ChemEurJ
Research Article
doi.org/10.1002/chem.202502647
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

Characterising Side Chain Conformations in Proteins Using ^{19}F-ENDOR and New Fluorinated Amino Acid Spin Labels

Martyna Judd,*[a, b] Elwy H. Abdelkader,[c] Haocheng Qianzhu,[a] Edan Habel,[a]
Caitlin Lindsay,[a] Anthony Brancewicz,[a] Thomas Huber,[a] Gottfried Otting,[c]
and Nicholas Cox*[a]

It has recently been shown that distances > 20 Å can be measured between Gd^{3+} spin tags and ^{19}F labels in proteins using ^{19}F-ENDOR at 94 GHz. Here we examine the precision with which this method can locate the positions of (aromatic) amino acid side chains and, with the help of established simulation tools, determine the conformational space sampled by solvent-exposed chains in solution. First, using a novel set of fluorinated phenylalanine amino acids incorporated into the metalloprotein Calbindin D_{9k} binding Gd^{3+}, we show that triangulation of the corresponding ^{19}F-ENDOR determined distances can precisely identify the conformation of buried side chains. The obtained conformation agrees with the same (single) orientation seen in the crystal structure of the native protein. In a second set of proteins labeled with a Gd^{3+} spin tag and noncanonical ^{19}F-labeled amino acids, the splittings, lineshapes, and integrated intensities of the ^{19}F-ENDOR signals were used to constrain the conformational space of the side chains initially identified by comprehensive rotamer simulations. This work provides the first example where ^{19}F-ENDOR constraints have been used to accurately pinpoint side chain conformations in a protein with the help of differently fluorinated amino acids.

1. Introduction

The 3D structure of a protein is critically influenced by the flexibility and packing of its amino acid side chains. These structural features are fundamental to key biological processes such as ligand recognition, protein–protein interactions, and enzymatic activity. Therefore, accurately determining side chain conformations under native conditions is essential for structure-activity studies and rational drug design.[1–3]

Traditional methods like X-ray crystallography and cryo-electron microscopy (cryo-EM) provide high-resolution structural data, but typically require high protein concentrations and the use of precipitants or detergents, which create conditions deviating from physiological relevance. In contrast, nuclear magnetic resonance (NMR) spectroscopy allows the study of proteins in solution under near-physiological conditions. However, the low sensitivity and natural abundance of many NMR-active nuclei often necessitates isotope enrichment.

The conformational space of flexible amino acid side chains has long been explored by computational techniques.[4] Most straightforwardly, the conformational space can be defined by the use of rotamer libraries, which use systematic bond rotations to define the range of possible side chain conformations. These libraries can further be informed by statistical analyses of crystal structures and refined to exclude conformations that cause steric clashes or are energetically unfavorable.[5] However, the experimental verification of these refined libraries is still challenging.

A promising new experimental method in protein structure analysis is high-field (>1.2 T, > 34 GHz) electron nuclear double resonance (ENDOR) spectroscopy combined with ^{19}F labelling. This technique detects the magnetic coupling between a ^{19}F nucleus and a paramagnetic center, such as a nitroxide or trityl radical,[6–9] or a metal ion featuring a long electron spin relaxation time. The present study used Gd^{3+} due to its demonstrated high sensitivity in ENDOR experiments.[10–14]

Electron–nuclear distances (r), can be determined from the magnitude of the ENDOR signal splitting, which is centered about the nuclear Larmor frequency, as well as from the lineshape and intensity of the signal.[10–12,15,16] When the electron and nuclear spins are separated by more than ~3 Å, the magnitude of the ENDOR splitting (T) follows a $1/r^3$ dependence, while the signal intensity scales with $1/r^6$. Compared to NMR, ^{19}F-ENDOR offers significantly higher intrinsic sensitivity, making it well-suited for low-concentration samples. It also imposes no upper limit on the protein size that can be

---

[a] M. Judd, H. Qianzhu, E. Habel, C. Lindsay, A. Brancewicz, T. Huber, N. Cox
Research School of Chemistry, The Australian National University, Canberra, ACT 2601, Australia
E-mail: martyna.judd@northwestern.edu
nick.cox@anu.edu.au

[b] M. Judd
Department of Chemistry, Weinberg College of Arts and Sciences, Northwestern University, Evanston, IL 60208, USA

[c] E. H. Abdelkader, G. Otting
ARC Centre of Excellence for Innovation in Peptide and Protein Science, The Australian National University, Canberra, ACT 2601, Australia

