Sam Stewart
Department of Pure and Applied Chemistry
Thomas Graham Building
295 Cathedral Street
Glasgow
G1 1XL

GlaxoSmithKline
Medicines Research Centre
Gunnels Wood Road
Stevenage
Herts
SG1 2NY

"Studies towards the synthesis and labelling of bioactive peptides"

Sam Stewart

Thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Department of Pure and Applied Chemistry, University of Strathclyde

Therapeutic Peptide Chemistry Performance Unit, GlaxoSmithKline

April 2016

Date: 11.04.2016

GSK Supervisor: Robert Harris
Academic Supervisor: Craig Jamieson

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: Sam Stewart Date: 11.04.2016

Table of Contents

Acknowledgments

First and foremost, I would like to extend my heartfelt thanks to Dr Robert Harris for his role as my GSK industrial supervisor. Without his constant support, encouragement, and tolerance (especially during difficult periods) this work would not have been possible.

Additional warm thanks to Prof. Harry Kelly and Prof. William Kerr for giving me the opportunity to take part in the GSK/UoS collaborative Industrial PhD programme, and Dr Craig Jamieson for his valuable guidance as my academic supervisor and for giving me the opportunity to work with him over the past few years.

I would also like to thank...

The various technicians, analysts, and administrative staff at GSK and Strathclyde University; in particular Sean Lynn, Nick Taylor, and Bill Leavens for their assistance with NMR, purification, and mass spectrometry, respectively.

The many people at GSK who have not only helped me with my PhD but also with my personal and professional development.

The members of the Therapeutic Peptide CPU, past and present, for their friendship, support, and welcome advice throughout my PhD. Dr Graham Simpson and Dr Praew Thansandote for their constant enthusiasm and mentoring, Andrew Mason for his brilliantly bad jokes, Dr Albert-Isidro Llobet for his peptide expertise and quiet humour, and Gail Fisher for her constant moral support and ability to acquire every contact or piece of equipment imaginable. Thanks also, to every one of them for putting up with my random/scary taste in music!

...and finally, a special thanks to my family and friends, in particular Julie Halvorson, Andrea Jessie Adams, Joella Hope, and Sarah-Louise Sexton for their constant support and unwavering encouragement.

Abbreviations

αHA	α-Hydroxyamide
αKA	α-Ketoamide
AA	Amino Acid
AAC	Alkyne-Azide Cycloaddition
p-ABSA	*para*-Acetamidobenzenesulfonyl Azide
Ac	Acetyl
ADHP	2-Amino-4,6-dihydroxypyrimidine
AIDS	Acquired Immune Deficiency Syndrome
Al	Allyl
Alloc	Allyloxycarbonyl
AMP	Antimicrobial Peptide
App	Apparent
Aq	Aqueous
ATP	Adenosine Triphosphate
Boc	*tert*-Butoxycarbonyl
BODIPY	Boron-dipyrromethene
BOP	Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium Hexafluorophosphate
Bpoc	2-(*p*-Biphenyl-4-yl)prop-2-yloxycarbonyl
br	Broad

BSA	Bovine Serum Albumin
Bn/Bzl	Benzyl
C	Celsius
Cat.	Catalytic
Cbz	Carboxybenzyl
CKD	Chronic Kidney Disease
CLEAR	Cross-Linked Ethoxylate Acrylate Resin
COPD	Chronic Obstructive Pulmonary Disease
COSY	Correlation Spectroscopy
CPP	Cell Permeable Peptide
CPU	Chemistry Performance Unit
Ct	Cyclotheonamide
CV	Column Volumes
d	Doublet
DABCO	1,4-Diazabicyclo[2.2.2]octane
Dap	2,3-Diaminopropionic acid
DAPI	4’,6-diamidino-2-phenylindole
DCC	*N,N*’-Dicyclohexylcarbodiimide
DCM	Dichloromethane
DIPEA	*N,N*-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine

DMF	*N,N*-Dimethylformamide
DMP	Dess-Martin Periodinane
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDG	Electron Donating Group
ELT	Encoded Library Technology
Em	Emission
Ex	Excitation
EPO	Erythropoietin
EPOR	Erythropoietin Receptor
Et	Ethyl
et. al.	*Et Alia*
Equiv.	Equivalents
ESI	Electrospray Ionisation
eV	Electronvolt
EWG	Electron Withdrawing Group
FBS	Fetal Bovine Serum
FDA	Food and Drug Administration
FITC	Fluorescein Isothiocyanate
FMO	Frontier Molecular Orbital

Fmoc	Fluorenylmethoxycarbonyl
FRET	Fluorescence Resonance Energy Transfer
FT-IR	Fourier-Transform Infrared
g	Grams
GRITS	Global Reagent Inventory Tracking System (GSK)
h	Hours
HAG	Homoallylglycine
HATU	1-[*Bis*(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium-3-oxid Hexafluorophosphate
HIV	Human Immunodeficiency Virus
HMBC	Heteronuclear Multiple Bond Correlation
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	1-Hydroxybenzotriazole
HOMO	Highest Occupied Molecular Orbital
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HSQC	Heteronuclear Single Quantum Correlation
HWE	Horner-Wadsworth-Emmons
Hz	Hertz
IED	Inverse Electron Demand
IPA	Isopropyl Alcohol

IR	Infrared
J	Coupling Constant
KLK	Kallikrein
LC-MS	Liquid Chromatography Mass Spectrometry
LUMO	Lowest Unoccupied Molecular Orbital
m	Multiplet
MBHA	(α-Amino-α-*p*-xylyl hydrochloride)polystyrene crosslinked with divinylbenzene
MDAP	Mass Directed Automated Preparation
Me	Methyl
MeBMT	(4*R*)-4-((*E*)-Butenyl)-4-*N*-dimethyl-L-threonine
2-MeTHF	2-Methyltetrahydrofuran
min	Minutes
mM/μM	Millimolar/Micromolar
mmol	Millimole
Mtr	4-Methoxy-2,3,6-trimethylbenzenesulfonyl
Mts	Mesitylene-2-sulfonyl
m/z	Mass-to-Charge Ratio
Nal	L-1-Naphthylalanine
nm	Nanometre
NMM	*N*-Methyl Morpholine

NMR	Nuclear Magnetic Resonance
NS3	Non-structural Protein 3
Orn	Ornithine
P^X	Protecting Group (where X = 1, 2, 3...*etc*)
Pac	Phenacyl
PADAM	Passerini-amine deprotection-acyl migration
PAMPA	Parallel Artificial Membrane Permeability Assay
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PBS	Phosphate Buffered Saline
PEG	Polyethylene Glycol
PEGA	Poly[acryloyl-*bis*(aminopropyl)ethylene glycol]
Ph	Phenyl
Pht	Phthalimide
PPI	Protein-Protein Interaction
ppm	Parts Per Million
psi	Pounds Per Square Inch
PyBOP	(Benzotriazol-1-yloxy)tripyrrolidinophosphonium Hexafluorophosphate
q	Quartet
RNA	Ribonucleic Acid
rpm	Revolutions Per Minute

rt	Room Temperature
U-Pa	Urokinase-type Plasminogen Activator
s	Singlet
Sar	Sarcosine
SAR	Structure Activity Relationship
Sat.	Saturated
sec	Seconds
SEM	[2-(Trimethylsilyl)ethoxy]methyl Acetal
sIBX	Stabilised 2-Iodoxybenzoic Acid
SPPS	Solid Phase Peptide Synthesis
STFI-1	Sunflower Trypsin-1
Su	Succinimide
t	Triplet
TBAB	Tetrabutylammonium Bromide
TBAF	Tetrabutylammonium Fluoride
TBDMS/TBS	*tert*-Butyldimethylsilyl Ether
TBME	*tert*-Butyl Methyl Ether
TBTU	*O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium Tetrafluoroborate
Teoc	Trimethylsilylethoxycarbonyl
TFA	Trifluoroacetic Acid

TFE	Trifluoroethanol
THP	Tetrahydropyran
TIPS	Triisopropylsilyl ether
TLC	Thin Layer Chromatography
TLSP	Trypsin-Like Serine Protease
TM	Trademark
TMS	Trimethylsilyl
TPPO	Triphenylphosphine Oxide
Trx	Thioredoxin
Ts	Tosyl
UV	Ultraviolet
vTyr	Vinyl Tyrosine
v/v	Volume/Volume Ratio

Amino acid abbreviations:

Amino Acid	Abbreviation	Single Letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Abstract

The photochemical synthesis of fluorescent pyrazoline adducts was investigated as a potential methodology for intracellular visualisation of cell permeable peptides. Alkene-tagged amino acids were reacted with 2,5-diaryl tetrazoles. These small tags were chosen as they would minimise alterations to the physicochemical properties when incorporated into peptides. The resulting pyrazoline adducts fluoresced at concentrations suitable for exploitation in cells; however, the addition of *N*-acetyl cysteine, a glutathione substitute, resulted in additional side reactions. A number of competition reactions and cell-based trials were conducted to determine the cause, and the methodology was determined to be unsuitable for intracellular peptide visualisation.

From this work several interesting side reactions were identified and were investigated further. The photochemical irradiation of lone 2,5-diaryl tetrazoles was shown to provide access to a number of azoles including: 2,4,5-trisubstituted 1,2,3-triazoles; 1,3,5-pyrazolines; 1,3,5-pyrazoles; and 3,5-indazoles. Irradiation in the presence of carboxylic acids was shown to yield hydrazides. This latter reaction was determined to be selective over both alkene and alkyne labelling and could represent a novel means of derivatising functionalised carboxylic acids.

Kallikrein 5 (KLK5) is a Trypsin-Like Serine Protease (TLSP) which is involved in the development of skin diseases such as atopic dermatitis. The cyclotheonamide (Ct) family of natural products has shown inhibitory activity against analogous TLSPs. Using previous learnings from literature syntheses of these macrocyclic peptides a flexible and efficient solid phase peptide synthesis (SPPS) has been developed. A number of protecting group strategies were investigated and the solid phase assembly of macrocyclic cyclotheonamide precursors has successfully been performed to enable SAR analysis with Trypsin-Like Serine Proteases (TLSPs) such as KLK5.

Chapter 1

Synthesis of fluorescent pyrazoline derivatives for the identification and subcellular localisation of cell permeable peptides

1.1. Introduction

1.1.1. Therapeutic peptides as alternatives to small molecules

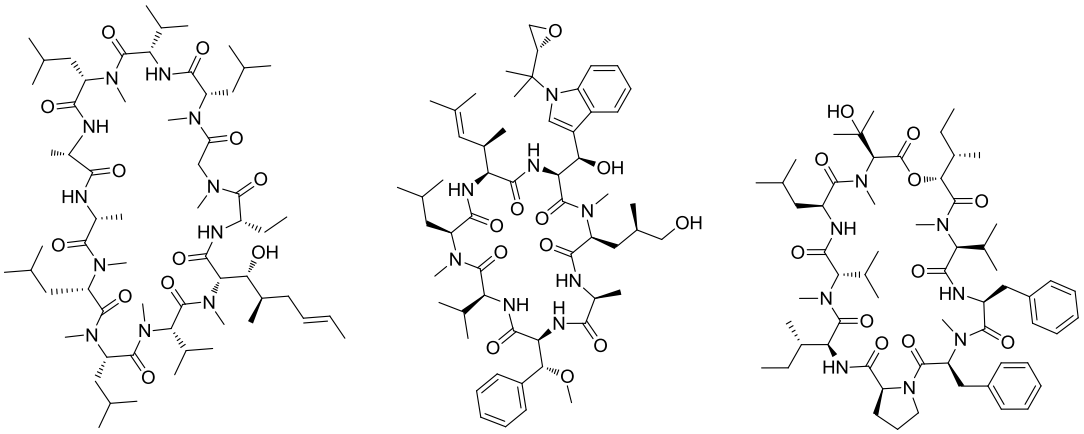
Verdine and Hilinski state that small molecules and biologics, the two major classes of pharmaceuticals, are collectively able to target only 20% of human proteins.[1] Biologics, such as recombinant proteins and antibodies, are effective against the 10% of protein-protein interactions that are extracellular in nature; however, in general, they are impermeable to cell membranes. Small molecules are able to target both intracellular and extracellular proteins, but only 10% of these proteins have suitable binding pockets.[1] Currently, intracellular therapeutic peptide drug molecules make up a very small proportion of marketed drug molecules.[2] Therefore, the design, synthesis, and progression of drug-like cell permeable peptides (CPPs) for the currently 'undruggable' targets is becoming increasingly prevalent.[3]

Peptides were traditionally avoided for therapeutic use because of expensive syntheses, poor oral bioavailability, limited tissue penetration, and vulnerability to protease degradation.[4] More recently, the efficiency of peptide synthesis has increased, particularly since the introduction of automated solid-phase synthesis. This can be carried out at room temperature or under microwave conditions which has enabled industrial scale synthesis with simplified purification, resulting in greater overall purity at each step.[5] Furthermore, efforts have been made to develop peptides with improved pharmacokinetic and pharmacodynamic properties.[6] During this period a number of natural peptides were determined to have therapeutic properties and have been developed as pharmaceuticals.

One of the most noted and first FDA approved peptide drugs was cyclosporin A **1** (Figure 1). Initially marketed as Sandimune™, and later as Neoral™,[7,8] this natural cyclic peptide consists of 11 amino acids, including the non-proteinogenic MeBMT residue and multiple *N*-methylated residues, and is isolated from *Beauveria nivea*. The MeBMT residue, containing a 1,2-*trans*-substituted alkene is important in the role of the peptide as an immunological suppressant which is able to inhibit the

activation of T-cells within the body. As a result it has been successful in preventing organ rejection after transplants.

Additional cyclic peptides have been identified from natural sources, many of which also contain non-proteinogenic amino acid residues. For example, Cyclomarin A **2** is involved in immunomodulation and is marketed as an anti-inflammatory drug with topical applications.[9] This cyclic heptapeptide contains four non-proteinogenic residues, including an epoxide, and can be obtained as a metabolite from a marine bacterium. A further example is Aureobasidin A **3**, a depsipeptide in which one of the amide bonds in the backbone is replaced with an ester. This cyclic peptide can be used in the prevention of fungal growth and has also been investigated for inhibiting the growth of parasites associated with AIDS.[10]



Cyclosporin A
1

Cyclomarin A
2

Aureobasidin A
3

Figure 1: Structure of natural therapeutic cyclic peptides containing non-proteinogenic residues.

Non-cyclic natural peptides also show potential to act as therapeutic agents. For example a large and varied class of antimicrobial peptides (AMPs) is known to be present in bacteria, fungi, animals, and plants.[11] These provide protection from both Gram-positive and Gram-negative pathogens, and have been investigated for their anti-cancer and anti-viral properties. Perhaps the most commonly administered non-cyclic natural peptide therapeutic is insulin.[3] Insulin is a 51 residue peptide hormone composed of two polypeptide chains, A and B. It exists in equilibrium

between a zinc-chelated hexameric form, suited for long term storage, and the active monomeric form (Figure 2).

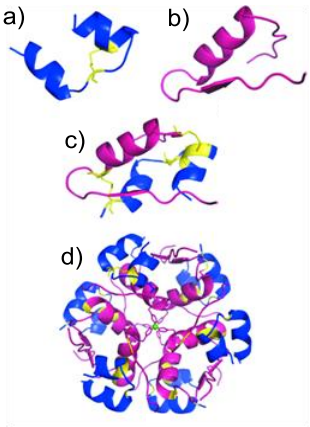


Figure 2: Structure of (a) the A chain; (b) the B chain; (c) monomeric bovine insulin; (d) hexameric bovine insulin.[12]

Lispro™, a short-acting analogue of insulin, was developed by swapping the positions of the penultimate lysine and proline residues on the B chain.[13] This two-point change prevents the aggregation of the insulin monomers, enabling faster activation, and was found to provide improved control of glucose levels in type 2 diabetic patients.[14] Importantly, a small molecule alternative to insulin has not been found, emphasising the fact that peptides represent a powerful class of drug molecules.

Despite the limitations already outlined, peptide therapeutics offer some significant advantages over the more traditional small molecule approach.[15] The amino acids commonly used for peptide synthesis are natural building blocks, and as a result are unlikely to cause toxicological issues; though it is worth noting that non-proteinogenic amino acids are becoming increasingly utilised and some may introduce toxic properties to potential drug candidates.[16] Peptides also have high specificity for their targets due to their typically large binding surfaces which results in lower promiscuity compared to small molecules.

Perhaps most importantly, peptides have the potential for the inhibition of protein-protein interactions (PPIs) in targets which do not have a suitable binding site for small molecules.[1] To date, the most successful small molecule inhibitors of PPIs are those that are able to interact with the large surface area of the proteins, either within the pocket or on the surface itself. These molecules tend to fall outside the typical ideal properties, such as low molecular weight and lipophilicity, as they often need to be greater than 500 Da to interact with the large hydrophobic pockets.[17] Molecules that fall outside of the ideal properties for an oral drug candidate are often associated with attrition during drug development.[18] While peptides also fall outside of these ideal properties, they represent an opportunity to target the PPI interactions with their increased surface area relative to small molecules.

As such, a number of fully synthetic peptides have been developed for therapeutic use. For example, Peginesatide **4** (Figure 3) was approved in 2012 and marketed as Omontys™ for use in the treatment of anaemia in patients suffering from chronic kidney disease (CKD).[19] This dimeric peptide consists of two identical 21 amino acid chains and a large polyethylene glycol (PEG) chain to aid solubility and reduce degradation. The peptide is an erythropoietin (EPO) mimic which binds to the EPO receptor (EPOR) and induces red blood cell differentiation.

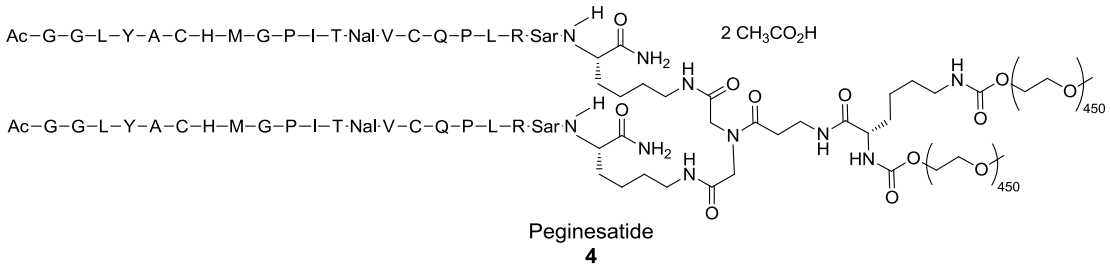


Figure 3: Structure of synthetic peptide Peginesatide **4**.

In 2003 the release of Enfuvirtide **5**, marketed as Fuzeon™, represented a second example of a successful therapeutic peptide programme where small molecules had not yet succeeded (Figure 4).[20] Enfuvirtide is the first of a new class of HIV fusion inhibitors which inhibit late stage fusion to uninfected cells by mimicking, and

subsequently displacing, the fusion apparatus required for viral penetration of the cell membrane. It consists of 36 proteinogenic amino acids with an acetylated *C*-terminus and has been validated for use as an anti-HIV-1 therapy.

Ac-Y-T-S-L-I-H-S-L-I-E-E-S-Q-N-Q-Q-E-K-N-E-Q-E-L-L-E-L-D-K-W-A-S-L-W-N-W-F-NH_2

Enfuvirtide

5

Figure 4: Structure of HIV fusion inhibitor Enfuvirtide **5**.

A common problem with peptide therapeutics is the expensive synthesis resulting from either slow and laborious solution phase syntheses or from the high costs associated with solid phase peptide synthesis. Initially, this resulted in Enfuvirtide being reserved as a last resort combination treatment. However, Enfuvirtide represents an interesting case study in the scalability of peptide therapeutics. By combining the high efficiency of SPPS for assembly of short chain peptides with solution phase manipulations of the final product the synthesis of Enfuvirtide has been achieved on 'multi-tonne per year' scales.[21] This process involves assembly of Ac-AA_1-AA_{16}-OH, Fmoc-AA_{17}-AA_{26}-OH and Fmoc-AA_{27}-AA_{36}-NH_2 as protected fragments in high yield and purity using 2-chlorotrityl chloride resin. Solution phase assembly of the fragments is followed by precipitation and HPLC purification to yield the final product. By performing this hybridised synthesis the purity and efficiency of the peptide assembly is maximised and the commercial scale of the synthesis significantly reduces the cost of goods.

Despite the success of these synthetic peptides against their respective targets, it should be noted that, in both applications, the target receptor is on a cell surface. In contrast to this, the work conducted in our laboratories focuses on the development of new cell permeable peptides to target intracellular receptors, leading to an additional set of challenges.

1.1.2. Monitoring the cell permeability of peptides

A number of known synthetic strategies have been investigated for the development of active cell permeable peptides with activity against a range of targets. These strategies include *N*-methylation,[22] α-alkylation,[23] cyclisation,[24] intramolecular hydrogen bonding,[25] stapling,[26] effect of charge,[27] and receptor mediated active uptake.[28] The cell permeability of peptide tool sets has been investigated using multiple methods including the commonly used PAMPA (Parallel Artificial Membrane Permeability Assay), which measures passive permeability across lipid bi-layers, and Caco-2 cell lines.[29] Other assays being developed elsewhere in our laboratories include incorporation of non-natural functionality, such as halogen atoms and alkynes, into peptides which can be utilised for analysis of cell permeability by IR and Raman spectroscopy. Mass spectrometry can also be used to confirm the presence of peptides in cell lysates.

While these techniques provide evidence for the potential cell permeability of peptides, they do not allow for direct visualisation of the subcellular location of the molecules within the cell, nor do they all allow for quantification of the permeability. Furthermore, techniques such as PAMPA can only be used to assess passive permeability and do not take into account the effects of active uptake mechanisms such as endocytosis, resulting in a limited analysis of the overall cell permeability. Endocytosis is thought to be a key mechanism for uptake of charged peptides, as suggested by inhibition of internalisation observed upon decreasing cell temperatures to 4°C or reducing the internal ATP content.[30] Therefore, the use of fluorescent tags in combination with fluorescence imaging techniques such as confocal microscopy is highly beneficial and gives a more comprehensive analysis of the cell permeability of peptides. Such techniques have previously been employed in the visualisation of peptide binding to opioid receptors on the surface of living cells.[31]

Developing this technique for intracellular use could not only confirm if the peptides are cell permeable, but also that they reach the desired intracellular site of

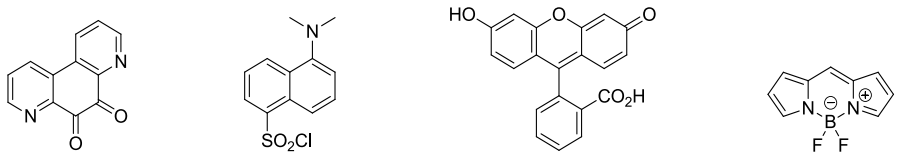
action for the targeted receptors. This is an important consideration since peptides that are able to penetrate cell membranes may still become trapped in endocytic organelles, rendering them effectively useless as therapeutic agents.[32] In cases where good permeability and target binding are indicated in extracellular assays but activity is not detected within cells, this technique could be adopted to identify potential causes of peptide leads with known activity being inactive in functional assays. With this in mind, a literature survey was conducted to identify suitable reactions for intracellular bioorthogonal labelling in order to identify a potential approach to monitoring cell permeability of peptide drug candidates.

1.1.3. Methods of fluorescent labelling

Significant work has been conducted on the observation of large biomolecules and small organic molecules in cells using fluorescence labelling. A number of potential methods were identified and researched for the labelling of peptides developed within our laboratories.

Possibly the most common method is the use of fluorophores as tags for intracellular visualisation of biomolecules. Fluorophores are molecules which exhibit fluorescence and a number of organic probes have been developed either as fluorophores, or as fluorophore precursors which undergo a reaction to become fluorescent. A review by Gonçalves has covered the wide range of available probes and their many uses in detail.[33]

Organic fluorophores tend to consist of a number of substituted aromatic rings linked together. Substituents can be modified to manipulate properties such as the wavelength of excitation and emission, fluorescence lifetime, and intensity of absorption. The wavelengths that can be achieved using these organic probes range from the near ultraviolet to the near infrared. These include probes such as 4,7-phenanthroline-5,6-diones **6** and dansyl chloride **7**, which can fluoresce below 500 nm, or probes such as fluorescein **8** and boron-dipyrromethene (BODIPY) dyes **9** which fluoresce in the range of 500-900 nm (Figure 5).[33]



6 **7** **8** **9**

Figure 5: Structures of common organic probes used in the tagging of biomolecules.[33]

The 4,7-phenanthroline-5,6-diones, such as **6**, have previously been in used in the labelling of L and D amino acids.[34] The molecules are fluorophore precursors in which the electrophilic carbonyl groups are attacked by nucleophilic side chains, such as those in lysine, to give a fluorescent adduct. Dansyl chloride **7** contains a sulfonyl chloride group which is able to undergo nucleophilic attack by amine groups to form sulfonamides. The resulting dansyl amino acid is fluorescent and one of its most popular uses has been for the identification of *N*-terminus amino acids in peptide chains during Edman degradation.[35] Fluoresceins **8** have found application in the area of peptide chemistry where they have been applied as a fluorescent tag to a number of polycationic peptoid oligomers for analysis of their cell penetration by flow cytometry.[36] Finally, BODIPY dyes, such as **9**, have been used for proteomic analysis to confirm protein structure using mass fingerprinting. Thiol-reactive BODIPY analogues containing electrophilic iodoalkyl moieties were shown to be highly specific for the binding to cysteine residues in proteins, allowing for accurate mass analysis.[37]

Another common molecular tag associated with therapeutic peptides is biotin **10**, also known as Vitamin H (Figure 6). Biotin itself is not fluorescent, but is particularly useful because of its affinity to a protein known as streptavidin which is produced by *Streptomyces Avidinii*.[38] The Biotin-streptavidin interaction is non-covalent and selective which results in very little non-specific binding. This is one of the strongest known non-covalent interactions with a binding association (k_a) of approximately 10^{13} M^{-1}, two orders of magnitude greater than the strongest protein-ligand interactions.[39] The ureido moiety of biotin is able to interact with the protein via five hydrogen bonding interactions which, when combined, give greater stabilisation than expected for the sum of the individual hydrogen bonds.

There are several advantages to this system; for example, once bound, any change in pH does not affect the complexation. Additionally, the system can be washed repeatedly without interrupting the interaction. A fluorescently conjugated biotin-flubida-2 derivative **11** can be tagged to a peptide and used to visually confirm the cell permeability of a molecule using confocal microscopy.[40] Once fluorescence is observed inside the cell it can be lysed and the fluorescent biotin-peptide adduct extracted with streptavidin for additional confirmation by mass spectrometry.

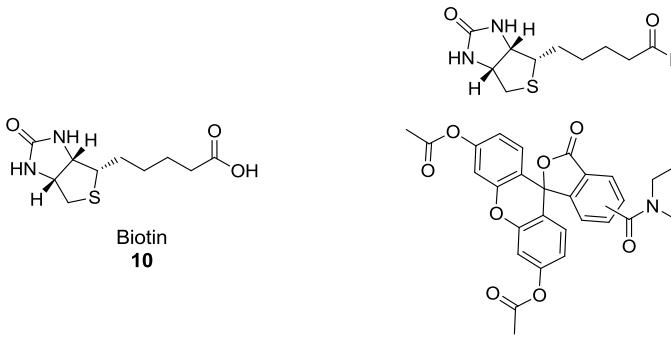


Figure 6: Structures of Biotin **10** and a Biotin-flubida-2 adduct **11**.[40]

In some cases, the permeability of a peptide can be inferred from functional cell-based assays. For example, if a peptide known to have activity against a particular target is introduced to a cell line, the permeability can be inferred from any functional effects observed, such as increased or decreased regulation of cytokines associated with the specific target. However, for general permeability assays where no receptor response can be monitored, the use of fluorescent tags is necessary. While the aforementioned classes of organic probes could be used for the labelling of cyclic and helical peptides for intracellular visualisation, they have major disadvantages; their size and lipophilicity.

The addition of large tags with multiple aromatic rings will introduce additional lipophilicity to the molecule, thus automatically increasing its passive permeability and overall promiscuity. While the comparison of peptides labelled with the same tag is possible, as the tag can be presumed to confer the same change in properties

on each peptide, it is not possible to state whether the peptide would still be permeable with no tag present at all.

Accordingly, a number of reactions have been developed in which a relatively small functional group is added to the peptide to act as a precursor for a fluorophore. This is advantageous as a small functional group will have less effect on the physicochemical properties of the peptide than a fluorophore. Of particular importance is the influence of the tags on the solubility and permeability of the peptide which would interfere with the accurate optimisation of these properties in SAR studies. In this case, the use of a minimalistic tag is preferable in order to give a structure with properties as similar to the original peptide as possible, without affecting any biological activity.

Furthermore, the use of high resolution visualisation techniques like confocal microscopy could give additional information about the subcellular location of peptides. This technique has been utilised for a number of years and its ability to differentiate between concentrations of anthracyclines in the nucleus or cytoplasm of tumour cells was demonstrated in 1993 by Coley *et. al.*[41] More recently, the technique has been used in the training of computer systems for the automated detection of proteins identified previously by human interpretation of captured images.[42]

Bioorthogonal reactions are known to be compatible with intracellular applications where adding chemical reagents could potentially result in a number of undesired reactions, leading to programmed cell death, known as apoptosis. As discussed in the next section, bioorthogonal reactions have been developed for use in the labelling of large biomolecules such as cell surface glycoproteins, carbohydrates and intracellular proteins using fluorophores. The labelling allowed for the introduction of a fluorescent tag after the molecule had permeated the cell membrane.

The principle of this methodology is outlined in Figure 7. A biomolecule within a cell is pre-tagged with an appropriate functional group. A second reagent bearing a

fluorophore moiety is introduced and a reaction conducted within the cell. The resulting biomolecule-fluorophore adduct and its subcellular location can then be visualised.

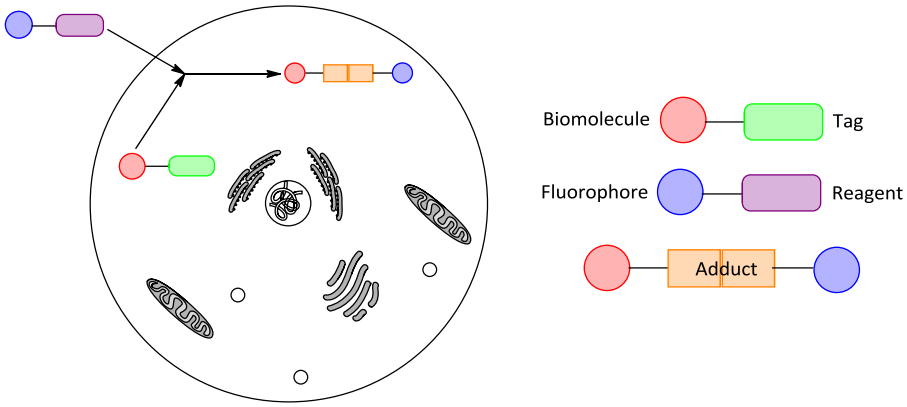


Figure 7: A general depiction for the labelling of intracellular biomolecules with fluorescent tags.

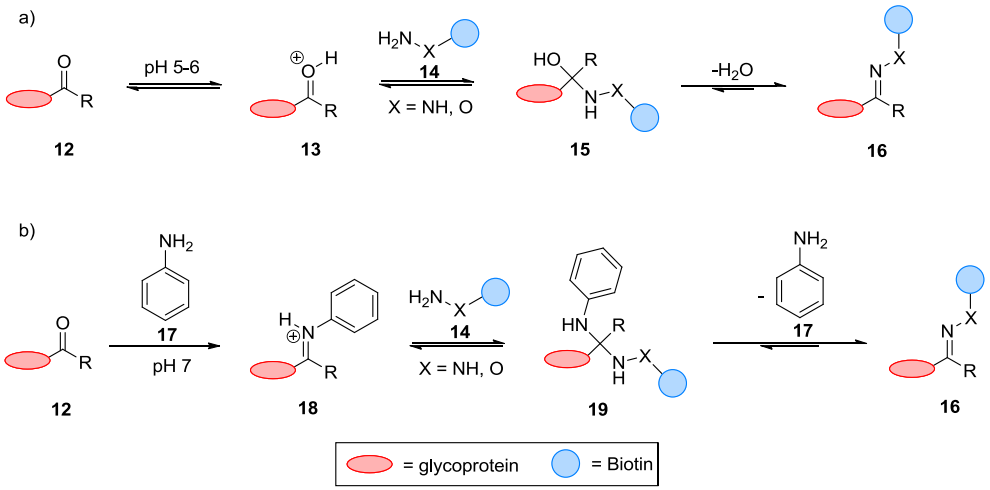
The incorporation of a tag can be achieved in three main ways: modified building blocks are incorporated into some of the target proteins via the cells natural processes; for example, the incorporation of unnatural oligosaccharides into cell glycoconjugates via a metabolic process;[43] genetic modification of the DNA/RNA bio-synthetic pathway to build proteins which always incorporate a specific tag, a common method used to introduce a range of unique amino acids;[44] and finally, a known protein can be visualised by forming a complex with a known substrate bearing an appropriate fluorescent tag.[45]

Regardless of which method is used for its incorporation, the tag can consist of a number of different functional groups. The various reactions that can be conducted with different tags have been reviewed in recent years,[46,47] and are assessed below for their potential use in a peptide cell permeability assay. The most desirable properties for a suitable tag include: small size and low lipophilicity to minimise perturbations of the physicochemical properties of the tagged molecule; chemoselectivity to prevent non-specific labelling; rapid kinetics to allow for quick screening and minimise undesirable side reactions; and applicability to an intracellular environment.

1.1.4. Bioorthogonal reactions for fluorescent labelling

1.1.4.1. Condensation reactions with carbonyls

The reaction of aldehydes and ketones **12** with nucleophilic amines proceeds under acid catalysis at pH 5-6 (Scheme 1a). The carbonyl group is protonated to give **13**, followed by nucleophilic addition of the amine to give a hemi-aminal **15** and finally, loss of water to produce an imine **16**. However, due to the excess water present in cells the equilibrium lies towards the carbonyl. Therefore, amines with an adjacent α-heteroatom to increase nucleophilicity are often required to push the equilibrium towards the product. Highly nucleophilic species such as hydrazines and alkoxyamines **14** are required for good reactivity under physiological conditions.



Scheme 1: (a) Acid or (b) Aniline catalysed formation of glycoprotein-Biotin labelled imines.[47]

However, this reaction is not suitable for intracellular labelling for several reasons: the physiological conditions inside a cell do not have the required pH range; high concentrations (mmol) of reagents are required; if the reaction kinetics are not sufficient the use of reactive electrophiles and nucleophiles could lead to reactions with other molecules present inside cells, such as carbonyl containing pyruvates; and finally, some combinations of nucleophilic amines and electrophilic carbonyls have been shown to be cytotoxic and have even been utilised in the treatment of cancer.[48]

The issue of pH was addressed by using aniline **17** as a catalyst for the formation of imine **18** at pH 7 (Scheme 1b). Weakly protonated carbonyl groups can be replaced with highly protonated aniline iminium groups. Subsequent transaminations, such as formation of imine **16**, have been shown to proceed via diamines **19**, with a 40 fold increase in reaction rate at pH 7.[49] This reaction system was adopted by Zeng *et. al.* who applied it to the labelling of glycoproteins on cell surfaces (Figure 8).[50]

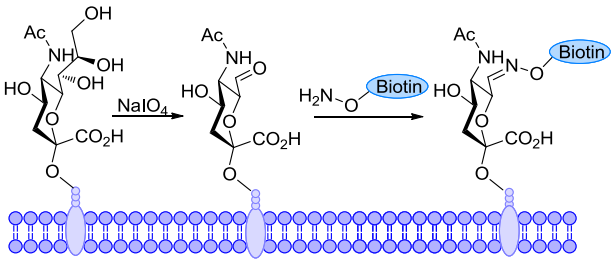


Figure 8: Aniline catalysed oxime ligation of cell surface sialyated glycoproteins.[50]

Cell surfaces do not naturally contain ketone or aldehyde moieties. Therefore, sialic acid in surface glycoproteins could be labelled at pH 7 via an oxidative cleavage of a 1,2-diol with sodium periodate and subsequent aniline catalysed aminooxy-biotin ligation. Cell surfaces were visualised by utilising the aforementioned biotin-streptavidin affinity with a fluorescein label. While useful for such extracellular labelling experiments, this methodology was evidently not suitable for an intracellular peptide permeability assay where the innate nucleophiles and electrophiles present could cause undesired reactions.

1.1.4.2. Palladium catalysed cross-coupling

The development of palladium chemistry in recent decades has resulted in efficient methods for coupling in aqueous media.[51] A number of examples have been reviewed in the literature including functionalisation of proteins using the Heck, Sonogashira and Suzuki reactions.[46,47] In each case a protein was pre-tagged with an aryl iodide or aryl boronic acid which was subsequently reacted bioorthogonally

with an appropriate alkynyl or aryl coupling partner bearing a fluorescent tag, such as biotin or a BODIPY dye, under palladium catalysed conditions.

At the time of writing, there was some doubt as to the suitability of these reagents for bioorthogonal systems. This was due to the requirement of basic conditions, high catalyst loading, temperature manipulation, and poor reaction yields. Furthermore, the reactions had not been tested in a cellular environment and the permeability and toxicity of the water soluble palladium catalysts had not been investigated substantially.

Having stated this, recent work by Spicer and Davis has shown that the Suzuki coupling can be adopted for site-selective incorporation of boronic acid linked carbohydrates **20**-**22** to *E.coli* surfaces (Figure 9).[52]

a)

b)

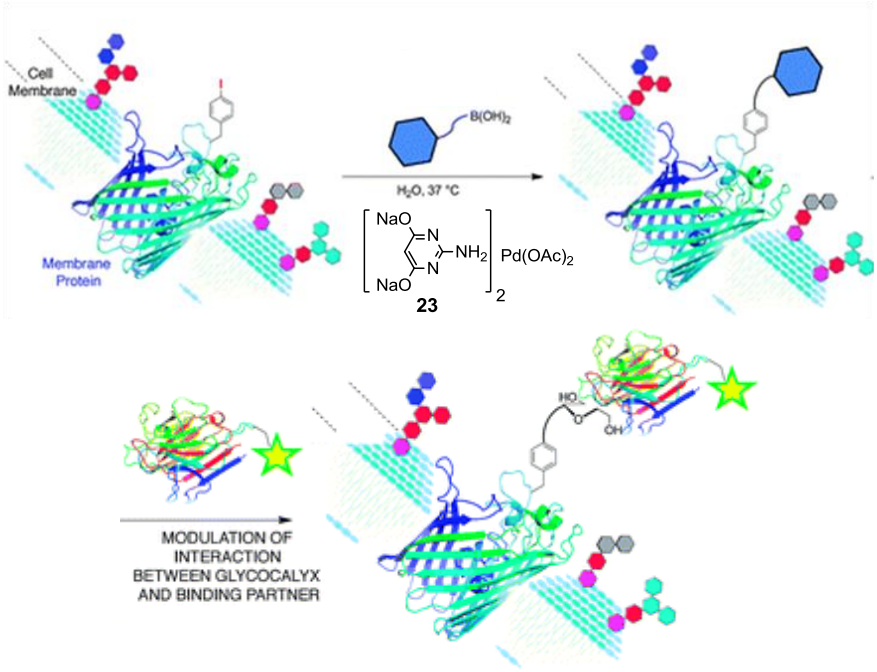


Figure 9: (a) Structures of boronic acid modified *D*-glucose **20**, *D-Galactose* **21**, and *D-Mannose* **22**; (b) The palladium catalysed labelling of cell surface glycocalyx in E.coli cells.[52]

The group used water soluble catalyst $Pd(OAc)_2(ADHP)_2$ **23** to conduct the reactions with modified amino acids, such as *p*-iodophenyl alanine, which were genetically incorporated into proteins on the cell surface of *E.coli*. Once labelled with sugar tags, the cells were used to investigate the sugar-selective binding of lectin proteins. Fluorescein probes were attached to the lectin proteins to allow for visualisation of the binding.

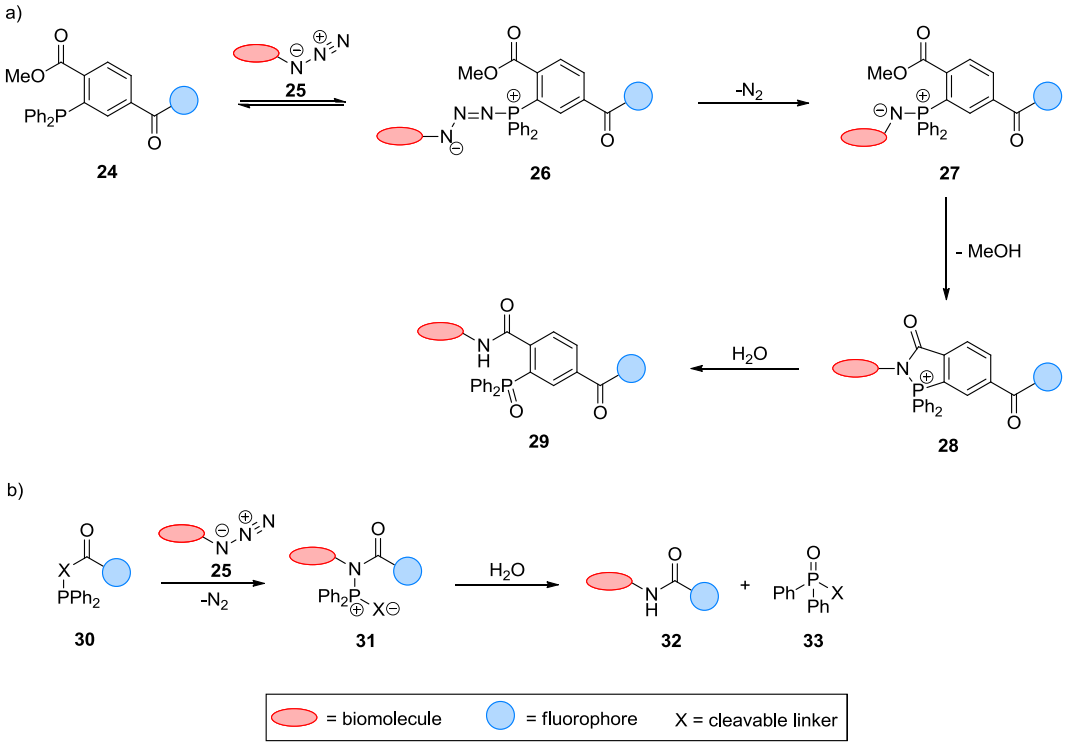
Importantly, this work demonstrated the successful use of palladium catalysed reactions in cellular environments without significant evidence of toxicity, despite the high concentrations of palladium catalyst and boronic acid reagents required (0.5 mmol and 3.5 mmol, respectively). Theoretically, this could be modified for the palladium catalysed coupling of fluorescent boronic acid derivatives with iodine labelled amino acids in peptides. However, the authors did not present yields for the coupling reactions, and more importantly, did not comment on the ability of the catalyst to penetrate cell membranes. In order to utilise this methodology, a catalyst permeability and solubility screen would need to be conducted prior to investigating the reaction itself.

1.1.4.3. The Staudinger ligation

Azide groups have become a popular tag for introducing fluorescent labels into biological systems. These small tags are bioorthogonal as they are not found in cells naturally and, due to their small size, they have negligible effects on the physicochemical properties of the substrate which could be potentially beneficial for use with cell permeable peptides.[53] The first bioorthogonal application of the azide group was in the Staudinger ligation which is a modification of the original Staudinger reaction.[54]

The Staudinger reaction involves the formation of primary amines by treating an azide with a phosphorus reagent. In 2004 Bertozzi *et. al.* demonstrated the use of the Staudinger reaction for the labelling of cell surfaces carrying azide functionality

with a phosphine reagent bearing a fluorescent biotin conjugate. This modification became known as the Staudinger Ligation.[55] The terminal nitrogen of azide **25** is initially attacked by the phosphorus group in the fluorophore containing molecule **24** to give phophazide intermediate **26** (Scheme 2a). An intramolecular rearrangement and loss of nitrogen results in aza-ylide intermediate **27**. At this point, the presence of water can result in the cleavage of the P-N bond to give a primary amine. However, addition of an adjacent ester group allows the aza-ylide to be trapped as intermediate **28**, which hydrolyses to give fluorescently labelled amide-linked biomolecule **29**.[56]



Scheme 2: (a) The Staudinger ligation; (b) The traceless Staudinger ligation.[47]

The traceless Staudinger reaction was developed in order to eliminate phosphine oxide from the resulting products (Scheme 2b). A fluorophore containing a cleavable phosphine **30** was reacted with azide **25** to give phosphoramide **31.** Upon hydrolysis the P-N bond is cleaved, leaving only the desired amide-linked fluorescent conjugate **32** and a phosphine oxide byproduct **33**.[57,58] Interestingly, this modification can be used for amide bond formation during solid phase peptide

synthesis due to the loss of gaseous nitrogen and formation of triphenyl phosphine oxide as a byproduct, which can be washed off the resin.[59]

The same group applied the Staudinger ligation to the labelling of cell surfaces using azide labelled sialic acid groups.[60] The group developed a phosphine reagent containing a fluorescent coumarin moiety and a quenching group. The fluorescence of the fluorophore was initially quenched by Fluorescence Resonance Energy Transfer (FRET). Upon conducting a Staudinger ligation the quenching group is cleaved and the fluorescence is switched on, enabling the labelled sialic amide to be observed (Figure 10).

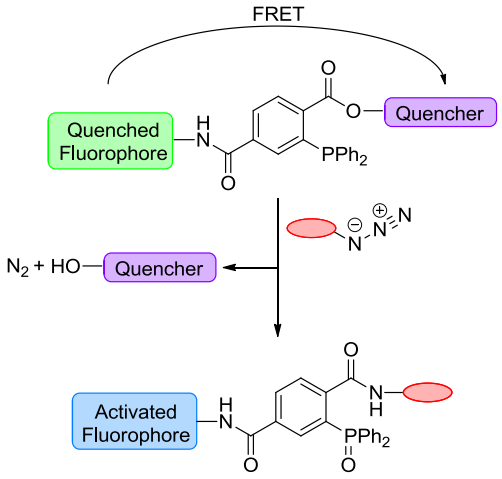
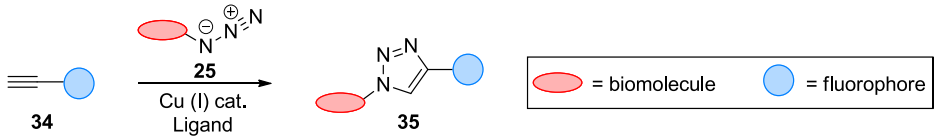


Figure 10: Application of the Staudinger ligation for the labelling of azide-linked biomolecules.[60]

While this type of fluorescence activation is potentially useful for the observation of peptides, there is a number of problems associated with the reaction. Firstly, the applications described above are extracellular and its use in an intracellular assay may be limited due to the ability of glutathione to reduce the azide tag.[61] This reduction by the thiol group could also be caused by any cysteine groups in the peptides being investigated for cell permeability. Furthermore, the phosphine reagents are prone to oxidation. This results in the requirement of excess phosphine reagents which have been shown to be capable of cleaving disulfide bonds, giving them the potential to denature proteins.[62] Having stated this, the group was still able to conduct the Staudinger ligation *in vivo* with mice.[55]

1.1.4.4. Azide-alkyne cycloaddition

The bioorthogonal nature of azide reactions has resulted in the development of fluorogenic azide-alkyne click reactions for use in fluorescent labelling (Scheme 3). This 1,3-dipolar cycloaddition reaction between an azide labelled biomolecule **25** and a fluorophore labelled alkyne **34** is conducted with copper catalysis for the synthesis of 1,2,3-triazoles **35**, and is known as the Cu(I) catalyzed azide-alkyne cycloaddition (CuAAC). A number of different fluorescent precursors have been applied to this reaction including coumarins, 1,8-napthalimides, BODIPY dyes, and anthracene derivatives.[63] The reaction relies on the modification of the electronic properties or energy transfer properties of the system to switch on the fluorescence from non-fluorescent starting reagents.



Scheme 3: The fluorogenic CuAAC reaction for the synthesis of fluorescent triazoles.[47]

Importantly, the CuAAC reaction has been shown to work intracellularly for the labelling of fucosylated glycans (Figure 11).[64] An azide-tagged fucose derivative was introduced to cells for incorporation into glycans via endogenous biosynthetic pathways. The cells were treated with non-fluorescent 4-ethynyl-*N*-ethyl-1,8-naphthalimide and copper(I) bromide was used to induce the fluorogenic click reaction. The intracellular location of the fluorescent fucosylated glycoconjugates was visualized with confocal microscopy and found to be in the area of the Golgi apparatus. Tirell *et. al.* also applied the fluorogenic CuAAC reaction to proteins containing amino acids modified with terminal alkyne groups.[65] A non-fluorescent 3-azido-7-hydroxycoumarin precursor was attached to the proteins with copper(I) catalyst and the resulting fluorescent product could then be visualised.

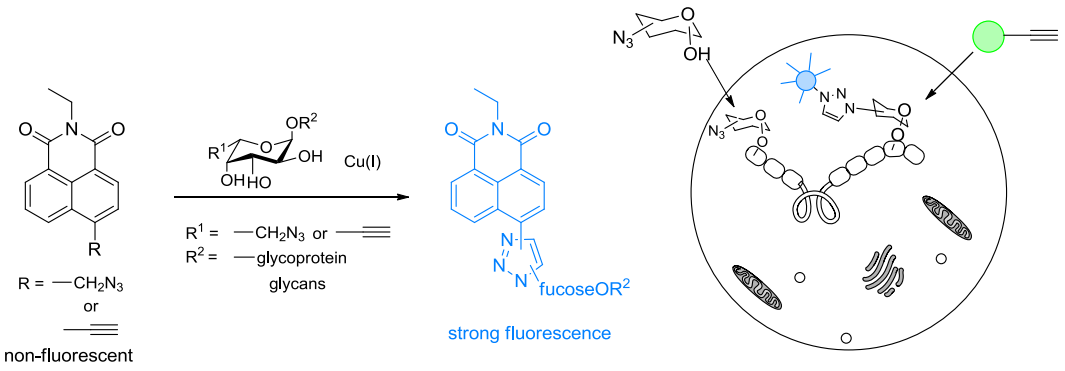


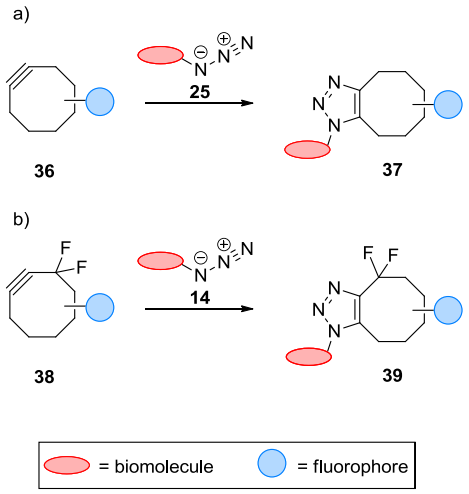
Figure 11: The intracellular CuAAC reaction for the labelling of fucosylated glycans.[64]

While these results are initially encouraging, it should be noted that the use of copper *in vivo* can be cytotoxic, leading to apoptosis. It is known that less than 500 µM concentrations of copper(I) catalyst can be tolerated by mammalian cells for up to 1 hour; however, the CuAAC reaction often requires concentrations of approximately 1 mM for optimal reaction kinetics.[66]

As a result of this significant disadvantage, Bertozzi *et. al.* took an alternative approach to the cycloaddition reaction between azides and alkynes which relied on strain activated alkynes to match the fast rates observed with copper catalysts.[46] Initial reactions were conducted with strained cyclooctyne derivatives **36** in the absence of copper catalysts but these failed to provide reaction rates comparable to the CuAAC reaction (Scheme 4a). Fluorination at the proparygylic positions to give difluorocyclooctyne derivatives **38** resulted in comparable reaction rates (Scheme 4b). However, the accessibility of these compounds is restricted by the number of steps required for their synthesis.[67]

Although there are more accessible cyclooctyne based tags, their use is still unlikely to be appropriate for cell permeable peptides. The size and lipophilicity of these tags could significantly affect the physicochemical properties of the peptides under study. Adopting the CuAAC reaction with small alkyne tags is also not optimal. The low concentrations of catalyst required to avoid cytotoxicity could result in poor reaction kinetics. Subsequently, this could increase the prevalence of azide

reduction by thiols in biomolecules present in the cell, such as glutathione, in addition to other possible side reactions.

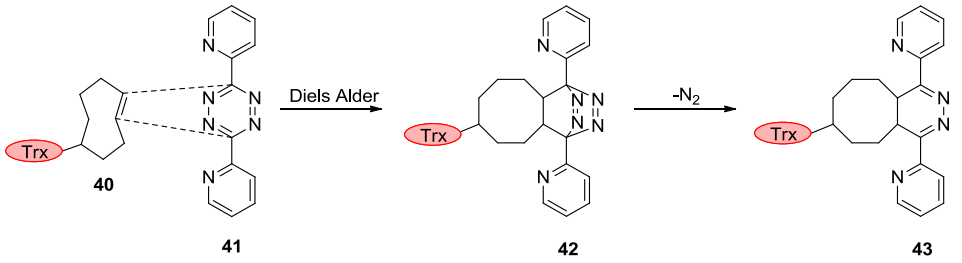


Scheme 4: Strained azide-alkyne click reaction with (a) cyclooctyne and (b) difluorocyclooctyne derivatives.[47]

1.1.4.5. Tetrazine-alkene cycloaddition

The reaction between a 1,2,4,5-tetrazine and an alkene is, at its core, an Inverse Electron Demand (IED) Diels Alder reaction.[68] The reaction proceeds through a Diels Alder intermediate which is followed by the irreversible extrusion of nitrogen to give dihydropyrazine products.

As with many other bioorthogonal techniques the starting materials are not native to cells and do not have any intrinsic reactivity with biomolecules. Early cell experiments were conducted with strained and bulky tags. Rapid rates of reaction could be achieved by using strained *trans*-cyclooctene or norborene as the reaction dienophile.[69,70] A range of tetrazine reagents have been investigated in SKBR-3 cells in order to optimise reaction rates.[71] This methodology was demonstrated by Fox *et. al.* for the IED Diels Alder reaction of cyclooctyne labelled thioredoxin protein (Trx) **40** with *s*-tetrazine **41** (Scheme 5).[72]



Scheme 5: The IED Diels Alder reaction of 1,2,4,5-tetrazines with strained-alkenes.[47]

Extrusion of nitrogen from intermediate **42** results in the rapid formation of conjugated dihydropyrazine **43** in good yield. The reaction was found to be faster than many of the other bioorthogonal reactions and tolerated a range of functionalities. Importantly, the concentration of starting materials required was in the micromolar range, significantly lower than for the other bioorthogonal reactions described above. While the labelling could be carried out at 37 °C with only nitrogen extruded as a byproduct, it should be noted that these tags are once again inappropriate for use in a peptide cell permeability assay. The large size and lipophilicity could again affect the physicochemical properties of the peptides under investigation.

Recent research conducted by Devaraj *et. al.* introduced the cyclopropene tag as an alternative for the larger norbornene or *trans*-cyclooctene tags.[73] The cyclopropene derivatives incorporate fewer atoms into the molecule, resulting in less perturbation of the physicochemical properties while maintaining high reactivity due to ring strain.

The group applied the cyclopropene-tetrazine system to fluorescence labelling of SKBR-3 cell membranes (Figure 12). The phospholipid bilayer of the cell was labelled with cyclopropene functionality and the IED Diels Alder reaction conducted at 20 °C to give the dihydropyrazine product. In addition, the conjugation of the tetrazine reagent to a BODIPY probe resulted in quenching of the intrinsic fluorescence of the probe. Upon conducting the reaction, the dihydropyrazine product was no longer able to quench the fluorescence, resulting in another turn-on fluorogenic reaction.

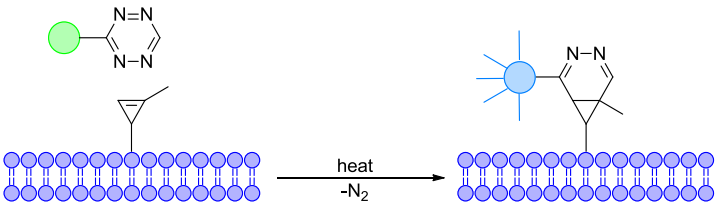
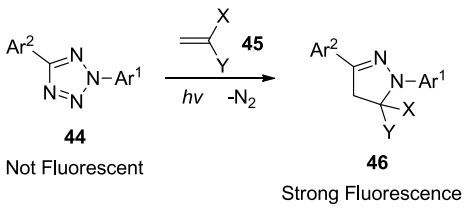


Figure 12: Turn-on fluorescence labelling of a cyclopropene labelled cell surface with a tetrazine-BODIPY conjugate.[73]

This technique offers a number of advantages that make it potentially suitable for use in a peptide cell permeability assay. These include small cyclopropene tags, giving minimal perturbation to peptide properties; low reagent concentrations; loss of nitrogen as the only byproduct; rapid reaction rates; mild reaction conditions; and turn-on fluorescence, which will minimise background interference.

1.1.4.6. Tetrazole-alkene cycloaddition

In addition to the tetrazine-alkene system outlined above, there is another related technique which appeared to be a viable method for the labelling of cell permeable peptides. An alkene tagged molecule **45** can react with 2,5-diaryl tetrazole **44** under photochemical conditions to give a fluorescent pyrazoline adduct **46** (Scheme 6). This method has been explored thoroughly, with multiple cell-based applications described by Lin *et. al*. in a recent review.[74]



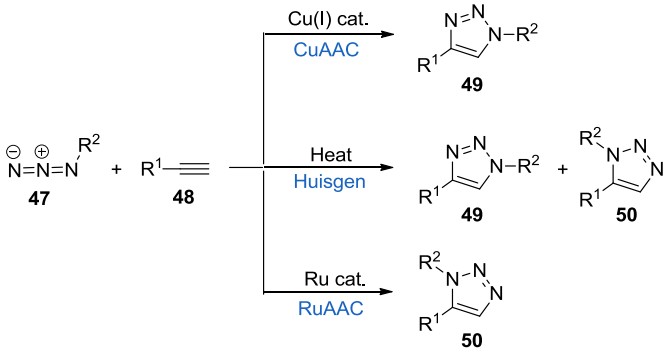
Scheme 6: The 1,3-dipolar cycloaddition reaction of 2,5-diaryl tetrazole **44** with alkene tags **45**.

This reaction is another example of a turn-on fluorogenic reaction which can be induced rapidly and under mild conditions with readily obtainable UV bulbs. The only reaction byproduct indicated by the authors is the loss of nitrogen, but more importantly, the alkene tag is even more minimalistic than those seen in the

tetrazine-alkene system. As a result of these factors, this methodology from Lin *et. al.* was chosen for further investigating the peptide cell permeability assay. The details of the reaction will be covered below in a more comprehensive manner.

1.1.5. The photochemical fluorogenic click reaction

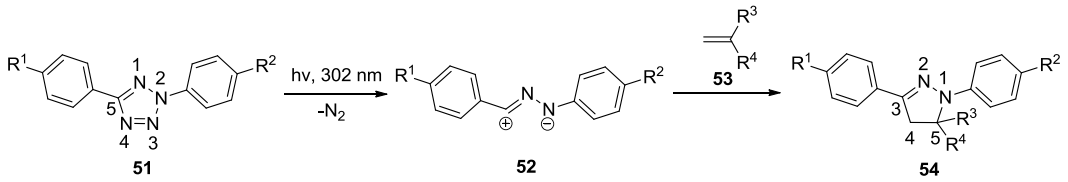
The 1,3-dipolar cycloaddition reaction between a 1,3-dipole and a dipolarophile was first reported by Huisgen *et. al.* in 1967.[75] The original temperature driven cycloaddition between substituted azides **47** and substituted alkynes **48** resulted in formation of both 1,4- and 1,5-regioisomers (Scheme 7). Modifications of the reaction have led to regiospecific cycloaddition reactions such as the aforementioned CuAAC reaction to give 1,4-substituted 1,2,3-triazoles **49**, or the analogous ruthenium based RuAAC reaction to give 1,5-substituted 1,2,3-triazoles **50**.[76,77] Since its origin, the substrate scope for the reaction has expanded to include reactive groups such as nitrile oxides, diazoalkanes, and ozone as 1,3-dipoles; with common dienophiles including alkenes, alkynes, and acrylates.[78]



Scheme 7: The CuAAC, Huisgen and RuAAC cycloaddition reactions.

In 2008, Lin *et. al.* introduced the regiospecific 1,3-dipolar cycloaddition of nitrile imines and alkenes for the fluorescent labelling of proteins in cells.[79] In this case, UV light (302 nm) is applied to a 2,5-diaryl tetrazole **51** which initiates a cycloreversion reaction and results in extrusion of nitrogen to form a reactive 1,3-dipolar nitrile imine intermediate **52** (Scheme 8). A [2+3] cycloaddition reaction will occur in the presence of an alkene **53** to form a 1,3,5-trisubstituted pyrazoline

adduct **54** with fluorescent properties.[80] This class of pyrazoline adducts is known to have strong UV absorption bands due to the π-π* transition of the conjugated backbone and emit fluorescence in the visible blue region with a Stokes shift of approximately 100 nm.[81]



Scheme 8: Synthesis of fluorescent pyrazoline derivatives **54** via photo-induced click chemistry.

Four potential electronic contributions have been postulated for the nitrile imine intermediate: propargylic **55**, allenic **56**, 1,3-dipolar **57**, and carbenic **58** (Figure 13).

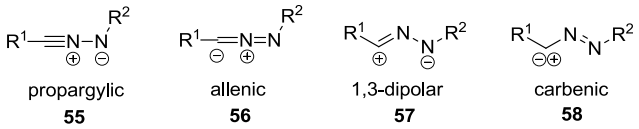


Figure 13: Postulated structures of the nitrile imine intermediate.

Direct observation of the reaction intermediate and its geometry was achieved by Zheng *et. al.* using solid state chemistry.[82] Crystals of a number of zinc-tetrazole complexes were formed using tetrazoles such as **59** (Figure 14a). These were irradiated with a He-Cd laser beam at 325 nm for 2 minutes. From the X-ray photodifference maps obtained it was observed that the nitrile imine backbone consisted of a bent geometry, narrowing the potential structure to either the 1,3-dipolar **57** or carbenic **58** resonance form (Figure 14b). The zinc atoms were found to coordinate with two carboxylate groups and two water molecules each (only one set of these coordinations is shown). Extruded nitrogen molecules were also observed, packed between the layers of tetrazoles in the crystal structure (Figure 14c).

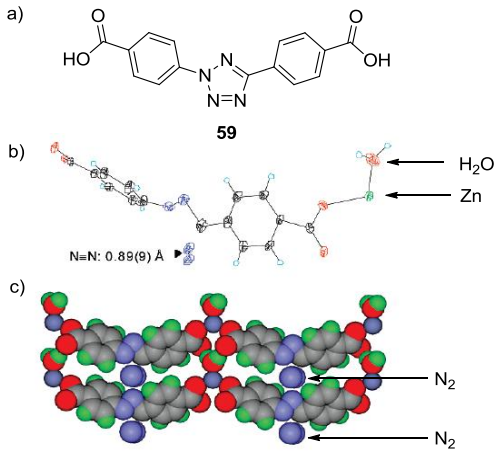
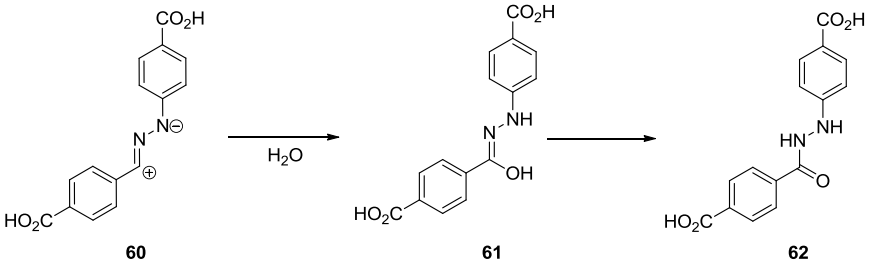


Figure 14: (a) Structure of tetrazole **59**; (b) Graphical representation of the geometry observed after nitrogen extrusion; (c) Observed packing of the nitrile imine intermediates with nitrogen molecules visible between the layers (perpendicular to the plane of view), Black = carbon, Red = oxygen, Blue = nitrogen, Green = hydrogen.[82]

Further experiments by the same group identified the 1,3-dipolar structure to be the most accurate representation of the nitrile imine geometry. Irradiation of tetrazole **59** in MeCN:H_2O (1:1 v/v) resulted in quenching of nitrile imine intermediate **60** to give hydrazonic acid **61**, followed by tautomerism to yield the thermodynamically stable hydrazide **62** (Scheme 9). Formation of hydrazide **62** via quenching with water is not possible via the carbenic resonance structure. The reactive nature of the nitrile imine intermediate can be partially attributed to the pre-organised bent geometry which gives a lower activation energy barrier for cycloaddition reactions.[83,84]



Scheme 9: Quenching of the nitrile imine intermediate by water to yield a stable hydrazide product.

This reaction highlights the potential for the nitrile imine to undergo quenching in aqueous media which would be detrimental in a cell based assay. However,

Molteni *et. al.* have demonstrated that, in the presence of a dipolarophile such as an alkene, the use of aqueous media will accelerate the rate of the desired reaction via the hydrophobic effect, provided that an organic co-solvent is used to achieve a homogenous solution.[85] Additionally, the relative reactivity of substrates for cycloaddition reactions is largely dominated by the electronic effects of their substituents. In this case, good overlap is required between the HOMO of the dipolar nitrile imine and the LUMO of the alkene dipolarophile (Figure 15). Maintaining high reaction rates for pyrazoline formation at the low concentrations present in cells should help prevent such side reactions.

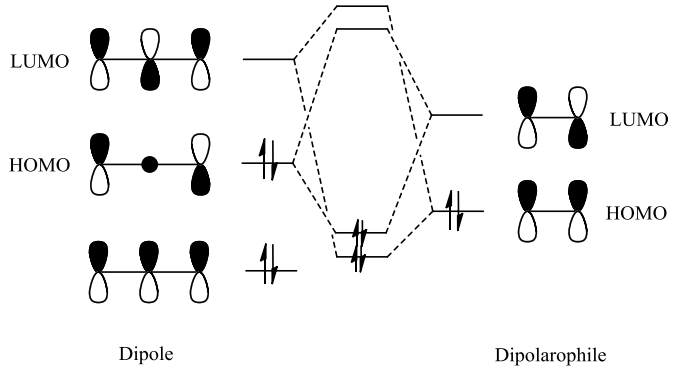
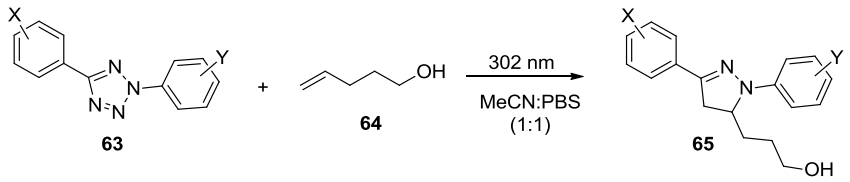


Figure 15: HOMO and LUMO matching between nitrile imine 1,3-dipole and alkene dipolarophiles.

A series of experiments conducted by Lin *et. al.* have demonstrated that a wide range of substituents are tolerated in the reaction for both the alkene and tetrazole. Clear dependence on the electronic character of both reagents was required for good HOMO/LUMO overlap.[86] A number of monosubstituted, 1,1-disubstituted, or 1,2-disubstituted alkenes bearing electron withdrawing moieties, including reactive aldehyde and nitrile groups, were reacted to give the corresponding pyrazoline adducts with quantitative conversions. The reaction of electron neutral oct-1-ene resulted in a moderate yield of 62%, while alkenes with electron donating substituents such as *N*-butyl vinyl ether were reported to give only byproducts, thus emphasising the importance of good orbital overlap. As is typical of reactions with concerted mechanisms, the nature of the solvent showed little or no effect on the reaction yields obtained.[86]

A more in-depth study was conducted to determine the relationship between the calculated HOMO energy of a number of tetrazoles **63** and their experimentally determined rate constants for cycloaddition with pent-4-en-1-ol **64** (Scheme 10, Figure 16).[80] As predicted, electron donating groups raise the HOMO energy leading to an increased rate constant, while electron withdrawing moieties had the opposite effect. Notably, the regioisomerism of the substituent was found to have a significant impact with *para* substitution having a larger effect than both *ortho* or *meta* substitution. Multiple substituents offered no further increase in reaction rate. Interestingly, while the same electronic relationship was observed for both rings, the substituents on the *C*5-aryl ring of the tetrazole showed significantly less effect on reaction rate compared to those on the *N*2-aryl ring.



Scheme 10: Synthesis of a range of pyrazoline derivates **65** via photochemical cycloaddition of a range of tetrazoles **63** with pent-4-en-1-ol **64**.

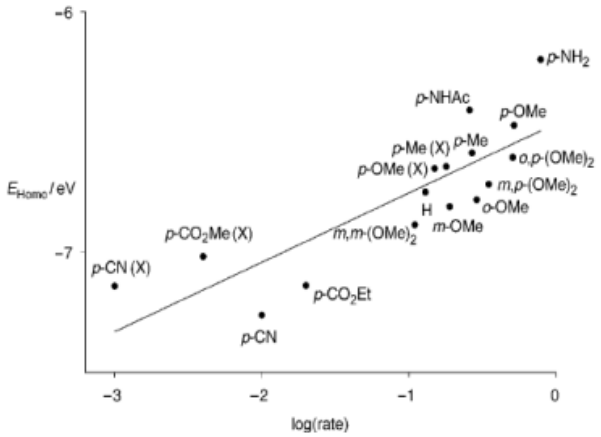
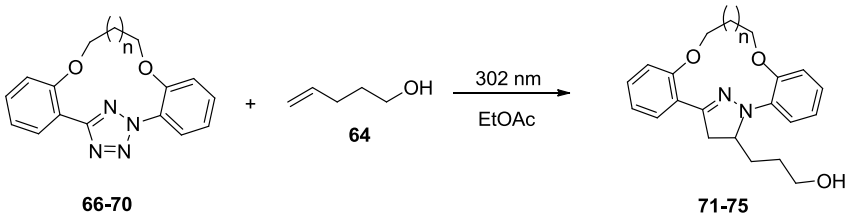


Figure 16: Observed relationship between the HOMO energy of the nitrile imine intermediate and its reaction rate with pent-4-en-1-ol **64**. Effect of substitution is shown for both the *N*-aryl ring and *C*-aryl ring, the latter of which are represented with (X).[80]

The group also devised a series of macrocyclic tetrazoles **66**-**70** which are pre-disposed to undergoing rapid pyrazoline formation.[87] The nitrile imine intermediate is held in the bent geometry required for the reaction by linking the aryl groups together via *ortho*-substitution (Scheme 11, Table 1).



Scheme 11: Reaction of tetrazoles **66**-**70** with 4-penten-1-ol **64** (100 equivs.) in ethyl acetate.

Entry	Tetrazole	Pyrazoline	n	Reaction Time/min	Reaction Yield/%
1	66	71	Untethered	120	58
2	67	72	1	120	59
3	68	73	2	75	71
4	69	74	3	84	60
5	70	75	4	180	70

Table 1: Reaction times and yields observed for the reaction of tetrazoles **66**-**70** with 4-penten-1-ol **64**.

The chain length was found to be an important factor for increasing the reaction rates and decreasing the reaction times relative to the untethered methoxy substituted tetrazole **66** (Entry 1). It was observed that, while a tether consisting of either three or six methylene units (Entry 2 and 5, respectively) gave no significant benefits to the reaction time and yield, a tether of four or five methylene units was able to give improved yields with significantly reduced reaction times (Entry 3 and 4, respectively). Of these, tetrazole **68** with a four methylene tether was found to be optimal; this was attributed to the optimal conformational flexibility instilled by the tether.

Evidently, careful consideration must be given to all of the rate determining factors when devising a suitable pair of reagents in order to maximise the rate of the desired reaction and minimise the rate of unwanted side reactions.

The electronic character of the starting materials not only affects the rate of reaction, but also its regioselectivity; or as is often observed for pyrazoline synthesis, its regiospecificity. For all of the above examples, regardless of whether the nitrile imine was photochemically induced from a tetrazole or chemically induced via hydrazonoyl chloride deprotonation, the *C*5-substituted pyrazoline is observed exclusively over the *C*4-substituted regioisomer.

The interpretation of the Frontier Molecular Orbitals (FMO) and mechanistic aspects of the reaction between nitrile imines and alkenes is a topic which has been under investigation with theoretical calculations since as early as 1974.[88] The formation of the *C*5-substituted pyrazoline over the *C*4-regioisomer is acknowledged to be an anomalous result based on the general understanding of cycloaddition reactions;[89] the negatively charged nitrogen atom of the nitrile imine **52** would be expected to attack the electropositive terminal carbon of an electron deficient dipolarophile **76** to give the *C*4–substituted pyrazoline **78** (Figure 17).

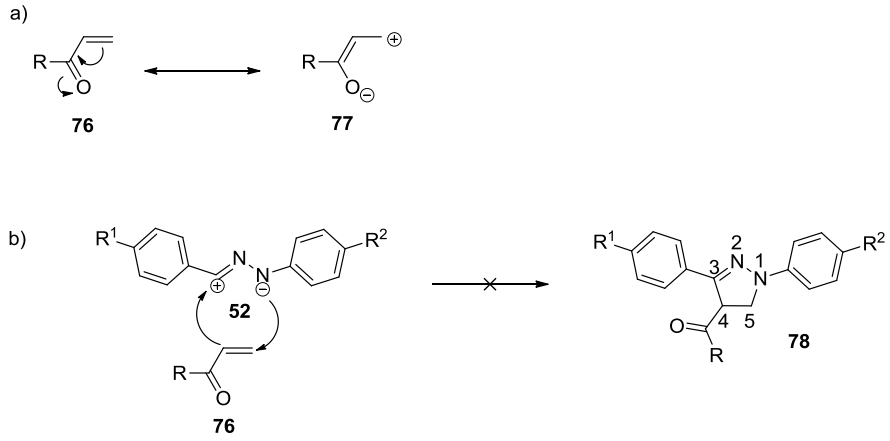


Figure 17: (a) Resonance structures of electron deficient dipolarophile **76**; (b) Expected regiochemistry of the 1,3-dipolar cycloaddition reaction with nitrile imine **52**.

In general, the *N*1 atom is accepted to be the nucleophilic terminus of the dipole, while the *C*3 atom acts as an electrophilic terminus. However, calculations conducted as early as 1976 by Caramella and Houk indicated that, contrary to established theoretical predictions, when a bent geometry is adopted the largest HOMO coefficient is actually present on the *C*3 atom and not the *N*1 atom.[89] The group therefore rationalised that the carbon atom of the nitrile imine acted as the nucleophilic terminus instead of the nitrogen atom, resulting in the formation of the *C*5-substituted regioisomer.

In agreement with these findings, the reaction of *N*-aryl *C*-ethoxycarbonyl nitrile imines with acrylic acid derivatives were found to give only the *C*5-substituted regioisomer. This regiospecificity is known to dominate, regardless of the substituents on the aromatic ring or the dipolarophile, and is a result of the nitrile imine HOMO and dipolarophile LUMO interactions (Figure 18).[90]

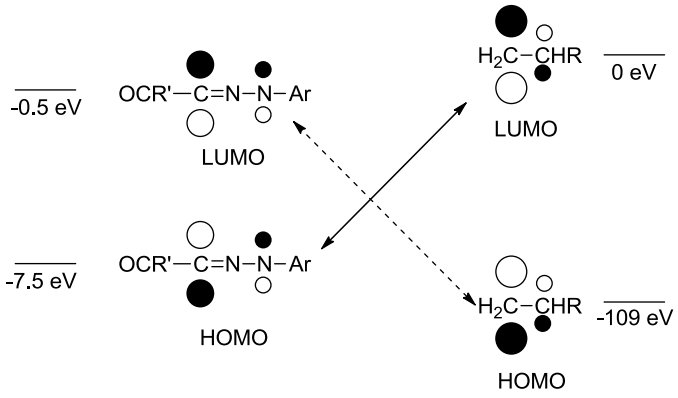


Figure 18: Calculated HOMO and LUMO energies for an *N*-aryl *C*-ethoxycarbonyl nitrile imine and alkene dipolarophile.[90]

As suggested previously, the largest coefficients were found to be present on the terminal carbon of the dipolarophile and on the *C*3 atom of the nitrile imine which resulted in asynchronous bond formation, leading to only the single regioisomer. The same regioselectivity has also been observed for 2,5-diaryl nitrile imines.[91]

Despite these initial understandings, the exact mechanism of the reaction is still a matter of debate. A more recent computational study conducted by

Mawhinney *et. al.* in 2005 led to new mechanistic insights.[92] Two mechanisms were hypothesised for [3+2] cycloaddition reactions: (1) a concerted mechanism in which both bonds are formed during the transition state, or (2) a two step radical process in which a single bond is formed, followed by bond rotation and ring closure.[93] Calculations were conducted from the geometric midpoint of both the alkene and the nitrile imine, as represented by the black dot (Figure 19).

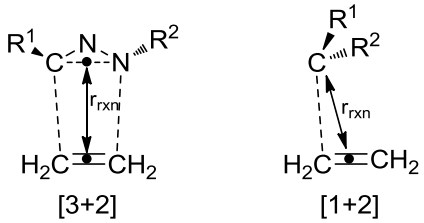


Figure 19: Proposed transition states for the [3+2] and [1+2] cycloaddition reactions of nitrile imines with ethene.[92]

The work carried out directly with nitrile imines by Mawhinney *et. al.* suggested that the former mechanism was most likely. The results of the computational study found that the [3+2] reaction is more kinetically and thermodynamically favourable than the [1+2] reaction. Most interestingly, the [3+2] reaction mechanism was found to be initiated via the [1+2] transition state. The authors suggest that the approach of the two molecules is initially dominated by [1+2] interactions until a critical bond distance is reached, at which point the [3+2] mechanism becomes dominant. This early interaction of the nitrile imine *C*3 atom and dipolarophile terminus, coupled with the asynchronous bond formation mentioned previously, can account for the exclusive formation of the *C*5-pyrazoline adduct.

This new evidence is an intriguing insight into the potential mechanistic pathways of cycloaddition reactions and will undoubtedly be subject to further scrutiny. Since the regiochemistry is fixed and the reactivity of the tetrazole and alkene starting materials can be predicted from the number of electronic factors outlined above, it should be possible to find a suitable pair of starting materials for the intracellular visualisation of cell permeable peptides. However, it should be noted that the

research conducted by Lin *et. al.* has principally focused around the incorporation of one of the reactions partners into proteins. While the group has shown that it is possible to incorporate mono-*N*-aryl tetrazoles into amino acids with either *C*-aryl, *C*-acyl or *C*-alkyl linkers, the majority of the work conducted has been carried out with alkene modified amino acids.[94,95] It should also be noted that the authors suggest that the presence of the *N*-aryl ring is vital for the photoactivation of the tetrazole.

The same authors have conducted a number of successful protein labelling experiments in both bacterial and mammalian cells using different combinations of alkene and tetrazole. Early experiments were conducted on a genetically modified Z-domain protein containing allyl tyrosine tags. The protein was labelled in *E.coli* cells using tetrazole **79**, with no labelling of the wild type protein observed (Figure 20).[96] The study was later repeated with tetrazole **81** with a 475 fold increase in reaction rate which was consistent with earlier observations of the reaction kinetics.[80]

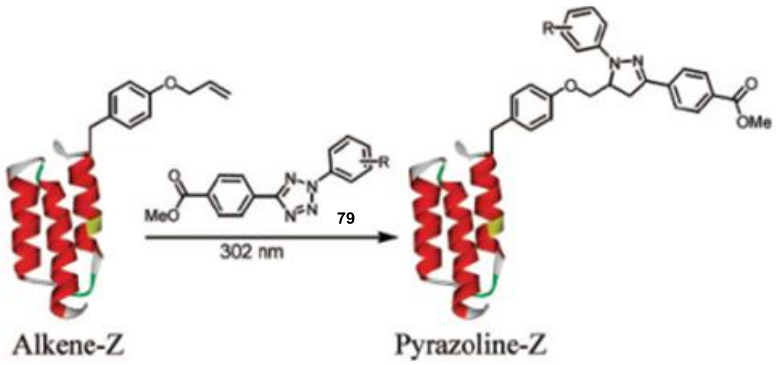


Figure 20: Photochemical labelling of allyltyrosine encoded Z-domain proteins in *E.coli* cells.[96]

Homoallyl glycine (HAG) **80** is an alternative tag to allyl tyrosine. HAG was incorporated into human HeLa cells by allowing the bio-synthetic pathways of the cells to synthesise proteins with HAG modified amino acids in place of methionine (Figure 21).[97] When labelled with tetrazole **81**, the HAG encoded proteins were visualised with significant increases in fluorescence over cells with only wild type

proteins. Interestingly, the authors also noted that if only the tetrazole is present in the cell a small increase in fluorescence is still observed. They imply that this weak fluorescence is indicative of the presence of the nitrile imine intermediate.

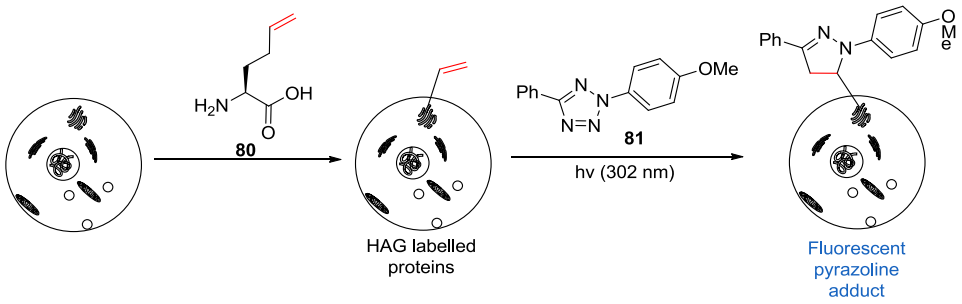


Figure 21: Labelling of HAG encoded proteins in HeLa cells.[97]

This research was encouraging for the development of the proposed cell permeability assay as both HAG and allyl tyrosine could provide the required peptide tag with minimal effects on the physicochemical properties. It had been noted that the use of 302 nm as the photoirradiation wavelength could be potentially harmful to cells due to the high energy of the electromagnetic radiation. Lin *et. al.* addressed this issue in subsequent papers by analysing several known tetrazoles for their reactivity under the longer and less energetic wavelength of 365 nm. This study showed that reactions yields varied greatly under these conditions. Therefore a number of alternative scaffolds were investigated, with increased conjugation for longer wavelength activation (Figure 22).[98,99] The experiments met with mixed results, with three of the new scaffolds **82**-**84** being unsuitable due to poor aqueous solubility. The remaining scaffold **85** showed reactivity in PBS:MeCN (1:1 v/v) at 365 nm but was not tested in cells. Evidently, additional research is required in order to obtain water soluble and long wavelength activated tetrazoles for biomolecule labelling in a cellular context.

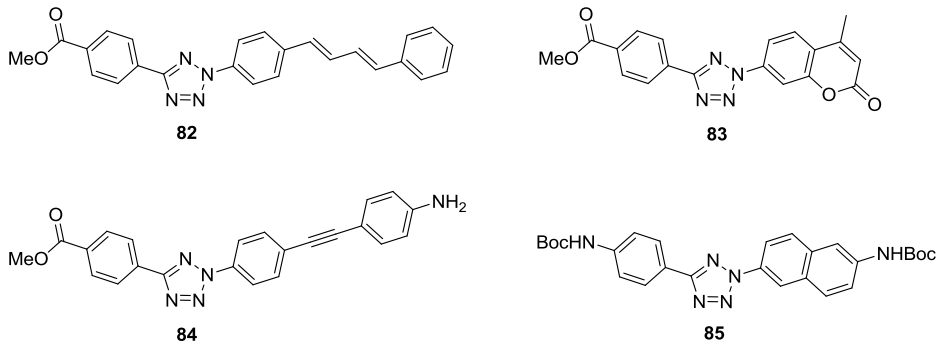


Figure 22: Structure of new scaffolds investigated for long wavelength induced cycloaddition reactions.

1.1.6. Summary and objectives

The fluorogenic photo-click reaction between an alkene and a 2,5-diaryl tetrazole is potentially a powerful technique for intracellular labelling and visualisation of alkene tagged peptides and proteins. The small alkene tag can be incorporated into helical or cyclic peptides during their synthesis using modified amino acid building blocks or via post-synthetic tagging. Upon identification of cell permeable peptides with activity against a target, it will be possible to determine the residues which are vital for activity and desirable physicochemical properties using alanine scanning. Those residues which are deemed to be of least importance can be replaced with alkene modified amino acids and any changes to the physicochemical properties can be measured with assays available in our laboratories.

The proposed methodology for the assay can be seen in Figure 23. Both the cell permeable peptide containing the alkene tag and the tetrazole reagent will be allowed to separately enter a cell, at which time irradiation with a 302 nm UV light will initiate the photo-click reaction. The fluorescence of the pyrazoline product will be monitored using appropriate excitation and emission wavelengths, which are to be determined. Comparison with background images and controls should show the fluorescence exhibited by this turn-on system. The possibility of obtaining quantitative measurements of cell permeability will be investigated, though this is

only likely to be possible if the pyrazoline adduct formed inside the cell is impermeable and unable to escape.

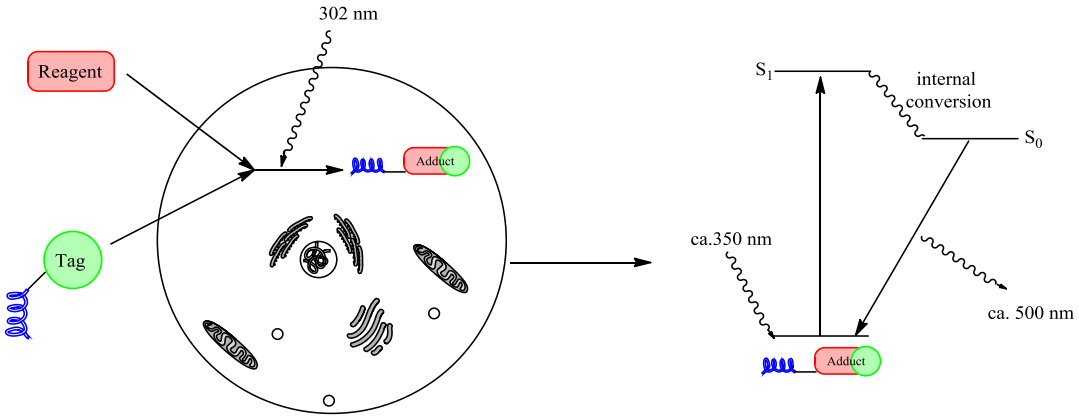


Figure 23: Proposed concept of the fluorogenic assay for visualisation of cell permeable peptides.

Based on all of the above, the objective of the project was to develop an intracellular labelling reaction and fluorescence assay which could determine the permeability of peptides synthesised in our laboratory and allow visualisation of their location within cells.

The successful development of this assay would require achievement of the following aims:

- To prepare suitable tool molecules for synthesis and analysis of reaction products.
- To test the bioorthogonal reaction under conditions appropriate for intracellular applications, with appropriate controls.
- To examine the bioorthogonal reaction intracellularly, with appropriate controls.
- To determine if subcellular localisation of cell permeable peptides is possible.
- To assess if a quantitative assessment of cell permeability can be conducted.

1.2. Identification, synthesis, and analysis of pyrazoline adducts

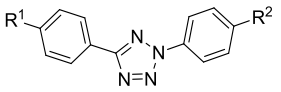
1.2.1. Identifying a tool set

The selection of suitable alkene tagged amino acid building blocks for incorporation into peptides was initiated by utilising a small group of tool molecules for exploratory reactions. Commercially available amino acids Fmoc-dehydro-Leu-OH **86** and Boc-allyl-Gly-OH **87** were identified as potential tool molecules due to the terminal alkene present in the side chain (Figure 24). Protection of the *N*-terminus was required for all amino acids since unprotected amino acids are known to undergo nucleophilic addition to 1,3 dipolar nitrile imines.[100]

86 **87**

Figure 24: Tool molecules for investigating the synthesis of pyrazoline derivatised amino acids.

Three tetrazole molecules were identified as reagents that could be used for the exploratory reactions based on the previously discussed kinetic data. Tetrazole **89** was available in our laboratory[101] and tetrazoles **88** and **81** had been obtained from commercial sources (Figure 25). Literature evidence indicated that pyrazoline molecules synthesised from tetrazole **89** were prone to rapid photobleaching, rendering them ineffective for bioorthogonal labelling; a potential problem that would need to be monitored during the assay development.[99] Therefore, this molecule was discarded as a potential tagging reagent. The two remaining tetrazoles were taken forward for photochemical cycloaddition with the alkene tagged amino acid tool molecules.

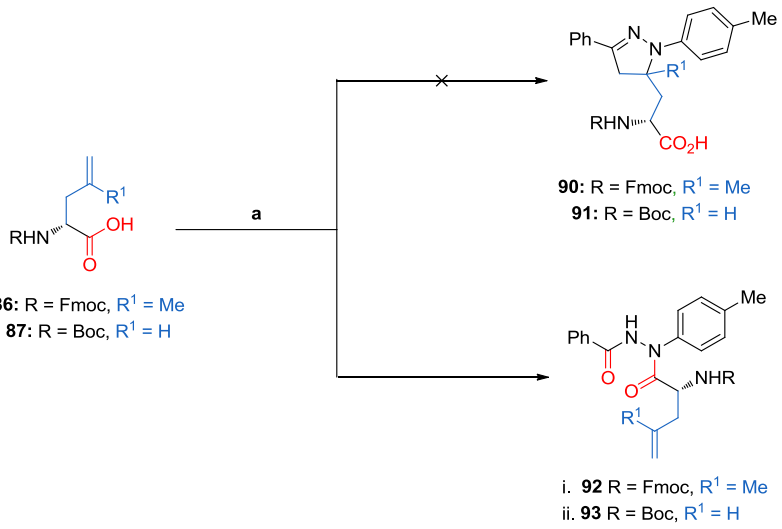


81 R^1 = H, R^2 = OMe
88 R^1 = H, R^2 = Me
89 R^1 = H, R^2 = NH_2

Figure 25: 2,5-diaryl tetrazole compounds available for reactions with alkene tagged amino acids.

1.2.2. Reaction of *N*-terminal protected amino acid tags

The preliminary photo-click reactions were conducted using tagged amino acids Fmoc-dehydro-Leu-OH **86** and Boc-allyl-Gly-OH **87** with tetrazole **88** as the tagging reagent (Scheme 12).



Scheme 12: Reactions of amino acids **86** and **87** with tetrazole **88**. Reagents and conditions: (a) i. Tetrazole **88** (1 equiv.), UV light (270-330 nm), MeCN, rt, 1 h, 51%; ii. Tetrazole **88** (1 equiv.), UV light (270-330 nm), MeCN, rt, 1 h, 63%.

The reaction of Fmoc-dehydro-Leu-OH **86** and tetrazole **88** appeared to proceed satisfactorily with 63% of the isolated product. A single peak was observed by LC-MS with the expected $[M+H]^+$ ion; however, analysis by ^{1}H NMR indicated that the structure of the product did not correspond to the expected pyrazoline, as shown by the presence of terminal alkene peaks in the region of 4.85-4.62 ppm.

Analysis by NMR was complicated due to the presence of rotamers which prevented an accurate ^{13}C NMR from being obtained. Proton signals in the ^{1}H NMR spectrum could be resolved by variable temperature NMR at 120 °C. The high temperature spectrum gave significantly improved dispersion and enabled identification of the product as hydrazide **92**; however, the high temperature also caused partial degradation of the Fmoc protecting group as observed by appearance of a singlet at

δ_H = 6.2 ppm, associated with the formation of dibenzofulvene. This is a known phenomenon and has been identified as a potential way of removing Fmoc protecting groups without the requirement of base or scavengers.[102]

The reaction of Boc-allyl-Gly-OH **87** with tetrazole **88** under identical conditions gave analogous results with the desired $[M+H]^+$ mass ion observed by LC-MS and analysis by ^{1}H NMR also indicating the presence of a terminal alkene with peaks at 5.75 and 5.05 ppm. Peak broadening was again observed and prevented full analysis by ^{13}C NMR but variable temperature ^{1}H NMR at 120 °C was able to give a single set of proton peaks without any cleavage of the Boc group observed.

The presence of the new carbonyl group in **93** was confirmed by HMBC which showed correlations between the ortho protons of the phenyl ring and the carbonyl centre (see appendix, Figure A1). Subsequent analysis provided further evidence of three different carbonyl signals observed by ^{13}C NMR (δ_C = 171.5, 165.6, 155.2 ppm) and IR (1696, 1666 and 1646 cm^{-1}). These data are consistent with literature values which report the carbonyl stretches of similar secondary carbamates to be in the region of 1736-1690 cm^{-1} with the two remaining peaks falling in the known region for amide bonds.[103,104]

A proposed mechanism for the rearrangement reaction which results in formation of hydrazide moieties is provided in Figure 26 using molecule **87** as a representative exemplar. The rearrangement reaction of carboxylic acids with nitrile imines to give hydrazides has little precedence in the literature but has been shown to occur intermolecularly, even in the presence of an intramolecular alkene coupling partner (Scheme 13).[105]

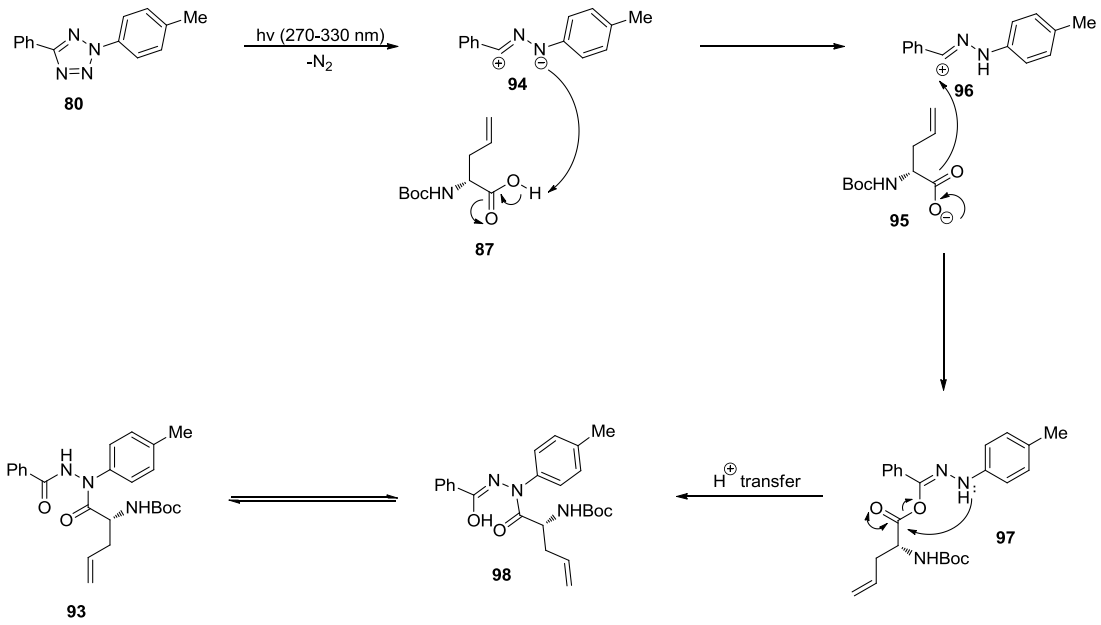
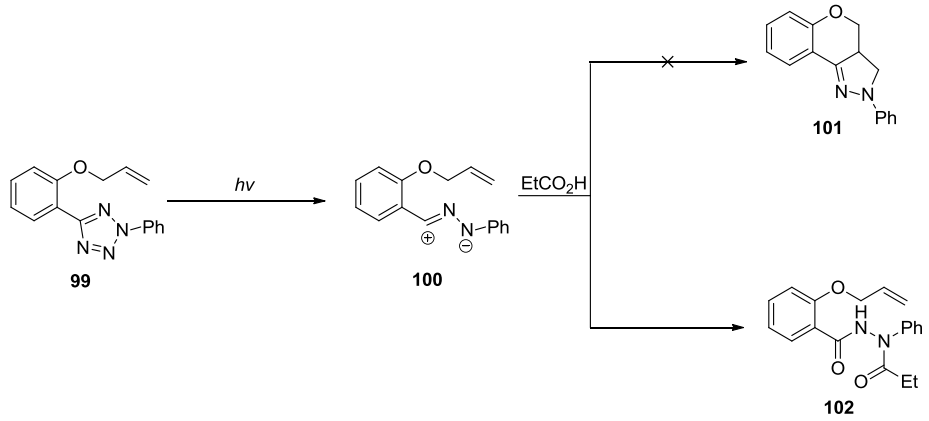


Figure 26: Proposed mechanism for the rearrangement reaction of carboxylic acids with nitrile imines.

While attempting to form tricyclic system **101** from 2,5-diaryl tetrazole **99**, Heimgartner and Heinzelmann observed unexpected intermolecular quenching of nitrile imine **100** by ethanoic acid to give hydrazide **102**, despite an alkene being present within the same molecule. This could be a result of poor HOMO-LUMO overlap associated with the use of a relatively electron neutral alkene and tetrazole.



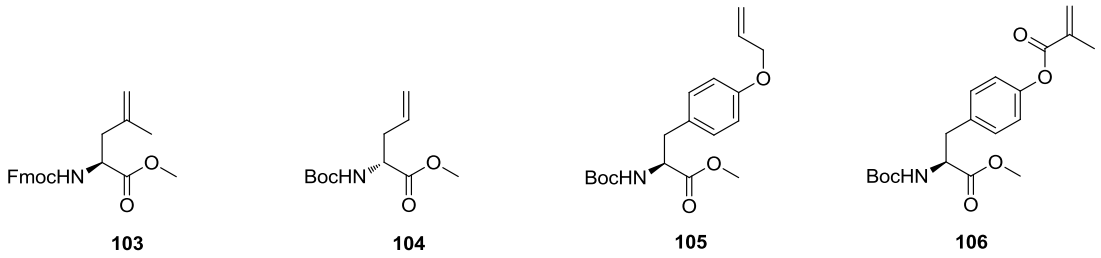
Scheme 13: Intermolecular hydrazide formation observed in the presence of an intramolecular alkene during irradiation of 2,5-diaryl tetrazole **99**.

The formation of hydrazides from carboxylic acids and 2,5-diaryl tetrazoles could represent a new technique for the labelling of peptides prior to their introduction to cells and will be discussed in further detail in Section 1.6.2. More importantly for the current study, the ability of the tetrazole to undergo this rearrangement reaction highlighted that side reactions might occur with biomolecules in cells. The biomolecule of primary concern in this instance was glutathione which warranted further investigation and is discussed in Section 1.2.5.

1.2.3. Reaction of *N*- and *C*- termini protected amino acid tags

After having confirmed that amino acid tags bearing unprotected *C*-termini were unsuitable for synthesis of pyrazoline adducts the focus of the research was redirected to amino acids bearing protecting groups on both the *C*- and *N*- termini.

A new set of tool molecules was identified for further investigations. These included Fmoc-dehydro-Leu-OMe **103**, Boc-allyl-Gly-OMe **104**, Boc-allyl-Tyr-OMe **105**, and Boc-methacryloyl-Tyr-OMe **106** (Figure 27). The latter two compounds contained more electron deficient terminal alkenes. It was reasoned that optimum reaction rates would be achieved with the use of electron poor dipolarophiles bearing electron withdrawing groups due to the improved HOMO-LUMO overlap, as discussed previously.

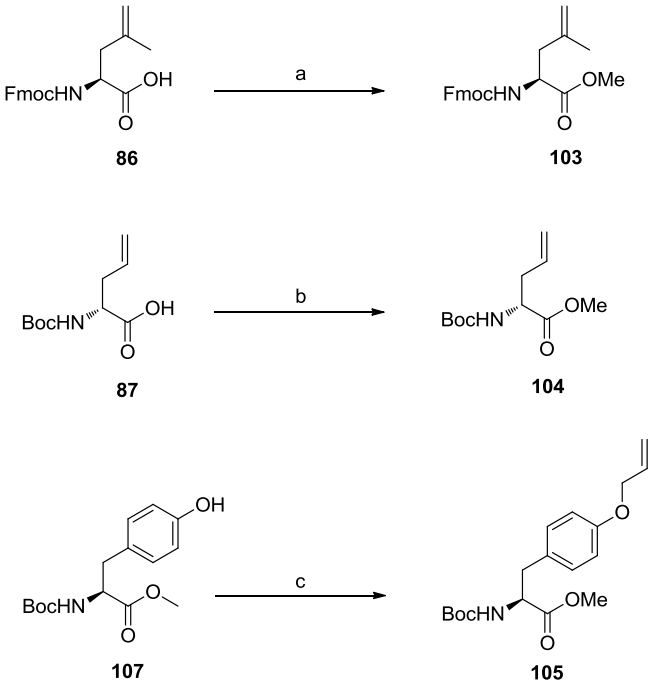


103 **104** **105** **106**

Figure 27: Additional tool molecules selected to investigate the synthesis of pyrazoline derivatised amino acids.

Commercially available Fmoc-dehydro-Leu-OH **86**, Boc-D-allyl-Gly-OH **87**, and Boc-Tyr(OH)-OMe **107** were used in the synthesis of Fmoc-dehydro-Leu-OMe **103**,

Boc-D-allyl-Gly-OMe **104**, and Boc-allyl-Tyr-OMe **105**, respectively (Scheme 14). Boc-methacryloyl-Tyr-OMe **106** was already available in our laboratories.[106] The use of Fmoc protection and *D*-stereochemistry for **103** and **104**, respectively, was based on their immediate commercial availability. It was reasoned that these alterations would not affect the alkene reactivity in the fluorogenic click reaction.

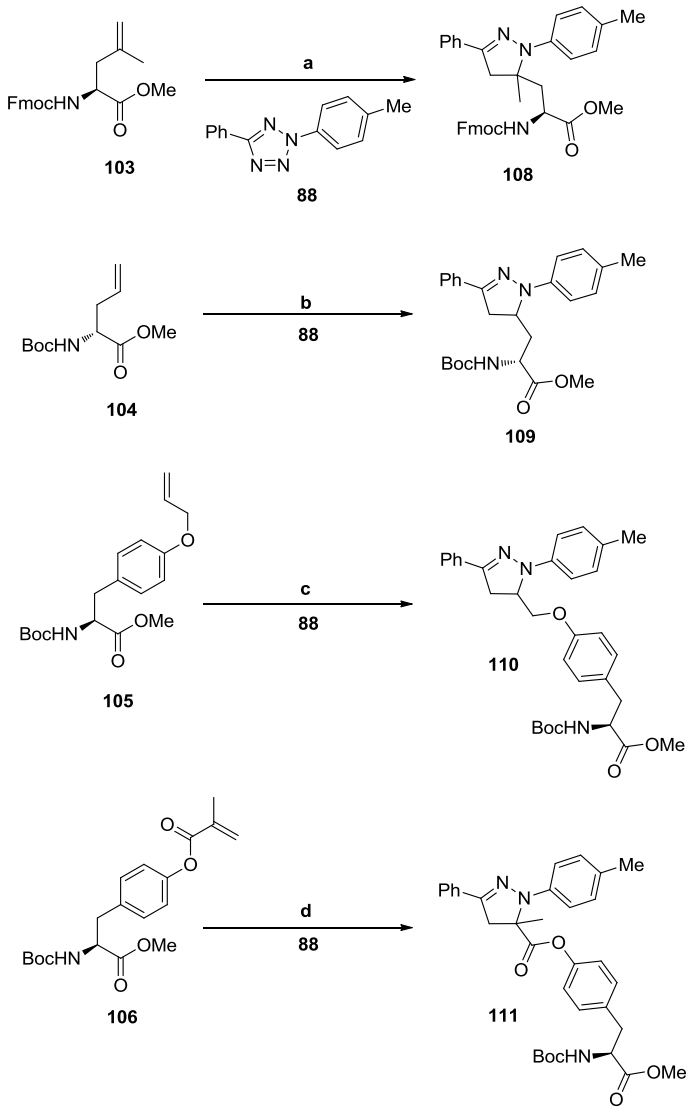


Scheme 14: Synthesis of modified amino acids for use in the fluorogenic click reaction. Reagents and conditions: (a) DL-10-camphorsulfonic acid (0.7 equiv.), MeOH, 70 °C, 3 h, 79%; (b) K_2CO_3 (2 equiv.), MeI (2 equiv.), Me_2CO, 60 °C, 16 h, 66%; (c) allyl bromide (1.2 equiv.), K_2CO_3 (2 equiv.), TBAB (0.1 equiv.), DMF, rt, 16 h, 75%.

The methyl ester analogues were prepared according to literature procedures.[107,108] Fmoc-dehydro-Leu-OMe **103** was prepared with a DL-10-camphorsulfonic acid catalyst in methanol to afford the desired product in 79% yield. In the presence of the acid labile Boc group of **87,** the methyl ester was formed by simple deprotonation of the carboxylic acid with potassium carbonate and S_N2 substitution with methyl iodide to afford **104** in 66% yield. Boc-allyl-Tyr-OMe **105** was prepared by the allylation of Boc-Tyr(OH)-OMe **107**. The S_N2' substitution of allyl bromide

using potassium carbonate and tetrabutylammonium iodide as an additive was chosen from the literature to afford **105** in 75% yield.[109]

The synthesis of pyrazoline adducts from amino acids **103-106** was investigated with tetrazole **88** (Scheme 15). These reactions were conducted to gauge the relative reactivity of the increasingly electron poor alkenes and for the isolation of the desired pyrazoline adducts.



Scheme 15: Reaction of tetrazole **88** with *N*- and *C*- terminus protected amino acid tags. Reagents and conditions: Tetrazole **88** (1 equiv.), UV light (270-330 nm), rt, 1 h; (a) MeCN, unisolated; (b) MeOH:MeCN (1:1 v/v), unisolated; (c) MeCN, 6%; (d) MeCN, 92%.

The reactions were conducted on a small scale and conversions were assessed by LC-MS. The syntheses of pyrazoline adducts **110** and **111** were carried out by other members of our laboratory.[110, 111] In each of the reactions a consistent byproduct arising from dimerisation of the nitrile imine intermediate generated from **88** (*vide infra)* was observed in varying amounts. The reactivity was found to vary significantly but followed the expected pattern, in the order **106** > **105** > **104** > **103**. This order of reactivity was determined by comparing the proportion of the product peak to that of the byproduct peak in the LC-MS chromatogram (Table 2).

Tetrazole	Tag	Adduct	Time (P)/ min	Abundance (P)/ %	Time (B)/ min	Abundance (B)/ %	Ratio (P/B+P)
88	**103**	**108**	1.52	4.6	1.58	22.6	0.17
88	**104**	**109**	1.42	13.9	1.59	31.9	0.30
88	**105**	**110**	1.50	33.1	1.57	21.0	0.61
88	**106**	**111**	1.34	100.0	-	0.0	1.00

Table 2: Relative proportions of pyrazoline product (P) to dimer byproduct (B) as observed by LC-MS.

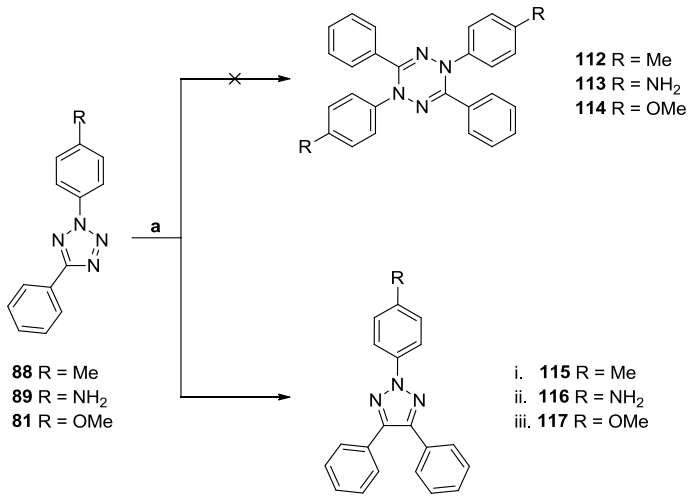
The reactivity can be explained as a result of both electronic and steric effects. The electron withdrawing carbonyl group in **106** increases the reactivity mesomerically and the electron withdrawing oxygen atom in **105** increases the reactivity inductively, to a lesser extent. The result is an improved HOMO-LUMO overlap between the photochemically induced nitrile imine and these alkene tags. The electronically neutral alkenes are both less reactive, with the additional methyl group in **103** providing greater steric hindrance than for **104**. This results in poorer HOMO-LUMO overlap and a lower ratio of product to byproduct for these alkenes.

These data suggested that, of the four tags, only *N*-Boc-allyl-Tyr-OMe **105** and *N*-Boc-methacryloyl-Tyr-OMe **106** had suitable reactivity for the synthesis of pyrazoline adducts, with the latter giving significantly greater conversion.

