Resubmitted to
Journal of Biomolecular NMR
(Date: August 30, 2016)

On the Use of Time-Averaging Restraints when Deriving Biomolecular Structure from ^3J-coupling Values Obtained from NMR Experiments

Lorna J. Smith[1], Wilfred F. van Gunsteren[2], Niels Hansen[3]*

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

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

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

*To whom correspondence should be addressed.
E-mail: hansen@itt.uni-stuttgart.de
Phone: +49 711 685 66112
Fax: +49 711 685 66140

Abstract

Deriving molecular structure from ^3J-couplings obtained from NMR experiments is a challenge due to (i) the uncertainty in the Karplus relation $^3J(\theta)$ connecting a ^3J-coupling value to a torsional angle θ, (ii) the need to account for the averaging inherent to the measurement of ^3J-couplings, and (iii) the sampling road blocks that may emerge due to the multiple-valuedness of the inverse function $\theta(^3J)$ of the function $^3J(\theta)$. Ways to properly handle these issues in structure refinement of biomolecules are discussed and illustrated using the protein hen egg white lysozyme as example.

Keywords:

Structure Refinement, Nuclear Magnetic Resonance, Force Field, Conformational Dynamics, Statistical Mechanics

1 Introduction

Structure determination of biomolecules at the atomic level of resolution has become possible through the availability of a variety of experimental measurement techniques such as X-ray or neutron diffraction and NMR and other spectroscopies in combination with the development of molecular models that represent the interactions between the atoms of molecules (Drenth, 1994; Wüthrich, 1986). The standard way to derive structure from measured data is to represent the latter by a biasing or restraining function $V^{Q,\mathrm{restr}}\left(Q(\vec{r}^N);Q^0\right)$ that limits the deviation of the value of an observable quantity Q that depends on the Cartesian coordinates $\vec{r}^N \equiv (\vec{r}_1, \vec{r}_2, ..., \vec{r}_N)$ of the system consisting of N atoms or particles, from the measured target value $Q^0 = Q^{\mathrm{exp}}$ of the quantity Q. This function $V^{Q,\mathrm{restr}}$ is added to the physical interaction function $V^{\mathrm{phys}}(\vec{r}^N)$ used to describe the system and

$$V(\vec{r}^N) = V^{\mathrm{phys}}(\vec{r}^N) + V^{Q,\mathrm{restr}}\left(Q(\vec{r}^N);Q^0\right) \tag{1}$$

is minimised or used in a simulation generating a trajectory or ensemble of structures.

However, the great majority of measured values Q^{exp} of observables Q corresponds to averages of Q over both space or molecules and time

$$Q^{\mathrm{exp}} = \langle\langle Q\rangle_{\mathrm{space}}\rangle_{\mathrm{time}} \tag{2}$$

where the angular brackets denote averaging over the distribution $P\left(\vec{r}^N\right)$ of configurations of the system, e.g. the Boltzmann distribution,

$$P(\vec{r}^N) = \frac{\exp\left(-V^{\mathrm{phys}}(\vec{r}^N)/(k_{\mathrm{B}}T)\right)}{\int \exp\left(-V^{\mathrm{phys}}(\vec{r}^N)/(k_{\mathrm{B}}T)\right)\mathrm{d}\vec{r}^N} \tag{3}$$

where k_{B} denotes the Boltzmann constant and T the temperature. This means that the averaging underlying the Q^{exp} values should be accounted for in the structure determination based on Equation 1.

Time-averaged restraining was introduced (Torda et al, 1989) using NOE derived atom-atom (i,j) distance, $r_{ij} = |\vec{r}_i - \vec{r}_j|$, information with $Q = r_{ij}^{-3}$ or $Q = r_{ij}^{-6}$. It was followed by use of time averages in crystallographic structure factor amplitude restraining (Gros et al, 1990), in chemical shift restraining (Harvey and van Gunsteren, 1993), ^3J-coupling restraining (Torda et al, 1993), and S^2 order-parameter restraining (Hansen et al, 2014). The time-averaging restraining method is characterised by two parameters, the force constant or weight K^{Q_r} of the restraining function $V^{Q,\mathrm{restr}}$ and the memory relaxation time τ_Q representing the time span over which $Q\left(\vec{r}^N(t)\right)$ is to be averaged. The exponentially damped time-averaged value of Q,

$$\left\langle Q(\vec{r}^N)\right\rangle_t = \left[\tau_Q\left(1-\exp\left(-t/\tau_Q\right)\right)\right]^{-1}\cdot\int_0^t \exp[-\left(t-t'\right)/\tau_Q]Q\left(\vec{r}^N(t')\right)\mathrm{d}t' \qquad (4)$$

is used in the restraining function $V^{Q,\mathrm{rest}}\left(Q;Q^0\right)$ instead of the instantaneous value $Q\left(\vec{r}^N(t)\right)$. The considerations leading to the introduction of the exponential memory damping factor into the time average have been given in Torda et al (1989). Time-averaging restraining has been investigated as function of K^{Q_r} and τ_Q for a variety of systems (Torda et al, 1990; Pearlman and Kollman, 1991; Schmitz et al, 1992, 1993; Pearlman, 1994a,b; Nanzer et al, 1994, 1995, 1997). The force constant K^{Q_r} should be taken as small as possible in order to avoid a restraining bias that puts strain into the molecule, while being large enough to force $\langle Q\rangle_{\mathrm{time}}$ close to Q^0. In addition, too strong restraints may destroy the proper Boltzmann weighting, Equation 3, of configurations of the trajectory. The memory relaxation time τ_Q should be of the order of magnitude 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 $Q\left(\vec{r}^N(t)\right)$ over t. In addition, the heating of the system due to the non-conservative force resulting from the time-averaging restraining potential energy term should be kept small.

Since most structures of proteins deposited in the protein data bank were determined using single-structure refinement, i.e. not accounting for averaging, one may wonder whether accounting for averaging is important. If the configurational ensemble of a molecule is not a very narrow one centered on a single most dominant structure, the average $\langle Q \rangle_{\vec{r}^N}$ over a Boltzmann distribution, Equation 3, of structures $\vec{r}^N$, will not be equal to the value of Q calculated for a single structure $\vec{r}^N$

$$\langle Q \rangle_{\vec{r}^N} \neq Q(\vec{r}^N) \tag{5}$$

and it will also not be equal to the value of Q calculated for a mean structure $\langle \vec{r}^N \rangle$,

$$\langle Q \rangle_{\vec{r}^N} \neq Q\left(\langle \vec{r}^N \rangle\right). \tag{6}$$

This due to the non-linearity of the Boltzmann weighting in Equation 3 and the non-linearity of the function $Q(\vec{r}^N)$ with respect to the configurations $\vec{r}^N$. As a consequence, single-structure refinement, i.e. not accounting for averaging, may lead to highly distorted structures (Daura et al, 1999) and artificially restricts the structural variability of the molecules (Torda et al, 1990; Schiffer et al, 1994, 1995; Schiffer and van Gunsteren, 1999).

The goal of adding a biasing or restraining function $V^{Q,\mathrm{restr}}\left(Q(\vec{r}^N); Q^0\right)$ to the physical interaction function $V^{\mathrm{phys}}(\vec{r}^N)$ representing the molecular model is to keep the value of the quantity $Q\left(\vec{r}^N\right)$ close to the target value Q^0. The functional form of $V^{Q,\mathrm{restr}}\left(Q; Q^0\right)$ should meet the following conditions, Figure 1, (van Gunsteren et al, 1985):

1. It should be a continuous function of Q with a continuous derivative in order to obtain a continuous energy and force in an MD simulation.

2. It should have two or three ranges that display a different behaviour of the restraining force as function of Q:

(a) a flat-bottom region of zero restraining energy and zero restraining force around Q^0, e.g. for $[Q^0 - \Delta Q^{\mathrm{fb}}, Q^0 + \Delta Q^{\mathrm{fb}}]$, representing the uncertainty in the Q^0 value or its relation $Q(\vec{r}^N)$ to structure;

(b) a region in which the restraining energy and force increase with increasing deviation of Q from Q^0, e.g. $[Q^0 - \Delta Q^{\mathrm{fb}} - \Delta Q^{\mathrm{h}}, Q^0 - \Delta Q^{\mathrm{fb}}]$ and $[Q^0 + \Delta Q^{\mathrm{fb}}, Q^0 + \Delta Q^{\mathrm{fb}} + \Delta Q^{\mathrm{h}}]$;

(c) a region in which the restraining force reaches a constant maximum in order to avoid too large restraining forces distorting the molecular structure if the deviation of Q from Q^0 becomes large, e.g. for $[-\infty, Q^0 - \Delta Q^{\mathrm{fb}} - \Delta Q^{\mathrm{h}}]$ and $[Q^0 + \Delta Q^{\mathrm{fb}} + \Delta Q^{\mathrm{h}}, \infty]$.

3. The restraining can be only attractive, i.e. apply when $Q > Q^0$, e.g. as in the case of NOE atom-atom distance upper-bound restraining (van Gunsteren et al, 1984; Kaptein et al, 1985), only repulsive, i.e. apply when $Q < Q^0$, e.g. when representing the absence of NOE intensity in a spectrum (de Vlieg et al, 1986) or both, e.g. as in the case of ^3J-coupling restraining (Torda et al, 1993).

The simplest representation of such a restraining function is a flat-bottom, then harmonic, then linear function. For an attractive restraint ($Q > Q^0$) it reads

$$\begin{aligned} V_{\mathrm{att}}^{Q,\mathrm{restr}} &\left(Q; K^{Q_{\mathrm{r}}}, Q^0, \Delta Q^{\mathrm{fb}}, \Delta Q^{\mathrm{h}}\right) = \\ &= \frac{1}{2} K^{Q_{\mathrm{r}}} \left(Q - Q^0 - \Delta Q^{\mathrm{fb}}\right)^2 \cdot H\left(Q; Q^0 + \Delta Q^{\mathrm{fb}}\right) \\ &\cdot \left(1 - H\left(Q; Q^0 + \Delta Q^{\mathrm{fb}} + \Delta Q^{\mathrm{h}}\right)\right) \\ &+ K^{Q_{\mathrm{r}}} \left(Q - Q^0 - \Delta Q^{\mathrm{fb}} - \frac{1}{2}\Delta Q^{\mathrm{h}}\right) \Delta Q^{\mathrm{h}} \\ &\cdot H\left(Q; Q^0 + \Delta Q^{\mathrm{fb}} + \Delta Q^{\mathrm{h}}\right) \end{aligned} \tag{7}$$

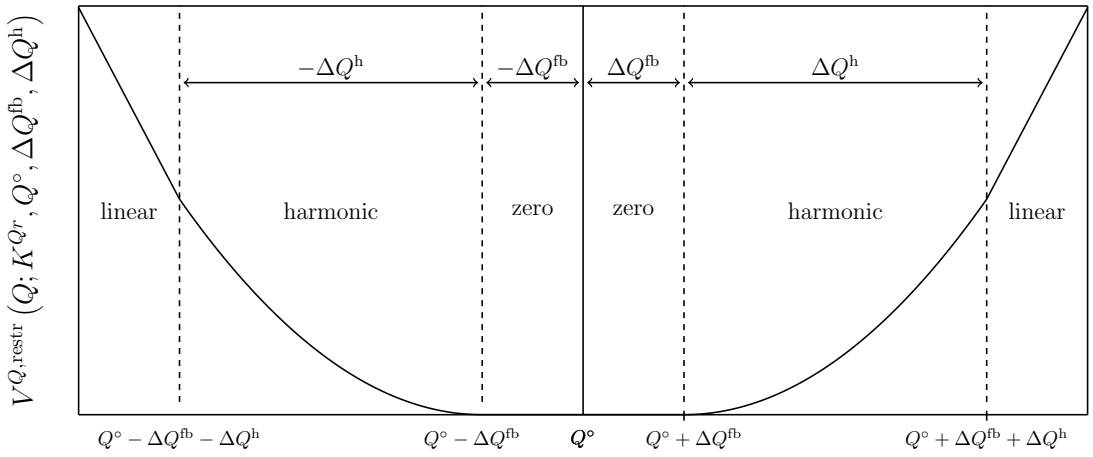

Figure 1: Potential energy terms $V^{Q,\mathrm{restr}}\left(Q; K^{Q_\mathrm{r}}, Q^0, \Delta Q^{\mathrm{fb}}, \Delta Q^{\mathrm{h}}\right)$ for restraining the quantity Q (Equation 7 and Equation 8).