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202502647

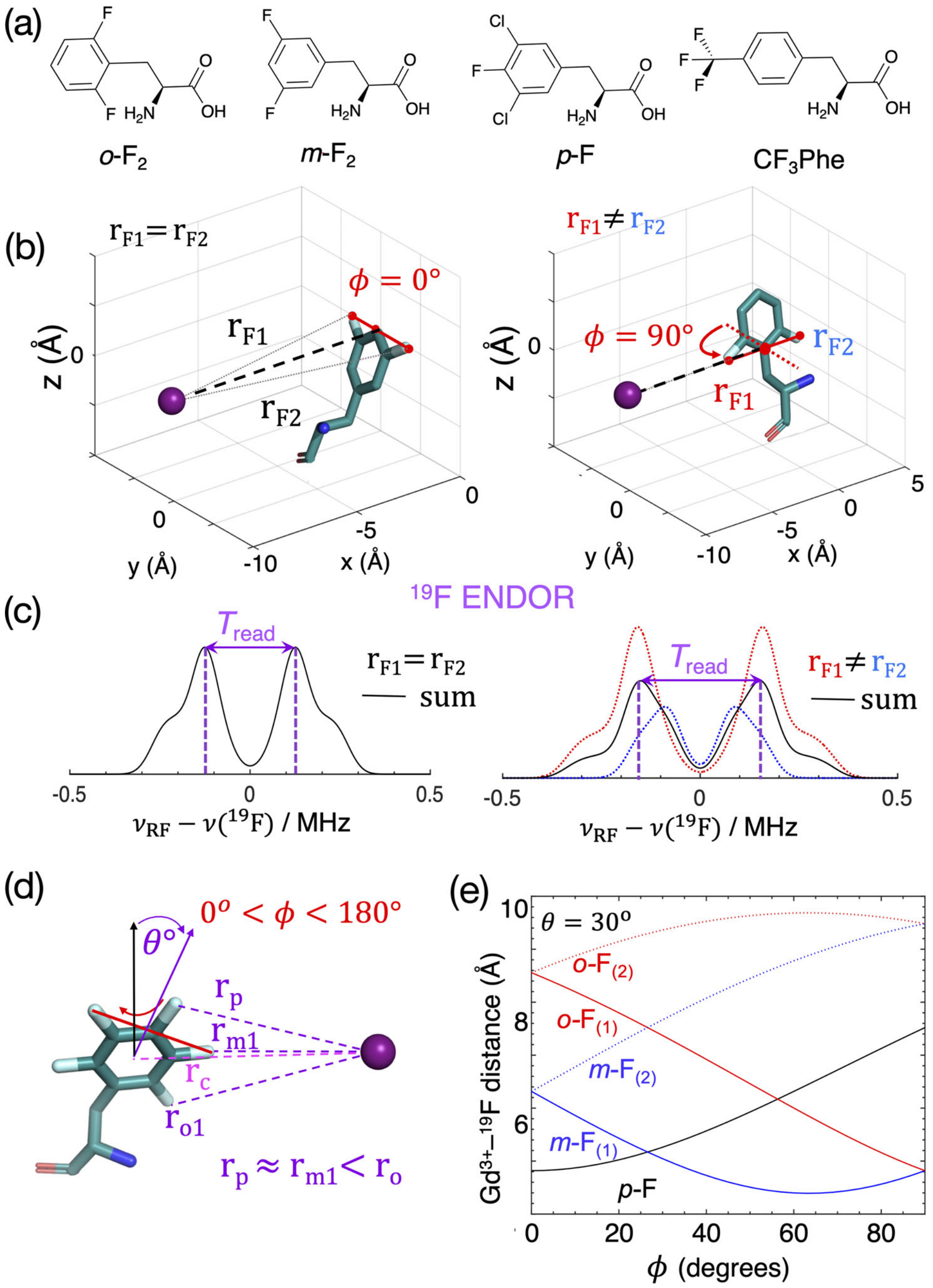

Figure 1. Parameterization of Phe ring geometry and distance to the Gd^{3+} ion used in the ^{19}F-ENDOR simulations. a) Chemical structures of fluorinated phenylalanine analogues used in the present work. b) Definition of the ring rotation angle ϕ in the xy plane. c) A change of ϕ changes the $^{19}F - Gd^{3+}$ distances of the two ^{19}F spins in o-F_2 and m-F_2 thus changing the ENDOR lineshape. The maximum difference $\Delta r_F = 4.7$ Å occurs for $\phi = 90°$ (assuming the case $\theta = 0°$), while the two r_F distances are equal for $\phi = 0°$. d) Definition of the Phe ring tilt angle θ and the effect that different combinations of θ and ϕ have on the relative ^{19}F–Gd^{3+} distances for the p-, o-, and m- positions of the ring. A positive θ corresponds to the phenyl ring facing towards the Gd^{3+} ion, while negative θ values correspond to the ring facing away. e) $^{19}F - Gd^{3+}$ distances for p-F, m-F_2, and o-F_2 for a fixed angle $\theta = 30°$ as a function of ϕ.

studied. These advantages make ^{19}F-ENDOR a powerful tool for probing the structure of biological systems under native-like conditions.

The present work shows that ^{19}F-ENDOR (at W-band, 94 GHz), when combined with novel fluorinated aromatic labels and strategic spin labelling, can be extended to precisely determine protein side chain conformations. This approach enables three new avenues of structural analysis:

1. Determining aromatic ring orientations of buried residues relative to a paramagnetic site, independent of a crystal structure.
2. Refining rotamer simulations using geometric constraints obtained from ^{19}F-ENDOR to determine the side chain dihedral angles (χ) of buried (i.e., fixed) residues.
3. Filtering rotamer ensembles to define the conformational space of solvent-exposed side chains, which have the

potential to exist in multiple conformations, providing experimental insight into dynamic regions.

2. Results and Discussion

2.1. Determining Well-Constrained Geometries in Calbindin D_{9k} at Short Inter-Spin Distances with ^{19}F-ENDOR

As our first test system, we used the metalloprotein calbindin D_{9k} (C_{D9k}, PDB: 4ICB),[17] introducing a paramagnetic center by replacing the native Ca^{2+} ion near residue 77 with a Gd^{3+} ion.[10] The resulting Gd^{3+} site is rigid and well-defined.[18] Previously, we used trifluoromethylphenylalanine (CF_3Phe) to replace Phe50 and Phe66 for ^{19}F-ENDOR measurements.[10] In the present work, we instead introduced three different fluorinated analogues of phenylalanine—4-F-3,5-Cl_2Phe (*p*-F), 3,5-F_2Phe (*m*-F_2), and 2,6-F_2Phe (*o*-F_2)[19]–(Figure 1a), which we term pseudocanonical amino acids (pcAAs) to reflect their close resemblance to native Phe. Phe50 and Phe66 are buried residues with limited flexibility,[20] making the pcAA-labelled C_{D9k}–Gd^{3+} complex ideal for benchmarking ^{19}F-ENDOR-derived structural restraints.

Figure 1 summarizes how geometric parameters can be extracted from ^{19}F-ENDOR data by exploiting the geometry of the pcAAs. First, the magnitude of the electron-nuclear dipolar coupling, observed as a resolved splitting (T_{read}) in the ENDOR spectrum, reports the ^{19}F–Gd^{3+} distance (r_F). Ring orientation is encoded in the lineshape of the spectrum. Since ENDOR is measured in frozen solution (10 K), the signal presents a solid-state Pake pattern. Accurate distance determination requires simulating the full ENDOR lineshape. For the singly fluorinated *p*-F, the spectrum reflects a single r_F value. For the difluorinated *o*-F_2 and *m*-F_2 residues, the lineshape of the overall Pake pattern is a superposition of two distinct r_F distances (Figure 1b, c). Importantly, the difference between the two distances—and thus the shape of the combined spectrum—is a consequence of both the θ (tilt) and ϕ (azimuthal) Euler angles (Figure 1d), describing the orientation of the Phe ring with respect to the r_F vector. By combining the experimental data of all three pcAAs and fitting these simultaneously, we can determine not only the ring angles θ and ϕ, but also the distance r_c between the ring centre and the Gd^{3+} ion, assuming that the different fluorination patterns do not change the side chain conformation.

