Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information

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

Martyna Judd,[a] Elwy H. Abdelkader,[b] Mian Qi,[c] Jeffrey R. Harmer,[d] Thomas Huber,[a] Adelheid Godt,[c] Anton Savitsky,[e] Gottfried Otting,[b] Nicholas Cox[a,]*

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

[c] *Faculty of Chemistry and Center for Molecular Materials (CM_2), 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

Table of Contents

SI 1. Protein Synthesis and Labelling

C_{D9k} Synthesis Samples of calbindin D_{9k} (**C_{D9k}**) loaded with a Gd^{3+} ion and containing CF_3-Phe residues at different sites were prepared from a construct (NT*-**C_{D9k}**) that contains an N-terminal domain consisting of the His_6-NT* solubility tag derived from spider silk protein,[1] a tobacco etch virus (TEV) protease recognition site (ENLYFQ/A, where the protease cleaves between Q and A) and the amino acid sequence of **C_{D9k}** (Fig. S1). The DNA sequences encoding NT*-**C_{D9k}** with the codon of residue F50 or F66 changed to the amber stop codon were synthesized by Integrated DNA Technologies (USA) and cloned between the *Nde*I and *BamH*I restriction sites of the pET-3a plasmid (Novagen, USA).

MGHHHHHHMS HTTPWTNPGL AENFMNSFMQ GLSSMPGFTA
SQLDKMSTIA QSMVQSIQSL AAQGRTSPND LQALNMAFAS
SMAEIAASEE GGGSLSTKTS SIASAMSNAF LQTTGVVNQP
FINEITQLVS MFAQAGMNDV SAGNSENLYF QAKKSPEELK
GIFEKYAAKE GDPNQLSKEE LKLLLQTEFP SLLKGGSTLD
EL**F**EELDKDG DGEVSFEE**F**Q VLVKKISQ

Fig. S1. Amino acid sequence of the NT*-**C_{D9k}** construct used. The residues of the N-terminal domain of the spider silk protein NT* and **C_{D9k}** are shown in red and blue, respectively. The TEV protease recognition site is underlined and residues F50 and F66 in **C_{D9k}** are highlighted in bold.

C_{D9k} samples with CF_3-Phe were produced by co-transforming *E. coli* B-95 ΔA cells[2] with pET-**C_{D9k}**-1 or pET-**C_{D9k}**-2 and pEVOL-CFRS plasmids as previously described in ref.[3] The recovered cells were grown at 37 °C in LB medium containing 100 mg/L carbenicillin and 33 mg/L chloramphenicol. A 10 mL overnight culture was used to inoculate 1 L LB medium supplemented with 100 mg/L carbenicillin, 33 mg/L chloramphenicol, 0.2% arabinose and 2 mM CF_3-Phe. The cells were grown at 37 °C to an OD_{600} of 0.6–1, where the temperature was reduced to 25 °C and protein expression induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG).

After protein expression for 16 h, the cells were harvested by centrifugation at 4000 *g* for 15 minutes at 4 °C. The harvested cells were resuspended in buffer A (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 5% glycerol, 10 mM imidazole) followed by lysis using an Avestin Emulsiflex C5 (Avestin, Canada) with two passes at 10,000–15,000 psi. The cell lysate was clarified by centrifugation for 1 h at 30,000 *g* at 4 °C. The

clarified cell lysate was loaded onto a 5 mL HisTrap FF column connected to an ÄKTA pure 25 chromatography system (Cytiva, USA). The column was washed with 20 column volumes buffer B (same as buffer A but with 20 mM imidazole) and the protein was eluted with 5 column volumes buffer C (same as buffer A but with 500 mM imidazole). The eluted protein was desalted using a HiPrep Desalting 26/10 column (Cytiva, USA) equilibrated with buffer D (50 mM Tris-HCl pH 7.5, 300 mM sodium chloride, 1 mM dithiothreitol (DTT)) and the TEV proteolytic cleavage was conducted by incubation with TEV protease at 4 °C for 16 h. The cleaved $\mathbf{C_{D9k}}$ was recovered by a reverse IMAC step.

To load $\mathbf{C_{D9k}}$ with a Gd^{3+} ion, 1 mM $CaCl_2$ was added to a 10 μM $\mathbf{C_{D9k}}$ sample, which was then dialyzed against (i) 50 mM Tris-HCl pH 7.5, 400 mM NaCl, 10 μM gadolinium acetate, (ii) 50 mM Tris-HCl pH 7.5, 400 mM NaCl and finally (iii) 25 mM MES pH 6.5 buffers. After loading with gadolinium acetate, the $\mathbf{C_{D9k}}$ samples were concentrated and exchanged to the EPR buffer (50 mM MES in D_2O, pH 6.5, 50 mM NaCl; pH value is uncorrected pH meter reading) using Amicon ultrafiltration centrifugal tubes (Merck Millipore, USA) with 3 kDa molecular weight cutoff (MWCO), and perdeuterated glycerol was added to a final concentration of 10% (v/v) to reach a final protein concentration of 0.1 mM.

GB1 Synthesis. **GB1** samples were produced from a construct containing an N-terminal MASMTG tag (corresponding to the first six codons of the T7 gene 5) and a C-terminal His_6 tag (Fig. S2). The DNA sequences encoding different **GB1** mutants with the codon for F52 changed to the amber stop codon were cloned between the *Nde*I and *BamH*I restriction sites of pET-3a and pCDF plasmids. **GB1** samples containing CF_3-Phe were expressed using the protocol used for $\mathbf{C_{D9k}}$.

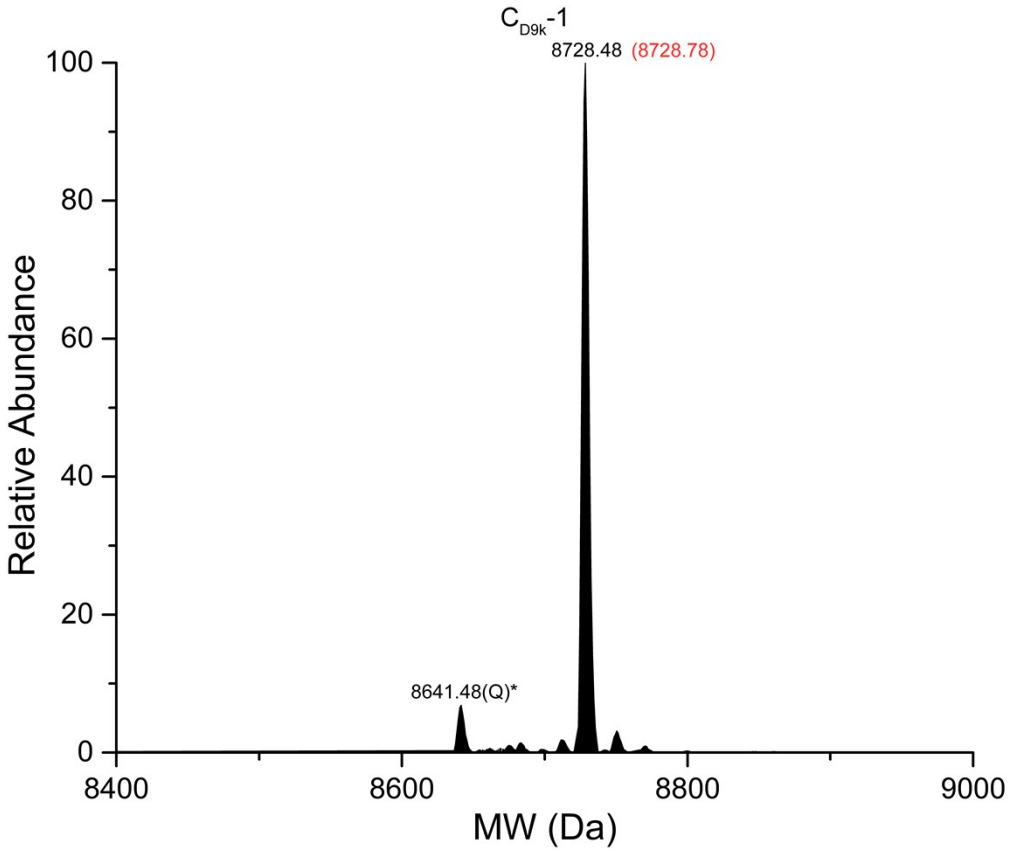
MASMTGMTYK LILNGKTLKG ETTTEAVDA**A** TAE**K**VFK**Q**YA
NDNGVDGEWT YDDATKTF̲TV TEHHHHHH

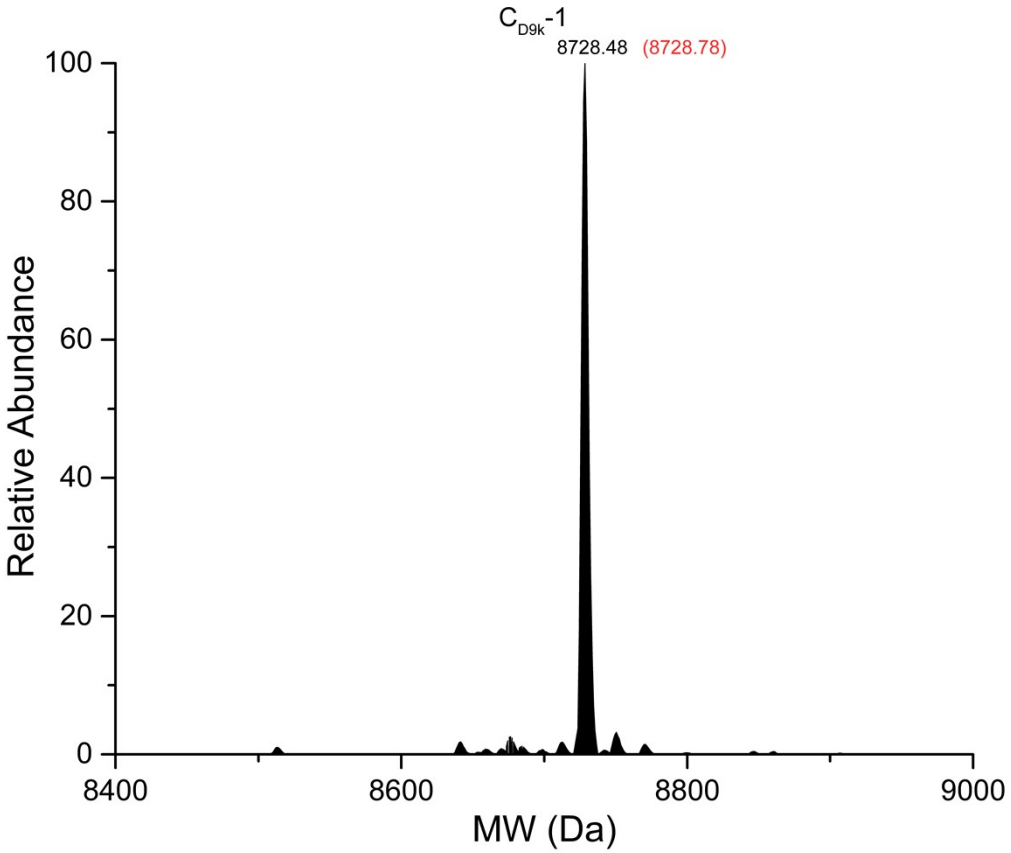
Fig. S2. Amino acid sequence of the **GB1** construct used. The N-terminal MASMTG tag is shown in red. Residues A24, K28 and Q32 are shown in bold and F52 is underlined.

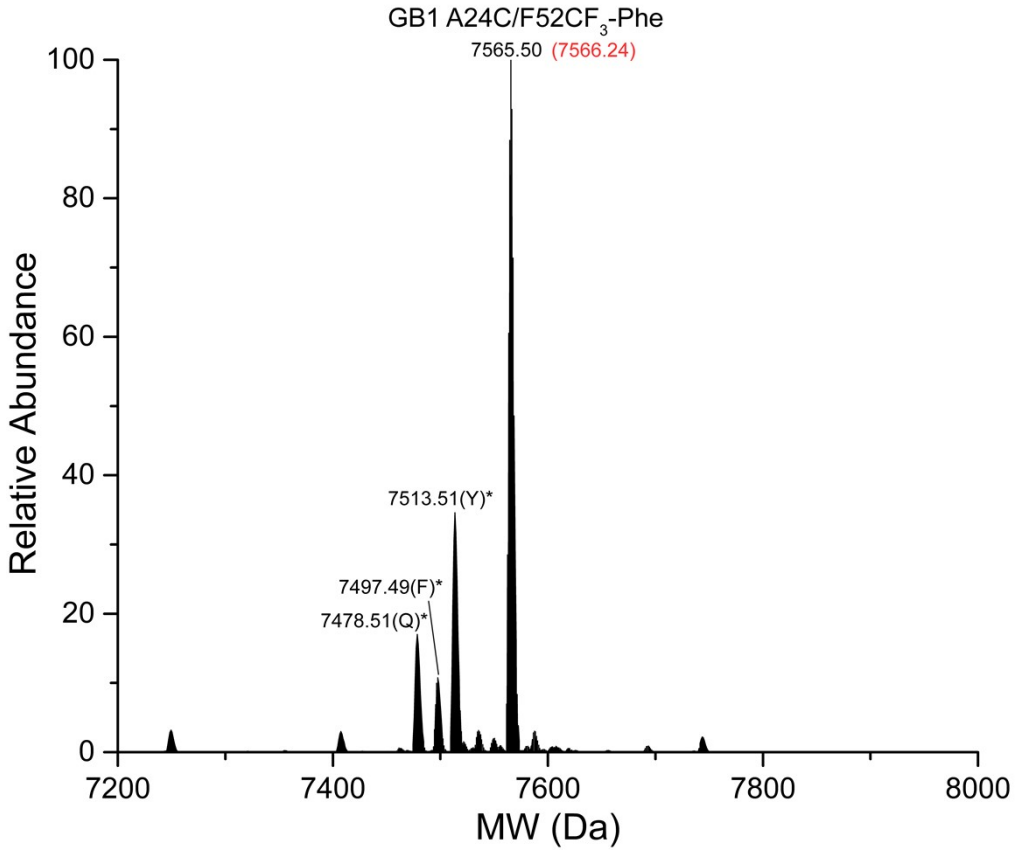
For cysteine ligation with **Gd.C1** or MTSL, 0.1 mM protein solution in 50 mM sodium phosphate, pH 7.5, was reduced with 2 mM of dithiothreitol (DTT) and the DTT was washed out using an Amicon ultrafiltration centrifugal tube with a molecular weight cutoff of 3 kDa. The reduced protein solution was

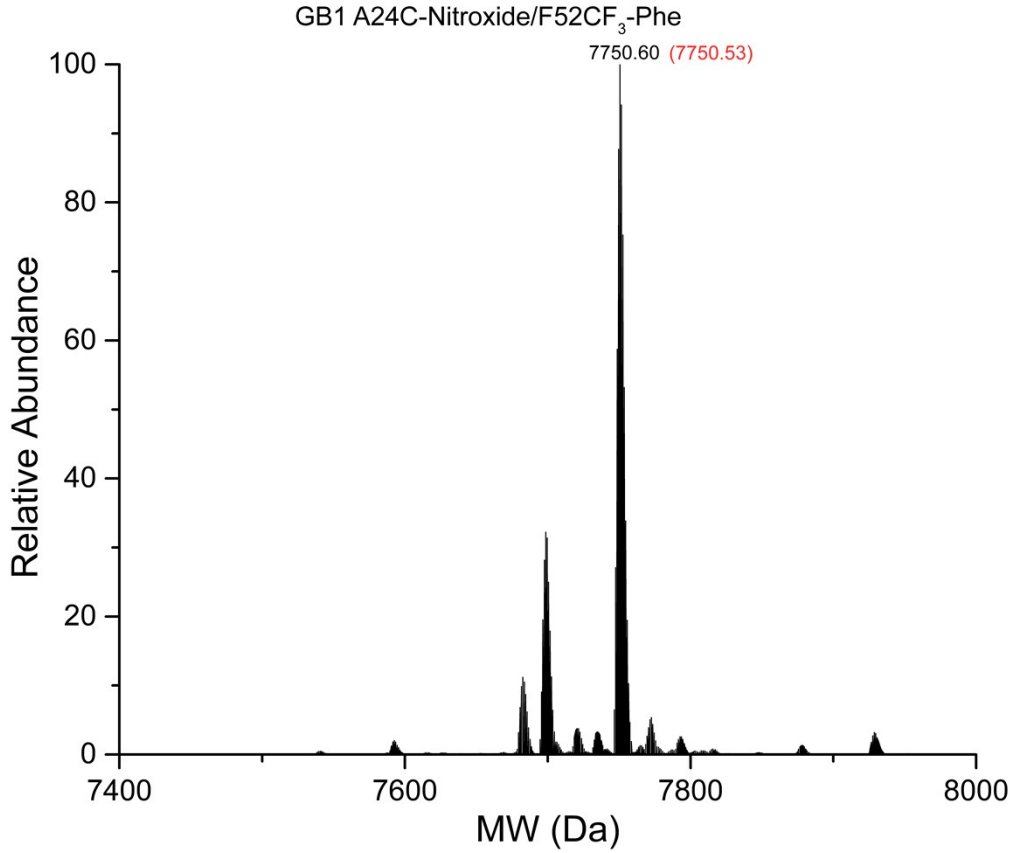
added slowly into a solution of 5 equivalents of **C1** tag or MTSL in the same buffer and kept at 25 °C overnight. Intact protein mass spectrometry indicated quantitative ligation yields. After the tagging reaction, the protein samples were concentrated and exchanged to EPR buffer (50 mM MES in D_2O, pH 6.5, 50 mM NaCl; pH value is uncorrected pH meter reading) using Amicon ultrafiltration centrifugal tubes with 3 kDa MWCO, and perdeuterated glycerol was added to a final concentration of 10% (v/v) to reach a final protein concentration of 0.1 mM.

