

Biosynthetic lanthanide-luminescent mini-proteins using genetic code expansion

Edan Habel,^a Haocheng Qianzhu,^a Elwy H. Abdelkader,^b Nathan Paul,^a Gottfried Otting,^b Thomas Huber^{a*}

[a] Research School of Chemistry, Australian National University
Canberra, ACT 2601(Australia)

E-mail:t.huber@anu.edu.au

[b] ARC Centre of Excellence for Innovations in Peptide & Protein Science, Research School of Chemistry, Australian National University

Table of contents:

Biosynthetic lanthanide-luminescent mini-proteins using genetic code expansion

Table of contents:

Supplementary methods

- a) General considerations
- b) Materials
- c) Expression and purification of LBT
- d) Intact protein mass spectrometry
- e) Luminescence measurements
- f) Integrated lifetime measurements
- g) Sequence design of EF2-derived mini-proteins
- h) q-value determination
- i) Lanthanide titration experiments
- j) Expression of isotope labeled RF2
- k) NMR measurements of ¹³C/¹⁵N-labeled RF2
- l) Spectral overlap calculations

Supplementary figures

Figure S1. Intact protein mass spectrometry of LBT.

Figure S2. Intact protein mass spectrometry of LBT 4AW.

Figure S3. Intact protein mass spectrometry of LBT 6AW.

Figure S4. Intact protein mass spectrometry of LBT 4CNW.

Figure S5. Intact protein mass spectrometry of LBT mCNP.

Figure S6. Intact protein mass spectrometry of LBT mCNF.

Figure S7. Intact protein mass spectrometry of the EF2 variant RF1.

Figure S8. Intact protein mass spectrometry of the EF2 variant RF2.

Figure S9. Intact protein mass spectrometry of the EF2 variant RF3.

Figure S10. Intact protein mass spectrometry of the RF2 42 6AW.

Figure S11. Intact protein mass spectrometry of the RF2 50 6AW.

Figure S12. Excitation scans of the LBT complexes with Tb³⁺ and Eu³⁺.

Figure S13. Absorbance spectra of the fluorescent ncAAs used in this study.

Figure S14. Binding affinity determination by titration of LBT variants with Tb³⁺.

Figure S15. Binding affinity determination by titration of LBT variants with Eu³⁺.

Figure S16. Spectral overlap of the Trp emission (free amino acid in Buffer E) with the absorption of Eu³⁺ and Tb³⁺ (LnCl₃ in MOPS buffer).

Figure S17. Spectral overlap of mCNP emission (free amino acid in Buffer E) with the absorption of Eu³⁺ and Tb³⁺ (LnCl₃ in MOPS buffer).

Figure S18. Spectral overlap of mCNF emission (free amino acid in Buffer E) with the absorption of Eu³⁺ and Tb³⁺ (LnCl₃ in MOPS buffer).

Figure S19. Spectral overlap of 4AW emission (free amino acid in Buffer E) with the absorption of Eu³⁺ and Tb³⁺ (LnCl₃ in MOPS buffer).

Figure S20. Spectral overlap of 4CNW emission (free amino acid in Buffer E) with the absorption of Eu³⁺ and Tb³⁺ (LnCl₃ in MOPS buffer).

Figure S21. Spectral overlap of 6AW emission (free amino acid in Buffer E) with the absorption of Eu³⁺ and Tb³⁺ (LnCl₃ in MOPS buffer).

Figure S22. HNCACB data of ¹⁵N/¹³C-labeled RF2.

Figure S23. Overlay of ¹³C-decoupled 3D NOESY-¹⁵N-HSQC (blue) and TOCSY-¹⁵N-HSQC (red) spectra of ¹⁵N/¹³C-labeled RF2.

Figure S24. Comparison of the secondary structure of the AlphaFold3 model and the prediction made from the assigned chemical shifts.

Figure S25. Comparison of excitation spectra for RF2 6AW/Eu³⁺, LBT 6AW/Eu³⁺ and LBT/Tb³⁺.

Figure S26. Excitation scans of the RF2 complexes with Tb³⁺ and Eu³⁺.

Figure S27. Fluorescence emission spectra of RF2 6AW with Sm³⁺ and Dy³⁺.

Figure S28. Binding affinity determination by titration of RF2 6AW complexes with Tb³⁺, Eu³⁺, and Dy³⁺.

Figure S30. Luminescence lifetime measurement of RF2 6AW with Eu³⁺.

Figure S31. Luminescence lifetime measurement of RF2 6AW with Sm³⁺.

Figure S32. Luminescence lifetime measurement of RF2 6AW with Dy³⁺.

Figure S33. D₂O titrations of RF2/Ln³⁺ complexes.

Figure S34. Competition experiment with M²⁺/M³⁺ ions.

Figure S36. Competition experiment with excess M²⁺ ions relative to Tb³⁺.

Figure S37. Quantification of cell lysis following incubation with 20 mM Eu³⁺ in 150 mM NaCl.

Figure S38. *In vivo* sensitization of lanthanide luminescence.

Figure S39. *In vivo* sensitization of europium luminescence in unwashed cells.

Supplementary tables

Table S1. Extinction coefficients of the amino acids used in this study

Table S2. Variables used in integrated lifetime calculations^a

Table S3. Amino acid sequences of the proteins used in this study

Table S4. Yields of different proteins co-expressed with G1(ncAA)RS in LB media with 1 mM ncAA^a

Table S5. EF2-RF variants with Tb³⁺

Table S6. Determination of *q*-values from D₂O titrations^a

References

Supplementary methods

a) General considerations

No unexpected or unusual safety hazards were encountered. All buffers used for lanthanide luminescence measurements were treated with Chelex 100 resin. Lanthanide ion stocks were made from $\text{LnCl}_3(\text{H}_2\text{O})_6$ in 10 mL volumetric flasks using water treated with Chelex 100 resin (BioRad), and a small amount of concentrated HCl to prevent hydroxide formation. All other metal ion stocks were prepared similarly using chloride salts. Buffers used for experiments with lanthanides were also treated with Chelex 100 resin. Concentrations of rare earth element stocks were validated using complexometric titration with EDTA and Xylenol Orange (Sigma). For experiments with protein- Ln^{3+} complexes, we used buffer conditions that have no or minimal affinity to lanthanide ions.¹

b) Materials

The non-canonical amino acids 6-aza-tryptophan (6AW, Cat. No.: A861240) and 4-aza-tryptophan (4AW, Cat. No.: A264672), were purchased from Ambeed (USA). 4-cyano-tryptophan (4CNW, Cat. ID: V152071) was purchased from Advanced ChemBlocks Inc. (USA). 3-(2-cyano-4-pyridyl)alanine (mCNP, Cat. ID: CSMB00010457344) was purchased from Chempace LLC (Kyiv, Ukraine). 3-cyano-phenylalanine (mCNF, Cat. ID: LC54307) was purchased from AK Scientific Inc., USA.

Buffer A: 50 mM Tris HCl pH 7.0, 300 mM NaCl

Buffer B: 50 mM Tris HCl pH 7.0, 20 mM imidazole, 300 mM NaCl

Buffer C: 50 mM Tris HCl pH 7.0, 500 mM imidazole, 300 mM NaCl

Buffer D: 10 mM Tris HCl pH 7.0, 50 mM KCl

Buffer E: 20 mM MOPS pH 7.0, 100 mM KCl.

LB media: 10 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, 5 g L⁻¹ NaCl

Autoinduction media: 20 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, 5 g L⁻¹ NaCl, 6 g L⁻¹ Na₂HPO₄, 3 g L⁻¹ KH₂PO₄, 0.5% glycerol, 0.05% glucose, 0.2% lactose

c) Expression and purification of LBT

The gene for the fusion protein was ordered in a low-copy number pCDF plasmid under the control of a T7 promoter, either as the wild-type sequence (LBT-Trp) or with the tryptophan codon as an amber codon (LBT-Amb). The genes for the EF2 variants RF1, RF2, and RF3, as well as the RF2 amber mutants, were also cloned into the pCDF vector under control of the T7 promoter. High-copy number pRSF plasmids were used to encode the genes of the G1 tRNA synthetase mutants under the T7 promoter and the paired G1-tRNA under the *proK* promoter. This two-plasmid system has previously been established as an optimized system² to genetically encode ncAAs in *E. coli* B-95.ΔΔΔ*fabR* cells.³ For LBT-Amb, the two plasmids were transformed into cells concurrently. For LBT-Trp, the pRSF plasmid was omitted. Transformed cells were grown in 15 mL LB media overnight supplemented with spectinomycin (25 μg mL⁻¹) and for cells with the pRSF plasmid, kanamycin (25 μg mL⁻¹) to produce a start culture. The start culture was diluted at a 1:100 ratio in LB media supplemented with antibiotic and 1 mM ncAA and grown at 37 °C with shaking, and protein expression was induced with 1 mM IPTG at an OD₆₀₀ of 0.6, at which time the culture was transferred to a shaker at room temperature for expression. For EF2 proteins not labelled with ncAA (EF2 RF1-Trp, EF2 RF2-Trp, EF2 RF3-Trp), the start culture was diluted instead at a 1:100 ratio in autoinduction media supplemented with antibiotics and grown for

3 hours at 37° C before being transferred to a room temperature shaker. After 16–18 hours of expression, the fusion protein was isolated from the cells using immobilized metal-ion affinity chromatography (IMAC). This typically yielded 20–30 mg L⁻¹ of fusion protein. Incorporation of the ncAA was confirmed by mass spectrometry. The fusion protein was then concentrated to 1 mL in TEV buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl) and cleaved at 4 °C overnight after the addition of TEV protease at a 50:1 ratio of protein:TEV and 10 mM DTT. The resultant LBT-Amb fusion proteins were then isolated via size-exclusion chromatography (SEC) in volatile buffer (200 mM ammonium bicarbonate pH 8.5). Fractions containing the peptide were then isolated and lyophilized before luminescence analysis. For the EF2 proteins, IMAC elution fractions were dialyzed into TEV buffer and cleaved at RT overnight after the addition of TEV protease at a 1:1 ratio of protein:TEV and 2 mM β-mercaptoethanol. The EF2 proteins were then isolated by SEC in buffer D.

d) 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 ZORBAX 300SB-C3, 3.5 μm, 4.6 x 50 mm HPLC column (Agilent Technologies, USA). Approximately 50 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% solvent B at the start to 80% after 12 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).

e) Luminescence measurements

Luminescence measurements were performed on a Cary Eclipse fluorescence spectrophotometer (Agilent, USA) equipped with a phosphorescence module. Precise instrument settings for fluorescence measurements are indicated with the associated spectra. Time-resolved emission and excitation spectra were recorded with a 100 μs lag time, followed by a 1 ms integration time.

f) Integrated lifetime measurements

Luminescence lifetimes were measured on a Cary Eclipse fluorescence spectrophotometer (Agilent, USA) equipped with a lifetime measurement module. Each measurement employed a 100 μs delay followed by a 50 μs gate time, with both excitation and emission slit widths set to 20 nm. A 360–1100 nm filter was used to avoid signal from harmonic doubling. A total of 25 excitation flashes was used for a 15 ms decay window, and each dataset reflects the average of 50 repeated measurement cycles. Samples had a total volume of 600 μL in a 1 cm quartz cuvette. Samples were excited at their respective excitation maxima with emission recorded at 545 nm for Tb³⁺ complexes and 615 nm for Eu³⁺ complexes with a 20 nm slit width. Each sample was excited at the protein's excitation maximum with a 20 nm slit width and a recording of 500 nM Ln³⁺ with no protein was subtracted. The instrument gain was set to 900. To account for variability in binding affinity, protein was iteratively added to the cuvette at small volumes (2 μL) until the initial intensity value did not change after addition of more protein. The lifetime measurement was then adjusted by subtracting the blank measurement, and then fit to the following exponential equation using GraphPad Prism, with the constraint on the fit that $K > 0$.

$$Y = Plateau + (Y_0 - Plateau) \cdot e^{-KX} \quad (1)$$

The values of Y_0 , Plateau, and K were then tabulated with a standard error of the fit. Residuals were inspected to assess the quality of fit. The integrated emission intensity was then calculated as:

$$I = \frac{Y_0 - \text{Plateau}}{K} \quad (2)$$

The standard error of the best-fit parameters was then propagated as follows:

$$\delta I = \sqrt{\left(\frac{1}{K}\right)^2 \cdot ((\delta Y_0)^2 + (\delta \text{Plateau})^2) + \left(\frac{Y_0 - \text{Plateau}}{K^2}\right)^2 \cdot (\delta K)^2} \quad (3)$$

Except for three very weakly luminescent samples, (RF2 6AW/Dy³⁺, RF2/Eu³⁺, LBT 4CNW/Eu³⁺) propagated error did not exceed 1% of the integrated value. Exact values are tabulated in Table S2.

g) Sequence design of EF2-derived mini-proteins

Using PyMOL,⁴ the coordinates of EF2 (residues 59–70) were extracted from the first frame of the lanmodulin NMR ensemble (PDB 6MI5).⁵ Thr65 was mutated to tryptophan and the isolated segment was reindexed to begin at residue 1. The structure was energy minimized using the PyRosetta MinMover and the *ref2015* score function⁶⁻⁷ The minimized coordinates were uploaded to the RFdiffusion Google Colab (v1.1.1).⁸ Two sets of RFdiffusion backbones and sequences were generated using the built-in ProteinMPNN module. RF1 and RF2 were taken from set A with contigs set to “35:A1-15:24,” and RF3 was taken from set B with contigs set to “20:A1-15:20.” For both sets, iterations=50, num_designs=32, and symmetry=“none” ProteinMPNN parameters were num_sequences=8, num_recycles=3, and mpnn_sampling_temp=0.1, with cysteine excluded from suggestions. Set A and B were downloaded and combined for processing. The final set included 64 sets of backbone coordinates, each with 8 associated sequences.

For each backbone, PyRosetta was used to mutate the coordinates to match each of the eight ProteinMPNN sequences. A Tb³⁺ ion was placed 2.5 Å from the tryptophan backbone carbonyl along the C-O bond vector. Each model was minimized with *ref2015* and the final score was recorded as “rosettascore.” The C_α alignment RMSD between lanmodulin EF2 residues 59–70 and the corresponding segment in each design was computed and recorded as “bb_RMSD.” Distances between side-chain oxygen atoms and the antenna tryptophan carbonyl were measured, the count of O-Tb³⁺ distances below 3.0 Å was recorded as “nO,” and an RMSD to a 2.5 Å target for these O-Tb³⁺ distances was calculated and recorded as “O_RMSD.” The 2.5 Å target approximates an Ln³⁺-O bond length in protein crystal structures containing lanthanides. The combination of rosettascore, bb_RMSD, nO, and O_RMSD was used to select one sequence per backbone. The resulting models were inspected in PyMOL and three designs were selected to represent different tertiary structures. For expression, the first residue was changed to glycine to introduce an N-terminal TEV site.

h) q-value determination

The q-value was determined using a version of the Horrock’s equation, originally:

$$q = A((k_{H_2O} - k_{D_2O}) - B) - C \quad (4)$$

where A is a proportionality constant for the quenching of the lanthanide ion by O-H oscillators, k is the rate constant for the radiative decay of the lanthanide emission, parameter B is the corrective factor, accounting for other deactivating vibrations surrounding the ion, and parameter C is the outer-sphere contribution of closely diffusing solvent molecules.⁹

As the outer-sphere contributions are expected to be equivalent between the complex in H₂O and D₂O, the difference between k_{H_2O} and k_{D_2O} is attributed primarily to the inner-sphere solvent. C was therefore omitted for all ions giving the equation:

$$q = A((k_{H_2O} - k_{D_2O}) - B) \quad (5)$$

For samarium, this equation is:¹⁰

$$q_{Sm} = 0.026(1/\tau_{H_2O} - 1/\tau_{D_2O}) \quad (6)$$

For europium, this equation is:¹¹

$$q_{Eu} = 1.2((1/\tau_{H_2O} - 1/\tau_{D_2O}) - 0.25) \quad (7)$$

For terbium, this equation is:¹¹

$$q_{Tb} = 5((1/\tau_{H_2O} - 1/\tau_{D_2O}) - 0.06) \quad (8)$$

For dysprosium, this equation is:¹⁰

$$q_{Dy} = 0.024(1/\tau_{H_2O} - 1/\tau_{D_2O}) \quad (9)$$

The lifetime in D₂O was determined by plotting the lifetime of the Ln³⁺-complex at mole fractions of D₂O (0.2, 0.4, 0.6 and 0.8) followed by a linear fit to this data. The τ_{D_2O} value was then determined by extrapolating to a D₂O mole fraction of 1.0. These parameters are empirically determined based on the assumption that O-D oscillators effectively do not contribute to non-radiative deactivation, and that all other deactivation paths are the same between D₂O and H₂O. The estimated uncertainties with using this technique are listed with the calculations of the q-values below.

