Supplementary Material (ESI) for Chemical Science

SUPPLEMENTARY INFORMATION

Genetically Encoding an Aliphatic Diazirine for Protein Photocrosslinking

Chungjung Chou[1], Rajendra Uprety[1], Lloyd Davis[2], Jason W. Chin[2,*], and Alexander Deiters[1,*]

[1]North Carolina State University, Department of Chemistry, Raleigh, NC 27695, USA; [2]Medical Research Council Laboratory of Molecular Biology, Hills Rd, Cambridge, CB2 0QH, UK.

alex_deiters@ncsu.edu, chin@mrc-lmb.cam.ac.uk

Protein Expression in *E. coli*

Top 10 competent cells were first transformed with the pMyo4TAGpylT plasmid[1] containing a tetracycline resistance marker, the spermwhale myoglobin gene with an amber stop codon at position 4 serine, and the gene for the pyrrolysyl-tRNA. Then the cells were made competent by the transform and storage solution (TSS) method.[2] The pBKpylS plasmid containing the pyrrolysyl tRNA synthetase was then transformed into these competent cell stocks. A single colony was picked and used to inoculate 2 mL of LB containing 25 µg / mL kanamycin / 50 µg / mL tetracycline. 250 µL of the overnight culture was used to inoculate two 25 mL cultures of 2XYT adjusted to pH = 7 in 500 mL flasks and supplemented with 50 µg / mL kanamycin / 50 µg / mL tetracycline. One culture was supplemented with 1 mM diazirine lysine **1**. The cultures were induced at OD_{600} = 0.6 with 20% arabinose to a final concentration of 0.04% and incubated at 37 °C shaker (250 rpm) for 16 hours. Induced cultures were then transferred and pelleted in 50 mL conical tubes and resuspended with 4 mL lysis buffer containing 20 mM Tris-HCl pH = 7.9, 500 mM NaCl, 10% glycerol and 4 µL protease inhibitor cocktail (Sigma-Aldrich). 40 µL of 10% Triton X-100 was added to the mixture with mild vortexing. The lysate was shaken at 250 rpm on an orbital shaker (New Brunswick) for 15 minutes, sonicated and then centrifuged in 1.5 mL tubes at 4 °C, 13000 g, for 15 minutes. The supernatant was transferred to a 15 mL conical tube and 100 µL of Ni-NTA resin slurry (Qiagen) was added to each tube, and the mixture was incubated at 4 °C for 1 hour on a rocking platform. The resin was pellet by brief a centrifugation, washed twice with 500 µL of lysis buffer, followed by two washes with 300 µL of wash buffer containing 20 mM Tris-HCl pH = 7.9, 500 mM NaCl, and 40 mM imidazole. The protein was then eluted with 400 µL of elution buffer containing 20 mM Tris-HCl pH = 7.9, 500 mM NaCl, and 250 mM imidazole. Eluted sample was mixed with 2X SDS loading buffer, heated

at 95°C for 5 minutes and loaded onto a 12% SDS PAGE gel. Gel electrophoresis was conducted at 60 V for 15 minutes and 150 V for 50 minutes. The gel was then stained overnight with Coomassie blue solution (0.1% Coomassie blue, 10% acetic acid, 40% ethanol) and subsequently destained (10% acetic acid, 40% ethanol). The protein was dialyzed against 1 L of 100 mM NH_4OAc overnight and concentrated with a Microcon filter column (Millipore, 10 kD cutoff). The concentrated protein was then adjusted to 20 μM. The sample was mass analysed using a LC-TOF (Agilent Technologies 6210).

pGSTF52TAGpylT was constructed by polymerase chain reaction using Phusion (NEB) polymerase and pBADGSTF52TAG as the template. Primers were standard pBAD forward (ATGCCATAGCATTTTTATCC) and pBAD reverse (GATTTAATCTGTATCAGG) primers. The PCR product was first treated with T4 polynucleotide kinase in T4 DNA ligase buffer (NEB) to phosphorylate the blunt end generated. The phosphorylated PCR product was then digested with NcoI to generate a 5' NcoI-3' blunt end insert. Backbone pMyo4TAGpylT was first digested with NdeI and then treated with T4 DNA polymerase to crease a blunt end. The treated linearized plasmid was then purified with PCR purification column (Qiagen) and digested with NcoI. The DNA were separated by a 0.7% agarose gel and the ~5.2 kb band was extracted by gel extraction kit (Qiagen). Purified insert was ligated to the backbone and transformed into GC5 ultracompetent cells (Genesee) and selected on LB agar supplemented with 50 μg / mL of tetracycline. Plasmids were isolated using a miniprep kit (Qiagen), restriction mapped, and sequenced to confirm the coding sequence.

