Supporting Information

A Highly Efficient Catalyst for Oxime Ligation and Hydrazone-Oxime Exchange Suitable for Bioconjugation

Mohammad Rashidian[†], Mohammad M. Mahmoodi[†], Rachit Shah[‡], Jonathan K. Dozier[†], Carston R. Wagner[‡] and Mark D. Distefano[*†]

†Department of Chemistry, University of Minnesota, Minneapolis, MN 55455 (USA)

‡Department of Medicinal Chemistry, University of Minnesota, Minneapolis, MN 55455 (USA)

* Author to whom correspondence should be addressed. Tel: (612) 624-0544; Fax (612) 626-7541.

Email: **diste001@umn.edu**

Contents

General. All synthetic reactions were carried out at rt and stirred magnetically unless otherwise noted. TLC was performed on precoated (250 mm) silica gel 60 F-254 plates (Merck). Plates were visualized by staining with $KMnO_4$ or with a hand-held UV lamp. Flash chromatography was performed using a Biotage® instrument. Deuterated NMR solvents were purchased from Cambridge Isotope Laboratories, Inc. ^{1}H NMR spectra were obtained at 500 MHz; ^{13}C NMR spectra were obtained at 125 MHz. All NMR spectra were acquired on Varian instruments at 25 °C. Chemical shifts are reported in ppm and *J* values are in Hz. Fluorescence assay data were obtained using a Varian Cary Eclipse Fluorescence Spectrophotometer. MS spectra for synthetic reactions were obtained on a Bruker BioTOF II instrument. Yeast PFTase was prepared as previously described.[1] Protein LC/MS analyses were performed using a Waters Synapt G2 Quadropole TOF mass spectrometer instrument. MALDI-MS analyses were performed with a Bruker MALDI TOF spectrometer instrument. Aminooxy alexafluor-488 (**5**) was from AnaSpec; hydrazide agarose beads were from Thermo Scientific. PEG aminooxy 10,000 MW was from NOF America Corporation. All solvents were of HPLC grade. All other reagents were from Sigma Aldrich.

Abbreviations.

mPDA, *m*-Phenylenediamine

FPP, Farnesyl diphosphate

DTT, Dithiothreitol

ESI-MS, Electrospray ionization mass spectrometry

RP-HPLC, Reversed-phase high-pressure liquid chromatography

PB, Phosphate buffer

PEG, Polyethylene glycol

DMAP, 4-Dimethylaminopyridine

PFTase, Protein farnesyl transferase

DMF, Dimethylformamide

DBU, 1,8-Diazabicycloundec-7-ene

TEA, Triethylamine

Tris, Tris(hydroxymethyl)aminomethane

EDTA, Ethylenediaminetetraacetic acid

GFP, Green fluorescent protein

CNTF, Ciliary neurotrophic factor

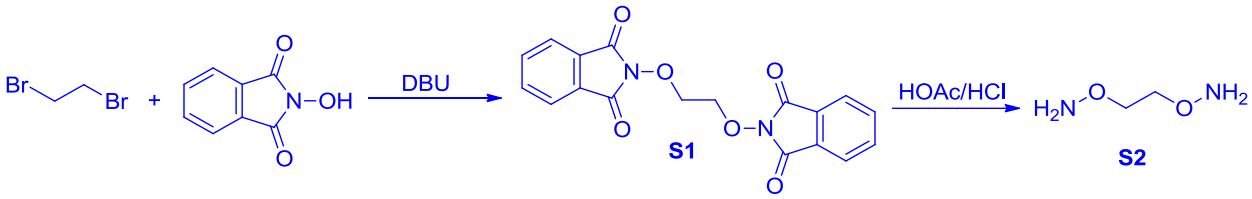
CD, Circular dichroism

DHFR, Dihydrofolate reductase

pAcF, *p*-Acetylphenylalanine

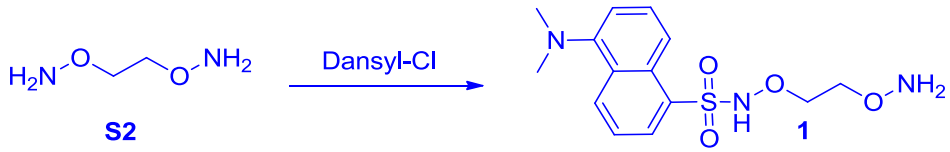
Synthesis of compound 1.

O,O'-(ethane-1,2-diyl)bis(hydroxylamine) S2.



Compound **S2** was synthesized according to a modified literature procedure.[2] To a solution of *N*-hydroxyphthalimide (20 g, 122.5 mmol) in DMF (120 mL), DBU (18.3 mL, 122.5 mmol) was added dropwise with stirring to give a very deep red solution. To this solution was added 1,2-dibromoethane (5.37 mL, 61.2 mmol) and the mixture was heated at 85 °C for 1 h during which time the deep red color faded to produce a colorless solution. The resulting solution was poured into ice and the resulting precipitate was filtered and washed with cold H_2O (50 mL) followed by cold CH_3CN (30 mL). The crude 1,2-diphthalimidooxyethane (**S1**) was recrystallized from n-butanol (3 L). A suspension of 1,2-diphthalimideooxyethane (16.0 g, 41.6 mmol) in glacial acetic acid/HCl (50 mL, 30:20, v/v) was heated at 115 °C for 3 h to give a clear solution. All the solvent and reagents were removed *in vacuo*. H_2O (10 mL) was added to the solid and the suspension was filtered and washed with HCl (6 M). The combined filtrate was collected and reduced to dryness. The crude product was recrystallized from EtOH/H_2O (5:1, v/v) to give the compound **1** ($C_2H_8N_2O_2$·2HCl salt) as white crystals (3.7 g, 22.4 mmol, 37 % over two steps and two recrystallizations); mp 228 °C; ^{1}H NMR (δ, DMSO): 11.16 (s, 6H), 4.26 (s, 4H). ^{13}C NMR (δ, DMSO): 72.21.

N-(2-(aminooxy)ethoxy)-5-(dimethylamino)naphthalene-1-sulfonamide (1).



To a solution of **S2** (61 mg, 0.37 mmol) in CH_2Cl_2 (3 mL), TEA (123 µL, 0.87 mmol) and DMAP (5 mg, 0.04 mmol) were added with stirring to give a clear solution. To this solution was added dansyl chloride (20 mg, 0.074 mmol) and the mixture was stirred at rt for 3 h. The solvent was removed *in vacuo* and the crude product was further purified by silica gel column chromatography with gradient elution (EtOAc:MeOH in 0.5% TEA) from 1:0 (v/v) going to 9:1 (v/v) to afford 14.2 mg of compound **1** (0.044 mmol, 59% yield) as a yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.600 (d, J = 8.0 Hz, 1H), 8.342 (d, J = 7.5 Hz, 1H), 8.298 (d, J = 9.0 Hz, 1H), 7.554 (m, 2H), 7.180 (d, J = 7.5 Hz, 1H), 4.083 (t, J = 4.5 Hz, 2H), 3.774 (t, J = 4.5 Hz,1H), 2.886 (s, 6H), ^{13}C NMR (125 MHz, $CDCl_3$) δ 132.18, 131.73, 128.85, 123.23, 118.18, 115.31, 75.03, 72.56, 45.4.

Kinetic analysis of oxime ligation reaction: rate constant analysis.

Dawson and coworkers originally established that the kinetics of oxime ligation reactions catalyzed by aniline fit a second order model that is first order in both aldehyde and alkoxyamine.[3] They also showed that the apparent second order rate constant for this process varied with aniline concentration. Wen-jun et al extended those observations by demonstrating that the apparent second order rate constant varied linearly with aniline.[4] Accordingly, we analyzed the rate data for the model reactions described here and confirmed that both the aniline- and mPDA-catalyzed reactions vary linearly with catalyst concentration. In order to obtain an

equation to curve fit the data, we started with the previously reported rate equation for the aniline catalyzed oxime reaction[3]:

Eq. S1 $V=k_{obs}\bullet[aldehyde][aminooxy]$

Where V is the rate of the reaction which is the rate of formation of the oxime as the end product, and k_{obs} is the observed rate constant of the reaction.

In general, for a second order reaction between A and B as reactants, the equation below for the kinetics of the reaction can be derived:

Eq. S2 $$\frac{[A]}{[B]} = \frac{[A]_0}{[B]_0} e^{([A]_0-[B]_0)kt}$$

By substituting [A] with {$[Aminooxy]_o$-[Oxime]} and [B] with {$[Aldehyde]_o$-[Oxime]} and solving the equation for [Oxime] the following relationship is obtained:

Eq. S3 $$[Oxime]=\frac{[Aminooxy]_0*[e^{([Aminooxy]_0-[Aldehyde]_0)kt}-1]}{\frac{[Aminooxy]_0}{[Aldehyde]_0}[e^{([Aminooxy]_0-[Aldehyde]_0)kt}]-1}$$

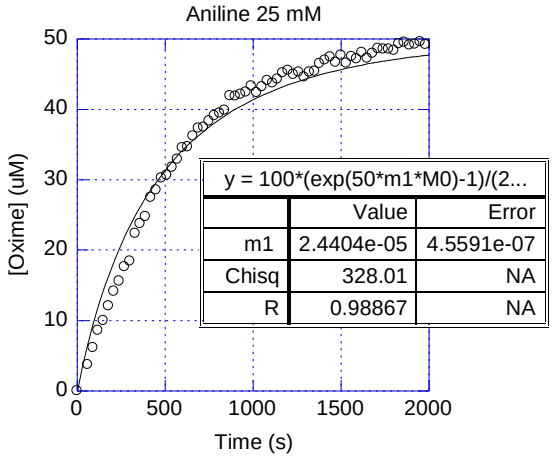
where [Oxime] is the concentration of product, $[Aminooxy]_o$ is the initial concentration of aminooxy-dansyl **1** and $[Aldehyde]_o$ is the initial concentration of aldehyde.

