Electronic Supplementary Information

A readily synthesized cyclic pyrrolysine analogue for site-specific protein "click" labeling

Ziyang Hao[1], Yanqun Song[1], Shixian Lin[1], Maiyun Yang[1], Yujie Liang[1], Jing Wang[2] and Peng R. Chen[1]*

[1]Beijing National Laboratory for Molecular Sciences (BNLMS), Department of Chemical Biology, College of Chemistry, Peking University, Beijing 100871 (China)
[2]Department of Chemistry, The University of Chicago, Chicago, IL 60637, USA.
Email:pengchen@uchicago.edu

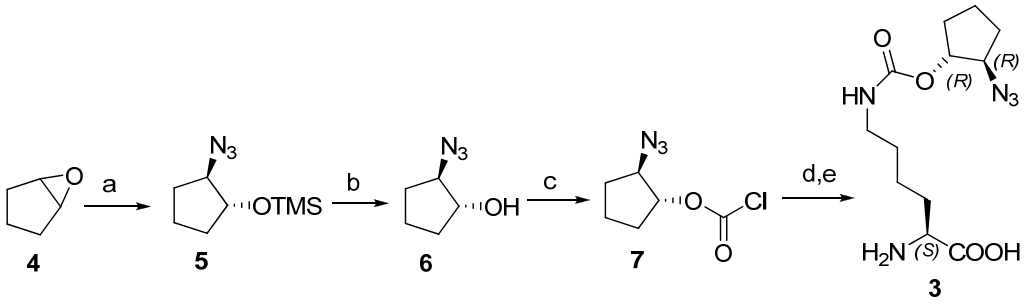
Experimental details

General materials and methods

Primers were ordered from Sangon Biotech (Shanghai) Co., Ltd. Enzymes were ordered from New England Biolabs (NEB). Anti-p53(Ab-1) was purchased from Calbiochem. For chemicals used in unnatural amino acid synthesis: 6-oxabicyclo[3.1.0] hexane, trimethylsilyl azide were purchased from Alfa Aesar, chiral catalyst (S,S)-N,N'-bis(3,5-di-tbutylsalicylidene) was purchased from Sigma Aldrich, Boc-Lys-OH was purchased from GL Biochem (Shanghai) Ltd, trifluoroacetic acid (TFA) was purchased from Bomaijie Technology Co., Ltd. For other chemical compouns, Heme was bought from Beijing Tide-Gene Science &Technology Co., Ltd. Tetramethylrhodamine (TMR) alkyne was bought from Invitrogen. DBCO-Alexa Fluor 488 was bought from Click Chemistry Tools. ^{1}H NMR spectra and high resolution mass spectra were recorded on a Bruker 400MHz instrument and Fourier Transform Ion Cyclotron Resonance Mass Spectrometer(APEX IV), rspectively. Protein purification was performed at AKTA UPC 900 system (GE healthcare). All pictures of protein gels including comassie stained SDS-PAGE gel, fluorescent dye labled SDS-PAGE gel and western blotting membrane were taken on ChemDoc

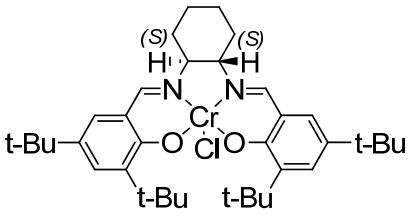
XRS+ (Bio-Rad).Fluorescence spectra were taken with Fluorescence Spectrometer (F-4600, HITACHI, Japan). UV-vis spectra were taken on UV-vis spectrometer (U-3010, HITACHI, Japan).

Chemical synthesis



Scheme S1. Synthesis of ACPK (3). (a):$TMSN_3$, chiral catalyst, yield: 92%; (b):MeOH, con.HCl, yield: 95%; (***c***) triphosgene, THF, yield: 94%; (d):Boc-Lys-OH, $NaHCO_3/H_2O$, yield: 92%; (e)TFA, CH_2Cl_2, yield: 92%.

a)



b)