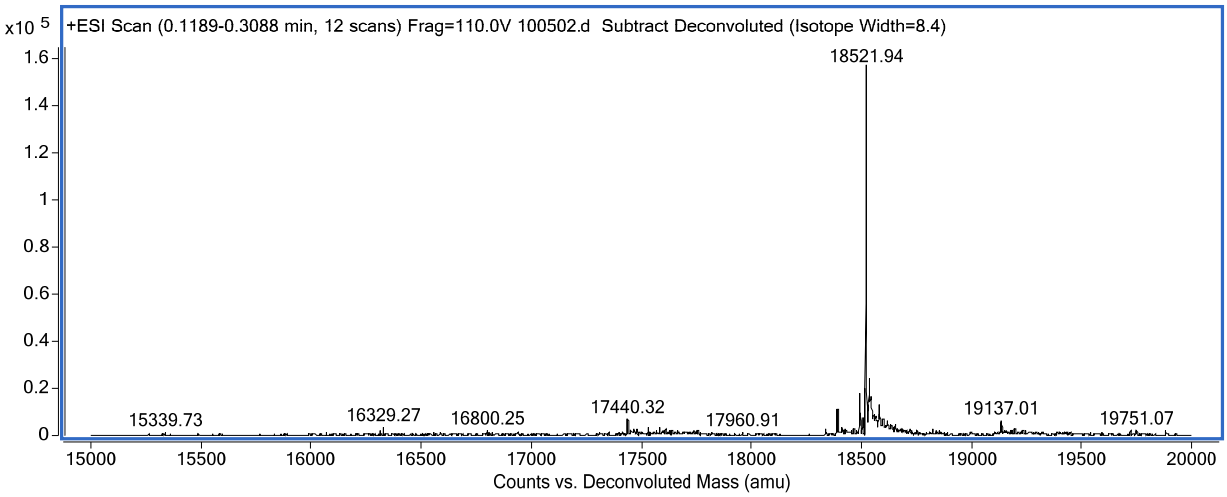
pGSTF52TAGpylT were co-transformed with pBKpylS into Top10 supercompetent cells (Invitrogen) and inoculated as described above. The cells were induced at OD = 0.6 with 0.04% arabinose and placed onto a 250 rpm shaker at room temperature overnight. For *in vitro* photocrosslinking, the harvested culture was lysed in ice cold lysis buffer (sodium phosphate pH = 8, 300 mM NaCl, 10 mM imidazole) with 1 mg / mL lysozyme, protease inhibitor cocktail (Sigma Aldrich), and 0.1% Triton X100. The mixture was incubated on ice for 1 hour and then sonicated on ice. Cell lysate was centrifuged and His-tagged protein was purified by addition of Ni-NTA slurry as described above. The resin was washed one time with lysis buffer and three times with wash buffer (sodium phosphate pH = 8, 300 mM NaCl, 20 mM imidazole) and eluted with elution buffer (sodium phosphate pH = 8, 300 mM NaCl, 250 mM imidazole). The eluted sample was kept in the dark or irradiated with 365 nm on a UV transluminator for 20 minutes before mixed with equal volume of 2X loading dye and analyzed by 8% SDS PAGE. One of two

identical gels was stained and the other was wet-blotted against PVDF membrane (Biorad, 0.2 µm pore size). The membrane was rinsed with TBST buffer (20mM Tris-HCl pH=7.6, 137mM NaCl, 0.1% Tween 20) and blocked with TBST/5% nonfat milk powder. Anti-His tag mouse antibody (Biorad) was diluted in TBST to 1/1500 and used to incubate the membrane at 4 °C overnight. The membrane was washed with TBST for 5 minutes three times and incubated with a secondary goat anti-mouse HRP conjugated antibody (dilution ratio 1/1000) (Biorad) for 1 hour at room temperature. The membrane was then washed with TBS buffer and 4CN colorimetric reagent (Biorad) was used to detect the tagged protein on the membrane. The membrane was rinsed briefly with DI water and scanned.

