THE UNIVERSITY
of EDINBURGH

Dyes, Linkers, Tags and Libraries – New Tools for Systems Chemical Biology

Thesis Submitted in Accordance with the Requirement of The University of Edinburgh for the Degree of Doctor of Philosophy

By

Gemma Mudd

MChem (Hons)

Supervised by

Prof. Manfred Auer

School of Biological Sciences

College of Science and Engineering

Dec 2014

Declaration

I hereby declare that, except where specific reference is made to other sources, the work contained within this thesis is the original work of my own research since the registration of the PhD degree in January 2011, and any collaboration is clearly indicated. This thesis has been composed by myself and has not been submitted, in whole or part, for any other degree, diploma or other qualification.

Signed

(Gemma Mudd)

Date

08/03/2015

Acknowledgements

First and foremost, I would like to thank Professor Manfred Auer for giving me this opportunity and his continued support and encouragement.

I would like to acknowledge our collaborators who carried out work on the GPCR54 homology model, Agnieszka Bronowska, Xavier Deupi and Steven Shave. Adam Belsom provided support and analysis for the cross-linking chapter of this thesis. Megan Gallant carried out biological evaluation of compounds. Rachel Milne provided assistance in the production of nanodiscs. Thanks also to my second supervisor, Professor Bob Millar for his support and advice.

I would also like to thank all members of the Auer lab, past and present, for providing such a lovely environment to work in. This process would have been a lot harder if it wasn't for all of you! In particular, I'd like to thank Angus and Christophe for guiding me through my first few years and being such inspirational people and chemists. Joanna, your teaching and support in the biological aspects of this project are greatly appreciated. Geordie, thanks for the many LOLs we had in the lab and for always having the time to listen to and help fix my chemistry problems. Irene, thank you for your help in the rhodamines project and good luck for the rest of your PhD – I know you'll do great! To Dave and Olivier, thank you for your support and help throughout this process. I would also like to express my gratitude to Nhan, for help with the imaging and biophysical aspects of this work and to Steve for computational support.

I would like to thank my family and the Fallan clan for all of their support and encouragement. To John and Juls, you guys are awesome; I can't wait to spend more time with you now. To my Mum and Dad, thank you for everything you have done for me, I am extremely lucky to have you. Also Dad - I hope this goes some small way to make up for the fact that I did not take A-level maths. To Charlene - as promised you helped me through every aspect of this journey. Thank you so much - I am so excited for whatever comes next for us! Finally, I would like to dedicate this thesis to the memory of two wonderful ladies, Jean and Judy.

List of Abbreviations

ACHN	1,1'-Azobis(cyclohexylnitrile)
AIDA	An indazole based drug discovery and analysis tool
ALBA	Acid labile benzofurnanone linker
Alloc	Allyloxycarbonyl
Ar	Aryl
Boc	*tert*-Butoxycarbonyl
br	Broad
Bzl	Benzyl
CONA	Confocal nanoscanning
CPP	Central precocious puberty
Cy5	Cyanine5 dye
CyHex	Cyclohexane
d	Doublet
Da	Daltons
DBCO	Dibenzocyclooctyl
DCC	*N*,*N*′-Dicyclohexylcarbodiimide
Dde	1(4,4dimethyl2,6 dioxocyclohexylidene)ethyl
DIC	*N*,*N*′-Diisopropylcarbodiimide
DIPEA	*N*,*N*-Diisopropylethylamine
DMAP	*N*,*N*-Dimethyl-4-aminopyridine
DMF	*N*,*N*-Dimethylformamide

DMSO	Dimethylsulfoxide
DOA	*N*-Fmoc-8-amino-3,6-dioxaoctanoic acid
EDC.HCl	*N*-(3-Dimethylaminopropyl)-*N′*-ethylcarbodiimide hydrochloride
ESI-MS	Electron spray ionisation – mass spectrometry
EtOAc	Ethyl acetate
EtOH	Ethanol
FACS	Fluorescence activated cell sorting
FCS	Fluorescence correlation spectroscopy
FIDA	Fluorescence intensity distribution analysis
FLAG	Epitope tag of motif DYKXXD
Fmoc	9-Fluorenylmethyloxycarbonyl
FSH	Follicle-stimulating hormone
g	Grams
GFP	Green fluorescent protein
GnRH	Gonadotropin-releasing hormone
GPCR/GPR	G-protein coupled receptor
h	Hour(s)
HA	Hypothalamic amenorrhea
HATU	*O*-(7-Azabenzotriazol-1-yl)-*N*,*N*,*N′*,*N′*-tetramethyluronium hexafluorophosphate
HBTU	*O*-(Benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate
HEK293	Human embryonic kidney 293 cells

HMBA	Hydroxymethylbenzoic acid
HPLC	High performance liquid chromatography
HPG	Hypothalamic–pituitary–gonadal axis
HTS	High throughput screen
IHH	Idiopathic hypogonadotropic hypogonadism
IVF	In vitro fertilisation
IP_3	Inositol trisphosphate
KP	Kisspeptin
LC-MS	Coupled liquid chromatography - mass spectrometry
LH	Leutinising hormone
M	Molarity
m	Multiplet
MALDI	Matrix assisted laser desorption ionization
MeCN	Acetonitrile
MeOH	Methanol
mg	Milligram
MHz	Megahertz
min	Minute(s)
mL	Millilitre
mmol	Millimole
MMP-9	Matrix metallopeptidase 9
MS	Mass spectrometry

m/z	Mass to charge ratio
NBS	*N*-Bromosuccinimide
NHS	*N*-Hydroxysuccinimide
NHSS	*N*-hydroxysulfosuccinimide
nm	Nanometer(s)
NMR	Nuclear magnetic resonance spectroscopy
NTA	Nitrilotriacetic acid
OBOC	One bead one compound
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
ppm	Parts per million
Pra	Propargylglycine
PS	Polystyrene
PS/PS	Post synthesis/post screening
q	Quartet
QSAR	Quantitative structure activity relationship
R_f	Retention factor
RG	Rhodamine green
R_t	Retention time
rt	Room temperature
s	Singlet
SAR	Structure activity relationship

SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
t	Triplet
TBAF	Tetrabutylammonium fluoride
^{t}Bu	*tert*-Butyl
TAMRA	Carboxytetramethylrhodamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TLC	Thin layer chromatography
TMR	Tetramethylrhodamine
TOBa	Trifunctional orthogonal backbone resin
TOBOC	Tagged one bead one compound
UV	Ultraviolet
WT	Wild type

Abstract

Chemical biology can be defined as the area of science where chemical tools are used to study biological systems. The simplest way this can be achieved is in the identification of compounds which inhibit or modulate a biological pathway and the consequences studied. However, novel tools are required to enable, for example, the development of assays to allow simpler screening of difficult targets such as membrane proteins and protein-protein interactions.

A series of kisspeptin analogues were synthesised for the development of a screening platform compatible with G-protein coupled receptors and tagged one bead one compound (OBOC) combinatorial libraries. Fluorescently labelled kisspeptin showed good affinity for GPCR54 and an on-bead version of the peptide, with the required C-terminal amide presented away from the bead was prepared and used for testing possible screening methods. GPCR54 was expressed in a number of formats and a kisspeptin based OBOC library designed and synthesised. Investigation into the C-terminal RF-amide motif of Kisspeptin was also carried out in order to assess the importance of the carbonyl moiety. The corresponding peptide amine was synthesised and the compound biologically assessed. This led to the development of a novel acid labile benzofuranone (ALBA) linker for anchoring amines to a solid support.

For the preparation of fluorescent kisspeptin ligands, a novel general synthetic route which gives direct access to single isomer functionalised rhodamine dyes from phthalides has been developed. This circumvents the arduous task of isomer separation usually associated with the synthesis of functionalised rhodamines. The route has been demonstrated with a range of linkage groups and rhodamine types.

This rhodamine material was used as a reporter group in various multifunctional reagents synthesised using a trifunctional orthogonally protected backbone (TOBa), which was prepared on a solid support and enables rapid synthesis of trifunctional reagents. This resin takes advantage of protecting group orthogonality and the high yields of peptide bond formation. A series of trifunctional reagents for screening use were prepared using this resin. A proof of concept study was carried out involving the simultaneous labelling and immobilisation of a protein for applications in probing protein-protein interactions.

Development of a trifunctional hydroxamic acid containing cross-linker was carried out which takes advantage of its reaction with boronic acids to enable reversible capture on solid

support for enrichment of cross-linked peptides. A new benzophenone based heterobifunctional reagent was prepared for protein cross-linking and mass spectrometry analysis. This was shown to give complimentary reactivity to existing cross-linkers, allowing more structural information to be extracted from protein samples.