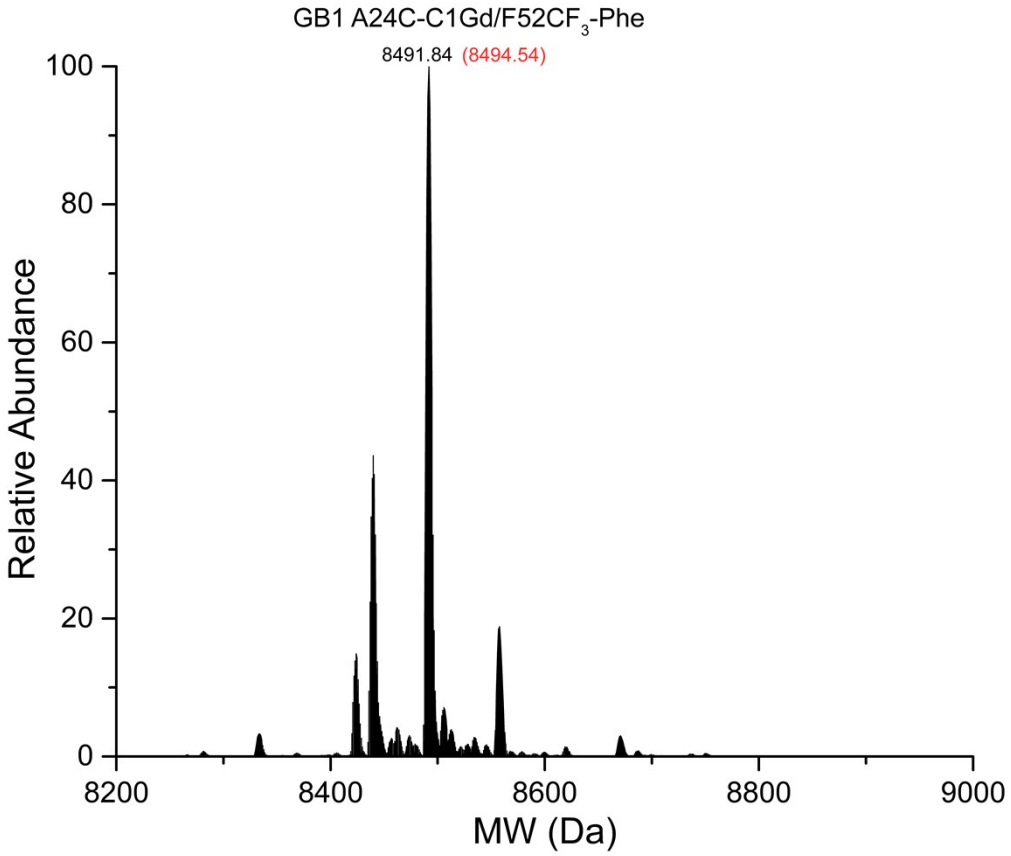
Intact protein mass spectrometry. Intact protein analysis was performed on an Orbitrap Fusion™ Tribrid™ mass spectrometer (Thermo Fisher Scientific, USA) connected to a Thermo Fisher Scientific UltiMate 3000 HPLC system equipped with a ZORBAX 300SB-C3, 3.5 μm, 4.6 x 50 mm HPLC column (Agilent Technologies, USA). Approximately 30 pmol of sample was injected using a 500 μL/min linear gradient of solvent A (0.1% (v/v) formic acid in water) and solvent B (0.1% (v/v) formic acid in acetonitrile), ramping solvent B from 5% at 2 min to 80% after 7 min. Data were collected using an electrospray ionization (ESI) source in positive ion mode. Protein intact mass was determined by deconvolution using the program Xcalibur 3.0.63 (Thermo Fisher Scientific, USA). The results of intact protein mass spectrometry are shown in Fig. S3.

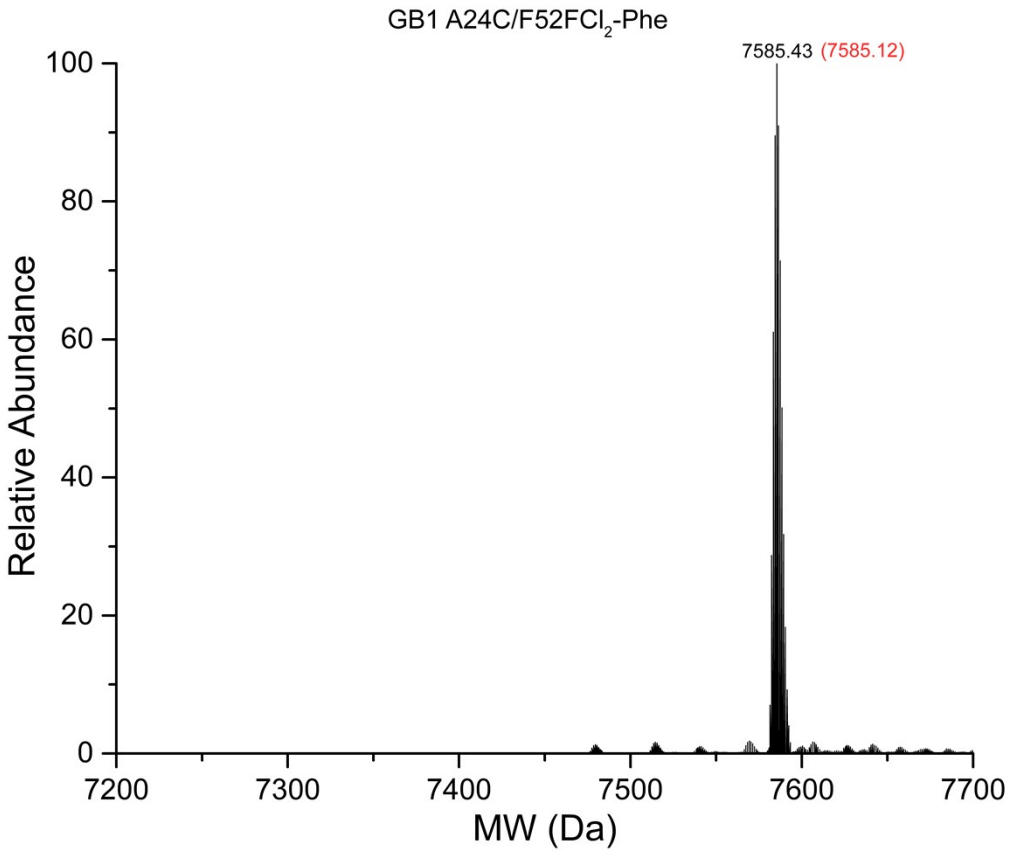

C D9k -1
8728.48 (8728.78)
8641.48(Q)*
Relative Abundance
MW (Da)
100
80
60
40
20
0
8400
8600
8800
9000

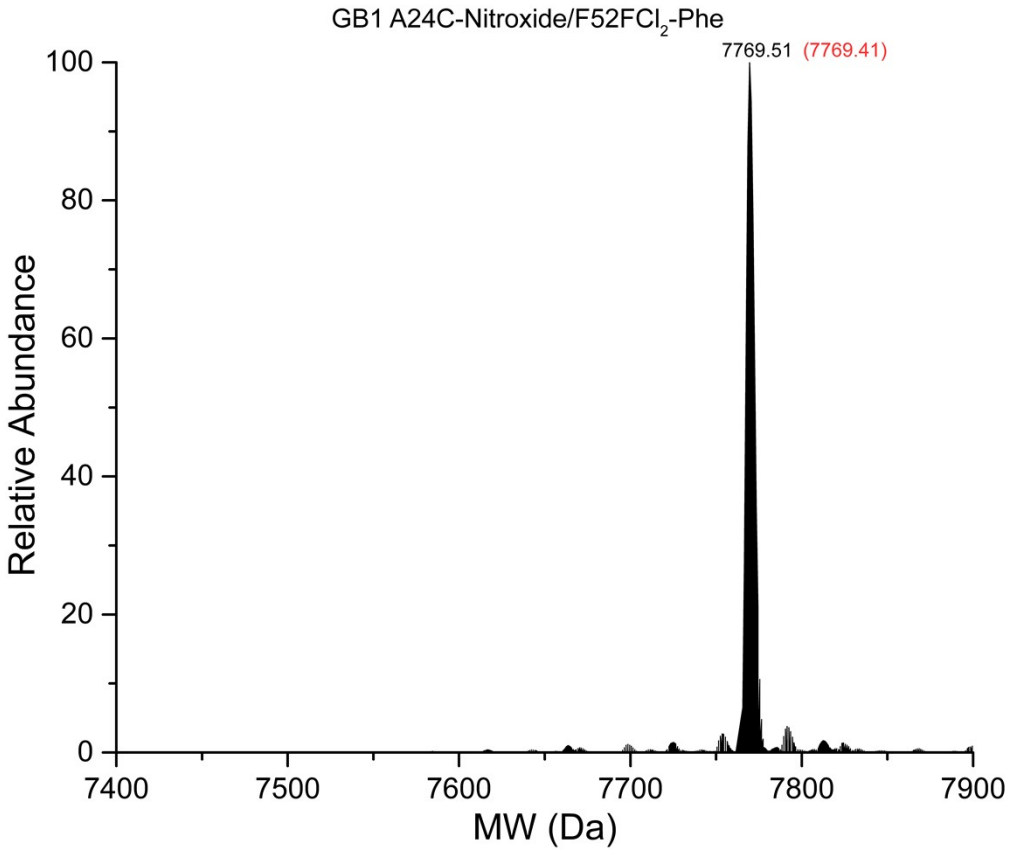

C D9k -1
8728.48 (8728.78)
Relative Abundance
MW (Da)
100
80
60
40
20
0
8400
8600
8800
9000

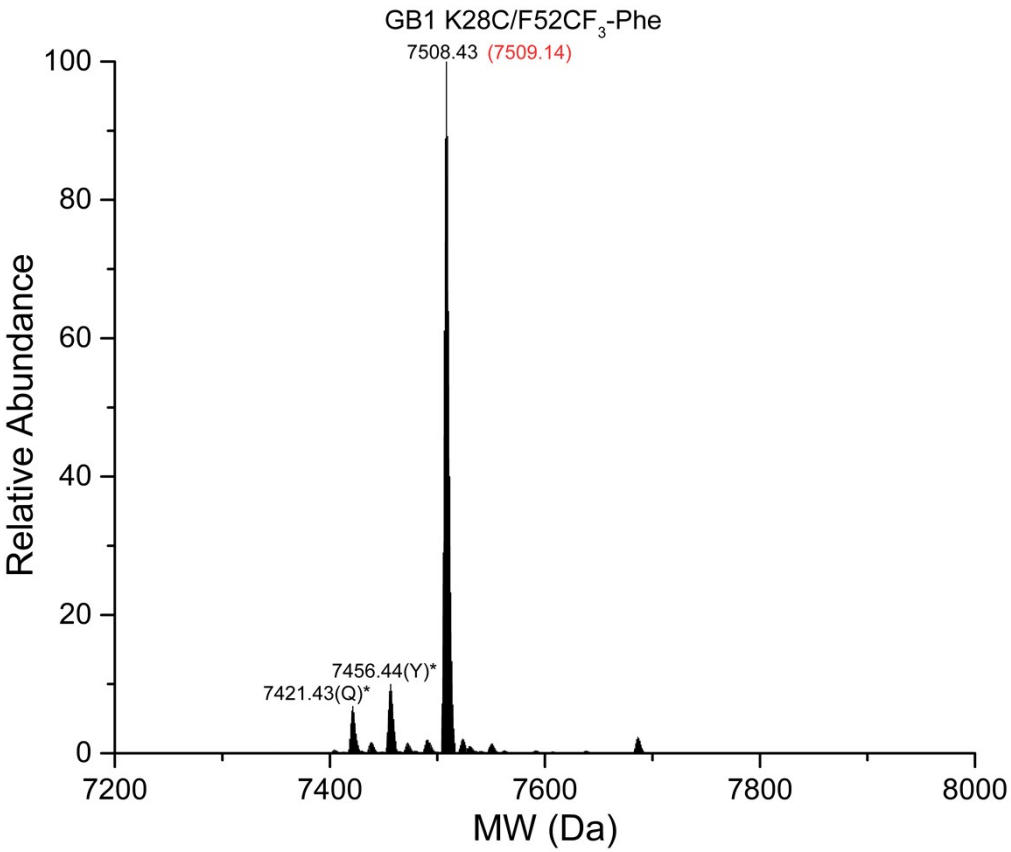

GB1 A24C/F52CF3-Phe
7565.50 (7566.24)
7513.51(Y)*
7497.49(F)*
7478.51(Q)*
Relative Abundance
MW (Da)
100
80
60
40
20
0
7200
7400
7600
7800
8000

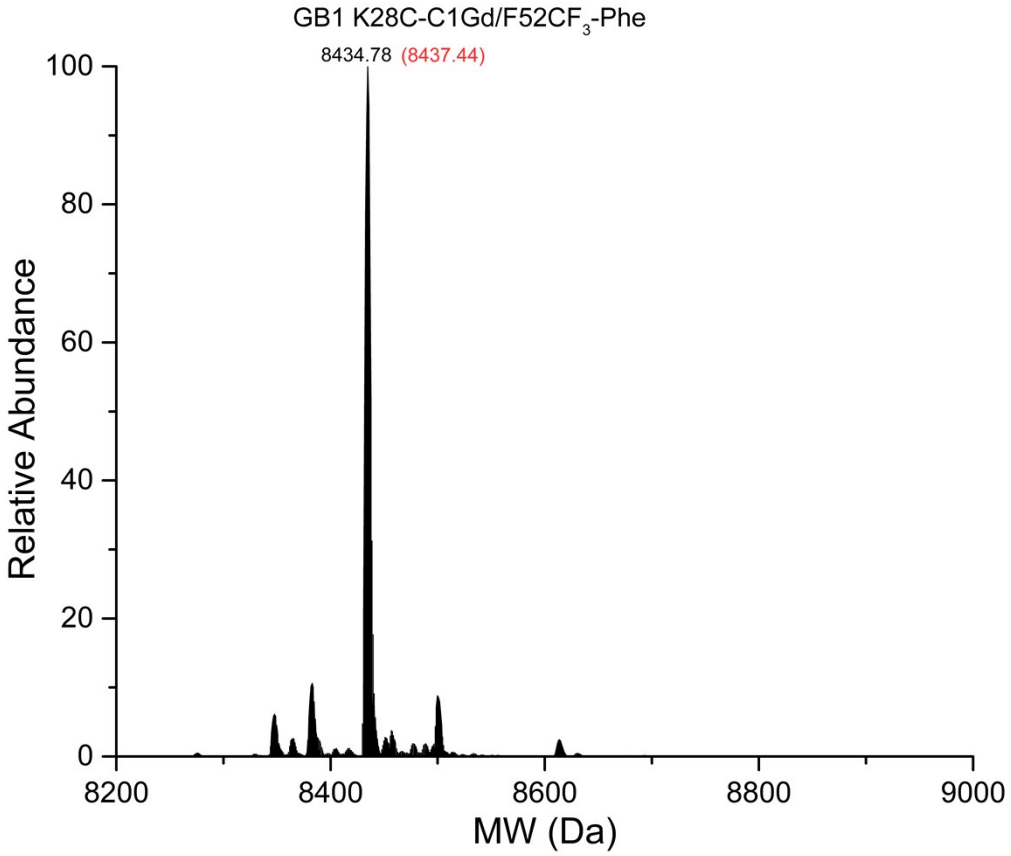

GB1 A24C-Nitroxide/F52CF3-Phe
7750.60 (7750.53)
Relative Abundance
MW (Da)
100
80
60
40
20
0
7400
7600
7800
8000

