Supporting Information

Regio- and Chemoselective Covalent Immobilization of Proteins through Unnatural Amino Acids

Cécile Gauchet, Guillermo R. Labadie and C. Dale Poulter*

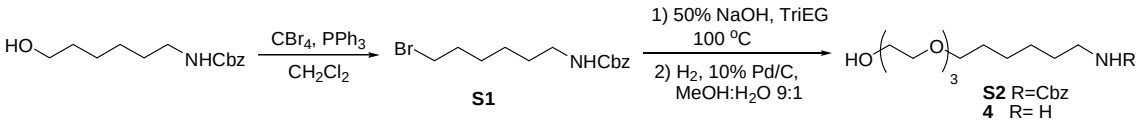
Department of Chemistry, University of Utah, Salt Lake City, Utah 84112

poulter@chem.utah.edu

Table of contents

General. Chemical and solvents were from Aldrich except for O-(2-azidoethyl)-*O'*-(*N*-diglycolyl-2-aminoethyl)heptaethyleneglycol from Novabiochem and NHS-C(O)-PEG-C(O)-NHS from Nektar Therapeutics. Anhydrous THF and CH_2Cl_2 were purified over activated silica. DMF was purchased from Aldrich and kept over dry 3 Å molecular sieves. Moisture-sensitive reactions were performed in a flame-dried flask, under N_2. TLC was performed with silica gel Merck 60 F_{254} plates and visualized under UV light, by $KMnO_4$ or by ninhydrin. Preparative TLC plates (2000 μM) with a UV_{254} Fluorescent indicator, were from Analtech. Amine-derivatized slides and Block It™ were purchased from Telechem International, Inc. (ArrayIt.com). Column chromatography was performed using Silicycle silica gel grade 60, 230-400 mesh. Water used in microarrays was 18 MΩ deionized. NMR spectra were recorded in $CDCl_3$ unless otherwise noted at 300.13 MHz for ^1H, 75.47 MHz for ^{13}C, 121.44 MHz for ^{31}P. Chemical shifts (δ) are in ppm (app = apparent, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, b = broad); coupling constants (*J*) are in Hz. Mass spectra were obtained using electrospray ionization (EI), chemical ionization (CI), or fast atom bombardment (FAB).

Synthesis of PEG-amine 4:



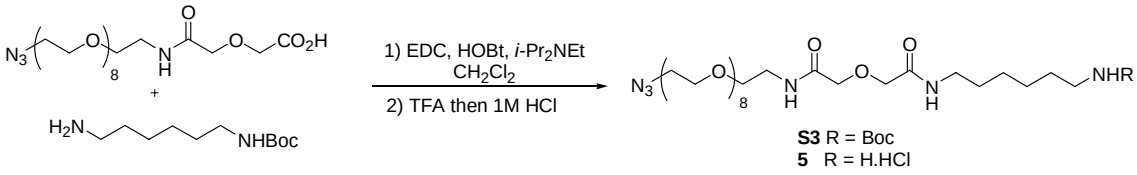
(6-Bromohexyl)carbamic acid benzyl ester (S1). To a solution of (6-hydroxyhexyl)-carbamic acid benzyl ester (1.85 g, 7.38 mmol) in CH_2Cl_2 (73 mL) was added CBr_4 (2.93 g, 8.85 mmol) at 0 °C. Triphenylphosphine (2.70 g, 10.3 mmol) was added in portions over 25 min, and the mixture stirred at rt for 6 h. Water was added and the organic phase was extracted with CH_2Cl_2. The combined organic extracts were washed with brine, dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by batch elution from silica gel (hexanes/EtOAc 1:0 to 4:1) to give 1.52 g (67 %) of a white solid; mp 43-45 ^oC; ^1H NMR δ 7.32-7.30 (m, 5H), 5.07 (s, 2H), 4.98 (bs, 1H), 3.37 (app t, *J* = 6.7 Hz, 2H), 3.17 (dd, *J* = 6.4, 12.9 Hz, 2H), 1.82 (app t, *J* = 6.8 Hz, 2H), 1.51-1.30 (m, 6H); ^{13}C NMR δ 156.5, 136.7, 128.5, 128.3, 128.1, 66.5, 40.9, 33.8, 32.6, 29.8, 27.8, 25.9; LRMS: (EI, m/z, %): 313.0 (M+1) (100); HRMS calcd for $C_{14}H_{20}O_2N^{79}Br$ 313.0677, found 313.0680.