and for a repulsive restraint ($Q < Q^0$) it reads

$$\begin{aligned} V_{\mathrm{rep}}^{Q,\mathrm{restr}}&\left(Q; K^{Q_\mathrm{r}}, Q^0, \Delta Q^{\mathrm{fb}}, \Delta Q^{\mathrm{h}}\right) = \\ &= \frac{1}{2} K^{Q_\mathrm{r}} \left(Q - Q^0 + \Delta Q^{\mathrm{fb}}\right)^2 \cdot \left(1 - H\left(Q; Q^0 - \Delta Q^{\mathrm{fb}}\right)\right) \\ &\cdot H\left(Q; Q^0 - \Delta Q^{\mathrm{fb}} - \Delta Q^{\mathrm{h}}\right) \\ &- K^{Q_\mathrm{r}} \left(Q - Q^0 + \Delta Q^{\mathrm{fb}} + \frac{1}{2}\Delta Q^{\mathrm{h}}\right) \Delta Q^{\mathrm{h}} \\ &\cdot \left(1 - H\left(Q; Q^0 - \Delta Q^{\mathrm{fb}} - \Delta Q^{\mathrm{h}}\right)\right), \end{aligned} \tag{8}$$

where $\Delta Q^{\mathrm{fb}} \geq 0$ and $\Delta Q^{\mathrm{h}} \geq 0$, and the Heaviside step function $H(x; x_0)$ is defined by

$$\begin{aligned} H(x; x_0) &= 0 \qquad \text{for} \quad x < x_0 \\ &= 1 \qquad \text{for} \quad x \geq x_0 \end{aligned} \tag{9}$$

The corresponding force along Q, i.e the negative of the derivative of $V^{Q,\mathrm{restr}}$

with respect to Q is then for the attractive case ($Q > Q^0$),

$$\begin{aligned} f_{\text{att}}^{Q,\text{restr}} &= 0 && \text{for } Q < Q^0 + \Delta Q^{\text{fb}} \\ &= -K^{Q_\text{r}} \left(Q - Q^0 - \Delta Q^{\text{fb}}\right) && \text{for } Q + \Delta Q^{\text{fb}} \leq Q \leq Q^0 + \Delta Q^{\text{fb}} + \Delta Q^{\text{h}} \\ &= -K^{Q_\text{r}} \Delta Q^{\text{h}} && \text{for } Q > Q^0 + \Delta Q^{\text{fb}} + \Delta Q^{\text{h}} \end{aligned} \tag{10}$$

and for the repulsive case ($Q < Q^0$)

$$\begin{aligned} f_{\text{rep}}^{Q,\text{restr}} &= 0 && \text{for } Q > Q^0 - \Delta Q^{\text{fb}} \\ &= -K^{Q_\text{r}} \left(Q - Q^0 + \Delta Q^{\text{fb}}\right) && \text{for } Q - \Delta Q^{\text{fb}} - \Delta Q^{h} \leq Q \leq Q^0 - \Delta Q^{\text{fb}} \\ &= +K^{Q_\text{r}} \Delta Q^{\text{h}} && \text{for } Q < Q^0 - \Delta Q^{\text{fb}} - \Delta Q^{\text{h}} \end{aligned} \tag{11}$$

When applied to ^3J-coupling restraining, the use of the time-average $\langle Q \rangle_t$, Equation 4, in the quadratic part of the restraining function $V^{Q,\text{restr}}$ (Equation 7 and Equation 8) led to large structural fluctuations (Nanzer et al, 1997). This is due to the fact that the average $\langle Q \rangle_t$ of Q is lagging the instantaneous value $Q(t)$ of Q. If the restraining force only depends on $\langle Q \rangle_t$, it may drive $Q(t)$ away from Q^0 in case $\langle Q \rangle_t$ and $Q(t)$ are at different sides of Q^0. This effect was not observed in time-averaging NOE distance bound restraining because (i) only attractive distance restraints are used, (ii) the r^{-3} or r^{-6} distance dependence of the NOE signal gives most weight to short distances when averaging, thereby reducing the difference between $Q(t)$ and $\langle Q \rangle_t$ for short distances r, and (iii) the van-der-Waals repulsion between atoms in V^{phys} counteracts too small atom-atom distances r.

This problem can be solved by using a biquadratic restraining function $V^{Q,\text{restr}}$ that depends on both $Q(t)$ and $\langle Q \rangle_t$ (Scott et al, 1998), for an attractive ($Q(t) >$

Q^0 and $\langle Q\rangle_t > Q^0$) restraint,

$$
\begin{aligned}
V_{\rm att}^{Q,{\rm restr}}&\left(Q(t),\langle Q\rangle_t;K^{Q_{\rm r}},Q^0,\Delta Q^{\rm fb}\right)=\\
&=\frac{1}{2}K^{Q_{\rm r}}\left(Q(t)-Q^0-\Delta Q^{\rm fb}\right)^2\cdot H\left(Q(t);Q^0+\Delta Q^{\rm fb}\right)\\
&\cdot\left(\langle Q\rangle_t-Q^0-\Delta Q^{\rm fb}\right)^2\cdot H\left(\langle Q\rangle_t;Q^0+\Delta Q^{\rm fb}\right),
\end{aligned}
\tag{12}
$$

and for a repulsive ($Q(t) < Q^0$ and $\langle Q\rangle_t < Q^0$) restraint,

$$
\begin{aligned}
V_{\rm rep}^{Q,{\rm restr}}&\left(Q(t),\langle Q\rangle_t;K^{Q_{\rm r}},Q^0,\Delta Q^{\rm fb}\right)=\\
&=\frac{1}{2}K^{Q_{\rm r}}\left(Q(t)-Q^0+\Delta Q^{\rm fb}\right)^2\cdot\left(1-H\left(Q(t);Q^0-\Delta Q^{\rm fb}\right)\right)\\
&\cdot\left(\langle Q\rangle_t-Q^0+\Delta Q^{\rm fb}\right)^2\cdot\left(1-H\left(\langle Q\rangle_t;Q^0-\Delta Q^{\rm fb}\right)\right).
\end{aligned}
\tag{13}
$$

The corresponding force along Q, i.e. the negative of the derivative of $V^{Q,{\rm restr}}$ with respect to Q, is then for the attractive restraint ($Q(t) > Q^0$ and $\langle Q\rangle_t > Q^0$),

$$
\begin{aligned}
f_{\rm att}^{Q,{\rm restr}}(t)&=0\qquad\text{for}\;\; Q(t)<Q^0+\Delta Q^{\rm fb}\;\text{or}\;\langle Q\rangle_t<Q^0+\Delta Q^{\rm fb}\\
&=-K^{Q_{\rm r}}\Big\{\left(Q(t)-Q^0-\Delta Q^{\rm fb}\right)\left(\langle Q\rangle_t-Q^0-\Delta Q^{\rm fb}\right)^2\\
&+\left(Q(t)-Q^0-\Delta Q^{\rm fb}\right)^2\left(\langle Q\rangle_t-Q^0-\Delta Q^{\rm fb}\right)\cdot\left(1-\exp(-\Delta t/\tau_Q)\right)\Big\}\\
&\qquad\text{for}\;\; Q(t)>Q^0+\Delta Q^{\rm fb}\;\text{and}\;\langle Q\rangle_t>Q^0+\Delta Q^{\rm fb}
\end{aligned}
\tag{14}
$$

and for the repulsive restraint ($Q(t) < Q^0$ and $\langle Q\rangle_t < Q^0$),

$$
\begin{aligned}
f_{\rm rep}^{Q,{\rm restr}}(t)&=0\qquad\text{for}\;\; Q(t)>Q^0-\Delta Q^{\rm fb}\;\text{or}\;\langle Q\rangle_t>Q^0-\Delta Q^{\rm fb}\\
&=-K^{Q_{\rm r}}\Big\{\left(Q(t)-Q^0+\Delta Q^{\rm fb}\right)\left(\langle Q\rangle_t-Q^0+\Delta Q^{b}\right)^2\\
&+\left(Q(t)-Q^0+\Delta Q^{\rm fb}\right)^2\left(\langle Q\rangle_t-Q^0+\Delta Q^{\rm fb}\right)\cdot\left(1-\exp(-\Delta t/\tau_Q)\right)\Big\}\\
&\qquad\text{for}\;\; Q(t)<Q^0-\Delta Q^{\rm fb}\;\text{and}\;\langle Q\rangle_t<Q^0-\Delta Q^{\rm fb}
\end{aligned}
\tag{15}
$$

where the derivative of the average $\langle Q \rangle_t$, Equation 4, with respect to $Q(t)$ has been expressed in terms of discrete MD time steps Δt,

$$\frac{\partial \langle Q \rangle_t}{\partial Q(t)} = (1 - \exp(-\Delta t/\tau_Q)) \,. \tag{16}$$

To obtain the force on particle i, these expressions are to be multiplied by $\frac{\partial Q(\vec{r}^N(t))}{\partial \vec{r}_i(t)}$. Using this biquadratic form, the restraining function only generates a force when both the instantaneous value $Q(t)$ of Q and the time-averaged value $\langle Q \rangle_t$ of Q lie both outside the flat bottom of the restraining function $V^{Q,\mathrm{restr}}$. If both, attractive and repulsive restraints are applied as in the case of ^3J-coupling restraining, the restraining is harmonic with a flat bottom. If $Q(t)$ and $\langle Q \rangle_t$ both lie outside the flat bottom and on the same side of Q^0, the two terms in Equation 14 and Equation 15, originating from the derivatives of $Q(t)$ and $\langle Q \rangle_t$ yield contributions to the force $f^{Q,\mathrm{restr}}$ that drive $Q(t)$ towards Q^0. If $Q(t)$ and $\langle Q \rangle_t$ lie on different sides of Q^0, i.e. $Q(t) < Q^0 - \Delta Q^{\mathrm{fb}}$ and $\langle Q \rangle_t > Q^0 + \Delta Q^{\mathrm{fb}}$ or $Q(t) > Q^0 + \Delta Q^{\mathrm{fb}}$ and $\langle Q \rangle_t < Q^0 - \Delta Q^{\mathrm{fb}}$, no restraining force is generated.

Eqs. 14 and 15 show that the relative weight of the two terms contributing to the biquadratic restraining force depends on a factor (Scott et al, 1998)

$$X = (1 - \exp(-\Delta t/\tau_Q)) \,. \tag{17}$$

When applying instantaneous restraining this factor is absent from the expression of the restraining force, and when applying non-biquadratic, i.e. standard quadratic time-averaging restraining this factor appears as an overall scaling factor in the expression for the restraining force. Since generally one has $\Delta t \ll \tau_Q$, one may approximate this scaling factor by

$$X = \Delta t/\tau_Q. \tag{18}$$

Eqs. 17 and 18 imply that the restraining force in standard quadratic time averaging restraining is inversely proportional to the exponential memory relaxation time τ_Q (for $\Delta t \ll \tau_Q$). In order to obtain comparable restraining forces for different values of τ_Q, one would have to choose the values of the restraining force constant K^{Q_r} proportional to $(1 - \exp(-\Delta t/\tau_Q))^{-1}$. Since the choice of the value for K^{Q_r} is empirical anyway, the factor X was set to 1 in previous work (Torda et al, 1993; Nanzer et al, 1995; Scott et al, 1998) and implementations (van Gunsteren et al, 1996; Christen et al, 2005; Schmid et al, 2011a) of the GROMOS software. In the current implementation (van Gunsteren, 2016) one may choose between $X = 1$, $X = (1 - \exp(-\Delta t/\tau_Q))$ and $X = 0$. The results presented here are obtained with the choice $X = (1 - \exp(-\Delta t/\tau_Q))$, Eq. 16.

