# Characterizing conformational ensembles of multi-domain proteins using anisotropic paramagnetic NMR restraints

Xue-Ni Hou and Hidehito Tochio

Department of Biophysics, Graduate School of Science, Kyoto University
Sakyo-ku, Kyoto, 606-8502, Japan

**Abstract**

It has been over two decades since paramagnetic NMR started to form part of the essential techniques for structural analysis of proteins under physiological conditions. Paramagnetic NMR has significantly expanded our understanding of the inherent flexibility of proteins, in particular, those that are formed by combinations of two or more domains. Here, we present a brief overview of techniques to characterize conformational ensembles of such multi-domain proteins using paramagnetic NMR restraints produced through anisotropic metals, with a focus on the basics of anisotropic paramagnetic effects, the general procedures of conformational ensemble reconstruction, and some representative reweighting approaches.

## Introduction

Multi-domain proteins, comprising more than one structurally well-folded domain connected by peptide linkers, widely exist in both prokaryotes and eukaryotes (Koonin et al. 2000; Chothia et al. 2003; Wriggers et al. 2005; Ekman et al. 2005). Characterizing the conformational states of these multi-domain proteins is critical to understanding their function in relevant biological events. Such structural characterization is nevertheless enormously challenging due to the flexibility of linkers as well as the complexity of inter-domain interactions (Reddy Chichili et al. 2013), even if the atomic resolution structures of individual domains are already available in the Protein Data Bank (PDB).

The orientation sensitive NMR measurements such as spin relaxation and residual dipolar couplings (RDCs) in alignment media, opened the way for determination of domain arrangement within a molecule. Works by Fushman et al. (2004), Ryabov and Fushman (2007), Walsh et al. (2010), Göbl et al. (2014) and Castañeda et al. (2016) have described these approaches in remarkable details. Thanks to the significant breakthroughs made recently in the site-specific incorporation of paramagnetic ions into proteins, e.g., substitution of metals in metalloproteins or attachment of a small tag coordinating a metal (Ravera et al. 2017; Nitsche and Otting 2017; Pell et al. 2019; Su and Chen 2019; Joss and Häussinger 2019; Softley et al. 2020), paramagnetic NMR has become one of the most attractive branches of biomolecular NMR, especially in the structural determination of multi-domain proteins,

due to the availability of long-range (~40Å) distance and/or orientation constraints.

Pseudocontact shifts (PCSs) and RDCs caused by paramagnetic self-alignment are two of the main valuable effects induced through anisotropic metals, e.g., cobalt, nickel, and most lanthanides. They can affect chemical shifts and coupling constants in standard NMR spectra, respectively (Bertini et al. 2005; Bertini et al. 2008). The values of PCSs and RDCs depend on the structural features and dynamics of proteins as well as the position of paramagnetic metals: PCSs are sensitive to metal-nucleus distance and orientation, and RDCs provide information on the orientation of internuclear vectors in the molecular frame. In addition to PCSs and RDCs, anisotropic metals also cause paramagnetic relaxation enhancements (PREs), which report on the distance between the metals and observed nuclei. At high magnetic fields and for molecules with a paramagnetic center, transverse relaxation is predominated by the Curie relaxation. This could be described by the isotropic magnetic susceptibility (Gueron 1975; Vega and Fiat 2006). Thus, PREs are preferably measured through isotropic metals, e.g., manganese and gadolinium. A number of applications of PREs in structural analysis are discussed thoroughly in Tang et al. (2007), Anthis et al. (2011), Liu et al. (2015), Chen et al. (2016), Wakamoto et al. (2019) and Lee et al. (2019). The use of PREs has also been extensively reviewed (Marius Clore and Iwahara 2009; Clore 2014). Here, we limit our discussion to two main anisotropic paramagnetic NMR restraints: PCSs and RDCs.

PCSs and RDCs are differently averaged depending on the interconversion rates over different conformational states (Fragai et al. 2013). The structural characterization then recovers a reasonable ensemble of such conformations with associated statistical weights, which should be consistent with averaged observables. Many approaches based on different algorithms for ensemble reconstruction have been proposed (Nodet et al. 2009; Bertini et al. 2010; Bertini et al. 2012; Berlin et al. 2013; Ihms and Foster 2015; Bonomi et al. 2017; Köfinger et al. 2019; Bottaro et al. 2020). However, reconstruction is inherently an underdetermined problem, as there are infinite combinations of structures that can potentially recapitulate the experimental data within certain uncertainties (Ravera et al. 2016; Medeiros Selegato et al. 2021). In other words, the dramatically distinct ensembles that represent experimental observables might be obtained by different approaches, or even by independent runs from the same approach. Therefore, discerning the similarities and differences in the different interpretations provided by these approaches is essential to selecting appropriate approaches to ensemble reconstruction.

In this review, we outline the basics, general procedures, and some approaches for ensemble reconstruction in multi-domain proteins by using anisotropic paramagnetic restraints. Here, we focus on the paramagnetic tags that can be attached to proteins with substantial rigidity (Yang et al. 2015;

Yang et al. 2016; Müntener et al. 2018; Lee et al. 2017; Pavlov et al. 2018; Joss and Häussinger 2019; Su and Chen 2019; Denis et al. 2020; Chen et al. 2020), for which the observables are mainly averaged by the internal dynamics of proteins other than the flexibility of paramagnetic tags.

**The anisotropic paramagnetic restraints**

PCSs and RDCs are two main anisotropic paramagnetic restraints, which can be observed through the same anisotropic metals. Their values depend on the positions of the observed nuclei in the frame of the magnetic susceptibility tensor ($\chi$ tensor) of the paramagnetic center, which are related to the conformational features of the systems investigated. The details of the $\chi$ tensor and its anisotropic component ($\Delta\chi$ tensor) can be found in several excellent reviews (Bertini et al. 2002; Otting 2010; Fragai et al. 2013; Nitsche and Otting 2017). It should be noted that, if conformation of the tag carrying the anisotropic metal is substantially mobile, PCSs and RDCs are severely affected. Hence, numerous studies have been dedicated to the tagging strategy. Shishmarev and Otting (2013), Hass et al. (2015), and Suturina and Kuprov (2016) are some of the excellent works that offer insightful details about the effect of the tag mobility on the paramagnetic effects.

PCSs arise from the spin-dipole interactions through space, and depend on the polar coordinates of nuclear spin $(r, \theta, \varphi)$ with respect to the $\Delta\chi$ tensor of metal and the axial ($\Delta\chi_{ax}$) and rhombic components ($\Delta\chi_{rh}$) of $\Delta\chi$ tensor (equation (1) and Figure 1a) (Bertini et al. 2002):

$$\Delta\delta^{pcs} = \frac{1}{12\pi r^3}\left[\Delta\chi_{ax}(3cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{rh}sin^2\theta cos2\varphi\right] \quad (1)$$

PCSs can be measured by comparing differences in chemical shifts (in ppm) of nuclei in biomolecules between paramagnetic and diamagnetic states (Figure 1c), in which a paramagnetic ion and a diamagnetic ion are rigidly bound to the molecule, respectively (John and Otting 2007). In the case that the perturbations in the structures and dynamics of molecules are insignificant with the incorporation of metal-chelating tags, i.e., the tag is peripherally bound to the surface of the molecules, the tag-free form of proteins may be regarded as a diamagnetic reference (Jensen and Led 2006; Otting 2010; Yang et al. 2016; Müntener et al. 2020).

In multi-domain systems, PCSs collected from the domain bearing the paramagnetic ion (the metal-bearing domain) can be used for the determination of the $\Delta\chi$ tensor and metal positions by fitting data against the available structure with firmly established programs, i.e., NUMBAT (John et al. 2005; Schmitz et al. 2008), FANTEN (Rinaldelli et al. 2015), and PARAMAGPY (Orton et al. 2020). There can be significant discrepancies between experimental and predicted PCSs if the structural model is not accurate. In this case, PCSs may be incorporated into structure refinement

tools, i.e., Xplor-NIH (Schwieters et al. 2003; Banci et al. 2004; Schwieters et al. 2006; Bermejo and Schwieters 2018) and CYANA (Banci et al. 1996; Bertini et al. 2001; Güntert 2004), as distance and angular constraints to obtain a refined model. If the structure of the metal-bearing domain had never been characterized in advance, the $\Delta\chi$ tensor and metal positions are concurrently determined during the structure calculation with Rosetta (Bowers et al. 2000; Schmitz et al. 2012; Yagi et al. 2013; Kuenze et al. 2019), where additional NMR restraints, such as the backbone dihedral-angle restraints, are ideally included. In addition, REFMAC5, part of the CCP4 suite, is a powerful tool for joint structural refinement in the combination of X-ray data and paramagnetic NMR restraints (Murshudov et al. 2011; Rinaldelli et al. 2014; Kovalevskiy et al. 2018; Carlon et al. 2019b).