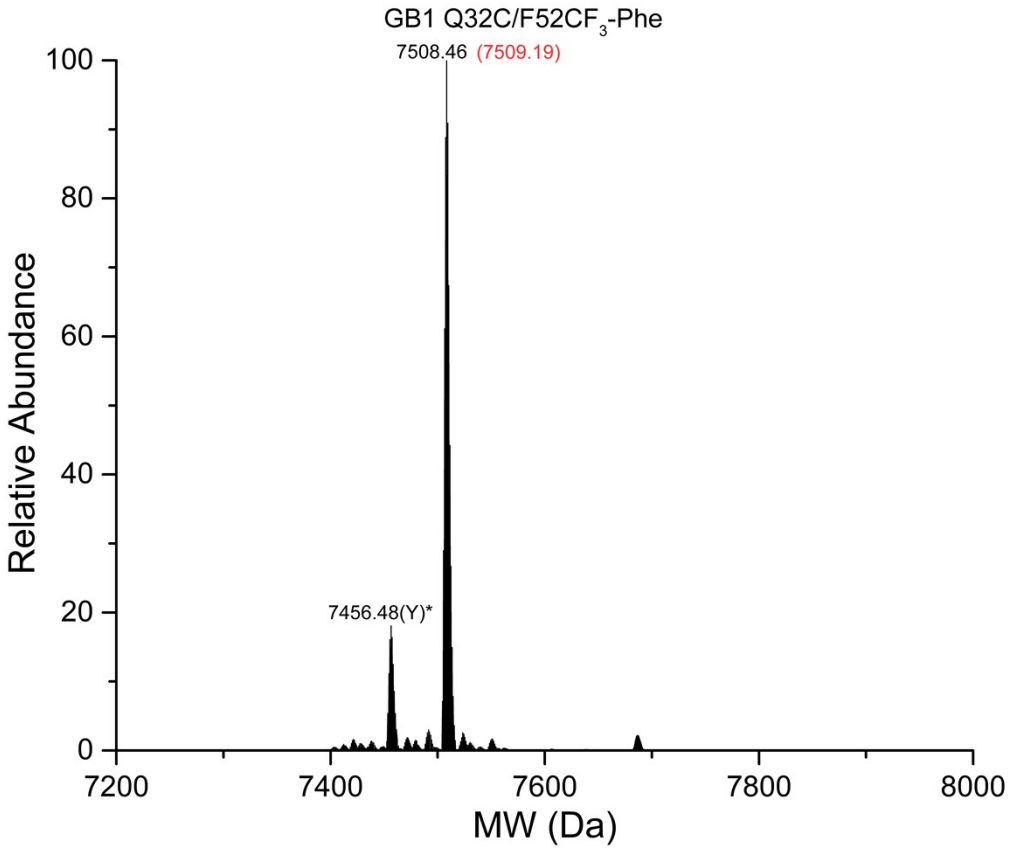

GB1 A24C-C1Gd/F52CF3-Phe
8491.84 (8494.54)
Relative Abundance
MW (Da)
100
80
60
40
20
0
8200
8400
8600
8800
9000


GB1 A24C/F52FCl2-Phe
7585.43 (7585.12)
Relative Abundance
MW (Da)
100
80
60
40
20
0
7200
7300
7400
7500
7600
7700


GB1 A24C-Nitroxide/F52FCl2-Phe
7769.51 (7769.41)
Relative Abundance
MW (Da)
0
20
40
60
80
100
7400
7500
7600
7700
7800
7900


GB1 K28C/F52CF3-Phe
7508.43 (7509.14)
7456.44(Y)*
7421.43(Q)*
Relative Abundance
MW (Da)
0
20
40
60
80
100
7200
7400
7600
7800
8000


GB1 K28C-C1Gd/F52CF3-Phe
8434.78 (8437.44)
Relative Abundance
MW (Da)
100
80
60
40
20
0
8200
8400
8600
8800
9000


GB1 Q32C/F52CF3-Phe
7508.46 (7509.19)
7456.48(Y)*
Relative Abundance
MW (Da)
100
80
60
40
20
0
7200
7400
7600
7800
8000

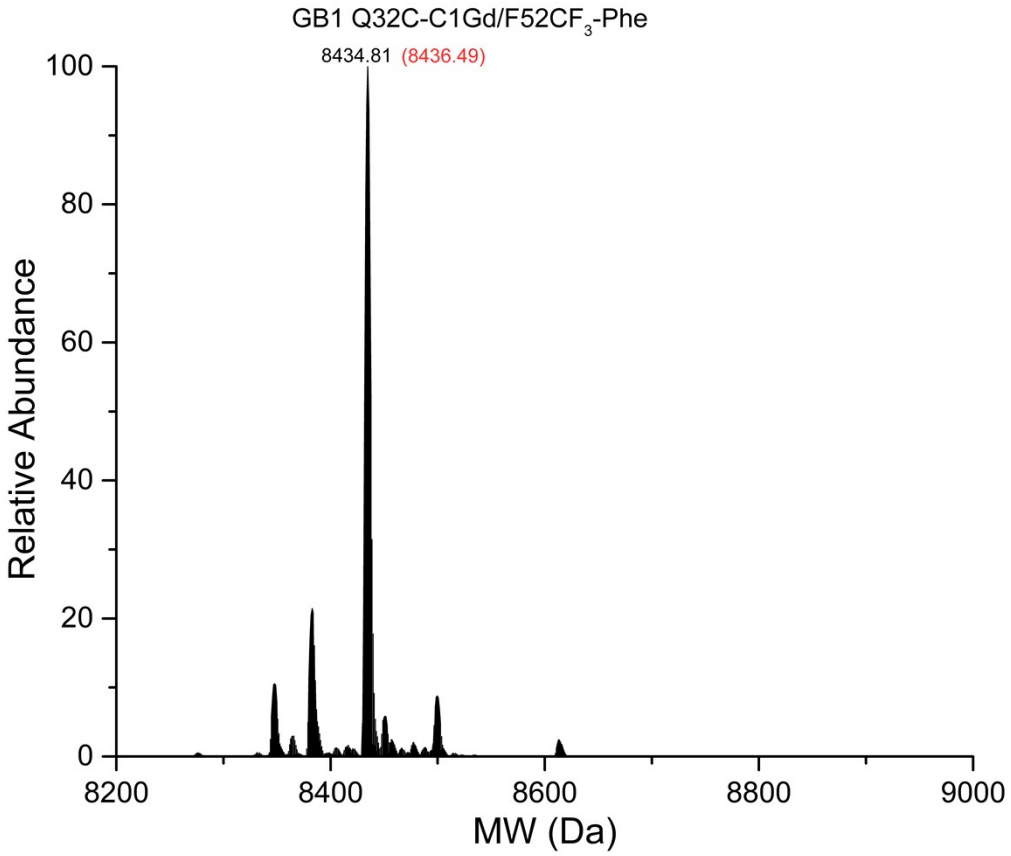
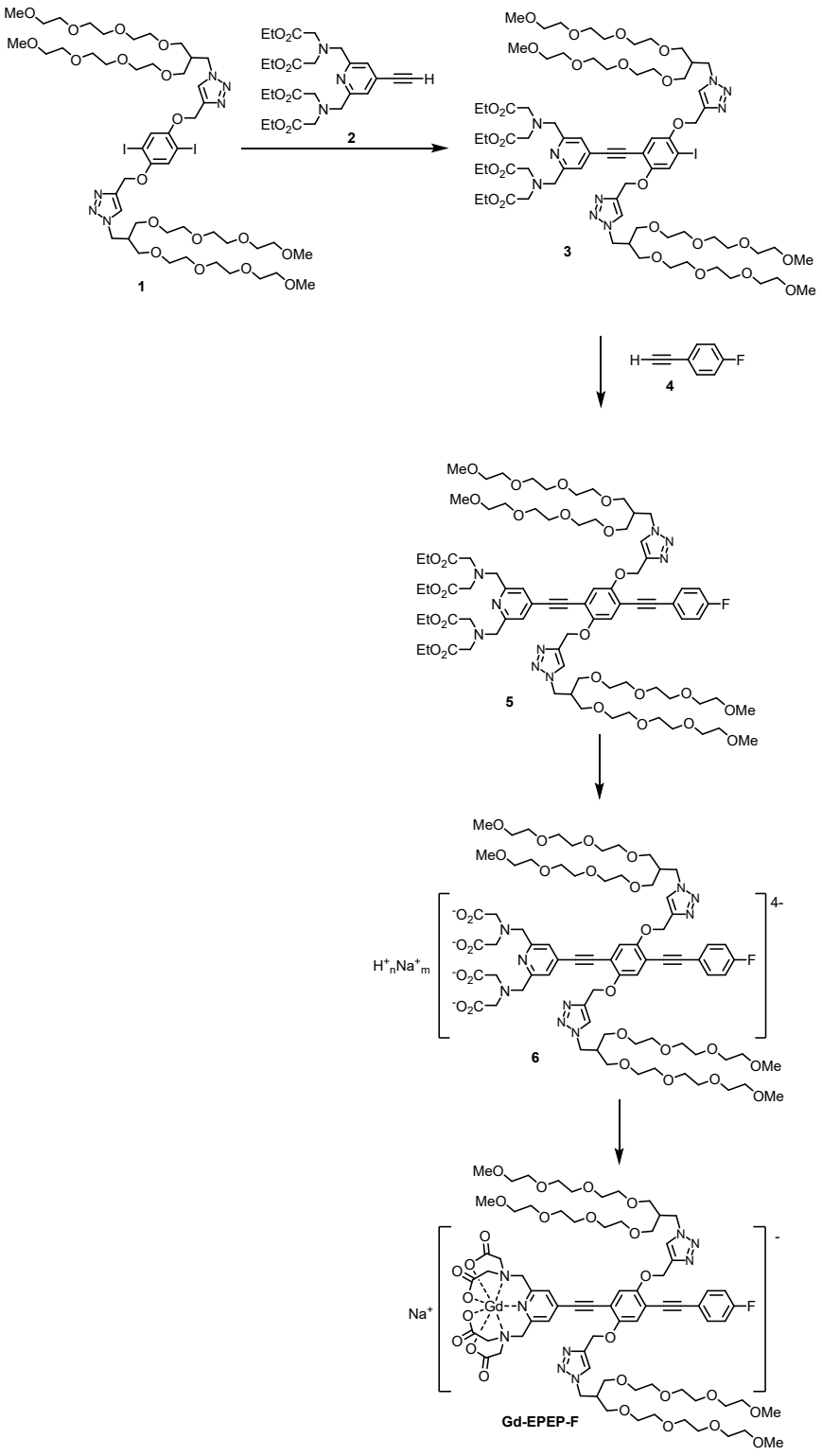


Figure S3. Intact protein mass spectrometric analysis results of the proteins used in this study. The calculated masses are shown in red. Masses marked with * are the products of amber codon suppression by canonical amino acids (F = phenylalanine, Q = glutamine, Y = tyrosine).

SI 2. Gd-EPEP-F Synthesis



Scheme S1. Synthesis of the Gd-^{19}F ruler **Gd-EPEP-F**.

To synthesize the Gd–^{19}F ruler **Gd-EPEP-F**, the pegylated diiodobenzene **1**[4] was coupled firstly with 4-ethynyl-PyMTA ester **2** and subsequently with 1-ethynyl-4-fluorobenzene **4**. After hydrolysis of the ester groups, the ruler precursor **6** was treated with $GdCl_3$ to obtain the rulers **Gd-EPEP-F**.

General

Unless otherwise stated, reactions were performed in dried glassware under argon using the Schlenk technique and commercial solvents and reagents. The argon was passed through anhydrous $CaCl_2$ prior to use. THF (HPLC grade) was dried with sodium/benzophenone. The solvents used for extraction and chromatography were of technical grade and were distilled prior to their use. Acetone-2-*^{13}C* (chemical purity 98.1%, isotopic enrichment 99.3%) was purchased from Cambridge Isotope Laboratories Inc. The proton-exchange resin (Dowex® 50WX4 hydrogen form, Sigma-Aldrich, 91 g) was subsequently washed with THF (3 × 200 mL), EtOH (2 × 100 mL), H_2O (2 × 150 mL), and EtOH (200 mL) and then dried over P_4O_{10} at 0.05 mbar for 5 days to obtain a pure and dry proton-exchange resin (30 g).

The temperatures given for the reactions refer to the bath temperatures. Solvents were removed at a bath temperature of ca. 40 °C and reduced pressure. The products were dried at room temperature at ca. 0.05 mbar. The pH values of the solutions were determined using pH indicator strips (resolution: 0.3 pH, Merck).

Column chromatography was carried out on silica gel 60 M (Macherey Nagel) applying slight pressure. In the procedures reported below the size of the column is given as diameter × length. The material was loaded onto the column dissolved in a small quantity of the eluent. Thin layer chromatography (TLC) was performed on silica gel coated aluminum foil (Merck, 60 F254). The spots were detected with UV light of λ = 254 and 366 nm. The compositions of solvent mixtures are given in volume ratios.

^1H NMR and ^{13}C NMR spectra were calibrated using the solvent signal as an internal standard [$CDCl_3$: $\delta(^1H)$ = 7.26, $\delta(^{13}C\{^1H\})$ = 77.16; CD_2Cl_2: $\delta(^1H)$ = 5.32, $\delta(^{13}C\{^1H\})$ = 54.00]. For ^{13}C NMR experiments in D_2O a drop of MeOH was added as the internal standard [$\delta(^{13}C)_{MeOH}$ = 49.5]. ^{19}F NMR spectra were calibrated using CCl_3F as an external standard. Signal assignments are supported by DEPT-135, COSY, HMBC and HMQC experiments.

EI mass spectra were recorded using an Autospec X magnetic sector mass spectrometer with EBE geometry (Vacuum Generators, Manchester, UK) equipped with a standard EI source. ESI mass spectra were recorded using an Esquire 3000 ion trap mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) equipped with a nano-ESI source. ESI accurate mass measurements were acquired using an Agilent 6220 time-of-flight mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) in extended dynamic range mode equipped with a Dual-ESI source or using a Q-IMS-TOF mass spectrometer Synapt G2Si (Waters GmbH, Manchester, UK) in resolution mode interfaced to a nano-ESI ion source.

The ratio of the components in a mixture was determined by ^1H NMR spectroscopy and is given in a molar ratio.

General procedure for alkynyl-aryl coupling. A solution of aryl iodide and terminal alkyne in THF and amine was degassed through three freeze-pump-thaw cycles. The solution was brought to room temperature. Then the catalysts were added, and the reaction mixture was stirred at the given temperature. Unless otherwise stated, shortly after the addition of the catalysts a precipitate formed.

Syntheses

Iodobenzene 3. See the general procedure for alkynyl-aryl coupling. Diiodobenzene **1**[4](1.01 g, 789 µmol), 4-ethynyl-PyMTA ethyl ester (**2**)[5] (200 mg, 396 µmol), $^{i}Pr_2NH$ (3.0 mL), THF (15 mL), catalysts: $Pd(PPh_3)_4$ (9.3 mg, 8.04 µmol), CuI (3.1 mg, 16.3 µmol). Room temperature. Reaction time: 22.5 h. After the reaction, in order to avoid oxidative damage of the PyMTA moiety, the copper ions were trapped with metal scavenger QuadraPure™ TU before the reaction mixture was exposed to air [6]. Under argon, all volatiles were evaporated and the residue was dissolved in degassed anhydrous CH_2Cl_2 (3 mL) and metal scavenger QuadraPure™ TU (80 mg) was added. The suspension was stirred at room temperature for 21 h. The suspension was filtered through a syringe filter (pore size: 0.45 µm, membrane: PTFE) and the resin was extracted with CH_2Cl_2 (3 × 1 mL). The filtrate and the extractants were combined. Removal of the solvents gave an orange-brown oil (1.34 g). Column chromatography (4.5 cm × 34 cm, CH_2Cl_2/EtOH 10:1) of the orange-brown oil gave iodobenzene **3** (207 mg, 31%; R_f = 0.45) as a yellow oil and a mixture of diiodobenzene **1** and iodobenzene **3**. Preparative HPLC of the mixed fraction (HPLC-column: phenomenex Luna Silica (2), particle size 5 µm, pore size 100 Å, length 250 mm, internal diameter 21.2 mm; elution: isocratic, 15 mL/min, CH_2Cl_2/EtOH 94:6; fraction between 7.4 min and 8.9 min was collected) gave iodobenzene **3** (120 mg, 18%) as a yellow oil. ^{1}H NMR (500 MHz, CD_2Cl_2): δ = 7.92 and 7.90 (2 s, 1 H each, $H_{triazole}$), 7.57 (s, 1 H, $H_{benzene}$ *ortho* to I), 7.53 (s, 2 H, $H_{pyridine}$), 5.28 and 5.18 (2 s, 2 H each, $C_{triazole}CH_2$), 4.52 and 4.50 (2 d, $^3J_{HH}$ = 6.4 Hz, 2 H each, $N_{triazole}CH_2$), 4.13 (q, $^3J_{HH}$ = 7.1 Hz, 8 H, C$\underline{H}_2$CH$_3$), 3.98 (s, 4 H, $C_{pyridine}CH_2$), 3.57 (s, 8 H, $NCH_2C{=}O$), 3.62 – 3.32 (m, 56 H, CH_2OCH_2), 3.30 and 3.29 (2 s, 6 H each, OCH_3), 2.48 (m, 2 H, C$\underline{H}(CH_2)_3$), 1.24 (t, $^3J_{HH}$ = 7.1 Hz, 12 H, $CH_2C\underline{H}_3$). ^{13}C NMR (126 MHz, CD_2Cl_2): δ = 171.6 (C=O), 159.3 ($C_{pyridine}$ *ortho* to N), 154.6 ($C_{benzene}$O *meta* to I), 152.4 ($C_{benzene}$O *ortho* to I), 143.38 and 143.32 ($\underline{C}_{triazole}CH_2$), 132.7 ($C_{pyridine}$ *para* to N), 126.0 ($C_{benzene}$H *ortho* to I), 125.66 and 125.53 ($C_{triazole}$H), 123.6 ($C_{pyridine}$H), 117.6 ($C_{benzene}$H *meta* to I), 113.8 ($C_{benzene}$ *para* to I), 92.8 (pyridine-$\underline{C}{\equiv}C$), 89.7 ($C_{benzene}$I), 89.2 (pyridine-C$\equiv\underline{C}$), 77.44, 77.42, 71.19, 71.12, 71.11, 71.07, 71.05, 71.04, 70.93, 70.90, 70.84, 70.78, 69.48, and 69.46 ($\underline{C}H_2O\underline{C}H_2$), 64.69 and 64.42 ($C_{triazole}\underline{C}H_2$), 61.1 ($\underline{C}H_2CH_3$), 60.3 (pyridine-$\underline{C}H_2$), 59.15 and 59.13 ($OCH_3$), 55.4 ($\underline{C}H_2C{=}O$), 49.1 ($N_{triazole}CH_2$), 41.2 ($\underline{C}H(CH_2)_3$), 14.6 ($CH_2\underline{C}H_3$). MS (ESI) *m/z* = 1684.8 $[M + Na]^+$, 1662.8 $[M + H]^+$, 853.9 $[M + 2 Na]^{2+}$. Accurate MS (ESI) *m/z* calcd. for $[M + H]^+$ $C_{73}H_{116}IN_9O_{26}H^+$: 1662.7149; found: 1662.7157.

