A Genetically Encoded Cyclobutene Probe for Labelling of Live Cells

Electronic Supporting Information, Section 1

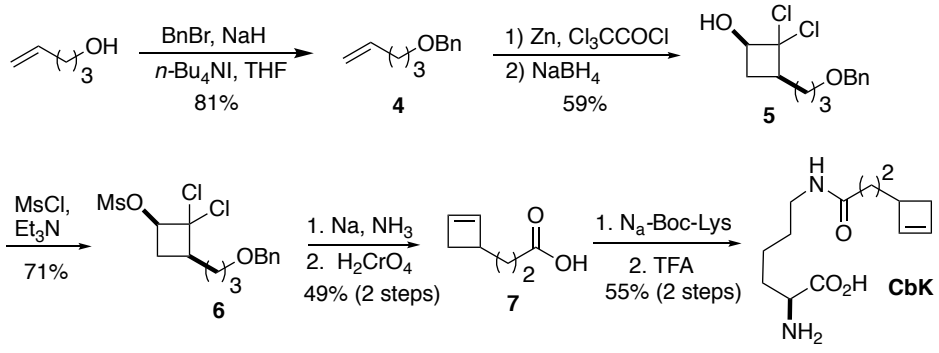
I. Materials and methods

Unless noted, starting materials, solvents, and reagents for chemical synthesis were obtained from commercial suppliers (Acros, Alfa Aesar, Sigma-Aldrich, Chem-impex) and used without further purification. Dry solvents were used as purchased except for use as reaction solvents: THF and ether (distilled from Na/benzophenone); CH_2Cl_2 and CH_3CN (distilled from CaH_2) Deuterated solvents were obtained from Sigma-Aldrich. Reactions are run at room temperature unless otherwise indicated. Flash chromatography was carried out on 230-400 mesh silica gel (Silicycle). Thin layer chromatography (TLC) was performed on glass-backed, precoated silica gel plates (0.25 mm GHLF, Analtech); plates were visualized using a handheld UV lamp and/or by one or more stains: ammonium molybdate/ceric sulfate in 10% sulfuric acid (general, after heating); aq. $KMnO_4$ (alkenes), or a solution of vanillin in ethanol and concentrated sulfuric acid (general, after heating). NMR spectra were recorded at room temperature at 400 MHz (^1H) or 100 MHz Bruker Advance III-HD 400 MHz NMR and are reported as: chemical shift (integration, multiplicity, J coupling constants if relevant). Chemical shifts are reported in ppm using ^1H or ^{13}C peaks with solvent used as internal standard: $CDCl_3$, residual H at 7.26, C 77.0; DMSO-d_6, residual H at 2.50, C 39.5;

D_2O, HOD at 4.79. Multiplicity was reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, b = broad. IR spectra were acquired on neat films (diamond, ATR mode) with selected absorbances reported in wavenumbers (cm^{-1}). UV absorbance measurements for kinetic studies were conducted on a Shimadzu UV2401-PC. Absorbance spectrum and intensity were measured on Shimadzu UV2401-PC and Thermo Scientific GENESYS 10S UV/Vis Spectrophotometer. Fluorescence spectra and intensities were recorded on Horiba FluoroMax 4 spectrometer and BioTek Synergy H1 Hybrid plate reader. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed on Bio-Rad mini-PROTEAN electrophoresis system. Bio-Rad Prestained Protein Ladder was applied to at least one lane of each gel for the estimation of apparent molecular weights. Protein gels were stained by Coomassie Brilliant Blue staining and visualized using Bio-Rad Molecular Imager ChemiDoc XRS+ System. For in-gel fluorescence imaging, either Bio-Rad Molecular Imager ChemiDoc XRS+ System or GE Typhoon FLA9500 was used. Live cells were imaged on an Olympus FV500 inverted (Olympus IX-81) confocal microscope. Abbreviations: tetrahydrofuran (THF), hexanes (Hex), ethyl acetate (EA), diethyl ether (ether), PBS (phosphate-buffered saline); rt (room temperature), rbf (round bottom flask).

II. Synthesis of 3-(cyclobuten-2-yl)propanoic acid (CbK)

The cyclobutene-substituted acid was prepared using a modification of a route described by Sittiwong, et al;[1] the current route carried in the carboxyl of **7** as a benzyl ether rather than the benzyl ester employed in the previous report (SI-1, Scheme 1).



SI-1, Scheme 1: Synthesis of 3-(cyclobuten-2-yl)propanoic acid (**CbK**)

((Pent-4-en-1-yloxy)methyl)benzene (4). A NaH suspension (1.1307 g; 28.28 mmol; 60% in oil) was placed in a flame-dried flask under N_2 and washed twice with pentane. The resulting powder was resuspended in THF (180 mL) and Bu_4NI (1.3921 g; 3.77 mmol) and 4-penten-1-ol (1.5152 g; 17.59 mmol) were sequentially added, the latter dropwise. The suspension was stirred at rt for 30 min, and then benzyl bromide (0.1500 g, 1.19 mmol) was added. The reaction was stirred for 12 h and then quenched with water (200 mL; initially dropwise, then more rapidly once evolution of heat ceased). The emulsion was extracted with Hex (2x100 mL) and then EA (3x100 mL). The combined organic layers were dried over Na_2SO_4 and the residue obtained after concentration under reduced pressure was purified by flash chromatography (1% ether/pentane) to give the benzyl ether **4** (2.5207 g, 81%) as a clear liquid. R_f = 0.77 (25% EA/Hex); ^{1}H NMR δ 1.74 (pentet, 2H, 7.0), 2.17 (dd, 2H, 13.6, 6.3), 3.507 (t, 2H, 6.4), 4.52 (s, 2H), 4.980 (dd, 1H, 10.2, 1.1), 5.04 (dd, 1H, 1.5, 17.1) 5.84 (ddt, 1H, 17.0, 11.1, 6.6), 7.29-7.37 (5H); ^{13}C NMR δ 29.1, 30.5, 69.9, 73.0, 114.9, 127.6, 127.8, 128.5, 138.4, 138.8; HRMS (EI): calcd for $C_{12}H_{16}O$ (M+): 176.1201; found: 176.1198; IR: 2850, 1449, 1201, 1092, 1068, 1027.

cis-**3-(3-(Benzyloxy)propyl)-2,2-dichlorocyclobutan-1-ol (5).** Into a solution of pentenyl ether **1** (3.1295 g; 17.76 mmol) in dry ether (130 mL) in a flame-dried 3N under N_2 was added Zn (fine mesh, 5.8342 g; 89.19 mmol). A reflux condenser was placed on the flask and the suspension was warmed to 30 $^\circ$C while a solution of trichloroacetyl chloride (9.72 g; 53.5 mmol) in dry ether (50 mL) was added dropwise via an addition funnel. The reaction was stirred for ~ 1 h, at which point it began to reflux. The heating bath was removed and the reaction was cooled in an ice bath. After stirring for an additional 10 min, the reaction was poured onto ice, and the resulting mixture was extracted with EA (3 x 200 mL). The combined organic layers were dried over Na_2SO_4 then concentrated under reduced pressure. The intermediate dichloroketone, although observable by TLC, was not isolated but applied directly to the next step.

