

PCCCP

Physical Chemistry Chemical Physics

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: M. Judd, E. H. Abdelkader, M. Qi, J. R. Harmer, T. Huber, A. Godt, A. Savitsky, G. Otting and N. Cox, *Phys. Chem. Chem. Phys.*, 2022, DOI: 10.1039/D2CP02889A.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Short-range ENDOR distance measurements between Gd(III) and trifluoromethyl labels in proteins.

Martyna Judd,^a Elwy Abdelkader Ali,^b Mian Qi,^c Jeffrey R. Harmer,^d Thomas Huber,^a Adelheid Godt,^c

Anton Savitsky,^c Gottfried Otting,^b Nicholas Cox^{a,}*

^a *Research School of Chemistry, The Australian National University, 137 Sullivans Creek Rd, 2601 Acton, Canberra ACT, Australia*

^b *ARC Centre of Excellence for Innovations in Peptide & Protein Science, Research School of Chemistry, The Australian National University, 137 Sullivans Creek Rd, 2601 Acton, Canberra ACT, Australia*

^c *Faculty of Chemistry and Center for Molecular Materials (CM₂), Bielefeld University, Universitätsstraße 25, Bielefeld 33615, Germany*

^d *Centre for Advanced Imaging, The University of Queensland, St Lucia, QLD 4072 Australia*

^e *Faculty of Physics, Technical University Dortmund, Otto-Hahn-Str. 4a, 44227 Dortmund, Germany*

AUTHOR INFORMATION

Corresponding Authors

Nicholas Cox: nick.cox@anu.edu.au;

ABSTRACT:

The measurement of distances in proteins can be challenging in the 5–20 Å range, which is outside the range accessible through conventional NMR and EPR methods. Recently it was demonstrated that distances in this range could be measured between a nitroxide as a paramagnetic spin label and a nearby fluorine atom (^{19}F) as a nuclear spin label using high-field (W-band / 3.4 T) ENDOR spectroscopy. Here we show that such measurements can also be performed using a gadolinium ion (Gd^{3+}) as the paramagnetic tag. Gd^{3+} has two advantages. i) A greater electronic spin ($S = 7/2$) and fast electronic spin-lattice (T_1) relaxation, improving sensitivity by allowing data to be collected at lower temperatures. ii) A narrow EPR signal for the $-1/2 \rightarrow 1/2$ transition, and therefore no orientation selection artefacts. Signal intensities can be further enhanced by using a trifluoromethyl (C^{19}F_3) group instead of a single ^{19}F atom. Using the protein calbindin $\text{D}_{9\text{k}}$ with a Ca^{2+} ion replaced by a Gd^{3+} ion and a trifluoromethylphenylalanine in position 50, we show that distances up to about 10 Å can readily be measured. Longer distances proved more difficult to measure due to variable electronic T_M relaxation rates, which lead to broader Lorentzian ENDOR lineshapes. Gd^{3+} complexes (Gd^{3+} tags), which reliably display longer T_M times, allow longer distances to be measured (8–16 Å). We also provide preliminary evidence that the intensity of ENDOR signals follows the predicted $1/r^6$ dependence, indicating that distances $r > 20$ Å can be measured by this method.

1. Introduction

Tagging proteins with paramagnetic labels provides a well-established avenue for the measurement of nanoscale structural restraints, also allowing local and global protein dynamics to be characterized. In combination with a nuclear spin, the measurement of distances between the introduced tag and nuclei of the protein 15–40 Å away can be achieved using NMR spectroscopy. For example, Paramagnetic Relaxation Enhancements (PRE) measurements are frequently used. Nuclear spins in the vicinity of the paramagnetic centre experience an enhanced relaxation rate, which has a $1/r^6$ dependence on the electron–nuclear distance r . The introduction of a second paramagnetic tag allows inter-tag distances to be measured using EPR spectroscopy. The best known technique is double electron-electron resonance (DEER), also known as PELDOR,^{1–4} but there exist a variety of methods.^{5,6} These techniques probe the electron–electron dipolar coupling between the paramagnetic labels, which features a $1/r^3$ distance dependence, and allow accurate measurements of distances in the 20–80 Å range.

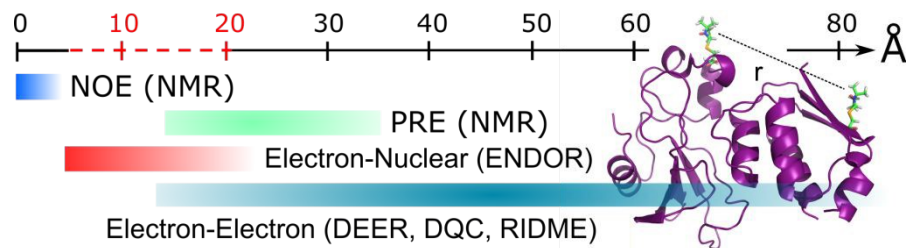


Figure 1. Introduction of paramagnetic tags allows the measurement of specific distances (r) in proteins. NMR spectroscopy exploits the nuclear Overhauser effect (NOE) to measure distances shorter than about 5 Å in proteins without spin labels. Paramagnetic relaxation enhancement (PRE) extends the distance range accessible by NMR spectroscopy to 40 Å, but the range between about 5–15 Å is difficult to address by PRE. Pulse EPR techniques, which exploit electron–electron interactions, extend the length scale to about 80 Å (and potentially beyond) but also lack broad applicability below 20 Å. EPR techniques probing electron–nuclear interactions (ENDOR) allow this difficult-to-access medium distance range (5–20 Å) to be interrogated.

When applying these EPR techniques to shorter distances, they are hampered by fast electronic $1/T_1$ (spin-lattice) relaxation rates.^{7–10} In the case of very short distances (<10 Å), the distance can be extracted by spectral simulation of the EPR lineshape^{11,12} or determining the relative intensity of half-field (spin forbidden) transitions observable in such systems.¹³ At a more practical level however, distances in this range begin to approach the size of the spin label including the overall tether (5–10 Å; for details see SI 3), which most often have a large degree of conformational flexibility, rendering such distance measurements between two spin labels less meaningful.

An alternative approach is to use EPR double resonance techniques such as Electron Nuclear Double Resonance (ENDOR).^{14–16} Like PREs, ENDOR exploits the electron–nuclear dipole interaction between the tag and the nearby nuclear spins, but its distance dependence corresponds to that of DEER ($1/r^3$).¹⁴ thus allowing short distances to be measured. Recently it was shown that the Mims ENDOR experiment¹⁷ can be used to measure distances on the 5–15 Å scale, between a nitroxide as the paramagnetic spin tag and an introduced fluorine nucleus (^{19}F).^{15,16,18} This measurement requires only a single paramagnetic tag, improving the accuracy of measured distances. The frequency resolution necessary to resolve electron-nuclear splittings on the order of kHz requires a nuclear spin with a large gyromagnetic ratio γ_n such as that of ^{19}F , as the magnitude of the ENDOR splitting scales linearly with γ_n . Fluorine is generally absent in biological systems and can be introduced into proteins easily through non-canonical amino acids,^{19,20} making it a selective and sensitive tag choice. To resolve ^{19}F from the background ^1H signal, the ENDOR experiment needs to be performed at fields above 1 T, i.e. at Q-band frequencies or higher.

The main disadvantage of the Mims ^{19}F -ENDOR experiment is its relatively low sensitivity. This is particularly an issue for measuring longer distances, which require longer pulse separations to suppress tau-dependent, blind-spotting artefacts and therefore lead to long accumulation times up to 30–60 h.^{15,18} In addition, paramagnetic tags associated with broad EPR spectra such as nitroxides require the ENDOR spectrum to be recorded at multiple field positions across the EPR spectrum to fully sample the dipolar Pake pattern (hyperfine couplings).

Preliminary attempts have already been made to overcome the limitations of nitroxide tags in the context of ^{19}F ENDOR. Bagryanskaya and coworkers investigated the use of triarylmethyl (trityl) radicals as alternative tags.²¹ Trityl radicals possess narrow EPR lines due to low g -anisotropy, and thus the ENDOR spectrum can be collected at a single magnetic field position without needing to account for orientation

selectivity. They also possess relatively long electronic phase-memory (T_M) relaxation times (4–5 μ s at 80 K), allowing the use of long delays and long RF pulses, t_{RF} as required to properly resolve small line splittings arising from longer distances.

To date, reliable ^{19}F -ENDOR distance measurements with trityl tags have only been achieved in the range of 8–12 \AA with model compounds.²¹ As trityl radicals are large and feature significant delocalization of the spin density, this introduces considerable uncertainty into the distance measurement, in particular if the trityl radical is attached to the target molecule via a flexible tether. Where a tether containing conjugated linkers is involved, a significant isotropic (fermi contact) hyperfine interaction at inter-spin distances shorter than about 8 \AA is introduced, which complicates distance determination. Finally, many possible conformational orientations of the trityl group result in heterogenous line broadening, which may compromise the hyperfine resolution required for longer distances. The use of the trityl radical as a generalizable tag in biomolecular systems is yet to be demonstrated.

To address the limitations of nitroxides and trityl radicals, Gd^{3+} complexes offer an attractive alternative. The Gd^{3+} ion has a half-filled valence f-shell ($4f^7$), which gives rise to an orbitally non-degenerate ground electronic configuration (^8S , $S = 7/2$). Most Gd isotopes carry no nuclear spin, so that the $-\frac{1}{2} \rightarrow \frac{1}{2}$ transition of the spin manifold gives rise to a sharp EPR transition ($\Delta\nu \approx 50$ MHz) that is ideal for sensitive detection. As the electron configuration of Gd^{3+} is high-spin, only modest microwave powers are required to achieve short microwave pulse lengths t_π that fully excite the EPR line ($t_\pi = <16$ ns with ca. 400 mW, $\Delta\nu = 60$ MHz). In addition, the unpaired electrons in a Gd^{3+} ion relax faster (their T_1 is up to 5–10 times shorter than for nitroxide^{15,16} and trityl radicals²¹), allowing data accumulation at lower temperatures than with nitroxides (10 K vs. 50 K), which further improves sensitivity. Like trityl labels, Gd^{3+} tags are also free from orientation selection artefacts, as shown earlier for DEER spectroscopy,^{22,23} i.e. the entire ENDOR Pake pattern can in principle be sampled in a single measurement. As opposed to rigid model compounds, Gd^{3+} tags designed for proteins often feature minimum tether lengths to the protein backbone of about 8 \AA . Notably, tether lengths are also a problem with nitroxide tags, as the most popular spin label MTSL, used in DEER, results in an overall tether of comparable length and flexibility (see SI 3). For benchmarking purposes, the present work introduced a Gd^{3+} ion into one of the Ca^{2+} binding sites of calbindin D_{9k} to increase the accuracy of short distance measurements.

