Reversible and size-controlled assembly and disassembly of reflectin proteins using a charged azobenzene photoswitch

Cassidy M. Tobin, Reid Gordon, Seren K. Tochikura, Bradley F. Chmelka, *Daniel E. Morse, and Javier Read de Alaniz**

C. M. Tobin, B. F. Chmelka.
Department of Chemical Engineering, University of California, Santa Barbara. Santa Barbara, California 93106, United States

R. Gordon, S. K. Tochikura, D. E. Morse.
Department of Molecular, Cellular, and Developmental Biology, University of California, Santa Barbara. Santa Barbara, California 93106, United States

J. Read de Alaniz.
Department of Chemistry and Biochemistry, University of California, Santa Barbara. Santa Barbara, California 93106, United States
E-mail: javier@chem.ucsb.edu

Table of Contents

1. Synthetic Procedures

1.1 Dryad Data Repository

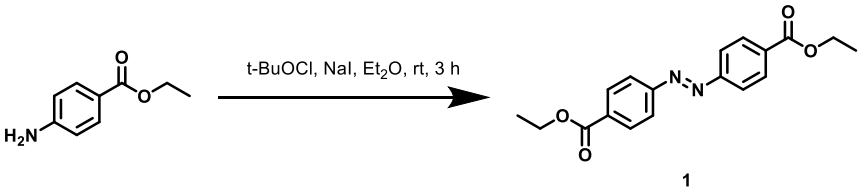
The data and methods for processing the data are described in the manuscript and SI can be found at https://doi.org/10.5061/dryad.fttdz092d.

1.2 Materials and Methods

All commercially obtained reagents were bought from Sigma Aldrich, Oakwood Chemicals, or Fisher Scientific and were used without purification. *Tert*-butyl hypochlorite was prepared using a literature procedure.[1] Anhydrous dichloromethane was dispensed from a solvent purification system immediately before use. Anhydrous dimethylformamide was purchased from Fisher Scientific. Room temperature reactions were carried out between 18–22 °C. Unless otherwise noted, reactions were performed under an inert atmosphere of nitrogen. Thin layer chromatography (TLC) was performed using Merck TLC plates (silica gel 60 F254 on aluminum) and visualized by UV light (254 nm) or staining with *p*-anisaldehyde. Silica gel chromatography was performed using normal phase silica gel from Geduran (technical grade, 60 Å pore size, 40–63 µm particle size). ^{1}H- and ^{13}C-NMR were recorded on Bruker spectrometers (400 MHz or 500 MHz) and are reported relative to $CDCl_3$, DMSO-d_6, or 10 vol% D_2O in water. Mass spectra were obtained from the UCSB DCB Mass Spectrometry Facility on a quadropole-time of flight mass analyzer with an ESI source. FTIR spectra were collected on a Thermo Nicolet iS10 FTIR Spectrometer equipped with a Smart Diamond attenuated total reflectance accessory.

1.3 Synthesis of azoEDTA

Oxidative dimerization of ethyl 4-aminobenzoate to **1**:



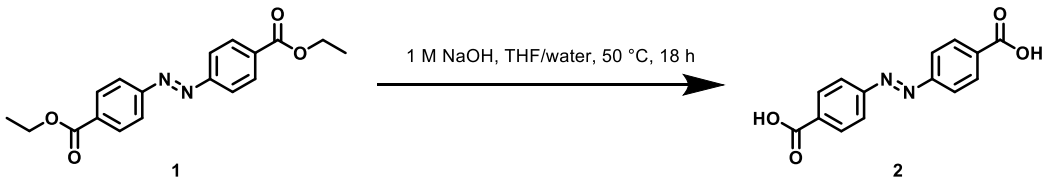
1

(*E*)-Diethyl 4,4'-(Diazene-1,2-diyl)dibenzoate (1):

The procedure to prepare compound **1** was modified from a literature procedure.[2] Diethyl ether (145 mL) was added to a 250 mL round bottom flask with ethyl 4-aminobenzoate (4.0 g, 24.2 mmol, 1.0 equiv) and sodium iodide (7.3 g, 48.4 mmol, 2.0 equiv) to form a slurry. The reaction was stirred at room temperature, and *tert*-butyl hypochlorite (5.3 mg, 48.4 mmol, 2.0 equiv) was added slowly. The reaction was quenched after 3 hours with 1 M $Na_2S_2O_3$ (500 mL) and transferred to a separatory funnel. The solution was extracted with CH_2Cl_2 (2x 300 mL). The organic layer was washed with water (4x 300 mL), dried over Na_2SO_4, and concentrated to give a bright orange solid. The crude product was subjected to flash column chromatography (hexanes:ethyl acetate 9:1) to yield compound **1** (2.9 g, 9.0 mmol, 74%). Spectral data matched that reported in literature.[2]

Compound 1: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.22 (d, 4H, *J* = 8.8 Hz), 7.99 (d, 4H, *J* = 8.8 Hz), 4.43 (q, 4H, *J* = 7.1 Hz), 1.44 (t, 6H, *J* = 7.1 Hz) ppm

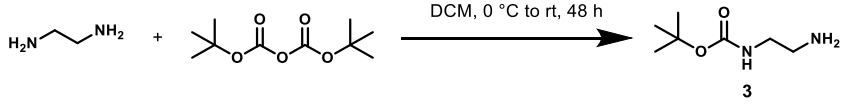
Saponification of **1** to **2**:



(*E*)-4,4'-(diazene-1,2-diyl)dibenzoic acid (2):
Compound **1** (2.3 g, 7.2 mmol, 1.0 equiv) was dissolved in THF (approximately 60 mL) and added to a round bottom flask with 1 M NaOH solution (900 mL, 900 mmol, 125 equiv). The reaction was stirred at ambient atmosphere for 18 h at 50 °C, cooled to room temperature, and acidified with concentrated HCl until the product precipitated (approximately 90 mL). The product was filtered and dried to yield compound **2** (1.8 g, 6.8 mmol, 94%) as an orange solid.

Compound 2: ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.16 (d, 4H, *J* = 8.7 Hz), 8.01 (d, 4H, *J* = 8.8 Hz), 4.43 (q, 4H, *J* = 7.1 Hz), 1.44 (t, 6H, *J* = 7.1 Hz) ppm

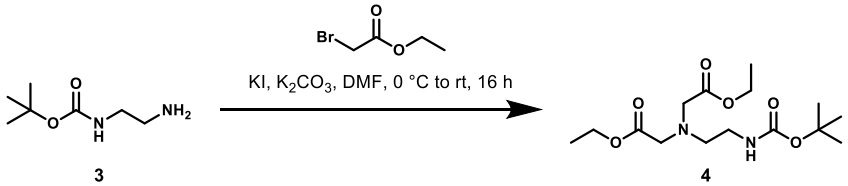
Boc-protection of ethylenediamine **3**:



***Tert*-butyl (2-aminoethyl)carbamate (3):**
1,2-Ethylenediamine (12.6 mL, 192.4 mmol, 6.9 equiv) was dissolved in 60 mL of anhydrous DCM in a round bottom flask and cooled to 0 °C. Di-*tert*-butyl dicarbonate (6.1 g, 27.8 mmol, 1.0 equiv) was dissolved in 140 mL of anhydrous DCM and the solution was added to the reaction flask dropwise. The reaction was then warmed up to room temperature. After 48 h, the reaction mixture was concentrated down to a yellow residue then redissolved in saturated K_2CO_3. The solution was transferred to a separatory funnel and extracted with CH_2Cl_2 (3x 150 mL). The organic layer was washed with brine (200 mL), dried with Na_2SO_4, and concentrated to a yellow oil (4.3 g, 26.7 mmol, 96%).

Compound 3: ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.84 (b, 1H), 3.17 (dt, 2H, *J* = 6.0, 6.0 Hz), 2.80 (t, 2H, *J* = 5.9 Hz), 1.45 (s, 9H), 1.13 (b, 2H) ppm

N-Alkylation/nucleophilic substitution of **3** with halogenated ester to **4**:



diethyl 2,2'-((2-((tert-butoxycarbonyl)amino)ethyl)azanediyl)diacetate (4):
Boc-protected ethylenediamine (410 mg, 2.6 mmol, 1.0 equiv), potassium iodide (1.2 g, 7.5 mmol, 2.9 equiv), and potassium carbonate (6.2 g, 45.0 mmol, 17.6 equiv) were dissolved in 5 mL of DMF in a round bottom flask. The reaction was cooled to 0 °C. Ethyl bromoacetate (0.83 mL, 7.5 mmol, 2.9 equiv) was added and the reaction was stirred at room temperature for 16 hours. The reaction was quenched with 3 mL of 1 M NaOH and transferred to a separatory funnel. The product was extracted with ethyl acetate

(3x 100 mL). The organic layer was washed with water (4x 150 mL), dried over Na_2SO_4, and concentrated to a yellow oil. The crude product was subjected to flash column chromatography (hexanes:ethyl acetate 4:1) to yield compound **4** as a clear oil (725 mg, 2.2, 85%). Spectral data match that reported in literature.[3]

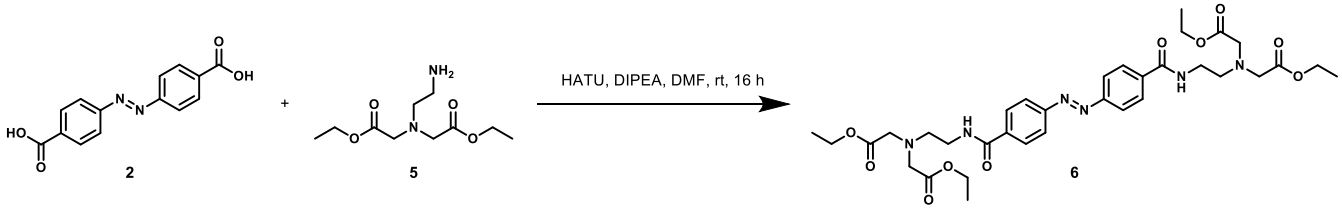
Compound 4: ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.53 (b, 1H), 4.17 (q, 4H, *J* = 7.1 Hz), 3.52 (s, 4H), 3.16 (dt, 2H, *J* = 5.5, 5.4 Hz), 2.85 (t, 2H, *J* = 4.8 Hz), 1.44 (s, 9H), 1.27 (t, 6H, *J* = 7.1 Hz) ppm

Boc-deprotection of **4** to make primary amine **5**:

diethyl 2,2'-((2-aminoethyl)azanediyl)diacetate (5):
Compound 5 (610 mg, 1.8 mmol, 1 equiv) was dissolved in 9 mL of CH_2Cl_2 in a round bottom flask and cooled to 0 ˚C under an ambient atmosphere. Trifluoroacetic acid (2.2 mL, 28.8 mmol, 15.6 equiv) was dissolved in 9 mL CH_2Cl_2 and the solution was added dropwise to the reaction. The reaction was warmed up to room temperature and stirred for 5 h. The solvent was removed and isolated as a trifluoroacetate salt (616 mg, 1.8 mmol, 97%).