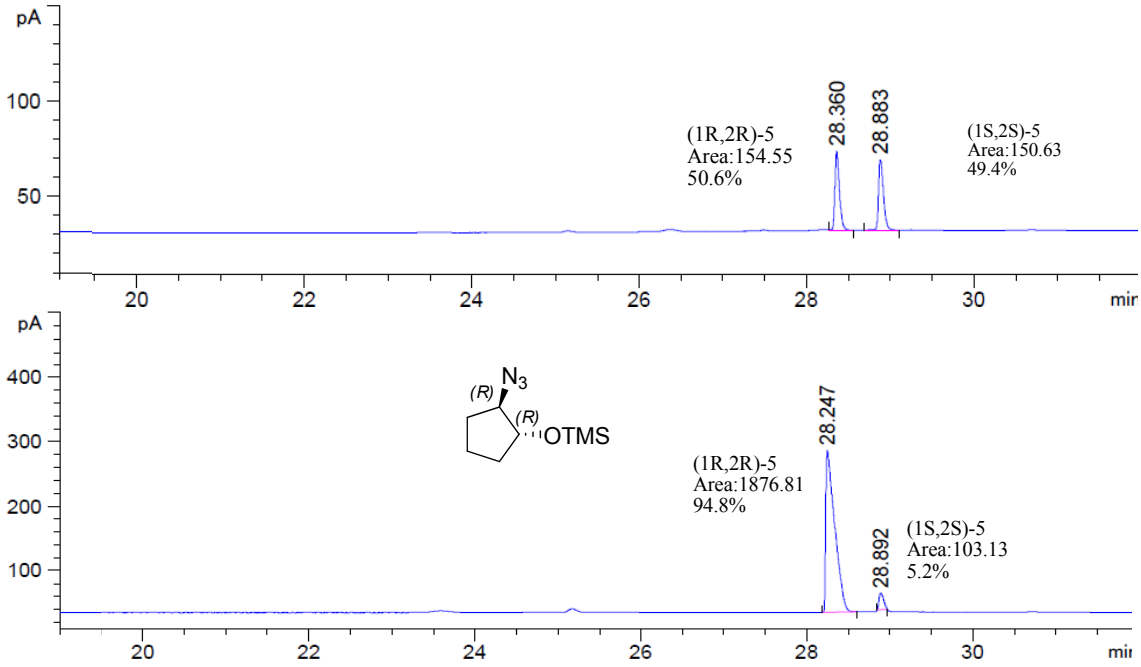


Fig. S1 Optical purity of the synthesized compound 5. a) Structure of chiral(salen)Cr(III) complex: (S,S)-N,N'-bis(3,5-di-t-butylsalicylidene)-1,2-cyclohexane-diaminochromium(III) chloride (comercial name) [1]. ***b)*** Optical purity of compound **5** as analyzed by GC chiral analysis (Varian's CP-Chirasil-Dex CB column ,25m*0.25mm). Top: racemic **5**; bottom: enantioenriched (1R,2R)-**5** (89.6% ee).

Synthesis of ((1R,2R)-2-azidocyclopentyloxy)trimethylsilane 5[1] Following a published procedure for the Asymmetric Ring Opening of 6-oxabicyclo[3.1.0] hexane, 320 mg of chiral catalyst (S,S)-N,N'-bis(3,5-di-t-butylsalicylidene) (0.5 mmol, 0.02 equiv) was added to a 25 mL flask, flushed with N_2 and sealed. 6-oxabicyclo[3.1.0] hexane (2.20 mL, 25.0 mmol) and trimethylsilyl azide (3.5 mL, 26.6 mmol, 1.05 equiv) were added sequentially at room temperature. The reaction mixture was allowed to stir for 12 h, and excess $TMSN_3$ was removed under reduced pressure. The final product **5** was isolated by vacuum distillation (≤40 mmHg, 84 °C) to afford a clear light white oil (4.58 g, 23.0 mmol, 92% yield). With the commercial chiral(salen)Cr(III) catalyst purchased from Sigma Aldrich, the optical purity of compound **5** was calculated to be near 90% ee (89.6% ee) as revealed by GC chiral analysis (Varian's CP-Chirasil-Dex CB column, 25 m*0.25 mm, 0.25 um. Constant flow: 1 mL/min. Temperature program: 60 °C for 2 min,then up to 70 °C at a rate of 10 °C/min and keep for 20 min； then up to 120 °C with 10 °C/min and keep for 10 min).

^{1}H NMR($CDCl_3$): δ =0.14(t ,9H),1.52-2.02 (m ,6H),3.63(q, 1H),3.98(q, 1H).

Synthesis of (1R,2R)-2-azidocyclopentanol 6. Compound **5** (4.50g, 22.6 mmol) was added into methanol (15 mL). Concentrated HCl (100 uL) was then added. The solution was refluxed for 4 h. Excess methanol was removed under reduced pressure and product **6** was isolated by vacuum distillation(≤40mmHg, 104 °C) to obtained a clear oil in 95% yield (2.59 g ,20.4 mmol).

^{1}H NMR($CDCl_3$):1.58-1.81(m ,4H),1.96-2.11(m ,2H),2.18(d ,1H), 3.70(q, 1H), 4.08(d, 1H)

Synthesis of (1R,2R)-2-azidocyclopentyl carbonochloridate 7. Compound **6** (1.43 g 11.25 mmol) was added to a solution of triphosgene (3.5 g 11.3 mmol) in dry THF (20 mL). The reaction was stirred for 10 h at room temperature, and the solvent was evaporated under vacuum. The residue was dried under vacuum for 1 h, affording 2.0 g (10.5 mmol) **7** in 94% conversion as a clear oil.

^{1}H NMR(CDCl3) δ = 1.80-1.87(m, 4H), 2.15 (m, 2H), 4.00 (m, 1H), 5.04 (m, 1H).

Synthesis of (S)-6-(((1R,2R)-2-azidocyclopentyloxy)carbonylamino)-2-(tert-butoxycarbonylamino)hexanoic acid 8. The chloroformate ***7*** (1.0 g, 5.3 mmol) dissolved in THF (10 mL) was added dropwise to a solution of Boc-lys-OH (1.4 g, 5.6 mmol) in aqueous $NaHCO_3$ (1.2 g,14.3 mmol) at 0 °C. The mixture was stirred for 18 h at room temperature, cooled to 0 ºC again, and washed with ice-cold Et_2O (30 mL). The waterlayer was subsequently acidified by 1 N HCl to pH 1~2 and extracted with ice-cold EtOAc (2×30 mL). The combined organic layers were dried over Na_2SO_4. Solvent was then evaporated to get the crude mixture. Purified **8** was obtained by silica column with PE: EtOAc:HAc 2:1:0.05 to afford 1.67 g (4.18 mmol) light yellow oil in 92% yield.

^{1}H NMR($CDCl_3$): δ =1.44-2.18(m ,21H), 3.18(m ,2H), 3.88(m ,1H), 4.28(m ,1H), 4.90(m ,1H),7.72(s, 2H).
HRMS: *m/z* calcd. for $C_{17}H_{29}N_5O_6$, (MH^+) 400.21961; found 400.21902.

Synthesis of (S)-2-amino-6-(((1R,2R)-2-azidocyclopentyloxy)carbonylamino)hexanoic acid 3. Compound ***8*** (1.67 g, 4.18 mmol) was dissolved in CH_2Cl_2 (15 mL), and TFA (15 mL) was slowly added to the solution. The reaction was stirred at room temperature for 40 min, and excess CH_2Cl_2 was evaporated under vacuum. The residue was re-dissolved in MeOH (5 mL) and precipitated into Et_2O, affording pure ***3*** in 92% yield (1.15 g, 3.84 mmol).

^{1}H NMR (CD_3OD): δ =1.402-2.02(m ,12H), 3.09(m ,2H),3.18(m ,1H), 3.89(m ,1H),4.81(m ,1H).
HRMS: *m/z* calcd. for $C_{12}H_{21}N_5O_4$, (MH^+) 300.16718; found: 300.16664.

