Synthesis and Applications of Triazole- and Triazine-containing Amino Acids

Katherine Anne Horner

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Biological Sciences

Astbury Centre for Structural Molecular Biology

School of Chemistry

September 2015

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 2 - *Use of τ-Phosphotriazolylalanine as a Molecular Probe* includes content from the publication: "Evaluation of the Interaction between Phosphohistidine Analogues and Phosphotyrosine-binding Domains" Chembiochem, 2014, volume 15, pages 1088-1091, T. E. McAllister, K. A. Horner and M. E. Webb. The candidate lysed and purified Grb2SH2 from a cell pellet and conducted all of the fluorescence anisotropy experiments. The candidate synthesised Fmoc-phosphotriazolylalanine following a route devised by T. E. McAllister. The candidate also synthesised the phosphotriazole and phosphotyrosine containing peptides and the fluorescent phosphotyrosine probe. All peptides synthesised for ITC experiments and the ITC experiments themselves were conducted by T. E. McAllister. In addition, T. E. McAllister synthesised all other phosphopeptides and control peptides not mentioned. Cell pellets of Grb2SH2 were generated by T. E. McAllister. Lysis and purification of Grb2SH2 for ITC experiments was undertaken by T. E. McAllister. M. E. Webb, T. E. McAllister and the candidate composed the manuscript.

Chapter 3 – *Synthesis and Functionalisation of 1,2,4-Triazines* and **Chapter 4** – *1,2,4-Triazine Cycloaddition Reactions* includes content from the publication "Strain-Promoted Reaction of 1,2,4 Triazines with Bicyclononynes" Chemistry-A European Journal, 2015, volume 41, 14376-14381, K. A. Horner, N. M. Valette and M. E. Webb. N. M. Valette devised the synthetic route for, and initially synthesised the triazinylalanine methyl ester. This synthetic route was subsequently optimised and extended by the candidate. All other work in this publication is attributable to the candidate. M. E. Webb and the candidate composed the manuscript.

Acknowledgements

I would primarily like to thank my supervisor, Mike, for his advice, support, patience, and guidance over the last four years, and for always being around when I have needed his help, even for monthly meetings at 8.30 pm! I would also like to thank my co-supervisor Stuart for his guidance and his ability to always make me smile, and my ex-co-supervisor Arwen for her 'pastoral' support. I would like to thank all members past and present of the Webb/Turnbull/Warriner groups: Bruce, Chadamas, Clare, Dan, Darren, Diana, Gemma, Heather, Ivona, James R, James W, Jeff, Kristian, Laura, Martin, Matt, Phil, Tom B, Tom M and Zoe for making our lab such a lovely environment to work in. I would especially like to thank Tom M for his help and guidance during the early years and for letting me work on his PhD project. Thanks also to Dan and Phil (the devil on my shoulder) for helpful discussions. I would like to thank James R, Darren and Diana for teaching me the wonders of molecular biology; and Clare, who kindly proof-read my introduction and has been indispensable for thesis-related advice. In addition, thank you to Matt and Chris, my partners in crime, who have made these last years far more enjoyable then they probably should have been. Particularly Matt who, especially in these last couple of months, I could not have coped without.

I would also like to acknowledge members of staff from the University: Martin Huscroft (LCMS), Tanya Marinko-Covell (Mass spec) and Chris Empson (Plate reader).

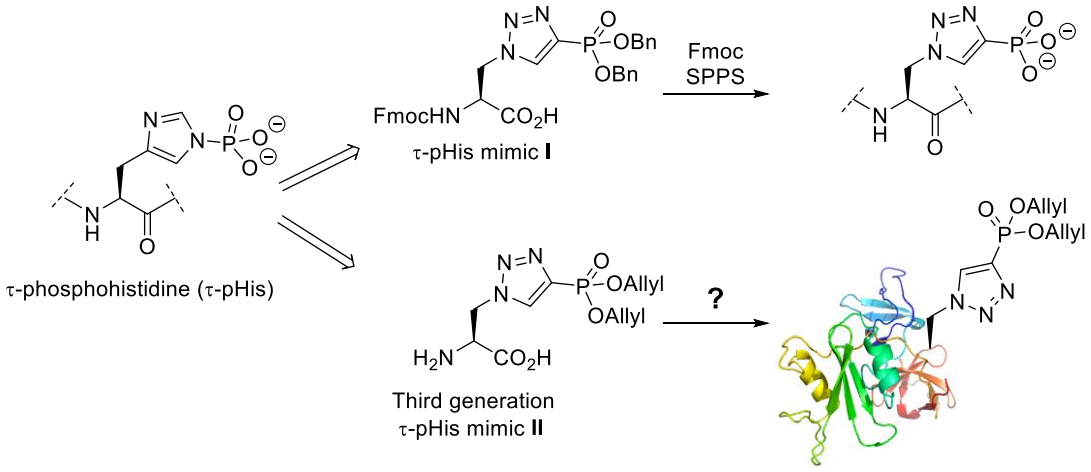
Thanks also to all my family for their love and support. In particular I would like the thank my best friend and sister, Annabel, who moved to Leeds at the start of my PhD and who has always been there for me, normally with a bottle of wine in her hand. Last but by no means least, I would like to thank my parents for their unfaltering support and belief in me, not just over the last four years, but throughout my entire life.

Abstract

Through their use as mimics of post-translational modifications (PTMs) and in bioorthogonal chemical reporting strategies, unnatural amino acids (UAAs) are vital tools for studying biological systems.

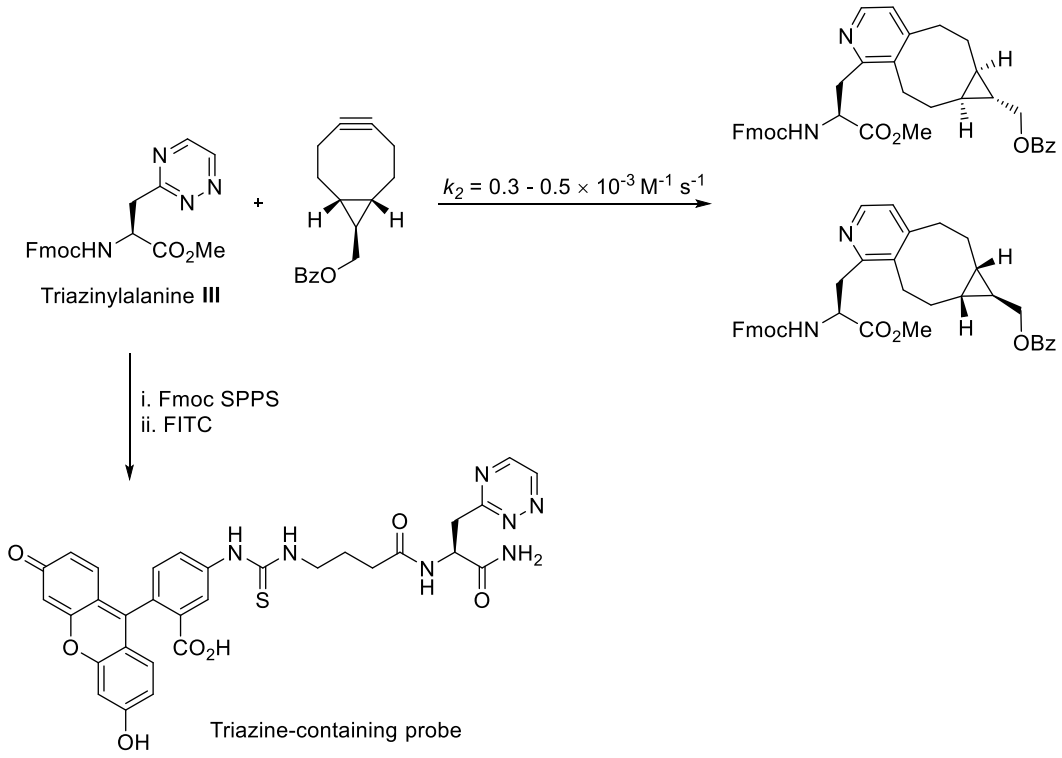
Unnatural Amino Acids as Phosphohistidine Mimics

τ-Phosphotriazolylalanine **I** can act as a non-hydrolysable analogue of phosphohistidine and is compatible with the Fmoc-strategy for peptide synthesis. Peptides containing either **I** or an alternative phosphoamino acid were synthesised and used to demonstrate the selectivity of the SH2 domain of the growth factor receptor-bound 2 protein (a phosphotyrosine binding protein) towards τ-phosphohistidine. In addition, peptides containing **I** were ineffectively used to study the binding interaction between a histidine-phosphorylated protein, phospho*enol*pyruvate synthase and its cognate regulatory protein, YdiA. It was concluded that mimicking the primary structure of one protein through peptide generation was not sufficient to study this protein-protein interaction. As a result, third generation τ-phosphotriazolylalanine **II** was synthesised, which has the potential to be genetically incorporated into proteins using amber suppression.



Unnatural Amino Acids as Bioorthogonal Probes

Bioorthogonal reactions can be used to selectively derivatise probes onto biomolecules. Although there are a number of chemical reactions that have been used for this purpose, many are limited in terms of biocompatibility and synthetic accessibility of bioorthogonal reagents. Therefore, a novel bioorthogonal reaction was developed based on the cycloaddition of 1,2,4-triazines to strained dienophiles. An accessible and robust synthesis of novel 1,2,4-triazin-3-yl-linked amino acid **III** was devised and its reactivity towards a range of strained dienophiles was investigated. It was determined that **III** reacted readily with bicyclononyne at 37 °C with a second order rate constant between 0.3 - 0.5 $\times 10^{-3}$ M^{-1} s^{-1}, depending on solvent mix. The utility of **III** towards late stage functionalisation of probe molecules was also demonstrated through generation of a fluorescent probe containing **III**. This triazine probe was used to demonstrate the cycloaddition of triazine to a strained dienophile *in vitro*.



Contents

Abbreviations

ADC	Aspartate α-decarboxylase
ADP	Adenosine 5'-diphosphate
ATP	Adenosine 5'-triphosphate
BCN	Bicyclononyne
Bn	Benzyl
Boc	*tert*-butoxycarbonyl
Bz	Benzoyl
C. elegans	*Caenorhabditis elegans*
CBT	Cyanobenzothiazole
Cbz	Carboxybenzyl
CD	Circular Dichroism
CoA	Coenzyme A
CP	Cyclopropene
CuAAC	Copper-catalysed Azide-Alkyne Cycloaddition
DCA	Dicyclohexylammonium
DCC	*N*-*N'*-Dicyclohexylcarbodiimide
DCE	1,2-Dichloroethene
DMAP	Dimethylaminopyridine
DMF	Dimethylformamide
E. coli	*Escherichia coli*
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EGF	Epidermal Growth Factor
EGR	Epidermal Growth-factor Receptor
eq.	Equivalents
FITC	Fluorescein isothiocyanate
Gaba	γ-aminobutyric acid
GDP	Guanosine 5'-diphosphate
Grb2	Growth Factor Receptor-bound 2
GTP	Guanosine 5'-triphosphate
HOMO	Highest Occupied Molecular Orbital
HPLC	High-Performance Liquid Chromtography
IC_{50}	Half maximal inhibitory concentration
IEDDA	Inverse Electron-demand Diels-Alder
IMAC	Immobolised Metal Affinity Chromatography
ITC	Isothermal Titration Calorimetry
LB	Lysogeny Broth
LCMS	Liquid Chromatography-mass Spectrometry
LUMO	Lowest Unoccupied Molecular Orbital
MBP	Maltose Binding Protein
Mbt	*Methanosarcina bakeri*
MD	Mass Directed
Mj	*Methanococcus jannaschi*
NCL	Native Chemical Ligation
NHS	*N*-Hydroxysuccinimide
NMR	Nuclear Magnetic Resonance Spectroscopy

Norb	Norbornene
NPDK	Nucleoside Diphosphate Dikinase
NTP	Nucleoside Triphosphate
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDRP	PPDK Regulatory Protein
PEP	Phospho*enol* Pyruvate
PGM	Phosphoglycerate Mutase
Pi	Inorganic Phosphate
Pip	Piperidine
PPDK	Pyruvate, phosphate dikinase
PPS	Phospho*enol*pyruvate synthase
PTM	Post-translational modification
Pyl	Pyrrolysine
r	Anisotropy
rds	Rate determining step
SDM	Site Directed Mutagenesis
SEC	Size Exclusion Chromatography
SH2	Src homology 2
SH3	Src homology 3
Shc	Src homology 2 domain containing
S_NAr	Nucleophilic aromatic substitution
Sos	Son of sevenless
SPAAC	Strain-promoted Azide-alkyne Cycloaddition
SPANC	Strain-promoted Azide-nitrone Cycloaddition
SPIEDAC	Strain-promoted Inverse Electron-demand Diels-Alder cycloaddition
SPPS	Solid Phase Peptide Synthesis
TCO	*trans*-cyclooctene
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIPS	Triisopropylsilyl ether
TIS	Triisopropylsilane
TLC	Thin-layer Chromatography
tRNA	Transfer RNA
Trz	Triazine
Tz	Triazole
wt	*wild type*

Chapter 1 Introduction

1.1 Unnatural Amino Acids as Molecular Mimics

Post-translational modifications (PTMs) of proteins are involved in the regulation of almost all aspects of cell biology. Protein PTMs expand the functional diversity of the proteome through covalent addition of functional groups; proteolysis of regulatory domains and protein degradation.[1] It has been found that approximately 5% of the genome is devoted to enzymes that catalyse protein PTMs in higher eukaryotes. The major types of modifications include phosphorylation, acetylation, methylation, ubiquitination, nitrosylation, glycosylation, lipidation and proteolysis; of these, protein phosphorylation is the most prevalent.

A major challenge in this area is to understand how post-translational modifications alter protein structure, function, activity and regulation.[2] In order to study them *in vitro*, a homogenous protein or a relevant peptide sequence needs to be altered on a specific residue to contain the covalent modification that is being investigated.[3] This can often be difficult as the enzyme that catalyses the modification *in vivo* may not be known or it cannot be used *in vitro*. Chemically synthesised unnatural amino acids (UAAs) can be designed to mimic amino acids that have been post-translationally modified and incorporated into proteins and peptides. As such, UAAs are increasingly becoming indispensable research tools as molecular probes for the study of PTMs of proteins in biological systems.

1.1.1 Protein Phosphorylation

Protein phosphorylation is essential to the roles of proteins in signal transduction and cellular regulation in prokaryotes and is becoming increasingly observed in eukaryotes. It is estimated that at any one time, approximately 30% of proteins are phosphorylated in the

cell.[4] Phosphorylation of proteins is a reversible process and as such involves two distinct classes of enzymes: kinases (for phosphorylation) and phosphatases (for dephosphorylation). Nearly all heteroatomic amino acids can be phosphorylated; the most well characterised of these are the phosphate esters phosphothreonine **1**,[5] phosphotyrosine **2**[6] and phosphoserine **3**,[5] and this is due to their stability under the acidic conditions routinely used to study them (Figure 1.1).[7] The phosphoroamidates derived from lysine **4**,[8] arginine **5**[9] and the two regioisomers of histidine (**6** and **7**)[10] have also been detected in proteins as well as phosphocysteine **8**[11] and the phosphoanhydrides derived from glutamate **9**[12] and aspartate **10**.[13]

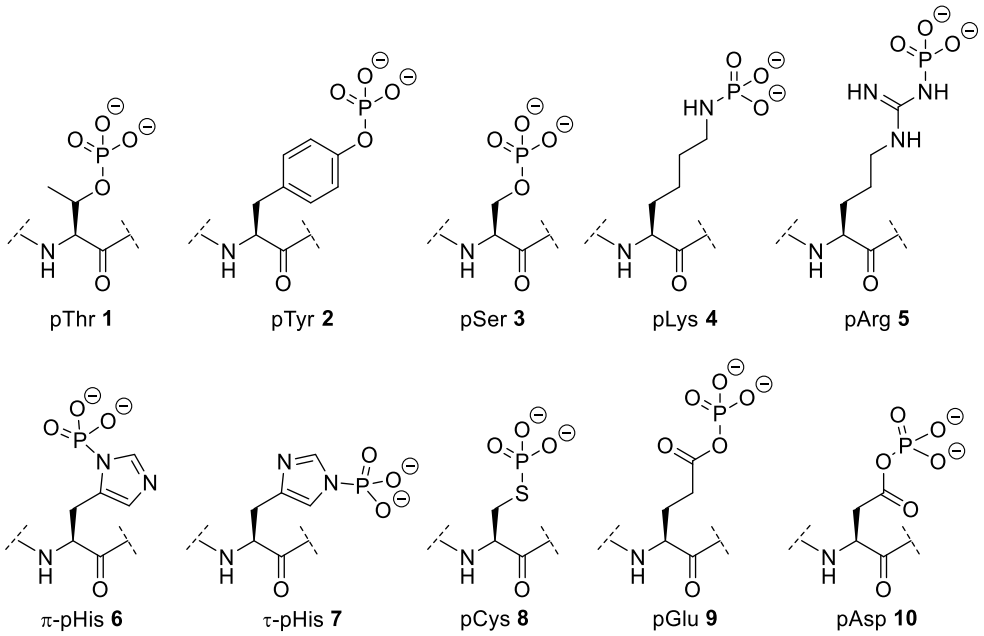


Figure 1.1: Phosphorylated amino acids that have been detected in proteins

1.1.2 Phosphohistidine

Protein histidine phosphorylation plays a crucial role in signalling processes in prokaryotes and is being increasingly observed in eukaryotes.[10] Phosphohistidine (pHis) is unusual compared to other phosphoamino acids in that phosphorylation can occur on either of the imidazole nitrogens to give two chemically distinct regioisomers: π(pros)-phosphohistidine **6** and τ(tele)-phosphohistidine **7**.[i] Phosphohistidine contains a high-energy P-N phosphoroamidate bond that is readily hydrolysed under acidic conditions. Hultquist[14] assessed the hydrolytic stability of chemically phosphorylated π-pHis **6** and τ-pHis **7** under a range of acidic conditions and found that in 1M HCl_{aq} at 49 °C, π-pHis **6** was hydrolysed

[i] There is some ambiguity in the nomenclature of the phosphohistidine isomers. Throughout this text the isomers will be referred to as π-(pros - near) **6** and τ-(tele - far) **7**.

faster than τ-pHis **7** (Figure 1.2). Under mildly acidic conditions (pH 5, 49 °C) the difference in hydrolytic stability is more pronounced with π-pHis **6** being dephosphorylated 10 times faster then τ-pHis **7**. Moreover, under mildly basic conditions, when facile dephosphorylation does not occur, slow conversion of **6** to **7** is observed; although it should be noted that τ-pHis **7** is stable in these conditions. These results correspond with the intrinsic kinetics of histidine phosphorylation as the kinetic product, π-pHis **6**, is more prone to hydrolysis and hence less thermodynamically stable than τ-pHis **7**.

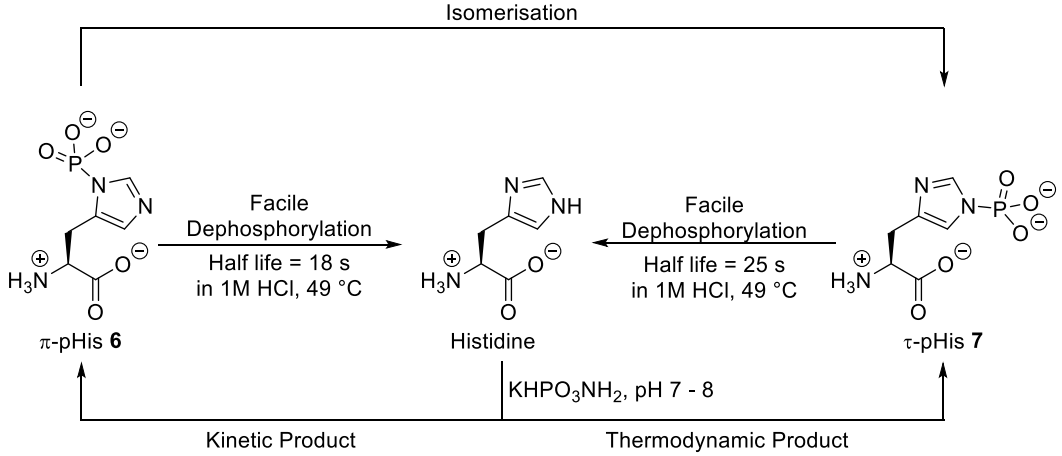


Figure 1.2: Relative stabilities of phosphohistidine regioisomers **6** and **7** in acidic conditions

1.1.3 Analysis of Phosphohistidine-containing Systems

Phosphohistidine has a higher ΔG of hydrolysis (-12 to -13 kcal mol^{-1}) in comparison to phosphohydroxyamino acids,[15] making it an excellent facilitator for the downstream transfer of phosphoryl groups to target molecules. However, the transient nature of phosphohistidines P-N bond is also why this modification has proven to be so difficult to detect and study.

1.1.3.1 Detection of Protein Histidine Phosphorylation

Historically, acidic conditions have been employed in protein phosphorylation analysis which fail to preserve the phosphohistidine modification.[16] Adaptations of such methodologies using milder conditions have now been reported. However, histidine phosphorylation can easily be unnoticed if it is not being specifically looked for as the detection strategy being used to identify other phosphoamino acids could result in the loss of phosphohistidines phosphoryl group modification.[15] Conventionally, a combination of phosphoamino acid analysis, radiolabelling, reverse-phase thin layer chromatography (TLC), reverse-phase electrophoresis, high performance liquid chromatography (HPLC), ^{31}P NMR spectroscopy and mass spectrometry (MS) is used to detect and localise the site of protein histidine phosphorylation. However, phosphoproteins are low in abundance and as

such, need to be enriched before analysis. This has traditionally been achieved through immunoaffinity chromatography or immobilized metal affinity chromatography (IMAC). There are no examples of IMAC enrichment on histidine-phosphorylated proteins, although Napper *et al*.[17] have successfully utilised Cu(II) to enrich phosphohistidine-containing peptides digested from an *in vitro* enzymatic phosphorylated HPr protein from *E. coli*.

1.1.3.2 Studying Protein Histidine Phosphorylation

In addition to detecting protein histidine phosphorylation, methods are needed to generate phosphohistidine containing proteins or peptides in order to study the modification. For histidine-containing proteins that autophosphorylate it is possible to use the substrate to selectively phosphorylate a defined histidine.[18] However, in many instances the kinase needed for, or the route towards, protein histidine phosphorylation is not known and as a result the use of enzymatic phosphorylation for this purpose is limited.[15] Histidines in proteins and peptides can also be selectively chemically phosphorylated using potassium phosphoroamidate.[19] Indeed, synthetic phosphopeptide generation has significantly facilitated investigation into serine, threonine and tyrosine phosphorylation in proteins.[20,21] However histidine rephosphorylation does not prevent hydrolysis and as a result, a basic pH must be maintained; restricting subsequent experiment type required for analysis.

1.1.3.3 Analogues of Phosphohistidine

Due to the difficulties in detecting and studying protein histidine phosphorylation, a number of unnatural amino acids acting as non-hydrolysable mimics of phosphohistidine have been generated. Arguably, the most important class of phosphohistidine analogues to date are phosphonotriazoles (pTz) **11** and **12**, which were designed to mimic τ-pHis **7** and π-pHis **6** respectively (Figure 1.3). These compounds maintain the 5-membered ring scaffold of the imidazole in phosphohistidine to ensure correct orientation of functional groups through replacement with a triazole moiety and incorporate the hydrogen bond accepting nitrogen present in phosphohistidine, whilst the hydrolytic P-N bond is replaced with a P-C bond. Their syntheses were independently described by the groups of Muir,[22] Webb[23] and Piggott[24]. Both the Muir and Piggott groups reported the generation of τ- and π-phosphotriazoles **13** and **16**; with Muir and co-workers demonstrating the compatibility of both regioisomers with the Boc-strategy for SPPS through their incorporation into peptides based on the sequence surrounding His18 of histone H4. McAllister *et al*.[23] synthesised Fmoc-protected τ-phosphotriazolylalanine derivative **14** and demonstrated its incorporation into a range of peptides using Fmoc-based SPPS and a two-step deprotection strategy using TMS-bromide for ethyl ester removal. Synthesis of Fmoc-protected pTz **14** obviates the need for a hydrogen fluoride facility, and more importantly, the use of this corrosive and dangerous gas, which was required for Boc deprotection when generating peptides

incorporating triazolylalanines **13** and **16**. McAllister *et al*.[25] went on to synthesise τ-pHis analogue **15** that included a dibenzyl-protected phosphonate, and demonstrated its full compatibility with the Fmoc-strategy for peptide synthesis with phosphoryl group deprotection being achieved under standard peptide cleavage conditions.

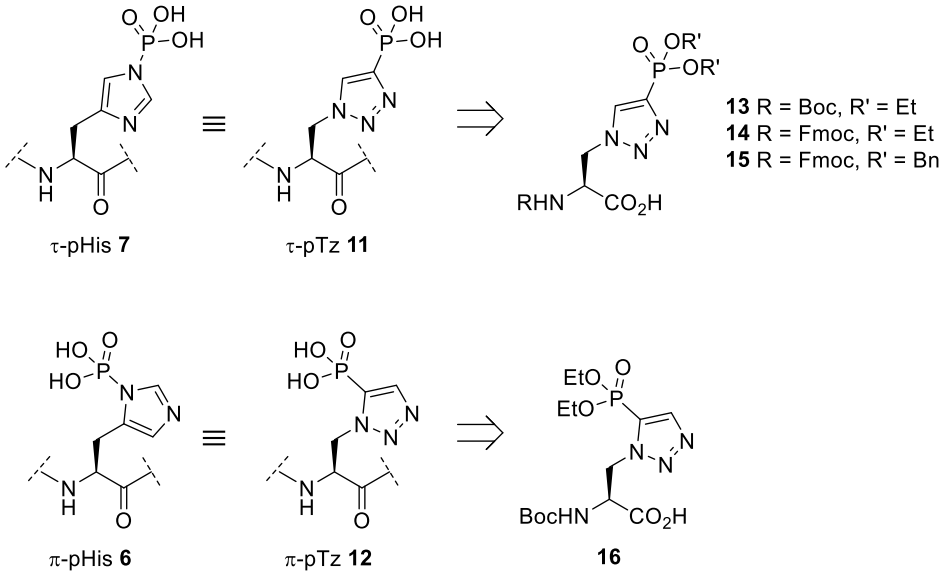


Figure 1.3: Stable analogues of τ- and π- phosphohistidine **7** and **6** based on a triazole scaffold.

Very recently, Hunter and co-workers[26] incorporated τ-pTz **11** and π-Tz **12** into degenerate peptide libraries and used these to develop monoclonal antibodies that specifically recognise either τ-pHis or π-pHis in a sequence independent manner. The development of such antibodies could enable the detection of histidine phosphorylation on a wide variety of phosphohistidine substrates. They could also be used to enrich samples containing phosphohistidine for further analysis. The successful generation of antibodies using τ-pTz **15** and π-Tz **16** as haptens also demonstrates the utility of these compounds as mimics of phosphohistidine for studying the modification.

1.1.4 Phosphohistidine Prevalence in Biology

While protein histidine phosphorylation is known to be vital for the survival of prokaryotes, observed occurrence in eukaryotes is relatively low.[27] It is possible that the reason for this is that higher organisms use different PTMs to relay biochemical information; but is most likely that the previous lack of biochemical tools available to study phosphohistidine is the cause. The latter proposition is reinforced when considered alongside the case of the slime mould *Physarum polycephalum*; where phosphohistidine was found to account for 6% of all phosphoamino acids in its basic nuclear proteins.[28] This prevalence is extremely high when compared with prevalence of phosphotyrosine in eukaryotic cellular phosphoproteins, which

is under 1%.[29] Therefore, it is possible that protein histidine phosphorylation is significantly more prevalent in eukaryotes and higher mammals than currently thought.

1.1.4.1 Prokaryotes

In bacteria, the lability of phosphohistidines phosphoroamidate bond is exploited in a number of ubiquitous systems including bacterial two-component signalling[30] and sugar phosphotransferase systems.[31] Furthermore, phosphohistidine is utilised as an intermediate in phosphoglycerate mutase (PGM) to alter the phosphorylation pattern of glycerate molecules[27] (PGM is an enzyme that is involved in gluconeogenesis and glycolysis).[32] In general, these systems have been extensively studied, and as such are well understood.

1.1.4.2 Eukaryotes

In eukaryotes, a small number of proteins have been identified to contain phosphohistidine. These include Histone H4, which is part of a family of histones that condense DNA into nucleosomes within cell nuclei.[33] In Histone H4, phosphorylation of both His18 and His75 has been observed, although the purposes of these phosphorylation events are unknown. The β-subunit of heterotrimeric G-proteins[34] have also been identified to contain a phosphohistidine residue as well as nucleoside phosphate kinases (NDPKs).[35] Heterotrimeric G-proteins transmit chemical signals across cell membranes and NDPKs sustain cellular levels of nucleoside triphosphates (NTPs). A crystal structure of succinyl-CoA synthetase, which catalyses the formation of succinyl-CoA for use in ketone body metabolism, was also found to contain a τ-phosphohistidine residue.[36] In addition, studies suggest that protein histidine phosphorylation may play a role in mammalian cell signalling with phosphorylated histidine residues being identified on both the cytoplasmic tail of P-selectin (a leukocyte adhesion molecule)[37] and Annexin I (a phospholipid-binding protein).[38] Protein histidine phosphorylation has also been shown to play an important role in C_4 carbon fixation of plants.

1.1.4.2.1 C_4 Carbon Fixation

C_4 photosynthesis is the means in which plants including maize, sorghum and sugar cane, obtain energy from the sun.[39] C_4 plants represent under 4% of all flowering plants; which could be thought insignificant if not considered alongside the fact that they dominate some of the most productive ecosystems on earth. In C_4 carbon fixation, carbon dioxide is captured in mesophyll cells through its conversion to hydrogen carbonate and reacted with phospho*enol* pyruvate (PEP) to form oxaloacetate (Figure 1.4). Oxaloacetate is then converted to C_4 acids, which are shuttled to the bundle sheath cell. Here, the organic acids are decarboxylated to form pyruvate, enabling the release of carbon dioxide to enter the Calvin cycle where it is used to make sugars via C_3 photosynthesis. Pyruvate is transported

back to the mesophyll cells where it is rephosphorylated by pyruvate, phosphate dikinase (PPDK) to regenerate PEP; a reaction that requires ATP and inorganic phosphate and produces AMP and diphosphate.

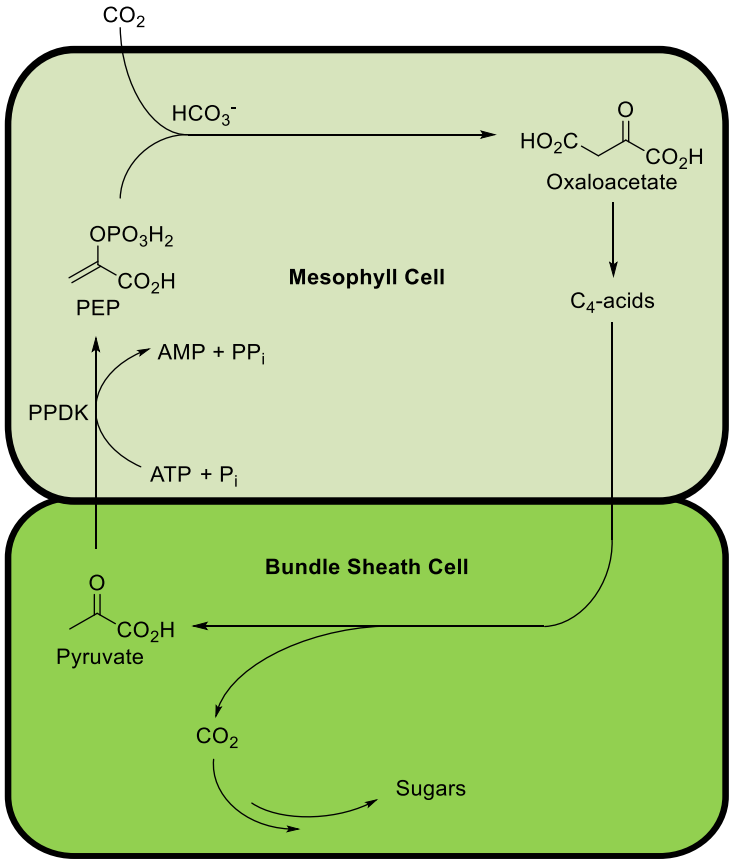


Figure 1.4: The mechanism of C_4 carbon fixation as used by a number of flowering plants. Carbon dioxide is fixed via reaction with PEP to form oxaloacetate, which is in turn converted to C_4 acids. The C_4 acids are transported to the bundle sheath and broken down into carbon dioxide and pyruvate by decarboxylases. β-phosphate is then transferred from ATP to Pyruvate to regenerate PEP; a reaction that is catalysed by PPDK.

PPDK consists of an ATP binding domain, a phosphohistidine binding domain and a pyruvate/PEP binding domain (Figure 1.5a).[40] The phosphohistidine binding domain swivels between the ATP domain where histidine becomes phosphorylated, and the pyruvate/PEP domain where the phosphoryl group is transferred to pyruvate to form PEP. C_4 carbon fixation is regulated by changes in PPDK activity which is achieved by the PPDK regulatory protein, PDRP.[41] PDRP binds to PPDK in its histidine phosphorylated form and phosphorylates a proximal threonine residue using ADP. This renders PPDK inactive by preventing phosphoryl group transfer to pyruvate to form PEP. After phosphorolytic cleavage of phosphohistidine, PDRP catalyses the dephosphorylation of phosphothreonine, a process that is inhibited by ADP, and PPDK is reactivated (Figure 1.5b). In the dark, stromal levels of ADP increase which serve to inhibit dephosphorylation of

phosphothreonine and PPDK is inactivated. In the light, stromal levels of ADP decrease, and hence dephosphorylation of phosphothreonine is favoured and PPDK is reactivated.[42]

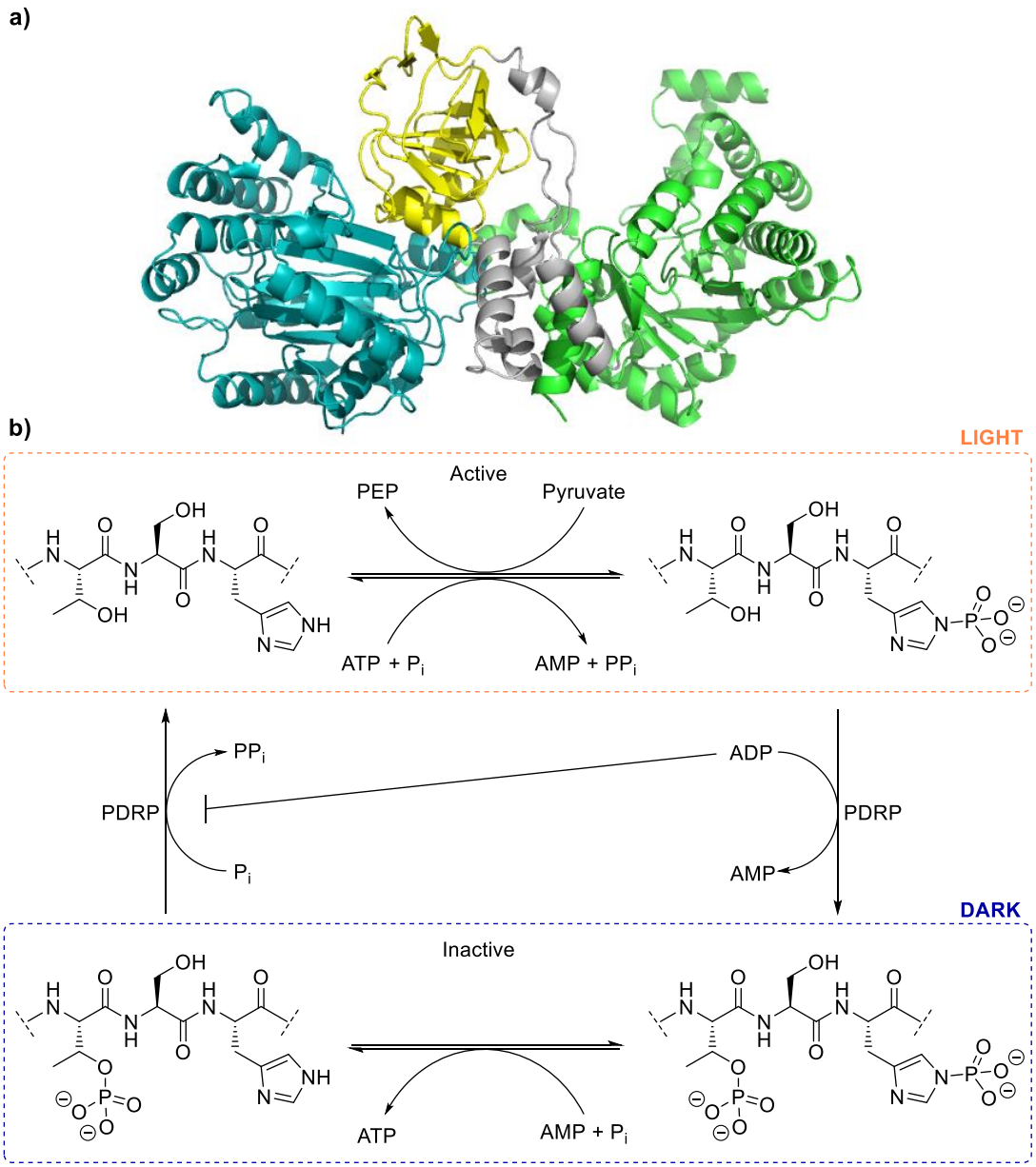


Figure 1.5: a) Crystal structure of PPDK with N-terminal ATP binding domain shown in blue, the phosphohistidine binding domain shown in yellow, and the C-terminal pyruvate/PEP binding domain shown in green. The two polypeptide segments that connect the pHis binding domain to the other two domains are shown in grey; b) Mechanism by which PPDP is regulated by PDRP: PDRP binds to the pHis binding domain (yellow in a)) of PPDK, when it contains a τ-phosphohistidine residue and uses ADP to phosphorylate a proximal threonine residue, rendering PPDK inactive. PPDK is reactivated by dephosphorylation of phosphothreonine, a process that is inhibited by ADP. PDB file: 1DIK

PDRP is an unusual enzyme in that it acts as both an ADP-dependent kinase and a phosphatase. Although the structure of PPDK is well characterised and has been extensively studied, not a lot is known about its regulatory protein PDRP and how this bifunctional enzyme regulates PPDK activity. In addition, it has not been elucidated how

phosphorylation of a threonine residue proximal to the catalytic histidine entirely eclipses PPDK activity. It is possible that this could be through electrostatic repulsion of negatively charged substrates (pyruvate, ATP and P_i) by the dianionic phosphoryl group of threonine.[40] Equally, steric interference of the proximal phosphothreonine residue could prevent association of the catalytic phosphohistidine domain to the pyruvate/PEP domain.[41]

1.2 Unnatural Amino Acids as Bioorthogonal Probes

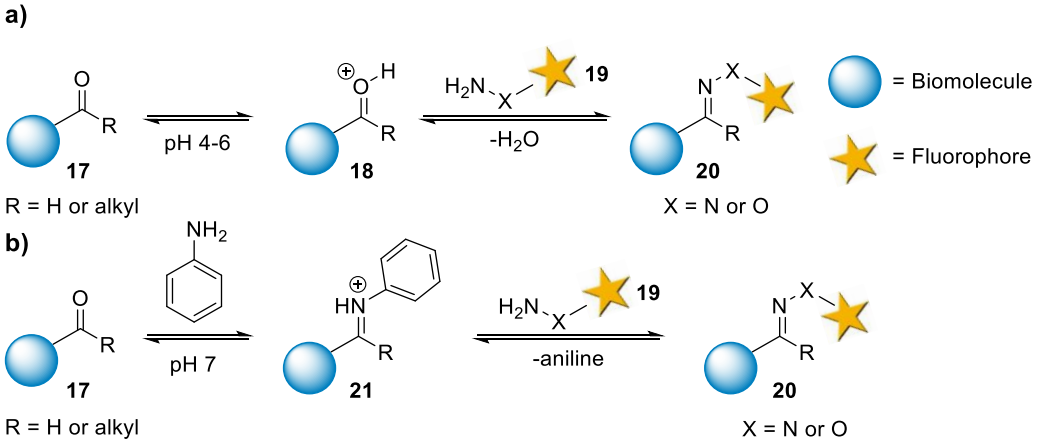
In addition to their use as molecular mimics, unnatural amino acids bearing reactive moieties can be incorporated into biomolecules and used to introduce biophysical probes and labels to study the function and dynamics of proteins.[43] A UAA that has unique chemical functionality is incorporated into a biomolecule through genetic or chemical methods or via hijacking a metabolic pathway. An external chemical probe is introduced which selectively and specifically reacts with the incorporated functionality, resulting in the covalent labelling of the biomolecule with the probe. For this to be possible, the associated chemical reaction must be bioorthogonal.

1.2.1 Bioorthogonal Reactions

A bioorthogonal reaction is a chemical reaction that can take place selectively inside a living system without interfering with any endogenous functional groups that are found in biomolecules.[44] The reactants, side-products, and products must not be toxic to the cell; be thermodynamically, kinetically and metabolically stable; and the resultant product must be covalently linked. Bioorthogonal reactions tend to be governed by second order reaction kinetics, and as such, their reaction rates depend upon the effective concentration of both of the reactive components. Because of this, the rate of reaction must be fast enough to obviate the need for a large excess of one reagent. This is to prevent solubility issues or off-target reactions of the excess reagent with abundant endogenous functional groups, the latter being possible when a reactive component is not infinitely chemoselective (which is unfortunately often the case). A number of chemical reactions have been successfully used both *in vitro* and *in vivo* as bioorthogonal chemical reporters, although many of these reactions have intrinsic limitations that put their biocompatibility into question.

1.2.1.1 Ketone/Aldehyde Condensations with Amine Nucleophiles

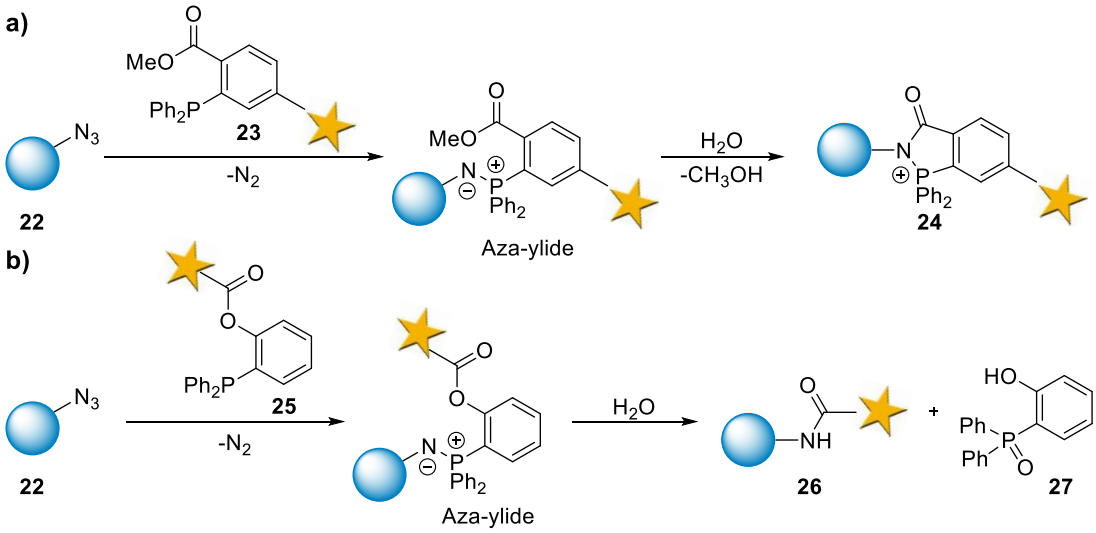
The reaction of aldehydes (**17**, R = H) or ketones (**17**, R = alkyl) with alkoxyamines (**19**, X = O) or hydrazides (**19**, X = N) was among the first reactions to be used as a bioorthogonal chemical reporter (Scheme 1.1a).[44] In acidic conditions (pH 4-6), the ketone or aldehyde **17** that has been previously incorporated into the biomolecule becomes protonated and reacts with amine nucleophile **19** to form Schiff base **20**.[45] Due to the acidic conditions needed, this reaction is only appropriate for *in vitro* systems.[46,47] It should also be noted that the reaction has slow reaction kinetics (k_2 = 10^{-4} - 10^{-3} M^{-1} s^{-1}).[48] However, these undesirable reaction kinetics and acidic conditions can be overcome by using aniline as a nucleophilic catalyst (Scheme 1.1b).[49,50] Unfortunately, the incorporated carbonyl functionality faces competition from carbonyl bearing metabolites such pyruvate, oxaloacetate and sugars.



Scheme 1.1: Condensation of ketones or aldehydes with alkoxyamines (**19**, X = O) or hydrazides (**19**, X = N) to fluorescently label a biomolecule under a) acid catalysis or b) aniline catalysis.

1.2.1.2 The Staudinger Ligation

First reported by Saxon and Bertozzi[51] in 2000, the Staudinger ligation, which involves the reaction of azide **22** to triarylphosphine reagent **23**, has been applied both *in vitro* and *in vivo* to covalently link biomolecules to phosphine derived probes (Scheme 1.2a).[52,53] This reaction has relatively slow reaction kinetics (k_2 = 10^{-3} M^{-1} s^{-1}),[54] and as a result of this and the intrinsic oxidation sensitivity of phosphine; high concentrations of the phosphine reagent are required. Other limitations include reduction of the incorporated azide functionality **22** by endogenous thiols or other reductants before reaction with **23** and cross-reactivity of the phosphine functionality with disulfides. There is also a traceless version of the Staudinger ligation in which the final amide-linked product **26** is released from the phosphine oxide moiety **27** (Scheme 1.2b).[55]

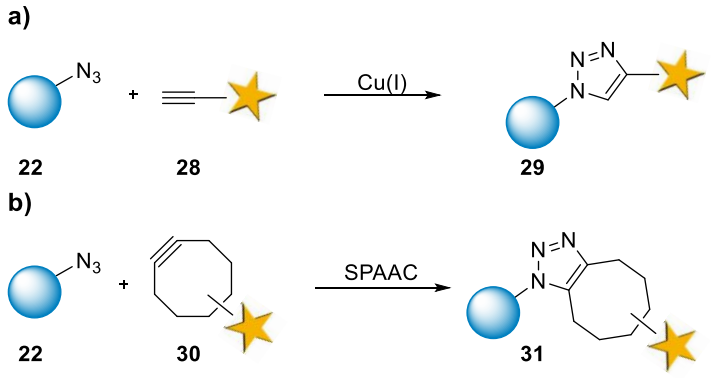


Scheme 1.2: a) The Staudinger ligation between azides **22** and triarylphosphines **23** via an aza-ylide intermediate as a chemoselective labelling strategy and b) a traceless version of the Staudinger ligation in which the fluorescently labelled product **26** is freed from the phosphine moiety **27**.

1.2.1.3 Reactions of Azides with Alkynes

The use of azides as 1,3-dipoles in (3 + 2) cycloadditions with terminal alkynes to generate stable triazoles was discovered by Rolf Huisgen in the 1950s,[56] with the addition of Cu(I) salts to accelerate reaction rate being independently demonstrated by the Sharpless and Meldal groups in 2002 (Scheme 1.3a).[57,58] The copper-catalysed azide-alkyne cycloaddition (CuAAC) reaction is approximately seven orders of magnitude faster than its analogous uncatalysed counterpart. Unfortunately, Cu(I) is toxic to living systems and this reaction is therefore not suitable for *in vivo* studies.[59]

Bertozzi's strain-promoted azide-alkyne cycloaddition (SPAAC) exploits the intrinsic reactivity of strained cyclooctynes towards azides, and thus circumvents the need for a metal catalyst (Scheme 1.3b).[60] Cycloadditions between a highly strained cyclooctyne such as **30** and azides have marginally improved reaction kinetics in comparison to CuAAC with reported second order rate constants between 0.1 - 1 M^{-1} s^{-1}.[61,62] Dubbed 'copper-free Click chemistry', this reaction has been successfully used in the imaging of live cells[63] and zebrafish embryos.[64] Despite this, cyclooctyne derivatives such as **30** often require complex, multistep synthesis and some have poor water solubility (although many are now commercially available).[61]

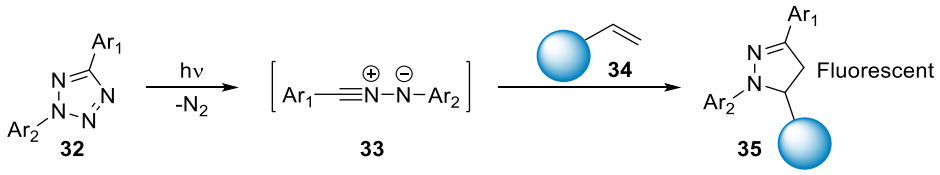


Scheme 1.3: Chemoselective labelling using azide-alkyne cycloaddition reactions with a) copper (I) catalysis, and b) a strained alkyne **30**, to promote reaction rate.

1.2.1.4 PhotoClick Reactions

The photoinduced dipolar cycloaddition of alkenes to nitrile imines (generated via *in situ* photoactivation of the corresponding tetrazole) to yield pyrazoline products such as **35** has relatively fast reaction kinetics (k_2 < 6 M^{-1} s^{-1}) (Scheme 1.4). Known as a PhotoClick reaction, its utility as a bioorthogonal probe has been demonstrated both *in vitro* and in living systems.[65,66] This chemistry is particularly useful as a bioorthogonal labelling strategy as the pyrazoline products are fluorescent, whilst the starting reagents are not. In addition, the requirement of UV light for reaction initiation allows for spatiotemporal photoactivation

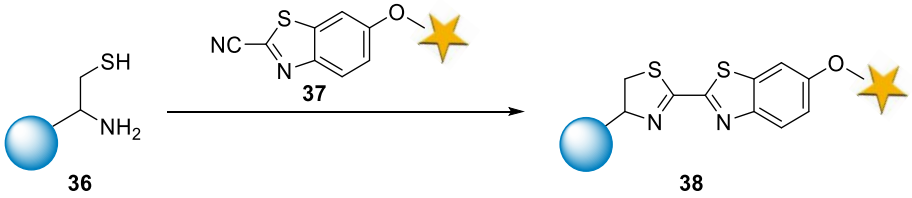
of the reaction. There is some debate as to whether the nitrile imine may cross-react with endogenous alkenes; although the decreased reactivity of *cis*-alkenes (which are the most abundant form found in biomolecules) in comparison to terminal alkenes suggests that this may not be an issue.[44] Development of tetrazoles that can be activated with light at wavelengths not harmful for living cells will make the PhotoClick reaction highly attractive for use in the chemoselective labelling of biomolecules.



Scheme 1.4: Photoclick reaction to label a biomolecule; *in situ* photoactivation of tetrazole **32** generates nitile imine intermediate **33** which undergoes a (3 + 2) dipolar cycloaddition to alkene **34** and fluorescent pyrazoline product **35** is generated.

1.2.1.5 1,2-Aminothiol Condensations

The condensation of 1,2-aminothiol **36** with 2-cyanobenzothiazole (CBT) **37** takes place under physiological conditions with a reported second order rate constant of ~9 M^{-1} s^{-1}, and its utility as a bioorthogonal probe has been demonstrated *in vitro* (Scheme 1.5).[67] Moreover, 1,2-aminothiol derivatives can now be genetically encoded into proteins using amber suppression and CBT **37** displays no cross-reactivity with any endogenous functional groups.[68] The 1,2-aminothiol condensation is yet to be demonstrated in living systems.

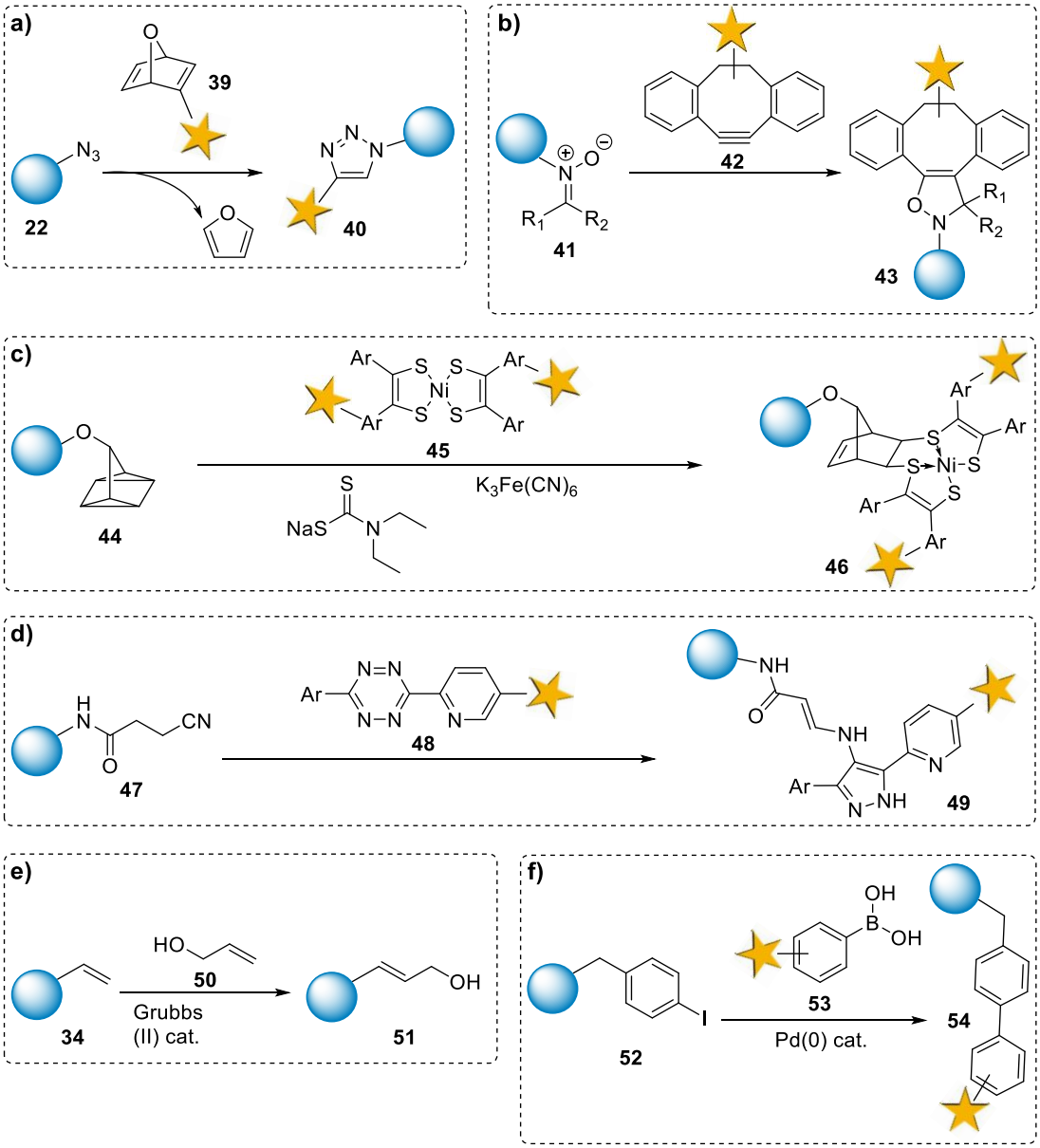


Scheme 1.5: Condensation of fluorescently tagged CBT **37** with 2-aminothiol **36** to label a biomolecule.

1.2.1.6 Other Bioorthogonal Reactions

Other chemistries that have been used in bioorthogonal chemical reporting strategies include the cycloaddition of azides to strained alkenes such as norbornadiene **39** (Scheme 1.6a).[69] This reaction has sluggish reaction rates (k_2 = 10^{-3} M^{-1} s^{-1}) and, due to the addition of cellular nucleophiles to **39** to form addition products, the reaction is not strictly chemoselective.[70] Nitrones such as **41** can also be used as 1,3-dipoles and added to cyclooctynes in a strain-promoted azide-nitrone cycloaddition (SPANC) reaction to yield *N*-alkylated isoxazolines like **43** (Scheme 1.6b).[71] This reaction was demonstrated *in vivo* through the cellular imaging of epidermal growth factor receptors.[72] Unfortunately, the hydrolytic instability of

nitrones limits its applicability as a bioorthogonal probe.[44] Other cycloadditions that have been used for protein labelling include the (2 + 2 + 2) cycloaddition of highly strained quadricyclane **44** to π-systems[73] and the (4 + 1) cycloaddition between isonitriles such as **47** and aromatic tetrazines (Scheme 1.6c and d respectively).[74] In the former case, the cycloaddition can only be applied *in vitro* due to instabilities of both **44** and **45** in living systems. The latter cycloaddition has successfully been used to label cell surface glycans through the metabolic incorporation of isonitrile-modified monosaccharides and subsequent addition of fluorescently tagged tetrazine derivative **48** (Scheme 1.6d).[75] In addition, a number of carbon-carbon bond forming reactions have shown utility as bioorthogonal probes. These include the cross metathesis of alkenes (Scheme 1.6e)[76] and a number of palladium-catalysed cross-coupling reactions (Scheme 1.6f).[44] Investigations into the stability and solubility of the catalysts required for these reactions will demonstrate whether they have broad compatibility with biological systems.



Scheme 1.6: Other reactions that have been used to label biomolecules; a) (3 + 2) cycloaddition of azide **22** and norbornadiene **39** to form stable triazole **40**; b) (3 + 2) cycloaddition of nitrone **41** and dibenzocyclooctyne **42**; c) (2 + 2 + 2) cycloaddition of quadricyclane **44** to nickel bis(dithiolene) **45** with stabilising adducts; d) (4 + 1) cycloaddition of isonitile **47** to aromatic tetrazine **48**; e) olefin cross-metathesis of alkenes **34** and **50** using a Hoveyda-Grubbs second-generation ruthenium catalyst and f) Suzuki-type cross-coupling reaction of aryl halide **52** and boronic acid **53** to generate **54**.

A reaction that has not yet been mentioned but has widespread use in the chemoselective labelling of biomolecules both *in vitro* and *in vivo*, is the inverse electron-demand Diels-Alder (IEDDA) cycloaddition between tetrazines and strained dienophiles.

1.2.2 The Inverse Electron-demand Diels-Alder Cycloaddition

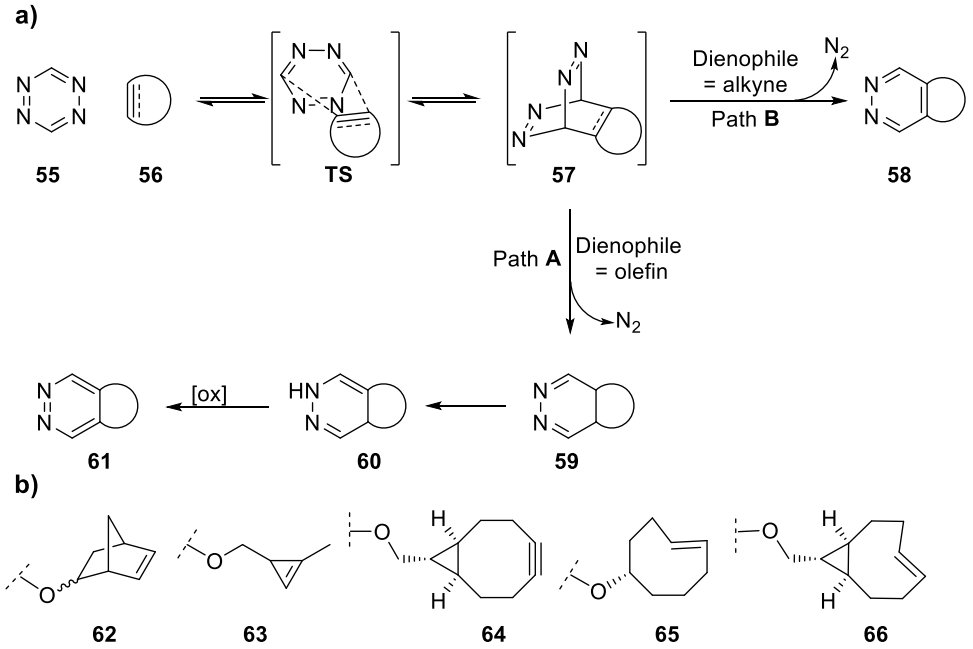
The IEDDA cycloaddition between aromatic *N*-heterocycles and dienophiles was first described by Carboni and Lindsey[77] in 1952 and since then has been used to facilitate the synthesis of pyridazine and pyridine derivatives of varying complexity.[78–81] It was not until 2008 that two groups simultaneously recognised that the strain-promoted inverse electron-demand Diels-Alder cycloaddition (SPIEDAC) between 1,2,4,5-tetrazine and strained dienophiles had application as a bioorthogonal probe.[82,83] While many nitrogen heterocycles can undergo cycloaddition reactions, tetrazine is the most reactive towards dienophiles. As a result, efforts have focused on the development of tetrazine derivatives and dienophiles with increased reactivity towards one another and not on cycloaddition reactions concerning other, less reactive nitrogen-containing aromatic compounds.