Compound 5: ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.59 (b, 3H), 4.10 (q, 4H, *J* = 7.1 Hz), 3.58 (s, 4H), 2.86 (s, 4H), 1.20 (t, 6H, *J* = 7.1 Hz) ppm

Amide formation via HATU-coupling of **2** and **5** to **6**:

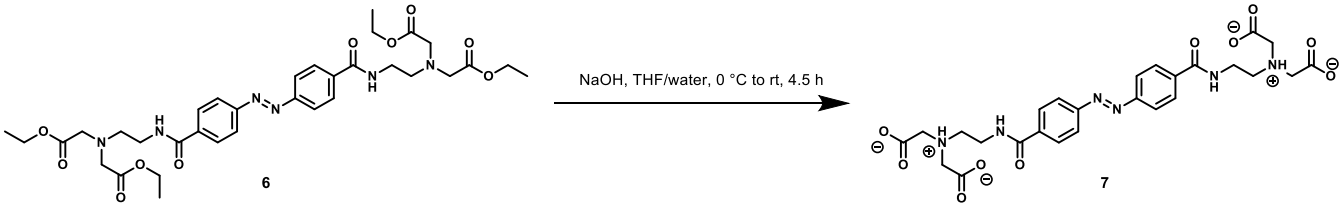


tetraethyl 2,2',2'',2'''-((((4,4'-(diazene-1,2-diyl)bis(benzoyl))bis(azanediyl))bis(ethane-2,1-diyl))-bis(azanetriyl))(*E*)-tetraacetate (6):
The procedure to prepare compound **6** was modified from a literature procedure.[4] To a round bottom flask equipped with a stir bar was added compound **2** (150 mg, 0.6 mmol, 1.0 equiv) and HATU (640 mg, 1.7 mmol, 3.0 equiv). DMF was added (10 mL) to form an orange slurry and the reaction was stirred at room temperature. N,N-diisopropylethylamine (0.8 mL, 4.59 mmol, 8.2 equiv) was added to the reaction and the solution turned dark red and transparent. The reaction was stirred for 15 minutes. Compound **5** (454 mg, 2.0 mmol, 3.5 equiv) was dissolved in 2 mL DMF and 0.39 mL of N,N-diisopropylethylamine (1.1:1 DIPEA:amine) was added to deprotonate the amine. The amine solution was added dropwise to the reaction mixture. The reaction was quenched with water (10 mL) after 16 hours and transferred to a separatory funnel. The product was extracted with ethyl acetate (2x 100 mL). The organic layer was washed with water (4x 150 mL), dried over Na_2SO_4, and concentrated. The crude product was purified

via plug (CH_2Cl_2:ethyl acetate 2:1) to yield compound **6** as a bright orange solid (263 mg, 0.4 mmol, 68%). Spectral data match that reported in literature.[5]

Compound 6: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.15 (d, 4H, *J* = 8.4 Hz), 8.12 (t, 2H, *J* = 4.4 Hz), 7.99 (d, 4H, *J* = 8.2 Hz), 4.20 (q, 8H, *J* = 7.2 Hz), 3.58 (s, 8H), 3.50 (dt, 4H, *J* = 7.7, 4.4 Hz), 2.99 (t, 4H, *J* = 5.4 Hz), 1.28 (t, 12H, *J* = 7.1 Hz) ppm; ^{13}C-NMR (125 MHz, $CDCl_3$): δ 171.9, 166.5, 154.0, 136.37, 128.4, 122.9, 60.9, 55.2, 52.2, 38.0, 14.2 ppm.

Acid formation via ester saponification of **6** to **7**:



(*E*)-2,2',2'',2'''-((((4,4'-(diazene-1,2-diyl)bis(benzoyl))bis(azanediyl))bis(ethane-2,1-diyl))-bis(azanetriyl))tetraacetic acid (7):

Compound **6** (101 mg, 0.1 mmol, 1.0 equiv) was dissolved in minimal THF (~1.5 mL), diluted with deionized water (5.5 mL), and cooled to 0 °C. Sodium hydroxide (458 mg, 11.4 mmol, 81.7 equiv) was added and the reaction stirred for 30 minutes. The ice bath was removed and the reaction stirred for 4 hours at room temperature. The organic solvent was evaporated and the aqueous mixture was transferred to dialysis tubing (Spectrum Labs, 0.1-0.5 kDa Biotech CE Dialysis Kit). The product was dialyzed against MilliQ water to remove salt (1,000-fold excess of dialyte). The solvent was removed to give an orange solid (81.9 mg, 0.1 mmol, 97%).

Compound 7: ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.62 (t, 2H, *J* = 5.1 Hz), 8.07 (d, 4H, *J* = 8.5 Hz), 7.98 (d, 4H, *J* = 8.4 Hz), 3.52 (s, 8H), 3.37 (q, 4H, *J* = 6.1 Hz), 2.87 (t, 4H, *J* = 6.4 Hz) ppm; ^{13}C-NMR (125 MHz, $CDCl_3$): δ 173.0, 165.3, 153.2, 137.1, 128.43, 122.6, 54.7, 52.5, 38.0 ppm; FTIR (ATR) 3266, 3008, 2970, 1742, 1610, 1548, 1493, 1400, 1294, 1141, 973, 861, 773 cm^{-1}; HRMS (ES+) Exact mass cald. For $C_{26}H_{30}N_6O_{10}$ $[M+H]^+$: 587.2101, found: 587.2102.

2. Reflectin Protein System Preparation

The pH of solutions was measured with an Orion Star A111 Benchtop pH meter and an Orion PerpHecT ROSS combination pH microelectrode (Thermo Fisher Scientific, Waltham, MA). Nanopure water (Thermo Fisher Scientific, Waltham, MA) was used for all solution preparation.

2.1 Reflectin Purification and Dialysis

Reflectin A1 wildtype protein was expressed and purified according to literature procedures.[6] Lyophilized protein was solubilized with the addition of 20 mM, pH 4.50 sodium acetate buffer. The buffer solution was filtered with a sterile, 0.2 µm Steritop bottle top filter and centrifuged at 14,000 rpm for 10 min to remove dust contaminants. The absence of dust was confirmed by DLS. To remove contaminants and excess salt, the protein was dialyzed 1,000-fold against a buffer solution (with pH and chemical composition as described) using a 10,000 MWCO dialysis cup. Protein concentration was determined by

measuring absorbance at 280 nm using a previously determined molar extinction coefficient of 120,865 M^{-1} cm^{-1}. Protein was stored at 4 °C before use.

2.2 Sample Preparation

Stock solutions of azoEDTA or other multivalent small molecules were prepared volumetrically with a pH 4.50, 20 mM acetate buffer, unless otherwise specified. The pH was adjusted with 0.5 mM solutions of NaOH or HCl in NanoPure water. Total ionic strength of the solutions was calculated after pH adjustment according to

$$\mathrm{I} = \frac{1}{2}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{c_i z_i^2} \quad \text{(S1)}$$

where I is the total ionic strength of the solution, calculated from the concentration, c_i, and the valency, z_i, of each species i. Prior to use with reflectin, all solutions were filtered with a sterile, 0.2 µm Acrodisc syringe filter and were centrifuged to remove dust contaminants. The photoswitch solution was left to thermally equilibrate in the dark for two days.

Unless otherwise specified, solutions for experimentation were prepared by diluting reflectin in 20 mM, pH 4.50 acetate buffer to 8 µM of reflectin monomer in a total volume of 110 µL. A stock solution of multivalent small molecules or photoswitch was added to the dilute protein sample at a specified concentration. The solution was mixed with a micropipette immediately and equilibrated for two hours in the dark prior to experimentation.

3. Characterization of Reflectin-azoEDTA Complexes

3.1 Dynamic Light Scattering

Dynamic light scattering was performed on a Malvern Zetasizer Nano ZS with pre-equilibrated 110 µL samples. The cuvette was equilibrated to 20 °C for 30 s prior to measurement. Five measurements were taken for each sample, each consisting of 5 runs for 10 s each.

3.2 Circular Dichroism

Circular dichroism spectroscopy was performed on a Jasco J-1500 circular dichroism spectrophotometer using a 0.5 mm pathlength demountable cuvette (Aireka Scientific Co., Hong Kong, China). A background measurement was collected using 110 µL of 20 mM, pH 4.50 acetate buffer. Data shown are the average of 4 measurements in the wavelength range of 400 nm to 185 nm with scanning speed of 50 nm/min and data integration time (DIT) of 4 s. For irradiation experiments, one measurement was collected and averaged over 4 independent experiments, due to structural dynamics of the system that change on the timescale of the measurements. High tension values are < 480 V for all data shown. Spectra were smoothed using the Savitzky-Golay function in Matlab with order 7 and frame length 11.

3.3 Protein Assembly Precipitation and Turbidity Experiments

After equilibration for two hours, the absorbance spectrum of azoEDTA-reflectin solutions with 0–480 µM photoswitch was taken using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Absorbance at 400 nm was used to calculate turbidity according to

$$T = 100\% \times (1 - 10^{-\mathrm{A}_{400}}) \quad \text{(S2)}$$

The contribution of azoEDTA to total absorption at 400 nm was subtracted from the measured absorption using the molar extinction coefficient of azoEDTA at 400 nm and the known concentration. The turbidity data was fit to a sigmoidal function in Matlab,

$$T = 100\% \times \left[0.0595 + \frac{0.734}{1+\exp\left(\frac{161-c_{azoEDTA}}{28.5}\right)}\right] \quad \text{(S3)}$$

The R^2 of the fits are all > 0.998 with root mean squared error < 0.02.

For protein precipitation measurements, samples were centrifuged at 14,000 rpm for 40 min at 20 °C and the absorbance spectrum of the supernatant was measured. DLS measurements of the supernatant confirmed that all azoEDTA-reflectin assemblies were precipitated and only free (i.e., non-interacting) reflectin monomers and azoEDTA were present in the supernatant. The concentration of soluble reflectin and azoEDTA in the supernatant was calculated using absorbances at 280 nm and 330 nm, respectively.

$$\mathrm{c_{azoEDTA}} = \frac{\mathrm{A_{331}}}{\varepsilon_{331}}\left(\frac{10^6\ \mu M}{1\ M}\right) \quad \text{(S4)}$$

The percentage of free, or unbound, azoEDTA was calculated using the concentration of photoswitch in the supernatant.

$$\text{Free AzoEDTA} = \left(\frac{\mathrm{c_{azoEDTA}}}{\mathrm{c_{initial}}}\right) \times 100\% \quad \text{(S5)}$$

AzoEDTA has non-negligible absorbance at 280 nm (Fig. S25). The absorbance of the photoswitch was subtracted from the total absorbance at 280 nm to calculate reflectin concentration.

$$\mathrm{c_{reflectin}} = \frac{\mathrm{A_{280}} - \varepsilon_{\mathrm{azo,\,280}} * \mathrm{c_{azoEDTA}}}{\varepsilon_{\mathrm{reflectin,\,280}}} \quad \text{(S6)}$$

$$\text{Soluble protein} = \frac{\mathrm{c_{reflectin}}}{8\ \mu\mathrm{M}} \times 100\% \quad \text{(S7)}$$

The concentration of reflectin and azoEDTA in the pellet was calculated by subtracting the concentration of soluble protein and azoEDTA from the concentration of each component added during sample preparation.

$$\mathrm{m_{reflectin,\,pellet}} = \frac{8\ \mu\mathrm{M} - \mathrm{c_{reflectin}}}{110\ \mu\mathrm{L}} \quad \text{(S8)}$$

$$\mathrm{m_{azo,\,pellet}} = \frac{\mathrm{c_{initial}} - \mathrm{c_{azoEDTA}}}{110\ \mu\mathrm{L}} \quad \text{(S9)}$$

$$\frac{\text{Moles photoswitch}}{\text{Moles protein}} = \frac{\mathrm{m_{azoEDTA,\,pellet}}}{\mathrm{m_{reflectin,\,pellet}}} \quad \text{(S10)}$$

Each measurement was repeated three times. The results were averaged and the error shown (**Fig. 4**) represents the standard deviation from 3 separate calculations.

For precipitation experiments with the photostationary state (PSS) of azoEDTA, the photoswitch was irradiated with 365 nm light for 40 min prior to sample preparation to ensure the PSS had been reached. Control measurements were performed with a reflectin-only and *trans*-azoEDTA-only solution; no change in absorbance was measured following centrifugation (Fig. S25).

3.4 Sample Irradiation

A 2 mm path length quartz cuvette with sample solution was irradiated for 20 min with 365 nm collimated LED lamp and 8 min with 470 nm collimated LED lamp (Thorlabs M365L2-C1 andM470L3-C1, Newton,

NJ). The 365 nm and 470 nm LED bulbs had light intensities of 178 mW cm^{-2} and 108 mW cm^{-2}, respectively, and were measured with a Newport Power Meter (Model 843-R, Irvine, CA). Since continuous mixing leads to aggregation, the solution was gently pipette-mixed every 5 min during irradiation.

3.5 Photostationary State and Half-Life of *cis*-azoEDTA

A 3.1 mM solution of azoEDTA in 20 mM acetate buffer (pH 4.50) with 10 vol% D_2O was prepared and stored in the dark for 3 days. The solution was irradiated with 365 nm light for 30 min, or until the photostationary state had been reached by UV-Vis. ^{1}H-NMR (500 MHz) spectra were collected before and immediately after irradiation. Integration of the peaks for aromatic and amide protons were used to quantify the fraction of *trans*-azoEDTA and *cis*-azoEDTA in mixture. The sample was stored in the dark and ^{1}H-NMR spectra were recorded for several weeks to determine the half-life of the *cis* isomer. The rate constant of thermal relaxation was calculated by fitting the data to a first order rate equation in Matlab.

