Engineered Recognition of Tetravalent Zirconium and Thorium by Chelator-Protein Systems: Toward Flexible Radiotherapy and Imaging Platforms

Ilya Captain,[†] Gauthier J.-P. Deblonde,[†] Peter B. Rupert,[‡] Dahlia D. An,[†] Marie-Claire Illy,[†] Emeline Rostan,[†] Corie Y. Ralston,[§] Roland K. Strong,[,‡] and Rebecca J. Abergel*[,†]*

[†]Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

[‡]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

[§]Berkeley Center for Structural Biology, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Supporting Information

Experimental Details

General Considerations. ^{1}H NMR spectra were recorded on Bruker instruments; ^{13}C NMR spectra were recorded on Bruker instruments with tetramethylsilane as an internal reference. SilicaFlash G60 (particle size 60-200 μm) was used for flash column chromatography. LC-MS was performed on an Agilent LC/MS system consisting of an Agilent 1200 binary LC pump, a temperature-controlled autosampler, a PDA UV detector, and a 6530 Accurate Mass Q-TOF mass spectrometer (Wilmington, DE, USA). The mass spectrometer was equipped with a JetStream® ESI probe operating at atmospheric pressure. The ESI source parameter settings were: mass range m/z 100–1000, gas temperature 350°C, gas flow 10 L/min, nebulizer 50 psi, sheath gas temperature 400°C, sheath gas flow 12 L/min, capillary voltage (Vcap) 3500 V, nozzle voltage 500 V, fragmentor 200 V, skimmer 65 V, octopole RF (OCT 1 RF Vpp) 750 V. Reverse phase preparatory HPLC was performed on a Varian ProStar system with a Vydac C18 column. High-resolution mass spectra were acquired using a Waters Xevo G2 QTof mass spectrometer. Absorption spectra were recorded on a Varian Cary G5 double beam absorption spectrometer or a NanoDrop 2000C, using quartz cells of 10 and 2 mm path lengths, respectively.

Methyl 2,3-dihydroxybenzoate (2). A stirred suspension of **1** (8.06 g, 52.3 mmol) in 100 mL of MeOH was treated with 2.00 ml of concentrated sulfuric acid. The suspension warmed and clarified 2 minutes after the addition. The reaction was equipped with a reflux condenser and was heated to 65°C overnight. The next morning the conversion was verified by LC-MS and the volatiles were removed under reduced pressure. The crude was partitioned between H_2O (100 mL) and ethyl acetate (100 mL) and the aqueous layer was extracted with ethyl acetate (3x50 mL). The organic extracts were combined, dried over $MgSO_4$, and concentrated under reduced pressure. The crude was passed through a plug of silica using 10% ethyl acetate in hexanes as eluent. The eluent was concentrated under reduced pressure and dried under high vacuum for 2 hours to yield **2** (7.66 g, 45.6 mmol, 88%) as a white solid, the spectral properties of which matched previous reports.[1]

Methyl 2,2-diphenylbenzo[d][1,3]dioxole-4-carboxylate (3). Precursor **2** (5.00 g, 29.7 mmol) was mixed with dichlorodiphenylmethane (8.56 mL, 44.6 mol) under an argon atmosphere; the resulting suspension was stirred and heated to 160°C for 1 hour. The mixture was allowed to cool to room temperature and was diluted with 100 mL of ethyl acetate. The solution was washed with sat. $NaHCO_3$ (30 mL), brine (30 mL), dried over $MgSO_4$, and then concentrated under reduced pressure. The ensuing greyish oil was dissolved in 30 mL of hot MeOH (65°C) and was slowly cooled to 5°C, which resulted in the formation of white crystals. The crystals were a mixture of **3** and benzophenone that could not be easily separated; the crude product was used as is for the subsequent step.

2,2-diphenylbenzo[d][1,3]dioxole-4-carboxylic acid (4). The mixture from the previous step was dissolved in 100 mL of THF and was treated with 100 mL of 0.9 M LiOH. The emulsion was rapidly stirred and heated to reflux for 5 hours. Conversion was verified by LC-MS and the reaction was cooled to room temperature. The solution was neutralized with 10% v/v aqueous acetic acid and was extracted with ethyl acetate (3x50 mL). The organic extracts were combined, dried over $MgSO_4$, and concentrated under reduced pressure. The crude was chromatographed using 25% ethyl acetate in hexanes as eluent.

Volatiles were then removed under reduced pressure followed by high vacuum to yield **4** (7.6 g, 24.06 mmol, 81% over 2 steps) as a white solid, the spectral properties of which matched previous reports.[1]