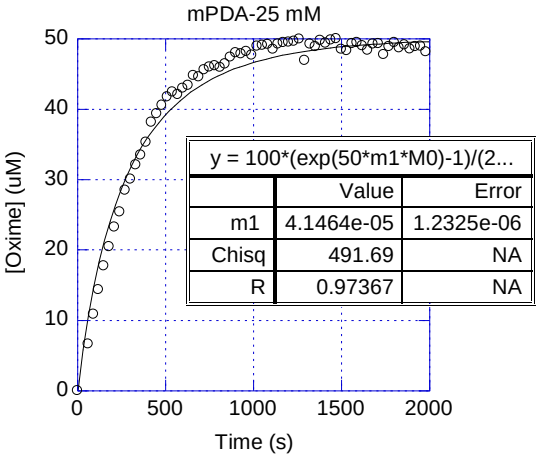
Fluorescence assay data analysis. Oxime-forming reactions were monitored by recording the increase in dansyl group fluorescence (λ_{ex}=340 nm, λ_{em}=505 nm) over time that occurred upon product formation. Hence, the raw experimental data was obtained in "fluorescence intensity/time" units. In order to obtain that data in units of "μM/time", it was necessary to convert the fluorescence values to concentrations. That was accomplished by first calculating the difference between the fluorescence intensity of the product and the starting reagents. Assuming complete reaction (100% conversion), that difference corresponds only to the fluorescence of the

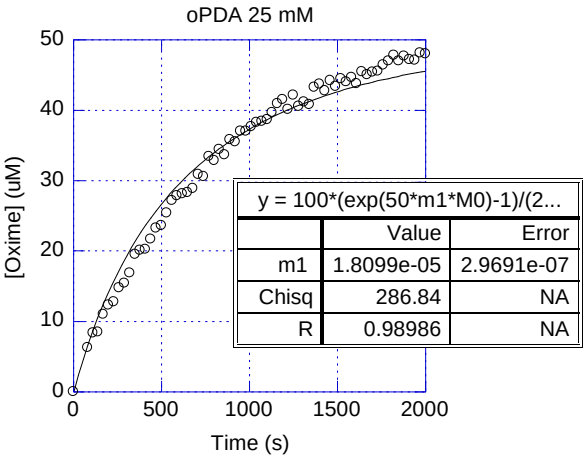
total amount of product. Thus, the fluorescence of the starting reagents was subtracted from the raw fluorescence data followed by dividing by the fluorescence difference between the final product and starting material and then multiplied by the total concentration of the limiting reagent (in μM) which then gives the data in units of "μM/time". That calculation is summarized in equation S4:

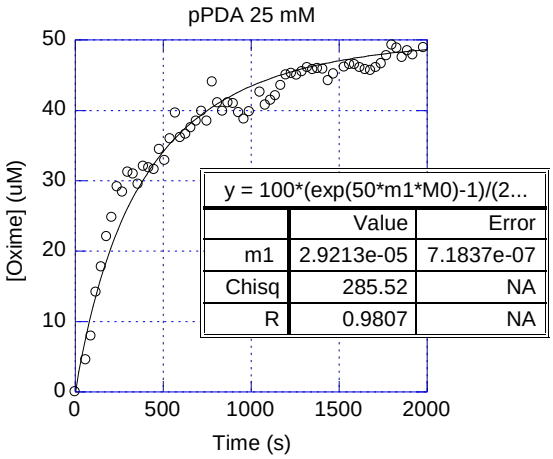
Eq. S4 $$[Oxime]_t = \frac{[F]_t - [F]_0}{[F]_{max} - [F]_0} \times [limiting\ reagent]_0$$

In that expression, $[Oxime]_t$ is the concentration of product at each given time, $[F]_t$ is the fluorescence of the reaction mixture at each given time, $[F]_0$ is the initial fluorescence of starting reaction mixture, $[F]_{max}$ is the maximum fluorescence of the reaction mixture at the end of the reaction and $[limiting\ reagent]_0$ is the initial concentration of the limiting reagent.


Aniline 25 mM
[Oxime] (uM)
Time (s)
y = 100*(exp(50*m1*M0)-1)/(2...
Value Error
m1 2.4404e-05 4.5591e-07
Chisq 328.01 NA
R 0.98867 NA


mPDA-25 mM
[Oxime] (uM)
Time (s)
y = 100*(exp(50*m1*M0)-1)/(2...
Value Error
m1 4.1464e-05 1.2325e-06
Chisq 491.69 NA
R 0.97367 NA


oPDA 25 mM
[Oxime] (uM)
Time (s)
y = 100*(exp(50*m1*M0)-1)/(2...
Value Error
m1 1.8099e-05 2.9691e-07
Chisq 286.84 NA
R 0.98986 NA


pPDA 25 mM
[Oxime] (uM)
Time (s)
y = 100*(exp(50*m1*M0)-1)/(2...
Value Error
m1 2.9213e-05 7.1837e-07
Chisq 285.52 NA
R 0.9807 NA

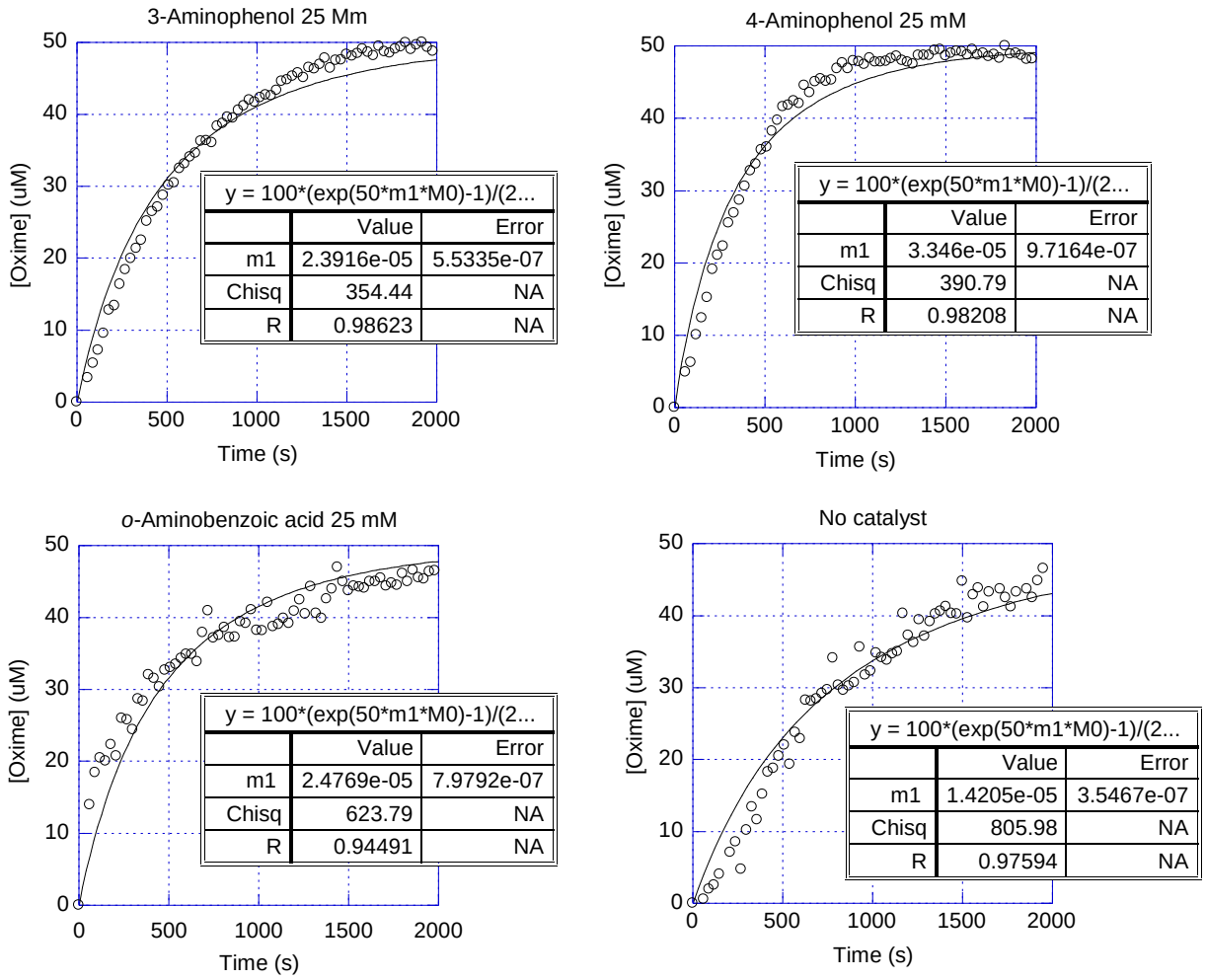


Figure S1. Kinetic analysis of oxime ligation reaction, using 50 μM aldehyde (citral), 100 μM aminooxy-dansyl **1**, and 25 mM of catalysts. The k_{obs} values were obtained by fitting the experimental data to equation S3.

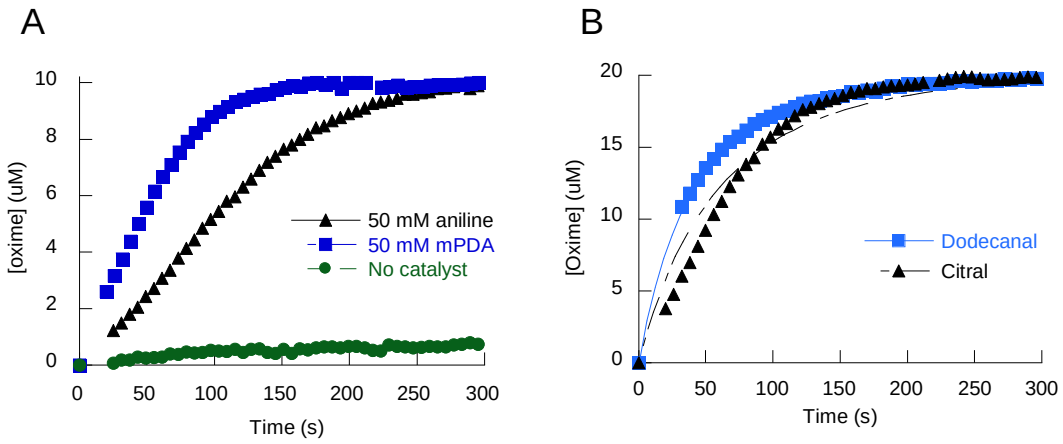


Figure S2. A) Analysis of the reaction of aminooxy-dansyl **1** (10 µM) with citral (50 µM), a conjugated hydrophobic aldehyde, in presence of either 50 mM mPDA (squares) or 50 mM aniline (triangles) or no catalyst (circles). B) Kinetic analysis of oxime ligation reaction, using 50 µM aldehyde (citral or dodecanal), 20 µM aminooxy-dansyl **1** and 50 mM mPDA.

