# A bifunctional non-canonical amino acid: synthesis, expression and residue-specific proteome-wide incorporation

Jan-Erik Hoffmann (1), Dmytro Dziuba (2), Frank Stein (2) and Carsten Schultz (1,2)

(1) Oregon Health & Science University, Department of Physiology & Pharmacology, L334, 3181 SW Sam Jackson Park Rd, Portland, OR 97239-3098 (USA)

(2) European Molecular Biology Laboratory, Cell Biology & Biophysics Unit, Meyerhofstr.1, 69117 Heidelberg (Germany)

Corresponding author Carsten Schultz; schulcar@ohsu.edu

**Supplementary Methods**

**Incorporation of non-canonical amino acids in bacterial proteins**

Test expressions were performed with MBP-GFP$^{39TAG}$-6His as a target protein for incorporation of the non-canonical amino acids. The plasmids pBAD-MBP-GFP$^{39TAG}$-6His and pEvol-MmPylRS$^{Y306A,Y384F}$-PylT$_{CUA}$[14] were transformed into electro competent *E. coli* BL21 DE3 cells and grown in TB medium containing 100 µg/ml ampicillin and 30 µg/ml chloramphenicol for selection. An overnight culture was inoculated by picking a single colony from an agar plate or directly from a DMSO stock.

The next day, a 5 ml (test expression) or 50 ml (medium scale expression) culture was inoculated 1:50 with the overnight culture and incubated at 37°C in a shaker for about 2 h until it reached an optical density at 600 nm of 0.3. The non-canonical amino acid was dissolved in 0.1 M NaOH at a concentration of 80 mM and added to the expression culture 1:80 for a final concentration of 1 mM amino acid. In control experiments, the corresponding amount of 0.1 M NaOH was added. After approximately 1 h further incubation at $OD_{600}$=0.5 the expression was induced by addition of 0.02 % arabinose (1:1000 from a 20 % stock solution in water). The expression culture was shaken over night at 37°C.

Cells were harvested by centrifugation at 4000 rpm and 4°C. All following steps were conducted on ice or in a cold room with pre-chilled buffers. Cell pellets were resuspended in 0.5 ml (test expression) or 1 ml (medium scale expression) lysis buffer (PBS with 5 mM imidazole and 1 mM PMSF or commercial protease inhibitor) and the cells were transferred into 2 mL reaction tubes and lysed by sonication (30*1 s pulses). Cell debris was separated by centrifugation (14000 rpm, 4°C, 30 min) and discarded. 25 µL (test expression) or 50 µL (medium scale expression) Ni-NTA agarose bead suspension (Qiagen) were washed 1x with water, 1x with lysis buffer and then resuspended in lysis buffer and incubated with the supernatant of the cell lysis for 3 h at 4°C under gentle agitation. After this binding step, the

beads were collected in centrifuge filter columns (test expression; 1000 rpm) or gravity filter columns (medium scale expression) and washed three times with washing buffer (PBS with 10 mM imidazole). The washing volume for each step was 300 µL or 3 mL, respectively. The flowthrough was discarded. Finally, the protein was eluted from the beads in two steps with 100 + 50 µL (test expression) or 200 +100 µL (medium scale expression) after 1 min incubation for each step in elution buffer (PBS with 500 mM imidazole). The purity of the eluate was confirmed by polyacrylamide gel electrophoresis.

***In vitro* crosslinking of purified proteins**

Crosslinking was performed with a 1000 W HXR-R1 Mercury/Xenon lamp (Newport Spectra-Physics GmbH, Darmstadt, Germany) with a 345 nm longpass filter (20CGA-345, Newport). The lamp was switched on for 30 min until the power output stabilized at 1000 W and the cone of light was centered. 10-20 µL of concentrated protein solution in elution buffer or PBS was placed in open microcentrifuge tubes in metal blocks on ice so that they were directly in the center of the 10 cm wide light cone. The sample was irradiated for 5 min and analyzed by gel electrophoresis. Strong eye protection is required for this experiment. Ideally, the lamp is mounted in a way that direct eye contact is prevented.

