



Cite this: *Phys. Chem. Chem. Phys.*,  
2019, 21, 18850

# Quantifying proton NMR coherent linewidth in proteins under fast MAS conditions: a second moment approach†

Alexander A. Malär,<sup>a</sup> Susanne Smith-Penzel,<sup>a</sup> Gian-Marco Camenisch,<sup>a</sup>  
Thomas Wiegand,<sup>a</sup> Ago Samoson,<sup>bc</sup> Anja Böckmann,<sup>d</sup> Matthias Ernst<sup>id</sup>\*<sup>a</sup>  
and Beat H. Meier<sup>id</sup>\*<sup>a</sup>

Proton detected solid-state NMR under fast magic-angle-spinning (MAS) conditions is currently redefining the applications of solid-state NMR, in particular in structural biology. Understanding the contributions to the spectral linewidth is thereby of paramount importance. When disregarding the sample-dependent inhomogeneous contributions, the NMR proton linewidth is defined by homogeneous broadening, which has incoherent and coherent contributions. Understanding and disentangling these different contributions in multi-spin systems like proteins is still an open issue. The coherent contribution is mainly caused by the dipolar interaction under MAS and is determined by the molecular structure and the proton chemical shifts. Numerical simulation approaches based on numerically exact direct integration of the Liouville–von Neumann equation can give valuable information about the lineshape, but are limited to small spin systems (<12 spins). We present an alternative simulation method for the coherent contributions based on the rapid and partially analytic calculation of the second moments of large spin systems. We first validate the method on a simple system by predicting the <sup>19</sup>F linewidth in CaF<sub>2</sub> under MAS. We compare simulation results to experimental data for microcrystalline ubiquitin (deuterated 100% back-exchanged at 110 kHz and fully-protonated at 125 kHz). Our results quantitatively explain the observed linewidth per-residue basis for the vast majority of residues.

Received 17th June 2019,  
Accepted 6th August 2019

DOI: 10.1039/c9cp03414e

rsc.li/pccp

## Introduction

Due to technological and conceptual advancements over the past two decades, solid-state NMR spectroscopy has become an important technique in structural biology.<sup>1–6</sup> Most of these studies have used <sup>13</sup>C detection, but recent improvements in magic-angle spinning (MAS) lead to sufficiently narrow proton-resonance lines in order to obtain high spectral resolution of perdeuterated and full back-exchanged and also fully-protonated proteins.<sup>7–15</sup> Still, even at the fastest spinning frequencies available today, the dipolar interaction is not perfectly averaged out, and the resonance lines are, at least for small proteins, broader than in

solution state.<sup>16</sup> The limits of spectral resolution remain, therefore, an open question.<sup>17–22</sup>

It is well-known<sup>23,24</sup> that the total linewidth  $\Delta^{\text{tot}}$  of a peak, measured as full width at half maximum (FWHM) can be separated into different contributions. For protons in proteins and organic solids we distinguish: (i) the coherent contribution  $\Delta^{\text{coh}}$ , which is described by the spin-system Hamiltonian and arises largely due to homonuclear dipolar couplings as well as isotropic and anisotropic chemical shifts; (ii) the incoherent contribution  $\Delta^{\text{incoh}}$  due to stochastic processes (relaxation and chemical exchange); (iii) the inhomogeneous contribution  $\Delta^{\text{inhomo}}$  which arises from sample and magnetic field heterogeneities. The first two can be reduced by increasing the MAS frequency  $\nu_r$  and are usually collected under the term homogeneous linewidth  $\Delta^{\text{homo}}$ . The third one can only be influenced by sample preparation techniques, probe design or shim coil currents. In summary we can write

$$\Delta^{\text{tot}}(\nu_r) = \Delta^{\text{homo}}(\nu_r) + \Delta^{\text{inhomo}} = \Delta^{\text{coh}}(\nu_r) + \Delta^{\text{incoh}}(\nu_r) + \Delta^{\text{inhomo}} \quad (1)$$

In a perfectly homogeneous magnetic field and in the absence of dynamical effects and sample heterogeneity, the proton

<sup>a</sup> Physical Chemistry, ETH Zürich, Vladimir-Prelog-Weg 2, 8093 Zurich, Switzerland.  
E-mail: maer@ethz.ch, beme@ethz.ch

<sup>b</sup> School of Information Technologies, Tallinn University of Technology, Tallinn, Estonia. E-mail: ago.samoson@ttu.ee

<sup>c</sup> NMR Institute MTÜ, Tallinn, Estonia

<sup>d</sup> Institut de Biologie et Chimie des Protéines, Bases Moléculaires et Structurales des Systèmes Infectieux, Labex Ecofect, UMR 5086 CNRS, Université de Lyon, 7 passage du Vercors, 69367 Lyon, France. E-mail: a.boeckmann@ibcp.fr

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9cp03414e



linewidth under magic-angle spinning in proteins is given by the residual dipolar couplings and their interplay with the chemical-shift differences. In principle, the resulting linewidth can be calculated, for a known molecular or crystal structure, by numerical simulation using established computer programs.<sup>25–29</sup> For most numerically exact approaches, it is, however, difficult to handle more than about a dozen spins, which turns out to be insufficient to describe the experimental observations, in particular for fully protonated proteins. Larger spin systems can be approximated based on the concept of reducing the dimension of the Liouville state space by excluding unimportant and unpopulated states.<sup>30–34</sup> In fact, detailed studies showed that a significant amount of states are not essential in magnetic resonance simulations of large dipolar-coupled spin systems and can be dropped from the state space, increasing the feasible size for the simulation of spin system dynamics by an order of magnitude.<sup>30</sup> Such approaches may open up the possibility to approximately simulate also very large spin systems.<sup>35,36</sup>

In this work, we present an alternative approach, which computes analytically or semi-analytically the linewidth using a moment expansion. This approach goes back to the moment expansion formulas derived by van Vleck.<sup>37,38</sup> Such a moment expansion has later been used both in solution-state NMR<sup>39</sup> and in solid-state NMR. In solids, it has been used to analyze the  $\text{CaF}_2$  lineshape with a high precision.<sup>40</sup> It has been proposed that in combination with slow spinning it could be used to extract chemical-shift tensor information from resolved sideband structure<sup>41</sup> and it was used in more recent times for REDOR curve fitting<sup>42</sup> and as a systematic tool for NMR powder pattern analysis, in particular in the field of material sciences.<sup>43,44</sup> In a related approach “effective dipolar couplings” for given sites have been calculated instead of second moments, and it has been shown that the two measures are proportional.<sup>19</sup> Applications of the van Vleck approach for MAS have been discussed by Brunner *et al.*, Zorin *et al.* and Schnell and Spiess.<sup>19,45–48</sup>

Here we adapt the method of moments to proton linewidth calculations in proteins in the fast spinning limit where the isotropic chemical-shift differences are larger than the remaining dipolar linewidth for the majority of the (coupled) spins, leading to a well-resolved proton spectrum. We will use the van Vleck formula in combination with second-order Floquet theory to derive analytical expressions for second moments of a homonuclear dipolar-coupled three-spin system under MAS, which can afterwards be used to predict second moment and linewidth for each individual resonance.

We will validate the approach, by comparing to experimental results for  $\text{CaF}_2$  and apply the method to deuterated (and fully back-exchanged) and fully-protonated (UL) ubiquitin where the homogeneous contribution from residual dipolar couplings has been shown to dominate the linewidth.<sup>15</sup> The proton network is very different between the two protein samples. In the deuterated (and 100% back-exchanged) case it is significantly diluted, since only  $\text{H}_\text{N}$  protons and exchangeable sidechain protons are present, which has a strong impact on the coherent linewidth.

The paper is organized as follows: (i) a theoretical section (“Theoretical concepts (I)–(IV)”), where we introduce the second-moment approach under MAS for the purely-dipolar case, validate it on  $\text{CaF}_2$ , extend it to include the effects of chemical shifts and compare its site-specific predictions with analogous results, using standard Liouville–von Neumann-based methods on ubiquitin. (ii) An experimental application section (“Results and discussion (I) and (II)”), where we apply the method to quantify the coherent linewidth for ubiquitin and compare it to experimental results. Combining these findings with an estimation of incoherent contributions, obtained from experimental relaxation measurements, the goal of this section consists in gaining further quantitative insight into the different linewidth contributions for a model protein.

## Theoretical concepts

### (I) Analytical calculation of second moments for the pure dipole case

For proton-detected protein spectra, the experimental spectral line is often close to a Lorentzian or Gaussian and will, in the following, be approximated by a Gaussian line. In this case, the spectral lineshape of spin  $k$ ,  $S(\omega_k)$ , is fully characterized by the second moment  $M_2$ :<sup>38</sup>

$$S(\omega_k) = \frac{1}{\sqrt{2\pi M_2^k}} \exp\left(-\frac{\omega_k^2}{2M_2^k}\right), \quad (2)$$

The index  $k$  refers to the observed spin. The full width at half maximum in angular frequency units is

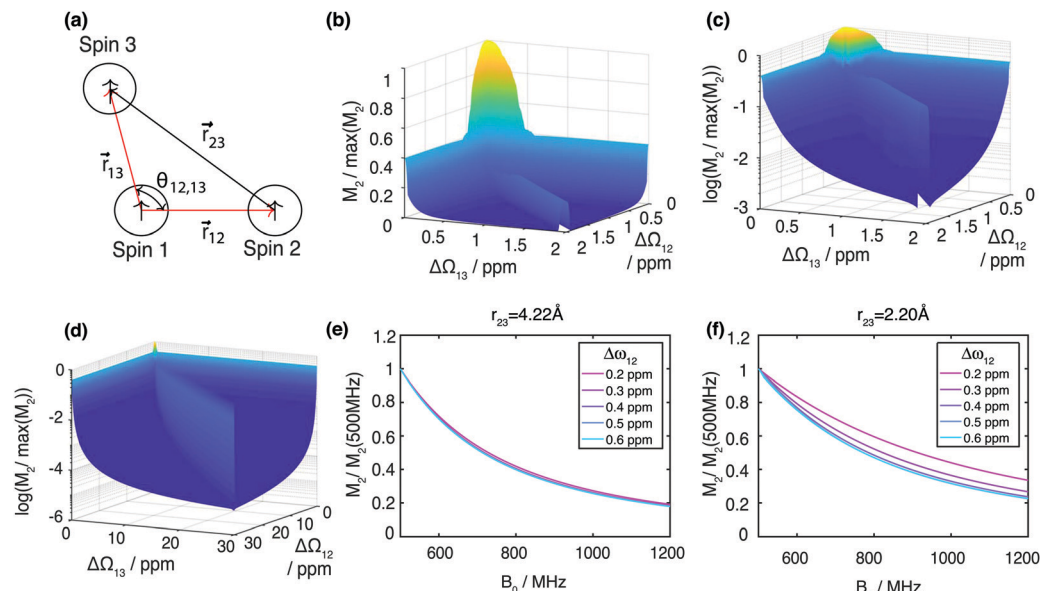
$$\text{FWHM} = 2\sqrt{2 \ln(2) M_2^k}, \quad (3)$$

We assume initially that the Hamiltonian is dominated by the homonuclear dipolar couplings (strong coupling regime) and calculate the total second moment of the system  $M_2$  for the static case, using the van Vleck formula:

$$M_2 = -\frac{\text{tr}\{[\mathcal{H}, F_x]^2\}}{\text{tr}\{F_x F_x\}} = \frac{9}{16N_1} \left(\frac{\mu_0}{4\pi}\right)^2 \gamma^4 \hbar^2 \sum_{\substack{k,j \\ j \neq k}} \frac{(1 - 3 \cos^2 \theta_{kj})^2}{r_{kj}^6}, \quad (4a)$$

where  $N_1$  denotes the number of spins in the system,  $r_{kj}$  is the length of the internuclear vector between spins  $(k,j)$ ,  $\theta_{kj}$  the inclination of the internuclear vector with respect to the external magnetic field. Due to the trace formation in eqn (4a), the operators can be evaluated in an arbitrary basis system and typically the Zeeman (uncoupled) rather than the eigenbasis of the Hamiltonian is used. The dipolar Hamiltonian itself is not diagonal in this basis and the eigenvalues of  $I_{nz}$  of the individual spins are in general not good quantum numbers. For the second-moment calculation we use  $F_x = \sum_{n=1}^N I_{nx}$  both as the initial density operator and detection operator.





