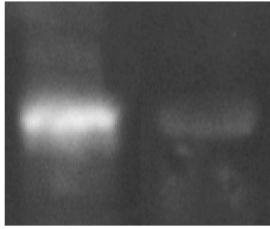
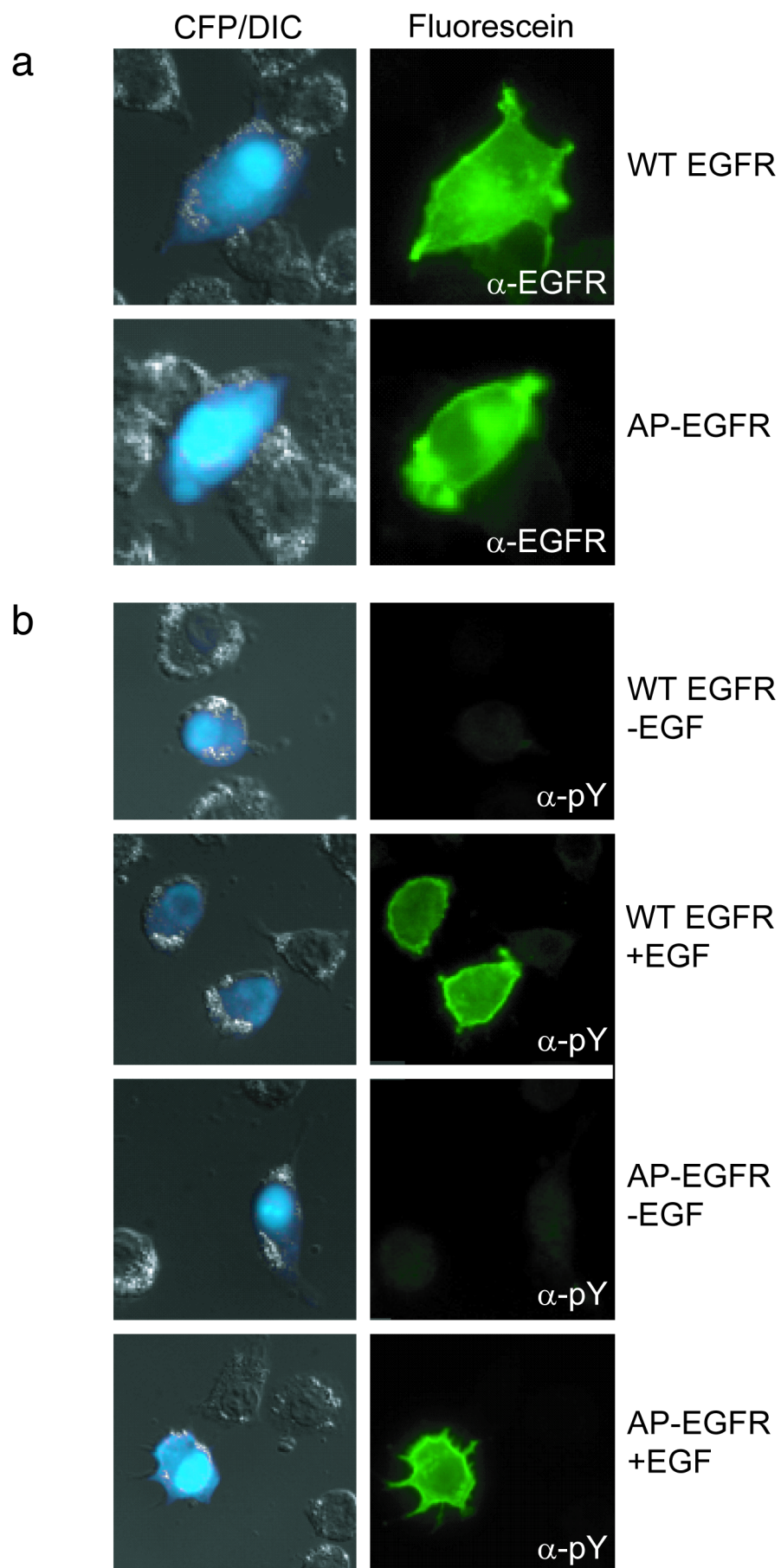


*ketone + FH*

*$\alpha$ -His5 antibody*



**Figure 1.** Comparison of our labeling method (ketone 1 followed by FH) to antibody detection (anti-pentahistidine mouse monoclonal followed by fluorescein-conjugated secondary antibody) for visualization of CFP-AP in lysate. Shown is a fluorescence image of denatured protein sample transferred to nitrocellulose membrane. Our method appears more sensitive than antibody detection.