The formation of tetrazine byproducts via dimerisation of the nitrile imine intermediate had been reported to be the major byproduct observed during

pyrazoline adduct formation.[80,86] From this evidence, the observed byproduct was predicted to be tetrazine **112** (Scheme 16).

However, despite claiming to have isolated the tetrazine dimer from reactions involving electron rich alkenes, no analytical data were presented in these earlier reports. When the byproduct of the reaction with tetrazole **88** was intentionally synthesised by irradiating the lone tetrazole it was instead determined to be 2,4,5-triaryl-1,2,3-triazole **115**, with no evidence of the proposed tetrazine byproduct present in the reaction mixture.



Scheme 16: Synthesis of *N*2-substituted 1,2,3-triazoles via the self coupling of nitrile imine intermediates. Reagents and conditions: (a) i. UV light (270-330 nm), EtOAc, rt, 12 h, 84%; ii. UV light (270-330 nm), DCM:MeOH (1:1 v/v), rt, 28 h, 84%; iii. UV light (270-330 nm), EtOAc, rt, 3 h, 43%.

The formation of 2,4,5-triaryl-1,2,3-triazoles from 2,5-diaryl tetrazoles was not limited to tetrazole **88**, as is evident from the synthesis of triazoles **116** and **117** from available tetrazoles **89** and **81**, respectively. Isolation and analysis by NMR confirmed that the structure of the molecules did not correspond to the tetrazines **112**-**114** but were indeed elucidated as triazoles **115**-**117**.

The $[M+H]^+$ values observed by LC-MS for each triazole was the initial indication that the expected tetrazine molecules were not the products. Additionally, the ratio of aromatic protons to methyl protons for **115** and **117** were 14:3 instead of the

expected 18:3. The aromatic peaks were indicative of a symmetrical structure with a single *para* disubstituted ring and two phenyl rings.

A number of key interactions were observed by multinuclear NMR (see Appendix, Figure A2). Analysis by $^{1}H^{15}N$ HMBC NMR showed a correlation between the *N*2 atom and the protons in the *ortho* positions of the linked aromatic ring.[112] For **116**, an additional correlation was noted at approximately 55.5 ppm, representing the interaction between the aniline nitrogen atom and the nearby *ortho* protons of the ring (Figure 28). Finally, a $^{1}H^{1}H$ COSY spectrum showed a correlation between these two sets of *ortho* protons which confirmed that the aryl group attached to the *N*2 atom was indeed the aniline ring.

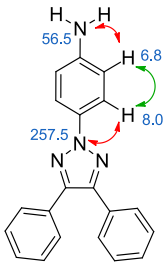


Figure 28: NMR interactions observed for 2,4,5-triaryl-1,2,3-triazole **116** by $^{15}N^{1}H$ HMBC NMR (red) and $^{1}H^{1}H$ COSY NMR (green).

Data obtained for triazole **117** were compared to a literature report of the same compound and were almost identical in the majority of techniques utilised.[113] A discrepancy between the observed and literature *N*2 chemical shifts was noted; however, the similar ^{15}N HMBC values obtained for the other triazoles suggests that the value obtained in the current study is correct. A second literature report relating to triazole **117** also agreed with the data obtained, although a ^{15}N NMR analysis was not conducted.[114] Literature comparisons for triazole **115** and **117** were also consistent with the data generated, except for the $^{15}N^{1}H$ HMBC.[113,115] This interesting synthetic approach to 2,4,5-triaryl-1,2,3-triazole systems warranted further investigation and is discussed in greater depth in Section 1.6.1.

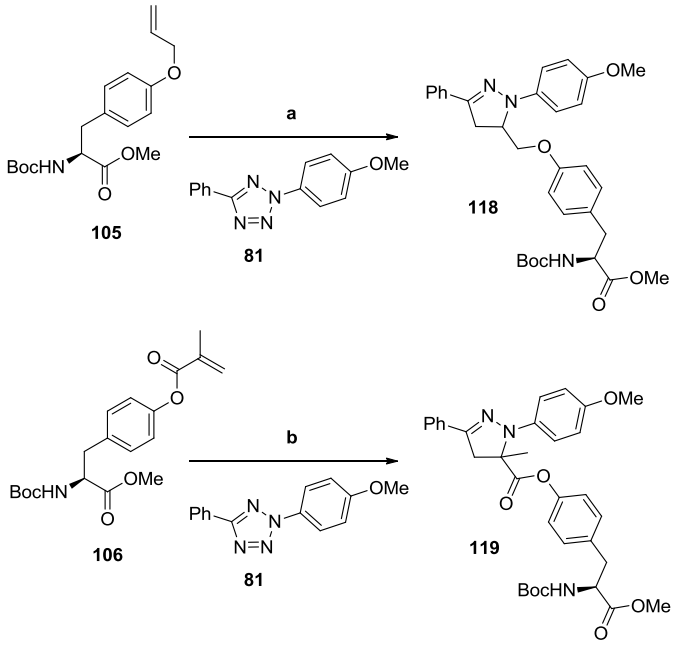
For the current study, the formation of such triazole byproducts during the fluorogenic click reaction further highlighted the possibility of additional side reactions competing with the desired reaction pathway. It was predicted that the byproduct formation could be minimised by switching to a more electron rich tetrazole which would have improved HOMO-LUMO overlap with the alkene tagged amino acids. The reactivity of these alkene tags with the more electron rich tetrazole **81** was next considered.

1.2.4. [3+2] cycloaddition reactions with an electron rich tetrazole

The failure to synthesise the model pyrazoline adducts **109** and **110** was attributed to the poor HOMO-LUMO overlap of the electronically neutral alkenes with tetrazole **88**, which has only limited electron donation from the methyl substituent. Tetrazole **81** was investigated next due to its increased mesomeric electron donation.

As discussed previously, tetrazole **81** showed the second fastest rate of reaction, based on kinetic studies conducted by Lin *et. al*.[80] Furthermore, this analogue had already been used in a reaction with protein-tagged allyl-Tyr in an *E.coli* cell and, therefore, demonstrated potential for use in a cellular environment.[80] As a result, tetrazole **81** was initially reacted *N*-Boc-allyl-Tyr-OMe **105** (7.5 equiv.) and *N*-Boc-methacryloyl-Tyr-OMe **106** (1 equiv.) to give pyrazoline adducts **118** and **119**, respectively (Scheme 17). The use of excess alkene for conversion of **105** to **118** was necessitated by the low reactivity of the allyl-Tyr tag.

In both reactions the pyrazoline adduct was successfully isolated. The high reactivity of *N*-Boc-methacryloyl-Tyr-OMe **106** gave pyrazoline **119** in 80% conversion by LC-MS, though an isolated yield of 29% was obtained due to loss of material during purification. The lower reactivity of *N*-Boc-allyl-Tyr-OMe **105** resulted in only 46% conversion by LC-MS and 41% isolated yield even with a large excess of the alkene tag. The products were easily identifiable by LC-MS and NMR.



Scheme 17: Reaction of allyl and methacryloyl tyrosine derivatives with tetrazole **81**. Reagents and conditions: (a) Boc-allyl-Tyr-OMe **105** (7.5 equiv.), tetrazole **81** (1 equiv.), UV light (270-330 nm), MeCN, rt, 1 h, 41%; (b) Boc-methacryloyl-Tyr-OMe **106** (1 equiv.), tetrazole **81** (1 equiv.), UV light (270-330 nm), MeCN, rt, 1 h, 29%.

The protons in the newly formed ring were characteristic of the pyrazoline adducts and a number of key interactions were clearly identifiable from the $^{1}H^{1}H$ COSY NMR spectra (see Appendix, Figure A3). For adduct **118** a clear correlation can be seen between the ring-bound methylene group and the adjacent methane (Figure 29a). Additionally, the correlation between the ring-bound methine proton and the exocyclic methylene group was evident. The splitting pattern is as expected, with small $^{3}J_{cis}$ and large $^{3}J_{trans}$ coupling values observed as a result of the axial/equatorial configuration of the ring bound protons and the geminal nature of the exocyclic methylene protons. The chemical shifts of the protons are also consistent with the expected regioisomer as evident by the downfield shift of the ring-bound CH proton relative to the ring-bound CH_2 protons as a result of de-shielding by the adjacent nitrogen atom.

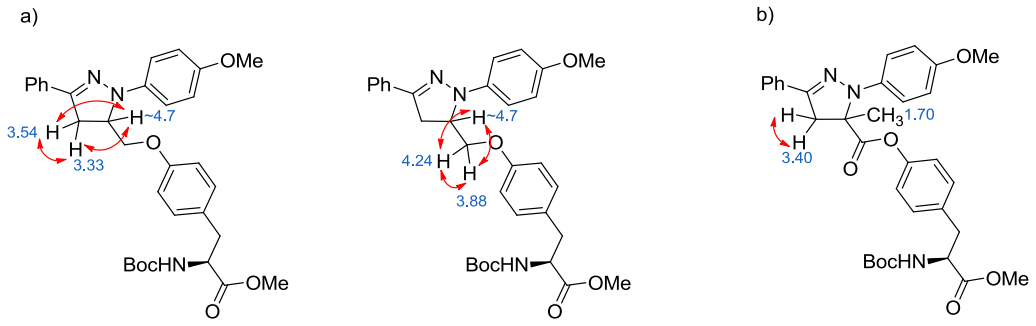
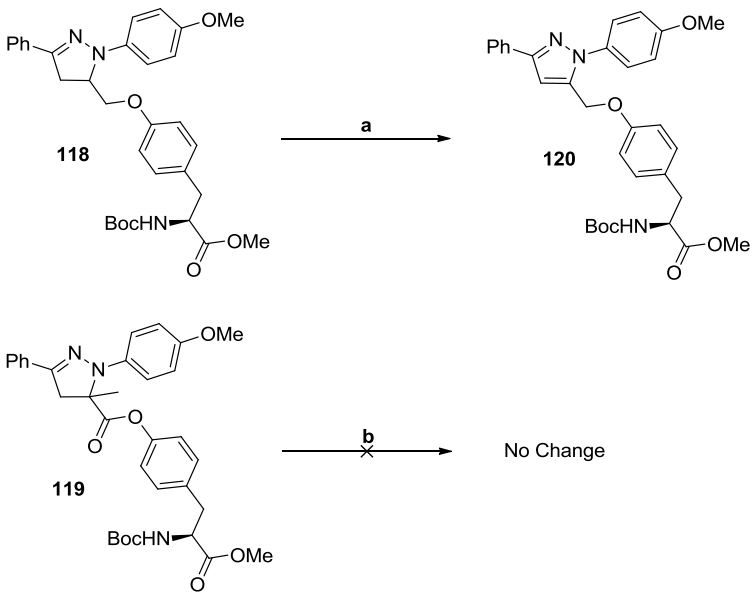


Figure 29: $^{1}H^{1}H$ COSY NMR interactions observed for (a) pyrazoline adduct **118** and (b) pyrazoline adduct **119**.

For pyrazoline adduct **119** quaternising the *C*5-position results in a simplified $^{1}H^{1}H$ COSY spectrum (see Appendix, Figure A4). Characteristic correlations could be observed between the axial and equatorial ring-bound methylene protons in addition to the presence of a singlet for the methyl group (Figure 29b). The expected regioisomer was obtained, as indicated by the chemical shifts which are similar to those observed for **118**.

During NMR analysis of the pyrazoline adducts a ^{13}C NMR spectrum of adduct **119** was easily obtained whereas an accurate ^{13}C NMR spectrum could not be obtained for adduct **118** due to the appearance of new peaks within a few hours of dissolution in deuterated chloroform. This suggested that the pyrazoline adducts may not be stable over short periods of time, which could adversely affect any fluorescence observed during a cell based assay. Therefore, additional NMR spectra of each pyrazoline adduct were taken one hour after dissolution in deuterated chloroform, followed by another NMR after a specified period of time. The results of this study are depicted in Scheme 18.



Scheme 18: Observed changes to pyrazoline adducts **118** and **119**. Reagents and conditions: (a) ambient light, air, $CDCl_3$, rt, 24 h, 98%; (b) ambient light, air, $CDCl_3$, rt, 10 days.

Pyrazoline **118** was determined to undergo oxidation to pyrazole **120** within a 24 hour time period under ambient conditions. Analysis of the sample by LC-MS showed a loss of m/z = 2 whilst analysis by NMR indicated a clear loss of both the pyrazoline methylene and methine signals. Instead, a new single was observed at δ_H = 6.85 ppm for the pyrazole aromatic proton, accompanied by a shift of most of the original aromatic signals (Figure 30). Furthermore, a singlet observed at δ_H = 4.98 ppm is attributable to the OCH_2 group, which was no longer diastereotopic.

Comparison of the ^{13}C NMR spectra before and after showed the appearance of a peak at δ_C = 105.8 ppm, confirmed as a new pyrazole aromatic moiety by $^{1}H^{13}C$ HSQC NMR, and another peak at δ_C = 117.6 ppm, confirmed as the quaternary carbon of the pyrazole by its absence in a DEPT-135 NMR spectrum. A similar experiment carried out with **119** showed no change after 10 days which is consistent with the methyl group preventing oxidation of the pyrazoline core.

a)

b)

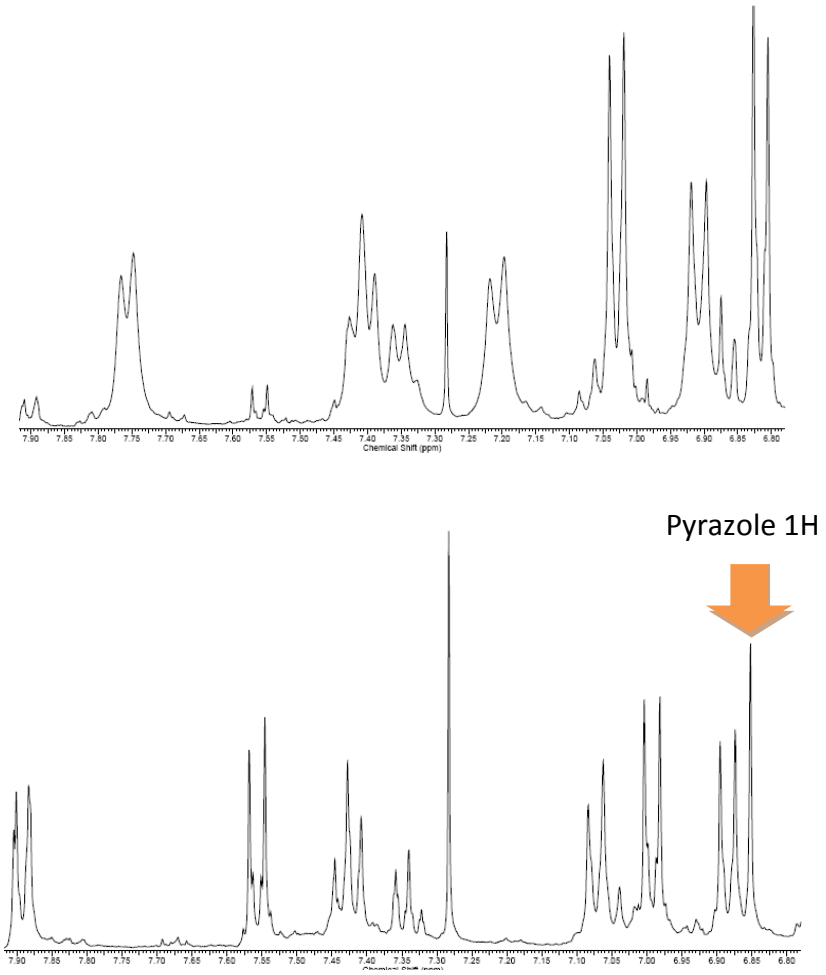


Figure 30: Observed change in aromatic peaks representing the conversion of pyrazoline **118** to pyrazole **120** after (a) 1 h and (b) 24 h.

As can be noted from the spectra taken for the conversion of pyrazoline **118** after 1 hour there are small peaks present corresponding to the pyrazole byproduct. This indicated that the oxidation began immediately but conversion to pyrazole **120** was slow. Due to the slow nature of the oxidation it was determined that the reaction between tetrazole **81** and alkene tagged amino acids **105** and **106** could still be used for the intracellular observation of cell permeable peptides, provided that any fluorescence measurements could be carried out within an appropriate time frame. To this end, the experiment would need to be conducted at a concentration suitable for a cellular environment and the fluorescence monitored over a period of time in order to determine if the product could be detected by fluorescence measurements

within a suitable time frame to prevent the oxidation reaction from interfering with the results.

The successful synthesis of pyrazolines **118** and **119** confirmed that tetrazole **81**, *N*-Boc-allyl-Tyr-OMe **105** and *N*-Boc-methacryloyl-Tyr-OMe **106** were suitable tool molecules to further investigate the intracellular fluorogenic click reaction. The next step was to identify a final set of tool molecules which could be used to investigate the potentially problematic hydrazide formation that could occur with biomolecules such as glutathione.

1.2.5. Glutathione: A potential inhibitor of the [3+2] cycloaddition reaction

Glutathione **121** (Figure 31) is an important anti-oxidant involved in cell regulatory processes such as the quenching of free radicals, reduction of peroxides, recycling of Vitamin C from oxidised to reduced form, and a range of conjugation reactions.[116]

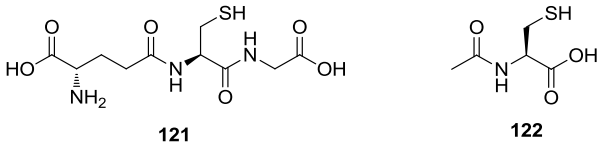
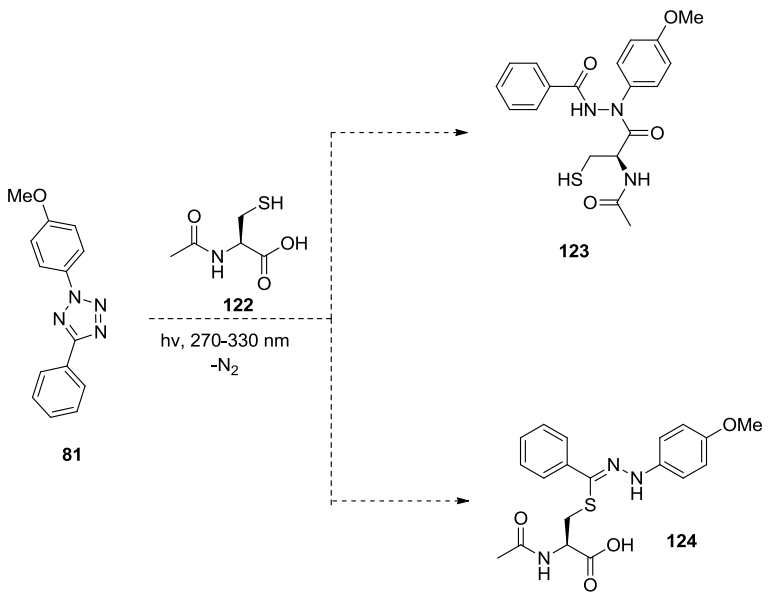


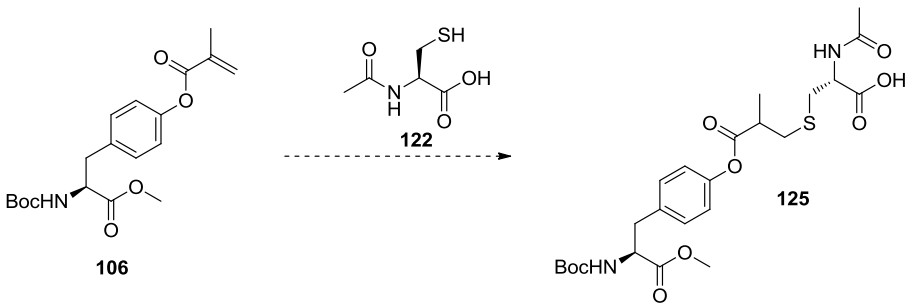
Figure 31: Structures of glutathione **121** and *N*-acetyl cysteine **122**.

N-acetyl cysteine **122** is a simplified glutathione analogue which contains the reactive thiol and carboxylic acid groups, and therefore, can be used as a mimic of the tripeptide. Due to its low molecular weight and improved organic solubility, relative to glutathione, this mimic was selected as a suitable reagent to investigate undesirable side reactions prior to cell experiments. The possible quenching reactions between *N*-acetyl cysteine **122** and tetrazole **81** are shown in Scheme 19.



Scheme 19: Predicted side reactions of tetrazole **81** with *N*-Ac-Cys-OH **122**.

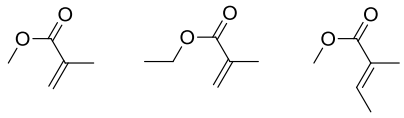
Quenching of the photochemically generated nitrile imine via hydrazide formation to give **123** or via nucleophilic attack of the thiol to give hydrazonothioate **124** could occur with *N*-Acetyl cysteine **122**. These potentially problematic side reactions were examined in greater detail and are discussed in Section 1.4.1. Of more immediate concern was the potential Michael addition of the thiol group to the alkene tags, particularly for electron poor Boc-methacryloyl-Tyr-OMe **106** (Scheme 20).



Scheme 20: Potential Michael addition reaction between Boc-methacryloyl-Tyr-OMe **106** and *N*-Acetyl Cysteine **122**.

Having stated this, methacrylate groups are known to be relatively unreactive towards Michael addition with glutathione, compared to other α,β-unsaturated carbonyl compounds. For example, the reactivity of methyl methacrylate **126** and

ethyl methacrylate **127** has been investigated along with 64 other α,β-unsaturated carbonyl compounds (Figure 32).[117] The experimentally determined rate of conjugate addition with glutathione was log k_{GSH} = -1.14 M^{-1} min^{-1} and -1.24 M^{-1} min^{-1}, respectively (at 25 °C, pH 7.4), marking these groups as two of the least reactive α,β-unsaturated carbonyl compounds. Of the remaining examples only methyl tiglate **128** had a slower reaction rate of -2.15 M^{-1} min^{-1}. While 10 fold less reactive, the methyl tiglate group was not considered as a potential tag due to the additional methyl group which may retard the rate of the cycloaddition reaction for both steric and electronic reasons.

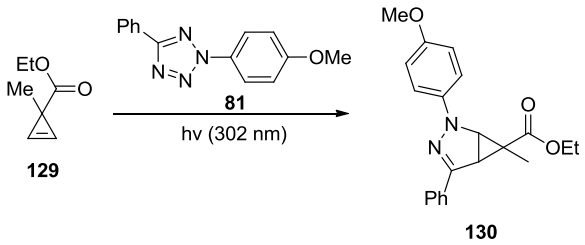


126 **127** **128**

Figure 32: Structures of methyl methacrylate **126**, ethyl methacrylate **127**, and methyl tiglate **128**.

Despite these data, glutathione can be present in cells in high concentration, thus increasing the likeliness of the undesired quenching of the alkene tag. The level of glutathione varies between cell types but is often in the millimolar region, reaching up to 10 mM in liver cells, for example.[116,117]

The use of cyclopropene groups as tags for intracellular labelling reactions has been investigated independently by Lin *et. al.*, Devaraj *et. al.* and Patterson *et. al.* [73,118,119] These moieties are attractive as they are resistant to Michael addition and can be used as a control to determine if this reaction is sequestering the alkene from the desired cycloaddition reaction. Furthermore, the Lin group have demonstrated a successful fluorogenic click reaction between cyclopropene **129** and tetrazole **81** to give pyrazoline adduct **130** (Scheme 21); a reaction which was found to be sixty times more rapid than the reaction of **81** with allyl-Tyr (k = 58 M^{-1} s^{-1} and 0.95 M^{-1} s^{-1}, respectively).[118]



Scheme 21: Literature synthesis of pyrazoline adducts from strained cyclopropenes.[118]

Encouragingly, it was demonstrated that cyclopropene **129** could be attached to amino acids such as lysine via hydrolysis and subsequent amide coupling reactions. Once formed, this tagged amino acid was determined to be highly resistant to glutathione, with only 5% being conjugated by Michael addition in a 60 hour period.[118] This resistance, and the increased cycloaddition reactivity, result from ring strain and the orbital overlap of the cyclopropene atoms. The alkene carbon atoms *C*1 and *C*2 are sp^2 hybridised with one *p*-orbital contributing to the ring and the other contributing to the double bond, while the *C*3 atom is also sp^2 hybridised but only contributes to the orbital overlap within the ring (Figure 33).[120]

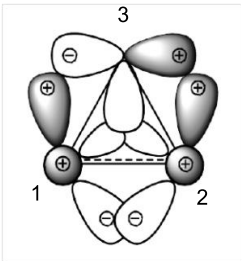
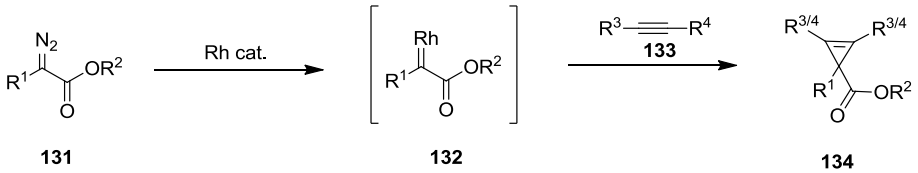


Figure 33: Graphical representation of the molecular orbitals in cyclopropene derivatives.

Functionalised cyclopropenes **134** are often synthesised using a cyclopropenation reaction between alkynes and metal activated diazo species **131** (Scheme 22).[121] For example, a catalytic rhodium dimer can be applied to initiate loss of nitrogen from the diazo group and form a Rh(II) carbene **132**.[122] This carbene is able to undergo cyclisation with a range of alkene or alkyne groups **133**, a reaction that is reported to proceed in high to excellent yields.[121]

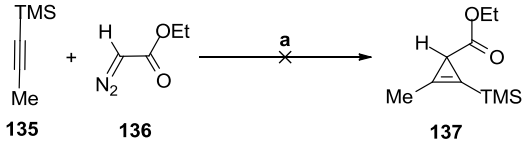


131 132 134

Scheme 22: Synthesis of cyclopropenes **134** from diazo species **131** via a rhodium carbene **132**.

1.2.5.1. Synthesis of cyclopropene tagged tool molecules

The initial synthesis of a 1,2,3-substituted cyclopropene **137** was attempted from readily available 1-(trimethylsilyl)-propyne **135** and ethyl 2-diazoacetate **136** (Scheme 23). The position of the ester would allow for subsequent conjugation to a peptide via hydrolysis and amide bond coupling.



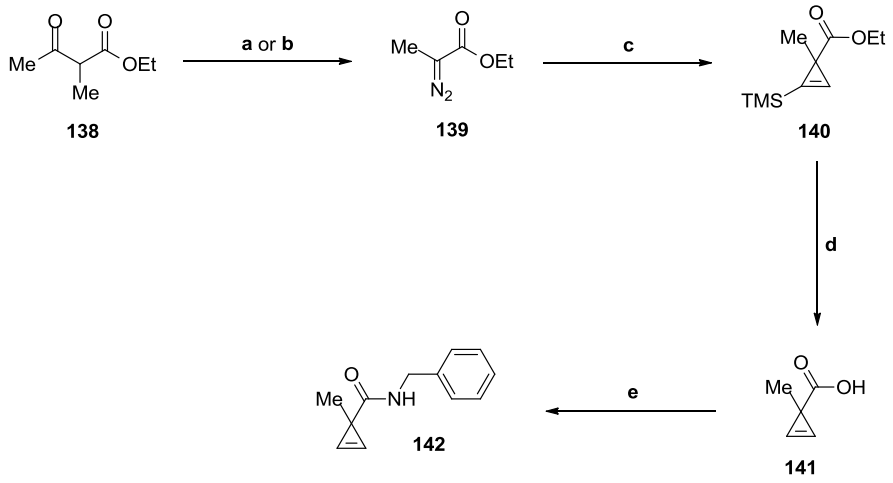
135 136 137

Scheme 23: Unsuccessful synthesis of cyclopropene **137** via a rhodium catalysed cyclopropenation. Reagents and conditions: (a) $Rh_2(OAc)_4$ (1 mol %), trimethyl(prop-1-yn-1-yl)silane (3 equiv.), DCM, rt, 16 h.

The formation of cyclopropene **137** was observed by LC-MS along with a significant number of byproducts. Isolation was unsuccessful due to poor separation of the byproducts and volatility of the desired product. The synthesis of small cyclopropene molecules is known to be challenging due to their inherent volatility and instability.[123]

The poor reaction profile observed for the cyclopropenation may be due to the instability of the cyclopropene resulting from only a single substitution at the *C*3 position. The importance of substitution at this position was first reported in 1963 by Closs *et. al.* who noticed an increase in room temperature stability of cyclopropenes bearing *C*3 substitution in the order of *geminal*-H,H < H,Me < *geminal*-Me,Me.[124] Therefore the synthesis of the analogous

cyclopropene carboxylic acid **141** was reasoned to give a greater probability of success via a previously established route (Scheme 24).[118]



Scheme 24: Synthetic route to 3,3-disubstituted cyclopropene **142**. Reagents and conditions: (a) Tf_2O (2 equiv.), NaN_3 (4 equiv.), TBAB (0.02 equiv.), 2M $NaOH_{(aq)}$ (2 equiv.), Hexanes:MeCN (1:1 v/v), 0 °C, 2 h, crude; (b) *p*-ABSA (1.5 equiv.), DBU (1.5 equiv.), MeCN, 0-25 °C, 2.5 h, 50%; (c) $Rh_2(OAc)_4$ (1 mol %), $HCC(SiMe_3)$ (4 equiv.), DCM, rt, 24 h, crude; (d) 1.5M $KOH_{(aq)}$ (7.7 equiv.), MeOH, rt, 20 h, 4% (two steps); (e) EDC (1.2 equiv.), HOAt (1.2 equiv.), benzylamine (2.4 equiv.), DIPEA (1 equiv.), DCM, rt, 16 h, 59%.

For the purpose of investigating the [3+2] cycloaddition reaction, the amide coupling with benzylamine to give cycloprop-2-ene amide **142** was desirable to give a strong chromophore for analysis by LC-MS.

The addition of a 3-methyl substituent for improved stability required the synthesis of ethyl diazopropanoate **139** from commercially available ethyl-2-methylacetoacetate **138** via a diazo transfer reaction. Synthesis via route (a) in Scheme 24 led to formation of the desired diazo product, as observed by NMR; however, the work up procedure was unsuccessful as the product was found to be volatile and could not be separated from the mixture of solvents used.

Gratifyingly, the synthesis and purification of ethyl diazopropanoate **139** by route (b) was successful, giving the desired product in 50% yield, similar to that previously reported.[125] This less complicated synthetic method involved the use of stable

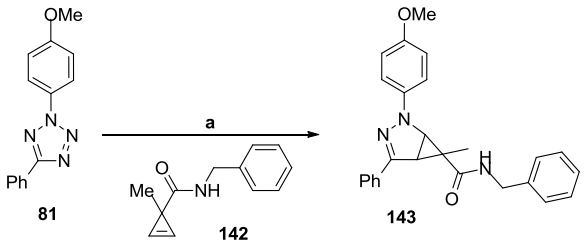
p-acetamidobenzenesulfonyl azide as a diazo source, without requiring its generation *in situ*. Acetonitrile was still required as a reaction solvent; however, with the isocratic solvent present it was possible to isolate the product with minimal loss of yield resulting from its volatility.

The synthesis of **141** was carried out by combining steps (c) and (d), as suggested for an analogous cyclopropene carboxylic acid preparation.[126] This allowed the synthesis of the stable solid **141** without isolation of intermediate cyclopropene ester **140**. Removal of the catalyst was immediately followed by ester hydrolysis with MeOH:DCM (1:1 v/v) and aqueous potassium hydroxide. As predicted, the product was obtained as a white solid, albeit with a yield of only 4% for the two steps. This is likely due to the unclean cyclopropenation reaction and the resulting presence of impurities during the hydrolysis step.

The amide bond formation proceeded smoothly to afford cyclopropene **142** which was confirmed to be stable in $CDCl_3$ under ambient conditions for a period of at least 72 hours by NMR analysis. Despite the low yield, sufficient quantities of **142** were prepared to test the [3+2] cycloaddition reaction.

1.2.5.2. [3+2] cycloaddition reactions with the cyclopropene tag

Having successfully synthesised pyrazoline adducts **118** and **119** by reaction of alkenes **105** and **106** with tetrazole **81** the analogous reaction was carried out with cyclopropene tag **142** (Scheme 25). The synthesis of **143** was carried out successfully in 88% yield. Excess cyclopropene **142** was maintained by adding small quantities of tetrazole **81** until the cyclopropene reagent was fully consumed.



Scheme 25: Synthesis of pyrazoline adduct **143** from cyclopropene **142**. Reagents and conditions: (a) cyclopropene **142** (3 equiv.), UV light (270-330 nm), DCM:MeOH (1:1 v/v), rt, 2.5 h, 88%.

As expected, the pyrazoline protons could be observed as an isolated unit (Figure 34), only showing correlations to each other by $^{1}H^{1}H$ COSY (see Appendix, Figure A5).

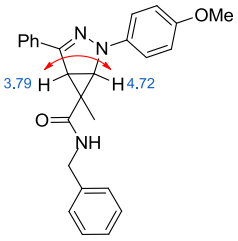
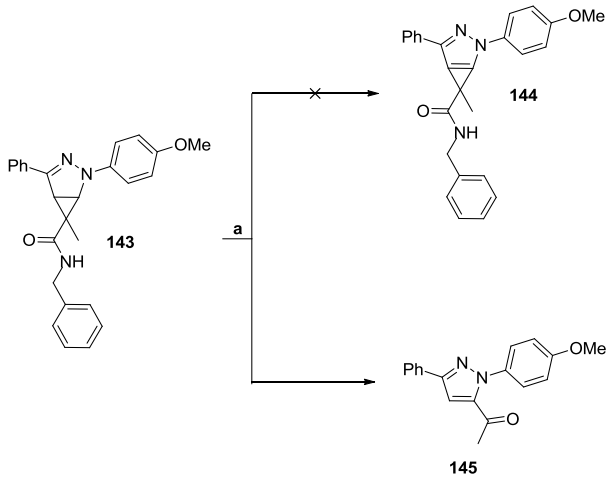


Figure 34: ^{1}H COSY NMR interactions observed for pyrazoline adduct **143.**

Similarly to pyrazoline adduct **118**, the pyrazoline adduct formed from cyclopropene tag **142** and tetrazole **81** was observed to oxidise in deuterated chloroform over an extended period of time (Scheme 26). However, instead of forming the expected pyrazole **144** the molecule was observed to degrade to pyrazole **145**.[127]



Scheme 26: Observed changes to pyrazoline adduct **143**. Reagents and conditions: (a) ambient light, air, $CDCl_3$, rt, 32 h, 80%.

This conversion was confirmed by observed changes in the aromatic region of the NMR spectra (Figure 35). A new singlet peak is observed at δ_H = 8.38 ppm for the pyrazole CH group, in addition to which all other non-aromatic peaks are lost, with the exception of the methoxy and methyl protons. The latter of which was observed to shift from δ_H = 0.8 ppm to 2.39 ppm.

a)

b)

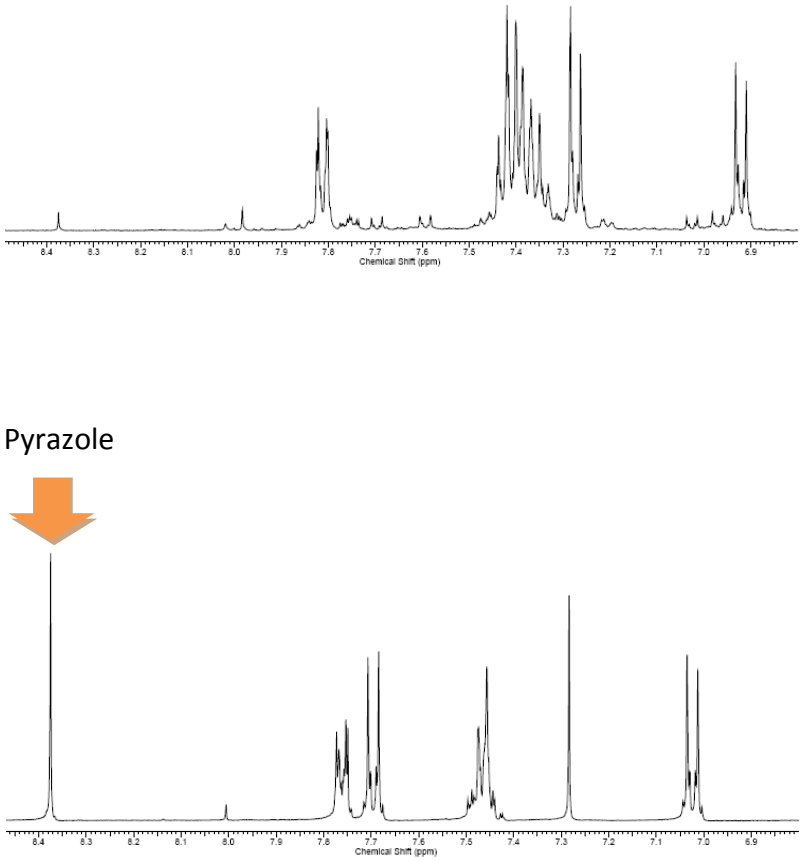


Figure 35: Observed change in aromatic peaks representing the conversion of pyrazoline **142** to pyrazole **145** after (a) 1 h and (b) 32 h.

The structure of the pyrazole **145** was confirmed by 2D NMR analysis (see Appendix, Figure A6). As can be seen from Figure 36 a number of key interactions were identified. Firstly, the methoxy protons were linked to their bonded carbon atom and the *ipso* carbon of the aromatic ring. As expected, the pyrazole proton showed two bond correlations to the *C*3 atom in the pyrazoline ring. A three bond correlation to the carbonyl signal could not be observed;

however, a weak three bond correlation between the acetyl protons and the *C*5 atom confirmed the structure since a two bond correlation between the pyrazole proton and C*5* atom could be observed, thus connecting these key signals.

a) b)

Figure 36: NMR characterisation of pyrazole **145**. (a) Assigned structure indicating single bond $^{1}H^{13}C$ couplings; (b) Assigned structure indicating two (red) and three (green) bond $^{1}H^{13}C$ couplings.

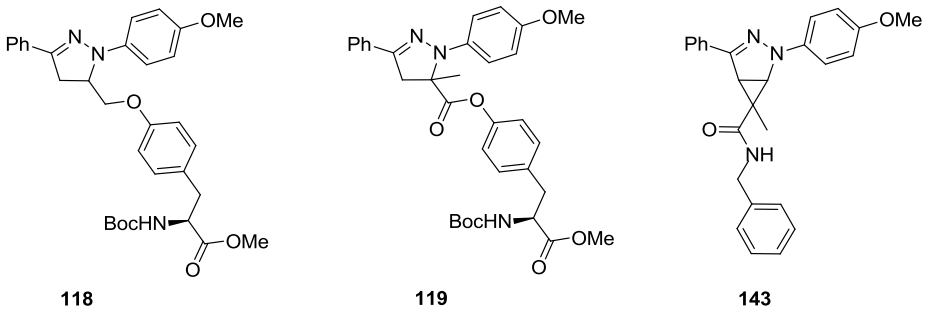
As was observed for pyrazoline adduct **118**, the oxidative degradation of pyrazoline adduct **143** occurred over an extended period of time. This again indicated that a cell based reaction would have to occur within a timescale that would not be deleteriously affected by this side reaction. Having successfully obtained the requisite tool molecules it was then necessary to confirm their suitability for use as reagents in the fluorogenic click reaction by conducting a detailed fluorescence analysis.

1.2.6. Fluorescence study

1.2.6.1. Excitation and emission profile of pyrazoline adducts

The fluorescence of isolated pyrazoline adducts **118**, **119** and **143** was observed by obtaining an excitation (λ_{ex} = 210-450 nm, λ_{em} = 470 nm) and emission (λ_{ex} = 350 nm, λ_{em} =360-700 nm) spectrum for each adduct in MeCN:PBS (1:1 v/v) at 0.1 mM concentrations (Figure 37). It should be noted that each of the pyrazoline adducts was stable in its isolated form and was used immediately after dissolution to ensure that oxidation to the corresponding pyrazole did not give anomalous results.

a)



b)

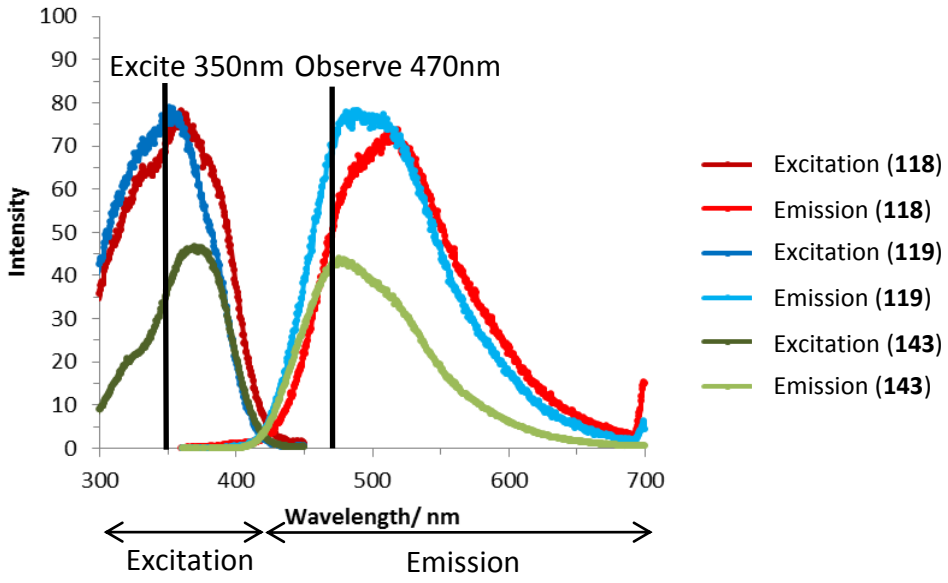


Figure 37: (a) Structure of pyrazoline adducts **118**, **119**, and **143**. (b) Excitation and emission spectra of pyrazoline adducts **118**, **119**, and **143** in MeCN:PBS (1:1 v/v) at 0.1 mM concentrations. Scan width = 3.0 nm, scan speed = 100 nm min^{-1}. Excitation spectra: λ_{em} = 500 nm, λ_{ex} scan = 210-450 nm. Emission spectra: λ_{ex} = 350 nm, λ_{em} scan = 360-700 nm.

The excitation/emission spectra of the pyrazoline adducts show emission occurring in the range of 400-700 nm and excitation occurring in the range below 300-400 nm. Each curve follows the same general profile with pyrazoline **118** and **119** exhibiting similar fluorescence intensity which is larger than that observed for pyrazoline **143**. This is potentially a result of the strained cyclopropane ring inducing slight changes in the backbone of the pyrazoline, leading to a change in its

absorbance profile. The spectra obtained are similar to those obtained for known diaryl substituted pyrazoline adducts.[81]

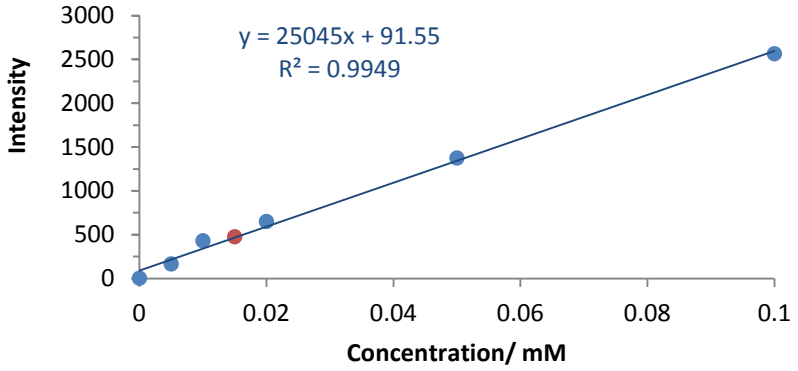
Appropriate wavelengths for excitation (350 nm) and observation (470 nm) were chosen for monitoring the formation of these pyrazoline adducts under suitable conditions for a cell experiment. The wavelengths were chosen based on readily available confocal microscope filters with wavelengths closest to the maxima of the excitation and emission profile of all three pyrazolines.

The fluorescence intensity of pyrazoline adducts **118** and **119** was subsequently investigated at these wavelengths and under varying concentrations to determine if a quantitative analysis would be feasible. The fluorescence intensity exhibited a linear concentration dependence for both pyrazolines with correlations of R^2 = >0.95 each (Figure 38). Adduct **143** was not submitted to this analysis due to the shortage of material available and its difficult synthesis.

The published reactions carried out in *E.coli* cells by Lin *et. al*. were conducted with a large excess of tetrazole (100-250 μM) in the presence of an allyl-Tyr group embedded into a protein (15 μM). Since the assay proposed in this work will require the alkene motif to be attached to a peptide, it would seem logical to adopt the same system in order to minimise the quantity of substrate required for the assay.

From the data presented in Figure 38 it can be seen that the fluorescence intensity at 15 μM is readily observable for both molecules. This indicated that if 100% conversion of the alkene tag to the pyrazoline adduct was achieved at this concentration then the pyrazoline adduct could be detected, provided that any background fluorescence was less prominent. In order to investigate the potential background fluorescence that may be present during the reaction, the fluorescence profiles of the triazole, pyrazole and hydrazide byproducts were also investigated.

a)



b)

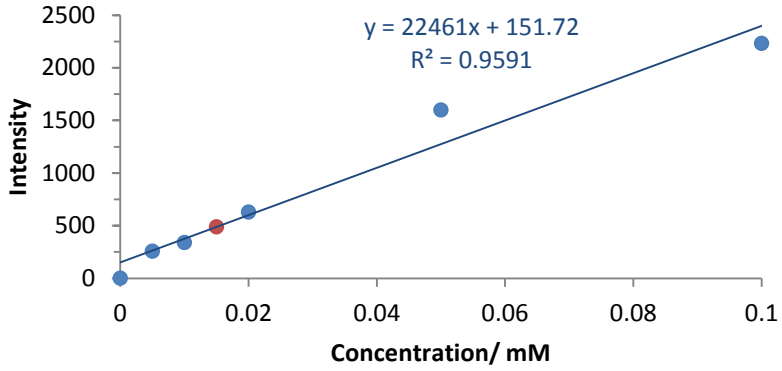


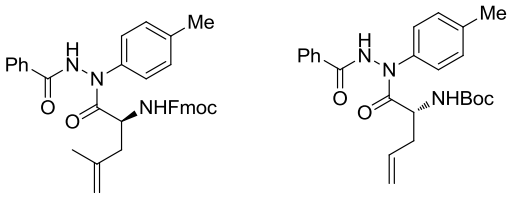
Figure 38: Concentration dependent fluorescence of pyrazoline adducts. (a) Spectrum obtained for pyrazoline **118** in MeCN:PBS (1:1 v/v). λ_{ex} = 350 nm, λ_{em} = 470 nm; (b) Spectrum obtained for pyrazoline **119** in MeCN:PBS (1:1 v/v). λ_{ex} = 350 nm, λ_{em} = 470 nm. Fluorescence intensities at 15 µM are denoted in red.

1.2.6.2. Excitation and emission profile of byproducts

Analysis of the excitation and emission spectra of hydrazides **92** and **93** indicated that only very low levels of fluorescence were observable (Figure 39). No significant difference was observed between the molecules indicating that the nature of the side chain employed did not affect the fluorescence intensity. As such, it can be presumed that hydrazides formed from reactions with cellular carboxylic acids would also not exhibit significant levels of fluorescence, and therefore, should not

provide background fluorescence capable of preventing observation of pyrazoline adducts.

a)



92 **93**

b)

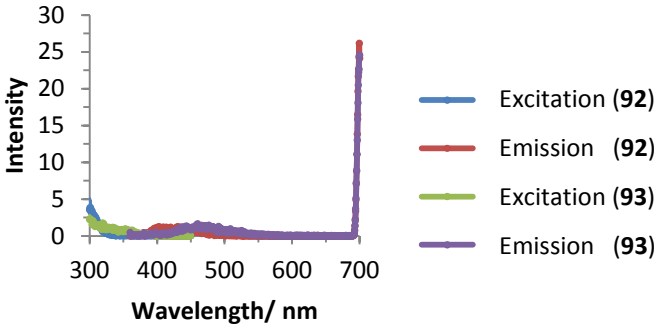
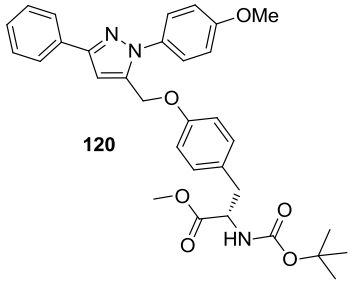


Figure 39: Excitation and emission spectra of hydrazides **92** and **93** in MeCN:PBS (1:1 v/v) at 0.1 mM concentrations. (a) Structures of hydrazide byproducts **92** and **93**; (b) Spectra obtained. Scan width = 3.0 nm, scan speed = 100 nm min^{-1}. Excitation spectra: λ_{em} = 500 nm, λ_{ex} scan = 210-450 nm. Emission spectra: λ_{ex} = 350 nm, λ_{em} scan = 360-700 nm.

The oxidation of the pyrazoline core of **118** and **143** under atmospheric conditions could be problematic for an assay if the formation of the pyrazoline adduct is slow. Importantly, the excitation and emission spectra of pyrazole **120** (Figure 40) indicated that the loss of the pyrazoline core results in an almost complete drop in absorption and emission observed at the wavelengths chosen for the assay. It is difficult to predict if the pyrazoline core will undergo rapid oxidation in a cellular environment. Monitoring of the reaction at regular time intervals will be necessary to determine the significance of this side reaction.

a)



b)

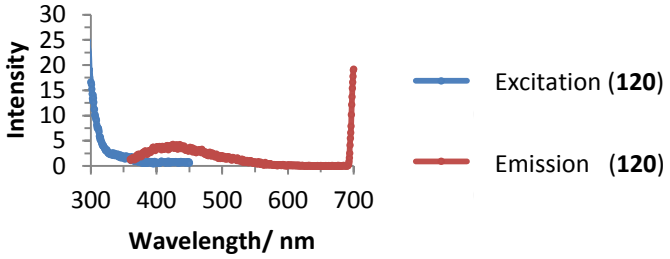
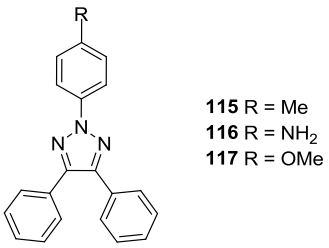


Figure 40: Excitation and emission spectra of pyrazole **120** in MeCN:PBS (1:1 v/v) at a concentration of 0.1 mM. (a) Structure of pyrazole **120**; (b) Spectrum obtained. Scan width = 3.0 nm, scan speed = 100 nm min^{-1}. Excitation spectra: λ_{em} = 500 nm, λ_{ex} scan = 210-450 nm. Emission spectra: λ_{ex} = 350 nm, λ_{em} scan = 360-700 nm.

Finally, 2,4,5-triaryl-1,2,3-triazole byproducts **115**-**117** were analysed. Excitation and emission spectra obtained for the three symmetrical triazoles indicated that the nature of the substituent on the *N*2-aryl ring has a significant effect on the fluorescence intensity observed (Figure 41); increasing the electron donating ability of the substituent results in increasing intensity. The intensity could be seen to follow the order Ar-NH_2 > Ar-OMe > Ar-Me with a bathochromic shift also observed. The difference in intensity could also be observed qualitatively by spotting each triazole onto a TLC plate and examining the surface at 365 nm.

a)



b)

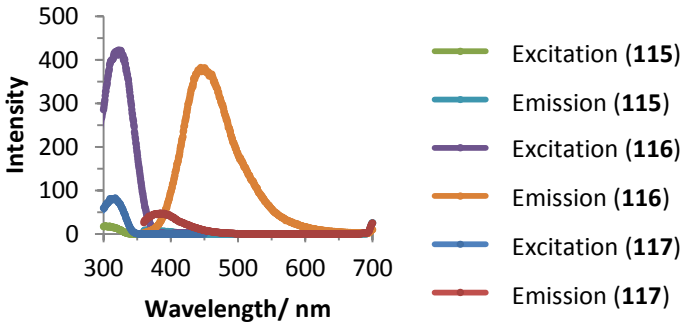


Figure 41: Excitation and emission spectra of triazoles **115**-**117** in MeCN:PBS (1:1 v/v) at 0.1 mM concentrations. (a) Structure of triazoles **115-117**; (b) Spectra obtained. Scan width = 3.0 nm, scan speed = 100 nm min^{-1}. Excitation spectra: λ_{em} = 500 nm, λ_{ex} scan = 210-450 nm. Emission spectra: λ_{ex} = 350 nm, λ_{em} scan = 360-700 nm.

These data indicated that some triazole byproducts may have sufficient levels of fluorescence at 470 nm to compete with fluorescence from the desired reaction. In this case, the fluorescence intensity of triazole **117** is approximately equal to that observed for pyrazolines **118** and **119**, and greater than that observed for pyrazoline **143**. Consequently, the reactions conducted at low concentration in cells will need to be monitored for background fluorescence. Additionally, any alternative tetrazoles identified as reagents for future reactions will need to be monitored for varying background fluorescence and shifting excitation and emission maxima.

1.3. Studies towards an intracellular fluorogenic click reaction

Following the successful synthesis and fluorescence analysis of the required tool molecules and their pyrazoline adducts, it was possible to begin to develop the intracellular fluorogenic click reaction for the visualisation of peptides inside cells.

To conduct reactions in a cellular environment the reagents would need to be dissolved in phosphate buffered saline (PBS) solution at concentrations of 100-250 μM of tetrazole **81** and approximately 15 μM of tags **105**, **106** or **142**, with up to 10% DMSO (v/v) to aid solubility if necessary. Greater ratios of DMSO are undesirable as this may affect the permeability and viability of the cell itself.

When attempting dissolution of the tetrazole reagent in 10% DMSO:PBS a suspension of the tetrazole was observed and full dissolution could not be achieved. The Lin group report that, while this lack of solubility was also observed during their reactions, it did not adversely affect the results of their experiments.[128] Even so, this was of concern; therefore, the reactions were first conducted in DMSO at ten times the desired concentration (140 μM **105**, **106** and **142**; 2.5 mM **81**) to monitor the reaction with full dissolution, and to ensure the fluorescence intensity was sufficient for direct observation.

Subsequently, these solutions were further diluted to give the desired concentrations of reagents and tags in 10% DMSO:MeCN and 10% DMSO:PBS (14 μM **105**, **106** and **142**; 250 μM **81**). As expected, poor dissolution of the tetrazole was only observed with the 10% DMSO:PBS solution; therefore, results obtained in the 10% DMSO:MeCN solution are not affected by this factor.

1.3.1. Control experiments

Due to the fluorescence profiles observed for the various reaction byproducts, a number of control experiments were conducted to monitor the levels of background fluorescence during the reaction. All controls and reaction mixtures were monitored both with and without irradiation to ensure that any positive results obtained were a direct result of the photochemical formation of the nitrile imine intermediate from the 2,5-diaryl tetrazole.

The relevant solutions were prepared in 384 well plates and a background measurement taken at time (t) = 0 minutes. The appropriate samples were illuminated for 3 minutes with a 270-330 nm UV lamp and the fluorescence intensity monitored (λ_{ex} = 350 nm, λ_{em} = 470 nm) every 30 seconds for 10 minutes and then every 30 minutes for an additional 14 hours to give a time-resolved fluorescence spectrum.

Firstly, the three solvents used in the experiment were irradiated and monitored for background fluorescence over the 14 hour period (Figure 42). No significant increase in intensity was observed during this period, although a very minor increase in intensity was observed for the DMSO control overnight.

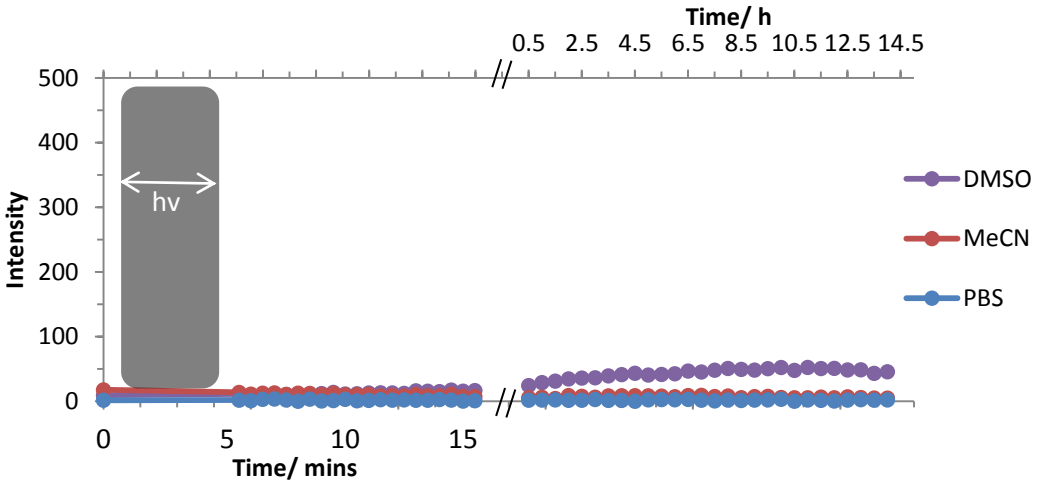


Figure 42: Background fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for solvents over a 14 hour period after irradiation (270-330 nm, 3 min).

The three alkene tags and two reagents (Figure 43) were then monitored individually for background fluorescence prior to irradiation (Figure 44) and after irradiation (Figure 45).

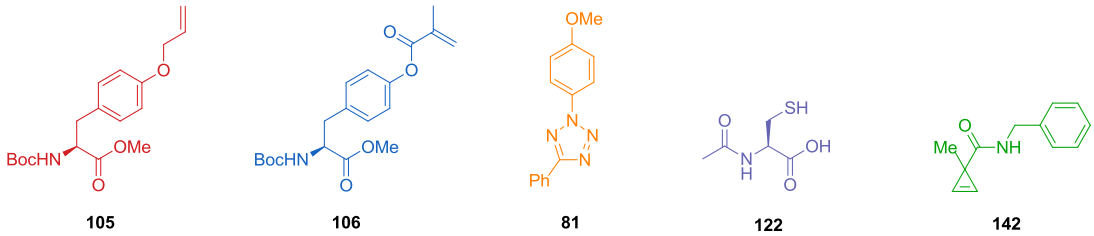


Figure 43: Structure of each reagent used in the investigation of the fluorogenic click reaction.

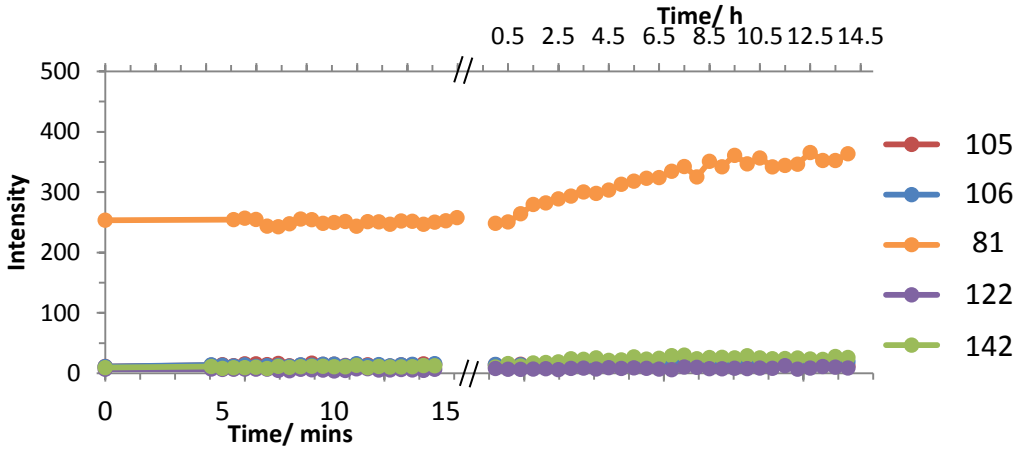


Figure 44: Background fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for individual reagents over a 14 hour period without irradiation. Concentrations: **105** (140 µM), **106** (140 µM), **81** (2.5 mM), **122** (50 mM) and **142** (140 µM) in DMSO.

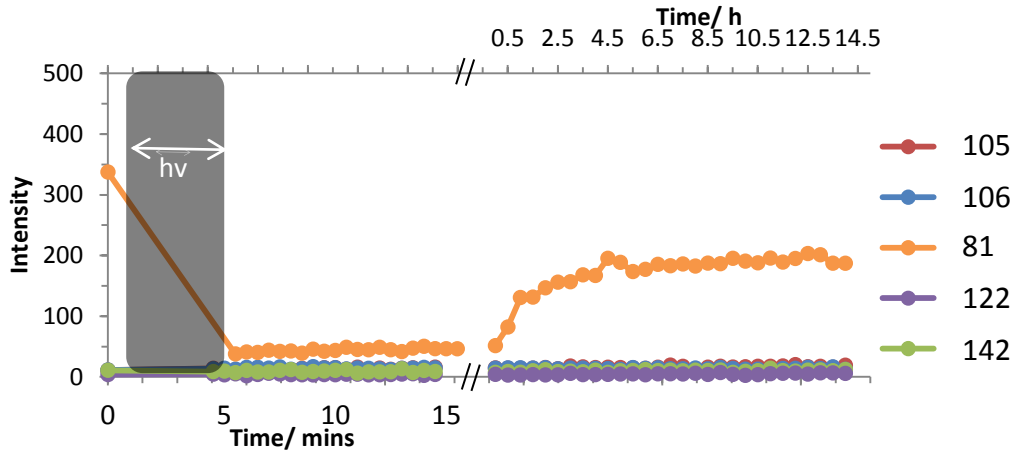


Figure 45: Background fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for individual reagents over a 14 hour period after irradiation (270-330 nm, 3 min). Concentrations: **105** (140 µM), **106** (140 µM), **81** (2.5 mM), **122** (50 mM) and **142** (140 µM) in DMSO.

The spectra indicated that over the 14 hour time period the controls for all of the alkene tags and for *N*-acetyl cysteine had minimal levels of fluorescence, comparable to those observed for the solvent controls. Furthermore, they showed no change in fluorescence intensity when irradiated.

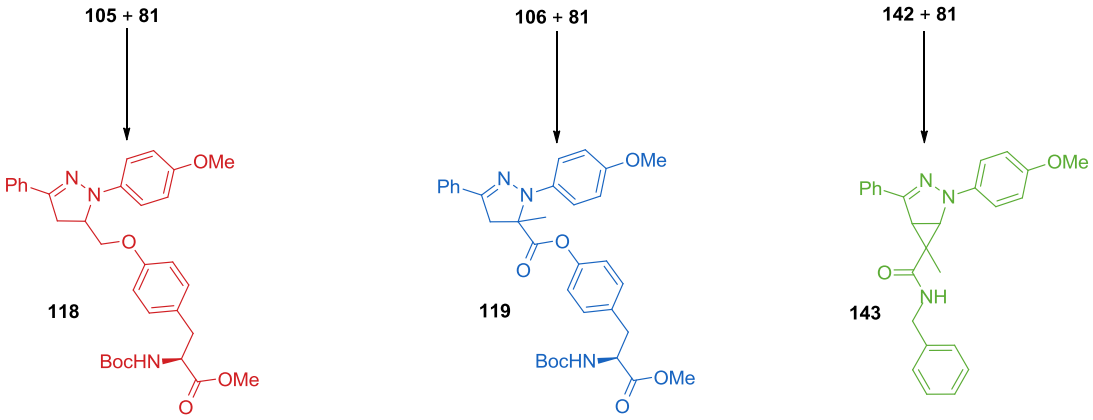
On the other hand, the tetrazole control shows a baseline level of fluorescence at approximately 250 units with no significant change observed over a 14 hour period without irradiation. Irradiation of the tetrazole results in a sharp drop in fluorescence, indicating activation of the tetrazole is successful under these conditions. A low level of fluorescence is observed initially, likely residual fluorescence from the nitrile imine intermediate formed, which then increases rapidly within the first couple of hours. This is most likely indicative of the formation of triazole **117** from tetrazole **81**. These data indicate that a background level of fluorescence may be observed due to formation of the triazole. This could be problematic if the rate of the pyrazoline formation is slower than the rate of triazole formation under these conditions, or if the fluorescence intensity observed from the pyrazoline is lower than the intensity of fluorescence emitted by any triazole formed from the excess tetrazole reagent.

1.3.2. Time-resolved fluorescence analysis (high concentration)

The reaction of the three alkene tagged tool molecules with tetrazole **81** to give the corresponding pyrazoline adducts was first monitored in DMSO at ten times the proposed operational concentration to give a comparison to the control experiments and to ensure the fluorescence would be observable (Scheme 27).

The reaction mixtures were initirally monitored over the 14 hour period without irradiation (Figure 46). The results indicated that each reaction mixture has a baseline level of fluorescence of approximately 300 units with no significant variation between the mixtures. The intensity and profile of the time-resolved

fluorescence spectrum are consistent with the results obtained for the tetrazole control, indicating that no reaction takes place prior to irradiation of the tetrazole.



Scheme 27: Structures of pyrazoline adducts formed during reactions between the corresponding reagent and tag.

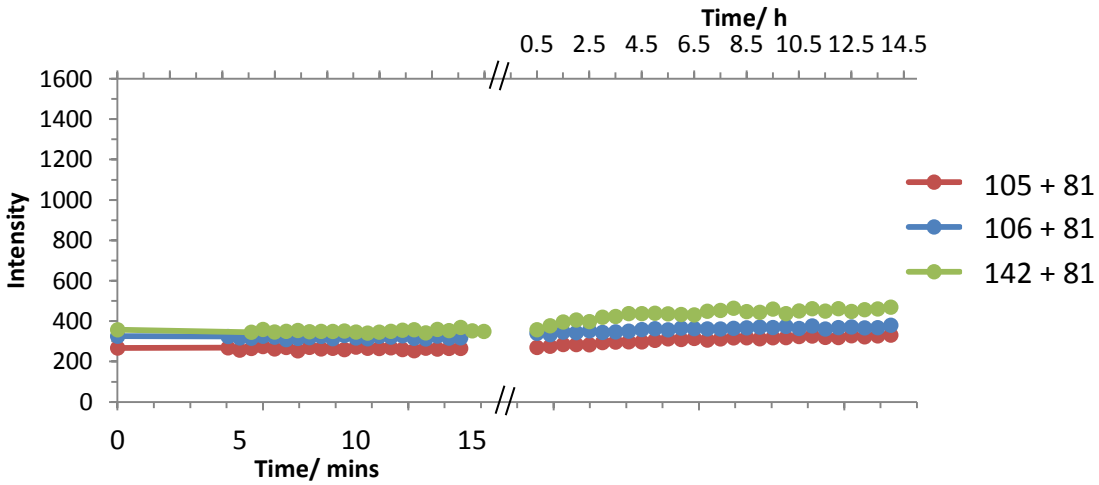


Figure 46: Fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for reaction mixtures over a 14 hour period without irradiation. Concentrations: **105** (140 µM), **106** (140 µM), **81** (2.5 mM) and **142** (140 µM) in DMSO.

Irradiation of the reaction mixtures containing both Boc-allyl-Tyr-OMe **105** and cyclopropene **142** showed the expected drop in fluorescence associated with photoactivation of the tetrazole reagent to form the nitrile imine intermediate (Figure 47). However, the profile of the reaction was similiar to that of the irradiated control for tetrazole **81** (Figure 41) suggesting that the desired pyrazoline formation did not occur. Analysis of these reaction mixtures by LC-MS was unable to

confirm the presence of the pyrazoline adducts due to the low concentrations at which the study was performed.

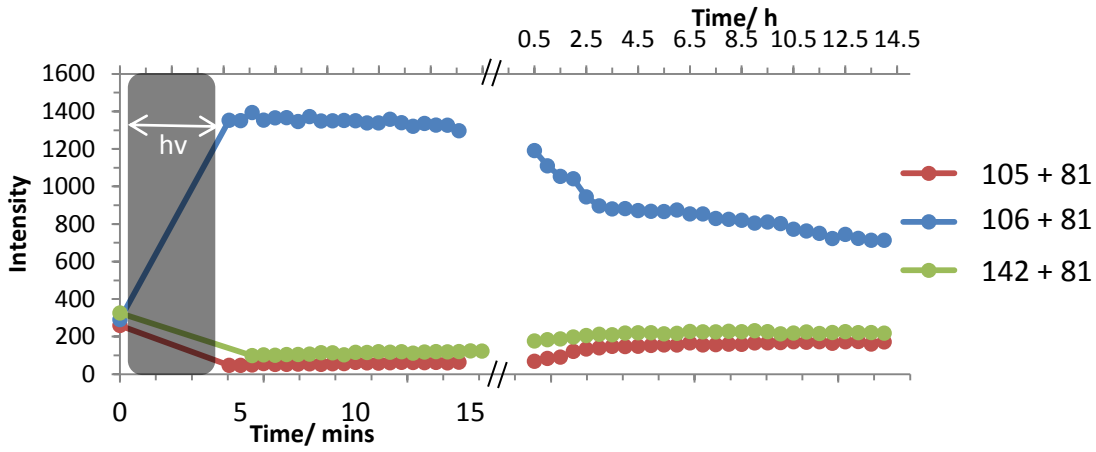


Figure 47: Fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for reaction mixtures over a 14 hour period after irradiation (270-330 nm, 3 min). Concentrations: **105** (140 µM), **106** (140 µM), **81** (2.5 mM) and **142** (140 µM) in DMSO.

Gratifyingly, the reaction mixture containing Boc-methacryloyl-Tyr-OMe **106** showed a significant change in fluorescence intensity when irradiated. In contrast to the other reaction mixtures, the characteristic drop in fluorescence after irradiation was not observed. Instead, the formation of pyrazoline adduct **119** occurs rapidly within the first 5 minutes of the reaction. This is not an unexpected result as the reactivity of Boc-methacryloyl-Tyr-OMe **106** was shown to be significantly greater than the other tags.

Interestingly, the fluorescence intensity of the adduct decreases gradually overnight. It is possible that this may be a result of photobleaching of the pyrazoline adduct. As stated previously, the aniline substituted 2,5-diaryl tetrazole reagents were not investigated as reagents due to rapid photobleaching of the pyrazoline adducts formed. It may be that photobleaching of the anisole substituted pyrazoline adducts can also occur on a slower timescale.

Although these data showed some promise for formation of adduct **119**, the effect of *N*-acetyl cysteine **122** on the reaction mixtures remained to be determined. As such, the reactions were repeated in the presence of 50 mM **122**, with the addition

of two further controls: a mixture containing *N*-acetyl cysteine **122** with tetrazole **81** and a mixture containing *N*-acetyl cysteine **122** with alkene tag **142** (Figure 48 and 49).

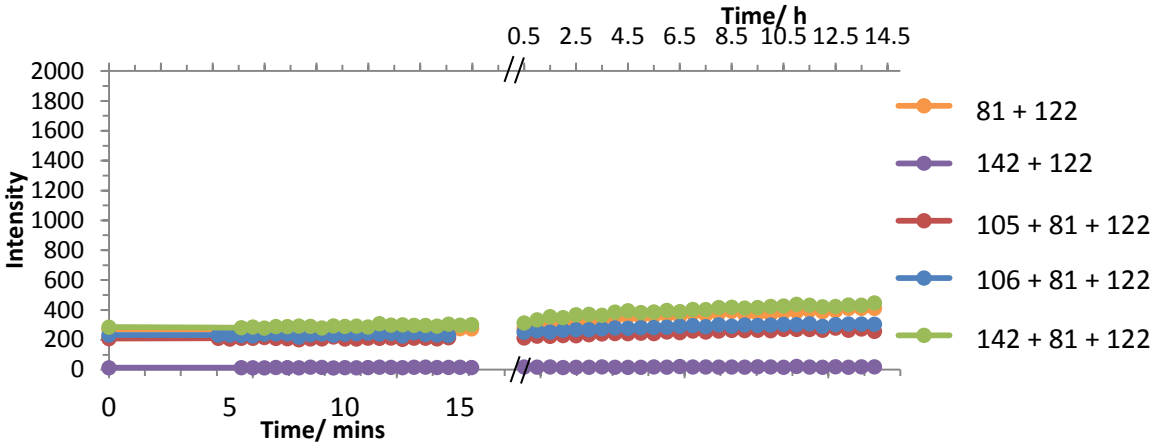


Figure 48: Fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for reaction mixtures containing *N*-acetyl cysteine **122** over a 14 hour period without irradiation. Concentrations: **122** (50 mM), **105** (140 µM), **106** (140 µM), **81** (2.5 mM) and **142** (140 µM) in DMSO.

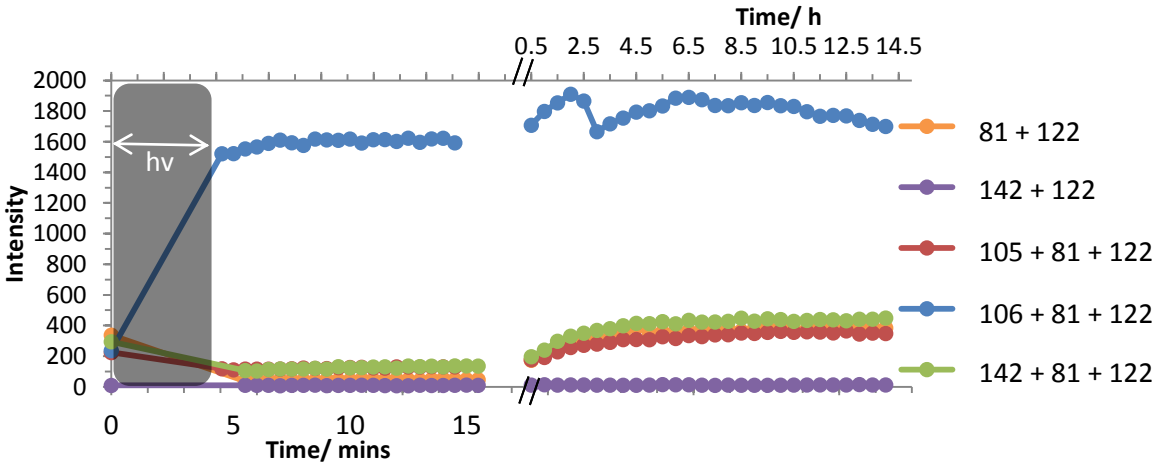


Figure 49: Fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for reaction mixtures containing *N*-acetyl cysteine **122** over a 14 hour period after irradiation (270-330 nm, 3 min). Concentrations: **122** (50 mM), **105** (140 µM), **106** (140 µM), **81** (2.5 mM) and **142** (140 µM) in DMSO.

The control containing both the alkene tag **142** and *N*-acetyl cysteine **122** showed no significant fluorescence intensity, indicating that this combination of reagents exhibits no background fluorescence. The control containing tetrazole **81** and *N*-acetyl cysteine **122** showed the same profile as observed for lone tetrazole **81**

(Figure 44 and 45), but with higher fluorescence intensity observed. When irradiated, the reaction mixtures containing Boc-allyl-Tyr-OMe **105** and cyclopropene **142** also followed the same reaction profile as the control, again suggesting that the desired pyrazoline formation did not occur with these tags.

Importantly, the fluorescence profile for the reaction containing Boc-methacryloyl-Tyr-OMe **106** still gives suitable fluorescence intensity in the presence of *N*-acetyl cysteine **122**, indicating that the desired pyrazoline formation still occurs in the presence of this glutathione mimic. Furthermore, these data indicated that Michael addition of *N*-acetyl cysteine **122** to the alkene tag was not a major side reaction since no inhibition of the reaction was observed in its presence.

Interestingly, the overnight decrease in fluorescence previously observed does not occur here, possibly due to the prevention of photobleaching by the presence of excess *N*-acetyl cysteine which acts as an anti-oxidant.

1.3.3. Time-resolved fluorescence analysis (low concentration)

With the successful observation of fluorescence for the reaction of Boc-methacryloyl-Tyr-OMe **106** and tetrazole **81** it was necessary to repeat the experiment at the desired concentrations of reagents and tags in 10% DMSO:MeCN and 10% DMSO:PBS (14 μM **105**, **106** and **142**; 250 μM **81**; 5 mM **122**). The experiment was also repeated for Boc-allyl-Tyr-OMe **105** and cyclopropene **142** to confirm the same lack of reactivity would be observed at the lower concentration.

The reaction mixtures were first irradiated in the absence of *N*-acetyl cysteine **122** in 10% DMSO:MeCN to monitor the reactions at lower concentrations while maintaining full solubility of the tetrazole reagent (Figure 50).

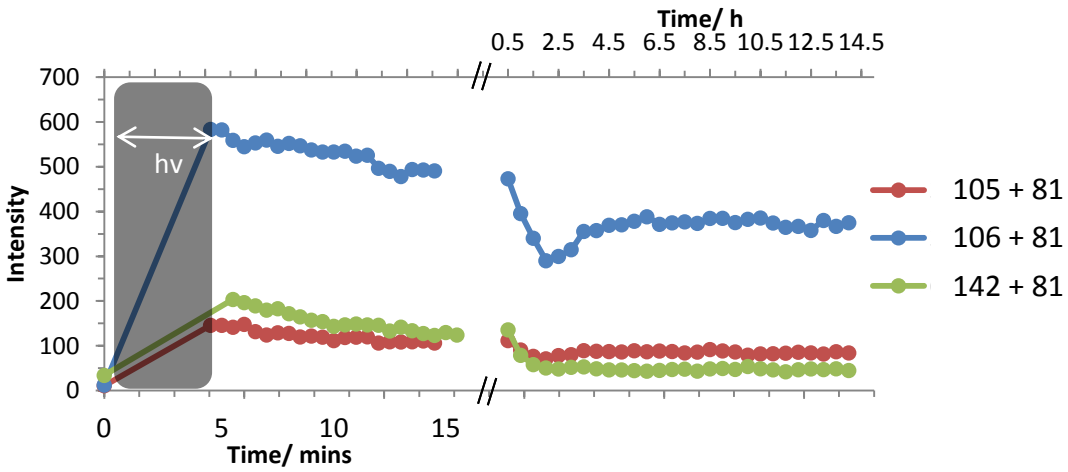


Figure 50: Fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for reaction mixtures over a 14 hour period after irradiation (270-330 nm, 3 min). Concentrations: **105** (14 µM), **106** (14 µM), **81** (250 µM) and **142** (14 µM) in 10 % DMSO:MeCN.

As expected, the Boc-methacryloyl-Tyr-OMe **106** reaction mixture showed a characteristic profile for formation of the pyrazoline adduct. Most interestingly, however, was the apparent formation of pyrazoline adducts from the other tags, as indicated by the characteristic rapid induction of fluorescence and subsequent decrease in intensity over time. Furthermore, the intensity of fluorescence observed for the Boc-methacryloyl-Tyr-OMe **106** reaction mixture was not decreased by a factor of 10, as would be expected from its concentration dependent fluorescence profile (Figure 38b). These results indicated that high volumes of DMSO may be inhibiting the reaction, and that all three adducts may still exhibit fluorescence in a cell based assay.

The experiment was repeated in the presence of *N*-acetyl cysteine **122** to monitor its effects under these conditions (Figure 51). The fluorescence intensity significantly decreased for all three reaction mixtures, particularly the mixture containing Boc-methacryloyl-Tyr-OMe **106**. These results indicate that under these conditions the presence of *N*-acetyl cysteine **122** significantly inhibits the formation of pyrazoline adducts and the low fluorescence intensity observed is unlikely to be distinguishable from the background intensity of the solvents or any byproducts formed.

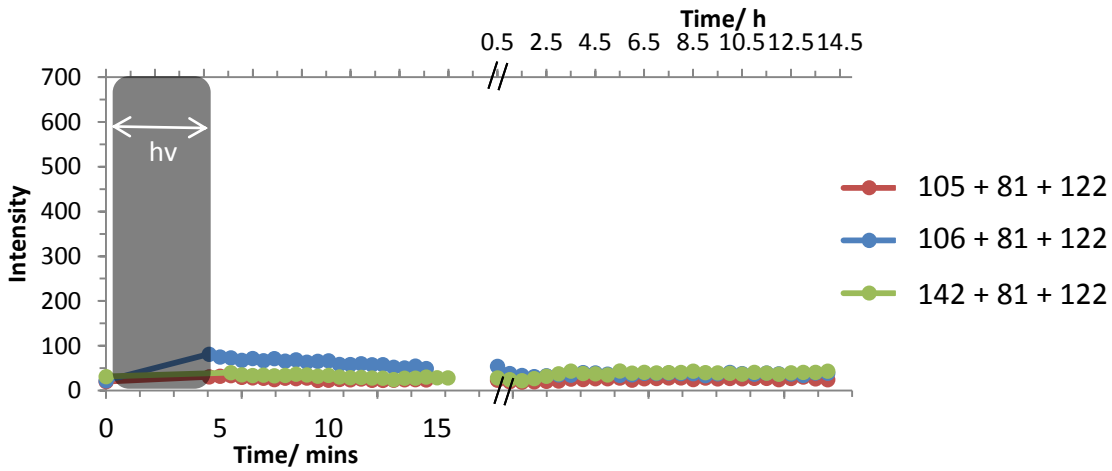


Figure 51: Fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for reaction mixtures containing *N*-acetyl cysteine **122** over a 14 hour period after irradiation (270-330 nm, 3 min). Concentrations: **122** (5 mM), **105** (14 µM), **106** (14 µM), **81** (250 µM) and **142** (14 µM) in 10 % DMSO:MeCN.

The cause of this unexpected result is unclear. Since the rate of each reaction is likely to be first order in respect to any of the individual reagents, the reduction of concentration by a factor of 10 should, in theory, result in directly proportional reduction of the reaction rate in each reaction. Therefore, the observed results cannot be attributed to the change in concentration but are most likely caused by the change in solvent system. While at higher concentrations the reactions are conducted in DMSO the solvent system is changed to 10% DMSO in MeCN at the lower concentration. It is possible that in the modified solvent system the side reactions, such as hydrazide formation between the tetrazole and *N*-acetyl cysteine or Michael addition between the alkene and *N*-acetyl Cysteine, become dominate and inhibit the pyrazoline formation.

Finally, the reaction was repeated in the absence of *N*-acetyl cysteine **122** and in 10% DMSO:PBS in order to monitor the effects of limited tetrazole solubility on the reaction (Figure 52).

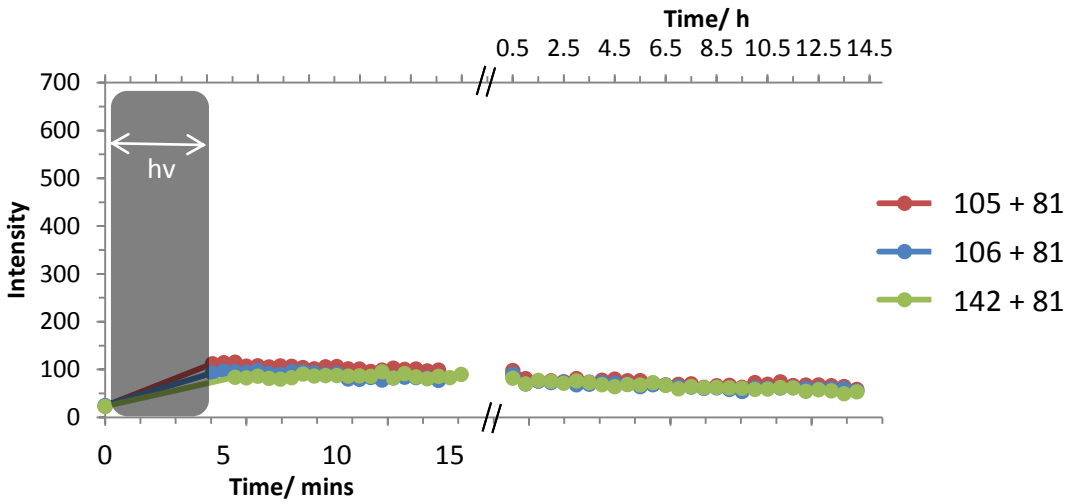


Figure 52: Fluorescence observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for reaction mixtures over a 14 hour period after irradiation (270-330 nm, 3 min). Concentrations: **105** (14 µM), **106** (14 µM), **81** (250 µM) and **142** (14 µM) in 10 % DMSO:PBS.

Again, the fluorescence intensity of each reaction mixture was significantly decreased and would likely be indistinguishable from other sources of background fluorescence. This result could indicate that the tetrazole is being quenched by the water solvent or that good solubility of the tetrazole reagent is essential for the reaction to proceed under cellular conditions.

Overall, these data indicate that, in its current form, the reaction is unlikely to be suitable for cellular reactions. While the tetrazole solubility could be improved with some additional synthesis, such as preparation of PEGylated analogues, the possible reactions that can occur in the presence of *N*-acetyl cysteine **122** and their effect on the reaction system warranted further scrutiny. Therefore, a series of competition reactions was conducted to further investigate the reaction system and the tetrazole component, in particular.

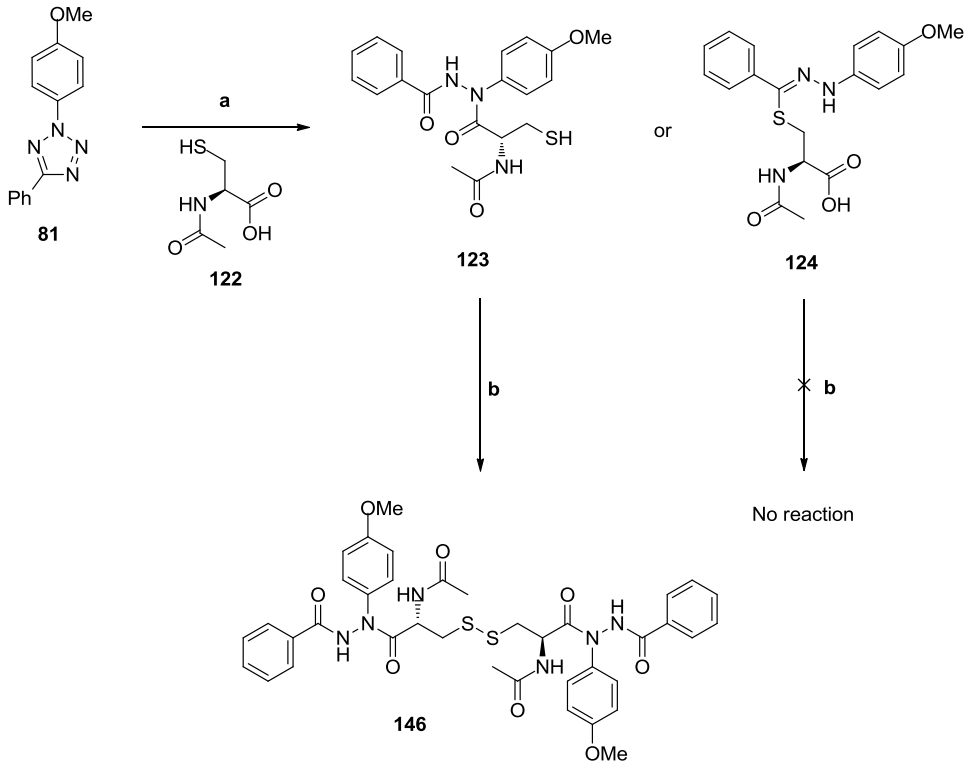
1.4. Investigation of side reactions and byproducts

1.4.1. Competition reactions

The side reactions that can occur during the synthesis of pyrazoline adducts identified so far have included the formation of hydrazides in the presence of carboxylic acids, the dimerisation of nitrile imines to give triazoles, and the oxidation of pyrazolines to pyrazoles (see Section 1.2). The results of the low concentration fluorogenic click reactions indicated that the presence of water or *N*-acetyl cysteine **122** can either inhibit the desired reaction, or perhaps cause further side reactions (see Section 1.3.3). Based on these observations, further investigation was required.

As intimated previously, the reaction of tetrazole **81** with the carboxylic acid moiety in *N*-acetyl cysteine **122** would be expected to give hydrazide **123** which could explain the loss of fluorescence observed in the low concentration reactions (Scheme 28). However, the presence of a nucleophilic thiol in *N*-acetyl cysteine **122** presents an alternative side reaction in which nucleophilic addition of the sulfur to the nitrile imine carbocation and proton transfer to quench the nitrogen anion could feasibly give hydrazonothioate **124** as a byproduct.

The reaction proceeded smoothly under the previously established conditions with peak to peak conversion of tetrazole **81** to a product with an $[M+H]^+$ of 388, consistent with either hydrazide **123** or hydrazonothioate **124**. The structure of the molecule was difficult to determine by ^{1}H NMR due to the similarity of the two compounds, exacerbated by the presence of rotamers. A ^{1}H variable temperature NMR experiment gave a single set of peaks; however, this resulted in partial formation of an additional byproduct with a $[M+H]^+$ of 773, consistent with dimerised structure **146** and further complicated the NMR analysis.

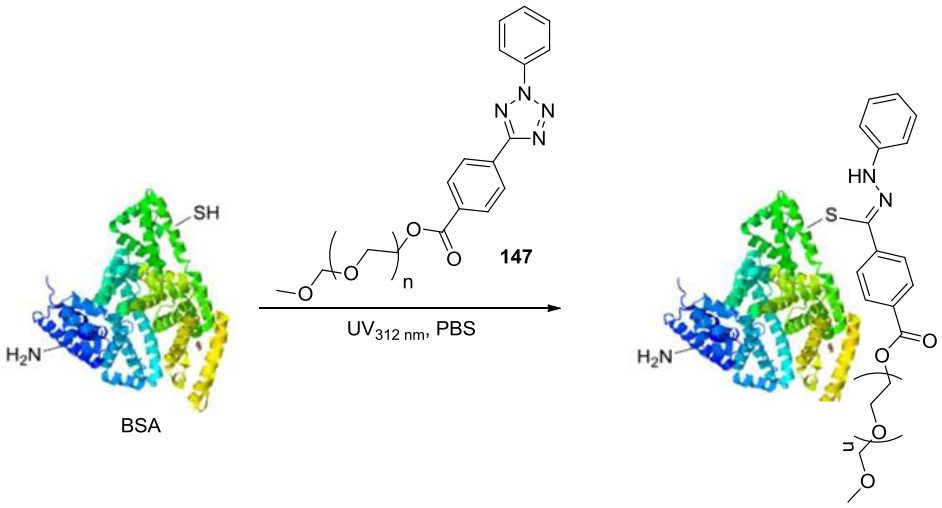


Scheme 28: Investigation of side reactions occurring between tetrazole **81** and *N*-acetyl cysteine **122**. Reagents and conditions: (a) *N*-acetyl cysteine (3 equiv.), UV light (270-330 nm), MeCN:EtOH (3:1 v/v), rt, 1 h, 20%; (b) 30% $H_2O_{2(aq)}$ (1.8 equiv.), NaI (0.4 equiv.), EtOAc, rt, 0.5 h, 58%.

Based on the above observations, deliberate conversion of the monomer to the hypothesised dimer gave clear identification of the byproduct. The use of 30% H_2O_2 and catalytic NaI to oxidise the thiol group under mild conditions successfully converted all of the remaining intermediate to disulfide linked dimer **146**.[129] The conversion of the intermediate to the dimer under oxidative conditions, as observed by LC-MS, along with ^{1}H variable temperature NMR analysis of **146** confirmed that the reaction with the tetrazole had occurred through the carboxylic acid moiety to give hydrazide **123** as the byproduct. The low yield obtained for **123** was a result of loss of the dimerised material when initially purifying the monomer. The formation of the dimer indicated that under these conditions the nitrile imine would preferentially react with the carboxylic acid moiety in preference to the thiol group.

Levkin *et. al.* have recently demonstrated some of the possible applications of nitrile imine quenching by thiols, including the synthesis of hydrazonothioate molecules, modification of proteins and surface labelling.[130] Of particular note was the labelling of bovine serum albumin (BSA), a protein which contains a single accessible thiol cysteine group, under aqueous conditions. BSA was labelled with water soluble PEG-modified tetrazole **147** with the predicted product being the hydrazonothioate (Figure 53).

a)



b)

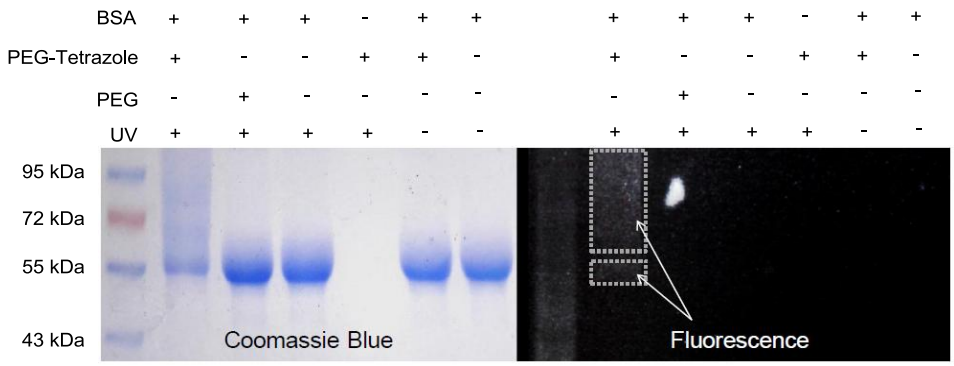


Figure 53: (a) Photochemical labelling of BSA with PEG-modified tetrazole **147** under aqueous conditions; (b) Coomassie blue stain and in-gel fluorescence analysis of the reaction and its controls.[130]

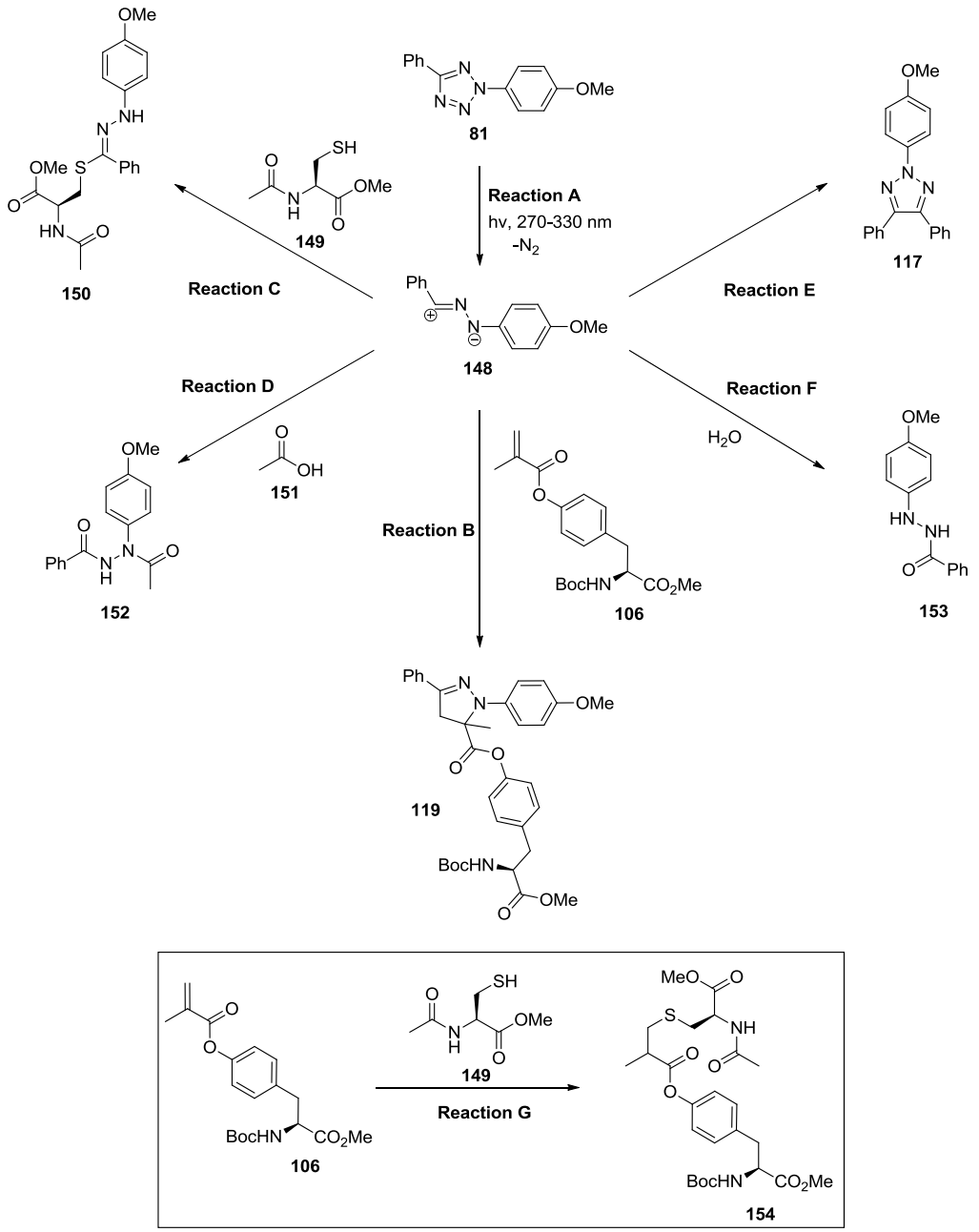
The reaction progress was analysed by Coomassie Blue staining and in-gel fluorescence. The results showed clear labelling of the protein when BSA, PEG-tetrazole **147** and UV light were all present. The group commented on the

apparent multiple addition of the nitrile imine intermediate to the protein, as evident from the fluorescent band spanning 55-95 kDa, and attributed it to multiple additions to reactive amines in the free lysine side chains. However, based on the reactivity established between nitrile imines and carboxylic acids in this work, it is more likely that the multiple additions observed were a result of hydrazide formation with free carboxylic acids in the side chains of glutamic acid and aspartic acid residues. In addition, it is entirely possible that the reactivity observed occurred only with the carboxylic acid functionality in the protein and the thiol cysteine group remained unreacted.

The rapid reactivity of photochemically generated nitrile imines with carboxylic acids in the presence of thiols (*vide supra*) or alkenes (see Section 1.2.2) could form the basis of a new bioorthogonal reaction and was investigated further in Section 1.6.2.

A series of competition reactions was conducted in order to identify the most likely cause of the negative results obtained for the fluorogenic click reaction and to determine the relative rates of the desired reaction compared to each side reaction. In order for the fluorescent labelling to work effectively, the rate of pyrazoline formation must be greater than the rate of each side reaction. For the three alkene tags tested the relative reactivity appeared to follow the order: Boc-methacryloyl-Tyr-OMe **106** > cyclopropene **142** > Boc-allyl-Tyr-OMe **105**. This order was expected based on the steric and electronic effects of the alkene substituents. Tag **106** was chosen for the competition reactions as it is the most reactive; therefore, if the side reactions could compete with this tag then they would present an ever greater issue with the less reactive alkenes.

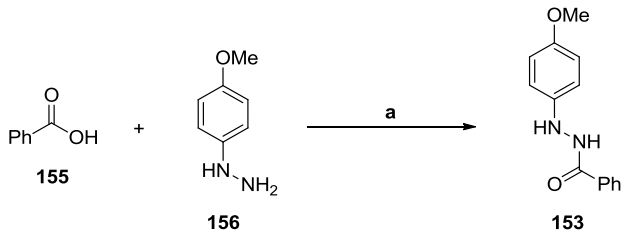
Scheme 29 details the reactions that may occur in a reaction mixture containing tetrazole **81**, Boc-methacryloyl-Tyr-OMe **106** and various reagents representing some of the functional groups present in cells.



Scheme 29: Overview of the reactions that can be carried out between tetrazole **81** and the various components of the fluorogenic click reaction under photochemical conditions.

In each competition reaction nitrile imine **148** was first generated by photochemical irradiation of tetrazole **81** (Reaction A). Five possible side reactions were identified that could compete with the desired formation of the pyrazoline adduct (Reaction B). *N*-acetyl-Cys-OMe **149** and acetic acid **151** were identified as suitable molecules to independently assess the relative rate of reaction of thiols

(Reaction C) and carboxylic acids (Reaction D) with the nitrile imine intermediate. *N*-acetyl-Cys-OMe **149** could simultaneously be used to observe if Michael addition to *N*-Boc-methacryloyl-Tyr-OMe **106** occurred (Reaction G). The formation of triazole byproduct **117** required no additional reagents (Reaction E). Unfortunately, the quenching of nitrile imine intermediate **148** with water to form hydrazide **153** (Reaction F) could not be conducted as a competition reaction due to the insolubility of the tetrazole in aqueous media, even when 10% DMSO (v/v) was used to aid solubility. Instead, **153** was synthesised from molecule **155** and **156** (Scheme 30) for use as a control in further experiments (see Section 1.5.).



Scheme 30: Synthesis of hydrazide **153**. Reagents and conditions: (a) Phenyl acetic acid **155** (1 equiv.), ethyl chloroformate (1.1 equiv.), triethylamine (2.4 equiv.), (4-methoxy)phenyl hydrazine **156** (1.1 equiv.), 0-25 °C, 1 h, 22%.

The relative rates of Reactions B-E and Reaction G were investigated by conducting three competition reactions. The proportions of pyrazoline adduct **119** and the relevant byproducts were monitored by LC-MS analysis.

Competition Reaction 1:

Tetrazole **81** (1 equiv.) and Boc-methacryloyl-Tyr-OMe **106** (1 equiv.) were reacted together to compare the relative rates of Reaction B and Reaction E (Scheme 29).

The reaction was monitored by LC-MS and the proportion of pyrazoline **119** to triazole **117** taken at approximately 50% completion (Figure 54). As can be noted from the LC-MS spectrum obtained, the formation of pyrazoline **119** is rapid for the electron rich nitrile imine and electron poor alkene, as expected. While a peak is

observed for the triazole **117** byproduct, the amount formed was insignificant. This result clearly indicates the dominance of Reaction B over Reaction E.

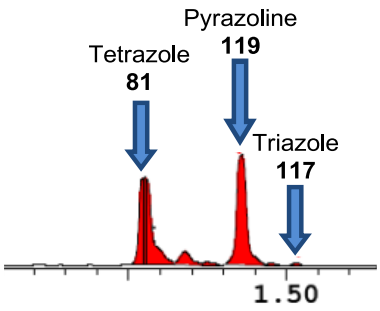


Figure 54: LC-MS spectrum for the competition reaction of tetrazole **81** (1 equiv.) with *N*-Boc-methacryloyl-Tyr-OMe **106** (1 equiv.).

Competition Reaction 2:

Tetrazole **81** (1 equiv.) and Boc-methacryloyl-Tyr-OMe **106** (1 equiv.) were reacted together in the presence of commercially available *N*-acetyl-Cys-OMe **149** (1 equiv.) to compare the relative rates of Reactions B, C, E and G (Scheme 29).

This competition reaction was designed to monitor the relative rate of hydrazonothioate **150** formation to pyrazoline adduct **119** formation when no carboxylic acid moiety is present. The reaction mixture was also monitored for signs that the thiol group could inhibit the desired reaction by undergoing Michael addition to *N*-Boc-methacryloyl-Tyr-OMe **106**. The reaction was allowed to proceed to completion and analysed by LC-MS (Figure 55).

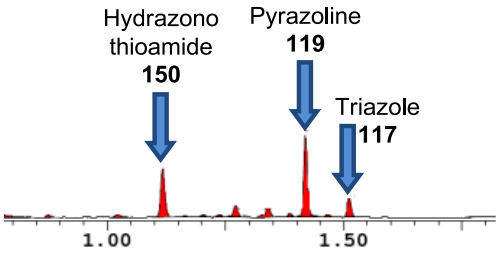


Figure 55: LC-MS spectrum for the competition reaction of tetrazole **81** (1 equiv.) with Boc-methacryloyl-Tyr-OMe **106** (1 equiv.) and *N*-acetyl-Cys-OMe **149** (1 equiv.).

Hydrozonothioate **150** was isolated from the crude mixture of the competition in 10% yield and the structure confirmed by LC-MS and NMR analysis. The relative rate of thioate formation was also shown to be significantly more rapid than triazole formation. Importantly, it was also shown to compete with the formation of the desired pyrazoline adduct **119**, thereby appropriating some of the tetrazole reagent from the desired reaction pathway. This would be even more problematic for less reactive alkenes such as Boc-allyl-Tyr-OMe **105**.

No evidence of Michael addition product **154** was observed, indicating that this side reaction is slower even than the formation of the triazole byproduct. It is, therefore, unlikely to be the cause of the negative results obtained for the low concentration reactions, indicating that the tetrazole side reactions are most likely the cause.

Competition Reaction 3:

Tetrazole **81** (1 equiv.) and Boc-methacryloyl-Tyr-OMe **106** (1 equiv.) were reacted together in the presence of acetic acid **151** (1 equiv.) to compare the relative rates of Reactions B, D, and E (Scheme 29) and determine the inhibitory effects of a carboxylic acid moiety on formation of the desired pyrazoline adduct (Figure 56).

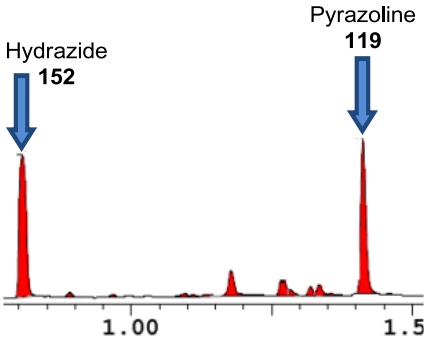


Figure 56: LC-MS spectrum for the competition reaction of tetrazole **81** (1 equiv.) with Boc-methacryloyl-Tyr-OMe **106** (1 equiv.) and Acetic acid **151** (1 equiv.).

When acetic acid was added to the reaction mixture a 45:35 ratio of hydrazide **152** to pyrazoline **119** was observed by LC-MS. Hydrazide **152** was successfully isolated from the crude mixture of the competition reaction in 24% yield and the structure

confirmed by LC-MS and NMR analysis. These data indicated that the carboxylic acid is able to compete with the most reactive alkene tag and can appropriate some of the tetrazole reagent from the desired reaction pathway. Again, this would be even more problematic for less reactive alkenes.

1.4.2. Summary of potential side reactions

Reaction scheme 29 and the subsequent data obtained indicate that formation of hydrazide **152** (Reaction D) and formation of hydrazonothioate **150** (Reaction C) can compete with the synthesis of pyrazoline adduct **119** (Reaction B). Dimerisation to triazole **117** (Reaction E) and formation of a Michael addition product (Reaction G) were too slow to interfere. Furthermore, the previous reaction between tetrazole **81** and *N*-acetyl cysteine **122** had indicated that the reactivity of nitrile imines was greater with carboxylic acids than thiols (Scheme 28). It can, therefore, be inferred that the inhibition of the photoclick reaction in the presence of *N*-acetyl cysteine **122** was most likely quenching of the nitrile imine by the carboxylic acid group (Reaction D). It was not possible to determine the relative rate of Reaction F; however, it is possible that quenching of the nitrile imine by water was the cause of inhibition observed when the reaction was conducted in PBS.

The formation of pyrazoline adduct **119** from Boc-methacryloyl-Tyr-OMe **106** was shown to be feasible in the presence of functional groups that can compete for reactivity with tetrazole **81**. However, the rate of Reaction B is dependent on the reactivity of the alkene tag, whereas the relative rate of reaction D would presumably remain the same in a cell based assay. Therefore, the rate of reaction with Boc-allyl-Tyr-OMe **105** and cyclopropene tag **142** is likely to be insufficient to overcome this side reaction. Furthermore, in this experiment the relative concentrations of each reagent were maintained as equimolar. As has been previously discussed, the presence of nucleophile and carboxylic acid containing molecules in a cell is likely to be in excess of both the tetrazole reagent and alkene

tag. The rate of reaction D would, therefore, be significantly greater than reaction B in the presence of 5 mM glutathione, for example.

Overall, the results obtained suggest that, despite the literature precedent, the formation of a fluorescent pyrazoline adduct of small amino acid tags is unlikely to work in a cellular environment.

1.5. Conducting a cell-based experiment

All of the data obtained suggest that the conditions present in a cell are likely to inhibit the formation of the pyrazoline adduct and, as such, led to the prediction that a cell based experiment is likely to be unsuccessful. Having stated this, the design of this methodology was based on published literature experiments conducted with an allyl-Tyr or Homo-allyl-Gly residue present on a protein in *E.coli* or HeLa cells, suggesting some probability of success when applied in a cellular environment.[80,97] Therefore, in order to confirm this prediction, a cell based experiment was conducted using Boc-allyl-Tyr-OMe **105** and Boc-methacryloyl-Tyr-OMe **106**.