i) Lanthanide titration experiments

Lanthanide binding affinity titrations were performed on a TECAN Infinite 200 Pro M Plex plate reader (Tecan, Switzerland) using black 384-well flat bottom plates (Corning, USA). Protein samples were prepared in Chelex-treated buffer E. Each well was a series of 1:2 serial dilutions of the corresponding lanthanide chloride stock, whose concentration was determined colorimetrically as described above. For each well, a background blank containing the same concentration of lanthanide ion was used for blank subtraction. Time-resolved luminescence measurements were performed for Eu³⁺ and Tb³⁺ with a 100 μ s delay followed by a 1 ms integration window. For Dy³⁺, time-resolved settings were not used, instead collecting the emission directly under steady-state conditions. Excitation wavelengths were matched to the known absorption maxima of each fluorinated amino acid residue, with emission measured at 545 nm for terbium, 615 nm for europium and 570 nm for dysprosium. Samples were equilibrated for 15 minutes at room temperature prior to measurement. Each point reflects the average of three replicate wells from replicate titrations. The averaged dataset was fit to the Hill binding equation using GraphPad Prism:

$$I = \frac{B_{max} \cdot [Ln^{3+}]}{K_D + [Ln^{3+}]} \quad (10)$$

j) Expression of isotope labeled RF2

¹⁵N/¹³C-labeled RF2 was produced in a bioreactor (Infors-HT, Switzerland) at 300 mL media volume using a protocol by Klopp *et al.*¹²

k) NMR measurements of ¹³C/¹⁵N-labeled RF2

All NMR measurements were carried out at 30 °C using a Bruker 800 MHz NMR spectrometer. The backbone resonance assignments were obtained from the analysis of 3D HNCO, HN(CA)CO, HNCA, HNCACB, together with NOESY-¹⁵N-HSQC and TOCSY-¹⁵N-HSQC. Chemical shift assignments were deposited in the BMRB under the accession number 53251 and the title EF2-RF2. Spectra are available on Zenodo (DOI: 10.5281/zenodo.1578834).

l) Spectral overlap calculations

The spectral overlap calculations were performed using pre-processed emission and absorption spectra. Spectra were truncated to the range of 280–600 nm. The area under each spectrum was normalized to 1 using trapezoidal integration. Spectral overlap was calculated as the integral of the scalar product between the normalized Ln³⁺ absorption spectrum, and the normalized amino acid emission spectrum, resulting in a dimensionless overlap value between 0 and 1. These calculations were performed in MATLAB.

The absorption spectra of Eu³⁺ and Tb³⁺ were recorded on a Cary 50 UV/Vis spectrophotometer using a 400 μL quartz cuvette. Spectra were acquired in dual beam mode and baseline corrected by blanking with water. Both lanthanide ions were provided as 1 M chloride salt solutions.

Fluorescence emission spectra of the amino acids were recorded in a 400 μL quartz cuvette using a Cary Eclipse fluorescence spectrophotometer. The excitation slit width was set to 10 nm and the emission slit width to 1.5 nm. For all amino acids except mCNP and mCNF, a 295–1100 nm long-pass emission filter was used to suppress scattered excitation light and harmonic doubling. For mCNP and mCNF, the emission filter was left open.

Supplementary figures

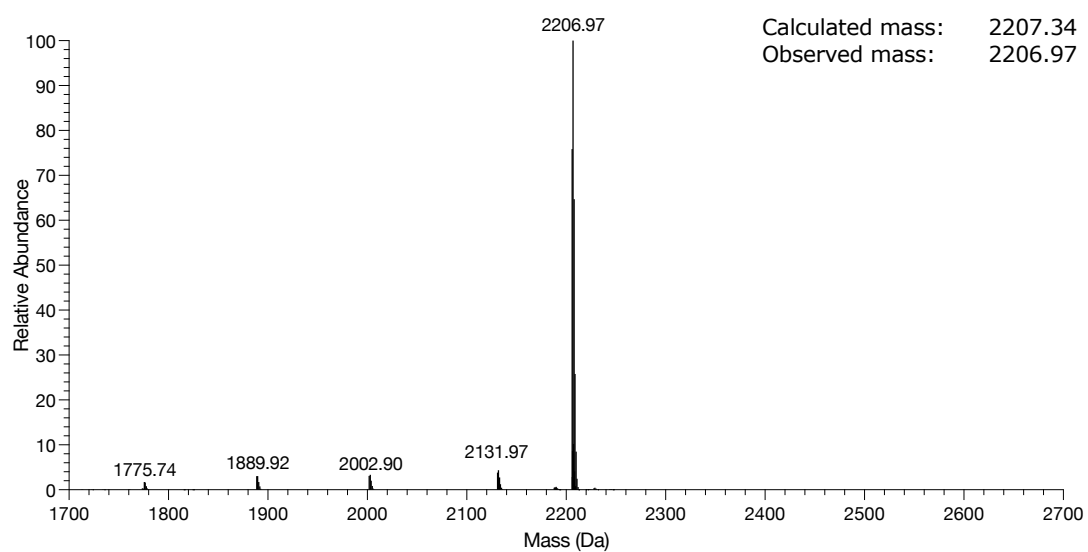


Figure S1. Intact protein mass spectrometry of LBT.

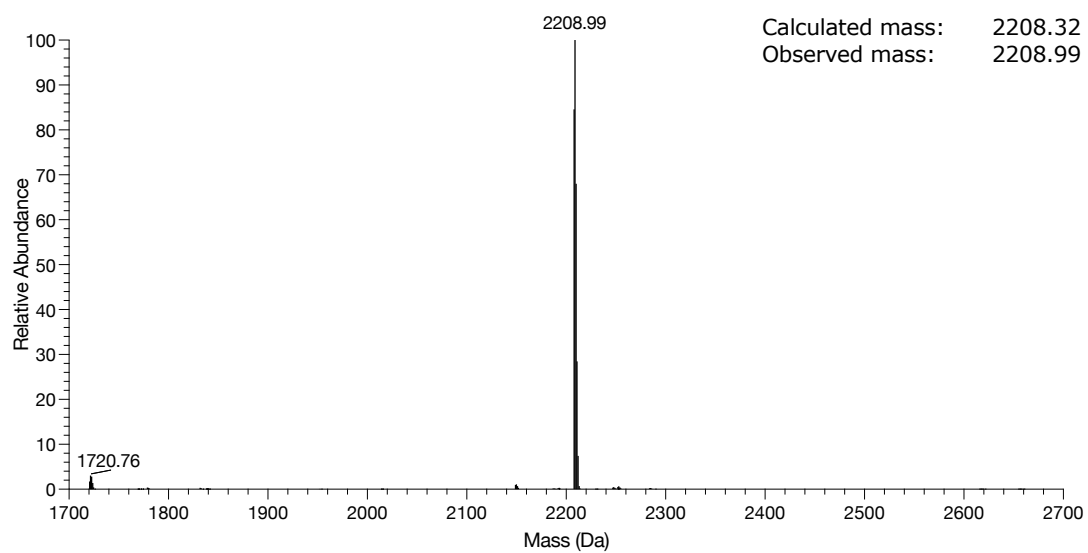


Figure S2. Intact protein mass spectrometry of LBT 4AW.

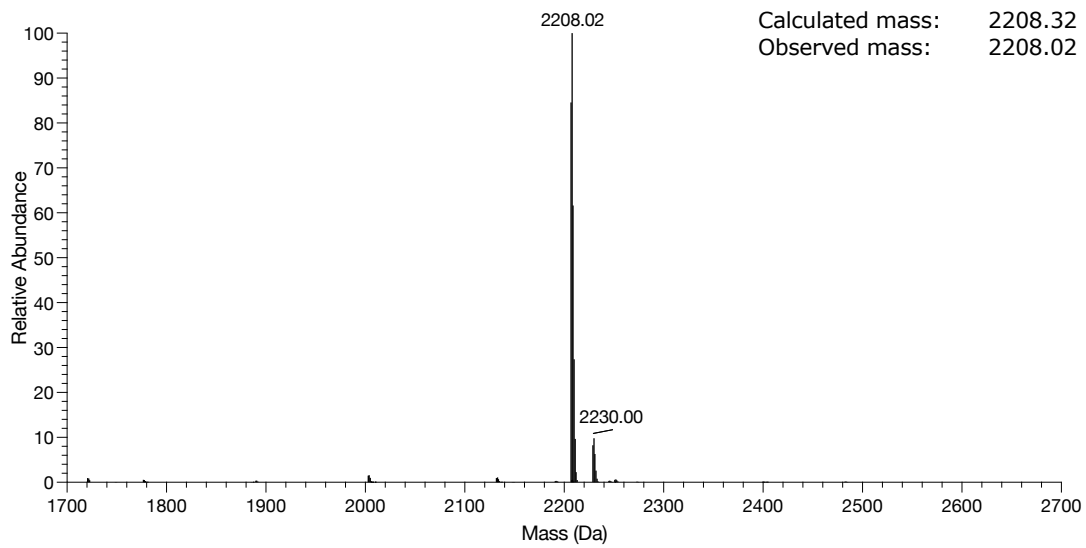


Figure S3. Intact protein mass spectrometry of LBT 6AW.

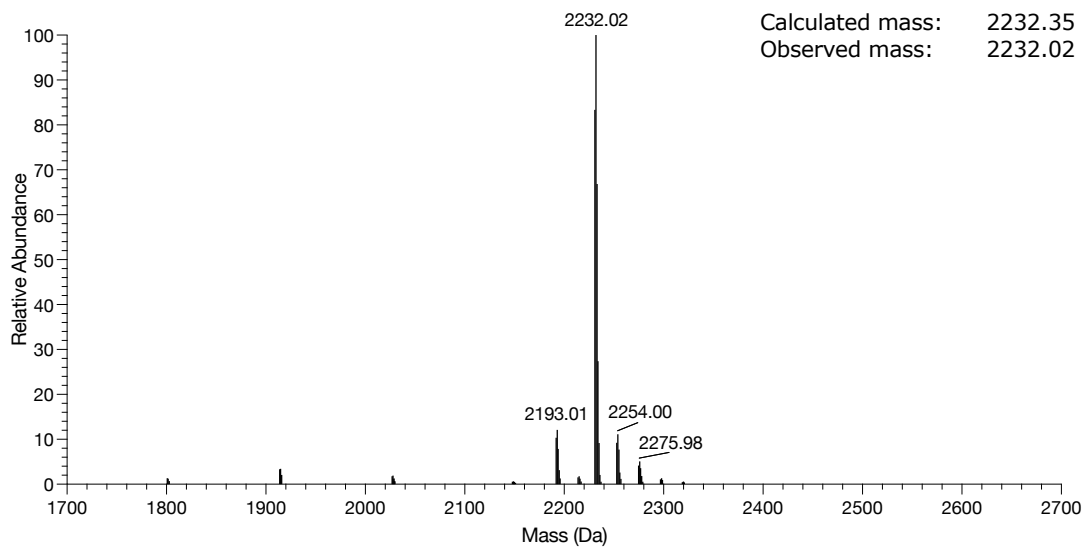


Figure S4. Intact protein mass spectrometry of LBT 4CNW.

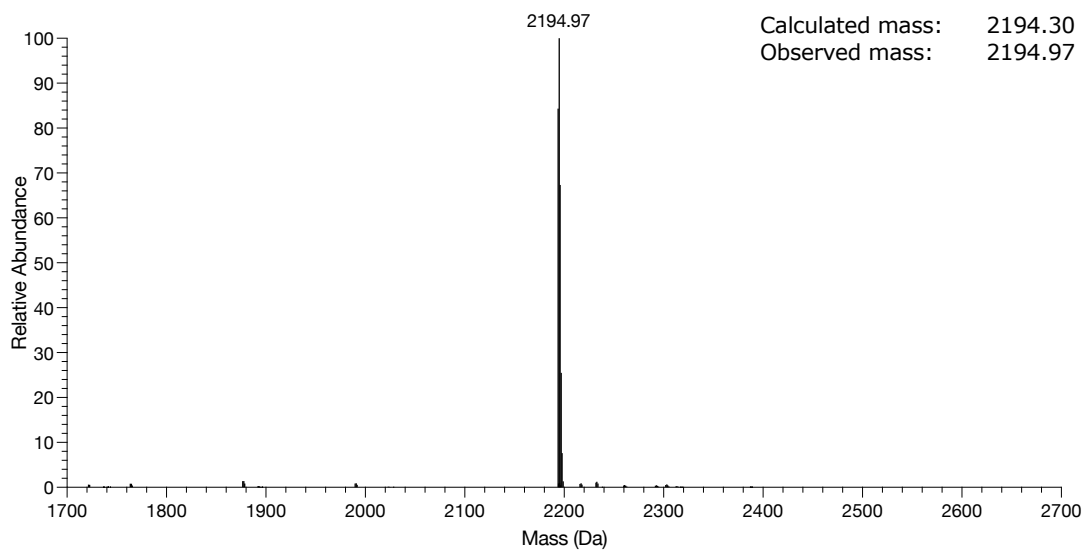


Figure S5. Intact protein mass spectrometry of LBT mCNP.

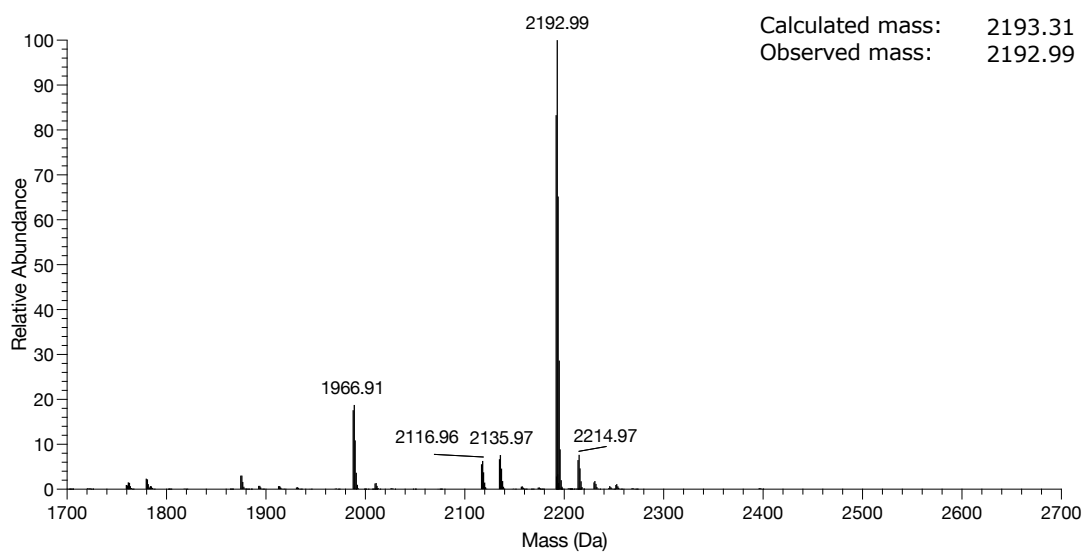


Figure S6. Intact protein mass spectrometry of LBT mCNF.

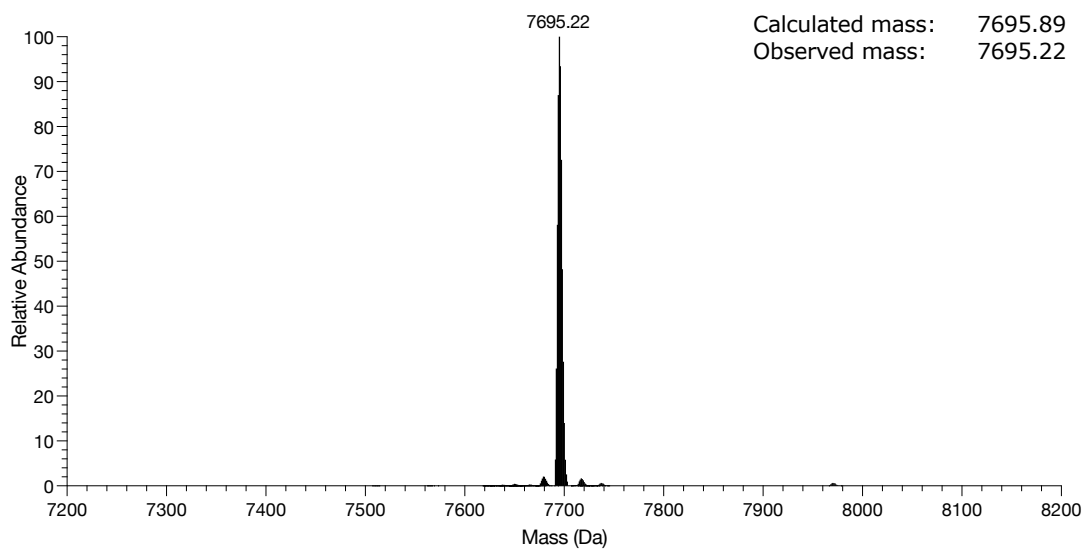


Figure S7. Intact protein mass spectrometry of the EF2 variant RF1.