**Click chemistry and gel electrophoresis**

Purified proteins or whole extracts in PBS were incubated with 1 mM sodium ascorbate (from 10 mM stock), 100 µM TBTA (from 1 mM stock), 1 mM $CuSO_4$ (from 10 mM stock) and 2 µM TAMRA-azide (from 20 µM stock; Azide-fluor 545, Sigma-Aldrich #760757) for 3 h at 37°C. The samples were supplemented with 4x sample buffer and heated at 95°C for 10 min. Gel electrophoresis was performed using Novex 4-12% Bis-Tris gels and MES running buffer (ThermoFisher Scientific, Waltham, MA, USA) at 150 V. Fluorescent gel were imaged before fixation, then stained with Coomassie blue in 20% methanol.

**Cell culture and microscopy**

HEK293T and HeLa Kyoto cells were cultured in DMEM with 4.5 g/l glucose and 10% FBS and split three times a week. For microscopy, cells were seeded in 8 well chambered 170 µm coverglass (iBiDi, Munich, Germany); 15000 cells per well in complete medium. After 24 h they were transfected with Lipofectamine 2000 (ThermoFisher Scientific) and 100 nmol each of the $tRNA^{pyl}$/$pylRS^{Y306A,\ Y384F}$ plasmid[41] and the respective Amber suppression construct. Shortly before the transfection mixture was added to cells, 500 µM ncAA was added from a 80 mM Stock in 0.1 M NaOH. After overnight incubation the transfection medium was replaced with fresh full growth medium and incubated for another night. The growth medium was exchanged for live cell imaging solution (ThermoFisher Scientific) and the cells were imaged on an Olympus FluoView 1200 confocal microscope inside an incubation chamber at 37°C. The eGFP constructs were excited at 488 nm and fluorescence was collected between 500 and 550 nm.

**SORT labeling with PrDiAzK in bacteria**

*Plasmid constructs.* The plasmid pEvol-$MmPylRS^{Y306A,Y384F}$-$PylT_{CUA}$ was mutagenized using primers targeting the anticodon loop of the Pyl-tRNA gene. The CUA anticodon sequence was replaced by anticodons that decode corresponding sense codons of Met, Arg, Val, Tyr and Lys (CAT, CCT, CAC, GTA and CTT, respectively). The sequences of the obtained plasmids pEvol-$MmPylRS^{Y306A,Y384F}$-$PylT_{xxx}$ were confirmed using the Sanger sequencing.

*E.coli for SORT experiments*. Chemically competent *E.coli* XL-1 Blue cells were transformed with the pEvol-$MmPylRS^{AF}$-$PylT_{xxx}$ constructs. The cells were recovered in SOC medium for 1 h at 37°C and then aliquots were plated on LB agar plates supplemented with 33 µg/ml of chloramphenicol and incubated overnight in an incubator at 37°C. Single

colonies from each plate were used to inoculate 5 mL of LB medium containing 25 µg/ml of chloramphenicol and the precultures were grown overnight at 37°C.

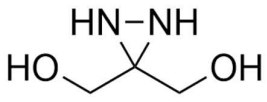
*Proteome labeling in E.coli using SORT.* 200 µL of overnight *E. coli* preculture were used to inoculate 5 ml of LB medium containing 25 µg/ml of chloramphenicol and optionally 0.1 mM of PrDiAzK. The bacteria were incubated with shaking (37°C, 200 rpm). At $OD_{600} \approx$ 0.3, L-arabinose (2.5 µL, 20% w/v solution in water) was added to the suspension and then the tubes were incubated overnight in the incubator with shaking (37°C, 200 rpm). The cells were harvested by centrifugation (4000 rpm, 30 min) and pellets were washed with ice-cold PBS. The pellets were resuspended in 0.9 ml of PBS and lysed by sonication (4°C, Branson Sonifier 250, output control 1, duty cycle 50%, 1 min). The lysates were cleared by centrifugation (4°C, 14000 rpm, 15 min) and the supernatants were aspirated into new tubes.