**Figure 2.** Analysis of the effect of the AP tag on EGFR distribution and function. **(a)** HeLa cells were transfected with either wild-type EGFR or AP-EGFR (cytoplasmic CFP was co-introduced as a transfection marker), fixed, and then stained with anti-EGFR antibody. The localization patterns of EGFR and AP-EGFR are indistinguishable. Note that because CFP-expressing cells consistently display higher levels of EGFR staining under these conditions, we conclude that our protocol detects mostly EGFR expressed from introduced plasmids rather than endogenous EGFR. **(b)** Immunofluorescence staining of fixed cells with anti-phosphotyrosine antibody before and after EGF stimulation. EGFR and AP-EGFR-expressing cells give the same response to EGF treatment.

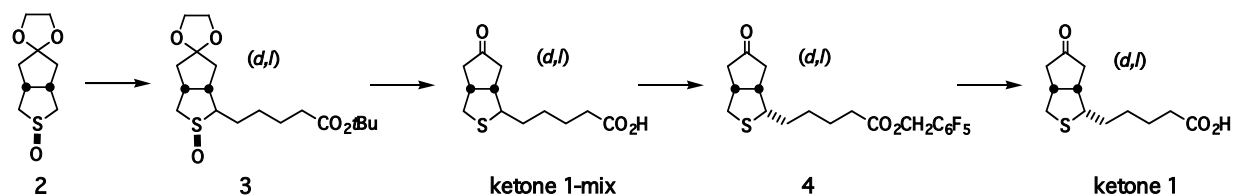
## Supplementary Methods

### Site-Specific Labeling of Cell Surface Proteins with Biophysical Probes using Biotin Ligase

Irwin Chen, Mark Howarth, Weiyang Lin, and Alice Y. Ting\*

Department of Chemistry, Massachusetts Institute of Technology,  
Cambridge, MA 02139, USA.

#### Synthesis of racemic ketone 1.



**General Methods.** All chemicals were purchased from Sigma-Aldrich or Alfa Aesar and used without further purification. Anhydrous tetrahydrofuran (THF) was distilled from sodium benzophenone ketyl and transferred with oven-dried syringes and cannulae. Analytical thin-layer chromatography (TLC) was performed using 0.25 mm silica gel 60 F<sub>254</sub> plates and visualized with *p*-anisaldehyde. Flash chromatography was carried out using silica gel (ICN SiliTech 32-63D). Solvents for chromatography are described as percent by volume. Infrared (IR) spectra were recorded on a Perkin-Elmer Model 2000 FT-IR spectrometer. Proton nuclear magnetic resonance (<sup>1</sup>H NMR) spectra were recorded using a Varian Unity 300 (300 MHz), Varian Mercury 300 (300 MHz), or Bruker Avance 400 (400 MHz) spectrometer. Chemical shifts are reported in delta (δ) units, parts per million (ppm) referenced to the deuteriochloroform singlet at 7.27 ppm. Coupling constants (*J*) are reported in Hertz (Hz). The following abbreviations for multiplicities are used: s, singlet; bs, broad singlet; t, triplet; dt, doublet of triplets; m, multiplet. Carbon nuclear magnetic resonance (<sup>13</sup>C NMR) spectra were recorded with broadband decoupling using a Varian Mercury 300 (75 MHz) spectrometer. Chemical shifts are reported in delta (δ) units, parts per million referenced to the center line of the deuteriochloroform triplet at 77.0 ppm. High resolution mass spectra (HRMS) were obtained on a Bruker Daltonics APEXII 3 Tesla Fourier Transform Mass Spectrometer using electrospray ionization.