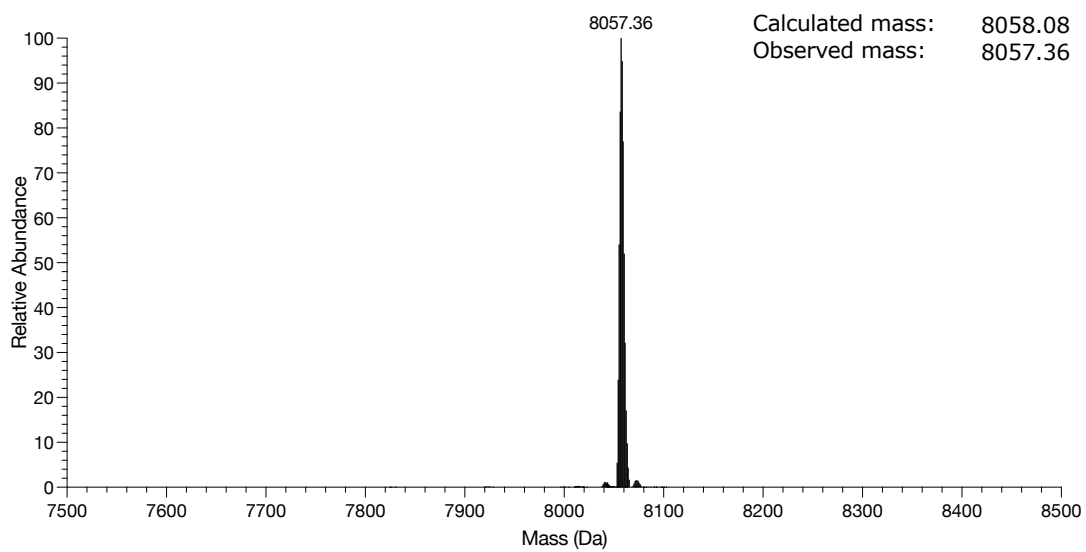


Figure S8. Intact protein mass spectrometry of the EF2 variant RF2.

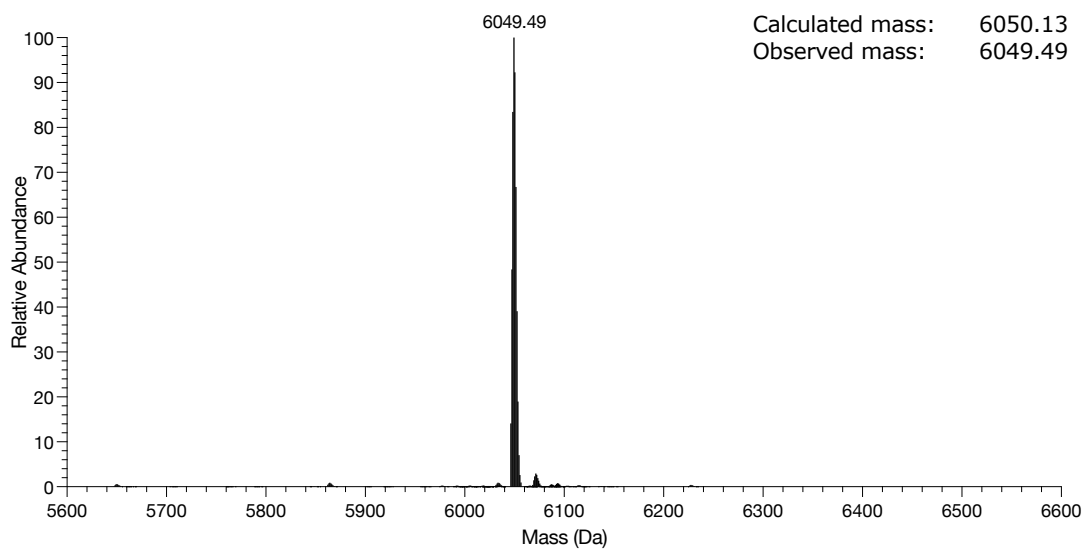


Figure S9. Intact protein mass spectrometry of the EF2 variant RF3.

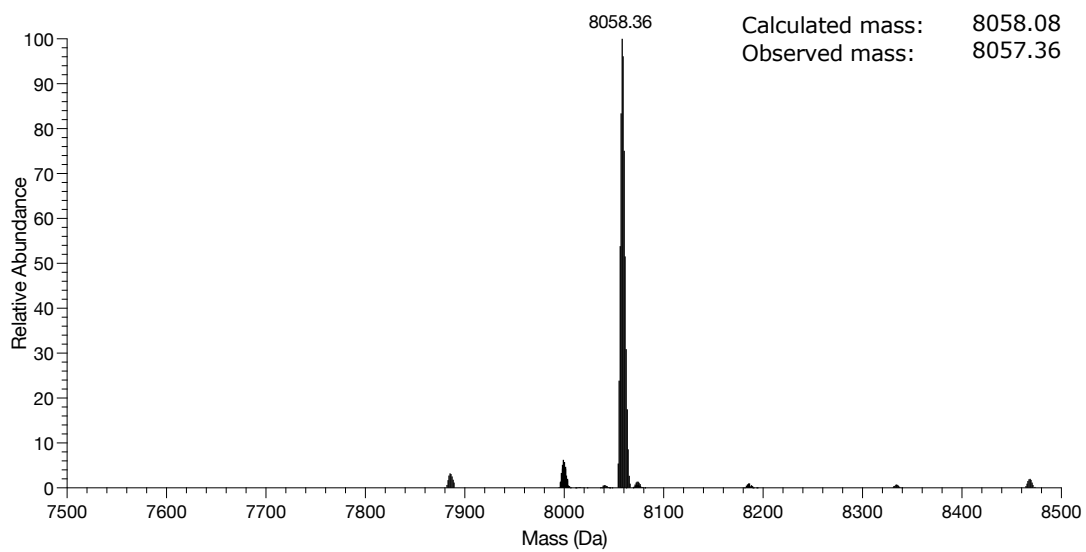


Figure S10. Intact protein mass spectrometry of the RF2 42 6AW.

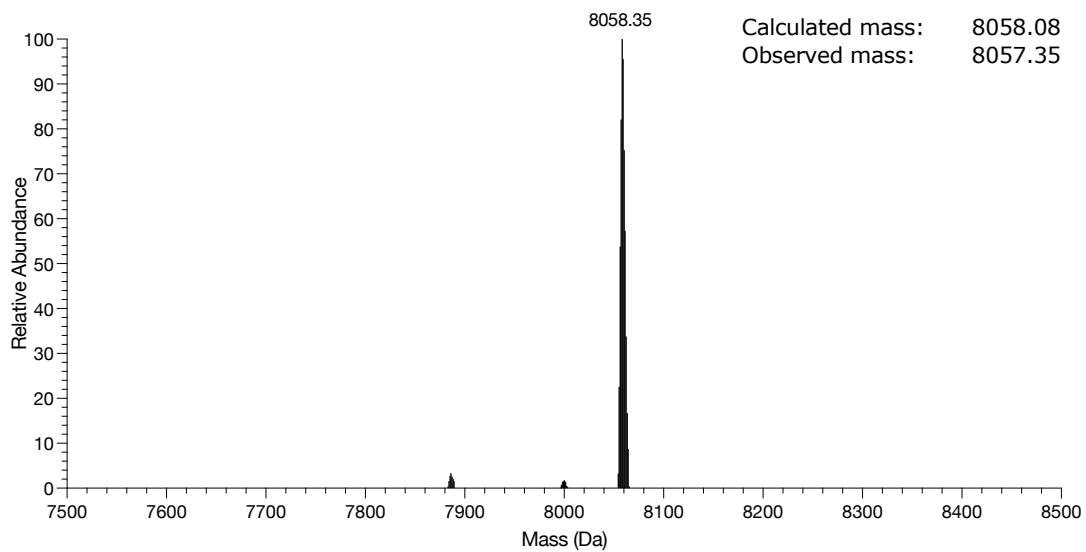


Figure S11. Intact protein mass spectrometry of the RF2 50 6AW.

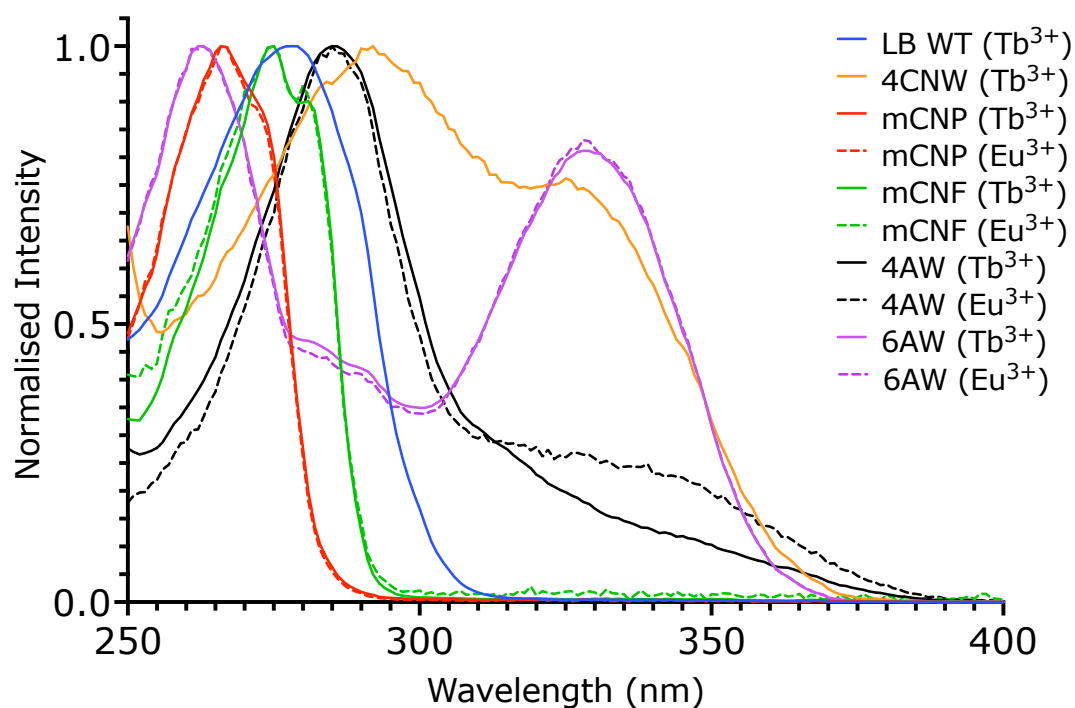


Figure S12. Excitation scans of the LBT complexes with Tb³⁺ and Eu³⁺. Emission at 545 nm or 615 nm was integrated for 5000 μ s after a delay of 100 μ s, using an excitation slit width of 5 nm and an emission slit width of 10 nm. Protein and lanthanide chloride were at 1 μ M and 2 μ M, respectively. The samples were prepared in buffer E. Each spectrum is normalized to the maximum intensity.

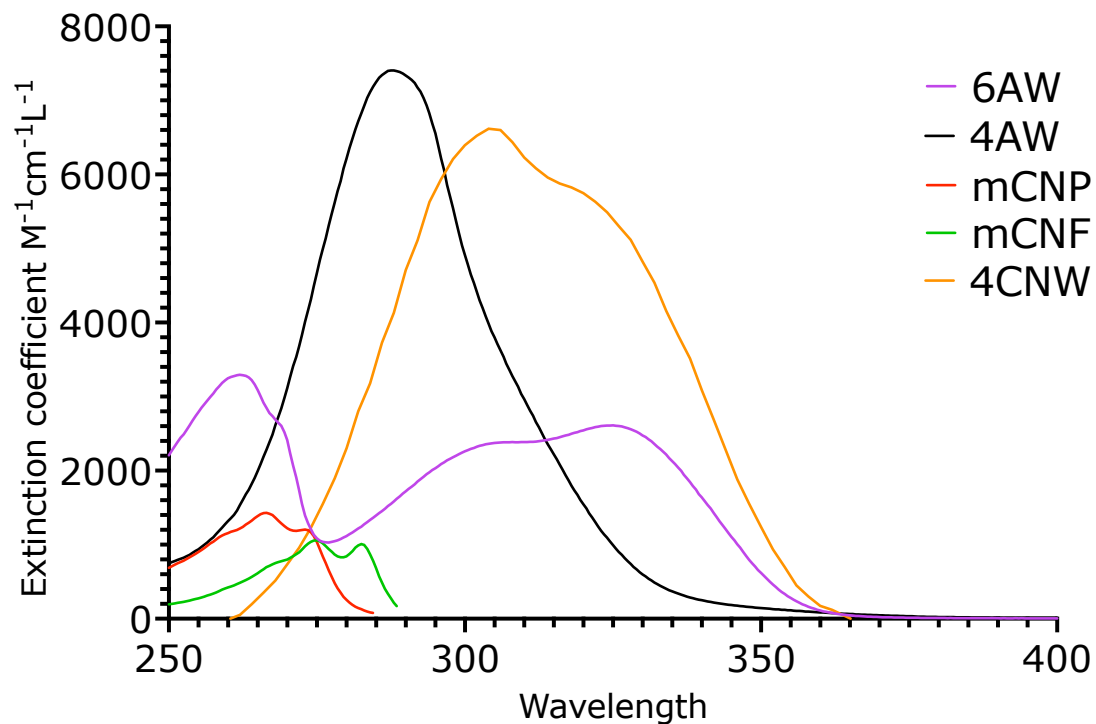


Figure S13. Absorbance spectra of the fluorescent ncAAs used in this study.

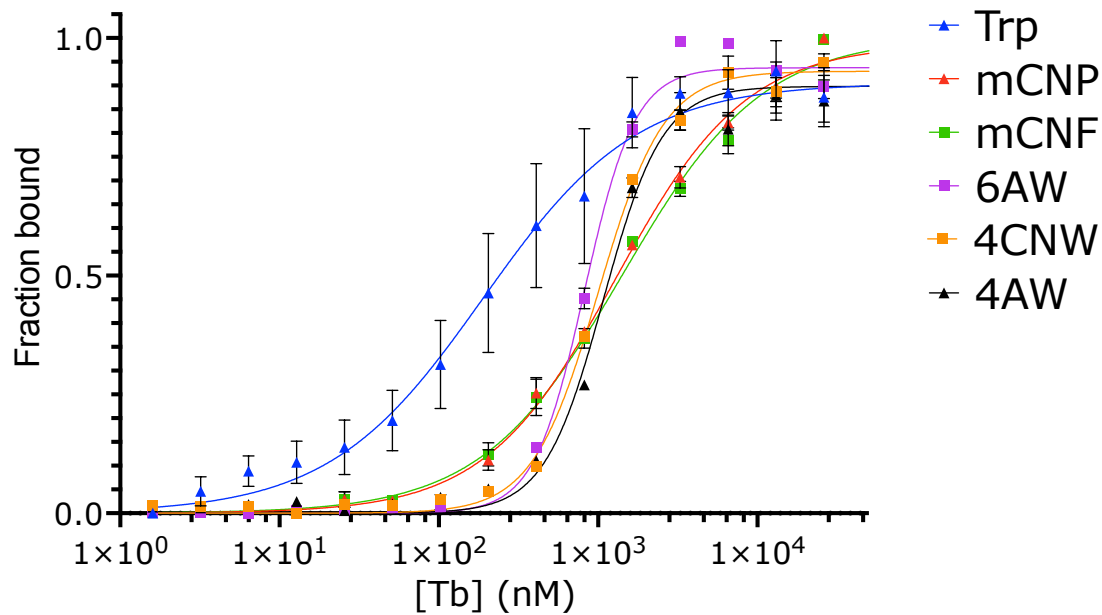


Figure S14. Binding affinity determination by titration of LBT variants with Tb^{3+} .