The Karplus relation (Karplus, 1959)

$$^3J(\theta) = a\cos^2(\theta) + b\cos(\theta) + c \tag{19}$$

relates a ^3J-coupling to a torsional angle θ, see Figure 2. There are up to four values of the angle θ that may correspond to a single ^3J-coupling value. This means that the inverse of the function $^3J(\theta)$, the function $\theta(^3J)$, is multiple-valued. Depending on the initial value of the angle θ, the restraining function $V^{Q,\mathrm{restr}}$ may drive the angle θ in different directions to one of the various angle values that correspond to $Q^0 = {}^3J^0$, the target value. This is illustrated in the upper panels of Figure 3. If the target $^3J^0$-value lies beyond one maximum or minimum of the function $^3J(\theta)$, but not beyond another maximum or minimum, the restraining function $V^{Q,\mathrm{restr}}$ may drive the angle θ in the wrong direction depending on the initial θ-value, as illustrated in the lower panels of Figure 3. For residues having two H_β-atoms with almost equal values of the $^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-coupling the restraining forces acting on the χ_1-angle may point in opposite directions such that the conformation may be locked in

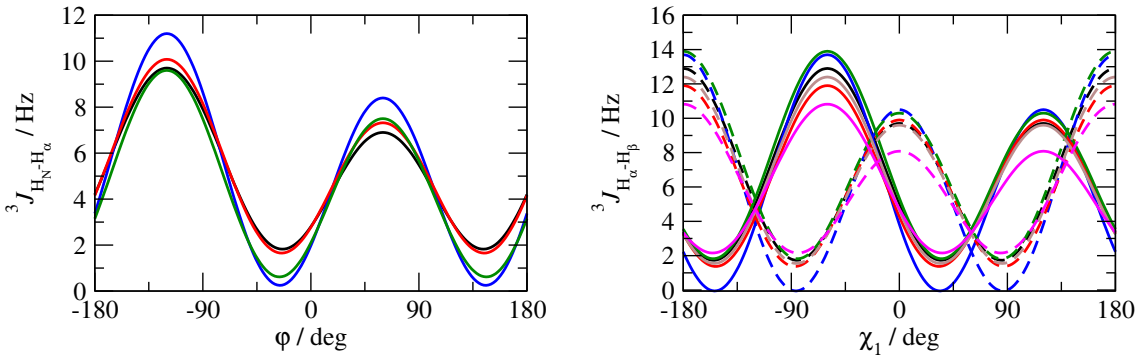

Figure 2: The Karplus curves describing the relation between ^3J-coupling and the torsional angles φ(C-N-C$_\alpha$-C) and χ_1, respectively. $\theta = \varphi + \delta$ for backbone $^3J_{\mathrm{H_N H_\alpha}}$ with $\delta = -60°$ (left panel) and $\theta = \chi_1 + \delta$ for side-chain $^3J_{\mathrm{H_\alpha H_\beta}}$ (right panel) with $\delta_{\beta_2} = -120°$ (full lines) and $\delta_{\beta_3} = 0°$ (dashed lines). The curves shown for the $^3J_{\mathrm{H_N H_\alpha}}$ couplings are based on the calibration constants reported by Pardi et al (1984) (black), Brüschweiler and Case (1994) (blue), Wang and Bax (1996) (red) and Schmidt et al (1999) (green). The curves shown for the $^3J_{\mathrm{H_\alpha H_\beta}}$ couplings are based on the calibration constants reported by deMarco et al (1978) (black), Abraham and McLauchlan (1968) (blue), Deber et al (1971) (red), Cung and Marraud (1982) (green), Fischman et al (1980) (brown), and Pérez et al (2001) (magenta).

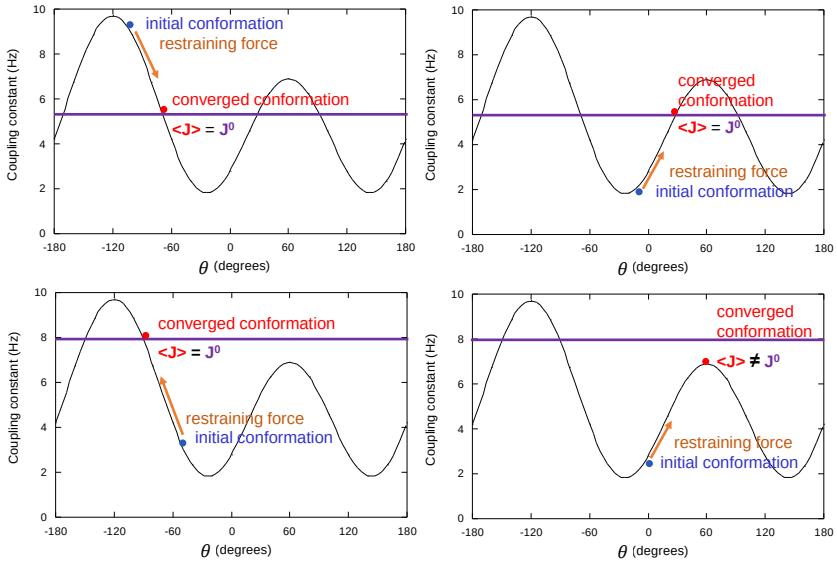

Figure 3: Consequences of a multiple-valued function $\theta(^3J)$ for restraining an angle θ to a particular ^3J-coupling value J^0 when starting from different initial values of θ in structure refinement.

the wrong intersection of the two Karplus curves, see right panel of Figure 2. Thus if the function $\vec{r}^N(Q)$ is multiple-valued, the restraining techniques discussed so far may not lead to structures that match the target Q^0 values. This problem can be solved in different ways:

1. The restraining function, and the corresponding torsional-angle dependent potential energy term are multiplied by a factor $\cos^2(\omega^{Q_r} t)$, which introduces oscillations in the forces due to these terms, thereby allowing the torsional angle to escape from local minima (Keller et al, 2007).

2. Hamiltonian replica-exchange MD simulation in which the restraining term $V^{Q,\mathrm{restr}}$ is made dependent on a coupling parameter λ such that for $\lambda = 0$ the restraining term is fully active, while for $\lambda = 1$ $V^{Q,\mathrm{restr}} = 0$, is applied. This allows escape from local minima through replicas with weak or zero restraining

forces (Schmid et al, 2011a).

3. A more straightforward way to avoid getting stuck in a local energy minimum of structures that do not match the experimental data during structure refinement is the use of local-elevation searching and sampling (Huber et al, 1994) in which the potential energy at already visited parts of configuration space is raised in order to avoid repetitive sampling of the same parts of configuration space in a simulation. The mathematical and algorithmic adaptation of local-elevation sampling for use in deriving structures from experimental data has been given for the case of ^3J-couplings serving as the quantity for which measured values are available (Christen et al, 2007).

The basic idea of local-elevation structure refinement based on adaptive restraints (Christen et al, 2007) is that whenever the simulated average value $\langle {}^3J \rangle_t$ of the ^3J-coupling and the current value $^3J(t)$ of the ^3J-coupling do not match the measured target value $^3J^0 = {}^3J^{\mathrm{exp}}$ to within an uncertainty Δ^3J^{fb}, the force constant or weight of a penalty or restraining function $V^{J,\mathrm{restr}}(\varphi(\vec{r}^N(t)))$, acting on the torsion angle φ (or χ_1 for a side-chain ^3J-coupling) corresponding to the ^3J-coupling and its current value $\varphi(t)$, is increased. The restraining potential energy term of a given ^3J-coupling ^3J_k is a sum of N_{le} terms, local with respect to a particular value φ_i^0 of φ_k,

$$
\begin{aligned}
V_k^{J,\mathrm{restr}} &\left(\varphi_k(\vec{r}^N(t)); K^{J_\mathrm{r}}, N_{\mathrm{le}}, {}^3J_k^0, \Delta^3J^{\mathrm{fb}}\right) = \\
&= \sum_{i=1}^{N_{\mathrm{le}}} V_{ki}^{J,\mathrm{restr}} \left(\varphi_k(\vec{r}^N(t)); K^{J_\mathrm{r}}, N_{\mathrm{le}}, {}^3J_k^0, \Delta^3J^{\mathrm{fb}}\right) \\
&= \sum_{i=1}^{N_{\mathrm{le}}} K^{J_\mathrm{r}} \omega_{\varphi_{ki}}(t) \exp\left(-(\varphi_k(t) - \varphi_i^0)^2/(2(\Delta\varphi^0)^2)\right), \qquad (20)
\end{aligned}
$$

with the weight $\omega_{\varphi_{ki}}(t)$ of the i-th penalty function changing during the simulation according to

$$\omega_{\varphi_{ki}}(t) = t^{-1} \int_0^t \delta_{\varphi_k(\vec{r}^N(t'))\varphi_i^0} \cdot \left({}^3J_k(t') - {}^3J_k^0 - \Delta^3 J^{\mathrm{fb}}\right)^2 \cdot H\left({}^3J_k(t'); {}^3J_k^0 + \Delta^3 J^{\mathrm{fb}}\right) \cdot \left(\langle {}^3J_k\rangle_{t'} - {}^3J_k^0 - \Delta^3 J^{\mathrm{fb}}\right)^2 \cdot H\left(\langle {}^3J_k\rangle_{t'}; {}^3J_k^0 + \Delta^3 J^{\mathrm{fb}}\right) \mathrm{d}t' \tag{21}$$

for an attractive (${}^3J(t) > {}^3J^0$ and $\langle {}^3J_k\rangle_t > {}^3J_k^0$) ${}^3J$-coupling restraint, and according to

$$\omega_{\varphi_{ki}}(t) = t^{-1} \int_0^t \delta_{\varphi_k(\vec{r}^N(t'))\varphi_i^0} \cdot \left({}^3J_k(t') - {}^3J_k^0 + \Delta^3 J^{\mathrm{fb}}\right)^2 \cdot \left(1 - H\left({}^3J_k(t'); {}^3J_k^0 - \Delta^3 J^{\mathrm{fb}}\right)\right) \cdot \left(\langle {}^3J_k\rangle_{t'} - {}^3J_k^0 + \Delta^3 J^{\mathrm{fb}}\right)^2 \cdot \left(1 - H\left(\langle {}^3J_k\rangle_{t'}; {}^3J_k^0 - \Delta^3 J^{\mathrm{fb}}\right)\right) \mathrm{d}t' \tag{22}$$

for a repulsive (${}^3J(t) < {}^3J^0$ and $\langle {}^3J_k\rangle_t < {}^3J_k^0$) ${}^3J$-coupling restraint. Compared to Equation 7 and Equation 8, the parameter ΔQ^{h} is omitted, or taken equal to infinity, because a ${}^3J$-coupling has a limited range of values, and thus no excessive restraining forces are generated if the restraining function is harmonic with a flat bottom. The centers φ_i^0 of the Gaussian functions in Equation 20 are equally distributed over the range of possible values of φ_k,

$$\varphi_i^0 = 2\pi i N_{\mathrm{le}}^{-1}, \qquad i = 1, 2, ..., N_{\mathrm{le}} \tag{23}$$

and their width is determined by

$$\Delta\varphi^0 = 2\pi N_{\mathrm{le}}^{-1}. \tag{24}$$

The overall weight or force constant of the penalty function is indicated by $K^{J_{\mathrm{r}}}$ and the Kronecker delta δ is defined using finite differences,

$$\begin{aligned} \delta_{\varphi_k(t)\varphi_i^0} &= 1 \qquad \text{if } \ (\varphi_i^0 - \Delta\varphi^0/2) \le \varphi_k(t) < (\varphi_i^0 + \Delta\varphi^0/2) \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{25}$$

Equation 21 and Equation 22 ensure that the torsional angle φ_k is pushed away from the value φ_i^0 when both the current value ${}^3J_k(t)$ and the averaged value $\langle {}^3J_k \rangle_t$ are deviating more than $\pm\Delta^3 J^{\mathrm{fb}}$ from the target value ${}^3J_k^0$.

