Extending the range of distances accessible by $^{19}$F electron-nuclear double resonance in proteins using high-spin Gd(III) labels.

Alexey Bogdanov$^1$*, Veronica Frydman$^2$, Manas Seal$^1$, Leonid Rapatskiy$^3$, Alexander Schnegg$^3$, Wenkai Zhu$^4$, Angela M. Gronenborn$^4$, Daniella Goldfarb$^1$*.

1 Department of Chemical and Biological Physics, The Weizmann Institute of Science, P. O. Box 26, Rehovot, 7610001, Israel
2 Department of Chemical Research Support, The Weizmann Institute of Science, P. O. Box 26, Rehovot, 7610001, Israel
3 Max Planck Institute for Chemical Energy Conversion, 34-36 Stiftstraße, Mülheim an der Ruhr, 45470, Germany
4 Department of Structural Biology, University of Pittsburgh, 4200 Fifth Ave, Pittsburgh, PA 15260, United States

*E-mail: daniella.goldfarb@weizmann.ac.il, alexey.bogdanov@weizmann.ac.il

$^{19}$F ENDOR

https://doi.org/10.26434/chemrxiv-2023-hg5zc ORCID: https://orcid.org/0000-0003-1044-4783 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0
Abstract
Fluorine electron-nuclear double resonance (\(^{19}\)F ENDOR) has recently emerged as a valuable tool in structural biology for distance determination between F atoms and a paramagnetic center, either intrinsic or conjugated to a biomolecule via spin labeling, yielding distances beyond those accessible by double electron-electron resonance (DEER). To further extend the accessible distance range we exploit the high-spin properties of Gd(III) and focus on transitions other than the central transition (\(|-1/2>|\leftrightarrow|+1/2>|\)), that become more populated at high magnetic fields and low temperatures. This increases the spectral resolution up to ca. 7 times, thus raising the long-distance limit of ENDOR almost twofold. We first demonstrate quantitative agreement between the experimental spectra and theoretical predictions for a model fluorine containing Gd(III) complex, whose \(^{19}\)F spectrum is well resolved in conventional central transition measurements. We then validate our approach on two proteins labeled with \(^{19}\)F and Gd(III), in which the Gd-F distance is too long to produce a well resolved \(^{19}\)F ENDOR doublet when measured at the central transition. By focusing on the \(|-5/2>|\leftrightarrow|-3/2>|\) and \(|-7/2>|\leftrightarrow|-5/2>|\) EPR transitions, a resolution enhancement of 4.5 and 7 fold was obtained, respectively. We also present data analysis strategies to handle contributions of different electron spin manifolds to the ENDOR spectrum. Our new extended \(^{19}\)F ENDOR approach should be applicable to Gd-F distances as large as 20Å, widening the traditional ENDOR distance window.

Introduction
Pulse electron paramagnetic resonance (EPR) methods, particularly those relying on electron-electron dipolar interactions (pulse dipolar EPR, PD-EPR) have emerged as effective tools for providing structural information on proteins and nucleic acids.\(^1\)\(^-\)\(^2\) PD-EPR experiments, usually carried out on frozen solutions, yield distance distributions between two paramagnetic centers.\(^3\) Since most biomolecules are diamagnetic, paramagnetic spin labels have to be introduced.\(^4\) The positions for installing labels are primarily selected based on the biological/structural question under consideration, and the choice of the spin label is dictated by the needs for chemical stability and/or compatibility with the biomolecule in its environment. In the last two decades, the number and variety of spin labels has expanded significantly and, at present, comprise standard nitroxide spin
labels, trityl radicals,\textsuperscript{5,6} Gd(III),\textsuperscript{7} Cu(II)\textsuperscript{8} and Mn(II)\textsuperscript{9} complexes. For routine experimental setups, the distance range accessible by PD-EPR methodology is 2-6 nm. The long range limit can be extended by applying rather complicated pulse sequences, like the 7-pulse DEER sequence,\textsuperscript{10} or by deuterating the protein.\textsuperscript{11} The low limit can be extended down to 1.5 nm by applying sufficiently short microwave pulses, preferably with single-resonance techniques\textsuperscript{12,13} or reverting to continuous wave (CW) EPR, where broadening induced by a pair of labels is isolated via comparison with the width of two singly labeled proteins.\textsuperscript{14}

Alternatively, rather than measuring electron-electron dipolar interactions, electron-nuclear dipolar interactions can be targeted by applying electron-nuclear double resonance (ENDOR) approaches. In ENDOR, hyperfine interactions between the electron spin in a paramagnetic center and surrounding nuclei are measured. Traditionally, it has been applied to extract the local spatial and electronic structure of intrinsic paramagnetic metal ions or metal clusters in proteins,\textsuperscript{15-20} while, at present, ENDOR experiments for distance determination on spin labeled biomolecules are being developed.\textsuperscript{21-27} For measuring weak hyperfine interactions, which are dipolar in nature, Mims ENDOR\textsuperscript{28} is the technique of choice, as demonstrated for a nitroxide label situated 1 nm away from a $^{31}$P nucleus in a membrane bilayer.\textsuperscript{29}