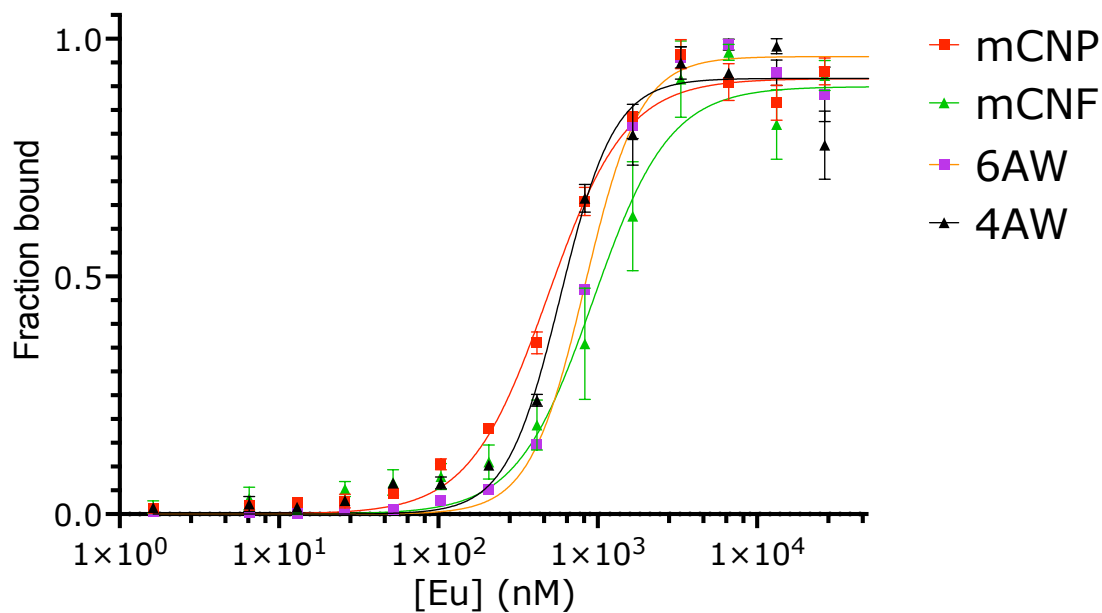


Figure S15. Binding affinity determination by titration of LBT variants with Eu^{3+} .

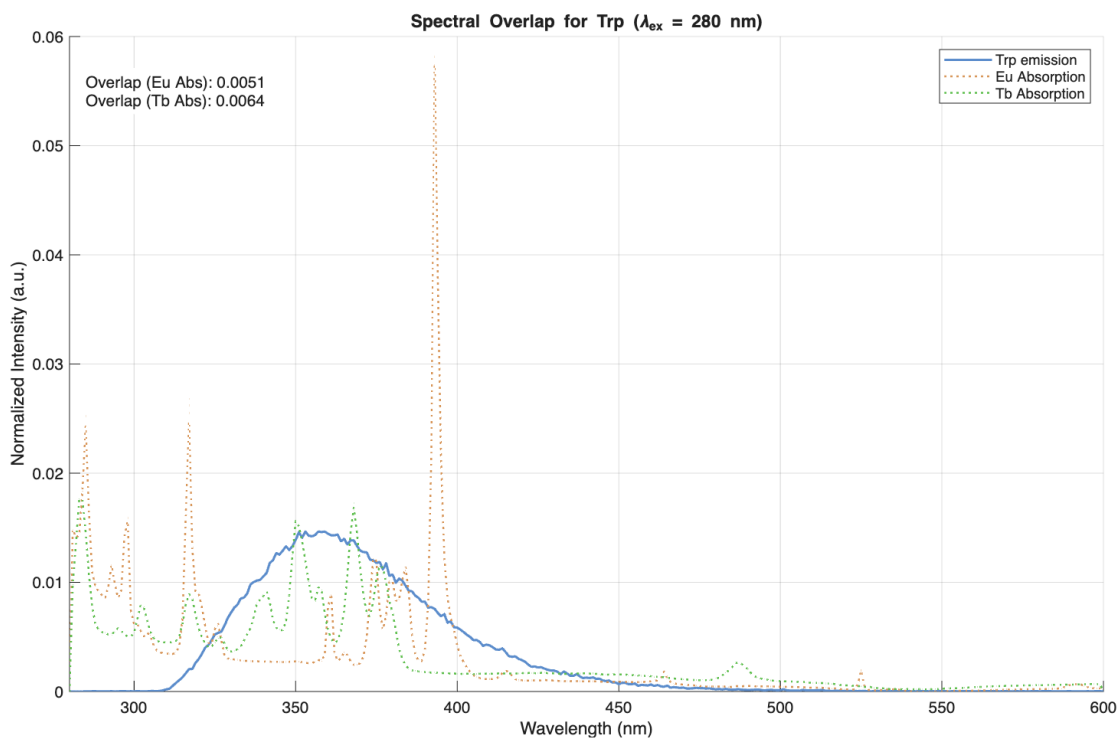


Figure S16. Spectral overlap of the Trp emission (free amino acid in Buffer E) with the absorption of Eu^{3+} and Tb^{3+} (LnCl_3 in MOPS buffer). Areas of each spectrum are normalized to 1 before plotting.

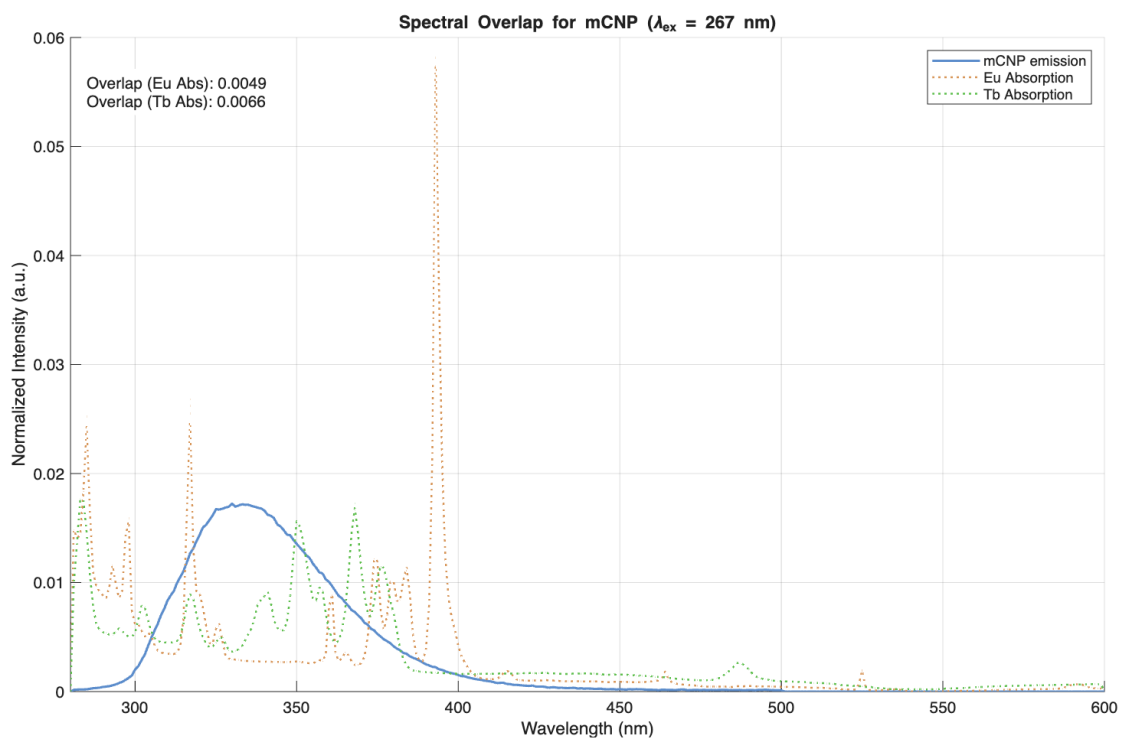


Figure S17. Spectral overlap of mCNP emission (free amino acid in Buffer E) with the absorption of Eu^{3+} and Tb^{3+} (LnCl_3 in MOPS buffer). Areas of each spectrum are normalized to 1 before plotting.

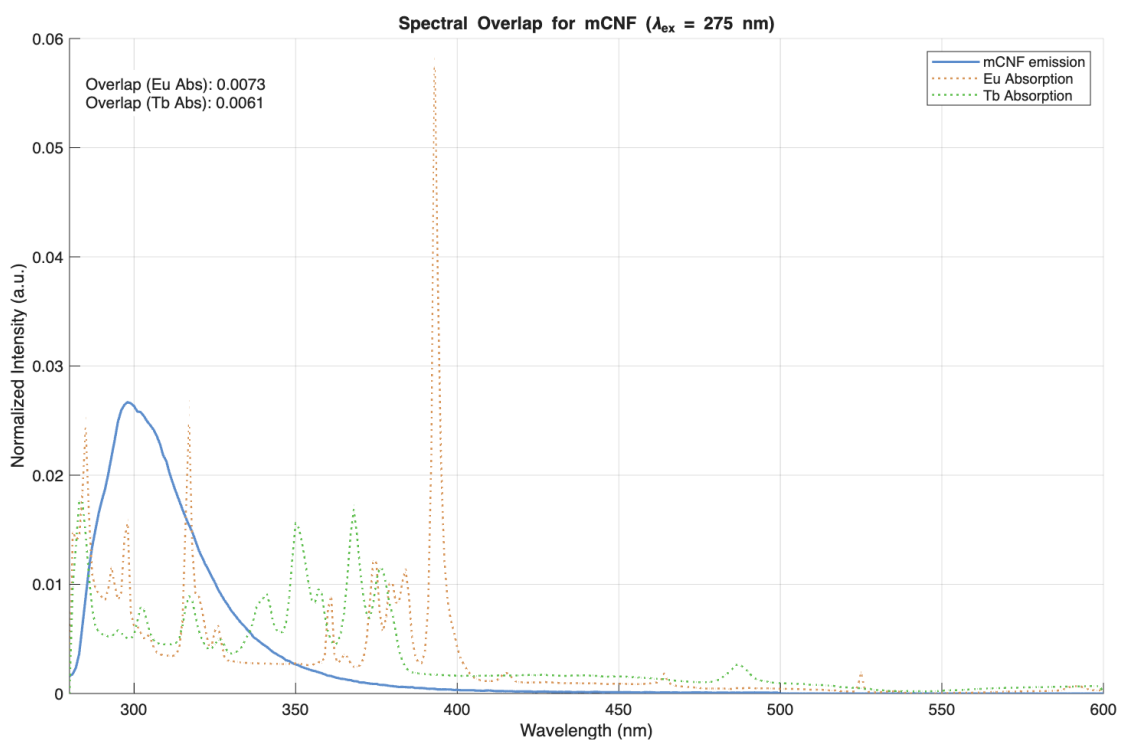


Figure S18. Spectral overlap of mCNF emission (free amino acid in Buffer E) with the absorption of Eu^{3+} and Tb^{3+} (LnCl_3 in MOPS buffer). Areas of each spectrum are normalized to 1 before plotting.

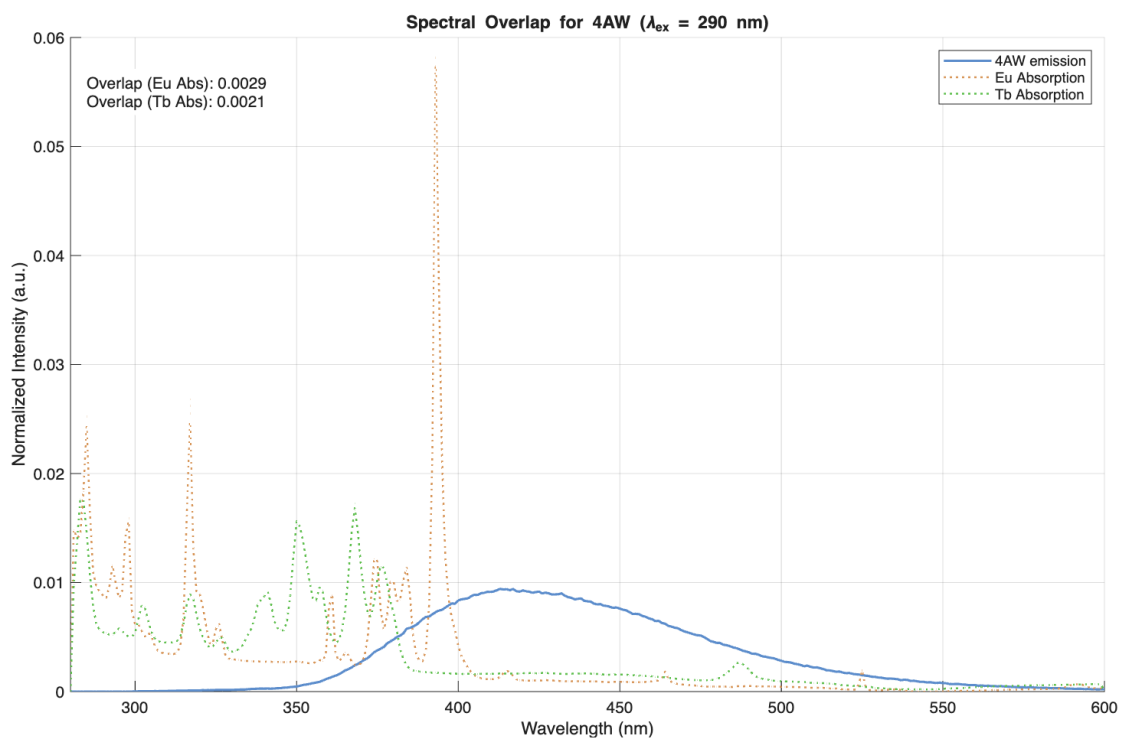


Figure S19. Spectral overlap of 4AW emission (free amino acid in Buffer E) with the absorption of Eu^{3+} and Tb^{3+} (LnCl_3 in MOPS buffer). Areas of each spectrum are normalized to 1 before plotting.

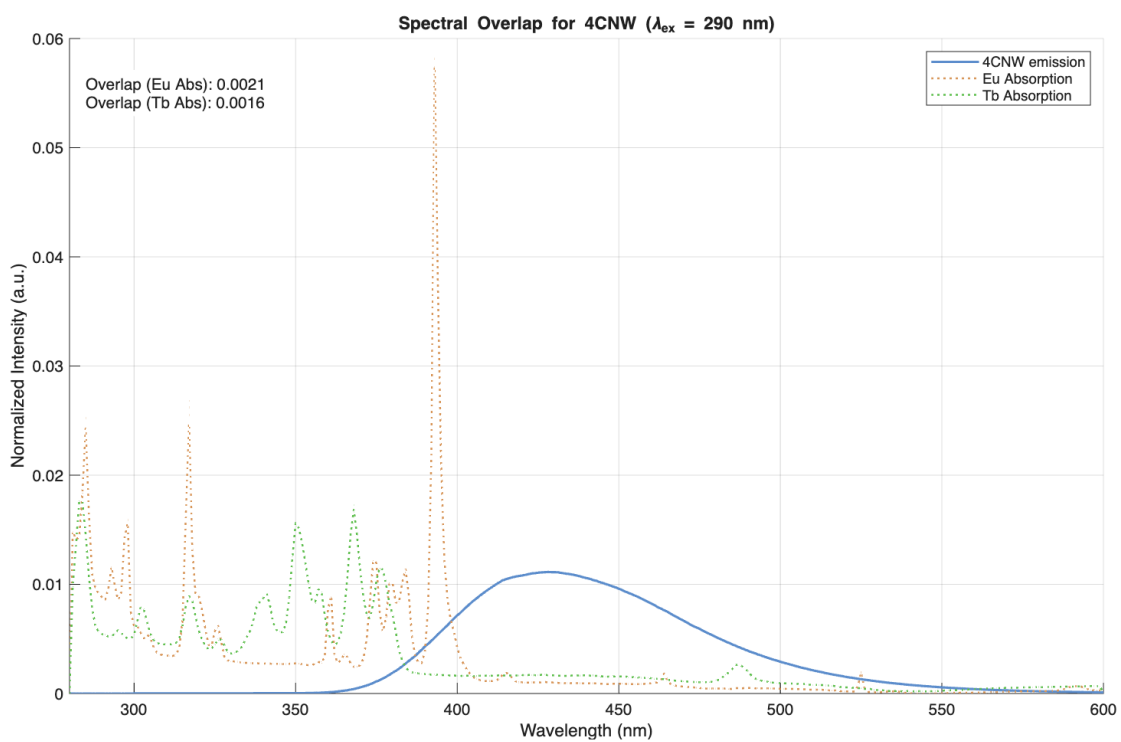


Figure S20. Spectral overlap of 4CNW emission (free amino acid in Buffer E) with the absorption of Eu^{3+} and Tb^{3+} (LnCl_3 in MOPS buffer). Areas of each spectrum are normalized to 1 before plotting.