The crude product of the cycloaddition was diluted in isopropanol (260 mL) and the solution cooled in an ice bath. Sodium borohydride (1.2112 g; 32.02 mmol) was added and then the ice bath was removed. The reaction was allowed to stir for 30 min and then quenched with sat. aq.

$NaHCO_3$ (250 mL). The mixture was extracted with Hex (6 x 100 mL) and the combined organic layers were dried over Na_2SO_4. The residue obtained by concentration *in vacuo* was purified by flash chromatography (15% EA/hexane) to give mainly the *cis*-dichlorocyclobutanol (**5**, 3.0203 g, 59% yield) as a thick yellow oil. The NMR showed traces of inseparable impurities which appeared to be the *trans* dichlorocyclobutanol generated during the hydride reduction and/or the regiosiomeric dichlorocyclobutanol resulting from an alternate mode of cycloaddition. The presence of traces of these species are unimportant as both will eventually converted to the same intermediate (**7**) as the major dichlorocyclobutanol (**5**). R_f = 0.38 (25% EA/Hex); ^{1}H NMR δ 1.48 (m, 1H), 1.642 (m, 2H), 1.69 (m, 2H), 1.78 (m, 2H), 2.462 (m, 2H), 2.51 (d, 1H, 10.8), 3.51 (dd, 2H, 9.9, 5.6), 4.30 (m, 1H), 4.52 (s, 2H), 7.35 (5H); ^{13}C NMR δ 26.8, 27.0, 34.8, 45.6, 69.9, 73.0, 75.6, 93.0, 127.7, 127.8, 128.5, 138.6; HRMS (EI): calcd for $C_{14}H_{18}Cl_2O_2$ (M+): 288.0684; found: 288.0697; IR: 3383, 2943, 2856, 1495, 1452, 1362, 1192, 1163, 1094, 1074, 1027.

cis-**3-(3-(Benzyloxy)propyl)-2,2-dichlorocyclobutyl methanesulfonate (6).** To a solution of cyclobutanol **2** (5.2748 g; 18.24 mmol) in dry CH_2Cl_2 (100 mL) in a flame-dried flask under N_2 was added methanesulfonyl chloride (2.96 g; 25.8 mmol). The mixture was cooled to -78 °C and triethylamine (3.15 g; 31.1 mmol) was added dropwise. The suspension was allowed to warm to rt and stirred for 1 h, then quenched with water (30 mL), and 1M aq. HCl (100 mL). The mixture was extracted with CH_2Cl_2 (4 x 100 mL). The combined organic layers were dried over Na_2SO_4 then evaporated under reduced pressure. The residue was purified by flash chromatography (12% EA/hexane) to give methanesulfonate **6** (4.7337 g, 71% yield) as a thick yellow oil: R_f = 0.25 (25% EA/Hex); ^{1}H NMR δ 1.789 (m, 5H), 2.32 (m, 2H), 3.20 (s, 3H), 3.51 (ddd, 2H, 5.9, 3.0, 2.6), 4.52 (s, 2H), 5.12 (t, 1H, 8.7), 7.35-7.36 (5H); ^{13}C NMR δ 26.9, 26.9, 31.7, 39.6, 45.9, 69.7, 73.1, 78.5, 88.4, 127.8, 128.5, 138.5; HRMS (ESI): calcd for $C_{15}H_{20}Cl_2O_4S$ (M+): 366.0460; found: 366.0455; IR: 2940, 2860, 1453, 1361, 1361, 1178, 1147, 1097, 1061, 1027.

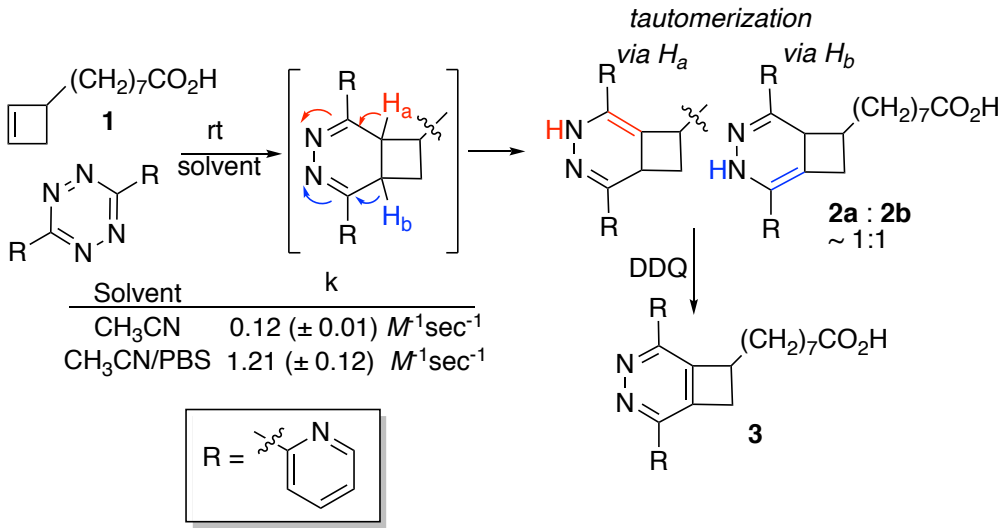
3-(Cyclobut-2-en-1-yl)propanoic acid (7). Sodium (3.53 g; 153 mmol; cut from a block and washed with hexane) was added to a -78 °C solution of liquid ammonia (172 mL) at -78 °C in a

flame-dried flask under N_2. To the resulting blue solution was added a solution of methanesulfonate **3** (3.1667 g; 8.62 mmol) in dry THF (2 mL). The suspension was stirred at -78 °C for 1 h, after which additional sodium (2.06 g; 89.6 g) was added. The reaction was stirred for an additional 30 min and then slowly quenched with sat. aq. NH_4Cl. (100 mL). The mixture was allowed to warm to rt over 5 h, during which time the ammonia was allowed to evaporate. The resulting suspension was diluted with aq. 1M HCl (100 mL), and the mixture extracted with ether (4 x 50 mL). The combined organic layers were dried over Na_2SO_4 and the majority of the solvent was carefully removed by concentration at 80 mm. The majority of the crude alcohol was subjected to oxidation (next step) without purification. A small portion of the volatile alcohol was purified by flash chromatography (20% ether/pentane) and/or microscale bulb-to-bulb distillation (75 °C at 30 mm) for NMR analysis: R_f = 0.31 (25% EA/Hex); ^{1}H NMR δ 1.57 (m, 4H), 2.07 (d, 1H, 13.4), 2.67 (dd, 1H, 13.7, 4.2), 2.82 (ddd, 1H, 6.1, 4.2), 3.66 (t, 2H, 6.3), 6.06 (d, 1H, 2.6), 6.11 (d, 1H, 2.4); ^{13}C NMR δ 30.8, 31.3, 36.9, 43.9, 63.2, 135.5, 140.8; IR: 3314, 3037, 2914, 2846, 1451, 1325, 1288, 1052, 1008, 918, 902, 694.

