ARTICLE

# Protein proton–proton dynamics from amide proton spin flip rates

**Daniel S. Weaver · Erik R. P. Zuiderweg**

**Abstract** Residue-specific amide proton spin-flip rates $K$ were measured for peptide-free and peptide-bound cal-modulin. $K$ approximates the sum of NOE build-up rates between the amide proton and all other protons. This work outlines the theory of multi-proton relaxation, cross relaxation and cross correlation, and how to approximate it with a simple model based on a variable number of equi-distant protons. This model is used to extract the sums of $K$-rates from the experimental data. Error in $K$ is estimated using bootstrap methodology. We define a parameter $Q$ as the ratio of experimental $K$-rates to theoretical $K$-rates, where the theoretical $K$-rates are computed from atomic coordinates. $Q$ is 1 in the case of no local motion, but decreases to values as low as 0.5 with increasing domina-tion of sidechain protons of the same residue to the amide proton flips. This establishes $Q$ as a monotonous measure of local dynamics of the proton network surrounding the amide protons. The method is applied to the study of proton dynamics in $Ca^{2+}$-saturated calmodulin, both free in solution and bound to smMLCK peptide. The mean $Q$ is $0.81 \pm 0.02$ for free calmodulin and $0.88 \pm 0.02$ for peptide-bound calmodulin. This novel methodology thus reveals the presence of significant interproton disorder in this protein, while the increase in $Q$ indicates rigidification of the proton network upon peptide binding, confirming the known high entropic cost of this process.

D. S. Weaver
Biophysics, The University of Michigan, 930 N. University Avenue, Ann Arbor, MI 48109, USA
e-mail: dswvr@umich.edu

E. R. P. Zuiderweg (✉)
Biological Chemistry, The University of Michigan, 1150 W. Medical Center Drive, Ann Arbor, MI 48109, USA
e-mail: zuiderwe@umich.edu

## Introduction

Several decades of NMR relaxation studies and computa-tional molecular dynamics studies have firmly established that proteins are involved in dynamical processes spanning many time scales (Mittermaier and Kay 2006). Much of the current attention is directed towards milli- and micro-sec-ond motions (Loria et al. 1999), since those processes are often directly related to functional conformational changes (Wang et al. 2001; Eisenmesser et al. 2005; Tang et al. 2006).

NMR spin relaxation generated by sub-nanosecond dynamics currently attracts somewhat less attention. The likely reason is that the most accessible of sub-nanosecond NMR dynamics measurements, $^{15}N$ relaxation, appears to disclose only the comparatively restricted motions of the protein backbone. More precisely, $^{15}N$ relaxation discloses only the angular motion of the imaginary vectors perpen-dicular to the NH bond-vectors in the protein backbone (Fischer et al. 1997). In contrast, protein sidechains display extensive dynamics at the sub-nano second timescale. Using methyl-relaxation methods (Yang et al. 1998), Wand et al. (Lee et al. 2000) have shown that the amplitudes of the motions of (mostly buried) protein sidechains in the protein calmodulin are large and highly disperse. Modern computational molecular dynamics studies concur with these findings (Best et al. 2005) and indicate, according to fundamental physico-chemical rules (Lindemann 1910), that some protein interiors classify rather as a liquid than as a solid, except for the backbone (Zhou et al. 1999; Best et al. 2005).

The widespread sub-nanosecond mobility in proteins indicated by sidechain dynamics experiments (Wand et al. 1995; Millet et al. 2002) cannot help but serve as a reservoir of configurational entropy to stabilize the folded protein relative to its highly entropic unfolded state. Local motional hotspots may serve as joints for slower domain motions (Labeikovsky et al. 2007) or allow induced fit in intermolecular interaction, including substrate recognition. Global changes in sub-nanosecond dynamics have been related (Akke et al. 1993; Yang and Kay 1996) to sizable entropic penalties to the affinity of ligand–protein binding (Frederick et al. 2007) and to allosteric free energy in positive cooperativity (Cooper and Dryden 1984; Mäler et al. 2000). Hence (changes in), sub-nano second dynamics have high relevance to biological functionality.

In this report, we refine our NMR method (Weaver and Zuiderweg 2008) to probe the sub-nanosecond dynamics of protein sidechains in the vicinity of the amide proton. Briefly, we measure the cross relaxation between the two components of $NH$ longitudinal spin order of the amide-nitrogen spin system: $N_zH_a$ and $N_zH_b$. The cross relaxation between these states is dominated by zero quantum spin-flip interactions of the amide proton with nearby other protons. We refer to the relaxation rate governing this process as $K$. The pulse sequence (Weaver and Zuiderweg 2008) allows the measurement of both $\eta_z$ and $K$. In our previous paper (Weaver and Zuiderweg 2008), we focused on the extraction of $\eta_z$ out of the multi-exponential cross relaxation rates. Here, we carry out the opposite: we extract from the data precise $K$-rates using mathematical models that can take into account up to 60 equidistant neighboring protons. This approximate model was developed to make $K$-rate extraction and statistical analysis computationally feasible for large systems with many protons.

We record and analyze data for two samples of human calmodulin, one without and one with the peptide (smMLCKp) corresponding to the calmodulin binding domain of the smooth muscle myosin light chain kinase (Vogel 1994). Obtained $K$-rates vary widely for the different amide protons in both samples. This is not surprising, since each amide proton has a unique environment of other protons. The resulting $K$-rates are directly related to the classical NOE build-up rates, and can be predicted from the (precise) coordinates of the structures in the protein data base. We use this property to show that our method of extracting the $K$-rates with a model based on equidistant protons is precise and accurate.

Detailed analysis of our experimental data shows that amide proton relaxation environments dominated by other backbone protons are rather rigid, and that those dominated by side chain protons are more mobile. This follows common sense (Best et al. 2005), hence, it establishes $Q$ as a monotonous measure of local dynamics of the proton network surrounding the amide protons. Nevertheless, $Q$ ratios cannot (yet) be formally related to order parameters. This has as a disadvantage that it is not possible to associate a single precise motion with $Q$. This same limitation has also a positive side: $Q$ is not so "myopic" as a real order parameter which reports on the motions of a single vector type.

The average ratio $Q$ of measured and predicted $K$-rates, denoted by $<Q>$, is found to be smaller than one, for both liganded and unliganded calmodulin. Excitingly, we also find that $<Q>$ is larger for liganded than for unliganded calmodulin, suggesting that the motions are quenched upon ligand binding. This result corresponds with the conclusions of Wand et al. who studied the sub-nanosecond dynamical differences between the two states for the same protein/peptide complex using NMR methyl-director dynamics experiments (Lee et al. 2000).

Our experimental method (Weaver and Zuiderweg 2008) is based on $^{15}$N–$^1$H TROSY detection. Since $K$-rates become faster with increasing molecular size, the method holds considerable promise for investigation of sub-nanosecond dynamics in larger proteins.

## Materials and methods

We carried out $\eta_z/\kappa$ TROSY experiments (Weaver and Zuiderweg 2008) for a 1 mM sample of calcium-saturated chicken calmodulin in 10% D$_2$O and for a 1 mM sample of calcium-saturated chicken calmodulin in complex with the calmodulin-binding domain of chicken gizzard smooth muscle myosin light chain kinase (CaM/smMLCKp), also in 10% D$_2$O.

Each $\eta_z/\kappa$ experiment consisted of acquiring sets of four two-dimensional $^1$H$_N$–$^{15}$N TROSY-type spectra at several different time points. $\eta_z/\kappa$ experiments were run at 35°C on a Bruker Avance II 500 MHz spectrometer equipped with a triple-resonance gradient probe. The $\eta_z/\kappa$ experimental pulse sequence is essentially an S$^3$E-filtered clean TROSY sequence with relaxation decays included. The S$^3$E filter selects either $\sqrt{2}N_zH_\alpha$ or $\sqrt{2}N_zH_\beta$ to pass through to the relaxation period $\tau$. Relaxation diffuses the initial pure state of $\sqrt{2}N_zH_\alpha$ or $\sqrt{2}N_zH_\beta$ into a mixture of $\sqrt{2}N_zH_\alpha$ and $\sqrt{2}N_zH_\beta$ magnetization. The end of the experiment selects for either the $\sqrt{2}N_zH_\alpha$ and $\sqrt{2}N_zH_\beta$ component. When the $\sqrt{2}N_zH_\alpha$ is selected, a TROSY readout follows. When the $\sqrt{2}N_zH_\beta$ component is selected, it is first converted to $\sqrt{2}N_zH_\alpha$ before the TROSY readout follows. The experiment is carried out corresponding to the four symmetric reconversion (Pelupessy et al. 2003) sub-experiments: (I) $\sqrt{2}N_zH_\alpha$ select $\rightarrow$ $\sqrt{2}N_zH_\alpha$ detect; (II) $\sqrt{2}N_zH_\alpha$ select $\rightarrow$ $\sqrt{2}N_zH_\beta$ detect; (III) $\sqrt{2}N_zH_\beta$ select $\rightarrow$ $\sqrt{2}N_zH_\beta$ detect; (IV) $\sqrt{2}N_zH_\beta$ select $\rightarrow$ $\sqrt{2}N_zH_\alpha$ detect.

$\eta_z/\kappa$ experiments at 35°C and 500 MHz were conducted with 16 scans per spectrum, with 160 complex $t_1$ values, at 17 relaxation time points (20, 30, 40,60, 60, 80, 90, 100, 120, 120, 140, 150, 160, 180, 180, 200, 210 ms). 160 complex $t_1$, with a recovery delay of 1 s. Experimental time was about 1 h 50 min per spectrum leading to a total experiment time of 270 h over both samples.

The first step in fitting the data is extracting $\kappa$ rates and initial state purity information (see "Theory") by a Levenberg–Marquardt nonlinear least squares fit of the symmetric reconversion ratio (Pelupessy et al. 2003)

$$X(t) = \sqrt{\frac{II(t)III(t)}{I(t)IV(t)}} \tag{1}$$

These rates are then incorporated into a second nonlinear least squares fit of residue-specific longitudinal $^{15}$N CSA/$^{15}$N–$^{1}$H dipolar cross correlation ($\eta_z$) rate and relative efficiency of state observation for $N_zH_a$ and $N_zH_b$ using the ratio

$$Y(t) = \frac{IV(t)}{I(t)} \tag{2}$$

as described in Weaver and Zuiderweg (2008).

In order to determine the rotational diffusion tensor, the $\eta_z$ rates were complemented with transverse $^{15}$N CSA/$^{15}$N–$^{1}$H dipolar cross correlation ($\eta_{xy}$) rates, determined using techniques as described by Wang et al. (2003) using a recovery delay of 1 s, 16 scans, and 160 complex points in $t_1$, with five relaxation time points at 5.55, 11.1, 16.65, 22.2 and 27.75 ms.

Experimental site-specific rotational diffusion constants were calculated from Kroenke et al. (1998)

$$D_i^{\text{exp}} = \frac{\omega_N}{3}\left(6\frac{\eta_{xy}}{\eta_z} - 7\right)^{-1/2} \tag{3}$$

Next, the site-specific rotational diffusion constants were plotted against the function $\frac{1}{2}(3\cos^2\theta_i - 1) \equiv P_2\cos\theta_i$ to

obtain the fitted site-specific rotational diffusion constants as defined by

$$D_i^{\text{Fitted}} = D_{\text{iso}} - \frac{(D_P - D_\perp)}{3}P_2\cos\theta_i \tag{4}$$

Here, $\Theta_i$ is the angle of the principle axis of the $^{15}$N CSA tensor with the principal axis of the (axially symmetric) diffusion tensor. PDB orientation and diffusion tensor fits were carried out using the programs pdb_inertia and quadric_diffusion (Palmer 2009). The diffusion tensor fits are shown in Fig. 1, and the fitted values are shown in Table 1. For free calmodulin, it was necessary to use different diffusion tensors for the N and C-terminal domains. smMLCKp—bound calmodulin was found to diffuse nearly isotropically, with a $\tau_c$ of 7.7 ns. All obtained values are slightly smaller than those obtained for the same protein from $^{15}$N $R_2/R_1$ values (Lee et al. 2000). This is to be expected, since the average $R_2$ value is systematically increased by exchange broadening (Lee et al. 2000).

In the further data analysis, the fitted site-specific rotational diffusion constants were used in the orientation dependent spectral density functions

$$J_i(\omega) = \frac{2}{5}\frac{\tau_{Ci}}{1 + (\omega\tau_{C_i})^2} \tag{5}$$

where $\tau_{C_i} = 1/6D_i^{\text{Fitted}}$.

The calculation of $Q$ depends on a proton–proton distance benchmark provided by a computationally protonated X-ray structure. Structures used in this study are a 1.08 Å resolution structure of Ca$^{2+}$-saturated smMLCKp-bound calmodulin (2O5G) (Valentine et al. 2009) and a 1.0 Å resolution structure of Ca$^{2+}$-saturated free calmodulin (1EXR) (Wilson and Brunger 2000). For the purposes of this work 1EXR was edited into N-terminal (1–76) and C-terminal (82–148) domains and the rotational diffusion tensor of each domain was determined independently. Protonation was accomplished with Mol-Probity (Lovell et al. 2003). Methyl group correction for rapid motional averaging was performed according to the



Fig. 1 Determination of the rotational diffusion constants. *Left* free calmodulin with closed symbols for the N terminus, and open symbols for the C terminus. Lines are fits from the program Quadric (Lee et al. 1997). *Solid* and *dashed lines* are the fit for the N and C terminus, respectively. *Right* peptide-bound calmodulin (N terminus = residues 1–76; C terminus = residues 82–148)
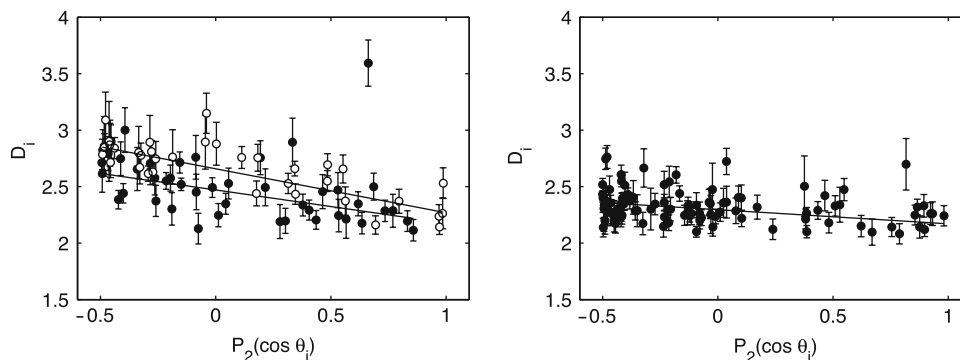
**Table 1** Rotational correlation times for calmodulin, 308°K

|  | $\tau_c \parallel$ (ns) | $\tau_c \perp$ (ns) |
|---|---|---|
| CaM N-term free | $5.5 \pm 0.2$ | $7.7 \pm 0.2$ |
| Cam C-term free | $4.9 \pm 0.3$ | $7.3 \pm 0.3$ |
| CaM bound | $6.4 \pm 0.2$ | $7.8 \pm 0.1$ |

method of Tropp (1980) as detailed in Neuhaus and Williamson (2000).

