ARTICLE

# On the calculation of $^3J_{\alpha\beta}$-coupling constants for side chains in proteins

Denise Steiner · Jane R. Allison · Andreas P. Eichenber-
ger · Wilfred F. van Gunsteren

**Abstract** Structural knowledge about proteins is mainly derived from values of observables, measurable in NMR spectroscopic or X-ray diffraction experiments, i.e. absorbed or scattered intensities, through theoretically derived relationships between structural quantities such as atom positions or torsional angles on the one hand and observable quantities such as squared structure factor amplitudes, NOE intensities or $^3J$-coupling constants on the other. The standardly used relation connecting $^3J$-couplings to torsional angles is the Karplus relation, which is used in protein structure refinement as well as in the evaluation of simulated properties of proteins. The accuracy of the simple and generalised Karplus relations is investigated using side-chain structural and $^3J_{\alpha\beta}$-coupling data for three different proteins, Plastocyanin, Lysozyme, and FKBP, for which such data are available. The results show that the widely used Karplus relations are only a rough estimate for the relation between $^3J_{\alpha\beta}$-couplings and the corresponding $\chi_1$-angle in proteins.

D. Steiner · J. R. Allison · A. P. Eichenberger ·
W. F. van Gunsteren (✉)
Laboratory of Physical Chemistry, ETH, Swiss Federal Institute
of Technology, 8093 Zurich, Switzerland
e-mail: wfvgn@igc.phys.chem.ethz.ch

*Present Address:*
J. R. Allison
Centre for Theoretical Chemistry and Physics, Institute for
Natural Sciences, Massey University Albany, Albany Highway,
Auckland 0632, New Zealand

## Introduction

A precise determination of the structural properties of proteins is still one of the major challenges in molecular biology, although thousands of protein structures in a crystalline environment or in aqueous solution at a particular thermodynamic state point, i.e. temperature, pH, ionic strength, etc., have been determined through X-ray diffraction or NMR spectroscopic experiments (Berman et al. 2000). The protein structures derived from X-ray diffraction intensities are generally of relatively high precision, because the ratio of the number of observable (independent) intensities $N^{obs}$ and the number of spatial degrees of freedom of the protein $N^{df}$ is larger than one. Moreover, the relation between the intensity of the diffracted beam and the structure of the protein is simple and well-known: the intensity of a diffraction peak is proportional to the square of the amplitude of the corresponding spatial Fourier transform of the electron density (Hendrickson and Konnert 1981), which is in turn directly related to the structure of the protein in terms of atom positions. The precision of a protein structure derived from X-ray diffraction data is mainly determined by the spatial resolution of the latter, which determines the ratio $N^{obs}/N^{df}$.

NMR experiments can deliver measured values for a variety of observable quantities, including intensities of nuclear Overhauser effect (NOE) peaks, residual dipolar couplings (RDCs), $^3J$-coupling constants, or chemical shifts (Wüthrich 1986). The precision of protein structures derived from NMR data is generally much lower than that of protein structures derived from X-ray diffraction data. This relatively low precision is caused by various aspects of the methodology used to derive protein structural information from NMR data (van Gunsteren et al. 1991, 1994, 1999; Salmon et al. 2011), and of the NMR data itself.

With respect to the NMR data, the first issue is that the number of measured values of observable quantities at a particular thermodynamic state point is much smaller than the number of protein degrees of freedom. Even if different observables, such as NOEs, RDCs, or $^3J$-couplings, are combined, the ratio $N^{obs}/N^{df}$ is still lower than one. Moreover, correlation between different data may reduce the number of independent data, the data may originate from different experiments at different thermodynamic state points at which the protein's structure may not be the same, and the quantities observable in an NMR experiment are in general related to particular subsets of the atoms of a protein.

A second problem is that the relation between an observable quantity $Q(\mathbf{r}^N)$ and the structure of a protein represented by the Cartesian coordinates $\mathbf{r}^N \equiv (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ of its $N$ atoms is generally not very precisely known for the aforementioned observables measurable by NMR. An NOE intensity depends not only on the distance between the two atoms involved, but also on the rotational motion of the protein and on the distances to other protein atoms surrounding the atom pair due to spin diffusion effects. An RDC depends not only on the angle $\theta$ between the vector connecting two atoms and the magnetic field direction for a single protein structure, but also on the spatial distribution of these vectors, i.e. on the distribution of protein orientations in a medium that induces a slight deviation from a uniform spherical distribution. A $^3J$-coupling constant depends not only on the dihedral angle $\theta$ between the four atoms involved and their types, but also on the substituents at the two central atoms of the dihedral angle. A chemical shift depends not only on the relative position of the atom involved with respect to its covalently bound neighbour atoms and their types, but also on the distance to non-bonded neighbour atoms and their types. In principle, the value of an NMR observable $Q$ can be calculated from $\mathbf{r}^N$ using quantum-chemical methods, but the accuracy that currently can be reached is rather low due to the approximations made during the calculations that are required by the finite computing power available. Therefore, semi-empirical, approximate functions $Q(\mathbf{r}^N)$ are generally used to relate protein structure to observable quantities.

A further difficulty in determining structural properties from NMR data is that the relations $Q(\mathbf{r}^N)$ for NMR observables are highly non-linear: $Q$ depends on $r^{-3}$ or $r^{-6}$ for a NOE distance $r$ between two atoms, it depends on the cosine of the angle $\theta$ for an RDC, and on the cosine of the angle $\theta$ and its square for a $^3J$-coupling. Together, these aspects make a precise determination of protein structural properties based on NMR data a challenging task.

Protein structure determination, be it based on X-ray, NMR or other experimental data, should also account for the motion or conformational variability of a protein, because all but a few experiments involve averaging over time and over the ensemble of protein structures in the sample. Due to the crystalline packing and the linear character of the Fourier transform, the neglect of properly accounting for conformational averaging in procedures to derive protein structure from measured data is much less aggravating when using X-ray diffraction data than when using NMR spectroscopic data. In particular, for $^3J$-coupling data measured for protein side chains, it is essential to properly account for averaging because of the strong non-linearity of the function $Q(\mathbf{r}^N)$ and the variety of possible side-chain conformations (Allison and van Gunsteren 2009). $^3J$-coupling data for protein side chains are less often used to determine protein structures than the corresponding backbone data. Yet, the information they provide about the distribution of side-chain dihedral angles is essential for characterising protein structures in view of the tight spatial packing of side chains in the interior of a protein.

A first approximation of the relation $Q(\mathbf{r}^N)$ between a vicinal $^3J$-coupling constant $^3J(A_1, A_4) = {}^3J(A_1-A_2-A_3-A_4)$ between two atoms $A_1$ and $A_4$ that are covalently connected through three bonds involving the atoms $A_2$ and $A_3$ and the dihedral angle $\theta = A_1-A_2-A_3-A_4$ is given by the Karplus relation (Karplus 1959, 1963)

$$^3J_{A_1 A_4}(\theta) = a\cos^2(\theta) + b\cos(\theta) + c, \qquad (1)$$

in which the coefficients $a$, $b$, and $c$ are parameters that depend on the types of the atoms $A_1$ to $A_4$ and in principle on the number and types of substituents at atoms $A_2$ and $A_3$. Using Eq. 1, the dependence of $^3J_{A_1 A_4}$ on the geometry of the configuration of atoms $A_1$ to $A_4$ and their substituents is reduced to a simple function of one dihedral angle $\theta$.

In proteins, different types of $^3J$-couplings can be observed, which are related to particular torsional angles, for example for $\theta_\alpha = H-N-C_\alpha-H_\alpha$, $^3J_{H_N H_\alpha}$ is related to $\varphi = C-N-C_\alpha-C$, and for $\theta_\beta = H_\alpha-C_\alpha-C_\beta-H_\beta$, $^3J_{H_\alpha H_\beta} = {}^3J_{\alpha\beta}$ is related to $\chi_1 = N-C_\alpha-C_\beta-C_\gamma/O_\gamma/S_\gamma$.

The relation between the angles $\theta_\alpha$ and $\theta_\beta$ and the angles $\varphi$ and $\chi_1$ depends on the configuration of the atoms involved and is generally approximated by

$$\theta_\alpha = \varphi + \delta_\alpha \qquad (2)$$

and

$$\theta_\beta = \chi_1 + \delta_\beta, \qquad (3)$$
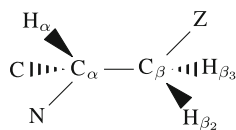
with $\delta_\alpha = -60°$, $\delta_{\beta_{H_2}} = -120°$, and $\delta_{\beta_{H_3}} = 0°$ for an L-amino acid residue and $\delta_\alpha = +60°$, $\delta_{\beta_{H_2}} = 0°$, and $\delta_{\beta_{H_3}} = +120°$ for a D-amino acid residue.

A more complex description of the relation $Q(\mathbf{r}^N)$ between measured proton-proton scalar $^3J_{HH}$-couplings and

the corresponding dihedral-angle values is given by the generalised Karplus equation proposed by Haasnoot et al. (1979, 1981) and used by Mádi et al. (1990), which takes into account the substituents of atoms $A_2$ and $A_3$. It applies to $^3J_{HH}$-couplings, i.e. $A_1 = H$ and $A_4 = H$, for which $A_2 = C$ and $A_3 = C$, and for which the fragment H–$A_2$–$A_3$–H bears *three* non-hydrogen substituents, see Fig. 1 and "Method" section. This generalised Karplus relation can be used to calculate $^3J_{\alpha\beta}$-couplings that depend on the $\theta_\beta$ torsional angle related to the $\chi_1$ torsional angle for 15 of the 20 amino acids naturally occurring in proteins. The exceptions are Ile, Thr and Val, for which the fragments H–$C_\alpha$–$C_\beta$–H bear four non-hydrogen substituents, and the residues Ala and Gly, which do not have $\chi_1$-angles.

The parameters $a$, $b$, and $c$ for the standard Karplus relation are generally determined empirically. A variety of different sets of Karplus parameters have been determined using different molecules and methodologies (Abraham and McLauchlan 1962; Deber et al. 1971; Kopple et al. 1973; Bystrov 1976; de Marco et al. 1978; Fischman et al. 1980; Pardi et al. 1984; Brüschweiler and Case 1994; Wang and Bax 1996; Schmidt et al. 1999; Pérez et al. 2001; Lindorff-Larsen et al. 2005; Vögeli et al. 2007). Often, the value of a particular torsional angle $\varphi$ or $\chi_1$ in the X-ray diffraction structure of a molecule in crystal form is assumed to be related through Eqs. 1–3 to the corresponding $^3J$-coupling measured for the same molecule in aqueous solution (de Marco et al. 1978; Pardi et al. 1984; Wang and Bax 1996). Using all available $^3J$-couplings for one or more molecules, a set of Karplus parameters can be obtained that minimises the sum of the squared differences of the measured $^3J$-couplings $^3J^{exp}$ to the ones calculated from the $\varphi$- or $\chi_1$-angles using Eqs. 1–3. Such a procedure rests upon the assumption that the value of a torsional angle

in the crystal is a good approximation of the value of the same angle in solution, and that conformational averaging plays a similar role in both environments. The approximate nature of these assumptions is illustrated by the variation of the Karplus parameters obtained using different sets of data. For $^3J_{H_N H_\alpha}$, the parameter ranges found in the literature (Bystrov 1976; Pardi et al. 1984; Brüschweiler and Case 1994; Wang and Bax 1996; Schmidt et al. 1999; Lindorff-Larsen et al. 2005; Vögeli et al. 2007) are 3.1 Hz for $a$, 0.64 Hz for $b$, and 1.6 Hz for $c$. For $^3J_{\alpha\beta}$, the variety of parameter values is even larger (Abraham and McLauchlan 1962; de Marco et al. 1978; Deber et al. 1971; Kopple et al. 1973; Fischman et al. 1980; Pérez et al. 2001), 4.87 Hz for $a$, 0.6 Hz for $b$, and 2.2 Hz for $c$, see Table 1, leading to quite some variation in the resulting Karplus curves (Fig. 2).

The precision of $a$, $b$, and $c$ values may be affected by the fact that the values of the $\varphi$- or $\chi_1$-angles in the crystal structures of the proteins or other molecules used to obtain the Karplus parameters do not cover the whole 360°
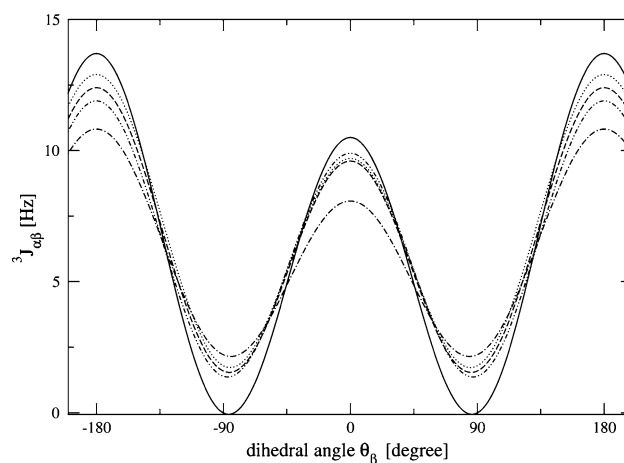


**Fig. 2** Karplus curves of Eq. 1 for $^3J_{\alpha\beta_3}$ as a function of $\theta_{\beta_3}$ using different values of the parameters $a$, $b$, and $c$ from the literature. The *solid line* was generated using the parameters of Abraham and McLauchlan (1962), the *dot-dot-dashed line* using the parameters of Deber et al. (1971), the *dashed line* using the parameters of Kopple et al. (1973), the *dotted line* using the parameters of de Marco et al. (1978), and the *dot-dashed line* using the parameters of Pérez et al. (2001)

**Fig. 1** Fragment H–C–C–H bearing *three* non-hydrogen substituents as found in the L-amino acids suitable for application of the generalised Karplus relation



**Table 1** Karplus relation parameters $a$, $b$, and $c$ from the literature

| Source | Molecule ($\theta_\beta$ determination method) | $a$ | $b$ | $c$ |
|---|---|---|---|---|
| Abraham and McLauchlan (1962) | Hydroxy-L-proline (theoretical) | 12.1 | −1.6 | 0 |
| Deber et al. (1971) | Cyclo(tri-L-prolyl) and derivatives (simulation) | 9.5 | −1.0 | 1.4 |
| Kopple et al. (1973) | Several molecules (X-ray and theoretical) | 9.4 | −1.4 | 1.6 |
| de Marco et al. (1978) | $\chi_1$–$\chi_3$ dihedral angles of ornityl residues in a cyclohexapeptide (X-ray) | 9.5 | −1.6 | 1.8 |
| Pérez et al. (2001) | Flavodoxin (self-consistent fitting) | 7.23 | −1.37 | 2.22 |

The molecules for which the $^3J_{\alpha\beta}$-couplings were measured and the methodology of $\theta_\beta$ determination are indicated. All values are in Hz
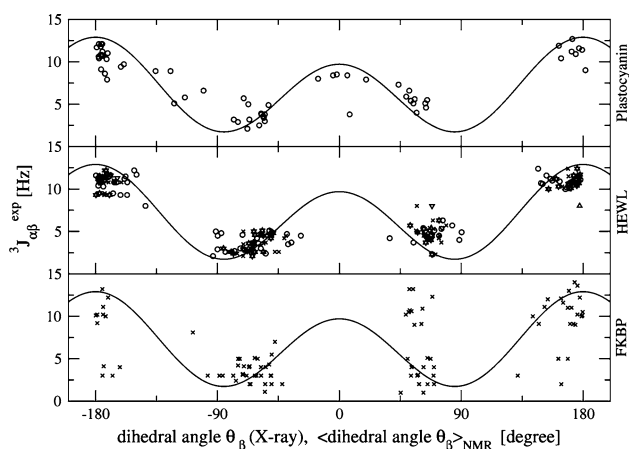
**Fig. 3** Stereospecifically assigned $^3J_{\alpha\beta}$-coupling constants as measured by NMR as a function of the corresponding $\theta_\beta$-angle values in the X-ray or NMR model structures for three proteins, plus the Karplus curve using the parameters from de Marco et al. (1978) (*solid line*). The *top panel* shows data for the NMR model structures 9PCY of Plastocyanin (*circles*). Data for the X-ray structures 1AKI (*crosses*), 193L A (*triangles up*), 193L B (*triangles down*), and the NMR model structures 1E8L (*circles*) of HEWL are shown in the *middle panel*. In the *bottom panel*, data for the X-ray structure 1FKF (*crosses*) of FKBP is plotted

domain of dihedral-angle values. This situation is illustrated in Fig. 3, which shows the stereospecifically assigned $^3J_{\alpha\beta}$-values as obtained from NMR experiments in solution for three different proteins and the corresponding $\theta_\beta$-angles in the X-ray crystal or NMR model structures of each protein.

An alternative procedure to obtain Karplus parameters that avoids the use of crystal data, which are characterised by low atom mobility and a particular environment, is to use structural data from molecular dynamics (MD) simulations of proteins in aqueous solution (Brüschweiler and Case 1994; Lindorff-Larsen et al. 2005; Vögeli et al. 2007). Least-squares fitting of the calculated $^3J$-couplings averaged over an ensemble or trajectory of structures $\langle ^3J^{calc}\rangle$ to the measured couplings $^3J^{exp}$ results in values of $a$, $b$, and $c$ optimised for that particular combination of NMR data and protein structures. Given a high accuracy protein force field and sufficient conformational sampling, such a procedure may lead to a more accurate set of Karplus parameters than the currently available ones. The question remains, however, as to how robust such fitted parameters are.

Whilst several groups (Brüschweiler and Case 1994; Lindorff-Larsen et al. 2005; Vögeli et al. 2007) have optimised the Karplus parameters for backbone or side-chain $^3J$-couplings using MD simulations, none of them made use of $^3J_{\alpha\beta}$-couplings. One possible reason may be the limited amount of data available, as 15 out of the 20 types of amino acids naturally occurring in proteins have two $H_\beta$ atoms, meaning that stereospecific assignment is required. Additionally, the measurement of $^3J_{\alpha\beta}$-coupling constants

becomes more difficult with increasing molecule size due to resonance overlaps and larger line-widths in the spectra. However, the $^3J_{H_N H_\alpha}$-coupling depends on the $\varphi$-angle, which is in turn correlated to the $\psi$-angle of the previous residue in a protein via the peptide bond. Since a variation of the orientation of the peptide plane is easily obtained without changing the spatial fold of the polypeptide backbone as long as the value of the sum $\psi + \varphi$ is constant, the $^3J_{H_N H_\alpha}$-couplings do not unambiguously determine the fold and are therefore less useful for protein structure determination. For these reasons, we shall not consider $^3J_{H_N H_\alpha}$-couplings, and instead concentrate on $^3J_{\alpha\beta}$-couplings, which are related to $\chi_1$-angle value distributions.

