Methods 138–139 (2018) 85–92

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Utilizing dipole-dipole cross-correlated relaxation for the measurement of angles between pairs of opposing CαHα-CαHα bonds in anti-parallel β-sheets

T. Michael Sabo[a,*,1], Vytautas Gapsys[b,1], Korvin F.A. Walter[c], R. Bryn Fenwick[d], Stefan Becker[c], Xavier Salvatella[e,f], Bert L. de Groot[b], Donghan Lee[a,*], Christian Griesinger[c,*]

[a] Department of Medicine, James Graham Brown Cancer Center, University of Louisville, 505 S. Hancock St., Louisville, KY 40202, USA
[b] Department for Computational Biomolecular Dynamics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
[c] Department for NMR-Based Structural Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
[d] Department of Integrative Structural and Computational Biology, Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA
[e] Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac 10, 08028 Barcelona, Spain
[f] Institució Catalana de Recerca i Estudis AvanÅats (ICREA), Barcelona, Spain

A B S T R A C T

Dipole-dipole cross-correlated relaxation (CCR) between two spin pairs is rich with macromolecular structural and dynamic information on inter-nuclear bond vectors. Measurement of short range dipolar CCR rates has been demonstrated for a variety of inter-nuclear vector spin pairs in proteins and nucleic acids, where the multiple quantum coherence necessary for observing the CCR rate is created by *through-bond* scalar coupling. In principle, CCR rates can be measured for any pair of inter-nuclear vectors where coherence can be generated between one spin of each spin pair, regardless of both the distance between the two spin pairs and the distance of the two spins forming the multiple quantum coherence. In practice, however, long range CCR (lrCCR) rates are challenging to measure due to difficulties in linking spatially distant spin pairs. By utilizing *through-space* relaxation allowed coherence transfer (RACT), we have developed a new method for the measurement of lrCCR rates involving CαHα bonds on opposing anti-parallel β-strands. The resulting lrCCR rates are straightforward to interpret since only the angle between the two vectors modulates the strength of the interference effect. We applied our lrCCR measurement to the third immunoglobulin-binding domain of the streptococcal protein G (GB3) and utilize published NMR ensembles and static NMR/X-ray structures to highlight the relationship between the lrCCR rates and the CαHα-CαHα inter-bond angle and bond mobility. Furthermore, we employ the lrCCR rates to guide the selection of sub-ensembles from the published NMR ensembles for enhancing the structural and dynamic interpretation of the data. We foresee this methodology for measuring lrCCR rates as improving the generation of structural ensembles by providing highly accurate details concerning the orientation of CαHα bonds on opposing anti-parallel β-strands.

1. Introduction

Biophysical processes such as protein folding [1,2], molecular recognition [3,4], allostery [5,6], and enzyme catalysis [7–9] depend on the fine-tuned interplay of biomolecular structural dynamics. A technique suitable for the characterization of such dynamic processes over a broad range of time-scales and with atomic resolution is nuclear magnetic resonance (NMR) spectroscopy [10,11]. Two complementary NMR observables, residual dipolar couplings (RDCs) [3,12–14] and cross-correlated relaxation (CCR) rates [13,15–19], encode information pertaining to the ensemble averaging of bond vector orientations on physiologically important time-scales (picosecond to millisecond). Whereas RDCs provide structural [20] and dynamic [21] insight into local bond motions, CCR rates report on the angle [15] and the degree of synchronicity between a pair of bond vectors [16].

The phenomenon of CCR originates from the relaxation of spin coherences by either the dipolar coupling, defined as the *through-space* magnetic interaction between two nuclei, or the chemical shift

* Corresponding authors.
E-mail addresses: mike.sabo@louisville.edu (T.M. Sabo), donghan.lee@louisville.edu (D. Lee), cigr@nmr.mpibpc.mpg.de (C. Griesinger).
[1] T. Michael Sabo and Vytautas Gapsys contributed equally to this work.

https://doi.org/10.1016/j.ymeth.2018.04.007
Received 9 February 2018; Received in revised form 2 April 2018; Accepted 10 April 2018
Available online 12 April 2018

anisotropy (CSA) [22,23]. The dipolar couplings and CSA result in the stochastic modulation of local magnetic fields that interferes with another relaxation mechanism (another unique dipolar coupling or CSA relaxation mechanism). The strength of the interference is proportional to the rotational correlation time (τ_c) of the biomolecule and dependent on both the relative orientation of the two interfering relaxation mechanisms and the distance between the nuclei involved in one of the two relaxation phenomena. The advantage of dipole-dipole CCR, or dipolar CCR, is the straightforward interpretation of the measured rate in terms of an angle between the two vectors participating in the CCR whereas CCR involving CSA, whether CSA-CSA CCR or CSA-dipole CCR, is harder to interpret due to difficulties in translating chemical shift data into structural information and accurately determining the CSA tensor. This direct interpretation has enabled dipolar CCR rates to yield information about torsion angles in proteins [15,16,24–26] and nucleic acids [27,28]. Furthermore, the CCR rate is a time-averaged observable that is scaled by ensemble-averaged dynamic processes occurring up to the msec time-scale [29]. Accordingly, the measured rate provides details concerning the level of correlation of the vectors involved in the CCR [16] and has been utilized to investigate correlated motions amongst the polypeptide backbone of the third immunoglobulin-binding domain of the streptococcal protein G (GB3) [19].

In theory, CCR rates can be measured between any two relaxation mechanisms within a macromolecule, however in practice, it is difficult to connect pairs of spatially distant spins in an efficient manner exclusively with *through-bond* scalar coupling (J-coupling). That said, the relaxation mechanisms themselves offer a route for creating the multi-spin coherences necessary for measuring long range CCR rates [30–33]. Recently, Boisbouiver and Bax utilized CCR amongst CαHα(*i*), Hα(*i*) Hα(*j*), and CαHα(*j*), where *i* and *j* denote two spatially distant bond vectors, to transfer net magnetization *through-space* across the anti-parallel β-sheets of GB3 and the HIV protease [34]. For this secondary structural element, it is the nearly linear arrangement of the opposing CαHα vectors coupled with the short distance between the non-bonded Hα atoms (between 2 and 3 Å) that enables the efficient transfer of magnetization *through-space*, termed relaxation allowed coherence transfer (RACT) [30] (Fig. 1). The measured RACT CCR rates provide both distance and angular information that can be utilized in the refinement of solution structures and biomolecular ensemble generation.