1.2.2.1 The 1,2,4,5-Tetrazine Cycloaddition to Strained Dienophiles

Tetrazines conjugate to strained dienophiles through an inverse electron-demand hetero-Diels-Alder retro-Diels-Alder type cascade to form pyridazine products (Scheme 1.7).[84] The highly strained bicyclic adduct **57**, formed from (4 + 2) cycloaddition with inverse electron-demand of tetrazine **55** and strained dienophile **56**, undergoes a rapid cycloreversion to give the corresponding 4,5-dihydropyridazine **59** (for olefin dienophiles (Path **A**, Scheme 1.7a)) or pyridazine **58** (for alkyne dienophiles (Path **B**, Scheme 1.7a) with the liberation of dinitrogen. When the dienophile is an olefin, 1,3-prototropic isomerisation normally gives the corresponding 1,4-dihydropyridazine **60**,[ii] which, dependent upon alkene type, may or may not be oxidised to the fully conjugated pyridazine **61**. Strained dienophiles reported to react with tetrazine derivatives included norbornene **62**,[85] cyclopropene **63**,[86] bicyclononyne **64**[87] and *trans*-cyclooctenes **65** and **66** (Scheme 1.7b).[87]

The tetrazine cycloaddition to strained dienophiles proceeds without a catalyst, is high yielding, produces no toxic by-products and has rate constants ranging from $1 - 10^5$ M^{-1} s^{-1}.[85–87] The possible biological applications of this cycloaddition are extensive and it has now been applied in many: including intracellular imaging;[88] *in vivo* imaging;[64,89] live labelling of cell-surface antigens[90] as well as the modification of cells with nanomaterials for clinical diagnostics.[91]

[ii] If an appropriately good leaving group is present at this point, it will eliminate to give the fully aromatised tetrazine **61**.



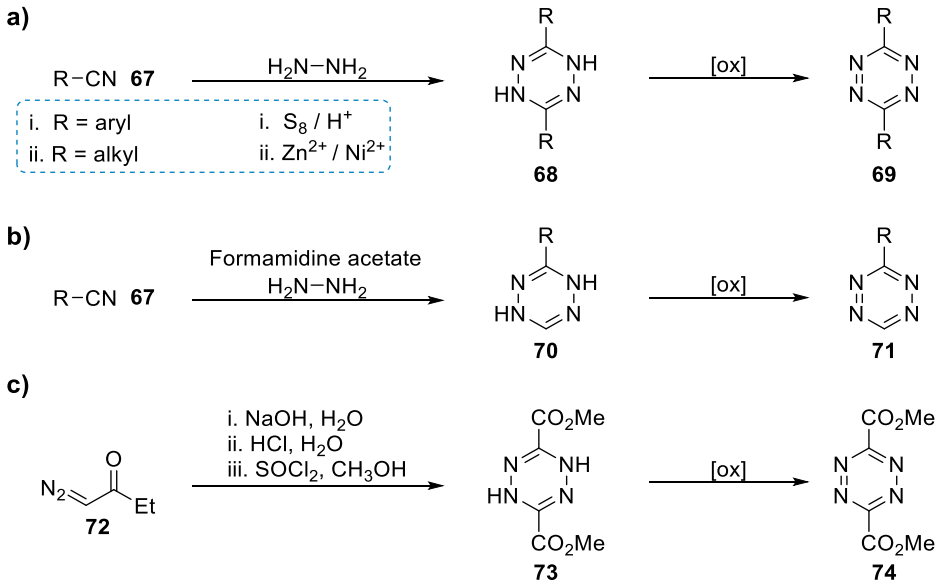
Scheme 1.7: a) SPIEDAC of tetrazine **55** to a cycloalkene (Path **A**) or cycloalkyne (Path **B**). Rate determining (4 + 2) cycloaddition between **55** and **56** leads to a highly strained bicyclic adduct **57** which undergoes a cycloreversion yielding pyridazine derivative **58** (Path **B**) or **59** (Path **A**) and releasing dinitrogen; b) examples of strained dienophiles that react with tetrazine.

1.2.2.2 Limitations of the Tetrazine-SPIEDAC

Although tetrazine-SPEIDAC reactions are rapid and efficient, production of functionalised tetrazine scaffolds needed for derivatisation onto probes remains synthetically challenging.[92] Furthermore, some tetrazines are prone to either hydrolysis[93] or decomposition into the corresponding pyrazoles or thiazoles when exposed to endogenous cellular nucleophiles.[94]

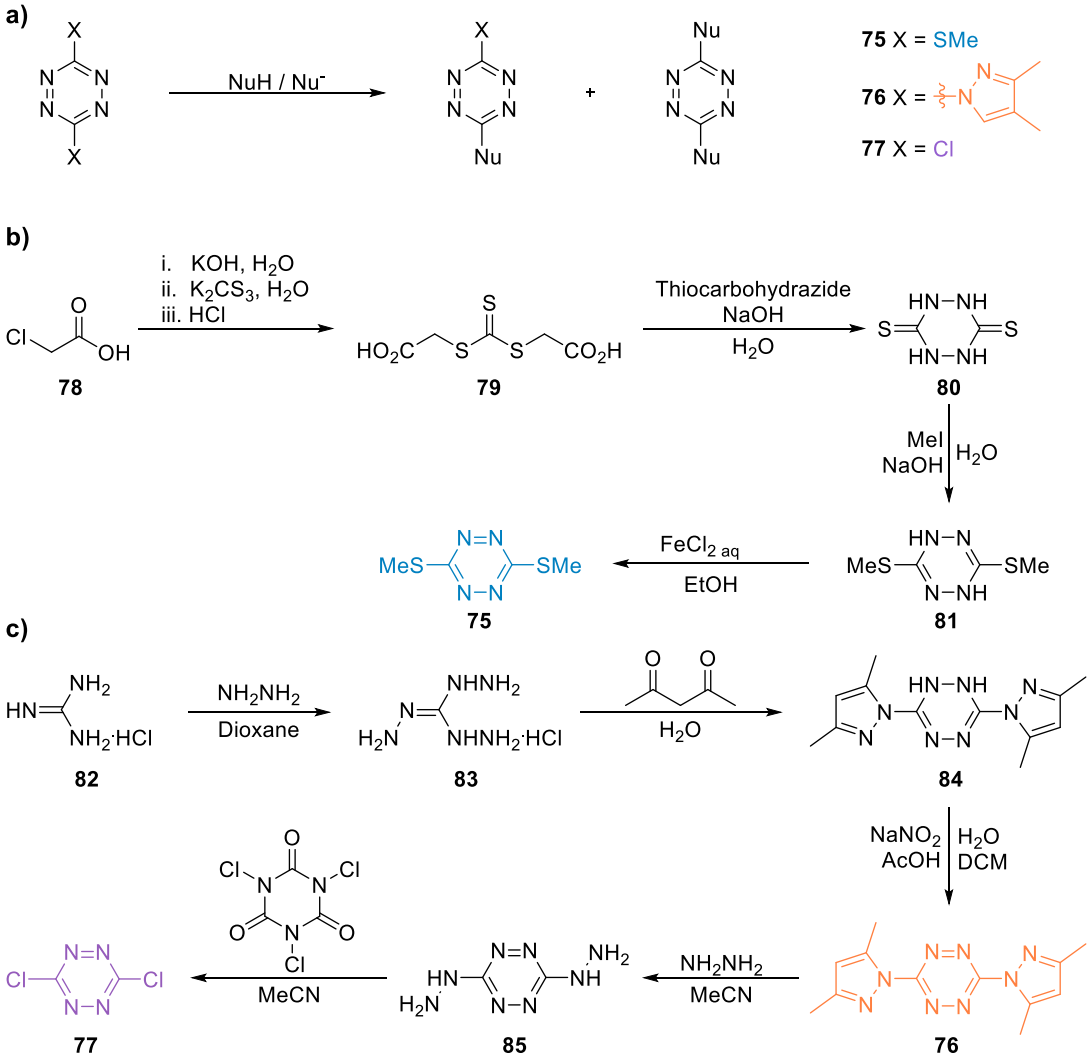
Historically, substituted aromatic tetrazines have been accessed inefficiently via the Pinner synthesis.[95] The Pinner synthesis involves the dimerisation of aromatic nitriles (or their analogues such as nitrile imines or aldehydes) and hydrazine to form dihydrotetrazine derivatives such as **68**, using acid or sulphur as co-catalysts (Scheme 1.8ai).[95,96] Oxidation of **68** affords the corresponding disubstituted tetrazine **69** together with 1,2,4-triazoles or thiadiazoles (when sulphur is used as a co-catalyst) as side-products. Use of formamidine acetate in place of one equivalent of hydrazine will lead to the generation of mono-substituted tetrazines (Scheme 1.8b).[93] Yields for the generation of tetrazine derivatives using this methodology vary considerably.[93,97] For some time, this route was limited to the generation of aromatic substituted tetrazines and if used to synthesise aliphatic substituted tetrazines, would not work, or would work but in low yields.[92] This was until Yang *et al*.[98] found that if a Lewis acid was used to activate alkyl nitrile **67** (R = alkyl), alkyl substituted tetrazines could be generated in good yields (Scheme 1.8aii). It should be noted that in this synthesis anhydrous hydrazine is required, which is volatile and has limited commercial

availability. An alternative strategy to generate functionalised tetrazines is the base-promoted dimerization of ethyl diazoacetate **72** to form dimethyl tetrazine-3,6-dicarboxylate **74**, after oxidation of dihydrotetrazine **73** (Scheme 1.8c).[99] However, **74** is prone to acid promoted rearrangement and slowly decomposes on warming.



Scheme 1.8: Reported syntheses of functionalised 1,2,4,5-tetrazines; a) generation of 3,6-disubstituted (i) aromatic-[95,97] or (ii) alkyl-[98] tetrazines via the Pinner synthesis; b) generation of mono-substituted tetrazine **71** using formamidine acetate and c) synthesis of dimethyl tetrazine-3,6-dicarboxylate **74** through base-catalysed dimerization of ethyldiazoacetate **72**.[99]

Tetrazines can also be functionalised through the nucleophilic aromatic substitution (S_NAr) of dimethylthio-,[100] dipyrazolyl-,[101] and dichloro-[102,103] tetrazines **75-77** with nucleophiles (Scheme 1.9a). In order to perform S_NAr on the tetrazine ring, tetrazines **75-77** must first be synthesised. This is relatively straightforward when dimethylthiotetrazine **75** (Scheme 1.9b)[104] or dipyrazolyltetrazine **76 (**Scheme 1.9c)[101] is required; but increases in number of synthetic steps in the case of dichlorotetrazine **77** (Scheme 1.9c).[92] Furthermore, dihydrazinotetrazine **85**, an intermediate in the synthetic route to **77**, is explosive.[103] Dichlorotetrazine **77** is the most electrophilic, and hence most reactive towards nucleophilic aromatic substitution, of the three cores. The added synthetic complexity required to access this derivative may be discouraging to synthetic chemists. Substitution reactions of **75-77** have been reported with nucleophiles such as alcohols,[105] thiols,[106] and amines/amides.[100,103] However the tetrazine cores of **75-77** are readily decomposed by reductive metals and are therefore not compatible in substitution reactions involving organometallic species.[92]

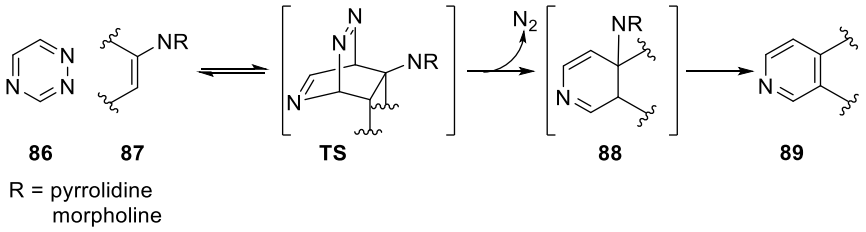


Scheme 1.9: a) General approach for the nucleophilic substitution of methylthio-, dipyrazolyl- or dichloro- tetrazines **75-77**; b) & c) Reported generation of 3,6-disubsituted 1,2,4,5-tetrazines **75-77** for use in S_NAr reactions; a) Synthesis of 3,6-dimethylthio-1,2,4,5-tetrazine **75** through base catalysed dimerization of thiocarbohydrazide and trithiocarbodiglycolic acid **79**;[104] b) Generation of 3,6-pyrazolyl-1,2,4,5-tetrazine **76**[101] through the condensation of triaminoguanidine hydrochloride **83** with pentenedione and subsequent generation of 3,6-dichloro-1,2,4,5-tetrazine **77** via nucleophllic aromatic substitution of dipyrazolyl-tetrazine **76**.[92]

1.2.2.3 The 1,2,4-Triazine Cycloaddition

An effective alternative to the use of 1,2,4,5-tetrazine in chemoselective labelling strategies might be 1,2,4-triazine. Cross-linking reactions have been reported to occur between 1,2,4-triazines and acyclic dienophiles to form dihydropyridine and pyridine derivatives.[107–109] The earliest example of a 1,2,4-triazine cycloaddition dates back to 1969 when Neunhoeffer *et al*[110] conjugated a range of 3-substituted-1,2,4-triazines to simple alkenes and alkynes. Poor regiocontrol combined with modest yields led to limited interest in this transformation for some time. It was not until 1981 that Boger and Panek[111] discovered that the cycloaddition of pyrrolidine or morpholine enamines **87** to 1,2,4-triazine **86** generated

pyridine products with the expected regioselectively in moderate to good yields (Scheme 1.10). This, and the establishment of robust synthetic routes[112,113] to access 1,2,4-triazine derivatives has led to the reaction being exploited to access a range of polycyclic and fused heterocycles through tethered triazine -alkyne/-alkene scaffolds.[108,112,114] Although this cycloaddition is now commonly used as a participant in elegant synthetic routes to provide complex pyridyl-containing structures, the need for elevated temperatures (mostly exceeding 100 °C), and extended reaction times means that it has never been considered for cellular applications. To date, nearly all examples of 1,2,4-triazine cycloaddition reactions involve open-chain and unstrained cyclic dienophiles. There are just two reports concerning 1,2,4-triazine-SPIEDAC reactions between a range of tri-substituted triazines[115] or tethered triazines[116] and a strained dienophile, norbornadiene.



Scheme 1.10:[111] The 1,2,4-triazine-IEDDA to pyrrolidine or morpholine enamines **87** to yield pyridine derivative **89**.

1.2.2.4 Kinetics of Tetrazine and Triazine Cycloadditions

The concerted (4 + 2) cycloaddition is the rate determining step (rds) in both 1,2,4,5-tetrazine and 1,2,4-triazine $LUMO_{diene}$ - $HOMO_{dienophile}$ controlled conjugations.[117] The rate of conjugation is controlled by an interplay of steric and electronic effects of both the diene and dienophile. Electron-donating dienophiles raise the energy of the dienophile HOMO, resulting in a smaller energy difference between the frontier molecular orbitals and an increase in reaction rate. Dienophiles with a high degree of ring strain also reduce the activation energy of the rds by raising the energy of the dienophile HOMO, and decreasing the distortion energy needed to reach the cycloaddition transition state.[117,118] Conversely, dienophile hydrogen exchange for a sterically demanding electron-donating substituents can have an impeding steric effect and raise the distortion energy needed to reach the transition state;[119] this effect is prevalent in alkynes and alkenes bearing large methylthio-, methoxy- and ethoxy- substituents.[117] Therefore, exchange of hydrogen for an electron-donating substituent can lead to an increase or decrease in dienophile reactivity depending upon the electron-donating power and steric bulk of the substituents, and the inherent strain of the dienophile.

Similarly, addition of electron deficient substituents onto the triazine and tetrazine rings lowers the energy of the diene LUMO and increases reaction rate.[107] Wang *et al.*[120] assessed

the rate of cycloaddition of a range of 3,6-disubstituted tetrazines to bicyclononyne **64** and found significant variations in the bimolecular rate constant dependent upon the electron-withdrawing nature of the tetrazine substituents. For instance, 3,6-dipyridinyl-tetrazine was found to cross-link to bicyclononyne 32 times faster than 3,6-diphenyl-tetrazine. Tetrazine is more electron deficient than triazine due to an additional nitrogen in the tetrazine core; this means that the diene LUMO of tetrazine is lower than that of triazine. Cycloaddition reactions involving tetrazines and a specific dienophile will therefore be intrinsically faster than the corresponding triazine-dienophile cycloaddition.

1.3 Incorporation of Unnatural Amino Acids into Proteins

1.3.1 Incorporation via Chemical or Enzymatic Methods

In order to use any of the aforementioned chemoselective reactions to label proteins, one of the reagents needs to be incorporated into the protein. This can be achieved chemically through bioconjugation reactions of an appropriately functionalised reagent with either the side chains of residues in the protein, or the N-terminal of the protein.[43] However, modification of a protein on a single residue is difficult to achieve unless there is a single reactive residue on the surface of a protein. In addition, N-terminal modification of proteins tends to rely on the presence of a specific residue at the N-terminal and is therefore limited in application. It is also possible to synthetically generate moderately sized proteins with an UAA bearing the desired reagent through the transthioesterification of two synthetic polypeptides containing an N-terminal cysteine and a C-terminal thioester respectively- a technique known as native chemical ligation (NCL).[121] This methodology is time-consuming and limited to proteins containing a cysteine residue at an optimal position for ligation.[iii]

It is also possible to incorporate some bioorthogonal reagents into proteins enzymatically. This is achieved through engineering ligases to recognise unnatural modifications instead of their natural small-molecule substrates. This was elegantly demonstrated by Ting and co-workers through the site-specific ligation of a *trans*-cyclooctene derivative[94] and an alkyl azide[122] onto both cell surface and intracellular proteins using engineered lipoic acid ligases. Site-specific incorporation of probe molecules using this technology is growing in popularity. However, for each new substrate to be incorporated, a lipoic acid ligase with an altered active site needs to be engineered; a process which is costly and time-consuming.[44]

[iii] It is possible to use a thiol containing removable auxillary or NCL followed by desulfurisation of the cysteine residue to alanine in order to generate a protein without a cysteine residue; although this adds to the complexity of an already complicated process.[176]

1.3.2 Incorporation via Genetic Methods: Amber Suppression

UAAs can also be incorporated into specific sites into proteins by exploitation of the protein's natural translational machinery using amber suppression.[123] Currently, over 30 UAAs have been genetically incorporated using this methodology.[44] Amber suppression uses the amber stop codon UAG, a codon that normally directs the termination of protein synthesis, to encode an unnatural amino acid loaded onto a complimentary tRNA. Due to the tolerance of the ribosome for UAAs, the unnatural amino acid can be incorporated into proteins during normal protein synthesis.[43]

To do this, employment of an orthogonal $tRNA_{CUA}$ and an aminoacyl-tRNA synthetase (RS) that will explicitly recognise the amber stop codon and the UAA respectively are needed. The amber suppression technique is depicted in Figure 1.6: initially the codon (XXX) for the specific gene is mutated to the amber codon (TAG) by PCR facilitated site directed mutagenesis. The RS then loads the orthogonal $tRNA_{CUA}$-unnatural amino acid pair, and on recognition of the amber codon (UAG) by the ribosome, a 21st amino acid will be incorporated at a specific site into the developing peptide. In order to do this, substrate recognition properties of a natural aminoacyl-tRNA synthetase must be altered so that the synthetase will selectively acylate its cognate $tRNA_{CUA}$ with the desired unnatural amino acid.[124] This can be achieved through a variety of methodologies, all involving repetitive rounds of positive and negative selections in order to amplify synthetase variants selective towards the UAA, and eliminate the variants selective towards endogenous amino acids or both the UAA and natural amino acids. The use of this strategy to generate unnatural RSs to genetically encode distinct unnatural amino acids has been tremendously successful; however, practises for evolving an unnatural RS are costly and time-consuming.

Revolutionary work by Mehl and co-workers[125] has shown that some amino acyl tRNA synthetases have broad specificity for families of structurally (and electronically) similar unnatural amino acids; these synthetases are termed permissive. It is thus possible to use these evolved RSs to genetically encode similar UAAs in addition to the UAA the unnatural amino acyl tRNA synthetase was originally evolved to incorporate. Successful screening of existing tRNA synthetases that have been evolved to incorporate similar UAAs for incorporation of the desired UAA has saved the time and cost that is associated with the generation of novel mutant RSs, or for that matter, the generation of engineered ligases for enzyme mediated incorporation of probes into proteins.

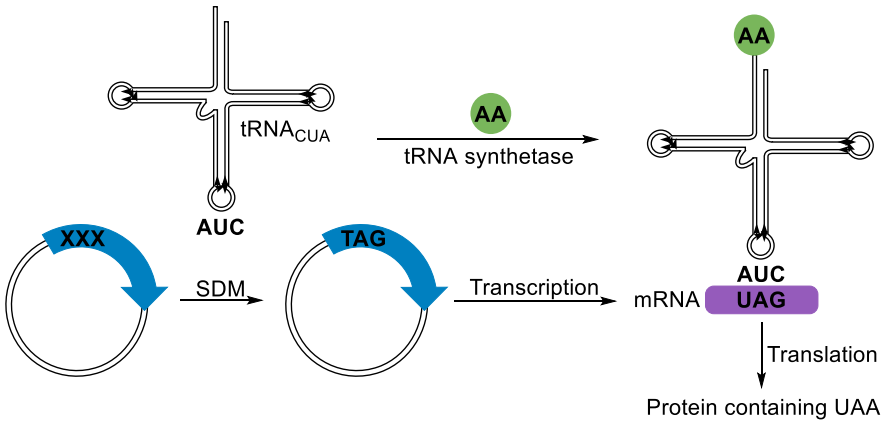


Figure 1.6: The amber suppression technique for incorporation of an UAA. The codon (XXX) for a gene is mutated to the amber stop codon TAG. After loading of the UAA and tRNA by an orthogonal tRNA synthetase and recognition of TAG by the ribosome, the UAA will be genetically incorporated at a specific site in the developing protein.

1.4 Aims and Objectives

1.4.1 Phosphotriazolylalanine as a Phosphohistidine Mimic

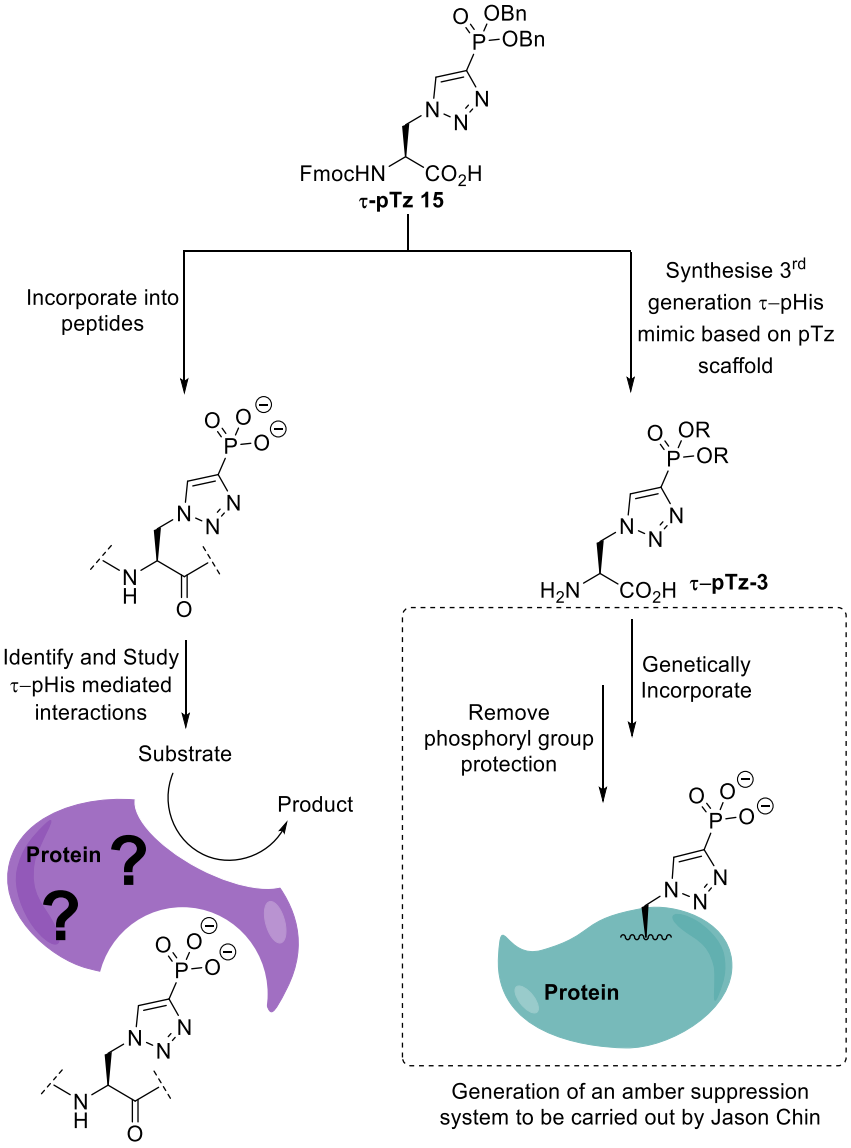


Figure 1.7: Outline of the method to achieve first set of research objectives. It was proposed that τ-pTz **15** would be incorporated into peptides and used to investigate and identify phosphohistidine mediated interactions of proteins. A 3rd generation pTz analogue, **pTz-3**, would also be synthesised and screened as a substrate for existing tRNA synthetases for incorporation into proteins or a new amber suppression system would be evolved to incorporate it.

Following the successful generation of non-hydrolysable τ-phosphohistidine mimic **15** in the group,[25] the aim of this project was to use **15** as a probe to identify phosphohistidine binding proteins and to study known phosphohistidine-mediated interactions that are dependent upon binding of phosphohistidine but not on hydrolysis. It was proposed that this would be achieved through incorporation of τ-pTz **15** into peptides that mimic the binding site of the protein in question and the subsequent use of biophysical techniques to determine

binding parameters of interactions between peptides containing τ-pTz **15** with relevant proteins. In addition, it was envisaged that a third generation phosphohistidine mimic, **pTz-3** would be synthesised with alternative phosphoryl group protection to τ-pTz **15** suitable for incorporation into proteins using amber suppression. The ability to incorporate a phosphohistidine mimic into proteins as well as peptides will further advance studies into protein histidine phosphorylation. These objectives are outlined in Figure 1.7.

1.4.2 The Triazine Cycloaddition to Strained Dienophiles as a Novel Bioorthogonal Probe

Based on the knowledge that strained dienophiles increase the rate of cycloaddition towards *N*-heterocycles, the aim of this project was to generate a novel bioorthogonal probe using the cycloaddition of triazine to a strained dienophile. It was proposed that this cross-linking reaction would proceed without the need for elevated temperatures, and could thus offer an alternative to the 1,2,4,5-tetrazine-SPIEDAC for use in the chemoselective labelling of biomolecules. Although inherently slower than its tetrazine counterpart, the triazine cycloaddition to strained dienophiles would prevent the use of toxic and volatile precursors and offer improved synthetic accessibility to functionalised triazine scaffolds. In addition, reduced reactivity of triazine in comparison to tetrazine may offer enhanced stability *in vivo*. To achieve this, it was envisaged that the reactivity of 1,2,4-triazines would be evaluated against a range of strained dienophiles at physiological temperatures to decide upon a suitable reaction partner. Following this, either a triazine- or strained dienophile- containing UAA would be synthesised, screened for genetic incorporation against existing tRNA synthetases and co-translationally incorporated into a model protein. The resultant mutant protein would then be incubated with the other reactive partner, which would have been derivatised onto a fluorescent tag, to yield a covalently labelled fluorescent protein. These objectives are summarised in Figure 1.8.

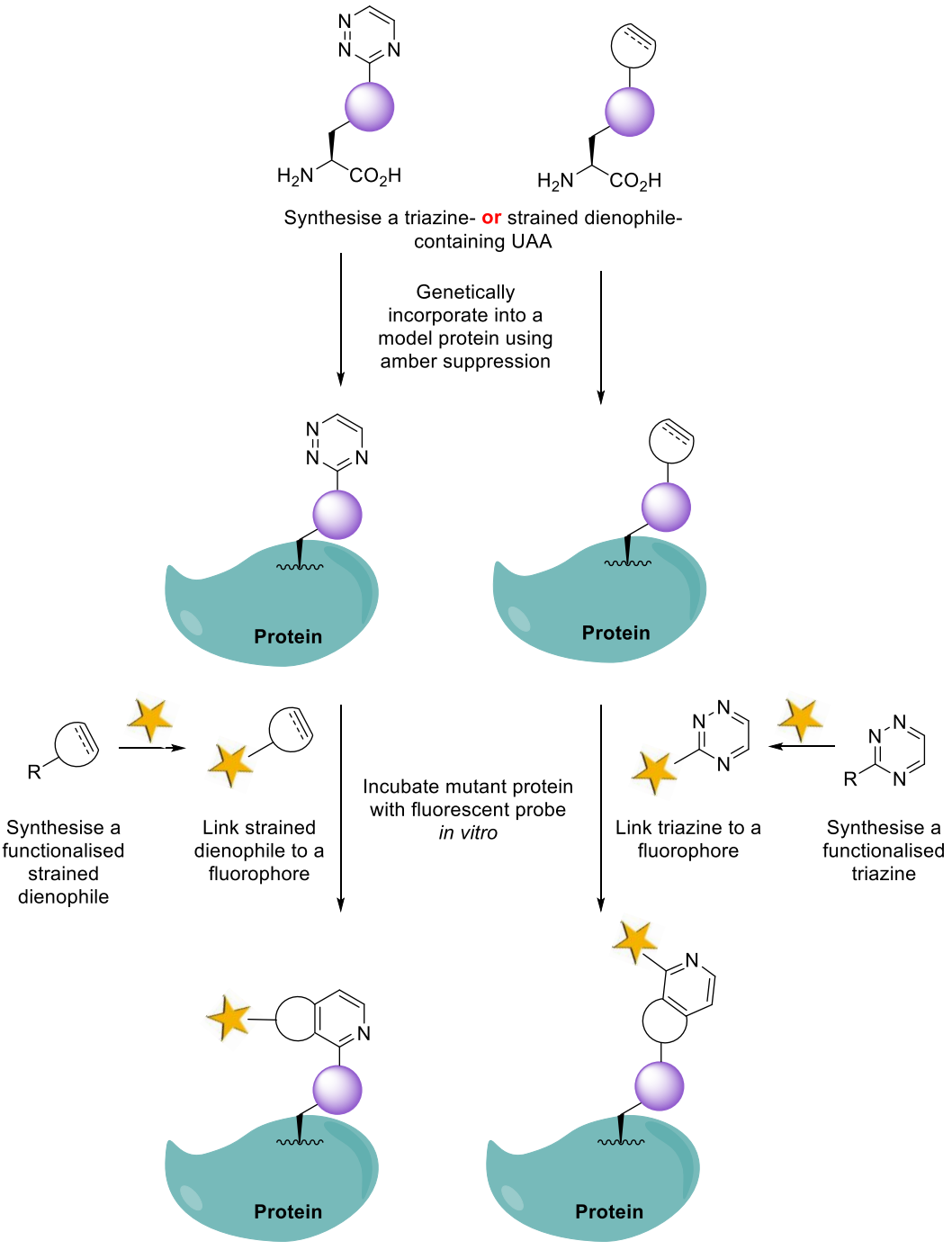


Figure 1.8: Outline of method to achieve second set of research objectives. An unnatural amino acid containing either a triazine (left) or strained dienophile (right) would be genetically incorporated into a model protein and incubated with a fluorescent probe containing the other reactive entity to generate a fluorescently labelled protein. If successful, this would demonstrate the utility of the triazine-SPIEDAC as a bioorthogonal labelling strategy.

Chapter 2 Use of τ-phosphotriazolylalanine as a Molecular Probe

2.1 Introduction

Following the generation of non-hydrolysable τ-phosphohistidine analogue, pTz **15** in the group, a natural progression was to incorporate **15** into peptides and use these peptides in a biophysical context to study proteins that have affinity for τ-phosphohistidine and protein-protein interactions that are contingent upon binding of τ-phosphohistidine.

2.2 Investigations into the Promiscuity of a Phosphotyrosine-Binding Protein

Previous work in the group using ITC had demonstrated that the Src Homology 2 (SH2) domain of the Growth Factor Receptor-bound 2 (Grb2) protein (a phosphotyrosine binding protein) binds to a peptide containing τ-pTz with micromolar affinity.[126] Following these preliminary results, an independent biophysical assay was required to determine whether the τ-phosphotriazole peptide bound to the phosphotyrosine-binding site of Grb2-SH2; and if it did, whether the protein simply binds phosphorylated amino acids or has selectivity towards τ-pTz (and phosphotyrosine). The role of Grb2-SH2 is typical to that of most phosphotyrosine binding modules. Selective binding of Grb2-SH2 to the τ-phosphotriazole peptide and not to analogous peptides containing other phosphoamino acids (apart from phosphotyrosine), would indicate that τ-phosphohistidine may recognise these domains *in vivo*. In this sense, Grb2-SH2 is being used as a prototype to study the interaction of τ-phosphohistidine to SH2 protein modules.

2.2.1 Growth Factor Receptor-Bound 2

Grb2 is a homodimeric adaptor protein that consists of one SH2 domain flanked by two SH3 domains and is known to play a crucial role in signal transduction through the linking of

receptor and cytoplasmic protein tyrosine kinases to the Ras signalling pathway (Figure 2.1).[127]

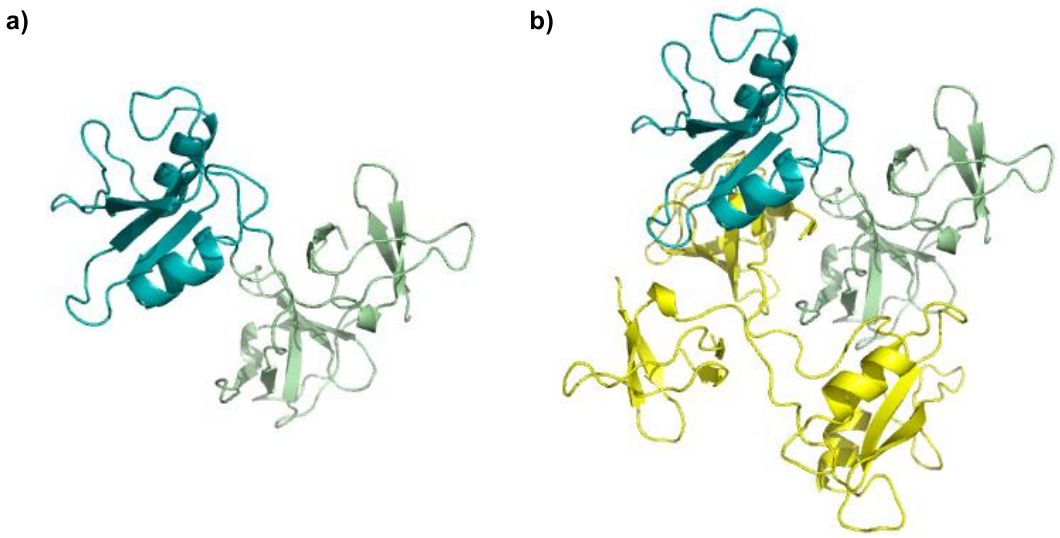


Figure 2.1: Crystal structure of Grb2 consisting of one SH2 domain (shown in blue) flanked by two SH3 domains (shown in green); b) Grb2 as its homodimer with the same colouring as a) for one monomer and the other monomer shown in yellow. PDB file: a) & b) -1GR1.

Grb2 links tyrosine kinases to the Ras signalling pathway through a translocation mechanism of activation.[128] As shown in Figure 2.2a, each SH3 domain of Grb2 binds to a short proline-rich sequence on the guanine nucleotide releasing factor for the Ras protein (Sos (Son of Sevenless)) in the cytoplasm. The resultant Grb2/Sos complex is then recruited to the cytoplasmic membrane where the Ras protein is located. The SH2-domain binds to a phosphotyrosine residue in the epidermal growth factor receptor (EGR) that has been autophosphorylated via activation by an epidermal growth factor (EGF). Increase in the effective concentration of the Grb2-Sos complex at the plasma membrane results in interaction of Sos with the Ras protein. The Ras protein becomes activated through an exchange of GDP to GTP, initiating a downstream kinase cascade; relaying signals to both the cytoplasm to control metabolic processes, and the nucleus to control gene expression. Grb2-SH2 can also directly and indirectly interact with other receptor and non-receptor kinases to stimulate the Ras signalling pathway.[127] For instance, association of the SH2 domain of Grb2 with the SH2-domain containing oncogenic protein Shc (which has been phosphorylated on a tyrosine residue by the oncogenic kinase *v*-Src) is also believed to activate Sos-mediated guanine nucleotide exchange on Ras (Figure 2.2b).[6] The use of SH2 domains to localise the proteins they are contained in through binding to short phosphotyrosine sequences is typical to many adaptor proteins, although many are less well characterised than Grb2-SH2.

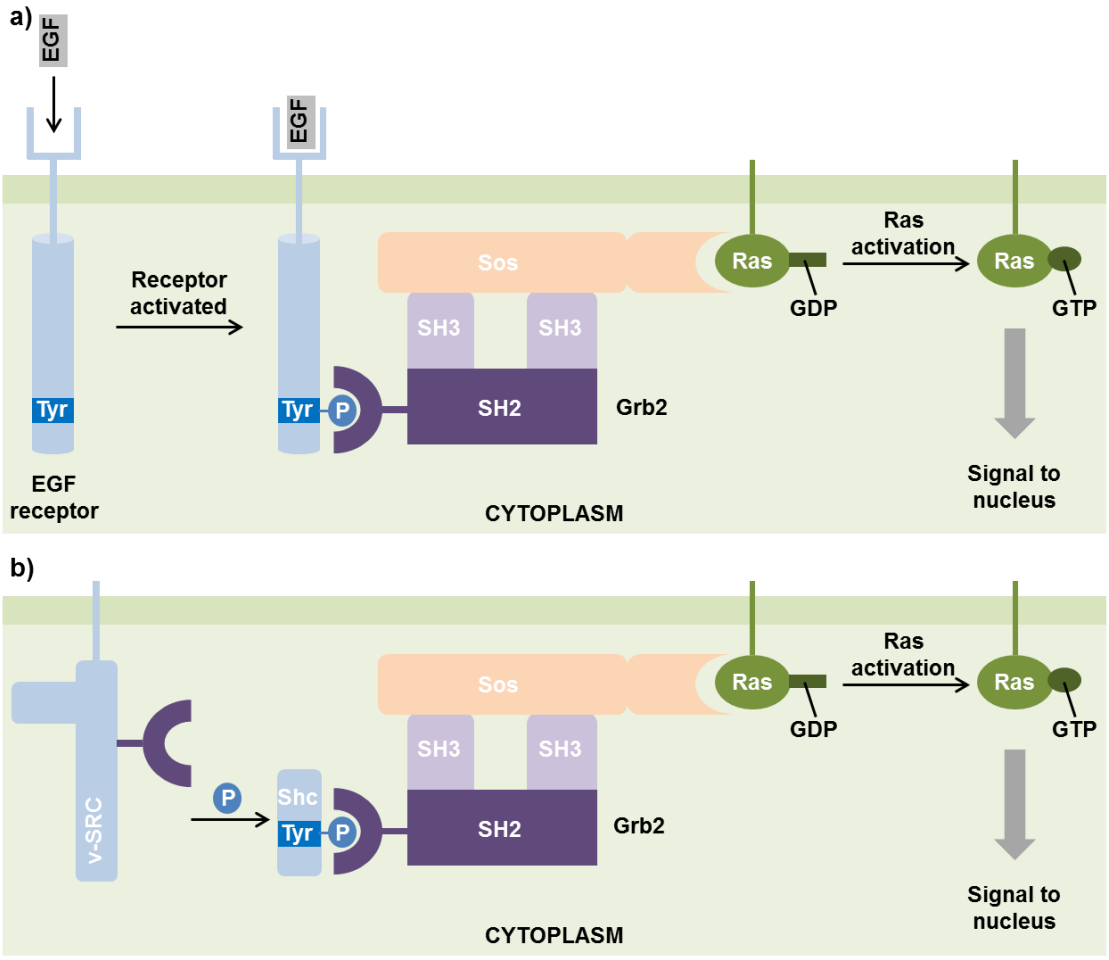


Figure 2.2: The role of Grb2 in signal transduction a) The SH2 domain binds a tyrosine residue on the EGF receptor that has been autophosphorylated through activation by an EGF. In doing so, Sos, which is bound to Grb2 through interaction with its SH3 domain, is recruited to the plasma membrane. Association of Sos and Ras stimulates the exchange of GDP to GTP, which in turns initiates a downstream signalling cascade ending in protein transcription; b) indirect interaction of the SH2 domain of Grb2 with non-receptor kinase Src (through a phosphorylated Shc intermediate) can also activate the Ras signalling pathway.

2.2.1.1 Binding of Grb2-SH2 to Phosphotyrosine-containing Peptides

The SH2 domain of Grb2 binds with high affinity to target proteins through short phosphotyrosine-containing sequences of the type pY(I/V)NX; with a hydrophobic residue at the +1 position and asparagine at the +2 position in the C-terminal direction from phosphotyrosine.[129] Weber and co-workers[130] have shown that an Shc-derived phosphotyrosine-containing synthetic peptide of the sequence Ac-SpYVNVQ-NH2 binds to Grb2-SH2 with a K_d of ca. 200 nM. The crucial residues involved in the binding interaction of Grb2-SH2 and phosphotyrosine-containing proteins have been elucidated by Ogura *et al.*[131] through NMR experiments, and are highlighted in Figure 2.3. As well as the cluster of positively charged residues that constitute the phosphate binding pocket of phosphotyrosine; the asparagine residue at the +2 position C-terminal to phosphotyrosine forms hydrogen bonds to Lys109 and Leu120 and is essential for specific binding to the SH2 domain of

Grb2. Moreover, the bulky side chain of Trp121 interacts with asparagine at the +2 position, causing the binding peptide to turn and preventing it from forming an extended structure. The strong recognition of asparagine (pTyr +2) as well as phosphotyrosine by Grb2-SH2 demonstrates that the protein recognises target phosphotyrosine-containing peptides in a sequence specific manner.

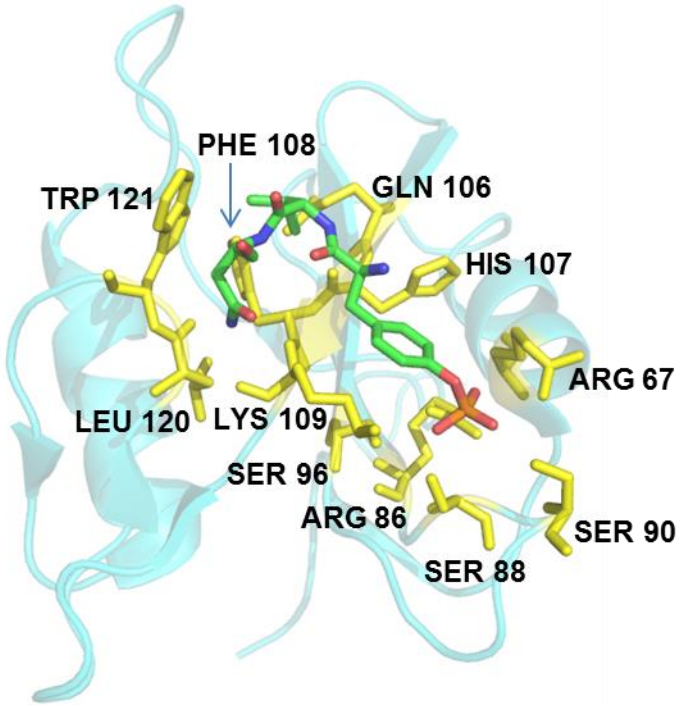


Figure 2.3: Structure of the SH2 domain (blue), with residues identified as being involved in binding depicted as yellow sticks.[131] The critical peptide sequence needed for binding (pTyr-Val-Arg) is shown in green, for simplicity, the rest of the peptide sequence has been omitted. PDB file 1QG1.

2.2.2 Previous Work in the Group

Work carried out by Dr Tom McAllister using ITC demonstrated that τ-phosphotriazolylalanine peptide **91** (Figure 2.4b) binds to the SH2 domain of Grb2 with ~2000 fold lower affinity then the analogous phosphotyrosine-containing peptide **90** (Figure 2.4a).[126] pTyr and pTz peptides **90** and **91** of consensus sequence Ac-Sp**X**VNVQ-NH_2 were synthesised using standard Fmoc SPPS and their binding to an N-terminally His-tagged SH2 domain of Grb2 (subcloned into a pET28a vector from a full-length Grb2-GST fusion protein)[i] was measured using ITC. Control pY peptide **90** was titrated into His_6–Grb2-SH2 to yield a sigmoidal binding curve that corresponded to 1:1 binding with K_d = 385 ± 41 nM (Figure 2.4a). Titration of pTz-containing peptide **91** into the protein gave a binding curve

[i] The Grb2-GST fusion protein aggregated at concentrations needed to measure binding of pTza peptide **91** by ITC.

with an affinity of 719 ± 28 μM (Figure 2.4b). To determine whether the reduced binding affinity of pTz peptide **91** to Grb2-SH2 was due to differing protonation states of the phosphoryl groups of **90** and **91**, the pK_a of both peptides were determined via NMR titrations. These experiments revealed a pK_a of 5.8 for pY peptide **90** and 5.95 for pTz peptide **91** (data not shown); hence both phosphoryl groups will be dianionic under the conditions of the ITC experiments (pH 7.4). It is therefore unlikely that the change in binding affinity is due to different protonation states.

The discovery that Grb2-SH2 binds to pTz peptide **91** was serendipitously made whilst investigating whether a peptide containing an analogue of phosphotriazole, phosphohomotriazole (phTz) peptide **92** could act as a mimic of phosphotyrosine and hence have affinity for the SH2 domain of Grb2 (Figure 2.4c). The extension of the linkage between the backbone of the peptide and the triazole by one carbon in comparison to parent pTz peptide **91** was expected to present the phosphoryl group in a similar orientation to that of phosphotyrosine. Moreover, Hofmann *et al.*[132] had previously demonstrated that a synthetic peptide containing phosphoarginine **5** binds to the SH2 domain of Src with a reduced affinity of ~4000 fold less than the analogous pY peptide. Phosphohomotriazole is similar in structure to phosphoarginine **5** and therefore it was highly possible that phTz peptide **92** would show affinity towards SH2 domains. Surprisingly, titration of **92** (of the same consensus sequence of **90** and **91**) into His_6–Grb2-SH2 using ITC showed no binding (Figure 2.4c).

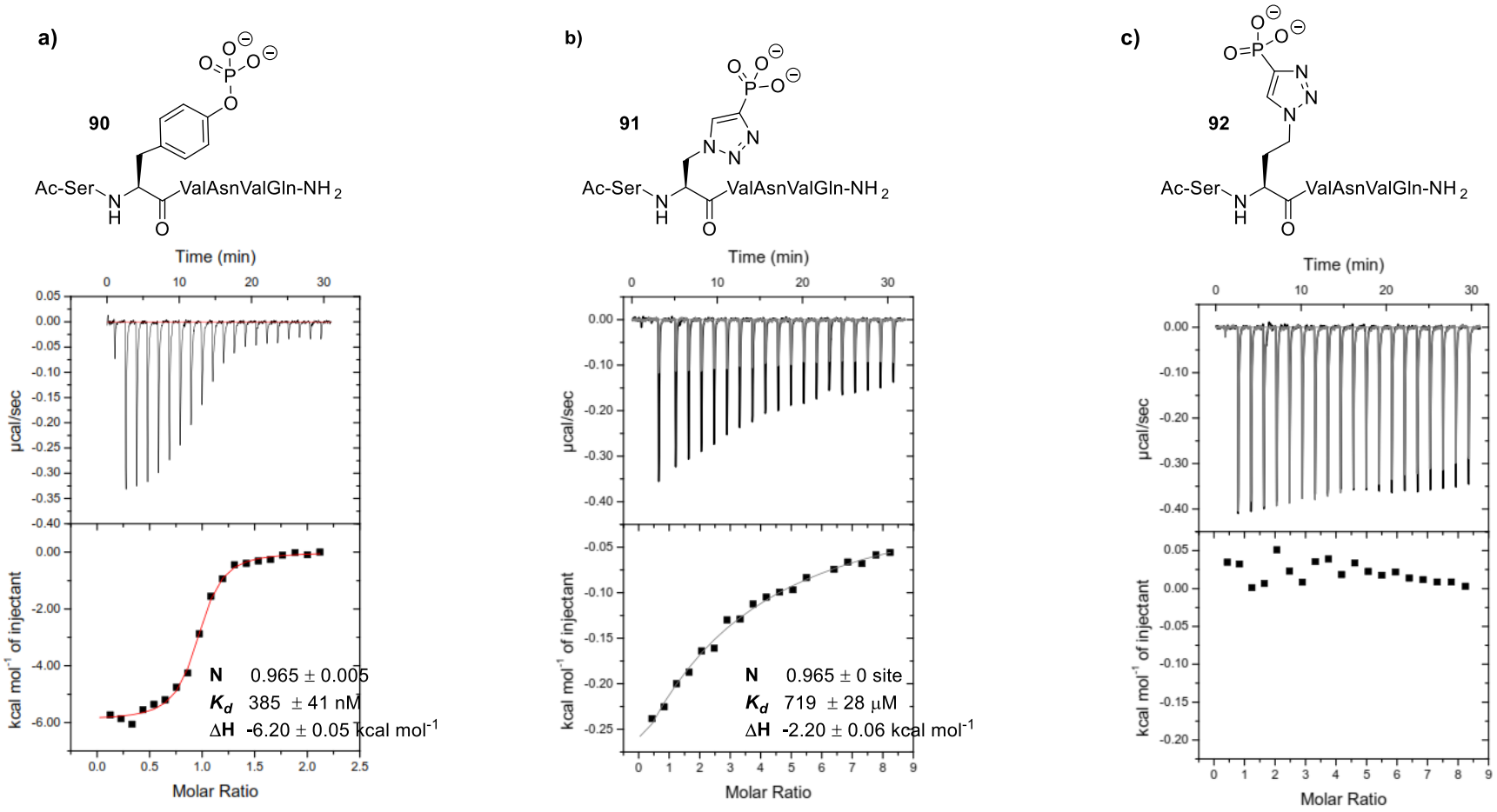


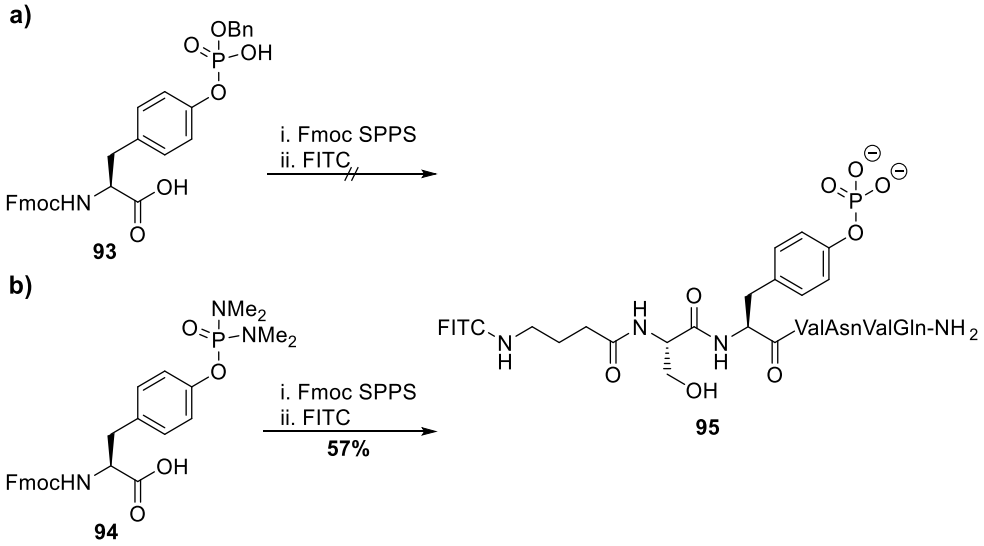
Figure 2.4: ITC experiments titrating a) pY peptide **90**, b) pTz peptide **91** and c) phTz peptide **92** into His6–Grb2-SH2 gave equilibrium binding constants of (385 ± 41) nM for **90** and (719 ± 28) µM for **91**. Binding of **92** to His6–Grb2-SH2 was not observed.

2.2.3 Further Investigations into the Phosphotyrosine-Binding site of Grb2-SH2

Following the discovery that pTz peptide **91** had affinity towards Grb2-SH2, the next step was to use an alternative biophysical technique to confirm that pTz peptide **91** was binding to the same site on Grb2-SH2 as phosphotyrosine peptide **90**, and to investigate the promiscuity of the protein to other phosphoamino acids. To this end, it was decided to conduct a series of fluorescence anisotropy competition experiments. For these assays, a phosphotyrosine-containing fluorescent probe of sequence FITC-GaSp**X**VNVQ-NH_2 and a series of phosphoamino acid containing peptides of consensus sequence Ac-Sp**X**VNVQ-NH_2 (p**X** = phosphoamino acid) were required.

2.2.3.1 Synthesis of a Fluorescent Probe

Phosphotyrosine-containing peptide with N-terminal FITC **95** was synthesised using standard Fmoc-based SPPS and reaction with FITC overnight in the dark on Rink amide resin (Scheme 2.1). The peptide was cleaved from the resin using a standard cleavage cocktail (TFA/H_2O/TIS 95:2.5:2.5). Interestingly, use of mono-benzyl phosphotyrosine **93** resulted in a mixture of unassignable peptide products; believed to be the result of reaction of the free phosphonate hydroxyl group and the highly electrophilic isothiocyanate carbon (Scheme 2.1a). A globally protected phosphotyrosine was therefore needed. *bis*(dimethylamino)phosphotyrosine **94** is commercially available and accordingly was used in place of monobenzyl phosphotyrosine **93**. The dimethylamino- protection for the phosphonate of **94** has reduced lability in acidic conditions and will not be removed when swelling in a standard cleavage cocktail for two hours. Accordingly, a two-step deprotection strategy was employed in which the peptide was cleaved from the resin using a standard cleavage cocktail for two hours, 10% (*v/v*) more H_2O was added and the mixture swelled overnight. After purification using anion-exchange chromatography, pY probe **95** was afforded in 57% yield.



Scheme 2.1: Synthesis of pY probe **95** by Fmoc SPPS; a) Incorporation of monobenzyl phosphotyrosine **93** resulted in a mixture of unassignable peptide products; b) Use of *bis*(dimethylamino) phosphotyrosine **94** gave the desired fluorescent peptide **95** in a yield of 57%.

2.2.3.2 Synthesis of Phosphotriazoles Compatible with the Fmoc-Strategy for SPPS[i]

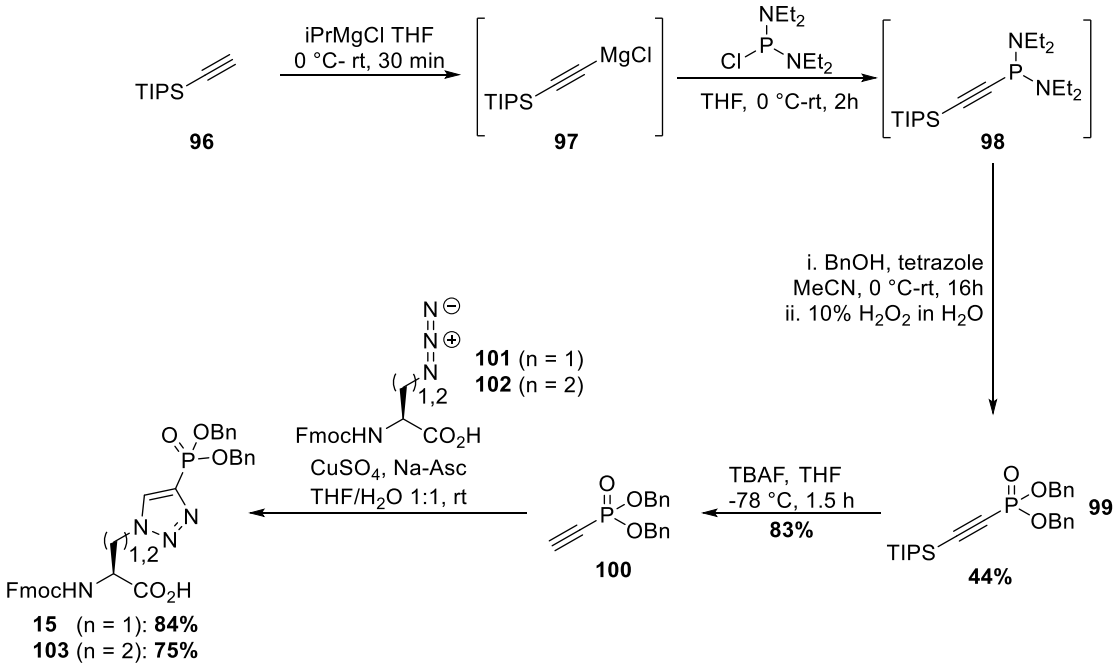
In order to generate pTz peptide **91**, the synthesis of Fmoc-protected pTz **15** was required. It seemed prudent to also use competitive fluorescence polarisation to conclusively confirm that phTz peptide **92** did not bind to the SH2 domain of Grb2 (as previously determined by ITC (Figure 2.4c)); thus Fmoc-protected phTz **103** was also synthesised.

Following the route of McAllister *et al*.,[25] triazoles **15** and **103** were synthesised in 6 steps. TIPS acetylene **96** was converted to reactive intermediate **97** using isopropylmagnesium chloride in THF at 0 °C; addition of *bis*(diethylamino)chlorophosphine to the reaction mixture and warming to room temperature afforded intermediate **98**. Protected alkyne **96** was used to prevent side reactions involving the terminal alkyne.[ii] Following a solvent switch to acetonitrile the diethylamino- groups of intermediate **98** were displaced using benzyl alcohol together with the dropwise addition of 1*H*-tetrazole at 0 °C. The reaction was allowed to warm to room temperature and stirred overnight. 1*H*-tetrazole was required for phosphoroamidate activation (through formation of the tetrazolide). Oxidation of the corresponding benzyl phosphite by washing the organics with 10% hydrogen peroxide gave benzyl phosphonate **99** in a yield of 44%. The TIPS group was removed using TBAF in

[i] Synthesis of phTz **103** was carried out by Dr. Tom McAllister

[ii] McAllister *et al*.[25] described the concurrent formation of Michael addition-type products when using the unprotected version of **96** with nucleophiles such as benzyl alcohol.

THF at -78 °C to yield phosphoalkyne **100** in 83% yield. Using copper(II) sulfate pentahydrate and sodium ascorbate in 1:1 THF/H_2O, triazoles **15** and **103** were generated through a cycloaddition reaction between phosphoalkyne **100** and azidoalanines **101** or **102** in yields of 84% and 75% respectively.



Scheme 2: Synthesis of pTz **15** and phTz **103** through the (3 + 2) cycloaddition of dibenzyl phosphoalkyne **100** to azidoalanine **101** or **102** respectively. Phosphoalkyne **100** was generated in 5 synthetic steps from commercially available TIPS-acetylene **96**.

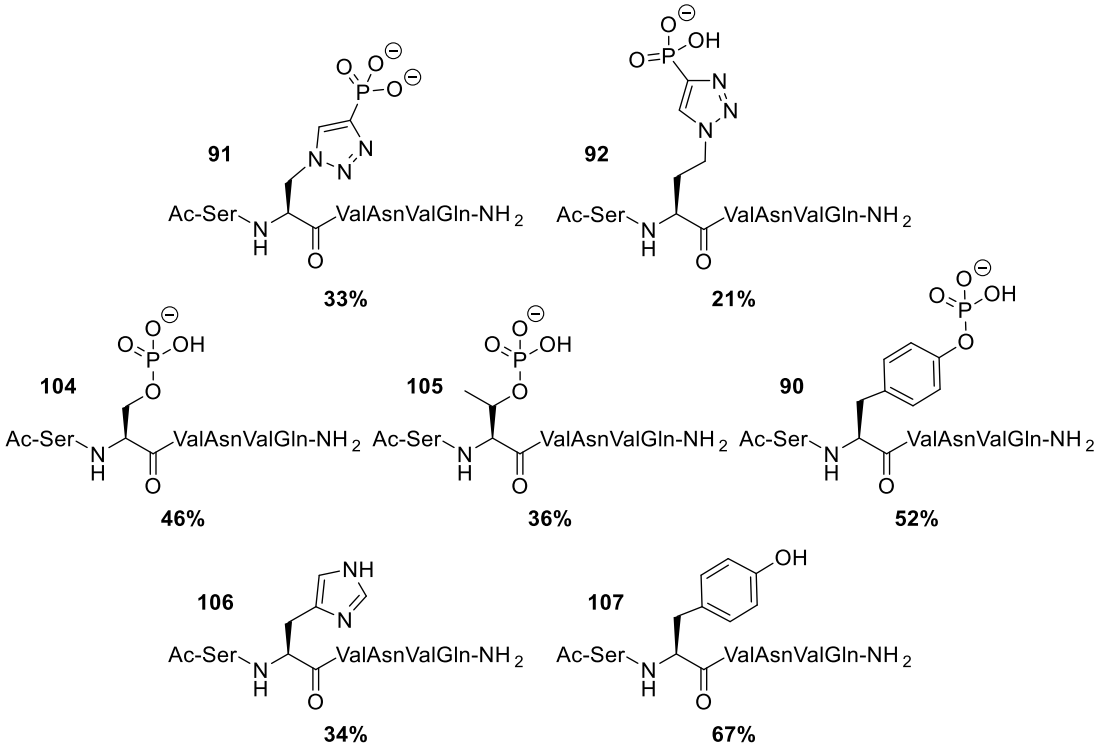
2.2.3.3 Synthesis of Phosphoamino Acid-containing Peptides[iii]

To investigate whether Grb2-SH2 selectively binds to τ-phosphohistidine sequences or to phosphorylated amino acids in general, the affinity of pS peptide **104** and pT peptide **105** towards Grb2-SH2 would also be determined using competitive fluorescence anisotropy. The affinity of pArg- and pK- containing peptides towards the SH2 domain of Grb2 was not investigated due to previous work (Figure 2.4c) indicating that the structurally similar phTza peptide **92** did not bind (to Grb2-SH2).

