Supplemental Information

Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific *In Vivo* Bioorthogonal Ligation with *trans*-Cyclooctenes

Jason L. Seitchik[†], Jennifer C. Peeler[†], Michael T. Taylor[‡], Melissa L. Blackman[‡], Timothy W. Rhoads[§], Richard B. Cooley[§], Christian Refakis[†], Joseph M. Fox[‡], Ryan A. Mehl[§]*

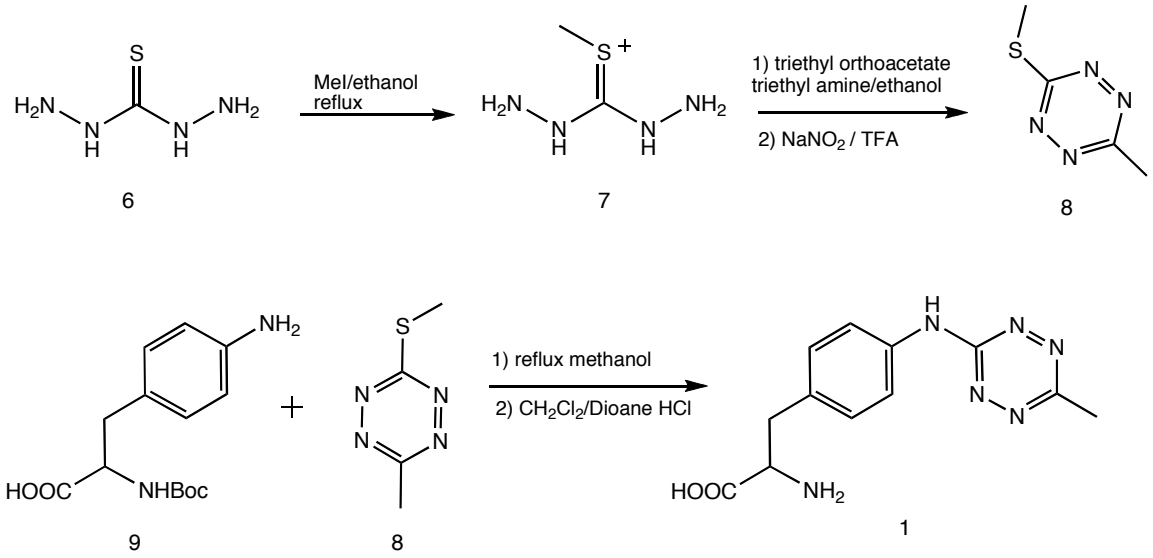
[†] Franklin & Marshall College, Department of Chemistry, Lancaster, PA 17604-3003, USA

[‡] Brown Laboratory, Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716, USA

[§] Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97331, USA.

Email corresponding author: ryan.mehl@oregonstate.edu

Synthesis of UAA tetrazine 4-(6-methyl-*s*-tetrazin-3-yl)aminophenylalanine, 1.



Scheme 1. Synthesis of 4-(6-methyl-*s*-tetrazin-3-yl)aminophenylalanine.

Methyl-thiocarbohydrazide (7)[1].

Commercially obtained thiocarbohydrazide **6** (24.09 g, 0.2273 mol) was dissolved in 700 mL of absolute ethanol and brought to reflux. MeI (16 mL, 1.1 eq.) in 100 mL of absolute ethanol was added in dropwise over 30 min with vigorous stirring. The solution was allowed to reflux for 1 hr and then filtered hot using a C type filter crucible. The filtrate was allowed to cool to room temperature and the product was allowed to

precipitate out of solution over 16 hr. The solution was decanted away and the product was dried *in vacuo* to obtain 27.02 g of an off-white solid (47.9% yield). ^{1}H-NMR (DMSO, 90 MHz): δ 2.40 (s, methyl, 3 H).

Methyl-thiomethyl-tetrazine (8)[2]

Methyl-thiocarbohydrazide **7** (5.62 g, 0.0227 mol) was dissolved in 150 mL of absolute ethanol. Triethyl orthoacetate (4.65 mL, 1.1 eq.) was added, forming a yellow solution. After five minutes, triethyl amine (3.12 mL, 1.0 eq.) was added, turning the solution a light pink. The solution was refluxed for 30 min, until an orange color had formed. $NaNO_2$ (3.20 g, 1.0 eq.) and TFA (1.74 mL, 1.0 eq.) were added and heating was continued for an additional 30 min, turning the solution a deep red. Hexane (150 mL) was added and a stream of air was used to purge gases. The solution was cooled to room temperature. The solution was diluted with water (300 mL) and extracted with ether. The product was evaporated under reduced pressure to 5-10 mL and then separated in a 5% diethyl ether: hexane silica gel column. Fractions containing the product were evaporated under reduce pressure to obtain 0.72 g of a red oil (22.4% yield). ^{1}H-NMR ($CDCl_3$, 90 MHz): δ 2.98 (s, methyl, 3 H), δ 2.73 (s, methyl, 3 H). ^{13}C-NMR ($CDCl_3$, 500 MHz): δ 175.26 (tetrazine C), δ 165.02 (tetrazine C), δ 20.62 (S-CH_3), δ 13.29 (methyl).

