Supplementary Information

Small Molecule Control of Protein Function through Staudinger Reduction

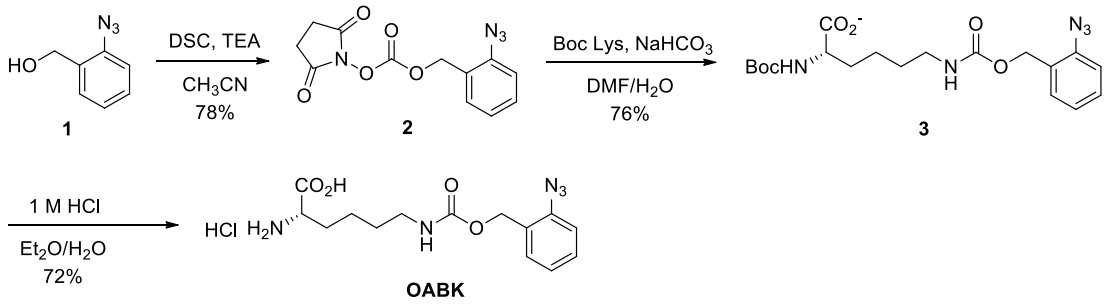
Ji Luo, Qingyang Liu, Kunihiko Morihiro, and Alexander Deiters*

University of Pittsburgh, Department of Chemistry, Pittsburgh, PA 15260, deiters@pitt.edu.

Table of Content

Supplementary Methods

Synthesis of *ortho*-azidobenzyloxycarbonyl lysine (OABK) HCl salt



2-Azidobenzyl succinimidyl carbonate (2). *N,N'*-Disuccinimidyl carbonate (2 eq., 5.2 g, 20 mmol) was added to a solution of the alcohol **1** (1 eq., 1.5 g, 10 mmol) in CH_3CN (44 ml), followed by TEA (3 eq., 4.2 ml, 30 mmol). The reaction mixture was stirred overnight at room temperature and concentrated under reduced pressure. The crude product was purified by silica gel chromatography with DCM/acetone (19:1) as the eluent, affording **2** as a white solid in 78% yield (2.28 g, 7.8 mmol). ^{1}H NMR (300 MHz, $CDCl_3$): δ = 2.66 (s, 4H), 5.11 (s, 2H), 7.13-7.19 (m, 2H), 7.39-7.50 (m, 2H). ^{13}C NMR (400 MHz, $CDCl_3$): δ = 25.6, 73.6, 118.5, 125.0, 125.1, 131.2, 132.4, 139.8, 171.2. HRMS: *m/z* calcd for $C_{12}H_9N_4O_5$ $[M-H]^-$: 289.0573; found: 289.0591.

***N*^{6}-(((2-Azidobenzyl)oxy)carbonyl)-*N*^{2}-(*tert*-butoxycarbonyl)-*L*-lysine (3).** Boc-protected lysine (1.2 eq., 2.3 g, 9.4 mmol) was added to a solution of the carbonate **2** (1 eq., 2.28 g, 7.8 mmol) in 80 ml of DMF/water (1:1), followed by potassium carbonate (3 eq., 3.2 g, 23.4 mmol). The mixture was stirred overnight at room temperature, concentrated, acidified to pH 3-4 with 1 M aq. HCl, and extracted with Et_2O (3 × 20 ml). The combined organic layers were washed with water (30 ml) and brine (30 ml), and dried over Na_2SO_4. After filtration, the solvent was removed under reduced pressure, affording the product **3** as a yellow foam in 76% yield (2.6 g, 5.9 mmol). ^{1}H NMR (300 MHz, $CDCl_3$): δ = 1.42 (s, 9H), 1.42-1.90 (m, 6H), 3.15-3.23 (m, 2H), 4.23-4.50 (m, 1H), 5.06 (s, 2H), 6.95-7.02 (m, 2H), 7.35-7.37 (m, 2H). ^{13}C NMR (400 MHz, $CDCl_3$): δ = 13.7, 22.1, 27.8, 28.3, 28.4, 38.7, 52.9, 59.9, 72.4, 78.9, 117.6, 124.3, 126.5, 129.4, 130.9, 155.3, 161.9, 174.2. HRMS: *m/z* calcd for $C_{19}H_{26}N_5O_6$ $[M-H]^-$: 420.1918; found: 420.1904.

***N*^{6}-(((4-Azidobenzyl)oxy)carbonyl)-*L*-lysine hydrochloride salt (OABK).** A solution of 1 M aq. HCl (2.4 ml) was added to **3** (1 eq, 100 mg, 0.24 mmol) in Et_2O (2.4 ml). The reaction mixture was stirred for 2 days and washed with Et_2O (3 × 5 ml). The aqueous layer was concentrated under reduced pressure, affording the product **OABK** as a white solid in 72% yield (62 mg, 0.17 mmol). ^{1}H NMR (400 MHz, DMSO): δ = 1.23-1.49 (m, 4H), 1.69-1.71 (m, 2H), 2.97-2.99 (q, *J* = 5.7 Hz, 2H), 3.80-3.85 (t, *J* = 5.7 Hz, 1H), 4.96 (s, 2H), 7.18-7.46 (m, 4H). ^{13}C NMR (400 MHz, DMSO): δ = 21.31, 26.17, 29.29, 38.13, 51.90, 60.78, 118.58, 124.91. 127.95, 129.54, 129.59, 137.34, 155.85, 170.80. HRMS: *m/z* calcd for $C_{14}H_{19}N_5O_4$ $[M+H]^+$: 322.1515; found: 322.1544.