Bennati and co-workers demonstrated on synthetic models and RNA molecules that distances up to 1.5 nm (15 Å) can be determined by combining nitroxide and $^{19}$F labeling.\textsuperscript{21,30} The use of $^{19}$F provides high sensitivity, approaching that of $^1$H, owing to its high gyromagnetic ratio, as well as excellent selectivity, since $^{19}$F is absent in natural proteins and nucleic acids, in contrast to the abundant $^1$H. This feature was previously harnessed to investigate the binding of fluorinated substrate analogues to intrinsic paramagnetic centers in metalloenzymes by ENDOR.\textsuperscript{31,32} In addition, $^{19}$F labeling is notably more benign than attaching spin labels and can be performed at any pre-determined location in a protein, even in cores. This is in contrast to a large spin labels that are attached to surface residues, are flexible and therefore significantly enlarge the sampled distance distribution, lowering resolution, especially for short distances. Carrying out $^{19}$F ENDOR measurements at W-band (95 GHz) has the advantage that the separation between $^{19}$F and $^1$H signals is sufficiently large to avoid the overlap encountered at Q-band (34 GHz).\textsuperscript{33} ENDOR measurements of nitroxides at W-band are complicated by the resolved g-anisotropy which
leads to orientation selection, *i. e.* preferential excitation of spin labels with certain orientations at different magnetic field positions within the EPR spectrum. Therefore, distance determination requires acquiring a series of ENDOR spectra at different fields, which is time consuming, particularly for distances larger than 10 Å. Nevertheless, orientation selectivity permits an accurate determination of the parallel component of the dipolar interaction and thereby provides access to longer distances. In the case of rigid spin labels, additional structural information can be obtained by determining the orientation of the spin label in the structure of the biomolecule. Trityl\textsuperscript{22, 25} and Gd(III)\textsuperscript{23, 27} labels have been used for distance determination by \(^{19}\text{F}\) ENDOR, with the advantage of not requiring a set of orientation selection measurements owing to the isotropic nature of their EPR spectrum. Recently, phenoxyl\textsuperscript{24} and Cu(II)\textsuperscript{26} have also been used for distance measurements. Going beyond *in vitro* measurements, in-cell Gd(III)–\(^{19}\text{F}\) ENDOR significantly expanded the scope of this technique.\textsuperscript{27}

When applied to the same systems, \(^{19}\text{F}\) ENDOR data is complimentary to results obtained by \(^{19}\text{F}\) paramagnetic relaxation enhancement (PRE) and pseudo contact shift (PCS) solution nuclear magnetic resonance (NMR) techniques.\textsuperscript{34-36} The latter cover similar distance ranges, are carried out near room temperature and provide average electron-nuclear distances, while ENDOR measurements are performed in the frozen state and report on the conformational distribution. Furthermore, at least for solution NMR, the rotational correlation time of the molecule presents a limitation, excluding studies on very large systems, in contrast to ENDOR and magic angle solid state NMR measurements.

At present, the longest distance determined from an \(^{19}\text{F}\) ENDOR spectrum featuring a resolved doublet with a splitting of 20 kHz was 15 Å in an RNA molecule with a semi-rigid nitroxide spin label.\textsuperscript{21} In general, distances extracted from ENDOR measurements are limited by the intrinsic widths of the \(^{19}\text{F}\) ENDOR lines, which generally range from 20 to 35 kHz. The linwidths are determined by either the transverse relaxation of the \(^{19}\text{F}\) nucleus or by the width of electron-nuclear distance distribution. Smaller hyperfine couplings (longer distances) with unresolved doublets can, in principle, be estimated by measuring absolute Mims ENDOR efficiency,\textsuperscript{23} however, the reliability of such intensity measurements is yet to be demonstrated.
Here, we put forward a different approach for extending the long-distance range of ENDOR measurements by increasing the frequency resolution of the $^{19}$F doublet. We exploit high-spin Gd(III) labels and accessing EPR transitions other than the central transition (CT, $|{-1/2}\rangle \leftrightarrow |{+1/2}\rangle$), facilitated by high magnetic field and low temperature. We demonstrate the viability of this approach on a specially designed and synthesized model compound 1 that serves as a molecular “ruler”, shown in Fig. 1A, along with the two model proteins labeled with $^{19}$F and Gd(III): ubiquitin (Fig. 1B), where the standard $^{19}$F ENDOR doublet is barely resolved and the B1 domain of immunoglobulin-binding protein G (GB1) (Fig. 1C) with an unresolved $^{19}$F signal. We show that by focusing on the Gd(III) $|{-7/2}\rangle \leftrightarrow |{-5/2}\rangle$ transition, the spectral resolution can be increased by a factor of ~7, resulting in extending the distance range by a factor of 1.9, potentially reaching 20–25 Å.

![Fig. 1.](https://doi.org/10.26434/chemrxiv-2023-hg5zc)

**Fig. 1.** (A) Chemical structure of a molecular ruler for ENDOR Gd–F distance measurements (B, C) Backbone structures in ribbon representation of ubiquitin T66C (pdb id 1UBQ) (B) and GB1 Q32C (pdb id 1GB1) (C). The cysteine and fluorine containing side chains are shown in stick representation with the sulfur atoms in yellow and the fluorine atoms in green.

**Experimental details**

The synthesis of the Gd–F ruler 1 is described in detail in the Supporting Information (SI, Section S1). In brief, the corresponding PyMTA (pyridine-2,6-diyl)bis(methyleneitrilo)-tetrakis acetate) methylcarboxylate precursor was synthesized
according to a literature procedure,\textsuperscript{40} hydrolyzed to the carboxylic acid and amidated with 4-fluoropiperidine. Upon removal of protecting groups it was complexed with Gd(III), yielding complex 1. The Gd–F distance for this compound is 9.5 Å, based on quantum chemical calculations (SI, Section S2).

Proteins were prepared and spin-labeled as described previously.\textsuperscript{27} Ubiquitin T66C possesses 4-trifluoromethyl phenylalanine (tFmPhe) at position 45 and GB1 Q32C contains 5-fluorotryptophan (5F-Trp) at position 43. The BrPSPyDO3A-Gd(III) tag was attached to the single cysteines on both proteins.\textsuperscript{41} Chemical structures of the Gd(III) tag and F-labeled amino acids are shown in Fig. S1 (SI).