PyMTA ester-EPEP-F (5). See the general procedure for alkynyl-aryl coupling. Iodobenzene **3** (41 mg, 25 µmol), 1-ethynyl-4-fluorobenzene (**4**) (10.2 mg, 84.9 µmol), $^{i}Pr_2NH$ (0.5 mL), THF (3 mL), catalysts:

$Pd(PPh_3)_4$ (1.44 mg, 1.25 µmol), $Pd_2(dba)_3$ (1.22 mg, 1.33 µmol), CuI (0.95 mg, 4.99 µmol). Room temperature. Reaction time: 3 d. In order to avoid oxidative damage of the PyMTA moiety, the copper ions were trapped with metal scavenger QuadraPure™ TU before the reaction mixture was exposed to air [6]. Under argon, all volatiles were evaporated and the residue was dissolved in degassed anhydrous CH_2Cl_2 (3 mL) and metal scavenger QuadraPure™ TU (89 mg) was added. The suspension was stirred at room temperature for 3 d. The suspension was filtered through a syringe filter (pore size: 0.45 µm, membrane: PTFE) and the resin was extracted with CH_2Cl_2 (3 × 1 mL). The filtrate and the extractants were combined. Removal of the solvents gave a pale-brown oil with precipitate. Column chromatography (2.8 cm × 24 cm, CH_2Cl_2/EtOH 10:1) gave PyMTA ester-EPEP-F **5** (37 mg, 89%, R_f = 0.40) with small amount of $O=PPh_3$ as a yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.87 and 7.86 (2 s, 1 H each, $H_{triazole}$), 7.55 (s, 2 H, $H_{pyridine}$), 7.44 (m, 2 H, $H_{benzene}$ *meta* to F), 7.17 and 7.15 (2 s, 1 H each, $H_{benzene}$ *para* to O), 7.02 (m, 2 H, $H_{benzene}$ *para* to F), 5.32 and 5.24 (2 s, 2 H each, $C_{triazole}CH_2$), 4.49 and 4.48 (2 d, $^3J_{HH}$ = 6.7 Hz, 2 H each, $N_{triazole}CH_2$), 4.14 (q, $^3J_{HH}$ = 7.1 Hz, 8 H, $C\underline{H}_2CH_3$), 4.02 (s, 4 H, $C_{pyridine}$-CH_2), 3.60 (s, 8 H, $NCH_2C=O$), 3.64 – 3.30 (m, 56 H, CH_2OCH_2), 3.307 and 3.305 (2 s, 6 H each, OCH_3), 2.45 (m, 2 H, $C\underline{H}(CH_2)_3$), 1.24 (t, $^3J_{HH}$ = 7.1 Hz, 12 H, $CH_2C\underline{H}_3$). ^{19}F NMR (471 MHz, $CDCl_3$): δ = -110.2. ^{13}C NMR (126 MHz, $CDCl_3$): δ = 171.2 (C=O), 162.7 (d, $^1J_{CF}$ = 250 Hz, $C_{benzene}F$), 158.8 ($C_{pyridine}$ *ortho* to N), 153.5 and 153.3 ($C_{benzene}O$), 143.44 and 143.41 ($\underline{C}_{triazole}CH_2$), 133.63 (d, $^3J_{CF}$ = 8.4 Hz, $C_{benzene}$ *meta* to F), 132.5 ($C_{pyridine}$ *para* to N), 124.9 and 124.8 ($C_{triazole}H$), 123.2 ($C_{pyridine}H$), 119.3 (d, $^4J_{CF}$ = 3.5 Hz, $C_{benzene}$ *para* to F), 118.6 and 118.4 ($C_{benzene}H$), 115.8 (d, $^2J_{CF}$ = 21.7 Hz, $C_{benzene}$ *ortho* to F), 115.5 and 113.6 ($\underline{C}_{benzene}$-C≡C), 94.7, 93.1, 89.6, and 85.2 (C≡$\underline{C}$), 77.36, 72.0, 71.96, 70.69, 70.60, 70.57, 70.56, 70.40, 69.03 ($\underline{C}H_2O\underline{C}H_2$), 64.0 and 63.9 ($C_{triazole}\underline{C}H_2$), 60.6 ($\underline{C}H_2CH_3$), 59.8 (pyridine-$\underline{C}H_2$), 59.1 ($OCH_3$), 55.0 ($\underline{C}H_2C=O$), 48.72 and 48.67 ($N_{triazole}CH_2$), 40.62 and 40.59 ($\underline{C}H(CH_2)_3$), 14.3 ($CH_2\underline{C}H_3$). MS (ESI) *m/z* = 1676.8 $[M + Na]^+$, 1654.8 $[M + H]^+$, 849.9 $[M + 2\ Na]^{2+}$, 838.9 $[M + H + Na]^{2+}$. Accurate MS (ESI) *m/z* calcd. for $[M + H]^+$ $C_{81}H_{120}FN_9O_{26}H^+$: 1654.84014; found: 1654.84042.

H_nNa_m[PyMTA-EPEP-F-] 6. PyMTA ester-EPEP-F **5** (36 mg, 21.8 µmol) was dissolved in EtOH (1 mL). H_2O (800 µL) and 1 M aqueous NaOH solution (218 µL, 218 µmol) were added. The pale yellow solution was stirred at room temterature for 27 h. Solvents of the reaction solution were removed. The residue was dissolved in H_2O (3 mL) and this aqueous solution was washed with CH_2Cl_2 (2 × 3 mL). The washing was performed in a centrifuge tube: CH_2Cl_2 and the aqueous solution were mixed well. Centrifugation of the resulting yellow emulsion at 6500 rpm for 2 min separated the mixture into two phases, an organic phase and an aqueous phase. The organic phase was removed with the help of a glass pipette. To the well-washed aqueous solution, proton-exchange resin (40 mg) was added. This way, the pH of the solution was lowered to ca. 3.9. The solution was separated from the resin through filtration through a syringe filter (pore size: 0.45 µm, membrane: PVDF) and the resin was extracted with a mixture of EtOH/H_2O 1:1 (4 × 1 mL). The filtrate and the extractants were combined. Removal of the solvents gave H_nNa_m[PyMTA-EPEP-F] **6** (33 mg, 85%) as a yellow film. The content of the structural motive $[PyMTA\text{-}EPEP\text{-}F]^{4-}$ was determined by quantitative NMR spectroscopy to be 86.7 wt.%. ^{1}H NMR (500 MHz, D_2O): δ = 8.04 and 8.00 (2 s, 1 H each, $H_{triazole}$),

7.60 (s, 2 H, $H_{pyridine}$), 7.44 (m, 2 H, $H_{benzene}$ *ortho* to F), 7.42 and 7.14 (2 s, 1 H each, $H_{benzene}$ *ortho* to O), 5.38 and 5.30 (2 s, 2 H each, $C_{triazole}CH_2$), 4.65 (br. s, 4 H, $C_{pyridine}CH_2$), 4.52 and 4.50 (2 d, $^3J_{HH}$ = 5.8 Hz, 2 H each, $N_{triazole}CH_2$), 3.89 (br. s, 8 H, $NCH_2C{=}O$), 3.62 – 3.16 (m, 56 H, CH_2OCH_2), 3.28 (s, 12 H, OCH_3), 2.43 (m, 2 H, $C\underline{H}(CH_2)_3$).

Na[{Gd^{III}(PyMTA)}-EPEP-F] (Gd-EPEP-F). H_nNa_m[PyMTA-EPEP-F] **6** (33 mg) was dissolved in D_2O (1500 µL). The concentration of the structural motif $[PyMTA\text{-}EPEP\text{-}F]^{4-}$ was determined as 12.4 mM by quantitative ^1H NMR spectroscopy using an external standard. A part of the obtained solution (212.8 µL, contains 2.639 µmol of the structural motif $[PyMTA\text{-}EPEP\text{-}F]^{4-}$) was mixed with a solution of $GdCl_3 \bullet 6\ H_2O$ in D_2O (50 mM, 50.0 µL, 2.50 µmol). A solution of NaOD in D_2O (0.10 M, 75 µL, 7.5 µmol) and a solution of DCl (0.10 M, 3 µL, 0.3 µmol) was added to adjust the pH of the solution to pH 6.8. The solution was diluted with D_2O (up to a total volume of 500 µL) to obtain a 5.0 mM solution of Na[{Gd^{III}(PyMTA)}-EPEP-F] (**Gd-EPEP-F**) in D_2O. Accurate MS (ESI) *m/z* calcd. for [{Gd^{III}(PyMTA)}-EPEP-F]$^-$ $C_{73}H_{100}N_9O_{26}FGd^-$: 1695.6010; found: 1695.5972.

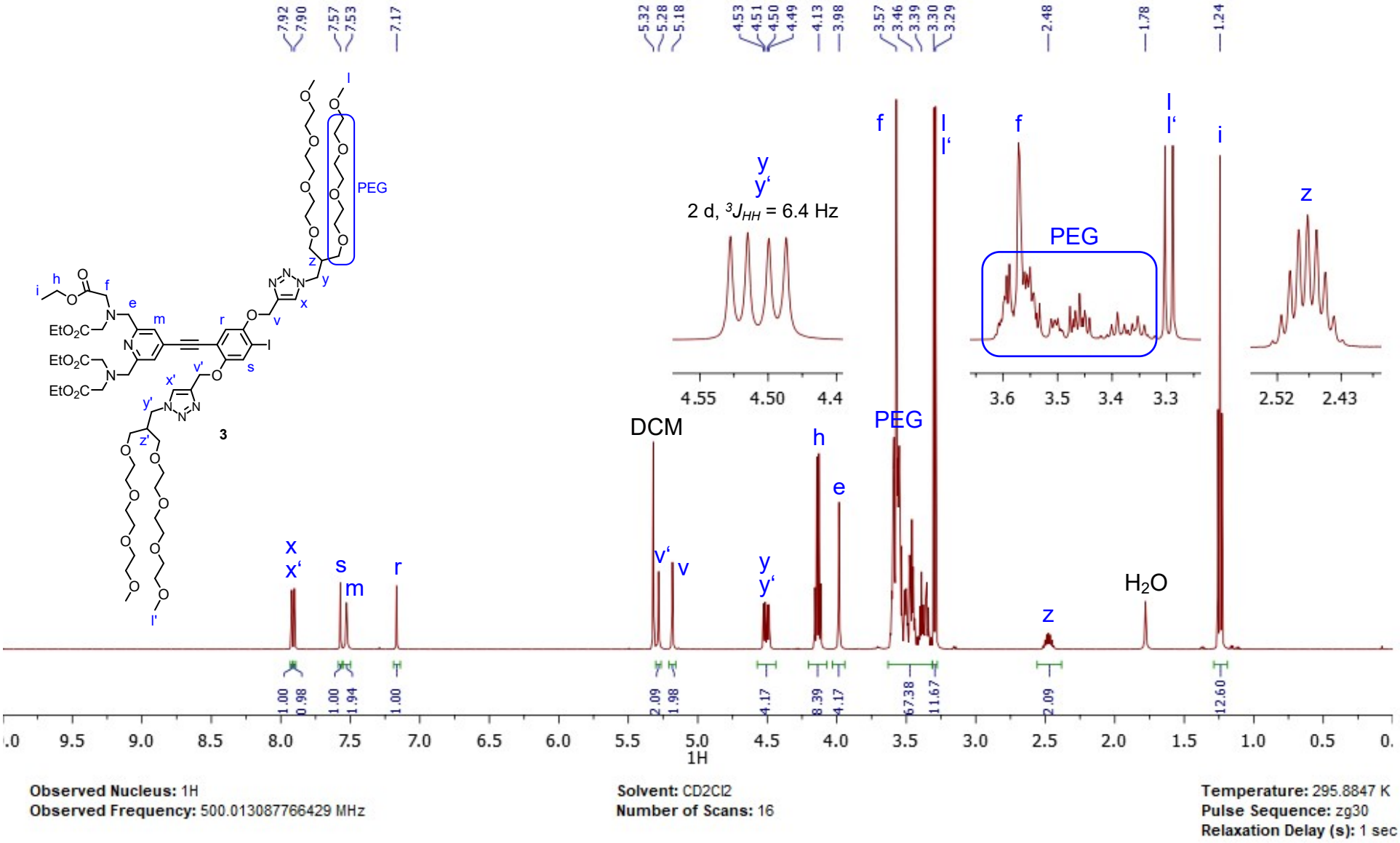


Fig. S4.1 ^{1}H-NMR spectrum (500 MHz, $CDCl_3$) of iodobenzene **3**.

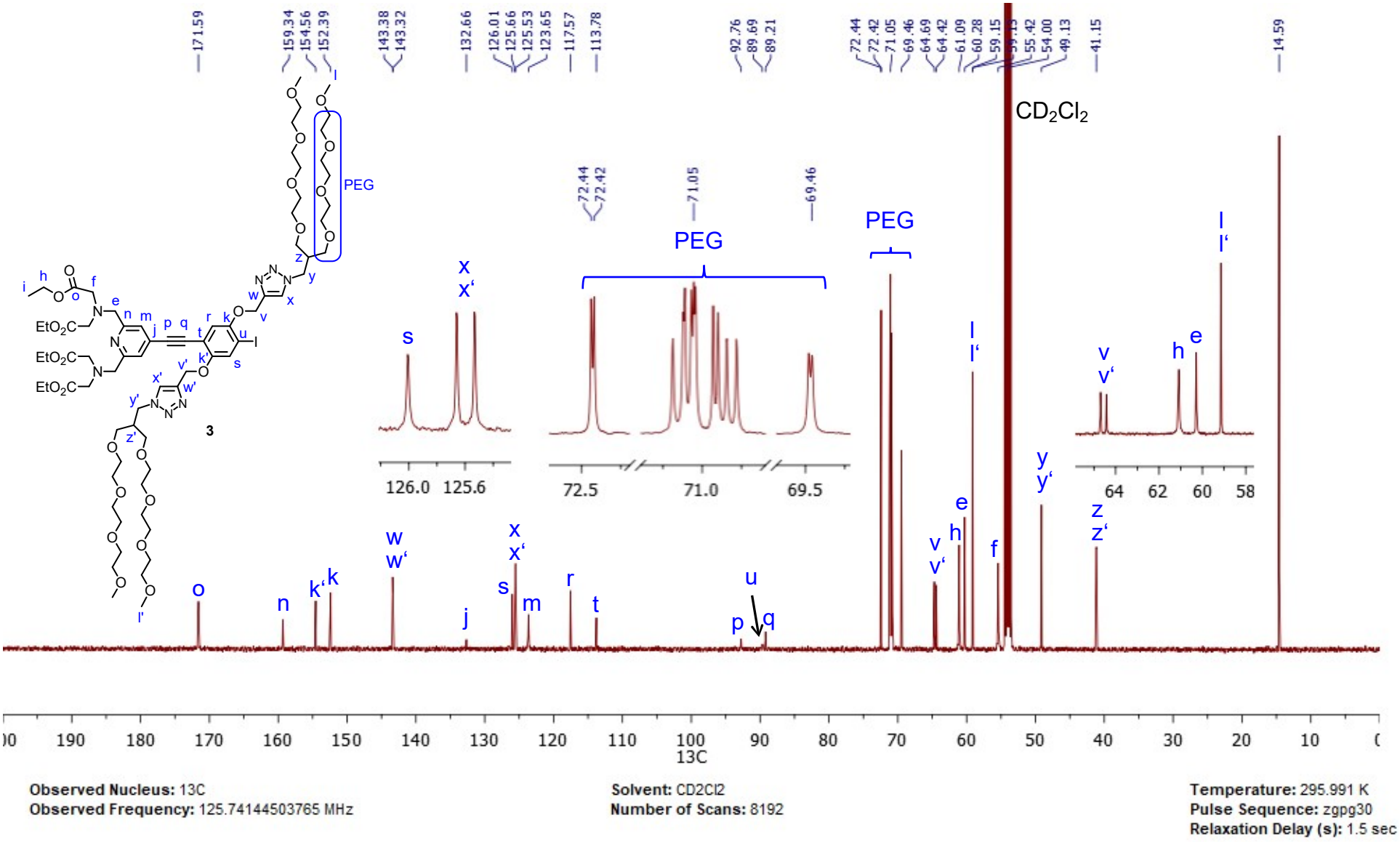


Fig. S4.2 ^{13}C-NMR spectrum (126 MHz, CD_2Cl_2) of iodobenzene **3**.