[6-(2-Hydroxy-(tripolyethyleneglycol)hexyl]carbamic acid *tert*-butyl ester (S2). To a solution of triethyleneglycol (6 mL, 47.7 mmol) in THF (6 mL) was added NaOH (50% in H_2O). The solution was stirred at 100 °C for 30 min before **S1** (1.5 g, 4.77 mmol) was added. The mixture was heated at the same temperature for 24 h before equal volumes of H_2O and EtOAc were added. The organic phase was extracted and dried over Na_2SO_4, filtrated and finally concentrated *in vacuo*. The residue was chromatographed on silica gel (hexanes/EtOAc 1:4 to 0:1) to afford 0.814 g (44%) of a pale yellow oil; ^1H NMR δ 7.37-7.32 (m, 5H), 5.09 (s, 2H), 4.83 (bs, 1H), 3.71-3.45 (m, 12H), 3.47 (t, *J* = 6.6 Hz, 2H), 3.19 (dd, *J* = 6.7, 13.3 Hz, 2H), 2.64 (bs, 1H), 1.58 (t, *J* = 6.7 Hz, 2H), 1.50 (t, *J* = 7.0 Hz, 2H), 1.34 (m, 4H); ^{13}C NMR δ 156.4, 136.8, 128.7, 128.3, 128.3, 72.6, 71.5,

70.7, 70.5, 70.2, 66.7, 61.9, 41.1, 30.0, 29.6, 26.6, 25.9; LRMS: (CI, m/z, %): 384.2 (M+1) (100); HRMS calcd for $C_{20}H_{34}O_6N$ 384.23861, found 384.23889.

2-(6-Aminohexyloxy)-(triethyleneglycol)ethanol) (4). S2 (814 mg, 2.12 mmol) was dissolved in MeOH/H_2O 9:1 (50 mL). A catalytic amount of 10 % Pd/C was added, and the reaction was stirred for 21 h under H_2. The mixture was filtered through Celite, the Celite was rinsed with MeOH, and the filtrate was concentrated *in vacuo*. Residual water was removed by lyophilization to give 518 mg (98%) of a pale yellow oil; ^{1}H NMR (CD_3OD) δ 3.66-3.60 (m, 8H), 3.57-3.52 (m, 4H), 3.46 (app t, *J* = 6.4 Hz, 2H), 2.63 (t, *J* = 6.8 Hz, 2H), 1.62-1.27 (m, 8H); ^{13}C NMR (CD_3OD) δ 72.8, 72.4, 71.7, 71.5, 71.2, 62.3, 42.4, 33.4, 30.7, 27.9, 27.2; LRMS: (CI, m/z, %): 250.2 (M+1) (100); HRMS calcd for $C_{12}H_{28}O_4N$ 250.20183, found 250.20199.

Synthesis of linker 5:

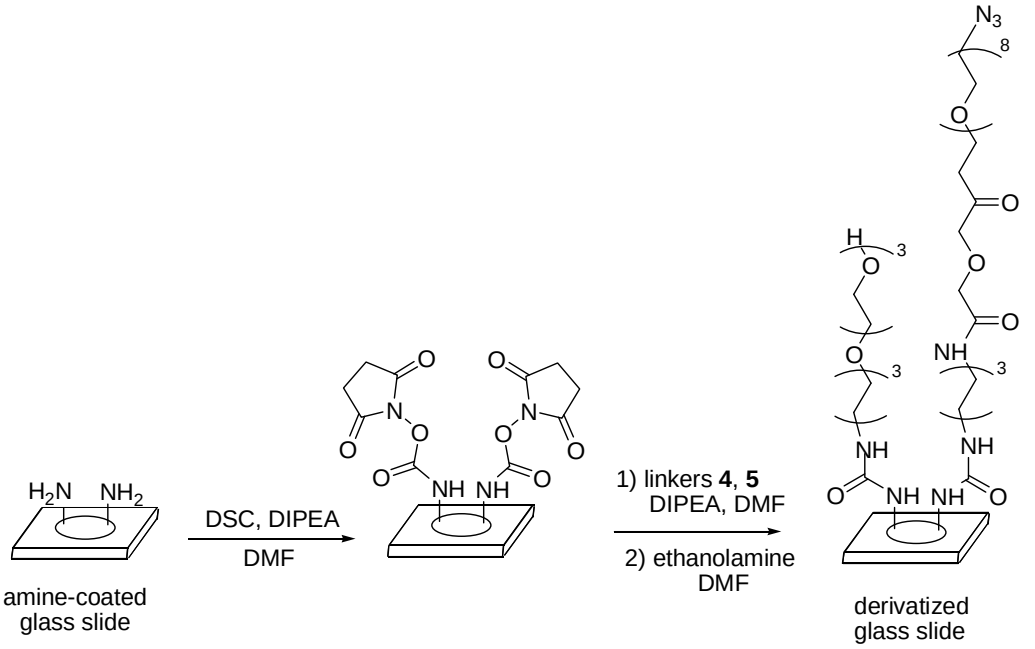


6-(2-{[2-(2-Azidooctaethylenoxy)ethylcarbamoyl]methoxy}acetylamino)hexyl carbamic acid *tert*-butyl ester (S3). To mono BOC-protected hexyl diamine[1] (151 mg, 1.2 eq) dissolved in CH_2Cl_2 (4 mL) at 0 °C was added Hünig base (0.25 mL, 2.5 eq), HOBt (109 mg, 1.4 eq) and O-(2-azidoethyl)-O'-(N-diglycolyl-2-aminoethyl)-heptaethyleneglycol (320 mg, 0.577 mmol) dissolved in CH_2Cl_2 (2 mL). The mixture was stirred for 15 min before EDC (143 mg, 1.3 eq) was added. After stirring for 24 h at rt, CH_2Cl_2 was added (5 mL), and the organic phase was washed in succession with saturated $NaHCO_3$, 1M HCl and brine. The CH_2Cl_2 layer was dried over Na_2SO_4 and

concentrated *in vacuo*. The resulting oil was chromatographed over silica gel (CH_2Cl_2/MeOH gradient from 1:0 to 4:1) to give 289 mg (67 %) of a colorless oil; ^{1}H NMR δ 7.27 (bt, 1H), 7.00 (bt, 1H), 4.63 (bs, 1H), 4.00 (d, *J* = 4.5 Hz, 4H), 3.64-3.58 (m, 28H), 3.54 (m, 2H), 3.53 (dd, *J* = 10.8, 5.6 Hz, 2H), 3.35 (app. t, *J* = 10.7, 5.6 Hz, 2H), 3.05 (m, 2H), 3.25 (dd, *J* = 13.3, 6.6 Hz, 2H), 1.53-1.41 (m, 4H), 1.39 (s, 9H), 1.35-1.25 (m, 4H); ^{13}C NMR δ 168.8, 168.7, 156.2, 79.1, 76.8, 71.1, 70.7, 70.68, 70.65, 70.5, 70.5, 70.2, 70.0, 69.6, 50.7, 40.2, 38.9, 38.7, 30.0, 29.4, 28.5, 26.2, 26.1; LRMS: (FAB, NBA, m/z, %): 753.3 (M+1) (45), 653.3 (100); HRMS calcd for $C_{33}H_{65}O_{13}N_6$ 753.46096, found 753.46118.

2-[(6-Aminohexylcarbamoyl)methoxy]-*N*-[2-(2-azido-ethoxy)ethyl]acetamide (5). S3 (274 mg, 0.364 mmol) was dissolved in CH_2Cl_2 (4 mL) and cooled to 0 ^{o}C before TFA (270 μL, 10 eq) was slowly added. The solution was stirred for 24 h and then concentrated. To the resulting oil was added 1M HCl (10 mL) and the mixture was extracted with EtOAc. The aqueous phase was lyophilized to give 193 mg (82 %) of a colorless oil; ^{1}H NMR (D_2O) δ 4.11 (m, 1H), 3.95 (d, *J* = 4.3 Hz, 3H), 3.59-3.46 (m, 28H), 3.31 (m, 3H), 3.06 (app t, *J* = 6.7 Hz, 2H), 2.81 (app t, *J* = 7.0 Hz, 2H), 1.48 (m, 2H), 1.37 (m, 2H), 1.20 (m, 4H); ^{13}C NMR (D_2O) δ 171.7, 171.3, 69.6, 50.2, 39.5, 38.9, 38.5, 28.2, 26.6, 25.5, 25.3; LRMS: (FAB, NBA, m/z, %): 653.4 (M+1) (100); HRMS calcd for $C_{28}H_{57}O_{11}N_6$ 653.40853, found 653.41040.