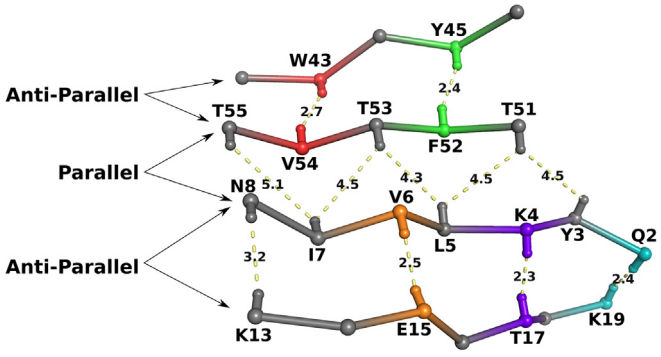

Fig. 1. Illustration of the distances between the β-strand Hα atoms of GB3 (PDB: 2OED). The efficiency of the relaxation allowed coherence transfer across the β-strands depends on the contribution of both the angle between the CαHα(*i*),Hα(*i*)Hα(*j*) vectors and the distance between the two Hα protons, $r^{-3}_{H\alpha(i)H\alpha(j)}$, to the CCR rate $\Gamma^{c}_{C\alpha H\alpha(i),H\alpha(i)H\alpha(j)}$, where *i* and *j* denote two spatially distant bond vectors [34]. The Hα atoms situated in parallel β-strands are too far apart for efficient transfer of magnetization (4.57 ± 0.29 Å), whereas the Hα atoms in anti-parallel β-strands are sufficiently close enough in space (2.59 ± 0.34 Å). The CαHα bond vector pairs that are the focus of this investigation are color coded throughout the manuscript: Q2-K19 (cyan), K4-T17 (purple), V6-E15 (orange), W43-V54 (red), and Y45-F52 (green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Here, we build upon the previously described RACT measurement to create the necessary *through-space* coherence to measure the long range CCR rate (lrCCR) between opposing backbone CαHα bonds in the anti-parallel β-sheets of GB3. The advantage of the lrCCR measurement described in the following is that it exclusively encodes the ensemble-averaged angular dependency of the CαHα(*i*) and CαHα(*j*) CCR rate, simplifying the analysis of the resulting NMR observable. We then compare our experimentally measured lrCCR rate to available NMR ensembles and NMR/X-ray structures of GB3 illustrating the structural and dynamic information contained by the lrCCR rate on the inter-bond angle and bond mobility. Finally, we utilize the lrCCR rates to generate sub-ensembles, demonstrating how lrCCRs can be used for structural ensemble generation.

2. Theory

The pulse sequence depicted in Fig. 2 was developed in order to measure the lrCCR rate ($\Gamma^{c}_{C\alpha H\alpha(1),C\alpha H\alpha(2)}$) resulting from the dipole-dipole interaction of two distinct CαHα vectors (represented by 1 and 2) pointing at each other across the anti-parallel β-sheet of GB3 (Fig. 1). After chemical shift evolution of Hα(1) and creation of the anti-phase coherence $H^1_y C^1_z$, the first relaxation allowed coherence transfer element (RACT1) at point *a* in the pulse sequence is implemented to link the two CαHα vectors through space as previously described [30,34]. An INEPT block converts the single-quantum term $H^1_y H^2_z$ into $H^1_x C^1_z H^2_z$ then followed by an *x*-pulse on ^{1}H to create $H^1_x C^1_z H^2_y$, which at point *b* is then ready to proceed through the CCR block in order to measure $\Gamma^{c}_{C\alpha H\alpha(1),C\alpha H\alpha(2)}$. The master equation describing evolution under CCR, chemical shifts and J couplings of the four multiple quantum coherences expressed as single element operators for the transverse operators that are active during the CCR block is as follows:

$$\frac{d}{dt}\begin{pmatrix} H^1_- H^2_-(t) \\ 2H^1_- C^1_z H^2_-(t) \\ 2H^1_- H^2_- C^2_z(t) \\ 4H^1_- C^1_z H^2_- C^2_z(t) \end{pmatrix} = -(L + iK)\begin{pmatrix} H^1_- H^2_-(t) \\ 2H^1_- C^1_z H^2_-(t) \\ 2H^1_- H^2_- C^2_z(t) \\ 4H^1_- C^1_z H^2_- C^2_z(t) \end{pmatrix}$$

where,

$$L = \begin{pmatrix} R_2(2H^1_y H^2_y) & 0 & 0 & \Gamma^{c}_{C\alpha H\alpha(1),C\alpha H\alpha(2)} \\ 0 & R_2(4H^1_x C^1_z H^2_y) & \Gamma^{c}_{C\alpha H\alpha(1),C\alpha H\alpha(2)} & 0 \\ 0 & \Gamma^{c}_{C\alpha H\alpha(1),C\alpha H\alpha(2)} & R_2(4H^1_y H^2_x C^2_z) & 0 \\ \Gamma^{c}_{C\alpha H\alpha(1),C\alpha H\alpha(2)} & 0 & 0 & R_2(8H^1_x C^1_z H^2_x C^2_z) \end{pmatrix}$$

$$K = \begin{pmatrix} \omega_{H\alpha(1)} + \omega_{H\alpha(2)} & \pi J_{C\alpha H\alpha(1)} & \pi J_{C\alpha H\alpha(2)} & 0 \\ \pi J_{C\alpha H\alpha(1)} & \omega_{H\alpha(1)} + \omega_{H\alpha(2)} & 0 & \pi J_{C\alpha H\alpha(2)} \\ \pi J_{C\alpha H\alpha(2)} & 0 & \omega_{H\alpha(1)} + \omega_{H\alpha(2)} & \pi J_{C\alpha H\alpha(1)} \\ 0 & \pi J_{C\alpha H\alpha(2)} & \pi J_{C\alpha H\alpha(1)} & \omega_{H\alpha(1)} + \omega_{H\alpha(2)} \end{pmatrix}.$$

Here, $R_2(M)$ represents the autorelaxation rate for each coherence *M*, ω is the Larmor frequency, and $J_{C\alpha H\alpha}$ is the scalar coupling between spins Cα and Hα. For the rest of the discussion, we convert the evolution of the single spin transverse operators into Cartesian operators. It should be noted that $R_2(4H^1_x C^1_z H^2_y) \approx R_2(4H^1_y H^2_x C^2_z) \approx R_2(2H^1_y H^2_y) \approx R_2(8H^1_x C^1_z H^2_x C^2_z)$ and, in the absence of internal dynamics, will be considered to be all equal, R_2.

During the mixing time of the CCR block (T), the coherence $4H^1_y H^2_x C^2_z$ is created only through CCR when $\Delta = 0$ (I^{cross}) and through both scalar coupling and CCR when $\Delta = 1/2J_{CH}$ (I^{ref}):