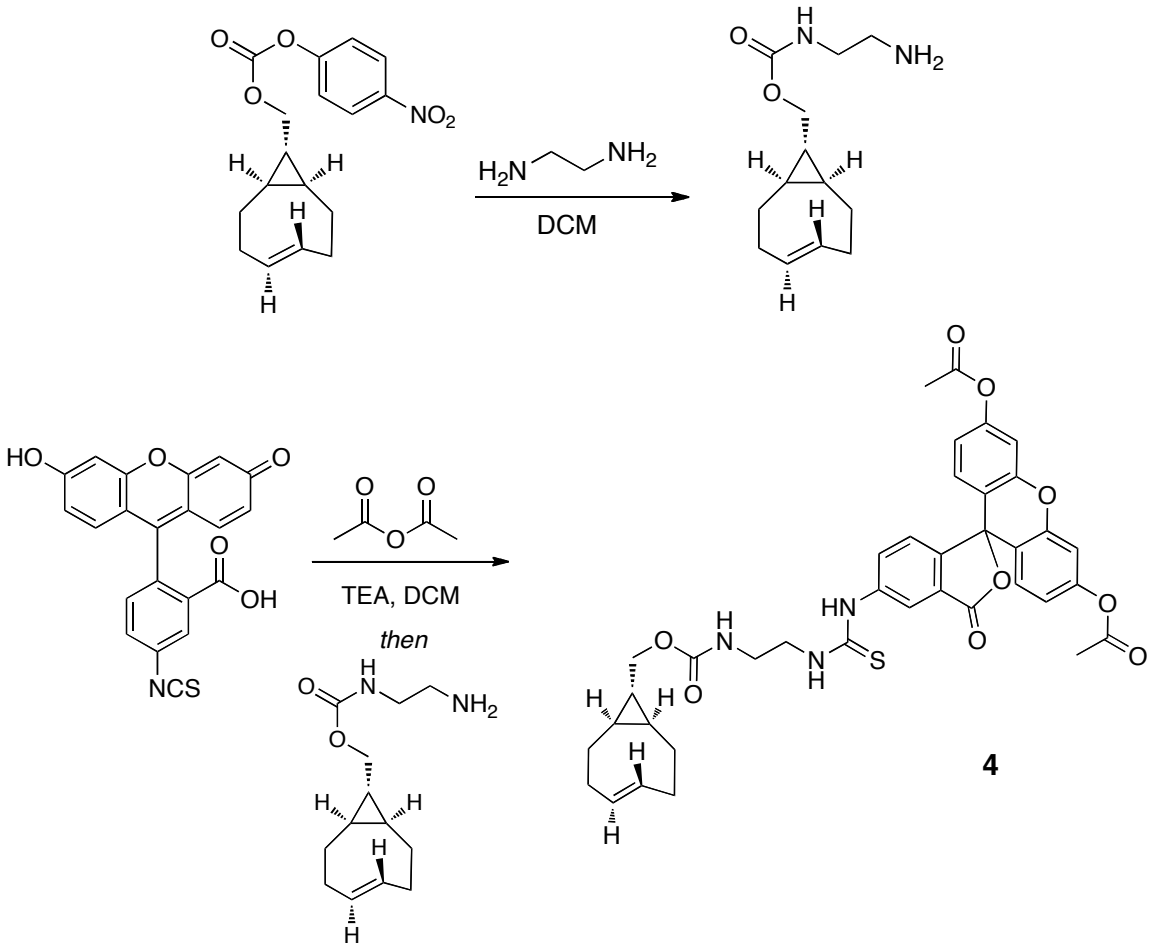
N-Boc-4-(6-methyl-*s*-tetrazin-3-yl)-aminophenylalanine.

Commercially obtained N-Boc 4-aminophenylalanine **9** (1.0211 g, 0.00408 mol) and methyl-thiomethyl-tetrazine **8** (810 mg, 0.00570 mol) were dissolved in 10 mL of dry methanol. The reaction was refluxed at 110 °C and complete after 72 h (monitored by TLC). The product was separated by a silica gel column, starting with 100% $CHCl_3$ and then increasing the presence of methanol from 2.5% to 5% with 0.25% acetic acid to move the starting material and product, respectively. The fraction containing the product was evaporated under reduced pressure and dried *in vacuo* to obtain 0.45 g of the orange solid (33.0% yield). ^{1}H-NMR ($CDCl_3$, 500 MHz): δ 7.57 (d, aromatic, 2 H), δ 7.17 (d, aromatic, 2 H), δ 4.50 (t, CH, 1 H), δ 3.1 (d of q, CH_2, 2 H), δ 2.58 (s, CH_3, 3 H), δ 1.39 (s, boc, 9 H). ^{13}C-NMR ($CDCl_3$, 500 MHz): δ 173.72 (tetrazine C), δ 162.22 (tetrazine C), δ 162.15 (COOH), δ 160.61 (ester), δ 131.83 (phenyl), δ 131.47 (phenyl), δ 130.06 (phenyl), δ 129.91 (phenyl), δ 119.71 (C-N), δ 54.20 (C-O), δ 37.24 (CH_2), δ 28.15 (boc methyl), δ 20.01 (tetrazine methyl).

4-(6-methyl-*s*-tetrazin-3-yl)-aminophenylalanine (1).

N-Boc-4-(6-methyl-*s*-tetrazin-3-yl)-aminophenylalanine (0.45 g, 0.0013 mol) was dissolved in 30 mL 1:1 CH_2Cl_2 and dry 4M HCl in dioxane. The reaction mixture was stirred at room temperature for ~4 h and then was evaporated under reduced pressure to a final volume of 5-10 mL. Diethyl ether was then added to the solution, and the precipitate was filtered using a C type filter crucible and dried *in vacuo* as the HCl salt (0.3087 g, 93.6% yield). ^{1}H-NMR (DMSO, 500 MHz): δ 7.67 (d, aromatic, 2 H), δ 7.27 (d, aromatic, 2 H), δ 4.14 (t, CH, 1 H), δ 3.11 (d, CH_2, 2 H), δ 2.78 (s, methyl, 3 H). ^{13}C-NMR (DMSO, 500 MHz): δ 170.37 (tetrazine C), δ 161.60 (tetrazine C), δ 160.75 (COOH), δ 137.60 (phenyl), δ 129.96 (phenyl), δ 129.30 (phenyl), δ 119.68 (phenyl), δ 53.19 (C-N), δ 35.17 (CH_2), δ 19.80 (methyl).