Kinetic analysis of oxime ligation between 2-pentanone and aminooxy-dansyl 1.

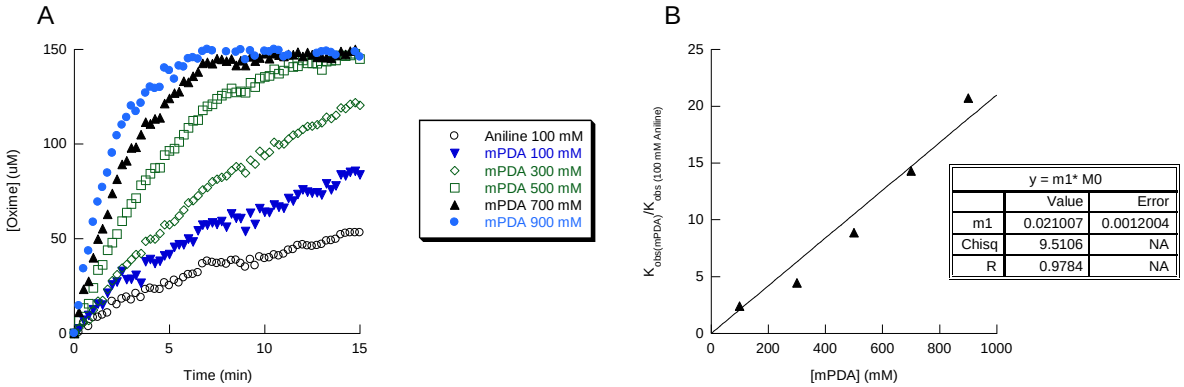
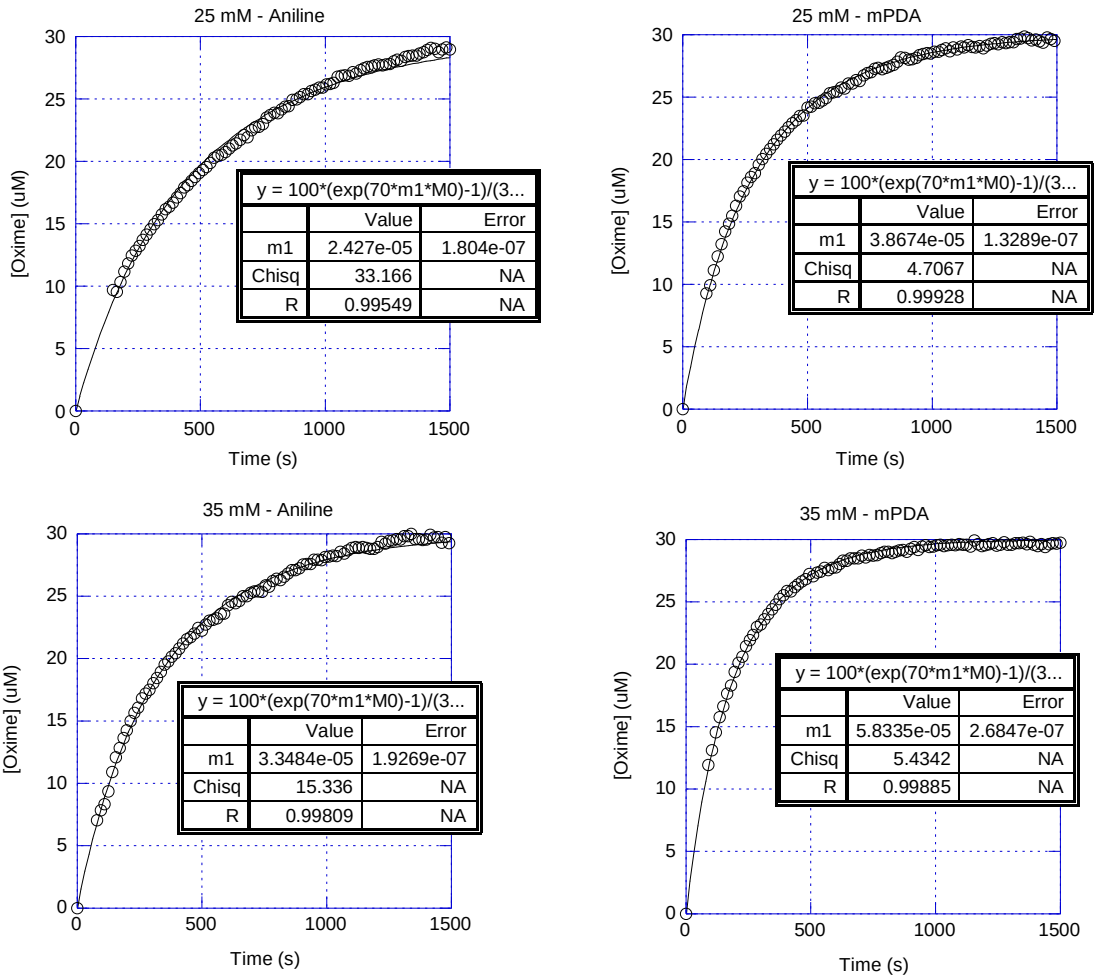


Figure S3. A) Kinetic analysis of oxime ligation between a ketone (2-pentanone) and aminooxy-dansyl **1**. Reaction mixtures contained Tris·HCl (50 mM, pH 7.5), aminooxy-dansyl **1** (150 μM), 0.4 % (*w/v*) *n*-dodecyl-*ß*-D-maltoside, 5 mM ketone (2-pentanone) and either 100 mM aniline or varying concentrations of mPDA, in a final volume of 200 μL. The reaction mixtures were equilibrated at rt for 1 min, initiated by the addition of the aminooxy, and monitored for an increase in fluorescence (λ_{ex}=340 nm, λ_{em}=505 nm) for approximately 4 h. B) Plot of ratio of rate constant of mPDA catalyzed oxime ligation over aniline (100 mM) catalyzed oxime reaction vs concentration of mPDA used. The k_{obs} values were obtained by fitting the experimental data to equation S3.

Effect of the catalyst concentration on the k_{obs}. In order to analyze the effect of the catalyst concentration on the k_{obs}, reactions with different catalyst concentrations were performed (Figure S4). After analyzing the kinetics of the reactions and curve-fitting the data to the second order kinetic equation described above (Eq. S3), we observed that in the cases of both aniline and mPDA as the catalyst, a good fit was obtained to the kinetic expression derived above. Additional analysis revealed that reaction rate is linear with respect to the catalyst concentration (Figure S5) which is in agreement with previous reports.[4]



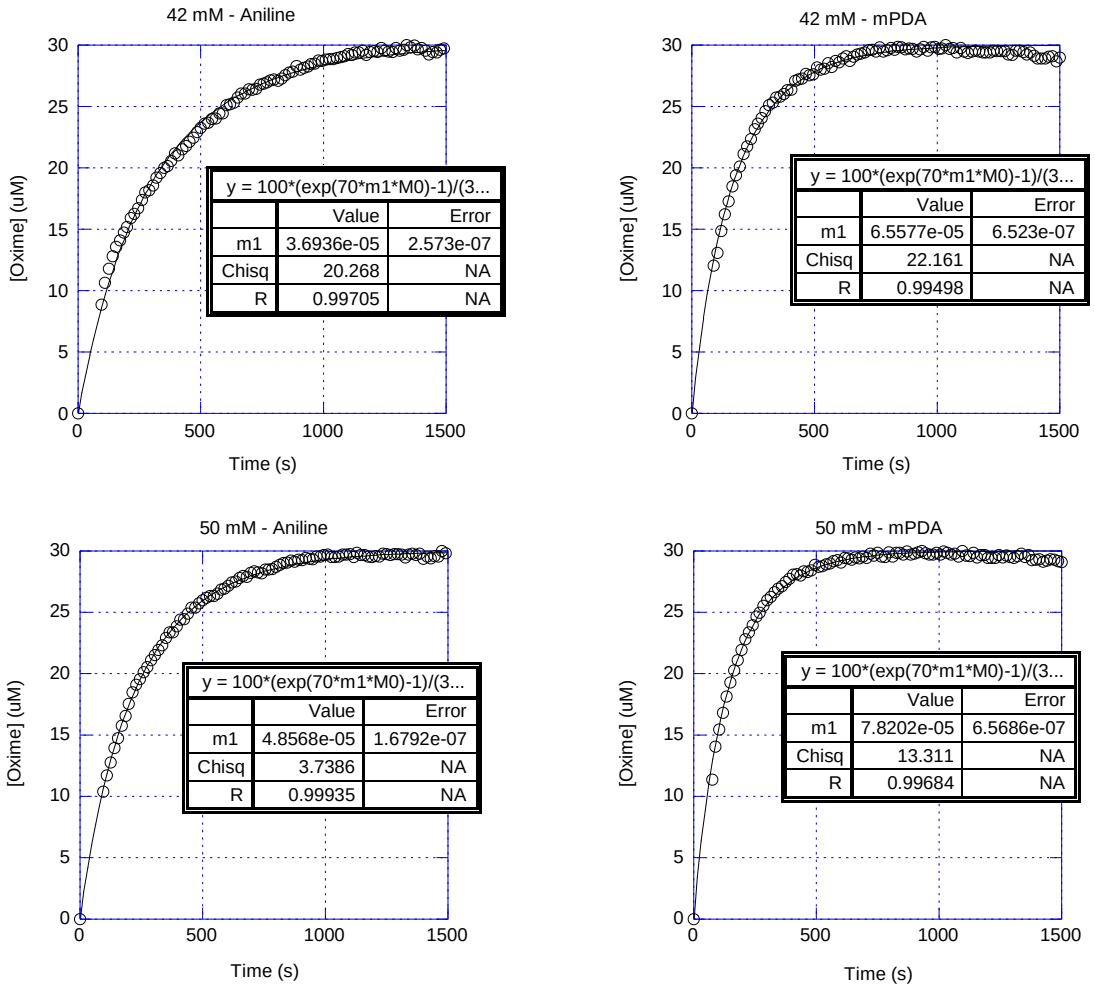
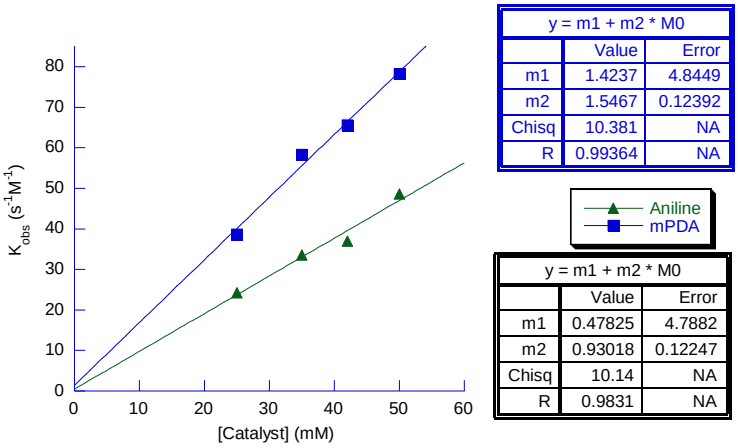


