Supporting Information

Rational design of a genetically encoded NMR zinc sensor

Zhuangyu Zhao,‡[a] Mingyang Zhou,‡[a] Serge D. Zemerov,[a] Ronen Marmorstein[ab] and Ivan J. Dmochowski*[a]

[a]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104–6323, United States.

[b]Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104–6323, United States.

Table of Contents

Experimental Procedures

Site-directed mutagenesis. Mutations were introduced to the pET His6 MBP TEV LIC cloning vector, a gift from Scott Gradia acquired via Addgene (# 29656), to remove His-tag and construct zinc binding site using the forward and reverse primers listed in Table S2. The mutated plasmids were amplified in NEB 5α competent *E. coli* cells (New England Biolabs) and purified using a miniprep kit (Qiagen). All mutated sequences were verified at University of Pennsylvania DNA Sequencing Facility.

ZS-GFP cloning. The GFP insert was amplified from the pET GFP LIC cloning vector, a gift from Scott Gradia acquired via Addgene (# 29772), using the primers listed in Table S3, and then added to the ZS vector by ligation independent cloning (LIC). The resulting ZS-GFP gene was sequenced at University of Pennsylvania DNA Sequencing Facility.

Protein expression and purification. MBP-K15D and MBP-D14K/K15D were expressed and purified using the same procedures for wt-MBP.[1] The plasmid encoding the sequence of ZS, ZS-GFP or ZS-D14H was transformed into BL21(DE3) competent *E. coli* cells (New England Biolabs) and grown in three flasks containing 1 L of LB media supplemented with 50 μg/mL kanamycin to an OD_{600} of 0.6 – 1.0. Protein expression was induced by adding 1 mM isopropyl-β-D-thiogalactoside (IPTG). The induced cells were incubated overnight at 25 °C, harvested by centrifugation and frozen at -80 °C. The cell pellets were resuspended in 20 mM Tris (pH 8.0), lysed with egg white lysozyme (Alfa Aesar), and treated with benzonase nuclease (Sigma). The lysate was clarified by centrifugation at 15,000 rpm for 1 h, and the supernatant was loaded onto a HiTrap DEAE FF anion exchange column (GE life Sciences) pre-equilibrated with 20 mM Tris (pH 8.0). Zinc sensor protein was eluted via a NaCl gradient of 0 – 500 mM. The eluate containing zinc sensor protein was concentrated and further purified by size-exclusion chromatography in 20 mM Tris (pH 7.4), 100 mM NaCl using a HiLoad 16/600 Superdex 75 pg column (GE Life Sciences). All pure protein solutions (purity >95% as evidenced by SDS-PAGE, Figure S2) were pooled and concentrated. Protein concentrations were determined from the absorbance at 280 nm using the extinction coefficient 67,840 M^{-1} cm^{-1} for MBP-K15D and MBP-D14K/K15D, 62,340 M^{-1} cm^{-1} for ZS and ZS-D14H and 81,250 M^{-1} cm^{-1} for ZS-GFP, as calculated by PROTPRAM.[2]

^{129}Xe hyper-CEST. Hyper-CEST z-spectra were acquired as described previously. In short, HP ^{129}Xe was generated using the spin-exchange optical pumping (SEOP) method with a home-built ^{129}Xe polarizer based on the IGI.Xe.2000 commercial model by GE. A Shark 65 W tunable ultra-narrow band diode laser (OptiGrate) set to 795 nm was used for optical pumping of Rb vapor. A gas mixture of 88% helium, 10% nitrogen, and 2% natural abundance xenon (Linde Group, NJ) was used as the hyperpolarizer input. ^{129}Xe hyperpolarization level was roughly 10-15%. For each data point in the hyper-CEST z-spectrum, hp ^{129}Xe was bubbled into the NMR tube through capillaries for 20 s, followed by a 3-s delay to allow bubbles to collapse. A d-SNOB saturation pulse with 690 Hz bandwidth was used. Pulse length, τ_{pulse} = 3.80 ms; field strength $B_{1,max}$ = 77 μT; number of pulses, n_{pulse} = 600; saturation time, T_{sat} = 2.28 s. NMR experiments were performed using a Bruker BioDRX 500 MHz NMR spectrometer and 10-mm PABBO probe, at 300 K. A 90° hard pulse with this probe has a pulse length of 22 μs. Protein concentration used was 80 μM, with 0.1% (v/v) Pluronic L81 (Aldrich) added to mitigate foaming. Data were acquired from -100 ppm

to +165 ppm at a step size of 5 ppm. A single z-spectrum has 54 data points and typically 30 to 40 minutes were required to collect all data for one z-spectrum. For the time-dependent saturation transfer experiments, 1 μM protein was used. Pulse length, τ_{pulse} = 1.0496 ms; field strength, $B_{1,max}$ = 279 μT. Both on-resonance and off-resonance data were fitted with first-order exponential decay curves. Saturation contrast was calculated as described before.[3]

