Paramagpy: Software for Fitting Magnetic Susceptibility Tensors Using Paramagnetic Effects Measured in NMR Spectra

Henry William Orton[1], Thomas Huber[1], and Gottfried Otting[1]

[1]Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

Correspondence: Gottfried Otting (gottfried.otting@anu.edu.au)

Abstract. Paramagnetic metal ions with fast relaxing electrons generate pseudocontact shifts (PCS), residual dipolar couplings (RDC), paramagnetic relaxation enhancements (PRE) and cross-correlated relaxation (CCR) in the NMR spectra of the molecules they bind to. These effects offer long-range structural information in molecules equipped with binding sites for such metal ions. Here we present the new open-source software Paramagpy, which has been written in Python 3 with a graphic user

5   interface. Paramagpy combines the functionalities of different currently available programs to support the fitting of magnetic susceptibility tensors using PCS, RDC, PRE and CCR data and molecular coordinates in Protein Data Bank (PDB) format, including a convenient graphical user interface. Paramagpy uses efficient fitting algorithms to avoid local minima and supports corrections to back-calculated PCS and PRE data arising from cross-correlation effects with chemical shift tensors. The source code is available from https://github.com/henryorton/paramagpy.

10 # 1 Introduction

Paramagnetic metal ions with fast relaxing electrons produce a number of spatially dependent effects in NMR spectra of biomolecules which are useful for probing molecular structure and interactions. These effects arise from the magnetic susceptibility of unpaired electrons, which manifests in NMR spectra most notably as pseudocontact shifts (PCS), paramagnetic relaxation enhancements (PRE) and residual dipolar couplings (RDC), but also as cross-correlated relaxation (CCR) effects.

15 PCSs and RDCs only arise when the magnetic susceptibility is anisotropic, which is the case for all trivalent paramagnetic lanthanide ions except Gd^{3+}.

A number of programs have been developed for fitting the parameters of magnetic susceptibility tensors, χ, to atomic coordinates of biomolecules using the paramagnetic effects experimentally observed in NMR spectra. The program Numbat supports calculation and fitting of the magnetic susceptibility anisotropy tensor, $\Delta\chi$, from experimental PCS data with corrections for

20 residual anisotropic chemical shifts (John et al., 2005) in a convenient graphical user interface (GUI) (Schmitz et al., 2008). The Python module PyParaTools offers similar functionality to Numbat but in a scripting environment and adds methods for fitting χ tensors and alignment tensors using PREs and RDCs, respectively (Stanton-Cook et al., 2014). The software FANTEN offers a convenient web-based GUI for fitting $\Delta\chi$ and alignment tensors from PCS and RDC data sets, respectively (Rinaldelli et al., 2015).

RDCs arise not only from paramagnetism, but also in the presence of alignment media such as dilute liquid crystals. The programs PALES (Zweckstetter and Bax, 2000) and REDCAT (Valafar and Prestegard, 2004) fit alignment tensors to atomic coordinates using RDCs. The program Module can use RDCs to fit alignment tensors for molecular structure refinement (Dosset et al., 2001). PCS and RDC restraints have also been implemented in the software packages CYANA (Balayssac et al., 2006), XPLOR-NIH (Banci et al., 2004), Rosetta (Schmitz et al., 2012; Raman et al., 2010) and HADDOCK (Dominguez et al., 2003; de Vries et al., 2010) for structure determination and refinement.

The coordinates of paramagnetic centres can also be determined from PREs and suitable fitting programs include the programs RelaxGUI (Bieri et al., 2011) and Spinach (Hogben et al., 2011). CCR is observed as a difference in relaxation rate within the multiplet structure of scalar coupled resonances which arises from interference between Curie spin relaxation and dipole-dipole relaxation pathways (Ghose and Prestegard, 1997; Bertini et al., 2002a). The software FANTACROSS supports calculation of CCR, but does not allow fitting of the χ tensor position (Bertini et al., 2001b).

NMR spectra of biomolecules labelled with paramagnetic metal ions with fast electronic relaxation rates, as afforded by lanthanide tags, simultaneously display PCS, RDC, PRE and CCR effects in the same spectrum (Pintacuda et al., 2004b). In principle, the $\Delta\chi$ tensor determined from PCS measurements allows predicting the RDCs, and RDCs arising from paramagnetic alignment allow predictions of some of the $\Delta\chi$-tensor parameters. In addition, the $\Delta\chi$ tensor is reflected in anisotropic PREs. At present, only the software PyParaTools offers integration of all of these effects, but lacks a GUI, does not compute CCR effects and misses some of the options for refined PCS and PRE calculations. Here we present a new Python-based program, Paramagpy, which offers a graphical interface for fitting magnetic susceptibility tensors using PCS, RDC, PRE and CCR data, and seamless transition between these calculations. The fitting routine of Paramagpy for determining $\Delta\chi$ tensors from PCSs employs an efficient grid search algorithm as previously implemented in GPS-Rosetta (Schmitz et al., 2012). The algorithm is adept at overcoming the local minima problem that sometimes compromises the results obtained with Numbat and PyParaTools. Paramagpy uses both Curie spin (Guéron, 1975) and Solomon-Bloembergen-Morgan (Solomon, 1955) theory to calculate PREs, and it includes cross-correlation effects with anisotropic chemical shift tensors (Pintacuda et al., 2004a), which have not been taken into account by any previous tensor-fitting software. Paramagpy can be installed as a Python module and scripted for efficient calculations, or run via an intuitive GUI.

Calculations using Paramagpy have been verified with data from previous publications. This includes fitting of $\Delta\chi$ tensors to amide PCS data of lanthanide-loaded calbindin D_{9k} and calculating PREs for amide ^{1}H spins (Orton and Otting, 2018). Paramagpy has also been used successfully to predict cross-correlated CSA-Curie spin relaxation giving rise to negative PREs for amide ^{15}N spins (Orton et al., 2016). CCR calculations have been verified with data from high- and low-spin paramagnetic myoglobin (Pintacuda et al., 2003). Paramagpy has been shown to fit alignment tensors consistent with previous results for lanthanide-tagged ubiquitin (Pearce et al., 2017), but may also be applied to datasets arising from alignment media where paramagpy reports alignment and Saupe tensors alongside $\Delta\chi$ tensors. Paramagpy can thus be used with RDC data obtained by any means of weak molecular alignment in the magnetic field, substituting softwares like Module. (Dosset et al., 2001).