Contents

DECLARATION ..... **I**

ACKNOWLEDGEMENTS ..... **II**

LIST OF ABBREVIATIONS ..... **III**

ABSTRACT ..... **VIII**

CONTENTS ..... **X**

1 HISTORY AND OVERVIEW OF THE AUER LAB SCREENING PROCESS **1**

1.1 ON-BEAD SCREENING AND OBOC LIBRARIES 2

1.2 CONFOCAL NANOSCANNING (CONA) 4

1.3 PICKOSCREEN-02 (PS02) 6

1.4 AIDA TECHNOLOGY 7

1.5 POST SYNTHESIS/POST SCREENING (PS/PS) LABELLING 10

1.6 DYES, LINKERS, TAGS AND LIBRARIES – NEW TOOLS FOR SYSTEMS CHEMICAL BIOLOGY 14

2 TOWARDS A HIGH THROUGHPUT SCREENING PLATFORM FOR GPCR54 USING TAGGED ONE BEAD ONE COMPOUND LIBRARIES **16**

2.1 ABSTRACT 17

2.2 INTRODUCTION 17

2.3 AIMS 36

2.4 RESULTS AND DISCUSSION 36

2.5 CONCLUSIONS AND FUTURE WORK 67

3 INVESTIGATION OF KISSPEPTIN *C*-TERMINUS & DEVELOPMENT OF THE ALBA LINKER **69**

3.1 ABSTRACT 70

3.2 INTRODUCTION 70

3.3 AIMS 82

3.4 RESULTS AND DISCUSSION 82

3.5 CONCLUSIONS AND FUTURE WORK 92

4 DEVELOPMENT OF A NOVEL SYNTHETIC ROUTE TO SINGLE ISOMER FUNCTIONALISED RHODAMINE DYES **95**

4.1 ABSTRACT 96

4.2 Introduction 96
4.3 Aims 106
4.4 Results and Discussion 106
4.5 Conclusions and Future Work 117

5 TOBA RESIN FOR GENERATION OF TRIFUNCTIONAL REAGENTS 120

5.1 Abstract 121
5.2 Introduction 121
5.3 Aims 138
5.4 Results and Discussion 138
5.5 Conclusions and Future Work 154

6 NOVEL PROTEIN CROSS-LINKING REAGENTS 157

6.1 Abstract 158
6.2 Introduction 158
6.3 Aims 182
6.4 Results and Discussion 183
6.5 Conclusions and Further Work 204

7 CONCLUDING DISCUSSION 205

8 EXPERIMENTAL 210

8.1 General Information 211
8.2 Chapter 2 Experimental 213
8.3 Chapter 3 Experimental 230
8.4 Chapter 4 Experimental 240
8.5 Chapter 5 Experimental 247
8.6 Chapter 6 Experimental 261

APPENDIX 292

Curriculum Vitae 292

CHAPTER 1

HISTORY AND OVERVIEW OF THE AUER LAB SCREENING PROCESS

1.1 On-Bead Screening and OBOC Libraries

Solid phase synthesis was first reported by Merrifield in 1963.[1] In this seminal report, an amino acid was covalently linked to a solid resin support *via* the *C*-terminus. *N*-Terminally protected amino acids were then sequentially added in order to build up polypeptide chains, discarding excess reagents *via* filtration. Liberation of the compound from the solid support yielded the free peptide in solution. This technique allowed the convenient synthesis of individual peptides up to high molecular weights, previously unattainable by solution based methods. The importance of this development in the field of peptide chemistry was recognised in 1984, with Merrifield receiving the Nobel Prize for 'his development of methodology for chemical synthesis on a solid matrix'. In the mid-80s this technique was developed to generate peptide libraries, although only small libraries of pre-determined sequence were achievable.[2, 3]

In 1991, Lam *et al.*[4] and Furka *et al*.[5] almost simultaneously developed a peptide library synthesis method named 'one-bead, one-peptide' or 'one bead, one compound' (OBOC). Here, a split synthesis approach was taken, whereby a pool of resin was split into several reaction vessels and into each one was added a different single amino acid (Figure 1-1). The resin was then re-pooled and mixed, then split again. This cycle was repeated until the full length of peptide required was built up, allowing an extraordinary number of compounds to be generated in a short space of time. Lam *et al*. used this technique to synthesise a pentamer library using 19 of the natural amino acids (all except cysteine), which can statistically result in 2,476,099 (19^5) individual compounds within a few days.

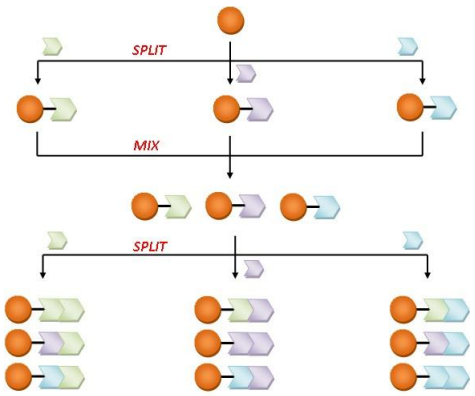


Figure 1-1: Split and mix synthesis leading to a one-bead one-compound library. Here, 3 building blocks are used over to generate a dimer which gives $3^2 = 9$ compounds. If a pentamer were to be made, this would give $3^5 = 243$ theoretical compounds.

In the same report, the concept of screening compound libraries against targets directly on bead was introduced. This involved fluorescently labelling target 'acceptor molecules' with fluorescein, and incubating this material with the prepared one-bead one-compound libraries. Beads presenting compounds which interacted strongly with the fluorescent target were identified either by the naked eye or by using a low-power dissecting microscope (Figure 1-2). These beads were isolated by hand using forceps, the fluorescent target washed off and the peptidic compound identified using a microsequencer. This combined approach presented a way to greatly enhance the speed at which compound libraries with millions of members could be produced and rapidly screened in a miniaturised format, with only 50 – 200 pmol of each compound required.

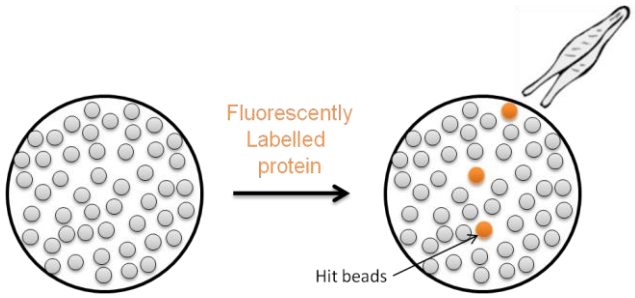


Figure 1-2: On-bead screening and hit bead retrieval of one-bead one-compound library according to Lam *et al.*

Within Novartis, in the Innovative Screening Technologies Unit (IST) headed by Manfred Auer, this basic concept was considered to have great potential as an alternative to vast screening archives that are normally employed within large pharmaceutical companies.

The large compound archives built up and run by pharmaceutical companies are associated with very high costs. It has been reported that a typical high throughput screen (HTS) of 1 million compounds can cost anywhere between $500 000 to $1 000 000.[6] This number does not even include the costs associated with the synthesis of compounds, storage in large archive buildings or maintenance of the archive.

Therefore the OBOC library synthesis and on-bead screening approach, which presented the possibility to generate millions of compounds in a small time scale, with no purification required and very low reagent costs, was highly appealing. The miniaturised format of the screens also means that an entire archive of billions of compounds could be stored in single freezer with no maintenance costs involved. A comparison between solid supported and solution phase libraries is shown in Table 1-1.

	On-bead Library	**Conventional Library for Solution Screening**
80 000 Compounds	€ 10 000	€ 3 500 000
Time	2 – 4 weeks	>> 4 weeks

Table 1-1: Comparison of time and cost to generate libraries in solution, or in a combinatorial fashion on solid support.[7]

However, in order to develop the idea of on-bead screening into a reliable and quantitative high-throughput screening technique which would be considered adequate in the pharmaceutical industry, significant improvements to the workflow were required. An overview of the contributions made by the Auer group to this field of research is detailed *vide infra*.

1.2 Confocal Nanoscanning (CONA)

The first on-bead screens used standard fluorescence microscopes which only allow imaging at the entire bead volume. However, some resins and compounds with extended π-systems can cause autofluorescence. This can affect the detection of binding events of fluorescently labelled target proteins, which can be masked by fluorescent background intensity.

Correct choice of bead material that exhibits minimal autofluorescence can reduce problems associated with hit detection. PEG based resins exhibit very little autofluorescence and therefore a resin based on this material, TentaGel, was selected (Figure 1-3). TentaGel beads are made of a divinyl benzene cross-linked polystyrene core with long PEG grafts (~75 units).[8] The PEG units represent the majority of the mass of the resin, therefore the properties of the resin closely resembles that of PEG. As well as limiting autofluorescence,

this feature also allows its use with organic solvents and aqueous buffers, which is required for screening in the presence of proteins. TentaGel resin also has a high mechanical stability making it an ideal choice for library production.

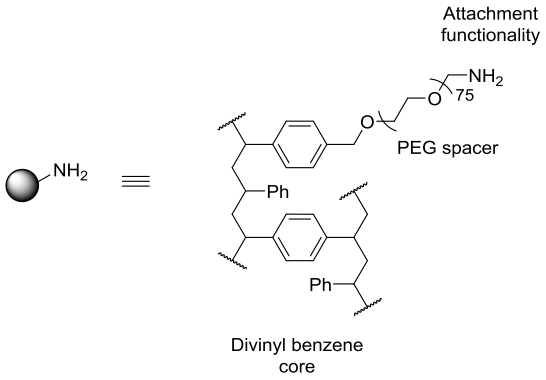


Figure 1-3: TentaGel resin is made of a divinyl benzene cross-linked polystyrene core with long PEG grafts

Binding of fluorescent biomolecules is generally restricted to the outer few micrometres of a bead, as proteins are too large to diffuse into the resin matrix in the time frame of an on-bead screen. Therefore, Confocal Nanoscanning (CONA) was developed as a method to focus only on the outer 5 µm of library beads, where binding events take place (Figure 1-4). In a typical screen, library resin is placed in wells of a 96 well microtiter plate (~1 mg/2000 beads) to form a monolayer and the beads are incubated with 1 – 500 nM, typically 20 – 50 nM, fluorescently labelled protein for a few hours. Next, by moving the stage, the confocal laser focus is used to scan the whole well just below the beads equatorial plane (~40 nm high), generating a cross sectional image of the beads. In these images, hit beads appear as 'rings', where fluorescent protein is bound to the outer surface layer of the bead.[9]

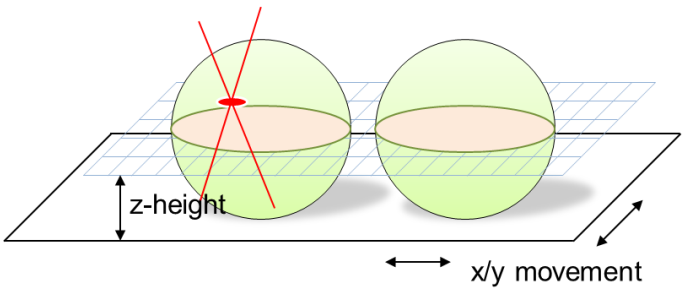


Figure 1-4: CONA scanning of resin monolayer. Laser beam scans equatorial plane of beads (~ 40 µm high for 90 µm TentaGel resin). This gives cross section of beads: top right - beads with no fluorescent material on surface; bottom right - beads with fluorescent material on surface displaying 'fluorescent rings'.