Literature conditions were adopted with 100 µM tetrazole **81** and 14 µM of each tag in HeLa cells using 5% DMSO in PBS as the solvent system.[80,97] The cells were prepared with a DRAQ5 nuclear stain; incubated at 37 °C with the reagents and controls for 30 minutes; irradiated with a UVB lamp (270-330 nm) for 3 minutes; and finally, the fluorescence monitored by confocal microscopy. A separate plate which was not exposed to UV light was used as a primary control. As a second control all experiments were repeated with cells that had been exposed to 0.01% Triton, a detergent which permeabilises the cell membrane; the purpose of which was to ensure that all reagents were able to penetrate the cell.[131]

The results are depicted in Table 3 and were normalised to represent the non-irradiated solvent control in untreated HeLa cells as a value of 1.0, with all remaining fluorescence intensities being relative to this.

a)

Plate	5% DMSO:PBS	105	106	81	Reaction 105 + 81	Reaction 106 + 81
No light	1.0	3.4	1.3	4.2	3.0	1.0
Irradiated	0.6	1.6	0.7	4.7	5.0	4.8

b)

Plate	5% DMSO:PBS	105	106	81	Reaction 105 + 81	Reaction 106 + 81
No light	0.9	4.8	0.6	0.7	1.3	2.0
Irradiated	3.2	7.0	1.3	5.3	7.3	4.4

Table 3: Fluorescence intensity observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for solvent controls, reagent controls, and reaction mixtures after irradiation (270-330 nm, 3 min) in (a) HeLa cells and (b) HeLa cells permeablised with 0.01% Triton. Concentrations: **81** (100 µM), **105** (14 µM), and **106** (14 µM) in 5 % DMSO:PBS.

The initial results obtained for the reagent controls and the two reaction mixtures showed significant variability. This was observed between the data sets obtained for those cells not exposed to the detergent and the cells that were exposed to it, and also between data sets for the non-irradiated and irradiated controls, which had previously been shown to exhibit the same fluorescence intensity under both conditions (see Section 1.3). A particularly anomalous result was obtained for the Boc-allyl-Tyr-OMe **105** control which showed higher levels of fluorescence intensity than Boc-methacryloyl-Tyr-OMe **106**, despite having previously been shown to give the same low level of background intensity.

Broadly speaking, the trends observed between the cells treated with detergent and those untreated were similar, indicating that the reagents had likely managed to penetrate the cell membranes. Most noticeably, in both sets of data the reaction

mixtures show a small increase in fluorescence when irradiated; however, the intensity of fluorescence observed was similar to that observed for the tetrazole control. Observation of the cells by confocal microscopy did not show any clear differentiation between the tetrazole control and the two reaction mixtures (Figure 57); however, it did appear to confirm the cell permeability of the tetrazole reagent.

a) b) c)

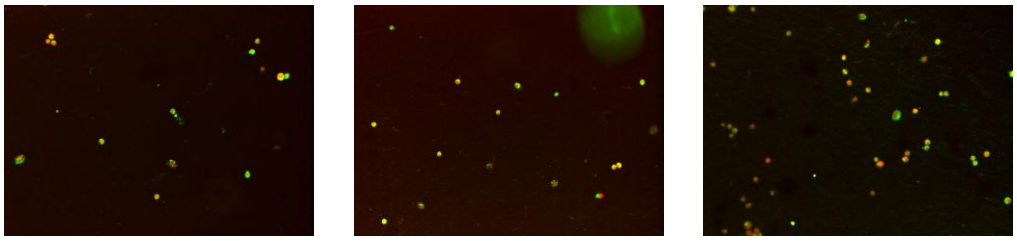


Figure 57: Confocal microscope images of cells exposed to (a) Tetrazole **81** control; (b) Reaction mixture containing Boc-allyl-Tyr-OMe **105** and tetrazole **81**; (c) Reaction mixture containing Boc-methacryloyl-Tyr-OMe **106** + tetrazole **81**.

This indicated that, if the desired pyrazoline formation was occurring, it was not exhibiting sufficient intensity to overcome the background fluorescence observed from the control reactions. This negative result was exacerbated by the observation of poor dissolution of the tetrazole reagent in the control prior to irradiation (Figure 58). Therefore, although the tetrazole had been confirmed as being able to penetrate the cell membrane, the amount present inside the cell itself at the time of irradiation was unquantifiable.

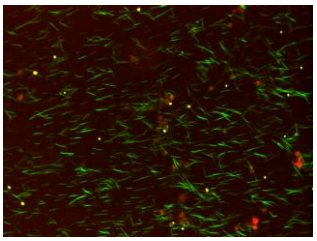


Figure 58: Crystals of tetrazole **81** observed as a precipitate from the 5% DMSO:PBS solution applied to the cell experiment.

In order to confirm the negative outcome of the cell based experiment, the two pyrazoline adducts **118** and **119** were introduced to the cells under the same conditions at concentrations of 3.0 µM (20 % conversion), 7.5 µM (50 % conversion), and 15.0 µM (100 % conversion). The results are depicted in Table 4.

a)

Plate	**118** (3.0 µM)	**118** (7.5 µM)	**118** (15.0 µM)	**119** (3.0 µM)	**119** (7.5 µM)	**119** (15.0 µM)
Control	14.8	20.5	16.7	21.9	21.2	19.3
Irradiated	13.7	16.8	17.0	21.6	21.3	18.3

b)

Plate	**118** (3.0 µM)	**118** (7.5 µM)	**118** (15.0 µM)	**119** (3.0 µM)	**119** (7.5 µM)	**119** (15.0 µM)
Control	18.5	19.0	15.3	21.9	21.6	17.2
Irradiated	21.4	20.6	17.0	21.7	20.5	15.7

Table 4: Fluorescence intensity observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for pyrazoline adduct **118** and **119** controls at concentrations of 3.0 µM, 7.5 µM, and 15.0 µM in (a) HeLa cells and (b) HeLa cells permeabilized with 0.01% Triton.

Comparison of these data to those obtained for the reaction mixtures indicated that if the pyrazoline adducts had been synthesised inside the cell then the intensity of the fluorescence would be clearly distinguishable from any background fluorescence, even at only 20% conversion. The fluorescence intensity was not concentration dependant as expected due to saturation of the confocal microscope, even at the lower concentrations. The higher intensities of the product controls clearly indicated that the formation of the pyrazoline adducts in the cell based experiment had been unsuccessful and that the use of this technique for quantitative measurement of peptide permeability would be unfeasible. Images taken of the cells containing 15 µM concentrations of the pyrazoline adducts are also clearly distinguishable from those taken for the reaction mixture (Figure 59).

a) b)

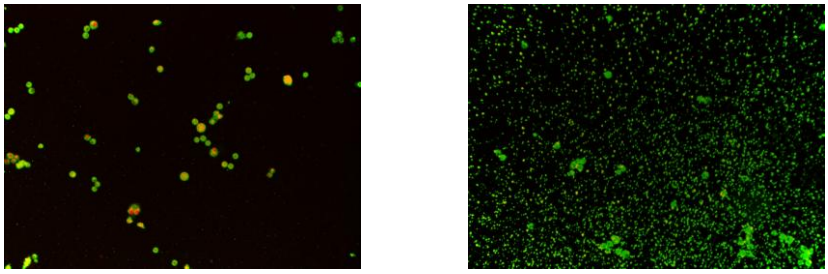


Figure 59: Confocal microscope images for (a) pyrazoline adduct **118** (15 μM) and (b) pyrazoline adduct **119** (15 μM).

Finally, the key byproducts identified from the competing side reactions discussed in Section 1.4. were also applied to the cells under the same conditions, assuming 100% conversion (Table 5).

a)

Plate	**Triazole 117** (100 μM)	**Pyrazole 120** (15 μM)	**Hydrazonothioate 150** (100 μM)	**Hydrazide 152** (100 μM)	**Hydrazide 153** (100 μM)
control	5.4	0.8	1.2	0.5	1.3
Irradiated	2.7	0.7	1.2	0.5	1.4

b)

Plate	**Triazole 117** (100 μM)	**Pyrazole 120** (15 μM)	**Hydrazonothioate 150** (100 μM)	**Hydrazide 152** (100 μM)	**Hydrazide 153** (100 μM)
control	2.4	0.7	2.0	0.7	0.7
Irradiated	2.8	0.7	1.1	0.7	0.8

Table 5: Fluorescence intensity observed (λ_{ex} = 350 nm, λ_{em} detection = 470 nm) for byproduct controls in (a) HeLa cells and (b) HeLa cells permeabilized with 0.01% Triton. Concentrations: **120** (15 μM), **117** (100 μM), **152** (100 μM), **150** (100 μM), and **153** (100 μM).

Of the five byproducts analysed the triazole shows slightly higher levels of fluorescence than the others. This would suggest that the formation of the triazole byproduct dominated over the other side reactions in the cell. However, the competition reactions conducted previously identify hydrazide **152** as the most likely byproduct. It is possible that accumulation of the tetrazole and the hydrophobic effect results in triazole formation dominating instead.

As a result of all of the investigations discussed above, and despite promising literature precedent, it can be concluded that the fluorogenic click reaction is not a suitable tool for the intracellular visualisation of alkene tagged peptides.

1.6. Exploring the potential applications of side reactions

1.6.1. Synthesis of 2,4,5-trisubsituted 1,2,3-triazoles

Despite the unsuccessful intracellular fluorogenic click reaction, the unexpected formation of 2,4,5-trisubstituted 1,2,3-triazoles from photochemical activation of 2,5-diaryl tetrazoles presented an opportunity to explore this unexploited reaction. A literature survey revealed that 1,2,3-triazoles are an important class of heterocycles due to a number of useful properties. *N*2-alkyl and acyl 1,2,3-triazoles act as PPAR agonists[132] and as selective CB1 receptor antagonists **157** (Figure 60).[133] *N*2-aryl 1,2,3-triazoles have application as β_3-Adrenergic receptor agonists **159**;[134] but more uniquely, have photonic properties which make them suitable as biocompatible UV/blue-light emitting dyes,such as **158**, and inhibitors of polymer degradation.[135,136] Furthermore, there are currently few examples of converting 2,5-diaryl tetrazoles to triazoles under photochemical conditions.[105,137] Issues of regioselectivity between the *N*1 and *N*2 position in many synthetic routes means that the synthesis of *N*2-substituted 1,2,3-triazoles remains a challenge.

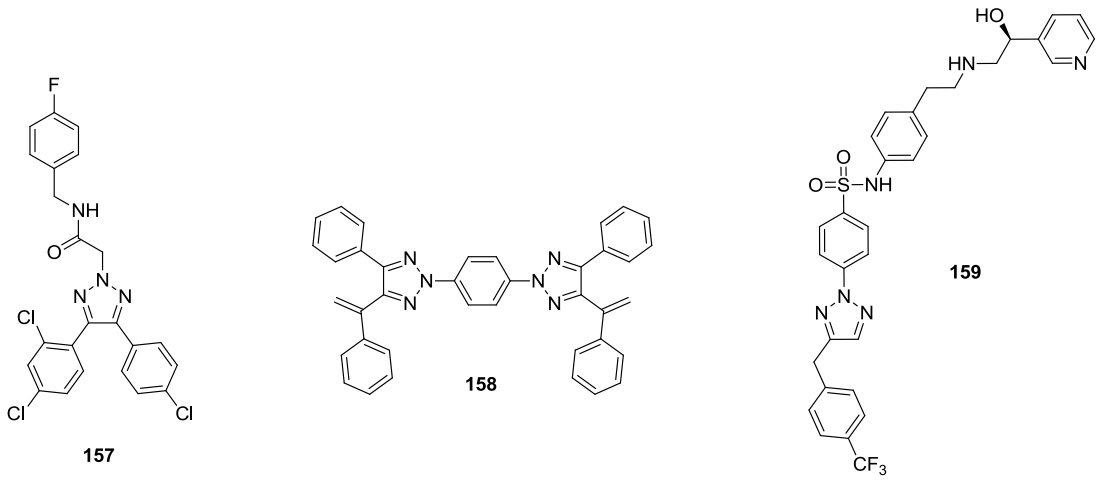
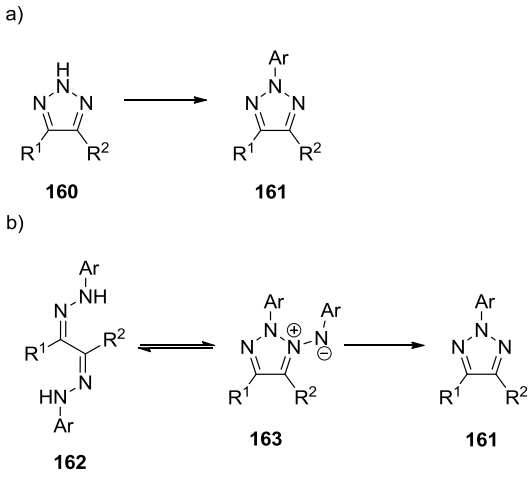


Figure 60: Exemplary structures of 1,2,3-triazoles as CB1 receptor antagonists **157**, UV/Blue light emitting fluorophores **158**, and β_3-Adrenergic receptor agonists **159**.

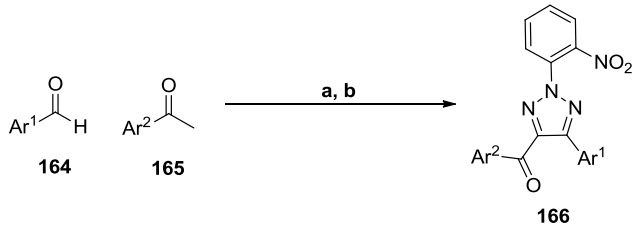
Currently, *N*2-aryl 1,2,3-triazoles **161** are synthesized via two main reactions: (a) arylation of NH-1,2,3-triazoles **160** or (b) rearrangement of bis(arylhydrazones) **162** via 1,2,3-triazolium 1-aminides **163** (Scheme 31).



Scheme 31: Common routes to *N*2-aryl 1,2,3-triazoles **161** via (a) post-triazole arylation of NH-free 1,2,3-triazoles **160** and (b) cyclisation of bis(arylhydrazone) intermediates **162**.

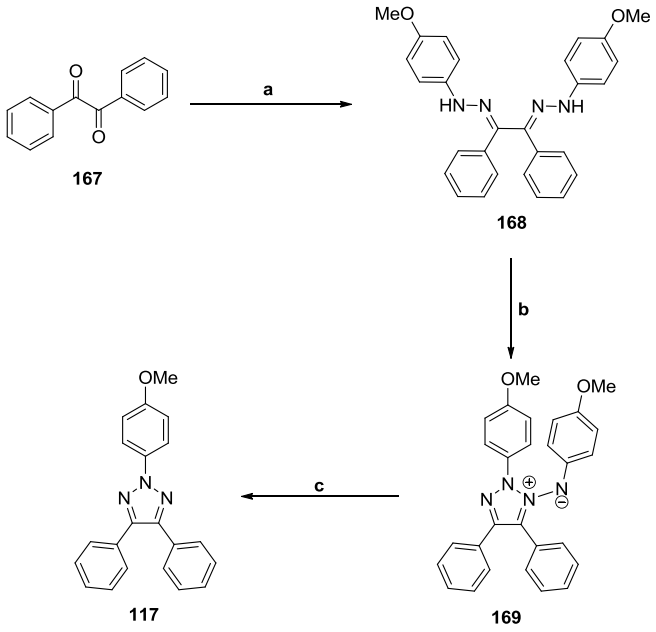
Shi *et. al.* reported regioselective copper catalysed, or S_NAr mediated, *N*2-arylation utilising bulky ligands or bulky 4,5-substituents.[138] Palladium catalysis has also been reported for this application.[139] Building on this work, Chen *et. al.* have demonstrated the regioselective generation of NH-1,2,3-triazoles followed by subsequent arylation using one-pot multi-component reactions starting from a

variety of aldehydes **164** and ketones **165** (Scheme 32).[140,141,142] However, the scope of these S_NAr reactions was limited to substrates with electron poor *N2*-aryl substituents such as *o*-NO_2 substituted triazoles **166**.



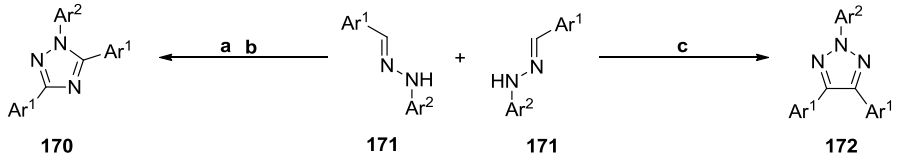
Scheme 32: Multi-component synthesis of electron poor *N*2-nitroaryl substituted 1,2,3-triazoles. Reagents and conditions: (a) CuO, NaN_3, $KHCO_3$, DMSO, 80 °C; (b) 2-$NO_2C_6H_4F$.

Syntheses from bis(arylhydrazones) **162** can proceed via acid catalysed rearrangement to 1,2,3-triazolium-1-aminides **163**. Subsequent photolysis or reduction with carbon disulfide and hydrogen sulfide yields the triazole.[115,143] Butler *et. al.* favoured formation of bis(arylhydrazone) **168** via condensation of benzil **167** and aryl hydrazines (Scheme 33). This was followed by oxidation to the 1,2,3-triazolium-1-aminide **169** and subsequent reduction to *N2*-aryl 1,2,3-triazole **117** using lead oxide and phosphorous trichloride, two highly toxic chemicals.[113]



Scheme 33: Synthesis of *N*2-aryl-1,2,3-triazole **117** via bis(arylhydrazone) **168**. Reagents and conditions: (a) *p*-anisylhydrazine hydrochloride, $NaOAc.3H_2O$, EtOH, reflux, 17 h, 64%; (b) PbO_2, DCM, rt, 24 h, 79%; (c) PCl_3, reflux, 1 h, 79%.

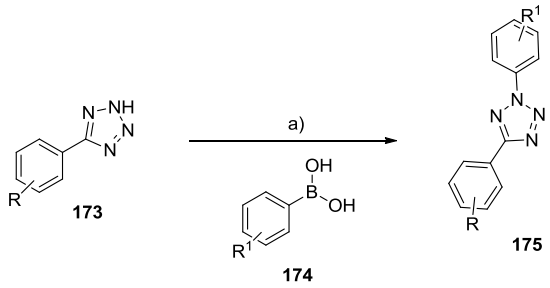
Most recently, Guru *et. al.* have shown efficient and regiospecific copper catalysed formation of triaryl-1,2,3-triazoles **170** or 1,2,4-triazoles **172** from arylhydrazones **171** under aerobic conditions (Scheme 34).[114] This synthesis provides a simple route to 1,2,3-triazoles under relatively mild conditions.



Scheme 34: Regiospecific, copper-catalysed synthesis of 1,2,3- **172** or 1,2,4-triazoles **170**. Reagents and conditions: (a) DABCO, dioxane, 60 °C, 30 h; (b) Cu(II) catalyst, air, rt; (c) Cu(II) catalyst, toluene, air, 60 °C.

It was desirable to investigate if the photochemical activation of tetrazoles could provide a similarly mild route to access 1,2,3-triazoles. In order to investigate the substrate scope of the reaction it was first necessary to identify a means of synthesising the requisite 2,5-diaryl tetrazoles. Opportunely, a copper catalysed synthesis of 2,5-diaryl tetrazoles has recently been developed which facilitated the synthesis of 2,5-diaryl tetrazoles **175a-q** from commercially available 5-aryl NH-tetrazoles **173** and aryl boronic acids **174** in 49-93% yield (Scheme 35, Table 6).[144]

The reaction proceeded efficiently with the successful isolation of the desired 2,5-diaryl tetrazoles **175a-q**. The reaction was found to be suitable for electron withdrawing or donating substituents on the phenyl rings; however, attempting the reaction with heteroaromatic 2-(2*H*-tetrazol-5-yl)pyridine failed to yield the desired substrate, presumably due to coordination with the copper catalyst.



Scheme 35: Synthesis of 2,5-diaryl tetrazoles **175**. Reagents and conditions: (a) 5 mol % Cu_2O, Boronic acid **150** (2 equiv.), O_2 (1 atm), DMSO, 110 °C.

Entry	R	R^1	Yield/ %	Time/ h
175a	H	H	86	6
175b	*p*-Cl	H	64	72
175c	*m*-Cl	H	84	24
175d	*o*-Cl	H	85	24
175e	*p*-OMe	H	-[a]	-
175f (88)	H	*p*-Me	93	6
175g (81)	H	*p*-OMe	-[a]	-
175h	H	*m*-OMe	77	16
175i	H	*o*-OMe	49	16
175j (89)	H	*p*-NH_2	-[b]	-
175k	H	*p*-NH(CO)Me	-[c]	-
175l	H	*p*-CO_2Et	75	24
175m	H	*m*-CN	60	6
175n	H	*p*-NO_2	55	24
175o	*p*-OMe	*p*-OMe	-[a]	-
175p	*p*-Cl	*p*-OMe	70	22
175q	*p*-SO_2NMe_2	*p*-OMe	79	16

[a]Commercially available

[b]Synthesised by reduction of tetrazole **175n** (86 %)

[c]Synthesised by acetylation of tetrazole **175j** (92%)

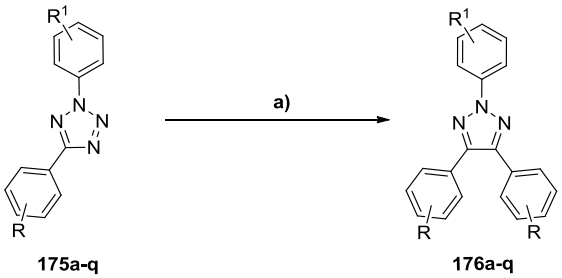
Table 6: Substituted 2,5-diaryl tetrazole analogues **175a**-**q** synthesized via the copper catalyzed coupling of **173** and **174**.

With these substrates in hand it was then possible to investigate the effect of substituents on the synthesis of 2,4,5-triaryl-1,2,3-triazoles. Firstly, a solvent screen was conducted and a number of solvents were shown to be suitable, including DCM, MeCN, IPA, TBME, 2-MeTHF, EtOAc, toluene and DMF. Further experimentation revealed that a mixture of 2-MeTHF:IPA (1:1 v/v) was a suitable environmentally sustainable solvent system for the reaction.[145] Furthermore, this combination of solvents enabled precipitation of some of the products for easier isolation and purification.

Tetrazoles **175a**-**q** were then irradiated with a UVB lamp (270-330 nm) in 2-MeTHF:IPA under atmospheric conditions until all starting material had been consumed (Scheme 36, Table 7).

Tetrazole **175a** was chosen as a baseline for comparison of substituent effects and gave triazole **176a** in 19% yield. Initial substituent effects were investigated with *C*-aryl substitution (Entries **176b**-**176e**). Electron donating substituents appeared to improve yields slightly relative to **176a**. Substitution in the *meta* and *para* positions was tolerated; however, *o*-Cl substitution failed to yield triazole **176d**. It was reasoned that this substitution was not tolerated due to steric effects in the nitrile imine intermediate disfavouring the dimerisation reaction.

The electronic donating effect of substituents on the *N*-aryl ring were then probed (Entries **176f**-**176k**). In general, addition of electron donating groups favoured formation of the desired triazoles in significantly improved yields, particularly in the *para* position. Interestingly, *ortho* substitution is also tolerated, albeit in moderate yield (entry **176i**). Substitution with *p*-NH_2 was expected to give excellent reactivity; however, triazole **176j** could only be isolated in 17% yield. Protecting the reactive aniline moiety as an acetamide derivative afforded triazole **176k** in an improved yield of 41%. Alternatively, irradiating tetrazole **175j** under more dilute conditions also afforded an improved yield of 46% by reducing byproduct formation.



175a-q 176a-q

Scheme 36: Photochemical synthesis of 2,4,5-triaryl-1,2,3-triazoles **176**. Reagents and conditions: (a) Compound **175a-q** (1 equiv.), UV light (270-330 nm), 2-MeTHF:IPA (1:1 v/v), rt, air.

Entry	R	R^1	Yield/ %[a]	Time/ h
176a	H	H	19	24
176b	*p*-Cl	H	27	16
176c	*m*-Cl	H	28	31
176d	*o*-Cl	H	0	26
176e	*p*-OMe	H	38	30
176f (115)	H	*p*-Me	47	19
176g (117)	H	*p*-OMe	62	10
176h	H	*m*-OMe	25	35
176i	H	*o*-OMe	33	40
176j (116)	H	*p*-NH_2	17 (46)[b]	48 (7)[b]
176k	H	*p*-NH(CO)Me	41	6
176l	H	*p*-CO_2Et	14	23
176m	H	*m*-CN	0	24
176n	H	*p*-NO_2	0	>100
176o	*p*-OMe	*p*-OMe	90	7
176p	*p*-Cl	*p*-OMe	67	16
176q	*p*-SO_2NMe_2	*p*-OMe	0	6

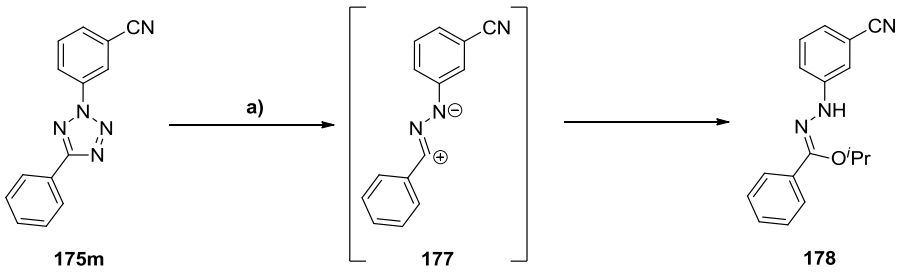
[a]Yield recalculated from a maximum yield of 50%

[b]2-MeTHF:IPA solvent (5x volume)

Table 7: Substituted 2,4,5-triaryl-1,2,3-triazoles **176a-q** synthesised via the dimerisation of nitrile imines photochemically generated from 2,5-diaryl tetrazoles **175a-q**.

As expected, electron deficient *N*-aryl substituents were poorly tolerated (Entries **176l**-**176n**). Of these, only *p*-CO_2Et tetrazole **175l** yielded the desired triazole in a low yield of 14%. Photoactivation of *p*-NO_2 tetrazole **175n** was inefficient with starting material predominating after irradiation for 100 hours and no discernible product evident after this time.

Incorporation of a *m*-CN substituent in tetrazole **175m** did not afford the expected triazole **176m**. Instead, hydrazonate byproduct **178**, formed by trapping of the nitrile imine **177** with IPA, was isolated as the major product with a yield of 54% (Scheme 37). Such quenching reactions have been reported previously with nucleophilic amines and carboxylic acids.[100,146] Repeating the reaction in non-nucleophilic ethyl acetate also failed to yield the desired product, presumably due to poor reactivity resulting from the electron withdrawing CN substituent. It should be noted that solvent trapping of the nitrile imine was not observed with the more reactive tetrazole substrates.

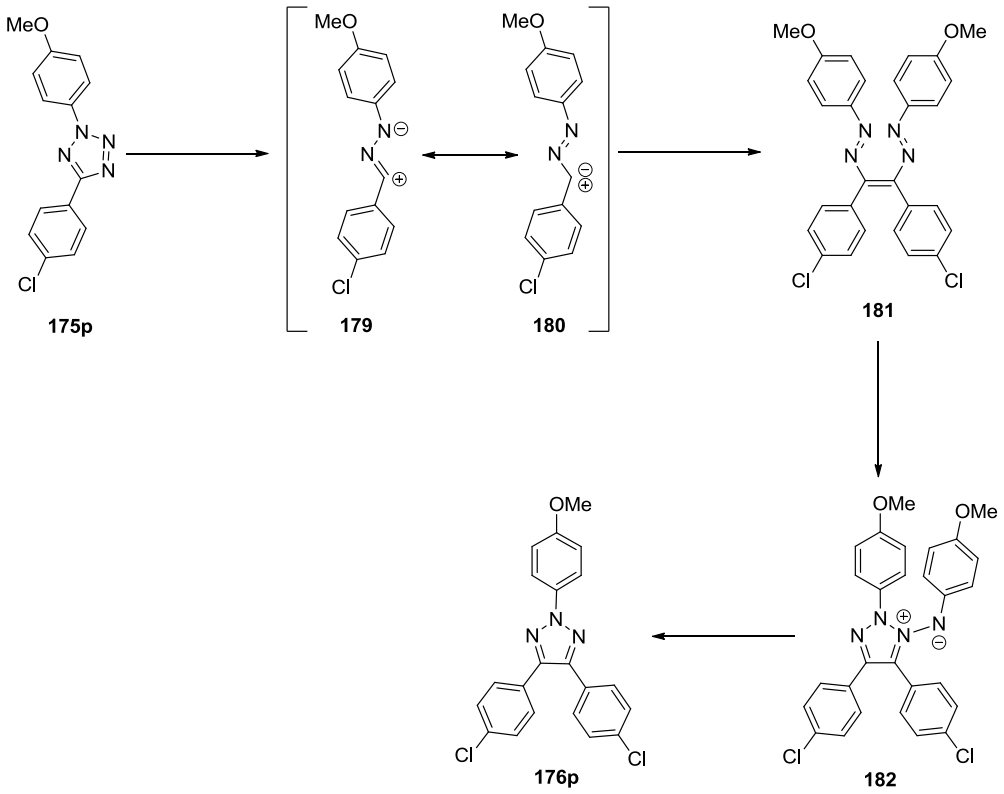


Scheme 37: Trapping of nitrile imine **177** with IPA. Reagents and conditions: (a) Compound **175m** (1 equiv.), UV light (270-330 nm), 2-MeTHF:IPA (1:1 v/v), air, rt, 24 h, 54%.

Finally, the observed trends in reactivity were combined when studying triazoles **176o**-**q**. As predicted, combining an electron rich *N*-aryl ring with electron donating groups on the *C*-aryl ring gave significant improvements in yields. However, the deleterious electron withdrawing effects of a sulfonamide moiety could not be surmounted (entry **176q**).

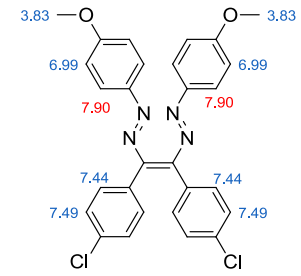
Interestingly, during the synthesis of triazole **176p** a precipitate was formed and, when isolated, Wanzlick dimer **181** was obtained in 57% yield instead of the

expected triazole (Scheme 38). Previously, photochemical formation of this intermediate had only been observed spectroscopically.[105] Further exposure to UV light yielded the desired product by rearrangement to 1,2,3-triazolium-1-aminide **182** and subsequent N-N bond cleavage.



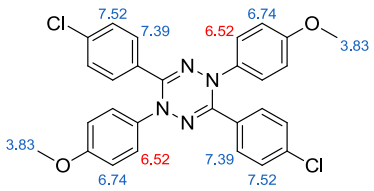
Scheme 38: Proposed mechanism of *N*2-aryl 1,2,3-triazole formation.

Formation of the Wanzlick dimer may occur through the carbenic resonance form of the nitrile imine **180** which explains the tolerance of *o*-substituents on the *N*-aryl ring but not the *C*-aryl ring.[147] Based on these observations, it may also be inferred that the tetrazine byproducts previously reported during pyrazoline syntheses may have been Wanzlick dimers as their NMR profile would appear similar to a tetrazine,[80,86] with the exception of the protons in the *meta* position relative to the methoxy groups. For example, Figure 61 represents the close similarity between Wanzlick dimer **181** and the alternative tetrazine structure **183**.



181

Chemical Formula: $C_{28}H_{22}Cl_2N_4O_2$
Exact Mass: 516.1120



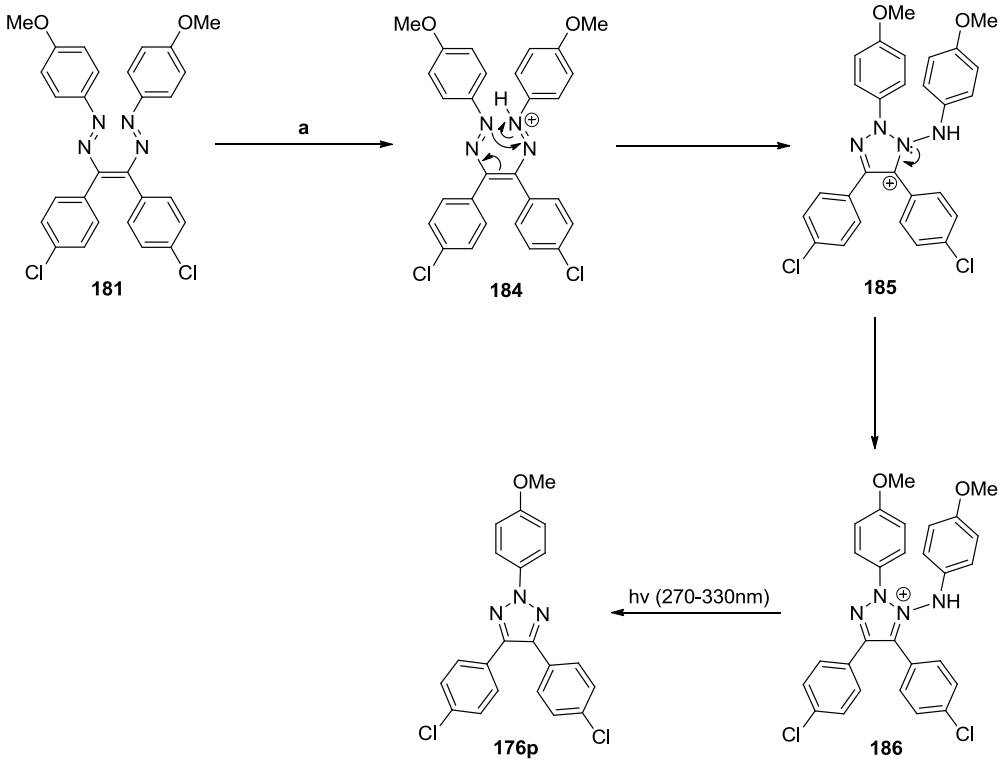
183

Chemical Formula: $C_{28}H_{22}Cl_2N_4O_2$
Exact Mass: 516.1120

Figure 61: Similarities between Wanzlick dimer **181** and tetrazine **183** with calculated chemical shifts (BioChemDraw Ultra 12.0) for the proton NMR of each molecule represented in blue. The clear difference in chemical shifts for the *meta* protons of the anisole ring are represented in red.

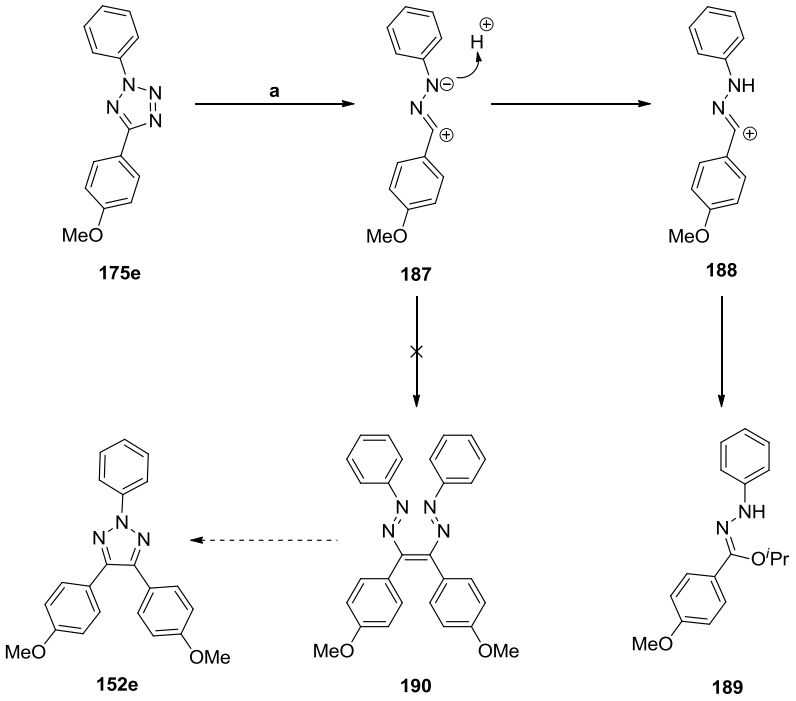
The mass of the two structures can also be seen to be identical. Interestingly, during the synthesis of the pyrazoline adducts the mass associated with these structures could be observed by LC-MS as part of the mass spectrum of the triazole byproducts. It was concluded after positive confirmation of the triazole structures by NMR that these observed masses were a result of small quantities of residual Wanzlick dimer which appear to have a strong ionisation.

During NMR analysis complete conversion of isolated intermediate **181** to triazole **176p** was observed in $CDCl_3$ over a period of 23 hours while a sample prepared in DMSO-d_6 only showed 43% conversion during the same time period. It was hypothesised that the mildly acidic nature of $CDCl_3$ could account for the more rapid conversion in this solvent. It was subsequently discovered that addition of a small amount of hydrochloric acid to isolated intermediate **181** resulted in immediate conversion to triazole **176p**. It is postulated that this is due to an acid catalysed rearrangement to protonated 1,2,3-triazolium-1-aminide **185** via protonated Wanzlick dimer **184** followed by cleavage of the N-N bond in the ambient light (Scheme 39). Such acid catalysed rearrangements have been previously documented.[143]



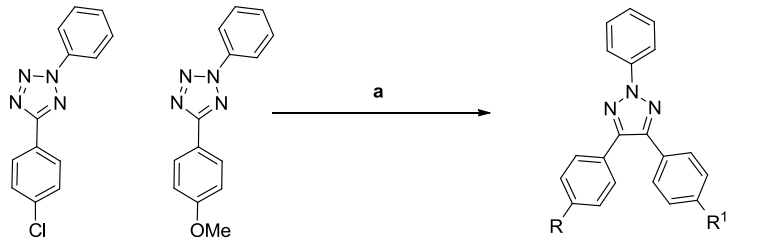
Scheme 39: Acid catalysed rearrangement of intermediate **181** to triazole **176p**. Reagents and conditions: (a) Bis(arylhydrazone) **181** (1 equiv.), cat. 4M HCl in dioxane, 2-MeTHF:IPA (1:1 v/v), ambient light, rt, air, unisolated.

Based on this knowledge, the photochemical irradiation of tetrazole **175e** was repeated in the presence of catalytic hydrochloric acid in an attempt to improve the previously obtained 38% yield of triazole **176e**. However, the photochemically induced loss of N_2 from tetrazole **175e** in the presence of acid failed to provide the desired triazole and the nitrile imine was instead quenched by IPA to give byproduct **189**. This is most likely due to the quenching of the nitrile imine anion by the acid, thus preventing carbenic formation of dimer **190**, and leaving a carbocation available for nucleophilic attack by the solvent (Scheme 40).



Scheme 40: Photochemical formation of nitrile imine **187** in the presence of small quantities of acid, resulting in nucleophilic quenching instead of triazole formation. Reagents and conditions: (a) Tetrazole **175e** (1 equiv.), cat. 4M HCl in dioxane, UV light (270-330 nm), 2-MeTHF:IPA (1:1 v/v), air, rt, 1.5 h, unisolated.

Having synthesised a range of symmetrical triazoles with a range of substituents, it was next desirable to attempt the synthesis of unsymmetrical examples. Indeed, such reactions had also been conducted by Guru *et. al.* using the copper catalysed coupling of aryl hydrazones described above to give triazole **191**.[114] The photochemical synthesis of this molecule was attempted with tetrazoles **175b** and **175e** in order to compare the utility of the methodologies (Scheme 41, Table 8).



Scheme 41: Synthesis of mixed 2,4,5-triaryl-1,2,3-triazole **191**. Reagents and conditions: (a) Tetrazole **175b** (1, 2 or 5 equiv.), tetrazole **175e** (1 equiv.), UV light (270-330 nm), 2-MeTHF:IPA (1:1 v/v), rt, air.

		LC-MS Proportions			
Equiv. 175b	**Equiv. 175e[b]**	**176b[a]**	**176e[a]**	**191[a]**	**Isolated Yield of 191 (%)**
1	1	0.4	1	0.75	0
2	1	1.11	1	1.11	5
5	1	1.43	1	1.57	18

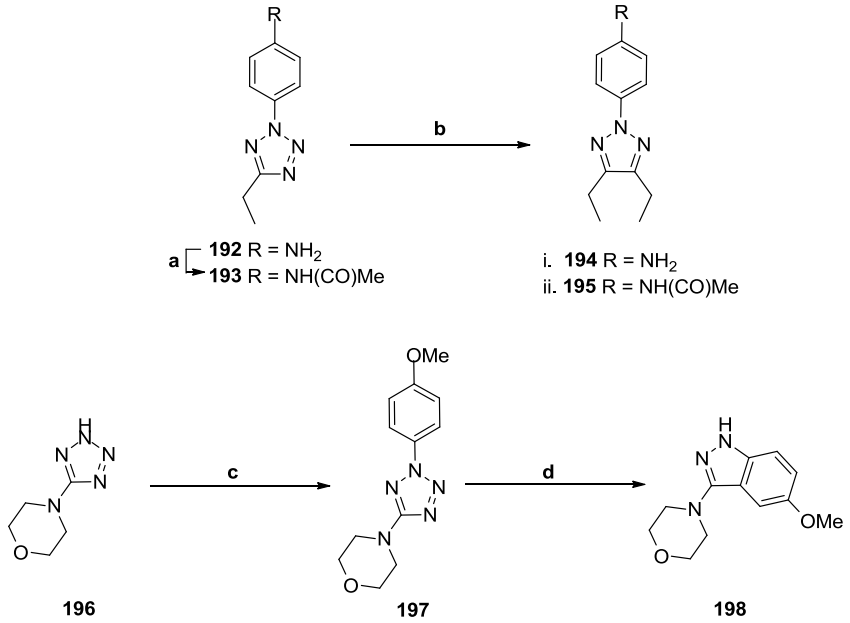
[a]Ratios determined by LC-MS analysis

[b]Mixture irradiated until tetrazole **175e** fully consumed

Table 8: Proportions of 2,4,5-triaryl-1,2,3-triazoles **176b**, **176e**, and **191** observed during photochemical irradiation of mixtures of tetrazole **175b** and **175**e.

Guru *et. al.* conducted their synthesis in the presence of one equivalent of each hydrazone to give triazole **191** in 70% yield. However, when the photochemical synthesis of **191** was attempted using one equivalent of **175b** and **175e** the desired triazole could be observed by LC-MS but could not be isolated as the reaction mixture was found to contain multiple components which prevented successful purification. Triazole **191** could be isolated by increasing the equivalents of tetrazole **175b** from one to five but only in a low yield of 18%. As can be noted from the data obtained, the dimerisation of the nitrile imine from tetrazole **175e** is more rapid than the coupling with the nitrile imine intermediate from tetrazole **175b**. It is possible that the proportion of cross coupling could be increased by attempting to match the HOMO-LUMO energies of the two tetrazole starting materials; however, these reactions were not attempted as it would be unlikely to compete with the methodology developed by Guru and co-workers.

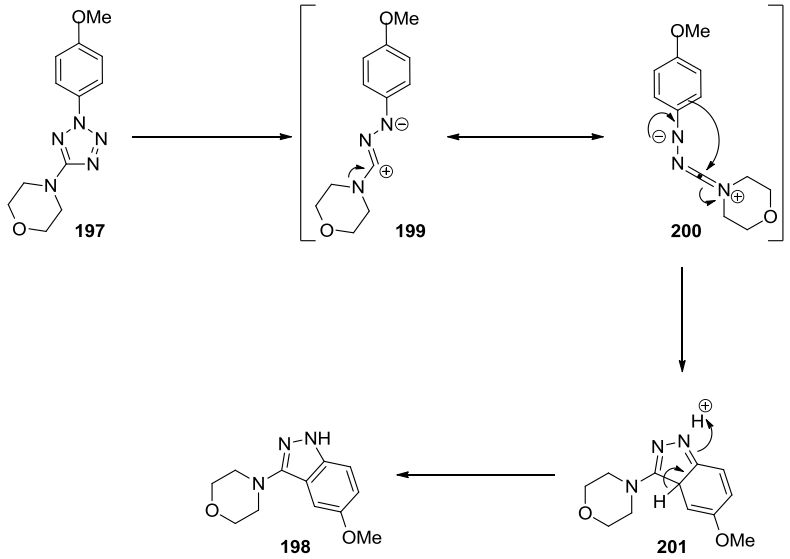
Having investigated the synthesis of 2,4,5-triaryl-1,2,3-triazoles, the scope of the reaction was expanded to determine if *C*-heterocyclic or *C*-alkyl groups could be tolerated (Scheme 42). Commercially available tetrazole **192** was irradiated to give triazole **194** in a modest 23% yield. In this case, masking the aniline with an acetamide moiety to give tetrazole **193** failed to improve the yield of the corresponding triazole **195**, indicating that *C*-alkyl groups are unfavourable. This could be attributed to a less stable carbocation in the nitrile imine intermediate, compared to *C*-aryl systems, resulting in a complex mixture of products.



Scheme 42: Synthesis of a *N*2-aryl 4,5-dialky 1,2,3-triazoles and unexpected formation of substituted indazoles. Reagents and conditions: (a) Ac_2O (11 equiv.), DCM, rt, 5 h, 98%; (b) i. UV light (270-330 nm), 2-MeTHF:IPA (1:1 v/v), rt, 20 h, 23%; ii. UV light (270-330 nm), 2-MeTHF:IPA (1:1 v/v), rt, 20 h, 21%; (c) 5 mol % Cu_2O, (4-methoxyphenyl)boronic acid (2 equiv.), O_2 (1 atm), DMSO, 110 °C, 16 h, 48%; (d) UV light (270-330 nm), 2-MeTHF:IPA (1:1 v/v), rt, 7 h, 79%.

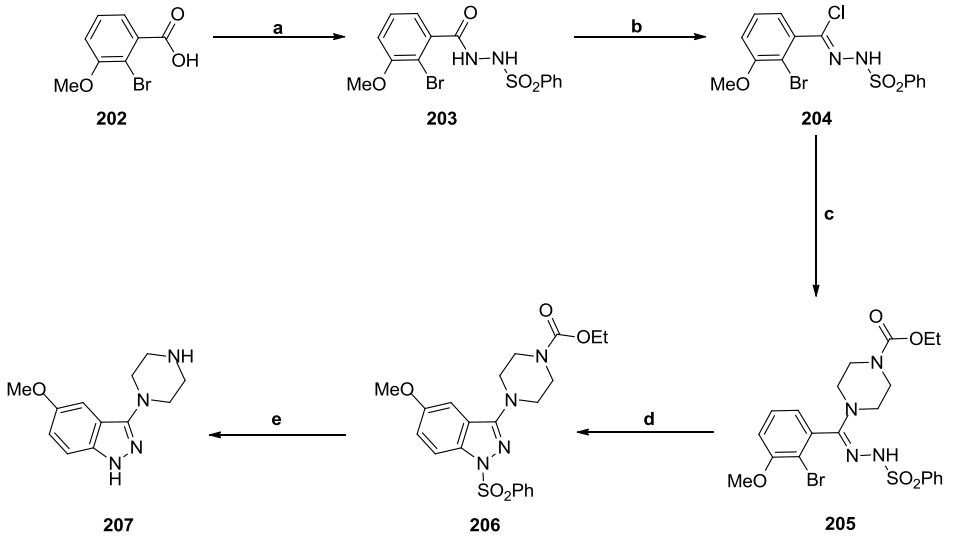
When irradiating *C*-morpholine tetrazole **197** the unexpected conversion to 5-methoxy-3-morpholine-*1H*-indazole **198** was observed with an isolated yield of 79% without optimisation. Such transformations from 2,5-diaryl tetrazoles have only been observed previously using forcing conditions such as temperatures of up to 400 °C.[148,149] A possible mechanism for this conversion is shown in Scheme 43.

The presence of the additional nitrogen atom in photochemically induced nitrile imine **199** could stabilise the carbocation by resonance, resulting in carbodiimide **200**. While this functional group is unable to undergo the Wanzlick dimer formation required for triazole formation, an intramolecular rearrangement can give bicycle **201** which after aromatisation gives indazole 198.



Scheme 43: Mechanism for the conversion of *C*-morpholine substituted tetrazole **197** to 5-methoxy-3-morpholine-1*H*-indazole **198**.

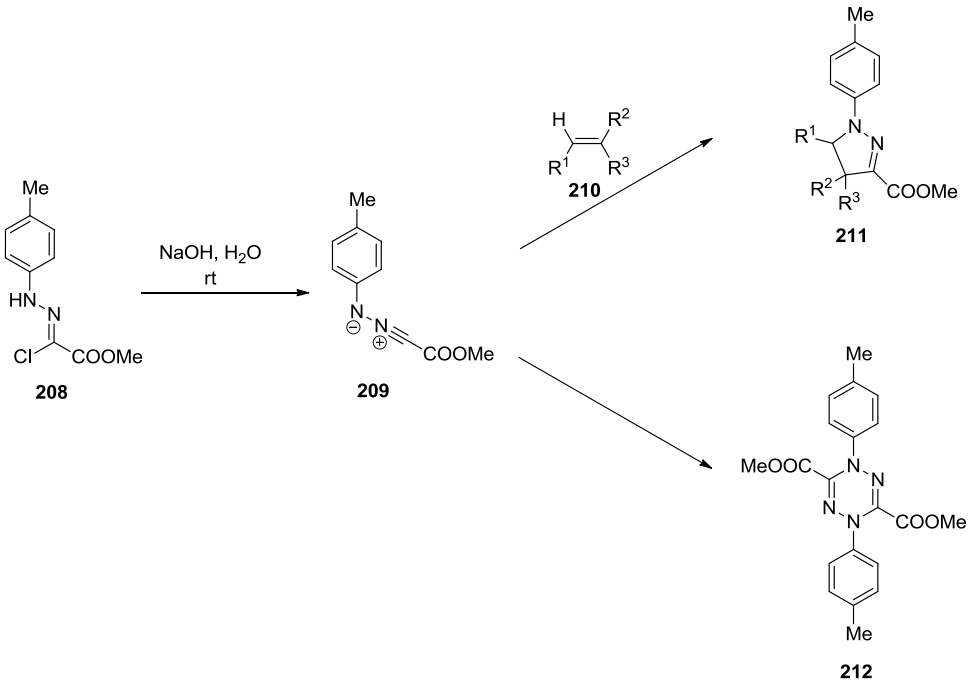
A literature preparation of the analogous indazole **207** has been reported in 5 steps and 38% yield from 2-bromo-3-methoxybenzoic acid **202** (Scheme 44).[150] Such scaffolds are of use in the pharmaceutical industry, for example in the preparation of treatments for obsessive compulsive disorder and attention deficit disorder.[150]



Scheme 44: Established process route to the synthesis of 5-methoxy-3-piperazinyl-*1H*-indazole **207**. Reagents and conditions: (a) i. H_2SO_4, EtOH, reflux, 20 h; ii. $NH_2NH_2.H_2O$, EtOH, reflux, 7 h; iii. py., $PhSO_2Cl$, -10-25°C, 2.5 h, 86% (three steps); (b) $SOCl_2$, reflux, 94%; (c) ethyl-1-piperazinecarboxylate, THF, 5-25°C, 8 h, 94%; (d) K_2CO_3, CuI, IPA, reflux, 5.5 h, 99%; (e) KOH, EtOH, reflux, 18 h, 51%. Overall yield: 38% (five steps).[150]

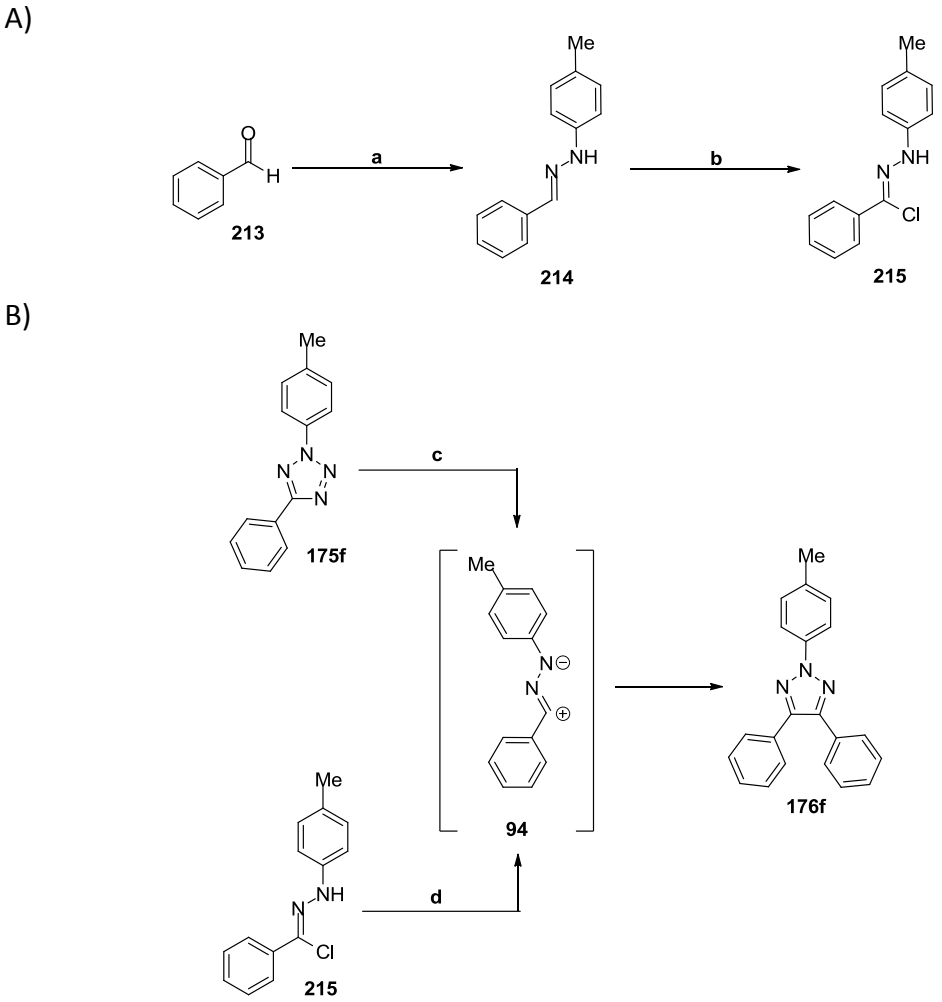
Although the yields are generally high the process requires high temperatures and careful manipulation of pressure for the use thionyl chloride on a large scale. Tetrazole **197** can be prepared in 48% yield, giving access to **198** in 79% yield and an overall unoptimised yield of 37% over the two steps, thus providing an efficient alternative route to this useful class of molecule.

Interestingly, the formation of tetrazine byproducts **212** isolated during the attempted synthesis of pyrazoline adducts **211** from *C*-acyl, *N*-Aryl hydrazonoyl chlorides **208** have been reported by Molteni *et. al.* (Scheme 45).[85]



Scheme 45: Synthesis of *N*2-aryl *C*5-acyl tetrazine **212** from substituted hydrazonoyl chloride **208**.

To determine if *C*-aryl substitution would instead yield triazole products hydrazonoyl chloride **215** was synthesised from benzaldehyde **213** via formation of hydrazone **214** with *p*-tolylhydrazine, followed by chlorination with the Corey-Kim reagent (Scheme 46a).[151] Nitrile imine **94** was generated from the resulting hydrazonoyl chloride using sodium hydroxide. Gratifyingly, the main product of the reaction was triazole **176f** (29% yield, unoptimised conditions), as confirmed by comparison to the previously obtained product (Scheme 46b).



Scheme 46: (A) Synthesis of hydrazonoyl chloride **215** from benzaldehyde. Reagents and conditions: (a) *p*-tolylhydrazine (1 equiv.), EtOH, 70 °C, 1.5 h, 71%; (b) NCS (1.7 equiv.), DMS (3 equiv.), DCM, -78-25 °C, 3 h, 68%. (B) Synthesis of triazole **176f** from 2,5-diaryl tetrazoles or hydrazonoyl chlorides. Reagents and conditions: (c) Compound **175f** (1 equiv.), UV light (270-330 nm), 2-MeTHF:IPA (1:1 v/v), air, rt, 19 h, 47%; (d) 2M $NaOH_{(aq)}$ (10 equiv.), THF, rt, 30 min, 29%.

As indicated by these experiments, the synthesis of 2,4,5-triaryl-1,2,3-triazoles is possible by forming nitrile imine intermediates from 2,5-diaryl tetrazoles or *N*-aryl, *C*-aryl hydrazonoyl chorides under photochemical or basic conditions, respectively. The reaction shows strong dependence on substituent effects, with electron donating groups giving the best reactivity and is primarily useful for the synthesis of electron rich 2,4,5-triaryl-1,2,3-triazoles. This mild, two step synthesis is, therefore, complementary to some of the other syntheses discussed previously, particularly those that are only efficient for electron withdrawing substituents. Furthermore,

the synthesis of *N*2-aryl 4,5-alkyl 1,2,3-triazoles was also shown to be viable, albeit in poor yield. Finally, the unexpected synthesis of 5-methoxy-3-morpholine indazole was shown to be possible under mild conditions and offers a more expedient means of accessing this template. The work conducted throughout Section 1.6.1 has been recently published.[152]

1.6.2. Photochemical labelling of peptide carboxylic acids

Despite having been noted as an interesting side reaction by Meier and Heimgartner during intramolecular pyrazoline syntheses in the presence of ethanoic acid (see Scheme 13, Section 1.2.2),[146] the formation of hydrazides via the photochemical reaction of 2,5-diaryl tetrazoles with carboxylic acids does not appear to have been exploited to date. While attempting the synthesis of pyrazoline adducts from alkene labelled amino acids with unprotected *C*-termini, the unexpected and selective formation of hydrazides was observed instead (see Section 1.2.2). During a literature survey of bioorthogonal labelling reactions a number of different approaches were identified for a variety of functional groups (see Section 1.1); however, carboxylic acids had not been identified as a potential functionality suitable for labelling.

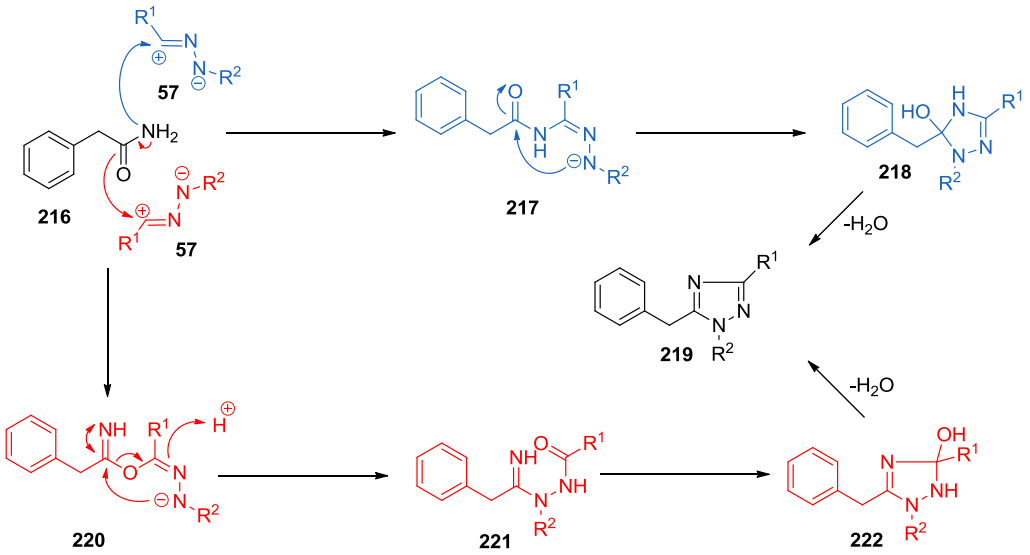
Indeed, the importance of this reaction was overlooked in the previously discussed publication by Levkin *et. al.* in which the group claimed to have selectively labelled a thiol chain in bovine serum albumin, even in the presence of residues with carboxylic acid side chains.[130] Multiple addition of the nitrile imine to the protein could be observed and are likely a result of hydrazide formation with these side chains, thus demonstrating the potential of this reaction as a labelling methodology.

With the majority of current labelling techniques requiring the synthesis of non-natural amino acids containing suitable functionality, the development of a labelling technique which could be conducted with any amino acid bearing a free *C*-terminus, or presumably any molecule bearing a carboxylic acid, would be practical. While this

may not be beneficial in a cellular environment where a large number of competing carboxylic acid groups would be present, it may be functional for use in *in vitro* experiments or for peptide and protein modifications.

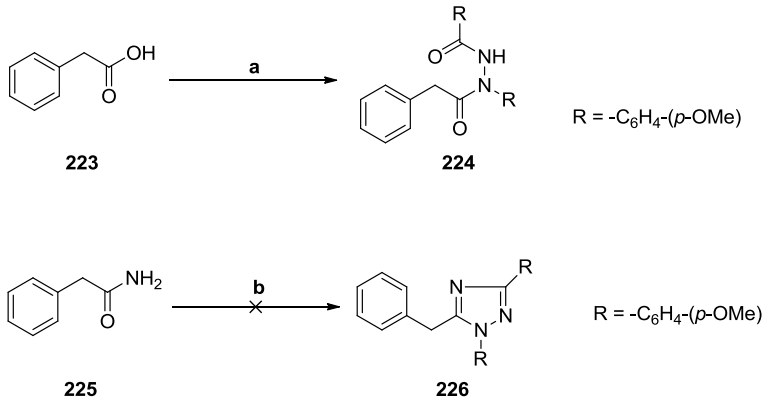
For this technique to be practical it would need to be high yielding and possess a high degree of chemoselectivity over other functional groups. Therefore, the synthesis of several model peptides consisting of natural amino acids with aliphatic, aromatic, acidic, basic, and nucleophilic functionality was carried out in order to test the practicality of the technique. Common tagging moieties, such as alkenes and alkynes, were also incorporated for chemoselectivity analysis.

The synthesis of the peptides was conducted using solid phase peptide synthesis (SPPS) techniques. Due to its immediate availability, Rink Amide MBHA 100-200 mesh was selected as the solid support. Cleavage of this resin leaves a primary amide at the *C*-terminus of the peptide and it was initially unclear if a photochemical reaction could be conducted in the presence of this functional group. This could result in additional undesired side reactions during hydrazide formation or could lead to the formation of 1,2,4-triazoles via plausible rearrangement mechanisms shown in Scheme 47. Nucleophilic quenching of the photochemically generated nitrile imine through either C-O or C-N bond formation followed by cyclisation and loss of water, driven by aromatisation of the ring, could eventually lead to such products.



Scheme 47: Two plausible pathways for the formation of 1,2,4-triazoles via the reaction of primary amides with 2,5-diaryl tetrazoles under photochemical conditions.

Control reactions were conducted to simultaneously investigate the formation of 1,2,4-triazoles and this possible chemoselectivity issue (Scheme 48). The reaction of phenyl acetic acid **223** with tetrazole **175o** under photochemical conditions gave the expected hydrazide **224** in 77% yield. The solvent trifluoroethanol (TFE) was used for the reaction as it is often used for solubilising peptides; no adverse effects on the reaction were observed under these conditions.

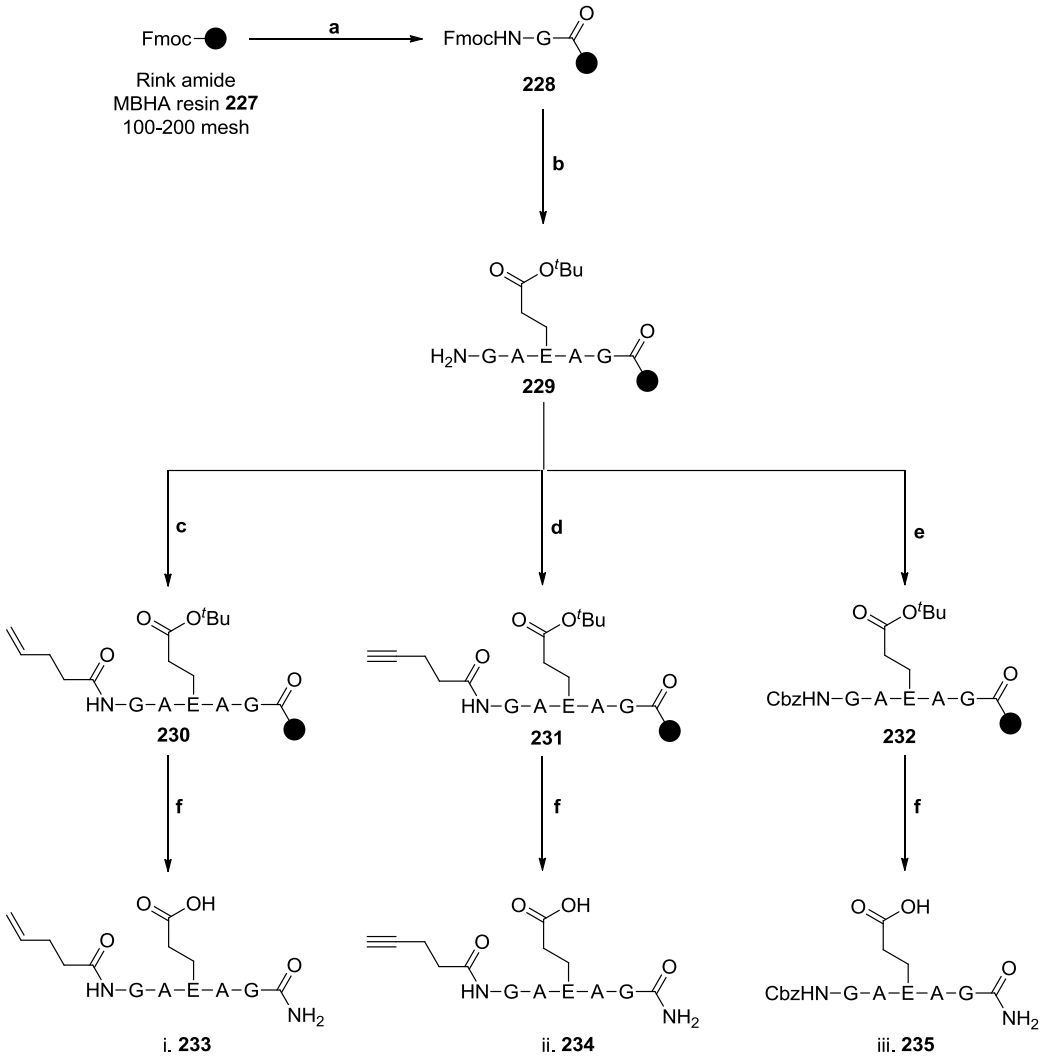


Scheme 48: Photochemical reaction of tetrazole **175o** with phenyl acetic acid **223** and phenyl acetamide **225**. Reagents and conditions: (a) Tetrazole **175o** (1 equiv.), UV light (270-330 nm), TFE, rt, 5 h, 77%; (b) Tetrazole **175o** (1 equiv.), UV light (270-330 nm), TFE, rt, 1 h.

The analogous primary amide phenyl acetamide **225** was reacted under identical conditions and gave a complex mixture of products, as observed by LC-MS, with only the 2,4,5-triaryl-1,2,3-triazole **176o** byproduct being dominant. When the tetrazole was irradiated in the presence of 1 equivalent of both acid **223** and amide **225** the exclusive formation of hydrazide **224** was observed with 97% LC-MS conversion (unisolated).

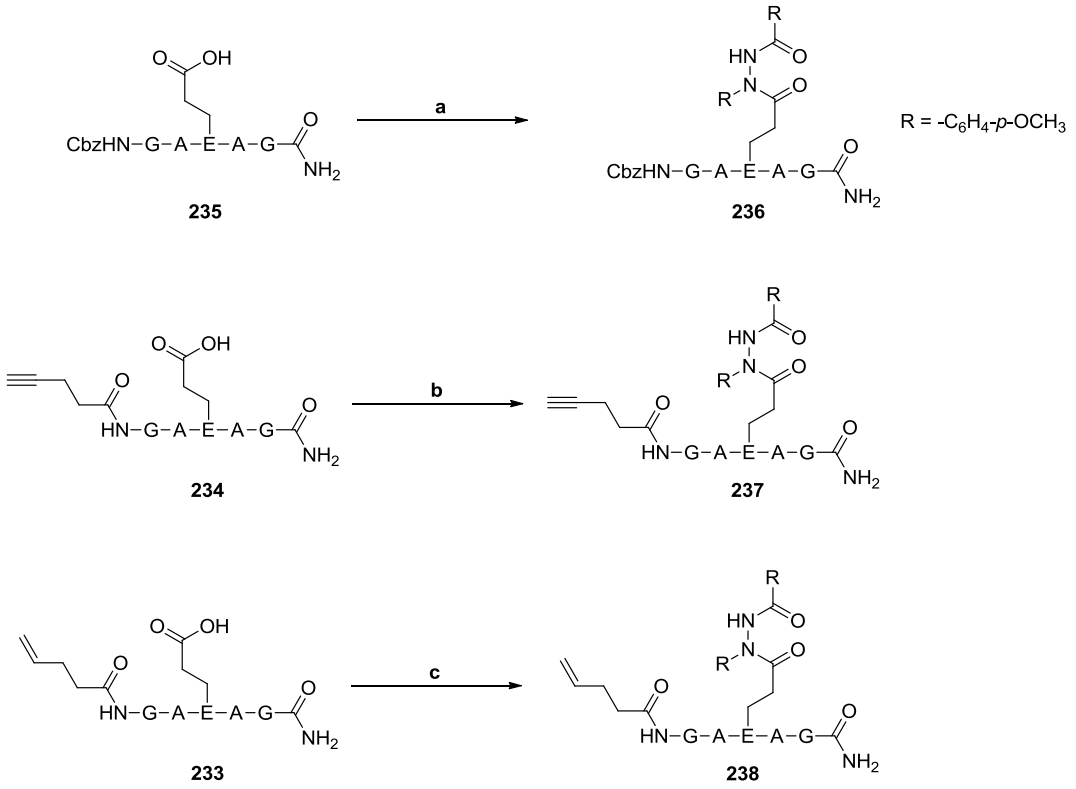
The results obtained from these control experiments clearly indicated that the nitrile imine intermediate was largely unreactive to primary amides and hydrazide formation was selective in the presence of these functional groups. Therefore, the use of MBHA resin was suitable for the solid phase peptide synthesis. Furthermore, the result indicated that amino acids such as asparagine and glutamine would likely be stable to labelling of peptides with photochemically generated nitrile imines. Unfortunately, the formation of 1,2,4-triazoles from primary amides was not observed and no further work was conducted on this aspect of the chemistry.

Peptides **233**-**235** were chosen to test the chemoselectivity of this reaction in the presence of functionality commonly used for peptide labelling (Scheme 49). These were designed to provide the required carboxylic acid in combination with established bioorthogonal tags, such as an alkene and alkyne, and the Cbz protecting group as carbamates are commonly used as protecting groups in peptide synthesis. Unreactive alanine and glycine residues were used for all other amino acids. Peptides **233**-**235** were synthesised by SPPS under standard conditions using HATU mediated amide formation, purified by reverse phase MDAP chromatography, and isolated by lyophilisation.



Scheme 49: Synthesis of a series of peptides containing functionality appropriate for orthogonal labelling. Reagents and conditions: (a) i. 20% piperidine in DMF (v/v), rt, 3 x 5 min; ii. Fmoc-Gly-OH/HATU/DIPEA (5/5/10 equiv.), DMF, rt, 4 h; (b) i. 20% piperidine in DMF (v/v), rt, 5 min; ii. Fmoc-Ala-OH/HATU (5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; iii. 20% piperidine in DMF (v/v), rt, 5 min; iv. Fmoc-Glu(^{t}Bu)-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; v. 20% piperidine in DMF (v/v), rt, 5 min; vi. Fmoc-Ala-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; vii. 20% piperidine in DMF (v/v), rt, 5 min; viii. Fmoc-Gly-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; ix. 20% piperidine in DMF (v/v), rt, 5 min; (c) 4-pentenoic acid/HATU/DIPEA (1/1/1 equiv.), DMF, rt, 4 h; (d) 4-pentynoic acid/HATU/DIPEA (5/5/10 equiv.), DMF, rt, 2 x 1 h; (e) Cbz-Osu (3 equiv.), DIPEA (3 equiv.), DMF, rt, 4 h; (f) TFA/H_2O/TIPS (95:2.5:2.5 v/v), rt, 4 h; i. **233** 8%; ii. **234** 82%; iii. **235** 29%, 78% LC-MS purity.

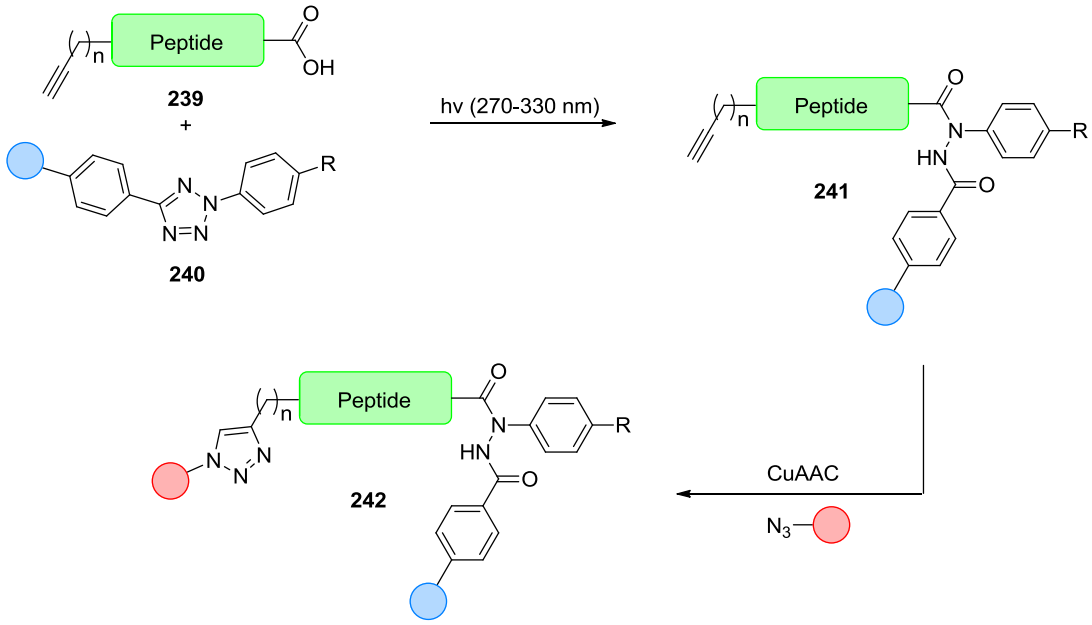
With the requisite peptides in hand the next step was to investigate the formation of hydrazide substituted peptides. Peptides **233**-**235** were irradiated in the presence of tetrazole **175o**, selected for its superior reactivity (Scheme 50).



Scheme 50: Synthesis of hydrazide labelled peptides **236**-**238** under photochemical conditions. Reagents and conditions: (a) Tetrazole **175o** (1 equiv.), UV light (270-330 nm), TFE, rt, 3.5 h, 45%; (b) Tetrazole **175o** (1 equiv.), UV light (270-330 nm), TFE, rt, 1 h, 64%; (c) Tetrazole **175o** (2 x 1 equiv.), UV light (270-330 nm), TFE, rt, 2 x 1 h, 84%.

Peptide **235** was irradiated with tetrazole **175o** to test the chemoselectivity of the hydrazide formation in the presence of primary amide and carbamate moieties. As expected, the major product was peptide **236** with the only byproduct being minor formation of triazole **176o**. This result indicated that under these conditions the formation of hydrazide labelled peptides was possible. The peptide itself was stable to UV light and the functional groups did not undergo any side reactions with the nitrile imine intermediate.

The photochemical generation of hydrazide labelled peptide **237** from peptide **234** was next examined. The presence of the alkyne moiety introduced greater complexity due to a possible [3+2] cycloaddition with the nitrile imine to give a pyrazole byproduct, in addition to potential labelling on both the acid and alkyne. Gratifyingly, conversion of the starting material to the hydrazide labelled peptide was observed selectively with no byproducts observed by LC-MS. The selectivity of this reaction could allow for sequential addition of labels to a peptide of this nature by applying chemoselective hydrazide formation on the carboxylic acid followed by CuAAC labelling on the alkyne (Scheme 51).



Scheme 51: Potential use of alkyne and carboxylic acid tagged peptides for sequential labelling reactions.

Attempts to demonstrate the chemoselectivity of the reaction by locating the alkyne peak in IR spectra of both the starting material and product were unsuccessful due to the weakness of this absorbance. However, analysis using HSQC 2D NMR (see Appendix, Figure A7) clearly showed correlations between the protons and carbons for the alkyne in both the starting material and the product (Figure 62).

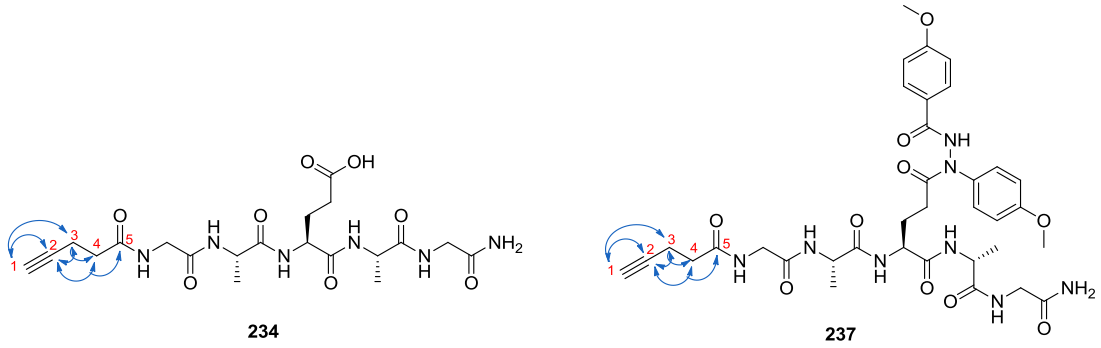


Figure 62: Structures of alkyne labelled peptide starting material **234** and hydrazide labelled peptide **237** after labelling with tetrazole **175o**. Interactions observed by 2D NMR are indicated in blue.

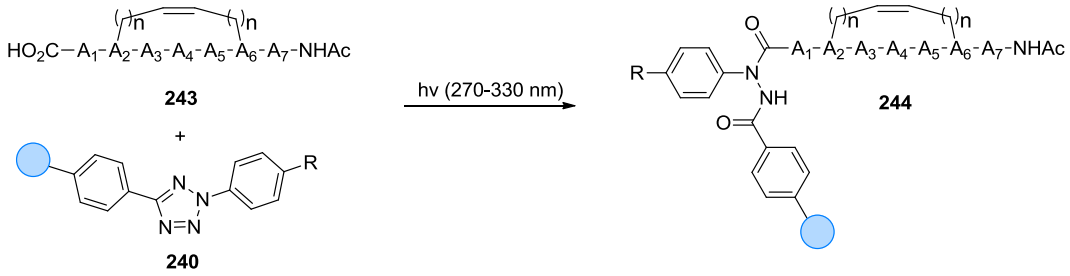
Despite being exposed to different chemical environments protons *H*3 and *H*4 formed a multiplet at approximately δ_H = 2.36 ppm, as is evident from HSQC interactions observed at δ_H = 2.36 ppm to carbon atoms *C*3 and *C*4 at δ_C = 14.0 and 33.9 ppm, respectively. Proton *H*1 was identified from the HSQC interaction to the characteristic chemical shift of *C*1 at δ_H = 2.73 and δ_C = 71.2 ppm, respectively. The peaks can be seen to be identical in both the alkyne labelled peptide starting material **234** and the hydrazide labelled product **237**.

Furthermore, analysis by HMBC 2D NMR (see Appendix, Figure A8) shows further interactions between *H*3/*C*1 and *H*1/*C*3); *H*3/*C*4 and *H*4/*C*3; *H*3 and *C*2; *H*1 and *C*2; *H*4 and *C*2 and finally, *H*4 and *C*5. Again, identical interactions can be observed in both the starting material and product, confirming the chemoselectivity of the hydrazide labelling reactions in the presence of alkyne groups.

The last model reaction was conducted with photochemical generation of hydrazide labelled peptide **238** from peptide **233**. The alkene tag again introduced a potential [3+2] cycloaddition reaction to give a pyrazoline byproduct. Additionally, a potential oxidation of this pyrazoline byproduct could potentially occur to give the corresponding pyrazole. Pleasingly, formation of the desired hydrazide substituted peptide **238** was successful with a yield of 84% and no byproducts observed.

The characteristic alkene protons allowed the chemoselectivity of the reaction to be confirmed by NMR analysis. Peaks representing the carboxylic acid (δ_H = 12.0 ppm) and the alkene protons [5.81 ppm (ddd, *J* = 17.1, 10.5, 1.5 Hz, 1 H), 5.03 ppm (dd, *J* = 17.1, 1.5 Hz, 1 H), 4.95 ppm (dd, *J* = 10.0, 1.5 Hz, 1 H)] were clearly observed during analysis of peptide **233**. While the alkene peaks remained consistent during analysis of the hydrazide labelled peptide **238** the carboxylic acid peak was no longer present, as would be expected. While some double addition was observed by LC-MS, the amount was minor even with two equivalents of tetrazole.

These data indicated that sequential labelling reactions of a carboxylic acid followed by an alkene tag could possibly be conducted on peptides. Additionally, the labelling of carboxylic acids in alkenyl stapled peptides with fluorophore labelled 2,5-diaryl tetrazoles could potentially be conducted with good chemoselectivity (Scheme 52).



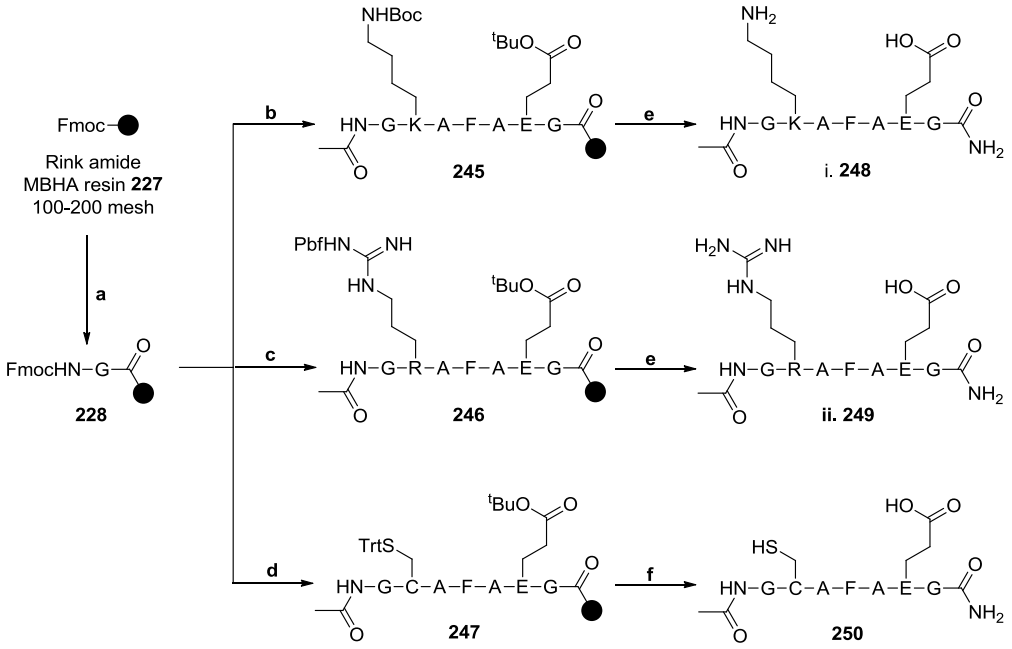
Scheme 52: Chemoselective hydrazide labelling of alkenyl stapled peptides.

These initial results clearly indicated that the use of 2,5-diaryl tetrazoles for the photochemical reaction with peptides containing free carboxylic acid moieties represents a potentially useful tool for peptide labelling. The results suggest that peptides containing alkyl, alkenyl, alkynyl, carbamate and amide (1° or 2°) functionality should all be successfully and efficiently labelled.

With this minimal functionality confirmed to be stable to the reaction conditions further investigation was conducted with peptides containing additional functional groups. As noted during the synthesis of 2,4,5-triaryl-1,2,3-triazoles, the presence of nucleophiles or acid in reaction mixtures can cause quenching of the nitrile imine intermediate; however, the effect of basic functionality remained untested.

Therefore, it was decided to prepare the analogous peptides containing either cysteine, arginine or lysine.

Peptides **248**-**250** were designed to incorporate the more complex functionality found in amino acids, including nucleophilic and basic moieties, and were synthesised using the same procedures as the previous tool peptides (Scheme 53).



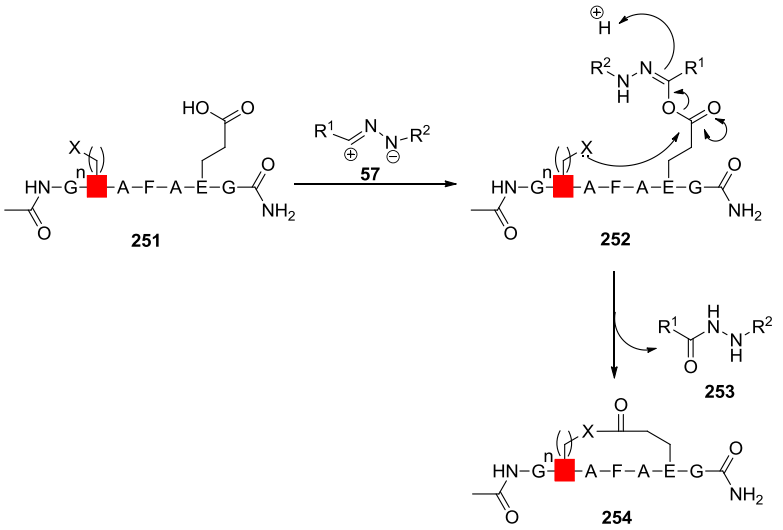
Scheme 53: Synthesis of a series of peptides containing functionality appropriate for orthogonal labelling. Reagents and conditions: (a) i. 20% piperidine in DMF (v/v), rt, 3 x 5 min; ii. Fmoc-Gly-OH/HATU/DIPEA (5/5/10 equiv.), DMF, rt, 4 h; (b) i. 20% piperidine in DMF (v/v), rt, 5 min; ii. Fmoc-Ala-OH/HATU (5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; iii. 20% piperidine in DMF (v/v), rt, 5 min; iv. Fmoc-Glu(^{t}Bu)-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; v. 20% piperidine in DMF (v/v), rt, 5 min; vi. Fmoc-Ala-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; vii. 20% piperidine in DMF (v/v), rt, 5 min; viii. Fmoc-Phe-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; ix. 20% piperidine in DMF (v/v), rt, 5 min; x. Fmoc-Ala-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; xi. 20% piperidine in DMF (v/v), rt, 5 min; xii. Fmoc-Lys(Boc)-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; xiii. 20% piperidine in DMF (v/v), rt, 5 min; xiv. Fmoc-Gly-OH/HATU 5/5 equiv.), 17% NMM in DMF (v/v), rt, 20 min; xv. 20% piperidine in DMF, rt, 5 min; xvi. 0.5M Ac_2O in DMF (v/v), rt, 5 min; (c) As above with Fmoc-Arg(Pbf)-OH in place of Fmoc-Lys(Boc)-OH; (d) As above with Fmoc-Cys(Trt)-OH in place of Fmoc-Lys(Boc)-OH; (e) TFA/H_2O/TIPS (95:2.5:2.5 v/v), rt, 4 h; i. **248** 80%; ii. **249** 76%; (f) TFA/H_2O/TIPS/DODT (92.5:2.5:2.5/2.5 v/v), rt, 4 h, **250** 65% (total for two diastereoisomers).

The only difference was for peptide **250** where 2.5% DODT was included to prevent oxidative dimerisation of the thiol side chains present in cysteine. Each peptide was purified by reverse phase high pH chromatography to give the ammonia salt of the carboxylic acid. This was to ensure the free amine or guanidine residue was present during the photochemical labelling so that the reaction could be monitored for nucleophilic quenching of the nitrile imine intermediate.

Synthesis of peptides **248** and **249** proceeded smoothly; however, the synthesis of cysteine containing peptide **250** showed two peaks with identical masses for the product. This is presumably a result of epimierisation of the cysteine stereocentre. Despite this unforeseen problem, the two components were successfully separated and a single diastereoisomer used for the hydrazide forming reaction.

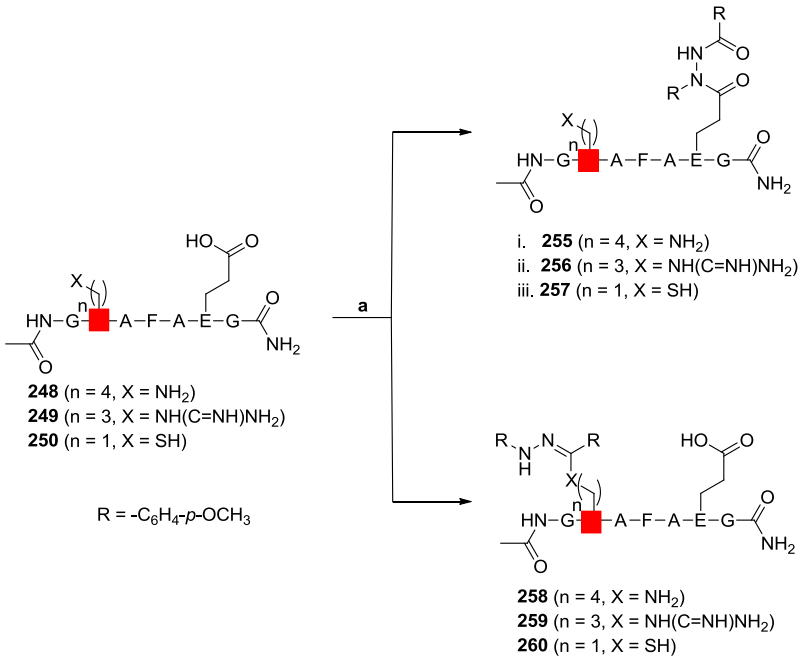
Glutamic acid was placed in the *i+4* position, a common position for peptide stapling.[153] It was hypothesised that a nucleophilic residue in the *i* position of peptide **251** could react with the ester intermediate **252** present during the hydrazide formation, resulting in quenching of the desired reaction and formation of stapled peptides **254** (Scheme 54). While this side reaction is plausible, it is unlikely to occur as the chains are separated by multiple residues and, therefore, is likely to be slow compared to the intramolecular 1,5-rearrangement required for hydrazide formation.

In a similar experiment, Lin *et. al.* have demonstrated the synthesis of stapled helical peptides with fluorescent properties. An alkene and an *N*-aryl tetrazole were incorporated into the amino acid side chains at the *i* and *i*+4 positions which were subsequently stapled using the fluorogenic click reaction.[154]



Scheme 54: Plausible mechanism for the formation of *i+4* stapled peptides during photochemical labelling of glutamic acid with 2,5-diaryl tetrazoles when X is a nucleophile.

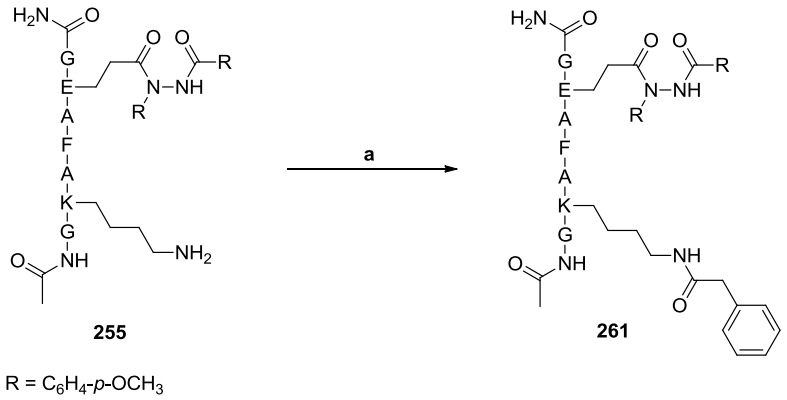
The photochemical addition of the nitrile imine intermediate to these new tool peptides was first conducted with peptide **248**, bearing a primary amine side chain, to give hydrazide labelled peptide **255** in 38% yield (Scheme 55).



Scheme 55: Chemoselective synthesis of hydrazide labelled peptides from peptides **248**-**250** under photochemical conditions. Reagents and conditions: (a) i. Tetrazole **175o** (1 equiv.), UV light (270-330 nm), TFE, rt, 1 h, 38%; ii. Tetrazole **175o** (2.2 equiv.), UV light (270-330 nm), TFE, rt, 1.5 h, isolated with poor purity; iii. Tetrazole **175o** (1 equiv.), UV light (270-330 nm), TFE, rt, 1 h, (9% after oxidation, *vide infra*).

While significant byproduct formation was observed by LC-MS, the mass ion of these byproducts was different to the hydrazide labelled product, indicating it was unlikely to be a result of nucleophilic quenching by the amine to give peptide **258**. In addition, none of the byproducts matched the predicted mass for the formation of a peptide stapled with an amide bond. Due to the small scale of the reaction the byproducts could not be isolated.

Chemoselective formation of **255** was confirmed by conducting an amide coupling reaction with phenyl acetic acid to give peptide **261** (Scheme 56). The small scale formation of the amide was successful upon isolation of peptide **261** in 25% yield after purification by MDAP chromatography. The structure was confirmed by LC-MS and HRMS.



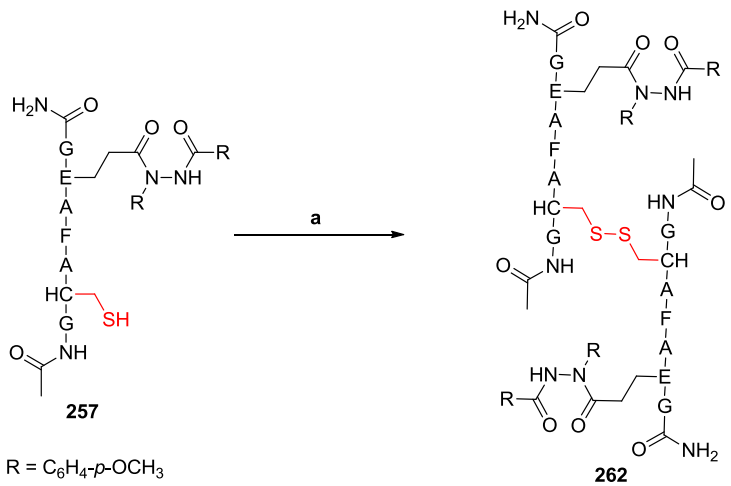
Scheme 56: Conducting an amide coupling to prove the chemoselectivity of the hydrazide formation from peptide **248**. Reagents and conditions: (a) Phenyl acetic acid **223** (1.1 equiv.), HATU (1.1 equiv.), DIPEA (3 equiv.), DMF, rt, 2.5 h, 25%.

Reaction of peptide **249** with tetrazole **175o** gave a product with a mass corresponding to predicted product **256**; however, the yield of the reaction was 9% with only 50% purity by LC-MS. The reaction mixture itself showed a complex spectrum containing a significant number of byproducts which could not feasibly be isolated. Due to the poor purity of the material the exact structure could not be confirmed to be that of peptide **256**. This result indicated that highly basic groups, such as those present in arginine, could cause significant problems for the labelling

of molecules bearing carboxylic acid moieties; although further experimentation would be required to confirm this.

Finally, the labelling reaction was conducted with a single diastereoisomer of cysteine containing peptide **250**. The reaction was cleaner than its previous examples, giving few byproducts. Two significant products were observed in a 2.2:1.0 ratio by LC-MS, both of which had the correct mass for the desired product. It was predicted that this was likely due to formation of desired product **257** as the major component with the minor component representing predicted byproduct **260**.

During isolation of the major component of the reaction mixture, conversion to a second peak with no clearly identifiable mass ion was observed. It was hypothesised that this may have been due to oxidation of the thiol to give dimerised peptide **262** (Scheme 57).



Scheme 57: Oxidative disulfide bond formation to prove the chemoselectivity of the hydrazide formation from peptide **250**. Reagents and conditions: (a) NaI (0.5 equiv.), 28% $H_2O_{2(aq)}$ (1.9 equiv.), EtOAc, rt, 30 min, quantitative conversion by LC-MS, unisolated.

The oxidation of peptide **257** to dimer **262** was confirmed to be responsible for the mixture of peaks obtained during purification by conducting an intentional oxidation on an LC-MS scale reaction. Analysis of a small amount of material by LC-MS, prior to, and after the oxidation with sodium iodide and hydrogen peroxide

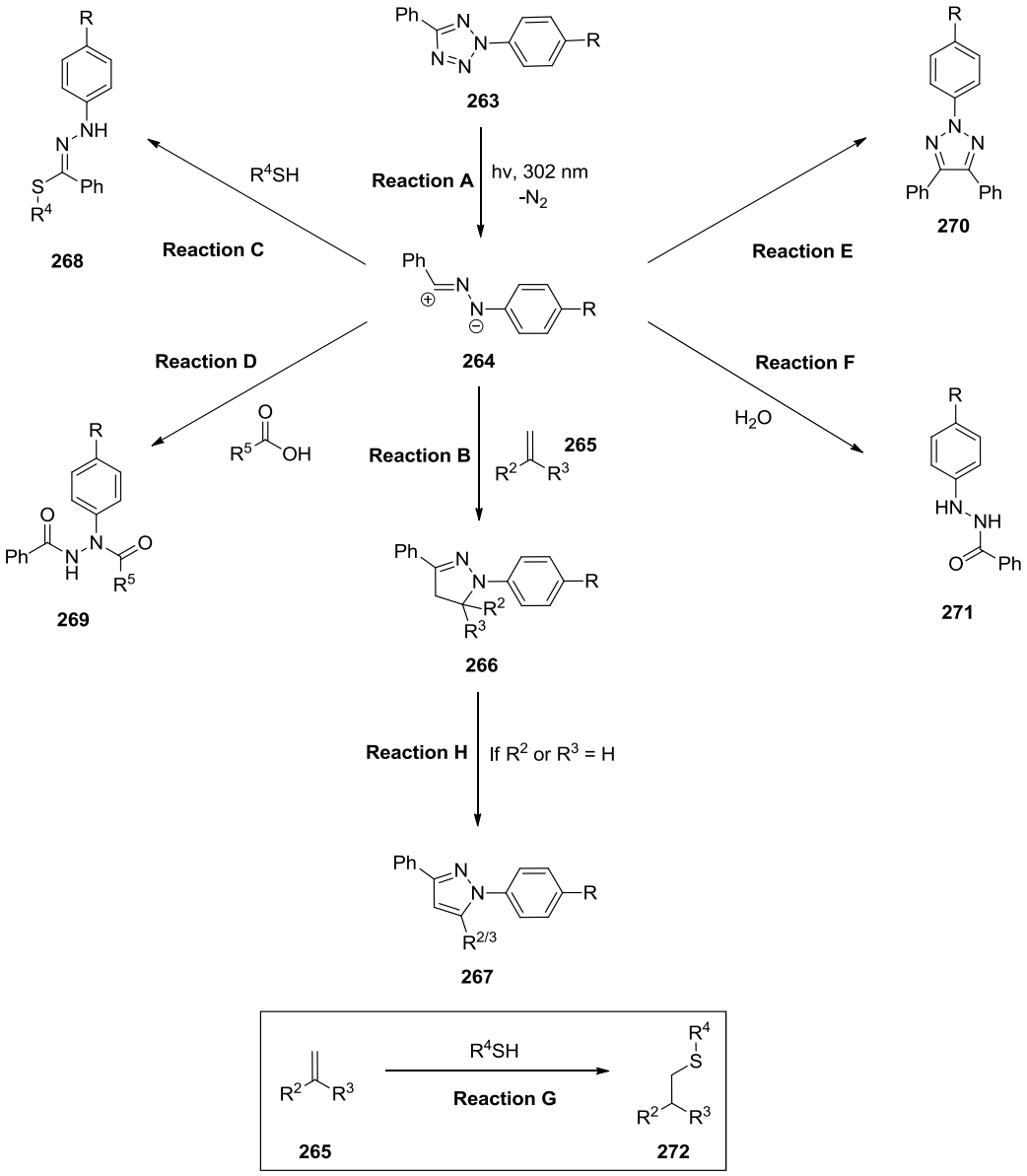
revealed peak to peak conversion of the monomer to the byproduct peak. This both provided evidence for the chemoselectivity of the hydrazide formation and explained the poor yield of the reaction despite good conversions; some product may have been lost during MDAP purification due to the formation of the dimer byproduct.

With the hydrazide formation conducted on a number of suitable tool molecules of varying complexity, a good understanding of the practicality of the process for labelling reactions has been obtained. The labelling reaction can be conducted in excellent yield and chemoselectivity for simple systems; most noticeably those peptides containing alkene and alkyne tags could be easily labelled, which would allow orthogonal labelling reactions to be conducted. However, addition of more reactive functionality such as nucleophilic or basic groups partially reduces the utility of the reaction, resulting in some byproduct formation.

As a result, the reaction may be unsuitable for intracellular labelling of peptides where specificity would be poor. However, throughout this study as a whole, the selectivity and yield of hydrazide formation in the presence of functionalities such as alkenes, alkynes, protected amines/carboxylic acids, and amides was shown to be excellent. While possibly not applicable to intracellular labelling of peptides, further research is warranted for potential non-cellular applications, such as labelling of surfaces, orthogonal peptide stapling, and solid phase purification or extraction.

1.7. Conclusions

The aim of this study was to conduct low concentration fluorogenic click reactions between 2,5-diaryl tetrazoles and alkene labelled amino acids inside cells; the purpose of which was to determine the permeability and intracellular location of peptides. During the synthesis of pyrazoline adducts for intracellular visualisation of peptides, a number of possible competing reactions were identified (Scheme 58).



Scheme 58: Summary of the reactions that can potentially occur with 2,5-diaryl tetrazoles and alkenes under photochemical conditions.

Photochemical activation of the tetrazole to give a nitrile imine intermediate (Reaction A) and the subsequent cycloaddition with alkenes to give pyrazoline adducts (Reaction B) was demonstrated to be successful at appropriate concentrations in MeCN. Oxidation of some pyrazoline adducts to corresponding pyrazoles (Reaction H) was shown to occur, but on a timescale which did not affect the immediate observation of fluorescence. However, when repeated in PBS the reaction was inhibited, a result also observed in the presence of the glutathione mimic *N*-acetyl cysteine. After a detailed investigation of the possible side reactions (Reactions C-G) and their competitive rates, it was determined that the formation of hydrazides from Reactions D and F were the most likely cause of the unsuccessful cell experiment. At no point was Michael addition of *N*-acetyl cysteine to the alkene tags (Reaction G) observed.

While briefly considered as the potential cause of the unsuccessful experiment, the formation of a 2,4,5-triaryl triazole (Reaction E) was shown to not be the major reaction inhibitor due to its slow rate, relative to the other side reactions. Even so, after an extensive study this reaction pathway proved to be an efficient means of synthesising symmetrical electron rich 2,4,5-triaryl-1,2,3-triazoles. A two step synthesis of 5-methoxy-3-morpholine-1*H*-indazole was also demonstrated, offering a short and efficient route to this important scaffold.

Hydrazide formation via reaction pathway D was also investigated for its potential use as a new bioorthogonal labelling reaction. The hydrazide labelling of peptides bearing primary amides, protected amines, alkenes and alkynes were shown to be selective and high yielding. In general, the chemoselectivity of the reaction was high, although, some byproduct formation in the presence of bases and nucleophiles may limit its utility as a bioorthogonal labelling reaction. However, this methodology still represents a potential new technique for non-cellular orthogonal labelling of simple peptides and also has potential in other fields such as surface chemistry, purification and labelling of small molecules. Further research into each of these areas will be conducted.

Overall, this work has concluded that, despite the literature precedent for the labelling of alkene tagged proteins in cells, the photochemical labelling of alkene tagged amino acids and peptides would require significant further research if it were to be a viable method for the intracellular visualisation of peptides. It is not clear as to why the reaction is more efficient for use with intracellular proteins but may be a result of local tetrazole accumulation and the hydrophobic effect. Precise repetition of the literature work would be required for further analysis; however, due to unavailable resources this was not possible. The evidence suggests that high concentrations of tetrazole were used to account for the array of side reactions that, fortuitously, produce byproducts with only low levels of fluorescence and do not prevent visualisation of the desired protein-pyrazoline adducts formed.

Chapter 2

Developing a solid phase synthesis of cyclotheonamide B and analogues as inhibitors of Trypsin-Like Serine Proteases

2.1. Introduction

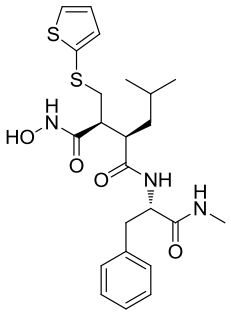
2.1.1. Protease class and structure

As discussed in Chapter 1, peptides represent an under exploited class of molecules that have the potential to target the 80% of proteins that are considered undruggable by small molecules and biologics.[1] Continually improving synthetic methodology and improvements in the drug-like properties of peptides has resulted in an increase in their popularity as potential drug candidates. The ability to utilise solid phase peptide synthesis (SPPS) for rapid synthesis of peptide analogues has made SAR analysis for disease targets significantly easier, while significant research has been conducted to gain an understanding of the mechanisms by which peptides are able to enter cells.[155] Peptides also have high specificity for their targets due to their typically large binding surfaces which results in lower promiscuity compared to small molecules.[156]

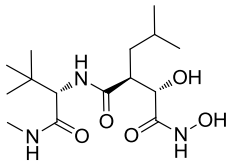
Despite the increasing prevalence of peptide based drug molecules, the number of marketed peptide drugs is still relatively few, even for protein targets with well established structures and known mechanisms of action. Proteases, for example, are a well known type of enzyme that have been studied for over one hundred years and represent a target family for which few peptide drugs are available.[157]

A protease, also known as a peptidase, is an enzyme that breaks down amide bonds, a process known as proteolysis. Proteases are present in all biological organisms and a number of different classes have evolved, each of which have similarities between them but ultimately perform proteolysis via different catalytic mechanisms. Some of the main sub-families include serine proteases, cysteine proteases, aspartyl proteases and metalloproteases.[158] Each protease is separated into one of these sub-families based on their mechanism of action, with each family's mechanism differing from the others due to evolutionary divergence. The development of drugs for the manipulation of these protease families is a continually evolving field of study with significant effort traditionally placed on discovery of small molecule inhibitors, and more recently, peptide inhibitors.[159]

Some peptidic protease inhibitors have even reached clinical trials. In an early example, clinical trials of hydroxamic acid peptide derivatives Batimastat **273** and Marimastat **274** were conducted in the 1990s for inhibition of matrix metalloproteases (Figure 63). These hydroxamic acid class of inhibitors were designed to bind to the catalytic Zn^{2+} ion in the protease, thus deactivating it. However, these were broad spectrum metalloprotease inhibitors and bound to zinc with high affinity with very little selectivity for a specific protease.[160] The phase III trials were ultimately unsuccessful due to the observation of peritonitis for Batimastat, while Marimastat failed to show progression free survival in cancer patients.[161,162] Despite this setback, peptides are still under investigation for the inhibition of metalloproteases.



273
Batimastat



274
Marimastat

Figure 63: Structure of hydroxamic acid derived metalloprotease inhibitors Batimastat **273** and Marimastat **274**.

In the 1970's the natural epoxysuccinyl peptide E64 **275** was discovered (Figure 64). The epoxide warhead allows this molecule to act as an irreversible covalent inhibitor through conjugation of a cysteine reside in the active site. E64 was subsequently found to be an inhibitor of a range of cysteine proteases, leading to significant research into analogues with specific binding to each of the different proteases.[163] CA074 **276** and E64-c **277** are two such derivatives. The former has been shown to act as a Cathepsin B specific inhibitor[164] while the latter was developed as a cell permeable analogue of E64. E64-c reached phase III clinical trials for the treatment of muscular dystrophy, though it was ultimately unsuccessful as it

was found to covalently modify non-Cathepsin proteins.[165] A range of alternative warheads has been investigated for use in cysteine protease inhibitors including functional groups such as Michael acceptors, aziridines, aldehydes, ketones and cyclopropenones.[166]

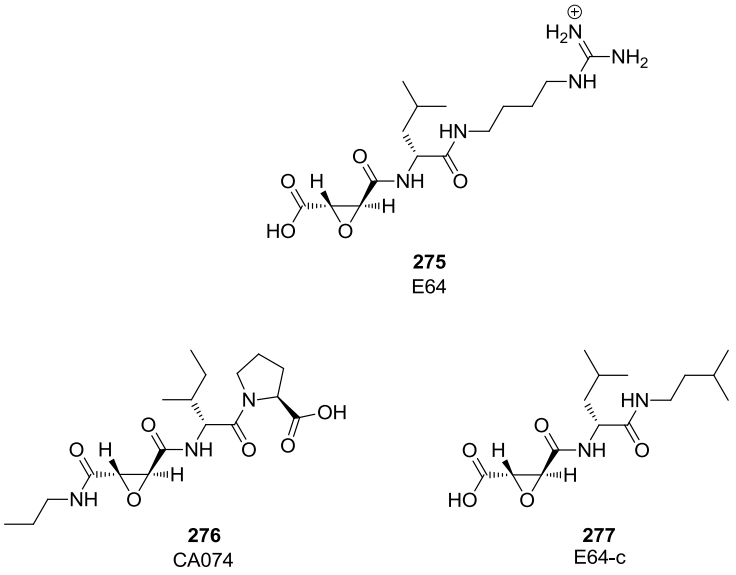


Figure 64: Structures of epoxysuccinyl peptide inhibitor E64 **275** and its derivatives E64-c **276** and CA-074 **277**.

