the potential energy of the unrestrained simulation is slightly smaller than for the restrained ones. Except for restraining case R3, all bonded and electrostatic energies are very similar to the unrestrained simulation. The reason for the deviation in case R3 is the different conformations the peptide adopted in terms of the ψ_1 and φ_2 angles. Finally, we note that the energies

Table 1 Torsional-angle fluctuations and N–H order parameters as a function of simulation parameters

Simulation[b]	Residue	Torsional-angle fluctuations[a]		N–H order parameters			S_0^2
		$\sigma_{\psi_{i-1}}$	σ_{φ_i}	0.5 ns[c]	1 ns[c]	100 ns[c]	
Unrestrained	2	13.4 (16.5)	10.5 (10.8)	0.80	0.79	0.75	–
	3	23.1 (24.5)	27.1 (29.8)	0.62	0.60	0.56	–
	4	11.7 (13.4)	10.6 (10.9)	0.82	0.81	0.81	–
Case R1	2	15.0 (55.2)	11.3 (11.9)	0.79	0.77	0.70	0.8
$K^{sr} = 100$	3	22.2 (26.4)	23.4 (26.6)	0.62	0.61	0.58	0.6
$\tau_{sr} = 20$	4	12.0 (14.7)	10.4 (11.8)	0.80	0.79	0.77	0.8
Case R2	2	12.0 (13.0)	9.5 (9.7)	0.85	0.85	0.83	0.9
$K^{sr} = 200$	3	16.0 (19.0)	17.4 (20.2)	0.85	0.84	0.81	0.9
$\tau_{sr} = 20$	4	10.1 (10.6)	9.6 (10.0)	0.87	0.87	0.86	0.9
Case R3	2	18.3 (98.5)	15.2 (61.3)	0.60	0.59	0.26	0.5
$K^{sr} = 200$	3	24.6 (24.6)	19.7 (22.8)	0.56	0.56	0.45	0.5
$\tau_{sr} = 20$	4	18.7 (18.7)	15.7 (18.0)	0.56	0.56	0.53	0.5
Case R4	2	13.8 (22.6)	11.8 (12.3)	0.72	0.70	0.63	0.7
$K^{sr} = 100$	3	21.6 (25.3)	22.4 (27.2)	0.66	0.64	0.55	0.7
$\tau_{sr} = 20$	4	12.3 (15.7)	11.1 (11.5)	0.71	0.70	0.66	0.7

[a] σ_α (in degree) was calculated as $\frac{1}{n_b}\sum_{j=1}^{n_b}\sqrt{\frac{1}{n_\alpha/n_b}\sum_{i=(j-1)n_\alpha/n_b+1}^{jn_\alpha/n_b}\left(\alpha_i - \langle\alpha\rangle_j\right)^2}$ where n_α is the total number of angles sampled, split into n_b subsets of equal number of samples. Here we used $n_\alpha = 10{,}000$ and $n_b = 100$, corresponding to trajectory segments of 1 ns length. The values in parentheses were obtained for $n_b = 1$, i.e. averaged over the entire 100 ns trajectory

[b] The units of K^{sr} and τ_{sr} are kJ mol^{-1} and ps, respectively. A flat bottom restraining function with $\Delta S^2 = 0.01$ was used in cases R1–R4.

[c] Averaging time window used in Eq. 42 to obtain S^2 values

Table 2 Average energies in unrestrained (U) and order-parameter restrained (R) simulations of 100 ns (in kJ mol^{-1})

	U	R1	R2	R3	R4
Potential energy[a]	−495.7	−493.6	−495.1	−495.7	−489.2
Fluctuation of potential energy	22.5	24.1	23.5	25.8	24.0
Bond angle energy	118.6	118.8	117.7	126.6	118.9
Improper dihedral energy	37.8	37.9	37.3	41.6	37.9
Proper dihedral energy	64.3	64.0	65.6	65.1	65.4
Electrostatic energy	−655.0	−656.1	−654.2	−662.4	−652.3
van der Waals energy	−61.5	−58.4	−61.6	−66.6	−59.1
Restraining energy	0.0	0.21	0.22	0.82	0.20

[a] Not including the restraining energy

of the restraints are two orders of magnitude smaller than the potential energies of the physical force field terms, which indicates that the energetic bias due to the restraining is very small.