Figure S4. Kinetic analysis of oxime ligation reaction, using 30 µM aldehyde (citral), 100 µM aminooxy-dansyl **1**, varying concentrations of either *m*-phenylenediamine or aniline (from 25 to 50 µM). The data was fit to the equation S3.



y = m1 + m2 * M0		
	Value	Error
m1	1.4237	4.8449
m2	1.5467	0.12392
Chisq	10.381	NA
R	0.99364	NA

y = m1 + m2 * M0		
	Value	Error
m1	0.47825	4.7882
m2	0.93018	0.12247
Chisq	10.14	NA
R	0.9831	NA

Figure S5. Kinetic analysis of oxime ligation reaction to determine the relationship between the observed rate constant and the catalyst concentration. Experiments were performed using 30 μM aldehyde (citral), 100 μM aminooxy-dansyl **1** and varying concentrations of either *m*-phenylenediamine or aniline (from 25 to 50 μM).

GFP-CVIA preparation. Protein was prepared as previously described with one modification.[5] In the final phenyl sepharose chromatography step, after the protein was loaded onto the column and washed with buffer as explained in the original work, the protein was eluted from column by adding water instead of buffer.

LC-MS analysis of proteins for determination of prenylation and labeling efficiency. Purified prenylated GFP (**4a**) and pure GFP-CVIA (**3**) were analyzed by LC-MS to ensure complete prenylation. Proteins were stored in Tris·HCl (50 mM, pH 7.5) prior to injection into the LC-MS instrument. Crude reaction mixtures of GFP-aldehyde **4a** and aminooxy **1** catalyzed by either aniline or *m*-phenylenediamine were analyzed by LC-MS to ensure complete ligation in both cases of the catalysts. The LC-MS method used was a gradient of 0–100% solvent A (H_2O, 0.1% HCO_2H) to B (CH_3CN, 0.1% HCO_2H) in 25 min.

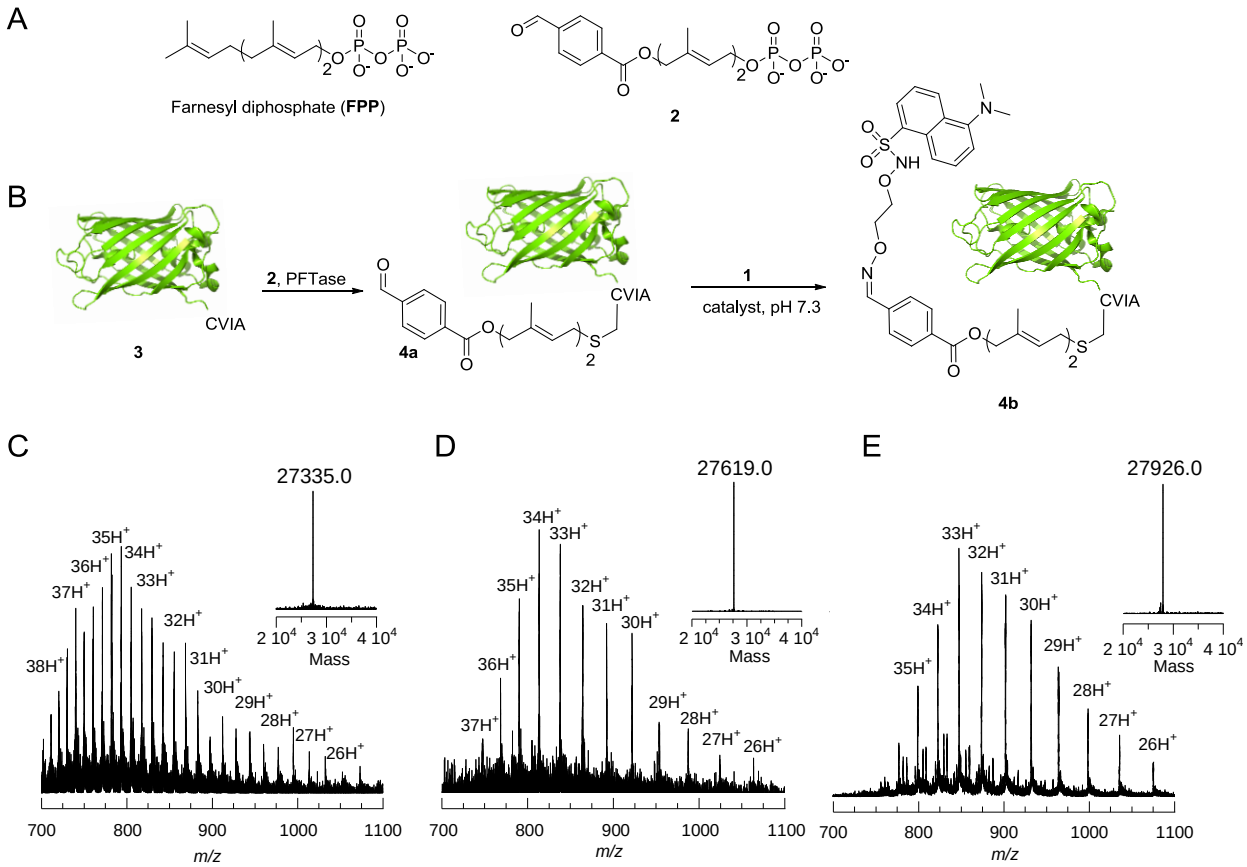


Figure S6. A) Structures of farnesyl diphosphate (**FPP**) and formylbenzoyl-oxy geranyl diphosphate (**2**). B) Schematic representation of prenylation of GFP-CVIA **3** with aldehyde-containing analog **2** to yield the prenylated product **4a** and subsequent oxime ligation with **1** to yield oxime **4b**. C, D and E) ESI MS analysis of **3**, **4a** and **4b** with the deconvoluted mass spectra shown in the insets, respectively.

Kinetic analysis of protein labeling via oxime ligation. Reaction mixtures contained PB (100 mM, pH 7.0), 10 μM GFP-aldehyde **3**, 50 μM aminooxy-dansyl **1** and varying concentrations of *m*-phenylenediamine or 100 mM aniline, in a final volume of 250 μL. The reaction mixtures were equilibrated at rt for 1 min, initiated by the addition of the catalyst, and monitored for an increase in fluorescence (λ_{ex}=340 nm, λ_{em}=505 nm) for approximately 100 min.

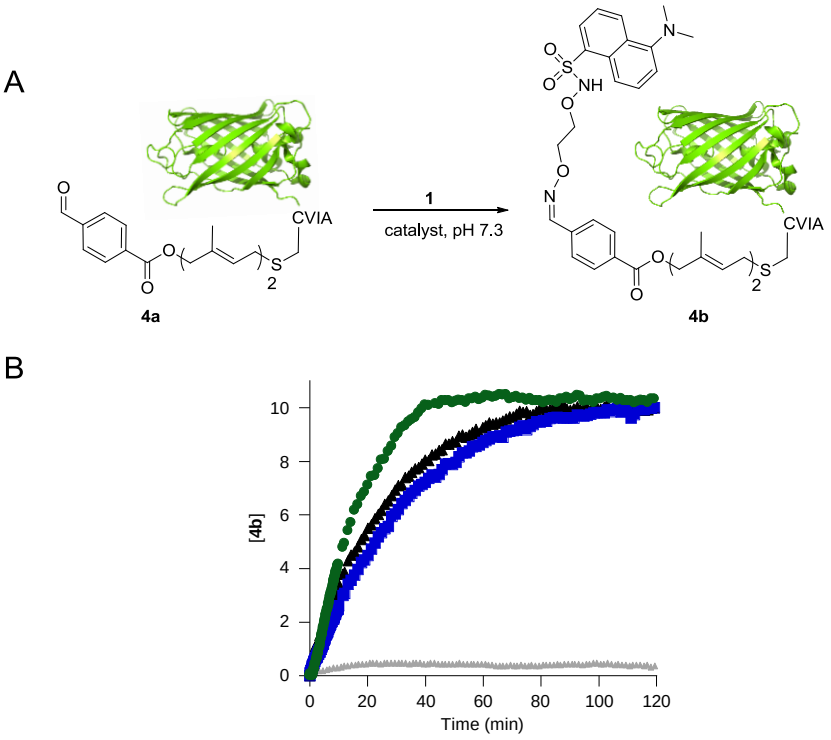


Figure S7. Schematic representation of labeling of GFP-aldehyde **4a** with aminooxy **1** to yield oxime **4b**. B) Kinetic analysis of oxime ligation reaction between **4a** (10 µM) and **1** (50 µM) using 50 mM *m*-phenylenediamine (green circles), 50 mM aniline (blue squares) or no catalyst (black triangles). Gray triangles: GFP-CVIA **3** (10 µM, with no aldehyde attached to it) is treated with aminooxy **1** (50 µM) in presence of 50 mM *m*-phenylenediamine to further confirm that reaction is truly bioorthogonal.