Preparation of cell lysates containing ZS-GFP. The plasmid encoding ZS-GFP sequence was transformed into BL21(DE3) competent *E. coli* cells (New England Biolabs). Cells were grown in LB media with 50 μg/mL kanamycin until OD_{600} reached 0.6 – 1.0. The non-induced and IPTG-induced cells were incubated overnight at 25 °C. TPEN-treated cells were incubated for another 2 h after addition of 20 μM TPEN. Zn^{2+}-supplemented cells were incubated for 2 h after addition of 20 μM pyrithione and 100 μM Zn^{2+}. Then, the cells (5 mL) were washed with cold PBS five times to remove extracellular content and resuspended in 2 mL cold PBS supplemented with protease inhibitor (Thermo Scientific). Cell lysis was performed by three freeze-thaw cycles, after which the total ZS-GFP concentration was determined by fluorescence measurements using a standard curve generated with purified ZS-GFP. Supernatants of cell lysates were used for hyper-CEST experiments. Labile Zn^{2+} concentration was determined as follows: the concentration of Zn^{2+}-ZS-GFP was calculated using the linear correlation with hyper-CEST saturation contrast[1, 4] and a slope of 0.17 saturation contrast per micromolar protein was obtained from purified ZS-GFP (Figure 2). Then $[Zn^{2+}]_{labile}$ was calculated using the Hill equation, $[Zn^{2+}\text{-ZS-GFP}] / [\text{ZS-GFP}] = [Zn^{2+}]_{labile} / ([Zn^{2+}]_{labile} + K_d)$.

Molecular dynamics (MD). MD simulations were performed with NAMD 2.14 software[5] using the Bridges-2 Regular Memory system at the Pittsburgh Supercomputing Center (PSC).[6, 7] Initial structure of ZS was modified from the crystal structure of MBP_{closed} (PDB: 1ANF). Each protein was solvated in a TIP3P water box[8] and 150 mM NaCl was used to balance charge. Each system was minimized with 1,000 steps, after which the system was equilibrated at a temperature of 300 K and a pressure of 1 atm using the Langevin thermostat and barostat methods. The CHARMM36 force field was used to perform all MD simulations.[9] Zn^{2+} was treated with electrostatic forces and van der Waals interactions.[10] Periodic boundary conditions were employed, and SHAKE algorithm[11] was applied to constrain the lengths of all bonds that involve a hydrogen atom. Salt-bridge distances and RMSF data were calculated using VMD 1.9.3.[12]

Fluorescence spectroscopy. For ZS-GFP fluorescence measurements, the excitation wavelength was 488 nm, and fluorescence emission was recorded at 511 nm. Protein concentration varied from 1 μM to 6 μM to give a linear standard curve, from which ZS-GFP concentration in cell lysates was calculated. All measurements were obtained at 298 K using a Varian Cary Eclipse fluorescence spectrophotometer.

Circular dichroism (CD) spectroscopy. CD spectra of 10 μM protein sample were acquired in a 1-mm quartz cuvette using an Aviv Model 425 CD spectrometer. Data were recorded from 260 nm to 195 nm with a wavelength step of 1 nm at 298 K.

Isothermal titration calorimetry (ITC). Protein samples for ITC were prepared at 30 μM concentration in 20 mM Tris (pH 7.4), 100 mM NaCl, and titrants were at 300 to 500 μM

concentration in the same buffer. Binding measurements were carried out using a MicroCal iTC_{200} system (GE Life Sciences) at 298 K.