Selection procedure for evolving ACPK (3) specific aminoacyl-tRNA synthetase

In order to identify MbPylRS mutants that can selectively aminoacylate $tRNA^{Pyl}_{CUA}$ with ACPK (**3**), three rounds of selections were performed as previously described[2, 3]. For the positive selection, the pBK-PylRS plasmids encoding the MbPylRS active site library were transformed into E. coli DH10B competent cells harboring pRep-$tRNA^{Pyl}_{CUA}$ to yield a library greater than 1×10^9 CFU, ensuring complete coverage of the original diversity. Cells were grown on LB agar plates containing 25 μg /mL tetracycline (Tet), 50 μg/mL kanamycin (Kan), 68 μg/mL chloramphenicol (Cm) and 1 mM ACPK After incubation at 37 °C for 48 h, colonies on the plates were pooled, total plasmids were isolated and pBK-PylRS plasmids were separated by agarose gel electrophoresis. The extracted pBK-PylRS plasmids were transformed into DH10B harboring pNEG-$tRNA^{Pyl}_{CUA}$ to start the negative selection. After eletroporation, the cells were allowed to recover for 2 h at 37 °C before being transferred to LB agar plates containing 50 μg/mL Kan, 100 μg/mL ampicillin (Amp) and 0.2% arabinose. The plates were incubated

for 12 h at 37 °C at which point the cells were pooled and the pBK-PylRS plasmids were extracted. Single colonies after the third positive round of selection were picked and their ability to survive upon Cm challenge was tested with increasing concentrations of Cm with and without 1 mM ACPK. Ten colonies that successfully passed through these selection procedures were sequenced and the results are summarized in **Table S1**.

Protein MS-spec

To acquire intact protein mass spectra, the purified HDEA-His_6 protein containing **ACPK** at residue V58 (HdeA-V58-ACPK) was buffer-exchanged by Hitrap desalting column (GE healthcare) into desalting buffer (20 mM tris-HCl, pH 7.8) and concentrated to ~ 0.3 mg /mL. Intact protein mass spectrum was acquired on an automated ESI/MS system (CapLC-ESI-Q-TOF).

Sequence of wt- HDEA-His_6 is:

ADAQKAADNK KPVNSWTCED FLAVDESFQP TAVGFAEALN NKDKPEDAVLDVQGIAT**V**TP
AIVQACTQDK QANFKDKVKG EWDKIKKDMK LGPEQKLISE EDLNSAVDHH HHHH

The calculation of the ESI-MS data to confirm the correct ACPK incorporation into HdeA are as follows:

	MW (Da)
wt-$HdeA_{(His6)}$	12631
Val *(free)*	117
ACPK *(free)*	298
HdeA-V58-$ACPK_{(His6)}$ (*expected*)	12631-117+298 =12812
Obtained from our ESI-MS	12811

Protein expression and purification in E. coli

Green fluorescent protein expression was carried out in *E. coli* BL21(DE3) cells co-transformed with plasmids expressing ACPK-RS-$tRNA^{Pyl}_{CUA}$ pair and GFP-N149TAG. Cells were grown in LB medium containing ampicilline (50 mg /mL) and chloramphenicol (50 mg/mL) with shaking overnight at 37 ºC. After 1:100 dilution in LB medium containing ampicilline (50 mg/mL) and chloramphenicol (50 mg/mL), the culture was grown at 37 °C to an OD_{600} ~0.5. Unnatural amino acid (100 mM) was added by 1:100 dilution to a final concentration of 1mM and incubated with cell culture for 30 min. Protein expression was induced by the addition of arabinose and IPTG to the final concentration of 0.2% and 1mM, respectively. After expression for 12 h, cells were harvested by centrifugation

(10000 g, 5 min), and resuspend in lysis buffer (Tris-Hcl pH 7.8 300 mM NaCl). Bacterial lysate after sonication was loaded onto a Ni-NTA column (Histrap 5 mL, GE healthcare). The column was washed with 30 ml washing buffer (20 mM tris-HCl, pH 7.8, 300 mM NaCl, 40 mM imidazole) and then eluted with elution buffer (20 mM tris-HCl, pH 7.8, 300 mM NaCl and 250 mM imidazole). The final yield of full-length GFP protein containing **3** at the 149 position (GFP-N149-**3**) is ~10mg/L.

For the expresion of HdeA protein, plasmids expressing ACPK-RS-$tRNA^{Pyl}_{CUA}$ pair and HdeA-V58TAG were co-transfomed into DH10B competent cells. Rest procedures of protein expression and purification were similar as above.

Native PAGE gel analysis

The native PAGE gel was prepared using Bio-Rad Mini-PROTEAN Tetra Electrophoresis System (see Figure S2). SDS was removed from the ingredients of both the stacking gel(pH6.8, 4%) and the resolving gel(pH8.8, 15%). All the protein samples were prepared in TRIS buffer containing bromophenol blue, native running buffer (1 L) was consisted of 14.4g glycine and 3.03 g TRIS base plus ddH_2O. Samples were run under ice-cold condition (150 V, 400 mA, 60 min). Soybean trypsin inhibitor (from Beijing Biodee Biotechnology Co.,Ltd, pI=4.6, MW~21 kDa) was used as the marker protein (HdeA protein pI=5.2)

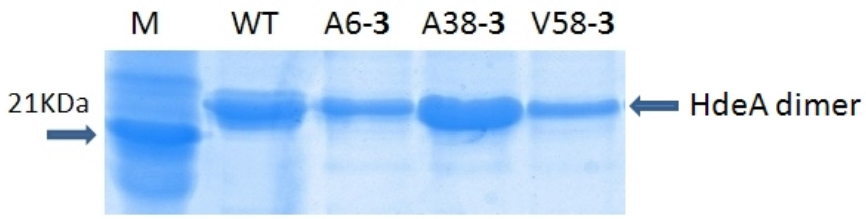


Fig. S2 Protein Native PAGE gel confirming HdeA dimer formation. ACPK was introduced at different sites (Ala 6, Ala 38 and Val 58) on HdeA, which didn't interfere HdeA's dimer formation.