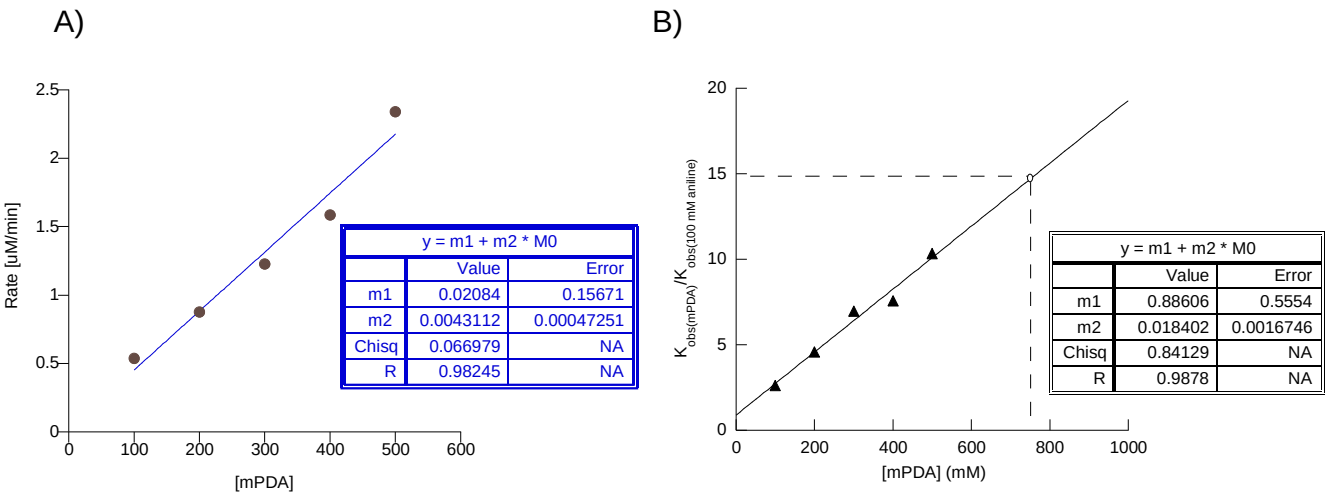


Figure S8. A) Kinetic analysis of oxime ligation reaction between aminooxy **1** and aldehyde-GFP **4a** to determine the relationship between the observed initial rate of the reaction and the catalyst concentration. Experiments were performed using 10 µM aldehyde-GFP **4a**, 50 µM aminooxy-dansyl **1** and varying concentrations of *m*-phenylenediamine. In this case when higher concentrations of the catalyst (>500 mM) were used (where the ratio [cat]/[aminooxy] was high), the Schiff base between the catalyst and aldehyde becomes a significant product, complicating the kinetic analysis; hence that data was omitted from this analysis. It should be noted if it is necessary to employ very high catalyst concentrations, Schiff base formation can be countered by raising concentration of the aminooxy reagent. B) Plot of ratio of rate constant of mPDA catalyzed oxime ligation over aniline (100 mM) catalyzed oxime reaction vs concentration of mPDA used. Extrapolation of the line shows that at 750 mM of mPDA reaction would be 15 times more efficient in case of mPDA relative to 100 mM aniline.

Table S1. Kinetic analysis of oxime ligation reaction between aldehyde-GFP **4a** and aminooxy-dansyl **1**, using aniline or mPDA as catalysts.

[Catalyst][a] mM	Initial rate ($\mu M\ s^{-1}$)	$\frac{\text{Rate (x mM mPDA)}}{\text{Rate (100 mM aniline)}}$	k_{obs} [b] ($s^{-1}\ M^{-1}$)	$k_{obs(\text{mPDA})}/k_{obs(\text{aniline})}$	$t_{1/2}$ (min)
100 aniline	0.226	-	10.35±0.05	-	23.7
100 mPDA	0.539	2.38	27.04±0.33	2.61	9.06
200 mPDA	0.878	3.87	48.45±0.74	4.68	5.18
300 mPDA	1.228	5.42	71.84±1.53	6.94	3.41
400 mPDA	1.586	7.00	78.11±1.01	7.54	3.14
500 mPDA	2.343	10.34	107.36±2.13	10.37	2.3

[a]Reactions were performed using 10 μM aldehyde-GFP **4a**, 50 μM aminooxy-dansyl **1**, and the catalyst concentrations given.
[b]The k_{obs} values were obtained by fitting the experimental data to Equation S3 using Kaleidagraph v4.1.3. The values are provided ± the standard error obtained from the curve fit.

Table S2. Kinetic analysis of the release of the hydrazone immobilized protein into fluorescently labeled oxime protein via hydrazone-oxime exchange reaction.

[Catalyst][a] mM	k_{obs} [b] (h^{-1})	$t_{1/2}$ (h)	$k_{obs(\text{mPDA})}/k_{obs(\text{aniline})}$
100 aniline	0.016±0.001	43.6	-
750 mPDA	0.237±0.009	2.9	15

[a]Immobilized protein was incubated with aminooxy fluorophore **5** (1 mM) and catalyst, followed by analysis of the amount of released protein in the solution via SDS-PAGE.
[b]The k_{obs} values were obtained by fitting the experimental data to a pseudo first order reaction. The values are provided ± the standard error obtained from the curve fit.

CNTF-CVIA preparation. The gene for Ciliary Neurotrophic Factor (CNTF) was purchased from DNA 2.0 on a pJexpress414 vector. The sequence of this synthetic gene (Menlo Park, CA) is given below. The CVIA portion of the protein was added using an Invitrogen QuikChange Site Directed Mutagenesis kit (catalogue #200523 Menlo Park, CA) following the manufacturer's instructions.

The forward primer was tatggtgcaaaagataaacaaatgtgcgtgattgcgtaactcgagccccctag, and the reverse primer was ctaggggctcgagttacgcaatcacgcacatttgtttatcttttgcaccata. A plasmid containing the CNTF-CVIA was transformed into BL21(DE3)pLysS *E. coli* cells.

BL21(DE3)pLysS *E. coli* cells containing the CNTF-CVIA plasmid were plated on LB-Agar plates containing 100 μg/mL ampicillin. These plates were grown overnight at 37 ^{o}C. Single colonies were then picked and used to innoculate 50 mL of LB media containing 100 μg/mL ampicillin. These flasks were grown overnight with shaking at 250 rpm at 37 ^{o}C. 10 mL of the overnight growth was added to 1 L of LB media containing 100 μg/mL of ampicillin and incubated at 37 ^{o}C with shaking at 250 rpm. This culture was grown to an OD_{600} of 0.8 at which time protein expression was induced by the addition of 1 mL of 1M IPTG. Cultures were then incubated for an additional 4 h by shaking at 250 rpm at 37 ^{o}C followed by harvesting by centrifugation at 5,400g for 10 min. *E. coli* cell pellets were stored at -80 ^{o}C. A cell pellet corresponding to 1 L of cell growth was suspended in 50 mL of buffer containing 50 mM Tris•HCl, pH 7.5 and 5 mM 2-mercaptoethanol. This was then subjected to pulse sonication (10 sec sonication / 10 sec off) at 50 W for a total sonication time of 5 min. The sonicated solution was then centrifuged at 13,000g for 30 min. The supernatant was removed and the insoluble protein pellet containing the CNTF-CVIA was suspended in 30 mL of buffer containing 10 mM Tris•HCl, pH 7.5, 6 M guanidine•HCl, 5 mM 2-mercaptoethanol, 100 mM NaH_2PO_4, and 20 mM imidiazole. To thoroughly suspend the protein, the solution was subjected again to pulse sonication (10 sec sonication / 10 sec off) at 50 W for a total sonication time of 5 min. This solution was then added at 3 mL/h to 500 mL of refolding buffer (50 mM Tris•HCl, pH 7.5, 0.5 M NaCl, 10 mM CHAPS, 2 mM DTT). This solution was then left to stir overnight at 4 ^{o}C and then concentrated using an Amicon Centriprep centrifugation (10,000 MW cut-off) according to

manufacturer's instructions. This yielded 9.2 mL of a 6.5 mg/mL protein solution. This solution was diluted in half to store in 40% glycerol at -80 ^{o}C.

CNTF gene from DNA 2.0 containing N-terminal His-tag.

Protein sequence:

MHHHHHHLVP • RGSMAFAEQT • PLTLHRRDLC • SRSIWLARKI • RSDLTALMES • YVKHQGLNKN • INLDSVDGVP • VASTDRWSEM • TEAERLQENL • QAYRTFQGML • TKLLEDQRVH • FTPTEGDFHQ • AIHTLMLQVS • AFAYQLEELM • VLLEQKIPEN • EADGMPATVG • DGGLFEKKLW • GLKVLQELSQ • WTVRSIHDLR • VISSHQMGIS • ALESHYGAKD • KQM

DNA sequence (with ATG start underlined):

```
aggagatatctagaatgcaccatcatcatcaccacctggttccacgcggtagcatggccttc
   E   I   S   R   M   H   H   H   H   H   H   L   V   P   R   G   S   M   A   F
gctgaacaaaccccgctgacgctgcaccgtcgcgatctgtgctcccgtagcatctggctg
 A   E   Q   T   P   L   T   L   H   R   R   D   L   C   S   R   S   I   W   L
gcccgcaagattcgtagcgacctgaccgcattgatggaatcttacgttaagcatcaaggt
 A   R   K   I   R   S   D   L   T   A   L   M   E   S   Y   V   K   H   Q   G
ctgaacaaaaacattaatctggatagcgtggatggtgttccggtcgcgagcacggaccgt
 L   N   K   N   I   N   L   D   S   V   D   G   V   P   V   A   S   T   D   R
tggagcgaaatgaccgaagcggagcgcctgcaggagaacctgcaggcatatcgtaccttc
 W   S   E   M   T   E   A   E   R   L   Q   E   N   L   Q   A   Y   R   T   F
caaggtatgctgaccaaactgctggaggatcaacgcgtgcactttacgccgaccgaaggt
 Q   G   M   L   T   K   L   L   E   D   Q   R   V   H   F   T   P   T   E   G
gattttcatcaggcgatccacaccctgatgctgcaagttagcgcttttgcttaccagctg
 D   F   H   Q   A   I   H   T   L   M   L   Q   V   S   A   F   A   Y   Q   L
gaagagctgatggtgttgttggaacagaagattccggagaatgaggccgacggtatgccg
 E   E   L   M   V   L   L   E   Q   K   I   P   E   N   E   A   D   G   M   P
gcgaccgtcggcgacggtggcctgttcgaaaagaagctgtggggcctgaaagttctgcag
 A   T   V   G   D   G   G   L   F   E   K   K   L   W   G   L   K   V   L   Q
gagctgagccagtggacggtccgttccattcatgacctgcgtgtgattagcagccaccaa
 E   L   S   Q   W   T   V   R   S   I   H   D   L   R   V   I   S   S   H   Q
atgggtatcagcgcactggaatctcattatggtgcaaaagataaacaaatgtaactcgag
 M   G   I   S   A   L   E   S   H   Y   G   A   K   D   K   Q   M   -   L   E
```

Primers for CVIA mutant construction (CVIA sequences underlined).