3.6 UV-Vis Spectroscopy

UV-Vis measurements for samples with volume > 350 µL were conducted on an Agilent UV-Vis Spectrometer in the wavelength range of 200 nm to 800 nm using a quartz cuvette. UV-Vis measurements for sample volumes < 350 µL were taken using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA). A sample volume of 2 µL was used for each measurement. All UV-Vis measurements were blanked with the appropriate buffer solution.

3.7 Time-dependent UV-Vis Spectroscopy

Photoinduced optical absorbance kinetics were measured on a home-built pump-probe set-up. The probe beam was generated by a Deuterium-Tungsten Halogen UV-Vis light source (Ocean Optics DH-MINI) and the pump beam was generated by LED light sources at 365 nm and 470 nm (Thorlabs M365FP1 and M470F4, respectively). The 365 nm and 470 nm LED bulbs had light intensities of 0.26 mW cm^{-2} and 0.31 mW cm^{-2}, respectively, and were measured with a Newport Power Meter (Model 843-R, Irvine, CA). Both beams were coupled into multimode optical fibers terminated with an output collimator. The experiments were controlled by a LabVIEW program that measured optical absorption spectra of the sample and controlled the pump and probe light sources. Further details can be found in previous literature.[7]

3.8 Kinetic Modelling

Data collected from the pump-probe experiments were fit to first-order exponential models in Matlab using the absorbance at the maximum absorbance wavelength. The photoisomerization of azoEDTA (λ_{max} = 333 nm) was modeled using a first order, monoexponential function

$$Abs_{t,\lambda_{max}} = Abs_{PSS,\lambda_{max}} + (Abs_0 - Abs_{PSS,\lambda_{max}})e^{-k_{A,\lambda_{irr}}*t} \quad \text{(S11)}$$

where $Abs_{PSS,\lambda}$ is the absorbance at the photostationary state after irradiation with λ = 365 or 470 nm light, Abs_0 is the sample absorbance prior to irradiation, and $k_{A,\lambda}$ is the first-order rate coefficient of the photoisomerization after irradiation with λ. The fitted rate constants are given in Table S1.

The reflectin-azoEDTA system (λ_{max} = 338 nm) was modeled using a first-order, biexponential function to obtain rate coefficients for the photoisomerization and protein assembly processes. The assembly-disassembly of reflectin-azoEDTA complexes was assumed to be a pseudo-first order process. This strategy has been shown for biological and inorganic systems that are stimuli-responsive and involve molecular or conformational rearrangement.[8–12] The reversible photoswitching and assembly processes contribute separately to the measured absorbance (via light absorption and light scattering, respectively) with different timescales. Therefore, we assume they can be approximated as independent processes, thus making the biexponential model given by Equation S12 suitable to estimate rate coefficients.[13]

$$Abs_{t,\lambda_{max}} = Abs_{PSS,\lambda_{max}} + \left(Abs_0 - Abs_{PSS,\lambda_{max}}\right)\left(e^{-k_{A,\lambda_{irr}}*t} + e^{-k_{R,\lambda_{irr}}*t}\right) \quad \text{(S12)}$$

In this equation, $k_{R,\lambda}$ is the pseudo-first-order rate coefficient associated with protein (dis)assembly after irradiation with λ. The fitted rate constants are given in Table S2. Due to noise in the data, the R^2 of the fits from Equations S11 and S12 are all > 0.94 with root mean squared error < 0.004 for all fits.

4. Results and Discussion

4.1 Physical Crosslinking of small molecules with reflectin A1 protein

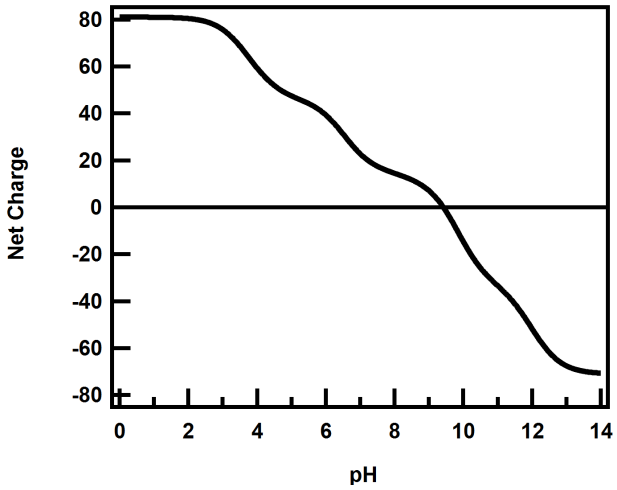


Figure S1. Calculated net charge of reflectin A1 wildtype sequence from pH 0–14 using the Henderson-Hasselbalch equation. All ionizable pKa values can be found in literature.[14] Calculations were performed in Matlab with pH increments of 0.01.

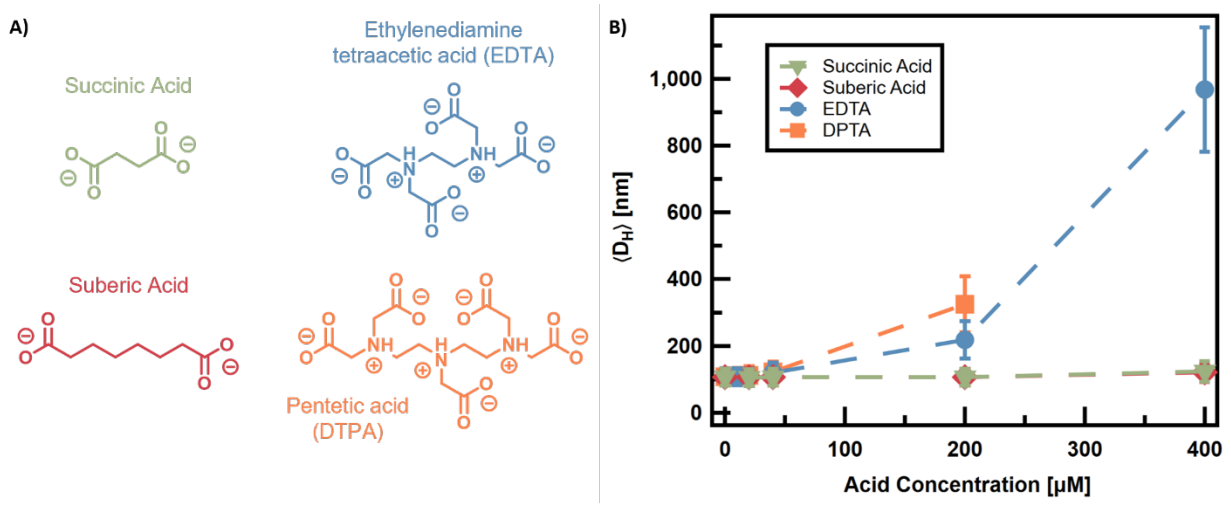


Figure S2. A) The valency and size of chelating agents with deprotonated carboxylate groups was varied. **B)** The hydrodynamic diameter (‹D_H›) of reflectin increased with concentration of chelating agents with more than two acid units. DLS measurements were taken with 50 µL solutions with 20 µM reflectin at pH 7.50 in a 10 mM MOPS buffer. The error bars represent one standard deviation of average diameter.

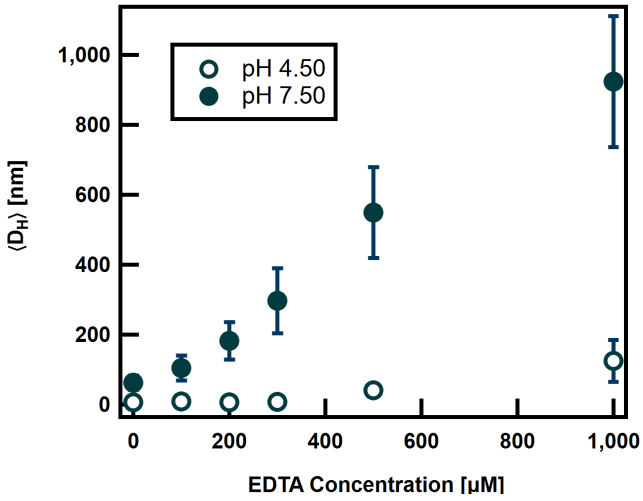


Figure S3. The mean hydrodynamic diameter (‹D_H›) of reflectin assemblies is dependent on pH and EDTA concentration. All samples were prepared with 20 µM reflectin. At pH 4.50, a 10 mM acetate buffer was used for protein dialysis and all sample preparation. At pH 7.50, a 10 mM MOPS buffer was used for protein dialysis and all sample preparation.

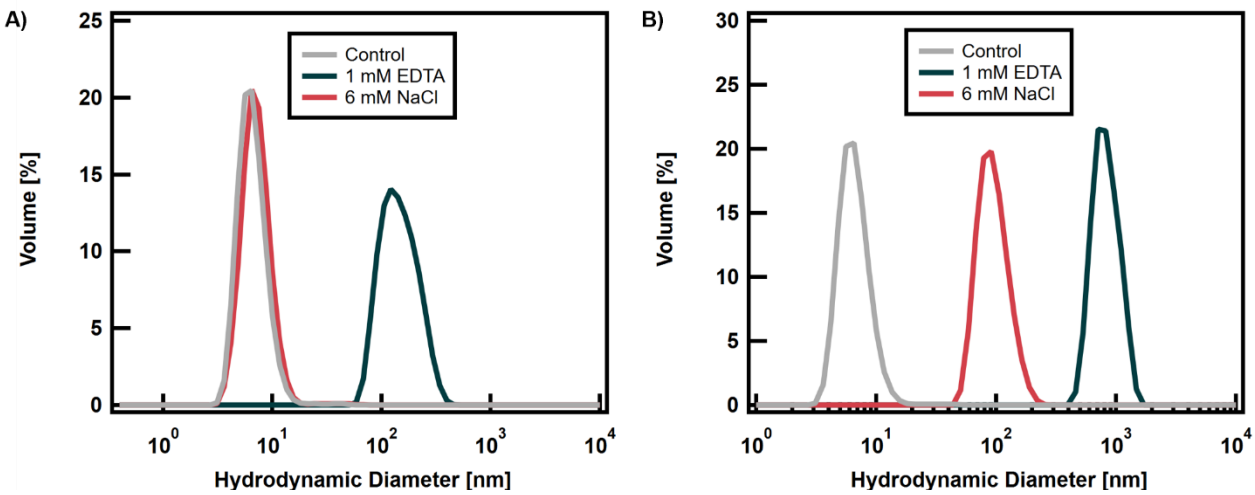


Figure S4. In equivalent ionic strength solutions (6 mM anion concentration), ethylenediamine tetraacetic acid (EDTA) and sodium chloride (NaCl) show different capacity to drive reflectin assembly formation at **A)** pH 4.50 (10 mM acetate buffer) and **B)** pH 7.50 (10 mM MOPS buffer), supporting the hypothesis that multivalent molecules such as EDTA form physical crosslinks with reflectin proteins. All samples were prepared with 20 µM reflectin and the control sample is a reflectin-only solution.

4.2 Photochemical properties of azoEDTA

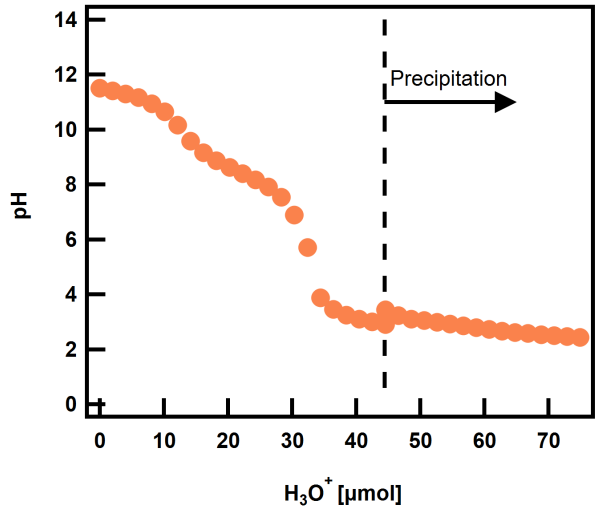


Figure S5. Titration on dark-equilibrated azoEDTA was performed by dissolving it (to 2 mM concentration) in a standardized solution of 16 mM NaOH and titrating with a standardized 25 mM solution of HCl. Two apparent pK_a values occur at pH 11.1 and 8.2. We attribute these to deprotonation of the tertiary amines, which is consistent with literature pK_as for EDTA.[15] Precipitation of the photoswitch was observed at pH 3.0, which we attribute to protonation of a carboxylate group.