Optimizations of the ^{19}F -based nuclear spin tag have also been explored in recent literature.^{15,16} Replacing the single fluorine atom with multiple fluorine atoms has been shown to increase the ENDOR signal response linearly with the number of fluorine atoms.¹⁶ The trifluoromethyl (CF_3) group is most suitable to position all ^{19}F spins at approximately the same distance to the paramagnetic spin tag.

Here we show that a tag pair consisting of Gd^{3+} and the genetically encoded non-canonical amino acid *p*-trifluoromethylphenylalanine ($\text{CF}_3\text{-Phe}$)²⁴ allows the reliable measurement of short-range distances up to 16 Å. We also show that shorter distances can be resolved in rigid biomolecules, such as proteins with a metal ion binding site, in addition to the orientation of the aromatic ring of the $\text{CF}_3\text{-Phe}$ residue relative to the Gd^{3+} site. Finally, we show preliminary evidence that the intensity of the ENDOR signal follows the predicted $1/r^6$ dependence. Using an integration approach to processing the data could potentially give access to much longer distances between 20 and 25 Å. Within these points of investigation, the broader aim of this paper is also to provide a practical assessment, including the strengths and limitations, of using of Gd^{3+} in ENDOR distance measurements in biological samples.

2. Results

2.1 Modelling Considerations. In this study we used the non-canonical amino acid $\text{CF}_3\text{-Phe}$ as the nuclear spin tag. The CF_3 group undergoes only weak intermolecular interactions^{25,26} and thus should interact little with the protein structure in the absence of steric clashes. ^{19}F NMR spectra of protein-bound $\text{CF}_3\text{-Phe}$, show degenerate ^{19}F resonances as expected for fast rotation of the CF_3 group around the $\text{C}_{\text{benzene}}\text{-C}_{\text{F}}$ bond at room temperature.²⁴ However, DFT calculations for temperatures below 50 K report a marked energy barrier for rotation about the C-C bond, predicting a barrier >5 kJ/mol at 10 K.¹⁶ We therefore assumed that the CF_3 group adopts one of two staggered orientations with respect to the plane of the aromatic ring (Fig. 2a), and modelled the resulting six possible positions of the three fluorine atoms as a cone, with the Gd^{3+} centre at the apex of the cone and the fluorine atoms defining the base of the cone (Fig. 2b).

To model systems where the location of the paramagnetic centre is well defined such as in metalloproteins, we took an approach similar to that described by Goldfarb and co-workers.¹⁴ The simplest model for simulating our Gd³⁺-CF₃ data requires three main fitting parameters: (i) the interspin distance r between the Gd³⁺ centre and the base of the cone; (ii) the angle ϕ (tilt angle) of the plane of the cone base relative to the interspin distance vector; and (iii) the ENDOR linewidth $1W$, from contributions such as relaxation processes and power broadening, which we account for by the single fitting parameter $1W$, assuming a Lorentzian lineshape. In the most extreme cases, the tilt angle is either 0° or 90°. For $\phi = 0^\circ$, all fluorine atoms are at the same distance to the Gd³⁺ centre, giving rise to an undistorted Pake pattern (Fig. 2c-d), the same as for the case of a single fluorine atom. If instead $\phi = 90^\circ$, the maximum and minimum interspin distances differ by about 2 Å. If the mean distance is 10 Å, this can lead to a coupling variation of up to 50 kHz (Fig. 2c), which can theoretically give rise to a structured ENDOR spectrum (Fig. 2d). Above a Gd³⁺-¹⁹F distance of 15 Å, the variation is comparable with the ENDOR linewidth, $1W$, and therefore the ENDOR signal yields essentially the same Pake pattern lineshape for both orientations (0 and 90°).

However, to allow fitting of the tilt angle, the interspin vector needs to be well defined, which only holds if the paramagnetic spin tag is not tethered to the protein via a flexible linker. In the case of a flexible tether such as for the Gd³⁺ label used in the herein reported study (Fig. 3b), the parameter ϕ can be omitted. Here, the lineshape of the ENDOR spectrum is dominated by the distance distribution arising from the inherent flexibility of the overall tether by which the Gd³⁺ tag is attached to the protein. This line broadening is larger than the intrinsic inhomogeneously broadened ENDOR linewidth $1W$. In this case of Gd³⁺ tags with flexible linkers, a tilt angle ϕ cannot be constrained, and we instead fit the Gd-F₃ distance distribution to a mean distance \bar{r} , and a Gaussian function for the distance distribution with a

standard deviation of σ . For both approaches a least-squares simplex minimization algorithm employed in MATLAB was used to arrive at the best-fit parameters.

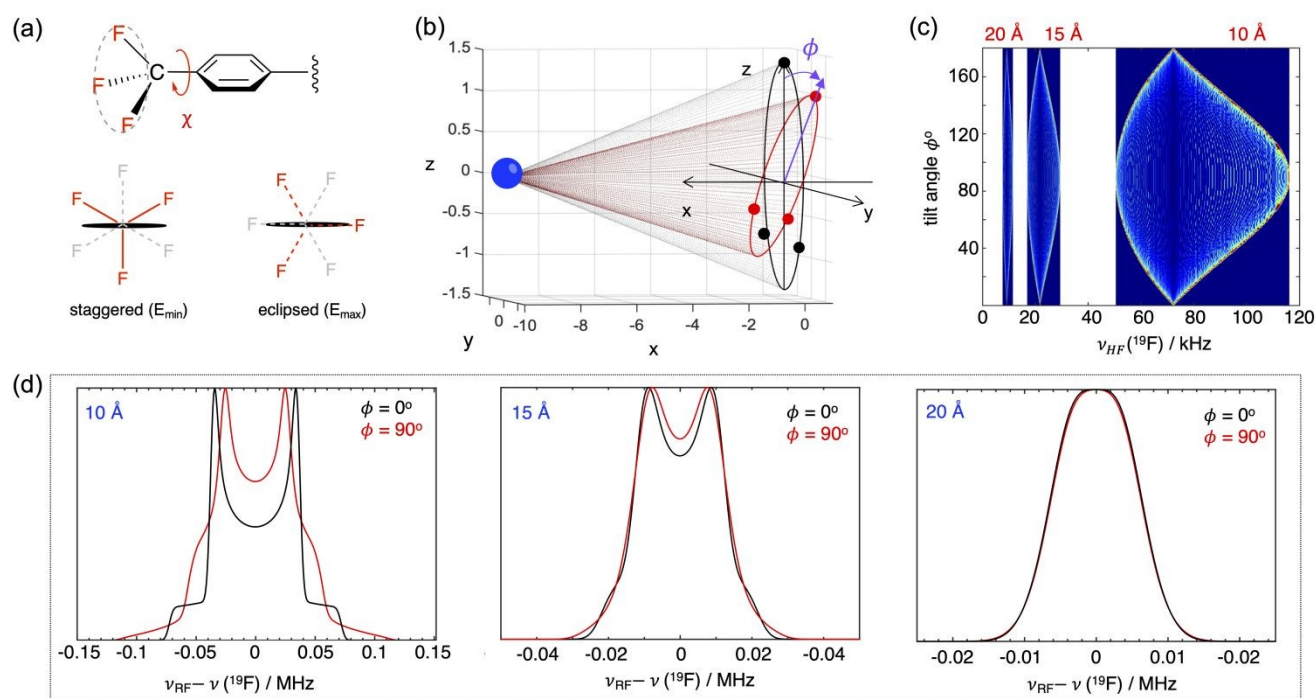


Fig. 2 Summary of the modelling approach used to simulate ^{19}F -ENDOR spectra in systems with a CF_3 group and a Gd^{3+} tag **(a)** Schematic representation of the CF_3 group of CF_3 -Phe. The model assumes two different rotameric states of the CF_3 group defined by the dihedral angle χ . To accelerate the ^{19}F -ENDOR simulations, we considered only the two staggered conformations, which present states of minimum energy, amounting to sampling six points on the CF_3 arc of rotations to calculate the range of $\text{Gd}-^{19}\text{F}$ distances. **(b)** Parametrisation of the CF_3 -Phe tag in terms of the distance r between the Gd^{3+} centre and the base of the center of the F_3 plane of the CF_3 group, and the tilt angle ϕ of the trifluoro plane (cone base) towards or away from the $\text{Gd}-\text{F}_3$ vector, which can be parametrised as a cone of $\text{Gd}-^{19}\text{F}$ interspin vectors. **(c)** Line broadening of the EPR spectrum caused by the interspin distance distribution resulting from $\phi \neq 0^\circ$. Line broadening is particularly pronounced in the wings of the ENDOR spectrum, as shown in **(d)**. **(d)** Line shapes (simulation linewidth = 6 kHz) predicted for different distances and $\phi = 0^\circ$ or 90° . As the hyperfine coupling is proportional to r^{-3} , the $\text{Gd}-\text{CF}_3$ interaction for distances longer than 15 Å is adequately represented by a single ^{19}F position, neglecting the effects from different ϕ angles.

2.2 Accurate distance measurements in model proteins containing a fixed Gd^{3+} ion position and a CF_3 -Phe residue. We first examined the viability of using Gd^{3+} as a spin tag in the ^{19}F -ENDOR experiment with the model protein calbindin $\text{D}_{9\text{k}}$ ($\text{C}_{\text{D}9\text{k}}$). $\text{C}_{\text{D}9\text{k}}$ possesses two Ca^{2+} -binding sites, where one of the Ca^{2+} ions can be exchanged for a Gd^{3+} ion. A Gd^{3+} ion bound in this way represents an effectively rigidly bound, i.e. conformationally unambiguous paramagnetic label. We used two mutants of $\text{C}_{\text{D}9\text{k}}$, Phe50 CF_3 -Phe and Phe66 CF_3 -Phe, loaded with a Gd^{3+} ion, here referred to as $\text{C}_{\text{D}9\text{k}-1}$ and $\text{C}_{\text{D}9\text{k}-2}$ respectively. The two proteins have a predicted $\text{Gd}^{3+}-^{19}\text{F}$ distances of about 7 Å ($\text{C}_{\text{D}9\text{k}-1}$) and 15 Å ($\text{C}_{\text{D}9\text{k}-2}$).