The force on particle j due to the k-th penalty function, Equation 20, is then

$$\begin{aligned}
\vec{f}_j^{J,\mathrm{restr}}(t) &= -\frac{\partial V_k^{J,\mathrm{restr}}}{\partial \vec{r}_j} \\
&= -\frac{\partial V_k^{J,\mathrm{restr}}}{\partial \varphi_k} \cdot \frac{\partial \varphi_k}{\partial \vec{r}_j} \\
&= \sum_{i=1}^{N_{\mathrm{le}}} V_{ki}^{J,\mathrm{restr}} \left(\varphi_k(t); K^{J_{\mathrm{r}}}, N_{\mathrm{le}}, {}^3J_k^0, \Delta^3 J^{\mathrm{fb}} \right) \\
&\cdot \frac{(\varphi_k(t) - \varphi_i^0)}{(\Delta\varphi^0)^2} \cdot \frac{\partial \varphi_k(t)}{\partial \vec{r}_j(t)}.
\end{aligned} \tag{26}$$

Expressions for $\partial\varphi_k(t)/\partial\vec{r}_j(t)$ can for example be found in (van Schaik et al, 1993; van Gunsteren et al, 1996; van Gunsteren, 2016).

Here, the properties of the different ways to account for the averaging inherent to the measurement of values of observable quantities when deriving biomolecular structures from NMR ${}^3J$-coupling data are investigated using as test system the protein hen egg white lysozyme, for which measured ${}^3J$-coupling values for backbone and side chains are available.

2 Methods and Model

The goal is to show the difference in performance between three structure refinement methods when applied to ^3J-couplings of a protein: (i) instantaneous restraining (IR), i.e. single-structure refinement, (ii) biquadratic time-averaging restraining (BQ-TAR), which allows accounting for the averaging inherent to the measurement of ^3J-coupings, and (iii) biquadratic time-averaging local-elevation restraining (BQ-TA-LER), which in addition induces torsional-angle transitions, i.e. wider sampling than gradient-based methods, in case the target $^3J^0$ value is not matched by the average $\langle {}^3J \rangle$. The restraining strength $K^{Q,\mathrm{restr}}$ and the flat-bottom range ΔQ^{fb}, which accounts for the uncertainties in the target $^3J^0$ value and induced by the approximation involved in the Karplus relations $^3J(\varphi)$ and $^3J(\chi_1)$, are varied and for the methods involving time averaging, the memory relaxation time τ_Q is also varied.

The protein hen egg white lysozyme in aqueous solution is used as test system, because for this protein ample $^3J_{\mathrm{H_N H_\alpha}}$- and $^3J_{\mathrm{H_\alpha H_\beta}}$-coupling values are available, for the latter even stereospecifically assigned for side chains with two H_β atoms. A selection of backbone and side-chain ^3J-couplings was made to illustrate the performance of the three methods. Considering the shape of the Karplus curves $^3J(\varphi)$ and $^3J(\chi_1)$ in Figure 2, it is clear that intermediate target values $^3J^0$ can be reached for four angle values and through averaging of ^3J-couplings, whereas for low or high target values $^3J^0$ only two, or one or no angle values – depending on the particular Karplus relation used – match the $^3J^0$ values, and averaging may not help. Secondly the potential energy function or force field $V^{\mathrm{phys}}(\vec{r}^N)$ used to model the protein and its aqueous environment will determine the range of φ- and χ_1-angle values that have a relatively low energy, i.e. that are accessible at the given temperature and pressure. Based on these considerations the following ^3J-couplings were selected for analysis.

(a) 7 Backbone $^3J_{\mathrm{H_N H_\alpha}}$-couplings, see Tables S1 – S3, three, with a high (9.3 Hz, Thr 43), a low (3.7 Hz, Ala 9), and an intermediate (7.0 Hz, Asn 19) $^3J^0$ target value, that were reproduced in MD simulation without any ^3J-coupling restraining, three ^3J-couplings, with a high (9.3 Hz, Thr 69), a low (3.6 Hz, Lys 33), and an intermediate (7.4 Hz, Met 105) $^3J^0$ target value, that were not reproduced in MD simulation without any ^3J-coupling restraining, and one ^3J-coupling with a high value that lies outside the Karplus curve used (Pardi et al, 1984) (10.6 Hz, Ile 124).

(b) 5 Side-chain $^3J_{\mathrm{H_\alpha H_\beta}}$-couplings, see Tables S4 – S6, two for residues with one H_β atom, a large (10.8 Hz, Val 2) and a small (2.6 Hz, Thr 47) target $^3J^0$ value that were reproduced in unrestrained MD simulation, and three ^3J-couplings that were not reproduced in unrestrained MD simulation, a high value (9.5 Hz, Thr 89), and two intermediate (6.3 Hz, Val 99 and 8.0 Hz, Val 109) values for residues with a short side chain for which averaging may be important.

(c) 10 Side-chain $^3J_{\mathrm{H_\alpha H_\beta}}$-couplings, see Tables S7 – S9, for 5 residues with two H_β atoms, two residues with a large and a small value ($\mathrm{H}_{\beta_2} = 11.2\,\mathrm{Hz}, \mathrm{H}_{\beta_3} = 3.6\,\mathrm{Hz}$ for Asp 52 and $\mathrm{H}_{\beta_2} = 11.2\,\mathrm{Hz}, \mathrm{H}_{\beta_3} = 4.7\,\mathrm{Hz}$ for Asn 46), a residue with two rather small values ($\mathrm{H}_{\beta_2} = 5.1\,\mathrm{Hz}, \mathrm{H}_{\beta_3} = 4.5\,\mathrm{Hz}$ for Asp 66) and two residues with almost equal (not stereospecifically assigned) values (6.7 Hz, 6.4 Hz for Glu 7 and 6.9 Hz, 6.7 Hz for Arg 45).

The Karplus relation used for $^3J_{\mathrm{H_N H_\alpha}}$ is the one from Pardi et al (1984) with $a = 6.4\,\mathrm{Hz}$, $b = -1.4\,\mathrm{Hz}$ and $c = 1.9\,\mathrm{Hz}$. The relation between the torsional angle θ connecting the hydrogens and the φ-angle (N-C-C_α-C) is $\theta = \varphi + \delta$ with $\delta = -60°$. For $^3J_{\mathrm{H_\alpha H_\beta}}$ the Karplus relation from deMarco et al (1978) was used. The relation $\theta = \chi_1 + \delta$ for side chain $^3J_{\mathrm{H_\alpha H_\beta}}$-couplings was used with $\delta = -120°$ for Ile and Thr and with $\delta = 0°$ for Val. When two protons are present at the C_β-atom, we have

$\delta_{\beta_2} = -120°$ and $\delta_{\beta_3} = 0°$.

The protein was modelled using the GROMOS 54A7 biomolecular force field (Schmid et al, 2011b) and the water using the simple point charge (SPC) model (Berendsen et al, 1981) that is compatible with it.

The starting coordinates for the simulations were taken from the structures of hen egg-white lysozyme determined at pH 4.5 with Protein Data Bank code 1AKI (Artymiuk et al, 1982). In the simulations the Asp and Glu side chains in the protein were unprotonated and His 15 was singly protonated at Nε2. The protein was solvated in a rectangular box and minimum image periodic boundary conditions were applied. The minimum solute-box wall distance was set to 1.4 nm giving 14365 simple point charge (SPC) water molecules (Berendsen et al, 1981). Eight chloride ions were added to achieve overall neutrality of the system.

For each simulation an initial equilibration scheme comprising five 20 ps simulations at temperatures of 60 K, 120 K, 180 K, 240 K and 300 K was used. During the first 80 ps of this equilibration the solute atoms were harmonically restrained to their positions in the starting structure with force constants of 25000, 2500, 250 and 25 $\mathrm{kJ\,mol^{-1}nm^{-2}}$ at temperatures of 60 K, 120 K, 180 K and 240 K respectively. Following equilibration the simulations were run at 300 K. An unrestrained simulation was run for 10 ns. This was then branched into the different simulations with ^3J-coupling restraining, each of which was run for 2 ns. Separate simulations were performed applying the different restraining functions (IR, BQ-TAR, BQ-TAR-LE) to either 95 backbone $^3J_{\mathrm{H_N H_\alpha}}$ or to 62 side-chain $^3J_{\mathrm{H_\alpha H_\beta}}$-couplings taken from Smith et al (1991). Joint application of both data sets would have made it impossible to delineate cause / effect relationships in regard to backbone and side-chain torsional angles. The protein configurations and energies were saved for analysis every 2500 steps (5 ps).

All simulations were performed at a constant pressure of 1 atm, the temperature and pressure being maintained using the weak coupling algorithm (Berendsen et al, 1984), with relaxation times of $\tau_{\mathrm{T}} = 0.1\,\mathrm{ps}$ and $\tau_{\mathrm{p}} = 0.5\,\mathrm{ps}$ and an isothermal compressibility of $4.575 \times 10^{-4}\,(\mathrm{kJ\,mol^{-1}nm^{-3}})^{-1}$. Protein and solvent were separately coupled to the heat bath. The SHAKE algorithm (Ryckaert et al, 1977) was used to constrain bond lengths and the geometry of the water molecules, with a relative geometric tolerance of 10^{-4} allowing for an integration time step of 2 fs. In simulations using BQ-TAR for restraining side-chain ^3J-coupling constants a time step of 1 fs was used to avoid numerical instabilities in cases where the configuration was locked e.g. in the intersection of the two Karplus curves. The centre of mass motion was removed every 1000 time steps. Non-bonded interactions were calculated using a triple-range cutoff scheme with cutoff radii of 0.8 and 1.4 nm. Interactions within 0.8 nm were evaluated every time step and intermediate range interactions were updated every fifth time step. To account for the influence of the dielectric medium outside the cutoff sphere, a reaction-field force (Tironi et al, 1995) with a relative dielectric permittivity ε of 61 was used (Heinz et al, 2001). The width $\Delta\varphi^0$ of the Gaussian function in Equation 20 was set to 10° corresponding to $N_{\mathrm{le}} = 36$.

The trajectories were analysed in terms of average values and root-mean-square fluctuations of the ^3J-couplings and the corresponding φ- and χ_1- angles, the intra-protein potential energy reflecting the strain in the protein induced by the restraining and the restraining energy $V^{Q,\mathrm{restr}}$, representing the effort required to satisfy the target $^3J^0$ values. In view of the uncertainty of 1-2 Hz of the Karplus relation (see Figure 2), a deviation between $\langle {}^3J \rangle$ and $^3J^0$ smaller than 1 Hz is considered insignificant. Deviations larger than 2 Hz seem unsatisfactory.