**Synthesis of intermediate 3.** Under an atmosphere of dry nitrogen, a solution of compound 2<sup>1</sup> (590 mg, 2.92 mmol) in 6.8 mL THF and 2.8 mL hexamethylphosphoramide (HMPA) was cooled to -78 °C. Methylolithium (2.2 mL of a 1.6 M solution in diethyl ether, 3.5 mmol) was added dropwise. The resulting yellow solution was stirred at -78 °C for 20 minutes before dropwise addition of neat *t*-butyl iodovalerate (3.89 g, 13.7 mmol). The reaction was stirred at -30 °C for 5 hours before quenching with water. The product was extracted with dichloromethane and the organic layer was dried over sodium sulfate and concentrated *in vacuo*. Purification on silica (0-6% methanol/ethyl acetate) afforded the desired product 3 (800 mg, 76% crude yield) as a diastereomeric mixture.

**Synthesis of ketone 1-mix.** Crude compound 3 (800 mg, 2.23 mmol) and triphenylphosphine (995 mg, 3.79 mmol) were dissolved in 20 mL carbon tetrachloride. The resulting solution was heated at reflux for 2 hours. The reaction mixture was decanted from the precipitated triphenylphosphine oxide and concentrated. After purification on silica (5-10% ethyl

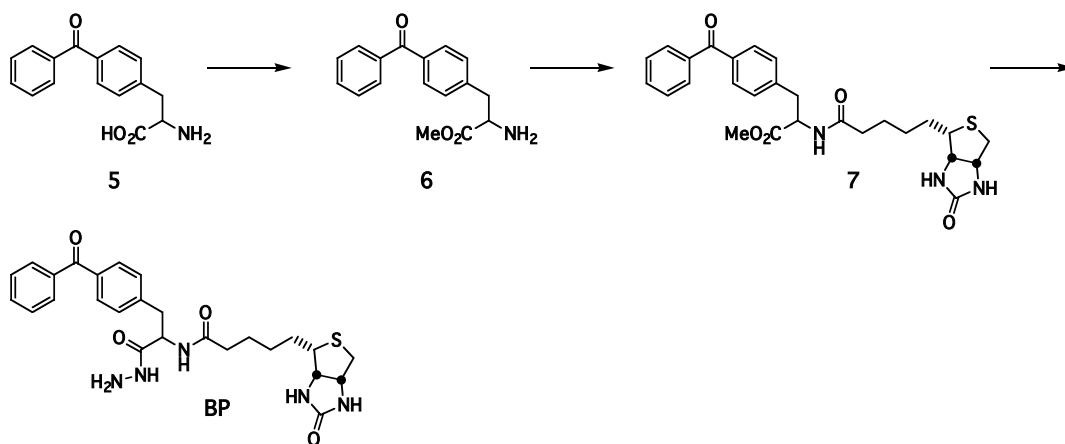
acetate/hexanes), the product was immediately dissolved in 6 mL glacial acetic acid and 3 mL water. Three drops of concentrated hydrochloric acid were added, and the resulting solution was heated at reflux for 16 hours. The reaction mixture was diluted with water, saturated with sodium chloride, and extracted with ethyl acetate. The combined organic layers were dried over magnesium sulfate and concentrated *in vacuo*. Purification on silica (20-50% ethyl acetate/hexanes with 0.5% acetic acid) afforded 181 mg (33% yield) of ketone **1** as a mixture of diastereomers.

**Synthesis of intermediate 4.** Ketone **1** was derivatized to its pentafluorobenzyl ester **4** to facilitate HPLC purification. To a solution of ketone **1**-mix (179 mg, 0.739 mmol) in 7 mL dichloromethane was added diisopropylethylamine (DIPEA) (0.14 mL, 0.80 mmol), followed by pentafluorobenzyl bromide (0.13 mL, 0.86 mmol). The resulting solution was stirred at ambient temperature for 48 hours. The reaction mixture was diluted with water, and the organic layer was separated. The aqueous layer was re-extracted with dichloromethane. The combined organic layers were washed with brine, dried over magnesium sulfate, and concentrated *in vacuo*. After purification on silica (10-20% ethyl acetate/hexanes), the diastereomeric esters were separated on a semi-preparative silica HPLC column (Microsorb-MV 100 Si; 1.5% isopropanol/hexanes; 5.0 mL/min); retention times for the diastereomers were 12.0 (undesired) and 13.1 (desired) minutes. The desired diastereomer **4** was obtained as a white solid (142 mg, 46% yield). <sup>1</sup>H NMR (CDCl<sub>3</sub>, 300 MHz) 5.20 (s, 2H), 3.70 (dt, *J* = 8.2, 5.7, 1H), 2.87-3.16 (m, 3H), 2.17-2.57 (m, 7H), 1.32-1.71 (m, 6H).