2 Pseudocontact shifts

The magnetic susceptibility tensor χ associated with a paramagnetic centre creates a dipolar shielding tensor σ at a given position $\boldsymbol{r}$ and distance r from the paramagnetic centre as shown in equation (1), where $\mathbb{I}_3$ is the 3×3 identity matrix and $\otimes$ denotes the Kronecker product.

$$\sigma = \frac{1}{4\pi}\left[3\frac{\boldsymbol{r} \otimes \boldsymbol{r}^T}{r^5} - \frac{\mathbb{I}_3}{r^3}\right].\chi \tag{1}$$

$$= \frac{1}{4\pi r^5}\begin{bmatrix} (3x^2 - r^2) & 3xy & 3xz \\ 3xy & (3y^2 - r^2) & 3yz \\ 3xz & 3yz & (3z^2 - r^2) \end{bmatrix} . \begin{bmatrix} \chi_{xx} & \chi_{xy} & \chi_{xz} \\ \chi_{xy} & \chi_{yy} & \chi_{yz} \\ \chi_{xz} & \chi_{yz} & \chi_{zz} \end{bmatrix} \tag{2}$$

The PCS is given by the trace of the shielding tensor as shown in the PCS equation (3). The $\Delta\chi$ tensor is given by the traceless part of the χ tensor. Considering only the $\Delta\chi$ tensor, a linear form of the PCS equation can be obtained, which characterises the $\Delta\chi$ tensor by 5 explicit parameters as shown in equation (4). Including the 3 position parameters represented by the coordinates of the metal centre (x, y, z), solving the PCS equation requires determining 8 parameters in total.

$$\delta^{\text{PCS}} = \frac{1}{3}Tr[\sigma] \tag{3}$$

$$= \frac{1}{4\pi r^5}\begin{bmatrix} x^2 - z^2, & y^2 - z^2, & 2xy, & 2xz, & 2yz \end{bmatrix} . \begin{bmatrix} \Delta\chi_{xx} \\ \Delta\chi_{yy} \\ \Delta\chi_{xy} \\ \Delta\chi_{xz} \\ \Delta\chi_{yz} \end{bmatrix} \tag{4}$$

2.1 Singular value decomposition (SVD) grid search

Equation (4) can be rewritten in matrix form to give equation 5, where B is a column vector of length n of the calculated PCS values, x is a column vector of length 5 of the $\Delta\chi$-tensor parameters and A is a $n \times 5$ matrix as defined by the row vector in equation 4 containing coordinate parameters.

$$A.x = B \tag{5}$$

$$x = A^+.B \tag{6}$$

$$x = (W.A)^+.(W.B) \tag{7}$$

Populating vector B with many experimental PCS values and the matrix A with atomic coordinates from a molecule of known structure, the system is likely over-determined and a least-squares solution for the $\Delta\chi$-tensor parameters x can be obtained analytically by considering the singular values of the matrix A and constructing the pseudo-inverse A^+. This allows calculation of the best-fitting tensor at a given position by equation 6 (Schmitz et al., 2012). A weighted least-squares fit can be obtained using equation 7, where the square matrix W contains the weights along the diagonal $W_{ii} = 1/S_{\text{PCS},i}$, which may be sourced from the experimental standard deviations $S_{\text{PCS},i}$ of the ith spin.

Since this calculation is fast, a grid search over many positions of the paramagnetic centre is feasible, providing a robust initial guess prior to iterative refinement of the tensor position by non-linear gradient-descent methods. Paramagpy can evaluate 5000 grid points for 50 PCS values in under one second using a 2 GHz Intel i5 2016 processor of a typical laptop computer.

2.2 Non-linear gradient descent

5 When fitting of the position of the paramagnetic centre is required, the PCS equation becomes non-linear. A fit can be found iteratively by minimising the sum of squares of the differences between experimental and back-calculated PCS values. An efficient method for minimisation is by non-linear gradient descent. We chose the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (Fletcher, 1988) for non-linear least-squares minimisation of the cost function in equation 8 . Here, $\text{PCS}_i^{\text{exp}}$ and $\text{PCS}_i^{\text{cal}}$ are, respectively, the experimental and back-calculated PCSs for spin i, and $S_{\text{PCS},i}$ is the experimental uncertainty in the 10 PCS of spin i.

$$cost = \sum_i \frac{(\text{PCS}_i^{\text{cal}} - \text{PCS}_i^{\text{exp}})^2}{S_{\text{PCS},i}^2} \tag{8}$$

2.3 Multiple PCS data-sets

Often there are multiple PCS data-sets available for different metal ions bound at the same position, obtained from multiple samples prepared with different metal ions. A simultaneous fit of the common position is possible, independently fitting the 15 tensor magnitude and orientation for each data set, and can lead to a more accurate overall position of the paramagnetic centre. Paramagpy supports multiple data sets for simultaneous fitting of a common metal position by both the SVD grid-search and non-linear gradient-descent algorithms.

2.4 Corrections to PCS calculations

An anisotropic magnetic susceptibility causes alignment of the molecule in the external magnetic field. As molecular orienta-20 tions are no longer sampled uniformly, shielding tensors may no longer average to their isotropic values. In this situation, the chemical shift actually observed in the paramagnetic sample contains contributions from residual anisotropic chemical shifts (RACS) arising from non-zero averaging of the chemical shift anisotropy (CSA) tensor. Paramagpy supports PCS calculations that include RACS corrections (John et al., 2005). Paramagpy provides standard CSA tensors for amide [1]H-spins and back-bone amide [15]N- and carbonyl ^{13}C-spins (Cornilescu and Bax, 2000). Customized CSA tensors may also be set for any of the 25 nuclear spins.

In addition to the CSA tensor, there is also a dipolar shielding tensor σ at the site of a nuclear spin, which arises from the magnetic susceptibility of the paramagnetic centre. In analogy to the RACS effect, this can lead to a residual anisotropic dipolar shift (RADS), which is a small perturbation to the observed PCS in paramagnetic samples arising from molecular alignment (Bertini et al., 2002b). Paramagpy includes RADS as an option in the PCS calculation and $\Delta\chi$-tensor fitting routines.

Systematic errors in experimental PCS values can arise due to variations in the carrier frequency or calibration of the NMR spectra recorded of the diamagnetic and paramagnetic species. This offset can be included as a parameter during the fitting of $\Delta\chi$ tensors, although doing so is meaningful only, if a sufficient number of PCS data are available to avoid over-fitting.

3 Residual dipolar couplings

An anisotropic magnetic susceptibility tensor induces a coincident alignment tensor $\mathbf{A}$, giving rise to RDCs between nuclear spins. The alignment tensor can be found from the $\Delta\chi$ tensor using equation 9, where B_0 is the magnetic field, μ_0 the vacuum permeability, k_B the Boltzmann constant and T the temperature (Bertini et al., 2002b).

$$\mathbf{A} = \frac{B_0^2}{15\mu_0 k_\mathrm{B} T}\Delta\chi \tag{9}$$

The RDC values can be calculated using equation 10, where $\boldsymbol{r}_{AB}$ is the inter-nuclear vector and r_{AB} the distance between the two nuclei A and B (Kramer et al., 2004). This can be expanded into the vector equation 11, where x, y and z are the Cartesian coordinates of the inter-nuclear vector $\boldsymbol{r}_{AB}$.

$$RDC = \frac{3\gamma_A\gamma_B\mu_0\hbar}{8\pi^2 r_{AB}^5}\boldsymbol{r}_{AB}^T.\mathbf{A}.\boldsymbol{r}_{AB} \tag{10}$$

$$= \frac{3\gamma_A\gamma_B\mu_0\hbar}{8\pi^2 r_{AB}^5}\begin{bmatrix} x^2-z^2, & y^2-z^2, & 2xy, & 2xz, & 2yz \end{bmatrix}.\begin{bmatrix} \mathbf{A}_{xx} \\ \mathbf{A}_{yy} \\ \mathbf{A}_{xy} \\ \mathbf{A}_{xz} \\ \mathbf{A}_{yz} \end{bmatrix} \tag{11}$$

Unlike the PCS tensor, the RDC tensor does not require parameters for position and can therefore be described by 5 parameters for magnitude and orientation. Fitting can therefore be achieved by a linear least squares fit.