Following (Cruickshank 1999), the precision of the position of an atom of 'typical' B-factor in an X-ray structure is (very) approximately

$$\sigma(r, B_{avg}) = 3^{1/2}(N_i/p)^{1/2}C^{-1/3}R\,d_{min} \tag{6}$$

where $N_i$ is the total number of atoms in the structure, $p$ is $n_{obs} - n_{params}$, $n_{obs}$ is the number of reflections, $n_{params}$ is the number of parameters ($\sim 4n_{atoms}$ for isotropic refinement, $\sim 9n_{atoms}$ for fully anisotropic refinement as in 1EXR and 2O5G), $C$ is the completeness for range of the data used in refinement, $R$ is the crystallographic $R$ value, and $d_{min}$ is the resolution of the structure. Assuming that the 'typical' B-factor can be used for computationally added protons, that any additional error introduced by MolProbity protonation is subsumed in Cruickshank's generous error estimation, and that distance errors between protons are uncorrelated (6), leads to a point-to-point distance error estimate for X-ray structures of

$$\sigma(r, B_{avg}) = 6^{1/2}(N_i/p)^{1/2}C^{-1/3}R\,d_{min} \tag{7}$$

based on propagating the error of the difference between two position estimates. A point-to-point distance precision of $\sim 0.055$ Å for 1EXR and $\sim 0.075$ Å for 2O5G are indicated by (7). If proton–proton distances are assumed to be a Gaussian distribution with a mean established by the protonated X-ray structure and a standard deviation given by the point-to-point distance error given above, we expect X-ray structural imprecision to contribute $\sim 5.5\%$ error in theoretical $K$ calculations for the 1.00 Å 1EXR structure and $\sim 7.5\%$ error in theoretical $K$ calculations for the 1.08 Å 2O5G structures. When calculating $Q$, the error introduced by this evaluation of theoretical $K$ imprecision is larger by a factor of 1.8 than the error introduced by imprecision in experimental $K$ fitting.

**Theory**

I. Rigorous description of the relaxation network

Consider the longitudinal magnetization of an amide $NH$ spin system in a protein at the large-molecule limit. We prepare $\sqrt{2}N_zH_a$ and $\sqrt{2}N_zH_b$ magnetization from linear combinations of nitrogen one-spin order $N_z$ and two-spin order $2N_zH_z$.

$$\sqrt{2}N_zH_\alpha = \frac{\sqrt{2}}{2}(N_z + 2N_zH_z)$$
$$\sqrt{2}N_zH_\beta = \frac{\sqrt{2}}{2}(N_z - 2N_zH_z) \tag{8}$$

These are the proton-up and proton-down components of nitrogen $Z$ magnetization, respectively. We will refer to them collectively as LCLO major states, LCLO standing for linear combinations of longitudinal order. We are interested in modeling the auto- and cross-relaxation behavior of the LCLO major states. For an isolated amide two-spin system, auto and cross-relaxation behavior are described by the following Liouville representation:

$$\frac{d}{dt}\begin{bmatrix} \sqrt{2}N_zH_\alpha \\ \sqrt{2}N_zH_\beta \end{bmatrix} = -\begin{bmatrix} \lambda + \eta_z & -\zeta \\ -\zeta & \lambda - \eta_z \end{bmatrix}\begin{bmatrix} \sqrt{2}N_zH_\alpha \\ \sqrt{2}N_zH_\beta \end{bmatrix} \tag{9}$$

with

$$\lambda = \frac{d_{NH}^2}{8}\begin{pmatrix} 6J(\omega_N) + J(\omega_H - \omega_N) \\ +3J(\omega_H) + 6J(\omega_H + \omega_N) \end{pmatrix} + \frac{c_N^2}{3}J(\omega_N)$$
$$+ \frac{c_N^2}{6}J(\omega_H) \tag{10}$$

and

$$\eta_z = c_N d_{NH} J_X^{NH/N}(\omega_N) \tag{11}$$

and

$$\xi = -\frac{d_{NH}^2}{8}(J(\omega_H - \omega_N) - 3J(\omega_H) + 6J(\omega_H + \omega_N))$$
$$- \frac{c_H^2}{6}J(\omega_H) \tag{12}$$

where $c_X = B_0\gamma_X\Delta_X$, $d_{XY} = (\mu_0/4\pi)\gamma_X\gamma_Y\,(h/2\pi)r_{XY}^{-3}$, $B_0$ is the spectrometer static magnetic field, $\gamma_X$ is the gyromagnetic moment of nucleus $X$, $\Delta_X$ is the magnitude of the axial CSA of nucleus $X$, $\mu_0$ is the permeability of free space, $r_{XY}$ is the distance between nuclei $X$ and $Y$, $\omega_X$ is the Larmor frequency for nucleus $X$, $J(\omega_X)$ is the autocorrelation spectral density function for angular motion of a particular vector at Larmor frequency $X$, and $J_X^{XY/X}(\omega)$ is the cross-correlation spectral density function for combined motion of the internuclear vector between nuclei $X$ and $Y$ and the CSA principal axis of nucleus $X$ at Larmor frequency $\omega$.

For the isolated two-spin system, all relaxation pathways are dominated by spectral densities at frequencies $\omega_N$ or larger, and the relevant rates are on the order of 0.01–0.1 s$^{-1}$. However, the LCLO states experimentally (see "Results") cross-relax at rates one to two orders of magnitude faster due to interactions with extraneous protons.
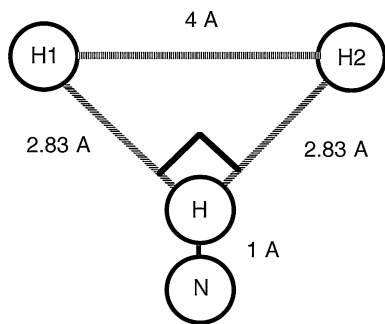
**Fig. 2** Relaxation network corresponding to the relaxation matrices and Table 2 in the theory section of this work. Note that a much extended network with up to 60 neighbor protons is used for the actual data fitting

## Ia. Symbolic representation of the relaxation process

The explanation for this fast cross-relaxation behavior lies in the network of protons surrounding the *NH* spin system. The presence of these protons, each of which may be in an *a* or *b* state, splits each of the two LCLO major states into several substates. $\sqrt{2}N_zH_a$ splits into $N_zH_aH'_aH''_a...$, $N_zH_aH'_aH''_b...$, $N_zH_aH'_bH''_a...$ and so on, while $\sqrt{2}N_zH_b$ splits into $N_zH_bH'_aH''_a...$, $N_zH_bH'_aH''_b...$, $N_zH_bH'_bH''_a...$ and so on (for clarity of notation, we will neglect the normalizing coefficients of LCLO substates). The total population in each LCLO state is the sum of their substate populations (see Fig. 3).

Figure 2 depicts an amide *NH* spin system with two neighbor protons, and Fig. 3 shows the associated energy level diagram. As is apparent from Fig. 3, the inclusion of the proton network surrounding the *NH* spin system presents another pathway for cross-relaxation between the LCLO major states. This pathway consists of zero-quantum 'spin flip' transitions between one proton in a given spin state and a neighbor proton in the opposite spin state. For example, $N_zH_aH'_bH''_a...$ may undergo a zero-quantum transition to $N_zH_bH'_aH''_a...$, or to $N_zH_aH'_aH''_b...$. For LCLO network protons *i* and *j*, zero-quantum cross-relaxation will take place between states $N_z...H_{ia}...H_{jb}...$ and $N_z...H_{ib}...H_{ja}...$. Density matrix calculations give the relaxation rate governing this spin-flip process as:

$$\kappa_{ij} = \frac{1}{4}d^2_{H_iH_j}J_{H_iH_j}(0) \tag{13}$$

where the spectral density function $J_{H_iH_j}(0)$ is

$$J_{H_iH_j}(0) = S^2R\frac{2}{5}\tau_c \tag{14}$$

as defined in Post (1992) and Brüschweiler et al. (1992). This definition of $J_{H_iH_j}(0)$ incorporates the effects of radial (*R*) and angular ($S^2$) motion of the relaxing protons *i* and *j* with respect to one another, assuming that these

motions are not correlated. This is a necessary extension in situations where the relaxing nuclei are not constrained by a connecting chemical bond. Further information on systems of this kind can be found in Olejniczak et al. (1984).

The zero-quantum transition rate defined in (13) is $\sim 1$–$3\ s^{-1}$ for close interproton distances in proteins with characteristic global diffusion times $\tau_c = 4$–$7$ ns, and faster for larger proteins with slower global diffusion times. These relaxation rates match the order of magnitude of experimentally observed LCLO cross-relaxation behavior. We will first develop the fitting model in the large molecule/high magnetic field limit, defined such that zero-quantum relaxation terms dwarfs all other relaxation mechanisms. Subsequently, we will bring back all small omitted terms and assess their cumulative effect on the relaxation processes.

In this work, we use the $\eta_z/\kappa$ pulse sequence (Weaver and Zuiderweg 2008) incorporating symmetrical reconversion techniques (Pelupessy et al. 2003). As described in the original publication describing the experiment (Weaver and Zuiderweg 2008), the sequence allows for the separation of $\eta_z$ and $\kappa$. In the following, we assume that this separation has been made; hence we will exclusively focus on the other mechanisms that contribute to $\kappa$.

## Ib. Cross-relaxation and cross-correlated cross-relaxation

We now construct the system of first order differential equations governing cross-relaxation in the LCLO substate system depicted in Fig. 2, considering only zero-quantum spin-flip terms in the relaxation matrix. In the following we denote the amide proton as $H^A$, the first neighbor proton as $H^1$, and the second neighbor proton as $H^2$. The matrix is related to a NOE relaxation matrix, in which the diagonal $R_1$ terms and "leakage" terms are removed. Rate matrices of this type distribute initial perturbations of the Boltzmann distribution throughout the system, without driving the system back to equilibrium. Below, we will refine the LCLO matrix by adding additional terms that do drive the system back to thermodynamic equilibrium.

$$\frac{d}{dt}\begin{bmatrix} N_zH^A_\alpha H^1_\alpha H^2_\alpha \\ N_zH^A_\alpha H^1_\alpha H^2_\beta \\ N_zH^A_\alpha H^1_\beta H^2_\alpha \\ N_zH^A_\alpha H^1_\beta H^2_\beta \\ N_zH^A_\beta H^1_\alpha H^2_\alpha \\ N_zH^A_\beta H^1_\alpha H^2_\beta \\ N_zH^A_\beta H^1_\beta H^2_\alpha \\ N_zH^A_\beta H^1_\beta H^2_\beta \end{bmatrix} = -\bar{R}\begin{bmatrix} N_zH^A_\alpha H^1_\alpha H^2_\alpha \\ N_zH^A_\alpha H^1_\alpha H^2_\beta \\ N_zH^A_\alpha H^1_\beta H^2_\alpha \\ N_zH^A_\alpha H^1_\beta H^2_\beta \\ N_zH^A_\beta H^1_\alpha H^2_\alpha \\ N_zH^A_\beta H^1_\alpha H^2_\beta \\ N_zH^A_\beta H^1_\beta H^2_\alpha \\ N_zH^A_\beta H^1_\beta H^2_\beta \end{bmatrix} \tag{15}$$

where *R* is the relaxation matrix

$$\bar{R} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & K_{12}+K_{A2} & -K_{12} & 0 & -K_{A2} & 0 & 0 & 0 \\ 0 & -K_{12} & K_{12}+K_{A1} & 0 & -K_{A1} & 0 & 0 & 0 \\ 0 & 0 & 0 & K_{A1}+K_{A2} & 0 & -K_{A1} & -K_{A2} & 0 \\ 0 & -K_{A2} & -K_{A1} & 0 & K_{A1}+K_{A2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -K_{A1} & 0 & K_{12}+K_{A1} & -K_{12} & 0 \\ 0 & 0 & 0 & -K_{A2} & 0 & -K_{12} & K_{12}+K_{A2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{16}$$

Terms equal to the negative sum of the off-diagonal cross-relaxation terms have been added to the diagonal (auto-relaxation) rate entry of the relaxation matrix for each LCLO substate in order to preserve detailed balance.

In order to clearly delineate the difference between true auto-relaxation terms and diagonal terms which arise as a result of detailed balance of cross-relaxation, the detailed balance terms will be added during consideration of the cross-relaxation elements of the matrix. Although they are located on the diagonal of the matrix, they are not true auto-relaxation processes.

As can be seen from the abbreviated form of the matrix, LCLO substates may be conveniently represented by binary numbers, with 0 and 1 s replacing $a$'s and $b$'s. Bitwise logic operations may be used to explore the operational relaxation pathways between LCLO substates. For instance, if the bitwise XOR of the binary representations of two LCLO substates contains two and only two 1's, $a$ 0 in the first state has changed into $a$ 1 in the second state ($a$

to $b$) while $a$ 1 in the first state has changed into $a$ 0 in the second state ($b$ into $a$). This indicates the presence of a zero-quantum spin flip relaxation pathway between these two states. Shortcuts such as these are convenient in the computational treatment of LCLO relaxation.