In vivo photocrosslinking was carried out by washing harvested cells with 10 mL PBS and then re-suspending pellets in 50 mL of PBS. The culture was then poured into a 150 mm petri dish and placed on a transiluminator for 10 and 20 minutes (365 nm). 5 mL of suspended cells were harvested and PBS removed. 250µL of 1X SDS loading buffer was added to re-suspend the pellet. The sample was heated at 95 °C for 10 minutes before loading to SDS PAGE gel. The residual cells were stored in -80°C and subjected to the same purification and immunoblotting steps as described above.

MS Analysis of Myo-1

Deconvoluted spectrum of electrospray ionization mass spectrum of myoglobin with **1** incorporated at position 4. The protein has an expected molecular weight of 18522.26 Da and a peak at 18521.94 Da confirms its presence.



Mammalian Cell Culture

Human Embryonic Kidney 293 cells (HEK293) were cultured in DMEM supplemented with 10% Fetal Bovine Serum and 1X Penicillin/Streptomycin. The cells were grown at 5% carbon dioxide.

Mammalian Cell Transfection

Cells were seeded in 24 well plates and transfected with one vector containing the MmPylS and the mCherry-eGFP-HA repoter, and another with the MmPyl tRNA. The cells were transfected using the GeneJuice (Novagen) transfection reagent using the manufacturer's protocol. At the same time as transfection the media was removed from the wells and replaced with antibiotic free medium supplemented with 1 mM of **1**. Some cells were transfected and grown without amino acid as a negative control. Other cells were transfected with the synthetase-reporter construct and/or containing human Tyrosyl tRNA as a positive control.

Mammalian Cell Lysis and Western Blotting

After 20 hours cells were washed in ice cold 1XPBS and then lysed in 125 ul universal lysis buffer (Roche) for 10 minutes, this was carried out on ice. Lysate was then extracted from the wells and spun down to pellet cell debris, 90 ul of supernatant was then removed and added to 30ul of 4x LDS Sample Buffer (Invitrogen). Western blotting was done using anti-HA (Sigma) and anti-FLAG (Cell Signalling) antibodies.

Mammalian Cell Imaging

Cells were seeded and transfected on 24 well plates as described above. Laser-scanning confocal microscopy was performed using a Bio-Rad Radiance 2100 system mounted on a Nikon Eclipse TE300 inverted microscope with a PanFluor 20x/0.45 objective. Fluorescence emission from 515-530 nm was measured for eGFP with excitation at 488 nm, and for mCherry fluorescence emission was measured above 560 nm with excitation at 543nm.

Synthesis of the Diazirine Amino Acid 1

2-(3-Methyl-3*H*-diazirin-3-yl)-ethanol (3). 4-hydroxy-2-butanone (**2**, 10 g, 0.11 mol) was added slowly into stirring liquid NH_3 (60 mL) at –78 ºC. After stirring for 5 h at –78 ºC, hydroxylamine *O*-sulfonic acid (14.1 g, 0.12 mol) dissolved in methanol (100 mL) was added slowly to the reaction mixture. Then the reaction mixture was allowed to warm up to room temperature and

was stirred overnight. The resulting white precipitate was filtered off and the volume of the filtrate was reduced to 100 mL. The filtrate was cooled to 0 ºC and methanol (100 mL) and Et_3N (15 mL) were added followed by slow addition of iodine (~14 g) until the iodine color was persistent. The reaction mixture was then allowed to warm up to room temperature and was stirred for 2 h. The volume of the reaction mixture was then reduced to 100 mL under vacuum. The reaction mixture was diluted with brine (200 mL) followed by extraction with ether (3 x 20 mL). The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and volume was reduced to 10 mL. Vacuum distillation (65ºC, 9.0 mmHg) was performed to obtain **3** (4.5 g, 39% overall yield) as a yellow oil.[3]