In many cases, PCSs observed from the metal-free domain (i.e., the domain bearing no paramagnetic ion) in multi-domain proteins are smaller than those from the metal-bearing domain due to the larger metal-nucleus distance. However, if the motion of the metal-free domain relative to the other is completely rigid, the determined $\Delta\chi$ tensor and metal positions should be similar to those calculated by using PCSs collected from the metal-bearing domain. This similarity, however, decreases according to the existence and extent of domain rearrangement motions due to motional averaging (Chen et al. 2016). In systems with higher mobility, e.g., calmodulin (Bertini et al. 2004), PCSs from the domain without a paramagnetic metal failed to fit any single model, as nuclei in the domain have considerably fluctuating PCS values in the reference frame. Therefore, PCSs from the metal-free domain are, in general, only used for selecting the optimal ensemble by minimizing the discrimination between experimental and back-calculated PCS data.

RDCs result from partial alignment of observed molecules caused by the anisotropic magnetic susceptibility of the metals, and they provide information regarding the orientation of the internuclear vector in the reference frame (equation (2) and Figure 1b) (Banci et al. 1998).

$$\Delta v_{ij}^{rdc} = -\frac{1}{4\pi}\frac{{B_0}^2}{15kT}\frac{\gamma_i\gamma_j}{2\pi {r_{ij}}^3}h\left[\Delta\chi_{ax}(3cos^2\theta' - 1) + \frac{3}{2}\Delta\chi_{rh}sin^2\theta' cos2\varphi'\right] \quad (2)$$

where $\gamma_i$ and $\gamma_j$ denote the magnetogyric ratios of nuclear spin i and j, respectively, $r_{ij}$ is the interaction vector connecting the coupled nuclei i and j, $B_0$ is the strength of external magnetic field, $\kappa$ is Boltzmann constant, $T$ is the temperature, $h$ is Planck's constant, and $\theta'$ and $\varphi'$define the orientation of $r_{ij}$ in the frame.

Paramagnetic-induced RDCs are typically small; thus, observables are often acquired for covalently bonded nuclei, e.g., backbone $^{1}$H-$^{15}$N, by comparing the coupling constants of partially aligned (paramagnetic states) and unaligned molecules (diamagnetic states) in the same solution

through IPAP-HSQC experiments (Figure 1d) (Yao et al. 2009; Ottiger et al. 1998).

In contrast to PCSs, RDCs collected from each domain can be represented by an effective tensor, as they are independent of metal-nucleus distance (Bertini et al. 2004). RDCs from the metal-bearing domain can be analyzed together with PCSs through fitting against the available structure model with the programs FANTEN (Rinaldelli et al. 2015) and PARAMAGPY (Orton et al. 2020). However, the PCSs-derived and RDCs-derived $\Delta\chi$ tensors are somewhat different for two reasons. First, they are averaged differently (Shishmarev and Otting 2013). That is, RDCs can always be described by a single averaged effective $\Delta\chi$ tensor, independent of the metal position (Figure 1(b)). On the other hand, PCSs depend not only on the $\Delta\chi$ tensor but also on the metal-nuclei distances (Figure 1(a)). This difference can cause discrepancies in the experimentally derived tensors when conformational mobility of the paramagnetic center exists. Second, the structures registered in PDB may not be sufficiently accurate in the orientation of some internuclear bond vectors, even if they are determined in similar solution conditions using NMR spectroscopy. Since RDCs are more sensitive than PCSs to such bond vector orientations, the tensors obtained from both can also differ. This structural accuracy can be improved by the joint PCSs/RDCs refinement procedure included in some of the programs mentioned above. RDC and PCS data from mobile residues should not be incorporated into the structural analysis, which can be judged through NMR relaxation data, including longitudinal relaxation rate ($R_1$), transverse relaxation rate ($R_2$), heteronuclear NOE (hnNOE) (Bertini et al. 2009). To be excluded, the selection criteria for the residues may differ between RDCs and PCSs because of the different effects of local motions.

Due to the domain mobility, RDCs from the metal-free domain might be smaller than those from the metal-bearing domain, since the aligning force produced by the metal and transmitted to the domain would be weaker in such a case. The extent of motions could be easily estimated by the ratio of magnitudes of effective tensors calculated from the two domains (Carlon et al. 2016). The ratio is approximately close to 1 in the complete absence of domain motion.

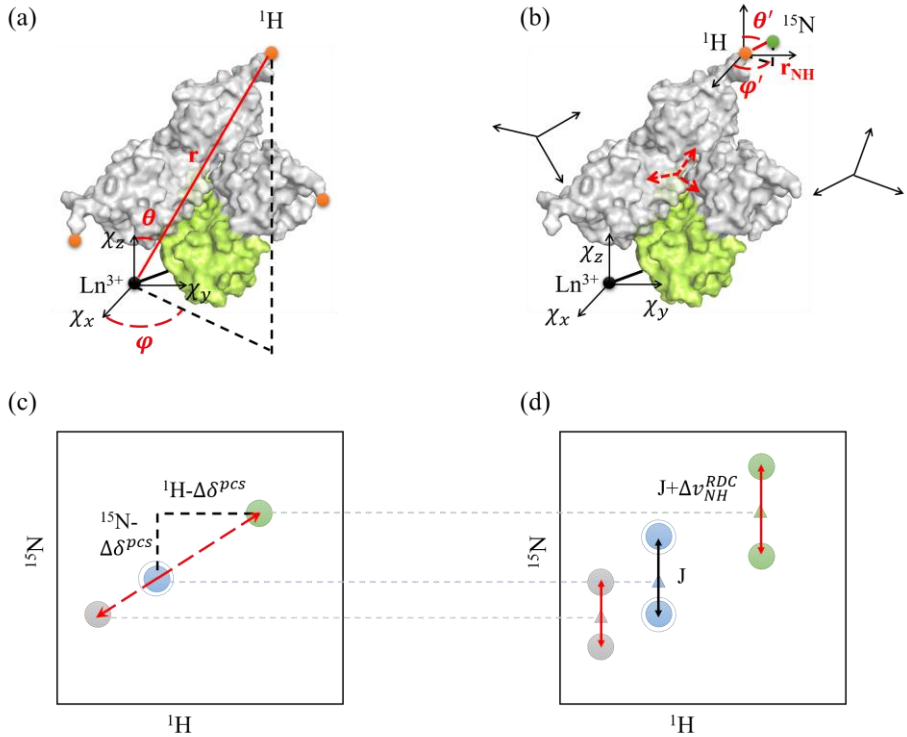


Figure 1. Schematic representation of structural information in a two-domain protein derived from PCSs and RDCs. (a) PCSs are described by equation (1); the observed nucleus ($^{1}$H, orange circle) of the metal-free domain in three different arrangements (surface, gray) have different PCSs in the fame of the $\Delta\chi$ tensor of the metal ($Ln^{3+}$, circle, black); PCSs from the metal-bearing domain (surface, limon) can be used to determine the $\Delta\chi$ tensor; (b) RDCs are described by equation (2); the averaged effective tensor (red frame in the center) from the metal-free domain depends on the exchange rate of all conformations (here, three are depicted); (c, d) illustration of paramagnetic effects PCSs (c) and RDCs(d) in 2D $^{1}$H-$^{15}$N correlation NMR spectra; for large biomolecules, there might be some overlaid peaks in diamagnetic NMR spectra (concentric circles, blue), while in paramagnetic states, they can be separated due to different metal-nucleus distances (solid circles, green and gray).