Apo-ZS/Zn^{2+}-bound ZS Crystallization and Data Collection. ZS was crystallized by vapor diffusion using the hanging-drop method. ZS at 5 mg/mL concentration in 10 mM MES, pH 6.2, 100 mM NaCl was mixed with an equal volume of reservoir solution. For apo-ZS, the reservoir solution was comprised of 0.2 M sodium formate, pH 7.0 and 20% PEG 3350. For Zn^{2+}-bound ZS, the reservoir solution was comprised of 0.2 M $MgCl_2$, 0.1 M Bis-Tris, pH 6.5 and 25% PEG 3350. Crystals were transferred into mother liquor, supplemented with 25% glycerol and flash-frozen in liquid nitrogen. Apo-ZS data were collected at the Advanced Photon Source (beamline 24-ID-C) and processed using the RAPD automated processing suite, which utilized XDS for integration and scaling.[13] Data for the Zn^{2+}-bound ZS crystal was collected at the Highly Automated Macromolecular Crystallography (AMX) beamline 17-ID-1 at the National Synchrotron Light Source II, Brookhaven National Laboratory (Upton, NY), using a wavelength measured at Zn K-absorption edge (1.28215 Å). Data were indexed, integrated, and scaled with FastDP.[13-16]

Apo-ZS/Zn^{2+}-bound ZS Structure Determination and Refinement. The crystal structures of apo-ZS and Zn^{2+}-ZS were determined by molecular replacement using unliganded MBP (PDB 1OMP) as a search model. Molecular replacement was done using Phaser in Phenix (Adams et al., 2010). Coot (Emsley and Cowtan, 2004) was used to generate the first manual model, and Phenix refine and Coot were both used interchangeably for all future rounds of refinement. Refinement statistics can be found in Table 2. The final model and structure factors were submitted to the Protein Data Bank with code PDB: 8F23 and 8ETB (Research Collaboratory for Structural Bioinformatics). PyMOL (http://www.pymol.org/) was used to prepare structural images, structural alignments and calculate distances.

