

Residue Choice Defines Efficiency and Influence of Bioorthogonal Protein Modification Via Genetically Encoded Strain Promoted Click Chemistry

Samuel Reddington,^{a,b} Eric M. Tippmann^a and D. Dafydd Jones^b

^a School of Chemistry, Cardiff University, Cardiff, UK. Tel: 02920 02920 870452; E-mail: tippmann@cardiff.ac.uk

^b School of Biosciences, Cardiff University, Cardiff, UK. Tel: 02920 20874290; E-mail: jonesdd@cardiff.ac.uk

Supporting Information

Chemicals

Synthetic mutagenic oligonucleotides were custom synthesised by Integrated DNA Technologies. The lyophilised samples were resuspended in water to a final concentration of 100 µM. Ampicillin (Melford) and tetracycline (Fluka) were used at final concentrations of 100 and 25 µg/mL, respectively. *Para*-azido-*L*-phenylalanine (AzPhe) was purchased from Bachem and dissolved in 0.25 M NaOH prior to use.

Dibenzylcyclooctyne-Fluor 585 (dye **2**) was purchased from Click Chemistry Tools and dissolved in DMSO to a stock concentration of 2.5 mM.

Generation of Amber Stop Codon Mutants

The gene encoding superfolder GFP (sfGFP), resident in the pBAD/His A plasmid (Invitrogen) between the NcoI and XhoI sites with a C-terminal hexahistidine tag, was mutated using the Phusion[®] site-directed mutagenesis PCR protocol (New England Biolabs). The amber stop codon (TAG) mutations were introduced using the mutagenic oligonucleotide primer pairs as outlined in Supporting Table 1. Following PCR, linear DNA fragments were phosphorylated using T4 Polynucleotide Kinase (New England Biolabs) to allow recircularisation of the vector by Quick DNA Ligase (New England Biolabs). The resulting ligation mix was used to transform *Escherichia coli* DH5α cells by electroporation. After selection on LB agar plates supplemented with ampicillin, plasmid was isolated from a number of different colonies and the gene encoding sfGFP was sequenced to confirm the presence and position of the desired TAG mutation. Sequencing was performed by the Cardiff University Molecular Biology Support Unit.

Supporting Table 1 – Mutagenetic oligonucleotide primer pairs used to introduce amber stop codon mutations into sfGFP by site directed mutagenesis PCR. TAG mutations are underlined.

Position of TAG mutation	Forward primer (5' - 3')	Reverse primer (5' - 3')
26	<u>TAG</u> TTTAGCGTTCGTGGCGAAGGC	ATGGCCATTCACATCACCATCCAGTTCC
34	<u>TAG</u> GGTGATGCGACCAACGGTAAACTGA	GCCTTCGCCACGAACGCTAA
132	<u>TAG</u> GATGGCAACATTCTGGGTCATAACTGG	TTTAAAATCAATACCTTTTCAGTTCAATGCGGTTTC
204	<u>TAG</u> AGCGTTCTGAGCAAAGATCCGAATG	GGTGCTCAGATAATGATTATCCGGCAGCA

Protein Production and Purification

E. coli Top10 cells (Invitrogen) were used to produce sfGFP containing AzPhe using a reprogrammed genetic code system similar to that described previously.¹ Cells were co-transformed with both:

- (1) pBAD-sfGFPXXTAG (TAG-mutated GFP genes) or pBAD-wtsfGFP (wild-type GFP gene). The basic pBAD backbone contains the *bla* gene that confers bacterial resistance to ampicillin and the *araBAD* arabinose inducible promoter immediately upstream of the sfGFP gene.
- (2) pDULEazidoRS which harbours the aminoacyl-tRNA-synthetase and tRNA_{CUA} required for AzPhe incorporation during cellular protein synthesis.¹ The plasmid also contains the *tet^r* gene that confers bacterial resistance to tetracycline.