3,4,3-LI(2,2-diphenylbenzo[d][1,3]-2,3-catecholamide) (5). Precursor **4** (746 mg, 2.33 mmol) was suspended in 10 mL of dry toluene under an argon atmosphere and was treated with oxalyl chloride (220 µL, 2.55 mmol). Catalytic N,N-dimethylformamide was added and the suspension was heated to 40°C. The solution was stirred until the evolution of gas ceased and was concentrated on the manifold vacuum at the same temperature. The resulting brown oil was dissolved in 10 mL of dry THF. In a separate container a solution of spermine (118 mg, 0.583 mmol), triethylamine (356 µL, 2.56 mmol), and THF (5 mL) was prepared. The solutions were combined and heated to 50°C overnight in a sealed flask. The following day the reaction was filtered and concentrated under reduced pressure. The resulting crude oil was chromatographed using 3% MeOH in CH_2Cl_2 as eluent. The volatiles were then removed under reduced pressure and dried under vacuum, yielding **5** as a white foam (641 mg, 0.457 mmol, 78 % yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.88 (1H, t, J = 5.7 Hz), 7.66-7.76 (6H, br t), 7.60 (1H, br s), 7.57 (1H, br s), 7.43-7.53 (10H, br s), 7.33-7.40 (4H, br s), 7.19-7.31 (20H, br s), 7.01 (2H, d, J = 7.6 Hz), 6.91 (4H, dd, J = 12.1 Hz, 8.0 Hz), 6.80 (2H, br s), 6.72 (2H, br s), 3.85 (4H, br s), 3.43 (2H br s), 3.21 (2H, br s), 3.06 (1H, br s), 2.96 (1H, br s), 2.80 (2H, br s), 1.81 (4H, br s), 1.54 (1H, br s), 1.43 (1H, br s), 1.19 (1H, br s), 0.89 (2H, br s). ^{13}C NMR (75 MHz, $CDCl_3$) δ 167.5, 163.7, 147.3, 147.1, 145.0, 142.8, 139.7, 139.4, 138.9, 129.7, 129.2, 128.4, 128.3, 126.4, 126.3, 126.1, 126.0, 122.3, 122.2, 121.7, 120.4, 118.4, 118.1, 116.0, 111.8, 111.4, 111.3, 109.4, 47.9, 41.8, 36.5, 27.9, 25.5 (Fig. S4).

3,4,3-LI(CAM) (6). The protected ligand **5** (411 mg, 2.93×10^{-4} mol) was dissolved in a mixture of 5 mL acetic acid, 0.5 mL H_2O, and 0.1 mL concentrated HCl. The solution was stirred in a sealed container at 60°C overnight. The next day the conversion was confirmed by LC-MS and the volatiles were removed under vacuum. A portion of the crude was purified using reverse-phase prep-HPLC using at 10→50% MeOH in H_2O + 0.1 % trifluoroacetic acid as eluent. The solvent was removed on a Genevac centrifugal evaporator followed by lyophilization of residual H_2O. CAM was obtained as a pure white powder (~90% yield). ^{1}H NMR (600 MHz, DMSO-d_6) δ 12.82 (1H, br s), 12.69 (1H, br s), 9.52 (2H, br s), 9.11 (2H, br s), 8.78 (1H, br s), 8.60 (3H, br s), 7.26 (1H, br s), 7.12 (1H, br s), 6.90 (2H, br s), 6.77 (1H, br s), 6.66 (4H, br s), 6.56 (2H, br s), 6.44 (1H, br s), 2.88-3.52 (12H, overlapping aliphatic signals), 1.16-1.83 (8H, overlapping aliphatic signals); ^{13}C NMR (125 MHz, MeOD-d_4) δ 172.9, 171.5, 150.4, 147.3, 146.6, 125.6, 125.4, 121.0, 119.6, 119.1, 118.8, 118.6, 116.9, 116.6, 47.7, 44.9, 43.2, 37.8, 37.5, 29.3, 28.2, 26.5, 25.5 (Fig. S5). MS-ESI (m/z) [M + H] Calcd for $C_{38}H_{43}N_4O_{12}$, 747.2878; found 747.2922 and [M – H] Calcd. for $C_{38}H_{41}N_4O_{12}$, 745.2721; found 745.2774 (Fig. S6).

Solution Thermodynamics. All titrant solutions were degassed by boiling for 1 h while being purged under Ar. Carbonate-free 0.1 M KOH was prepared from concentrate (J.T Baker Dilut-It) and was standardized by titrating against 0.1 M potassium hydrogen phthalate (99.95%, Sigma Aldrich). Solutions of 0.1 M HCl were similarly prepared and were standardized by titrating against TRIS (99.9%, J.T. Baker). The glass electrode (Metrohm - Micro Combi - response to $[H^+]$) used for the pH measurements was calibrated at 25.0°C and at an ionic strength of 0.1 M (KCl) before each potentiometric or spectrophotometric titration. The calibration data were analyzed using the program GLEE[2] to refine for the E° and slope. All thermodynamic measurements were conducted at 25.0°C, in 0.1 M KCl supporting

electrolyte under positive Ar gas pressure. The automated titration system was controlled by a *867 pH Module* (Metrohm). Two-milliliter *Dosino 800* burets (Metrohm) dosed the titrant (0.1 M KOH or 0.1 M HCl) into the thermostated titration vessel (5−90 mL). UV−visible spectra were acquired with an Ocean Optics USB4000-UV−vis spectrometer equipped with a TP-300 dip probe (Ocean Optics; path length of 10 mm), fiber optics and a DH-2000 light source (deuterium and halogen lamps). The fully automated titration system and the UV-vis spectrophotometer were coordinated by *LBNL titration system*, a computer program developed in house.

Incremental Spectrophotometric Titrations. This method was used to determine the protonation constants of 3,4,3-LI(CAM) as well as the stability constants of its complexes formed with Eu(III), Zr(IV) and ^{232}Th(IV). The experimental titration setup is similar to previously described systems.[3] For the 3,4,3-LI(CAM) protonation (and Eu(III)-3,4,3-LI(CAM) complexes), titrations were performed with an initial concentration of 50 µM of 3,4,3-LI(CAM) (and 50 µM of Eu(III)) resulting in absorbance values comprised between 0 and 1.0 throughout the titration. Typically, 9 mL of a sample containing 3,4,3-LI(CAM) (and Eu(III)) and the supporting electrolyte (KCl/HCl) were incrementally perturbed by addition of 0.025 mL of carbonate-free 0.1 M KOH followed by a time delay of 80 s. Buffering of the solution was ensured by the addition of 10 mM of HEPES, 10 mM of CHES and 10 mM of MES. Between 130 and 250 data points were collected per titration, each data point consisting of a pH measurement and a UV-Vis spectrum (250-450 nm) over the pH range 1.5 to 12.0. All spectra were corrected for dilution before data fitting. The entire procedure (electrode calibrate, titration and data treatment) was performed independently five times for the protonation constants and four times for the Eu(III)-3,4,3-LI(CAM) complexes. For the Zr(IV) and Th(IV) complexes, titrations were performed similarly but in the presence of DTPA to avoid the formation of metal hydroxides at low pH, before the uptake by 3,4,3-LI(CAM). For each metal, three titrations were performed independently in the presence of 1.1 to 40 equivalents of DTPA. Examples of titrations are displayed in the Supporting Information (Fig. S1-S3).