Complex 1 was dissolved in 50:50 v/v D$_2$O/glycerol-d$_8$ solution at a final concentration of 380 μM. Proteins were dissolved in 25 mM D$_2$O-based phosphate buffer (pD 7.0), 150 mM NaCl, and 20 vol. % glycerol-d$_8$ was added as a cryoprotectant, yielding a final protein concentration of 40 μM. For EPR measurements, solutions were placed in fused silica capillaries (inner diameter 0.6 mm) and sealed at one end.

Pulsed EPR and ENDOR measurements were performed using two pulsed W-band EPR spectrometers, a home-built spectrometer of previously described design,\textsuperscript{42} permitting EPR measurements at 1.7–300 K and ENDOR measurements at 6–300 K, and a Bruker Elexsys E680 spectrometer equipped with a home-built W-band MW extension and a cryogen free Cryogenic 6 T magnet with a variable temperature insert for ENDOR measurements at 2.2–300 K. Detailed descriptions of ENDOR experimental conditions are provided in the SI (Section S4). As shown earlier,\textsuperscript{43} the use of an adiabatic chirp pulse prior to the Mims sequence allows for the transfer of spin polarization to the observed electron spin transition. For the systems studied here, at low temperatures, this results in a signal enhancement of ca. 30% (Fig. S2) at 6 K. Unfortunately, this complicates the quantitative analysis of the contribution from each EPR transition to the ENDOR spectra. Therefore, all spectra in the present work were obtained in the absence of adiabatic chirp pulses.

Details of the numerical simulations of ED-EPR and Mims ENDOR spectra are presented in the SI (Section S6), and the best fit parameters obtained from the simulations are listed in Tables S3 and S4.
Results and discussion

Theoretical background

We first consider the Mims ENDOR powder line shape for a high-spin paramagnetic $S=7/2$ center, coupled to a nuclear spin with $I=1/2$. For Gd(III) with weakly coupled nuclei in a high magnetic field, the Larmor frequency of the electron spin, $\nu_S$, is much larger than the zero field splitting (ZFS) components, and the Larmor frequency of the nuclear spin, $\nu_I$, is much larger than the hyperfine splitting. Accordingly, the projections $m_S$ and $m_I$ of the electron and nuclear spins on the external magnetic field axis are good quantum numbers, and the ENDOR resonance frequencies for allowed NMR transitions ($|\Delta m_I|=1$) are given by\(^{37}\)

$$\nu(\beta, m_S) = \nu_I - m_S \cdot a(\beta)$$

where $\beta$ is the angle between the magnetic field and the vector connecting the Gd(III) ion and $^{19}\text{F}$ nucleus. Here we consider only long Gd–F distances in non-conjugated systems, hence the hyperfine splitting can be assumed to be purely dipolar, given by

$$a(\beta) = \left(3\cos^2 \beta - 1\right) \frac{\mu_0 g_e \mu_B g_n H_N}{2\hbar r^3} = \left(3\cos^2 \beta - 1\right) a_\perp,$$

where $\mu_0$ is vacuum magnetic permeability, $g_e$ and $g_n$ are electron and nuclear $g$-values, $\mu_B$ and $\mu_N$ are Bohr magneton and nuclear magneton, respectively, $\hbar$ is the Planck constant and $r$ is the Gd-F distance. The energy level diagram for $S=1/2$ coupled to a nucleus with $I=1/2$ focusing on the $m_S=-7/2, -5/2, -1/2$ and $+1/2$ levels is shown in Fig. 2A.

For each allowed EPR transition $|m_S\rangle \leftrightarrow |m_S+1\rangle$ and each orientation $\beta$, the ENDOR spectrum consists of a doublet separated by the dipolar splitting constant $a(\beta)$. In a homogenous frozen solution, all values of $\beta$ are equally probable and the resulting powder patterns for the EPR transitions $|{-1/2}\rangle \leftrightarrow |{+1/2}\rangle$ and $|{-7/2}\rangle \leftrightarrow |{-5/2}\rangle$ are illustrated in Fig. 2B. In the simulated spectra, the singularities corresponding to parallel ($a_\parallel, \beta = 0$) and perpendicular ($a_\perp, \beta = \pi/2$) orientations of electron-nuclear dipolar...
vector are highlighted in Fig. 2B. Note that the singularities corresponding to $a_\perp$ and $a_\parallel$ that belong to the same electron spin manifold $m_S$, appear at different sides of Larmor frequency, $\nu_I$.

**Fig. 2.** (A) Schematic illustration of energy levels corresponding to different projections $m_S$ and $m_I$ of electron and nuclear spins for Gd(III) coupled to $^1$H or $^{19}$F nucleus. Allowed EPR and NMR (ENDOR) transitions are shown with green and violet arrow, respectively. Nuclear sublevels are shown for clarity only for $m_S=\pm7/2, \pm5/2, \pm1/2$ and $+1/2$. Note, the splittings are not shown to scale. The populations of the energy levels are represented by different line thickness. (B) Simulated powder ENDOR spectra for $|\pm1/2\rangle \leftrightarrow |\mp1/2\rangle$ and $|\pm7/2\rangle \leftrightarrow |\mp5/2\rangle$ transitions; vertical blue and red lines mark parallel and perpendicular contributions for $|\pm1/2\rangle \leftrightarrow |\mp1/2\rangle$ and $|\pm7/2\rangle \leftrightarrow |\mp5/2\rangle$, respectively. (C) Simulated Mims ENDOR patterns for the $|\pm1/2\rangle \leftrightarrow |\mp1/2\rangle$ (blue) and $|\pm7/2\rangle \leftrightarrow |\mp5/2\rangle$ (red) transitions. For all spectra the following parameters were used: $a_\perp = 100$ kHz, $\tau = 2$ $\mu$s and a Lorentzian line width of 50 kHz.
In order to correctly describe the Mims ENDOR spectrum, one has to account for blind spots and scale each of the ENDOR doublets by the coefficient:

\[ C_{\text{Mims}}(a) = 1 - \cos(2\pi \cdot a\tau) = 2 \cdot \sin^2 \left( \frac{2\pi \cdot a\tau}{2} \right) \],

where \( \tau \) is the time delay between the first and the second \( \pi/2 \) pulses in the Mims sequence.