Protein GB3

For GB3 the restrained simulations were carried out with force constants of 100, 200, 300, 400 kJ mol^{-1} and memory relaxation times of 20, 50, 100 ps, restraining 50 out of 56 residues using experimentally derived order parameters reported by Hall and Fushman (Hall and Fushman 2003). Each trajectory was analyzed according to Eq. 42 using two different averaging time windows of 0.5 and 1.0 ns, respectively. Figure 9 shows a comparison of calculated order parameters determined from unrestrained and restrained simulations, the latter carried out with $K^{sr} = 400$ kJ mol^{-1} and $\tau_{sr} = 100$ ps. For the unrestrained simulation the averaging time window has some influence on the order parameters for residues 11–15, showing larger S^2 values for the shorter averaging time, as expected. For the restrained simulation no difference between the two

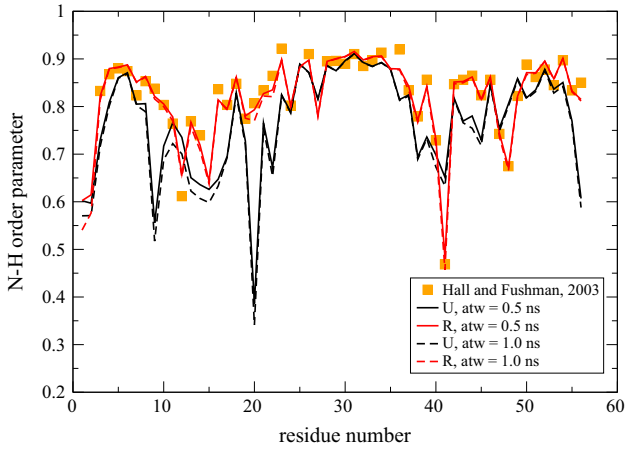

Fig. 9 Comparison of backbone N–H order parameters for protein GB3, determined from unrestrained (U) and restrained (R) MD simulations using different averaging time windows (atw) in the analysis according to Eq. 42. The experimentally derived order parameters used for restraining were taken from the work of Hall and Fushman (2003)) (anisotropic model). The restraining parameters were $K^{sr} = 400\,\mathrm{kJ\,mol^{-1}}$ and $\tau_{sr} = 100\,\mathrm{ps}$. A flat bottom restraining function (Eqs. 19 and 43) with a flat range of $\Delta S^2 = 0.01$ was used

averaging time windows is observed. In the restrained simulations the order parameters are close to the target values except for residues 7, 12 and 36 that show a deviation of about 0.04. However, we expect that the use of individual force constants for these residues would lead to better agreement. The results for the other combinations of K^{sr} and τ_{sr} (not displayed) show that a relaxation time τ_{sr} of 20 ps seems to be too short for some of the residues. Force constants K^{sr} smaller than 400 also lead to good agreement with experimental data, with some deviations around residues 7 and 22. In all cases the restraining energies were at least two orders of magnitude smaller than the intra-solute potential energies, again showing the very small energetic bias of the restraining. A detailed analysis involving cross validation with other NMR observables will be reported separately.

Conclusion

The aim of the present work was to provide a mathematical framework for time-averaged order parameter restraints for use in MD simulations and to report an investigation of the role of the two parameters that characterize these restraints.

As the order parameter is not an instantaneous observable, the memory relaxation time should be chosen such that the plateau value of the internal autocorrelation function of the vector connecting the two atoms is reached. However, its value should also be considerably smaller than the rotational correlation time for overall molecular tumbling. The peptide and protein simulations suggest that values between 100 and 400 $\mathrm{kJ\,mol^{-1}}$ for the force constant are appropriate. In general one should use the smallest force constant that leads to agreement with the experimental data.