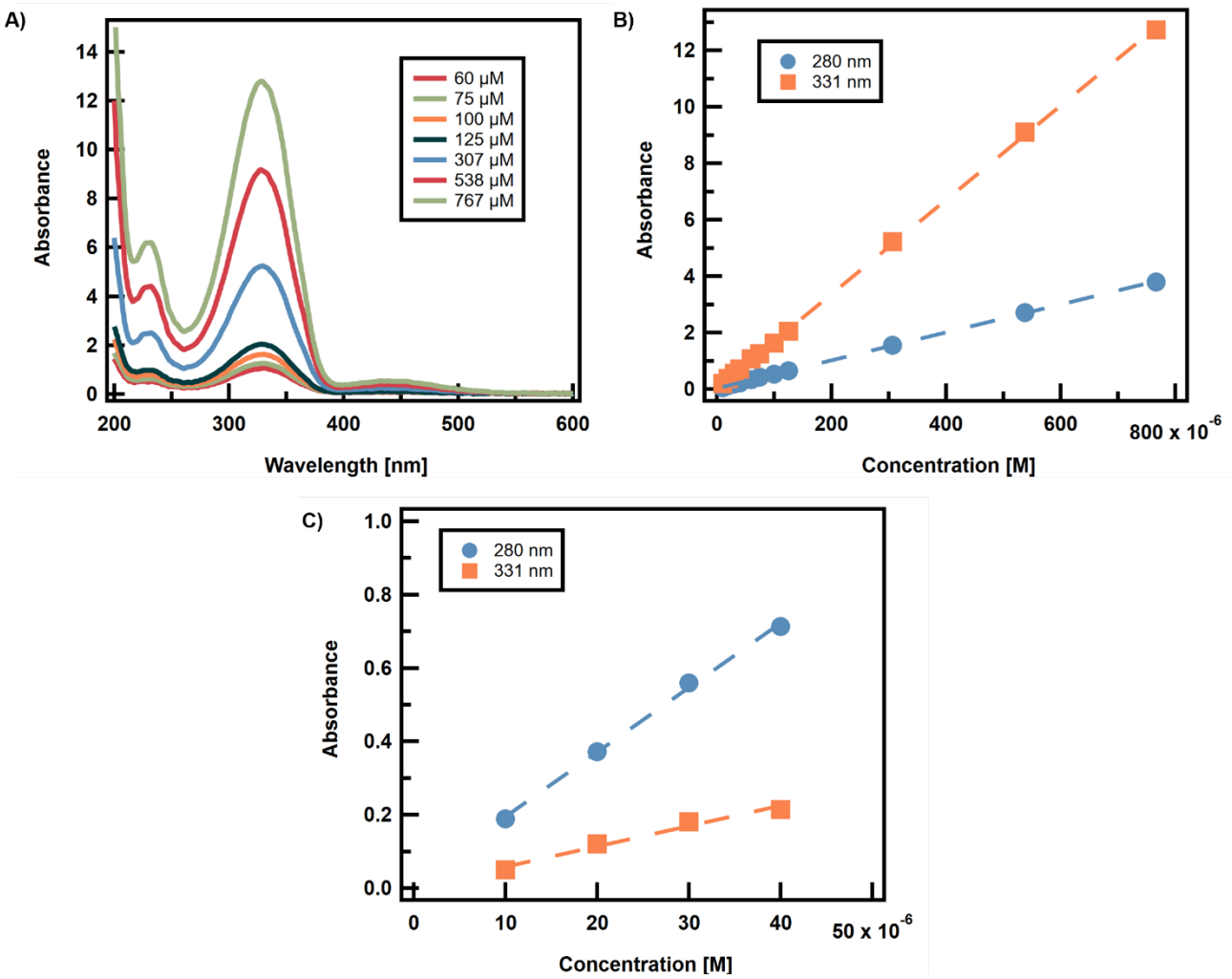


Figure S6. Molar extinction coefficient of azoEDTA in 20 mM, pH 4.50 acetate buffer. Samples were prepared and thermally equilibrated for 3 days. **A)** AzoEDTA solutions had concentrations between 10 and 767 µM [not all shown], and the maximum absorbance wavelength, λ_{max}, is 331 nm. **B)** The absorbance at λ_{max} was fit using Matlab and the molar extinction coefficient ~16,670 1/(M*cm). The absorbance at 280 nm was fit and the molar extinction coefficient ~4,940 (1/M*cm). **C)** For absorbance values between 0 and 1, the absorbance at λmax was fit and the molar extinction coefficient ~17,620 1/(M*cm). The absorbance at 280 nm was fit and the molar extinction coefficient ~5,560 (1/M*cm).

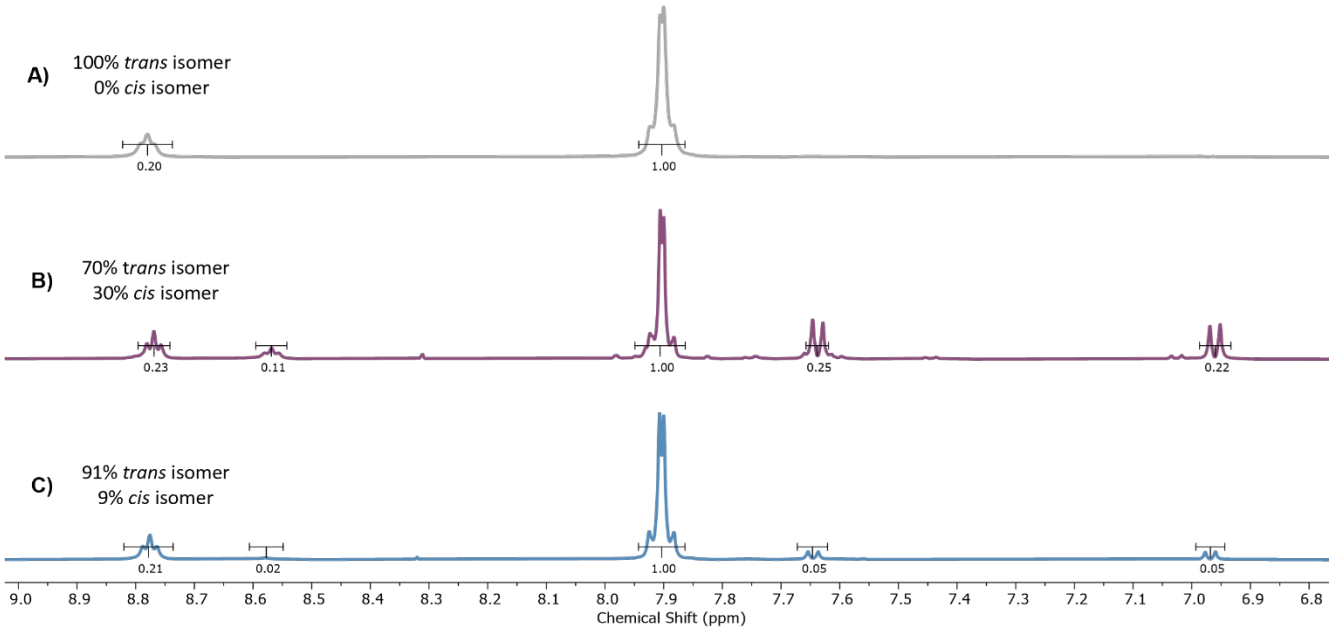


Figure S7. Solution-state ^{1}H NMR spectrum (500 MHz, 10 vol% D_2O) spectra of 3.1 mM azoEDTA in 20 mM acetate buffer (pH 4.50) after **A)** thermal equilibration in the dark for 3 days, **B)** irradiation with 365 nm light for 40 min, and **C)** irradiation with 470 nm light for 15 min.

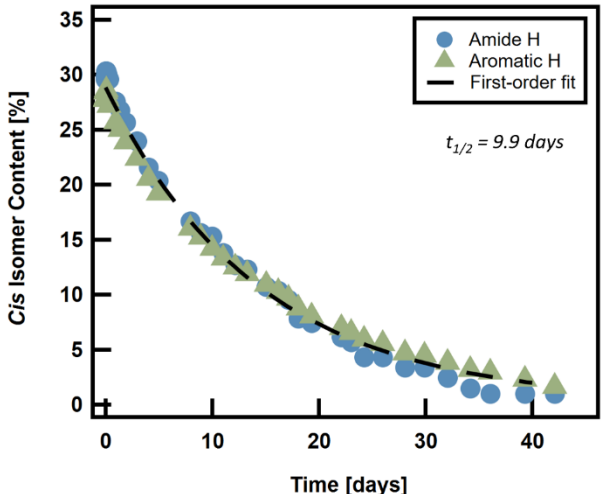


Figure S8. Thermal relaxation of *cis*-azoEDTA to *trans*-azoEDTA follows first-order rate kinetics where k = 0.069 day^{-1}. Percent of *cis* isomer present in the solution was calculated using relative ^{1}H-NMR integrations from a 3.1 mM azoEDTA solution in 20 mM acetate buffer (pH 4.50) with 10 vol% D_2O.

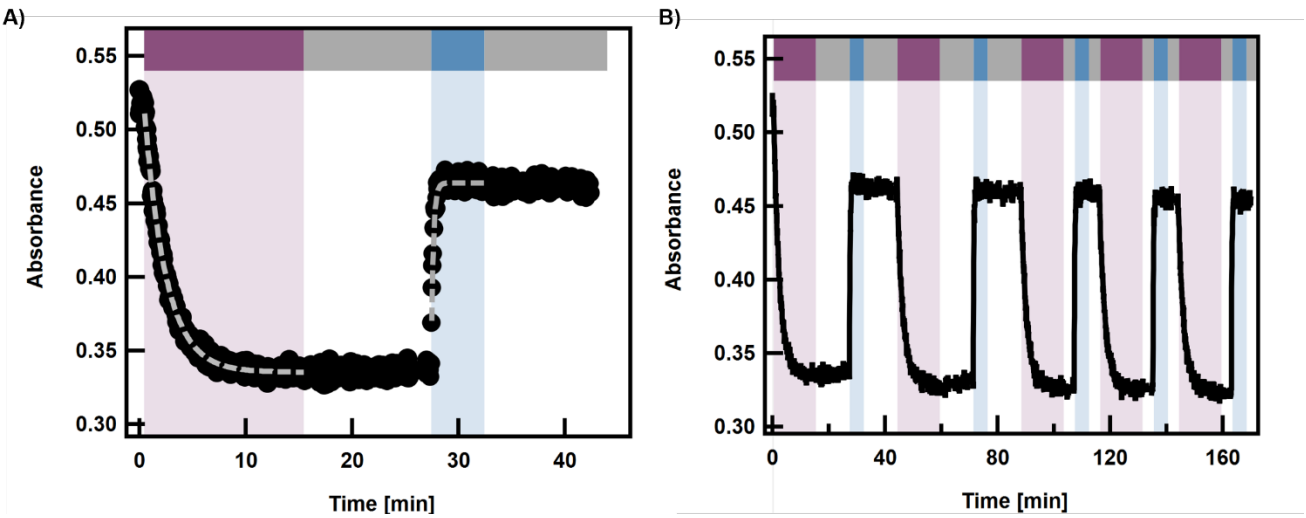


Figure S9. Pump-probe measurements of a 240 µM azoEDTA solution prepared in 20 mM acetate buffer (pH 4.50). The sample was irradiated in a quartz cuvette with 2 mm path length. The purple box represents irradiation with 365 nm light and the blue box represents 470 nm. **A)** First photoisomerization cycle of azoEDTA using pump-probe. The *trans*→*cis* and the *cis*→*trans* absorption data (black dots) were fit in Matlab (gray lines) to extract first-order rate coefficients: $k_{A,365}$ = 0.00082 ± 0.0001 s^{-1} and $k_{A,470}$ = 0.083 ± 0.008 s^{-1}, respectively. **B)** AzoEDTA photoresponse does not show photofatigue over 5 cycles. Rate coefficients are provided in **Table S1** and, for all cycles, were within the uncertainty of those calculated from the first cycle.

Table S1. Photoswitching kinetics of azoEDTA in 20 mM, pH 4.50 acetate buffer. The data were fit to a first-order rate equation in Matlab given by Equation S11. The error given was estimated from the fit of the model.

