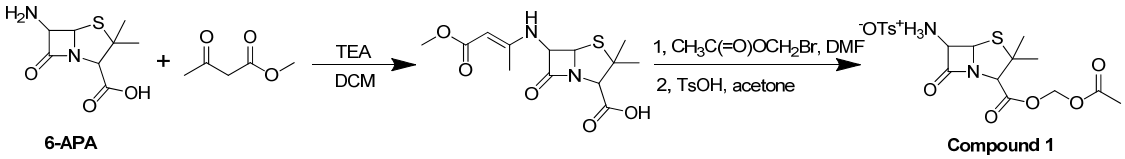
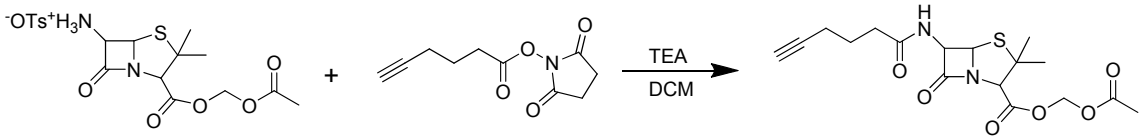
Experiment Procedure and Materials

Synthesis of CP-AM



6-APA Compound 1

To a suspension of 6-APA (1.08 g, 5 mmol) in DCM (25 mL) was added triethylamine (1.38 mL, 10 mmol), and then methyl acetoacetate (0.58 g, 5 mmol) drop-wise in 10 min. The reaction was stirred at room temperature for 3 h. The solvent and excess TEA were evaporated, and the residue was taken up in DMF (12 mL). Bromomethyl acetate (0.49 mL, 5 mmol) was added dropwise while maintaining the temperature below 30 ^{o}C. The mixture was stirred at room temperature for 3 h and partitioned between ethyl ether (50 mL) and brine (50 mL). The organic layer was washed further with brine (2 x 50 mL), dried over anhydrous Na_2SO_4. After the volatile solvent was evaporated, the residue was redissolved in acetone (5 mL) and then *p*-toluenesulfonic acid monohydrate (0.80 g, 5 mmol) was added. The titled **compound 1** was precipitated from ethyl ether (25 mL) as the beige powder (1.80 g, 78%): ^{1}H NMR (MeOD, 600 MHz) δ = 7.73 (d, 2H, *J* = 7.8 Hz), 7.27 (d, 2H, *J* = 7.8 Hz), 5.88 (d, 1H, *J* = 5.4 Hz), 5.80 (d, 1H, *J* = 5.4 Hz), 5.65 (s, 1H), 5.02 (s, 1H), 4.69 (s, 1H), 2.40 (s, 3H), 2.12 (s, 3H), 1.73 (s, 3H), 1.56 (s, 3H). MS (ESI) calc for $C_{11}H_{16}N_2O_5S$ 288.1, found 288.8 $(M+H)^+$.



Compound 1 Compound 2

A solution of **compound 1** (87 mg, 0.2 mmol) in DCM (2 mL) was treated sequentially with 5-hexynoic acid 2, 5-dioxo-1-pyrrolidinyl ester[1] (42 mg, 0.2 mmol) and triethylamine (42 μL, 0.3 mmol). The reaction mixture was stirred at room temperature overnight and partitioned between DCM (20 mL) and water (20 mL). The organic layer was washed with brine (20 mL), dried over anhydrous Na_2SO_4, and concentrated under reduced pressure. Chromatography (SiO_2, 2 x 6 cm, hexane: ethyl acetate = 3:1 then 2:1 and finally 1:1) afforded **compound 2** (CP-AM, 50 mg, 65%) as clear wax: ^{1}H NMR ($CDCl_3$, 600 MHz) δ = 6.13 (d, 1H, *J* = 9.0 Hz), 5.83 (d, 1H, *J* = 5.4 Hz), 5.77 (d, 1H, *J* = 5.4 Hz), 5.73 (dd, 1H, *J* = 4.2, 9 Hz), 5.55 (d, 1H, *J* = 4.2 Hz), 4.45 (s, 1H), 2.41 (m, 2H), 2.29 (m, 2H), 2.13 (s, 3H), 1.99 (s, 1H), 1,88 (m, 2H), 1.67 (s, 3H), 1.52 (s, 3H). ^{13}C NMR (151 MHz) δ = 174.1, 171.9, 166.8, 164.0, 83.5, 79.8, 70.2, 69.8, 67.8, 64.9, 58.7, 34.9, 31.8, 27.4, 24.3, 24.2, 20.9. MS (ESI) calc for $C_{17}H_{23}N_2O_6S$ $(M+H)^+$ 383.12768, found 383.12732.