Scheme 2. Synthesis of sTCO-diacetyl-fluorescein 4.



Synthesis of sTCO-diacetyl-Fluorescein 4.

(1*R*,8*S*,9*R*,4*E*)-Bicyclo[6.1.0]non-4-ene-9-ylmethyl (*N*-aminoethyl) carbamate

To a stirring solution of ethylene diamine (0.150 mL, 2.21 mmol), in methylene chloride (3 mL) was added (1*R*,8*S*,9*R*,4*E*)-Bicyclo[6.1.0]non-4-ene-9-ylmethyl (4-nitrophenyl) carbonate[3] (69.4 mg, 0.221 mmol) in methylene chloride (1.5 mL) over 1 hour via syringe pump. The reaction solution was diluted with methylene chloride and extracted 3 times with water. The organic layer was dried with K_2CO_3, filtered and the solvent was removed using a rotary evaporator to yield 45 mg (87%) of the title compound as a faint yellow oil.

^{1}H-NMR (400 MHz, CD_3OD, δ): 5.82-5.74 (m, 1H), 5.09-5.00 (m, 1H), 3.91-3.80 (m, 2H), 3.07 (t, *J*= 6.2 Hz, 2H), 2.61 (t, , *J*= 6.2 Hz, 2H), 2.27 (app d, *J*= 14.4 Hz, 1H), 2.20-2.09 (m, 3H), 1.89-1.77 (m, 2H), 0.86-0.76 (m, 1H), 0.59-0.49 (m, 2H), 0.41-0.32 (m, 2H).

^{13}C-NMR (100 MHz, CD_3OD, δ): 156.6 (C), 136.3 (CH), 129.3 (CH), 67.5 (CH_2), 41.4 (CH_2), 39.5 (CH_2), 36.8 (CH_2), 31.8 (CH_2), 30.7 (CH_2), 25.7 (CH_2), 23.1 (CH), 20.3 (CH), 29.3 (CH)

IR ($CHCl_3$, cm^{-1}): 3452, 3018, 2933, 2850, 1707, 1515, 1447, 1219, 1144, 1011, 930.

HRMS (ESI) [M+H] calcd. for $C_{13}H_{22}N_2O_2$, [M+H]: 239.1759 ; found: 239.1758.

(1*R*,8*S*,9*R*,4*E*)-Bicyclo[6.1.0]non-4-ene-9-ylmethyl (*N*-ethyl thioureidofluorescein diacetate) carbamate (4)

A round bottomed flask was charged fluorescein isothiocyanate (26.6 mg, 0.0683 mmol). The flask was evacuated and refilled with N_2. Anhydrous methylene chloride (0.5 mL) was added, followed by triethylamine (40 µL, 0.285 mmol). Acetic anhydride (13 µL, 0.133 mmol) was added and the reaction allowed to stir for 1.5 hours. (1*R*,8*S*,9*R*,4*E*)-Bicyclo[6.1.0]non-4-ene-9-ylmethyl (*N*-aminoethyl) carbamate (21.7 mg, 0.0911 mmol) in methylene chloride (0.5 mL) and the reaction allowed to stir for an additional 1.5 hrs. The reaction solution was transferred directly to a TLC plate and purified by preparatory TLC (30% acetone/toluene) to yield 14.2 mg (29%) of the title compound as a solid.

HRMS (ESI) [M+H] calcd. for $C_{38}H_{37}N_3O_9S$, [M+H]: 712.2323 ; found: 712.2328.

Selection of an aminoacyl-tRNA synthetase specific for 4-(2'-bromoisobutyramido)phenylalanine 1.

The library of aminoacyl-tRNA synthetases was encoded on a kanamycin (Kn) resistant plasmid (pBK, 3000 bp) under control of the constitutive *Escherichia coli* GlnRS promoter and terminator. The aminoacyl synthetase library (3D-Lib) was randomized as follows: Leu65, His70, Gln155, and Ile159 were randomized to all 20 natural amino acids; Tyr32 was randomized to 15 natural amino acids (less Trp, Phe, Tyr, Cys, and Ile); Asp158 was restricted to Gly, Ser, or Val; Leu162 was restricted to Lys, Ser, Leu, His, and Glu; and Phe108 and Gln109 were restricted to the pairs Trp-Met, Ala-Asp, Ser-Lys, Arg-Glu, Arg-Pro, Ser-His, or Phe-Gln. The library plasmid, pBK-3D-Lib, was moved between cells containing a positive selection plasmid (pCG) and cells containing a negative selection plasmid (pNEG).

The positive selection plasmid, pCG (10000 bp), encodes a mutant *Methanococcus jannaschii* (Mj) tyrosyl-$tRNA_{CUA}$, an amber codon-disrupted chloramphenicol acetyltransferase, an amber codon-disrupted T7 RNA polymerase that drives the production of green fluorescent protein, and the tetracycline (Tet) resistance marker. The negative selection plasmid, pNEG (7000 bp), encodes the mutant tyrosyl-$tRNA_{CUA}$, an amber codon-disrupted barnase gene under control of an arabinose promoter and rrnC terminator, and the ampicillin (Amp) resistance marker. pCG electrocompetent cells and pNEG electrocompetent cells were made from DH10B cells carrying the respective plasmids and stored in 100 μL aliquots at −80 °C for future rounds of selection.

