

# Residual dipolar couplings as a tool to study molecular recognition of ubiquitin

Nils-Alexander Lakomek\*, Oliver F. Lange†, Korvin F.A. Walter\*, Christophe Farès‡, Dalia Egger§, Peter Lunkenheimer§, Jens Meiler¶, Helmut Grubmüller†, Stefan Becker\*, Bert L. de Groot† and Christian Griesinger\*<sup>1</sup>

\*Department for NMR-based Structural Biology, Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany, †Department for Theoretical and Computational Biophysics, Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany, ‡University Health Network, Max Bell Research Center, Toronto, Canada, §Department for Experimental Physics V, University of Augsburg, Germany, and ¶Center of Structural Biology, Department of Chemistry, Vanderbilt University, Nashville, TN, U.S.A.

## Abstract

RDCs (residual dipolar couplings) in NMR spectroscopy provide information about protein dynamics complementary to NMR relaxation methods, especially in the previously inaccessible time window between the protein correlation time  $\tau_c$  and 50  $\mu$ s. For ubiquitin, new modes of motion of the protein backbone could be detected using RDC-based techniques. An ensemble of ubiquitin based on these RDC values is found to comprise all different conformations that ubiquitin adopts upon binding to different recognition proteins. These conformations in protein-protein complexes had been derived from 46 X-ray structures. Thus, for ubiquitin recognition by other proteins, conformational selection rather than induced fit seems to be the dominant mechanism.

## Protein dynamics studied using NMR spectroscopy

Knowledge about protein dynamics is crucial for understanding protein function. The dynamic processes observed in biomolecules cover a large time scale from femtoseconds to hours or even longer. On the femtosecond to picosecond time scale, very fast fluctuations occur that have been studied time-resolved, e.g. after temperature jumps with optical methods, electric methods, IR spectroscopy and fluorescence correlation spectroscopy [1]. When it comes to the experimental study of internal dynamics in a protein with atomic resolution (sub-Å resolution) and time resolution from femtosecond to hours, NMR spectroscopy is the tool of choice. With multidimensional NMR spectroscopy, most of the atoms in a molecule are accessible for NMR-detected observation and, except for replacing the naturally occurring carbon 12 (<sup>12</sup>C) by carbon 13 (<sup>13</sup>C) and the naturally occurring nitrogen 14 (<sup>14</sup>N) by nitrogen 15 (<sup>15</sup>N), no modification of the protein is necessary and it can be observed in a near-physiological environment. Also, by contrast with time-resolved measurements, kinetics can be measured in equilibrium. The equilibrium needs not to be disturbed.

Motion on time scales that are faster than the overall tumbling time correlation  $\tau_c$  of the whole molecule in solution (for ubiquitin 4 ns at room temperature) can be measured in amplitude and rate by NMR relaxation. The amplitude is measured as a Lipari-Szabo order parameter that is 1 for a rigid vector and 0 for a totally flexible vector

[2]. Rates of interconversion faster than the correlation time  $\tau_c$  (for ubiquitin 4 ns) are observed, i.e. motion on the sub- $\tau_c$  range [3]. Those motions are considered to mainly contribute to the entropy of proteins [4].

Motion between 50  $\mu$ s and 50 ms can be measured by the so-called relaxation dispersion [5]. The motion in this time window has been shown to be relevant for enzyme function [6].

That means, relaxation methods are only sensitive to protein motion in the sub- $\tau_c$  range and in the range between 50  $\mu$ s and 50 ms. Thus there is a window between the correlation time  $\tau_c$  and 50  $\mu$ s that has been unexplored (compare Figure 1). Functionally relevant dynamic processes in proteins might have remained undetected so far because of the inaccessibility of this time window. To our knowledge, the only report in which such motions were described with atomic detail is the photodissociation of carbon monoxide from myoglobin that has been studied by Moffat and co-workers [7] in the crystal with Laue diffraction. Here, conformational rearrangements in the small microsecond range were reported. However, such investigations have not been possible in solution so far.