From the images of wells, the BeadEval software (Perkin Elmer) recognises each individual bead and generates x/y coordinates for their positions. Two parameters are then measured: the average fluorescence intensity of the bead ring interior, and the mean intensity on the surface of the bead. Hit beads have a much higher fluorescence intensity on the surface of the bead than the interior, whereas autofluorescent beads appear bright throughout (Figure 1-5). The relative fluorescent ring intensity was found to be proportional to the amount of fluorescent protein bound to the resin.[10]

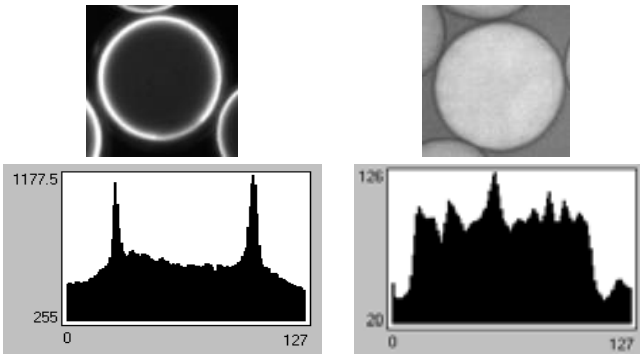


Figure 1-5: Example intensity profiles for a hit bead (left) and an autofluorescent bead (right) after CONA screening[7]

1.3 PickoScreen-02 (PS02)

The PS02 is a confocal scanning and bead picking instrument developed in the Auer group in collaboration with Evotec Technologies (now Perkin Elmer, Germany). It consists of a confocal microscope for CONA screening, along with a motorised scanning compartment with sample holder. This allows the sample plate to move so that the relatively large area of a whole well can be scanned, therefore eliminating distortion of images which would be encountered should the laser move instead of the plate. In addition to this, the instrument contains an automated bead-picker, which consists of a robotic arm controlled by a

computer. Attached to this is a hydraulic-driven polyimide capillary for isolating selected beads. The PS02 instrument represented the first example of an automated platform for on-bead screening of one-bead one-compound combinatorial libraries.[11]

1.4 AIDA Technology

Although CONA screening can identify and, to a point, quantify affinity of a resin bound compound for a target protein, this process can only be used as a preliminary screen. This is due to several reasons. Firstly, the concentration of ligand is very high at the bead surface. The loading capacity of TentaGel resin is around 0.3 mmol/g, which corresponds to a ligand concentration of around 100 mM.[12] This means that even low affinity binders can appear as hits. The affinity may be so low that binding may not be detectable in solution based assays. Secondly, the localisation of a fluorescent target around a bead may not be due to molecular recognition of the bead bound ligand, but rather factors such as unspecific binding, aggregation or ion exchange effects.[9]

Therefore it was clear that although on-bead CONA screening delivered a way to screen a vast number of compounds rapidly and can eliminate many inactive compounds, a secondary in-solution assay was required in order to confirm identified potential hits before any resynthesis was started to generate enough compound for follow up assays. Such resynthesis work can bind up a lot of chemical capacity and must be focused on the most promising compounds only. As a vast range of ligands and targets were intended to be screened, a universal assay was required which would allow solution based assessment of the target binding affinity to any of them. The generation of fluorescently labelled on-bead libraries presented an ideal solution, as following resynthesis, hit library members could be taken into fluorescence based assays using fluorescence anisotropy or fluorescence resonance energy transfer (FRET) as a detection method. For this purpose, so called AIDA (an indazole based drug discovery and analysis tool) UV dyes were developed (Figure 1-6).[13] These contained two linkage functionalities so that they could be easily incorporated into every compound at the start of synthesis.

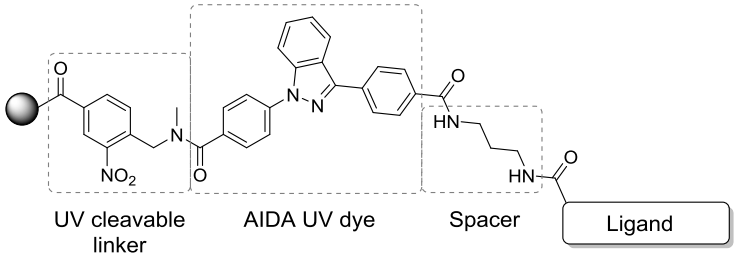


Figure 1-6: Setup of on-bead libraries using AIDA dyes

AIDA dyes allowed the generation of libraries based on a range of peptidic and small molecule scaffolds, due to their chemical stability. Their spectral properties (excitation at 350 nm) allowed the use of more red shifted dyes (emission of 488 nm to 633 nm) to be used to label biomolecular targets (proteins, DNA, RNA) for CONA screening, wavelengths at which AIDA is practically undetected. The dye also acted as a decoding tool, showing predicted fragmentation patterns in mass spectrometry analysis. It also provided a chromophore for HPLC based quality control, purification and affinity determination in solution.

In a typical AIDA based screen, beads identified with a strong fluorescent ring from CONA were cleaved from the resin *via* a UV sensitive linker and decoding of the compounds carried out by MS. Hit compounds were then re-synthesised and taken into solution based assays to determine quantitative binding data against unlabelled target proteins *via* AIDA fluorescence intensity or anisotropy changes using an ensemble averaging fluorometer or plate reader.

The AIDA process allowed successful on-bead to solution phase screens to be carried out, with hit compounds found against more than 60 biological targets, including a small molecule inhibitor of importin β mediated nuclear import.[14] Despite this success, there were some drawbacks to the AIDA approach which required improvements to be made in the screening approach. The UV linker used to tether compounds to the solid support proved unsuitable for long term storage of compounds. The planar nature of the dye also caused aggregation effects, which interfered with protein binding experiments. Also, the hydrophobic nature of the AIDA dye compound led to resynthesised compounds being very poorly soluble in buffers. Following screening of a library of various scaffolds against Hu protein R (HuR), resynthesised hits were not soluble at physiological conditions, making the usual in-solution affinity confirmation experiments impossible.[15] In this case, competition assays were performed where labelled protein was dissociated from the most intensely

fluorescent bead by unlabelled hits. This screen led to the identification of the first RRM3 (RNA recognition motif 3) targeted low molecular weight HuR inhibitors and allowed the discovery of an ATP-binding pocket within HuR RRM3, associated with enzymatic activity.

Another drawback of this approach was that the large AIDA dye on each compound presented an entity with which target biomolecules could interact. Indeed, following the screening of a diazepanone library against fluorescently labelled integrin leucocyte function-associated antigen 1 I domain (LFA-1 ID), it was found that resulting hits beads were not due to the synthesised library compounds, but to the AIDA dye itself This serendipitous finding led to IBE667, shown in Figure 1-7, the first ever reported LFA-1 activator.[16]

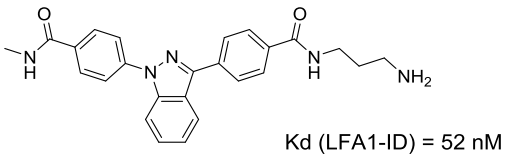


Figure 1-7: ICAM-1 binding enhancer 667 (IBE-667)

As this compound was found to be an activator of LFA-1-ICAM-1 binding, it was named ICAM-1 binding enhancer 667 (IBE-667). Its interaction with LFA-1 ID has a K_d of 52 nM and the affinity of the LFA-1 ID/ IBE-667 complex with ICAM-1 was found to be 76 nM (K_d), whereas without the small molecule activator, LFA-1 binds to ICAM-1 with an K_d of 1.5 mM.

In order to prove solution binding of on-bead hit compounds after cleavage using the amount released from an individual bead, highly parallelised and miniaturised single molecule spectroscopic methods could be employed. The AIDA dyes suffer from short excitation wavelength, low brightness and low photo stability, properties which prevent their use in such assays. Therefore, each hit from a CONA screen was required to be re-synthesised in higher amounts before confirmation of the hit could be carried out in solution. Without single-bead based solution confirmation, on-bead screening often results in hit compounds with too low affinity to target proteins in solution, or too low cellular activity. In the pharmaceutical industry, such discrepancies between a surface K_d and a solution K_d might be referred to as artefacts, or false positives, while in fact they are a thermodynamic reality usually caused by differences in binding enthalpy. With different isoelectric points of proteins and different charge states of ligands, these affinity differences are unavoidable. This means that resources are wasted in the re-synthesis of compounds which turned out to show too low solution binding affinity to the target protein.

Therefore, a second generation screening method was developed by the Auer lab with the aim to combat these problems. The goal was to develop a method which allows sorting out compounds based on solution binding, rather than on-bead binding affinity without any re-synthesis needs, and to include purity check and structure decoding from the same single bead batch of compound.

1.5 Post Synthesis/Post Screening (PS/PS) Labelling

In 2004, the development of a second generation of an on-bead to solution screening process began. Here, rather than incorporating a dye from the beginning, fluorescent labelling was carried out after library synthesis and CONA screening, exclusively to isolated, picked or sorted hit beads. It was therefore named post synthesis/post screening labelling.