Peptides of sequence AcS**X**VNVQ-NH_2 (where **X** is either pTz **91**, phTz- **92**, pSer- **104**, pThr- **105**, pTyr- **90**, His **106** or Tyr **107**) were synthesised using a standard Fmoc SPPS protocol (5 eq. of the commercially available amino acids, 5 eq. HCTU and 10 eq. DIPEA and 3/3/6 for pTza **15** and phTz **103**); cleaved from the resin using a standard cleavage cocktail (TFA:H_2O:TIS 95:2.5:2) and purified using anion-exchange chromatography. pTyr

[iii] Peptides **92** and **104**-**107** were synthesised by Dr Tom McAllister.

peptide **90** and His and Tyr peptides **106** and **107** were to be used as positive and negative control(s) respectively.



Scheme 2.2: Peptides synthesised by Fmoc SPPS to be used in fluorescence anisotropy competition assays. pTz **15** and phTz **103** were incorporated into peptides **91** and **92**. All other amino acids were commercially available.

2.2.3.4 Lysis and Purification of His_6–Grb2-SH2

Following overexpression from pET28a-Grb2-SH2 in *E. coli* C41(DE3) cells, the cells were lysed and the inclusion bodies that were collected with the cellular debris solubilised in 8M urea overnight. The solubilised protein was applied to a nickel affinity column, refolded through washing with *Phosphate Buffer 1*, and purified using a standard Ni-NTA protocol. Fractions containing His_6–Grb2-SH2 were determined by measuring UV absorbance at 280 nm and those containing protein were combined, concentrated and purified further using size exclusion chromatography (SEC) (Figure 2.5a). The SEC trace showed two peaks eluting at 175 ml and 220 ml; the former peak corresponded to a domain-swapped dimer and the latter was the correctly folded His_6–SH2Grb2 as previously shown by Benfield *et al.*[133,iv] Thus, fractions corresponding to the later peak were concentrated. Analysis by SDS-

[iv] Dr Tom McAllister had previously ran ITC experiments with protein corresponding to both of these peaks shown on the SEC trace to obtain K_d values for their binding to pY peptide **90**. He found that the peak eluting at 220 ml had a K_d corresponding to correctly folded His_6–Grb2-SH2.

PAGE (Figure 2.5b) and mass spectrometry (Figure 2.5c) revealed a protein corresponding to a mass of 12999.6 consistent with the mass of His_6–Grb2-SH2 after cleavage of the N-terminal methionine (expected mass 13000.6 Da).

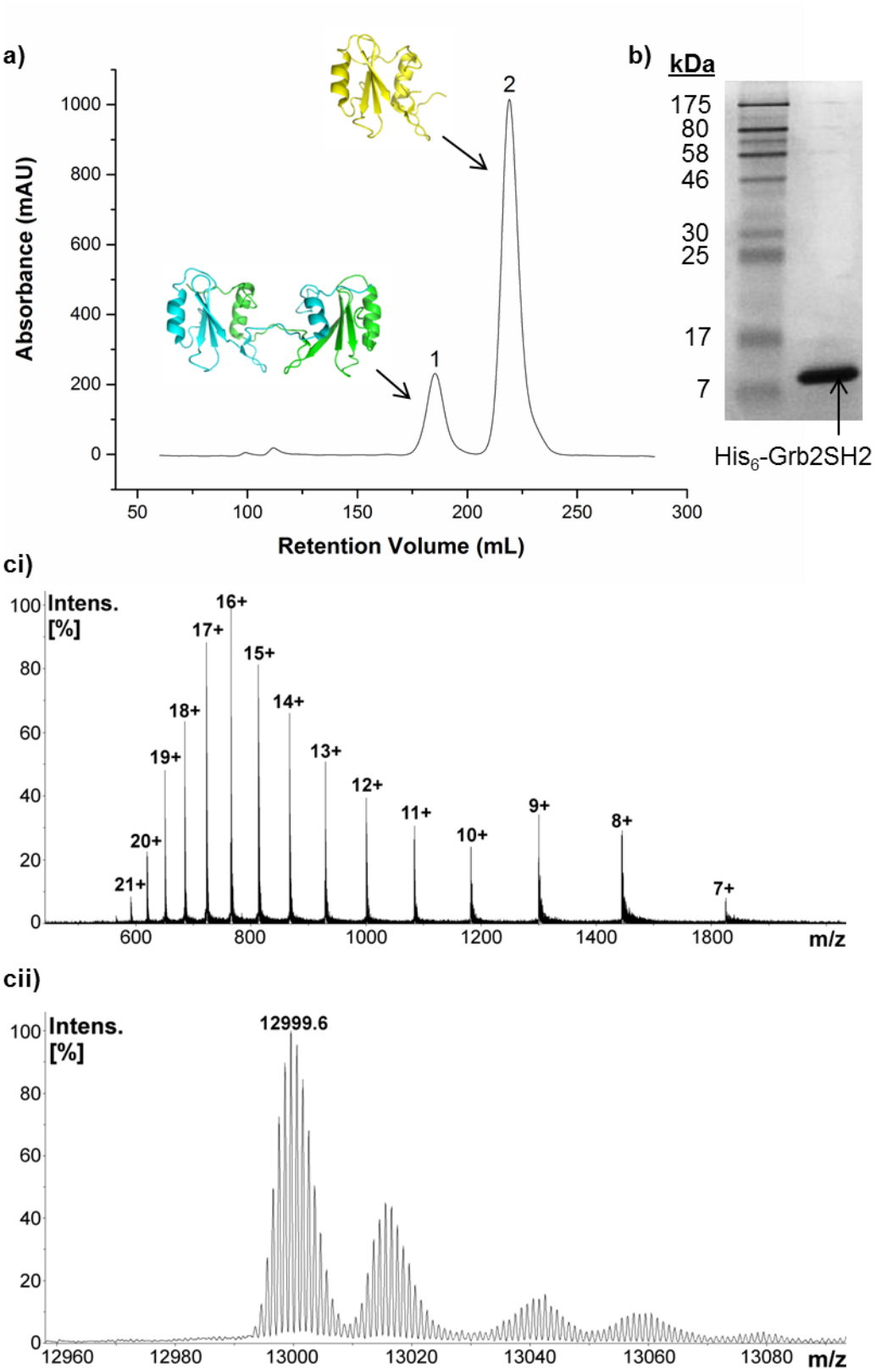


Figure 2.5: a) SEC trace of fractions containing His_6–SH2-Grb2. Peak 1 corresponds to a domain swapped dimer, peak 2 is monomeric His_6–Grb2-SH2; b) 12% SDS-PAGE gel of His_6–SH2-Grb2 after purification by SEC; ci) HRMS trace and cii) deconvoluted HRMS trace for His_6–SH2-Grb2 gave a measured mass 12999.6 Da, consistent with the mass of His_6–SH2-Grb2 after cleavage of the N-terminal methionine (expected mass 13000.6 Da). PDB file: 2H46

2.2.3.5 Development of a Fluorescence Anisotropy Competition Assay

2.2.3.5.1 An Introduction to Fluorescence Anisotropy

Fluorescence anisotropy determines the extent of decorrelation of plane polarised light of a sample and can be used to obtain information about protein binding interactions.[134] It is particularly appropriate for the studying of protein-protein or protein-peptide interactions due to its sub-nanomolar sensitivity and its applicability at low concentrations; providing a suitable fluorophore is used for the system in question.

If a sample of fluorescent molecules is illuminated with plane polarised light, only the subset of molecules that have absorption dipoles aligned to the plane of polarisation are raised to an excited electronic state.[135] The excited state will exist normally for a few to tens of nanoseconds, and after emission of a photon, the molecules will return to their ground electronic state. The probability of excitation (P_{ex}) that a fluorescent molecule will absorb a photon is proportional to the angle between the polarisation of exciting light and the excitation transition dipole (φ).

$$P_{ex} \propto \cos^2\varphi \qquad \textbf{(1)}$$

The probability of excitation is maximal if both the plane of polarisation and the transition dipole moment of excitation are aligned ($\varphi = 0$). If both the plane of polarisation and the transition dipole moment of excitation are perpendicular ($\varphi = 90°$), P_{ex} will be zero. This is known as selective photoexcitation and can be used to look at the relative anisotropies of associated and dissociated protein complexes through the differences in their plane polarised orientated subpopulations. If during the excited state lifetime, there is no rotation of the probe molecule and excitation and emission dipoles are aligned; the emitted fluorescence will be polarised in the same direction. Conversely, if the fluorescent molecule rotates during the excited state lifetime, the plane of polarisation of emitted fluorescence will rotate and be depolarised relative to the excitation light.

The rotation of a molecule on a characteristic fluorescent timescale can be theoretically calculated.[134] Assuming a protein is globular and rigid, the rotational correlation time (Θ) can be approximately linked to its molecular weight (MW) by the following equation:

$$\theta = \frac{\eta MW}{RT}(\nu + h) \qquad \textbf{(2)}$$

Where η = solvent viscosity (P), R = 8.31 x 10^7 erg Mol^{-1} K^{-1}, T = absolute temperature (K), ν = partial specific volume and h = the degree of hydration. If a spherical rigid protein monomer of 30 kDa at 298 K had typical values of ν = 0.74 ml/g and h = 0.2 $g_{water}/g_{protein}$ and was in a solvent of η = 0.01 P, its rotational correlation time would be ~11 ns. Similarly, in these conditions a 1 kDa substrate would have Θ = ~0.4 ns. Fluorescein has an excited state

lifetime of 4 ns; shorter than the rotational correlation time for the 30 kDa protein but longer than that for the 1 kDa substrate. Hence, excitation of a fluorophore attached to a small peptide or protein with plane polarised light will result in rotation of the molecule a number of times throughout the excited state lifetime as its rotational correlation time is below that of the fluorescence timescale. Consequently a low level of light relative to the excitation plane of polarisation will be emitted (Figure 2.6a). The corresponding peptide-protein complex, which is significantly larger, will have a slower rotation than the excited state lifetime and a higher level of plane polarised light will be emitted (Figure 2.6b). Therefore, an increase in anisotropy signal will show an increase in ligand-protein association. Exciting a sample with polarised light both parallel and perpendicular to the plane of excitation will mean that the extent of a molecules rotation can be assessed and anisotropy values can be calculated. Observing a system over a range of concentrations of one of the binding partners, or a competitive binder, will enable the generation of binding curves required to attain parameters such as the K_d and IC_{50}.

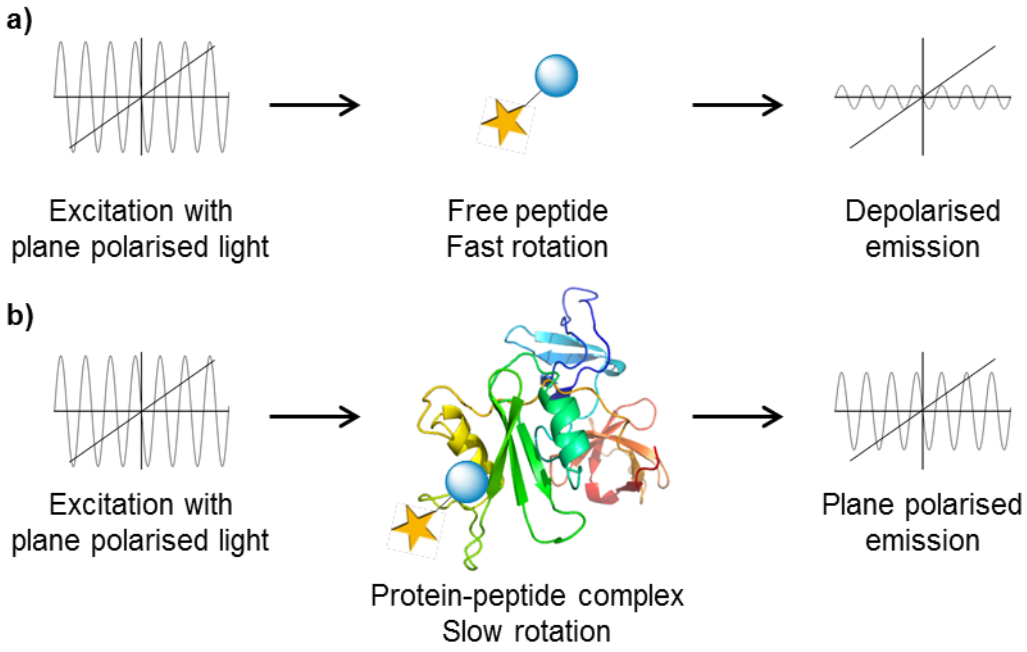


Figure 2.6: Fluorescence anisotropy: Excitation with plane polarised light of a) a fluorescent peptide and b) a fluorescent peptide-protein complex. Fast rotation of the free peptide on the fluorescent timescale will result in depolarised emission. Slow rotation of the protein-peptide complex will lead to plane polarised emission.

2.2.3.5.2 Fluorescence Anisotropy Experiments

*2.2.3.5.2.1 Determination of the Binding Affinity of FITC-pY Peptide **95** to Grb2-SH2*

To determine the binding constant (K_d) of fluorescent pY probe **95** to the SH2 domain of Grb2, a 2–fold dilution series of purified His_6–Grb2-SH2 in ITC buffer containing 100 nM **95** was prepared and the fluorescence intensity was measured in both parallel and

perpendicular channels. The total intensity I_T and anisotropy r were calculated for each data point through application of equations **(3)** and **(4)**: where $I_{parallel}$ is the intensity of light emitted with parallel polarisation to the laboratory source, $I_{perpendicular}$ is the intensity of light emitted with perpendicular polarisation, and G is the instrumental grating factor which allows for the variation in responsivity of an instrument to light polarised in the perpendicular direction. G was approximated under the assumption that the free peptide **95** had anisotropy $r = 0$. A logistic model was then used to fit data points and their standard deviation (averaged from each triplicate measurement after subtraction from a blank average) to give maximum and minimum limiting values for I_T and r (Figure 2.7a and b respectively).

$$I_T = I_{parallel} + 2GI_{perpendicular} \quad \textbf{(3)}$$

$$r = \frac{I_{parallel} - GI_{perpendicular}}{I_T} \quad \textbf{(4)}$$

Under the assumption that plane polarised light is exciting an ensemble of randomly orientated dipoles, the anisotropy must be averaged over all possible orientations and weighted by the probability of a dipole being excited at that orientation. Therefore the maximal fluorescence anisotropy value for one photon should be equal to 0.4.[134] Binding of fluorescent pY peptide **95** to His_6–SH2-Grb2 resulted in a maximum anisotropy value of 0.08 (Figure 2.7b), a significant deviation from the theoretical maximum value. Hence, considerable fluorescence quenching is occurring upon complex formation. Anisotropy values were therefore converted to the fraction bound (F) by application of equation **(5)** where $\lambda = I_{bound} / I_{free}$ using maximal and minimal anisotropy values (r_{max} and r_{min}), before being fit with a hyperbolic binding model to give a K_d of 283 ± 18 nM (Figure 2.7c). This result is consistent with the binding constant measured previously in ITC experiments (Figure 2.4).

$$F = \frac{r - r_{min}}{\lambda (r_{max} - r - r_{min})} \text{ where } \lambda = \frac{I_{bound}}{I_{free}} \quad \textbf{(5)}$$

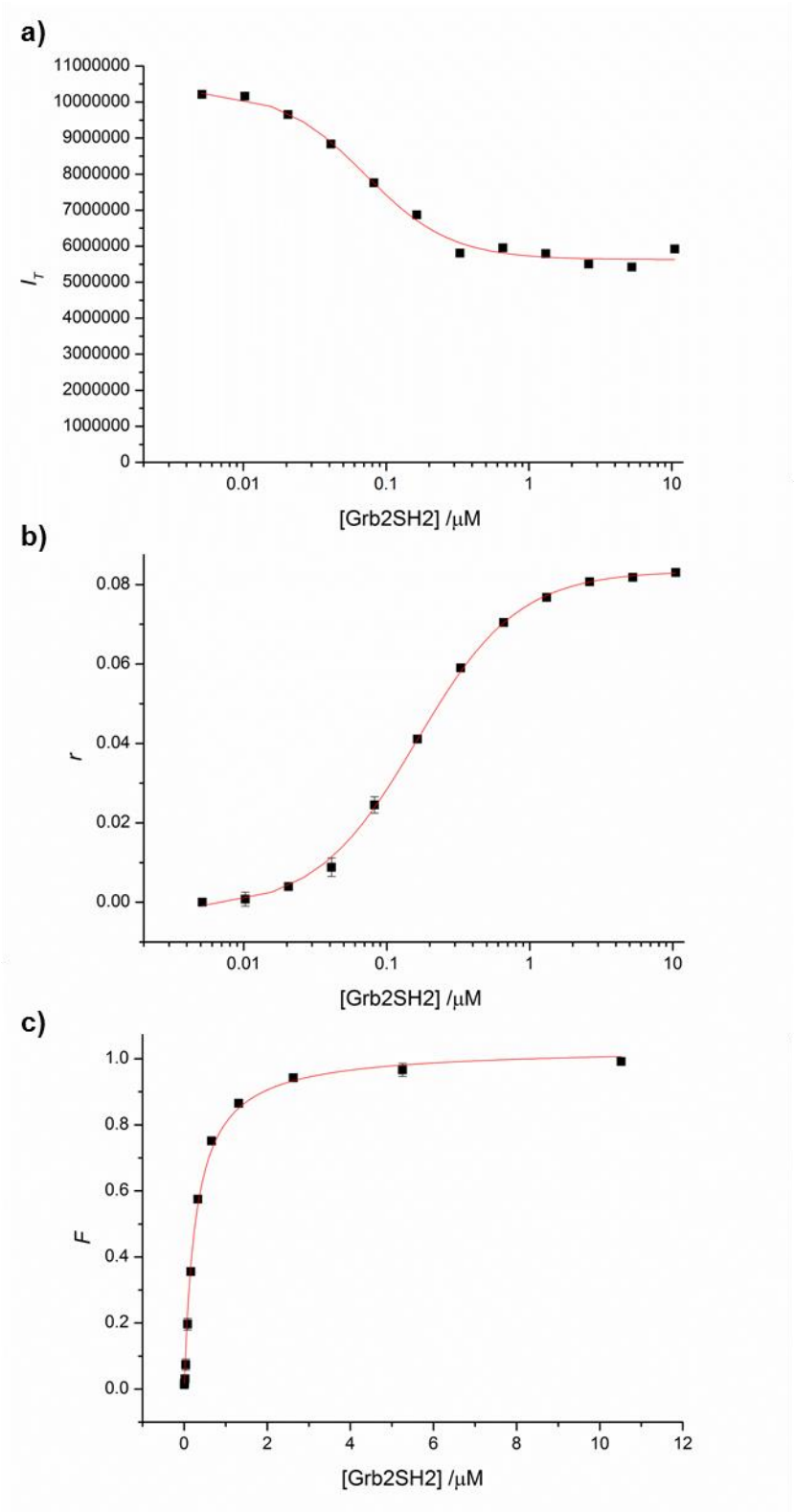


Figure 2.7: Fluorescence anisotropy analysis of the binding of fluorescent peptide **95** to His_6–Grb2-SH2: a) Total fluorescence as a function of [Grb2-SH2] with 100 nM **95**; data were fit using a logistic model to give limiting values of I_{bound} 10.34 abu and I_{free} 5.63 abu; b) Dependence of anisotropy upon [Grb2-SH2] revealing r_{min} 0.0021 abu and r_{max} 0.083 abu; c) Fraction bound **95** as a function of [Grb2-SH2] gave a binding constant of 283 ± 18 nM.

2.2.3.5.2.1.1 Competition Assays

For the fluorescence anisotropy competition experiments, peptides pTz **91**, phTza **92**, pSer **104**, pThr **105**, pTyr **90**, His **106** and Tyr **107** were dissolved in ITC buffer and a 2.5–fold or 3–fold dilution series of each peptide was prepared in triplicate containing 200 nM of peptide probe **95** and 175 nM of His_6–Grb2-SH2. A binding response (displacement of peptide probe **95**) was observed only for pTz peptide **91** and pY peptide **90**. Accordingly the anisotropy for each data point was converted to the fraction bound via application of equation **(5)** using the maximum and minimum values for fluorescence intensity and anisotropy determined for binding of fluorescent pY probe **95** to Grb2-SH2. Data were globally fit using a logistic binding model to yield half maximal inhibitory concentrations of 363 ± 25 μM and 442 ± 66 nM for pTz peptide **91** and pY peptide **90** respectively (Figure 2.8a). The IC_{50} value calculated for pTz peptide **91** was slightly lower than anticipated when compared to the binding constant of 700 μM calculated from ITC experiments (Figure 2.4); indicating that initial ITC analysis might have underestimated the affinity of pTz peptide **91** for Grb2-SH2.

As can be seen in Figure 2.8b, no displacement of fluorescent peptide **95** was observed for peptides pS **104**, pT **105**, phTz **92**, Tyr **106** or His **107**. These results indicate that the SH2 domain of Grb2 does not simply have affinity for phosphoryl groups orientated at the appropriate position in a peptide sequence and that Grb2-SH2 does not just bind to an aromatic ring. It is noteworthy that phTz peptide **92** does not have affinity to Grb2-SH2; especially when considered alongside reports from Hoffman *et al.*,[132] who demonstrated that the structurally similar pArg **5** binds to the SH2 domain of Src kinase. Moreover, it is interesting that extension of the linkage between the triazole and peptide backbone by just one carbon atom results in total elimination of binding.

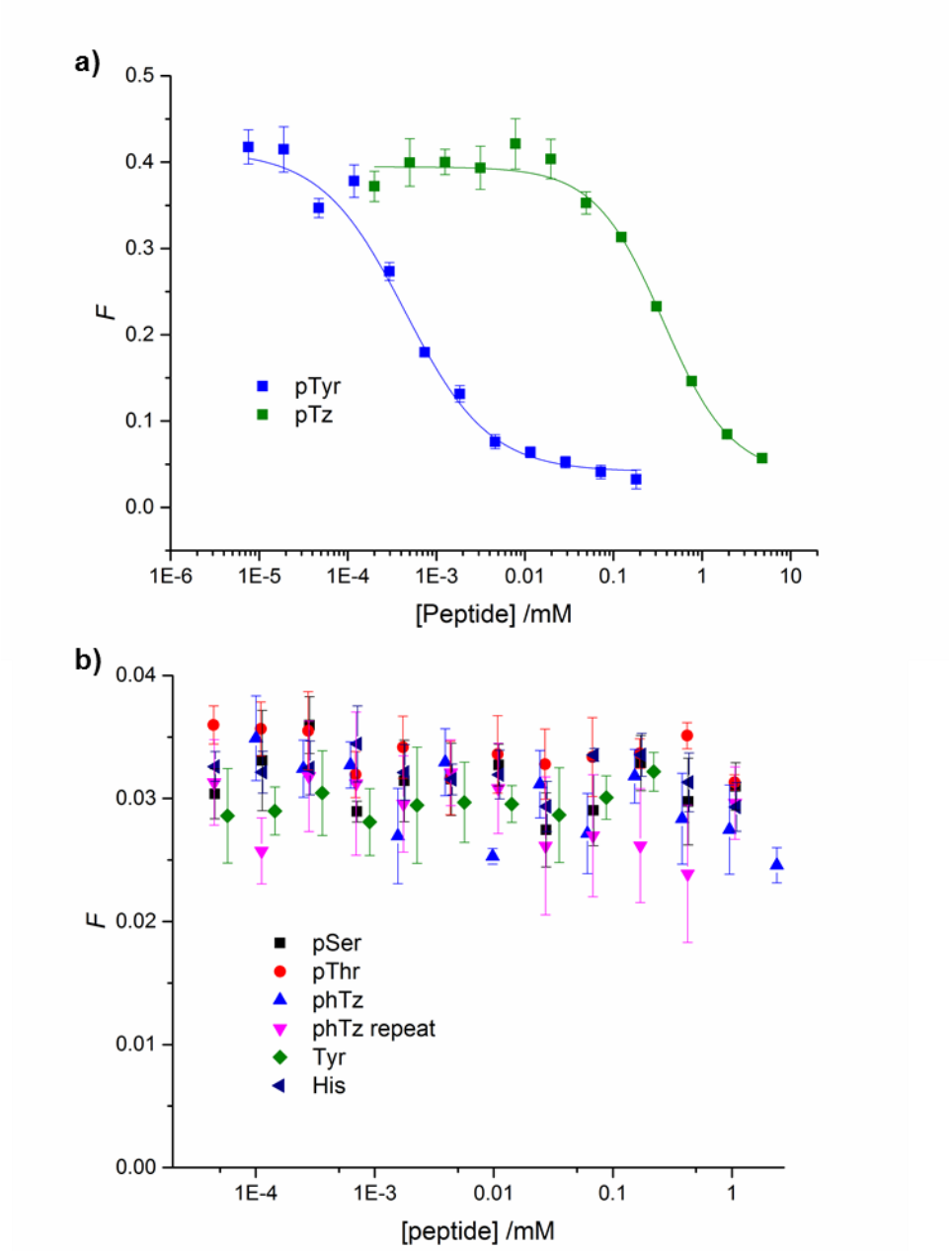


Figure 2.8: Fluorescence anisotropy competition experiments. All peptides serially diluted and mixed with 175 nM His_6–Grb2-SH2 and 200 nM pY probe **95**. a) Competition of pTz and pY peptides **91** and **90** for binding of Grb2-SH2 over pY probe **95** reveal half maximal inhibitory concentrations of 363 ± 25 μM and 442 ± 66 nM respectively; b) No competition is evident for peptides **92** and **104**-**107**.

Results of the fluorescence anisotropy competition experiments confirm that a known τ-phosphohistidine mimic binds to a canonical phosphotyrosine binding domain with ~1000 fold weaker affinity then the analogous phosphotyrosine-containing sequence. The mechanism by which Grb2-SH2 links receptor tyrosine kinases to the Ras signalling pathway is well established and it is therefore unlikely that the SH2 domain of Grb2 interacts with τ-phosphohistidine *in vivo*. However, it is a possible that a subset of other SH2 domains that have less well-defined targets may bind τ-phosphohistidine-containing

sequences. The prevalence of phosphotyrosine binding modules in human signalling proteins is relatively high, and the binding affinity of phosphotyrosine peptide sequences to these target protein modules can vary from high micromolar to low nanomolar.[136] Hence, the micromolar affinity of pTz peptide **91** to the SH2 domain of Grb2 does not mean that it is an insignificant interaction. Binding of τ-pHis-containing sequences could therefore be a characteristic feature of many phosphotyrosine binding modules. Further investigations to establish the incidence of this are needed.

This result contrasts studies by Senderowicz *et al.*[137] who demonstrated that a peptide derived from FGFR1 containing a τ-phosphohistidine residue did not bind to the SH2 domain of phospholipase C-γ1 (FGFR1 is known to bind to the SH2 domain of PLC-γ1 when its tyrosine residue is phosphorylated). However, it should be noted that in this study, the conditions used prevented observation of low affinity binding.

2.2.4 Conclusions

It has been unequivocally demonstrated that peptide **91** containing a known τ-phosphohistidine analogue, interacts with the SH2 domain of Grb2 using two distinct biophysical techniques. Conversely, it has been shown that with the exception of phosphotyrosine peptide **90**, a number of other phosphoamino acid-containing peptides do not bind to Grb2-SH2. This includes phosphohomotriazole peptide **92**, an analogue conjectured to mimic phosphotyrosine (a known, high affinity substrate of Grb2-SH2). From a functional perspective, the fact that τ-pTz peptide **91** binds selectively to Grb2-SH2 suggests that τ-phosphohistidine-containing sequences may be recognised by a subset of SH2 binding modules *in vivo*.

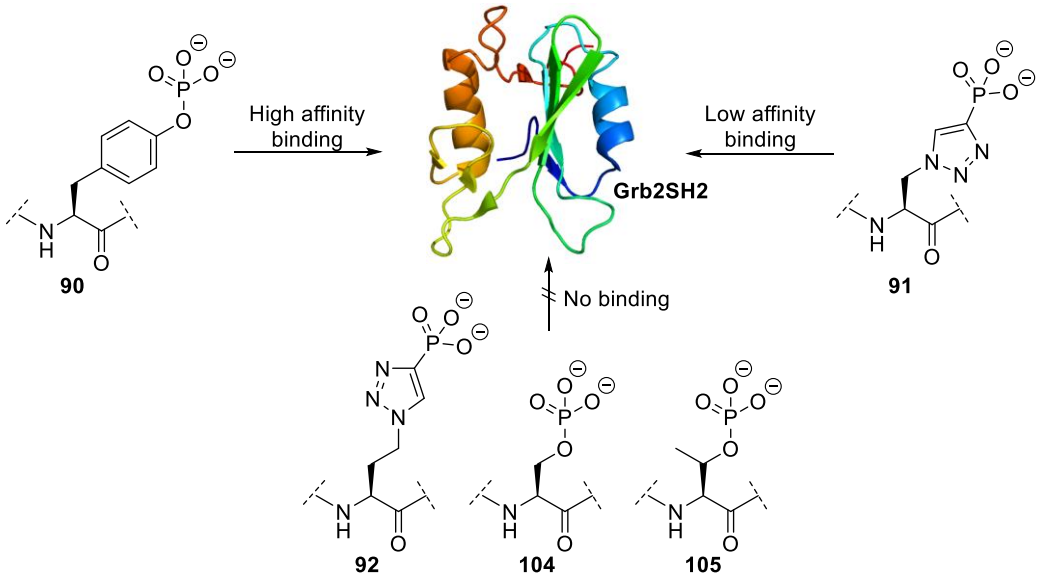


Figure 2.9: Summary of binding interactions of Grb2-SH2 to phosphoamino acid-containing peptides.

2.3 Examining the Interaction between PPDK and its Regulatory Protein, PDRP

τ-Phosphotriazole can also be used as a probe molecule to study known τ-phosphohistidine-meditated protein-protein interactions. As described in Section 1.1.4.2.1, interaction of pyruvate, orthophosphate dikinase (PPDK) and PDRP (PPDKs regulatory protein) is dependent upon phosphorylation of a catalytic histidine residue.[41] Work carried out by Dr. Zhenlian Ling demonstrated that PDRP is unstable and prone to precipitation and is therefore unsuitable for use in experiments to study binding interactions of this protein. Fortunately, an analogous phosphohistidine-mediated regulatory system exists in *E. coli* in which the protein equivalent to PPDK is known as phospho*enol*pyruvate synthase (PPS), and its regulatory protein, analogous to PDRP, is YdiA.[138] PPS is known to catalyse the formation of phospho*enol*pyruvate (PEP) from pyruvate in a similar mechanism to that of PPDK. Furthermore, there is a high level of sequence homology around the active sites in PPDK and PPS.[139] The following section focuses on the use of biophysical techniques to study the PPS-YdiA regulatory system in an attempt to further understanding of the binding interaction between these two proteins and the analogous PPDK/PDRP system.

2.3.1 Studying the Binding Interaction between YdiA and PPS

2.3.1.1 Previous Work in the Group[i]

Preliminary investigations into the binding interaction between PPS and YdiA had previously been carried out by Dr Tom McAllister and Dr Jeff Hollins in which they used ITC to determine the binding parameters between YdiA and two related peptides that were designed to mimic the binding site of PPS.[126,140] τ-pTz peptide **108** (Ac-RGGRTSpTzAA-NH_2), which contains τ-phosphohistidine analogue pTz, was synthesised to study the binding interaction of catalytic histidine-phosphorylated PPS and YdiA (Figure 2.10a). pT peptide **109** (Ac-RGGRpTSHAA-NH_2), which contains a phosphothreonine residue, was synthesised to investigate binding of regulatory threonine-phosphorylated PPS to YdiA (Figure 2.10b). ITC experiments were conducted in which PPS peptides **108** and **109** were titrated into purified YdiA (expressed as a fusion protein with MBP) to yield binding constants of 11.4 ± 6.9 μM for pTz peptide **108** (Figure 2.10a) and 39.8 ± 5.5 μM for pT peptide **109** (Figure 2.10b).

[i] Synthesis of peptides for ITC and subsequent ITC experiments were carried out by Dr Tom McAllister and Dr Jeff Hollins respectively. Dr Jeff Hollins generated MBP–YdiA.

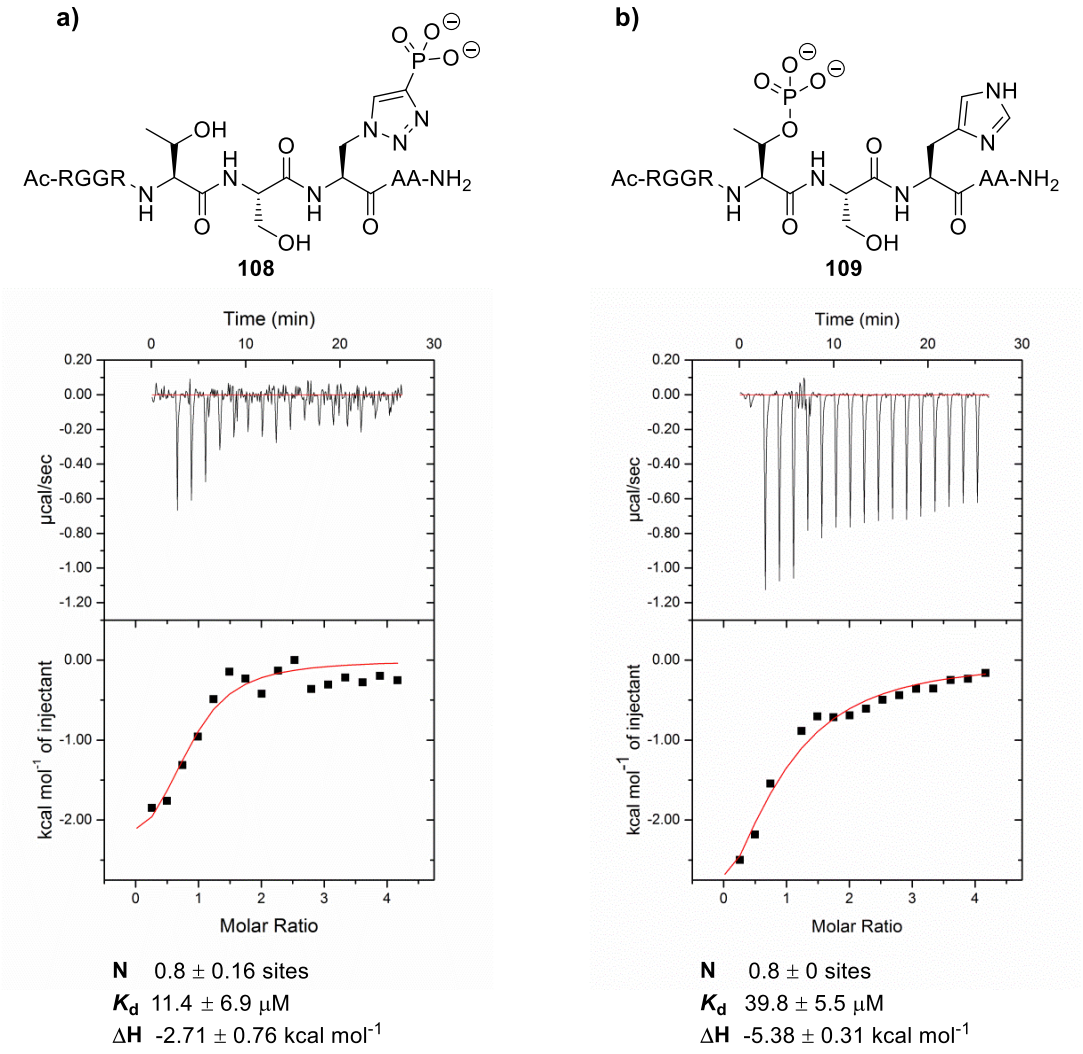


Figure 2.10: ITC experiments were carried out titrating PPS-derived a) pTz peptide **108** or b) pT peptide **109** into MBP–YdiA to give dissociation constants of 11.4 ± 6.9 µM and 39.8 ± 5.5 µM respectively.

2.3.1.2 Fluorescence Polarisation Experiments[ii]

Following the successful use of competitive fluorescence polarisation to study the interaction between Grb2-SH2 and phosphoamino acid-containing sequences (Section 2.2.3.5.2), it was decided to use an analogous assay to further elucidate the binding interaction between YdiA and a variety of PPS-derived peptides. For these assays, fluorescent τ-pTz peptide **110** comprising an N-terminal FITC and an analogous peptide sequence to τ-pTz peptide **108** (that had been used in preceding ITC experiments) would be used as the competitive probe (Figure 2.11a). After using fluorescence polarisation to obtain

[ii] Fluorescence anisotropy experiments and purification of MBP–YdiA were performed by Ieva Drulyte (Wellcome Trust rotation student) under the supervision of KAH. Fluorescence polarisation analysis was performed by KAH.

the binding affinity of **110** to YdiA, pTz probe **110** would be used in a series of competition experiments against a range of PPS derived peptides of differing lengths containing either phosphothreonine **1** and/or τ-phosphotriazolylalanine **11**. Variation of the length of peptide sequence surrounding the PPS-binding site would provide insight into the crucial residues and length of sequence needed for binding to YdiA. It would also be interesting to see whether a PPS-derived peptide containing both phosphothreonine **1** and τ-phosphohistidine analogue **11** had affinity for the protein

pTz peptide probe **110** was synthesised using Fmoc SPPS with reaction of the peptide N-terminal with fluorescein isothiocyanate (Figure 2.11a). To investigate binding of pTz probe **110** to YdiA, a 3–fold dilution series of MBP–YdiA in *SEC Buffer* containing 45 nM of **110** was prepared in triplicate. Fluorescence intensity was measured in both parallel and perpendicular channels and the resultant data points were converted to anisotropy *r* through application of equations **(3)** and **(4)** (Section 2.2.3.5.2.1). Surprisingly, this revealed a pattern inconsistent with a binding event (Figure 2.11b).

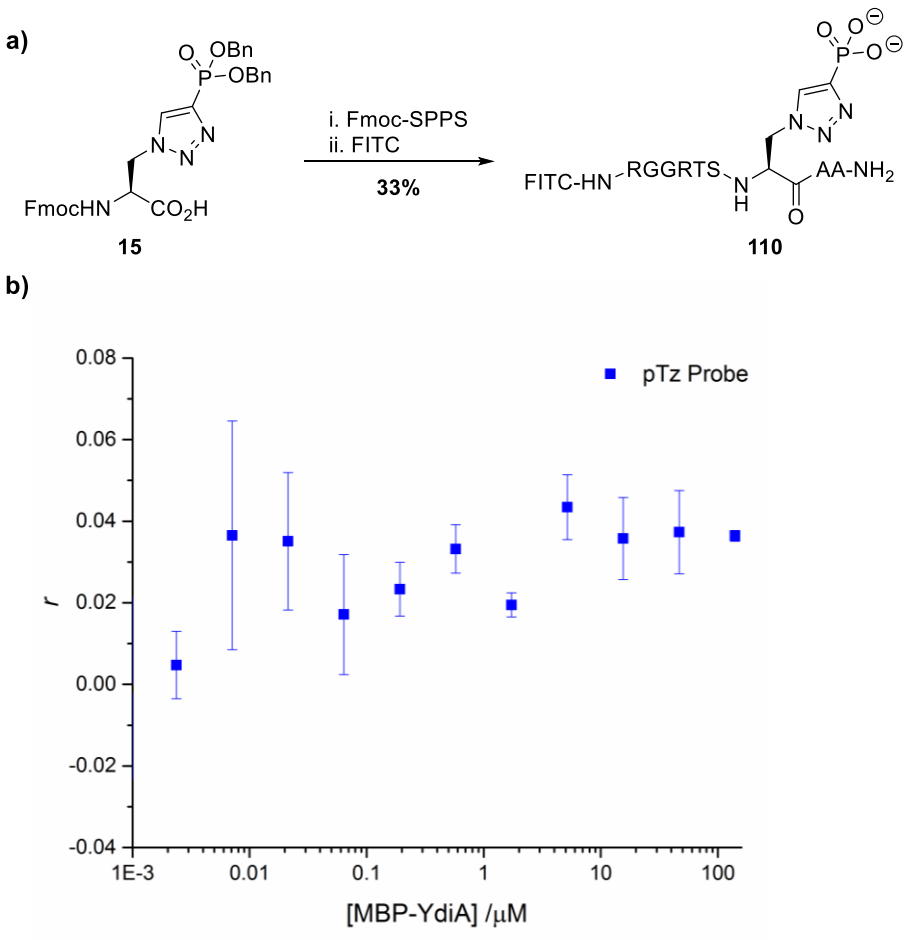


Figure 2.11: a) Synthesis of PPS-derived pTz- probe **110** via Fmoc SPPS; b) Fluorescence polarisation analysis of the binding **110** to serially diluted MBP-YdiA (140 μM – 2.4 nM) shows no binding interaction.

2.3.2 Discussion and Conclusions

Clearly, there is contradiction in the data obtained from ITC and fluorescence anisotropy experiments. It is possible that the presence of a relatively bulky fluorophore in pTz probe **110** could decrease binding affinity towards YdiA. However, this is unlikely since no low affinity binding was observed in the polarisation assays, even at protein concentrations as high as 140 μM. It is also doubtful that MBP-YdiA had unfolded before use in the fluorescence anisotropy assay as circular dichroism (CD) experiments conducted by Ieva Drulyte demonstrated that MBP-YdiA retains its secondary structure over time, even after being stored for 30 days (data not shown). It is possible that ITC data indicative of a binding event between pTz peptide **108** and MBP-YdiA (Figure 2.10) was due to peptide disaggregation and not a protein-peptide interaction. It is probable that the peptide sequences generated to mimic the binding site of PPS are not extended to the crucial length of residues needed for interaction with YdiA.

From the disappointing results of these experiments, it is evident that use of a peptide chain to mimic the primary structure of one protein counterpart to study a binding interaction is often not sufficient. This is because protein-protein interactions can be the result of association of residues over a considerable surface area of a protein. It may be possible to study binding of PPS to YdiA if the whole PPS protein could be generated with a τ-phosphohistidine mimic (or phosphothreonine) incorporated at the desired position. In fact, the ability to site-specifically incorporate a τ-pHis analogue into any protein would be of considerable use in studying a wide array of τ-phosphohistidine-mediated systems. The following section describes current progress towards this goal.

2.4 A Third Generation τ-phosphohistidine Analogue

As described in Section 1.3.2, unnatural amino acids can be incorporated into specific sites in proteins by exploitation of the protein's natural machinery using amber suppression.[123] Previous attempts to incorporate UAAs bearing phosphate groups have been unsuccessful as the highly charged nature of the phosphoryl moiety has prevented transportation across the cell membrane.[124,126] τ-phosphotriazolylalanine (pTz-3) **111**, that has allyl-protected phosphoryl groups, should have reduced polarity sufficient to penetrate cellular membranes (Figure 2.12).

It should be possible to synthesise **111** using the established synthetic procedure previously used to generate pTz **15** and phTz **103** (Section 2.2.3.2). Once synthesised, pTz-3 **111** will be sent to Jason Chin's laboratory at the University of Cambridge and screened as a substrate for a vast library of pyrrolysyl- and tyrosyl-tRNA synthetases. If a compatible synthetase is not found, a novel amber suppression system will be evolved to genetically encode pTz-3 **111** into proteins. On discovery or evolution of a system such as this, it will be possible to site-specifically incorporate **111** into a range of proteins including PPS. It should be possible to remove the allyl protection of **111** after incorporation into proteins using palladium(0).

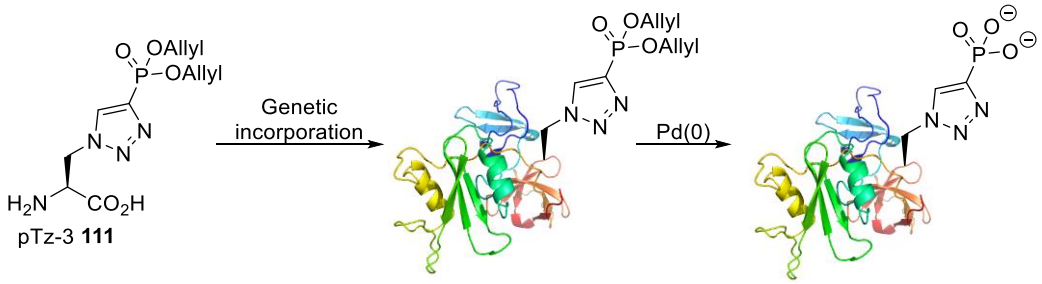


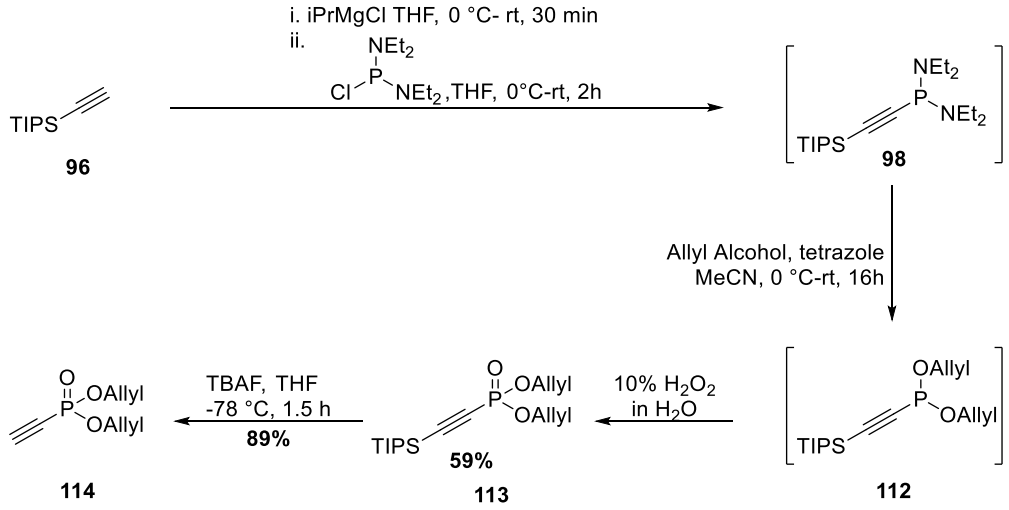
Figure 2.12: Genetic incorporation of a third generation τ-phosphohistidine mimic **111**. After incorporation, the phosphoryl groups can be removed using Pd(0).

2.4.1 Synthesis of Second Generation τ-Phosphotriazole

2.4.1.1 Phosphoalkyne Synthesis

Protected phosphoalkyne **113** was generated in 4 steps from TIPS-acetylene **96** in a good overall yield of 59% (Scheme 2.3).[25] Phosphine **112** was found to be volatile and therefore prior to oxidation with hydrogen peroxide to yield phosphate **113**, concentration of **112** *in vacuo* was performed at 50 mbar and 40 °C; this served to minimise loss of intermediate **112**. Deprotection of **113** using freshly opened TBAF in THF at -78 °C gave phosphoalkyne **114** in 89% yield. Repetition of this step with TBAF that had been stored in the fridge for ca. 1 month reduced the yield of the final step to 50%. Generation of novel allyl protected alkynylphosphonate **113** demonstrates the robustness of the synthetic route devised by

McAllister *et al*.,[25] who synthesised analogous tBu- and Bn- versions of **113**. The ability to incorporate a variety of phosphoryl protecting groups via reaction of the corresponding alcohol with intermediate **98** makes this route adaptable to the synthesis of a range of different phosphonates (and phosphonites).



Scheme 2.3: Synthesis of target phosphoalkyne **114** in an overall yield of 52% via alkynylphosphine intermediate **98**. Use of intermediate **98** allows for any organic alcohol to be used to form the corresponding ethynylphosphonite **112**. After oxidation of **112** to form phosphonate **113**, simple removal of the TIPS protecting group using TBAF gave the free phosphoalkyne **114**.

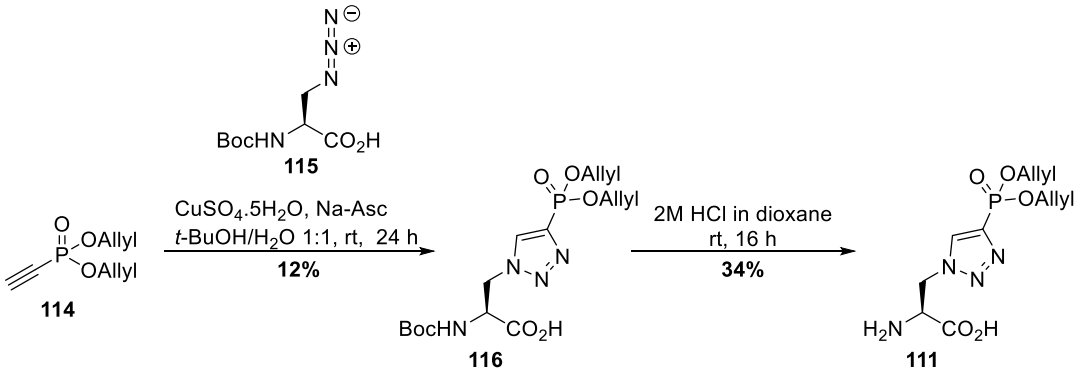
2.4.1.2 Triazole Synthesis

2.4.1.2.1 Synthesis through a Boc-protected Intermediate

Boc-protected triazole **116** was synthesised through a Cu(I) catalysed cycloaddition reaction between phosphoalkyne **114** and Boc-protected azidoalanine **115** (Scheme 2.4).[25] Azidoalanine **115** was purchased as the dicyclohexylammonium (DCA) salt. Cu(I) was generated *in situ* from the reduction of copper (II) sulfate pentahydrate by sodium ascorbate in a 1:1 mixture of H_2O and tBuOH. Work-up of the crude mixture resulted in the product being distributed between the aqueous and organic layers; acidification to pH 1 and subsequent extraction with ether did not lead to total extraction of **116** into the ethereal layer. A monoallylated version of triazole **116** was also observed in the aqueous layers through LCMS analysis. The reaction was carried out a second time and crude **116** was applied directly to a column after concentration without work up. Purification by column chromatography gave diallyl-phosphotriazole **116** together with a number of contaminants that were revealed to be the DCA salt from azidoalanine **115**, and both monoallyl- and fully deprotected- versions of triazole **116**. Purification of crude **116** by anion-exchange chromatography, eluting the Q-sepharose resin with increasing concentrations of ammonium bicarbonate (10-500 mM), did serve to remove the DCA salt but lead to increased formation of the monoallylated version of **116**. Purification by mass-directed

HPLC gave phosphotriazole **116** in a poor yield of 12%; NMR analysis showed that **116** still had low levels of contamination by impurities.

Nevertheless, Boc-protected phosphotriazole **116** was used in a trial deprotection reaction to generate third generation phosphotriazole **111** by stirring in 2M HCl in dioxane overnight. Analysis of the crude reaction mixture revealed the presence of the desired product **111** together with the corresponding monoallylated derivative. Purification by mass-directed HPLC yielded pTz-3 **111** in a moderate yield of 34%.



Scheme 2.4: Synthesis of third generation phosphotriazole **111** via Boc-protected triazole **116**, adapted from a route described by McAllister *et al*.[25] Protected triazole **116** was synthesised via the (2+3) cycloaddition of alkyne **114** and azidoalanine **115** to yield **116** in 12% yield. The poor yield can be attributed to *in situ* allyl-group deprotection of **116** and purification by MD-HPLC. Acid-catalysed Boc-deprotection of **116** yielded desired product **111** together with the corresponding monoallylated derivative.

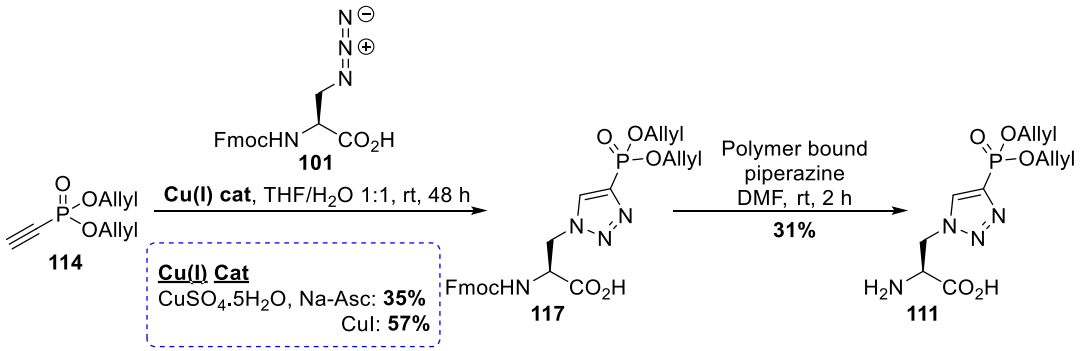
Use of ammonium bicarbonate in anion-exchange chromatography in attempted purification of protected triazole **116** and use of 2M HCl to form triazole **111** both lead to removal of one or both of the allyl protecting groups. Commercially available azidoalanine **115** is purchased as the dicyclohexylammonium salt and this has been proven difficult to remove via manual column chromatographic methods. Although purification by mass-directed HPLC served to remove the DCA-salt, it may have also served to reduce the yield of protected triazole **116**. To circumvent these problems, azidoalanine with alternative amino-group protection was required.

2.4.1.2.2 Synthesis through an Fmoc-protected Intermediate

Fmoc-protected triazole **117** was initially generated from the cycloaddition of commercially available Fmoc-azidoalanine **101** and alkyne **114** following the same procedure as described for the generation of Boc-triazole **116** (Scheme 2.5).[25] Azidoalanine **101** was found to be poorly soluble in H_2O/tBuOH and hence a 1:1 mixture of H_2O/THF was used instead. After the reaction had reached completion, the aqueous reaction mixture was extracted with CH_2Cl_2, concentrated, and purified by column chromatography to give triazole **117** in a moderate yield of 35%. Without acidification of the aqueous layer, extraction of triazole **117**

into the organics proved to be a lengthy procedure. Furthermore, LCMS analysis of the aqueous layer revealed presence of the monoallylated derivative of triazole **117**, similar to the formation of the monoallylated triazole in the synthesis of Boc-triazole **116** when using an acidic work up. This suggests that acid-catalysed elimination of the allylic alcohol is not the reason for formation of mono-allylated side-products. It is possible that the mildly basic sodium ascorbate used to aid generation of the Cu(I) species through reduction of Cu_2SO_4 may be removing allyl protection of triazoles **116** and **117**. Accordingly, protected triazole **117** was synthesised using CuI as a source of Cu(I), and, after purification using automated column chromatography, **117** was generated in an improved yield of 57%. With the optimised procedure for formation of protected triazole **117** in place, attention was turned to the removal of the fluorenylmethyl group. Deprotection of triazole **117** was attempted through addition of a fluoride anion by stirring of **117** in 0.1M TBAF at room temperature.[141] Unfortunately, it was not possible to separate desired product **117** from the alkyl fluorenylcarbonate side-product via automated or manual chromatographic procedures.

Lang *et al.*[87] demonstrated the effective use of polymer-bound piperazine in the removal of an Fmoc group from a bicyclononyne derivative. It was anticipated that if reacted with triazole **117**, the supported piperazine would be a weak enough base to leave the allyl-protecting groups of triazole **117** intact, whilst the piperazine-linked solid support would capture the dibenzofulvene leaving group; eliminating the need for work-up and purification. Thus, Fmoc-protected triazole **117** was added to a suspension of polymer-bound piperazine in dry DMF and gravity filtered after 2.5 hours. Analysis of the filtrate showed a detectable amount of the monoallylated version of **117** alongside the desired triazole **117**. Crude triazole **111** was therefore purified by mass-directed HPLC to give the final product in 30% yield. The analogous reaction was performed in wet DMF to generate the desired pTz-3 **111** in a comparable yield of 31%. It should be noted that in the second attempt, the reaction was stopped after 2 hours and no trace of the corresponding monoallylated triazole was observed. Unfortunately, reduction of reaction time coincided with trace amounts of protected triazole **117**; hence, purification by mass-directed HPLC was still necessary. It should be noted that in both instances, the crude yield of **111** before HPLC purification was significantly higher in comparison to the yield of the purified product.

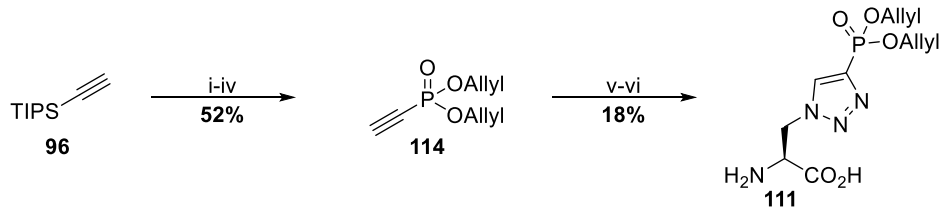


Scheme 2.5: Synthesis of pTz-3 **111** through Fmoc-protected triazole **117**. Triazole **117** was generated in 57% yield through optimised cycloaddition reaction conditions utilising Cu(I) before Fmoc-deprotection with polymer-bound piperazine to give the desired product **111**.

Deprotection of Fmoc-pTz-3 **117** to give **111** is dissapointing in terms of yield, seemingly due to the need for automated HPLC methodology for purification of crude **111**. In general, removal of amino group protecting groups should involve no purification. However, in this case, concomitant formation of the monoallylated derivative of the desired triazole **111** has meant that purification of **111** was neccasary.

2.5 Conclusions

Third generation triazolylalanine, τ-pTz-3 **111** has been generated from TIPS acetylene **96** in 7 synthetic steps with an overall yield of 9% (Scheme 2.6). Synthesis of diallyl alkynylphosphate **114** has demonstrated that the synthetic strategy devised by Dr. Tom McAllister,[25] can indeed be adapted to incorporation of alternative protection groups for the phosphonate. pTz-3 **111** has the potential to be site-specifically incorporated into proteins using amber suppression. Scale-up of the synthetic route to pTz-3 **111** devised in this chapter is currently on going.



i. iPrMgCl, THF, 0 °C- rt; ii. *bis*(diethylamino)chlorophosphine, THF, 0 °C-rt; iii. Allyl alcohol, tetrazole, MeCN, 0 °C-rt; iv. 10% H_2O_2 in H_2O; iv.TBAF, THF, -78 °C; v. Fmoc-azidoalanine-OH **101**, CuI, THF/H_2O (1:1), rt; vi. Polymer bound piperazine, DMF, rt.

Scheme 2.6: Optimised synthetic route to novel τ-pTz-3 **111**.

Chapter 3 Synthesis of Functionalised 1,2,4-triazines

In addition to their use as mimics of post-translational modifications, unnatural amino acids can also be used in bioorthogonal chemical reporting strategies. As discussed in Section 1.2, there are a number of chemical reactions that have been used to label biomolecules. However, many of these reactions are limited in terms of biocompability or synthetic accessibility of reagents. The following chapters discuss progress made towards the generation of a novel bioorthogonal probe involving the cycloaddition of triazines to strained dienophiles.

3.1 Synthesis of Functionalised 1,2,4-Triazines

To test the utility of the 1,2,4-triazine-SPIEDAC as a novel bioorthogonal probe, a triazine with a functional group handle was required that could either be incorporated into an unnatural amino acid (to be genetically encoded into a protein), or derivatised onto a fluorescent reporter. Accordingly, a range of 3-amino-1,2,4-triazines were to be synthesised with the intention of generating target molecules through amide bond formation of the exocyclic amino-group handle (Figure 3.1). In order to acquire a comprehensive spectrum of 1,2,4-triazine reactivity in SPIEDAC reactions, it was deemed necessary to generate a range of substituted triazine derivatives differing in size and electronic nature. Balcar *et al.*[107] measured the second order rate constant for the cycloaddition of cyclooctene to 1,2,4-triazine, 3-phenyl-1,2,4-triazine and 3-methyl-1,2,4-triazine. Calculated rate constants decreased in the order H > Ph > Me. Therefore, 3-amino-1,2,4-triazine **118a**, 6-phenyl-3-amino-1,2,4-triazines **118b** and 6-methyl-3-amino-1,2,4-triazine **118c** would all be synthesised (Figure 3.1).