## RDCs (residual dipolar couplings) as a tool to study protein dynamics

In the last few years, the group of Joel Tolman at Johns Hopkins University and ourselves have developed NMR methods to access this 'hidden' time window by using RDCs [8–11]. Dipolar couplings can be observed as a splitting of the resonance of a given spin that is dipolar coupled to another spin. Since this dipolar coupling depends on  $(3\cos^2\theta - 1)$  with respect to the external magnetic field, it is generally averaged

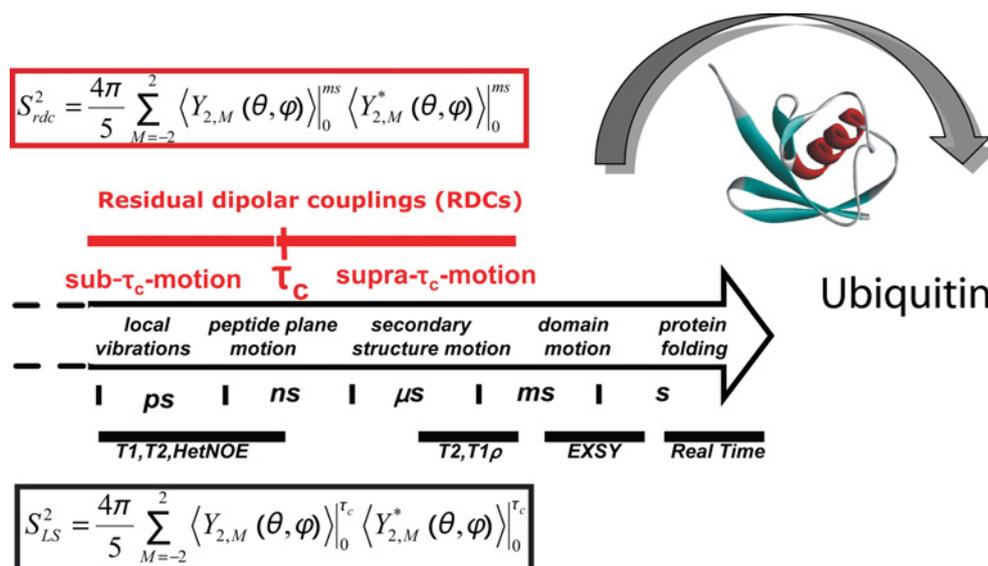
**Key words:** Ensemble Refinement with Orientational restraints (EROS), NMR, residual dipolar coupling (RDC), ubiquitin ensemble.

**Abbreviations used:** EROS, Ensemble Refinement with Orientational restraints; RDC, residual dipolar coupling; RMSD, root mean square deviation.

<sup>1</sup>To whom correspondence should be addressed (email [cigr@nmr.mpibpc.mpg.de](mailto:cigr@nmr.mpibpc.mpg.de)).

**Figure 1 | NMR measurements are suitable to detect motion on time scales from sub-picoseconds to hours**

Compared with other methods, RDCs cover a particular large time scale, including the supra- $\tau_c$  time window that was so far inaccessible for experimental observation by NMR. The ticks under the arrow should not be considered as upper and lower bounds for the respective motion.



out in solution NMR spectroscopy because of the isotropic tumbling of the protein in solution and the isotropic distribution of orientations of the molecule. However, proteins can be aligned in solution by weak liquid crystal media which lead still to isotropic tumbling of the molecule but induce an anisotropic distribution of orientations reintroducing a small amount of the dipolar coupling interaction, the ‘RDC’. The alignment happens in very non-invasive manner so that the resulting RDC is about 1000 times smaller than the original dipolar coupling interaction, providing orientation information without affecting the shape and dynamics of the protein. The anisotropy of the orientations is the so-called alignment of the molecule by the ‘alignment’ medium, e.g. the liquid crystal [12,13]. That RDCs are sensitive to protein dynamics has been recognized early on [14].