Preparation of slides for Click ligation

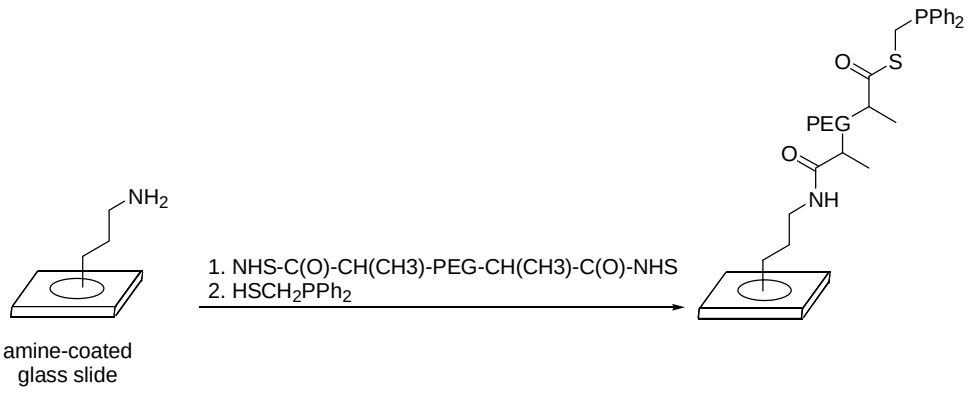


Amine-coated glass slides were treated with a solution of *N,N'*-disuccinimidyl carbonate (DSC, 15 mM) and *N,N*-diisopropylethylamine (DIPEA, 15 mM) in DMF overnight at rt with shaking. The slides were washed five times (5 min each) with DMF and dried under N_2. The succinimidyl derivatized slides were treated with a 1:5 solution of **5** and **4**, (5 mM total concentration) and DIPEA (5 mM) in DMF overnight at rt. The slides were rinsed with EtOAc (5 times - 5 min each), dried under N_2 and then capped by treatment with a solution of ethanolamine (50 mM) in DMF for 3 h at rt. The blocked slides were rinsed with EtOAc (5 times - 5 min each) and dried under N_2.

Click ligation. One microliter of a solution of alkyne-derivatived proteins and Cu^+ complex (1.7 mM final concentration, prepared by mixing equal volumes of *tris*-benzyltriazolylmethylamine (TBTA),[2] $CuSO_4$ and TCPE (20mM each)), in 1:3 glycerol/water was spotted into wells on azido slides masked with a silicon membrane.

The slides were kept in a humid chamber at 4 °C for 3 h before Block It™ solution (4 μL) was added to each well and then were returned to the chamber for 5 h. The slides were washed with PBST (5 min each), appropriate fluorescently-tagged antibodies (4 μg/mL) were added to the wells (4 μL), and the plates were incubated at 4 °C. The slides were then washed with PBST (5 times - 5 min each) and scanned using a Typhoon 8600 Variable Mode Imager.

Staudinger ligation. Diphenylphosphinothiol was attached to amine-derivatized slides according to the procedure of Raines and coworkers[3] to give a surface-bound phosphinothioester. Azido-derivatized proteins in DMF/H_2O (50:1) were added to



phosphinothioester slides by spotting 1 μL with into the appropriate well. The Staudinger ligation was allowed to proceed for 1.5 h in an enclosed chamber saturated with DMF. Block It™ solution (4 μL) was added to each well and the slides were allowed to stand for 2.5 h. The slides were washed once with DMF and four times with PBST (5 min each). Four microliters of a solution of the appropriate antibodies (4 μg/mL) were added to the wells and the plates were incubated at 4 °C. The slide was then washed with PBST (5 times - 5 min each) and scanned with a Typhoon 8600 Variable Mode Imager.

Antibody Stripping. A slide containing immobilized proteins was immersed in a stripping solution (125 mM glycine, 500 mM NaCl, 2.5% Tween 20®, pH 2) at 80 °C for 2 h. The slide was then washed twice with 25 mM phosphate buffer, pH 7.0, and scanned to verify that the fluorescent antibody had been removed. The "stripped" slide was treated with anti-GFP for 16 h at 4 °C, followed by 5 washes with PBST, and then scanned. No cross-reactivity with GST was observed. The same procedure was repeated with anti-GST, and again no cross-reactivity was seen.