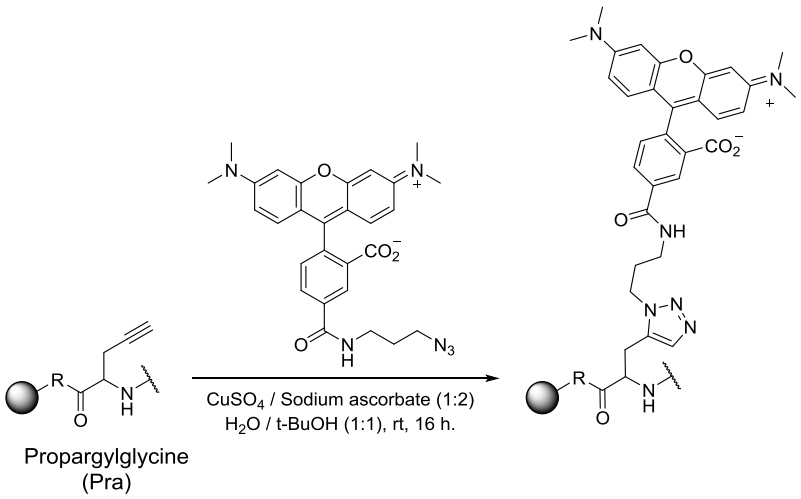
This second generation approach has a few inherit benefits compared to the AIDA technology. Firstly, a very small 'tagging group' can be incorporated into each compound for attachment of a dye following screening, thus reducing the risk of non-specific interaction with the target proteins compared with the large AIDA dye. Secondly, the dye is introduced following synthesis, and therefore is not required to withstand synthetic conditions, allowing more sensitive fluorophores to be used. Finally, the choice of dye is limitless, allowing access to single molecule confocal spectroscopy for confirmation of binding in solution, using techniques such as fluorescence correlation spectroscopy (FCS) or fluctuation intensity distribution analysis (FIDA/2D-FIDA). These experiments require very little material (< 1 pmol), hence the small amount of compound present on a single bead (~ 40 – 100 pmol) is sufficient for direct use in solution based hit confirmation. Therefore re-synthesis decisions can be based on K_d data, eliminating the risk of exerting efforts and resources on low affinity compounds.[7] It also allows sensitive detection of compounds using fluorescence based RP-HPLC for qc of compounds and aids in structure determination by generating compounds with masses high enough to exceed that of MALDI matrix signals.

The reaction to be chosen for PS/PS labelling was required to fulfil the following conditions:

- Bioorthogonal reactivity so as not to interfere with deprotected library compounds
- Stability of reactive partners to withstand conditions of synthesis
- Small reactive partners, to minimise interaction of tagging group with the protein target during CONA screening

- Efficient, to ensure the transformation of compound to fluorescent ligand is high yielding

The copper catalysed Huisgen 1,3-dipolar addition 'click' reaction between terminal alkynes and azides to form a triazole linkage was chosen (Scheme 1-1).[17-20] To label using this chemistry, a terminal alkyne is required to be incorporated into library compounds. This was done simply by using alkyne functionalised amino acid propargylglycine (Pra). This discreet alkyne group is stable to many chemical conditions and allows labelling of the compounds *via* an azide functionalised dye.[9] Tetramethylrhodamine was used, as it has good fluorescence properties and is relatively stable, and so can withstand linker cleavage conditions.[21]



Scheme 1-1: Click chemistry employed for PS/PS labelling. A terminal alkyne bearing amino acid is incorporated into library compounds. Following screening, copper catalysed 1,3-dipolar cycloaddition allows labelling using an azide functionalised dye such as tetramethylrhodamine azide (TMR-N_3).

The PS/PS labelling approach also required a linker which would be stable to synthesis conditions and side chain removal, but allow cleavage of compounds from the resin in a fashion that would not damage dye compounds. HMBA linker (4-hydroxymethylbenzoic acid) satisfies these requirements and allows linkage of compounds to the bead *via* an ester, which is stable to acid mediated side chain deprotection and allows cleavage of compounds using a variety of nucleophiles to give various *C*-terminally modified peptides (Figure 1-8).[22]

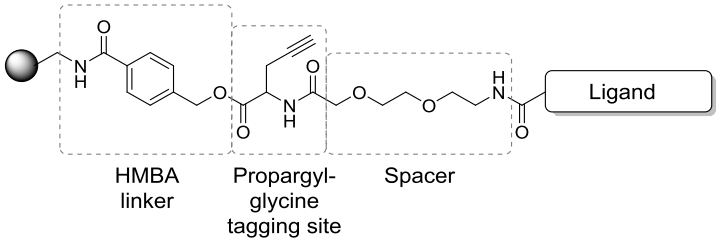


Figure 1-8: Library compound setup for PS/PS labelling

This setup proved successful, with an average of around 60 pmol of clean fluorescent material recovered from single beads before and following CONA, indicating that material was not lost during on-bead screening (Figure 1-9).[7]

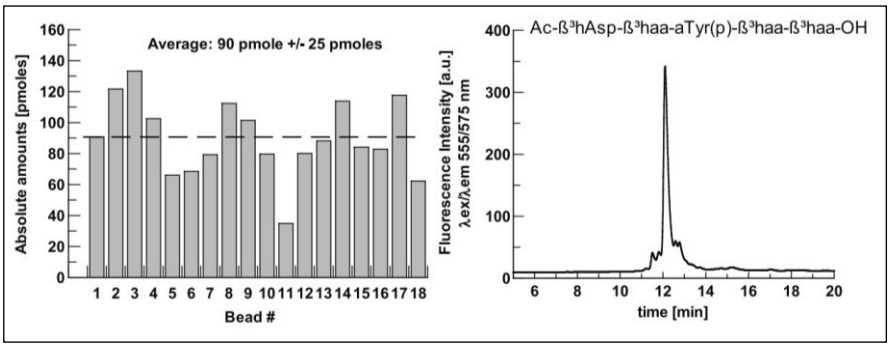


Figure 1-9: Analysis of fluorescently labelled single beads shows ~90 pmol material is recovered in high purity following PS/PS labelling.[9]

In addition to the chemical modifications made to the screening process, some technical advances were also made. The PS04 was developed as a faster confocal scanning microscope (due to a Nipkow scanning disc) for CONA screening to enable rapid on-bead screening. The COPAS (Union Biometrica) is a flow cytometer which can sort beads at a rate of 20 beads per second. This is used as a faster alternative to the capillary picker arm on the PS02. Finally, a high resolution imaging and micro-spectroscopy instrument (PS03), was introduced in order to perform cellular imaging experiments with hit compounds in either the labelled or unlabelled form. This allowed a mechanistic investigation element to be incorporated into the screening process.[11]

PS/PS labelling was proven to work reliably in the synthesis of a variety of libraries, including a β-peptide library based on the (M)-3_{14}-helix which provided a 102 ± 7 nM inhibitor of the p53/hDM2 interaction.[23] A mixed α,β-peptidomimetic library designed to bind SLAM-associated protein (SAP) was designed and synthesised. This led to rich SAR knowledge of the peptidic ligands and compounds which showed ~ 30 μM (K_d) binding

affinities to the protein.[24] A summary of this elegant approach to on-bead screening is depicted in Figure 1-10.

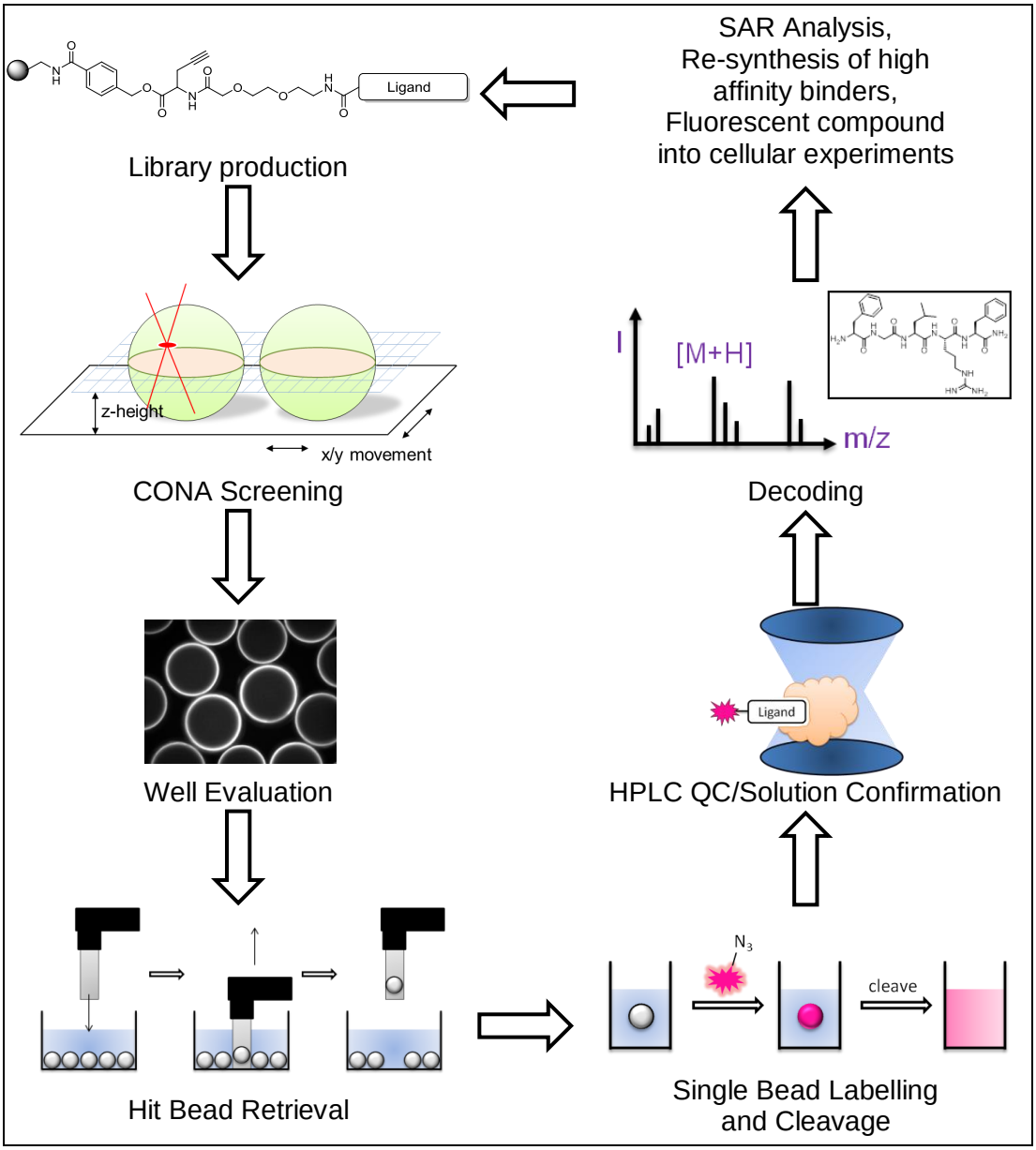


Figure 1-10: PS/PS labelling and CONA screening

The PS/PS labelling – CONA approach has provided an ultra-high throughput screening platform which allows quantitative solution-derived binding data to be obtained for potential hit compounds identified by on-bead screening. This confirmation is carried out using the material present on a single bead, therefore eliminating the wasteful cost and time associated with decoding and resynthesising compounds which are revealed to be too weak binders in solution. Vast compound libraries incorporating expensive reagents such as β-amino acids, which would otherwise be too costly to produce, are easily accessible using this method. The

ability to generate fluorescent probes from each library member even allows highly sensitive quality control and follow-up mechanistic assays such as cellular imaging to be carried out from the same single bead.