Cytotoxicity studies of phosphines

HEK 293T cells were seeded at 10,000 cells per well in clear bottom 96-well plates and incubated for 24 h. Cells were treated in triplicate with phosphine derivatives or 0.1% DMSO for 20 h. Cells were washed with 150 µL of fresh DMEM, and then 100 µL of DMEM and 40 µL of activated XTT-menadione reagent (1 mg/mL XTT in PBS containing 8 µL/mL menadione activation reagent, Alfa Aesar) were added to each well according to the manufacturer's instructions. Absorbance was read out at 450 nm (XTT) and 630 nm (background) using a Tecan M1000 Pro microplate reader. After 4 h incubation, absorbance was measured again. The cell viability was determined by subtracting 630 nm readings from 450 nm absorbances, then subtracting the averages of background wells (starting time point) from experimental wells (4 h), followed by normalization to the 0.1% DMSO control.[1]

Plasmid construction for mammalian cell applications

1) p*Mb*PylOABKRS-mCherry-TAG-EGFP: The plasmid (pPylRS_Y271A-mCherry-TAG-EGFP) was obtained by converting the TAC (Y271) codon of wild type PylRS to a GGC (A271) codon using primers P1 & P2 (table below) and a QuikChange site-directed mutagenesis kit (Agilent). The plasmid (pPylRS_Y271A_Y349F-mCherry-TAG-EGFP) was generated by converting the TAC (Y349) codon of PylRS_Y271A into a TTC (F349) codon using primers P3 & P4 (table below) and the same mutagenesis method.

2) p*Mb*PylOABKRS: The *E. coli* codon-optimized plasmid (*Mb*PylOABKRS_Y271A_Y349F) was constructed using the same mutagenesis method as above. A and F mutations were introduced into wild-type *Mb*PylRS at Y271 and Y349 sites with two pairs of primers (P5 & P6 and P7 & P8; see table below), respectively.

3) p*Mb*PylOABKRS-PylT: The plasmid was obtained by ligating the p4CMVE-U6-PylT fragment from p*Mb*4PylT between the *Nhe*I and *Mfe*I restriction sites of p*Mb*OABKRS.

4) p*Mb*PylOABKRS-wtCre and p*Mb*PylOABKRS-Cre-K201TAG: The wtCre and Cre-K201TAG fragments were PCR amplified from p*Mb*PCKRS-wtCre/-CreK201TAG using the primers P9 & P10, digested with *Nhe*I and *Mfe*I, and ligated into the p*Mb*PylOABKRS-mCherry-TAG-EGFP vector in place of the mCherry-TAG-EGFP, generating the pOABKRS-wtCre and pOABKRS-Cre-K201TAG plasmids.

5) p*Mb*PylOABKRS-wtCas9 and p*Mb*PylOABKRS-Cas9-K866TAG: The pwtCas9 and pCas9-K866TAG fragments were generated from p*Mb*CKRS-wtCas9 and p*Mb*CKRS-Cas9-K866TAG[2] using the *Nhe*I and *Mfe*I restriction sites, respectively. p*Mb*PylOABKRS-mCherry-TAG-EGFP was digested with the same restriction enzymes (*Nhe*I and *Mfe*I) to remove the mCherry-TAG-EGFP-HA reporter. The wtCas9 and Cas9-K866TAG genes were ligated into the p*Mb*PylOABKRS-mCherry-TAG-EGFP backbone replacing the mCherry-EGFP using Quick ligase (NEB) to create p*Mb*PylOABKRS-wtCas9 and pMbPylOABKRS-Cas9-K866TAG, respectively.

Primer 1	ctatgctggcccccaccctggccaactacctgcggaaactg
Primer 2	cagtttccgcaggtagttggccagggtgggggccagcatag
Primer 3	gcgacagctgcatggtgttcggcgacaccctggacatc
Primer 4	gatgtccagggtgtcgccgaacaccatgcagctgtcgc
Primer 5	ctatgctggccgccaccctggccaactacctgcggtaactg
Primer 6	cagttaccgcaggtagttggccagggtggcggccagcatag
Primer 7	gcgacagctgcatggtgttcggcgacaccctggacatc
Primer 8	gatgtccagggtgtcgccgaacaccatgcagctgtcgc
Primer 9	gtcagatccgctagcacc
Primer 10	cgatcgatatcaattgtggtttgtccaaactcatca

Plasmid maps are shown in **Supplementary Figure S4**.

Protein deprotection via Staudinger reduction

After lysis of a bacterial expression culture (25 mL), a Staudinger reaction was performed with **5** (1 mM) in cell lysates (5 mL) for 4 h at room temperature under shaking (250 rpm). The reaction mixture was centrifuged (5,000 rpm, 5 min) and the supernatant was transferred to a 15 mL conical tube and 75 µL of Ni-NTA resin (Qiagen) was added. The mixture was incubated at 4 °C for 2 h under mild shaking. The resin was then collected by centrifugation (500 g, 10 min), washed three times with 400 µL of lysis buffer, followed by two washes with 400 µL of wash buffer containing 20 mM imidazole. The protein was eluted with 300 µL of elution buffer containing 250 mM imidazole. The purified protein was analyzed by 10% SDS PAGE and stained with Coomassie Blue. ESI-MS analysis confirmed that sfGFP-**OABK** could be reduced and deprotected via a Staudinger reaction, since the observed mass of 28226.41 Da is in agreement with the expected mass of 28227.89 Da of the deprotected protein.