The partially concentrated solution of alcohol obtained in the previous step was dissolved in acetone (105 mL). Into the 0 °C solution was dropwise added chromic acid in aqueous acetone (Jones' reagent, ~2M, 21 mL). The suspension was stirred for 10 min and then extracted with pentane (8 x 75 mL). The combined extracts were washed with 1M NaOH (4x50 mL). The combined alkaline washes were cooled in an ice bath and acidified (6 M HCl) to a pH of ~ 1. The solution was extracted with ether (5 x 150 mL) and the combined organic layers were dried over Na_2SO_4 and then concentrated under reduced pressure. The residue was purified by flash chromatography on acidified silica (230-400 mesh silica gel rinsed with 0.1M sulfuric acid, and then dried overnight at 120 °C) with 50% ether/pentane to provide 3-(cyclobut-2-en-1-yl)propanoic acid **7** (0.5343 g, 49% over two steps) as a light yellow liquid: R_f = 0.22 (25% EA/Hex); ^{1}H NMR δ 1.83 (q, 2H, 7.2), 2.10 (d, 1H, 13.9), 2.39 (t, 2H, 7.2), 2.68 (dd, 1H, 13.9, 4.5), 2.86 (ddd, 1H, 7.2, 4.6, 4.6), 6.084 (m, 2H), 11.51 (s, 1H); ^{13}C NMR δ 29.4, 32.5, 36.4, 43.2, 136.0, 139.9, 180.6; HRMS (EI): calcd for $C_7H_{10}O_2$ (M+): 126.0681; found: 126.0683; IR: 2915, 1702, 1410, 1274, 1216.

***N*$_6$-(3-(Cyclobut-2-en-1-yl)propionyl)-*L*-lysine (CbK).** To an ice-bath cooled solution of cyclobutene derived acid (252 mg, 2 mmol, 1 eq.) in anhydrous CH_2Cl_2 (10 mL) was slowly added oxalyl chloride (171 uL, 5 mmol, 2.5 eq.), followed by 1 drop of *N*,*N*-dimethylformamide. The solution was allowed to warm to rt and stirred for 2 h. After careful removal of solvent and volatile byproducts under reduced pressure (~ 50 mm), the residue was dissolved in 1 mL THF and added to a 0 °C solution of N_α-Boc-*L*-lysine (592 mg, 2 mmol, 1 eq.) in a mixture of 1 *N* NaOH (5 mL) and THF (5 mL). The reaction was stirred overnight at rt and then recooled to 0°C and washed with cold ether (3x5 mL). The aqueous layer was then acidified to between pH 1 and 2 and extracted with EA (3x20 mL). The combined organic layers were dried over Na_2SO_4 and concentrated under reduced pressure. The residue was dissolved in a mixture of CH_2Cl_2 (5 mL) and trifluoroacetic acid (5 mL) and the reaction stirred for 1 h. The residue obtained after concentration was dissolved in minimal amount of methanol and the product precipitated by addition of ether. The collected precipitate was dried under vacuum to furnish (279 mg, 55% yield): ^{1}H NMR (D_2O/NaOH) δ 6.03-6.01 (m, 2H), 3.07-3.05 (m, 3H), 2.71-2.68 (m, 1H), 2.58-2.54 (m, 1H), 2.18-2.15 (m, 2H), 2.01-1.97(m, 1H), 1.70-1.62(m, 2H), 1.53-1.43(m, 4H), 1.30-1.25 (m, 2H); ^{13}C NMR (D_2O/NaOH) δ 183.6, 176.9, 140.6, 136.3, 56.1, 43.0, 39.4, 36.1, 34.5, 30.4, 28.4, 22.6; HRMS (ESI) calc. for [$C_{13}H_{22}N_2O_3$+H]: 255.1703; found: 255.1702.

III. Model studies of inverse electron-demand Diels-Alder (IEDDA)



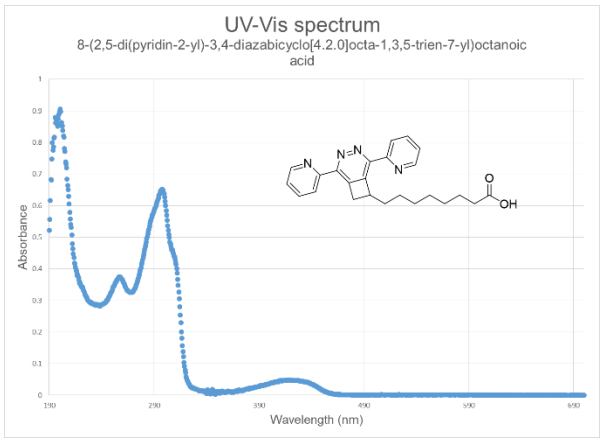
SI-1, Scheme 2. Model IEDDA

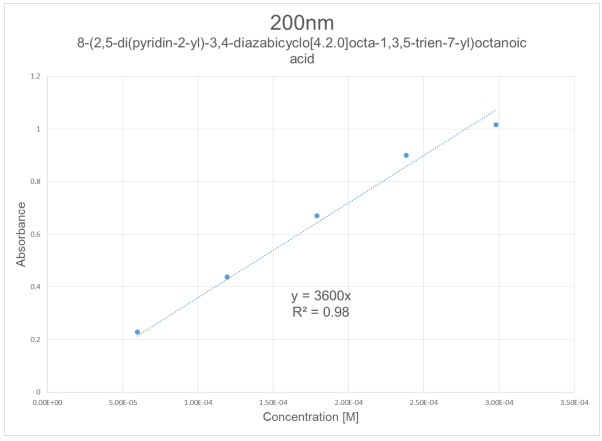
3-(Cyclobut-2-en-1-yl)octanoic acid (1) was prepared using a published procedure.[1]

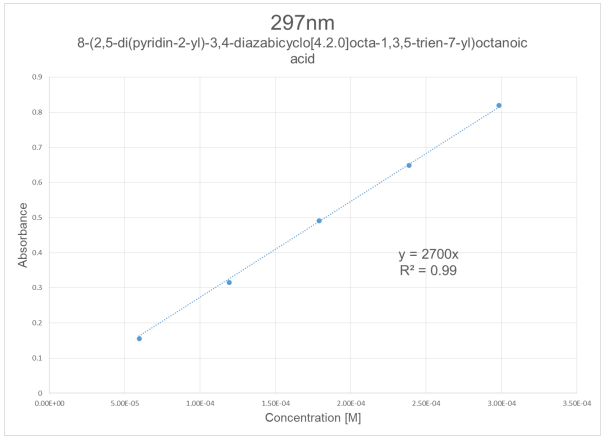
8-(2,5-Di(pyridin-2-yl)-3,4-diazabicyclo[4.2.0]octa-1,4-dien-7-yl)octanoic acid & 8-(2,5-di(pyridin-2-yl)-3,4-diazabicyclo[4.2.0]octa-2,5-dien-7-yl)octanoic acid (**2a**, **2b**). Commercially available 3,6-di(pyridin-2-yl)-1,2,4,5-tetrazine (0.1299 g, 0.55 mmol, 1.1 equiv, Sigma-Aldrich) was dissolved in 5 mL of a 1:1 mixture of THF and CH_3CN (both distilled) in a round bottom flask. A magnetic stir bar was added and the flask was capped with a septum. Dissolution of the tetrazine resulted in a clear and bright pink solution. Cyclobutene **1** (0.5mmol, 0.0981g 1.0 eq.)[1] was added and the reaction was allowed to stir overnight, gradually darkening to a red-orange solution containing a white precipitate. The solid (0.0460 g, 23%) was collected by vacuum filtration and determined by NMR to be an inseparable mixture of dihydrodiazine tautomers (**2a**, **2b**): ^{1}H (MeOD, 400MHz) δ 1.23 (s, 8H), 1.35 (t, 3H, 7.4), 1.47 (t, 2H, 6.4), 1.57 (q, 1H, 7.52), 1.76 (m, 2H), 2.09 (t, 2H 7.4), 2.17 (t, 2H, 7.4), 2.57 (t, 1H, 10.8), 3.13 (m, 2H), 3.60 (m, 1H), 4.26 (d, 1H, 9.4), 6.96 (m, 1H), 7.01 (d, 1H, 8.04), 7.08 (m, 1H), 7.33 (m 2H), 7.56 (m, 1H), 7.72 (m, 2H), 7.82 (dt, 1H, 7.7, 1.5), 7.94 (d, 1H, 8), 8.07 (d, 1H, 7.9), 8.10 (d, 1H, 4.1), 8.31 dd, 2H, 19.1, 4.3), 8.43 (d, 1H, 4.4), 8.64 (s, 1H) δ 12.01 (s, 2H).