Cycle	$k_{A,365}$ [min^{-1}]	$k_{A,470}$ [min^{-1}]
1	(8.59±0.1) x 10^{-3}	(9.10±0.7) x 10^{-2}
2	(8.48±0.1) x 10^{-3}	(8.42±0.5) x 10^{-2}
3	(7.85±0.1) x 10^{-3}	(9.03±0.6) x 10^{-2}
4	(8.74±0.1) x 10^{-3}	(8.23±0.4) x 10^{-2}
5	(8.16±0.1) x 10^{-3}	(8.64±0.5) x 10^{-2}

4.3 Size, structure, and properties of reflectin assemblies

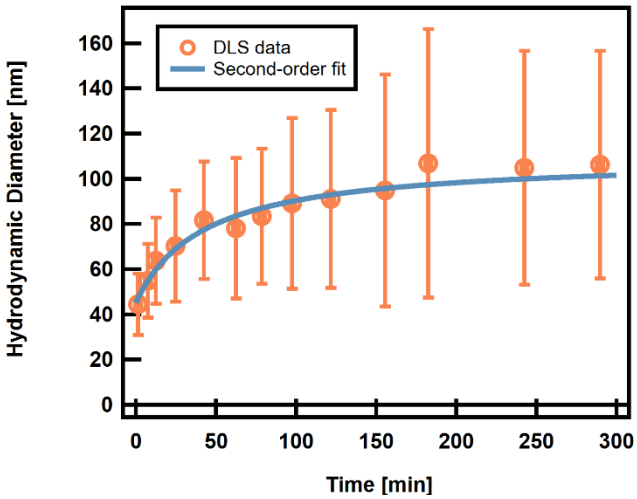


Figure S10. The increase in hydrodynamic diameter ‹D_H› appears to follow second-order kinetics, which is consistent with some IDP-ligand binding models.[16] More rigorous experiments are necessary before confirming this is a second-order process. No significant change in ‹D_H› is observed after 120 min, but the standard deviation continues to increase, indicating protein aggregation. After several days, solution turbidity decreased as larger assemblies/aggregates settled out of solution. Long equilibration times prior to light irradiation limited cyclability and size change of azoEDTA-reflectin assemblies.

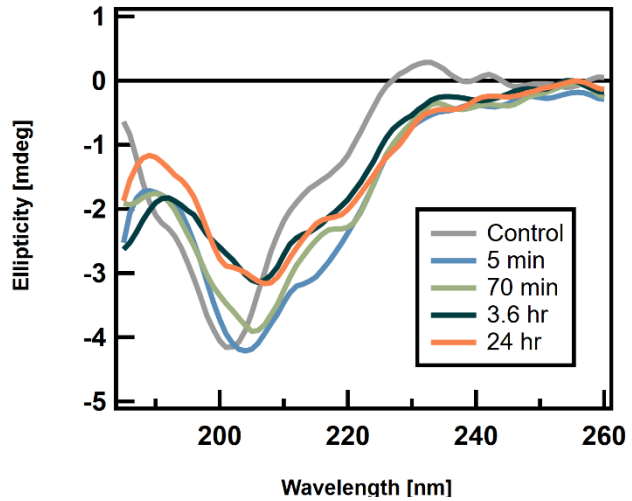


Figure S11. Time-dependent evolution of secondary structure in azoEDTA-reflectin assemblies prepared with 8 µM reflectin and 240 µM azoEDTA in a 20 mM acetate buffer (pH 4.50).

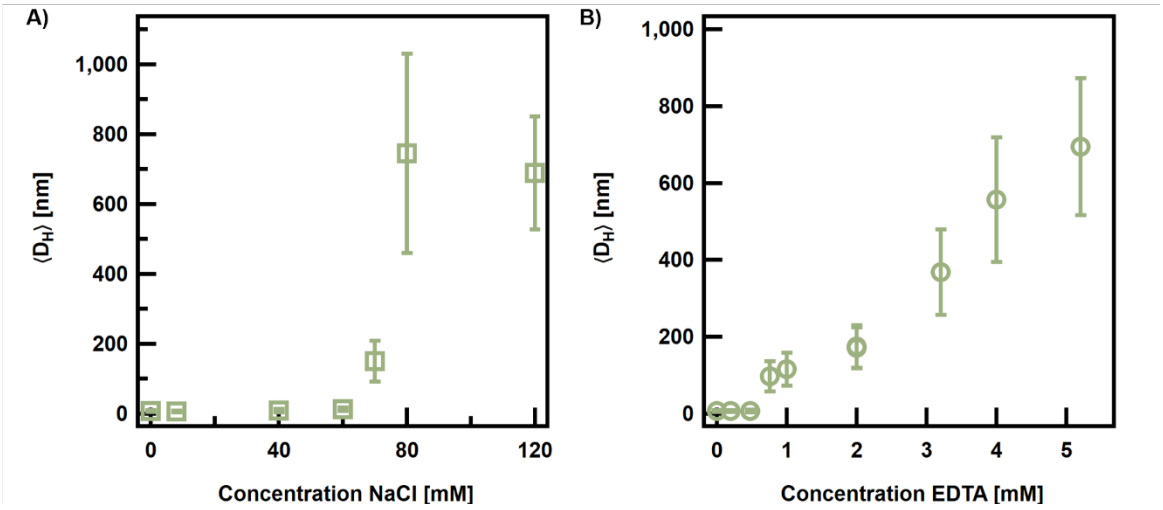


Figure S12. Assembly size for 8 µM reflectin and sodium chloride (NaCl) or ethylenediaminetetraacetic acid (EDTA) in 20 mM acetate buffer (pH 4.50). Hydrodynamic diameter was measured by DLS. **A)** Reflectin-NaCl assemblies form between 40 mM and 60 mM NaCl, which is > 300 times the concentration of azoEDTA needed for assembly and is consistent with reflectin literature.[17] Hydrodynamic diameter was measured by DLS. **B)** Reflectin-EDTA assemblies form between 0.46 mM and 0.78 mM EDTA, which is > 5 times the concentration of azoEDTA needed for assembly. The error bars in both plots represent standard deviation of average hydrodynamic diameter.

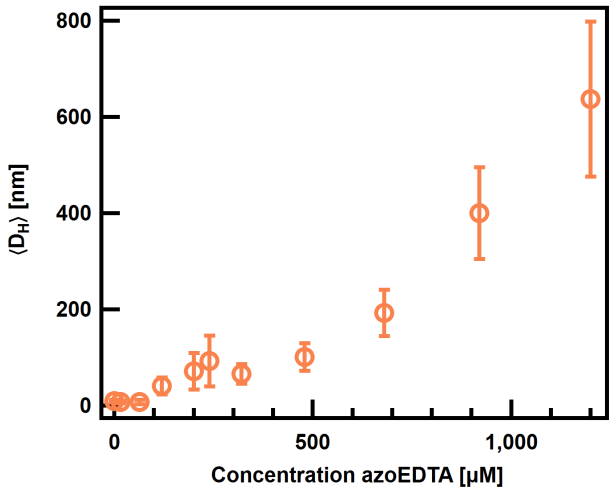


Figure S13. The mean hydrodynamic diameter (‹D_H›) of azoEDTA-reflectin assemblies with concentrations of azoEDTA between 16 – 1,200 µM. Assembly size is controlled at a size of ~100 nm for concentrations < 480 µM. The error bars in both plots represent standard deviation of average hydrodynamic diameter.

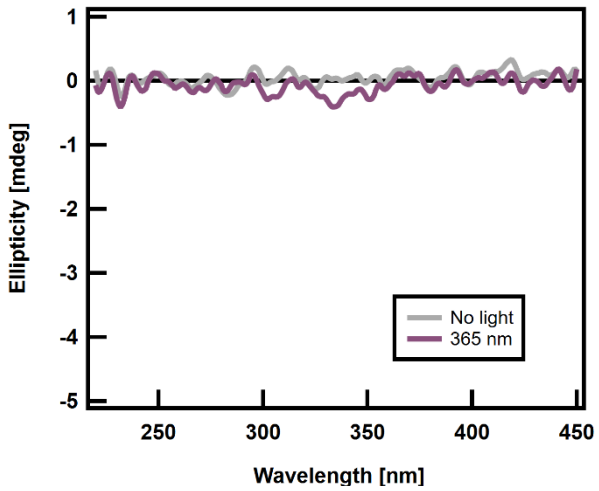


Figure S14. Circular dichroism of 240 µM azoEDTA in 20 mM acetate buffer does not show ellipticity in the near UV or visible light region. This behavior is also consistent after irradiation of 365 nm light.

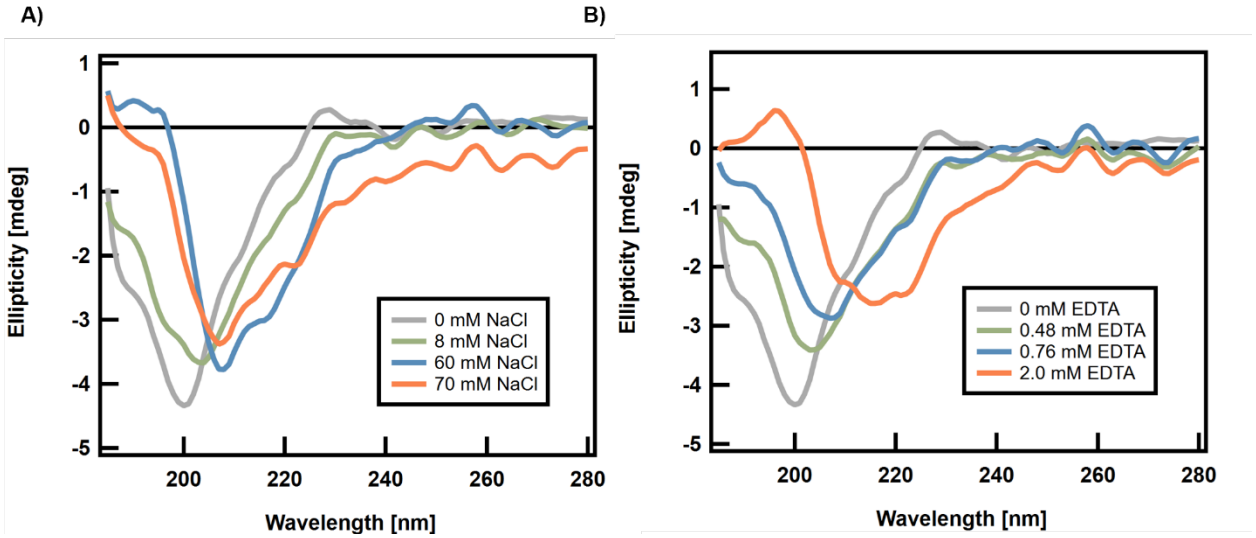


Figure S15. Secondary structure formation for 8 µM reflectin and sodium chloride (NaCl) or ethylenediaminetetraacetic acid (EDTA) in 20 mM acetate buffer (pH 4.50). Measurements were taken with circular dichroism spectroscopy. **A)** As concentration of NaCl increases, the α helix and β sheet character of the system increases. **B)** As concentration of EDTA increases, the α helix and β sheet character of the system increases. When [EDTA] > 2.0 mM, β sheets are the only structural featured measured by CD.