We investigate whether the agreement between calculated $^3J_{\alpha\beta}$-couplings $^3J_{\alpha\beta}^{calc}$ and experimentally measured couplings $^3J_{\alpha\beta}^{exp}$ in proteins can be improved either by using the generalised Karplus relation or by fitting calculated $\langle ^3J_{\alpha\beta}^{calc}\rangle$-values to measured $^3J_{\alpha\beta}^{exp}$-values using conformational ensembles of proteins generated by MD simulation or X-ray or NMR model structures to find optimal values for the parameters $a$, $b$, and $c$ of the standard Karplus relation. We use three proteins, Plastocyanin (Moore et al. 1991), hen egg white Lysozyme (HEWL) (Smith et al. 1991), and FK506 binding protein (FKBP) (Xu et al. 1992), for which measured, stereospecifically assigned $^3J_{\alpha\beta}$-couplings are available, as test proteins and for calibration of the Karplus parameters.

## Method

Generalised Karplus relation

For a fragment H–C–C–H in which each of the C atoms carries three substituents, the generalised Karplus relation takes the form (Haasnoot et al. 1979, 1981; Mádi et al. 1990)

$$^3J_{HH}(\theta) = a_1 \cos^2(\theta) + a_2 \cos(\theta) + a_3$$
$$+ \sum_{\substack{=1 \\ \text{substituents}}}^{6} \Delta x_i'\left(a_4 + a_5 \cos^2(\xi_i \theta + a_6 |\Delta x_i'|)\right),$$

(4)

in which $\theta$ is the H–C–C–H dihedral angle (IUPAC convention of 1970), $\Delta x_i'$ are the effective electronegativity differences between the substituent atoms at the two C-atoms and an H-atom as given by the expression

$$\Delta x_i' = \Delta x_i - a_7 \sum_{\substack{=1 \\ \text{substituents}}}^{3} \Delta x_k,$$

(5)

in which $\Delta x_i$ is the electronegativity difference between the substituent $i$ on the C-atom and an H-atom, and the $\Delta x_k$ are

these quantities for the atoms $k$ bound to the substituent $i$, representing the secondary substituent effect. Table II of Huggins (1953) gives

$$\Delta x_H = x_H - x_H = 2.20 - 2.20 = 0.00$$
$$\Delta x_C = x_C - x_H = \Delta x_S = x_S - x_H = 2.60 - 2.20 = 0.40$$
$$\Delta x_N = x_N - x_H = 3.05 - 2.20 = 0.85$$
$$\Delta x_O = x_O - x_H = 3.50 - 2.20 = 1.30.$$
(6)

The quantity $\xi_i$ depends on the orientation of the substituent $i$ with respect to its geminal coupled proton. Since in an L-amino acid fragment as in Fig. 1 the substituents X = N and Y = C on the $C_\alpha$ atom are the same for all residues, only substituent Z on the $C_\beta$ atom varies, which is $C_\gamma$ for Arg, Asn, Asp, Glu, Gln, His, Leu, Lys, Met, Phe, Pro, Trp, and Tyr, $S_\gamma$ for Cys, and $O_\gamma$ for Ser. We have for the pair $(H_\alpha, H_{\beta_2})$ the values $\xi_N = +1$, $\xi_C = -1$, $\xi_Z = +1$, and $\delta_{\beta_2} = -120°$, and for the pair $(H_\alpha, H_{\beta_3})$ the values $\xi_N = +1$, $\xi_C = -1$, $\xi_Z = -1$, and $\delta_{\beta_3} = 0°$, with $Z = C_\gamma$, $O_\gamma$, or $S_\gamma$ and $\chi_1 = N-C_\alpha-C_\beta-Z$, with $\delta_\beta$ defined by Eq. 3 and $\theta = \theta_\beta = H_\alpha-C_\alpha-C_\beta-H_{\beta_2/\beta_3}$. The coefficients $a_i$ are (Mádi et al. 1990) $a_1 = 13.22$ Hz, $a_2 = -0.99$ Hz, $a_3 = 0$ Hz, $a_4 = 0.87$ Hz, $a_5 = -2.46$ Hz, $a_6 = 19.9$, and $a_7 = 0$. Because $a_7 = 0$, $\Delta x_i' = \Delta x_i$ and so

$$
\begin{aligned}
^3J_{HH}(\theta) = {} & a_1 \cos^2(\theta) + a_2 \cos(\theta) \\
& + \Delta x_N \left( a_4 + a_5 \cos^2(\xi_N \theta + a_6 |\Delta x_N|) \right) \\
& + \Delta x_C \left( a_4 + a_5 \cos^2(\xi_C \theta + a_6 |\Delta x_C|) \right) \\
& + \Delta x_Z \left( a_4 + a_5 \cos^2(\xi_Z \theta + a_6 |\Delta x_Z|) \right).
\end{aligned}
$$
(7)

Considering the $\beta_2/\beta_3$ protons and $Z = C_\gamma/S_\gamma/O_\gamma$, but with the simplification $\Delta x_C = \Delta x_S$, this yields four different expressions for $^3J_{\alpha\beta}(\theta_\beta)$. The four corresponding generalised Karplus curves $^3J_{\alpha\beta}(\theta_\beta)$, two for Ser (Z = O) and two for the other 14 residues with two $\beta$-protons, are displayed in Fig. 4.

### Determination of the parameters of the Karplus relation

As an alternative to using the generalised Karplus relation of Eq. 4 we may optimise the parameters $a$, $b$, and $c$ of the standard Karplus relation in Eq. 1 by fitting MD trajectory averaged $^3J_{\alpha\beta}$-couplings $\langle ^3J_{\alpha\beta}^{calc} \rangle$ to the corresponding experimental values $^3J_{\alpha\beta}^{exp}$ (Brüschweiler and Case 1994; Lindorff-Larsen et al. 2005; Vögeli et al. 2007). Using ensemble-averaged values

$$\langle \cos\theta_{\beta_i} \rangle = \langle \cos\chi_i \rangle \cos\delta_{\beta_i} - \langle \sin\chi_i \rangle \sin\delta_{\beta_i}$$
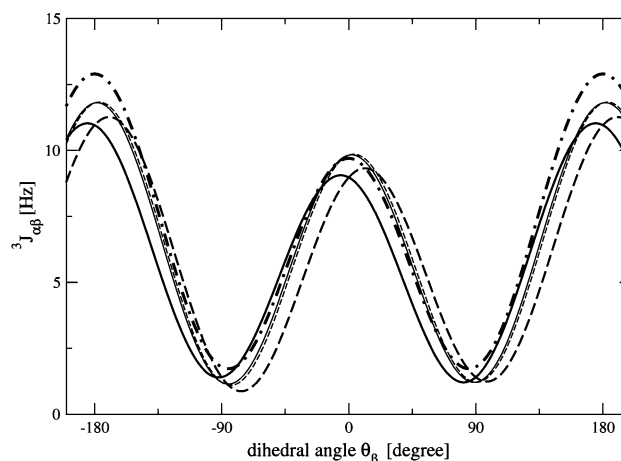(8)

and



**Fig. 4** The four generalised Karplus relations for different substituents Z (O: *thick lines*, C/S: *thin lines*) and different $H_\beta$ types ($H_{\beta_2}$: *dashed lines*, $H_{\beta_3}$: *solid lines*), and the curve obtained using the standard Karplus relation with the parameters of de Marco et al. (1978) (*dash-dotted*)

$$
\begin{aligned}
\langle \cos^2\theta_{\beta_i} \rangle = {} & \langle \cos^2\chi_i \rangle \cos^2\delta_{\beta_i} - 2\langle \cos\chi_i \sin\chi_i \rangle \cos\delta_{\beta_i} \sin\delta_{\beta_i} \\
& + \langle \sin^2\chi_i \rangle \sin^2\delta_{\beta_i}
\end{aligned}
$$
(9)

obtained from an MD trajectory of a particular $H_\alpha$, $H_\beta$ torsional angle $\theta_{\beta_i}$ in the Karplus relation in Eq. 1, ensemble-averaged values of $\langle ^3J_{\alpha\beta}^{calc} \rangle$ can be obtained. By using all or a particular subset, e.g. $^3J_{\alpha\beta}$, $^3J_{\alpha\beta_2}$, or $^3J_{\alpha\beta_3}$, of the $N_J$ experimental values $^3J_{\alpha\beta}^{exp}$ measured for a protein, optimal values for $a$, $b$, and $c$ can be obtained by least-squares fitting of $\langle ^3J_{\alpha\beta}^{calc} \rangle$-values to the corresponding $^3J_{\alpha\beta}^{exp}$-values. In doing so the quantity

$$Q^2 = \frac{1}{N_J} \sum_{i=1}^{N_J} \left( a\langle \cos^2\theta_{\beta_i} \rangle + b\langle \cos\theta_{\beta_i} \rangle + c - ^3J_i^{exp} \right)^2$$
(10)

is minimised with respect to variation of the parameters $a$, $b$, and $c$. Their values follow from the equations $\partial Q/\partial a = \partial Q/\partial b = \partial Q/\partial c = 0$, or

$$\sum_{i=1}^{N_J} \left( a\langle \cos^2\theta_{\beta_i} \rangle + b\langle \cos\theta_{\beta_i} \rangle + c - ^3J_i^{exp} \right)\langle \cos^2\theta_{\beta_i} \rangle = 0$$
(11)

$$\sum_{i=1}^{N_J} \left( a\langle \cos^2\theta_{\beta_i} \rangle + b\langle \cos\theta_{\beta_i} \rangle + c - ^3J_i^{exp} \right)\langle \cos\theta_{\beta_i} \rangle = 0$$
(12)

$$\sum_{i=1}^{N_J} \left( a\langle \cos^2\theta_{\beta_i} \rangle + b\langle \cos\theta_{\beta_i} \rangle + c - ^3J_i^{exp} \right) = 0,$$
(13)

which can be solved using Cramer's rule.

Analysis of the structural and $^3J_{\alpha\beta}$-coupling data

For each of the three proteins we use three different subsets of $^3J_{\alpha\beta}$-couplings, i.e.

1. the stereospecifically assigned $^3J_{\alpha\beta_2}$ and $^3J_{\alpha\beta_3}$ for the side chains with two stereospecifically assigned $H_\beta$ protons,
2. the $^3J_{\alpha\beta}$ for the side chains with one $H_\beta$ proton (Ile, Thr and Val), and
3. the non-stereospecifically assigned $^3J_{\alpha\beta_2}$ and $^3J_{\alpha\beta_3}$.

Two types of structural data sets for the three proteins were used: (i) X-ray or NMR model structures, and (ii) trajectories of protein structures obtained from MD simulations of the proteins in aqueous solution. The simulations were carried out using the GROMOS (Christen et al. 2005; Schmid et al. 2011a) software and different GROMOS biomolecular force fields, namely the force field parameter sets 45A3 of the year 2001 (Schuler et al. 2001), 53A6 of 2004 (Oostenbrink et al. 2004), and 54A7 of 2011 (Schmid et al. 2011b), see Table 2. The non-bonded interaction parameters of the 45A3 force field were obtained by fitting the heat of vaporisation, density and solvation free energy in water and in cyclohexane for a set of compounds representing apolar side chains in proteins. In the 53A6 force field this set was extended to compounds representing polar side chains in proteins. The 54A7 force field contains a

slight modification of protein backbone non-bonded and torsional-angle parameters compared to 53A6. The corresponding force field parameter sets for simulations of proteins in vacuo are denoted as 45B3, 53B6 and 54B7. The X-ray and NMR model structures were taken from the Protein Data Bank (PDB, Berman et al. 2000): 9PCY (Moore et al. 1991) (16 NMR model structures) for Plastocyanin, 1AKI (Artymiuk et al. 1982; Carter et al. 1997) and 193L (Vaney et al. 1996) (X-ray structures) and 1E8L (Schwalbe et al. 2001) (50 NMR model structures) for HEWL, and 1FKF (Xu et al. 1992) for FKBP. The setups of the MD simulations are described in earlier studies of Plastocyanin (Steiner and van Gunsteren 2012), HEWL (Schmid et al. 2011b), and ascomycin bound to FKBP (Allison and van Gunsteren 2009).

For the evaluation of the generalised Karplus relation in Tables 3, 4, 5, 6, 7, S1, S2, and S3 simulation trajectories of lengths 1 ns for Plastocyanin, 20 ns for HEWL, and 8 ns for FKBP were used. $\langle^3J_{\alpha\beta}^{calc}\rangle$-couplings and $Q$-values, defined as the square-root of the $Q^2$ obtained from Eq. 10, were calculated for the set of stereospecifically assigned $^3J_{\alpha\beta_2}$- and $^3J_{\alpha\beta_3}$-coupling constants (subset 1). The same trajectories were used to calculate the $Q$-values for the least-squares fitted Karplus relations in Tables 8, 11, S4, and S5. The values of $a$, $b$, and $c$ obtained from least-squares fitting of subsets 1 and 2 of the $^3J_{\alpha\beta}^{exp}$ for one protein were used to determine $Q$-values for subsets 1 and 2, or 1,

**Table 2** Proteins and structure sets investigated

| | Plastocyanin Steiner and van Gunsteren (2012) | HEWL Schmid et al. (2011b) | FKBP/asc Allison and van Gunsteren (2009) |
|---|---|---|---|
| Number of residues | 99 | 129 | 107 |
| PDB code of X-ray (NMR) structures | (16 9PCY) | 1AKI, 193L, (50 1E8L) | 1FKF |
| Number of water molecules | 3553 | 14365/14355/14378 | 6285 |
| Box type, length (nm) | t, 6.26 | r, 7.72 | r, 5.94 |
| Force field | 45B3 (Schuler et al. 2001), 53A6 (Oostenbrink et al. 2004) | 54A7 (Schmid et al. 2011b), 54B7 (Schmid et al. 2011b), 53A6 (Oostenbrink et al. 2004), 45A3 (Schuler et al. 2001) | 45A3 (Schuler et al. 2001), 45B3 (Schuler et al. 2001) |
| Initial structure | 9PCY model 16 | 1AKI | 1FKF |
| Temperature (K) | 300 | 300 | 303 |
| Simulation length (ns) | 1 | 20 | 8 |
| Number of $^3J_{\alpha\beta}$-couplings total | 108 | 100 | 94 |
| Number of entries in subset 1 | 42 | 46 | 37+37 |
| Number of entries in subset 2 | 20 | 14 | 20 |
| Number of entries in subset 3 | 46 | 40 | 0 |

The structure sets are either molecular dynamics simulation (MD) trajectories from a simulation in a vacuum or water environment with a particular force field, or experimental X-ray or NMR model structures from the PDB (Berman et al. 2000). For MD simulations in water, a rectangular (r) or truncated octahedron (t) box was used. The number of NMR model structures is indicated before the PDB code name of the NMR structure set. Subset 1 contains the measured, stereospecifically assigned $^3J_{\alpha\beta_{2/3}}$-values, subset 2 the measured, assigned $^3J_{\alpha\beta}$-values of amino acids with only one $H_\beta$, and subset 3 contains the measured, non-stereospecifically assigned $^3J_{\alpha\beta_{2/3}}$-values

**Table 3** $^3J_{\alpha\beta}$-coupling constants of subset 1 measured experimentally $^3J_{\alpha\beta}^{exp}$ (Moore et al. 1991) and the average and rmsd of the $^3J_{\alpha\beta}^{calc}$ calculated using the standard Karplus relation with the parameters of de Marco et al. (1978) and using the generalised Karplus relation from the simulation of Plastocyanin in water (53A6) and the corresponding averaged dihedral-angle value $\langle\theta_\beta\rangle$

| Residue | Proton | $^3J_{\alpha\beta}^{exp}$ | $\langle\theta_\beta\rangle$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{DeMarco}$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{genKarplus}$ |
|---|---|---|---|---|---|
| 4 LEU | $H_{\beta_2}$ | 11.7 | 158 ± 31 | 10.7 ± 3.5 | 9.6 ± 3.3 |
| 4 LEU | $H_{\beta_3}$ | 2.5 | 278 ± 31 | 3.8 ± 2.9 | 3.2 ± 2.8 |
| 7 SER | $H_{\beta_2}$ | 5.0 | 295 ± 11 | 3.1 ± 1.1 | 1.6 ± 0.8 |
| 7 SER | $H_{\beta_3}$ | 5.4 | 55 ± 11 | 4.1 ± 1.4 | 3.0 ± 1.2 |
| 11 SER | $H_{\beta_2}$ | 4.6 | 221 ± 75 | 7.0 ± 4.8 | 5.4 ± 4.3 |
| 11 SER | $H_{\beta_3}$ | 9.1 | 341 ± 75 | 4.3 ± 2.8 | 3.7 ± 2.3 |
| 12 LEU | $H_{\beta_2}$ | 12.1 | 178 ± 12 | 12.5 ± 1.0 | 11.3 ± 1.1 |
| 12 LEU | $H_{\beta_3}$ | 3.8 | 298 ± 12 | 3.4 ± 1.2 | 2.9 ± 1.2 |
| 14 PHE | $H_{\beta_2}$ | 11.9 | 169 ± 9 | 12.3 ± 0.7 | 11.0 ± 0.8 |
| 14 PHE | $H_{\beta_3}$ | 3.2 | 289 ± 9 | 2.5 ± 0.7 | 2.0 ± 0.7 |
| 19 PHE | $H_{\beta_2}$ | 3.0 | 299 ± 7 | 3.4 ± 0.9 | 2.7 ± 0.9 |
| 19 PHE | $H_{\beta_3}$ | 5.5 | 59 ± 7 | 3.5 ± 0.9 | 3.5 ± 1.0 |
| 22 PRO | $H_{\beta_2}$ | 5.1 | 226 ± 24 | 7.8 ± 3.7 | 7.3 ± 3.6 |
| 22 PRO | $H_{\beta_3}$ | 8.5 | 346 ± 24 | 7.8 ± 0.9 | 7.7 ± 1.0 |
| 37 HISB | $H_{\beta_2}$ | 11.8 | 166 ± 7 | 12.1 ± 0.7 | 10.8 ± 0.8 |
| 37 HISB | $H_{\beta_3}$ | 3.6 | 286 ± 7 | 2.2 ± 0.5 | 1.7 ± 0.5 |
| 42 ASP | $H_{\beta_2}$ | 4.0 | 72 ± 8 | 2.4 ± 0.6 | 2.3 ± 0.8 |
| 42 ASP | $H_{\beta_3}$ | 11.6 | 192 ± 8 | 12.2 ± 0.7 | 11.2 ± 0.6 |
| 43 GLU | $H_{\beta_2}$ | 5.7 | 299 ± 11 | 3.5 ± 1.3 | 2.8 ± 1.3 |
| 43 GLU | $H_{\beta_3}$ | 5.9 | 59 ± 11 | 3.6 ± 1.3 | 3.6 ± 1.4 |
| 47 PRO | $H_{\beta_2}$ | 8.9 | 233 ± 27 | 6.9 ± 4.0 | 6.3 ± 4.0 |
| 47 PRO | $H_{\beta_3}$ | 8.4 | 353 ± 27 | 7.8 ± 1.0 | 7.7 ± 1.1 |
| 51 ASP | $H_{\beta_2}$ | 5.1 | 64 ± 10 | 3.0 ± 1.1 | 3.1 ± 1.3 |
| 51 ASP | $H_{\beta_3}$ | 10.9 | 184 ± 10 | 12.5 ± 0.5 | 11.5 ± 0.4 |
| 56 SER | $H_{\beta_2}$ | 10.6 | 171 ± 11 | 12.4 ± 1.2 | 10.1 ± 1.3 |
| 56 SER | $H_{\beta_3}$ | 3.9 | 291 ± 11 | 2.8 ± 1.0 | 3.0 ± 1.0 |
| 58 PRO | $H_{\beta_2}$ | 8.9 | 227 ± 25 | 7.7 ± 3.6 | 7.1 ± 3.6 |
| 58 PRO | $H_{\beta_3}$ | 8.0 | 347 ± 25 | 7.8 ± 0.9 | 7.8 ± 1.0 |
| 63 LEU | $H_{\beta_2}$ | 12.1 | 157 ± 38 | 10.4 ± 4.2 | 9.4 ± 4.0 |
| 63 LEU | $H_{\beta_3}$ | 3.8 | 277 ± 38 | 4.7 ± 3.3 | 4.1 ± 3.1 |
| 66 PRO | $H_{\beta_2}$ | 6.6 | 228 ± 25 | 7.4 ± 3.9 | 6.9 ± 3.8 |
| 66 PRO | $H_{\beta_3}$ | 7.9 | 348 ± 25 | 7.8 ± 0.9 | 7.8 ± 1.1 |
| 70 TYR | $H_{\beta_2}$ | 6.6 | 59 ± 8 | 3.5 ± 1.0 | 3.8 ± 1.2 |
| 70 TYR | $H_{\beta_3}$ | 11.2 | 179 ± 8 | 12.7 ± 0.4 | 11.6 ± 0.4 |
| 74 LEU | $H_{\beta_2}$ | 12.1 | 158 ± 18 | 11.0 ± 2.4 | 9.6 ± 2.4 |
| 74 LEU | $H_{\beta_3}$ | 3.4 | 278 ± 18 | 2.6 ± 1.4 | 2.1 ± 1.4 |
| 80 TYR | $H_{\beta_2}$ | 12.7 | 159 ± 7 | 11.5 ± 0.9 | 10.1 ± 1.0 |
| 80 TYR | $H_{\beta_3}$ | 2.1 | 279 ± 7 | 2.0 ± 0.3 | 1.4 ± 0.3 |
| 84 CYS | $H_{\beta_2}$ | 7.3 | 66 ± 8 | 2.8 ± 0.8 | 2.9 ± 1.0 |
| 84 CYS | $H_{\beta_3}$ | 10.4 | 186 ± 8 | 12.6 ± 0.4 | 11.5 ± 0.3 |
| 86 PRO | $H_{\beta_2}$ | 5.8 | 216 ± 20 | 9.2 ± 3.1 | 8.7 ± 3.0 |
| 86 PRO | $H_{\beta_3}$ | 8.4 | 336 ± 20 | 7.4 ± 0.9 | 7.2 ± 1.1 |
| $Q$: | | | | 1.8 | 2.0 |