*Labeling of PrDiAzK-containing proteins using Cu-catalyzed azide-alkyne cycloaddition (CuAAC)*. The following stock solutions were prepared and stored at –20°C: 25 mM ascorbic acid in $H_2O$; 2.5 mM tris(benzyltriazolylmethyl)amine (TBTA) in DMSO; 25 mM $CuSO_4$ in $H_2O$; 2 mM Azide-fluor 545 (TAMRA-PEG-azide, SigmaAldrich) in DMSO. For the click labeling, a click mix was freshly prepared by mixing equal volumes of PBS, TBTA, $CuSO_4$, ascorbic acid and Azide-fluor 545. For negative controls, the same mixture lacking $CuSO_4$ was prepared. The click labeling reaction was initiated by mixing 20 µL of the bacterial lysate and 5 µL of the click mix. The reaction mixtures were incubated at 23 °C for 1 hour, than 25 µL of NuPAGE LDS Sample Buffer (2x) were added and the proteins were denatured by heating at 75 °C for 10 min. Aliquots (1 µL) were analyzed by SDS-PAGE (precast NuPAGE 4-12% Bis-Tris protein gels, MOPS SDS running buffer). PageRuler Prestained Protein Ladder, 10 to 180 kDa (ThermoFisher) was used as a reference. The gel was washed with water multiple times on a rocker at room temperature to wash out unreacted dye. The in-gel fluorescence of TAMRA-labeled proteins was recorded using a Typhoon FLA 9500 imager.

The gel was further stained with Coomassie blue, destained and then imaged using a Bio-Rad ChemiDoc Touch Imaging System.

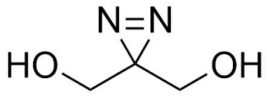
**Synthesis of PrDiAzK**

Diaziridine-3,3-diyldimethanol (**6**):




This synthesis step was adapted from a protocol by Shigdel et al., 2008[43]. Dry ammonia gas was condensed in a three-way flask at -78°C in a dry ice/acetone bath. 6.35 g dihydroxyacetone (**5**; 70.5 mmol) were dissolved in 70 ml dry methanol and added to about 200 ml of liquid ammonia and stirred for 3 h. Hydroxylamine O-sulfonic acid (9.16 g, 81 mmol, 1.1 eq) was dissolved in 60 ml dry methanol and added drop wise to the mixture over the course of 1 h. During this time an overpressure of ammonia was maintained and the reaction mixture cooled to -78°C. After that the mixture was stirred over night while the dry ice bath evaporated and the mixture slowly reached room temperature. On the next day the remaining ammonia was removed by bubbling air through the mixture, the white precipitate was filtered off and the methanol was evaporated to obtain **6**.

(3H-Diazirine-3,3-diyl)dimethanol (**7**):

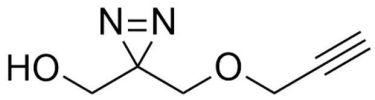



Without further purification the diaziridine was oxidized to diazirine. For this 50 ml of methanol and 14.8 ml (106 mmol) triethylamine were added to the crude intermediate **6** at 0°C. Under stirring 8 g (32 mmol) of solid iodine were added slowly until the mixture stayed red for over 5 min. The reaction was stirred for an additional 2 h, after which the methanol was evaporated and the mixture was extracted with ethyl acetate and dried over $Na_2SO_4$. The

solvent was evaporated and the crude product purified by flash chromatography with CyH/EtOAc 3:1 to obtain 1.06 g of **7** as colorless oil. (10.4 mmol; 15% yield over two steps).

$^1$H NMR (400 MHz, MeOD) δ 4.88 (s, 2H), 3.37 (s, 4H).