RDC  $D(\theta, \varphi)$  values report on the direction of inter-nuclear vector orientations, for example, the direction  $(\theta, \varphi)$  of the backbone amide group in a protein:

$$D(\theta, \varphi) = D_a[(3 \cos^2 \theta - 1) + 1.5 R \sin^2 \theta \cos 2\varphi] \quad (1)$$

Here,  $D_a$  is the so-called axial component of the alignment tensor and  $R$  is its rhombicity. An alignment tensor can be considered as being like an American football where the lengths of the axes represent the probability of the molecule to adopt that given orientation. The angles  $(\theta, \varphi)$  are the polar co-ordinates in the alignment frame. The alignment frame is fixed to the molecular frame of the molecule.

Now, coming back to dynamics: if there are internal dynamics of an inter-nuclear vector within the molecular frame, these will have an effect on the orientation of the inter-nuclear vector in the molecular frame, and therefore also in the alignment frame, and because then the angles  $(\theta, \varphi)$

vary this will also affect the value of the RDC. Thus the experimentally measured RDC values are dynamically averaged and can function as a probe for protein dynamics. Since RDCs are fixed to the molecular frame independently of the tumbling of the molecule, they pick up motions faster and slower than rotational tumbling correlation time  $\tau_c$  of the molecule up to the same limit that was active for chemical shift modulations used for NMR relaxation dispersion measurements. Thus RDCs are especially also sensitive to the hidden time window from the correlation time  $\tau_c$  up to 50  $\mu$ s (Figure 1).

In order to convert the measurement of RDCs into angular fluctuations, another step is needed. From the above formula, it is apparent that if we knew the alignment tensor and the structure of the molecule, i.e. the averaged vector orientations  $(\theta, \varphi)$ , we could directly infer the dynamics. However, those are unknown. They are only accessible in a model-free way (not assuming any motional model) by measuring the RDCs in at least five linear independent alignment media because each dipolar coupling depends on the five linearly independent spherical harmonics  $Y_{2,M}(\theta, \varphi)$  [8,9].

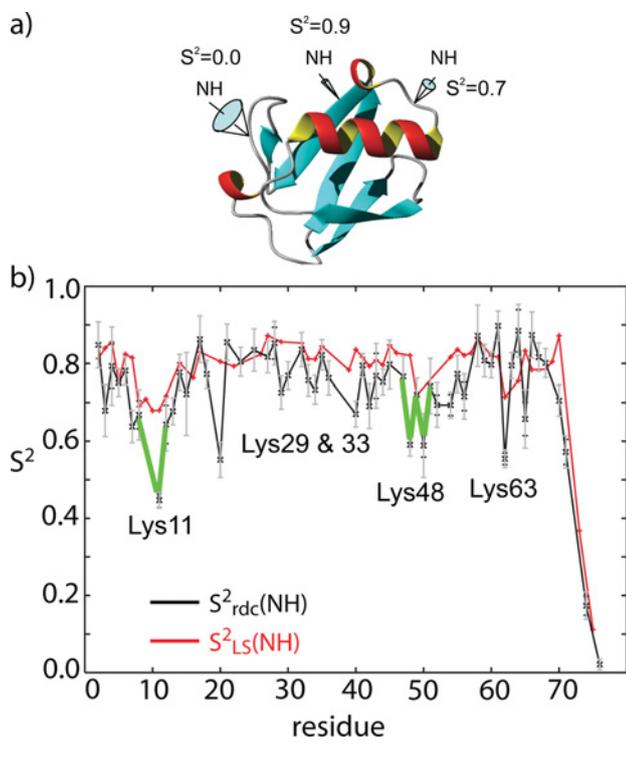
We have measured a very broad collection of different RDC datasets for the protein ubiquitin in order to enable a systematic study of the basics of protein dynamics [10,11]. Mathematical methods have been developed that deliver order parameters  $S_{rdc}^2$  which describe the amplitudes of motion up to the millisecond time scale:

$$S_{rdc}^2 = \frac{4\pi}{5} \sum_{M=-2}^2 \langle Y_{2,M}(\theta, \varphi) \rangle_0^{ms} \langle Y_{2,M}^*(\theta, \varphi) \rangle_0^{ms}$$

The  $S_{rdc}^2$  order parameters characterize the amplitude of motions up to the millisecond time scale for the studied individual inter-nuclear vector.