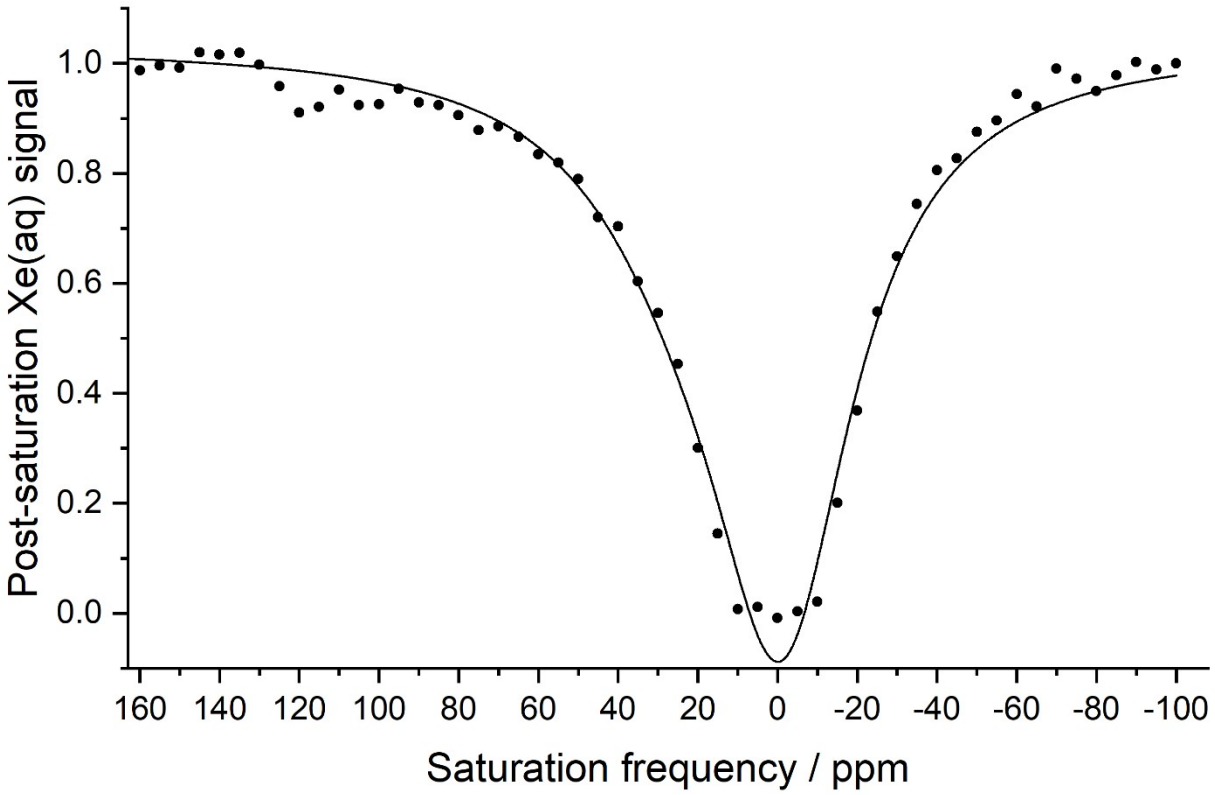

Figure S1. Hyper-CEST z-spectrum of a previously reported zinc sensor (MBP-A63H/R66H/Y155E/W340E)[17] at 80 μM concentration in presence of 400 μM $ZnCl_2$ in 20 mM Tris (pH 7.4), 100 mM NaCl at 300 K.

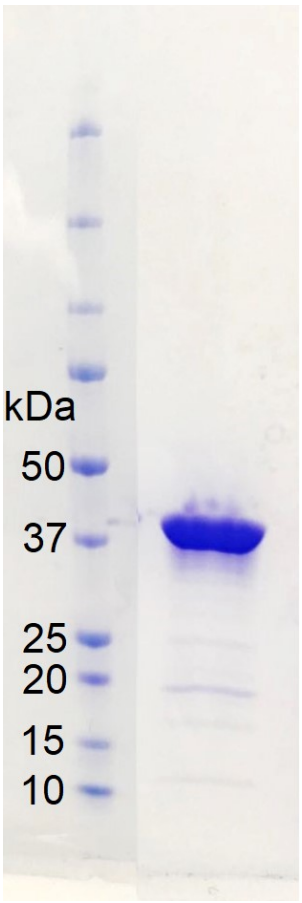

Figure S2. Identity and purity of ZS protein (MBP-K15H/E111H/W230H) were confirmed by SDS-PAGE analysis. Calculated molecular weight of ZS is 40.3 kDa.