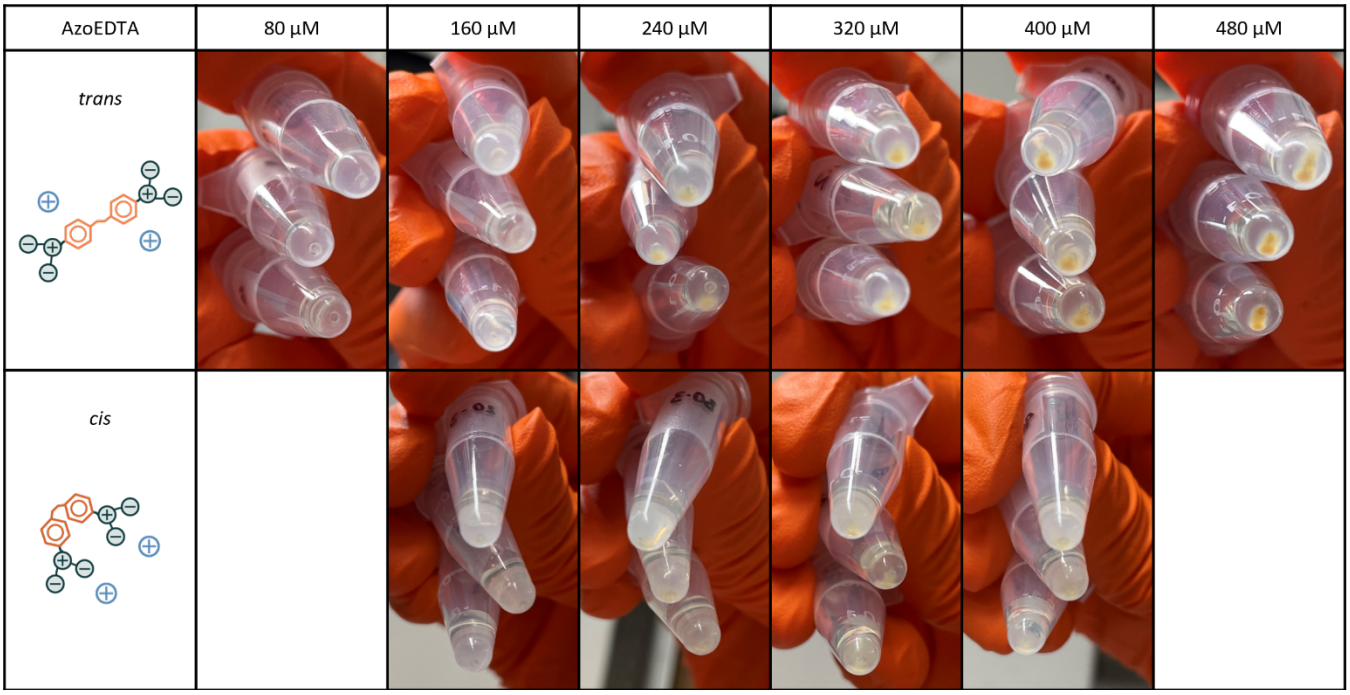


Figure S16. Centrifugation of azoEDTA-reflectin samples precipitated protein complexes as orange pellets.

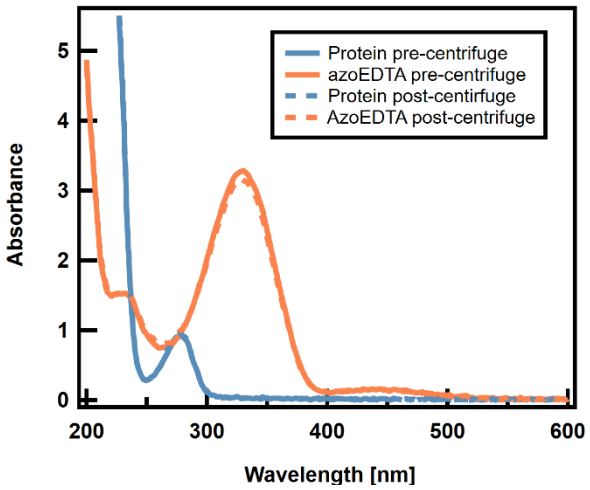


Figure S17. The concentrations of separate solutions of 8 µM reflectin protein and 240 µM azoEDTA in 20 mM acetate buffer do not change after centrifugation. Monomeric reflectin and azoEDTA do not precipitate prior to forming protein assemblies.

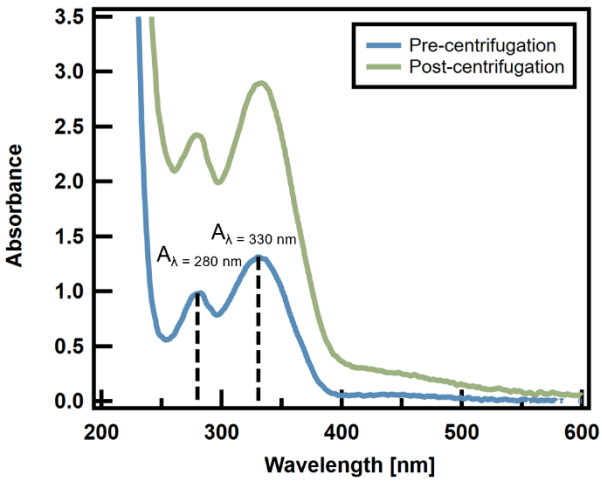


Figure S18. Centrifugation of 8 µM reflectin protein and 160 µM azoEDTA removes solution turbidity by precipitating protein complexes. The absorbance at 280 and 330 nm were used to calculate concentrations of each species in the supernatant and pellet.

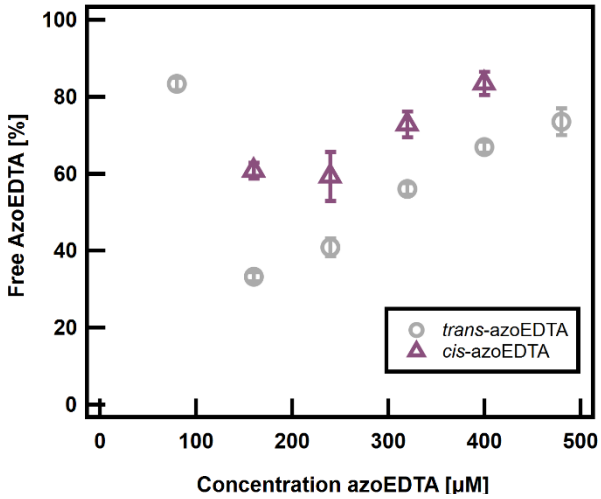


Figure S19. The percentage of unbound, or free, azoEDTA was calculated from the concentration of photoswitch in the supernatant after centrifugation.

4.4 Photoresponsive properties of the reflectin-azoEDTA system

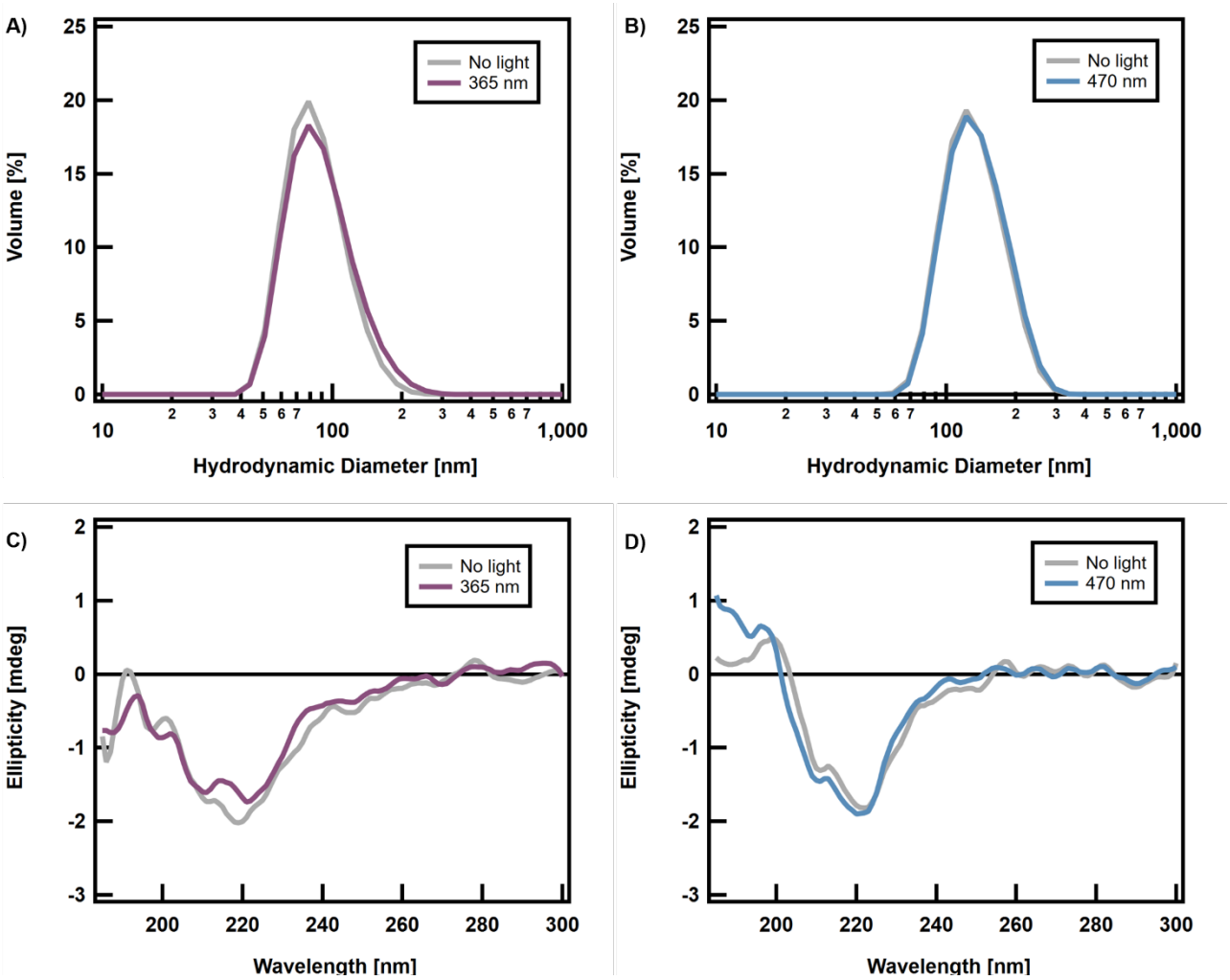


Figure S20. Size and structure of reflectin assemblies are not disrupted by light irradiation in the absence of a photoresponsive molecule. Control experiments were performed on reflectin assemblies dialyzed into 20 mM MOPS buffer (pH 7.50), and samples were diluted to 12 µM reflectin and irradiated for 20 minutes. **A)** The hydrodynamic diameters before and after irradiation with 365 nm light were 87 ± 27 nm and 93 ± 33 nm, respectively. **B)** The hydrodynamic diameters before and after with 470 nm light were 132 ± 44 nm and 145 ± 48 nm, respectively. Samples for experiments **A)** and **B)** were prepared independently. More base was needed to adjust pH of the buffer used to prepare **B)** – larger ionic strength caused a larger ‹D_H›. **C)** No change in secondary structure of reflectin assembly is observed after irradiation with UV light. **D)** No change in secondary structure of reflectin assembly is observed after irradiation with blue light.

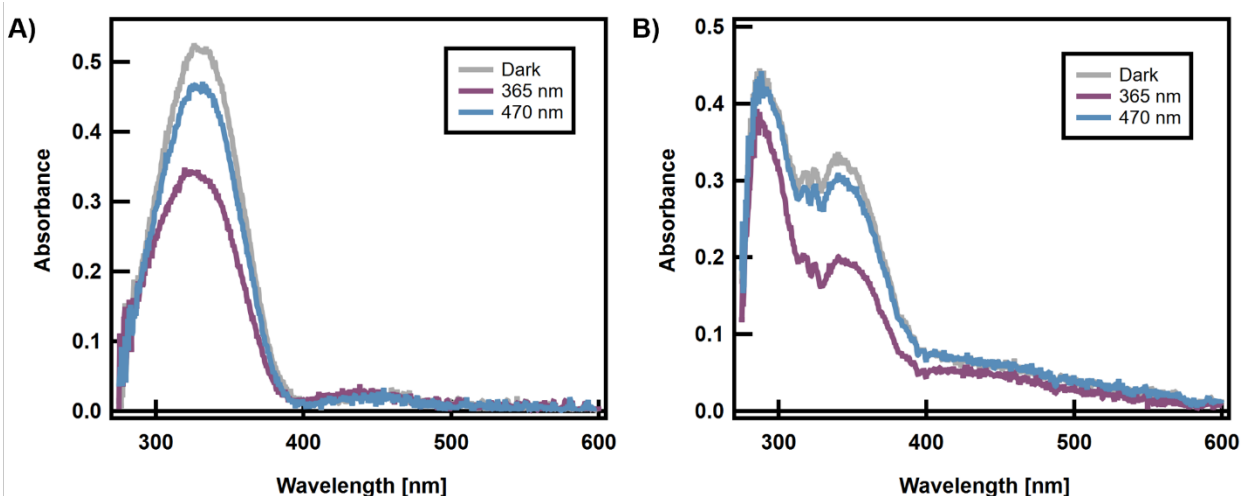


Figure S21. A) Light absorbance spectra for 240 µM azoEDTA in 20 mM acetate buffer (pH 4.50) using the pump-probe instrument. **B)** Absorbance spectra for 240 µM azoEDTA and 8 µM reflectin on the pump probe instrument. The maximum absorbance wavelength is 338 nm and turbidity was observed visually and is shown by the raised baseline for λ > 400 nm. All spectra were acquired using a quartz cuvette with a path length of 2 mm.