It is instructive to compare the ENDOR-derived parameter set {r_c, θ, and ϕ} to rotamer simulations. In rotamer simulations, the side chain dihedral angles χ_1 and χ_2 (Figure 2a) are varied independently. Figure 2b shows unconstrained rotamer simulations for Phe50 in C_{D9k}, plotting χ_1 versus χ_2, with the color scale indicating the r_c distance, which strongly depends on χ_1. Comparing Figures 2a and 2b, we see that the rotamers cluster around staggered χ_1 angles of 180° (*t*), -60° (*g*-), and + 60° (*g*+). χ_1 is much more tightly restrained by the ENDOR measurements than χ_2. This is expected given the symmetry of the fluorinated phenyl rings of the pcAAs about the χ_2 ring-flip axis. While $\chi_2 \approx 90°$ minimizes steric clashes between ring substituents and the $C^{\beta}H_2$ group,[21] other χ_2 angles are also possible in proteins.

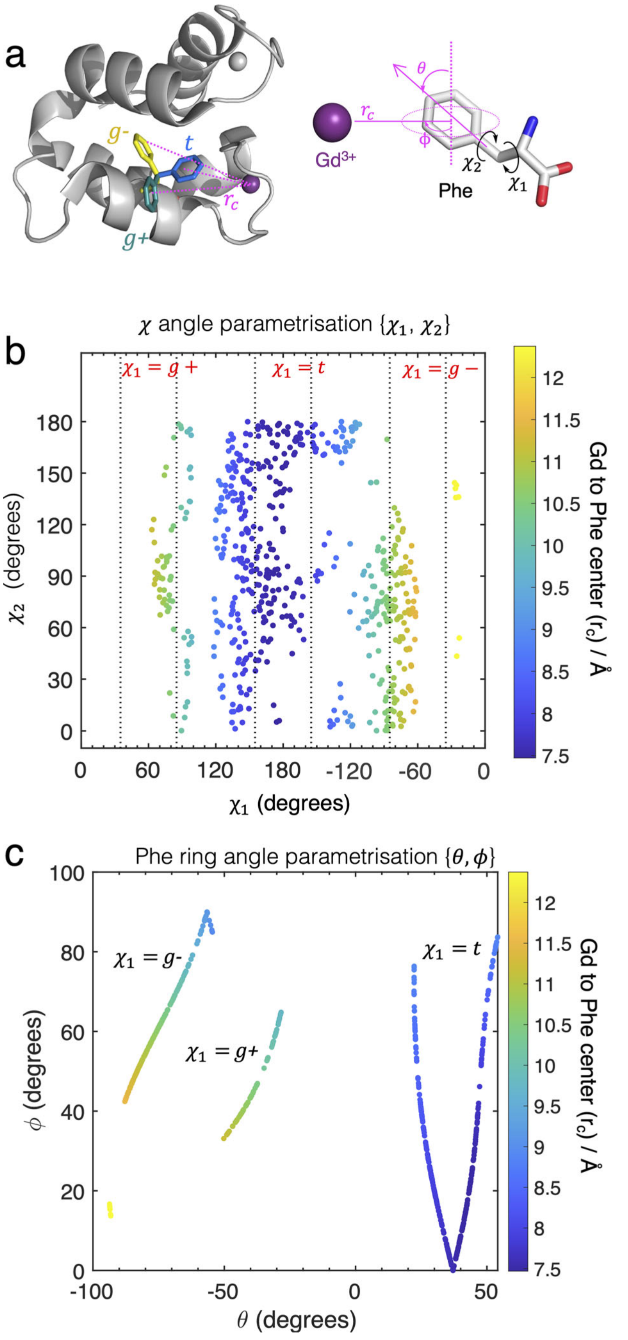

Figure 2. Rotamer analysis of Phe50 in C_{D9k}. a) Crystal structure of C_{D9k} with models of the $\chi_1 = t, g+$ and *g*- conformations of Phe50. The right-hand side defines the geometric parameters χ_1, χ_2, and θ, ϕ, and r_c, the latter set describing the orientation of the Phe ring relative to the Gd^{3+} centre. b) Rotamer library of Phe50 C_{D9k} showing χ_1 versus χ_2 distributions along with a colour bar indicating the distances r_c associated with each rotamer. c) The same rotamer library plotted in terms of the Phe ring geometry parameters θ and ϕ, with the colour bar indicating r_c.

The T_{read} values measured for the variants labeled with *p*-F, *m*-F_2, or *o*-F_2 (Figure 3a) provided initial estimates of the most highly populated distances: 7.1 ± 0.3 Å (*p*-F), 7.4 ± 0.3 Å (*m*-F_2), and 8.9 ± 0.6 Å (*o*-F_2). For the difluorinated pcAAs, T_{read} reflects the closer one of the two ^{19}F spins, as the ENDOR signal intensity scales with $1/r^6$. The observed distance trend r_F(*p*-F)