Gene mutation and plasmid construction

Plasmid for the expression of NLacN-SH3 protein

SH3 domain coding sequence from yeast gene *Sho1p* was codon-optimized for *E. coli.* expression and a serial of five primers OSL082, OSL083, OSL084, OSL085 and OSL086 were synthesized to cover the modified gene. PCR reactions were used to join these primers to obtain whole SH3 domain DNA fragment. NLacN fragment was generated by PCR reaction with primer OSL020 and OSL039 with Tem-1 β-lactamase gene as template. E166N mutation was introduced with primer OSL109 and OSL110. DNA fragment generated from mutation was further extended by two rounds of PCR reaction with primer pair OSL039-OSL111 and OSL039-OSL082 to join NLacN-Linker with SH3 coding DNA fragment to obtain

NLacN-SH3 fusion gene. The fusion gene was digested by SacII and BamHI and cloned into pET11(c)+ derived vector with 6×his tag and Ubiquitin gene (HisUBQ vector) as described before.[2]

Plasmid for the expression of LacC-PPLP protein

Two primers OSL087 and OSL088 were designed to cover Sho1p SH3 domain binding peptide QQIVNKPLPPLPVAGSS from Pbs2p (amino acid 88-104),[3] and codons were also optimized for *E. coli.* expression. β-lactamase C-terminal (amino acid 198-286) DNA fragment was amplified with primer OSL035-OSL021 and extended with primers OSL088 and OSL087 to obtain fusion gene SH3 binding peptide and β-lactamase C-terminal joined by a linker, named LacC-PPLP. This fusion gene was cloned into the same HisUBQ vector mentioned above for protein expression.

Plasmid for mammalian cell expression of Zip-NLacN and Zip-LacC

Two plasmids for mammalian cell expression of Zip-Bla(1) and Zip-Bla(2) in vector pcDNA3.1/zeo were requested and received from a previous report.[4] E166N mutation was introduced into Zip-Bla(1) expression plasmid by the same method as that for NLacN-SH3 mentioned above and primers including OSL100, OSL109, OSL110 and OSL105. Mutated gene was cloned back into the same vector and obtained plasmid Zip-NLacN pcDNA3.1/zeo. In order to increase Zip-NLacN and Zip-LacC co-expression efficiency in mammalian cells, Zip-LacC (the same as Zip-Bla(2) , renamed here) fusion gene with its regulatory elements was transferred to the plasmid Zip-NLacN pcDNA3.1/zeo. Primers OSL136 and OSL137, each of which contain a BglII site, were used to amplify the Zip-LacC fusion gene with its regulatory elements and add BglII sites at both ends. Generated DNA fragment was digested by BglII and inserted into the only BglII site in plasmid Zip-NLacN pcDNA3.1/zeo. In the end, we obtained the plasmid with two sets of regulatory elements to control the expression of both Zip-NLacN and Zip-LacC in one pcDNA3.1/zeo vector. After construction, *E. coli* strain DH5α was used to amplify and produce high amount of plasmids, and QIAGEN Plasmid Midi Kit was used to prepare highly pure plasmids for mammalian cell transfection experiment.

Plasmid for mammalian cell expression of NLS-Zip-NLacN and NLS-Zip-LacC

The sequence of the NLS is DPAALKRARNTEAARRSRARKLQRMK, which is predicted by software predictNLS[5] and cNLS Mapper.[6] Primer OSL166, OSL167 and OSL168 were used to generate the fusion gene with this NLS from GCN4 gene. Specifically, primer OSL168 and OSL105 (BGH reverse primer) were firstly used to amplify the Zip-NLacN and Zip-LacC fusion genes, and OSL166-OSL105 and OSL167-OSL105 were used as second round of PCR to obtain the fusion gene NLS-Zip-NLacN and NLS-Zip-LacC repectively and add restriction site BamHI. Generated DNA fragments were cloned back into the vector pcDNA3.1/zeo. These two plasmids were further combined to make one plasmid that will express both NLS-Zip-NLacN and NLS-Zip-LacC in mammalian cells, and the procedure was exactly the same as mentioned above.

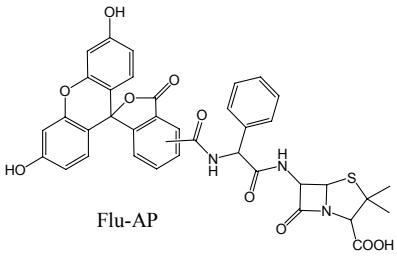
Protein expression in *E. coli.* and purification with Ni-Resin

Proteins were expressed in *E. coli.* Strain BL21 (DE3). Plasmid was transformed into competent cells, and bacterial cells were grown overnight on plates with appropriate antibiotics accordingly. Next day, bacterial cells were inoculated into LB media with antibiotics and cell density was determined by the measurement of OD at 600nm. When OD_{600} reached 0.6, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a finial concentration of 1mM. Cells were allowed to grow for additional 3 h and then collected and stored at -70 ^{o}C for future use.