1.6 Dyes, Linkers, Tags and Libraries – New Tools for Systems Chemical Biology

In 2009, the Auer group left Novartis and established an academic group at the University of Edinburgh. Expansion of the screening platform remained a significant priority due to the vast array of biological targets which academic groups requested to be screened and the various models in which the compounds would be tested. New tools were also required to generate multifunctional probes from hit compounds in order to elucidate their modes of action.

Although the screening platform developed in Novartis enables the screening of almost any target stable in solution, this one prerequisite excludes integrated membrane proteins, such as GPCRs. Chapter 2 of this thesis describes efforts towards expanding the capabilities of the CONA – PS/PS screening platform to include this highly important protein class using GPCR-54 and its natural ligands, the kisspeptins, as a case study. Fluorescently labelled kisspeptin and a kisspeptin resin were developed as tool compounds for this investigation. Part of the SAR exploration into the kisspeptins looked into their highly conserved C-terminal RFamide motif and the generation of an C-terminally aminated analogue. This led to the development of a novel linker for the attachment of amines to a solid support, which was designed and synthesised for application in this study.

The synthesis of fluorescently labelled Kisspeptin probes for assay development required the production of suitable fluorescent dyes. Rhodamine material was generated in house, however methods followed from the literature rely on arduous and unreliable separation of regioisomeric intermediates. Inspired by the phthalide compounds used in Chapter 3 for the synthesis of a novel amine linker, Chapter 4 describes a novel synthetic route developed as part of this thesis project that gives easy and direct access to valuable single isomer rhodamine dyes.

This rhodamine material was subsequently used as a reporter group in labelled peptides and in multipurpose trifunctional reagents which aimed to expand the CONA PSPS screening platform. The generation of a resin which allows rapid parallel synthesis of trifunctional reagents is detailed in Chapter 5. Here, an orthogonally protected trifunctional backbone

(TOBa) resin is utilised in the generation of tools for new on-bead screening methods, such as reverse CONA, where proteins are simultaneously fluorescently labelled and immobilised onto resin in order to allow the probing of protein-protein interactions. The reagents also allow the tagging of alkyne functionalised library members *via* an azide functionality on one arm, allowing the single reactive site to accommodate two useful groups.

Trifunctional reagents are also proving useful in the elucidation of protein structure through their use in the area of chemical cross-linking teamed with mass spectrometry. During the use of this technique, difficulty is faced in identifying cross-linked peptides in the myriad of non-modified peptide fragments present in a sample. To this end, a trifunctional reagent was developed which allowed a purification group to be incorporated into a homobifunctional cross-linker and is detailed in Chapter 6. The synthesis and characterisation of a novel benzophenone containing heterobifunctional cross-linker which shows specificity complimentary to examples already reported is also described.

CHAPTER 2

TOWARDS A HIGH THROUGHPUT SCREENING PLATFORM FOR GPCR54 USING TAGGED ONE BEAD ONE COMPOUND LIBRAIRES

2.1 Abstract

Although the CONA PS/PS screening platform enables the screening of almost any soluble protein, this one prerequisite excludes integrated membrane proteins. This chapter describes efforts towards expanding the capabilities of the CONA – PS/PS screening platform to include G protein coupled receptors. GPCR-54 and its natural ligands, the kisspeptins, were used as a case study. Active fluorescent probes based on KP-10 were synthesised and a resin presenting KP-10 in the correct orientation was developed. GPCR54 was expressed in various formats which mimic its natural environment in an effort to retain its activity during screening.

2.2 Introduction

The CONA-PS/PS platform introduced in Chapter 1 allows rapid, ultra high throughput screening of any protein which can be solubilised and fluorescently labelled. Membrane proteins reside within lipid bilayers, causing their endogenous activity to be dependent upon their location within this specialised biochemical environment.[25] Therefore the need for solubilised protein in CONA appears to preclude the use of membrane proteins in this screening setup. As it has been estimated that around 26% of all the human protein coding genes translate to membrane proteins,[26] it follows that roughly a quarter of all human proteins are excluded from surface based screening. This important protein class, which includes enzymes, receptors, transporters and ion channels, represents the targets for over half of known drugs,[27] therefore their exclusion from this screening setup is a major drawback. The single largest protein class within this family is the G-Protein Coupled Receptors (GPCRs), of which predictions show more than 800 are coded for in the human genome.[28]

2.2.1 G-Protein Coupled Receptors (GPCRs)

Ligands recognised by GPCRs are wide-ranging and include neurotransmitters, hormones, proteins and even electromagnetic radiation. They are found ubiquitously throughout the human body and are involved in most physiological processes,[29] therefore it is not surprising that >40% of currently marketed drugs are modulators of GPCRs.[30]

2.2.1.1 GPCR Structure and Function

GPCRs all have a common motif of 7 transmembrane α-helical regions arranged in a barrel-like confirmation, with intracellular and extracellular loops joining them. The protein

N-terminus is on the cytosolic side of the cell membrane and the *C*-terminus is on the intracellular side. GPCRs are classified into 6 main groups based on sequence homology, shown in Table 2-1.[29]

Class	Name	Number in human
Class A (or 1)	Rhodopsin-like	662
Class B (or 2)	Secretin receptor family	15
Class C (or 3)	Metabotropic glutamate/pheromone	22
-	Adhesion	33
-	Frizzled	11
-	Taste type 2	25
-	Others	23

Table 2-1: Classes of human GPCRs

The largest class by far is the Rhodopsin-like family of receptors, which itself represents a widespread family of proteins made up of 19 sub-groups.

Upon extracellular binding of ligands, GPCRs undergo a conformational change, enabling them to transduce signals into the cell *via* activation of a G-protein bound to their *C*-terminus (Figure 2-1). The G-protein is a heterotrimeric protein made up of an α-, β- and γ- subunit. GPCR activation catalyses guanine nucleotide exchange at the G-protein α-subunit, where GDP is exchanged with GTP, which causes the whole trimeric protein to dissociate from the GPCR. The α-subunit separates from the βγ-dimer, both of which go on to active downstream effectors. Following signal propagation, the Gα-GTP complex is hydrolysed back to GDP, causing the whole system to re-associate. In addition to this, GPCRs can signal through G-protein independent pathways.[31]

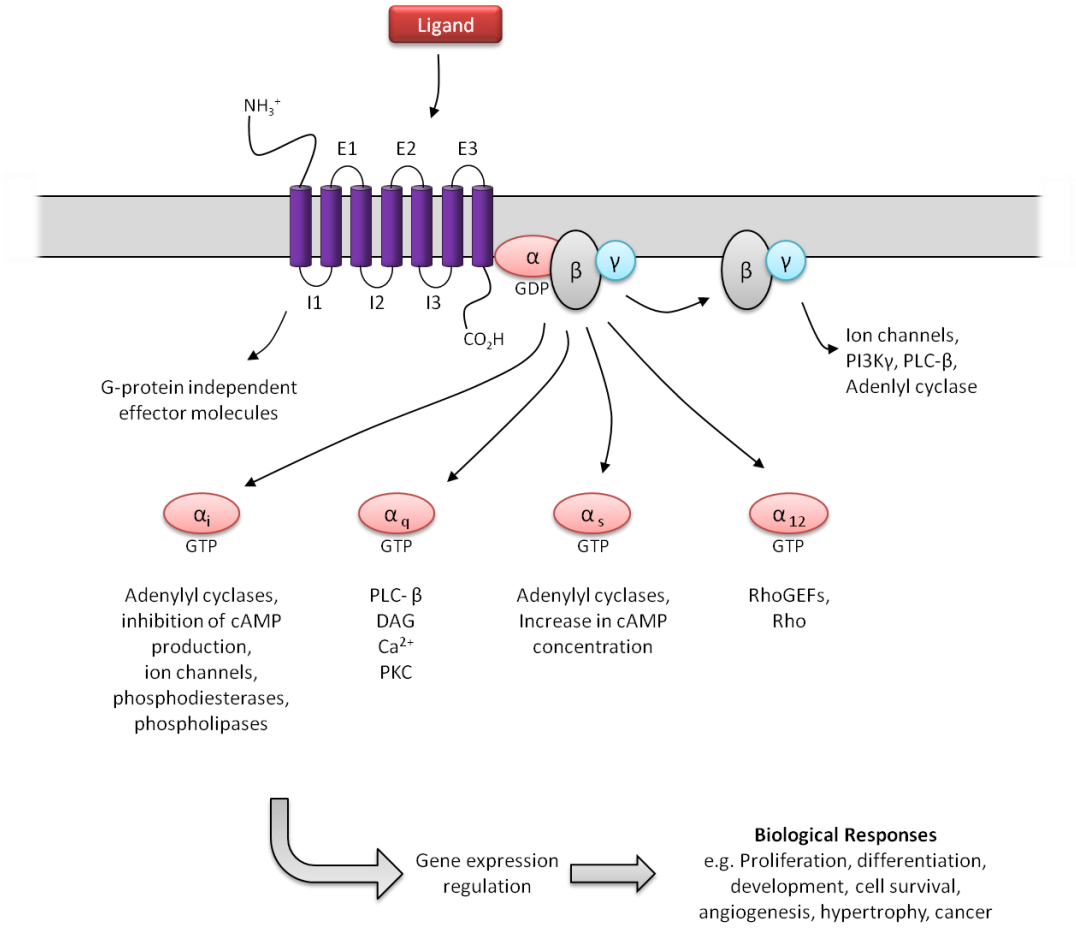


Figure 2-1: GPCR signalling[32]

There are four main classes of G-protein; G_i (inhibits adenylyl cyclase), G_q (activates phospholipase C), G_s (activates adenylyl cyclase) and G_{12} of little known function.[33] As well as these distinct pathways for the various α-subunits of G-proteins, there are believed to be six gene products which encode β-subunits and 12 for γ-subunits.