**Fig. 1** (a) Sketch of the geometry of the three-spin system, as used for the calculation of the scaling factors  $\varepsilon = \frac{M_2^{1,23}}{\max(M_2^{1,23})}$  of the contributions to the resonance linewidth of spin 1. (b) Surface plot of the scaling factor  $\varepsilon$  calculated for  $(r_{12}, r_{13}, r_{23}) = (2.79 \text{ \AA}, 2.85 \text{ \AA}, 4.22 \text{ \AA})$ . For  $\Omega_1 = \Omega_2 = \Omega_3$  (coincident spectral lines),  $M_2$  assumes its maximum value and decays more strongly, the further  $\Omega_2, \Omega_3$  move away from  $\Omega_1$ . Surface plot for the same geometry but with logarithmic z axis for the regions (c)  $\Delta\Omega_{12}, \Delta\Omega_{13} = 0-2 \text{ ppm}$  (d)  $\Delta\Omega_{12}, \Delta\Omega_{13} = 0-30 \text{ ppm}$ . Magnetic field dependence from 500–1200 MHz of the second moment calculated with (e)  $r_{23} = 4.22 \text{ \AA}$ ,  $\Delta\Omega_{13} = 0.8 \text{ ppm}$  and (f)  $r_{23} = 2.20 \text{ \AA}$ ,  $\Delta\Omega_{13} = 0.8 \text{ ppm}$ .

If all  $N_1$  spins are equivalent, the double sum in eqn (4a) can be reduced to:<sup>49</sup>

$$M_2 = -\frac{\text{tr}\{[\mathcal{H}, F_x]^2\}}{\text{tr}\{F_x F_x\}} = \frac{9}{16} \left(\frac{\mu_0}{4\pi}\right)^2 \gamma^4 \hbar^2 \sum_j \frac{(1 - 3 \cos^2 \theta_{kj})^2}{r_{kj}^6}$$

$$= -\frac{\text{tr}\{[\mathcal{H}, F_x][\mathcal{H}, I_{kx}]\}}{\text{tr}\{F_x I_{kx}\}} \quad (4b)$$

where the choice of  $k$  does not influence the result for an infinitely extended system of equivalent nuclei. The detection operator is now given by  $I_{kx}$ .

Under sample spinning the Hamiltonian of eqn (4a) becomes time dependent and can be expanded in a Fourier series:

$$\mathcal{H}(t) = \sum_{n=-\infty}^{+\infty} \mathcal{H}^{(n)} e^{in\omega_r t}, \quad (5)$$

where  $\omega_r$  is the spinning frequency and the  $\mathcal{H}^{(n)}$  are the Fourier coefficients of the Hamiltonian. In order to calculate analytical expressions for  $M_2$ , we replace the time-dependent Hamiltonian by an equivalent time-independent second-order effective Hamiltonian  $\mathcal{H}_{\text{eff}}$ , which can be calculated based on average Hamiltonian theory<sup>50</sup> or Floquet theory:<sup>51–53</sup>

$$\mathcal{H}_{\text{eff}}^{(2)} = -\frac{1}{2} \sum_{n \neq 0} \frac{[\mathcal{H}^{(-n)}, \mathcal{H}^{(n)}]}{n\omega}. \quad (6)$$

The commutator vanishes for the dipolar-coupling Hamiltonian except for three-spin terms from commutators of dipolar couplings with one spin being common. We thus subdivide a

dipolar-coupled  $N$ -spin Hamiltonian into individual three-spin terms with a geometry defined as in Fig. 1(a).<sup>45,46,54</sup> One can then calculate the contribution of each three-spin system to the second moment and then sum over all possible three-spin terms present in the system.

The Fourier expansion of the dipolar Hamiltonian of a homonuclear three-spin system under magic-angle spinning is

$$\mathcal{H}_{\text{DD}}^{123}(t) = \sum_{k=1}^2 \sum_{q=k+1}^3 \mathcal{H}_{kq}(t)$$

$$= \sum_{k=1}^2 \sum_{q=k+1}^3 \sum_{n=-2}^2 \omega_{kq}^{(n)} e^{in\omega_r t} \cdot (2I_{kz}I_{qz} - (I_{kx}I_{qx} + I_{ky}I_{qy})) \quad (7)$$

For the calculation of the second-order effective Hamiltonian in such a three-spin system only the cross terms  $[\mathcal{H}_{12}, \mathcal{H}_{13}]$ ,  $[\mathcal{H}_{12}, \mathcal{H}_{23}]$ , and  $[\mathcal{H}_{13}, \mathcal{H}_{23}]$  between the non-commuting interactions need to be considered. The same considerations are valid for other terms  $\mathcal{H}_{\text{DD}}^{klm}(t)$ . The spatial components  $\omega_{kq}^{(n)}$  are calculated by sequential coordinate transformations from the principal-axis systems (PAS) of the dipolar couplings into the rotor-fixed frame (MAS) into the laboratory frame (LAB), using Wigner rotation matrices  $\mathcal{D}^{(2)}$ .

$$\omega_{kq}^{(n)} = \frac{1}{\sqrt{6}} \sum_{l=-2}^2 d_{n,0}^{(2)}(-\theta_m) \mathcal{D}_{l,n}^{(2)}(\alpha, \beta, \gamma) \mathcal{D}_{0,l}^{(2)}(0, \theta_{kq,kj}, 0) \sqrt{\frac{3}{2}} \delta_{kq}$$

$$= \frac{1}{\sqrt{6}} \sum_{l=-2}^2 d_{n,0}^{(2)}(-\theta_m) e^{-il\alpha} e^{-in\gamma} d_{l,n}^{(2)}(\beta) d_{0,l}^{(2)}(\theta_{kq,kj}) \sqrt{\frac{3}{2}} \delta_{kq}. \quad (8)$$



Note that computation of the cross terms only requires a pairwise treatment of the three dipolar couplings. This fact was exploited to simplify the theoretical treatment, introducing the intermediate Wigner rotation  $\mathcal{D}_{0,j}^{(2)}(0, \theta_{kq,kj}, 0)$ , which transforms the PAS value of one dipolar interaction, into the PAS system of the second interaction. Here  $\theta_{kq,kj}$  denotes the opening angle between the two dipolar couplings, which can be calculated by means of the scalar product between the inter-spin vectors  $\vec{r}_{kj}$  of the interactions

$$\theta_{kq,kj} = \arccos\left(\frac{\vec{r}_{kj} \cdot \vec{r}_{kq}}{|\vec{r}_{kj}| |\vec{r}_{kq}|}\right). \quad (9)$$

The analytical expressions  $\omega_{kq}^{(n)}$  can be combined with the expressions for the effective Hamiltonians in eqn (7) leading to the second-moment contribution to spin  $k$ ,  $M_2^{kqj}(\alpha, \beta)$  for the three-spin system  $(k, q, j)$ . Analytical integration over all possible crystallite orientations, leads to compact powder-averaged expressions for the second-moment contributions for the three-spin tuple  $(k, q, j)$  to the linewidth of spin  $k$ :

$$M_2^k = \sum_{\substack{q,j \\ q \neq j}} M_2^{kqj} \quad (10)$$

$$M_2^{kqj} = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi M_2^{k,qj}(\alpha, \beta) \sin(\beta) d\beta d\alpha.$$

We note that this second moment denotes the linewidth of the center band of the corresponding resonance and is dependent on the MAS frequency. It is well-known, that the second moment over all sidebands is invariant under magic-angle spinning if one takes into account the location and the intensity of the sidebands.<sup>41</sup> Note as well that we have considered only homonuclear dipole–dipole coupling terms, neglecting isotropic chemical shift, CSA and heteronuclear dipolar coupling contributions. The exact form of the analytical expressions for  $M_2^{kqj}$  and the formula for the summation over all three-spin systems are discussed in more detail in the next section.

## (II) Validation on calcium fluoride (CaF<sub>2</sub>)

Calcium fluoride was frequently used for the validation of moment-based linewidth simulation approaches and contains magnetically equivalent <sup>19</sup>F nuclei arranged in a simple cubic lattice. The second moment of static CaF<sub>2</sub> single crystals has been calculated in the original van Vleck paper and found to coincide well with experimental measurements.<sup>37,55</sup> We have measured a set of CaF<sub>2</sub> spectra for a powder sample at MAS frequencies ranging from 8 to 55 kHz and extracted their respective linewidths using a Gaussian fit (blue squares in Fig. 2). We now use the approach explained in the previous section to obtain the analytical powder-averaged expressions for the total second moment  $M_2^{kqj}$  of a three-spin system in frequency units squared (Hz<sup>2</sup>). Using a symbolic mathematical program (Mathematica, Wolfram, Champaign, USA), we obtain:

$$\begin{aligned} M_2^{kqj} = & \frac{1}{163840\nu_r^2} (61\delta_{13}^2\delta_{23}^2 + 14\delta_{12}\delta_{13}\delta_{23}(\delta_{13} + \delta_{23}) \\ & + \delta_{12}^2(61\delta_{13}^2 + 14\delta_{13}\delta_{23} + 61\delta_{23}^2) \\ & + \delta_{12}\delta_{13}(- (28\delta_{12}\delta_{13} + 14(\delta_{12} + \delta_{13})\delta_{23} - 47\delta_{23}^2) \cos(2\theta_{1213}) \\ & - 33\delta_{12}\delta_{13} \cos(4\theta_{1213})) - \delta_{23}(14\delta_{13}(-\delta_{12}^2 + 2\delta_{13}\delta_{23} + \delta_{12} \\ & (\delta_{13} + \delta_{23})) \cos(2\theta_{1213}) + \delta_{12}((14\delta_{12} - 47\delta_{13})\delta_{13} + 14(2\delta_{12} + \delta_{13})\delta_{23} \\ & + 33\delta_{13}^2 \cos(4\theta_{1213}))) \cos(2\theta_{1223}) - 33\delta_{23}^2(\delta_{12}^2 + \delta_{12}\delta_{13} \cos(2\theta_{1213}) \\ & + \delta_{13}^2 \cos(4\theta_{1213})) \cos(4\theta_{1223}) - 2\delta_{13}\delta_{23}(-40\delta_{12}^2 + 14\delta_{13}\delta_{23} \\ & + 7\delta_{12}(\delta_{13} + \delta_{23}) + 33\delta_{12}\delta_{13} \cos(2\theta_{1213})) \sin(2\theta_{1213}) \sin(2\theta_{1223}) \\ & - 33\delta_{13}\delta_{23}^2(\delta_{12} + 2\delta_{13} \cos(2\theta_{1213})) \sin(2\theta_{1213}) \sin(4\theta_{1223})). \end{aligned} \quad (11)$$

Here,  $\nu_r$  is the magic-angle-spinning frequency and  $\delta_{kj} = -2\frac{\mu_0\gamma_k\gamma_j\hbar}{4\pi r_{kj}^3} \frac{1}{2\pi}$  is the anisotropy of the dipolar coupling where the factor of  $2\pi$  comes from the conversion of angular frequency to frequency. Summing over all possible three-spin systems and dividing by the number of resonant spins  $N_i$  yields the second moment:

$$M_2 = \frac{1}{N_1} \sum_{k=1}^{N_1} \sum_{\substack{q=1 \\ q \neq k}}^{N_1-1} \sum_{\substack{j > q \\ q \neq k}}^{N_1} M_2^{kqj}. \quad (12)$$