Before calculating the averages of φ- and χ_1- angles these were mapped to values between -180° and $+180^\circ$. This may obscure the relation between the angle value and the corresponding ^3J-coupling constant in cases where the conformational

fluctuations lead to frequent crossings of the interval boundaries. This happend especially for the side chain restraining of residue Val 2 using IR and BQ-TA-LER. Therefore, unless noted otherwise, the average χ_1- angles reported for this residue in Tables 1, S4 and S6 are not mapped while for all other residues averages were obtained from mapped angle values while the root mean square fluctuations were calculated from the unmapped time series to emphasize the occurrence of torsional transitions.

3 Results

The numerical results of the various restraining procedures are summarized in detail in Tables S1 – S9 in the Supplementary Material. These data are visualised in Fig. 4 (IR), Fig. 5 (BQ-TAR) and Fig. 6 (BQ-TA-LER), respectively. In each of these figures each ^3J-coupling is represented by a different colour, while each line type represents a particular set of restraining parameters. Black lines display the intra-protein energy.

3.1 Backbone $^3J_{\mathrm{H_N H_\alpha}}$-couplings

Table S1 shows the results for the seven backbone $^3J_{\mathrm{H_N H_\alpha}}$-couplings using instantaneous restraining (IR). The target $^3J^0$ values of the first three residues are reproduced by the force field without any restraining ($K^{J_\mathrm{r}} = 0$), while for the next four residues deviations up to 3.9 Hz are observed. For the first three residues, restraining is not needed, it only may reduce the fluctuations of the ^3J-couplings and φ-angles. For the next three residues instantaneous restraining with a force constant of $K^{J_\mathrm{r}} = 10\,\mathrm{kJmol^{-1}Hz^{-2}}$ is sufficient to bring the averaged ^3J-couplings within 1 Hz from the target $^3J^0$ values be it at the price of a reduction of the conformational variability, i.e. the φ-angle fluctuations. Too strong restraining will increase the strain in the protein, as is illustrated in Figure 4 (upper panel) by an increase of the intra-protein energy upon increasing K^{J_r}, an artefact that is to be avoided (van Gunsteren, 1990).

The values obtained with $\Delta^3J^{\mathrm{fb}} = 0$, i.e. without a flat-bottom in the restraining function $V^{Q,\mathrm{restr}}$, were included in Table S1, because this corresponds to the shape of the restraining function commonly used in single-structure refinement. In view of the inaccuracy of the Karplus relation a value $\Delta^3J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$ should at least be used in order to avoid artificially reducing the conformational variability.

If a measured ${}^3J^0$-coupling results from averaging over different φ-angle values, accounting for averaging in the restraining may lead to a more faithful conformational representation of the measured data. Table S2 shows the results for the seven backbone ${}^3J_{\mathrm{H_N H_\alpha}}$-couplings using biquadratic time-averaging restraining (BQ-TAR). Since for backbone φ-angles the conformational variability is limited, only peptide-plane rotational fluctuations are allowed given a particular fold of the protein backbone, accounting for ${}^3J$-coupling averaging in restraining leads to only slightly enhanced fluctuations of ${}^3J$-couplings and φ-angles compared to applying instantaneous restraining. Figure 5 shows that the intra-protein strain is reduced when accounting for averaging.

As is illustrated in Figure 3, the BQ-TAR restraining function is driving the φ-angle along the (negative) gradient of the Karplus curve, thus will not be able to reach those ranges of φ-angle values that lie beyond barriers in $V^{Q,\mathrm{restr}}(\varphi(\vec{r}^N))$. Its sampling domain depends on the initial φ-angle value and does not comprise 360°. Allowing for sampling of all φ-angle values, if needed because of a too large deviation of $\langle {}^3J \rangle$ from ${}^3J^0$, as is done in biquadratic time-averaging local-elevation restraining (BQ-TA-LER), yields the results presented in Table S3. Using a relatively short averaging time of $\tau_Q = 5\,\mathrm{ps}$ and a $\Delta^3 J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$ flat bottom, a restraining force constant $K^{J_\mathrm{r}} = 5 \times 10^{-4}\,\mathrm{kJmol^{-1}Hz^{-4}}$ already suffices to get $\langle {}^3J \rangle$ close to ${}^3J^0$ for the 6 residues with ${}^3J^0_{\mathrm{H_N H_\alpha}}$ values within the range of the Karplus curve, while showing more conformational variability than in the IR and BQ-TAR cases. The results of BQ-TA-LER for Met 105 illustrate that there may exist more than one set of conformations that brings $\langle {}^3J \rangle$ close to ${}^3J^0$; one with $\langle \varphi \rangle \approx -80°$ and one with $\langle \varphi \rangle \approx +50°$.

3.2 Side-chain $^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-couplings (one H_β atom)

Threonine and Valine possess relatively short side chains which may allow a greater range of χ_1-angle values to be sampled than for backbone φ-angles. This is illustrated in Table S4, which shows larger χ_1-angle fluctuations than φ-angle fluctuations (Table S1) in unrestrained ($K^{J_\mathrm{r}} = 0$) MD simulations. As was observed for the backbone ^3J-couplings, instantaneous restraining artificially reduces the ^3J-coupling and corresponding χ_1-angle fluctuations in cases where the unrestrained MD simulations already reproduced the target $^3J^0$ values (Val 2, Thr 47). For residues Thr 89 and Val 99 only using the larger restraining force constant $K^{J_\mathrm{r}} = 10\,\mathrm{kJmol^{-1}Hz^{-2}}$ forced $\langle^3J\rangle$ close to $^3J^0$. For residue Val 109 the target value is reached for one particular parameter combination ($K^{J_\mathrm{r}} = 1\,\mathrm{kJmol^{-1}Hz^{-2}}$, $\Delta^3J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$) through averaging over two populated conformations, one with χ_1 around $-50°$ and one with χ_1 around $-175°$ leading to $\langle^3J\rangle$ of 8 Hz.

Allowing for ^3J-coupling averaging as in BQ-TAR simulation does not bring the $\langle^3J\rangle$ values closer to the $^3J^0$ target values, as is illustrated in Table S5. Yet the results improve significantly when local-elevation sampling is added to the time-averaging in BQ-TA-LER simlations (Table S6). The target $^3J^0$ values can be reached while maintaining a sizeable conformational variability. In particular for Val 99 and Val 109 averaging over a large range of χ_1-angle values leads to the measured $^3J^0$ values. An averaging time of only $\tau_Q = 5\,\mathrm{ps}$ combined with an $\Delta^3J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$ flat-bottom and a force constant of only $K^{J_\mathrm{r}} = 5 \times 10^{-4}\,\mathrm{kJmol^{-1}Hz^{-4}}$ reproduces the target $^3J^0$ values.

The power of the BQ-TA-LER method lies in its ability to let the torsional angle χ_1 escape from a minimum in the restraining function that does not correspond to the target $^3J^0$ value, see Figure 3. To show this ability of BQ-TA-LER more

clearly, one could give the χ_1-angle an initial value that lies outside the range of values for which the gradient of $V^{Q,\mathrm{restr}}$ would drive the χ_1-angle directly towards a value corresponding to ${}^3J^0$. However, such a situation can alternatively be generated by shifting the Karplus curve by 120° while keeping the initial χ_1-angle value unchanged. The results for two residues, Val 2 and Val 29, are shown in Table 1. Using the standard Karplus relation, the IR procedure can get $\langle {}^3J \rangle$ on target (within 1 Hz) for Val 2, but not for Val 29, while the BQ-TA-LER procedure manages both cases. Using a Karplus curve shifted by +120° the IR procedure does not get $\langle {}^3J \rangle$ on target, while the BQ-TA-LER procedure does.

3.3 Side-chain ${}^3J_{\mathrm{H}_\alpha\mathrm{H}_{\beta_2}}$- and ${}^3J_{\mathrm{H}_\alpha\mathrm{H}_{\beta_3}}$-couplings (two H_β atoms)

The results of the three restraining methods for the five selected residues with two H_β atoms are shown in Tables S7 – S9. The ${}^3J_{\mathrm{H}_\alpha\mathrm{H}_{\beta_{2/3}}}$-couplings for Asp 52 are reproduced by the force field used, so no restraining would be needed. Instantaneous restraining (IR) with a large force constant reduces the ${}^3J$-coupling and χ_1-angle fluctuations. This effect is less strong when averaging is accounted for (BQ-TAR) and is absent in the case of local-elevation sampling (BQ-TA-LER). For the other four residues, IR does not bring both $\langle {}^3J \rangle$-values within 1 Hz from the ${}^3J^0$-target values, except for Asp 66 with the larger force constant $K^{J_\mathrm{r}} = 10\,\mathrm{kJmol^{-1}Hz^{-2}}$. Accounting for averaging as in BQ-TAR (Table S8) does not change this picture much, although larger fluctuations are obtained. For Arg 45 with $\tau_Q = 5\,\mathrm{ps}$, $K^{J_\mathrm{r}} = 50\,\mathrm{kJmol^{-1}Hz^{-4}}$ and $\Delta^3J^{\mathrm{fb}} = 1.0\,\mathrm{Hz}$ even a few torsional angle transitions are observed, but these do not bring $\langle {}^3J_{\mathrm{H}_\alpha\mathrm{H}_{\beta_{2/3}}} \rangle$ to the target value. Using BQ-TA-LER simulation does improve the results significantly. For all five residues both target ${}^3J^0$ values are reproduced within 1 Hz except for Asn 46 with $\tau_Q = 50\,\mathrm{ps}$, $K^{J_\mathrm{r}} = 5 \times 10^{-4}\,\mathrm{kJmol^{-1}Hz^{-4}}$ and $\Delta^3J^{\mathrm{fb}} = 1.0\,\mathrm{Hz}$ where $\langle {}^3J_{\mathrm{H}_\alpha\mathrm{H}_{\beta_2}} \rangle$ is a bit too large (Table S9). For Asp 66 a larger range of χ_1-angle values is explored and for Glu 7 and Arg 45 extensive averaging

leads to the correct ${}^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$ values. It seems that an averaging time of 5 ps combined with the lower force constant of $5 \times 10^{-4}\,\mathrm{kJmol^{-1}Hz^{-4}}$ and a 0.5 Hz flat-bottom suffices to bring the average ${}^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-couplings on target while maintaining the conformational variability.