**Synthesis of ketone 1.** Compound **4** (108 mg, 0.256 mmol) was dissolved in 0.8 mL THF and 0.8 mL methanol. A solution of lithium hydroxide (32.2 mg, 0.767 mmol) in 0.8 mL water was added dropwise. The resulting yellow solution was stirred at ambient temperature for 12 hours. The reaction was partitioned between ethyl acetate and 1 N HCl that had been saturated with sodium chloride. The layers were separated and the aqueous layer was re-extracted with ethyl acetate. The combined organic layers were dried over magnesium sulfate and concentrated *in vacuo*. Purification of the crude oil on silica (40% ethyl acetate/hexanes with 0.5% acetic acid) afforded ketone **1** as a white solid (45 mg, 72% yield from **4**, 6.4% yield from **2**). IR (neat) 3300-2500 (broad), 2934, 1739, 1707, 1405, 1244, 1169, 949, 747 cm<sup>-1</sup>; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 300 MHz) 10.58 (bs, 1H), 3.71 (dt, *J* = 8.2, 5.7, 1H), 2.88-3.17 (m, 3H), 2.18-2.58 (m, 7H), 1.33-1.72 (m, 6H); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 75 MHz) 217.2, 179.3, 52.1, 48.5, 44.5, 44.5, 37.0, 36.0, 33.8, 32.6, 29.0, 24.5; HRMS calc'd. for (M+Na)<sup>+</sup> C<sub>12</sub>H<sub>18</sub>O<sub>3</sub>Na: 265.0869; found: 265.0875.

**HPLC separation of ketone 1 enantiomers.** The enantiomers of pentafluorobenzyl ester **4** were resolved on a semi-preparative Daicel CHIRALPAK AD-H column (10% isopropanol/hexanes; 3.0 mL/min). The enantiomeric excess (ee) after separation was determined using an analytical Daicel CHIRALPAK AD column (10% isopropanol/hexanes; 1.0 mL/min); retention times of the enantiomers were 15.7 minutes (most likely *d*) and 24.2 minutes (most likely *l*). *d*-**4** was obtained in >99% ee, while *l*-**4** was obtained in 85% ee. Each enantiomer was subsequently hydrolyzed to its acid as described above. The free acids *d*-ketone **1** and *l*-ketone **1** were purified using reverse-phase HPLC (Microsorb-MV 300 C18; 10-43% acetonitrile/water with 0.1% TFA over 20 minutes; flow rate 4.7 mL/min; retention time of product 16.0 minutes).

**Synthesis of benzophenone-biotin hydrazide (BP).**



Synthesis of intermediate **6**. Amino acid **5**<sup>2</sup> (2.3 g, 8.5 mmol) was suspended in 2,2-dimethoxypropane (120 mL) and concentrated hydrochloric acid (10 mL) was added. The mixture was stirred at room temperature overnight. The volatile components were removed under reduced pressure. The residue was purified by silica gel column chromatography (20% methanol/ethyl acetate with 1% DIPEA) to afford a colorless oil (1.8 g, 75%). <sup>1</sup>H NMR (CDCl<sub>3</sub>, 300 MHz) δ 7.33-7.80 (m, 9H), 3.80 (t, 1H), 3.76 (s, 3H), 3.18 (dd, 1H), 2.94 (dd, 1H), 1.58 (s, 2H).

Synthesis of intermediate **7**. To a stirred solution of **6** (0.97 g, 3.5 mmol) in *N,N*-dimethylformamide (50 mL) were added biotin-*N*-hydroxysuccinimidyl ester (1.2 g, 3.5 mmol) and DIPEA (3 mL, 17.5 mmol). After stirring at room temperature overnight, the solvent was removed under vacuum. The residue was purified by silica gel column chromatography (10% methanol/ethyl acetate) to afford a slightly yellow solid (0.86 g, 48%). <sup>1</sup>H NMR (CD<sub>3</sub>OD, 300 MHz) δ 7.40-7.78 (m, 9 H), 4.78 (t, 1H), 4.42 (m, 1H), 4.22 (m, 1H), 3.74 (s, 3H), 3.34 (m, 2H), 3.08 (m, 2H), 2.82 (m, 1H), 2.20 (t, 2H), 1.20-1.62 (m, 6H).