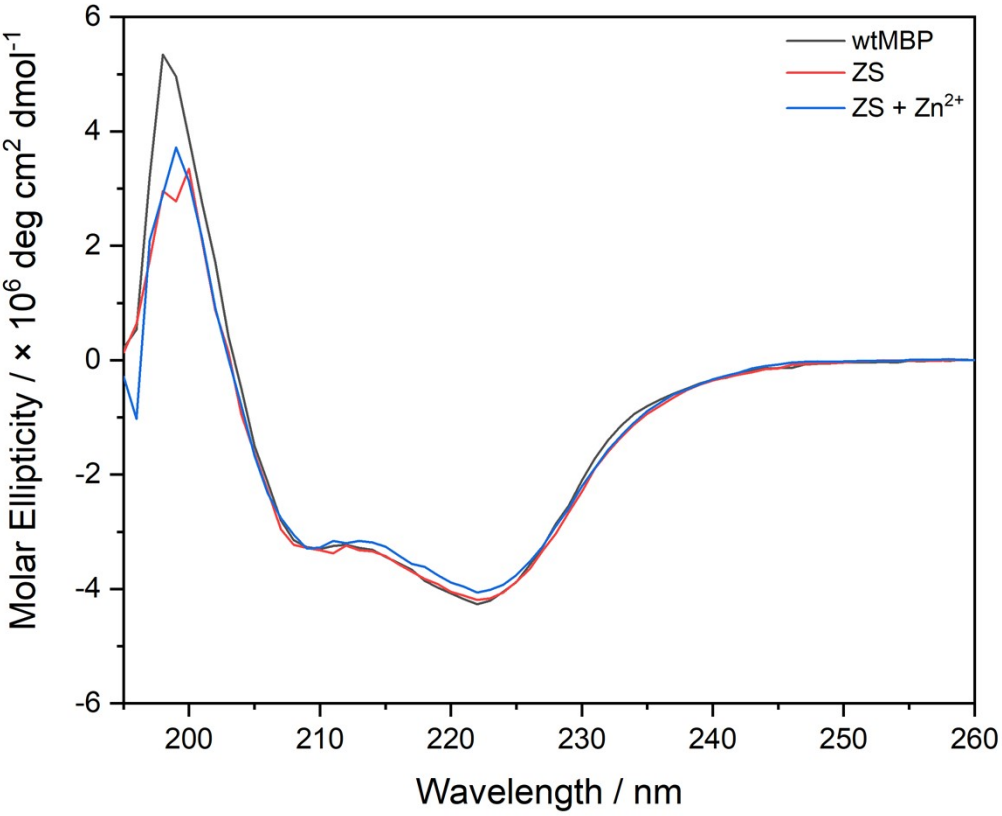

Figure S3. CD spectra of 10 µM wt-MBP in PBS (grey), 10 µM ZS in 20 mM Tris (pH 7.4), 100 mM NaCl (red) and 10 µM ZS with 50 µM $ZnCl_2$ in 20 mM Tris (pH 7.4), 100 mM NaCl (blue). Comparison of the three spectra revealed almost identical structures of wt-MBP, apo-ZS and Zn^{2+}-bound ZS.