The $Q$-values quantifying the agreement between $^3J_{\alpha\beta}^{exp}$ and each set of $\langle^3J_{\alpha\beta}^{calc}\rangle$ are given at the bottom. All $^3J$-couplings and $Q$-values are given in Hz and the $\langle\theta_\beta\rangle$-angle values are given in degrees

2, and 3 of the $^3J_{\alpha\beta}^{exp}$ in all other proteins as a jack-knife test. The stereospecific assignment of subset 3 giving the lowest $Q$-value was chosen by assigning the lower of the two $^3J_{\alpha\beta}^{exp}$-values given for a residue to the lower calculated $\langle^3J_{\alpha\beta}^{calc}\rangle$-value of this residue. The degree of convergence of the trajectory averages of quantities such as $^3J_{\alpha\beta}$, $\theta$, $\cos(\theta)$,

**Table 4** $^3J_{\alpha\beta}$-coupling constants of subset 1 measured experimentally $^3J_{\alpha\beta}^{exp}$ (Smith et al. 1991) and the average and rmsd of the $^3J_{\alpha\beta}^{calc}$ calculated using the standard Karplus relation with the parameters of de Marco et al. (1978) and using the generalised Karplus relation from the simulation of HEWL in water (45A3) and the corresponding averaged dihedral-angle value $\langle\theta_\beta\rangle$

| Residue | Proton | $^3J_{\alpha\beta}^{exp}$ | $\langle\theta_\beta\rangle$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{DeMarco}$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{genKarplus}$ |
|---|---|---|---|---|---|
| 3 PHE | $H_{\beta_2}$ | 10.0 | 166 ± 19 | 11.7 ± 1.9 | 10.4 ± 1.9 |
| 3 PHE | $H_{\beta_3}$ | 3.0 | 286 ± 19 | 2.7 ± 1.3 | 2.1 ± 1.3 |
| 6 CYS1 | $H_{\beta_2}$ | 11.5 | 173 ± 10 | 12.4 ± 0.6 | 11.2 ± 0.7 |
| 6 CYS1 | $H_{\beta_3}$ | 3.5 | 293 ± 10 | 2.8 ± 0.9 | 2.3 ± 1.0 |
| 15 HISB | $H_{\beta_2}$ | 11.2 | 163 ± 27 | 11.3 ± 3.0 | 10.2 ± 2.8 |
| 15 HISB | $H_{\beta_3}$ | 2.6 | 283 ± 27 | 3.5 ± 2.6 | 2.9 ± 2.5 |
| 18 ASP | $H_{\beta_2}$ | 4.2 | -37 ± 52 | 3.2 ± 1.1 | 2.6 ± 1.2 |
| 18 ASP | $H_{\beta_3}$ | 11.0 | 82 ± 52 | 5.6 ± 3.5 | 5.4 ± 3.2 |
| 20 TYR | $H_{\beta_2}$ | 2.3 | 119 ± 37 | 6.3 ± 4.3 | 5.4 ± 3.9 |
| 20 TYR | $H_{\beta_3}$ | 11.7 | 239 ± 37 | 6.5 ± 4.3 | 5.8 ± 4.2 |
| 23 TYR | $H_{\beta_2}$ | 10.9 | 179 ± 10 | 12.6 ± 0.5 | 11.5 ± 0.6 |
| 23 TYR | $H_{\beta_3}$ | 2.7 | 299 ± 10 | 3.5 ± 1.2 | 3.0 ± 1.2 |
| 27 ASN | $H_{\beta_2}$ | 10.3 | 180 ± 11 | 12.6 ± 0.8 | 11.5 ± 0.8 |
| 27 ASN | $H_{\beta_3}$ | 2.4 | 300 ± 11 | 3.6 ± 1.2 | 3.1 ± 1.2 |
| 30 CYS1 | $H_{\beta_2}$ | 5.3 | 92 ± 17 | 2.7 ± 1.3 | 2.0 ± 1.2 |
| 30 CYS1 | $H_{\beta_3}$ | 10.8 | 212 ± 17 | 9.5 ± 2.4 | 8.8 ± 2.3 |
| 34 PHE | $H_{\beta_2}$ | 10.7 | 162 ± 14 | 11.6 ± 1.7 | 10.2 ± 1.7 |
| 34 PHE | $H_{\beta_3}$ | 5.0 | 282 ± 14 | 2.4 ± 1.1 | 1.8 ± 1.1 |
| 39 ASN | $H_{\beta_2}$ | 4.5 | 72 ± 16 | 2.8 ± 1.5 | 2.7 ± 1.5 |
| 39 ASN | $H_{\beta_3}$ | 10.8 | 192 ± 16 | 11.9 ± 1.7 | 11.0 ± 1.6 |
| 46 ASN | $H_{\beta_2}$ | 11.2 | 179 ± 10 | 12.6 ± 0.7 | 11.5 ± 0.7 |
| 46 ASN | $H_{\beta_3}$ | 4.7 | 299 ± 10 | 3.5 ± 1.1 | 3.0 ± 1.2 |
| 48 ASP | $H_{\beta_2}$ | 2.6 | 305 ± 9 | 4.2 ± 1.2 | 3.4 ± 1.2 |
| 48 ASP | $H_{\beta_3}$ | 3.7 | 65 ± 9 | 3.0 ± 1.0 | 2.9 ± 1.2 |
| 52 ASP | $H_{\beta_2}$ | 11.6 | 177 ± 10 | 12.5 ± 0.6 | 11.4 ± 0.7 |
| 52 ASP | $H_{\beta_3}$ | 3.6 | 297 ± 10 | 3.3 ± 1.1 | 2.7 ± 1.1 |
| 53 TYR | $H_{\beta_2}$ | 10.4 | 163 ± 9 | 11.8 ± 1.0 | 10.4 ± 1.1 |
| 53 TYR | $H_{\beta_3}$ | 3.0 | 283 ± 9 | 2.2 ± 0.5 | 1.6 ± 0.5 |
| 59 ASN | $H_{\beta_2}$ | 5.4 | 74 ± 7 | 2.2 ± 0.4 | 2.0 ± 0.6 |
| 59 ASN | $H_{\beta_3}$ | 11.3 | 194 ± 7 | 12.1 ± 0.7 | 11.1 ± 0.6 |
| 61 ARG | $H_{\beta_2}$ | 5.7 | 104 ± 35 | 4.8 ± 3.9 | 4.1 ± 3.5 |
| 61 ARG | $H_{\beta_3}$ | 10.8 | 224 ± 35 | 8.2 ± 4.2 | 7.5 ± 4.0 |
| 66 ASP | $H_{\beta_2}$ | 5.1 | 59 ± 28 | 3.0 ± 1.0 | 3.0 ± 1.1 |
| 66 ASP | $H_{\beta_3}$ | 4.5 | 179 ± 28 | 11.9 ± 2.4 | 11.0 ± 2.2 |
| 75 LEU | $H_{\beta_2}$ | 12.4 | 113 ± 46 | 6.4 ± 4.7 | 5.8 ± 4.2 |
| 75 LEU | $H_{\beta_3}$ | 2.1 | 233 ± 46 | 7.8 ± 4.7 | 7.0 ± 4.5 |
| 87 ASP | $H_{\beta_2}$ | 5.1 | 69 ± 25 | 3.3 ± 2.0 | 3.3 ± 2.0 |
| 87 ASP | $H_{\beta_3}$ | 11.5 | 189 ± 25 | 11.8 ± 2.1 | 10.8 ± 1.9 |
| 94 CYS2 | $H_{\beta_2}$ | 4.0 | 67 ± 9 | 2.8 ± 0.8 | 2.9 ± 1.0 |
| 94 CYS2 | $H_{\beta_3}$ | 12.2 | 187 ± 9 | 12.5 ± 0.6 | 11.5 ± 0.6 |
| 119 ASP | $H_{\beta_2}$ | 4.9 | 69 ± 15 | 2.8 ± 1.3 | 2.8 ± 1.4 |
| 119 ASP | $H_{\beta_3}$ | 11.7 | 189 ± 15 | 12.2 ± 1.2 | 11.2 ± 1.1 |
| 123 TRP | $H_{\beta_2}$ | 10.6 | 162 ± 13 | 11.6 ± 1.4 | 10.2 ± 1.4 |
| 123 TRP | $H_{\beta_3}$ | 2.9 | 282 ± 13 | 2.3 ± 0.9 | 1.8 ± 0.9 |
| 127 CYS2 | $H_{\beta_2}$ | 11.6 | 182 ± 10 | 12.6 ± 0.6 | 11.5 ± 0.6 |
| 127 CYS2 | $H_{\beta_3}$ | 4.8 | 302 ± 10 | 3.9 ± 1.2 | 3.4 ± 1.2 |
| $Q$: | | | | 2.5 | 2.4 |

The $Q$-values quantifying the agreement between $^3J_{\alpha\beta}^{exp}$ and each set of $\langle^3J_{\alpha\beta}^{calc}\rangle$ are given at the bottom. All $^3J$-couplings and $Q$-values are given in Hz and the $\langle\theta_\beta\rangle$-angle values are given in degrees

**Table 5** $^3J_{\alpha\beta}$-coupling constants of subset 1 measured experimentally $^3J_{\alpha\beta}^{exp}$ (Smith et al. 1991) and the average and rmsd of the $^3J_{\alpha\beta}^{calc}$ calculated using the standard Karplus relation with the parameters of de Marco et al. (1978) and using the generalised Karplus relation from the simulation of HEWL in water (53A6) and the corresponding averaged dihedral-angle value $\langle\theta_\beta\rangle$

| Residue | Proton | $^3J_{\alpha\beta}^{exp}$ | $\langle\theta_\beta\rangle$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{\text{DeMarco}}$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{\text{genKarplus}}$ |
|---|---|---|---|---|---|
| 3 PHE | $H_{\beta_2}$ | 10.0 | $158 \pm 16$ | $10.9 \pm 1.9$ | $9.6 \pm 2.0$ |
| 3 PHE | $H_{\beta_3}$ | 3.0 | $278 \pm 16$ | $2.5 \pm 0.9$ | $1.9 \pm 1.0$ |
| 6 CYS1 | $H_{\beta_2}$ | 11.5 | $177 \pm 25$ | $12.0 \pm 2.3$ | $10.9 \pm 2.2$ |
| 6 CYS1 | $H_{\beta_3}$ | 3.5 | $297 \pm 25$ | $3.5 \pm 1.7$ | $3.0 \pm 1.7$ |
| 15 HISB | $H_{\beta_2}$ | 11.2 | $167 \pm 10$ | $12.1 \pm 0.9$ | $10.8 \pm 1.0$ |
| 15 HISB | $H_{\beta_3}$ | 2.6 | $287 \pm 10$ | $2.5 \pm 0.8$ | $1.9 \pm 0.8$ |
| 18 ASP | $H_{\beta_2}$ | 4.2 | $-13 \pm 66$ | $2.9 \pm 1.0$ | $2.4 \pm 1.0$ |
| 18 ASP | $H_{\beta_3}$ | 11.0 | $106 \pm 66$ | $7.2 \pm 4.1$ | $6.9 \pm 3.6$ |
| 20 TYR | $H_{\beta_2}$ | 2.3 | $107 \pm 40$ | $5.4 \pm 4.5$ | $4.8 \pm 4.0$ |
| 20 TYR | $H_{\beta_3}$ | 11.7 | $227 \pm 40$ | $8.3 \pm 4.6$ | $7.5 \pm 4.5$ |
| 23 TYR | $H_{\beta_2}$ | 10.9 | $178 \pm 11$ | $12.5 \pm 0.6$ | $11.4 \pm 0.7$ |
| 23 TYR | $H_{\beta_3}$ | 2.7 | $298 \pm 11$ | $3.4 \pm 1.1$ | $2.9 \pm 1.2$ |
| 27 ASN | $H_{\beta_2}$ | 10.3 | $175 \pm 14$ | $12.3 \pm 1.1$ | $11.1 \pm 1.2$ |
| 27 ASN | $H_{\beta_3}$ | 2.4 | $295 \pm 14$ | $3.3 \pm 1.3$ | $2.8 \pm 1.3$ |
| 30 CYS1 | $H_{\beta_2}$ | 5.3 | $106 \pm 50$ | $6.9 \pm 4.2$ | $6.4 \pm 3.6$ |
| 30 CYS1 | $H_{\beta_3}$ | 10.8 | $226 \pm 50$ | $8.0 \pm 5.1$ | $7.1 \pm 4.9$ |
| 34 PHE | $H_{\beta_2}$ | 10.7 | $170 \pm 12$ | $12.3 \pm 1.1$ | $11.0 \pm 1.1$ |
| 34 PHE | $H_{\beta_3}$ | 5.0 | $290 \pm 12$ | $2.8 \pm 1.1$ | $2.2 \pm 1.1$ |
| 39 ASN | $H_{\beta_2}$ | 4.5 | $66 \pm 32$ | $2.7 \pm 1.3$ | $2.6 \pm 1.4$ |
| 39 ASN | $H_{\beta_3}$ | 10.8 | $186 \pm 32$ | $11.5 \pm 2.3$ | $10.6 \pm 2.1$ |
| 46 ASN | $H_{\beta_2}$ | 11.2 | $181 \pm 14$ | $12.4 \pm 1.4$ | $11.4 \pm 1.4$ |
| 46 ASN | $H_{\beta_3}$ | 4.7 | $301 \pm 14$ | $3.9 \pm 1.3$ | $3.4 \pm 1.4$ |
| 48 ASP | $H_{\beta_2}$ | 2.6 | $305 \pm 10$ | $4.2 \pm 1.3$ | $3.5 \pm 1.3$ |
| 48 ASP | $H_{\beta_3}$ | 3.7 | $65 \pm 10$ | $3.0 \pm 1.1$ | $2.9 \pm 1.3$ |
| 52 ASP | $H_{\beta_2}$ | 11.6 | $183 \pm 9$ | $12.6 \pm 0.5$ | $11.6 \pm 0.4$ |
| 52 ASP | $H_{\beta_3}$ | 3.6 | $303 \pm 9$ | $3.9 \pm 1.2$ | $3.4 \pm 1.2$ |
| 53 TYR | $H_{\beta_2}$ | 10.4 | $166 \pm 10$ | $12.1 \pm 1.0$ | $10.7 \pm 1.1$ |
| 53 TYR | $H_{\beta_3}$ | 3.0 | $286 \pm 10$ | $2.4 \pm 0.6$ | $1.8 \pm 0.7$ |
| 59 ASN | $H_{\beta_2}$ | 5.4 | $76 \pm 7$ | $2.1 \pm 0.4$ | $1.9 \pm 0.6$ |
| 59 ASN | $H_{\beta_3}$ | 11.3 | $196 \pm 7$ | $11.9 \pm 0.7$ | $11.0 \pm 0.7$ |
| 61 ARG | $H_{\beta_2}$ | 5.7 | $95 \pm 29$ | $3.8 \pm 3.1$ | $3.2 \pm 2.8$ |
| 61 ARG | $H_{\beta_3}$ | 10.8 | $215 \pm 29$ | $9.2 \pm 3.7$ | $8.4 \pm 3.5$ |
| 66 ASP | $H_{\beta_2}$ | 5.1 | $66 \pm 10$ | $2.9 \pm 0.8$ | $2.9 \pm 1.0$ |
| 66 ASP | $H_{\beta_3}$ | 4.5 | $186 \pm 10$ | $12.5 \pm 0.7$ | $11.5 \pm 0.6$ |
| 75 LEU | $H_{\beta_2}$ | 12.4 | $149 \pm 32$ | $10.0 \pm 3.6$ | $8.8 \pm 3.3$ |
| 75 LEU | $H_{\beta_3}$ | 2.1 | $269 \pm 32$ | $3.8 \pm 3.3$ | $3.2 \pm 3.2$ |
| 87 ASP | $H_{\beta_2}$ | 5.1 | $71 \pm 15$ | $2.8 \pm 1.3$ | $2.7 \pm 1.4$ |
| 87 ASP | $H_{\beta_3}$ | 11.5 | $191 \pm 15$ | $11.9 \pm 1.4$ | $11.0 \pm 1.3$ |
| 94 CYS2 | $H_{\beta_2}$ | 4.0 | $74 \pm 12$ | $2.4 \pm 0.7$ | $2.3 \pm 0.9$ |
| 94 CYS2 | $H_{\beta_3}$ | 12.2 | $194 \pm 12$ | $11.9 \pm 1.3$ | $11.0 \pm 1.2$ |
| 119 ASP | $H_{\beta_2}$ | 4.9 | $-42 \pm 164$ | $2.9 \pm 1.3$ | $2.9 \pm 1.4$ |
| 119 ASP | $H_{\beta_3}$ | 11.7 | $77 \pm 164$ | $12.2 \pm 1.3$ | $11.2 \pm 1.2$ |
| 123 TRP | $H_{\beta_2}$ | 10.6 | $171 \pm 12$ | $12.2 \pm 0.9$ | $11.0 \pm 1.1$ |
| 123 TRP | $H_{\beta_3}$ | 2.9 | $291 \pm 12$ | $2.9 \pm 1.1$ | $2.3 \pm 1.1$ |
| 127 CYS2 | $H_{\beta_2}$ | 11.6 | $186 \pm 51$ | $10.7 \pm 3.6$ | $9.7 \pm 3.4$ |
| 127 CYS2 | $H_{\beta_3}$ | 4.8 | $306 \pm 51$ | $3.9 \pm 2.4$ | $3.4 \pm 2.3$ |
| $Q$: | | | | 2.0 | 2.0 |