Examination of this relaxation matrix (16) leads to the following observations:

$a$   When only the zero-quantum terms are considered, the relaxation matrix may be blocked into submatrices representing isoenergetic levels in Fig. 3 (labeled block 0, block 1, etc.). Since only zero-quantum terms are being considered, each block exchanges magnetization only within itself. Each block is characterized by a constant number of $a$ and $b$ spins, corresponding to the sum of the spin quantum number of all protons considered, $m_J$. Zero-quantum relaxation processes conserve the total number of $a$ and $b$ spins while rearranging their location within the proton network. The energetic blocks in (16) are as follows:
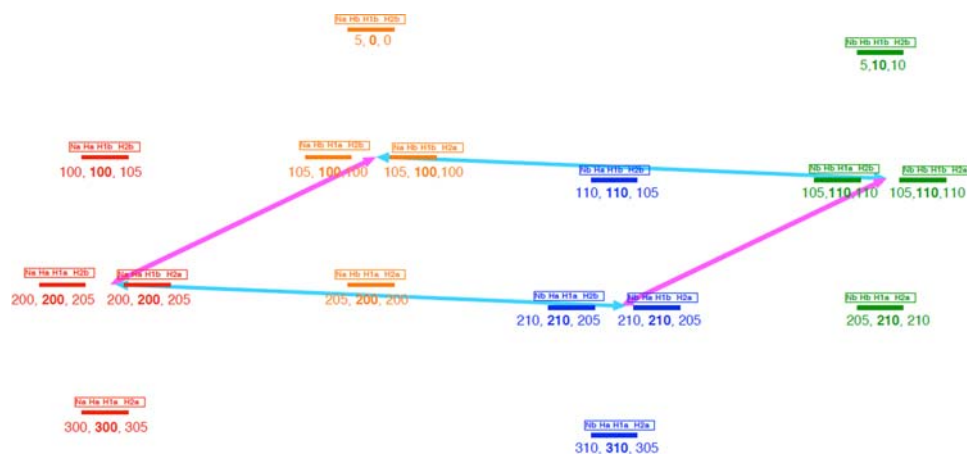


**Fig. 3** Definition of the energy levels in the proton relaxation network shown in Fig. 2, and used for the relaxation matrices and Table 2 in the theory section of this work. Note that a much extended network with upto 60 neighbor protons is used for the actual data fitting. The NaHa levels are in *red*; NaHb levels in *orange*; NbHa in *blue*; NbHb in *green*. The *cyan arrows* symbolize 4 $^{15}$N transitions each; the *magenta arrows* symbolize four amide hydrogen transitions each. The *numbers* indicate the relative populations in the $N_zH_a$, equilibrium and $N_zH_b$ state, respectively

Block 0 : $\dfrac{\mathrm{d}}{\mathrm{d}t} N_z H_\alpha^A H_\alpha^1 H_\alpha^2 = 0$

Block 1 : $\dfrac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} N_z H_\alpha^A H_\alpha^1 H_\beta^2 \\ N_z H_\alpha^A H_\beta^1 H_\alpha^2 \\ N_z H_\beta^A H_\alpha^1 H_\alpha^2 \end{bmatrix} = -\begin{bmatrix} K_{12}+K_{A1} & -K_{12} & -K_{A1} \\ -K_{12} & K_{12}+K_{A2} & -K_{A2} \\ -K_{A1} & -K_{A2} & K_{A1}+K_{A2} \end{bmatrix}\begin{bmatrix} N_z H_\alpha^A H_\alpha^1 H_\beta^2 \\ N_z H_\alpha^A H_\beta^1 H_\alpha^2 \\ N_z H_\beta^A H_\alpha^1 H_\alpha^2 \end{bmatrix}$

(17)

Block 2 : $\dfrac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} N_z H_\alpha^A H_\beta^1 H_\beta^2 \\ N_z H_\beta^A H_\alpha^1 H_\beta^2 \\ N_z H_\beta^A H_\beta^1 H_\alpha^2 \end{bmatrix} = -\begin{bmatrix} K_{A1}+K_{A2} & -K_{A1} & -K_{A2} \\ -K_{A1} & K_{12}+K_{A1} & -K_{12} \\ -K_{A2} & -K_{12} & K_{12}+K_{A2} \end{bmatrix}\begin{bmatrix} N_z H_\alpha^A H_\beta^1 H_\beta^2 \\ N_z H_\beta^A H_\alpha^1 H_\beta^2 \\ N_z H_\beta^A H_\beta^1 H_\alpha^2 \end{bmatrix}$

Block 3 : $\dfrac{\mathrm{d}}{\mathrm{d}t} N_z H_\beta^A H_\beta^1 H_\beta^2 = 0$

*b* A system containing $N$ neighbor protons contains $2^{N+1}$ substates in total, arranged among $N+2$ energy level blocks. Let us refer to these blocks as Block 0, Block 1, Block 2, …, Block $N+1$, as above. In our labeling system, block 0 contains the single substate with all protons in the *a* state, block 1 contains substates having one *b* proton, block 2 contains substates having two *b* protons, and so on, all the way up to block $N+1$ which contains the single substate having $N+1$ *b* protons. In general, block $m_J$ contains $\begin{pmatrix} N+1 \\ m_J \end{pmatrix}$ substates, $\begin{pmatrix} N \\ m_J \end{pmatrix}$ of which will be *a* substates and $\begin{pmatrix} N \\ m_J-1 \end{pmatrix}$ of which will be *b* substates. Here we use the common definition $\begin{pmatrix} N \\ m_J \end{pmatrix} \equiv \frac{N!}{(N-m_J)! m_J!}$.

We next consider several other terms that contribute to the cross-relaxation of the amide *NH* LCLO substate network. These terms are united in the fact that they all contain spectral density functions on the order of $J(\omega_H)$. These spectral density functions are at least two orders of magnitude smaller in magnitude than the zero-frequency

spectral density functions $J(0)$ that govern the zero-quantum spin flip transitions. For this reason, the non-$J(0)$ terms that we will examine here should be considered second-order contributions to the overall relaxation behavior of the system.

Single-quantum homonuclear proton–proton dipole–dipole interactions (hereafter abbreviated SQHH interactions) cause LCLO substates to exchange a single *a* spin for a single *b* spin, or vice versa. For instance, $N_z H_a H'_b H''_a \ldots$ may transition to $N_z H_b H'_b H''_a \ldots$ or to $N_z H_a H'_a H''_a \ldots$ due to SQHH interactions. In general, SQHH interactions operating on the $i$th proton in the LCLO system connect states $N_z \ldots H_{ia} \ldots$ and $N_z \ldots H_{ib} \ldots$. We may sum the SQHH interactions for the state-switching proton $i$ over every other proton $j$ in the system, leading to a total rate:

$$\mathrm{SQHH}_i = \sum_j \mathrm{SQHH}_{ij} = \sum_j \frac{3}{8} d^2_{H_i H_j} J_{H_i H_j}(\omega_H) \qquad (18)$$

In order to preserve legibility, we label SQHH interactions with the single letter *A* when illustrating their placement in the Liouville relaxation matrix. SQHH interactions enter the Liouville-space relaxation matrix as:

$$\bar{R} = \begin{bmatrix} \sum A & -A_2 & -A_1 & 0 & -A_A & 0 & 0 & 0 \\ -A_2 & K_{12}+K_{A2}+\sum A & -K_{12} & -A_1 & -K_{A2} & -A_N & 0 & 0 \\ -A_1 & -K_{12} & K_{12}+K_{A1}+\sum A & -A_2 & -K_{A1} & 0 & -A_A & 0 \\ 0 & -A_1 & -A_2 & K_{A1}+K_{A2}+\sum A & 0 & -K_{A1} & -K_{A2} & -A_N \\ -A_N & -K_{A2} & -K_{A1} & 0 & K_{A1}+K_{A2}+\sum A & -A_2 & -A_1 & 0 \\ 0 & -A_A & 0 & -K_{A1} & -A_2 & K_{12}+K_{A1}+\sum A & -K_{12} & -A_1 \\ 0 & 0 & -A_A & -K_{A2} & -A_1 & -K_{12} & K_{12}+K_{A2}+\sum A & -A_2 \\ 0 & 0 & 0 & -A_A & 0 & -A_1 & -A_2 & \sum A \end{bmatrix}$$

(19)

wherein $\sum A = \sum_i A_i$.

Heteronuclear NOE contributions to the LCLO relaxation matrix have already been mentioned. Heteronuclear NOE processes connect LCLO substates along the same pathways as SQHH interactions, allowing cross-relaxation between single $a$ spins and single $b$ spins, or vice versa. For instance, $N_z H_a H'_b H''_a \ldots$ may transition to $N_z H_b H'_b H''_a \ldots$ or to $N_z H_a H'_A H''_a \ldots$ due to heteronuclear NOE interactions.

$$B_i = \sum_j B_{ij} = \sum_j \frac{1}{8} d^2_{H_i X_j} \left( J_{H_i X_j}(\omega_H - \omega_X) - 3J_{H_i X_j}(\omega_H) \right.$$
$$\left. + 6J_{H_i X_j}(\omega_H + \omega_X) \right) \tag{20}$$

When these summed heteronuclear NOE interactions are added to the Liouville relaxation matrix for the LCLO substate system, the relaxation matrix becomes:

$$\bar{R} = \begin{bmatrix} \sum A + \sum B & -A_2 + B_2 & -A_1 + B_1 & 0 & -A_A + B_A & 0 & 0 & 0 \\ -A_2 + B_2 & K_{12} + K_{A2} + \sum A + \sum B & -K_{12} & -A_1 + B_1 & -K_{A2} & -A_A + B_A & 0 & 0 \\ -A_1 + B_1 & -K_{12} & K_{12} + K_{A1} + \sum A + \sum B & -A_2 + B_2 & -K_{A1} & 0 & -A_A + B_A & 0 \\ 0 & -A_1 + B_1 & -A_2 + B_2 & K_{A1} + K_{A2} + \sum A + \sum B & 0 & -K_{A1} & -K_{A2} & -A_A + B_A \\ -A_A + B_A & -K_{A2} & -K_{A1} & 0 & K_{A1} + K_{A2} + \sum A + \sum B & -A_2 + B_2 & -A_1 + B_1 & 0 \\ 0 & -A_A + B_A & 0 & -K_{A1} & -A_2 + B_2 & K_{12} + K_{A1} + \sum A + \sum B & -K_{12} & -A_1 + B_1 \\ 0 & 0 & -A_A + B_A & -K_{A2} & -A_1 + B_1 & -K_{12} & K_{12} + K_{A2} + \sum A + \sum B & -A_2 + B_2 \\ 0 & 0 & 0 & -A_A + B_A & 0 & -A_1 + B_1 & -A_2 + B_2 & \sum A + \sum B \end{bmatrix} \tag{21}$$

However, there is a complication caused by the mixed-basis representation presented here, in which heteronuclear NOE processes are operating between heteroatom spins in the $z$ state and proton spins in the $a/b$ state. This causes the familiar double-quantum/zero-quantum relaxation processes produced by the heteronuclear NOE to be split across cross- and auto-relaxation entries in the Liouville-space relaxation matrix.

For each proton in the relaxation network, we may sum all NOE interactions with heteronuclei over the entirety of the protein. This sum yields the total rate at which NOE-induced spin transitions take place for each individual proton in the network.

The structure of heteronuclear NOE-type interactions in the Liouville relaxation matrix consists of a split of equal terms (here abbreviated $B$) between the auto and cross relaxation components of a given single proton $a$–$b$ state pair. For instance: the state $N_z H_a H'_a H''_b$ will have an NOE term $B$ with which it relaxes to itself (autorelaxation along the diagonal into $N_z H_a H'_a H''_b$) and another NOE term $B$ with which it enters each other state available by a single-quantum proton transition. These terms $B$ both have the same sign.

The heteronuclear NOE term $B_i$ affecting the state-switching behavior of the $i$th proton is the sum over all $j$ of the $i$th state-switching proton's heteronuclear NOE interaction with the $j$th heteronucleus (atom type $X$):

where $\sum B = \sum_i B_i$.

Single-quantum proton CSA relaxation operates along the same relaxation pathways as the previous two relaxation mechanisms, allowing cross-relaxation between single $a$ spins and single $b$ spins. It depends on the strength of the proton CSA, however, a quantity much less rigorously quantified than the dipolar interaction distances on which previous relaxation mechanisms have been based. A CSA of 10 ppm for the amide proton is consistent with orders of magnitude reported in the literature and will serve to illustrate the size of the interaction.

Furthermore, since proton CSA relaxation is based on a monic and not a dyadic interaction, the size of the proton CSA relaxation rate does not scale with the number of protons in the system. This causes proton CSA relaxation to be a particularly minor contributor to LCLO relaxation; in general, relaxation rates that do not scale with the number of other atoms in the protein should be considered third-order in the LCLO system. For the $i$th proton in the LCLO network, the proton CSA relaxation (here abbreviated $C_i$)

$$C_i = \frac{1}{6} c^2_{H_i} J(\omega_H) \tag{22}$$

is added to the Liouville relaxation matrix as follows:

$$\bar{R} = \begin{bmatrix} \cdots & -A_2 + B_2 - C_2 & -A_1 + B_1 - C_1 & 0 & -A_N + B_N - C_N & 0 & 0 & 0 \\ -A_2 + B_2 - C_2 & \cdots & -K_{12} & -A_1 + B_1 - C_1 & -K_{N2} & -A_N + B_N - C_N & 0 & 0 \\ -A_1 + B_1 - C_1 & -K_{12} & \cdots & -A_2 + B_2 - C_2 & -K_{N1} & 0 & -A_N + B_N - C_N & 0 \\ 0 & -A_1 + B_1 - C_1 & -A_2 + B_2 - C_2 & \cdots & 0 & -K_{N1} & -K_{N2} & -A_N + B_N - C_N \\ -A_N + B_N - C_N & -K_{N2} & -K_{N1} & 0 & \cdots & -A_2 + B_2 - C_2 & -A_1 + B_1 - C_1 & 0 \\ 0 & -A_N + B_N - C_N & 0 & -K_{N1} & -A_2 + B_2 - C_2 & \cdots & -K_{12} & -A_1 + B_1 - C_1 \\ 0 & 0 & -A_N + B_N - C_N & -K_{N2} & -A_1 + B_1 - C_1 & -K_{12} & \cdots & -A_2 + B_2 - C_2 \\ 0 & 0 & 0 & -A_N + B_N - C_N & 0 & -A_1 + B_1 - C_1 & -A_2 + B_2 - C_2 & \cdots \end{bmatrix} \tag{23}$$