3.1 SVD fitting-algorithm

Paramagpy uses the SVD algorithm similar to the original implementation in the program REDCAT (Valafar and Prestegard, 2004). It is functionally the same as the algorithm applied to solving the PCS equation in section 2.1. A $n \times 5$ matrix as defined by the row vector in equation 11 containing coordinate parameters is constructed. From this, a pseudo-inverse matrix is calculated and applied to the experimental RDC values, yielding the best-fitting alignment tensor.

4 Paramagnetic relaxation enhancements

PREs describe the relaxation rates of longitudinal magnetisation, $R_1 = 1/T_1$, or transverse magnetisation, $R_2 = 1/T_2$, of nuclear spins, where T_1 and T_2 are the longitudinal and transverse relaxation times, respectively. For PREs of paramagnetic molecules in solution, the relaxation rates are governed by dipole-dipole interactions as described by the Solomon-

Bloembergen-Morgan (SBM) equations or the shielding tensor anisotropy as described by the Curie spin equations (Solomon, 1955; Guéron, 1975).

4.1 Solomon-Bloembergen-Morgan theory

The SBM equations for R_1 and R_2 are shown in equations 12 and 13, respectively, where γ is the nuclear gyromagnetic ratio,
5 r the distance of the nucleus from the paramagnetic centre, and ω and ω_S are the nuclear and electronic Larmor frequencies, respectively. τ_c is the correlation time calculated as $1/\tau_c = 1/\tau_r + 1/T_{1e}$, where τ_r is the rotational correlation time of the molecule and T_{1e} is the electronic relaxation time. μ_{eff} is the effective magnetic moment of the paramagnetic centre, which can be predicted from the Landé g-factor, the Bohr magneton μ_{B} and the total angular momentum quantum number J (equation 14).

$$R_1^{\text{SBM}} = \frac{2}{15}\left(\frac{\mu_0}{4\pi}\frac{\gamma\mu_{\text{eff}}}{r^3}\right)^2\left[\frac{3\tau_c}{1+\tau_c^2\omega^2} + \frac{7\tau_c}{1+\tau_c^2\omega_S^2}\right] \tag{12}$$

$$R_2^{\text{SBM}} = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\frac{\gamma\mu_{\text{eff}}}{r^3}\right)^2\left[4\tau_c + \frac{3\tau_c}{1+\tau_c^2\omega^2} + \frac{13\tau_c}{1+\tau_c^2\omega_S^2}\right] \tag{13}$$

$$\mu_{\text{eff}} = g\mu_{\text{B}}\sqrt{J(J+1)} \tag{14}$$

An extension to the SBM theory which accounts for anisotropy of the dipolar spectral density is described by equations 15 and 16 where $\mathbf{G}(\omega)$ describes the spectral power density tensor (Suturina et al., 2018b). $\hat{r}$ is the unit vector from the
15 paramagnetic centre to the nuclear spin. The spectral power density tensor usually cannot be derived theoretically, but is instead fitted to experimental data.

$$R_1^{\text{SBM-aniso}} = \frac{2}{3}\left(\frac{\mu_0}{4\pi}\frac{\gamma}{r^3}\right)^2 Tr\left[(3\hat{r}\otimes\hat{r}-\mathbb{I}_3)^2.\mathbf{G}(\omega)\right] \tag{15}$$

$$R_2^{\text{SBM-aniso}} = \frac{1}{3}\left(\frac{\mu_0}{4\pi}\frac{\gamma}{r^3}\right)^2 Tr\left[(3\hat{r}\otimes\hat{r}-\mathbb{I}_3)^2.(\mathbf{G}(0)+\mathbf{G}(\omega))\right] \tag{16}$$

4.2 Curie spin theory

20 Curie-spin relaxation is governed by the dipolar shielding tensor σ as calculated in equation 1, which must include the isotropic component of the χ tensor, χ_{iso}, which can be predicted using equation 17. The first invariant Λ and second invariant Δ of the shielding tensor are calculated by equations 18 and 19, where σ_{ij} denotes the i,j-th component of the shielding tensor σ (Suturina et al., 2018a). This allows calculation of the R_1 and R_2 PREs by equations 20 and 21, respectively. These equations

account for anisotropy of the magnetic susceptibility, provided equation 1 is used to calculate σ (Vega and Fiat, 1976).

$$\chi_{\text{iso}} = \frac{\mu_0 \mu_{\text{eff}}^2}{3k_{\text{B}}T} \tag{17}$$

$$\Lambda^2 = (\sigma_{xy} - \sigma_{yx})^2 + (\sigma_{xz} - \sigma_{zx})^2 + (\sigma_{yz} - \sigma_{zy})^2 \tag{18}$$

$$\Delta^2 = \sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2 - \sigma_{xx}\sigma_{yy} - \sigma_{xx}\sigma_{zz} - \sigma_{yy}\sigma_{zz} + \frac{3}{4}\left[((\sigma_{xy} + \sigma_{yx})^2 + (\sigma_{xz} + \sigma_{zx})^2 + (\sigma_{yz} + \sigma_{zy})^2]\right. \tag{19}$$

$$R_1^{\text{Curie}} = \frac{1}{2}\Lambda^2\omega^2\left[\frac{\tau_r}{1 + 9\tau_r^2\omega^2}\right] + \frac{2}{15}\Delta^2\omega^2\left[\frac{\tau_r}{1 + \omega^2\tau_r^2}\right] \tag{20}$$

$$R_2^{\text{Curie}} = \frac{1}{4}\Lambda^2\omega^2\left[\frac{\tau_r}{1 + 9\tau_r^2\omega^2}\right] + \frac{1}{45}\Delta^2\omega^2\left[4\tau_r + \frac{3\tau_r}{1 + \tau_r^2\omega^2}\right] \tag{21}$$

When PREs due to Curie-spin relaxation are cross-correlated with CSA relaxation, the CSA tensor is added to the dipolar shielding tensor to give an effective shielding tensor σ_{eff}. The PRE including CSA cross-correlation $R_{\text{Curie}\times\text{CSA}}$ is determined as the difference in relaxation rates in the paramagnetic and diamagnetic state as shown in equation 22. This can give rise to negative PREs as shown previously and confirmed by experiment (Pintacuda et al., 2004a; Orton et al., 2016).

$$R^{\text{Curie}\times\text{CSA}} = R^{\text{Curie}}(\sigma_{\text{eff}}) - R^{\text{Curie}}(\sigma_{\text{CSA}}) \tag{22}$$