2). In the crystal structure of the wild-type protein (PDB ID: 4ICB²⁷), the conformational space of the aromatic rings of CF₃Phe is restricted, although the CF₃ group of CF₃-Phe in position 50 is expected to be partially solvent-exposed (Fig. 3a) and therefore orientated away from the protein core. In position 66 on the other hand, the CF₃-Phe is buried in the protein core, although as shown in Fig. 2c at longer distances the orientation information becomes less resolved and important.

Fig. 3 shows the ¹⁹F Mims ENDOR spectra of C_{D9k-1}. ENDOR data for C_{D9k-2} are shown in the Supporting Information (SI 3). Mims ENDOR is subject to τ -dependent blind-spot artefacts, given by:

$$F_{\text{ENDOR}} = \frac{1}{2} \sin^2\left(\frac{A}{2\tau}\right) \approx \frac{1}{2} \left(\frac{A}{2\tau}\right)^2 \propto \frac{1}{\tau^6} \quad (1)$$

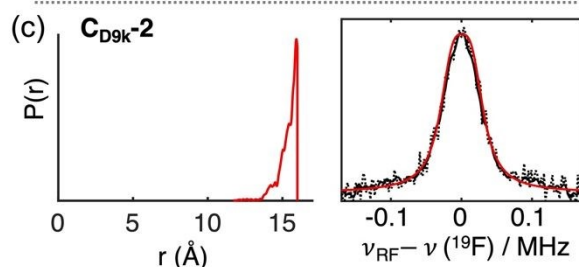
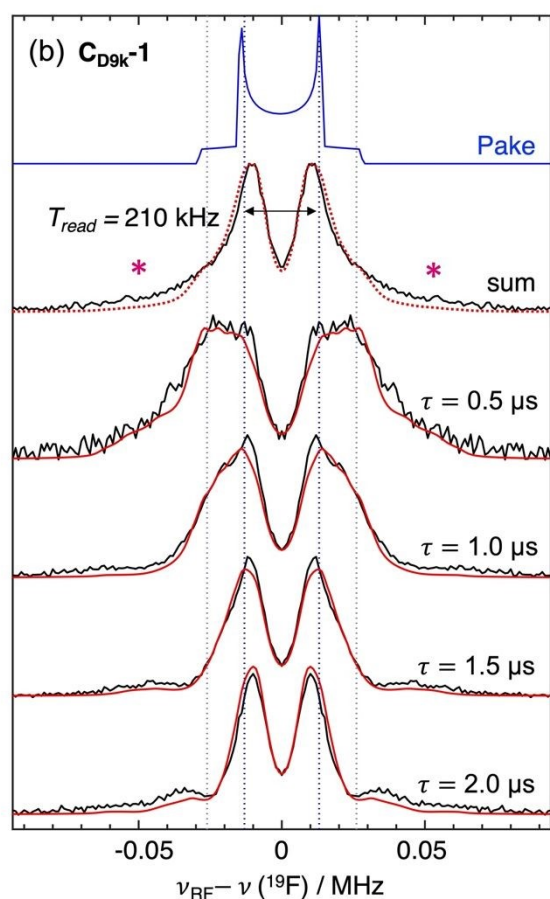
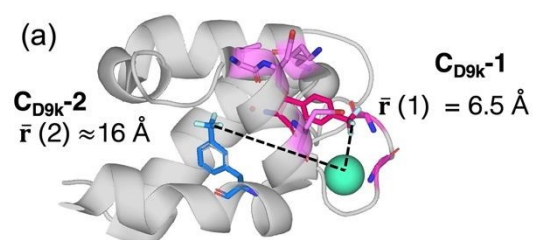
where A is the hyperfine coupling and τ is the delay between the $\pi/2$ pulses of the detection sequence (see experimental section). When $\frac{A}{2\tau}$ is small, i.e. in the “small-angle limit” (in the sense of the variable of the \sin^2 function), the relation can be approximated by the right-hand side of Eq. 1, and the ENDOR efficiency, F_{ENDOR} , is proportional to $1/\tau^6$.¹⁴ To fully resolve the ENDOR lineshape, the spectrum of C_{D9k-1} was averaged over four different τ values: 0.5, 1.0, 1.5 and 2.0 μ s. A resolved ¹⁹F-ENDOR line splitting of ≈ 210 kHz was observed, as read from the sum spectrum in Fig. 2a. This was interpreted as the perpendicular component of the dipolar interaction (T)ⁱ between the two spins, referred to by Meyer et al. as T_{read} ,¹⁵ and is consistent with an interspin distance of ≈ 7.1 Å.

Interestingly, the ENDOR lineshape of C_{D9k-1} and C_{D9k-2} is Lorentzian (≈ 60 kHz), which manifests in terms of broad wings that extend up to 50 kHz either side about the centre. This stands in contrast to the results obtained for the rigid compounds incorporating the nitroxide and ¹⁹F label pair reported by Bennati and co-workers,¹⁵ which instead exhibited narrower Gaussian linewidths of between 10–20 kHz. We also observed similar narrow linewidths in the MTSL-labelled proteins in this work (SI 5). The increase in linewidth correlates with faster T_M relaxation rates observed for the Gd³⁺ ion as compared to nitroxide labels, 7.0 vs. 13–15 μ s. As a consequence of the larger linewidth of the fast-relaxing C_{D9k}, the longer distance in C_{D9k-2} did not yield a resolved splitting (see SI 5) and, therefore the distance parameter could not be constrained reliably.

ⁱ Note that in this text we use the unitalicized symbol T to denote dipolar couplings in Hz in line with the notation used previously,^{15,16,18} whereas italicised T values indicate time constants, such as the relaxation times T_1 and T_M .

For C_{D9k-1} , the overall profile of the ENDOR signal was not consistent with expectations for a single distance. As described in section 2.1, we needed to account for a spread of Gd– ^{19}F distances introduced by the tilt of the CF_3 group with respect to the Gd– $\text{C}_{3\text{F}}$ interspin vector. This was particularly important to fit the broad wings of the ENDOR spectra. The fitting parameters shown in Table 1 yielded a mean distance \bar{r}_{sim} of $6.5 \pm 0.3 \text{ \AA}$ (very close to the 7 \AA distance described above) and tilt angle of $\sim 90^\circ$, with the CF_3 group of the CF_3 -Phe residue lying adjacent to the protein surface as opposed to pointing into the polar solvent, as expected from the crystal structure.

As described above, we did not resolve a splitting for C_{D9k-2} . Protein-bound Gd^{3+} ions therefore likely have more limited applications as paramagnetic tags in the ^{19}F -ENDOR measurement, being only able



to detect shorter distances. We can estimate the shortest distance that can be measured using the fitted linewidth. This is 35 kHz , limiting our maximum ruler length to 12 \AA . We do however note that the C_{D9k-2} lineshape is at least consistent with the engineered distance of $\approx 15 \text{ \AA}$. The fitting shown in Fig. 3c includes a modelled distance distribution ($14.5\text{--}16 \text{ \AA}$) that was obtained by rotamer-state simulations using the program PyParaTools.²⁸ While we are unable to detect a splitting of the ^{19}F -ENDOR signal, we can readily detect it. As a result, protein-bound Gd^{3+} ions can be used to definitively show a fluorine label at least 12 \AA from the metal site.

Fig. 3 ^{19}F -Mims ENDOR experimental data and simulations of the C_{D9k-1} and C_{D9k-2} systems. (a) Structure of C_{D9k} showing the CF_3 -Phe residue in position 50 (magenta) in C_{D9k-1} and position 66 (blue) in C_{D9k-2} , and the gadolinium ion (cyan) in the N-terminal calcium binding site. The hydrophobic residues around the 50CF_3 -Phe are shown in violet. (b) Normalised 94 GHz ^{19}F -ENDOR spectra of C_{D9k-1} measured at τ -values of $0.5\text{--}2.0 \text{ \mu s}$. Experimental data are shown in black. T_{read} was determined from the weighted-sum spectrum. The asterisks (red) annotated identify the broad wing features extending 50 kHz either side of the centre of the spectrum. Simulations (red traces) were performed using a global four-parameter fit, including the mean distance \bar{r} (shown in panel a), the tilt angle ϕ of the trifluoro plane with respect to the Gd– $\text{C}_{3\text{F}}$ vector, the Lorentzian linewidth and the

Gaussian ENDOR linewidth. All five spectra were simultaneously fitted using a least-squares simplex minimization algorithm. The best fit parameters are listed in Table 1. The acquisition time amounted to approximately one hour per spectrum or 5 hours in total for the sum spectrum. Further experimental details are provided in the methods section and SI 5. (c) $\text{CF}_3\text{-Gd}^{3+}$ distance distribution predicted for $\text{C}_{\text{D9k-2}}$ (first panel) predicted using the rotamer-state simulation program PyParaTools. The second panel is the ^{19}F -ENDOR experimental spectrum of $\text{C}_{\text{D9k-2}}$ (black trace 2-point smoothing, dotted trace unsmoothed data), overlaid with the *Easyspin* (saffron) Mims ENDOR simulation using the PyParaTools-predicted distance distribution to calculate the ^{19}F hyperfine coupling inputs for the *EasySpin* calculation.

2.3 Distance measurements in proteins containing a tethered Gd^{3+} spin label and $\text{CF}_3\text{-Phe}$ residue. To examine the viability of using Gd^{3+} as a spin tag in a more generalizable way, we turned to using the **Gd.C1** tag (a Gd^{3+} complex with a DOTAM-derived ligand) in the immunoglobulin-binding protein G (B1 domain, here referred to as **GB1**) as a model system (Fig. 5).^{28,29} The **Gd.C1** tag was installed on the protein via a surface exposed cysteine residue with formation of a disulfide linkage (SI 1). Three cysteine mutants with $\text{CF}_3\text{-Phe}$ in place of Phe52 were prepared and reacted with **Gd.C1** tags to yield $\text{Gd}^{3+}\text{-}^{19}\text{F}$ distances of approximately 10 Å (A24C, **GB1-a**), 15 Å (K28C, **GB1-b**) and 20 Å (Q32C, **GB1-c**). Although the use of the **Gd.C1** spin tag results in a fairly long and flexible linker, its conformational sampling can be limited by hydrophobic interactions between its pendant arms and the protein surface (see SI 6).