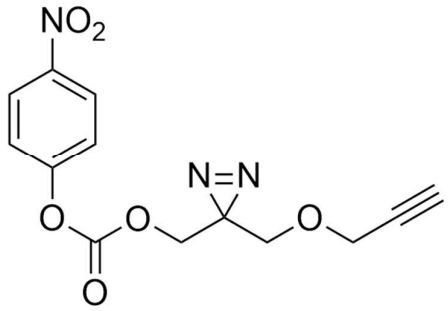
(3-((Prop-2-yn-1-yloxy)methyl)-3H-diazirin-3-yl)methanol (**8**):




3.37 g of **7** (33 mmol) were dissolved in 78 ml anhydrous THF. 832 mg (34.7 mmol) sodium hydride were added and the mixtures was stirred at 0°C for 30 min. Then, propargyl bromide 80% in toluene (3.57 mL; 32 mmol) was added drop wise and the resulting mixture stirred at room temperature. After 6h, 5 mL methanol were added, the reaction mixture was evaporated at reduces pressure and the residue was dissolved in ethyl acetate. The organic layer was washed with saturated ammonium chloride. The aqueous layer was extracted two times with ethyl acetate and the combined organic layers were evaporated under vacuum. The dry residue (2.8 g) was purified by flash chromatography (Ethyl acetate/cyclohexane 1:4) to obtain 520 mg propargyl-diazirine-alcohol 7 (3.71 mmol; 11.23 % yield) as a colorless oil. In addition to the desired product, 63.4 mg of double substituted side product were isolated and 1.98 g (59%) starting material was recovered.

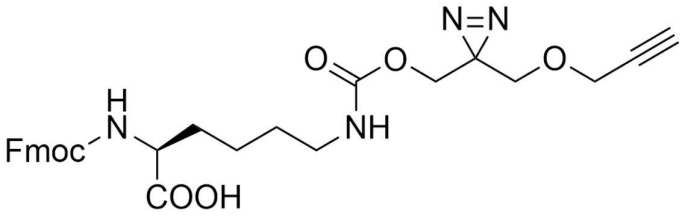
$^1$H NMR (400 MHz, $CDCl_3$) δ 4.19 (d, $J$ = 2.2 Hz, 2H), 3.57 (s, 2H), 3.47 (s, 2H), 2.49 (t, $J$ = 2.2 Hz, 1H), 1.88 (s, 1H). HR-ESI (pos. mode)= 167.0677 g/mol (calc. $+Na^+$: 167.0684 g/mol).

Propargyl-diazirine active ester [4-nitrophenyl ((3-((prop-2-yn-1-yloxy)methyl)-3H-diazirin-3-yl)methyl) carbonate] (**9**):



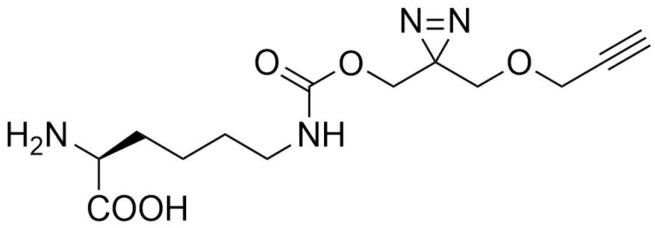
The following steps were adapted from Hoffmann et al., 2015[45]. Under Argon, 110 mg (1 mmol) of propargyl-diazirine-alcohol **8** were dissolved in 8 mL anhydrous DCM and pyridine (0.19 mL, 2.5 eq.) under argon and stirred at RT. 214 mg (1.1 eq) para-nitrophenyl chloroformate dissolved in 8 ml anhyd. DCM under argon were added and the reaction was stirred for 2 h. Then, the mixture was quenched with sat. $NH_4Cl$, extracted two times with DCM, washed with brine and dried over $Na_2SO_4$. The solvent was evaporated to yield 220 mg (0.7 mmol, 70% yield) of **9** as a yellow solid.

N2-(((9H-Fluoren-9-yl)methoxy)carbonyl)-N6-(((3-((prop-2-yn-1-yloxy)methyl)-3H-diazirin-3-yl)methoxy)carbonyl)-L-lysine (**10**)



Without further purification, 220 mg (0.7 mmol) of **9** were dissolved in 10 ml anhydrous DMSO under argon atmosphere. 350 mg (0.86 mmol; 1.2 eq.) Fmoc-lysine-OH were dissolved in DIEA (0.19 ml; 1.1 mmol; 1.5 eq.) and anhydrous DMSO under an argon atmosphere and stirred at room temperature. The solution of **9** was added drop wise over the course of 2 h. After stirring the mixture over night at room temperature, H2O and ethyl acetate were added, the aqueous phase adjusted to pH 3 with HCl and extracted with ethyl acetate. The solvent was evaporated to yield **6**.