Synthesis of **BP**. Ester **7** (90 mg, 177 μmol) was added to a solution of hydrazine (2 mL) in ethanol (5 mL). After heating at reflux for 10 hours, the volatile components were removed under reduced pressure. The residue was triturated in diethyl ether. The precipitate was removed by filtration, washed with diethyl ether, and dried under vacuum to afford a white solid (70 mg, 78%). IR (KBr): 3275, 1684 cm<sup>-1</sup>; <sup>1</sup>H NMR (CD<sub>3</sub>OD, 400 MHz) δ 7.20-7.78 (m, 9H), 4.56-4.78 (m, 2H), 4.24-4.34 (m, 1H), 3.10 (m, 2H), 2.88 (m, 2H), 2.72 (m, 1H), 2.18 (m, 2H), 1.22-1.60 (m, 6H). HRMS calc'd. for C<sub>26</sub>H<sub>31</sub>N<sub>5</sub>O<sub>4</sub>S: 510.2170; found: 510.2157.

**Bacterial expression and purification of biotin ligase.** pBTac2 plasmid containing the gene for *E. coli* BirA with a C-terminal hexahistidine tag (a gift from D. Beckett) was introduced into *E. coli* strain JM109 by heat-shock transformation. The cells were cultured in Luria Bertani (LB) media supplemented with ampicillin (100 μg/mL) at 37 °C until OD<sub>600</sub> 0.9. Enzyme expression was induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM. After incubation with shaking at 30 °C for three hours, the cells were harvested by centrifugation. The cells were lysed by sonication at 4 °C (three 10-second pulses at half-maximal power with 1 minute between each pulse) in lysis buffer (50 mM Tris pH 7.8, 300 mM NaCl) containing 5 mM phenylmethylsulfonyl fluoride (PMSF) and protease inhibitor cocktail (Calbiochem). The His<sub>6</sub>-tagged enzyme was purified from the lysate using Ni-NTA agarose (Qiagen). Fractions containing enzyme were consolidated and dialyzed against PBS pH 7.4. To remove biotin-AMP ester that co-purified with BirA, the enzyme was incubated with 100 μM synthetic AP at 25 °C for one hour and then re-dialyzed against PBS pH 7.4.

**Construction of the CFP-AP gene.** The CFP gene was PCR-amplified using the primers BamHI-N-CFP.F (5'-GAT AAG GAT CCG AGC TCG ATG GTG AGC AAG GGC GAG GA; incorporates a *BamHI* site) and CFP-C-AP.R (5'-CAT CAT GAA TTC CAT TTA CTC GTG CCA CTC GAT CTT CTG GGC CTC GAA GAT GTC GTT CAG GCC GCC GGA GGA CTC CTT GTA CAG CTC GTC CAT GCC; incorporates the AP sequence GLNDIFEAQKIEWHE and an *EcoRI* site). The insert was digested with *EcoRI* and *BamHI* and ligated in-frame to similarly digested pRSETB vector (Invitrogen), which introduces an N-terminal His<sub>6</sub> tag. For mammalian expression, the same insert was also ligated in-frame to a modified pcDNA3 vector (Invitrogen) that also introduces an N-terminal His<sub>6</sub> tag. CFP-Ala with a lysine to alanine mutation in the AP was constructed using the QuikChange kit (Stratagene) with the primers 5'-CGA CAT CTT CGA GGC CCA GGC GAT CGA GTG GCA CG and 5'-CGT GCC ACT CGA TCG CCT GGG CCT CGA AGA TGT CG.