Mass spec analysis of proteins

High-resolution (≥ 10,000 R FWHM) exact mass measurement of sfGFP-**OABK** was carried out using Electrospray Ionization (ESI) on an Agilent Technologies (Santa Clara, California) 6210 LC-TOF mass spectrometer. Samples were analyzed via a 1 µL flow injection at 300 µL/min in a water:methanol mixture (25:75 v/v) with 0.1% formic acid. The mass spectrometer was operated in positive-ion mode with a capillary voltage of 4 kV, nebulizer pressure of 35 psig, and a drying gas flow rate of 12 L/min at 350 °C. The fragmentor and skimmer voltages were 200 and 60 V respectively. Reference ions of purine at *m/z* 121.0509 and HP-0921 at *m/z* 922.0098 were simultaneously introduced via a second orthogonal sprayer, and used as internal calibrants.

Thermo Q-Exactive Orbitrap mass spectrometer (Bremen, Germany), connected to a Dionex Ultimate 3000 UHPLC system was used for intact protein analysis (phosphine reduction and deprotection of sfGFP-**OABK**). The sample was analyzed through a ProSwift RP-10R, 1 mm by 5 cm column, flow rate of 200 µL/min and ACN gradient (0.1% formic acid) 26-80% for 30 min. The mass spectrometer was operated in ESI positive-ion mode with a capillary voltage of 3.5 kV and resolution at 17,500. Sheath gas, aux gas, and sweep gas flow rates were 35, 10, and 5, respectively. Source temperature was 250 °C. The S-lens RF level voltage was 50 V and the ion transfer tube temperature was 250 °C. The instrument was tuned and calibrated with standard thermo mix solution and the data was collected at the range of m/z 500-3000. The Xcaliber 3.0.63 and Protein Deconvolution 3.0 software was used for the data analysis.

Small molecule activation of Cre-OABK in live cells

HEK 293T cells (~20,000 cells/well) were plated into a 96-well cell culture microplate (Greiner), and incubated overnight in 200 µL of DMEM growth media supplemented with 10% FBS and 1% penicillin/streptomycin at 37 °C in 5% CO_2. At ~75% confluency, HEK293T cells were co-transfected with the p*Mb*OABKRS-wtCre or p*Mb*OABKRS-CreK201TAG, p4CMVE-U6-PylT,[3] and pC-SL[4] plasmids (100 ng of each) using linear PEI (2 µL per well, 0.323 mg/mL) in the presence or absence of **OABK** (0.25 mM) overnight at 37 °C. The media supplemented with **OABK** (0.25 mM) was replaced with 200 µL of fresh DMEM and incubation was continued for 2 h, followed by addition of **5** (0.5 mM) at 37 °C. Fluorescence imaging of the Cre activation reporter was performed after 24 h incubation. Media was replaced with DMEM without phenol red (Thermo Scientific) for imaging on a Zeiss Axio Observer Z1 microscope (10X objective, NA 0.8 plan-apochromat) with DsRed (E_x: BP550/25; E_m: BP605/70) and EGFP (E_x: BP470/40; E_m: BP525/50) filter cubes, then processed in Zen Pro 2012 imaging software. Fluorescent cell counting was performed on a Nikon A1 confocal microscope (10X objective) and analyzed using Elements software.

Small molecule activation of Cas9-OABK in live cells

HEK 293T cells (~20,000 cells/well) were seeded into a 96-well cell culture microplate (Greiner), and incubated overnight in 200 µL of DMEM growth media supplemented with 10% FBS and 1% penicillin/streptomycin at 37 °C in 5% CO_2 humidified atmosphere. Before transfection, media was replaced with antibiotics-free DMEM supplemented with or without **OABK** (0.25 mM). Quadruple transfections were carried out with linear PEI (2 µL per well, 0.323 mg/mL) using p*Mb*OABKRS-wtCas9 or p*Mb*OABKRS-Cas9-K866TAG, p4CMVE-U6-PylT,[3] gRNA plasmids,[2] and pRG dual reporter plasmid[2] (100 ng of each). After a 24 h incubation, cells were washed with fresh DMEM and incubated for 2 h. For regulating gene editing in live cells, HEK 293T cells were treated with **5** (0.5 mM) or without **5**. After overnight incubation, media was replaced with 200 µL of DMEM without phenol red, and cells were imaged on a Zeiss Axio Observer Z1 microscope (10X objective, NA 0.8 plan-apochromat) with EGFP (E_x: BP470/40; E_m: BP525/50) and DsRed (E_x: BP550/25; E_m: BP605/70) filter cubes. Fluorescent cell counting was performed on a Nikon A1 confocal microscope (10X objective) and analyzed using Elements software.

LC-MS analysis of OABK deprotection with 2DPBA

Time course extracted LC-ESI-MS chromatograms for the reaction of Boc-protected **OABK** (0.1 mM) with 2-DPBA (5 equiv.) in PBS buffer (pH 7.4):DMSO = 9:1 at 37 °C (**Supplementary Figure 1**). The reaction mixture was analyzed at different time points using a Hypersil GOLD C18 column, 1.9 um, 2.1 mm i.d. x 100 mm (Thermo Scientific, Waltham, MA) on a Prominence HPLC system (Shimadzu, Kyoto, Japan). The LC system was directly coupled to an electrospray ionization mass spectrometer (LCMS-2020, Shimadzu). LC gradient = 5-60% MeCN in 0.1% aqueous formic acid, Flow rate = 0.2 mL/min, column temperature = 40 °C.