8-(2,5-Di(pyridin-2-yl)-3,4-diazabicyclo[4.2.0]octa-1,3,5-trien-7-yl)octanoic acid (**3**). 3,6-Di(pyridin-2-yl)-1,2,4,5-tetrazine (0.0520 g, 0.22 mmol, 1.1 eq.) was placed in a flame-dried, N_2-backfilled 10 ml RBF equipped with a magnetic stir bar and a capped with a rubber septum. A solution of 8-(cyclobut-2-en-1-yl)octanoic acid (**1**, 0.0393g, 0.2 mmol, 1.0 eq.) in 3 mL of distilled tetrahydrofuran was added and the stirred mixture was heated to 65 °C (oil bath) for 2.5 h. Dissolution of the tetrazine resulted in a bright pink solution which gradually took on an orange-red color as reaction proceeded. 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone (DDQ, 0.0499 g, 0.22 mmol, 1.1eq) was added, and the reaction solution immediately became a deep red color. TLC (10% CH_3OH/ CH_2Cl_2) revealed a UV active spot with an R_f just lower than the tetrazine starting material. The reaction was concentrated in vacuo and the residue submitted directly to flash column chromatography (~30 cm Si in a column of 2.5 cm diameter) using a gradient of 0 % (200mL) to 2%

(400 mL) to 6% (400 mL) CH_3OH/CH_2Cl_2 which incorporated 2 drops of glacial acetic acid per 100 mL volume. The residue obtained after concentration was redissolved in 0.5mL of methanol and placed in a -20 °C freezer. The dipyridyl diazine **3** crystallized out of this solution after 18 h and was collected by filtration (0.0471g 58%): R_f = 0.15 (10% CH_3OH/CH_2Cl_2); ^1H ($CDCl_3$ 400MHz) δ 1.32 (m 8H), δ 1.62 (m 2H), δ 2.33 (t, 3H, 6.8), δ 3.42 (dd, 1H, 16.0, 2.2), 3.85 (dd 15.8, 4.84), 4.08 (m 1H) δ 7.35 (m, 2H), 7.87 (dq, 2H, 11.74, 1.68), 8.64 (t, 2H, 9.12), 8.73 (t, 2H, 4.9); ^{13}C ($CDCl_3$, 100MHz, 10 sec delay, 0 Hz line broadening) δ 24.7, 26.9, 28.7, 28.8, 29.1, 31.6, 34.1, 38.8, 47.1, 122.0, 122.1, 124.3, 124.4, 136.9, 137.1, 146.9, 149.6, 149.7, 151.7, 153.1, 153.3, 154.2, 154.4, 178.7.; UV-Vis (distilled CH_3CN; λ_{max} = 200 nm with significant absorption at 297 nm); IR (thin film) 2927 (s, alkane), 2854, 2397, 1708 (COOH); 1379. HRMS (ESI) calc. for $C_{24}H_{26}N_4O_2Na$ $(M+Na)^+$: 425.1593; found: 425.1960 (1.6 ppm).





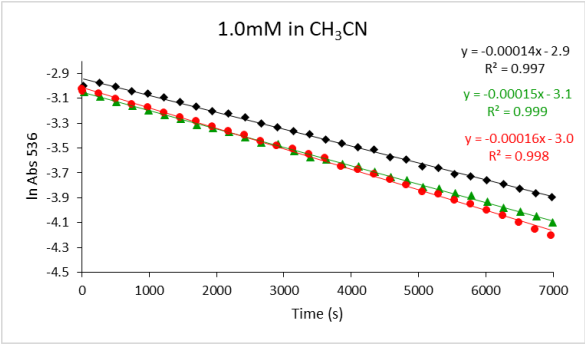


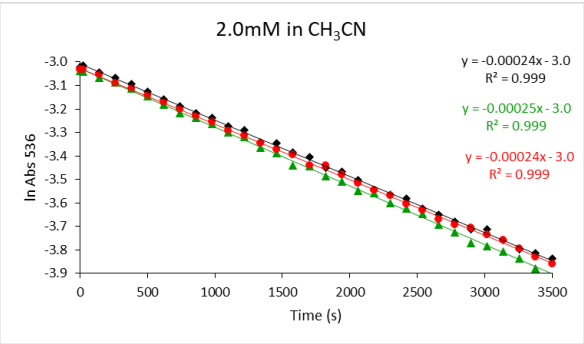
SI-1, Figure 1. UV-Vis spectra of cyclobutene/tetrazine adduct, including determination of extinction coefficients for absorptions at 200 and 297 nm.

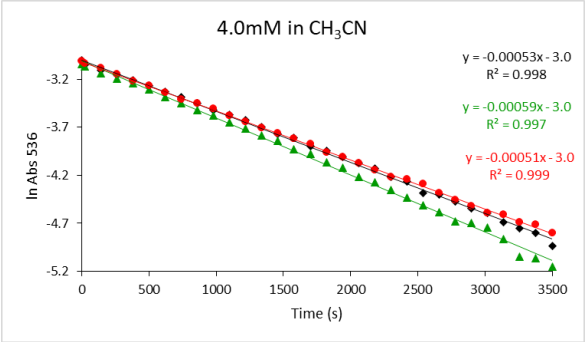
IV. Kinetic analysis of the inverse electron demand Diels-Alder (IEDDA):[2]

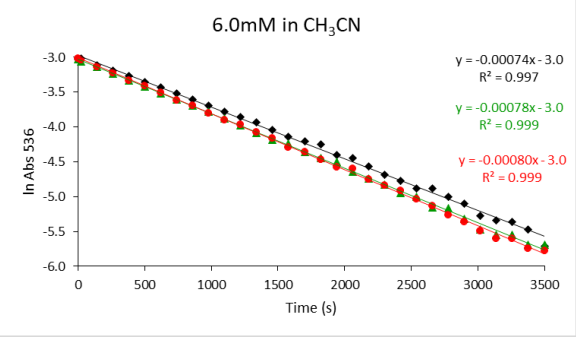
Kinetic analyses of the cycloaddition were conducted in both CH_3CN and in 1:1 PBS/CH_3CN by holding the concentration of 3,6-di(2-pyridyl)tetrazine at 0.1 m*M* and varying the concentration of