Fluorescent labeling of proteins by click reations in vitro and in vivo.

CuAAC in vitro: 60 μL HdeA V58TAG protein solution ([conc]=320 μM) was mixed with 3 μL 19 mM alk-TMR, 1 μL 500 mM $CuSO_4$ (1 equiv) and 10 μL 100 mM ascorbic acid (2 equiv) in a tube filled with nitrogen.[4] Then the mixture were allowed to react at 37 ^oC for 4 h. Excessl dye was removed by passing through a desalting column.

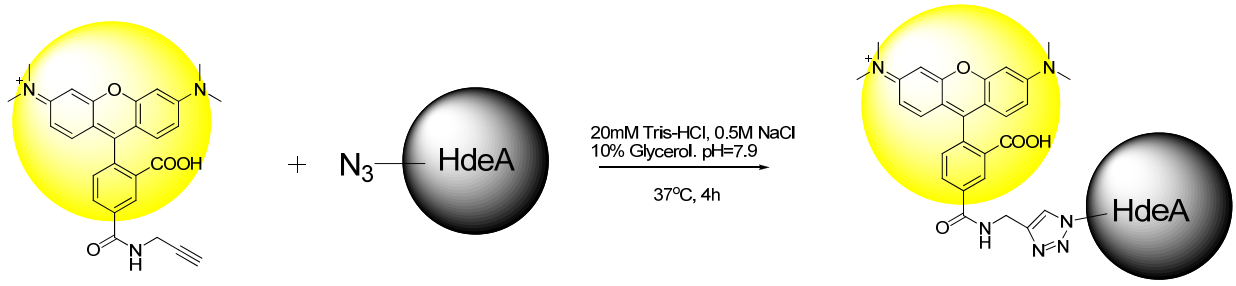


Fig. S3 Conjugation of alk-TMR with HdeA-3 by CuAAC

Copper-free click reaction: 5 μL DBCO-488(100 mM, 5 eq, DMSO) was added into 1mL NEAT-K75-**3** (100 μM) protein in PBS buffer, and the reaction was performed at 4 °C for 2 h[5]. Excess dye was seperated by passing through desalting column. For the calculation of click labeling efficiency, heme was removed after fluorescent labeling and further purification of the NEAT-K75-488 protein. Then BCA assay was performed to calculate the concentration of the protein while the concentration of Fluor-488 was obtained according to its UV absorption at 520 nm. With the optimized reaction and heme removal condition, the click labeling efficiency is 90%.

Biocompatible CuAAC in living E. coli cells: 1 ml *E. coli* bacteria cells expressing HdeA-V58-ACPK was washed twice by PBS buffer and resuspended in 200 μL the same 1X PBS as the labeling buffer (pH 7.4). Alk-coumarin was added to the buffer to 100 μM final concentration together with 2.5 mM sodium ascorbate, 80 μM $CuSO_4$ and 15 μM BTTES. The reaction was allowed to proceed at room temperature for 30-45 min before being quenched by excess Cu(I) chelator BCS. The cells were then washed three times with 500 μL of ice-colde PBS buffer before being applied for flow cytometry analysis.

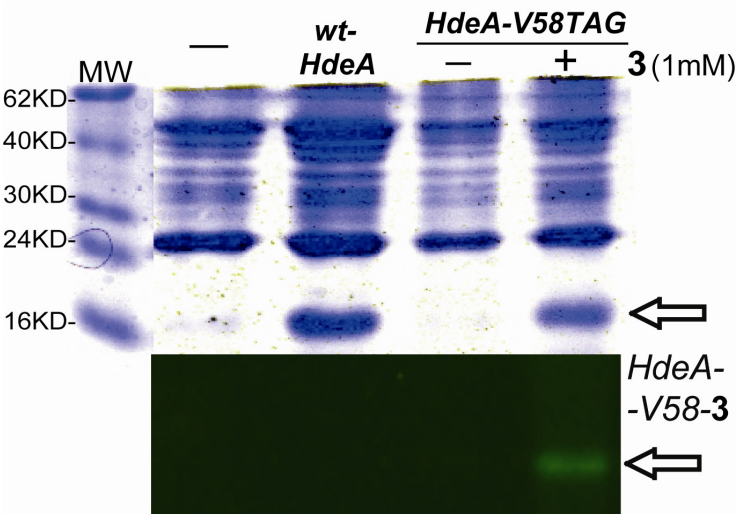


Fig. S4. SDS-PAGE analysis on lysates of E. coli cells expressing wt-HdeA, HdeA-V58-ACPK or without HdeA after being reacted with alk-coumarin by biocompatible CuAAC in vivo.

IsdA NEAT expression and purification

We amplified gene of IsdA NEAT Domain[6, 7] (363 base pair) by PCR from *S. aureus* chromosome (primers. forward: GGAATTC CATATG GCAACATCACAACCAATTAAT. Reverse: CCG CTCGAG TTA AGCTAATGTAGGAATTGC), with NdeI and XhoI cutting sites. Then it was incorporated into pET28b with N-terminal 6X-His tag with NdeI and XhoI restriction enzyme cutting sites. The plasmid was then transformed into *E.coli* strain BL21(DE3) and grew on Kanamycin plate for 16 h. Single colony was picked up into 10 mL LB broth with 50 μg/mL Kanamycin and cultured for overnight at 37 ^{o}C, 220 rpm. Then the medium was enlarged to 1 L and cultured for about another 3 h when the OD_{600} was about 0.8, IPTG was added into medium to 0.5 mM to induce protein expression. Then cells were put into shaker for 16 h at 30 ^{o}C, 220 rpm. Cells were centrifuged down at 4 ^{o}C and resuspended in lysis buffer (PBS buffer with 1 mM PMSF (phenylmethylsulfonyl fluoride), and the solution was lysated by sonicator. Then insoluble materials were removed by centrifuge at 13000 rpm for 40 min. The protein was separated by loading the binding buffer across 5 mL HiTrap column, washed with 50 mL washing buffer (PBS with 40 mM imidazole) and eluted with elution buffer (PBS with 250 mM imidazole). All the processes were performed on ice. Based on this protein sample, we got the UV-vis of this protein with clear absorption peak of Hemoprotein at 405 nm.