**Obtaining conformational states from paramagnetic data**

Once the domain mobility is assessed in the system, PCSs and RDCs can be simultaneously used for characterizing the conformational space sampled by domain rearrangements. This is tackled according to a procedure applied through the following steps (Figure 2):

(a) Refine the structures of the individual domains and assemble the entire molecule.

(b) Generate a conformational pool with a sufficient number of structures.

(c) Calculate the magnetic susceptibility parameter of the paramagnetic metal for all the structures considered in (b) by using PCSs and RDCs from the metal-bearing domains.

(d) Back-calculate the PCSs and RDCs for the metal-free domains of all the structures considered in (b) by using the corresponding magnetic susceptibility parameters obtained in (c).

(e) Reconstruct the conformational ensembles satisfying the experimental PCSs and RDCs.

In step (a), the accurate structure of the metal-bearing domain, especially with refined orientation of internuclear vectors related to RDCs, is obtained by using PCSs and RDCs from this domain. In addition, orientations of internuclear vectors in the metal-free domain are adjusted by using RDCs from this domain. Some programs have been implemented in structural refinement by using PCSs and/or RDCs, starting without or with available structures from X-ray crystallography or NMR (For the use of Xplor-NIH, see Ref. (Banci et al. 2004; Bertini et al. 2009); for the use of CYANA, see Ref. (Banci et al. 1996; Bertini et al. 2001); for the use of Rosetta, see Ref. (Schmitz et al. 2012; Kuenze et al. 2019), and for the use of REFMAC5, see Ref. (Rinaldelli et al. 2014; Carlon et al. 2019b)). Once the structure of each domain is obtained, the AIDA program provides fast docking for domain assembly (Xu et al. 2015; Xu et al. 2014). Some programs, e.g., pyDockTET (Cheng et al. 2008) and Rosetta (Wollacott et al. 2007), are also able to assemble structures of isolated domains into a multi-domain molecule. If the structure of the entire protein is available, replacing each domain with a refined structure is an alternative.

Generation of the conformational pool described in step (b) is typically achieved by treating all domains as rigid domains and randomizing the backbone torsion angles of several residues in the linkers. The unbiased and sufficient sampling from entire conformational space is essential to reliable ensemble reconstruction. All generated conformations should be located in the topologically allowed space and maintain chain connectivity. The programs RanCh (Bernadó et al. 2007; Tria et al. 2015) and PDB Generator module of MESMER (Ihms and Foster 2015) are the two simplest tools for pool generation in the case that residues comprising the linkers are placed neither in alpha helices nor in beta sheets (Bernadó et al. 2005). Otherwise, the quasi-Ramachandran space of residues in those secondary structures should be carefully considered, such as in the native-like model of program RanCh (Bernadó and Svergun 2012). Furthermore, the utilization of Molecular Dynamics (MD) is helpful in sampling more physically realistic set of conformations. If available, the incorporation of complementary high-resolution constraints, e.g., data from NMR, small-angle X-ray scattering (SAXS), or dipolar electron-electron resonance (DEER), is preferable (Ihms and Foster 2015). In order to simplify the further calculation and visualization of conformational ensemble, the metal-bearing domain is considered fixed in the reference frame in this step.

Next, in step (c), PCSs and RDCs observed from the metal-bearing domain are fitted by a single set of effective parameters, $\Delta\chi$ tensors and the metal position (only for PCSs). Several programs have been developed for such calculations. e.g., PALES (Zweckstetter 2008), PATI (Berlin et al. 2009), REDCAT (Valafar and Prestegard 2004) and MODULE (Dosset et al. 2001) for RDC analysis, and NUMBAT (John et al. 2005; Schmitz et al. 2008) for PCS analysis. The program FANTEN (Rinaldelli et al. 2015) and PARAMAGPY (Orton et al. 2020) can fit $\Delta\chi$ tensors and metal coordinates to the atomic coordinates of biomolecules by using PCSs and RDCs, respectively. These parameters can be subsequently used in step (d) to predict PCSs and RDCs for the metal-free domains of all conformations in the generated pool. For this purpose, PyParaTools python library (http://comp-bio.anu.edu.au/mscook/PPT/) is a powerful tool that has been involved in some reweighting programs (see next section for details), such as MESMER (Ihms and Foster 2015). The above outlines the basic processes of data preparation in ensemble reconstruction with any reweighting approaches.

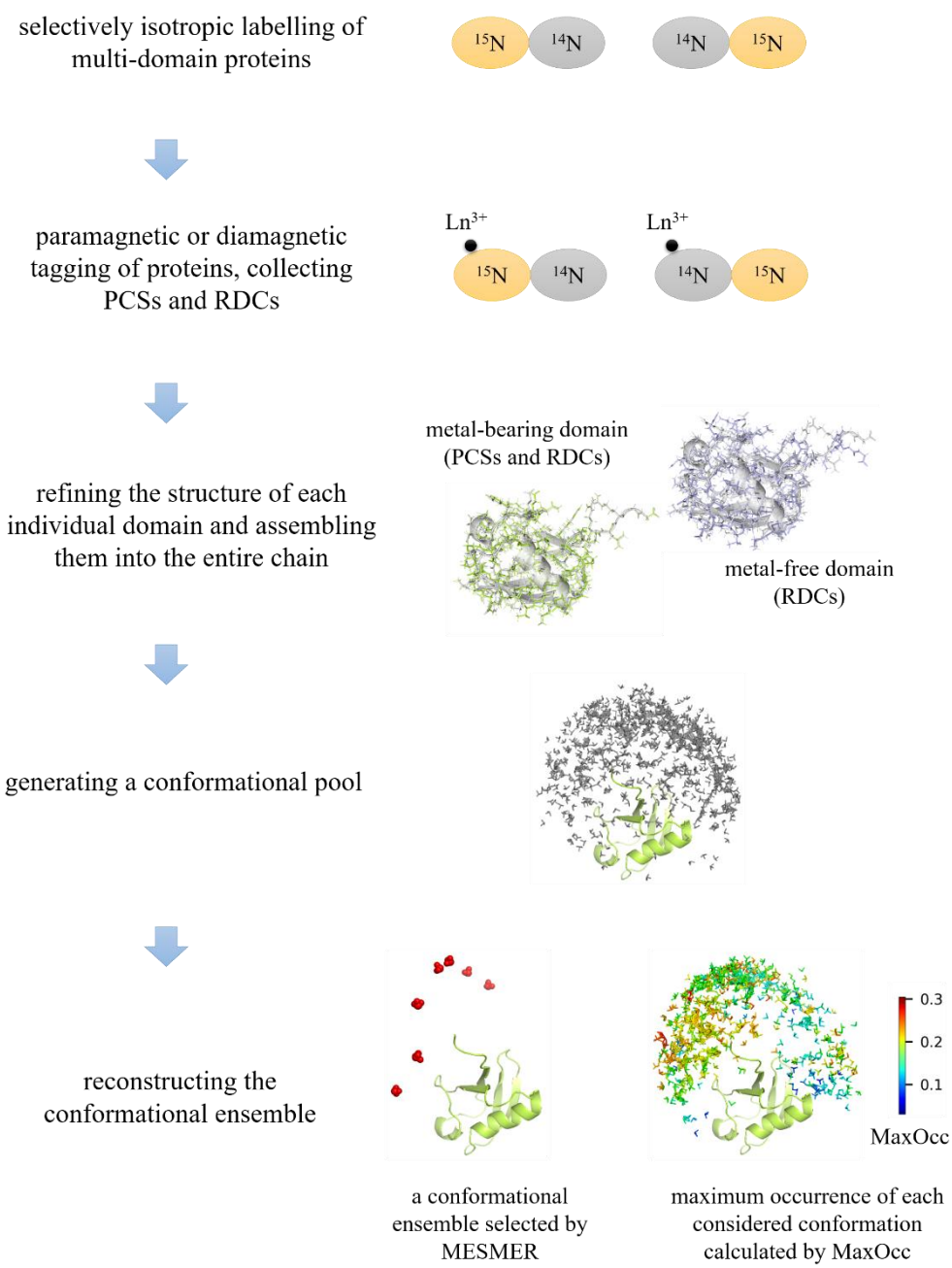


Figure 2 Overview of ensemble reconstruction in a two-domain protein, linear diubiquitin ($Ub_2$), by using PCSs and RDCs: to enable unambiguous chemical shift assignments, the N- or C-terminal Ub was selectively enriched with $^{15}$N-nuclei; the structure of each individual Ub was refined by using paramagnetic data with Xplor-NIH, and the structure model of linear $Ub_2$ was obtained by assembling two Ubs with AIDA; the conformational pool was generated by MESMER; and two approaches were employed for the conformational ensemble reconstruction: MESMER selected a minimal ensemble comprising seven conformations with associated weights that are proportional to their putative population, while MaxOcc depicted the conformational distributions. Almost all conformations selected by MESMER had higher MaxOcc values. Adapt from (Hou et al. 2021). Selected programs available for ensemble reconstruction from PCSs and RDCs are listed in Table 1.