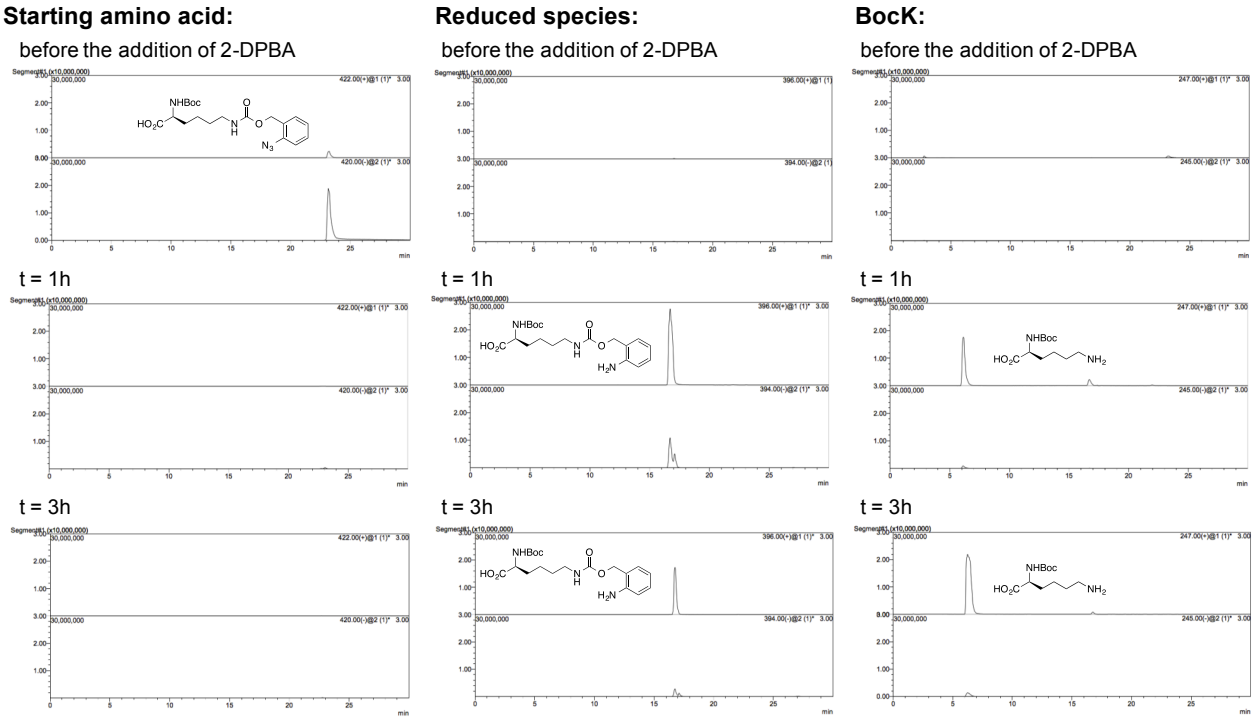
Supplementary Figures and Movies:

→ see *Nature Chemistry* website

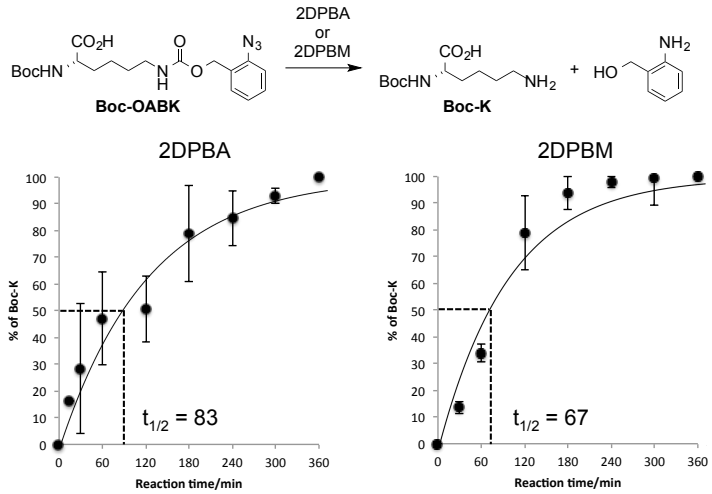
Supplementary Movie 1. Small molecule activation of EGFP fluorescence. HEK 293T cells expressing EGFP-K85TAG-mCherry in the presence of OABKRS/$tRNA_{CUA}$ and 0.5 mM of **OABK** were treated with 2DPBA (500 µM) at time t = 0 min. The EGFP and mCherry fluorescence was followed through time-lapse imaging.