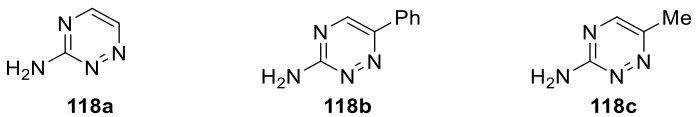
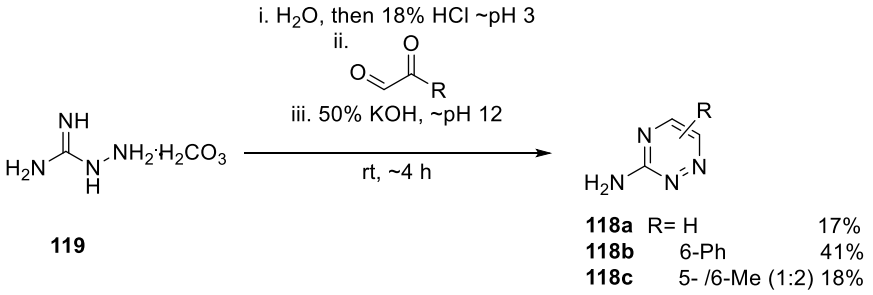


Figure 3.1: 1,2,4-triazines with exocyclic amino handles for further derivatisation: 3-amino-1,2,4-triazine **118a** and phenyl- and methyl-3-amino-1,2,4-triazines **118b** and **118c** respectively.

3.1.1 Synthesis of Substituted 3-Amino-1,2,4-triazines

As reported by Erickson,[142] substituted (and non-substituted) 3-amino-1,2,4-triazines can be accessed through the condensation and subsequent cyclocondensation of aminoguanidine to 1,2-dicarbonyl compounds. A synthetic route towards aminotriazines **118a-c** was devised in the group that involved the controlled condensation of aminoguanidine bicarbonate **119** and glyoxal derivatives under acidic conditions (Scheme 3.1). On formation of the corresponding aminoguanidine intermediate, the pH of the reaction mixture is increased to pH ~12 to allow for the subsequent cyclocondensation reaction to take place. Accordingly, aminotriazines **118a-c** were synthesised through the condensation of aminoguanidine bicarbonate **119** and glyoxal, phenylglyoxal or methylglyoxal at pH ~3; the pH of the reaction mixture was increased to ~pH 12 through the addition of 50% KOH, and an ensuing cyclocondensation gave the corresponding products **118a-c** in yields of 17%, 41% and 18% respectively. Methyl-aminotriazine **118c** was recovered as a mixture of regioisomers in a 2:1 ratio, with the 6-substituted product being the major isomer as deduced by NMR. Efforts to improve yields of **118a-c** based on more gentle cyclocondensation methods of stirring in either water,[142] or phosphate buffer,[143] resulted in no significant improvement. Synthetic routes to triazines **118a-c** require optimisation; this could possibly have been achieved via the use of freshly distilled glyoxal derivatives.[144] However, at this time, 3-amino-1,2,4-triazine **118a** became commercially available and was therefore used as the basis for the majority of subsequent reactions.



Scheme 3.1: Synthesis of 5/6-substituted 3-amino-1,2,4-triazines **118a-c** through the controlled condensation of aminoguanidine bicarbonate **119** and glyoxal derivatives. In the case of methyl-aminotriazine **118c**, a mixture of regioisomers was recovered.

3.2 Incorporation of 3-Aminotriazine into Unnatural Amino acids

With aminotriazine **118a** in hand, attention was turned to incorporation of **118a** into unnatural amino acids. Once generated, the triazine-containing UAAs would be assessed against a range of amber suppression systems. The use of permissive aminoacyl tRNA-synthetases for the genetic incorporation of UAAs that are structurally related to the UAAs they were evolved to incorporate is described in Section 1.3.2.

3.2.1 tRNA Synthetase-tRNA$_{CUA}$ Pairs used for Genetic Code Expansion

In general, four tRNA synthetase-tRNA pairs are used for genetic code expansion in various organisms. The *Methanococcus jannaschi* Tyrosyl-tRNA synthetase (*Mj*TyrRS)-tRNA$_{CUA}$ pair is orthogonal to natural synthetases and tRNAs in prokaryotes but not in eukaryotes and incorporates UAAs based on tyrosine scaffold **120** (Figure 3.2).[145] Both *E. coli* Tyrosyl-tRNA synthetase (*Ec*TyrRS)-[146] and *E. coli* Leucyl-tRNA synthetase (*Ec*LeuRS)[147]- tRNA$_{CUA}$ pairs are orthogonal in yeast and mammalian cells but not in *E. coli*. These synthetase-tRNA pairs genetically encode analogues of tyrosine **120** and leucine **122** respectively. The pyrrolysyl-tRNAsynthetase (PylRS)-tRNA$_{CUA}$ pair from *Methanosarcina* species has been used to genetically encode a variety of UAAs based on pyrrolysine scaffold **121** and is orthogonal to natural synthetases and tRNAs in *E. coli*,[148] yeast,[149] mammalian cells[150] and *C.elegans*.[151]

The PylRS-tRNA$_{CUA}$ pair has two advantages over the other three synthetase pairs. The first and most obvious advantage is that this system can be used to incorporate UAAs into more than one organism. Secondly, pyrrolysine **121** (the amino acid the synthetase naturally encodes) is not one of the 20 canonical amino acids; as such natural synthetase activity does not need be destroyed when creating specificity for a new amino acid.[3] This has also made it possible to use the unmodified PylRS to incorporate a range of unnatural amino acids.[150] There have been many reports of permissive PylRS-tRNA$_{CUA}$ pairs that have genetically encoded additional UAAs to the UAA they were originally evolved to incorporate.[44] Accordingly a triazine-containing UAA based on pyrrolysine scaffold **121** would be synthesised and screened as a substrate for existing pyrrolysyl-tRNA synthetases.

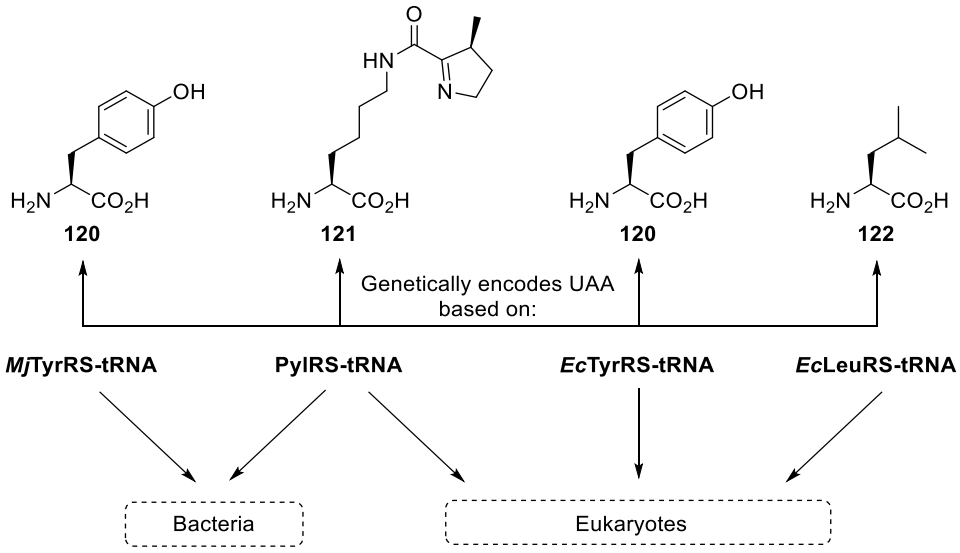
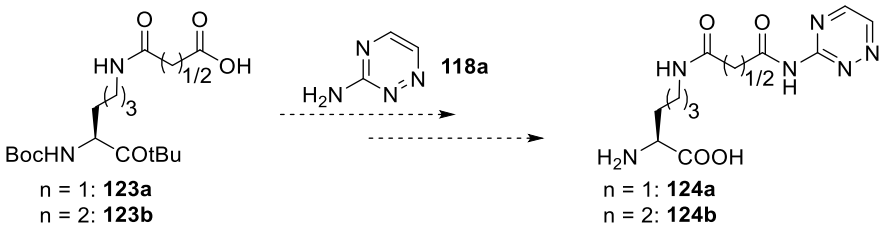


Figure 3.2: Orthogonality of synthetases and their cognate tRNA$_{CUA}$s that are used for genetic code expansion in different organisms.

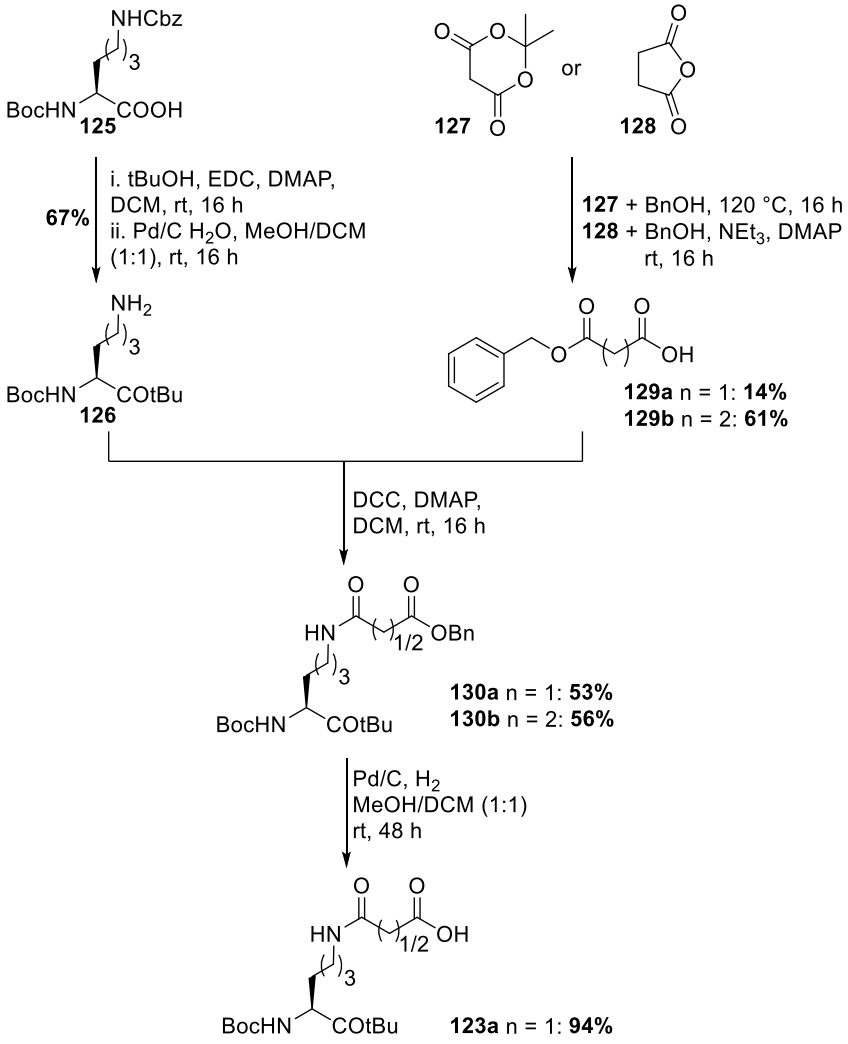
3.2.2 Attempted Synthesis of Triazine-containing Pyrrolysine Analogues

Built on methodology by Du *et al*,[152] triazine-containing pyrrolysine analogues **124a** and **124b** were to be synthesised via the derivatisation of malonyl- (n = 1) or succinyl- (n = 2) lysines **123a** and **123b** to aminotriazine **118a** (Scheme 3.2). **124a** and **124b** differ in the length of linkage between lysine and triazine by one carbon atom; it is possible that this difference in flexibility could result in one analogue being a substrate for a pyrrolysyl-tRNA synthetase.



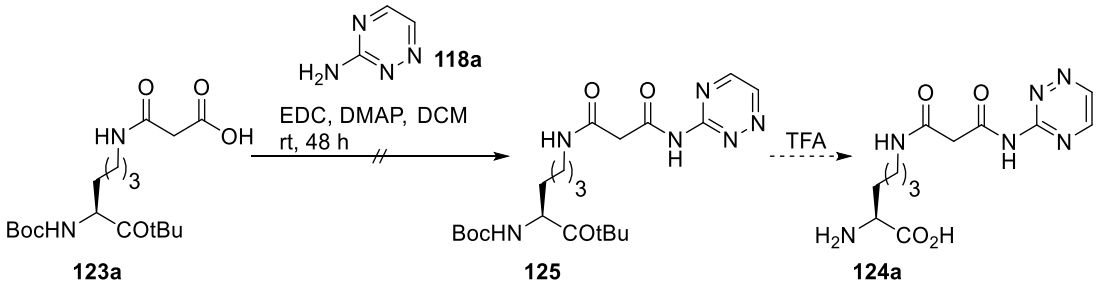
Scheme 3.2: Overview of proposed synthetic route towards triazinyl-lysine derivatives **124a** and **124b**.

In order to generate **124a** and **124b**, protected malonyl- and succinyl-lysines **123a** and **123b** needed to be synthesised. This was achieved as described in Scheme 3.3: Initially Boc-Lys(Z)-OH **125** was converted to the fully protected Boc-Lys(Z)-OtBu using tBuOH in a yield of 68%. The Cbz group was subsequently removed through palladium catalysed hydrogenation to give Boc-Lys-OtBu **126** in 99% yield.[153] Hydrogenation of the Cbz group of Boc-Lys(Z)-OtBu went to completion in 6 hours; leaving the hydrogenation for longer periods resulted in catalyst poisoning. Generation of benzyl malonate/succinate coupling partners **129a** and **129b** proceeded via the ring-opening of Meldrum's acid **127** or succinic anhydride **128** by benzyl alcohol to afford benzyl malonate **129a**[154] and benzyl succinate **129b**,[155] in yields of 14% and 61% respectively. The 14% yield obtained on formation of benzyl malonate **129a** is attributed to the need for a base to promote the reaction. Formation of benzyl-protected malonyl- and succinyl-lysines **130a** and **130b** was achieved through coupling of benzyl malonate **129a** or benzyl succinate **129b** to Boc-Lys-OtBu **126** using a DCC/DMAP activation strategy. This afforded benzyl malonyl- and benzyl succinyl-lysines **130a** & **130b** in yields of 53% and 56% respectively. The benzyl group of malonyl-lysine **130a** was removed by hydrogenation to yield crude malonyl-lysine **123a** in 94% yield. Before hydrogenation of benzyl succinyl-lysine **130b** to form **123b** was performed, malonyl-lysine **123a** was used in a trial coupling reaction with aminotriazine **118a** in an attempt to form triazinyl-lysine **124a** (Scheme 3.4).



Scheme 3.3: Synthetic route towards malonyl- succinyl-lysines **123a** and **123b** (**123b** was not synthesised): Benzyl malonate **129a** and benzyl succinate **129b** were generated through reaction of benzyl alcohol with meldrums acid **127** or succinic anhydride **128** respectively. Reaction of protected lysine **127** with **129a** or **129b** resulted in lysines **130a** and **130b**. Hydrogenation of **130a** to remove the benzyl-protection yielded succinyl-lysine **123a**.

Attempts to couple malonyl-lysine **123a** and aminotriazine **118a** to generate triazinyl-lysine **125** using an EDC/DMAP activation strategy at room temperature failed (Scheme 3.4). The use of higher temperatures and stronger bases may facilitate this reaction.



Scheme 3.4: Failed generation of Boc-triazinyl-lysine derivative **124a** (to generate triazinyl-lysine **124a**) through coupling of malonyl-lysine **123a** to aminotriazine **118a**.

At this time, independent reports of the genetic encoding of norbornyl-carboxylysine **126**[85], bicyclononyl-carboxylysine **128**[87,156], and *trans*-cyclooctenyl-carboxylysines **129**[85] and **130**[87] into proteins using *wt* or evolved pyrrolysl tRNA synthetases were published.[85,87,156] Co-translational incorporation of cyclopropenyl-carboxylysine **127** into proteins has since been achieved.[86] Due to the possibility of not finding a permissive tRNA synthesis that could be used for genetic incorporation of a novel triazine-containing amino acid, it seemed judicious to utilise amber suppression systems reported to encode strained dienophile-containing amino acids instead. Hence, synthesis of a functionalised triazine that could be appended to a probe molecule was now needed. As a result, no further attempts were made to generate triazinyl-lysines **124a** or **124b**.

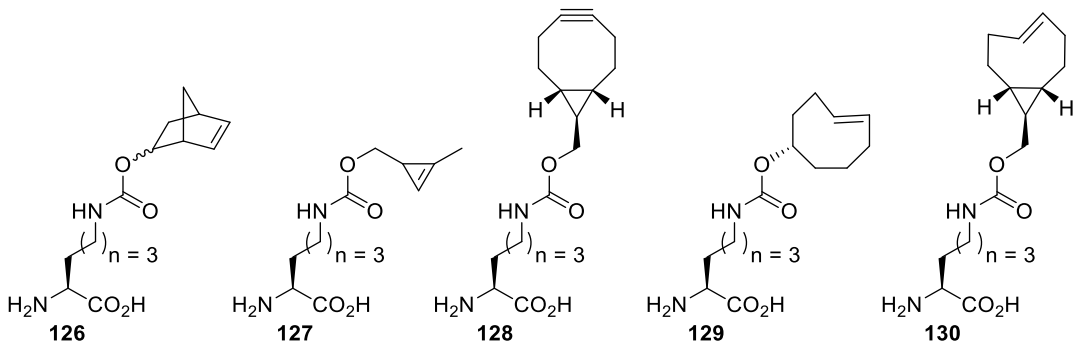


Figure 3.3: Strained dienophile-containing lysine derivatives that have been genetically encoded into proteins in response to an amber codon using *wt* or evolved pyrrolysyl-tRNAsynthetases.

3.3 Generation of 1,2,4-triazine Derivatives for Probe Functionalisation

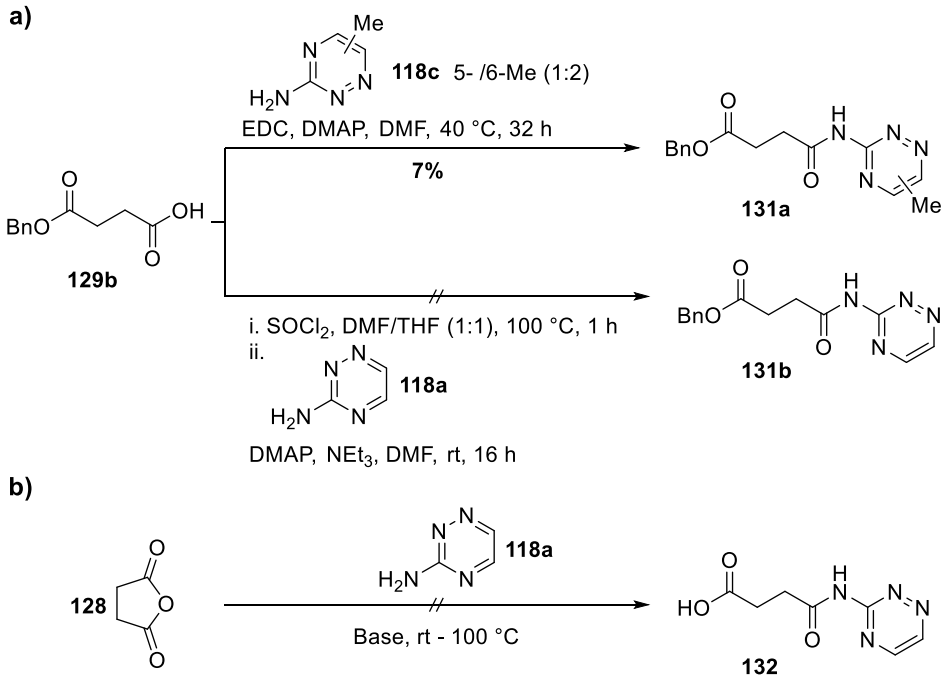
3.3.1 Through Amide-bond Formation

Methods to generate functionalised triazine compounds suitable for late-stage derivatisation onto probe molecules initially focused on the generation of benzyl-protected succinylamidotriazine derivatives of the form **131a** and **131b** through the coupling of aminotriazines **118c** or **118a** to benzyl succinate **129b** (Scheme 3.5a). Removal of the benzyl-protecting group of **131a** or **131b** by hydrogenation would generate acid functionalised triazines suitable for reaction with small molecule- and peptide-probes containing nucleophilic functionalities.

For the synthesis of protected succinylamidotriazine **131a**, an EDC/DMAP activation strategy was used, with stirring at room temperature for 16 h to give **131a** in a poor yield of 1%. LCMS analysis of the reaction mixture revealed presence of the *N*-acylurea cyclic displacement product, formed from rearrangement of the corresponding *O*-acylurea intermediate (generated through *in situ* activation of benzyl succinate **129b** by EDC). To

promote nucleophilic attack of methyl aminotriazine **118c** onto the activated *O*-acylisourea intermediate, the reaction was heated to 40 °C. This served to increase the yield of **131a** to 7% (Scheme 3.5a, top). A failed attempt to generate benzyl succinylamidotriazine **131b** through conversion of **129b** to the corresponding acid chloride, and subsequent coupling to aminotriazine **118a** was also made (Scheme 3.5a, bottom).

Efforts to access succinylamidotriazine **132** directly through the ring-opening of succinic anhydride **128** with aminotriazine **118a** were also made (Scheme 3.5b). On a number of occasions, formation of **132** was observed by LCMS; however the reaction was too inefficient to yield suitable quantities of material for purification. It was also noted that at high temperatures aminotriazine **118a** was consumed. It is possible that the high temperatures used in this reaction may have led to the degradation of triazine **118a** and contribute to the insufficient quantities of crude succinylamidotriazine **132**.



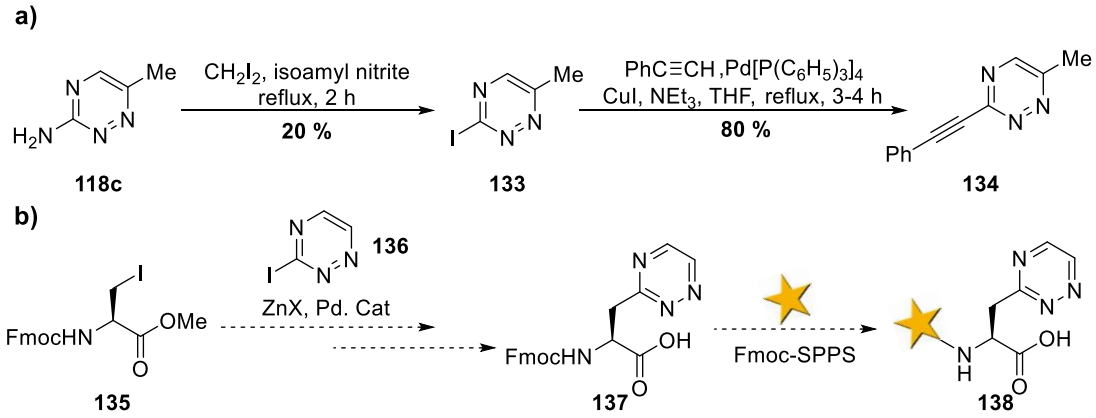
Scheme 3.5: a) Syntheses of benzyl succinylamidotriazines **131a** (top) and **131b** (bottom) through an *O*-acylisourea or acid chloride intermediate respectively. **131a** was synthesised in a poor yield of 7%, whilst synthesis of **131b** was unsuccessful; b) attempted formation of desired acid **132** through reaction of aminotriazine **118a** and succinic anhydride **128**.

Syntheses utilising the amino-group of triazines **118a** and **118c** as reactive handles for further functionalisation have generally been unsuccessful (and when successful, as in the synthesis of benzyl succinylamidotriazine **131a**, have generated products with inadequate yields to be of use). This is presumably due to the highly electron-withdrawing nature of the triazine core which will reduce the nucleophilicity of the exocyclic amino group of **118a** and

118c. It was therefore necessary to alter synthetic strategy and convert the amino-group of triazine **118a** into an alternative functional handle.

3.3.2 Through Palladium-catalysed Cross-coupling

Metal-mediated carbon-carbon bond forming reactions involving triazine have been reported using an appropriately functionalised substituted triazine and unsaturated aliphatic compounds.[113,157] Carroll *et al.*[157] exploited methyl iodotriazine **133** in a metal-mediated Sonogashira-type reaction with phenylacetylene to afford phenylethynyltriazine **134** in 80% yield (Scheme 3.6a). Carroll *et al.*[157] generated methyl iodotriazine **133** through diazotization and ensuing halogenation of methyl aminotriazine **118c**, using isoamyl nitrite and diiodomethane, in a modest yield of 20%. Jackson and co-workers[158,159] have reported the synthesis of enantiomerically pure pyridylalanine amino acids through the palladium catalysed cross-coupling of serine-derived organozinc reagents with halopyridyl derivatives. It was postulated that this Negishi-type cross-coupling could be extended to include other *N*-heterocycles such as iodotriazine **136** (Scheme 3.6). Therefore a synthetic strategy was devised based on the metal-mediated cross-coupling of **136** and protected iodoalanine **135** to generate (after removal of the methyl ester) Fmoc-triazinylalanine (TrzA) **137** (Scheme 3.6b). Triazinylalanine **137** would be suitable for use in the Fmoc-strategy for SPPS and hence be particularly appropriate for rapid functionalisation of peptide probe molecules. Alternatively, it would be possible to use the acid functionality of **137** to derivatise onto small molecule probes.



Scheme 3.6: a) Reported transformation of methyl aminotriazine **118c** into methyl iodotriazine **133** and subsequent carbonic addition to form alkynyl-triazine **134** using metal-mediated cross-coupling;[157] b) proposed synthetic route to Fmoc-triazinylalanine **137** via a Negishi-type cross-coupling of iodoalanine **135** and iodotriazine **136**, and subsequent methyl ester removal. **137** could be used in an Fmoc SPPS strategy to generate a fluorescent triazine-containing fluorescent probe such as **138**.

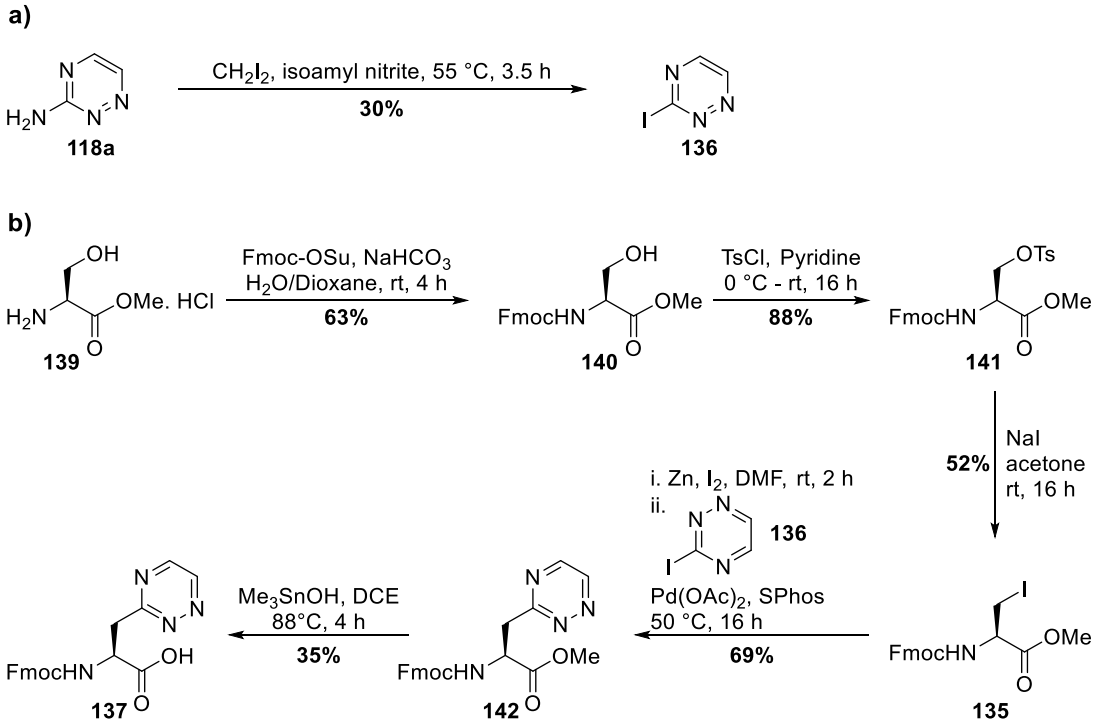
Iodotriazine **136** was formed in 30% yield through the diazotisation of aminotriazine **118a** using isopentyl nitrite in diiodomethane at 55 °C (Scheme 3.7a). Although low, this yield is a 10% increase from the reported yield for formation of methyl iodotriazine **133** (Scheme 3.6).[157] It is possible that this difference is due to degradation of triazines **118c** and **133** at the higher temperatures employed by Carrol *et al*.,[157] as observed previously when using aminotriazine **118a** (Scheme 3.5b, Section 3.3.1).

Fmoc-iodoalanine methyl ester **135** was synthesised in three synthetic steps starting from serine methyl ester **139** (Scheme 3.7b).[158] L-Serine methyl ester **139** was sequentially reacted with Fmoc-OSu and *p*-toluenesulfonyl chloride to give Fmoc-Ser(OTs)-OMe **140** in an overall yield of 55%. S_N2-substitution of the tosyl group of **140** with sodium iodide in a Finkelstein-type reaction generated iodoalanine methyl ester **135** in 52% yield. In this instance, the moderate yield was attributed to incomplete conversion of tosyloxyalanine methyl ester **140**.[160]

For zinc insertion of iodoalanine methyl ester **135** to form the corresponding organozinc reagant, the procedure reported by Jackson and co-workers was initially followed.[158] Commercial zinc dust and I_2 (0.3 eq) were weighed into an oven dried flask which was evacuated and purged with nitrogen three times at 0 °C. Iodoalanine methyl ester **135** was dissolved in anhydrous DMF (freshly opened bottle), transferred to the reaction mixture, and stirred at 0 °C. After 2 hours, no formation of the activated organozinc reagent was observed by TLC. It was reasoned that a higher temperature may be needed for formation of the alkyl zincate of **135**, and activation of zinc dust by catalytic amounts of iodine may be more effective in solution. Accordingly, formation of the organozinc reagent was attempted a second time. After evacuation and purging of an oven-dried flask containing zinc dust, anhydrous DMF (freshly opened bottle) and I_2 (0.15 eq) were added in quick succession and stirred at room temperature. After 15 minutes, iodoalanine **135** was added followed by a further 0.15 equivalents of I_2.[i] Formation of the corresponding organozinc reagent was observed after 2 hours; at which time iodotriazine **136**, Pd(II) acetate and 2-dicyclohexylphosphino-2′,6′-dimethoxybiphenyl (SPhos) were added in quick succession. The reaction was stirred at 50 °C for 5 hours and after purification by column chromatography, triazinylalanine methyl ester **142** was recovered in 69% yield (Scheme 3.7b). In attempts to optimise Negishi-coupling, the palladium catalyst loading was altered and an alternative palladium catalyst (tris(dibenzylideneacetone)dipalladium(0)) was used; neither of these changes resulted in a significant change in yield.

---

[i] After discussions with Christian Hedberg, catalytic I_2 was added once for zinc activation, and a second time for activation of the organozinc reagent.



Scheme 3.7: a) Diazotisation of aminotriazine **118a** to give iodotriazine **136**; b) successful synthesis of triazinylalanine **137** via Negishi-type cross coupling of zinc activated iodoalanine methyl ester **135** to 3-iodo-1,2,4-triazine **136** to generate **142** using palladium acetate and SPhos as a co-ligand. Methyl ester removal of **142** using trimethyltin hydroxide gave the desired product **137**. Iodoalanine methyl ester **135** was synthesised in three steps from serine methyl ester **139**.

In the optimised method for the metal-mediated cross-coupling of **135** and **136** to generate triazinylalanine methyl ester **142**, the total 0.3 equivalents of iodine were added in two stages. This was to ensure consistent catalytic activation of zinc dust by iodine and subsequent formation of iodide. Iodide promotes *in situ* formation of dianionic zincate species $RZnI_3^{2-}$ (where R = alanine methyl ester). $RZnI_3^{2-}$ is formed from $RZnI_2^-$, which in turn is generated from the organozinc reagent (RZnI). This doubly charged zincate species has been shown to be the active transmetallating agent in Negishi-type cross coupling reactions involving alkyl halides.[161] Hence, iodine was exploited in this reaction in two ways: initially for activation of zinc towards nucleophilic addition, and secondly for the catalytic formation of a higher order zincate species. It should also be noted that Negishi coupling was only successful when using a freshly opened bottle of anhydrous DMF. This is due to the inherent sensitivity of organozinc reagents to oxygen and water.

Demethylation of Fmoc-TrzAla-OMe **142** to give the free acid **137** was achieved by stirring in DCE with trimethyltin hydroxide at 88 °C (Scheme 3.7b).[162] Initial yields of 13% were reported, although alterations to the purification strategy of the crude product increased yields of **137** to 27%. In this case, the overall isolated yield of triazinylalanine **137** is limited by observable degradation of the triazinylalanines **137** and **142** at the elevated temperatures

required for deprotection; the methyl ester of **142** could not be removed at moderate temperatures of 55 °C. Base-catalysed deprotection of triazinylalanine methyl ester **142** was also attempted using LiOH, unfortunately, this led to concomitant Fmoc-group removal.

3.4 Conclusions

A novel strategy to generate an Fmoc-compatible triazine, triazinylalanine **137** has been developed (Scheme 3.7). The synthetic route towards **137** is robust and uses readily available and inexpensive starting materials. The route to **137** includes an optimised Negishi-type cross-coupling reaction between iodotriazine **136** and an organozinc reagent. It should be possible to couple iodotriazine **136** to a variety of alkyl- and aryl-organozinc reagents, which will ultimately increase the diversity of available (and synthetically accessible) triazine scaffolds. It should be possible to use triazinylalanine **137** in the Fmoc-strategy for SPPS to rapidly and efficiently functionalise a range of probes such as fluorescein isothiocyanate or carboxyfluorescein. The fully deprotected version of triazinylalanine **137** is similar in structure to a range of tyrosine-based scaffolds that have been genetically incorporated into proteins in response to an amber codon using evolved tyrosyl-tRNA synthetases.[145,163] Due to the promiscuity of these synthetases it is possible that fully deprotected triazinylalanine could be genetically incorporated into proteins. Therefore, it should be possible to use triazinylalanine **137** and its derivatives as either counterpart of a bioorthogonal probe.

Chapter 4 1,2,4-Triazine Cycloaddition Reactions

4.1 Introduction

With triazinylalanine **137** in hand, attention was turned to investigating the reactivity of this *N*-heterocycle and its derivatives towards strained dienophiles. To assess the utility of the triazine cycloaddition as a bioorthogonal probe, it was ultimately envisaged that strained dienophiles displaying reactivity towards **137** would be genetically encoded into a model protein using amber suppression and subsequently reacted with a triazinylalanine-containing fluorescent probe. Norbornyl-carboxylysine (Norb-K) **126**,[85] cyclopropenyl-carboxylysine (CP-K) **127**,[86] bicyclononyl-carboxylysine (BCN-K) **128** [87,156], *trans*-cyclooctenyl-carboxylysine (TCO-K) **129**[85] and strained *trans*-cyclooctenyl-carboxylysine (sTCOK) **130**[87] have all been incorporated into proteins in response to an amber codon using *wt* or engineered pyrrolysyl-tRNA synthetases (Figure 4.1a). Figure 4.1b shows calculated bimolecular rate constants between these strained dienophiles and disubstituted tetrazine derivatives.[86,87] Their rate of reaction to tetrazine increases in the order Norb-K < CP-K < BCN-K < TCO-K < sTCO-K.

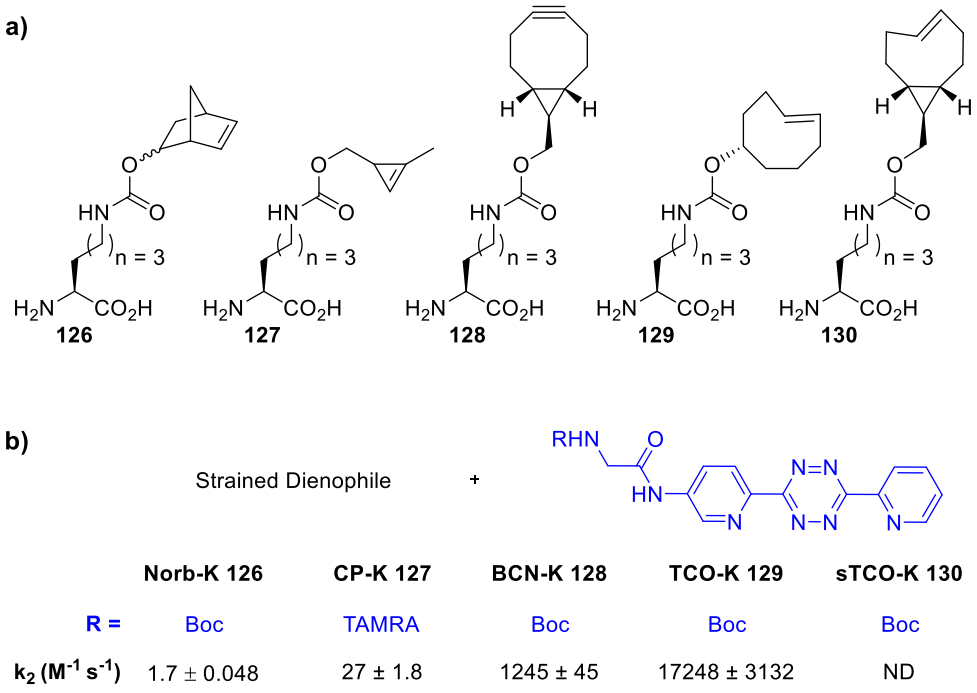


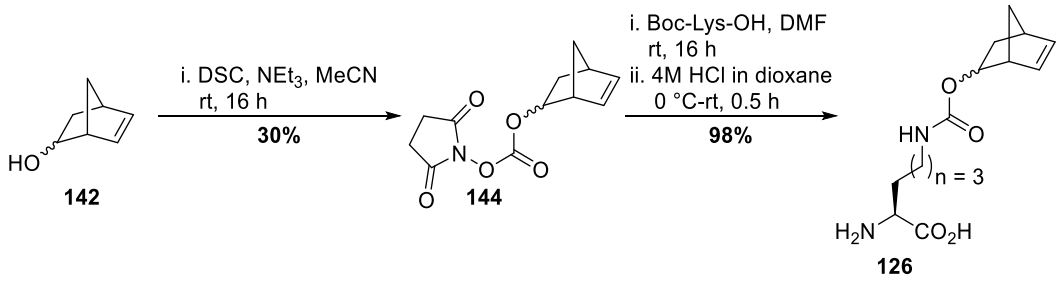
Figure 4.1: a) Strained dienophiles **126** – **130** have been genetically encoded into proteins in response to an amber codon and b) measured bimolecular rate constants of **126** - **130** in cycloaddition reactions towards tetrazine derivatives.

4.2 Norbornene as the Dienophile

4.2.1 Synthesis

In order to acquire a comprehensive spectrum of triazine reactivity, norbornene, which has a relatively slow rate of reaction (k_2 1-10 M^{-1} s^{-1})[85] towards tetrazine (Figure 4.1b) was initially selected for triazine cycloaddition studies. To investigate norbornene's reactivity towards triazine, generation of norbornyl-carboxylysine **126** was required.

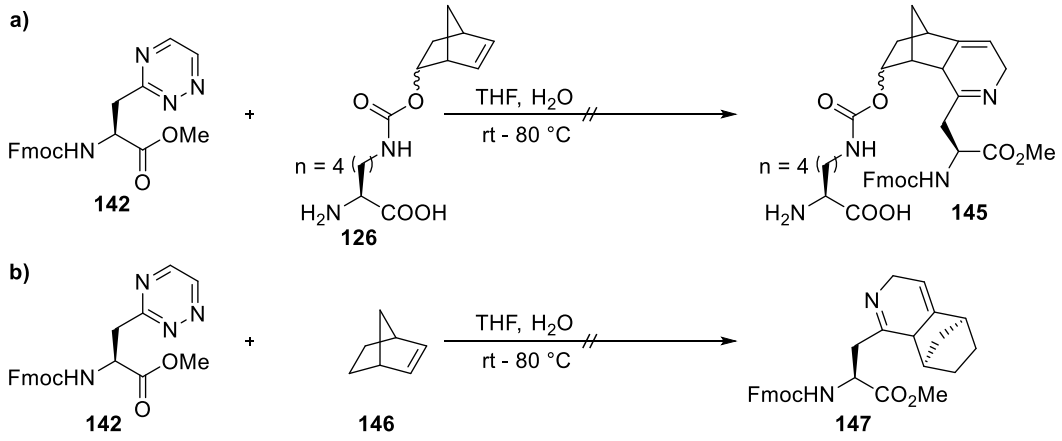
Norbornyl-lysine **126** was synthesised in three steps from 5-norbornene-2-ol in an overall yield of 29% in accordance with the procedure described by Lang *et al* (Scheme 4.1).[85] A yield of 30% was obtained for the generation of the activated ester **144**; this can be attributed to partial ester hydrolysis on exposure to silica during purification. Displacement of the succinimide ester of **144** by Boc-Lys-OH[85] and subsequent Boc-deprotection using 4M HCl in dioxane[164] yielded norbornyl-lysine **126** in a near quantitative yield of 98%.



Scheme 4.1: Synthesis of Norbornyl-lysine **126** via NHS-ester **144**.

4.2.2 Reactivity of Norbornene to Triazinylalanine

To investigate the reactivity of norbornyl-carboxylysine **126** to triazine, equimolar amounts of **126** and triazinylalanine methyl ester **142** were placed in a reaction flask at a concentration of 40 mM and stirred for three hours at room temperature in THF/H_2O (1:1); after which time, no formation of pyridyl derivative **145** was observed (Scheme 4.2a). Incubation of the reaction mixture at 37 °C and 50 °C for three hours and overnight at 80 °C (below the temperature that degradation of triazines **137** and **142** had been observed to occur) did not lead to formation of cross-linking product **145**. Triazine methyl ester **142** was used in this study to prevent possible interference of the free acid. However it is conceivable that unprotected norbornyl-carboxylysine **126** could have been interfering with the analysis. Accordingly, the potential conjugation of triazine methyl ester **142** to unfunctionalised norbornene **146** was investigated through incubating **142** and **146**, at 40 mM with temperatures of up to 80 °C and stirring in THF/H_2O (1:1) overnight (Scheme 4.2b). Unfortunately, no detectable formation of cycloaddition product **147** was observed.



Scheme 4.2: Attempted reaction of a) triazine methyl ester **142** to norbornyl-carboxylysine **126** and b) **142** to unfunctionalised norbornene **146**. Even at elevated temperatures of 80 °C, cycloaddition reactions were unsuccessful.

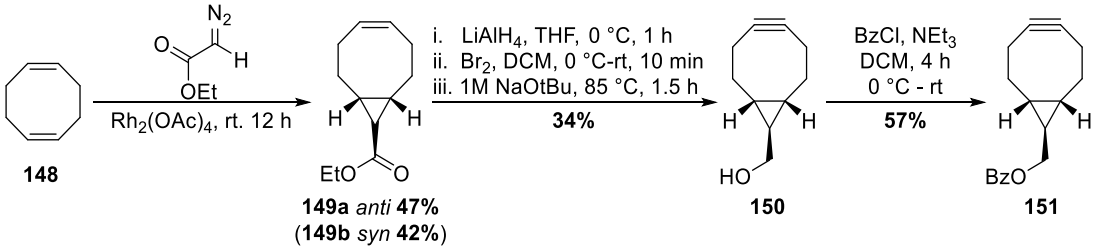
4.3 Bicyclononyne as the Dienophile

4.3.1 Synthesis

Due to the failure of the norbornene-triazine cycloaddition, attention was directed towards a coupling partner with a higher degree of ring strain. Unhindered strained bicyclo[6.1.0]nonynes cross-link to tetrazines more rapidly than norbornene with approximate rate constants of 10^2 - 10^4 M^{-1} s^{-1} (Figure 4.1b).[87]

To assess the reactivity of BCN to TrzAla methyl ester **142**, benzoyl-protected bicyclononyne (BCN-Bz) **151** was generated in five steps from 1,5-cycloctadiene (Scheme 4.3).[165] Rhodium catalysed cyclopropanation of cyclooctadiene **148** gave a mixture of bicyclononene ethyl esters **149a** (*anti*) and **149b** (*syn*) in an overall yield of 89%. *anti*-Bicyclononene ethyl ester **149a** was converted to the corresponding alkyne by sequential ester reduction, dibromination and double elimination to give **150** in 34% yield. **150** was protected using benzoyl chloride to generate cycloalkyne **151** in a yield of 57%.[166]

The moderate yield for the three-step transformation of *anti* bicyclononene ethyl ester **149a** to alkyne **150** can be attributed to formation of an unknown side-product during the base-catalysed elimination of the corresponding dibromo-bicyclononane species. Analysis of the unknown species using NMR spectroscopy suggests presence of an unsymmetrical diene, although it was not possible to assign the spectra.

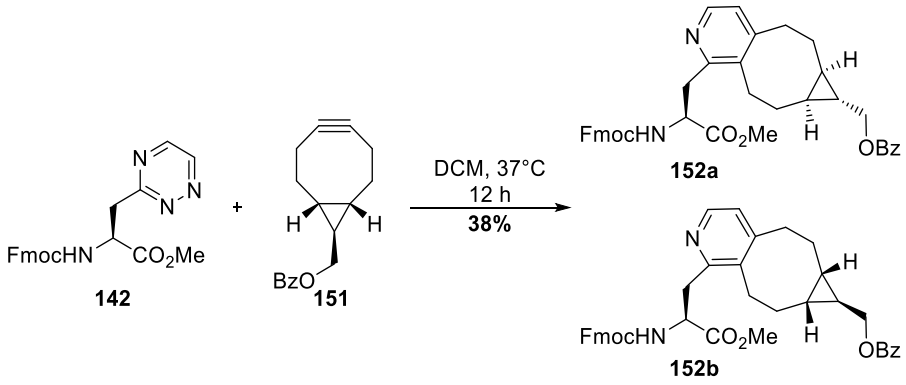


Scheme 4.3: Generation of benzoyl-protected BCN **151** through the stepwise reduction, dibromination and double elimination of a single diastereoisomer of bicyclononene, **149a**. Subsequent protection with benzoyl chloride yielded BCN-Bz **151**.

4.3.2 Cross-linking of Bicyclononyne to Triazinylalanine

To assess the reactivity of BCN-Bz **151** to TrzAla methyl ester **142**, an equimolar mixture of **151** and **142** was incubated at 37 °C for 12 hours at high concentration (65 mM in dichloromethane, (Scheme 4.4)). Diasteromeric bicyclononapyridyls **152a** and **152b** were isolated together by column chromatography in 38% yield; unreacted triazine **142** was also isolated. In this case, the moderate yield can be attributed to the instability of bicyclononyne

151, which was confirmed by apparent degradation of BCN-Bz **151** over time by NMR spectroscopy.



Scheme 4.4: Cycloaddition reaction of Fmoc-TrzAla-OMe **142** to BCN-Bz **151** at 37 °C yielded a mixture of diastereoisomers **152a** and **152b** in an overall yield of 38% yield.

4.3.3 Rate Determination

Encouraged by the successful cross-linking of triazinylalanine methyl ester **142** to BCN-Bz **151**, a series of experiments were conducted to determine the bimolecular rate constant for the reaction. This was achieved by measuring the rate of product formation by HPLC at 37 °C, using a mixture of authentic **152a** and **152b** as a concentration standard.

152a and **152b** were serially diluted 3–fold and analysed by HPLC using a linear gradient (5% H_2O in MeCN to 95% H_2O). The resultant data points were plotted as a function of concentration to generate a concentration gradient that was used to determine the rate of formation of **152a** and **152b** in all subsequent experiments. Product formation in MeCN was measured by sequential injection (onto the same HPLC column and gradient as for the calibration) over a reaction time of approximately 18 hours at 37 °C using 1 mM Trz methyl ester **142** and a range of BCN-Bz **151** concentrations (1 mM, 2 mM, 4 mM and 8 mM). After every analysis, 2 µl of MeCN was injected onto the column and the column washed using the same linear gradient. The data points obtained for each concentration variant were then fitted against time (Figure 4.2a); the resultant calculated gradients were then plotted as a function of [BCN-Bz] and fitted to 2nd order kinetics (Rate = k_2[BCN-Bz][Fmoc-TrzAla-OMe]), under the assumption of low substrate consumption during the measured time-course, to yield a 2nd order rate constant of $3.95 \pm 0.22 \times 10^{-4}$ M^{-1} s^{-1} (Figure 4.2a inset).

It has been suggested that the rate of the analogous tetrazine cycloaddition reaction can be significantly increased when in aqueous media.[167] This phenomenon is believed to mainly be the result of stabilising interactions between water and the activated transition state, which reduces the Gibbs energy of activation. **142** and **151** are poorly water-soluble which prevented the analogous experiment from being carried out in water; however the reaction

rate measured in 10% H_2O in MeCN increased slightly to $5.14 \pm 0.37 \times 10^{-4}$ M^{-1} s^{-1}, suggesting that the rate in biological media will be slightly higher (Figure 4.2b and inset).

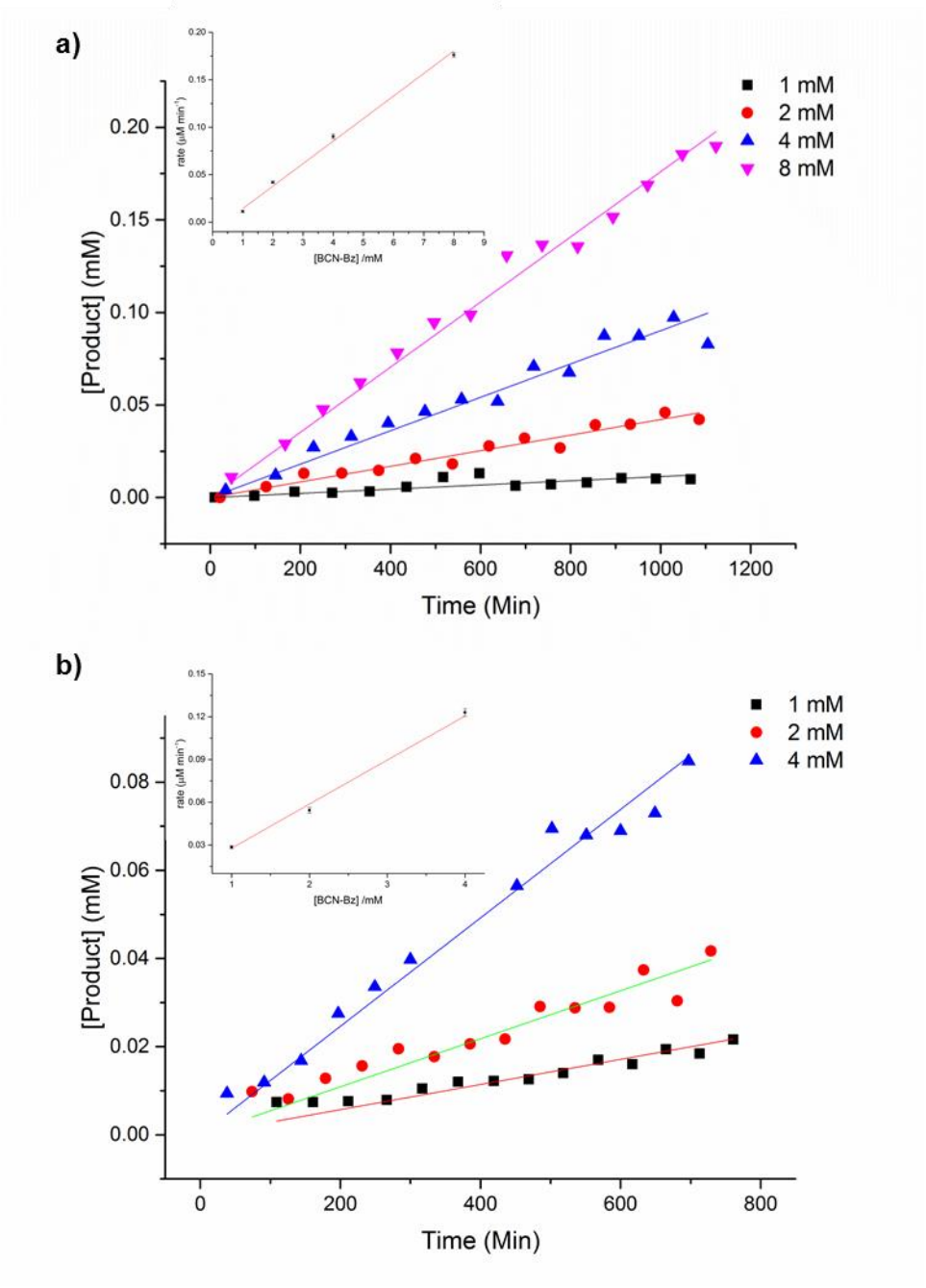


Figure 4.2: Rate of reaction of Fmoc-TrzAla-OMe **142** with increasing concentrations of BCN-Bz **151** at 37 °C determined by HPLC measurement of formation of bicyclononapyridyls **152a** and **152b** in a) pure MeCN and b) 10% H_2O in MeCN. Rate data were fitted to 2nd order kinetics (Rate = k_2[BCN-Bz][Fmoc-TrzAla-OMe]) under the assumption of low substrate concentration to give second order rate constants of a) $3.95 \pm 0.22 \times 10^{-4}$ M^{-1} s^{-1} and b) $5.14 \pm 0.37 \times 10^{-4}$ M^{-1} s^{-1}.

Comparison of this rate of reaction (0.3 - 0.5 × 10^{-3} M^{-1} s^{-1}) to other reactions in this class[44] suggests that although the cycloaddition of bicyclononyne **151** to triazine **142** is significantly slower than the corresponding reaction with tetrazines, it is comparable with other known bioorthogonal reactions such as the Staudinger ligation, which has second order rate constants in the low 10^{-3} M^{-1} s^{-1} range.[168]

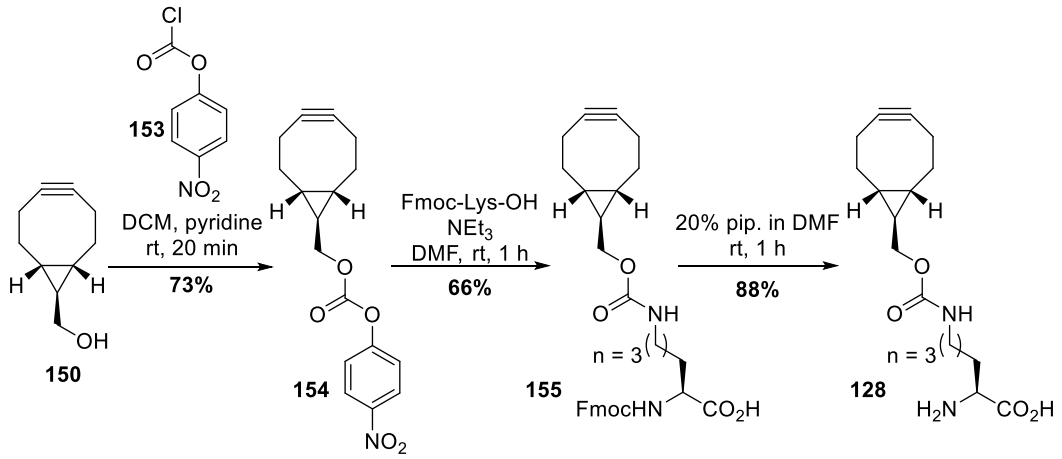
4.4 Application of 1,2,4-Triazine Cycloaddition Reactions *in vitro*

To investigate whether the cycloaddition of triazines to bicyclononynes has utility as a bioorthogonal probe, the cross-linking reaction needed to be demonstrated *in vitro*. For this purpose the synthesis and subsequent genetic incorporation of bicyclononyl-carboxylysine **128** (Figure 4.1a) into a model protein was required.

4.4.1 Genetic Incorporation of Bicyclononyne into a Model Protein

4.4.1.1 Bicyclononyl-lysine Synthesis

Bicyclononynol **150** (previously synthesised in Section 4.3.1) was converted to its corresponding activated ester **154** using 4-nitrophenyl chloroformate **153** and reacted with Fmoc-Lys-OH to give Fmoc-protected BCN-lysine **155** in an overall yield of 48% (Scheme 4.5). Fmoc-deprotection using 20% piperidine in DMF yielded the desired BCN-carboxylysine **128** in 88% yield.



Scheme 4.5: Synthesis of BCN-Lysine **128** via 4-nitrophenylester **154**. Displacement of activated ester **154** with Fmoc-Lys-OH gave **155** in 66% yield. Subsequent deprotection with 20% piperidine in DMF yielded the desired product **128** in 88% yield.

4.4.1.2 Aspartate α-decarboxylase (ADC) as a Model Protein

Aspartate α-decarboxylase (ADC) is responsible for the decarboxylation of L-aspartate to form β-alanine as part of the pantothenate biosynthetic pathway. As shown in Figure 4.3, ADC exists as a tetramer and is known to bind to and be activated by PanZ in a coenzyme A

(CoA) dependent manner.[169] An ADC construct, pCDF-PylT-ADC(K9X), containing a gene encoding ADC with Lys9 mutated to the amber codon (ADC(K9X)), together with an $MbtRNA_{CUA}$ gene from *Methanosarcina bakeri* encoding pyrrolysyl-tRNA(PylT) was previously generated in the group by Michael Rugen. Once co-transformed with a cognate pyrrolysyl-tRNA synthetase, the resultant expression system can be used to incorporate a corresponding UAA into ADC in place of Lys9. The utility of this amber suppression system has been demonstrated by the co-translational incorporation of norbornyl-carboxylysine **126** using *wt* pyrrolysyl-tRNAsynthetase (PylRS) into this position in ADC (data not shown). Hence, this system will be used to incorporate bicyclononyl-carboxylysine **128** in place of lysine 9 into ADC using a commercially available cognate pyrrolysyl-tRNA synthetase, BCN-PylRS, which was specifically developed to recognise BCN-lysine **128**.

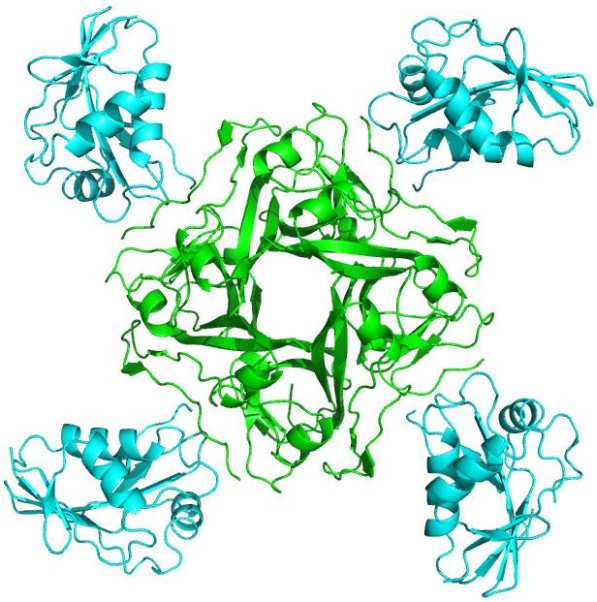


Figure 4.3: Crystal structure of tetrameric ADCT57V (shown in green) with each ADC monomer bound to a PanZ monomer (shown in blue) to give a heterooctameric complex. PDB file: 4CRZ.

4.4.1.3 Genetic Incorporation of Bicyclononyl-lysine into ADC

pCDF-pylT-ADC(K9X) and BCN-PylRS were co-transformed into *E. coli* ΔDZ(DE3) competent cells. ΔDZ(DE3) cells contain chromosomal deletions in both PanD (inactive ADC) and PanZ. Dual resistance agar plates (spectinomycin and ampicillin) were inoculated with cells that had been concentrated and resuspended and the plates were incubated overnight at 37 °C. Previous co-transformations utilising dilute cells had resulted in no colonies.

A mini culture of ΔDZ(DE3) pCDF-pylT-ADC(K9X) BCN-PylRS was used to inoculate 1 L of LB media containing ampicillin and spectinomycin (100 mg/L). Protein overexpression was induced after 4 hours using IPTG when an OD_{600} between 0.4-0.6 was reached, at which time BCN-lysine **128** was also added (100 μM). Cells were harvested by centrifugation, protease inhibitor (EDTA free) was added, and 2/3 of the cells were lysed by sonication. The resultant lysate was cleared of cellular debris and purified using standard Ni-NTA column chromatography. Analysis by SDS-PAGE showed two faint bands corresponding to the mass of His_6–ADC(K9BCN), demonstrating that BCN-lysine had been incorporated into the protein at the desired position (Figure 4.4). Fractions containing the protein were concentrated using a spin filter to ~300 μl and the concentration was determined to be ~1 μM by measuring absorbance at 280 nM. It was observed that some protein had precipitated on the filter. The lower than expected concentration obtained after purification of putative ADC(K9BCN) and the observation of protein precipitation both indicate that incorporation of BCN-lysine **128** in place of Lys9 into ADC has affected stability of the protein. In retrospect, this is unsurprising as Lys9 is found in the active site of ADC and is involved in critical hydrogen bonding with His11 and Tyr58 (Figure 4.5a). Clearly, disruption of this interaction has rendered the protein unstable and prone to precipitation.