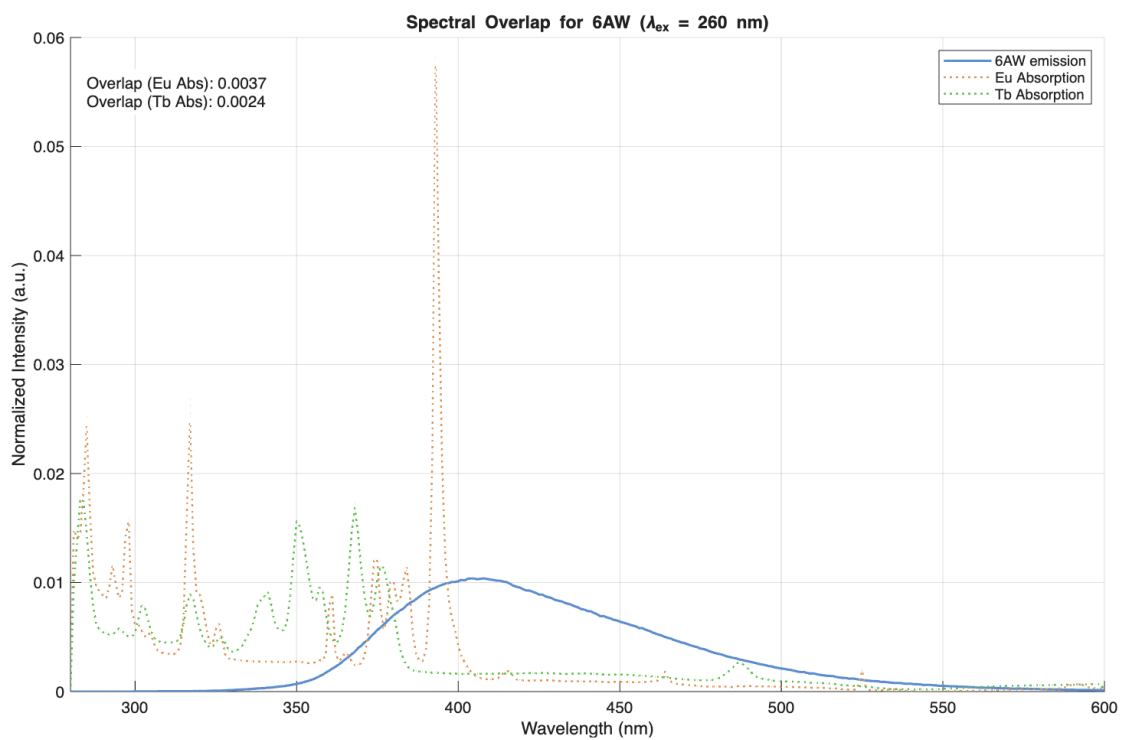


Figure S21. Spectral overlap of 6AW emission (free amino acid in Buffer E) with the absorption of Eu^{3+} and Tb^{3+} (LnCl_3 in MOPS buffer). Areas of each spectrum are normalized to 1 before plotting.

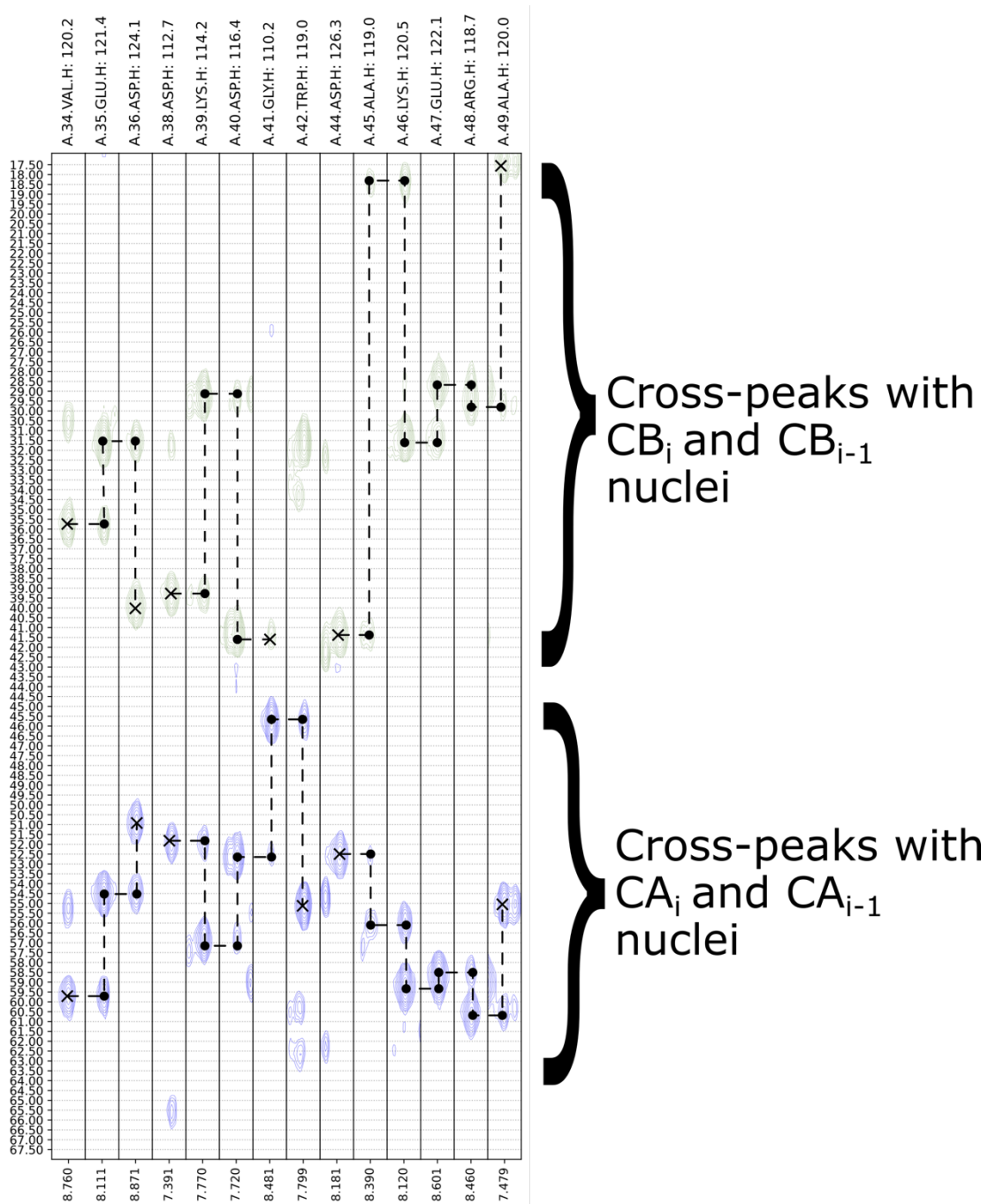


Figure S22. HNCACB data of $^{15}\text{N}/^{13}\text{C}$ -labeled RF2. The strip plots are of the 3D HNCACB spectrum of the lanthanide-binding EF-Hand of RF2 including 4 flanking residues (residues 34–49): VEDPDKDGWLDAKERA, where underlined residues do not have assigned NH peaks. The strips illustrate the sequential CB_i - CB_{i-1} peaks along the sequence.

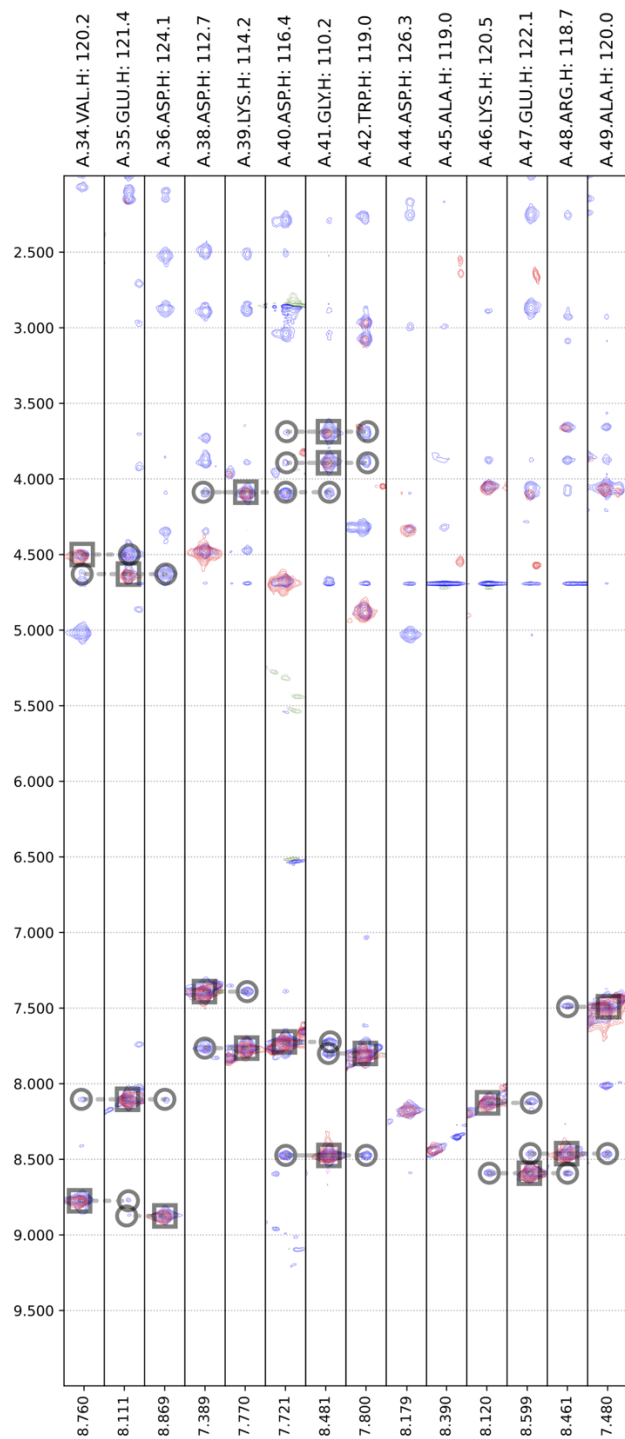


Figure S23. Overlay of ^{13}C -decoupled 3D NOESY- ^{15}N -HSQC (blue) and TOCSY- ^{15}N -HSQC (red) spectra of $^{15}\text{N}/^{13}\text{C}$ -labeled RF2. The strips belong to the lanthanide-binding EF-Hand of RF2 and 4 flanking residues (res 34–49) VEDPDKGWLDAKERA, where underlined residues do not have assigned NH peaks. Intra- and inter-residual cross-peaks are identified by squares and circles, respectively.

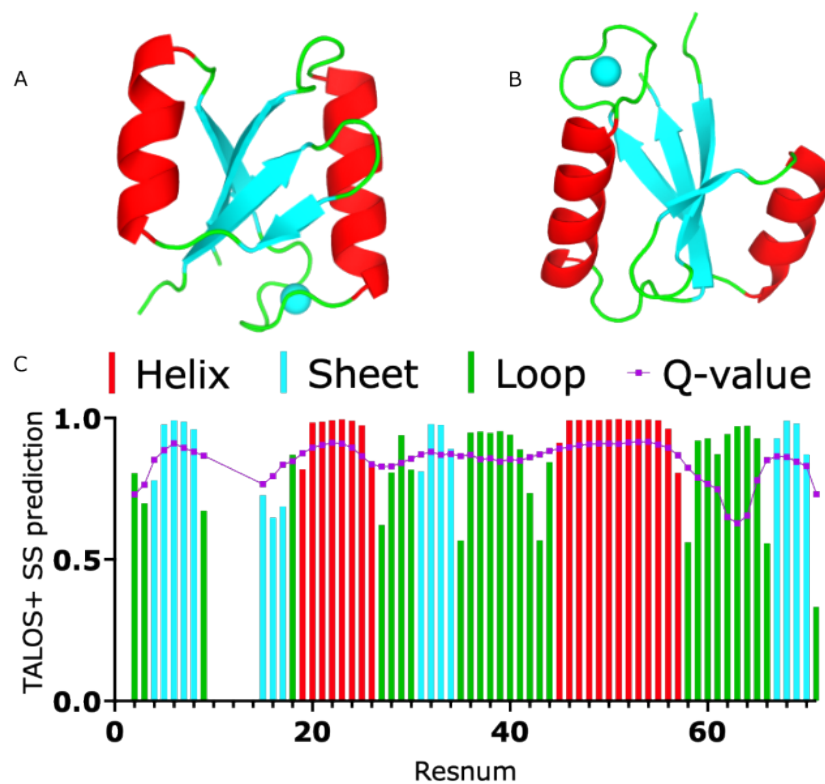


Figure S24. Comparison of the secondary structure of the AlphaFold3 model and the prediction made from the assigned chemical shifts. (A) and (B) Two different orientations of the AlphaFold3 model of RF2 generated with a Ca^{2+} ion (cyan). The cartoon representation is colored by secondary structure: red-helix, cyan-sheet, green-loop. (C) Output from the TALOS+ webserver.¹³ Colored bars indicate the secondary structure prediction associated with the residue number. The Q-value (purple) indicates the confidence for that prediction.

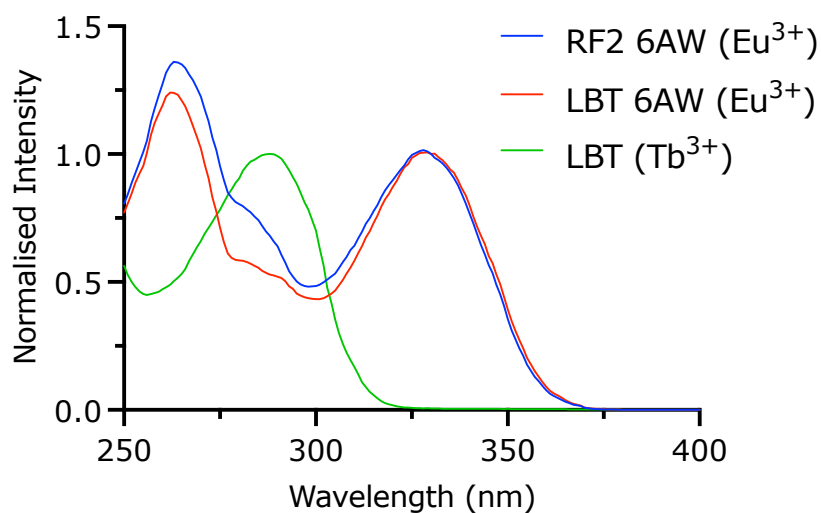


Figure S25. Comparison of excitation spectra for RF2 6AW/Eu³⁺, LBT 6AW/Eu³⁺ and LBT/Tb³⁺. RF2 6AW/Eu³⁺ (blue) and LBT 6AW/Eu³⁺ have been normalized to the intensity of 330 nm, where the contribution of canonical amino acids to sensitization is negligible. LBT/Tb³⁺ has been normalized to the intensity at 280 nm, as in Figure S21.

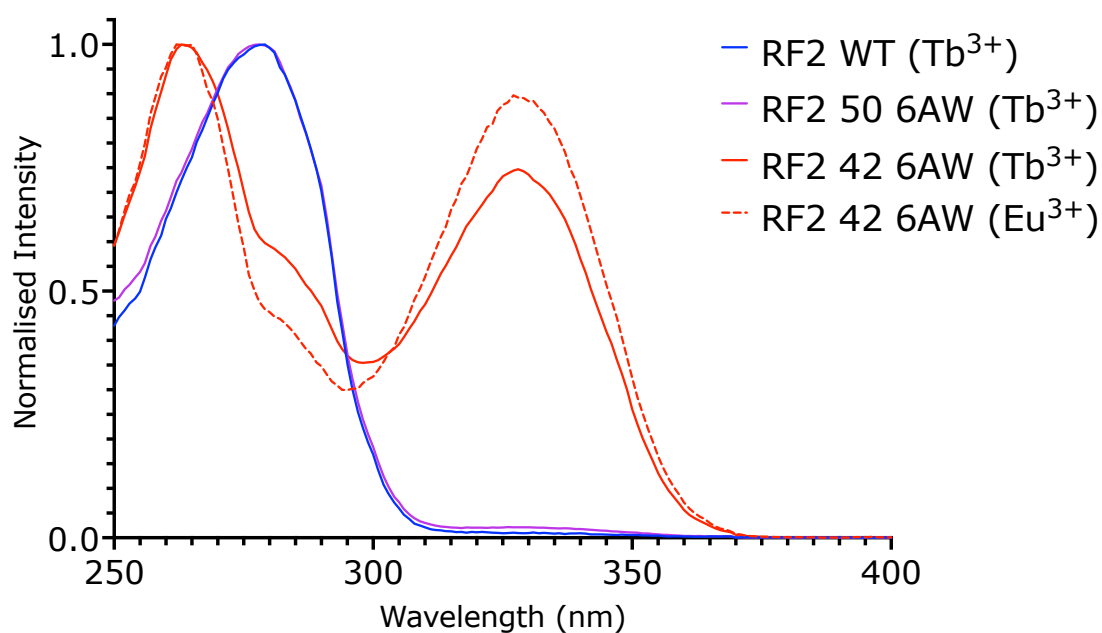


Figure S26. Excitation scans of the RF2 complexes with Tb³⁺ and Eu³⁺. Emission at 545 nm or 615 nm was integrated for 5000 μ s after a delay of 100 μ s, using an excitation slit width of 5 nm and an emission slit width of 10 nm. Protein and lanthanide chloride were at 1 μ M and 2 μ M, respectively. The samples were prepared in buffer E. Each spectrum is normalized to the maximum intensity.