PrDiAzK (**1**):




The crude protected amino acid was dissolved in DCM with 20% piperidine *v/v* and stirred for 30 min at RT. Water was added and the solvent evaporated. The product was purified by flash chromatography with acetone/MeOH/$H_2O$ 65/25/10 *v/v/v* to yield 155 mg (0.5 mmol) of PrDiAzKs **1** (70% yield over two steps).

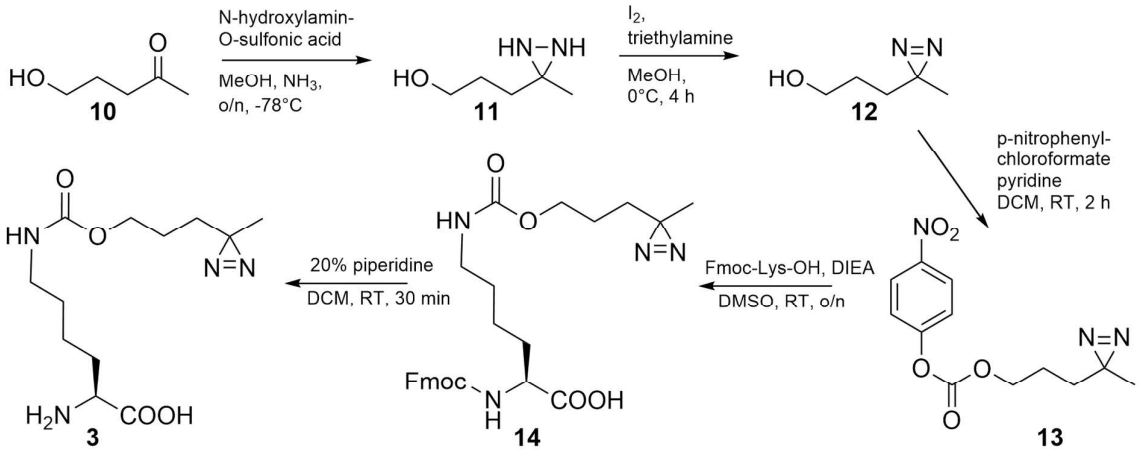
$^{1}$H NMR (400 MHz, MeOD) δ 7.90 (s, 1H), 4.15 (d, *J* = 2.4 Hz, 2H), 3.96 (s, 2H), 3.52 (dd, *J* = 6.9, 5.2 Hz, 1H), 3.47 (s, 2H), 3.11 (t, *J* = 6.7 Hz, 2H), 2.88 (t, *J* = 2.3 Hz, 1H), 1.85 (dt, *J* = 22.4, 11.4, 4.5 Hz, 2H), 1.52 (m, 2H), 1.45 (m, 2H).

$^{13}$C NMR (MeOD) δ 174.63, 79.88, 76.66, 69.81, 64.84, 58.89, 56.23, 41.55, 32.12, 30.60, 28.29, 23.55.

HR-ESI (neg. mode)= 311.13585 g/mol (calc. $-H^+$: 311.13609 g/mol). HR-ESI (pos. mode) = 313.15156 g/mol (calc. $+H^+$: 313.15065 g/mol).

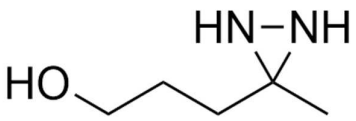
**Synthesis of DiAzK and DiAzKs**

The short version DiAzKs (**2**) was first reported by Hancock et al.[42] and later used by Koehler et al.[11]. The synthesis was modified to yield an elongated variation with an additional $CH_2$-group, called DiAzK (**3**).



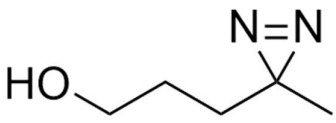

**Scheme S1.** Synthesis scheme of DiAzK

Diaziridine-pentanol [3-(3-methyldiaziridin-3-yl)propan-1-ol] (**12**):



The conversion of the keton 5-hydroxypentan-2-one (**11**) into the diazirine was done analogous to the procedure described for compound **5**. 10 ml (98.6 mmol) 5-hydroxypentan-2-one were dissolved in 100 ml liquid $NH_3$ and 12.8 g (113.4 mmol) hydroxylamine O-sulfonic acid in 50 ml ethanol were added drop wise. After stirring over night the product was filtered and the solvent evaporated.