The next cross-relaxation pathway we will take note of, is, unlike the previous pathways, in that its characteristic spectral density function is $J(2\omega_H)$. Double-quantum proton–proton dipolar interactions cause two protons in an LCLO substate to change at once from either an $a$ state to a $b$ state, or vice versa. For example, $N_zH_aH'_bH''_a\ldots$ may transition to $N_zH_bH'_bH''_b\ldots$ due to a double-quantum proton–proton dipolar interaction. For LCLO network protons $i$ and $j$, double-quantum cross-relaxation, which we will label $D_{ij}$, will take place between states $N_z\ldots H_{ia}\ldots H_{ja}\ldots$ and $N_z\ldots H_{ib}\ldots H_{jb}\ldots$ at the rate:

$$D_{ij} = \frac{3}{2}d^2_{H_iH_j}J_{H_iH_j}(2\omega_H) \tag{24}$$

These double-quantum interactions modify the Liouville relaxation matrix in the following manner:

The proton–proton dipolar/proton CSA terms will have this proton involved as one member of the dipolar pair, and also as the CSA contributor. Next, consider the other member of the dipolar pair. If the other member of the dipolar pair is in the $a$ state during the transition, the sign will be negative; if the other member of the dipolar pair is in the $b$ state during the transition, the sign will be positive. The contribution from the dipolar pair with state-switching proton $i$ and dipolar partner proton $j$ will therefore be:

$$XC^{DD/CSA}_{HH/H} = \frac{1}{2}c_{H_i}d_{H_iH_j}J^X_{H_iH_j/H_i}(\omega_H) \tag{26}$$
(dipolar partner in $a$ state)

$$XC^{DD/CSA}_{HH/H} = -\frac{1}{2}c_{H_i}d_{H_iH_j}J^X_{H_iH_j/H_i}(\omega_H) \tag{27}$$
(dipolar partner in $b$ state)

$$\bar{R} = \begin{bmatrix} \cdots & -A_2+B_2-C_2 & -A_1+B_1-C_1 & -D_{12} & -A_N+B_N-C_N & -D_{N2} & -D_{N1} & 0 \\ -A_2+B_2-C_2 & \cdots & -K_{12} & -A_1+B_1-C_1 & -K_{N2} & -A_N+B_N-C_N & 0 & -D_{N1} \\ -A_1+B_1-C_1 & -K_{12} & \cdots & -A_2+B_2-C_2 & -K_{N1} & 0 & -A_N+B_N-C_N & -D_{N2} \\ -D_{12} & -A_1+B_1-C_1 & -A_2+B_2-C_2 & \cdots & 0 & -K_{N1} & -K_{N2} & -A_N+B_N-C_N \\ -A_N+B_N-C_N & -K_{N2} & -K_{N1} & 0 & \cdots & -A_2+B_2-C_2 & -A_1+B_1-C_1 & -D_{12} \\ -D_{N2} & -A_N+B_N-C_N & 0 & -K_{N1} & -A_2+B_2-C_2 & \cdots & -K_{12} & -A_1+B_1-C_1 \\ -D_{N1} & 0 & -A_N+B_N-C_N & -K_{N2} & -A_1+B_1-C_1 & -K_{12} & \cdots & -A_2+B_2-C_2 \\ 0 & -D_{N1} & -D_{N2} & -A_N+B_N-C_N & -D_{12} & -A_1+B_1-C_1 & -A_2+B_2-C_2 & \cdots \end{bmatrix} \tag{25}$$

## Cross-correlated cross-relaxation

Cross-correlated proton–proton cross-relaxation mechanisms also exist in the LCLO system. These mechanisms take two forms. The first consists of cross-correlation between proton–proton dipolar relaxation and proton CSA relaxation, while the second term consists of cross-correlation between one proton–proton dipolar interaction and another proton–proton dipolar interaction.

Other dipolar/CSA terms consist of cross-correlation between heteroatom–proton dipolar relaxation and proton CSA relaxation, as well as cross-correlation between one heteroatom-proton dipolar interaction and another nitrogen–proton dipolar interaction. Spectroscopic means (180 pulses during relaxation) are used to average, and consequently eliminate, the effects of these heteroatom cross-correlated cross-relaxation mechanisms on relaxation during the $\eta_z/\kappa$ experiment.

Although the mechanisms vary, both active types of cross-correlated cross-relaxation operate along the same pathway as single-quantum cross relaxation. They cause LCLO substates to switch a single $a$ spin for a single $b$ spin, or vice versa. The sign of each cross-correlated cross-relaxation term changes depending on the spin states of the other proton states coupled by that term.

The rules for these sign changes are different for the dipolar/CSA and dipolar/dipolar terms. First, identify the proton whose state is being switched. In the case of a switch from $N_zH_aH'_bH''_a\ldots$ to $N_zH_aH'_aH''_a\ldots$, this is the $H'$ proton.

The proton–proton dipolar/proton–proton dipolar cross-correlated cross-relaxation terms follow a different rule. The proton whose state is being switched, will be involved in both dipolar interactions. Again, we label this proton $i$. Two other protons, $j$ and $k$, are the dipolar partners of proton $i$, one in each of the dipolar interactions. If the states of proton $j$ and $k$ are the same, the sign of the dipolar/dipolar cross-correlated cross-relaxation term will be negative; if the states of proton $j$ and $k$ are different, the sign of the term will be positive. The contribution from the dipolar/dipolar terms with state-switching proton $i$ and dipolar partner protons $j$ and $k$ will therefore be:

$$XC^{DD/DD}_{H_iH_j/H_iH_k} = \frac{3}{4}d_{H_iH_j}d_{H_iH_k}J^X_{H_iH_j/H_iH_k}(\omega_H) \tag{28}$$
($j$ and $k$ same state ($a/a$, $b/b$))

$$XC^{DD/DD}_{H_iH_j/H_iH_k} = -\frac{3}{4}d_{H_iH_j}d_{H_iH_k}J^X_{H_iH_j/H_iH_k}(\omega_H) \tag{29}$$
($j$ and $k$ different state ($a/b$, $b/a$))

## Ic. Auto-relaxation and cross-correlated contributions to auto-relaxation

The auto-relaxation behavior of the LCLO substates consists of a common longitudinal relaxation rate shared between all LCLO substates, with LCLO substate-specific modifications based on cross-correlated and auto-relaxation processes. We will first describe the common longitudinal relaxation rate and then move through the various cross-

**Table 2** Sample longitudinal relaxation rates for the network of Figs. 2 and 3

| Term | $J(\omega)$ term | Rate (s$^{-1}$) | $J(\omega)$ term | $J(w)$ | Rate (s$^{-1}$) |
|---|---|---|---|---|---|
| $K_{A1}$ (ZQ A-1) | $J(0)$ | 0.779 | $^1$H–$^{15}$N NOE | $J(N+H)$ | 0.007 |
| | | | | $J(N-H)$ | |
| $K_{12}$ (ZQ 1–2) | $J(0)$ | 0.097 | SQ A (1α,2α)$^a$ | $J(H)$ | 0.0034 |
| $^{15}$N CSA | $J(N)$ | 0.3568 | SQ A (1 α,2β)$^a$ | $J(H)$ | 0.0082 |
| $^{15}$N–($^1$H) Dip | $J(N) J(N+H)$ | 2.07 | SQ 1 (Aα,2α)$^a$ | $J(H)$ | 0.0031 |
| | $J(N-H)$ | | | | |
| $\eta_z$ | $J(N)$ | 1.71 | SQ 1 (A α,2β)$^a$ | $J$H, $J$H2 | 0.0023 |
| | | | DQ A-1 | $J(2H)$ | 0.0024 |
| | | | DQ 1–2 | $J(2H)$ | 0.0003 |
| | | | $^1$H-amide CSA | $J(H)$ | 0.002 |
| | | | $\zeta$ | $J(H)$ | 0.003 |
| | | | | $J(N+H)$ | |
| | | | | $J(N-H)$ | |

Longitudinal relaxation rates for a $^{15}$N–$^1$H amide spin system with two extraneous protons, as shown in Fig. 2. $B_0 = 11.47$ T (500 MHz $^1$H), and $\tau_c = 7$ ns, $S^2 = 1$. $A = {}^1$HN, $1 = {}^1$H$^1$, $2 = {}^1$H$^2$ as in the relaxation matrices. Because the network is symmetric, some values for $^1$H$^2$ are not listed. $^{15}$N CSA was 150 ppm, $^1$H-amide CSA was 10 ppm

$^a$ Symbols in the parenthesis show the spin states of the other two protons (effect of dipole–dipole/dipole–dipole cross correlation). SQ A (1β,2β) = SQ A (1α,2α), SQ A (1α,2β) = SQ A (1β,2α), etc.

correlation effects on auto-relaxation within the LCLO system.

The auto-relaxation rate common to all LCLO substates is:

$$R = \frac{1}{3}c_N^2 J_{\text{NCSA}}(\omega_N) + \sum_i \text{DDR}_i$$
$$= \frac{1}{3}c_N^2 J_{\text{NCSA}}(\omega_N) + \sum_i \frac{1}{4}d_{NH_i}^2 (J_{NH_i}(\omega_N - \omega_H)$$
$$+ 3J_{NH_i}(\omega_N) + 6J_{NH_i}(\omega_N - \omega_H)) \quad (30)$$

where $N$ is the amide nitrogen involved in the $NH$ spin system being probed.

Several cross-correlated auto-relaxation terms exist in the LCLO system. These terms change sign depending on the $a/b$ substate of the protons within each substate, causing each LCLO substate to display varying auto- as well as cross-relaxation behavior. Their origin lies in cross-correlation between nitrogen-proton dipolar and nitrogen CSA relaxation, or between the nitrogen-proton dipolar relaxation mechanisms of two different protons.

The $i$th proton in an LCLO substate will contribute a nitrogen-proton dipolar/nitrogen CSA cross-correlated auto-relaxation rate of the form:

$$\eta_{z_i} = c_N d_{NH_i} J_{NH_i/N}^X(\omega_N) \quad \text{(proton state } a) \quad (31)$$

$$\eta_{z_i} = -c_N d_{NH_i} J_{NH_i/N}^X(\omega_N) \quad \text{(proton state } b) \quad (32)$$

where the sign of the contribution depends on the $a$ or $b$ state of the $i$th proton.

The spectroscopic combination of the LCLO major states ensures that only the amide proton $H_N$'s DD/CSA cross-correlated auto-relaxation contribution will be observed in the final analysis of the experiment.

Meanwhile, the $i$th and $j$th proton in an LCLO substate will contribute a nitrogen–proton dipolar/nitrogen–proton dipolar cross-correlated auto-relaxation rate of the form:

$$\text{DDDD}_{ij} = \frac{3}{2}d_{NH_i}d_{NH_j} J_{NH_i/NH_j}^X(\omega_N)$$
$$(i,j \text{ states same } (a,a \text{ or } b,b)) \quad (33)$$

$$\text{DDDD}_{ij} = -\frac{3}{2}d_{NH_i}d_{NH_j} J_{NH_i/NH_j}^X(\omega_N)$$
$$(i,j \text{ states different } (a,b \text{ or } b,a)) \quad (34)$$

Numerical values for all different longitudinal relaxation rates for the spin system shown in Figs. 2 and 3, were computed for a $^{15}$N–$^1$H amide spin system with $B_0 = 11.47$ T (500 MHz $^1$H), and $\tau_c = 7$ ns, $S^2 = 1$. These values are presented in Table 2. It is clear that, individually, the rates governed by spectral densities at frequencies higher than $\omega_N$ are all very small compared to the $J(0)$-driven spin flips and $J(\omega_N)$-driven relaxation of the diagonal terms. For larger proteins, $\kappa$ rates will dominate the spectral density terms. At higher magnetic field, all $J(\omega)$ terms other than $J(0)$ will be even smaller, making the $\kappa$ domination even larger. In the "Results" section, we will show that the cumulative effect of the rates driven by $J(\omega)$ terms other than $J(0)$ can be neglected in practice, even when working with calmodulin on a 500 MHz spectrometer. Those interested in further investigations into the theory of relaxation in $AX_n$ spin

systems should consult the classic review (Werbelow and Grant 1977).

## II. Approximating the LCLO system

The detailed LCLO system contains ten different relaxation mechanisms operating over literally thousands of proton–proton and heteroatom–proton interactions, each with its own set of spectral density functions. These interactions cannot be practically separated and individually characterized with the current state of the art of NMR spectroscopy.

$$\text{Block } 0 : \frac{d}{dt} N_z H_\alpha^A H_\alpha^1 H_\alpha^2 = -(R + \eta_z) N_z H_\alpha^A H_\alpha^1 H_\alpha^2$$

$$\text{Block } 1 : \frac{d}{dt} \begin{bmatrix} N_z H_\alpha^A H_\alpha^1 H_\beta^2 \\ N_z H_\alpha^A H_\beta^1 H_\alpha^2 \\ N_z H_\beta^A H_\alpha^1 H_\alpha^2 \end{bmatrix} = - \begin{bmatrix} R + \eta_z + K_{12} + K_{A1} & -K_{12} & -K_{A1} \\ -K_{12} & R + \eta_z + K_{12} + K_{A2} & -K_{A2} \\ -K_{A1} & -K_{A2} & R - \eta_z + K_{A1} + K_{A2} \end{bmatrix} \begin{bmatrix} N_z H_\alpha^A H_\alpha^1 H_\beta^2 \\ N_z H_\alpha^A H_\beta^1 H_\alpha^2 \\ N_z H_\beta^A H_\alpha^1 H_\alpha^2 \end{bmatrix}$$

$$\text{Block } 2 : \frac{d}{dt} \begin{bmatrix} N_z H_\alpha^A H_\beta^1 H_\beta^2 \\ N_z H_\beta^A H_\alpha^1 H_\beta^2 \\ N_z H_\beta^A H_\beta^1 H_\alpha^2 \end{bmatrix} = - \begin{bmatrix} R + \eta_z + K_{A1} + K_{A2} & -K_{A1} & -K_{A2} \\ -K_{A1} & R - \eta_z + K_{12} + K_{A1} & -K_{12} \\ -K_{A2} & -K_{12} & R - \eta_z + K_{12} + K_{A2} \end{bmatrix} \begin{bmatrix} N_z H_\alpha^A H_\beta^1 H_\beta^2 \\ N_z H_\beta^A H_\alpha^1 H_\beta^2 \\ N_z H_\beta^A H_\beta^1 H_\alpha^2 \end{bmatrix}$$

$$\text{Block } 3 : \frac{d}{dt} N_z H_\beta^A H_\beta^1 H_\beta^2 = -(R - \eta_z) N_z H_\beta^A H_\beta^1 H_\beta^2$$

$$(35)$$

We can, however, approximate the detailed LCLO system by a simpler system that retains the important features of the LCLO system while being accessible to experiment. We will develop this approximation in two stages.