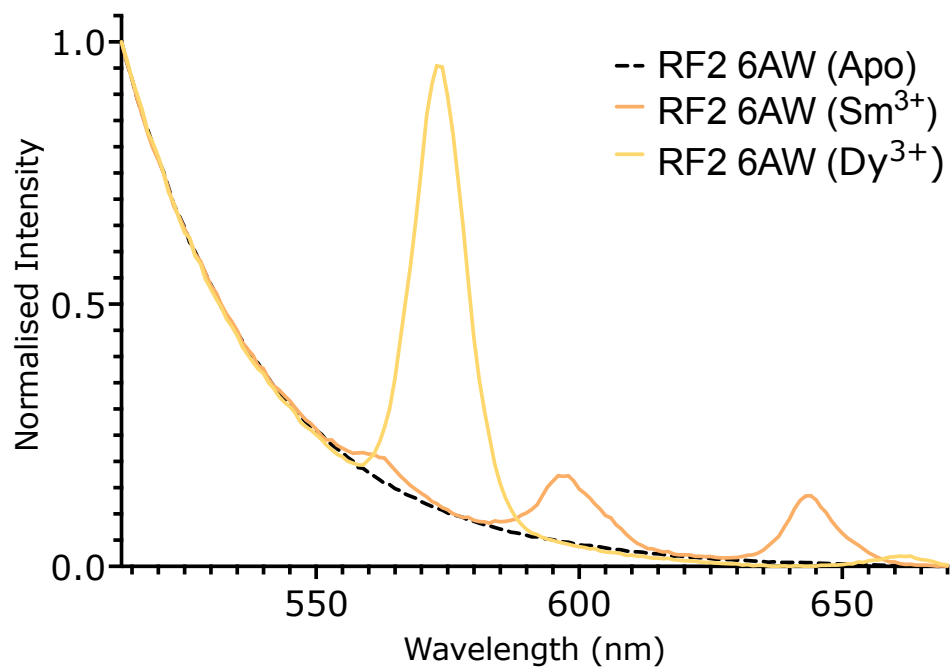


Figure S27. Fluorescence emission spectra of RF2 6AW with Sm^{3+} and Dy^{3+} . Each spectrum is normalized to the intensity at 520 nm.

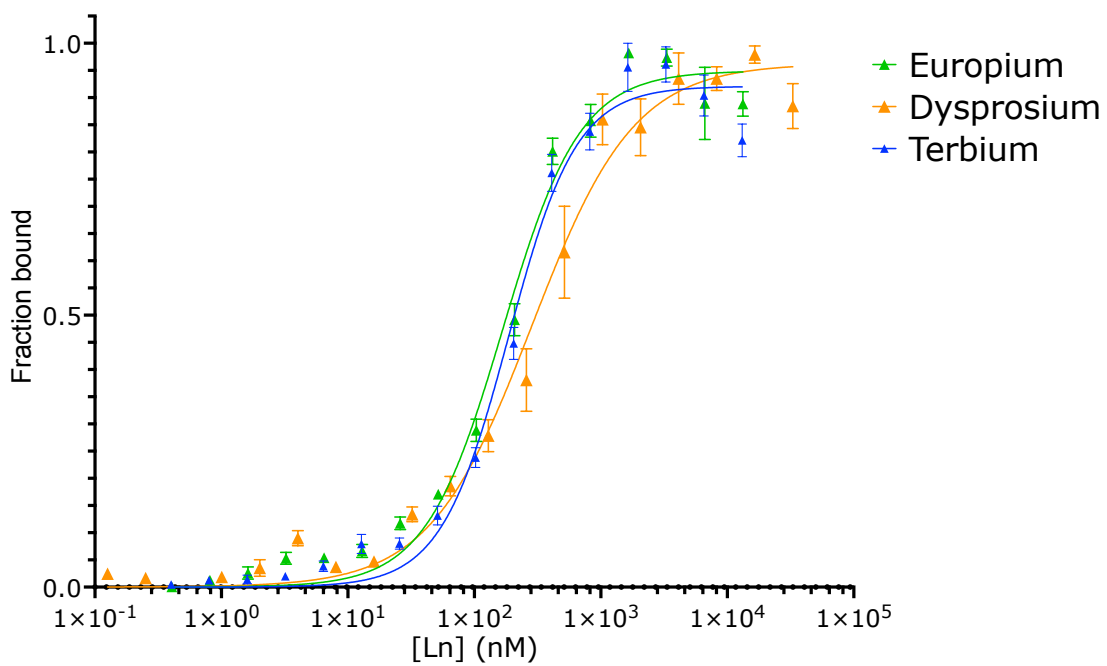


Figure S28. Binding affinity determination by titration of RF2 6AW complexes with Tb^{3+} , Eu^{3+} , and Dy^{3+} .

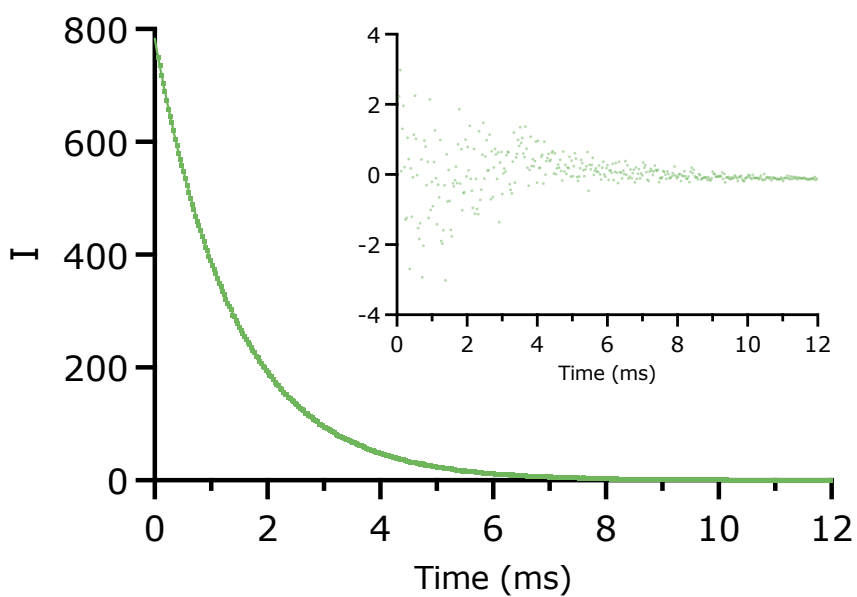


Figure S29. Luminescence lifetime measurement of RF2 with Tb^{3+} . The excitation wavelength was 280 nm and emission was recorded at 545 nm. Inset: residuals of the experimental values from the curve fit. The best-fit value for τ_{Tb} is 1.4 ms.

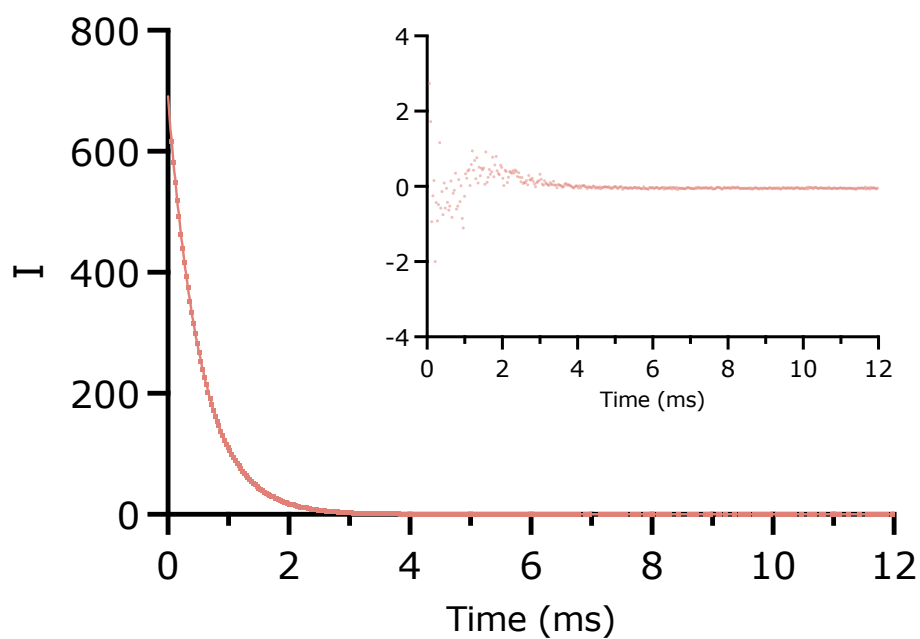


Figure S30. Luminescence lifetime measurement of RF2 6AW with Eu^{3+} . The excitation wavelength was 330 nm and emission was recorded at 615 nm. Inset: residuals of the experimental values from the curve fit. The best-fit value for τ_{Eu} is 0.54 ms.

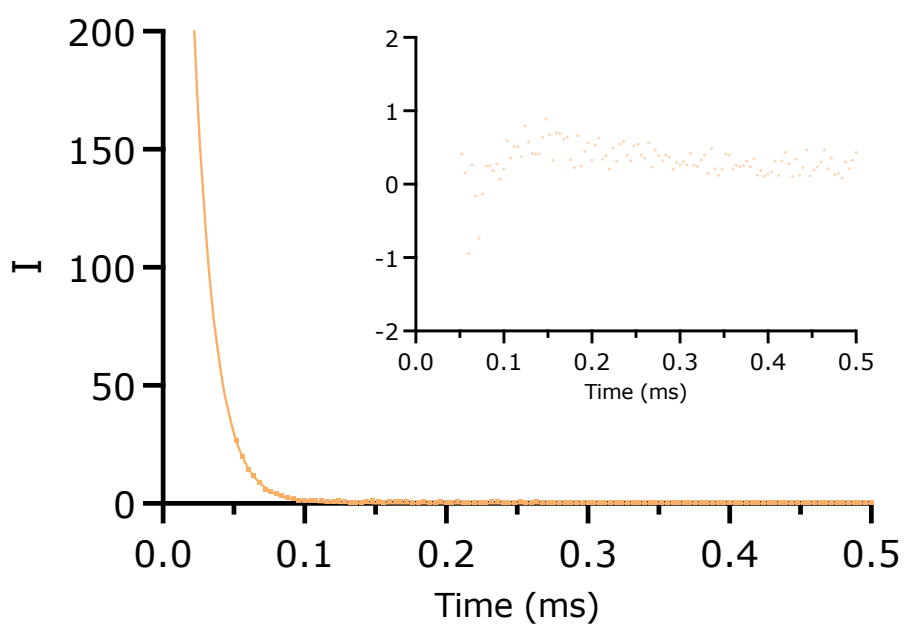


Figure S31. Luminescence lifetime measurement of RF2 6AW with Sm^{3+} . The excitation wavelength was 265 nm and emission was recorded at 650 nm. The instrument gain was set to 1000. Inset: residuals of the experimental values from the curve fit. The best-fit value for τ_{Sm} is 12 μs .

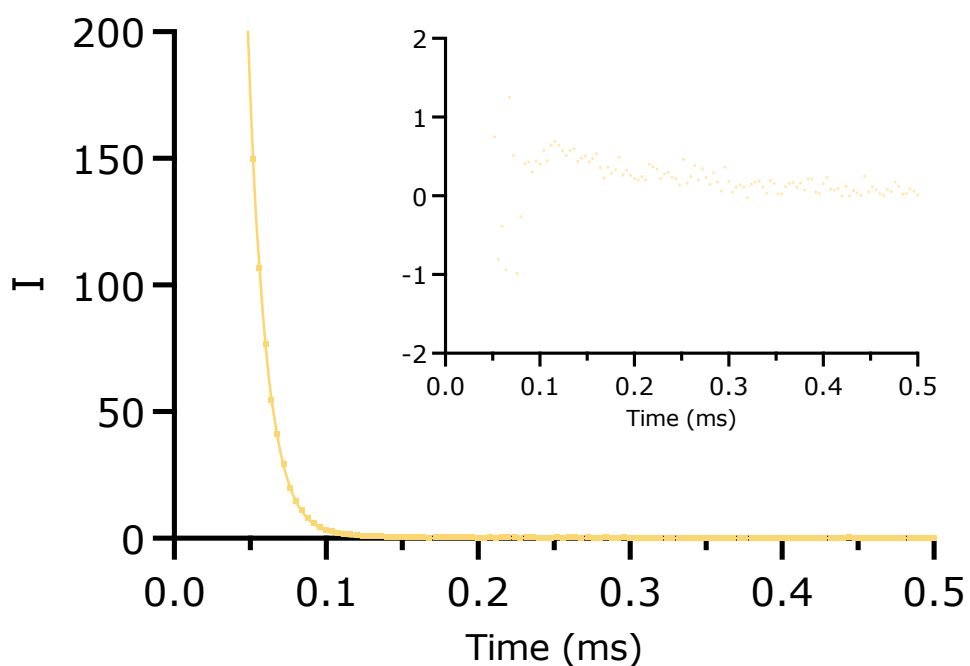


Figure S32. Luminescence lifetime measurement of RF2 6AW with Dy^{3+} . The excitation wavelength was 265 nm and emission was recorded at 570 nm. Inset: residuals of the experimental values from the curve fit. The best-fit value for τ_{Dy} is 14 μs .

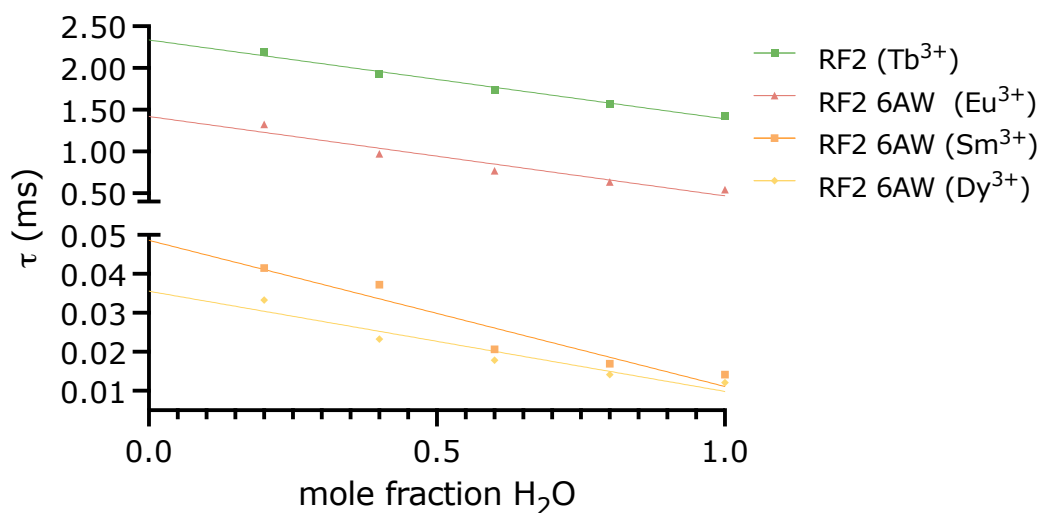


Figure S33. D_2O titrations of RF2/ Ln^{3+} complexes. The luminescence lifetimes are plotted versus increasing H_2O concentrations. Extrapolation to a H_2O mole fraction of 0 gives the lifetime in D_2O .