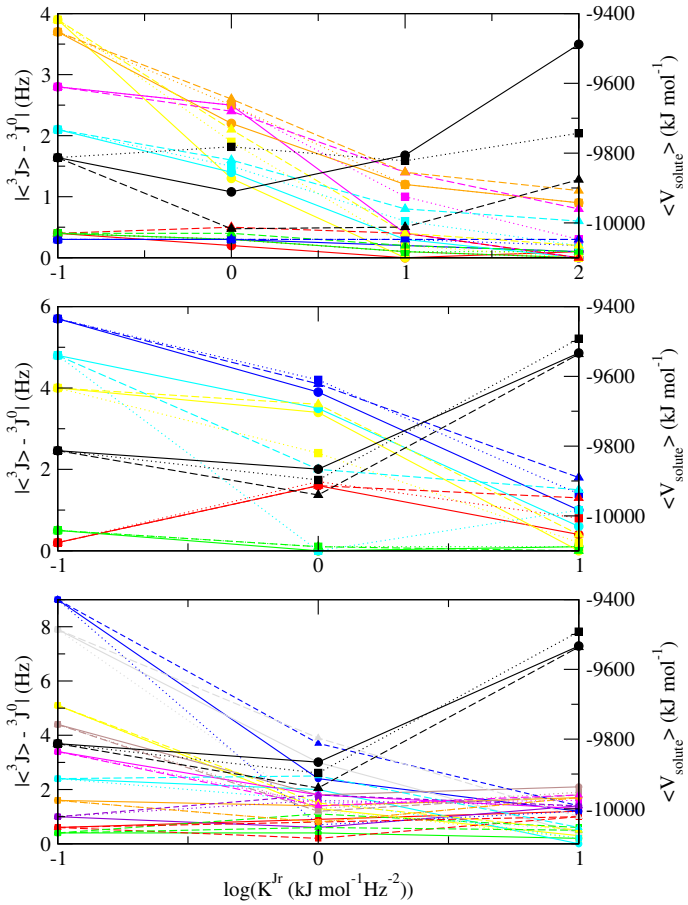

Figure 4: Size of the deviation of the average ^3J-coupling value $\langle {}^3J \rangle$ from the target value $^3J^0$ (left scale, coloured lines) and the intra-protein energy (bonded and non-bonded, right scale, black lines) from 2 ns MD simulations without ^3J-coupling restraining ($K^{J_r} = 0$, $\log K^{J_r} = -\infty$, but put at -1) or with instantaneous restraining (IR, $K^{J_r} > 0$), one with restraining 95 main chain $^3J_{\mathrm{H_N H_\alpha}}$-couplings and one with restraining 62 side chain $^3J_{\mathrm{H_\alpha H_\beta}}$-couplings as function of the restraining force constant $K^{J_r} = K^{Q,\mathrm{restr}}$ and for different values of the flat-bottom range $\Delta^3 J^{\mathrm{fb}} = \Delta Q^{\mathrm{fb}}$ (0 Hz: circles and solid lines; 0.5 Hz: squares and dotted lines; 1 Hz: triangles and dashed lines) of the restraining function. Upper panel: main chain $^3J_{\mathrm{H_N H_\alpha}}$ for 7 residues: Thr 43 (red), Ala 9 (green), Asn 19 (blue), Thr 69 (yellow), Lys 33 (light blue), Met 105 (magenta), Ile 124 (orange). Middle panel: side chain (one H_β) $^3J_{\mathrm{H_\alpha H_\beta}}$ for 5 residues: Val 2 (red), Thr 47 (green), Thr 89 (blue), Val 99 (yellow), Val 109 (light blue). Lower panel: side chain (two H_β) $^3J_{\mathrm{H_\alpha H_\beta}}$ for 5 residues: Asp 52 (H_{β_2}: red, H_{β_3}: green), Asn 46 (H_{β_2}: blue, H_{β_3}: yellow), Asp 66 (H_{β_2}: light blue, H_{β_3}: grey), Glu 7 (H_{β_2}: orange, H_{β_3}: violet), Arg 45 (H_{β_2}: brown, H_{β_3}: magenta).

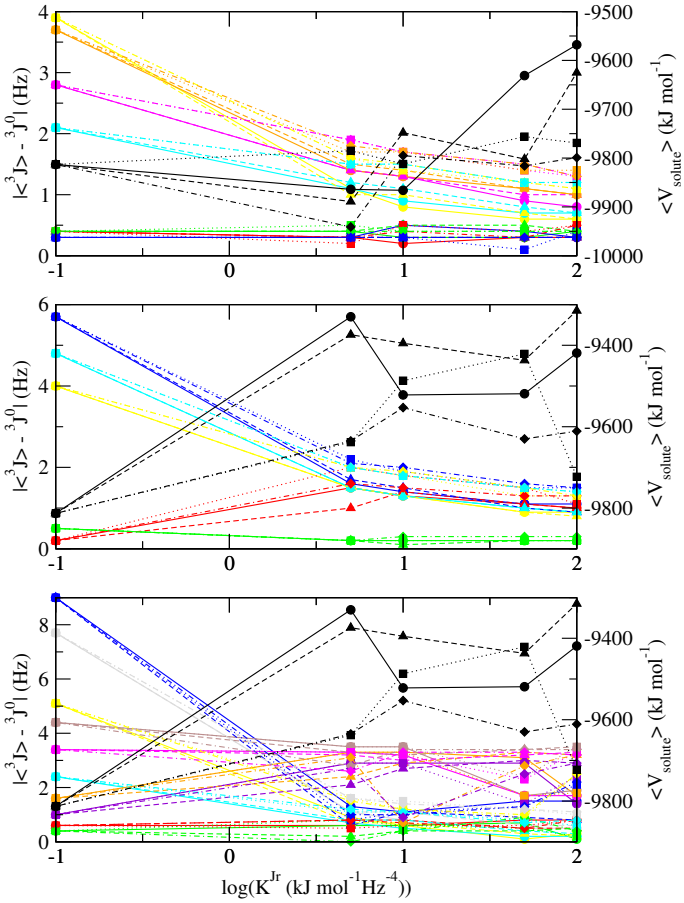

Figure 5: Size of the deviation of the average ^3J-coupling value $\langle {}^3J \rangle$ from the target value $^3J^0$ (left scale, coloured lines) and the intra-protein energy (bonded and non-bonded, right scale, black lines) from 2 ns MD simulations without ^3J-coupling restraining ($K^{J_r} = 0$, $\log K^{J_r} = -\infty$, but put at -1) or with biquadratic time-averaging restraining (BQ-TAR, $K^{J_r} > 0$), one with restraining 95 main chain $^3J_{\mathrm{H_N H_\alpha}}$-couplings and one with restraining 62 side chain $^3J_{\mathrm{H_\alpha H_\beta}}$-couplings as function of the restraining force constant $K^{J_r} = K^{Q,\mathrm{restr}}$ and for different values of the restraining relaxation time τ_Q (5 and 50 ps) and of the flat-bottom range $\Delta^3J^{\mathrm{fb}} = \Delta Q^{\mathrm{fb}}$ (0.5 and 1 Hz) of the restraining function. Data for $\tau_Q = 5\,\mathrm{ps}$: $\Delta^3J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$ (circles and solid lines), $\Delta^3J^{\mathrm{fb}} = 1\,\mathrm{Hz}$ (squares and dotted lines). Data for $\tau_Q = 50\,\mathrm{ps}$: $\Delta^3J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$ (triangles and dashed lines), $\Delta^3J^{\mathrm{fb}} = 1\,\mathrm{Hz}$ (diamonds and dot-dash lines). Upper panel: main chain $^3J_{\mathrm{H_N H_\alpha}}$ for 7 residues: Thr 43 (red), Ala 9 (green), Asn 19 (blue), Thr 69 (yellow), Lys 33 (light blue), Met 105 (magenta), Ile 124 (orange). Middle panel: side chain (one H_β) $^3J_{\mathrm{H_\alpha H_\beta}}$ for 5 residues: Val 2 (red), Thr 47 (green), Thr 89 (blue), Val 99 (yellow), Val 109 (light blue). Lower panel: side chain (two H_β) $^3J_{\mathrm{H_\alpha H_\beta}}$ for 5 residues: Asp 52 (H_{β_2}: red, H_{β_3}: green), Asn 46 (H_{β_2}: blue, H_{β_3}: yellow), Asp 66 (H_{β_2}: light blue, H_{β_3}: grey), Glu 7 (H_{β_2}: orange, H_{β_3}: violet), Arg 45 (H_{β_2}: brown, H_{β_3}: magenta).

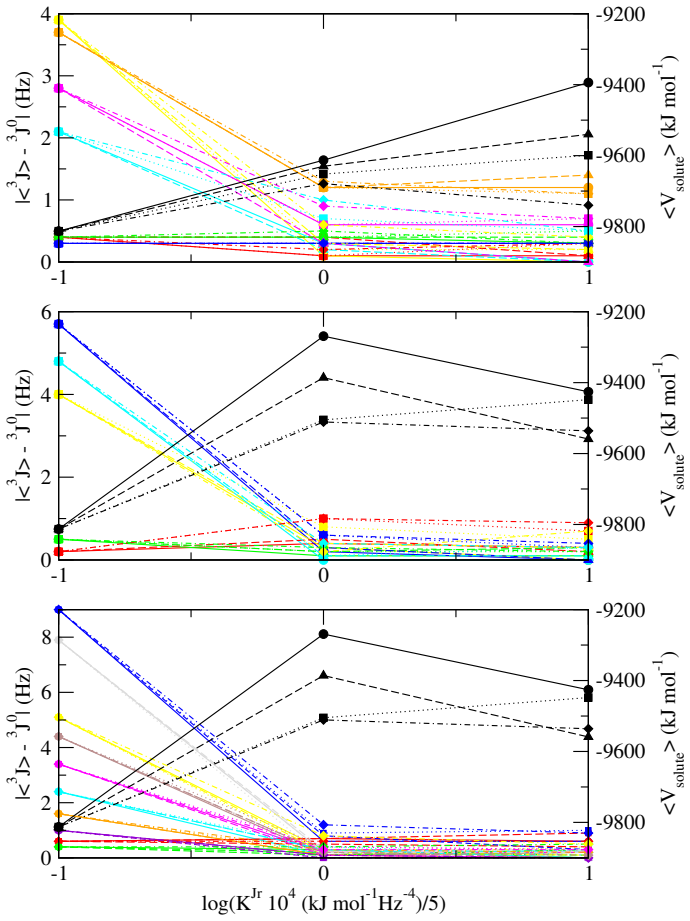

Figure 6: Size of the deviation of the average ^3J-coupling value $\langle {}^3J \rangle$ from the target value $^3J^0$ (left scale, coloured lines) and the intra-protein energy (bonded and non-bonded, right scale, black lines) from 2 ns MD simulations without ^3J-coupling restraining ($K^{J_r} = 0$, $\log K^{J_r} = -\infty$, but put at -1) or with biquadratic time-averaging local-elevation restraining (BQ-TA-LER, $K^{J_r} > 0$, $\Delta\varphi^0 = 10°$, $N_{le} = 36$), one with restraining 95 main chain $^3J_{H_N H_\alpha}$-couplings and one with restraining 62 side chain $^3J_{H_\alpha H_\beta}$-couplings as function of the restraining force constant $K^{J_r} = K^{Q,restr}$ and for different values of the restraining relaxation time τ_Q (5 and 50 ps) and of the flat-bottom range $\Delta^3J^{fb} = \Delta Q^{fb}$ (0.5 and 1 Hz) of the restraining function. Data for $\tau_Q = 5$ ps: $\Delta^3J^{fb} = 0.5$ Hz (circles and solid lines), $\Delta^3J^{fb} = 1$ Hz (squares and dotted lines). Data for $\tau_Q = 50$ ps: $\Delta^3J^{fb} = 0.5$ Hz (triangles and dashed lines), $\Delta^3J^{fb} = 1$ Hz (diamonds and dot-dash lines). Upper panel: main chain $^3J_{H_N H_\alpha}$ for 7 residues: Thr 43 (red), Ala 9 (green), Asn 19 (blue), Thr 69 (yellow), Lys 33 (light blue), Met 105 (magenta), Ile 124 (orange). Middle panel: side chain (one H_β) $^3J_{H_\alpha H_\beta}$ for 5 residues: Val 2 (red), Thr 47 (green), Thr 89 (blue), Val 99 (yellow), Val 109 (light blue). Lower panel: side chain (two H_β) $^3J_{H_\alpha H_\beta}$ for 5 residues: Asp 52 (H_{β_2}: red, H_{β_3}: green), Asn 46 (H_{β_2}: blue, H_{β_3}: yellow), Asp 66 (H_{β_2}: light blue, H_{β_3}: grey), Glu 7 (H_{β_2}: orange, H_{β_3}: violet), Arg 45 (H_{β_2}: brown, H_{β_3}: magenta).

4 Discussion

The conformational sampling behavior of the three different ^3J-coupling restraining techniques is illustrated in Figs 7 and 8 for the side chains of Val 29 and Val 109. The IR, BQ-TAR and BQ-TA-LER restraining simulations for Val 29 start from a χ_1-angle value of $-63°$ that corresponds to a $^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-coupling of 3 Hz. This value is quite far from the target $^3J^0$ value of 11.1 Hz for this $^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-coupling. The upper panels of Figure 7 show that by applying instantaneous restraining (IR) the target $^3J^0$ value is not reached and thus the restraining energy does not become zero, which is due to the very limited sampling of the IR technique. This picture does not change much if BQ-TAR is applied. In both cases (IR and BQ-TAR) the χ_1-angle is driven to the lower maximum of the Karplus curve. Applying BQ-TA-LER (lower panels) a very wide range of χ_1-angle values and thus $^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-couplings is sampled, leading to a dominant conformation with χ_1-angle values between $-150°$ and $-200°$, $^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$ close to the target $^3J^0$ value and a restraining energy close to zero.