2.2.2 GPCRs and OBOC Libraries

2.2.2.1 Screening in Solution

OBOC libraries have previously been used to identify novel ligands for GPCRs, although in most reported cases, compounds are cleaved from the resin and assays carried out in solution. Competitive radioligand binding assays have been used for screening library compounds.[34-37] Here, compounds are first screened in bulk as whole sub-libraries, due to the expense and time consuming nature of the assay. Only active sub-libraries are taken on for single compound analysis. Although this approach allows the successful screening of OBOC libraries against GPCRs, it risks missing potent compounds in a sub-library of low affinity ones.

In 1994, Jayawickreme and co-workers synthesised a 31,360 member peptidic library based on α-melanocyte-stimulating hormone (α-MSH) and it was assayed for inhibition of α-MSH receptor.[38] Here, a functional cellular assay observing pigment translocation was used to screen library compounds. Resynthesis of apparently active compounds was required to obtain quantitative activity data, which in this case led to the identification of an antagonist compound with an IC_{50} of 11 ± 7 nM. Similarly, Iuga *et al*. synthesised a OBOC 1,000,000 member hexapeptide library and used it to screen against the melanocortin receptor expressed in *Xenopus laevis* (African clawed frog).[39] Pigment dispersion was again used as readout. This assay allowed the identification of peptides able to stimulate pigment dispersion with EC_{50} ~ 0.1 – 1 μM. Although useful, this functional assay could not be used as a general screen for GPCRs, as the readout is unique to this family of receptor.

2.2.2.2 Screening on Bead

On-bead screening using whole cells was first reported by the Lam group, who used the method to identify binders of $\alpha_4\beta_1$ integrin.[40] The Kodadek group developed this method to generate an On-Bead Two-Colour (OBTC) cell screen,[41-43] which was used to identify antagonists of oxerin receptor 1 (OxR1) from a 250 000 member peptoid library in the first report of an on-bead screen against a GPCR (Figure 2-2).[44] Here, a wild-type cell line was labelled with green quantum dots and cells differing only by the expression of OxR1 were labelled with red quantum dots. Both cell types were then mixed together and incubated with library compounds. Upon visual inspection, beads which had bound both red and green cells were likely to contain compounds which have affinity for other cell surface molecules, however beads which bound only the red cells were assumed to have selective affinity for OxR1. These hit beads were extracted from the bulk of the resin by addition of anti-OxR1 polyclonal antibody from rabbit, followed by bead washing. Sheep antirabbit IgG secondary antibody coated iron oxide particles were then added, allowing magnetic retrieval of beads which bind OxR1. Visual inspection then allowed manual isolation of the selective (red cell only) beads from the mixture.

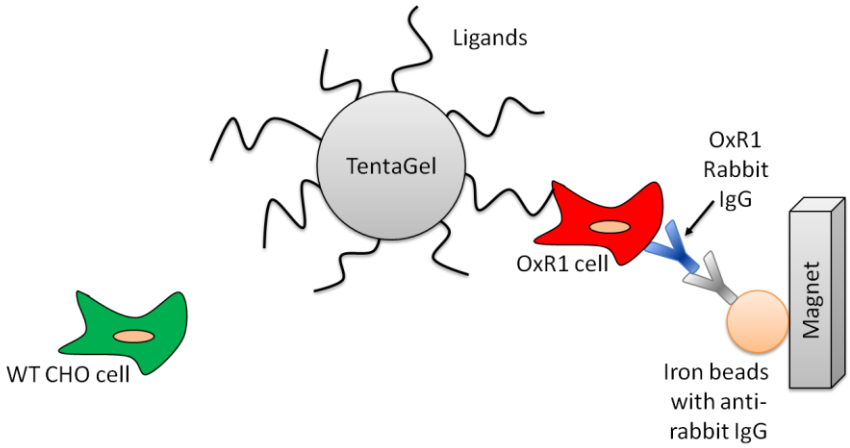


Figure 2-2: OBTC screen used by Qi *et al.* to identify ligands of OxR1 from an OBOC library[44]

Following the screen, a peptoid which antagonises OxR1 with an IC_{50} of 51 ± 3 µM was identified. This on-bead screening approach represents a universal, inexpensive and rapid way to screen GPCRs in their natural environment against OBOC libraries. However, for the purpose of hit confirmation, compounds from the on-bead screen require re-synthesis.

2.2.2.3 Towards GPCR Screening using the CONA – PS/PS platform

The aim of this project was to enable this highly important protein class to be screened using the CONA and PS/PS technology already established. GPCR54 and its natural peptidic ligands the kisspeptins were used for this investigation.

2.2.3 GPCR54

GPCR54 was an orphan receptor when discovered during a PCR search of rat brain cDNA in 1999 by Lee *et al.*[45] The 396 amino acid protein shares significant identities (~45%) with the Galanin receptors, although no binding to Galanin is observed. The group demonstrated that the receptor is localised in the CNS and peripheral area and suggested that the natural ligand would be peptidergic in nature. A blast search using the rat GRP54 sequence revealed a human orthologue which encoded a 398 amino acid protein which had 81% translated amino acid homology. 100% Amino acid identity was found in the transmembrane regions.

In 2001, Clements *et al.* cloned and sequenced rat, mouse and human GPCR54.[46] The rat and mouse sequence share 95% homology and both share around 85% with human. A series of neuropeptides were screened against the receptor and it was found that three, all of which bear an RF-amide motif at the *C*-terminus, activated GPCR54 with micromolar affinity. It

was found that GPCR54 couples through the $G_{\alpha q}$ pathway. GPCR54 is a member of the Rhodopsin family of GPCRs in Class A.

2.2.4 The KiSS-1 Gene

In 1996, Lee at al discovered a novel melanoma and breast cancer metastasis suppressor gene, which was named KiSS-1. [47, 48] The name was intended to acknowledge its discovery in Hershey, Pennsylvania (named after Hershey's kisses, a confectionary made there), as well as to include nomenclature indicating it as a putative suppressor sequence. The structure and localisation of KiSS-1 was reported by West *et al.* who showed that the gene maps to chromosome 1 on band 1q32.[49] The whole coding sequence is made up of 438bp and this gives rise to a 145 amino acid peptide. KiSS-1 contains two non-coding exons and two partially coding exons.

2.2.5 The Kisspeptins

The KiSS-1 gene gives rise to a 145 amino acid precursor protein, prepro-kisspeptin (Figure 2-3). This contains a 19-amino acid signal peptide and a main 54-amino acid region flanked by cleavage sites. Upon cleavage, kisspeptin-54 (KP-54, metastin) is liberated and amidated. This is further cleaved to give lower molecular weight kisspeptins (14-, 13- and 10 amino acid lengths). The kisspeptins bind GPCR54 with nanomolar affinity and are members of the RF-amide family of neuropeptides (for more information see chapter 3), all of which contain an RF-amide motif at the *C*-terminus, which is essential for receptor binding.

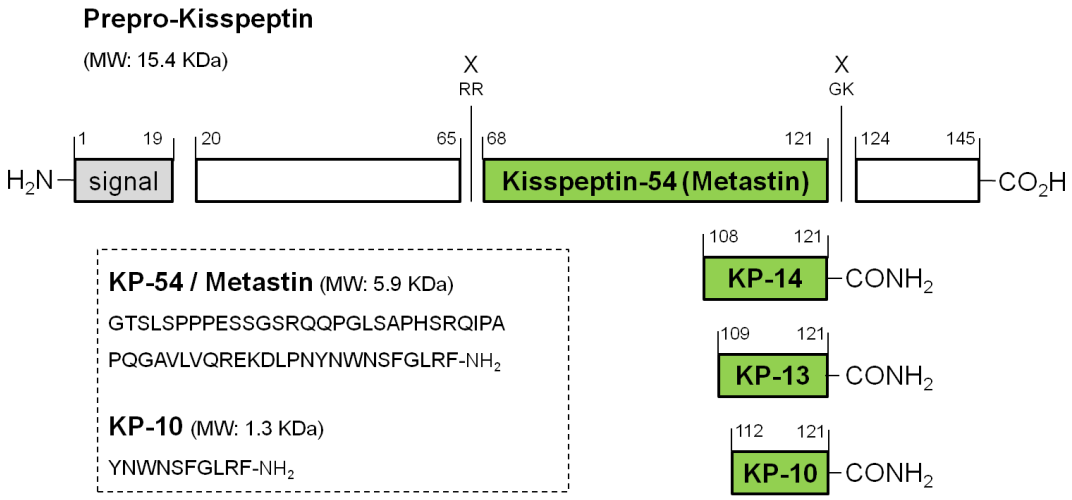


Figure 2-3: The KiSS-1 gene encodes the Propro-kisspeptin, which is cleaved and modified to give amidated KP-54, KP-14, KP-13 and KP-10.

In 2001, three groups made the discovery that the natural ligands for GPCR54 are the peptidic products of the KiSS-1 gene almost simultaneously. Ohtaki *et al.*[50] and Kotani *et al.*[51] both observed that human placental extracts induced a response when exposed to cells transfected with GPCR54. The peptide responsible for this effect was isolated and identified as kisspeptin-54. Muir *et al.* created synthetic versions of the peptides to test against the receptor.[52]

Ohtaki *et al.* focused their studies on the ability of kisspeptin and GPCR54 to prevent metastasis in tumours (which led to the alternative name for KP-54: 'Metastin') whereas Katoni *et al.*, based on localisation of the receptor and KiSS-1 in the body, suggested it would play a role in endocrine function. Immunohistochemical localisation and mRNA analysis of receptor and ligand within the brain and peripheral tissues showed GPCR54 and KiSS-1 within the hypothalamus, brain stem, spinal cord, pituitary, ovary, prostate and placenta.