The adequacy of the values found for the force constants for other proteins in solution will be evaluated in future work. It may make sense to use different force constants and memory relaxation times for different residues as a measure to minimize perturbing the dynamics resulting from the physical force field. One may think of using relaxation times for different residues derived from NMR experiments. Yet, we refrained from pursuing this option for various reasons: (1) The so-called "model free analysis" by Lipari and Szabo (1982) is de facto based on a model, as is any analysis; (2) The uncertainty in relaxation times for individual bond vectors is large due to the assumptions and approximations of the procedure (called function f in Eq. 2) used to derive relaxation times and order parameters from measured NMR data; (3) Their values depend on an estimate of the rotational tumbling time of the molecule considered; (4) We prefer a restraining procedure that is as simple as possible, with a small number of adjustable parameters. Deriving values for these parameters from experiments particular to the molecule considered may introduce noise, i.e. enhances the chance of fitting the configurational distribution to noise in the experimental data.

Acknowledgments This work was financially supported by the National Center of Competence in Research (NCCR) in Structural Biology and by Grant Number 200020-137827 of the Swiss National Science Foundation, and by grant number 228076 of the European Research Council, which is gratefully acknowledged. N.H. thanks the German Research Foundation (DFG) for financial support within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.

References

Barker JA, Watts RO (1973) Monte Carlo studies of the dielectric properties of water-like models. Mol Phys 26:789–792

Beauchamp KA, Lin YS, Das R, Pande VS (2012) Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. J Chem Theory Comput 8:1409–1414

Berendsen HJC (1985) Treatment of long-range forces in molecular dynamics. In: Hermans J (ed) Molecular dynamics and protein structure. Polycrystal Book Service, Western Springs, pp 18–22

Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) Interaction models for water in relation to protein hydration. In: Pullmann B (ed) Intermolecular forces. Reidel, Dordrecht, pp 331–342

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690

Best RB, Vendruscolo M (2004) Determination of protein structures consistent with NMR order parameters. J Am Chem Soc 126:8090–8091

Brainard JR, Szabo A (1981) Theory for nuclear magnetic relaxation of probes in anisotropic systems: application to cholesterol in phospholipid vesicles. Biochemistry 20:4618–4628

Braun W, Bösch C, Brown LR, Gō N, Wüthrich K (1981) Combined use of proton–proton Overhauser enhancements and a distance geometry algorithm for determination of polypeptide conformations. Application to micelle-bound glucagon. Biochim Biophys Acta 667:377–396

Brüschweiler R, Wright PE (1994) NMR order parameters of biomolecules: a new analytical representation and application to the Gaussian axial fluctuation model. J Am Chem Soc 116:8426–8427

Brüschweiler R, Roux B, Blackledge M, Griesinger C, Karplus M, Ernst RR (1992) Influence of rapid intramolecular motion on NMR cross relaxation rates. A molecular dynamics study of antamanide in solution. J Am Chem Soc 114:2289–2302

Buck M, Bouguet-Bonnet S, Pastor RW, MacKerell AD (2006) Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. Biophys J 90:L36–L38

Cavalli A, Camilloni C, Vendruscolo M (2013) Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. J Chem Phys 138:094112

Chandrasekhar I, Clore GM, Szabo A, Gronenborn AM, Brooks BR (1992) A 500 ps molecular dynamics simulation study of interleukin-1β in water. Correlation with nuclear magnetic resonance spectroscopy and crystallography. J Mol Biol 226:239–250

Christen M, Hünenberger PH, Bakowies D, Baron R, Bürgi R, Geerke DP, Heinz TN, Kastenholz MA, Kräutler V, Oostenbrink C, Peter C, Trzesniak D, van Gunsteren WF (2005) The GROMOS software for biomolecular simulation: GROMOS05. J Comput Chem 26:1719–1751

Christen M, Keller B, van Gunsteren WF (2007) Biomolecular structure refinement based on adaptive restraints using local-elevation simulation. J Biomol NMR 39:265–273

Daura X, Mark AE, van Gunsteren WF (1999) Peptide folding simulations: No solvent required? Comput Phys Commun 123:97–102

d'Auvergne EJ, Gooley PR (2003) The use of model selection in the model-free analysis of protein dynamics. J Biomol NMR 25:25–39

Dolenc J, Missimer JH, Steinmetz MO, van Gunsteren WF (2010) Methods of NMR structure refinement: molecular dynamics simulations improve the agreement with measured NMR data of a C-terminal peptide of GCN4-p1. J Biomol NMR 47:221–235

Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999–2012

Eichenberger AP, Allison JR, Dolenc J, Geerke DP, Horta BAC, Meier K, Oostenbrink C, Schmid N, Steiner D, Wang D, van Gunsteren WF (2011) The GROMOS++ software for the analysis of biomolecular simulation trajectories. J Chem Theory Comput 7:3379–3390

Evenäs J, Forsén S, Malmendal A, Akke M (1999) Backbone dynamics and energetics of a calmodulin domain mutant exchanging between closed and open conformations. J Mol Biol 289:603–617

Feenstra KA, Peter C, Scheek RM, van Gunsteren WF, Mark AE (2002) A comparison of methods for calculating NMR cross-relaxation rates (NOESY and ROESY intensities) in small peptides. J Biomol NMR 23:181–194

Fennen J, Torda AE, van Gunsteren WF (1995) Structure refinement with molecular dynamics and a Boltzmann-weighted ensemble. J Biomol NMR 6:163–170

Fukunishi H, Watanabe O, Takada S (2002) On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction. J Chem Phys 116:9058–9067

Gapsys V, de Groot BL (2013) Optimal superpositioning of flexible molecule ensembles. Biophys J 104:196–207

Gáspári Z, Perczel A (2010) Protein dynamics as reported by NMR. Annu Rep NMR Spectrosc 71:35–75

Gattin Z, Schwartz J, Mathad RI, Jaun B, van Gunsteren WF (2009) Interpreting experimental data by using molecular simulation instead of model building. Chem Eur J 15:6389–6398

Gniewek P, Kolinski A, Jernigan RL, Kloczkowski A (2012) How noise in force fields can affect the structural refinement of protein models. Proteins Struct Funct Bioinf 80:335–341

Gros P, van Gunsteren WF (1993) Crystallographic refinement and structure-factor time-averaging by molecular dynamics in the absence of a physical force field. Mol Sim 10:377–395

Hall JB, Fushman D (2003) Characterization of the overall and local dynamics of a protein with intermediate rotational anisotropy: differentiating between conformational exchange and anisotropic diffusion in the B3 domain of protein G. J Biomol NMR 27:261–275

Harvey TS, van Gunsteren WF (1993) The application of chemical shift calculation to protein structure determination by NMR. In: Angeletti RH (ed) Tech Protein Chem, vol 4. Academic Press, New York, pp 615–622

Heinz TN, van Gunsteren WF, Hünenberger PH (2001) Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. J Chem Phys 115:1125–1136

Henry ER, Szabo A (1985) Influence of vibrational motion on solid state line shapes and NMR relaxation. J Chem Phys 82:4753–4761

Hess B, Scheek RM (2003) Orientation restraints in molecular dynamics simulations using time and ensemble averaging. J Magn Reson 164:19–27

Hockney RW (1970) The potential calculation and some applications. Methods Comput Phys 9:136–211

Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins Struct Funct Bioinf 65:712–725

Huber T, van Gunsteren WF (1998) SWARM-MD: searching conformational space by cooperative molecular dynamics. J Phys Chem A 102:5937–5943

Huber T, Torda AE, van Gunsteren WF (1996) Optimization methods for conformational sampling using a Boltzmann-weighted mean field approach. Biopolymers 39:103–114

Jardetzky O (1980) On the nature of molecular conformations inferred from high-resolution NMR. Biochim Biophys Acta 621:227–232

Jarymowycz VA, Stone MJ (2006) Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. Chem Rev 106:1624–1671

Johnson E (2012) Separability between overall and internal motion: a protein folding problem. Proteins Struct Funct Bioinf 80:2645–2651

Johnson E, Showalter SA, Brüschweiler R (2008) A multifaceted approach to the interpretation of NMR order parameters: a case study of a dynamic α-helix. J Phys Chem B 112:6203–6210

Kim DE, Blum B, Bradley P, Baker D (2009) Sampling bottlenecks in de novo protein structure prediction. J Mol Biol 393:249–260

Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J Am Chem Soc 104:4546–4559

Luginbühl P, Wüthrich K (2002) Semi-classical nuclear spin relaxation theory revisited for use with biological macromolecules. Prog Nucl Magn Reson Spectrosc 40:199–247

MacKerell AD (2004) Empirical force fields for biological macromolecules: overview and issues. J Comput Chem 25:1584–1604