Figure 8 shows a case, Val 109, in which conformational averaging leads to reproduction of the target $^3J^0$ value. Again, IR (and BQ-TAR) are not able to match the target $^3J^0$ value and show no (IR) or very little (BQ-TAR) conformational sampling (upper and middle panels), while BQ-TA-LER allows wide sampling leading to reproduction of the $^3J^0$ target value. We note that the restraining energy very slowly increases with time. This may indicate that the relaxation time $\tau_Q = 5\,\mathrm{ps}$ is too short to allow wide sampling or that the flat bottom of the restraining function is with a value of $\Delta^3J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$ too small. Indeed, using either $\tau_Q = 50\,\mathrm{ps}$ or $\Delta^3J^{\mathrm{fb}} = 1.0\,\mathrm{Hz}$ the slow increase in restraining energy disappears.

A comparison of the average values of the intra-protein potential energy in the IR (Figure 4) and the BQ-TA-LER (Figure 6) simulations shows that the latter type of restraining leads to a higher intra-protein energy. This may be due to the wider conformational sampling of φ- or χ_1-angles using the BQ-TA-LER technique which

involves crossings of intra-protein energetic barriers that contribute to the average intra-protein potential energy. The addition of a restraining term to the potential energy or interaction function of a system will lead to an increase of the energy value of the other potential energy terms in case the system is in a minimum energy configuration in regard to the interaction function. By adding a restraining potential energy term the energy hypersurface of the system is modified and the restraining forces move the system away from the configuration for which the potential energy without restraints is minimal. If the system before restraining is not in a minimal potential energy region of the hypersurface, wider sampling using restraints may lead to a lower potential energy, as is illustrated in Figure 6 (lower panels). If the intra-protein energy rises substantially upon restraining, this may have different origins:

1. The molecular interaction function disfavours configurations that satisfy the restraints;

2. The NMR data originate from multiple protein conformations and the MD sampling is insufficiently wide to grasp all those configurations;

3. The Karplus relation linking protein conformations to ^3J-coupling values is insufficiently accurate;

4. The NMR data are erroneous, i.e. there exists no configuration or set of configurations that do satisfy the restraints, see (Allison and van Gunsteren, 2009).

The BQ-TA-LER simulations involving restraining of the χ_1-angles display for $\Delta^3 J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$ a decrease of the average intra-protein potential energy upon increasing the restraining force constant from $5\times10^{-4}\,\mathrm{kJ\,mol^{-1}Hz^{-4}}$ to $50\times10^{-4}\,\mathrm{kJ\,mol^{-1}Hz^{-4}}$. Increasing the restraining force constant will induce wider conformational sampling if the flat bottom $\Delta^3 J^{\mathrm{fb}}$ is relatively small and the target $^3J^0$ value is not reached.

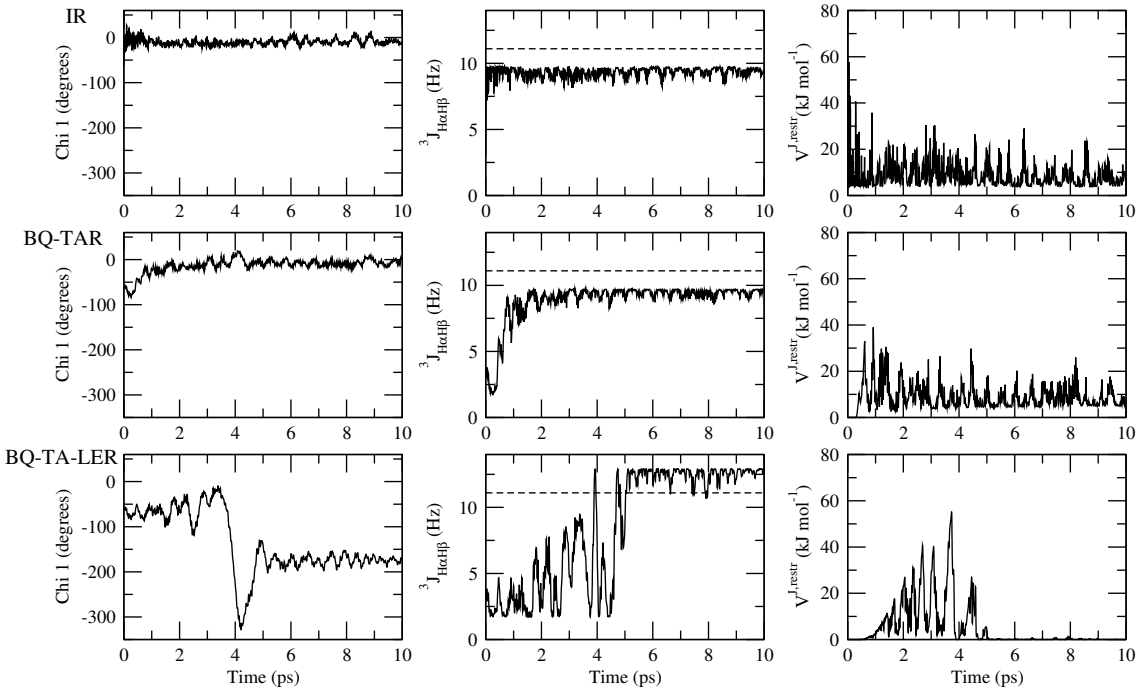

Figure 7: Time series of the side chain χ_1 torsional angle (left panels), the corresponding $^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-coupling (middle panels) and the corresponding restraining potential energy function term (right panels, Eqs. 7 and 8 for IR, Eqs. 12 and 13 for BQ-TAR, Eq. 20 for BQ-TA-LER) for Val 29 from 2 ns MD simulations. Upper panels: instantaneous restraining (IR) with $K^{J_\mathrm{r}} = 10\,\mathrm{kJ\,mol^{-1}Hz^{-2}}$. Middle panels: biquadratic time-averaging restraining (BQ-TAR) with $K^{J_\mathrm{r}} = 10\,\mathrm{kJ\,mol^{-1}Hz^{-4}}$ and $\tau_Q = 5\,\mathrm{ps}$. Lower panels: biquadratic time-averaging local elevation restraining (BQ-TA-LER) with $K^{J_\mathrm{r}} = 5 \times 10^{-4}\,\mathrm{kJ\,mol^{-1}Hz^{-4}}$, $\tau_Q = 5\,\mathrm{ps}$, $N_\mathrm{le} = 36$, and $\Delta\varphi^0 = 10°$. In all cases $\Delta^3J^\mathrm{fb} = 0.5\,\mathrm{Hz}$, $^3J^0 = 11.1\,\mathrm{Hz}$.

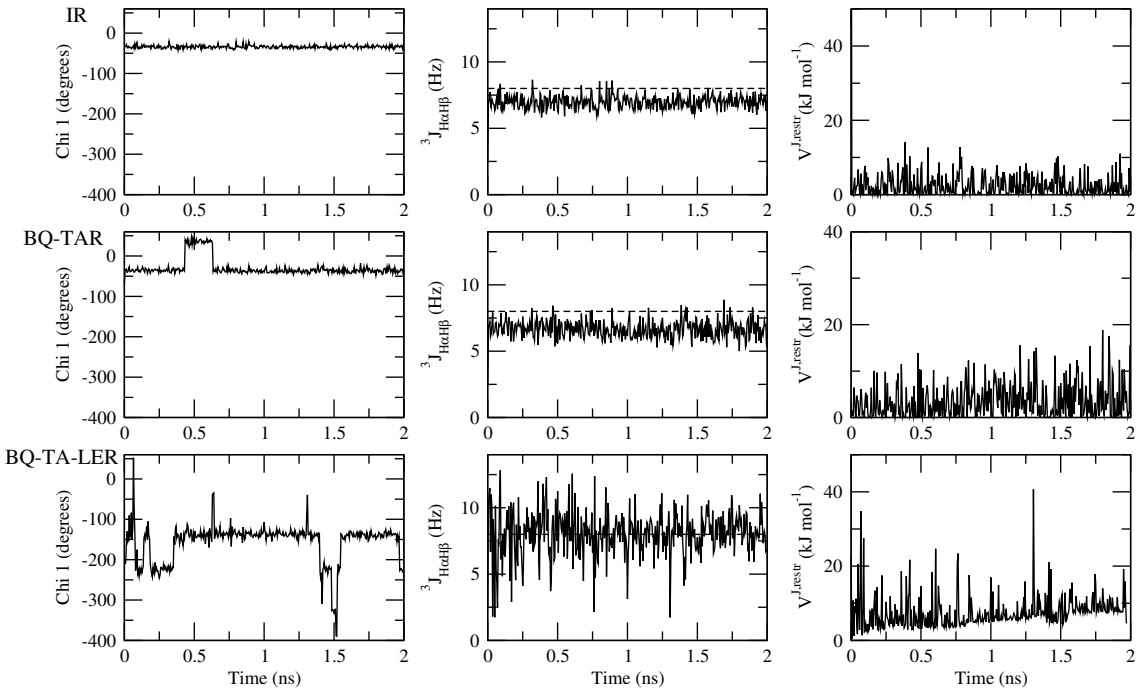

Figure 8: Time series of the side chain χ_1 torsional angle (left panels), the corresponding $^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-coupling (middle panels) and the corresponding restraining potential energy function term (right panels, Eqs. 7 and 8 for IR, Eqs. 12 and 13 for BQ-TAR, Eq. 20 for BQ-TA-LER) for Val 109 from 2 ns MD simulations. Upper panels: instantaneous restraining (IR) with $K^{J_\mathrm{r}} = 10\,\mathrm{kJ\,mol^{-1}Hz^{-2}}$. Middle panels: biquadratic time-averaging restraining (BQ-TAR) with $K^{J_\mathrm{r}} = 10\,\mathrm{kJ\,mol^{-1}Hz^{-4}}$ and $\tau_Q = 5\,\mathrm{ps}$. Lower panels: biquadratic time-averaging local elevation restraining (BQ-TA-LER) with $K^{J_\mathrm{r}} = 5 \times 10^{-4}\,\mathrm{kJ\,mol^{-1}Hz^{-4}}$, $\tau_Q = 5\,\mathrm{ps}$, $N_\mathrm{le} = 36$, and $\Delta\varphi^0 = 10°$. In all cases $\Delta^3J^\mathrm{fb} = 0.5\,\mathrm{Hz}$, $^3J^0 = 8.0\,\mathrm{Hz}$.

5 Conclusion

The use of ^3J-couplings in structure refinement of biomolecules is a challenge due to the inaccuracy of the Karplus relation $^3J(\theta)$ connecting a ^3J-coupling value to the corresponding torsional angle θ and the multiple-valuedness of the inverse relation $\theta(^3J)$. Standard single-structure refinement or instantaneous restraining (IR) is unable to meet this challenge because it does not account for the averaging inherent to the measurement of ^3J-couplings and it suffers from limited conformational sampling. IR restricts the conformational fluctuations artificially for larger restraining force constants, while for lower ones the target $^3J^0$ values may not be reached. Allowing for conformational averaging, as in biquadratic time-averaging restraining (BQ-TAR), does reduce these problems, but not satisfactorily. When using time-averaging and local-elevation sampling restraining (BQ-TA-LER) the problems are solved through wide conformational sampling and the ability of the LE method to surmount energy barriers. The results show that there may be multiple different torsional-angle ranges that lead to reproduction of the target ^3J-coupling values in structure refinement simulations.