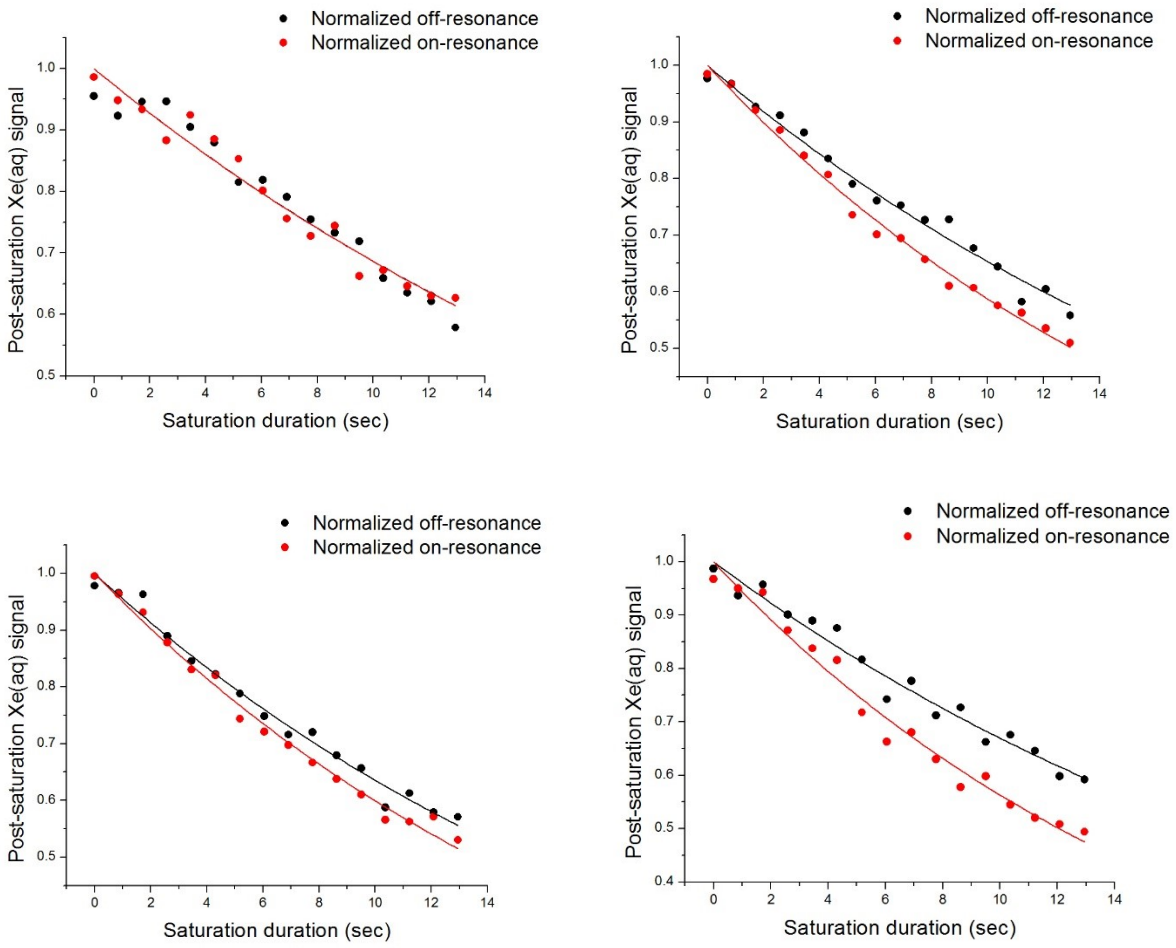

Figure S4. Time-dependent saturation transfer data for *E. coli* cell lysate at 300 K. Saturation pulses were positioned at +50 ppm and -50 ppm referenced to Xe@aq peak for on- and off-resonance. Top left: non-induced, T_{1on} = 26.6 ± 1.4 s, T_{1off} = 26.6 ± 1.6 s, saturation contrast = 0 ± 0.02. Top right: IPTG-induced, T_{1on} = 18.8 ± 0.4 s, T_{1off} = 23.5 ± 0.9 s, saturation contrast = 0.08 ± 0.01. Bottom left: IPTG-induced and incubated with TPEN, T_{1on} = 19.5 ± 0.5 s, T_{1off} = 22.1 ± 0.7 s, saturation contrast = 0.05 ± 0.01. Bottom right: IPTG-induced and incubated with zinc pyrithione, T_{1on} = 17.4 ± 0.7 s, T_{1off} = 24.9 ± 1.2 s, saturation contrast = 0.14 ± 0.01.