2.2.6 Signalling in the Cell

GPCR54 signals through G protein ($G_{\alpha q/11}$) cascades (Figure 2-4). This involves activation of phospholipase C which in turn hydrolyses phosphatidylinositol-4,5-biphosphate (PIP_2) into diacyl glycerol (DAG) and inositol triphosphate (IP_3). IP_3 activates calcium channels, which leads to intracellular Ca^{2+} mobilisation. Protein kinase C is also activated by IP_3 and DAG. For this reason, calcium mobilisation assays are frequently used as functional assays when assessing GPCR54 activation.[45, 50, 52]

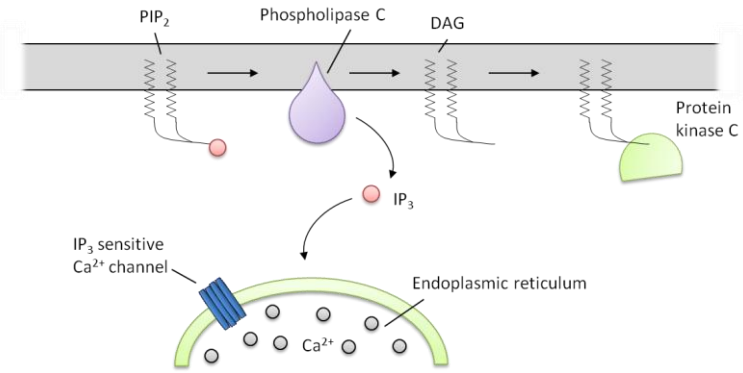


Figure 2-4: G_q coupled GPCR downstream signalling in the cell

Additionally, GPCR54 has been shown to activate pathways related to mitogen activated protein kinases (MAPK), especially ERK1/2 and p38 and phosphatidylinositol-3-kinase (PI3K)/Akt. GPCR54 can also interact with other receptors, including chemokine receptor CXCR4, and GnRH receptor.[53]

2.2.7 Function in the Body

Although the kisspeptins were initially discovered as anti-metastasis agents,[47, 48, 54] kisspeptin and GPCR54 were soon found to have a much more vital role in the hypothalamic-pituitary-gonadal (HPG) axis.

Discovered in the 1970's, Gonadotropin-releasing hormone (GnRH) is the major hormone in the human reproductive system (Figure 2-5). It is produced by neurones in the hypothalamus and stimulates cells in the pituitary gland named gonadotrophs. This consequently produces the gonadotrophin hormones: follicle-stimulating hormone (FSH) and luteinising hormone (LH). These stimulate the gonadal secretion of sex steroids testosterone and estrogen, which go on to stimulate the gonads (ovaries and testes) causing them to mature and generate sperm or eggs.[55] In a feedback loop, the sex steroids inhibit the secretion of GnRH and the gonadotrophes, leading to the secretion of LH in pulses.

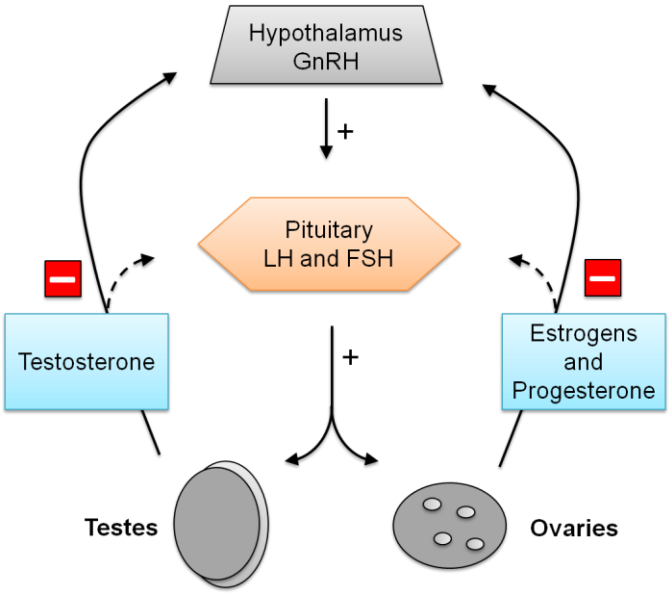


Figure 2-5: Regulation of human reproduction *via* Gonadotrophin Releasing Hormone (GnRH)

Isolated hypogonadotropic hypogonadism (IHH) is a disease defined as the impairment of pubertal maturation and reproductive function due to a deficiency of the pituitary secretion of FSH and LH.[56] In 2003, it was reported that a case of IHH in one consanguineous family was due to an inactivating mutation in GPCR54,[57, 58] with several more cases having since been reported.[59-63] Generation of GPCR54 knockout mice was found to have the same effect, with the mice exhibiting symptoms of IHH, including underdeveloped genitalia and low levels of gonadotropin hormones.[64] KiSS1 null mice displayed the same phenotypes as those of GPCR54 knockout mice and exhibited the same failure to go through sexual

maturation.[65] An inactivating mutation in the KiSS-1 gene was also identified as the cause of IHH in one family.[66]

Conversely, activating mutations of GPCR54 have been identified as the cause of central precocious puberty (CPP) in some patients,[67, 68] which is defined as the early onset of puberty due to a cause which can be traced to the hypothalamus or pituitary. A mutation in the KiSS-1 gene leading to a more stable kisspeptin mutant has also been shown to cause CPP[69] and elevated kisspeptin levels were found in patients with premature thelarche.[70]

In 2005, it was proven that kisspeptin directly stimulates GnRH release *via* GPCR54,[71] which was found to co localise with GnRH neurons in the mouse hypothalamus. Injection of kisspeptin resulted in a robust rise in plasma LH and FSH levels in mice, whereas administration of a GPCR54 antagonist inhibited the firing if GnRH neurons and the release of LH in rats and mice.[72] In GPCR54 knockout mouse models, administration of GnRH was able to rescue the function of the reproductive system, indicating that GPCR54 is not required for the development of downstream messengers.[55] Conversely, the use of GnRH antagonists such as acycline inhibits the ability of kisspeptin to induce the release of LH and FSH in mice and rats,[73, 74] therefore proving that the role of GPCR54 and the kisspeptins is to activate the HPG axis *via* regulation of GnRH secretion.

In humans, kisspeptin-54 has been shown to stimulate the release of reproductive hormones in women, even those suffering from Hypothalamic Amenorrhea.[75-77] It has also been used to trigger egg maturation in women undergoing IVF treatment.[78] In men, increased LH, FSH and testosterone levels have been observed following administration of kisspeptin-54[79] and kisspeptin-10.[80] For these reasons, GPCR54 and the kisspeptins have been named the gatekeepers of reproduction.[56]

As well as stimulating GPCR54, it has also been shown that human kisspeptin-13 has the ability to activate Neuropeptide FF2 Receptor (NPFF2R) with an EC_{50} value of around 100 nM, indicating that the kisspeptins may have a role in the regulation of pain, for example.[81] A recent study has indicated that kisspeptin signalling may be involved in metabolism and obesity.[82]

These findings suggest that the development of novel compounds which modulate GPCR54 could be used as therapeutics for disorders of reproductive function, such as IHH, hypothalamic amenorrhoea (HA) and prostate and breast malignancies.[83] Recent studies

may also reveal novel uses for kisspeptin analogues as therapeutics in areas such as modulation of pain and obesity.

2.2.8 Kisspeptin Structure Activity Relationship

Since their emergence as major players in the regulation of the reproductive system, the kisspeptins have received a widespread attention from numerous research groups, interested in elucidating their mode of binding and in the generation of potent and stable analogues which may have therapeutic potential. Below is a review of the works carried out in this field to date.

2.2.8.1 GPCR54 Agonists

Before the kisspeptins were known to be the natural ligands of GPCR54, Clements *et al.* carried out a screen of compounds. They reported a series of invertebrate neuropeptides of the RFamide and RWamide family which acted as micromolar agonists of the receptor.[84] Through substitution analysis, it was demonstrated that the compound giving best activation of GPCR54 in a calcium mobilisation assay was a doubly substituted neuropeptide NF1 analogue, with the *C*-terminal sequence of Gly-Leu-Arg-Trp-NH_2 (Figure 2-6).

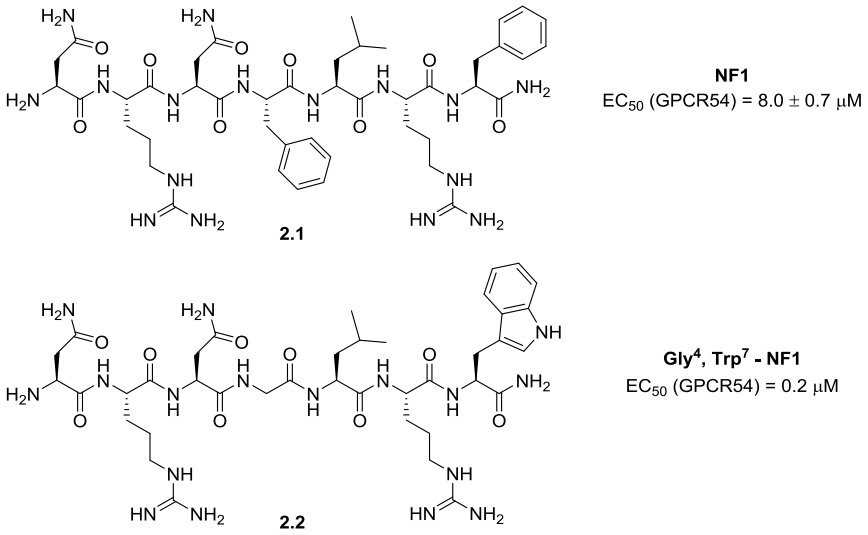
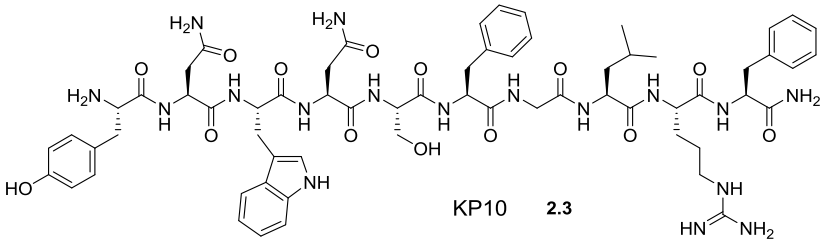


Figure 2-6: Optimal *C*-terminal sequence for GPCR54 activation suggested by Clements *et al.*

Muir *et al.* carried out a screen of 1500 putative ligands of the receptor, which also revealed a series of invertebrate neuropeptides to be the best hits, all incorporating an amidated LRF or LRW *C*-terminal motif (Table 2-2).[52] A literature search led them to assay the kisspeptins which showed very high activation of GPCR54, and with KP-10 proving to be the most potent in a calcium mobilisation assay.