In regard to the choice of the three parameters of the BQ-TA-LER method, theoretical considerations dictate that (i) the force constant $K^{Q,\mathrm{restr}} = K^{J_\mathrm{r}}$ should be chosen as small as possible while large enough to bring the average ^3J-couplings close to the target values, in order to approximate as good as possible the properly $(V^{\mathrm{phys}}(\vec{r}^N))$ Boltzmann weighted conformational probability distribution, (ii) the averaging time τ_Q should match the experimental one but should not be larger than $\frac{1}{10}$ of the total simulation time in order to secure sufficient statistics when averaging, and (iii) the flat-bottom $\Delta Q^{\mathrm{fb}} = \Delta^3 J^{\mathrm{fb}}$ should represent the uncertainty or inaccuracy of the Karplus relation $^3J(\theta)$. The data obtained here suggest that $K^{J_\mathrm{r}} = 5\times10^{-4}\,\mathrm{kJmol^{-1}Hz^{-4}}$, $\tau_Q = 5\,\mathrm{ps}$ and $\Delta^3 J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$ are generally appropriate to generate a conformational ensemble that reproduces the experimental $^3J_{\mathrm{H_N H_\alpha}}$- and

$^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-couplings of hen egg white lysozyme.

Whether the combined restraining of different sets of values for observable quantities of the same type, e.g. backbone and side-chain ^3J-couplings, or of different type, e.g. ^3J-couplings with NOE distance bounds or residual dipolar couplings (RDCs), will change the sampling properties of the restraining MD simulations will depend on the system (protein), on the number and type of the restraints and their distribution over the protein, on which degrees of freedom are directly related to the restraints and on how the restraining is executed (flat-bottom, allowing for averaging, choice of values for the restraining parameters, etc.), see e.g. (Smith et al, 2016). This is currently under investigation.

Acknowledgements

LJS would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out some of this work. http://dx.doi.org/10.5281/zenodo.22558. WFvG thanks the Swiss National Science Foundation, grant number 200020-137827, and the European Research Council, grant number 228076, for financial support. NH thanks the German Research Foundation (DFG) for financial support within the Cluster of Excellence in Simulation Technology (EXC 310/2) at the University of Stuttgart.

References

Abraham RJ, McLauchlan KA (1968) The proton resonance spectra and conformations of the prolines. Mol Phys 5:513–523

Allison JR, van Gunsteren WF (2009) A method to explore protein side chain conformational variability using experimental data. ChemPhysChem 10:3213–3228

Artymiuk PJ, Blake CCF, Rice DW, Wilson KS (1982) The structures of the monoclinic and orthorhombic forms of hen egg-white lysozyme at 6 Ångstroms resolution. Acta Crystallogr Sect B 38:778–783

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, the Netherlands, 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

Brüschweiler R, Case DA (1994) Adding harmonic motion to the Karplus relation for spin-spin coupling. J Am Chem Soc 116:11,199–11,200

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

Cung MT, Marraud M (1982) Conformational dependence of the vicinal proton coupling constant for the C^{α}-C^{β} bond in peptides. Biopolymers 21:953–967

Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE (1999) Peptide folding: When simulation meets experiment. Angew Chemie Intl Ed 38:236–240

de Vlieg J, Boelens R, Scheek RM, Kaptein R, van Gunsteren WF (1986) Restrained molecular dynamics procedure for protein tertiary structure determination from NMR data: A lac repressor headpiece structure based on information on J-coupling and from presence and absence of NOE's. Isr J Chem 27:181–188

Deber CM, Torchia DA, Blout ER (1971) Cyclic peptides. I. Cyclo(tri-L-propyl) and derivatives. Synthesis and molecular conformation from nuclear magnetic resonance. J Am Chem Soc 93:4893–4897

deMarco A, Llinás M, Wüthrich K (1978) Analysis of the ^{1}H-NMR spectra of ferrichrome peptides. I. The non-amide protons. Biopolymers 17:617–636

Drenth J (1994) Principles of Protein X-ray Crystallography. Springer, New York

Fischman AJ, Live DH, Wyssbrod HR, Agosta WC, Cowburn D (1980) Torsion angles in the cystine bridge of oxytocin in aqueous solution. Measurements of circumjacent vicinal couplings between ^{1}H, ^{13}C, and ^{15}N. J Am Chem Soc 102:2533–2539

Gros P, van Gunsteren WF, Hol WG (1990) Inclusion of thermal motion in crystallographic structures by restrained molecular dynamics. Science 249:1149–1152

Hansen N, Heller F, Schmid N, van Gunsteren WF (2014) Time-averaged order parameter restraints in molecular dynamics simulations. J Biomol NMR 60:169–187

Harvey TS, van Gunsteren WF (1993) The application of chemical shift calculation to protein structure determination by NMR. In: Angeletti RH (ed) Techniques in Protein Chemistry IV, Academic Press, San Diego, pp 615–615

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

Huber T, Torda AE, van Gunsteren WF (1994) Local elevation: a method for improving the searching properties of molecular dynamics simulation. J Comput-Aided Mol Des 8:695–708

Kaptein R, Zuiderweg ERP, Scheek RM, Boelens R, van Gunsteren WF (1985) A protein structure from nuclear magnetic resonance data: lac repressor headpiece. J Mol Biol 182:179–182

Karplus M (1959) Contact electron-spin coupling of nuclear magnetic moments. J Chem Phys 30:11–15

Keller B, Christen M, Oostenbrink C, van Gunsteren WF (2007) On using oscillating time-dependent restraints in MD simulation. J Biomol NMR 37:1–14

Nanzer AP, Poulsen FM, van Gunsteren WF, Torda AE (1994) Reassessment of the structure of chymotrypsin inhibitor 2 (CI-2) using time-averaged NMR restraints. Biochemistry 33:14,503–14,511

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. J Mol Biol 267:1012–1025

Pardi A, Billeter M, Wüthrich K (1984) Calibration of the angular dependence of the amide proton-C^{α} proton coupling constants, $^{3}J_{\mathrm{HN}\alpha}$, in a globular protein. Use of $^{3}J_{\mathrm{HN}\alpha}$ for identification of helical secondary structure. J Mol Biol 180:741–751

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 nuclear magnetic resonance refinement?: A model study for DNA. J Mol Biol 220:457–479

Pérez C, Löhr F, Rüterjans H, Schmidt JM (2001) Self-consistent Karplus parametrization of ^{3}J couplings depending on the polypeptide side-chain torsion χ_1. J Am Chem Soc 123:7081–7093

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

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, Huber R, Wüthrich K, van Gunsteren WF (1994) 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 241:588–599

Schiffer CA, Gros P, van Gunsteren WF (1995) Time-averaging crystallographic refinement: Possibilities and limitations using alpha-cyclodextrin as a test system. Acta Cryst D51:82–92

Schmid N, Allison JR, Dolenc J, Eichenberger AP, Kunz AP, 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 Gun-

steren WF (2011b) Definition and testing of the GROMOS force-field versions: 54A7 and 54B7. Eur Biophys J 40:843–856

Schmidt JM, Blümel M, Löhr F, Rüterjans H (1999) Self-consistent ^3J coupling analysis for the joint calibration of Karplus coefficients and evaluation of torsion angles. J Biomol NMR 14:1–12

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 NB, Kumar A, James TL (1993) Molecular dynamics with weighted time-averaged restraints for a DNA octamer: dynamic interpretation of nuclear magnetic resonance data. J Mol Biol 234:373–389

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

Smith LJ, Sutcliffe MJ, Redfield C, Dobson CM (1991) Analysis of ϕ and χ_1 torsion angles for hen lysozyme in solution from ^{1}H NMR spin-spin coupling constants. Biochemistry 30:986–996

Smith LJ, Rought Whitta G, Dolenc J, Wang D, van Gunsteren WF (2016) A molecular dynamics simulation investigation of the relative stability of the cyclic peptide octreotide and its deprotonated and its (CF_3)-Trp substituted analogs in different solvents. Bioorg Med Chem http://dx.doi.org/10.1016/j.bmc.2016.08.001

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, Scheek RM, van Gunsteren WF (1990) Time-averaged nuclear Overhauser effect distance restraints applied to tendamistat. J Mol Biol 214:223–235

Torda AE, Brunne RM, Huber T, Kessler H, van Gunsteren WF (1993) Structure

refinement using time-averaged J-coupling constant restraints. J Biomol NMR 3:55–66

van Gunsteren WF (1990) On testing theoretical models by comparison of calculated with experimental data. In: Rivail JL (ed) Studies in Physical and Theoretical Chemistry, Vol 71, Modelling of Molecular Structures and Properties, Elsevier, Amsterdam, pp 463–478

van Gunsteren WF (2016) http://www.gromos.net, accessed on May 13, 2016

van Gunsteren WF, Kaptein R, Zuiderweg ERP (1984) Use of molecular dynamics computer simulations when determining protein structure by 2D NMR. In: Olsen WK (ed) Proceedings of the NATO/CECAM Workshop on Nucleic Acid Conformation and Dynamics, Orsay, pp 79–82

van Gunsteren WF, Boelens R, Kaptein R, Scheek RM, Zuiderweg ERP (1985) An improved restrained molecular dynamics technique to obtain protein tertiary structure from nuclear magnetic resonance data. In: Molecular Dynamics and Protein Structure, Polycrystal Book Service, Western Springs, Ill., USA, pp 92–99

van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG (1996) Biomolecular Simulation: The GROMOS96 Manual and User Guide. Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Groningen

van Schaik RC, Berendsen HJC, Torda AE, van Gunsteren WF (1993) A structure refinement method based on molecular dynamics in four spatial dimensions. J Mol Biol 234:751–762

Wang AC, Bax A (1996) Determination of the backbone dihedral angles ϕ in human ubiquitin from reparametrized empirical Karplus equations. J Am Chem Soc 118:2483–2494

Wüthrich K (1986) NMR of proteins and nucleic acids. Wiley, New York

Tables

Table 1: Average values and root-mean-square fluctuations of ${}^3J_{\mathrm{H}_\alpha\mathrm{H}_\beta}$-couplings (in Hz) and the corresponding χ_1-angle values (in degree) from 2 ns MD simulations with instantaneous restraining (IR; $K^{J_\mathrm{r}} = 10\,\mathrm{kJmol^{-1}Hz^{-2}}$, $\Delta^3J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$) or with biquadratic time-averaging local-elevation restraining (BQ-TA-LER; $K^{J_\mathrm{r}} = 5 \times 10^{-4}\mathrm{kJ\,mol^{-1}Hz^{-4}}$, $\tau_Q = 5\,\mathrm{ps}$, $\Delta^3J^{\mathrm{fb}} = 0.5\,\mathrm{Hz}$), for two residues of HEWL. The standard Karplus curve (deMarco et al, 1978) (upper part) and this curve shifted over +120° (lower part) is used in restraining. The latter mimics the case in which the initial value of χ_1 lies outside the region from which the target value can be reached by following the gradient of the restraining function.

	Val 2		Val 29	
	$\langle {}^3J \rangle$	$\langle \chi_1 \rangle$	$\langle {}^3J \rangle$	$\langle \chi_1 \rangle$
Target ${}^3J^0$	10.8	-	11.1	-
Initial value	12.9	-177	3.0	-63
Standard Karplus curve				
IR	11.6 (0.7)	-168 (17)	9.5 (0.4)	-5 (9)
BQ-TA-LER	11.2 (1.5)	-167 (26)	11.6 (1.0)	-175 (21)
+120° shifted Karplus curve				
IR	9.5 (0.4)	-121 (9)	9.6 (0.3)	-121 (8)
BQ-TA-LER	11.2 (1.2)	+80 (14)	11.4 (1.1)	+78 (18)