Table 1. Selected programs available for capturing structural information from PCSs and RDCs induced through anisotropic metals.

| Programs | Description | Reference |
|---|---|---|
| NUMBAT | determination of anisotropy tensors from PCSs for a given 3D molecular structure | (Schmitz et al. 2008) |
| FANTEN | determination of anisotropy tensors related to PCSs and RDCs for a given 3D molecular structure | (Rinaldelli et al. 2015) |
| PARAMAGPY | determination of magnetic susceptibility tensors related to PCSs, RDCs, paramagnetic relaxation enhancements, and cross-correlated relaxation data | (Orton et al. 2020) |
| PCS-ROSETTA | determination of 3D structures using PCSs as the only restraints; estimation of anisotropy tensors from PCSs during the structure prediction | (Schmitz et al. 2012) |
| RosettaNMR | determination of 3D structures using paramagnetic data (PCSs, RDCs, and PREs) with CSs and NOEs; determination of magnetic susceptibility tensors | (Kuenze et al. 2019) |
| Xplor-NIH | refinement of protein structures based on available structures having the smallest change to fulfil paramagnetic restraints | (Banci et al. 2004; Bertini et al. 2009) |
| CYANA | | (Güntert 2004) |
| REFMAC5 | refining protein structure simultaneously using X-ray crystallographic data and paramagnetic data | (Rinaldelli et al. 2014) |
| AIDA | modelling and assembling multi-domain proteins | (Xu et al. 2015; Xu et al. 2014) |
| pyDockTET | | (Cheng et al. 2008) |
| RanCh | generating a conformational pool | (Bernadó and Svergun 2012) |
| MESMER | identifying conformational ensembles to simultaneously fulfil multiple observables | (Ihms and Foster 2015) |
| SES | | (Berlin et al. 2013; Andrałojć et al. 2015) |
| MaxOcc/MaxOR | calculating maximum occurrence of any considered conformations or combinations | (Bertini et al. 2010; Andrałojć et al. 2016; Gigli et al. 2018) |

**Reweighting approaches for ensemble reconstruction**

In this section, we review some approaches for conformational ensemble reconstruction based on reweighting (step (e) in the previous section). Reweighting means that the experimental data are used as a *posterior* to optimize the weights of conformations in a pre-calculated unbiased ensemble with the aim of minimizing the discrimination of experimental and back-calculated data. It is distinct from "restraining" approaches, in which the additional energy terms, as functions of experimental data, are directly incorporated into classical MD force fields during the simulations to generate and analyze possible conformational states (Roux and Weare 2013). Such approaches are preferable for structural characterization of molecules with extensive conformational heterogeneity, such as intrinsically disordered proteins (Bonomi et al. 2017). The broader understanding of this subject may be found in other existing works (Boomsma et al. 2014; Ravera et al. 2016; Bonomi et al. 2017; Rangan et al. 2018; Cárdenas et al. 2020). Here, we focus on reweighting approaches, in which PCS and RDC restraints could be incorporated.

Reweighting approaches have the goal of either finding an optimal ensemble with the minimal subset of conformations or calculating the maximum allowed probability (MAP) or the maximum occurrence (MaxOcc) of all considered structures. The former can be achieved with several software packages. First, we selected minimal ensemble solutions for the multiple experimental restraints (MESMER) approach because it is user-friendly and available for nearly any type of observable (Ihms and Foster 2015).

MESMER has been developed for identifying and selecting ensembles that can simultaneously fulfil multiple experimental data, e.g., SAXS, paramagnetic NMR, DEER, etc. In simultaneous fitting, the relative scale for each experimental dataset is typically pre-set to the inverse of the average fitness obtained from individual fits. The optimal ensemble is iteratively selected via a genetic algorithm with the following steps:

(1) $K$ types of predicted data (PCSs and RDCs, etc.) for each structure in the conformational pool ($Z$ structures) are calculated as described in the previous section, and they are compiled into a series of components ($Z$) as input.

(2) A "parent" ensemble pool, comprised of $N$ ensembles with $M$ components, is generated randomly from $Z$ components. This pool is then duplicated to form a "child" ensemble pool with some diversifications through different replacement mechanisms.

(3) $2N$ ensembles generated in (2) are ranked according to their consistency with experimental data.

(4) *N* best-fitting ensembles are selected and used as the "parent" ensemble pool for the next generation.

Steps (2), (3) and (4) are iterated until even the poorest-fitting ensemble in step (4) has a reasonable agreement with all *K* types of experimental data, or the residual standard deviation (RSD) is less than the pre-set value. Solutions with the total sum of weight in one ensemble slightly less or more than 1 are acceptable.

MESMER has been applied to characterize the conformational states of some biomolecules, e.g., calmodulin (Ihms and Foster 2015), PDZ domains (Delhommel et al. 2017) and linear $Ub_2$ (Hou et al. 2021), by using paramagnetic data. It provides a graphical user interface to streamline all processes from the generation of conformational pool to the visualization of selected best ensembles as well as representation of correlation between experimental data and predicted data. However, users should be cautious when simultaneously fitting multiple datasets obtained from different experiments, as they might contain significantly distinct information. The inappropriately pre-set scales for those input datasets would lead to severe overestimation or underestimation of conformational diversity, as the weight for components involved in ensembles that fit one dataset well would easily increase during the iteration. Thus, the relative scales for each dataset need to be carefully tuned. This problem is distinct even when using multiple sets of PCSs collected from protein with incorporated metals at different positions (Hou et al. 2021). In addition, the solution is not deterministic. In essence, it is hard to obtain one single specific solution (ensemble) by executing MESMER repeatedly. Even so, major components (conformers) in *N* non-unique ensembles should be similar in any meaningful executions.

The Sparse Ensemble Selection (SES) method is an alternative global-fit approach, which can recover a representative conformational ensemble from multiple experimental datasets. Distinct from MESMER, SES requires no problem-specific tuning parameters and provides a deterministic solution, as only the structures comprised of the ensemble, which fits the experimental observables best, are cloned and subsequently complemented during each generation (Berlin et al. 2013; Ravera et al. 2016). In addition, it is important to stress that certain ensembles selected by these approaches only provide possible solutions to recapitulate the experimental data, which are sensitive to certain aspects of conformational sampling. No ensembles can be considered unique and represent the full and complete conformational states in practice (Ozenne et al. 2012).

Rather than finding a solution (the best ensemble) to fit the datasets, some approaches try to calculate reasonable existence probabilities of any considered structures in the ensemble. MAP

(Longinetti et al. 2006; Bertini et al. 2007), MaxOcc (Bertini et al. 2010; Bertini et al. 2012), the maximum and minimum occurrence of defined regions, MaxOR and MinOR (Andrałojć et al. 2014; Andrałojć et al. 2016) have been sequentially developed with the aim of estimating the maximum probabilities of single conformation or regions comprised of multiple conformations. The MAP approach utilizes free domain movements (Ravera et al. 2016), while the others calculate corresponding values of conformations in a pre-defined pool. Here, we focus on the MaxOcc, as MaxOR and MinOR are natural extensions of the MaxOcc approach for estimation of specific combinations of conformations.