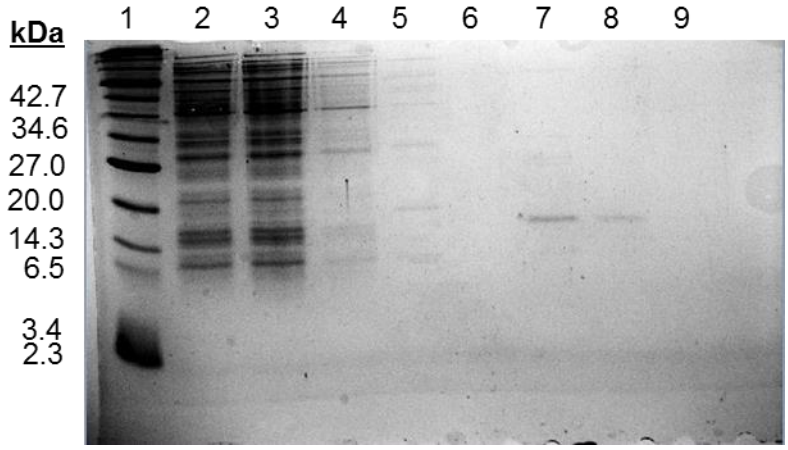


Figure 4.4: 15% SDS PAGE gel of purification of putative His_6–ADCK9BCN-K by nickel column; lanes: 1 MW marker, 2 flow through, 3 wash with *Lysis Buffer 2*, 4 wash with *Wash Buffer 2*, 5-9 wash with *Elution Buffer 2*. Protein corresponding to the correct mass is present as bands in lanes 7 & 8.

Interestingly, HRMS analysis of the small amount of protein that remained in solution after concentration revealed a mass of 15314.5 Da (Figure 4.5b), which corresponds to the mass of *wt* ADC (the mass of ADC(K9BCN) was calculated to be 15492.3 Da). It should be noted that it would not have been possible for endogenous lysine to have been incorporated instead of BCN-lysine **128** using the amber suppression expression system described.[87] Therefore, presence of lysine at this position must be the result of hydrolysis of the

BCN-carbamate modification. It is possible that His11, which is found adjacent to the site at which BCN-lysine **128** has been incorporated (Figure 4.5a), could catalyse the hydrolysis of the BCN-carbamate to yield lysine. It appears that the small amount of protein obtained that had not precipitated was stable as it resembled *wt* ADC and lacked the BCN-carbamate modification. Hence, His_6–ADC(K9BCN) was successfully generated but the mutant protein is too unstable to be of any use. Clearly, if BCN-lysine **128** is to be genetically incorporated into a protein and remain stable, a different amber suppression system that will either incorporate **128** at an alternative position away from the active site of ADC, or incorporate BCN-lysine **128** in place of a residue in a different protein, is needed.

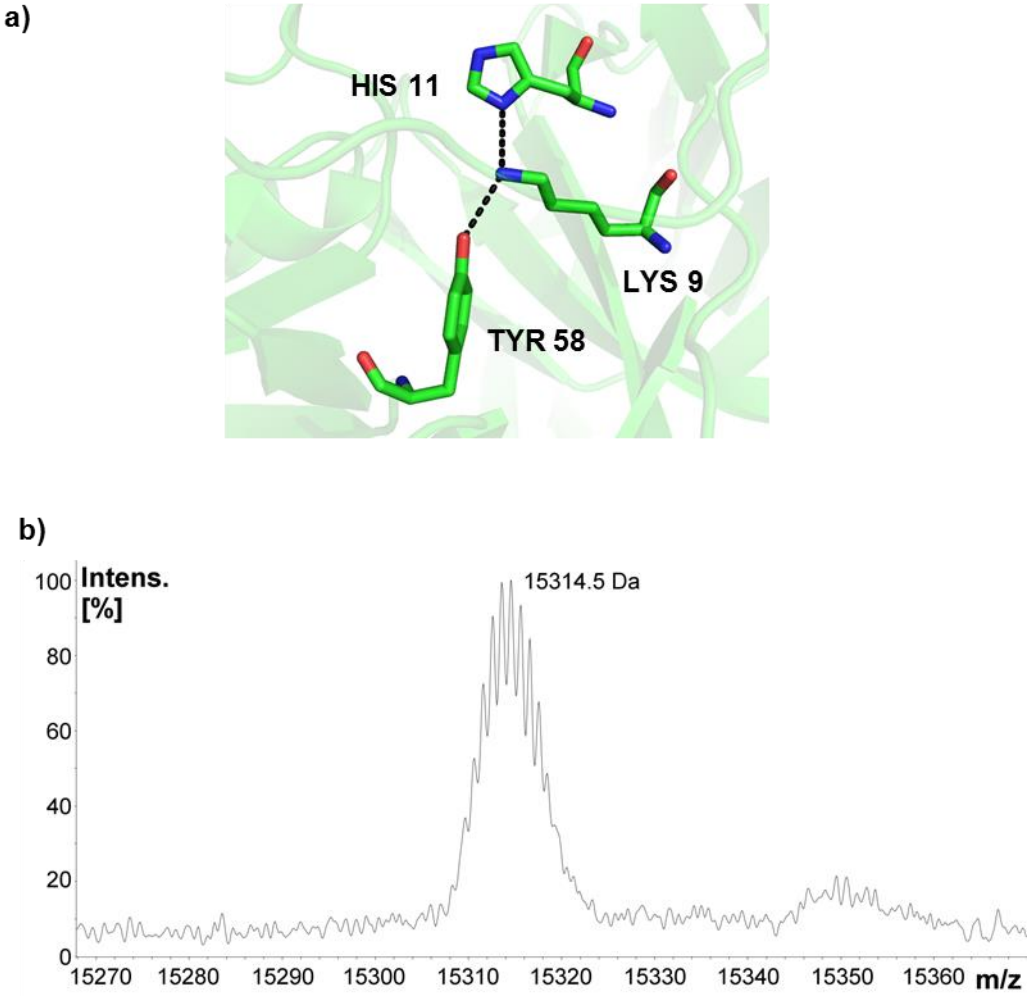


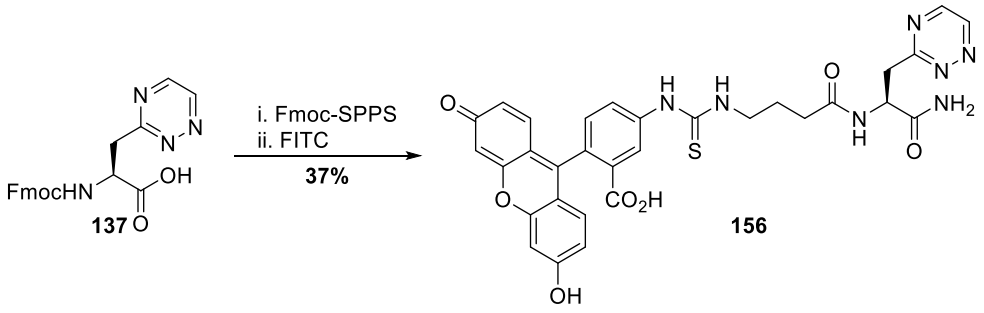
Figure 4.5: a) Critical hydrogen bonding interactions between Lys9 and His11/Tyr58 in the active site of ADC, PDB file: 4CRZ; b) Deconvoluted HRMS of putative His_6–ADC(K9BCN) gives a mass of 15314.5 Da, which corresponds to a mass of 15315.2 Da calculated for *wt* His_6–ADC (a mass of 15492.3 Da is expected for His_6–ADC(K9BCN)). This result is indicative of BCN-carbamate hydrolysis.

4.4.2 Cycloaddition of 1,2,4-Triazine to *trans*-cyclooctene *in vitro*

Due to the successful reaction of triazinylalanine methyl ester **142** with BCN-Bz **151**, attention was turned to the cycloaddition of triazine with highly strained *trans*-cyclooctene (TCO) *in vitro*. This dienophile reacts with tetrazine derivatives with rate constants between 10^3 - 10^4 M^{-1} s^{-1}.[43,44] TCO therefore needed to be incorporated into a model protein and a fluorescent triazine-containing probe needed to be generated.

4.4.2.1 Synthesis of a Triazine-containing Fluorescent Probe

Incorporation of triazinylalanine **137** into FITC-labelled peptide **156** was achieved on Rink amide resin using standard Fmoc SPPS and on-resin fluorescent labelling (Scheme 4.6). Following purification by mass-directed HPLC, TrzAla probe **156** was obtained in 37% yield. Incorporation of TrzA **137** into a peptide demonstrates the compatibility of **137** with the Fmoc-strategy for peptide synthesis. Moreover, the rapid and straightforward generation of triazine probe **156** demonstrates that triazinylalanine **137** can indeed rapidly functionalise probe molecules at a late stage.



Scheme 4.6: Incorporation of Fmoc-compatible building block **137** into short fluorescent peptide **156** using Fmoc SPPS and N-terminal FITC ligation.

4.4.2.2 Incorporation of *trans*-cyclooctene into ADC[i]

TCO-PEG_3-maleimide **157** is commercially available and can be derivatised onto cysteine residues in proteins through conjugate addition of the maleimide functionality (Figure 4.6a). ADC contains two cysteine residues located at positions 26 and 78 (Figure 4.6b). Cys78 is exposed on the surface of the protein whilst Cys26 is buried in the active site; as a result, it should be possible to selectively react Cys78 with TCO-PEG_3-maleimide **157**.

[i] ADCS25A was generated and expressed by Diana Monteiro; the cell pellet was lysed and purified by Zoe Arnott

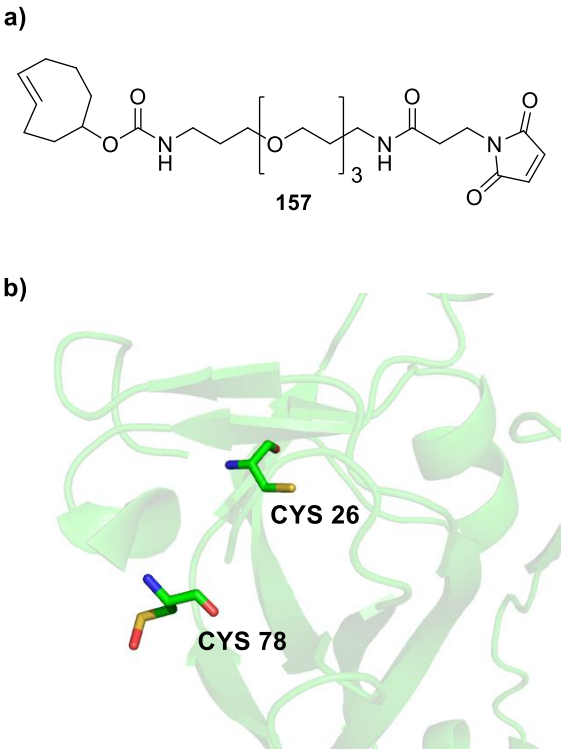


Figure 4.6: a) Chemical structure of TCO-PEG$_3$-maleimide **157**; b) Orientations of Cys26 and Cys78 in ADC. PDB file: 4CRZ.

ADC mutant, ADCS25A (200 μM), was incubated with 2.2 equivalents of TCO-PEG$_3$-maleimide **157** for 1 hour at room temperature in *SEC Buffer* and applied directly to a desalting column to remove unreacted **157**. Analysis by mass spectrometry revealed two peaks at 16294.3 Da and 16817.6 Da; which corresponded to the single labelled and double labelled protein respectively (Figure 4.7). Use of one equivalent of **157** may prevent labelling of both cysteine residues in future. Nevertheless, the mixture of single and double-labelled ADC-TCO that had been generated was used in the following *in vitro* cycloaddition reaction, with the intention to optimise labelling of **157** to ADC in the future.

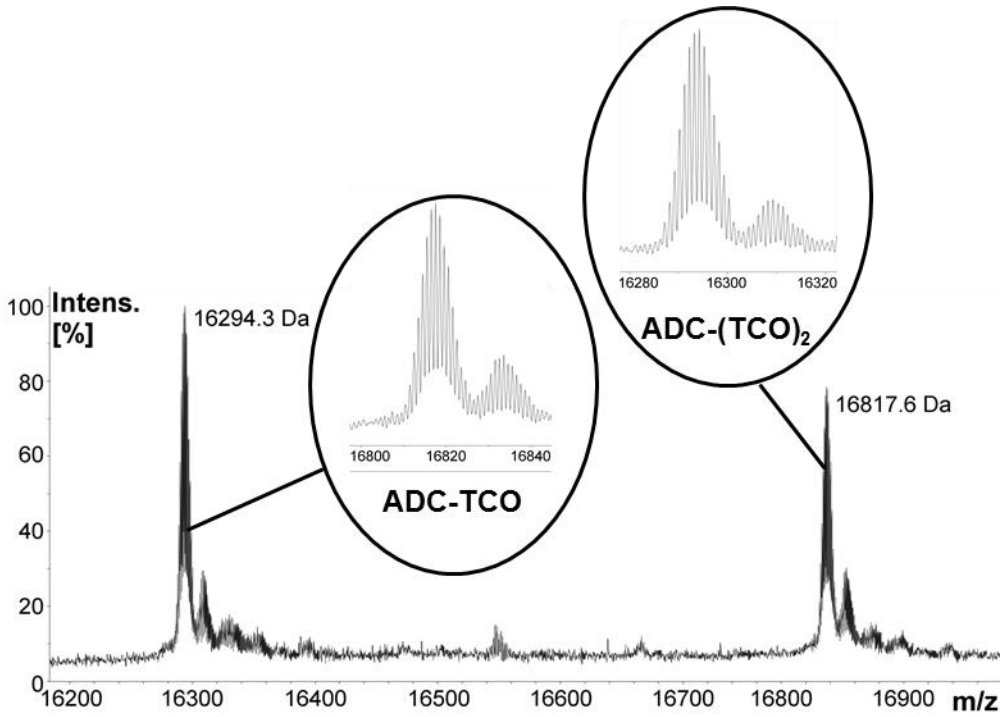


Figure 4.7: Labelling of ADCS25A with TCO-PEG_3-maleimide **157**. Deconvoluted HRMS reveals two peaks corresponding to single labelled (ADC-TCO) and double labelled (ADC-$(TCO)_2$). Expected masses were calculated as 16293.11 Da for (ADC-TCO) and 16818.34 Da for (ADC-$(TCO)_2$), which correspond to experimental masses of 16294.3 Da and 16817.6 Da respectively.

4.4.2.3 Labelling of *trans*-cyclooctene with a Triazinylalanine-containing Fluorescent Probe

A mixture of single labelled (ADC-TCO) and double labelled (ADC-$(TCO)_2$) (combined concentration 30 μM) was incubated overnight at 37 °C with 10 equivalents of TrzAla probe **156** in ITC buffer. UV Visualisation of the subsequent SDS PAGE gel revealed a fluorescent band at an approximate mass of 66 kDa for the unboiled reaction sample (Lane 3, Figure 4.8a). This corresponds to the fluorescently labelled ADC tetramer, which should have a mass of ~69 kDa. Staining of the gel using coomassie stain revealed that the unboiled reaction sample had partially unfolded into both the dimeric and monomeric ADC (Lane 8). Unreacted TrzAla probe **156** is displayed on the gel as two bands (representing the interconverting quinoid and lactone forms of fluorescein) at ~6 Kda and ~20 kDa (Lanes 2 and 3); saturating the point at which fluorescent labelling of monomeric ADC (~17 kDa) would be observed (Lane 3). Analysis of the reaction sample by HRMS revealed a peak at 17430.8 Da corresponding to fluorescently labelled ADC-$(TCO)_2$, which has a calculated mass of 17430.9 Da; indicating that the cross-linking of *trans*-cyclooctene to triazine **156** was successful (Figure 4.8b). However, it is evident from the mass spectrometry trace that a significant amount of unreacted ADC-TCO and ADC-$(TCO)_2$ is still present, suggesting that

the reaction procedure needs to be optimised by incubating ADC-TCO with a higher concentration of the fluorescent triazine probe **156**.

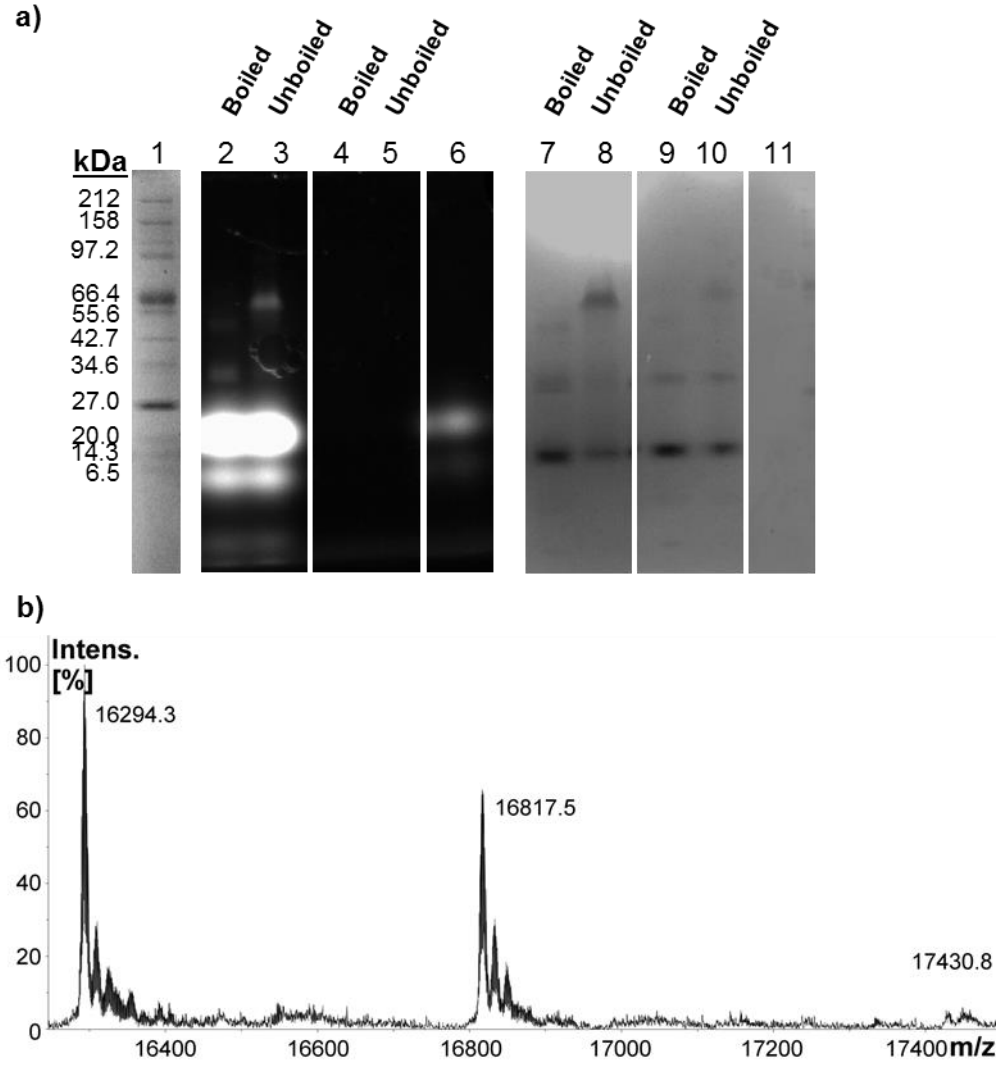


Figure 4.8: Results of incubation of a mixture of ADC-TCO and ADC-$(TCO)_2$ with fluorescent triazine probe **156**; a) 10% Tris-trycine SDS-PAGE gel; lanes: 1 MW marker, 2 & 3 UV visualisation of boiled and unboiled reaction mixture, 4 & 5 UV visualisation of boiled and unboiled unlabelled ADCS25A, 6 UV visualisation of TrzA probe **156** on its own. Lanes 7-11 correspond to lanes 2-6 but coomassie stained. b) Deconvoluted ESMS of reaction mixture: peaks at 16294.3 Da and 16817.5 Da are unlabelled ADC-TCO and ADC-$(TCO)_2$ respectively, the peak at 17430.8 Da corresponds to fluorescently labelled ADC-$(TCO)_2$ which has a calculated mass of 17430.9 Da.

Shortly after this labelling reaction was carried out, Kamber *et al.*[170] reported a complementary study of the cycloaddition reaction between a range of 1,2,4-triazin-6-yl derivatives and *trans*-cyclooctene. This reaction was demonstrated *in vitro* through the genetic incorporation of triazinylphenylalanine into green fluorescent protein (GFP) and subsequent incubation with 100 equivalents of *trans*-cyclooctene.

4.5 Conclusions

The reactivity of 1,2,4-triazine to a range of strained dienophiles, including norbornene, bicyclononyne and *trans*-cyclooctene has been investigated. Alkyl triazine **142** was found to react readily with bicyclononyne at 37 °C with rate constants between 0.3 - 0.5 10^{-3} M^{-1} s^{-1}. Kamber *et al*. reported reaction rates between 1 and 7 × 10^{-2} M^{-1} s^{-1} for the cycloaddition of *trans*-cyclooctene to a range of triazine substrates- a rate approximately 30-fold higher than determined in Section 4.3.3. This ratio is similar to the approximately 15-fold difference in rate observed for the analogous reaction of tetrazine with bicyclononyne and *trans*-cyclooctene substrates by Lang *et al.*[87] Kamber *et al*.'s observations are therefore fully consistent with observed reaction rates between triazine and bicyclononyne. A small triazinylalalnine containing peptide **156** was synthesised, demonstrating the applicability of triazinylalanine **137** to the late stage functionalization of probe molecules. Furthermore, probe **156** was utilised in an *in vitro* context where its reaction with *trans*-cyclooctene was demonstrated

Norbornene does not react with triazine at a temperature appropriate for application in a biological system; this has since been confirmed in Kamber *et al*.'s[170] complimentary study. Since triazines do not react with norbornene, the triazine cycloaddition to bicyclononyne or *trans*-cyclooctenes could be used in a mutually orthogonal labelling sequence alongside a tetrazine-norbornene cross-linking reaction. The 1,2,4-triazine cycloaddition to bicyclononynes and *trans*-cyclooctenes has potential to be used as an alternative to the tetrazine-cycloaddition for applications in cellular and biochemical studies.

Chapter 5 Conclusions and Future Work

5.1 Application and Generation of Phosphohistidine Mimics

5.1.1 Summary

A known τ-phosphohistidine mimic, τ-phosphotriazolylalanine was used to assess the affinity of phosphohistidine towards two targets, Grb2-SH2 and YdiA. τ-pTz **15** was synthesised following an established synthetic route[25] and incorporated into peptides that either mimicked the binding sequence of a canonical pY-peptide or the binding site of PPS (Figure 5.1a).

Using fluorescence anisotropy and ITC, It was demonstrated that pTz peptide **91** bound to the SH2 domain of Grb2 with ~ 1000 fold weaker affinity than pY peptide **90**.[i] Analogous peptides containing phosphohomotriazole, phosphoserine and phosphothreonine were generated and their affinity for Grb2-SH2 was also assessed (**92**, **104** and **105** respectively). It was determined that none of these phosphopeptides were capable of interacting with the protein, indicating that SH2 domains do not merely have affinity for appropriately positioned phosphoryl groups. The mechanism by which Grb2-SH2 links tyrosine kinases to the Ras signalling pathway is well established[171] and it is therefore unlikely that histidine phosphorylation plays a role. However, it is possible that histidine phosphorylation plays a transient role in binding to one of the ~110 other phosphotyrosine binding modules; some of which have not been well characterised.[136]

A fluorescence polarisation assay was also used to assess the binding affinity of pTz peptide probe **110** towards YdiA, it was found that **110** did not interact with the protein. Since it is

[i] ITC experiments performed by Dr. Tom McAllister

established that YdiA binds to PPS when a histidine residue is phosphorylated, it is unlikely that these results accurately reflect the binding interaction of the two proteins.[138] It was decided that generating a peptide containing a phosphohistidine analogue to mimic the primary structure of PPS was not adequate to study this protein-protein interaction and the whole protein containing the post-translational modification was required. Consequently, a third generation τ-phosphohistidine mimic, τ-pTz-3 **111** was synthesised through the adaptation of a synthetic route previously devised in the group to incorporate allyl protection for the phosphoryl group. **111** has potential to be genetically incorporated into proteins using amber suppression (Figure 5.1b).

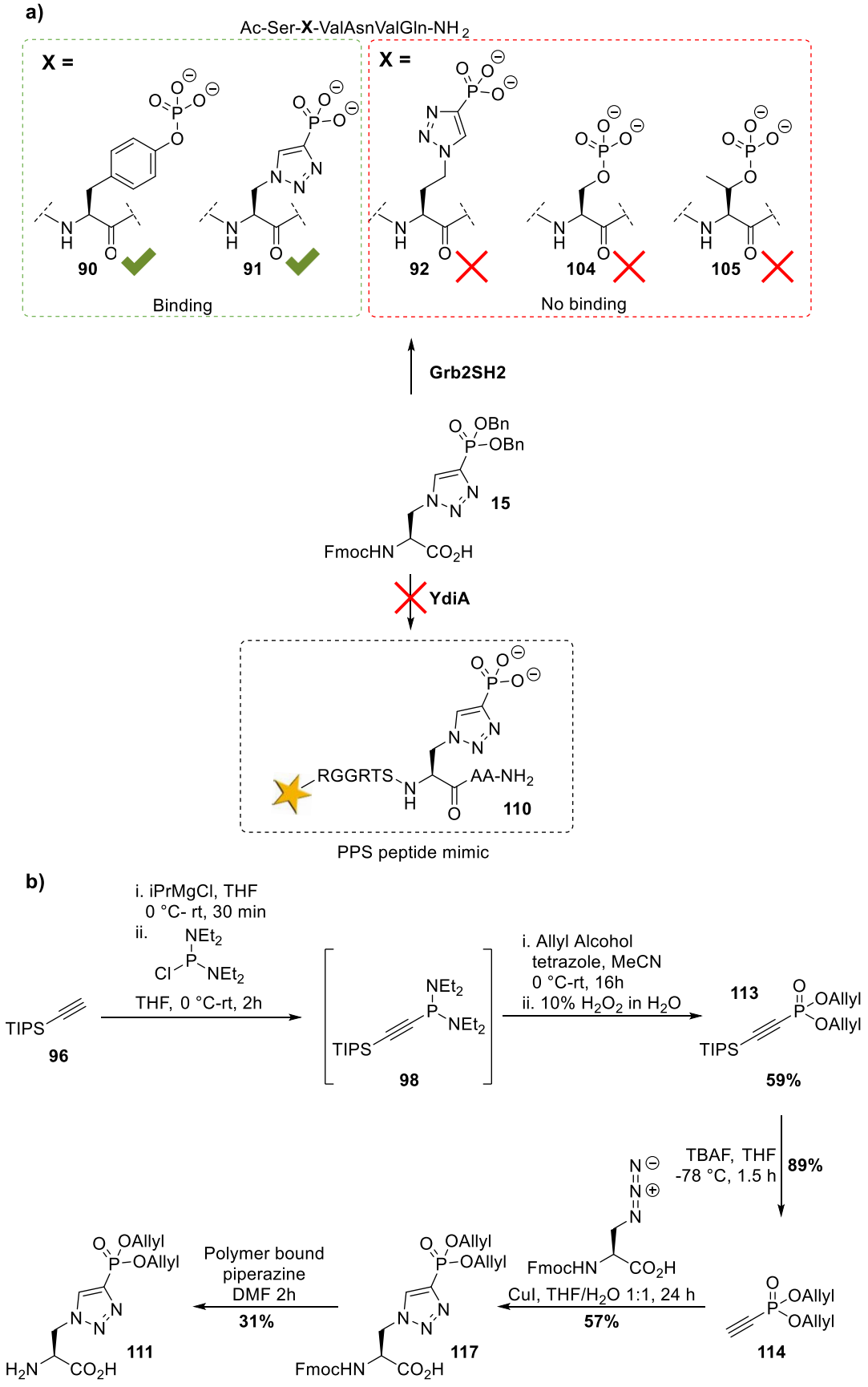


Figure 5.1: Summary of work achieved through a) using τ-pTz **15** in peptide sequences to probe binding interactions involving τ-phosphohistidine and b) the generation of third generation τ-phosphohistidine mimic **111** suitable for genetic encoding into proteins using amber suppression.

5.1.2 Future Work

The most obvious next-step as regards this part of the project is the scale-up of the synthesis to make pTz-3 **111** (Figure 5.1b). This is currently on-going and when achieved **111** will be initially screened against a variety of established amber suppression systems. If **111** is not found to be a substrate for any existing tRNA synthetases, a synthetase will be evolved to specifically incorporate pTz-3 **111**. Screening of existing amber suppression systems and potential evolution of a novel mutant synthetase will be carried out by Jason Chin and co-workers (MRC lab, University of Cambridge).

5.1.2.1 Assessing the Interaction of τ-Phosphohistidine analogues with other SH2 Domains

Following the discovery that τ-phosphotriazole peptide **91** has affinity towards Grb2-SH2, it would be interesting to establish whether τ-phosphotriazole interacts with other SH2 domains. Plasmids containing genes encoding over 40 different hexaHis-tagged SH2 domains are available in the group; the majority of which have been reported to bind to various phosphotyrosine-containing peptide sequences with high affinity. A number of these phosphotyrosine binding domains could be expressed and purified simultaneously, and used in fluorescence anisotropy competition assays, with a cognate pY-containing fluorescent peptide and a competing, analogous pTz peptide (Figure 5.2). In general, the crucial residues involved in binding to SH2 domains tend to be small in number; it should therefore be possible to generate a library of pY probes and analogous pTz peptides relatively rapidly by tandem Fmoc SPPS. Use of fluorescence anisotropy competition assays for this purpose would provide a rapid means in which to assess interaction of τ-pTz to a large number of SH2 domains. Furthermore, fluorescence polarisation assays do not require a large amount of material; therefore, each SH2 domain would only need to be expressed in small quantities on initial screening.

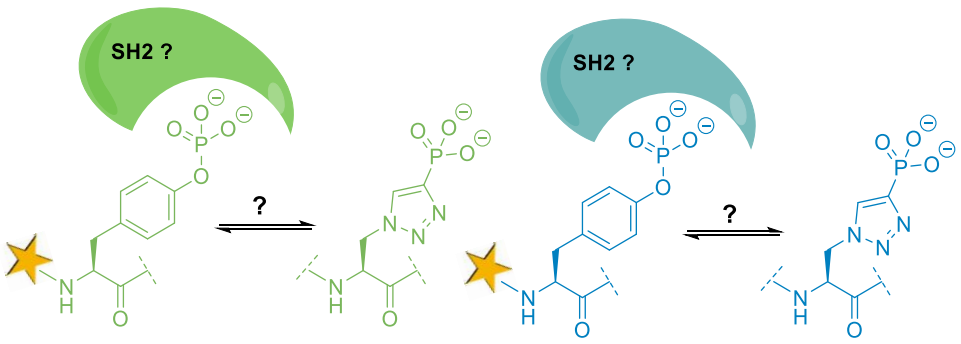


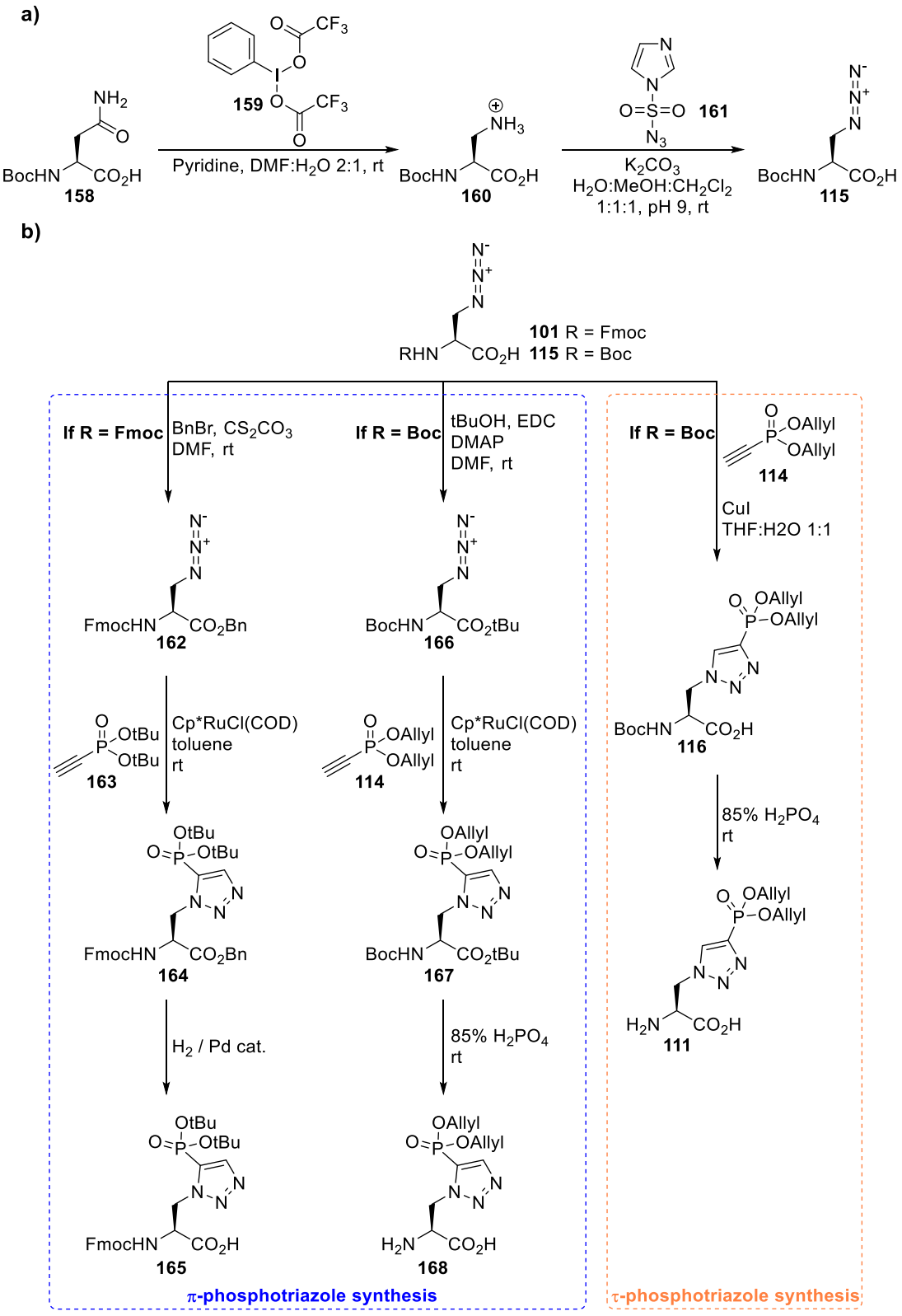
Figure 5.2: Proposed fluorescent anisotropy competition assay to rapidly assess binding interactions of pTz peptides to SH2 domains. A number of SH2 domains could be purified and their interaction with cognate pTz peptides could be evaluated.

5.1.2.2 Synthesis of π-Phosphotriazolylalanine

As mentioned in Section 1.1.3.3, π-phosphotriazolylalanine **12** has been shown to be an effective mimic of the other regioisomer of phosphohistidine, π-phosphohistidine **6**.[26,172] To date, there are no reported syntheses of either a π-pHis analogue fully compatible with the Fmoc-strategy for SPPS or an analogue suitable for genetic incorporation into proteins.

It should be possible to synthesise Fmoc SPPS compatible π-pTz **165** through the ruthenium catalysed (3 + 2) cycloaddition of azidoalanine **162** and phosphoalkyne **163** and subsequent hydrogenation of the benzyl group of **164** (Scheme 5.1b, left). Use of an Ru(II) catalyst in place of a Cu(I) catalyst has been reported to generate the desired π-isomer in high yields, providing the acid functionality of azidoalanine **162** is protected with a bulky protecting group.[22] tBu-protected phosphoalkyne **163** can be efficiently synthesised using the same strategy previously used to generate both allyl- and benzyl- protected phosphoalkynes **100** and **114** (Sections 2.2.3.2 and 2.4.1.1 respectively), but with *tert*-butyl alcohol rather than allyl alcohol or benzyl alcohol.[25]

Allyl protected π-pTz **168**, which has potential to be genetically incorporated into proteins, could be generated through the (3 + 2) Ru(II) catalysed cycloaddition of suitably protected azidoalanine **166** and allyl protected phosphoalkyne **114**, and subsequent removal of both Boc- and tBu- protecting groups using aqueous phosphoric acid (Scheme 5.1b, middle). Li *et al*.[173] reported the use of aqueous phosphoric acid to remove *tert*-butyl esters, *tert*-butyl carbamates and *tert*-butyl ethers in the presence of other acid sensitive functional groups. Although the orthogonality of aqueous phosphoric acid has not been demonstrated towards allyl functionalities, it is possible that this deprotection strategy is mild enough to effectively remove both tBu- and Boc- protection of π-pTz **167**, whilst leaving the allyl protection for the phosphoryl groups intact. Commercially available Boc-azidoalanine-OH **115** is purchased as the dicyclohexylammonium (DCA) salt. Low yields were reported on using this reagent in a Cu(I) catalysed cycloaddition with phosphoalkyne **114** to generate τ-pTz-3 **111** (See section 2.4.1.2.1). This was due to problems with the removal of the DCA salt during purification of crude τ-pTz-3 **111**. It is, however, possible to synthesise Boc-azidoalanine-OH **115** from Boc-asparagine **158** in a two-step synthesis (Scheme 5.1a).[174] Synthesised **115** could then be used in the ensuing cycloaddition reaction to generate π-pTz **167** (after protection of the acid functionality of **115**). If the synthetic route to allyl protected π-pTz **168** is particularly high-yielding, the same synthetic strategy could be used as an alternative route to access allyl protected τ-pTz-3 **111** (Scheme 5.1b, right).



Scheme 5.1: a) Synthesis of Boc-azidoalanine-OH **N** from commercially available Boc-asparagine-OH **158**.[174] b) Left and middle: Suggested synthetic routes towards π-pTz **165** and **168** suitable for either the Fmoc strategy for SPPS (left) or genetic encoding into proteins (middle); right: Proposed alternative synthesis of allyl protected τ-pTz-3 **111**.

5.2 The Triazine Cycloaddition to Strained Dienophiles as a Novel Bioorthogonal Reaction

5.2.1 Summary

In summary, a synthetic route to novel 1,2,4-triazin-3-yl-linked amino acids compatible with conventional peptide synthesis strategies has been developed using readily available and inexpensive starting materials (Figure 5.3a). The cycloaddition of an alkyl triazine **142** to a range of strained dienophiles including norbornene **146**, bicyclononyne **151** and *trans*-cyclooctene **157** was investigated. It was found that whilst cross-linking of **142** to norbornene **146** at ambient physiological temperatures was not possible, triazinylalanine methyl ester **142** reacted readily with strained BCN-Bz **151** at 37 ºC with second order rate constants between 0.3 and 0.5 $\times 10^{-3}$ M^{-1} s^{-1}, depending on solvent mix (Figure 5.3b, top). The triazine cycloaddition to strained dienophiles was also demonstrated *in vitro* through incubation of fluorescent triazine probe **156** with ADC, which had previously been derivatised with *trans*-cyclooctene (Figure 5.3b bottom).

The synthetic strategy towards triazinylalanine **137** involves an optimised Negishi-type cross-coupling between iodotriazine **136** and an alkyl organozinc reagent (Figure 5.3a). With this optimised synthetic route in place, it should be possible to cross-link iodotriazine **136** to a variety or alkyl- or aryl- zincate reagents which, as well as allowing for the tuning of cycloaddition rates, will allow easy access to a plethora of functionalised triazines. Use of triazinylalanine **137** in the rapid generation of triazine-containing fluorescent peptide probe **156** demonstrates the effective use of **137** to functionalise a probe molecule at a late stage. Furthermore, the stability of the triazine core in the harsh conditions used for cleavage in the Fmoc SPPS strategy indicates that triazines may show enhanced stability over other bioorthogonal reagents in physiological environments. This has been confirmed by Kamber *et al.*[170] who found that a range of monosubstitued triazines incubated with both PBS and cysteine at 37 °C were stable for over one week.

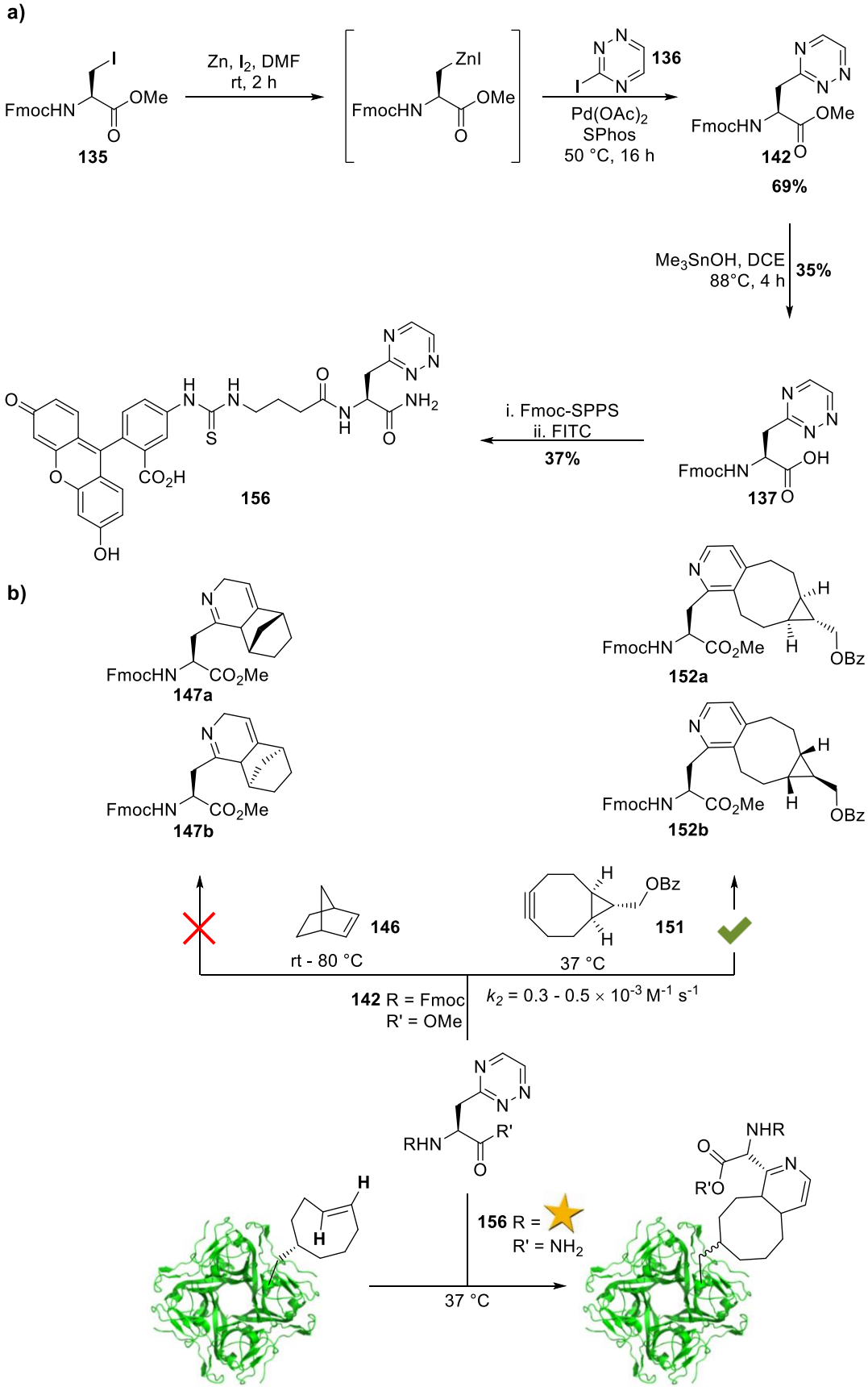


Figure 5.3: Summary of work achieved: a) Synthetic route to novel functionalised triazinylalanine derivative **137** and subsequent incorporation into fluorescent peptide **156** by Fmoc SPPS; b) Use of triazinylalanine **142** and triazine probe **156** as a reactive component in cycloaddition reactions with strained dienophiles in the reaction flask (top) and *in vitro* (bottom).

5.2.2 Future Work

5.2.2.1 Demonstration of the Triazine Cycloaddition to Bicyclononynes *in vitro*

Clearly, the most vital piece of future work in regards to this project is to demonstrate the cycloaddition of triazine to BCN on a protein. This is to be achieved by the genetic incorporation of BCN-lysine **128** into a model protein and subsequent incubation with triazinylalanine probe **156**. Previous efforts to do this through the genetic encoding of BCN-Lysine **128** into ADC in place of lysine 9 were unsuccessful (Section 4.4.1.3). This was due to the instability of the modified protein; therefore, an alternative model protein is required. PanZ is a small protein that is known to bind to PanD (the inactive form of ADC) and promote its activation to form ADC in a CoA dependent manner (Figure 5.4).[169] This protein has been well characterised in the group, is relatively stable and consistently shows high levels of protein expression. It is therefore a viable choice of protein for use in the generation of an alternative amber suppression system.

Codons normally encoding Ser29, Lys10 and Asp35 in PanZ will be mutated by site directed mutagenesis to the amber stop codon TAG, after incorporation of the *wild type* protein into a vector containing pyrrolysyl $tRNA_{CUA}$. Hence, 3 separate amber suppression systems will be generated to enable the incorporation of BCN-Lysine **128** into 3 different positions in PanZ; this will serve to minimise the risk of perturbing protein stability. As can be seen in Figure 5.4, the residues selected are part of a loop region and exposed on the surface of PanZ. They also do not appear to be involved in any critical interactions in the protein and as such, incorporation of BCN-lysine **128** in place of any of these residues should not affect protein stability. Work to generate the PanZ constructs is being undertaken by Matt Balmforth (PhD student) and is nearly complete.

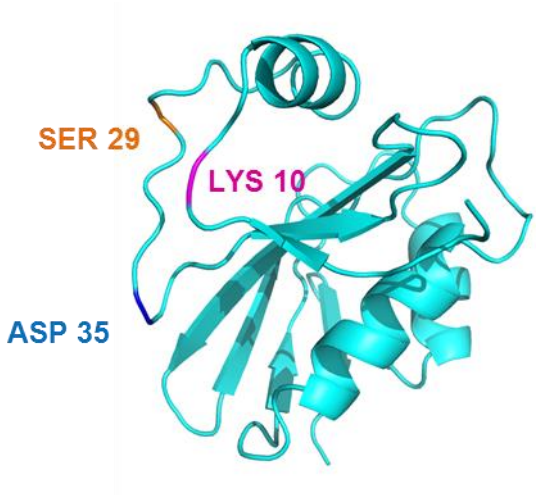


Figure 5.4: Crystal structure of PanZ with residues that will be mutated to TAG highlighted. PDB file: 4CRZ.

5.2.2.2 Genetic Incorporation of Triazinylalanine

The fully deprotected version of triazinylalanine, TrzA **169** is similar in structure to a range of tyrosine-based scaffolds that have been genetically incorporated into proteins in response to an amber codon using evolved tyrosyl-tRNA synthetases (Figure 5.5a).[145,163] It should be possible to screen **169** for compatibility towards one of these amber suppression systems. If **169** is found to be a substrate for an available tyrosyl tRNA synthetase, the corresponding amber suppression system could be used to incorporate **169** into proteins. Triazinylalanine **169** could easily be generated through the piperidine based Fmoc-group removal of **137**. Kamber *et al*[170] very recently demonstrated the site-specific incorporation of triazinylphenylalanine **170** into GFP using a mutant *mj*TyrRS in response to an amber codon (Figure 5.5b). Triazinylalanine **169** is smaller in size then **170** and therefore the capability of genetically encoding **169** over triazinylphenylalanine **170** into proteins will be advantageous. The ability to incorporate either bioorthogonal reacting partner of the triazine-strained dienophile cycloaddition into a protein will enable its use alongside other bioorthogonal reporting strategies in the same protein system.

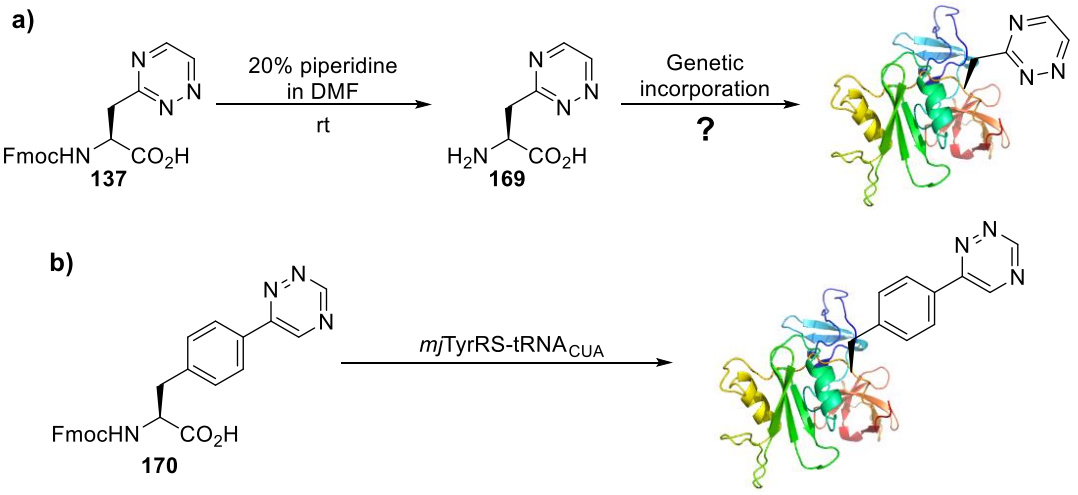


Figure 5.5: a) Deprotection of **137** using 20% piperidine in DMF should yield triazinylalanine **169**, which could then be screened as a substrate for a library of tyrosyl tRNA synthetases. b) triazinylalaninephenylalanine **170** has been genetically incorporated into GFP using suppression and a mutant TyrRS.[170]

5.2.2.3 Development of an Antagonistic Drug-Binding Screen

The slower rate of the triazine-BCN cycloaddition ($k_{2 =}$ 0.3-0.5 × 10^{-3} M^{-1} s^{-1}) in comparison to some of the other reported bioorthogonal cycloaddition reactions may lend itself for use in the development of an *in vivo* antagonistic drug-binding screen. A screen such as this could be used to demonstrate that a drug candidate has bound to its protein target in a living system. A prototype of this screen is illustrated in Figure 5.6. In this prototype, BCN-lysine **128** would be genetically encoded into the required protein near to (but not obstructing) the

binding site that is under investigation. The modified protein would then be incubated with a known substrate, which would have been previously derivatised with triazinylalanine or another functionalised triazine. The effective concentration between the incorporated BCN functionality and triazine will be increased when the substrate binds to the protein binding site; as a result, a cycloaddition reaction between the two components will occur and the substrate and the protein will be covalently linked (Figure 5.6a). If the drug candidate is also present and does bind to its target, then the triazine-containing substrate will face competition for the protein binding site. This will reduce the effective concentration of triazine and BCN as they will no longer be in close proximity to one another. Consequently, cross-linking between triazine and BCN will not occur (Figure 5.6b). Hence, if the drug candidate does bind to its intended protein target, no cycloaddition reaction will take place between triazine and the incorporated BCN-functionality. Conversely, if the drug candidate does not bind to its protein target, a cycloaddition reaction will occur and yield a covalently linked substrate-protein product. The products of this assay could simply be analysed by mass spectrometry.

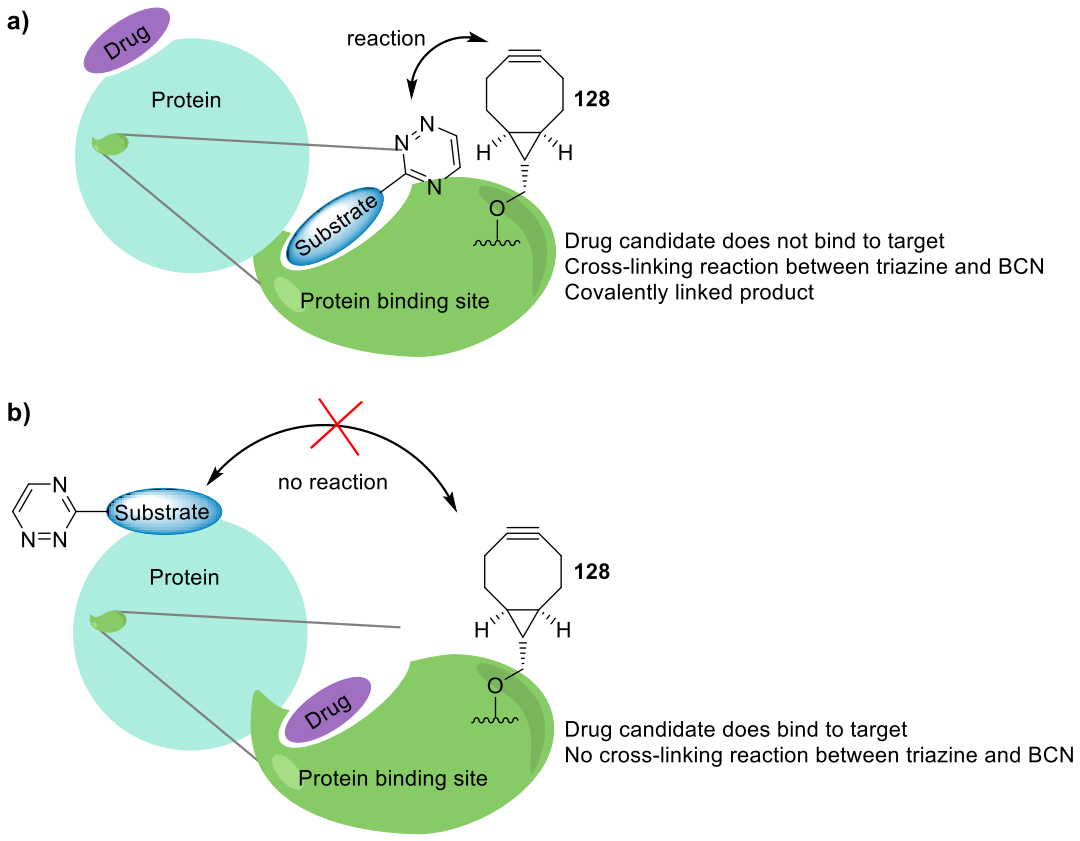


Figure 5.6: Use of the triazine cycloaddition to BCN as an antagonistic drug-binding screen: a) drug candidate does not bind to target binding site; b) drug candidate binds to target binding site.

The triazine-BCN cycloaddition is a perfect choice for application in this technology as its intrinsically slow reaction rate indicates that the reaction is not under diffusional control.

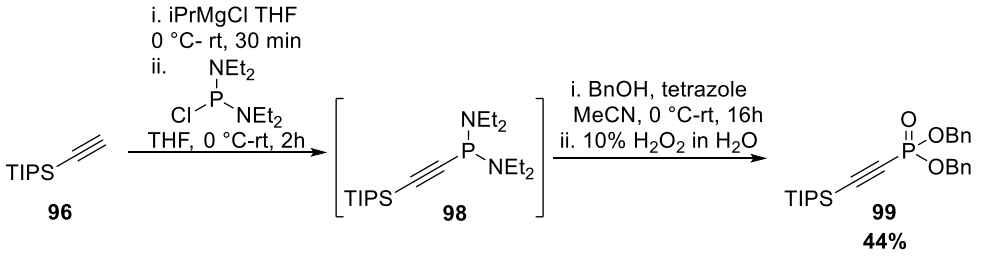
Hence if triazine-containing substrate is not in close proximity with the incorporated BCN-functionality, a cross-linking reaction should not occur. In addition, the enhanced stability of triazine over other bioorthogonal reagents makes it suitable for application in living systems. Work to establish whether this will be possible is needed.

Chapter 6 Experimental

6.1 Preparation of Small Molecules

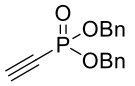
All reagents were purchased from Sigma Aldrich, Alfa Aesar, Merck or Fisher Scientific and were used without further purification unless otherwise stated. All solvents used were HPLC grade and mixtures are *v/v*. NMR data were collected using a Bruker DPX300 and analysed using MestReNova software. The following abbreviations are used to describe the multiplicity of signals: s = singlet, d = doublet, dd = doublet of doublets, qn = quintet, m = multiplet. *J*-values are given in Hz. IR spectra were recorded using a PerkinElmer spectrum one FTIR spectrometer, vibrational frequencies are reported in wavenumbers (cm^{-1}). Optical rotations were measured using a Schmidt+Haensch polartronic H532, $[\alpha]_D$ values are given in 10^{-1} deg cm^2 g^{-1}. High resolution mass spectrometry (HRMS) was carried out on a Bruker Daltonics micrOTOF using electrospray ionisation (ES), a Bruker maXis impact using ES or on a Micromass GCT Premier, using electron impact ionisation (EI). Column chromatography was carried out using silica gel or on an automated Biotage Isolera 1.3.3 using SNAP prepacked flash column silica cartridges from RediSep®Rf. TLC was performed on silica gel 60-F_{254} (Merck) with detection by fluorescence upon irradiation with UV light or staining by immersion in a solution of potassium permanganate, bromocresol green or ninhydrin and heating. Melting points were determined on a Stuart SMP3 melting point apparatus. Melting points obtained were uncorrected. Dropwise additions were performed manually or using a KD scientific syringe pump (KDS-100-CE). HPLC analyses were carried out on a Dionex HPLC system using an Ace UltraCore 2.5 Super C18, (50 x 2.1mm) and diode array as a detector; either on gradient A: mobile phase (H_2O/MeCN, 0.1% TFA), gradient (5-95% MeCN) or B: mobile phase (H_2O/MeCN), gradient (5-95% MeCN), at a flow rate of 0.5 ml/min. mass-directed HPLC purification was performed on an Agilent Techonologies 1260 infinity system. Lyophilisation was performed using a Virtis Benchtop K freezedryer or a BenchTop Pro with Omnitronics™ from VirTis SP Scientific.

Dibenzyl (triisopropylsilyl)ethynylphosphonate 99[25]



Isopropylmagnesium chloride (2.0 M in diethyl ether, 10 mL, 20 mmol, 1.1 eq.) was added dropwise to a solution of TIPS-acetylene (4.4 mL, 19.4 mmol, 1.02 eq.) in anhydrous THF (150 mL) at 0 °C. The reaction mixture was stirred over ice for 5 min, allowed to warm to room temperature and stirred for a further 30 min. The mixture was cooled once more to 0 °C, *bis*(diethylamino)chlorophosphine (4 mL, 19.0 mmol, 1 eq.) was added dropwise, the reaction mixture was warmed to room temperature and stirred for a further 3 hours; at which time complete consumption of *bis*(diethylamino)chlorophosphine was observed by TLC. The volatiles were concentrated *in vacuo* to leave an amorphous white solid, the apparatus back-filled with N_2, the flask quickly removed and purged with N_2. Anhydrous acetonitrile (150 mL) was added to the flask followed by anhydrous benzyl alcohol (4.5 mL, 44 mmol, 2.3 eq.). The stirred reaction mixture was cooled to 0 °C and tetrazole (0.45 M in MeCN, 93 mL, 42 mmol, 2.2 eq.) was added dropwise. The reaction was stirred at 0 °C for 15 min, allowed to warm to room temperature and then stirred at 30 °C for 20 min. The reaction mixture was concentrated to ca. 30 mL *in vacuo*, quenched with sat. NH_4Cl_{aq} (100 mL) and extracted with EtOAc (3 × 200 mL). The organics were combined and washed with H_2O_2 (10% (*v/v*) in H_2O, 2 × 30 mL), H_2O (2 × 30 mL), $Na_2S_2O_{3\,(aq)}$ (10% (*w/v*) H_2O, 2 × 100 mL) and brine (2 × 100 mL). The organics were dried ($MgSO_4$), concentrated *in vacuo* and the crude material was purified by column chromatography on silica gel, eluting the sample with 0.25% MeOH in CH_2Cl_2. *dibenzyl (triisopropylsilyl)ethynylphosphonate* **99**[25] (3.67 g, 8.3 mmol, 44%) was recovered as a colourless oil. $\boldsymbol{R_F}$ (0.25% MeOH in CH_2Cl_2) 0.41; $\boldsymbol{\delta_H}$ (300 MHz, $CDCl_3$): 7.42 – 7.33 (10 H, m, 2 × Ph-*H*), 5.14 (4 H, d, $^3J_{H\text{-}P}$ 8.5, 2 × OCH_2Ph), 1.14 – 1.07 (21H, m, Si(CH(CH_3)$_2$)$_3$ & Si(CH(CH$_3$)$_2$)$_3$); $\boldsymbol{\delta_C}$ (75 MHz, *d*6-DMSO): 130.3 (d, $^2J_{C\text{-}P}$ 7.6, Ph-C_1), 123.3 (s), 123.2 (s) & 122.2 (s) (2 × Ph-$C_{2\text{-}5}$) 102.6 (d, $^2J_{C\text{-}P}$ 38.2, PC*C*Si), 90.6 (d, $^1J_{C\text{-}P}$ 274.8, P*C*CSi), 63.2 (d, $^2J_{C\text{-}P}$ 5.2, 2 × OCH_2Ph), 13.2 (s) & 5.6 (s) (Si(CH(CH$_3$)$_2$)$_3$ & Si(CH(CH$_3$)$_2$)$_3$); $\boldsymbol{\delta_P}$ (121 MHz, $CDCl_3$): -8.01 (qn, $^3J_{P\text{-}H}$ 8.6); ***m/z*** (ES): Found: M*(+Na)* 465.1992 $C_{25}H_{35}O_3PSiNa$ requires 465.1985; **HPLC** (5-95% A): retention time 4.92 min, 100%.