On a more successful note, nine aspartyl protease inhibitors (saquinavir, nelfinavir **278**, lopinavir, amprenavir **279**, darunavir, tipranavir **280**, indinavir **282**, atazanavir **283**) have been developed and approved for the inhibition of Human Immunodeficiency Virus (HIV) related proteases (Figure 65). Recent research has shown that the combination of Nelfinavir **278** with Bortezomib **281**, an approved antineoplastic drug, has been successful in augmenting protease inhibition and overcoming bortezomib resistance in myeloma cells. The study has indicated that further clinical trials may be beneficial for the development of anti-resistant cancer therapies.[167]

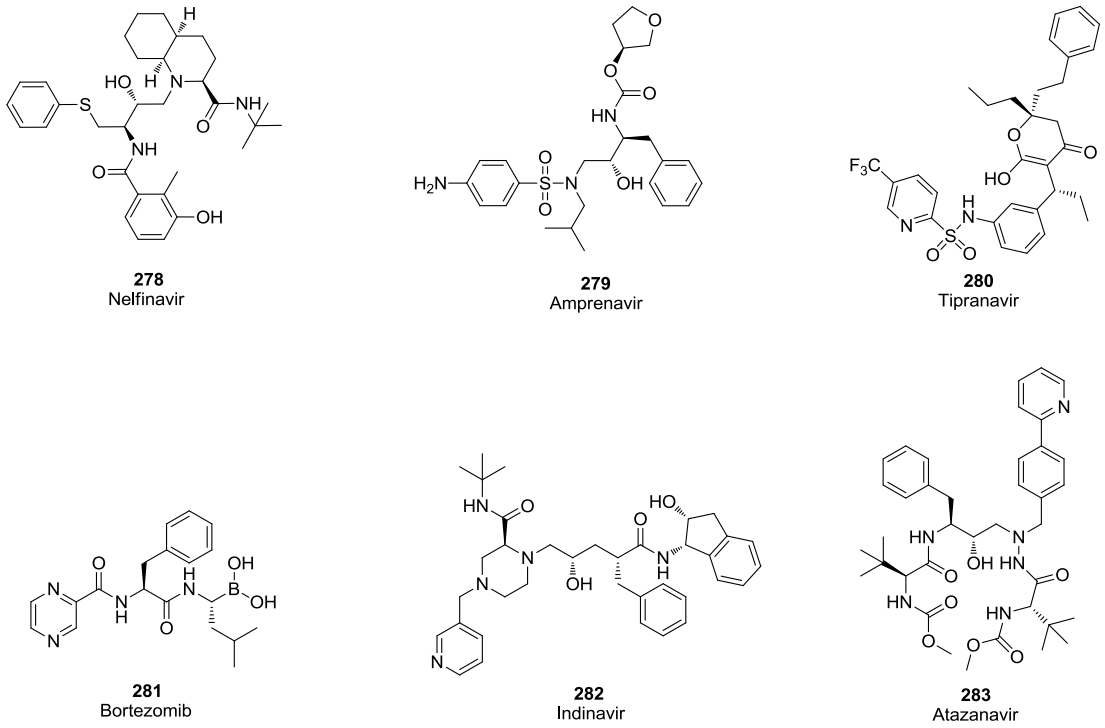
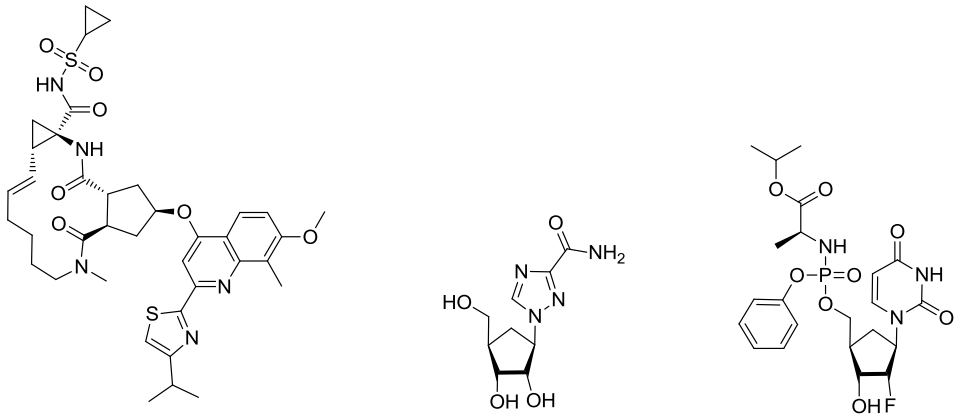


Figure 65: Examples of marketed aspartyl protease inhibitors **278**-**280** and **282**-**283**, in addition to the structure of antineoplastic drug Bortezomib **281**.

Over one third of proteolytic enzymes are serine proteases making this an important target class in medicinal chemistry.[168] Aralast™, Prolastin™, and Zemaira™, a trio of serine protease inhibitors, have been marketed for the treatment of emphysema, a symptom associated with Chronic Obstructive Pulmonary Disease (COPD).[169] Patients with this disease have a deficiency of α-antitrypsin, a natural serine protease inhibitor, often as a result of smoking. The three treatments are intravenous α-antitrypsin doses, all of which are sourced from human plasma, differing only in their purification methods.

The synthetic molecule Simeprevir **284** is an FDA approved inhibitor of the Non-structural protein 3 (NS3) serine protease, an enzyme associated with viral replication (Figure 66).[170] Simeprivir can therefore be used as a treatment of Hepatitis C and is often delivered in combination with nucleoside inhibitor Ribavirin **285** and with pegylated interferon alpha. The once daily dose of Simeprivir and its significantly shortened treatment time make this an effective second generation

medicine for the treatment of Hepatitis C. Released at a similar time was Sofosbuvir **286**, a nucleotide analogue and RNA polymerase inhibitor.[171] This molecule was designed to allow for treatment of Hepatitis C in the absence of interferon alpha, which is known to have severe side effects.



284
Simeprivir

285
Ribavirin

286
Sofosbuvir

Figure 66: Structures of serine protease inhibitor Simeprivir **284**, nucleoside inhibitor Ribavirin **285**, and RNA polymerase inhibitor Sofosbuvir **286**.

The structure of serine proteases and their mechanism of action are well known. The individual residues of a peptide and the pockets of the protease itself are given specific nomenclature relative to the location of the scissile bond (Figure 67). Residues on the *N*-terminal side of the scissile bond are designated as S1, S2, S3 and so forth, while those on the *C*-terminal side of the scissile bond are designated as S1', S2', S3' *etc*. Trypsin-like Serine Proteases (TLSPs) are classified by specific structural features. The active site consists of four main pockets, designated as S1, S2, S3 and S4. Additional pockets, such as the S1' pocket, are present in the active site of some TLSPs.

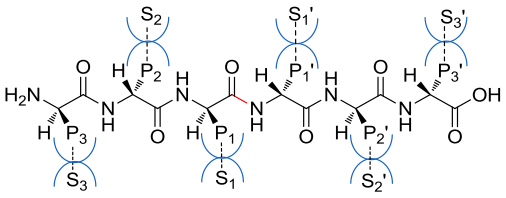


Figure 67: Nomenclature of peptide-protease interactions with protease pockets S1-3/S1'-S3' and peptide residues P1-3/P1'-3' relative to the scissile bond, represented in red.

TLSPs are often classified by the nature of the residue in the S1 pocket which is normally an Asp residue. This negatively charged residue forms strong bonding interactions with positively charged species such as found in protonated arginine or lysine residues of peptides (Figure 68a). Additionally they each have a common catalytic triad which consists of Asp-His-Ser residues (Figure 68b).

a) b)

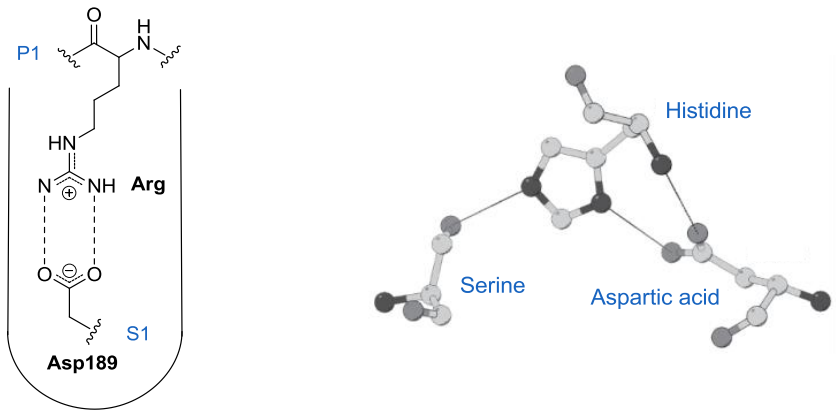
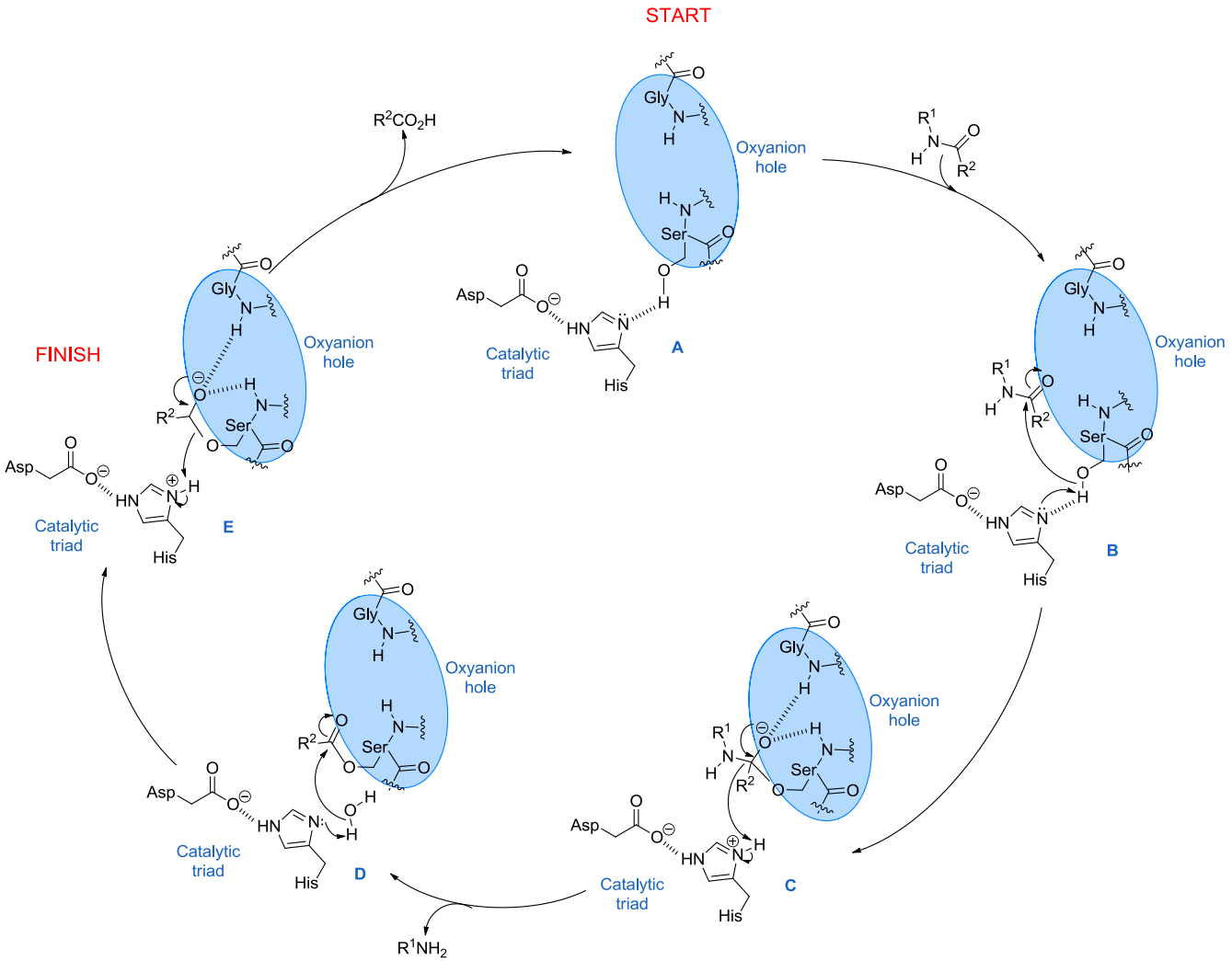


Figure 68: (a) Binding of arginine amino acids to the Asp residue in the P1 pocket of a serine protease and (b) General structure of the Ser-His-Asp catalytic triad of serine proteases.

While the Asp residue binds basic ligands, making it useful for selectivity over other proteases, it is the catalytic triad which is key to the activity of the protease itself; the purpose of which is to break amide bonds (Scheme 59).

The alcohol side chain of the Ser residue initially binds to the carbonyl group, forming an anionic tetrahedral transition state which is stabilised by the oxyanion hole. This stabilisation consists of hydrogen bonding interactions from the Ser N-H bond and a nearby Gly N-H bond present in the peptide backbone (A-C). The carbonyl group is reformed, the amine leaving group expelled, and an ester bond formed to the Ser residue. This is followed by hydrolysis of the ester bond to give the corresponding carboxylic acid, simultaneously restoring the catalytic site to its initial state (C-E). During this process the His residue acts as an acid-base transfer vehicle for the protonation and deprotonation of key intermediates while the Asp residue plays a key part in facilitating these transfers via hydrogen bonding interactions.



Scheme 59: Mechanism of catalytic amide bond cleavage in serine proteases.

2.1.2. Kallikrein 5 (KLK5) as a medicinally relevant serine protease

The kallikrein (KLK) family consists of 15 proteases under the serine protease class and are generally located in the epidermal tissue. Disregulation of the KLK proteases results in under-expression or over-expression of these important proteins, leading to a number of disease states. For example, over-expression of KLK4, 5, 6, and 7 can increase the prevalence of malignant ovarian cancer.[172] In 2013, MDPK67b, an inhibitor of serine protease KLK2, was approved for phase II clinical trials as a potential treatment for prostate cancer.[173]

The regulation of KLK5 is associated with the process of cell shedding (desquamation) and is expressed in a number of tissue types, including the brain, thyroid, and salivary glands, but is particularly prevalent in the skin.[174] KLK5 is, therefore, associated with disease states such as Atopic Dermatitis which can be considered an area of unmet medical need.[175] Accordingly, the development of an inhibitor for the KLK5 protein is desirable from a therapeutic standpoint.

The structure of KLK5 is typical of a serine protease with a conserved Ser195-His57-Asp102 catalytic triad and an oxyanion hole.[176] Furthermore the S1 pocket contains an Asp189 residue, also typical of the serine protease class of receptors. The development of inhibitors will, therefore, focus around the synthesis of peptides bearing arginine or lysine residues to interact with the S1 pocket and a suitable electrophilic warhead to target the catalytic serine residue. This combination should instil strong binding of the inhibitor to the protease.

Further affinity can be gained from the remaining S2, S3 and S1’ pockets. In this case, the S1’ and S2 pockets were thought to be key to obtaining selectivity over other serine proteases due to the differences in the polarity of the residues within the pockets of the various proteases. Tables 9 and 10 show the residues present in the S1’ and S2 pockets of KLK5 and a selection of serine proteases over which selectivity is desired.

The S1’ pocket of the selected proteases consists of key residues 41, 57, and 60. There is significant variation between the proteases with the pockets of KLK1 and

KLK5 being more polar than those of urokinase-type plasminogen activator (U-Pa) and thrombin. The KLK5 pocket almost uniquely contains a charged Lys60 and Tyr41 residue. Therefore, selectivity could be obtained by forming hydrogen bonding interactions to the positively charged lysine or π-interactions to the tyrosine.

S1'	41	57	60
KLK5	Tyr	His	Lys
KLK1	Gln	His	Ser
U-Pa	Val	His	Tyr
Thrombin	Leu	His	Phe
Factor Xa	Tyr	His	-

S2	57	94	99
KLK5	His	Tyr	His
KLK1	His	Phe	Tyr
U-Pa	His	Tyr	His
Thrombin	His	Tyr	Leu
Factor Xa	His	Phe	Tyr

Tables 9 and 10: Residues present in the S1' and S2 pockets of a selection of serine proteases. Green = Non-polar residues, Yellow = Polar (Uncharged) residues, Blue = Polar (Charged) residues.

Similarly, the S2 pocket for KLK5 shows greater polarity than observed for the other proteases, with the exception of U-Pa which has a conserved S2 pocket identical to KLK5, indicating that selectivity over this protease will be difficult to obtain. Having stated this, the selectivity for KLK5 over the other proteases could be achieved by increasing the polarity within the S2 pocket.

The affinity of a peptide for these pockets alone may be enough to develop a non-covalently binding inhibitor. Such non-covalent peptide inhibitors have been developed for use against proteases; for example, Ritonavir and its analogues have been demonstrated to inhibit the chymotrypsin-like human 20S proteasome.[177] However, a more common approach is to develop binding and selectivity for the pockets in addition to the use of an electrophilic warhead. Peptidic protease inhibitors have been developed using a range of electrophilic warheads including, but not limited to aldehydes,[178] α-keto aldehydes,[179] α-ketoamides,[180] fluoromethyl ketones,[181] hydroxamic acids,[157] epoxides,[167] cyclopropanones,[182] vinyl esters,[183] vinyl sulfones,[184] and boronic acids.[185] Each of these functionalities can act as an electrophile for the nucleophilic attack from Ser195 in the KLK5 protease.

2.1.2.1. Identifying a suitable class of protease inhibitor

To date, only a single example of a peptide based inhibitor of KLK5 has been disclosed. Sunflower Trypsin-1 (STFI-1) **287** is a 14 amino acid cyclic peptide and is homologous to the Bowman-Birkman class of natural peptide inhibitors (Figure 69).[186] The structure consists of a Cys3-Cys11 chain containing a Lys5 residue which binds to the S1 pocket of TLSPs, a non-reactive loop consisting of the remaining amino acid residues Thr4-Ile10, and a disulfide bond between the cysteine residues to give a bicylic structure. The synthesis of STFI-1 analogues was achieved by Macmillan *et. al.* using both synthetic and recombinant approaches. The analogues were found to inhibit KLK5 with IC_{50} values in the sub-micromolar range.[187]

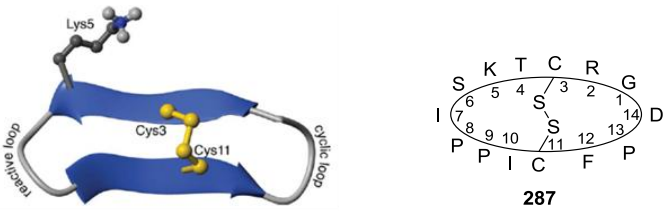
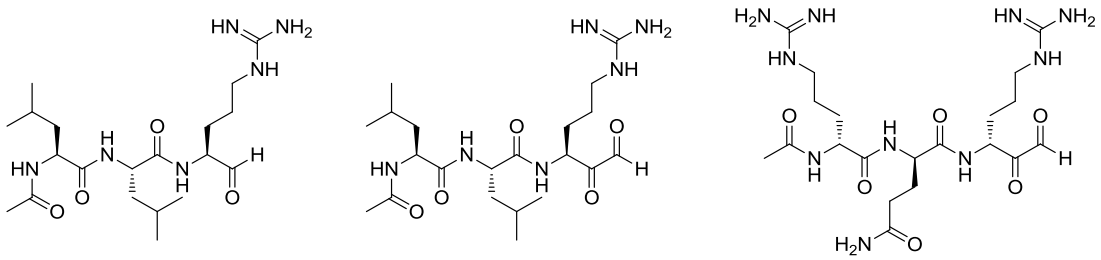


Figure 69: Structure of Sunflower Trypsin Inhibitor (STFI-1) **287**.[186]

This encouraging precedent prompted us to conduct a literature survey to identify peptide based inhibitors of related TLSPs as potential starting points from which to develop KLK5 inhibitors.

Lynas *et. al.* have demonstrated the inhibition of thrombin and trypsin using tripeptides **289** and **290** with α-keto aldehyde warheads based on the structure of Leupeptin **288**, one of the first known inhibitors of TLSPs (Figure 70).[179] Interestingly, while the inhibitory action of Leupeptin is reversible, the use of the α-keto aldehyde warhead resulted in irreversible inhibition of thrombin and trypsin for peptides **289** and **290**, respectively.



288
Leupeptin

289

290

Figure 70: Structures of natural serine protease inhibitor Leupeptin **288** and synthesised inhibitors for thrombin **289** and trypsin **290**.

The crystal structure of KLK5 with leupeptin indicated that the P1 arginine residue resides in the S1 pocket of KLK5, binding to the Asp189 residue as expected, while the aldehyde forms a covalent interaction with Ser195 (Figure 71).[188] The two remaining leucine residues occupy space near the S2 and S3 pockets of KLK5 with the acetamide moiety within the S3 pocket itself.

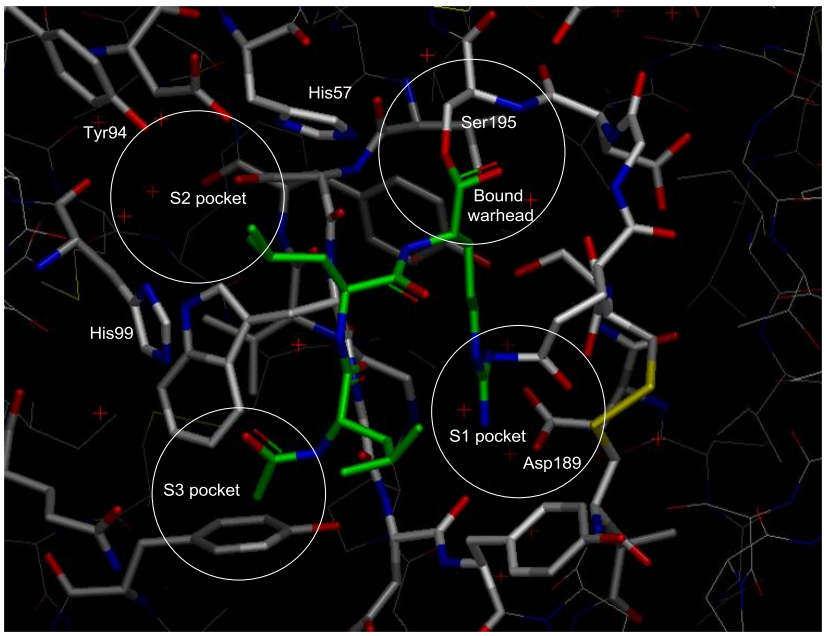
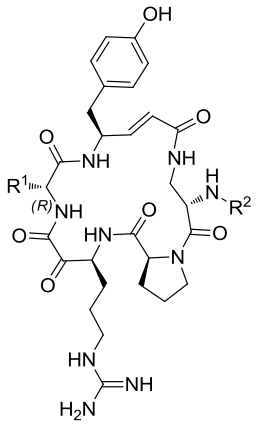


Figure 71: Crystal structure of the natural peptide inhibitor Leupeptin **288** in KLK5.

The cyclotheonamide (Ct) family **291**-**295** is a class of natural cyclic peptide inhibitors with similar structural features to leupeptin (Figure 72). The structures consist of two natural amino acids and three non-proteinogenic amino acids, the latter of which consist of a vinyl tyrosine residue, an α-ketoamide arginine residue,

and a substituted diaminopropionic acid (Dap) residue. The complex functionality of the molecule is contained within a cyclic backbone. The cyclic nature of the molecule was of interest due to the greater stability and improved pharmacokinetic properties of cyclised peptides (see Section 1.1). The natural amino acids include proline and either *D*-phenylalanine or *D*-isoleucine.



291 CtA: R^1 = Bn, R^2 = CHO
292 CtB: R^1 = Bn, R^2 = Ac
293 CtE: R^1 = ^{i}Bu, R^2 = -Ala-OBn
294 CtE_2: R^1 = ^{i}Bu, R^2 = -Ala-O^{i}Pr
295 CtE_3: R^1 = ^{i}Bu, R^2 = -Ala-O^{i}Pr

Figure 72: Representative examples of the cyclotheonamide (Ct) family **291**-**295** of natural products.

Further examination of the literature revealed that the representative cyclothoneamides have been tested for their activity against the TLSPs thrombin and trypsin (Table 11).[189] These data revealed that the Ct family shows significant activity against both targets. Furthermore, analysis of the inhibitory activity of Cyclotheonamide A (CtA) **291** against a number of serine proteases, including an unspecified kallikrein, showed strong inhibitory activity, further confirming the potential for cyclotheonamides to act as inhibitors of KLK5 (Table 12).[190]

Protease	CtA	CtE	CtE_2	CtE_3
Thrombin	23	2.9	13	9.5
Trypsin	16	30	55	52

Table 11: Inhibitory activity of cyclotheonamides **291** and **293**-**295** against Thrombin and Trypsin (IC_{50}/nM).

Enzyme	CtA	Enzyme	CtA
Thrombin	0.18	Streptokinase	0.035
Plasmin	0.37	Factor Xa	130
Kallikrein	0.51	Trypsin	0.023
tPa	6.5	Chymotrypsin	6.9
Urokinase	0.37	Protein Ca	3.1

Table 12: Inhibitory activity of CtA **291** against a selection of serine proteases (Ki/μM).

Using molecular modelling techniques the structure of cyclotheonamide A **291** was overlaid with the structure of leupeptin **288** in the KLK5 crystal structure to determine if the binding mode was similar (Figure 73a).[191] The structures of the two ligands overlay well with the Arg-Leu-Leu chain of leupeptin **288** overlapping with the Arg-Pro-Dap chain of CtA **291**. Additionally, the acetyl and formaldehyde groups are shown to overlap, as are the electrophilic warheads of both ligands. The docked structure of CtA **291** in KLK5 shows the expected interactions between the Arg residue and Asp189 in the S1 pocket and the α-ketoamide with the Ser195 residue (Figure 73b). As expected, the proline and Dap residues appear to be in close proximity to the S2 and S3 pockets, respectively, offering good opportunities for SAR exploration. Interestingly, the *D*-phenylalanine residue is in close proximity to the S1' pocket, suggesting that the use of a cyclic peptide approach may offer access to pockets not possible with linear peptides. Finally, the vinyl tyrosine residue is exposed to the solvent, indicating that simplification of the structure may be possible if this residue were to be replaced with a simpler alternative.

a)

b)

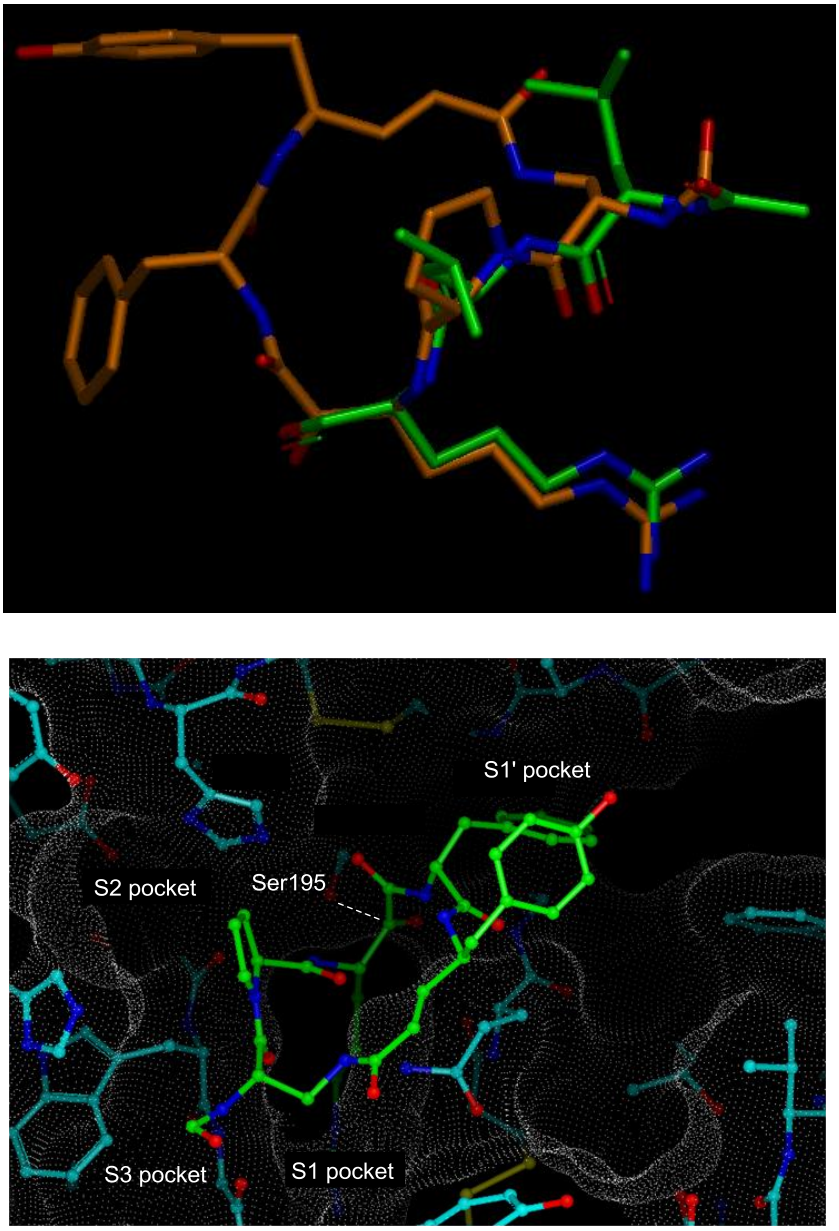


Figure 73: (a) Overlay of the structures of Leupeptin **288** (green) and CtA **291** (orange); (b) Structure of CtA **291** (green) in the KLK5 crystal structure.

2.1.2.2. Project aims

The development of a robust and flexible synthetic route to cyclotheonamide derivatives was investigated with the aim of exploring the SAR of the S1', S2 and S3 pockets, in addition to confirming activity at KLK5 as suggested from the modelling

data. The cyclotheonamide family appeared to be promising as a starting point for the development of inhibitors of the kallikrein family and the crystal structures of the ligand docked to KLK5 using computational modelling indicated that these natural products were worth investigating further.

In order to identify a suitable route for the synthesis of cyclotheonamide analogues a number of factors were taken into consideration. These included:

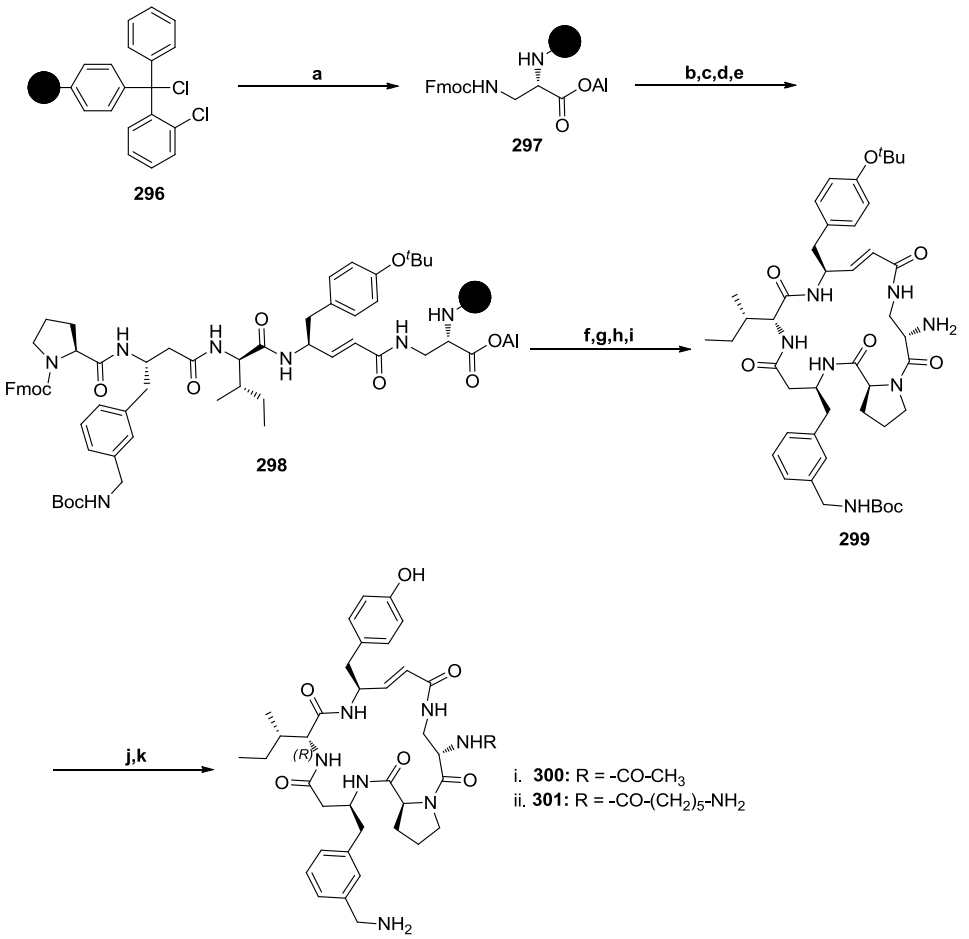
- Short and efficient preparation of the requisite amino acid residues
- Suitable methodology for preparation of the key α-ketoamide moiety
- Protecting groups suitable for orthogonal or global deprotection, as required.
- Flexible approach for easy modification of the P1', P2 and P3 residues
- Suitable position for the cyclisation of the linear pentapeptide
- Short and efficient route to the full cyclotheonamide structure

Additionally, an ideal synthetic route would use a solid phase peptide synthesis (SPPS). Solid phase synthesis offers a number of advantages over solution phase synthesis, such as: generally high yields; an insoluble polymeric support which allows for easy separation of the growing peptide from byproducts and excess reagents; minimal loss of material during physical manipulations of the resin; physical manipulations of reaction mixtures which are amenable to automation; rapid syntheses; simultaneous preparation of multiple peptides; orthogonal cleavage of resins without affecting side chain protecting groups; and global cleavage to give fully deprotected peptides.

With these factors in mind, a literature review of previous cyclotheonamide syntheses was conducted in order to identify the synthetic options already available and determine how these could inform the design of a synthetic route to meet the overall objectives of this work.

2.1.3. Evaluating previous syntheses of the cyclotheonamide family

In two recent publications Schaschke *et. al.* have demonstrated a flexible synthesis of simplified cyclotheonamide derivatives using SPPS in order to facilitate the exploration of the SAR of a human β-tryptase (Scheme 60).[192,193] The key aspects of the synthesis are summarised in Table 13 (Entry 1).



Scheme 60: Solid phase synthesis of cyclotheonamide analogues for the inhibition of Human β-tryptase. Reagents and conditions: (a) i. H-Dap(Fmoc)-OAl x HCl (2 equiv.), DIPEA, CH_2Cl_2; ii. MeOH/CH_2Cl_2, DIPEA; (b) i. piperidine/DMF (1:4 v/v); ii. Fmoc-vTyr(^{t}Bu)-OH/TBTU/HOBt/DIPEA (1:1:1:1, 2 x 1 equiv.), DMF; (c) i. piperidine/DMF (1:4 v/v); ii. Fmoc-D-*allo*-Ile-OH/TBTU/HOBt/DIPEA (1:1:1:1, 4 equiv.), DMF; (d) i. piperidine/DMF (1:4 v/v); ii. Fmoc-β^3hPhe(3-BocHN-CH_2)-OH/TBTU/HOBt/DIPEA (1:1:1:1, 2 x 1 equiv.), DMF; (e) i. piperidine/DMF (1:4 v/v); ii. Fmoc-Pro-OH/TBTU/HOBt/DIPEA (1:1:1:1, 4 equiv.), DMF; (f) $Pd(PPh_3)_4$ (3 x 0.4 equiv.), $PhSiH_3$ (3 x 24 equiv.),CH_2Cl_2; (g) piperidine/DMF (1:4 v/v); (h) PyBOP/HOBt/DIPEA (1:1:2, 3 equiv.), DMF; (i) TFA/TIS/CH_2Cl_2 (1:1:98 v/v/v), over eight steps 96%; (j) BocNH$(CH_2)_5$COOH/TBTU/HOBt/DIPEA (1:1:1:1, 2 equiv.), DMF or Ac_2O/pyridine (1:1, 10 equiv.), CH_2Cl_2; (k) TFA/H_2O (95:5 v/v) i. **300**, 30% (two steps); ii. **301**, 33% (two steps).

This research demonstrated the utility of solid phase methodology for rapid synthesis of analogues with modified P1 residues. Each residue could be prepared individually and coupled to the resin sequentially. This would also allow for easy and rapid modification of the P1', P2 and P3 residues, as required for the current study. The P3 residue would be particularly easy to modify for late stage analogue production due to the late stage deprotection of the Dap amine group.

Each residue could be prepared with *N*-Fmoc protection which could be orthogonally cleaved under mildly basic conditions in the presence of acid labile protecting groups. Additionally, the use of *O*-allyl protection for the *C*-terminus allowed for orthogonal deprotection of the two termini immediately prior to cyclisation between the Pro and Dap residues. Notably, the use of the highly acid labile 2-chlorotrityl resin **296** provided a means of highly selective deprotection of the resin over other acid labile groups, enabling late stage modification of the P3 residue.

Overall, this route appeared to be practicable for the synthesis of cyclotheonamide analogues; however, the key α-ketoamide had been omitted, substituted instead with a simple methylene group. Such non-covalently binding analogues would be anticipated to have lower inhibitory activity against KLK5. The solid phase synthesis of full cyclotheonamide analogues, inclusive of the α-ketoamide functionality, was therefore deemed a necessary challenge.

Additionally, the authors demonstrated that the synthesis of the Dap and vTyr residues was straightforward. As such, the remaining literature review was focused on identifying a suitable method of installing the α-ketoamide functionality in a manner that was compatible with the solid phase synthesis. These additional synthetic strategies are summarised in Table 13 for reference and then examined individually throughout the emainder of this section.

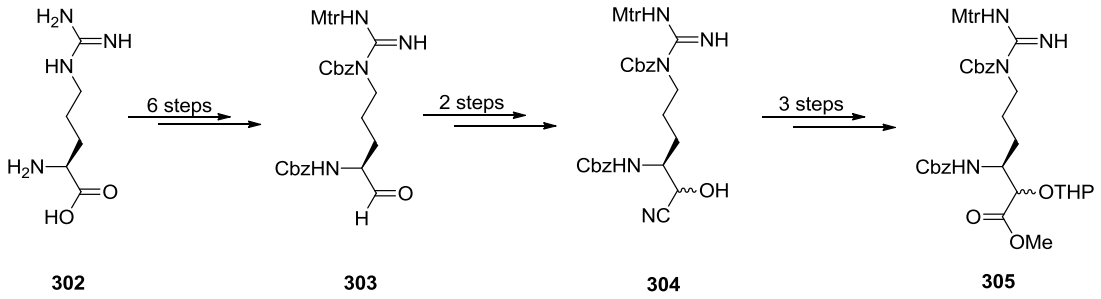
Table 13: Summary of several previous cyclotheonamide syntheses.

Entry[(Ref.)]	Ct	Cyclisation	Key Reaction	Linear Intermediate	OH Protection	OH Oxidation	vTyr Protection	Arg Protection	Deprotection
1[192]	Methylene analogues	Pro/Dap	-	-	-	-	^{t}Bu	Pmc	TFA:H_2O (95:5 v/v), rt
2[194]	A	*D*-Phe/vTyr	α-hydroxyester formation	α-hydroxyamide	THP	DMP, MeCN 80 °C	TIPS	Boc/Mtr	HF.py, anisole, rt then TFA, thioanisole, rt
3[196]	A/B	Pro/Dap	α-hydroxyacid formation	α-hydroxyamide	SEM	DMP, MeCN 65 °C	TBDMS	Ts	HF.py, anisole, rt
4[197]	B	Orn/Pro[1]	5-(furan-2-yl) oxazolidine formation	α-hydroxyamide	Oxazolidine	DMP MeCN/DCM 55 °C	$Bzl(Cl)_2$	Boc[1]	TFA, thioanisole, rt
5[198]	C	Dap/vTyr	Masked acyl cyanide	α-hydroxyamide	TBS	DMP, MeCN 60 °C	TIPS	Cbz/Cbz	HF.py, anisole, rt
6[199]	B	Pro/Dap	α-hydroxyester formation	α-hydroxyamide	None	DMP ^{t}BuOH/DCM rt	^{t}Bu	Boc/Boc	TFA, thioanisole, rt
7[200]	B	Arg/Pro	α-hydroxyester formation	α-hydroxyamide	None	DMP, DCM -	Bn	Boc/Mtr	TFA, thioanisole, rt
8[201]	C	Dap/vTyr	Passereni	α-hydroxyamide	None	DMP, MeCN 70 °C	TIPS	Cbz/Cbz	-
9[203]	E2/E3	Dap/vTyr	α-keto cyanophosphorane oxidation	α-ketoamide	-	-	TIPS	Cbz/Cbz	HF.py, anisole, rt
10[204]	C	Dap/vTyr	α-keto cyanophosphorane oxidation	α-ketoamide	-	-	TIPS	Cbz/Cbz	HF.py, anisole, rt

[1] Boc protected ornithine used as a precursor for late stage incorporation of the arginine residue.

2.1.3.1. Syntheses via α-hydroxyamide intermediates

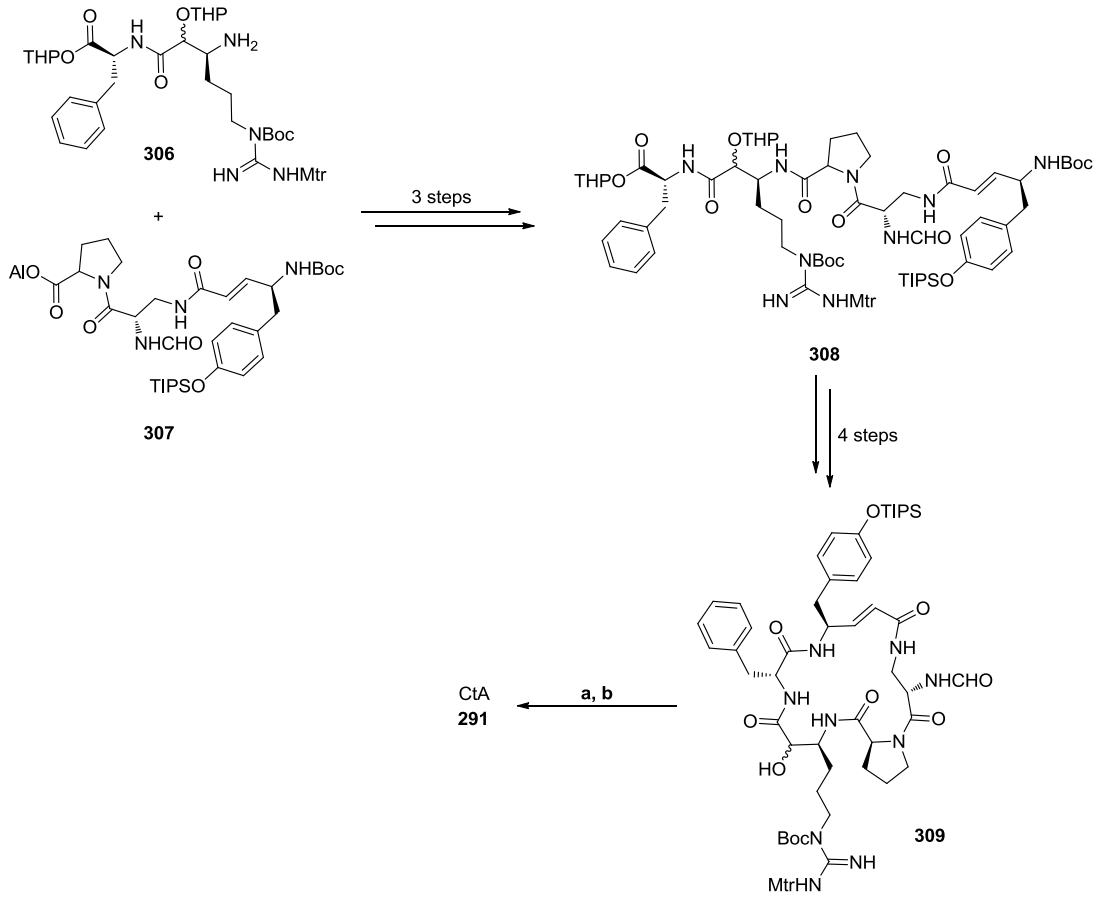
In an early total synthesis of a cyclotheonamide, Kim and Wipf prepared Cbz-Arg(Mtr,Cbz)-CH(OTHP)CO_2Me **305** as a precursor of the α-ketoamide backbone of CtA **291** (Table 13, Entry 2).[194] A lengthy 11 step synthesis yielded the final building block from commercially available *L*-arginine (Scheme 61).[195] After initial protection of the natural arginine residue **302** a Weinreb amide mediated reduction to aldehyde **303** was achieved with $LiAlH_4$. The key formation of protected α-hydroxyester **305** was achieved via formation and acid catalysed hydrolysis of cyanohydrin **304** in methanol followed by THP protection of the resulting alcohol.



Scheme 61: Synthesis of the key α-hydroxyester intermediate for the synthesis of CtA **291**.

The protecting group strategy for this synthesis was complex, with four orthogonal protecting groups employed. Furthermore, some of the protecting groups proved to be labile during the synthesis of the cyclic peptide, while others had to be swapped for more labile alternatives. Overall, the lengthy synthesis and protecting group strategy were inefficient.

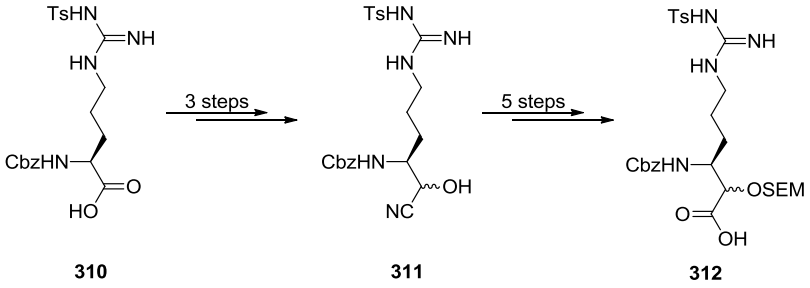
Linear pentapeptide **308** was prepared from α-hydroxyamide (αHA) dipeptide H-Arg-αHA-D-Phe-OTHP **306** and tripeptide Boc-vTyr(OTIPS)-Dap(NHCHO)-Pro-OAl **307** for the final stages of the synthesis (Scheme 62). This convergent approach was efficient for the synthesis of a single cyclotheonamide analogue but would require a lengthy synthesis of complicated building blocks for any subsequent SAR modifications.



Scheme 62: Synthesis of CtA **291** from linear pentapeptide **308**. Reagents and conditions: (a) DMP, MeCN, 80 °C, 1 h, 83%; (b) i. HF/Pyridine, THF; ii. TFA, thioanisole, rt, 36%.

Cyclisation was achieved in 42% yield between the *D*-Phe and vTyr residue via a pentafluorophenol ester during which THP deprotection of the alcohol was achieved simultaneously. DMP oxidation at 80 °C furnished the key α-ketoamide while global deprotection required both HF.pyridine and TFA to yield CtA **291**. The authors noted that oxidation of the alcohol could not be achieved at room temperature, possibly due to conformational shielding of the alcohol. Overall, this route showed that the synthesis of CtA **291** could be achieved via a protected α-hydroxyamide intermediate.

In a similar approach, Maryanoff *et. al.* completed a total synthesis of CtA **291** and CtB **292** with an improved protecting group strategy and alternative cyclisation position (Table 13, Entry 3).[196] In an analogous manner to the previous synthesis, Cbz-Arg(Ts)-OH **310** was reduced to an aldehyde and converted to α-hydroxyacid **312** via cyanohydrin intermediate **311** (Scheme 63).



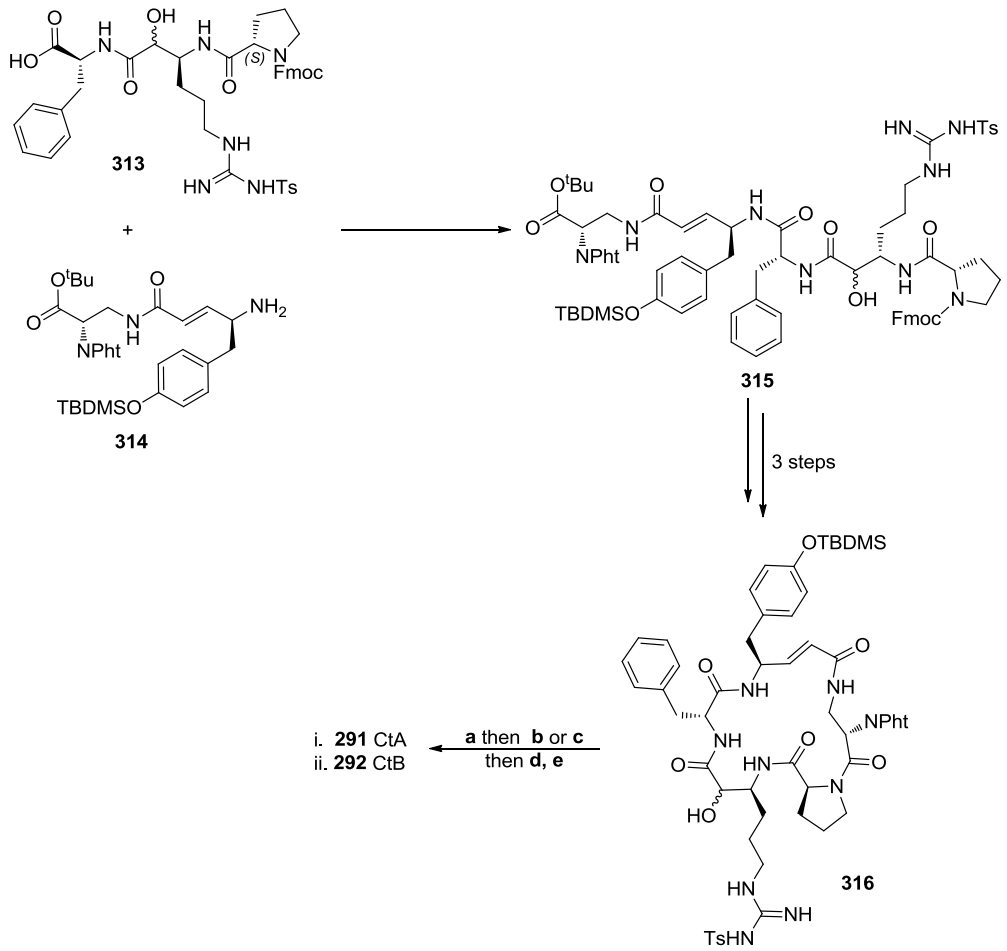
310 **311** **312**

Scheme 63: Synthesis of protected α-hydroxyacid **312** as a precursor for the synthesis of CtA **291** and CtB **292**.

In this case, the use of the silicon based SEM protecting group serendipitously revealed that protection of an α-hydroxyamide was not necessary for the synthesis of cyclotheonamide analogues. The SEM group was found to be labile during a TFA deprotection step and the remainder of the synthesis could be carried out in the presence of the free alcohol (Scheme 64).

Linear, unprotected α-HA pentapeptide **315** was prepared in a convergent manner from Fmoc-Pro-Arg(Ts)-D-Phe-OH **313** and H-vTyr-Dap(NPht)-O^{t}Bu **314**. Cyclisation was achieved between the Pro and Dap residues using DCC mediated amide bond formation to give cyclic pentapeptide **316** in 41% yield, offering another possible cyclisation strategy.

A distinct advantage of this synthesis is the late stage deprotection of the pthalimide protecting group which allowed for late stage modification of the Dap side chain, enabling the synthesis of both CtA **291** and CtB **292** from a common intermediate. This would be beneficial for the synthesis of P3 residue analogues proposed in the current study.

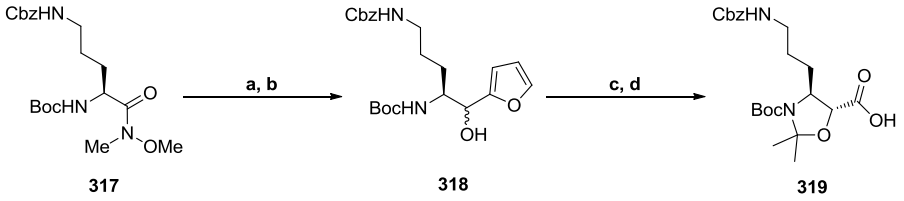


Scheme 64: Synthesis of CtA **291** or CtB **292** via intermediates containing a free α-hydroxy group. Reagents and conditions: (a) N_2H_4, (*Z*)-2-butenol, MeOH, rt, 24 h, 68%; (b) ethyl formate, 55 °C, 3 h, 86%; (c) pentafluorophenyl acetate, DMF, rt, 3 h, 90%; (d) DMP, MeCN, 65 °C, 1-2 h, crude; (e) HF.pyridine, ansiole, -78-25 °C, 3 h; i. CtA **291** 33% (two steps); ii. CtB **292** 33% (two steps).

This synthesis by Maryanoff *et. al.* offers a somewhat more efficient approach to the synthesis of cyclotheonamide analogues for SAR investigations, particularly with regards to the shorter synthesis of the α-hydroxyacid building block and improved deprotection strategy which allowed for global deprotection with HF.pyridine and anisole in a single step.

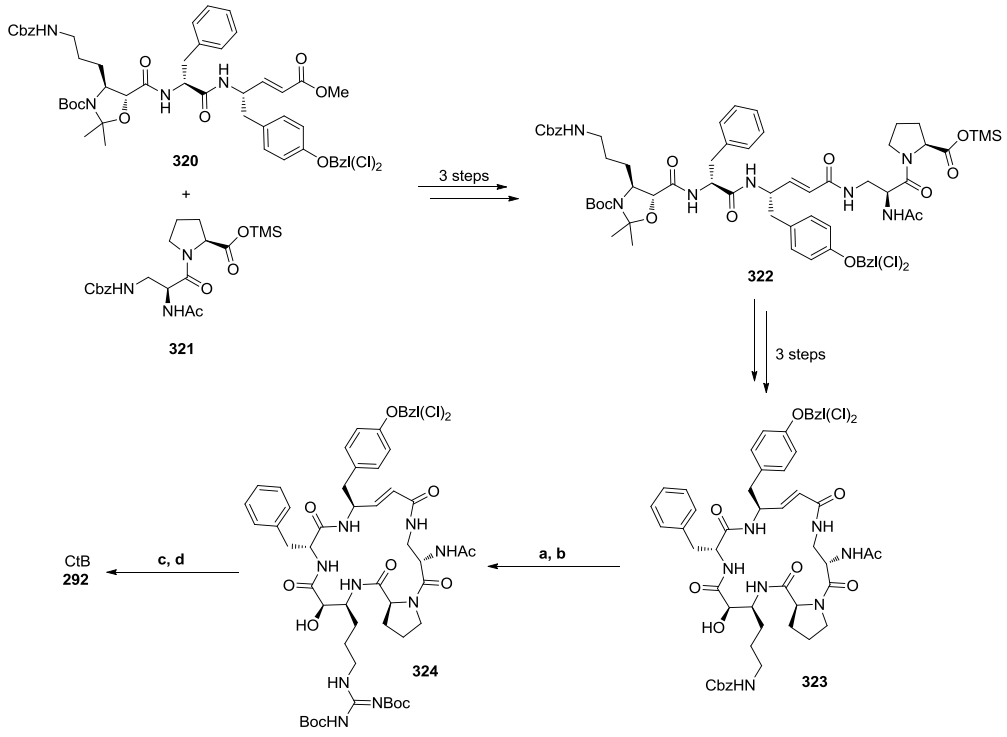
An interesting and more unique synthesis of CtB **292** was performed by Shiori *et. al.* in 1996 using a 5-(furan-2-yl)-oxazolidine protecting group strategy coupled with late stage incorporation of the arginine residue using an ornithine precursor throughout the majority of the synthesis (Table 13, Entry 4).[197] The synthesis of

oxazolidine protected α-hydroxyacid ornithine derivative **319** was achieved from a previously obtained source of Weinreb amide **317** via α-hydroxy furan **318** (Scheme 65).



Scheme 65: Preparation of oxazolidine protected α-hydroxyacid **319** from Weinreb amide **317**. Reagents and conditions: (a) Furyllithium, THF, -78 °C, 35 min; (b) K-selectride, THF, -78 °C, 30 min; (c) 2,2-dimethoxypropane, pyridine-*p*-TsOH, CH_2Cl_2, rt, 60 h; (d) $NaIO_4$, $RuCl_3$ (cat.), MeCN-EtOAc-H_2O (2:2:3 v/v), rt, 1.5 h.

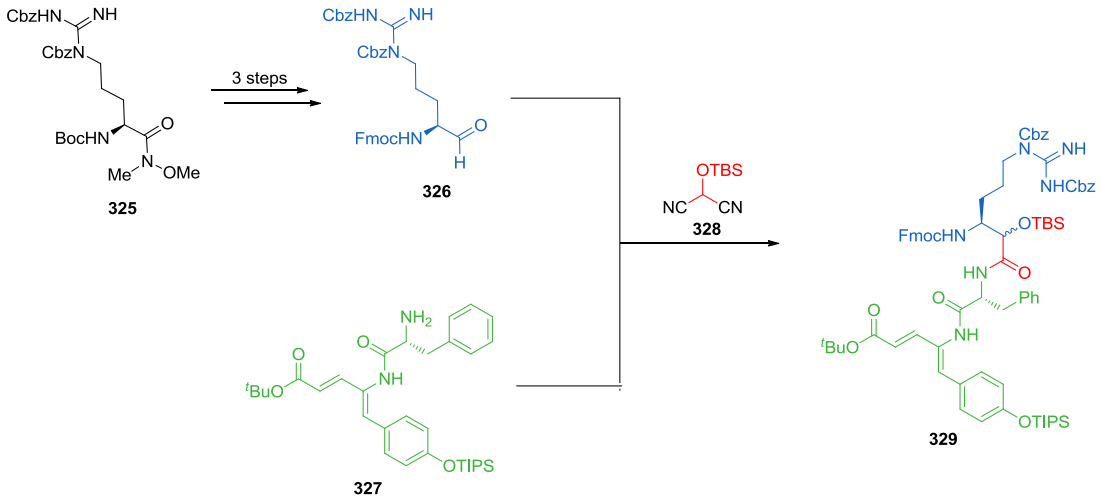
Linear pentapeptide **322** was prepared from tripeptide Boc/Oxazolidine-Arg(Cbz)-αHA-D-Phe-vTyr(Bzl(Cl)$_2$)-OMe **320** and dipeptide Cbz-Dap(NHAc)-Pro-OTMS **321** (Scheme 66).



Scheme 66: Synthesis of CtB **292** via the late stage modification of an ornithine side chain. Reagents and conditions: (a) 31% HBr-HOAc, rt, 25 min; (b) *N,N*'-Di-(*tert*-butoxycabonyl)thiourea, $HgCl_2$, Et_3N, DMF, 0-25 °C, 40 min; (c) DMP, MeCN-CH_2Cl_2 (1:2 v/v), 55 °C, 7 h; (d) TFA, thioanisole, rt, 13 h.

Cyclisation was successful between the Pro and Orn residues in 57% yield, possibly indicating that the absence of the full guanidine moiety is beneficial at this stage of the synthesis. The use of ornithine removes the complex guanidine functionality from the majority of the synthesis, thereby reducing the likelihood of side reactions occurring. Furthermore, this strategy could potentially provide a suitable route for the investigation of ornithine derived P1 residue analogues. However, the yield for this transformation was poor, potentially due to the incompatibility of HBr/HOAc with the vinyl amide. Additionally, the use of toxic reagents such as mercury chloride would make scale up difficult.

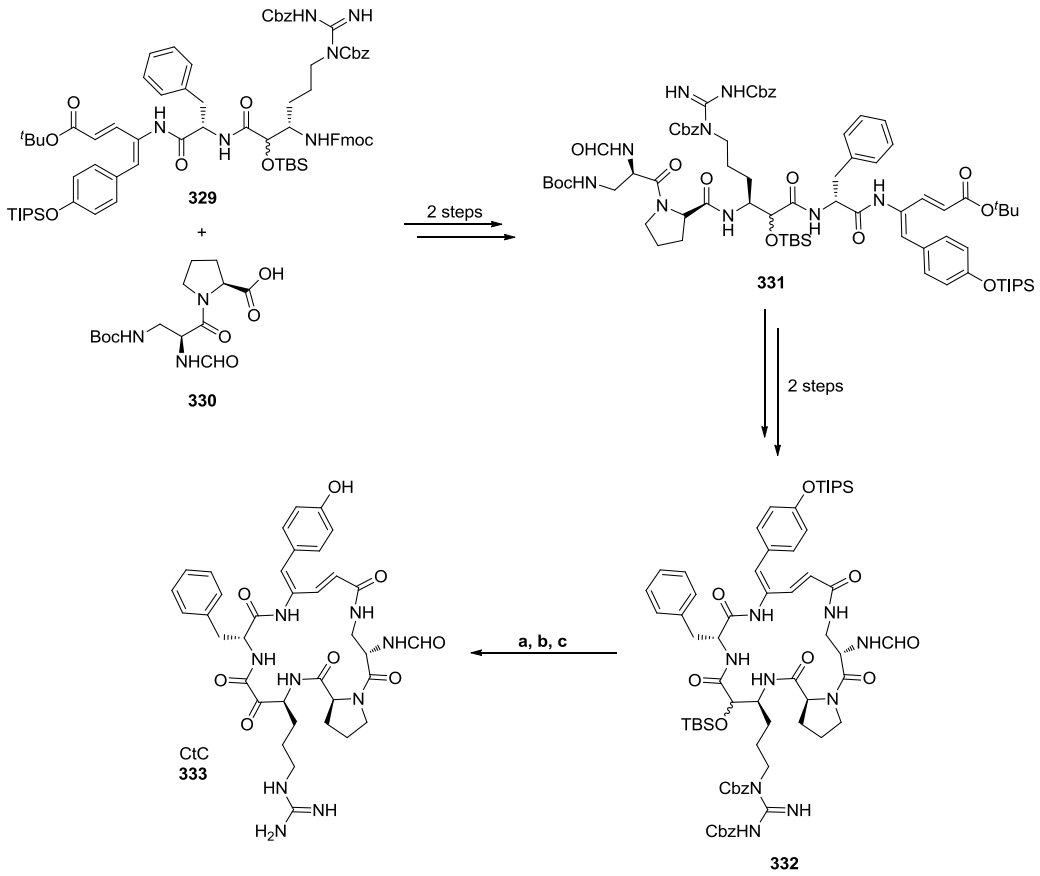
A more recent example of an α-hydroxyamide based cyclotheonamide synthesis was conducted by Aitken *et. al.* in 2008 (Table 13, Entry 5). The group demonstrated the synthesis of CtC **333** using a one-pot three-component masked acyl cyanide reaction to form key α-hydroxyamide intermediate Fmoc-Arg(Cbz)$_2$-αHA(TBS)-D-Phe-vTyr-O^{t}Bu **329** in 53% yield (Scheme 67).[198] Notably, the structure of CtC **333** contains a vinyl dehydrotyrosine residue in place of the usual vinyl tyrosine residue. However, with this residue exposed to the water in the KLK5 site, this feature would have little impact on the SAR in this study.



Scheme 67: Masked acyl cyanide reaction to synthesise TBDMS protected α-hydroxyamide intermediate **329**.

The use of *N*-Fmoc protection on the arginine residue was tolerated during the reaction, suggesting that this methodology could be suitable for developing building blocks for SPPS. However, the route requires preassembly of intermediates **326** and **327**, increasing the number of synthetic steps and reducing its flexibility for SAR studies.

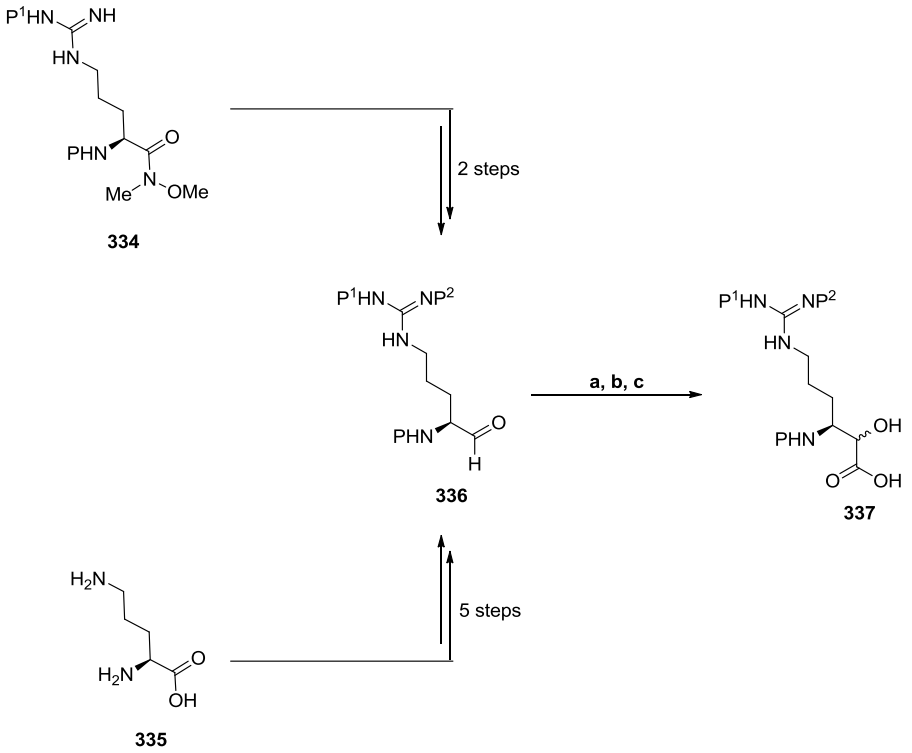
The remaining residues required to form linear pentapeptide **331** are added sequentially, allowing for modification of both the Pro and Dap residues individually (Scheme 68). Efficient simultaneous deprotection of both termini was carried out with TFA prior to macrolactamisation between the Dap and vTyr residues with TBTU. The cyclisation occurs in an excellent yield of 69% with no reported loss of the Cbz protecting groups in TFA.



Scheme 68: Synthesis of CtC **333** from TBDMS protected α-hydroxyamide **329**. Reagents and conditions: (a) HCl/EtOH, 30 °C, 2 h, 35%; (b) DMP, MeCN, 60 °C, 1 h, 71%; (c) HF.pyr, anisole, rt, 12 h, 51%.

The required deprotection of the *O*-TBS group prior to oxidation decreases the efficiency of this route as protection of this alcohol has been demonstrated to be unnecessary (*vide supra*). Serendipitously, the group found that the oxidation could not be achieved in the absence of a protecting group for the phenol. Loss of the TIPS protecting group was observed while conducting a TBAF deprotection of the TBS group and the subsequent oxidation resulted in extensive degradation, indicating that protection of the phenol group is essential. Selective deprotection of TBS could be achieved with HCl/EtOH, allowing the subsequent oxidation and global deprotection to proceed in a similar manner to previous syntheses.

Two independent syntheses of CtB **292** by Schreiber *et. al.* and Ottenheijm *et. al.* (Table 13, Entries 6 and 7) utilised similar methods for synthesis of the α-hydroxyacid building blocks **337** from differently protected arginine derived aldehydes **336** (Scheme 69).[199,200]



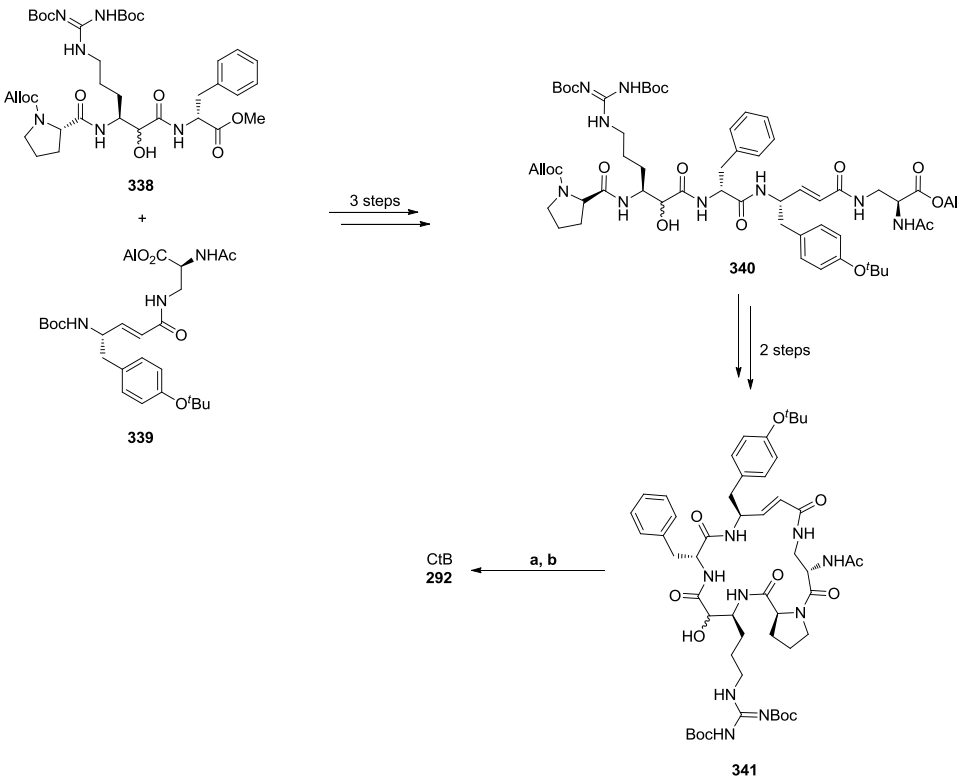
Scheme 69: Synthesis of α-hydroxyacid analogues **337** from Weinreb amide modified arginine residue **334** or ornithine **335**. Reagents and conditions: (a) $LiC(SMe)_3$; (b) HgO, $HgCl_2$, H_2O; (c) LiOH, H_2O.

In both cases, nucleophilic attack on the aldehyde by the anion of tris(methylthio)methane and oxidation with HgO/$HgCl_2$ was followed by hydrolysis to give the protected α-hydroxyacid **337.** These routes offer a useful method of preparing arginine derivatives with multiple protecting group options; however, they are either several steps or require the synthesis of the starting material. Additionally, both routes use toxic mercury reagents.

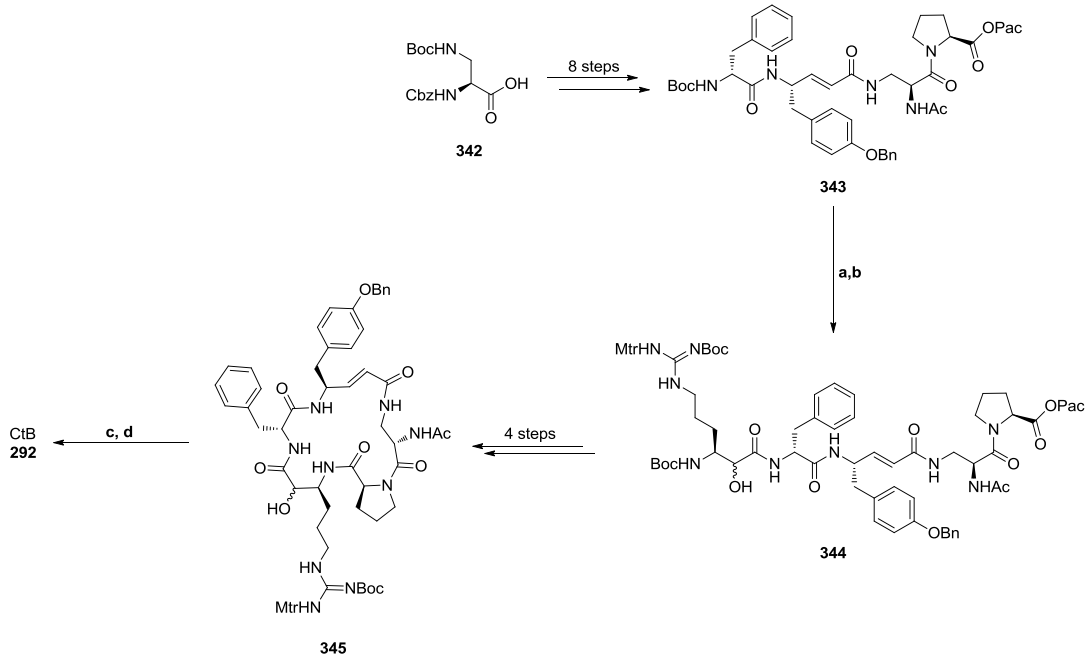
The remaining steps of these two syntheses were significantly different from each other. The synthesis by Ottenjheim *et. al.* followed a similar pathway to previous syntheses with linear pentapeptide **340** being prepared from tripeptide Alloc-Pro-Arg(Boc)$_2$-αHA-D-Phe-OMe **338** and dipeptide Boc-vTyr(^{t}Bu)-Dap(NHAc)-OAl **339** (Scheme 70).

The formation of the linear pentapeptide **340** was followed by efficient simultaneous deprotection of the *N*-alloc and *C*-allyl protected termini. Cyclisation between the Dap and Pro residues with TBTU and HOBt followed to give cyclised pentapeptide **341** in 45% yield over the two steps. Oxidation of the α-hydroxyamide was once again achieved with DMP, though in this case the reaction was reported to proceed at room temperature. The combination of the *tert*-butyl ether and Boc/Boc protecting groups allowed for global deprotection with TFA/thioanisole instead of the more hazardous HF.pyridine reagent.

The synthesis of CtB **292** by Schreiber *et. al.* was one of the earliest reports of a cyclotheonamide synthesis and utilised a more linear approach in which each amino acid residue was added in sequence (Scheme 71). While an inefficient approach for solution phase synthesis, the linear addition of each amino acid residue would be suitable for SPPS and would allow for modification of each individual residue for SAR studies. Importantly, this synthesis demonstrated that a free α-hydroxyacid building block Boc-Arg(Mtr,Boc)-CHOHCO$_2$H could be coupled individually without protection of the hydroxyl group or prior formation of the α-ketoamide. This would allow for modification of the P1 residue for further SAR studies.



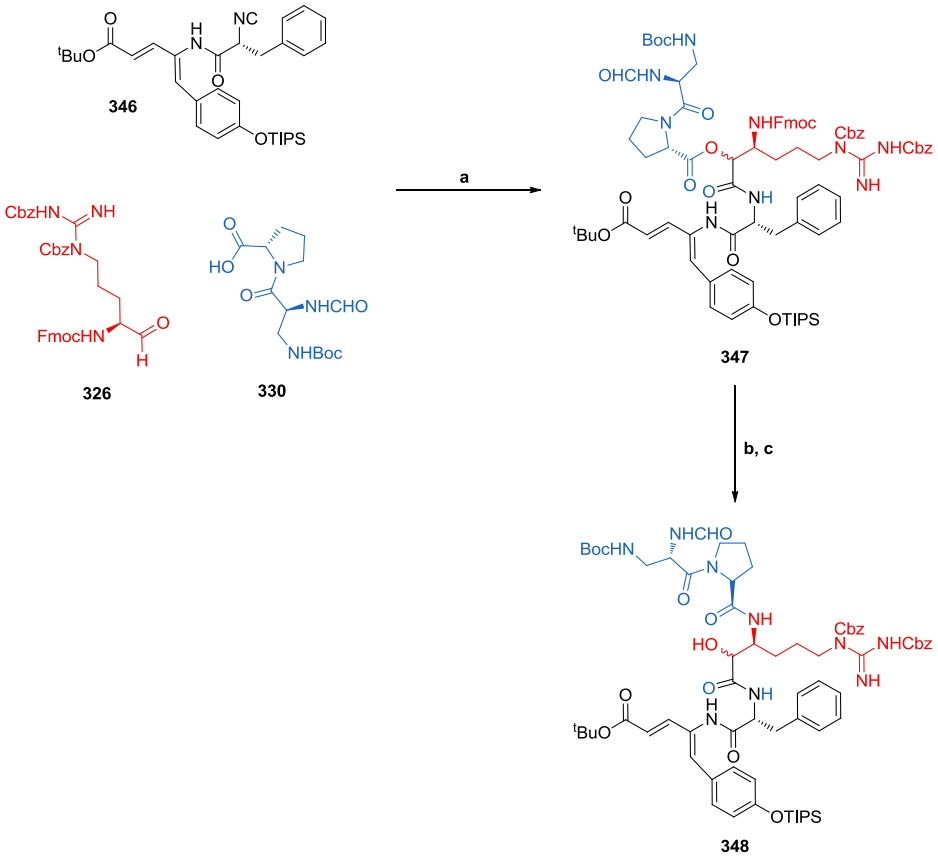
Scheme 70: Synthesis of CtB **292** from tripeptide Alloc-Pro-Arg(Boc)$_2$-αHA-D-Phe-OMe **338** and dipeptide Boc-vTyr(^{t}Bu)-Dap(NHAc)-OAl **339**. Reagents and conditions: (a) DMP, ^{t}BuOH, DCM, rt, 24 h, crude; (b) TFA, thiosanisole, 80 min, 51% (two steps).



Scheme 71: Sequential synthesis of linear pentapeptide **344** and subsequent formation of CtB **292**. Reagents and conditions: (a) *p*-TsOH, MeCN, DCM, crude; (b) Boc-Arg(Mtr,Boc)-CHOHCO$_2$H, BOP, DMAP, DCM, rt, 71% (two steps); (c) DMP, DCM, 74%; (d) TFA, thioanisole, rt, 3 h, 58%.

Cyclisation between the Arg and Pro residues was achieved in four steps by deprotection of the phenylacyl bromide (Pac) group and activation with pentafluorophenol, followed by Boc deprotection with *p*-TsOH and subsequent cyclisation in the presence of base. Cyclic pentapeptide **345** was obtained in 31% yield over the four steps. Oxidation with DMP once again afford the α-ketoamide while the combination of benzyl ether, Mtr and Boc groups again allowed for global deprotection with TFA in place of the more hazardous HF.pyridine.

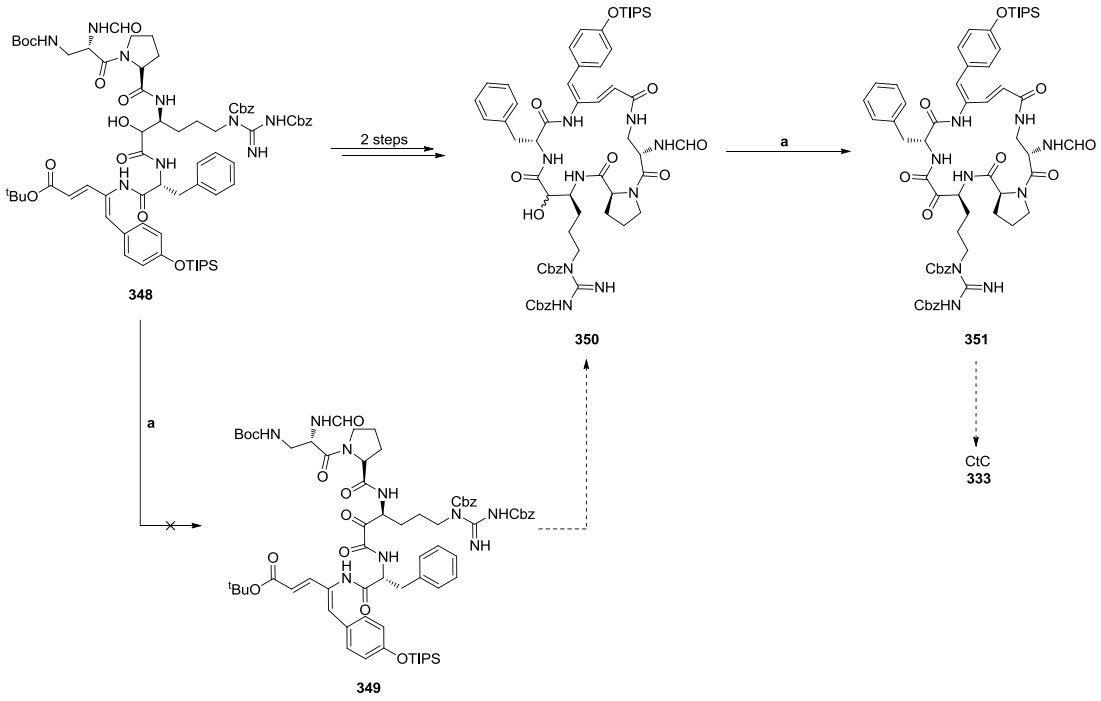
In 2009 Aitken *et. al.* again demonstrated the synthesis of CtC **333** (Table 13, Entry 8), this time by forming unprotected α-hydoxy amide linear pentapeptide **348** using the Passerini-amine deprotection-acyl migration (PADAM) reaction (Scheme 72).[201]



Scheme 72: Synthesis of linear pentapeptide **348** via α-acyloxy amide intermediate **347** using the PADAM reaction. Reagents and conditions: (a) DCM, rt, 6 days, 44%; (b) $NHEt_2$-MeCN, 0 °C, 1 h, crude; (c) NEt_3, DCM, 0 °C, 18 h, 88% (two steps).

Intermediate depsipeptide **347** was formed from the three pre-formed components aldehyde **326**, acid **330**, and isocyanide **346**. The reaction proceeds in 44% yield under mild conditions over 6 days. The *N*-terminus was then deprotected to allow initiation of an *O,N*-dipeptide migration in the presence of triethylamine.

As with their previous synthesis, the group used the Boc/*tert*-butyl combination of protecting groups for orthogonal deprotection with TFA and cyclisation in the presence of TBTU and HOBt to yield the cyclised pentapeptide in 52% yield (Scheme 73). This formal synthesis may be completed with DMP oxidation followed by global deprotection of the previously applied TIPS and *bis*-Cbz groups in the presence of HF. Notably, attempts to oxidise linear pentapeptide **348** failed to yield α-ketoamide linear pentapeptide **349** due to significant degradation of the molecule, attesting to the instability of uncyclised peptides containing α-ketoamide groups.

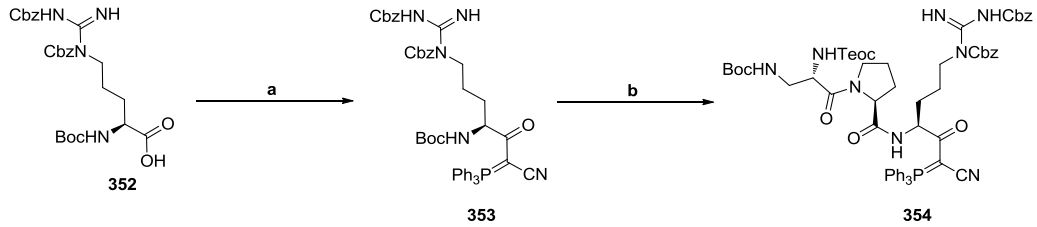


Scheme 73: Formal synthesis of CtC **333** from linear pentapeptide **348**. Reagents and conditions: (a) DMP, MeCN, 70 °C, 1 h, 65%.

On initial inspection this route appears to be highly suitable for the synthesis of cyclotheonamide analogues for SAR studies, especially since the PADAM reaction has been previously utilised for the solid phase synthesis of peptidomimetics.[202] However, the long reaction time required to synthesise the linear pentapeptide is unsuitable for the rapid synthesis of analogues for SAR studies. Furthermore, the requirement of two dipeptide building blocks reduce the flexibility of the reaction to allow for modification of the individual residues.

2.1.3.2. Syntheses of α-ketoamide intermediates

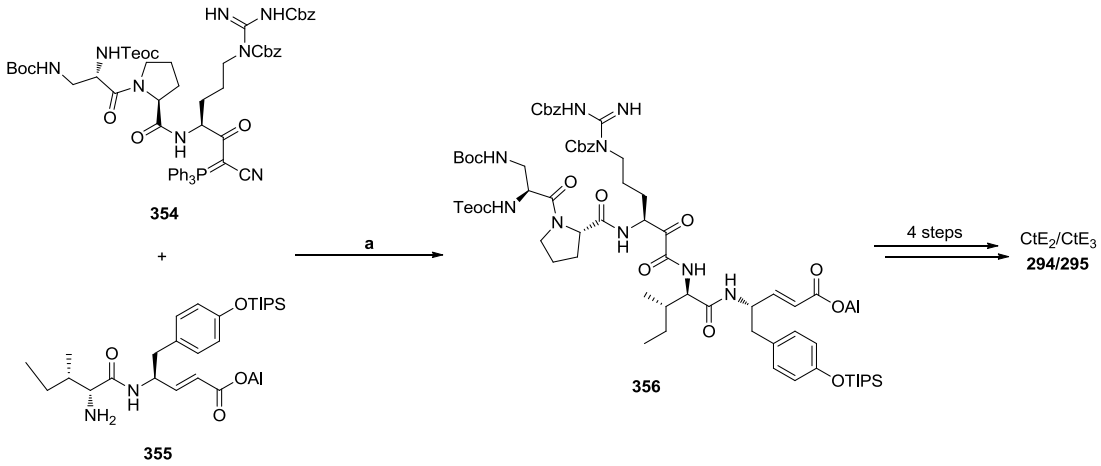
The synthesis of CtE_2 **294** and CtE_3 **295** by Wasserman and Zhang in 2002 demonstrated an alternative method for the installation of α-ketoamides in peptide chains, again without the need to proceed via an α-hydroxyamide intermediate (Table 13, Entry 9).[203] Wasserman's method involved the oxidative activation of α-ketocyanophosphorane intermediate **353** which is easily formed from commercially available Boc-Arg$(Cbz)_2$-OH **352** under conditions akin to an amide coupling reaction (Scheme 74). The α-ketocyanophosphorane was also shown to be stable under acidic conditions and peptide chain extension was possible to give α-ketocyanophosphorane tripeptide **354**.



Scheme 74: Synthesis of α-ketocyanophosporane modified tripeptide **354**. Reagents and conditions: (a) EDC, DMAP, Ph_3P=CHCN, DCM, rt, 4 h, 86%; (b) i. DCM:TFA (1:1 v/v), 1 h, crude; ii. Boc-Dap(Teoc)-Pro-Arg$(Cbz)_2$-COC(PPh_3)CN, EDCI, HOBt, DCM, overnight, 78% (two steps).

Oxidative activation of α-ketocyanophosporane **354** with ozone formed an α-keto acyl cyanide intermediate which acts as a good leaving group and was substituted with the unprotected amine of H-D-Ile-vTyr(OTIPS)-OAl **355** to yield linear

pentapeptide **356** (Scheme 75). The use of Boc, Teoc and CBz protecting groups were shown to be stable under the ozonolysis conditions.

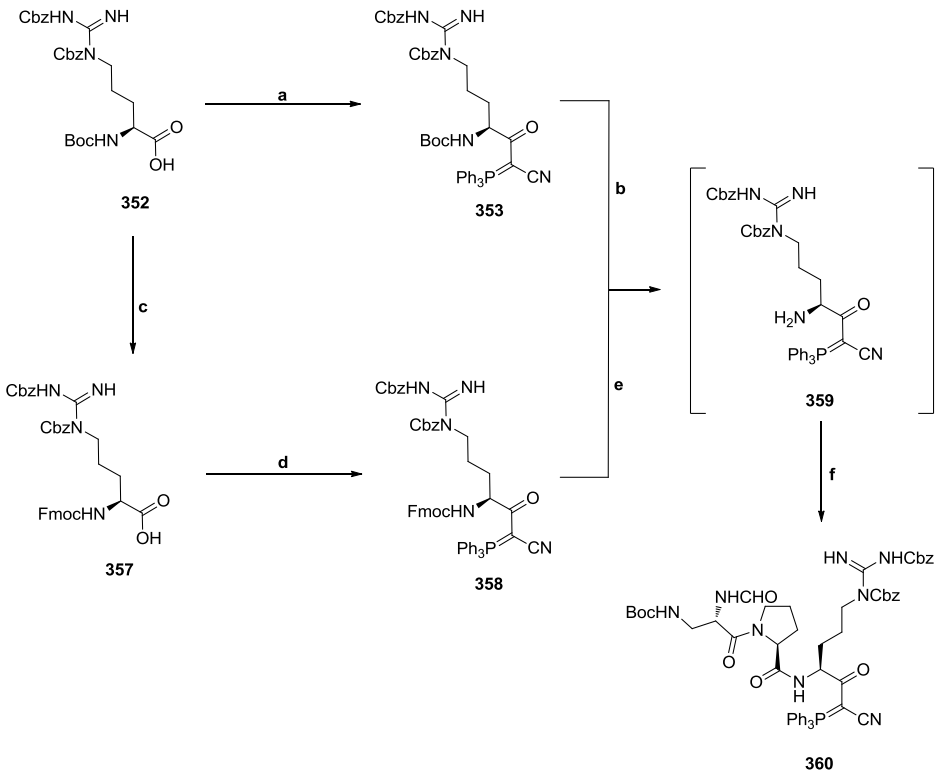


Scheme 75: Synthesis of CtE_2 **294** and CtE_3 **295** via the oxidative activation of α-ketocyanophosphorane **354**. Reagents and conditions: (a) O_3, DCM, -78 °C; then **355**, -78 °C, 15 min, 75%.

While the coupling of a tripeptide and dipeptide fragment would not be efficient for modification of the individual residues, this method could potentially be adapted for the synthesis of α-ketoamide dipeptides as building blocks for the solid phase synthesis. Importantly, the α-ketoamide functionality is already present, removing the need for additional oxidation steps or protecting group chemistry.

Cyclisation between the Dap and vTyr residues could be achieved in 46% yield over two steps but required an inefficient swap of the allyl ester to a more activated pentafluorophenol ester. In addition, the use of Teoc allowed for orthogonal deprotection of the Dap residue and late stage modification of the P3 residue. Once again, global deprotection could be carried out with HF.pyridine to yield either CtE_2 **294** or CtE_3 **295**.

In a more recent synthesis Aitken *et. al.* have utilised the same α-ketocyanophosphorane route for the synthesis of CtC **333** (Table 13, Entry 10).[204] The group investigated the use of α-ketocyanophosporane modified arginine building blocks in more depth during the synthesis of tripeptide **360** (Scheme 76).

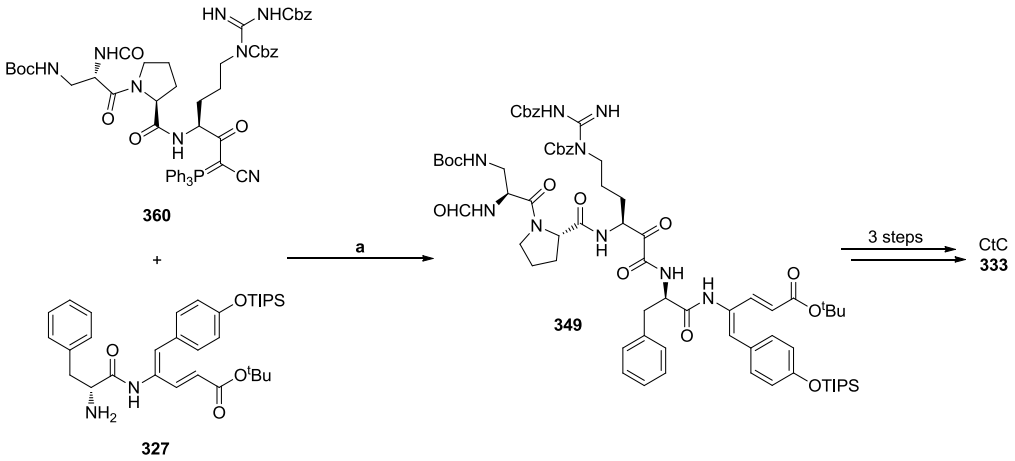


Scheme 76: Investigation of protecting groups for the synthesis of tripeptide α-ketocyanophosphorane **360**. Reagents and conditions: (a) Ph_3P=CHCN, EDCI, DMAP, DCM, 0-25 °C, 12 h, 78%; (b) TFA:DCM (1:1 v/v), 0 °C, 1 h, quench with 10% $Na_2CO_{3(aq)}$, crude; (c) TFA:DCM (1:1 v/v), 0 °C, 1 h, then Fmoc-Cl, Na_2CO_3, 1,4-dioxane:H_2O (3:2 v/v), 0-25 °C, 12 h, 68% (two steps); (d) Ph_3P=CHCN, EDCI, DMAP, DCM, 0-25 °C, 5 h, 41%; (e) $NHEt_2$, MeCN, 0-25 °C, 1 h, crude; (f) Boc-Dap(NHCHO)-Pro-OH, EDCI, DMAP, DCM, 0-25 °C, 12 h; from **353**, 22% (two steps); from **358**, 75% (two steps).

Initially, tripeptide **360** was synthesised from Boc-Arg(Cbz)$_2$-OH **352** in 17% yield over three steps. The group attributed the low yield to degradation observed during the acid deprotection of the Boc group with TFA. As a result, Fmoc-Arg(Cbz)$_2$-OH **357** was synthesised from the original starting material and subsequently used to prepare tripeptide **360** in 31% yield over three steps. The intermediates were found to be more stable under the basic conditions. This alternative protecting group strategy with an Fmoc protected *N*-terminus would be more suitable for the development of building blocks for the SPPS of cyclotheonamide analogues.

Tripeptide Boc-Dap(NHCHO)-Pro-Arg(Cbz)$_2$-COC(PPh_3)CN **360** was oxidatively coupled with dipeptide H-D-Phe-vTyr(OTIPS)-O^{t}Bu **327** (Scheme 77) to afford

α-ketoamide peptide **349**. Notably, the group did not report any degradation under these conditions in comparison to the original attempt to synthesise this molecule (see Scheme 73).



Scheme 77: Synthesis of CtC **333** via the oxidative activation of α-ketocyanophosphorane **360**. Reagents and conditions: (a) O_3, DCM, -78 °C; then **327**, -78-25 °C, 18 h, 48%.

The remainder of the synthesis was carried out in an analogous manner to previous syntheses. Simultaneous deprotection of the termini with TFA and cyclisation between the Dap and vTyr residues was followed by global deprotection with HF pyridine to yield CtC **333**.

2.1.3.3. Summary of reported works and retrosynthesis of CtB

A number of different approaches have been demonstrated for the synthesis of cyclotheonamide products, as summarised in Table 13 (Section 2.1.3).

In general, the solution phase syntheses were convergent and would allow for some modification of the residues as required for the planned SAR studies. However, in most cases a dipeptide or tripeptide fragment was a key intermediate which would require a repeated synthesis for modifications of any of the residues during SAR studies. The use of mainly individual residues for the solid phase synthesis would be more efficient for this purpose.

Cyclisation of the peptide backbone was achieved at multiple points in the ring using a variety of strategies. The use of acid labile *tert*-butyl esters and Boc groups were particularly prevalent; however, Cbz carbamate moieties could partially deprotect during deprotection of these groups. This issue can be avoided with the use of the base labile Fmoc group and the palladium labile allyl groups which were also demonstrated to be suitable for the cyclisations.

Several strategies were investigated for the synthesis of the key α-ketoamide functionality, the majority of which involved an α-hydroxyamide intermediate. Protection of the alcohol group was found to be unnecessary and detrimental to the efficiency and success of the synthesis.

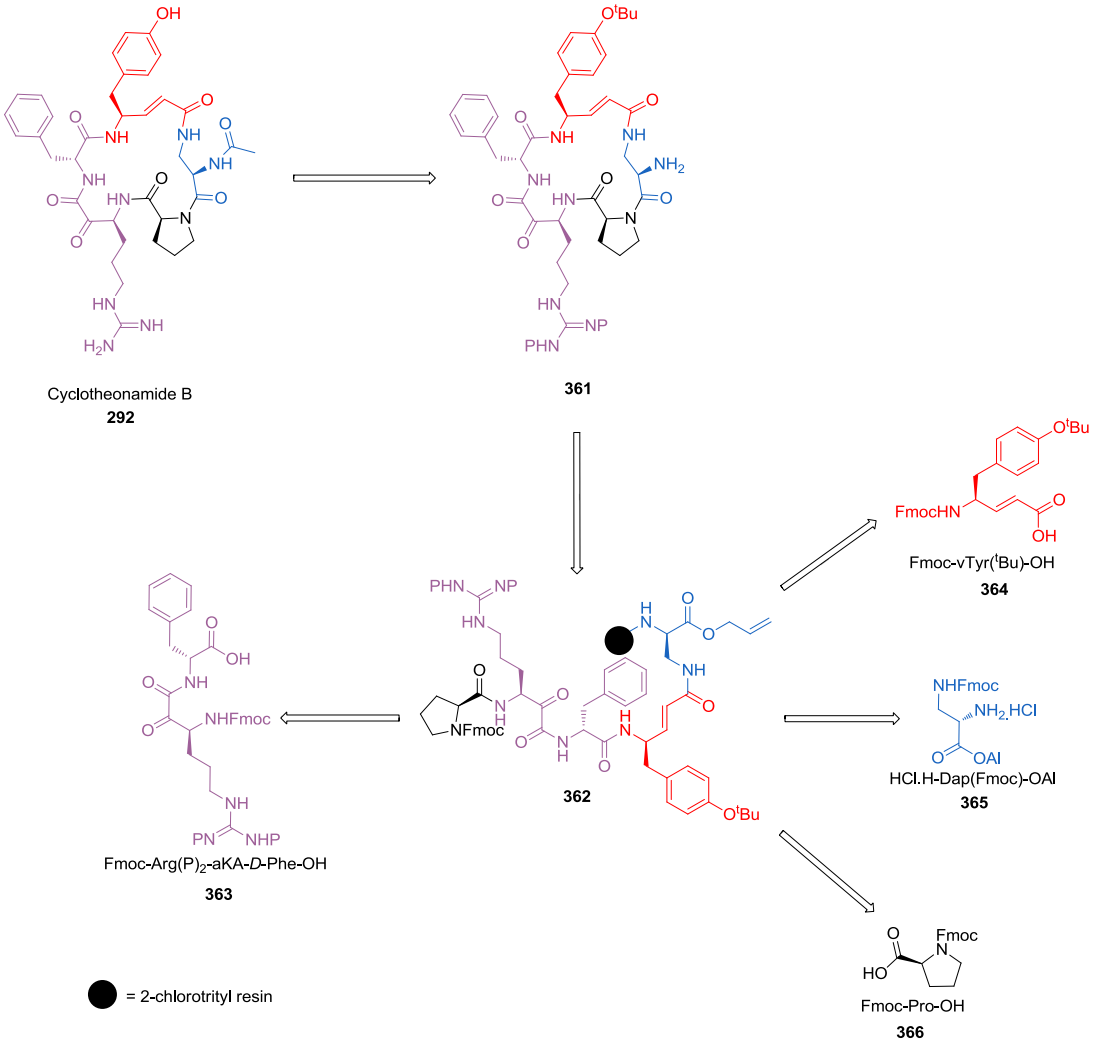
Protection of the vTyr residue was necessary during oxidation of α-hydroxyamide intermediates and TIPS was the predominant protecting group used. However, this required the use of hazardous HF.pyridine to be removed and as such the use of a TFA labile protecting group such as *tert*-butyl ether could be investigated.

The most efficient synthesis of the α-ketoamide functionality was via α-ketocyanophosphorane intermediates which bypassed the need for α-hydroxyamide intermediates. The key functional group could be synthesised in just two steps from a protected Arg-OH residue. Only a limited examination of protecting groups for the guanidine moiety and *N*-terminus was carried out. The utility of this reaction will depend on the ability to conduct the ozonolysis reaction in the presence of *N*-Fmoc moiety in order to develop suitable building blocks for solid phase synthesis. Furthermore, only Cbz protection was investigated for the Arg residue. Therefore, an investigation of a range of acid labile protecting groups will be necessary if this approach is to be adopted.

The combination of the Arg residue with the *D*-Phe residue to build such a building block would be detrimental for the synthesis of S1’ analogues as new dipeptides would need to be prepared each time; however, the simplification of the later

stages of the synthesis of the cyclotheonamide is likely to significantly compensate for this potential disadvantage.

Taking all of this into account, a retrosynthetic plan was proposed for the initial synthesis using CtB **292** as an example (Scheme 78).



Scheme 78: Retrosynthetic route to CtB **292** including retrosynthetic analysis of linear pentapeptide **362**.

This route offers a flexible synthesis of linear pentapeptide **362** from four building blocks. The P2, P3, and vTyr residues are individual building blocks, allowing for easy modification at these positions. The Fmoc-Arg(P)$_2$-αKA-D-Phe-OH **363** building block

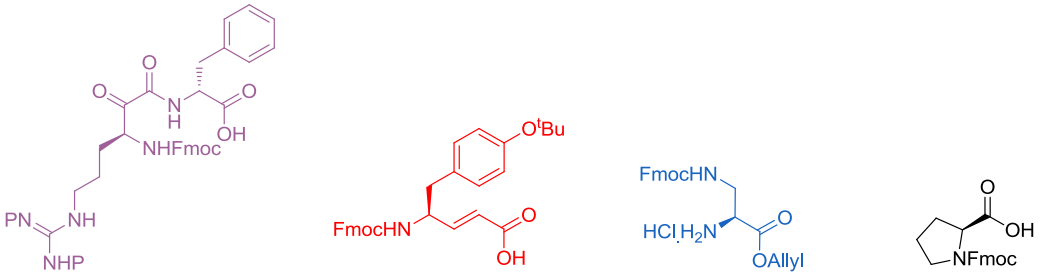
will require resynthesis for modifications of the P1' residue. However, this can be achieved rapidly using the α-ketocyanophosphorane methodology.

Furthermore, the application of SPPS for the preparation of the cyclotheonamide analogues should be rapid with straightforward work up and purification of the reactions, further increasing the flexibility of the approach. Cyclisation between the Dap and Pro residues has been demonstrated on solid phase previously and the late stage cleavage of the resin will allow for facile modification of the P3 residue side chain. Global deprotection of the selected protecting groups should be achievable with TFA instead of HF.pyridine.

Once the requisite building blocks were prepared it was anticipated that the synthesis could be completed following the methodology applied in the previous solid phase synthesis of cyclotheonamide analogues discussed in Section 2.1.3.

2.2. Solid Phase Peptide Synthesis of CtB

Based on the retrosynthetic analysis outlined above, the synthesis of CtB **292** requires four key fragments (Figure 74).



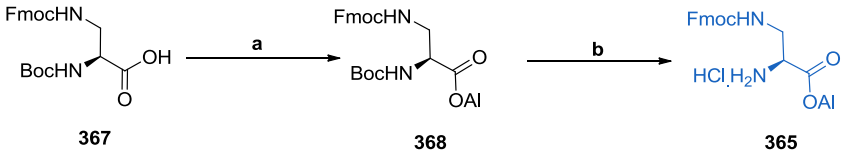
Fmoc-Arg(P)$_2$-aKA-*D*-Phe-OH **363** | Fmoc-vTyr(^{t}Bu)-OH **364** | Fmoc-Dap(NH$_2$.HCl)-OAl **365** | Fmoc-Pro-OH **366**

Figure 74: Requisite amino acids for the synthesis of CtB **292**.

Of these, only Fmoc-Pro-OH **366** was commercially available while the remaining three required assembly. Both Fmoc-Dap(NH$_2$.HCl)-OAl **365** and Fmoc-vTyr(^{t}Bu)-OH **364** had been synthesised previously.[192] The more complicated dipeptide Fmoc-Arg(P)$_2$-αKA-D-Phe-OH **363** required further investigation. The synthesis of each building block is expanded upon in the following sections.

2.2.1. Synthesis of HCl.H-Dap(Fmoc)-OAl

The synthesis of HCl.H-Dap(Fmoc)-OAl **365** was carried out according to an existing literature procedure (Scheme 79).[192]



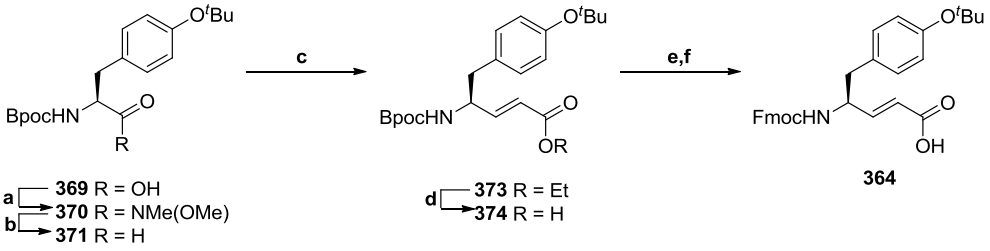
367 **368** **365**

Scheme 79: Synthesis of HCl.H-Dap(Fmoc)-OAl **365**. Reagents and conditions: (a) i. Cs_2CO_3 (0.5 equiv.), 1,4-dioxane:H_2O (1.0:0.3 v/v), rt, 30 min; ii. allyl bromide (1.2 equiv.), DMF, rt, 4.5 h, 72%; (b) i. TFA:H_2O (95:5 v/v), 0-25 °C, 3 h; ii. 4M HCl/dioxane, rt, 95%.

Commercially available Boc-Dap(Fmoc)-OH **367** was stirred with half an equivalent of cesium carbonate and reacted with allyl bromide via an S_N2' reaction to give allyl protected analogue **368** in 72% yield. The Boc protection was removed with TFA:H_2O (95:5 v/v) and the crude material dried azeotropically with 4M HCl in dioxane to afford HCl.H-Dap(Fmoc)-OAl **365** in 95% yield.

2.2.2. Synthesis of Fmoc-vTyr(*t*Bu)-OH

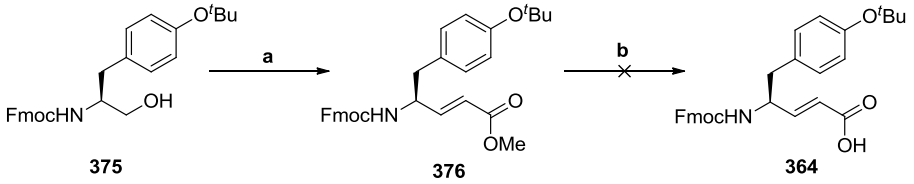
Schaschke *et. al.* reported the synthesis of Fmoc-vTyr(*t*Bu)-OH **364** from Bpoc-Tyr(*t*Bu)-OH via formation of an aldehyde and a subsequent Horner-Wadsworth-Emmons (HWE) olefination reaction to give the desired *E*-alkene (Scheme 80).[192]



Scheme 80: Literature synthesis of Fmoc-vTyr(*t*Bu)-OH **184**. Reagents and conditions: (a) HCl.HNMe(OMe), TBTU, HOBt, DIPEA, $CHCl_3$, rt, overnight, 50%; (b) $LiAlH_4$, Et_2O, 0-25 °C, 2 h, 97%; (c) $(EtO)_2P(O)$-CH_2CO_2Et, NaH, THF, -50-25 °C, 2.5 h, 65%; (d) LiOH, 1,4-dioxane:H_2O (4:1 v/v), rt, overnight, 93%; (e) TFA:DCM (1.5:98.5 v/v), rt, 3 h, crude; (f) Fmoc-OSu, $NaHCO_3$, dioxane:H_2O (2:1 v/v), rt, overnight, 53% (two steps).

The significant acid lability of Bpoc allowed this group to be replaced by Fmoc in the presence of the *tert*-butyl ester. However, the starting material and Bpoc protecting reagent were not immediately available; therefore, an alternative route was sought.

McMurray *et. al.* have shown that the synthesis of vinyl esters is also possible in the presence of the Fmoc group.[205] A literature survey indicated that the hydrolysis of a methyl ester protecting group is still possible under basic conditions in the presence of the Fmoc group.[206,207] As such, Fmoc-vTyr(^{t}Bu)-OMe **376** was synthesised with the intention of selectively deprotecting the methyl ester group (Scheme 81).

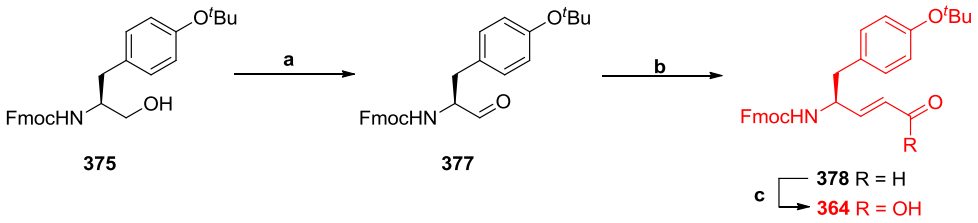


Scheme 81: Attempted synthesis of Fmoc-vTyr(^{t}Bu)-OH **364** via intermediate vinyl ester Fmoc-vTyr(^{t}Bu)-OMe **376**. Reagents and conditions: (a) i. DMP (1.5 equiv.), DCM, rt, 5 h, crude; ii. $Ph_3P{=}CHCO_2Me$ (1.5 equiv.), DCM, rt, 2 h, 28% (2 steps); (b) LiOH (12 equiv.), $CaCl_2$ (15 equiv.), IPA:THF (4:1 v/v), rt.