MaxOcc analysis can be performed by the MATLAB script (Andrałojć et al. 2016; Gigli et al. 2018). In this case, some further data preparation should be performed in advance. $K$ types of experimental data should be concatenated into a length-$J$ column vector representing all observables. Then it is essential to normalize this experimental vector through some methods, i.e., dividing the data by the square of their sum (Medeiros Selegato et al. 2021). The $K$ types of predicted data from $Z$ structures should also be concatenated into a $J \times Z$ prediction matrix and normalized using the same procedure.

Then, the fitting calculation is repeated by increasing the weight of a single conformation until there is no acceptable ensemble comprising this conformation with one certain weight to explain the experimental data reasonably. This certain weight is defined as the MaxOcc of this conformation. Namely, MaxOcc is the maximum allowed weight for a conformation in any possible ensembles, which does not violate the experimental data. If a conformation has a higher MaxOcc value, this implies that it is more likely to be visited by molecules due to their intrinsic dynamics in solution. However, it should be noted that in principle, even if a conformation has higher-MaxOcc, it can be absent in reality, as MaxOcc doesn't guarantee minimum occurrence. In this sense, one may rather safely exclude the existence of lower-MaxOcc conformations in the ensemble.

MaxOcc has been used to evaluate the conformational states on the basis of paramagnetic data, SAXS data, or DEER data in many systems, e.g., calmodulin (Bertini et al. 2010), MMP1(Cerofolini et al. 2013), transactivation response element (TAR) RNA from the HIV-1 virus (Andrałojć et al. 2016), the capsid of human immunodeficiency virus type 1 (Carlon et al. 2019a), and linear $Ub_2$ (Hou et al. 2021). One excellent work reported a direct relation between the results from SES and MaxOcc: SES are more likely to select conformations with highest-MaxOcc values (Medeiros Selegato et al. 2021).

As stated above, MESMER and SES provide an optimal ensemble of several discrete conformations with associated statistical weights that are proportional to their population for recapitulating the experimental data. It is unrealistic that those conformations truly exist with the calculated weights in solution. Rather, they can be viewed as representative snapshots taken from a large continuum of states; each of them represents a group of similar conformations, namely a certain region of the conformational space. Degree of conformational variability manifested in the obtained ensemble gives us functional insights into the molecule. As an example, we have captured a very compact conformer of free linear $Ub_2$ by performing a MESMER analysis, which proposed that the dynamics of linear $Ub_2$ is even more complicated than previously considered (Hou et al. 2021). In this study, MaxOcc further allowed us to discern which conformations or conformational regions (MaxOR) are more likely to be visited by the proteins, having expanded our insight into their intrinsic dynamics. In addition, existence of any structure considered can be evaluated by MaxOcc, which is helpful for probing particular aspects of the conformational fluctuations. For example, the introduction of two distinct mutations, E16Rp or E18Rp (where "p" represents the mutation introduced at C-terminal (proximal) Ub), in linear $Ub_2$ drastically changed the PCSs, strongly implying the perturbation of conformational space. Notably, the rank of MaxOcc of a target-bound conformation changed significantly among the whole structural pool. Such perturbed conformational sampling of free linear $Ub_2$ was shown to correlate with binding affinities for Ub-binding proteins, HOIL-1L-NZF (Figure 3), thus providing a profound insight into the binding mode of linear $Ub_2$ (Hou et al. 2021).

In the linear $Ub_2$ example described above, both the high flexibility of the linker and the weak interaction between domains result in a certain degree of domain movement continuity among several conformations with similar energies. Such a situation probably presents one of the most difficult cases for reconstructing conformational ensembles based on paramagnetic restraints. If a protein takes only a few stable conformations, then domain motion can be regarded as an exchange between these states. In such a case, the methods reviewed in this paper are likely to work more robustly. This is because the complexity of PCS and RDC data is greatly reduced when the number of stable conformational states is small.

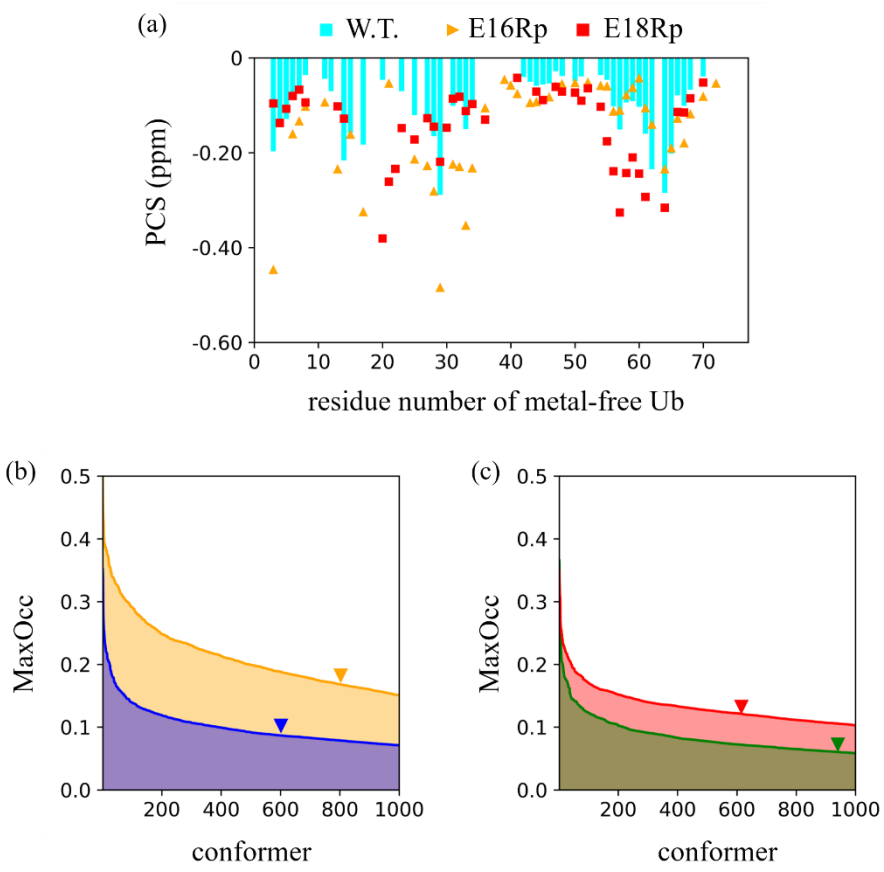


Figure 3 (a) An example of dramatically changed PCSs by introduction of mutations in free linear $Ub_2$: PCSs collected from the metal-free Ub of wild-type (cyan bar), E16Rp (orange triangle), and E18Rp (red square) of linear $Ub_2$, with paramagnetic tag (PSPy-6M-DO3MA-$Tm^{3+}$)(Yang et al. 2016) at D39C of N-terminal (distal) Ub. (b, c) MaxOcc of top 1,000 conformers calculated using PCSs from wild-type (purple or green), E16Rp (orange), and E18Rp (red) of linear $Ub_2$, are sorted and plotted in descending order. The rank of MaxOcc of the bound conformer of linear $Ub_2$ in complex with HOIL-1L-NZF (PDB code: 3b0a) (Sato et al. 2011), shown as triangles by corresponding color, dropped (elevated) in E16Rp (E18Rp), implying the E16Rp (E18Rp) decreased (increased) the probability for free linear $Ub_2$ to adapt the bound state. Correspondingly, dissociation constant ($K_d$) increased by over 14-fold for E16Rp and decreased by fivefold for E18Rp, implying contribution from the conformational selection mechanism. Adapted from (Hou et al. 2021).

## Conclusions

To conclude, paramagnetic effects can be a rich source of structural restraints for characterizing the conformational states of multi-domain proteins. The anisotropic paramagnetic restraints are very sensitive to metal-nuclear distances (PCSs) and relative orientations of metal-nuclear (PCSs) or internuclear vectors (RDCs). Therefore, they can identify the conformations that are more likely to visit due to their intrinsic dynamics, thus providing a structural insight into their physiological behaviors. A number of programs have been developed for determination of magnetic anisotropy susceptibility tensors, refinement of solution structure, assembly of multiple individual domains, generation of conformational pools, prediction of paramagnetic data for other nuclei in biomolecules, and reconstruction of conformational ensembles. In combination with these programs, visualization of populated conformational space is possible from averaged paramagnetic experimental data. Certainly, in the near future, the use of paramagnetic NMR restraints induced through anisotropic paramagnetic ions would be a powerful approach for visualizing the dynamic behavior of proteins under physiological conditions, or even in living cells.

## Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

## References


A. Suturina E, Ilya Kuprov (2016) Pseudocontact shifts from mobile spin labels. Phys Chem Chem Phys 18:26412–26422. https://doi.org/10.1039/C6CP05437D

Andrałojć W, Berlin K, Fushman D, et al (2015) Information content of long-range NMR data for the characterization of conformational heterogeneity. J Biomol NMR 62:353–371. https://doi.org/10.1007/s10858-015-9951-6

Andrałojć W, Luchinat C, Parigi G, Ravera E (2014) Exploring Regions of Conformational Space Occupied by Two-Domain Proteins. J Phys Chem B 118:10576–10587. https://doi.org/10.1021/JP504820W

Andrałojć W, Ravera E, Salmon L, et al (2016) Inter-helical conformational preferences of HIV-1 TAR-RNA from maximum occurrence analysis of NMR data and molecular dynamics simulations. Phys Chem Chem Phys 18:5743–5752. https://doi.org/10.1039/c5cp03993b

Anthis NJ, Doucleff M, Clore GM (2011) Transient, Sparsely Populated Compact States of Apo and Calcium-Loaded Calmodulin Probed by Paramagnetic Relaxation Enhancement: Interplay of Conformational Selection and Induced Fit. J Am Chem Soc 133:18966–18974. https://doi.org/10.1021/JA2082813

Banci L, Bertini I, Bren KL, et al (1996) The use of pseudocontact shifts to refine solution structures of paramagnetic metalloproteins: Met80Ala cyano-cytochrome c as an example. J Biol Inorg Chem 1:117–126. https://doi.org/10.1007/s007750050030

Banci L, Bertini I, Cavallaro G, et al (2004) Paramagnetism-based restraints for Xplor-NIH. J Biomol NMR 28:249–261. https://doi.org/10.1023/B:JNMR.0000013703.30623.f7

Banci L, Bertini I, Huber JG, et al (1998) Partial orientation of oxidized and reduced cytochrome b5 at high magnetic fields: Magnetic susceptibility anisotropy contributions and consequences for protein solution structure determination. J Am Chem Soc 120:12903–12909. https://doi.org/10.1021/ja981791w

Berlin K, Castañeda CA, Schneidman-Duhovny D, et al (2013) Recovering a representative conformational ensemble from underdetermined macromolecular structural data. J Am Chem Soc 135:16595–16609. https://doi.org/10.1021/ja4083717

Berlin K, O'Leary DP, Fushman D (2009) Improvement and analysis of computational methods for prediction of residual dipolar couplings. J Magn Reson 201:25–33. https://doi.org/10.1016/j.jmr.2009.07.028

Bermejo GA, Schwieters CD (2018) Protein structure elucidation from NMR data with the program Xplor-NIH. In: Methods in Molecular Biology. Humana Press Inc., pp 311–340

Bernadó P, Blanchard L, Timmins P, et al (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. Proc Natl Acad Sci U S A 102:17002–17007. https://doi.org/10.1073/pnas.0506202102

Bernadó P, Mylonas E, Petoukhov M V., et al (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc 129:5656–5664. https://doi.org/10.1021/ja069124n

Bernadó P, Svergun DI (2012) Analysis of intrinsically disordered proteins by small-angle X-ray scattering. Methods Mol Biol 896:107–122. https://doi.org/10.1007/978-1-4614-3704-8_7

Bertini I, Del Blanco C, Gelis I, et al (2004) Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. Proc Natl Acad Sci U S A 101:6841–6846. https://doi.org/10.1073/pnas.0308641101

Bertini I, Donaire A, Jiménez B, et al (2001) Paramagnetism-based versus classical constraints: An analysis of the solution structure of Ca Ln calbindin D9k. J Biomol NMR 21:85–98. https://doi.org/10.1023/A:1012422402545

Bertini I, Ferella L, Luchinat C, et al (2012) MaxOcc: A web portal for maximum occurrence analysis. J Biomol NMR 53:271–280. https://doi.org/10.1007/s10858-012-9638-1

Bertini I, Giachetti A, Luchinat C, et al (2010) Conformational space of flexible biological macromolecules from average data. J Am Chem Soc 132:13553–13558. https://doi.org/10.1021/ja1063923

Bertini I, Gupta YK, Luchinat C, et al (2007) Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. J Am Chem Soc 129:12786–12794. https://doi.org/10.1021/ja0726613

Bertini I, Kursula P, Luchinat C, et al (2009) Accurate solution structures of proteins from X-ray data and a minimal set of NMR Data: Calmodulin-peptide complexes as examples. J Am Chem Soc 131:5134–5144. https://doi.org/10.1021/ja8080764

Bertini I, Luchinat C, Parigi G (2002) Magnetic susceptibility in paramagnetic NMR. Prog Nucl Magn Reson Spectrosc 40:249–273. https://doi.org/10.1016/S0079-6565(02)00002-X

Bertini I, Luchinat C, Parigi G, Pierattelli R (2005) NMR spectroscopy of paramagnetic metalloproteins. ChemBioChem 6:1536–1549

Bertini I, Luchinat C, Parigi G, Pierattelli R (2008) Perspectives in paramagnetic NMR of metalloproteins. Dalt Trans 3782–3790. https://doi.org/10.1039/b719526e

Bonomi M, Heller GT, Camilloni C, Vendruscolo M (2017) Principles of protein structural ensemble determination. Curr. Opin. Struct. Biol. 42:106–116

Boomsma W, Ferkinghoff-Borg J, Lindorff-Larsen K (2014) Combining Experiments and Simulations Using the Maximum Entropy Principle. PLOS Comput Biol 10:e1003406. https://doi.org/10.1371/JOURNAL.PCBI.1003406

Bottaro S, Bengtsen T, Lindorff-Larsen K (2020) Integrating Molecular Simulation and Experimental Data: A Bayesian/Maximum Entropy Reweighting Approach. In: Methods in Molecular Biology. Humana Press Inc., pp 219–240

Bowers PM, Strauss CEM, Baker D (2000) De novo protein structure determination using sparse NMR data. J Biomol NMR 18:311–318. https://doi.org/10.1023/A:1026744431105

Cárdenas R, Martínez-Seoane J, Amero C (2020) Combining Experimental Data and Computational Methods for the Non-Computer Specialist. Mol 2020, Vol 25, Page 4783 25:4783. https://doi.org/10.3390/MOLECULES25204783

Carlon A, Gigli L, Ravera E, et al (2019a) Assessing Structural Preferences of Unstructured Protein Regions by NMR. Biophys J 117:1948–1953. https://doi.org/10.1016/J.BPJ.2019.10.008

Carlon A, Ravera E, Andrałojć W, et al (2016) How to tackle protein structural data from solution and solid state: An integrated approach. Prog. Nucl. Magn. Reson. Spectrosc. 92–93:54–70

Carlon A, Ravera E, Parigi G, et al (2019b) Joint X-ray/NMR structure refinement of multidomain/multisubunit systems. J Biomol NMR 73:265–278. https://doi.org/10.1007/s10858-018-0212-3

Castañeda CA, Chaturvedi A, Camara CM, et al (2016) Linkage-specific conformational ensembles of non-canonical polyubiquitin chains. Phys Chem Chem Phys 18:5771–5788. https://doi.org/10.1039/c5cp04601g

Cerofolini L, Fields GB, Fragai M, et al (2013) Examination of Matrix Metalloproteinase-1 in

Solution: A PREFERENCE FOR THE PRE-COLLAGENOLYSIS STATE *. J Biol Chem 288:30659–30671. https://doi.org/10.1074/JBC.M113.477240

Chen JL, Li B, Li XY, Su XC (2020) Dynamic Exchange of the Metal Chelating Moiety: A Key Factor in Determining the Rigidity of Protein-Tag Conjugates in Paramagnetic NMR. J Phys Chem Lett 31:9493–9500. https://doi.org/10.1021/acs.jpclett.0c02196