The $Q$-values quantifying the agreement between $^3J_{\alpha\beta}^{exp}$ and each set of $\langle^3J_{\alpha\beta}^{calc}\rangle$ are given at the bottom. All $^3J$-couplings and $Q$-values are given in Hz and the $\langle\theta_\beta\rangle$-angle values are given in degrees

**Table 6** $^3J_{\alpha\beta}$-coupling constants of subset 1 measured experimentally $^3J_{\alpha\beta}^{exp}$ (Smith et al. 1991) and the average and rmsd of the $^3J_{\alpha\beta}^{calc}$ calculated using the standard Karplus relation with the parameters of de Marco et al. (1978) and using the generalised Karplus relation from the simulation of HEWL in water (54A7) and the corresponding averaged dihedral-angle value $\langle\theta_\beta\rangle$

| Residue | Proton | $^3J_{\alpha\beta}^{exp}$ | $\langle\theta_\beta\rangle$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{DeMarco}$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{genKarplus}$ |
|---|---|---|---|---|---|
| 3 PHE | $H_{\beta_2}$ | 10.0 | 149 ± 23 | 10.0 ± 3.1 | 8.6 ± 2.9 |
| 3 PHE | $H_{\beta_3}$ | 3.0 | 269 ± 23 | 3.0 ± 2.5 | 2.4 ± 2.4 |
| 6 CYS1 | $H_{\beta_2}$ | 11.5 | 165 ± 17 | 11.6 ± 1.5 | 10.3 ± 1.6 |
| 6 CYS1 | $H_{\beta_3}$ | 3.5 | 285 ± 17 | 2.7 ± 1.5 | 2.2 ± 1.5 |
| 15 HISB | $H_{\beta_2}$ | 11.2 | 164 ± 10 | 11.9 ± 1.1 | 10.5 ± 1.2 |
| 15 HISB | $H_{\beta_3}$ | 2.6 | 284 ± 10 | 2.3 ± 0.7 | 1.7 ± 0.7 |
| 18 ASP | $H_{\beta_2}$ | 4.2 | 67 ± 17 | 3.1 ± 1.3 | 3.1 ± 1.4 |
| 18 ASP | $H_{\beta_3}$ | 11.0 | 187 ± 17 | 12.2 ± 1.5 | 11.2 ± 1.4 |
| 20 TYR | $H_{\beta_2}$ | 2.3 | 92 ± 30 | 3.9 ± 3.0 | 3.3 ± 2.6 |
| 20 TYR | $H_{\beta_3}$ | 11.7 | 212 ± 30 | 9.4 ± 3.9 | 8.6 ± 3.7 |
| 23 TYR | $H_{\beta_2}$ | 10.9 | 176 ± 10 | 12.5 ± 0.5 | 11.4 ± 0.6 |
| 23 TYR | $H_{\beta_3}$ | 2.7 | 296 ± 10 | 3.1 ± 1.0 | 2.6 ± 1.1 |
| 27 ASN | $H_{\beta_2}$ | 10.3 | 168 ± 11 | 12.1 ± 1.1 | 10.8 ± 1.2 |
| 27 ASN | $H_{\beta_3}$ | 2.4 | 288 ± 11 | 2.5 ± 0.9 | 2.0 ± 0.9 |
| 30 CYS1 | $H_{\beta_2}$ | 5.3 | 58 ± 13 | 3.8 ± 1.3 | 4.0 ± 1.5 |
| 30 CYS1 | $H_{\beta_3}$ | 10.8 | 178 ± 13 | 12.4 ± 0.9 | 11.3 ± 0.8 |
| 34 PHE | $H_{\beta_2}$ | 10.7 | 146 ± 26 | 9.6 ± 3.3 | 8.3 ± 3.1 |
| 34 PHE | $H_{\beta_3}$ | 5.0 | 266 ± 26 | 3.3 ± 2.8 | 2.7 ± 2.7 |
| 39 ASN | $H_{\beta_2}$ | 4.5 | 72 ± 11 | 2.5 ± 0.9 | 2.4 ± 1.1 |
| 39 ASN | $H_{\beta_3}$ | 10.8 | 192 ± 11 | 12.0 ± 1.0 | 11.1 ± 0.9 |
| 46 ASN | $H_{\beta_2}$ | 11.2 | 114 ± 39 | 4.8 ± 4.5 | 4.0 ± 4.3 |
| 46 ASN | $H_{\beta_3}$ | 4.7 | 234 ± 39 | 8.2 ± 3.1 | 7.5 ± 3.0 |
| 48 ASP | $H_{\beta_2}$ | 2.6 | 306 ± 8 | 4.3 ± 1.1 | 3.5 ± 1.1 |
| 48 ASP | $H_{\beta_3}$ | 3.7 | 66 ± 8 | 2.8 ± 0.9 | 2.7 ± 1.0 |
| 52 ASP | $H_{\beta_2}$ | 11.6 | 178 ± 12 | 12.4 ± 0.8 | 11.3 ± 0.9 |
| 52 ASP | $H_{\beta_3}$ | 3.6 | 298 ± 12 | 3.4 ± 1.2 | 2.9 ± 1.2 |
| 53 TYR | $H_{\beta_2}$ | 10.4 | 168 ± 9 | 12.2 ± 0.7 | 10.9 ± 0.8 |
| 53 TYR | $H_{\beta_3}$ | 3.0 | 288 ± 9 | 2.4 ± 0.6 | 1.9 ± 0.7 |
| 59 ASN | $H_{\beta_2}$ | 5.4 | 73 ± 7 | 2.3 ± 0.5 | 2.2 ± 0.7 |
| 59 ASN | $H_{\beta_3}$ | 11.3 | 193 ± 7 | 12.2 ± 0.6 | 11.3 ± 0.5 |
| 61 ARG | $H_{\beta_2}$ | 5.7 | 98 ± 39 | 4.7 ± 4.0 | 4.2 ± 3.6 |
| 61 ARG | $H_{\beta_3}$ | 10.8 | 218 ± 39 | 9.2 ± 4.3 | 8.4 ± 4.1 |
| 66 ASP | $H_{\beta_2}$ | 5.1 | 65 ± 20 | 2.8 ± 0.9 | 2.8 ± 1.1 |
| 66 ASP | $H_{\beta_3}$ | 4.5 | 185 ± 20 | 12.2 ± 1.6 | 11.2 ± 1.4 |
| 75 LEU | $H_{\beta_2}$ | 12.4 | 168 ± 15 | 12.0 ± 1.7 | 10.7 ± 1.7 |
| 75 LEU | $H_{\beta_3}$ | 2.1 | 288 ± 15 | 2.8 ± 1.2 | 2.2 ± 1.2 |
| 87 ASP | $H_{\beta_2}$ | 5.1 | 60 ± 12 | 3.5 ± 1.3 | 3.7 ± 1.5 |
| 87 ASP | $H_{\beta_3}$ | 11.5 | 180 ± 12 | 12.4 ± 0.7 | 11.3 ± 0.7 |
| 94 CYS2 | $H_{\beta_2}$ | 4.0 | 69 ± 8 | 2.6 ± 0.7 | 2.6 ± 0.9 |
| 94 CYS2 | $H_{\beta_3}$ | 12.2 | 189 ± 8 | 12.4 ± 0.5 | 11.4 ± 0.5 |
| 119 ASP | $H_{\beta_2}$ | 4.9 | 65 ± 11 | 3.0 ± 1.1 | 3.1 ± 1.3 |
| 119 ASP | $H_{\beta_3}$ | 11.7 | 185 ± 11 | 12.4 ± 0.8 | 11.4 ± 0.7 |
| 123 TRP | $H_{\beta_2}$ | 10.6 | 163 ± 16 | 11.5 ± 1.6 | 10.1 ± 1.7 |
| 123 TRP | $H_{\beta_3}$ | 2.9 | 283 ± 16 | 2.6 ± 1.3 | 2.0 ± 1.3 |
| 127 CYS2 | $H_{\beta_2}$ | 11.6 | 66 ± 155 | 11.6 ± 2.5 | 10.4 ± 2.4 |
| 127 CYS2 | $H_{\beta_3}$ | 4.8 | 186 ± 155 | 2.9 ± 1.2 | 2.4 ± 1.3 |
| Q: | | | | 2.0 | 2.0 |

The Q-values quantifying the agreement between $^3J_{\alpha\beta}^{exp}$ and each set of $\langle^3J_{\alpha\beta}^{calc}\rangle$ are given at the bottom. All $^3J$-couplings and Q-values are given in Hz and the $\langle\theta_\beta\rangle$-angle values are given in degrees

**Table 7** $^3J_{\alpha\beta}$-coupling constants of subset 1 measured experimentally $^3J_{\alpha\beta}^{exp}$ (Xu et al. 1992) and the average and rmsd of the $^3J_{\alpha\beta}^{calc}$ calculated using the standard Karplus relation with the parameters of de Marco et al. (1978) and using the generalised Karplus relation from the simulation of FKBP in water (45A3) and the corresponding averaged dihedral-angle value $\langle\theta_\beta\rangle$

| Residue | Proton | $^3J_{\alpha\beta}^{exp}$ | $\langle\theta_\beta\rangle$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{\mathrm{DeMarco}}$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{\mathrm{genKarplus}}$ |
|---|---|---|---|---|---|
| 3 GLN | $H_{\beta_2}$ | 9.2 | $122 \pm 51$ | $7.5 \pm 4.9$ | $6.9 \pm 4.4$ |
| 8 SER | $H_{\beta_3}$ | 2.0 | $221 \pm 57$ | $9.7 \pm 4.3$ | $8.3 \pm 3.3$ |
| 8 SER | $H_{\beta_2}$ | 4.0 | $101 \pm 57$ | $5.7 \pm 4.3$ | $5.7 \pm 3.2$ |
| 11 ASP | $H_{\beta_3}$ | 2.0 | $240 \pm 75$ | $8.9 \pm 4.4$ | $8.1 \pm 4.1$ |
| 11 ASP | $H_{\beta_2}$ | 5.5 | $120 \pm 75$ | $4.9 \pm 4.1$ | $4.4 \pm 3.7$ |
| 13 ARG | $H_{\beta_3}$ | 3.0 | $63 \pm 156$ | $4.6 \pm 3.6$ | $4.1 \pm 3.4$ |
| 13 ARG | $H_{\beta_2}$ | 5.1 | $-56 \pm 156$ | $8.4 \pm 4.6$ | $7.5 \pm 4.3$ |
| 15 PHE | $H_{\beta_3}$ | 2.0 | $277 \pm 35$ | $3.8 \pm 3.2$ | $3.2 \pm 3.1$ |
| 15 PHE | $H_{\beta_2}$ | 9.0 | $157 \pm 35$ | $11.1 \pm 3.2$ | $10.0 \pm 2.8$ |
| 17 LYSH | $H_{\beta_2}$ | 11.0 | $184 \pm 46$ | $11.1 \pm 3.3$ | $2.9 \pm 1.9$ |
| 17 LYSH | $H_{\beta_3}$ | 3.2 | $304 \pm 46$ | $3.4 \pm 1.9$ | $10.0 \pm 3.1$ |
| 20 GLN | $H_{\beta_3}$ | 4.1 | $217 \pm 140$ | $3.9 \pm 3.2$ | $3.3 \pm 3.0$ |
| 26 TYR | $H_{\beta_2}$ | 2.0 | $309 \pm 6$ | $4.7 \pm 0.9$ | $4.0 \pm 0.9$ |
| 26 TYR | $H_{\beta_3}$ | 4.0 | $69 \pm 6$ | $2.5 \pm 0.5$ | $2.3 \pm 0.6$ |
| 29 MET | $H_{\beta_2}$ | 2.0 | $205 \pm 61$ | $9.7 \pm 4.3$ | $8.7 \pm 4.1$ |
| 29 MET | $H_{\beta_3}$ | 2.0 | $325 \pm 61$ | $3.5 \pm 1.9$ | $3.1 \pm 1.9$ |
| 30 LEU | $H_{\beta_3}$ | 4.2 | $257 \pm 40$ | $5.5 \pm 4.1$ | $4.8 \pm 3.9$ |
| 30 LEU | $H_{\beta_2}$ | 11.6 | $137 \pm 40$ | $8.4 \pm 4.9$ | $7.4 \pm 4.5$ |
| 32 ASP | $H_{\beta_2}$ | 3.0 | $198 \pm 105$ | $4.0 \pm 3.2$ | $3.5 \pm 3.0$ |
| 32 ASP | $H_{\beta_3}$ | 4.0 | $318 \pm 105$ | $6.9 \pm 4.2$ | $6.5 \pm 3.8$ |
| 34 LYSH | $H_{\beta_2}$ | 3.0 | $69 \pm 132$ | $7.8 \pm 4.8$ | $7.0 \pm 4.4$ |
| 34 LYSH | $H_{\beta_3}$ | 8.1 | $189 \pm 132$ | $5.9 \pm 4.3$ | $5.3 \pm 4.0$ |
| 36 PHE | $H_{\beta_2}$ | 4.0 | $305 \pm 9$ | $4.1 \pm 1.2$ | $3.4 \pm 1.2$ |
| 36 PHE | $H_{\beta_3}$ | 4.0 | $65 \pm 9$ | $3.0 \pm 0.9$ | $2.9 \pm 1.0$ |
| 37 ASP | $H_{\beta_2}$ | 5.0 | $78 \pm 10$ | $2.2 \pm 0.5$ | $1.9 \pm 0.7$ |
| 37 ASP | $H_{\beta_3}$ | 10.0 | $198 \pm 10$ | $11.5 \pm 1.3$ | $10.7 \pm 1.2$ |
| 38 SER | $H_{\beta_2}$ | 3.0 | $57 \pm 23$ | $3.4 \pm 1.1$ | $4.4 \pm 1.4$ |
| 38 SER | $H_{\beta_3}$ | 9.1 | $177 \pm 23$ | $12.3 \pm 1.7$ | $10.4 \pm 1.5$ |
| 39 SER | $H_{\beta_3}$ | 2.0 | $66 \pm 7$ | $2.8 \pm 0.7$ | $1.9 \pm 0.6$ |
| 39 SER | $H_{\beta_2}$ | 3.0 | $306 \pm 7$ | $4.2 \pm 1.0$ | $2.3 \pm 0.9$ |
| 40 ARG | $H_{\beta_3}$ | 2.0 | $288 \pm 11$ | $2.6 \pm 0.9$ | $2.0 \pm 0.9$ |
| 40 ARG | $H_{\beta_2}$ | 14.0 | $168 \pm 11$ | $12.2 \pm 1.0$ | $10.9 \pm 1.1$ |
| 43 ASN | $H_{\beta_3}$ | 3.1 | $46 \pm 159$ | $4.3 \pm 3.2$ | $3.8 \pm 3.1$ |
| 43 ASN | $H_{\beta_2}$ | 5.0 | $-73 \pm 159$ | $9.5 \pm 4.3$ | $8.5 \pm 4.0$ |
| 46 PHE | $H_{\beta_2}$ | 5.0 | $64 \pm 14$ | $3.2 \pm 1.2$ | $3.3 \pm 1.5$ |
| 46 PHE | $H_{\beta_3}$ | 9.1 | $184 \pm 14$ | $12.3 \pm 1.0$ | $11.3 \pm 1.0$ |
| 47 LYSH | $H_{\beta_3}$ | 5.0 | $25 \pm 166$ | $6.7 \pm 4.3$ | $6.0 \pm 4.1$ |
| 47 LYSH | $H_{\beta_2}$ | 10.1 | $-94 \pm 166$ | $7.4 \pm 4.9$ | $6.7 \pm 4.5$ |
| 48 PHE | $H_{\beta_2}$ | 3.0 | $317 \pm 10$ | $5.7 \pm 1.4$ | $5.1 \pm 1.5$ |
| 48 PHE | $H_{\beta_3}$ | 3.0 | $77 \pm 10$ | $2.2 \pm 0.6$ | $1.9 \pm 0.7$ |
| 49 MET | $H_{\beta_3}$ | 3.0 | $212 \pm 37$ | $9.7 \pm 4.1$ | $8.8 \pm 3.9$ |
| 49 MET | $H_{\beta_2}$ | 11.0 | $92 \pm 37$ | $4.4 \pm 3.6$ | $4.0 \pm 3.2$ |
| 50 LEU | $H_{\beta_3}$ | 3.0 | $215 \pm 37$ | $9.8 \pm 3.8$ | $9.0 \pm 3.6$ |
| 50 LEU | $H_{\beta_2}$ | 10.0 | $95 \pm 37$ | $4.2 \pm 4.0$ | $3.7 \pm 3.6$ |