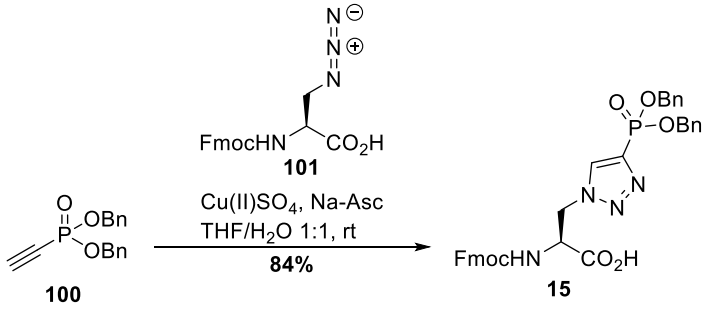
Dibenzyl ethynylphosphonate 100[25]



100

To a stirred solution of dibenzyl (triisopropylsilyl)ethynylphosphonate **99** (1.53 g, 3.23 mmol, 1 eq.) in THF, freshly opened TBAF (1.0 M in THF, 3.6 mL, 3.6 mmol, 1.1 eq.) was added with stirring at -78 °C and the reaction was stirred for 1.5 h; after which time complete consumption of **99** was observed by TLC. The reaction was warmed to 0 °C and quenched with sat. $NH_4Cl_{\ aq}$ ((*w/v*) 70 mL), generating a thick white suspension. H_2O (50 mL) was added and the resultant homogenous mixture was extracted with diethyl ether (4 × 150 mL). The combined ethereal layers were washed with brine (1 × 150 mL), dried ($MgSO_4$) and the volatiles were removed to leave a brown oil. The crude product was purified by column chromatography on silica gel eluting the sample with 2:1 hexanes:EtOAc to yield *dibenzyl ethynylphosphonate* **100**[25] (764 mg, 2.67 mmol, 83%) as a colourless oil. $\boldsymbol{R_F}$ (2:1 hexanes:EtOAc) 0.28; $\boldsymbol{\delta}_{\ \mathbf{H}}$ (300 MHz, $CDCl_3$): 7.41 – 7.30 (10 H, m, 2 × Ph-*H*), 5.11 (4 H, d, $^3J_{H\text{-}P}$ 8.6, 2 × OC*H*$_2$Ph), 2.91 (1 H, d $^3J_{H\text{-}P}$ 13.5, PCC*H*); $\boldsymbol{\delta}_{\ \mathbf{C}}$ (75 MHz, $CDCl_3$): 135.2 (d, $^3J_{C\text{-}P}$ 7.5, 2 × Ph-*C*$_1$), 128.7 (s), 128.6 (s) & 128.0 (s) (2 × Ph-*C*$_{2\text{-}5}$), 88.4 (d, $^2J_{C\text{-}P}$ 51.8, PC*C*H), 74.0 (d, $^1J_{C\text{-}P}$ 294.9, P*C*CH), 68.8 (d, $^2J_{H\text{-}P}$ 5.3, 2 × O*C*H$_2$Ph); $\boldsymbol{\delta}_{\ \mathbf{P}}$ (121 MHz, $CDCl_3$): -7.87 (dqn, $^3J_{P\text{-}H}$ 13.5, $^3J_{P\text{-}H}$ 8.6); ***m/z*** (ES): Found: M*(+H)* 287.0834 $C_{16}H_{16}O_3P$ requires 287.0832; **HPLC** (5-95% A): retention time 3.04 min, 100%.

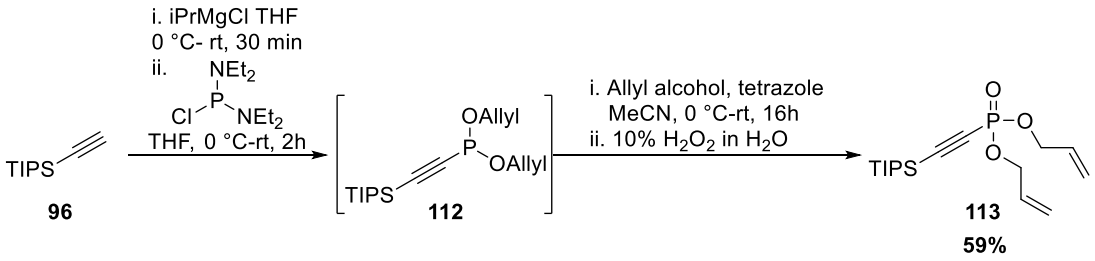
(2*S*)-3-(4-Dibenzylphosphonyl-[1,2,3]-triazol-1-yl)-2-(9H-fluoren-9-ylmethoxycarbonyl)-amino-propionate 15[25] (*Fmoc-pTz(OBn)$_2$-OH*)



A freshly prepared solution of copper (II) sulphate (84 mg, 0.52 mmol, 0.1 eq.) and sodium ascorbate (623 mg, 3.2 mmol, 0.6 eq.) in H_2O (10 mL) was added to a solution of dibenzyl ethynylphosphonate **100** (1.50 g, 5.2 mmol, 1 eq.) in anhydrous THF (10 mL). After 1 min, a solution of Fmoc-AzaAla-OH (1.67 g, 4.7 mmol, 0.9 eq.) dissolved in 1:1 THF:H_2O (40 mL) was added in one portion and the reaction mixture was stirred for 1.5 h at room temperature, until complete consumption of the azide was observed by TLC. The reaction mixture was diluted with $Na_2CO_{3\ aq}$ (10% *w/v* in H_2O, 100 mL) and extracted with diethyl

ether (2 × 130 mL). The aqueous phase was acidified to pH 1 by dropwise addition of conc. HCl_{aq} and extracted with EtOAc (5 × 130 mL). The organic extracts were combined, dried ($MgSO_4$) and the volatiles were removed *in vacuo* to leave an off-white foam, which was purified by column chromatography on silica (90:8:2 CH_2Cl_2:MeOH:AcOH). The resultant amorphous solid was dissolved in H_2O/MeCN and lyophilised to yield *Fmoc-pTz(OBn)$_2$-OH* **15**[25] (1.68 g, 2.63 mmol, 84%) as an amorphous white solid. $\boldsymbol{R_F}$ (90:8:2 CH_2Cl_2:MeOH:AcOH) 0.28; $\boldsymbol{\delta_H}$ (500 MHz, *d*6-DMSO): 8.53 (1 H, s, Tz-H_5), 7.84 (2 H, d, $^3J_{H\text{-}H}$ 7.5, Fmoc-H_4), 7.59 & 7.54 (2 H, 2 × d, $^3J_{H\text{-}H}$ 7.5 & 7.6, 2 × Fmoc-H_1), 7.40 – 7.33 (2 H, m, 2 × Fmoc-H_3), 7.33 – 7.24 (12 H, m, 2 × Ph-$H_{2\text{-}6}$ & 2 × Fmoc-H_2), 6.88 (1 H, s, N*H*), 5.10 – 4.91 (5 H, m, 2 × OCH_2Ph & H_β), 4.73 – 4.58 (1 H, m, $H_{\beta'}$), 4.23 – 3.99 (4 H, m, Fmoc-C*H*, Fmoc-CH_2 & H_α). $\boldsymbol{\delta_C}$ (125 MHz, *d*6-DMSO): 143.8 (s, Ph-C_1), 140.6 (s, 2 × Fmoc-C_{1a}), 136.1 – 135.9 (m, Tz-C_4), 132.3 (d, $^2J_{H\text{-}H}$ 32.6 Tz-C_5) 128.4, 128.2, 127.8, 127.6, 127.1, 125.1 & 120.0 (s, 2 × Ph-$C_{2\text{-}5}$ & 2 × Fmoc-$C_{1\text{-}4a}$), 67.29 (s, POCH_2Ph), 66.33 (s, Fmoc-CH_2), 65.6 (s, C_β), 52.02 – 51.66 (m, C_α & Fmoc-*C*H); $\boldsymbol{\delta_P}$ (121 MHz, *d*6-DMSO): 8.45 (qn, $^3J_{P\text{-}H}$ 7.3); ***m/z*** (ES): Found: M*(+H)* 639.2003 $C_{34}H_{32}N_4O_7P$ requires 639.1925; **HPLC** (5-95% A): retention time 3.58 min, 81%

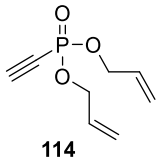
Diallyl (triisopropylsilyl)ethynylphosphonate 113



The reaction procedure was the same as was used to generate dibenzyl (triisopropylsilyl)ethnylphosphonate **99** apart from allyl alcohol (3.9 mL, 28.0 mmol, 2.3 eq.) was used instead of benzyl alcohol. Allyl alcohol was dried by stirring in $CaSO_4$ under anhydrous conditions for 16 h and distilled before use. Phosphine **112** was found to be volatile and therefore concentration of **112** *in vacuo* was performed at 50 mbar and 40 °C. The crude product was purified by column chromatography on silica gel, eluting the sample with 4:1 hexanes:EtOAc to give *diallyl (triisopropylsilyl)ethnylphosphonate* **113** (5.68 g, 16.5 mmol, 59%) as a colourless oil. R_F (4:1 hexanes:EtOAc) 0.31; $\boldsymbol{\delta_H}$ (500 MHz, CDCl3): 5.94 – 5.85 (2 H, m, 2 × allyl-H_2), 5.36 – 5.30 (2 H, m, 2 × allyl-H_{3a}), 5.20 (2 H, dd $^3J_{H\text{-}H}$ 10.4, $^2J_{H\text{-}H}$ 0.7, 2 × allyl-H_{3b}), 4.54 (4 H, dd, $^3J_{H\text{-}P}$ 7.7 & $^4J_{H\text{-}H}$ 3.5, 2 × allyl-H_1), 1.06 – 1.00 (21 H, m, Si(CH(CH_3)$_2$)$_3$ & Si(C*H*(CH$_3$)$_2$)$_3$); $\boldsymbol{\delta_C}$ (125 MHz, $CDCl_3$): 132.2 (d, $^3J_{C\text{-}P}$ 7.21, 2 × allyl-C_2), 118.4 (s, 2 × allyl-C_3), 107.5 (d, $^2J_{C\text{-}P}$ 38.3, PC*C*Si), 74.0 (d, $^1J_{C\text{-}P}$ 274.4, P*C*CSi), 67.4 (d, $^2J_{C\text{-}P}$ 5.2, 2 × allyl-C_1), 18.4 (s) & 10.9 (s) (Si(*C*H(CH$_3$)$_2$)$_3$ & Si(CH(*C*H$_3$)$_2$)$_3$); $\boldsymbol{\delta_P}$

(121 MHz, $CDCl_3$): -8.11 (qn, $^3J_{P\text{-}H}$ 9.01); ν_{max} (solid)/cm^{-1}: 2131 (disubstituted alkyne), 1268 (PO); ***m/z*** (ES): Found: M*(+H)* 343.1858 $C_{17}H_{32}O_3PSi$ requires 343.1853;

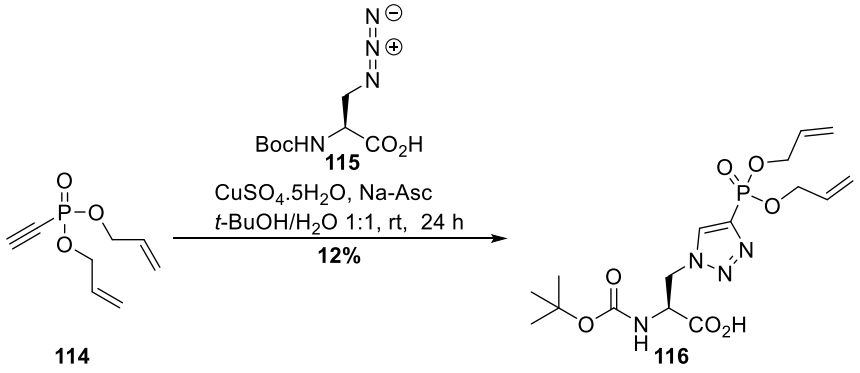
Diallyl ethnylphosphonate 114



114

Diallyl ethnylphosphonate **114** was synthesised in the same manner as dibenzyl ethnylphosphonate **100**, however diallyl (triisopropylsilyl)ethnylphosphonate **113** (4.96 g, 14.5 mmol, 1 eq.) was used instead of dibenzyl (triisopropylsilyl)ethnylphosphonate **99**. The crude product was purified by column chromatography (2:1 hexanes:EtOAc) to yield *diallyl ethnylphosphonate* **114** (2.40 g, 12.8 mmol, 89%) as a colourless oil. $\boldsymbol{R_F}$ (2:1 hexanes:EtOAc) 0.20; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$): 5.96 (2 H, dd, $^3J_{H\text{-}H}$ 10.8 & $^3J_{H\text{-}H}$ 5.6, 2 × allyl-H_2), 5.44 – 5.36 (2 H, m, 2 × allyl-H_{3a}), 5.31 – 5.26 (2 H, m, 2 × allyl-H_{3b}), 4.71 – 4.49 (4 H, m, 2 × allyl-H_1) 2.94 (1 H, d, $^3J_{H\text{-}P}$ 13.4, PCC*H*); $\boldsymbol{\delta_C}$ (75 MHz, $CDCl_3$): 135.2 (d, $^3J_{C\text{-}P}$ 7.5, 2 × allyl-C_2), 118.8 (s, 2 × allyl-C_3), 88.4 (d, $^2J_{C\text{-}P}$ 51.8, PC*C*H), 74.0 (d, $^1J_{C\text{-}P}$ 294.9, P*C*CH), 68.8 (d, $^2J_{C\text{-}P}$ 5.3, 2 × allyl-C_1); $\boldsymbol{\delta_P}$ (121 MHz, $CDCl_3$): -7.96 (dqn, $^3J_{P\text{-}H}$13.5, $^3J_{P\text{-}H}$ 8.9); ν_{max} (solid)/cm^{-1}: 3176 (C-H alkyne), 1255 (PO); ***m/z*** (ES): Found: M*(+Na)* 209.0338 $C_8H_{11}O_3PNa$ requires 209.0346.

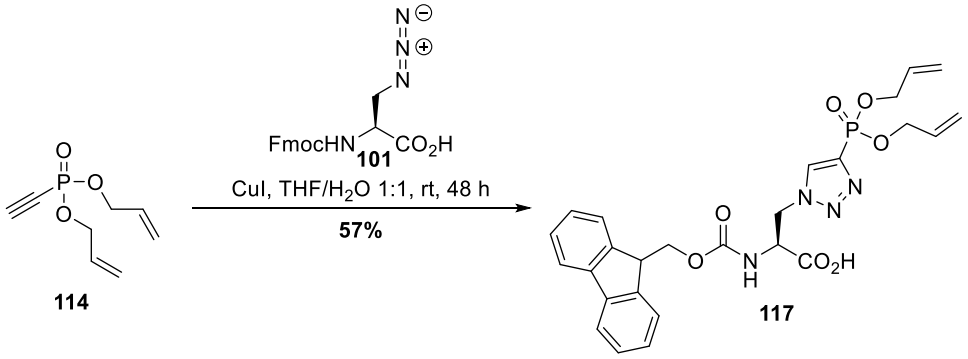
(2*S*)-3-(4-Diallylphosphonyl-[1,2,3]-triazol-1-yl)-2-(tertbutoxycarbonyl)amino-propionate 116 (*Boc-pTz(OAllyl)$_2$-OH*)



The reaction procedure to generate **116** was similar to that use to generate *Fmoc-pTz(OBn)$_2$-OH* **15**, however diallyl ethnylphosphonate **114** (299 mg, 1.6 mmol, 1.2 eq.) was used instead of dibenzyl ethnylphosphonate **100** and Boc-AzaAla-OH **115** (purchased as a dicyclohexylammonium salt, 535 mg, 1.3 mmol, 1.0 eq.) was used instead of Fmoc-AzaAla-OH **101**. The reaction was stirred for 24 h and concentrated as far as possible. Dioxane was added to dissolve the precipitate and the solution was lyophilised. The crude product was purified by mass-directed HPLC (50 – 95% MeCN) and lyophilised to yield

Boc-pTz(OAllyl)$_2$-OH **116** (68 mg, 0.16 mmol, 12%) as a pale yellow solid. $\boldsymbol{R_F}$ (10% MeOH in CH_2Cl_2) 0.06; $\boldsymbol{\delta}_{\mathbf{H}}$ (500 MHz, MeOD): 8.46 (1 H, s, Tz-H_5), 6.06 – 5.93 (2 H, m, 2 × allyl-H_2), 5.46 – 5.36 (2 H, m, 2 × allyl-H_{3a}), 5.32 – 5.24 (2 H, m, 2 × allyl-H_{3b}), 5.08 – 4.99 (1 H, m, H_β), 4.82 (2 H, dd, $^2J_{H\text{-}H}$ 13.9, $^3J_{H\text{-}H}$ 8.4, $H_{\beta'}$), 4.74 – 4.59 (5 H, m, 2 × allyl-H_1 & H_α), 1.47 – 1.38 (9 H, m, $CO_2(CH_3)_2$); $\boldsymbol{\delta}_{\mathbf{C}}$ (75 MHz, MeOD): 171.9 (s, *C*OOH), 157.4 (s, NH*C*(O)O), 136.9 (d, $^1J_{C\text{-}P}$ 243.7, Tz-C_4), 133.7 – 133.2 (m, Tz-C_5 & 2 × allyl-C_2), 118.8 (s, allyl-C_3), 80.9 (s, O*C*$(CH_3)_2$), 68.7 (d, $^2J_{C\text{-}P}$ 5.5, 2 × allyl-C_1), 54.8 (s, C_α), 51.9 (s, C_β), 28.6 (s, OC($CH_3)_2$); $\boldsymbol{\delta}_{\mathbf{P}}$ (121 MHz, MeOD): 8.18 (qn, $^3J_{P\text{-}H}$ 8.5); ν_{max} (solid)/cm^{-1}: 3337 (OH), 1711 (CO) 1237 (PO); ***m/z*** (ES): Found: M*(+2Na-H)* 461.1174 $C_{16}H_{24}N_4O_7PNa_2$ requires 461.1173.

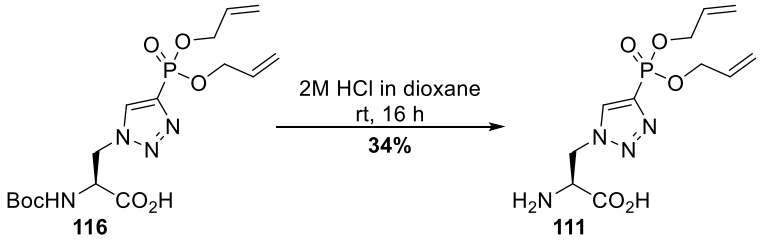
(2*S*)-3-(4-Diallylphosphonyl-[1,2,3]-triazol-1-yl)-2-(9H-fluoren-9-ylmethoxycarbonyl)-amino-propionate 117 (*Fmoc-pTz(OAllyl)$_2$-OH*)



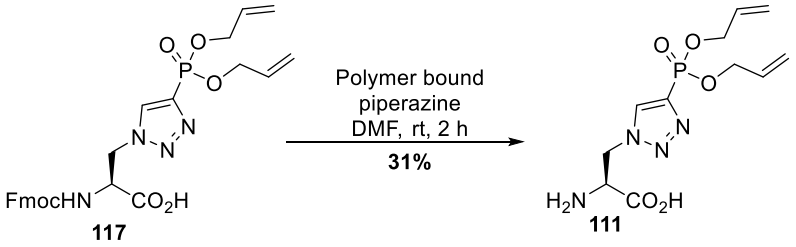
Fmoc-AzaAla-OH (235 mg, 0.61 mmol, 1 eq.) was dissolved in H_2O (10 mL) and added to a stirred solution of diallyl ethnylphosphonate **114** (150 mg, 0.80 mmol, 1.2 eq.) in THF (10 mL). Copper(I) iodide (29 mg, 0.15 mmol, 0.25 eq.) was dissolved in 1:1 THF/H_2O (5 mL), added to the reaction mixture and the white suspension was stirred for 48 h. The reaction was filtered at the pump, washed with 1:1 MeOH/DCM (2 × 30 mL), and the filtrate was concentrated *in vacuo*. The recovered solid was dissolved in H_2O:dioxane and lyophilised to leave a pale green flocculent solid which was purified by automated column chromatography (0 – 4% MeOH in CH_2Cl_2 (*v/v*)). *Fmoc-pTz(OAllyl)$_2$-OH* **117** (188 mg, 0.35 mmol, 57%) was recovered as an amorphous white solid. $\boldsymbol{R_F}$ (10% MeOH in CH_2Cl_2 (*v/v*)) 0.08; $\boldsymbol{\delta}_{\mathbf{H}}$ (500 MHz, *d*6-DMSO): 8.51 (1 H, s, Tz-H_5), 7.89 (2 H, d, $^3J_{H\text{-}H}$ 7.5 2 × Fmoc-H_4), 7.66 & 7.63 (2 H, 2 × d, $^3J_{H\text{-}H}$ 7.6 & 7.5, 2 × Fmoc-H_1), 7.41 (2 H, t, $^3J_{H\text{-}H}$ 7.4, 2 × Fmoc-H_3), 7.33 (2 H, t, $^3J_{H\text{-}H}$ 7.3, 2 × Fmoc-H_2), 6.95 (1 H, s, N*H*), 5.94 – 5.81 (2 H, m, 2 × allyl-H_2), 5.31 – 5.24 (2 H, m, 2 × allyl-H_{3a}), 5.17 (2 H, d, $^3J_{H\text{-}H}$ 10.5, 2 × allyl-H_{3b}), 4.98 (1 H, d, $^2J_{H\text{-}H}$ 11.7, H_β), 4.67 (1 H, dd, $^2J_{H\text{-}H}$ 12.0, $^3J_{H\text{-}H}$ 7.9, $H_{\beta'}$), 4.58 – 4.42 (4 H, m, 2 × allyl-H_1), 4.28 – 4.07 (4 H, m, Fmoc-C*H*, Fmoc-CH_2 & H_α); $\boldsymbol{\delta}_{\mathbf{C}}$ (125 MHz, *d*6-DMSO): 170.4 (s, *C*OOH), 155.52 (s, NH*C*(O)O), 143.8 (s, 2 × Fmoc-C_{1a}), 140.6 (s, 2 × Fmoc-C_{4a}), 135.3 (d, $^1J_{C\text{-}P}$ 238.8, Tz-C_4) 133.0 (d, $^3J_{C\text{-}P}$ 6.7, 2 × allyl-C_2), 131.9 (d, $^2J_{C\text{-}P}$ 35, Tz-C_5),

127.6 (s, 2 × Fmoc-C_3), 127.1 (s, 2 × Fmoc-C_2), 125.1 (s, 2 × Fmoc-C_1), 120.0 (s, 2 × Fmoc-C_4), 117.7 (s, 2 × allyl-C_3), 91.4 (s, 2 × Fmoc-C_{4a}), 66.4 (d, $^2J_{C\text{-}P}$ 5.1, 2 × allyl-C_1), 65.7 (s, Fmoc-CH_2), 56.2 (s, C_α), 51.7 (s, C_β), 46.6 (s, Fmoc-*C*H); **δ $_P$** (121 MHz, *d*6-DMSO): 8.26 (qn, $^3J_{P\text{-}H}$ 8.5); ν_{max} (solid)/cm^{-1}: 3334 (OH), 1702 (CO), 1232 (PO); ***m/z*** (ES): Found: M*(+H)* 539.1697 $C_{26}H_{28}N_4O_7P$ requires 539.1690; **HPLC** (5-95% A): retention time 3.07 min, 75%

(2*S*)-3-(4-Diallylphosphonyl-[1,2,3]-triazol-1-yl)-2-amino-propionate **111**
(*pTz(OAllyl)$_2$-OH*)



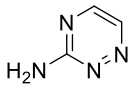
Boc-pTz(OAllyl)$_2$-OH **116** (24 mg, 0.06 mmol) was stirred in 2M HCl in dioxane (2 mL) for 16 h under N_2, the volatiles were removed *in vacuo* and the resultant white solid was purified by mass-directed HPLC (5 – 95% MeCN in H_2O) to yield pTz(OAllyl)$_2$-OH **111** (6.1 mg, 0.02 mmol, 34%) as a viscous orange oil.



pTz(OAllyl)$_2$-OH **111** was also synthesised from *Fmoc-pTz(OAllyl)$_2$-OH* **117**: *Fmoc-pTz(OAllyl)$_2$-OH* **117** (160 mg, 0.30 mmol, 1 eq.) was added to a solution of piperazine (polymer bound, 200-400 mesh, extent of labelling 1.0 – 2.0 mmol/g 2% cross-linked with divinylbenzene, 1.18 g, 1.18 mmol, 4 eq.) in DMF (10 mL) and the reaction was stirred for 2 h. The reaction was vacuum filtered, the solid support washed with 1:1 MeOH:CH_2Cl_2 (4 × 10 mL) and the filtrate concentrated *in vacuo*. The resultant white solid was dissolved in CH_2Cl_2, extracted with H_2O (3 × 15 mL) and the combined aqueous layers were lyophilised. The crude product was purified by mass directed HPLC (5 – 95% MeCN in H_2O) to yield *pTz(OAllyl)$_2$-OH* **111** (29 mg, 0.09 mmol, 31%) as a brittle white solid; **δ $_H$** (500 MHz, *d*6-DMSO) 8.58 (1H, s, Tz-H_5), 5.94 (2 H, ddtd, $^3J_{H\text{-}H}$ 18.2, $^3J_{H\text{-}H}$ 7.8, $^3J_{H\text{-}H}$ 5.3, $^4J_{H\text{-}P}$ 2.5, 2 × allyl-H_2), 5.38 – 5.31 (2 H, m, 2 × allyl-H_{3a}), 5.25 – 5.19 (2 H, m, 2 × allyl-H_{3b}), 4.92 – 4.83 (1 H, m, H_β) 4.73 – 4.60 (1 H, m, H_β), 4.60 – 4.48 (5 H, m, 2 × allyl-H_1 & H_β). 3.74 (1 H, s, COO*H*); δ C (125 MHz, CDCl3): 175.0 (s, *C*OOH), 135.6 (d, $^1J_{C\text{-}P}$ 240.0,

Tz-C_4), 133.1 (d, $^3J_{C\text{-}P}$ 6.5, 2 × allyl-C_2), 132.4 (d, $^2J_{C\text{-}P}$ 34.1, Tz-C_5), 117.7 (s, 2 × allyl-C_3), 66.4 (d, $^2J_{C\text{-}P}$ 5.4, 2 × allyl-C_1), 54.1 (s, C_α), 52.7 (s, C_β); **δ $_P$** (121 MHz, *d*6-DMSO): 8.11 (qn, $^3J_{H\text{-}P}$ 8.3); **ν_{max}** (solid)/cm^{-1}: 3100 & 2982 (NH_2), 2740 (OH) 1593 (CO), 1225 (PO); ***m/z*** (ES): Found: M*(+H)* 317.1015 $C_{11}H_{18}N_4O_5P$ requires 317.1009.

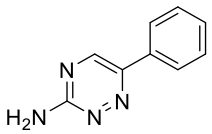
1,2,4-Triazin-3-amine 118a[142]



118a

Aminoguanidine bicarbonate (3.00 g, 22 mmol, 1 eq.) was stirred in H_2O (40 ml), the pH was dropped to 3-4 on addition of small portions of 18% HCl (~5 ml) until CO_2 gas evolution seized and the solution became clear. Glyoxal (40% in H_2O) (2.5 ml, 22 mmol, 1 eq.) was added and the pH increased to 12 with addition of 50% KOH $_{aq}$ ((*w/v*) 5 mL). The mixture was stirred for 2 hours and lyophilised. The resultant brown powder was triturated in 1:3 hexanes/EtOAc and hot filtered. The solution was concentrated and the crude product was recrystallised from 1:3 hexanes:EtOAc to give *1,2,4-triazin-3-amine* **118a**[142] (365 mg, 3.91 mmol, 17%) as pale orange needles. **m.p.** 179.2-179.9 ºC; ***R*$_F$** (5% MeOH in CH_2Cl_2) 0.17; **δ $_H$** (300 MHz, $CDCl_3$) 8.50 (1H, d, $^3J_{H\text{-}H}$ 2.4, H_6), 8.20 (1H, d, $^3J_{H\text{-}H}$ 2.4, H_5), 7.17 (2H, s, NH_2); **δ $_C$** (125 MHz, $CDCl_3$) 163.3 (C_3), 149.88 (C_5), 140.63 (C_6); **ν_{max}** (solid)/cm^{-1}: 3016 & 2970 (NH stretch); ***m/z*** (EI) Found: M(+*H*) 93.0 $C_3H_5N_4$ requires 93.0436; **HPLC** (5-95% A): retention time 0.43 min, 100%.

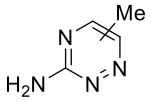
6-Phenyl-3-amino-1,2,4-triazine 118b[144]



118b

The reaction procedure was similar used to synthesise *1,2,4-triazin-3-amine* **118a** however phenylglyoxal monohydrate (8.5 g, 56 mmol, 1 eq.) was used instead of glyoxal. The crude product was recrystallized twice from boiling isoproponal to give *6-phenyl-1,2,4-triazin-3-amine* **118b**[144] (3.96 g, 23 mmol, 41%) as a flocculent orange solid; ***R*$_F$** (5% MeOH in CH_2Cl_2) 0.10; **δ $_H$** (500 MHz, $CDCl_3$): 9.34 (1H, s, H_5), 8.19 (2H, dd, $^3J_{H\text{-}H}$ 8.2, $^4J_{H\text{-}H}$ 1.4, Ph-H_{2+6}), 7.63 – 7.56 (3H, m, Ph-$H_{3\text{-}5}$), 7.25 (2H, s, NH_2); **δ $_C$** (125 MHz, $CDCl_3$): 163.8 (C_3), 150.4 (C_5), 135.7 (C_6), 130.3 (Ph-C_1), 126.7 (Ph-$C_{2\text{-}6}$); **ν_{max}** (solid)/cm^{-1}: 3301 (NH stretch); *m/z* (ES) Found: M*(+H)* 173.0822 $C_9H_9N_4$ requires 173.1866; **HPLC**: retention time 1.20 min, 90%.

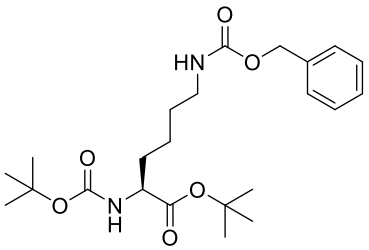
6-Methyl-1,2,4-triazin-3-amine & 5-methyl-1,2,4-triazin-3-amine 118c[144]



118c

The reaction procedure was the same as used to synthesise *3-amino-1,2,4-triazine* **118a**, however methyl glyoxal (40% in H_2O, 10 mL, 65 mmol, 1 eq.) was used instead of glyoxal. The crude product was recrystallized twice from boiling isopropanol to give a mixture of *6-methyl-1,2,4-triazin-3-amine & 5-methyl-1,2,4-triazin-3-amine* (2:1) **118c**[144] (1.28 g, 12 mmol, 18%) as a yellow solid; **δ $_H$** (300 MHz, MeOD): 8.46 (1H, s, 5-methyl-H_6), 8.23 (1H, s, 6-methyl-H_5), 5.46 (2H, s, NH_2), 2.48 (3H, s, 6-methyl-CH_3), 2.38 (3H, s, 5-methyl-CH_3).

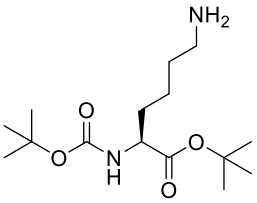
***N*$_ε$-(Carboxybenzyl)-*N*$_α$-(1,1-dimethylethoxy-carbonyl)-L-lysine-tert-butyl ester**[153] (*Boc-Lys(Z)-OtBu*)



Boc-Lys(Z)-OH (1.86 g, 4.88 mmol, 1 eq.), 4-dimethylaminopyridine (358 mg, 2.93 mmol, 0.6 eq.) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (1.87 g, 9.76 mmol, 2 eq.) were dissolved in anhydrous CH_2Cl_2 (10 mL) before addition of anhydrous *tert*-butanol (4.8 mL, 48.8 mmol, 10 eq.) and the mixture was stirred for 16 h. The reaction mixture was diluted with CH_2Cl_2 (~30 mL), washed with H_2O (2 × 30 mL) and brine (2 × 20 mL); dried ($MgSO_4$) and concentrated *in vacuo*. The crude product was purified by column chromatography (2:1 hexanes:EtOAc) to give *Boc-Lys(Z)-OtBu*[153] (1.44 g, 3.19 mmol, 68%) as a colourless oil; R_F 0.27 (2:1 hexanes:EtOAc); **[α]$_D$**^{21} -0.1 (*c* 2.2, ($CHCl_3$); **δ $_H$** (300 MHz, $CDCl_3$); 7.40 – 7.33 (5H, m, Bn-H_{2-6}); 5.15 (2H, s, Bn-CH_2); 4.81 (1H, br s, N*H*C(O)OCH(CH$_3$)$_3$); 4.21 – 4.06 (1H, m, $H_α$); 3.21 (2H, dd, $^3J_{H-H}$ 12.6 & $^3J_{H-H}$ 6.4, $H_ε$); 1.60 – 1.51 (4H, m, $H_β$ & $H_δ$); 1.47 (10H, s, NHC(O)OCH(CH_3)$_3$ & NHC(O)OC*H*(CH$_3$)$_3$); 1.45 (10H, s, C(O)OCH(CH_3)$_3$ & C(O)OC*H*(CH$_3$)$_3$); 1.41 – 1.33 (2H, m, $H_γ$); **δ $_C$** (125 MHz, $CDCl_3$); 171.9 (Bn-O*C*(O)); 156.4 & 155.5 (*C*(O)OCH(CH$_3$)$_3$ & NH*C*(O)OCH(CH$_3$)$_3$); 136.6 (Bn-C_1); 128.5, 128.1 & 128.1 (Bn-C_{2-6}); 81.9 & 79.7 (NHC(O)O*C*H(CH$_3$)$_3$ & C(O)O*C*H(CH$_3$)$_3$); 66.6 (Bn-*C*H$_2$); 53.7 ($C_α$); 40.8 ($C_ε$); 32.7 ($C_β$); 29.4 ($C_δ$); 28.3, 28.3 & 28.0 (NHC(O)OCH(*C*H$_3$)$_3$ & C(O)OCH(*C*H$_3$)$_3$); 22.3 ($C_γ$); **ν$_{max}$** (solid)/cm^{-1}: 3334 (NH); 1695 (CO); ***m/z*** (ES) Found: M*(+Na)*, 459.2468 $C_{23}H_{36}N_2O_6$ requires 459.2466. **HPLC** (5-95% A): retention time 3.86 min, 87%.

N_α-(1,1-Dimethylethoxy-carbonyl)-L-lysine-*tert*-butyl ester 126[153]

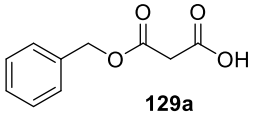
(*Boc-Lys-OtBu*)



126

Boc-Lys(Z)-O^{t}Bu (1.44 g, 3.3 mmol) and palladium on activated carbon (10%, 144 mg) were suspended in MeOH:CH_2Cl_2 (1:1) (10 mL) and stirred under H_2 for 6 h. The catalyst was removed by vacuum filtration through celite and the filtrate concentrated to give the *Boc-Lys-OH* **126**[153] (985 mg, 3.26 mmol, 99%) as a yellow oil. $\boldsymbol{R_F}$ 0.04 (10% MeOH in CH_2Cl_2); $\boldsymbol{[\alpha]_D^{21}}$ +1.0 (*c* 3.4, $CHCl_3$); $\boldsymbol{\delta_H}$ (300 MHz, $CDCl_3$); 7.38 (2H, br s, NH_2); 5.16 (1H, d, $^3J_{H\text{-}H}$ 8.8, N*H*); 4.20 – 4.09 (1H, m, H_α); 3.03 (2H, ap s, H_ε); 1.90 – 1.74 (2H, m, H_β); 1.67 – 1.56 (2H, m, H_δ); 1.48 (10H, s, NHC(O)OCH(CH_3)$_3$ & NHC(O)OCH(CH_3)$_3$); 1.46 (10H, s, C(O)OCH(CH_3)$_3$ & C(O)OCH(CH_3)$_3$); 1.27 (2H, dd, $^3J_{H\text{-}H}$ 9.9 & $^3J_{H\text{-}H}$ 4.6, H_γ); $\boldsymbol{\delta_C}$ (125 MHz, DMSO-*d6*); 156.0 (*C*(O)OCH(CH$_3$)$_3$); 141.8 (NH*C*(O)OCH(CH$_3$)$_3$); 81.9 & 79.7 (NHC(O)O*C*H(CH$_3$)$_3$ & C(O)O*C*H(CH$_3$)$_3$); 53.8 (C_α); 40.1 (C_ε); 32.4 (C_β & C_δ); 28.3 & 28.0 (NHC(O)OCH(*C*H$_3$)$_3$ & C(O)OCH(*C*H$_3$)$_3$); 22.4 (C_γ); $\boldsymbol{\nu_{max}}$ (solid)/cm^{-1}: 3354 (NH); 2977 (NH$_2$); 1700 (CO); ***m/z*** (ES) Found: M*(+H)* 303.2278 $C_{15}H_{31}N_2O_4$ requires 302.2274; **HPLC** (5-95% A): retention time 4.07 min, 91%.

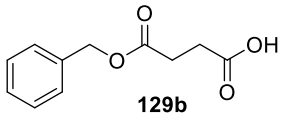
3-(Benzyloxy)-3-oxopropanoic acid 129a[154]



129a

A mixture of benzyl alcohol (3.00 g, 20.8 mmol, 1 eq.) and 2,2-dimethyl-1,3-dioxane-4,6-dione (Meldrum's acid) (2.15 ml, 20.8 mmol, 1 eq.) was heated to 120 ºC and stirred for 16 h. The reaction was quenched with a solution of Na_2CO_3 (5% (*w/v*) 5 mL), acidified with to pH 3-4 with 1M HCl $_{aq}$, extracted with EtOAc (3 × 10 mL), washed with brine (2 × 10 mL), dried ($MgSO_4$) and concentrated. The crude product was purified by column chromatography (2:1 hexanes:EtOAc) to give the *3-(Benzyloxy)-3-oxopropanoic acid* **129a**[154] (566 mg, 2.9 mmol, 14%) as an amorphous white solid; $\boldsymbol{R_F}$ 0.16 (2:1 hexanes:EtOAc); ; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$); 7.43 – 7.36 (5H, m, Bn-$H_{2\text{-}6}$); 5.25 (2H, s, Bn-CH_2); 3.52 (2H, s, C(O)CH_2C(O)); $\boldsymbol{\delta_C}$ (125 MHz, $CDCl_3$); 170.0 (*C*OOH 167.1 (O*C*(O)CH$_2$); 134.8 (Bz-C_1); 128.7, 128.7 & 128.5 (Bn-$C_{2\text{-}6}$); $\boldsymbol{\nu_{max}}$ (solid)/cm^{-1}: 2980 (OH); 1700 (CO); ***m/z*** (ES) Found: M*(+Na)* 217.0475. $C_{10}H_{10}O_4Na$ requires 217.0471. **HPLC** (5-95% A): retention time 3.00 min, 100%.

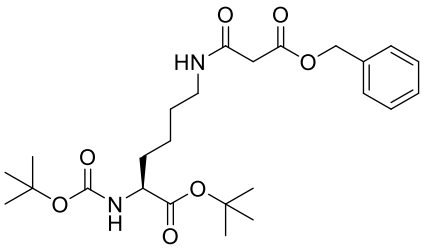
3-((Benzyloxy)carbonyl)propanoic acid 129b[155]



129b

A mixture of succinic anhydride (2.00 g, 20 mmol, 1 eq.), benzyl alcohol (2.1 mL, 20 mmol, 1 eq.), 4-dimethylaminopyridine (2.48 g, 20 mmol, 1 eq.) and triethylamine (3 mL, 1.1 eq., 22 mmol) in CH_2Cl_2 (30 mL) was stirred at room temperature for 16 h.[154] The reaction was quenched with Na_2CO_3 ((5% *w/v*) 30 mL) and the aqueous layer was acidified to pH 3-4 with 1M HCl_{aq}, extracted with EtOAc (3 × 30 mL). The combined organics were washed with brine (2 × 20 ml), dried ($MgSO_4$) and the volatiles evaporated *in vacuo* to give *3-((Benzyloxy)carbonyl)propanoic acid* **129b**[155] (2.52 g, 12.1 mmol, 61%) as an amorphous white solid; $\boldsymbol{R_F}$ (2:1 hexanes:EtOAc) 0.21; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$): 7.39 – 7.26 (5H, m, Bn-$H_{2\text{-}6}$), 5.16 (2H, s, Bn-CH_2), 2.84 – 2.56 (4H, m, C(O)CH_2CH_2C(O)NH); $\boldsymbol{\delta_C}$ (125 MHz, $CDCl_3$): 178.0 (*C*OOH), 172.0 (O*C*(O)CH_2), 135.7 (Ar-C_1), 128.6, 128.3 & 128.2 (Bn-$C_{2\text{-}6}$), 66.7 (Bn-*C*H_2), 28.9 (O*C*H_2*C*H_2O); $\boldsymbol{\nu_{max}}$ (solid)/cm^{-1}: 3181 (OH); 1736 (CO); ***m/z*** (ES): Found *M(+Na)* 231.0628. $C_{11}H_{12}O_4Na$ requires 231.0555); **elemental analysis** (Found: C, 62.5; H, 5.8; O, 30.7; $C_{11}H_{12}O_4$ requires C, 63.4; H, 5.8; O, 30.7); **HPLC** (5-95% A): retention time 2.14 min, 100%.

N_ε-(Carbamoylacetylbenzyl)-N_α-(1,1-dimethylethoxy-carbonyl)-L-lysine-*tert*-butyl ester 130a

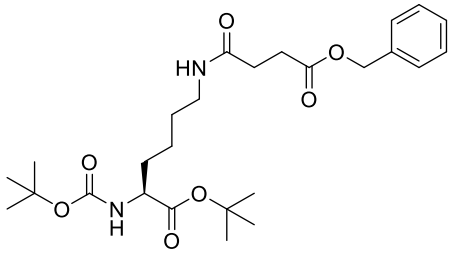


130a

3-(Benzyloxy)-3-oxopropanoic acid **N** (269 mg, 1.37 mmol, 1 eq.), *Boc-Lys-OtBu* **N** (414 mg, 1.37 mmol, 1 eq.), *N,N'*-Dicyclohexylcarbodiimide (0.2 mL, 1.37 mmol, 1 eq.) and 4-dimethylaminopyridine (167 mg, 1.37 mmol, 1 eq.) were suspended in dry CH_2Cl_2 (5 mL) and stirred for 16 h. The resultant mixture was washed with sat. NH_4HCl_{aq} (2 × 25 mL) and sat. $NaHCO_{3\ aq}$ (2 × 20 mL); and the combined organics were dried ($MgSO_4$) and concentrated. The crude product was purified by column chromatography (5% MeOH in CH_2Cl_2) to give *N_ε-(carbamoylacetylbenzyl)-N_α-(1,1-dimethylethoxy-carbonyl)-L-lysine-tert-butyl ester* **130a** (350 mg, 0.73 mmol, 53%) as a yellow oil; $\boldsymbol{R_F}$ 0.45 (5% MeOH in DCM); $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$); 7.40 – 7.33 (5H, m, Bn-$H_{2\text{-}6}$); 7.06 (1H, d, $^3J_{H\text{-}H}$ 5.4, malonyl-N*H*); 5.18 (2H, s, Bn-CH_2); 5.04 (1H, d, $^3J_{H\text{-}H}$ 6.3, Boc-N*H*); 4.18 – 4.13 (1H, m, H_α); 3.35 (2H, s,

C(O)CH_2C(O)); 3.27 (2H, dd, $^3J_{H\text{-}H}$ 13.0 & $^3J_{H\text{-}H}$ 7.0, H_ε); 1.87 – 1.72 (2H, m, H_β); 1.66 – 1.50 (2H, m, H_δ); 1.47 (10H, s, NHC(O)OCH(CH_3)$_3$ & NHC(O)OCH(CH$_3$)$_3$); 1.45 (10H, s, C(O)OCH(CH_3)$_3$ & C(O)OCH(CH$_3$)$_3$); 1.42 – 1.31 (2-H, m, H_γ); **δ_C** (125 MHz, $CDCl_3$); 171.9 (*C*(O)CH$_2$OC(O) & C(O)CH$_2$O*C*(O)); 169.4 (*C*(O)OCH(CH$_3$)$_3$); 155.4 (NH*C*(O)OCH(CH$_3$)$_3$); 135.0 (Bn-C_1); 128.7, 128.6 & 128.4 (Bn-$C_{2\text{-}6}$); 81.9 & 79.7 (NHC(O)O*C*H(CH$_3$)$_3$ & C(O)O*C*H(CH$_3$)$_3$); 67.3 (Bn-*C*H$_2$); 53.7 (C_α); 41.1 (C(O)*C*H$_2$C(O)); 39.4 (C_ε); 32.7 (C_β); 28.9 (C_δ); 28.3 & 28.0 (NHC(O)OCH(*C*H$_3$)$_3$ & C(O)OCH(*C*H$_3$)$_3$); 22.6 (C_γ); **ν_{max}** (solid)/cm^{-1}: 3322, 2977 & 2866 (NH); 1660 (CO); ***m/z*** (ES) Found: M*(+Na)* 501.2591, $C_{25}H_{38}N_2O_7Na$ requires 501.2571; **HPLC** (5-95% A): retention time 3.51 min, 85%.

N_ε-(Benzylcarbamoylpropanoyl)-N_α-(1,1-dimethylethoxy-carbonyl)-L-lysine-*tert*-butyl ester 130b

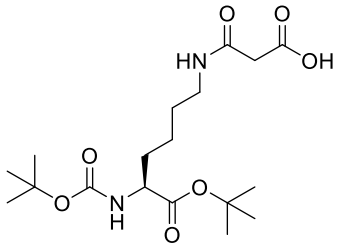


130b

The reaction procedure was similar to that used to synthesise *N_ε-(carbamoylacetylbenzyl)-N_α-(1,1-dimethylethoxy-carbonyl)-L-lysine*-tert-*butyl ester* **130a** apart from *3-((Benzyloxy)carbonyl)propanoic acid* **129b** (354 mg, 1.7 mmol, 1 eq.) was used instead of *3-(Benzyloxy)-3-oxopropanoic acid* **129a**. The resultant mixture was washed with sat. $NH_4HCl_{\ aq}$ (2 × 25 mL), sat. $NaHCO_{3\ aq}$ (2 × 20 mL), and the combined organics were dried ($MgSO_4$) and concentrated to leave the *title compound* **130b** (471 mg, 0.96 mmol, 56%) as a viscous orange oil which did not require further purification; ***R*$_F$** 0.56 (5% MeOH in CH_2Cl_2); **$[\alpha]_D^{21}$** +1.8 (*c* 2.2 ($CHCl_3$); **δ_H** (500 MHz, $CDCl_3$); 7.39 – 7.31 (5H, m, Bn-$H_{2\text{-}6}$); 5.74 (1H, br s, succinyl-N*H*); 5.13 (2H, s, Bn-CH_2); 5.12 (1H, s, Boc-N*H*); 4.18 – 4.11 (1H, m, H_α); 3.26 – 3.19 (2H, m, H_ε); 2.73 (2H, t, $^3J_{H\text{-}H}$ 6.9, C(O)CH$_2$CH_2C(O)NH); 2.47 (2H, t, $^3J_{H\text{-}H}$ 6.9, OCH_2CH$_2$ONH); 1.66 – 1.48 (4H, m, H_β & H_δ); 1.46 (10H, s, NHC(O)OCH(CH_3)$_3$ & NHC(O)OCH(CH$_3$)$_3$); 1.44 (10H, s, C(O)OCH(CH_3)$_3$ & C(O)OCH(CH$_3$)$_3$); 1.36 (2H, dd, $^3J_{H\text{-}H}$ 14.4 $^3J_{H\text{-}H}$ 8.4, H_γ); **δ_C** (125 MHz, $CDCl_3$); 172.8 & 171.3 (*C*(O)CH$_2$CH$_2$OC(O) & C(O)CH$_2$CH$_2$O*C*(O)); 160.1 (*C*(O)OCH(CH$_3$)$_3$; 150.5 NH*C*(O)OCH(CH$_3$)$_3$); 135.8 (Bz-C_1); 128.6, 128.3 & 128.2 (Bz-$C_{2\text{-}6}$); 81.9 (NHC(O)O*C*H(CH$_3$)$_3$ & C(O)O*C*H(CH$_3$)$_3$); 66.5 (Bn-*C*H$_2$); 53.7 (C_α); 39.3 (C_ε); 32.7 (C_β); 31.1 (C(O)CH$_2$*C*H$_2$OC(O) & C(O)*C*H$_2$CH$_2$OC(O)); 29.7 (C_β); 29.0 (C_δ); 28.4 & 28.0 (NHC(O)OCH(*C*H$_3$)$_3$ & C(O)OCH(*C*H$_3$)$_3$); 22.5 (C_γ); **ν_{max}** (solid)/cm^{-1}: 3332, 2977 & 2851 (NH); 1713 & 1660 (CO); ***m/z*** (ES) Found: M*(+Na)*

515.2739, $C_{26}H_{40}N_2O_7Na$ requires 5145.2728; **HPLC** (5-95% A): retention time 3.56 min, 81%.

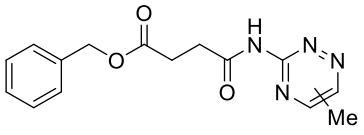
N_ε-(Carbamoylacetate)-N_α-(1,1-dimethylethoxy-carbonyl)-L-lysine-*tert*-butyl ester 123a



123a

N_ε-(carbamoylacetylbenzyl)-N_α-(1,1-dimethylethoxy-carbonyl)-L-lysine-tert-*butyl ester* **130a** (282 mg, 0.59 mmol) was dissolved in 1:1 MeOH:CH_2Cl_2; palladium on activated carbon (10%, 28 mg) was added and the reaction was stirred at room temperature under a nitrogen atmosphere for 48 h. The reaction mixture was filtered through a celite pad and the filtrate was concentrated *in vacuo* to leave *N_ε-(carbamoylacetate)-N_α-(1,1-dimethylethoxy-carbonyl)-L-lysine*-tert-*butyl ester* **123a** (215 mg, 0.55 mmol, 94%) as a crude brittle white foam. **δ_H** (500 MHz, $CDCl_3$); 6.86 (1 H, br s, succinyl-N*H*), 5.18 (1 H, d, $^3J_{H\text{-}H}$ 5.0, Boc-N*H*), 4.14 (1 H, app. S, H_α), 3.50 – 3.21 (4 H, m, H_ε & C(O)CH_2C(O)), 1.84 – 1.67 (2 H, m, H_β), 1.62 (2 H, s, H_δ), 1.47 (10 H, s, NHC(O)OCH(CH_3)$_3$ & NHC(O)OCH(CH_3)$_3$), 1.45 (10 H, s, C(O)OCH(CH_3)$_3$ & C(O)OCH(CH_3)$_3$), 1.27 (2 H, d, $^3J_{H\text{-}H}$ 10.0, H_γ); ***m/z*** (ES) Found: M*(+Na)* 411.2109, $C_{18}H_{32}N_2O_7Na$ requires 411.2102; **HPLC** (5-95% A): retention time 3.05 min, 82%.

Benzyl 3-(6-methyl-1,2,4-triazin-3-ylcarbamoyl)propanoate and benzyl 3-(5-methyl-1,2,4-triazin-3-ylcarbamoyl)propanoate 131a

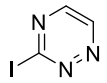


131a

3-((benzyloxy)carbonyl)propanoic acid **129b** (297 mg, 2.7 mmol, 1 eq.), *6-methyl-1,2,4-triazin-3-amine & 5-methyl-1,2,4-triazin-3-amine* (2:1 mixture) **118c** (561 mg, 2.7 mmol, 1 eq.), 4-dimethylaminopyridine (660 mg, 5.4 mmol, 2 eq.) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (1.03 g, 5.4 mmol, 2 eq.) were stirred at 40 °C under a nitrogen atmosphere in DMF (10 mL) for 32 hours. The reaction was quenched with H_2O (20 mL) and the aqueous layer extracted with EtOAc (4 × 30 mL). The combined organics were washed with H_2O (1 × 30 mL) and brine (2 × 40 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The crude yellow oil was purified by column chromatography on silica gel (2:1 hexanes:EtOAc) to give a mixture of *benzyl 3-(6-methyl-1,2,4-triazin-3-*

ylcarbamoyl)propanoate and *benzyl 3-(5-methyl-1,2,4-triazin-3-ylcarbamoyl)propanoate* (1:1) **131a** (45 mg, 0.15 mmol, 7%) as an orange solid. $\boldsymbol{R_F}$ (2:1 hexanes:EtOAc) 0.27; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$): 8.87 (1 H, s, 5-methyl- H_6), 8.36 (1 H, s, 6-methyl-H_5), 7.38 – 7.35 (5 H, m, Bn-$H_{2\text{-}6}$), 3.22 – 3.18 (2 H, m, 5-methyl-C(O)CH_2CH$_2$OH), 3.12 (2 H, t, $^3J_{H\text{-}H}$ 6.5, 6-methyl- C(O)CH_2CH$_2$OH), 2.86 – 2.81 (2 H, m, C(O)CH$_2$CH_2OH), 2.68 (3H, s, 6-methyl-CH_3), 2.54 (3 H, s, 5-methyl-CH_3); $\boldsymbol{\nu_{max}}$ (solid)/cm^{-1}: 3430 (NH stretch), 1709 (CO); ***m/z*** (ES): Found *M(+H)* 301.1294, $C_{15}H_{17}N_4O_3$ requires 301.1295; **HPLC** (5-95% A): retention time 1.88 min, 84%.

3-Iodo-1,2,4-triazine 136

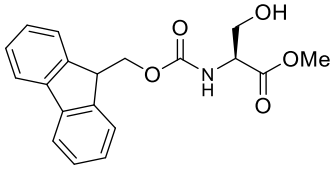


136

Isoamyl nitrite (42 ml, 300 mmol, 14 eq.) was added to a stirred solution of 3-amino-1,2,4-triazine (2 g, 20 mmol, 1 eq.) in diiodomethane (~40 mL). The turbid orange mixture was stirred at 55 °C for 4 h, allowed to cool, filtered at the pump and concentrated as far as possible *in vacuo*. The remaining filtrate was applied to a silica column and purified by column chromatography eluting with 3:1 hexanes/EtOAc. The resultant orange solid was dissolved in 1,4-dioxane and lyophilised to yield *3-Iodo-1,2,4-triazine* **136** (1.25g, 6.04 mmol, 30%) as a flocculent orange solid. $\boldsymbol{R_F}$ (3:1 hexanes/EtOAc) 0.29; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$): 9.26 (1H, d, $^3J_{H\text{-}H}$ 2.1 Hz, H_6), 8.38 (1H, d, $^3J_{H\text{-}H}$ 2.2 Hz, H_5); δ_C (125 MHz, $CDCl_3$); 149.2 (C_5), 148.4 (C_6); ν_{max} (solid)/cm^{-1}: 3450 & 3417 (NH_2 stretch) ***m/z*** (ES): Found: *M(+H)* 207.9366, $C_3H_3IN_3$ requires 207.9293; **HPLC** (5-95% A): retention time 1.41 min, 100%.

(9H-Fluoren-9-yl)methyl (*S*)-1-(methoxycarbonyl)-2-hydroxyethylcarbamate 140

(*Fmoc-Ser-OMe*)

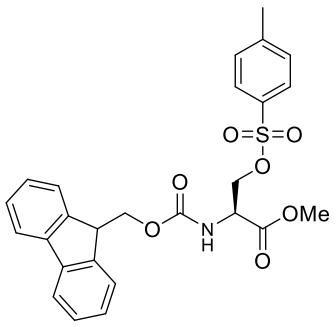


140

Sodium bicarbonate (6.7 g, 60 mmol, 2.5 eq.) was added to a vigorously stirred solution of L-serine methyl ester hydrochloride (5.0 g, 32 mmol, 1 eq.) in H_2O (60 ml) at 0 °C; the turbid solution immediately became clear. Fmoc *N*-hydroxysuccinimide ester (10.8 g, 32 mmol, 1 eq.), dissolved in H_2O (60 mL) was added dropwise and a colourless precipitate formed. The solution was allowed to warm to room temperature and stirred for 4 h, at which time TLC showed complete consumption of both reagents. The reaction was quenched with

H_2O (1 × 50 mL) and extracted with EtOAc (3 × 60 mL). The combined organics were washed with H_2O (2 × 50 mL) and brine (2 × 50 mL); dried ($MgSO_4$) and concentrated on the cold finger. The crude colourless solid was recrystalised from EtOAc:hexanes, cooled to -20 °C, and the resultant crystals collected by vacuum filtration, before being washed with ice-cold hexanes (2 × 50 mL) to yield *Fmoc-Ser-OMe* **140** (9.66 g, 28 mmol, 88%) as large, colourless crystals. **m.p.** 128.2 – 130.6 °C; $[\alpha]_D^{27}$ +4.35 (*c* = 0.12, $CHCl_3$); R_F (3:1 hexanes:EtOAc) 0.17; δ_H (500 MHz, $CDCl_3$): 7.70 (2H, d, $^3J_{H\text{-}H}$ 7.5, 2 × Fmoc-H_4), 7.57 – 7.50 (2H, m, 2 × Fmoc-H_1), 7.34 (2H, t, $^3J_{H\text{-}H}$ 7.5, 2 × Fmoc-H_3), 7.25 (2H, t, $^3J_{H\text{-}H}$ 7.5, 2 × Fmoc-H_2), 5.62 (1H, br s, N*H*), 4.45 – 4.27 (3H, m, 2 × Fmoc-CH_2 & H_α), 4.16 (1 H, t, $^3J_{H\text{-}H}$ 6.8, 2 × Fmoc-C*H*), 3.94 (1 H, d, $^2J_{H\text{-}H}$ 8.9, H_β), 3.86 (1 H, d, $^2J_{H\text{-}H}$ 9.1, H_β), 3.73 (3 H, s, OCH_3); δ_C (125 MHz, $CDCl_3$): 171.1 (*C*(O)OCH_3), 156.2 (O*C*(O)NH), 143.7, 141.4 & 141.3 (2 × Fmoc-$C_{4a/1a}$), 127.77 (2 × Fmoc-C_3), 127.1 (2 × Fmoc-C_2), 125.1 (2 × Fmoc-C_1), 120.0 (2 × Fmoc-C_4), 67.2 (2 × Fmoc-*C*H_2), 63.3 (C_β), 56.03 (C_α), 52.8 (O*C*H_3), 47.2 (Fmoc-*C*H); ν_{max} (solid)/cm^{-1}: 3470 & 2431 (NH stretch), 3314 (OH) 1749 (CO); ***m/z*** (ES): Found: *M(+H)* 342.1339, $C_{19}H_{20}NO_5$ requires 342.1336; **HPLC** (5-95% B): retention time 2.94 min, 100%.