Figure 4 shows the structure of **GB1** modelled with $\text{CF}_3\text{-Phe}$, indicating the positions of the **Gd.C1** tags in the three mutants, together with the measured ^{19}F -ENDOR data. As seen for C_{D9k} , the ENDOR line shapes of the **GB1-a-c** are characteristically Lorentzian, leading to broad wings extending out either side of the centre. Unlike in the case of C_{D9k} , however, the spectra of **GB1-a-c** displayed significantly narrower linewidths (≤ 35 kHz), which allows the measurement of longer distances. This can be attributed to longer, invariant T_M relaxation times (see SI 4), which resulted in better linewidth resolution. As expected, the spectrum **GB1-a** resolves a splitting of ≈ 100 kHz, consistent with an interspin distance of about 9.4 ± 0.3 Å and **GB1-b** spectrum resolved a splitting of ≈ 20 kHz, consistent with an interspin distance of 16 ± 1 Å. The uncertainties result from different runs of the ^{19}F -ENDOR experiment, which gave slightly different readings of the splitting (± 4 kHz variation), (SI 6b contains more details on the

uncertainty analysis). The spectrum of **GB1-c** did not resolve a splitting, consistent with a distance >17 Å.

GB1-b and **GB1-c** were measured at the standard concentration of 100 μM and at 500 μM . The data shown in Fig. 4d for **GB1-c** were obtained at a protein concentration of 500 μM . For both mutants **GB1-b** and **Gb1-c** a higher concentration yielded a slight improvement in sensitivity (and therefore faster accumulation) compared to the 100 μM data. We did not observe the 5-fold increase in sensitivity that was expected, perhaps due to some concentration-induced sample degradation via protein unfolding or aggregation which may have accelerated the $1/T_M$ relaxation rate. Notably, however, the ^{19}F -ENDOR spectra were indistinguishable at both concentrations, for both **GB1-b** and **GB1-c**.

Simulation of the **GB1** data by the fitting procedure established for **C_{D9K}** required additional modelling of the flexibility of the tag tether, which we approximated by a Gaussian-shaped interspin distance distribution. As the tether flexibility obscures any information about the tilt angle of the CF_3 group with respect to the $\text{Gd}^{3+}-\text{C}_{3\text{F}}$ vector, the tilt angle was excluded from the fitting procedure. In practice, any choice of the initial value of the ϕ angle resulted in convergence to a local minimum. Therefore, the fit was restricted to three parameters, including (i) the mean interspin distance \bar{r} , (ii) the standard deviation (σ) of the distance distribution and (iii) a Lorentzian linewidth lw . Similarly to **C_{D9K}**, the fitted mean interspin distance obtained for **GB1-a** and **GB1-b** matched the estimate from the peak splitting. It also matched the peak of the interspin distance distribution predicted from molecular modelling of the rotamer space sampled by the Gd^{3+} tag (see SI 6a). We note that the distance distribution and linewidth (lw) parameters are correlated and therefore not individually determinable.

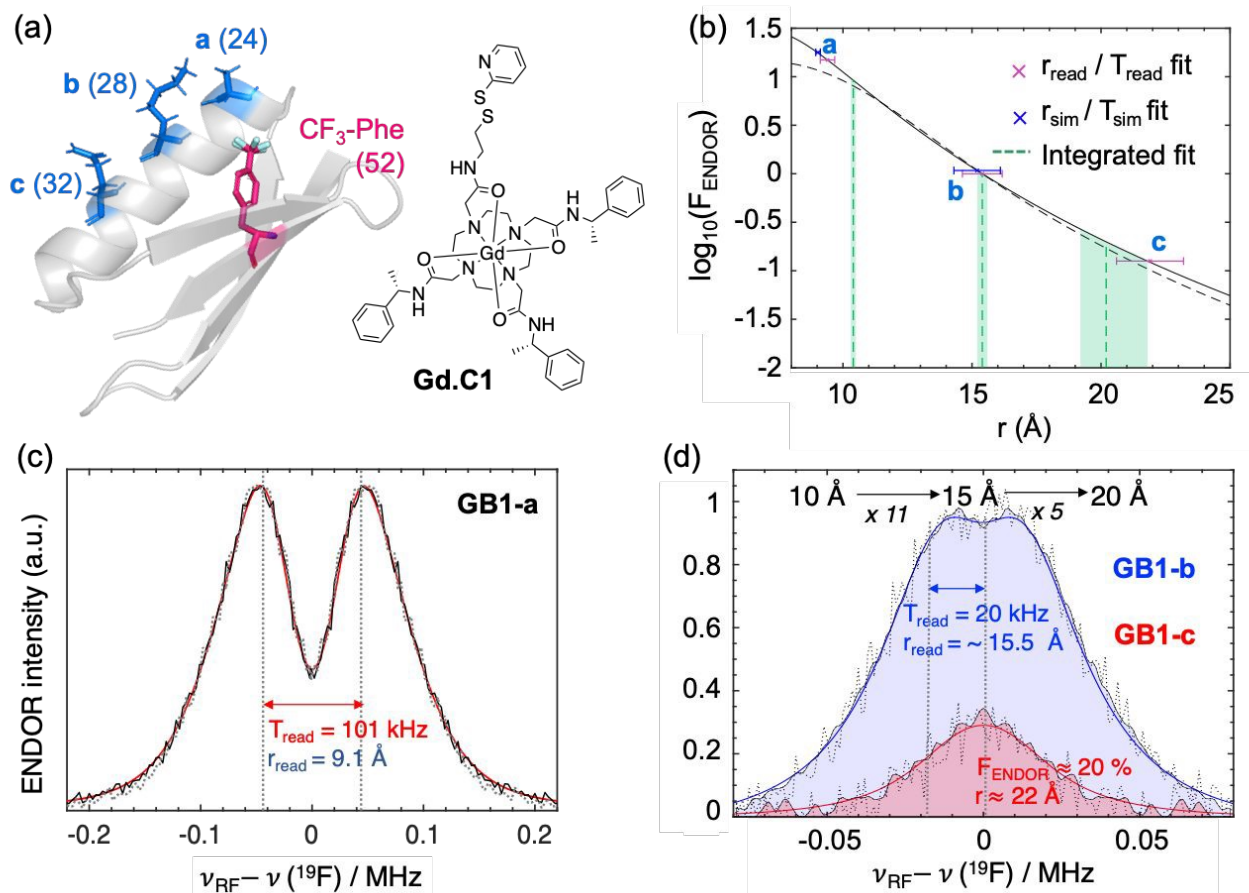


Fig. 4 ¹⁹F-ENDOR distance measurements on **GB1** containing CF₃-Phe and a **Gd.C1** tag. **(a)** Structure of **GB1-a,b,c**. The positions of the cysteine residues used for installing the Gd³⁺ tag **Gd.C1** (indicated beside the protein), are highlighted in blue and the CF₃-Phe residue in position 52 in magenta. **(b)** F_{ENDOR} vs Gd³⁺-¹⁹F₃ distance from Eq. 1 assuming singular distances (black trace) or a normal distribution about each distance with a standard deviation of 2.0 Å (black dotted trace). Gd³⁺-¹⁹F₃ distances were determined by the following methods: (i) r_{read} from directly measuring the ¹⁹F ENDOR line splitting T_{read} from each experimental spectrum (magenta scatter points), (ii) \bar{r}_{sim} from the coupling and \bar{T}_{sim} value modelled using the 3-parameter fitting procedure (blue scatter points), and (iii) the distance modelled via the integration method using Eq. 1 and scaling to the F_{ENDOR} value of the **GB1-b** spectrum ($F_{\text{ENDOR}}(\text{GB1-b}) = 1$), green dashed line). The error bars were estimated using either 3-fold lesser frequency resolution (magenta) or the rmsd value of the fit (blue) or the noise level percentage (green area). More details are given in the SI 6b. **(c)** and **(d)** are the Mims ENDOR experimental spectra (black traces) of **GB1-a** (panel c), **GB1-b** and **GB1-c** (panel d), measured using the parameters described in the experimental section. Accumulation times were 2 h for all **GB1**. The spectra were simulated (red/blue traces) using the approach described in the experimental section, with the parameters listed in Table 1. T_{read} refers to the hyperfine coupling in kHz determined from the resolved experimental splitting. In **(c)** the spectrum of **GB1-c** (red area) is scaled by the number of scans and echo intensity to **GB1-b**, resulting in a 0.19:1 amplitude ratio corresponding to a distance of 21.9 Å, assuming the ENDOR signal intensity scales directly with the ENDOR efficiency.

Interestingly, the integrated intensity of the ENDOR signal appears to follow the predicted $1/r^6$ Å distance dependence as described by Zänker et al.¹⁴ (Fig. 4b). A $1/r^6$ dependence is expected from substituting the hyperfine coupling (A) with $1/r^3$ (dipolar approximation) in the expression for the

ENDOR efficiency in the small angle limit, i.e. the right hand side of Eq. 1. Comparing the integrated intensities of the **GB1-a** and **GB1-b** ^{19}F -ENDOR spectra (measured with the respective optimized parameters), we find that the latter is approximately 9-fold smaller. This ratio is notably insensitive with regard to the uncertainty in the splitting found for **GB1-b**. If the distance in **GB1-b** is about 15 Å, the difference in signal intensity is consistent with a distance of 10.4 Å in **GB1-a**, which is close to that inferred from the line-splitting and 3-parameter modelling (9.4 ± 0.5 Å). Applying the same approach to the **GB1-b** and **GB1-c** data sets, we find that the ENDOR signal for the latter is 5-fold smaller, corresponding to an estimated distance of 21.9 Å for **GB1-c**, which is consistent with the maximum of the distance distribution obtained by molecular modelling (SI 5b). A more detailed analysis of this integration approach for ^{19}F -ENDOR distance measurement is currently in progress in our laboratory.

We note that in our hands, the integrated intensities of ^{19}F signals in the set of GB1 constructs were more reproducible than exact line splittings (see SI 6). Measured as a percentage of the echo, the integral values were well reproduced when using different operating frequencies/magnetic fields and for scans of different duration. The integration approach most likely works well in these systems as the ENDOR line is very narrow and orientation selection artefacts are absent.

Table 1. Fitting parameters used for simulations of the ^{19}F -ENDOR data for C_{D9k} and **GB1**. T_{read} is defined as the splitting taken from the experimental spectra corresponding to the ^{19}F - Gd^{3+} dipolar coupling in kHz, and r_{read} is the corresponding distance. \bar{r}_{sim} is the distance mean fitted by the simplex algorithm (and corresponding dipolar coupling \bar{T}_{sim}), which is fit alongside a Lorentzian Linewidth at full width half maximum (Lorentz. $lw.$) and either the F_3 -plane tilt angle ϕ for C_{D9k} , or a normal distribution width σ about \bar{r}_{sim} for **GB1-a-c**.