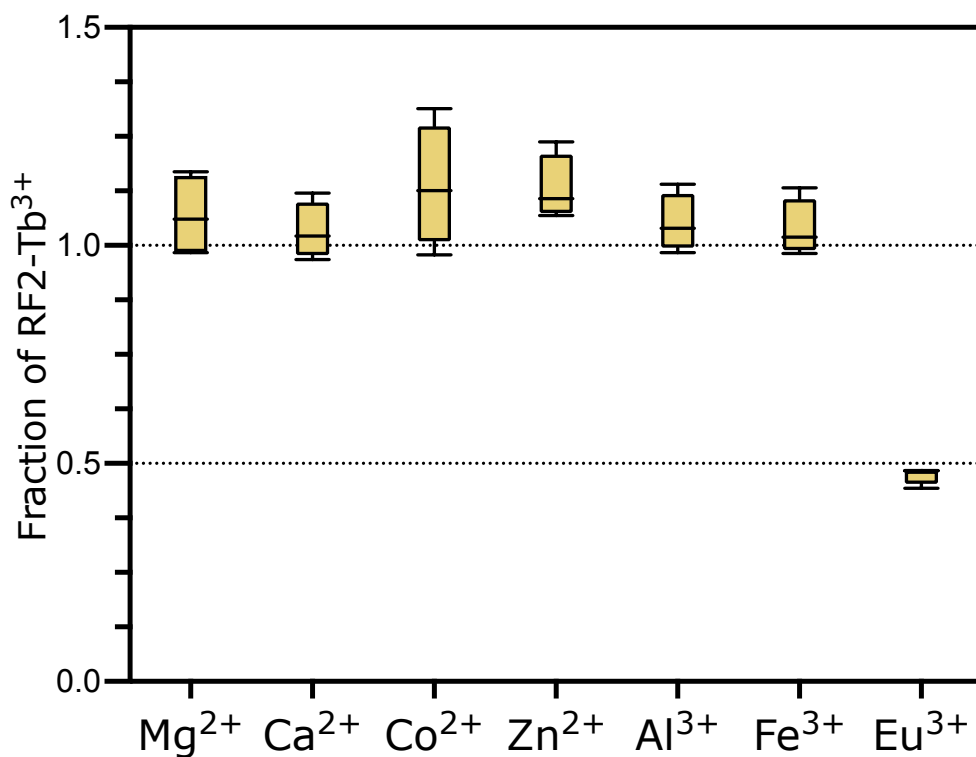


Figure S34. Competition experiment with M²⁺/M³⁺ ions. Fold-change in the luminescence intensity (gold) of RF2 (1 μ M) samples with Tb³⁺ (2 μ M) and competitor (2 μ M) in 20 mM MOPS pH 7.0, 50 mM KCl. Each data point represents the ratio of integrated emission intensity (collected at 545 nm emission with a 100 μ s delay and 1 ms integration window, following excitation at 280 nm) normalized to the intensity of a sample containing only Tb³⁺. Boxes indicate the mean and range of four technical replicates.

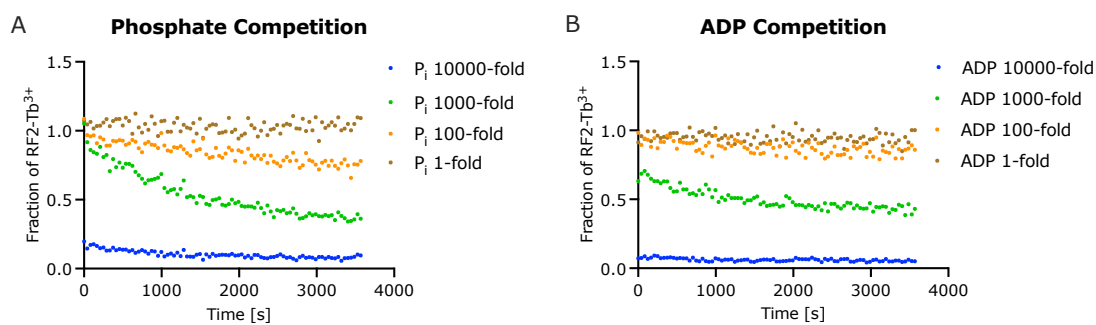


Figure S35. Competition experiment with excess phosphate and ADP anions relative to Tb³⁺. Fold-change in the luminescence intensity of RF2 (1 μM) samples with Tb³⁺ (2 μM) in the presence of competitor at 2 μM, 200 μM, 2 mM and 20 mM. The concentrations of anions represent a 1-fold, 100-fold, 1000-fold and 10000-fold excess, respectively. All samples and anion stocks were prepared in 20 mM MOPS pH 7.0, 50 mM KCl. Each data point represents the ratio of integrated emission intensity (collected at 545 nm emission with a 100 μs delay and 1 ms integration window, following excitation at 280 nm) normalized to the intensity of a sample containing only Tb³⁺. Points indicate the mean of three technical replicates measured at regular intervals for 1 hour.

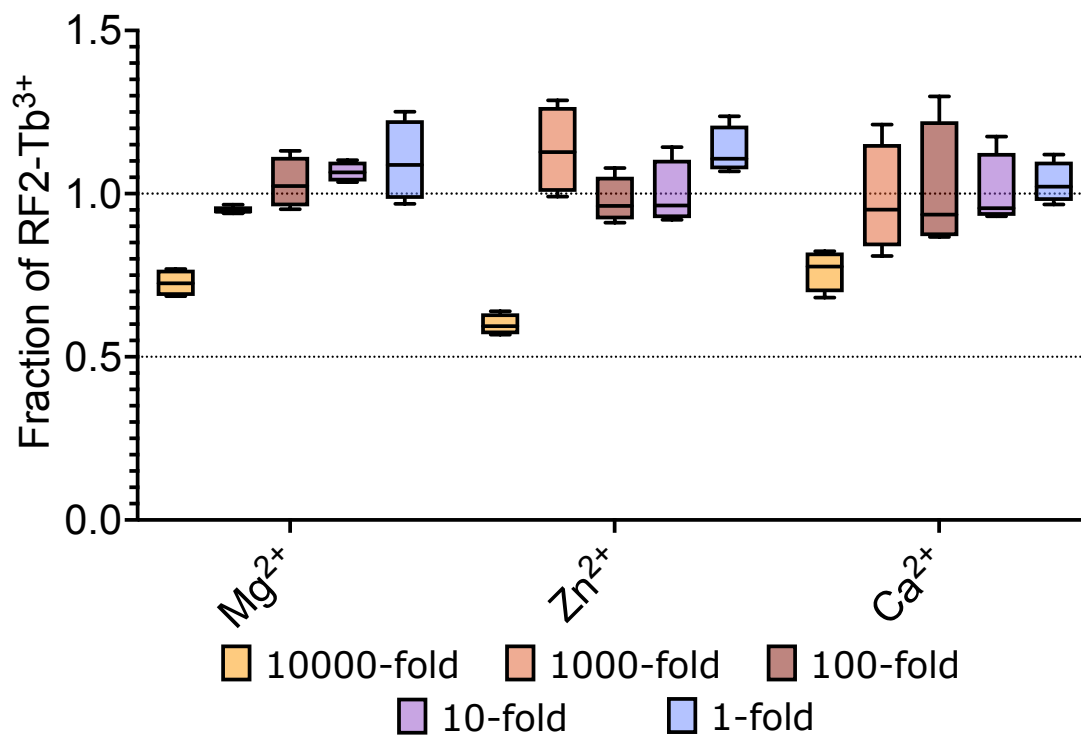


Figure S36. Competition experiment with excess M^{2+} ions relative to Tb^{3+} . Fold-change in the luminescence intensity of RF2 (1 μ M) samples with Tb^{3+} (2 μ M) in the presence of competitor at 2 μ M, 20 μ M, 200 μ M, 2 mM and 20 mM. The concentrations of M^{2+} ions represent a 1-fold, 10-fold, 100-fold, 1000-fold and 10000-fold excess, respectively. All samples and stocks of metal ions were prepared in 20 mM MOPS pH 7.0, 50 mM KCl. Each data point represents the ratio of integrated emission intensity (collected at 545 nm emission with a 100 μ s delay and 1 ms integration window, following excitation at 280 nm) normalized to the intensity of a sample containing only Tb^{3+} . Boxes indicate the mean and range of four technical replicates.

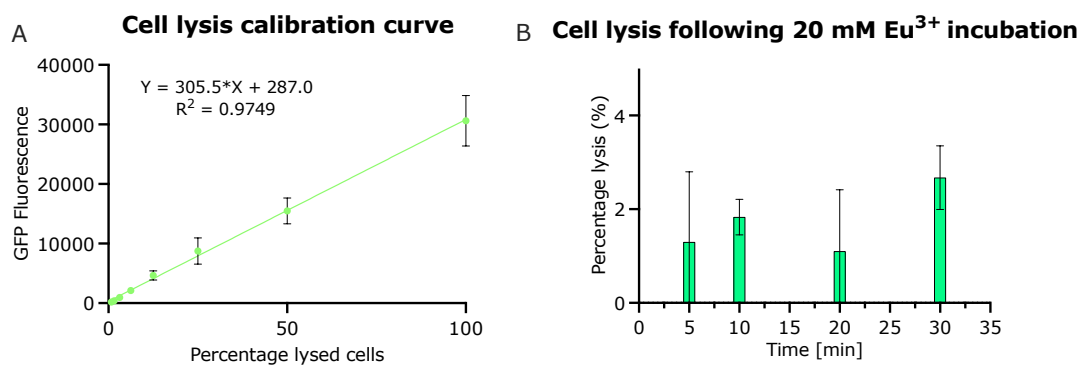


Figure S37. Quantification of cell lysis following incubation with 20 mM Eu³⁺ in 150 mM NaCl. (A) B95 cells which had overexpressed green fluorescent protein (GFP) were resuspended to OD₆₀₀ = 1 in 150 mM NaCl and sonicated. A calibration curve using GFP fluorescence ($\lambda_{\text{ex}} = 482 \text{ nm}$, $\lambda_{\text{em}} = 528 \text{ nm}$) was established using the supernatant of centrifuged mixtures of cells before and after lysis. A linear regression provides an equation that can be used to estimate cell lysis in supernatant from observed GFP fluorescence (inset). (B) Percentage lysis following incubation with 20 mM Eu³⁺. Cells were resuspended to OD₆₀₀ = 1 in 150 mM NaCl and incubated with 20 mM Eu³⁺ (as EuCl₃) after 5, 10, 20 and 30 minutes. y-values are calculated using the equation in (A) and plotted as means with standard deviations for three technical replicates.

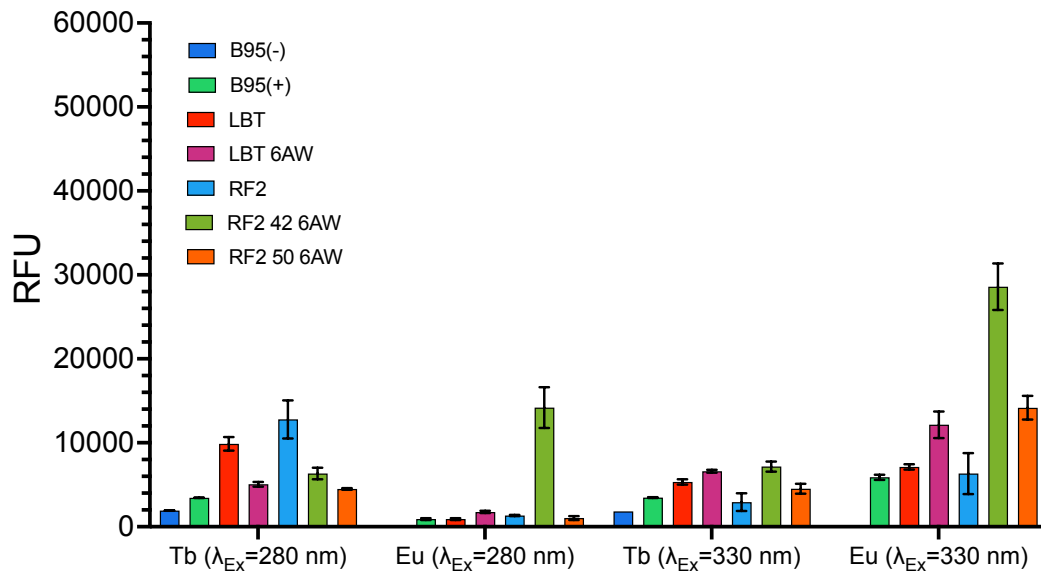


Figure S38. *In vivo* sensitization of lanthanide luminescence. Untransformed B95 cells with and without 6AW (B95(+), B95(-)), transformed B95 cells with a WT LBT or RF2 plasmid, or the two-plasmid expression system for installing 6AW were grown to $OD_{600} = 0.6-0.8$ in LB media and supplemented with 1 mM IPTG. Cultures with cells containing the G1(6AW)RS plasmid, as well as the B95(+) cells were supplemented with 1 mM 6AW. Cells were cultured overnight at RT and resuspended to $OD_{600} = 1.0$ in 150 mM NaCl. The standardized cell resuspensions were then resuspended again in 150 mM NaCl containing 20 mM $EuCl_3$ or 20 mM $TbCl_3$. The pH of the resuspensions was measured using litmus paper to be pH = 6.0. After 15 minutes, cells were washed with 150 mM NaCl. Readings are integrated emission intensities over 1000 μs following a 600 μs delay. Emission intensities were $\lambda_{Em} = 545$ nm for Tb, and $\lambda_{Em} = 615$ nm for Eu. Datasets for Eu- and Tb-incubated cells were recorded with different gain settings for the detector. Therefore, the values between them are not directly comparable.

***in vivo* sensitization in unwashed cells**

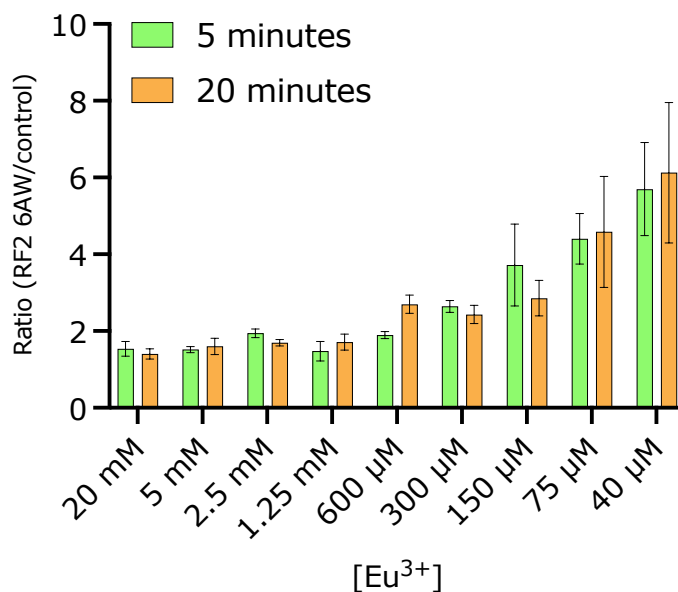


Figure S39. *In vivo* sensitization of europium luminescence in unwashed cells. Untransformed B95 cells incubated with 1 mM 6AW were used as a control, B95 cells with the two-plasmid expression system for installing 6AW at position 42 in RF2 were grown to $\text{OD}_{600} = 0.6\text{--}0.8$ in LB media and supplemented with 1 mM IPTG. Cells were cultured overnight at RT and resuspended to $\text{OD}_{600} = 0.5$ in 150 mM NaCl. Cells were then diluted in a 2x concentration of Eu^{3+} (as EuCl_3) in 150 mM NaCl to achieve the final concentrations represented above. The pH of the resuspensions was measured using litmus paper to be $\text{pH} = 6.0\text{--}6.5$. Samples were then shaken continuously for 5 and 20 minutes, and at each time the luminescence reading was recorded. Readings are integrated emission intensities over 1000 μs following a 600 μs delay. Emission intensities were for $\lambda_{\text{Ex}} = 280$ nm, and $\lambda_{\text{Em}} = 615$ nm. Plotted values are the ratio of luminescence intensities RF2 6AW-containing cells to B95 cells at that Eu^{3+} concentration and are the mean of three biological replicates. Error bars are standard deviations.

Supplementary tables

Table S1. Extinction coefficients of the amino acids used in this study

Amino acid	λ_{ex} (nm)	Extension coefficient ^a (M ⁻¹ cm ⁻¹)
4-CN-Trp ^{a,b}	290	4708 ± 58 (at 290 nm)
		2304 ± 48 (at 280 nm)
6-Aza-Trp	265	3087 ± 8 (at 265 nm)
		1013 ± 4 (at 280 nm)
4-Aza-Trp	290	7335 ± 4 (at 290 nm)
		6272 ± 4 (at 280 nm)
<i>m</i> -cyano- pyridylalanine	265	1386 ± 9 (at 265 nm)
		255 ± 9 (at 280 nm)
<i>m</i> -cyano- phenylalanine	275	1059 ± 30 (at 275 nm)
		843 ± 19 (at 280 nm)

^a Extinction coefficients at 280 nm are useful for determining protein concentrations based on their amino acid sequence. For comparison, the ϵ_{280} value of tryptophan is 5422 M⁻¹ cm⁻¹.