**Figure 2 | RDC-derived order parameters for ubiquitin**

(a) Order parameters can have values between 0 and 1. The smaller the order parameter, the more mobile the respective amino acid. (b) RDC-based  $S_{\text{rdc}}^2$  (NH) order parameters (black) are compared with the Lipari-Szabo  $S_{\text{LS}}^2$  (NH) (red) from relaxation for each amino acid in the ubiquitin sequence (residues 1–76). The  $S_{\text{rdc}}^2$  (NH) order parameters are sensitive to a much wider time scale than the  $S_{\text{LS}}^2$  (NH) order parameters and reveal additional modes of motion indicated by lower-order parameter values. Functionally important lysine residues appear highly mobile in the protein backbone. The loop region around Lys<sup>11</sup> and an alternating pattern of  $S_{\text{rdc}}^2$  (NH) in the fourth  $\beta$ -strand including Lys<sup>48</sup> are marked in green.



This order parameter is different from the sub- $\tau_c$  Lipari-Szabo order parameter  $S_{\text{LS}}^2$  from relaxation measurements that is defined as:

$$S_{\text{LS}}^2 = \frac{4\pi}{5} \sum_{M=-2}^2 \langle Y_{2,M}(\theta, \varphi) \rangle_0^{\tau_c} \langle Y_{2,M}^*(\theta, \varphi) \rangle_0^{\tau_c}$$

In principle, these order parameters can be measured for all inter-nuclear vectors. So far, for reasons of experimental simplicity, we have focused on the backbone amide vectors.

A comparison between RDC-based order parameters  $S_{\text{rdc}}^2$  (NH) and Lipari-Szabo order parameters  $S_{\text{LS}}^2$  (NH) from relaxation [2] is shown in Figure 2 [11]. The RDC-based order parameters  $S_{\text{rdc}}^2$  (NH) of several residues are substantially lower than the  $S_{\text{LS}}^2$  (NH) order parameters from relaxation. Since Lipari-Szabo  $S_{\text{LS}}^2$  (NH) order parameters are only sensitive for motion faster than the overall tumbling correlation time  $\tau_c$  in the order of a few nanoseconds, while the RDC-based order parameters  $S_{\text{rdc}}^2$  (NH) are sensitive up to

milliseconds, additional dynamics slower than  $\tau_c$  must exist. NMR relaxation dispersion has detected mobility only for very few residues in ubiquitin. In conclusion, the additionally observed dynamics occur in the previously ‘hidden’ time window between  $\tau_c$  and 50  $\mu\text{s}$  which we will call supra- $\tau_c$  time window in the following. Interestingly, several lysine residues, such as Lys<sup>11</sup> and Lys<sup>48</sup>, known to play a major role in the polyubiquitination process, show substantial dynamics in the supra- $\tau_c$  time scale. A possible correlation between backbone dynamics and functional importance of residues in recognition events has also been observed for protein G [15] and the aggregation of  $\alpha$ -synuclein [16].

Furthermore, a correlation between backbone mobility indicated by the  $S_{\text{rdc}}^2$  (NH) and side-chain orientation could be observed [11,17]. Those residues that have solvent-exposed side chains appear more mobile in the protein backbone. Those that have side chains pointing towards the hydrophobic core of the protein appear more rigid in the protein backbone. That finding has been independently observed by Blackledge and co-workers for protein G [15] and earlier by Palmer and co-workers for ribonuclease H in the sub- $\tau_c$  time scale [18].