4.3 Fitting algorithm

Paramagpy includes routines to calculate PREs and fit all parameters for each of the above relaxation theories, including cross-correlated relaxation with CSA effects. This is achieved by non-linear gradient descent to minimise the cost function of equation 23. Here, $\text{PRE}_i^{\text{exp}}$ and $\text{PRE}_i^{\text{cal}}$ are, respectively, the experimental and back-calculated PREs for spin i, and $S_{\text{PRE},i}$ is the experimental uncertainty in the PRE of spin i. The user can choose to fit or constrain different parameters, such as the magnetic susceptibility or power spectral density tensor position, magnitude, correlation time τ_c, etc. Parameter templates for lanthanide ions are also provided, based on tensor magnitudes and anisotropies previously reported for lanthanide complexes of calbindin D_{9k} (Bertini et al., 2001a). These may be used to give a quick estimate of expected PRE values.

$$cost = \sum_i \frac{(\text{PRE}_i^{\text{cal}} - \text{PRE}_i^{\text{exp}})^2}{S_{\text{PRE},i}^2} \tag{23}$$

5 Curie spin dipole-dipole cross-correlated relaxation

Interference of the internuclear dipole-dipole (DD) relaxation with Curie-spin relaxation provides a mechanism for differential relaxation rates of multiplet components by cross-correlated relaxation (CCR) (Ghose and Prestegard, 1997). This effect is readily observed and measured as the difference in the relaxation rate R_2 of the two doublet components of an amide $^1\text{H}-^{15}\text{N}$ spin pair. In this case, the shielding tensor arising at the ^1H spin due to the ^{15}N dipole is given by equation 24, where $\boldsymbol{r}_{\text{HN}}$ is the H$-$N bond vector, r_{HN} is the internuclear distance, γ_{N} is the gyromagnetic ratio of ^{15}N and $I = \frac{1}{2}$ is the spin of ^{15}N. The factor of $1/B_0$ is necessary to express the ^{15}N shielding tensor with units of ppm to match the units of the Curie spin shielding tensor. The effective shielding tensor for the ^1H spin due to both the Curie spin and the ^{15}N dipole in either the up

or down spin state is given by equations 25 and 26 respectively. The relaxation rate R_2^{Curie} is then calculated using equation 21 for both the up and down effective shielding tensors $\sigma_\uparrow$ and $\sigma_\downarrow$, and their difference is taken to represent the Curie×DD differential line-broadening $R^{\text{Curie}\times\text{DD}}$. In this way the auto-correlated relaxation mechanisms arising from the separate DD and Curie mechanisms are subtracted out, leaving the pure cross-correlated term.

$$\sigma_{\text{N}} = \frac{1}{B_0}\frac{\mu_0}{4\pi}\gamma_{\text{N}}\hbar I\left[3\frac{\boldsymbol{r}_{\text{HN}}\otimes\boldsymbol{r}_{\text{HN}}^T}{r_{\text{HN}}^5} - \frac{\mathbb{I}_3}{r_{\text{HN}}^3}\right] \tag{24}$$

$$\sigma_\uparrow = \sigma + \sigma_{\text{N}} \tag{25}$$

$$\sigma_\downarrow = \sigma - \sigma_{\text{N}} \tag{26}$$

$$R^{\text{Curie}\times\text{DD}} = R^{\text{Curie}}(\sigma_\uparrow) - R^{\text{Curie}}(\sigma_\downarrow) \tag{27}$$

Paramagpy uses the above equations for all DD×Curie relaxation calculations. By using equation 1 for calculating the Curie-spin shielding tensor σ, these equations also account for anisotropy of the magnetic susceptibility χ. CCR values can be calculated between any two atoms in the specified Protein Data Bank (PDB) file. The calculations have been shown to agree with previous experimental CCR data on high- and low-spin myoglobin (Pintacuda et al., 2003).

5.1 Fitting algorithm

Paramagpy includes routines to fit all parameters of the χ tensor, including position, magnitude and anisotropy to experimentally measured CCR data. This is achieved by non-linear gradient descent to minimise the cost function of equation 28. Here, $\text{CCR}_i^{\text{exp}}$ and $\text{CCR}_i^{\text{cal}}$ are, respectively, the experimental and back-calculated CCRs for spin i, and $S_{\text{CCR},i}$ is the experimental uncertainty in the CCR of spin i.

$$cost = \sum_i \frac{(\text{CCR}_i^{\text{cal}} - \text{CCR}_i^{\text{exp}})^2}{S_{\text{CCR},i}^2} \tag{28}$$

6 Uncertainty calculations

To judge the quality of a $\Delta\chi$ or χ tensor fitted using PCS, RDC, PRE or CCR data, Paramagpy offers two methods to test the robustness of the fit, bootstrap and Monte-Carlo. The bootstrapping method repeats the fit many times, with each iteration randomly sampling a specified proportion of the data, and subsequently reports the standard deviation in the fitted parameters. The Monte-Carlo method repeats the fit using all the data, but each time adds noise to the experimental values. The noise is sourced from a uniform distribution that has been scaled by values provided by the user for each atom. These scaling values are ideally calculated from noise in the spectrum to reflect uncertainty in peak positions or amplitudes (Kontaxis et al., 2000). The standard deviations in the fitted tensor parameters are then reported.

7   Molecular structures with multiple models

The PDB file format for the structures of biomolecules determined by solution NMR is often represented by an ensemble of conformers to capture the experimental uncertainties of structure determination. Paramagpy accounts for these conformer ensembles by calculating ensemble-averaged paramagnetic effects at all stages of calculations and fitting. Effects calculated for the atoms of different models are averaged, if the atoms have the same atom number in the PDB file. Custom ensemble averaging behaviour can be changed by the user in the scripted environment. In the implementations of the SVD algorithm, ensemble averaging involves summation of rows for common atoms of the matrix A of equation 5 before calculation of the singular values. In the implementations of the non-linear gradient descent algorithm, the values calculated for the common atoms are averaged prior to calculating the sum of squares of differences. This is shown in equation 29 where a^{cal} and a^{exp} are the calculated and experimental PCS, RDC, PRE or CCR values, respectively. The index m is for atoms that are common between models, and the index i is over all atoms in the structure. The effect of ensemble-averaged fitting is particularly noticeable for RDCs, as the ensemble average of RDCs can be much smaller than the corresponding RDC of a single conformer.

$$cost_{\mathrm{ensemble}} = \sum_i \frac{\left(\sum_m \left[\mathrm{a}^{\mathrm{cal}}_{m,i} - \mathrm{a}^{\mathrm{exp}}_i\right]\right)^2}{\sigma^2_{\mathrm{a(i)}}} \qquad (29)$$

8   Quality factors

To judge the agreement of tensor fits with the experimental data, a Q-factor can be assigned to a given fit, which Paramagpy calculates using equation 30. Here, the experimental and calculated PCS, RDC, PRE or CCR values are denoted a^{exp} and a^{cal}, respectively, the index m is for ensemble averaging of common spins between models and the index i is for summation over all spins of the molecule. A low Q-factor signifies a good quality fit.

$$Q = \sqrt{\frac{\sum_i \left[\left(\sum_m \left[a^{\mathrm{exp}}_i - a^{\mathrm{cal}}_{m,i}\right]\right)^2\right]}{\sum_i \left[\left(\sum_m \left[a^{\mathrm{exp}}_i\right]\right)^2\right]}} \qquad (30)$$