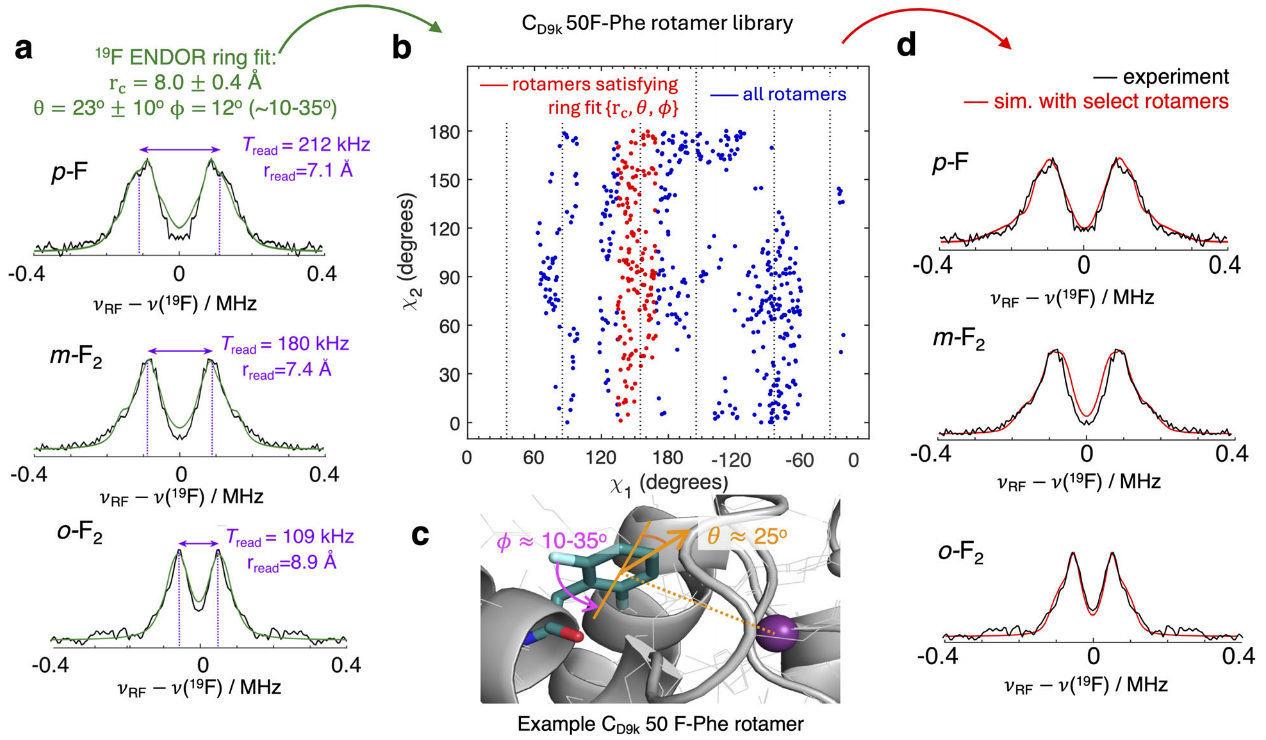

Figure 3. a) Experimental ^{19}F-ENDOR data for C_{D9k} labelled with *p*-F, *m*-F$_2$, and *o*-F$_2$ in position 50 (black traces) together with an *EasySpin* simulation of the ENDOR lineshapes fitted simultaneously to a common parameter set {θ, ϕ, and r$_c$}, (green traces). The T_{read} splittings are estimated as the separation between the ENDOR peak maxima, which correspond to the perpendicular components of the electron-nuclear dipolar coupling tensor. For the spectrum of *p*-F where the peaks look structured, the mean peak positions were taken. b) Rotamer library presented in terms of a χ_1 versus χ_2 plot. Red points represent rotamers satisfying the {θ, ϕ, and r$_c$} parameters from (a). c) Example rotamer in C_{D9k} crystal structure satisfying the ENDOR fit parameters. d) Experimental data (black trace) compared to ENDOR simulations (red) using filtered set of rotamers as input.

$\approx r_F$(*m*-F$_2$) $< r_F$(*o*-F$_2$), indicates that the Phe ring is tilted toward the Gd^{3+} ion (i.e., with $\theta > 0°$). This narrows the possible ring orientations to 27° $< \theta <$ 69°, 5.4° $< | \phi | <$ 34°, and 7.5 Å $< r_c <$ 9.4 Å (see Supporting Information 3.1). These approximate ranges were used to constrain simulations of the ENDOR lineshapes for all three C_{D9k}-50 variants, performed using the *saffron* module in *EasySpin*.[22] By varying {r_c, θ, ϕ} within these bounds, the best-fit simulation yielded r_c = 8.0 ± 0.4 Å, θ = 23° ± 10°, and ϕ = 17° ± 17° (Figure 3b, see Supporting Information for more detail). Simulations within these uncertainty ranges reproduced the experimental spectra across all three variants. Notably, the orientation of the Phe ring determined by ^{19}F-ENDOR closely matches the native conformation in the C_{D9k} crystal structure (r_c = 7.6 Å, θ = 47°, ϕ = 28°).[17] Previous crystallographic studies of fluorinated aliphatic residues have indicated that substituting ^{1}H to ^{19}F mostly does not change the side chain orientation, and even CF$_3$ and O-CF$_3$ substitutions can preserve the native rotamer.[24,25] We also note that the fit did not change when data from the previously reported CF$_3$Phe-labeled C_{D9k}-50 were also included,[10] indicating that the side chain geometry is conserved despite the different fluorination of this buried residue.

2.2. Filtering Rotamer Simulations for C_{D9k}-50 Using ^{19}F-ENDOR Constraints

We next compared the independently fitted ^{19}F-ENDOR data to conformations predicted by rotamer simulations. When we use ^{19}F–Gd^{3+} distances from the full rotamer libraries as input for *EasySpin* ENDOR simulations, this overestimates the conformational space of the Phe residue (Supporting Information 4.3), giving poor agreement with experimental lineshapes. Testing individual χ_1 ranges revealed that only *t* conformers matched the experimental data (Figure 3), consistent with their positive tilt angle (θ > 0°) and a short average $\bar{r}_c \approx$ 7.9 Å.

Using the ENDOR-derived parameters as a reference, we filtered rotamers to include only those with θ > 0° and 7.4 Å $< r_c <$ 8.4 Å, excluding rotamers with minimum r_F distances below thresholds set by the ENDOR fit of each pcAA, that is, <6.3 Å (*p*-F), <7.2 Å (*m*-F$_2$), <8.4 Å (*o*-F$_2$). The resulting filtered rotamer set (red population in Figure 3b) improved the fits to the experimental spectra. It is unclear whether the fit could be improved further by data with better signal-to-noise ratio (SNR), or if it is limited by the conformational sampling of the rotamer library which depends only on an inter-atomic clash criterion.

2.3. Filtering Rotamer Simulations for C_{D9k}-66 Using ^{19}F-ENDOR Constraints

The same analysis was performed for C_{D9k}-66 variants labeled with *p*-F, *m*-F$_2$, and *o*-F$_2$. In this position, the Phe–Gd^{3+} distances are longer, resulting in lower spectral resolution and lower SNR. Previous work using CF$_3$Phe in position 66 yielded a ^{19}F-ENDOR spectrum corresponding to a mean r_F value of about 16 Å,[10]