The synthetase library in pBK-3D-Lib was transformed by electroporation into DH10B cells containing the positive selection plasmid, pCG. The resulting pCG/pBK-3D-Lib-containing cells were amplified in 1 L of 2×YT with 50 μg/mL Kn and 25 μg/mL Tet with shaking at 37 °C. The cells were grown to saturation, then pelleted at 5525 rcf, resuspended in 30 mL of 2×YT and 7.5 mL of 80% glycerol, and stored at −80 °C in 1 mL aliquots for use in the first round of selections.

For the first positive selection, 2 mL of pCG/pBK-3D-Lib cells were thawed on ice before addition to 1.2 L of room temperature 2×YT media containing 50 μg/mL Kn and 25 μg/mL Tet. After incubation (11 h, 250 rpm, 37 °C), a 200 μL aliquot of these cells was plated on eleven 15 cm GMML-agar plates containing 50 μg/mL Kn, 25 μg/mL Tet, and 60 μg/mL chloramphenicol (Cm). The positive selection agar medium also contained 1 mM **1**. After spreading, the surface of the plates was allowed to dry completely before incubation (37 °C, 15 h). To harvest the surviving library members from the plates, 10 mL of 2×YT (50 μg/mL Kn, 25 μg/mL Tet) was added to each plate. Colonies were scraped from the plate using a glass spreader. The resulting solution was incubated with shaking (60 min, 37 °C) to wash cells free of agar. The cells were then pelleted, and plasmid DNA was extracted. For the first positive selection a Qiagen midiprep kit was used to purify the plasmid DNA. For all other plasmid purification steps a Qiagen miniprep kit was used to purify the plasmid DNA. The smaller pBK-3D-Lib plasmid was separated from the larger pCG plasmid by agarose gel electrophoresis and extracted from the gel using the Qiagen gel extraction kit.

The purified pBK-3D-Lib was then transformed into pNEG-containing DH10B cells. A 100 μL sample of pNEG electrocompetent cells was transformed with 50 ng of

purified pBK-3D-Lib DNA. Cells were rescued in 1 mL of SOC for 1 h (37 °C, 250 rpm) and the entire 1 mL of rescue solution was plated on three 15 cm LB plates containing 100 μg/mL Amp, 50 μg/mL Kn, and 0.2% L-arabinose. Cells were collected from plates and pBK-3D-Lib plasmid DNA was isolated in the same manner as described above for positive selections.

For the second round of positive selection, 50 ng of purified library DNA was transformed into 100 μL of pCG competent cells. The transformants were rescued for 1.5 h in 1 mL of SOC (37 °C, 250 rpm). A 50 μL sample of these cells was plated on three plates prepared as described in the first positive selection on LB agar plates.

For the second negative selection, one plate was spread with 250 μL of rescued cells, and two plates were spread with 50 μL of rescued cells and then incubated (12−16 h, 37 °C). For this round, the cells were plated on LB agar containing 100 μg/mL Amp, 50 μg/mL Kn, and 0.04% L-arabinose.

In order to evaluate the success of the selections based on variation in synthetase efficacy (as opposed to traditional survival/death results), the synthetases resulting from the selection rounds were tested with the pALS plasmid. This plasmid contains the sfGFP reporter with a TAG codon at residue 150 as well as tyrosyl-$tRNA_{CUA}$. When a pBK plasmid with a functional synthetase is transformed with the pALS plasmid and the cells are grown in the presence of the appropriate amino acid on autoinduction agar, sfGFP is expressed and the colonies are visibly green.

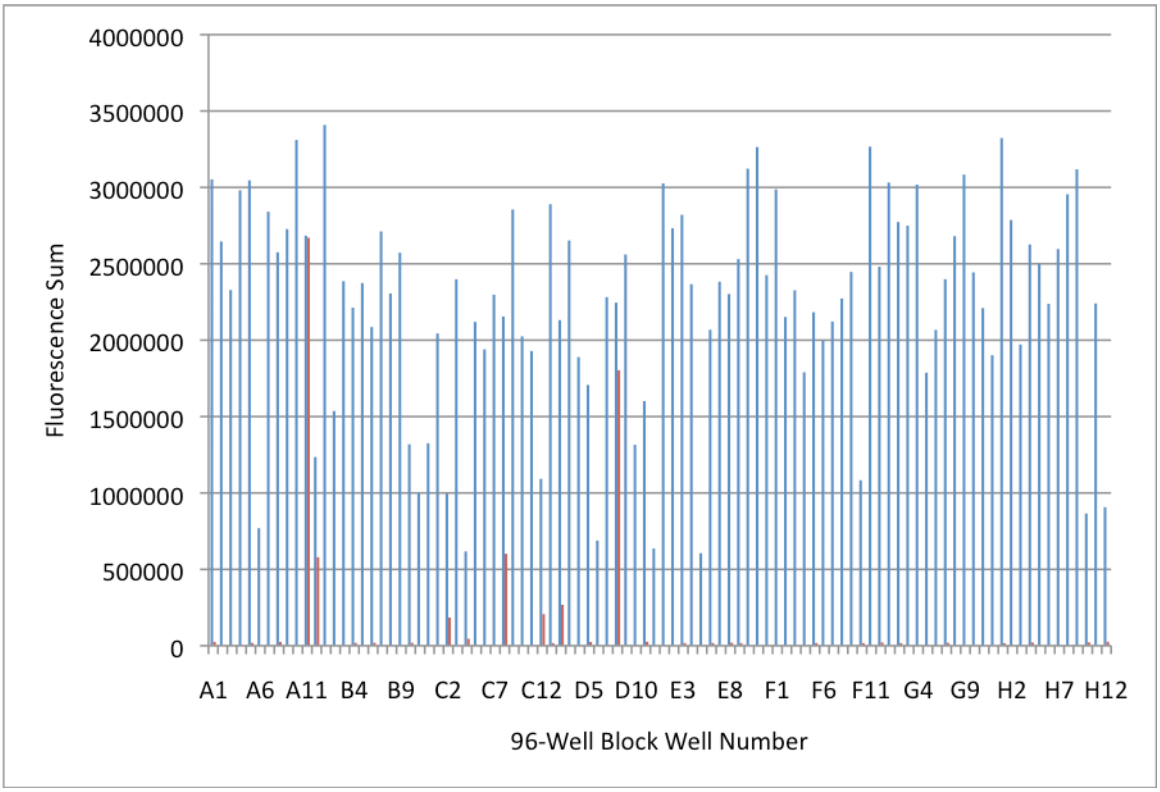
One microliter of each library resulting from the second positive and the second negative rounds of selection was transformed with 60 μL of pALS-containing DH10B cells. The cells were rescued for 1 hr in 1 mL of SOC (37 °C, 250 rpm). A 250 μL and 50 μL of cells from each library were plated on autoinducing minimal media with 25 μg/mL Kn, 25 μg/mL Tet, and 1 mM **1**. Plates were grown at 37 °C for 24 hours and then grown on the bench top, at room temperature, for an additional 24 hours.