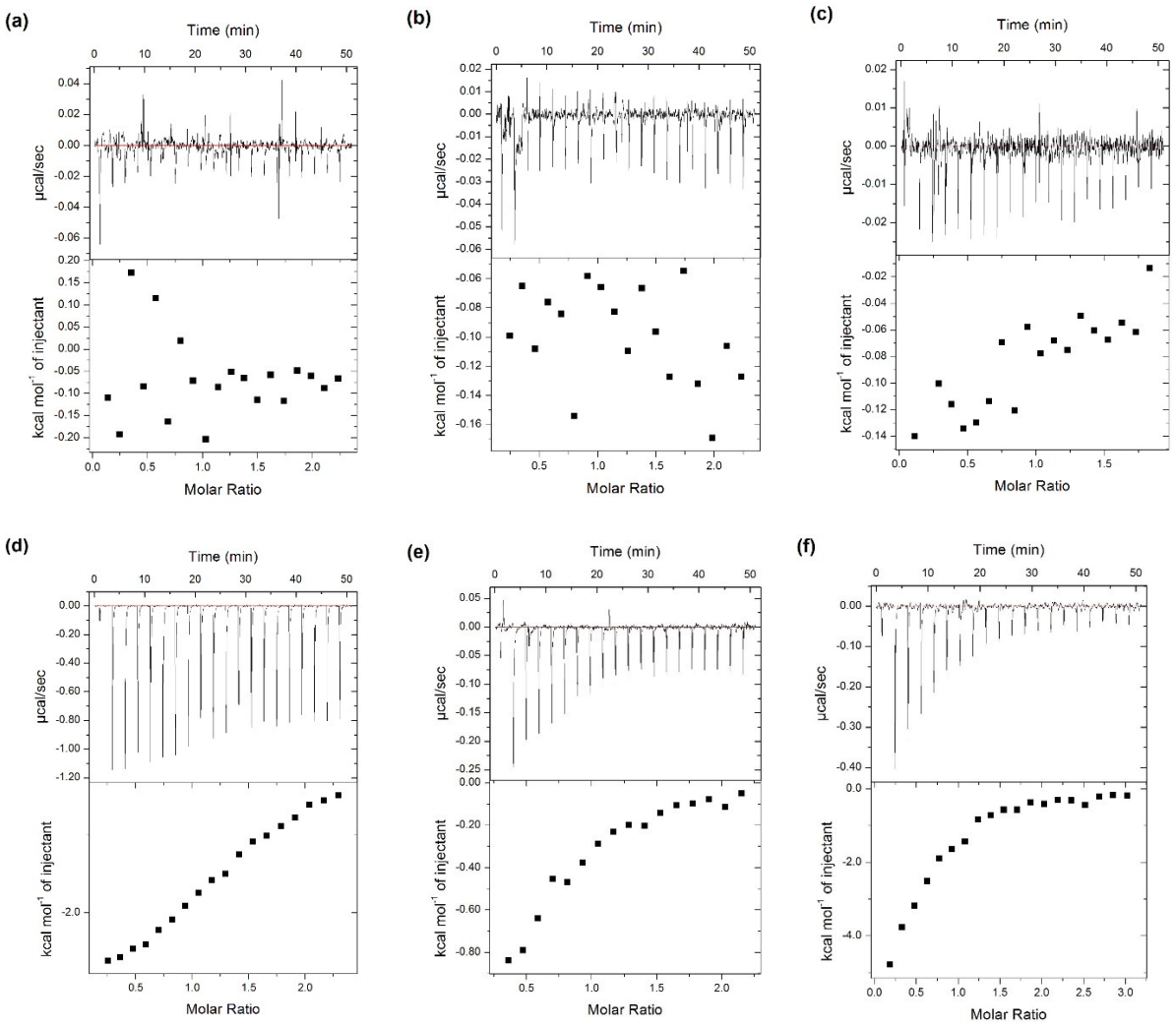

Figure S5. ITC enthalpograms of ZS with different titrants at 298 K. (a) $MgCl_2$; (b) $CaCl_2$; (c) $FeCl_3$; (d) $CuSO_4$ (K_d = 21.2 ± 1.5 μM); (e) $ZnCl_2$ (K_d = 21.7 ± 4.9 μM). (f) ITC enthalpograms of $ZnCl_2$ titrated into ZS-D14H at 298 K (K_d = 9.35 ± 1.21 μM).

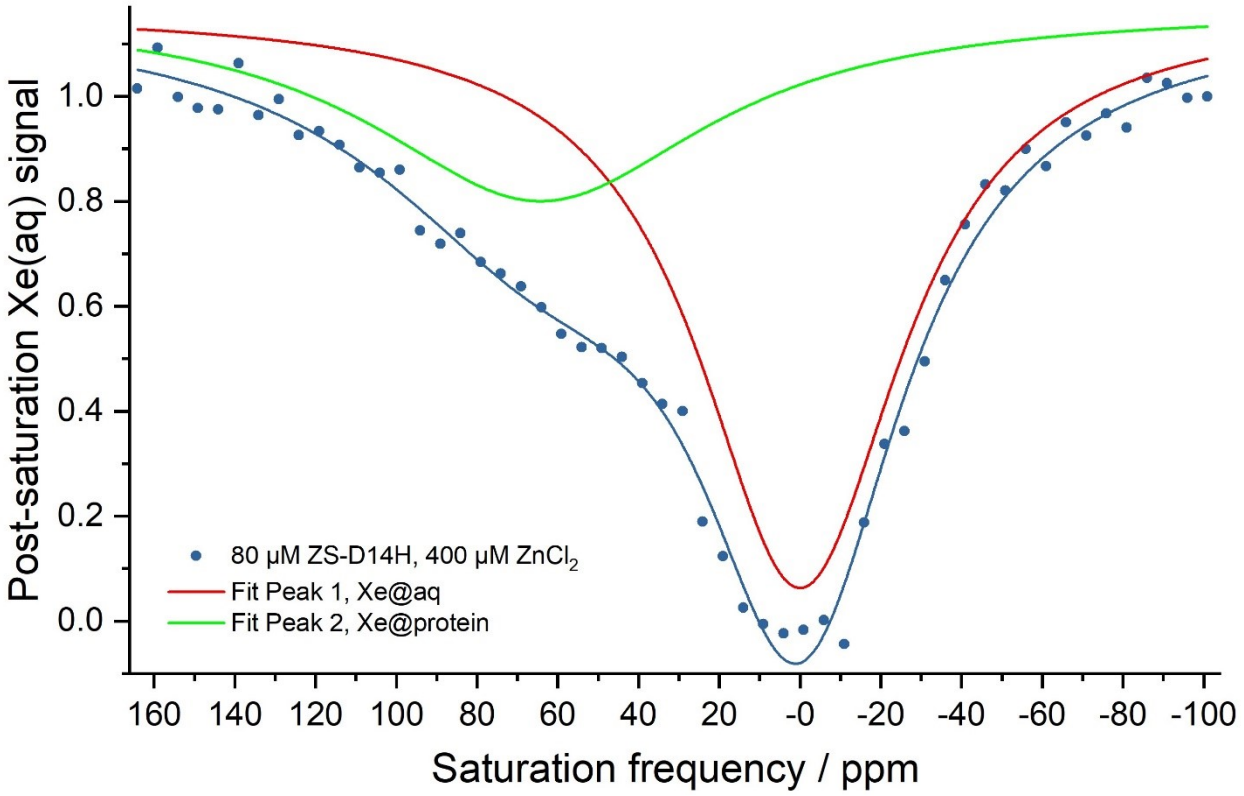

Figure S6. Hyper-CEST z-spectrum of 80 μM ZS-D14H in presence of 400 μM $ZnCl_2$ in 20 mM Tris (pH 7.4), 100 mM NaCl at 300 K. Red and green lines show Lorentzian fits to the Xe@aq and Xe@protein peaks. Chemical shift of ^{129}Xe@aq is referenced as zero and ^{129}Xe@protein peak is 64 ppm downfield-shifted from the ^{129}Xe@aq peak.