Data Treatment. Thermodynamic constants and spectral deconvolution were refined using the nonlinear least-squares fitting program HypSpec.[4] All equilibrium constants were defined as cumulative formation constants, β_{mlh} according to Equation (1), where the metal and ligand are designated as M and L, respectively. All metal and ligand concentrations were held at estimated values determined from the volume of standardized stock solutions. All species formed with 3,4,3-LI(CAM) were considered to have significant absorbance to be observed in the UV-vis spectra and were therefore included in the refinement process. The refinements of the overall formation constants β included in each case with previously determined ligand protonation constants and the metal hydrolysis products, whose equilibrium constants were fixed to the literature values.[5] The speciation diagrams were calculated using the modeling program Hyss.[6] Errors on log β_{mlh} and pKa values presented in this paper correspond to the standard deviation observed over the n replicates (n = 3 to 5) of the entire procedure (electrode calibrate, titration and data treatment).

$$m\mathrm{M} + l\mathrm{L} + h\mathrm{H} \rightleftharpoons [\mathrm{M}_m\mathrm{L}_l\mathrm{H}_h];\ \beta_{mlh} = \frac{[\mathrm{M}_m\mathrm{L}_l\mathrm{H}_h]}{[\mathrm{M}]^m[\mathrm{L}]^l[\mathrm{H}]^h} \qquad (1)$$

Fluorescence Quenching Binding Assay. Equimolar amounts of metal and chelator were used to constitute metal-chelator solutions (2 μM, pH 7.4, 5% DMSO) in Tris-buffered saline (TBS). Then, a solution of recombinant wild-type Scn (50 nM, 3 mL, 10 μg/mL ubiquitin, TBS pH 7.4, 5% DMSO) was titrated with the metal-chelator solution. Fluorescence quenching of Scn was measured after each titrant addition on a HORIBA Jobin Yvon IBH FluoroLog-3 spectrofluorimeter, with 3 nm slit band-pass, using the characteristic excitation and emission wavelengths λ_{exc} = 280 and λ_{em} = 320-360 nm. The intrinsic fluorescence in proteins is generally attributed to tryptophan residues; two residues W31 and W79 are found in the proximity of the Scn binding site. Fluorescence values were corrected for dilution upon addition of titrant. Fluorescence data were analyzed by nonlinear regression analysis of fluorescence response versus ligand concentration using a one-site binding model as described elsewhere.[7] The K_d values are the results of at least three independent titrations were determined according to Equation (2). Control experiments were performed with $[Fe^{III}(Ent)]^{3-}$ to ensure the stability of the protein under experimental conditions.

$$\mathrm{Scn} + \mathrm{ML} \leftrightarrow \mathrm{Scn:ML};\ \frac{1}{K_d} = \frac{[\mathrm{Scn:ML}]}{[\mathrm{Scn}][\mathrm{ML}]} \qquad (2)$$

Crystallography. For crystallization, 1 mM solutions of equimolar metal/chelator complexes (prepared as above) were mixed in a 2:1 molar ratio with Scn, which was then buffer-exchanged into 25 mM PIPES (pH = 7.0), 150 mM NaCl, 1 mM EDTA, and 0.01% w/w NaN_3, and concentrated to ~10 mg/ml protein. Diffraction-quality crystals were grown by vapor diffusion from drops containing 1 μl of ternary metal-chelator-protein complex plus 1 μl of well solution (50 mM NaCl, 200 mM Li_2SO_4, 100 mM NaOAc (pH = 4.3-4.5), 1.2-1.4 M $(NH_4)_2SO_4$). Crystals were cryo-preserved by transfer to 50 mM NaCl, 200 mM Li_2SO_4, 100 mM NaOAc (pH=4.3-4.5), 1.2 M $(NH_4)_2SO_4$, and 20% v/v glycerol. Diffraction data were collected on beamline 5.0.2 at the Advanced Light Source (ALS, Berkeley, CA). Diffraction data were integrated and scaled with HKL-2000.[8] Initial phases were determined by rigid body positional refinement with Refmac[9] using 3FW5.pdb as a starting structure, or molecular replacement with MolRep[10] using 3FW5.pdb as a search model. Structures were refined through iterative rounds of positional refinement using Refmac[9] alternating with model building using COOT,[11] followed by a final round of TLS refinement.[12] Residues or side-chains that did not exhibit clear electron density in $2F_{obs}$-F_{calc} Fourier syntheses when contoured at 0.7σ were removed or truncated to the Cβ atom. The quality of the final model was assessed using ProCheck[13] and Molprobity.[14] Crystallographic statistics are reported in Supporting Information Data Table S1. Final models have been deposited in the PDB.[15]