The overall Mims ENDOR line shape, \( F_{\text{ENDOR}} \), is obtained by summation over all orientations and EPR transitions:

\[ F_{\text{ENDOR}}(v_{RF}; B_0) \propto \sum_{m_S} \int_{\beta=0}^{\pi/2} \rho(\beta; B_0) \sin \beta d\beta \cdot \sin^2 \left( \frac{\pi \cdot a(\beta) \tau}{2} \right) \cdot w_{\text{EPR}}(|m_S\rangle \leftrightarrow |m_S+1\rangle; B_0) \times \\
\times \left[ F[v_I(B_0) - m_S \cdot a(\beta)] + F[v_I(B_0) - (m_S+1) \cdot a(\beta)] \right] \]

(4)

where \( w_{\text{EPR}} \) is the excitation probability of a given EPR transition and \( F \) is the line shape of the individual ENDOR line, which, in the general case, can be described by convolution of Gaussian and Lorentzian line shapes. \( \rho(\beta; B_0) \) is an orientation selection function that represents the number density of Gd–F pairs with orientation \( \beta \), excited at a particular magnetic field position, \( B_0 \), which is uniform (\( \rho \equiv 1 \), no orientation selection) in the simplest case.

In Fig. 2C we present Mims ENDOR line shapes simulated according to eq. (4) for the CT, \( |-1/2\rangle \leftrightarrow |+1/2\rangle \) and the \( |-7/2\rangle \leftrightarrow |-5/2\rangle \) transition of Gd(III). For spectra recorded at CT the most pronounced splitting is \( a_\perp \), located between the perpendicular singularities of the \( m_S = \pm 1/2 \) manifolds. For other transitions, a different behavior is noted: the dominant splitting is observed between parallel and perpendicular features of the spectrum, separated by a strong blind spot in the region of Larmor frequency. The splitting is of the order of \( 3|m_S'|a_\perp \), with \( |m_S'| = \min \{|m_S|, |m_S+1|\} \). Thus, for Gd(III), \( S = 7/2 \), at low temperature where the lowest \( m_S \) is predominantly populated, the expected splitting can be as high as \( 7.5a_\perp \) and the longest measurable distance is increased by a factor close to two, compared to conventional spin 1/2 labels.
Using the synthetic Gd–F “ruler” for ENDOR

In order to assess the validity of the proposed approach, complex 1 (Fig. 1A) with a fixed molecular geometry was synthesized, serving as a Gd–F “ruler”. The echo detected EPR (ED-EPR) spectrum of 1 consists of a sharp peak corresponding to the Gd(III) CT transition, superimposed on a broad envelope that corresponds to all other transitions (Fig. S3A). The contributions of the individual transitions to the ED-EPR Gd(III) spectrum can be deconvoluted by simulations shown in Figs. 3A, S3A and parameters listed in Table S3. To ensure that contributions of the individual transitions to the ENDOR spectra can be correctly determined, the ED-EPR was acquired with the Mims ENDOR pulse sequence, the same time delays and the RF frequency located far from away from the $^1$H and $^{19}$F resonances (Fig. S4). This was necessary because different transitions can have different phase memory times and consequently ED-EPR spectra acquired with different pulse sequences may have somewhat different lineshapes.46 Echo decay and spin lattice relaxation measurements carried out on the central transition are presented in Fig. S3.

$^{19}$F Mims ENDOR spectra of 1 recorded at different magnetic field positions are presented in Fig. 3B. The top spectrum, recorded at the CT, shows a doublet with a resolved splitting of $a_\perp$. This spectrum is well reproduced by simulations, assuming that it arises only from the CT, yielding $a_\perp = 73.7\pm 0.9$ kHz and a Lorentzian line width $\Delta_L = 19.0\pm 0.9$ kHz.
Fig. 3. (A) Experimental (black trace) and simulated (red trace) ED-EPR spectra of 1 at 6 K with contributions of individual transitions in different colors. Colored dots correspond to the positions where ENDOR spectra were measured and indicate the transition intensities obtained from simulations of ENDOR spectra shown in panel B. (B) Mims ENDOR spectra (6 K, τ=2 μs) recorded at different field positions relative to the maximum of the CT (black traces). Simulated spectra using relative contributions of the various transitions determined from simulation of the ED-EPR spectrum (blue traces) superimposed on simulations with those determined from the ENDOR spectra (red traces). (C) Contributions of individual transitions to the ENDOR spectrum recorded at +50 mT. The color coding for individual transitions is identical in panels A and C and each of the spectral singularities are indicated by dotted lines in C.

The other spectra, recorded away from the CT (Fig. 3B) exhibit complex shapes with several peaks. We simulated these using eq. (4), with the $a_{\perp}$ and line width parameters from the simulation of the CT spectrum, and the contributions of the individual EPR transitions from the simulation of the ED-EPR spectrum (see colored lines in Fig. 3A). The contributions of the individual transitions to the ENDOR spectra are depicted in Fig. 3C for the ENDOR spectrum recorded at +50 mT. At this juncture, we would like to emphasize that the agreement between simulation and experiment is remarkable, given that
the simulations of ENDOR spectra recorded off the CT were performed without any fitting parameters and that all parameters were taken from independent experiments. Further improvements can be obtained by a global fitting of $a_{\perp}$, the line width and the relative contributions of each EPR transition, and these improved fits are shown in Fig. 3B with best fit parameters listed in Table S4, yielding $a_{\perp} = 76.6 \pm 1.1$ kHz which corresponds to $r = 9.9 \pm 0.05$ Å. This Gd–F distance is in excellent agreement with the distance obtained by DFT optimization, $r = 9.98$ Å for the axial and $r = 10.15$ Å for the equatorial conformer of 1 (see Section S2, SI).