Diazirine-pentanol [3-(3-methyl-3H-diazirin-3-yl)propan-1-ol] (**13**):

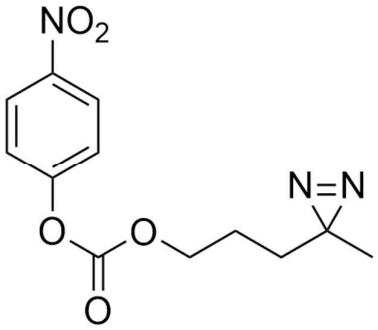


The crude diaziridine was dissolved in 50 ml methanol and 30 ml triethylamine at 0°C. 20 g iodine were added slowly and the mixture stirred at air for 2 h. The solvent was evaporated and the mixture was extracted with ethyl acetate and dried over $Na_2SO_4$. The solvent was

evaporated and the crude product purified by flash chromatography with cyclohexane/ethyl acetate 9:1 (RF=0.6) to obtain 1.89 g of **13** (16.6 mmol; 17% yield).

$^1$H NMR (400 MHz, $CDCl_3$) δ = 3.95 – 3.84 (m, 2H), 2.01 (d, *J* = 8.9 Hz, 2H), 1.43 – 1.40 (m, 2H), 1.03 (s, 3H).

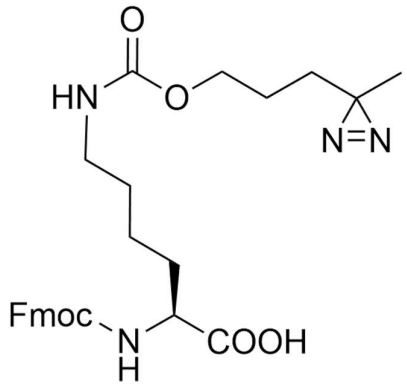
Diazirine active ester [3-(3-methyl-3H-diazirin-3-yl)propyl (4-nitrophenyl) carbonate] (**14**):




1.89 g (16.6 mmol) of diazirine-pentanol **13** were dissolved in 20 ml anhydrous DCM and pyridine (3.34 ml, 2.5 eq.) under argon and stirred at room temperature. 3.67 g (1.1 eq) para-nitrophenyl chloroformate dissolved in 30 ml anhydrous DCM under argon were added and the reaction was stirred for 2 h. The mixture was quenched with sat. $NH_4Cl$, extracted two times with DCM, washed with brine and dried over $Na_2SO_4$. The solvent was evaporated to yield 5.8 g of crude **14** as a yellow wax.

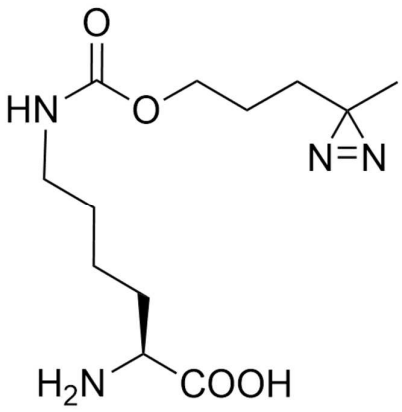
$^1$H NMR (400 MHz, $CDCl_3$) δ 8.37 – 8.24 (m, 2H), 7.34 – 7.27 (m, 2H), 4.29 (t, *J* = 6.4 Hz, 2H), 1.67 (dt, *J* = 9.0, 6.3 Hz, 2H), 1.59 – 1.49 (m, 2H), 1.08 (s, 3H).