Autoinducing agar plates were prepared by combining the reagents in Sup. Table 1A with an autoclaved solution of 40 g of agarose in 400 mL water. Sterile water was added to a final volume of 500 mL. Antibiotics were added to a final concentration of 25 μg/mL Tet and 25 μg/mL Kan. Nine plates were poured with 1 mM **1**, and nine plates were maintained as controls without UAA.

A total of 96 visually green colonies were selected from the two 1 mM **1** plates and used to inoculate a 96-well plate containing 0.5 mL per well non-inducing minimal media (Sup. Table 1B, with sterile water added to a final volume of 500 mL) with 25 μg/mL Kn, 25 μg/mL Tet. After 24 hours of growth (37 °C, 250 rpm), 5 μL of these non-inducing samples were used to inoculate 96-well plates with 0.5 mL autoinduction media (Sup. Table 1C, with sterile water added to a final volume of 500 mL) containing 25 μg/mL Kn, 25 μg/mL Tet with and without 1 mM **1**. Fluorescence measurements of the cultures were collected 40 hours after inoculation using a HORIBA Jobin Yvon FluoroMax®-4. The emission from 500 to 520 nm (1 nm bandwidth) was summed with excitation at 488 nm (1 nm bandwidth). Samples were prepared by diluting suspended cells directly from culture 100-fold with phosphate buffer saline (PBS). Results of fluorescence measurements are shown in Supplemental Figure 1.

Sup. Table 1. Components for autoinducing and non-inducing mediums, for final volume of 500 mL.

	A) Autoinduction medium	**B)** Non-inducing medium	**C)** Autoinducing plates
5% aspartate, pH 7.5	25 mL	25 mL	25 mL
10% glycerol	25 mL	-	25 mL
25× 18 amino acid mix	20 mL	20 mL	20 mL
50× M	10 mL	10 mL	10 mL
leucine (4 mg/mL), pH 7.5	5 mL	5 mL	5 mL
20% arabinose	1.25 mL	-	1.25 mL
1 M $MgSO_4$	1 mL	1 mL	1 mL
40% glucose	625 µL	6.25 mL	125 µL
Trace metals	100 µL	100 µL	100 µL



Sup. Figure 1. Fluorescence measurements of 96 synthetases with GFP clones. Blue represents colonies induced in media containing 1 mM **1** while red represents colonies induced in the absence of UAA. Expressions of 500 µL were grown for 40 hours before dilution of suspended cells directly from culture 100-fold with phosphate buffer saline (PBS). Fluorescence measurements were collected using a HORIBA Jobin Yvon FluoroMax®-4. The emission from 500 to 520 nm (1 nm bandwidth) was summed with excitation at 488 nm (1 nm bandwidth).

Fluorescence analysis of highest-fluorescing clones

Non-inducing cultures (3 mL) with 25 μg/mL Kn and 25 μg/mL Tet were grown to saturation (37 °C with shaking at 250 rpm) from the 20 highest-fluorescing colonies. Autoinduction cultures (3 mL) with 25 μg/mL Kn and 25 μg/mL Tet were inoculated with 30 μL of non-inducing cultures and grown with and without 1 mM **1** at 37 °C with shaking at 250 rpm. After approximately 40 hours, fluorescence was assessed as described above. Theses top 20 performing clones were sequence revealing one unique mtaF gene (R9I, Y32E, L65A, A107E, F108P, Q109S, D158G, L162G).

The top performing clone was moved from the *pBK-G2* plasmid to the pDule plasmid (*pDule-BIBAF*). *pDule* plasmid was generated by amplifying the *MjYRS* gene from the *pBK* plasmid isolated from the library using primers RSmovef (5'-CGCGCGCCATGGACGAATTTGAAATG-3') and RSmover (5'-GACTCAGTCTAGGTACCCGTTTGAAACTGCAGTTATA-3'). The amplified DNA fragments were cloned in to the respective sites on the *pDule* plasmids using the incorporated NcoI and KpnI sites.

Expression and purification of GFP-1.