Protein	$r_{\text{read}} (\sigma) \pm \text{err.}^* / \text{Å}$	$\bar{r}_{\text{sim}} (\sigma) \pm \text{err.}^* / \text{Å}$	Lorentz. $lw.$ /kHz	ϕ /degrees	T_{read} /kHz	\bar{T}_{sim} /kHz
C_{D9k-1}	7.1 ± 0.4	$6.5 (-) \pm 0.3$	46.5	93	210 ± 3	270
C_{D9k-2}[†]	-	$16.0 (-)$	35	-	-	-
GB1-a	9.0 ± 0.1	$9.4 (1.2) \pm 0.3^*$	33	-	101 ± 3	90.9
GB1-b	15.5 ± 0.9	$16 (2.2) \pm 1.0^*$	32	-	20 ± 3	22.9
GB1-c	-	$21.9 (2-3.5) \pm 3.0^*$	30	-	$< 8.0^{**}$	-

* The uncertainty was determined as three times the experimental frequency resolution (as per the approach used by Bennati and co-workers^{15,16}) as a percentage of the resolved splitting. The resulting uncertainties in the distance domain are consistent with the variation of fitted/measured $\bar{T}_{\text{sim}}/T_{\text{read}}$ values across different batches of measurements on the same sample. See SI 6b for a full description of the uncertainty analysis. Note that the actual uncertainty is likely larger, as properly attributing the errors to the modelled values would need to consider the potential covariance of the fitting parameters.

** We attempted to fit a mean distance and distance distribution for the **GB1-c** data using the same simplex procedure as for the shorter distances associated with **GB1-a** and **GB1-b**, but fitting data with unresolved splitting in this way is unreliable as the distance distribution and linewidth could be arbitrarily adjusted to fit any distance. By constraining the linewidth to be no larger than the linewidth of **GB1-b**, the distance obtained for **GB1-c** was within the estimate made by integrating and scaling the ENDOR intensities as described in Section 2.3.

† The **C_{D9k}-2** mean distance \bar{r}_{sim} (and corresponding \bar{T}_{sim}) and linewidth lw , were fit using the same simplex algorithm as for **C_{D9k}-1**. The fit was also validated against the distance distribution modelled using PyParaTools. Two considerations were made in the simplex fitting: (i) the tilt angle ϕ was excluded from the simulation given the $> 15 \text{ \AA}$ interspin distance; (ii) the linewidth was constrained to be less than the linewidth fit for the shorter **C_{D9k}-1** distance.

2.4 Enhanced sensitivity with a Gd³⁺ versus nitroxide tag. To compare the sensitivities of the ¹⁹F-ENDOR experiment using Gd³⁺ or nitroxide tags, we used the **GB1-a** protein prepared in 100, 25 and 10 μM concentrations, and an analogue protein **GB1-aN** at 100 μM , where the **Gd.C1** tag was replaced with a nitroxide (MTSL) at the same cysteine residue (Fig. 5a). The interspin distance is predicted to be shorter in **GB1-aN**, by about 2 \AA (8 vs. 10 \AA). To directly compare the sensitivities of the measurements using the two tags, the ¹⁹F-ENDOR spectra were collected using exactly the same conditions and parameters ($\tau = 2.0 \mu\text{s}$, $t_{\pi/2} = 16 \text{ ns}$, $t_{\text{RF}} = 54 \mu\text{s}$, integration window = 20 ns, shot repetition time = 5 ms, $B_o = B_{\text{max}}$ (field position maximum of the Gd³⁺ EPR spectrum), resolution = 5 kHz), with the exception of the temperature (10 K vs 50 K for **GB1-a** and **GB1-aN** respectively). All spectra (Fig. 5b-c) were collected with 500 scans, amounting to about 45 minutes of measurement time. The **GB1-aN** ENDOR spectra were collected at the maximum of the EPR spectrum. In Fig. 5b, a baseline subtraction was applied to each spectrum, and they are all presented normalized to the largest signal, such that the maximum of the 100 μM **GB1-a** ENDOR signal is 1. Fig. 5b shows that the Gd³⁺ label indeed afforded a 2.7-fold gain in signal-to-noise ratio over the nitroxide label for the same concentration, even though the Gd–CF₃ distance

is expected to be longer than the nitroxide–CF₃ distance.ⁱⁱ In terms of concentration sensitivity, we find that Gd³⁺ measurements allow approximately 10 x lower sample concentrations to be used. Fig. 5c compares the **GB1-a** dilutions shows that the nitroxide-tagged **GB1-aN** has a S/N comparable to the 10 μM Gd³⁺-tagged protein. Notably, the detection settings used were optimal for **GB1-aN** ($S=1/2$). Using the settings optimized for Gd–F₃ distance measurement, i.e. $t_{\pi/2} = 12$ ns and a shot repetition time of 1.5 ms, a spectrum of approximately the same quality can be measured for **GB1-a** in about 16 min. The lowest concentration of Gd³⁺-tagged protein that could be measured whilst still reliably resolving a splitting was 5 μM. The S/N achieved in Fig. 5d resulted from 11 hrs of measurement under the same pulse sequence settings as described above, but over a smaller 0.8 MHz frequency range.

ⁱⁱ We expect about a 4-fold loss in signal intensity from this effect for **GB1-a** with a designed interspin of ~10 Å (9.4 ± 0.3 Å fit) as compared **GB1-aN**, with a designed interspin distance of ~±8 Å. Consequently, for the same interspin distance, one would expect about a 10-fold signal-to-noise improvement from using Gd³⁺ labels as compared to nitroxides. An order of magnitude enhancement is consistent with the gain we estimate by considering the lower temperature of the measurements with the Gd³⁺ tag, the higher spin of the Gd³⁺ ion ($S=7/2$) and similar T_M times for both tags.

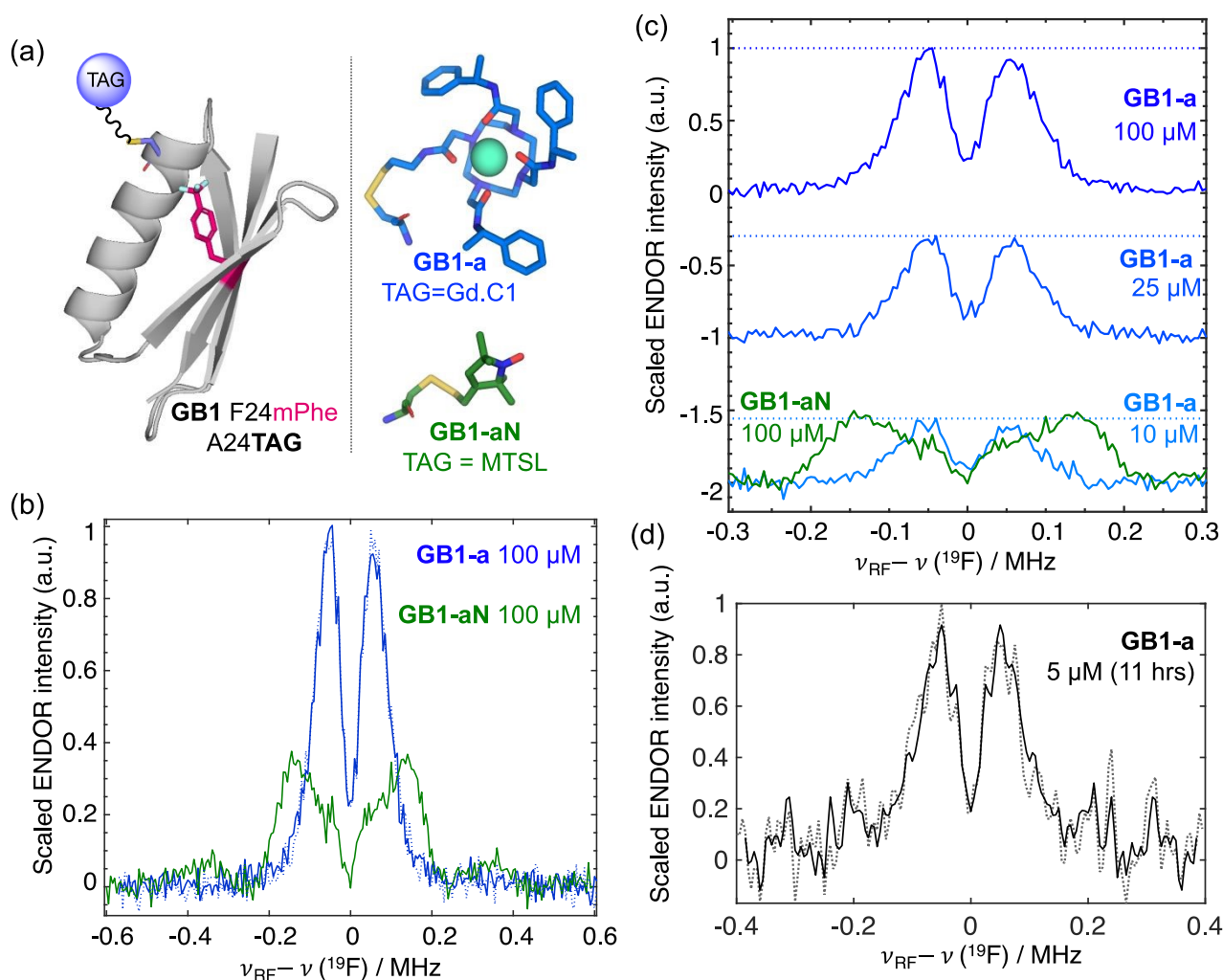


Fig. 5 Sensitivity comparison between ^{19}F -ENDOR distance measurements using **Gd.C1** and **MTSL** tags. **(a)** Structure of **GB1-a** and **GB1-aN** showing the CF_3 -Phe residue in position 52 (magenta) and the paramagnetic tag in position 24. The right panel shows the chemical structures of the tags including the cysteine residue to which the tags are attached on the protein. **(b)** Mims ENDOR spectra of **GB1-a** (blue trace) and **GB1-aN** (green trace), scaled to the number of scans, noise level and microwave power. The ENDOR spectra were both measured using the interpulse (detection) delay $\tau = 2.0 \mu\text{s}$ and $t_{\text{RF}} = 30 \mu\text{s}$, collected at the respective field maxima of the gadolinium and nitroxide EPR spectra. The $\pi/2$ detection pulses were 12 ns and 16 ns for **GB1-a** and **GB1-b** respectively. The **GB1-a** and **GB1-aN** data were collected at 10 and 50 K, respectively. **(c)** Raw data showing the scaled comparison of dilutions of **GB1-a** (all measured under the same conditions as described for **b**) compared to the $100 \mu\text{M}$ **GB1-aN** protein, which is shown against the $10 \mu\text{M}$ **GB1-a** spectrum having the most similar S/N. **(d)** The lowest concentration at which a splitting could be resolved for the 10 \AA distance in **GB1-a** was $5 \mu\text{M}$. The data was collected over 15,000 scans, amounting to 11 h of measurement time for a 0.8 MHz frequency sweep. The symmetrized smoothed data is shown in the solid trace, the raw data in the dotted trace.