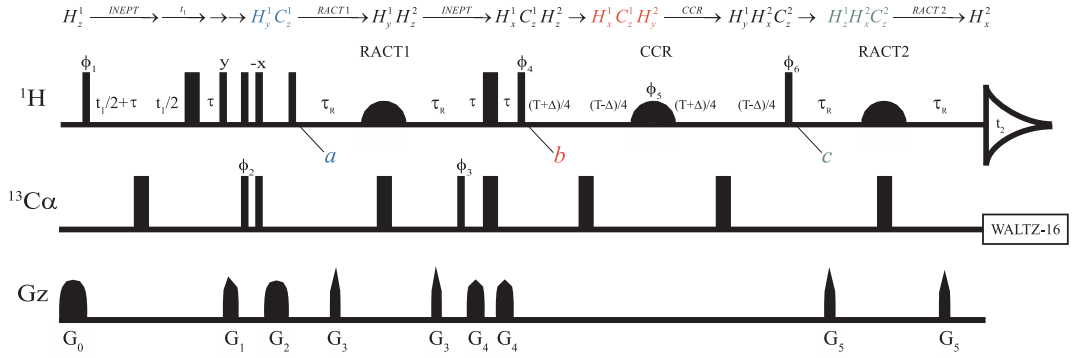

Fig. 2. Across β-sheet cross-correlated relaxation (CCR) Pulse Sequence. Above the sequence, the relevant product operators are indicated. The narrow bars have a flip angle of 90° and the wide bars have a flip angle of 180°. All pulses are phase x unless marked otherwise. ϕ_1 = x, −x; ϕ_2 = 2(x), 2(−x); ϕ_3 = 16(y), 16(−y); ϕ_4 = 4(x), 4(−x); ϕ_5 = 32(x), 32(y); ϕ_6 = 8(x), 8(−x), and ϕ_{REC} = −x, 2(x), −x, x, 2(−x), x, x, 2(−x), x, −x, 2(x), −x. Shaped ^{1}H pulses are of the 180° REBURP type (2.67 ms duration at 700 MHz for a 2.1 ppm bandwidth inversion). The shaped pulse is necessary for suppressing $^3J_{HH}$ dephasing. The carriers are centered at 4.7 ppm for ^{1}H and 57 ppm for ^{13}C. Delays utilized are $\tau = 1/4J_{C\alpha H\alpha}$ = 1.74 ms, $\tau_R = 3/2J_{C\alpha H\alpha}$ = 10.5 ms, and T = 13 ms. Δ for the cross measurement is 0 ms and for the reference measurement is 3.49 ms ($1/2J_{C\alpha H\alpha}$). Quadrature detection in the F1 dimension is obtained by States-TPPI of ϕ_1. The duration and peak strengths of the sine-shaped gradients are G_0 = 1.0 ms, 80%; G_1 = 0.8 ms, 30%; G_2 = 1.0 ms, 65%; G_3 = 1.0 ms, 15%; G_4 = 1.0 ms, 37%; G_5 = 1.0 ms, 17%.

$$
\begin{aligned}
4H_x^1C_z^1H_y^2 \rightarrow\ & 4H_x^1C_z^1H_y^2[cosh(\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}T)cos^2(\pi J_{C\alpha H\alpha}) \\
& -sinh(\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}T)sin^2(\pi J_{C\alpha H\alpha}\Delta)] \\
& -4H_y^1H_x^2C_z^2[sinh(\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}T)cos^2(\pi J_{C\alpha H\alpha}\Delta) \\
& -cosh(\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}T)sin^2(\pi J_{C\alpha H\alpha}\Delta)].
\end{aligned}
$$

At point *c*, the RACT2 element converts $4H_y^1H_x^2C_z^2$ to H_x^2 for detection. In the end, the intensity ratio of the cross to the reference measurement will provide $\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}$ [27]:

$$\frac{I^{cross}}{I^{ref}} = tanh(\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}T)$$

3. Materials and Methods

3.1. NMR sample preparation and measurement

Uniformly ^{13}C-labeled recombinant GB3 was expressed in *E. coli* in minimal media supplemented with ^{13}C-glucose as the sole carbon source and purified as described before [35]. NMR experiments were performed on a 5 mM GB3 sample in 20 mM sodium phosphate buffer at pH 5.5 and 0.1% sodium azide. The long range CCR rates (lrCCR) were measured at 288 K with the pulse sequence depicted in Fig. 2 on a 700 MHz Bruker AVANCE spectrometer equipped with a 5-mm TCI cryogenic probe head.

3.2. Structures and structural ensembles

For the structural interpretation of the lrCCR measurements, we investigated four NMR ensembles ERMD [36], EM8 [37], eNOE [38], and CCR16 [19], two X-ray structures 1IGD and 1IGC [39] and five single structures refined or determined by means of NMR: structure by Li et al. (2015) [40]; 2OED, 1P7E and 1P7F [41]; 2N7J [42]. All the NMR ensembles and structures have been generated utilizing residual dipolar couplings (RDC). In addition to the RDCs, the CCR16 ensemble used backbone CCR restraints, while the eNOE ensemble used exact NOEs, scalar couplings and chemical shifts. 2N7J is a 20-conformer ensemble. In the current work, however, we refer to 2N7J as a single NMR derived structure, because every conformer in the ensemble has identical backbone atom positions and the difference between the conformers comes only from the side-chain rotameric states. Prior to calculating RDCs and order parameters (S^2), the structural ensembles were superimposed on the backbone atoms by minimizing variance over the ensembles [43]. The average angles between the bond vectors in an ensemble were obtained by calculating the angle between the average bond vectors in the investigated ensembles. To analyze the CαHα and NH bond directions in the X-ray structures, the hydrogen atom positions were constructed based on a set of geometric rules. The Hα atoms were placed on a line going through the Cα atom and an averaged position of the N, C' and Cβ atoms [44]. For the NH bond vectors, the H atoms were placed in a plane defined by the N, C' and Cα atoms and on a line bisecting the angle C'-N-Cα, such that the angles C'-N-H and Cα-N-H are larger than 90°. The CαHα and NH bond lengths were set to 0.1 nm, however, for the subsequent RDC and CCR calculations the bond vectors of the structures and structural ensembles were normalized, thus having no impact on the calculated NMR observables. The 1IGD structure contains GB3 in its unbound form, while in the case of 1IGC, GB3 is bound to an antibody fragment.

3.3. RDC, CCR and S^2 order parameters

Every structure and structural ensemble were analyzed by calculating the RDCs and CCR rates and comparing them to the experimentally measured values. The experimental RDC values were collected from the literature [40,41,45]. In total, RDC measurements for the CαHα and NH bond vectors in 17 alignment media were used. For every structure/ensemble, 17 alignment tensors were calculated by means of a singular value decomposition (SVD) fit [46]. Subsequently, these alignment tensors were used to calculate the CαHα RDCs for the bonds of interest and, later, we refer to these local structural assessments as the local quality of structures or structural ensembles.

The experimental CCR rates over the backbone CαHα and NH bonds were reported in [47]. In this work, we were interested in the structural and dynamic properties of the CαHα bond vectors for which the lrCCR rates were measured. Therefore, three backbone CCR rate sets involving CαHα bond vectors were used: intra-residue and sequential CαHα-NH, as well as sequential CαHα-CαHα rates. For every structure/ensemble and each CCR set, the contribution of the libration order parameters, $S^{libr}_{XH}S^{libr}_{YH}$, was taken from a fit to the experimental data (the influence of this procedure is investigated in more detail in the SI Fig. 3). Subsequently, only those local backbone CCR rates reported for vectors participating in the lrCCR measurement were considered. The lrCCRs were calculated following the same procedure as for the backbone CCR rates. For the multi-structure NMR ensembles, $S^2_{ensemble}$ order parameters were calculated and compared to the model free order parameters, S^2_{RDC}, obtained from the ORIUM [48] and DIDC [45] procedures.

3.4. Q Values

The quality of the agreement between the structural GB3

representation and the experiment was quantified by Q value:

$$Q = \sqrt{\frac{\sum (x_i - d_i)^2}{\sum d_i^2}}$$

where x_i is a calculated RDC or CCR value and d_i is an experimentally measured corresponding value. The RDC Q values were calculated for every CαHα bond vector of interest by summing over the measurements in 17 alignment media. For the backbone CCR Q values, the summation was performed over the intra-residue and sequential CαHα-NH, as well as sequential CαHα-CαHα measurements by considering those entries which had a contribution from the analyzed CαHα bonds. For the lrCCR Q values, a Q value was calculated for every investigated CαHα bond vector pair by considering a single lrCCR value measured in this work. For the cases where two lrCCR values have been measured for one bond vector pair, an average experimental value was used.