**Table 7** continued

| Residue | Proton | $^3J_{\alpha\beta}^{exp}$ | $\langle\theta_\beta\rangle$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{\text{DeMarco}}$ | $\langle^3J_{\alpha\beta}^{calc}\rangle_{\text{genKarplus}}$ |
|---|---|---|---|---|---|
| 52 LYSH | $H_{\beta_3}$ | 3.0 | $250 \pm 73$ | $7.8 \pm 4.6$ | $7.0 \pm 4.3$ |
| 52 LYSH | $H_{\beta_2}$ | 5.0 | $130 \pm 73$ | $5.8 \pm 4.5$ | $5.2 \pm 4.1$ |
| 59 TRP | $H_{\beta_2}$ | 9.1 | $124 \pm 19$ | $6.2 \pm 2.7$ | $4.8 \pm 2.4$ |
| 61 GLU | $H_{\beta_2}$ | 13.0 | $143 \pm 46$ | $9.4 \pm 4.5$ | $4.8 \pm 4.1$ |
| 61 GLU | $H_{\beta_3}$ | 4.1 | $263 \pm 46$ | $5.5 \pm 4.3$ | $8.5 \pm 4.1$ |
| 67 SER | $H_{\beta_3}$ | 3.0 | $56 \pm 15$ | $3.8 \pm 1.1$ | $2.7 \pm 1.0$ |
| 67 SER | $H_{\beta_2}$ | 5.1 | $296 \pm 15$ | $3.4 \pm 1.4$ | $1.7 \pm 1.2$ |
| 71 ARG | $H_{\beta_3}$ | 4.1 | $285 \pm 35$ | $4.0 \pm 2.9$ | $3.5 \pm 2.8$ |
| 73 LYSH | $H_{\beta_3}$ | 3.0 | $212 \pm 43$ | $9.9 \pm 4.2$ | $9.0 \pm 4.0$ |
| 73 LYSH | $H_{\beta_2}$ | 11.1 | $92 \pm 43$ | $5.0 \pm 4.0$ | $4.6 \pm 3.5$ |
| 74 LEU | $H_{\beta_3}$ | 3.0 | $250 \pm 43$ | $5.7 \pm 4.1$ | $5.0 \pm 4.0$ |
| 74 LEU | $H_{\beta_2}$ | 12.2 | $130 \pm 43$ | $7.9 \pm 4.1$ | $6.9 \pm 3.8$ |
| 77 SER | $H_{\beta_3}$ | 1.0 | $62 \pm 8$ | $3.2 \pm 0.9$ | $2.3 \pm 0.8$ |
| 77 SER | $H_{\beta_2}$ | 4.0 | $302 \pm 8$ | $3.8 \pm 1.0$ | $1.9 \pm 0.8$ |
| 79 ASP | $H_{\beta_3}$ | 1.0 | $51 \pm 10$ | $4.6 \pm 1.3$ | $4.7 \pm 1.5$ |
| 79 ASP | $H_{\beta_2}$ | 5.0 | $291 \pm 10$ | $2.7 \pm 0.9$ | $2.0 \pm 0.9$ |
| 80 TYR | $H_{\beta_3}$ | 3.0 | $233 \pm 28$ | $6.7 \pm 4.0$ | $6.0 \pm 3.9$ |
| 80 TYR | $H_{\beta_2}$ | 12.0 | $113 \pm 28$ | $5.2 \pm 3.6$ | $4.2 \pm 3.1$ |
| 82 TYR | $H_{\beta_3}$ | 1.1 | $234 \pm 45$ | $7.1 \pm 5.1$ | $6.3 \pm 4.9$ |
| 82 TYR | $H_{\beta_2}$ | 13.2 | $114 \pm 45$ | $6.9 \pm 4.3$ | $6.2 \pm 3.6$ |
| 97 LEU | $H_{\beta_3}$ | 4.0 | $292 \pm 15$ | $3.1 \pm 1.2$ | $2.5 \pm 1.3$ |
| 97 LEU | $H_{\beta_2}$ | 11.1 | $172 \pm 15$ | $12.1 \pm 1.3$ | $10.9 \pm 1.4$ |
| 99 PHE | $H_{\beta_3}$ | 2.4 | $286 \pm 10$ | $2.4 \pm 0.6$ | $1.8 \pm 0.7$ |
| 99 PHE | $H_{\beta_2}$ | 12.1 | $166 \pm 10$ | $12.1 \pm 0.9$ | $10.7 \pm 1.0$ |
| 100 ASP | $H_{\beta_2}$ | 2.0 | $67 \pm 9$ | $2.8 \pm 0.9$ | $2.9 \pm 1.1$ |
| 100 ASP | $H_{\beta_3}$ | 10.0 | $187 \pm 9$ | $12.5 \pm 0.8$ | $11.4 \pm 0.7$ |
| 106 LEU | $H_{\beta_3}$ | 3.0 | $273 \pm 40$ | $4.2 \pm 3.6$ | $3.6 \pm 3.4$ |
| 106 LEU | $H_{\beta_2}$ | 12.2 | $153 \pm 40$ | $10.3 \pm 3.7$ | $9.2 \pm 3.3$ |
| 107 GLU | $H_{\beta_2}$ | 3.0 | $166 \pm 41$ | $10.7 \pm 3.6$ | $9.6 \pm 3.3$ |
| 107 GLU | $H_{\beta_3}$ | 7.0 | $286 \pm 41$ | $3.8 \pm 2.9$ | $3.3 \pm 2.8$ |
| $Q$: | | | | 3.5 | 3.4 |

The $Q$-values quantifying the agreement between $^3J_{\alpha\beta}^{exp}$ and each set of $\langle^3J_{\alpha\beta}^{calc}\rangle$ are given at the bottom. All $^3J$-couplings and $Q$-values are given in Hz and the $\langle\theta_\beta\rangle$-angle values are given in degrees

and $\cos^2(\theta)$ was investigated using cumulative ensemble averages $\langle\ldots\rangle_t$. To follow the evolution of the ensemble averages $\langle\ldots\rangle$ over time, the total simulation time was divided into 10 equal time periods, i.e. 100 ps for Plastocyanin, 2 ns for HEWL and 800 ps for FKBP.

The determination of an average value $\langle\theta_\beta\rangle$ for a dihedral angle $\theta_\beta$ will depend on the range of values considered, e.g. $[-180°, +180°]$ or $[0°, +360°]$. Therefore, as well as calculating the mean, we calculated the probability distribution $P(\theta_\beta)$ for a given range, i.e. $[-180°, +180°]$, from a MD trajectory and used the $\theta_\beta$-value for which $P(\theta_\beta)$ is largest, i.e. the median, rather than the mean. For the ensemble of NMR structures of Plastocyanin, 9PCY,

and HEWL, 1E8L, and the two X-ray structures of HEWL, 193L, the mean dihedral angle was always used. The root-mean-square fluctuation of $\theta_\beta$ was calculated from the $P(\theta)$ obtained from the MD trajectory in which the $\theta_\beta$-values were not mapped onto a finite range.

## Results

### Calculation of $^3J_{\alpha\beta}$-values

A diverse range of parameters for the standard Karplus relation have been proposed (Abraham and McLauchlan

1962; Deber et al. 1971; Kopple et al. 1973; de Marco et al. 1978; Fischman et al. 1980; Pérez et al. 2001), see Fig. 2 and Table 1, based on different parametrisation methods and molecules. This creates uncertainty as to how to choose the optimal parameter set for Eq. 1 for use in protein structure determination. The parameters of de Marco et al. (1978) are most widely used, and also result in Karplus curves that lie between the curves generated by the other parameter sets, thus we use these for our initial investigations. The distribution of the measured $^3J_{\alpha\beta}^{exp}$-couplings over the protein structures is shown in Fig. 5. For all three proteins the couplings are well spread over the residues and throughout the space occupied by the protein. Figure 3 shows the measured $^3J_{\alpha\beta}^{exp}$-coupling constants of subsets 1 and 2 versus the value of $\theta_\beta$ (or $\langle\theta_\beta\rangle$ for a set of NMR model structures) in the X-ray or NMR structures along with the Karplus relation using the parameters of de Marco et al. (1978). Several $^3J_{\alpha\beta}^{exp}$-values deviate considerably from the value suggested by the curve for the corresponding $\theta_\beta$ or $\langle\theta_\beta\rangle$. Moreover, the values of $\theta_\beta$ or $\langle\theta_\beta\rangle$ for the dihedral angles for which $^3J_{\alpha\beta}$-couplings were measured do not cover the whole dihedral-angle value range. For HEWL and FKBP in particular, they are clustered around the canonical rotamer positions. This is a known issue in the determination of Karplus parameters.

Firstly, the averaged values $\langle^3J_{\alpha\beta}^{calc}\rangle$ calculated from the X-ray or NMR model structures or from MD simulation trajectories are compared to those measured experimentally for each protein, see Fig. 6. The $\langle^3J_{\alpha\beta}^{calc}\rangle$-values were computed using the standard Karplus relation with the de Marco parameters, and using the generalised Karplus relation as described in the "Method" section. A deviation of ± 1 Hz is considered to be acceptable, given the uncertainty in the Karplus parameters.

For Plastocyanin, the $\langle^3J_{\alpha\beta}^{calc}\rangle$ calculated from the water simulation using the 53A6 force field using the standard Karplus relation lie slightly closer to the experimental

values than those calculated from the vacuum simulation using the 45B3 force field, giving rise to Q-values of 1.8 Hz (Table 3) and 2.1 Hz (Table S1) respectively. Many of the $\langle^3J_{\alpha\beta}^{calc}\rangle$ calculated from the set of NMR model
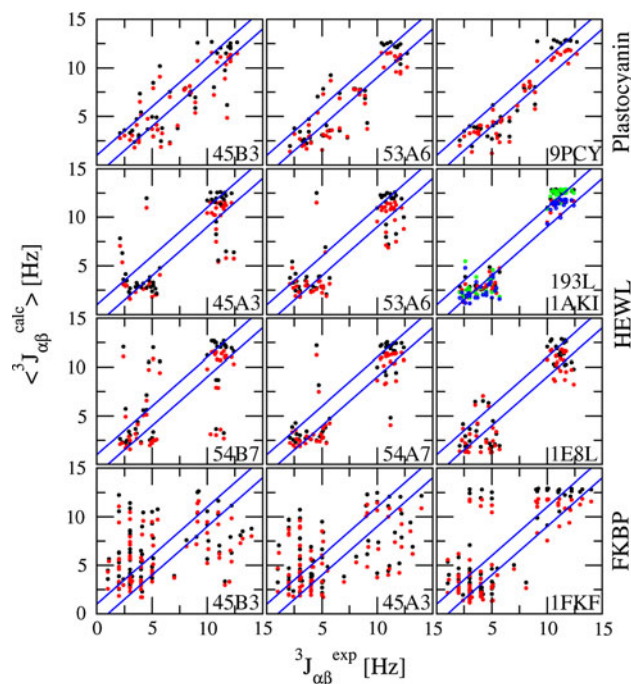


**Fig. 6** Comparison of the stereospecifically assigned $^3J_{\alpha\beta}$-couplings (subset 1) measured experimentally $^3J_{\alpha\beta_{2/3}}^{exp}$ and calculated $\langle^3J_{\alpha\beta_{2/3}}^{calc}\rangle$ using the parameters of de Marco et al. (1978) and the standard Karplus relation (*black*) and using the generalised Karplus relation (*red*) for Plastocyanin (*top row*), HEWL (*second* and *third rows*), and FKBP (*bottom row*). The *blue lines* indicate a deviation of ±1 Hz. The $\langle^3J_{\alpha\beta}^{calc}\rangle$ are calculated from the simulations in vacuum (*left panels* except *second row*) and in water (*middle panels* and *left panel second row*) and from the experimental model structures (*right panels*). In the *second row*, results from two water simulations (45A3, *left panel*, and 53A6, *middle panel*) and two X-ray structure sets 1AKI (*green* and *blue dots* for the standard and generalised Karplus relation respectively) and 193L (*black* and *red dots*) are given
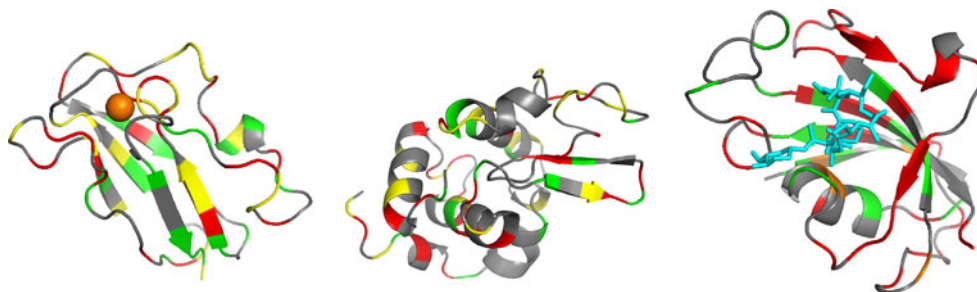


**Fig. 5** Cartoon pictures of Plastocyanin (NMR model structure 16 of 9PCY (Moore et al. 1991) with the coordinated copper ion in orange, *left panel*), HEWL (X-ray structure 1AKI (Artymiuk et al. 1982; Carter et al. 1997), *middle panel*), and FKBP (X-ray structure 1FKF (Xu et al. 1992) with the bound ascomycin in blue, *right panel*). The amino acids for which $^3J_{\alpha\beta}$-values are available are shown in *red*

(stereospecifically assigned $^3J_{\alpha\beta_{2/3}}$), *green* (residues Val, Ile and Thr with only one $H_\beta$), and *yellow* (non-stereospecifically assigned $^3J_{\alpha\beta_{2/3}}$). In the *right panel* the orange amino acids are the ones for which only one of the two stereospecifically assigned $^3J_{\alpha\beta_{2/3}}$ is available

**Table 8** Karplus relation parameters $a$, $b$, and $c$ and the corresponding $Q$-value obtained by fitting the values of $\langle {}^3J_{\alpha\beta}^{calc} \rangle$ calculated for the indicated structures to the stereospecifically assigned measured ${}^3J_{\alpha\beta}^{exp}$-coupling constants (subsets 1 and 2)

| Protein | Structure set | Source | Number of structures (simulation length) | Least-squares fitted parameters | | | |
|---|---|---|---|---|---|---|---|
| | | | | $a$ | $b$ | $c$ | $Q$ |
| Plasto- | 45B3 | MD (vac) | 200 (1 ns) | 5.66 | −1.41 | 4.51 | 2.00 |
| cyanin | 53A6 | MD (wat) | 200 (1 ns) | 6.61 | −0.93 | 3.96 | 1.86 |
| | 9PCY | NMR | 16 | 6.61 | −1.07 | 3.36 | 1.13 |
| HEWL | 54B7 | MD (vac) | 4,000 (20 ns) | −0.21 | −3.72 | 6.32 | 2.87 |
| | 54A7 | MD (wat) | 4,000 (20 ns) | 5.00 | −2.56 | 4.10 | 1.69 |
| | 53A6 | MD (wat) | 4,000 (20 ns) | 4.80 | −2.42 | 4.27 | 1.98 |
| | 45A3 | MD (wat) | 4,000 (20 ns) | 2.40 | −2.67 | 5.11 | 2.69 |
| | 1AKI | X-ray | 1 | 3.99 | −2.68 | 4.42 | 1.10 |
| | 193L | X-ray | 2 | 5.87 | −1.59 | 3.50 | 0.99 |
| | 1E8L | NMR | 50 | 3.55 | −3.38 | 4.51 | 1.10 |
| FKBP | 45B3 | MD (vac) | 1,600 (8 ns) | 0.60 | −2.72 | 5.60 | 3.63 |
| | 45A3 | MD (wat) | 1,600 (8 ns) | −1.14 | −4.48 | 6.25 | 3.38 |
| | 1FKF | X-ray | 1 | 5.34 | −0.65 | 3.70 | 3.15 |

All values are given in Hz. The structure sets are either MD trajectories from simulations in a vacuum (vac) or water (wat) environment with a particular force field, or experimental X-ray or NMR model structures from the PDB (Berman et al. 2000). The structure set is denoted either by the PDB entry or by the code of the force field used

structures, shown in Fig. 6, deviate more than ±1 Hz from the measured ${}^3J_{\alpha\beta}^{exp}$-values, but the overall agreement is better than for the MD simulations ($Q = 1.5$ Hz). Use of the generalised Karplus relation of Eq. 4 results in $Q$-values of 2.0 Hz for the water simulation, 2.2 Hz for the vacuum simulation and 1.4 Hz for the set of NMR model structures. Except for the latter, the deviation from the measured data is even larger than for the $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-values calculated with the standard Karplus relation of Eq. 1. It is also obvious that with the generalised Karplus relation, the calculated $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-couplings shift to lower values.

For HEWL, Fig. 6 shows a similar situation: the vacuum simulation (54B7 force field) yields the highest $Q$ (3.4 Hz, Table S2), the water simulations perform better ($Q = 2.5$ Hz for 45A3, Table 4, and 2.0 Hz for force fields 53A6 and 54A7, Tables 5 and 6), and the $\langle {}^3J_{\alpha\beta}^{calc} \rangle$ obtained from the set of NMR model structures (1E8L) agree best with the experimental data ($Q = 1.7$ Hz). Upon application of the generalised Karplus relation, the same shift downwards of the $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-values is observed, but the $Q$-value improves slightly for the vacuum simulation ($Q = 3.3$ Hz) and for the water simulation using the 45A3 force field ($Q = 2.4$ Hz), stays the same for the two other water simulations using the force fields 53A6 and 54A7 ($Q = 2.0$ Hz), and becomes worse for the NMR model structures ($Q = 2.0$ Hz). The $Q$-value for the X-ray structure 1AKI decreases from 1.6 Hz for the standard Karplus relation to 1.5 Hz for the generalised Karplus relation. The same tendency is observed for the average over the

two X-ray structures from 193L, where the $Q$-value is 1.4 Hz for the standard Karplus relation and 1.2 Hz for the generalised Karplus relation.

For FKBP, the calculated $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-values diverge much more from the measured ones than for the other two proteins, see Fig. 6. This is also evident in the high $Q$-values obtained using the standard Karplus relation: 3.9 Hz for the vacuum simulation using the 45B3 force field (Table S3) and 3.5 Hz for the water simulation using the 45A3 force field (Table 7). Even the ${}^3J_{\alpha\beta}^{calc}$-values from the X-ray structure yield a $Q$-value of 3.0 Hz. Using the generalised Karplus relation, the agreement improves slightly, but with $Q$-values of 3.6, 3.4, and 2.8 Hz for the vacuum simulation, water simulation and X-ray structure respectively it is still worse than for the other two proteins. As for Plastocyanin and HEWL, the calculated $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-couplings shift towards lower values when the generalised Karplus relation is used. It is noteworthy that for FKBP, most of the experimental ${}^3J_{\alpha\beta}^{exp}$-values are close to integer values, suggesting that this data may be of limited precision.