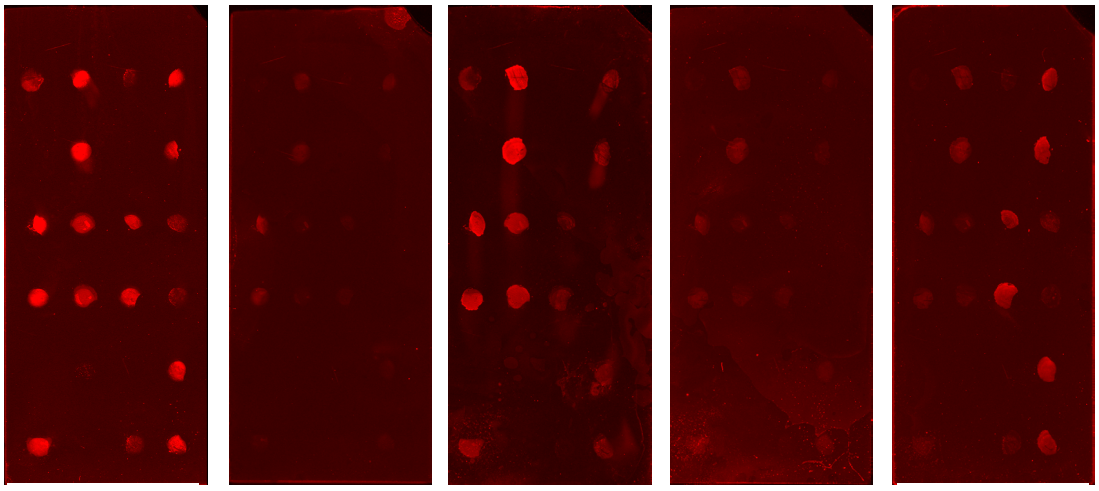


Figure S1. Phosphorimager scans of immobilized GFP-C_2 and GST-C_2 (Rows 1-4 and Columns a-f). 1a and 1f, 15 μM GFP-F; 1b and 1e, no protein; 1c and 1d, 10 μM GFP-C_2; 2a and 2b, 20 μM GFP-C_2; 2c and 2d, 5 μM GFP-C_2; 2e and 2f, 15 μM GFP-F; 3a and 3f, 15 μM GST-F; 3b and 3e, no protein; 3c and 3d, 10 μM GST-C_2; 4a and 4b, 20 μM GST-C_2; 4c and 4d, 15 μM GST-F; 4e and 4f, 5 μM GST-C_2.

Fluorescence of immobilized GFP. **GFP-C_2** (25-100 μM) was immobilized by the "Click" procedure described above. The slide was then immersed in BlockIt solution and washed with PBST, and the fluorescence of GFP was measured by phosphorimaging. The slide was then stored in phosphate buffer at 4 ^{o}C and the fluorescence was measured at 24 and 48 hours.

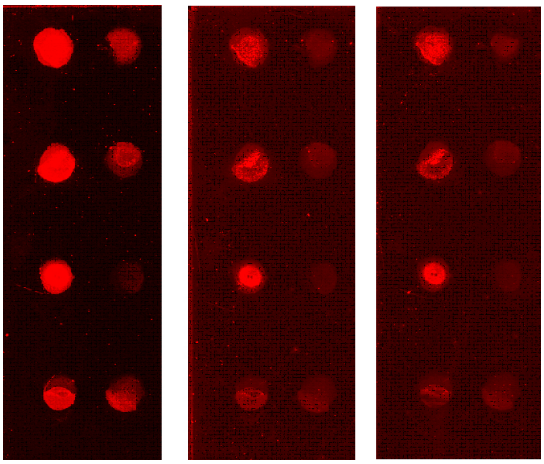


Figure 2S. Phosphorimager scans of immobilized GFP-C_2 at t = 0 (left), t = 24 h (center), and t = 48 h (right). Columns: 1 (left), GFP-C_2; 2, GFP-F. Rows: a (top), 100 μM; b, 75 μM; c, 50 μM; d, 25 μM.

Construction of pQE-GFP-CVIA. The gene for the GFPuv protein was subcloned into a pQE-30Xa expression vector in order to append a His_6 sequence to the *N*-terminus and the *C*-terminal RTRCVIA PFTase recognition site. A 714 bp fragment from pGFPuv containing GFPuv was amplified by PCR. The forward primer contained 28 bases, 17 of which were complementary to the vector sequence upstream of the 5' end, FP-GFP: 5'-CCGGTAGCATGC**ATG**AGTAAAGGAGAAG-3'. The GFP start codon is indicated in bold, and a *SphI* site is underlined. The reverse primer contained 57 bases, 21 of which were complementary to the of the 3' end of RP-GFP: 5'-