Commercially available Fmoc-Tyr(^{t}Bu)-CH_2OH **375** was oxidised under Dess-Martin conditions to give the corresponding aldehyde **377** which was immediately reacted with methyl 2-(triphenylphosphoranylidene)acetate to give vinyl methyl ester **376**. The *trans* stereochemistry of the alkene was confirmed by ^{1}H NMR spectroscopy with the observation of an alkenyl proton at 5.86 ppm (d, J = 14.6 Hz, 1 H). The coupling value of 14.6 Hz is indicative of the *trans* stereochemistry of the double bond.[208] The product was isolated in a yield of only 28% due to partial degradation and difficulty in separating the triphenylphosphine oxide byproduct. No further work was conducted on this reaction as the subsequent hydrolysis failed to selectively deprotect the methyl ester in the presence of the Fmoc group.

Instead of screening a number of conditions to selectively deprotect the methyl ester in the presence of the Boc group it was more efficient to alter the route

slightly. A literature survey provided an alternative route to the product via Pinnick oxidation of the analogous vinyl aldehyde (Scheme 82).[209]



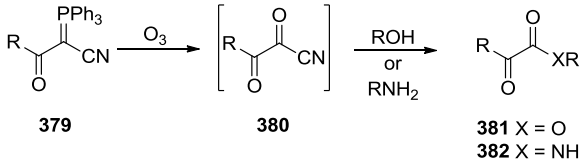
Scheme 82: Synthesis of Fmoc-vTyr(^{t}Bu)-OH **364** via intermediate vinyl aldehyde Fmoc-vTyr(^{t}Bu)-H **377**. Reagents and conditions: (a) DMP (1.3 equiv.), DCM, rt, 5 h, 91%, crude; (b) Ph_3P=H(CHO) (1.2 equiv.), DCM, 76 h, 50%; (c) $NaClO_2$ (10 equiv.), $NaH_2PO_4.H_2O$ (7.5 equiv.), *tert*-butanol:2-methyl-2-butene:H_2O (2.8:1.0:1.8 v/v), rt, 3 h, 62%.

Fmoc-Tyr(^{t}Bu)-CH_2OH **375** was once again oxidised under DMP conditions to give crude aldehyde **377** in 91% yield. Synthesis of vinyl aldehyde **378** was achieved in 50% yield using Wittig reagent 2-(triphenylphosphoranylidene)acetaldehyde. In hindsight, the use of the corresponding HWE reagent would be beneficial as a simple water wash could be used for purification. Once again the stereochemistry was confirmed by identification of a doublet alkenyl signal at 6.77 ppm with a coupling constant of 13.9 Hz, indicative of a *trans* double bond. Pinnick oxidation was achieved in 62% yield to afford the required Fmoc-vTyr(^{t}Bu)-OH **364** building block.

2.2.3. Synthesis of Fmoc-Arg(P)$_2$-αKA-D-Phe-OH

Based on the examination of previous cyclotheonamide syntheses (Section 2.1.3) it was determined that the preferred method for synthesising the α-ketoamide group was via the use of Wasserman's reaction (Scheme 83). Oxidation of α-ketocyanophosphorane **379** results in the *in-situ* formation of intermediate acyl cyanide **380** with triphenylphosphine oxide as the only byproduct. Subsequent addition of a nucleophile leads to rapid reaction with this highly reactive

intermediate. The reaction is most commonly used for the synthesis of α-ketoesters **381** and α-ketoamides **382**.



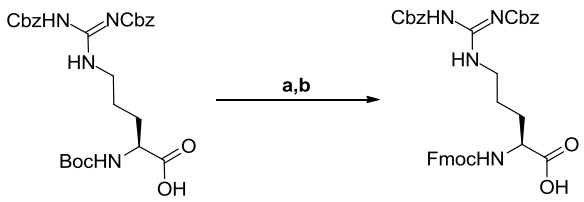
379 **380** **381** X = O **382** X = NH

Scheme 83: Formation of α-ketoamides or esters from α-ketocyanophosphoranes.

As part of the overall strategy for the SPPS of CtB **292** it was desirable to use protecting groups for the arginine side chain which could be concomitantly removed with all other protecting groups in the cyclic peptide. With the current synthetic strategy, the only other protecting group would be the acid labile *tert*-butyl ether present in the vTyr residue. The range of available and suitable arginine protecting groups was also limited by the presence of the sensitive alkene and α-ketoamide moieties. For example, protecting groups which require nucleophilic reagents, oxidising conditions or reducing conditions would not be suitable.

Wasserman's reaction had been successfully applied to the synthesis of CtE_2 **294** and CtE_3 **295** in the presence of Boc and Cbz protecting groups but had not been demonstrated in the presence of Fmoc protecting groups.[204]

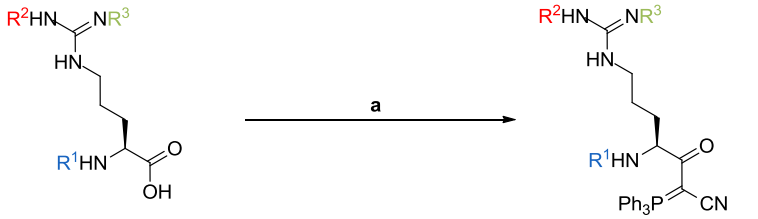
A number of commercially available Fmoc-Arg$(P)_2$-OH derivatives were identified to further explore this chemistry as the Fmoc protection was desirable for the solid phase synthesis. Of these, several bore sulfonamide moieties (*e.g.* Mtr, Pbf, Ts *etc*) and Fmoc-Arg(Pbf)-OH **383** was selected to represent this class of protecting groups. Fmoc-Arg$(Boc)_2$-OH **384** and Fmoc-Arg$(Cbz)_2$-OH **357** were also selected for their acid labile protecting groups. Although it was not commercially available, Fmoc-Arg$(Cbz)_2$-OH **357** could be synthesised from commercial supplies of Boc-Arg$(Cbz)_2$-OH **352** in 45% yield via a TFA driven Boc deprotection and subsequent Fmoc protection with Fmoc-Cl and sodium carbonate (Scheme 84).



352 **357**

Scheme 84: Synthesis of Fmoc-Arg(Cbz)$_2$-CO$_2$H **357** from Boc-Arg(Cbz)$_2$-CO$_2$H **352**. Reagents and conditions: (a) TFA:DCM (1:1 v/v), rt, 3 h, crude; (b) Fmoc-Cl (1.1 equiv.), 10% $Na_2CO_{3(aq)}$ (6.2 equiv.), 1,4-dioxane, 0-25 °C, 2 h, 45% (2 steps).

The first stage of the building block synthesis required the formation of the α-ketocyanophosphoranes of each respective arginine starting material (Scheme 85). Based on the literature, the α-ketocyanophosphoranes were synthesised under conditions similar to an amide coupling using EDC, (triphenylphosphoranylidene)acetonitrile, and catalytic DMAP.[204]



383 R^1 = Fmoc, R^2 = Pbf, R^3 = H
384 R^1 = Fmoc, R^2 = Boc, R^3 = Boc
357 R^1 = Fmoc, R^2 = Cbz, R^3 = Cbz
352 R^1 = Boc, R^2 = Cbz, R^3 = Cbz

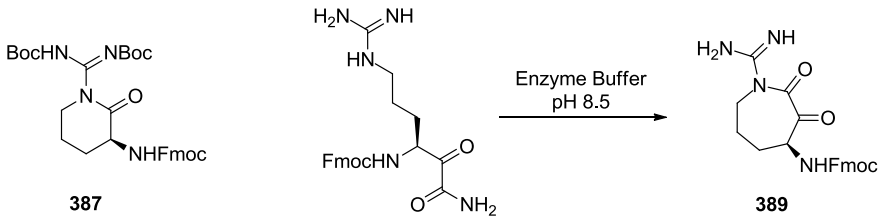
i. **385** R^1 = Fmoc, R^2 = Pbf, R^3 = H
ii. **386** R^1 = Fmoc, R^2 = Boc, R^3 = Boc
iii. **358** R^1 = Fmoc, R^2 = Cbz, R^3 = Cbz
iv. **353** R^1 = Boc, R^2 = Cbz, R^3 = Cbz

Scheme **85**: Synthesis of α-ketocyanophosphorane derivatives of protected arginine building blocks. Reagents and conditions: (a) DMAP (0.1 equiv.), EDC.HCl (1.3 equiv.), $PPh_3CH(CN)$ (1.3 equiv.), DCM, 0-25 °C, 16 h; i. **385**, 82%; ii. **386**, required $PPh_3CH(CN)$ (3 equiv.), 5%; iii. **358**, 5%; iv. **353**, 45%.

The Fmoc/Pbf protected arginine was the most successful starting material for these reactions, giving Fmoc-Arg(Pbf)-COC(PPh_3)CN **385** in 82% yield. The α-ketocyanophosphorane group was identified by a characteristic ^{13}C NMR peak at δ_C = 194.4 ppm. The product contained a 12% triphenylphosphine oxide (TPPO) impurity as this could not be completely separated by chromatography; however, this was not deemed to be problematic as the following ozonolysis reaction would

result in additional TPPO as a byproduct. Therefore, this material was used without further purification.

Synthesis of Fmoc-Arg(Boc)$_2$-COC(PPh$_3$)CN **386** was successful but in only 5% yield, with two possible explanations for this. Firstly, the reaction did not proceed to completion, even when additional equivalents of the PPh$_3$CH(CN) reagent were used, and secondly, the desired α-ketocyanophosphorane **386** was observed to be the minor product of the reaction by LC-MS. The unisolated major product is suggested to be cyclised byproduct **387** by LC-MS ($[M+H]^+$ = 579.3) which can be formed during the activation of the carboxylic acid with EDC (Scheme 86). Similar cyclisations have been observed with primary α-ketoamide modified arginine **388** under enzymatic conditions.[185] While the Boc protecting combination was used for protection of the arginine during previous cyclotheonamide syntheses, it was only employed in combination with the bulky Mtr protecting group or installed post cyclisation of the peptide backbone. [197,199] There exists the possibility for further cyclisation to occur at later stages of the synthesis; therefore, no further work was conducted with this starting material.

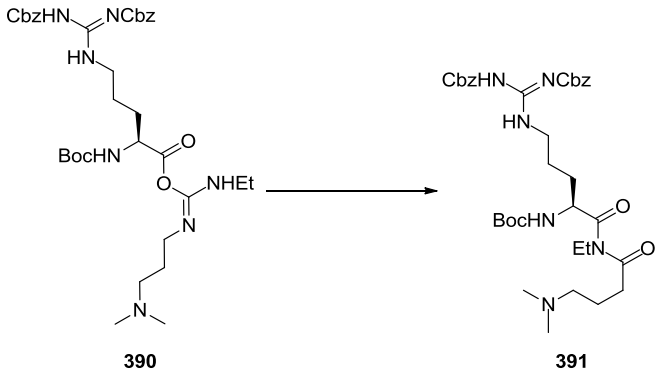


Scheme 86: Potential structure of cyclised byproduct **387** and previously observed cyclisation of α-ketoamide modified arginine **388**.

While the synthesis of Fmoc-Arg(Cbz)$_2$-COC(PPh$_3$)CN **358** from Fmoc-Arg(Cbz)$_2$-OH **357** was shown to be feasible in 41% yield by Aitken *et. al*. only a comparatively poor yield of 5% could be obtained in our hands.[204] The reaction profile itself was poor, as determined by LC-MS, with a number of byproducts observed. Multiple purifications were required, each of which appeared to decrease the purity of the material suggesting that it was unstable on silica. The structure of the

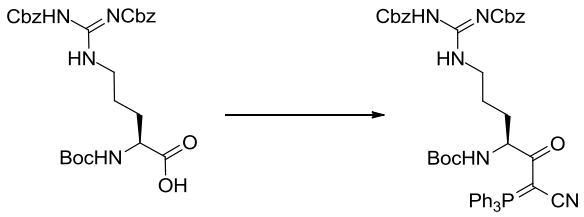
product was confirmed by ^{13}C NMR spectroscopy using the characteristic α-ketocyanophosphorane peak (δ_C = 193.7 ppm). This product, with its desirable combination of protecting groups, was tested in subsequent reactions despite the poor yield (*vide infra*).

In the same report, Aitken *et. al.* had also demonstrated that the synthesis of Boc-Arg(Cbz)$_2$-COC(PPh$_3$)CN **353** could be achieved in an improved yield of 78% in comparison to the Fmoc protected analogue. This material was prepared in the event that the other α-ketocyanophosphoranes proved to be untenable to further reactions. Again, a comparatively lower yield of 45% was obtained compared to that obtained in the literature. However, in this case the lower yield was determined to be due to the rapid formation of the EDC activated ester **390** but slow formation of the desired product, resulting in the formation of N-acyl urea **391** via an intramolecular rearrangement (Scheme 87).[210]



Scheme 87: Byproduct **391**, observed during the synthesis of α-ketocyanophosphorane derivatives of arginine building blocks.

Additives have been developed to be used in combination with EDC with the advantageous effects of preventing formation of this byproduct, increasing the rate of coupling, and reducing racemisation; one such example is HOAt.[211] Accordingly, the reaction was optimised (Scheme 88, Table 14).



352 **353**

Scheme 88: Synthesis of α-ketocyanophosphorane **353** from Boc-Arg(Cbz)$_2$-OH **352**.

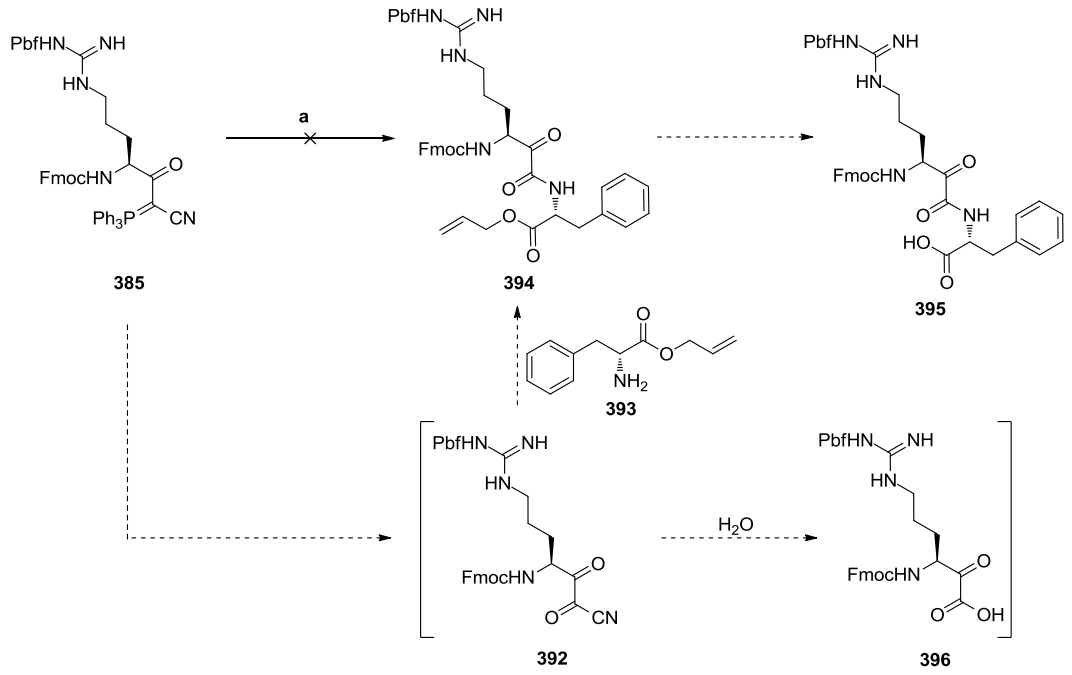
		Reagent Equivalents					
Entry	**Scale/ g**	**$PPh_3CH(CN)$**	**DMAP**	**EDC**	**HOAt**	**Time/h**	**Yield/%**
1	5	1.3	0.1	1.3	0	16	45
2	0.1	2	0.1	1	1	22	51[a]
3	0.1	3	0.1	1.3	1.3	5	74
4	0.1	3	0	1.3	1.3	6	76
5	1	3	0	1.3	1.3	6.5	89
6	5	3	0	1.3	1.3	6.5	95

[a]Reaction did not go to completion but reaction profile significantly improved

Table 14: Optimisation of the coupling reaction with triphenylphosphoranylidene acetonitrile.

Repeating the reaction with a combination of HOAt (1 equiv.), DMAP (0.1 equiv.), $PPh_3CH(CN)$ (2 equiv.) and EDC (1 equiv.) resulted in an immediate improvement in reaction profile and a slight improvement in yield on a small scale (Entry 2). However, the reaction failed to go to completion after 22 hours. Increasing the stoichiometry of the coupling reagent and additive allowed the reaction to proceed to completion within a much shorter time frame (Entry 3) and removal of the base under the same conditions gave no deleterious effect on the yield or reaction time (Entry 4). With these new reaction conditions it was possible to conduct the reaction on both 1 g and 5 g scales with a significant improvement in yield each time (Entries 5 and 6, respectively).

With three α-ketocyanophosphorane reagents now available, the remaining steps towards the synthesis of protected α-ketoamide dipeptides were tested. Firstly, the synthesis of Pbf protected α-ketoamide dipeptide **395** from Fmoc-Arg(Pbf)-COC(PPh_3)CN **385** was attempted (Scheme 89).



Scheme 89: Attempted synthesis of Fmoc-Arg(Pbf)-αKA-D-Phe-OH **395** via α-ketocyanophosphorane **385**. Reagents and conditions: (a) O_3, DCM, then H-D-Phe-OAl (1 equiv.) in DCM:IPA (3:1 v/v), -78 °C, degraded.

The synthesis of Fmoc-Arg(Pbf)-αKA-D-Phe-OH **395** was envisaged to proceed by ozonolysis of **385** to give the acyl cyanide intermediate **392** followed by nucleophilic substitution with H-D-Phe-OAl **393** to give Fmoc-Arg(Pbf)-αKA-D-Phe-OAl **394**. Orthogonal deprotection of the allyl ester protecting group under standard palladium conditions would then yield the desired dipeptide building block. The reaction progress was monitored by LC-MS and evidence of acyl cyanide intermediate **392** was expected to be observed in the form of water quenched α-ketoacid **396**.

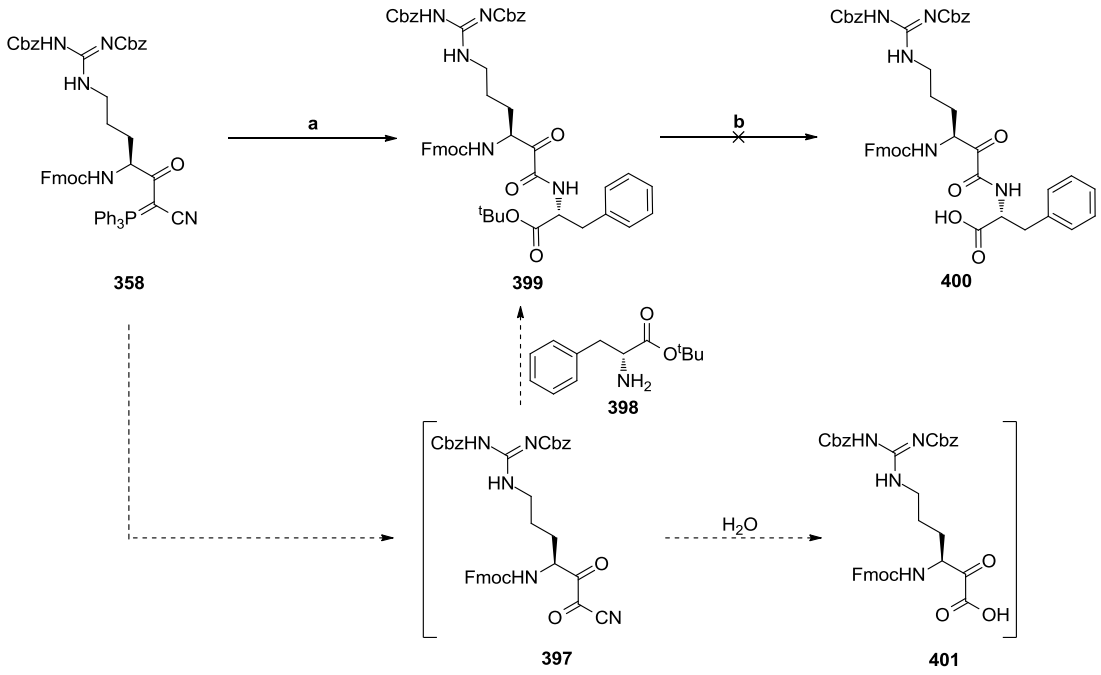
When conducting the ozonolysis of **385** in DCM at -78 °C the resulting reaction mixture showed a complex mixture of products. The expected α-ketoacid **396** was not observed by LC-MS after quenching with water, and subsequent addition of poorly soluble H-D-Phe-OAl **393** in DCM:IPA (3:1 v/v) had no effect on the reaction mixture. This result was unexpected since the Wasserman reaction had been previously reported for analogous arginine derivatives.[203,204]

Subjecting Fmoc-Arg(Pbf)-OH **383** to the same ozonolysis conditions resulted in a very similar LC-MS profile, indicating that the combination of Fmoc and Pbf protecting groups was not stable under these reaction conditions; consequently, this route was abandoned.

While the sulfonamide family of protecting groups has been employed in several of the aforementioned syntheses of cyclotheonamide products, the decomposition of Fmoc-Arg(Pbf)-OH **383** suggested that they would likely be incompatible with this particular reaction. This ruled out the use of commercially available Arg(Ts), Arg(Mtr) and Arg(Mts) protected analogues for this synthetic strategy.

Despite its low yielding synthesis, Fmoc-Arg(Cbz)$_2$-COC(PPh$_3$)CN **358** was next subjected to the same reaction conditions in an effort to synthesise the corresponding dipeptide Fmoc-Arg(Cbz)$_2$-αKA-D-Phe-OH **400** (Scheme 90). In this case H-D-Phe-OAl **393** was replaced with the more soluble H-D-Phe-O^{t}Bu **398**. The acid mediated deprotection of the *tert*-butyl ester was predicted to occur more rapidly than the less labile Cbz protecting groups.

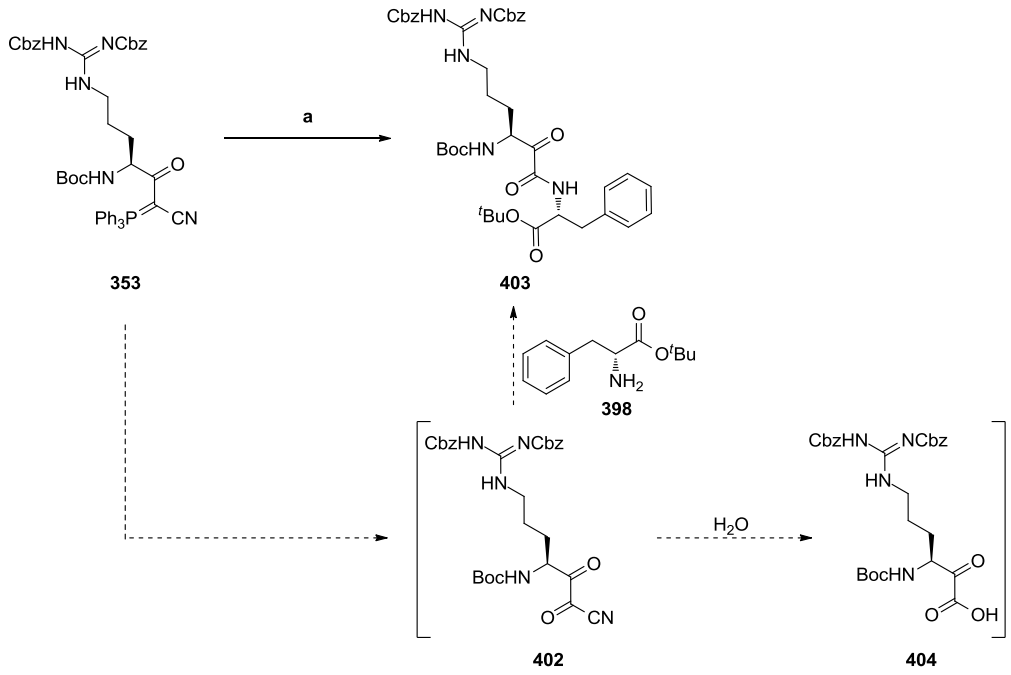
The ozonolysis of Fmoc-Arg(Cbz)$_2$-COC(PPh$_3$)CN **358** also gave a poor reaction profile with a number of degradation peaks observed. Interestingly, the two major products of the reaction were observed in a 1:1 ratio by LC-MS with $[M+H]^+$ of 693.4 and 665.4. The former is consistent with the mass of the expected α-ketoacid **401** while the latter was consistent with the retention time and mass of Fmoc-Arg(Cbz)$_2$-OH **357**, suggesting that the cyanophosphorane itself can be oxidatively converted back to this starting material. This indicated that the combination of Fmoc and Cbz protecting groups was somewhat more stable to ozonolysis than the Fmoc and Pbf combination.



Scheme 90: Attempted synthesis of Fmoc-Arg(Cbz)$_2$-αKA-D-Phe-OH **400** via α-ketocyanophosphorane **358**. Reagents and conditions: (a) O_3, H-D-Phe-O^{t}Bu (1 equiv.), DCM, -78 °C, 22% (poor purity); (b) TFA:DCM (1:1 v/v), rt, 72 h, unisolated.

Upon addition of H-D-Phe-O^{t}Bu **398** the α-keto acyl cyanide **397** was successfully converted to α-ketoamide **399**. The product could not be sufficiently purified and was telescoped to deprotection of the *tert*-butyl ester using TFA:DCM (1:1 v/v); however, the synthesis was ultimately unsuccessful.

Having experienced problems with each reaction conducted, it was evident that neither the combination of Fmoc/Cbz or Fmoc/Pbf groups was optimal. Therefore the final α-ketocyanophosphorane Boc-Arg(Cbz)$_2$-COC(PPh$_3$)CN **353** was investigated for the synthesis of intermediate Boc-Arg(Cbz)$_2$-αKA-D-Phe-O^{t}Bu **403** (Scheme 91). Using H-D-Phe-O^{t}Bu **398** was reasoned to be the most efficient strategy as the Boc and *tert*-butyl ester could be deprotected simultaneously, thus reducing the number of steps required to convert the Boc group to an Fmoc group at a later stage (*vide infra)*.



Scheme 91: Synthesis of Boc-Arg(Cbz)$_2$-αKA-D-Phe-O^{t}Bu **403** via α-ketocyanophosphorane **353**. Reagents and conditions: (a) O_3, H-D-Phe-O^{t}Bu (1 equiv.), DCM, -78-25 °C, 18 h, 43%.

Boc-Arg(Cbz)$_2$-COC(PPh$_3$)CN **353** was observed to be more stable to ozonolysis than its alternatively protected analogues. Some degradation was still observed, though to a lesser extent in comparison to the previous attempts. Importantly, α-ketoacid **404** was now observed to be the major product, formed by quenching of the acyl cyanide intermediate **402** with water during LC-MS analysis. Addition of H-D-Phe-O^{t}Bu **398** resulted in conversion of the acyl cyanide intermediate to the desired product Boc-Arg(Cbz)$_2$-D-Phe-O^{t}Bu **403** which proved stable to isolation by column chromatography. However, it was noted that removal of the purification solvents must be conducted at ambient temperature as the product was observed to be unstable if heated to 30 °C.

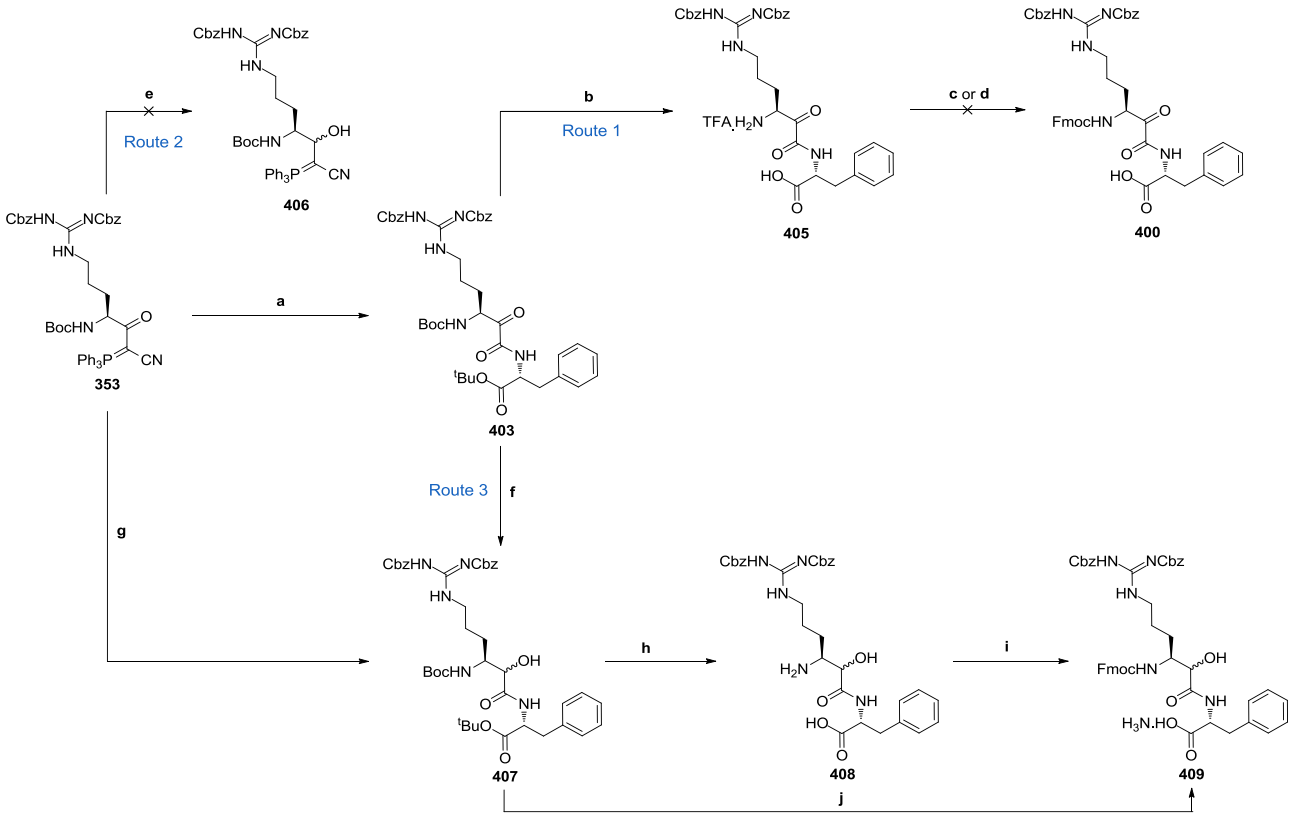
The results for the synthesis of α-ketoamide dipeptides from protected arginine derivatives are summarised in Table 15.

Protecting groups			Yield of α-keto cyanophosphorane	Yield of α-ketoamide	Conclusion
R^1	R^2	R^3			
Fmoc	Pbf, H	Allyl	**385** 82%	**394** 0%	Unstable to ozonolysis
Fmoc	Boc, Boc	Allyl	**386** 5%	- -	Minor product
Fmoc	Cbz, Cbz	^{t}Bu	**358** 5%	**399** 22%	Degradation
Boc	Cbz, Cbz	^{t}Bu	**353** 95%	**403** 43%	Stable (at room temp.)

Table 15: Results from the investigation of various protecting groups for the synthesis of α-ketoamide dipeptides via α-ketocyanophosphoranes.

Of the four protecting group combinations only two gave suitable yields from the initial formation of the α-ketocyanophosphorane. The stability of the various protecting groups to ozonolysis conditions followed the order: Boc/Cbz > Fmoc/Cbz > Fmoc/Pbf. The synthesis of stable α-ketoamide analogue **403** was encouraging for the formation of a dipeptide building block; however, to be applicable to solid phase synthesis a Boc to Fmoc exchange and deprotection of the *C*-terminus would have to be conducted (Scheme 92, Route 1).

Simultaneous deprotection of the Boc/*tert*-butyl ester protecting groups was achieved with TFA in DCM; however, the reaction profile was poor, as determined by LC-MS, with several close running impurities precluding isolation of the product. The material was therefore taken forward in crude form. Attempts to reprotect the amine with an Fmoc group failed under multiple conditions. The desired product **400** was not observed by LC-MS and all attempts at column chromatography resulted in further degradation. Exposure of the free amino intermediate **405** to a solution of sodium carbonate in 1,4-dioxane:H_2O in the absence of an Fmoc reagent also resulted in degradation, indicating that the synthesis of the building block was unlikely to be successful via this route.



Scheme 92: Attempted synthesis of Fmoc-Arg(Cbz)$_2$-αKA-D-Phe-OH **400** and synthesis of Fmoc-Arg(Cbz)$_2$-αHA-D-Phe-OH.NH$_3$ **405** from Boc-Arg(Cbz)$_2$-CO(PPh$_3$)CN **353**. Reagents and conditions: (a) O_3, H-D-Phe-O^{t}Bu **398** (1 equiv.), DCM, -78-25 °C, 18 h, 43%; (b) TFA:DCM:H_2O (1.5:1.0:0.1 v/v), rt, 16 h, crude; (c) Fmoc-Cl (1.3 equiv.), DIPEA (4.5 equiv.), DMF, 0-25 °C, 1 h, unisolated; (d) Fmoc-OSu (1.1 equiv.), Na_2CO_3 (2.5 equiv.), 1,4-dioxane:H_2O (1:1 v/v), rt, 16 h, unisolated; (e) Multiple reducing conditions; (f) $NaBH_4$ (1 equiv.), MeOH:DCM (4:1 v/v), 0-25 °C, 2.5 h, 92%; (g) O_3, H-D-Phe-O^{t}Bu **398** (1.05 equiv.), DCM, -78-25 °C, 18 h, then $NaBH_4$ (1 equiv.), DCM, 0 °C, 2 h, 49% (two steps); (h) TFA:DCM (6:94 v/v), rt, 4 h, then 4M HCl/dioxane, rt, 51 h, 10%; (i) Fmoc-OSu (1 equiv.), DIPEA (5 equiv.), 1,4-dioxane, rt, 5.5 h, 19%; (j) 4M HCl/dioxane, rt, 5 h, crude; then Fmoc-OSu (1 equiv.), DIPEA (5 equiv.), 1,4-dioxane, rt, unisolated.

It was evident from the various reactions conducted thus far that α-ketoamide dipeptides have some inherent instability. Furthermore, it was noted by Aitken *et. al.* that attempting to synthesise a linear pentapeptide precursor of cyclotheonamide also resulted in some instability issues when an α-ketoamide was present.[204] As discussed previously, there have been several successful syntheses of the Ct family using an α-hydroxyamide intermediate. While the use of this functionality would require additional steps for the overall synthesis, it could potentially stabilise the dipeptide building blocks which would be beneficial in enabling a flexible route for SAR studies.

It was reasoned that reduction of the α-keto group in α-ketocyanophosphorane **353** to give α-hydroxycyanophosphorane **406** could increase the yield of the ozonolysis step by forming the presumably more stable α-hydroxy acyl cyanide instead of the α-keto acyl cyanide (Scheme 92, Route 2). Unfortunately, attempts at reducing the α-keto group failed with a range of common reducing agents; sodium borohydride, borane and diisobutylaluminium hydride proved to be unreactive, yielding only starting material, whilst lithium aluminium hydride resulted in degradation of the starting material.

Instead, α-hydroxyamide dipeptide **407** was prepared from α-ketoamide dipeptide **403** (Scheme 92, Route 3). Reduction of the α-keto group was achieved with $NaBH_4$ to give two diastereoisomers of dipeptide **407** in an approximate ratio of 1:1 and 92% yield without need for purification. Isolation at 30 °C did not result in degradation of the product, suggesting that this moiety did indeed afford greater stability than the α-ketoamide. Product **407** was also isolated in an improved yield of 49% over two steps via a one-pot ozonolysis and reduction reaction.

With this optimised route in place, the remaining steps of the synthesis were investigated. The simultaneous deprotection of the *tert*-butyl ester and Boc group was first attempted in dilute TFA and the reaction profile was significantly improved, presumably due to removal of the α-ketoamide.

Unexpectedly, these conditions resulted in significant chemoselective deprotection of a single Cbz group, as indicated by an LC-MS peak with the corresponding mass ($[M+H]^+$ = 642.5). In contrast to the majority of previous cyclotheonamide syntheses, this indicated that at least one of the Cbz groups was unstable in acidic conditions. An alternative deprotection of the Boc group with 4M HCl in dioxane resulted in the crude amine **408** with only 5% Cbz deprotection observed.

The final Fmoc protection also gave a significantly improved reaction profile over that observed for the α-ketoamide. Use of Fmoc-Cl resulted in protection of both *O*-Fmoc and *N*-Fmoc protection, even when applied in sub-stoichiometric amounts. However, the use of 9-fluorenylmethyl *N*-succinimidyl carbonate (Fmoc-OSu) on a small scale yielded the mono *N*-Fmoc protected α-hydroxyamide **409** in a 1:1 ratio of diastereoisomers and 19% yield. Unfortunately, it was not possible to replicate the result on a larger scale. Although the NMR spectrum of these molecules was complex, the chemoselectivity of the reaction was confirmed with a high degree of confidence.

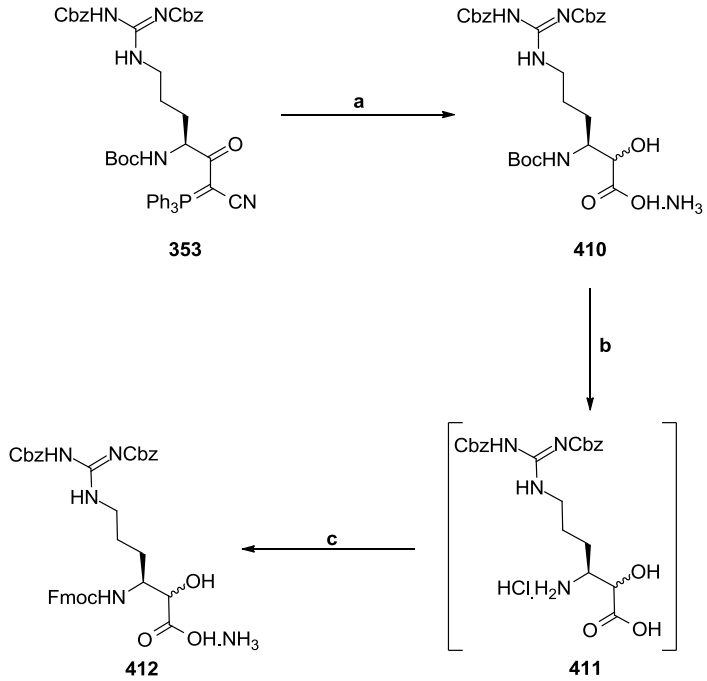
No further work was conducted on this product due to the development of an alternative route which was investigated in parallel and is discussed further in subsequent sections.

2.2.4. Synthesis of Fmoc-Arg(Cbz)$_2$-CH(OH)CO$_2$H

A survey of the literature indicated that synthesis of α-ketoacids or α-hydroxyacids was also possible from α-ketocyanophosphoranes.[212] Since this would separate the P1' and P1 residues it would be efficient for SAR studies. The learnings made during the synthesis of the dipeptide analogues discussed above were subsequently applied to this alternative route.

Based on experience with the above syntheses, it was reasoned that the α-hydroxyacid would likely exhibit greater stability than an α-ketoacid. Therefore the synthesis of monomeric Fmoc-Arg(Cbz)$_2$-CH(OH)CO$_2$H.NH$_3$ **412** from

α-ketocyanophosphorane **353** was proposed as an alternative approach (Scheme 93).

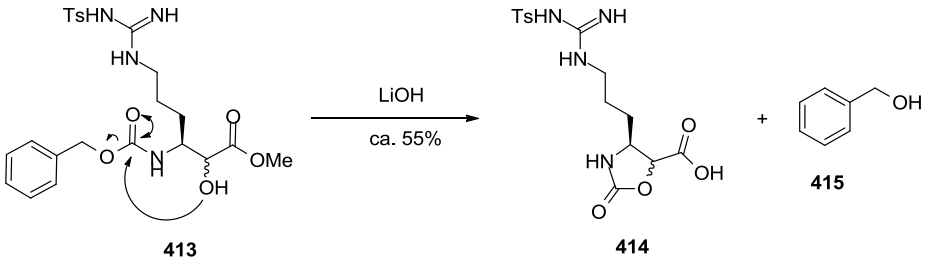


Scheme 93: Synthesis of Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **412** from α-ketocyanophosphorane **353**. Reagents and conditions: (a) O_3, DCM, -78°C, 10 min; then, THF:H_2O (9:1 v/v), -78-0 °C, 30 min; then, $NaBH_4$ (1.2 equiv.), 0-25 °C, 45 min; column chromatography with $(NH_4)_2CO_3$ modifier, 76% (two steps); (b) 4M HCl/dioxane, rt, 1 h, used *in situ*; (c) Na_2CO_3 to pH 10, then Fmoc-OSu (1 equiv.), various solvents, 0 °C, 2 h; column chromatography with $(NH_4)_2CO_3$ modifier, 16%.

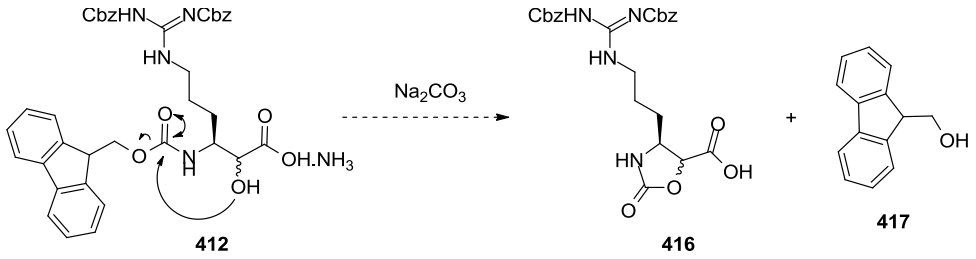
The ozonolysis of α-ketocyanophosphorane **353** was conducted as previously described except the reaction was now quenched with THF:H_2O (9:1 v/v) to form the α-ketoacid. Again, this could be reduced *in situ* with $NaBH_4$ to give diastereoisomers of α-hydroxyacid **410** in an approximate 1:1 ratio. Applying the learnings from the synthesis of the α-ketoamide dipeptide allowed the Boc deprotection of **410** with 4M HCl in dioxane to proceed without additional loss of the acid labile Cbz group. The free amine **411** formed *in situ* was used immediately for Fmoc protection. Initial attempts to synthesise Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **412** by quenching with sodium carbonate and Fmoc-OSu were met with incomplete reactions and formation of an unknown byproduct with no clearly identifiable mass by LC-MS. The product could only be obtained in 16% yield after multiple attempts.

A brief literature survey showed that, in the presence of aqueous lithium hydroxide, an intramolecular cyclisation of the α-hydroxy group could occur with a CBz group to give oxazolidinone **414** (Scheme 94a).[196] Based on this, it was predicted that an analogous reaction could be occurring with the Fmoc protected monomer **412** to give oxazolinone **416** (Scheme 94b).

a)



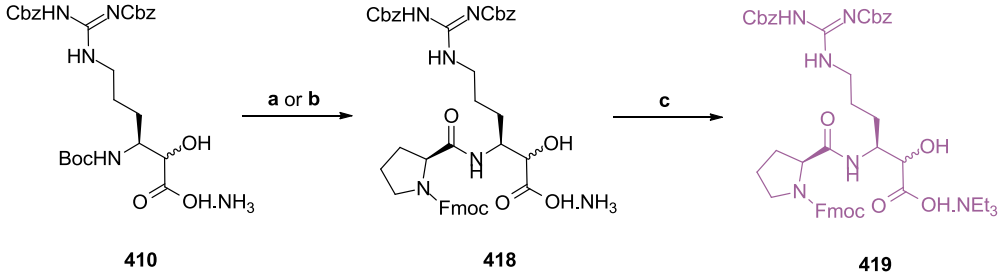
b)



Scheme 94: (a) Reported side reaction observed for Cbz-Arg(Ts)-$CHOHCO_2Me$ **413** in the presence of lithium hydroxide and (b) predicted side reaction for Fmoc-$Arg(Cbz)_2$-$CHOHCO_2H.NH_3$ **412** in the presence of sodium carbonate.

A number of conditions were examined to try and reduce the hypothesised cyclisation. These included a range of solvents in a 1:1 mixture with water, combined with either sodium hydrogen carbonate or sodium carbonate, at either 0 °C or room temperature. The ratio of byproduct to product varied little under all conditions tested. Eventually, a commercially available sample of fluorenyl methanol **417** was obtained and spiked in to the reaction mixtures. While the NMR and LC-MS peaks all closely matched and almost overlapped with the byproduct peaks it was evident that the byproduct was in fact not fluorenyl methanol but was likely to be a related species.

Fortuitously, an analogous reaction to synthesise Fmoc-Pro-Arg(Cbz)$_2$-CH(OH)CO$_2$H.NH$_3$ **418** revealed the nature of the side reaction (Scheme 95). Owing to the possibility of intramolecular cyclisation occurring with an Fmoc group in close proximity to the α-hydroxy group, it had been deemed prudent to investigate the synthesis of a building block where an amide replaced the carbamate moiety. Additionally, this allowed direct access to a dipeptide fragment for the synthesis of cyclotheonamide analogues, without requiring further manipulation of protecting groups for the *N*-terminus of the arginine derivative.

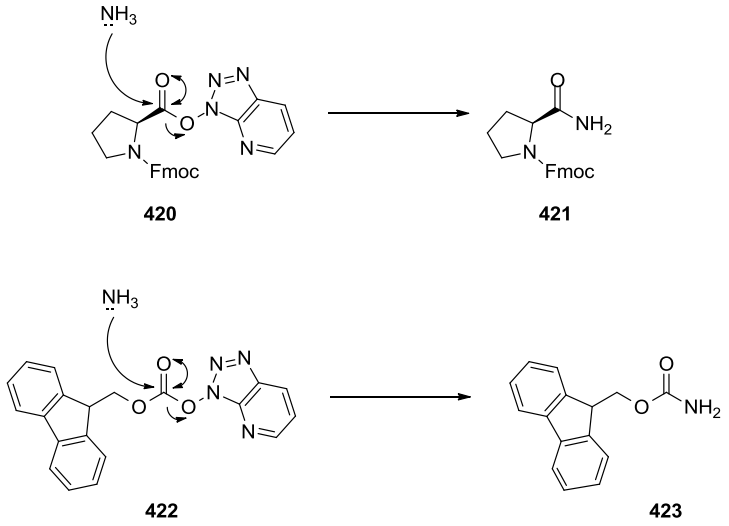


410 **418** **419**

Scheme 95: Synthesis of Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **419** as an alternative building block to Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **412**. (a) i. 4M HCl/dioxane, rt, 1.5 h, used *in situ*; ii. DIPEA (2.6 equiv.), HATU (1.2 equiv.), Fmoc-Pro-OH (1.2 equiv.), DMF:1,4-dioxane (3:1 v/v), rt, 4.5 h; column chromatography with $(NH_4)_2CO_3$ modifier, 22% (two steps); (b) i. 5% citric acid, lyophilisation; ii. 4M HCl/dioxane, rt, 1 h, used *in situ*; iii. DIPEA (2.6 equiv.), HATU (1.3 equiv.), HOAt (1.3 equiv.) Fmoc-Pro-OH (1.3 equiv.), DMF:1,4-dioxane (5:3 v/v), rt, 3.5 h; column chromatography with $(NH_4)_2CO_3$ modifier, 66% (two steps); (c) aminopropyl column, 5% Et_3N-MeOH, quantitative.

The initial Boc deprotection was followed by quenching with DIPEA and addition of a slight excess of the pre-activated HOAt ester of Fmoc-Pro-OH in DMF. Similarly to the attempted Fmoc protection, only a low yield of 22% could be obtained. The reaction profile observed was similar to that of the reaction with Fmoc-OSu but with one key difference; the byproduct peak observed for this reaction had a clearly identifiable LC-MS mass of $[M+H]^+$ = 337.2 which was consistent with Fmoc-Pro-NH$_2$ **412** (Scheme 96). Surprisingly, this evidence indicated that, despite an acidic deprotection taking place, the residual ammonia present in Boc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **410** was likely quenching the activated HOAt ester **420** of Fmoc-Pro-OH **366**. This was confirmed by repeating the initial protecting group

exchange with α-hydroxyacid **410** using Fmoc-OSu and spiking in a commercially obtained sample of fluorenyl carbamate **423** which was confirmed to be the byproduct by both LC-MS and NMR analysis.



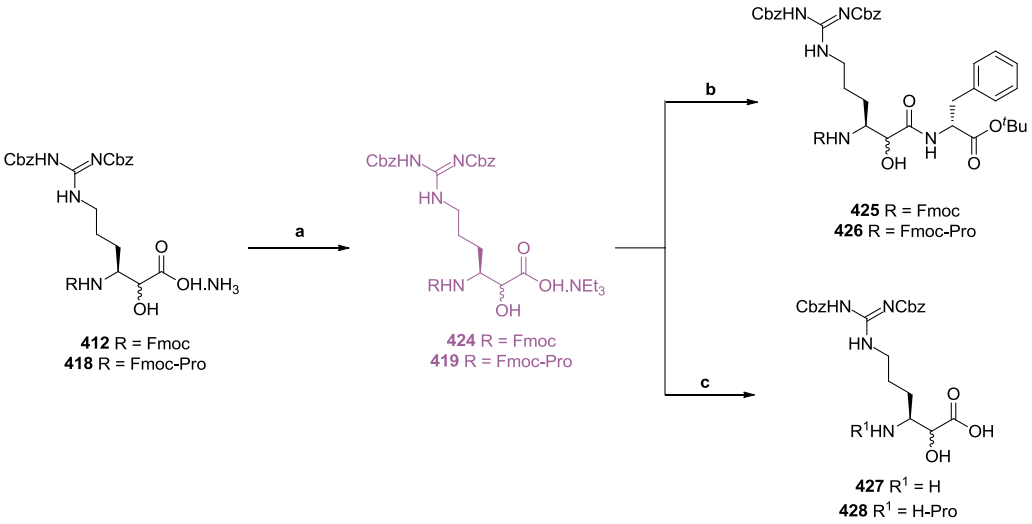
Scheme 96: Quenching of the HOAt activated esters of Fmoc-OSu and Fmoc-Pro-OH by residual ammonia present in Boc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **410**.

The remainder of the starting material Boc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **410** was washed with 5% citric acid, redissolved in 1,4-dioxane, and lyophilised twice to remove the excess ammonia. The dipeptide synthesis was then repeated with optimised conditions to give the desired dipeptide Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **418** as the ammonia salt in 66% yield. Careful addition of a slight excess of reagents was required to ensure that double addition to the α-hydroxy group was minimised; however, a small amount (*ca.* 8%) of double addition was still observed by LC-MS.

Attempting to lyophilise Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **418** to remove the ammonia prior to the solid phase synthesis resulted in partial degradation. An alternative method of removing the ammonia salt was required. Gratifyingly, an aminopropyl column could be used for catch-and-release of the acidic dipeptide. The basic amino groups on the modified silica sequestered the acid, then the ammonia washed away with methanol and the desired product eluted with a

solution of 1-10% triethylamine in methanol to afford the triethylamine salt. The non-nucleophilic nature of triethylamine allowed subsequent reactions to proceed without quenching of the activated ester. However, residual triethylamine was enough to cause partial Fmoc deprotection if stored for a period of two days; therefore, immediate use of the building block after the salt exchange was optimal.

This method of activating the building block was investigated to test the stability of both the monomer Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **424** and the dipeptide Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **419** in the two reactions that would take place on the solid phase resin (Scheme 97). Firstly, the *C*-terminus amide coupling of the dipeptide H-D-Phe-O^{t}Bu **398** was tested. Secondly, the deprotection of the Fmoc group using 20% piperidine in DMF was also tested. These reactions were conducted on an LC-MS scale to conserve stocks of the starting materials.

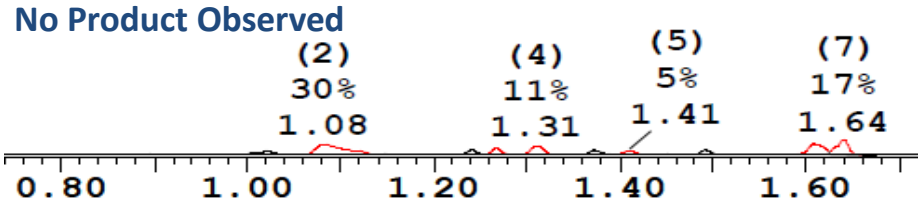


Scheme 97: LC-MS scale screening reactions used to assess the reactions that will be conducted with these building blocks during SPPS. (a) aminopropyl column, 10% Et$_3$N-MeOH, quantitative; (b) HATU (1.3 equiv.), DIPEA (1.9 equiv.), H-D-Phe-CO^{t}Bu (1.9 equiv.), DMF, rt, 1.5 h, unisolated; (c) 20% piperidine/DMF, rt, 10 min, unisolated.

The monomer Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **424** and dipeptide Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **419** were pre-activated with a solution of HATU and

DIPEA and then H-D-Phe-O^{t}Bu **398** added. Observation of tripeptides **425** and **426** was expected. LC-MS analysis of the results is depicted in Figure 75.

(a)



(b)

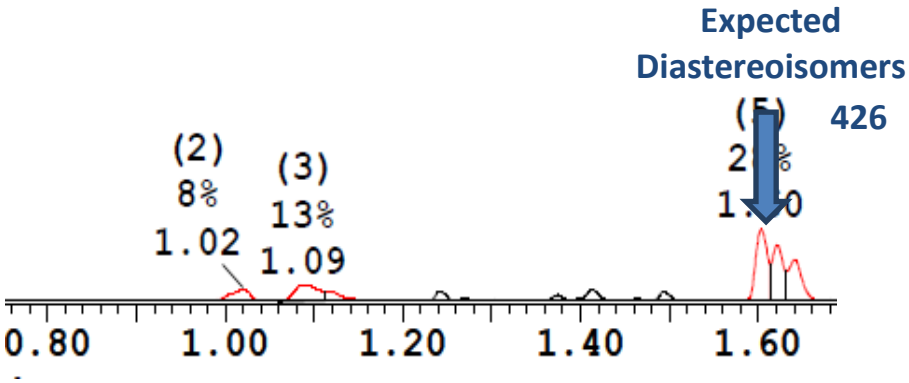


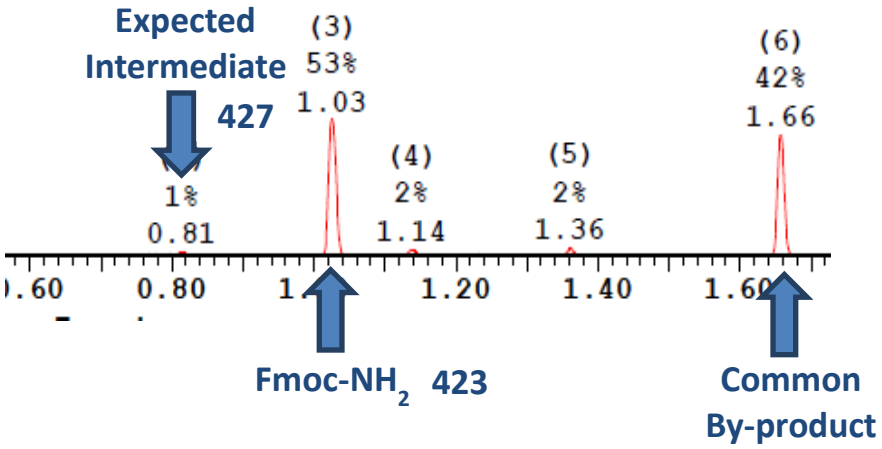
Figure 75: LC-MS spectra from the amide coupling of (a) monomer Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **424** and (b) dipeptide Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **419**.

The amide coupling conducted on monomer **424** resulted in no evidence of the formation of the expected tripeptide product **425** (Figure 75a). The reaction profile was poor with a number of byproducts observed. In contrast, the amide coupling conducted on the dipeptide **419** gave significantly more promising results (Figure 75b). Tripeptide **426** was observed as a pair of diastereoisomers with $[M+H]^+$ = 995.8. An additional impurity was observed running closely to the products but could not be identified by LC-MS. It was anticipated that optimisation of the reaction conditions during solid phase synthesis could improve the reaction profile.

Having examined compatibility for *C*-terminal amide couplings, the building blocks were next tested under *N*-terminal deprotection conditions. The monomer

Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **424** and the dipeptide Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **419** were exposed to a solution of 20% piperidine in DMF in order to monitor the Fmoc deprotection reaction. Observation of the respective amine intermediates **427** and **428** was expected. LC-MS analysis of these results is depicted in Figure 76.

a)



b)

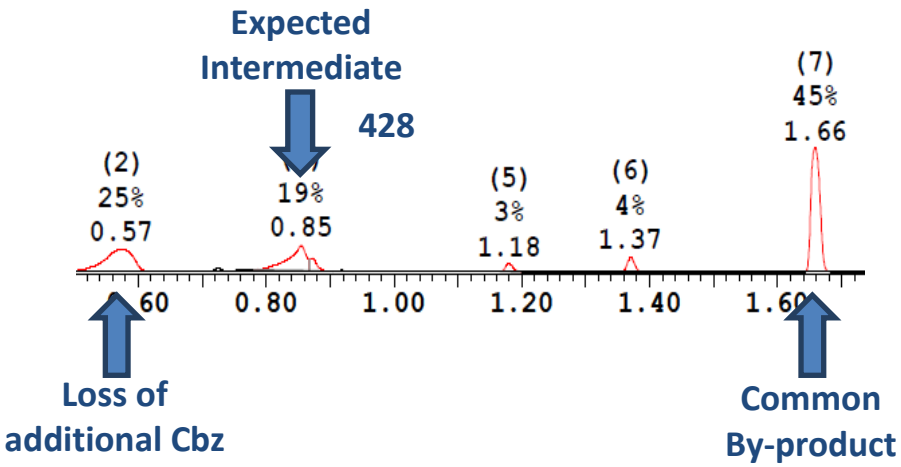


Figure 76: LC-MS spectra from the Fmoc deprotection of (a) monomer Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **424** and (b) dipeptide Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **419**.

The deprotection of the monomer did not result in significant amounts of the expected amine intermediate **427** (Figure 76a). A peak at 1.66 min was commonly observed for both deprotections which could possibly be the piperidine adduct of

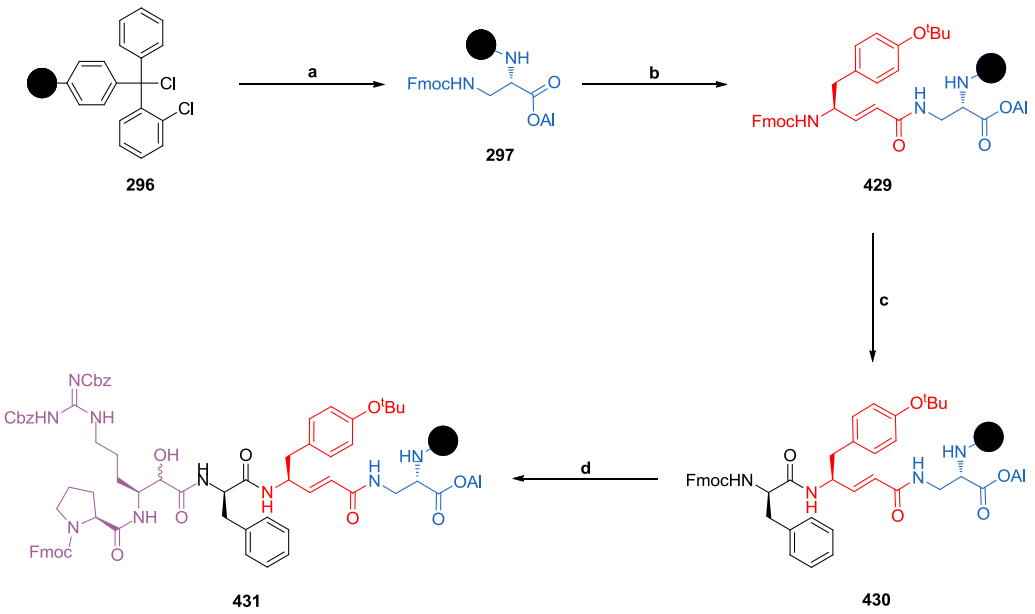
the fulvene byproduct. The peak at 1.03 min is consistent with that observed for fluorenyl carbamate and likely indicated insufficient removal of the ammonia on the aminopropyl column on this occasion; however, this is unlikely to have affected the deprotection reaction. In contrast, Figure 76b shows that the deprotection of the dipeptide building block **419** to give amine intermediate **428** was successful. Additionally, it highlighted the potential for unwanted deprotection of a Cbz group; a factor which would require further investigation (*vide infra*). Evidently, the dipeptide building block was more stable to the conditions required for solid phase synthesis.

In both the Fmoc deprotection and the amide coupling reactions the monomer **424** showed poor reaction profiles with little or no evidence of the desired products. Gratifyingly, dipeptide **419** showed more encouraging reaction profiles by LC-MS with the expected products observed for both reactions. These data clearly indicated that the dipeptide building block **419** was the preferred choice for the solid phase synthesis of CtB **292**. With the successful synthesis of the three requisite building blocks in hand, the solid phase synthesis of the target peptide was then examined.

2.2.5. Solid phase assembly: Synthesis of CtB

Although it is possible to automate SPPS, the solid phase synthesis of CtB **292** was carried out manually in order to minimise reagent use and allow for analysis at each stage. The first stage of the synthesis required the assembly of the resin attached linear pentapeptide chain **431** (Scheme 98).

The first step of the solid phase synthesis was addition of the first building block HCl.H-Dap(Fmoc)-OAl **365** to 2-chlorotrityl chloride resin **296** by nucleophilic substitution of the tertiary C-Cl bond in the presence of DIPEA, followed by capping of any unreacted bonds with excess methanol. This reaction yielded the Dap-loaded resin **297** with a loading of 0.36 mmol g^{-1}, as determined by Fmoc loading analysis.

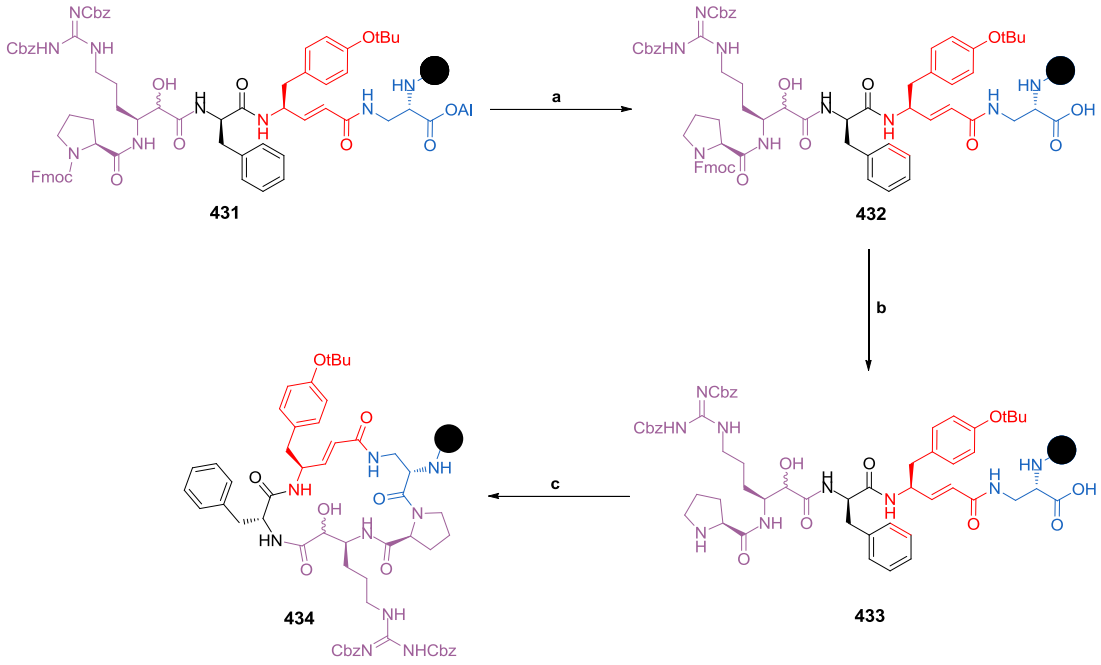


Scheme 98: Synthesis of linear pentapeptide **431** using solid phase synthesis with 2-chlorotrityl resin **296**. Reagents and conditions: (a) i. 2-chlorotrityl resin **296** (1.7 mmol g^{-1}), H-Dap(Fmoc)-OAl **365** (2 equiv.), DIPEA (4 equiv.), DCM, rt, 3 h; ii. DIPEA (2 equiv.), MeOH:DCM (1:4 v/v), rt, 1 h, 0.36 mmol g^{-1} loading; (b) i. 20% piperidine in DMF (v/v), rt, 2 x 10 min; ii. Fmoc-vTyr(^{t}Bu)-OH **364**/HATU/HOAt/DIPEA (1.5/1.5/1.5/1.5 equiv.), DMF, rt, 2 h; (c) i. 20% piperidine in DMF (v/v), rt, 2 x 10 min; ii. Fmoc-D-Phe-OH/HATU/HOAt/DIPEA (4/4/4/4 equiv.), DMF, rt, 2 h; (d) i. 20% piperidine in DMF (v/v), rt, 2 x 10 min; ii. Fmoc-Pro-Arg(Cbz)$_2$-CH(OH)CO2H.NH$_3$ **418**, aminopropyl column, 5% Et_3N-MeOH, crude; iii. HATU/HOAt/DIPEA (1.5/1.5/1.5 equiv.), DMF, rt, 2 h.

Subsequent chain extensions were conducted on a 0.2 mmol scale using alternating Fmoc deprotection with 20% piperidine in DMF and amide couplings of the amino acid building blocks after pre-activation with a combination of HATU, HOAt and DIPEA. In this way, the sequential addition of Fmoc-vTyr(^{t}Bu)-OH **364**, Fmoc-D-Phe-OH, and Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **419** (freshly prepared from Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **418**) was completed successfully. The completion of the reactions was first confirmed using a standard chloranil test[213] followed by a small-scale cleavage for analysis by LC-MS. The chloranil test detects the presence of the primary amines in any unreacted peptide on the resin as indicated by a colour change from yellow to blue/green beads. If no colour change is observed then the reaction is deemed complete.

Synthesis of **429**, **430**, and **431** proceeded efficiently with an approximate 2:1 ratio of diastereoisomers of the pentapeptide observed. Notably, no major byproducts were observed during the coupling of the dipeptide fragment despite having been observed during the test reactions conducted previously (Section 2.2.4). Furthermore, no reaction with the α-hydroxyl group was observed. Overall, the solid phase synthesis of the linear pentapeptide chain proceeded smoothly with only minimal byproduct formation.

With the linear chain assembled, the next stage of the synthesis required the deprotection of both the *N*-termini and *C*-termini of the linear peptide and cyclisation to give the resin bound macrocyclic pentapeptide **434** (Scheme 99).



Scheme 99: Solid phase macrocyclisation of linear pentapeptide **431** to give cyclic pentapeptide **434**. Reagents and conditions: (a) $Pd(PPh_3)_4$ (0.4 equiv.), $PhSiH_3$ (24 equiv.), DCM, rt, 5 x 10 min; (b) 20% piperidine/DMF, rt, 2 x 45 sec; (c) PyBOP/HOAt/DIPEA (3/3/6 equiv.), DMF, rt, 2 h.

The deprotection of the *C*-terminus was carried out with a sub-stoichiometric amount of palladium tetra(triphenylphosphine) and a phenyl silane scavenger. The deprotection proceeded efficiently and was followed immediately by the deprotection of the *N*-terminus using 20% piperidine in DMF. As can be seen from

Table 16, careful control of the reaction time for this latter reaction was found to be essential for achieving an effective deprotection.

Entry	Time/sec	Repetitions	Conversion Observed by LC-MS/ %		
			Starting Material	Fmoc Deprotection	Cbz Deprotection
1	300	2	0	54	12
2	30	2	5	60	0
3	45	2	0	70	0

Table 16: Optimisation for the Fmoc deprotection of pentapeptide **432** on solid phase resin.

Under standard deprotection conditions (20% piperidine in DMF, 2 x 5 min) the Fmoc removal was successful; however, an additional 12% mono Cbz deprotection was observed. When the reaction time was reduced to 30 seconds per round the conversion to the free amine was increased by 6% with no loss of Cbz. This reaction failed to proceed to completion, an issue which was easily solved by increasing the reaction time to 45 seconds per round, giving 70% conversion to the desired product. This optimisation exemplifies the advantages of solid phase peptide synthesis since such specific control of the reaction time would have been challenging in solution phase. The use of SPPS allowed for much more accurate control of the reaction time, thereby minimising byproduct formation.

It should be noted that, on reversing the order of the *C*- and *N*- terminal deprotection steps, the palladium catalysed deprotection of the allyl group gave no conversion to the desired product, presumably due to coordination of the amine and α-hydroxy group to the palladium centre, rendering it inactive.

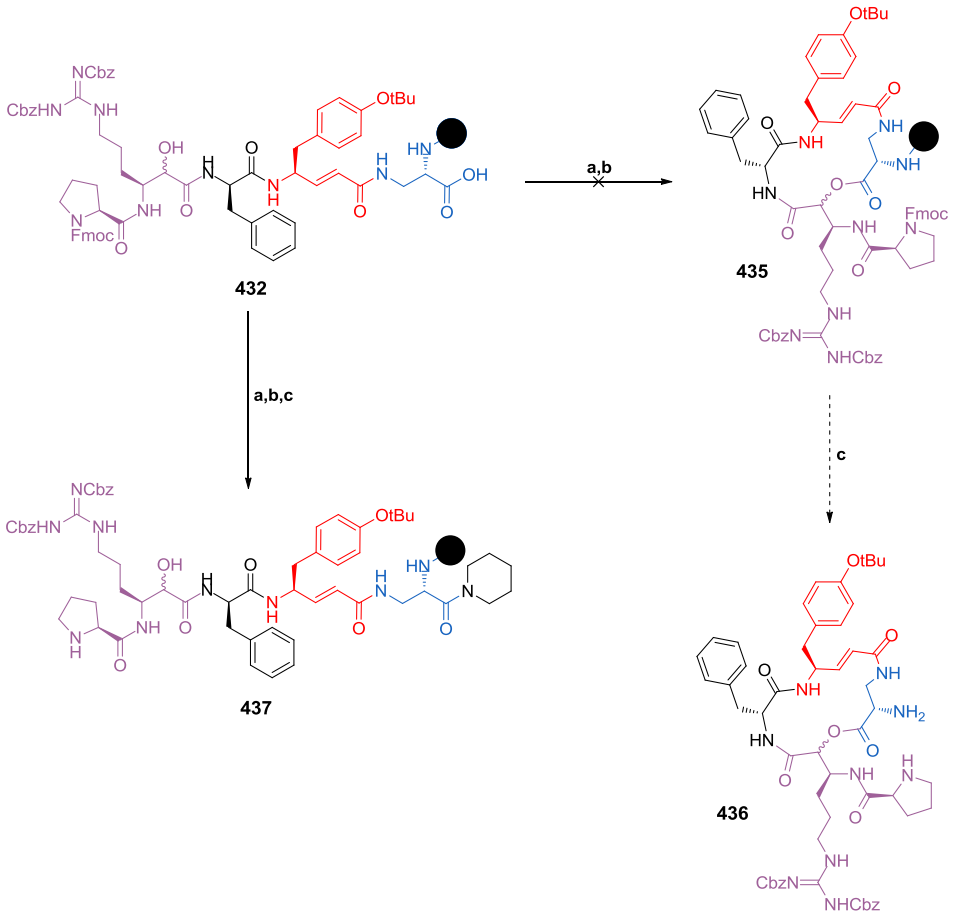
With both termini deprotected, the remaining step in this sequence was the cyclisation with PyBOP and HOAt in the presence of DIPEA. PyBOP was selected for the cyclisation due to the precedence for cyclisation with this reagent between these residues.[193] As expected, the cyclisation proceeded to completion to give peaks corresponding to the diastereoisomeric macrocylic peptides with

$[M+H]^+$ = 1030.8. A single significant byproduct was observed with $[M+H]^+$ = 1115.9 which was subsequently identified with further experimentation (*vide infra*).

While the cyclisation of the *C*- and *N*-termini was predicted to give diastereoisomers of cyclic pentapeptide **434** it was also possible that the free alcohol could undergo cyclisation to give cyclic tetrapeptide **435** (Scheme 100). Since both of these products would have identical mass by LC-MS it was important to establish if this alternative cyclisation could occur.

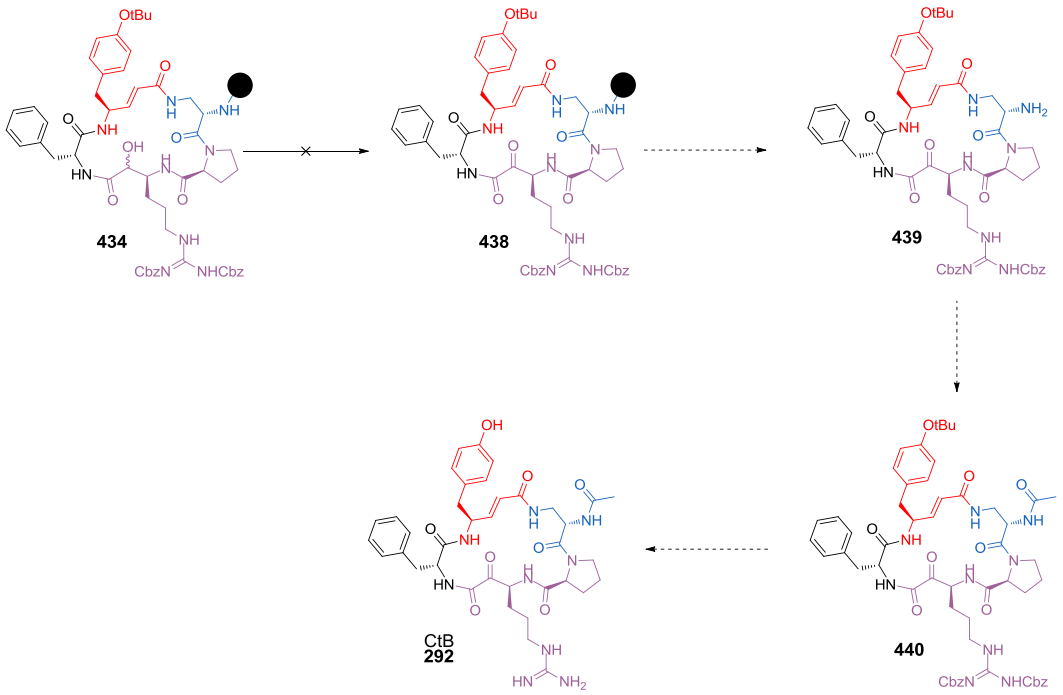
In order to determine if lactone formation could be a potential side reaction the allyl deprotection was conducted first, the *N*-terminal Fmoc group left in place, and the cyclisation conducted under identical conditions to those used for the synthesis of the diastereoisomers of cyclic peptide **434**. After two hours none of macrocyclic tetrapeptide **435** was observed by LC-MS. A pair of peaks with $[M+H]^+$ = 1367.8 were observed instead which appeared to match the expected mass of the PyBOP activated ester. This was confirmed when addition of a 20% piperidine in DMF solution followed by a small-scale cleavage failed to yield cyclic tetrapeptide **436** and instead gave two diastereoisomeric peaks corresponding to piperidine quenched linear tetrapeptide **437** with $[M+H]^+$ = 1115.8.

These data indicated that the byproduct observed during formation of cyclic pentapeptide **434** was in fact linear tetrapeptide **437**. It appears that this impurity is formed by quenching of the PyBOP activated ester with residual piperidine from the Fmoc deprotection step, despite the resin having been washed multiple times. Furthermore, since the macrolactone was not observed within a two hour period, cyclic pentapeptide **434** was confirmed to be the only product formed during the macrocylisation reaction.



Scheme 100: Monitoring for an alternative cyclisation to give macrocyclic ester **436**. Reagents and conditions: (a) PyBOP/HOAt/DIPEA (3/3/6 equiv.), DMF, rt, 2 h; (b) 20% piperidine in DMF, rt, 2 x 45 sec; (c) TFA:TIPS:DCM (1:1:98), rt, 45 min, unisolated.

With the cyclised pentapeptide successfully synthesised, the solid phase oxidation of the α-hydroxyacid was attempted on resin for the synthesis of cyclic α-ketoamide **438** (Scheme 101). All previous cyclotheonamide syntheses had utilised DMP as an oxidising agent and a brief literature search indicated that the reagent has also been applied to oxidation reactions on $PEGA_{800}$ resin.[214] This would be followed by cleavage from the resin to give free amine **439**, formation of the acetamide **440** and final global deprotection to give CtB **292**.



Scheme 101: Attempted solid phase oxidation of macrocyclic pentapeptide **434** and planned synthesis of CtB **292**.

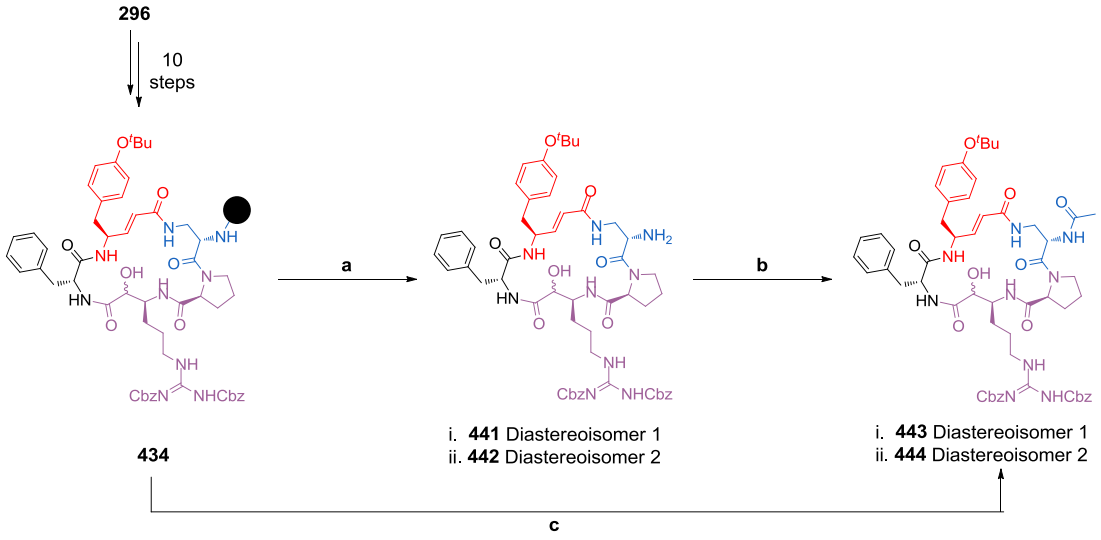
However, despite the earlier precedent, no conversion to α-ketoamide **438** was observed after a number of attempts with multiple oxidising agents and bases (Table 17). Therefore, this approach was halted.

Entry	Reagent	Equivalents	Base	Solvent	Time/ h	Observation
1	DMP	7.5	-	DMF	3	No Reaction
2	DMP	3	5% pyridine	DMF	3	No Reaction
3	DMP	7.5	-	DCM	3	No Reaction
4	DMP	3	5% pyridine	DCM	3	No Reaction
5	sIBX	7.5	-	DMSO	3	No Reaction
6	SO_3.pyridine	7.5	50% Et_3N	DMSO	3	No Reaction

Table 17: Conditions used for the attempted oxidation of cyclic peptide **434** on resin.

It has been reported that polystyrene based 2-chlorotrityl resin **296** may be incompatible with this oxidising reagent as the reactions are sluggish and acetic acid generated from DMP can result in partial resin cleavage.[215] PEG based resin such as CLEAR and $PEGA_{800}$ were suggested to be more appropriate for this class of reaction as the improved swelling properties allow for more efficient oxidation and they are not acid labile, thereby reducing unwanted resin cleavage.

With the oxidation unsuccessful on the solid phase, it was necessary to conduct this step of the synthesis at a later stage. A number of the previous cylcotheonamide syntheses reported the oxidation of the α-hydroxyamide as the penultimate step in the synthesis, immediately prior to the global deprotection. Based on this, the formation of the macrocyclic pentapeptide **434** was followed by cleavage from the resin and acetamide formation of the free amine on the Dap residue (Scheme 102).

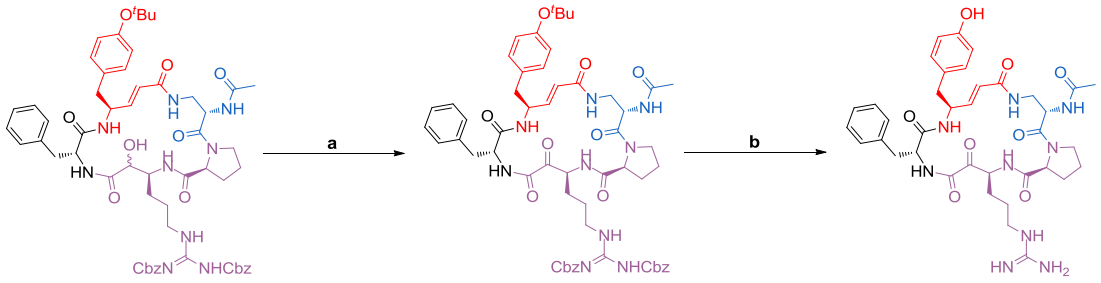


Scheme 102: Solution phase synthesis of α-hydroxy precursors to CtB **292**. Reagents and conditions: (a) TFA:TIPS:DCM (1:1:98 v/v), rt, 3 x 20 min; then pyridine, DCM, rt; i. Diastereoisomer **441**, 8%, 90% LC-MS purity (over 12 steps); ii. Diastereoisomer **442**, 9%, 78% LC-MS purity (over 12 steps); (b) pentafluorophenyl acetate (5 equiv.), DMF, rt, 3 h; i. Diastereoisomer **443**, 16%, 84% LC-MS purity; ii. Diastereoisomer **444**, 38%, 89% LC-MS purity; (c) TFA:TIPS:DCM (1:1:98 v/v), rt, 3 x 20 min; then pyridine, DCM, rt; then pentafluorophenyl acetate (6 equiv.), DCM, rt, 42 h; i. Diastereoisomer **443**, 5%, 100% LC-MS purity (over 12 steps); ii. Diastereoisomer **444**, 5%, 100% LC-MS purity (over 12 steps).

Cleavage of the 2-chlorotrityl resin was achieved orthogonally using dilute 1% TFA in DCM with a TIPS scavenger. The TFA was quenched with pyridine and the cleaved peptide purified by high pH reverse phase chromatography. During the purification diastereoisomers **441** and **442** were separated and isolated in 8% and 9% yield, respectively. Over the 11 steps conducted on solid phase this equated to an acceptable average yield of approximately 85% per step to give an overall yield of 17%. Selective acetylation of the unmasked amine has been demonstrated previously in the presence of the α-hydroxyl group using pentafluorophenyl acetate.[196] Synthesis and separation of diastereoisomers **443** and **444** was achieved with pentafluorophenyl acetate in 16% and 38% yield, respectively. This equated to an overall yield of 9.2% over the 12 steps.

The resin cleavage and acetylation reactions could also be telescoped to remove one purification step while maintaining high yields. Diastereoisomers **443** and **444** were obtained in 5% yield each, giving an overall yield of 10% over 12 steps which equated to an average of 81% yield per step. Furthermore, the purity of the acetylated products was also greater when the reactions were telescoped.

With this material in hand, the solution phase oxidation of the α-hydroxyamide and the global deprotection of the remaining acid labile protecting groups was investigated under a range of conditions (Scheme 103, Table 18).



443 Diastereoisomer 1
444 Diastereoisomer 2

440

CtB
292

Scheme 103: Solution phase oxidation and global deprotection for the synthesis of CtB **292**.

Table 18: Conditions applied to the solution phase oxidation and global deprotection of diastereoisomers **443** and **444**.

Entry	Scale/mg	Reagent	Equiv.	Solvent	Temp/°C	Time/h	Result	Deprotection	Observation
1	1.5	sIBX	4	MeCN:$CHCl_3$ (1:1 v/v)	65	4.5	No reaction	-	-
2	1.5	sIBX	20	DMSO	100	2.5	Degraded	-	-
3	7	DMP	5	MeCN	75	7	Degraded	-	-
4	1.8	DMP	5	MeCN	70	1	Crude	TFA:H_2O:Thioanisole (95:5:15 v/v)	Degraded on work-up
5	10	DMP	10	MeCN	75	2.5	Crude	TFA:H_2O:Thioanisole (95:5:15 v/v)	Degraded on work-up
6	7	DMP	4.5	MeCN	75	1.5	Crude	TFA:H_2O:Thioanisole (95:5:15 v/v)	Degraded during reaction

Initial oxidation attempts were carried out with stabilised IBX; however this was found to be unsuitable as it either failed to oxidise the hydroxyl group (Entry 1) or caused degradation (Entry 2). The oxidation of the α-hydroxyamide was successful under a range of conditions when DMP was used instead (Entries 4-6) except when extended reaction times were applied (Entry 3). The α-ketoamide was observed as an acetal formed by nucleophilic attack of water present in the analytical system with $[M+H]^+$ = 1088.4. The equivalents of DMP, temperature, and concentration had little effect on the reaction time and complete conversion was observed under all conditions.

Initially, the crude α-ketoamide **440** was taken forward for the global deprotection in TFA with water and thioanisole scavengers. Under these conditions the final product could be observed by LC-MS analysis as an acetal adduct of the α-ketoamide with $[M+H]^+$ = 764.5. However, there was evidence that the product could degrade during work-up (Entry 5 and 6) and as a result it was not possible to isolate a sample of CtB **292** using the limited amount of intermediates available.

It was evident from these attempted experiments that the stability of the final cyclotheonamide product was limited under the attempted reaction and work up conditions. In order to complete the synthesis, further optimisation of the work-up would be required. For example, an immediate MDAP purification could be conducted without isolation of the crude product, or the product could be partitioned between water and an organic solvent to extract the excess TFA, or the mixture could be diluted with water and the solvent removed by lyophilisation. In order to carry out this optimisation, significantly larger amounts of material would be required. Unfortunately, due to time constraints this was not possible. However, with further optimisation it should be possible to develop an efficient synthesis of CtB **292** using this route.

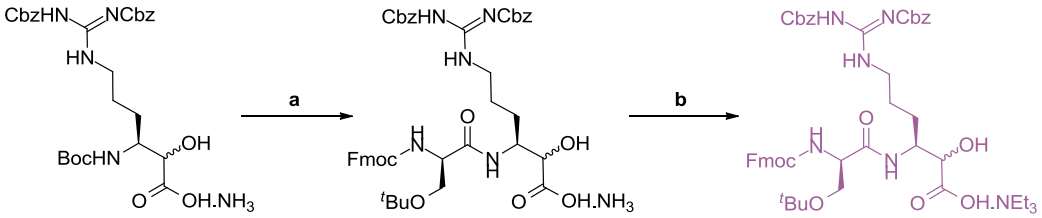
2.3. Synthesis of analogues

While the oxidation and deprotection sequence of the cyclotheonamide B precursor still required further attention, it was important to confirm that the solid phase synthesis would be suitable for the preparation of cyclotheonamide analogues.

The purpose of this research was to prepare cyclotheonamide analogues which could potentially inhibit the trypsin-like serine protease KLK5. The S1' and S2 pockets had been identified as suitable regions of the protease for further investigation of SAR with cyclotheonamide analogues. As such, the preparation of analogues with modified P1' and P2 residues was also a major goal of this synthetic campaign and was investigated in parallel with the original synthesis of CtB **292**.

The synthesis of analogues to probe the SAR of the S2 pocket required modification of the P2 residue. As discussed previously (Section 2.1.2), more polar residues in the S2 pocket could result in improved potency and selectivity for the KLK5 protease. Accordingly, modification of the P1-P2 building block would be required for this. Therefore, commercially available Fmoc-Ser(^{t}Bu)-OH was utilised to test the synthesis of cyclotheonamide analogues with an alternative P2 residue.

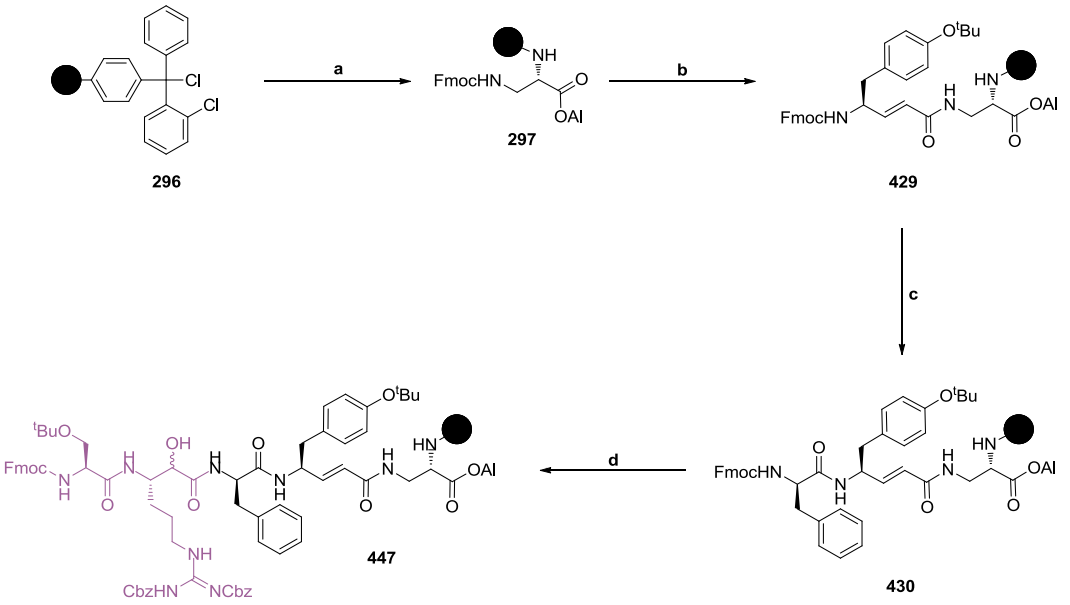
Preparation of the requisite dipeptide Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **446** from Boc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **410**, which had been prepared previously, was completed in three steps (Scheme 104).



410 445 446

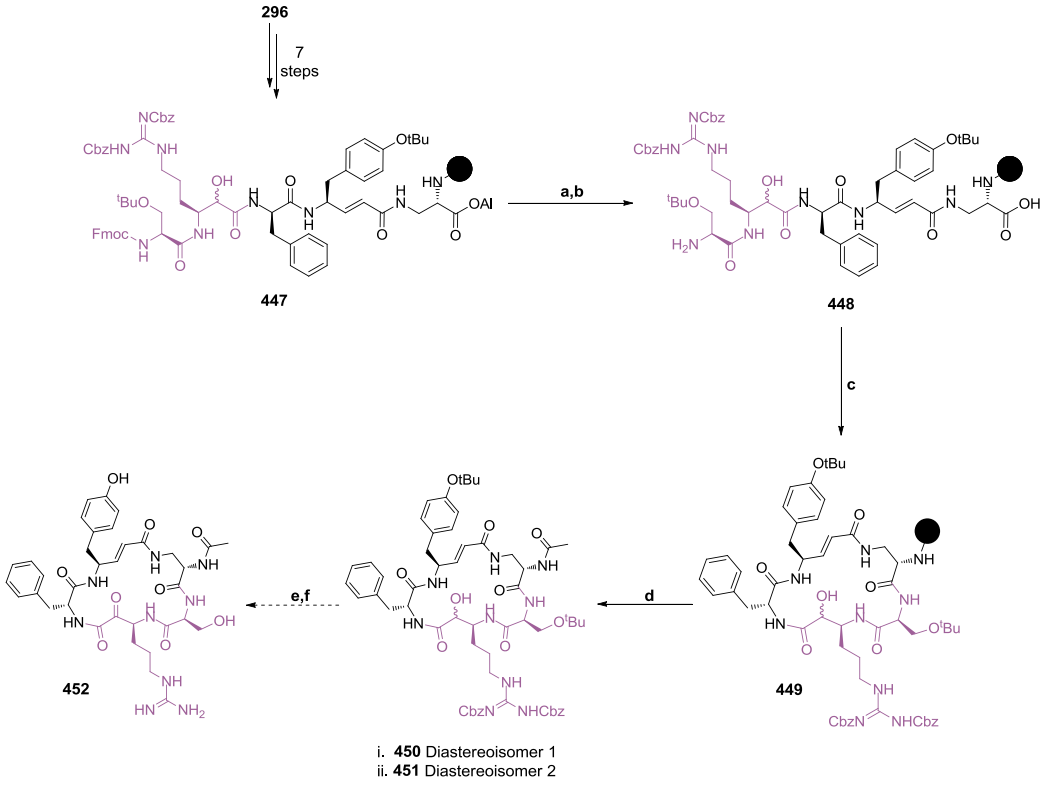
Scheme 104: Synthesis of Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **446** as a modified building block for the SAR analysis of the S2 pocket. Reagents and conditions: (a) i. 4M HCl/dioxane, rt, 1 h, used *in situ*; ii. DIPEA, HATU (1.3 equiv.), HOAt (1.3 equiv.), Fmoc-Ser(^{t}Bu)-OH (1.3 equiv.), DMF:1,4-dioxane (1:1 v/v), rt, 4.5 h; column chromatography with $(NH_4)_2CO_3$ modifier, 43%; (b) aminopropyl column, 5% Et_3N-MeOH, quantitative.

The *N*-Boc group was deprotected with 4M HCl in dioxane and the acid quenched with DIPEA. Addition of a pre-activated solution of HATU, HOAt, DIPEA, and Fmoc-Ser(^{t}Bu)-OH in DMF afforded Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ **445** after purification by high pH reverse phase chromatography. Exchange of the ammonia salt for the triethylamine salt afforded the desired building block Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ **446** in 43% yield over the two steps. Having prepared the requisite P1-P2 building block the solid phase synthesis could be repeated using the conditions established for the wild type CtB in order to synthesise the S2 pocket analogue of cyclotheonamide B (Scheme 105).



Scheme 105: Synthesis of linear pentapeptide **447** using solid phase synthesis with 2-chlorotrityl resin **296**. Reagents and conditions: (a) i. 2-chlorotrityl resin **296** (1.7 mmol g^{-1}), H-Dap(Fmoc)-OAl (2 equiv.), DIPEA (4 equiv.), DCM, rt, 3 h; ii. DIPEA (2 equiv.), MeOH:DCM (1:4 v/v), rt, 1 h, 0.36 mmol g^{-1} loading; (b) i. 20% piperidine in DMF (v/v), rt, 2 x 10 min; ii. Fmoc-vTyr-OH/HATU/HOAt/DIPEA (1.5/1.5/1.5/1.5 equiv.), DMF, rt, 2 h; (c) i. 20% piperidine in DMF (v/v), rt, 2 x 10 min; ii. Fmoc-D-Phe-OH/HATU/HOAt/DIPEA (4/4/4/4 equiv.), DMF, rt, 2 h; (d) i. 20% piperidine in DMF (v/v), rt, 2 x 10 min; ii. Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$, aminopropyl column, 5% Et$_3$N-MeOH, quantitive; iii. HATU/HOAt/DIPEA (1.5/1.5/1.5 equiv.), DMF, rt, 2 h.

Linear pentapeptide **447** was next converted to cyclic pentapeptide **449** via sequential deprotection of the *C*-terminus and *N*-terminus followed by cyclisation under amide coupling conditions (Scheme 106). The modification of the P2 residue had no deleterious effects on these steps.



Scheme 106: Synthesis of P2 residue modified α-hydroxyamide precursors of cyclotheonamide analogues. Reagents and conditions: (a) $Pd(PPh_3)_4$ (0.4 equiv.), $PhSiH_3$ (24 equiv.), DCM, rt, 5 x 10 min; (b) 20% piperidine/DMF, rt, 2 x 45 sec; (c) PyBOP/HOAt/DIPEA (3/3/6 equiv.), DMF, rt, 2 h; (d) TFA:TIPS:DCM (1:1:98 v/v), rt, 3 x 20 min; then pyridine, DCM, rt; then pentafluorophenyl acetate (3 equiv.), DMF, rt, 16.5 h; i. Diastereoisomer **450**, 9%, 100% LC-MS purity (over 12 steps); ii. Diastereoisomer **451**, 8%, 100% LC-MS purity (over 12 steps); (e) DMP, DCM, 75 °C; (f) TFA:H_2O:Thioanisole (95:5:15 v/v), rt.

The solid phase synthesis of the cyclised pentapeptide **449** was completed successfully and cleavage from the resin was required for the remaining steps. As noted during the synthesis of the cyclotheonamide B precursor the cleavage and subsequent acetamide formation could be telescoped without adversely affecting the product yields. Diastereoisomers **450** and **451** were obtained in 9% and 8%

yield, respectively, with an overall yield of 17% over 12 steps, equating to an average yield of 85% per step.

Further optimisation of the oxidation and deprotection steps used in the attempted synthesis of wild type CtB **292** were required prior to completion of this P2 modified CtB analogue **452**. Despite this, the synthesis of a precursor to cyclotheonamide analogues for the investigation of SAR of the S2 pocket has been demonstrated using SPPS. This method therefore represents a straightforward, rapid and flexible approach towards analogues of this complex class of natural products.

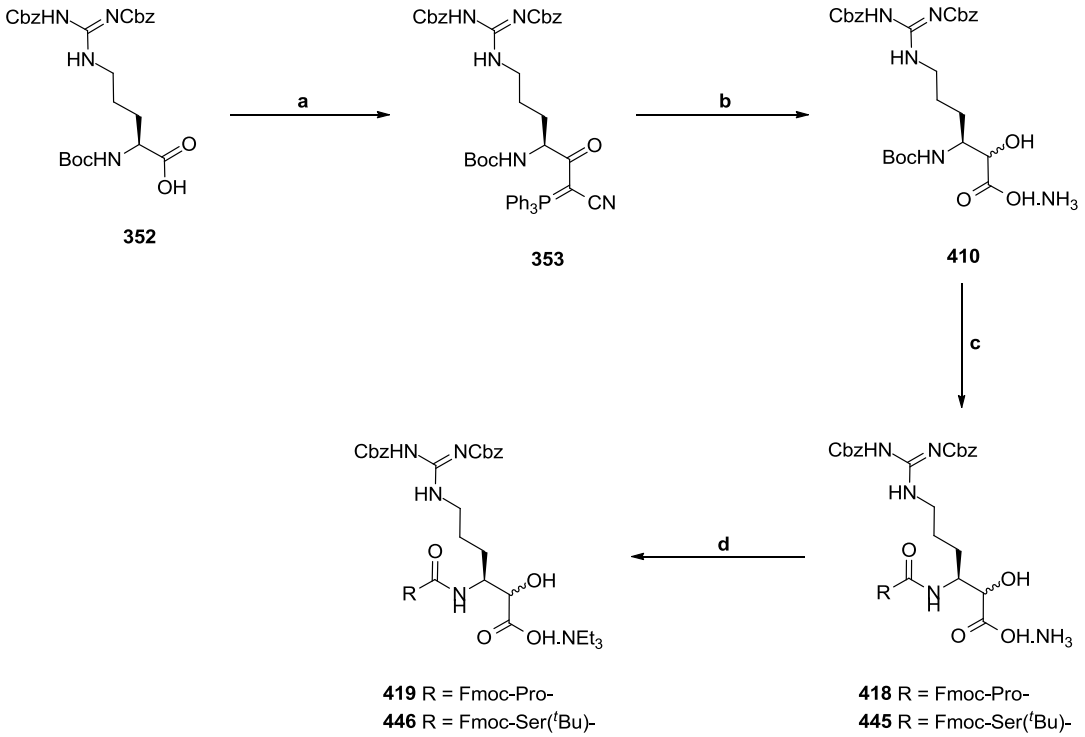
2.4. Conclusions

The cyclotheonamide family of natural products is a well established group of Trypsin-Like Serine Protease inhibitors and was of interest for conducting an SAR study with KLK5, a serine protease with indications in skin diseases such as atopic dermatitis. In order to study the SAR of analogues an efficient route to the synthesis of these complex structures was investigated.

Firstly, the synthesis of a number of building blocks for the synthesis of cyclotheonamide analogues was achieved. Of these, the synthesis of the building block for the S1 pocket was the most complex. The synthesis then oxidative activation of an α-ketocyanophosphorane moeity and substitution of the resulting acyl cyanide intermediate was investigated for the synthesis of the α-ketoamide functionality.

The combination of the P1 and P1' residues as an α-ketoamide dipeptide building block was found to be unsuitable due to inherent stability issues, while its analogous α-hydroxyamide structure had a tendency to undergo side reactions. Gratifyingly, the combination of the P1 and P2 residues was found to be suitable for the synthesis of α-hydroxyacid dipeptide building blocks which were suitable for use in SPPS (Scheme 107). The synthesis could be carried out in just six steps with generally high yields.

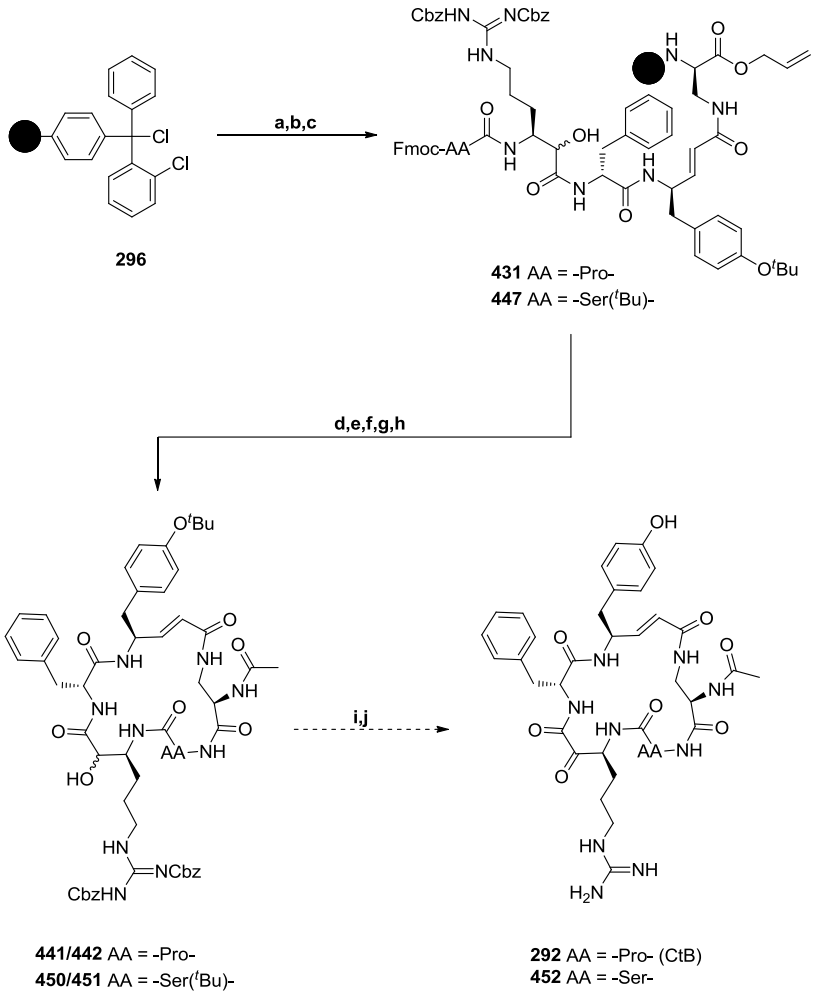
This new synthesis combined the benefits of the solid phase synthesis demonstrated by Schaske *et. al.* with a new, short, and high yielding preparation of α-hydroxyacid dipeptide building blocks for installation of the α-ketoamide warhead. The linear preparation of each building block is six steps or less and the subsequent couplings to prepare the linear pentapeptides are rapid.



Scheme 107: General synthesis of modified P1-P2 dipeptides as building blocks in the synthesis of cyclotheonamide analogues. Reagents and conditions: (a) EDC (1.3 equiv.), HOAt (1.3 equiv.), Ph_3P=CH(CN) (3 equiv.), DCM, 0-25 °C, 6.5 h, 95%; (b) O_3, DCM, -78°C, 10 min; then THF:H_2O (9:1 v/v), -78-0 °C, 30 min; then, $NaBH_4$ (1.2 equiv.), 0-25 °C, 45 min; column chromatography with $(NH_4)_2CO_3$ modifier, 76% (two steps); (c) i. 4M HCl/dioxane, rt, 1 h, used *in situ*; ii. DIPEA, HATU, HOAt, Fmoc-AA-OH, DMF:1,4-dioxane (1:1 v/v), rt; column chromatography with $(NH_4)_2CO_3$ modifier; (d) aminopropyl column, 5% Et_3N-MeOH, quantitative.

The solid phase syntheses of precursors of CtB **292** and analogue **452** were carried out successfully to give α-hydroxyamide precursors to cyclotheonamide analogues (Scheme 108).

Deprotection of the *C*- and *N*- termini were shown to be efficient, with the use of solid phase resin allowing for excellent control over the reaction timing for minimal byproduct formation, a task which would be difficult to achieve in solution phase. Cyclisation is achieved under standard amide coupling conditions with no significant byproduct formation observed. The potential macrolactonisation side reaction with the free α-hydroxyl group was demonstrated to not occur under these conditions.



Scheme 108: General SPPS of cyclotheonamide analogues. (a) **365**, DIPEA, DCM, rt; (b) MeOH, DIPEA, rt; (c) 3 x sequential **AA** (**364** then Fmoc-D-Phe-OH then **419** or **446**), HATU, HOAt, DIPEA, DMF, rt; (d) $Pd(PPh_3)_4$, $PhSiH_3$, DCM, rt; (e) 20% piperidine in DMF (1:4 v/v), rt; (f) PyBOP, HOAt, DIPEA, DMF, rt; (g) TFA:TIPS:DCM (1:1:98), rt; (h) pentafluorophenyl acetate, pyridine, DMF, rt; (i) DMP, MeCN, 60-70 °C; (j) TFA:H_2O:thioanisole (0.95:0.05:0.15 v/v).

Oxidation of the cyclotheonamide B precursor was successful under a variety of conditions and the final deprotected product was also observed by LC-MS. However, the product was found to be relatively unstable and could not be successfully isolated under the conditions attempted so far.

In conclusion, a rapid and flexible SPPS had been developed and the synthesis of advanced intermediates of cyclotheonamide analogues has been demonstrated. Unfortunately due to time constraints it was not possible to complete the final step

of the synthesis; however, with minimal further optimisation it should be possible to access a number of Ct analogues for an SAR study with KLK5. It is anticipated that these compounds and related analogues will be of value in further delineating the role of KLK5 in diseases such as Atopic Dermatitis and could potentially provide a pathway towards new treatment units for such diseases.

Chapter 3

Experimental, Appendix and References

3.1. General

All chemicals and solvents were obtained from commercial suppliers or internal sources and used without further purification unless otherwise stated. Reactions were conducted under an atmosphere of nitrogen with anhydrous solvents where stated. Products were characterised by melting point, LC-MS, IR, αD, HRMS, ^{1}H NMR and ^{13}C NMR where applicable.

3.2. Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR spectra were recorded using a Bruker DPX400, AV400 or AVIII600 (with cyroprobe) referenced to tetramethylsilane. Chemical shifts are reported in parts per million (ppm) to the nearest 0.01 ppm (^{1}H NMR) or 0.1 ppm (^{13}C). Coupling constants (*J*) are reported in Hz to the nearest 0.1 Hz. Spectra were recorded at room temperature unless otherwise stated.

3.3. Liquid Chromatography Mass Spectrometry (LC-MS)

LC conditions: The UPLC analysis was conducted on an Acquity UPLC BEH C_{18} column (50 mm x 2.1 mm, i.d. 1.7 µm packing diameter) with a flow rate of 1 mL/min at 40 °C. The UV detection was a summed signal from wavelengths of 210 nm to 350 nm.

MS conditions: Mass spectra were conducted on a Waters ZQ mass spectrometer, with an ionisation mode of alternate–scan positive and negative electrospray. The scan range was 100 to 1000 AMU, the scan time was 0.27 sec and the inter–scan delay was 0.10 sec.

- *Method A Using Formic Acid Modifier*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of formic acid in water; B = 0.1% v/v solution of formic acid in acetonitrile. The gradient (A:B) employed was from 97:3 to 3:97 over 2 min.

- *Method B Using TFA Modifier*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of trifluoroacetic acid in water; B = 0.1% v/v solution of trifluoroacetic acid in acetonitrile. The gradient (A:B) employed was from 97:3 to 3:97 over 2 min.

- *Method C Using Ammonium Bicarbonate Modifier*

LC and MS conditions as reported above. The solvents employed were: A = ammonium hydrogen carbonate in water adjusted to pH 10 with ammonia solution; B = acetonitrile. The gradient (A:B) employed was from 99:1 to 0:100 over 2 min.

- *Method D Using Formic Acid Modifier (Xselect CSH C18 column)*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of formic acid in water; B = 0.1% v/v solution of formic acid in acetonitrile. The gradient (A:B) employed was from 98:2 to 2:98 over 2 or 3 min.

- *Method E Using TFA Modifier (Xselect CSH C18 column)*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of trifluoroacetic acid in water; B = 0.1% v/v solution of trifluoroacetic acid in acetonitrile. The gradient (A:B) employed was from 98:2 to 2:98 over 2 or 3 min.