For protein purification, bacterial pellet with protein expressed was thawed at room temperature, and Bugbuster (Novagen) was added to break down cells according to manufacture's instruction. Cell lysate was centrifuged at 40,000×g and supernatant was kept for purification. Ni-resin (Bio-Rad) was used to purify these proteins according to the instruction for purification of 6×his tagged protein. Briefly, cell lysate was diluted with 3 volumes of binding solution (50mM Tris buffer, pH7.5, 500mM NaCl, 5mM

imidazole) and loaded to Nickel column equilibrated with large amount of binding solution. Column was rotated at 4°C to allow the binding to occur. The column was then washed intensively with wash buffer (50mM Tris buffer, pH7.5, 500mM NaCl, 10 to 20mM imidazole). Next, target protein was eluted with elution buffer (50mM Tris buffer, pH7.5, 500mM NaCl, 250mM imidazole), and fractions were collected and resolved by SDS-PAGE. Protein obtained was further dialyzed against appropriate reaction buffer and concentrated by ultra-filtration tubes (Millipore). Protein concentration was estimated by either SDS-PAGE or Bradford assay (Sigma).

in vitro PCA experiment with purified proteins using Flu-AP as substrate



Flu-AP

For *in vitro* PCA experiment, purified NLacN-SH3 and LacC-PPLP proteins were mixed with equal molar ratio and Flu-AP was added to the reaction to 100μM. Two control experiments with the same concentration of NLacN-SH3 and LacC-PPLP individually were also set up at the same time. All reactions were maintained at 37°C for 3 h. After reaction, all samples were precipitated with cold acetone, dissolved in urea and resolved by SDS-PAGE. The gel imaging system (Alpha Innotech) was used to take a fluorescent picture by using 365 UV lamp and filter for SYBR Green. The same gel was then stained with Coomassie blue and documented again with white light.

Mammalian cell culture and transfection

Human cell line HeLa and HEK293T were grown in complete DMEM, supplemented with 10 mM HEPES, 10% fetal calf serum, 1% L-glutamine, nonessential amino acids (Invitrogen). Cells were grown in a 5% CO_2 incubator at 37°C. One day before the transfection, cells were seeded to a proper density. For live cell imaging, HeLa cells were used and seeded at a low density; for protein enrichment experiment, HEK293T cells were used and seeded to a relatively higher density so that the next day the confluence could reach 90%. At the day one of the experiment, purified plasmid DNA was transfected into mamanlian cells by Effectene reagent (Qiagen). Briefly, DNA was firstly mixed with DNA-condensation buffer EC and Enhancer was added at a mass to volume ratio of 1:8. The sample was incubated at room temperature for 5min. Then Effectene reagent was added to the mixture, followed by vortex for 10 seconds, and the sample was kept at room temperature for 10 min. Cells were prepared during this 10 min and fresh medium was added. After 10 min incubation, transfection complex was first mixed with fresh media and then added to cells in a drop-wise way. After transfection, cells were maintained in 5% CO_2 incubator at 37°C until further use.

Live cell imaging with CCF2-AM substrate

Live cell imaging was carried out with HeLa cells because of their exceptional adhesive property and clear morphological feature. Poly-lysine coated cell culture chamber slides (BD) were used to grow cells. Similar transfection protocol was used as mentioned above but at a smaller scale proportional to the size of the cell culture chamber. 48 h after transfection, cells were washed with 1X HBSS (Hyclone) twice. CCF2-AM dissolved in DMSO was prepared by following manufacture's instruction (LiveBLAzer™ FRET-B/G loading kit, Invitrogen), and added to the cells at a final concentration of 1μM. Cells were kept at room temperature for 2 h and then examined with fluorescent microscope (Zeiss) with filter setting specified by manufacture's document with the substrate CCF2-AM. Pictures obtained from live cell imaging were further organized with ImageJ software (Developed by National Institutes of Health, Bethesda, MD, US.)

Application of cell permeable substrate CP-AM to label and enrich the target protein in live cells with PCA