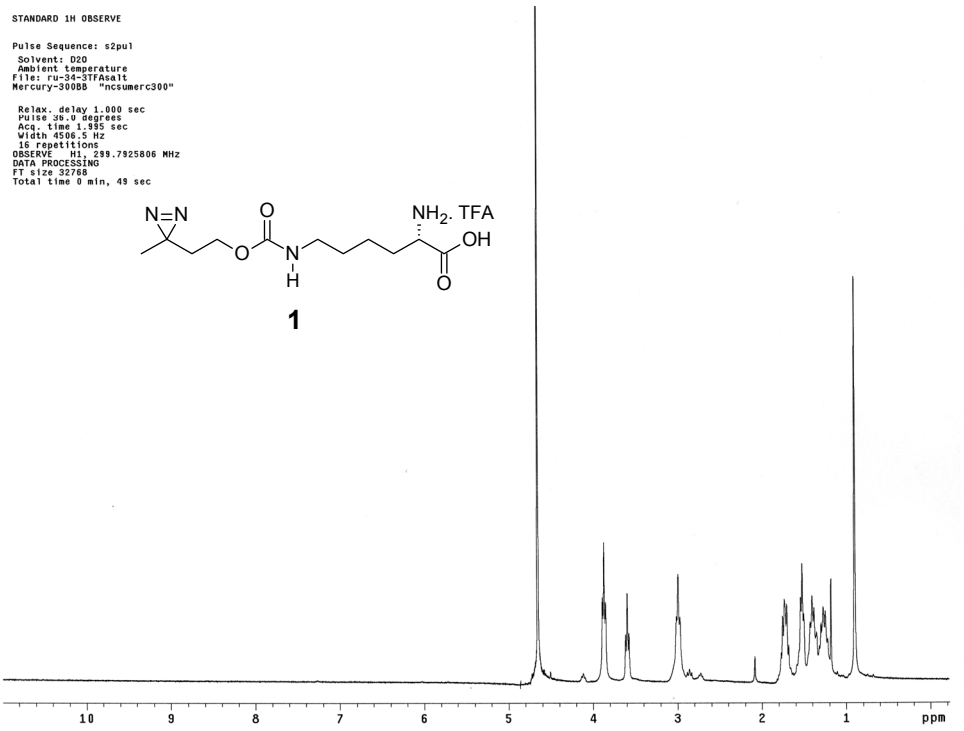
1-({[2-(3-methyl-3*H*-diaziren-3-yl)ethoxy]carbonyl}oxy)pyrrolidine-2,5-dione (4). The alcohol **3** (770 mg, 7.7 mmol) was dissolved in dry acetonitrile (20 mL) under an argon atmosphere. *N*,*N*′-disuccinimidyl carbonate (2.95 g, 11.5 mmol) and Et_3N (4 mL) were added to the reaction mixture and the solution was stirred for 16 h at room temperature. The volatilities were evaporated and the residue was purified by column chromatography using an acetone/chloroform mixture (20:1) over silica gel to obtain **4** (2.5 g, 83%) as a white solid. ^{1}HNMR (400 MHz, $CDCl_3$) δ = 4.17 (t, *J* = 6.4, 2H, -OCH_2), 3.25 (s, 4H, -CH_2), 1.66 (t, *J* = 6.4, 2H, -CH_2), 0.99 (s, 3H, CH_3). ^{13}CNMR (100 MHz, $CDCl_3$) δ = 168.9 (-CON), 151.4 (-OCOO), 66.5 (-CNN), 33.8 (-OCH_2), 25.6 (-CH_2), 19.9 (-CH_2). LRMS (ESI^+) calcd for $C_9H_{11}N_3O_5$ $(M+Na)^+$ 264.0596, found 264.1000.