- *Method F Using Ammonium Bicarbonate Modifier (Xselect CSH C18 column)*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of ammonium hydrogen carbonate in water adjusted to pH 10 with ammonia solution; B = 0.1% v/v solution of ammonium hydrogen carbonate in acetonitrile. The gradient (A:B) employed was from 98:2 to 2:98 over 2 or 3 min.

3.4. Mass Directed Auto-Preparative HPLC (MDAP)

LC conditions: The HPLC analysis was conducted on either a Sunfire C_{18} column (100 mm x 19 mm, i.d 5 µm packing diameter) or a Sunfire C_{18} column (150 mm x 30 mm, i.d. 5 µm packing diameter) at ambient temperature. The UV detection was a summed signal from wavelength of 210 nm to 350 nm.

MS conditions: The mass spectrometry was conducted on a Waters ZQ mass spectrometer, with an ionisation mode of alternate–scan positive and negative electrospray. The scan range was 100 to 1000 AMU, the scan time was 0.50 sec and the inter–scan delay was 0.20 sec.

- *Method A Using Formic Acid Modifier*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of formic acid in water; B = 0.1% v/v solution of formic acid in acetonitrile. The purification was run as a gradient (A:B) over either 15 min or 25 min, with a flow rate of 20 mL/min (100 mm x 19 mm, i.d 5 µm packing diameter) or 40 mL/min (150 mm x 30 mm, i.d. 5 µm packing diameter).

- *Method B Using Ammonium Bicarbonate Modifier*

LC and MS conditions as reported above. The solvents employed were: A = 10 mM ammonium bicarbonate in water, adjusted to pH 10 with ammonia solution; B = acetonitrile. The purification was run as a gradient (A:B) over either 15 min or 25 min, with a flow rate of 20 mL/min (100 mm x 19 mm, i.d 5 µm packing diameter) or 40 mL/min (150 mm x 30 mm, i.d. 5 µm packing diameter).

- *Method C Using Trifluoroacetic Acid Modifier*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of trifluoroacetic acid in water; B = 0.1% v/v solution of trifluoroacetic acid in acetonitrile. The purification was run as a gradient (A:B) over either 15 min or 25 min, with a flow rate of 20 mL/min (100 mm x 19 mm, i.d 5 µm packing diameter) or 40 mL/min (150 mm x 30 mm, i.d. 5 µm packing diameter).

- *Method D using Ammonium Bicarbonate modifier (Xselect CSH C18 column)*

LC and MS conditions as reported above. The solvents employed were: A = 10 mM ammonium bicarbonate in water, adjusted to pH 10 with ammonia solution; B = acetonitrile. The purification was run as a gradient (A:B) over either 15 min or 25 min, with a flow rate of 20 mL/min (100 mm x 19 mm, i.d 5 µm packing diameter) or 40 mL/min (150 mm x 30 mm, i.d. 5 µm packing diameter).

- *Method E Using Formic Acid Modifier (Xselect CSH C18 column)*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of formic acid in water; B = 0.1% v/v solution of formic acid in acetonitrile. The purification was run as a gradient (A:B) over either 15 min or 25 min, with a flow rate of 20 mL/min (100 mm x 19 mm, i.d 5 µm packing diameter) or 40 mL/min (150 mm x 30 mm, i.d. 5 µm packing diameter).

- *Method F Using Trifluoroacetic Acid Modifier (Xselect CSH C18 column)*

LC and MS conditions as reported above. The solvents employed were: A = 0.1% v/v solution of trifluoroacetic acid in water; B = 0.1% v/v solution of trifluoroacetic acid in acetonitrile. The purification was run as a gradient (A:B) over either 15 min or 25 min, with a flow rate of 20 mL/min (100 mm x 19 mm, i.d 5 µm packing diameter) or 40 mL/min (150 mm x 30 mm, i.d. 5 µm packing diameter).

3.5. High Resolution Mass Spectrometry (HRMS)

ESI (+) high resolution mass spectra (HRMS) were obtained on a Micromass Q-T of 2 hybrid quadrupole time-of-flight mass spectrometer, equipped with a Z-spray interface, over a mass range of 100 – 1100 Da, with a scan time of 0.9 s and an interscan delay of 0.1 s. Reserpine was used as the external mass calibrant ($[M+H]^+$ = 609.2812 Da). The Q-Tof 2 mass spectrometer was operated in W reflectron mode to give a resolution (FWHM) of 16000-20000. Ionisation was achieved with a spray voltage of 3.2 kV, a cone voltage of 50 V, with cone and desolvation gas flows of 10-20 and 600 L/h, respectively. The source block and desolvation temperatures

were maintained at 120 °C and 250 °C, respectively. The elemental composition was calculated using MassLynx v4.1 for the $[M+H]^+$ and the mass error quoted as ppm.

An Agilent 1100 Liquid Chromatograph equipped with a model G1367A autosampler, a model G1312A binary pump and a HP1100 model G1315B diode array detector was used. The method used was generic for all experiments. All separations were achieved using a Phenomenex Luna C18 (2) reversed phase column (100 x 2.1 mm, 3 µm particle size). Gradient elution was carried out with the mobile phases as (A) water containing 0.1 % (v/v) formic acid and (B) acetonitrile containing 0.1 % (v/v) formic acid. The conditions for the gradient elution were initially 5 % B, increasing linearly to 100 % B over 6 minutes, remaining at 100 % B for 2.5 min then decreasing linearly to 5 % B over 1 min, followed by an equilibration period of 2.5 min prior to the next injection. The flow rate was 0.5 mL/min, temperature controlled at 35 °C with an injection volume of between 2 to 5 µL. All samples were diluted with DMSO (99.9 %) prior to LC-MS analysis.

3.6. Purification by column chromatography

Column chromatography was conducted on a Combiflash® Rf automated flash chromatography system, from Teledyne Isco using disposable, normal or reverse phase, SPE Redisep cartridges (4 g to 330 g). The CombiFlash® Rf uses RFID (Radio Frequency Identification) technology to automate setting the parameters for purification runs and fraction collection. The system is equipped with a UV variable dual-wavelength and a Foxy® fraction collector enabling automated peak cutting, collection, and tracking. Reverse phase chromatography employing modified water (high pH) used a pH 10 aqueous ammonium bicarbonate solution.

3.7. Phase separators

Isolute® phase separator cartridges are fitted with a hydrophobic Teflon frit. They were used to separate chlorinated solvent from aqueous phase under gravity.

3.8. Melting point

Melting points were measured on a Stuart automatic melting point apparatus, SMP40. For compounds that decomposed over a wide temperature range, it was possible to watch a recorded video of the experiment to manually determine the melting point range.

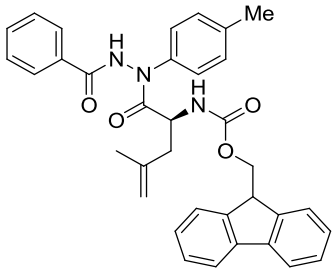
3.9. IR and fluorescence measurements

IR spectra were obtained on a PerkinElmer spectrum one FT-IR with peaks reported in cm^{-1}. Excitation and Emission spectra were recorded on a Perkin Elmer: LS50B Luminescence Spectrometer using 100 µL of the sample solution in a cuvette with path length of 10 mm. Concentration dependent fluorescence and monitoring of reactions was conducted on a Molecular Devices: Spectra Max Gemini XS. All wavelengths used are reported in the body of this document.

3.10. Experimental procedures

3.10.1. Chapter 1: Labelling of alkene tagged amino acids

(*S*)-(9*H*-fluoren-9-yl)methyl(1-(2-benzoyl-1-(p-tolyl)hydrazinyl)-4-methyl-1-oxopent-4-en-2-yl)carbamate (92)

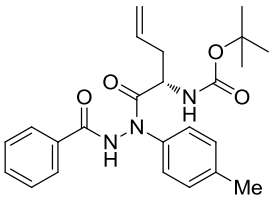


A solution of Fmoc-dehydro-Leu-OH **86** (30 mg, 0.085 mmol) and tetrazole **88** (20 mg, 0.085 mmol) in MeCN:MeOH (1:1 v/v, 2 mL) was irradiated with a UVB lamp (270-330 nm) in a flow reactor with a residence time of 60 min. The solvent was removed *in vacuo* and the crude mixture purified by automated flash column chromatography (Si, 5 g, 0-25% EtOAc-cyclohexane, 60 min) to afford **92** (30 mg, 63%) as a white solid.

LC-MS Method B (2 min): rt = 1.35 min, $[M+H]^+$ = 560.0, $[M+NH_4]^+$ = 577.0 and $[M+Na]^+$ = 582.0. mp 124-125 °C. IR (cm^{-1}): 3288, 2925, 2851, 1688, 1665, 1510, 1448, 1392, 1246, 1030. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 10.20 (br. s, 1 H), 8.95 (br. s, 1 H), 7.90 (d, *J* =5.8 Hz, 1 H), 7.83 (d, *J* = 7.3 Hz, 1 H), 7.77 (app. d, *J* = 7.6 Hz, 2 H), 7.62 - 7.53 (m, 4 H), 7.52 - 7.45 (m, 1 H), 7.40 (m, 4 H), 7.35 - 7.27 (m, 2 H), 7.24 (d, *J* = 7.3 Hz, 1 H), 7.16 (d, *J* = 7.1 Hz, 1 H), 4.58 - 4.47 (m, 1 H), 4.35 - 4.10 (m, 3 H), 4.35 - 4.10 (m, 1 H), 2.30 (s, 3 H), 2.40 - 2.18 (m, 2 H), 1.40 (s, 3 H) ppm. HRMS (ESI^+): m/z calcd for $C_{35}H_{34}N_3O_4$: 560.2544; found 560.2520.

Note: A ^{13}C NMR spectrum could not be accurately obtained due to the significant number of peaks present from rotamers. A ^1H NMR spectrum was successfully obtained by using Variable Temperature ^1H NMR.

(*S*)-*tert*-butyl (1-(2-benzoyl-1-(*p*-tolyl)hydrazinyl)-1-oxopent-4-en-2-yl)carbamate (93)

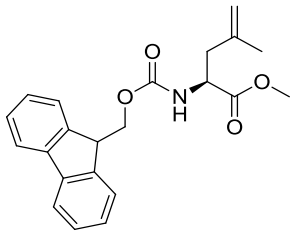


A solution of Boc-D-allyl-Gly-OH **87** (20 mg, 0.090 mmol) and tetrazole **88** (22 mg, 0.090 mmol) in MeCN:MeOH (1:1 v/v, 2 mL) was irradiated with a UVB lamp (270-330 nm) in a flow reactor with a residence time of 60 min. The solvent was removed *in vacuo* and the crude mixture purified by automated flash column chromatography (Si, 20 g, 0-25% EtOAc-cyclohexane, 60 min) to afford **93** (20 mg, 51%) as a white solid.

LC-MS Method E (3 min): rt = 1.90 min, $[M+H]^+$ = 424.1, $[M+NH_4]^+$ = 441.1 and $[M+Na]^+$ = 446.1. mp 114-117 °C. IR (cm^{-1}): 3411, 3357, 3324, 3230, 2977, 2930, 1696, 1666, 1646, 1509, 1408, 1300, 1248, 1165. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 11.20 - 10.54 (br.s., 1 H) 7.88 (app. d, *J* = 7.3 Hz, 2 H), 7.58 (t, *J* = 7.4 Hz, 1 H), 7.50 (t, *J* = 7.5 Hz, 2 H), 7.41 (app. d, *J* = 8.3 Hz, 2 H), 7.20 (app. d, *J* = 8.3 Hz, 2 H), 6.04 (d, *J* = 7.0 Hz, 1 H), 5.75 (dddd, *J* = 17.0, 10.0, 8.0, 7.0 Hz, 1 H), 5.05 (dq, *J* = 17.1, 1.7 Hz, 1 H), 5.02 (dq, *J* = 10.3, 1.5 Hz, 1 H), 4.59 (dt, *J* = 8.0, 4.9 Hz, 1 H), 2.59 - 2.47 (m, 1 H), 2.41 - 2.35 (m, 1 H), 2.33 (s, 3 H), 1.39 (s, 9 H) ppm. HRMS (ESI^+): m/z calcd for $C_{24}H_{30}N_3O_4$: 424.2236; found 424.2239.

Note: A ^{13}C NMR could not be accurately obtained due to the significant number of peaks present from rotamers.

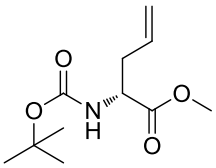
Fmoc-dehydro-Leu-OMe (103)[216]



DL-10-camphorsulfonic acid (93 mg, 0.40 mmol) was added to Fmoc-dehydro-Leu-OH **86** (200 mg, 0.57 mmol) in MeOH (5 mL). The reaction mixture was stirred at 70 °C for 3 h, cooled to room temperature and left standing for 24 h. The reaction mixture was concentrated *in vacuo* and the residue washed with sat. $NaHCO_{3(aq)}$ (20 mL). The product was extracted with EtOAc (2 x 20 mL), washed with sat. $NaCl_{(aq)}$, dried through a hydrophobic frit and concentrated *in vacuo*. The crude material was purified by automated flash column chromatography (Si, 20 g, 0-25% TBME-cyclohexane, 40 min) to afford Fmoc-dehydro-Leu-OMe **103** (164 mg, 79%) as a white solid.

LC-MS Method A (2 min): rt = 1.26 min, $[M+H]^+$ = 366.3. mp 87-89 °C. IR (cm^{-1}): 3330, 3064, 2948, 1737, 1694, 1532, 1446, 1267, 1215, 1085. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 7.76 (d, *J* = 7.5 Hz, 2 H), 7.62 - 7.54 (m, 2 H), 7.40 (t, *J* = 7.5 Hz, 2 H), 7.31 (td, *J* = 7.4, 1.0 Hz, 2 H), 5.21 (d, *J* = 8.1 Hz, 1 H), 4.88 (s, 1 H), 4.78 (s, 1 H), 4.51 (dt, *J* = 8.2, 5.6 Hz, 1 H), 4.39 (d, *J* = 7.4 Hz, 2 H), 4.23 (t, *J* = 7.4 Hz, 1 H), 3.76 (s, 3 H), 2.57 (dd, *J* = 14.0, 5.5 Hz, 1 H), 2.41 (dd, *J* = 14.0, 8.3 Hz, 1 H), 1.76 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 172.8, 155.8, 143.8 (2 C), 141.3 (2 C), 140.3, 127.7 (2 C), 127.1 (2 C), 125.1 (2 C), 120.0 (2 C), 114.7, 67.1, 52.3, 52.2, 47.2, 40.8, 21.9 ppm. HRMS (ESI^+): m/z calcd for $C_{22}H_{24}NO_4$: 366.1700; found 366.1701. The data obtained were consistent with literature values.

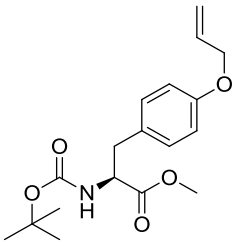
Boc-D-allyl-Gly-OMe (104)[217]



To a suspension of Boc-allyl-Gly-OH **87** (185 mg, 0.859 mmol) in acetone (15 mL) was added K_2CO_3 (238 mg, 1.72 mmol) and methyl iodide (0.11 mL, 1.80 mmol). The reaction mixture was stirred at 60 °C for 16 h. The solvent was removed *in vacuo* and the residue washed with sat. $NaHCO_{3(aq)}$ (20 mL). The product was extracted with EtOAc (2 x 20 mL), washed with sat. $NaCl_{(aq)}$ (20 mL), dried through a hydrophobic frit and concentrated *in vacuo*. The crude material was purified by automated flash column chromatography (Si, 20 g, 0-25% TBME-cyclohexane, 40 min) to afford Boc-D-allyl-Gly-OMe **104** (130 mg, 66%) as a colourless oil.

LC-MS Method E (3 min): rt = 1.52 min, $[M+H]^+$ = 230.1, $[M+Na]^+$ = 252.1. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 5.76 - 5.62 (m, 1 H), 5.15 (dd, *J* = 2.8, 1.3 Hz, 1 H), 5.11 (dd, *J* = 2.8, 1.3 Hz, 1 H), 5.08 - 4.97 (br. s, 1 H), 4.39 (m, 1 H), 3.74 (s, 3 H), 2.63 – 2.43 (m, 2 H), 1.49 - 1.40 (m, 9 H) ppm. The data obtained were consistent with literature values.

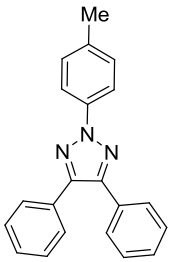
Boc-allyl-Tyr-OMe (105)[218]



To a solution of Boc-Tyr-OMe **107** (1.0 g, 3.4 mmol) in dry DMF (10 mL) was added K_2CO_3 (0.95 g, 6.8 mmol), tetrabutylammonium iodide (0.12 g, 0.33 mmol) and allyl bromide (0.35 mL, 4.0 mmol). The reaction mixture was stirred under an atmosphere of nitrogen at room temperature for 16 h. The reaction mixture was washed with a solution of 1M $KHSO_{4(aq)}$ (50 mL) then extracted with EtOAc (3 x 25 mL). The combined organic phase was washed with sat. $NaHCO_{3(aq)}$ (20 mL), sat. NaCl (20 mL) and sat. LiCl (2 x 20 mL). The organic phase was dried through a hydrophobic frit and the solvent removed *in vacuo*. The crude material was purified by automated flash column chromatography (Si, 20 g, 0-25% EtOAc-cyclohexane, 60 min) to afford **105** (848 mg, 75 %) as a pale yellow oil.

LC-MS Method A (2 min): rt = 1.18 min, $[M+H]^+$ = 335.9, $[M+Na]^+$ = 357.9. IR (cm^{-1}): 3367, 2977, 2929, 1743, 1711, 1612, 1509, 1365, 1242, 1220, 1162, 1056. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 7.04 (app. d, *J* = 8.6 Hz, 2 H), 6.85 (app. d, *J* = 8.6 Hz, 2 H), 6.06 (ddt, *J* = 17.4, 10.5, 5.3 Hz, 1 H), 5.41 (dq, *J* = 17.4, 1.5, 1 H), 5.29 (dq, *J* = 10.6, 1.5 Hz, 1 H), 4.96 (d, 1 H), 4.53 (dt, *J* = 5.3, 1.5 Hz, 2 H), 4.60 - 4.46 (m, 1 H), 3.72 (s, 3 H), 3.11 - 2.83 (m, 2 H), 1.43 (s, 9 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 172.4, 157.7, 155.1, 133.3, 130.3 (2 C), 128.1, 117.6, 114.8 (2 C), 80.0, 68.8, 54.5, 52.1, 37.5, 28.3 (3 C) ppm. HRMS (ESI^+): m/z calcd for $C_{18}H_{26}N_3O_5$: 336.1806; found 336.1806. The data obtained were consistent with literature values.

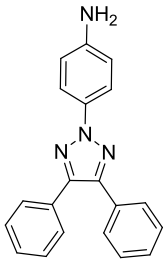
4,5-diphenyl-2-(*p*-tolyl)-2*H*-1,2,3-triazole (115)



A solution of tetrazole **88** (90 mg, 0.38 mmol) in EtOAc (50 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 7 h. The solvent was removed *in vacuo* until *ca.* 10 mL remained and the reaction mixture irradiated for a further 5 h. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 10 g, 0-50% DCM-cyclohexane, 40 min). The solvent was removed *in vacuo* and residual solvents removed azeotropically with MeOH to afford **115** (50 mg, 84%) as a yellow-orange solid.

LC-MS Method A (2 min): rt = 1.59 min, $[M+H]^+$ = 312.0. mp 107-111 °C. IR (cm^{-1}): 2920, 1509, 1439, 1292, 1265. ^1H NMR (600 MHz, $CDCl_3$): δ_H = 8.07 (app. d, *J* = 8.4 Hz, 2 H), 7.65 (dd, *J* = 7.7, 1.8 Hz, 4 H), 7.43 - 7.36 (m, 6 H), 7.31 (app. d, *J* = 8.4 Hz, 2 H), 2.43 (s, 3 H) ppm. ^{13}C NMR (151 MHz, $CDCl_3$): δ_C = 145.7 (2 C), 137.6, 137.3, 130.9 (2 C), 129.8 (2 C), 128.6 (4 C), 128.5 (2 C), 128.4 (4 C), 118.7 (2 C), 21.2 ppm. $^{15}N^1H$ HMBC NMR (60 MHz, $CDCl_3$) δ_N = 256.8 (*N*2, *ortho*-H) ppm. HRMS (ESI^+): m/z calcd for $C_{21}H_{18}N_3$: 312.1495; found 312.1496.

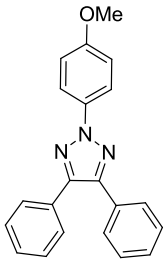
4-(4,5-diphenyl-2*H*-1,2,3-triazol-2-yl)aniline (116)



A solution of tetrazole **89** (100 mg, 0.42 mmol) in MeCN:EtOAc (1:1 v/v, 30 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 6 h. The solvent was removed *in vacuo* until *ca.* 10 mL remained followed by irradiation for a further 3 h. The solvent was removed *in vacuo* and the residue dissolved in DCM:MeOH (1:1 v/v, 25 mL) followed by irradiation for an additional 19 h. The reaction mixture was pre-absorbed onto Florisil® and purified by automated flash column chromatography (Si-NH_2, 10 g, 0-50% DCM-cyclohexane, 60 min) to afford **116** (55 mg, 84%) as a yellow-orange solid.

LC-MS Method A (2 min): rt = 1.28 min, $[M+H]^+$ = 313.1. mp 138-141 °C. IR (cm^{-1}): 3459, 3362, 2924, 1620, 1514, 1290. ^{1}H NMR (600 MHz, $CDCl_3$): δ_H = 7.96 (app. d, *J* = 8.8 Hz, 2 H), 7.65 (dd, *J* = 7.7, 1.8 Hz, 4 H), 7.44 - 7.36 (m, 6 H), 6.79 (app. d, *J* = 8.8 Hz, 2 H), 3.80 (br. s., 2 H) ppm. ^{13}C NMR (151 MHz, $CDCl_3$): δ_C = 145.9 (2 C), 145.1, 132.1, 131.1 (2 C), 128.5 (4 C), 128.4 (4 C), 128.4 (2 C), 120.3 (2 C), 115.2 (2 C) ppm. $^{15}N^1$H HMBC NMR (60 Hz, $CDCl_3$) δ_N = 257.1 (*N*2, *ortho*-H), 55.5 (NH_2) ppm. HRMS (ESI^+): m/z calcd for $C_{20}H_{17}N_4$: 313.1448; found 313.1445.

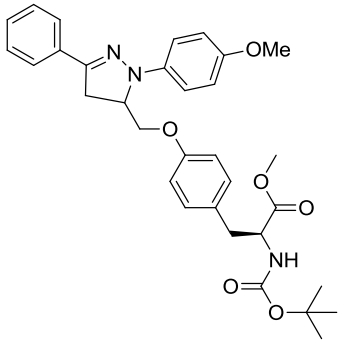
2-(4-methoxyphenyl)-4,5-diphenyl-2*H*-1,2,3-triazole (117)[114]



A solution of tetrazole **81** (100 mg, 0.396 mmol) in MeCN:EtOAc (1:1 v/v, 10 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 1 h, precipitation was observed. Additional EtOAc (5 mL) was added followed by further irradiation for 1 h, further precipitation was observed. Additional EtOAc (60 mL) was added to achieve full dissolution followed by further irradiation for 1 h. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 10 g, 0-100% DCM-cyclohexane) to afford **117** (28 mg, 43%) as a yellow solid.

LC-MS Method D (3 min): rt = 2.45 min, $[M+H]^+$ = 328.1. mp 97-101 °C. IR (cm^{-1}): 3012, 1511, 1461, 1441, 1253, 1175, 1025. ^{1}H NMR (600 MHz, $CDCl_3$): δ_H = 8.10 (app. d, *J* = 9.2 Hz, 2 H), 7.65 (dd, *J* = 7.3, 2.2 Hz, 4 H), 7.43 – 7.37 (m, 6 H), 7.02 (app. d, *J* = 9.2 Hz, 2 H), 3.89 (s, 3 H) ppm. ^{13}C NMR (151 MHz, $CDCl_3$): δ_C = 159.0, 145.5 (2 C), 133.6, 131.0 (2 C), 128.6 (4 C), 128.5 (2 C), 128.4 (4 C), 120.2 (2 C), 114.3 (2 C), 55.6 ppm. ^{15}N NMR (60 MHz, $CDCl_3$) δ_N = 256.3 (*N*2, *ortho*-H) ppm. HRMS (ESI^+): m/z calcd for $C_{21}H_{18}N_3O$: 328.1444; found 328.1445. The data obtained were consistent with literature values.

(2*S*)-methyl-2-((*tert*-butoxycarbonyl)amino)-3-(4-((1-(4-methoxyphenyl)-3-phenyl-4,5-dihydro-1*H*-pyrazol-5-yl)methoxy)phenyl)propanoate (118)

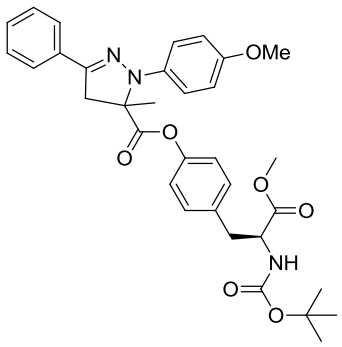


A solution of Boc-allyl-Tyr-OMe **105** (299 mg, 0.89 mmol) and tetrazole **81** (30 mg, 0.12 mmol) in MeCN (4 mL) was irradiated with a UVB lamp (270-330 nm) in a flow reactor with a residence time of 60 min. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 5g, 0-50% TBME-cyclohexane, 60 min) to afford **118** (27 mg, 41 %) as a yellow solid.

LC-MS Method E (3 min): rt = 2.33 min, $[M+H]^+$ = 559.1. mp 91-97 °C. IR (cm^{-1}): 3369, 2926, 1744, 1712, 1508, 1444, 1365, 1238, 1164, 1038. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 7.75 (app. d, *J* = 7.3 Hz, 2 H), 7.41 (t, *J* = 7.4 Hz, 2 H), 7.34 (t, *J* = 7.3 Hz, 1 H), 7.21 (app. d, *J* = 8.1 Hz, 2 H), 7.03 (app. d, *J* = 8.6 Hz, 2 H), 6.91 (app. d, *J* = 9.1 Hz, 2 H), 6.82 (app. d, *J* = 8.6 Hz, 2 H), 4.94 (d, *J* = 7.1 Hz, 1 H), 4.77 - 4.65 (m, 1 H), 4.63 – 4.48 (m, 1 H), 4.24 (dd, *J* = 8.5, 3.8 Hz, 1 H), 3.88 (dd, *J* = 11.9, 8.5 Hz, 1 H), 3.81 (s, 3 H), 3.72 (s, 3 H), 3.54 (dd, *J* = 16.1 Hz, 11.5 Hz, 1 H), 3.33 (dd, *J* = 16.8, 4.7, 1 H), 3.12 - 2.95 (m, 2 H), 1.44 (s, 9 H) ppm. HRMS (ESI^+): m/z calcd for $C_{32}H_{38}N_3O_6$: 560.2761; found 560.2756.

Note: A ^{13}C NMR spectrum could not be accurately obtained due to oxidation of the compound within the timeframe required to obtain the spectrum. See compound **120**.

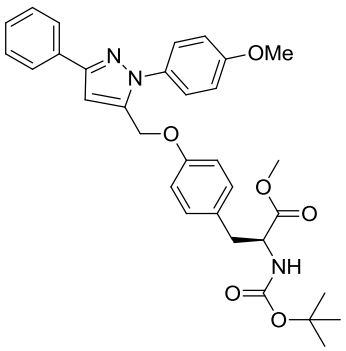
4-((*S*)-2-((*tert*-butoxycarbonyl)amino)-3-methoxy-3-oxopropyl)phenyl-1-(4-methoxyphenyl)-5-methyl-3-phenyl-4,5-dihydro-1*H*-pyrazole-5-carboxylate (119)



A solution of Boc-methacryloyl-Tyr-OMe **106** (43 mg, 0.12 mmol) and tetrazole **81** (40 mg, 0.16 mmol) in MeCN (2 mL) was stirred under irradiation with a UV lamp (270-330 nm) for 30 min. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 5g, 0-25% EtOAc-cyclohexane, 60 min) to afford **119** (20 mg, 29 %) as a pale yellow solid.

LC-MS Method C (2 min): rt = 1.43 min, $[M+H]^+$ = 588.3. mp 70-77 °C. IR (cm^{-1}): 3362, 2927, 1747, 1711, 1507, 1444, 1366, 1243, 1164. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 7.74 (app. d, *J* = 7.3 Hz, 2 H), 7.43 (t, *J* = 6.8, 2 H), 7.38 (t, *J* = 7.3 Hz, 1 H) 7.21 (app. d, *J* = 9.1 Hz, 2 H), 7.11 (app. d, *J* = 8.3 Hz, 2 H), 6.89 (app. d, *J* = 9.1 Hz, 2 H), 6.89 (app. d, *J* = 8.8, 2 H), 4.95 (d, *J* = 7.3 Hz, 1 H), 4.56 (d, *J* = 6.0 Hz, 1 H), 3.94 (d, *J* = 16.6 Hz, 1 H), 3.81 (s, 3 H), 3.69 (s, 3 H), 3.40 (d, *J* = 16.6 Hz, 1 H), 3.16 – 2.99 (m, 2 H), 1.70 (s, 3 H), 1.42 (s, 9 H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6): δ_C = 172.9, 172.0, 155.8, 154.5, 149.3, 146.5, 137.9, 136.0, 132.4, 130.6 (2 C), 129.2, 129.1 (2 C), 126.1 (2 C), 121.6 (2 C), 117.5 (2 C), 115.0 (2 C), 78.8, 70.2, 55.8, 55.5, 52.2, 47.7, 36.2, 28.6 (3 C), 21.9 ppm. HRMS (ESI^+): m/z calcd for $C_{33}H_{38}N_3O_7$: 588.2704; found 588.2683.

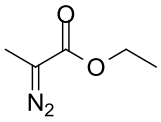
(2*S*)-methyl-2-((*tert*-butoxycarbonyl)amino)-3-(4-((1-(4-methoxyphenyl)-3-phenyl-1*H*-pyrazol-5-yl)methoxy)phenyl)propanoate (120)



A solution of **118** (10.0 mg) in $CDCl_3$ (0.5 mL) was left to stand at room temperature for 24 h and the solvent removed *in vacuo* to afford **120** (9.8 mg, 98%) as a yellow gum.

LC-MS Method E (3 min): rt = 2.25 min, $[M+H]^+$ = 558.0. IR (cm^{-1}): 3371, 2922, 1745, 1711, 1510, 1444, 1364, 1247, 1164, 1020. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 7.89 (app. d, *J* = 7.3 Hz, 2 H), 7.55 (app. d, *J* = 9.1 Hz, 2 H), 7.43 (t, *J* = 7.4 Hz, 2 H), 7.34 (t, *J* = 7.3 Hz, 1 H), 7.07 (app. d, *J* = 8.6 Hz, 2 H), 7.00 (app. d, *J* = 9.1 Hz, 2 H), 6.88 (app. d, *J* = 8.6 Hz, 2 H), 6.85 (s, 1 H), 4.98 (s, 2 H), 5.02 - 4.95 (br. s, 1 H), 4.64 - 4.46 (m, 1 H), 3.87 (s, 3 H), 3.73 (s, 3 H), 3.16 - 2.88 (m, 2 H), 1.44 (s, 9 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 172.4, 159.4, 157.1, 155.1, 151.6, 139.3, 133.0, 132.5, 130.4 (2 C), 130.3, 128.6 (2 C), 127.9, 126.3 (2 C), 125.8 (2 C), 115.1 (2 C), 114.4 (2 C), 105.8, 68.8, 60.9, 55.6, 54.6, 52.2, 37.6, 28.3 (3 C) ppm. HRMS (ESI^+): m/z calcd for $C_{32}H_{36}N_3O_6$: 558.2599; found 558.2597.

ethyl 2-diazopropanoate (139)[118]



Synthesis (a):

A suspension of tetrabutylammonium bromide (20 mg, 0.062 mmol) and sodium azide (902 mg, 13.8 mmol) in hexanes (20 mL) was cooled to 0 °C. A solution of 2M $NaOH_{(aq)}$ (3.5 mL, 7.0 mmol) was added and the reaction mixture stirred for 10 min to give a colourless solution. Triflic anhydride (1.2 mL, 7.1 mmol) was added dropwise over 5 min and the solution stirred for a further 10 min. A solution of ethyl 2-methyl-3-oxobutanoate **138** (0.50 mL, 3.5 mmol) in MeCN (20 mL) was added in one portion to give a yellow solution. The reaction mixture was stirred at 0 °C for a further 2 h. The reaction mixture was partitioned between ice water (20 mL) and cold Et_2O (20 mL), the layers separated and aqueous layer extracted with cold Et_2O (3 x 20 mL). The solvent was removed *in vacuo* at 25 °C (100 mbar) to give **139** as a crude yellow liquid.

Note: The product was volatile and some solvent could not be removed. The yield was not determined.

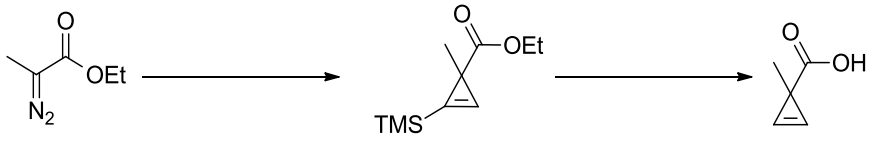
Synthesis (b):

A solution of *p*-ABSA (6.25 g, 26.0 mmol) in MeCN (35 mL) was cooled to 0 °C and ethyl 2-methyl-3-oxobutanoate **138** (2.50 g, 17.3 mmol) added in one portion. The reaction mixture was stirred under a nitrogen atmosphere and DBU (3.9 mL, 26.0 mmol) added dropwise over 10 min. The solution was stirred at 0 °C for 1 h, warmed to room temperature and stirred for a further 1.5 h. The solvent was removed *in vacuo* at 30 °C and the crude material was purified by automated flash column chromatography (Si, 100 g, 100% DCM). The solvent was removed *in vacuo* at 25 °C (125 mbar) to give **139** (1.12 g, 50%) as a yellow liquid which was stored at -8 °C.

Note: The product was volatile and some solvent could not be removed. Product yield was determined by calculating the mass of **139** relative to the mass of residual solvents by ^{1}H NMR.

^{1}H NMR (400 MHz ,$CDCl_3$) δ_H = 4.19 (q, *J* = 7.1 Hz, 2 H), 1.94 (s, 3 H), 1.24 (t, *J* = 7.1 Hz, 3 H) ppm. The data obtained were consistent with literature values.

1-methylcycloprop-2-enecarboxylic acid (141)[118]



139 **140** **141**

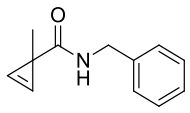
A solution of 1 mol % $Rh_2(OAc)_4$ (34 mg, 0.078 mmol) and ethynyltrimethylsilane (3.07 g, 31.2 mmol) in DCM (30 mL) was stirred under a nitrogen atmosphere at room temperature for 10 min. A solution of **139** (1.0 g, 7.8 mmol) in DCM (6 mL) was added at room temperature over 24 h (rate = 0.25 mL/h) and the reaction mixture stirred for a further 2 h. The reaction mixture was filtered through silica (Si, 2 x 5 g, 100% DCM) to afford crude **140** in DCM (80 mL).

Methanol (80 mL) was added and the reaction mixture was cooled to 0 °C followed by slow addition of 1.5M $KOH_{(aq)}$ (40 mL). The reaction mixture was stirred at room temperature for 20 h. The organic solvents were removed *in vacuo* and the aqueous layer acidified to pH 1 with 2M $HCl_{(aq)}$ then extracted with DCM (3 x 50 mL). The combined organic phase was washed with $NaCl_{(aq)}$ (50 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 5 g, 0-100% EtOAc-DCM + 0-20% MeOH-DCM, 30 min) to afford **141** (20 mg, 4%) as a pale yellow solid.

mp 81-84 °C. IR (cm^{-1}): 3154, 3110, 2965, 2934, 2870, 2823, 2675, 2546, 1686, 1656, 1417, 1306, 1130. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 7.01 (s, 2 H), 1.40 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 183.0, 109.4 (2 C), 21.5, 21.2 ppm. The data obtained were consistent with literature values.

Note: Exchangeable CO_2H proton not observed in ^{1}H NMR spectrum. Additionally, product was not analysed by LC-MS or HRMS due to lack of UV chromophore and lack of available material, respectively.

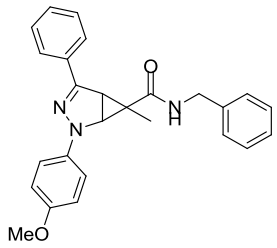
_N_-benzyl-1-methylcycloprop-2-enecarboxamide (142)



To a solution of **141** (30 mg, 0.31 mmol) in DCM (2 mL) was added a solution of EDC (70 mg, 0.37 mmol) and HOAt (50 mg, 0.37 mmol) in DCM (3 mL), followed by benzylamine (84 µL, 0.77 mmol) and DIPEA (53 µL, 0.31 mmol). The reaction mixture was stirred under a nitrogen atmosphere at room temperature for 16 h. The reaction mixture was washed with sat. $NaHCO_{3(aq)}$ (10 mL) and the aqueous layer extracted with DCM (3 x 20 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (20 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 5 g, 0-50% TBME-cyclohexane, 20 min) to afford **142** (35 mg, 59%) as a white solid.

LC-MS Method A (2 min): rt = 0.72 min, $[M+H]^+$ = 188.1. mp 80-83 °C. IR (cm^{-1}): 3310, 3132, 3082, 3031, 2965, 2928, 2855, 1650, 1616, 1529, 1354, 1286, 1241, 1206, 1153. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 7.37 - 7.23 (m, 5 H), 7.14 (s, 2 H), 5.77 - 5.31 (m, 1 H), 4.41 (d, *J* = 5.8 Hz, 2 H), 1.45 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 177.1, 138.8, 128.6 (2 C), 127.7 (2 C), 127.4, 112.1 (2 C), 43.9, 41.9, 21.2 ppm. HRMS (ESI^+): m/z calcd for $C_{12}H_{14}NO$: 188.1070; found 188.1068.

_N_-benzyl-2-(4-methoxyphenyl)-6-methyl-4-phenyl-2,3-diazabicyclo[3.1.0]hex-3-ene-6-carboxamide (143)

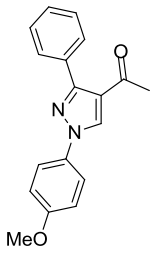


A solution of tetrazole **81** (7.0 mg, 0.028 mmol) and cyclopropene **142** (17 mg, 0.091 mmol) was stirred under irradiation with a UVB lamp (270-330 nm) for 30 min. Additional tetrazole (7.0 mg, 0.028 mmol) was added, the reaction mixture stirred for a further 2 h, and the solvent removed *in vacuo*. The crude material was dissolved in MeOH:DMSO (1:1 v/v, 1 mL) and purified by MDAP Method A (Si C_{18}, 50-99% MeCN-H_2O, 25 mins) to afford **143** (10 mg, 88%) as a yellow gum.

LC-MS Method A (2 min): rt = 1.31 min, $[M+H]^+$ = 412.1. IR (cm^{-1}): 3348, 1643, 1509, 1453, 1244. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 7.81 (dt, *J* = 8.5, 1.4 Hz, 2 H), 7.45 - 7.32 (m, 8 H), 7.27 (td, *J* = 8.6, 1.8 Hz, 2 H), 6.92 (app. d, *J* = 9.1 Hz, 2 H), 6.25 (t, *J* = 5.3 Hz, 1 H), 4.74 (d, *J* = 7.1 Hz, 1 H), 4.58 (d, *J* = 5.6 Hz, 2 H), 3.83 (s, 3 H), 3.80 (d, *J* = 6.8 Hz, 1 H), 0.80 (s, 3 H) ppm. HRMS (ESI^+): m/z calcd for $C_{26}H_{26}N_3O_2$: 412.2020; found 412.2017.

Note : A ^{13}C NMR spectrum could not be obtained as the product was found to degrade within the time frame required to run the spectrum (See compound **145**).

1-(1-(4-methoxyphenyl)-3-phenyl-1*H*-pyrazol-4-yl)ethanone (145)



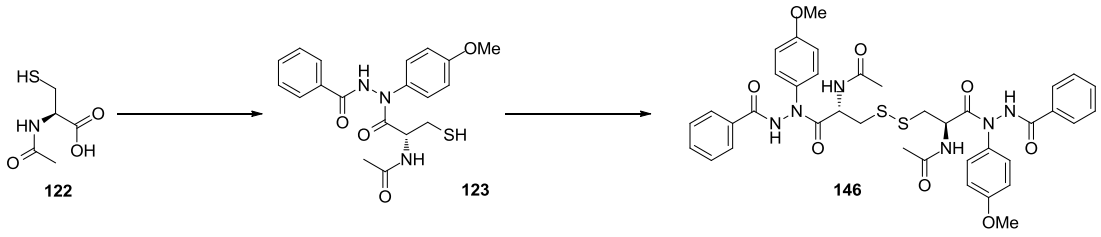
A solution of **143** (7.0 mg, 0.017 mmol) in $CDCl_3$ (0.5 mL) was left to stand at room temperature for 32 h and the solvent was removed in a Radleys blowdown apparatus under a stream of nitrogen at 50 °C. The crude material was dissolved in MeOH:DMSO (1:1 v/v, 1 mL) and purified by MDAP Method A (Si C_{18}, 30-85% MeCN-H_2O, 25 min) to afford **145** as a yellow gum (4 mg, 80%).

LC-MS Method A (2 min): rt = 1.16 min, $[M+H]^+$ = 293.1. IR (cm^{-1}): 3060, 1673, 1525, 1449, 1358, 1255. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.38 (s, 1 H), 7.75 (dd, *J* = 7.8, 2.3 Hz, 2 H), 7.70 (td, *J* = 9.1, 2.3 Hz, 2 H), 7.47 (app. d, *J* = 7.1 Hz, 3 H), 7.04 (dd, *J* = 8.8, 2.0 Hz, 2 H), 3.89 (s, 3 H), 2.39 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 192.5, 159.1, 153.3, 132.8, 132.6, 131.4, 129.4 (2 C), 128.8, 128.1 (2 C), 122.7, 121.3 (2 C), 114.7 (2 C), 55.6, 29.3 ppm. HRMS (ESI^+): m/z calcd for $C_{18}H_{17}N_2O_2$: 293.1285; found 293.1280.

General procedure for excitation and emission profiling

A 0.1M solution of sample was prepared in MeCN and diluted with MeCN:PBS (1:1 v/v) to give a 0.01 mM solution. A blank sample of MeCN:PBS (1:1 v/v) was prepared and scanned over a range of 0-800 nm followed by an excitation and emission spectrum with the following properties: Scan width = 3.0 nm, scan speed = 100 nm min^{-1}. Excitation spectra: λ_{em} = 500 nm, λ_{ex} scan = 210-450 nm. Emission spectra: λ_{ex} = 350 nm, λ_{em} scan = 360-700 nm.

_N_-((_R_)-3-(((_S_)-2-acetamido-3-(2-benzoyl-1-(4-methoxyphenyl)hydrazinyl)-3-oxopropyl)disulfanyl)-1-(2-benzoyl-1-(4-methoxyphenyl)hydrazinyl)-1-oxopropan-2-yl)acetamide (146)



A solution of *N*-acetyl cysteine **122** (60 mg, 0.37 mmol) and tetrazole **81** (30 mg, 0.12 mmol) in MeCN:EtOH (3:1 v/v, 4 mL) was stirred under irradiation with a UVB lamp (270-330 nm) at room temperature for 1 h. The solvent was removed under a stream of nitrogen in a Radleys blowdown apparatus at 50 °C and the residue partitioned between water (20 mL) and DCM (20 mL). The layers were separated and the aqueous layer was extracted with DCM (2 x 20 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$, separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material was purified by automated flash column chromatography (Si, 5 g, 0-100% EtOAc-DCM, 40 min) followed by MDAP Method B (Si C_{18}, 15-55% MeCN-H_2O, 25 min) to afford **123** (9 mg, 20%) as a white solid.

LC-MS Method F (3 min): rt = 1.37 min, $[M+H]^+$ = 388.1.

Note: Partial conversion of **123** to **146** occurred rapidly and further characterisation of **123** could not be obtained.

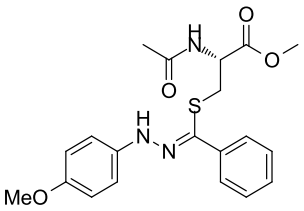
To a solution of **123** (7.0 mg, 0.018 mmol) and sodium iodide (1.0 mg, 6.7 µmol) in EtOAc (1 mL) was added 28% $H_2O_{2(aq)}$ (1.0 µL, 0.033 mmol) and the reaction mixture was stirred at room temperature for 30 min. The reaction was quenched with sat. $Na_2S_2O_{3(aq)}$ (0.5 mL) and the aqueous layer extracted with EtOAc (2 x 5 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$, separated, and dried

through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by MDAP Method C (Si C_{18}, 15-55% MeCN-H_2O, 25 min) to afford **146** (4 mg, 58%) as a white solid.

LC-MS Method A (2 min): rt = 0.99 min, $[M+H]^+$ = 773.2. IR (cm^{-1}): 1665, 1510, 1250. ^1H NMR (400 MHz, DMSO-d_6): δ_H = 11.30 - 10.52 (m, 2 H), 7.90 (app. d, *J* = 7.3 Hz, 4 H), 7.79 - 7.67 (m, 2 H), 7.57 (tt, *J* = 7.3, 1.3 Hz, 2 H), 7.48 (t, *J* = 7.6 Hz, 6 H), 6.97 (app. d, *J* = 8.6 Hz, 4 H), 6.16 - 5.64 (m, 1 H), 5.05 (m, 2 H), 3.79 (s, 6 H), 3.15 - 3.01 (m, 2 H), 2.78 - 2.69 (m, 2 H), 1.85 (s, 6 H) ppm. HRMS (ESI^+): m/z calcd for $C_{38}H_{41}N_6O_8S_2$: 773.2422; found 773.2425.

Note: Shortage of available material limited the data that could be obtained. IR and ^{13}C NMR spectra were weak and could not be effectively analysed.

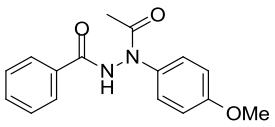
(*R*,*Z*)-methyl-2-acetamido-3-(((2-(4-methoxyphenyl)hydrazono)(phenyl)methyl)thio)propanoate (150)



A solution of Boc-methacryloyl-Tyr-OMe **106** (48 mg, 0.13 mmol), *N*-acetyl-Cys-OMe (23 mg, 0.13 mmol) and tetrazole **81** (33 mg, 0.13 mmol) in MeCN (2 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 3 h. The solvent was removed under a stream of Nitrogen in a Radleys blowdown apparatus at 50 °C and the crude material purified by automated flash column chromatography (Si, 10 g, 0-100% TBME-cyclohexane + 0-20% MeOH-DCM, 30 min) followed by MDAP Method A (Si C_{18}, 50-99% MeCN-H_2O, 15 min) to afford **150** (5.5 mg, 10%) as a yellow gum.

LC-MS Method A (2 min): rt = 1.13 min, $[M+H]^+$ = 402.1. IR (cm^{-1}): 3271, 3059, 2952, 2834, 1744, 1656, 1510, 1440, 1228, 1175, 1143, 1033. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.75 (s, 1 H), 7.88 (app. d, *J* = 8.8 Hz, 2 H), 7.42 (t, *J* = 7.8 Hz, 2 H), 7.36 (t, *J* = 8.1 Hz, 1 H), 7.22 (d, *J* = 8.8 Hz, 2 H), 6.93 (app. d, *J* = 8.8 Hz, 2 H), 6.24 (d, *J* = 7.3 Hz, 1 H), 4.90 - 4.81 (m, 1 H), 3.83 (s, 3 H), 3.74 (s, 3 H), 3.32 (dd, *J* = 14.1, 3.8 Hz, 1 H), 3.14 (dd, *J* = 14.4, 4.5 Hz, 1 H), 1.87 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 170.8, 169.6, 154.5, 137.8, 136.5, 132.5, 128.6 (2 C), 128.4, 126.8 (2 C), 114.9 (2 C), 114.7 (2 C), 55.7, 53.2, 52.9, 35.3, 22.9 ppm. HRMS (ESI^+): m/z calcd for $C_{20}H_{24}N_3O_4S$: 402.1482; found 402.1487.

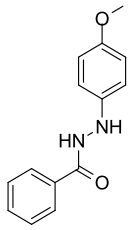
_N_-acetyl-_N_-(4-methoxyphenyl)benzohydrazide (152)



A solution of Boc-methacryloyl-Tyr-OMe **106** (32 mg, 0.087 mmol), glacial AcOH (5.0 µL, 0.087 mmol) and tetrazole **81** (23 mg, 0.090 mmol) in MeCN (2 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 1 h. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 10 g, 0-100% TBME-cyclohexane + 0-20% MeOH-DCM, 40 min) followed by MDAP Method A (Si C_{18}, 15-55% MeCN-H_2O, 15 min) to afford **152** (12 mg, 24%) as a yellow gum.

LC-MS Method A (2 min): rt = 0.81 min, $[M+H]^+$ = 285.0. IR (cm^{-1}): 3529, 3004, 2838, 1656, 1605, 1507, 1377, 1298, 1247, 1181, 1032. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 9.25 (s, 1 H), 7.77 (app. d, *J* = 7.6 Hz, 2 H), 7.51 - 7.42 (m, 3 H), 7.34 (d, *J* = 7.6 Hz, 2 H), 6.93 (app. d, *J* = 8.6 Hz, 2 H), 3.83 (s, 3 H), 2.06 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 170.5, 166.4, 159.6, 135.0, 132.1, 131.8, 129.2 (2 C), 128.5 (2 C), 127.4 (2 C), 114.6 (2 C), 55.5, 22.1 ppm. HRMS (ESI^+): m/z calcd for $C_{16}H_{17}N_2O_3$: 285.1234; found 285.1238.

_N_'-(4-methoxyphenyl)benzohydrazide (153)[219]



A solution of benzoic acid (200 mg, 1.6 mmol) in dry DCM (5 mL) was cooled to 0 °C under a nitrogen atmosphere. Triethylamine (0.27 mL, 2.0 mmol) and ethyl chloroformate (0.17 mL, 1.8 mmol) were added sequentially and the reaction mixture was stirred at 0 °C for 30 min. A white precipitate was observed. Solid (4-methoxyphenyl)hydrazine, hydrochloride (315 mg, 1.8 mmol) and triethylamine (0.27 mL, 2.0 mmol) were added, the reaction mixture allowed to warm to room temperature and then stirred for a further 1 h. The reaction mixture was sequentially washed with sat. citric $acid_{(aq)}$ (10 mL), sat. $NaHCO_{3(aq)}$ (10 mL) and sat. $NaCl_{(aq)}$ (10 mL). The organic layer was separated and dried through a hydrophobic frit. The solvent removed *in vacuo* and the crude material was purified by automated flash column chromatography (Si, 20 g, 0-50% EtOAc-cyclohexane, 60 min) then MDAP method A (Si C_{18}, 15-55% MeCN-H_2O, 25 min) to afford **153** (88 mg, 22%) as a white solid.

LC-MS Method A (2 min): rt = 0.79 min, $[M+H]^+$ = 243.1. mp 136-138 °C. IR (cm^{-1}): 3261, 1640, 1506, 1476, 1461, 1322, 1294, 1238, 1034. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 8.28 (s, 1 H), 7.83 (dd, *J*= 7.1, 1.3 Hz, 2 H), 7.56 (t, *J* = 7.3 Hz, 1 H), 7.46 (t, *J* = 7.8 Hz, 2 H), 6.90 (app. d, *J* = 8.8 Hz, 2 H), 6.81 (td, *J* = 8.8, 3.3 Hz, 2 H), 6.39 (br. s., 1 H), 3.76 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 167.8, 154.9, 141.6, 132.4, 132.2, 128.7 (2 C), 127.1 (2 C), 115.8 (2 C), 114.7 (2 C), 55.7 ppm. HRMS (ESI^+): m/z calcd for $C_{14}H_{15}N_2O_2$: 243.1128; found 243.1123. The data obtained were consistent with literature values.

Fluorogenic click reaction carried out in HeLa cell line and monitored by confocal microscopy

The suspension of HeLa cells were sourced from BioCat (# 81308), expanded and frozen down 12 months previously under GRITS #49806. The cells were suspended in MR1-4 (Sigma Aldrich Fine Chemicals manufactured media) supplemented with 10% FBS (Gibco Life Technologies, #10099-141). They were grown at 37 °C, 5% CO_2, 80% humidity, with shaking at 100 rpm and plated in 96 well plates with a cell density of 25000 cells mL^{-1} with 100 µL of cell suspension per well.

Solutions of Boc-allyl-Tyr-OMe **105** (280 µM), Boc-methacryloyl-Tyr-OMe **106** (280 µM) and tetrazole **81** (2 mM) were prepared in DMSO. From these stock solutions the following diluted solutions were prepared:

1) 5% DMSO in PBS control
2) Boc-allyl-Tyr-OMe **105** control (14 µM in 5% DMSO/PBS)
3) Boc-methacryloyl-Tyr-OMe **106** control (14 µM in 5% DMSO/PBS)
4) Tetrazole **81** control (100 µM in 5% DMSO/PBS)
5) Reaction Mixture 1 (in 5% DMSO/PBS): Boc-allyl-Tyr-OMe **105** (14 µM) + tetrazole **81** (100 µM)
6) Reaction Mixture 2 (in 5% DMSO/PBS): Boc-methacryloyl-Tyr-OMe **106** (14 µM) + tetrazole **81** (100 µM)

From previously obtained products a set of controls were made at three concentrations for each:

7) Pyrazoline adduct **118** (15 µM in 5% DMSO/PBS)
8) Pyrazoline adduct **118** (7.5 µM in 5% DMSO/PBS)
9) Pyrazoline adduct **118** (3 µM in 5% DMSO/PBS)
10) Pyrazoline adduct **119** (15 µM in 5% DMSO/PBS)
11) Pyrazoline adduct **119** (7.5 µM in 5% DMSO/PBS)
12) Pyrazoline adduct **119** (3 µM in 5% DMSO/PBS)

From previously obtained byproducts from various side reactions a number of control solutions were prepared:

13) Pyrazole **120** (15 μM in 5% DMSO/PBS)
14) Triazole **117** (50 μM in 5% DMSO/PBS)
15) Hydrazide **152** (100 μM in 5% DMSO/PBS)
16) Hydrazonothioate **150** (100 μM in 5% DMSO/PBS)
17) Hydrazide **153** (100 μM in 5% DMSO/PBS)

These solutions were utilised in a cell based assay conducted with HeLa cells using the following protocol on two 96 well plates:

1) Remove growth media.
2) Add PBS containing DRAQ5 (5 μM) to stain the cell nuclei. Incubate for 15 min in the dark at room temperature.
3) Spin down at 2000 rpm to remove supernatant. Wash with PBS.
4) Spin down at 2000 rpm to remove supernatant. Add PBS to columns 1-6, add PBS plus 0.1% triton in PBS to remaining columns 7-12. Note: Triton is a detergent that permeabalizes the cells.
5) Incubate for 10 min at room temperature in the dark.
6) Spin down at 2000 rpm to remove supernatant.
7) Add 100 μL of reagent mixtures to appropriate wells in columns 1-6 and in the same positions in wells 7-12.
8) Incubate cells with reagents for 30 min at 37 °C.
9) Spin down at 2000 rpm to remove supernatant and wash with PBS (repeat twice).
10) Irradiate for 3 minutes using UVB lamp (270-330 nm).
11) Cell imaging conducted immediately on a Perkin Elmer Opera confocal microscope.
    - Nucleus Observation – Cy 5 (Excitation = 645 nm, Emission = 705 nm)
    - Reaction Observation 2 – DAPI (Excitation 350 nm), FITC (Emission = 470 nm)

3.10.2. Chapter 1: Synthesis of 2,4,5-trisubstituted 1,2,3-triazoles

General Procedure A: Synthesis of 2,5-disubstituted tetrazoles[144]

A solution of tetrazole **173** (1 equiv.), boronic acid **174** (2 equiv.) and Cu_2O (5 mol %) in DMSO was stirred under O_2 (1 atm) at 110 °C until the tetrazole was consumed. The reaction mixture was allowed to cool, diluted with EtOAc (200 mL) and washed successively with 1M $HCl_{(aq)}$ (50 mL) and sat. $NaCl_{(aq)}$ (50 mL) three times. The organic layer was separated and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material was purified as stated in each case.

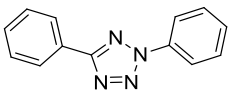
General Procedure: Solvent Screen

A solution of tetrazole **175e** (30 mg, 0.059 mmol) in the appropriate solvent (3 mL) was stirred under irradiation with a UVB lamp (270-330 nm) at room temperature for 4 h. The samples were analysed by LC-MS and the reaction profile used to determine the optimum solvent.

General Procedure B: Synthesis of *N*2-aryl 4,5-disubstituted 1,2,3-triazoles

A solution of tetrazole **175a-q** in 2-MeTHF:IPA (1:1 v/v, 10 mL) was stirred under irradiation with a UVB lamp (270-330 nm) at room temperature until all tetrazole was consumed. The solvent was removed *in vacuo* and the crude material purified as stated in each case.

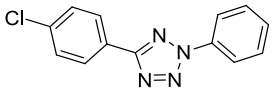
2,5-diphenyl-2*H*-tetrazole (175a)[220]



Synthesised by general procedure A using 5-phenyl-1*H*-tetrazole (1.0 g, 6.8 mmol), phenylboronic acid (1.67 g, 13.7 mmol), and Cu_2O (49 mg, 0.34 mmol) over 6 h. Purified by automated flash column chromatography (Si, 50 g, 0-10% EtOAc-cyclohexane, 60 min) to afford tetrazole **175a** (1.31 g, 86%) as a white solid.

LC-MS Method A (2 min): rt = 1.27 min, $[M+H]^+$ = 223.0. mp 101-102 °C. IR (cm^{-1}): 1595, 1530, 1495, 1471, 1448, 1369, 1215, 1073, 1016. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.29 (dd, *J* = 2.0, 8.1 Hz, 2 H), 8.24 (dd, *J* = 1.3, 8.8 Hz, 2 H), 7.65 - 7.58 (m, 2 H), 7.58 - 7.50 (m, 4 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 165.2, 137.0, 130.5, 129.7, 129.6 (2 C), 128.9 (2 C), 127.2, 127.1 (2 C), 119.9 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{13}H_{11}N_4$: 223.0978; found 223.0971. The data obtained were consistent with literature values.

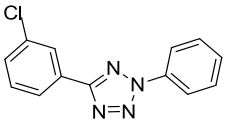
5-(4-chlorophenyl)-2-phenyl-2*H*-tetrazole (175b)[221]



Synthesised by general procedure A using 5-(4-chlorophenyl)-1*H*-tetrazole (1.0 g, 5.54 mmol), phenylboronic acid (1.35 g, 11.1 mmol), and Cu_2O (40 mg, 0.28 mmol) over 72 h. Purified by automated flash column chromatography (Si, 50 g, 0-10% EtOAc-cyclohexane, 60 min) followed by recrystallisation from hot MeOH to afford tetrazole **175b** (915 mg, 64%) as a white crystalline solid.

LC-MS Method A (2 min): rt = 1.40 min, $[M+H]^+$ = 257.1. mp 109-111 °C. IR (cm^{-1}): 1604, 1494, 1454, 1414, 1365, 1272, 1211. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.29 - 8.15 (m, 4 H), 7.61 (tt, *J* = 1.5, 7.1 Hz, 2 H), 7.57 - 7.47 (m, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 164.4, 136.8, 136.6, 129.8, 129.7 (2 C), 129.3 (2 C), 128.3 (2 C), 125.7, 119.9 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{13}H_{10}ClN_4$: 257.0589; found 257.0579.

5-(3-chlorophenyl)-2-phenyl-2*H*-tetrazole (175c)

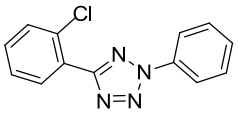


Synthesised by general procedure A using 5-(3-chlorophenyl)-1*H*-tetrazole (1.0 g, 5.54 mmol), phenylboronic acid (1.35 g, 11.1 mmol), and Cu_2O (40 mg, 0.28 mmol) over 24 h. Purified by automated flash column chromatography

(Si, 50 g, 0-5% EtOAc-cyclohexane, 60 min) to afford tetrazole **175c** (1.20 g, 84%) as a white solid.

LC-MS Method A (2 min): rt = 1.41 min, $[M+H]^+$ = 257.0. mp 95-96 °C. IR (cm^{-1}): 1596, 1576, 1522, 1497, 1469, 1448, 1374, 1348, 1300, 1265, 1214, 1074. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.29 (dd, *J* = 1.5, 2.3 Hz, 1 H), 8.23 (dd, *J* = 2.3, 7.6 Hz, 2 H), 8.18 (ddd, *J* = 1.5, 2.8, 6.0 Hz, 1 H), 7.61 (t, *J* = 7.3 Hz, 2 H), 7.56 (td, *J* = 1.3, 7.6 Hz, 1 H), 7.50 (dd, *J* = 1.3, 4.3 Hz, 2 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 164.1, 136.8, 135.1, 130.6, 130.3, 129.8, 129.7 (2 C), 128.9, 127.1, 125.1, 119.9 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{13}H_{10}ClN_4$: 257.0589; found 257.0579.

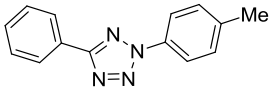
5-(2-chlorophenyl)-2-phenyl-2*H*-tetrazole (175d)



Synthesised by general procedure A using 5-(2-chlorophenyl)-1*H*-tetrazole (1.0 g, 5.5 mmol), phenylboronic acid (1.35 g, 11.1 mmol), and Cu_2O (40 mg, 0.28 mmol) over 24 h. Purified by automated flash column chromatography (Si, 50 g, 0-5% EtOAc-cyclohexane, 60 min) to afford tetrazole **175d** (1.21 g, 85%) as a white solid.

LC-MS Method A (2 min): rt = 1.29 min, $[M+H]^+$ = 257.0. mp 73-75 °C. IR (cm^{-1}): 1599, 1491, 1469, 1427, 1261, 1208. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.25 (dd, *J* = 1.3, 7.6 Hz, 2 H), 8.12 - 8.07 (m, 1 H), 7.68 - 7.58 (m, 3 H), 7.54 (tt, *J* = 1.0, 7.1 Hz, 1 H), 7.51 - 7.42 (m, 2 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 163.5, 136.9, 133.3, 131.4, 131.3, 131.0, 129.8, 129.7 (2 C), 126.9, 126.3, 120.0 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{13}H_{10}ClN_4$: 257.0589; found 257.0575.

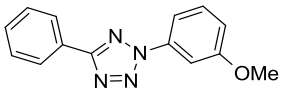
5-phenyl-2-(*p*-tolyl)-2*H*-tetrazole (175f)[144]



Synthesised by general procedure A using 5-phenyl-1*H*-tetrazole (1.0 g, 6.8 mmol), *p*-tolylboronic acid (1.86 g, 13.7 mmol), and Cu_2O (49 mg, 0.34 mmol) over 6 h. Purified by automated flash column chromatography (Si, 50 g, 0-10% EtOAc-cyclohexane, 60 min) to afford tetrazole **175f** (1.51 g, 93%) as a white solid.

LC-MS Method A (2 min): rt = 1.36 min, $[M+H]^+$ = 237.0. mp 103-104 °C. IR (cm^{-1}): 1507, 1448, 1369, 1212, 1073, 1012. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.26 (dd, *J* = 2.0, 8.1 Hz, 2 H), 8.08 (app. d, *J* = 8.6 Hz, 2 H), 7.57 - 7.47 (m, 3 H), 7.38 (d, *J* = 8.6 Hz, 2 H), 2.46 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 165.1, 139.9, 134.8, 130.5, 130.2 (2 C), 128.9 (2 C), 127.3, 127.0 (2 C), 119.8 (2 C), 21.2 ppm. HRMS (ESI^+): m/z calcd for $C_{14}H_{13}N_4$: 237.1135; found 237.1128. The data obtained were consistent with literature values.

2-(3-methoxyphenyl)-5-phenyl-2*H*-tetrazole (175h)[80]

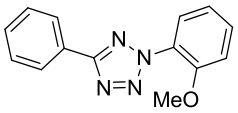


Synthesised by general procedure A using 5-phenyl-1*H*-tetrazole (1.0 g, 6.8 mmol), (3-methoxyphenyl)boronic acid (2.08 g, 13.7 mmol), and Cu_2O (49 mg, 0.34 mmol) over 16 h. Purified by automated flash column chromatography (Si, 50 g, 0-50% DCM-cyclohexane, 60 min) to afford tetrazole **175h** (1.33 g, 77%) as a white solid.

LC-MS Method A (2 min): rt = 1.29 min, $[M+H]^+$ = 253.3. mp 89-91 °C. IR (cm^{-1}): 3012, 2973, 2768, 1611, 1594, 1496, 1450, 1246, 1209, 1172, 1018. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.33 - 8.24 (m, 2 H), 7.82 (ddd, *J* = 0.9, 2.0, 8.0 Hz, 1 H), 7.78 (t, *J* = 2.1 Hz, 1 H), 7.59 - 7.52 (m, 3 H), 7.49 (t, *J* = 8.2 Hz, 1 H), 7.07 (ddd, *J* = 0.8, 2.5, 8.3 Hz, 1 H), 3.96 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 165.2, 160.5, 137.9, 130.6, 130.5, 128.9 (2 C), 127.2, 127.1 (2 C), 115.8, 112.1, 105.3, 55.7 ppm. HRMS

(ESI$^+$): m/z calcd for $C_{14}H_{13}N_4O$: 253.1084; found 253.1081. The data obtained were consistent with literature values.

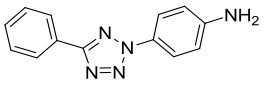
2-(2-methoxyphenyl)-5-phenyl-2*H*-tetrazole (175i)[80]



Synthesised by general procedure A using 5-phenyl-1*H*-tetrazole (1.0 g, 6.8 mmol), (2-methoxyphenyl)boronic acid (2.08 g, 13.7 mmol), and Cu_2O (49 mg, 0.34 mmol) over 16 h. After work-up the crude material was dissolved in DCM, pre-absorbed onto Florisil®, and purified by automated flash column chromatography (Si, 50 g, 0-25% EtOAc-cyclohexane, 60 min) to afford tetrazole **175i** (853 mg, 49%) as a viscous yellow oil.

LC-MS Method A (2 min): rt = 1.15 min, $[M+H]^+$ = 253.3. IR (cm^{-1}): 2934, 2841, 1604, 1505, 1466, 1449, 1285, 1257. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.30 - 8.24 (m, 2 H), 7.61 (dt, *J* = 1.5, 7.6 Hz, 1 H), 7.58 - 7.49 (m, 4 H), 7.16 (t, *J* = 8.6 Hz, 2 H), 3.91 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 164.9, 153.7, 132.0, 130.4, 128.9 (2 C), 127.4, 127.1 (2 C), 127.0, 126.5, 120.7, 112.8, 56.3 ppm. HRMS (ESI$^+$): m/z calcd for $C_{14}H_{13}N_4O$: 253.1084; found 253.1072. The data obtained were consistent with literature values.

4-(5-phenyl-2*H*-tetrazol-2-yl)aniline (175j)[80]

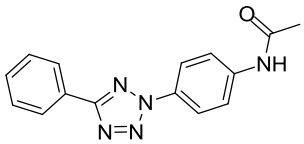


A solution of tetrazole **175n** (500 mg, 1.9 mmol) in THF (30 mL) was stirred under a hydrogen atmosphere (55 psi) at room temperature for 2 h. Note: longer exposure may cause further hydrogenation of the aromatic ring. The reaction mixture was filtered through Celite®, washed with THF, and the solvent removed *in vacuo.* The

crude material was purified by automated flash column chromatography (Si, 20 g, 0-25% EtOAc-cyclohexane, 40 min) to afford tetrazole **175j** (381 mg, 86%) as a yellow solid.

LC-MS Method A (2 min): rt = 1.04 min, $[M+H]^+$ = 238.3. mp 158-160 °C. IR (cm^{-1}): 3407, 3329, 3221, 1635, 1602, 1514, 1450, 1298, 1206, 1184, 1172. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.26 (dd, *J* = 7.6, 1.5 Hz, 2 H), 7.97 (app. d, *J* = 8.6 Hz, 2 H), 7.60 - 7.45 (m, 3 H), 6.83 (app. d, *J* = 8.6 Hz, 2 H), 4.25 - 3.80 (br. s, 2 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 164.7, 147.8, 130.3, 128.9 (2 C), 128.7, 127.5, 126.9 (2 C), 121.5 (2 C), 115.0 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{13}H_{12}N_5$: 238.1087; found 238.1079. The data obtained were consistent with literature values.

_N_-(4-(5-phenyl-2_H_-tetrazol-2-yl)phenyl)acetamide (175k)[80]

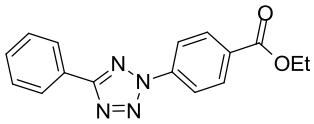


A solution of tetrazole **175j** (100 mg, 0.42 mmol) and acetic anhydride (0.5 mL, 5.3 mmol) in DCM (4 mL) was stirred under a nitrogen atmosphere at room temperature for 5 h. The solvent was removed *in vacuo* and the residue partitioned between EtOAc (10 mL) and sat. $NaHCO_{3(aq)}$ (10 mL). The layers were separated and the aqueous layer extracted with EtOAc (2 x 10 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$, separated, and dried through a hydrophobic frit to afford tetrazole **175k** (108 mg, 92%) as a yellow solid.

LC-MS Method A (2 min): rt = 1.06 min, $[M+H]^+$ = 280.3. mp 178-180 °C. IR (cm^{-1}): 3320, 3213, 3073, 2926, 1677, 1610, 1546, 1508, 1450, 1416, 1368, 1312, 1260, 1205, 1184, 1171. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 10.30 (s, 1 H), 8.17 (dd, *J* = 2.3, 7.8 Hz, 2 H), 8.09 (app. d, *J* = 9.1 Hz, 2 H), 7.89 (app. d, *J* = 9.1 Hz, 2 H), 7.65 - 7.55 (m, 3 H), 2.11 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 169.2, 164.8, 141.4, 131.5, 131.3 (2 C), 129.7 (2 C), 127.0 (2 C), 121.1 (2 C), 120.1 (2 C), 24.6 ppm. HRMS

(ESI$^+$): m/z calcd for $C_{15}H_{14}N_5O$: 280.1193; found 280.1191. The data obtained were consistent with literature values.

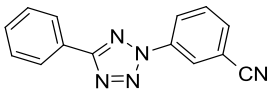
ethyl 4-(5-phenyl-2*H*-tetrazol-2-yl)benzoate (175l)[80]



Synthesised by general procedure A using 5-phenyl-1*H*-tetrazole (1.0 g, 6.8 mmol), (4-(ethoxycarbonyl)phenyl)boronic acid (2.65 g, 13.7 mmol), and Cu_2O (49 mg, 0.34 mmol) over 24 h. After work-up the crude material was dissolved in DCM, pre-absorbed onto Florisil®, and purified by automated flash column chromatography (Si, 50 g, 0-10% TBME-cyclohexane, 60 min). The solvent was removed *in vacuo* and the residue triturated with MeOH, filtered, washed with MeOH, and dried *in vacuo* to afford tetrazole **175l** (1.51 g, 75%) as an off-white solid.

LMCS Method A (2 min): rt = 1.37 min, $[M+H]^+$ = 295.1. mp 102-104 °C. IR (cm^{-1}): 2973, 2932, 2905, 1713, 1607, 1449, 1368, 1271, 1210, 1106, 1071. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.36 - 8.24 (m, 6 H), 7.62 - 7.50 (m, 3 H), 4.46 (q, *J* = 7.1 Hz, 2 H), 1.46 (t, *J* = 7.1 Hz, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 165.6, 165.4, 139.7, 131.4, 131.2 (2 C), 130.8, 129.0 (2 C), 127.2 (2 C), 126.9, 119.5 (2 C), 61.5, 14.3 ppm. HRMS (ESI$^+$): m/z calcd for $C_{16}H_{15}N_4O_2$: 295.1190; found 295.1183. The data obtained were consistent with literature values.

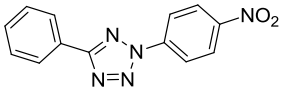
3-(5-phenyl-2*H*-tetrazol-2-yl)benzonitrile (175m)



Synthesised by general procedure A using 5-phenyl-1*H*-tetrazole (1.0 g, 6.8 mmol), (3-cyanophenyl)boronic acid (2.01 g, 13.7 mmol), and Cu_2O (49 mg, 0.34 mmol) over 6 h. After work-up the crude material was dissolved in DCM, pre-absorbed onto Florisil®, and purified by automated flash column chromatography (Si, 50 g, 0-10% EtOAc-cyclohexane, 60 min). The solvent was removed *in vacuo* and the residue triturated with MeOH, filtered, washed with MeOH, and dried *in vacuo* to afford tetrazole **175m** (1.02 g, 60%) as a white solid.

LC-MS Method A (2 min): rt = 1.20 min, $[M+H]^+$ = 248.0. mp 117-119 °C. IR (cm^{-1}): 3076, 2231, 1583, 1486, 1471, 1447, 1382, 1209, 1018. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.56 (t, *J* = 1.6 Hz, 1 H), 8.52 (ddd, *J* = 1.3, 2.0, 8.1 Hz, 1 H), 8.31 - 8.27 (m, 2 H), 7.82 (td, *J* = 1.5, 7.8 Hz, 1 H), 7.76 (t, *J* = 7.8 Hz, 1 H), 7.59 - 7.55 (m, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 165.8, 137.3, 132.8, 131.0, 130.8, 129.1 (2 C), 127.2 (2 C), 126.6, 123.7, 123.1, 117.3, 114.3 ppm. HRMS (ESI^+): m/z calcd for $C_{14}H_{10}N_5$: 248.0931 ; found 248.0916.

2-(4-nitrophenyl)-5-phenyl-2*H*-tetrazole (175n)[222]

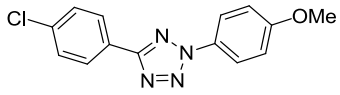


Synthesised by general procedure A using 5-phenyl-1*H*-tetrazole (1.0 g, 6.8 mmol), (4-nitrophenyl)boronic acid(2.28 g, 13.7 mmol), and Cu_2O (49 mg, 0.34 mmol) over 24 h. Purified by recrystallisation from hot MeOH to afford tetrazole **175n** (1.01 g, 55%) as a white solid.

LC-MS Method A (2 min): rt = 1.24 min, $[M+H]^+$ = 268.2. mp 175-176 °C. IR (cm^{-1}): 3121, 3093, 1596, 1524, 1495, 1468, 1448, 1339, 1212, 1109, 1014. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.53 - 8.44 (m, 4 H), 8.33 - 8.26 (m, 2 H), 7.60 - 7.55 (m, 3 H)

ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 166.0, 147.9, 140.6, 131.1, 129.1 (2 C), 127.2 (2 C), 126.4, 125.5 (2 C), 120.3 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{13}H_{10}N_5O_2$: 268.0829; found 268.0828. The data obtained were consistent with literature values.

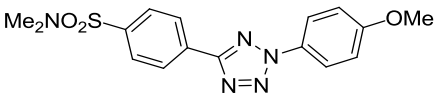
5-(4-chlorophenyl)-2-(4-methoxyphenyl)-2*H*-tetrazole (175p)[223]



Synthesised by general procedure A using 5-(4-chlorophenyl)-1*H*-tetrazole (1.0 g, 5.54 mmol), (4-methoxyphenyl)boronic acid (1.68 g, 11.1 mmol), and Cu_2O (40 mg, 0.28 mmol) over 22 h. Purified by trituration in MeOH, filtering, washing with MeOH, and drying *in vacuo* to afford tetrazole **175p** (1.11 g, 70%) as a white solid.

LC-MS Method A (2 min): rt = 1.40 min, $[M+H]^+$ = 287.2. mp 153-155 °C. IR (cm^{-1}): 2999, 2964, 2836, 1599, 1509, 1468, 1455, 1433, 1415, 1254, 1208, 1182, 1092. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.21 (app. d, *J* = 8.8 Hz, 2 H), 8.12 (app. d, *J* = 9.1 Hz, 2 H), 7.52 (app. d, *J* = 8.6 Hz, 2 H), 7.09 (app. d, *J* = 9.1 Hz, 2 H), 3.93 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 164.1, 160.6, 136.5, 130.4, 129.2 (2 C), 128.3 (2 C), 125.8, 121.4 (2 C), 114.7 (2 C), 55.7 ppm. HRMS (ESI^+): m/z calcd for $C_{14}H_{12}ClN_4O$: 287.0694; found 287.0695.

4-(2-(4-methoxyphenyl)-2*H*-tetrazol-5-yl)-*N*,*N*-dimethylbenzenesulfonamide (175q)

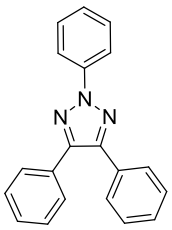


Synthesised by general procedure A from *N*,*N*-dimethyl-4-(1*H*-tetrazol-5-yl)benzenesulfonamide (350 mg, 1.4 mmol), (4-methoxyphenyl)boronic acid (420 mg, 2.8 mmol), and Cu_2O (10 mg, 0.069 mmol) over 16 h. Purified by

trituration in MeOH, filtering, washing with MeOH, and drying *in vacuo* to afford tetrazole **175q** (393 mg, 79%) as a brown solid.

LC-MS Method A (2 min): rt = 1.17 min, $[M+H]^+$ = 360.2. mp 187-189 °C. IR (cm^{-1}): 3052, 3014, 2978, 2841, 1512, 1464, 1334, 1254, 1161. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.45 (app. d, *J* = 8.3 Hz, 2 H), 8.14 (td, *J* = 2.0, 9.1 Hz, 2 H), 7.96 (app. d, *J* = 8.3 Hz, 2 H), 7.11 (td, *J* = 2.3, 9.1 Hz, 2 H), 3.93 (s, 3 H), 2.79 (s, 6 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 163.6, 160.8, 137.2, 131.4, 130.3, 128.3 (2 C), 127.5 (2 C), 121.5 (2 C), 114.8 (2 C), 55.7, 37.9 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{16}H_{18}N_5O_3S$: 360.1125; found 360.1125.

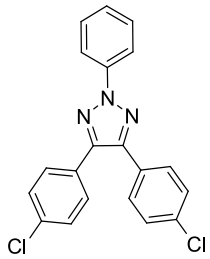
2,4,5-triphenyl-2*H*-1,2,3-triazole (176a)[114]



Synthesised by general procedure B from tetrazole **175a** (200 mg, 0.90 mmol) over 24 h. Purified by automated reverse phase chromatography (Si C_{18}, 50-95% MeCN:H_2O, HCO_2H modifier, 30 min) to afford triazole **176a** (25 mg, 19%) as a yellow gum.

LC-MS Method A (2 min): rt = 1.54 min, $[M+H]^+$ = 298.1. mp 113-116 °C. IR (cm^{-1}): 1596, 1494, 1460, 1440, 1292, 1267, 1074. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.23 (d, *J* = 7.7 Hz, 2 H), 7.69 (dd, *J* = 1.9, 7.7 Hz, 4 H), 7.54 (t, *J* = 8.0 Hz, 2 H), 7.47 - 7.42 (m, 6 H), 7.40 (t, *J* = 7.4 Hz, 1 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 146.0 (2 C), 139.7, 130.8 (2 C), 129.3 (2 C), 128.7 (2 C), 128.6 (4 C), 128.5 (4 C), 127.4, 118.8 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{20}H_{16}N_3$: 298.1339; found 298.1339. The data obtained were consistent with literature values.

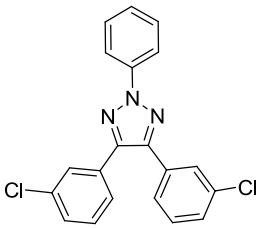
4,5-bis(4-chlorophenyl)-2-phenyl-2*H*-1,2,3-triazole (176b)[114]



Synthesised by general procedure B from tetrazole **175b** (200 mg, 0.78 mmol) over 16 h. Purified by automated reverse phase chromatography (Si C_{18}, 50-95% MeCN:H_2O, HCO_2H modifier, 30 min) to afford triazole **176b** (39 mg, 27%) as a yellow solid.

LC-MS Method A (2 min): rt = 1.67 min, $[M+H]^+$ = 366.1. mp 143-146 °C. IR (cm^{-1}): 3066, 1662, 1597, 1494, 1453, 1284, 1265, 1091. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.18 (td, *J* = 1.3, 7.6 Hz, 2 H), 7.59 (td, *J* = 2.0, 8.6 Hz, 4 H), 7.52 (dd, *J* = 7.3, 8.3 Hz, 2 H), 7.42 (td, *J* = 2.0, 8.6 Hz, 5 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 144.9 (2 C), 139.5, 134.9 (2 C), 129.7 (4 C), 129.3 (2 C), 129.0 (4 C), 129.0 (2 C) 127.7, 118.8 (2 C) ppm. HRMS (ESI+): m/z calcd for $C_{20}H_{14}Cl_2N_3$: 366.0559; found 366.0561. The data obtained were consistent with literature values.

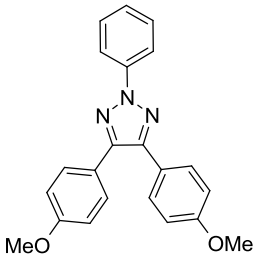
4,5-bis(3-chlorophenyl)-2-phenyl-2*H*-1,2,3-triazole (176c)



Synthesised by general procedure B from tetrazole **175c** (200 mg, 0.78 mmol) over 31 h. Purified by automated reverse phase chromatography (Si C_{18}, 10 g, 50-95% MeCN:H_2O, HCO_2H modifier, 30 min) and automated flash column chromatography (Si, 0-10% EtOAc-cyclohexane, 20 min) to afford triazole **176c** (40 mg, 28%) as a brown gum.

LC-MS Method A (2 min): rt = 1.67 min, $[M+H]^+$ = 366.1. IR (cm^{-1}): 3064, 1597, 1569, 1495, 1449, 1291, 1258, 1079. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.20 (dd, *J* = 1.0, 7.6 Hz, 2 H), 7.73 (t, *J* = 1.5 Hz, 2 H), 7.54 (t, *J* = 8.0 Hz, 2 H), 7.49 (dd, *J* = 1.0, 7.6 Hz, 2 H), 7.45 - 7.39 (m, 3 H), 7.36 (d, *J* = 7.8 Hz, 2 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 144.7 (2 C), 139.5, 134.7 (2 C), 132.2 (2 C), 129.9 (2 C), 129.3 (2 C), 129.0 (2 C), 128.4 (2 C), 127.8, 126.5 (2 C), 118.9 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{20}H_{14}Cl_2N_3$: 366.0559; found 366.0562.

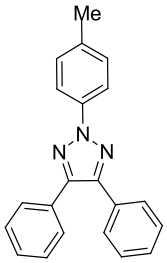
4,5-bis(4-methoxyphenyl)-2-phenyl-2*H*-1,2,3-triazole (176e)[114]



Synthesised by general procedure B from tetrazole **175e** (98 mg, 0.38 mmol) over 30 h. Purified by automated reverse phase chromatography (Si C_{18}, 30-80% MeCN:H_2O, HCO_2H modifier, 30 min) to afford triazole **176e** (25 mg, 38%) as a brown solid.

LC-MS Method A (2 min): rt = 1.48 min, $[M+H]^+$ = 358.2. mp 129-132 °C. IR (cm^{-1}): 2932, 1613, 1526, 1492, 1459, 1273, 1246, 1176, 1023. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.19 (d, *J* = 8.3 Hz, 2 H), 7.61 (app. d, *J* = 8.6 Hz, 4 H), 7.51 (t, *J* = 7.6 Hz, 2 H), 7.37 (t, *J* = 7.3 Hz, 1 H), 6.96 (app. d, *J* = 8.6 Hz, 4 H), 3.87 (s, 6 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 159.9 (2 C), 145.5 (2 C), 139.8, 129.7 (4 C), 129.2 (2 C), 127.1, 123.4 (2 C), 118.6 (2 C), 114.1 (4 C), 55.3 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{22}H_{20}N_3O_2$: 358.1550; found 358.1549. The data obtained were consistent with literature values.

4,5-diphenyl-2-(*p*-tolyl)-2*H*-1,2,3-triazole (176f)



Synthesised by general procedure B from tetrazole **175f** (200 mg, 0.84 mmol) over 19 h. Purified by automated reverse phase chromatography (Si C_{18}, 50-95% MeCN:H_2O, HCO_2H modifier, 30 min) to afford triazole **176f** (62 mg, 47%) as a brown gum.

LC-MS Method A (2 min): rt = 1.58 min, $[M+H]^+$ = 312.2. IR (cm^{-1}): 3029, 2920, 1510, 1460, 1441, 1375, 1293, 1268. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.10 (app. d, *J* = 8.6 Hz, 2 H), 7.72 - 7.65 (m, 4 H), 7.43 (dd, *J* = 1.9, 4.9 Hz, 6 H), 7.33 (app. d, *J* = 8.6 Hz, 2 H), 2.45 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 145.7 (2 C), 137.6, 137.3, 130.9 (2 C), 129.8 (2 C), 128.6 (4 C), 128.6 (2 C), 128.5 (4 C), 118.7 (2 C), 21.1 ppm. $^{15}N^1H$ HMBC NMR (60 Hz, $CDCl_3$): δ_N = 256.8 ppm. HRMS (ESI^+): m/z calcd for $C_{21}H_{18}N_3$: 312.1495; found 312.1496.

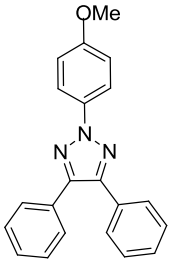
From hydrazonoyl chloride **215**

To a solution of hydrazonoyl chloride **215** (73 mg, 0.30 mmol) in THF (2 mL) was added 2M $NaOH_{(aq)}$ (1.5 mL, 3.0 mmol) and the reaction mixture stirred at room temperature for 30 min. The solvent was removed *in vacuo* and the residue partitioned between DCM (10 mL) and 1M $HCl_{(aq)}$ (10 mL). The layers were separated and the aqueous layer extracted with DCM (2 x 10 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$, separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 10 g, 0-10% TBME-cyclohexane,

60 min) and automated reverse phase chromatography (Si C_{18}, 80-99% MeCN-H_2O, HCO_2H modifier, 30 min) to afford triazole **176f** (13 mg, 29%) as a brown gum.

LC-MS Method A (2 min): rt = 1.60 min, $[M+H]^+$ = 312.1. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.08 (app. d, *J* = 8.6 Hz, 2 H), 7.71 - 7.63 (m, 4 H), 7.48-7.39 (m, 6 H), 7.33 (app. d, *J* = 8.3 Hz, 2 H), 2.45 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 145.7 (2 C), 137.6, 137.3, 130.9 (2 C), 129.8 (2 C), 128.6 (4 C), 128.6 (2 C), 128.4 (4 C), 118.7 (2 C), 21.1 ppm. Data agrees with that obtained from general procedure B.

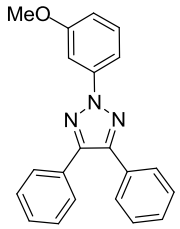
2-(4-methoxyphenyl)-4,5-diphenyl-2*H*-1,2,3-triazole (176g)[114]



Synthesised by general procedure B from tetrazole **175g** (200 mg, 0.79 mmol) over 10 h. Purified by automated flash column chromatography (Si, 10 g, 0-100% TBME-cyclohexane, 30 min) to afford triazole **176g** (80 mg, 62%) as an off-white solid.

LC-MS Method A (2 min): rt = 1.50 min, $[M+H]^+$ = 328.1. mp 97-100 °C. IR (cm^{-1}): 3055, 3013, 2935, 1511, 1461, 1440, 1297, 1252, 1175, 1074. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.11 (app. d, *J* = 9.1 Hz, 2 H), 7.70 - 7.62 (m, 4 H), 7.41 (dd, *J* = 1.8, 4.8 Hz, 6 H), 7.01 (app. d, *J* = 9.1 Hz, 2 H), 3.90 (s, 3 H) ppm. ^{13}C NMR (400 MHz, $CDCl_3$): δ_C = 159.0, 145.5 (2 C), 133.7, 131.0 (2 C), 128.6 (4 C), 128.5 (2 C), 128.4 (4 C), 120.2 (2 C), 114.4 (2 C), 55.6 ppm. $^{15}N^1H$ HMBC NMR (60 Hz, $CDCl_3$): δ_N = 256.3 ppm. HRMS (ESI^+): m/z calcd for $C_{21}H_{18}N_3O$: 328.1445; found 328.1444. The data obtained were consistent with literature values.

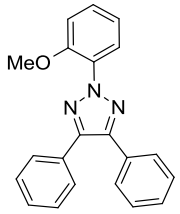
2-(3-methoxyphenyl)-4,5-diphenyl-2*H*-1,2,3-triazole (176h)



Synthesised by general procedure B from tetrazole **175h** (200 mg, 0.79 mmol) over 35 h. Purified by automated reverse phase chromatography (Si C_{18}, 50-99% MeCN:H_2O, TFA modifier, 30 min) to afford triazole **176h** (32 mg, 25%) as a brown gum.

LC-MS Method A (2 min): rt = 2 min, $[M+H]^+$ = 328.1. IR (cm^{-1}): 3063, 2980, 2836, 1610, 1595, 1494, 1463, 1438, 1295, 1249. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 7.81 (ddd, *J* = 1.0, 2.0, 8.1 Hz, 1 H), 7.78 (t, *J* = 2.1 Hz, 1 H), 7.71 - 7.63 (m, 4 H), 7.47 - 7.38 (m, 7 H), 6.94 (ddd, *J* = 1.0, 2.5, 8.3 Hz, 1 H), 3.94 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 160.4, 146.0 (2 C), 140.8, 130.8, 130.1 (2 C), 128.7 (2 C), 128.6 (4 C), 128.5 (4 C), 113.7, 111.1, 104.2, 55.6 ppm. HRMS (ESI^+): m/z calcd for $C_{21}H_{18}N_3O$: 328.1444; found 328.1446.

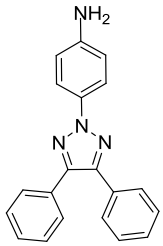
2-(2-methoxyphenyl)-4,5-diphenyl-2*H*-1,2,3-triazole (176i)



Synthesised by general procedure B from tetrazole **175i** (214 mg, 0.85 mmol) over 40 h. Purified by automated reverse phase chromatography (Si C_{18}, 50-99% MeCN:H_2O, TFA modifier, 30 min) to afford triazole **176i** (46 mg, 33%) as a colourless gum.

LC-MS Method A (2 min): rt = 1.32 min, $[M+H]^+$ = 328.3. IR (cm^{-1}): 3054, 2972, 2834, 1602, 1505, 1465, 1439, 1285, 1256, 1024. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 7.71 - 7.63 (m, 5 H), 7.47 (ddd, *J* = 1.1, 7.2, 8.3 Hz, 1 H), 7.43-7.38 (m, 6 H), 7.16 - 7.08 (m, 2 H), 3.93 (s, 3 H). ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 153.7, 145.5 (2 C), 130.9, 130.4 (2 C), 129.8, 128.5 (4 C), 128.5 (4 C), 128.4 (2 C), 127.3, 120.7, 112.9, 56.4 ppm. HRMS (ESI^+): m/z calcd for $C_{21}H_{18}N_3O$: 328.1444; found 328.1441.

4-(4,5-diphenyl-2*H*-1,2,3-triazol-2-yl)aniline (176j)[224]



Low dilution experiment

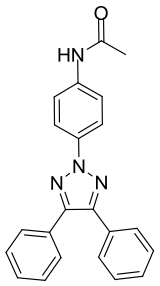
Synthesised by general procedure B from tetrazole **175j** (200 mg, 0.82 mmol) over 48 h. Purified by automated flash column chromatography (Si, 5 g, 0-25% EtOAc-cyclohexane, 40 min) and the solvent removed *in vacuo*. The residue was triturated with cyclohexane, filtered, washed with cyclohexane, and dried *in vacuo* to afford triazole **176j** (23 mg, 17%) as a pale yellow solid.

High dilution experiment

Synthesised by general procedure B from a solution of tetrazole **175j** (50 mg, 0.21 mmol) in 2-MeTHF:IPA (1:1 v/v, 25 mL) over 7 h. Purified by automated flash column chromatography (Si KPNH, 11 g, 0-25% TBME-cyclohexane, 60 min) to afford triazole **176j** (15 mg, 46%) as an off-white solid.

LC-MS Method A (2 min): rt = 1.26 min, $[M+H]^+$ = 313.3. mp 158-161 °C. IR (cm^{-1}): 3459, 3361, 1620, 1513, 1290. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.00 (app. d, *J* = 8.8 Hz, 2 H), 7.71 - 7.60 (m, 4 H), 7.46 - 7.35 (m, 6 H), 6.86 (app. d, *J* = 8.8 Hz, 2 H), 4.47 (br. s, 2 H). The data obtained agreed with those obtained for compound **116**.

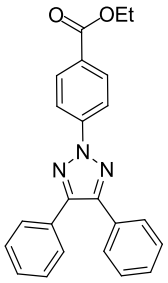
***N*-(4-(4,5-diphenyl-2*H*-1,2,3-triazol-2-yl)phenyl) acetamide (176k)**



Synthesised by general procedure B from tetrazole **175k** (100 mg, 0.36 mmol) over 6 h. Purified by automated reverse phase chromatography (Si C_{18}, 40-60% MeCN:H_2O, TFA modifier, 30 min) to afford triazole **176k** (26 mg, 41%) as a yellow solid.

LC-MS Method A (2 min): rt = 1.27 min, $[M+H]^+$ = 355.2. mp 198-200 °C. IR (cm^{-1}): 3278, 3267, 3069, 1668, 1612, 1561, 1512, 1416, 1313, 1268. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.15 (app. d, *J* = 8.8 Hz, 2 H), 7.74 - 7.58 (m, 6 H), 7.48 (br. s, 1 H), 7.45 - 7.37 (m, 6 H), 2.23 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 168.5, 145.9 (2 C), 137.1, 136.1, 130.8 (2 C), 128.6 (2 C), 128.6 (4 C), 128.2 (4 C), 120.4 (2 C), 119.4 (2 C), 24.6 ppm. HRMS (ESI^+): m/z calcd for $C_{22}H_{19}N_4O$: 355.1559; found 355.1555.

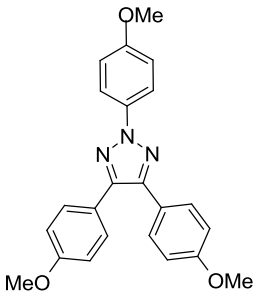
ethyl 4-(4,5-diphenyl-2*H*-1,2,3-triazol-2-yl)benzoate (176l)



Synthesised by general procedure B from tetrazole **175l** (200 mg, 0.68 mmol) over 23 h. Purified by automated reverse phase chromatography (Si C_{18}, 50-99% MeCN:H_2O, HCO_2H modifier, 20 min) to afford triazole **176l** (17 mg, 14%) as a yellow solid.

LC-MS Method A (2 min): rt = 1.38 min, $[M+H]^+$ = 370.2. mp 92-94 °C. IR (cm^{-1}): 2996, 1715, 1606, 1510, 1443, 1378, 1266, 1170, 1103. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.29 (app. d, *J* = 8.8 Hz, 2 H), 8.21 (app. d, *J* = 8.8 Hz, 2 H), 7.75 - 7.64 (m, 4 H), 7.52 - 7.39 (m, 6 H), 4.44 (q, *J* = 7.1 Hz, 2 H), 1.46 (t, *J* = 7.1 Hz, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 165.9, 146.8 (2 C), 142.6, 130.9 (2 C), 130.4 (2 C), 129.1, 128.9 (2 C), 128.7 (4 C), 128.5 (4 C), 118.2 (2 C), 61.2, 14.4 ppm. HRMS (ESI^+): m/z calcd for $C_{23}H_{20}N_3O_2$: 370.1550; found 370.1554.

2,4,5-tris(4-methoxyphenyl)-2*H*-1,2,3-triazole (176o)[113]

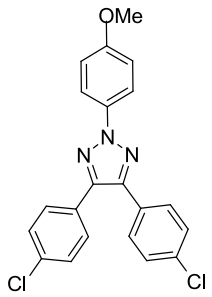


Synthesised by general procedure B from tetrazole **175o** (200 mg, 0.71 mmol) over 7 h, during which time the product precipitated from solution. The product was

filtered, washed with MeOH, and dried *in vacuo* to afford triazole **176o** (123 mg, 90%) as a brown solid.

LC-MS Method A (2 min): rt = 1.45 min, $[M+H]^+$ = 388.1. mp 145-147 °C. IR (cm^{-1}): 3001, 2957, 2835, 1602, 1579, 1512, 1499, 1459, 1323, 1299, 1249, 1177, 1145, 1030. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.07 (td, *J* = 2.0, 9.0 Hz, 2 H), 7.57 (td, *J* = 2.0, 8.8 Hz, 4 H), 7.01 (td, *J* = 2.2, 9.0 Hz, 2 H), 6.93 (td, *J* = 2.2, 9.0 Hz, 4 H), 3.88 (s, 3 H), 3.86 (s, 6 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 159.8 (2 C), 158.8, 145.0 (2 C), 133.7, 129.7 (4 C), 123.5 (2 C), 120.1 (2 C), 114.3 (2 C), 114.0 (4 C), 55.6, 55.3 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{23}H_{22}N_3O_3$: 388.1656; found 388.1648. The data obtained were consistent with literature values.

4,5-bis(4-chlorophenyl)-2-(4-methoxyphenyl)-2*H*-1,2,3-triazole (176p)[114]

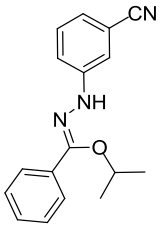


Synthesised by general procedure B from tetrazole **175p** (200 mg, 0.70 mmol) over 16 h, during which time the product precipitated from solution. The product was filtered, triturated with MeOH, filtered, and dried *in vacuo* to afford triazole **176p** (93 mg, 67%) as a brown solid.

LC-MS Method A (2 min): rt = 1.66 min, $[M+H]^+$ = 396.2. mp 134-136 °C. IR (cm^{-1}): 2957, 2924, 2837, 1603, 1508, 1450, 1394, 1250, 1090. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 8.08 (td, *J* = 2.0, 9.1 Hz, 2 H), 7.58 (td, *J* = 1.8, 8.3 Hz, 4 H), 7.41 (app. d, *J* = 8.6 Hz, 4 H), 7.04 (app. d, *J* = 9.1 Hz, 2 H), 3.90 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 159.2, 144.4 (2 C), 134.7 (2 C), 133.3, 129.6 (4 C), 129.2 (2 C), 129.0 (4 C), 120.3 (2 C), 114.4 (2 C), 55.6 ppm. HRMS (ESI^+): m/z calcd for

$C_{21}H_{16}Cl_2N_3O$: 396.0665; found 396.0669. The data obtained were consistent with literature values.

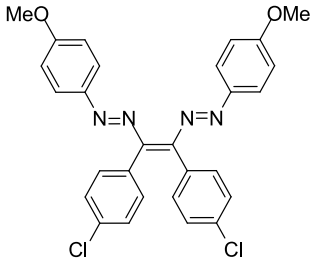
(*Z*)-isopropyl *N*'-(3-cyanophenyl)benzohydrazonate (178)



Isolated as a byproduct from tetrazole **175m** (200 mg, 0.81 mmol) using general procedure B over 24 h. Purified by automated reverse phase chromatography (Si C_{18}, 40-80% MeCN:H_2O, HCO_2H modifier, 10 CV) to afford **178** (61 mg, 54%) as a white solid.

LC-MS Method A (2 min): rt = 1.32 min, $[M+H]^+$ = 280.2. mp 83-86 °C. IR (cm^{-1}): 3307, 2977, 2228, 1599, 1509, 1491, 1439, 1315, 1280, 1266, 1085, 1066. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.20 - 8.16 (br. s, 1 H), 7.68 (dd, *J* = 2.3, 7.8 Hz, 2 H), 7.49 - 7.41 (m, 4 H), 7.35 - 7.23 (m, 2 H), 7.08 (td, *J* = 1.5, 7.1 Hz, 1 H), 4.50 (spt, *J* = 6.3 Hz, 1 H), 1.37 (d, *J* = 6.1 Hz, 6 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 147.3, 145.5, 131.6, 129.8, 129.4, 128.6 (2 C), 127.0 (2 C), 122.6, 119.4, 116.9, 115.6, 112.9, 73.5, 22.6 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{17}H_{18}N_3O$: 280.1444; found 280.1440.

(*Z*)-1,2-bis(4-chlorophenyl)-1,2-bis((*E*)-(4-methoxyphenyl)diazenyl)ethane (181)

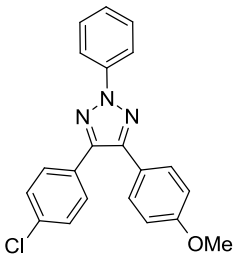


Intermediate isolated from tetrazole **175p** (200 mg, 0.70 mmol) using general procedure B over 23 h, during which time the intermediate precipitated from solution. The product was filtered, washed with MeOH, and dried *in vacuo* to afford **181** (103 mg, 57%) as a brown solid.

LC-MS Method A (2 min): rt = 1.67 min, $[M+H]^+$ = N/A. mp (decomposed). IR (cm^{-1}): 2937, 2385, 1600, 1579, 1493, 1402, 1324, 1296, 1252, 1145, 1087. ^{1}H NMR (600 MHz, $CDCl_3$): δ_H = 7.50 (app. d, *J* = 8.8 Hz, 4 H), 7.33 (app. d, *J* = 8.1 Hz, 4 H), 7.21 - 7.19 (m, 4 H), 6.84 (app. d, *J* = 8.6 Hz, 4 H), 3.79 (s, 6 H) ppm. ^{13}C NMR (151 MHz, $CDCl_3$): δ_C = 162.6 (2 C), 155.3 (2 C), 148.1 (2 C), 133.5 (4 C), 133.5 (4 C), 131.9 (2 C), 127.3 (4 C), 125.6 (4 C), 114.1 (2 C), 55.6 (2 C) ppm.

Note: The intermediate did not give a clear value for accurate mass determination by LC-MS or HRMS.

4-(4-chlorophenyl)-5-(4-methoxyphenyl)-2-phenyl-2*H*-1,2,3-triazole (191)[114]

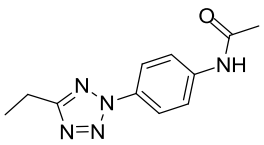


Synthesised by general procedure B from a mixture of tetrazole **175b** (340 mg, 1.3 mmol) and tetrazole **175e** (70 mg, 0.28 mmol) over 24 h. An LC-MS ratio of 7:11:10 was observed for the three triazole products **176b**, **176e** and **191**,

respectively. The solvent was removed *in vacuo* and the residue dissolved in MeOH:DMSO (1:1 v/v, 1 mL) then purified by automated reverse phase chromatography (Si C_{18}, 12 g, 80-99% MeCN-H_2O, HCO_2H modifier, 10 CV) to afford triazole **191** (18 mg, 18%) as an off-white gum.

LC-MS Method A: rt = 1.61 min, $[M+H]^+$ = 362.3. IR (cm^{-1}): 2834, 1613, 1597, 1495, 1457, 1371, 1299, 1246, 1175, 1091. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 8.18 (dd, *J* = 8.7, 1.1 Hz, 2 H), 7.62 (td, *J* = 8.6, 2.0 Hz, 2 H), 7.57 (td, *J* = 8.8, 2.0 Hz, 2 H), 7.52 (t, *J* = 8.0 Hz, 2 H), 7.42 - 7.37 (m, 3 H), 6.97 (td, *J* = 8.8, 2.3 Hz, 2 H), 3.88 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 160.1, 145.9, 144.5, 139.7, 134.5, 129.8 (2 C), 129.6 (2 C), 129.5, 129.3 (2 C), 128.8 (2 C), 127.4, 122.9, 118.7 (2 C), 114.2 (2 C), 55.3 ppm. HRMS (ESI^+): m/z calcd for $C_{21}H_{17}ClN_3O$: 362.1053; found 362.1055. The data obtained were consistent with literature values.

_N_-(4-(5-ethyl-2_H_-tetrazol-2-yl)phenyl)acetamide (193)

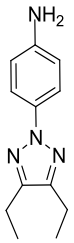


To a solution of tetrazole **192** (92 mg, 0.49 mmol) in dry DCM (4 mL) was added acetic anhydride (0.5 mL, 5.3 mmol). The reaction mixture was stirred under a nitrogen atmosphere at room temperature for 5 h. The reaction mixture was washed with sat. $NaHCO_{3(aq)}$ (10 mL), the layers separated, and the aqueous layer extracted with DCM (2 x 10 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (10 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* to afford tetrazole **193** (110 mg, 98%) as a white solid.

LC-MS Method A (2 min): rt = 0.78 min, $[M+H]^+$ = 232.2. mp 133-135 °C. IR (cm^{-1}): 3275, 3210, 3144, 3077, 2980, 2937, 1669, 1609, 1546, 1516, 1419, 1370, 1311, 1262, 1178, 1002. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 8.08 (app. d, *J* = 8.8 Hz, 2 H), 7.72 (app. d, *J* = 8.8 Hz, 2 H), 7.48 - 7.33 (br. s, 1 H), 3.03 (d, *J* = 7.6 Hz, 2 H), 2.25 (s, 3 H),

1.46 (t, *J* = 7.7 Hz, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ_C = 167.5, 167.1, 138.0, 131.9, 119.5 (2 C), 119.3 (2 C), 23.6, 18.1, 11.3 ppm. HRMS (ESI^+): m/z calcd for $C_{11}H_{14}N_5O$: 232.1193; found 232.1188.

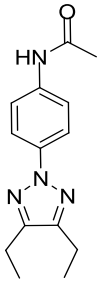
4-(4,5-diethyl-2*H*-1,2,3-triazol-2-yl)aniline (194)



Synthesised by general procedure B from tetrazole **192** (100 mg, 0.53 mmol) over 20 h. Purified by automated reverse phase chromatography (Si C_{18}, 10-30% MeCN:H_2O, TFA modifier, 20 min). The solvent was removed *in vacuo*, the residue partitioned between DCM (20 mL) and 2M $NaOH_{(aq)}$ (20 mL), and the organic layer separated. The organic phase was washed with sat. $NaCl_{(aq)}$ (20 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* to afford triazole **194** (13 mg, 23%) as a brown gum.

LC-MS Method A (2 min): rt = 0.90 min, $[M+H]^+$ = 217.4. IR (cm^{-1}): 3353, 2972, 2932, 1627, 1520, 1457, 1285, 1172, 1064. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 7.77 (app. d, *J* = 8.8 Hz, 2 H), 6.75 (app. d, *J* = 8.8 Hz, 2 H), 3.73 (br. s., 2 H), 2.72 (q, *J* = 7.6 Hz, 4 H), 1.32 (t, *J* = 7.7 Hz, 6 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 147.5 (2 C), 145.1, 132.6, 119.8 (2 C), 115.3 (2 C), 18.2 (2 C), 13.7 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{12}H_{17}N_4$: 217.1448; found 217.1444.

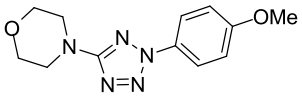
_N_-(4-(4,5-diethyl-2_H_-1,2,3-triazol-2-yl)phenyl) acetamide (195)



Synthesised by general procedure B from tetrazole **193** (105 mg, 0.46 mmol) over 20 h. Purified by automated reverse phase chromatography (Si C_{18}, 10-70% MeCN:H_2O, TFA modifier, 30 min) to afford triazole **195** (12 mg, 21%) as a brown gum.

LC-MS Method A (2 min): rt = 1.02 min, $[M+H]^+$ = 259.3. IR (cm^{-1}): 3299, 3143, 3076, 2973, 2934, 1667, 1611, 1549, 1518, 1417, 1310, 1266. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 7.95 (app. d, *J* = 8.8 Hz, 2 H) 7.60 (app. d, *J* = 9.1 Hz, 2 H) 7.29 - 7.34 (br. s, 1 H) 2.73 (q, *J* = 7.6 Hz, 4 H) 2.21 (s, 3 H) 1.33 (t, *J* = 7.6 Hz, 6 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 168.2, 148.4 (2 C), 136.4, 136.2, 120.3 (2 C), 118.8 (2 C), 24.6, 18.3 (2 C), 13.5 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{14}H_{19}N_4O$: 259.1553; found 259.1545.