DH10B *E. coli* cells co-transformed with the *pBad-sfGFP-150TAG* vector and the machinery plasmid *pDule-mtaF* were used to inoculate 5 mL of non-inducing medium containing 100 μg/mL Amp and 25 μg/mL Tet. The non-inducing medium culture was grown to saturation with shaking at 37 °C, and 5.0 mL was used to inoculate 0.5 L autoinduction medium with 100 μg/mL Amp, 25 μg/mL Tet, and 1 mM **1** (0.5 L of media grown in 2 L plastic baffled flasks). After 40 hours of shaking at 37 °C, cells were collected by centrifugation.

The protein was purified using BD-TALON cobalt ion-exchange chromatography. The cell pellet was resuspended in wash buffer (50 mM sodium phosphate, 300 mM sodium chloride, pH 7) containing 1 mg/mL chicken egg white lysozyme, and sonicated 3 × 1 min while cooled on ice. The lysate was clarified by centrifugation, applied to 6-9 mL bed-volume resin, and bound for 30 min. Bound resin was washed with >50 volumes wash buffer.

Protein was eluted from the bound resin with 2.5 mL aliquots of elution buffer (50 mM sodium phosphate, 300 mM sodium chloride, 150 mM imidazole pH 7) until the resin turned pink and the color of the eluent the column was no longer green. The elusions concentrations were check with a Bradford protein assay. The protein was desalted into PBS using PD10 columns.

Rapid in vitro labeling of tetrazine containing protein with strained cycolooctene, 2

Pure GFP-1 at 3 μM in 20 mM ammonium acetate pH 7 was incubated with **2** at a final concentration of 39 μM. The reaction was run for 5 min. at room temperature before desalting with PD10 columns into 20 mM ammonium acetate pH 7. These samples were frozen and vacuum dried. Reactions of GFP and GFP-1 with no addition of **2** were run in parallel as controls.

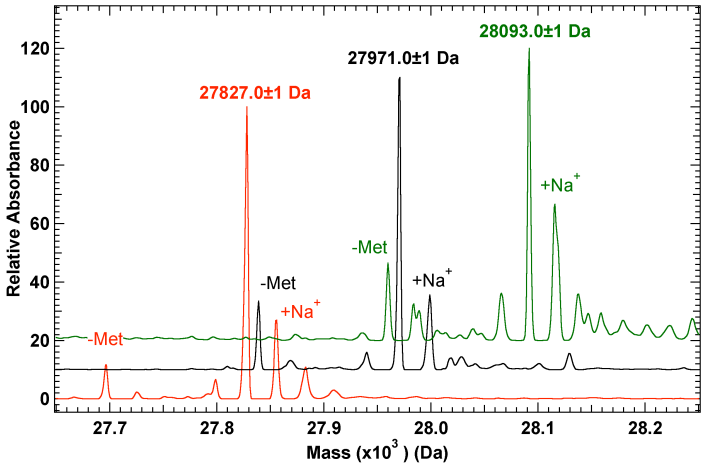
Rapid in vivo labeling of tetrazine containing protein with strained cycolooctene, 2

Cells (50 mL) containing expressed GFP-1 were pelleted, washed twice with 10 mL 1x wash buffer, and suspended in 5 mL PBS buffer pH 7.5. Washed cells (400 μL) containing GFP-1 in PBS buffer was incubated with **2** at a final concentration of 120 μM

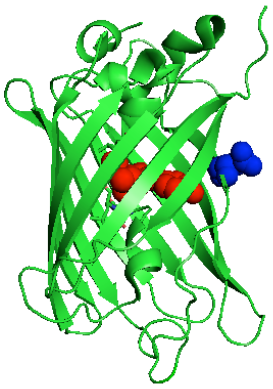
(final total reaction volume was 1 mL). The reaction was run for 45 min. at room temperature. The unreacted **2** was remove by pelleting the cells before the GFP-1 was purifying with BD-TALON cobalt ion-exchange chromatography as described above. The pure protein samples were frozen and vacuum dried. Control reactions of GFP and GFP-1 (no **2** added**)** were run in parallel under identical conditions.

MS Analysis of GFP, GFP-1 and GFP-3

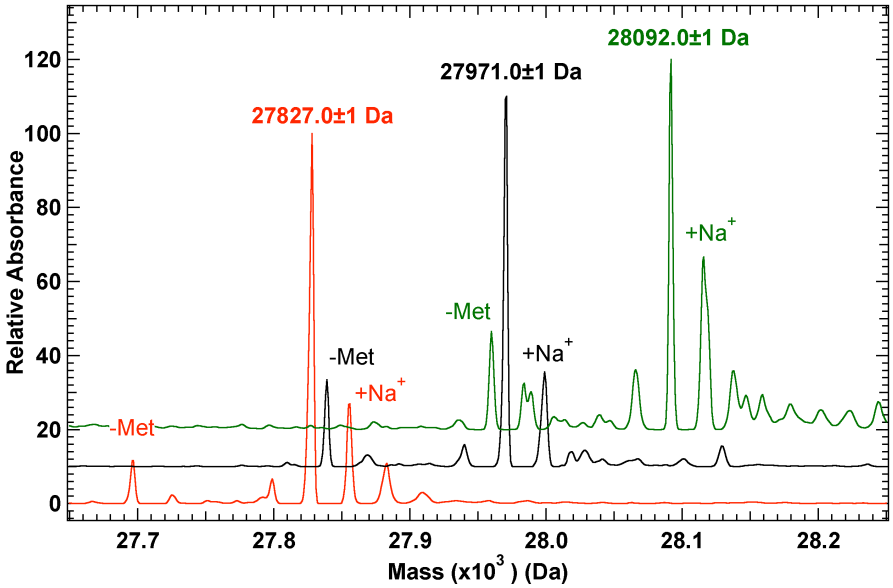
The samples in 20 mM ammonium acetate buffer pH 7 were dried overnight on a vacuum-line. Protein samples for full protein mass spectrometry were resuspended in 1:1 water:acetonitrile with 0.2% formic acid. The samples were analyzed at the Mass Spectrometry Facility at University of Illinois Urbana-Champaign under the direction of Dr. Furong Sun using their ESI-Q-Tof Ultima.