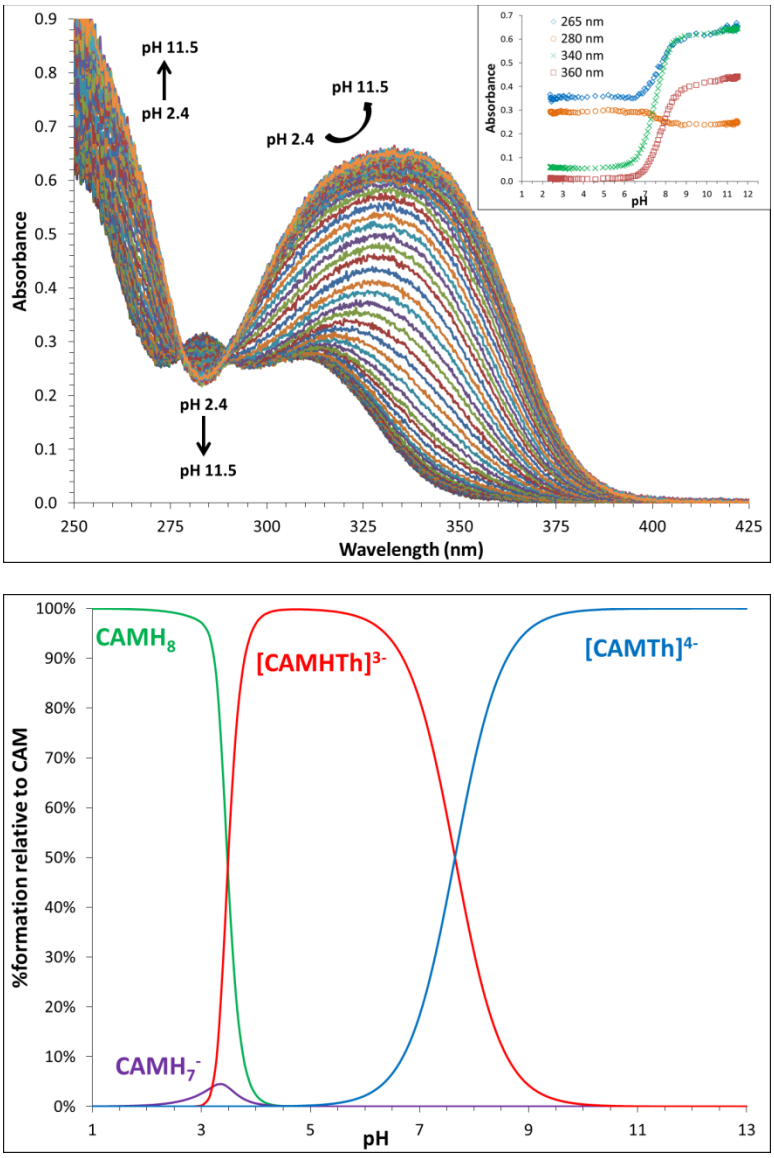

Figure S1. Top: Example of spectrophotometric competition titration of **Th(IV)-CAM** complexes. Starting conditions: 50 µM 3,4,3-LI(CAM), 50 µM Th(IV), 112 µM DTPA, 3 mM CHES, 3 mM TRIS, 3 mM MES, 10 mM HCl. I = 0.1 M (KCl). T = 25 °C. 130 spectra measured between pH 2.4 and 11.5. Path length = 10 mm. Spectra corrected for dilution. Inset: Change in absorbance 360 nm (red squares), 340 nm (green crosses), 280 nm (orange circles) and 265 nm (blue lozenges) as a function of pH. **Bottom:** Speciation diagram of the 3,4,3-LI(CAM) ligand in the presence of Th(IV). [Th] = ^{16} = 10 µM. T = 25°C, I = 0.1 M. Species: $CAMH_8$ (green), $CAMH_7^-$ (purple), $[CAMHTh]^{3-}$ (red) and $[CAMTh]^{4-}$ (blue). Calculations performed with Hyss software.