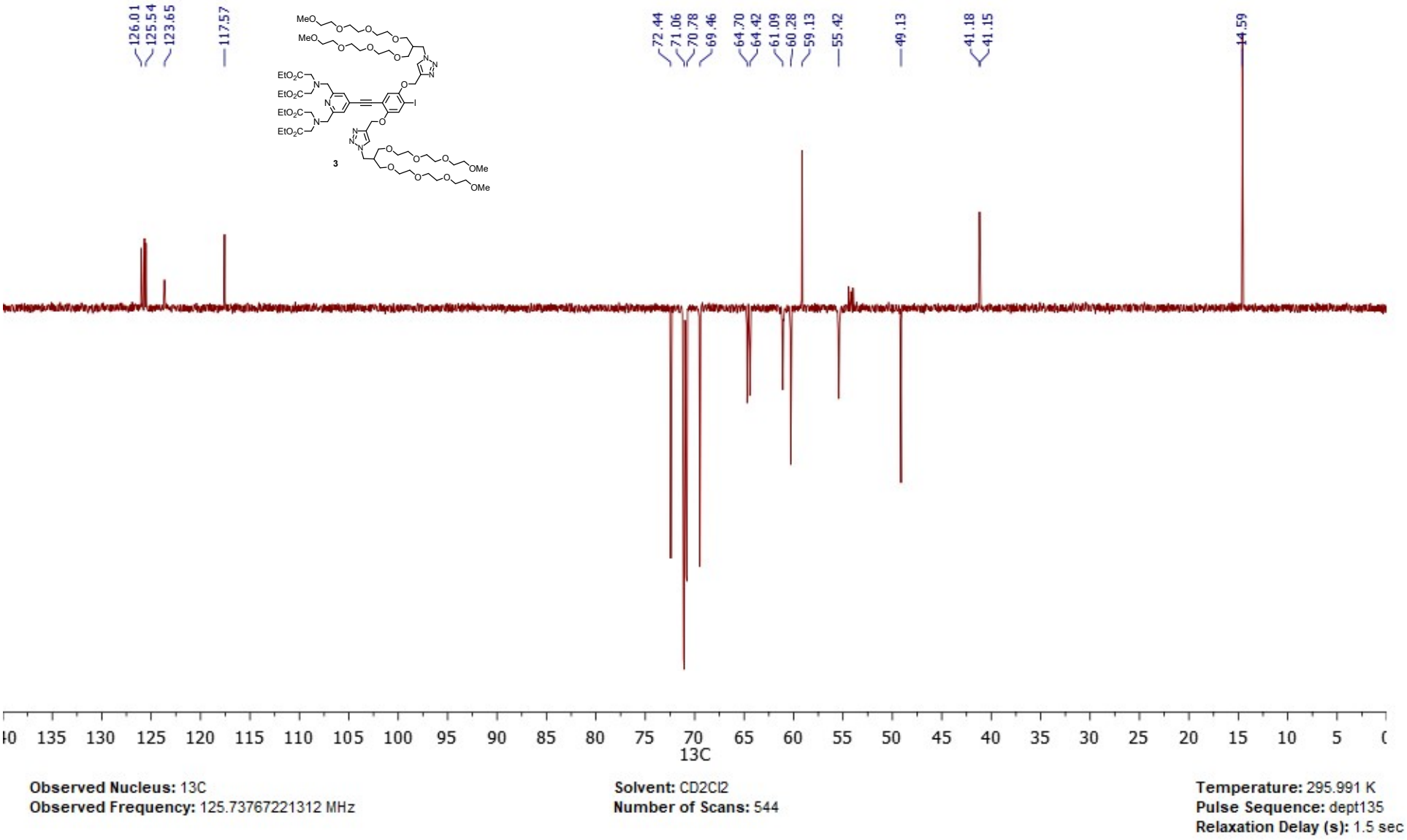


Fig. S4.3 ^{13}C DEPT 135 spectrum (126 MHz, CD_2Cl_2) of iodobenzene **3**.

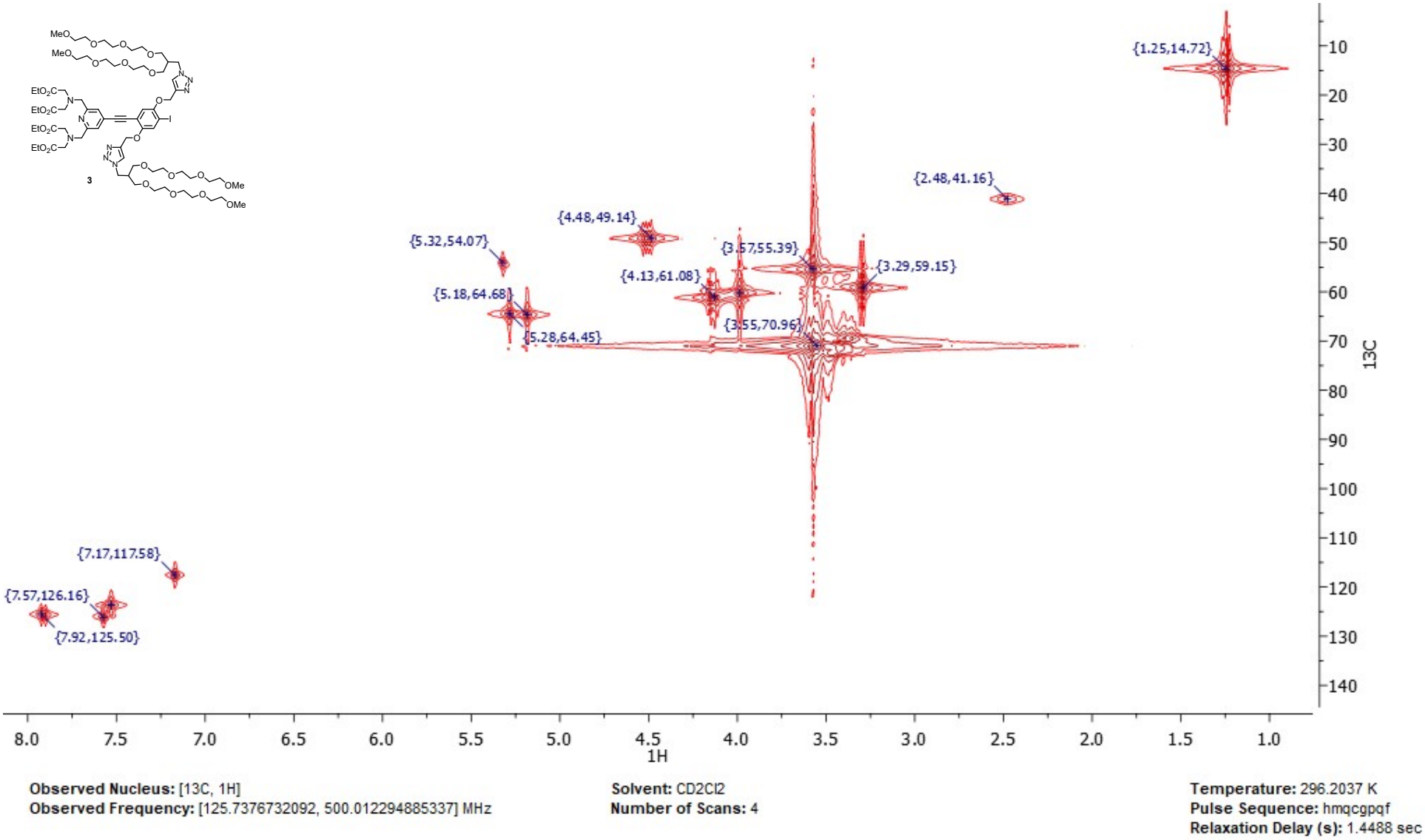


Fig. S4.4. HMQC spectrum (126 MHz, 500 MHz, CD_2Cl_2) of iodobenzene **3**.

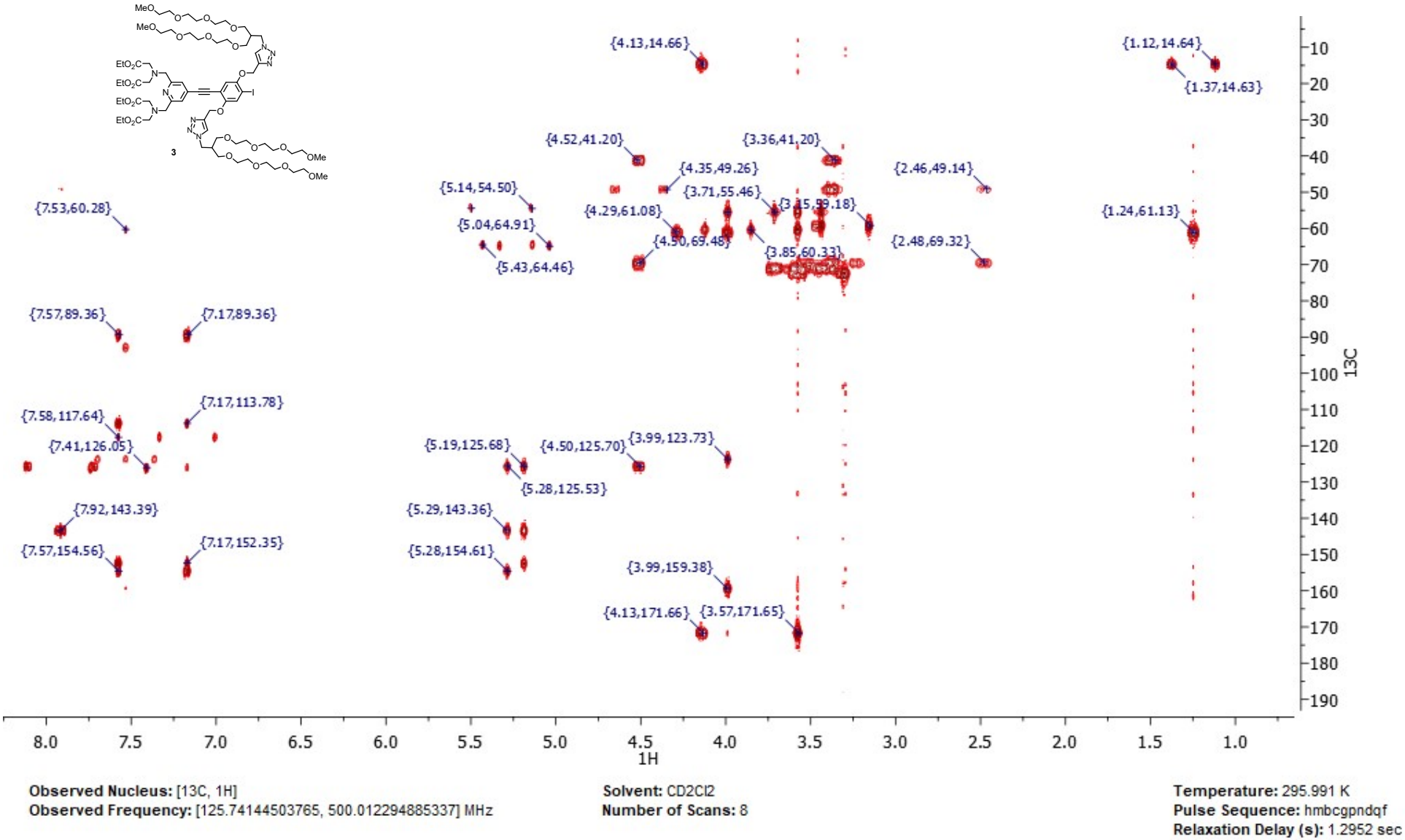


Figure S4.5. HMBC spectrum (126 MHz, 500 MHz, CD_2Cl_2) of iodobenzene **3**.

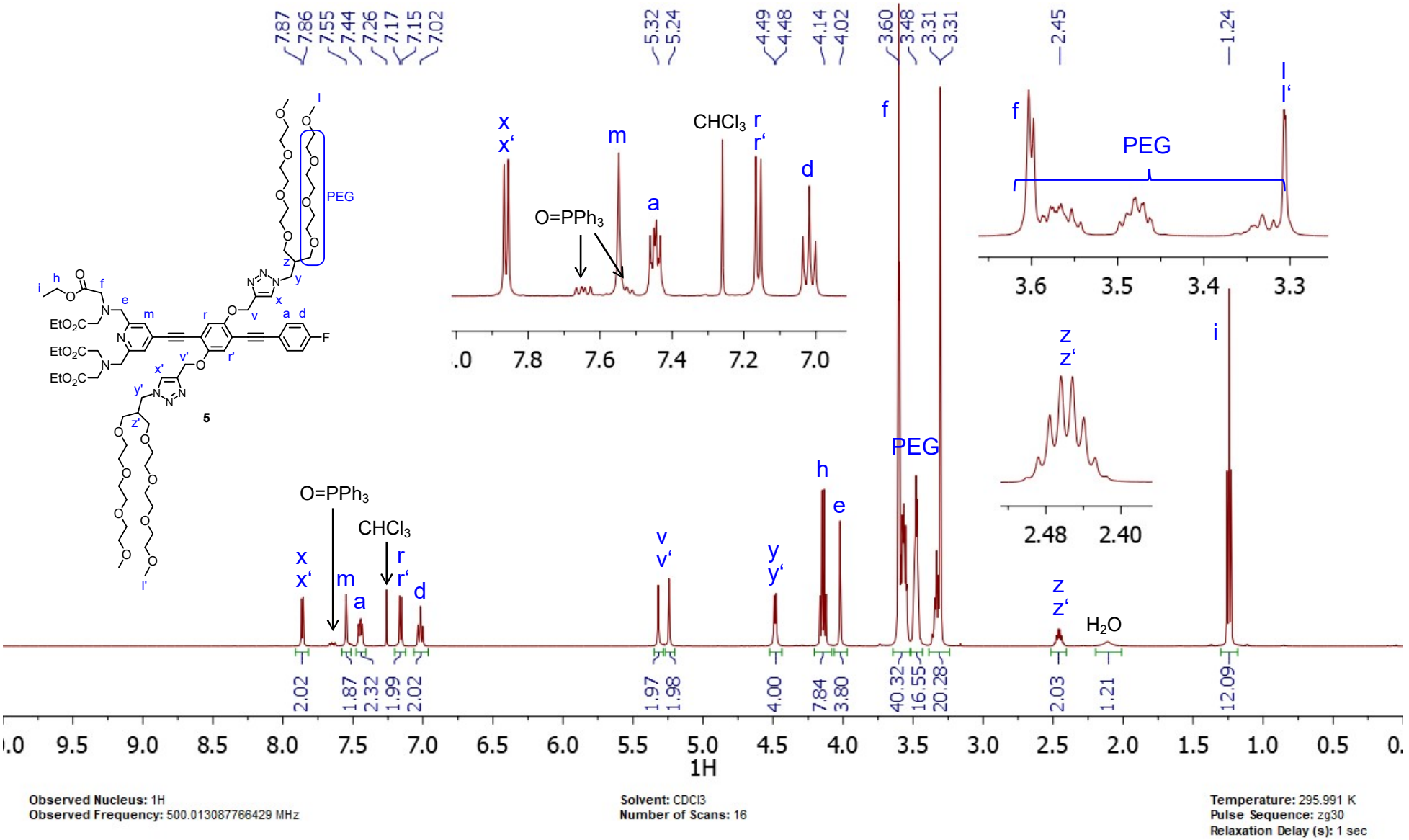


Figure S4.6. ^{1}H-NMR spectrum (500 MHz, $CDCl_3$) of PyMTA ester-EPEP-F **5**.

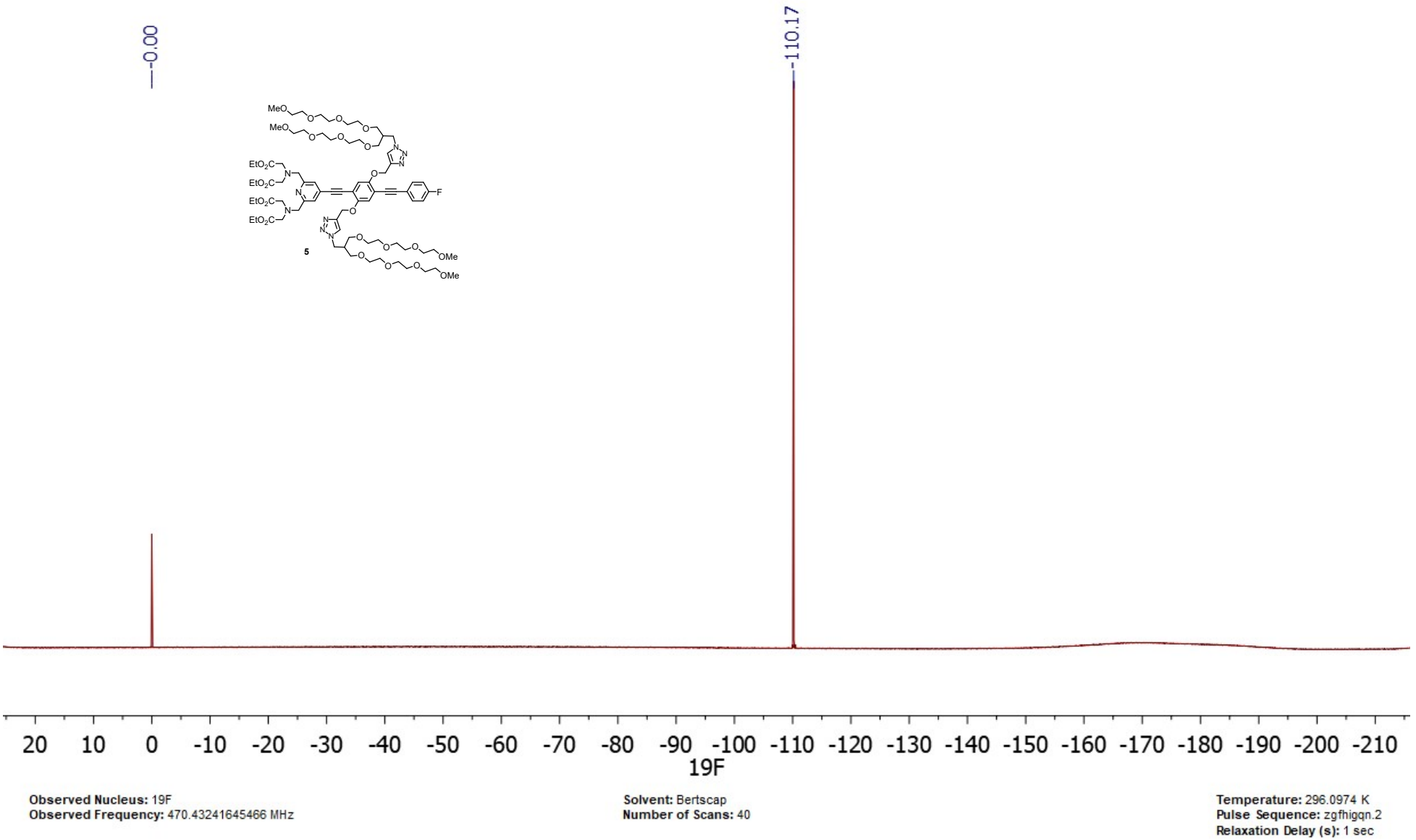


Figure S4.7. ^{19}F-NMR spectrum (471 MHz, $CDCl_3$) of PyMTA ester-EPEP-F **5** with CCl_3F as an external standard for the calibration.