Single colonies grown on ampicillin and tetracycline LB agar were used to inoculate 5 mL LB broth cultures with the same antibiotics. After 16 hours of growth at 37°C, the saturated 5 mL cultures were used to inoculate 50 mL ZYM-5052 arabinose autoinduction media² (1/200 dilution of the expression culture) supplemented with ampicillin and tetracycline. ZYM-5052 autoinduction media was composed of 1% (w/v) Tryptone, 0.5% (w/v) Yeast Extract, 0.5% (v/v) glycerol, 0.05% (w/v) glucose, 0.2% (w/v) lactose, 25 mM Na₂HPO₄, 25 mM KH₂PO₄, 50 mM NH₄Cl, 5 mM Na₂SO₄, 2 mM MgSO₄, trace metals and 0.05% (w/v) L-arabinose in ultra-pure water. The trace metal mixture contained 4 μM CaCl₂, 2 μM MnCl₂, 2 μM ZnSO₄, 0.4 μM CoCl₂, 0.4 μM CuCl₂, 0.4 μM NiCl₂, 0.4 μM Na₂MoO₄, 0.4 μM H₃BO₃, and 10 μM FeCl₃ in ultra-pure water. Cultures were grown at 37°C with shaking for 1 hour before 1 mM *p*-azido-L-phenylalanine was added, cultures were grown for a further 23 hours. Cells were harvested by centrifugation and lysed in sodium phosphate buffer (100 mM, 300 mM NaCl; pH 8) containing 0.1 mg/mL lysozyme and 1 mM PMSF by sonication. Proteins were purified by Ni-affinity chromatography using His SpinTrap™ columns in a microcentrifuge (GE Life Sciences). Protein concentration was quantified using the D_C Protein Assay (BioRad) and protein production yields were calculated in mg of purified protein per litre of growth medium. The yields for each sfGFP variant were: Lys26AzPhe-sfGFP, 69 mg/L; Glu34AzPhe-sfGFP, 132 mg/L; Glu132AzPhe-sfGFP, 106 mg/L; Gln204AzPhe-sfGFP, 296 mg/L; wt-sfGFP, 524 mg/L.

Strain-Promoted Azide-Alkyne Cycloaddition (SPAAC)

SPAAC reactions were performed using both pure protein samples and crude cell lysates (to mimic intracellular conditions more closely) as the source of AzPhe-sfGFP. With regards to pure protein, SPAAC was performed using 10 μM protein and various concentrations of dye **2** in the range of 10 to 100 μM in 100 mM sodium phosphate buffer (pH 8) at 25 °C for up to 48 hours. SPAAC was also performed using cell lysates from 50 mL expression cultures standardised to an A₆₀₀ of 1. Reactions were performed with 100 μM of dye **2** at 25 °C for 20 hours. Reaction mixtures were centrifuged at 15,000 x g for 2 minutes to remove any insoluble material. Unreacted probe was removed from the protein component by applying the reaction mixture to a His SpinTrap™ Ni-affinity chromatography column and separated using a microcentrifuge.

Fluorescent SDS-PAGE Analysis

SPAAC reactions were analysed using fluorescent imaging after SDS-PAGE. The reaction components were separated by SDS-PAGE (15 % (w/v) acrylamide) before and after removal of unreacted dye **2** using established protocols. Gel bands were imaged and analysed using a Typhoon 9400 Variable Mode Imager with a 532 nm excitation laser and a 610 (BP 30) nm emission filter. Images were processed and fluorescent bands were quantified using ImageJ software.³ SDS-PAGE gels were subsequently stained with Coomassie Blue stain (50% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) R250 Coomassie blue) and imaged to view all proteins.