### RDC-refined structural ensemble of ubiquitin shows conformational sampling

While an order parameter analysis gives a quantitative insight into the present amplitudes of mobility, only a structural representation of the dynamics, i.e. a structure ensemble, can give insight into the atomic details of the dynamics. Therefore, based on RDCs, a structural ensemble of ubiquitin dynamics in solution, the EROS (Ensemble Refinement with Oriental restraintS) ensemble, has been derived [19]. The EROS ensemble uses RDCs as additional restraints and, by that, covers dynamics up to milliseconds, especially the supra- $\tau_c$  time window (cf. Figure 3). The ensemble has an average backbone RMSD (root mean square deviation) value of 1.2  $\text{\AA}$  (1  $\text{\AA}$  = 0.1 nm). A more detailed description of that work is published in [19].

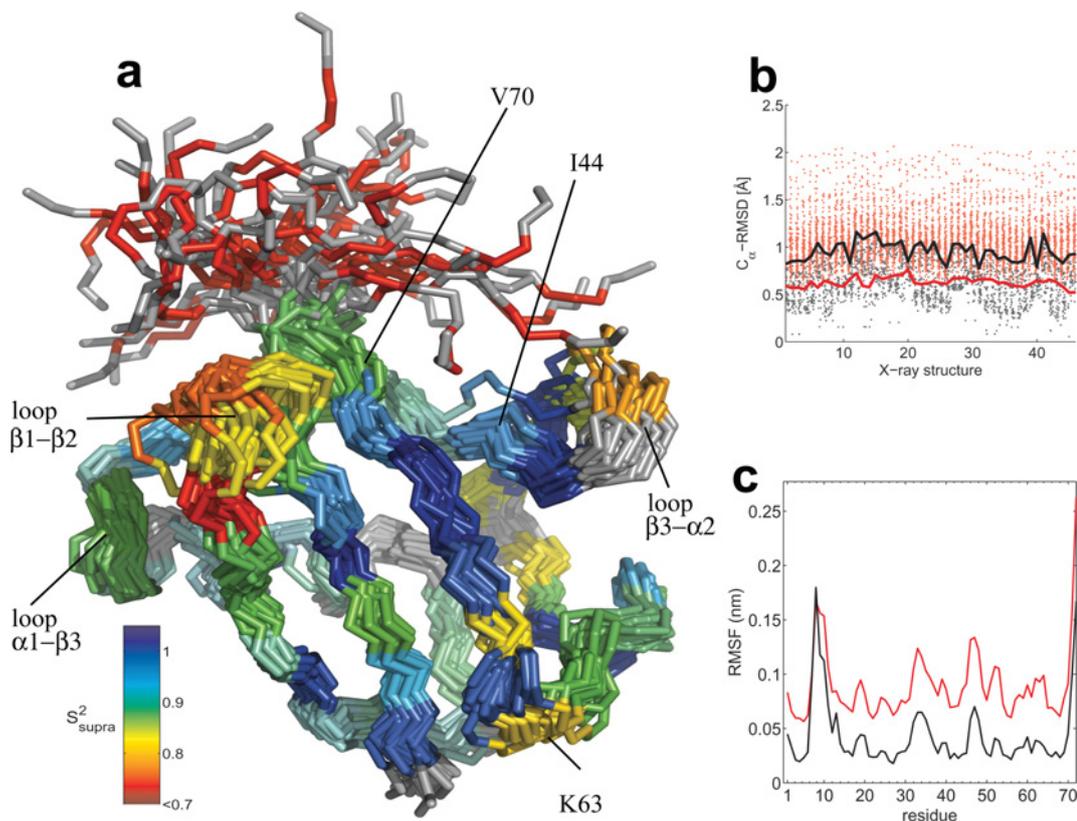
The ensemble also reveals new dynamics on the previously hidden time scale. As mentioned, the faster time scale is normally considered to be the entropic reservoir since order parameters change, for example, when a protein is engaged in a complex with another protein. The 50  $\mu\text{s}$  to 50 ms time scale has been found to be important for enzyme reactions that are often limited in reaction speed by intrinsic protein motions occurring in this time scale [6,20].