→ see *Nature Chemistry* website

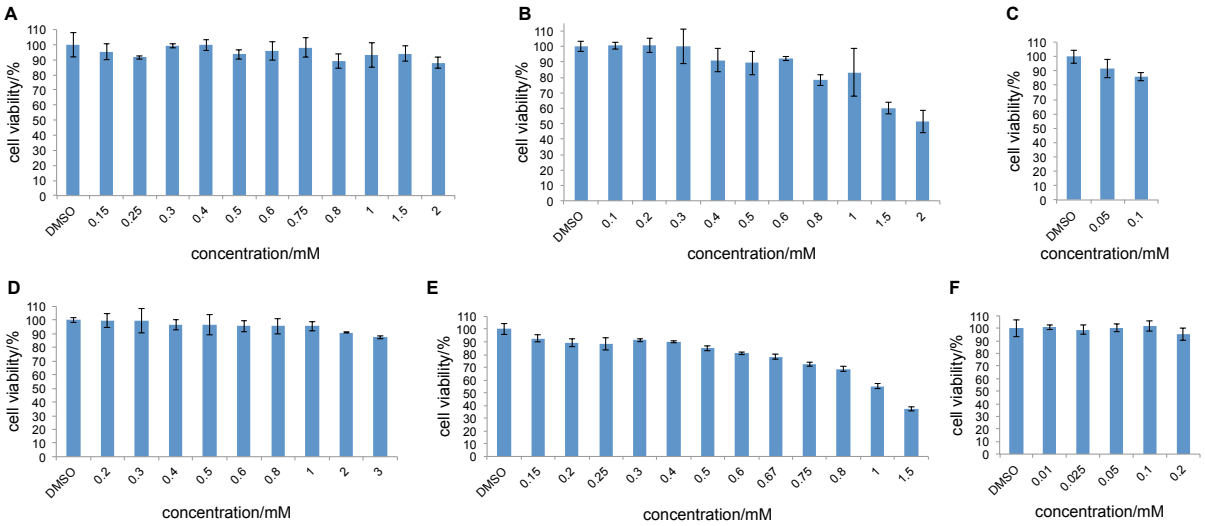
Supplementary Movie 2. Small Molecule activation of protein nuclear translocation. HEK 293T cells expressing EGFP-SatB1-K29TAG-mCherry in the presence of OABKRS/$tRNA_{CUA}$ and 0.5 mM of **OABK** were treated with 2DPBA (500 µM) at time t = 0 min. The EGFP and mCherry fluorescence is followed.



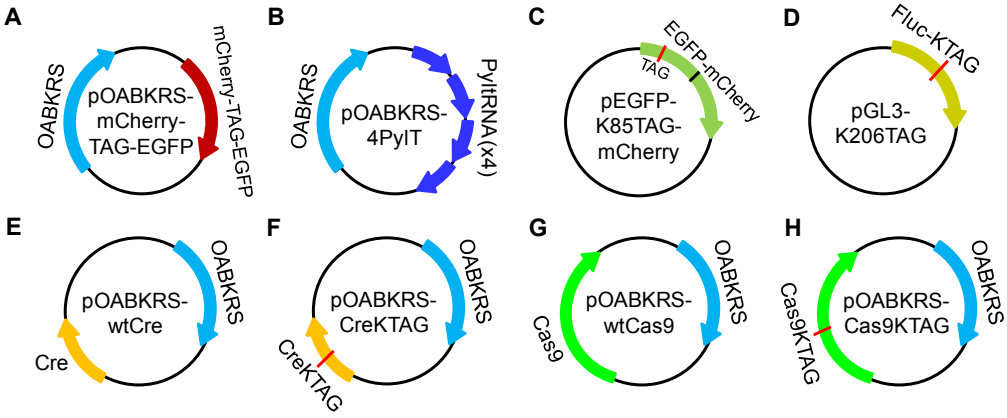
Supplementary Figure S1. LC/MS analysis of the reduction of Boc-**OABK** with **2DPBA**. Time-course extracted LC-ESI-MS chromatograms for the starting amino acid (left), the reduced species (center), and Boc-lysine (right). Upper chromatograms were collected in positive ion mode and lower ones in negative ion mode. The presence of the aniline species indicates that the Staudinger reduction occurs faster than the protecting group fragmentation.



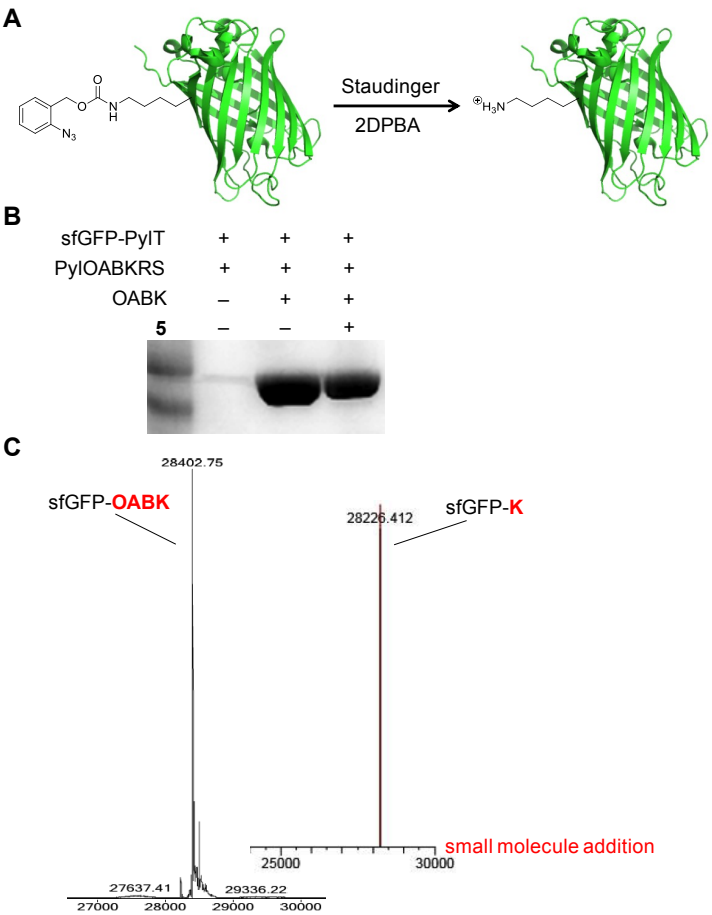
Supplementary Figure S2. LC/MS studies and identification of the half life for the deprotection of **OABK**. **2DPBA** or **2DPBM** (500 μM) was added to a solution of Boc-**OABK** (5 μM) in DMEM/DMSO = 9:1 and the reaction mixture was incubated at 37 ˚C. Aliquots of the reaction mixture were directly analyzed by LC/MS at indicated time points and the conversion into of Boc-K was calculated from the ESI-MS peak area (positive ion mode).