Analytical Gel Filtration

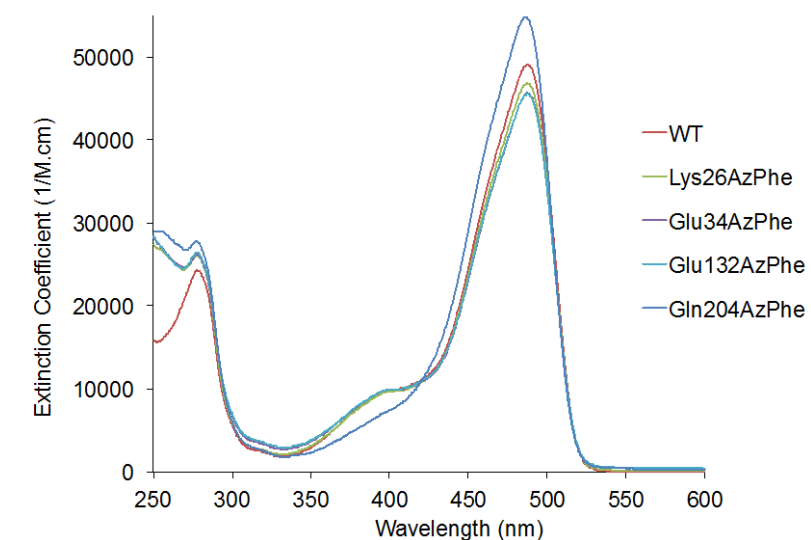
AzPhe-sfGFP variants and wt sfGFP were analysed for the presence of any oligomeric species by analytical gel filtration using a SuperdexTM 75 GL column (GE Healthcare). Proteins were loaded at 10 μ M and separated at a flow rate of 0.5 mL/min with sodium phosphate buffer (50 mM, 150 mM NaCl; pH 8). Elution volume was monitored by absorption at 280 and 485 nm. The column was calibrated using protein standards of known molecular mass (1.4, 17, 44, 158 and 670 kDa) and used to estimate the molecular mass of samples. All protein variants eluted from the column within 0.2 mL of each other with an average molecular mass of 25 kDa (pred. 27.8 Da).

Fluorescence and Absorption Spectroscopy

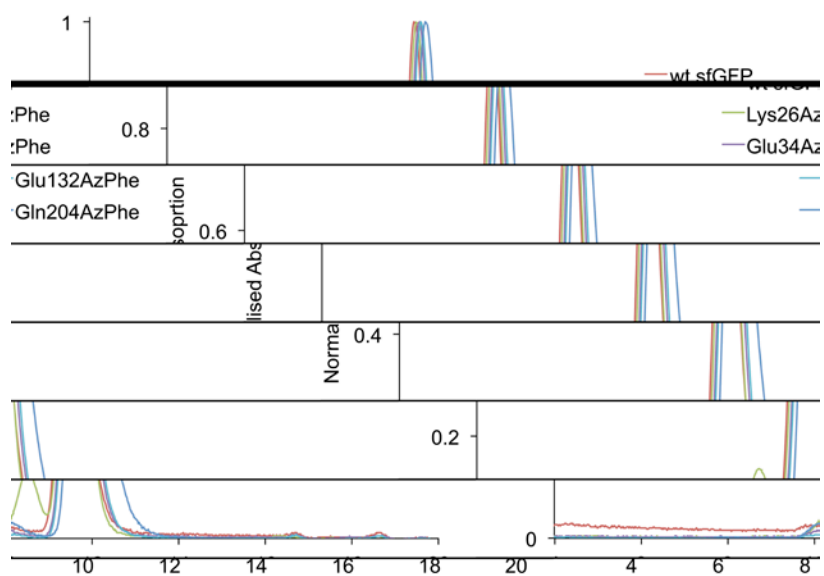
SPAAC reactions were also analysed using fluorescence and UV-visible absorption spectroscopy after removal of unreacted dye **2** by Ni-affinity chromatography as described above. Samples were prepared in 100 mM sodium phosphate buffer (pH 8) and measurements performed at 25°C. Fluorescence measurements were performed in a Varian Cary Eclipse spectrophotometer in a 5 mm pathlength QS quartz cuvette. Emission spectra of sfGFP were measured using a constant excitation wavelength of 485 nm and excitation spectra of sfGFP were measured by monitoring emission at 511 nm. Emission spectra of dye **2** were measured using a constant excitation wavelength of 595 nm and excitation spectra of dye **2** were measured by monitoring emission at 611 nm. Absorbance measurements were performed in a JASCO V-660 spectrophotometer in a 10 mm pathlength QS quartz cuvette. Extinction coefficients were calculated for AzPhe-sfGFP variants and wt-sfGFP using Beer-Lambert's law.⁴ Briefly, the UV-visible absorption of 12.5 μ M of the purified protein variants was measured; protein samples were then re-quantified using the Bio-Rad D_C assay to confirm the protein concentration as 12.5 μ M. The absorbance value at the λ_{max} of the sfGFP variants (\sim 485 nm) was divided by the concentration (in M) to give the extinction coefficient. The full absorption spectra of the sfGFP variants converted to extinction coefficient can be seen in Supp. Fig. 1. Reaction yields were calculated by comparing the absorption peak intensity of sfGFP (\sim 485) with dye **2** (\sim 594). Absorption values were converted to concentrations using the extinction coefficients of the sfGFP variants (eg. Gln204AzPhe-sfGFP; $\epsilon_{485} = 55,000 \text{ M}^{-1}\text{cm}^{-1}$) and dye **2** ($\epsilon_{594} = 102,000 \text{ M}^{-1}\text{cm}^{-1}$) then used to calculate molar ratios with [dye **2**] / [Gln204AzPhe-sfGFP].