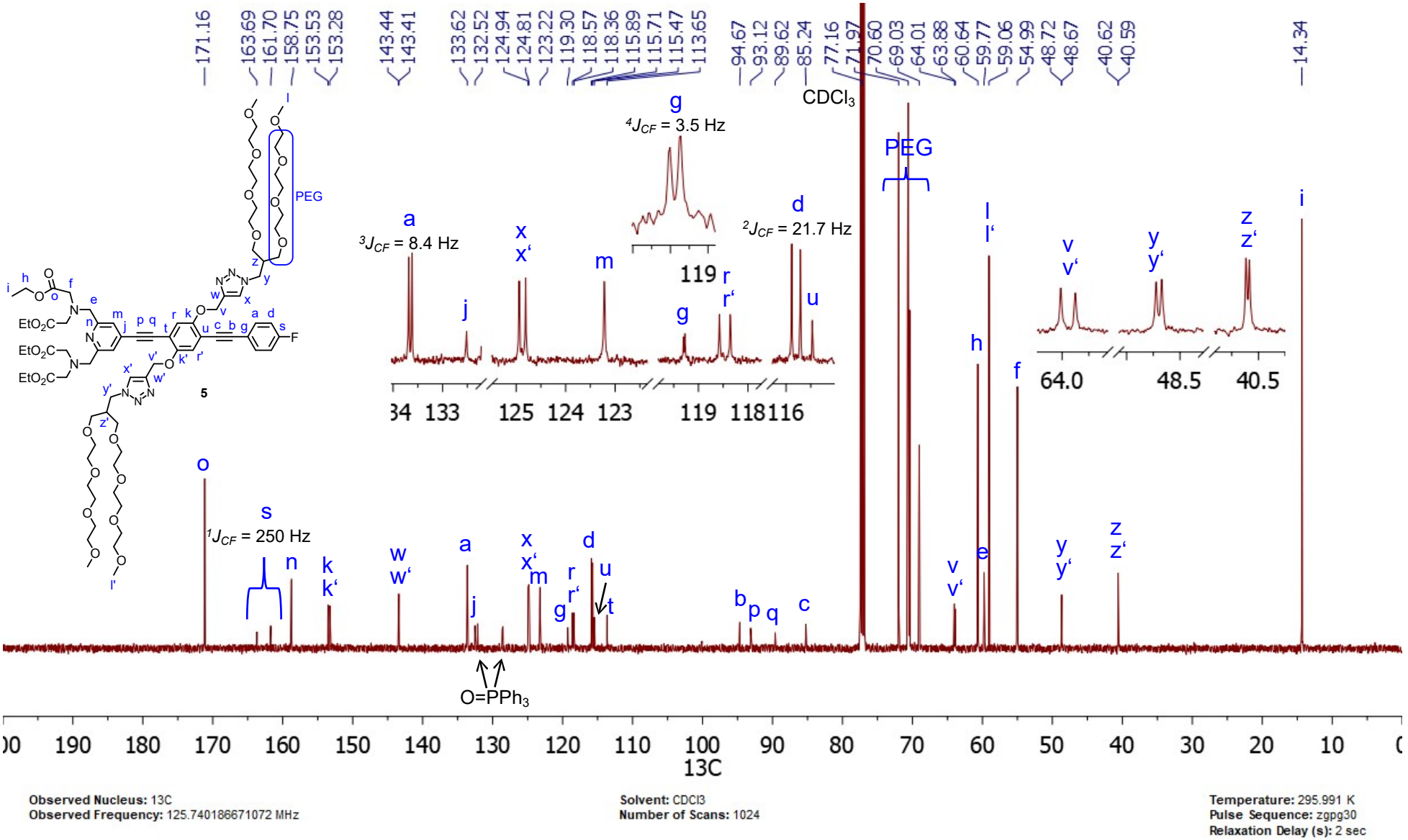


Figure S4.8. ^{13}C-NMR spectrum (126 MHz, $CDCl_3$) of PyMTA ester-EPEP-F **5**.

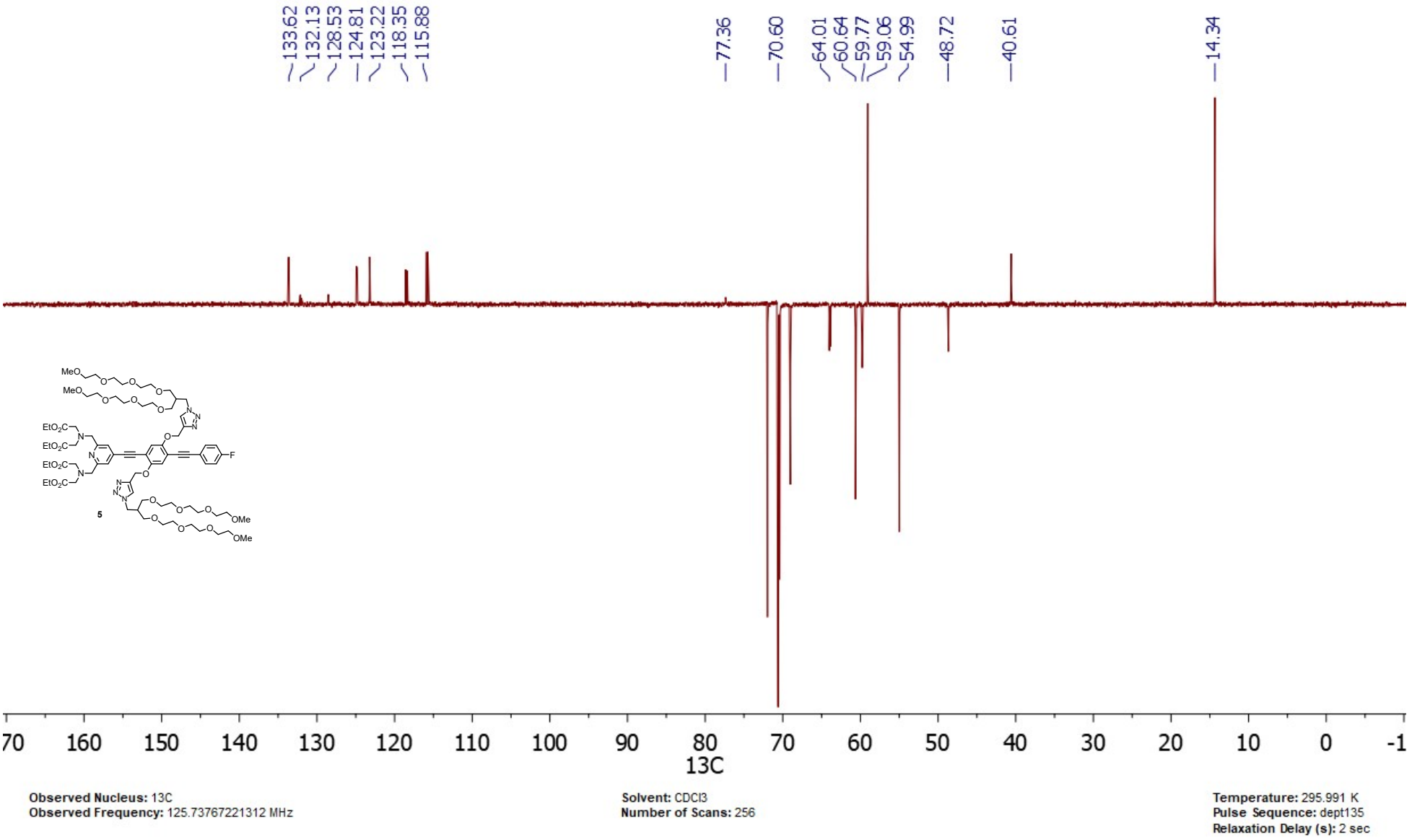


Figure S4.9. ^{13}C-DEPT 135 NMR spectrum (126 MHz, $CDCl_3$) of PyMTA ester-EPEP-F **5**.

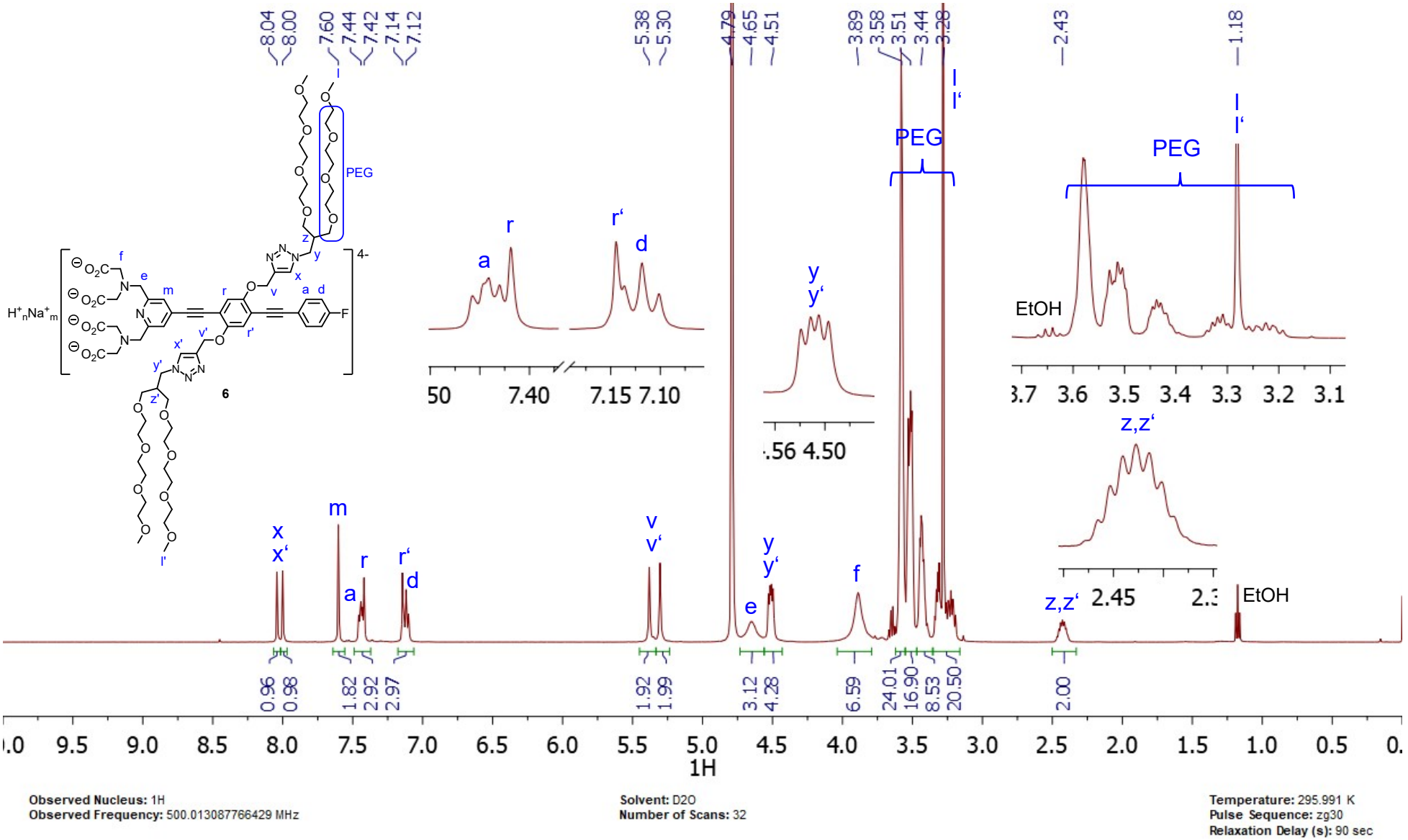


Figure S4.10. ^{1}H-NMR spectrum (500 MHz, D_2O) of H_nNa_m[PyMTA-EPEP-F] **6**.

SI 3. Tag structures used for rotamer simulations and comparison of tether flexibility

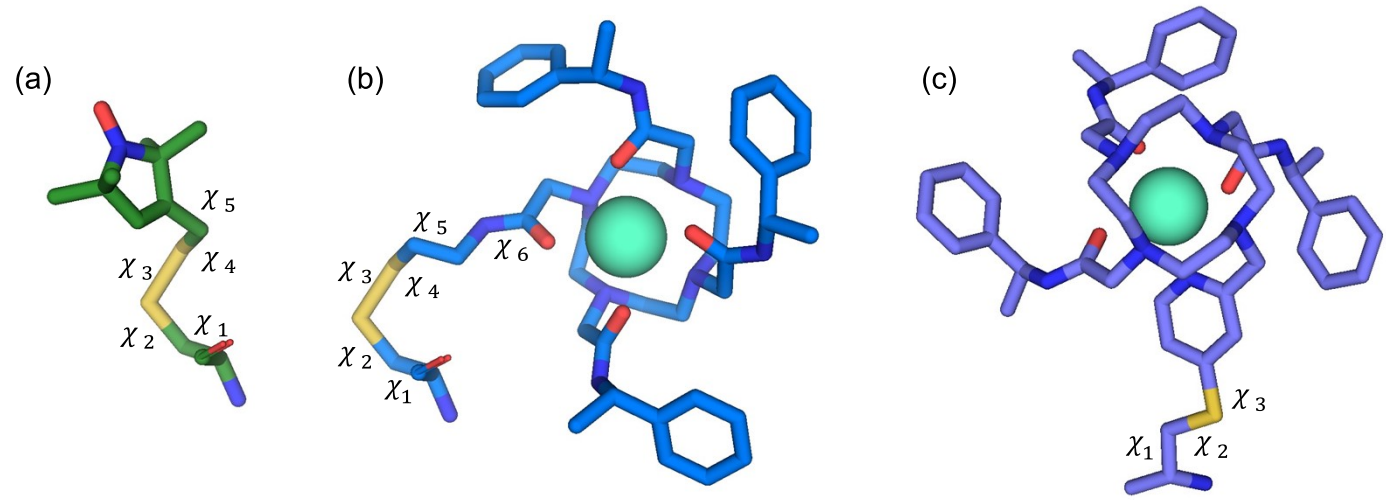


Fig. S5. Heavy atom representations of cysteine residues with different tags. **(a)** MTSL, used for the nitroxide-labelled systems in this work. **(b) Gd.C1** used for Gd^{3+}-labelled systems in this work. **(c)** An example of a short-tether tag, **Gd.C12**, forming a thioether bond with cysteine and a shorter tether, reported by Butler and co-workers.[7] The structures are annotated with the dihedral angles χ_n corresponding to the freely rotatable bonds in each tag. These χ_n angles were varied randomly to generate rotamer libraries of each tag at the labelling site on the surface of the protein using the program PyParaTools.[8] Conformations generating steric clashes with the protein were excluded, using a van der Waals radius (r_{vdW}) scaling factor of s = 0.9 for every heavy atom ($s \cdot r_{vdW}^2 < d_{atom}^2$), where d_{atom} is the distance between neighbouring atoms.

Although both the **C1** and **C12** tags are larger than the MTSL nitroxide tag (~7.5 Å from the Gd^{3+} ion to the furthest point of the lanthanide cage, vs ~2.5 Å across the MTSL pyrrol ring), the MTSL tag is comparable to the **Gd.C1** tag in terms of the flexibilities of both tags' tethers, with 5 and 6 rotatable bonds respectively, and will therefore introduce a similar degree of distance distribution, and therefore linewidth broadening, to the ENDOR spectrum. **Gd.C12** on the other hand has only three freely rotatable bonds.