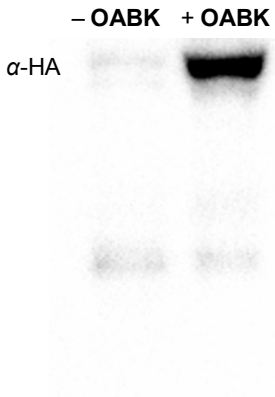
Supplementary Figure S3. Cytotoxicity measurements of phosphine derivatives using XTT assays in HEK 293T cells. **A**) tris(2-carboxyethyl)phosphine (TCEP, **1**), **B**) TCEP methylester (**2**), **C**) triphenylphosphine (TPP, **3**), **D**) phosphanetriyltris(benzenesulfonic acid) trisodium salt (TPPTS, **4**), **E**) 2-(diphenylphosphino) benzoic acid (2DPBA, **5**), **F**) 2-(diphenylphosphino)benzamide (2DPBM, **6**). Error bars represent standard deviations from three independent experiments.



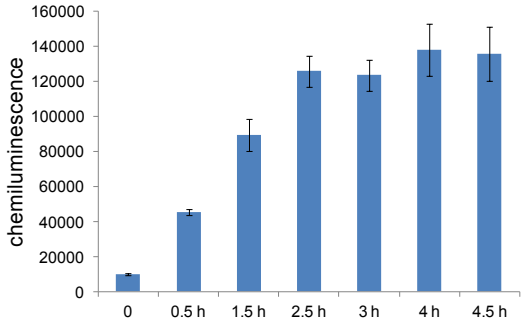
Supplementary Figure S4. Newly constructed plasmids for mammalian cell experiments.



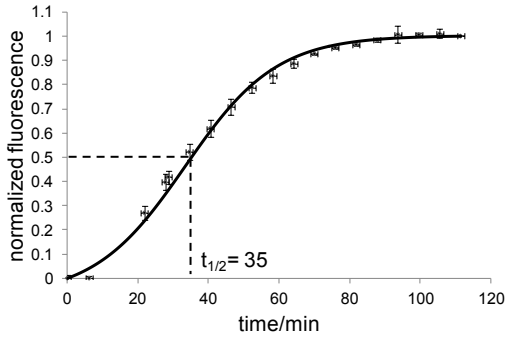
Supplementary Figure S5. A) Staudinger reduction of sfGFP-**OABK** to sfGFP-K. **B**) SDS-PAGE confirmation of protein expression in *E. coli* and reaction with **5**. **C**) ESI-MS analysis of sfGFP-**OABK** and **5**-reduced sfGFP-**OABK**. sfGFP-**OABK** (left): observed MS: 28402.75 Da ± 0.28 Da, expected MS: 28403.03 Da; **5**-reduced sfGFP-**OABK** (right):,observed MS: 28226.41 Da ± 1.48 Da, expected MS: 28227.89 Da.



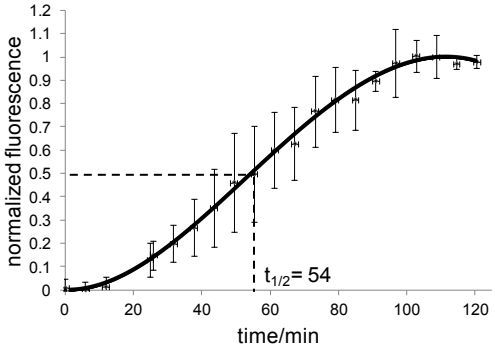
Supplementary Figure S6. Western blot confirmation of **OABK** incorporation into mCherry-TAG-EGFP-HA in mammalian cells.



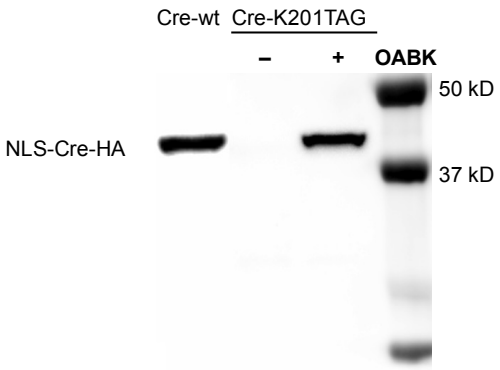
Supplementary Figure S7. Time course of luciferase activation with **5** (0.5 mM) in HEK293T cells. A plateau in enzyme activity is reached after 2.5 h. Error bars represent standard deviations from three independent experiments.