Chen JL, Yang Y, Zhang LL, et al (2016) Analysis of the solution conformations of T4 lysozyme by paramagnetic NMR spectroscopy. Phys Chem Chem Phys 18:5850–5859. https://doi.org/10.1039/c5cp07196h

Cheng TMK, Blundell TL, Fernandez-Recio J (2008) Structural assembly of two-domain proteins by rigid-body docking. BMC Bioinformatics 9:1–13. https://doi.org/10.1186/1471-2105-9-441

Chothia C, Gough J, Vogel C, Teichmann SA (2003) Evolution of the protein repertoire. Science (80-. ). 300:1701–1703

Clore GM (2014) Interplay between conformational selection and induced fit in multidomain protein–ligand binding probed by paramagnetic relaxation enhancement. Biophys Chem 186:3–12. https://doi.org/10.1016/J.BPC.2013.08.006

Delhommel F, Cordier F, Bardiaux B, et al (2017) Structural Characterization of Whirlin Reveals an Unexpected and Dynamic Supramodule Conformation of Its PDZ Tandem. Structure 25:1645-1656.e5. https://doi.org/10.1016/j.str.2017.08.013

Denis M, Softley C, Giuntini S, et al (2020) The Photocatalyzed Thiol-ene reaction: A New Tag to Yield Fast, Selective and reversible Paramagnetic Tagging of Proteins. ChemPhysChem 21:863–869. https://doi.org/10.1002/cphc.202000071

Dosset P, Hus JC, Marion D, Blackledge M (2001) A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. J Biomol NMR 20:223–231. https://doi.org/10.1023/A:1011206132740

Ekman D, Björklund ÅK, Frey-Skött J, Elofsson A (2005) Multi-domain Proteins in the Three Kingdoms of Life: Orphan Domains and Other Unassigned Regions. J Mol Biol 348:231–243. https://doi.org/10.1016/J.JMB.2005.02.007

Fragai M, Luchinat C, Parigi G, Ravera E (2013) Conformational freedom of metalloproteins revealed by paramagnetism-assisted NMR. Coord. Chem. Rev. 257:2652–2667

Fushman D, Varadan R, Assfalg M, Walker O (2004) Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar coupling measurements. Prog Nucl Magn Reson Spectrosc 3–4:189–214. https://doi.org/10.1016/J.PNMRS.2004.02.001

Gigli L, Andrałojć W, Dalaloyan A, et al (2018) Assessing protein conformational landscapes: IntegraGigli L, Andrałojć W, Dalaloyan A, et al (2018) Assessing protein conformational landscapes: Integration of DEER data in Maximum Occurrence analysis. Phys Chem Chem Phys 20:27429–27438. https://doi.org/1. Phys Chem Chem Phys 20:27429–27438. https://doi.org/10.1039/c8cp06195e

Göbl C, Madl T, Simon B, Sattler M (2014) NMR approaches for structural analysis of multidomain proteins and complexes in solution. Prog Nucl Magn Reson Spectrosc 80:26–63. https://doi.org/10.1016/J.PNMRS.2014.05.003

Gueron M (1975) 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

Güntert P (2004) Automated NMR structure calculation with CYANA. Methods Mol Biol 278:353–378. https://doi.org/10.1385/1-59259-809-9:353

Hass MAS, Liu W-M, Agafonov R V., et al (2015) A minor conformation of a lanthanide tag on adenylate kinase characterized by paramagnetic relaxation dispersion NMR spectroscopy. J Biomol NMR 2015 612 61:123–136. https://doi.org/10.1007/S10858-014-9894-3

Hou X-N, Sekiyama N, Ohtani Y, et al (2021) Conformational Space Sampled by Domain Reorientation of Linear Diubiquitin Reflected in Its Binding Mode for Target Proteins. ChemPhysChem 22:1505–1517. https://doi.org/10.1002/CPHC.202100187

Ihms EC, Foster MP (2015) MESMER: minimal ensemble solutions to multiple experimental restraints. Bioinformatics 31:1951–1958. https://doi.org/10.1093/bioinformatics/btv079

Jensen MR, Led JJ (2006) Metal-protein interactions: Structure information from Ni 2+-induced pseudocontact shifts in a native nonmetalloprotein. Biochemistry 45:8782–8787. https://doi.org/10.1021/bi0604431

John M, Otting G (2007) Strategies for measurements of pseudocontact shifts in protein NMR spectroscopy. ChemPhysChem 8:2309–2313

John M, Park AY, Pintacuda G, et al (2005) 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

Joss D, Häussinger D (2019) Design and applications of lanthanide chelating tags for pseudocontact shift NMR spectroscopy with biomacromolecules. Prog. Nucl. Magn. Reson. Spectrosc. 114–115:284–312

Köfinger J, Stelzl LS, Reuter K, et al (2019) Efficient Ensemble Refinement by Reweighting. J Chem Theory Comput 15:3390–3401. https://doi.org/10.1021/acs.jctc.8b01231

Koonin E V., Aravind L, Kondrashov AS (2000) The impact of comparative genomics on our understanding of evolution. Cell 101:573–576

Kovalevskiy O, Nicholls RA, Long F, et al (2018) Overview of refinement procedures within REFMAC 5: Utilizing data from different sources. Acta Crystallogr Sect D Struct Biol 74:215–227. https://doi.org/10.1107/S2059798318000979

Kuenze G, Bonneau R, Leman JK, Meiler J (2019) Integrative Protein Modeling in RosettaNMR from Sparse Paramagnetic Restraints. Structure 27:1721-1734.e5. https://doi.org/10.1016/j.str.2019.08.012

Lee K-Y, Fang Z, Enomoto M, et al (2020) Two Distinct Structures of Membrane-Associated Homodimers of GTP- and GDP-Bound KRAS4B Revealed by Paramagnetic Relaxation Enhancement. Angew Chemie 132:11130–11138. https://doi.org/10.1002/ANGE.202001758

Lee MD, Dennis ML, Graham B, Swarbrick JD (2017) Short two-armed lanthanide-binding tags for paramagnetic NMR spectroscopy based on chiral 1,4,7,10-tetrakis(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane scaffolds. Chem Commun 53:13205–13208. https://doi.org/10.1039/c7cc07961c

Liu Z, Dong X, Yi H-W, et al (2019) Structural basis for the recognition of K48-linked Ub chain by proteasomal receptor Rpn13. Cell Discov 2019 51 5:1–15. https://doi.org/10.1038/s41421-019-0089-7

Liu Z, Gong Z, Jiang WX, et al (2015) Lys63-linked ubiquitin chain adopts multiple conformational states for specific target recognition. Elife 4:e05767. https://doi.org/10.7554/eLife.05767.001

Longinetti M, Luchinat C, Parigi G, Sgheri L (2006) Efficient determination of the most favoured orientations of protein domains from paramagnetic NMR data. Inverse Probl 22:1485–1502. https://doi.org/10.1088/0266-5611/22/4/019

Marius Clore G, Iwahara J (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. Chem Rev 109:4108–4139. https://doi.org/10.1021/cr900033p

Medeiros Selegato D, Bracco C, Giannelli C, et al (2021) Comparison of Different Reweighting Approaches for the Calculation of Conformational Variability of Macromolecules from Molecular Simulations. ChemPhysChem 22:127–138. https://doi.org/10.1002/cphc.202000714

Müntener T, Böhm R, Atz K, et al (2020) NMR pseudocontact shifts in a symmetric protein homotrimer. J Biomol NMR 74:413–419. https://doi.org/10.1007/s10858-020-00329-7

Müntener T, Kottelat J, Huber A, Häussinger D (2018) New Lanthanide Chelating Tags for PCS NMR Spectroscopy with Reduction Stable, Rigid Linkers for Fast and Irreversible Conjugation to Proteins. Bioconjug Chem 29:3344–3351. https://doi.org/10.1021/acs.bioconjchem.8b00512

Murshudov GN, Skubák P, Lebedev AA, et al (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr Sect D Biol Crystallogr 67:355–367. https://doi.org/10.1107/S0907444911001314

Nitsche C, Otting G (2017) Pseudocontact shifts in biomolecular NMR using paramagnetic metal tags. Prog. Nucl. Magn. Reson. Spectrosc. 98–99:20–49