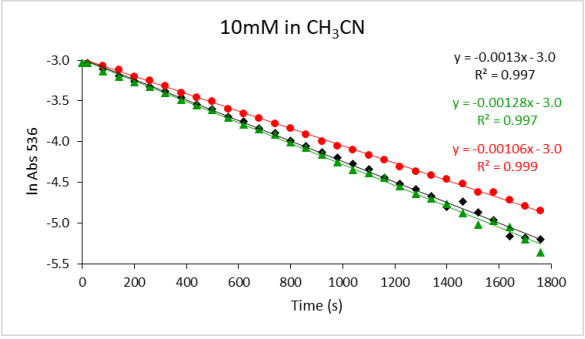
cyclobutene **1** (1, 2, 4, 6, 10 m*M* in CH_3CN; 1, 2, 3, 4, 5, 6 m*M* in 1:1 PBS:CH_3CN). The analyses, conducted in triplicate, monitored the disappearance of the tetrazine absorbance at 536 nm using the following intervals/observation periods: CH_3CN, 10 m*M*: each 1 min over 30 min; CH_3CN, 2/4/6 mM: each 2 min over 60 min; CH_3CN, 1 m*M*: each 4 min over 120 min. 1:1 PBS:CH_3CN, 1/2/3/4/5 m*M*: each 1 min over 15 min; 1:1 PBS:CH_3CN, 6 m*M*: each 15 sec over 5 min. The pseudo-first order rate constants (k_{obs}) for reactions in CH_3CN (Figure 2) and PBS:CH_3CN (Figure 3) were calculated by plotting the natural log of tetrazine absorbance versus time in a series of reactions varying concentrations of cyclobutene **1**.



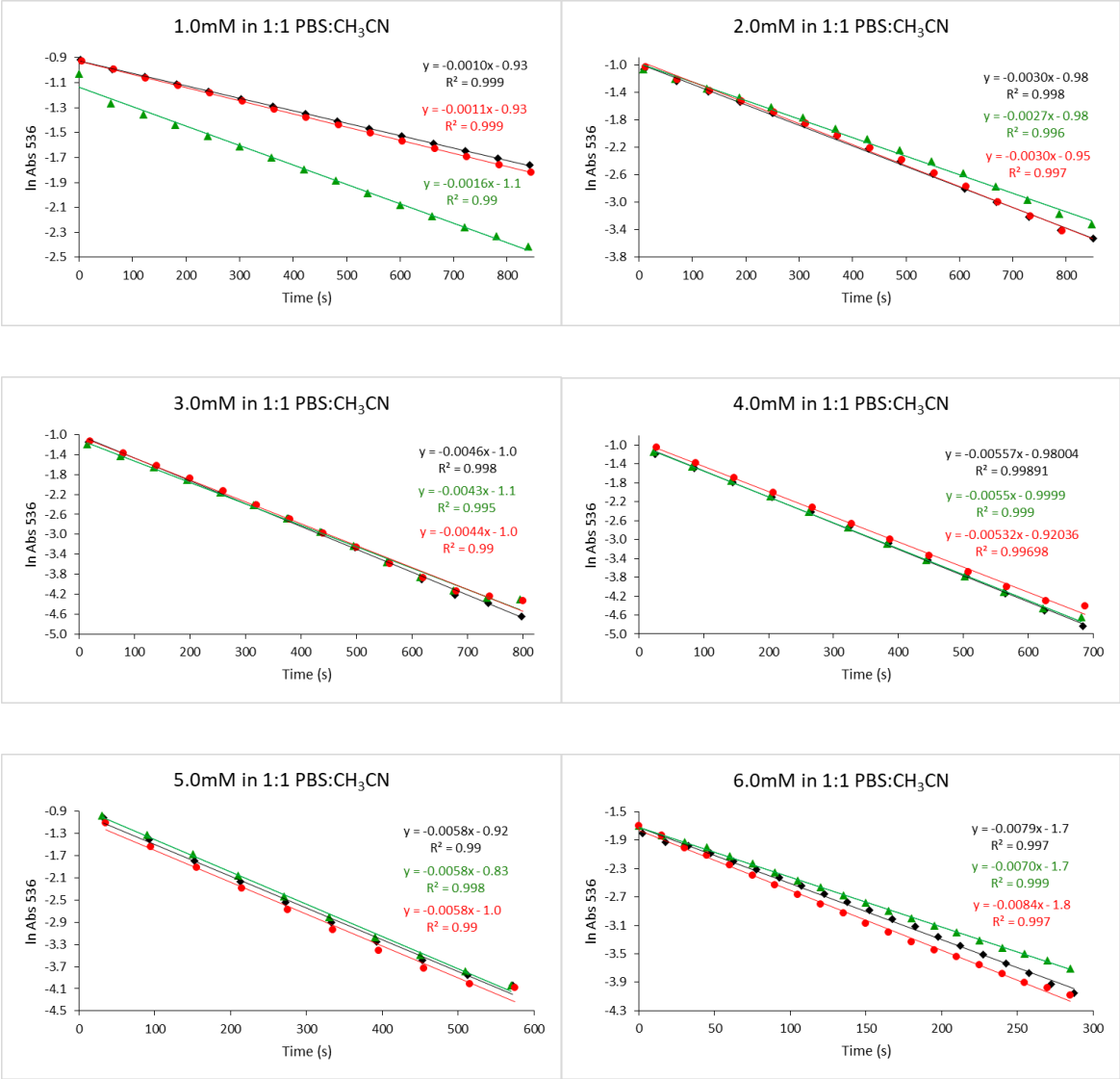




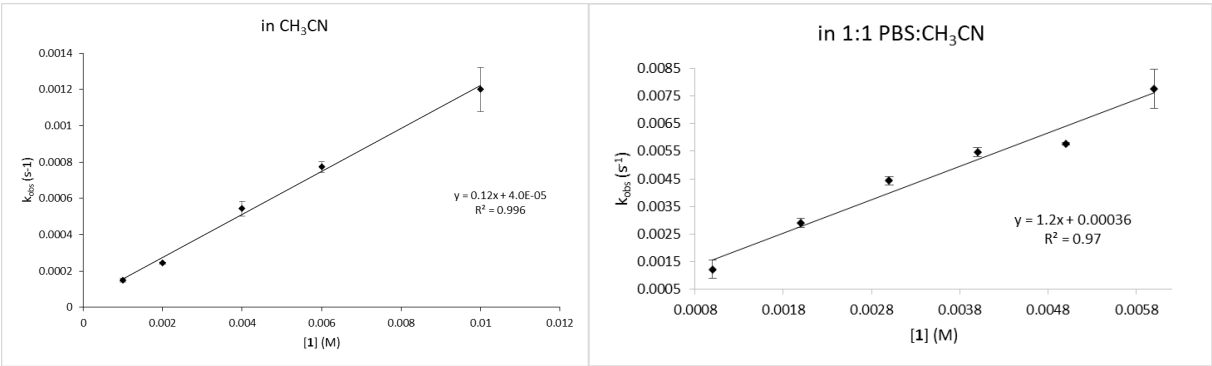




SI-1, Figure 2. Plots of psuedo first-order reactions between dipyridyl tetrazine and cyclobutene **1** in CH_3CN.