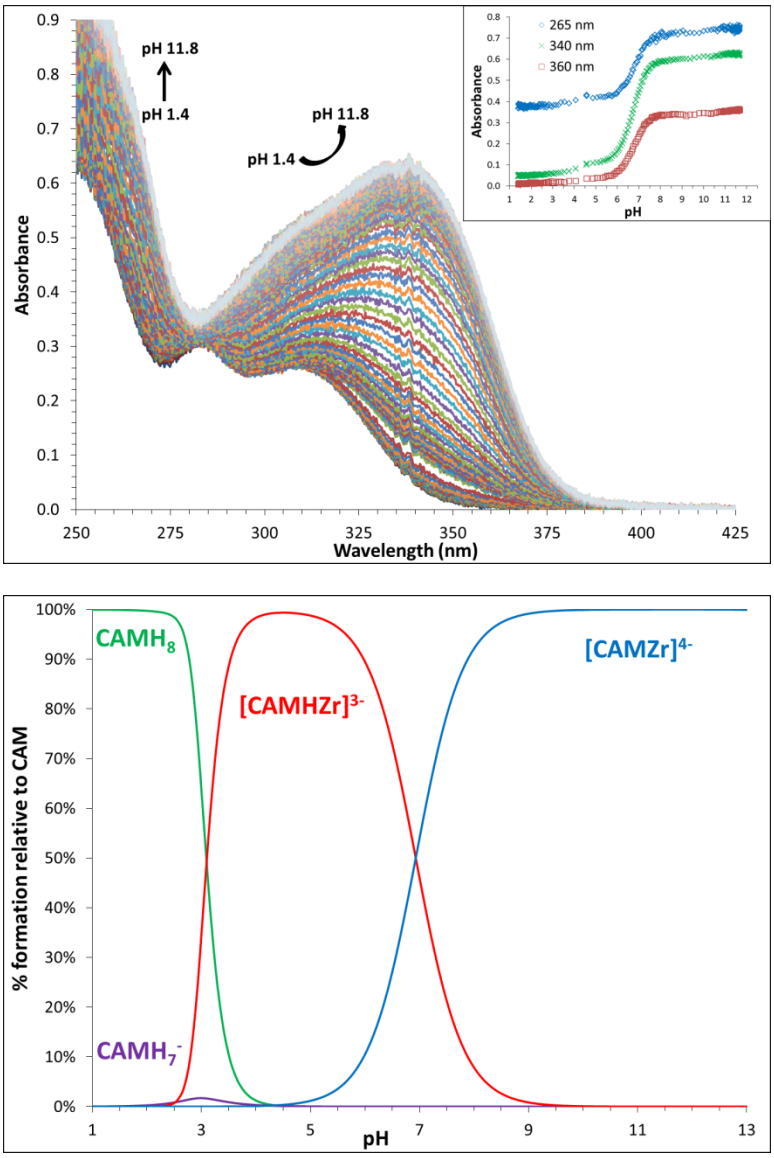

Figure S2. Top: Example of spectrophotometric competition titration of **Zr(IV)-CAM** complexes. Starting conditions: 50 µM 3,4,3-LI(CAM), 50 µM Zr(IV), 56 µM DTPA, 5 mM CHES, 5 mM TRIS, 5 mM MES, 45 mM HCl. I = 0.1 M (KCl). T = 25 °C. 230 spectra measured between pH 1.4 and 11.8. Path length = 10 mm. Spectra corrected for dilution. Inset: Change in absorbance 360 nm (red squares), 340 nm (green crosses) and 265 nm (blue lozenges) as a function of pH. **Bottom:** Speciation diagram of the 3,4,3-LI(CAM) ligand in the presence of Zr(IV). [Zr] = [16] = 10 µM. T = 25°C, I = 0.1 M. Species: $CAMH_8$ (green), $CAMH_7^-$ (purple), $[CAMHZr]^{3-}$ (red) and $[CAMZr]^{4-}$ (blue). Calculations performed with Hyss software.

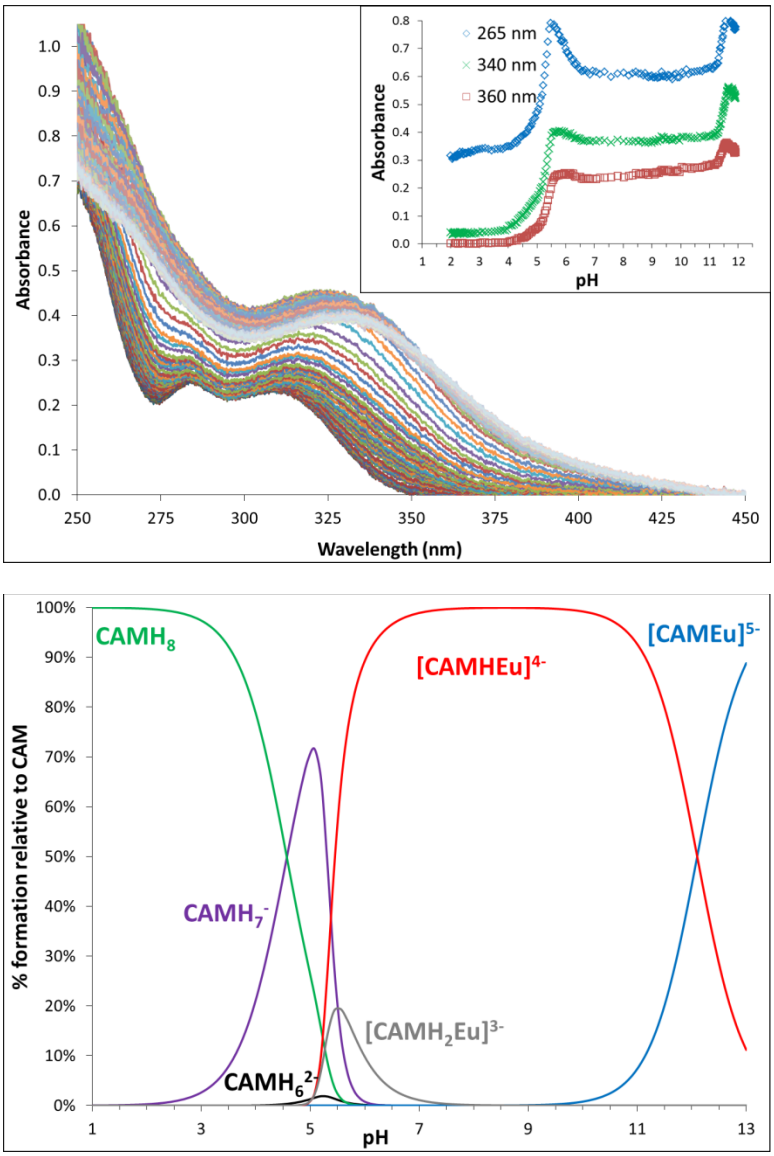

Figure S3. Top: Example of spectrophotometric competition titration of **Eu(IV)-CAM** complexes. Starting conditions: 50 µM 3,4,3-LI(CAM), 50 µM Eu(IV), 10 mM CHES, 10 mM MES, 10 mM acetic acid, 10 mM HCl. I = 0.1 M (KCl). T = 25 °C. 215 spectra measured between pH 2.0 and 11.9. Path length = 10 mm. Spectra corrected for dilution. Inset: Change in absorbance 360 nm (red squares), 340 nm (green crosses) and 265 nm (blue lozenges) as a function of pH. **Botttom:** Speciation diagram of the 3,4,3-LI(CAM) ligand in the presence of Eu(III). [Eu] = ^{16} = 10 µM. T = 25°C, I = 0.1 M. Species: $CAMH_8$ (green), $CAMH_7^-$ (purple), $CAMH_6^{2-}$ (black), $[CAMH_2Eu]^{3-}$ (grey), $[CAMHEu]^{4-}$ (red) and $[CAMEu]^{5-}$ (blue). Calculations performed with Hyss software.