Importantly, the **GB1-aN** spectrum cannot immediately be used to determine the interspin distance as it does not display the full Pake pattern. Sampling over the full Pake pattern would considerably increase the accumulation time needed. Nonetheless, the CF_3 label appears to be insensitive to orientation effects

at short interspin distances ($< 10 \text{ \AA}$), as the dipolar vectors between the ^{19}F spins of the CF_3 group are orientated differently ($\pm 120^\circ$) relative to the g -tensor frame of the nitroxide. As a result, an ENDOR spectrum measured at the centre of the EPR spectrum must resolve both the parallel and perpendicular components of the hyperfine tensors, because at least two of the electron– ^{19}F vectors contain projections onto both the perpendicular and parallel axes of the g -frame. Therefore, both the edges and central peaks of the ENDOR Pake pattern can be extracted at a single field position. The experimental details and results demonstrating the use of a CF_3 tag as a strategy for rapid Pake pattern accumulation are further discussed in the SI 7.

2.4 Well defined Gd^{3+} – ^{19}F models for benchmarking ^{19}F -ENDOR distance measurements. The reliability of DEER and RIDME experiments has been extensively validated using sets of spectroscopic rulers, which are small molecules capped by EPR spin labels specifically constructed to achieve well-defined interspin distances.^{30–32} We performed a similar evaluation of the ^{19}F -ENDOR experiment with the Gd^{3+} – ^{19}F ruler Gd^{III} -EPEP-F carrying the Gd^{3+} complex Gd^{III} (PyMTA) and a ^{19}F atom connected via a spacer consisting of ethynylene (E) and *para*-phenylene (P) units.³¹ The ruler was expected to result in a very narrow Gd^{3+} – ^{19}F distance distribution. In addition, it exhibits a relatively long electronic T_M time of about 11 μs , which should allow the resolution of long distances by the ^{19}F -ENDOR spectrum.

Disappointingly, the ruler did not yield sensible distance measurements. The Gd^{3+} – ^{19}F ruler has an interspin distance of 19.7 Å , which corresponds to a 10 kHz splitting. However, the experimentally determined splitting was instead 75 kHz, corresponding to a distance of 10 Å . We suspect that the conjugation that affords the molecule its rigidity also leads to a large through-bond isotropic hyperfine coupling component of the order of the dipole coupling, as was found for the compounds described by Stoll and co-workers³³ and in the trityl ^{19}F -ENDOR examples.²¹ The wings of the spectrum in Fig. 6b are also partially suppressed and do not resolve the T_\perp feature as would be expected for a single distance associated with a purely through-space dipolar coupling, again pointing to a significant isotropic contribution to the hyperfine interaction. This parallels the problems encountered by the authors of the recent publication on ^{19}F -ENDOR with a trityl radical paramagnetic spin tag,²¹ which required more complicated computational modelling to validate the observed splittings. Whether interactions other than conjugation contribute to the splitting observed in the ^{19}F -ENDOR spectrum of the Gd^{3+} – ^{19}F ruler is open

to further investigation. We are currently trying to synthesize new molecules in which the conjugation is disrupted without affecting the conformational rigidity of the ruler, as achieved in refs.^{15, 16}

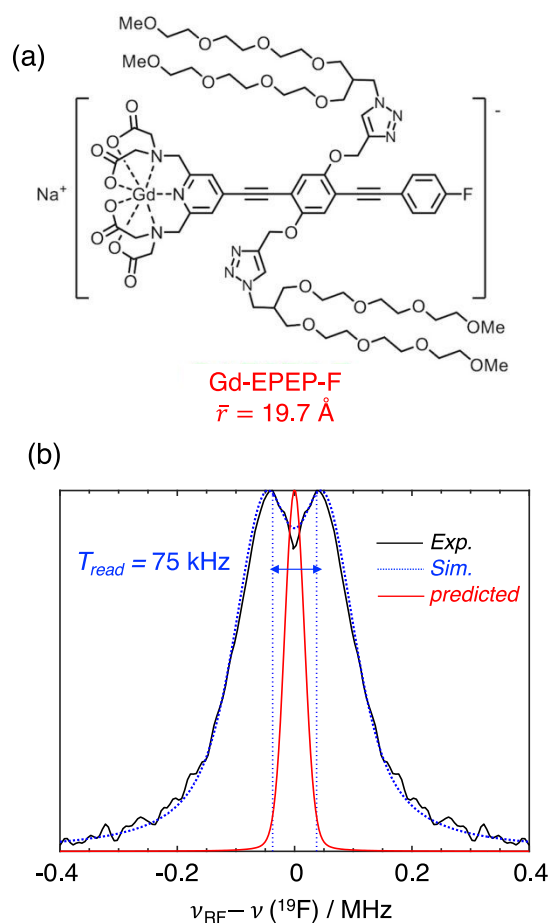


Fig. 6 Mims ^{19}F -ENDOR measurements on the Gd-EPEP-F molecular ruler. **(a)** Chemical structure of the $\text{Gd}^{3+}\text{-}^{19}\text{F}$ ruler used to test the upper detection limit of the ^{19}F -ENDOR distance measurement. **(b)** Experimental Mims ENDOR spectrum (black trace) of the $\text{Gd}^{3+}\text{-}^{19}\text{F}$ ruler, measured using $\tau = 3.0$ s, $t_{\text{RF}} = 45$ μs , $\pi/2 = 12$ ns, with the RF power (100 W source) optimized by nutation referenced to the ^1H line, in stochastic mode at 10 K. The blue trace shows a least-squares best fit, resulting in the parameters $\bar{r} = 8.9$ \AA , lw (Lorentzian convolution) = 33 kHz, lw (Gaussian EasySpin parameter) = 89 kHz. The red trace is a simulation of the predicted Mims ENDOR spectrum for a 19.7 \AA distance with the same linewidth parameters.

3. Discussion

Overall, our results demonstrate the different advantages and disadvantages of using Gd^{3+} ions as paramagnetic centres for ^{19}F -ENDOR distance measurements. Importantly, significant sensitivity gains are achieved for both short and long distance measurements by using Gd^{3+} , both in terms of measurement time and flexibility in sample concentration.

Compared to other tags reported in the literature, Gd^{3+} tags arguably possess better relaxation properties than nitroxides and trityls (similar T_M times $\approx 10 \mu\text{s}$, faster T_1 relaxation $\approx 100 \mu\text{s}$), and are also free of orientation selection effects. These considerations combined, shorten the accumulation time dramatically. A fairly direct comparison can be made between the 15 Å distance measurements on the nitroxide-tagged proteins in ref.,¹⁵ and the Gd^{3+} -tagged **GB1-b** protein reported here. The nitroxide-tagged proteins reported in the literature were measured at $\sim 250 \mu\text{M}$ concentrations used 30 h measurement time for a 14-15 Å distance (and 18 h for 10-12 Å distances).¹⁵ Using Gd^{3+} we can achieve comparable signal-to-noise in 2-4 h for a 15-16 Å distance (1-2 h for shorter distances), at much lower (25-100 μM) concentrations. The sensitivity enhancement of ^{19}F -ENDOR by Gd^{3+} detection allows us to reliably resolve signals using concentrations down to 5 μM , which nears typical physiological ($<1 \mu\text{M}$ ¹⁸) protein concentrations. In practice, setting up low concentration EPR measurements (at low temperatures), will always be time consuming as it requires careful stabilization of helium flow over time and minimizing RF heating artefacts to prevent phase drifting and signal cancellation. Nevertheless, amongst the tags in the currently reported literature, Gd^{3+} labelling is the most promising avenue for efficiently measuring ^{19}F -ENDOR distance experiments at physiological concentrations.

3.1 Short distance measurement and orientation information using Gd^{3+} labels. Our results show that in the case of conformationally unambiguous labelling with Gd^{3+} , such as a Gd^{3+} ion bound at an intrinsic metal binding site of a metalloprotein, the Gd^{3+} - CF_3 pair can be used to measure short distances

and gain information on the relative orientation of the CF_3 nuclear spin tag. The maximum accessible distance by this approach is about 12 Å, where this upper limit is constrained by the relatively short electronic T_M of protein-bound Gd^{3+} .

However, as proteins rarely contain natural binding sites with good affinity for Gd^{3+} ions, this approach may not be widely generalizable. Nevertheless, there is a host of other metal ions that can be used instead of Gd^{3+} , such as Mn^{2+} and Cu^{2+} , which can functionally replace EPR-silent metal ions in biological systems (Mg^{2+} , Ca^{2+} and low-spin Fe^{2+} etc.) that are also widely used as DEER tags.^{34–38} These different metal tags have the potential to be used for the determination of short distances and orientation information, offering different advantages and disadvantages. While Gd^{3+} will always be the most sensitive tag, as it has the highest electronic spin state ($S = 7/2$) in addition to a narrow EPR line which can be fully excited, we envisage other metal ions could be used to implement the ^{19}F -ENDOR experiment albeit with lower overall sensitivity.

3.2 Gd^{3+} labels as sensitive probes for longer distances. Attaching Gd^{3+} tags to the protein allowed us to efficiently access longer distances in the 15–20 Å range, comparable to those reported for nitroxide tags¹⁵ (≈ 15 – 16 Å). There is still significant scope for achieving more accurate distance measurements by using Gd^{3+} tags with shorter or more rigid tethers,³⁹ to reduce the contribution of the tether flexibility to the distance distributions.

We observed that the integrated intensities of the experimental ^{19}F -ENDOR signals in this study broadly follow the predicted $1/r^6$ dependence. This suggests that the ^{19}F -ENDOR sensitivity need not be constrained by the ENDOR linewidth, allowing much longer distances to be measured, here up to 20 Å, but potentially extending to ≈ 25 Å with the right tags and optimisation. The distances measured using the integration method are approximate, as without a resolved ENDOR splitting, the distance, distance distribution and linewidth parameters become correlated in the fit and cannot be determined on their own as individual parameters. Nonetheless, the uncertainties can be expected to be limited to a couple of Ångströms for the longest distances accessible by this measurement.