3.5. Error estimation

The standard errors for the Q values, S^2 order parameters and angles of the NMR ensembles were calculated by means of bootstrapping. For the lrCCR Q values, the standard errors represent a combined error of bootstrapping and experimental uncertainty. Structural sub-ensembles were generated from the ERMD, EM8 and eNOE ensembles minimizing the lrCCR Q values for the five investigated CαHα bond vectors, as well as the overall backbone CCR and RDC agreement with the experimental measurements.

4. Results and discussion

4.1. NMR measurement of long-range CCR rates

We report CCR rates for 5 pairs of CαHα bonds, Q2-K19, K4-T17, V6-E15, W43-V54, and Y45-F52 (Table 1 and Fig. 3). The magnitude of $\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}$ depends on the projection angle (θ_{12}) between the two CαHα vectors [15]:

$$\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)} = \left(\frac{\mu_0 h \gamma_H \gamma_C}{8\pi^2 r^3_{C\alpha H\alpha}}\right)^2 \left\langle \frac{3\cos^2\theta_{12} - 1}{2} \right\rangle \frac{2\tau_c}{5} S^{libr}_{C\alpha H\alpha(1)} S^{libr}_{C\alpha H\alpha(2)}$$

where μ_0 is the permeability of a vacuum, h is the Planck constant, γ_H and γ_C are the gyromagnetic ratios of ^{1}H and ^{13}C, respectively, $r_{C\alpha H\alpha}$ is the CαHα bond length (1.09 Å), and τ_c is the rotational correlation time. Angular brackets represent ensemble-averaging.

An advantage of the dipolar CCR rates is that they scale with τ_c. Thus, the time for the RACT transfer decreases with increasing molecular weight, opening up the possibility of applying our across β-sheet measurement to larger systems, as demonstrated with the RACT transfer for the HIV protease [34]. The measured lrCCR rates range in value from ~29 Hz to 41 Hz and were used in the following structural and dynamic analysis of the CαHα vector pairs in GB3.

Table 1
$\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}$ measured for GB3.

		$\Gamma^c_{C\alpha H\alpha(1),C\alpha H\alpha(2)}$ (Hz)	error
Q2	K19	37.6	1.5
K19	Q2	41.0	2.0
K4	T17	31.8	1.3
T17	K4	34.2	1.3
V6	E15	34.2	3.1
E15	V6	28.8	3.0
W43	V54	30.9	2.8
V54	W43	30.1	3.2
Y45	F52	37.4	2.7
F52	Y45	n/a	n/a

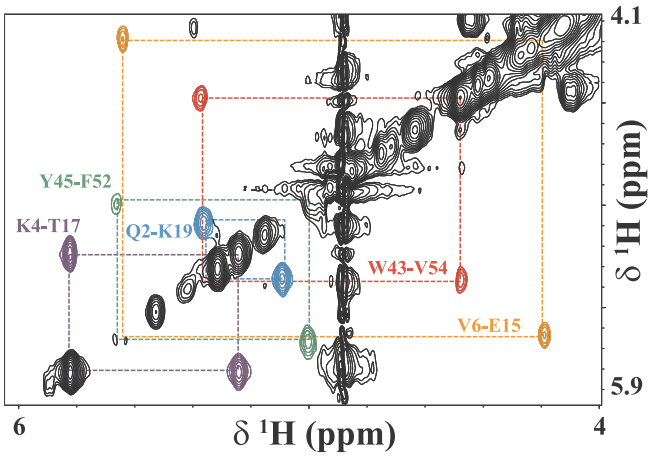

Fig. 3. Representative spectrum of the across β-sheet CCR pulse sequence measured with uniformly ^{13}C-labeled GB3, recorded at a proton frequency of 700 MHz and a temperature of 288 K. The concentration of GB3 was 5 mM in D_2O, containing 20 mM sodium phosphate at pH 5.5 and 0.1% NaN_3. The spectrum was recorded with 1024 and 43 complex points in the direct (t_2) and indirect (t_1) dimensions, respectively, with 1600 scans per t_1 increment. For the cross measurement, 3200 scans per t_1 were recorded. The $t_{1,max}$ and $t_{2,max}$ were 15.3 ms and 122 ms, respectively. The reference experiment was recorded in 2 days, while the cross experiment was recorded in 4 days. Frequency discrimination in the indirectly detected dimension was achieved with the States-TPPI scheme. The spectrum was processed with NMRPipe software. The average signal to noise for the reference and cross measurements were 50 and 19, respectively.

4.2. Structural and dynamic interpretation

For the structural interpretation of the measured CCR rates, we investigated a number of NMR conformational ensembles as well as single structures resolved by means of X-ray crystallography or NMR (see Methods for the description and literature references). First, we calculated the angle between the CαHα bond vector pairs of interest in every structure or structural ensemble (Fig. 4, left panel). Furthermore, we provide estimates of quality at which a particular bond vector is represented. The quality is assessed by means of the Q values: one Q value is calculated for the lrCCR rates per every bond vector pair; two Q values are reported for every bond vector in a pair when considering RDCs and backbone CCR rates. While the lrCCRs depend solely on the bond vector angles of interest, the RDC and backbone CCR rates provide information about the structural quality of the bond pair in a context of other bond vectors: the RDC calculation depends on the alignment tensor; backbone CCR rates depend on the neighboring NH or CαHα bond vectors.

The largest of the analyzed NMR ensembles, ERMD (640 structures) has been derived by fulfilling RDC restraints. The ensemble shows an overall good agreement with the experimental CCR rates and the 10 RDCs for all five bond vector pairs: in all cases the Q value does not exceed 0.12. These 10 RDCs, 10 backbone CCR rates and the 5 lrCCR rates measure the local quality of structures or structural ensembles. Regarding this local quality, ERMD is similar to the single structures resolved or refined by NMR. Consistent with the comparable local quality of the ERMD ensemble and single NMR structures, also the five inter-bond vector angles (see Methods for the details on the ensemble angle calculations) attain similar values. Another RDC-based ensemble, EM8, performs consistently slightly worse in the local quality measures than ERMD and the single NMR structures. An exception is the V6-E15 CαHα bond pair where EM8 shows the lowest long range CCR Q value. The third multi-structure NMR ensemble was constructed by utilizing exact NOE-based distance restraints (eNOEs), RDCs, scalar couplings and chemical shifts. This eNOE ensemble performs particularly poorly for bond vectors W43-V54 and Y45-F52, while for the vectors in the