For the study of the reaction rates of Glu34AzPhe-sfGFP and Gln204AzPhe-sfGFP (Fig. 2C), protein labelling was monitored using FRET between sfGFP and the attached dye **2**. The reaction was performed in the fluorimetry cuvette (as above) and started by the addition of dye **2**. The reaction was followed by monitoring dye **2** emission (611 nm) after excitation of sfGFP (485 nm) until the reaction reached completion. The endpoint was quantified as above by removal of unreacted dye **2** followed by measuring UV-visible absorption. Reactions were performed with 10 μ M AzPhe-sfGFP and 20 μ M dye **2**. Rate constants (k) were calculated by plotting the integrated form of the second order rate equation; $1/(A_0 - B_0) [\ln (A_0 B)/(B_0 A)] = kt$, where A_0 and B_0 are the initial concentrations (in M) of sfGFP and dye **2**, respectively, and A and B are the concentrations of sfGFP and dye **2** at time t (in sec), respectively.⁵ The rate constant (units $\text{M}^{-1}\cdot\text{sec}^{-1}$) was given by the gradient of the line and was calculated from 3 repeats.

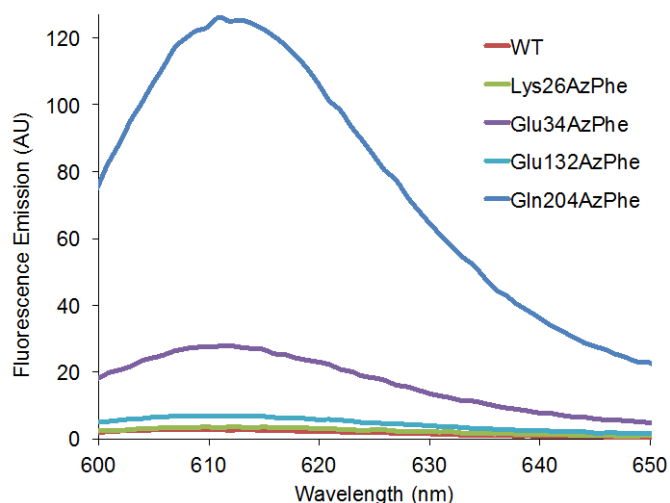
Supporting Figures



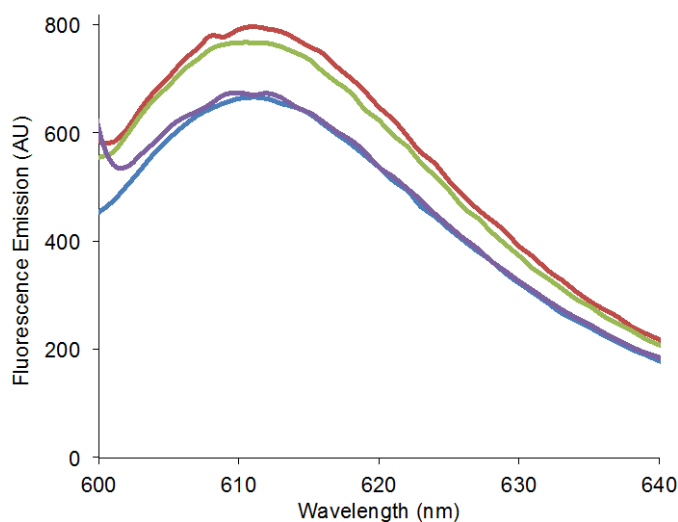
Supporting Fig S1. – UV-visible absorption spectra of AzPhe-sfGFP variants. Measurements were taken with 12.5 μ M protein and converted to molar extinction coefficients as described in the Supporting Methods.



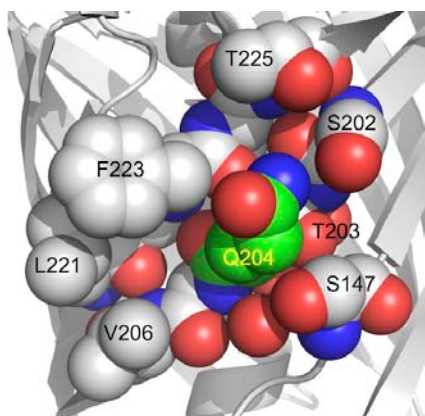
Supporting Figure S2. – Analytical gel filtration of AzPhe-sfGFP variants. Variants were analysed using a Superdex™ 75 GL column (GE Healthcare) at 10 μ M as described in the Supporting Methods. No significant dimerization was observed. The average molecular weight of protein variants was calculated from the elution volume as 25 kDa (pred. 27.8 Da). The small side peak from Lys26AzPhe-sfGFP (elution volume = 10.5 mL) was calculated as 45 kDa.