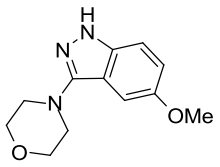
4-(2-(4-methoxyphenyl)-2_H_-tetrazol-5-yl)morpholine (197)



Synthesised by general procedure A using 4-(1*H*-tetrazol-5-yl)morpholine (1.16 g, 7.5 mmol), (4-methoxyphenyl)boronic acid (2.27 g, 15.0 mmol), and Cu_2O (53 mg, 0.37 mmol) over 16 h. Purified by automated flash column chromatography (Si NH_2, 50 g, 0-25% DCM-cyclohexane, 40 min) to afford tetrazole **197** (932 mg, 48%) as a white solid.

LC-MS Method A (2 min): rt = 0.99 min, $[M+H]^+$ = 262.0. mp 153-156 °C. IR (cm^{-1}): 2968, 2901, 2864, 1567, 1513, 1441, 1400, 1381, 1280, 1251, 1173, 1112, 1020. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 7.96 (dt, *J* = 9.1, 2.0 Hz, 2 H) 7.02 (dt, *J* = 9.1, 2.0 Hz, 2 H) 3.90 (s, 3 H) 3.88 (t, *J* = 5.5 Hz, 4 H) 3.59 (t, *J* = 4.8 Hz, 4 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 169.2, 160.0, 130.7, 120.8 (2 C), 114.5 (2 C), 66.3 (2 C), 55.6, 47.0 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{12}H_{16}N_5O_2$: 262.1299; found 262.1292.

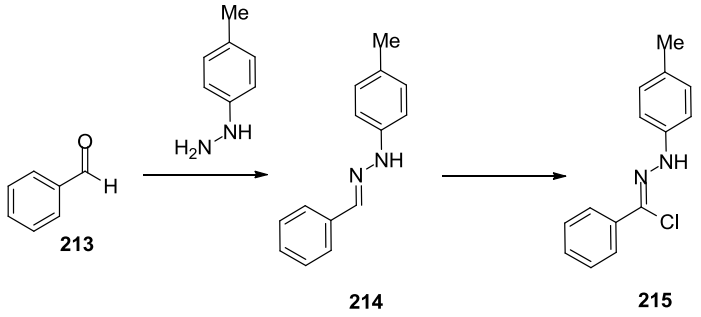
4-(5-methoxy-1*H*-indazol-3-yl)morpholine (198)[225]



Isolated as a byproduct from tetrazole **197** (204 mg, 0.78 mmol) using general procedure B over 7 h. Purified by automated flash column chromatography (Si, 10 g, 0-50% EtOAc-cyclohexane, 40 min) to afford **198** (144 mg, 79%) as a pink glass.

LC-MS Method A (2 min): rt = 0.68 min, $[M+H]^+$ = 234.3. IR (cm^{-1}): 3196, 3051, 2960, 2911, 2855, 2830, 1514, 1491, 1449, 1329, 1296, 1275, 1251, 1224, 1183, 1111, 1025. ^1H NMR (600 MHz, $CDCl_3$): δ_H = 9.14 (br. s., 1 H), 7.27 (d, *J* = 9.0 Hz, 1 H), 7.06 (dd, *J* = 9.0, 2.2 Hz, 1 H), 7.03 (d, *J* = 2.2 Hz, 1 H), 3.93 - 3.97 (m, 4 H), 3.86 (s, 3 H), 3.40 - 3.43 (m, 4 H) ppm. ^{13}C NMR (151 MHz, $CDCl_3$): δ_C = 153.7, 152.8, 138.5, 118.9, 115.1, 111.0, 100.8, 66.8 (2 C), 55.9, 50.3 (2 C) ppm. HRMS (ESI^+): m/z calcd for $C_{12}H_{16}N_3O_2$: 234.1243; found 234.1236. The data obtained were consistent with literature values.

(*Z*)-*N*'-(*p*-tolyl)benzohydrazonoyl chloride (215)



(*E*)-1-benzylidene-2-(p-tolyl)hydrazine **214**[226]

To a solution of benzaldehyde **213** (0.50 mL, 4.9 mmol) in EtOH (10 mL) was added *p*-tolylhydrazine (600 mg, 4.9 mmol) portionwise. The reaction mixture was stirred at 70 °C for 1.5 h, cooled to room temperature and water (20 mL) added to form a precipitate. The precipitate was filtered, washed with water, and dried *in vacuo* to afford **214** (731 mg, 71%) as a white solid.

LC-MS Method A (2 min): rt = 1.33 min, $[M+H]^+$ = 211.1. mp 106-109 °C. IR (cm^{-1}): 3308, 3022, 2915, 1614, 1596, 1564, 1522, 1444, 1354, 1257, 1134, 1105, 1071. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 7.70 (s, 1 H), 7.67 (app. d, *J* = 7.3 Hz, 2 H), 7.39 (t, *J* = 7.3 Hz, 2 H), 7.32 (t, *J* = 7.3 Hz, 1 H), 7.11 (d, *J* = 8.1 Hz, 2 H), 7.05 (app. d, *J* = 8.3 Hz, 2 H), 2.32 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 142.5, 136.8 (2 C), 135.4, 129.8 (2 C), 128.6 (2 C), 128.2, 126.1 (2 C), 112.8 (2 C), 20.6 ppm. HRMS (ESI^+): m/z calcd for $C_{14}H_{15}N_2$: 211.1230; found 211.1230. The data obtained were consistent with literature values.

Note: Exchangeable NH proton missing in the ^1H NMR spectrum.

(*Z*)-*N*'-(p-tolyl)benzohydrazonoyl chloride **215**[227]

A solution of *N*-chlorosuccinimide (216 mg, 1.6 mmol) in DCM (12 mL) was cooled to 0 °C. Dimethyl sulfide (0.21 mL, 2.9 mmol) was added under a nitrogen atmosphere over 10 min, during which time a white precipitate formed. The reaction mixture was cooled to -78 °C and a solution of **214** (200 mg, 0.95 mmol) in DCM (1.5 mL)

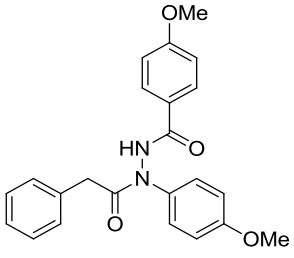
was added dropwise. The reaction mixture was stirred at - 78 °C for 2 h, warmed to room temperature, and stirred for a further 1 h. The reaction mixture was partitioned between DCM (15 mL) and water (15 mL). The organic layer was separated and washed with water (15 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 10 g, 0-25% DCM-cyclohexane, 20 min) to afford **215** (158 mg, 68%) as a pale pink solid.

LC-MS Method A (2 min): rt = 1.43 min, $[M+H]^+$ = N/A. mp 120-121 °C. IR (cm^{-1}): 3317, 3054, 3031, 2915, 2855, 1612, 1567, 1516, 1472, 1446, 1315, 1263, 1229, 1135, 1102, 1075. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 8.00 (s, 1 H), 7.94 (dd, *J* = 1.5, 8.3 Hz, 2 H), 7.46 - 7.34 (m, 2 H), 7.46 - 7.34 (m, 1 H), 7.17 - 7.08 (m, 4 H), 2.34 (s, 3 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 141.1, 134.6, 130.5, 129.9 (2 C), 129.0, 129.0, 128.4 (2 C), 126.3 (2 C), 113.4 (2 C), 20.6 ppm. The data obtained were consistent with literature values.

Note: The product did not give a clear value for accurate mass determination by LC-MS or HRMS.

3.10.3. Chapter 1: Photochemical labelling of peptide carboxylic acids

4-methoxy-*N*-(4-methoxyphenyl)-*N*-(2-phenylacetyl)benzohydrazide (224)



A solution of 2-phenylacetic acid **223** (25 mg, 0.18 mmol) and tetrazole **175o** (52 mg, 0.18 mmol) in TFE (3 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 5 h. The solvent was removed *in vacuo* and the

crude material dissolved in DMF (1 mL) then purified by MDAP Method D (Si C_{18}, 15-85% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min) to afford **224** (55 mg, 77%) as a fluffy white solid.

LC-MS Method F (2 min): rt = 1.09 min, $[M+H]^+$ = 391.3. IR (cm^{-1}): 3260, 3008, 2964, 2889, 1656, 1654, 1606, 1579, 1508, 1455, 1442, 1375, 1300, 1249, 1175, 1029. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 9.17 (br. s., 1 H), 7.73 (app. d, *J* = 8.3 Hz, 2 H), 7.43 - 7.32 (m, 3 H), 7.31-7.27 (m, 2 H), 7.16 (app. d, *J* = 6.8 Hz, 2 H), 6.88 (app. d, *J* = 8.8 Hz, 2 H), 6.80 (t, *J* = 8.1 Hz, 2 H), 3.84 (s, 3 H), 3.82 (s, 3 H), 3.63 (s, 2 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$): δ_C = 171.4, 165.7, 162.5, 159.7, 134.6, 134.5, 129.7 (2 C), 129.4 (2 C), 129.3 (2 C), 128.4 (2 C), 126.9, 123.9, 114.5 (2 C), 113.6 (2 C), 55.5, 55.3, 40.4 ppm. HRMS (ESI^+): m/z calcd for $C_{23}H_{23}N_2O_4$: 391.1662; found 391.1677.

General synthesis of HMBA bound peptides using SPPS

Peptides were synthesised using solid phase synthesis on the stated resin. Manual couplings were performed in a Merrifield bubbler with vacuum and nitrogen lines attached to a three way valve. Automated couplings were carried out using a Protein Technologies Tribute Automated Synthesiser. All reagents and solvents were of peptide synthesis grade and were obtained from Rathburn Chemicals.

Precipitation of the peptides was conducted using an Eppendorf Minispin or Eppindorf 5804 centrifuge for small-scale and full-scale cleavage mixtures, respectively. Some peptide purification was conducted with a reverse-phase HPLC Gilson system with a 322 pump, coupled to a 151 UV/Vis Spectrometer, 234 Autoinjector and a GX-271 liquid handler using an Agilent Zorbax SB-C18 21.2x100mm column at room temperature and a 10 mL min^{-1} flow rate. All other purification was conducted using MDAP purification (See Section 3.4.). Lyophilisation was carried out with a Christ Alpha 1-2 LD freeze drying system.

Fmoc Loading Test

To two 10 mL volumetric flasks was added an accurately weighed portion (*ca.* 10 mg) of loaded resin. Fmoc deprotection was carried out by accurately filling the flasks with 20% piperidine in DMF (10 mL) and sonicating the suspension for 15 minutes. The UV absorption of the samples was measured at 302 nm (Varian Cary 50 probe UV-visible spectrometer, Varian Cary software), relative to a blank solution of 20% piperidine in DMF. The loading of the resin was calculated using the equation below. The two values were averaged to give the resin loading.

$$\text{Loading} = \frac{\text{A} \times 10}{\text{m} \times 7.8}$$

where A = absorption at 302 nm

m = mass of resin/ mg

Small-scale Cleavage

To a sample of resin (*ca.* 5 mg) was added a cleavage mixture consisting of TFA:H_2O:TIPS (95:2.5:2.5 v/v, 250 µL) and the resin shaken vigorously for 1 h. The solution was decanted into a microcentrifuge tube containing cold Et_2O (1 mL) to precipitate the peptide. The precipitate was centrifuged at 11,000 rpm for 3 min and the supernatant discarded. The precipitate was washed with cold Et_2O (1 mL) and centrifuged twice more to give the crude peptide which was dissolved in water:MeCN (1:1 v/v) and filtered for analysis by LC-MS.

Note: Peptides containing cysteine residues were cleaved using a mixture consisting of TFA:H_2O:TIPS:DODT (92.5:2.5:2.5:2.5) and DMSO was avoided as a solvent for all subsequent manipulations.

Full-scale Cleavage

A solution of TFA:H_2O:TIPS (95:2.5:2.5 v/v, 10 mL) was added to the resin and the suspension shaken vigorously for 4 h. The cleavage mixture was filtered through a frit with dropwise addition to cold Et_2O (30 mL) to precipitate the peptide. The

precipitate was centrifuged at 5000 rpm for 5 min and the supernatant carefully discarded by pipette. The precipitate was washed with cold Et_2O (30 mL) and centrifuged twice more to give the crude peptide which was dissolved in the minimum amount of water:MeCN (varied ratios) for purification by automated reverse phase chromatography. The appropriate fractions were combined and the solvent removed by lyophilisation to give the desired peptide as a white solid.

Note: Peptides containing cysteine residues were cleaved using a mixture consisting of TFA:H_2O:TIPS:DODT (92.5:2.5:2.5:2.5) and DMSO was avoided as a solvent for all subsequent manipulations to minimise disulfide formation.

Table 19: Manual loading of Rink-amide HMBA resin.

Cycle	Reagents	Solvent	Volume/ mL	Time/ min	Mixing	Iterations
Swell	-	DCM	5	10	N_2 Bubbling	1
Wash	-	DMF	5	0.5	N_2 Bubbling	5
Deprotection	20% piperidine in DMF	-	5	5	N_2 Bubbling	3
Wash	-	DMF	5	0.5	N_2 Bubbling	5
Coupling	HATU (5 equiv.) Fmoc-Gly-OH (5 equiv.) DIPEA (10 equiv.)	DMF	5	240	N2 Bubbling	1
Wash	-	DMF	5	0.5	N_2 Bubbling	5
Wash	-	DCM	5	0.5	N_2 Bubbling	5
Vacuum dry	-	-	-	10	-	-

Table 20: Automated addition of first amino acid to loaded resin.

Cycle	Reagents	Solvent	Volume/ mL	Time/ min	Mixing	Iterations
Swell	-	DCM	5	10	Shaking	1
Wash	-	DCM	5	0.5	Shaking	5
Wash	-	DMF	5	0.5	Shaking	5
Deprotection	20% piperidine in DMF	-	5	5	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Coupling	HATU (5 equiv.) Amino acid (5 equiv.)	DMF	5	20	Shaking	1
	17% NMM in DMF	-	5			
Wash	-	DMF	5	0.5	Shaking	5

Table 21: Automated addition of subsequent amino acids.

Cycle	Reagents	Solvent	Volume/ mL	Time/ min	Mixing	Iterations
Deprotection	20% piperidine in DMF	-	5	5	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Coupling	HATU (5 equiv.) Amino acid (5 equiv.)	DMF	5	20	Shaking	1
	17% NMM in DMF		5			
Wash	-	DMF	5	0.5	Shaking	5

Table 22: Automated addition of final amino acid with final deprotection step.

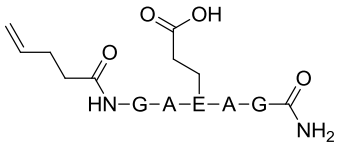
Cycle	Reagents	Solvent	Volume/ mL	Time/ min	Mixing	Iterations
Deprotection	20% piperidine in DMF	-	5	5	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Coupling	HATU (5 equiv.) Amino acid (5 equiv.) 17% NMM in DMF	DMF -	5 5	20	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Deprotection	20% piperidine in DMF	-	5	5	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Wash	-	DCM	5	0.5	Shaking	5
Drain dry	-	-	-	10	-	-

Table 23: Automated addition of final amino acid with final capping step.

Cycle	Reagents	Solvent	Volume/ mL	Time/ min	Mixing	Iterations
Deprotection	20% piperidine in DMF	-	5	5	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Coupling	HATU (5 equiv.) Amino acid (5 equiv.) 17% NMM in DMF	DMF -	5 5	20	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5

Deprotection	20% piperidine in DMF	-	5	5	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Capping	0.5 M Ac_2O in DMF	-	5	5	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Wash	-	DCM	5	0.5	Shaking	5
Drain dry	-	-	-	10	-	-

Alkene labelled Peptide (233)



100-200 mesh MBHA resin **227** (0.59 mmol g^{-1}, 1.20 g, 0.71 mmol) was loaded using the appropriate resin loading procedure (see Table 19) with a pre-activated solution of Fmoc-Gly-OH (1.05 g, 3.5 mmol), DIPEA (1.23 mL, 7.2 mmol) and HATU (1.35 g, 3.5 mmol) in DMF (5 mL). The resin loading was determined to be 0.32 mmol g^{-1} by UV analysis of the fulvene-piperidine adduct (see Fmoc Loading Test).

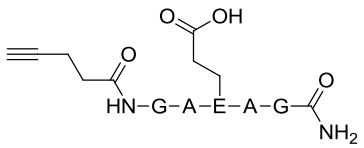
The resin was split into two batches (0.32 mmol g^{-1}, 600 mg, 0.19 mmol) and underwent automated peptide synthesis using the appropriate procedures (see Table 20-22) with Fmoc-Ala-OH (296 mg, 0.95 mmol) Fmoc-Glu(^{t}Bu)-OH (404 mg, 0.95 mmol), Fmoc-Gly-OH (283 mg, 0.95 mmol) and HATU (362 mg, 0.95 mmol). A small-scale cleavage confirmed the successful synthesis of the peptide for both batches, which were subsequently re-combined.

Two thirds of the resin (0.32 mmol g^{-1}, 800 mg, 0.25 mmol) was washed with DCM (5 x 30 sec), allowed to swell for 10 min then washed with DMF (5 x 30 sec). A pre-activated solution of 4-pentenoic acid (25 mg, 0.25 mmol), HATU

(97 mg, 0.25 mmol) and DIPEA (24 µL, 0.25 mmol) in DMF (5 mL) was added to the resin and the reaction mixture shaken for 4 h. After a full cleavage the precipitated peptide was dissolved in the minimal amount of MeCN-H_2O (1:1 v/v) and purified by reverse phase chromatography (Si C_{18}, 10-40% MeCN-H_2O, TFA modifier, 10 min, HPLC Gilson system) and the solvent removed by lyophilisation to afford peptide **233** (10 mg, 8%) as a white solid.

LC-MS Method F (3 min): rt = 0.40 min, $[M+H]^+$ = 485.3. ^{1}H NMR (400 MHz, DMSO-d_6): δ_H = 12.00 (br. s., 1 H), 8.11 (t, *J* = 5.8 Hz, 1 H), 8.08 - 7.98 (m, 3 H), 7.91 (d, *J* = 6.5 Hz, 1 H), 7.15 (br. s., 1 H), 7.06 (br. s., 1 H), 5.81 (ddd, *J* = 17.1, 10.5, 1.5 Hz, 1 H), 5.03 (dd, *J* = 17.1, 1.5 Hz, 1 H), 4.95 (dd, *J* = 10.0, 1.5 Hz, 1 H), 4.33 - 4.14 (m, 3 H), 3.70 (d, *J* = 5.5 Hz, 2 H), 3.65 (d, *J* = 5.8, 1 H), 3.62 (d, *J* = 5.6, 1 H), 2.30 - 2.18 (m, 6 H), 2.01 - 1.69 (m, 2 H), 1.24, (d, *J* = 7.0 Hz, 3 H) 1.22 (d, *J* = 7.0 Hz, 3 H) ppm. HRMS (ESI^+): m/z calcd for $C_{20}H_{33}N_6O_8$: 485.2354; found 485.2372.

Alkyne labelled Peptide (234)



100-200 mesh MBHA resin **227** (0.59 mmol g^{-1}, 1.20 g, 0.71 mmol) was loaded using the appropriate resin loading procedure (see Table 19) with a pre-activated solution of Fmoc-Gly-OH (1.05 g, 3.5 mmol), DIPEA (1.23 ml, 7.1 mmol) and HATU (1.35 g, 3.5 mmol) in DMF (5 mL). The resin loading was determined to be 0.30 mmol g^{-1} by UV analysis of the fulvene-piperidine adduct (see Fmoc Loading Test).

A portion of the resin (0.30 mmol g^{-1}, 600 mg, 0.18 mmol) underwent automated peptide synthesis using the appropriate procedures (see Table 20-22) with Fmoc-Ala-OH (277 mg, 0.89 mmol) Fmoc-Glu(^{t}Bu)-OH (379 mg, 0.89 mmol),

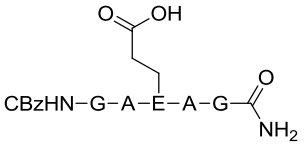
Fmoc-Gly-OH (265 mg, 0.89 mmol) and HATU (338 mg, 0.89 mmol). A small-scale cleavage confirmed the successful synthesis of the peptide.

The resin (0.30 mmol g^{-1}, 600 mg, 0.18 mmol) was washed with DCM (5 x 30 sec), allowed to swell for 10 min then washed with DMF (5 x 30 sec). A pre-activated solution of 4-pentynoic acid (87 mg, 0.89 mmol), HATU (338 mg, 0.89 mmol) and DIPEA (180 µL, 1.8 mmol) in DMF (5 mL) was added to the resin and the reaction mixture shaken for 1 h. The resin was filtered and shaken with a fresh solution of reagents for an additional 1 h. After a full cleavage the precipitated peptide was dissolved in the minimal amount of DMSO and purified by MDAP Method D (Si C_{18}, 0-20% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min) and the solvent removed by lyophilisation to afford peptide **234** (69 mg, 82%) as a white solid.

LC-MS Method D (3 min): rt = 0.39 min, $[M+H]^+$ = 483.2. IR (cm^{-1}): 3281, 3072, 2982, 2937, 1644, 1625, 1526, 1445, 1419, 1239, 1173, 1024. ^1H NMR (600 MHz, DMSO-d_6): δ_H = 2.73 (s, 1 H), 2.34 - 2.36 (m, 2 H), 2.34 - 2.36 (m, 2 H) ppm. ^{13}C NMR (151 MHz, DMSO-d_6): δ_C = 170.7, 83.8, 71.2, 33.9, 14.0 ppm. HRMS (ESI^+): m/z calcd for $C_{20}H_{31}N_6O_8$: 483.2198; found 483.2201.

Note: The alkyne peak was not observed in the weak IR spectrum. 2D $^1H^{13}C$ HSQC and HMBC spectra were obtained to identify the presence of the key alkyne peaks but were not analysed further.

CBz protected peptide (235)



100-200 mesh MBHA resin **227** (0.59 mmol g^{-1}, 603 mg, 0.36 mmol) was loaded using the appropriate resin loading procedure (see Table 19) with a pre-activated solution of Fmoc-Gly-OH (529 mg, 1.8 mmol), HATU (677 mg, 1.8 mmol) and DIPEA (0.62 mL, 3.6 mmol) in DMF (5 mL). The resin loading was determined to be 0.30 mmol g^{-1} by UV analysis of the fulvene-piperidine adduct (see Fmoc Loading Test).

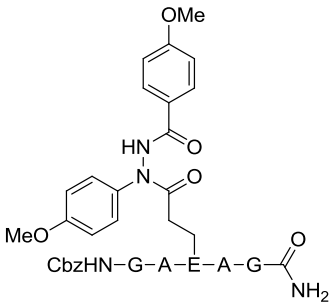
The resin (0.30 mmol g^{-1}, 603 mg, 0.18 mmol) underwent automated peptide synthesis using the appropriate procedures (see Table 20-22) with Fmoc-Ala-OH (266 mg, 0.86 mmol) Fmoc-Glu(^{t}Bu)-OH (364 mg, 0.86 mmol), Fmoc-Gly-OH (254 mg, 0.86 mmol) and HATU (325 mg, 0.86 mmol). A small-scale cleavage confirmed the successful synthesis of the peptide.

The resin (0.30 mmol g^{-1}, 603 mg, 0.18 mmol) was washed with DCM (5 x 30 sec), allowed to swell for 10 min then washed with DMF (5 x 30 sec). A pre-activated solution of CBz-OSu (128 mg, 0.51 mmol) and DIPEA (89 µL, 0.51 mmol) in DMF (15 mL) was added to the resin and the reaction mixture shaken for 4 h. After a full cleavage the precipitated peptide was dissolved in the minimal amount of MeCN-H_2O (1:1 v/v) and purified by automated reverse phase chromatography (Si C_{18}, 10-60% MeCN-H_2O, TFA modifier, 15 min, HPLC Gilson system) and the solvent removed by lyophilisation to afford peptide **235** (27 mg, 29%) as a white solid.

LC-MS Method F (2 min): rt = 0.54 min, $[M+H]^+$ = 537.3, HRMS (ESI^+): m/z calcd for $C_{23}H_{33}N_6O_9$: 537.2309; found 537.2314.

Note: The product was isolated in only 78% purity, as determined by LC-MS.

Hydrazide labelled peptide (236)

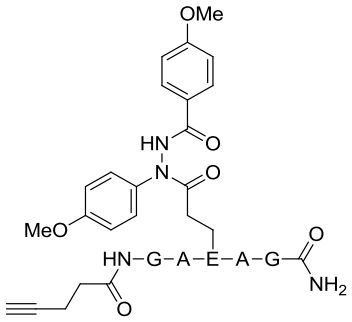


A solution of tetrazole **175o** (3.7 mg, 0.013 mmol) and Peptide **235** (7.0 mg, 0.013 mmol) in TFE (2 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 4 h. The solvent was removed *in vacuo* and the crude material

purified by MDAP Method B (Si C_{18}, 30-85% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 15 min). The solvent was removed by lyophilisation to afford peptide **236** (4.6 mg, 45%) as a fluffy white solid.

LC-MS Method F (3 min): rt = 1.38 min, $[M+H]^+$ = 791.4. HRMS (ESI^+): m/z calcd for $C_{38}H_{47}N_8O_{11}$: 791.3359; found 791.3382.

Hydrazide labelled peptide (237)

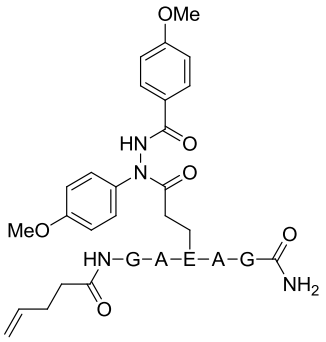


A solution of tetrazole **175o** (5.8 mg, 0.021 mmol) and peptide **234** (10.0 mg, 0.020 mmol) in TFE (1 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 1 h. The solvent was removed *in vacuo* and the crude material purified by MDAP method D (Si C_{18}, 20-40% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min). The solvent was removed by lyophilisation to afford peptide **237** (9.4 mg, 64%) as a fluffy white solid.

LC-MS Method F (2 min): rt = 0.79 min, $[M+H]^+$ = 737.6. IR (cm^{-1}): 3282, 3064, 1696, 1641, 1625, 1513, 1511, 1445, 1422, 1315, 1250, 1176, 1029. ^1H NMR (600 MHz, DMSO-d_6): δ_H = 8.19 - 8.12 (m, 1 H), 8.12 - 7.97 (m, 3 H), 7.93 - 7.80 (m, 3 H), 7.44 - 7.38 (m, 1 H), 7.35 (app. d, *J* = 7.3 Hz, 2 H), 7.16 (br. s., 1 H), 7.08 – 7.03 (m, 2 H), 7.02 – 6.98 (m, 1 H), 6.92 (app. d, *J* = 8.8 Hz, 2 H), 4.30 - 4.09 (m, 3 H), 3.85 - 3.80 (s, 3 H), 3.79 – 3.76 (m, 1 H), 3.74 (s, 3 H), 3.73 – 3.68 (m, 1 H), 3.68 - 3.63 (dd, *J* = 17.2, 6.2 Hz, 1 H), 3.61 - 3.55 (m, 1 H), 2.76 – 2.74 (s, 1 H), 2.36 – 2.34 (m, 2 H), 2.36 – 2.34 (m, 2 H), 2.40 – 2.20 (m, 2 H), 2.11 - 1.70 (m, 2 H), 1.23 - 1.09 (m, 6 H) ppm. HRMS (ESI^+): m/z calcd for $C_{35}H_{45}N_8O_{10}$: 737.3253; found 737.3268.

Note: The alkyne peak was not observed in the weak IR spectrum. 2D $^{1}H^{13}C$ HSQC and HMBC spectra were obtained to identify the presence of the key alkyne peaks but were not analysed further.

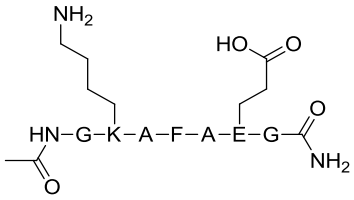
Hydrazide labelled peptide (238)



A solution of tetrazole **175o** (2.5 mg, 8.8 µmol) and peptide **233** (4.3 mg, 8.8 µmol) in TFE (2 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 1 h. Additional tetrazole **175o** (2.5 mg, 8.8 µmol) was added and the mixture stirred under irradiation for a further 1 h. The solvent was removed *in vacuo* and the crude material purified by MDAP method D (Si C_{18}, 20-40% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min). The solvent was removed by lyophilisation to afford peptide **238** (5.5 mg, 84%) as a fluffy white solid.

LC-MS Method F (2 min): rt = 0.83 min, $[M+H]^+$ = 739.6. ^{1}H NMR (400 MHz, DMSO-d_6): δ_H = 8.05 (m, 4 H), 7.87 (m, 3 H), 7.45 - 7.28 (m, 2 H), 7.19 - 7.12 (br. s., 1 H), 7.04 (m, 4 H), 6.95 - 6.87 (m, 1 H), 5.87 - 5.73 (m, 1 H), 5.01 (d, *J* = 17.9 Hz, 1 H), 4.94 (d, *J* = 10.0 Hz, 1 H), 4.32 - 4.05 (m, 4 H), 3.83 (s, 3 H), 3.74 (s, 3 H), 3.69 (m, 2 H), 3.64 (d, *J* = 5.9 Hz, 1 H), 3.60 (d, *J* = 5.6 Hz, 1 H), 2.22 (m, 6 H), 2.14 - 1.68 (m, 2 H), 1.32 - 1.08 (m, 6 H) ppm. HRMS (ESI^+): m/z calcd for $C_{35}H_{47}N_8O_{10}$: 739.3410; found 739.3414.

Lysine containing Peptide (248)



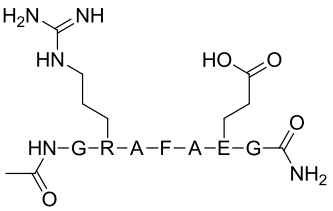
100-200 mesh MBHA resin **227** (0.59 mmol g^{-1}, 1.80 g, 1.1 mmol) was loaded using the appropriate resin loading procedure (see Table 19) with a pre-activated solution of Fmoc-Gly-OH (1.58 g, 5.3 mmol), HATU (2.02 g, 5.3 mmol) and DIPEA (1.85 mL, 10.6 mmol) in DMF (5 mL). The resin loading was determined to be 0.37 mmol g^{-1} by UV analysis of the fulvene-piperidine adduct (see Fmoc Loading Test).

A portion of the resin (0.37 mmol g^{-1}, 600 mg, 0.22 mmol) underwent automated peptide synthesis using the appropriate procedures (see Table 20, 21, and 23) with Fmoc-Glu(^{t}Bu)-OH (473 mg, 1.1 mmol), Fmoc-Ala-OH (346 mg, 1.1 mmol), Fmoc-Phe-OH (430 mg, 1.1 mmol), Fmoc-Lys(Boc)-OH (520 mg, 1.1 mmol), Fmoc-Gly-OH (330 mg, 1.1 mmol), and HATU (422 mg, 1.1 mmol). A small-scale cleavage confirmed the successful synthesis of the peptide.

After a full cleavage the precipitated peptide was dissolved in the minimal amount of DMSO and purified by MDAP method D (Si C_{18}, 0-20% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 30 min). The solvent was removed by lyophilisation to afford peptide **248** (132 mg, 80%) as a fluffy white solid.

LC-MS Method F (3 min): rt = 0.68 min, $[M+H]^+$ = 720.4. HRMS (ESI^+): m/z calcd for $C_{32}H_{50}N_9O_{10}$: 720.3681; found 720.3688.

Arginine containing Peptide (249)



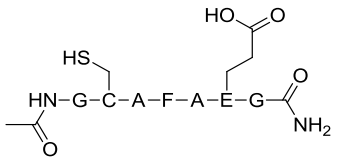
100-200 mesh MBHA resin **227** (0.59 mmol g^{-1}, 1.80 g, 1.1 mmol) was loaded using the appropriate resin loading procedure (see Table 19) with a pre-activated solution of Fmoc-Gly-OH (1.58 g, 5.3 mmol), HATU (2.02 g, 5.3 mmol) and DIPEA (1.85 mL, 10.6 mmol) in DMF (5 mL). The resin loading was determined to be 0.37 mmol g^{-1} by UV analysis of the fulvene-piperidine adduct (see Fmoc Loading Test).

A portion of the resin (0.37 mmol g^{-1}, 600 mg, 0.22 mmol) underwent automated peptide synthesis using the appropriate procedures (see Table 20, 21, 23) with Fmoc-Glu(^{t}Bu)-OH (473 mg, 1.1 mmol), Fmoc-Ala-OH (346 mg, 1.1 mmol), Fmoc-Phe-OH (430 mg, 1.1 mmol), Fmoc-Arg(Pbf)-OH (720 mg, 1.1 mmol), Fmoc-Gly-OH (330 mg, 1.1 mmol), and HATU (422 mg, 1.1 mmol). A small-scale cleavage confirmed the successful synthesis of the peptide.

After a full cleavage the precipitated peptide was dissolved in the minimal amount of DMSO and purified by MDAP method D (Si C_{18}, 0-20% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 30 min). The solvent was removed by lyophilisation to afford peptide **249** as a fluffy white solid (126 mg, 76%).

LC-MS Method F (3 min): rt = 0.70 min, $[M+H]^+$ = 748.5. HRMS (ESI^+): m/z calcd for $C_{32}H_{50}N_{11}O_{10}$: 748.3742; found 748.3744.

Cysteine containing Peptide (250)



100-200 mesh MBHA resin **227** (0.59 mmol g^{-1}, 1.80 g, 1.1 mmol) was loaded using the appropriate resin loading procedure (see Table 19) with a pre-activated solution of Fmoc-Gly-OH (1.58 g, 5.3 mmol), HATU (2.02 g, 5.3 mmol) and DIPEA (1.85 mL, 10.6 mmol) in DMF (5 mL). The resin loading was determined to be 0.37 mmol g^{-1} by UV analysis of the fulvene-piperidine adduct (see Fmoc Loading Test).

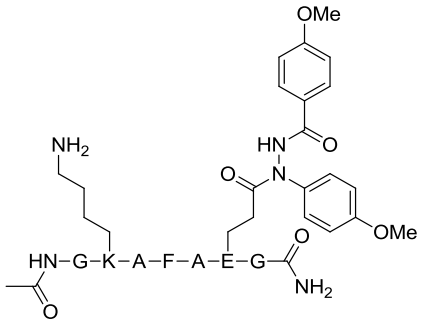
A portion of the resin (0.37 mmol g^{-1}, 600 mg, 0.22 mmol) underwent automated peptide synthesis using the appropriate procedures (see Table 20, 21 and 23) with Fmoc-Glu(^{t}Bu)-OH (473 mg, 1.1 mmol), Fmoc-Ala-OH (346 mg, 1.1 mmol), Fmoc-Phe-OH (430 mg, 1.1 mmol), Fmoc-Cys(Trt)-OH (650 mg, 1.1 mmol), Fmoc-Gly-OH (330 mg, 1.1 mmol), and HATU (422 mg, 1.1 mmol). A small-scale cleavage confirmed the successful synthesis of the peptide.

After a full cleavage the precipitated peptide was dissolved in the minimal amount of DMF and purified by MDAP Method D (Si C_{18}, 0-20% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min). Two diastereoisomers were obtained, due to partial racemisation during the reaction and each was treated separately. The solvent was removed by lyophilisation to give diastereoisomer 1 (66.4 mg, 43%) and diastereoisomer 2 (33.9 mg, 22%) of peptide **250** as fluffy white solids.

Diastereoisomer 1: LC-MS Method F (2 min): rt = 0.45 min, $[M+H]^+$ = 695.5. HRMS (ESI^+): m/z calcd for $C_{29}H_{43}N_8O_{10}S$: 695.2817; found 695.2826.

Diastereoisomer 2: LC-MS Method F (2 min): rt = 0.53 min, $[M+H]^+$ = 695.5. HRMS (ESI^+): m/z calcd for $C_{29}H_{43}N_8O_{10}S$: 695.2817; found 695.2814.

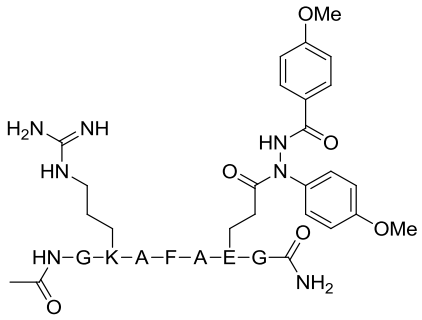
Hydrazide labelled Peptide (255)



A solution of tetrazole **175o** (4.0 mg, 0.014 mmol) and peptide **248** (10.0 mg, 0.014 mmol) in TFE (1 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 1 h. The solvent was removed *in vacuo* and the crude material purified by MDAP method D (Si C_{18}, 20-40% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min). The solvent was removed by lyophilisation to afford peptide **255** (5.0 mg, 38%) as a fluffy white solid.

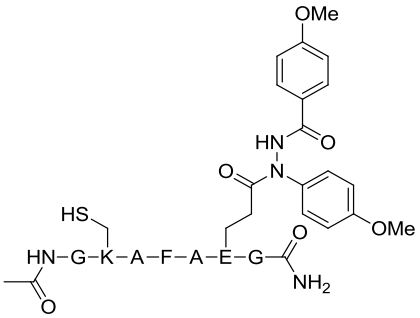
LC-MS Method F (2 min): rt = 0.82 min, $[M+H]^+$ = 974.9. HRMS (ESI^+): m/z calcd for $C_{47}H_{64}N_{11}O_{12}$: 974.4730; found 974.4730.

Hydrazide labelled Peptide (256)



A solution of tetrazole **175o** (4.8 mg, 0.016 mmol) and peptide **249** (10.0 mg, 0.013 mmol) in TFE (3 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 2 h. The solvent was removed *in vacuo* and the crude material purified by MDAP method D (Si C_{18}, 20-40% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 15 min). The solvent was removed by lyophilisation to afford a mixture of products which was not purified further.

Hydrazide labelled Peptide (257)



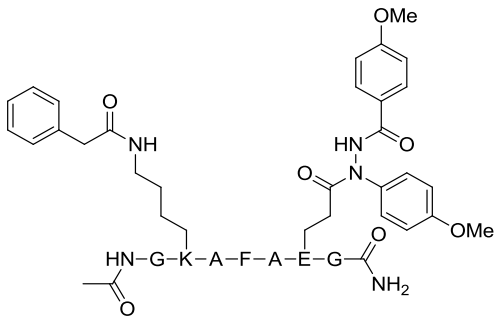
A solution of tetrazole **175o** (4.4 mg, 0.015 mmol) and peptide **250** (10.2 mg, 0.014 mmol) in TFE (4 mL) was stirred under irradiation with a UVB lamp (270-330 nm) for 1 h. The solvent was removed *in vacuo* and the crude material purified by MDAP method E (Si C_{18}, 20-40% MeCN-H_2O, HCO_2H modifier, 30 min). The solvent was removed by lyophilisation to afford a white solid (1.3 mg) containing a mixture of two products.

LC-MS Method F (2 min): Peak 1 - rt = 0.81 min, $[M+H]^+$ = 949.5.

LC-MS Method F (2 min): Peak 2 - rt = 0.97 min, $[M+H]^+$ not observed.

Note: HRMS data could not be obtained in this experiment due to further conversion from product peak 1 to product peak 2 (see peptide dimer **262**).

Hydrazide and amide coupled peptide (261)

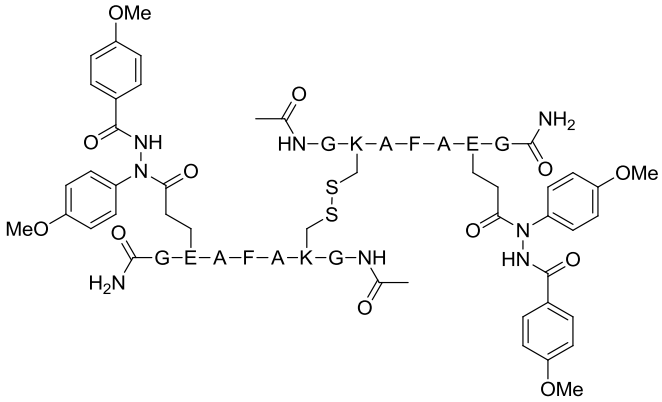


To a solution of peptide **255** (4.0 mg, 4.1 µmol) in DMF (0.5 mL) was added a pre-activated solution of 2-phenylacetic acid **223** (0.62 mg, 4.5 µmol), DIPEA (2.2 µl, 0.012 mmol) and HATU (1.7 mg, 4.5 µmol) in DMF (0.1 mL). The reaction

mixture was stirred at ambient temperature for 2 h. The reaction mixture was diluted with DMF (1 mL) and purified by MDAP Method D (Si C_{18}, 30-50% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min). The solvent was removed by lyophilisation to afford peptide **261** (1.1 mg, 25%) as a fluffy white solid.

LC-MS Method F (2 min): rt = 0.93 min, $[M+H]^+$ = 1092.7. HRMS (ESI^+): m/z calcd for $C_{55}H_{70}N_{11}O_{13}$: 1092.5149; found 1092.5171.

Hydrazide labelled Peptide dimer (262)

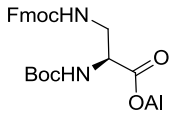


An LC-MS sample of peptide **257** and the unidentified byproduct in EtOAc (1 mL) was prepared. A solution of 33% $H_2O_{2(aq)}$ (0.3 µL, 0.98 mmol) and sodium iodide (3.9 mg, 0.26 mmol) in EtOAc (10 mL) was prepared and 0.1 mL of this solution added to the LC-MS sample. The sample was shaken at room temperature for 30 min and analysis by LC-MS showed peak to peak conversion of monomer **257** to dimer **262**.

LC-MS Method F (2 min): rt = 0.97, (2M+2)/2 = 949.5

3.10.4. Chapter 2: Synthesis of CtB

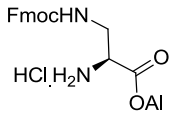
Boc-Dap(Fmoc)-OAl (368)



To a solution of Boc-Dap(Fmoc)-OH **367** (2.5 g, 5.9 mmol) in 1,4-dioxane (100 mL) was added a solution of Cs_2CO_3 (1.0 g, 3.1 mmol) in water (30 mL). The reaction mixture was stirred at room temperature for 30 min, the solvent removed *in vacuo* and the residue dried azeotropically with toluene (5 x 20 mL). The resulting salt was dissolved in DMF (50 mL) and allyl bromide (0.61 mL, 7.0 mmol) added dropwise. The reaction mixture was stirred at room temperature for 4.5 h. The solvent was removed *in vacuo* and the residue partitioned between EtOAc (200 mL) and water (150 mL). The layers were separated and the aqueous layer extracted with EtOAc (2 x 50 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (100 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material was purified by automated flash column chromatography (Si, 80 g, 0-25% EtOAc-cyclohexane, 16 CV) to afford **368** (1.97 g, 72%) as a white solid.

LC-MS Method C (2 min): rt = 1.31 min, $[M+H]^+$ = 467.3. mp 121-123 °C. IR (cm^{-1}): 3346, 2981, 1740, 1690, 1543, 1448, 1368, 1262, 1162. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 7.79 (d, *J* = 7.6 Hz, 2 H), 7.60 (d, *J* = 7.6 Hz, 2 H), 7.42 (t, *J* = 7.1 Hz, 2 H), 7.34 (t, *J* = 7.6 Hz, 2 H), 5.93 (ddd, *J* = 16.7, 10.6, 6.1 Hz, 1 H), 5.46 (br. s., 1 H), 5.36 (dd, *J* = 17.3, 1.4 Hz, 1 H), 5.27 (dd, *J* = 10.4, 1.0 Hz, 1 H), 5.21 (br. s., 1 H), 4.68 (dd, *J* = 5.6, 1.0 Hz, 2 H), 4.41 (d, *J* = 6.8 Hz, 2 H), 4.51-4.41 (m, 1 H) 4.23 (t, *J* = 6.8 Hz, 1 H), 3.73 - 3.55 (m, 2 H), 1.48 (s, 9 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ_C = 170.3, 156.7, 155.5, 143.8 (2 C), 141.3 (2 C), 131.4, 127.7 (2 C), 127.1 (2 C), 125.1 (2 C), 120.0 (2 C), 119.1, 80.4, 67.1, 66.4, 54.2, 47.2, 43.1, 28.3 (3 C) ppm. HRMS (ESI^+): m/z calcd for $C_{26}H_{31}N_2O_6$: 467.2177; found 467.2185.

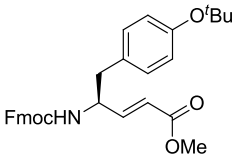
HCl.H-Dap(Fmoc)-OAl (365)[192]



To Boc-Dap(Fmoc)-OAl **368** (1.97 g, 4.2 mmol) was added a solution of TFA:H_2O (19:1 v/v, 20 mL) at 0 °C. The reaction mixture was stirred under a nitrogen atmosphere at room temperature for 3 h. The solvent was removed *in vacuo* and the crude material dried azeotropically with 4M HCl/dioxane (20 mL) twice to afford **365** (1.61 g, 95%) as a white solid.

LC-MS Method C (2 min): rt = 1.06 min, $[M+H]^+$ = 367.2. mp 210-212 °C. IR (cm^{-1}): 3337, 2902, 2884, 2857, 1747, 1697, 1540, 1448, 1264. ^{1}H NMR (400 MHz, DMSO-d_6) δ_H = 8.73 (br. s., 3 H), 7.90 (d, *J* = 7.6 Hz, 2 H), 7.70 (d, *J* = 7.6 Hz, 2 H), 7.68 (t, *J* = 5.8 Hz, 1 H), 7.43 (t, *J* = 7.3 Hz, 2 H), 7.34 (t, *J* = 7.6 Hz, 2 H), 5.89 (ddd, *J* = 16.2, 10.6, 5.6 Hz, 1 H), 5.36 (dd, *J* = 17.2, 1.5 Hz, 1 H), 5.21 (dd, *J* = 10.4, 1.0 Hz, 1 H), 4.68 (dd, *J* = 13.4, 6.6 Hz, 1 H), 4.59 (dd, *J* = 13.4, 5.1 Hz, 1 H), 4.39 - 4.17 (m, 3 H), 4.13 - 4.07 (m, 1 H), 3.78 - 3.62 (m, 2 H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6) δ_C = 168.1, 156.8, 144.2 (2 C), 141.2 (2 C), 132.2, 128.1 (2 C), 127.5 (2 C), 125.6 (2 C), 120.6 (2 C), 119.0, 66.6, 66.4, 52.7, 47.0, 40.9 ppm. HRMS (ESI^+): m/z calcd for $C_{21}H_{23}N_2O_4$: 367.1658; found 367.1651. $[\alpha_D^{24}]$ - 3.6 (*c* 1.00, DMSO-d_6). The data obtained were consistent with literature values.

Fmoc-vTyr(^{t}Bu)-OMe (376)



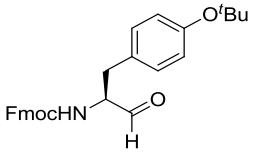
To a solution of **375** (130 mg, 0.29 mmol) in DCM (3 mL) was added DMP (186 mg, 0.44 mmol) in one portion. The reaction mixture was stirred under a nitrogen atmosphere at room temperature for 5 h. The reaction mixture was diluted with DCM (10 mL) and stirred with sat. $Na_2CO_{3(aq)}$ (10 mL) for 20 min. The

layers were separated and the aqueous layer extracted with DCM (2 x 10 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (10 mL), separated, and dried through a hydrophobic frit. Note: A white precipitate was filtered and discarded. The solvent was removed *in vacuo*.

To a solution of the crude material in DCM (5 mL) was added methyl 2-(triphenylphosphoranylidene)acetate (146 mg, 0.44 mmol) in one portion. The reaction mixture was stirred under a nitrogen atmosphere at room temperature for 2 h. The reaction mixture was partitioned between water (20 mL) and DCM (20 mL). The layers were separated and the aqueous layer extracted with DCM (2 x 40 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (20 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated reverse phase chromatography (Si C_{18}, 30 g, 55-85% MeCN:H_2O, $(NH_4)_2CO_3$ modifier, 20 min) to afford **376** (41 mg, 28%) as a white gum.

LC-MS Method C (2 min): rt = 1.42 min, $[M+H]^+$ = 500.4. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 7.78 (app. d, *J* = 7.6 Hz, 2 H), 7.75 – 7.68 (m, 1 H), 7.56 (dd, *J* = 6.9, 4.2 Hz, 2 H), 7.51 – 7.47 (m, 1 H), 7.42 (t, *J* = 7.3 Hz, 2 H), 7.33 (dt, *J* = 7.6, 1.0 Hz, 2 H), 7.11 - 7.04 (m, 2 H), 6.94 (d, *J* = 8.3 Hz, 1 H), 5.86 (d, *J* = 14.6 Hz, 1 H), 4.76 (br. s, 1 H), 4.65 (br. s, 1 H), 4.52 - 4.31 (m, 2 H), 4.20 (t, *J* = 6.3 Hz, 1 H), 3.75 (s, 3 H), 3.01 - 2.76 (m, 2 H), 1.35 (s, 9 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ_C = 166.4, 155.5, 154.4, 147.3, 143.8 (2 C), 141.3 (2 C), 129.8 (2 C), 127.7 (2 C), 127.1 (2 C), 125.0 (2 C), 124.2 (2 C), 121.0 (2 C), 120.0 (2 C), 78.4, 66.8, 52.9, 51.7, 47.2, 40.0, 28.8 (3 C) ppm.

Fmoc-Tyr(^{t}Bu)-H (377)[228]

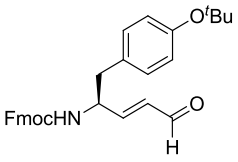


To a solution of **375** (1.87 g, 4.2 mmol) in DCM (50 mL) was added DMP (2.3 g, 5.5 mmol) portionwise. The reaction mixture was stirred under a nitrogen atmosphere at room temperature for 5 h. The reaction mixture was quenched with

sat. $Na_2CO_{3(aq)}$ (30 mL) with stirring for 30 min. The reaction mixture was partitioned between DCM (100 mL) and water (100 mL). The layers were separated and the aqueous layer extracted with DCM (2 x 50 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (50 mL), separated, and dried through a hydrophobic frit. Note: A white precipitate was filtered and discarded. The solvent was removed *in vacuo* and the residue partitioned between EtOAc (100 mL) and 10% $Na_2CO_{3(aq)}$ (100 mL). The layers were separated and the organic phase was washed with water (50 mL) and sat. $NaCl_{(aq)}$ (2 x 50 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* to afford **377** (1.70 g, 91%) as a white foam. This was used directly in the next experiment.

LC-MS Method C (2 min): rt = 1.20-1.40 min, $[M+H]^+$ = 444.4. IR (cm^{-1}): 3325, 2977, 1713, 1506, 1450, 1366, 1238, 1161, 1057. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 9.66 (s, 1 H), 7.79 (app. d, *J* = 7.6 Hz, 2 H), 7.59 (d, *J* = 7.3 Hz, 2 H), 7.43 (t, *J* = 7.6 Hz, 2 H), 7.34 (t, *J* = 7.6 Hz, 2 H), 7.05 (app. d, *J* = 7.6 Hz, 2 H), 6.94 (d, *J* = 8.1 Hz, 2 H), 5.31 (d, *J* = 6.3 Hz, 1 H), 4.57 - 4.38 (m, 3 H), 4.24 (t, *J* = 6.7 Hz, 1 H), 3.13 (d, *J* = 4.0 Hz, 2 H), 1.35 (s, 9 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ_C = 198.8, 155.9, 154.6, 143.7 (2 C), 143.7, 141.4 (2 C), 129.8 (2 C), 127.8 (2 C), 127.1 (2 C), 125.0 (2 C), 124.3 (2 C), 120.0 (2 C), 78.5, 67.0, 61.1, 47.2, 34.8, 28.9 (3 C) ppm. The data obtained were consistent with literature values.

Fmoc-vTyr(^{t}Bu)-H (378)

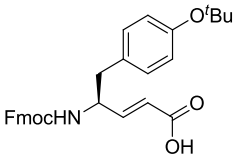


To a solution of **377** (1.70 g, 3.8 mmol) in DCM (60 mL) was added 2-(triphenylphosphoranylidene) acetaldehyde (1.40 g, 4.6 mmol) in one portion. The reaction mixture was stirred at room temperature for 76 h. The reaction mixture was washed with water (50 mL) and the layers separated. The aqueous layer was extracted with DCM (2 x 50 mL), the layers separated, and the combined

organic phase dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 120 g, 0-50% TBME-cyclohexane, 17 CV) to afford **378** (864 mg, 50%) as a yellow foam.

LC-MS Method C (2 min): rt = 1.37 min, $[M+H]^+$ = 470.3. IR (cm^{-1}): 3334, 2982, 1687, 1505, 1450, 1365, 1236, 1160. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 9.65 (br. s, 1H), 7.79 (app. d, *J* = 7.3 Hz, 2 H), 7.56 (dd, *J* = 7.3 Hz, 2.8 Hz, 2 H), 7.43 (t, *J* = 7.5 Hz, 2 H), 7.34 (t, *J* = 7.3 Hz, 2 H), 7.11 - 7.00 (m, 2 H), 6.95 (app. d, *J* = 8.1 Hz, 2 H), 6.77 (d, *J* = 13.9 Hz, 1 H), 6.19 - 6.02 (m, 1 H), 4.87 - 4.68 (m, 2 H), 4.53 - 4.39 (m, 2 H), 4.21 (t, *J* = 6.3 Hz, 1 H), 2.99 - 2.81 (m, 2 H), 1.35 (s, 9 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ_C = 192.9, 155.6, 154.6, 143.7 (2 C), 141.4 (2 C), 131.6, 129.7 (2 C), 127.8, 127.8, 127.1 (2 C), 124.9 (2 C), 124.8 (2 C), 124.3 (2 C), 120.0 (2 C), 78.5, 66.7, 53.2, 47.3, 39.8, 28.9 (3 C) ppm. HRMS (ESI^+): m/z calcd for $C_{30}H_{32}NO_4$: 470.2326; found 470.2323.

Fmoc-vTyr(*t*Bu)-OH (364)[192]



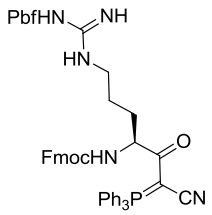
To a solution of **378** (0.845 g, 1.8 mmol) in *tert*-butanol (85 mL) and 2-methyl-2-butene (30 mL) was added a solution of $NaClO_2$ (2.17 g, 18.0 mmol) and $NaH_2PO_4.H_2O$ (1.86 g, 13.5 mmol) in H_2O (55 mL) dropwise. The reaction mixture was stirred at room temperature for 3 h. The solvent was removed *in vacuo* to approximately 1/3 of the volume and additional water (50 mL) added. The pH was adjusted to pH 3 with 25% $HCl_{(aq)}$ solution and the aqueous layer extracted with EtOAc (100 mL). The layers were separated and the aqueous layer extracted with EtOAc (2 x 50 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (50 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated flash column

chromatography (Si, 80 g, 0-50% TBME-cyclohexane, AcOH modifier, 16 CV) to afford (539 mg, 62%) as a white solid.

LC-MS: Method F (2 min): rt = 1.02 min, $[M+H]^+$ = 486.2. mp 97-100 °C. IR (cm^{-1}): 3338, 2977, 1692, 1508, 1447, 1366, 1238, 1158. ^{1}H NMR (400 MHz, DMSO-d_6) δ_H = 7.89 (app. d, *J* = 7.6 Hz, 2 H), 7.70 - 7.59 (m, 2 H), 7.42 (t, *J* = 7.8 Hz, 2 H), 7.37 - 7.26 (m, 2 H), 7.18 - 7.10 (m, 3 H), 6.88 - 6.77 (m, 3 H), 5.78 (d, *J* = 15.7 Hz, 1 H), 4.37 (m, 1 H), 4.27 - 4.10 (m, 3 H), 2.83 (dd, *J* = 13.6, 5.3 Hz, 1 H), 2.70 (dd, *J* = 13.9, 8.8 Hz, 1 H), 1.20 (s, 9 H) ppm. ^{13}C NMR (151 MHz, $CDCl_3$) δ_C = 170.2, 155.3, 154.2, 149.3, 143.6, 143.5, 141.1 (2 C), 130.4, 129.5 (2 C), 127.5 (2 C), 126.9 (2 C), 124.7 (2 C), 124.0 (2 C), 120.4, 119.8 (2 C), 78.3, 66.6, 52.7, 47.0, 39.7, 28.6 (3 C) ppm. HRMS (ESI^+): m/z calcd for $C_{30}H_{32}NO_5$: 486.2258; found 486.2280. $[\alpha_D^{24}]$ -29.5 (*c* 1.00, $CDCl_3$). The data obtained were consistent with literature values.

Note: Exchangeable CO_2H proton was not observed in the ^{1}H NMR spectrum.

Fmoc-Arg(Pbf)-COC(PPh$_3$)CN (385)



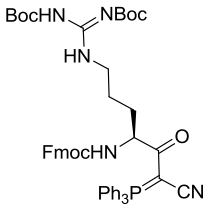
A solution of **383** (5.05 g, 7.8 mmol) in DCM (100 mL) was cooled to 0 °C then DMAP (95 mg, 0.78 mmol) and EDC.HCl (1.94 g, 10.1 mmol) added sequentially. The reaction mixture was stirred for 15 min then 2-(triphenylphosphoranylidene) acetonitrile (3.05 g, 10.1 mmol) added in one portion. The reaction mixture was allowed to warm to room temperature and stirred for 16 h. The reaction mixture was washed with water (100 mL), the layers separated, and the aqueous layer extracted with DCM (2 x 50 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (50 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated flash column

chromatography (Si, 120 g, 25-75% EtOAc-cyclohexane, 16 CV) to afford **385** (5.94 g, 82%) as a white solid.

LC-MS Method C (2 min): rt = 1.44 min, $[M+H]^+$ = 932.4. IR (cm^{-1}): 3334, 2973, 2968, 2176, 1718, 1550, 1438, 1248, 1162, 1107. ^{1}H NMR (400 MHz, DMSO-d_6) δ_H = 7.89 (d, *J* = 7.6 Hz, 2 H), 7.78 - 7.48 (m, 18 H), 7.45 - 7.34 (m, 3 H), 7.30 (t, *J* = 6.8 Hz, 2 H), 6.70 (br. s, 1 H), 6.42 (br. s, 1 H), 4.61 - 4.52 (m, 1 H), 4.34 - 4.18 (m, 3 H), 3.16 - 3.03 (m, 2 H), 2.93 (s, 2 H), 2.44 (s, 3 H), 1.98 (s, 3 H), 1.90 - 1.58 (m, 2 H), 1.57 - 1.44 (m, 2 H), 1.39 (s, 6 H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6) δ_C = 194.4, 157.9, 156.5, 156.5, 144.4, 144.2, 141.2, 137.8, 133.8, 133.7, 132.7, 132.5, 132.5, 132.0, 131.9, 129.8, 129.7, 129.3, 129.2, 128.1, 127.5, 125.9, 125.8, 124.8, 123.4, 122.5, 120.5, 116.7, 86.7, 66.2, 60.2, 55.4, 47.2, 42.9, 28.8 , 21.2, 19.4, 18.1, 14.6, 12.7 ppm. HRMS (ESI^+): m/z calcd for $C_{54}H_{55}N_5O_6PS$: 932.3610; found 932.3611.

Note: For the ^{1}H NMR spectrum a Me peak is coincident with the DMSO solvent signal.

Fmoc-Arg(Boc)$_2$-COC(PPh$_3$)CN (386)



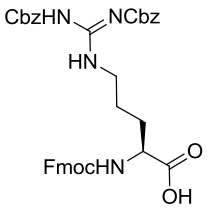
A solution of **384** (200 mg, 0.36 mmol), HOAt (59 mg, 0.44 mmol) and 2-triphenylphosphoranylidene)acetonitrile (303 mg, 1.0 mmol) in DCM (5 mL) was cooled to 0 °C and stirred for 15 min. EDC.HCl (84 mg, 0.44 mmol) was added and the reaction mixture stirred under a nitrogen atmosphere for 18 h. The reaction mixture was partitioned between DCM (50 mL) and water (50 mL), the organic layer separated, and the aqueous layer extracted with DCM (2 x 50 mL). The combined organic phase was dried through a hydrophobic frit, the solvent removed *in vacuo,* and the crude material purified by automated flash column chromatography (Si, 20 g, 0-50% EtOAc-cyclohexane, 30 min) and MDAP Method D

(Si C_{18}, 80-99% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 15 min). The solvent was removed by lyophilisation to afford **386** (14 mg, 5%) as a fluffy white solid.

LC-MS Method F (2 min): rt = 1.63 min, $[M+H]^+$ = 880.7. IR (cm^{-1}): 3329, 2979, 2178, 1719, 1638, 1615, 1586, 1485, 1439, 1415, 1366, 1331, 1230, 1155, 1133, 1108, 1050, 1026. ^1H NMR (400 MHz, DMSO-d_6) δ_H = 11.50 (s, 1 H), 8.33 (t, *J* = 5.1 Hz, 1 H), 7.89 (d, *J* = 7.3 Hz, 2 H), 7.78 - 7.68 (m, 5 H), 7.65 - 7.53 (m, 12 H), 7.45 - 7.36 (m, 3 H), 7.30 (t, *J* = 7.3 Hz, 2 H), 4.66 - 4.55 (m, 1 H), 4.31 - 4.18 (m, 3 H), 3.57 (s,1 H), 3.35 (dd, *J* = 6.1, 12.5 Hz, 1 H), 1.90 - 1.53 (m, 4 H), 1.47 (s, 9 H), 1.39 (s, 9 H) ppm. HRMS (ESI^+): m/z calcd for $C_{51}H_{55}N_5O_7P$: 880.3834; found 880.3848.

Note: An accurate ^{13}C spectrum could not be collected as the molecule was found to be unstable within the timescale of the analysis.

Fmoc-Arg(Cbz)$_2$-OH (357)



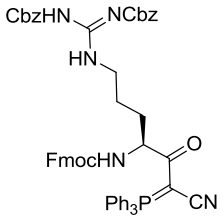
A suspension of **352** (7.5 g, 13.8 mmol) in DCM (25 mL) was cooled to 0 °C and TFA (24 mL, 312 mmol) added dropwise with stirring. The reaction mixture was allowed to warm to room temperature and stirred for 3 h. The solvent was removed *in vacuo* and the residual TFA removed azeotropically with DCM (3 x 10 mL). The residue was dissolved in 1,4-dioxane (100 mL), cooled to 0 °C, and 10% $Na_2CO_{3(aq)}$ (85 mL, 85 mmol) added. The reaction mixture was stirred for 10 min and Fmoc-Cl (3.93 g, 15.2 mmol) added portionwise. The reaction mixture was allowed to warm to room temperature and stirred for 5 h. The reaction mixture was partitioned between 10% citric $acid_{(aq)}$ (100 mL) and EtOAc (250 mL). The layers were separated and the aqueous layer extracted with EtOAc (2 x 200 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (50 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and

the crude material purified by automated flash column chromatography (Si, 220 g, 0-25% (3:1 EtOAc-EtOH)-cyclohexane, AcOH modifier, 17 CV) and then (Si, 120 g, 20-50% EtOAc-cyclohexane, AcOH modifier, 17 CV) to afford **357** (4.13 g, 45%) as a white solid.

LC-MS Method F (2 min): rt = 1.11 min, $[M+H]^+$ = 665.3. mp 170-173 °C. IR (cm^{-1}): 3316, 2960, 1719, 1688, 1608, 1498, 1449, 1377, 1253, 1099. ^1H NMR (400 MHz, DMSO-d_6) δ_H = 12.55 (br. s, 1 H), 9.16 (br. s, 2 H), 7.88 (d, *J* = 7.6 Hz, 2 H), 7.70 (d, *J* = 7.6 Hz, 2 H), 7.62 (d, *J* = 8.1 Hz, 1 H), 7.44 - 7.25 (m, 14 H), 5.24 (s, 2 H), 5.05 (s, 2 H), 4.31 - 4.24 (m, 2 H), 4.20 (q, *J* = 5.8 Hz, 1 H), 3.99 - 3.79 (m, 3 H), 1.75-1.51 (m, 4 H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6) δ_C = 174.1, 163.3, 160.2, 156.6, 155.4, 144.3, 144.2, 141.2 (2 C), 137.6, 135.7, 129.0 (2 C), 128.8 (2 C), 128.7, 128.4 (2 C), 128.3 (2 C), 128.2, 128.1 (2 C), 127.5 (2 C), 125.7 (2 C), 120.5 (2 C), 68.7, 66.6, 66.0, 54.2, 47.2, 28.5, 25.8 ppm. HRMS (ESI^+): m/z calcd for $C_{37}H_{37}N_4O_8$: 665.2611; found 665.2622.

Note: One 1 C not observed in the ^{13}C NMR spectrum as it is coincident with the DMSO-d_6 peak.

Fmoc-Arg(Cbz)$_2$-COC(PPh$_3$)CN (358)[204]



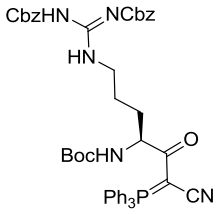
A solution of **357** (4.0 g, 6.0 mmol) in DCM (225 mL) was cooled to 0 °C then DMAP (74 mg, 0.61 mmol) and EDC.HCl (1.5 g, 7.8 mmol) added. The reaction mixture was stirred at 0 °C for 30 min then 2-(triphenylphosphoranylidene)acetonitrile (2.4 g, 7.8 mmol) added in one portion. The reaction mixture was stirred at 0 °C for a further 30 min, allowed to warm to room temperature, and stirred for 16 h. The

reaction mixture was washed with water (50 mL), the organic layer separated, and the aqueous layer extracted with DCM (2 x 50 mL). The combined organic phase was washed with sat. $NaCl_{(aq)}$ (50 mL), separated, and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated flash column chromatography (Si, 220 g, 50-100% TBME-cyclohexane, 16 CV) and then (Si, 80 g, 20-60% EtOAc-cyclohexane, 16 CV) to afford **358** (293 mg, 5%) as a white gum.

LC-MS Method F (2 min): rt = 1.62 min, $[M+H]^+$ = 948.5. IR (cm^{-1}): 3385, 3062, 2950, 2176, 1715, 1604, 1498, 1438, 1378, 1246, 1197, 1098. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 9.54 (br. s, 1 H), 9.32 (br. s, 1 H), 7.77 (d, *J* = 7.6 Hz, 2 H), 7.69 - 7.55 (m, 10 H), 7.51 (dt, *J* = 7.8, 3.3 Hz, 6 H), 7.45 - 7.18 (m, 15 H), 5.74 (d, *J* = 6.8 Hz, 1 H), 5.21 (d, *J* = 4.8 Hz, 2 H), 5.13 (s, 2 H), 5.01 (br. s., 1 H), 4.39 (dd, *J* = 10.1, 7.8 Hz, 1 H), 4.28 (dd, *J* = 10.3, 7.3 Hz, 1 H), 4.24 - 4.09 (m, 2 H), 3.97 (br. s, 1 H), 2.07 (s, 1 H), 1.94 - 1.60 (m, 3 H) ppm. HRMS (ESI^+): m/z calcd for $C_{57}H_{51}N_5O_7P$: 948.3521; found 948.3555. The data obtained were consistent with literature values.

Note: Product degradation was observed on silica and accurate ^{13}C spectrum could not be obtained due to degradation observed during the NMR timescale.

Boc-Arg(Cbz)$_2$-COC(PPh$_3$)CN (353)[204]



General procedure for the optimisation reaction:

A solution of **352**, 2-(triphenylphosphoranylidene)acetonitrile, and HOAt in DCM was cooled to 0 °C under a nitrogen atmosphere and stirred for 15 min. EDC was added, the reaction mixture allowed to warm to room temperature, and then stirred for a specified amount of time. The reaction mixture was washed with water

(100 mL), the organic layer separated, and the aqueous layer extracted with DCM (2 x 50 mL). The combined organic phase was dried with $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was pre-absorbed onto diatomaceous earth and purified by automated flash column chromatography.

Entry 1: A solution of 2-(triphenylphosphoranylidene)acetonitrile (3.70 g, 12.3 mmol), **352** (5.0 g, 9.2 mmol), DMAP (115 mg, 0.94 mmol), and EDC (2.35 g, 12.3 mmol) in DCM (100 mL) was stirred for 16 h. Purified by automated flash column chromatography (Si, 120 g, 20-60% EtOAc-cyclohexane, 16 CV) to afford **353** (3.52 g, 45%) as a white foam.

Entry 2: A solution of 2-(triphenylphosphoranylidene)acetonitrile (111 mg, 0.37 mmol), **352** (100 mg, 0.18 mmol), DMAP (2.3 mg, 0.018 mmol), HOAt (25 mg, 0.18 mmol), and EDC (33 µL, 0.18 mmol) in DCM (2 mL) was stirred for 22 h. Purified by automated reverse phase chromatography (Si C_{18}, 30 g, 65-95% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 20 CV) to afford **353** (77 mg, 51%) as a white foam.

Entry 3: A solution of 2-(triphenylphosphoranylidene)acetonitrile (167 mg, 0.55 mmol), **352** (100 mg, 0.18 mmol), DMAP (2.3 mg, 0.018 mmol), HOAt (33 mg, 0. 24 mmol), and EDC (42 µL, 0.24 mmol) in DCM (2 mL) was stirred for 5 h. Purified by automated reverse phase chromatography (Si C_{18}, 30 g, 65-95% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 20 CV) to afford **353** (112 mg, 74%) as a white foam.

Entry 4: A solution of 2-(triphenylphosphoranylidene)acetonitrile (167 mg, 0.55 mmol), **352** (100 mg, 0.18 mmol), HOAt (33 mg, 0.24 mmol), and EDC (42 µL, 0.24 mmol) in DCM (2 mL) was stirred for 6 h. Purified by automated reverse phase chromatography (Si C_{18}, 30 g, 65-95% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 20 CV) to afford **353** (115 mg, 76%) as a white foam.

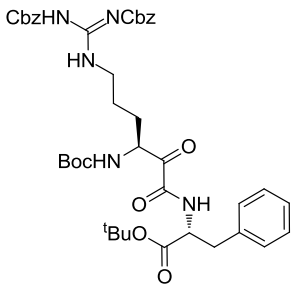
Entry 5: A solution of 2-(triphenylphosphoranylidene)acetonitrile (1.67 g, 5.5 mmol), **352** (1.0 g, 1.8 mmol), HOAt (326 mg, 2.4 mmol), and EDC (0.42 mL, 2.4 mmol) in

DCM (20 mL) was stirred for 6.5 h. Purified by automated reverse phase chromatography (Si C_{18}, 150 g, 65-95% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 20 CV) to afford **353** (1.35 g, 89%) as a white foam.

Entry 6: A solution of 2-(triphenylphosphoranylidene)acetonitrile (8.3 g, 27.6 mmol), **352** (5.0 g, 9.2 mmol), HOAt (1.6 g, 12.0 mmol), and EDC (2.1 mL, 12.0 mmol) in DCM (100 mL) was stirred for 6.5 h. Purified by automated reverse phase chromatography [3 x (Si C_{18}, 150 g, 65-95% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 20 CV)] to afford **353** (7.21 g, 95%) as a white foam.

LC-MS Method F (2 min): rt = 1.52 min, $[M+H]^+$ = 826.5. IR (cm^{-1}): 3385, 2974, 2176, 1709, 1587, 1487, 1438, 1366, 1247, 1162, 1098. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 9.52 (br. s, 1 H), 9.29 (br. s, 1 H), 7.68 - 7.22 (m, 25 H), 5.31 (m, 1 H), 5.25 (s, 2 H), 5.14 (s, 2 H), 4.93 - 4.75 (m, 1 H), 4.19 - 3.89 (m, 2 H), 2.07 - 1.67 (m, 4 H), 1.44 (s, 9 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ_C = 194.6, 164.0, 160.7, 156.0, 155.6, 137.0, 134.8, 133.6, 133.5, 133.2, 129.3, 129.1, 128.8, 128.4, 128.3, 127.8, 127.6, 123.3, 122.3, 79.0, 68.9, 67.1, 56.0, 47.0, 44.7, 30.4, 28.4, 24.9 ppm. HRMS (ESI^+): m/z calcd for $C_{47}H_{49}N_5O_7P$: 826.3370; found 826.3370. The data obtained were consistent with literature values.

Boc-Arg(Cbz)$_2$-αKA-D-Phe-O^{t}Bu (403)



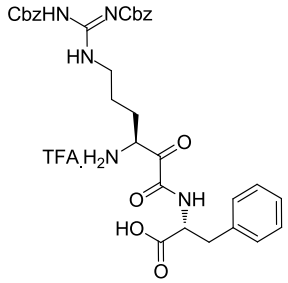
A solution of **353** (717 mg, 0.87 mmol) in DCM (30 mL) was cooled to -78 °C under a nitrogen atmosphere. A stream of ozone (generated from oxygen using an Oz502 ozone generator, 50 L h^{-1}, 89 W) was passed through the reaction mixture for *ca.*

10 min until a green-yellow solution was observed. The ozone was cleared under a stream of oxygen then nitrogen. Note: A sat. $Na_2S_2O_{3(aq)}$ scrubber was present to quench excess ozone and a 10% $NaClO_{(aq)}$ scrubber was present to quench any hydrogen cyanide released. A solution of H-D-Phe-O^{t}Bu **398** (202 mg, 0.91 mmol) in DCM (10 mL) was added dropwise over 10 min and the reaction mixture allowed to warm to room temperature with stirring over 16 h. The solvent was removed *in vacuo* at room temperature and the crude material was purified by automated flash column chromatography (Si, 40 g, 20-50% TBME-cyclohexane, 16 CV) to afford **403** (290 mg, 43%) as a sticky white solid.

Note: Careful treatment of this material is required as it will degrade if exposed to heat.

LC-MS Method F (2 min): rt = 1.61 min, $[M+H]^+$ = 774.5. IR (cm^{-1}): 3388, 2979, 2970, 1715, 1644, 1609, 1498, 1455, 1368, 1247, 1155, 1097. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 9.32 (br. s, 3 H), 7.51 - 7.10 (m, 15 H), 5.35 - 5.30 (m, 1 H), 5.26 (s, 2 H), 5.17 (s, 2 H), 5.05 - 4.95 (m, 1 H), 4.72 - 4.63 (m, 1 H), 4.08 - 3.90 (m, 2 H), 3.12 (t, *J* = 4.5 Hz, 2 H), 1.77 - 1.67 (m, 2 H), 1.59 - 1.49 (m, 2 H), 1.42 (s, 9 H), 1.40 (s, 9 H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ_C = 195.9, 169.4, 163.8, 160.5, 158.7, 155.8, 137.0, 135.5, 134.7, 129.4, 128.8, 128.5, 128.4, 127.8, 127.7, 127.2, 119.6, 82.8, 77.2, 68.9, 67.0, 53.7, 44.2, 37.9, 28.3, 27.9, 26.9, 25.1 ppm. HRMS (ESI^+): m/z calcd for $C_{41}H_{52}N_5O_{10}$: 774.3714; found 774.3693.

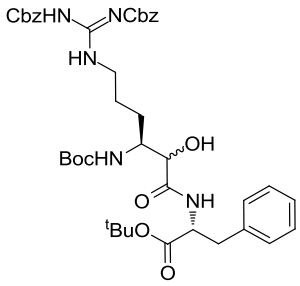
TFA.H-Arg(Cbz)$_2$-αKA-D-Phe-OH (405)



To **403** (48 mg, 0.062 mmol) was added a solution of DCM:TFA:H_2O (1.5:1.0:0.1 v/v, 2.6 mL) and the reaction mixture stirred at room temperature for 16 h. The solvent was removed *in vacuo* and the residual TFA removed azeotropically with DCM (3 x 10 mL). The crude material was taken forward without further purification.

LC-MS Method F (2 min): rt = 0.94 min, $[M+H]^+$ = 618.2.

Boc-Arg(Cbz)$_2$-αHA-D-Phe-O^{t}Bu (407)



Reduction from isolated **403**:

A solution of **403** (290 mg, 0.37 mmol) in DCM:MeOH (4:1 v/v, 5 mL) was cooled to 0 °C and $NaBH_4$ (15 mg, 0.39 mmol) added. The reaction mixture was stirred at 0 °C for 30 min, allowed to warm to room temperature, and stirred for 2 h. The reaction mixture was adjusted to pH 4/5 with 2M $HCl_{(aq)}$ then partitioned between water (25 mL) and DCM (25 mL). The organic layer was separated and the aqueous layer extracted with DCM (2 x 20 mL). The combined organic phase was dried over

$MgSO_4$, filtered, and the solvent removed *in vacuo* at 25 °C to afford **407** (267 mg, 92%) as a colourless gum.

Telescoped ozonolysis and reduction from **353**:

A solution of **353** (2.5 g, 3.03 mmol) in DCM (100 mL) was cooled to -78 °C under nitrogen. A stream of ozone (generated from oxygen using an Oz502 ozone generator, 50 L h^{-1}, 89 W) was passed through the reaction mixture for *ca.* 10 min until a green-yellow solution was observed. The ozone was cleared under a stream of oxygen then nitrogen. Note: A sat. $Na_2S_2O_{3(aq)}$ scrubber was present to quench excess ozone and a 10% $NaClO_{(aq)}$ scrubber was present to quench any hydrogen cyanide released. A solution of H-D-Phe-O^{t}Bu **398** (703 mg, 3.2 mmol) in DCM (25 mL) was added dropwise over 20 min. The reaction mixture was allowed to warm to room temperature with stirring overnight, during which time the solvent evaporated.

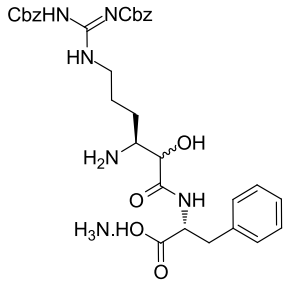
The residue was redissolved in DCM (50 mL), cooled to 0 °C, and $NaBH_4$ (115 mg, 3.0 mmol) added. The reaction mixture was allowed to warm to room temperature with stirring over 2 h. The solvent was removed *in vacuo* and the residue redissolved in EtOAc (50 mL). The pH was adjusted to pH 4/5 with 2M $HCl_{(aq)}$ and the organic layer washed with water (50 mL). The organic layer was separated and the aqueous layer extracted with EtOAc (2 x 50 mL). The combined organic phase was dried through a hydrophobic frit, pre-absorbed onto diatomaceous earth and purified by automated flash column chromatography (Si, 120 g, 30-60% TBME-cyclohexane, 20 CV) to afford **407** (1.16 g, 49%) as a white solid.

LC-MS Method F (2 min): rt = 1.55/1.57 min; $[M+H]^+$ = 776.5. IR (cm^{-1}): 3389, 2977, 2970, 1715, 1607, 1500, 1498, 1454, 1367, 1247, 1197, 1154, 1098. ^{1}H NMR (600 MHz, $CDCl_3$) δ_H = 9.54 - 9.10 (m, 2 H), 7.44 – 7.17 (m, 15 H), 5.61 - 5.50 (m, 1 H), 5.29 – 5.21 (m, 2 H), 5.18 – 5.07 (m, 2 H), 4.78 – 4.65 (m, 1 H), 4.29 - 4.04 (m, 2 H), 4.02 - 3.81 (m, 2 H), 3.90 - 3.79 (m, 2 H), 3.14 – 2.94 (m, 2 H), 1.77 - 1.56

(m, 4 H), 1.36 - 1.41 (m, 18 H) ppm. ^{13}C NMR (151 Hz, $CDCl_3$) δ_C = 171.8, 171.1, 170.3, 163.7, 163.6, 160.7, 160.5, 158.1, 157.0, 155.9, 155.9, 136.8, 136.8, 136.2, 136.2, 134.7, 134.7, 129.5, 129.4, 129.4, 129.4, 128.9, 128.8, 128.8, 128.5, 128.4, 128.4, 128.4, 128.4, 127.9, 127.9, 127.9, 127.8, 127.0, 126.9, 82.2, 82.0, 80.2, 79.9, 75.9, 74.0, 69.0, 68.9, 67.0, 67.0, 54.5, 53.5, 53.3, 53.3, 53.1, 53.0, 44.3, 44.1, 38.7, 38.1, 28.3, 28.3, 28.2, 27.9, 27.9, 26.8, 25.2, 25.1, 24.8 ppm. HRMS (ESI$^+$): m/z calcd for $C_{41}H_{54}N_5O_{10}$: 776.3871; found 776.3872.

Note: ^{13}C NMR is a mixture of diastereoisomers and all observed peaks are reported.

H-Arg(Cbz)$_2$-αHA-D-Phe-OH (408)



Deprotection and isolation with TFA then HCl:

To a solution of **407** (247 mg, 0.32 mmol) in DCM (1.5 mL) was added TFA (50 µL, 0.64 mmol) and the reaction mixture stirred at room temperature for 3 h. Additional TFA (50 µL, 0.64 mmol) was added and the reaction mixture stirred for a further 1 h. The solvent was removed *in vacuo* and the residual TFA removed azeotropically with toluene (3 x 10 mL).

Note: Partial CBz deprotection was observed instead of the desired Boc/*tert*-butyl ester deprotection.

The residue was dissolved in 1,4-dioxane (2 mL), 4M HCl in dioxane (0.40 mL, 1.6 mmol) added, and the reaction mixture stirred at room temperature

for 24 h. Additional 4M HCl in dioxane (0.40 mL, 1.6 mmol) was added and the reaction mixture stirred for an additional 28 h. The solvent was removed *in vacuo* and the residue redissolved in neat 4M HCl in dioxane (3 mL, 12 mmol) and stirred at room temperature for 3 h. The solvent was removed *in vacuo* and the crude material pre-absorbed onto diatomaceous earth then purified by automated reverse phase chromatography (Si C_{18}, 30 g, 25-75% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 20 CV) to afford **408** (20 mg, 10%) as an off white gum.

LC-MS: Method F (2 min): rt = 0.92 min, $[M+H]^+$ = 620.3. IR (cm^{-1}) : 3385, 3032, 2952, 1607, 1511, 1455, 1381, 1246, 1215, 1098. ^1H NMR (400 MHz, $CDCl_3$) δ_H = 9.41 (br. s., 1 H), 9.27 (br. s., 1 H), 7.47 - 7.30 (m, 7 H), 7.27 - 7.00 (m, 8 H), 5.29 - 5.14 (m, 2 H), 5.06 (s., 2 H), 4.57 - 4.39 (m, 1 H), 4.39 - 4.24 (m, 1 H), 3.82 - 3.45 (m, 4 H), 3.30 (m, 2 H), 3.08 - 2.77 (m, 2 H), 1.85 - 1.36 (m, 4 H) ppm. HRMS (ESI^+): m/z calcd for $C_{32}H_{37}N_5O_8$: 620.2709; found 620.2715.

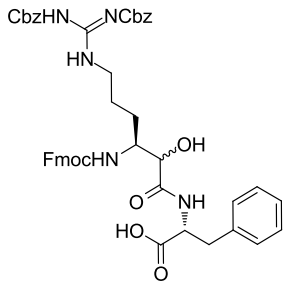
Note: The exchangeable CO_2H proton was not observed in the ^1H NMR spectrum. A ^{13}C spectrum was not obtained for this product but was obtained for the final molecule in the sequence (see **409**).

Deprotection with HCl:

A solution of **407** (1.16 g, 1.5 mmol) in 4M HCl in dioxane (11.5 mL, 46 mmol) was stirred at room temperature for 5 h. The solvent was removed *in vacuo* and the residual HCl removed azeotropically with DCM (3 x 10 mL). The crude material was taken forward without further purification.

LC-MS: Method F (2 min): rt = 0.90 min, $[M+H]^+$ = 620.4.

Fmoc-Arg(Cbz)$_2$-αHA-D-Phe-OH (409)

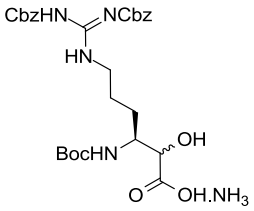


To a solution of **408** (20 mg, 0.031 mmol) in 1,4-dioxane (2 mL) was added a solution of DIPEA (20 mg, 0.16 mmol) in 1,4-dioxane (0.5 mL). The mixture was stirred at 30 °C for 5 min and the solvent removed *in vacuo.* The process was repeated once more. The residue was dissolved in 1,4-dioxane (2 mL) and a solution of Fmoc-OSu (10.6 mg, 0.031 mmol) in 1,4-dioxane (0.5 mL) added. The reaction mixture was stirred at room temperature for 5.5 h. The solvent was removed *in vacuo* and the residue partitioned between EtOAc (10 mL) and water (10 mL) then adjusted to pH 4/5 with 2M $HCl_{(aq)}$. The organic layer was separated and the aqueous layer extracted with EtOAc (2 x 10 mL). The combined organic phase was dried through a hydrophobic frit, the solvent removed *in vacuo* and the crude material purified by MDAP Method A (Si C_{18}, 80-99% MeCN-H_2O, HCO_2H modifier, 15 min) to afford **409** (5 mg, 19%) as a white gum.

LC-MS Method E (2 min): rt = 1.48/1.49 min, $[M+H]^+$ = 842.5/842.5. IR (cm^{-1}): 3390, 1717, 1608, 1512, 1451, 1253, 1104. ^1H NMR (600 MHz, $CDCl_3$) δ_H = 9.72 – 9.03 (m, 2 H), 7.95 – 6.91 (m, 24 H), 5.99 – 5.59 (m, 1 H), 5.27 – 4.94 (m, 4 H), 4.78 – 4.61 (m, 1 H), 4.44 – 4.21 (m, 2 H), 4.11 – 4.04 (m, 1 H), 4.21 – 3.98 (m, 1 H), 3.95 – 3.66 (m, 2 H), 3.95 – 3.66 (m, 1 H), 3.24 – 2.94 (m, 2 H), 1.66 – 1.20 (m, 4 H) ppm. ^{13}C NMR (151 MHz, $CDCl_3$) δ_C = 179.4, 176.8, 173.1, 160.7, 157.9, 155.8, 143.7, 141.3, 136.6, 134.6, 131.6, 129.2, 129.1, 128.9, 128.8, 128.7, 128.6, 128.4, 128.4, 128.3, 128.1, 128.0, 127.9, 127.7, 127.1, 127.1, 126.9, 125.0, 124.9, 119.9, 68.9, 67.0, 66.7, 54.8, 53.3, 47.2, 47.1, 44.1, 37.2, 26.7, 25.1, 24.9, 24.1 ppm. HRMS (ESI^+): m/z calcd for $C_{47}H_{47}N_5O_{10}$: 842.3396; found 842.3428.

Note: For the ^{1}H NMR spectrum the exchangeable CO_2H proton was not observed. The ^{13}C NMR is a mixture of diastereoisomers and all observed peaks are reported

Boc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ (410)



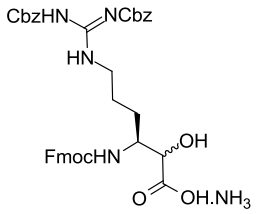
A solution of **353** (1.50 g, 1.8 mmol) in DCM (50 mL) was cooled to -78 °C under a nitrogen atmosphere. A stream of ozone (generated from oxygen using an Oz502 ozone generator, 50 L h^{-1}, 89 W) was passed through the reaction mixture for *ca.* 10 min until a green-yellow solution was observed. The ozone was cleared under a stream of oxygen then nitrogen. Note: A sat. $Na_2S_2O_{3(aq)}$ scrubber was present to quench excess ozone and a 10% $NaClO_{(aq)}$ scrubber was present to quench any hydrogen cyanide released.

A solution of THF:H_2O (9:1 v/v, 50 mL) was added dropwise over 20 min then the reaction mixture was allowed to warm to 0 °C and stirred for a 20 min. Note: The solvent initially freezes but melts upon warming during which the yellow solution becomes colourless. $NaBH_4$ (80 mg, 2.1 mmol) was added in one portion and the reaction mixture stirred at 0 °C for 45 min. The solvent was removed *in vacuo* at 35 °C and the residue partitioned between EtOAc (50 mL) and 10% citric $acid_{(aq)}$ (50 mL). The organic layer was separated and the aqueous layer extracted with EtOAc (2 x 50 mL). The combined organic phase was dried through a hydrophobic frit, the solvent removed *in vacuo* and the crude material purified by automated reverse phase chromatography (Si C_{18}, 400 g, 20-70% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 45 min) to afford **410** (810 mg, 76%) as a white solid.

LC-MS Method F (2 min): rt = 0.94 min, $[M+H]^+$ = 573.3. IR (cm^{-1}): 3394, 2976, 1721, 1683, 1607, 1497, 1455, 1379, 1247, 1194, 1171, 1098. ^1H NMR (400 MHz, DMSO-d_6): δ_H = 9.13 (br. s., 2 H), 7.48 - 7.27 (m, 10 H), 6.45 (d, *J* = 9.0 Hz, 1 H), 6.03 (d, *J* = 8.6 Hz, 1 H), 5.24 (s, 2 H), 5.06 (s, 2 H), 3.90 - 3.41 (m, 4 H), 1.72 - 1.41 (m, 4 H), 1.34 (m, 9 H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6): δ_C = 174.7, 163.4, 160.1, 155.7, 155.4, 137.6, 135.8, 129.0 (2 C), 128.8 (2 C), 128.7, 128.3 (2 C), 128.2 (2 C), 128.1, 77.8, 73.8, 72.0, 68.6, 66.6, 45.8, 28.7 (3 C), 28.6, 25.9 ppm. HRMS (ESI^+): m/z calcd for $C_{28}H_{37}N_4O_9$: 573.2555; found 573.2559.

Note: For the ^1H NMR spectrum the exchangeable CO_2H proton was not observed.

Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ (412)



Experiment 1: A solution of **410** (25 mg, 0.044 mmol) in 4M HCl in dioxane (0.5 mL, 2.0 mmol) was stirred at room temperature for 1 h. The pH was adjusted to pH 10 with sat. $Na_2CO_{3(aq)}$ solution. The solvent was removed *in vacuo* and the residue redissolved in water:MeCN (1:1 v/v, 4 mL). The reaction mixture was cooled to 0 °C and Fmoc-OSu (15.5 mg, 0.045 mmol) added. The reaction mixture was stirred for 2 h.

Experiment 2: A solution of **410** (25 mg, 0.044 mmol) in 4M HCl in dioxane (0.5 mL, 2.0 mmol) was stirred at room temperature for 1 h. The pH was adjusted to pH 10 with sat. $Na_2CO_{3(aq)}$ solution. The solvent was removed *in vacuo* and the residue redissolved in water:acetone (1:1 v/v, 4 mL). The reaction mixture was cooled to 0 °C and Fmoc-OSu (15.5 mg, 0.045 mmol) added. The reaction mixture was stirred for 2 h.

Experiment 3: A solution of **410** (10 mg, 0.018 mmol) in 4M HCl in dioxane (0.25 mL, 1.0 mmol) was stirred at room temperature for 1 h. The pH was adjusted to pH 10 with sat. $NaHCO_{3(aq)}$ solution. The solvent was removed *in vacuo* and the residue redissolved in water:THF (1:1 v/v, 2 mL). The reaction mixture was cooled to 0 °C and Fmoc-OSu (11.9 mg, 0.035 mmol) added. The reaction mixture was stirred for 4 h.

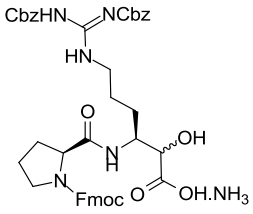
Experiment 4: A solution of **410** (10 mg, 0.018 mmol) in 4M HCl in dioxane (0.25 mL, 1.0 mmol) was stirred at room temperature for 1 h. The product was precipitated with cold Et_2O (1 mL), centrifuged at 2500 rpm, and the solvent decanted. This process was repeated twice more. The residue was dissolved in water:THF (1:1 v/v, 1 mL) then $NaHCO_3$ (4.5 mg, 0.053 mmol) and Fmoc-OSu (11.9 mg, 0.035 mmol) added. The reaction mixture was stirred at room temperature for 4 h.

Purification: The samples from experiments 1-4 were combined, the solvent removed *in vacuo* and the crude material purified by automated reverse phase chromatography (Si C_{18}, 30 g, 30-85% MeCN-H_2O, HCO_2H modifier, 30 min) to afford **412** (14 mg, 16%) as a white solid.

LC-MS Method F (2 min): rt = 1.09 min, $[M+H]^+$ = 695.5. IR (cm^{-1}): 3381, 1729, 1681, 1616, 1536, 1498, 1450, 1372, 1258, 1192, 1101. ^{1}H NMR (400 MHz, DMSO-d_6) δ_H = 9.15 (br. s., 2H), 7.87 (d, *J* = 7.6 Hz, 2 H), 7.75 - 7.62 (m, 2 H), 7.45 - 7.24 (m, 15 H), 5.25 - 5.19 (m, 2 H), 5.04 (s, 2 H), 4.28 - 4.10 (m, 3 H), 3.99 - 3.70 (m, 4 H), 1.72 - 1.40 (m, 4 H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6): δ_C = 174.4, 163.4, 160.1, 156.2, 155.5, 144.4 (2 C), 144.2 (2 C), 141.2 (2 C), 137.6, 135.8, 129.0 (2 C), 128.8 (2 C), 128.3 (2 C), 128.3 (2 C), 128.1, 128.0, 127.5 (2 C), 125.8 (2 C), 120.5 (2 C), 68.6, 66.8, 66.6, 65.9, 53.9, 47.2, 45.1, 25.7, 25.7 ppm. HRMS (ESI^+): m/z calcd for $C_{38}H_{39}N_4O_9$: 695.2712; found 695.2703.

Note: The exchangeable CO_2H was not observed in the ^{1}H NMR spectrum. An additional proton was not observed and may be coincident with the DMSO-d_6 peak.

Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ (418)



Unoptimised conditions:

A solution of **410** (25 mg, 0.044 mmol) in 4M HCl in dioxane (0.25 mL, 1.0 mmol) was stirred at room temperature for 1.5 h then quenched with DIPEA (0.2 mL, 1.2 mmol) with stirring over 20 min. A yellow precipitate was observed. A pre-activated solution of HATU (20 mg, 0.053 mmol), DIPEA (30 µL, 0.023 mmol) and Fmoc-Pro-OH **366** (18 mg, 0.053 mmol) in DMF (0.75 mL) was added dropwise and the mixture stirred at room temperature for 5 h. The solvent was removed *in vacuo* and the crude material purified by MDAP Method D (Si C_{18}, 30-85% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 15 min). The solvent was removed by lyophilisation to afford **418** (7.7 mg, 22%) as a yellow solid.

Optimised conditions:

Solid **410** (300 mg, 0.52 mmol) was partitioned between EtOAc (100 mL) and 5% citric acid$_{(aq)}$ (100 mL). The organic layer was separated and the aqueous layer extracted with EtOAc (2 x 50 mL). The combined organic phase was dried through a hydrophobic frit and the solvent removed *in vacuo*. The residue was dissolved in 1,4-dioxane (20 mL) and lyophilised twice.

4M HCl in dioxane (3.0 mL, 12.0 mmol) was added and the reaction mixture stirred at room temperature for 1 h then quenched with DIPEA (2.3 mL, 13.2 mmol) with stirring over 20 min. A yellow precipitate was observed. A pre-activated solution of HATU (229 mg, 0.60 mmol), HOAt (82 mg, 0.60 mmol), DIPEA (0.23 mL, 1.3 mmol) and Fmoc-Pro-OH **366** (203 mg, 0.60 mmol) in DMF (5 mL) was added dropwise. The reaction mixture was stirred at room temperature for 2 h. Additional pre-activated

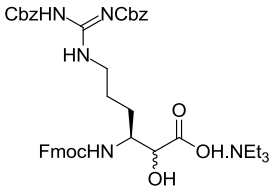
HATU (29.9 mg, 0.078 mmol), HOAt (11 mg, 0.078 mmol), DIPEA (30 µL, 0.17 mmol) and Fmoc-Pro-OH **366** (27 mg, 0.080 mmol) in DMF (0.5 mL) was added dropwise and the reaction mixture stirred for a further 1.5 h.

The reaction mixture was partitioned between EtOAc (200 mL) and 10% $NaHCO_{3(aq)}$ (200 mL). The organic layer was separated and the aqueous layer extracted with EtOAc (2 x 50 mL). The combined organic phase was washed with 10% citric $acid_{(aq)}$ (200 mL) and sat. $LiCl_{(aq)}$ (200 mL). The organic phase was separated and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated reverse phase chromatography (Si C_{18}, 120 g, 30-60% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 40 min). The solvent was removed *in vacuo* and the residual water removed azeotropically with toluene (3 x 25 mL) and DCM (3 x 25 mL) to afford **418** (280 mg, 66%) as a white solid.

LC-MS Method F (2 min): rt = 1.06/1.06 min, $[M+H]^+$ = 792.5/792.6. IR (cm^{-1}): 2986, 1712, 1606, 1500, 1450, 1414, 1353, 1246, 1050. ^1H NMR (400 MHz, DMSO-d_6): δ_H = 9.12 (m, 2H), 7.97 - 7.77 (m, 2 H), 7.72 - 7.48 (m, 2 H), 7.46 - 7.21 (m, 13 H), 5.23 (m., 1 H), 5.08 - 4.91 (m, 2 H), 4.45 - 2.92 (m, 15 H), 2.23 - 0.92 (m, 8 H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6): δ_C = 174.7, 173.8, 163.4, 160.2, 155.5, 154.5, 154.4, 144.4, 143.0, 141.2, 141.1, 139.9, 137.9, 137.6, 135.8, 129.4, 129.0, 129.0, 128.8, 128.8, 128.7, 128.6, 128.4, 128.3, 128.3, 128.2, 128.2, 128.1, 128.1, 127.8, 127.6, 125.6, 121.8, 120.6, 120.5, 110.2, 73.6, 68.6, 67.5, 67.0, 66.6, 60.3, 60.2, 60.1, 55.4, 51.8, 51.2, 47.7, 47.1, 47.0, 45.3, 32.0, 30.8, 30.5, 25.9, 25.7, 24.3, 23.4 ppm. HRMS (ESI^+): m/z calcd for $C_{43}H_{46}N_5O_{10}$: 792.3239; found 792.3258.

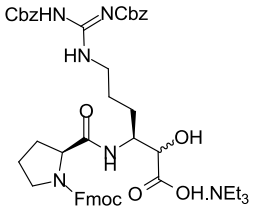
Note: The ^{13}C NMR spectrum shows a mixture of diastereoisomers and all observed peaks are reported.

Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ (424)



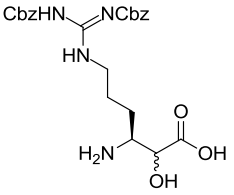
A solution of **412** (5-10 mg) in the minimal amount of MeOH was loaded onto an aminopropyl column (1 g) under gravity. The column was washed with MeOH (3 CV) then eluted with 1-10% Et_3N:MeOH (6 CV). The solvent was removed *in vacuo* to afford **424** as a white solid which was used immediately in subsequent reactions.

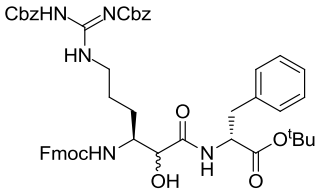
Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ (419)



A solution of **418** (243 mg) in the minimal amount of MeOH was loaded onto an aminopropyl column (1-5 g) under gravity. The column was washed with MeOH (3 CV) then eluted with 5-10% Et_3N:MeOH (6 CV). The solvent was removed *in vacuo* to afford **419** as a white solid which was used immediately in subsequent reactions.

Investigative reactions with monomer Fmoc-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ (424)





Monitoring for the formation of H-Arg(Cbz)$_2$-CHOHCO$_2$H **427** or Fmoc-Arg(Cbz)$_2$-αHA-D-Phe-O^{t}Bu **425**:

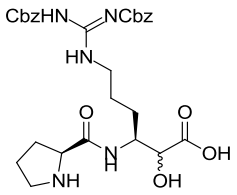
Deprotection:

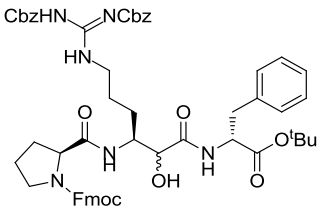
To solid **424** (0.5 mg, 0.72 µmol) was added a solution of 20% piperidine in DMF (0.5 mL) and the reaction mixture stirred at room temperature for 10 min. The reaction mixture was analysed by LC-MS which indicated that **427** was only the minor component of the reaction mixture.

Amide coupling:

To a solution of **424** (10.0 mg, 14 µmol) in DMF (0.25 mL) was added a solution of HATU (8.4 mg, 22 µmol) and DIPEA (5.0 µL, 28 µmol) in DMF (0.1 mL). The mixture was pre-activated with stirring for 10 minutes before addition of H-D-Phe-O^{t}Bu **398** (4.9 mg, 21 µmol) in DMF (0.1 mL). The reaction mixture was stirred at room temperature for 30 min then analysed by LC-MS which indicated that **427** was not present.

Investigative reactions with dipeptide Fmoc-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ (419)





Monitoring for the formation of H-Pro-Arg(Cbz)$_2$-CHOHCO$_2$H **428** or Fmoc-Pro-Arg(Cbz)$_2$-αHA-D-Phe-O^{t}Bu **426**:

Deprotection:

To **419** (5.0 mg, 6.3 µmol) was added a solution of 20% piperidine in DMF (0.15 mL) and the reaction mixture stirred at room temperature for 10 min. The reaction mixture was analysed by LC-MS which indicated that **428** was present. Additional mono-Cbz deprotection was observed.

Amide coupling:

To a solution of **419** (10.0 mg, 13 µmol) in DMF (0.25 mL) was added a solution of HATU (6.2 mg, 16 µmol) and DIPEA (4.0 µL, 22 µL) in DMF (0.1 mL). The mixture was pre-activated with stirring for 15 minutes before addition of H-D-Phe-O^{t}Bu **398** (4.9 mg, 21 µmol) in DMF (0.1 mL). The reaction mixture was stirred at room temperature for 1.5 h then analysed by LC-MS which indicated that **426** was present as two diastereoisomers.

General synthesis of cyclotheonamide and analogues using SPPS

Peptides were synthesised using solid phase synthesis on the stated resin. All solid phase chemistry was carried out manually using a shaker and fishtank or a Merrifield resin bubbler. All reagents and solvents were of peptide synthesis grade.

Fmoc Loading Test

To two 10 mL volumetric flasks was added an accurately weighed portion (*ca.* 10 mg) of loaded resin. Fmoc deprotection was carried out by accurately filling the flasks with 20% piperidine in DMF (10 mL) and sonicating the suspension for 15 minutes. The UV absorption of the samples was measured at 302 nm (Varian Cary 50 probe UV-visible spectrometer, Varian Cary software), relative to a blank solution of 20% piperidine in DMF. The loading of the resin was calculated using the equation below. The two values were averaged to give the resin loading.

$$\text{Loading} = \frac{\text{A} \times 10}{\text{m} \times 7.8}$$

where A = absorption at 302 nm

m = mass of resin/ mg

Chloranil Test[213]

A few beads of resin were separated and washed with MeOH. A solution of acetaldehyde (200 µL) and sat. chloranil in toluene (50 µL) was added and the reaction mixture shaken at room temperature for 5 min. Positive result (indicates an amine is present): Blue/green colour. Negative result (indicates all amine has reacted): Yellow/orange/amber colour.

Small-scale Cleavage

To a sample of resin (*ca.* 2.5 mg) was added a solution of TFA:TIPS:DCM (1:1:98 v/v, 250 µL). The resin was shaken vigorously for 45 min then the reaction mixture quenched with DIPEA (13 µL). The sample was analysed by LC-MS.

Full-scale Cleavage

A solution of TFA:TIPS:DCM (1:1:98 v/v, 5 mL) was added to the resin, the suspension shaken vigorously for 20 min, filtered, and washed with DCM (5 mL). This was repeated three times. The filtrates were combined, quenched with pyridine (180 µL) and the solvent removed *in vacuo*. The crude peptide was dissolved in the minimum amount of DMF for purification by automated reverse phase chromatography. The solvent was reduced to *ca.* 1/3 volume *in vacuo* and extracted with EtOAc (3 x 25 mL). The organic phase was combined and the solvent removed *in vacuo* to afford the desired peptide as a white solid.

Table 24: Manual loading of 2-chlorotrityl chloride resin.

Cycle	Reagents	Solvent	Volume/ mL	Time/ min	Mixing	Iterations
Swell	-	DCM	5	10	N_2 Bubbling	-
Coupling	DIPEA (4 equiv.) HCl.H-Dap (Fmoc)-OAl **365** (2 equiv.)	DCM	5	180	N_2 Bubbling	1
Wash	-	DCM	5	0.5	N_2 Bubbling	3
Capping	DIPEA (2 equiv.) MeOH:DCM (1:4 v/v)	-	5	60	N_2 Bubbling	1
Wash	-	DMF	5	0.5	N_2 Bubbling	3
Wash	-	2-Propanol	5	0.5	N_2 Bubbling	3
Wash	-	TBME	5	0.5	N_2 Bubbling	1
Wash	-	Et_2O	5	0.5	N_2 Bubbling	1
Vacuum dry	-	-	-	30	-	-

Table 25: Manual addition of each subsequent amino acid.

Cycle	Reagents	Solvent	Volume/ mL	Time/ min	Mixing	Iterations
Swell	-	DCM	5	10	Shaking	-
Wash	-	DMF	5	0.5	Shaking	5
Deprotection	20% piperidine in DMF	-	5	10	Shaking	2
Wash	-	DMF	5	0.5	Shaking	5
Coupling	HATU (1.5-2.0 equiv.) HOAt (1.5-2.0 equiv.) DIPEA (1.5-2.0 equiv.) Amino acid (1.5-2.0 equiv.)	DMF	5	120 to 180	Shaking	1
Wash	-	DMF	5	0.5	Shaking	5
Wash	-	DCM	5	0.5	Shaking	5
Vacuum dry	-	-	-	10	-	-

Table 26: Manual deprotections and cyclisation of linear peptide.

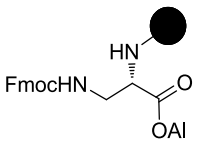
Cycle	Reagents	Solvent	Volume/ mL	Time/ min	Mixing	Iterations
Swell	-	DCM	5	10	Shaking	-
Deprotection	$Pd(PPh_3)_4$ (0.4 equiv.) Phenylsilane (24 equiv.)	DCM	5	10	Shaking	1
Wash	-	DCM	5	0.5	Shaking	5

Repeat as above	-	-	-	-	-	4
Wash	-	IPA	5	0.5	Shaking	5
Wash	-	Et_2O	5	0.5	Shaking	5
Vacuum dry	-	-	-	10	-	-

Swell	-	DCM	5	10	Shaking	-
Wash	-	DMF	5	0.5	Shaking	5
Deprotection	20% piperidine in DMF	-	5	0.5-5	Shaking	2
Wash	-	DMF	5	0.5	Shaking	5
Wash	-	DCM	5	0.5	Shaking	5
Vacuum dry	-	-	-	10	-	-

Swell	-	DCM	5	10	Shaking	-
Wash	-	DMF	5	0.5	Shaking	5
Cyclisation	PyBOP (3 equiv.) HOAt (3 equiv.) DIPEA (6 equiv.)	DMF	5	120	Shaking	-
Wash	-	DMF	5	0.5	Shaking	5
Wash	-	DCM	5	0.5	Shaking	5
Vacuum dry	-	-	-	10	-	-

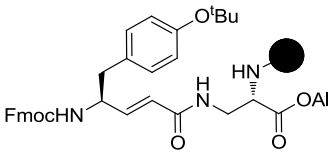
FmocHN-CH_2-CH(NH®)-CO_2Al (297)



2-chlorotrityl chloride resin **296** (1.7 mmol g^{-1}, 0.70 g, 1.2 mmol) was loaded using the appropriate resin loading and capping procedure (see Table 24) with a solution of **365** (960 mg, 2.4 mmol) and DIPEA (0.83 mL, 4.8 mmol) followed by a solution of DCM-MeOH (4:1 v/v, 5 mL) and DIPEA (0.40 mL, 2.4 mmol). The resin loading was

determined to be 0.36 mmol g^{-1} by UV analysis of the fulvene-piperidine adduct (see Fmoc Loading Test).

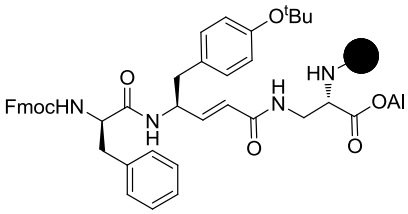
Fmoc-vTyr(^{t}Bu)-HN-CH_2-CH(NH®)-CO_2Al (429)



The loaded resin **297** (0.36 mmol g^{-1}, 562 mg, 0.20 mmol) underwent an Fmoc deprotection followed by a 120 minute manual coupling (See Table 25) with a pre-activated solution of Fmoc-vTyr-OH (146 mg, 0.30 mmol), HATU (114 mg, 0.30 mmol), DIPEA (52 µL, 0.30 mmol), and HOAt (41 mg, 0.30 mmol) in DMF (5 mL). The reaction was shown to be complete with a negative 'chloranil test' followed by a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.31 min, $[M+H]^+$ = 612.2.