A major drawback of the integration method is the need for calibration, i.e., a protein with a long distance needs to be referenced against a protein engineered with a shorter distance (10–15 Å). In the cases of **GB1-a-c**, we could directly compare the integrated intensities of the ^{19}F -ENDOR signals at different distances because i) the EPR lineshape and electronic relaxation properties of the Gd^{3+} tag were conserved and ii) the nuclear (T_{1n}) relaxation time of the ^{19}F nuclear label could be assumed to be conserved as well. If either of these conditions fails, the comparison of the integrated intensities becomes unreliable. Previous experience from using **Gd.C1** in DEER applications^{28,39} suggests that its electronic

relaxation properties are invariant in a variety of proteins. It is less clear what governs the relaxation properties of the nuclear ^{19}F label (see section 3.3 below). In the tagged proteins measured here, we observe no change in the nuclear relaxation rate as a function of its distance from the Gd^{3+} tag on the timescale of the ENDOR experiment, (at least over the 0–25 Å range). These observations would suggest that the **GB1-a-c** proteins could be used as models to calibrate measurements performed on other labelled proteins. Work is in progress to test this hypothesis. We anticipate that the ideal protein for calibration has an interspin distance of about 15 Å, which is the limit for distance measurements based on the peak splitting. Proteins with shorter interspin distances are expected to deviate from the $1/r^6$ dependence, as they are necessarily further away from the small angle approximation. They will also be more sensitive to any distance distribution, with shorter distances of the distribution contributing more to the overall integrated intensity (Fig. 2c) leading to increasing deviation from the $1/r^6$ relationship (Fig. 4b). As noted in the Results section 2.3, further assessment of the integration method to measure Gd^{3+} – ^{19}F interspin distances is needed to test its reliability and general applicability. This forms ongoing work of our laboratory.

It is important to note that the integrated intensity of the ENDOR signal is referenced to the **EPR** signal intensity. Therefore, there is no need to measure different samples at exactly the same concentration. Similarly, incomplete tagging of a protein or the presence of unbound Gd^{3+} tag will not affect these measurements, so long as the relative concentration of the unbound tag is known. Higher protein concentrations (>100 μM) can be used in the ENDOR experiment than those typically used in DEER experiments, boosting sensitivity, which can be advantageous for measuring isolated protein systems. Closer intermolecular spacing of labelled molecules can be tolerated because each protein molecule only carries a single paramagnetic tag. In our hands, however, raising the concentration much past 250 μM appeared to increase (speed up) the electronic $1/T_1$ rates significantly.

3.3 ^{19}F nuclear relaxation. For a useful ^{19}F -ENDOR experiment, the nuclear ^{19}F label needs to feature spin-lattice relaxation (T_{1n}) times of the order of a millisecond. This is short for a nuclear spin in the solid state at cryogenic temperatures (10 K).⁴⁰ At present, we have no evidence that the $T_1(^{19}\text{F})$ relaxation time significantly depends on the distance of the nuclear label from the spin tag, suggesting that the fast nuclear relaxation is not driven by *direct* coupling to the electronic spin. A likely origin of the fast $1/T_{1n}$ rates is in the dipolar couplings of the ^{19}F spins with nearby ^1H spins, which are not saturated during the experiment. In these circumstances, the $T_1(^{19}\text{F})$ relaxation can be much faster than that measured for ^1H , as it is a T_1 value of a single spin, as opposed to where all spins of the system are saturated or inverted.

In solid-state NMR of proteins, methyl (CH_3) group rotation presents the main source of nuclear ^1H spin-lattice relaxation.⁴¹ In contrast, CF_3 groups are not expected to rotate at low temperature, but as the paramagnetic tag enhances the relaxation rates of the protons in its vicinity, ^1H – ^1H dipolar couplings transmit this relaxation enhancement to the protons near the ^{19}F label. The ^{19}F label is thus embedded in a rapidly relaxing ^1H spin bath. Work is in progress to better understand the factors that determine the relaxation properties of the nuclear label by using a modified *Tidy* ENDOR sequence.⁴² Rapid nuclear relaxation is important because slow nuclear relaxation at long distances from the electronic tag would lead to an underestimation of the ENDOR efficiency and, as a consequence, an overestimation of the interspin distance.

4. Conclusions. This work has shown that Gd^{3+} complexes are an attractive substitute for nitroxide radicals as the paramagnetic label for the ^{19}F -ENDOR experiment, providing benefits in terms of sensitivity and avoiding orientation selection effects at high magnetic fields. Using different labelling strategies, distances up to 15 Å can readily be measured using the resolved line splitting. There is also potential for extending this method to distances up to 25 Å through signal integration. With further optimization, we believe the ^{19}F -ENDOR experiment could be developed into a useful screening tool for monitoring binding interactions and conformational changes. For example, drug development frequently encounters difficult to crystallize target proteins and tagging of the target protein with a paramagnetic tag and the drug with ^{19}F could present a practical strategy for experimental verification of molecular models by quick ^{19}F -ENDOR measurements. We are currently exploring these applications.

5. Materials and Methods

Sample synthesis and labelling. Details of the protein and small molecule syntheses and labelling protocols along with characterization by mass spectrometry and NMR are provided in the supporting information.

EPR characterisation. Samples were prepared for W-band EPR measurements in 130–150 μM concentrations in deuterated buffer (50 mM MES in D_2O , pH 6.5, 50 mM NaCl) and 10% deuterated glycerol, then transferred to W-band quartz capillaries (0.5 mm ID, 0.9 mm OD). EPR measurements were performed on a modified E680 Bruker spectrometer⁴³ equipped with a TE_{011} -mode EPR/ENDOR resonator in a helium gas flow cryostat (Oxford Instruments) to run measurements at 50 K for nitroxide-tagged systems and 10 K for gadolinium-tagged systems. The EPR lineshapes were recorded using a 2-pulse Hahn echo sequence $\pi/2 - \tau - \pi - \tau - \text{echo}$, with $\pi/2 = 40$ ns, $\pi = 80$ ns and $\tau = 500$ ns. The electronic T_1 relaxation was measured by the inversion–recovery sequence $\pi - t + dt - \pi/2 - \tau - \pi - \text{echo}$, recording the integrated echo intensity as a function of the delay time t (incremented with the time

step dt), with the same pulse lengths as above. T_M relaxation was measured by recording the decay of the integrated echo intensity with time, using the pulse sequence $\pi/2 - t + dt - \pi - t + 2dt - \text{echo}$. Details of the EPR characterization and relaxation properties of all the tagged protein systems are provided in SI 4.

ENDOR measurements. ^{19}F -ENDOR data were measured using the Mims ENDOR pulse sequence $\pi/2 - \tau - \pi/2 - t_{\pi, \text{RF}} - \pi/2 - \tau - \text{echo}$, with $\pi/2 = 12$ ns, $\tau = 2.0$ μs (for **C_{D9k}**) or 2.5 μs (for the **GB1** samples) and $t_{\pi, \text{RF}} = 45$ μs for short distances in **C_{D9k}**, 60 μs for short distances in **GB1**, and 100-120 μs for long (> 15 Å) distances. The RF power (source 100 W) was adjusted to achieve the maximum ENDOR response. The attenuation (0-3 dB for short distance measurements settings and 6 dB for long distance settings) was determined by microwave nutation, using the proton line for calibration. The ENDOR spectra were acquired in stochastic mode, with 1 shot per (sample) point and a 4 ms shot repetition time in the Gd^{3+} systems, and 5 ms shot repetition time in the nitroxide system. Each ENDOR spectrum was conducted in batches, as per the approach by Kehl et al.¹⁶ due to phase drifts over extended periods of time. Each batch took 1–2 h with phase readjustment in between. The accumulation time was 1 h for **GB1-a**, 3 h for **GB1-b**, 7 h for **GB1-c**, 5 h in total for **C_{D9k-1}** (about 1 h per τ value), 1.5 h for **C_{D9k-2}** and 45 min – 1.5 h per field position in the nitroxide echo-detected EPR spectrum in **GB1-aN**.

Processing of ^{19}F -ENDOR spectra. When performing scaled intensity comparisons of the experimental ENDOR data, the maximum intensity of the ^{19}F -ENDOR intensity was scaled a percentage of the EPR echo intensity, i.e. the ^{19}F -ENDOR baseline value. The spectra were then symmetrised about the central ^{19}F frequency corresponding to the given magnetic field position. The spectra are presented unsmoothed unless stated otherwise in the figure caption.

Simulations and fitting. *EasySpin simulations:* Mims ENDOR simulations were performed using the *EasySpin* program *saffron*.⁴⁴ Simulations were performed using a grid size of 200 knots, with the simulated spectrum resolution set to match that of the corresponding experimental data. To speed up calculations, the excitation bandwidth was omitted from the Gd ENDOR simulations. To include the effect of an interspin distance distribution, the *saffron* calculation was looped over the distance axis, scaling each individual distance contribution by the distribution function $P(r)$. The current version of *saffron* only allows Gaussian line broadening, which did not fit our experimental data well. Therefore, to account for Lorentzian line broadening of the ENDOR lineshape, we fixed the *saffron* Gaussian linewidth parameter *Sys.lwEndor* to 1 kHz and convoluted the sum ENDOR spectrum with a fitted Lorentzian linewidth using the function *convspec* in MATLAB.

The systematic search for the best fit used the MATLAB simplex minimisation procedure *fminsearch* to vary the set of parameters (\bar{r} , σ , ϕ and lw) as discussed in the main text, minimising the standard deviation of the simulated spectrum to converge at a cut-off of a 10^{-4} root-mean-square deviation (RMSD). Where a system was measured at multiple τ values (as for the short distance construct **C_{D9K-1}**), the fit was optimised to all n spectra simultaneously.

Rotamer simulations (PyParaTools). To simulate the inter-spin distance distributions between the **Gd.C1** and CF₃-Phe labels in the **GB1** model protein systems, the program PyParaTools²⁸ was used. The protein coordinates (PDB ID: 1PGB⁴⁵) were kept fixed, while the dihedral angles of the Gd.C1 tag and CF₃-Phe non-canonical amino acid were varied in random combinations (the conformations generating steric clashes with the protein were excluded. More details are included in SI 6a.

Conflicts of Interest

There are no conflicts of interest to declare.