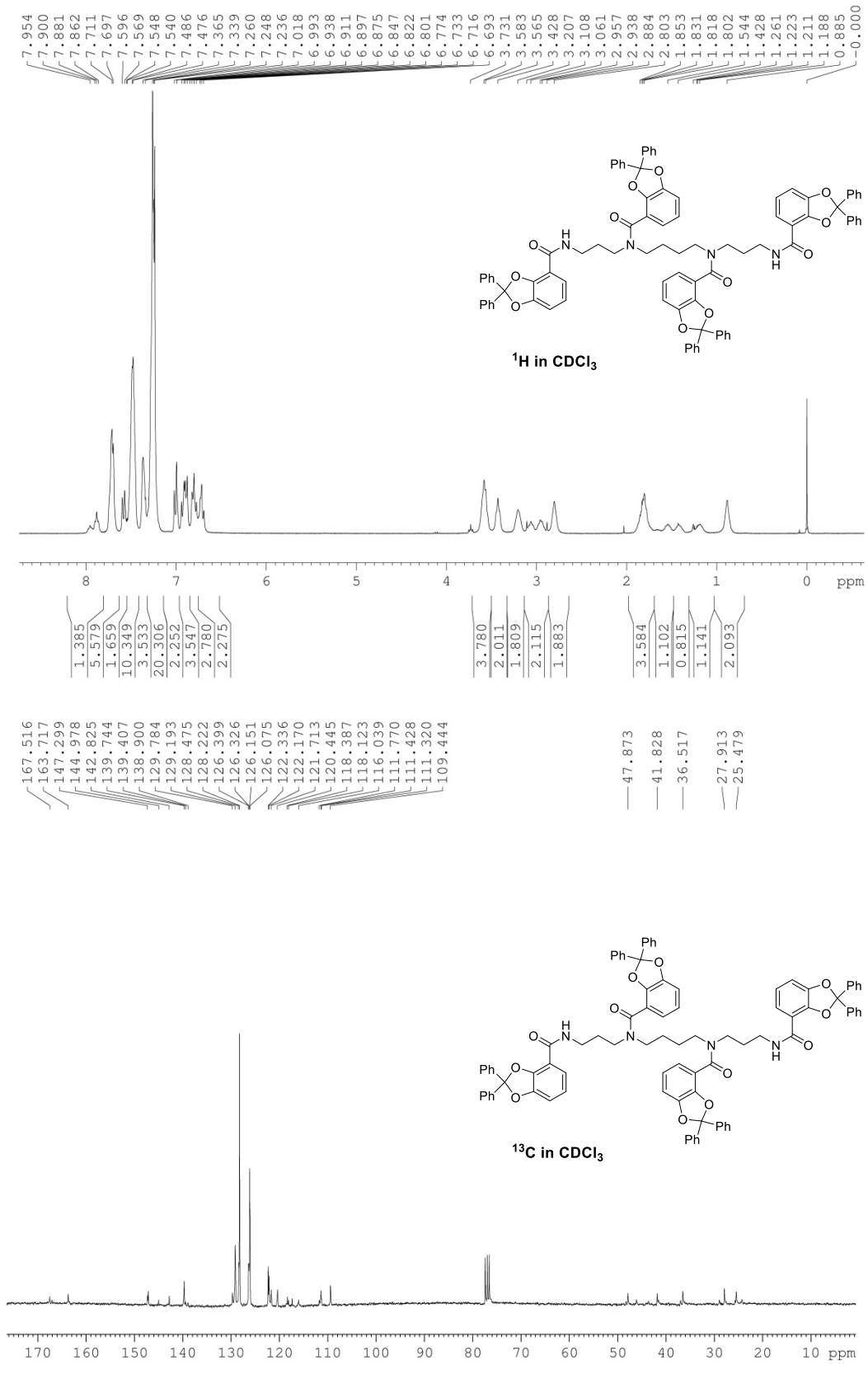

Figure S4. 3,4,3-LI(2,2-diphenylbenzo[d][1,3]-2,3-catecholamide) (**5**) – NMR Spectra.