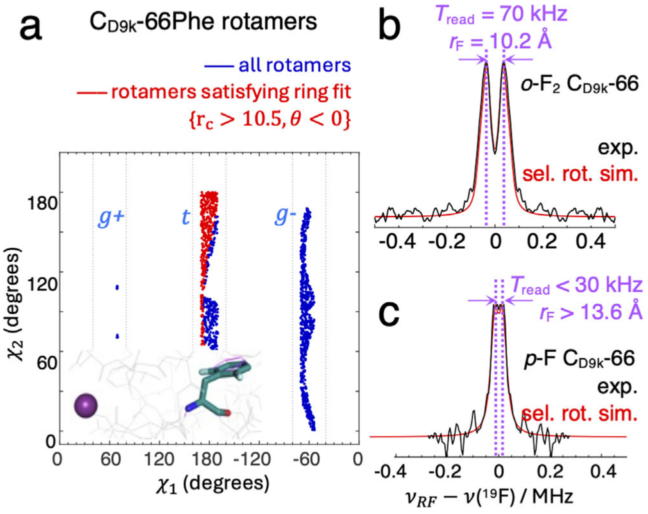

Figure 4. ENDOR-based rotamer analysis of C_{D9k} Phe66. a) Rotamer library (constrained χ_1 sampling–see Supporting Information) showing the populations (red points) consistent with the r_c and θ parameters determined from the ^{19}F-ENDOR data. b) Experimental ^{19}F-ENDOR spectra of C_{D9k} Phe66 labelled with *o*-F_2 and *p*-F, along with the simulations performed using the filtered rotamer populations as input into *Easyspin* calculations.

owing to the linewidth being too broad to resolve a splitting. The set of *p*-F, *m*-F_2, and *o*-F_2 residues provide access to a broader range of ^{19}F–Gd^{3+} distances, increasing the chance of detecting a resolved ENDOR signal. For C_{D9k}-66, we observed ENDOR signals for *p*-F and *o*-F_2, but not for *m*-F_2 (Figure 4). Only C_{D9k}-66 with *o*-F_2 showed a resolved splitting ($T_{read} = 70$ kHz, corresponding to $r_{read} = 10.2$ Å). The *p*-F spectrum was less sensitive and displayed limited fine structure. Its total linewidth indicated a maximum possible coupling of 30 kHz, corresponding to $r_F \geq 13.6$ Å. Since $r_o < r_p$, the phenyl ring must be tilted away from the Gd^{3+} ion ($\theta < 0°$), and its center must be further from the Gd^{3+} ion than the fluorine atoms in *o*-F_2 ($r_c > 10.2$ Å). This constrains the rotamer space to a subset of *t* conformers with $160° < |\chi_1| < 200°$ and $\bar{r}_c \approx 11.8$ Å, which agrees with the native Phe66 conformation (Figure 4a, inset).

Although not used in the analysis above, we note that the signal intensities of the pcAA-labeled proteins for both C_{D9k}-50 and C_{D9k}-66 approximately follow the expected trend, in that longer distances produce weaker ENDOR signals. When experimental conditions are carefully controlled such that the relative ENDOR intensities of different samples using the same labelling system can be quantitatively controlled,[11] this additional constraint of scaled intensity provides valuable information when moving to more dynamic systems where multiple conformations are sampled.

2.4. Analyzing Side Chain Orientations in GB1 Proteins Labelled with Flexible Spin Labels

As a next step, we investigated whether ^{19}F-ENDOR could be applied to more general cases involving spin-labelled proteins, where the Gd^{3+} ion is part of a flexible tag, while the aromatic residue may be buried or solvent-exposed, and in the latter case can populate multiple conformations. We synthesized five GB1 mutants, each labelled with a **Gd.C1** tag attached to a cysteine residue in position 28 and containing CF$_3$Phe installed in one of five positions: Q32 (**2a**), F30 (**2b**), V21 (**2c**), F52 (**2d**), or G14 (**2e**). All constructs displayed the same narrow EPR spectrum (see Supporting Information Figure S2.1), apart from **2b**, which displayed distinct populations (see Supporting Information 2). Based on the crystal structure (PDB: 1PGB),[26] the mean ^{19}F–Gd^{3+} distances range from 8 to 18 Å. Previously, we reported ENDOR spectra for all these proteins except for **2b**, showing that rotamer simulations poorly matched experimental data,[10,11] though no further conformational analysis was performed. As the conformational variability of the **Gd.C1** tag is much larger than that of the Phe side chain, we cannot restrain its conformation by a set of precise Euler angles, rendering the pcAAs an ineffective label choice. Therefore, we reverted to CF$_3$Phe, where the presence of three fluorine atoms (which can be represented by a single point dipole) increases the ENDOR sensitivity towards long ^{19}F–Gd^{3+} distances.[7,10]

To analyze the CF$_3$Phe conformations across all five GB1 mutants, we first fit the experimental ENDOR spectra using a lineshape model that is independent of rotamer simulations. This model assumes a Gaussian distribution of ^{19}F–Gd^{3+} distances, parameterized by the mean distance ($\bar{r}_{ls}$), distance distribution width (δ_{ls}), and spectral linewidth (*lw*).[10] Simulated lineshapes and corresponding distance distributions are shown in Figures 5a and 5b respectively (red traces). The Gaussian distance distribution width δ_{ls}, was fairly consistent across all GB1 variants, suggesting that the **Gd.C1** tag flexibility remains similar regardless of the position of the fluorinated Phe analogues. Differences in distance distributions can thus be attributed solely to variations in CF$_3$Phe side chain conformations (see Supporting Information 6.4 for further discussion).

We then generated rotamer libraries for both the **Gd.C1** tag at position 28 and for each CF$_3$Phe position, to calculate the rotamer-predicted r_F distributions, and compared their fits to the experimental ^{19}F-ENDOR data (Figure 5a, b). Details on χ_1 and χ_2 restraints are provided in Supporting Information 6. As with C_{D9k}, using unfiltered rotamer libraries in ENDOR simulations overestimates the conformational space, especially for solvent-exposed residues (Figure 5b)—a limitation which has also been demonstrated for rotamer libraries of other spin-labeled proteins.[13] The exception was **2d**, where the CF$_3$Phe residue is buried and adopts a single conformation. In this case, because the positions of the CF$_3$ is essentially fixed in space, the rotamer library *must* reflect only the flexibility of the **Gd.C1** tag. Indeed, the rotamer-predicted r_F distribution aligns very well with the Gaussian fit from the lineshape model, which supports the fact that the Gaussian lineshape model provides a valid description of the flexibility of a surface-tethered Gd^{3+} tag.