**Bacterial expression and purification of CFP-AP.** pRSETB plasmid containing the CFP-AP gene was introduced into *E. coli* strain BL21(DE3) by heat-shock transformation. The cells were cultured in LB supplemented with ampicillin (100 µg/mL) at 22-37 °C until OD<sub>600</sub> 0.5. Protein expression was induced by the addition of IPTG to a final concentration of 0.4 mM. After incubation at 30 °C for three hours, the cells were harvested by centrifugation. The cells were lysed by resuspension in bacterial protein extraction reagent (B-PER, Pierce) supplemented with 5 mM PMSF, 1 M pepstatin, and protease inhibitor cocktail (Calbiochem). After incubation at room temperature for 10 minutes, the lysate was cleared by centrifugation, and the supernatant was loaded onto a Ni-NTA agarose column (Qiagen). Fractions containing CFP-AP were collected and consolidated. Biotinylated CFP-AP (a result of endogenous BirA activity in the BL21(DE3) strain) was removed using a streptavidin-agarose column (Novagen). Following debiotinylation, the protein was concentrated by exchanging into pure water by gel filtration and lyophilizing to dryness. The solid protein was redissolved in 50 mM bicine pH 8.3, 5 mM Mg(OAc)<sub>2</sub>. The concentration of protein was determined from its absorbance at 434 nm using the extinction coefficient for enhanced CFP (34,000 M<sup>-1</sup>cm<sup>-1</sup>).

**Antibody detection of CFP-AP in HEK cell lysate.** Lysate was prepared by hypotonic lysis as described in the "Methods" section under "Fluorescent labeling of CFP-AP in mammalian cell lysates." After resolution by SDS-PAGE and transfer to nitrocellulose membrane, the CFP-AP was detected by blotting with anti-pentahistidine mouse antibody (1:1000, Qiagen) followed by fluorescein-conjugated anti-mouse antibody (1:20, Calbiochem). The blot was imaged with a Storm 860 instrument (Amersham Biosciences) to generate **Supplementary Figure 1**.

**Construction of AP-tagged pDisplay-CFP (AP-CFP-TM).** The CFP gene was PCR-amplified using the primers AP-N-CFP.F (5'-TGA AGA TCT GGC GGC GGC CTG AAC GAC ATC TTC GAG GCC CAG AAG ATC GAG TGG CAC GAG GGC GCG CCG ATG GTG AGC AAG GGC GAG GAG; incorporates the AP sequence GLNDIFEAQKIEWHE and a *BglII* site) and CFP-C-SacII.R (5'-TCC CCG CGG CCG CTT GTA TAG CTC GTC CAT GCC GAG AGT; incorporates a *SacII* site). The insert was digested with *BglII* and *SacII* and ligated in-frame to similarly digested pDisplay vector (Invitrogen). The point mutant Ala-CFP-TM was constructed using the QuikChange kit (Stratagene) with the primers 5'-CGA CAT CTT CGA GGC CCA GGC GAT CGA GTG GCA CG and 5'-CGT GCC ACT CGA TCG CCT GGG CCT CGA AGA TGT CG.

**Construction of AP-tagged EGFR.** The human EGFR gene in pcDNA3 (Invitrogen) was a gift from K. D. Wittrup. The acceptor peptide (GLNDIFEAQKIEWHE) was introduced at the *NheI* site between residues A21 and S22 of the EGFR signal sequence using the primers 5'-CTA GTC GGG CTG GCC TGA ACG ATA TCT TCG AGG CCC AGA AGA TCG AGT GGC ACG AGG and 5'-CTA GCC



TCG TGC CAC TCG ATC TTC TGG GCC TCG AAG ATA TCG TTC AGG CCA GCC CGA. These primers were annealed, phosphorylated with T4 polynucleotide kinase (New England Biolabs), and ligated into *NheI*-digested EGFR plasmid. Ala-EGFR with a lysine to alanine mutation in the AP was constructed using the QuikChange kit (Stratagene) with the primers 5'-CGA CAT CTT CGA GGC CCA GGC GAT CGA GTG GCA CG and 5'-CGT GCC ACT CGA TCG CCT GGG CCT CGA AGA TGT CG.