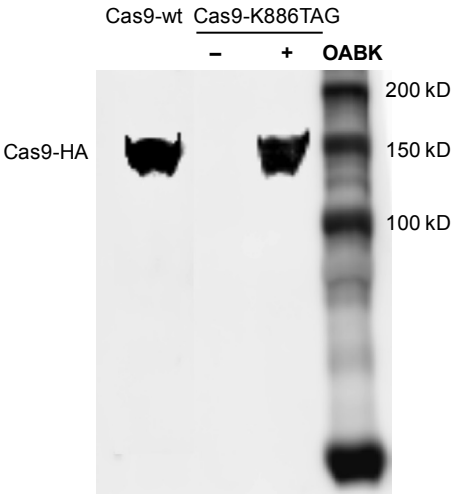
Supplementary Figure S8. Time-course EGFP fluorescence measurements after light stimulation (405 nm, 20 s) of a coumarin-caged EGFP-K85**HCK** mutant. EGFP fluorescence quantification indicates a $t_{1/2}$ of 35 min. Error bars represent standard deviations from fluorescence imaging of three independent cells.



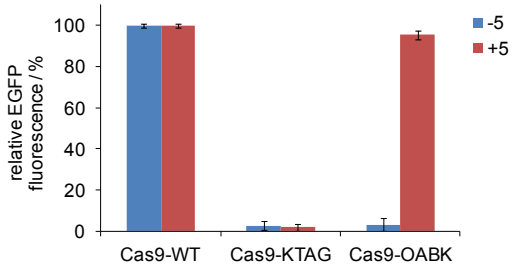
Supplementary Figure S9. Time-course of SATB1 nuclear import after light-stimulation (405 nm, 20 s) of a coumarin-caged SatB1-**HCK**. mCherry fluorescence measurements in the nucleus indicate a $t_{1/2}$ of protein translocation of 54 min. Error bars represent standard deviations from fluorescence imaging of three independent cells.



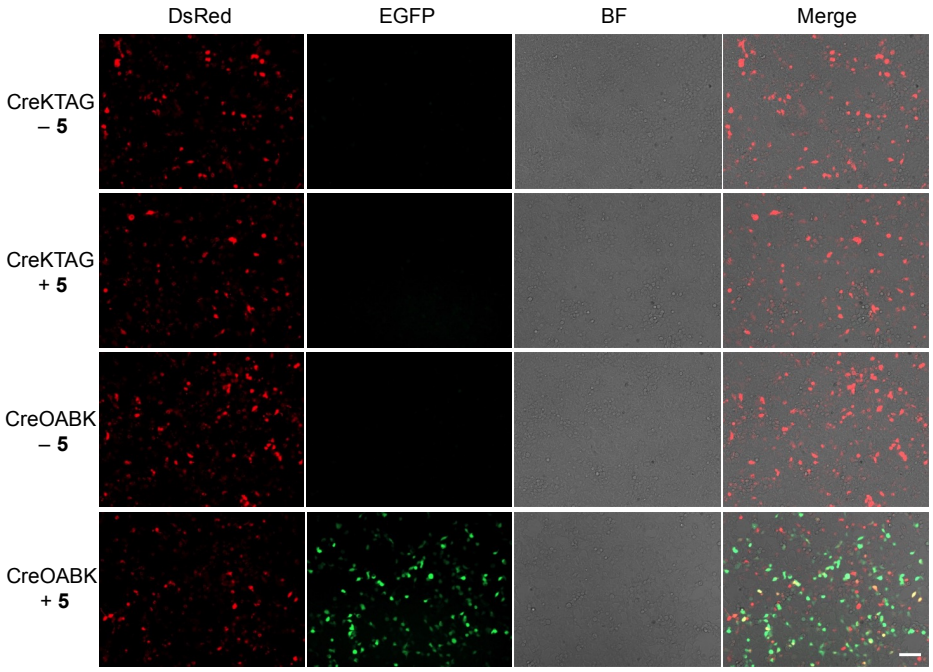
Supplementary Figure S10. Confirmation of **OABK** incorporation into Cre recombinase by Western blot.



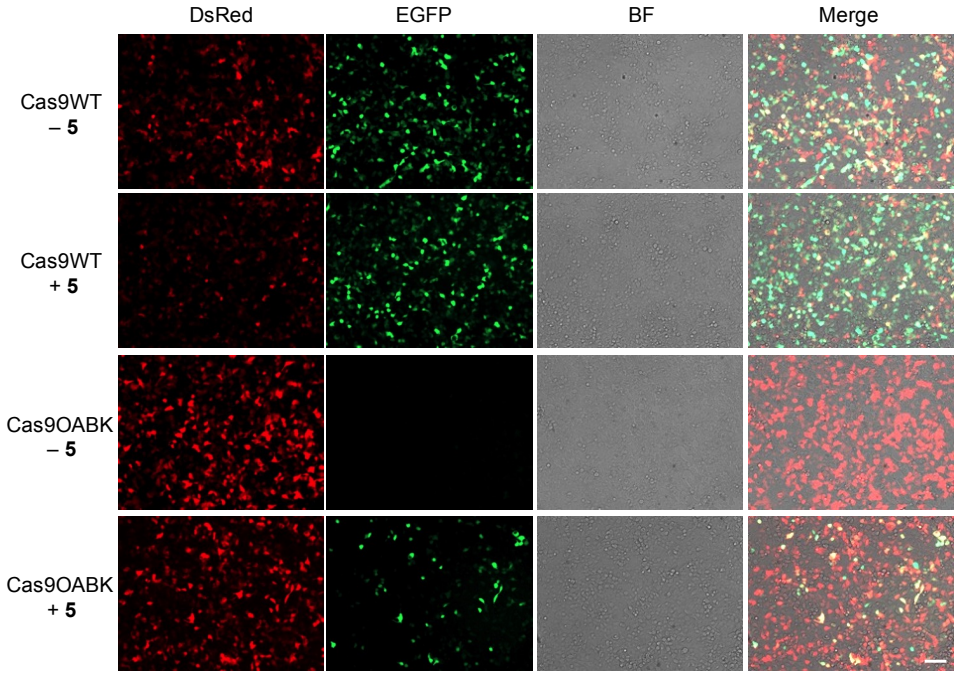
Supplementary Figure S11. Confirmation of **OABK** incorporation into Cas9 by Western blot.