To refine the rotamer simulations for the GB1 mutants **2a**–**c** and **2e**, we filtered the rotamer libraries using the ^{19}F–Gd^{3+} distance distributions from the ENDOR lineshape model ($\bar{r}_{ls}$, δ_{ls}). Specifically, we retained only CF$_3$Phe rotamers, where at least one ^{19}F spin is at a distance *r* to at least 10% of the Gd^{3+} rotamers, and *r* falls within the ^{19}F–Gd^{3+} distance distribution predicted

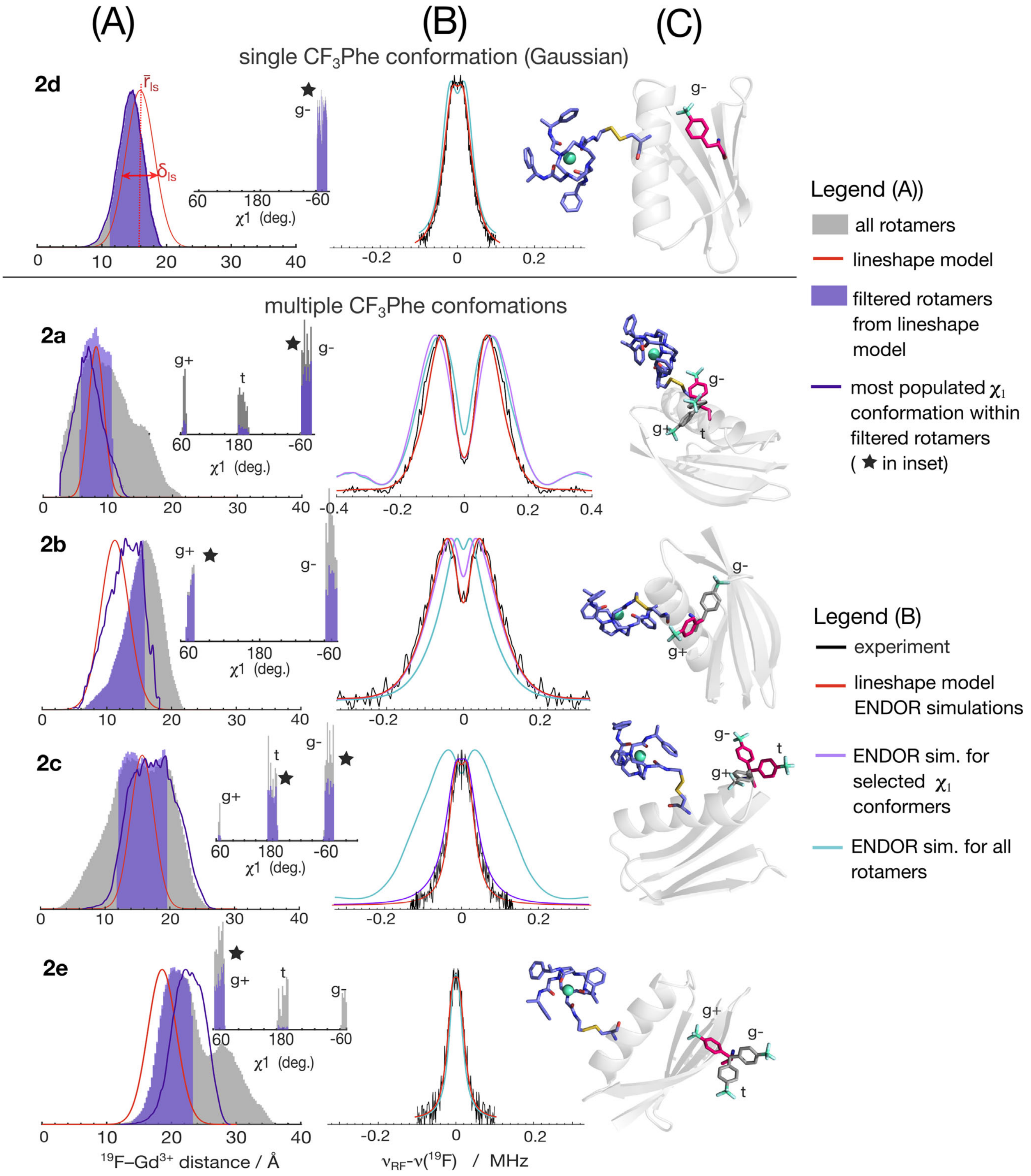

Figure 5. Conformational analysis of the CF$_3$Phe residue in the GB1 constructs **2a**–**2e**, using ENDOR lineshape analysis to identify the most likely conformations of the CF$_3$Phe side chains. a) ^{19}F–Gd^{3+} distance distributions calculated from rotamer libraries (grey histograms) overlayed with the distance distributions predicted by lineshape simulations in *Easyspin* (red trace). The portions of rotamers selected to match the lineshape model distribution are highlighted in purple. The χ_1 angles of the ENDOR-filtered rotamers are shown in red, overlayed over the χ_1 distribution of the whole rotamer library (grey). b) Experimental ENDOR data (symmetrised and smoothed; black trace) together with *Easyspin* Mims ENDOR simulations based on the lineshape model (red trace), the full rotamer library (light blue) and the filtered rotamers (purple). c) CF$_3$Phe conformers (stick representations) compatible with the experimental ENDOR data in the proteins **2a**–**2e**. The CF$_3$Phe conformations that best match the experimental data are coloured magenta. Alternative CF$_3$Phe rotamers predicted by the rotamer simulations but found to be incompatible with the experimental ENDOR data are coloured grey.

by the lineshape model. This filtered rotamer library (purple in Figure 5a) delivers significantly improved fits to the experimental ENDOR spectra compared to using the full, unadjusted rotamer library. The χ_1 populations sampled in the ENDOR filter are shown as insets in Figure 5a, indicating that the dominant χ_1 rotamers (marked with a star in each inset) are different for each CF_3Phe site, sometimes clustering around a single conformation. Due to the symmetry of the CF_3Phe residue, χ_2 angles remain unresolved.

The experimentally determined χ_1 rotamers suggest that a set of simple criteria can help in predicting the most likely conformations of solvent-exposed aromatic residues. The criteria includes: (i) conservation of native side-chain dihedral angles, (ii) maintaining hydrophobic interactions by keeping the phenyl ring close to the protein surface or core, and (iii) minimizing steric clashes. For example, in **2a**, although *g*+ would place the ring closer to the protein core, it deviates from the χ_1 angle of the native glutamine residue and could produce clashes with the nearby Val29 residue. Here, although fitting the ENDOR data suggests that all three staggered rotamers could reasonably be present in variable populations, the *g*- population seem to be preferred, which preserves the native χ_1 angle. In **2b**, only the *g*+ conformer fits the experimental ^{19}F ENDOR data, keeping the CF_3Phe side-chain in close proximity to the hydrophobic regions of the protein. Interestingly this is *not* the *g*- rotamer of the native Phe30 residue, which may be because the g- rotamer would result in a clash with the Thr18 methyl group in the protein core. The deviation from the native χ_1 angle for this buried residue may be the reason two protein populations are seen in the EPR spectrum of this sample, where one is likely a destabilized, misfolded population (see Supporting Information 6.4 for further discussion). In **2c**, where the CF_3Phe residue is on a flexible loop, both *g*- (native χ_1) and *t* conformers are equally populated in our filtered rotamer model, while the *g*+ conformer may clash with the nearby Thr25 and Asp22. For **2e**, where the CF_3Phe residue substitutes Gly14, the experimental data cannot be modelled without including the short r_F distances produced by the *g*+ rotamer which also minimizes solvent exposure of the side chain.