Acknowledgements

The authors would like to thank the ANU Research School of Chemistry workshop staff for their ongoing support with hardware maintenance. We acknowledge the support of the Australian Research Council DP210100088 (NC, TH, JH, GO) and the Westpac Foundation Future Leaders Scholarship (MJ). GO thanks the Australian Research Council for a Laureate Fellowship (FL170100019) and support through the Centre of Excellence for Innovations in Peptide & Protein Science (CE200100012).

References

- 1 A. D. Milov, A. B. Ponomarev and Y. D. Tsvetkov, *Chem. Phys. Lett.*, 1984, **110**, 67–72, DOI: 10.1016/0009-2614(84)80148-7.
- 2 A. D. Milov, A. G. Maryasov and Y. D. Tsvetkov, *Appl. Magn. Reson.*, 1998, **15**, 107–143, DOI: 10.1007/BF03161886.
- 3 A. D. Milov, K. M. Salikhov and M. D. Shirov, *Fiz. Tverd. Tela*, 1981, **23**, 975–982.
- 4 M. Pannier, S. Veit, A. Godt, G. Jeschke and H. W. Spiess, *J. Magn. Reson.*, 2000, **142**, 331–340, DOI: 0.1006/jmre.1999.1944.
- 5 S. Milikisyants, F. Scarpelli, M. G. Finiguerra, M. Ubbink and M. Huber, *J. Magn. Reson.*, 2009, **201**, 48–56, DOI: 10.1016/j.jmr.2009.08.008.
- 6 P. P. Borbat, A. J. Costa-Filho, K. A. Earle, J. K. Moscicki and J. H. Freed, *Science (80-.)*, 2001, **291**, 266–269, DOI: 10.1126/science.291.5502.266.
- 7 S. S. Eaton and G. R. Eaton, *J. Amer. Chem. Soc.*, 1982, **104**, 5002–5003, DOI: 10.1021/ja00382a068.
- 8 S. S. Eaton, K. M. Moore, B. M. Sawant and G. R. Eaton, *J. Am. Chem. Soc.*, 1983, **105**, 6560–6567, DOI: 10.1021/ja00360a005.
- 9 P. P. Borbat, E. R. Georgieva and J. H. Freed, *EPR Newsl.*, 2007, **17**, 21–33.
- 10 G. Jeschke, *Annu. Rev. Phys. Chem.*, 2012, **63**, 419–446, DOI: 10.1146/annurev-physchem-032511-143716.
- 11 D. T. Edwards, Z. Ma, T. J. Meade, D. Goldfarb, S. Han and M. S. Sherwin, *Phys. Chem. Chem. Phys.*, 2013, **15**, 11313–11326, DOI: 10.1039/C3CP43787F.
- 12 J. A. Clayton, M. Qi, A. Godt, D. Goldfarb, S. Han and M. S. Sherwin, *Phys. Chem. Chem. Phys.*, 2017, **19**, 5127–5136, DOI: 10.1039/C6CP07119H.
- 13 M. D. Rabenstein and Y. K. Shin, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, **92**, 8239–8243, DOI: 10.1073/pnas.92.18.8239.
- 14 P.-P. Zänker, G. Jeschke and D. Goldfarb, *J. Chem. Phys.*, 2004, **122**, 24515, DOI: 10.1063/1.1828435.

- 15 A. Meyer, S. Dechert, S. Dey, C. Höbartner and M. Bennati, *Angew. Chemie Int. Ed.*, 2020, **59**, 373–379, DOI: 10.1002/anie.201908584.
- 16 A. Kehl, M. Hiller, F. Hecker, I. Tkach, S. Dechert, M. Bennati and A. Meyer, *J. Magn. Reson.*, 2021, **333**, 107091, DOI: 10.1016/j.jmr.2021.107091.
- 17 W. B. Mims, *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.*, 1965, **283**, 452–457, DOI: 10.1098/rspa.1965.0034.
- 18 A. Meyer, A. Kehl, C. Cui, F. A. K. Reichardt, F. Hecker, L.-M. Funk, M. K. Ghosh, K.-T. Pan, H. Urlaub, K. Tittmann, J. Stubbe and M. Bennati, *J. Am. Chem. Soc.*, , DOI:10.1021/jacs.2c02906.
- 19 J. T. Gerig, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1994, **26**, 293–370, DOI: 10.1016/0079-6565(94)80009-X.
- 20 J. L. Kitevski-LeBlanc and R. S. Prosser, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2012, **62**, 1–33, DOI: 10.1016/j.pnmrs.2011.06.003.
- 21 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 and E. G. Bagryanskaya, *Phys. Chem. Chem. Phys.*, 2022, **24**, 5982–6001, DOI: 10.1039/D1CP05445G.
- 22 A. Potapov, H. Yagi, T. Huber, S. Jergic, N. E. Dixon, G. Otting and D. Goldfarb, *J. Am. Chem. Soc.*, 2010, **132**, 9040–9048, DOI: 10.1021/ja1015662.
- 23 A. Potapov, Y. Song, T. J. Meade, D. Goldfarb, A. V Astashkin and A. Raitsimring, *J. Magn. Reson.*, 2010, **205**, 38–49, DOI: 10.1016/j.jmr.2010.03.019.
- 24 H. W. Orton, H. Qianzhu, E. H. Abdelkader, E. I. Habel, Y. J. Tan, R. L. Frkic, C. J. Jackson, T. Huber and G. Otting, *J. Am. Chem. Soc.*, 2021, **143**, 19587–19598, DOI: 10.1021/jacs.1c10104.
- 25 C. Jäckel, M. Salwiczek and B. Koksich, *Angew. Chemie Int. Ed.*, 2006, **45**, 4198–4203, DOI: 10.1002/anie.200504387.
- 26 J. A. Olsen, D. W. Banner, P. Seiler, B. Wagner, T. Tschopp, U. Obst-Sander, M. Kansy, K. Müller and F. Diederich, *Chem. Bio. Chem.*, 2004, **5**, 666–675, DOI: 10.1002/cbic.200300907.
- 27 L. A. Svensson, E. Thulin and S. Forsén, *J. Mol. Biol.*, 1992, **223**, 601–606, DOI: 10.1016/0022-

2836(92)90976-Q.

- 28 H. Yagi, D. Banerjee, B. Graham, T. Huber, D. Goldfarb and G. Otting, *J. Am. Chem. Soc.*, 2011, **133**, 10418–10421, DOI: 10.1021/ja204415w.
- 29 B. Graham, C. T. Loh, J. D. Swarbrick, P. Ung, J. Shin, H. Yagi, X. Jia, S. Chhabra, N. Barlow, G. Pintacuda, T. Huber and G. Otting, *Bioconjug. Chem.*, 2011, **22**, 2118–2125, DOI: 10.1021/bc200353c.
- 30 A. Dalaloyan, M. Qi, S. Ruthstein, S. Vega, A. Godt, A. Feintuch and D. Goldfarb, *Phys. Chem. Chem. Phys.*, 2015, **17**, 18464–18476, DOI: 10.1039/c5cp02602d.
- 31 G. Jeschke, M. Sajid, M. Schulte, N. Ramezani, A. Volkov, H. Zimmermann and A. Godt, *J. Am. Chem. Soc.*, 2010, **132**, 10107–10117, DOI: 10.1021/ja102983b.
- 32 S. Razzaghi, M. Qi, A. I. Nalepa, A. Godt, G. Jeschke, A. Savitsky and M. Yulikov, *J. Phys. Chem. Lett.*, 2014, **5**, 3970–3975, DOI: 10.1021/jz502129t.
- 33 S. Pribitzer, D. Mannikko and S. Stoll, *Phys. Chem. Chem. Phys.*, 2021, **23**, 8326–8335, DOI: 10.1039/D1CP00229E.
- 34 A. Martorana, Y. Yang, Y. Zhao, Q.-F. Li, X.-C. Su and D. Goldfarb, *Dalt. Trans.*, 2015, **44**, 20812–20816, DOI: 10.1039/C5DT04123F.
- 35 F. D. Breitgoff, K. Keller, M. Qi, D. Klose, M. Yulikov, A. Godt and G. Jeschke, *J. Magn. Reson.*, 2019, **308**, 106560, DOI: 10.1016/j.jmr.2019.07.047.
- 36 A. Gamble Jarvi, K. Ranguelova, S. Ghosh, R. T. Weber and S. Saxena, *J. Phys. Chem. B.*, 2018, **122**, 10669–10677, DOI: 10.1021/acs.jpccb.8b07727.
- 37 A. M. Bowen, M. W. Jones, J. E. Lovett, T. G. Gaule, M. J. McPherson, J. R. Dilworth, C. R. Timmel and J. R. Harmer, *Phys. Chem. Chem. Phys.*, 2016, **18**, 5981–5994, DOI: 10.1039/C5CP06096F.
- 38 K. Keller, M. Zalibera, M. Qi, V. Koch, J. Wegner, H. Hintz, A. Godt, G. Jeschke, A. Savitsky and M. Yulikov, *Phys Chem Chem Phys*, 2016, **18**, 25120–25135, DOI: 10.1039/C6CP04884F.
- 39 I. D. Herath, C. Breen, S. H. Hewitt, T. R. Berki, A. F. Kassir, C. Dodson, M. Judd, S. Jabar, N. Cox, G. Otting and S. J. Butler, *Chemistry*, 2021, **27**, 13009–13023, DOI: 10.1002/chem.202101143.

- 40 K. R. Thurber and R. Tycko, *J. Magn. Reson.*, 2008, **195**, 179–186, DOI: 10.1016/j.jmr.2008.09.015.
- 41 P. A. Beckmann and A. L. Rheingold, *J. Chem. Phys.*, 2016, **144**, 154308, DOI: 10.1063/1.4944981.
- 42 J. J. L. Morton, N. S. Lees, B. M. Hoffman and S. Stoll, *J. Magn. Reson.*, 2008, **191**, 315–321, DOI: 10.1016/j.jmr.2008.01.006.
- 43 A. Nalepa, K. Möbius, W. Lubitz and A. Savitsky, *J. Magn. Reson.*, 2014, **242**, 203–213, DOI: 10.1016/j.jmr.2014.02.026.
- 44 S. Stoll and A. Schweiger, *J. Magn. Reson.*, 2006, **178**, 42–55, DOI: 10.1016/j.jmr.2005.08.013.
- 45 T. Gallagher, P. Alexander, P. Bryan and G. L. Gilliland, *Biochemistry*, 1994, **33**, 4721–4729, DOI: 10.1021/bi00181a032.