MacKerell AD, Feig M, Brooks CL (2004) Improved treatment of the protein backbone in empirical force fields. J Am Chem Soc 126:698–699

Marchand S, Roux B (1998) Molecular dynamics study of calbindin D_{9k} in the apo and singly and doubly calcium-loaded state. Proteins Struct Funct Bioinf 33:265–284

Misura KMS, Baker D (2005) Progress and challenges in high-resolution refinement of protein structure models. Proteins Struct Funct Bioinf 59:15–29

Nanzer AP, van Gunsteren WF, Torda AE (1995) Parametrisation of time-averaged distance restraints in MD simulations. J Biomol NMR 6:313–320

Nanzer AP, Torda AE, Bisang C, Weber C, Robinson JA, van Gunsteren WF (1997) Dynamical studies of peptide motifs in the plasmodium falciparum circumsporozoite surface protein by restrained and unrestrained MD simulations. J Mol Biol 267:1012–1025

Olsson S, Frellsen J, Boomsma W, Mardia KV, Hamelryck T (2013) Inference of structure ensembles of flexible biomolecules from sparse, averaged data. PLoS One 8:e79439

Olsson S, Vögeli BR, Cavalli A, Boomsma W, Ferkinghoff-Borg J, Lindorff-Larsen K, Hamelryck T (2014) Probabilistic determination of native state ensembles of proteins. J Chem Theory Comput 10:3484–3491

Palmer AG III, Williams J, McDermott A (1996) Nuclear magnetic resonance studies of biopolymer dynamics. J Phys Chem 100:13,293–13,310

Pearlman DA (1994a) How is an NMR structure best defined? An analysis of molecular dynamics distance based approaches. J Biomol NMR 4:1–16

Pearlman DA (1994b) How well do time-averaged J-coupling restraints work? J Biomol NMR 4:279–299

Pearlman DA, Kollman PA (1991) Are time-averaged restraints necessary for NMR refinement? A model study for DNA. J Mol Biol 220:457–479

Pepermans H, Tourwé D, van Binst G, Boelens R, Scheek RM, van Gunsteren WF, Kaptein R (1988) The combined use of NMR, distance geometry, and restrained molecular dynamics for the conformational study of a cyclic somatostatin analogue. Biopolymers 27:323–338

Peter C, Daura X, van Gunsteren WF (2001) Calculation of NMR-relaxation parameters for flexible molecules from molecular dynamics simulations. J Biomol NMR 20:297–310

Pfeiffer S, Fushman D, Cowburn D (2001) Simulated and NMR-derived backbone dynamics of a protein with significant flexibility: a comparison of spectral densities for the βARK1 PH domain. J Am Chem Soc 123:3021–3036

Pitera JW, Chodera JD (2012) On the use of experimental observations to bias simulated ensembles. J Chem Theory Comput 8:4335–3451

Raval A, Piana S, Eastwood MP, Dror RO, Shaw DE (2012) Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. Proteins 80:2071–2079

Richter B, Gsponer J, Várnai P, Salvatella X, Vendruscolo M (2007) The mumo (minimal under-restraining minimal over-restraining) method for the determination of native state ensembles of proteins. J Biomol NMR 37:117–135

Roux B, Weare J (2013) On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. J Chem Phys 138:084107

Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J Comput Phys 23:327–341

Sapienza PJ, Lee AL (2010) Using NMR to study fast dynamics in proteins: methods and applications. Curr Opin Pharmacol 10:723–730

Scheek RM, Torda AE, Kemmink J, van Gunsteren WF (1991) Structure determination by NMR: the modelling of NMR parameters as ensemble averages. In: Hoch JC, Poulsen FM, Redfield C (eds) Computational aspects of the study of biological macromolecules by nuclear magnetic resonance spectroscopy, NATO ASI Series A225. Plenum Press, New York, pp 209–217

Schiffer CA, van Gunsteren WF (1999) Accessibility and order of water sites in and around proteins: a crystallographic time-averaging study. Proteins 36:501–511

Schiffer CA, Gros P, van Gunsteren WF (1995) Time-averaging crystallographic refinement: possibilities and limitations using alpha-cyclodextrin as a test system. Acta Cryst D51:85–92