To further probe the conformations of flexible side chains, we also examined the relative ENDOR signal intensities, which have a $1/r^6$ dependence on the ^{19}F–Gd^{3+} distance. In **2c**, the observed splitting ($T_{\mathrm{read}} \approx 15 \pm 3$ kHz) corresponds to a ^{19}F–Gd^{3+} distance of 17.0 + 1.3/-1.0 Å, but the signal is more intense than for **2d**. This suggests that **2c** features a small population of a short-distance conformer, which adds intensity to the shoulders of the ENDOR spectrum, while the more highly populated longer-distance conformers dominate the splitting. Indeed, modeling the full lineshape and intensity via the lineshape model predicts a mean distance of 15.5 ± 0.9 Å. Peaks around both ~15 Å and 17 Å appear both in the full and filtered rotamer libraries, consistent with a bimodal distribution of r_F distances. Thus, an ENDOR signal that resolves a small splitting but high intensity can be seen as a marker for multiple conformers and multimodal distance distributions.

At present, reliable conformer assignments using both the ENDOR line splitting and intensity are only feasible for distance measurements in the 10–18 Å range, where the $1/r^6$ dependence of ENDOR intensity is close to linear. At short distances ($\lesssim 8$ Å), intense signals from low-population conformers can mask the contributions from more highly populated, longer-distance conformers. This is the case for **2a**, where the simulation using selected rotamers barely outperforms the simulation using the full library (Figure 5b). Similarly, at longer distances (>20 Å) where the $1/r^6$ dependence flattens, leads to long-distance conformers providing vanishingly small contributions to the overall signal. This situation is encountered for **2e**, where again, the full rotamer library explains the ENDOR data similarly well as the set of selected rotamers. Nonetheless, a lower-distance bound can still be placed on the fit of the data of **2e** based on the wings of the ENDOR spectrum, which represent the shortest Gd^{3+}-^{19}F distances sampled. In this instance, the shortest distances are produced by the rotamer *g*+ (see SI 5).

3. Conclusion

In conclusion, this work demonstrates the utility of ^{19}F-ENDOR spectroscopy for determining protein side-chain conformations. First, using the *p*-F, *m*-F_2, and *o*-F_2 tag set, ENDOR lineshape analysis accurately resolves the side chain conformations of buried fluorinated Phe residues in C_{D9k}, which are consistent with its crystal structure. We note that concurrently and independently of this work, a similar approach has been demonstrated by Bogdanov et al.,[27] using sets of fluorinated benzylamine guest molecules in a paramagnetic cavitand. Our present work has furthered the generalisability of this approach, not only by providing the first example of triangulating ^{19}F-ENDOR constraints within a real protein structure, but also demonstrated how conformational information can be construed for residues >11 Å away from the Gd^{3+} center by cross-validating ^{19}F-ENDOR modeling with rotamer simulations, even when the sensitivity of the ENDOR experiment becomes limited at these longer distances.

Second, in the more flexible GB1 system, where the Gd^{3+} ion is attached via a flexible tag, we show that distances determined from simulating the ENDOR lineshape and intensity–combined with rotamer simulations—similarly identify the dominant conformations of solvent-exposed CF_3Phe residues. With the rapid advancement of genetic encoding technologies for site-specific ^{19}F substitution,[28,29] ^{19}F-ENDOR is poised to become a powerful tool for probing side-chain conformations and their variability under physiological conditions, at low concentrations, and without molecular weight limitations. We further note that recently reported time domain ^{19}F-ENDOR approaches, which resolve narrower spectral linewidth,[16] potentially allows the same conformational analysis described here to be performed for distances longer than 18 Å.

4. Experimental

Experimental detail describing genetic encoding of noncanonical amino acids and protein expression is provided in Supporting Information 1. Details of EPR measurements (EPR echo-detected lineshape, electron spin relaxation data) for all the protein samples are provided in Supporting Information 2. Supporting Information 3 provides a thorough description of fitting the ENDOR data of the C_{D9k} proteins by modeling the phenyl ring orientation. Supporting Information 4 describes the rotamer simulations performed for C_{D9k} and includes more detail about how these were filtered against the ENDOR fits. Supporting Information 6 provides more detail on the rotamer simulations performed for the GB1 spin-labeled proteins, tabulates the lineshape model parameters $\{r_c, \theta, \phi\}$ for each sample and includes additional conformation-filtering results.

Supporting Information

The data supporting this article, including experimental details on protein synthesis and labelling, EPR relaxation measurements, as well as additional analysis of the ENDOR spectra and rotamer libraries, have been included as part of the Supplementary Information. The authors have cited additional references within the Supporting Information.[6,10,11,17,19,22,26,30–35]

Acknowledgments

This work was funded by the Australian Research Council (grants CE200100012, DP210100088, and DP240100273).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: ^{19}F-ENDOR · protein structure · rotamers · sidechain conformations