(2S)-2-(*tert*-Butoxycarbonylamino)-6-({[2-(3-methyl-3*H*-diaziren-3-yl)ethoxy]carbonyl} amino)hexanoic acid (5). The NHS-carbonate **4** (1.8 g, 7.4 mmol) was dissolved in dry DMF (25 mL) under argon atmosphere. Boc-Lys-OH (2.7 g, 11.2 mmol) was added to the reaction mixture and the solution was stirred for 30 h at room temperature. The reaction mixture was poured into water (100 mL) and stirred for 10 minutes. The product was extracted with ether (3 x 20 mL); the combined organic layers were washed with brine (2 x 15 mL) and dried over anhydrous Na_2SO_4. The solvent was removed and the product was dried under *vacuo* to obtain **5** (2.4 g, 88%) as a colorless, viscous liquid. ^{1}HNMR (300 MHz, $CDCl_3$) δ = 10.30 (s, br, 1H, -OH), 6.31 (s, br, 0.5 -NH), 5.21 (s, 1H, -NH), 4.89 (s, br, 0.5 H,-NH), 4.24 (s, br, 1H, -CH), 3.99-3.92 (m, 2H, -OCH_2), 3.11 (m, br 2H, -NCH_2), 1.77-1.37 (m, 17H, -CH_2, $C(CH_3)_3$), 0.98 (s, 3H, -CH_3); ^{13}CNMR (100 MHz, $CDCl_3$) δ = 176.6 (-COOH), 156.7 (–CONH), 156.0 (-CONH), 80.2 (-$C(CH_3)_3$, 60.1 (-CNN), 53.3 (-CH), 40.7 (-OCH_2), 34.3 (-CH_2), 32.2 (-CH_2), 29.4 (CH_2), 28.5

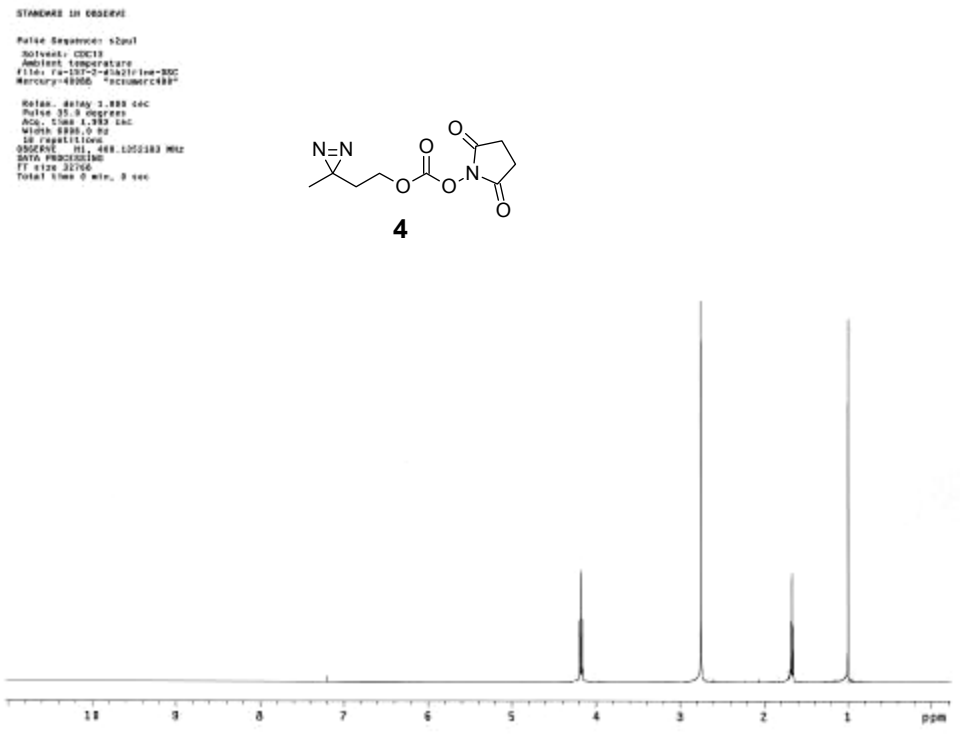
$(CH_3)_3C$, 22.5 ($-CH_3$), 19.9 ($-CH_2$). LRMS (ESI^+) calcd for $C_{16}H_{28}N_4O_6$ $(M+Na)^+$ 395.1906, found 395.1920.