Forward Primer

5′tatggtgcaaaagataaacaaatgtgcgtgattgcgtaactcgagccccctag-3′

Translated 5′-3′ Frame 1
tatggtgcaaaagataaacaaatgtgcgtgattgcgtaactcgagccccctag
 Y G A K D K Q M C V I A - L E P P

Reverse Primer

5′-ctaggggctcgagttacgcaatcacgcacatttgtttatcttttgcaccata-3′

Translated 3′-5′ Frame 1
tatggtgcaaaagataaacaaatgtgcgtgattgcgtaactcgagccccctag
 Y G A K D K Q M C V I A - L E P P

Coupling reaction between aldehyde-labeled CNTF-CVIA (9) with alexafluor-488 (5). Alexafluor-488 (**5**) (4.2 µL of 3.2 mM solution in DMSO) was added to 42 µL of **9** (stock solution of 60 µM in Tris·HCl (50 mM, pH 7.5)). PB (2 M, pH 7, 2.5 µL) was added and the reaction was initiated by adding 50 mM *m*-phenylenediamine (stock solution of 1.5 M in 0.3 M PB, pH 7.0) and was allowed to proceed for 2 h at rt. LS-MS analysis of the sample showed only oxime ligated protein and no free aldehyde was detected indicating a complete reaction in both prenylation and oxime ligation reactions.

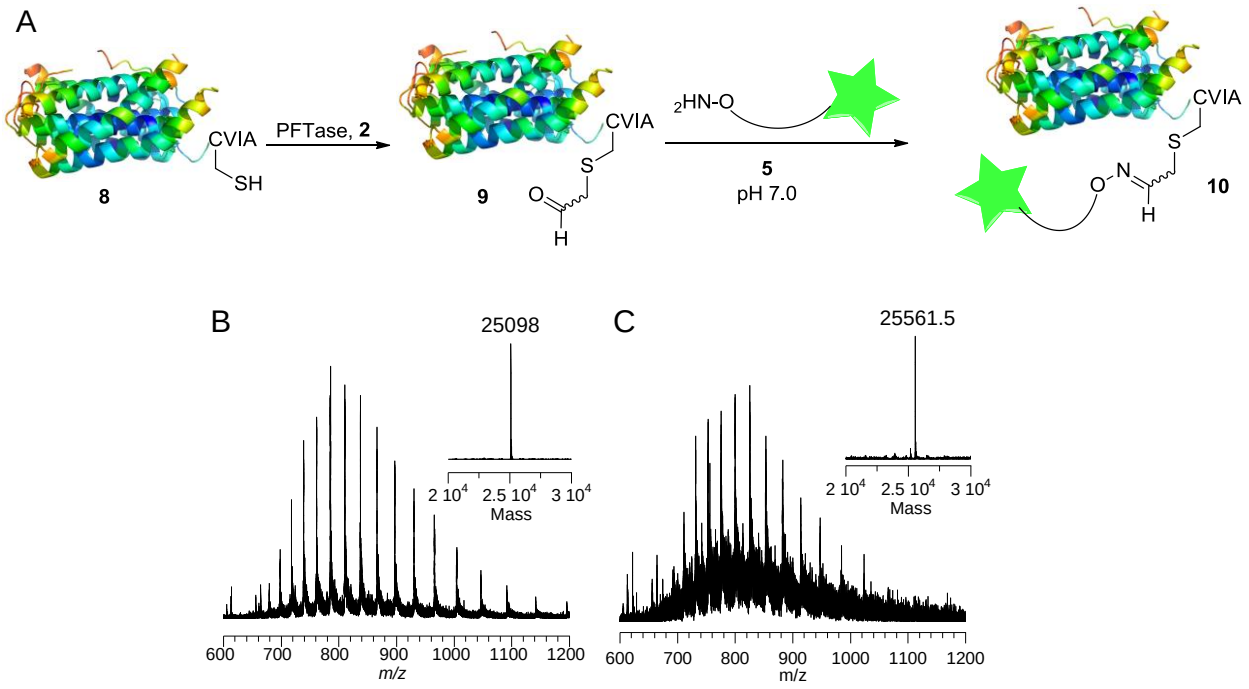


Figure S9. A) Schematic representation of prenylation following by oxime ligation of **8** to yield **9** and **10** respectively. B and C) ESI MS spectra of prenylated-CNTF **9** (spectrum B) and oxime product **10** (spectrum C) with the deconvoluted mass spectra shown in the insets.

Preparation of DHFR^2 M174pAcF (11). The unnatural amino acid *p*-Acetyl phenylalanine was synthesized and characterized as described previously.[6] The BL21 (DE3) competent cells were purchased from Invitrogen™. Experimental details of the plasmid encoding two cysteine free DHFR (DHFR^2) fusion proteins connected with 13 amino acid linker has been described previously.[7] The site for the unnatural amino acid incorporation (M174TAG) was chosen on the basis of the surface accessibility of the residue. The site was mutated to an amber (TAG) stop codon with a Quick change® site directed mutagenesis kit (Strategene). The primers used for the mutation were RS_M174TAG F and RS_M174TAG R as shown below. To facilitate the isolation of the fully expressed DHFR^2 M174 pAcF protein, a C-terminus 6xHis sequence was appended. The primers used for the His_6 insertion were RS_M174TAG His_6 F & RS_M174TAG His_6 R as shown below. The plasmid pEVOL_pAcF encoding amino acyl-tRNA synthetase

(MjTyrRS) and $tRNA_{CUA}$ evolved from *M. jannaschi* was provided by Dr. Schultz group and has been described previously.[8]

Primer	Seuquence
RS_M174TAG F	GGT GGT TAG GTT CCG CGT GGT
RS_M174TAG R	ACC ACG CGG AAC CTA ACC ACC
RS_M174TAG His_6 F	C GAA ATC CTC GAG CGT CGT TAG CAC CAC CAT CAC CAT CAT **TAA GGA** TCC TAA TTA ATT AAT TCA C
RS_M174TAG His_6 R	G CTA ACG ACG CTC GAG GAT TTC GAA ACT ATA GCT ATG CGA G

To express the $DHFR^2$ fusion protein with *p*-acetyl phenylalanine, the plasmid encoding $DHFR^2$ M174TAG was co-transformed with pEVOL_pAcF into BL21 (DE3) competent *E. coli* cells. 50 ml overnight cultures in LB (Luria-Bertani) media were used to inoculate 500 mL of M9 minimal media containing chloramphenicol (34 mg/ml) and ampicillin (50 mg/ml). Cultures were then grown at 37 °C until the O.D. reached 0.57 after which the protein expression was induced by adding 0.3 mM IPTG, 0.04% arabinose and 1 mM pAcF. The cultures were then transferred to 30 °C and incubated with shaking for an additional 18 h. The cells were then harvested via centrifugation at 7,500 rpm for 10 min. The cells pellet was resuspended into lysis buffer containing 1 mg/ml of lysozyme, 20 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole with gentle shaking for 30 min. The partially lysed cells were then cooled and

sonicated. The lysate was then centrifuged at 16,000 rpm for 45 min and the supernatant was then loaded onto a Ni-NTA agarose column. The column was washed with buffer A (20 mM sodium phosphate, 300 mM NaCl, 25 mM imidazole) and eluted with gradient buffer B (20 mM sodium phosphate, 300 mM NaCl, 500 mM imidazole). The protein concentration and purity in the fractions was determined using Bradford assay and SDS PAGE, respectively. After pooling the fractions containing the desired protein, it was concentrated and exchanged into 0.1 M PB pH 7.0 using an Amicon ultrafilitration device equipped with a 30 kDa membrane. The final yield of the protein was 8 mg/L and was stored in - 80 °C. The production of the final $DHFR^2$ M174 pAcF protein was confirmed by ESI-MS (Calculated m/z 38061.5, found 38062.6).

PEGylation of DHFR2 M174pAcF (11) with aminooxy-PEG (12, 3 kDa) using mPDA. Aminooxy-PEG **12** (5 mM) was reacted with $DHFR^2$ M174pAcF **11** (10 μM) in PB (0.1 M, pH 7) in presence of either 100 mM aniline, 500 mM mPDA or no catalyst. The amounts of PEGylated protein in the solutions were analyzed at different times via SDS-PAGE. Gels were visualized by staining with Coomassie blue. Densitometry analysis on the gels was performed using the program ImageJ v1.46.