Supporting Figure S3. – SPAAC using cell lysate as the source of the AzPhe-sfGFP variants (cell lysate standardised to A_{600} of 1.0) and dye **2** (100 μM). Fluorescent emission spectra after excitation at 595 nm that measures the amount of dye **2** conjugated to each sfGFP variant. Unreacted dye **2** was removed by Ni-purification. Reactions were allowed to proceed for 20 hours.



Supporting Figure S4. – Optimisation of protein:probe molar ratio for the reaction between Gln204AzPhe-sfGFP and **2**. Reactions were assessed by fluorescence measuring amount of **2** (after excitation at 595 nm). Gln204AzPhe-sfGFP was kept at a constant concentration of 10 μM and DBCO-585 concentration was varied from 10 μM (1:1, blue line), 20 μM (1:2, red line), 50 μM (1:5, green line) and 100 μM (1:10, purple line). Reactions were allowed to proceed for 20 hours.



Supporting Figure S5. – Local surface environment of residue Gln204 (carbons coloured green). Proximal residues are shown with spacefill representation (carbons coloured grey) and labelled on the diagram.

References

1. S. J. Miyake-Stoner, C. A. Refakis, J. T. Hammill, H. Lusic, J. L. Hazen, A. Deiters and R. A. Mehl, *Biochemistry*, 2010, **49**, 1667-1677.
2. F. W. Studier, *Protein expression and purification*, 2005, **41**, 207-234.
3. M. D. Abramoff, P. J. Magalhães and S. J. Ram, *Biophotonics international*, 2004, **11**, 36-42.
4. J. Pedelacq, S. Cabantous, T. Tran, T. Terwilliger and G. Waldo, *Nature biotechnology*, 2005, **24**, 79-88.
5. E. V. Anslyn and D. A. Dougherty, *Modern physical organic chemistry*, University Science Books, 2006.