(9H-Fluoren-9-yl)methyl (*S*)-1-(methoxycarbonyl)-2-(*p*-toluenesulfonate)ester 140 (*Fmoc-Ser(OTs)-OMe*)



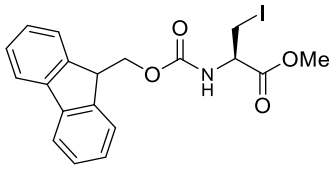
141

p-toluenesulfonyl chloride (16.4 g, 86 mmol, 2 eq.) was added to a solution of *Fmoc-Ser-OMe* **140** (14.7 g, 43 mmol, 1 eq.) in pyridine (40 mL) at 0 °C and the colourless solution immediately turned pale orange. The reaction mixture was allowed to warm to room temperature and stirred for 16 h. The resultant clear ruby solution was quenched with H_2O (1 × 40 mL) and extracted with EtOAc (3 × 60 mL). The combined organics were washed with 1M citric acid solution (3 × 50 mL), saturated $NaHCO_3$ (2 × 50 mL), brine (1 × 40 mL), 1M HCl solution (3 × 50 mL), sat. $NaHCO_3$ (2 × 50 mL) and brine (2 × 40 ml). The organic layer was dried ($MgSO_4$) and concentrated on the cold finger. The colourless solid was recrystallised from 1:1 EtOH:Petroleum ether and the crystals collected by vacuum filtration. The resultant pale yellow crystals were triturated in ice-cold petroleum ether for 30 minutes and collected to yield *Fmoc-Ser(OTs)-OMe* **141** (13.38 g, 27 mmol, 63%) as

colourless crystals. **m.p.** 119.7 – 120.2 °C; $[\alpha]_D^{27}$ +3.25 (*c* = 0.14, CH_2Cl_2); R_F (3:1 hexanes/EtOAc) 0.29; δ_H (500 MHz, $CDCl_3$): 7.74 – 7.64 (4 H, m, 2 × Fmoc-H_4 & OTs-H_2), 7.56 – 7.49 (2 H, m, 2 × Fmoc-H_1), 7.34 (2 H, t, $^3J_{H\text{-}H}$ 7.3, 2 × Fmoc-H_3), 7.29 – 7.23 (2 H, m, 2 × Fmoc-H_2), 7.21 (2 H, d, $^3J_{H\text{-}H}$ 8.2, OTs-H_3), 5.55 (1 H, d, $^3J_{H\text{-}H}$ 7.9, N*H*), 4.53 – 4.48 (1 H, m, H_α), 4.39 (1 H, d, $^2J_{H\text{-}H}$ 7.0, H_β), 4.35 (1H, d, $^2J_{H\text{-}H}$ 7.5, $H_{\beta'}$), 4.32 – 4.15 (2 H, m, Fmoc-CH_2), 4.12 (1 H, t, $^3J_{H\text{-}H}$ 7.2, Fmoc-C*H*), 3.67 (3 H, s, OCH_3), 2.29 (3 H, s, OTsCH_3). δ_C (125 MHz, $CDCl_3$): 168.7 (*C*(O)OCH_3) 155.6 (O*C*(O)NH), 145.2 (OTs-C_4) 143.6 (2 × Fmoc-C_{4a}) 141.3 (2 × Fmoc-C_{1a}) 132.3 (OTs-C_2), 130.0 (OTs-C_3), 128.0 & 127.8 (2 × Fmoc-C_3), 127.2 (2 × Fmoc-C_2), 125.2, & 125.1 (2 × Fmoc-C_1), 120.0 (2 × Fmoc-C_4), 69.0 (C_β), 67.5 (Fmoc-*C*H_2), 53.4 (C_α), 53.1 (O*C*H_3), 47.02 (Fmoc-*C*H), 21.58 (OTs-*C*H_3); ν_{max} (solid)/cm^{-1}: 3428 & 3338 (NH stretch), 1749 (CO); ***m/z*** (ES): Found: *M(+H)* 496.1428, $C_{26}H_{26}NO_7S$ requires 496.1424; **HPLC** (5-95% B): retention time 4.10 min, 61%.

(9H-Fluoren-9-yl)methyl (*R*)-1-(methoxycarbonyl)-2-iodoethylcarbamate 135

(*Fmoc-IodoAla-OMe*)



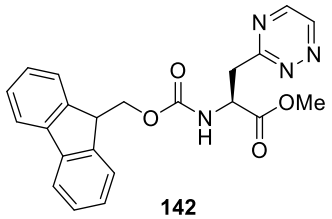
135

Sodium iodide (7.20 g, 48 mmol, 3 eq.) in acetone (80 ml) was added to a stirred solution of *Fmoc-Ser(OTs)-OMe* **141** (8.00 g, 16 mmol, 1 eq.) in acetone (160 mL). The colourless solution turned yellow. After stirring for 16 h at room temperature the mixture was filtered at the pump and concentrated *in vacuo* to obtain a yellow oil. The oil was dissolved in CH_2Cl_2 (40 mL) and washed with H_2O (3 × 50 mL), sat. sodium thiosulfate solution (2 × 50 mL) and brine (2 × 40 mL). The organic layer was dried ($MgSO_4$), concentrated and purified by column chromatography on silica gel (2:1 CH_2Cl_2/hexanes) to leave a colourless viscous oil. The oil was triturated in an ice-cold solution of 1:1 hexanes:EtOH and collected at the pump to yield *Fmoc-IodoAla-OMe* **135** (3.74 g, 8.2 mmol, 52%) as a flocculent, colourless solid. **m.p.** 148.8 – 149.4°C; $[\alpha]_D^{27}$ + 4.85 (*c* = 0.11, CH_2Cl_2); R_F (2:1 CH_2Cl_2:hexanes) 0.21; δ_H (500 MHz, $CDCl_3$): 7.78 (2 H, d, $^3J_{H\text{-}H}$ 7.5, 2 × Fmoc-H_4), 7.63 (2 H, d, $^3J_{H\text{-}H}$ 7.4, 2 × Fmoc-H_1), 7.42 (2 H, t, $^3J_{H\text{-}H}$ 7.5, 2 × Fmoc-H_3), 7.34 (2 H, td, $^3J_{H\text{-}H}$ 7.5, $^4J_{H\text{-}H}$ 3.0, 2 × Fmoc-H_2), 5.69 (1 H, d, $^3J_{H\text{-}H}$ 7.3, N*H*), 4.60 (1 H, dt, $^3J_{H\text{-}H}$ 7.6, $^3J_{H\text{-}H}$ 3.8, H_α), 4.48 – 4.36 (2 H, m, Fmoc-CH_2), 4.26 (1 H, t, $^3J_{H\text{-}H}$ 7.2, Fmoc-C*H*), 4.00 (1 H, dd, $^3J_{H\text{-}H}$ 11.4, $^3J_{H\text{-}H}$ 3.8, H_β); 3.91 (1 H, dd, $^3J_{H\text{-}H}$ 11.4, $^3J_{H\text{-}H}$ 3.8, $H_{\beta'}$); 3.84 (3 H, s, OCH_3); δ_C (125 MHz, $CDCl_3$); 169.7 (*C*(O)OCH_3), 155.4 (O*C*(O)NH), 143.8 & 143.7 (2 × Fmoc-C_{4a}), 141.3 (2 × Fmoc-C_{1a}), 127.8 (2 × Fmoc-C_3), 127.1 (2 × Fmoc-C_2), 125.2 & 125.1 (2 × Fmoc-C_1), 120.0 (2 × Fmoc-C_4), 67.4 (Fmoc-*C*H_2), 54.1 (C_α) 53.1 (O*C*H_3), 45.2 (Fmoc-*C*H), 7.4 (C_β);

ν_{max} (solid)/cm^{-1}: 3743 & 3338 (NH stretch), 1749 (CO); ***m/z*** (ES) Found: *M(+H)*, 452.0349, $C_{19}H_{19}INO_4$ requires 452.0353; **HPLC** (5-95% B): retention time 3.91 min, 72%.

(9H-Fluoren-9-yl)methyl (*S*)-1-(methoxycarbonyl)-2-(1,2,4-triazin-3-yl)ethylcarbamate 142

(*Fmoc-TrzAla-OMe*)

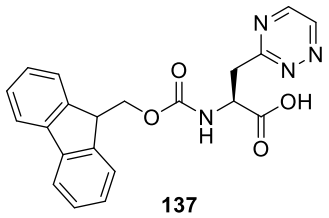


142

To an oven dried two-neck flask, zinc dust (1.16 g, 18 mmol, 3 eq.) was added; the flask was evacuated, dried with a flame and purged with nitrogen three times. The flask was allowed to cool to room temperature and freshly opened, dry DMF (18 mL) and I_2 (225 mg, 0.89 mmol, 0.15 eq.) were added in quick succession. The solution became orange, and after two min returned to grey. After 15 min, *Fmoc-IodoAla-OMe* **135** (2.67 g, 5.9 mmol, 1 eq.) was added, followed immediately by I_2 (225 mg, 0.89 mmol, 0.15 eq.) and the reaction mixture was stirred at room temperature. After two h, formation of the organozincate was shown to be complete by TLC (2:1 hexanes:EtOAc) and *3-iodo-1,2,4-triazine* **136** (1.59 g, 7.68 mmol, 1.3 eq.), palladium (II) acetate (33 mg, 0.15 mmol, 0.025 eq.) and 2-Dicyclohexylphosphino-2′,6′-dimethoxybiphenyl (SPhos, 121 mg, 0.30 mmol, 0.05 eq.) were added to the flask in quick succession. The flask was heated to 50 °C and stirred for five h, the reaction was allowed to cool and filtered through a celite pad, which was washed several times with CH_2Cl_2. The resultant solution was concentrated *in vacuo* and the pale orange solid was purified by column chromatography on silica gel, eluting initially with 4:1 hexanes:EtOAc and then EtOAc. The combined fractions were concentrated *in vacuo* and lyophilised to give *Fmoc-TrzAla-OMe* **142** (1.66 g, 4.12 mmol, 69%) as a flocculent, pale orange solid. $[\alpha]_D^{27}$ +4.12 (*c* = 0.19, CH_2Cl_2); R_F (EtOAc) 0.72; δ_H (500 MHz, $CDCl_3$): 9.36 (1 H, app s, Tz-H_6), 8.81 (1 H, d, $^3J_{H\text{-}H}$ 2.4, Tz-H_5), 7.79 – 7.73 (2 H, m, 2 × Fmoc-H_4), 7.58 (2 H, t, $^3J_{H\text{-}H}$ 7.4, 2 × Fmoc-H_1), 7.43 – 7.37 (2 H, m, 2 × Fmoc-H_3), 7.31 (2 H, t, $^3J_{H\text{-}H}$ 7.3, 2 × Fmoc-H_2), 5.98 (1 H, d, $^3J_{H\text{-}H}$ 8.5, N*H*), 5.07 – 4.96 (1 H, m, H_α), 4.45 – 4.35 (2 H, m, Fmoc-CH_2), 4.21 (1 H, t, $^3J_{H\text{-}H}$ 7.0, Fmoc-C*H*), 3.79 – 3.72 (5 H, m, H_β & OCH_3); δ_C (125 MHz, $CDCl_3$): 171.4 (*C*(O)OCH$_3$), 166.8 (Trz-C_3), 155.9 (O*C*(O)NH), 148.7 (Trz-C_6), 148.1 (Trz-C_5), 143.8, (2 × Fmoc-C_{4a}), 141.3, (2 × Fmoc-C_{1a}), 127.7 & 127.1 (2 × Fmoc-$C_{3/2}$), 125.1 (2 × Fmoc C_1), 120.0 (2 × Fmoc C_4), 67.2 (C_β), 67.1 (Fmoc-*C*H$_2$), 53.1 (C_α), 51.9 (O*C*H$_3$), 47.0 (Fmoc-*C*H); ν_{max} (solid)/cm^{-1}: 3049 & 2950 (NH stretch), 1715 (CO); ***m/z*** (ES) Found: *M(+H)* 405.1560 $C_{22}H_{21}N_4O_4$ requires 405.1557; **HPLC** (5-95% B): retention time 2.93 min, 100%.

(9H-Fluoren-9-yl)methyl (*S*)-1-(carboxy)-2-(1,2,4-triazin-3-yl)ethylcarbamate 137

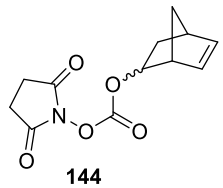
(*Fmoc-TrzAla-OH*)



137

Fmoc-TrzAla-OMe **142** (300 mg, 0.74 mmol, 1 eq.) and trimethyltin hydroxide (402 mg, 2.2 mmol, 3 eq.) were dissolved in anhydrous DCE (9 mL) and heated to reflux temperature for 2.5 h, when methyl deprotection was shown to be complete by TLC (EtOAc). The reaction was allowed to cool to room temperature and quenched with H_2O (15 mL). The organic layer was extracted with CH_2Cl_2 (3 × 20 mL) and the combined organics were washed with H_2O (1 × 20 mL) and brine (2 × 20 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The orange oil was purified by column chromatography on silica gel (95:4:1 CH_2Cl_2:MeOH:AcOH) and lyophilised to yield *Fmoc-TrzAla-OH* **137** (78 mg, 0.19 mmol, 27%) as a pale yellow, flocculent solid**.** $\mathbf{[\alpha]_D}^{27}$ +4.90 (*c* = 0.10, CH_2Cl_2); R_F (95:4:1 CH_2Cl_2:MeOH:AcOH) 0.26; $\boldsymbol{\delta}_{\mathbf{H}}$ (500 MHz, $CDCl_3$): 9.18 (1 H, app s, Tz-H_6), 8.63 (1 H, app s, Tz-H_5), 7.75 (2 H, d, $^3J_{H\text{-}H}$ 7.5, 2 × Fmoc-H_4), 7.57 (2 H, dd, $^3J_{H\text{-}H}$ 7.4, $^4J_{H\text{-}H}$ 3.1, 2 × Fmoc-H_1), 7.39 (2 H, t, $^3J_{H\text{-}H}$ 7.3, 2 × Fmoc-H_3), 7.30 (2 H, t, $^3J_{H\text{-}H}$ 7.4, 2 × Fmoc-H_2), 6.05 (1 H, d, $^3J_{H\text{-}H}$ 8.3. N*H*), 5.04 – 4.97 (1 H, m, H_α), 4.44 – 4.35 (2 H, m, Fmoc-CH_2), 4.22 (1 H, t, $^3J_{H\text{-}H}$ 7.1, Fmoc-C*H*), 3.79 (1 H, d, $^3J_{H\text{-}H}$ 5.5, H_β), 3.76 (1 H, d, $^3J_{H\text{-}H}$ 5.0, H_β); $\boldsymbol{\delta}_{\mathbf{C}}$ (125 MHz, $CDCl_3$); 174.3 (*C*OOH), 167.0 (Trz-C_3), 156.2 (O*C*(O)NH), 149.2 (C_6), 148.0 (C_5), 143.7 (2 × Fmoc-C_{4a}), 141.3 (2 × Fmoc-C_{1a}), 127.7 (2 × Fmoc C_3), 127.1 (2 × Fmoc C_2), 125.1 (2 × Fmoc C_1), 120.0 (2 × Fmoc C_4), 67.3 (C_β), 67.1 (Fmoc-*C*H_2), 52.2 (C_α), 39.0 (Fmoc-*C*H); $\boldsymbol{\nu_{max}}$ (solid)/cm^{-1}: 3379 (OH stretch), 1714 (CO); ***m/z*** (ES) Found: *M(+H)* 391.1401 $C_{21}H_{19}N_4O_4$ requires 391.1401; **HPLC** (5-95% A): retention time 2.87 min, 100%.

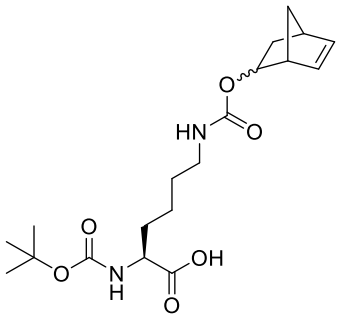
(1*R*, 2*R)-(norborn-5-ene-2-yloxy-carbonyl)-succinimide (*endo/exo* (2:1)) 144**



144

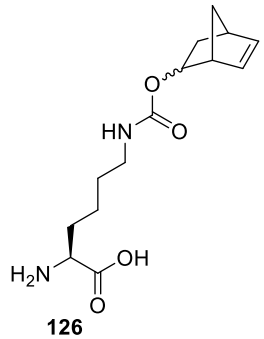
Disuccinimide carbonate (17.6 g, 68.8 mmol, 1.7 eq.) was added to a stirred solution of (1*R*, 2*R**)-5-norbornene-2-ol (*endo/exo* mixture, 4.46 g, 40.5 mmol, 1 eq.) and triethylamine (dried over 4Å MS and freshly distilled before use, 16.9 ml, 121.5 mmol, 3 eq.) in anhydrous MeCN (50 mL) at room temperature. After 16 h, the volatiles were evaporated and the crude product was purified by column chromatography (1% diethyl ether in CH_2Cl_2) to leave the title compound *(1R, 2R*)-(5-norbornene-2-yloxy-carbonyl)-succinimide* **144** (3.05g, 12.1 mmol 30%, 7:3 *endo*/*exo*) as a pale yellow solid; $\boldsymbol{R_F}$ (1% diethyl ether in CH_2Cl_2) 0.50; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$): 6.40 & 6.31 (1 H, dd, $^3J_{H\text{-}H}$ 5.7, $^3J_{H\text{-}H}$ 3.1, $H_{5\ endo}$ & dd, $^3J_{H\text{-}H}$ 5.7, $^3J_{H\text{-}H}$ 2.9, $H_{5\ exo}$), 6.04 & 5.96 (1 H, dd, $^3J_{H\text{-}H}$ 5.7, $^3J_{H\text{-}H}$ 2.9, $H_{6\ endo}$ & dd, $^3J_{H\text{-}H}$ 5.7, $^3J_{H\text{-}H}$ 3.3, $H_{6\ exo}$), 5.36 & 4.77 – 4.73 (1 H, ddd, $^3J_{H\text{-}H}$ 8.1, $^3J_{H\text{-}H}$ 3.9, $^3J_{H\text{-}H}$ 2.6, $H_{2\ endo}$ & m, $H_{2\ exo}$), 3.27 & 3.08 (1 H, d, $^3J_{H\text{-}H}$ 2.0, $H_{1\ endo}$ & d, $^3J_{H\text{-}H}$ 1.5, $H_{1\ exo}$), 2.91 (1 H, dd, $^3J_{H\text{-}H}$ 2.5, $^3J_{H\text{-}H}$ 1.8, H_4), 2.82 (4 H, s, 2 × succinimide-CH_2), 2.19 & 1.80 (1 H, ddd, $^2J_{H\text{-}H}$ 12.9, $^3J_{H\text{-}H}$ 8.1, $^3J_{H\text{-}H}$ 3.7, $H_{3a\ endo}$ & ddd, $^2J_{H\text{-}H}$ 13.0, $^3J_{H\text{-}H}$ 6.9, $^3J_{H\text{-}H}$ 2.7 $H_{3a\ exo}$), 1.73 (0.3 H, d, $^2J_{H\text{-}H}$ 8.9, $H_{7a\ exo}$), 1.69 – 1.65 (0.3 H, m, $H_{7b\ exo}$), 1.62 (0.3 H, ddd, $^2J_{H\text{-}H}$ 13.0, $^3J_{H\text{-}H}$ 3.5, $^3J_{H\text{-}H}$ 2.2, $H_{3b\ exo}$), 1.55 – 1.51 (0.7 H, m, $H_{7a\ endo}$), 1.34 (0.7 H, dd, $^2J_{H\text{-}H}$ 9.1, $^3J_{H\text{-}H}$ 0.4, $H_{7b\ endo}$), 1.15 (0.7 H, ddd, $^3J_{H\text{-}H}$ 13.0, $^3J_{H\text{-}H}$ 3.9, $^3J_{H\text{-}H}$ 2.6, $H_{3b\ endo}$). $\boldsymbol{\delta_C}$ (125 MHz, CDCl3): 168.69 (2 × succinimide-*C*O), 151.32 (O*C*(O)O), 142.13 ($C_{5\ exo}$), 139.14 ($C_{5\ endo}$), 131.74 ($C_{6\ exo}$), 130.95 ($C_{6\ endo}$), 83.30 ($C_{2\ exo}$), 82.84 ($C_{2\ endo}$), 47.61 ($C_{7\ endo}$), 47.29 ($C_{1\ exo}$), 46.28 ($C_{7\ exo}$), 45.78 ($C_{1\ endo}$), 42.21 (C_4), 34.47 ($C_{3\ exo}$), 34.35 ($C_{3\ endo}$), 25.48 (2 × succinimide-*C*H_2); ***m/z*** (ES): Found: 274.0697 *M(+Na)* $C_{12}H_{13}NO_5Na$ requires 274.0686; **HPLC** (5-95% A): retention times 3.12 min (*endo*), 3.19 min (*exo*), 99%.

(1*R*, 2*R)-N_ε-(norborn-5-ene-2-yloxy-carbonyl)-N_α-(*tert*butoxycarbonyl)-L-lysine (*endo/exo* 2:1)**



Boc-Lys-OH (1.16 g, 4.71 mmol, 1.3 eq.) and *(1R, 2R*)-(norborn-5-ene-2-yloxy-carbonyl)-succinimide (endo/exo (2:1))* **144** (914 mg, 3.64 mmol, 1 eq.) were stirred in anhydrous DMF (30 mL). After 16 h the reaction mixture was quenched with H_2O (30 mL) and extracted with EtOAc (3 × 30 mL). The combined organic layers were washed with brine (2 x 20 mL), dried ($MgSO_4$) and concentrated to give the *title compound* (1.37g, 3.58 mmol, 98%, 7:3 *endo*/*exo*) as a brittle, off-white foam; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$); 6.39 & 6.34 – 6.29 (1H, dd, $^3J_{H\text{-}H}$ 6.4, $^3J_{H\text{-}H}$ 3.1, $H_{5'\ exo}$ & m, $H_{5'\ endo}$); 6.03 & 5.96 (1H, dd, $^3J_{H\text{-}H}$ 5.8, $^3J_{H\text{-}H}$ 2.9, $H_{6'\ exo}$ & dd, $^3J_{H\text{-}H}$ 5.6, $^3J_{H\text{-}H}$ 3.0, $H_{6'\ endo}$); 5.36 & 5.27 – 5.22 (1H, ddd, $^3J_{H\text{-}H}$ 8.1, $^3J_{H\text{-}H}$ 3.9, $^3J_{H\text{-}H}$ 2.6, $H_{2'\ exo}$ & m, $H_{2'\ endo}$); 4.69 – 4.59 (1H, m, H_α); 4.44 – 4.24 (1H, br, N*H*); 3.31 – 3.24 & 3.20 – 3.04 (3H, m, $H_{1'\ exo}$ & m, H_ε & $H_{1'\ endo}$); 2.84 – 2.83 & 2.82 – 2.81 (1H, m, $H_{4'\ exo}$ & m, $H_{4'\ endo}$); 2.19 & 2.15 – 2.08 (1H, ddd, $^2J_{H\text{-}H}$ 12.9, $^3J_{H\text{-}H}$ 8.1 & $^3J_{H\text{-}H}$ 3.7, $H_{3a'\ exo}$ & m, $H_{3a'\ endo}$); 1.93 – 1.74 (2H, m, H_β); 1.74 –1.56 (2H, m, H_δ); 1.53 – 1.36 (11H, m, 7'-*H* & OC(CH_3)$_2$); 1.35 – 1.29 (2H, m, H_γ); 1.14 & 0.97 – 0.92 (1H, ddd, $^2J_{H\text{-}H}$ 13.0, $^3J_{H\text{-}H}$ 4.0, $^3J_{H\text{-}H}$ 2.6, $H_{3b'\ exo}$ & m, $H_{3b'\ endo}$); ***m/z*** (ES): Found: 405.2001 *M(+Na)* $C_{19}H_{30}N_2O_6Na$ requires 405.1996.

(1*R*, 2*R)-N_ε-(norborn-5-ene-2-yloxy-carbonyl)-L-lysine (*endo/exo*) 126**

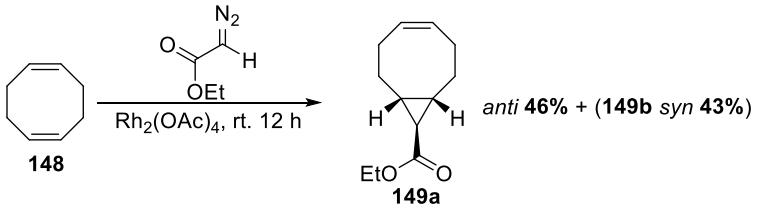


126

4M HCl in dioxane (anhydrous, 10 mL) was added to *(1R, 2R*)-N_ε-(norborn-5-ene-2-yloxy-carbonyl)-N_α-(*tert*butoxycarbonyl)-L-lysine (endo/exo 2:1)* (1.68 g, 4.39 mmol, 1 eq.) and the mixture was stirred at rt under nitrogen for 30 min. The volatiles were evaporated and the resultant cream powder was triturated in ice cold ether for 30 min. A pale yellow solid

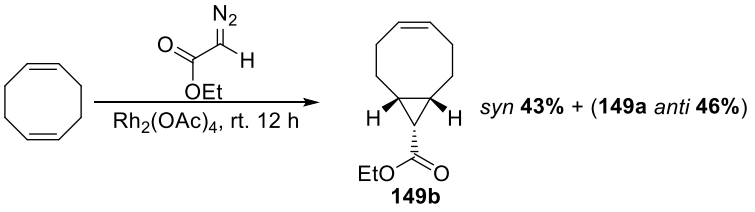
was collected by gravity filtration and lyophilised (H_2O) to give *(1R, 2R*)-N_ε-(norborn-5-ene-2-yloxy-carbonyl)-L-lysine* **126** (1.24 g, 4.39 mmol, 100%, 7:3 *endo*/*exo*) as an off-white flocculent solid; **δ $_H$** (500 MHz, MeOD); 6.35 & 6.31 (1H, dd, $^3J_{H\text{-}H}$ 5.2, $^3J_{H\text{-}H}$ 2.9, $H_{5'\ endo}$ & dd, $^3J_{H\text{-}H}$ 5.7, $^3J_{H\text{-}H}$ 2.6, $H_{5'\ exo}$); 6.04 & 6.02 – 5.99 (1H, dd, $^3J_{H\text{-}H}$ 5.7, $^3J_{H\text{-}H}$ 3.2, $H_{6'\ exo}$ & m, $H_{6'\ endo}$); 5.25 – 5.19 & 4.60 – 4.55 (1H, m, $H_{2'\ endo}$ & m, $H_{2'\ exo}$); 4.02 – 3.97 (1H, m, H_α); 3.20 – 3.11 & 2.93 – 2.82 (4H, m, H_ε + $H_{1'\ endo}$ & m, $H_{1'\ exo}$, $H_{4'\ exo}$ & $H_{4'\ endo}$); 2.15 (0.7H, ddd, $^2J_{H\text{-}H}$ 12.0, $^3J_{H\text{-}H}$ 8.2, $^3J_{H\text{-}H}$ 3.6, $H_{3a'\ endo}$); 2.07 – 1.88 (2H, m, H_β); 1.76 – 1.67 (0.6H, m, $H_{3'\ exo}$); 1.56 – 1.42 (5.3H, m, H_δ, H_γ, $H_{7a'\ endo}$ & $H_{7'\ exo}$); 1.42 – 1.37 (0.7H, m, $H_{7b'\ endo}$); 0.99 – 0.93 (0.7H, m, $H_{3b'\ endo}$); **δ $_C$** (125 MHz, $CDCl_3$); 172.0 (*C*OOH); 159.4 (NH*C*(O)O); 140.3 ($C_{5'exo}$); 139.3 ($C_{5'\ endo}$); 133.7 ($C_{6'\ exo}$); 132.7 ($C_{6'\ endo}$); 76.6 (C-$_{2'}$); 54.1 (C_α); 47.3 ($C_{1'}$ & $C_{7'}$); 42.0 ($C_{4'}$); 41.3 (C_ε); 35.6 ($C_{3'}$); 31.3 (C_β); 30.5 (C_γ); 23.3 (C_δ); ***m/z*** (ES) (Found: MNa^+, 305.1473. $C_{14}H_{22}N_2O_4$ requires 305.1472); **HPLC** (5-95% A): retention times 2.07 min (*endo*), 2.16 min (*exo*), 100%.

(Z,1*R*,8*S*,9*r*)-9-(Ethoxymethyl)bicyclo[6.1.0]non-4-ene 149a[165]



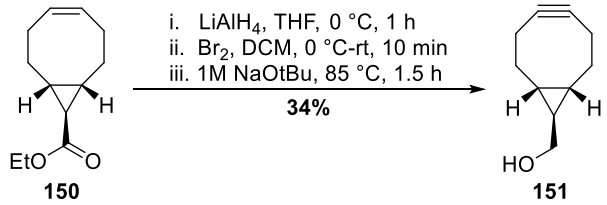
1,5-cyclooctadiene (35 mL, 283 mmol, 8 eq.) and rhodium (II) acetate (94 mg, 0.22 mmol, 0.01 eq.) were added to a flask which was evacuated and charged with nitrogen three times.[165] The reactants were vigorously stirred in dry CH_2Cl_2 (20 mL) and ethyl diazoacetate (85% in CH_2Cl_2, 4.4 mL, 35 mmol, 1 eq.) was added dropwise via a syringe pump over 12 h. The resultant orange solution was concentrated *in vacuo*, and purified using flash chromatography on silica eluting with hexanes, and after collection of excess 1,5-cyclooctadiene, 1% Et_2O in hexanes, to leave *(Z,1R,8S,9r)-9-(Ethoxymethyl)bicyclo[6.1.0]non-4-ene* **149a**[165] (3.17 g, 46%, 18 mmol) (combined isomers: 6.11 g, 34 mmol, 89%) as a colourless oil. **[α]$_D$**^{27} +1.0 (*c* = 0.07, CH_2Cl_2); ***R*$_F$** (1 % Et_2O in hexanes) 0.24; **δ $_H$** (500 MHz, $CDCl_3$): 5.68 – 5.60 (2 H, m, H_4), 4.14 – 4.07 (2 H, m, CH_2CH$_3$), 2.34 – 2.27 (2 H, m, H_3), 2.23 – 2.16 (2 H, m, H_2), 2.13 – 2.05 (2 H, m, $H_{3'}$), 1.59 – 1.54 (2 H, m, H_{2a}), 1.48 (2 H, dddd, $^2J_{H\text{-}H}$ 11.0, $^3J_{H\text{-}H}$ 8.7, $^3J_{H\text{-}H}$ 5.4, $^3J_{H\text{-}H}$ 2.9, $H_{2'}$), 1.25 (2 H, 2, $^3J_{H\text{-}H}$ 7.2, CH$_2$CH_3), 1.19 (1 H, t, $^3J_{H\text{-}H}$ 4.6, H_1); **δ $_C$** (125 MHz, $CDCl_3$): 174.4 (*C*OOEt), 129.9 (C_4), 60.2 (*C*H$_2$CH$_3$), 28.3 (C_2), 27.9 (C_{2a}), 27.7 (C_1), 26.6 (C_3), 14.20 (CH$_2$*C*H$_3$); **ν$_{max}$** (solid)/cm^{-1}: 2934 (alkyne), 1095 (ether); ***m/z*** (EI) (Found: MH^+, 195.1388. $C_{12}H_{19}O_2$ requires 195.1380; **HPLC** (5-95% B): retention time 3.77 min, 99%.

(Z,1*R*,8*S*,9*s*)-9-(Ethoxymethyl)bicyclo[6.1.0]non-4-ene 149b[165]



(Z,1R,8S,9s)-9-(ethoxymethyl)bicyclo[6.1.0]non-4-ene **149b** was synthesised alongside *(Z,1R,8S,9r)-9-(ethoxymethyl)bicyclo[6.1.0]non-4-ene* **149a** and gave the *title compound* **149b**[165] (2.94 g, 16 mmol, 43%) as a colourless oil. $[\alpha]_D^{27}$ +3.2 (*c* = 0.03, CH_2Cl_2); R_F (1 % Et_2O in hexanes) 0.33; δ_H (500 MHz, $CDCl_3$): 5.65 – 5.58 (2 H, m, H_4), 4.12 (2 H, q, $^3J_{H\text{-}H}$ 7.1, CH_2CH_3), 2.55 – 2.46 (2 H, m, H_3), 2.25 – 2.17 (2 H, m, H_2), 2.10 – 2.02 (2 H, m, $H_{3'}$), 1.88 – 1.79 (2 H, m, H_{2a}), 1.71 (1 H, t, $^3J_{H\text{-}H}$ 8.8, H_1), 1.44 – 1.35 (2 H, m, $H_{2'}$), 1.26 (2 H, t, $^3J_{H\text{-}H}$ 7.1, CH_2CH_3). δ_C (125 MHz, $CDCl_3$): 172.31 (CH_2OEt), 129.47 (C_4), 59.72 (CH_2CH_3), 27.09 (C_3), 24.18 (C_{2a}), 22.68 (C_2), 21.27 (C_1), 14.41 (CH_2CH_3); ν_{max} (solid)/cm^{-1}: 2937 (alkyne), 1010 (ether); ***m/z*** (EI) (Found: MH^+, 195.1383. $C_{12}H_{19}O_2$ requires 195.1380); **HPLC** (5-95% B): retention time 4.00 min, 100%.

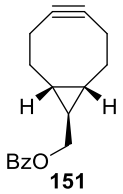
(Z,1*R*,8*S*,9*r*)-Bicyclo[6.1.0]non-4-ene-9-ylmethanol 150[165]



Under a nitrogen atmosphere, a solution of *(Z,1S,8S,9r)-9-(ethoxymethyl)bicyclo[6.1.0]non-4-ene* **149a** (4.58 g, 23.6 mmol, 1 eq.) in CH_2Cl_2 (100 mL) was added dropwise to a stirred suspension of lithium aluminium hydride (2.7 g, 71 mmol, 3 eq.) in CH_2Cl_2 (100 mL) at 0 °C. The suspension was allowed to warm to room temperature and stirred for 2 h until TLC showed complete reduction of the ester (3:1 hexanes:EtOAc). The mixture was cooled to 0 °C and quenched with sodium sulphate decahydrate (~20 g), filtered through a celite pad, concentrated and azeotroped with toluene to leave a colourless oil. Under a nitrogen atmosphere the reduction product was dissolved in CH_2Cl_2 (100 mL) and stirred at 0 °C before a solution of bromine (1.6 mL, 30.7 mmol, 1.3 eq.) in DCM (18 mL) was added dropwise until a persistent red colour was observed. The reaction was quenched with 10% sodium thiosulphate solution (100 mL), the organics were combined and washed with H_2O (2 × 80 mL), brine (1 × 70 mL), concentrated and azeotroped with toluene. The resultant yellow oil was dissolved in THF (100 mL) under a nitrogen atmosphere and cooled to 0 °C before potassium *tert*-butoxide (1M in THF, 71 mL, 71 mmol, 3 eq.) was added dropwise. The solution was heated to reflux and stirred for 16 h before being cooled to room

temperature, quenched with sat. NH_4Cl_{aq} solution (100 mL), extracted with DCM (3 × 100 mL), brine (1 × 100 mL), dried ($MgSO_4$) and concentrated to leave an orange oil. The crude residue was purified by column chromatography on silica gel eluting with 1% Et_2O in CH_2Cl_2 to leave *(Z,1R,8R,9r)-bicyclo[6.1.0]non-4-ene-9-ylmethanol* **150**[165] (1.2 g, 8.0 mmol, 34 %) as a viscous yellow oil. $[\alpha]_D^{27}$ +2.9 (*c* = 0.12, CH_2Cl_2); R_F (1 % Et_2O in hexanes) 0.17; δ_H (500 MHz, $CDCl_3$): 3.56 (2 H, d, $^3J_{H\text{-}H}$ 6.3, CH_2OH), 2.42 (2 H, dd, $^2J_{H\text{-}H}$ 13.3, $^3J_{H\text{-}H}$ 2.7, H_2), 2.33 – 2.25 (2 H, m, H_3), 2.20 – 2.13 (2 H, m, $H_{3'}$), 1.45 – 1.35 (2 H, m, $H_{2'}$), 0.74 – 0.64 (3 H, m. H_{2a+1}). δ_C (125 MHz, $CDCl_3$): 98.8 (C_4), 67.2 (CH_2OH), 33.4 (C_2), 27.3 (C_1), 22.6 (C_{2a}), 21.4 (C_3); ν_{max} (solid)/cm^{-1}: 3350 (OH stretch), 2923 (alkyne); **HPLC** (5-95% B): retention time 1.85 min, 86%.

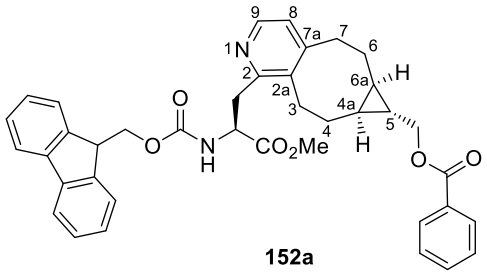
(1*R*,8*S*,9*r*)-9-((Benzoyloxy)methyl)bicyclo[6.1.0]non-4-yne 151[166]



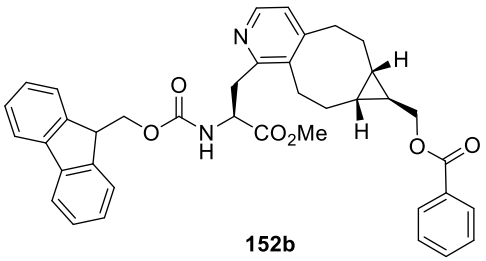
151

(Z,1R,8S,9r)-bicyclo[6.1.0]non-4-ene-9-ylmethanol **150** (159 mg, 1.1 mmol, 1 eq.) and 4-dimethylaminopyridine (8 mg, 0.06 mmol, 0.06 eq.) were stirred in CH_2Cl_2 (10 mL) under a nitrogen atmosphere and freshly distilled triethylamine (dried with 4Å MS, 0.45 mL, 1.32 mmol, 3 eq.) was added.[166] The solution was cooled to 0 °C and benzoyl chloride (0.24 ml, 2.11 mmol, 2 eq.) was added dropwise, the colourless solution became amber. The reaction mixture was allowed to warm to room temperature and stirred for 4 h until TLC showed complete consumption of starting material (4:1 hexanes:EtOAc). The reaction was quenched with sat. NH_4Cl_{aq} (10 mL) and extracted with CH_2Cl_2 (3 × 20 mL), the combined organics were washed with brine (2 × 15 mL), dried ($MgSO_4$) and concentrated to leave a brown oil, which was purified by column chromatography on silica (25:1 hexanes:EtOAc) to leave *(Z,1R,8S,9r)-Bicyclo[6.1.0]non-4-ene-9-ylmethanol* **151**[166] (144 mg, 0.59 mmol, 57 %) as a white solid. $[\alpha]_D^{27}$ +1.2 (*c* = 0.07, CH_2Cl_2); R_F (25:1 hexanes:EtOAc) 0.09; δ_H (500 MHz, $CDCl_3$): 8.08 – 8.04 (2 H, m, Bz-H_2), 7.59 – 7.55 (1 H, m, Bz-H_4), 7.45 (2 H, t, $^3J_{H\text{-}H}$ 7.8, Bz-H_3), 4.27 (2 H, d, $^3J_{H\text{-}H}$ 6.4, OCOCH_2), 2.48 – 2.41 (2 H, m, H_2), 2.36 – 2.27 (2 H, m, H_3), 2.18 (2 H, dd, $^2J_{H\text{-}H}$ 9.4, $^3J_{H\text{-}H}$ 5.6, $H_{3'}$), 1.42 (2 H, d, $^2J_{H\text{-}H}$ 11.5, $H_{2'}$), 0.90 – 0.80 (3 H, m, H_{2a+1}); δ_C (125 MHz, $CDCl_3$): 132.83 (Bz-C_4), 130.5 (Bz-C_1), 129.6 (Bz-C_2), 128.3 (Bz-C_3), 98.8 (C_4), 69.1 (CH_2OH), 33.3 (C_2), 23.6 (C_1), 23.1 (C_{2a}), 21.4 (C_3); ν_{max} (solid)/cm^{-1}: 2227 (alkyne), 1263 (ether); **HPLC** (5-95% B): retention time 3.93 min, 100%.

Methyl (2*S*)-3-[2-((1*S*,8*R*,9*r*)-9-benzoyloxymethylbicyclo[6.1.0]nona[4,5-c]pyridyl)-2-((1*S*)*N*-(9-fluorenylmethoxycarbonyl)amino)propionate & Methyl (2*S*)-3-[2-((1*R*,8*S*,9*s*)-9-benzoyloxymethylbicyclo[6.1.0]nona[4,5-c]pyridyl)-2-((1*S*)*N*-(9-fluorenylmethoxycarbonyl)amino)propionate 152a & 152b



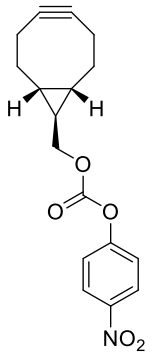
152a



152b

To a solution of *(Z,1R,8S,9r)-Bicyclo[6.1.0]non-4-ene-9-ylmethanol* **151b** (66 mg, 0.26 mmol, 1 eq.) in CH_2Cl_2 (2 mL), *Fmoc-TrzAla-OMe* **N** (105 mg, 0.26 mmol, 1 eq.) in CH_2Cl_2 (2 mL) was added and the reaction was stirred at 37 °C for 16 h, at which time, complete consumption of (Z,1S,8R,9r)-Bicyclo[6.1.0]non-4-ene-9-ylmethanol **151** was observed (4:1 hexanes:EtOAc). The orange solution was concentrated *in vacuo* and purified by column chromatography on silica gel (5% MeOH in CH_2Cl_2). The resultant pale yellow solid was dissolved in dioxane and lyophilised to leave *Methyl (2S)-3-[2-((1S*,8R*,9r*)-9-Benzoyloxymethylbicyclo[6.1.0]nona[4,5-c]pyridyl)-2-(1S)N-(9-fluorenylmethoxycarbonyl)amino)propionate* (1:1) **152a** and **152b** (62 mg, 0.09 mmol, 38%) as a pale yellow, flocculent solid. $[\alpha]_D^{27}$ +4.7 (*c* = 0.11, CH_2Cl_2); R_F (25:1 hexanes:EtOAc) 0.09; δ_H (CD_2Cl_2, 500 MHz); 8.26 – 8.18 (1H, m, H_9), 8.08 – 7.96 (2H, m, Bz-H_2), 7.79 (2H, dd, $^3J_{H\text{-}H}$ 7.6, $^4J_{H\text{-}H}$ 4.0 Hz, 2 × Fmoc-H_4), 7.67 – 7.55 (3H, m, Bz-H_4 & 2 × Fmoc-H_1), 7.53 – 7.43 (2H, m, Bz-H_3), 7.43 – 7.38 (2H, m, 2 × Fmoc-H_3), 7.37 – 7.28 (2H, m, 2 × Fmoc-H_2), 7.03 – 6.93 (1H, m, H_8), 6.64 – 6.55 (1H, m, N*H*), 4.83 – 4.73 (1H, m, H_α), 4.43 – 4.29 (2H, m, Fmoc-CH_2), 4.25 (1H, t, $^3J_{H\text{-}H}$ 7.1 Hz, Fmoc-C*H*), 4.12 – 3.98 (2H, m, Bz-C(O)OCH_2), 3.67 (3H, s, OCH_3), 3.54 (1H, ddd, $^2J_{H\text{-}H}$ 17.0, $^4J_{H\text{-}H}$ 12.1, $^3J_{H\text{-}H}$ 5.6 Hz, H_β), 3.31 (1H, ddd, $^2J_{H\text{-}H}$ 16.0, $^4J_{H\text{-}H}$ 11.4, $^3J_{H\text{-}H}$ 4.3 Hz, $H_{\beta'}$), 3.05 – 2.92 (2H, m, H_{3+7}), 2.87 – 2.72 (2H, m, $H_{3'+7'}$), 2.64 – 2.44 (2H, m, H_{4+6}), 1.48 – 1.33 (2H, m, $H_{4'+6'}$), 0.95 – 0.88 (1H, m, H_5), 0.85 – 0.70 (2H, m, H_{4a}). δ_C (125 MHz, CD_2Cl_2): 172.8 (*C*OOCH$_3$), 166.8 (Bz-*C*(O)OCH$_2$), 156.5 (Fmoc-*C*(O)ONH), 155.0 (C_2), 152.3 (C_{2a}), 146.1 (C_9), 144.4 (2 × Fmoc-C_{4a}), 141.6 (2 × Fmoc-C_{1a}), 136.6 (C_{7a}), 133.1 (2 × Fmoc-C_1), 131.0 (Bz-C_1), 129.8 (Bz-C_2), 128.7 (Bz-C_3), 128.0 (2 × Fmoc-C_3), 127.4 (2 × Fmoc-C_2), 125.5 (Bz-C_4), 124.5 (C_8), 120.3 (2 × Fmoc-C_4), 68.8 ($C_{5'}$), 67.2 (Fmoc-*C*H$_2$), 53.1 (COO*C*H$_3$), 52.5 (C_α), 47.6 (Fmoc-*C*H), 35.9 (C_β), 33.9 (C_3), 29.0 (C_4), 26.7 (C_5), 22.4 (C_{4a}); ν_{max} (solid)/cm^{-1} *m/z*: 3335 (NH stretch), 1714 (CO); ***m/z*** (ES): found *M(+H)* 631.2814, $C_{39}H_{39}N_2O_6$ requires 631.2803, **HPLC** (5-95% B): retention time 4.75 min, 100%.

(Z,1*R*,8*S*,9*r*)-Bicyclo[6.1.0]non-4-ene-9-ylmethanol (4-nitrophenyl)carbonate 154[165] (*BCN-ONP*)

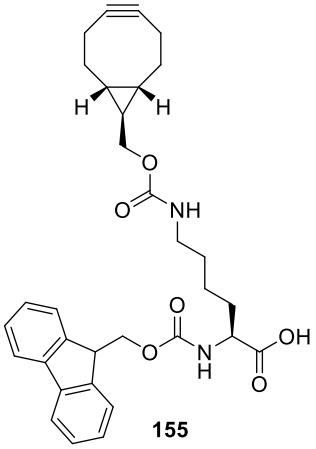


154

(1R,8S,9r)-9-((Benzoyloxy)methyl)bicyclo[6.1.0]non-4-yne **150** (1.09 g, 7.3 mmol, 1 eq.) and *p*-nitrophenylchloroformate (1.76 g, 8.7 mmol, 1.2 eq.) were dissolved in anhydrous CH_2Cl_2 (200 mL) and pyridine (wet, 1.5 mL, 18 mmol, 2.5 eq.) was added. The reaction mixture was stirred for 20 minutes, until consumption of **150** was observed by TLC (6:1 hexanes:EtOAc), quenched with $NH_4Cl_{\ aq}$ (200 mL), extracted with CH_2Cl_2 (3 × 150 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The crude product was purified by column chromatography on silica gel, (8:1 hexanes:EtOAc) to give *BCN-ONP* **154**[165] (1.66 g, 5.26 mmol, 73%) as an amorphous white solid. $\boldsymbol{R_F}$ (8:1 hexanes:EtOAc) 0.24; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$): 8.30 – 8.26 (2 H, m, Ph-H_3), 7.41 – 7.37 (2 H, m, Ph-H_2), 4.22 (2 H, d, $^3J_{H\text{-}H}$ 6.9, CH_2OC(O)Ph), 2.45 (2 H, dd, $^2J_{H\text{-}H}$ 13.3, $^3J_{H\text{-}H}$ 2.5, H_2), 2.36 – 2.26 (2 H, m, H_3), 2.22 – 2.15 (2 H, m, $H_{3'}$), 1.48 – 1.36 (2 H, m, $H_{2'}$), 0.91 – 0.78 (3 H, m, H_{2a+1}); $\boldsymbol{\delta_C}$ (125 MHz, $CDCl_3$): 155.61 (O*C*(O)Ph), 152.6 (Ph-C_4), 145.4 (Ph-C_1), 125.3 (Ph-C_3), 121.8 (Ph-C_2), 98.7 (C_4), 74.0 (*C*H_2O(C)OPh), 33.1 (C_2), 23.3 & 23.0 (C_1 & C_{2a}), 21.3 (C_3); ***m/z*** (ES): Found: 631.2294 *2M(+H)* $C_{34}H_{35}N_2O_{10}$ requires 631.2286; HPLC (5-95% A): retention time 3.79 min, 100%.

(2*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)-6-[((Z,1*R*,8*S*,9*r*)-Bicyclo[6.1.0]non-4-ene)oxycarbonylamino]hexanoic acid 155[87]

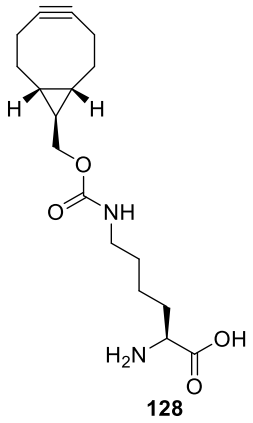
(*Fmoc-BCN-K*)



155

BCN-ONP **154** (158 mg, 0.50 mmol, 1 eq.) and Fmoc-Lys-OH (243 mg, 0.60 mmol, 1.2 eq.) were dissolved in anhydrous DMF (12 mL) and triethylamine (dried with 4Å MS and freshly distilled before use, 121 µL, 0.87 mmol, 1.7 eq.) was added. The reaction was stirred at rt for 45 minutes (at which time consumption of *BCN-ONP* **154** was observed by TLC (3:1 hexanes:EtOAc)), quenched with sat. $Na_2CO_{3\ aq}$ (10 mL) and extracted with CH_2Cl_2 (3 × 15 mL). The combined organics were dried ($MgSO_4$) and concentrated to leave a white foam which was purified by column chromatography on silica, eluting the sample with 10% MeOH in CH_2Cl_2 to leave *Fmoc-BCN-K* **155**[87] (158 mg, 0.33 mmol, 66%) as a brittle white foam. $\boldsymbol{R_F}$ (10% MeOH in CH_2Cl_2) 0.26; $\boldsymbol{\delta_H}$ (500 MHz, $CDCl_3$): 7.90 (2 H, d, $^3J_{H\text{-}H}$ 7.5, Fmoc-H_4), 7.70 (2 H, d, $^3J_{H\text{-}H}$ 7.5, Fmoc-H_1), 7.43 (2 H, t, $^3J_{H\text{-}H}$ 7.3, Fmoc-H_3), 7.34 (2 H, dd, $^3J_{H\text{-}H}$ 7.1, $^4J_{H\text{-}H}$ 4.9, Fmoc-H_2), 7.06 (1 H, br s, BCN-N*H*), 6.71 (1 H, br s, Fmoc-N*H*), 4.30 (1 H, d, $^4J_{H\text{-}H}$ 10.5, Fmoc-C*H*), 4.23 (2 H, d, $^3J_{H\text{-}H}$ 7.0, Fmoc-CH_2), 3.84 (2 H, d, $^3J_{H\text{-}H}$ 6.9, BCN-CH_2), 3.70 (1 H, br s, H_α), 2.94 (2 H, d, $^3J_{H\text{-}H}$ 6.4, H_ε), 2.33 – 2.16 (4 H, m, H_2 & H_3), 2.07 (2 H, d, $^2J_{H\text{-}H}$ 15.1, $H_{3'}$), 1.69 (1 H, br s, H_β), 1.55 (1 H, br s, H_β), 1.43 – 1.20 (5 H, m, H_γ, H_δ & $H_{2'}$), 0.71 – 0.58 (3 H, m, H_{2a+1}).

(2*S*)-2-amino-6-[((Z,1*R*,8*S*,9*r*)-Bicyclo[6.1.0]non-4-ene)oxycarbonylamino]hexanoic acid[87]



128

Fmoc-BCN-K **155** (374 mg, 0.69 mmol) was stirred in 20% piperidine in DMF (40 mL) for 30 min at room temperature, the volatiles were removed as far as possible *in vacuo* and the resultant solid azeotroped with toluene three times. The off-white solid was triturated in ice-cold diethyl ether for 1 h and gravity filtered. The precipitate was dissolved in H_2O and lyophilised to leave *BCN-lysine* **128**[87] (181 mg, 0.56 mmol, 81%) as a flocculent white solid. $\boldsymbol{\delta}_{\mathbf{H}}$ (500 MHz, *d*6-DMSO): 3.86 (2 H, d, $^3J_{H\text{-}H}$ 6.2, CH_2OC(O)NH), 3.12 – 3.06 (1 H, s, H_α), 2.95 (2 H, d, $^3J_{H\text{-}H}$ 6.0, H_ε), 2.35 – 2.19 (4 H, m, H_2 & H_3), 2.09 (2 H, d, $^2J_{H\text{-}H}$ 16.4, $H_{3'}$), 1.76 – 1.66 (1 H, m, H_β), 1.59 – 1.48 (1 H, m, H_β), 1.41 – 1.26 (5 H, m, H_γ, H_δ & $H_{2'}$), 0.72 – 0.60 (3 H, m, H_{2a+1}); ***m/z*** (ES): Found: 323.1973 *M(+H)* $C_{17}H_{26}N_2O_4$ requires 323.1965; **HPLC** (5-95% A): retention time 0.26 min, 100%.

6.2 Peptide Synthesis

Resins and amino acids were purchased from Novabiochem: Fmoc-Asn(Trt)-OH, Fmoc-Gaba-OH[i], Fmoc-His(Trt)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Val-OH, Fmoc-Asn(Trt)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Ala-OH, Fmoc-Thr(tBu)-OH, Fmoc-Ser(PO(OH)(OBn))-OH, Fmoc-Thr(PO(OH)(OBn))-OH, Fmoc-Tyr(PO(OH)(OBn))-OH and Fmoc-Tyr(PO(NMe_2)$_2$)-OH. Q-Sepharose FF resin was purchased from GE Healthcare. Other reagents were purchased from Sigma Aldrich, Alfa Aesar, Merck or Fisher Scientific and were used without further purification.

Peptides were generated manually using a standard solid phase peptide synthesis protocol: Resin (amount specified in each reaction) was swollen in DMF (2-5 mL per 100 mg of resin – the same volume was used for the wash steps) for 30-90 min with agitation (Stuart Rotator SB2). After removal of DMF by vacuum filtration and the first coupling mixture added.

[i] Gaba = γ-aminobutyric acid

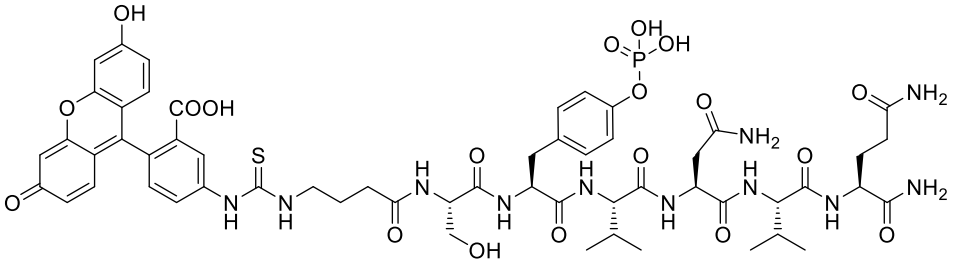
Coupling reactions using commercially available amino acids were performed using with respect to resin 5 eq. Fmoc-amino acid, 5 eq. HCTU and 10 eq. DIPEA in DMF (minimal volume) for 60 min with agitation – unless otherwise stated. After each coupling reaction the resin was drained by vacuum filtration and washed with DMF (3 × 2 min), 20% piperidine in DMF (5 × 2 min) and DMF (5 × 2 min). Subsequent couplings were carried out in the same manner.

Peptide N-termini were acetylated by mixing the drained resin with a solution containing acetic anhydride (5 eq.) and DIPEA (5 eq.) in DMF (2-5 mL per 100 mg of resin) for 30 min with agitation, the solution was removed from the resin by vacuum filtration and washed with DMF (3 × 2 min).

To prepare the resin for cleavage it was washed with CH_2Cl_2 (3 × 2 min) and MeOH (3 × 2 min) then dried for a minimum of 2 h *in vacuo*. The peptide was cleaved from the resin by mixing with a cleavage cocktail (2-5 mL per 100 mg of resin) consisting of TFA (95%), H_2O (2.5%) and triisopropylsilane (2.5%). The cleavage cocktail was applied to the resin and mixed for 2 h, then then precipitated into ice-cold diethyl ether (10 × volume of cleavage cocktail – same volume used in later ethereal washes) and the precipitate collected by centrifugation (4,000 × *g*, 10 minutes). The ethereal supernatant was decanted, the peptide pellet resuspended in ice-cold diethyl ether and centrifuged again. This was repeated 2 more times before residual ether was removed under a stream of nitrogen. The resultant amorphous solid was dissolved in the minimum volume H_2O (with the addition of small portions of dioxane), frozen and lyophilized before further purirfication.

If stated, peptides were purified by ion exchange chromatography: the lyophilised peptide was dissolved in H_2O, loaded onto a bed of Q-Sepharose FF resin or SP-Sepharose FF resin (1 mL), washed with with H_2O (10 mL), then eluted with a stepwise gradient of $NH_4HCO_{3(aq)}$ (10, 20, 50, 100, 200 and 500 mM, 10 mL per elution) under gravity. Fractions containing the peptide were identified by LCMS, combined (as appropriate) and lyophilised. Q-Sepharose FF resin was cleaned by washing with 1 M $NH_4HCO_{3\ aq}$ (10 mL) then H_2O (10 mL) and stored in 20% EtOH at 4 °C.

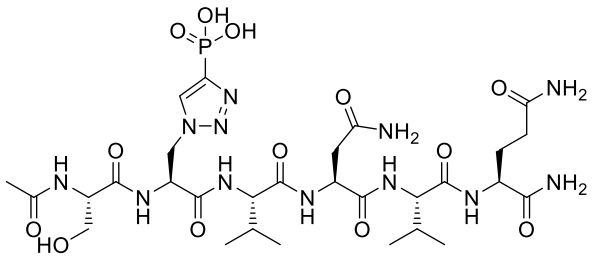
FITC-Gaba-Ser-pTyr-Val-Asn-Val-Gln-NH_2 95



95

FITC-Gaba-Ser-pTyr-Val-Asn-Val-Gln-NH_2 **95** was synthesised using Rink Amide Novagel™ resin (0.64 mmol/g loading; 100 mg, 0.064 mmol). Coupling of Fmoc-Tyr($PO(NMe_2)_2$)-OH was carried out using 5 eq. amino acid, 5 eq. HCTU and 10 eq. DIPEA in DMF for 1 hour. Coupling of Fmoc-GA-OH was carried out using 3 eq. amino acid, 3 eq. HCTU and 6 eq. Coupling of FITC was carried out using 6 eq. amino acid, 6 eq. DIPEA with agitation in the dark for 16 h. Everything subsequent to coupling of FITC was carried out in the dark. For cleavage, a standard cleavage cocktail was applied to the resin and mixed for 2 h, 10% H_2O (*v/v*) was added and the resin was swelled overnight. The peptide was then precipitated according to standard procedure and lyophilised. The crude peptide was dissolved in H_2O and purified by anion-exchange chromatography (Q-Sepharose). After lyophilisation of relevant fractions *FITC-Gaba-Ser-pTyr-Val-Asn-Val-Gln-NH_2* **95** was obtained as a flocculent orange solid (46 mg, 0.036 mmol, 57%). ***m/z*** (ES): Found $M(-2H)^{2+}$ 629.6998, $C_{56}H_{66}N_{11}O_{19}PS$ requires 629.7003; **HPLC** (5-95% A): retention time 2.16 min, 100%.

AcHN-Ser-pTzAla-Val-Asn-Val-Gln-NH_2 91

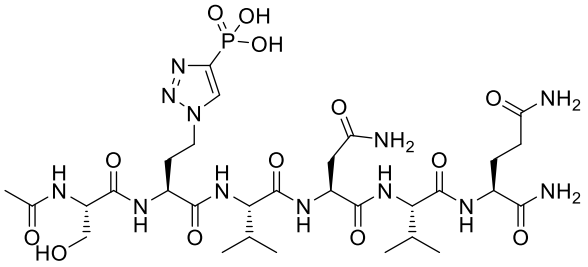


91

AcHN-Ser-pTzAla-Val-Asn-Val-Gln-NH_2 **91** was synthesised using Rink Amide Novagel™ resin (0.64 mmol/g loading; 100 mg, 0.064 mmol). Coupling of *Fmoc-pTz$(OBn)_2$-OH* **15** was carried out using 3 eq. amino acid (123 mg, 0.192 mmol), 2.9 eq. HCTU and 6 eq. DIPEA in DMF for 1 h. After cleavage and lyophilisation, the crude peptide was dissolved in H_2O and purified by anion-exchange chromatography (Q-Sepharose). After lyophilisation of relevant fractions, *AcHN-Ser-pTzAla-Val-Asn-Val-Gln-NH_2* **91** was obtained as a

flocculent white solid (6.81 mg, 0.0085 mmol, 13%). δ_H[ii] (500 MHz, CD_3OD): 8.02 (1H, s, pTzAla-H_5), 5.01 – 4.91 (1H, m, Asn4-H_α), 4.41 (1H, t, $^3J_{H\text{-}H}$ 5.6, Ser1-H_α), 4.34 (1H, dd, $^3J_{H\text{-}H}$ 9.5, $^3J_{H\text{-}H}$ 5.1, Gln6-H_α), 4.19 – 4.11 (2H, m, Val3-H_α & Val5-H_α), 3.77 (2H, d, $^3J_{H\text{-}H}$ 5.9, Ser1-H_β), 2.94 – 2.78 (2H, m, Asn4-H_β), 2.45 – 2.39 (2H, m, Gln6-H_γ), 2.33 – 2.11 (2H, m, Val3-H_β & Val5-H_β), 1.01 – 0.91 (12H, m, Val3-H_γ & Val5-H_γ), 2.10 (3H, s, Acetyl-CH_3), 2.09 – 2.09 (2H, m, Gln6-H_β). ***m/z*** Found *M(-H)* 803.3241, $C_{29}H_{48}N_{12}O_{13}P$ requires 803.3207; **HPLC** (5-95% A): retention time 1.12 min, 92%.