GACGATAAGCTTTTA**AGCAATAACGCACCTAGTTCG**TTTGTAGAGCTCATCC ATGCC-3'. The primer contained a *HindIII* site (underlined), and a *C*-terminal yeast RTRCVIA farnesylation motif (**bold**). The PCR product was purified on a 1% agarose gel and extracted from the gel using the GFX™ PCR DNA Gel band purification kit (Amersham Biosciences). The purified DNA and pQE-30Xa (Qiagen) were separately digested with *SphI* and *HindIII* and purified by GFX™ DNA purification kit (Amersham Biosciences) to obtain the 740 bp GFP-CVIA fragment and the 3477 bp linear pQE-30Xa vector, respectively. The two fragments were ligated using Quick T4 DNA ligase (New England Biolabs) and transformed into *Epicurian coli* XL1-Blue electrocompetent cells by electroporation. Cells were selected for growth on ampicillin/tetracyclin. Restriction analysis and sequencing confirmed the desired construct. Sequence analysis of pQE-GFP-CVIA indicated that the GFP-CVIA module was free of mutations. Ligation of the GFP-CVIA module into pQE-30Xa gave pQE-GFP-CVIA (**Figure S3**) with GFP expression under control of the T5 promoter. The recombinant protein contained an *N*-terminal His_6 affinity tag and a *C*-terminal RTRCVIA PFTase recognition site.

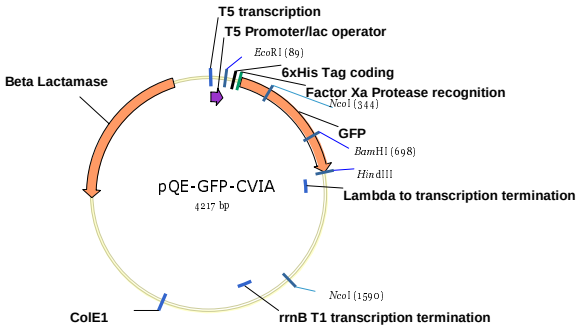


Figure S3. Expression plasmid for GFP-CVIA

Expression and purification from pQE-GFP-CVIA. *pQE-GFP-CVIA* was transformed into chemically competent *Epicurian coli* M15(PREP4) cells (Qiagen). Starter cultures (10 mL of LB, 100 μg/mL ampicillin) were grown overnight at 37 °C, with shaking. Three cultures, each containing 800 mL of LB (100 μg/mL ampicillin), were inoculated with 8 mL of the overnight culture and were grown at 37 °C, with shaking at 240 rpm, until OD_{600} = 0.6 when IPTG was added to 1 mM (final concentration), incubation was continued for 5 h, and cells were harvested by centrifugation and stored at -80 °C. Cell paste (9 g) was resuspended in 36 mL sonication buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 20 mM imidazole, 1 mM β-mercaptoethanol (BME), and 1 mM phenylmethanesulfonyl fluoride -PMSF). Cells were lysed by sonication, three periods of 30 s each. The sample was centrifuged for 30 min at 18000 rpm in a Beckman J20 rotor. The supernatant was mixed with 5 mL of Ni-NTA agarose resin (Qiagen) and shaken for 1 h at 4 °C. GFP-CVIA was eluted from the resin with 250 mM imidazole. Fractions were screened by fluorescence and by SDS-PAGE. Pooled fractions were dialyzed against four changes of 50 mM Tris, pH = 7.0, and 10 mM BME. Dialyzed enzyme was concentrated in a Centriprep centrifugal filter (10,000 cutoff, Millipore) and stored at -80 °C in 3:7 glycerol/water. A 1 L fermentation gave ~4.5 g of wet cell paste and ~10 mg of GFP-CVIA that was > 95% pure by SDS-PAGE.

Construction of pQE-GST-CVIA. The gene for the GST protein was subcloned into a pQE-30Xa in order to append an *N*-terminal His_6 affinity tag and a *C*-terminal RTRCVIA PFTase recognition site to the enzyme. A 660 bp fragment from pET-42b containing GST

was amplified by PCR. The forward primer contained 33 bases, 21 of which were complementary to the vector sequence upstream of the 5'GST. FP_GST: 5'-ATACATAAGCTT**ATG**TCCCCTATACTAGGTTAT-3'. The GST start codon is indicated in bold, and a *BamHI* site is underlined. The reverse primer contained 51 bases, 15 of which were complementary to the 3'end of GST RP-GST**:** 5'-TGAACCAAGCTT**TTAGGCTATAACACAGCGCGTACG**ATCCGATTTTGGAGG -3'. A *HindIII* site (underlined) and the *C*-terminal yeast farnesylation site (**bold** - enconding RTRCVIA) were introduced. The PCR product was purified on a 1% agarose gel and extracted using the GFX™ DNA purification kit. The purified DNA and pQE-30Xa (Qiagen) were separately digested with *BamHI* and *HindIII* and purified using a GFX™ DNA purification kit to obtain the 690 bp GST-CVIA fragment and the 3362 bp linear pQE-30Xa vector, respectively. The two fragments were ligated using Quick T4 DNA ligase (New England Biolabs) and transformed into XL1-Blue electrocompetent cells by electroporation. Cells were selected for growth on ampicillin/tetracyclin. Restriction analysis and sequencing confirmed the desired construct. Sequence analysis of pQE-GST-CVIA indicated that the GST-CVIA module was free of mutations. Ligation of the GST-CVIA module into pQE-30Xa yielded pQE-GST-CVIA (**Figure S4**) with GST under control of the T5 promoter.