Ubiquitin is not an enzyme, but it is a protein whose function is to be recognized by other proteins and then be ligated. Thus the function of ubiquitin is to be recognized. So, does the supra- $\tau_c$  ensemble reflect recognition dynamics?

The EROS ensemble that has been determined using the RDCs mentioned above (and by that including the supra- $\tau_c$  motion) comprises all different X-ray structures of conformations that ubiquitin adopts when it binds to different recognition proteins. (The structures of those proteins are

**Figure 3 | Structure ensemble of ubiquitin**

(a) Backbone trace of 40 randomly chosen structures from the EROS ensemble. Residues are coloured by the amount of additional (supra- $\tau_c$ ) mobility as compared with the Lipari-Szabo order parameters. (b) For each X-ray structure, the backbone RMSD of residues 1–70 is shown for superpositions with each EROS structure (red) and each (other) X-ray structure (black). The minimal RMSD for EROS structures (red) and the maximal RMSD for X-ray structures (black) are highlighted as a line in the respective colour to guide the eye. (c)  $C_\alpha$  root mean square fluctuations (RMSF) of EROS structures (red) and of 46 known ubiquitin X-ray structures (black). Reprinted from [19] with permission from AAAS.



deposited in the PDB.) Indeed, each of the X-ray structures is similar to members of the solution ensemble within less than 0.8 Å RMSD backbone, despite the fact that no crystallographic data have been used during the EROS refinement. That means that the different conformations ubiquitin can adopt when it binds to different recognition proteins are already present in the equilibrium ground state in solution. This supports the possibilities that a recognition protein selects the fitting conformation from the ground state equilibrium when it binds ubiquitin and that no major structural changes need to be ‘induced’ in the binding process.

This finding has implications for the long-standing question of how proteins recognize each other. Since the 1960s, there are two major hypotheses discussed in the literature that try to explain the conformational changes of proteins upon complex formation [21]. The ‘induced-fit’ hypothesis proposes that the initial interaction between a protein and a binding partner ‘induces’ a conformational change of the protein through a step-wise process. In contrast, the conformational selection model favours a pre-existing equilibrium of protein conformations prior to molecular in-

teraction, and the binding event represents a ‘conformational selection’ of the most favourable interacting state with the structural ensemble. How the solution dynamics relate to contacts that ubiquitin forms when it engages in recognition complexes is described in [19]. Whereas in many biochemistry textbooks, the induced-fit mechanism is championed, there is growing support for the conformational selection mechanism. According to our results [19], for ubiquitin recognition, conformational selection rather than induced fit seems to be the dominant mechanism. Since conformational sampling on a time scale in the micro- to millisecond range has been found to be rate limiting in many catalyses [6,20,22,23] it would be intriguing if conformational sampling on the nano- to micro-second time scale, i.e. 1000 times faster than the time scale where most catalytic events occur, would prove to be essential for protein–protein recognition dynamics in general.

## References

- 1 Kolano, C., Helbing, J., Kozinski, M., Sander, W. and Hamm, P. (2006) Watching hydrogen-bond dynamics in a  $\beta$ -turn by transient two-dimensional infrared spectroscopy. *Nature* **444**, 469–472