As all fluorine atoms in CaF<sub>2</sub> are crystallographically equivalent it is again easier to calculate the  $M_2$  (similar to the difference between eqn (4a) and (4b)) by using as a starting density operator  $F_x$ , but  $I_{kx}$  as a detection operator, where  $k$  denotes the spin selected. The van-Vleck equation then reads

$$M_2^k = -\frac{\text{tr}\{\mathcal{H}, F_x\}[\mathcal{H}, I_{kx}]}{\text{tr}\{F_x I_{kx}\}} \quad (13)$$

and (again using Mathematica)

$$\begin{aligned} M_2^{kqj} = & -\frac{3}{655360\nu_r^2} (-122\delta_{12}^2\delta_{13}^2 - 28\delta_{12}\delta_{13}(\delta_{12} + \delta_{13})\delta_{23} \\ & - 61(\delta_{12}^2 + \delta_{13}^2)\delta_{23}^2 \\ & + 28\delta_{13} \cos(2\theta_{1213})(\delta_{12}(2\delta_{12}\delta_{13} + (\delta_{12} + \delta_{13})\delta_{23}) \\ & + \delta_{23}(-\delta_{12}^2 + \delta_{12}\delta_{13} + \delta_{13}\delta_{23}) \cos(2\theta_{1223})) \\ & + 33\delta_{13}^2 \cos(4\theta_{1213})(2\delta_{12}(\delta_{12} + \delta_{23} \cos(2\theta_{1223})) \\ & + \delta_{23}^2 \cos(4\theta_{1223})) \\ & + \delta_{23}(\delta_{12}(-94\delta_{13}^2 + 28\delta_{12}(\delta_{13} + \delta_{23})) \cos(2\theta_{1223}) \\ & + 33\delta_{12} \cos(4\theta_{1223})) \\ & + 4\delta_{13}(-40\delta_{12}^2 + 7\delta_{12}\delta_{13} + 7\delta_{13}\delta_{23}) \sin(2\theta_{1213}) \sin(2\theta_{1223}) \\ & + 66\delta_{13}^2(\delta_{12} + \delta_{23} \cos(2\theta_{1223})) \sin(4\theta_{1213}) \sin(2\theta_{1223})). \end{aligned} \quad (14)$$





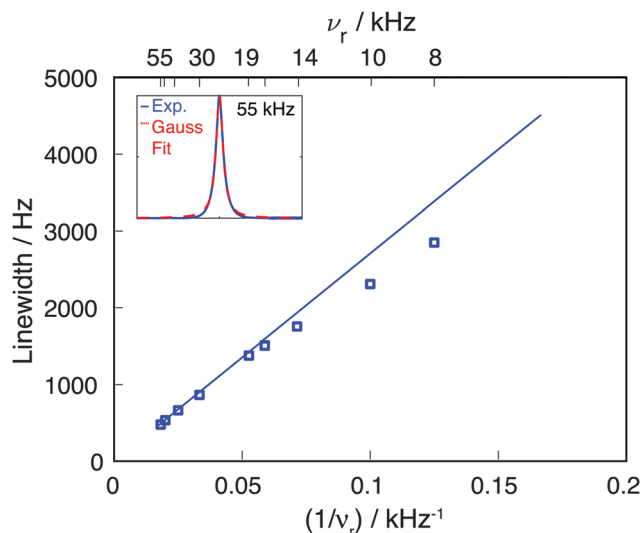


Fig. 2 Comparison of experimental  $^{19}\text{F}$   $\text{CaF}_2$  linewidth as function of the MAS frequency (blue squares) and the linewidth obtained by using an  $M_2(F_x I_{kx})$  approach (blue line). The inset shows the experimental line at 55 kHz MAS and the corresponding Gaussian fit (the fits for the remaining MAS frequencies are given in Fig. S1, ESI†).

The  $F_x I_{kx}$  case is much easier to evaluate since we now only need to sum over all three-spin systems containing the spin  $k$  of interest

$$M_2^k(F_x I_{kx}) = \frac{1}{3} \sum_{q \neq k} \sum_{\substack{j \neq k \\ j > q}} M_2^{k,qj}. \quad (15)$$

A more detailed discussion of the formula, including a motivation for the factor 3 appearing in the denominator is given in the ESI†, Section S1.

Calculating  $M_2^k(F_x I_{kx})$  for the central  $^{19}\text{F}$  spin in the  $\text{CaF}_2$  crystal containing  $12 \times 12 \times 12$  unit cells, leads to the blue curve in Fig. 2, which agrees well with the experimental data. The somewhat larger deviations between experimental and simulated  $\nu_r$  dependence of the linewidth at slow spinning frequencies arise from the fact that the Gaussian fit of the experimental  $^{19}\text{F}$  line is biased by an increased overlap between center-band and spinning sidebands (as it can be seen in Fig. S1, ESI†).

### (III) Inclusion of isotropic chemical-shift information in the $M_2$ simulations

Since the isotropic chemical-shift term is not modulated by the MAS rotation it can be simply added to the effective Hamiltonian, leading to  $\mathcal{H}_{\text{DD}+\text{CS}} = \mathcal{H}_{\text{eff}} + \sum_k \Omega_k I_{kz}$ . We neglect contributions to the second moment by the chemical-shift anisotropy (CSA) and heteronuclear dipolar-coupling terms, as well as by heteronuclear dipolar/CSA homonuclear dipolar cross terms. Indeed, heteronuclear couplings and CSA terms are inhomogeneous interactions<sup>41</sup> and can be neglected for fast spinning as they lead to sharp sidebands outside the window considered, whereas their cross terms with homonuclear dipolar couplings might have a significant contribution for large heteronuclear dipolar couplings/CSA (larger than the ones generally observed in the systems

considered here), and are harder to evaluate in practice, since information about CSA tensor orientation is required.

The signal  $S(\omega)$  under combined dipole and isotropic shift Hamiltonian can be calculated for a particular set of Euler angles  $(\alpha, \beta, \gamma)$  by numerically diagonalizing its matrix representation according to  $\tilde{H} = V^\dagger H_{\text{DD}+\text{CS}} V$ , where  $\tilde{H}$  denotes the diagonalized Hamiltonian. The intensities and the frequencies of the transitions of spin  $k$  in the eigenbase of the Hamiltonian are given by  $I_{il} = (\tilde{F}_{kx})_{il}(\tilde{F}_{kx})_{li}$  and  $\omega_{il} = \tilde{H}_{ii} - \tilde{H}_{ll}$ , respectively. Here,  $\tilde{F}_{kx}$  represents the  $F_{kx}$  operator in the eigenbase of the Hamiltonian.

The second moment  $M_2$  relative to the isotropic chemical shift of the spin of interest  $k$  is calculated from the (stick) spectrum as<sup>23</sup>

$$M_2^k(\omega, \alpha, \beta, \gamma) = \int (\omega - \Omega_k)^2 I_k(\omega, \alpha, \beta, \gamma) d\omega = \sum_{l,i < l} (\omega_{il} - \Omega_k)^2 I_{il}, \quad (16)$$

where  $\Omega_k$  is the resonance of interest, while the summation goes only over the dipolar-split transitions of spin  $k$  with frequency  $\omega_{il}$  and intensities  $I_{il}$ . Because the operators  $I_{kz}$  do not commute with the Hamiltonian, this approach involves a weak coupling approximation and becomes more exact at faster MAS (which reduces the effective dipolar coupling) and higher magnetic field (which increases the spectral separation). From now on, we will refer to a second moment that is calculated by taking into account only transitions with a frequency near the isotropic chemical shift of spin  $k$  as the local second moment of spin  $k$ . Here we include, for proteins, only transitions within  $\pm 50$  Hz (0.06 ppm for a proton Larmor frequency of 850 MHz) around the isotropic chemical shift of spin  $k$  to calculate  $M_2^k$ .

In the weak-coupling case (effective couplings smaller than the chemical-shift difference of the coupled spins), it is thus possible to sort out dipolar-split transitions belonging to spin  $k$ . With intermediate and strong couplings, this is no longer the case. The strong-coupling regime applies if the transitions of spin  $k$  overlap with transitions of another resonance, which is, in addition, strongly coupled to  $k$ . In such a case, the calculation of individual linewidths of different spins becomes meaningless. However, under fast MAS conditions ( $\nu_r \geq 100$  kHz), the weak coupling approximation requires that the dipolar-coupled spins have a chemical-shift difference exceeding 0.05–0.1 ppm. Fig. S2 in the ESI† shows that at 110 kHz MAS, the majority of spins in deuterated 100% back-exchanged ubiquitin lies in the weak-coupling regime and only 966 out of 150 348 three-spin systems (0.6%) fall into the intermediate coupling regime, where the chemical-shift inclusion may be biased. Note that the statistics is not expected to get worse for larger proteins because only spins in the immediate neighborhood can lead to strong coupling effects.

Calculating the local second moment and performing a numerical powder average over the Euler angles  $(\alpha, \beta, \gamma)$  gives a scaling factor  $\varepsilon = \frac{M_{2,\Omega}}{M_{2,\Omega=0}}$  for the reduction of a given three-spin second moment contribution due to the chemical-shift offset effect. The corresponding scaling of the second moment for a



three-proton system with the spatial geometry  $(r_{12}, r_{13}, r_{23}) = (2.79 \text{ \AA}, 2.85 \text{ \AA}, 4.23 \text{ \AA})$  as function of the chemical-shift differences  $\Delta\Omega_{12}, \Delta\Omega_{13}$  is shown in Fig. 1(b–d). We observe a significant reduction in the second moment of the line of spin one for non-vanishing chemical-shift differences compared to the case of no chemical-shift differences, which is more pronounced if the two passive spins two and three are also spectrally separated from each other. Such findings are in agreement with Liouville–von Neumann (LvN) simulations both for three- and for five-spin systems (see examples in Fig. S3, ESI†) and have also been noticed by Brunner *et al.*<sup>46</sup> Within the above geometry and setting  $\Delta\Omega_{13} = 0.3 \text{ ppm}$ , we obtain  $\varepsilon = 0.14$  (for  $\Delta\Omega_{12} = 0.01 \text{ ppm}$ ) and  $\varepsilon = 0.04$  (for  $\Delta\Omega_{12} = 1 \text{ ppm}$ ) for a Larmor frequency of 850 MHz. In the case where both  $\Delta\Omega_{12}$  and  $\Delta\Omega_{13}$  are very large (in the order of MHz) we find  $\varepsilon \approx 0$  (extreme weak coupling/heteronuclear case).