For all three proteins, the inclusion of solvent in the MD simulations improves the agreement with experimental data when either the standard or generalised Karplus relation is used. Despite including substituent effects, use of the generalised Karplus relation does not significantly improve the agreement with the measured ${}^3J_{\alpha\beta}^{exp}$-couplings. In all cases, however, the $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-values calculated from the X-ray and NMR model structures agree better with the measured ${}^3J_{\alpha\beta}^{exp}$. This may be a
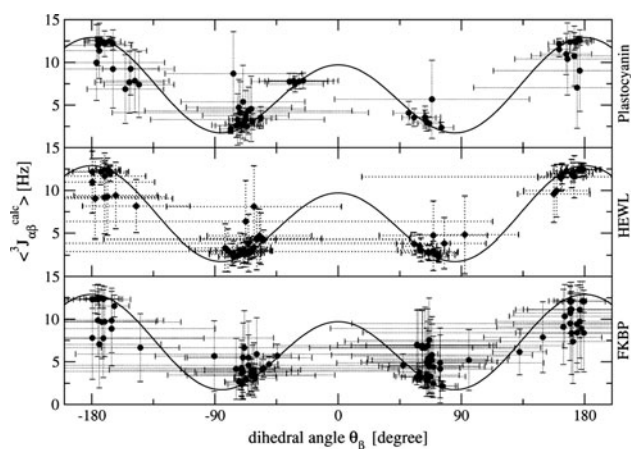
**Fig. 7** Median (*black circles*) and rms variation (*bars*) of the dihedral-angle values $\theta_\beta$ and corresponding $\langle {}^3J_{\alpha\beta_{2/3}}^{calc} \rangle$-values of subsets 1 and 2 calculated from the MD simulations in water of Plastocyanin (53A6 force field, *upper panel*), HEWL (54A7 force field, *middle panel*) and FKBP (45A3 force field, *lower panel*) using the standard Karplus relation with the parameters of de Marco et al. (1978). The Karplus curves generated using the de Marco parameters are shown as *black lines*
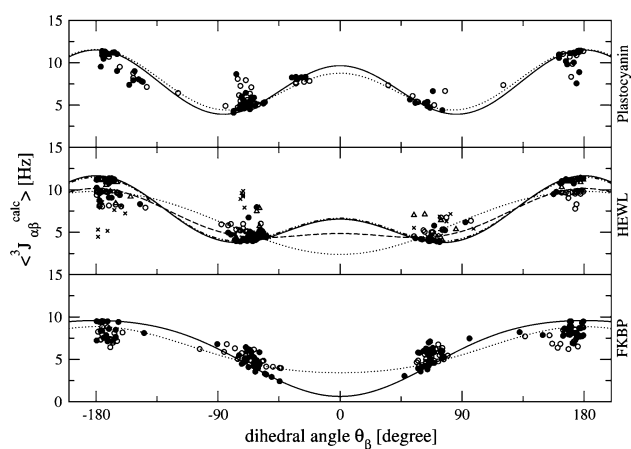


**Fig. 8** Median of the dihedral-angle values $\theta_\beta$ and the average $\langle {}^3J_{\alpha\beta_{2/3}}^{calc} \rangle$-values of subsets 1 and 2 calculated from the MD simulations of Plastocyanin (*upper panel*), HEWL (*middle panel*), and FKBP (*lower panel*) and the Karplus curves obtained using the parameters optimised by least-squares fitting of the plotted $\langle {}^3J_{\alpha\beta_{2/3}}^{calc} \rangle$-values to the measured ${}^3J_{\alpha\beta_{2/3}}^{exp}$-values of the subsets 1 and 2. For Plastocyanin and FKBP, the *filled circles* and *solid lines* are for the simulations in water (53A6 or 45A3 force field, respectively) and the *open circles* and *dotted lines* are for the simulations in vacuum (45B3 force field). For HEWL, the *filled circles* and *solid line* are for the simulation in water (54A7 force field), the *open circles* and *dotted line* for the simulation in vacuum (54B7 force field), the *crosses* and *dashed line* are for the simulation in water (45A3 force field), and the *triangles* and *dot-dot-dashed line* are for the simulation in water (53A6 force field)

consequence of using parameters for the Karplus relations, both standard and generalised, that were derived using rather rigid small molecules.
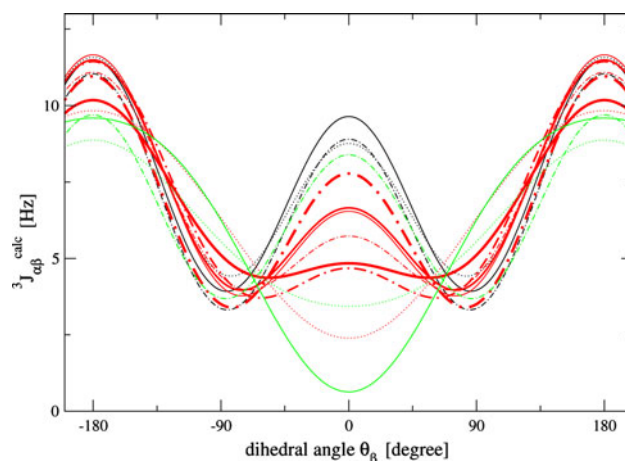


**Fig. 9** Karplus curves generated using the optimised parameters obtained by least-squares fitting of $\langle {}^3J_{\alpha\beta_{2/3}}^{calc} \rangle$-values calculated from different structure sets to the measured ${}^3J_{\alpha\beta_{2/3}}^{exp}$ for subsets 1 and 2 for Plastocyanin (*black*), HEWL (*red*), or FKBP (*green*). For Plastocyanin and FKBP, the *solid lines* correspond to the simulation in water (53A6 and 45A3 force field, respectively), the *dotted lines* to the simulation in vacuum (45B3 force field) and the *dash-dotted lines* to the NMR model structures (9PCY) or X-ray structure (1FKF), respectively. For HEWL, three different X-ray or NMR model structures (*dot-dashed lines*) and MD trajectories (*solid lines*) in water were analysed: the 1AKI X-ray structure and the simulation using the 54A7 force field (*thin lines*), the 1E8L NMR model structures and the simulation using the 53A6 force field (*normal lines*), and the 193L X-ray structures and the simulation using the 45A3 force field (*thick lines*). The Karplus curve generated using the parameters optimised against the vacuum simulation of HEWL (54B7 force field) is shown as a *dotted red line*
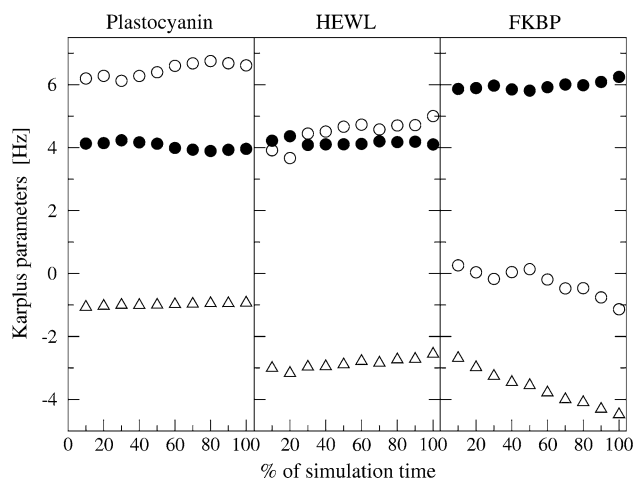


**Fig. 10** Karplus parameters *a* (*open circles*), *b* (*triangles*), and *c* (*filled circles*) as a function of the proportion of the simulation period used to calculate the $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-values used in the fitting procedure for Plastocyanin (53A6 force field, 100 % = 1 ns), HEWL (54A7 force field, 100 % = 20 ns), and FKBP (45A3 force field, 100 % = 8 ns)

Indeed, both the standard and the generalised Karplus relation link experimentally measured ${}^3J_{\alpha\beta}^{exp}$-couplings to a single angle value, implying a static structure. Because
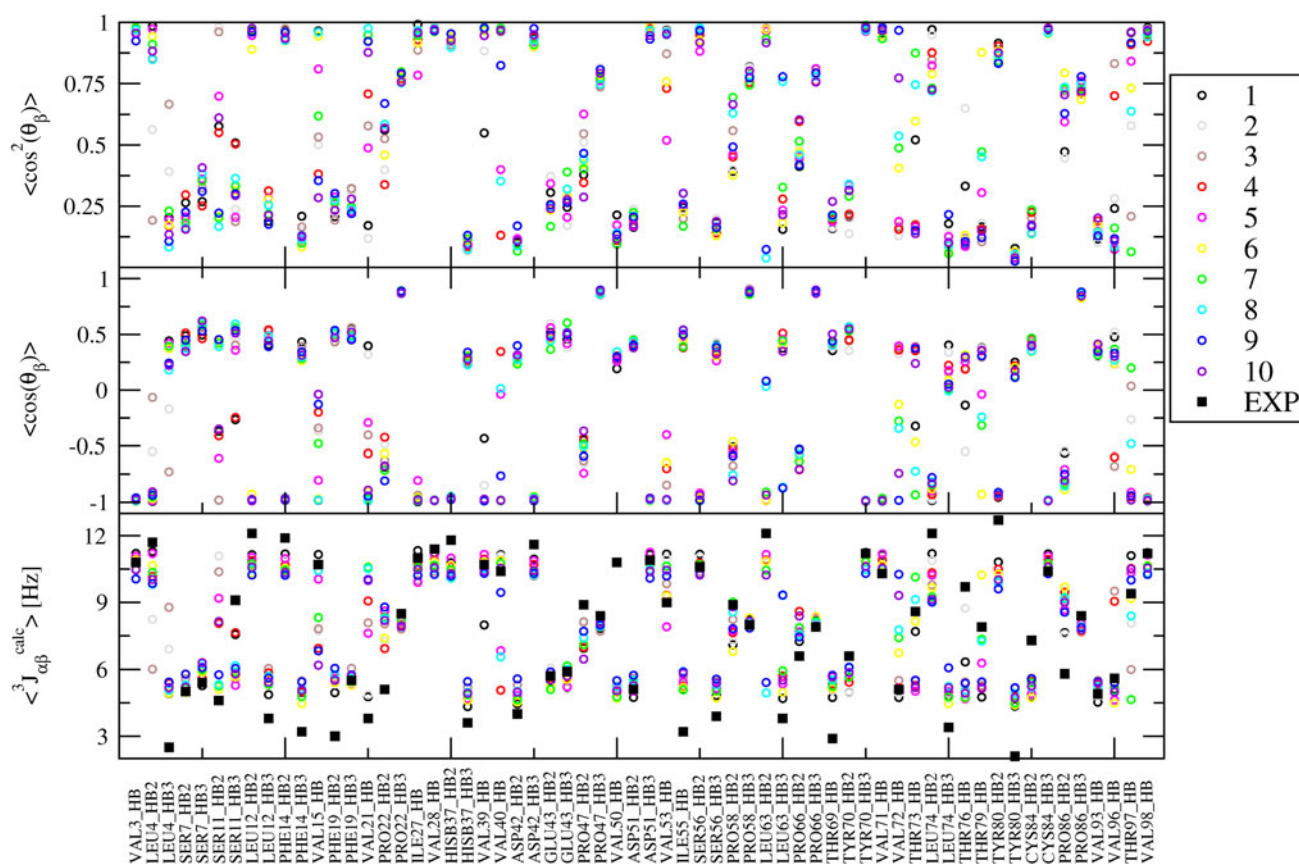
**Fig. 11** $\langle cos^2\theta_\beta\rangle$, $\langle cos\theta_\beta\rangle$, and $\langle ^3J_{\alpha\beta}^{calc}\rangle$-values calculated from 10 time windows (each 100 ps) of the MD simulation of Plastocyanin in water (53A6 force field). The measured $^3J_{\alpha\beta}^{exp}$-values (Moore et al. 1991) are shown as *black squares*. $\langle ^3J_{\alpha\beta}^{calc}\rangle$-values were obtained using

the Karplus parameters *a*, *b*, and *c* from the least-squares fit to the measured $^3J_{\alpha\beta}^{exp}$-values using the averaged $\langle cos^2\theta_\beta\rangle$ and $\langle cos\theta_\beta\rangle$ of the corresponding time window

NMR experiments take place in solution or are measured from a powder, however, it is expected that the proteins are mobile on the time-scale of the measurement or that they point in different directions in the powder, meaning that the measured $^3J_{\alpha\beta}^{exp}$-couplings are averages over an ensemble of structures as well as over the timescale of the experiment.

Least-squares fitting of Karplus parameters

One way to overcoming this discrepancy is to employ least-squares fitting of the $\langle ^3J_{\alpha\beta}^{calc}\rangle$, averaged over MD trajectories or a set of experimental X-ray or NMR model structures, to the measured $^3J_{\alpha\beta}^{exp}$-couplings to determine *a*, *b*, and *c* parameters for the standard Karplus relation in Eq. 1. Table 8 lists the parameters *a*, *b*, and *c* obtained in this manner using only subsets 1 and 2 of the $^3J_{\alpha\beta}$-couplings for each simulation of each protein and from the experimentally determined structure(s), along with the Q-values for each fit. For all three proteins, the same result is seen as before, with the $\langle ^3J_{\alpha\beta}^{calc}\rangle$-values from the vacuum simulations resulting in the highest Q-values, and the

$\langle ^3J_{\alpha\beta}^{calc}\rangle$-values from the experimental structures yielding a better fit to the $^3J_{\alpha\beta}^{exp}$-values than those from the MD simulation trajectories. The enhanced conformational flexibility in the MD trajectories seems to complicate the fitting to a Karplus relation of the form of Eq. 1. This may be due to the fact that not all of the dihedral angles necessarily undergo the same degree of conformational averaging in an MD simulation, meaning that some of the $\langle ^3J_{\alpha\beta}^{calc}\rangle$-values used for the fitting are averages over a wide distribution and other arise from dihedral angles that are nearly rigid. This degree of conformational averaging for a specific dihedral angle may or may not correspond to the one occurring in experiment.

The degree of conformational sampling that the side-chain dihedral angles $\theta_\beta$ with stereospecifically assigned $^3J_{\alpha\beta}^{exp}$-values (subset 1 and 2) undergo during MD simulation is shown in Fig. 7 for the simulations of Plastocyanin (53A6), HEWL (54A7), and FKBP (45A3) in water, along with the corresponding $^3J_{\alpha\beta}^{calc}$-values calculated using the de Marco parameters. Both the dihedral-angle values $\theta_\beta$ and the $^3J_{\alpha\beta}^{calc}$-values show significant variation, with the $^3J_{\alpha\beta}^{calc}$
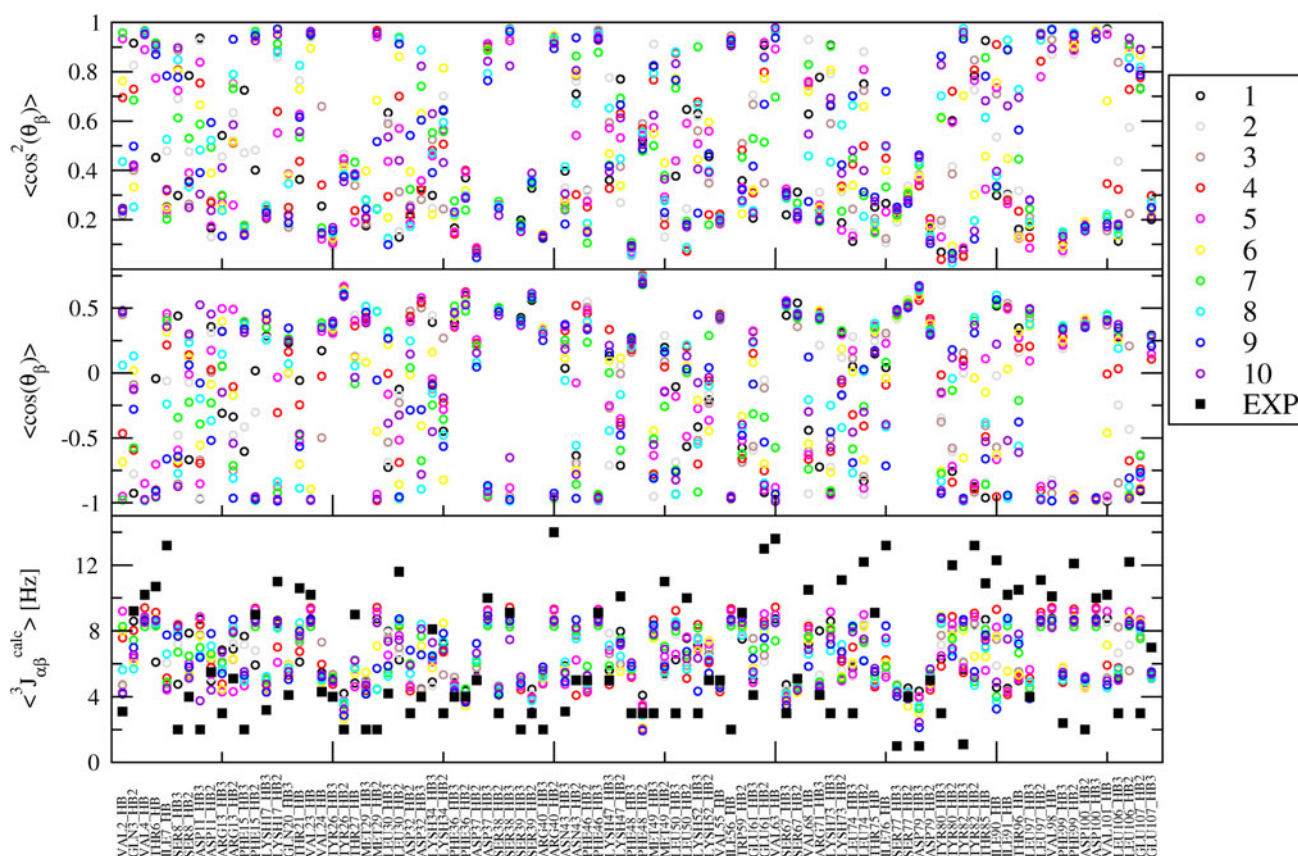
**Fig. 12** $\langle cos^2\theta_\beta \rangle$, $\langle cos\theta_\beta \rangle$, and $\langle ^3J_{\alpha\beta}^{calc} \rangle$-values calculated from 10 time windows (each 2 ns) of the MD simulation of HEWL in water (54A7 force field). The measured $^3J_{\alpha\beta}^{exp}$-values (Smith et al. 1991) are shown as *black squares*. $\langle ^3J_{\alpha\beta}^{calc} \rangle$-values were obtained using the Karplus parameters *a*, *b*, and *c* from the least-squares fit to the measured $^3J_{\alpha\beta}^{exp}$-values using the averaged $\langle cos^2\theta_\beta \rangle$ and $\langle cos\theta_\beta \rangle$ of the corresponding time window
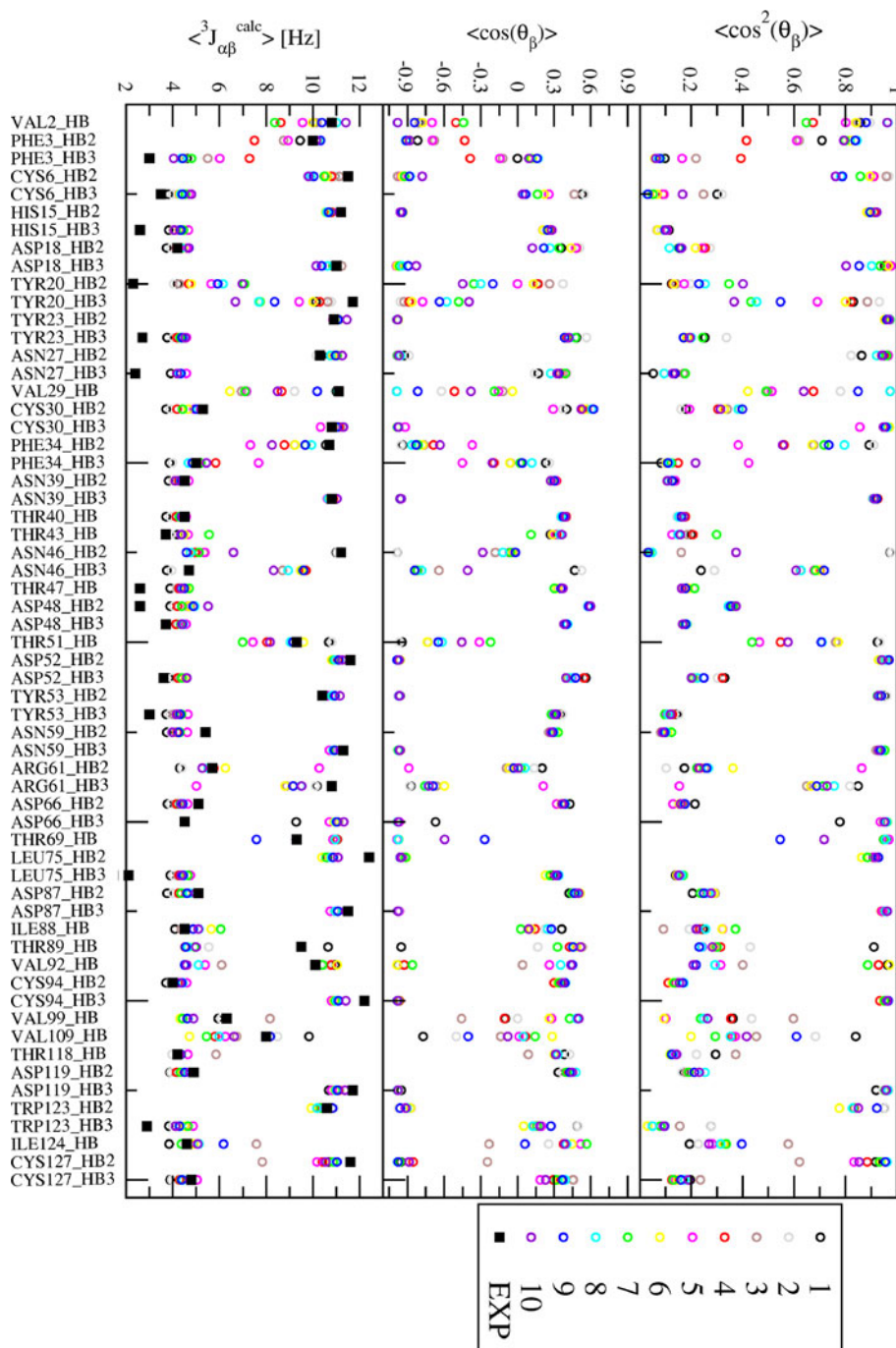


varying by up to 10 Hz. Not all dihedral angles undergo the same degree of conformational sampling. Moreover, the range of corresponding $^3J_{\alpha\beta}^{calc}$-values depends on where on the Karplus curve the dihedral-angle value $\theta_\beta$ lies: variation in dihedral-angle values situated in flat parts of the Karplus curve has relatively little effect on the $^3J_{\alpha\beta}^{calc}$-values, whereas a small change in a $\theta_\beta$ located in a steep part of the Karplus curve results in a comparably large change in the $^3J_{\alpha\beta}^{calc}$-value. Because of this, the $\langle ^3J_{\alpha\beta}^{calc} \rangle$-values calculated from the parameters obtained in the fitting procedure are often different from the $^3J_{\alpha\beta}$-values predicted by the Karplus relation using the same parameters for the corresponding $\langle \theta_\beta \rangle$. Together, these effects cause a large variation in the parameters in Table 8 obtained using least-squares fitting to subsets 1 and 2 for the different simulations of the three proteins studied here and in the resulting Karplus curves shown in Figs. 8 and 9.