9   Graphical User Interface

Paramagpy has a graphical user interface (GUI) written for the inbuilt Tk/Tcl interface of Python 3, which can run on Mac OSX, Windows and Linux operating systems. The GUI offers a user-friendly environment for loading and visualising PDB files and experimental PCS, RDC, PRE and CCR data. Two frames display the initial and fitted tensor. The fitted tensor is calculated and displayed by the push of a button. An overview of the PCS fitting tab is shown in Figure 1. Hovering the mouse over any element in the window displays a useful tool tip to help the user.
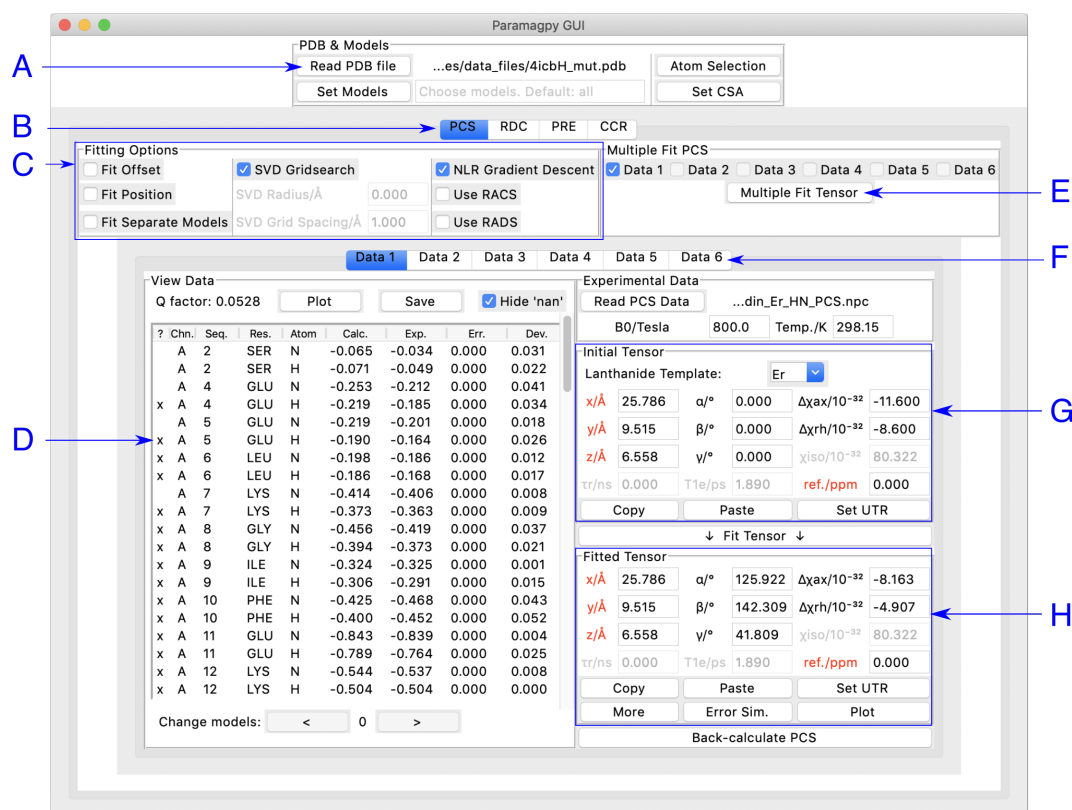
Figure 1. Paramagpy GUI running on Mac OSX. A: Frame for loading PDB coordinates. The atoms and models (conformers) of interest can be selected and CSA tensor parameters set by the user. B: The user can switch between PCS, RDC, PRE and CCR tabs. C: Fitting options can be specified by selecting the relevant check box. The SVD grid search searches for the best-fit tensor within a sphere about the initial tensor origin with radius and grid spacing as specified. NLR Gradient Descent refines the tensor using non-linear least squares minimisation. D: Experimental data for atoms in the PDB file are displayed here. The first column contains an 'x' if the datum will be used during fitting and may be toggled by pressing the 'x' key on the keyboard. Experimental and back-calculated PCS values are also reported and their correlation can be displayed by clicking the 'Plot' button above. E: To utilise multiple PCS data sets to fit different tensors to a common position, the Multiple Fit Tensor button can be clicked after selecting the desired data sets. F: Each tab can contain a different PCS data set, allowing up to 6 to be loaded. If more data sets are required, Paramagpy supports this through the scripted module. G: The initial tensor parameters can be specified here to define a starting point before fitting. For convenience, the paramagnetic centre can be positioned at any atom in the PDB file by double-clicking on a row of the data view in the frame to the left. Parameters in red are constrained during fitting. Greyed out parameters are not relevant to PCS or RDC calculations, but are used in PRE and CCR calculations. H: The fitted tensor is displayed here. Clicking the Copy button allows the tensor to be pasted into other tabs of the program (see B and F above). The Plot button will prompt the user to save an isosurface file for opening in PyMOL. Error Sim. will assess the quality of the fit by bootstrap or Monte-Carlo methods. Set UTR converts the tensor parameters to the unique tensor representation defined by Numbat (Schmitz et al., 2008).

10 Visualisation

Paramagpy offers a number of plot options to quickly visualise tensors and quality of fit. The scalar PCS or PRE field can be written to a CCP4 (McNicholas et al., 2011) density map, which can then be visualised as a three-dimensional contour plot in the program PyMOL (Schrödinger, 2015). The fit quality can be visualized in correlation plots of back-calculated PCS, RDC,

5 PRE and CCR values versus the experimental values. Finally, a scatter plot of the principle axes of the tensors can be viewed in a Sanson–Flamsteed-style projection following Monte-Carlo or bootstrap error analyses. Example plots are summarised in Figure 2.

11 Scripting

Paramagpy is a python module and can be imported into a scripting environment. The module is split into 4 major submod-

10 ules. (i) The 'metal' submodule deals with the paramagnetic centre, tensor representations and methods for calculating PCS, RDC, PRE and CCR values. (ii) The 'protein' submodule handles the atomic coordinates from the PDB file and CSA tensor definitions. (iii) The 'dataparse' submodule manages the reading and writing of data files. (iv) The 'fit' submodule contains functions for fitting tensors to experimental data. An example script for fitting of a $\Delta\chi$ tensor to experimental PCS data for calbindin D_{9k} is shown in Figure 3. It uses only 9 lines of code. Some more advanced features of paramgpy, such as fitting of

15 power spectral density tensors in equation 15 and 16 are only availble in the scripted environment. The scripted environment also offers control over which parameters are included for fitting routines and allows calculations for coordinates other than PDB formats.

12 NMR software integration

Paramagpy includes macro scripts to interface with popular NMR software: CCPNMR analysis and Sparky (Vranken et al.,

20 2005; Lee et al., 2014). Currently, these macros allow for the rapid calculation of experimental PCS values from NMR spectra with up to 3 dimensions, fitting of $\Delta\chi$ tensors and plotting of back-calculated PCS values onto paramagnetic spectra.