Fmoc-DiAzK [N2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N6-((3-(3-methyl-3H-diazirin-3-yl)propoxy)carbonyl)-L-lysine] (**15**):




Without further purification, 2.03 g (7.3 mmol) of **14** were dissolved in 20 ml anhydrous DMSO. 2.94 g (1 eq) Fmoc-lysine-OH (Sigma-Aldrich 17290) were dissolved in DIEA (1.95 ml, 1.5 eq.) and anhydroous DMSO under argon and stirred at RT. The solution of **14** was added dropwise over the course of 2 h. After stirring the mixture over night at room temperature, $H_2O$ and ethyl acetate were added, the aqueous phase adjusted to pH 3 with HCl and extracted with EtOAc. The solvent was evaporated to yield **15**.

DiAzK (**3**):




The crude protected amino acid was dissolved in DCM with 20% piperidine *v/v* and stirred for 30 min at room temperature. Water was added and the solvent evaporated. The product was purified by flash chromatography with acetone/MeOH/$H_2O$ 65/25/10 *v/v/v* to yield 332 mg (1.2 mmol) of DiAzK **3** (59% yield over three steps). The synthesis was repeated on a larger scale to yield another 2.8 g of **3** with similar yields.

$^{1}$H NMR (400 MHz, $D_2O$) δ 3.95 (t, *J* = 6.1 Hz, 2H), 3.63 (t, *J* = 6.1 Hz, 1H), 3.26 – 3.18 (m, 2H), 3.10 – 2.98 (m, 2H), 1.86 – 1.19 (m, 8H), 0.92 (s, 3H).

$^{13}$C NMR (MeOD) δ 129.97, 128.29, 122.14, 120.77, 79.57, 79.22, 56.22, 41.40, 31.86, 24.01, 19.84.

HR-ESI (neg. mode) = 285.15658 g/mol (calc. -$H^+$: 285.15683 g/mol).