Even for the same protein, fitting the Karplus parameters to different simulations with different force fields, in water or in vacuum, yields different parameter sets, most

**Fig. 13** $\langle cos^2\theta_\beta \rangle$, $\langle cos\theta_\beta \rangle$, and $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-values calculated from 10 time windows (each 800 ps) of the MD simulation of FKBP in water (45A3 force field). The measured ${}^3J_{\alpha\beta}^{exp}$-values (Xu et al. 1992) are shown as *black squares*. $\langle {}^3J_{\alpha\beta}^{calc} \rangle$-values were obtained using the Karplus parameters *a*, *b*, and *c* from the least-squares fit to the measured ${}^3J_{\alpha\beta}^{exp}$-values using the averaged $\langle cos^2\theta_\beta \rangle$ and $\langle cos\theta_\beta \rangle$ of the corresponding time window

noticeably for HEWL and FKBP, see Figs. 8 and 9 and Table 8. The three curves obtained for FKBP are all rather different, with the curves obtained from the MD trajectories exhibiting only one maximum. For HEWL, the curve for the vacuum simulation also has only one maximum. The remainder of the curves for HEWL, along with those for Plastocyanin, display the expected two maxima, but their heights vary considerably. For HEWL, the simulations in water using the 54A7 and 53A6 force fields produce very similar Karplus parameters and curves, but the simulation carried out using the 45A3 force field yields a curve almost without a second maximum. The two different X-ray structure sets, 1AKI and 193L, and the NMR model structures 1E8L give rise to three quite different curves in terms of the height of the maximum centred at $\theta_\beta = 0°$, and the curves obtained from fitting to the MD simulation data using the 54A7 and 53A6 force fields lie in between.

The large variation around $\theta_\beta = 0°$ between the differently fitted Karplus curves in Fig. 9 is due to the lack of dihedral-angle values in the range $-60° < \theta_\beta < 60°$ in the X-ray or NMR structures or MD simulation trajectories, as shown in Figs. 3 and 7. Chemically this makes sense, as eclipsed conformations are generally disfavored compared to staggered conformations. In experimental structure refinement, often only staggered conformations, the rotamers $g^+$, $g^-$ and $t$, are considered to be energetically favourable. In contrast, in the MD simulations, quite a wide range of angle values is sampled outside of $-60° < \theta_\beta < 60°$, although the median, i.e. the most populated dihedral-angle values $\theta_\beta$, are concentrated around the classical rotamer positions $\theta_\beta = \pm60°$ and $\pm180°$. Ultimately, however, it is angle values around $0° \pm60°$ that determine the shape of the curve, as the minima and maxima of the Karplus relation are mainly defined by the $cos^2$ function and the *b* parameter of the cos part of the Karplus relation determines the shape of the curve around $0°$.

The conformational motion that takes place during the MD simulations means that the least-squares fitted Karplus parameters will depend on the length of the simulation used
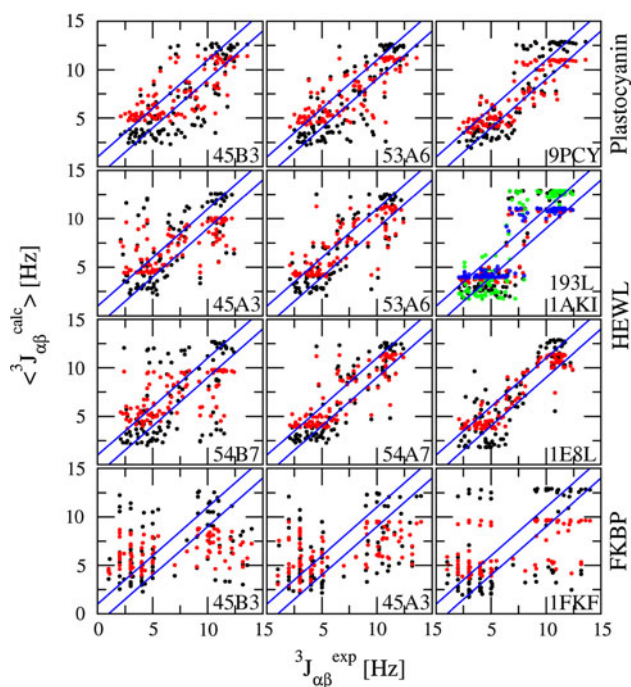
**Fig. 14** Comparison of all (subsets 1-3) measured $^3J_{\alpha\beta}^{exp}$-values for Plastocyanin (*top row*), HEWL (*second* and *third rows*), and FKBP (*bottom row*) with those calculated using the parameters of de Marco et al. (1978) (*black*) or the least-squares fitted parameters (*red*) optimised for that structure set. Note that the optimised Karplus parameters were obtained using subsets 1 and 2. The *blue lines* indicate a deviation of ±1 Hz. The $\langle^3J_{\alpha\beta}^{calc}\rangle$ are calculated from the simulations in vacuum (*left panels*) and in water (*middle panels*) and from the experimental model structures (*right panels*). In the *second row*, results from two water simulations (45A3, *left panel*, and 53A6, *middle panel*) and two X-ray structure sets 1AKI (*green* and *blue dots* for using the de Marco and the least-squares fitted parameters respectively) and 193L (*black* and *red dots*) are given

in the fitting procedure, i.e. on the range of different structures that are included. The dependence of the parameters $a$, $b$, and $c$ on the size of the time range considered for fitting is shown in Fig. 10 for the three proteins. The $a$, $b$, and $c$ values differ between the proteins and vary over the whole simulation period, even for the 20 ns simulation of HEWL in water. The origin of this variation can be seen in Figs. 11, 12, and 13, where the averages $\langle\cos^2\theta_\beta\rangle$ and $\langle\cos\theta_\beta\rangle$ over each 10 % window of the total simulation time together with the $\langle^3J_{\alpha\beta}^{calc}\rangle$ calculated using the least-squares fitted parameters for this time window are given for the three proteins. For Plastocyanin, Fig. 11, the side-chain angles $\theta_\beta$ of many of the residues, e.g. Ser 11, Val 15, Val 21, Pro 22, Val 40, Pro 47, Val 53, Pro 58, Leu 63, Val 72, Thr 73, Thr 79, Pro 86 , Val 96, and Thr 97 show considerable motion during the simulation, resulting in quite different $\langle\cos^2\theta_\beta\rangle$, $\langle\cos\theta_\beta\rangle$ and $\langle^3J_{\alpha\beta}^{calc}\rangle$-values for each window. In the case of HEWL, Fig. 12, less motion

occurs, but still some residues, e.g. Val 2, Phe 3, Tyr 20, Val 29, Phe 34, Asn 46, Thr 51, Thr 69, Val 92, 99, and 109, Trp 123 and Ile 124 show different values of the averages in each time window. In the FKBP simulation, nearly half of the residues exhibit quite different values of $\langle\cos^2\theta_\beta\rangle$, $\langle\cos\theta_\beta\rangle$, and $\langle^3J_{\alpha\beta}^{calc}\rangle$ over time as shown in Fig. 13.

The calculated $\langle^3J_{\alpha\beta}^{calc}\rangle$-values for all three subsets 1, 2 and 3 for all structure sets of all proteins calculated using the least-squares fitted Karplus parameters of the corresponding simulation or the de Marco parameters are compared to the measured $^3J_{\alpha\beta}^{exp}$ in Fig. 14. For subset 3, the assignment was chosen to minimise the Q-value, i.e. by assigning the larger of the two $\langle^3J_{\alpha\beta}^{calc}\rangle$-couplings to the larger of the two $^3J_{\alpha\beta}^{exp}$-coupling constants.

The robustness of a given set of parameters may be quantified by conducting jack-knife tests, in which the parameters obtained from fitting to one particular structure set of one protein are used to back-calculate $\langle^3J_{\alpha\beta}^{calc}\rangle$ for another structure set, possibly of another protein, and the goodness of fit (Q-value) is compared to the one obtained for the structure set used in the fitting procedure. Jack-knife tests were carried out for all possible combinations of fitted Karplus parameters, structure sets and proteins.

The assignment of the unassigned $^3J_{\alpha\beta}^{exp}$ of subset 3 adds some complication to this procedure. Two possible assignment protocols were tested:

1. Using a particular structure set and Karplus parameters $a$, $b$, and $c$ optimised using the $^3J_{\alpha\beta}$-values of subset 1 and 2 and that particular structure set, the assignment of the $^3J_{\alpha\beta}$-values of subset 3 is chosen such that the Q-value is minimal. Subsequently, this assignment of subset 3 is used for all calculations of Q-values for that particular structure set using all the different sets of $a$, $b$, and $c$ parameters, see Table S4.
2. For every combination of structure set and Karplus parameters $a$, $b$, and $c$, the assignment of the $^3J_{\alpha\beta}$-values of subset 3 is chosen such that the Q-value is minimal for that combination, see Table S5.

The Q-values obtained using the second procedure are, as expected, lower, but the differences are mostly small or nonexistent.

A robust parameter set might be expected to perform similarly in terms of Q-values for all structure sets of all proteins, not just for the one it was optimised for. Applying this criterion is complicated, however, by the fact that the $\langle^3J_{\alpha\beta}^{calc}\rangle$-values calculated from the various structure sets of the different proteins do not all match the experimental data equally well, even when the Karplus parameters optimised for that structure set of that protein are used.

**Table 9** $Q$-values in Hz quantifying the similarity between the measured $^3J_{\alpha\beta}^{exp}$-values and the calculated $\langle^3J_{\alpha\beta}^{calc}\rangle$-values of the two stereospecifically assigned subsets 1 and 2 for each structure set for each of the three proteins using the Karplus parameters obtained using the same structure set of that protein (bold) and using each structure set of all proteins

| | | Parameters fitted to | | | | | | | | | | | | | de Marco et al. (1978) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Plastocyanin | | | HEWL | | | | | | | FKBP | | | |
| $Q$ calculated for | | 45B3 | 53A6 | 9PCY | 54B7 | 54A7 | 53A6 | 45A3 | 1AKI | 193L | 1E8L | 45B3 | 45A4 | 1FKF | |
| Plasto-cyanin | 45B3 | **2.00** | 2.02 | 2.13 | 2.61 | 2.17 | 2.14 | 2.34 | 2.23 | 2.19 | 2.36 | 2.74 | 2.95 | 2.35 | 2.47 |
| | 53A6 | 1.88 | **1.86** | 1.95 | 2.57 | 2.07 | 2.03 | 2.26 | 2.13 | 2.03 | 2.29 | 2.68 | 2.91 | 2.18 | 2.30 |
| | 9PCY | 1.31 | 1.27 | **1.13** | 2.13 | 1.39 | 1.33 | 1.62 | 1.43 | 1.19 | 1.69 | 2.22 | 2.58 | 1.44 | 1.78 |
| HEWL | 54B7 | 3.04 | 3.04 | 3.03 | **2.87** | 3.10 | 3.05 | 2.89 | 3.00 | 3.04 | 3.08 | 2.95 | 2.89 | 2.97 | 3.57 |
| | 54A7 | 1.81 | 1.78 | 1.76 | 2.03 | **1.69** | 1.70 | 1.89 | 1.72 | 1.76 | 1.70 | 2.29 | 2.03 | 2.08 | 1.99 |
| | 53A6 | 2.06 | 2.03 | 2.02 | 2.20 | 1.99 | **1.98** | 2.11 | 2.00 | 2.03 | 2.00 | 2.44 | 2.24 | 2.27 | 2.28 |
| | 45A3 | 2.84 | 2.82 | 2.75 | 2.71 | 2.82 | 2.79 | **2.69** | 2.74 | 2.75 | 2.81 | 2.80 | 2.73 | 2.75 | 3.14 |
| | 1AKI | 1.42 | 1.41 | 1.23 | 1.53 | 1.17 | 1.14 | 1.28 | **1.10** | 1.17 | 1.16 | 1.86 | 1.66 | 1.55 | 1.79 |
| | 193L | 1.32 | 1.25 | 1.02 | 1.46 | 1.12 | 1.08 | 1.21 | 1.03 | **0.99** | 1.12 | 1.79 | 1.53 | 1.36 | 1.64 |
| | 1E8L | 1.57 | 1.64 | 1.54 | 1.66 | 1.18 | 1.20 | 1.40 | 1.16 | 1.39 | **1.10** | 1.94 | 1.76 | 1.89 | 1.91 |
| FKBP | 45B3 | 3.92 | 3.89 | 3.78 | 3.69 | 3.85 | 3.83 | 3.69 | 3.77 | 3.76 | 3.82 | **3.63** | 3.67 | 3.68 | 4.08 |
| | 45A3 | 3.63 | 3.61 | 3.51 | 3.41 | 3.52 | 3.51 | 3.42 | 3.46 | 3.48 | 3.48 | 3.42 | **3.38** | 3.48 | 3.74 |
| | 1FKF | 3.38 | 3.34 | 3.25 | 3.34 | 3.41 | 3.36 | 3.24 | 3.32 | 3.26 | 3.45 | 3.28 | 3.45 | **3.15** | 3.79 |

For comparison, the $Q$-values obtained when calculating the $\langle^3J_{\alpha\beta}^{calc}\rangle$ with the parameters of de Marco et al. (1978) are also given

**Table 10** Residues of FKBP for which the assignment was changed (c) in order to optimise the $Q$-value comparing the measured $^3J_{\alpha\beta}^{exp}$ of the subsets 1-3 to those back-calculated from the MD simulation of FKBP in water (45A3) using the parameters of de Marco et al. (1978) (DMopt) or the least-squares fitted Karplus parameters (LSFopt)

| Residue | DMopt | LSFopt | COMopt |
|---|---|---|---|
| 8 SER | c | c | c |
| 11 ASP | c | c | c |
| 34 LYS | c | c | c |
| 39 SER | - | c | c |
| 49 MET | c | c | c |
| 50 LEU | c | c | c |
| 52 LYS | c | c | c |
| 67 SER | c | - | c |
| 73 LYS | c | c | c |
| 77 SER | - | c | c |
| 79 ASP | c | - | c |
| 80 TYR | c | c | c |
| 82 TYR | c | c | c |
| 107 GLU | c | c | c |

The two sets of assignment changes are merged to form the set "COMopt"

This is in fact the dominant factor determining the magnitude of the $Q$-values in the majority of cases, see Tables S4 and S5, that is, there is more variation between the $Q$-values obtained using a given set of Karplus parameters to calculate the $\langle^3J_{\alpha\beta}^{calc}\rangle$-values from each protein structure set (variation within columns) than between those obtained using each different set of Karplus parameters for a given structure set (variation within rows). Indeed, the $Q$-values obtained using a given set of Karplus parameters to calculate the $\langle^3J_{\alpha\beta}^{calc}\rangle$-values from a protein structure set other than the one to which the parameters were fitted is in several cases better than the $Q$-value obtained during the fitting procedure. The main exception to this trend is Plastocyanin, for which the Karplus parameters obtained from fitting to the Plastocyanin structure sets give quite different $Q$-values to the parameters obtained from the FKBP and some of the HEWL structure sets.

A somewhat surprising result is that the $Q$-value obtained using the parameters optimised for the same structure set is not the lowest $Q$-value for that structure set in all cases, although it is always among the lowest. For instance, the $Q$-value calculated for the simulation of HEWL in the 54A7 force field is lower when the Karplus parameters obtained from the NMR model structures of Plastocyanin ($Q = 1.63$ Hz), the simulation of HEWL in 53A6 ($Q = 1.63$ Hz), the 1AKI ($Q = 1.61$ Hz) and 193L ($Q = 1.63$ Hz) X-ray structures or the 1E8L ($Q = 1.63$ Hz) NMR model structures are used than when the Karplus parameters obtained from the fit to the simulation of HEWL in the 54A7 force field are used ($Q = 1.64$ Hz).