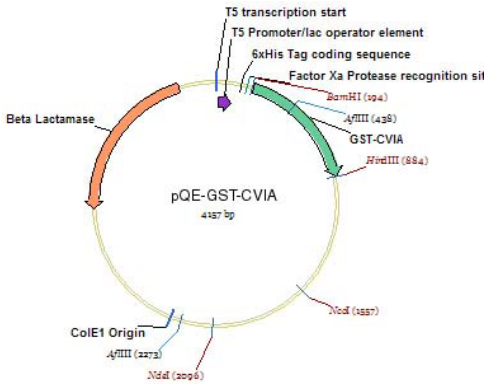


Figure S4. Expression plasmid for GST-CVIA

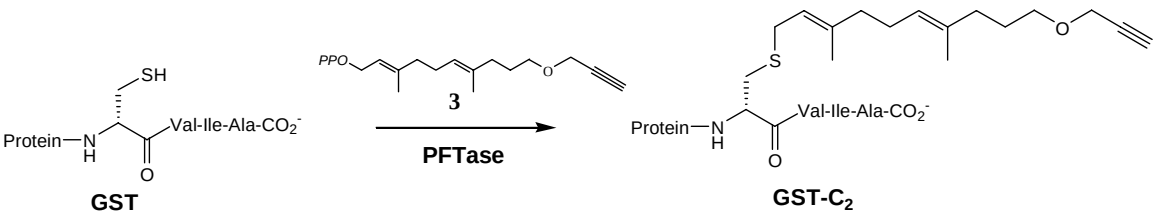
Expression and purification from pQE-GST-CVIA. pQE-GST-CVIA was transformed into chemically competent *Epicurian coli* M15(PREP4) cells (Qiagen). Starter cultures (10 mL of LB, 100 μg/mL ampicillin) were grown overnight at 37 °C, with shaking. Three cultures, each containing 800 mL of LB (100 μg/mL ampicillin), were inoculated with 8 mL of the overnight culture and were grown at 37 °C, with shaking at 240 rpm, until OD_{600} = 0.6 when IPTG was added to 1 mM (final concentration), incubation was continued for 5 h, and cells were harvested by centrifugation and stored at -80 °C. Cell paste (10 g) was resuspended in 50 mL of sonication buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 20 mM imidazole, 1 mM β-mercaptoethanol (BME), and 1 mM phenylmethanesulfonyl fluoride (PMSF). Cells were lysed by sonication, three periods of 30 s each. The sample was centrifuged for 30 min at 18000 rpm in a Beckman J20 rotor. The supernatant was mixed with 5 mL of Ni-NTA agarose resin (Qiagen) and shaken for 1 h at 4 °C. The resin was loaded onto a column. The flow trough was collected and the column was eluted with washing buffer

(sonication buffer minus PMSF) until the absorbance at 280 nm decreased to baseline. GST-CVIA was eluted with washing buffer containing 250 mM imidazole. Fractions were screened for GST activity using the GST.Tag Assay Kit (Novagen) and by SDS-PAGE. Pooled fractions were dialyzed four times against 15 mM Tris, pH 7.0, and 10 mM BME. Dialyzed enzyme was concentrated in a Centriprep centrifugal filter 10,000 cutoff (Millipore) and glycerol was added to 30%. The sample was flash frozen in liquid nitrogen and stored at -80 °C. One liter of cell culture produced ~5 g of wet cell paste and ~12 mg of GST-CVIA that was at least 95% pure by SDS-PAGE.