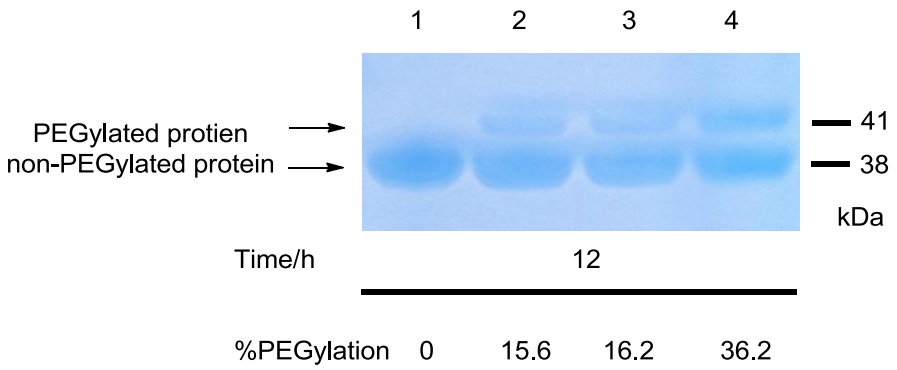


Figure S10. SDS-PAGE analysis of PEGylation of the protein DHFR[2] M174pAcF which contains the unnatural amino acid *p*-Acetyl phenylalanine with aminooxy-PEG **12** (3 kDa). Protein (10 µM) was incubated with aminooxy-PEG (5 mM) with or without catalyst for 12 h. Next the samples were analyzed via SDS-PAGE. Lane 1: pure protein; lane 2: no catalyst was used; lane 3: 100 mM aniline was used as the catalyst; lane 4: 400 mM mPDA was used as the catalyst. The bands were visualized by staining with Coomassie Blue. %PEGylation were obtained by densitometric analysis of the SDS-PAGE using the Coomassie blue stained gel.

Circular dichroism spectroscopy studies for analysis of the effect of mPDA on protein structure and function.

GFP-CVIA **3** (from stock solution of 92 µM diluted to final concentration of 18 µM) was treated with different concentrations of mPDA for ~30 min following by gel filtration chromatography using Zeba spin desalting column according to the manufacturer's instruction (Thermo Scientific), to remove the mPDA catalyst. Circular dichroism spectroscopy of the resulting samples showed no significant differences suggesting that exposure to high concentrations of mPDA does not cause substantial denaturation or irreversible protein unfolding.

Recovered GFP solutions from Zeba spin desalting columns were concentrated down using an Amicon Centriprep centrifugation device (10,000 MW cut-off). Assuming 100% recovery from Zeba spin desalting columns, concentrations were calculated based on the volume of recovered concentrated solutions: pure GFP (21 µM, 440 µL), GFP that has been treated with 100 mM mPDA (21.5 µM, 430 µL); GFP that has been treated with 200 mM mPDA (22 µM, 410 µL);

GFP that has been treated with 300 mM mPDA (17 µM, 540 µL); GFP that has been treated with 400 mM mPDA (18 µM, 500 µL); GFP that has been treated with 500 mM mPDA (17 µM, 540 µL); GFP that has been treated with 600 mM mPDA (17 µM, 540 µL). Spectra were normalized based on 21 µM as the concentration of GFP solutions.

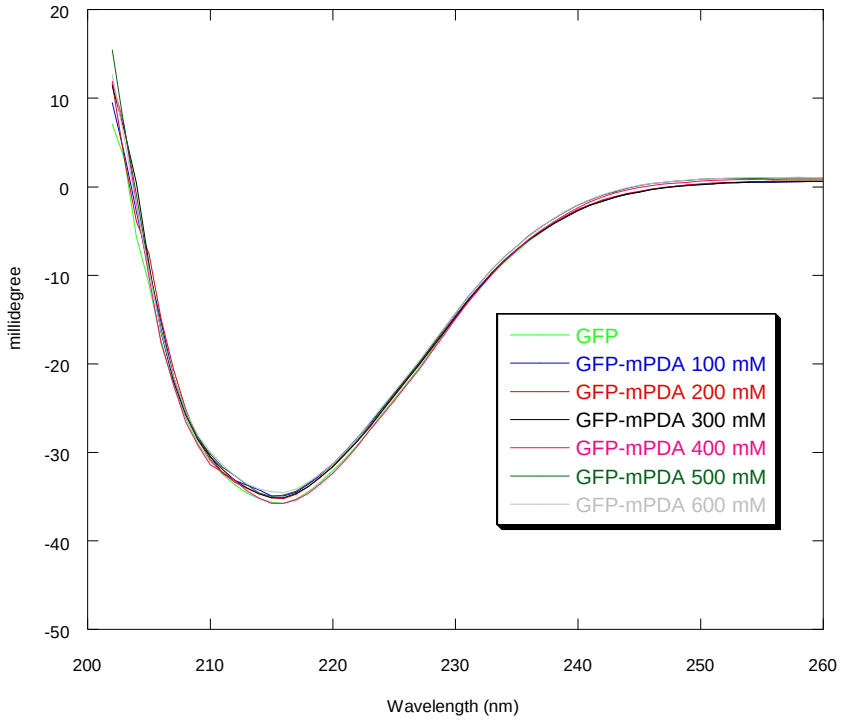


Figure S11. Circular dichroism spectroscopy studies for analysis of the effect of mPDA on protein structure and function. Spectra were normalized to an equal concentration of GFP-CVIA **3** (21 µM). Green line: pure GFP; blue line: GFP that has been treated with 100 mM mPDA for 30 min; red line: GFP that has been treated with 200 mM mPDA for 30 min; black line: GFP that has been treated with 300 mM mPDA for 30 min; pink line: GFP that has been treated with 400 mM mPDA for 30 min; dark green line: GFP that has been treated with 500 mM mPDA for 30 min; gray line: GFP that has been treated with 600 mM mPDA for 30 min.

Effect of catalyst concentration on the enzyme activity.

PFTase, as a model enzyme, was incubated with varying concentrations of either of the two catalysts, aniline or mPDA, for ~15 min, in PB (0.3 M, pH 7.0). Next, 10 μL of each of the PFTase solutions was added to a solution containing Tris•HCl (50 mM , pH 7.5) , $MgCl_2$ (10 mM), $ZnCl_2$ (10 μM), DTT (5.0 mM) and 2.0 μM *N*-dansyl-GCVIA in a final volume of 200 μL. The reaction mixtures were equilibrated at 30 °C for 5 min, initiated by the addition of FPP (10 μM), and monitored for an increase in fluorescence (λ_{ex}=340 nm, λ_{em}=505 nm) for approximately 25 min. The initial rates of formation of products were obtained as slopes in IU/s using least squares analysis. Corrections were applied to all the rate calculations based on the difference between the fluorescence intensity of the prenylated product and the starting peptide. Assuming 100% conversion, the difference corresponds only to the fluorescence of the total amount of the product. The slope was then divided by the fluorescence difference followed by multiplying by the total concentration of peptide (2.0 μM) which then gives the rate of formation of product in μM/s. The rates for the different reactions were normalized to the rate observed in the absence of any catalyst (set to 100%) to facilitate comparison.

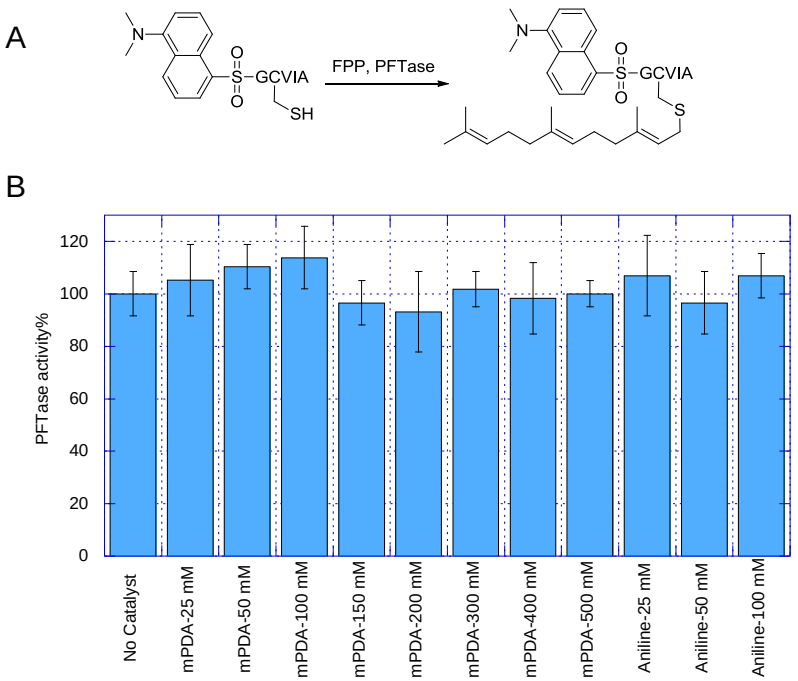


Figure S12. Effect of catalysts on the activity of PFTase. A) Schematic representation of the farnesylation reaction of *N*-dansyl-GCVIA using PFTase. B) Rate analysis of the enzymatic farnesylation reaction using varying concentrations of either aniline or mPDA. The rates for the different reactions were normalized to the rate observed in the absence of any catalyst (set to 100%) to facilitate comparison. Reactions were performed in duplicates.

General procedure for MALDI analysis of protein samples. The sample was adsorbed onto a zip-tip (C_4 column) via repeated cycles of aspiration and ejection (5-10 cycles of 10 µL each) using a pipettor. Next, in order to remove excess buffer and reagents, the zip-tip was washed 5x10 µL with solvent A (H_2O containing 0.1% TFA; v/v) and the proteins eluted with 2 µL of a mixture of A and B (25:75) (solvent B: CH_3CN containing 0.1% TFA; v/v). Next 0.7 µL of the eluted material was added to a MALDI plate and 0.7 µL of matrix was added on top of the sample plate and both were mixed thoroughly to form crystals. Saturated solution of sinapinic acid (3,5-dimeth-oxy-4-hydroxy-cinnamic acid) was used as the matrix.

References.

(1) Gaon, I., Turek, T. C., Weller, V. A., Edelstein, R. L., Singh, S. K., and Distefano, M. D. (1996) Photoactive Analogs of Farnesyl Pyrophosphate Containing Benzoylbenzoate Esters: Synthesis and Application to Photoaffinity Labeling of Yeast Protein Farnesyltransferase. *J. Org. Chem. 61*, 7738–7745.

(2) Shirayev, A., Thoo lin, P. K., and Moiseev, I. K. (1997) Synthesis of Novel Adamantylalkoxyurea Derivatives from 2-(1-Adamantylimino)-1,3-oxathiolane. *Synthesis 1997*, 38–40.