**Supplementary figures:**

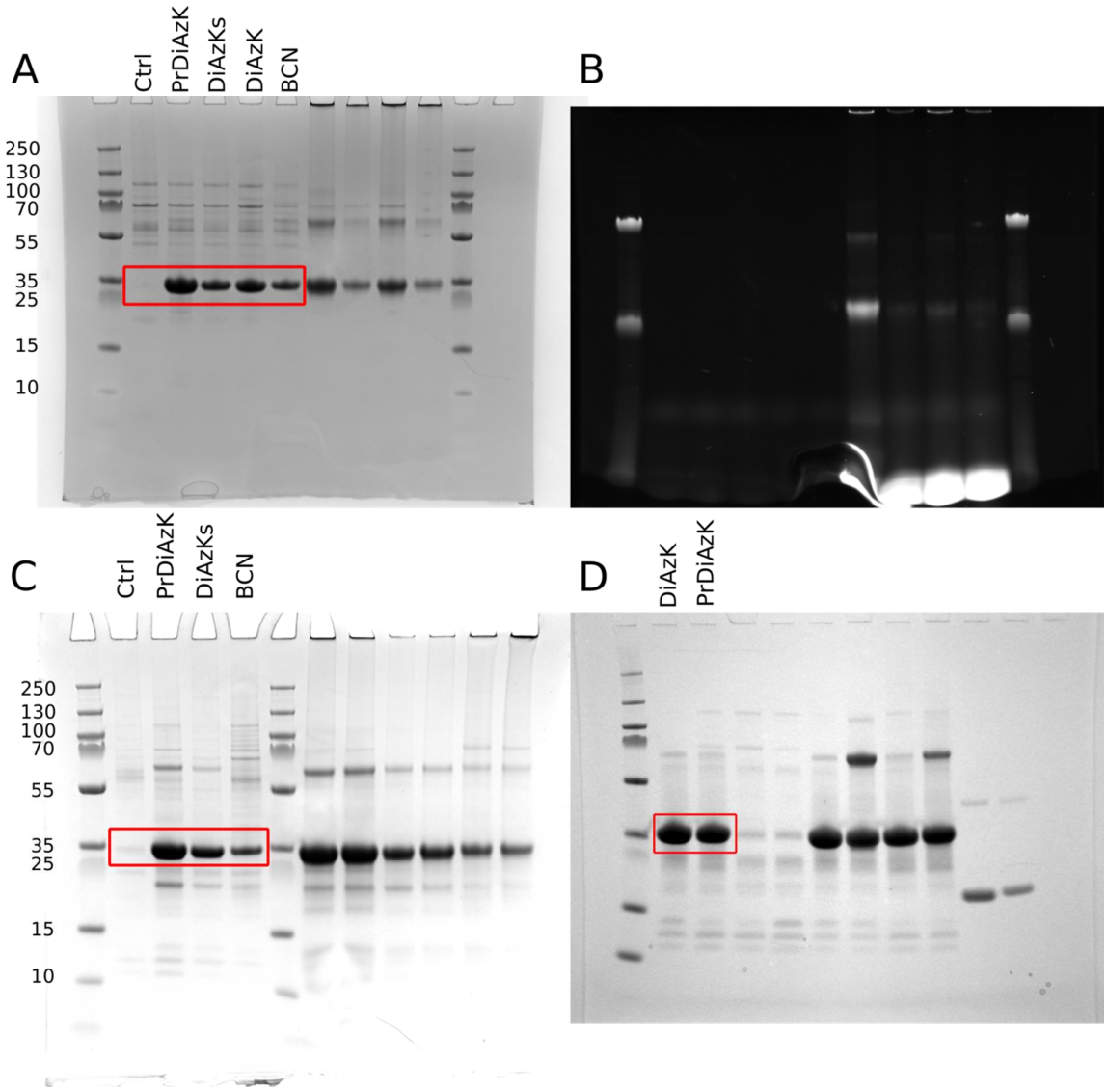


**Figure S1:** Unmodified gels. A) Coomassie image of the gel in figure 1C and 2C. B) Fluorescence image of gel in figure 2C. C) Coomassie image of additional expression test gel used for quantification. D) Coomassie image of gel shown in figure 2B. Red boxes indicate bands used for quantification in figure 1D.

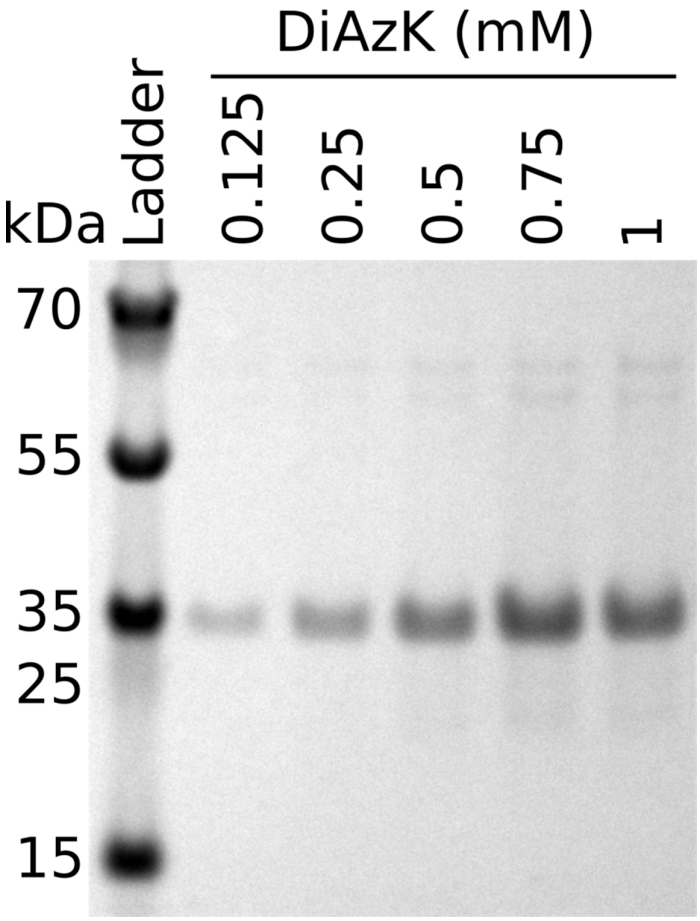



**Figure S2:** Concentration dependent incorporation of DiAzK (**3**) into eGFP. Concentrations of 0.125 mM to 1 mM are used, indicating a maximum protein yield between 0.75 and 1 mM.