**Table 11** Karplus parameters $a$, $b$, and $c$ and the corresponding $Q$-values obtained by least-squares fitting of the $\langle^3 J_{\alpha\beta}^{calc}\rangle$-values of all three subsets after the re-assignment of the measured $^3 J_{\alpha\beta}^{exp}$-values for FKBP

| Parameter set | Assignment | $a$ | $b$ | $c$ | $Q$ |
|---|---|---|---|---|---|
| DM | Xu | 9.5 | −1.6 | 1.8 | 3.74 |
| LSF | Xu | −1.14 | −4.48 | 6.25 | 3.38 |
| LSF-DMopt | DMopt | 2.75 | −3.85 | 4.47 | 2.95 |
| LSF-LSFopt | LSFopt | 0.97 | −4.74 | 5.19 | 2.94 |
| LSF-COMopt | COMopt | 2.29 | −4.08 | 4.66 | 2.95 |

DM refers to the Karplus parameters of de Marco et al. (1978) and LSF to the Karplus parameters obtained by least-squares fitting to the simulation of FKBP using the 45A3 force field. Xu refers to the original, published assignment of Xu et al. (1992), DMopt to the assignment optimised using the $\langle^3 J_{\alpha\beta}^{calc}\rangle$-values calculated using the DM parameters, LSFopt to the assignment optimised using the $\langle^3 J_{\alpha\beta}^{calc}\rangle$-values calculated using the LSF parameters and COMopt to the merged set of assignment changes in Table 10
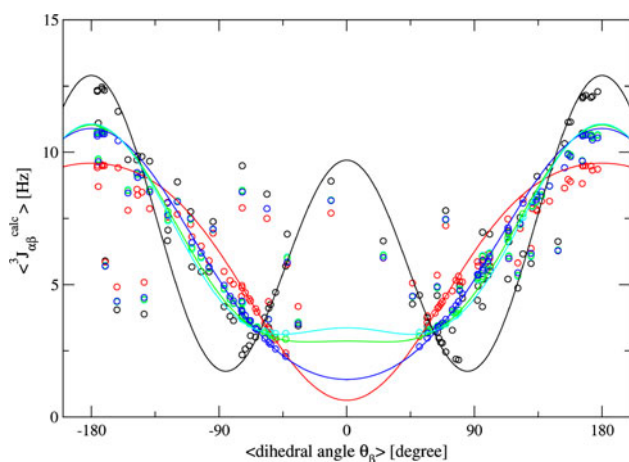


**Fig. 15** Karplus curves and $\langle^3 J_{\alpha\beta}^{calc}\rangle$-values calculated from the MD simulation of FKBP in water (45A3) using the parameters optimised for each different assignment given in Table 10 plotted against the unmapped mean of the corresponding dihedral angle $\langle\theta_\beta\rangle$ in the simulation, shifted to the value $\langle\theta_\beta\rangle - n \cdot 360$ in the range [−180°, +180°] for integer values of $n$. The assignments used in the optimisation and the Karplus parameter sets are LSF-DMopt/DMopt (*cyan*), LSF-LSFopt/LSFopt (*blue*), and LSF-COMopt/COMopt (*green*), see Table 11. The $\langle^3 J_{\alpha\beta}^{calc}\rangle$-values calculated using the assignments by Xu et al. (1992) and using the parameters of de Marco et al. (1978) (DM, *black*) and optimised (LSF, *red*) parameters and the corresponding Karplus curves are given for reference

Even more intriguing is the fact that the Karplus parameters obtained from fitting to the X-ray structure of FKBP perform well in the back-calculation of $\langle^3 J_{\alpha\beta}^{calc}\rangle$-values from the HEWL X-ray structures 1AKI ($Q = 1.51$ Hz) and 193L ($Q = 1.45$ Hz). These unexpected results may, however, occur due to the uncertainty introduced by the unassigned $^3 J_{\alpha\beta}$-couplings of subset 3. To avoid this uncertainty, a jack-knife test was carried out for subsets 1 and 2 only to calculate the $Q$-values in Table 9. With this approach, the $Q$-value of a specific structure set is always lowest when the Karplus parameters optimised for that structure set were used. The $Q$-values are now also more sensitive to the

set of Karplus parameters used, indicating that some of the apparent dominance of the structure set in the goodness of fit was due to assignment uncertainty.

On a related note, it should be remembered that the experimental NMR $^3 J_{\alpha\beta}$-coupling data for the three proteins were published two decades ago. They were measured in 2D $^1$H-$^1$H E.COSY experiments for Plastocyanin (Moore et al. 1991) and Lysozyme (Smith et al. 1991), and using 3D $^{15}$N-edited techniques for FKBP (Xu et al. 1992). Although no error bars for the $^3 J$-couplings are discussed in these papers, the resolution of the spectra is likely to have been of the order of 0.5 Hz to 1 Hz. Since $^3 J$-couplings are usually obtained from differences in peak positions, the uncertainty in the experimental data may easily be 1 Hz, which will contribute to the difficulty in obtaining good fits to the data (low $Q$-values) using all of the relations and parameter sets explored here.

### Reassignment of FKBP $^3 J_{\alpha\beta}^{exp}$-coupling constants

It is obvious from Tables 9, S4, and S5 that for FKBP, none of the parameter sets, even those fitted to the same structure set from which the $\langle^3 J_{\alpha\beta}^{calc}\rangle$-values were back-calculated, provides a good fit between the measured and calculated data. This is surprising given that FKBP is the only one of the three proteins for which all of the $^3 J_{\alpha\beta}^{exp}$-values were stereospecifically assigned. To check whether any of the couplings had been incorrectly assigned, the assignment of all $^3 J_{\alpha\beta_2}$- and $^3 J_{\alpha\beta_3}$-couplings for residues with two $H_\beta$ protons and two measured $^3 J_{\alpha\beta}^{exp}$-coupling constants was compared and changed according to the same exchange criterion as was described earlier for the fitting procedure. These comparisons were carried out using the average $\langle\cos^2\theta_\beta\rangle$ and $\langle\cos\theta_\beta\rangle$ values calculated from the MD simulation of FKBP in water (45A3) and the $\langle^3 J_{\alpha\beta}^{calc}\rangle$-values calculated using either the de Marco Karplus parameters or the fitted Karplus parameters. The resulting assignment changes are given in Table 10. A new

set of "re-fitted" Karplus parameters and corresponding $Q$-values were then calculated using the different optimised assignments, see Table 11. In Fig. 15, the Karplus curves obtained using these "re-fitted" Karplus parameters are compared.

In all cases, the "re-fitted" parameters lead to a decrease in the $Q$-values, from 3.38 Hz for using least-squares fitting with the original assignment (LSF/Xu) to 2.95, 2.94 and 2.95 Hz for LSF-DMopt, LSF-LSFopt and LSF-COMopt respectively, although they are still not particularly good. The main difference between the fitted and "re-fitted" Karplus parameters is the change of sign of $a$, although only for the parameter sets re-fitted using the assignments optimised for the de Marco parameters or the combined set of assignment changes, LSF-DMopt and LSF-COMopt, is this enough to induce a small maximum in the Karplus curve at $\theta_\beta = 0°$. The similar performance of all three "re-fitted" parameter sets is due to their similarity outside of the region $-60° < \theta_\beta < 60°$, which is where most of the time-averaged dihedral-angle values lie. To illustrate this, the mean of the dihedral angles $\langle\theta_\beta\rangle$ over the simulations, not the median, is shown in Fig. 15.

## Conclusions and discussion

The variety of available parameter values for the standard Karplus relation between a $^3J_{\alpha\beta}$-coupling and the corresponding $\theta_\beta$-angle hints at an insufficient description of this relation in the form of the standard Karplus equation, Eq. 1. Indeed, the $^3J_{\alpha\beta}^{exp}$-values measured experimentally for the three proteins studied here, Plastocyanin, HEWL, and FKBP, deviate considerably from the $\langle^3J_{\alpha\beta}^{calc}\rangle$-values predicted using the commonly-used de Marco parameters and the standard Karplus relation to calculate $\langle^3J_{\alpha\beta}^{calc}\rangle$-values from the X-ray or NMR model structures or from MD simulation trajectories in vacuum or water in various force fields. Therefore, we explored two avenues for improving the relation $^3J_{\alpha\beta}(\theta_\beta)$.

The first approach of calculating the $\langle^3J_{\alpha\beta}^{calc}\rangle$-coupling constants using the generalised Karplus relation in Eq. 4 yielded at best only slightly better agreement between the $\langle^3J_{\alpha\beta}^{calc}\rangle$ and measured $^3J_{\alpha\beta}^{exp}$-coupling constants than when using the simpler Karplus relation of Eq. 1 with the de Marco parameters. Thus accounting for substituent effects is not sufficient to improve the agreement between the calculated and measured $^3J_{\alpha\beta}$-couplings for these proteins. Moreover, use of dihedral-angle values $\theta_\beta$ from X-ray or NMR model structures to calculate $\langle^3J_{\alpha\beta}^{calc}\rangle$-values leads to better agreement with the $^3J_{\alpha\beta}^{exp}$-values than when using simulation trajectories, for which the $\langle^3J_{\alpha\beta}^{calc}\rangle$-values are

averages over a variety of conformations as they are in the NMR experiments. This may be related to the fact that the parameters of the standard Karplus relation are often fitted assuming a single structure.

To investigate the effect of conformational averaging, the parameters $a$, $b$, and $c$ of the standard Karplus relation were obtained by least-squares fitting of the $\langle^3J_{\alpha\beta}^{calc}\rangle$-coupling constants averaged over MD trajectories or X-ray or NMR model structures to measured $^3J_{\alpha\beta}^{exp}$-coupling constants. The parameters $a$, $b$, and $c$ and the $Q$-values quantifying the goodness of fit depend not only on the choice of protein, but on the particular structure set of each protein that is used. It is noticeable that the shape of the fitted Karplus curves is highly dependent on the values of the dihedral angles $\theta_\beta$ used to calculate the $\langle^3J_{\alpha\beta}^{calc}\rangle$ in the fitting procedure. A general lack of sampling of angle values in the range $-60° < \theta_\beta < 60°$ means that the fitted curves are not well defined in this region, with some lacking the maximum located here when parameter sets from the literature are used.

A further factor influencing the performance of the fitting procedure is how well the relative weights of the different conformations sampled during the MD simulations match the conformational probability density in the NMR experiment. This will depend on both the quality of the force field and the degree of sampling. Indeed, it was seen that the Karplus parameters obtained from the least-squares fitting procedure are rather sensitive to the part of the simulation, i.e. subset of the conformational ensemble, to which they are fitted.

It was observed that the goodness of fit between the measured $^3J_{\alpha\beta}^{exp}$-couplings and the calculated $\langle^3J_{\alpha\beta}^{calc}\rangle$-values depends as much on the $^3J_{\alpha\beta}^{exp}$ dataset as on the choice of structure set or Karplus parameters. In particular, the $Q$-values calculated for FKBP are always rather high. Optimisation of the assignment of the stereospecifically assigned $^3J_{\alpha\beta_{2/3}}^{exp}$-couplings and re-fitting of the Karplus parameters using the optimised assignment improved the $Q$-values, but only marginally.

Overall, the present study highlights the uncertainty inherent in the parameters of the Karplus relation used to link $^3J$-couplings to dihedral-angle values and in the relation itself in the case of side-chain $\theta_\beta$-angles. Similar conclusions are expected to hold for other dihedral angles, although for those that are less mobile, fewer problems are anticipated.

We note that the $^3J_{\alpha\beta}$-coupling constants reported for Plastocyanin (Moore et al. 1991), Lysozyme (Smith et al. 1991), and FKBP (Xu et al. 1992) were measured in order to identify which $\chi_1$-angle rotamer was preferred by a particular side chain or the presence of rotational averaging. It was

found that side chains buried inside the protein usually adopt a particular rotamer, while those at the surface mostly show conformational averaging. This general conclusion does not depend on the particular parametrisation of the relation between $^3J_{\alpha\beta}$-coupling and $\chi_1$ torsional angle that is used in the analysis, thus despite the uncertainties discussed here, it is still possible to obtain some structural information from side-chain $^3J_{\alpha\beta}$-couplings.

There remain some further possibilities for improving the quality of the function $^3J_{\alpha\beta}(\theta_\beta)$. An extended generalised Karplus equation as suggested by Imai and Osawa (1990) could be applied. This takes into account more substituent effects. Given the minimal improvement seen here for the generalised Karplus relation, however, it seems unlikely that the extended version will offer significant further improvement. Another possibility would be to consider asymmetric, amino-acid specific Karplus relations as done by Schmidt (2007). The experimental data in Fig. 3 show larger measured $^3J_{\alpha\beta}^{exp}$-coupling constants for dihedral-angle values $\theta_\beta$ around $+60°$ than around $-60°$, which would support the concept of asymmetric relations. Schmidt (2007) parametrised asymmetric Karplus relations for each amino acid type using a self-consistent method (Schmidt et al. 1999). A wide spread in the parameter sets obtained for $^3J_{\alpha\beta}$-couplings for different amino acid types was observed. Other approaches to calculate different side-chain vicinal coupling constants around $\chi_1$ (Suardíaz et al. 2007) of Valine also showed the highest deviation from the experimental values when considering $^3J_{\alpha\beta}$-couplings, illustrating the particular difficulty of finding an appropriate $^3J_{\alpha\beta}(\theta_\beta)$ relation compared to other types of side-chain $^3J(\chi_1)$-coupling constants.

## References

IUPAC-IUB (1970) Commission of biochemical nomenclature: abbreviations and symbols for the description of the conformation of polypeptide chains. Biochemistry 9:3471–3479

Abraham RJ, McLauchlan KA (1962) The proton resonance spectra and conformations of the prolines Part II. The conformations of trans hydroxy-L-proline and cis (allo) hydroxy-L-proline in solution. 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 structure of the monoclinic and orthorhombic forms of hen egg-white Lysozyme at 6Å resolution. Acta Cryst B 38:778–783

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank, www.pdb.org. Nucleic Acids Res 28:235–242

Brüschweiler R, Case DA (1994) Adding harmonic motion to the Karplus relation for spin–spin coupling. J Am Chem Soc 116:11199–11200

Bystrov VF (1976) Spin–spin coupling and the conformational states of peptide systems. Prog Nucl Mag Res Sp 10:41–81

Carter D, He J, Ruble JR, Wright B (1997) The structure of the orthorhombic form of hen egg-white lysozyme at 1.5Å resolution. PDB 1AKI

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 simulations: GROMOS05. J Comput Chem 26:1719–1751

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

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

van Gunsteren WF, Gros P, Torda AE, Berendsen H, van Schaik R (1991) On deriving spatial structure from NMR or X-ray diffraction data. In: Protein conformation Wiley-Interscience, Ciba Foundation symposium, vol 161, pp. 150–166

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, Bonvin AMJJ, Daura X, Smith L (1999) Aspects of modeling biomolecular structure on the basis of spectroscopic or diffraction data. In: Krishna and Berliner (eds) Structure computation and dynamics in protein NMR, Biol Magnetic Resonance. Plenum Publishers, New York, vol 17, pp 3–35

Haasnoot CAG, de Leeuw FAAM, de Leeuw HPM, Altona C (1979) Interpretation of vicinal proton-proton coupling constants by a generalised Karplus relation. Conformational analysis of the exocyclic C4'-C5' bond in nucleosides and nucleotides. Rec J Roy Neth Chem Soc 98:576–577

Haasnoot CAG, de Leeuw FAAM, de Leeuw HPM, Altona C (1981) Relationship between proton-proton NMR coupling constants and substituent electronegativities. III. Conformational analysis of proline rings in solution using a generalized Karplus equation. Biopolymers 20:1211–1245

Hendrickson WA, Konnert JH (1981) Stereochemical restrained crystallographic least-squares refinement of macromolecule structures. In: Srimivasan R (eds) Biomolecular structure, conformation, function and evolution, vol 1: diffraction and related studies, Pergamon Press, New York, pp 43–57

Huggins ML (1953) Bond energies and polarities. J Am Chem Soc 75:4123–4126

Imai K, Osawa E (1990) An empirical extension of the Karplus equation. Magn Reson Chem 28:668–674

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

Karplus M (1963) Vicinal proton coupling in Nuclear Magnetic Resonance. J Am Chem Soc 85:2870–2871

Kopple KD, Wiley GR, Tauke R (1973) A dihedral angle-vicinal proton coupling constant correlation for the $\alpha$-$\beta$ bond of amino acid residues. Biopolymers 12:627–636

Lindorff-Larsen K, Best RB, Vendruscolo M (2005) Interpreting dynamically-averaged scalar couplings in proteins. J Biomol NMR 32:273–280

Mádi ZL, Griesinger C, Ernst RR (1990) Conformational dynamics of proline residues in Antamadine. J coupling analysis of strongly coupled spin systems based on E.COSY spectra. J Am Chem Soc 112:2908–2914

de Marco 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

Moore JM, Lepre CA, Gippler GP, Chazin WJ, Case DA, Wright PE (1991) High-resolution solution structure of reduced french bean Plastocyanin and comparison with the crystal structure of poplar Plastocyanin. J Mol Biol 221:533–555

Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 25:1656–1676

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

Pérez C, Löhr F, Rüterjans H, Schmidt JM (2001) Self-consistent Karplus parametrization of $^3J$ couplings depending on the polypeptide side-chain torsion $\chi_1$. J Am Chem Soc 123:7081–7093

Salmon L, Bouvignies G, Markwick P, Blackledge M (2011) Nuclear Magnetic Resonance provides a quantitative description of protein conformational flexibility on physiologically important time scales. Biochemistry 50:2735–2747

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

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

Schmidt JM (2007) Asymmetric Karplus curves for the protein side-chain $^3J$ couplings. J Biomol NMR 37:287–301

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

Schuler LD, Daura X, van Gunsteren WF (2001) An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. J Comput Chem 22:1205–1218

Schwalbe H, Grimshaw SB, Spencer A, Buck M, Boyd J, Dobson CM, Redfield C, Smith LJ (2001) A refined solution structure of hen Lysozyme determined using residual dipolar couling data. Protein Sci 10:677–688

Smith LJ, Sutcliffe MJ, Redfield C, Dobson CM (1991) Analysis of $\varphi$ and $\chi_1$ torsion angles from hen Lysozyme in solution from $^1$H NMR spin-spin coupling constants. Biochemistry 30:986–996

Steiner D, van Gunsteren WF (2012) An improved structural characterisation of reduced french bean Plastocyanin based on NMR data and local-elevation molecular dynamics simulation. Eur Biophys J (accepted)

Suardíaz R, Pérez C, de la Vega JMG, Fabián JS, Contreras RH (2007) Theoretical Karplus relationships for vicinal coupling constants around $\chi_1$ in valine. Chem Phys Lett 442:119–123

Vaney MC, Maignan S, Riès-Kautt M, Ducruix A (1996) High-resolution structure (1.33Å) of a HEW Lysozyme tetragonal crystal grown in the APCF apparatus. Data and structural comparison with a crystal grown under microgravity from SpaceHab-01 mission. Acta Cryst D 52:505–517

Vögeli B, Ying J, Grishaev A, Bax A (2007) Limits on variations in protein backbone dynamics from precise measurements of scalar couplings. J Am Chem Soc 129:9377–9385

Wang AC, Bax A (1996) Determination of the backbone dihedral angles $\phi$ in human Ubiquitin from reparameterized empirical Karplus equations. J Am Chem Soc 118:2483–2494

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

Xu RX, Olejniczak ET, Fesik SW (1992) Stereospecific assignment and $\chi_1$ rotamers for FKBP when bound to Ascomycin from $^3J_{H_\alpha H_\beta}$ and $^3J_{NH_\beta}$ coupling constants. FEBS Lett 305:137–143