SI 1, Figure 3. Calculation of pseudo-first order rate constants of the reaction between dipyridyl tetrazine and cyclobutene **1** in 1:1 PBS:CH_3CN at 1, 2, 3, 4, 5, and 6 mM.



SI-1, Figure 4. Determination of the second order rate constant for the IEDDA between cyclobutene **1** and dipyridyl tetrazine in CH_3CN and 1:1 PBS:CH_3CN.

As illustrated in Figure 4, plotting the slope of the tetrazine absorption vs. concentration of cyclobutene provided second-order rate constants (k) of 0.12 ± 0.01 $M^{-1}s^{-1}$ (CH_3CN) and k = 1.21 ± 0.12 $M^{-1}s^{-1}$ (1:1 PBS:CH_3CN).

V. Plasmid construction

pOmpX. The *OmpX* gene was PCR amplified from *E. coli* genomic DNA using the following primers, The *OmpX* gene was then digested with *Nde*I and *Blp*I, and ligated into a pLei-GFP_{UV} vector,[3] which was treated with the same restriction enzymes, to afford plasmid pLei-OmpX-TAG.

pOmpX-TAG. A DNA fragment that encodes the Ala-Ala-Ala-X-Ala-Ala-Ala sequence (X denotes an amber mutation) was introduced into pOmpX by site-directed mutagenesis using the following primers.

VI. PylRS library construction

A PylRS (from *Methanosarcina barkeri*) library was created, in which residues Leu270, Tyr271, Leu274, and Cys313 were completely randomized. Overlapping polymerase chain reaction (PCR) was performed with synthetic oligonucleotide primers in which the randomized residues were encoded as NNK (N=A, C, T, or G, K=T or G) to generate a library with a theoretical diversity of 1.05 x 10^6. The quality of the library (>99% coverage) was validated by DNA sequencing. The following primers were used in the library construction:

5’-GGAATTCCATATGGATAAAAAACCATTAGATG-3’

5’-AGTCGGGGCAAGCATTGGC-3’

5’-GCCAATGCTTGCCCCGACT**NNKNNK**AACTAT**NNK**CGAAAACTCGATAGGATTTTA-3’

5’- GAAGTTCACCATAGTAAATTC-3’

5’-GAATTTACTATGGTGAACTTC**NNK**CAGATGGGTTCGGGATGTAC-3’

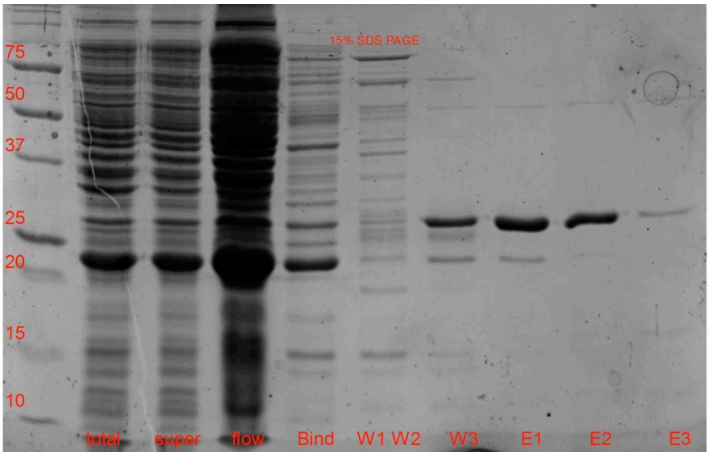
5’-AACTGCAGTTATAGATTGGTTGAAATCCC

VII. Directed evolution of PylRS

The resulting PylRS library was subjected to a negative selection to remove PylRS variants that could charge $tRNA^{Pyl}$ with natural amino acid as previosly described, followed by a positive selection to identify functional PylRS variants. Briefly, the negative selection uses the toxic barnase gene with amber mutations at permissive sites (Gln2TAG and Asp44TAG) and was carried out in the absence of CbK. The positive selection is based on resistance to chloramphenicol (Cm), which is conferred by the suppression of an amber mutation at a permissive site (Asp112) in the chloramphenicol acetyltransferase-encoding gene in the presence of $tRNA^{Pyl}$, CbK, and functional PylRS mutants. The surviving PylRS variants were subsequently screened for chloramphenicol resistance level in the presence and absence of CbK. A few clones that survived on 75 μg/mL chloramphenicol in the presence of CbK and did not grow on 35 μg/mL chloramphenicol in the absence of CbK were identified. Among these clones, CbKRS (Leu274Met and Cys313Ala), displayed the fastest growth rate in the presence of chloramphenicol.

VIII. Protein expression

E. coli GeneHogs cells harboring pBK-CbKRS and pLei-sfGFP-N149TAG[4] were cultured in 50 mL of LB medium containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37 ^{o}C with shaking. When the OD_{600} of the culture reached 0.6, the protein expression was induced by the addition of IPTG (0.5 mM) and CbK (1.0 mM). After 24 h of cultivation at room temperature, cells were collected by centrifugation (5,000g, 10 min), resuspended in lysis buffer (20 mM potassium phosphate, pH 7.4, 150 mM NaCl, and 20 mM imidazole), and disrupted by sonication. Cellular debris was removed by centrifugation (21,000g, 30 min, 4 ^{o}C). The cell-free lysate was applied to Ni Sepharose 6 Fast Flow resin (GE Healthcare). Protein purification followed manufacturer's instructions. Purified protein was desalted using the Amicon Econo-Pac® 10DG column (BioRad Laboratories, Inc) using PBS buffer.