Based on this, we did the K75TAG mutation and co-transformed this plasmid with plasmid expressing ACPK-RS-$tRNA^{Pyl}_{CUA}$ pair into *E.coli* strain BL21 DE3 and grew on plate with 50 μg/mL Kanamycin and 50 μg/mL chloramphenicol. Single colony was picked up, and cultured in 10 mL LB at 37 ^{o}C, 220 rpm overnight followed which was the same to wild-type. Then the pre-cultured cells was inoculated into 1 L LB with 50 μg/mL Kanamycin and 50 μg/mL chloramphenicol (final concentration). When bacterial grew to OD_{600}=0.5 at 37 ^{o}C, 220 rpm, 10 mL ACPK was added into medium and the bacterial was cultured for another 0.5 h, at which point 0.5 mL 1M IPTG and 5 mL 20% L-arabinose were added (the final concentrations were 0.5 μM IPTG and 0.1% L-Arabinose), then the medium was induced for16 h at 30 ^{o}C, 220 rpm. Then the process of purification was similar to wild-type proteins. SDS-PAGE was used to verify the incorporation of **3** at K75 site on IsdA-NEAT (see Figure S5a)

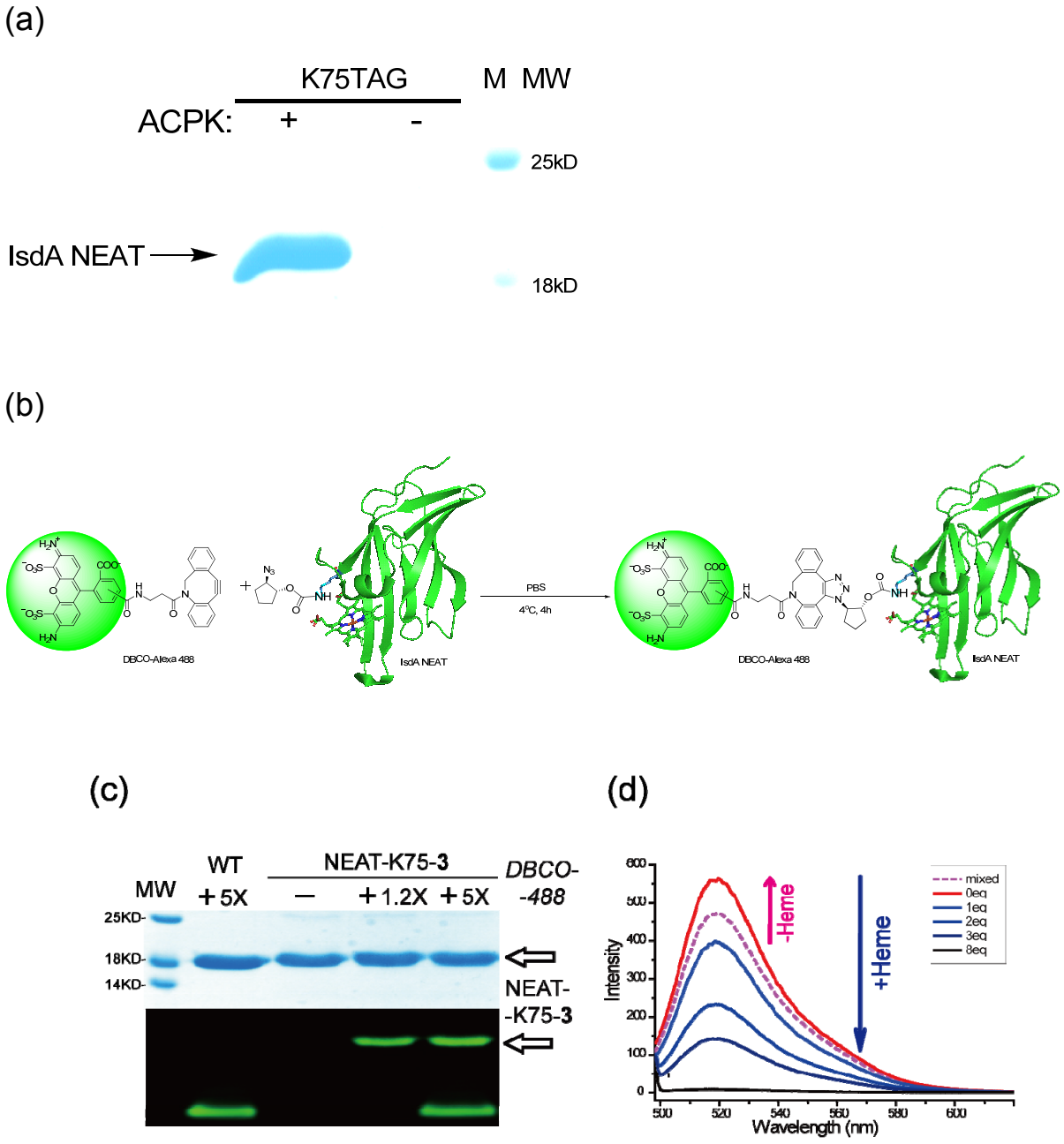


Fig. S5 Site-specific "click" labeling of a fluorophore near IsdA "heme-pocket" through ACPK (3). (a) SDS-PAGE showing the incorporation of ACPK(3) at K75 position on IsdA-NEAT. (b) Conjugation of DBCO-488 with IsdA-NEAT-3 by "copper-free" click chemistry.(c) (b) Visualizing the "click" labeled NEAT-K75-3 protein on SDS-PAGE gel by both coomassie blue stain (up) and UV excitation (below; the bottom fluorescent bands are residual DBCO-488 dye). (d) Fluorescence spectra of NEAT-K75-488 with and without Heme (excitation:495 nm; emission:518 nm).