13 Tensor conventions and conversions

Paramagpy offers a number of simple routines to convert between tensor representations. In addition to the 3×3 matrix representations of tensors, positions, rotation matrices, eigenvalues, axial/rhombic components and Euler angles, alignment

25 tensors and Saupe tensors are available upon clicking the 'More' button within the GUI. The axial and rhombic components are defined as follows (equations 31 and 32).

$$\Delta\chi_{\mathrm{ax}} = \Delta\chi_{zz} - \left(\frac{\Delta\chi_{xx} + \Delta\chi_{yy}}{2} \right) \tag{31}$$

$$\Delta\chi_{\mathrm{rh}} = \Delta\chi_{xx} - \Delta\chi_{yy} \tag{32}$$
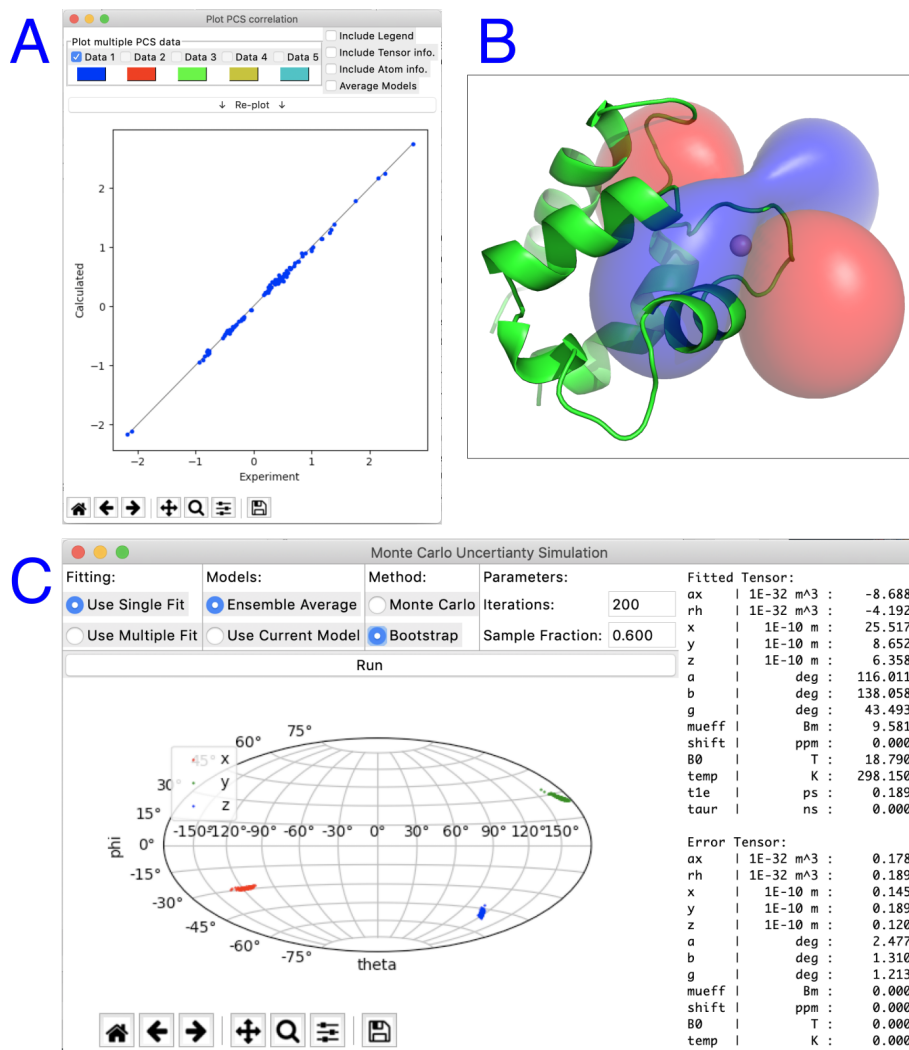
Figure 2. Plotting options available in Paramagpy illustrated with data of calbindin D_{9k} loaded with Er^{3+}. A: Correlation plot of calculated versus experimental PCS values after fitting of the $\Delta\chi$ tensor. B: PCS isosurface plot viewed in PyMOL. C: Sanson-Flamsteed plot showing the principle axes projections after bootstrap analysis. The Error-Tensor reports the standard deviation in fitted parameters.

```
 1  from paramagpy import protein, fit, dataparse, metal
 2  # Load the PDB file
 3  prot = protein.load_pdb('../data_files/4icbH_mut.pdb')
 4  # Load the PCS data
 5  rawData = dataparse.read_pcs('../data_files/calbindin_Er_HN_PCS.npc')
 6  # Associate PCS data with atoms of the PDB
 7  parsedData = prot.parse(rawData)
 8  # Define an initial tensor
 9  mStart = metal.Metal()
10  # Set the starting position to an atom close to the metal
11  mStart.position = prot[0]['A'][56]['CA'].position
12  # Calculate an initial tensor from an SVD gridsearch
13  mGuess, calc, qfac = fit.svd_gridsearch_fit_metal_from_pcs(
14      [mStart],[parsedData], radius=10, points=10)
15  # Refine the tensor using non-linear regression
16  mFit, calc, qfac = fit.nlr_fit_metal_from_pcs(mGuess, [parsedData])
17  # Save the fitted tensor to file
18  mFit[0].save('calbindin_Er_HN_PCS_tensor.txt')
```

```
 1  ax    | 1E-32 m^3 :    -8.688
 2  rh    | 1E-32 m^3 :    -4.192
 3  x     | 1E-10 m  :    25.517
 4  y     | 1E-10 m  :     8.652
 5  z     | 1E-10 m  :     6.358
 6  a     |      deg :   116.011
 7  b     |      deg :   138.058
 8  g     |      deg :    43.492
 9  mueff |       Bm :     0.000
10  shift |      ppm :     0.000
11  B0    |        T :    18.790
12  temp  |        K :   298.150
13  t1e   |       ps :     0.000
14  taur  |       ns :     0.000
```

Figure 3. Example python script for fitting a $\Delta\chi$ tensor to experimental PCS data. The output with fitted tensor parameters is displayed to the right.

By default, Paramagpy reports all fitted tensors in the unique tensor representation used by the program Numbat (Schmitz et al., 2008). This requires that the principle axes magnitudes of the $\Delta\chi$ tensor are ordered $|\Delta\chi_{zz}| \geq |\Delta\chi_{yy}| \geq |\Delta\chi_{xx}|$ and all Euler angles are in the range $[0, \pi]$ using the ZYZ convention.

14 Example PRE calculation

5   PRE calculations that include anisotropy effects and cross-correlation with CSA can be daunting to set up as they require the $\Delta\chi$ and CSA tensors to possess the correct orientations in the frame of the molecular coordinates. Paramagpy simplifies this for the user by allowing $\Delta\chi$ tensors fitted from PCS data to be transferred easily to the tab for PRE calculations. Furthermore, CSA tensor templates are provided for most protein backbone atoms.

As an example, Figure 4 shows the Paramagpy GUI with $R_1(^{15}\text{N})$ PRE data for calbindin D_{9k} loaded with Tb^{3+} (Orton
10 et al., 2016). A $\Delta\chi$ tensor was fitted using the PCS tab, then transferred to the PRE tab using the 'Copy' and 'Paste' buttons. Curie-spin/CSA cross-correlation is taken into account simply by checking the box "Use CSA". This greatly improves the correlation and allows the prediction of negative PREs. The small additional correction arising from the anisotropy of the Curie spin can be included by setting the $\Delta\chi_{\text{ax}}$ and $\Delta\chi_{\text{rh}}$ parameters to the non-zero values obtained from the $\Delta\chi$ tensor fitted with the help of PCS data.