The best -fit contributions of the various EPR transitions at each field are shown in Fig. 3A as dots. Only minor, albeit systematic, differences are observed between the contributions from the simulations of the ED-EPR spectrum and those from global fitting of the ENDOR spectra. In general, the values obtained from the ED-EPR simulation slightly underestimate the contribution of the $|-3/2\rangle \leftrightarrow |-1/2\rangle$ transition and overestimate the contribution of the $|-7/2\rangle \leftrightarrow |-5/2\rangle$ transition. The observed discrepancies most likely originate from the uncertainties inherent to ED-EPR spectra simulations, in particular pulse non-ideality and the uncertainty in the exact form of ZFS parameter distribution. In this context, we would like to point out that the transition weights obtained from the ENDOR spectra are independent of these shortcomings and can in principle be used to experimentally validate and refine the ZFS distribution models.

To further corroborate the above data analysis approach, we also applied it to the $^1$H ENDOR spectra of 1. $^1$H ENDOR spectra are complex since multiple types of hydrogens contribute. Therefore, we initially focused on the analysis of the ENDOR spectrum recorded at the CT for different $\tau$ values (Fig. S5A, B). Simulations of this series yielded a total of six types of hydrogens. Their tentative assignment based on the DFT optimized structure of 1, relative numbers of hydrogen atoms for each type and Gd–H distances are provided in Fig. S5C. We also recorded spectra at several field positions away from the CT (Fig. S6) and simulated these with the fixed $a_{\perp}$, line widths and relative numbers of hydrogens extracted from the CT measurements, using contributions of the various EPR transitions at each field position obtained from ED-EPR simulations. The
agreement between the experimental and simulated spectra is remarkably good (Fig. S6) and details of the simulation are provided in Section S8 of the SI.

In summary, the experiments carried out on the synthetic Gd–F “ruler” 1 demonstrated the validity of our approach both in terms of data collection and the theoretical model for analyzing the experimental data. We established that the most advantageous conditions for such measurements are far away from the CT and at very low temperatures, i.e. when the EPR spectrum is dominated by the $|{-7/2}\rangle \leftrightarrow |{-5/2}\rangle$ transition. Under these conditions, the spectral resolution is large and interpretation of the results is straightforward.

Given the excellent results for the synthetic ruler, we proceeded to apply the above detailed methodology to two protein samples with low resolution $^{19}$F ENDOR spectra at the CT arising from longer Gd–F distances and possibly broader Gd–F distance distributions.27

Distance measurements on proteins

Two model proteins, T66C ubiquitin and Q32C GB1 that possess tFmPhe and 5F-Trp as fluorinated amino acids, respectively, were tagged with Gd(III)-BrPy-DO3A spin label, referred to as Ub-T66C-DO3A and GB1-Q32C-DO3A, respectively (See Fig. 1). The chemical structure of the spin label (Fig. S1A) is characterized by a relatively short tether that restricts the conformational mobility of the tag and hence limits the distance distribution width.41 $^{19}$F ENDOR spectra for the same spin-labeled proteins measured at CT were reported by us previously, and the $^{19}$F doublet was barely resolved for Ub-T66C-DO3A and not resolved at all for GB1-Q32C-DO3A.27 The detailed comparison between the Gd–F distances obtained from $^{19}$F ENDOR measured on CT, PRE NMR and in silico modeling of the labeled protein structures is reported in our previous work.27

Ub-T66C-DO3A

The ED-EPR spectrum and spin relaxation characteristics of Ub-T66C-DO3A are shown in Fig. S7. As pointed out above, to enhance resolution, measurements are ideally carried out at the lowest possible temperature. This is clearly illustrated by the temperature dependence of the ED-EPR spectrum of Ub-T66C-DO3A recorded between 1.7 and 16K
The associated redistribution of the $m_s$ level populations (Fig. S8B) shows that at 2 K the $|{-7/2}\rangle \leftrightarrow |{-5/2}\rangle$ transition dominates the spectrum, also confirmed by ED-EPR spectrum simulation (Fig. S8C,D). Figure 4A shows the Mims $^{19}$F ENDOR spectra of Ub-T66C-DO3A recorded at 2.2 K at different field positions with respect to the CT. These can be compared to the spectrum recorded at the CT at 6 K (Fig. 4B, upper trace), which displays a poorly resolved doublet with a splitting of ca. 40 kHz. The off-CT spectra exhibit remarkably larger splittings of ca. 190 kHz, with a shape similar to the theoretical predictions illustrated in Fig. 2C and the spectral maxima are identified as $5a_\perp/2$ on the left and $5a_{\parallel}/2$ on the right. These are highlighted in Fig. 4 by dashed vertical lines for $a_\perp = 30.6$ kHz. The low-frequency maximum of the spectrum is well aligned with the perpendicular singularity, whereas the high-frequency one is somewhat shifted. This behavior is also apparent in the theoretical spectra (Fig. 2C) and is most likely due to the blind spot pattern of Mims ENDOR. According to the simulated spectrum in Fig. 2C, the observed splitting in the $|{-7/2}\rangle \leftrightarrow |{-5/2}\rangle$ spectrum corresponds to $7a_\perp$ instead of $7.5a_\perp$ as anticipated from the parallel and perpendicular singularity positions.

A striking feature of the off-CT spectra is the different relative intensities of the signals at the left and right sides of the Larmor frequency for different field positions, suggesting orientation selection,

non-uniform excitation of the differently oriented ZFS over the magnetic field range. This behavior is surprising, considering the broad distribution of ZFS parameters characteristic of Gd(III) complexes. Indeed, the off-CT ENDOR spectra of the complex 1, which has a different Gd(III) chelate, did not show orientation selection.