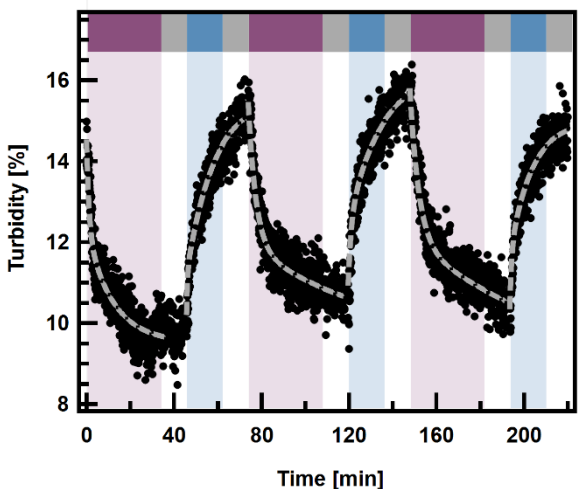


Figure S22. The turbidity of the azoEDTA-reflectin system was calculated using absorbance at 400 nm, where *trans*- and *cis*-azoEDTA have negligible absorbance (<0.01 units on the pump-probe instrument) and measured absorbance is attributed solely to turbidity. The difference in turbidity between this experiment and **Figure 3B** is attributed to the difference in cuvette pathlength.

Table S2. Photoswitching kinetics of 240 µM azoEDTA and 8 µM reflectin in 20 mM, pH 4.50 acetate buffer. The data were fit to a biexponential equation in Matlab given by Equation S12. The error given is from the fit of the model.

Cycle	$k_{A,365}$ [min^{-1}]	$k_{R,365}$ [min^{-1}]	$k_{A,470}$ [min^{-1}]	$k_{R,470}$ [min^{-1}]
1	(1.28±0.07) x 10^{-2}	(3.93±0.3) x 10^{-4}	(6.14±0.7) x 10^{-2}	(9.94±1) x 10^{-4}
2	(6.60±0.2) x 10^{-3}	(7.01±0.6) x 10^{-5}	(5.23±0.6) x 10^{-2}	(5.04±0.6) x 10^{-4}
3	(6.41±0.2) x 10^{-3}	(7.14±0.7) x 10^{-5}	(5.69±0.7) x 10^{-2}	(4.40±0.5) x 10^{-4}

Table S3. Diffusion coefficients of monomeric reflectin (8 µM) and reflectin-azoEDTA assemblies (240 µM azoEDTA, 8 µM reflectin) in 20 mM, pH 4.50 acetate buffer measured by DLS. Reflectin-azoEDTA assembly diffusion coefficients correspond to assembly size data for *in situ* irradiation experiments shown in **Figure 5C**.

Sample	Irradiation	D [$\mu m^2\ s^{-1}$]
Monomeric reflectin	None	12.06±0.84
Reflectin-azoEDTA assembly	None	4.40±0.04
Reflectin-azoEDTA assembly	365 nm	6.87±0.06
Reflectin-azoEDTA assembly	470 nm	5.07±0.05

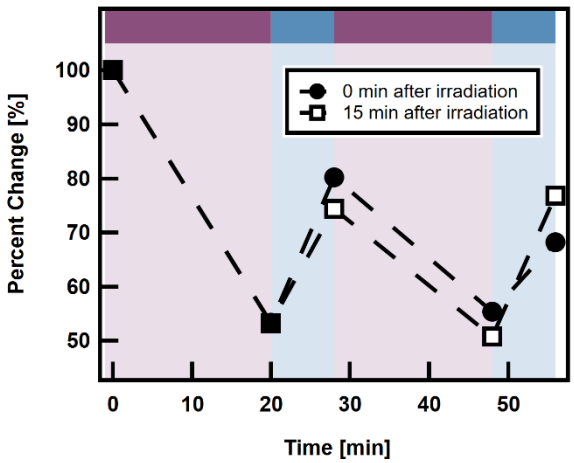


Figure S23. The change in hydrodynamic diameter is plotted as a percent change from the initial diameter size. When the sample is in the dark for 15 min after irradiation has completed, the system shows less decay in protein assembly size than if DLS measurements were taken immediately. The purple box represents irradiation with 365 nm light and the blue box represents 470 nm. The time between irradiation and DLS measurement is not shown for simplicity.

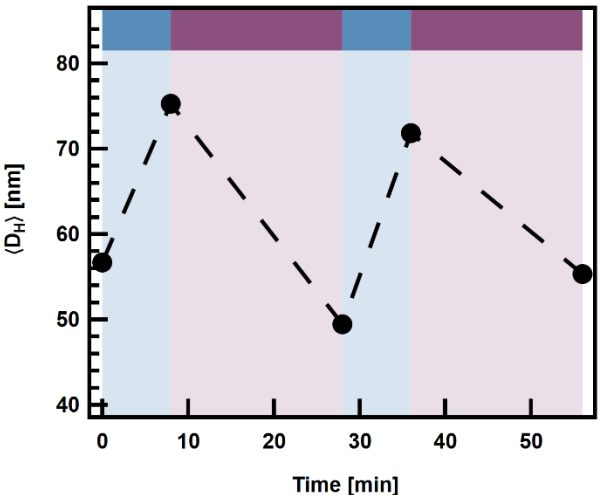


Figure S24. When protein assemblies are formed with reflectin and *cis*-azoEDTA, the mean hydrodynamic diameter, ‹D_H›, cycles twice with 470 nm (blue box) and 365 nm (purple box) light.

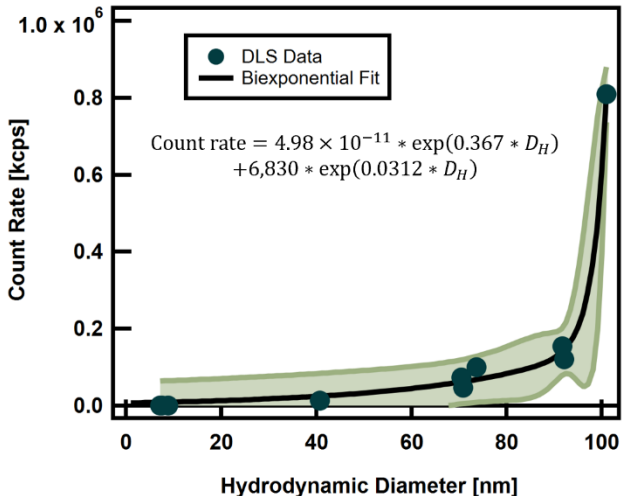


Figure S25. The nonlinear relationship between derived count rate and hydrodynamic diameter (D_H) for the azoEDTA-reflectin system was fit to a biexponential model in Matlab. The shaded green region is the 95% confidence interval of the fit, calculated by Matlab.

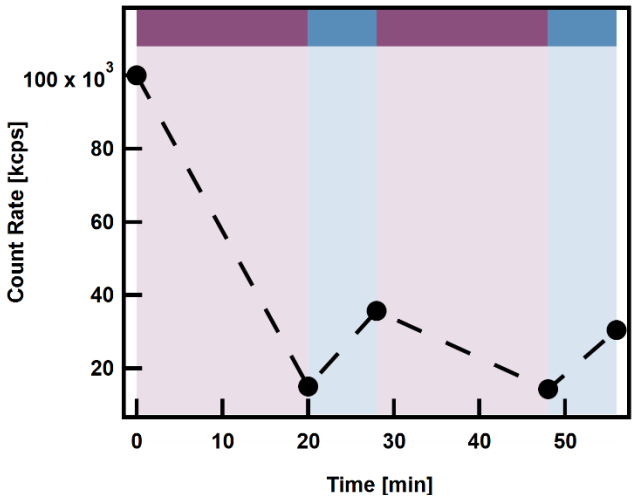


Figure S26. The derived count rate measured by DLS cycles with 365 nm (purple box) and 470 nm (blue box) light. The cycling behavior shown is consistent with the nonlinear relationship between count rate and diameter, indicating protein assemblies are not sedimenting out of solution.

5. Synthesis Characterization

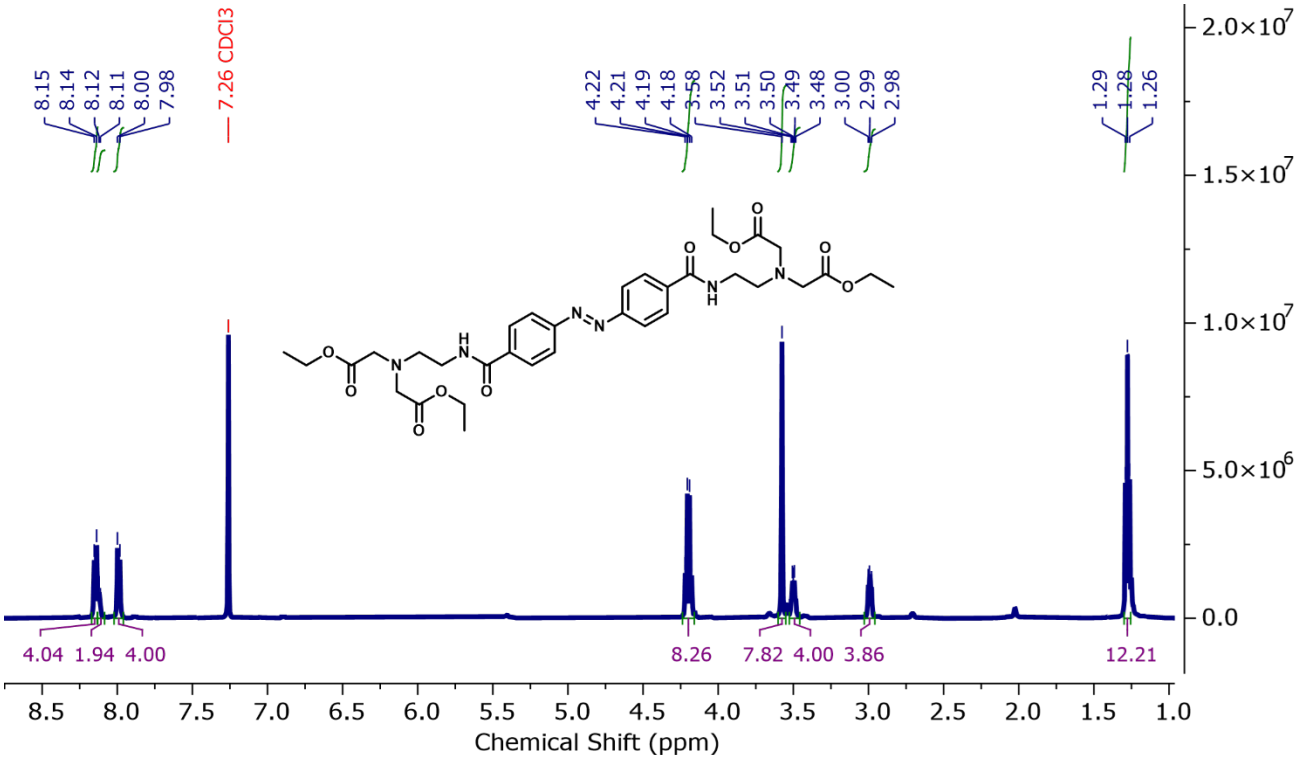


Figure S27. ^{1}H NMR (500 MHz, $CDCl_3$) compound **6**.