15 ## 15 Conclusions

Paramagpy is an easy-to-use program that integrates the related paramagnetic NMR phenomena of PCS, RDC, PRE and CCR. Paramagpy allows the rapid analysis of NMR spectra of samples containing a single paramagnetic centre, which is particularly useful for data recorded with different paramagnetic lanthanide ions. With an intuitive calculation flow, Paramagpy can be used, for example, to fit a $\Delta\chi$ tensor using experimental PCS data and then quickly report the expected PREs of the same
20 complex, informing the user which signals may be too broad to observe. Paramagpy uses efficient fitting algorithms and an
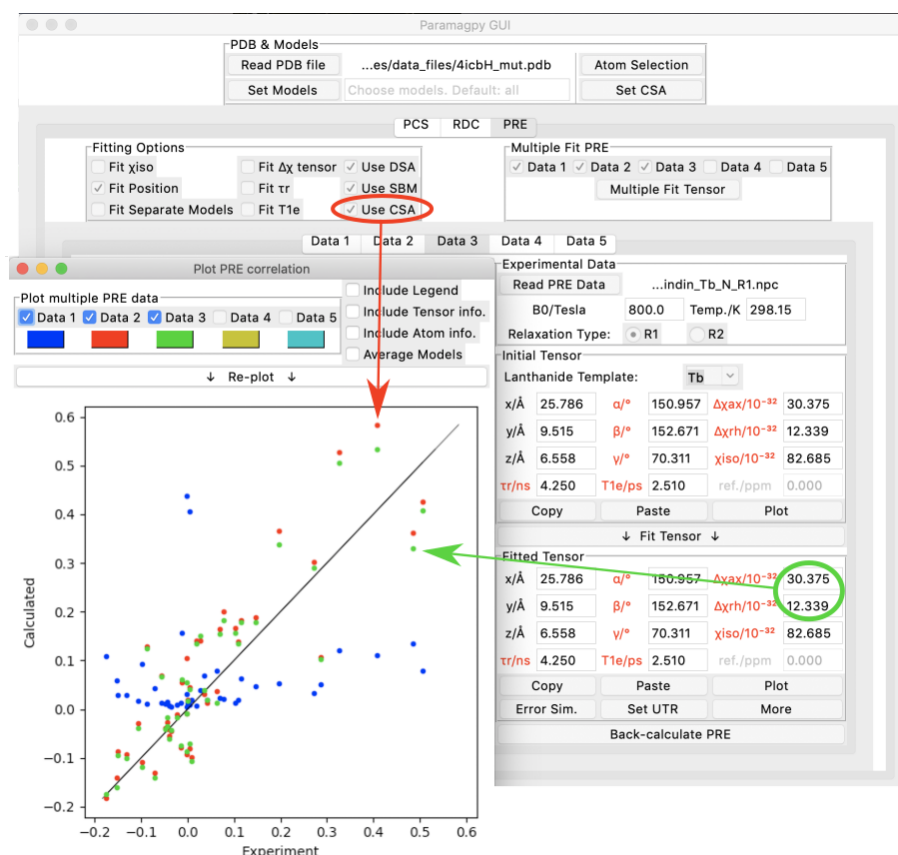
Figure 4. Paramagpy GUI showing $R_1(^{15}N)$ PRE data for calbindin D_{9k} loaded with Tb^{3+}. The correlation plot shows calculated vs. experimental values. Blue: SBM and isotropic Curie-spin theory are used for calculating PREs. Red: Taking into account also the cross-correlation between Curie spin and CSA relaxation. Green: Including the additional correction arising from the anisotropy of the χ tensor.

up-to-date implementation of paramagnetic NMR theory to capture subtle corrections arising from CSA and anisotropy effects in the PCS and PRE calculations.

Code availability. https://github.com/henryorton/paramagpy

Competing interests. There are no competing interests.

Acknowledgements. H. W. O. thanks the Westpac Bicentennial foundation for a Future Leaders Scholarship. Financial support by the Australian Research Council, including a Laureate Fellowship to G. O., is gratefully acknowledged.

References

Balayssac, S., Bertini, I., Luchinat, C., Parigi, G., and Piccioli, M.: ^{13}C Direct Detected NMR Increases the Detectability of Residual Dipolar Couplings, Journal of the American Chemical Society, 128, 15042–15043, https://doi.org/10.1021/ja0645436, 2006.

Banci, L., Bertini, I., Cavallaro, G., Giachetti, A., Luchinat, C., and Parigi, G.: Paramagnetism-Based Restraints for Xplor-NIH, J. Biomol. NMR, 28, 249–261, https://doi.org/10.1023/B:JNMR.0000013703.30623.f7, 2004.

Bertini, I., Janik, M. B. L., Lee, Y.-M., Luchinat, C., and Rosato, A.: Magnetic Susceptibility Tensor Anisotropies for a Lanthanide Ion Series in a Fixed Protein Matrix, J. Am. Chem. Soc., 123, 4181–4188, https://doi.org/10.1021/ja0028626, 2001a.

Bertini, I., Kowalewski, J., Luchinat, C., and Parigi, G.: Cross Correlation between the Dipole–Dipole Interaction and the Curie Spin Relaxation: The Effect of Anisotropic Magnetic Susceptibility, J. Magn. Reson., 152, 103–108, https://doi.org/10.1006/jmre.2001.2378, 2001b.

Bertini, I., Cavallaro, G., Cosenza, M., Kümmerle, R., Piccioli, M., and Poggi, L.: Cross correlation rates between Curie spin and dipole-dipole relaxation in paramagnetic proteins: The case of Cerium substituted Calbindin D9k, Journal of Biomolecular NMR, 23, 11, 2002a.

Bertini, I., Luchinat, C., and Parigi, G.: Magnetic susceptibility in paramagnetic NMR, Prog. NMR Spec., 40, 249–273, 2002b.

Bieri, M., d'Auvergne, E. J., and Gooley, P. R.: relaxGUI: a new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and μs motion of proteins, J. Biomol. NMR, 50, 147–155, https://doi.org/10.1007/s10858-011-9509-1, 2011.

Cornilescu, G. and Bax, A.: Measurement of Proton, Nitrogen, and Carbonyl Chemical Shielding Anisotropies in a Protein Dissolved in a Dilute Liquid Crystalline Phase, J. Am. Chem. Soc., 122, 10143–10154, https://doi.org/10.1021/ja0016194, 2000.

de Vries, S. J., van Dijk, M., and Bonvin, A. M. J. J.: The HADDOCK web server for data-driven biomolecular docking, Nature Protocols, 5, 883–897, https://doi.org/10.1038/nprot.2010.32, 2010.

Dominguez, C., Boelens, R., and Bonvin, A. M. J. J.: HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information, Journal of the American Chemical Society, 125, 1731–1737, 2003.

Dosset, P., Hus, J.-C., Marion, D., and Blackledge, M.: A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings, J. Biomol. NMR, 20, 223–231, 2001.

Fletcher, R.: Practical Methods of Optimization, John Wiley & Sons, New York, 2 edn., 1988.