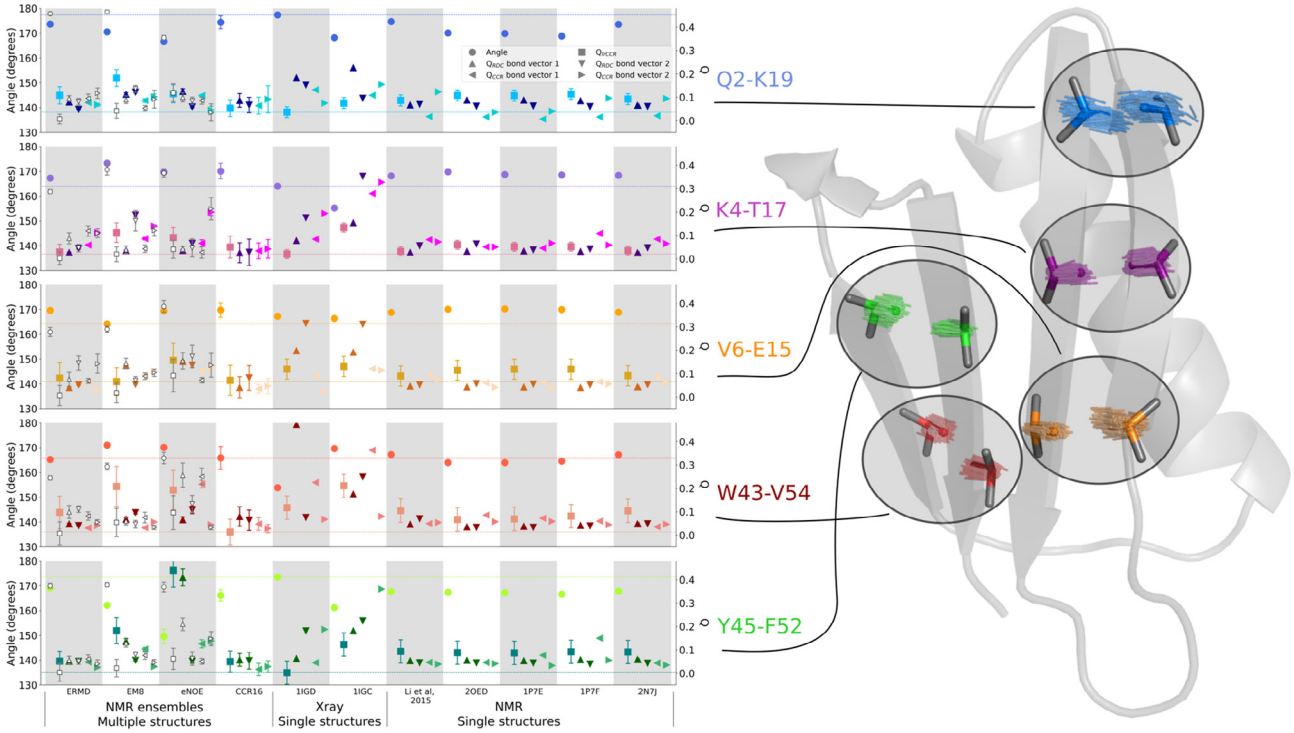

Fig. 4. CαHα bond vector representation in the NMR ensembles, single NMR and Xray structures and sub-selected ensembles. (Left) The angles between the CαHα bond vectors are depicted by circle symbols and marked on the left y-axis. The quality with which every CαHα bond pair reproduces the experimentally measured observables is expressed in terms of Q values and is marked on the right y-axis. The following Q values were calculated: one Q value for the lrCCR rates for every CαHα bond vector pair (square symbol), two Q values – one for each CαHα bond vector – for the RDCs measured in 17 alignment media (up/down triangle symbols), two Q values for the backbone intra-residue, sequential and CαHα-CαHα CCR rates (left/right triangle symbols). The lines in the figure mark the level of the lowest Q value for the lrCCR rate and the inter-vector angle corresponding to that structure or ensemble with the lowest lrCCR Q value. The colored symbols correspond to the values calculated for the NMR ensembles and NMR/X-ray structures, while white symbols denote the calculations for the sub-selected ensembles. (Right) A single GB3 structure is depicted in a cartoon representation. The CαHα bond vectors for which the lrCCR rates were measured are depicted in color: the bond vectors from the ERMD ensemble are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

other β-sheet it is comparable to EM8, but worse than ERMD ensemble. The poor performance of the eNOE ensemble falls in line with a recent study demonstrating that this ensemble could be improved by adding additional RDC and scalar coupling restraints [49]. CCR16 is the smallest (16 structures) of the multi-structure NMR ensembles analyzed in this work. It has been constructed by imposing restraints using both RDC and CCR experimental data sets. Overall, in our local quality comparison CCR16 performs on par with the much larger ERMD ensemble.

Of the two X-ray structures, 1IGC has been resolved in a bound state with an antibody fragment, while 1IGD was resolved without a binding partner. For the analysis of the bond vectors, we reconstructed hydrogen atom positions based on a set of simple geometric rules (see Methods), thus including no direct information derived from the NMR measurements on GB3. It is of no surprise that structures with the hydrogen atom positions reconstructed in this simplistic manner did not yield high local quality in terms of matching the experimental RDCs and backbone CCR rates for most of the bond pairs analyzed. The Q values for the lrCCR rates, however, were unexpectedly low for the 1IGD structure: for 3 bond vector pairs 1IGD performed better than any other ensemble or single structure. This suggests that while the observables that depend on the overall quality of the bond vector orientations in a structure (RDCs and backbone CCR rates) indicate only a low-level agreement with experiment, the relative geometry between some vectors may still be properly represented while the orientation with respect to the molecular frame (RDCs) does not fully agree to the experimental values. In most cases the unbound crystallographic structure shows better agreement with the NMR measurements than the bound GB3 conformer. It is an expected outcome considering that all the experimental observables used in the analysis have been measured for the protein free in solution. Furthermore, this observation implies that the local set of RDC and CCR values, if measured for a bound form of GB3, could potentially allow discriminating the subtle conformational changes between the bound and free protein structures.

The single structures based on NMR data were all constructed following similar procedures to one another and all utilized RDC based restraints. Therefore, their quality in reproducing CαHα bond vector orientations is high. Subsequently, the dynamically averaged inter vector angle values are similar for all of these structures which allows for an accurate reproduction of the lrCCR data. It needs to be noted that for the single X-ray or NMR structures that well reproduce the lrCCR values, the inter CαHα bond angles should not be interpreted as arithmetic averages of the true structural ensembles underlying the experimental measurements. Rather, a good agreement between the structure and experimental lrCCR indicates that the inter-bond angle represents well a dynamic average of the true ensemble (see Section 4.1).

So far, we have described the local quality of the ensembles and single structures, as well as the inter-bond geometry in terms of an angle subtended between the bond vectors. However, for four NMR ensembles containing multiple structures it is also possible to extract the dynamic properties of the CαHα bond motions. In the following, we will compare S^2 calculated from structural ensembles of GB3 ($S^2_{ensemble}$) versus experimentally determined order parameters (S^2_{RDC}). The definition for S^2 is given in references [11,48]. Briefly, the order parameter can take on any value between 0 and 1, where 1 denotes a rigid bond vector and 0 denotes unrestricted local motion. Both $S^2_{ensemble}$ and S^2_{RDC} represent bond vector dynamics occurring up to the msec time-scale, the time-scale encompassed by the RDCs, eNOEs, and/or CCR rates