Nodet G, Salmon L, Ozenne V, et al (2009) Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings. J Am Chem Soc 131:17908–17918. https://doi.org/10.1021/ja9069024

Orton HW, Huber T, Otting G (2020) Paramagpy: software for fitting magnetic susceptibility tensors using paramagnetic effects measured in NMR spectra. Magn Reson 1:1–12. https://doi.org/10.5194/mr-1-1-2020

Ottiger M, Delaglio F, Bax A (1998) Measurement of J and Dipolar Couplings from Simplified Two-Dimensional NMR Spectra. J. Magn. Reson. 131:373–378

Otting G (2010) Protein NMR Using Paramagnetic Ions. Annu Rev Biophys 39:387–405. https://doi.org/10.1146/annurev.biophys.093008.131321

Ozenne V, Bauer F, Salmon L, et al (2012) Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. Bioinformatics 28:1463–1470. https://doi.org/10.1093/BIOINFORMATICS/BTS172

Pavlov AA, Savkina SA, Belov AS, et al (2018) Very Large Magnetic Anisotropy of Cage Cobalt(II) Complexes with a Rigid Cholesteryl Substituent from Paramagnetic NMR Spectroscopy. ACS Omega 3:4941–4946. https://doi.org/10.1021/acsomega.8b00772

Pell AJ, Pintacuda G, Grey CP (2019) Paramagnetic NMR in solution and the solid state. Prog. Nucl. Magn. Reson. Spectrosc. 111:1–271

Rangan R, Bonomi M, Heller GT, et al (2018) Determination of Structural Ensembles of Proteins: Restraining vs Reweighting. J Chem Theory Comput 14:6632–6641. https://doi.org/10.1021/acs.jctc.8b00738

Ravera E, Parigi G, Luchinat C (2017) Perspectives on paramagnetic NMR from a life sciences infrastructure. J Magn Reson 282:154–169. https://doi.org/10.1016/j.jmr.2017.07.013

Ravera E, Sgheri L, Parigi G, Luchinat C (2016) A critical assessment of methods to recover information from averaged data. Phys. Chem. Chem. Phys. 18:5686–5701

Reddy Chichili VP, Kumar V, Sivaraman J (2013) Linkers in the structural biology of protein-protein

interactions. Protein Sci. 22:153–167
Rinaldelli M, Carlon A, Ravera E, et al (2015) 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
Rinaldelli M, Ravera E, Calderone V, et al (2014) Simultaneous use of solution NMR and X-ray data in REFMAC5 for joint refinement/detection of structural differences. Acta Crystallogr Sect D Biol Crystallogr 70:958–967. https://doi.org/10.1107/S1399004713034160
Roux B, Weare J (2013) On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. J Chem Phys 138:84107. https://doi.org/10.1063/1.4792208
Ryabov Y, Fushman D (2007) Structural Assembly of Multidomain Proteins and Protein Complexes Guided by the Overall Rotational Diffusion Tensor. J Am Chem Soc 129:7894–7902. https://doi.org/10.1021/JA071185D
Sato Y, Fujita H, Yoshikawa A, et al (2011) Specific recognition of linear ubiquitin chains by the Npl4 zinc finger (NZF) domain of the HOIL-1L subunit of the linear ubiquitin chain assembly complex. Proc Natl Acad Sci U S A 108:20520–20525. https://doi.org/10.1073/pnas.1109088108
Schmitz C, Stanton-Cook MJ, Su XC, et al (2008) Numbat: An interactive software tool for fitting δχ-tensors to molecular coordinates using pseudocontact shifts. J Biomol NMR 41:179–189. https://doi.org/10.1007/s10858-008-9249-z
Schmitz C, Vernon R, Otting G, et al (2012) Protein structure determination from pseudocontact shifts using ROSETTA. J Mol Biol 416:668–677. https://doi.org/10.1016/j.jmb.2011.12.056
Schwieters CD, Kuszewski JJ, Marius Clore G (2006) Using Xplor-NIH for NMR molecular structure determination. Prog. Nucl. Magn. Reson. Spectrosc. 48:47–62
Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160:65–73. https://doi.org/10.1016/S1090-7807(02)00014-9
Shishmarev D, Otting G (2013) How reliable are pseudocontact shifts induced in proteins and ligands by mobile paramagnetic metal tags? A modelling study. J Biomol NMR 56:203–216. https://doi.org/10.1007/s10858-013-9738-6
Softley CA, Bostock MJ, Popowicz GM, Sattler M (2020) Paramagnetic NMR in drug discovery. J Biomol NMR 74:287–309. https://doi.org/10.1007/s10858-020-00322-0
Su XC, Chen JL (2019) Site-Specific Tagging of Proteins with Paramagnetic Ions for Determination of Protein Structures in Solution and in Cells. Acc Chem Res 52:1675–1686. https://doi.org/10.1021/acs.accounts.9b00132
Tang C, Schwieters CD, Clore GM (2007) Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. Nat 2007 4497165 449:1078–1082. https://doi.org/10.1038/nature06232
Tria G, Mertens HDT, Kachala M, Svergun DI (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. IUCrJ 2:207–217. https://doi.org/10.1107/S205225251500202X
Valafar H, Prestegard JH (2004) REDCAT: A residual dipolar coupling analysis tool. J Magn Reson 167:228–241. https://doi.org/10.1016/j.jmr.2003.12.012
Vega AJ, Fiat D (2006) Nuclear relaxation processes of paramagnetic complexes The slow-motion case. http://dx.doi.org/101080/00268977600100261 31:347–355.

https://doi.org/10.1080/00268977600100261
Wakamoto T, Ikeya T, Kitazawa S, et al (2019) Paramagnetic relaxation enhancement-assisted structural characterization of a partially disordered conformation of ubiquitin. Protein Sci 28:1993–2003. https://doi.org/10.1002/PRO.3734
Walsh JD, Meier K, Ishima R, Gronenborn AM (2010) NMR Studies on Domain Diffusion and Alignment in Modular GB1 Repeats. Biophys J 99:2636–2646. https://doi.org/10.1016/J.BPJ.2010.08.036
Wollacott AM, Zanghellini A, Murphy P, Baker D (2007) Prediction of structures of multidomain proteins from structures of the individual domains. Protein Sci 16:165–175. https://doi.org/10.1110/ps.062270707
Wriggers W, Chakravarty S, Jennings PA (2005) Control of protein functional dynamics by peptide linkers. Biopolym. - Pept. Sci. Sect. 80:736–746
Xu D, Jaroszewski L, Li Z, Godzik A (2015) AIDA: Ab initio domain assembly for automated multi-domain protein structure prediction and domain-domain interaction prediction. Bioinformatics 31:2098–2105. https://doi.org/10.1093/bioinformatics/btv092
Xu D, Jaroszewski L, Li Z, Godzik A (2014) AIDA: Ab initio domain assembly server. Nucleic Acids Res 42:W308–W313. https://doi.org/10.1093/nar/gku369
Yagi H, Pilla KB, Maleckis A, et al (2013) Three-dimensional protein fold determination from backbone amide pseudocontact shifts generated by lanthanide tags at multiple sites. Structure 21:883–890. https://doi.org/10.1016/j.str.2013.04.001
Yang F, Wang X, Pan B Bin, Su XC (2016) Single-armed phenylsulfonated pyridine derivative of DOTA is rigid and stable paramagnetic tag in protein analysis. Chem Commun 52:11535–11538. https://doi.org/10.1039/c6cc06114a
Yang Y, Wang JT, Pei YY, Su XC (2015) Site-specific tagging proteins via a rigid, stable and short thiolether tether for paramagnetic spectroscopic analysis. Chem Commun 51:2824–2827. https://doi.org/10.1039/c4cc08493d
Yao L, Ying J, Bax A (2009) Improved accuracy of 15N-1H scalar and residual dipolar couplings from gradient-enhanced IPAP-HSQC experiments on protonated proteins. J Biomol NMR 43:161–170. https://doi.org/10.1007/s10858-009-9299-x
Zweckstetter M (2008) NMR: Prediction of molecular alignment from structure using the PALES software. Nat Protoc 3:679–690. https://doi.org/10.1038/nprot.2008.36