The first stage of the approximation is to neglect the operation of non-zero-quantum terms in the cross-relaxation elements of the matrix. Relevant auto-relaxation terms (the common LCLO auto-relaxation rate, and the dipole–dipole/CSA cross-correlated relaxation rate) are retained along the diagonal. This is equivalent to returning to (16) in the detailed description of the relaxation matrix given previously.

The qualitative basis for this approximation is that, in the large-molecule limit for which the experiment is designed, zero-quantum processes dominate the relaxation behavior of the LCLO system. Under conditions of relative molecular rigidity, the total non-zero-quantum cross-relaxation contribution from a given proton–proton interaction is 1–2% of the total zero-quantum relaxation contribution. This is due to the relative sizes of the spectral density function evaluated at zero frequency and at the proton Larmor frequency $\omega_H$; under semi-rigid molecular conditions, $J(0)$ is roughly $100\times$ larger than $J(\omega_H)$ at

500 MHz. Non-zero-quantum processes are consequently lost in the experimental error of measuring the zero-quantum processes.

In the "Results" section, we will return to this approximation and make a quantitative appraisal of its validity based on actual protein structures. For now, we will assume its validity and discuss its consequences.

As mentioned earlier, retaining only the zero-quantum terms in the LCLO relaxation matrix allows the matrix to be split into submatrices describing zero-quantum relaxation within a single energetic block. For the relaxation matrix these submatrices are, as illustrated previously:

The second step of the approximation is the replacement of the zero-quantum dipolar relaxation rates $K_{Aj}$ ($j = 1, \ldots, N$) between the amide proton and the $N$ neighbor protons involved in the LCLO system by the mean rate across the $N$ neighbor protons, $K = \sum_j K_{Aj}/N$. This is equivalent to replacing the real proton relaxation network with an approximate proton relaxation network in which all protons are equidistant from the amide proton. This equidistance is set such that the sum of zero-quantum relaxation rates between the amide proton and the virtual equidistant protons in the approximation is equal to the sum of zero-quantum relaxation rates between the amide protons and all true neighbor protons in the molecule.

We justify this approximation by making a brief detour into the conceptual similarities between the approximate LCLO system and the study of chemical dynamics.

In chemical dynamics, chemical reaction processes are examined for the effects of the energetic substate distributions of reactants and products. For instance, one vibrational energy state of a molecule may be much more likely to undergo a given reaction process than other vibrational energy states of the same molecule.

Consider a chemical system with an initial state $i$ and several possible final states $f$. The initial state converts to

each final state at a rate $k_{if}$. The chemical dynamics expression $K = \sum_j K_{Aj}/N$ for the total rate of decrease of the initial state is then simply:

$$k_i = \sum_f k_{if} \quad (36)$$

Now, consider an initial LCLO substate $i$ which has several possible final states $f$. For instance, $N_zH_aH'_bH''_a\ldots$ has among its possible final states $N_zH_bH'_aH''_a\ldots$ and $N_zH_aH'_aH''_b\ldots$. The initial state converts to each final state at $a$ rate $K_{if}$. The expression for the total rate of decrease of the initial LCLO substate is then:

$$K_i = \sum_f K_{if} \quad (37)$$

same LCLO major state, are removed from the system (an example would be the pathway connecting $N_zH_aH'_bH''_a$ and $N_zH_aH'_aH''_b$, labeled $K_{12}$ in the relaxation matrix equations). Because we can only observe changes in LCLO major state (i.e. changes in the amide proton spin state), we may neglect substate exchanges that do not cause a change of major state.

These simplifications drastically lower the number of parameters required to describe the relaxation behavior of the system. The system can now be described in terms of the sum of zero-quantum relaxation rates $K$ and the number of virtual approximation protons $N$. The zero-quantum relaxation submatrix blocks of the simplified system become:

$$\text{Block } 0 : \frac{\mathrm{d}}{\mathrm{d}t} N_z H_\alpha^A H_\alpha^1 H_\alpha^2 = -(R + \eta_z) N_z H_\alpha^A H_\alpha^1 H_\alpha^2$$

$$\text{Block } 1 : \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} N_z H_\alpha^A H_\alpha^1 H_\beta^2 \\ N_z H_\alpha^A H_\beta^1 H_\alpha^2 \\ N_z H_\beta^A H_\alpha^1 H_\alpha^2 \end{bmatrix} = -\begin{bmatrix} R + \eta_z + K & \cdots & -K \\ \cdots & R + \eta_z + K & -K \\ -K & -K & R - \eta_z + 2K \end{bmatrix}\begin{bmatrix} N_z H_\alpha^A H_\alpha^1 H_\beta^2 \\ N_z H_\alpha^A H_\beta^1 H_\alpha^2 \\ N_z H_\beta^A H_\alpha^1 H_\alpha^2 \end{bmatrix}$$

$$\text{Block } 2 : \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} N_z H_\alpha^A H_\beta^1 H_\beta^2 \\ N_z H_\beta^A H_\alpha^1 H_\beta^2 \\ N_z H_\beta^A H_\beta^1 H_\alpha^2 \end{bmatrix} = -\begin{bmatrix} R + \eta_z + 2K & -K & -K \\ -K & R - \eta_z + K & \cdots \\ -K & \cdots & R - \eta_z + K \end{bmatrix}\begin{bmatrix} N_z H_\alpha^A H_\beta^1 H_\beta^2 \\ N_z H_\beta^A H_\alpha^1 H_\beta^2 \\ N_z H_\beta^A H_\beta^1 H_\alpha^2 \end{bmatrix}$$

$$\text{Block } 3 : \frac{\mathrm{d}}{\mathrm{d}t} N_z H_\beta^A H_\beta^1 H_\beta^2 = -(R - \eta_z) N_z H_\beta^A H_\beta^1 H_\beta^2$$

$$(38)$$

Although we cannot observe the individual $K_{if}$ rates (the rate at which individual neighbor protons generate spin flips), we can observe their sum, the $K_i$ rate (the rate at which all neighbor protons together generate spin flips). This suggests that we can make an approximation to the real system (which has varying, unknowable $K_{if}$ rates) by constructing a virtual system with the same total $K_i$ rate, distributed equally across $N$ equidistant virtual neighbor protons. Each of these virtual neighbor protons has the same $K_{if}$ rate. The idea is that splitting the relaxation contributions equally among equidistant virtual neighbor protons allows the virtual system to approximate the equilibration behavior of a real system and preserve the same cross-relaxation rate $K_i$ while sidestepping the observationally intractable problem of dissecting $K_i$ into its component $K_{if}$s. We test this approximation in the "Results" section. For now, again, we will take it as given and move along with exploration of its consequences.

Under the equidistant virtual neighbor proton approximation, the interproton relaxation rates that involve zero-quantum spin flips between LCLO substates sharing the

This representation gives only two kinds of rows in each submatrix, one corresponding to substates belonging to the $N_zH_a$ major state and the other corresponding to substates belonging to the $N_zH_b$ major state. For the $m_J$th energetic block, each $N_zH_a$-type substate passes magnetization to $m_J$ $N_zH_b$-type substates, with all magnetization leaving at an individual pathway rate $K$. Each $N_zH_b$-type substate passes magnetization to $N + 1 - m_J$ $N_zH_a$-type substates, with all magnetization entering and leaving at $a$ rate $K$. Effectively, the behavior of the individual substates has been homogenized within the major states. This representation is closely related to dividing the system in "flippable" and "unflippable" states as described in our earlier work (Weaver and Zuiderweg 2008).

One may now remove the individual substates from the energetic block picture and replace them with two single states which represent substates belonging to $N_zH_a$ and $N_zH_b$ major states. The $N_zH_a$-type substate passes magnetization to the $N_zH_b$-type substate at $a$ rate $(N + 1 - m_J)K$, while the $N_zH_b$-type substate passes magnetization to the $N_zH_a$-type substate at $a$ rate $m_J K$. The dynamics of the $N_zH_a$- and $N_zH_b$-type substate in this representation are

identical to the dynamics of an individual $N_zH_a$- and $N_zH_b$-affiliated substate in the block picture above. This results in the replacement of every block, "no matter its size", by a $2 \times 2$ representative matrix:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} N_zH_\alpha \\ N_zH_\beta \end{bmatrix}_{N,m_J} = \\ -\begin{bmatrix} R+\eta_z+m_JK & -m_JK \\ -(N+1-m_J)K & R-\eta_z+(N+1-m_J)K \end{bmatrix}\begin{bmatrix} N_zH_\alpha \\ N_zH_\beta \end{bmatrix}_{N,m_J}$$

(39)

The intra-$N_zH_a$ and -$N_zH_b$ relaxation pathways—those which transfer magnetization from $N_zH_a$-affiliated substates to $N_zH_a$-affiliated substates, or from $N_zH_b$-affiliated substates to $N_zH_b$-affiliated substates—have disappeared from this representation. These relaxation pathways do not produce 'visible' changes in magnetization state in the $2 \times 2$ representative matrix, and can be therefore be ignored. The only productive pathways in the representative energetic block matrix are those leading from $N_zH_a$-affiliated substates to $N_zH_b$-affiliated substates, and vice versa.

For a given set of initial conditions, the solution of the system of equations represented by this $2 \times 2$ matrix yields the time evolution of a representative $N_zH_a$- and $N_zH_b$-type substate within the $m_J$th energetic block. This can be used to calculate the time evolution of the $N_zH_a$ and $N_zH_b$ major states within the $m_J$th energetic block after properly weighting the representative substates to account for the number of substates within each energetic block.

As mentioned previously, block $m_J$ contains $\binom{N+1}{m_J}$ substates, $\binom{N}{m_J}$ of which will be $a$ substates and $\binom{N}{m_J-1}$ of which will be $b$ substates. There are zero $a$ substates in the case of $m_J = N+1$, and zero $b$ substates in the case of $m_J = 0$ (in these cases, the binomial formulation is nonsensical). Multiplying the $N_zH_a$ and $N_zH_b$ solutions of the system of differential equations given in (39) by, respectively, $\binom{N}{m_J}$ and $\binom{N}{m_J-1}$ allows proper accounting for the LCLO substate multiplicity within each energetic block.

The final step is to divide the results of the system of differential equations by $2^{N/2}$ in order to normalize the total population of all substates. The time evolution of the complete approximate LCLO system can then be calculated by summing (39) over all $N+2$ energetic blocks. This leaves us with the final system of differential equations for the LCLO approximation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} N_zH_\alpha \\ N_zH_\beta \end{bmatrix}_N = \sum_{m_J=0}^{N+1}\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} N_zH_\alpha \\ N_zH_\beta \end{bmatrix}_{N,m_J}$$

$$= \sum_{m_J=0}^{N+1} -\begin{bmatrix} 2^{-N/2}\binom{N}{m_J} & 0 \\ 0 & 2^{-N/2}\binom{N}{m_J-1} \end{bmatrix}$$

$$\begin{bmatrix} R+\eta_z+m_JK & -m_JK \\ -(N+1-m_J)K & R-\eta_z+(N+1-m_J)K \end{bmatrix}$$

$$\begin{bmatrix} N_zH_\alpha \\ N_zH_\beta \end{bmatrix}_{N,m_J}$$

(40)

## Results

### Data fitting procedure

Synthetic or experimental data was fitted in several steps. The symmetric reconversion time evolution function $X(t)$ (see "Methods") for each residue is fit by Levenberg–Marquardt nonlinear least squares while varying the number of equidistant protons $N$ from 1 to 60 (Fig. 4). The number of neighbor protons needed for the fit is a function of the protein correlation time and isotopic labeling scheme. For $^{15}$N-labeled, protonated calmodulin, optimal $N$ values varied between 5 and 11. The fit also includes parameters governing the purity of initial magnetization state preparation (either $N_zH_a$ or $N_zH_b$, typically > 95%). The fitted $K$-rates and state purity parameters are then used to fit the symmetric reconversion time evolution function $Y(t)$ (see "Methods") for the value for $\eta_z$ for each residue and the efficiency of observing $N_zH_a$ or $N_zH_b$ magnetization (typically 10% more efficient for $N_zH_b$ than $N_zH_a$). Bootstrap statistical methods (Efron and Tibshirani 1994) are used to construct confidence intervals for the fitting parameters.

### Validation of the approximate LCLO system

In order to validate the equidistant proton network approximation, we generated synthetic relaxation data for the peptide-bound form of Ca$^{2+}$-bound calmodulin using the X-ray coordinate file 2O5G.pdb. The hydrogen atoms were placed with MolProbity (Lovell et al. 2003). For each amide proton, we took into account the eight nearest protons.
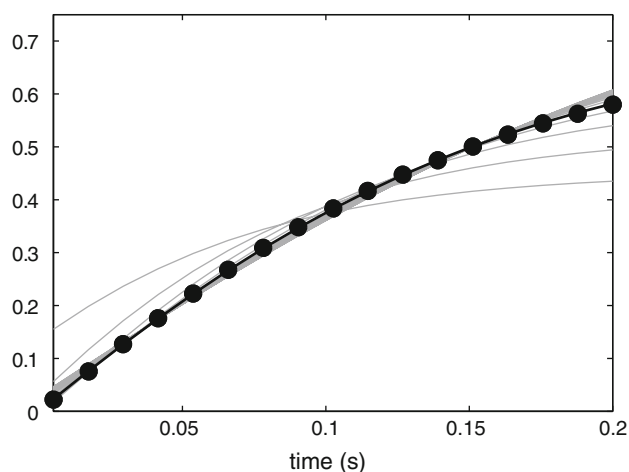
**Fig. 4** Fitting of the symmetric reconversion function $X(t)$ (see "Materials and methods") as a function of number of neighbor equidistant protons $N$. The symbols are synthetic data taking the eight closest protons in the PDB file for residue Ile 100 of bound calmodulin into account, using all possible relaxation mechanisms; axially symmetric rotational diffusion as in Table 1; $S^2 = 0.8$; $\tau_e = 1$ ps; 500 MHz. *Grey lines* are test fits using the equidistant proton approach, varying $N$ from 1 to 60. The lowest trace at time = 0.2 s is for a $N = 1$, the highest trace corresponds to $N = 60$. The *black line* is the best fit, which occurred for $N = 5$ equidistant protons

The system was set up for axially symmetric diffusion according to our experimentally determined diffusion parameters for smMLCKp/CaM and included all possible relaxation mechanisms including proton–proton dipolar/proton–proton dipolar cross correlations. We assumed an order parameter of 0.8 for all spectral density functions with a local correlation time of 1 ps. We then fit this synthetic data using the equidistant proton network approximation with only proton–proton spin-flip ($K$) rates and compare the $K$-rates resulting from this fit to the detailed $K$-rates which were used to generate the synthetic dataset. The comparison of the synthetic $K$-rates and back-fitted $K$-rates is shown in Fig. 5.