Expression and purification of yeast Protein Farnesyl Transferase. Yeast PFTase was obtained from pET42-RAM2/RAM1 according at our previously reported procedure.[4]

***In vitro* farnesylation of GFP-CVIA and GST-CVIA.** Purified GFP-CVIA or GST-CVIA (20 μl; 350 μM) and 1.8 mM of farnesyl diphosphate or the appropriate analog (4 μL) were added to 156 μL of 25 mM phosphate buffer, pH 7.0, containing 10 mM $MgCl_2$, and 10 μM $ZnSO_4$. The sample was incubated at 30 °C for 10 min before addition 10 μL of yeast PFTase (250 nM). After 1 hour at 30 °C an additional 20 μL of GFP-CVIA and 10 μL of PFTase were added and the incubation was continued for 1 h. The samples were concentrated using Centricon YM-10 (Millipore), and the concentration of total protein was determined by Bradford analysis. The samples were stored at -20 °C.

Mass spectrometric analysis. Mass spectra for GST and **GST-C_2** were obtained on a Micromass Quattro-II (Micromass, Inc.) triple quadrupole mass spectrometer equipped with a pneumatically assisted electrostatic ion source operating at atmospheric pressure and in a positive ion mode. Samples were purified using a Ziptip C18 device (Millipore). Analyses were made by loop injection into a stream of 60% aqueous acetonitrile flowing at 3 μL/min. Spectra were acquired in the multi-channel analyzer mode from *m*/*z* 600–1400 (scan time 5 s). Mass spectra of multiply-charged molecular ions obtained by ESI were processed into molecular mass spectra (i.e. neutral MW) using MaxEnt software (Micromass, Inc.). The spectra are shown in **Figure S5**.



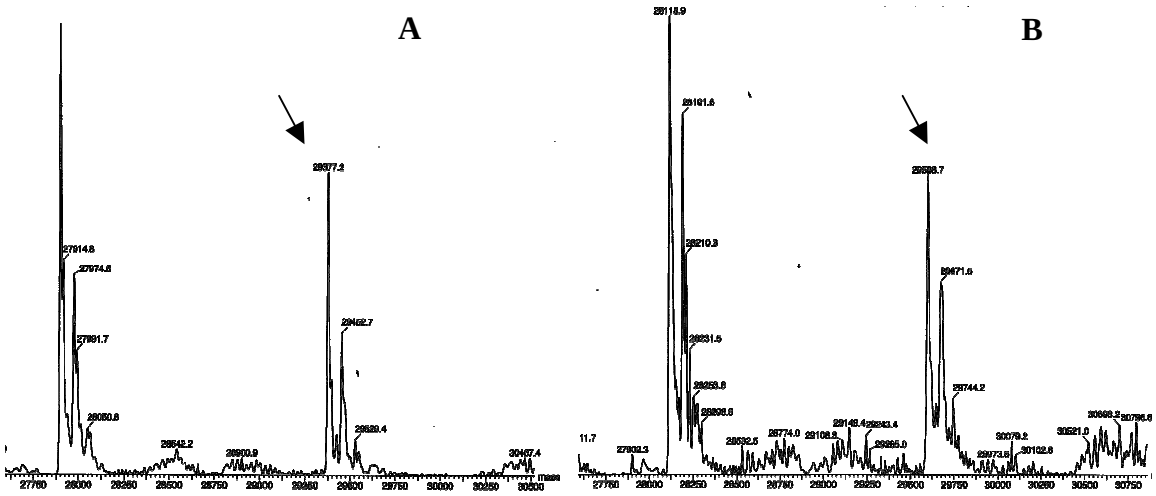


Figure S5 (A) Mass spectrum of purified GST. The peak marked with an arrow corresponds to the expected molecular ion at 29,377 Da; (B) Mass spectrum of **GST-C_2**. The peak marked with an arrow at 28,599 Da (observed) corresponds to the addition of 218 Da (calculated, 28,597).

References

(1) Prepared as described by Callahan, J.F.; Ashton-Shue, D.; Bryan, H.G.; Bryan, W. M.D.; Heckman, G.; Kinter, L.B.; McDonald, J.E.; Moore, M.L.; Schmidt, D.B. *J. Med. Chem.* **1989**, *32*, 391-396.

(2) Synthesized as reported by Wang, Q.; Chan, T.R.; Hilgraf, R.; Fokin, V.V.; Sharpless, K.B.; Finn, M.G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193; Chan, T.R.; Hilgraf, R.; Sharpless, K.B.; Fokin, V.V. *Org. Lett.* **2004**, 6, 2853-2855.

(3) Soellner, M.B. Dickson, K.A.; Nilsson, B.L.; Raines, R.T. *J. Am. Chem. Soc.* **2003**, *125*, 11790-11791.

(4) Harris, C.H.; Derdowski, A.M.; Poulter, C.D. *Biochemistry* **2002,** *41,* 10554-10562.