(3) Dirksen, A., Hackeng, T. M., and Dawson, P. E. (2006) Nucleophilic Catalysis of Oxime Ligation. *Angew. Chem. Int. Ed. 45*, 7581–7584.

(4) Wen-jun, W., Chen-ming, C., Chen, J., Xin, W., and George-peng, W. (2011) Kinetic Studies on Aniline-catalyzed Carbohydrate Oxime Formation via Real-time NMR. *Chem. Res. Chinese Universities 27*, 886–890.

(5) Yakhnin, A. V., Vinokurov, L. M., Surin, A. K., and Alakhov, Y. B. (1998) Green Fluorescent Protein Purification by Organic Extraction☆. *Protein Expres. Pur. 14*, 382–386.

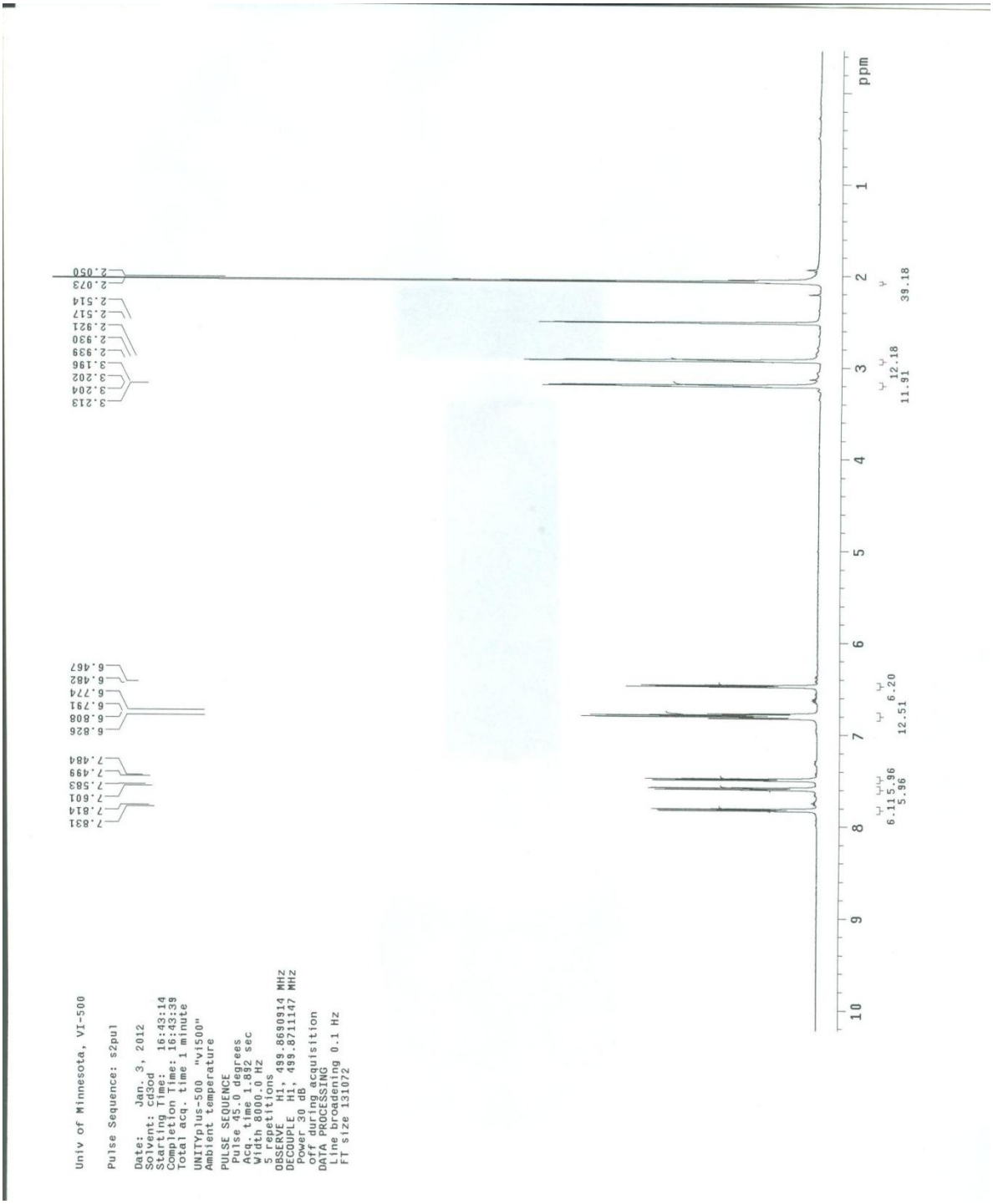
(6) Satyanarayana, M., Vitali, F., Frost, J. R., and Fasan, R. (2012) Diverse organo-peptide macrocycles via a fast and catalyst-free oxime/intein-mediated dual ligation. *Chem. Commun. (Camb.) 48*, 1461–1463.

(7) Li, Q., Hapka, D., Chen, H., Vallera, D. A., and Wagner, C. R. (2008) Self-assembly of antibodies by chemical induction. *Angew. Chem. Int. Ed. 47*, 10179–10182.

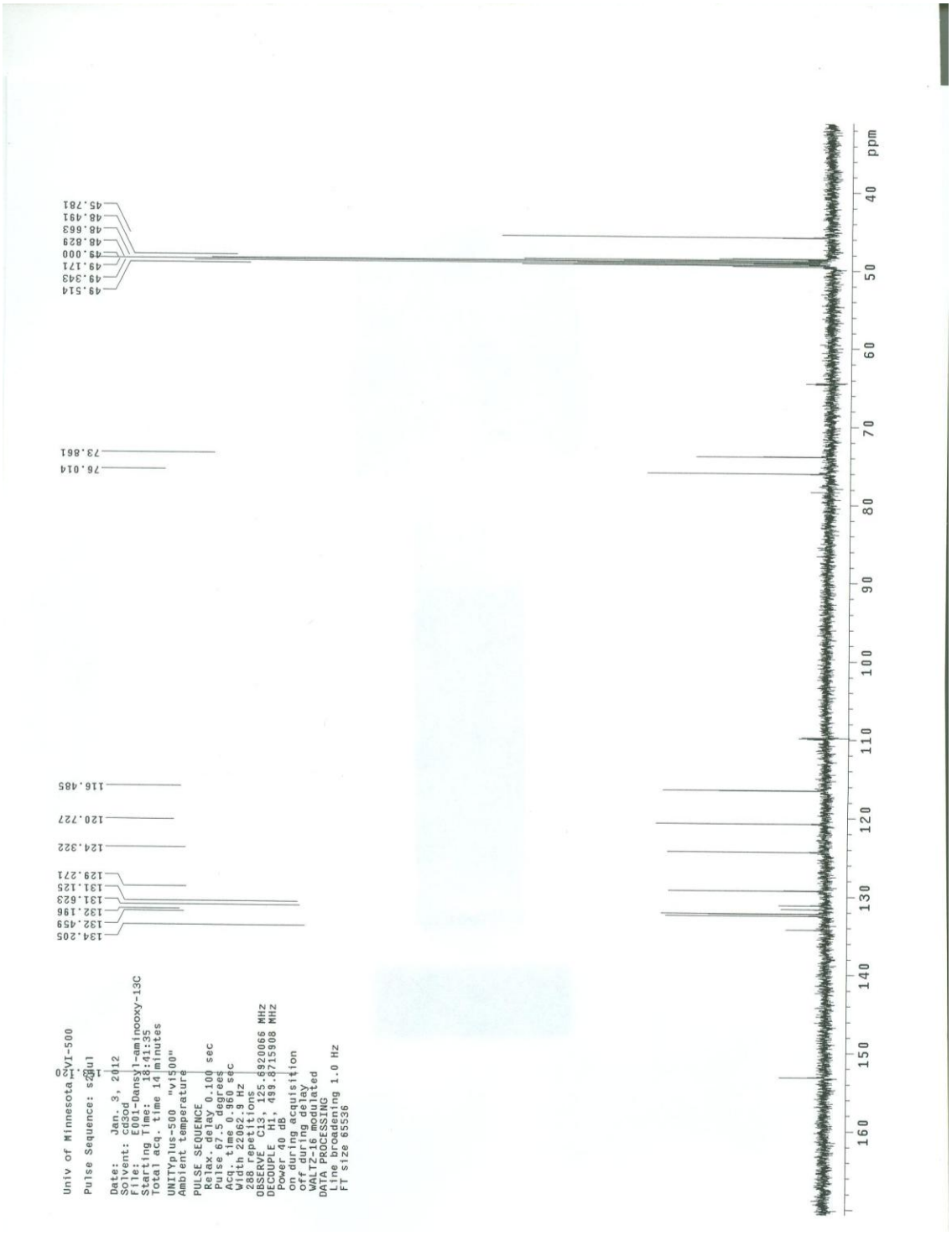
(8) Young, T. S., Ahmad, I., Yin, J. A., and Schultz, P. G. (2010) An enhanced system for unnatural amino acid mutagenesis in E. coli. *J. Mol. Biol. 395*, 361–374.

NMRs.

^{1}H NMR of compound **1**



^{13}C NMR of compound **1**.


Univ of Minnesota-NMR
Pulse Sequence: s2pul
Date: Jan. 3, 2012
Solvent: cd3od
File: cd01-Dansyl-aminooxy-13C
Starting Time: 13:41:35
Total acq. time 14 minutes
UNITYplus-500 "vi500"
Ambient temperature
PULSE SEQUENCE
Relax. delay 0.100 sec
Pulse 67.5 degrees
Acq. time 0.960 sec
Width 22621.8 Hz
288 repetitions
OBSERVE C13, 125.6920066 MHz
DECOUPLE H1, 499.8715908 MHz
Power 40 dB
on during acquisition
off during delay
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
134.205
132.459
132.196
131.623
131.125
129.221
124.322
120.727
116.485
76.014
73.861
49.514
49.343
49.171
49.000
48.829
48.663
48.491
45.781
160
150
140
130
120
110
100
90
80
70
60
50
40
ppm