The correspondence is excellent ($p = 0.995$) over the range of $K$-rates between 3 and 13 s$^{-1}$. This is strong proof that the neglect of higher frequency spectral density terms and use of the equidistant proton approximation is justified.

Data fitting for calmodulin

An example of the quality of data fits using the equidistant proton approximation for smMLCKp-bound calmodulin is shown in Fig. 6. Figure 7 shows all fitted $K$ values for free and bound calmodulin. As expected, the $K$ values vary strongly over the protein due to diversity in local proton environments. Error bars are obtained from 200 Monte Carlo simulations, and are seen to be smaller than the variation in $K$-rates. Several residues in free and bound
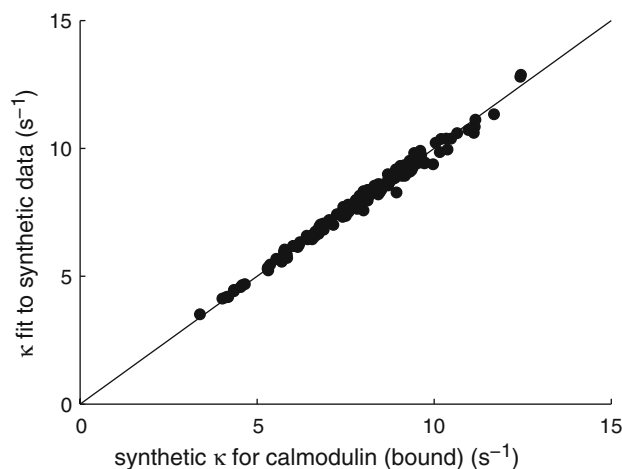


**Fig. 5** Synthetic relaxation data, $N_zH_a$ ($t = 0$) $\rightarrow N_zH_a(t)$, $N_zH_a$ ($t = 0$) $\rightarrow N_zH_b(t)$, $N_zH_b$ ($t = 0$) $\rightarrow N_zH_a(t)$ and $N_zH_b$ ($t = 0$) $\rightarrow N_zH_b$ ($t$) was generated from the PDB file for all residues of bound calmodulin, taking the eight closest neighbors for each amide into account and using all possible relaxation mechanisms with all spectral density terms; using axially symmetric rotational diffusion as in Table 1; $S^2 = 0.8$; $\tau_e = 1$ ps; 500 MHz. These synthetic relaxation data were fitted with the equidistant proton approach, using terms with spectral densities $J(0)$ only. The obtained fitted $K$-rates, averaged over 200 Monte Carlo trials, are shown on the *vertical axis*. On the *horizontal axis* are the input $K$-rates as derived from the crystal structure. The *solid line* is $y = x$. The correlation coefficient is 0.995

form demonstrate very fast $K$-rates. These outliers can likely be explained by amide proton exchange: exchange between amide protons and water protons will scramble the prepared $N_zH_a$ and $N_zH_b$ states, giving rise to large $K$ values.

Amide proton exchange has been extensively studied for $Ca^{2+}$-CaM free in solution and bound to M13 skeletal muscle MLCK (Spera et al. 1991), a peptide similar to the smooth muscle MLCK peptide used in this study (Crivici and Ikura 1995). Residues for which the free or M13-bound exchange rate was found to be larger than 1 s$^{-1}$ are rendered in grey in Fig. 7. The outliers in the raw $K$ data for the free/smMLCK system all correspond to amide protons with very fast exchange rates in the free/M13 system. For instance, the outlier for N42 at 30 s$^{-1}$, was reported (Spera et al. 1991) to have an amide proton exchange rate $k_{HX} = 120$ s$^{-1}$. In principle, one should be able to correct experimental $K$-rates for amide proton mass-exchange rates and water—amide NOEs values provided the experimental conditions of the latter experiments are identical. Currently we do not have such data available. For this reason, residues with reported amide exchange rates faster than 1 s$^{-1}$ in free or bound form are not included in our analysis of proton–proton dynamics.

Figure 8 illustrates $Q$, the ratio of the experimental $K$-rates to those computed from the structure. Due to X-ray
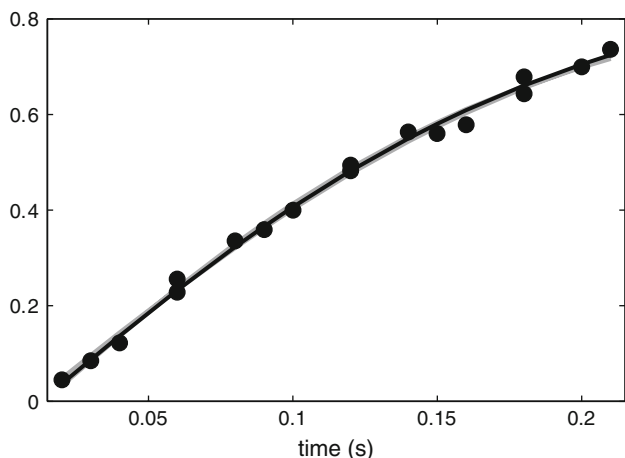
**Fig. 6** Examples of the quality of the equidistant proton fitting of experimental 500 MHz $X(t)$ cross relaxation data (see "Materials and methods") for Ile 100 of peptide-bound calmodulin. The *faint grey lines* give the fits $\pm 1$ SD in the ensemble of 200 Monte Carlo trials. $K$ value is $10.2 \pm 0.5$ s$^{-1}$. Note that the fitting process allows a small offset at the origin; this is to take into account imperfect preparation of the initial states

structural imprecision and its effects on theoretical $K$-rates (see "Methods" and "Discussion"), relative errors in $Q$ are considerably larger than relative errors in experimental

$K$-rates. Deviations of $Q$ from 1 are due to sub-nanosecond motion of the proton network in the vicinity of the amide protons. Local angular motions of the inter proton vectors will reduce $Q$. For these cases, $Q$ is similar in concept to a classic $S^2$ order parameter. However, in a non-covalent network, it is also necessary to consider dynamic changes in interproton distances. If we consider the coordinates used to compute the theoretical $K$-rates to represent the dynamic average, we must consider two possibilities. Translational fluctuations of a neighboring proton along the virtual vector connecting it with the amide proton will foreshorten the apparent interproton distance (by the $1/r^6$ mechanism), leading to higher $K$-rates, and hence $Q$ values larger than one. Translational fluctuations perpendicular to the virtual connection vector will make the interproton distance longer, reducing $Q$. In this case, $Q$ is further reduced by the apparent angular motion of the interproton vector.

Hence, it is possible to associate values of $Q$ different than one as indicative motion, but the amplitude of the motion does not necessarily change monotonously with increasing deviation of the value. Indeed, even $Q$ values of unity could hide motions that counteract each other in relaxation effects. For a detailed discussion of the $Q$

**Fig. 7** $K$-rates fitted to the experimental relaxation curves for calmodulin free (*top*) and calmodulin bound (*bottom*). The *error bars* were obtained from 200 Monte Carlo trials. In *grey* are the residues for which amide protons exchange faster than 1 s$^{-1}$ (see Spera et al. 1991). These residues are not included in the data analysis of the next figures. In this and subsequent figures, *thick horizontal bars denote* helical regions of CaM, *gray horizontal bars* denote Ca$^{2+}$-binding loops, and *thin horizontal lines* denote other loops of CaM
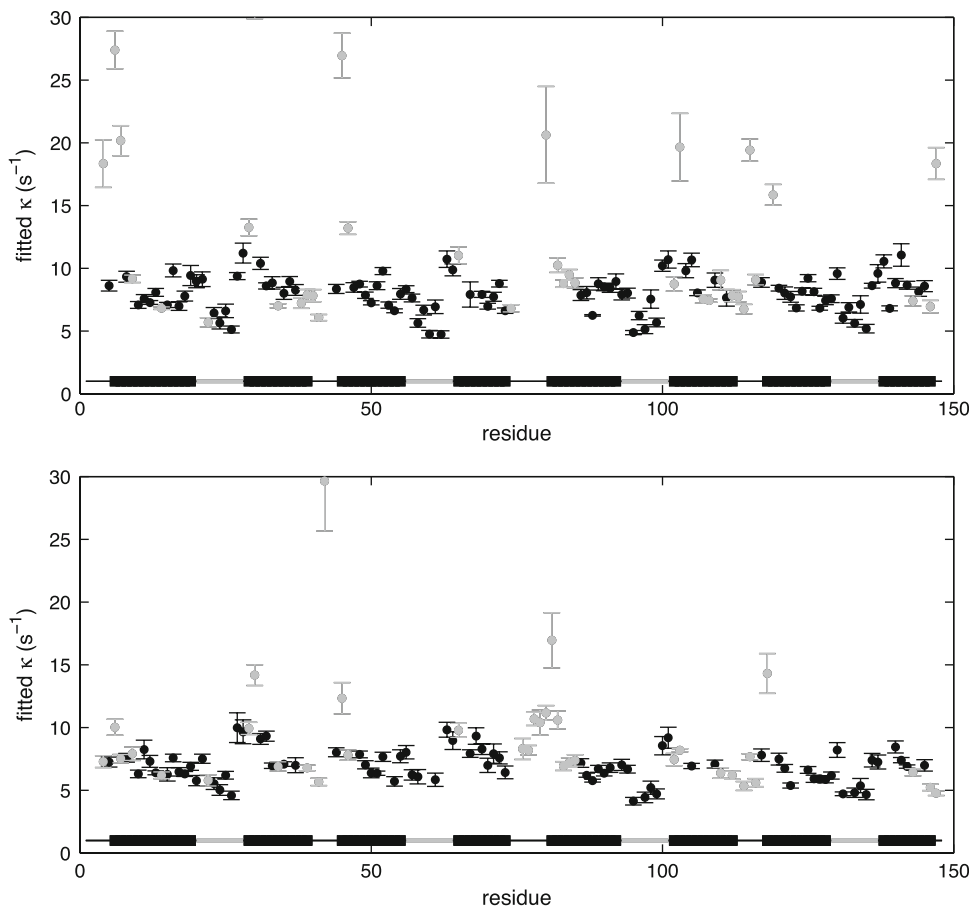
**Fig. 8** *Q*-values for
calmodulin. *Top panel* bound
calmodulin; *middle panel* free
calmodulin; *bottom panel*
bound—free. Average *Q* values
are: $<Q_{NT}> = 0.82 \pm 0.02$ and
$<Q_{CT}> = 0.80 \pm 0.02$ for free
calmodulin, and
$<Q_{NT}> = 0.88 \pm 0.02$ and
$<Q_{CT}> = 0.89 \pm 0.02$ for
smMLCKp—bound calmodulin
(N terminus = residues 1–76; C
terminus = residues 82–148)
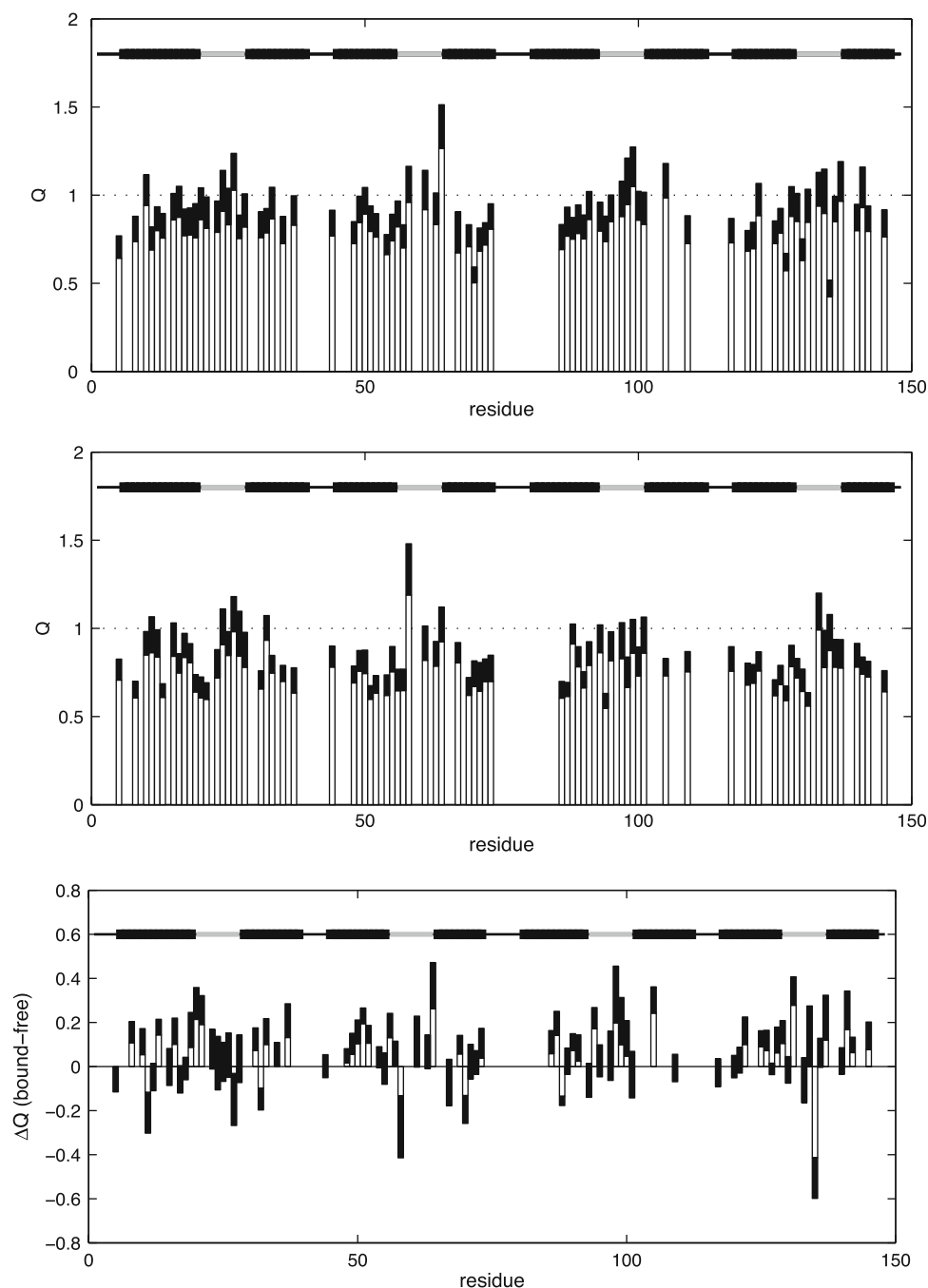


parameter, readers are referred to the work of Post et al. (1989) and Post (1992).