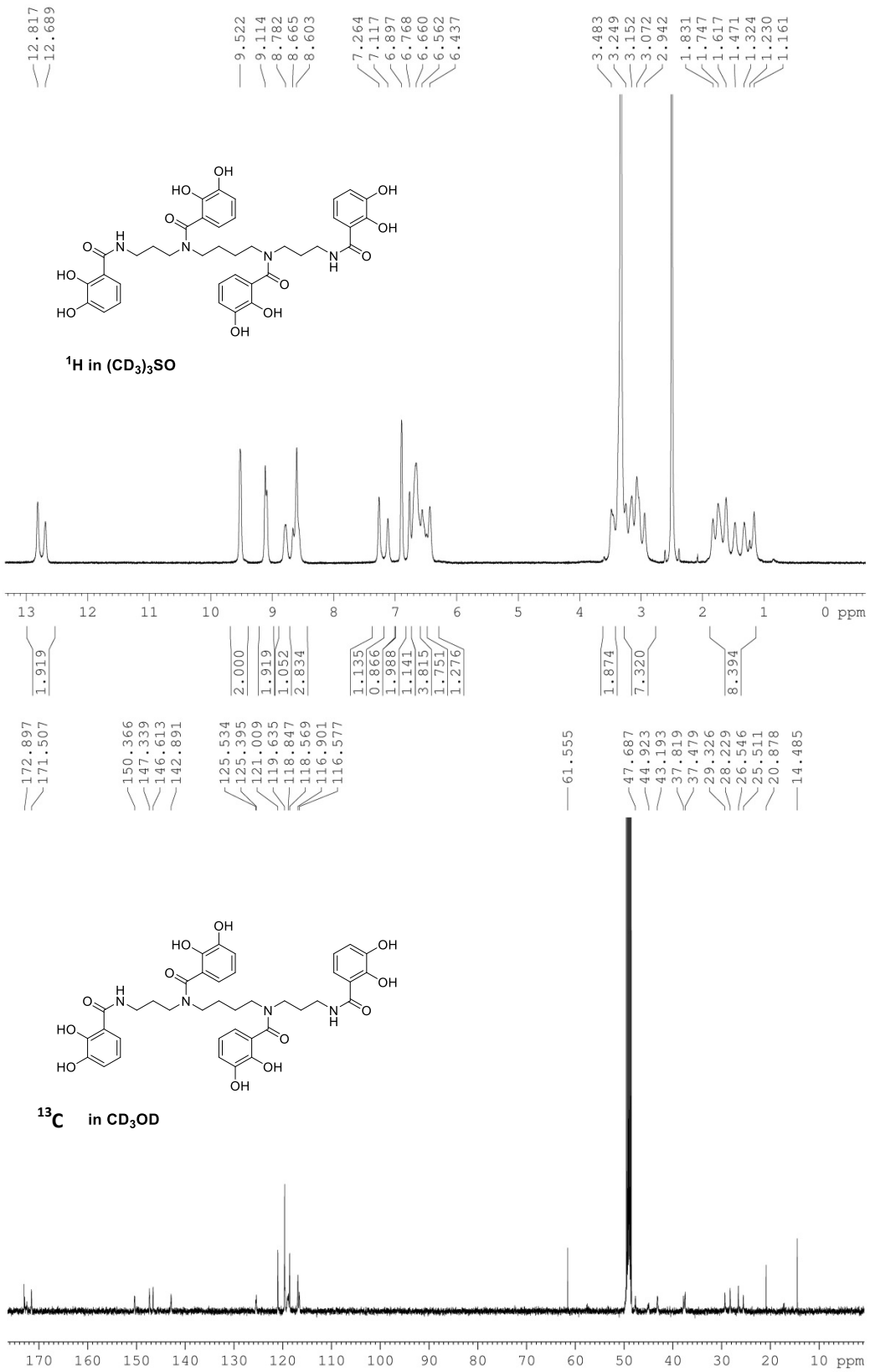

Figure S5. 3,4,3-LI(CAM) (**6**) – NMR Spectra.