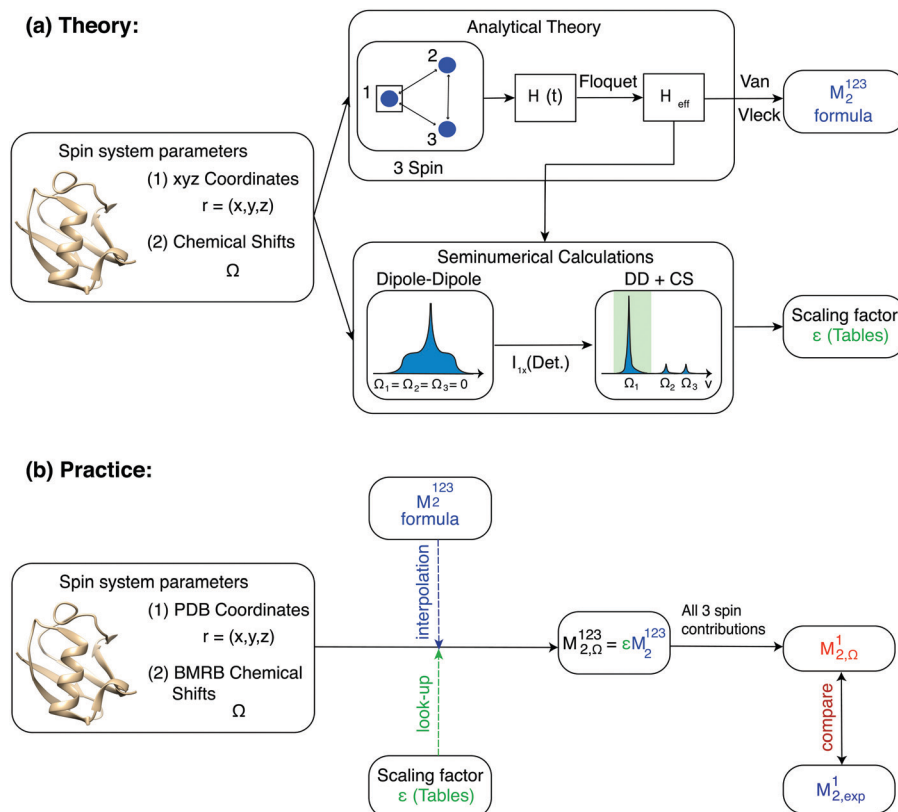
For increasing magnetic field strength, we expect a stronger truncation of the dipolar couplings. This is illustrated in Fig. 1(e), which shows the magnetic-field dependence of the

second moment, simulated for Larmor frequencies between 500 MHz to 1200 MHz, for the same geometry and  $\Delta\Omega_{13} = 0.8 \text{ ppm}$ . We observe the expected linewidth decrease with increasing field. Fig. 1(f) shows the same dependence for a stronger coupled geometry ( $r_{23} = 2.20 \text{ \AA}$ ). In particular we notice that the improvement with higher fields gets less steep the stronger the couplings are compared to chemical-shift differences (in particular refer to the  $\Delta\Omega_{13} = 0.2 \text{ ppm}$  curve).

For numerical efficiency we pre-calculated the scaling factors  $\varepsilon$  obtained by numerically diagonalizing the three-spin effective Hamiltonian as a function of the parameters  $(r_{12}, r_{13}, \theta_{12,13}, \Delta\Omega_{12}, \Delta\Omega_{13})$  (see Material and methods). A schematic overview of the second moment approach is given in Fig. 3.

#### (IV) Numerical results and comparison to brute-force Liouville–von Neumann simulations on ubiquitin

The  $M_2$  method is computationally much less demanding than exact LvN calculations as the calculation does not involve the diagonalization of many-spin matrices. For deuterated 100%



**Fig. 3** Schematic overview of the second moment method. (a) Analytical calculations are used to retrieve an effective Hamiltonian  $\mathcal{H}_{\text{eff}}$ , from which we obtain formulas for the second moment of a dipolar-coupled three-spin system under MAS rotation, using the van Vleck approach. To include the effect of chemical shifts on the second moment, such an effective Hamiltonian is as well used to compute semi-numerical spectra. Comparing the second moment obtained from the analytical reference with the one obtained from the numerical spectra with chemical shift inclusion gives chemical shift scaling factors  $\varepsilon$ , which are tabulated. Note that in the case of chemical shift inclusion, we calculate a “local” second moment, summing only over transitions with frequencies in a window of  $\pm 50 \text{ Hz}$  around the isotropic chemical shift of spin  $k$ . (b) In practice, we use the combination of structure coordinates and protein chemical shifts, to calculate  $M_{2,\Omega}^{123}$  for the spin of interest in a specific three-spin system in the protein. We perform a “look-up” to extract the scaling factor corresponding to the best-matching three-spin system parameters, while we use the analytical formula as interpolation function. Summation over all possible three spin system contributions containing spin 1, gives an  $M_{2,\Omega}^1$  for the spin of interest in the whole protein, which can be compared to second moments obtained from the experimental homogeneous linewidths.



back-exchanged ubiquitin, the simulated residue-specific linewidths obtained by summation over one protein molecule can be obtained in less than a minute on a laptop computer and even the simulations for the fully-protonated protein do not require more than ten minutes. The presented results have been obtained by including the 200–250 spins closest to the spin-of interest (usually lying within a sphere of 10 to 12 Å), within a calculation time of 15–20 minutes on a laptop. As a comparison, the numerical integration of the LvN equation for  $N = 12$  spins take approximately 17–18 hours, even being fully parallelized for the powder average loop (537 orientations) and distributed over 12 processors (Intel Xeon E5-2680v3 2.5–3.3 GHz, Matlab version 9.1). Furthermore, the  $M_2$  simulations only require the explicit construction of spin matrices for a three-spin system, while constructing the  $\varepsilon$  table, so no particular memory requirements arise. Convergence tests for  $M_2$  as a function of the number of spins are described in Fig. S3 and S4 (ESI†). In the deuterated case, we need to consider a sphere with  $r_s \approx 5$ –10 Å (6–10 spins), while in the fully-protonated case  $r_s$  needs to be around 5–6 Å, which corresponds already to 30–60 spins.

Fig. 4 shows a comparison between the linewidth obtained by site-specific second-moment analysis with and without chemical-shift terms, and LvN simulations for the  $H_N$  protons in deuterated 100% back-exchanged ubiquitin at 110 kHz MAS (Fig. 4(a)) and fully-protonated ubiquitin at 125 kHz MAS (Fig. 4(b)). For the deuterated molecule, the exchangeable side-chain protons were not included (*vide infra*). For the LvN approach  $N = 11$  spins were used in the simulations: inclusion of an additional spin broadens the linewidths by only a small amount (increasing  $N$  might however still have an effect on the lineshape).

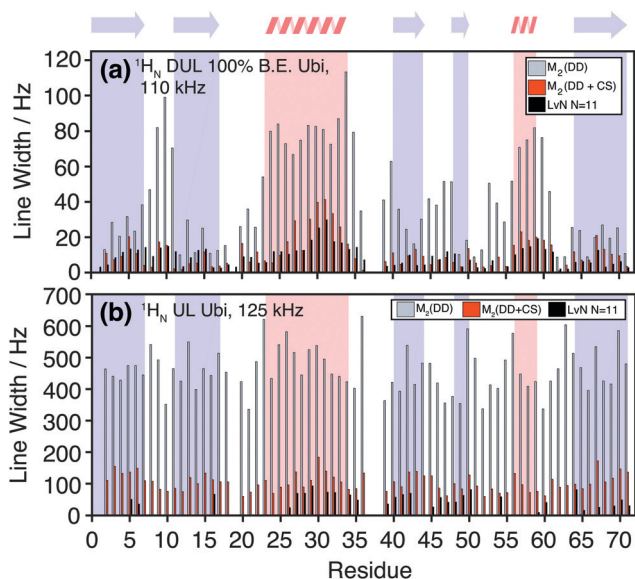


Fig. 4 Comparison of estimated coherent linewidths obtained by numerical integration of the LvN equation with  $N = 11$  (black) and second moment analysis with (red) and without chemical-shift inclusion (grey) in deuterated 100% back-exchanged ubiquitin at 110 kHz MAS (a) and UL ubiquitin at 125 kHz MAS (b). We performed LvN simulations only for residues where the chemical shifts for the closest  $N = 11$  are known. The blue background panels mark  $\beta$  sheets, while the red panels indicate  $\alpha$ -helix regions.

In the deuterated case, LvN and  $M_2 + CS$  simulations agree in general trends and magnitude over the entire protein with some deviations in one or the other direction. The pure dipolar  $M_2$  are, in all cases, significantly larger than the LvN and are clearly not suited for predictions in proteins. The linewidth is extracted from the LvN calculations by direct fitting of the simulated spectral line with a Gaussian model. When the numerically calculated line is broad, the corresponding peak becomes often asymmetric and a simple Gaussian or Lorentzian fit is not adequate (see examples in ESI†, Fig. S6). Situations in which the LvN linewidths are narrower than the  $M_2 + CS$  linewidths could hint that not enough spins have been included in the former approach. Fig. S8 in the ESI† shows a comparison of second moments obtained by  $M_2$ ,  $M_2 + CS$  simulation and by extracting it from the spectral lineshape of the numerical LvN spectra. Overall, we notice a good agreement between linewidths obtained by a Gaussian fit and by calculating the second moment from the simulated spectra. Even though,  $M_2$  and LvN-calculated second moments agree reasonably well, we note that the LvN lineshapes are often not Gaussian and asymmetric (Fig. 5 and Fig. S6, ESI†), in contrast to those obtained from an  $M_2$  approach, which are by definition Gaussian and symmetric as also typically observed in experiments (*vide infra*). The asymmetry seems to decrease and the linewidth to increase, the more spins are considered (Fig. S7, ESI†) and might vanish if