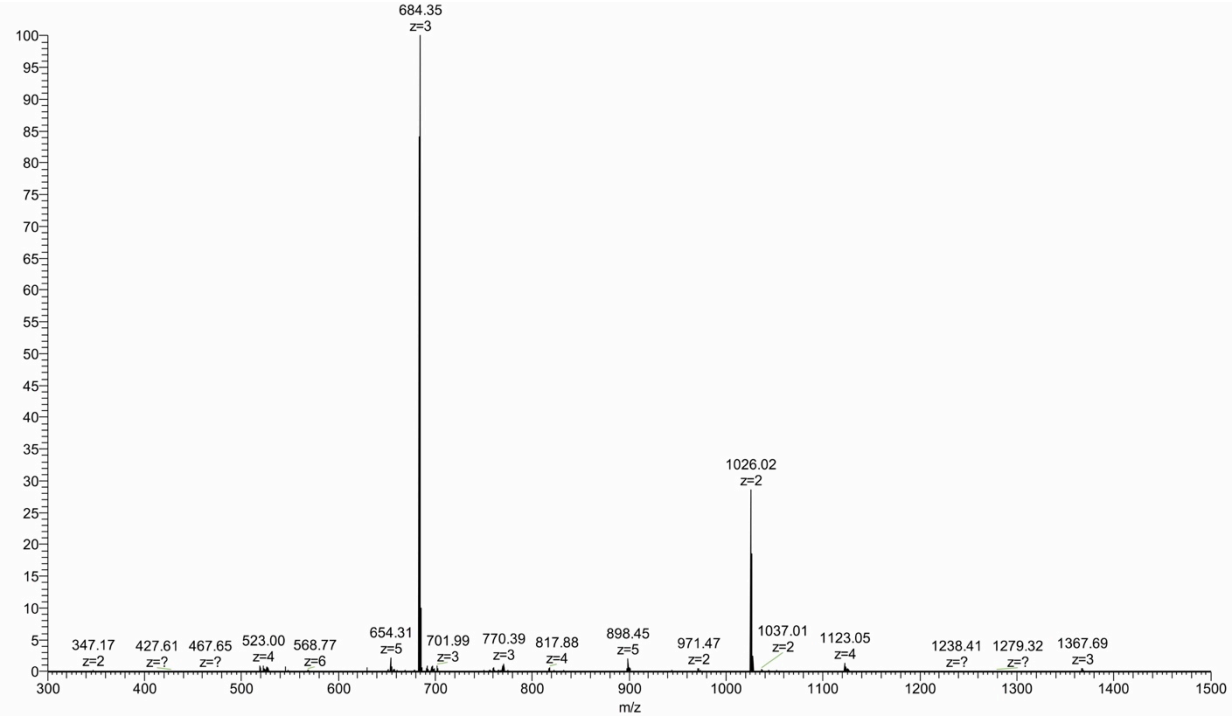


SI-1, Figure 5. SDS-PAGE analysis of protein purification.

IX. Mass spectroscopy analysis

For protein fragment from in-gel digestion, gel band containing sfGFP-N149CbK was cut from SDS-PAGE stained by Coomassie blue. After in-gel digestion with trypsin, the protein sample was dried down and re-dissolved in 120 μL of aqueous solution with 2.5% acetonitrile and 0.1% formic acid. A 5 μL of the digest sample was injected into a LC-MS that was equipped with a 0.075 mm x 250 mm C18 Dionex column and a Q-Exactive HF mass spectrometer. The mass spectrometry data was analyzed using Mascot (Matrix Science, London, UK; version 2.5.1). Mascot was set up to search Cbk-containing peptide, LEYNFNSH-CbK-VYITADK. Deamidation of asparagine and glutamine, oxidation of methionine was specified in Mascot as variable modifications.

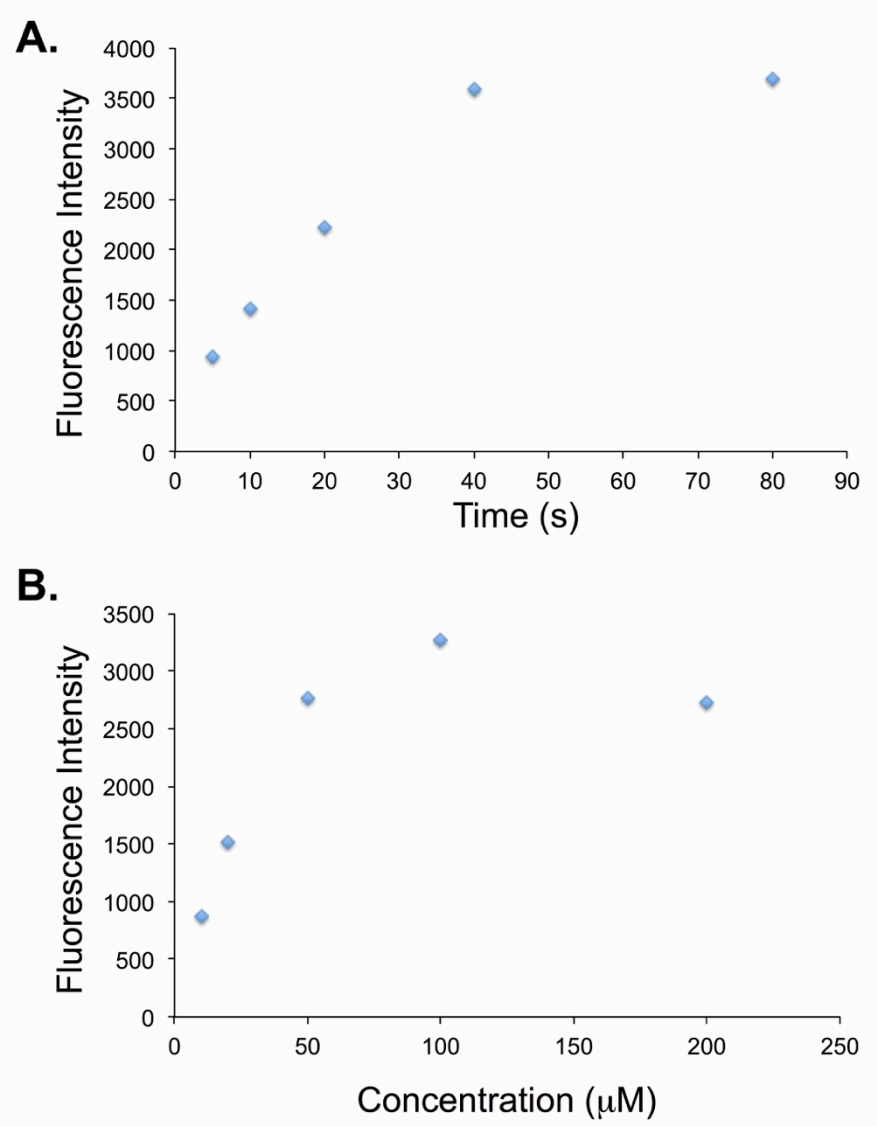
For protein fragment from solution digestion, sample was loaded onto Ni-Sap column after digestion, washed, and eluted with phosphate buffer containing 250 mM imidazole. Sample was buffer exchanged into water using 10K MWCO spin column (Amicon). LC-MS analysis followed the above procedure.