Schmid N, Allison JR, Dolenc J, Eichenberger AP, Kunz APE, van Gunsteren WF (2011a) Biomolecular structure refinement using the GROMOS simulation software. J Biomol NMR 51:265–281

Schmid N, Eichenberger A, Choutko A, Riniker S, Winger M, Mark AE, van Gunsteren WF (2011b) Definition and testing of the GROMOS force-field versions: 54A7 and 54B7. Eur Biophys J 40:843–856

Schmid N, Christ CD, Christen M, Eichenberger AP, van Gunsteren WF (2012) Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. Comput Phys Commun 183:890–903

Schmitz U, Kumar A, James TL (1992) Dynamic interpretation of NMR data: molecular dynamics with weighted time-averaged restraints and ensemble R-factor. J Am Chem Soc 114:10,654–10,656

Schmitz U, Ulyanov B, Kumar A, James TL (1993) Molecular dynamics with weighted time-averaged restraints for a DNA octamer: dynamic interpretation of NMR data. J Mol Biol 234:373–389

Scott WRP, Mark AE, van Gunsteren WF (1998) On using time-averaging restraints in molecular dynamics simulations. J Biomol NMR 12:501–508

Showalter SA, Brüschweiler R (2007) Validation of molecular dynamics simulations of biomolecules. J Chem Theory Comput 3:961–975

Smith LJ, Mark AE, Dobson CM, van Gunsteren WF (1995a) Comparison of MD simulations and NMR experiments for hen lysozyme: analysis of local fluctuations, cooperative motions and global changes. Biochemistry 34:10918–10931

Smith PE, van Schaik RC, Szyperski T, Wüthrich K, van Gunsteren WF (1995b) Internal mobility of the basic pancreatic trypsin inhibitor in solution: a comparison of NMR spin relaxation measurements and molecular dynamics. J Mol Biol 246:356–365

Stocker U, van Gunsteren WF (2000) Molecular dynamics simulation of hen egg white lysozyme: a test of the GROMOS96 force field against nuclear magnetic resonance data. Proteins Struct Funct Bioinf 40:145–153

Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 314:141–151

Sugita Y, Kitao A, Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. J Chem Phys 113:6042–6051

Tironi IG, Sperb R, Smith PE, van Gunsteren WF (1995) A generalized reaction field method for molecular dynamics simulations. J Chem Phys 102:5451–5459

Torda AE, Scheek RM, van Gunsteren WF (1989) Time-dependent distance restraints in molecular dynamics simulations. Chem Phys Lett 157:289–294

Torda AE, Brunne RM, Huber T, Kessler H, van Gunsteren WF (1993) Structure refinement using time-averaged J-coupling restraints. J Biomol NMR 3:55–66

Trbovic N, Kim B, Friesner RA, Palmer AG (2008) Structural analysis of protein dynamics by MD simulations and NMR spin-relaxation. Proteins Struct Funct Bioinf 71:684–694

van Gunsteren WF, Berendsen HJC (1988) A leap-frog algorithm for stochastic dynamics. Mol Sim 1:173–185

van Gunsteren WF, Berendsen HJC, Rullmann JAC (1981) Stochastic dynamics for molecules with constraints. Brownian dynamics of n-alkanes. Mol Phys 44:69–95

van Gunsteren WF, Brunne RM, Gros P, van Schaik RC, Schiffer CA, Torda AE (1994) Accounting for molecular mobility in structure determination based on nuclear magnetic resonance spectroscopic and x-ray diffraction data. In: James TL, Oppenheimer NJ (eds) Methods in enzymology: nuclear magnetic resonance, vol 239. Academic Press, New York, pp 619–654

van Gunsteren WF, Dolenc J, Mark AE (2008) Molecular simulation as an aid to experimentalists. Curr Opin Struct Biol 18:149–153

White AD, Voth GA (2014) Efficient and minimal method to bias molecular simulations with experimental data. J Chem Theory Comput 10:3023–3030

Wong V, Case DA (2008) Evaluating rotational diffusion from protein MD simulation. J Phys Chem B 112:6013–6024

Yun-Yu S, Lu W, van Gunsteren WF (1988) On the approximation of solvent effects on the conformation and dynamics of cyclosporin A by stochastic dynamics simulation techniques. Mol Sim 1:369–383