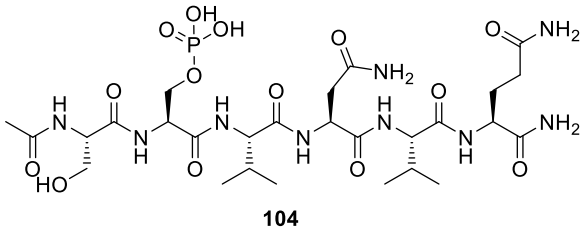
AcHN-Ser-phTzAla-Val-Asn-Val-Gln-NH_2 92



92

AcHN-Ser-phTzAla-Val-Asn-Val-Gln-NH_2 **92** was prepared by Dr Tom McAllister.[175]

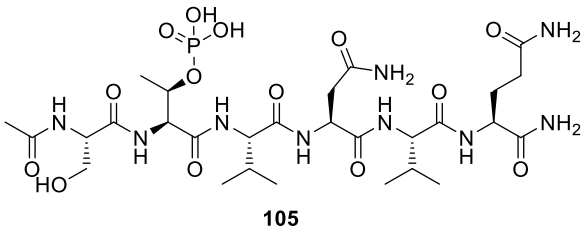
AcHN-Ser-pSer-Val-Asn-Val-Gln-NH_2 104



104

AcHN-Ser-pSer-Val-Asn-Val-Gln-NH_2 **104** was prepared by Dr Tom McAllister.[175]

AcHN-Ser-pThr-Val-Asn-Val-Gln-NH_2 105

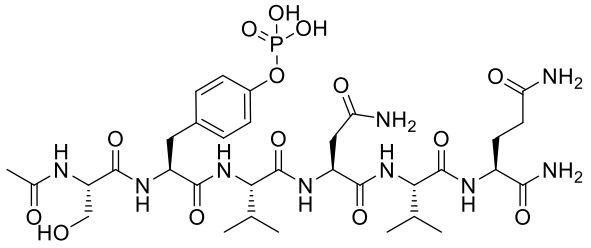


105

AcHN-Ser-pThr-Val-Asn-Val-Gln-NH_2 **105** was prepared by Dr Tom McAllister.[175]

[ii] Signals for pTzAla-H_α & H_β are coincident with the suppressed residual water peak.

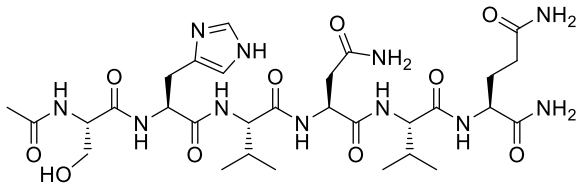
AcHN-Ser-pTyr-Val-Asn-Val-Gln-NH_2 90



90

AcHN-Ser-pTyr-Val-Asn-Val-Gln-NH_2 **90** was synthesised using Rink Amide Novagel™ resin (0.64 mmol/g loading; 100 mg, 0.064 mmol). Coupling of Fmoc-Tyr(PO(OH)(OBn))-OH was carried out using 5 eq. amino acid, 5 eq. HCTU and 10 eq. DIPEA in DMF for 1 hour. After cleavage and lyophilisation, the crude peptide was dissolved in H_2O and purified by anion-exchange chromatography (Q-Sepharose). After lyophilisation of relevant fractions, *AcHN-Ser-pTyr-Val-Asn-Val-Gln-NH_2* **90** was obtained as a flocculent white solid (26 mg, 0.031 mmol, 49%). ***m/z*** Found $M(-2H)^{2+}$ 828.3291, $C_{32}H_{51}N_9O_{14}P$ requires 828.3299; **HPLC** (5-95% A): retention time 1.26 min, 79%.

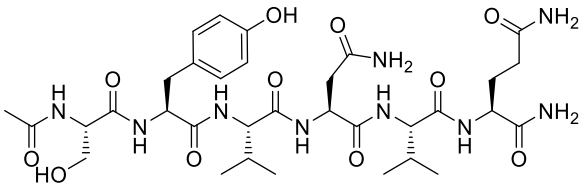
AcHN-Ser-His-Val-Asn-Val-Gln-NH_2 106



106

AcHN-Ser-His-Val-Asn-Val-Gln-NH_2 **106** was prepared by Dr Tom McAllister.[175]

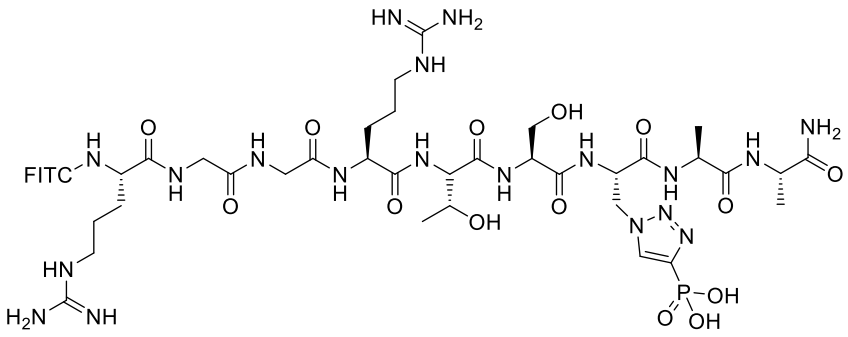
AcHN-Ser-Tyr-Val-Asn-Val-Gln-NH_2 107



107

AcHN-Ser-Tyr-Val-Asn-Val-Gln-NH_2 **107** was prepared by Dr Tom McAllister.[175]

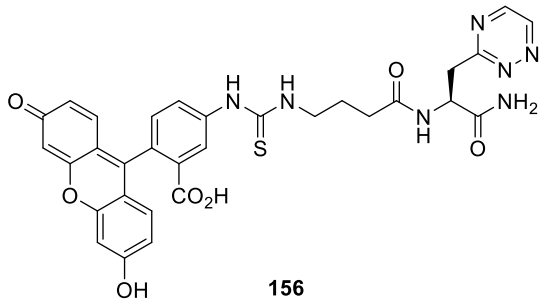
FITC-HN-Arg-Gly-Gly-Arg-Thr-Ser-His-pTzAla-Ala-NH_2 110



110

FITC-HN-Arg-Gly-Gly-Arg-Thr-Ser-His-pTzAla-Ala-NH_2 **N** was synthesised using Rink Amide Novagel™ resin (0.64 mmol/g loading; 100 mg, 0.064 mmol). Coupling of FITC was carried out using 5 eq. amino acid, 5 eq. DIPEA with agitation in the dark for 16 h. Everything subsequent to coupling of FITC was carried out in the dark. For cleavage, a standard cleavage cocktail was applied to the resin and mixed for 2 h, 10% H_2O (*v/v*) was added and the resin was swelled overnight. The peptide was then precipitated according to standard procedure and lyophilised. The crude peptide was dissolved in H_2O and purified by cation-exchange chromatography (SP-Sepharose) and further purified by UV-directed (HPLC 50-95% MeCN + 0.1% formic acid) After lyophilisation of relevant fractions *FITC-Gaba-Ser-pTyr-Val-Asn-Val-Gln-NH_2* **110** was obtained as a flocculent orange solid (31 mg, 0.021 mmol, 33%). ***m/z*** (ES): Found 733.7753, $C_{59}H_{82}N_{21}O_{20}PS$ requires 733.7746; **HPLC** (5-95% A): retention time 1.64 min, 100%.

FITC-Gaba-TrzAla-$CONH_2$ 156



156

FITC-Gaba-TrzA-$CONH_2$ **156** was synthesised using Rink Amide Novagel™ (0.64 mmol/g loading, 50 mg, 0.032 mmol). Coupling of *Fmoc-TrzAla-OH* **156** was carried out using 3 eq. amino acid, 3 eq. HCTU and 3 eq. DIPEA in DMF for 1 hour. Coupling of Fmoc-GA-OH was carried out using 5 eq. amino acid, 5 eq. HCTU and 10 eq. Coupling of FITC was carried out using 5 eq. amino acid, 5 eq. DIPEA with agitation in the dark for 16 h. Everything subsequent to coupling of FITC was carried out in the dark. After cleavage and lyophilisation the crude peptide was purified by mass-directed HPLC (5-95% MeCN + 0.1%

formic acid). After lyophilisation of relevant fractions *FITC-Gaba-TrzA-CONH$_2$* **156** was obtained as a flocculent orange solid compound (7.52 mg, 0.12 mmol, 37%). ***m/z*** (ES): Found *M(+H)* 642.1759 $C_{31}H_{28}N_7O_7S$ requires 642.1765; **HPLC** (5-95% A): retention time 1.11 min, 100%.

6.3 Assays

Fluorescence Polarisation Assays

Fluorescence polarization assays were carried out in triplicate at rt in black 96 well plates (160 μL per well). All experiments were performed in 50 mM Tris.HCl, 100 mM NaCl pH 7.4 (*SEC buffer*). A 2.5–fold or 3–fold dilution series of each peptide or protein were routinely prepared before mixing in the plate well. Data was acquired using a Perkin Elmer EnVision™ 2103 MultiLabel plate reader (fluorescein FP mirror, fluorescein FP 480 excitation filter and P-pol 535 and S-pol 535 emission filters).

Rate Determination Experiments

Fmoc-TrzAla-OMe **142** (0.73 mg, 1.80 μmol, 1 eq) was dissolved in 900 μL of the required solvent mix (either 100% MeCN or 10% H_2O). *BCN-Bz* **151** was dissolved in 900 μL of the required solvent mix ((0.46 mg, 1.8 μmol), (0.92 mg, 3.6 μmol), (1.83 mg, 7.20 μmol) or (3.66 mg (14.3 μmol) with respect to a final concentration (in 1.8 mL) of 1 mM, 2 mM, 4 mM or 8 mM respectively. The two solutions were combined, mixed thoroughly and capped before being immediately incubated at 37 °C in the auto loader of an Agilent technologies 1290 HPLC, fitted with an ACE Ultracore 2.5 Super C18 column (50 × 2.1 mm).

Each 2 μL sample was injected and analysed using a linear gradient (B, 5% H_2O, in MeCN to 95% H_2O) over 5 min. Typically 4 reactions of varying concentration in respect to BCN-Bz were assayed in parallel. After every analysis, 2 μL of MeCN was injected onto the column and the column washed using the same linear gradient over 9 min. HPLC analysis was carried out over a 20 hour period. Concentration of the cycloaddition product was measured against a previously measured concentration dilution series of the product. A typical calibration graph (Figure 6.2) and HPLC trace (Figure 6.1) are shown below.

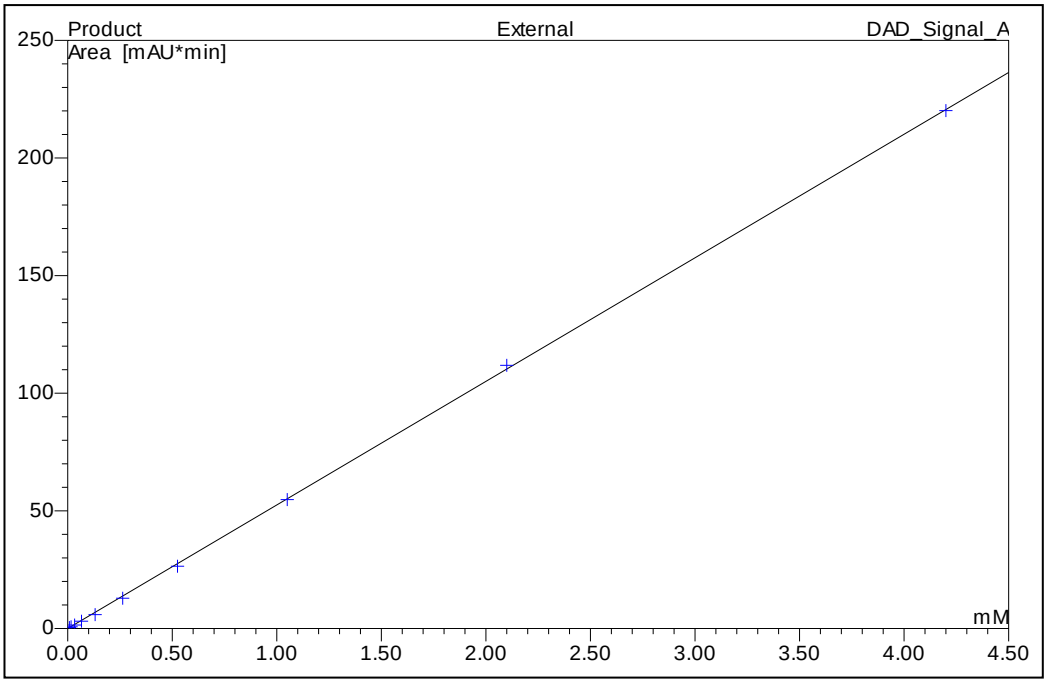


Figure 6.1: Calibration curve for determination of cycloaddition product formation by HPLC

.

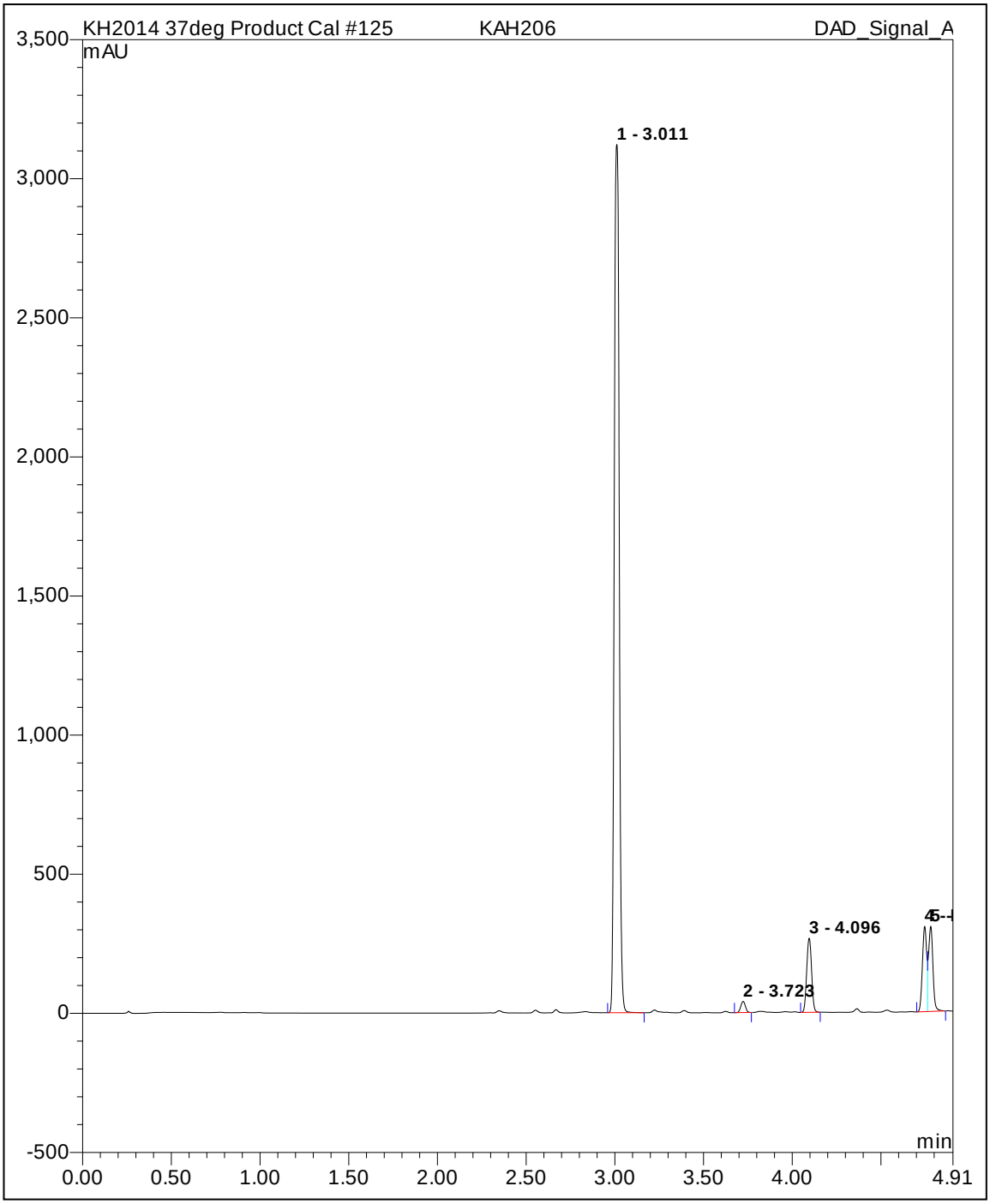


Figure 6.2: Representative chromatogram from cycloaddition rate determination. 1. Fmoc-TrzAla-OMe **142**, 3. BCN-Bz **151** and 4/5 cycloaddition products **152a** and **152b**

6.4 Procedures for Protein Production, Purification and Characterisation

6.4.1 Materials

Instrumentation

Media, buffer and other equipment were sterilised using either a Prestige Medical bench top autoclave or an LTE Touchclave-R autoclave. All contaminated glassware, plastic-ware and media containing genetically modified bacteria was disposed of by bleaching, using either Virkon or Presept for 24 h or autoclaved (120 °C for 20 min) before disposal as aqueous waster for incineration. When a sterile environment was req.uired, work was undertook in a Thermo electron corporation Microbiological safety cabinet SAFE 2010 Class II or on the benchtop under a blue flame. Bacterial cultures were incubated in a Kuhner ShakerX ISF1-X or a Stuart orbital incubator. LB Agar plates and small scale assays were incubated in a Binder BD23 incubator at 37°C. Centrifugation was carried out in a Beckman Coulter™ Avanti™ J-301 centrifuge and in bench-top centrifuges Pico™ and Fresco™ 17/21 from Thermo electron corporation. Cells were disrupted using a Bandelm Sonopuls HD2070 sonicator or a Constant Systems Cell Disrupter (T2/40/AA/AA). Size exclusion chromatography was performed using an Äkta Purifier FPLC from GE Healthcare with a 20/60 Superdex 75 column. Spectrophotometric readings were measured using a Thermo Scientific Nanodrop 2000. SDS-PAGE electrophoresis was carried out using a Bio-Rad mini-protean 3 apparatus and the polyacrylamide gels were imaged with both UV and white light using a Bio-Rad molecular imager® Gel Doc™ XR. Water was purified before use to 15 Mq using an ELGA PURLAB® Classic purification system. Protein mass spectrometry was performed in original buffers on a Bruker MaXis impact spectrometer and the raw mass spectrum were deconvoluted using a maximum entropy algorithm part of Bruker data analysis. Results are quoted as mass-charge ratios (m/z).

Materials

Analytical grade reagents were purchased from Sigma Aldrich, VWR International, Fisher Scientific and Alfa Aesar unless otherwise stated. Ni-NTA agarose was purchased from Qiagen. Enzymes, buffers and molecular weight markers were supplied by New England Biolabs. DNase I recombinant and Protease inhibitor cocktail (cOmplete™) were purchased from Roche Diagnostics. Instant blue was purchased from PAGEgel inc.

Media

Lysogeny Broth (LB): LB pre-mix freeze dried powder (consisting of tryptone (10 g/L), yeast extract (5 g/L), NaCl 10 (g/L)) was made up with purified water to a quantity of 25 g/L, autoclaved at 121 °C for 20 minutes as required and allowed to cool before use.

Agar

Agar plates were made with agar (15 g/L) and LB (25 g/L) before being autoclaved at 121 °C for 20 minutes and stored at 60 °C. Agar was allowed to cool for 15 minutes before being inoculated with the desired antibiotic(s), poured onto Petri dishes in 20 ml portions and left to set in a sterile environment. Plates were stored at 4 °C prior to use.

Antibiotics

All antibiotic stock solutions were made to 1000× working concentration. Kanamycin stock was made to 50 mg/mL (using kanamycin sulfate from BioChemica). Both ampicillin and spectinomycin stock were purchased from Melford and made up to 100 mg/mL (using the sodium salt of ampicillin and the hydrochloride salt of spectinomycin respectively).

Cell strains

Chemically competent cell lines of *E. coli* MG1655 ΔDZ DE3 (double deletion of PanD (ADC) and PanZ (yhhk)), C41(DE3) and BL21 (DE3) were available in house. *E. coli* MG1655 ΔDZ DE3 cells were used for all ADC work. *E .coli* C41(DE3) and BL21 (DE3) cells were used for all Grb2SH2 and YdiA work respectively.

Plasmids

Name	Resistance	Gene encoded
pET28a-Grb2SH2	Kanamycin	His_6–Grb2SH2
pCDF-PylT-ADC(K9X)	Spectinomycin	$tRNA_{CUA}$ and His_6–ADCK9X
BCN-PylRS	Ampicillin	Pyrrolysyl tRNA synthetase (Y271M, L274G, C313A)
pRSETA-ADC(S25A)	Ampicillin	His_6–ADCS25A
pMALc2x-YdiA	Ampicillin	MBP–YdiA

pET28a-Grb2SH2 was generated by Dr Tom McAllister;[175] pET-28a was purchased from Novagen. pMALc2x-YdiA was generated by Dr Jeff Hollins.[140] pCDF-PylT-ADC(K9X) was generated by Michael Rugen (MSc student); pCDF-PylT was previously obtained from Jason Chin (MRC lab, University of Cambridge). BCN-PylRS was purchased from GenScript. pRSETA-ADC(S25A) was generated by Diana Monteiro (PhD student).[169]

Buffers and Stains

Recipes where quantities are given in the parenthesis are for 1 L of buffer. Where no quantities are listed, buffers were available in house as stock solutions. Buffers for Ni-NTA chromatography were used without further sterilisation or filtration. Buffers for size exclusion were vacuum filtered through a 0.22 μM membrane.

Phosphate buffer 1: 50 mM sodium phosphate (NaH_2PO_4 1.4 g, Na_2HPO_4 5.42 g), 300 mM NaCl (17.5 g), pH 7.4 at 4 °C

Lysis Buffer 1: *Phosphate Buffer 1* with 10 mM imidazole (681 mg)

Wash Buffer 1: *Phosphate Buffer 1* with 20 mM imidazole (1.36 g)

Wash Buffer 1a: *Phosphate Buffer 1* with 50 mM imidazole (3.40 g)

Elution Buffer 1: *Phosphate Buffer 1* with 250 mM imidazole (17.02 g)

Phosphate Buffer 2: 50 mM potassium phosphate (KH_2PO_4 1.5 g, K_2HPO_4 6.78 g) 300 mM NaCl (17.5 g), pH 7.4 at 4 °C

Lysis Buffer 2: *Phosphate Buffer 2* with 10 mM imidazole (681 mg)

Wash Buffer 2: *Phosphate Buffer 2* with 20 mM imidazole (1.36 g)

Elution Buffer 2: *Phosphate Buffer 2* with 250 mM imidazole (17.02 g)

Inclusion Body Buffer 1: 20 mM Tris (Tris 970 mg, Tris.HCl 1.86 g), 200 mM NaCl (11.7 g), 2 mM EDTA (1.49 g of EDTA di-sodium salt), 1.5% (*v/v*) Triton X-100, (15 mL) pH 8.0 at 25 °C

Inclusion Body Buffer 2: 10 mM Tris (Tris 490 mg, Tris.HCl 930 mg), 1 M NaCl (58.5 g), 1 mM EDTA (750 mg of EDTA di-sodium salt), 1.5% (*v/v*) Triton X-100, (15 mL) pH 8.0 at 25 °C

SEC Buffer: 50 mM Tris (4.93 g Tris, 2.92 g Tris-HCl), 100 mM NaCl (5.84 g), pH 7.4 at 25°C.

2 × Loading Buffer (5 mL): 100 mM Tris HCl (0.6 g) pH 6.8, SDS (0.3 g), DTT (0.2 g), bromophenol blue (2 mg), 10% (*v/v*) glycerol (0.5 mL)

5 × SDS-PAGE Tris-glycine running Buffer: 125 mM Tris (15.15 g), 960 mM glycine (72 g), 0.5% (*w/v*) SDS (5 g)

SDS-PAGE Tris-tricine gel Buffer (100 mL): 3 M Tris (36 g), 0.3% (*w/v*) SDS (300 mg), pH 8.45 (the pH was adjusted by addition of 1 M HCl)

SDS-PAGE Tris-tricine cathode buffer: 0.1 M Tris (12.1 g), 0.1 M tricine (17.9 g), 0.1% (*w/v*) SDS (100 mg), pH 8.25 (the pH was adjusted by addition of 1 M HCl)

SDS-PAGE Tris-tricine anode buffer: 0.2 M Tris (24.2 g), pH 8.90 (the pH was adjusted by addition of 1 M HCl)

Coomassie Stain: Coomassie blue R-250 (1.0 g), 40% (*v/v*) methanol (400 mL), 10% (*v/v*) acetic acid (100 mL)

Coomassie Destain: 40% (*v/v*) methanol (400 mL), 10% (*v/v*) acetic acid (100 mL)

6.4.2 Methods

Transformations

Under sterile conditions chemically competent cells (stored in -80 °C freezer) were thawed on ice. Aliquots (50 µL) were dispensed into sterilised Eppendorf tubes that had been chilled on ice and 1 µL of the req.uired plasmid was mixed with the cells (for double transformations 1 µL of each plasmid was added). The Eppendorf tubes were incubated on ice for 10 minutes, at 42 °C for 30-45 seconds and returned to ice for 10 minutes, LB (900 µL) was added and the culture was incubated at 37 °C for 90 minutes and 100 µL of the aliquot was spread onto LB-agar plates containing the appropriate antibiotic(s). The remaining culture was centrifuged at 13,000*g* for 30 seconds, the Eppendorf tube inverted to remove the majority of the media and the remaining cells were resuspended and spread onto LB-agar plates, again, containing the appropriate antibiotic(s). The LB-agar plates were incubated overnight at 37 °C.

Starter Cultures

Under aseptic conditions, 5 mL of sterile LB media was inoculated with a single colony picked from an LB-agar plate or a scraping from a frozen glycerol stock in a 30 mL sterilin tube. The appropriate antibiotic(s) were added and culture was incubated overnight at 37 °C with shaking at 230 rpm.

Glycerol Cell Stocks

500 µL of starter culture was removed and mixed with 500 µL of sterile 80% glycerol (*v/v*) in a sterile Eppendorf and stored at -80 °C. Cell stocks were used to initiate subsequent starter cultures as described in 0.

Protein Overexpression[iii]

N-terminal hexaHis-tagged Grb2-SH2 was overexpressed in *E .coli* C41(DE3) cells harbouring a pET28a-GrbSH2 plasmid as described by McAllister *et al*.[175] N-terminal MBP–tagged YdiA was overexpressed in *E. coli* BL21 (DE3) harbouring a pMALc2x-YdiA plasmid.[140] N-terminal His_6–ADC(S25A) was overexpressed in *E. coli* MG1655 ΔDZ DE3 cells containing a pRSETA-ADC(S25A) plasmid as described by Monteiro *et al*.[169] N-terminal hexaHis-tagged ADC(K9BCN) was overexpressed in *E. coli* MG1655 ΔDZ DE3 cells harbouring pCDF-PylT-ADCK9X and BCN-PylRS plasmids.

Protein Overexpression of His_6–ADC(K9BCN) in LB Media

1 L of LB media was inoculated with 1 mL of each antibiotic stock (spectinomycin and ampicillin) followed by 3 mL of starter culture. The culture was incubated at 37 °C with shaking at 200 rpm and monitored by taking aliquots and measuring the optical density (OD_{600}). When an OD_{600} between 0.4-0.6 was reached (OD_{600} = 0.586 after ca. 4 h), expression was induced with IPTG (1 M, 1 mL) and *BCN-lysine* **128** (100 μM, 32 mg) was added. The culture was then incubated at 37 °C with shaking (200 rpm) for 16 h and the cells harvested by centrifugation at 6000 rpm at 4 °C for 10 minutes and, after decanting the supernatant, stored at -80 °C prior to lysis for protein purification or used immediately.

Cell Lysis

Pelleted cells were resuspended in *Lysis Buffer 1* or *2* for His_6–Grb2SH2 and His_6–ADC(K9BCN) respectively (50 – 100 mL per litre of cells) over ice and lysed using either a constant cell disrupter (20 psi, 5 mL injections), or by sonication, keeping the sample on ice (3 × 5 min at 40% power, 40% cycle). Before lysis of His_6–ADC(K9BCN), protease inhibitor cocktail was added. DNase was added to the cell lysate, which was clarified by centrifugation at 30,000*g* for 45 min (or 48,000*g* for 20 min) to collect the cellular debris. The supernatant was decanted and stored over ice for no more than 4 h before purification.

Inclusion Body Lysis of His_6–Grb2SH2

The cellular debris obtained from cell lysis was resuspended in *Inclusion Body Buffer 1* (50 mL) and stirred for 30 minutes at rt. The mixture was centrifuged (17,000*g*, 20 min) and the supernatant decanted. The insoluble fraction was resuspended in *Inclusion Body Buffer 2* (50 mL), stirred at room temperature for 20 min, centrifuged (17,000*g*, 20 min) and the supernatant again decanted. This was repeated a further 2 times using *Inclusion Body Buffer*

[iii] His_6-Grb2-SH2, MBP-YdiA and His_6-ADC(S25A) were overexpressed by Dr Tom McAllister, Ieva Drulyte and Diana Monteiro respectively. MBP-YdiA and His_6-ADC(S25A) cell pellets were lysed and purified by Ieva Drulyte and Zoe Arnott respectively.

2 and the insoluble fraction was solubilised into *Stock Buffer 1* + 8M urea (50-80 mL) by stirring at room temperature overnight.

Purification of hexaHis-tagged Proteins by Ni-NTA chromatography

Ni-NTA agarose resin (10 mL, stored in 20% EtOH at 4 °C) was washed under gravity (in this instance 1 CV is 10 mL) with:

- 10 CV of H_2O
- 10 CV of *Lysis Buffer 1* or *2.*
- The supernatant while collecting the flow through
- 5 CV of *Lysis Buffer 1* or *2*
- 5 CV of *Wash Buffer 1* or *2* while collecting the flow through
- 5 CV of *Wash Buffer 1a* (for His_6-Grb2SH2) while collecting the flow through
- 3 CV of *His Elution Buffer 1* or *2* while collecting 6 × 5 mL fractions
- 10 CV of *Lysis Buffer 1* or *2*

The presence of protein in the fractions was determined by SDS-PAGE and the elution fractions containing the protein combined and concentrated using a spin concentrator, to ca. 1 mL.

Colum Regeneration

After every 4 uses of the column and each time a different protein was purified the column was stripped and regenerated using the following procedure:

- 5 CV 0.1% SDS
- 5 CV 0.1M NaOH
- 5 CV 10 mM EDTA
- 20 CV H_2O
- 2 CV 500 mM $NiSO_4$
- 20 CV H_2O

Size Exclusion Chromatography

Following affinity purification, the concentrated protein was further purified by size exclusion chromatography using an Äkta Purifier FPLC from GE Healthcare with a 20/60 Superdex 75. The column was pre-eq.uilibrated with 1 CV *SEC buffer* before loading the protein sample and eluting with 1 CV of *SEC buffer* while collecting 5-10 mL fractions. Fractions containing protein were identified by their UV absorbance at 280 nm and analysed by SDS-PAGE.

Protein Concentration

Protein concentration was determined by loading 1.5 μL of protein sample onto the Nandrop and measuring the absorbance at 280 nm in triplicate. Protein concentration (μM) was then calculated by application of the Beer-Lambert equation: $A = \varepsilon \times C \times l$, where ε = molar extinction coefficient (M^{-1}, cm^{-1}) calculated from the amino acid seq.uences of each protein using the ProtParam online platform; C = concentration (M) and l = path length (cm).

SDS-PAGE analysis

Sample Preparation

Protein samples were mixed in a 1:1 ratio with 2 × *Loading Buffer* (typically 15 μM of each) and heated at 95 °C (100 °C for ADC) for 5 min (10 min for ADC) and allowed to cool before being loaded onto the gel – 10, 15, 20 μL for 0.75, 1.0 and 1.5 mm gels respectively. A molecular weight marker (10 μL) was also used in one lane. Unboiled protein samples were prepared immediately before loading onto the gel in order to minimize protein denaturation.

Tris-glycine SDS-PAGE

12% or 15% resolving gels were prepared according to Table 1; recipes shown produces enough to case two gels of any thickness. All components (except APS and TEMED) were mixed thoroughly by inversion. Prior to casting, APS was added followed by TEMED and the solution was mixed carefully between additions. The resultant mixture was poured so as to fill ca. 80% of each mould and isopropanol was added to fill the remaining space. The separating gel was prepared in the same manner as the resolving gel (with APS and TEMED added last). Once the resolving gel had set (ca. 20 min), isopropanol was blotted off and the stacking gel carefully poured on top of the resolving gel. Immediately after this, a comb of the desired amount of lanes was place in the mould and the gel was left to set (ca. 20 min). The comb was then removed and the gel used immediately or wrapped in a damp paper towel and stored at 4 °C.

The set gel was incorporated into a running vat and the inner vessel filled with 1 × *SDS Running Buffer* (made by 1 in 5 dilution of the 5 × concentrated stock). The gel was loaded with the desired protein samples and the outer vat filled up to the relevant mark on the side with 1 × *SDS Running Buffer* and the gel was ran at 180 V (120 V for 15% gels).

	Stacking Gel	Resolving Gel - 12%	Resolving Gel - 15%
H_2O (mL)	3.112	4.885	3.385
0.5 M Tris pH 6.8 (mL)	0.945	-	-
1.5 M Tris pH 8.8 (mL)	-	3.8	3.8
30% Acrylamide (mL)	0.833	6	7.5
10% SDS (mL)	0.05	0.15	0.15
15% APS (mL)	0.05	0.15	0.15
TEMED (mL)	0.01	0.015	0.015
Total (mL)	5	15	15

Table 1: Component quantities necessary to make two *Tris-glycine* SDS-PAGE gels

Tris-tricine SDS-PAGE

The 10% resolving and stacking gels were prepared according to Table 2. All components except for APS and TEMED were mixed thoroughly by vortexing and the solution was allowed to set for 20 minutes prior to casting. The stacking and resolving solutions were loaded onto the mould, and protein samples loaded in the same manner as described in for loading a Tris-glycine SDS-PAGE gel. *SDS-PAGE Tris-tricine Cathode Buffer* and *SDS-PAGE Tris-tricine Anode Buffer* were added to the outer and inner vessels respectively and the gel was ran at a constant current of 35 mA.

	Stacking Gel	Resolving Gel - 10%
H_2O (mL)	8.04	4.14
Glycerol (mL)	-	2
Tris-tricine gel buffer (mL)	3.10	5.00
40% Acrylamide (mL)	1.25	3.75
20% APS (mL)	0.1	0.1
TEMED (mL)	0.01	0.01
Total (mL)	12.6	15

Table 2: Component quantities necessary to make two *Tris-tricine* SDS-PAGE gels

Visualisation of SDS-PAGE gels

Gels were stained with enough of either *Coomassie stain* or Instant Blue to cover the gel. Gels stained with Instant Blue were incubated at rt with rocking for 1-12 h, then in H_2O for an hour before being photographed. Gels stained with *Coomassie blue stain* were warmed in the microwave (15 s, full 750 W), incubated for 1.5 h, then in *Coomassie destain* for 3 h (both at rt with rocking), and washed with H_2O before being photographed. Gels containing fluorescent probes were visualised by UV light and photographed prior to staining.

Chapter 7 References

1 C. T. Walsh, S. Garneau-Tsodikova and G. J. Gatto, *Angew. Chem. Int. Ed.*, 2005, **44**, 7342–7372.
2 D. Dougherty, *Curr. Opin. Chem. Biol.*, 2000, **4**, 645–652.
3 L. Davis and J. W. Chin, *Nat. Rev. Mol. Cell Biol.*, 2012, **13**, 168–82.
4 P. Cohen, *Trends Biochem. Sci.*, 2000, **25**, 596–601.
5 M. B. Yaffe and S. J. Smerdon, *Structure*, 2001, **9**, R33–R38.
6 E. Y. Skolnik, C. H. Lee, A. Batzer, L. M. Vicentini, M. Zhou, R. Daly, M. J. Myers, J. M. Backer, A. Ullrich and M. F. White, *EMBO J.*, 1993, **12**, 1929–36.
7 B. Duclos, S. Marcandier and A. J. Cozzone, *Protein Phosphorylation Part B: Analysis of Protein Phosphorylation, Protein Kinase Inhibitors, and Protein Phosphatases*, Elsevier, 1991, vol. 201.
8 J. Cieśla, T. Frączyk and W. Rode, *Acta Biochim. Pol.*, 2011, **58**, 137–48.
9 J. Fuhrmann, A. Schmidt, S. Spiess, A. Lehner, K. Turgay, K. Mechtler, E. Charpentier and T. Clausen, *Science*, 2009, **324**, 1323–7.
10 J. Puttick, E. N. Baker and L. T. J. Delbaere, *Biochim. Biophys. Acta*, 2008, **1784**, 100–5.
11 N. D. Meadow, D. K. Fox and S. Roseman, *Annu. Rev. Biochem.*, 1990, **59**, 497–542.
12 P. V Attwood, P. G. Besant and M. J. Piggott, *Amino Acids*, 2011, **40**, 1035–51.
13 H. Cho, *Curr. Opin. Struct. Biol.*, 2001, **11**, 679–684.
14 D. E. Hultquist, *Biochim. Biophys. Acta - Bioenerg.*, 1968, **153**, 329–340.
15 J.-M. Kee and T. W. Muir, *ACS Chem. Biol.*, 2012, **7**, 44–51.
16 Y.-F. Wei and H. R. Matthews, *Protein Phosphorylation Part A: Protein Kinases: Assays, Purification, Antibodies, Functional Analysis, Cloning, and Expression*, Elsevier, 1991, vol. 200.
17 S. Napper, J. Kindrachuk, D. J. H. Olson, S. J. Ambrose, C. Dereniwsky and A. R. S. Ross, *Anal. Chem.*, 2003, **75**, 1741–1747.
18 A. Teplyakov, K. Lim, P.-P. Zhu, G. Kapadia, C. C. H. Chen, J. Schwartz, A. Howard, P. T. Reddy, A. Peterkofsky and O. Herzberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 16218–23.
19 K. F. Medzihradszky, N. J. Phillipps, L. Senderowicz, P. Wang and C. W. Turck, *Protein Sci.*, 1997, **6**, 1405–11.
20 W. Kavanaugh, C. Turck and L. Williams, *Science*, 1995, **268**, 1177–1179.
21 W. J. Fantl, J. A. Escobedo, G. A. Martin, C. W. Turck, M. del Rosario, F. McCormick and L. T. Williams, *Cell*, 1992, **69**, 413–423.
22 J.-M. Kee, B. Villani, L. R. Carpenter and T. W. Muir, *J. Am. Chem. Soc.*, 2010, **132**, 14327–9.

23 T. E. McAllister, M. G. Nix and M. E. Webb, *Chem. Commun. (Camb).*, 2011, **47**, 1297–9.
24 S. Mukai, G. R. Flematti, L. T. Byrne, P. G. Besant, P. V Attwood and M. J. Piggott, *Amino Acids*, 2012, **43**, 857–74.
25 T. E. McAllister and M. E. Webb, *Org. Biomol. Chem.*, 2012, **10**, 4043–9.
26 S. R. Fuhs, J. Meisenhelder, A. Aslanian, L. Ma, A. Zagorska, M. Stankova, A. Binnie, F. Al-Obeidi, J. Mauger, G. Lemke, J. R. Yates and T. Hunter, *Cell*, 2015, **162**, 198–210.
27 T. E. McAllister, J. J. Hollins and M. E. Webb, *Biochem. Soc. Trans.*, 2013, **41**, 1072–7.
28 K. H. Pesis, Y. Wei, M. Lewis and H. R. Matthews, *FEBS Lett.*, 1988, **239**, 151–154.
29 D. R. Stover, J. Caldwell, J. Marto, K. Root, J. Mestan, M. Stumm, O. Ornatsky, C. Orsi, N. Radosevic, L. Liao, D. Fabbro and M. F. Moran, *Clin. Proteomics*, 2004, **1**, 069–080.
30 R. C. Stewart, *Curr. Opin. Microbiol.*, 2010, **13**, 133–41.
31 P. W. Postma, J. W. Lengeler and G. R. Jacobson, *Microbiol. Mol. Biol. Rev.*, 1993, **57**, 543–594.
32 C. S. Bond, M. F. White and W. N. Hunter, *J. Biol. Chem.*, 2001, **276**, 3247–53.
33 J. M. Fujitaki, G. Fung, E. Y. Oh and R. A. Smith, *Biochemistry*, 1981, **20**, 3658–3664.
34 T. Wieland, B. Nurnberg, I. Ulibarri, S. Kaldenberg-Stasch, G. Schultz and K. Jakobs, *J. Biol. Chem.*, 1993, **268**, 18111–18118.
35 S. Morera, M. Chiadmi, G. LeBras, I. Lascu and J. Janin, *Biochemistry*, 1995, **34**, 11062–11070.
36 M. E. Fraser, M. N. James, W. A. Bridger and W. T. Wolodko, *J. Mol. Biol.*, 2000, **299**, 1325–39.
37 C. S. Crovello, B. C. Furie and B. Furie, *Cell*, 1995, **82**, 279–286.
38 R. Muimo, Z. Hornickova, C. E. Riemen, V. Gerke, H. Matthews and A. Mehta, *J. Biol. Chem.*, 2000, **275**, 36632–6.
39 N. J. Brown, K. Parsley and J. M. Hibberd, *Trends Plant Sci.*, 2005, **10**, 215–21.
40 O. Herzberg, C. C. Chen, G. Kapadia, M. McGuire, L. J. Carroll, S. J. Noh and D. Dunaway-Mariano, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 2652–7.
41 C. J. Chastain and R. Chollet, *Plant Physiol. Biochem.*, 2003, **41**, 523–532.
42 J. N. Burnell and M. D. Hatch, *Biochem. Biophys. Res. Commun.*, 1984, **118**, 65–72.
43 E. M. Sletten and C. R. Bertozzi, *Angew. Chem. Int. Ed.*, 2009, **48**, 6974–6998.
44 K. Lang and J. W. Chin, *Chem. Rev.*, 2014, **114**, 4764–806.
45 E. G. Sander and W. P. Jencks, *J. Am. Chem. Soc.*, 1968, **90**, 6154–6162.
46 L. K. Mahal, *Science*, 1997, **276**, 1125–1128.
47 R. Sadamoto, K. Niikura, T. Ueda, K. Monde, N. Fukuhara and S.-I. Nishimura, *J. Am. Chem. Soc.*, 2004, **126**, 3755–61.
48 D. Rideout, *Science*, 1986, **233**, 561–563.
49 Y. Zeng, T. N. C. Ramya, A. Dirksen, P. E. Dawson and J. C. Paulson, *Nat. Methods*, 2009, **6**, 207–209.
50 J. Rayo, N. Amara, P. Krief and M. M. Meijler, *J. Am. Chem. Soc.*, 2011, **133**, 7469–7475.
51 E. Saxon, *Science*, 2000, **287**, 2007–2010.
52 D. J. Vocadlo, H. C. Hang, E.-J. Kim, J. A. Hanover and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9116–21.
53 E. Saxon, S. J. Luchansky, H. C. Hang, C. Yu, S. C. Lee and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2002, **124**, 14893–14902.
54 F. L. Lin, H. M. Hoyt, H. van Halbeek, R. G. Bergman and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2005, **127**, 2686–2695.
55 E. Saxon, J. I. Armstrong and C. R. Bertozzi, *Org. Lett.*, 2000, **2**, 2141–2143.
56 R. Huisgen, *Angew. Chem. Int. Ed.*, 1963, **75**, 604–+.
57 V. V Rostovtsev, L. G. Green, V. V Fokin and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2002, **41**, 2596–+.

58 C. W. Tornoe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057–3064.

59 F. Wolbers, P. ter Braak, S. Le Gac, R. Luttge, H. Andersson, I. Vermes and A. van den Berg, *Electrophoresis*, 2006, **27**, 5073–80.

60 N. J. Agard, J. A. Prescher and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2004, **126**, 15046–15047.

61 J. A. Codelli, J. M. Baskin, N. J. Agard and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2008, **130**, 11486–93.

62 C. G. Gordon, J. L. Mackey, J. C. Jewett, E. M. Sletten, K. N. Houk and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2012, **134**, 9199–208.

63 J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16793–16797.

64 S. T. Laughlin, J. M. Baskin, S. L. Amacher and C. R. Bertozzi, *Science*, 2008, **320**, 664–667.

65 W. Song, Y. Wang, J. Qu, M. M. Madden and Q. Lin, *Angew. Chem. Int. Ed.*, 2008, **47**, 2832–2835.

66 Y. Wang, W. Song, W. J. Hu and Q. Lin, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 5330–3.

67 H. Ren, F. Xiao, K. Zhan, Y.-P. Kim, H. Xie, Z. Xia and J. Rao, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 9658–62.

68 D. P. Nguyen, T. Elliott, M. Holt, T. W. Muir and J. W. Chin, *J. Am. Chem. Soc.*, 2011, **133**, 11418–11421.

69 S. S. van Berkel, A. T. J. Dirks, S. A. Meeuwissen, D. L. L. Pingen, O. C. Boerman, P. Laverman, F. L. van Delft, J. J. L. M. Cornelissen and F. P. J. T. Rutjes, *Chembiochem*, 2008, **9**, 1805–15.

70 S. S. van Berkel, A. T. J. Dirks, M. F. Debets, F. L. van Delft, J. J. L. M. Cornelissen, R. J. M. Nolte and F. P. J. T. Rutjes, *Chembiochem*, 2007, **8**, 1504–1508.

71 X. Ning, R. P. Temming, J. Dommerholt, J. Guo, D. B. Ania, M. F. Debets, M. A. Wolfert, G.-J. Boons and F. L. van Delft, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 3065–8.

72 C. S. McKay, J. A. Blake, J. Cheng, D. C. Danielson and J. P. Pezacki, *Chem. Commun. (Camb).*, 2011, **47**, 10040–2.

73 E. M. Sletten and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2011, **133**, 17570–3.

74 H. Stöckmann, A. A. Neves, S. Stairs, K. M. Brindle and F. J. Leeper, *Org. Biomol. Chem.*, 2011, **9**, 7303–5.

75 S. Stairs, A. A. Neves, H. Stöckmann, Y. A. Wainman, H. Ireland-Zecchini, K. M. Brindle and F. J. Leeper, *Chembiochem*, 2013, **14**, 1063–7.

76 Y. A. Lin, J. M. Chalker, N. Floyd, G. J. L. Bernardes and B. G. Davis, *J. Am. Chem. Soc.*, 2008, **130**, 9642–3.

77 R. A. Carboni and R. V. Lindsey, *J. Am. Chem. Soc.*, 1959, **81**, 4342–4346.

78 T. Koike, Y. Hoashi, T. Takai and O. Uchikawa, *Tetrahedron Lett.*, 2011, **52**, 3009–3011.

79 E. R. Bilbao, M. Alvarado, C. F. Masaguer and E. Raviña, *Tetrahedron Lett.*, 2002, **43**, 3551–3554.

80 A. M. Palmer, B. Grobbel, C. Brehm, P. J. Zimmermann, W. Buhr, M. P. Feth, H. C. Holst and W. A. Simon, *Bioorg. Med. Chem.*, 2007, **15**, 7647–60.

81 D. L. Boger and J. Hong, *J. Am. Chem. Soc.*, 2001, **123**, 8515–8519.

82 N. K. Devaraj, R. Weissleder and S. A. Hilderbrand, *Bioconjug. Chem.*, 2008, **19**, 2297–2299.

83 M. L. Blackman, M. Royzen and J. M. Fox, *J. Am. Chem. Soc.*, 2008, **130**, 13518–9.

84 A.-C. Knall and C. Slugovc, *Chem. Soc. Rev.*, 2013, **42**, 5131–42.

85 K. Lang, L. Davis, J. Torres-Kolbus, C. Chou, A. Deiters and J. W. Chin, *Nat. Chem.*, 2012, **4**, 298–304.

86 T. S. Elliott, F. M. Townsley, A. Bianco, R. J. Ernst, A. Sachdeva, S. J. Elsässer, L. Davis, K. Lang, R. Pisa, S. Greiss, K. S. Lilley and J. W. Chin, *Nat. Biotechnol.*, 2014, **32**, 465–72.

87 K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox and J. W. Chin, *J. Am. Chem. Soc.*, 2012, **134**, 10317–20.

88 N. K. Devaraj, S. Hilderbrand, R. Upadhyay, R. Mazitschek and R. Weissleder, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 2869–72.

89 R. Rossin, P. R. Verkerk, S. M. van den Bosch, R. C. M. Vulders, I. Verel, J. Lub and M. S. Robillard, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 3375–8.

90 N. K. Devaraj, R. Upadhyay, J. B. Haun, S. A. Hilderbrand and R. Weissleder, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 7013–6.

91 J. B. Haun, N. K. Devaraj, S. A. Hilderbrand, H. Lee and R. Weissleder, *Nat. Nanotechnol.*, 2010, **5**, 660–5.

92 G. Clavier and P. Audebert, *Chem. Rev.*, 2010, **110**, 3299–314.

93 M. R. Karver, R. Weissleder and S. A. Hilderbrand, *Bioconjug. Chem.*, 2011, **22**, 2263–70.

94 D. S. Liu, A. Tangpeerachaikul, R. Selvaraj, M. T. Taylor, J. M. Fox and A. Y. Ting, *J. Am. Chem. Soc.*, 2012, **134**, 792–795.

95 N. O. Abdel, M. A. Kira and M. N. Tolba, *Tetrahedron Lett.*, 1968, **9**, 3871–3872.

96 W. Chen, D. Wang, C. Dai, D. Hamelberg and B. Wang, *Chem. Commun. (Camb).*, 2012, **48**, 1736–8.

97 P. Audebert, S. Sadki, F. Miomandre, G. Clavier, M. Claude Verniéres, M. Saoud and P. Hapiot, *New J. Chem.*, 2004, **28**, 387.

98 J. Yang, M. R. Karver, W. Li, S. Sahu and N. K. Devaraj, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 5222–5.

99 D. L. Boger, R. S. Coleman, J. S. Panek, F. X. Huber and J. Sauer, *J. Org. Chem.*, 1985, **50**, 5377–5379.

100 D. L. Boger, R. P. Schaum and R. M. Garbaccio, *J. Org. Chem.*, 1998, **63**, 6329–6337.

101 M. D. Coburn, G. A. Buntain, B. W. Harris, M. A. Hiskey, K.-Y. Lee and D. G. Ott, *J. Heterocycl. Chem.*, 1991, **28**, 2049–2050.

102 D. E. Chavez and M. A. Hiskey, *J. Heterocycl. Chem.*, 1998, **35**, 1329–1332.

103 D. E. Chavez and M. A. Hiskey, *J. Energ. Mater.*, 1999, **17**, 357–377.

104 D. L. Boger and S. M. Sakya, *J. Org. Chem.*, 1988, **53**, 1415–1423.

105 Y.-H. Gong, F. Miomandre, R. Méallet-Renault, S. Badré, L. Galmiche, J. Tang, P. Audebert and G. Clavier, *European J. Org. Chem.*, 2009, **2009**, 6121–6128.

106 D. Andrade-Acuña, J. G. Santos, W. Tiznado, Á. Cañete and M. E. Aliaga, *J. Phys. Org. Chem.*, 2014, **27**, 670–675.

107 J. Balcar, G. Chrisam, F. X. Huber and J. Sauer, *Tetrahedron Lett.*, 1983, **24**, 1481–1484.

108 B. R. Lahue, S.-M. Lo, Z.-K. Wan, G. H. C. Woo and J. K. Snyder, *J. Org. Chem.*, 2004, **69**, 7171–82.

109 E. C. Taylor and L. G. French, *J. Org. Chem.*, 1989, **54**, 1245–1249.

110 von Hans Neunhoeffer and H.-W. Frühauf, *Tetrahedron Lett.*, 1969, **10**, 3151–3154.

111 D. L. Boger and J. S. Panek, *J. Org. Chem.*, 1981, **46**, 2179–2182.

112 N. Catozzi, M. G. Edwards, S. A. Raw, P. Wasnaire and R. J. K. Taylor, *J. Org. Chem.*, 2009, **74**, 8343–54.

113 E. Badarau, F. Suzenet, A.-L. Fînaru and G. Guillaumet, *European J. Org. Chem.*, 2009, **2009**, 3619–3627.

114 R. A. A. Foster and M. C. Willis, *Chem. Soc. Rev.*, 2013, **42**, 63–76.

115 B. Shi, W. Lewis, I. B. Campbell and C. J. Moody, *Org. Lett.*, 2009, **11**, 3686–8.

116 V. N. Kozhevnikov, D. N. Kozhevnikov, O. V. Shabunina, V. L. Rusinov and O. N. Chupakhin, *Tetrahedron Lett.*, 2005, **46**, 1791–1793.

117 J. Sauer, D. K. Heldmann, J. Hetzenegger, J. Krauthan, H. Sichert and J. Schuster, *European J. Org. Chem.*, 1998, **1998**, 2885–2896.

118 F. Thalhammer, U. Wallfahrer and J. Sauer, *Tetrahedron Lett.*, 1990, **31**, 6851–6854.

119 Y. Liang, J. L. Mackey, S. A. Lopez, F. Liu and K. N. Houk, *J. Am. Chem. Soc.*, 2012, **134**, 17904–7.
120 D. Wang, W. Chen, Y. Zheng, C. Dai, K. Wang, B. Ke and B. Wang, *Org. Biomol. Chem.*, 2014, **12**, 3950–5.
121 D. Schwarzer and P. A. Cole, *Curr. Opin. Chem. Biol.*, 2005, **9**, 561–9.
122 M. Fernández-Suárez, H. Baruah, L. Martínez-Hernández, K. T. Xie, J. M. Baskin, C. R. Bertozzi and A. Y. Ting, *Nat. Biotechnol.*, 2007, **25**, 1483–7.
123 L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498–500.
124 J. M. Xie and P. G. Schultz, *Curr. Opin. Chem. Biol.*, 2005, **9**, 548–554.
125 S. J. Miyake-Stoner, C. A. Refakis, J. T. Hammill, H. Lusic, J. L. Hazen, A. Deiters and R. A. Mehl, *Biochemistry*, 2010, **49**, 1667–77.
126 T. E. McAllister, University of Leeds, 2012.
127 T. Pawson and J. Schlessinger, *Curr. Biol.*, 1993, **3**, 434–442.
128 J. Schlessinger, *Curr. Opin. Genet. Dev.*, 1994, **4**, 25–30.
129 Z. Songyang, S. E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W. G. Haser, F. King, T. Roberts, S. Ratnofsky and R. J. Lechleider, *Cell*, 1993, **72**, 767–778.
130 C. McNemar, M. E. Snow, W. T. Windsor, A. Prongay, P. Mui, R. Zhang, J. Durkin, H. V Le and P. C. Weber, *Biochemistry*, 1997, **36**, 10006–14.
131 K. Ogura, S. Tsuchiya, H. Terasawa, S. Yuzawa, H. Hatanaka, V. Mandiyan, J. Schlessinger and F. Inagaki, *J. Mol. Biol.*, 1999, **289**, 439–45.
132 F. T. Hofmann, C. Lindemann, H. Salia, P. Adamitzki, J. Karanicolas and F. P. Seebeck, *Chem. Commun. (Camb).*, 2011, **47**, 10335–7.
133 A. P. Benfield, B. B. Whiddon, J. H. Clements and S. F. Martin, *Arch. Biochem. Biophys.*, 2007, **462**, 47–53.
134 D. M. Jameson and S. E. Seifried, *Methods*, 1999, **19**, 222–33.
135 J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer Science & Business Media, Third., 2006.
136 D. Kraskouskaya, E. Duodu, C. C. Arpin and P. T. Gunning, *Chem. Soc. Rev.*, 2013, **42**, 3337–70.
137 L. Senderowicz, J. X. Wang, L. Y. Wang, S. Yoshizawa, W. M. Kavanaugh and C. W. Turck, *Biochemistry*, 1997, **36**, 10538–44.
138 J. N. Burnell, *BMC Biochem.*, 2010, **11**, 1.
139 R. Tolentino, C. Chastain and J. Burnell, *Adv. Biol. Chem.*, 2013, **03**, 12–21.
140 J. J. Hollins, University of Leeds, 2012.
141 U. Jacquemard, V. Bénéteau, M. Lefoix, S. Routier, J.-Y. Mérour and G. Coudert, *Tetrahedron*, 2004, **60**, 10039–10047.
142 J. G. Erickson, *J. Am. Chem. Soc.*, 1952, **74**, 4706–4706.
143 M. A. Glomb and R. Tschirnich, *J. Agric. Food Chem.*, 2001, **49**, 5543–5550.
144 von Hans Neunhoeffer, H. Hennig, H.-W. Frühauf and M. Mutterer, *Tetrahedron Lett.*, 1969, **10**, 3147–3150.
145 J. Xie and P. G. Schultz, *Methods*, 2005, **36**, 227–38.
146 J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang and P. G. Schultz, *Science*, 2003, **301**, 964–7.
147 N. Wu, A. Deiters, T. A. Cropp, D. King and P. G. Schultz, *J. Am. Chem. Soc.*, 2004, **126**, 14306–7.
148 H. Neumann, S. Y. Peak-Chew and J. W. Chin, *Nat. Chem. Biol.*, 2008, **4**, 232–234.
149 S. M. Hancock, R. Uprety, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2010, **132**, 14819–14824.
150 T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto and S. Yokoyama, *Biochem. Biophys. Res. Commun.*, 2008, **371**, 818–822.
151 S. Greiss and J. W. Chin, *J. Am. Chem. Soc.*, 2011, **133**, 14196–9.
152 J. Du, Y. Zhou, X. Su, J. J. Yu, S. Khan, H. Jiang, J. Kim, J. Woo, J. H. Kim, B. H. Choi, B. He, W. Chen, S. Zhang, R. A. Cerione, J. Auwerx, Q. Hao and H. Lin, *Science*, 2011, **334**, 806–809.
153 A. M. Gellett, P. W. Huber and P. J. Higgins, *J. Organomet. Chem.*, 2008, **693**, 2959–2962.

154 D. Felder, M. Gutiérrez Nava, M. del Pilar Carreón, J.-F. Eckert, M. Luccisano, C. Schall, P. Masson, J.-L. Gallani, B. Heinrich, D. Guillon and J.-F. Nierengarten, *Helv. Chim. Acta*, 2002, **85**, 288–319.
155 S. Isomura, P. Wirsching and K. D. Janda, *J. Org. Chem.*, 2001, **66**, 4115–4121.
156 A. Borrmann, S. Milles, T. Plass, J. Dommerholt, J. M. M. Verkade, M. Wiessler, C. Schultz, J. C. M. van Hest, F. L. van Delft and E. A. Lemke, *Chembiochem*, 2012, **13**, 2094–9.
157 F. I. Carroll, S. V Kotturi, H. A. Navarro, S. W. Mascarella, B. P. Gilmour, F. L. Smith, B. H. Gabra and W. L. Dewey, *J. Med. Chem.*, 2007, **50**, 3388–3391.
158 S. Tabanella, I. Valancogne and R. F. W. Jackson, *Org. Biomol. Chem.*, 2003, **1**, 4254–61.
159 R. F. W. Jackson, N. Wishart, A. Wood, K. James and M. J. Wythes, *J. Org. Chem.*, 1992, **57**, 3397–3404.
160 A. Streitwieser, *Chem. Rev.*, 1956, **56**, 571–752.
161 L. C. McCann, H. N. Hunter, J. A. C. Clyburne and M. G. Organ, *Angew. Chem. Int. Ed. Engl.*, 2012, **51**, 7024–7.
162 R. L. E. Furlán, E. G. Mata and O. A. Mascaretti, *J. Chem. Soc. Perkin Trans. 1*, 1998, 355–358.
163 L. Wang, Z. Zhang, A. Brock and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 56–61.
164 G. Han, M. Tamaki and V. J. Hruby, *J. Pept. Res.*, 2001, **58**, 338–341.
165 J. Dommerholt, S. Schmidt, R. Temming, L. J. A. Hendriks, F. P. J. T. Rutjes, J. C. M. van Hest, D. J. Lefeber, P. Friedl and F. L. van Delft, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 9422–5.
166 T. Cruchter, K. Harms and E. Meggers, *Chemistry*, 2013, **19**, 16682–9.
167 J. W. Wijnen, S. Zavarise, J. B. F. N. Engberts and M. Charton, *J. Org. Chem.*, 1996, **61**, 2001–2005.
168 F. L. Lin, H. M. Hoyt, H. van Halbeek, R. G. Bergman and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2005, **127**, 2686–95.
169 D. C. F. Monteiro, V. Patel, C. P. Bartlett, S. Nozaki, T. D. Grant, J. A. Gowdy, G. S. Thompson, A. P. Kalverda, E. H. Snell, H. Niki, A. R. Pearson and M. E. Webb, *Chem. Biol.*, 2015, **22**, 492–503.
170 D. N. Kamber, Y. Liang, R. J. Blizzard, F. Liu, R. A. Mehl, K. N. Houk and J. A. Prescher, *J. Am. Chem. Soc.*, 2015.
171 J. Schlessinger and M. A. Lemmon, *Sci. STKE*, 2003, **2003**, RE12.
172 J.-M. Kee, R. C. Oslund, D. H. Perlman and T. W. Muir, *Nat. Chem. Biol.*, 2013, **9**, 416–21.
173 B. Li, M. Berliner, R. Buzon, C. K.-F. Chiu, S. T. Colgan, T. Kaneko, N. Keene, W. Kissel, T. Le, K. R. Leeman, B. Marquez, R. Morris, L. Newell, S. Wunderwald, M. Witt, J. Weaver, Z. Zhang and Z. Zhang, *J. Org. Chem.*, 2006, **71**, 9045–50.
174 Y. Lau and D. Spring, *Synlett*, 2011, **2011**, 1917–1919.
175 T. E. McAllister, K. A. Horner and M. E. Webb, *Chembiochem*, 2014, **15**, 1088–91.
176 P. Thapa, R.-Y. Zhang, V. Menon and J.-P. Bingham, *Molecules*, 2014, **19**, 14461–83.