^b Value from Abdelkader *et al.*¹⁴

Table S2. Variables used in integrated lifetime calculations^a

Protein	Ln³⁺	Y₀		Plateau		K		Integral		Error (%)
LBT	Tb	259.17	± 0.20	0.00	± 0.07	0.38	± 0.00	682.88	± 1.14	0.17
LBT	Eu	3.32	± 0.03	0.00	± 0.00	0.71	± 0.00	4.65	± 0.04	0.79
LBT mCNP	Tb	233.43	± 0.20	-0.15	± 0.07	0.37	± 0.00	630.27	± 1.17	0.19
LBT mCNP	Eu	253.05	± 0.18	-0.01	± 0.03	0.70	± 0.00	361.21	± 0.48	0.13
LBT mCNF	Tb	156.29	± 0.15	-0.14	± 0.05	0.37	± 0.00	423.25	± 0.85	0.20
LBT mCNF	Eu	35.33	± 0.06	0.00	± 0.01	0.70	± 0.00	50.39	± 0.16	0.32
LBT 4AW	Tb	468.61	± 0.92	2.16	± 0.18	0.64	± 0.00	728.16	± 2.67	0.37
LBT 4AW	Eu	61.71	± 0.09	0.01	± 0.02	0.72	± 0.00	85.77	± 0.24	0.28
LBT 6AW	Tb	402.22	± 0.30	-0.14	± 0.10	0.38	± 0.00	1072.23	± 1.71	0.16
LBT 6AW	Eu	268.10	± 0.24	0.01	± 0.04	0.71	± 0.00	377.86	± 0.62	0.16
LBT 4CNW	Tb	209.92	± 0.31	0.16	± 0.06	0.65	± 0.00	323.19	± 0.88	0.27
LBT 4CNW	Eu	3.77	± 0.02	-0.00	± 0.00	0.70	± 0.01	5.38	± 0.06	1.20
RF1	Tb	867.20	± 0.48	0.17	± 0.09	0.72	± 0.00	1207.33	± 1.24	0.10
RF2	Tb	1236.34	± 0.51	0.02	± 0.07	0.69	± 0.00	1783.83	± 1.20	0.07
RF2	Eu	0.67	± 0.02	-0.01	± 0.00	1.11	± 0.04	0.61	± 0.03	4.77
RF3	Tb	357.29	± 0.59	0.28	± 0.10	0.77	± 0.00	462.69	± 1.38	0.30
RF2 42 6AW ^b	Tb	1027.07	± 0.54	0.07	± 0.10	0.68	± 0.00	1507.04	± 1.46	0.10
RF2 42 6AW ^b	Eu	920.91	± 0.47	0.03	± 0.04	1.82	± 0.00	506.04	± 0.43	0.09
RF2 42 6AW ^b	Dy	1631.29	± 119.71	0.02	± 0.00	56.77	± 0.98	28.73	± 2.17	7.54
RF2 50 6AW ^b	Tb	1200.19	± 0.67	0.03	± 0.10	0.71	± 0.00	1700.41	± 1.60	0.09
RF2 50 6AW ^b	Eu	24.18	± 0.07	-0.02	± 0.01	1.88	± 0.01	12.84	± 0.06	0.48

^a Uncertainties were derived from the deviation of the best-fit equation to the data. Final values are rounded to whole numbers in the main text.

^b The notations used here are different from the main text to facilitate discrimination of the different 6AW positions.

Table S3. Amino acid sequences of the proteins used in this study

Protein	Amino acid sequence ^a	Nucleotide sequence
NT-LBT-Trp	MASMTGHHHHHSHTTPW TNPGLAENFMNSFMQGLSS MPGFTASQLDKMSTIAQSM VQSIQSLAAQGRTPNDLQA LNMAFASSMAEIAASEEGG GSLSTKTSSIASAMSNAFLQ TTGVVNQPFINEITQLVSMF AQAGMNDVSAGNSENLYF QGFDITNNDGWIEGDELLLE G	ATGGCAAGCATGACAGGTCATCATCAC CATCATCATAGCCATACCACACCGTGG ACCAATCCTGGTCTGGCAGAAAACCTT ATGAATAGCTTTATGCAGGGTCTGAGC AGCATGCCTGGTTTTACCGCAAGCCAG CTGGACAAAATGAGCACCATTGCACAG AGCATGGTTCAGAGCATTGAGCCTG GCAGCACAGGGTCGTACCAGTCCGAAT GATCTGCAGGCACTGAATATGGCATT GCAAGCAGCATGGCAGAAATTGCAGC AAGCGAAGAAGGTGGCGGTAGCCTGA GCACCAAACCAGCAGCATTGCAAGC GCAATGAGCAATGCATTTCTGCAGACA ACCGGTGTTGTTAATCAGCCGTTTATTA ACGAAATTACCCAGCTGGTTAGCATGT TTGCACAGGCAGGTATGAATGATGTTA GCGCAGGTAATAGCGAAAACCTGTATT TTCAGGGCTTTATCGATACCAATAACG ATGGTTGATTGAAGGTGATGAACTGC TGCTGGAAGGTAA
NT-LBT-Amb	MASMTGHHHHHSHTTPW TNPGLAENFMNSFMQGLSS MPGFTASQLDKMSTIAQSM VQSIQSLAAQGRTPNDLQA LNMAFASSMAEIAASEEGG GSLSTKTSSIASAMSNAFLQ TTGVVNQPFINEITQLVSMF AQAGMNDVSAGNSENLYF QGFDITNNDGZIEGDELLLE G	ATGGCAAGCATGACAGGTCATCATCAC CATCATCATAGCCATACCACACCGTGG ACCAATCCTGGTCTGGCAGAAAACCTT ATGAATAGCTTTATGCAGGGTCTGAGC AGCATGCCTGGTTTTACCGCAAGCCAG CTGGACAAAATGAGCACCATTGCACAG AGCATGGTTCAGAGCATTGAGCCTG GCAGCACAGGGTCGTACCAGTCCGAAT GATCTGCAGGCACTGAATATGGCATT GCAAGCAGCATGGCAGAAATTGCAGC AAGCGAAGAAGGTGGCGGTAGCCTGA GCACCAAACCAGCAGCATTGCAAGC GCAATGAGCAATGCATTTCTGCAGACA ACCGGTGTTGTTAATCAGCCGTTTATTA ACGAAATTACCCAGCTGGTTAGCATGT TTGCACAGGCAGGTATGAATGATGTTA GCGCAGGTAATAGCGAAAACCTGTATT TTCAGGGCTTTATCGATACCAATAACG ATGGTTAGATTGAAGGTGATGAACTGC TGCTGGAAGGTAA
EF2 RF1-Trp	MASMTGHHHHHHGGSSEN YFQKKVITISEDTPESFEVK MGKGLLKVTVPKTFIVED PKDGDWLDAKEIKALLAAA KAATGAKTIEIEDIPK	ATGGCAAGCATGACAGGTCATCATCAC CATCATCATGGTGGTAGCGAAAATCTG TATTTTCAGGGCAAAAAGGTGATTACC ATCAGCGAAGATACACCGGAAAGCTTT GAAGTGA AAAATGGGTAAAGGTCTGCTG AAAGTTACCGTTCGGGTAAAACCTTT ATTGTTGAAGATCCGGATAAAGATGGT TGGCTGGATGCCAAAGAAATTAAGCA CTGCTGGCAGCAGCAAAAAGCAGCAAC CGGTGCAAAAACCATTGAAATTGAAGA TATCCCGAAATAA

EF2 RF2-Trp	MASMTGHHHHHHGGSENLYFQRRRAYLLRVPDPTLVEISPEEAERLAKTRPVLEVEDPKDGDWLDAKERAWILEHLRETHPDASGAVIVVVVD	ATGGCAAGCATGACAGGTCATCATCACCATCATCATGGTGGTAGCGAAAATCTGTATTTTCAGGGTCGTCGTGCATATCTGTGCGTGTTGATCCGGATACTGGTTGAAATTAGTCCGGAAGAAGCAGAACGTCTGGCAAAAACCCGTCCGGTTCTGGAA GTTGAAGATCCTGATAAAGATGGTTGGCTGGATGCAAAAAGAACGTGCATGGATTCTGGAACATCTGCGTGAAACCCATCCTGATGCAAGCGGTGCAGTTATTGTTGTTGTTGACTAA
EF2 RF3-Trp	MASMTGHHHHHHGGSENLYFQKTTITTYKFATKEDEKKAKELDPKDGDWLDAKELK KLKKAGLKITKVTKTIK	ATGGCAAGCATGACAGGTCATCATCACCATCATCATGGTGGTAGCGAAAATCTGTATTTTCAGGGTAAAACCATCACCACCATAAAGTTTGCAACCAAAGAGGATGAGAAAAAGGCCAAAGAAGACTGGATCCGGA TAAAGATGGTTGGCTGGATGCAAAAAGA ACTGAAAAAGCTGAAGAAAGCCGGTCTGAAAATCACCAAAGTTAAAACCAAG ACCATTAAATAA
EF2 RF2-42Z	MASMTGHHHHHHGGSENLYFQRRRAYLLRVPDPTLVEISPEEAERLAKTRPVLEVEDPKDGDGZLDAKERAWILEHLRETHPDASGAVIVVVVD	ATGGCAAGCATGACAGGTCATCATCACCATCATCATGGTGGTAGCGAAAATCTGTATTTTCAGGGTCGTCGTGCATATCTGTGCGTGTTGATCCGGATACTGGTTGAAATTAGTCCGGAAGAAGCAGAACGTCTGGCAAAAACCCGTCCGGTTCTGGAA GTTGAAGATCCTGATAAAGATGGTTAGCTGGATGCAAAAAGAACGTGCATGGATTCTGGAACATCTGCGTGAAACCCATCCTGATGCAAGCGGTGCAGTTATTGTTGTTGTTGACTAA
EF2 RF2-50Z	MASMTGHHHHHHGGSENLYFQRRRAYLLRVPDPTLVEISPEEAERLAKTRPVLEVEDPKDGDWLDAKERAZILEHLRETHPDASGAVIVVVVD	ATGGCAAGCATGACAGGTCATCATCACCATCATCATGGTGGTAGCGAAAATCTGTATTTTCAGGGTCGTCGTGCATATCTGTGCGTGTTGATCCGGATACTGGTTGAAATTAGTCCGGAAGAAGCAGAACGTCTGGCAAAAACCCGTCCGGTTCTGGAA GTTGAAGATCCTGATAAAGATGGTTGGCTGGATGCAAAAAGAACGTGCATAGATTCTGGAACATCTGCGTGAAACCCATCCTGATGCAAGCGGTGCAGTTATTGTTGTTGTTGACTAA

^a Z indicates the positions of the amber sites. ^a TAG indicates the amber stop codon.

Table S4. Yields of different proteins co-expressed with G1(ncAA)RS in LB media with 1 mM ncAA^a

Amber site	G1(ncAA)RS	Yield (mg per 1 L cell culture)
NT-LBT	-	31 ^b
NT-LBT 4AW	4AWRS ¹⁵	16 ^c
NT-LBT 6AW	6AWRS ¹⁵	26 ^c
NT-LBT 4CNW	4CNWRS ¹⁴	20 ^c
NT-LBT mCNP	mCNPRS ¹⁶	10 ^c
NT-LBT mCNF	mCNPRS ¹⁶	15 ^c

EF2 RF1	-	50 ^{b,f}
EF2 RF2	-	>400 ^{b,f}
EF2 RF3	-	10 ^{b,f}
EF2 RF2 42 6AW	6AWRS	51 ^d
EF2 RF2 50 6AW	6AWRS	19 ^e

^a All proteins were expressed and purified as described in the methods section. Yields were calculated by measuring the protein concentration at a known volume and adjusting for expression volume. ^b From 1 L expression volume. ^c From 100 mL expression volume. ^d From 50 mL expression volume. ^e From 30 mL expression volume. ^f In Autoinduction media.

Table S5. EF2-RF variants with Tb³⁺

Protein	Intensity Tb*	Apparent K_D (nM)	Lifetime (τ_{Tb} , ms)
EF2 RF1	1207	320 ± 50	1.4
EF2 RF2	1734	20 ± 3	1.4
EF2 RF3	462	3300 ± 130	1.3

Table S6. Determination of q -values from D₂O titrations^a

Complex	τ_{H_2O}	τ_{D_2O}	q	Uncertainty on q
RF2 WT Tb	1.42 ms	2.34 ms	1.07	±0.3
RF2 42 6AW Eu	0.54 ms	1.42 ms	1.06	±0.3
RF2 42 6AW Sm	12.12 μ s	48.64 μ s	1.30	±0.5
RF2 42 6AW Dy	14.18 μ s	35.63 μ s	1.31	±0.5

^a The q values were determined using the modified Horrocks equation as described above. Uncertainties are from errors in the empirical parameters used to determine the q value.

References

- Mandal, P.; Kretschmar, J.; Drobot, B. Not just a background: pH buffers do interact with lanthanide ions—A Europium(III) case study. *J. Biol. Inorg. Chem.* **2022**, *27*, 249–260.
- Abdelkader, E. H.; Qianzhu, H.; Tan, Y. J.; Adams, L. A.; Huber, T.; Otting, G. Genetic encoding of N⁶-(((trimethylsilyl)methoxy)carbonyl)-L-lysine for NMR studies of protein–protein and protein–ligand interactions. *J. Am. Chem. Soc.* **2021**, *143*, 1133–1143.

3. Mukai, T.; Hoshi, H.; Ohtake, K.; Takahashi, M.; Yamaguchi, A.; Hayashi, A.; Yokoyama, S.; Sakamoto, K. Highly reproductive *Escherichia coli* cells with no specific assignment to the UAG codon. *Sci. Rep.* **2015**, *5*, 9699.
4. The PyMOL Molecular Graphics System, Version 3.0, Schrödinger, LLC.
5. Cook, E. C.; Featherston, E. R.; Showalter, S. A.; Cotruvo, J. A. Structural basis for rare earth element recognition by lanmodulin. *Biochemistry* **2019**, *58*, 120–125.
6. Chaudhury, S.; Lyskov, S.; Gray, J. J. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **2010**, *26*, 689–691.
7. Park, H.; Bradley, P.; Greisen, P., Jr.; Liu, Y.; Mulligan, V. K.; Kim, D. E.; Baker, D.; DiMaio, F. Simultaneous optimization of biomolecular energy functions on features from small molecules and macromolecules. *J. Chem. Theory Comput.* **2016**, *12*, 6201–6212.
8. Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Hanikel, N.; Pellock, S. J.; Courbet, A.; Sheffler, W.; Wang, J.; Venkatesh, P.; Sappington, I.; Torres, S. V.; Lauko, A.; De Bortoli, V.; Mathieu, E.; Ovchinnikov, S.; Barzilay, R.; Jaakkola, T. S.; DiMaio, F.; Baek, M.; Baker, D. De novo design of protein structure and function with RFdiffusion. *Nature* **2023**, *620*, 1089–1100.
9. Bünzli, J.-C. G. Lanthanide luminescence for biomedical analyses and imaging. *Chem. Rev.* **2010**, *110*, 2729–2755.
10. Kimura, T.; Kato, Y. Luminescence study on determination of the hydration number of Sm(III) and Dy(III). *J. Alloys Compd.* **1995**, *225*, 284–287.
11. Beeby, A.; Clarkson, I. M.; Dickins, R. S.; Faulkner, S.; Parker, D.; Royle, L.; de Sousa, A. S.; Williams, J. A. G.; Woods, M. Non-radiative deactivation of the excited states of europium, terbium and ytterbium complexes by proximate energy-matched OH, NH and CH oscillators: An improved luminescence method for establishing solution hydration states. *J. Chem. Soc., Perkin Trans. 2* **1999**, 493–504.
12. Klopp, J.; Winterhalter, A.; Gebleux, R.; Scherer-Becker, D.; Ostermeier, C.; Gossert, A. D. Cost-effective large-scale expression of proteins for NMR studies. *J. Biomol. NMR* **2018**, *71*, 247–262.
13. Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **2009**, *44*, 213–223.
14. Qianzhu, H.; Abdelkader, E. H.; Welegedara, A. P.; Habel, E.; Paul, N.; Frkic, R. L.; Jackson, C. J.; Huber, T.; Otting, G. Rendering proteins fluorescent inconspicuously: Genetically encoded 4-cyanotryptophan conserves their structure and enables the detection of ligand binding sites. *Angew. Chem. Int. Ed.* **2025**, *64*, e202421000.
15. Abdelkader, E. H.; Qianzhu, H.; Huber, T.; Otting, G. Genetic incorporation of azatryptophans enhances PETase activity. *submitted for publication*.
16. Abdelkader, E. H.; Qianzhu, H.; George, J.; Frkic, R. L.; Jackson, C. J.; Nitsche, C.; Otting, G.; Huber, T. Genetic encoding of cyanopyridylalanine for in-cell protein macrocyclization by the nitrile–aminothiols click reaction. *Angew. Chem. Int. Ed.* **2022**, *61*, e202114154.