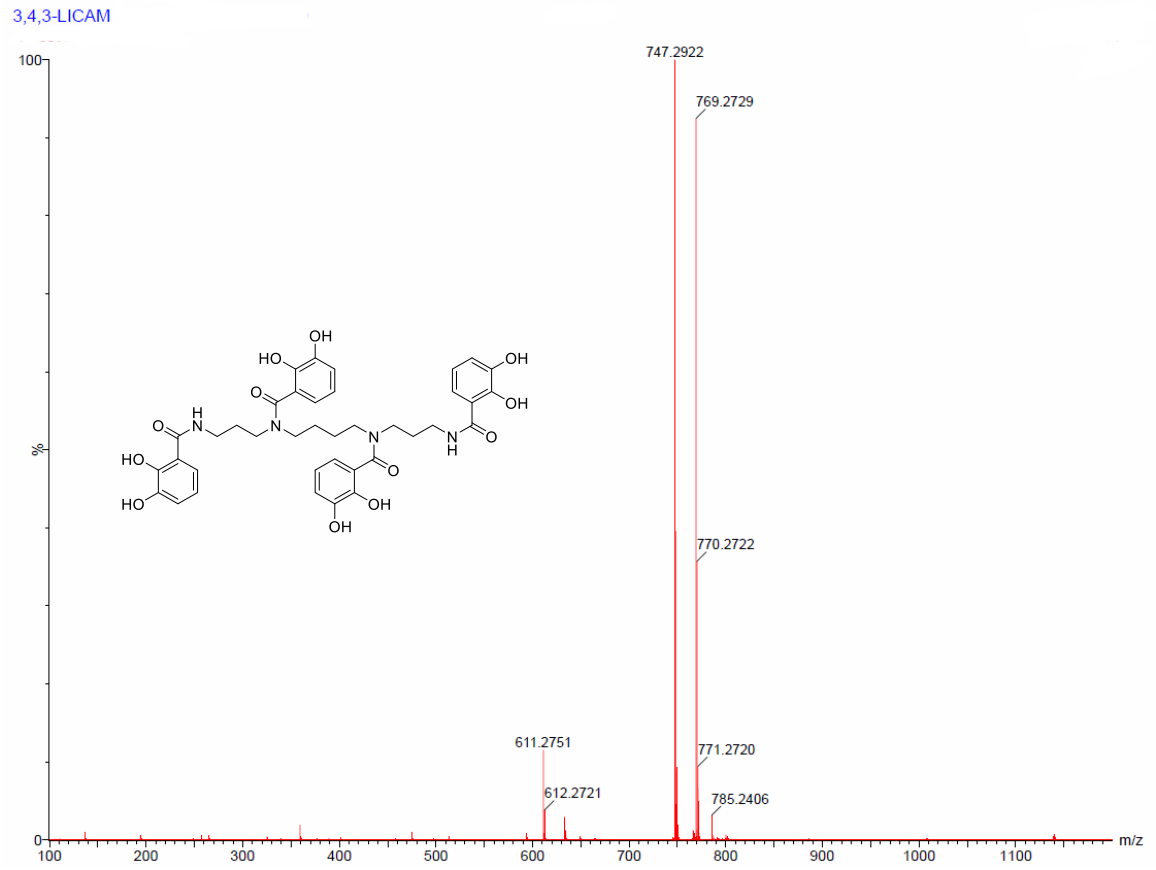

Figure S6. 3,4,3-LI(CAM) (**6**) – Mass Spectrum, Positive Mode**.**

Table S1. Crystallography data collection and refinement statistics.

Crystallization:		
Ligand	Zr-3,4,3-LI(CAM)	Th-3,4,3-LI(CAM)
Crystallization method	hanging drop	hanging drop
Crystallization conditions	1.2-1.4 M $(NH_4)_2SO_4$	1.2-1.4 M $(NH_4)_2SO_4$
	200 mM Li_2SO_4	200 mM Li_2SO_4
	100 mM NaAcetate	100 mM NaAcetate
	50 mM NaCl	50 mM NaCl
	pH = 4.1-4.3	pH = 4.1-4.3
Space group	$P4_12_12$	$P4_12_12$
Cell constants (Å)	a = b = 114.4	a = b = 115.5
	c = 117.1	c = 118.0
Data Collection:		
Cryopreservative	+ 20% v/v glycerol	+ 20% v/v glycerol
Beamline (Advanced Light Source in Berkeley, CA)		
	5.0.2	5.0.2
Wavelength (Å)	1.0000	1.0000
Resolution Range (Å)	50.0-2.65 (2.70-2.65)	50.0-2.15 (2.19-2.15)
Unique Reflections	22217 (1107)	42554 (2103)
Average Redundancy	8.1 (8.1)	8.3 (6.5)
R_{merge} (%)	11.0 (50.4)	7.5 (51.9)
I/σ(I)	28.2 (3.8)	25.7 (4.5)
Structure Refinement:		
Resolution (Å)	50.0-2.65	50.0-2.15
Number of reflections		
all / test	20976 / 1219	39580 / 2017
Phasing method	molecular replacement	molecular replacement
Search model	1L6M.pdb	1L6M.pdb
R_{cryst} / R_{free}	20.8 / 22.8	23.4 / 26.6
No. of non-hydrogen atoms (average B-factor ($Å^2$))		
Protein	4096 (58)	4147 (34)
Ligands	39 (48)	39 (59)
solvent	99 (48)	111 (58)
Rmsd		
Bonds (Å) / Angles (°)	0.01 / 1.24	0.01 / 1.48
Estimated coordinate error (Å)		
Maximum likelihood e.s.u.	0.196	0.157
Ramachandran values (MolProbity)		
Favored region (%)	96.1	98.2
Allowed region (%)	100.0	100.0
Outlier region (%)	0	0
MolProbity Score	1.06	0.88
PDB accession code	5KHP.pdb	5KID.pdb

Note: Numbers in parentheses are for reflections in the highest resolution shell.

1. F. L. Weitl and K. N. Raymond, *J. Am. Chem. Soc.*, 1980, **102**, 2289-2293.
2. P. Gans and B. O'Sullivan, *Talanta*, 2000, **51**, 33-37.
3. M. Sturzbecher-Hoehne, G.-P. Deblonde and R. Abergel, *Radiochimica Acta International journal for chemical aspects of nuclear science and technology*, 2013, **101**, 359-366.
4. P. Gans, A. Sabatini and A. Vacca, *Talanta*, 1996, **43**, 1739-1753.
5. R. Smith, A. Martell and R. Motekaitis, *NIST Critically selected stability constants of metal complexes database ver*, 2004, **2**.
6. L. Alderighi, P. Gans, A. Ienco, D. Peters, A. Sabatini and A. Vacca, *Coordination Chemistry Reviews*, 1999, **184**, 311-318.
7. B. E. Allred, P. B. Rupert, S. S. Gauny, D. D. An, C. Y. Ralston, M. Sturzbecher-Hoehne, R. K. Strong and R. J. Abergel, *Proc. Natl. Acad. Sci.*, 2015, **112**, 10342-10347.
8. Z. Otwinowski and W. Minor, in *Methods in Enzymology*, eds. C. W. Carter, Jr. and R. M. Sweet, Academic Press, New York, 1997, vol. 276, pp. 307-326.
9. G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr D Biol Crystallogr* 1997, **53**, 240-255.
10. A. A. Vagin and A. Teplyakov, *J. Appl. Cryst.*, 1997, **30**, 1022-1025.
11. P. Emsley and K. Cowtan, *Acta Crystallogr D Biol Crystallogr* 2004, **60**, 2126-2132.
12. M. D. Winn, M. N. Isupov and G. N. Murshudov, *Acta Cryst D*, 2001, **57**, 122-133.
13. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Cryst.*, 1993, **26**, 283-291.
14. I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, 3rd, J. Snoeyink, J. S. Richardson and D. C. Richardson, *Nucleic Acids Res.*, 2007, **35**, W375-383.
15. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235-242.
16. H. Heyerdahl, N. Abbas, E. M. Brevik, C. Mollatt and J. Dahle, *PloS one*, 2012, **7**, e42345.