Sup. Figure 2. *In vitro* incubation of GFP and GFP-1 with **2.** ESI-MS of proteins; GFP-wt, GFP-1, and GFP-3, demonstrates specific and quantitative labeling of GFP-1 and no background labeling of GFP. Red trace: ESI-MS-Tof analysis of sfGFP incubated with **2** shows a single major peak at 27827.0 Da ±1 Da showing that there is no background reactivity of **2** with proteins. Black trace: ESI-MS-Tof analysis of GFP-1 shows a single major peak at 27971.0 Da ±1 Da. This shows the expected molecular weigh difference of 144 Da from native indicating a single efficient incorporation of **1** at the expected site. Green trace: ESI-MS-Tof analysis of GFP-1 incubated with **2** shows a single major peak at 28093.0 Da ±1 Da. This shows the expected molecular weight difference of 123 Da from GFP-1 demonstrating specific and quantitative conversion to GFP-3. Each sample did show a small peak at -131 ±1 Da indicating minor amounts of peptidase-based removal of N-terminal methionines and +22 sodium adducts.



Sup. Figure 3. Structure of GFP-1. The location of incorporation of 1 at site N150 in GFP protein is indicated by the blue space-filling amino acid (previously asn). Altering the amino acid at surface site 150 unconnected to the chromophore (red space fill) does not affect the stability or fluorescence of GFP. The distance from the center of the chromophore to the ß-carbon on site 150 is 12.3 Å.



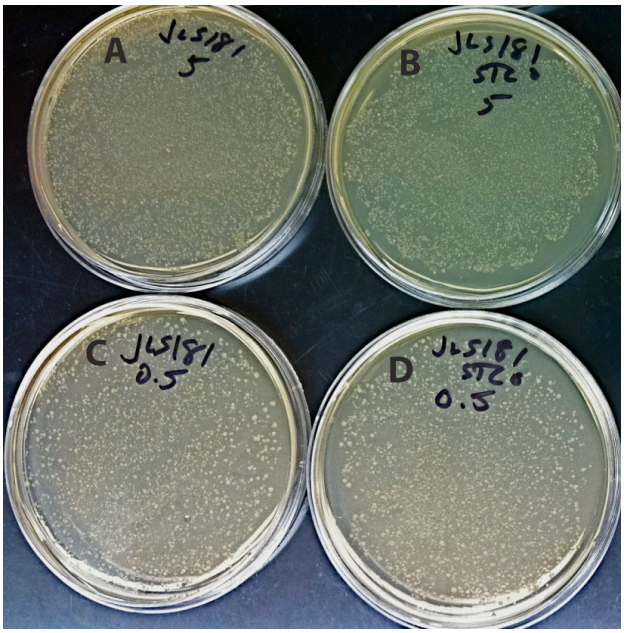
Sup. Figure 4. *In vivo* incubation of GFP and GFP-1 with **2.** ESI-MS of proteins; GFP-wt, GFP-1, and GFP-3, demonstrates specific and quantitative labeling of GFP-1 and no background labeling of GFP. Red trace: ESI-MS-Tof analysis of sfGFP incubated with **2** shows a single major peak at 27827.0 Da ±1 Da showing that there is no background reactivity of **2** with proteins. Black trace: ESI-MS-Tof analysis of GFP-1 shows a single major peak at 27971.0 Da ±1 Da. This shows the expected molecular weigh difference of 144 Da from native indicating a single efficient incorporation of **1** at the expected site. Green trace: ESI-MS-Tof analysis of GFP-1 incubated with **2** shows a single major peak at 28092.0 Da ±1 Da. This shows the expected molecular weight difference of 123 Da from GFP-1 demonstrating specific and quantitative conversion to GFP-3. Each sample did show a small peak at -131 ±1 Da indicating minor amounts of peptidase-based removal of N-terminal methionines and +22 sodium adducts.

***In vitro* kinetic analysis of GFP-1 with 2.**
A 200 mL GFP-1 cell pellet was resuspended in 15 mL 1x wash buffer and diluted in an additional 15 mL 1x wash buffer. A spatula tip of lysozyme (kept at -20 °C) was added, the cells were placed on a roller for five minutes, and then sonicated three times for 1 minute. The samples were centrifuged at 11,000 rcf for 30 minutes and the supernatant was collected in a 50 mL falcon tube. 1 mL bed volume of washed resin was added to the supernatant and then placed on a roller for 20 minutes to allow the protein to bind to the resin. The protein was filtered through a column and eluted off into 2.5 mL 1x elution buffer. The purified protein was desalted into 3.5 mL 1x PBS buffer and protein concentration was determined (0.08635 mM). The GFP-1 solution was diluted 10x to obtain a solution of 0.008635 mM. A solution of **2** in MeOH (4.50 mL 0.80 mg/mL) was dried down to obtain 3.24 mg and dissolved in 16 mL 1:4 EtOH: 1x PBS buffer to make a 1.35 mM solution. This stock solution was diluted to make five separate solutions of differing concentrations (1.35, 0.965, 0.685, 0.488, and 0.342 mM). Kinetic trials contained 30 µL of GFP-1 (0.008635 mM) in 2.5 mL of 1x PBS buffer at 22 °C and initiated by adding 50 µL of a **2** stock solution. Reactions were monitored by observing the fluorescence increase from product, GFP-3, forming (Excitation 488 nm with 2 nm slit, Emission 506 nm with 5 nm slit, 0.1 s integration time, and 2 s increments). Fluorescence measurements for each trial was run until a constant emission intensity. A unimolecular rate constant was obtained for each concentration and all five unimolecular rate constants were used to obtain a bimolecular rate constant.