[1] T. Kirys; A. M. Ruvinsky; A. V. Tuzikov; I. A. Vakser, *BMC Bioinformatics* **2012**, *13*, 236.
[2] J. Janin; S. Wodak; M. Levitt; B. Maigret, *J. Mol. Biol.* **1978**, *125*, 357.
[3] L. X. Peterson; X. Kang; D. Kihara, *Proteins Struct. Funct. Bioinforma.* **2014**, *82*, 1971.
[4] O. F. Lange; N.-A. Lakomek; C. Farès; G. F. Schröder; K. F. A. Walter; S. Becker; J. Meiler; H. Grubmüller; C. Griesinger; B. L. de Groot, *Science* **2008**, *320*, 1471.
[5] R. L. Dunbrack, *Curr. Opin. Struct. Biol.* **2002**, *12*, 431.
[6] A. Meyer; S. Dechert; S. Dey; C. Höbartner; M. Bennati, *Angew. Chem. Int. Ed.* **2020**, *59*, 373.
[7] A. Kehl; M. Hiller; F. Hecker; I. Tkach; S. Dechert; M. Bennati; A. Meyer, *J. Magn. Reson.* **2021**, *333*, 107091.
[8] M. Gauger; M. Heinz; A.-L. J. Halbritter; L. S. Stelzl; N. Erlenbach; G. Hummer; S. Th. Sigurdsson; T Prisner, *Angew. Chem. Int. Ed.* **2024**, *63*, e202402498.
[9] N. B. Asanbaeva; A. A. Sukhanov; A. A. Diveikina; O. Y. Rogozhnikova; D. V. Trukhin; V. M. Tormyshev; A. S. Chubarov; A. G. Maryasov; A. M. Genaev; A. V. Shernyukov; G. E. Salnikov; A. A. Lomzov; D. V. Pyshnyi; E. G. Bagryanskaya, *Phys. Chem. Chem. Phys.* **2022**, *24*, 5982.
[10] M. M. Judd; E. H. Abdelkader; M. Qi; J. Harmer; T. Huber; A. Godt; A. Savitsky; G. Otting; N. Cox, *Phys. Chem. Chem. Phys.* **2022**, *24*, 25214.
[11] M. Judd; M. Qi; E. H. Abdelkader; H. Qianzhu; A. Savitsky; T. Huber; J. R. Harmer; A. Godt; G. Otting; N. Cox, *J. Am. Chem. Soc.* **2025**, *147*, 16826.
[12] A. Bogdanov; V. Frydman; M. Seal; L. Rapatskiy; A. Schnegg; W. Zhu; M. Iron; A. M. Gronenborn; D. Goldfarb, *J. Am. Chem. Soc.* **2024**, *146*, 6157.
[13] A. Bogdanov; L. Gao; A. Dalaloyan; W. Zhu; M. Seal; X.-C. Su; V. Frydman; Y. Liu; A. M. Gronenborn; D. Goldfarb, *Phys. Chem. Chem. Phys.* **2024**, *26*, 26921.
[14] M. Seal; W. Zhu; A. Dalaloyan; A. Feintuch; A. Bogdanov; V. Frydman; X.-C. Su; A. M. Gronenborn; D. Goldfarb, *Angew. Chem. Int. Ed.* **2023**, *62*, e202218780.
[15] A. Kehl; L. Sielaff; L. Remmel; M. L. Rämisch; M. Bennati; A. Meyer, *Phys. Chem. Chem. Phys.* **2025**, *27*, 1415.
[16] L. Sielaff; A. Kehl; A. Aden; A. Meyer; M. Bennati, *Sci. Adv.* *11*, eady5665.
[17] L. A. Svensson; E. Thulin; S. Forsén, *J. Mol. Biol.* **1992**, *223*, 601.
[18] I. Bertini; I. Gelis; N. Katsaros; C. Luchinat; A. Provenzani, *Biochemistry* **2003**, *42*, 8011.
[19] H. Qianzhu; Y. J. Tan; E. H. Abdelkader; T. Huber; G. Otting, *ACS Sens.* **2025**, *10*, 3152.
[20] Z. Miao; Y. Cao, *Sci. Rep.* **2016**, *6*, 37024.
[21] D. F. Gauto; O. O. Lebedenko; L. M. Becker; I. Ayala; R. Lichtenecker; N. R. Skrynnikov; P. Schanda, *J. Struct. Biol. X* **2023**, *7*, 100079.
[22] S. Stoll; A. Schweiger, *J. Magn. Reson.* **2006**, *178*, 42.
[23] R. L. Frkic; Y. J. Tan; E. H. Abdelkader; A. Maleckis; E. Tarcoveanu; C. Nitsche; G. Otting; C. J. Jackson, *Biochemistry* **2024**, *63*, 1388.
[24] R. L. Frkic; Y. J. Tan; A. Maleckis; N. F. Chilton; G. Otting; C. J. Jackson, *Biochemistry* **2024**, *63*, 2602.
[25] H. W. Orton; H. Qianzhu; E. H. Abdelkader; E. I. Habel; Y. J. Tan; R. L. Frkic; C. J. Jackson; T. Huber; G. Otting, *J. Am. Chem. Soc.* **2021**, *143*, 19587.
[26] T. Gallagher; P. Alexander; P. Bryan; G. L. Gilliland, *Biochemistry* **1994**, *33*, 4721.
[27] A. Bogdanov; M. Seal; E. Goren; A. Bar-Shir; D. Goldfarb, *Phys. Chem. Chem. Phys.* **2025**, *27*, 3885.
[28] H. Qianzhu; E. H. Abdelkader; I. D. Herath; G. Otting; T. Huber, *ACS Sens.* **2022**, *7*, 44.
[29] H. Qianzhu; E. H. Abdelkader; G. Otting; T. Huber, *J. Am. Chem. Soc.* **2024**, *146*, 13641.
[30] A. Nalepa; K. Möbius; W. Lubitz; A. Savitsky, *J. Magn. Reson.* **2014**, *242*, 203.
[31] W. B. Mims, *Proc. R. Soc. Lond. Ser. Math. Phys. Sci.* **1965**, *283*, 452.
[32] M. Stanton-Cook; X.-C. Su; G. Otting; T. Huber, PyParaTools. https://compbio.anu.edu.au/mscook/PPT.
[33] H. Yagi; D. Banerjee; B. Graham; T. Huber; D. Goldfarb; G. Otting, *J. Am. Chem. Soc.* **2011**, *133*, 10418. https://doi.org/10.1021/ja204415w.
[34] S. C. Lovell; J. M. Word; J. S. Richardson; D. C. Richardson, *Proteins Struct. Funct. Bioinforma.* **2000**, *40*, 389. https://doi.org/10.1002/1097-0134(20000815)40:3(389::AID-PROT50>3.0.CO;2-2.
[35] A. Arseniev; P. Schultze; E. Wörgötter; W. Braun; G. Wagner; M. Vašák; J. H. R. Kägi; K. Wüthrich, *J. Mol. Biol.* **1988**, *201*, 637. https://doi.org/10.1016/0022-2836(88)90644-4.

Manuscript received: September 1, 2025
Revised manuscript received: October 12, 2025
Version of record online: ■■, ■■■