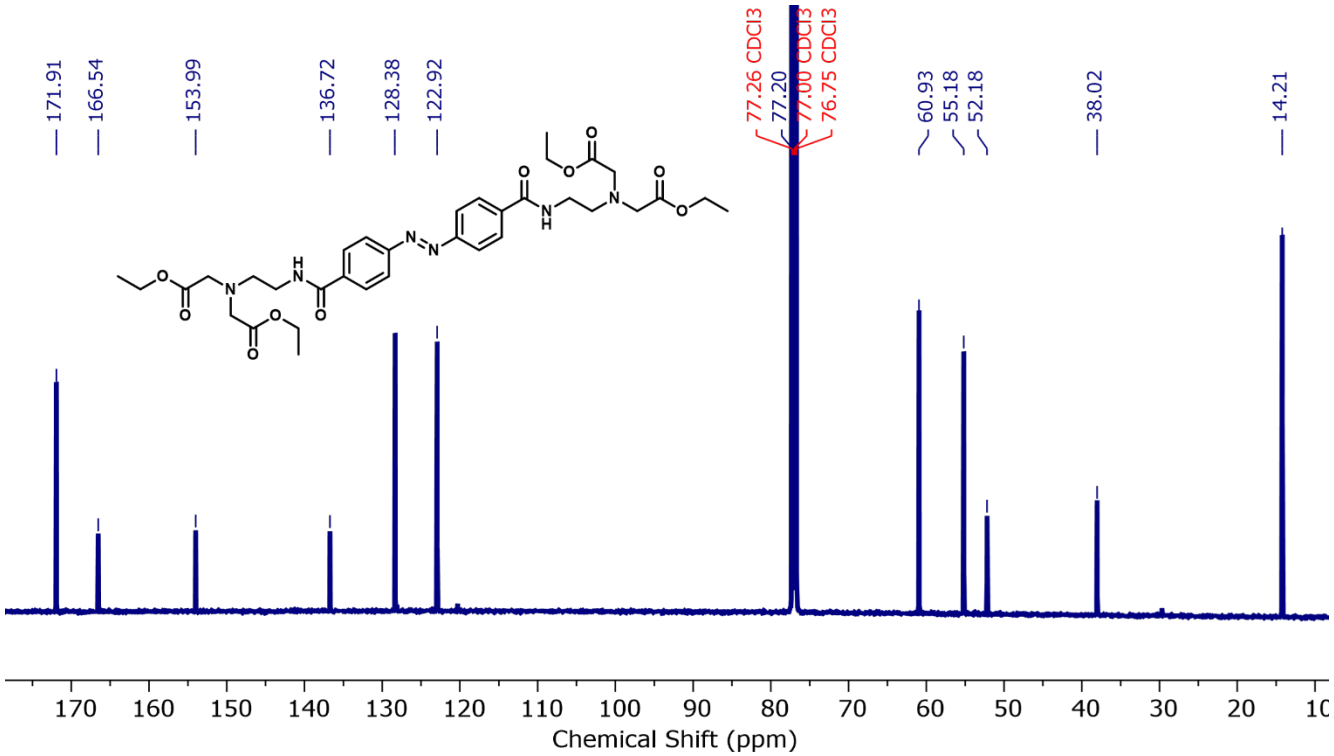


Figure S28. ^{13}C NMR (125 MHz, $CDCl_3$) compound **6**.

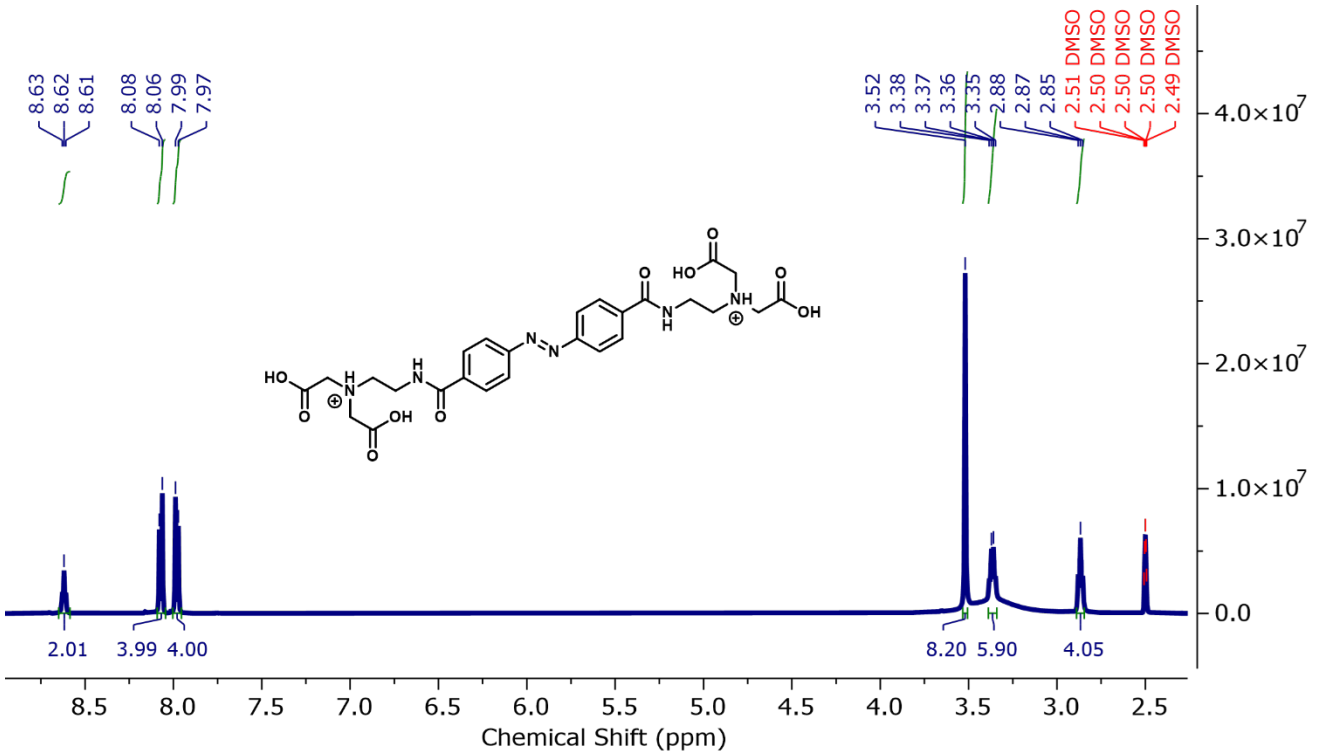


Figure S29. ^{1}H NMR (500 MHz, DMSO-d_6) compound **7**. (Peak at 3.37 ppm partially obscured by moisture in NMR solvent)

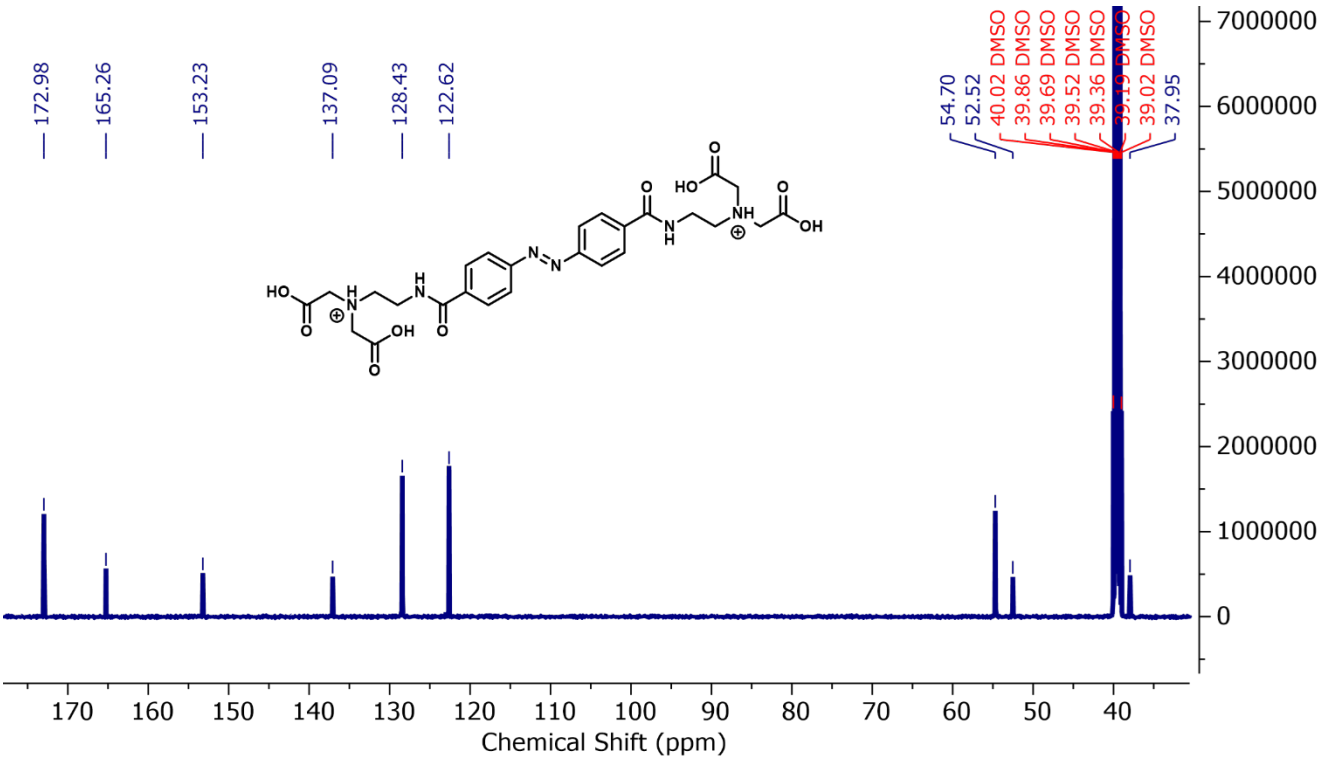


Figure S30. ^{13}C NMR (125 MHz, DMSO-d_6) compound **7**.

6. References

1 G. Bentzinger, W. De Souza, C. Mullié, P. Agnamey, A. Dassonville-Klimpt and P. Sonnet, *Tetrahedron Asymmetry*, 2016, **27**, 1–11.

2 S. Okumura, C. H. Lin, Y. Takeda and S. Minakata, *J. Org. Chem.*, 2013, **78**, 12090–12105.

3 J. Hatai, S. Pal and S. Bandyopadhyay, *Tetrahedron Lett.*, 2012, **53**, 4357–4360.

4 M. Dudek, Z. Pokładek, M. Deiana and K. Matczyszyn, *Dyes Pigments*, 2020, **180**, 108501.

5 J. Hatai and S. Bandyopadhyay, *Chem. Eur. J.*, 2014, **20**, 10020–10026.

6 R. Levenson, C. Bracken, C. Sharma, J. Santos, C. Arata, B. Malady and D. E. Morse, *J. Biol. Chem.*, 2019, **294**, 16804–16815.

7 F. Stricker, J. Peterson, S. K. Sandlass, A. de Tagyos, M. Sroda, S. Seshadri, M. J. Gordon and J. Read de Alaniz, *Chem.*, 2023, 9, 1994–2005.

8 A. M. Alfaraidi, B. Kudisch, N. Ni, J. Thomas, T. Y. George, K. Rajabimoghadam, H. J. Jiang, D. G. Nocera, M. J. Aziz and R. Y. Liu, *J. Am. Chem. Soc.*, 2023, **145**, 26720–26727.

9 T. A. Grusenmeyer, B. A. McClure, C. J. Ziegler and J. J. Rack, *Inorg. Chem.*, 2010, **49**, 4466–4470.

10 N. K. Marynich, M. G. Khrenova, A. V. Gavshina, I. D. Solovyev and A. P. Savitsky, *Sci. Rep.*, 2022, **12**.

11 W. Y. Yang and M. Gruebele, *Nature*, 2003, **423**, 193–197.

12 R. M. Dickson, A. B. Cubitt, R. Y. Tsien and W. E. Moerner, *Nature*, 1997, **388**, 355–358.

13 P. Lecca, M. Lecca, C. A. Maestri, A. Maestri and M. Scarpa, *Math. Models Methods Appl. Sci.*, 2021, **44**, 10154–10171.

14 R. L. Thurlkill, G. R. Grimsley, J. M. Scholtz and C. N. Pace, *Protein Sci.*, 2006, **15**, 1214–1218.

15 D. C. Elliot, W. H. Elliott and K. M. Jones, *Data for biochemical research*, Oxford University Publisher, New York, 2nd edn., 1959.

16 V. Iešmantavičius, J. Dogan, P. Jemth, K. Teilum and M. Kjaergaard, *Angew. Chem.*, 2014, **53**, 1548–1551.

17 R. Levenson, B. Malady, T. Lee, Y. Al, P. Kohl, Y. Li, D. E. Morse, R. Levenson, B. Malady, T. Lee, Y. Al Sabeh and P. D. Morse, *preprint bioRxiv*, 2021, DOI: 10.1101/2021.04.23.441158.