***In vivo* kinetic analysis of GFP-1 with 2.**
Identical solutions of **2** from the *in vitro* experiments were used for five *in vivo* experiments. A 50 mL GFP150tet cell pellet was resuspended in 1x PBS buffer and equilibrated to 22 °C. Five kinetic trials containing 500 µL of the cell solution added to 2 mL of 1x PBS buffer were initiated by adding 50 µL of one of the stock solutions of **2** (27.3, 19.5, 13.7, 9.76, and 6.84 µM). Kinetics trials stirred continuously and monitored by observing the fluorescence increase from product, GFP-3, forming (Excitation 488 nm with 2 nm slit, Emission 506 nm with 5 nm slit, 0.1 s integration time, and 2 s increments). Fluorescence measurements for each trial was run until a constant emission intensity. A unimolecular rate constant was obtained for each concentration and all five unimolecular rate constants were used to obtain a bimolecular rate constant.

Toxicity assessment of 2 on in vivo labeling of GFP-1
Cells (5 mL) expressing GFP-1 (at ~40 hr of expression) were pelleted and suspended in 5 mL autoinduction medium without arabinose added (Sup table 1). The live cells were incubated with **2** (120 µM). A control reaction containing no **2** was run in parallel. The reaction was run for 30 min. at room temperature. Samples of each reaction (without removing **2**) were plated on LB Agar plates and allowed to incubate at 37 °C for 24 hr to assess toxicity. For the remaining cells were used to purify protein and verify complete labeling by MS. The unreacted **2** was remove by pelleting the cells before the GFP-1 was

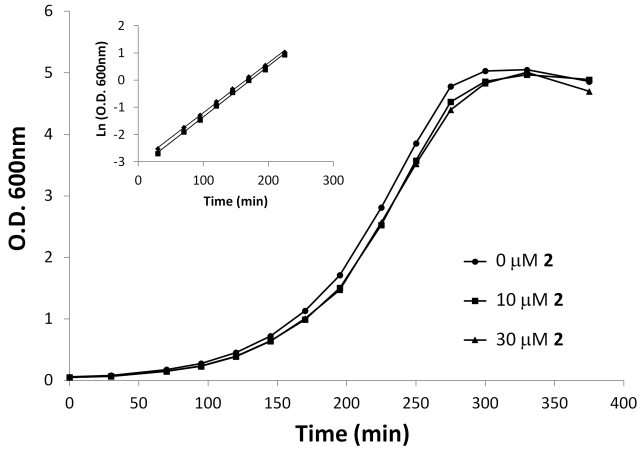
purifying with BD-TALON cobalt ion-exchange chromatography as described above. The pure protein samples were frozen and vacuum dried.



Sup Figure 5. Live cells containing GFP-1 reacted with **2** shows no effect on cell grow. Diluted samples from both reactions (5 μL and 0.5 μL volumes) were plated on LB Agar plates and allowed to incubate at 37 °C for 24 hr to assess toxicity. Diluted samples from both reactions (5 μL and 0.5 μL volumes) were plated on LB Agar plates and allowed to incubate at 37 °C for 24 hr to assess toxicity. No difference in colony forming units or growth rate was detected between cells treated with or without **2** indicating the *in vivo* reaction has minimal toxicity.

To verify that the labeling compound, **2**, is non-toxic to growing *E. coli*, it was added to growth media at concentrations used for protein labeling and kinetic experiments. DH10B *E. coli* cells co-transformed with the *pBad-sfGFP-150TAG* vector and the machinery plasmid *pDule-mtaF* were used to inoculate 25 mL of non-inducing medium containing 100 μg/mL Amp and 25 μg/mL Tet. The non-inducing medium cultures was grown to saturation with shaking at 37 °C, and 1 mL was used to inoculate 0.1 L non-inducing medium with 100 μg/mL Amp and 25 μg/mL Tet. Labeling compound, **2**, was dissolved in methanol and added to each culture at the time of inoculation (final concentrations 0, 10 and 30 μM). All three cultures received the same quantity of methanol. Optical densities of the cultures were monitored until the cultures reacted saturation. No toxicity was evident as indicated by identical cell

growth rates and levels of saturation (Sup. Figure 6.).



Sup. Figure 6: Growth curve of DH10B *E. coli* cells co-transformed with the *pBad-sfGFP-150TAG* vector and the machinery plasmid *pDule-mtaF* in the presence of labeling compound, **2**. Inset: The natural logarithm of the optical density plotted as a function of time during exponential growth phase. The identical slope of each curve demonstrates the cultures grew at the same rate (doubling time ~37 min).

References

1. Scott, E. S. and Audrieth, L. F. *J. Org. Chem*. **1954**, *19*, 733-741.

2. Fields, S. C., et. al. *J. Org. Chem*. **1994**, *59*, 8284- 8287.

3. Taylor, M. T.; Blackman, M. L.; Dmitrenko, O.; Fox, J. M. *J. Am. Chem. Soc.*, **2011**, *133*,9646–9649.

Full reference 6e from manuscript:

6e. Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. *ACS Chem Biol* **2008**, *3*, 373.