For mammalian cell enrichment experiment, HEK293T cells were used and transfected with the PCA plasmid. When the confluency reached 90% and 24 h after transfection, cells were washed with PBS, trypsinized, resuspended and seeded into 3 volumes of the original flask. After another 24~36 h, cells will be ready for next experiment. Cell permeable substrate CP-AM was dissolved in DMSO and loaded to cells transfected with PCA plasmid in the same way as CCF2-AM. After incubating at room temperature for 2 h, cells ($1\sim2\times10^8$ cells, collected from $6\times75cm^2$ flasks) were washed once with 1X HBSS to remove extra substrate and then lysed in 0.5% CHAPS, 1X PBS solution on ice for 30min. Next, cell lysate was separated from insoluble fraction by centrifuging at 40,000×g. The "Click Chemistry" reaction was carried out by using azide conjugated glass beads (binding capacity 136 μmol/g). In the reaction, 10 mg glass beads were mixed with the cell lysate sample, and 1mM copper sulphate ($CuSO_4$), 1mM ascorbic acid and 0.1mM TBTA were used as a compound of catalytic system. The reactants were mixed and rotated at room temperature for 1 hour. After the "Click" reaction, the sample was transferred into a column and washed intensively with 2% CHAPS, 8M urea, and acetonitrile sequentially. In the end, the glass beads were rinsed by water and equilibrated with ammonium bicarbonate (40 mM) and digested by trypsin. The resulting peptides were analyzed by a LC-coupled LTQ-Orbitrap and identified by Mascot.

Acknowledgement

We thank Prof. Stephen Michnick for providing Zip-Bla(1) and Zip-Bla(2) plasmids. We thank Dr. Yan Wang in Proteomics Core Facility at University of Maryland College Park for assistance in LC-MS/MS analysis.

Reference

1. Sarah S. Gallagher, Julia E. Sable, Michael P. Sheetz and Virginia W. Cornish, ACS Chemical Biology, 4(7), 547-556; 2009
2. Jiang J, Zeng D, Li S (2009) Photogenerated quinone methides as protein affinity labeling reagents. *Chembiochem* **10**(4)**:** 635-638
3. Marles JA, Dahesh S, Haynes J, Andrews BJ, Davidson AR (2004) Protein-protein interaction affinity plays a crucial role in controlling the Sho1p-mediated signal transduction pathway in yeast. *Mol Cell* **14**(6)**:** 813-823
4. Galarneau A, Primeau M, Trudeau LE, Michnick SW (2002) Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nat Biotechnol* **20**(6)**:** 619-622
5. Cokol M, Nair R, Rost B (2000) Finding nuclear localization signals. *EMBO Rep* **1**(5)**:** 411-415
6. Kosugi S, Hasebe M, Tomita M, Yanagawa H (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A* **106**(25)**:** 10171-10176

Table S1 Primers used in this research

OSL020	**5'-AACTCGAGGCAGAATTCGCTACCACCACCACCGCCAGTTAATAG-3'**
OSL021	**5'-GTGGTAGCGGTGGTGGTGGTAGCACTAGTCTACTTACTCTAGC -3**
OSL035	**5'- GCAGCCGGATCCTTACCAATGCTTAATCAG -3'**
OSL039	**5' –AAGGCTCCGCGGTGGTCACCCAGAAAC-3'**
OSL082	**5'- GTGCTTTTGCCTTGTAAATGCTACCACCACCACCGCTACCACCACCACCGCTACCACCAC -3'**
OSL083	**5'- CAAAAGCACTGTACCCATACGACGCTGATGACGATGATGCTTACGAAATCTCTTTTGAAC -3'**
OSL084	**5'- GCCTTCCACCAGCGGCCTTCAATGTCAGAGACTTGCAGGATTTCATTTTGTTCAAAAGAG -3'**
OSL085	**5'- GTGGAAGGCACGCCGTGCAAACGGTGAAACGGGTATTATTCCAAGCAATTATGTTCAACT-3'**
OSL086	**5'- GCAGCCGGATCCTTAACGATGCATTTCTTCTGGACCATCGATCAGTTGAACAT-3'**
OSL087	**5'- AAGGCTCCGCGGTGGTCAGCAGATTGTTAATAAGCCGCTGCCGCCGCTGCCGGTAGC-3'**
OSL088	**5'-CCGCTGCCGGTAGCAGGCAGCTCTGGTGGTGGTGGTAGCGGTGGTGGTGGTAGCGGT-3'**
OSL100	**5'-TATCGGGGATCCGAAATTAATACGACTCACTATAG-3'**
OSL105	**5'- CTAGAAGGCACAGTCGAGGC -3'**
OSL109	**5'-GATCGTTGGAACCCGGAGCTG-3'**
OSL110	**5'-CAGCTCCGGGTTCCAACGATC-3'**
OSL136	**5'-ATCGGGAGATCTCCCGATCCCCTATGG-3'**
OSL137	**5'-CAGCTGAGATCTCCAGCTGGTTCTTTCCGCC-3'**
OSL166	**5'-ATTCGCGGATCCACCATGGACTACAAAGACGATGACGACAAGGATCCTGCTGCTCTAA-3'**
OSL167	**5'- TCGCGGATCCACCATGTACCCATACGATGTTCCAGATTACGCTGATCCTGCTGCTCTAA-3'**
OSL168	**5'-GATCCTGCTGCTCTAAAACGTGCTAGAAACACTGAAGCCGCCAGG-3'**