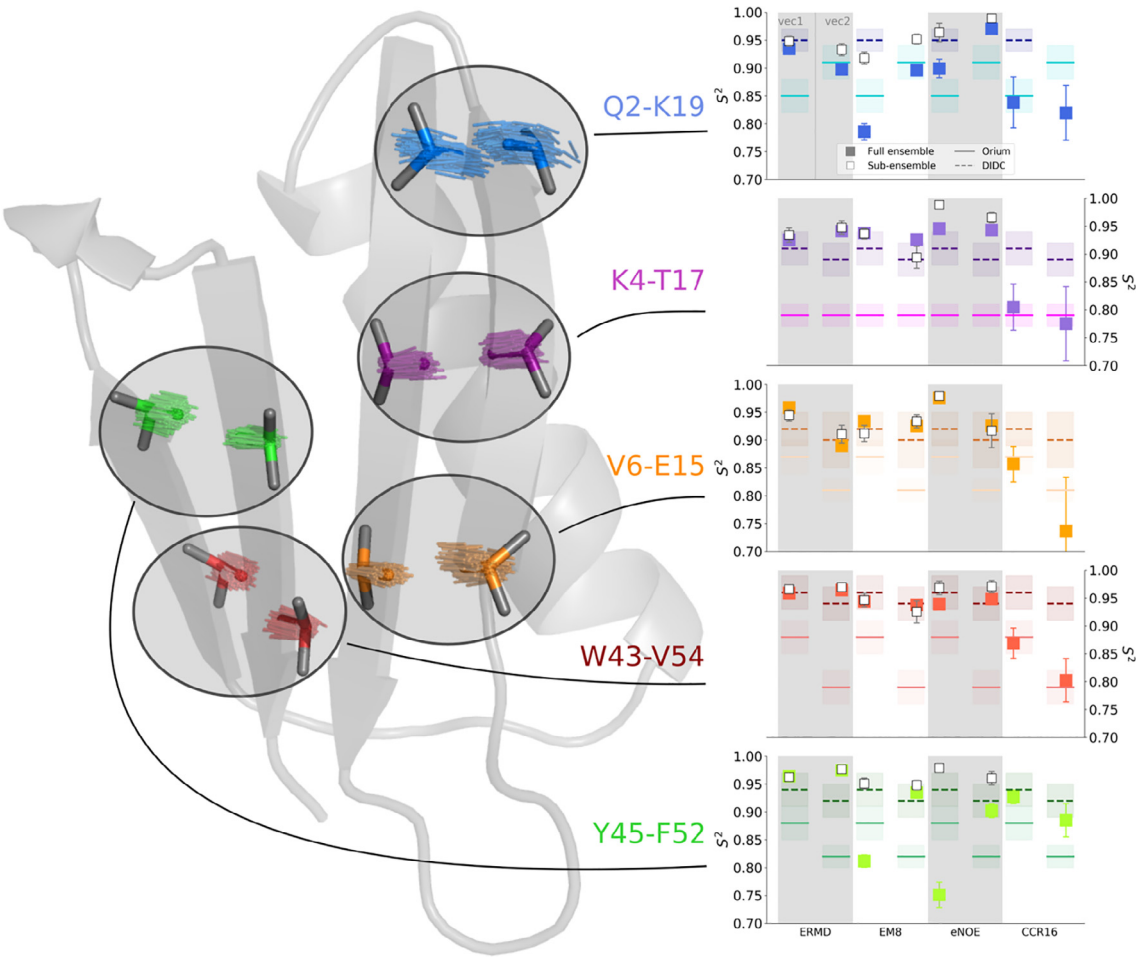

Fig. 5. (Left) A single GB3 structure is depicted in a cartoon representation. The CαHα bond vectors for which the lrCCR rates were measured are depicted in color: the bond vectors from the ERMD ensemble are shown. (Right) $S^2_{ensemble}$ for the ERMD, EM8, eNOE and CCR16 ensembles: one value for every CαHα bond vector of each investigated bond vector pairs are shown in colored symbols. The white symbols mark $S^2_{ensemble}$ for the sub-ensembles. Dotted lines denote S^2_{RDC} values obtained by the DIDC approach, solid lines denote ORIUM calculation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

utilized to generate the GB3 structural ensembles or calculate the RDC derived order parameters.

We summarize the order parameters, S^2, calculated for the 5 CαHα bond vector pairs of interest in Fig. 5, right panel. The lines in the figure denote S^2_{RDC} values obtained by the model free approaches: ORIUM [48] (solid line) and DIDC [45] (broken line). ORIUM S^2_{RDC} values are consistently smaller than those determined by DIDC, indicating that the ORIUM approach considers the CαHα bonds to be more mobile. The $S^2_{ensemble}$ values calculated for the ERMD, EM8 and eNOE structural ensembles appear to support DIDC S^2_{RDC} order parameters, suggesting that the 10 CαHα bond vector motions relevant for this work ought to be rather restricted. There are several exceptions to this observation: for the Y45 CαHα bond in EM8 and eNOE ensembles, the $S^2_{ensemble}$ are significantly lower than ORIUM estimates. However, the large local RDC Q values for these vectors indicate that the bond orientations and/or dynamics may not be optimal in these ensembles (see Fig. 4, SI Fig. 4). This argument does not hold for the Q2 CαHα bond vector in the EM8 ensemble, which shows higher mobility and still has a rather low RDC Q value.

The CCR16 ensemble is different from the other analyzed multi-structure NMR ensembles with respect to $S^2_{ensemble}$. CCR16 has consistently lower $S^2_{ensemble}$ values matching the ORIUM S^2_{RDC} values better than the DIDC S^2_{RDC}, except for the Y45-F52 CαHα bond vectors. The CCR16 ensemble allows for a larger mobility of the CαHα bond vectors, which corresponds well to the libration motion analysis in the SI Fig. 3: the calculated S^{libr} values for CCR16 are largest among all the ensembles and single structures in the study. This indicates that little motion is absorbed into the alignment tensor for CCR16, and in turn the bond vectors represented by the conformers in the ensemble are more dynamic.

4.3. Selected sub-ensembles confirm the observed structural and dynamic trends

The detailed investigation into geometry and mobility of the CαHα bond vectors in the NMR ensembles and X-ray/NMR structures provided basic structural and dynamic interpretation for the long range CCR measurements. To further explain the influence of the attained lrCCR values for the GB3 bond vector orientations, we used the respective lrCCR measurements to generate new structural ensembles from the ERMD, EM8 and eNOE ensembles. This approach allowed obtaining sub-selections of structures [50] from each of the NMR ensembles such that the lrCCR values were optimized in addition to the RDCs and backbone CCR rates (details described below).

The existing NMR ensembles (ERMD, EM8 and eNOE) served as pools of structures for the procedure. From each of the conformer pools, we selected a two-structure sub-ensemble, which would reproduce the experimentally measured lrCCR values as well as possible. The sub-ensemble size of two structures allowed introducing a dynamic component in contrast to using a single structure, which would only be able

to represent the mean angle direction for the CαHα bond vectors. Having obtained the best matching sub-ensemble, the two identified structures were removed from the conformer pool and the procedure was repeated to find the second best two-structure sub-ensemble to reproduce lrCCR. Iterating these steps, the NMR ensembles were divided into sub-ensembles of decreasing quality in representing lrCCR rates. Subsequently, we combined the sub-ensembles in a stepwise manner by appending them one-by-one, starting from the ensemble showing the best agreement with the experiment in terms of the lrCCR values.