**Determination of labeling sensitivity using AP-CFP-TM.** We used the wedge method to estimate the concentration of CFP in single cells expressing the AP-CFP-TM construct.<sup>3,4</sup> A wedge-shaped microchamber was constructed from three glass coverslips. The length along the x-direction was 7 mm and the height of the chamber (z-direction) increased linearly from 0 to 175  $\mu\text{m}$ . The chamber was filled with a solution of 27  $\mu\text{M}$  purified CFP in PBS pH 7.4. The fluorescence of the wedge was imaged under conditions identical to those used for cellular imaging. We assumed an average cell thickness of 5  $\mu\text{m}$  and therefore interpolated to the region of the wedge with thickness 5  $\mu\text{m}$  and used the fluorescence intensity measured there as a reference standard for CFP concentration measured in single cells. Using the wedge for comparison, we imaged a sample of HeLa cells expressing AP-CFP-TM that had been labeled with the BP probe. Examination of the CFP channel images for cells that displayed clear BP labeling (signal to background >4:1) showed that the CFP concentrations ranged from 10  $\mu\text{M}$  to >1 mM. We therefore concluded that cells expressing as little as 10  $\mu\text{M}$  AP-CFP-TM could be labeled by our method. We also performed deconvolution imaging to estimate the amount of AP-CFP-TM on the cell surface as opposed to internal. We acquired 20  $\mu\text{m}$  image stacks with 0.2  $\mu\text{m}$  step size. The stacks were deconvoluted using Volocity software (Improvision) and analysis showed that approximately half of the total CFP fluorescence was inside the cell, and half was at the plasma membrane. Based on this observation, and the estimate that a typical HeLa cell has a volume of 1 pL, the number of cell surface AP tags is  $\sim 10^6$  for a cell expressing 10  $\mu\text{M}$  of AP-CFP-TM. This represents an upper limit to the labeling sensitivity of our methodology.

**Immunofluorescence detection of AP-EGFR localization and EGF-induced tyrosine phosphorylation.** HeLa cells were co-transfected with a cytoplasmic CFP marker plasmid and either wild-type EGFR or AP-EGFR in pcDNA3, in a 1:8 ratio, using Lipofectamine 2000 (Invitrogen). After 24 hours of expression at 37  $^{\circ}\text{C}$ , the cells were fixed at 4  $^{\circ}\text{C}$  for 5 minutes, then 25  $^{\circ}\text{C}$  for 10 minutes, using 4% paraformaldehyde, 4% sucrose, and 1 mM sodium orthovanadate. After two rinses with PBS, the cells were incubated at 25  $^{\circ}\text{C}$  for 10 minutes with 0.1% Triton X-100, 1 mM sodium orthovanadate, and PBS to permeabilize the membranes. Anti-EGFR antibody (1  $\mu\text{g}/\text{mL}$ , Oncogene GR01L), followed by fluorescein-conjugated goat anti-mouse IgG (5  $\mu\text{g}/\text{mL}$ , Calbiochem 401244) were used to stain the EGF receptor. After mounting, the cells were imaged on the Zeiss Axiovert 200M microscope using a 40x lens and a fluorescein filter set (495DF10 excitation, 515DRLP dichroic, 530DF30 emission). Because we consistently observed higher levels of staining for cells expressing CFP, we concluded that the majority of EGFR staining on transfected cells resulted from introduced EGFR plasmid rather than endogenous EGFR. To detect tyrosine phosphorylation, transfected cells were serum-starved for 7-24 hours, then EGF (100 ng/mL) was added for 10 minutes at room temperature. The cells were fixed and stained with anti-phosphotyrosine antibody (2.5  $\mu\text{g}/\text{mL}$ , BD Transduction Laboratories PY20B). Fluorescein-conjugated goat anti-mouse antibody was used to detect the primary antibody, and the cells were imaged using the same fluorescein filter set as above at 40x magnification.

## References

1. Baraldi, P.G., Barco, A., Benetti, S., Gandolfi, C.A., Pollini, G.P., Polo, E. & Simoni, D. Synthesis of Sulfur-Containing Carbaprostacyclin Analogs. *Gazzetta Chimica Italiana* **114**, 177-183 (1984).
2. Kauer, J.C., Ericksonviitanen, S., Wolfe, H.R. & Degrado, W.F. Para-Benzoyl-L-Phenylalanine, A New Photoreactive Amino-Acid - Photolabeling of Calmodulin with A Synthetic Calmodulin-Binding Peptide. *J. Biol. Chem.* **261**, 695-700 (1986).
3. Adams, S.R., Campbell, R.E., Gross, L.A., Martin, B.R., Walkup, G.K., Yao, Y., Llopis, J. & Tsien, R.Y. New Biarsenical Ligands and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. *J Am. Chem. Soc.* **124**, 6063-6076 (2002).
4. Miyawaki, A., Griesbeck, O., Heim, R. & Tsien, R.Y. Dynamic and quantitative Ca<sup>2+</sup> measurements using improved cameleons. *Proc. Natl. Acad. Sci. USA* **96**, 2135-2140 (1999).