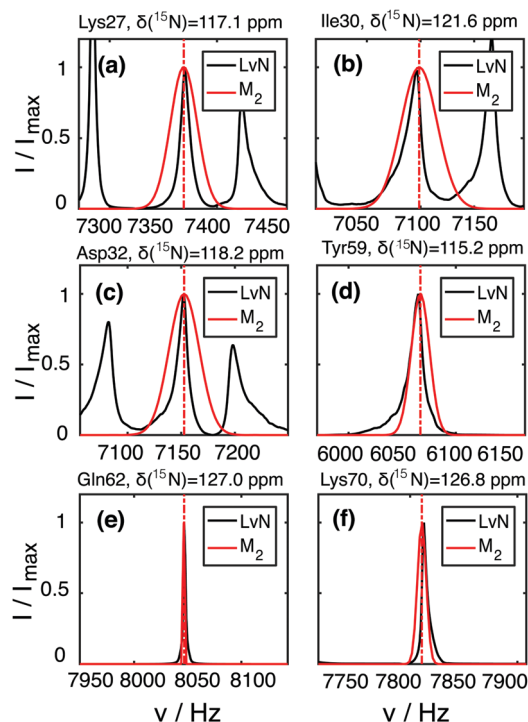


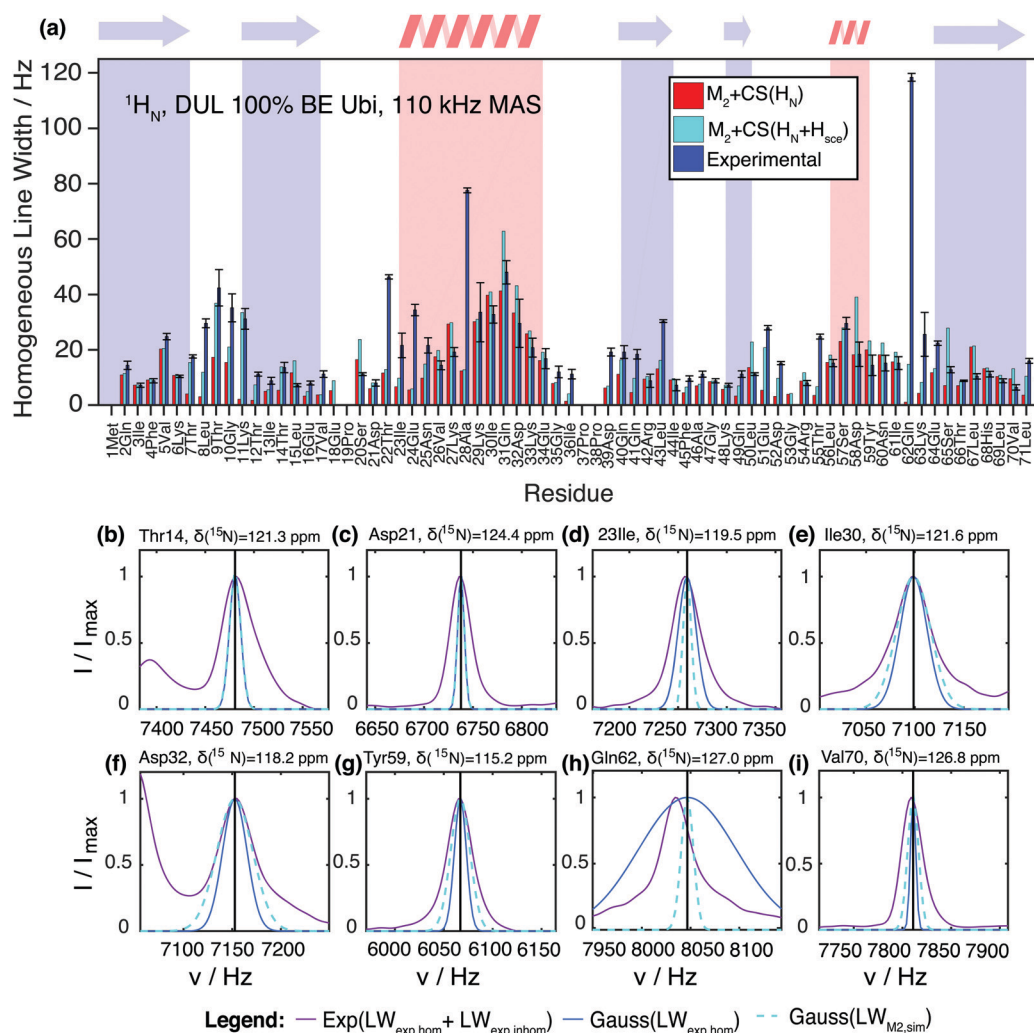
Fig. 5 Comparison of the  $M_2$  (red) and LvN-simulated (black) lineshapes for residues Lys27 (a), Ile30 (b), Asp32 (c), Tyr59 (d), Gln62 (e), Lys70 (f) in deuterated 100% back-exchanged ubiquitin at 110 kHz. The LvN lineshape is shown by directly displaying the numerically simulated spectrum. The  $M_2$  lineshape is obtained by plotting a Gaussian line defined by the calculated second moment and centered at the proton chemical shift known from the spectral assignment (marked by the vertical dashed red line).



many more spins could be included into the LvN calculation (though this cannot be verified due to the lack of computation power). It is, however, well known that the effect of spin diffusion, a many-body effect not easily described in smaller spin systems,<sup>56</sup> has a considerable effect on lineshapes even in the fast-spinning regime.<sup>57</sup> Furthermore, as shown in ref. 19 asymmetry can also arise as a consequence of the “termination” of small spin system clusters and lineshapes become symmetrical if the network is “closed” such that there are no edges.

In the fully-protonated case,  $M_2$  and LvN simulations do not agree so well with the LvN lines always being significantly narrower. Using the results from the  $M_2$  convergence tests

(*vide supra*), this behavior is most likely attributable to not considering enough spins to reach convergence in the calculation. Fig. 4 also shows the large effect that the inclusion of chemical shifts has on the predicted linewidths, both in the deuterated and in the UL case. In fact, as it is shown in Fig. 1(b), Fig. S2 and S3 (ESI<sup>†</sup>), as soon as spins are slightly detuned (0.05–0.1 ppm chemical-shift difference), most ubiquitin spins fall into a weak-coupling regime under fast MAS conditions, where chemical-shift scaling factors are smaller than 0.1. This is also illustrated in the scatter plots in Fig. S9 (ESI<sup>†</sup>), which shows how the pairs  $(M_{2,i})$  are distributed in the deuterated and fully-protonated case.



**Fig. 6** (a) Comparison of the experimentally-determined site-specific homogeneous linewidths (blue) in deuterated 100% back-exchanged ubiquitin at 110 kHz MAS and corresponding coherent linewidths obtained from the  $M_2$  simulations including chemical-shift effects (red). The cyan bars show the estimation of the coherent linewidth, if in addition to the amide protons also side-chain protons from exchangeable groups are taken into account. In general, these side-chain protons need to be considered unless they exchange with rate constants faster than the inverse of the dipolar coupling. In such a situation, “self-decoupling” of these side-chain protons would occur and these particular protons need not to be considered. (b–i) Comparison of the total (homogeneous + inhomogeneous) experimental (solid purple) lineshape and simulated coherent  $M_2$  (dashed cyan) linewidth represented by a Gaussian line for a selection of residues in deuterated 100% back-exchanged ubiquitin at 110 kHz (exchangeable side-chain atoms included in the analysis). Note that the shown experimental peaks are broader than the homogeneous linewidth given in Fig. 6(a), due to sample and magnetic-field inhomogeneities (a comparison between homogeneous and inhomogeneous linewidths is given in Fig. S12, ESI<sup>†</sup>). To give an impression of the size of the homogeneous contribution to the experimental linewidth, we plot as well a Gaussian with the FWHM as obtained from the site-specific  $T_2'$  measurements (solid blue), even though this of course does not correspond to the real experimental lineshape.





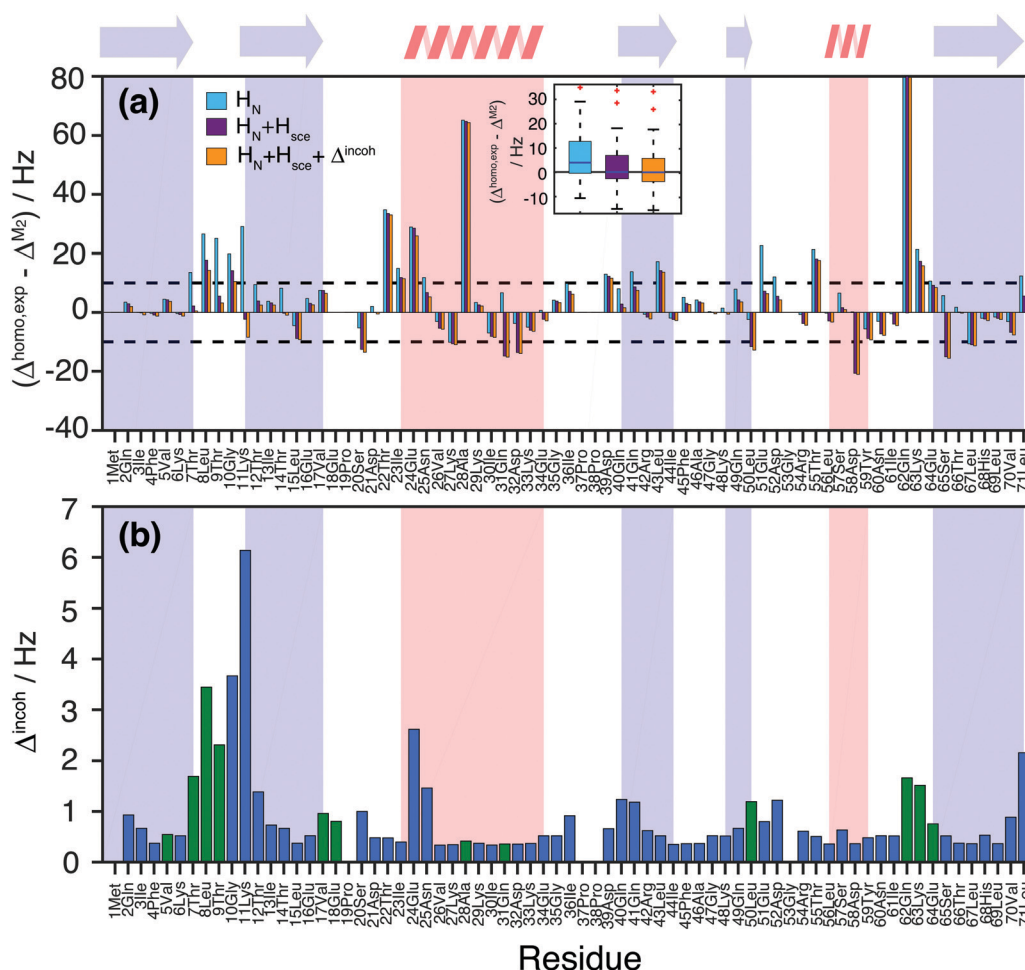
## Results and discussion

### (I) Deuterated 100% back-exchanged ubiquitin at 110 kHz MAS

Fig. 6 shows a comparison between the experimental site-specific homogeneous linewidths ( $T_2'$ ) of the  $H_N$  protons in deuterated 100% back-exchanged ubiquitin at 110 kHz MAS and the results from  $M_2$  simulations with chemical-shift inclusion. First of all, one may notice that the second moment method is able to predict in general qualitative and quantitative experimental trends surprisingly well without using any adjustable parameters. The simulations tend to give for most residues linewidths, which stay below the experimental ones, leaving room for additional line broadening effects (*vide infra* for further quantitative details). Coherent contributions are higher within  $\alpha$ -helices and loop regions, compared to  $\beta$ -sheets, as noticed earlier.<sup>11</sup> Convergence of the  $M_2$  method is reached by inclusion of all spins within a sphere of *ca.* 10 Å. In Fig. S10 (ESI<sup>†</sup>) we show the MAS dependence of the simulated coherent linewidth for a selection of representative residues. In agreement with an extrapolation from previously documented experimental findings for deuterated

100% back-exchanged ubiquitin, MAS frequencies of  $\sim 300$  kHz are required to reach amide proton linewidths comparable to those of a solution-state NMR sample (6–9 Hz).<sup>15</sup> These results are also in agreement with similar documented results obtained by simulations for selectively methyl-labeled proteins.<sup>21</sup> It should be noted that larger proteins have broader solution linewidth due to slower rotational tumbling.

Though the Gaussian assumption for the lineshape is a crude approximation, we can see from Fig. 6(b–i) that in many cases, it reproduces the experimental lineshape quite well. There are a few situations in which the  $M_2$  simulated linewidth slightly (*e.g.*, residues Leu15, Ser20, Val26, Ile30, Lys33, Asp58, Val70), or for residues Lys27 and Leu67 even significantly, exceeds the experimental linewidth. This could be due to imprecisions in the  $H_N$  coordinates of the PDB structure, which are introduced into the X-ray structure by molecular dynamic minimization (Fig. S13, ESI<sup>†</sup>). Another explanation could be fast dynamics that partially averages the dipolar interaction.



**Fig. 7** (a) Site-specific difference between experimental and  $M_2$  predicted linewidths for the deuterated 100% back-exchanged ubiquitin considering all  $H_N$  (light blue),  $H_N + H_{sca}$  (violet),  $H_N + H_{sca}$  and the incoherent contributions plotted in part (b) (orange). The black dashed horizontal lines mark the  $\pm 10$  Hz deviation limits, the inset contains boxplots illustrating the spread of the simulation vs. experiment deviations for all three simulated cases; (b) incoherent contributions  $\Delta^{incoh}$  estimated from experimental order parameters and correlation times, solving a stochastic LvN equation for a two-site hop model (see Materials and methods section). Blue bars have been obtained from the available literature data,<sup>58</sup> while green bars have been obtained from the previously missing experimental values that we measured in this work (Fig. S16 and S17, ESI<sup>†</sup>).