For the cumulative ensembles, the quality of lrCCR agreement with experiment decreases with an increase in the ensemble size (SI Fig. 1). The Q values for all the available RDC (all CαHα and NH RDCs obtained in 17 alignment media [40,41,45]) and all backbone CCR (intra NH-CαHα, sequential CαHα-NH, sequential CαHα-CαHα, and sequential NH-NH [47]) measurements initially decrease with the increasing ensemble size, but saturate quickly before reaching 10% of the respective NMR ensemble size. Finally, among the cumulative ensembles of different size we identified the ones minimizing the averaged Q value for the long range and backbone CCRs (all available intra NH-CαHα, sequential CαHα-NH, sequential CαHα-CαHα, and sequential NH-NH CCR rates), as well as the RDCs (all available CαHα and NH RDCs obtained in 17 alignment media). The final selected sub-ensembles contained 32 structures for ERMD (originally ERMD contains 640 structures), 38 for EM8 (originally 500 structures) and 6 for eNOE (originally 60 structures).

The analysis of the CαHα bond vector angles and quality of reproducing experimental measurements in terms of the Q values was performed for the sub-ensembles in analogy to the analysis described for the other NMR/X-ray structures (white symbols in the Fig. 4, left panel). It is interesting to note that for 12 out of 15 cases the sub-selection shifted the inter-vector angle towards the value previously identified from a structure or an ensemble best matching lrCCR measurements (marked by line in Fig. 4, left panel).

Similarly, the sub-ensemble $S^2_{ensemble}$ values (Fig. 5, right panel) shift the largest outliers in the NMR ensembles towards the S^2_{RDC} values obtained by DIDC. The largest change was observed in the $S^2_{ensemble}$ values for the Y45 CαHα bond in eNOE sub-ensemble. For this vector both, the angle and order parameter, have attained the values consistent with the calculations from the other ensembles. In addition, the eNOE sub-ensemble showed markedly lower RDC and CCR Q values for the Y45 CαHα bond vector. Another interesting case is the EM8 sub-ensemble's Q2 CαHα bond vector: here, the RDC and backbone CCR rate quality has not changed much, while the lrCCR quality has increased. While the angle in this case is altered, the main contribution to the change appears to come from the changes in the orientation distribution as illustrated by the order parameter: the $S^2_{ensemble}$ value for the sub-ensemble significantly shifts closer to the S^2_{RDC} calculated by the DIDC method.

In general, the sub-ensembles selected based on the lrCCR data are comparable in quality in terms of RDCs and backbone CCR rates to the full size structural ensembles that served as conformational pools for selection (SI Fig. 2). Interestingly, at the same time randomly selected sub-ensembles also show comparable quality, suggesting that it is not the feature of the lrCCRs to restrict the number of conformations required to reproduce RDCs and backbone CCR rates, but rather a feature of the large ensembles (ERMD, EM8 and eNOE): even random sub-selections of conformers match RDCs and backbone CCR rates well.

All in all, the lrCCRs appear to drive the ensembles towards a defined set of inter CαHα bond vector angles and order parameters. Admittedly, in the current work only 5 lrCCR values were used to guide the ensemble generation which imposes a relatively weak requirement for the sub-ensemble selection procedure that can be fulfilled with a low number of structures. This does not imply, however, that the lrCCR measurements support an overall decrease in the structural heterogeniety of the β-strands. To faithfully represent the dynamics of all the bond vectors in a protein, i.e. more than the currently investigated 5 CαHα bond vector pairs, larger ensembles constructed similarly to ERMD, EM8 or eNOE may be required. The lrCCR measurements could be considered as an additional set of restraints in the future ensemble generation. Another option is to utilize these observables as cross-validation of the structural ensembles constructed based on other structural and dynamic data.

5. Conclusions

In summary, we present a new method for measuring long range dipole-dipole CCR rates with NMR spectroscopy between opposing backbone CαHα bonds in the anti-parallel β-sheets. The acquired evidence from the NMR ensembles and NMR/X-ray structures shows that the measured lrCCR rates carry structural and dynamic information on the inter-bond angle and bond mobility. We foresee this methodology for measuring lrCCR rates as enhancing future structural ensemble generation efforts.

Acknowledgments

This work was supported by funds from the James Graham Brown Cancer Center (to D.L and T.M.S.), the National Center for Research Resources CoBRE 1P30GM106396, the Max Planck Society, and the EU (ERC grant agreement number 233227 to C.G.). We thank Claudia Schwiegk for excellent technical help.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ymeth.2018.04.007.

References

[1] D.M. Korzhnev, X. Salvatella, M. Vendruscolo, A.A. Di Nardo, A.R. Davidson, C.M. Dobson, et al., Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR, Nature 430 (2004) 586–590.
[2] K. Lindorff-Larsen, S. Piana, R.O. Dror, D.E. Shaw, How fast-folding proteins fold, Science 334 (2011) 517–520.
[3] O.F. Lange, N.-A. Lakomek, C. Farès, G.F. Schröder, K.F.A. Walter, S. Becker, et al., Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution, Science 320 (2008) 1471–1475.
[4] K.K. Frederick, M.S. Marlow, K.G. Valentine, A.J. Wand, Conformational entropy in molecular recognition by proteins, Nature 448 (2007) 325–329.
[5] A.J. Wand, Bringing disorder and dynamics in protein allostery into focus, Proc. Natl. Acad. Sci. U.S.A. 114 (2017) 4278–4280.
[6] H.N. Motlagh, J.O. Wrabl, J. Li, V.J. Hilser, The ensemble nature of allostery, Nature 508 (2014) 331–339.
[7] E.Z. Eisenmesser, O. Millet, W. Labeikovsky, D.M. Korzhnev, M. Wolf-Watz, D.A. Bosco, et al., Intrinsic dynamics of an enzyme underlies catalysis, Nature 438 (2005) 117–121.
[8] S.K. Whittier, A.C. Hengge, J.P. Loria, Conformational motions regulate phosphoryl transfer in related protein tyrosine phosphatases, Science 341 (2013) 899–903.
[9] D.D. Boehr, D. McElheny, H.J. Dyson, P.E. Wright, The dynamic energy landscape of dihydrofolate reductase catalysis, Science 313 (2006) 1638–1642.
[10] A.G. Palmer, NMR characterization of the dynamics of biomacromolecules, Chem. Rev. 104 (2004) 3623–3640.
[11] D. Ban, T.M. Sabo, C. Griesinger, D. Lee, Measuring dynamic and kinetic information in the previously inaccessible supra-τ(c) window of nanoseconds to microseconds by solution NMR spectroscopy, Molecules 18 (2013) 11904–11937.
[12] G. Bouvignies, P. Bernadó, S. Meier, K. Cho, S. Grzesiek, R. Brüschweiler, et al., Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 13885–13890.
[13] R.B. Fenwick, S. Esteban-Martín, B. Richter, D. Lee, K.F.A. Walter, D. Milovanovic, et al., Weak long-range correlated motions in a surface patch of ubiquitin involved in molecular recognition, J. Am. Chem. Soc. 133 (2011) 10336–10339.
[14] L. Salmon, G. Bascom, I. Andricioaei, H.M. Al-Hashimi, A general method for constructing atomic-resolution RNA ensembles using NMR residual dipolar couplings: the basis for interhelical motions revealed, J. Am. Chem. Soc. 135 (2013) 5457–5466.
[15] B. Reif, M. Hennig, C. Griesinger, Direct measurement of angles between bond vectors in high-resolution NMR, Science 276 (1997) 1230–1233.
[16] P. Pelupessy, S. Ravindranathan, G. Bodenhausen, Correlated motions of successive amide N-H bonds in proteins, J. Biomol. NMR 25 (2003) 265–280.
[17] B. Vögeli, L. Yao, Correlated dynamics between protein HN and HC bonds observed