Ub-T66C-DO3A possesses three $^{19}$F nuclei in the tri-fluoro methyl group and in the spectral analysis we consider them as identical, which is justified if the determined distance is substantially larger than the distance between the F atoms (ca. 2 Å). The simplest way to analyze the spectra is to abolish the orientation selection by adding all ENDOR spectra with weights equal to the EPR spectrum intensities at the corresponding positions (bottom trace Fig. 4B).

The observed splitting of 190 kHz corresponds to $a_\perp \approx 190/7 = 27$ kHz, yielding Gd–F distance of $r \approx 14$ Å. If a sufficient number of spectra are summed, the spectral line can be analyzed using eq. (4) without any orientation.
selection ($\rho \equiv 1$). The numerical simulation using eq. (4) (Fig. 4B) yielded $a_\perp = 30.1$ kHz and a distance $r = 13.5$ Å. Using a Gaussian distance distribution with $r_0 = 14.9\,\text{Å}$ as the center of the distance distribution and $\Delta r = 4.1\,\text{Å}$ as its width improves the agreement between the simulation and the experimental spectra (Fig. 4B). This distance distribution is in excellent agreement with the one derived from the central transition ENDOR, namely $r_0 = 14.8\,\text{Å}$ and $\Delta r = 6.3\,\text{Å}$.

**Fig. 4.** (A) Mims ENDOR spectra of Ub-T66C-DO3A ($\tau=4 \,\mu\text{s}$, 2.2 K) recorded at different field positions relative to the maximum of the CT (black trace). Simulated spectra with the orientation selection taken from simulations of the ED-EPR spectrum (red traces), and represented by the phenomenological function $\rho(\beta; B_0)$ in eq. (4) (blue traces). Note that the experimental spectra recorded at positions $-100$ mT and $+50$ mT CT exhibit non-uniform background, originating from either remote $^1$H transitions or instrumental artifact. This background was subtracted in the spectra shown (see Fig. S9, SI). (B) Top traces: Mims ENDOR spectrum of Ub-T66C-DO3A ($\tau=2 \,\mu\text{s}$, 6 K) recorded at the CT (black trace) and its simulation using a single Gd–F distance (red trace). Bottom traces: summation of all off-CT spectra, weighted according to the ED-EPR spectrum intensity (black trace) along with simulations according to eq. (4), assuming a single Gd–F distance (red trace) or a Gaussian distribution of Gd–F distances (blue trace). Vertical dashed lines correspond to parallel and perpendicular singularities of the powder spectra corresponding to $m_s=-5/2$ electron spin manifold and $a_\perp = 30.6$ kHz.
Table 1. Comparison of the $a_\perp$, kHz values and the associated distances using different analysis approaches for the ENDOR spectra of Ub-T66C-DO3A shown in Fig. 4. The Lorentzian $\Delta_L$ and Gaussian $\Delta_G$ line widths used in the simulations are listed as well.

<table>
<thead>
<tr>
<th>Method</th>
<th>$a_\perp$, kHz</th>
<th>$r$, Å</th>
<th>$\Delta_L$, kHz</th>
<th>$\Delta_G$, kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Splitting determination</td>
<td>27</td>
<td>14</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Fit of summed spectra; single Gd–F distance$^a$</td>
<td>30.1±1.4</td>
<td>13.5±0.2</td>
<td>42±8</td>
<td>0</td>
</tr>
<tr>
<td>Fit of summed spectra; Gaussian distribution of distances$^b$</td>
<td>–</td>
<td>$r_0 = 14.9 \pm 0.5$</td>
<td>$\Delta r = 4.2 \pm 0.6$</td>
<td>13±6</td>
</tr>
<tr>
<td>Fit of individual spectra orientation selection predicted from ED-EPR simulation$^c$</td>
<td>30.4±1.6</td>
<td>13.5±0.2</td>
<td>36±5</td>
<td>0</td>
</tr>
<tr>
<td>Fit of individual spectra; orientation selection modeled phenomenologically$^d$</td>
<td>31.4±2.0</td>
<td>13.3±0.3</td>
<td>38±3</td>
<td>0</td>
</tr>
</tbody>
</table>

$^a$ Fig. 4B, red line; $^b$ Fig. 4B, blue line; $^c$ Fig. 4A, red line; $^d$ Fig. 4A, blue line.

Despite the difference in the average distances obtained by both approaches (i.e., distribution vs single distance) similar spectral shapes are obtained, since shorter distances contribute more to the overall ENDOR spectrum due to the $\sim 1/r^6$ electron-nuclear distance dependence of the ENDOR efficiency.\(^{27}\) As a result, wide distance distributions introduce larger uncertainties in the measured distances (in the present case amounting to ca. 1.5Å with respect to the mean distance). These uncertainties can be reduced by choosing Gd(III) tags with short tethers which limits the conformational space sampled by the tag.