SI 4. Relaxation data

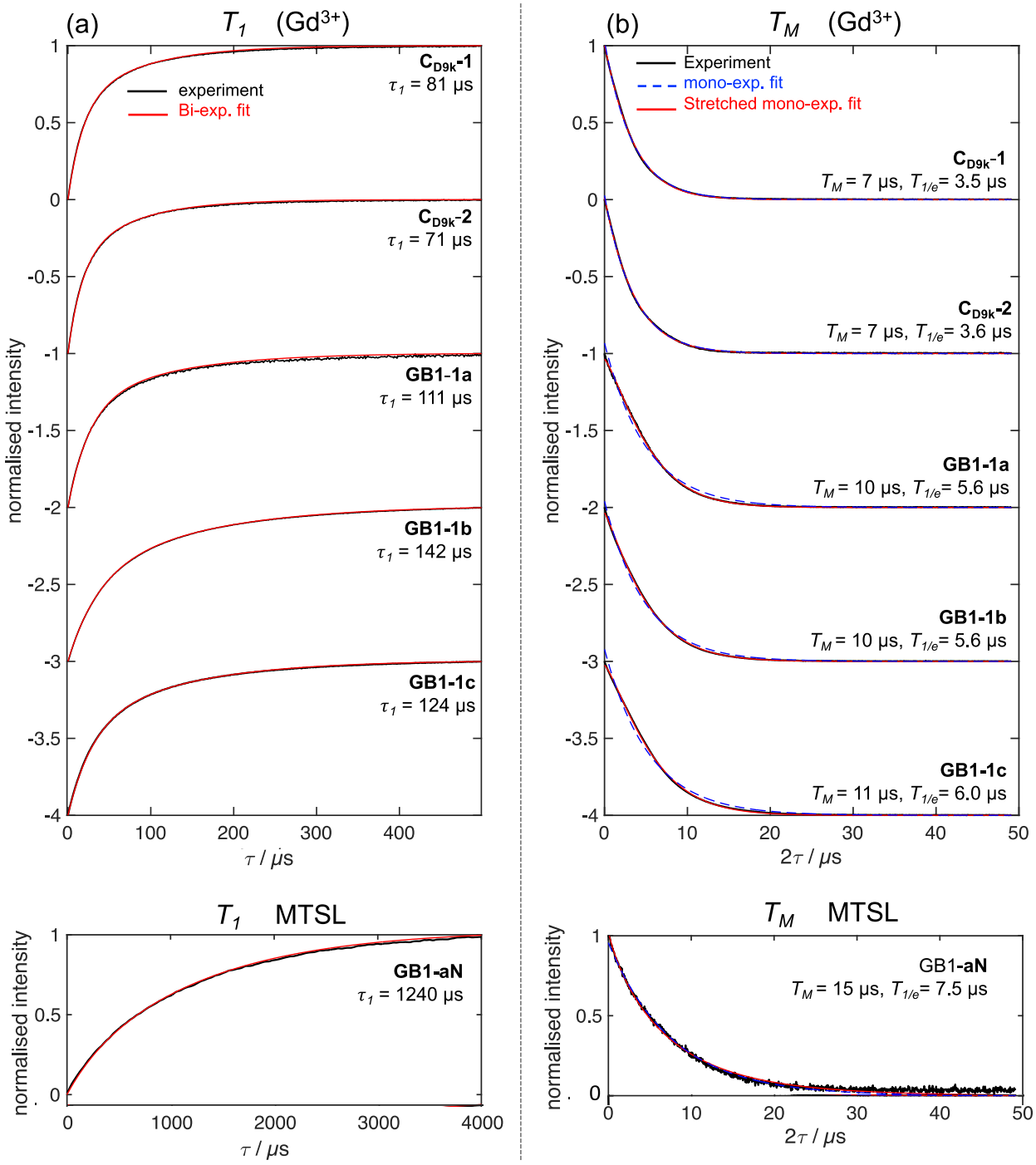


Fig. S6. EPR relaxation data of the Gd^{3+}-tagged $\mathbf{C_{D9k}}$ and **GB1** protein variants (top panels), and MTSL-tagged **GB1** protein variants (bottom panels), along with the associated fits. All fits were performed using a least-square fitting algorithm in MATLAB. The best-fit time constants are indicated next to each trace. **(a)** T_1 inversion–recovery measurements (black trace) were fitted with a biexponential decay function of the form $[Aexp(-t/\tau_1) + Bexp(-t/\tau_2)]$ (red trace). **(b)** Spin-echo decay measurements of the phase memory time T_M (black trace) with the corresponding mono-exponential fit $Aexp(-2t/T_M)$ (blue dashed trace), and stretched exponential function fit $exp(-2t/T_M^a)$ (red trace). The $T_{1/e}$ time denotes the time for the signal to reach 1/e of its initial intensity, and $T_M \approx 2T_{1/e}$.[8]

Table S1 Parameters used to fit the relaxation data of the Gd^{3+} and MTSL tagged $\mathbf{C_{D9k}}$ and **GB1** systems.

	T_1 - τ_1 (µs)	T_1 - τ_2 (µs)	T_M (µs) monoexp.	T_M (µs) / stretch	$T_{1/e}$ (µs)
C_{D9k}-1	80.5	18.0	6.7	7.0 / 1.06	3.46
C_{D9k}-2	70.5	16.7	7.0	7.2 / 1.03	3.55
GB1-1a	110.5	25.2	10.0	11.6 / 1.28	5.62
GB1-1b	142.7	35.6	9.7	10.8 / 1.18	5.28
GB1-1c	123.9	28.7	10.8	12.3 /1.27	6.05
GB1-aN*	1239.8	266.2	15.5	13.8 / 0.86	7.49

***** MTSL-labelled systems, see SI 7.

SI 5a. C_{D9k}-1: CF_3-Phe residue environment and orientation

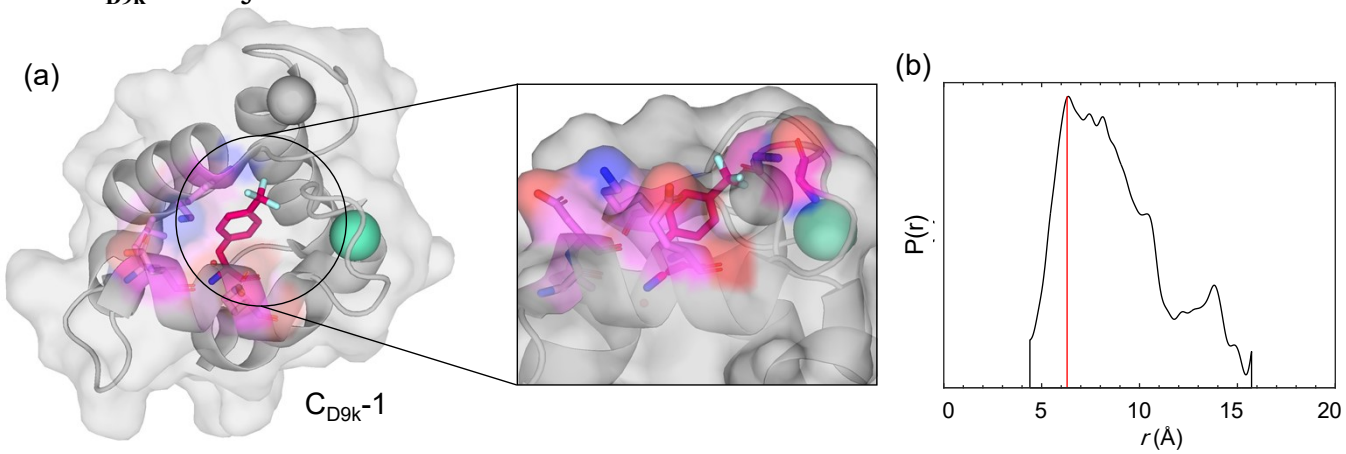


Fig. S7. Molecular modelling analysis of $\mathbf{C_{D9k}}$**-1**. **(a)** Surface and cartoon representation of $\mathbf{C_{D9k}}$**-1** (PDB ID: 4ICB[9]), with the CF_3-Phe tag at the site of F50, with the surrounding hydrophobic residues shown in magenta (coloured by element) along with a close-up of the tag region. Although the side chain of F50 is near the surface of the protein, it is hemmed in by the other protein residues, restricting rotation of the CF_3 group. **(b)** The theoretically possible Gd^{3+}–^{19}F distance distribution in $\mathbf{C_{D9k}}$**-1**(shaded grey) calculated using the PyParaTools program. This distance distribution does not take into account any hydrophobic interactions – only steric clashes. However, as discussed in the main text, our simulations fit mainly a short ~7 Å distance (and CF_3 tilt angle relative to the Gd^{3+}-C_{3F} dipolar axis) which aligns with only one of the peaks of the distribution (red line), suggesting the CF_3 group is practically rigid in conformation.

SI 5b. C_{D9k}-2 with a ~16 Å interspin distance: ^{19}F-ENDOR spectra and modelling

The structure of $\mathbf{C_{D9k}}$**-2** (PDB ID: 4ICB)[9] designed to have a long ~16 Å interspin Gd^{3+}–^{19}F distance, is shown in Fig. S8, alongside the rotamer library simulation of the Gd^{3+}–^{19}F distances calculated using PyParaTools (Fig. S8b), and the measured and simulated ^{19}F-ENDOR spectra (Fig, S8c). The experimental ENDOR spectrum is unstructured and unresolved, and can be reproduced by fitting with a Lorentzian lineshape function with a full linewidth at half maximum (fwhm) of 30 kHz as shown in Fig. S8b. Limited by the short electronic T_M time of $\mathbf{C_{D9k}}$, the smallest achievable linewidth is about ~30 kHz, which restricts the minimum resolvable dipolar coupling to 20 kHz, corresponding to an upper bound of 15 Å. Therefore in principle, any distance > 16 Å could account for the measured ENDOR spectrum.

Interestingly, the rotamer simulations (Fig. S8b, black trace) also indicate that the maximum physically possible distance between the Gd^{3+} ion and CF_3 group is 16 Å. Indeed, including the entire rotamer-predicted distance distribution in the Mims *EasySpin*[10] calculations, closely reproduced the simulated ENDOR spectrum. The simulated ^{19}F ENDOR spectrum shown in the red trace in Fig. S8c is the weighted sum of the ENDOR responses contributed by each distance in the distribution in Fig. S8b, convoluted with a Lorentzian linewidth of 35 kHz. Using a linewidth of 30 kHz or less, the simulation begins to resolve a splitting from the

shorter distance contributions, which are most likely oversampled in the rotamer simulation. The real distance distribution is thus most likely quite narrow with an inter-spin distance of about 16 Å, which is consistent with a mostly rigid CF_3 position.

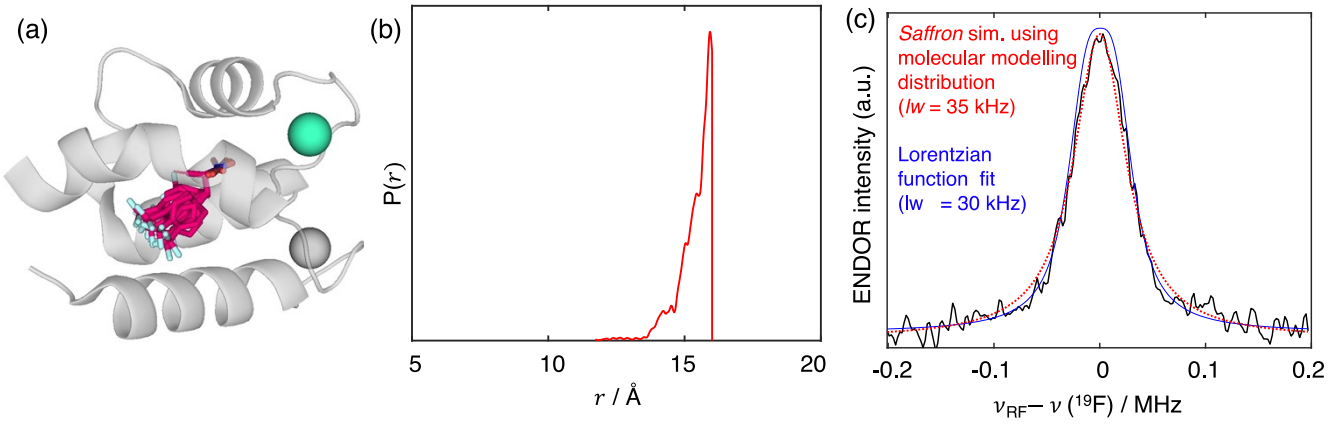


Fig. S8 Approaches to fitting the ^{19}F-ENDOR spectrum of $\mathbf{C_{D9k}}$**-2**. **(a)** Structure of $\mathbf{C_{D9k}}$**-2** showing the location of the CF_3-Phe residue in position 66 (magenta) and the Gd^{3+} ion bound to the N-terminal calcium binding site (cyan). **(b)** Distance distribution predicted from PyParaTools simulations of the CF_3-Phe rotameric states. **(c)** Normalised 94 GHz ^{19}F-ENDOR spectra of $\mathbf{C_{D9k}}$**-2** measured using τ = 2.0 µs. Experimental data are shown in black. The blue trace corresponds to the ENDOR EasySpin *saffron* simulation calculated from the modelled rotamer distance distribution in (b), convoluting the weighted sum with a Lorentzian linewidth of 35 kHz. The red trace is a Lorentzian line function with a linewidth of 30 kHz. The best fit parameters are listed in Table 1. The total recording time of the ^{19}F-ENDOR spectrum was 1.5 h. Further experimental details are provided in the Methods section.

SI 6a. Fitting ^{19}F-ENDOR spectra of GB1-a-c: accounting for flexibility of the Gd^{3+} tag tether

The contribution of the tether flexibility in the lanthanide-tag to the ^{19}F- ENDOR-detected distances and distance distributions could be examined by modelling the rotameric states of the **Gd.C1** tag and CF_3-Phe residue using the program PyParaTools, as described by Yagi et al.[8] In this approach, Gd^{3+}–^{19}F distance distributions are generated by random variation of the dihedral angles as indicated in Fig. S5. The protein coordinates (PDB ID: 1PGB[11]**)** were kept fixed, and conformers of the tags with steric clashes with the protein were excluded. To generate a library of between 500–2000 rotamers, the calculation explored about 10,000 random dihedral angle combinations. We compared this approach to the 3-parameter fit discussed in the main text (which included a mean distance, Gaussian distribution of width (standard deviation) σ and a Lorentzian linewidth *lw*). Simulations were performed using the EasySpin program *saffron* using τ= 2.0 µs with a grid sampling of 200 knots and assuming an infinite excitation bandwidth. The fit was performed using the simplex algorithm in MATLAB (*fminsearch*) to minimise the standard deviation between the simulation fit and symmetrised raw data.

Fig. S9b shows the calculated PyParaTools distributions (blue traces), compared to those fitted by the simplex minimization procedure (red traces). Encouragingly, the simplex Gaussian distributions are largely encompassed by the rotamer distance distribution, which validates that the simplex distributions are physically sensible. Importantly, we note that the PyParaTools simulation output represents the theoretically allowed (and physically feasible) conformational space, but it is not refined to account for preference of some conformations over others, e.g., if there are favourable hydrophobic interactions between the phenylethylamide arms of the C1 tag with the protein, which would retain the tag in one conformation over another. The true inter-tag distance distribution may thus be a subset of the rotamer-simulated distribution.

For **GB1-1b** we get almost identical agreement between the simplex and rotamer distance distribution models. For **GB1-1a,** the simplex fit, which best fits the ^{19}F-ENDOR data, is constrained mostly to the lower-distance shoulder of the rotamer model, at a local maximum of ~9.5 Å. We can see from Fig. S9c that when the entire distance distribution predicted by the rotamer simulations ($\bar{r}_{mean}$ = 13 Å, σ = 2.3 Å) is used to simulate the ^{19}F-ENDOR experimental data, it severely underestimates the splitting and cannot account for the real ENDOR lineshape. The overlap between the two distance distribution models for **GB1-1c** is much worse than for the other two complexes. The simplex fit includes longer distances than physically possible according to the rotamer simulation. As noted in the main text, however, there is significant uncertainty associated with the Gaussian fit to the **GB1-1c** data, as without the constraint of a resolved splitting, the linewidth and σ parameters are strongly correlated, compromising the determination of a distinct distribution width. This is reinforced by Fig. S9c, where the rotamer and simplex models produce comparable fits to the ^{19}F-ENDOR spectrum.

Although there are differences between the simplex fitting approach and rotamer simulation approach, molecular modelling simulations such as PyParaTools could be used as a starting point for analysing the distance distribution underlying the ^{19}F-ENDOR data without the need for reliable user-informed input, particularly for longer inter-tag distances. At shorter distances, where fitting a distance distribution presents less of an ill-posed problem, molecular modelling simulations can be used to check the validity of the Gaussian-based parameter fit.