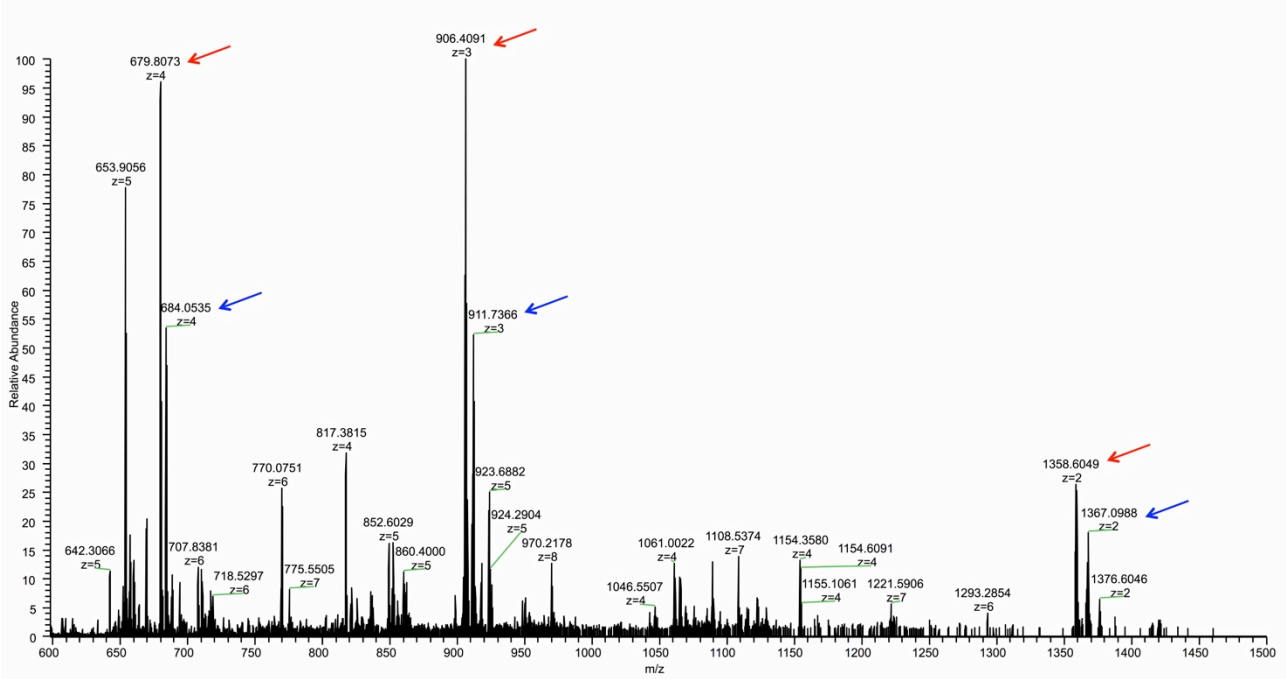
SI-1, Figure 6. LC-MS of in-gel digested sfGFP-N149CbK. Expected, 684.40 Da, z = 3; 1026.10 Da, z = 2. Observed 684.35 Da, z = 3; 1026.02 Da, z = 2.

X. Protein labeling in vitro

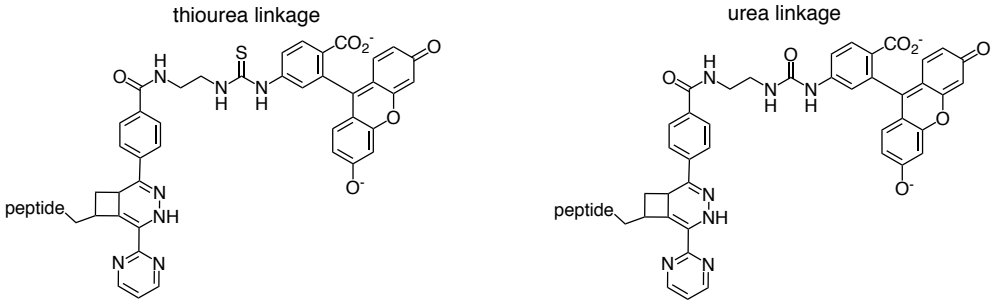
Purified sfGFP-N149CbK (10 μL, 0.24 mg/mL) in PBS buffer was treated with 1 μL of Fl-Tet (in DMSO) at indicated concentrations. The reaction mixture was incubated at room temperature with agitation for indicated period of time. Solution of 5-norbonene-2-methanol (1 μL, 100 mM) in DMSO was added to quench the reaction. The quenched reaction was then mixed with SDS-PAGE sample loading buffer (2x), heated at 95 ^{o}C for 15 min, and loaded onto wells of SDS-PAGE. Fluorescence detection was performed before staining by Coomassie blue. Protein gels were imaged using Bio-Rad Molecular Imager ChemDoc XRS+ or GE Typhoon FLA9500. Coomassie blue-stained gels were imaged using Bio-Rad Molecular Imager. As a control, wild-type sfGFP (10 μL, 0.24 mg/mL) was subjected to labeling and analysis under the same conditions.



SI-1, Figure 7. In vitro protein labeling of sfGFP variants with cyclobutene-tetrazine reaction. Fluorescence intensities of protein bands (Figure 3 of the maintext) were estimated using ImageJ software. (A) Reaction progress of sfGFP-N149CbK protein labeling with 50 μM of Fl-Tet; (B) Labeling of sfGFP-N149CbK protein with varied concentrations of Fl-Tet for 80 minutes.



SI-1, Figure 8. LC-MS of solution digested reaction product of sfGFP-N149CbK and fluorecein-labeled tetrazine. Arrows indicate masses corrsponding to the peptide fragment (LEYNFNSH**X**VYITADK) containing the modification (**X**). We observed hydrolysis of thiourea linkage (blue arrows) to urea linkage (red arrows). Other peaks are from other peptide fragments of the protein. Calculated and observed masses are listed in the following table.



	calculated	observed
thiourea linkage	684. 28 Da, z = 4 912.04 Da, z = 3 1367.56 Da, z = 2	684.05 Da, z = 4 911.74 Da, z = 3 1367.10 Da, z = 2
urea linkage	680.28 Da, z = 4 906.71 Da, z = 3 1359.57 Da, z = 2	679.81 Da, z = 4 906.41 Da, z = 3 1358.60 Da, z = 2

XI. Live cell labeling

E. coli cells expressing either wild-type OmpX or OmpX-CbK mutant proteins were harvested by centrifugation (21,000 g, 5 min, 4 °C). Collected cells were washed three times with PBS buffer. Washed cells were resuspended in PBS buffer. Fl-Tet stock solution was added into cell suspension to a final concentration of 100 μM. After incubation at room temperature with agitation for 60 min, cells were collected by centrifugation (21,000 g, 5 min, 4 °C) and resuspended in PBS buffer. Cell suspensions were placed on the surface of a glass slide and covered with a glass cover slip for imaging. Cells were imaged on an Olympus FV500 inverted (Olympus IX-81) confocal microscope. Excitation wavelength (495 nm) and emission filter (510 nm) were used in the imaging experiments.

References:

1 (a) W. Sittiwong, D. K. Zinniel, R. J. Fenton, D. D. Marshall, C. B. Story, B. Kim, J. Lee, R. Powers, R. G. Barletta, and P. H. Dussault, "Development of Cyclobutene- and Cyclobutane-Functionalized Fatty Acids with Inhibitory Activity against Mycobacterium tuberculosis." *ChemMedChem*, 2014, **9**, 1838; (b) W. Sittiwong. Synthesis and Application of Four-Membered Carbocycle Amphiphiles. Ph.D. Thesis, University of Nebraska-Lincoln, Lincoln, NE, 2014.

2 (a) F. Sicilio, and M.D. Peterson, "Ratio errors in pseudo first order reactions." *J. Chem. Ed.* 1961, **38**, 576; (b) J.F. Corbett, "Pseudo first-order kinetics." *J. Chem. Ed.* 1972, **49**, 663.

3 N. Wang, T. Ju, W. Niu and J. Guo, "Fine-tuning Interaction between Aminoacyl-tRNA Synthetase and tRNA for Efficient Synthesis of Proteins Containing Unnatural Amino Acids" *ACS Synth. Biol.*, 2014, **4**, 207-212.

4 X. Shang, X. Song, C. Faller, R. Lai, H. Li, R. Cerny, W. Niu and J. Guo, "Fluorogenic protein labeling using a genetically encoded unstrained alkene" *Chem. Sci.*, 2017, **8**, 1141-1145.