- 2 Lipari, G. and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546–4559
- 3 Kay, L.E., Torchia, D.A. and Bax, A. (1989) Backbone dynamics of proteins as studied by  $^{15}\text{N}$  inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979
- 4 Lee, A.L. and Wand, A.J. (2001) Microscopic origins of entropy, heat capacity and the glass transition in proteins. *Nature* **411**, 501–504
- 5 Akke, M. and Palmer, A.G. (1996) Monitoring macromolecular motions on microsecond to millisecond timescales by  $R_{1\rho}$ - $R_1$  constant relaxation time NMR spectroscopy. *J. Am. Chem. Soc.* **118**, 911–912
- 6 Wang, L., Pang, Y., Holder, T., Brender, J.R., Kurochkin, A.V. and Zuiderweg, E.R. (2001) Functional dynamics in the active site of the ribonuclease binase. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7684–7689
- 7 Srajer, V., Teng, T., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M. and Moffat, K. (1996) Photolysis of the carbon monoxide complex of myoglobin: nanosecond time-resolved crystallography. *Science* **274**, 1726–1729
- 8 Meiler, J., Prompers, J.J., Peti, W., Griesinger, C. and Brüschweiler, R. (2001) Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. *J. Am. Chem. Soc.* **123**, 6098–6107
- 9 Tolman, J.R. (2002) A novel approach to the retrieval of structural and dynamic information from residual dipolar couplings using several oriented media in biomolecular NMR spectroscopy. *J. Am. Chem. Soc.* **124**, 12020–12030
- 10 Lakomek, N.A., Carlomagno, T., Becker, S., Griesinger, C. and Meiler, J. (2006) A thorough dynamic interpretation of residual dipolar couplings in ubiquitin. *J. Biomol. NMR* **34**, 101–115
- 11 Lakomek, N.A., Walter, K.F., Fares, C., Lange, O.F., de Groot, B.L., Grubmüller, H., Brüschweiler, R., Munk, A., Becker, S., Meiler, J. and Griesinger, C. (2008) Self-consistent residual dipolar coupling-based model-free analysis for the robust determination of nanosecond to microsecond protein dynamics. *J. Biomol. NMR* **41**, 139–155
- 12 Tolman, J.R., Flanagan, J.M., Kennedy, M.A. and Prestegard, J.H. (1995) Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9279–9283
- 13 Tjandra, N. and Bax, A. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278**, 1111–1113
- 14 Tolman, J.R., Flanagan, J.M., Kennedy, M.A. and Prestegard, J.H. (1997) NMR evidence for slow collective motions in cyanometmyoglobin. *Nat. Struct. Biol.* **4**, 292–297
- 15 Bouvignies, G., Bernado, P., Meier, S., Cho, K., Grzesiek, S., Brüschweiler, R. and Blackledge, M. (2005) Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13885–13890
- 16 Bertocini, C.W., Jung, Y.S., Fernandez, C.O., Hoyer, W., Griesinger, C., Jovin, T.M. and Zweckstetter, M. (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured  $\alpha$ -synuclein. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 1430–1435
- 17 Lakomek, N.A., Carlomagno, T., Becker, S., Meiler, J. and Griesinger, C. (2005) Side-chain orientation and hydrogen-bonding imprint supra- $\tau_c$  motion on the protein backbone of ubiquitin. *Angew. Chem. Int. Ed. Engl.* **44**, 7776–7778
- 18 Mandel, A.M., Akke, M. and Palmer, III, A.G. (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI: correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144–163
- 19 Lange, O.F., Lakomek, N.A., Fares, C., Schroder, G.F., Walter, K.F., Becker, S., Meiler, J., Grubmüller, H., Griesinger, C. and de Groot, B.L. (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* **320**, 1471–1475
- 20 Kern, D. and Zuiderweg, E.R.P. (2003) The role of dynamics in allosteric regulation. *Curr. Opin. Struct. Biol.* **13**, 748–757
- 21 Boehr, D.D. and Wright, P.E. (2008) Biochemistry. How do proteins interact? *Science* **320**, 1429–1430
- 22 Boehr, D.D., McElheny, D., Dyson, H.J. and Wright, P.E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638–1642
- 23 Henzler-Wildman, K.A., Lei, M., Thai, V., Kerns, S.J., Karplus, M. and Kern, D. (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913–916

---

Received 1 August 2008  
doi:10.1042/BST0361433