Heme extraction from IsdA-NEAT

Cold protein solution in 200 mM NaCl buffer was mixed with equal amount of methyl ethyl ketone (MEK)[8], followed by suitable amount 0.1 M HCl to adjust pH to 2.0. And the tube was vigorous shaking till two phases were separated, aqueous phase was dialysed exhaustively against cold 200 mM NaCl solution for several times. Finally the protein solution was dialysed against the PBS buffer in cold room. With this method, over 95% heme can be removed from protein.

FRET analysis of IsdA-NEAT-K75-ACPK upon heme addition

We employed this newly evolved ACPK-RS- $tRNA^{Pyl}_{CUA}$ pair and **3** to monitor the reversible heme binding and release from IsdA, a crucial heme-transfer protein from the iron-regulated surface determinants (Isd) system in bacterial pathogens such as *Staphylococcus aureus* (Fig. S5)[7]. IsdA, by working together with other surface proteins IsdB, IsdC, IsdH and ABC transporter IsdDEF, constitutes the acquisition machinery for heme as a preferred source of the essential iron element for almost all organisms. During bacterial pathogenesis, it has been proposed that IsdA abstracts heme from bound hemoglobin that had been previously captured by IsdB. Then heme was transferred from IsdA to IsdC, and eventurally to ABC transporters IsdDEF. However, detailed mechanism of heme acquisition and transfer by this system remains to be elucidated.

The NEAr transporter (NEAT) domain of IsdA has been previously shown to be involved in direct heme binding and exhibits a higher solubility than the full-length IsdA protein.[7] We thus introduced **3** at K75 site on IsdA-NEAT domain near the heme pocket and the highly efficient "copper-free" click reaction was performed by incubating with Dibenzylcyclooctyne-Fluor 488 (DBCO-488) at 4 ^{o}C for 2 h (Figs. S5c). Since recombinant IsdA-NEAT expressed and purified from *E. coli* is a mixture of apo- and holo- forms, a partially quenched fluorescence was observed due to the energy transfer between heme and Fluor 488 in holo-NEAT-K75-488 (Fig. S5d). After removal of the bound heme by methyl ethyl ketone (MEK), an enhanced fluorescence signal was observed, corresponding to the apo-NEAT-K75-488. The efficiency of this *in vitro* click labeling calculated after protein purification and heme-removal is over 90%. Addition of heme led to a significantly decreased fluorescence signal and the homogeneous holo-NEAT-K75-488 was obtained by increment loading of heme. With the aforementioned biocompatible CuAAC being successfully applied for protein labeling *in vivo*, a cell-permeable fluorophore could be introduced onto the ACPK-incorporated IsdA-NEAT in living bacteria, which might allow the investigation of heme-transfer among heme-acquisition proteins when the bacterial pathogen invades host cells.

Plasmid transfection and protein expression in mammalian cells

Both H1299 and Hela cells were maintained in Gibco DMEM supplemented with glutamine, penicillin/streptomycin and 10% FBS at 37 °C in a humidified atmosphere of 5% CO_2. In all experiments, cells were transiently transfected using FuGENE HD reagent according to the manufacturer's instructions (Roche).

Western blot analysis

The pCMV-ACPK-RS plasmid was co-transfected into H1299 cells (p53-null) with a plasmid encoding p53 harboring a single amber mutation (TAG) at the N-terminal phosphorylation site Ser20 or at the central tetramerization domain residue Met340. After 18 h growth in the presence or absence of 1 mM ACPK, cells were harvested and washed in phosphate-buffered saline (PBS), followed by lysis in Nonidet P-40 (NP-40) buffer (30 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% NP-40, 5 mM ethylenediaminetetraacetic acid [EDTA], 1 mM sodium orthovanadate, 1 mM NaF, 1 mM PMSF, 1 mM dithiothreitol [DTT], and protease inhibitor cock[Roche]). Aliquots (50 ug) of cell extracts were resolved in 12% SDS-PAGE and transferred to a polyvinylidene difluoride membrane (Bio-Rad). After blocking at room temperature for 1 h in Blotto (5% nonfat dry milk in 1×TBST), the membranes were incubated with anti-p53 (Ab-1,epitope within amino acids 376-378 at the C-terminus of human p53) antibody (Calbiochem) overnight at 4 °C followed by incubation with horseradish peroxidase (HSP)-conjugated goat anti-mouse secondary antibody (Santa Cruz) at room temperature for 2 h. Antibodies were detected with ECL reagents (Pierce) and anti-actin was used as a control (Fig. S7)

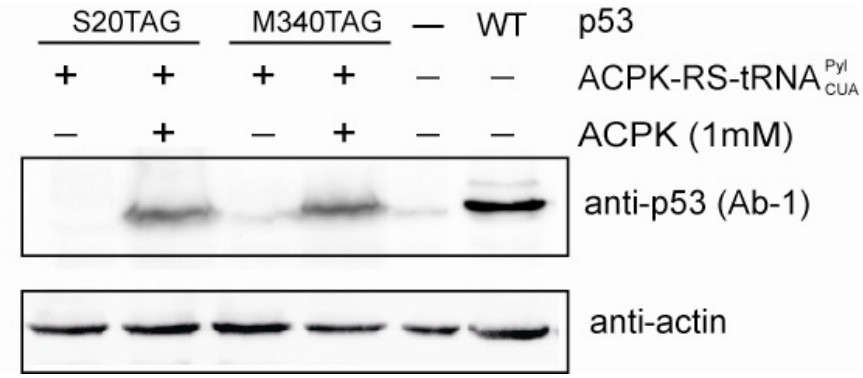


Fig. S6 Immunoblotting analysis of p53 containing ACPK at residue S20 or M340 in H1299 cells (p53-null) harboring ACPK-RS- $tRNA^{Pyl}_{CUA}$ ***pair with or without 1 mM ACPK.***