Table S1. Peak widths (FWHM) of hyper-CEST z-spectra

Z-spectrum	FWHM @aq / ppm	FWHM @protein /ppm
wtMBP, 1 mM maltose	47 ± 2	35 ± 2
ZS, 400 μM Zn^{2+}	42 ± 3	33 ± 3
ZS-D14H, 400 μM Zn^{2+}	62 ± 3	103 ± 17

Table S2. Oligonucleotide primers used in site-directed mutagenesis

ΔHis-tag	Forward	5'-GATATACCATGGGTTCTTCTGGTTCTTCTATGAAAATCG-3'
	Reverse	5'-CGATTTTCATAGAAGAACCAGAAGAACCCATGGTATATC-3'
K15H	Forward	5'-GATTAACGGCGATCATGGCTATAACGGTC-3'
	Reverse	5'-GACCGTTATAGCCATGATCGCCGTTAATC-3'
E111H	Forward	5'-ACCCGATCGCTGTTCATGCGTTATCGCTG-3'
	Reverse	5'- CAGCGATAACGCATGAACAGCGATCGGGT-3'
W230H	Forward	5'-CATCAACGGCCCGCATGCATGGTCCAACATC-3'
	Reverse	5'- GATGTTGGACCATGCATGCGGGCCGTTGATG-3'
D14H/K15H	Forward	5'-CTGGATTAACGGCCATCATGGCTATAACG-3'
	Reverse	5'- CGTTATAGCCATGATGGCCGTTAATCCAG-3'

Table S3. Oligonucleotide primers used for GFP insert amplification

GFP insert	Forward	5'-TACTTCCAATCCAATGCAAGCAAGGGCGAGGAGCTGTTC-3'
	Reverse	5'-TTATCCACTTCCAATGTTATTACTTGTACAGCTCGTCCATGCC-3'

References

1. B. W. Roose, S. D. Zemerov and I. J. Dmochowski, *Chem. Sci.*, 2017, **8**, 7631-7636.
2. E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch, in *The Proteomics Protocols Handbook*, ed. J. M. Walker, Humana Press, 2005, pp. 571-607.
3. Y. Wang and I. J. Dmochowski, *Acc Chem Res*, 2016, **49**, 2179-2187.
4. S. D. Zemerov, B. W. Roose, K. L. Farenhem, Z. Zhao, M. A. Stringer, A. R. Goldman, D. W. Speicher and I. J. Dmochowski, *Anal. Chem.*, 2020, **92**, 12817-12824.
5. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J Comput Chem*, 2005, **26**, 1781-1802.
6. J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott and N. Wilkins-Diehr, *Comput Sci Eng*, 2014, **16**, 62-74.
7. G. J. Stasiuk, F. Minuzzi, M. Sae-Heng, C. Rivas, H. P. Juretschke, L. Piemonti, P. R. Allegrini, D. Laurent, A. R. Duckworth, A. Beeby, G. A. Rutter and N. J. Long, *Chemistry*, 2015, **21**, 5023-5033.
8. W. L. Jorgensen, J. Chandrasekhar and J. D. Madura, *J Chem Phys*, 1998, **79**, 926.
9. C. Boutin, H. Desvaux, M. Carriere, F. Leteurtre, N. Jamin, Y. Boulard and P. Berthault, *NMR Biomed*, 2011, **24**, 1264-1269.
10. R. H. Stote and M. Karplus, *Proteins*, 1995, **23**, 12-31.
11. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
12. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33-38, 27-38.
13. W. Kabsch, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 125-132.
14. G. Winter and K. E. McAuley, *Methods*, 2011, **55**, 81-93.
15. R. W. Grosse-Kunstleve, N. K. Sauter, N. W. Moriarty and P. D. Adams, *J Appl Crystallogr*, 2002, **35**, 126-136.
16. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr D Biol Crystallogr*, 2011, **67**, 235-242.
17. J. S. Marvin and H. W. Hellinga, *Proc Natl Acad Sci U S A*, 2001, **98**, 4955-4960.