Ghose, R. and Prestegard, J. H.: Electron Spin–Nuclear Spin Cross-Correlation Effects on Multiplet Splittings in Paramagnetic Proteins, J. Magn. Reson., 128, 138–143, https://doi.org/10.1006/jmre.1997.1227, 1997.

Guéron, M.: Nuclear relaxation in macromolecules by paramagnetic ions: a novel mechanism, J. Magn. Reson., 19, 58–66, https://doi.org/10.1016/0022-2364(75)90029-3, 1975.

Hogben, H., Krzystyniak, M., Charnock, G., Hore, P., and Kuprov, I.: Spinach – A software library for simulation of spin dynamics in large spin systems, J. Magn. Reson., 208, 179–194, https://doi.org/10.1016/j.jmr.2010.11.008, 2011.

John, M., Park, A. Y., Pintacuda, G., Dixon, N. E., and Otting, G.: Weak Alignment of Paramagnetic Proteins Warrants Correction for Residual CSA Effects in Measurements of Pseudocontact Shifts, J. Am. Chem. Soc., 127, 17190–17191, https://doi.org/10.1021/ja0564259, 2005.

Kontaxis, G., Clore, G., and Bax, A.: Evaluation of Cross-Correlation Effects and Measurement of One-Bond Couplings in Proteins with Short Transverse Relaxation Times, J. Magn. Reson., 143, 184–196, https://doi.org/10.1006/jmre.1999.1979, 2000.

Kramer, F., Deshmukh, M., Kessler, H., and Glaser, S.: Residual dipolar coupling constants: An elementary derivation of key equations, Concepts Magn. Reson., 21A, 10–21, https://doi.org/10.1002/cmr.a.20003, 2004.

Lee, W., Tonelli, M., and Markley, J. L.: NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy, Bioinformatics, 31, 1325–1327, https://doi.org/10.1093/bioinformatics/btu830, 2014.

McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. M.: Presenting your structures: the *CCP4mg* molecular-graphics software, Acta Crystallographica Section D, 67, 386–394, https://doi.org/10.1107/S0907444911007281, 2011.

5   Orton, H. W. and Otting, G.: Accurate Electron–Nucleus Distances from Paramagnetic Relaxation Enhancements, J. Am. Chem. Soc., 140, 7688–7697, https://doi.org/10.1021/jacs.8b03858, 2018.

Orton, H. W., Kuprov, I., Loh, C.-T., and Otting, G.: Using Paramagnetism to Slow Down Nuclear Relaxation in Protein NMR, J. Phys. Chem. Lett., 7, 4815–4818, https://doi.org/10.1021/acs.jpclett.6b02417, 2016.

Pearce, B. J. G., Jabar, S., Loh, C.-T., Szabo, M., Graham, B., and Otting, G.: Structure restraints from heteronuclear pseudocontact shifts

10   generated by lanthanide tags at two different sites, J. Biomol. NMR, 68, 19–32, https://doi.org/10.1007/s10858-017-0111-z, 2017.

Pintacuda, G., Hohenthanner, K., Otting, G., and Müller, N.: Angular dependence of dipole-dipole-Curie-spin cross-correlation effects in high-spin and low-spin paramagnetic myoglobin, J. Biomol. NMR, 27, 18, 2003.

Pintacuda, G., Kaikkonen, A., and Otting, G.: Modulation of the distance dependence of paramagnetic relaxation enhancements by CSAxDSA cross-correlation, J. Magn. Reson., 171, 233–243, https://doi.org/10.1016/j.jmr.2004.08.019, 2004a.

15   Pintacuda, G., Keniry, M. A., Huber, T., Park, A. Y., Dixon, N. E., and Otting, G.: Fast Structure-Based Assignment of ^{15}N HSQC Spectra of Selectively ^{15}N-Labeled Paramagnetic Proteins, Journal of the American Chemical Society, 126, 2963–2970, https://doi.org/10.1021/ja039339m, 2004b.

Raman, S., Lange, O. F., Rossi, P., Tyka, M., Wang, X., Aramini, J., Liu, G., Ramelot, T. A., Eletsky, A., Szyperski, T., Kennedy, M. A., Prestegard, J., Montelione, G. T., and Baker, D.: NMR Structure Determination for Larger Proteins Using Backbone-Only Data, Science,

20   327, 1014–1018, https://doi.org/10.1126/science.1183649, 2010.

Rinaldelli, M., Carlon, A., Ravera, E., Parigi, G., and Luchinat, C.: FANTEN: a new web-based interface for the analysis of magnetic anisotropy-induced NMR data, J. Biomol. NMR, 61, 21–34, https://doi.org/10.1007/s10858-014-9877-4, 2015.

Schmitz, C., Stanton-Cook, M. J., Su, X.-C., Otting, G., and Huber, T.: Numbat: an interactive software tool for fitting $\Delta\chi$-tensors to molecular coordinates using pseudocontact shifts, J. Biomol. NMR, 41, 179–189, https://doi.org/10.1007/s10858-008-9249-z, 2008.

25   Schmitz, C., Vernon, R., Otting, G., Baker, D., and Huber, T.: Protein Structure Determination from Pseudocontact Shifts Using ROSETTA, J. Mol. Biol., 416, 668–677, https://doi.org/10.1016/j.jmb.2011.12.056, 2012.

Schrödinger: The PyMOL Molecular Graphics System, Version 1.8, 2015.

Solomon, I.: Relaxation processes in a system of two spins, Phys. Rev., 99, 559–565, 1955.

Stanton-Cook, M. J., Su, X.-C., Otting, G., and Huber, T.: PyParaTools, 2014.

30   Suturina, E. A., Mason, K., Geraldes, C. F. G. C., Chilton, N. F., Parker, D., and Kuprov, I.: Lanthanide-induced relaxation anisotropy, Phys. Chem. Chem. Phys., 20, 17 676–17 686, https://doi.org/10.1039/C8CP01332B, 2018a.

Suturina, E. A., Mason, K., Geraldes, C. F. G. C., Chilton, N. F., Parker, D., and Kuprov, I.: Lanthanide-induced relaxation anisotropy, Physical Chemistry Chemical Physics, 20, 17 676–17 686, https://doi.org/10.1039/C8CP01332B, 2018b.

Valafar, H. and Prestegard, J. H.: REDCAT: a residual dipolar coupling analysis tool, J. Magn. Reson., 167, 228–241,

35   https://doi.org/10.1016/j.jmr.2003.12.012, 2004.

Vega, A. J. and Fiat, D.: Nuclear relaxation processes of paramagnetic complexes The slow-motion case, Mol. Phys., 31, 347–355, https://doi.org/10.1080/00268977600100261, 1976.

Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D.: The CCPN data model for NMR spectroscopy: Development of a software pipeline, Proteins: Struct., Funct., Bioinf., 59, 687–696, https://doi.org/10.1002/prot.20449, 2005.

Zweckstetter, M. and Bax, A.: Prediction of Sterically Induced Alignment in a Dilute Liquid Crystalline Phase: Aid to Protein Structure
5   Determination by NMR, J. Am. Chem. Soc., 122, 3791–3792, https://doi.org/10.1021/ja0000908, 2000.