## Discussion

### *Q* and peptide binding in calmodulin

The great majority of *Q* values (Fig. 8) are smaller than 1 for both free and bound calmodulin. Average *Q* values are

$(<Q_{NTerm}> = 0.82 \pm 0.02$ and $<Q_{CTerm}> = 0.80 \pm 0.02)$ for free calmodulin and $(<Q_{NTerm}> = 0.88 \pm 0.02$ and $<Q_{CTerm}> = 0.89 \pm 0.02)$ for smMLCKp—bound cal-modulin. The N terminus comprises residues 1–76 and the C terminus comprises residues 82–148. A paired *t*-test of *Q* values in free and smMLCKp-bound calmodulin indicates that the means of these values differ with a *p* value of 0.001. The 95% confidence interval for the change in the mean *Q* is 0.03–0.11. For the N-terminal domain alone the paired *t*-test *p* value is 0.05 and the 95% confidence

interval is 0.00–0.12, while for the C-terminal domain alone the paired *t*-test *p* value is 0.01 and the 95% confidence interval is 0.02–0.14. These results clearly indicate global quenching of proton–proton dynamics in $Ca^{2+}$-saturated calmodulin upon smMLCKp binding. Our experimentally measured *Q* values agree well with values found in earlier theoretical and experimental work (Post 1992; Brüschweiler et al. 1992; Olejniczak et al. 1984).

The distribution of *Q* values for bound and free states is shown in Fig. 9. As discussed previously, mean and median *Q* values are both considerably <1. The shift in average *Q* between bound and free calmodulin constitutes a shift in the total distribution.

*Q* values are dependent on accurate measurements of the rotational diffusion tensor, since the theoretical *K*-rates are constructed from measurement of the diffusion tensor. In this work, we obtained axially symmetric rotational diffusion tensors and associated site-specific $\tau_c$ values from the ratio of the $^{15}N$ CSA/$^{15}N$–$^1HN$ dipolar cross correlation ratios $\eta_{xy}/\eta_z$ (see "Materials and methods"). We obtain values that are on average 10% smaller than the $^{15}N$ $R_2/R_1$ ratio-based $\tau_c$ values reported for this protein in the literature (Lee et al. 2000). Without evaluating the relative merits of these methods, we note that the use of the literature values for $\tau_c$ would cause <*Q*> to decrease by 10%.

The fact that <*Q*> is (significantly) less than one implies the presence of sub-nanosecond motion of the proton environment of the amide protons. Due to different and potentially counteracting effects of local motions on *Q* (Post 1992 and "Results" section), one may not (at this time) simply assign more motion to free calmodulin (<*Q*> = 0.81 ± 0.02) than to smMLCKp—bound calmodulin (<*Q*> = 0.88 ± 0.02).

Nevertheless, there are some trends in the statistics of *Q* factors that do suggest that we can, after all, associate lower *Q* values with more motion. Consider Fig. 10, which compiles the *Q*-data according to the percentage contribution of different types of protons driving the amide proton flip rates. Dominance of the $C_\alpha$-proton (mostly of the same residue) is associated with higher values of *Q*,

indicating little motion of the protein backbone itself, in accordance with $^{15}N$ relaxation measurements for this protein (Lee et al. 2000). Dominance of sidechain $\gamma$ protons (predominantly located on the residue's own sidechain) is associated with *Q* values tending to 0.5, indicating motions of very significant amplitude in these moieties. Dominance of $H_\beta$ does not show any particular trend, likely reflecting their intermediate position between the $H_\alpha$ and $H_\gamma$'s which have opposite trends. The statistics for the amide protons in the bound state shows similar trends as for the $C_\alpha$ protons. This would also indicate that the dynamics of the amide–amide proton vector are rather restricted.

Overall, trends in Fig. 10 are to be expected if *Q*, like an order parameter, does anticorrelate monotonically with motion. Potentially, one may consider *Q* values as a semiquantitative measure of the motions of the sidechain of the same residue. But before a quantitative relation can be established, more work, including comparison with computational molecular dynamics, is required.

If we take the observations above to indicate that <*Q*> is at the very least a "monotonic" function that anticorrelates with local sidechain motion, we may conclude that the protein relaxation network as seen by the amide proton is significantly more dynamic than that as felt by the amide nitrogen. This conforms to the common observation that all dynamical measures, such as $S^2CH_3$-detected dynamics (Lee et al. 2000; Tugarinov et al. 2007), backbone RDC-detected dynamics (Peti et al. 2002), backbone $^{13}CO$-detected dynamics (Pang et al. 2002), and cross correlated dynamics (Ferrage et al. 2006) such as backbone $^{13}C$–$^{15}N$ tensor reorientation detected dynamics (Pellecchia et al. 1999), indicate more motion than classical $^{15}N$-detected dynamics. It also supports results from molecular dynamics calculations, which show large amplitude sidechain motions even in the protein interior (Best et al. 2005).

If we assume that <*Q*> is a monotonic function that anticorrelates with local sidechain motion, we may associate the difference in *Q* values (Fig. 8) and/or the shift in the *Q* distribution between free and smMLCKp—bound calmodulin (Fig. 9), as a change in motion. Especially the



**Fig. 9** The distribution of *Q* values for free calmodulin (*open bars*) and bound calmodulin (*closed bars*). *Left* N terminus (residues 1–76) *Right* C terminus (residues 82–148)
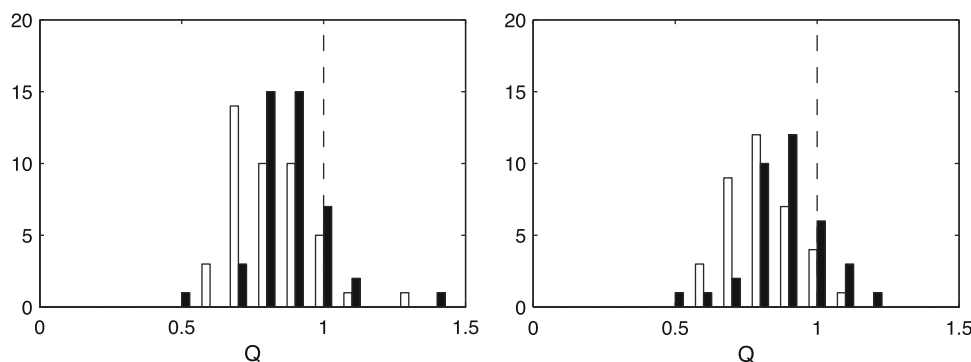
**Fig. 10** Structural statistics of
*Q*. The horizontal axis specifies
the percentage of the
contribution to theoretical *K*
values as expected from the
PDB coordinates of the protein
for the indicated type of
neighbor proton. The vertical
axis gives the *Q* value
associated with the proton. *Left*
calmodulin free; *right*
calmodulin bound. From *top* to
*bottom* HN, HA, HB (no
methyls), HG (no methyls)
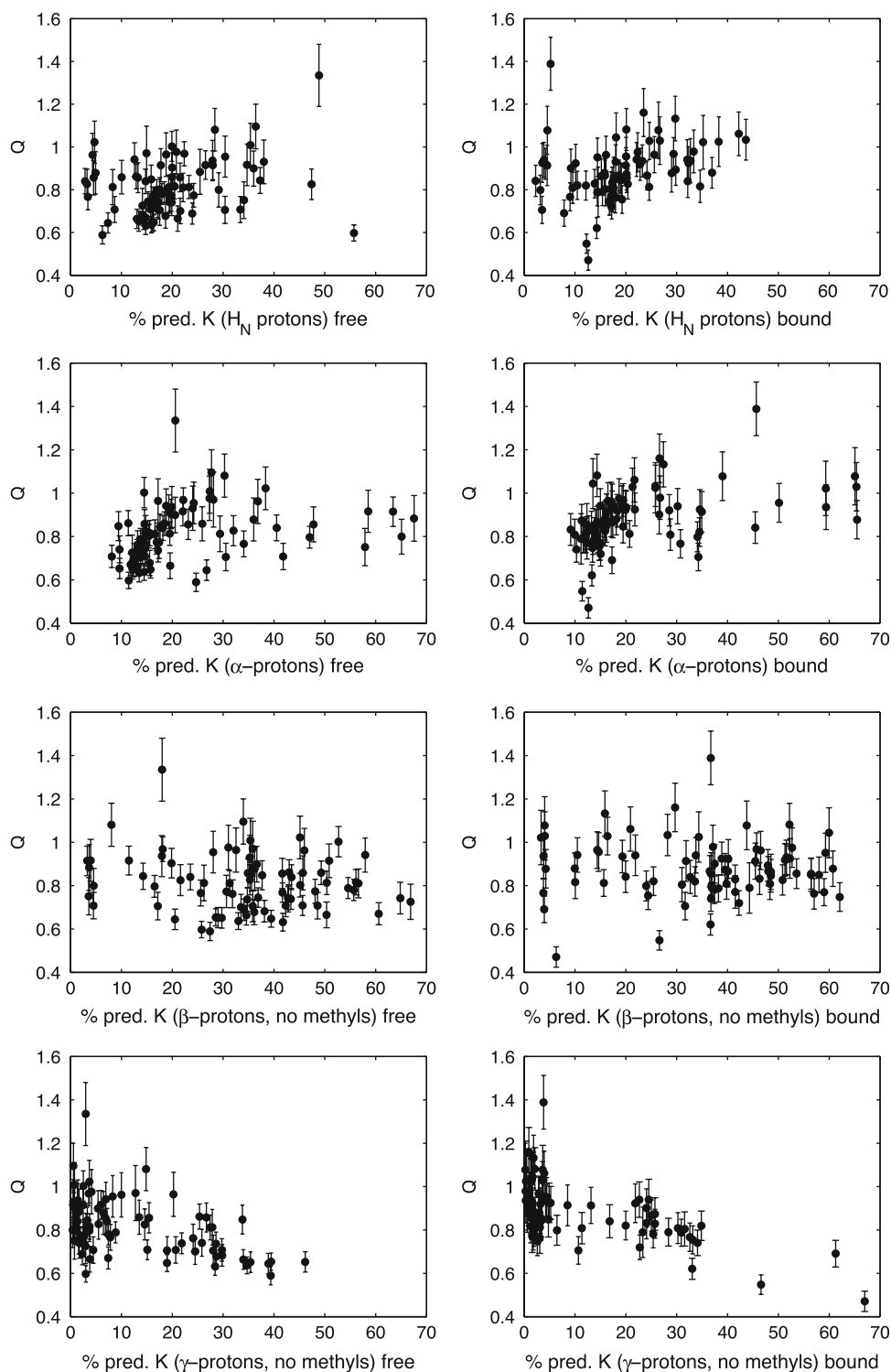


difference in $Q$ in Fig. 8 is very reminiscent of work presented by Lee et al. (2000), which shows a similar change in $S^2CH_3$ for the same protein upon binding of the same ligand. Surprisingly, and likely by chance, the increase in average $Q$, $<\Delta Q>$ of 0.07 upon ligand binding corresponds to the done $<\Delta S^2CH_3>$ of 0.07 as determined by Lee et al. (2000).

Effect of the limited precision of experimental protein structures

In this work, we have used the 1EXR.pdb and 2O5G.pdb X-ray crystal structures (1.0 and 1.08 Å resolution, respectively). As explained in the "Methods" section, this resolution formally translates into a 0.055 to 0.075 Å

point-to-point distance imprecision, contributing $\sim 5.5\%$ and error and $\sim 7.5\%$ error margins to theoretical $K$-rates. These errors only slightly larger than the errors of fit, and are much smaller than the individual differences in $K$-rates. Hence, the salient points of our discussion in terms of structure and dynamics are not affected by this.

However, more commonplace structures are of 1.5–2.0 Å resolution, and have a point-to-point distance precision of 0.3–0.4 Å. Distance errors of this size would render a calculation of $Q$ essentially meaningless under the assumptions made above, since the error in $Q$ measurements in combination with $r^{-6}$-dependent relaxation rates would cause large swings in theoretical $K$-rates.

However, this assertion may be overly harsh. Intuitively, the proton geometry immediately local to a given amide proton should be reasonably well defined by the restrictions placed upon it by just a few bond lengths and angles. Although estimates of a given atom's absolute position in space may be imperfect, the network of chemical bonds near an atom ensures that relative inter-atomic distances and orientations are preserved, and these may be known with high precision even when the overall precision of the structure is lower. This means that one may be able to make a more precise estimate of distances between nearby neighbors than would be expected from the error in the difference between two absolute position estimates.

Structure-free use of $K$-rates

In the absence of a high-resolution structure, possibilities remain for using $K$-rates in experimental structural biology. The key is the ability to compare two different molecular states.

Changes in $K$-rates, when corrected for changes in overall rotational correlation times, between states $A$ and $B$ indicate that the local proton network experiences a structural and/or dynamic perturbation during the transition from $A$ to $B$. In the case of calmodulin, changes in $K$ correlate (weakly) with changes in $Q$ (Fig. 11). Investigating this in more detail, we find that protons within 4 Å of the amide protons, reporting the ten largest changes are concentrated within the hydrophobic clefts of both domains and in the face of helix 1 packed against the bound peptide (Fig. 12). Hence, without knowledge of the structure, one would in this example be able to use the comparison of $K$-rates to help pinpoint the areas that change most in structure and dynamics.