by NMR cross relaxation, J. Am. Chem. Soc. 131 (2009) 3668–3678.
[18] R.B. Fenwick, B. Vögeli, Detection of correlated protein backbone and side-chain angle fluctuations, Chembiochem 18 (2017) 2016–2021.
[19] R.B. Fenwick, C.D. Schwieters, B. Vögeli, Direct investigation of slow correlated dynamics in proteins via dipolar interactions, J. Am. Chem. Soc. 138 (2016) 8412–8421.
[20] N. Tjandra, A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, Science 278 (1997) 1111–1114.
[21] J.R. Tolman, J.M. Flanagan, M.A. Kennedy, J.H. Prestegard, NMR evidence for slow collective motions in cyanometmyoglobin, Nat. Struct. Mol. Biol. 4 (1997) 292–297.
[22] B. Reif, A. Diener, M. Hennig, M. Maurer, C. Griesinger, Cross-correlated relaxation for the measurement of angles between tensorial interactions, J. Magn. Reson. 143 (2000) 45–68.
[23] D. Frueh, Internal motions in proteins and interference effects in nuclear magnetic resonance, Prog. Nucl. Magn. Reson. Spectrosc. 41 (2002) 305–324.
[24] P. Pelupessy, E. Chiarparin, R. Ghose, G. Bodenhausen, Efficient determination of angles subtended by C(alpha)-H(alpha) and N-H(N) vectors in proteins via dipole-dipole cross-correlation, J. Biomol. NMR 13 (1999) 375–380.
[25] E. Chiarparin, P. Pelupessy, R. Ghose, G. Bodenhausen, Relative orientation of CαHα-bond vectors of successive residues in proteins through cross-correlated relaxation in NMR, J. Am. Chem. Soc. 122 (2000) 1758–1761.
[26] P. Pelupessy, E. Chiarparin, R. Ghose, G. Bodenhausen, Simultaneous determination of Psi and Phi angles in proteins from measurements of cross-correlated relaxation effects, J. Biomol. NMR. 14 (1999) 277–280.
[27] I.C. Felli, C. Richter, C. Griesinger, H. Schwalbe, Determination of RNA sugar pucker mode from cross-correlated relaxation in solution NMR spectroscopy, J. Am. Chem. Soc. 121 (1999) 1956.
[28] C. Richter, B. Reif, C. Griesinger, H. Schwalbe, NMR Spectroscopic Determination of Angles α and ζ in RNA from CH-Dipolar Coupling, P-CSA Cross-Correlated Relaxation, J. Am. Chem. Soc. 122 (2000) 12728–12731.
[29] B. Vögeli, Comprehensive description of NMR cross-correlated relaxation under anisotropic molecular tumbling and correlated local dynamics on all time scales, J. Chem. Phys. 133 (2010) 014501.
[30] S. Wimperis, G. Bodenhausen, Relaxation-allowed transfer of coherence in NMR between spins which are not scalar coupled, Chem. Phys. Lett. 140 (1987) 41–45.
[31] S. Wimperis, G. Bodenhausen, Relaxation-allowed cross-peaks in two-dimensional NMR correlation spectroscopy, Mol. Phys. 66 (1989) 897–919.
[32] R. Brüschweiler, R.R. Ernst, Molecular dynamics monitored by cross-correlated cross relaxation of spins quantized along orthogonal axes, J. Chem. Phys. 96 (1998) 1758–1766.
[33] C. Dalvit, 1H to 15N polarization transfer via ^{1}H chemical-shift anisotropy — ^{1}H–^{15}N dipole-dipole cross correlation, J. Magn. Reson. 97 (1992) 645–650.
[34] J. Boisbouvier, A. Bax, Long-range magnetization transfer between uncoupled nuclei by dipole-dipole cross-correlated relaxation: a precise probe of beta-sheet geometry in proteins, J. Am. Chem. Soc. 124 (2002) 11038–11045.
[35] A.M. Gronenborn, D.R. Filpula, N.Z. Essig, A. Achari, The immunoglobulin binding domain of streptococcal protein G has a novel and highly stable polypeptide fold, Science 253 (1991) 657–661.
[36] S. Pratihar, T.M. Sabo, D. Ban, R.B. Fenwick, S. Becker, X. Salvatella, et al., Kinetics of the antibody recognition site in the third IgG-binding domain of protein G, Angew. Chem. Int. Ed. Engl. (2016).
[37] S. Olsson, B.R. Voegeli, A. Cavalli, W. Boomsma, J. Ferkinghoff-Borg, K. Lindorff-Larsen, et al., Probabilistic determination of native state ensembles of proteins, J. Chem. Theory Comput. 10 (2014) 3484–3491.
[38] B. Vögeli, S. Kazemi, P. Güntert, R. Riek, Spatial elucidation of motion in proteins by ensemble-based structure calculation using exact NOEs, Nat. Struct. Mol. Biol. 19 (2012) 1053–1057.
[39] J.P. Derrick, D.B. Wigley, The third IgG-binding domain from streptococcal protein G, J. Mol. Biol. 243 (1994) 906–918.
[40] F. Li, J.H. Lee, A. Grishaev, J. Ying, A. Bax, High accuracy of Karplus equations for relating three-bond J couplings to protein backbone torsion angles, Chemphyschem 16 (2015) 572–578.
[41] T.S. Ulmer, B.E. Ramirez, F. Delaglio, A. Bax, Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy, J. Am. Chem. Soc. 125 (2003) 9179–9191.
[42] F. Li, A. Grishaev, J. Ying, A. Bax, Side chain conformational distributions of a small protein derived from model-free analysis of a large set of residual dipolar couplings, J. Am. Chem. Soc. 137 (2015) 14798–14811.
[43] V. Gapsys, B.L. de Groot, Optimal superpositioning of flexible molecule ensembles, Biophys. J. 104 (2013) 196–207.
[44] S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, et al., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29 (2013) 845–854.
[45] L. Yao, B. Voegeli, D.A. Torchia, A. Bax, Simultaneous NMR study of protein structure and dynamics using conservative mutagenesis, J. Phys. Chem. B 112 (2008) 6045–6056.
[46] J.A. Losonczi, M. Andrec, M.W.F. Fischer, J.H. Prestegard, Order matrix analysis of residual dipolar couplings using singular value decomposition, J. Magn. Reson. 138 (1999) 334–342.
[47] B. Vögeli, Cross-correlated relaxation rates between protein backbone H-X dipolar interactions, J. Biomol. NMR 67 (2017) 211–232.
[48] T.M. Sabo, C.A. Smith, D. Ban, A. Mazur, D. Lee, C. Griesinger, ORIUM: optimized RDC-based Iterative and Unified Model-free analysis, J. Biomol. NMR 58 (2014) 287–301.
[49] B. Vögeli, S. Olsson, P. Güntert, R. Riek, The exact NOE as an alternative in ensemble structure determination, Biophys. J. 110 (2016) 113–126.
[50] S. Xiang, V. Gapsys, H.-Y. Kim, S. Bessonov, H.-H. Hsiao, S. Möhlmann, et al., Phosphorylation drives a dynamic switch in serine/arginine-rich proteins, Structure 21 (2013) 2162–2174.