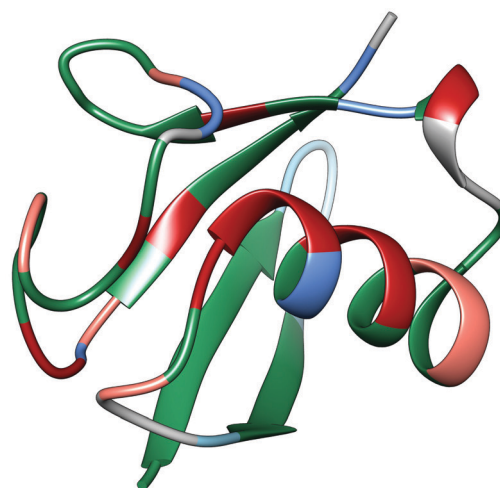
So far, we have neglected the exchangeable side-chain protons in the  $M_2$  calculations, which are in principle present in the experimental sample, due to exchange processes with the crystal water molecules, which are an essential part of the hydrated ubiquitin crystal. It is not entirely clear if such protons (namely those of hydroxyl, carboxyl, and amide groups) need to be considered for the calculation. In principle, one would have to include them in the  $M_2$  calculations; however, in the case of fast chemical exchange with water (fast compared to the dipole coupling), the exchange process decouples the interaction and the protons should not be included in the  $M_2$  calculations. In the intermediate exchange regime, more complex lineshape effects may appear. Therefore, two sets of residue-specific  $M_2$  values were calculated, one with and one without the exchangeable side-chain protons. Adding the Ser-OH, Thr-OH, Tyr-OH, His-NH, Asn-NH<sub>2</sub>, Gln-NH<sub>2</sub>, Arg-NH<sub>2</sub>, Lys-NH<sub>3</sub> side-chain protons to the  $M_2$  calculation and assuming slow exchange leads to the cyan bars in Fig. 6. For some residues, this leads to a significant contribution to the coherent linewidth. In particular, we notice that for Thr7(7), Thr9(9), Lys11(12), Asn25(25), Gln41(41), Asp52(27), Glu51(59), Glu64(65), Leu71(40) it could be a reasonable explanation for most of the previously missing linewidth. (The number in parenthesis indicates the residue number of the spatially closest  $H_{\text{sce}}$  for each  $H_N$ ; Section S8 in the ESI† gives a systematic overview over  $H_N$  and  $H_{\text{sce}}$  distance distributions.) In a few cases, namely Ser20(20), Gln31(41), Leu50(59), Asp58(55), Ser65(65), adding exchangeable side-chain protons leads to an overestimation of the linewidth for deuterated ubiquitin. The reasons could be fast exchange with solvent water or errors in side-chain proton coordinates, since these coordinates are difficult to predict.

Fig. 7(a) quantifies the remaining discrepancies between simulation and experiment. For 34 residues, the predicted coherent linewidth using  $M_2$  + CS simulations on all  $H_N$  protons differs by less than 10 Hz from the experimental value. In 15 cases, including  $H_N$  +  $H_{\text{sce}}$  contributions leads to deviations smaller than 10 Hz. In 6 cases the inclusion of  $H_{\text{sce}}$  makes the prediction worse than in the simple  $H_N$  case. Finally, 10 outliers remain, where prediction and measurement differ by more than 20 Hz.

We neglected until now incoherent contributions, as they are generally small in ubiquitin,<sup>58,59</sup> with the notable exception of few amino acids. In order to explicitly consider incoherent contributions to the linewidth, one can estimate them using order parameters and correlation times from  $^{15}\text{N}$   $R_{1\rho}(\nu_r)$  relaxation data. Part of the data was taken from the literature,<sup>58</sup> and completed with additional measurements of missing residues (see Materials and methods and Fig. S16 and S17 in ESI†). One can observe an enhanced relaxation contribution to the expected linewidth in the first loop region Thr7, Leu8, Thr9, Gly10, Lys11, Thr12, at the beginning of helix one, and in the last loop. Estimated incoherent contributions are on the order of 2–6 Hz (Fig. 7(b)), which can explain some of the missing linewidth for these residues.

It should be noted that additional dynamic processes can be present, which lead to a further modulation of the dipolar interaction and/or the isotropic chemical shift<sup>59</sup> and which are

not considered in the incoherent contribution of Fig. 7(b). In particular, it is known that residues Thr9, Ile23, Asn25, Lys33, Thr55, and Val70 undergo chemical exchange between two conformations.<sup>60,61</sup> This could account for the faster relaxation rates observed for residues Thr22, Ile23, Glu24 and Thr55. Gln62 is somewhat separated but at the other end of a loop contacting the exchanging region. It has the broadest experimental homogeneous linewidth but is one of the narrowest resonances in the simulation (Fig. 5(c) and 6). Still the broadening observed seems too large to be explained by the isotropic



Residues	$H_N$	$H_N + H_{\text{sce}}$
Gln2, Ile3, Phe4, Val5, Lys6, Ile13, Leu15, Glu16, Val17, Asp21, Val26, Lys27, Lys29, Ile30, Lys33, Glu34, Gly35, Ile36, Arg42, Ile44, Phe45, Ala46, Gly47, Lys48, Arg54, Leu56, Tyr59, Asn60, Ile61, Thr66, Leu67, His68, Leu69, Val70	✓	✓
Thr7, Thr9, Gly10, Lys11, Thr12, Thr14, Asn25, Gln40, Gln41, Gln49, Glu51, Asp52, Ser57, Glu64, Leu71	~	✓
Ser20, Gln31, Asp32, Leu50, Asp58, Ser65	✓	~
Leu8, Thr22, Ile23, Glu24, Ala28, Asp39, Leu43, Thr55, Gln62, Lys63	X	X

**Fig. 8** Classification of ubiquitin residues in groups and their location on the crystal structure. The classification is based on the bars in Fig. 7(a), residues have been classified as “explained”, when experimental and simulated values deviate by less than 10 Hz. Residues where we can explain the linewidth by coherent effects ( $H_N$ ) are marked in green (absolute value of light blue bars  $\leq 10$  Hz); residues where the inclusion of exchangeable side-chain protons explains the missing linewidth or substantially improves the deviance between experiment and simulation compared to the  $H_N$  case are marked in blue ( $H_N + H_{\text{sce}}$ , absolute value of violet bars  $\leq 10$  Hz, while for the light blue bars it is  $\geq 10$  Hz or the simulation improves compared to the  $H_N$  case by more than 3 Hz); residues, where including side-chain protons is detrimental are marked in light pink (absolute value of light blue bars  $\leq 10$  Hz and of the violet bars  $\geq 10$  Hz). Residues where we cannot explain the linewidth are marked in red (absolute value of all 3 bars in Fig. 7(a)  $\geq 10$  Hz).



chemical exchange effect. The additional contributions to the lineshape of residues Ala28, Asp39, Leu43, and Lys63 remain unexplained at this point in time.

Finally, Fig. 8 presents a summary of our current understanding of deuterated (100% back-exchanged) ubiquitin linewidths, based on coherent effects from  $H_N$ ,  $H_N + H_{sc}$  and shows in red the ten cases where additional effects need to be considered.

Finally, Fig. 9(a) shows how the simulated coherent ubiquitin  $H_N$  linewidths could still profit from a higher magnetic field (Larmor frequencies of 1200 MHz vs. 850 MHz) due to the larger spectral separation (see also Fig. 1(e and f)). Each site shows a different improvement (Fig. 9(b)), since the steepness of the curve encoding the ratio between the second moment at different fields depends on the parameters of the individual three-spin systems contributing to the linewidth (Fig. 1(e and f)). On average we observe a decrease of around 50% in the simulated homogeneous linewidths between these two fields.

## (II) Fully-protonated ubiquitin at 125 kHz MAS

Fig. 10(a) shows the comparison between the homogeneous experimental linewidth ( $T_2'$ ) and the simulated coherent

contributions to the linewidth for fully-protonated ubiquitin at 125 kHz MAS. Here the exchangeable side-chain protons are always included in the calculations. Still “self-decoupling” of some resonances through exchange may take place but such effects are expected to be minor due to the large number of non-exchangeable protons in fully-protonated molecules. In general, the  $M_2$  simulations are able to predict the order of magnitude and general behavior of the experimental linewidth fairly well. Simulated data for the HA resonances are provided in Fig. 10(b). The predicted linewidths are similar than for the amide protons.

Fig. 11 quantifies the differences between experimental and predicted coherent contribution for the fully-protonated case. The absolute deviations are larger than in the deuterated case, but the relative errors are mostly below 50%. The average deviation over the whole protein is around 38 Hz at an average linewidth of 99 Hz. It is important to note that in the fully-protonated case homonuclear  $^1H$ - $^1H$  dipolar couplings give a stronger contribution to not only the coherent but also the incoherent proton linewidth than for the perdeuterated case. Therefore, the interpretation in terms of dynamics is more

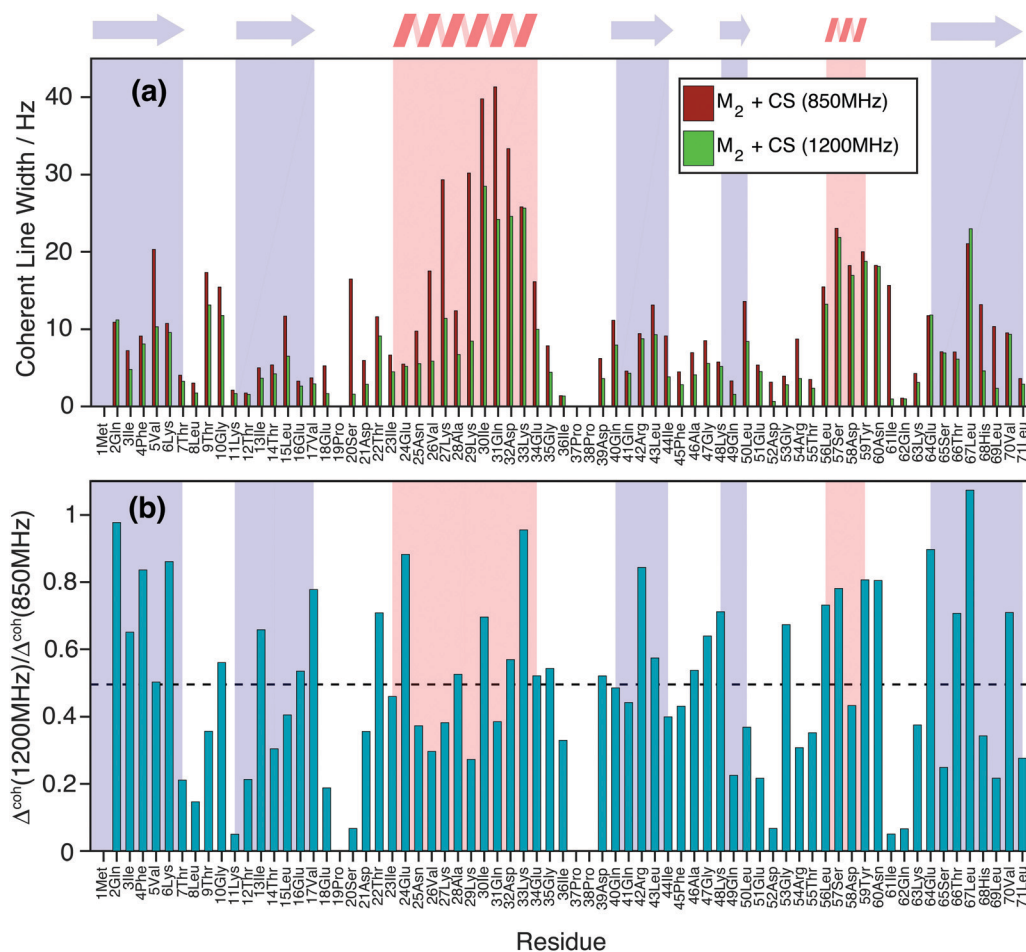


Fig. 9 (a) Comparison of simulated homogeneous linewidth at Larmor frequencies of 850 MHz (dark red) and 1200 MHz (green) for the  $H_N$  protons in DUL 100% back-exchanged ubiquitin. (b) Site-specific ratio between the simulated linewidths at the two fields. The black dotted line indicates the average gain over the molecule, which amounts to ca. 50%.