It is also possible to analyze the set of spectra in Fig. 4 individually, which is usually done when the anisotropy is determined by $g$-anisotropy.\(^{21, 26}\) In this case, information regarding the orientation of the electron-nuclear dipolar vector in the framework of the $g$-tensor is obtained and, in turn, provides additional important geometrical information. This approach, however, comes with several caveats for cases when the anisotropy is dominated by the Gd(III) ZFS. First, the predicted orientation selection behavior is sensitive to the details of the ZFS distribution, which can be obtained only to a limited extent from the ED-EPR spectra. Second, the orientation of the ZFS tensor within the Gd(III) moiety is likely to be broad, resulting in a wide orientational distribution of the Gd–F vectors in the
framework of ZFS tensor of Gd(III). Adding distributions to the orientations of the dipolar direction relative to the ZFS is impractical as it should be correlated with the distribution of the ZFS parameters. Nevertheless, we attempted using this approach, assuming a well-defined orientation of the Gd–F vector, with polar angles $(\theta_F, \varphi_F)$ in the reference frame of ZFS as fitting parameters, (red traces in Fig. 4A). This resulted in $a_\perp = 30.4$ kHz, $(\theta_F = 86^\circ, \varphi_F = 90^\circ)$, however with an unsatisfactory agreement for $-50$ and $-100$ mT (see Section S10, SI for details). Therefore, we preferred to account for the orientation selection by considering directly the orientation of the Gd–F vector relative to $B_0$. For this we defined a function $\rho(\beta; B_0)$ in eq. (4), which takes into account the relative number of molecules with an orientation $\beta$ between $B_0$ and the Gd–F vector. The orientation selection function $\rho(\beta)$ for each spectrum is fitted individually and details of this fitting procedure is provided in the SI (Section S10). Although this approach comprises an increased number of fitting parameters for each spectrum, these parameters are independent and information on the Gd–F distances can be extracted, as long as the hyperfine splitting and line widths are assumed to be the same for all simulated spectra. The result is shown as blue traces in Fig. 4A, revealing excellent agreement between the experiment and the simulation for $a_\perp = 31.4$ kHz. The resulting orientation selection functions $\rho(\beta)$ for different field positions are depicted in Fig. S12, revealing that for spectra recorded left of the CT there is preferential excitation for Gd–F vectors perpendicular to the static magnetic field $(\beta = 90^\circ)$, while to the right of the CT parallel orientations $(\beta = 0)$ are preferentially excited.

Table 1 lists all $a_\perp$ values obtained from the different analysis approaches, with an average of $a_\perp \approx 30.6 \pm 2.4$ kHz, corresponding to an average Gd–F distance of $r = 13.4 \pm 0.4$ Å. The small error in the Gd–F distance suggests that this approach can be exploited to extract distances with high precision.

For comparison, we also recorded $^{19}$F ENDOR spectra at 6 K, demonstrating that spectra recorded at positions $\pm 50$ mT away from the CT are similar to those recorded at 2.2 K (Fig. S13). Therefore, we posit that also at 6 K the contributions of low-lying electron
spin manifolds to the ENDOR spectrum are significant and that Gd–F distances can be estimated from such higher temperature spectra. This makes our approach more broadly applicable since low temperatures around 2K are not readily accessible in many spectrometers.

GB1-Q32C-DO3A

The $^{19}$F ENDOR spectra of GB1-Q32C-DO3A recorded at 6 K are shown in Fig. 5A, and exhibit a well resolved splitting of ca. 85 kHz for off-CT, in contrast to the unresolved CT spectrum. The ED-EPR spectrum and relaxation measurements on this protein are presented in Fig. S14.

Because of the higher temperature the contributions of the various transitions have to be taken into account in the analysis of the spectra. It is possible to estimate these contributions from the ED-EPR, although, as shown above for complex 1, the Gd–F “ruler”, this approach may introduce systematic errors. Fortuitously, the contributions of different spin manifolds can be readily estimated from an independent experiment, the $^1$H ENDOR spectrum of the same molecule at the same field positions which exhibit a much better SNR. Since the $^1$H and $^{19}$F ENDOR spectra are recorded with identical parameters (except for the $\tau$ values and the RF frequency range), a reliable estimate of the relative contributions for different transitions can be obtained. No orientation selection is expected for the $^1$H ENDOR because of the symmetrical arrangement of the protons around the Gd(III) in DO3A. Another advantage of using $^1$H ENDOR spectra to estimate the EPR transition probabilities is that the protons, being part of the Gd(III) chelate, are not sensitive to the conformation distribution of the spin label within the protein.

https://doi.org/10.26434/chemrxiv-2023-hg5zc ORCID: https://orcid.org/0000-0003-1044-4783 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0
Fig. 5. (A) Experimental (black traces) and simulated $^{19}$F ENDOR spectra of GB1-Q32C-DO3A ($\tau = 4 \mu s$, 6 K) (red traces) at different field positions with respect to the maximum of the CT. (B) Simulated $^{19}$F ENDOR spectra illustrating the shape dependence of the spectra for the $\left| -7/2 \right \rangle \leftrightarrow \left| -5/2 \right \rangle$ (upper panel) and $\left| -1/2 \right \rangle \leftrightarrow \left| +1/2 \right \rangle$ (lower panel) EPR transitions on the Gd–F distance for purely Lorentzian lines with $\Delta L = 30$ kHz. The intensity of spectrum at the shortest distance is scaled by 0.5.

In the simulation, two types of hydrogens as well as the matrix hydrogens were considered, consistent with previous ENDOR measurements for Gd complexes$^{50}$ (Fig. S15, S1). Hyperfine splittings of individual hydrogens and their relative contributions were determined independently from the spectrum recorded at the CT. The $^1$H ENDOR spectra recorded off CT were simulated to extract the relative contributions of the EPR transitions at different fields (Fig. S16) and these values were used in the simulation of the $^{19}$F ENDOR spectra shown in Fig. 5A. Parameters for the simulations are provided in Table S6 along with those obtained from ED-EPR simulations. As can be appreciated, the ENDOR spectra recorded farthest away from the CT feature orientation selection, similar to those of Ub-T66C-DO3A (compare Figs. 5A and S13). Therefore, we assumed that the orientation...
selection function for GB1 is equal to that previously determined for Ub-T66C-DO3A, and is the same for all EPR transitions. Despite such oversimplification, satisfactory simulations of the $^{19}$F ENDOR spectra were obtained with the following best fit parameter: $a_\perp = 21.1$ kHz that corresponds to Gd–F distance of 15.2 Å. Note that in these $^{19}$F ENDOR spectra, the observed splitting originates mainly from the $|{-5/2}\rangle \leftrightarrow |{-3/2}\rangle$ electron spin manifold (as illustrated in Fig. S17), and thus should be of the order of $(a_\perp + a_\parallel)3/2 = 4.5a_\perp$. Therefore, the measured 85 kHz splitting yields $a_\perp \approx 19$ kHz, corresponding to Gd–F distance of $r = 15.7$ Å, in excellent agreement with the value obtained from the simulation.