(2S)-2-amino-6-({[2-(3-methyl-3*H*-diaziren-3-yl)ethoxy]carbonyl}amino)hexanoic acid TFA salt, (1): The Boc-amine **5** (600 mg, 1.6 mmol) and triethylsilane (513 µL, 3.2 mmol) were added to a flame dried flask under an argon atmosphere. A mixture of 5% TFA in dichloroethane (25 mL) was slowly added to dissolve **5** at room temperature and the solution was stirred for 12 h. Volatile components were removed by rotavap, and the obtained residue was dissolved in MeOH (1.5 mL) and precipitated into ether (100 mL) through dropwise addition while stirring vigorously. The precipitation process was performed twice to obtain **1** (472 mg, 76%) as a white crystalline solid. ^{1}HNMR (300 MHz, D_2O) δ = 3.98 (t, *J* = 5.7, 2H, $-OCH_2$), 3.70 (t, *J* = 6.0, 1H, -CH), 3.10 (t, *J* = 6.0, 2H, $-NCH_2$), 1.75-1.70 (m, 2H, $-CH_2$), 1.53-1.50 (t, *J* = 5.7, 2H, $-CH_2$), 1.42-1.32 (m, 2H, $-CH_2$), 1.29-1.24 (m, 2H, $-CH_2$), 1.03 (s, 3H, $-CH_3$). ^{13}CNMR (75 MHz, D_2O) δ = 174.2 (-COOH), 158.6 (-CONH), 60.6 (-CNN), 54.6 (-CH), 40.0 ($-OCH_2$), 33.4 ($-NCH_2$), 30.1 ($-CH_2$), 28.7 ($-CH_2$), 26.5 ($-CH_2$), 23.3 ($-CH_3$), 21.7 ($-CH_2$),. LRMS (ESI^+) calcd for $C_{11}H_{20}N_4O_4$ $(M+H)^+$ 273.3089, found 273.1573.

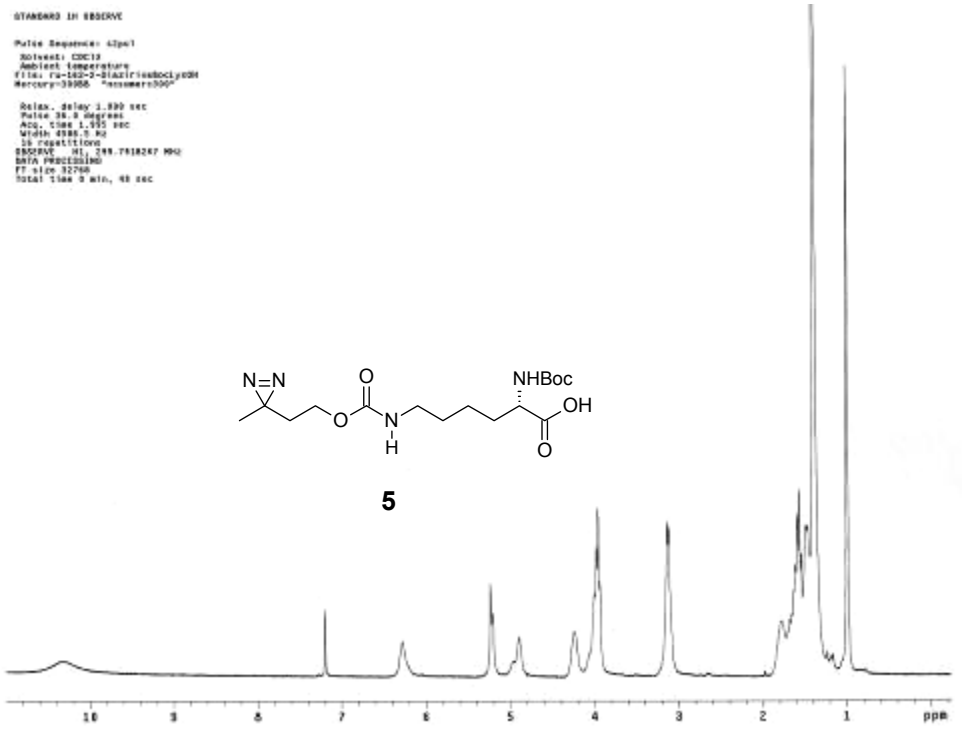
Proton NMR Spectra of Compounds 1, 4, and 5



1



4



5

[1] Neumann, H., Peak-Chew, S. Y., Chin, J. W. *Nat. Chem. Biol.* **2008**, *4*, 232-234.

[2] Chung, C. T., Niemela, S. L., Miller, R. H. *PNAS* **1989**, *86*, 2172-2175.

[3] a) Church R. F. R., Weiss M. J. *J. Org. Chem.* **1970**, *35*, 2465-2471. b) Shigdel U. K., Zhang J., He C., *Angew. Chem. Ind. Ed*. **2008**, *47*, 90-93.