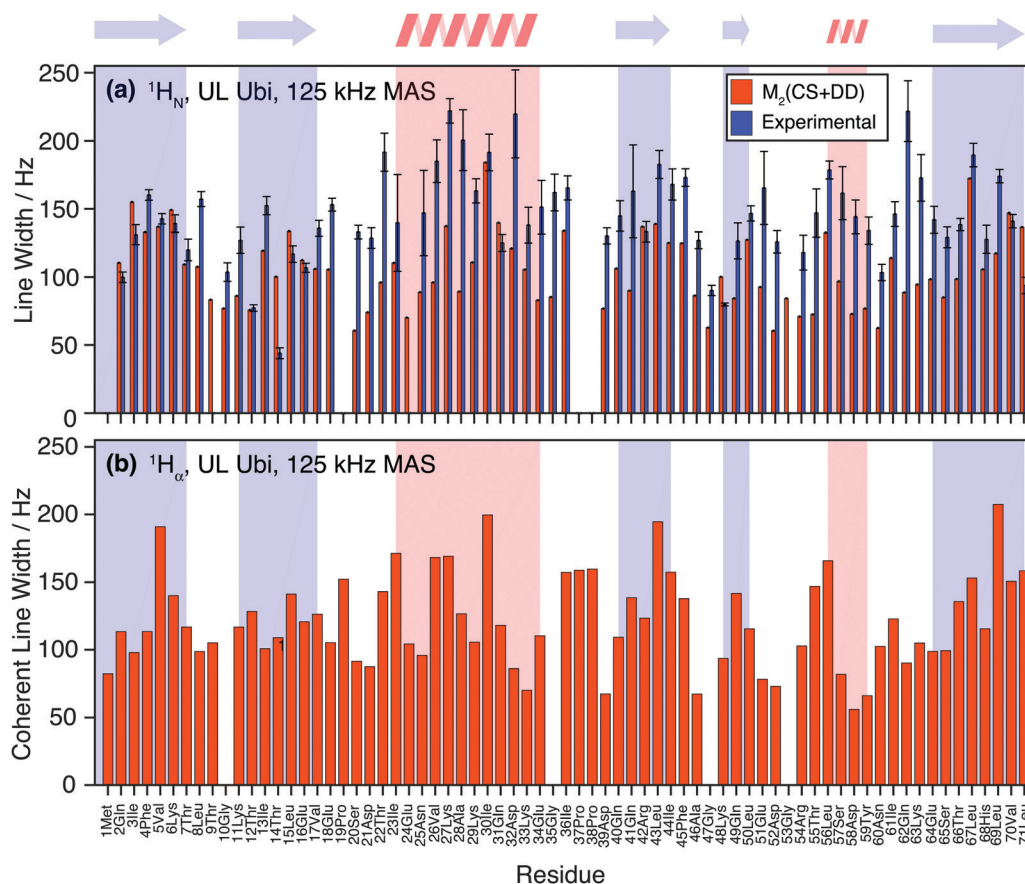


Fig. 10 (a) Comparison of the experimental site-specific homogeneous  $\text{H}_\text{N}$  linewidths (blue) in fully-protonated ubiquitin at 125 kHz MAS and corresponding coherent linewidths obtained from  $M_2$  simulations (red). (b)  $M_2$  simulated  $\text{H}_\alpha$  linewidth in fully-protonated ubiquitin.

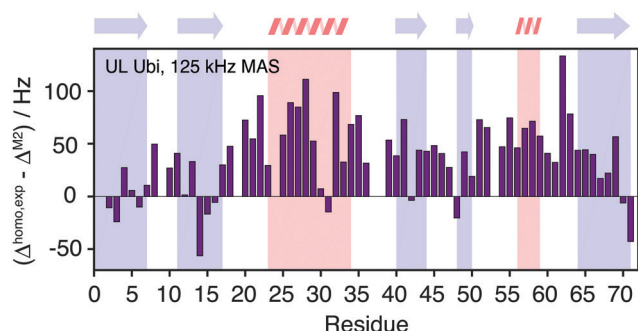


Fig. 11 Site-specific difference between experimental and calculated linewidths for fully protonated ubiquitin.

complicated.<sup>62</sup> Most simulated linewidths tend to lie below the experimental result with a few exceptions like Gln2, Lys6, Thr14, Leu15, Gln31, Lys48, Leu71. The discrepancies between simulated and experimental linewidth could be attributed to: (i) errors in the proton coordinates; (ii) lack of information on chemical shifts of side-chain protons. Since not all side-chain proton chemical shifts are known, the missing chemical shifts have been replaced by the average BMRB chemical shifts leading to degenerate chemical shifts; (iii) slow motions within the molecule can lead to additional line broadening; (iv) the incoherent contribution of  $^1\text{H}$ - $^1\text{H}$  interactions.

## Conclusions

In this work we presented an efficient simulation method for predicting coherent proton linewidths in proteins in MAS solid-state NMR under the assumption of a Gaussian lineshape. The method is based on semi-analytical second-moment calculations. The residual linewidth contribution evaluated here arises from second-order homonuclear dipolar interactions, which are significantly truncated by isotropic chemical-shift differences. The calculation of  $M_2$  is computationally much less expensive than standard simulation methods based on numerical integration of the LvN equation, and can easily be evaluated until convergence is reached, which makes it suitable for studying biological macromolecules, also in their fully protonated form. The simulated linewidths agree very well with the experimental linewidth in deuterated 100% back-exchanged and also rather well in fully protonated UL ubiquitin. The simulations indicate that the presence of back-exchanged side-chain protons is a significant source of coherent line-broadening in deuterated proteins. The broadening by incoherent relaxation effects is estimated as less than 6 Hz in ubiquitin (calculated based on  $T_{1\rho}$  data given in ref. 58) and significantly below this value for most residues (for details on the calculation refer to the Materials and methods section). For ten out of the 68 studied residues, the experimental lines are significantly broader than predicted, for two, Ala28 and





This journal is © the Owner Societies 2019

Note that the 3ONS structure has been determined by X-ray crystallography and does not contain precise information on proton positions. We inserted the proton positions in VMD<sup>70</sup> and post-optimized them using MD simulations with NAMD,<sup>71</sup> which slightly improves <sup>1</sup>H linewidth predictions (see ESI,<sup>†</sup> Section S7).

For simulation of a deuterated 100% back-exchanged sample we took into account only the other amide protons H<sub>N,q</sub> of the backbone or exchangeable sidechain protons when indicated, while for simulation of a protonated sample, we calculated the distances to all H<sub>q</sub>, including all aliphatic and aromatic protons.

The distances were then sorted in descending order to identify the  $N - 1$  spins closest to the spin of interest and then used to calculate  $\delta_{kq}$  for all  $\frac{N(N-1)}{2}$  spin pairs. The Euler angles of each dipolar interaction were extracted by reading the structure file into VMD and then used together with  $\delta_{kq}$  to calculate the spatial components  $\omega_{kq}^{(l)}$  based on eqn (9) and the full expression of the systems Hamiltonian  $\mathcal{H}(t)$  in the LAB frame (eqn (7)). Second moments  $M_2$  were calculated from the coupling constants  $\delta_{kq}$  and the opening angles  $\theta_{kq,kj}$ . The isotropic chemical shift  $\sigma_{iso,k}$  for each proton was obtained from the Biological Magnetic Resonance Data Bank (BMRB) file of ubiquitin (BMRB entry 17769<sup>72</sup>). In case of missing chemical shifts, we assigned the corresponding proton to its average chemical-shift value from the BMRB statistics. The three proton positions in CH<sub>3</sub> groups have been replaced by an averaged proton coordinate along the methyl-group principal axis, accounting for the fast CH<sub>3</sub> rotation.

For ubiquitin we ran LvN simulations with  $N = 10$ – $12$  spins programmed in MATLAB (The MathWorks 2016, version 9.1). Throughout this work, we used 537 orientations, calculated according to the method explained in ref. 73. Results have been validated by comparison with the output of an analogous implementation of the calculation in the C++ library GAMMA.<sup>25</sup>

For  $M_2$  calculations with chemical-shift information, we encoded the effect in chemical-shift scaling factors (CSSF), obtained by comparing the result for  $M_2$  calculations with and without chemical-shift differences between each resonance. We then generated CSSF tables for the five-dimensional parameter space spanned by  $(r_{12}, r_{13}, \theta_{12,13}, \Delta\Omega_{12}, \Delta\Omega_{13})$  in a general three-spin system and obtained tables for the spatial-spectral grid  $(\Delta\Omega_{12}, \Delta\Omega_{13}) = [0 : 0.025 : 0.02, 0.25 : 0.05 : 1, 1 : 0.2 : 10]$  ppm (in total  $71 \times 71$  points),  $(r_{12}, r_{13}) = [1 : 8 : 0.2 : 8.2, 9 : 16]$  Å; (in total  $45 \times 45$  points),  $\theta_{12,13} = 0.05 : \pi$  in 66 points. The bandwidth around the spin-of interest used for the tabulation process was 50 Hz. The actual calculation for  $M_2$  with chemical-shift inclusion computes the second moment for each three-spin system in the protein and corrects it by multiplication with the CSSF given by the tabulated value associated with the closest match to the input system parameters. This speeds up computations, compared to directly diagonalizing each three-spin system for every new calculation and makes application of  $M_2$  + CS simulations to new molecules easy and straightforward. All analytical computations for the  $M_2$  derivation have been performed with Mathematica (Wolfram Research, Inc., version 9.0,

Champaign, IL (2017)). All numerical simulation programs have been written in MATLAB (The MathWorks 2016, version 9.1). The most important software code for this work (calculation of  $M_2$  of deuterated 100% back-exchanged and UL ubiquitin as well as the calculation of chemical-shift scaling factors) have been commented and included in the ESI,<sup>†</sup> Section S10.

#### (IV) Estimation of incoherent contributions

Site-specific incoherent contributions have been estimated from the experimental order parameters and correlation times by numerically simulating the H<sub>N</sub> linewidth due to the N–H motion, interpreted in the context of a two-site hop model<sup>74</sup>

with opening angle  $\theta_i = \arccos\left(\sqrt{\frac{4S_i^2 - 1}{3}}\right)$  and correlation time  $\tau_i$ , solving a stochastic LvN equation according to the methods explained in ref. 75–77.

## Conflicts of interest

The authors declare no competing interests.

## Acknowledgements

We thank Riccardo Cadalbert for the ubiquitin sample preparation, as well as Andreas Hunkeler and Alexander Däpp for technical support. The brute-force LvN were performed on the Euler cluster of ETHZ. This research was supported by ETH Zürich and the Swiss National Science Foundation (SNF) (grant numbers 200020\_159707 and 20020\_178792). A. S. has been supported by Estonian research project PUT 1534. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 741863, FASTER).

## References

- 1 F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein and H. Oschkinat, *Nature*, 2002, **420**, 98–102.
- 2 C. Wasmer, A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek and B. H. Meier, *Science*, 2008, **319**, 1523–1526.
- 3 M. T. Colvin, R. Silvers, B. Frohm, Y. Su, S. Linse and R. G. Griffin, *J. Am. Chem. Soc.*, 2015, **137**, 7509–7518.
- 4 J. P. Demers, B. Habenstein, A. Loquet, S. Kumar Vasa, K. Giller, S. Becker, D. Baker, A. Lange and N. G. Sgourakis, *Nat. Commun.*, 2014, **5**, 4976.
- 5 M. D. Tuttle, G. Comellas, A. J. Nieuwkoop, D. J. Covell, D. A. Berthold, K. D. Kloepper, J. M. Courtney, J. K. Kim, A. M. Barclay, A. Kendall, W. Wan, G. Stubbs, C. D. Schwieters, V. M. Lee, J. M. George and C. M. Rienstra, *Nat. Struct. Mol. Biol.*, 2016, **23**, 409–415.
- 6 M. A. Walti, F. Ravotti, H. Arai, C. G. Glabe, J. S. Wall, A. Bockmann, P. Guntert, B. H. Meier and R. Riek, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E4976–E4984.
- 7 A. Böckmann, M. Ernst and B. H. Meier, *J. Magn. Reson.*, 2015, **253**, 71–79.