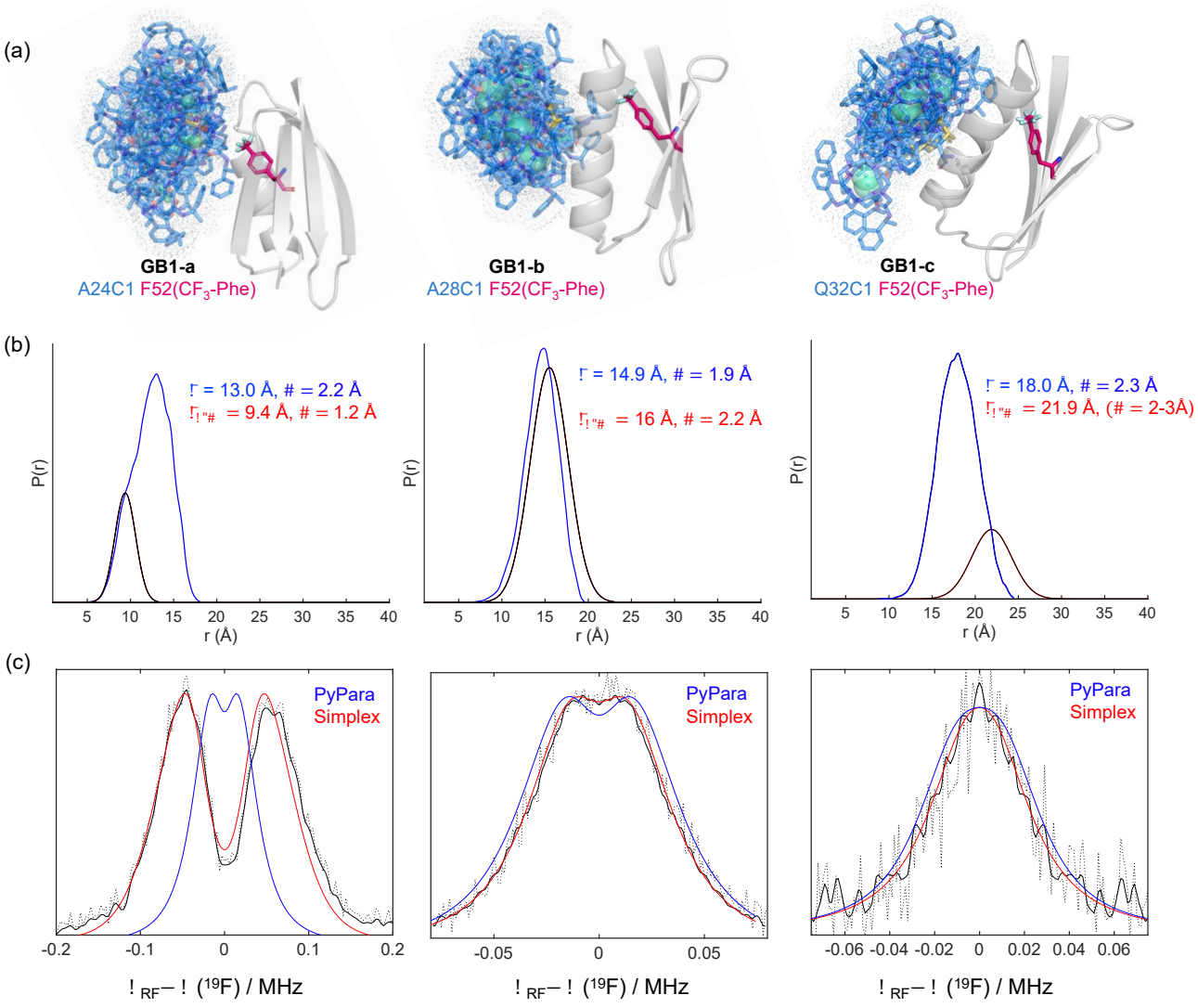


Fig. S9 Molecular modelling analysis and Mims ^{19}F-ENDOR spectra fitting of **GB1-a-c**. **(a)** Structures of the **GB1** rotamers (first 20 models) generated by PyParaTools. **(Left panel) GB1-a** (~10 Å predicted Gd^{3+}–C_{F3} distance). **(Middle panel) GB1-b** (~15 Å predicted Gd^{3+}–C_{F3} distance). **(Right panel) GB1-c** (~20 Å predicted distance). The positions of the **Gd.C1** tag (cyan) and CF_3-Phe residue (magenta) are shown and annotated in each figure. **(b)** Distance distributions predicted from the PyParaTools rotamer simulations (blue trace) for each protein using the procedure described above, compared to the Gaussian distribution (red shaded traces) fit the experimental data by the 3-parameter fitting procedure. **(c)** Experimental ENDOR spectra (also shown in Fig. 4 of the main text; smoothed, symmetrised data in black, unsmoothed data in grey dotted lines), overlayed with the Mims ENDOR (EasySpin, *saffron)*[10] simulations performed using the corresponding rotamer distance distribution (blue) or simplex distribution (red).

SI 6b. Uncertainty estimates

The simplest way of determining the distance from the ^{19}F-ENDOR spectra is simply by reading off the interspin dipolar coupling (T_{read}) as the peak-to-peak splitting of the ENDOR line. To arrive at the corresponding distance, r_{read}, we assume the dipolar approximation to the electron-nuclear hyperfine coupling : $\left(\frac{\mu_0}{4\pi}\right) g_e g_N \mu_e \mu_N \hbar \frac{1}{r^3}$.

Following the approach taken by Bennati and co-workers,[12,13] we estimate the uncertainties in T_{read} and r_{read} to be proportional to three times the resolution of the frequency axis, as a percentage of the resolved splitting T_{read}. This way of approximating the uncertainty is consistent with the range of resolved T_{read} values obtained from different batches of measurements on the same sample (e.g., on different days, under different measurement conditions and at slightly different operating frequencies). An example is shown below for the ^{19}F-ENDOR spectrum of **GB1-b**, where the splitting for the ~15−16 Å distance in **GB1-b** approaches the linewidth limit of the splitting resolution. In this case, we observed about ± 3 kHz variation in the peak-to-peak splitting, which corresponds to an about uncertainty of about ±1 Å in the distance domain. The variation in the measured ^{19}F hyperfine splitting in the ENDOR spectrum is smaller in the case of **GB1-a** where a similar variation in T_{read} corresponds to only about ± 0.3 Å uncertainty. The variations in the ENDOR splitting frequency could be attributed to several factors, including that at longer distances the magnitude of the splitting begins to match the linewidth resolution, the noise level being comparable to the resolution of the ENDOR splitting (how distinctly separated the peaks of the ENDOR pattern are), and that the uncertainty in a small splitting is more susceptible to slight frequency drifts and frequency dependent artefacts. We observed better resolved splittings at some values of the operating MW frequency (94±0.015 GHz) than others, so this parameter also had to be tested and optimized, particularly for the longer-distance proteins.

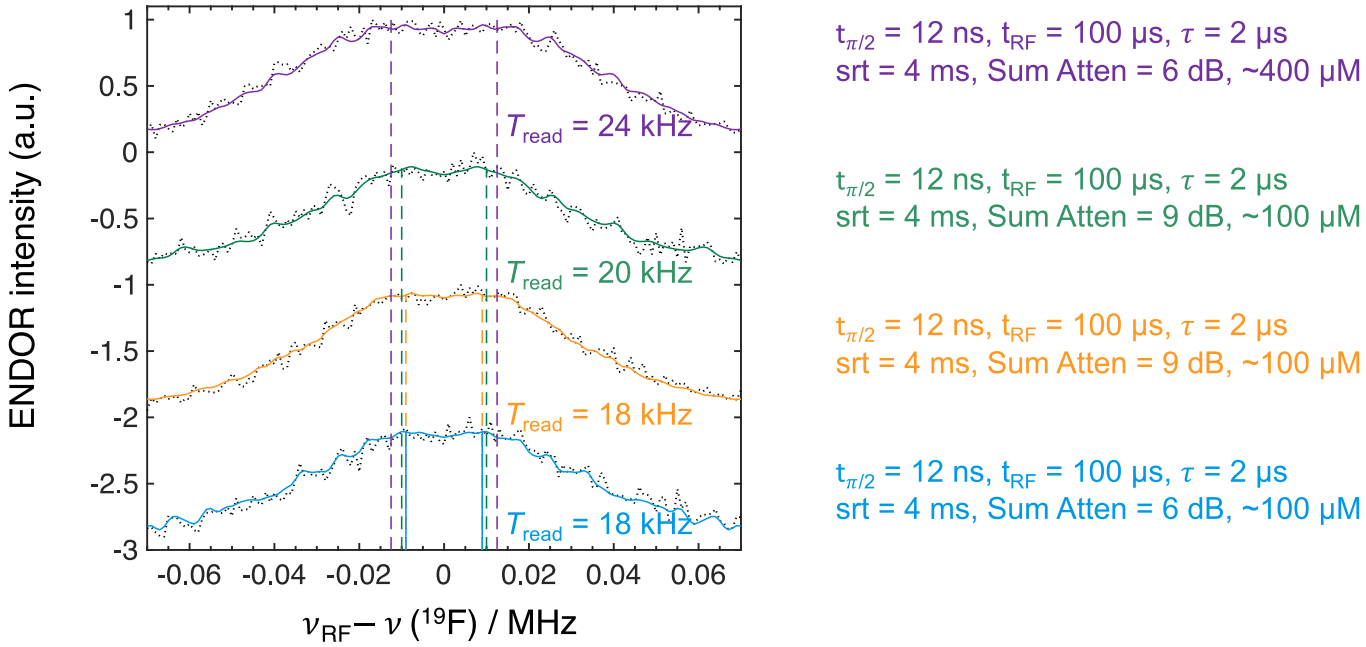


Fig. S10 ^{19}F-ENDOR spectra of **GB1-1b**. The annotations with T_{read} values illustrate the experimentally observed variability and uncertainty in the splitting value. The pulse parameters used for each measurement are given next to each spectrum: $t_{\pi/2}$ is the detection pulse length, t_{RF} is the RF pulse length, τ is the interpulse delay of the detection sequence, and Sum Atten stands for the sum attenuation applied to the RF pulse power. The concentration of the sample in each measurement is annotated at the end. The raw data are shown using grey dotted lines, while the symmetrized and 3-point smoothed data are shown by coloured solid lines.

We used the same approach to estimating the uncertainties for the fits obtained from the 3-parameter simplex algorithm. The actual uncertainty is most likely larger than estimated here, as the parameters can be strongly correlated and, as noted by Kehl et al.,[13] a more thorough error analysis would require calculating the covariances of the fit parameters. Particularly in the data sets without a resolved splitting, the distance distribution and Lorentzian linewidth are both effectively line-broadening parameters that cannot easily be distinguished in this case. To arrive at reasonable values, we imposed some constraints. For example, we set the linewidth of the **GB1-c** which has the longest Gd^{3+}–F_3 distance, to be equal to or less than that of **GB1-a** with the shortest distance, as the greater distance of the ^{19}F label from the paramagnetic centre in **GB1-c** must necessarily imply lesser paramagnetic enhancement of the nuclear T_2 relaxation. Other assumptions, however, such as symmetrical, Gaussian-shaped distance distributions, which may be reasonable for modelling long distances (15–20 Å) or very narrow distributions, may not always be justifiable in other circumstances. For example, a normal distribution fit may be too simplistic for broad distance distributions or shorter distances, where the $1/r^6$ dependence of the F_{ENDOR} intensity changes most dramatically with small changes in distances (eqn. 1 in main text). Nevertheless, a preliminary evaluation of ^{19}F-ENDOR data may start from the assumption of a normal distribution to estimate approximate interspin distances involving tethered tags and assess the range of Gd^{3+}–^{19}F distances.

SI 7. Suppressing orientation selection using CF_3 tags

The choice of a CF_3-Phe residue was advantageous because, as previously shown, the ENDOR sensitivity increases linearly with the number of probe nuclei,[13] and the amino acid was readily available and tested in ^{19}F-NMR experiments in our laboratories.[3] An important advantage associated with the CF_3 group is that it can allow more rapid accumulation of the ENDOR Pake pattern in orientation-selective systems such as nitroxide tags. We observed this effect for the nitroxide-labelled analogue **GB1-aN**, with an MTSL tag in position 24 and CF_3-Phe in position 52.

The dipolar vector connecting a single fluorine atom and the radical centre only ever selects a single component of the nitroxide *g*-frame, so that recovering both the $A_\perp$ and $A_\parallel$ edges of the ENDOR lineshape requires sampling multiple points in the nitroxide EPR spectrum. This represents a significant penalty in accumulation time. With a CF_3 group on the other hand, there are three non-orthogonal electron–fluorine dipolar vectors, where at least two vectors contain projections onto both the perpendicular and parallel axes of the *g*-frame (Fig. S11a). As a result, both the edges and central peaks of the ENDOR Pake pattern can be extracted at a single field position.

Fig. S11 below shows the experimental demonstration of this effect by comparing the Mims ENDOR lowest and highest field edges in the nitroxide spectrum, corresponding to the g_{xx} and g_{zz} components of the **g**-tensor orientations, respectively. Unlike the systems studied by Bennati and co-workers,[12] the hyperfine tensor **A,** and **g**-tensor frames are not colinear in our **GB1** systems. We therefore could not clearly distinguish the $A_\perp$ and $A_\parallel$ (in the main text we refer to T, the dipolar component of the **A**-tensor) components of the ENDOR spectra of our systems. Nevertheless, measuring at either field edge in the CF_3-labelled system resulted in almost the same ENDOR spectrum.

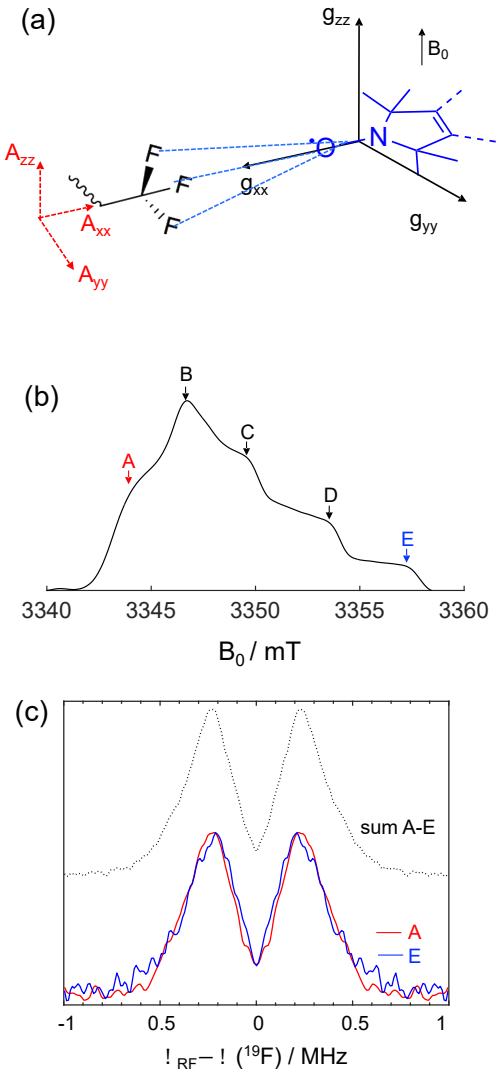


Fig. S11 Using a CF_3-Phe tag to minimize orientation selection in Mims ^{19}F -ENDOR. **(a)** Schematic representation of the orientation of the three dipolar vectors of a CF_3 group with respect to the nitroxide *g*-frame and external magnetic field in a colinear system. **(b)** Electron spin-echo spectrum of the nitroxide-labelled **GB1-aN** showing the five main turning points of the nitroxide EPR lineshape. **(c)** Mims ^{19}F - ENDOR spectrum (t_{RF} = 30 µs, τ = 0.5 µs, π/2 =2 ns) of the CF_3-Phe labelled **GB1-aN** system. The figure compares the ENDOR spectrum measured at the lowest field position (A, red trace) with the spectrum measured at the high field edge (E, blue trace), which correspond to the g_{xx} and g_{zz} components of the nitroxide **g**-tensor respectively. The sum the ENDOR spectra measured at all five turning points A–E is shown in the black dotted trace above.

References

1 N. Kronqvist, M. Sarr, A. Lindqvist, K. Nordling, M. Otikovs, L. Venturi, B. Pioselli, P. Purhonen, M. Landreh, H. Biverstål, Z. Toleikis, L. Sjöberg, C. V Robinson, N. Pelizzi, H. Jörnvall, H. Hebert, K. Jaudzems, T. Curstedt, A. Rising and J. Johansson, *Nat. Commun.*, 2017, **8**, 15504, DOI: 10.1038/ncomms15504.

2 T. Mukai, H. Hoshi, K. Ohtake, M. Takahashi, A. Yamaguchi, A. Hayashi, S. Yokoyama and K. Sakamoto, *Sci. Rep.*, 2015, **5**, 9699, DOI: 10.1038/srep09699.

3 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.

4 M. Qi, M. Hülsmann and A. Godt, *J. Org. Chem.*, 2016, **81**, 2549–2571, DOI: 10.1021/acs.joc.6b00125.

5 M. Kaushik, M. Qi, A. Godt and B. Corzilius, *Angew. Chem. Int. Ed. Engl.*, 2017, **56**, 4295–4299, DOI: 10.1002/anie.201612388.

6 M. Qi, M. Hülsmann and A. Godt, *Synthesis (Stuttg).*, 2016, **48**, 3773–3784, DOI: 10.1055/s-0035-1561660.

7 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.

8 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.

9 L. A. Svensson, E. Thulin and S. Forsén, *J. Mol. Biol.*, 1992, **223**, 601–606, DOI: 10.1016/0022-2836(92)90976-Q.

10 S. Stoll and A. Schweiger, *J. Magn. Reson.*, 2006, **178**, 42–55, DOI: 10.1016/j.jmr.2005.08.013.

11 T. Gallagher, P. Alexander, P. Bryan and G. L. Gilliland, *Biochemistry*, 1994, **33**, 4721–4729, DOI: 10.1021/bi00181a032.

12 A. Meyer, S. Dechert, S. Dey, C. Höbartner and M. Bennati, *Angew. Chemie Int. Ed.*, 2020, **59**, 373–379, DOI: 10.1002/anie.201908584.

13 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.