Fluorescence imaging

Hela cells were plated on six-well containing glass coverslips (Costar) to obtain 80-90% confluency. The next day, cells were transfected with p53-EGFP or p53EGFP-372TAG using HD transfection reagent (Roche; 8 μl Fugene HD + 0.8 μg of pCMV-ACPK-RS + 1.2 μg of p53EGFP or p53EGFP-K372TAG for 2 ml cell culture). After 24h, cells on the coverslips were washed three times with PBS and then fixed in 4% paraformaldehyde/PBS for 10 min at room temperature. After 3 washes with ice-cold PBS, cells were permeabilized in ice-cold PBS containing 0.2% Triton X-100 for 10 min, followed by blocking in PBS containing 1% bovine serum albumin and 1 ug/mL of DAPI (Sigma) at room temperature for 30 min in dark. Cells were washed three times with PBS, and the stained cells were mounted with mounting medium and the cover slips were sealed with nail polish. Fluorescence imgages were recorded under confocal microscopy (Olympus, FV-1000, **Fig. 2**).

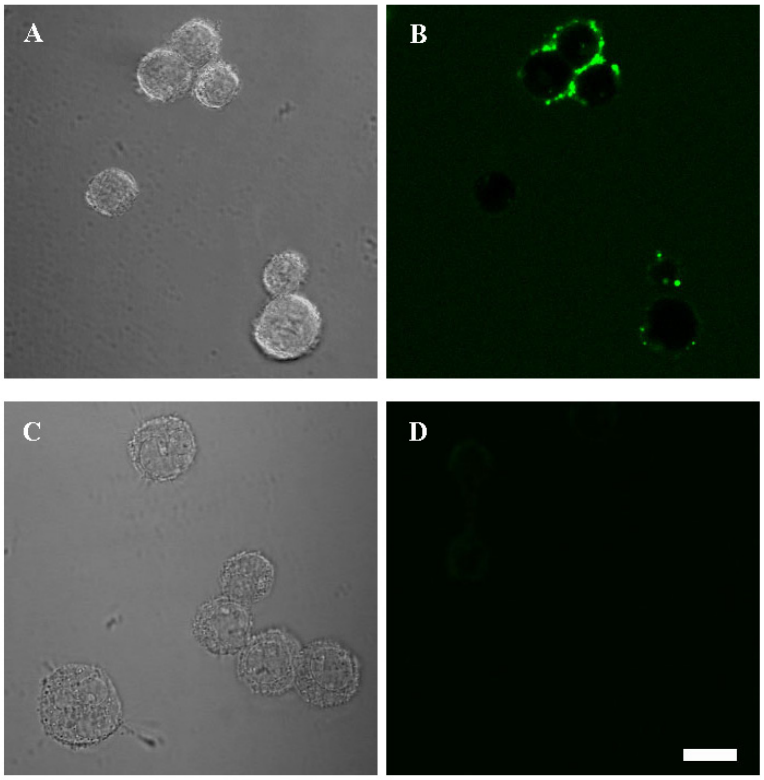
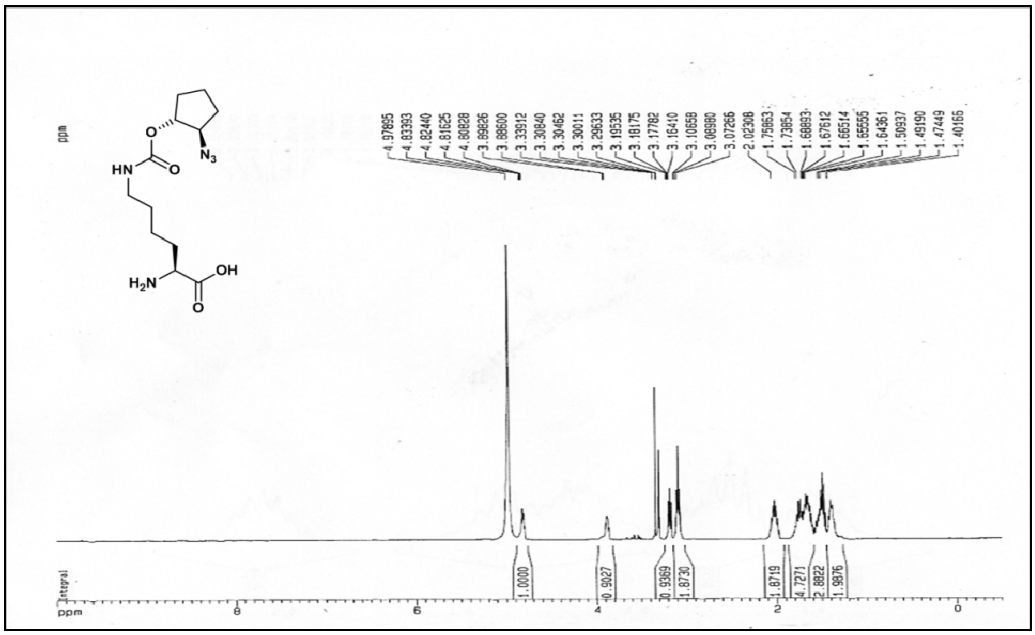
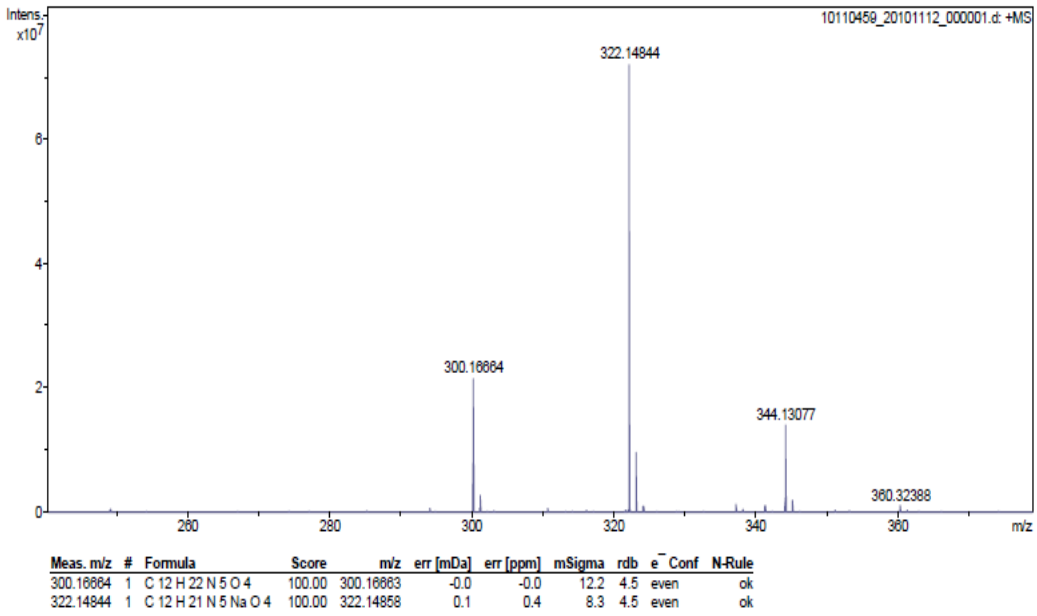