**Estimation of distance limits of the proposed approach**

To evaluate the limits of the above approach, we simulated $^{19}$F ENDOR spectra for a series of Gd–F distances in the range 10–25 Å with line width of 30 kHz, (Fig. 5B) characteristic of the experimental spectra reported here. The upper panel shows spectra recorded at $|{-7/2}\rangle \leftrightarrow |{-5/2}\rangle$, and the lower panel CT spectra (for a $S=\frac{1}{2}$ system). The CT spectra are resolved for distances up to ca. 12.5 Å, and for the $|{-7/2}\rangle \leftrightarrow |{-5/2}\rangle$ transition, resolved doublets can be detected up to ca. 20 Å. For a narrower ENDOR linewidth, ~20 kHz, the upper distance limit is expected to be even higher, as is observed for a semi-rigid nitroxide spin label in RNA.\(^\text{21}\) Indeed, distances as large as 25 Å maybe assessable (Fig. S18).

As is generally known, the Mims ENDOR efficiency scales according to $\sim 1/r^6$ reducing the sensitivity for long $r$ values (see Fig. 5B). Our approach also comes at the price of lower sensitivity, given the larger width of the EPR spectrum of $|ms|>1/2$ transitions. This is offset, to some degree, by the sensitivity gain afforded by the higher EPR excitation probability (4 times higher at the CT and 2.6 times for $|{-7/2}\rangle \leftrightarrow |{-5/2}\rangle$). Sensitivity enhancements can also be gained using rigid Gd(III) tags with small ZFS.\(^\text{51,52}\) In addition, improvements in RF efficiency and an increase in the repetition rate, currently limited by the RF amplifier duty cycle, can also be exploited for gaining sensitivity.
Finally, performing these measurements at Q-band, taking advantage of currently easily accessible pulsed EPR spectrometers are possible and $^{19}$F ENDOR distance measurements at Q-band have already been reported for trityl and Cu(II) labeled biomolecules. However, at lower frequencies, interference between $^{19}$F and $^1$H ENDOR lines may occur, especially for high spin electron transitions, since the corresponding $^1$H ENDOR lines can extend far out from the Larmor frequency according to eq. (1). If this is the case, subtraction of spectra obtained in the absence of $^{19}$F, is necessary, as recently shown for Cu(II)–$^{19}$F ENDOR. In addition, the smaller thermal spin polarization at Q-band may be limiting. To achieve the same thermal occupancy of the low-lying levels a 95/34 at Q-band ~ 3 times lower temperature has to be employed compared to W-band. Since we showed above for GB1-Q32C-DO3A that measurements at 6 K are possible, we envisage that collecting spectra at ~2 K at Q band or by enhancing the population of low-lying spin levels by polarization transfer are potentially feasible.

**Conclusions**

We have developed and presented an efficient approach for significantly extending the range of accessible ENDOR-derived Gd–F distances by taking advantage of the high spin of Gd(III) in combination with high field and low temperature measurements. Our approach includes measurement schemes as well as data analysis strategies, as illustrated for a model compound with a fixed Gd–F distance that serves as a molecular “ruler”, as well as for two model proteins containing fluorine atoms and Gd(III) tags. We demonstrate that a Gd–F distance of 15 Å can be extracted from resolved $^{19}$F ENDOR spectra recorded at the $|−7/2⟩ ↔ |−5/2⟩$ and $|−5/2⟩ ↔ |−3/2⟩$ transitions and that distances of up to 20–25 Å may be reachable.

Although a quantitative interpretation of such $^{19}$F ENDOR spectra in the context of Gd–F distance determination is complex and influenced by orientation selection and overlapping contributions from different electron spin manifolds of Gd(III), we demonstrate that effective strategies for overcoming the resulting shortcomings are possible. The contributions of different electron spin manifolds to the ENDOR spectrum can be determined either by EPR spectral simulations, or from the $^1$H ENDOR spectrum recorded on the same system. The presence of unexpected orientation selection for the
BrPy-DO3A-Gd(III) label in both ubiquitin and GB1 was efficiently treated phenomenologically and did not present a unsurmountable problem.

Acknowledgments

This work was funded by the National Science Foundation USA-Israel Science Foundation program through BSF 2021617 (to DG), and NSF-MCB 2116534 to (AMG), as well as NSF grant CHE 1708773 (to AMG) and made possible, in part, by support from the Helen and Martin Kimmel Institute for Magnetic Resonance Research, the Max Planck Institute for Chemical Energy Conversion and the historic generosity of the Harold Perlman Family (DG). The authors thank Mr. Ofer Aluf for his help in maintaining the spectrometer at WIS.

Associated content

Further details on synthesis, EPR and ENDOR measurements, spin relaxation properties of the samples, spectra simulations and DFT calculations.

References


(14) Altenbach, C.; Oh, K.-J.; Trabanino, R. J.; Hideg, K.; Hubbell, W. L. Estimation of Inter-Residue Distances in Spin Labeled Proteins at Physiological


(32) Tucci, F. J.; Jodts, R. J.; Hoffman, B. M.; Rosenzweig, A. C. Product Analogue Binding Identifies the Copper Active Site of Particulate Methane Monooxygenase. Nature Catalysis 2023. DOI: 10.1038/s41929-023-01051-x.