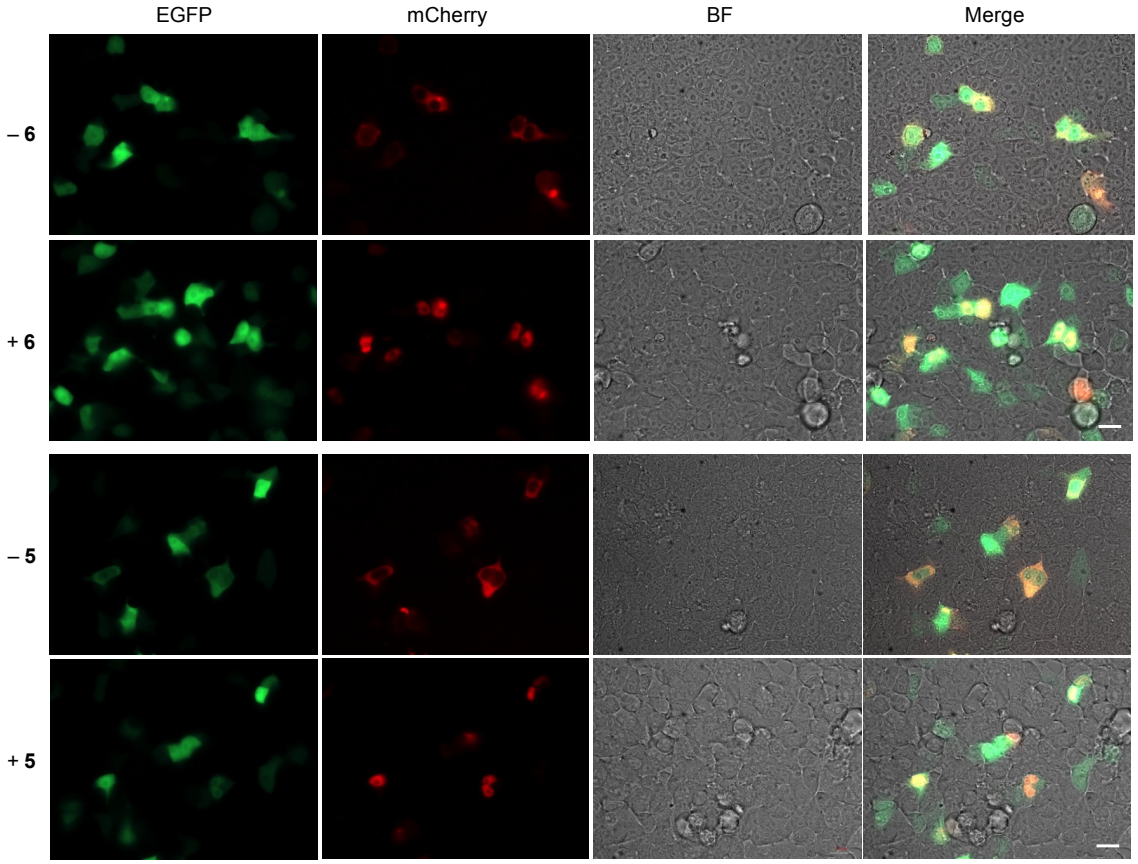
Supplementary Figure S12. Analysis of DsRed and EGFP expression after activation with **5** (500 uM) and comparison to the activity of wild-type (WT) Cas9 by imaging cytometry. Fluorescent cells were counted in the EGFP channel over six randomly selected fields of view per well. Data was analyzed using ImageJ. Error bars represent standard deviations from three independent experiments. Cas9-KTAG = expression in the absence of **OABK**; Cas9-OABK = expression in the presence of **OABK**.



Supplementary Figure S13. Representative brightfield (BF) and fluorescence-merged brightfield images for Cre activation. Scale bar represents 100 μm.



Supplementary Figure S14. Representative brightfield (BF) and fluorescence-merged brightfield images for Cas9 activation. Scale bar represents 100 μm.



Supplementary Figure S15. Representative brightfield (BF) and fluorescence-merged brightfield images for SatB1 activation with **6** (10 μM) and **5** (100 μM). Scale bar represents 20 μm.

References

1 Berridge, M. V., Herst, P. M. & Tan, A. S. Tetrazolium dyes as tools in cell biology: new insights into their cellular reduction. *Biotechnol. Annu. Rev.* **11**, 127-152 (2005).

2 Hemphill, J., Borchardt, E. K., Brown, K., Asokan, A. & Deiters, A. Optical control of CRISPR/Cas9 gene editing. *J. Am. Chem. Soc.* **137**, 5642-5645 (2015).

3 Gautier, A. *et al.* Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc.* **132**, 4086-4088 (2010).

4 Yang, Y. S. & Hughes, T. E. Cre stoplight: A red/green fluorescent reporter of Cre recombinase expression in living cells. *Biotechniques* **31**, 1036, 1038, 1040-1041 (2001).