- 8 J. R. Lewandowski, J.-N. Dumez, Ü. Akbey, S. Lange, L. Emsley and H. Oschkinat, *J. Phys. Chem. Lett.*, 2011, **2**, 2205–2211.
- 9 M. J. Knight, A. L. Webber, A. J. Pell, P. Guerry, E. Barbet-Massin, I. Bertini, I. C. Felli, L. Gonnelli, R. Pierattelli, L. Emsley, A. Lesage, T. Herrmann and G. Pintacuda, *Angew. Chem., Int. Ed.*, 2011, **50**, 11697–11701.
- 10 E. Barbet-Massin, A. J. Pell, J. S. Retel, L. B. Andreas, K. Jaudzems, W. T. Franks, A. J. Nieuwkoop, M. Hiller, V. Higman, P. Guerry, A. Bertarello, M. J. Knight, M. Felletti, T. Le Marchand, S. Kotelovica, I. Akopjana, K. Tars, M. Stoppini, V. Bellotti, M. Bolognesi, S. Ricagno, J. J. Chou, R. G. Griffin, H. Oschkinat, A. Lesage, L. Emsley, T. Herrmann and G. Pintacuda, *J. Am. Chem. Soc.*, 2014, **136**, 12489–12497.
- 11 S. Penzel, A. A. Smith, V. Agarwal, A. Hunkeler, M. L. Org, A. Samoson, A. Bockmann, M. Ernst and B. H. Meier, *J. Biomol. NMR*, 2015, **63**, 165–186.
- 12 J. Stanek, L. B. Andreas, K. Jaudzems, D. Cala, D. Lalli, A. Bertarello, T. Schubeis, I. Akopjana, S. Kotelovica, K. Tars, A. Pica, S. Leone, D. Picone, Z. Q. Xu, N. E. Dixon, D. Martinez, M. Berbon, N. El Mammeri, A. Noubhani, S. Saupe, B. Habenstein, A. Loquet and G. Pintacuda, *Angew. Chem., Int. Ed.*, 2016, **55**, 15504–15509.
- 13 L. B. Andreas, K. Jaudzems, J. Stanek, D. Lalli, A. Bertarello, T. Le Marchand, D. Cala-De Paepe, S. Kotelovica, I. Akopjana, B. Knott, S. Wegner, F. Engelke, A. Lesage, L. Emsley, K. Tars, T. Herrmann and G. Pintacuda, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 9187–9192.
- 14 G. David, M. L. Fogeron, M. Schledorn, R. Montserret, U. Haselmann, S. Penzel, A. Badillo, L. Lecoq, P. Andre, M. Nassal, R. Bartenschlager, B. H. Meier and A. Bockmann, *Angew. Chem., Int. Ed.*, 2018, **57**, 4787–4791.
- 15 S. Penzel, A. Oss, M.-L. Org, A. Samoson, A. Böckmann, M. Ernst and B. H. Meier, *J. Biomol. NMR*, 2019, **73**, 19–29.
- 16 V. Agarwal, S. Penzel, K. Szekely, R. Cadalbert, E. Testori, A. Oss, J. Past, A. Samoson, M. Ernst, A. Böckmann and B. H. Meier, *Angew. Chem., Int. Ed.*, 2014, **53**, 12253–12256.
- 17 S. Ray, E. Vinogradov, G. J. Boender and S. Vega, *J. Magn. Reson.*, 1998, **135**, 418–426.
- 18 C. Filip, S. Hafner, I. Schnell, D. E. Demco and H. W. Spiess, *J. Chem. Phys.*, 1999, **110**, 423–440.
- 19 V. E. Zorin, S. P. Brown and P. Hodgkinson, *J. Chem. Phys.*, 2006, **125**, 144508.
- 20 K. Xue, R. Sarkar, C. Motz, S. Asami, D. C. R. Camargo, V. Decker, S. Wegner, Z. Tosner and B. Reif, *Sci. Rep.*, 2017, **7**, 7444.
- 21 K. Xue, R. Sarkar, C. Motz, S. Asami, V. Decker, S. Wegner, Z. Tosner and B. Reif, *J. Phys. Chem. C*, 2018, **122**, 16437–16442.
- 22 A. A. Malär, S. Dong, G. Kehr, G. Erker, B. Meier and T. Wiegand, *ChemPhysChem*, 2019, **20**, 672–679.
- 23 A. Abragam, *The Principles of Nuclear Magnetism*, Oxford University Press, Oxford, 1961.
- 24 J. Cavanagh, W. J. Fairbrother, A. G. Palmer, M. Rance and N. J. Skelton, in *Protein NMR Spectroscopy*, ed. J. Cavanagh, W. J. Fairbrother, A. G. Palmer, M. Rance and N. J. Skelton, Academic Press, Burlington, 2nd edn, 2007, pp. 1–28.
- 25 S. A. Smith, T. O. Levante, B. H. Meier and R. R. Ernst, *J. Magn. Reson., Ser. A*, 1994, **106**, 75–105.
- 26 M. Bak, J. T. Rasmussen and N. C. Nielsen, *J. Magn. Reson.*, 2000, **147**, 296–330.
- 27 P. Hodgkinson and L. Emsley, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2000, **36**, 201–239.
- 28 H. J. Hogben, M. Krzystyniak, G. T. Charnock, P. J. Hore and I. Kuprov, *J. Magn. Reson.*, 2011, **208**, 179–194.
- 29 Z. Tosner, R. Andersen, B. Stevansson, M. Eden, N. C. Nielsen and T. Vosegaard, *J. Magn. Reson.*, 2014, **246**, 79–93.
- 30 I. Kuprov, N. Wagner-Rundell and P. J. Hore, *J. Magn. Reson.*, 2007, **189**, 241–250.
- 31 I. Kuprov, *J. Magn. Reson.*, 2008, **195**, 45–51.
- 32 M. C. Butler, J.-N. Dumez and L. Emsley, *Chem. Phys. Lett.*, 2009, **477**, 377–381.
- 33 J. N. Dumez, M. C. Butler and L. Emsley, *J. Chem. Phys.*, 2010, **133**, 224501.
- 34 H. J. Hogben, P. J. Hore and I. Kuprov, *J. Chem. Phys.*, 2010, **132**, 174101.
- 35 I. Kuprov, *J. Magn. Reson.*, 2016, **270**, 124–135.
- 36 U. Sternberg, R. Witter, I. Kuprov, J. M. Lamley, A. Oss, J. R. Lewandowski and A. Samoson, *J. Magn. Reson.*, 2018, **291**, 32–39.
- 37 J. H. Van Vleck, *Phys. Rev.*, 1948, **74**, 1168–1183.
- 38 M. Mehring, *Principles of high-resolution NMR in solids*, Springer-Verlag, Berlin, 2nd, rev. and enl. edn, 1983.
- 39 H. Dolhaine, M. Engelhardt, G. Hägele and W. Küchelhaus, *J. Magn. Reson.*, 1980, **41**, 1–16.
- 40 G. W. Parker and F. Lado, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1974, **9**, 22–28.
- 41 M. M. Maricq and J. S. Waugh, *J. Chem. Phys.*, 1979, **70**, 3300–3316.
- 42 M. Bertmer and H. Eckert, *Solid State Nucl. Magn. Reson.*, 1999, **15**, 139–152.
- 43 R. Goc, *Solid State Nucl. Magn. Reson.*, 1998, **13**, 55–61.
- 44 B. Herreros, A. W. Metz and G. S. Harbison, *Solid State Nucl. Magn. Reson.*, 2000, **16**, 141–150.
- 45 E. Brunner, *J. Chem. Soc., Faraday Trans.*, 1990, **86**, 3957–3960.
- 46 E. Brunner, D. Freude, B. C. Gerstein and H. Pfeifer, *J. Magn. Reson.*, 1990, **90**, 90–99.
- 47 E. Brunner, *J. Chem. Soc., Faraday Trans.*, 1993, **89**, 165–169.
- 48 I. Schnell and H. W. Spiess, *J. Magn. Reson.*, 2001, **151**, 153–227.
- 49 G. A. Webb, *Org. Magn. Reson.*, 1979, **12**, IV.
- 50 U. Haeberlen and J. S. Waugh, *Phys. Rev.*, 1968, **175**, 453–467.
- 51 M. Leskes, P. K. Madhu and S. Vega, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **57**, 345–380.
- 52 I. Scholz, J. D. van Beek and M. Ernst, *Solid State Nucl. Magn. Reson.*, 2010, **37**, 39–59.
- 53 G. Floquet, *Ann. Sci. Ecole Norm. Sup.*, 1883, **12**, 47–88.
- 54 R. Challoner and C. A. McDowell, *J. Magn. Reson.*, 1992, **98**, 123–133.
- 55 G. E. Pake and E. M. Purcell, *Phys. Rev.*, 1948, **74**, 1184–1188.
- 56 N. Bloembergen, *Physica*, 1949, **15**, 386–426.



- 57 M. Ernst, A. Verhoeven and B. H. Meier, *J. Magn. Reson.*, 1998, **130**, 176–185.
- 58 N. A. Lakomek, S. Penzel, A. Lends, R. Cadalbert, M. Ernst and B. H. Meier, *Chemistry*, 2017, **23**, 9425–9433.
- 59 D. Marion, D. F. Gauto, I. Ayala, K. Giandoreggio-Barranco and P. Schanda, *ChemPhysChem*, 2019, **20**, 276–284.
- 60 F. Massi, M. J. Grey and A. G. Palmer III, *Protein Sci.*, 2005, **14**, 735–742.
- 61 N. Salvi, S. Ulzega, F. Ferrage and G. Bodenhausen, *J. Am. Chem. Soc.*, 2012, **134**, 2481–2484.
- 62 P. Rovó and R. Linser, *J. Phys. Chem. B*, 2017, **121**, 6117–6130.
- 63 J. Sass, F. Cordier, A. Hoffmann, M. Rogowski, A. Cousin, J. G. Omichinski, H. Löwen and S. Grzesiek, *J. Am. Chem. Soc.*, 1999, **121**, 2047–2055.
- 64 T. I. Igumenova, A. E. McDermott, K. W. Zilm, R. W. Martin, E. K. Paulson and A. J. Wand, *J. Am. Chem. Soc.*, 2004, **126**, 6720–6727.
- 65 M. Schubert, T. Manolikas, M. Rogowski and B. H. Meier, *J. Biomol. NMR*, 2006, **35**, 167–173.
- 66 C. Jaeger and F. Hemmann, *Solid State Nucl. Magn. Reson.*, 2014, **57–58**, 22–28.
- 67 S. Penzel, A. A. Smith, M. Ernst and B. H. Meier, *J. Magn. Reson.*, 2018, **293**, 115–122.
- 68 A. A. Smith, *J. Biomol. NMR*, 2017, **67**, 77–94.
- 69 K. Momma and F. Izumi, *J. Appl. Crystallogr.*, 2011, **44**, 1272–1276.
- 70 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
- 71 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
- 72 K. Y. Huang, G. A. Amodeo, L. Tong and A. McDermott, *Protein Sci.*, 2011, **20**, 630–639.
- 73 V. B. Cheng, H. H. Suzukawa Jr. and M. Wolfsberg, *J. Chem. Phys.*, 1973, **59**, 3992–3999.
- 74 R. J. Wittebort and A. Szabo, *J. Chem. Phys.*, 1978, **69**, 1722–1736.
- 75 R. Kubo, *J. Math. Phys.*, 1963, **4**, 174–183.
- 76 D. Abergel and A. G. Palmer III, *Concepts Magn. Reson., Part A*, 2003, **19A**, 134–148.
- 77 A. A. Smith, E. Testori, R. Cadalbert, B. H. Meier and M. Ernst, *J. Biomol. NMR*, 2016, **65**, 171–191.