Fig. S7. Imaging of chemical attachment of EGFP-azide on living HEK 293T cell surface. 293T cells were incubated with (A and B) or without (C and D) 50 μM Ac4ManNAl for 3 days, followed by reacting with 15 μM EGFP-azide by Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) using BTTES ligand. Channels shown are differential interference contrast (A and C), and the green GFP channel (B and D). (Scale bar: 20 μm)

a)



b)



Meas. m/z	#	Formula	Score	m/z	err [mDa]	err [ppm]	mSigma	rdb	e⁻ Conf	N-Rule
300.16664	1	C 12 H 22 N 5 O 4	100.00	300.16663	-0.0	-0.0	12.2	4.5	even	ok
322.14844	1	C 12 H 21 N 5 Na O 4	100.00	322.14858	0.1	0.4	8.3	4.5	even	ok

Fig. S9 NMR and HRMS spectra of ACPK (3). *a)*.^{1}H NMR (CD_3OD) spectrum of ACPK (**3**). *b)* HRMS spectrum of ACPK (**3**).

Table S1. The sequence of 10 selected ACPK-specific MbPylRS mutants from directed evolution.

MbPylRS	Leu270	Tyr271	Leu274	Cys313	Tyr349	Frequency
ACPK-RS-A	Leu	Tyr	Ala	Ala	Phe	3 clones
ACPK-RS	Leu	Tyr	Ala	Val	Phe	7 clones

Construction of the MbPylRS active-site mutant library

The active-site mutant library was constructed by overlap extension polymerase chain reaction (OE-PCR) using synthetic oligonucleotide primers to introduce mutations.[2] The wild-type *Methanosarcina barkeri* PylRS (MbPylRS) gene served as the template to perform standard PCR reactions. Four pairs of primers were used to fragment MbPylRS gene into four pieces by which the codons for the intended mutations were replaced by NNK (N=A or C or G or T, K=G or T). OE-PCR was employed to assemble these PCR fragments and multiple rounds of PCR were conducted with the combination of primers listed above. The intact MbPylRS gene was generated by this strategy and the desired mutation sites were substituted by NNK, so that all 20 common amino acids were encoded. The final products were amplified and digested with *Nde*I and *Pst*I. The selection vector pBK-MJYRS[9, 10] was digested with the same two restriction enzymes and purified by Gel-extraction kit. DNA ligation was conducted at 16 °C for 16 h followed by yeast-tRNA enhanced ethanol precipitation. Finally, the precipitated ligation products were transformed into electrocompetent DH10B cells. The theoretical diversity for this five codon randomized library is 3.4×10^7; the final transformation yielded $\sim 1.0 \times 10^8$ mutants, estimated by CFU counting, which provided 3× coverage of the theoretical diversity. The quality of the library was validated by sequencing 20 individual clones, which revealed that there was no codon bias in the randomized region.

Notes and references

1. S. E. Schaus, J. F. Larrow and E. N. Jacobsen, *J. Org. Chem.*, 1997, **62**, 4197-4199.
2. P. R. Chen, D. Groff, J. T. Guo, W. J. Ou, S. Cellitti, B. H. Geierstanger and P. G. Schultz, *Angew. Chem. Int. Edn. Engl.*, 2009, **48**, 4052-4055.
3. H. Neumann, S. Y. Peak-Chew and J. W. Chin, *Nat. Chem. Biol.*, 2008, **4**, 232-234.
4. D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2009, **131**, 8720-8721.
5. J. C. Jewett, E. M. Sletten and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2010, **132**, 3688-3689.
6. J. C. Grigg, C. L. Vermeiren, D. E. Heinrichs and M. E. P. Murphy, *Mol. Microbiol.*, 2007, **63**, 139-149.
7. (a) M. Pluym, N. Muryol, D. E. Heinrichs and M. J. Stillman, *J. Inorg. Biochem.*, 2008, **102**, 480-488. (b) S. K. Mazmanian, E. P. Skaar, A. H. Gaspar, M. Humayun, P. Gornicki, J. Jelenska, A. Joachmiak, D. M. Missiakas and O. Schneewind, *Science*, 2003, **299**, 906-909; (*c*) H. Zhu, G. Xie, M. Y. Liu, J. S. Olson, M. Fabian, D. M. Dooley and B. F. Lei, *J. Biol. Chem.*, 2008, **283**, 18450-18460.
8. F. Ascoli, M. R. Fanelli and E. Antonini, *Methods Enzymol*, 1981, **76**, 72-87.
9. L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498-500.
10. J. M. Xie, W. S. Liu and P. G. Schultz, *Angew. Chem. Int. Edn.*, 2007, **46**, 9239-9242.