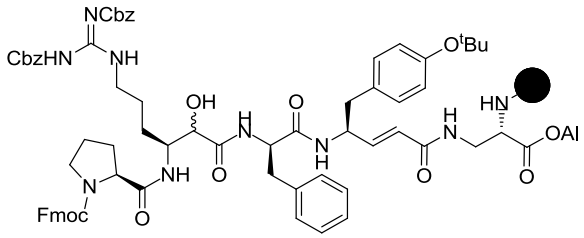
Fmoc-D-Phe-vTyr(^{t}Bu)-HN-CH_2-CH(NH®)-CO_2Al (430)



Dipeptide **429** (0.20 mmol) underwent an Fmoc deprotection followed by a 120 minute manual coupling (See Table 25) with a pre-activated solution of Fmoc-D-Phe-OH (116 mg, 0.30 mmol), HATU (114 mg, 0.30 mmol), DIPEA (52 µL, 0.30 mmol), and HOAt (41 mg, 0.30 mmol) in DMF (5 mL). The reaction was shown to be complete with a negative 'chloranil test' followed by a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.38 min, $[M+H]^+$ = 759.6.

Fmoc-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO$_2$Al (431)



Tripeptide **430** (0.20 mmol) underwent an Fmoc deprotection followed by a 120 minute manual coupling (See Table 25) with a pre-activated solution of **419** (243 mg, 0.30 mmol), which had been freshly prepared by salt exchange with **418** on an aminopropyl column, HATU (114 mg, 0.30 mmol), DIPEA (52 µL, 0.3 mmol), and HOAt (41 mg, 0.30 mmol) in DMF (5 mL). The reaction was shown to be complete with a negative 'chloranil test' followed by a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.54/1.56 min, $[M+H]^+$ = 1312.1/1312.0.

Optimisation of Fmoc Deprotection

Optimisation was conducted using the standard Fmoc deprotection procedure (see Table 26) with 250 µL of all solvents and deprotection times between 0.5-5 min.

Entry 1: Pentapeptide **431** (14 mg, 0.36 mmol/g, 5.0 µmol) was subjected to Fmoc deprotection conditions using a solution of 20% piperidine in DMF (250 µL, 2 x 5 min). Analysis was conducted by small-scale cleavage and LC-MS.

LC-MS Method F (2 min):
Fmoc deprotection: rt = 1.36 min, $[M+H]^+$ = 1088.8 (54% conversion).
Fmoc + Cbz deprotection: rt = 1.10 min, $[M+H]^+$ = 954.7 (13% conversion).

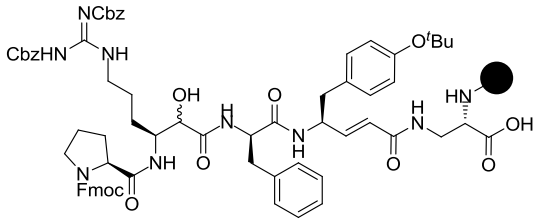
Entry 2: Pentapeptide **431** (14 mg, 0.36 mmol/g, 5.0 µmol) was subjected to Fmoc deprotection conditions using a solution of 20% piperidine in DMF (250 µL, 2 x 45 sec). Analysis was conducted by small-scale cleavage and LC-MS.

LC-MS Method F (2 min):
Fmoc deprotection: rt = 1.18 min, $[M+H]^+$ = 1048.8 (70% conversion).
Fmoc + mono-Cbz deprotection: Not observed.

Entry 3: Pentapeptide **431** (14 mg, 0.36 mmol/g, 5.0 µmol) was subjected to Fmoc deprotection conditions using a solution of 20% piperidine in DMF (250 µL, 2 x 30 sec). Analysis was conducted by small-scale cleavage and LC-MS.

LC-MS Method F (2 min):
Fmoc deprotection: rt = 1.36 min, $[M+H]^+$ = 1088.8 (60% conversion).
Fmoc + mono-Cbz deprotection: Not observed.

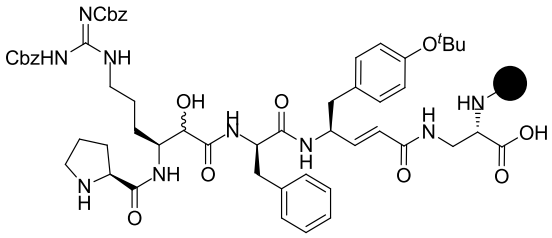
Fmoc-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO$_2$H (432)



Allyl Deprotection:
Pentapeptide **431** (0.20 mmol) was subjected to allyl deprotection conditions (see Table 26) using a solution of $Pd(PPh_3)_4$ (92 mg, 0.080 mmol) and phenylsilane (0.60 mL, 4.9 mmol) in dry DCM (5 mL) five times. The resin was washed with DCM (5 mL, 5 x 30 sec) after each treatment and a fresh solution of deprotection reagents used each time. The reaction was shown to be complete by conducting a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.36/1.39 min, $[M+H]^+$ = 1271.0/1271.0.

H-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO$_2$H (433)

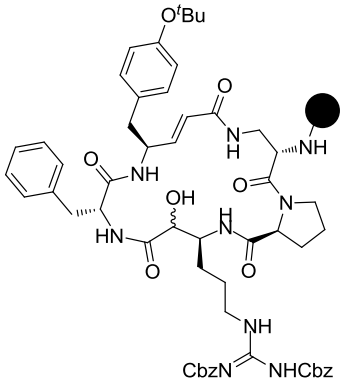


Fmoc Deprotection:

Pentapeptide **432** (0.20 mmol) was subjected to Fmoc deprotection conditions (see Table 26) using a solution of 20% piperidine in DMF (5 mL, 2 x 45 sec). Note: Care must be taken to minimise the reaction time to avoid additional Cbz deprotection. The reaction was shown to be complete by conducting a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.16/1.18 min, $[M+H]^+$ = 1048.8/1048.8.

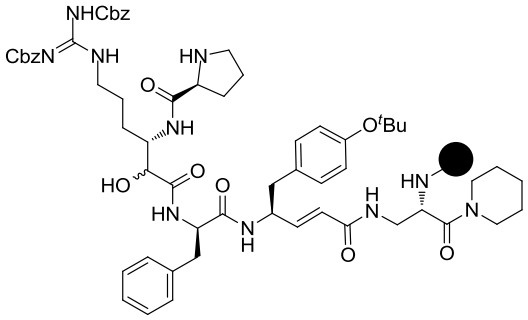
Cyclo [-HN-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO-] (434)



Unprotected pentapeptide **433** (0.20 mmol) was cyclised (see Table 26) using a solution of PyBOP (310 mg, 0.60 mmol), HOAt (82 mg, 0.60 mmol) and DIPEA (0.21 mL, 1.2 mmol) in DMF (5 mL). The reaction was shown to be complete by conducting a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.26/1.31 min, $[M+H]^+$ = 1030.8/1030.8.

H-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH$_2$)CO(piperidine) (437)



Allyl Deprotection:

Pentapeptide **431** (*ca.* 10 mg, *ca.* 5 µmol) was subjected to allyl deprotection conditions (see Table 26) using a solution of $Pd(PPh_3)_4$ (2.3 mg, 2.0 µmol) and phenylsilane (15 µL, 120 µmol) in dry DCM (250 µL) five times. The resin was washed with DCM (0.5 mL, 5 x 30 sec) after each treatment and a fresh solution of deprotection reagents used each time. The reaction was shown to be complete by conducting a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.37/1.39 min, $[M+H]^+$ = 1270.9/1270.9.

Attempted Cyclisation:

Pentapeptide **432** (*ca.* 10 mg, *ca.* 5 µmol) was subjected to cyclisation conditions (see Table 26) with a solution of PyBOP (7.8 mg, 15 µmol), HOAt (2.0 mg, 15 µmol) and DIPEA (5.2 µL, 30 µmol) in DMF (0.5 mL). The reaction was shown to be complete by conducting a small-scale cleavage and analysis by LC-MS. The observed product was determined to be the HOAt activated ester Fmoc-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO$_2$HOAt which was subsequently quenched with piperidine (see Fmoc deprotection).

LC-MS Method F (2 min): rt = 1.30/1.33 min, $[M+H]^+$ = 1367.9/1367.9.

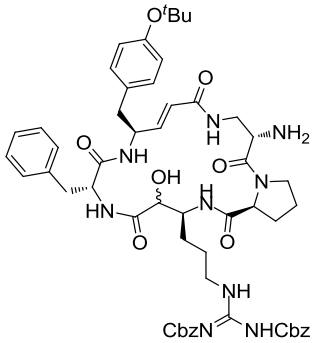
Fmoc Deprotection:

Pentapeptide Fmoc-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO$_2$HOAt (*ca.* 10 mg, *ca.* 5 µmol) was subjected to Fmoc deprotection conditions (see Table 26) using a solution of 20% piperidine in DMF (2 x 30 sec, 0.25 mL). The

reaction was shown to be complete by conducting a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.36/1.36 min, $[M+H]^+$ = 1115.9/1115.9.

Cyclo [-HN-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH$_2$)-CO-] (441/442)



Cyclic pentapeptide **434** (0.10 mmol) was swollen in DCM (5 mL) for 10 min. A solution of TFA:TIPS:DCM (1:1:98 v/v, 5 mL) was added to the resin, the suspension shaken vigorously for 20 min, filtered, and washed with DCM (2 x 2 mL). This was repeated three times. The filtrates were combined, quenched with pyridine (180 µL, 0.10 mmol) and the solution washed with water (25 mL). The organic layer was separated and the aqueous layer extracted with DCM (2 x 25 mL) and EtOAc (3 x 25 mL). The combined organic phase was dried through a hydrophobic frit, the solvent removed *in vacuo,* and the crude material purified by MDAP method D (Si C_{18}, 40-60% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min) to afford two separated diastereoisomers as white solids, with a total yield of 17% over 12 steps. [Diastereoisomer 1: **441** (8.3 mg, 8%, 90% purity by LC-MS); Diastereoisomer 2: **442** (9.2 mg, 9%, 78% purity by LC-MS)].

Diastereoisomer 1 **441**: LC-MS Method F (2 min): rt = 1.26 min, $[M+H]^+$ = 1030.8. IR (cm^{-1}): 3387, 3358, 2925, 1720, 1644, 1507, 1455, 1367, 1258, 1098. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 9.64 - 9.38 (m, 2 H), 7.99 - 7.84 (d, *J* = 10.1, 1 H), 7.48 – 7.30 (m, 14 H), 7.26 – 7.17 (m, 3 H), 6.95 - 6.65 (m, 5 H), 6.01 - 5.83 (dd, *J* = 15.6, 1.5 Hz, 1 H), 5.31 - 5.03 (m, 5 H), 4.94 - 4.80 (m, 1 H), 4.54 - 4.34 (m, 3 H), 4.26 - 4.00 (m, 2 H), 3.94 - 3.79 (m, 2 H), 3.78 - 3.68 (dd, *J* = 13.4, 8.3 Hz, 2 H), 3.61 - 3.47

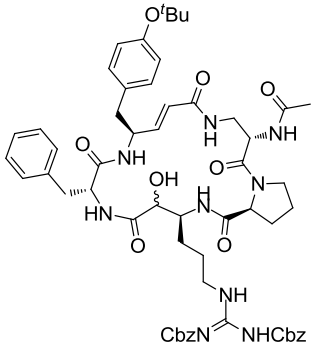
(m, 1 H), 3.36 (t, *J* = 6.8 Hz, 2 H), 2.92 (d, *J* = 8.1 Hz, 2 H), 2.78 - 2.46 (m, 3 H), 2.14 - 1.98 (m, 2 H), 1.33 (s, 9 H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ_C = 173.3, 172.7, 170.9, 169.5, 165.2, 163.4, 161.7, 156.0, 154.4, 141.1, 136.9, 136.7, 134.4, 130.2, 129.9, 129.5, 129.0, 128.9, 128.7, 128.5, 128.5, 128.0, 128.0, 127.2, 124.7, 124.0, 78.4, 74.7, 69.2, 66.9, 60.6, 54.6, 53.4, 51.7, 49.9, 47.6, 44.6, 39.8, 39.1, 31.9, 29.7, 29.7, 28.9, 25.7, 25.5, 24.6, 22.7, 14.1 ppm. HRMS (ESI^+): m/z calcd for $C_{55}H_{67}N_9O_{11}$: 1030.5038; found 1030.5022. [α^{22}_D] +31.1 (*c* 0.10, $CDCl_3$).

Note: For the ^{1}H NMR spectrum 6 protons were not observed (Proline CH_2 x 2 and exchangeable NH_2). These peaks are likely coincident with the water peak.

Diastereoisomer 2 **442**: LC-MS Method F (2 min): rt = 1.32 min, $[M+H]^+$ = 1030.9. IR (cm^{-1}): 3387, 3358, 2925, 1720, 1644, 1507, 1455, 1367, 1258, 1098. ^{1}H NMR (500 MHz, $CDCl_3$) δ_H = 9.67 - 9.58 (br.s, 1 H), 9.55 - 9.48 (br.s, 1 H), 7.92 (d, *J* = 10.4 Hz, 1 H), 7.81 (d, *J* = 9.1 Hz, 1 H), 7.51 - 7.31 (m, 15 H), 7.17 (d, *J* = 7.1 Hz, 2 H), 6.86 (app. d, *J* = 8.2 Hz, 2 H), 6.79 - 6.67 (m, 4 H), 5.92 (d, *J* = 15.4 Hz, 1 H), 5.31 - 5.18 (m, 5 H), 4.85 – 4.77 (m, 2 H), 4.58 (t, *J* = 13.7 Hz, 1 H), 4.49 (dd, *J* = 8.5, 4.1 Hz, 1 H), 4.36 - 4.17 (m, 3 H), 3.87 (td, *J* = 14.8, 4.4 Hz, 1 H), 3.73 (dd, *J* = 10.4, 6.0 Hz, 1 H), 3.62 - 3.47 (m, 1 H), 3.36 (m, 3 H), 3.03 (dd, *J* = 13.2, 4.4 Hz, 1 H), 2.82 (dd, *J* = 13.2, 9.9 Hz, 1 H), 2.69 (dd, *J* = 14.3, 4.7 Hz, 1 H), 2.61 (t, *J* = 11.5 Hz, 1 H), 2.49 (dd, *J* = 14.3, 6.0 Hz, 1 H), 1.33 (s, 9 H) ppm. ^{13}C NMR (151 MHz, $CDCl_3$) δ_C = 173.1, 173.0, 171.0, 169.1, 165.0, 163.5, 161.9, 156.4, 156.1, 154.9, 154.6, 140.4, 137.7, 136.9, 136.4, 134.5, 129.9, 129.8, 129.5, 128.9, 128.8, 128.6, 128.5, 128.1, 128.0, 127.4, 124.9, 124.0, 78.5, 72.5, 69.1, 66.9, 60.3, 56.2, 53.5, 53.3, 51.5, 50.1, 47.6, 44.6, 44.3, 40.4, 38.6, 31.9, 30.4, 29.7, 28.8, 26.0, 24.7, 22.7, 14.1 ppm. HRMS (ESI^+): m/z calcd for $C_{55}H_{67}N_9O_{11}$: 1030.5038; found 1030.5021. [α^{22}_D] -66.8 (*c* 0.10, $CDCl_3$).

Note: For the ^{1}H NMR spectrum 6 protons were not observed (Proline CH_2 x 2 and exchangeable NH_2). These peaks are likely coincident with the water peak.

Cyclo [-HN-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NHAc)-CO-] (443/444)



From Cyclo [-HN-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(O^{t}Bu)-HN-CH$_2$-CH(NH$_2$)-CO-] **(441/442)**

To a solution of diastereoisomers **441** and **442** (13.5 mg, 0.013 mmol) in DMF (1 mL) was added a solution of perfluorophenyl acetate (14.8 mg, 0.066 mmol) in DMF (0.25 mL). The reaction mixture was stirred at room temperature under a nitrogen atmosphere for 3 h. The solvent was removed *in vacuo* and the crude material purified by MDAP Method D (Si C_{18}, 40-60% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min) to give separated diastereoisomers. For each diastereoisomer, the solvent was reduced to 1/3 volume *in vacuo* and then extracted with EtOAc (3 x 25 mL) and DCM (1 x 25 mL). The combined organic phase was dried through a hydrophobic frit and the solvent removed *in vacuo* to afford two separated diastereoisomers as white solids, with an overall yield of 54%. [Diastereoisomer 1: **443** (2.2 mg, 16%, 84% purity by LC-MS); Diastereoisomer 2: **444** (5.4 mg, 38%, 89% purity by LC-MS)].

Diastereoisomer 1 **443**: LC-MS Method F (2 min): rt = 1.26 min, $[M+H]^+$ = 1072.5. HRMS (ESI^+): m/z calcd for $C_{57}H_{69}N_9O_{12}$: 1072.5144; found 1072.5129.

Diastereoisomer 2 **444**: LC-MS Method F (2 min): rt = 1.32 min, $[M+H]^+$ = 1072.4. HRMS (ESI^+): m/z calcd for $C_{57}H_{69}N_9O_{12}$: 1072.5144; found 1072.5132.

From Cyclo [-HN-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO-] **(434)**

Cyclic pentapeptide **434** (0.10 mmol) was swollen in DCM (5 mL) for 10 min. A solution of TFA:TIPS:DCM (1:1:98 v/v, 5 mL) was added to the resin, the suspension shaken vigorously for 20 min, filtered, and washed with DCM (2 x 3 mL). This was repeated three times. The filtrates were combined, quenched with pyridine (180 µL, 0.10 mmol). The solvent was removed *in vacuo* and the residue redissolved in DCM (5 mL). Solid perfluorophenylacetate (136 mg, 0.60 mmol) was added in three portions and the reaction mixture stirred at room temperature for 42 h. The solvent was removed *vacuo* and the crude material purified by MDAP Method D (Si C_{18}, 40-60% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 15 min). The solvent was removed to 1/3 volume *in vacuo* and then extracted with EtOAc (3 x 25 mL). The combined organic phase was dried through a hydrophobic frit, the solvent removed *in vacuo* and the residue triturated with cyclohexane (10 mL) then filtered to give two separated diastereoisomers as white solids, with a total yield of 10% over 12 steps. [Diastereoisomer 1: **443** (5.7 mg, 5%, 100% pure by LC-MS); Diastereoisomer 2: **444** (5.5 mg, 5%, 100% pure by LC-MS)].

Diastereoisomer 1 **443**: LC-MS Method F (2 min): rt = 1.28 min, $[M+H]^+$ = 1072.9. IR (cm^{-1}): 3298, 3278, 2931, 1720, 1645, 1508, 1454, 1368, 1259, 1098. ^{1}H NMR (500 MHz, $CDCl_3$) δ_H = 9.59 (br. s., 1 H), 9.47 (br. s., 1 H), 8.01 (d, *J* = 8.8 Hz, 1 H), 7.65 (br. s., 1 H), 7.48 - 7.25 (m, 20 H), 7.21 (d, *J* = 7.1 Hz, 2 H), 6.86 (app. d, *J* = 8.2 Hz, 2 H), 6.81 (dd, *J* = 15.6, 2.7 Hz, 1 H), 6.73 (d, *J* = 8.2 Hz, 2 H), 5.89 (d, *J* = 14.8 Hz, 1 H), 5.26 (s, 2 H), 5.22 (d, *J* = 8.8 Hz, 1 H), 5.15-5.10 (m, 1 H), 4.90 - 4.73 (m, 2 H), 4.48 - 4.35 (m, 2 H), 4.30 (td, *J* = 12.1, 4.9 Hz, 1 H), 4.05 (qd, *J* = 9.3, 3.6 Hz, 1 H), 3.80 (dt, *J* = 15.1, 5.5 Hz, 1 H), 3.76 - 3.70 (m, 1 H), 3.63 - 3.55 (m, 1 H), 3.46 - 3.39 (m, 1 H), 2.96 - 2.87 (m, 2 H), 2.76 - 2.67 (m, 2 H), 2.54 (dd, *J* = 13.9, 6.4 Hz, 1 H), 2.08-2.00 (m, 2 H), 1.96 (s, 3 H), 1.90 - 1.76 (m, 2 H), 1.46 (m, 2 H), 1.33 (s, 9 H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$): δ_C = 172.5, 170.9, 170.3, 169.7, 169.5, 165.1, 163.4, 161.7, 156.0, 154.5, 141.4, 136.9, 136.6, 134.5, 130.1, 129.9 (2 C), 129.5 (2 C), 129.0 (2 C), 128.9 (2 C), 128.7 (2 C), 128.5 (2 C), 128.5 (2 C),

128.0 (2 C), 127.2, 124.3, 124.0 (2 C), 78.5, 69.2, 66.9, 60.7, 54.5, 53.4, 50.0, 49.5, 48.1, 44.6, 40.8, 39.9, 38.9, 29.8, 28.9 (3 C), 25.7, 25.3, 24.7, 23.0, 23.0 ppm. HRMS (ESI^+): m/z calcd for $C_{57}H_{70}N_9O_{12}$: 1072.5144; found 1072.5129. $[\alpha^{22}_D]$ +88.1 (*c* 0.10, $CHCl_3$).

Diastereoisomer 2 **444**: LC-MS Method F (2 min): rt = 1.33 min, $[M+H]^+$ = 1072.9. IR (cm^{-1}): 3294, 3277, 2926, 1720, 1644, 1507, 1455, 1367, 1257, 1100. ^1H NMR (500 MHz, $CDCl_3$) δ_H = 9.65 (br. s., 1 H), 9.50 (br. s, 1 H), 7.99 - 7.85 (m, 2 H), 7.50 - 7.31 (m, 20 H), 7.16 (d, *J* = 7.8 Hz, 2 H), 6.87 (app. d, *J* = 8.3 Hz, 2 H), 6.74 (d, *J* = 15.9 Hz, 1 H), 6.72 - 6.65 (m, 2 H), 5.88 (d, *J* = 14.9 Hz, 1 H), 5.27 (s, 2 H), 5.24 (m, 2 H), 4.85-4.81 (m, 3 H), 4.60 - 4.50 (m, 1 H), 4.45 (dd, *J* = 8.3, 4.8 Hz, 1 H), 4.38 - 4.22 (m, 3 H), 3.94 (d, *J* = 5.5 Hz, 1 H), 3.82 (d, *J* = 14.9 Hz, 1 H), 3.62-3.54 (m, 1 H), 3.48 - 3.36 (m, 2 H), 3.04 (dd, *J*= 13.1, 4.3 Hz, 1 H), 2.82 (dd, *J* = 13.3, 10.3 Hz, 1 H), 2.75 - 2.65 (m, 2 H), 2.56 - 2.47 (m, 1 H), 2.05 - 1.97 (m, 1 H), 1.94 (s, 3 H), 1.92 - 1.79 (m, 2 H), 1.37 - 1.32 (s, 9H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$): δ_C = 172.7, 170.9, 170.0, 169.5, 169.1, 164.9, 163.5, 161.9, 156.0, 154.6, 140.7, 136.8, 136.3, 134.5, 129.8 (2 C), 129.5 (2 C), 128.9 (2 C), 128.8 (2 C), 128.8 (2 C), 128.7 (2 C), 128.5 (2 C), 128.2, 128.1 (2 C), 127.4, 124.5, 124.1 (2 C), 78.5, 69.2, 67.0, 60.5, 56.1, 53.8, 50.2, 49.2, 48.1, 44.2, 41.0, 40.4, 38.6, 29.7, 28.9 (3 C), 25.3, 24.7, 24.7, 23.0, 23.0 ppm. HRMS (ESI^+): m/z calcd for $C_{57}H_{70}N_9O_{12}$: 1072.5144; found 1072.5132. $[\alpha^{22}_D]$ -9.5 (*c* 0.10, $CHCl_3$).

Attempted conditions for the oxidation and deprotection of Cyclo [-HN-Pro-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NHAc)-CO-] (443/444)

Entry 1: To a solution of diastereoisomer 1 **443** (1.5 mg, 1.4 µmol) in MeCN:$CHCl_3$ (1:1 v/v, 1 mL) was added sIBX (3.5 mg, 5.6 µmol) and the reaction mixture stirred at 55 °C under a nitrogen atmosphere for 1.5 h, then 65 °C for 3 h, and finally room temperature for 16 h. Analysis by LC-MS at each stage of heating indicated that no reaction had occurred.

Entry 2: To a solution of diastereoisomer 1 **443** (1.5 mg, 1.4 μmol) in DMSO (0.75 mL) was added sIBX (15.5 mg, 24.8 μmol) and the reaction mixture stirred at 100 °C under a nitrogen atmosphere for 2.5 h. Analysis by LC-MS indicated that the material had degraded with none of the α-ketoamide observed.

Entry 3: To a suspension of mixed diastereoisomers 1 **443** and 2 **444** (1:1 w/w, 7.0 mg, 6.5 μmol) in MeCN (0.7 mL) was added DMP (6.9 mg, 0.016 mmol) and the reaction mixture stirred at 75 °C under a nitrogen atmosphere for 3 h. Additional DMP (6.9 mg, 0.016 mmol) was added and the reaction mixture stirred at 75 °C for a further 4 h. Analysis by LC-MS indicated the material had degraded with none of the α-ketoamide observed.

Entry 4: To a suspension of diastereoisomer 2 **444** (1.8 mg, 1.7 μmol) in MeCN (1 mL) was added DMP (3.0 mg, 7.0 μmol) and the reaction mixture stirred at 75 °C under a nitrogen atmosphere for 1 h. The solvent was removed *in vacuo* and the residue partitioned between DCM (10 mL) and sat. $NaHCO_{3(aq)}$ (10 mL). The organic layer was separated and the aqueous layer extracted with DCM (2 x 10 mL). The combined organic phase was dried through a hydrophobic frit and the solvent removed *in vacuo*. A solution of TFA:H_2O:thioanisole (95:5:15 v/v, 1.15 mL) was added and the reaction mixture stirred at room temperature for 17 h. The solvent was removed *in vacuo* and the residue washed with cold Et_2O (1 mL). The grey precipitate was collected by centrifugation (2500 rpm, 5 min), the solvent decanted and the crude material washed twice more. Analysis by LC-MS indicated that the final product had degraded.

LC-MS Method E (3 min):
Protected acetal: rt = 1.69 min, $[M+H]^+$ = 1088.5.
Protected α-ketoamide: rt = 1.92 min, $[M+H]^+$ = 1070.4.

Entry 5: To a suspension of mixed diastereoisomers 1 **443** and 2 **444** (1:1 w/w, 10.0 mg, 9.3 μmol) in MeCN (5 mL) was added DMP (19.8 mg, 0.047 mmol) and the reaction mixture stirred at 75 °C under a nitrogen atmosphere for 2.5 h. Additional DMP (19.8 mg, 0.047 mmol) was added and the reaction mixture stirred at 75 °C for a further 1 h. The solvent was removed *in vacuo* and the residue partitioned between DCM (25 mL) and sat. $NaHCO_{3(aq)}$ (25 mL). The organic layer was separated and the aqueous layer extracted with DCM (2 x 25 mL). The combined organic phase was dried through a hydrophobic frit and concentrated *in vacuo*. A solution of TFA:H_2O:thioanisole (95:5:15 v/v, 1.15 mL) was added and the reaction mixture stirred at room temperature for 3 h. The reaction mixture was poured into cold Et_2O (3 mL), the white precipitate collected by centrifugation (2500 rpm, 5 min) the solvent decanted, and the crude material washed twice more. Analysis by LC-MS indicated that the final product had degraded.

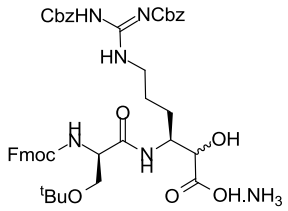
LC-MS Method F (3 min): rt = 0.94 min, $[M+H]^+$ = 1088.4.

Entry 6: To a suspension of mixed diastereoisomers 1 **443** and 2 **444** (7.0 mg, 6.5 μmol) in MeCN (1.4 mL) was added DMP (12.5 mg, 0.029 mmol) and the reaction mixture stirred at 75 °C under a nitrogen atmosphere for 1.5 h. The reaction mixture was quenched with 10% $Na_2S_2O_{3(aq)}$ (5 mL) with stirring over 10 min and extracted with DCM (3 x 5 mL) and EtOAc (3 x 5 mL). The combined organic phase was dried through a hydrophobic frit and the solvent removed *in vacuo*. A solution of TFA:H_2O:thioanisole (95:5:15 v/v, 1.15 mL) was added and the reaction mixture stirred at room temperature for 1.25 h. Analysis by LC-MS did not indicate the presence of the final product.

LC-MS Method E (3 min):
Protected α-ketoamide: rt = 1.68 min, $[M+H]^+$ = 1070.6
Protected acetal: rt = 1.92 min, $[M+H]^+$ = 1088.4.

3.10.5. Chapter 2: Synthesis of a P2 modified cyclotheonamide analogue

Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-CHOHCO$_2$H.NH$_3$ (445)

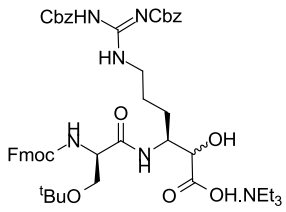


A solution of **410** (550 mg, 0.96 mmol) in 4M HCl in dioxane (5.0 mL, 20.0 mmol) was stirred at room temperature for 1 h. The reaction mixture was cooled to 0 °C and quenched with DIPEA (4.0 mL, 22.9 mmol) with stirring over 20 minutes. A yellow precipitate was observed. A pre-activated solution of HATU (420 mg, 1.1 mmol), HOAt (150 mg, 1.1 mmol), DIPEA (0.40 mL, 2.3 mmol) and Fmoc-Ser(^{t}Bu)-OH (424 mg, 1.1 mmol) in DMF (5 mL) was added dropwise and the mixture stirred at room temperature for 3 h. Additional pre-activated HATU (55 mg, 0.15 mmol), HOAt (20 mg, 0.15 mmol), DIPEA (50 µL, 0.29 mmol) and Fmoc-Ser(^{t}Bu)-OH (55 mg, 0.15 mmol) in DMF (0.5 mL) was added dropwise and the reaction mixture stirred for a further 1.5 h. The reaction mixture was partitioned between EtOAc (200 mL) and 10% $NaHCO_{3(aq)}$ (200 mL). The organic layer was separated and the aqueous layer extracted with EtOAc (2 x 50 mL). The combined organic phase was washed with 10% citric $acid_{(aq)}$ (200 mL) and sat. $LiCl_{(aq)}$ (200 mL). The organic phase was separated and dried through a hydrophobic frit. The solvent was removed *in vacuo* and the crude material purified by automated reverse phase chromatography (Si C_{18}, 120 g, 40-70% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 40 min). The solvent was removed *in vacuo* and the residual water removed azeotropically with toluene (3 x 25 mL) and DCM (3 x 25 mL) to afford **445** (356 mg, 43%) as a white solid.

LC-MS Method F (2 min): rt = 1.16 min, $[M+H]^+$ = 838.7. IR (cm^{-1}): 3376, 3317, 2965, 1718, 1647, 1617, 1500, 1496, 1467, 1392, 1244, 1229, 1192, 1098. ^1H NMR (500 MHz, DMSO-d_6): δ_H = 9.27 - 8.87 (br. s, 2 H), 7.89 (d, *J* = 7.4 Hz, 2 H), 7.78 – 7.67 (m, 2 H), 7.47 - 7.27 (m, 16 H), 5.30 – 5.21 (m, 2 H), 5.08 – 5.02 (m, 2 H), 4.33 - 3.92 (m, 4 H), 3.89 - 3.71 (m, 2 H), 3.60 – 3.35 (m, 4 H), 2.31 (s, 1 H), 1.77 - 1.17 (m, 4 H), 1.08 (s, 9 H) ppm. ^{13}C NMR (126 MHz, DMSO-d_6): δ_C = 174.7, 174.3, 169.5, 169.3, 163.4, 160.1, 156.4, 156.2, 155.5, 144.5, 144.4, 144.2, 143.0, 141.2, 139.9, 137.9, 137.8, 137.6, 135.8, 135.8, 129.4, 129.4, 129.0, 129.0, 128.8, 128.8, 128.7, 128.6, 128.3, 128.3, 128.3, 128.3, 128.1, 128.1, 127.8, 127.5, 125.8, 121.8, 120.5, 120.5, 110.2, 74.1, 73.3, 73.2, 73.1, 71.8, 68.6, 68.6, 66.6, 66.2, 62.6, 56.4, 55.9, 52.0, 51.7, 51.4, 47.1, 45.3, 27.6, 27.6, 27.6, 26.0, 25.8, 21.5 ppm. HRMS (ESI^+): m/z calcd for $C_{45}H_{52}N_5O_{11}$: 838.3657; found 838.3658.

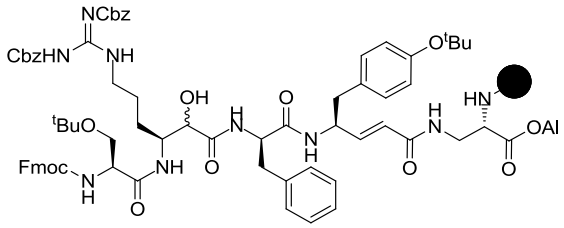
Note: The exchangeable CO_2H proton was not observed in the ^1H NMR spectrum. The ^{13}C NMR is a mixture of diastereoisomers and all observed peaks are reported.

Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-CHOHCO$_2$H.NEt$_3$ (446)



A solution of **445** (256 mg) in the minimum amount of MeOH was loaded onto an aminopropyl column (5 g) under gravity. The column was washed with MeOH (3 CV) then eluted with 5-10% Et_3N:MeOH (6 CV). The solvent was removed *in vacuo* to afford **446** as a white solid which was used immediately in subsequent reactions.

Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO$_2$Al (447)

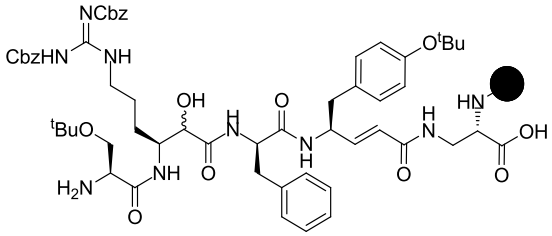


Tripeptide **430** was prepared as previously described (see Tables 24 and 25) with a loading of 0.36 mmol g^{-1}, as determined by an 'Fmoc Loading Test'.

Tripeptide **430** (521 mg, 0.36 mmol g^{-1}, 0.19 mmol) underwent an Fmoc deprotection followed by a 120 min manual coupling (See Table 25) with a pre-activated solution of **446** (256 mg, 0.30 mmol, 1.5 equiv.), which had been freshly prepared by salt exchange with **445** on an aminopropyl column, HATU (114 mg, 0.30 mmol), DIPEA (52 µL, 0.3 mmol), and HOAt (41 mg, 0.30 mmol) in DMF (5 mL). The reaction was shown to be complete with a negative 'chloranil test' followed by a small-scale cleavage and analysis by LC-MS.

LC-MS Method D (3 min): rt = 1.84/1.85 min, $[M+H]^+$ = 1357.0/1357.0.

H-Ser(^{t}Bu)-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NH®)-CO$_2$H (448)



Allyl Deprotection:

Pentapeptide **447** (0.19 mmol) was subjected to allyl deprotection conditions (see Table 26) using a solution of Pd(PPh$_3$)$_4$ (86 mg, 0.074 mmol) and phenylsilane (0.60 mL, 4.9 mmol) in dry DCM (5 mL) five times. The resin was washed with DCM (5 x 30 sec, 5 mL) after each treatment and a fresh solution of deprotection

reagents used each time. The reaction was shown to be complete by conducting a small-scale cleavage and analysis by LC-MS.

LC-MS Method D (3 min): rt = 2.16/2.20 min, $[M+H]^+$ = 1317.0/1317.0.

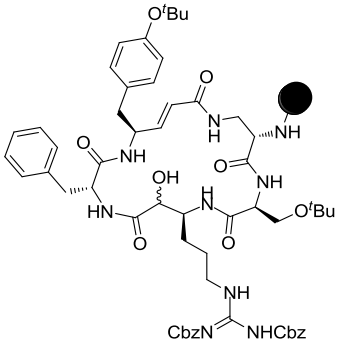
Fmoc Deprotection:

The pentapeptide Fmoc-Ser(^{t}Bu)-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH_2-CH(NH®)-CO_2H (0.19 mmol) was subjected to Fmoc deprotection conditions (see Table 26) using a solution of 20% piperidine in DMF (5 mL, 2 x 45 sec). The reaction was shown to be complete by conducting a scmall-scale cleavage and analysis by LC-MS.

Note: Care must be taken to minimise the reaction time to avoid additional Cbz deprotection.

LC-MS Method F (2 min): rt = 1.20/1.23 min, $[M+H]^+$ = 1094.8/1094.8.

Cyclo[-HN-Ser(^{t}Bu)-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH_2-CH(NH®)-CO-] (449)

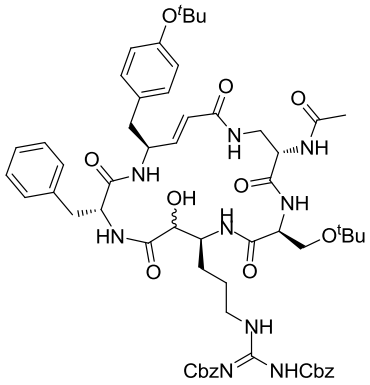


Prior to cyclisation, the resin was washed with 10% DIPEA in DMF solution (2 x 5 mL) to remove any residual piperidine.

Pentapeptide **448** (0.19 mmol) was cyclised (see Table 26) using a solution of PyBOP (290 mg, 0.56 mmol), HOAt (76 mg, 0.56 mmol) and DIPEA (0.19 mL, 1.1 mmol) in DMF (5 mL). The reaction was shown to be complete by conducting a small-scale cleavage and analysis by LC-MS.

LC-MS Method F (2 min): rt = 1.34/1.38 min, $[M+H]^+$ = 1076.5/1076.5.

Cyclo[-HN-Ser(^{t}Bu)-Arg(Cbz)$_2$-αHA-D-Phe-vTyr(^{t}Bu)-HN-CH$_2$-CH(NHAc)-CO-] (450/451)



Cyclic pentapeptide **449** (0.19 mmol) was swollen in DCM (5 mL) for 10 min. A solution of TFA:TIPS:DCM (1:1:98 v/v, 5 mL) was added to the resin, the suspension shaken vigorously for 20 min, filtered, and washed with DCM (2 x 3 mL). This was repeated three times. The filtrates were combined, quenched with pyridine (180 µL, 0.10 mmol) with stirring over 20 min. The solvent was removed *in vacuo* and the residue redissolved in DMF (2 mL). Solid perfluorophenylacetate (125 mg, 0.55 mmol) was added in one portion and the reaction mixture stirred at room temperature for 16.5 h. The solvent was removed *in vacuo* and the crude material purified by MDAP Method D (Si C_{18}, 50-70% MeCN-H_2O, $(NH_4)_2CO_3$ modifier, 25 min). The solvent was removed to 1/3 volume *in vacuo* and then extracted with EtOAc (3 x 25 mL) and DCM (25 mL). The combined organic phase was dried through a hydrophobic frit, the solvent removed *in vacuo* and the residue were triturated with Et_2O (10 mL) then filtered and dried *in vacuo* overnight to give two separated diastereoisomers as white solids, with a total yield of 17% over 12 steps. [Diastereoisomer 1: **450** (19.4 mg, 9%, 100% purity by LC-MS); Diastereoisomer 2: **451** (16.5 mg, 8%, 100% purity by LC-MS)].

Diastereoisomer 1 **450**: LC-MS Method F (2 min): rt = 1.40 min, $[M+H]^+$ = 1118.4. IR (cm^{-1}): 3378, 3285, 2977, 2928, 1721, 1653, 1508, 1455, 1367, 1240, 1195, 1100. ^1H NMR (500 MHz, $CDCl_3$) δ_H = 9.52 - 9.45 (br. s., 1 H), 9.44 - 9.36 (br. s., 1 H), 7.48 - 7.31 (m, 19 H), 7.16 (d, *J* = 7.1 Hz, 2 H), 6.87 - 6.79 (m, 3 H), 6.65 (d, *J* = 7.4 Hz, 2 H), 5.79 (d, *J* = 15.4 Hz, 1 H), 5.33 - 5.11 (m, 5 H), 4.87 (d, *J* = 5.2 Hz, 1 H), 4.53 – 4.45 (m, 1 H), 4.45 - 4.19 (m, 4 H), 4.05 - 3.99 (m, 1 H), 3.93 - 3.82 (m, 1 H), 3.67 - 3.53 (m, 2 H), 3.42 (t, *J* = 7.4 Hz, 1 H), 3.10 - 2.99 (m, 1 H), 2.87 - 2.64 (m, 3 H), 2.52 (dd, *J* = 13.6, 5.6 Hz, 1 H), 1.96 (s, 3 H), 1.66 - 1.47 (m, 4 H), 1.34 (s, 9 H), 1.06 (s, 9 H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$): δ_C = 171.0, 170.9, 170.6, 169.5, 169.3, 165.0, 163.5, 161.3, 155.9, 154.6, 141.3, 137.0, 136.5, 134.6, 129.9 (2 C), 129.4 (2 C), 128.9 (2 C), 128.8 (2 C), 128.8 (2 C), 128.6 (2 C), 128.5 (2 C), 128.0, 127.6 (2 C), 127.3, 124.4, 123.8 (2 C), 78.3, 74.0, 72.2, 69.0, 66.8, 61.5, 56.2, 53.5, 53.0, 52.5, 50.2, 44.1, 40.2, 40.0, 38.6, 29.1, 28.9 (3 C), 27.4 (3 C), 25.6, 23.1 ppm. HRMS (ESI^+): m/z calcd for $C_{59}H_{76}N_9O_{13}$: 1118.5563; found 1118.5559. $[\alpha^{24}{}_D]$ +7.5 (*c* = 0.1, MeOH).

Diastereoisomer 2 **451**: LC-MS Method F (2 min): rt = 1.35 min, $[M+H]^+$ = 1118.5. IR (cm^{-1}): 3386, 3276, 3070, 2976, 2929, 1720, 1656, 1612, 1551, 1508, 1455, 1367, 1251, 1200, 1100. ^1H NMR (500 MHz, $CDCl_3$) δ_H = 9.38 (br. s., 1 H), 9.29 (br. s., 1 H), 7.55 – 7.45 (m, 3 H), 7.45 - 7.36 (m, 9 H), 7.33 (t, *J* = 7.3 Hz, 2 H), 7.21 (t, *J* = 7.4 Hz, 2 H), 7.12 (t, *J* = 6.6 Hz, 2 H), 7.08 (d, *J* = 7.4 Hz, 2 H), 6.99 - 6.84 (m, 6 H), 5.81 (d, *J* = 15.1 Hz, 1 H), 5.30 - 5.09 (m, 5 H), 4.98 - 4.86 (m, 1 H), 4.57 – 4.40 (m, 3 H), 4.16 - 3.98 (m, 3 H), 3.86 – 3.74 (m, 2 H), 3.67 - 3.50 (m, 2 H), 2.95 - 2.73 (m, 5 H), 1.88 (s, 3 H), 1.63 - 1.38 (m, 4 H), 1.33 (s, 9 H), 1.17 (s, 9 H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$): δ_C = 172.3, 171.5, 170.8, 170.2, 169.8, 165.7, 163.4, 160.6, 155.7, 154.2, 136.7, 136.6, 134.7, 131.2, 129.8 (2 C), 129.1 (2 C), 128.9 (2 C), 128.8 (2 C), 128.6 (2 C), 128.5 (2 C), 128.4 (2 C), 128.2 (2 C), 128.1, 126.9, 124.0 (2 C), 123.4, 78.4, 74.6, 72.1, 69.0, 67.1, 60.8, 54.5, 53.9, 53.0, 52.2, 50.5, 44.6, 40.1, 39.2, 37.8, 28.8 (3 C), 27.3 (3 C), 26.9, 24.2, 22.9 ppm. HRMS (ESI^+): m/z calcd for $C_{59}H_{76}N_9O_{13}$: 1118.5563; found 1118.5554. $[\alpha^{24}{}_D]$ -6.3 (*c* = 0.1, MeOH).

3.11. Appendix

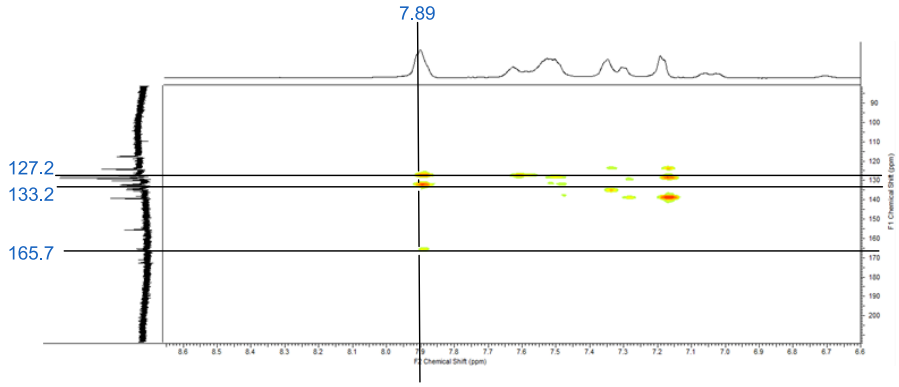
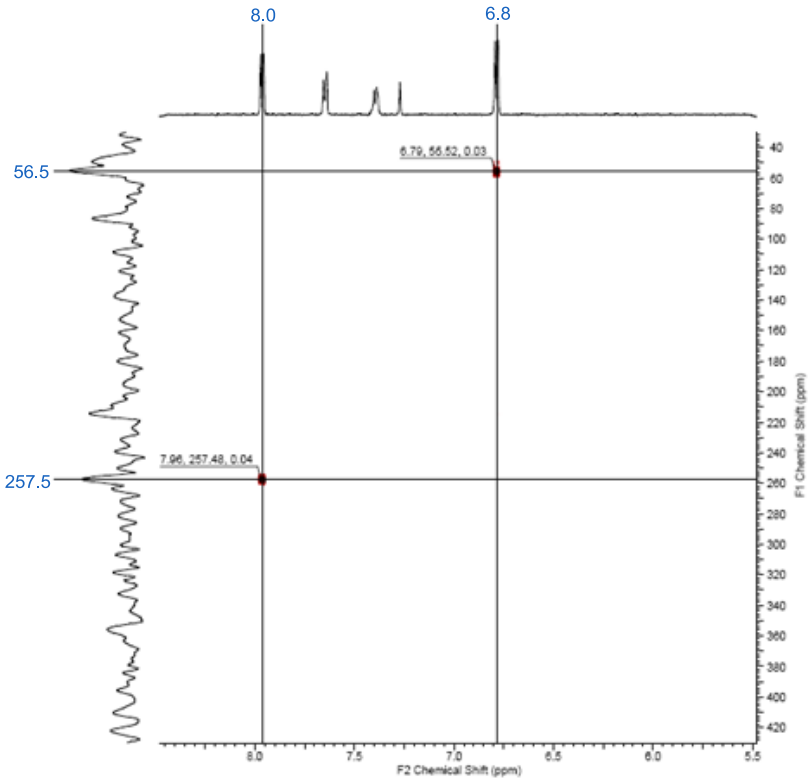


Figure A1: $^{1}H^{13}C$ HMBC 2D NMR spectrum for hydrazide **93**.

a)



b)

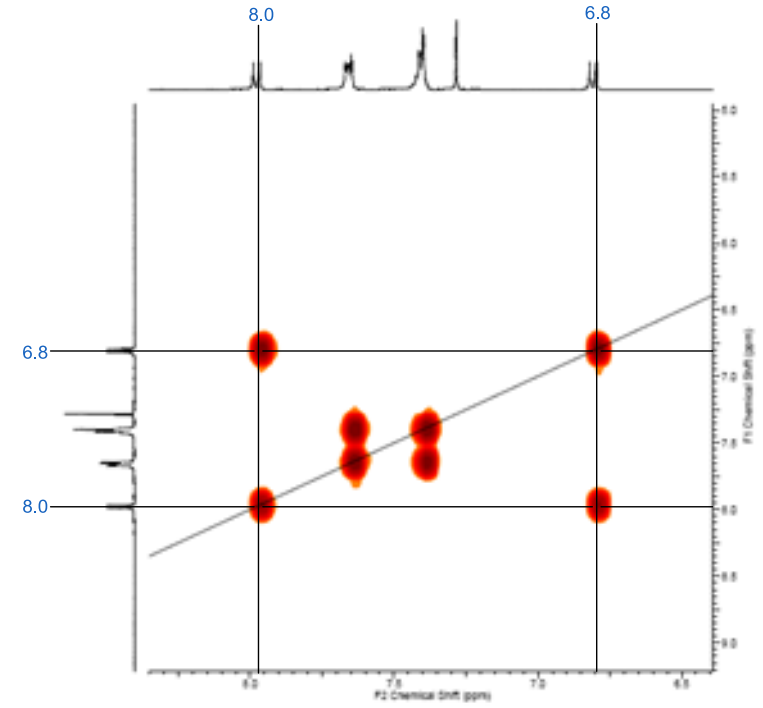
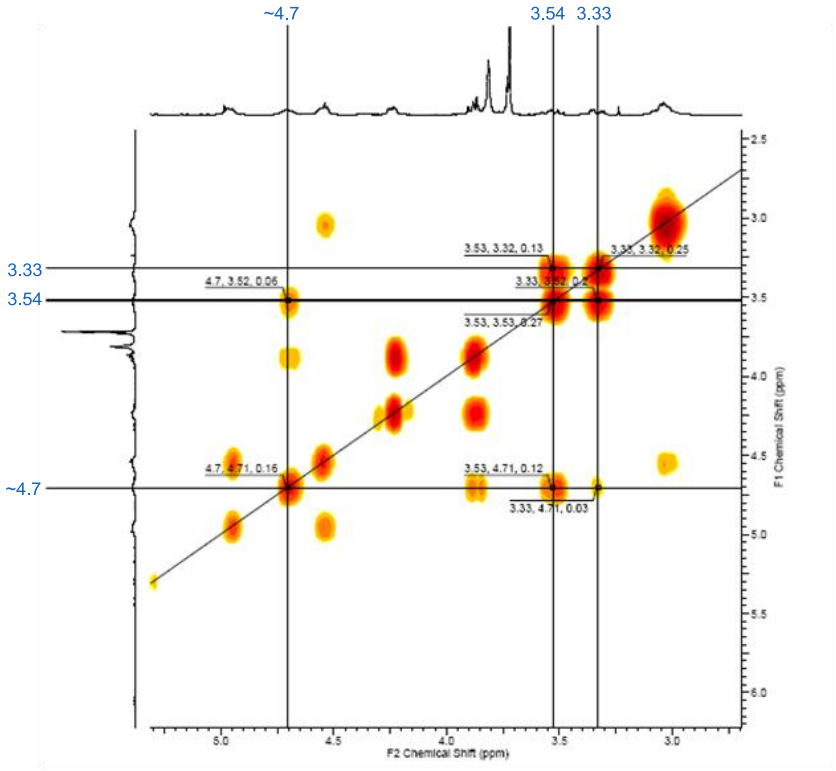


Figure A2: (a) $^{15}N^{1}H$ HMBC NMR spectrum and (b) $^{1}H^{1}H$ COSY NMR spectrum of 2,4,5-triaryl-1,2,3-triazole **116.**

a)



b)

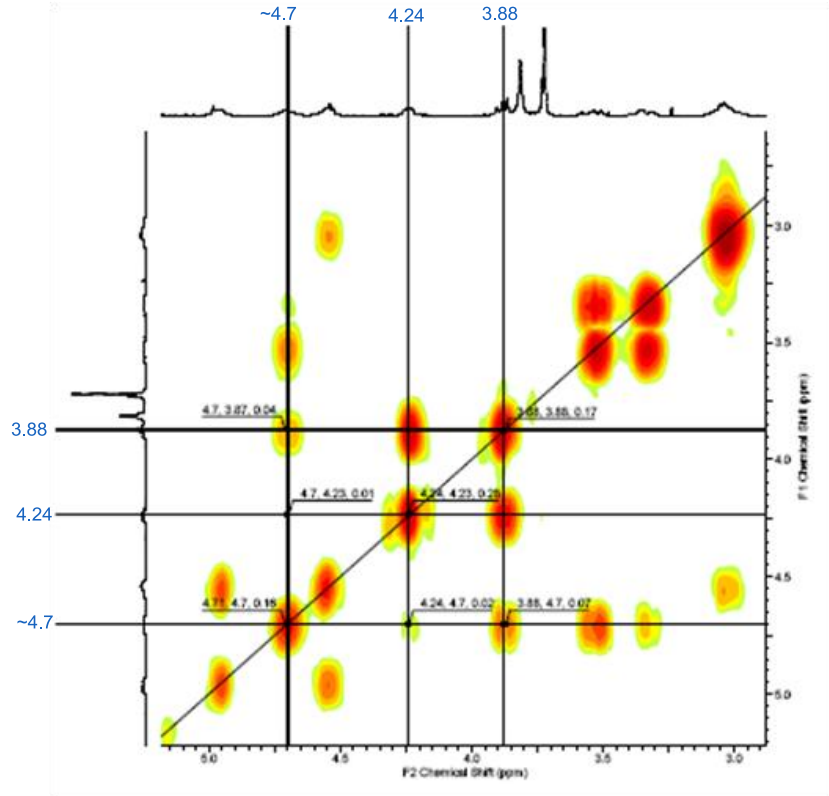


Figure A3: $^{1}H^{1}H$ COSY NMR spectrum of pyrazoline adduct **118**. (a) correlation between pyr-CH_2 and pyr-CH protons and (b) correlation between pyr-CH and OCH_2 protons.

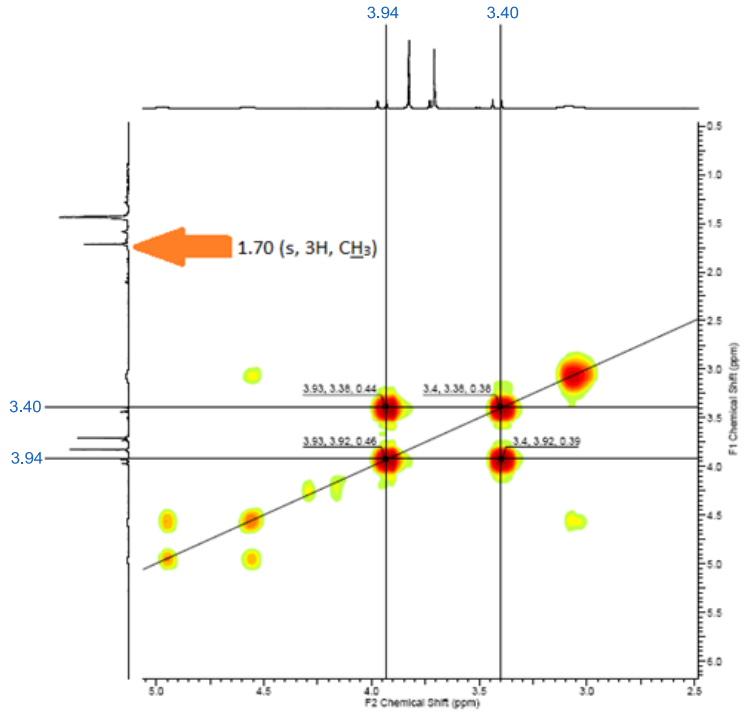


Figure A4: $^{1}H^{1}H$ COSY NMR spectrum of pyrazoline adduct **119** - observation of pyr-CH_2 and Me groups.

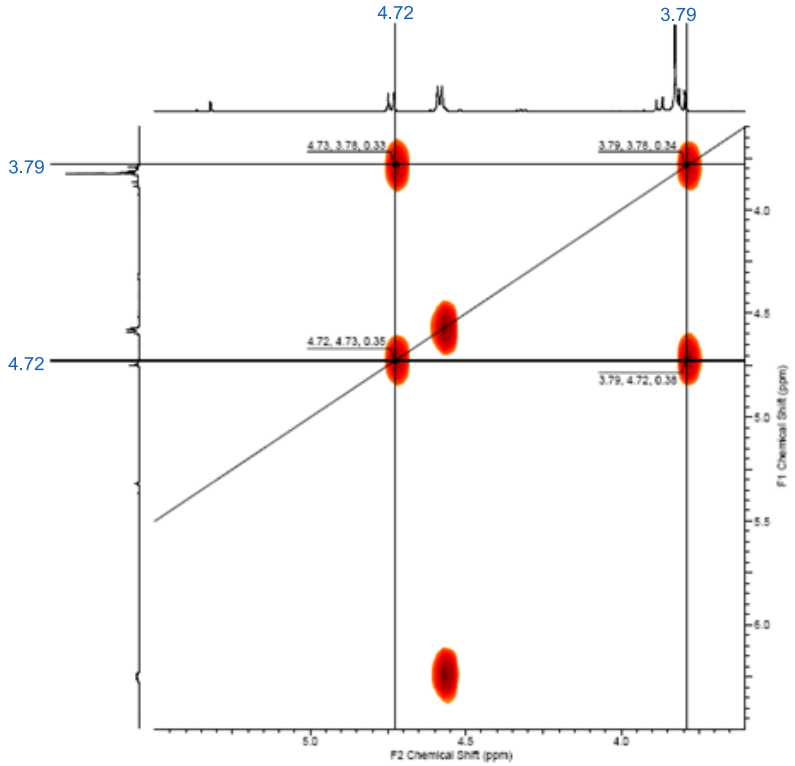
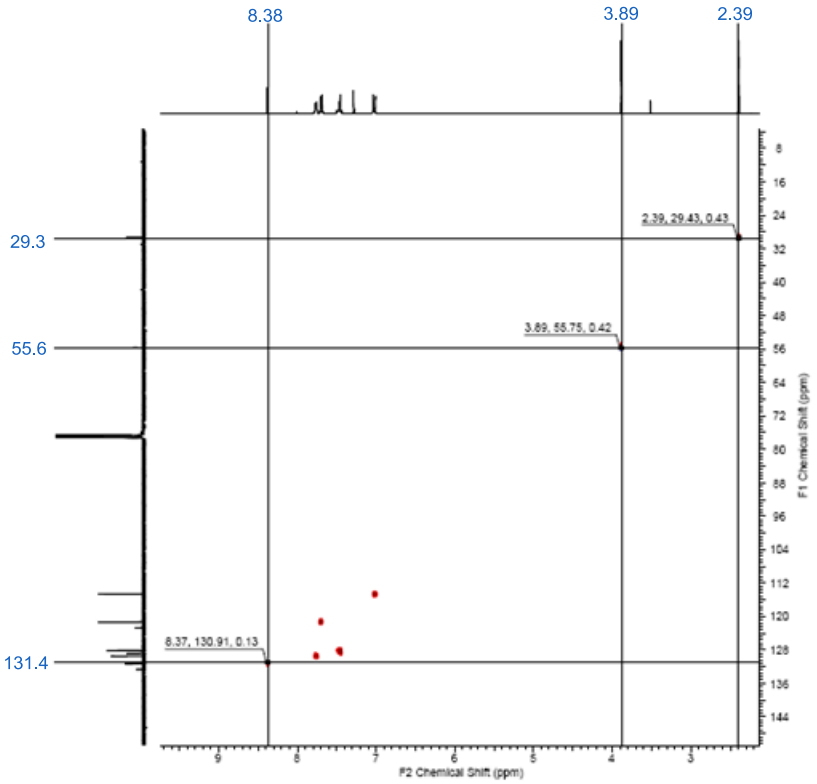


Figure A5: $^{1}H^{1}H$ COSY NMR spectrum of pyrazoline adduct **143** - observation of pyr-CH group.

a)



b)

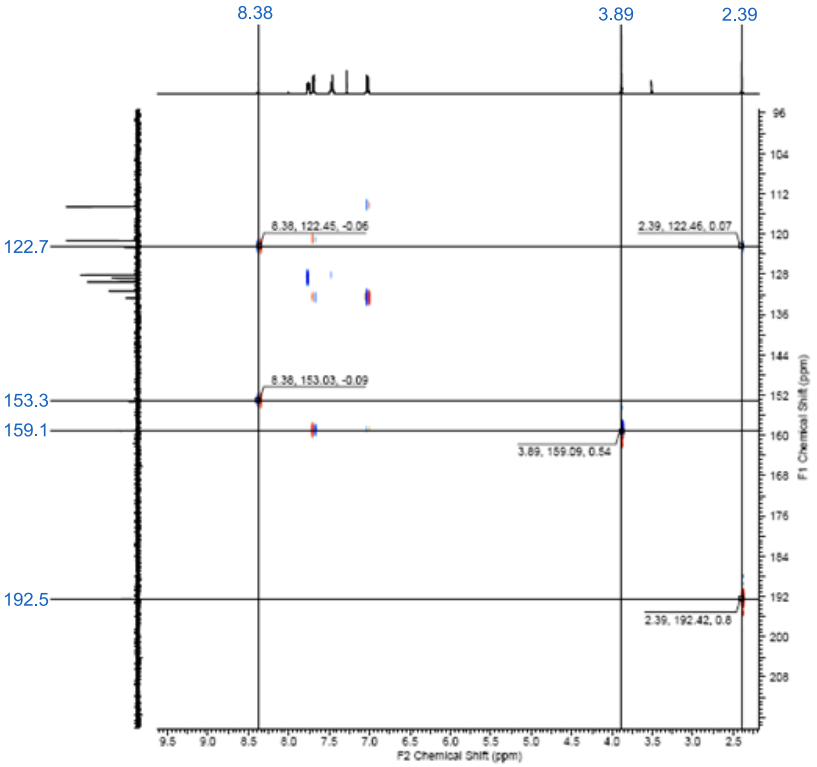
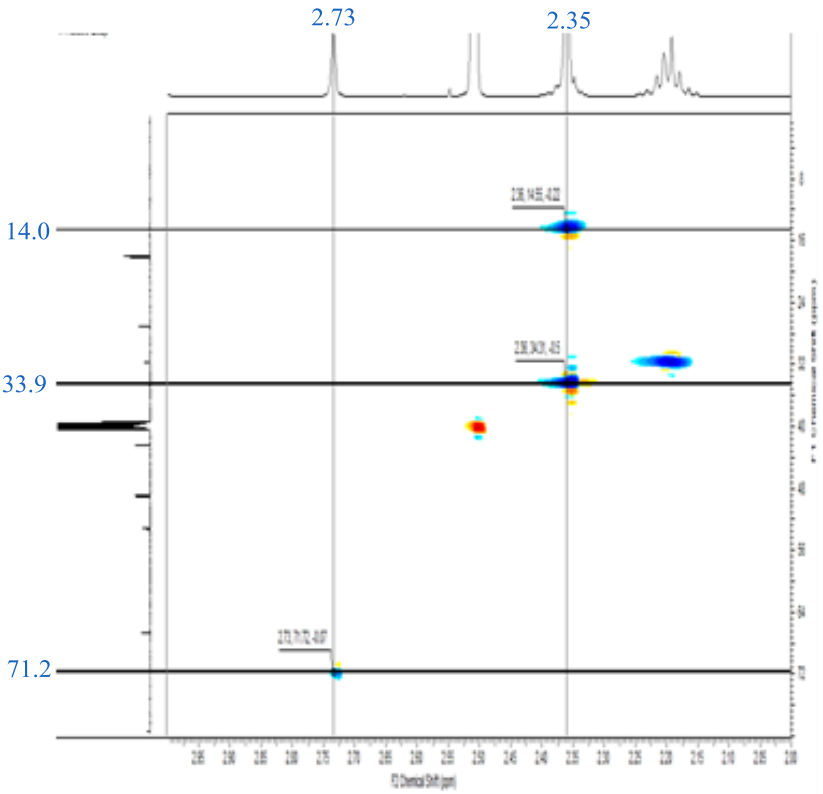


Figure A6: (a) $^{1}H^{13}C$ HSQC NMR spectrum and (b) $^{1}H^{13}C$ HMBC NMR spectrum of pyrazole **145**.

a)



b)

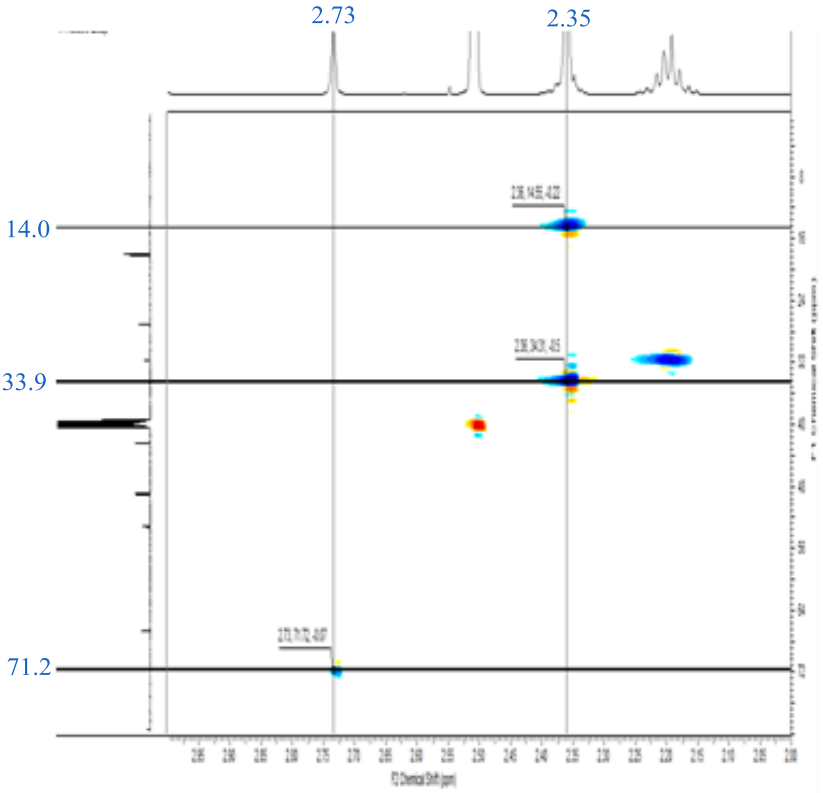
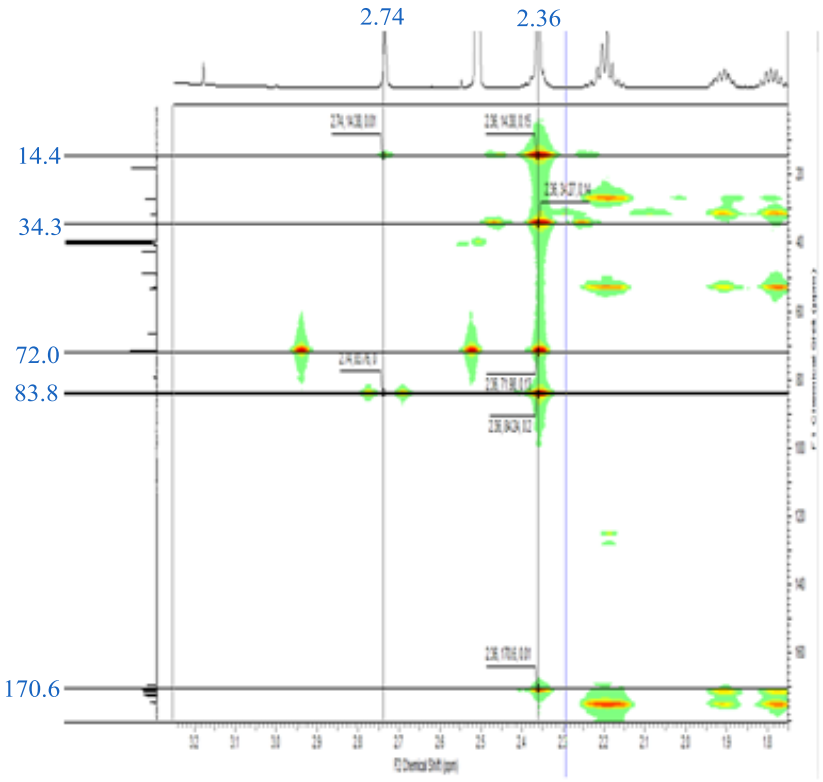


Figure A7: $^{1}H^{13}C$ HSQC NMR spectra for (a) starting material **234** and (b) hydrazide product **237**.

a)



b)

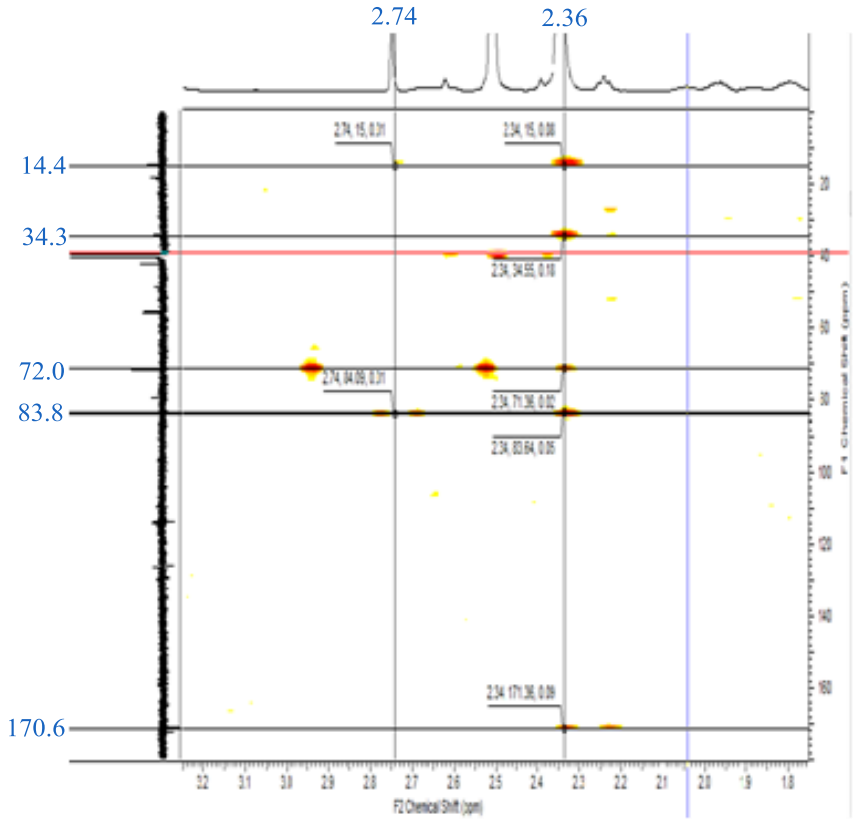


Figure A8: $^{1}H^{13}C$ HMBC NMR spectra for (a) starting material **234** and (b) hydrazide product **237**.

3.12. References

1. Verdine, G.L.; Hilinski, G.J.; *Methods Enzymo.*, **2012**, *503*, 3-33.
2. Milletti, F.; *Drug Discov. Today,* **2012**, *17*, 850-860.
3. Edwards, C.M.B.; Cohen, M.A.; Bloom, S.R. *Q JM,* **1999**; *92*, 1-4.
4. Lu, Y.; Yang, J.; Sega, E. *AAPS J.* **2006**, *8*, 466-478.
5. Pedersen, S.L.; Tofteng, A.P.; Malik, L.; Jensen, K.; *J. Chem. Soc. Rev*., **2012**, *41*, 1826-1844.
6. For a few examples see: (a) Rand, A.C.; Leung, S.S.F.; Eng, H.; Rotter, C.J.; Sharma, R.; Kalgutkar, S.; Zhang, Y.; Varma, M.V; Farley, K.A.; Khunte, B.; Limberakis, C.; Price, D.A.; Liras, S.; Mathiowetz, A.M.; Jacobson, P.; Lokey, R.S.; *Med. Chem. Commun.* **2012**, *3*, 1282-1289; (b) Ovadia, O.; Greenberg, S.; Laufer, B.; Gilon, C.; Hoffman, A.; Kessler, H.; *Expert Opin. Drug Discov.* **2010**, *5*, 655-671; (c) Bock, J.E.; Gavenonis, J.; Kritzer, J.A.; *ACS Chem. Biol.,* **2013**, *8*, 488-499; (d) Ovadia, O.; Linde, Y.; Haskell-luevano, C.; Dirain, M.L.; Sheynis, T.; Jelinek, R.; Gilon, C.; Hoffman, A.; *Bioorg. Med. Chem.,* **2010**, *18*, 580-589; (e) Bundgaard, H.; *Adv. Drug Deliv. Rev.,* **1992**, *8*, 1-38.
7. Tedesco, D.; Haragsim, L.; *J. Transplant*, **2012***, 2012*, 230386-230393.
8. Choc, M.G.; *Int. J. Dermatol.,* **1997**, *36*, 1-6.
9. Renner, M.K.; Shen, Y.C.; Cheng, X.C.; Jensen, P.R.; Frankmoelle, W.; Kauffman, C.A.; Fenical, W.; Lobkovsky, E.; Clardy, J.; *J. Am. Chem. Soc.,* **1999**, *121*, 11273-11276.
10. Sonda, S.; Sala, G.; Ghidoni, R.; Sonda, S.; Sala, G.; Ghidoni, R.; Hemphill, A.; *Antimicrob. Agents Chemother.,* **2005**, *49*, 1794-1801.
11. Hoskin, D.W.; Ramamoorthy, A.; *Biochim. Biophys. Acta,* **2008**, *1778*, 357-375.
12. Margiolaki, I.; Giannopoulou, A.E.; Wright, J.P.; Knight, L.; Norrman, M.; Schluckebier, G.; Fitch, A.N.; Von Dreele, R.B.; *Acta Crystallogr. D. Biol. Crystallogr.,* **2013**, *69*, 978-990.
13. Noble, S.L.; Johnston, E.; Walton, B.; *Am. Fam. Physician.,* **1998**, *57*, 279-286.
14. Mannucci, E.; Monami, M.; Marchionni, N.; *Diabetes Obes. Metab.,* **2009**, *11,* 53-59.
15. Mason, J.M.; *Future Med. Chem*., **2010**, *2*, 1813-1822.
16. Jaffe, I.A.; Altman, K.; Marryman, P.; *J. Clin. Investig.,* **1964**, *43*, 1869-1873.
17. Zinzalla, G., Thurston, D.E.; *Future Med. Chem.*, **2009**, *1*, 65-93.
18. Gleeson, M.P.; *J. Med. Chem.,* **2008**, *51*, 817-834.
19. Schmid, H.; *Expert Opin. Pharmacother.,* **2013**, *14*, 937-948.
20. Bird, G.H.; Madani, N.; Perry, A.F.; Princiotto, A.M.; Supko, J.G.; He, X.; Gavathiotis, E.;

Sodroski, J.G.; Walensky, L.D.; *Proc. Natl. Acad. Sci. U.S.A.,* **2010**, *107*, 14093-14098.

21. Bray, B.L.; *Nat. Rev.,* **2003**, *2*, 587-593.
22. Chatterjee, J.; Rechenmacher, F.; Kessler, H.; *Angew. Chem. Int. Ed.,* **2013**, *52*, 254-269.
23. Altmann, E.; Altmann, K.H.; Nebel, K.; Mutter, M.; *Int. J. Pept. Protein Res.,* **1988**, *32*, 344-351.
24. Kwon, Y.; Kodadek, T.; *Chem. Biol.,* **2007***, 14*, 671-677.
25. Rezai, T.; Bock, J.E.; Zhou, M.V.; Kalyanaraman, C.; Lokey, S.R.; Jacobson, M.P.; *J. Am. Chem. Soc.,* **2006***, 128*, 14073-14080.
26. Chu, Q.; Moellering, R.E.; Hilinski, G.J.; Kim, Y.; Grossman, T.N.; Yeh, J.; Verdine, G.L.; *Med. Chem. Comm.,* **2015**, *6*, 111-119.
27. Pauletti, G.M.; Okumu, F.W.; Borchardt, R.T.; *Pharm. Res.,* **1997**, *14*, 164-168.
28. Tan, M.C.; Mommaas, A.M.; Drijfhout, J.W.; Jordens, R.; Onderwater, J.J.; Verwoerd, D.; Mulder, A.A.; van der Heiden, A.N.; Scheidegger, D.; Oomen, L.C.; Ottenhoff, T.H.; Tulp, A.; Neefjes, J.J.; Koning, F.; *Eur. J. Immunol.,* **1997***, 27*, 2426-2435.
29. Masungi, C.; Mensch, J.; Van Dijck, A.; Borremans, C.; Willems, B.; Mackie, C.; Noppe, M.; Brewster, M.E.; *Pharmazie*, **2008**, *63*, 194-199.
30. Richard, J.P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B.; Gait, M.J.; Chernomordik, L.V.; Lebleu, B.; *J. Biol. Chem.,* **2003***, 271*, 585-590.
31. Arttamangkul, S.; Alvarez-Maubecin, V.; Thomas, G.; Williams, J. T.; Grandy, D.K.; *Mol. Pharmacol.,* **2000***, 58*, 1570-1580.
32. Erazo-Oliveras, A.; Muthukrishnan, N.; Baker, R.; Wang, T. Y.; Pellois, J. P.; *Pharmaceuticals*, **2012**, *5*, 1177-1209.
33. Gonçalves, M.S.T.; *Chem. Rev.,* **2009**, *109*, 190-212.
34. Gatti, M.; Gioia, M.G.; Andreatta, P.; Pentassuglia, G.; J. *Pharm. Biomed. Anal.,* **2004**, *35*, 339-348.
35. Walker, J.M.; *Methods Mol. Biol.,* **1984**, *1*, 203-212.
36. Unciti-Broceta, A.; Diezmann, F.; Ou-Yang, C.Y.; Fara, M.A.; Bradley, M.; *Bioorg. Med. Chem.,* **2009**, *17*, 959-966.
37. Tyagarajan, K.; Pretzer, E.; Wiktorowicz, J.E.; *Electrophoresis,* **2003**, *24*, 2348-2358.
38. Diamandis, E.P.; Christopoulos, T.K.; *Clin. Chem.,* **1991**, *37*, 625-636.
39. DeChancie, J.; Houk, K.N.; *J. Am. Chem. Soc.,* **2007**, *129*, 5419-5429.
40. Wu, M.M.; Llopis, J.; Adams, S.R.; McCaffery, J.M.; Teter, K.; Kulomaa, M.S.; Machen,

T.E.; Moore, H.P.; Tsien, R.Y.; *Methods Enzymol.,* **2000**, *327*, 546-564.

41. Coley, H.M.; Amos, W.B.; Twentyman, P.R.; Workman, P.; *Br. J. Cancer,* **1993**, *67*, 1316-1323.
42. Li, J.; Newberg, J.Y.; Uhlén, M.; Lundberg, E.; Murphy, R.F.; *PloS one,* **2012**, *7*, e50514.
43. Dube, D.; *Curr. Opin. Chem. Biol.,* **2003**, *7*, 616-625.
44. Hohsaka, T.; Sisido, M.; *Curr. Opin. Chem. Biol.,* **2002**, *6*, 809-815.
45. Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, B.F.; *Proc. Nat. Acad. Sci. U.S.A,* **2004**, *101*, 10000-10005.
46. Sletten, E.; Bertozzi, C.; *Angew. Chem. Int. Ed.,* **2009**, *48*, 6974-6998.
47. Lim, R.K.V; Lin, Q.; *Chem. Commun.,* **2010**, *46*, 1589-1600.
48. Rideout, D.; *Science*, **1986**, *233*, 561-563.
49. Dirksen, A.; Hackeng, T.M.; Dawson, P.E.; *Angew. Chem. Int. Ed.,* **2006***, 45*, 7581-7584.
50. Zeng, Y.; Ramya, T.; Dirksen, A.; *Nat. Methods,* **2009**, *6*, 207-209.
51. Abela, A.R.; Taft, B.R.; Lipshutz, B.H.; *Platin. Met. Rev.,* **2012**, *56*, 62-74.
52. Spicer, C.D.; Davis, B.G.; *Chem. Commun.,* **2013**, *49*, 2747-2749.
53. Köhn, M.; Breinbauer, R.; *Angew. Chem. Int. Ed.,* **2004**, *43*, 3106-3116.
54. Staudinger, H.; Meyer, J.; *Helv. Chim. Acta.,* **1919**, *2*, 635-646.
55. Prescher, J.; Dube, D.; Bertozzi, C.; *Nature,* **2004**, *430*, 873-877.
56. Saxon, E.; Bertozzi, C.R.; *Science*, **2000**, *287*, 2007-2010.
57. Nilsson, B.L.; Kiessling L.L.; Raines, R.T.; *Org. Lett*., **2000**, *2*, 1939-1941.
58. Saxon, E.; Armstrong J.I.; Bertozzi, C.R.; *Org. Lett.,* **2000**, *2*, 2141-2143.
59. De Rosa, L.; Russomanno, A.; Romanelli, A.; D'Andrea, L.D.; *Molecules,* **2013**, *18*, 440-465.
60. Hangauer, M.J.; Bertozzi, C.R.; *Angew. Chem. Int. Ed.,* **2008**, *47*, 2394-2397.
61. Handlon, A.L.; Oppenheimer, N.J.; *Pharm. Res.,* **1988**, *5*, 297-299.
62. Dmitrenko, O.; Thorpe, C.; Bach, R.D.; *J. Org. Chem.,* **2007**, *72*, 8298-8307.
63. Le Droumaguet, C.; Wang, C.; Wang, Q.; *Chem. Soc. Rev.,* **2010**, *39*, 1233-1239.
64. Sawa, M.; Hsu, T.L.; Itoh, T.; Sugiyama, M.; Hanson, S.R.; Vogt, P.K.; Wong, C.H.; *Proc. Nat. Acad. Sci. U.S.A,* **2006**, *103*, 12371-12376.
65. Beatty, K.E.; Xie,F.; Wang, Q.; Tirrell, D.A.; *J. Am. Chem. Soc.,* **2005**, *127*, 14150-14151.
66. Speers, A.E.; Adam, G.C.; Cravatt, B.F.; *J. Am. Chem. Soc.,* **2003**, *125*, 4686-4687.
67. Baskin, J.M.; Prescher, J.A.; Laughlin, S.T.; Agard, N.J.; Chang, P.V.; Miller, I.A.; Lo, A.;

Codelli J.A.; Bertozzi, C.R.; *Proc. Nat. Acad. Sci. U.S.A,* **2007**, *104*, 16793-16797.

68. Figeys, H.P.; Mathy, A.; *Tetrahedron Lett.,* **1981**, *22*, 1393-1396.
69. Sauer, J.; Wiest, H.; *Angew. Chem. Int. Ed.,* **1962**, *1*, 269-269.
70. Thalhammer, F.; Wallfahrer, U.; Sauer, J.; *Tetrahedron Lett.,* **1990**, *31*, 6851-6854.
71. Karver, M.R.; Weissleder, R.; Hilderbrand, S.; *Bioconjugate Chem.,* **2011**, *22*, 2263-2270.
72. Blackman, M.L.; Royzen, M.; Fox, J.M.; *J. Am. Chem. Soc.,* **2008**, *130*, 13518-13519.
73. Yang, J.; Šečkutė, J.; Cole, C.M.; Devaraj, N.K.; *Angew. Chem. Int. Ed.,* **2012**, *51*, 7476-7479.
74. Lim, R.K.V; Lin, Q.; *Acc. Chem. Res.,* **2011**, *44*, 828-839.
75. Clovis, J.S.; Eckell, A.; Huisgen, R.; Sustmann, R.; *Chem. Ber.,* **1967**, *100*, 60-70.
76. Bock, V.D.; Hiemstra, H.; van Maarseveen, J.H.; *Eur. J. Org. Chem.,* **2006**, *1*, 51-68.
77. Boren, B.C.; Narayan, S.; Rasmussen, L.K.; Zhang, L.; Zhao, H.; Lin, Z.; Jia, G.; Fokin, V.V.; *J. Am. Chem. Soc.,* **2008**, *130*, 8923-8930.
78. Kumar, K.; *Int. J. ChemTech Res.,* **2013**, *5*, 3032-3050.
79. Song, W.; Wang, Y.; Qu, J.; Madden, M.M.; Lin, Q.; *Angew. Chem. Int. Ed.,* **2008**, *47*, 2832-2835.
80. Wang, Y.; Song, W.; Hu, W.J.; Lin, Q.; *Angew. Chem. Int. Ed.*, **2009**, *48*, 5330-5333.
81. Aurangzeb, H.; Abbas, A.; Akhtar, M.N.; *Molecules*, **2011**, 16, 7789-7802.
82. Zheng, S.L.; Wang, Y.; Yu, Z.; Lin, Q.; Coppens, P.; *J. Am. Chem. Soc.,* **2009**, *131*, 18036-18037.
83. Ess, D.H.; Houk, K.N.; *J. Am. Chem. Soc.,* **2007**, *129*, 10646-10647.
84. Ess, D.H.; Houk, K.N.; *J. Am. Chem. Soc.,* **2008**, *130*, 10187-10198.
85. Broggini, G.; Orlandi, M.; Molteni, G.; *J. Chem. Soc. Perkin. Trans.,* **2000**, *1*, 3742-3745.
86. Wang, Y.; Vera, C.R.I.; Lin, Q.; *Org. Lett.*, **2007**, *9*, 4155-4158.
87. Yu, Z.; Lim, R.K.V.; Lin, Q.; *Chem. Eur. J.,* **2010**, *16*, 13325-13329.
88. Sustmann, R.; *Pure Appl. Chem.,* **1974**, *40,* 569-593.
89. Caramella, P.; Houk, K.; *J. Am. Chem. Soc.,* **1976**, *98*, 6397-6399.
90. Ezmirly, S.T.; Shawali, A.S.; *J. Heterocyclic Chem.*, **1988**, *25*, 257-263.
91. Houk, K.N.; *Acc. Chem. Res.*, **1975**, *8*, 361-369.
92. Mawhinney, R.C.; Muchall, H.M.; Peslherbe, G.H.; *Can. J. Chem.,* **2005**, *83*, 1615-1625.
93. Sakai, S.; Nguyen, M.T.; *J. Phys. Chem.* A, **2004**, *108*, 9169-9179.
94. Wang, J.; Zhang, W.; Song, W.; Wang, Y.; Yu, Z.; Li, J.; Wu, M.; Wang, L.; Zang, J.; Lin, Q.;

J. Am. Chem. Soc., **2010**, *132*, 14812-14818.

95. Wang, Y.; Lin, Q.; *Org. Lett.,* **2009**, *11*, 3570-3573.
96. Song, W.; Wang, Y.; Qu, J.; Lin, Q.; *J. Am. Chem. Soc.,* **2008**, *130*, 9654-9655.
97. Song, W.; Wang, Y.; Yu, Z.; Vera, C.; *ACS Chem. Bio.,* **2010**, *5*, 875-885.
98. Wang, Y.; Hu, W.; Song, W.; Lim, R.; Lin, Q.; *Org. Lett.,* **2008**, *10*, 3725-3728.
99. Yu, Z.; Ho, L.Y.; Wang, Z.; Lin, Q.; *Bioorg. Med.Chem. Lett.,* **2011**, *21*, 5033-5036.
100. Thaher, B.A.A.; Zahra, J.A.; El-Abadelah, M.M.; Voelter, W. *Z.; Naturforsch*, **2004**, *59b*, 930-933.
101. Compound **89** was synthesised by Dr Robert Bream (GlaxoSmithKline, Stevenage, UK).
102. Hock, S.; Marti, R.; Riedl, R.; Simeunovic, M.; *Chimia,* **2010**, *64*, 200-202.
103. Ben-Ishai, D.; Pinchas, S.; *J. Am. Chem. Soc.*, **1957**, *79*, 4099-4104.
104. Coates, J.; *Encyclopedia of Analytical Chemistry (R.A. Meyers Ed.)*, **2000**, 10815-10837.
105. Meier, H.; Heinzelmann, W.; Heimgartner, H.; *Chimia*, **1980**, *34*, 504-506.
106. Compound **106** was synthesised by Dr Robert Bream (GlaxoSmithKline, Stevenage, UK).
107. Zvilichovsky, G.; Gurvich, V.; *Tetrahedron*, **1997**, *53*, 4457-4468.
108. Rajapakse, H.A.; Walji, A.M.; Moore, K.P.; Zhu, H.; Mitra, A.W.; Gregro, A.R.; Tinney, E.; Burlein, C.; Touch, S.; Paton, B.L.; Caroll, S.S.; DiStefano, D.J.; Lai, MT.; Grobler, J.A.; Sanchez, R.I.; Williams, T.M.; Vacca, J.P.; Nantermet, P.G.; *Chem. Comm. Med. Chem.,* **2011**, *6*, 253-257.
109. Bayardon, J.; Sinou, D.; *Tetrahedron: Asymmetry*, **2005**, *16*, 2965-2972.
110. Reaction conducted by Hannah Dexter (GlaxoSmithKline, Stevenage, UK).
111. Reaction conducted by Dr Robert Bream (GlaxoSmithKline, Stevenage, UK).
112. Analysis conducted by Sean Lynn (GlaxoSmithKline, Stevenage, UK).
113. Butler, R.N.; Hanniffy, J.M.; Stephens, J.C.; Burke, L.A.; *J. Org. Chem.,* **2008**, *73*, 1354-1364.
114. Guru, M.M.; Punniyamurthy, T.; *J. Org. Chem.,* **2012**, *77*, 5063-5073.
115. Butler, R.; O'Shea, D.; *J. Chem. Res.,* **1994**, 350-351.
116. Kidd, P.; *Altern. Med. Rev.* **1997**, *2*, 155-176.
117. Schwöbel, J.A.H.; Wondrousch, D.; Koleva, Y.K.; Madden, J.C.; Cronin, M.T.D.; Schüürmann, G.; *Chem. Res. Tox.,* **2010**, *23*, 1576-1585.
118. Yu, Z.; Pan, Y.; Wang, Z.; Wang, J.; Lin, Q.; *Angew. Chem. Int. Ed.,* **2012**, *51*,

10600-10604.

119. Patterson, D.M.; Nazarova, L.A.; Xie, B.; Kamber, D.N.; Prescher, J.; *J. Am. Chem. Soc.,* **2012**, *134*, 18638-18643.

120. Allen, F.H.; *Tetrahedron*, **1982**, *38*, 645-655.

121. Liao, L.; Zhang, F.; Yan, N.; Golen, J.A.; Fox, J.M.; *Tetrahedron*, **2004**, *60*, 1803-1816.

122. Wong, F.M.; Wang, J.; Hengge, A.C.; Wu, W.; *Org. Lett.,* **2007**, *9*, 1663-1665.

123. Devaraj, N.K.; Houk. K.N.; Seckute, J.; Liang, Y.; Yang, J.; *Chem. Eur. J.*, **2014**, *20*, 3365-3375.

124. Closs, G.; Closs, L.; Boll, W.*; J. Am. Chem. Soc.,* **1963**, *85*, 3796-3800.

125. Huang, L.; Wulff, W.D.; *J. Am. Chem. Soc.,* **2011**, *133*, 8892-8895.

126. Yan, N.; Liu, X.; Pallerla, M.K.; Fox, J.M.; *J. Org. Chem.,* **2008**, *73*, 4283-4286.

127. Structure confirmed by NMR analyst Sean Lynn (GlaxoSmithKline, Stevenage, UK).

128. Unpublished communication (Lin, Q. *et. al*.)

129. Kirihara, M.; Asai, Y.; Ogawa, S.; Noguchi, T.; Hatano, A.; Hirai, Y.; *Synthesis,* **2007**, *21*, 3286-3289.

130. Levkin, P.A.; Trapp, O.; Welle, A.; Yang, C.; Li, L.; Feng, W.; *Angew. Chem. Int. Ed.,* **2015**, *54*, 8732-8735.

131. Koley, D.; Bard, A.J.; *Proc. Nat. Acad. Sci. U.S.A*, **2010**, *107*, 16783-16787.

132. Hadd, M.A.; Nichelson, B.J.; Zhu, Z.; WO 2005/000841A1, **2005**.

133. Hou, D.; Alam, S.; Kuan, T.; Ramanathan, M.; Lin, T.; Hung, M.; *Bioorg. Med. Chem. Lett.,* **2009**, *19*, 1022-1025.

134. Brockunier, L.L.; Parmee, E.R.; Ok, H.O.; Candelore, M.R.; Cascieri, M.A.; Colwell, L.F.; Deng, L.; Feeney, W.P.; Forrest, M.J.; Hom, G.J.; Macintyre, D.E.; Tota, L.; Wyvratt, M.J.; Fisher, M.H.; Weber, A.E.; *Bioorg. Med. Chem. Lett.,* **2000**, *10*, 2111-2114.

135. Yan, W.; Wang, Q.; Lin, Q.; Li, M.; Petersen, J.L.; Shi, X.; *Chem. Eur. J.,* **2011**, *17*, 5011-5018.

136. Maliakal, A.; Lem, G.; Turro, N.J.; Ravichandran, R.; Suhado-lnik, J.C.; DeBellis, A.D.; Wood, M.G.; Lau, J.J.; *Phys. Chem. A.,* **2002**, *106*, 7680-7689; (b) Heller, H.J.; Blattmann, H.R.; *Pure Appl. Chem.,* **1973**, *36*, 141-162.

137. Claus, P.; Doppler, T.; Gakis, N.; Georgarakis, M.; Giezendanner, H.; Gilgen, P.; Heimgartner, H.; Jackson, B.; Märky, M.; Narasimhan, N.S.; Rosenkranz, H.J.; Wunderli, A.; Hansen, H.J.; Schmid, H.; *Pure Appl. Chem.,* **1973**, *33*, 339-362.

138. Liu, Y.; Yan, W.; Chen, Y.; Petersen, J.L.; Shi, X.; *Org. Lett.,* **2008**, *10*, 5389-5392.

139. Ueda, S.; Su, M.; Buchwald, S.L.; *Angew. Chem. Int. Ed.,* **2011**, *50*, 8944-8947.

140. Zhang, Y.; Wang, D.; Wang, W.; Gao, T.; Wang, L.; Li, J.; Huang, G.; Chen, B.; *Synlett,* **2010**, *11*, 1617-1622.

141. Zhang, Y.; Li, X.; Li, J.; Chen, J.; Meng, X.; Zhao, M.; Chen, B.; *Org. Lett.,* **2012**, *14*, 26-29.

142. Liu, X.; Li, X.; Chen, Y.; Wang, D.; Chen, J.; Chen, B.; *Asian J. Org. Chem.,* **2013**, 212-215.

143. Sukumaran, K.B.; Satish, S.; George, M.V.; *Tetrahedron,* **1974**, *30*, 445-450.

144. Li, Y.; Gao, L.X.; Han, F.S.; *Chem. Commun.* **2012**, *48*, 2719-2721.

145. (a) Henderson, R.K.; Jiménez-González, C.; Constable, D.J.C.; Alston, S.R.; Inglis, G.G.; Fisher, G.; Sherwood, J.; Binks, S.P.; Curzons, A.D.; *Green Chem.,* **2011**, *13*, 854-862; (b) Pace, V.; Hoyos, P.; Castoldi, L.; Domínguez de María, P.; Alcántara, A.R.; *ChemSusChem,* **2012**, *5*, 1369-1379; (c) Yilgor, E.; Ekin Atilla, G.; Ekin, A.; Kurt, P.; Yilgor, I.; *Polymer,* **2003**, *44*, 7787-7793.

146. Meier, H.; Heinzelmann, W.; Heimgartner, H.; *Helv. Chim. Acta*, **1985**, *68*, 1283-1300.

147. Böhm, V.P.W.; Herrmann, W.A.; *Angew. Chem. Int. Ed.,* **2000**, *39*, 4036-4038.

148. Wentrup, C.; Benedikt, J.; *J. Org. Chem.,* **1980**, *45*, 1407-1409.

149. Wentrup, C.; Damerius, A.; Reichen, W.; *J. Org. Chem.,* **1978**, *43*, 2037-2041.

150. Watson, T.J.; Ayers, T.A.; Shah, N.; Wenstrup, D.; Webster, M.; Freund, D.; Horgan, S.; Carey, J.P.; *Org. Proc. Res. Dev.,* **2003**, *7*, 521-532.

151. Corey, E.J.; Kim, C.U.; *J. Am. Chem. Soc.,* **1972**, *94*, 7586-7587.

152. Stewart, S.; Harris, R.; Jamieson, C.; *Synlett*, **2014**, *25*, 2480-2484.

153. Verdine, G.L.; Kim, Y.; *Bioorg. Med. Chem. Lett.,* **2009**, *19*, 2533-2536.

154. Lin, Q.; Song, W.; Vera, C.I.R.; Madden, M.; *Chem. Commun.,* **2009**, *37*, 5588-5590.

155. Sagan, S.; Bechara, C.; *FEBS Lett.,* **2013**, *587*, 1693-1702.

156. Hoffmann, T.; Fosgerau, K.; *Drug Discov. Today,* **2015**, *20*, 122-128.

157. Levene, P.A.; *J. Biol. Chem.,* **1905**, *1*, 45-48.

158. Lopez-Otin, C.; Bond, J.S.; *J. Biol. Chem.,* **2008**, *283*, 30433-30437.

159. Burk, B.; *Nat. Review*, **2006**, *5*, 785-799.

160. Jacobsen, J.A.; Jourden, J.L.M.; Miller, M.T.; Cohen, S.M.; *Biochim. Biophys. Acta*, **2010,** *1803*, 72-94.

161. Rothenber, M.L.; Nelson, A.R.; Hande, K.R.; *The Oncologist*, **1998**, *3*, 271-274.

162. Sparano, J.A.; Bernardo, P.; Stephenson, P.; Gradishar, W.J.; Ingle, J.N.; Zucker, S.;

Davidson, N.E. *J.; Clin. Oncol.,* **2004**, *22*, 4683-4690.

163. Hanada, K.; Tamai, M.; Morimoto, S.; Adachi, T.; Ohmura, S.; Sawada, J.; Tanaka, I.; *Agric. Biol. Chem*., **1978**, *42*, 537-541.

164. Matsumoto, K.; Mizoue, K.; Kitamura, K.; Tse, W.; Huber, C.P.; Ishida, T.; *Peptide Science*, **1999**, *51*, 99-107.

165. Gonzalez-Perez, R.R.; Rueda, B.R. (eds); *Tumor Angiogenesis Regulators*, **2013**, CRC Press, 314-315.

166. Siklos,M.; BenAissa, M.; Thatcher, G.R.J.; *Acta Pharm. Sin. B*, **2015**, *5*, 506-519.

167. Kraus, M.; Bader, J.; Overkleeft, H.; Driessen, C.; *Blood Cancer J.*, **2013**, *4*, e103.

168. Cera, E.; *IUBMB Life*, **2009**, *61*, 510-515.

169. Mordwinkin, N.M.; Louie, S.G.; *Expert Opin. Pharmacother.,* **2007**, *8*, 2609-2614.

170. Izquierdo, L.l; Helle, F.; Francois, C.; Castelain, S.; Duverlie, G.; Brochot, E.; *Pharmacogenomics Pers. Med.*, **2014**, *7*, 241-249.

171. Ross, B.S.; Sofia, M.J.; Pamulapati, G.R.; Rachakonda, S.; Zhang, H.R.; Chun, B.K.; Wang, P.; WO 2010/135569, **2010**.

172. Prezas, P.; Arlt, M.J.; Viktorov, P.; Soosaipillai, A.; Holzscheiter, L.; Schmitt, M.; Talieri, M.; Diamandis, E.P.; Krüger, A.; Magdolen V.; *Biol. Chem*., **2006**, *387*, 807-811.

173. Deperthes, D.S.; Clouitier, S.; WO 2004/087912, **2004**.

174. Yousef, G.M.; Diamandis, E.P.; *Atlas Genet. Cytogenet. Oncol. Haematol.,* **2009**, *13*, 357-359.

175. Hovnanian, A.; Dubus, P.; Besson, C.; Robin, A.; Bonnart, C.; Lacroix, M.; Deraison, C.; Briot, A.; *J. Exp. Med.,* **2009**, *206*, 1135-1147.

176. Debela, M.; Beaufort, N.; Magdolen, V.; Schechter, N.M.; Craik, C.S.; Schmitt, M.; Bode, W.; Goettig, P.; *Biol. Chem.,* **2008**, *389*, 623-632.

177. Garcia-Echeverria, C.; Imbach, P.; France, D.; Furst, P.; Lang, M.; Noorani, M.; Scholz, D.; Zimmermann, J.; Furet, P.; *Bioorg. Med. Chem. Lett.,* **2001**, *11*, 1317-1319.

178. Zhu, L.; George, S.; Schmidt, M.F.; Al-Gharabli, S.I.; Rademann, J.; Hilgenfeld, R.; *Antiviral Res.,* **2011**, *92*, 204-212.

179. Lynas, J.F.; Martin, S.L.; Walker, B.; *J. Pharm. Pharmacol.,* **2001**, *53*, 473-480.

180. Liu, Y.; Stoll, V.S.; Richardson, P.L.; Saldivar, A.; Klaus, J.L.; Molla, A.; Kohlbrenner, W.; Kati, M.W.; *Arch. Biochem. Biophys.,* **2004**, *421*, 207-216.

181. Olson, J.E.; Lee, G.K.; Semenov, A.; Rosenthal, P.J.; *Bioorg. Med. Chem.,* **1999**, *7*,

633-638.

182. Gurjar, P.N.; Bercz, A.P.; Nguyen, T.N.; Vadukoot, A.K.; Litosh, V.A.; *Uni. J. Chem.,* **2015**, *3*, 16-26.
183. Baldisserotto, A.; Marastoni, M.; Fiorini, S.; Pretto, L.; Frretti, V.; Gavlioli, R.; Tomatis, R.; *Bioorg. Med. Chem. Lett.,* **2008**, *18*, 1849-1854.
184. Palmer, J.T.; Rasnick, D.; Klaus, J.L.; Bromme, D.*; J. Med. Chem.,* **1995**, *38*, 3193-3196.
185. Yin, Z.; Patel, S.J.; Wang, W.; Wang, G.; Chan, W.; Rao, R.K.R.; Alam, J.; Jeyaraj, D.A.; Ngew, X.; Patel, V.; Beer, D.; Lim, S.P.; Vasudevan, S.G.; Keller, T.H.; *Bioorg. Med. Chem. Lett.,* **2006***, 16*, 36-39.
186. Colgrave, M.L.; Korsinczky, M.J.L.; Clark, R.J.; Foley, F.; Craik, D.J.; *Peptide Science*, **2010**, *94*, 665-672.
187. Sharif, L.; Zhu, Y.; Cowper, B.; Di, W.; Macmillan, D.; *Tetrahedron*, **2014**, *70,* 7675-7680.
188. Debela, M.; Goettif, P.; Magdolen, V.; Huber, R.; Schechter,N.M.; Bode, W.; *J. Mol. Biol.,* **2007**, *373*, 1017-1031.
189. Yoichi, N.; Naoya, O.; Shigeki, M.; Nobuhiro, F.; *J. Nat. Prod*., **1998**, *61*, 667-670.
190. Maryanoff, B.E.; Qiu, X.; Padmanabhan, K.P.; Tulinsky, A.; Almond, H.R.; Andrade-Gordon, P.; Greco, M.N.; Kauffman, J.A.; Nicolaou, K.C.; Liu, A.; Brungs, P.H.; Fusetani, N.; *Proc. Nat. Acad. Sci. U.S.A*, **1993**, *90*, 8048-8052.
191. Molecular Operating Enivronment (MOE) software, MOE, Chemical Computing Group Inc. http://www.chemcomp.com, **2015**.
192. Schaschke, N.; Sommerhoff, C.P.; *ChemMedChem,* **2010**, *5*, 367-370.
193. Janke, D.; Sommerhoff, C.P.; Schaschke, N.; *Bioorg. Med. Chem.,* **2011**, *19*, 7236-7243.
194. Wipf, P.; Kim, H.; *J. Org. Chem.,* **1993**, *58*, 5592-5594.
195. Wipf, P.; Kim, H.; *Tetrahedron Lett.,* **1992**, *33*, 4275-4278.
196. Maryanoff, B.E.; Zhang, H.; Andrade-Gordon, P.; Greco, M.N.; Kauffman, J.A.; Nicolaou, K.C.; Liu, A.; Brungs, P.H.; *J. Am. Chem. Soc.,* **1995**, *117*, 1225-1239.
197. Deng, J.; Hamada, Y.; Shioiri, T.; *Tetrahedron Lett.,* **1996**, *37*, 2261-2264.
198. Roche, S.P.; Faure, S.; Aitken, D.J.; *Angew. Chem. Int. Ed.,* **2008**, *47*, *6840*-6842.
199. Hagihara, M.; Schreiber, S.L.; *J. Am. Chem. Soc.,* **1992**, *114*, 6570-6571.
200. Bastiaans, H.M.M.; van der Baan, J.L.; Ottenheijm, H.C.J.; *Tetrahedron Lett.,* **1995**, *36*, 5963-5966.
201. Faure, S.; Hjelmgaard, T.; Roche, S.P.; Aitken, D.J.; *Org. Lett.,* **2009**, *11*, 1167-1170.

202. Banfi, L.; Basso, A.; Guanti, G.; Riva, R.; *Mol. Divers.,* **2003**, *6*, 227-235.

203. Wasserman, H.H.; Zhang, R.; *Tetrahedron,* **2002**, *58*, 6277-6283.

204. Roche, S.P.; Faure, S.; Blidi, L.E.; Aitken, D.J.; *Eur. J. Org. Chem.,* **2008**, *30*, 5067-5078.

205. Mandal, P.K.; Ren, Z.; Chen, X.; Xiong, C.; McMurray, J.S.; *J. Med. Chem.,* **2009**, *52*, 6126-6141.

206. Adamson, J.; Blaskovich, M.A.; Groenevelt, H.; Lajoie, G.A.; *J. Am Chem. Soc.,* **1991**, 3447-3449.

207. Capicciotti, C.J.; Trant, J.F.; Leclère, M.; Ben, R.N.; *Bioconjugate Chem.*, **2011**, *22*, 605-616.

208. Jackman, L.M.; Sternhall, S. (eds.); *Application of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry*; Pergamon Press: Oxford, **1969**, 301-302.

209. Bey, P.; Gerhart, F.; Jung, M.; *J. Org. Chem.,* **1986**, *51*, 2835-2838.

210. Djebbar, S.; Gabouze, N.; Ozanam, F.; Moraillon, A.; Henry de Villeneuve, C.; Gouget-Laemmel, A.C.; Chazalviel, J.N.; Allongue, P.; Andresa, J.S.; Touahir, L.; Sam. S.; *Langmuir*, **2010**, *26*, 809-814.

211. Carpino, L.A.; *J. Am. Chem. Soc.,* **1993**, *115*, 4397-4398.

212. Wasserman, H.H.; Ho, W.; *J. Org. Chem.,* **1994**, *59*, 4364-4366.

213. Vojkovsky, T.; *Pept. Res.,* **1995**, *8*, 236-237.

214. Christensen, C.; Meldal, M.; *Peptides: The Wave of the Future,* American Peptide Society*,* **2001***,* 269-270.

215. Hasegawa, K.; Kawakami, T.; Aimoto, S.; Akaji, K.; Bang, J.K.; *Tetrahedron Lett.,* **2004**, *45*, 99-102.

216. Yu, Z.; Caldera, P.; McPhee, F.; De Voss, J.J.; Jones, P.R.; Burlingame, A.L.; Kuntz, I.D.; Craik, C.S.; Ortiz de Montellano, P.R.; *J. Am. Chem. Soc.,* **1996**, *118*, 5846-5856.

217. Danion, D.; Danion-Bougot, R.; Bauchat, P, Collet, S.; *Tetrahedron: Asymmetry*, **1998**, *9*, 2121-2131.

218. Scherkenbeck, J.; Van, C.T.; *Bioorg. Med. Chem.,* **2015**, *23*, 3278-3286.

219. Zhang, J.; Huang, G-B.; Weng, J.; Lu, G.; Chan, A.S.C.; *Org. Biomol. Chem.,* **2015**, *13*, 2055-2063.

220. Maegawa, T.; Nakamura, A.; Miki, Y.; Umemoto, H.; Onaka, T.; *J. Org. Chem.,* **2014**, *79*, 6703-6707.

221. Kakehi, A.; Tanaka, Y.; Iro, S.; *Bull. Chem. Soc. Jpn,* **1976**, *49*, 762-766.

222. Efimova, Y.A.; Artamonova,T.V.; Koldobskii, G.I.; *Russ. J. Org. Chem.,* **2010**, *46*, 612-614.

223. Ito, S.; Tanaka, Y.; Kakehi, A.; Kondo, K.; *Bull. Chem. Soc. Jpn,* **1976**, *49*, 1920-1923.

224. Scartoni, V.; Martini, C.; Lucacchini, A.; Livi, O.; Biaga, G.; *J. Pharma. Sci.,* **1992**, *81*, 543-545.

225. Burke, M.J.; Trantow, B.M.; *Tetrahedron Lett.,* **2008**, *49*, 4579-4581.

226. Lundgren, R.J.; Stradiotto, M.; *Angew. Chem. Int. Ed.,* **2010**, *49*, 8686-8690.

227. Wang, G.; Huang, T.; Kuang, Y.; Lin, L.; Feng, X.; *Org. Lett.,* **2013**, *15*, 76-79.

228. Iera, J.A.; Jenkins, L.M.M.; Kajiyama, H.; Kopp, J.B.; Appella, D.H.; *Bioorg. Med. Chem. Lett.,* **2010**, *20*, 6500-6503.