Quantitatively, the mean N-terminal corrected $K$-rate increases by 5%, while the mean C-terminal corrected $K$-rate increases by 8%. One may infer increased proton density or quenched proton dynamics for these areas from these changes, both of which would raise $K$-rates. Likely, one phenomenon goes hand in glove with the other (Zhang
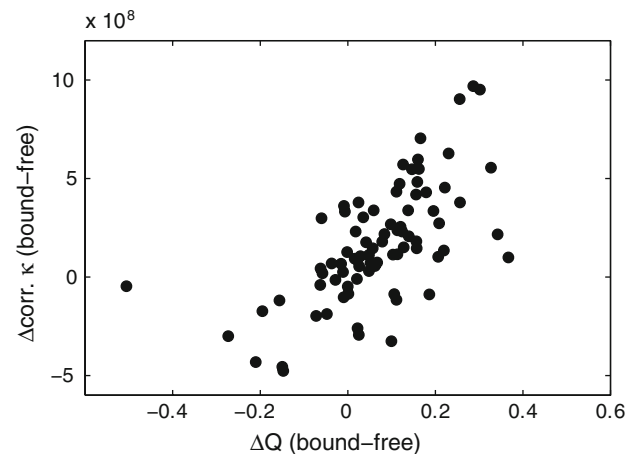


**Fig. 11** Correlation plot of change in corrected $K$ versus change in $Q$ upon smMLCKp binding to $Ca^{2+}$-saturated calmodulin. The value of the correlation coefficient is 0.68

and Bruschweiler 2002; Ming and Bruschweiler 2004). This may be compared with increases in mean $Q$ by 7 and 10% for the N and C-terminal areas, respectively. Hence, even without structure, one would correctly conclude quenching in dynamics for these areas. This example therefore suggests that changes in $K$ may provide leads on the location of binding sites, regions of structural plasticity, and other functionally interesting areas in the absence of detailed structural information.
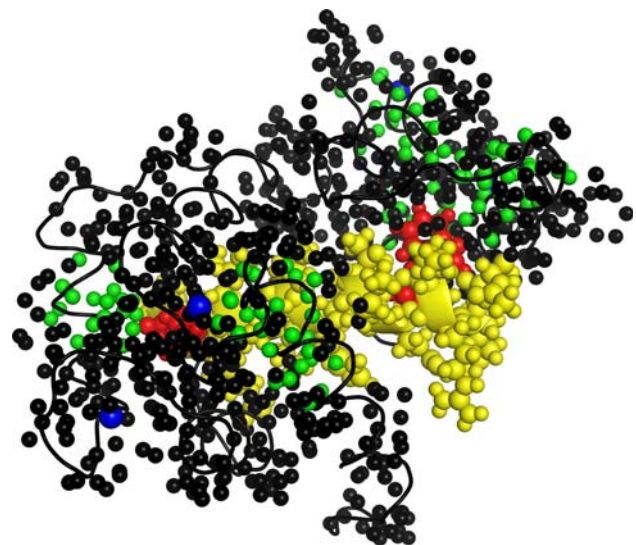


**Fig. 12** Illustration of ten largest corrected $K$ changes (by absolute value) upon smMLCKp binding to $Ca^{2+}$-saturated calmodulin. smMLCK peptide in *yellow*, evolutionarily conserved smMLCKp major hydrophobic anchors W5 and L18 in *red*. All smMLCK atoms are displayed. CaM protons are generally in *black* with CaM protons within 4 Å of amide protons reporting ten largest corrected $K$ changes in *green*. Protons belonging to residues excluded due to amide proton exchange are not included in the illustration. Sphere size is set to one-half the van der Waals radius for clarity of illustration

Relative merit of methods

The $Q$ ratios are not formally related to order parameters, since each residue's $K$-rate is affected by the radial and axial dynamics of several dipole–dipole vectors (see "Results"). This is a drawback in that it is not possible to associate a single precise motion with $Q$, but simultaneously, it is an advantage in that $Q$ is not as "myopic" as a true order parameter, which detects only the perpendicular angular motion of a single type of vector. As discussed above, even translational motions, classically opaque to NMR relaxation, will affect both distance and angle of the interproton vectors, and can be detected by the new method. As an advantage to ILV-directed methyl dynamics (Yang et al. 1998), $Q$ sidechain dynamics can be obtained for every residue in the sequence, as it just requires $^{15}$N labeling. Nevertheless, the method cannot replace more precise $S^2$CH$_3$ side chain order parameter determinations, and must be seen as a complement to those methods: less accurate, but possibly covering broader ground. The method holds advantages over $^{15}$N relaxation measurements, not only by the virtue that it is less "myopic", but also by the fact that the experiments can be carried out for larger proteins, since no insensitive $^1$H $\rightarrow$ $^{15}$N NOE experiment is required (Gong and Ishima 2007). In addition, the experiments do not challenge spectrometer hardware, as no CPMG, $T_{1\rho}$ or other decoupling sequences are required.

The method has as an inherent advantage that the key rate parameter, $K$, scales with $J(0)$ and hence increases linearly with molecular weight. Moreover, non-zero-frequency spectral density functions decrease in size as molecular weight and magnetic field strength increase, efficiently "cleaning up" the experiment. $Q$ is also independent of cross correlation, CSA and conformational exchange contributions, and is hence not affected by the increases in magnetic field needed for the studies of larger proteins.

Deuteration of part or all of the non-amide proton network provides an opportunity to continue $K$ experimentation in proteins larger than 50 kDa. Replacement of protons with deuterons removes protons out of the relaxation network, leading to slower $K$-rates, but the increase in molecular weight will compensate for this. The $K$-rates would report on inter-amide proton relaxation, which is rather limited, but is still better than having no backbone dynamics measurements at all. Deuteration of part or all of the non-amide proton network also provides an opportunity to sort the local dynamics of smaller proteins in backbone and sidechain contributions.

Lastly, the methods are equally applicable to imino protons in nucleic acids and to carbon-bound protons in both proteins and nucleic acids. Although TROSY effects

are diminished in such protons, one can still study the cross relaxation between the proton "up" and proton "down" states using the same $\eta_z/K$ NMR experiment, with $^{13}$C rather than $^{15}$N pulses, and using the same data interpretation and fitting methods.

## References

Akke M, Bruschweiler R, Palmer AG (1993) NMR order parameters and free-energy—an analytical approach and its application to cooperative Ca2 + binding by calbindin-D(9 k). J Am Chem Soc 115:9832–9833

Best RB, Clarke J, Karplus M (2005) What contributions to protein side-chain dynamics are probed by NMR experiments? A molecular dynamics simulation analysis. J Mol Biol 349:185–203

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

Cooper A, Dryden D (1984) Allostery without conformational change. A plausible model. Eur Biophys J 11:103–109

Crivici A, Ikura M (1995) Molecular and structural basis of target recognition by calmodulin. Annu Rev Biophys Biomol Struct 24:85–116

Cruickshank D (1999) Remarks about protein structure precision. Acta Cryst D55:583–601

Efron B, Tibshirani R (1994) An introduction to the bootstrap. Chapman & Hall/CRC, London

Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D (2005) Intrinsic dynamics of an enzyme underlies catalysis. Nature 438:117–121

Ferrage F, Pelupessy P, Cowburn D, Bodenhausen G (2006) Protein backbone dynamics through 13C'-13C alpha cross-relaxation in NMR spectroscopy. J Am Chem Soc 128:11072–11078

Fischer MWF, Zeng L, Pang Y, Hu W, Majumdar A, Zuiderweg ERP (1997) Experimental characterization of models for backbone pico-second dynamics in proteins. Quantification of NMR auto- and cross correlation relaxation mechanisms involving different nuclei of the peptide plane. J Am Chem Soc 119:12629–12642

Frederick KK, Marlow MS, Valentine KG, Wand AJ (2007) Conformational entropy in molecular recognition by proteins. Nature 448:325–329

Gong Q, Ishima R (2007) 15 N–1H NOE experiment at high magnetic field strengths. J Biomol NMR 37:147–157

Kroenke CD, Loria JP, Lee LK, Rance M, Palmer AGIII (1998) Longitudinal and transverse 1H–15 N dipolar/15 N chemical shift anisotropy relaxation interference: unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules. J Am Chem Soc 120:7905–7915

Labeikovsky W, Eisenmesser EZ, Bosco DA, Kern D (2007) Structure and dynamics of pin1 during catalysis by NMR. J Mol Biol 367:1370–1381

Lee LK, Rance M, Chazin WJ, Palmer AGIII (1997) Rotational diffusion anisotropy of proteins from simultaneous analysis of 15 N and 13C alpha nuclear spin relaxation. J Biomol NMR 9:287–298

Lee AL, Kinnear SA, Wand AJ (2000) Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. Nat Struct Biol 7:72–77

Lindemann F (1910) The calculation of molecular vibration frequencies. Physikalisches Zeitschrift 11:609–612

Loria JP, Rance M, Palmer AGIII (1999) A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. J Am Chem Soc 121:2331–2332

Lovell S, Davis I, Arendall W, De Bakker P, Word J, Prisant M, Richardson J, Richardson D (2003) Structure validation by C-alpha geometry: phi, psi, and C-beta deviation. Proteins Struct Funct Genet 50:437–450

Mäler L, Blankenship J, Rance M, Chazin WJ (2000) Site-site communication in the EF-hand Ca2 + -binding protein calbindin D9 k. Nat Struct Biol 7:245–250

Millet O, Muhandiram DR, Skrynnikov NR, Kay LE (2002) Deuterium spin probes of side-chain dynamics in proteins. 1. Measurement of five relaxation rates per deuteron in (13)C-labeled and fractionally (2)H-enriched proteins in solution. J Am Chem Soc 124:6439–6448

Ming D, Bruschweiler R (2004) Prediction of methyl-side chain dynamics in proteins. J Biomol NMR 29:363–368

Mittermaier A, Kay LE (2006) New tools provide new insights in NMR studies of protein dynamics. Science 312:224–228

Neuhaus D, Williamson M (2000) The nuclear Overhauser effect in stereochemical and conformational analysis. Wiley, New York

Olejniczak E, Dobson C, Karplus M, Levy R (1984) Motional averaging of proton nuclear Overhauser effects in proteins. Predictions from a molecular dynamics simulation of lysozyme. J Am Chem Soc 106:1923–1930

Palmer III AG (2009) Inertia and diffusion tensors. http://biochemistryhscolumbiaedu/labs/palmer/software/diffusionhtml

Pang A, Buck M, Zuiderweg ERP (2002) Backbone dynamics of the ribonuclease binase active site area using multinuclear ($^{15}$N and $^{13}$CO) NMR relaxation and computational molecular dynamics. Biochemistry 41:2655–2666

Pellecchia M, Pang Y, Wang L, Kurochkin AV, Kumar A, Zuiderweg ERP (1999) Quantitative measurement of cross-correlations between $^{15}$N and $^{13}$CO chemical shift anisotropy relaxation mechanisms by multiple quantum NMR. J Am Chem Soc 121:9165–9170

Pelupessy P, Espallargas GM, Bodenhausen G (2003) Symmetrical reconversion: measuring cross-correlation rates with enhanced accuracy. J Magn Reson 161:258–264

Peti W, Meiler J, Bruschweiler R, Griesinger C (2002) Model-free analysis of protein backbone motion from residual dipolar couplings. J Am Chem Soc 124:5822–5833

Post C (1992) Internal motional averaging and three-dimensional structure determination by nuclear magnetic resonance. J Mol Biol 224:1087–1101

Post CB, Dobson CM, Karplus M (1989) A molecular dynamics analysis of protein structural elements. Proteins 5:337–354

Spera S, Ikura M, Bax A (1991) Measurement of the exchange rates of rapidly exchanging amide protons: application to the study of calmodulin and its complex with a myosin light chain kinase fragment. J Biomol NMR 1:155–165

Tang C, Iwahara J, Clore GM (2006) Visualization of transient encounter complexes in protein-protein association. Nature 444:383–386

Tropp J (1980) Dipolar relaxation and nuclear Overhauser effects in nonrigid molecules: the effect of fluctuating internuclear distances. J Chem Phys 72:6035–6043

Tugarinov V, Sprangers R, Kay LE (2007) Probing side-chain dynamics in the proteasome by relaxation violated coherence transfer NMR spectroscopy. J Am Chem Soc 129:1743–1750

Valentine K, Ng H, Schneeweis J, Kranz J, Frederick K, Alber T, Wand A (2009) Ultrahigh resolution crystal structure of calmodulin-smooth muscle light kinase peptide complex. PDB (to be published)

Vogel HJ (1994) The Merck Frosst award lecture 1994. Calmodulin: a versatile calcium mediator protein. Biochem Cell Biol 72:357–376

Wand AJ, Bieber RJ, Urbauer JL, McEvoy RP, Gan Z (1995) Carbon relaxation in randomly fractionally 13C-enriched proteins. J Magn Reson B 108:173–175

Wang L, Pang Y, Holder T, Brender JR, Kurochkin AV, Zuiderweg ER (2001) Functional dynamics in the active site of the ribonuclease binase. Proc Natl Acad Sci USA 98:7684–7689

Wang C, Rance M, Palmer A (2003) Mapping chemical exchange in proteins with MW &gt; 50 kD. J Am Chem Soc 125:8968–8969

Weaver DS, Zuiderweg ER (2008) Eta(z)/kappa: a transverse relaxation optimized spectroscopy NMR experiment measuring longitudinal relaxation interference. J Chem Phys 128:155103

Werbelow L, Grant D (1977) Intramolecular dipolar relaxation in multispin systems. Adv Magn Reson 9:189–299

Wilson M, Brunger A (2000) The 1.0 A crystal structure of Ca(2+)-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity. J Mol Biol 301:1237–1256

Yang D, Kay LE (1996) Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. J Mol Biol 263:369–382

Yang D, Mittermaier A, Mok YK, Kay LE (1998) A study of protein side-chain dynamics from new $^2$H auto-correlation and $^{13}$C cross-correlation NMR experiments: application to the N-terminal SH3 domain from drk. J Mol Biol 276:939–954

Zhang F, Bruschweiler R (2002) Contact model for the prediction of NMR N-H order parameters in globular proteins. J Am Chem Soc 124:12654–12655

Zhou Y, Vitkup D, Karplus M (1999) Native proteins are surface-molten solids: application of the Lindemann criterion for the solid versus liquid state. J Mol Biol 285:1371–1375