Peptide	Sequence	pEC_{50}
Antho-RW-amide II ***(Sea anemone)***	pESLRW-amide	5.73
Antho-RW-amide I ***(Sea anemone)***	pEGLRW-amide	6.42
Peptide F1 ***(Lobster)***	TNRNFLRF-amide	5.63
KP-54	GTSLSPPPESSGSRQQPGLSAPHSRQIPAP QGAVLVQREKDLPNYNWNSFGLRF-NH_2	8.00
KP-10	YNWNSFGLRF-NH_2	9.30

Table 2-2: Activation potential of KP-54 and KP-10 against GPCR54 as reported by Muir *et al.* (top), structure of KP-10 (bottom).

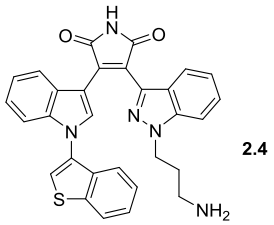
Orsini *et al.* carried out structural NMR experiments using KP-13 in an SDS micelle containing (membrane mimicking) environment,[85] which showed the peptide to have a stable helix structure from residues Asn^7 to Phe^{13}, in agreement with findings from other groups.[86, 87] Their interpretation of the NMR structure showed Phe^9 and Phe^{13} lying on top of each other, flanked by Arg^{12}. This information was used in conjunction with SAR data obtained experimentally to suggest a pharmacaphore for kisspeptin-GPCR binding and use that to screen *in silico* for possible small molecule binders.

The group performed alanine scanning on KP-10 in an attempt to elucidate some structure activity relationship information. From this it was deduced that Phe^6, Arg^9 and Phe^{10} are critical residues for binding, with significant loss of activity observed upon substitution. Leu^8 also appeared sensitive. Further substitutions were performed at the 'pharmacaphore site', Phe^6 – Phe^{10}, in order to find out what was tolerated (Table 2-3).

Residue	Tolerated	K_i (nM)	Not Tolerated	K_i (nM)	Comments
Phe[6]	Tyr	16.2	His	109.6	Aromatic ring required. Cyclohexyl may be ok.
	Trp	2.0	Arg	196.7	
	CHA	7.7	Leu	141.8	
	3-Pal	4.7	Gln	298.2	
Gly[7]			Leu	775.2	Even small changes not tolerated.
			Pro	10 000	
Leu[8]	Nle	15.8	Ile	51.5	Authors comment that Ile may be inducing turn structure which is not tolerated.
Arg[9]	Lys	9.4	Cit	88.1	Basic group required
			β-Phe	10 000	
Phe[10]	Tyr	2.2	His	161.8	Aromatic ring required. Cyclohexyl may be ok.
	Trp	1.6	Arg	10 000	
	CHA	11.1	Leu	80.8	
	3-Pal	9.3	Gln	494.9	
***C*-terminal amide**			CO_2H	551.8	Amide essential for binding

Table 2-3: Tolerances of substitutions made on KP-10
CHA = cyclohexylalanine, 3-Pal = 3-pyridinylalanine, Cit = citrulline

From this information, a pharmacaphore was proposed that involved Phe[6], Arg[9] and Phe[10]. The pharmacaphore was implemented in a small molecule screen *in silico*, aimed to search for compounds which mimic the features described. This resulted in the first small molecule agonists of GPCR54, though they possess reduced potency compared to the kisspeptins. The best compound, **2.4**, has a K_i of 0.708 μM (Figure 2-7). Although agonists were identified, perhaps the relatively low activity of the compounds could be attributed to the fact that the kisspeptins adopt a different conformation to the one taken up in a micelle environment when binding GPCR54, therefore leaving this pharmacaphore unsuitable.



2.4

K_i = 0.708 μM

Figure 2-7: Small molecule GPCR54 agonist reported by Orsini *et al.*

Over the past 9 years, the Fujii group have focused much attention on the development of GPCR54 modulating compounds. In 2006, they reported an SAR investigation of KP-10 including alanine and D-amino acid scanning.[88] The results suggested that the five *C*-terminal residues (Phe[6] – Phe[10]) are highly important for receptor binding, as replacement

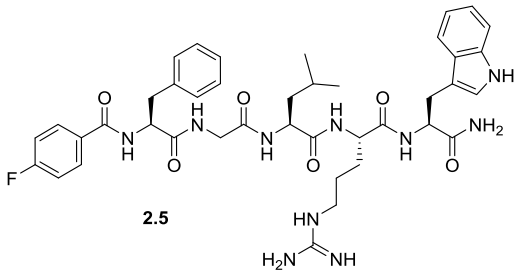
of these for their non natural D-amino acid counterparts shows a significant drop in activity. Replacement of Phe[6], Leu[8] or Phe[10] with alanine results in a complete loss of activity.

Highly potent downsized kisspeptin analogues were also developed,[89-91] by keeping the five important *C*-terminal residues (with Phe[5] substituted for Trp) and acylating the *N*-terminus with an aromatic group. An SAR and QSAR study was undertaken, firstly on the amino acid residues within the pentamer. Results are summarised in Table 2-4.

Residue	Tolerated	% Activity	Not Tolerated	% Activity	Comments
Phe[1]	Phe Trp NaI(2)	97 78 89	Cha Tyr	44 10	Aromatic groups required.
Leu[3]	Val Nva Cha NaI(2)	101 65 47 49	Abu Nle Ile Ser Hse Phe	32 0.3 9.6 0 0 22	β-methyl group could increase activity. Polar residues not tolerated. Hydrophobic groups required.
Arg[4]			Orn Lys Cit Lys(Ac) Gln Glu	1 1 1 0 0 1	Substitutions not well tolerated, Arg critical for receptor interaction
Trp[5]	Phe Tyr NaI(2)	70 60 68			Trp better than Phe. Naphthalene maybe too large.
***C*-terminal amide**			-Tryptamine -Trp-NHMe Trp-NMe_2 -Trp-$NHNH_2$	4 0 0 0	Amide essential for activity

Table 2-4: Amino acid substitutions on downsized GPCR54 agonists
Nva = norvaline, Nle = norleucine, NaI(2) = 3-(2-naphthyl)alanine, Hse = homoserine

Next, the acylating group was investigated, the results of which showed the most potent analogue to be a 4-fluorobenzolyl acylated compound, with an EC_{50} of 0.69 nM (Figure 2-8).



2.5

Figure 2-8: Peptidic Agonist of GPCR54 Reported by Tomita *et al.*

The group then carried out modifications to the backbone of this potent compound in order to counteract degradation carried out by peptidases such as MMP-9 between the leucine and

glycine residue. A variety of isosteres were generated and evaluated, including a series of alkenes, fluoroalkenes, allylic and alcohol-containing isosteres (Figure 2-9).[92, 93] The most promising of these were the trans alkene and hydroxyethylene type isosteres, which showed higher resistance to MMPs and higher stability in murine serum compared to KP-10, whilst maintaining biological activity.

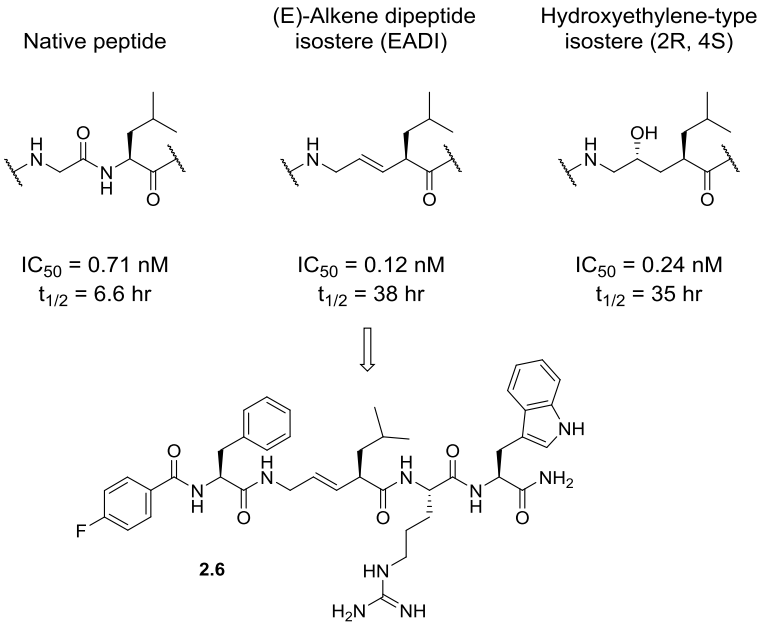


Figure 2-9: Activity and half-lives of isosteres of Gly-Leu peptide bond

KP-13 has been demonstrated to activate NPFF2R, therefore the selectivies of downsized GPCR54 agonist **2.6** was assessed against a range of RFamide peptide receptors. The compound demonstrated activity toward NPFFRs, but less than that of KP-10.[94]

During the final year of this PhD project, the Fujii group reported the design and synthesis of fluorescent probes for GPCR54 based on KP-52 and KP-14 (Figure 2-10).[95] Rhodamine green (RG) or tetramethylrhodamine (TMR) was attached to the *N*-terminus of the kisspeptin compounds, leaving the essential *C*-terminal amide free. The RG probes were synthesised using 1,3-dipolar cylcoaddition, as amide coupling resulted in substantial side products. All of the probes showed high activity in a radioligand competition assay against radiolabelled KP-10, proving that the bulky fluorophore groups do not interfere with binding to GPCR54.

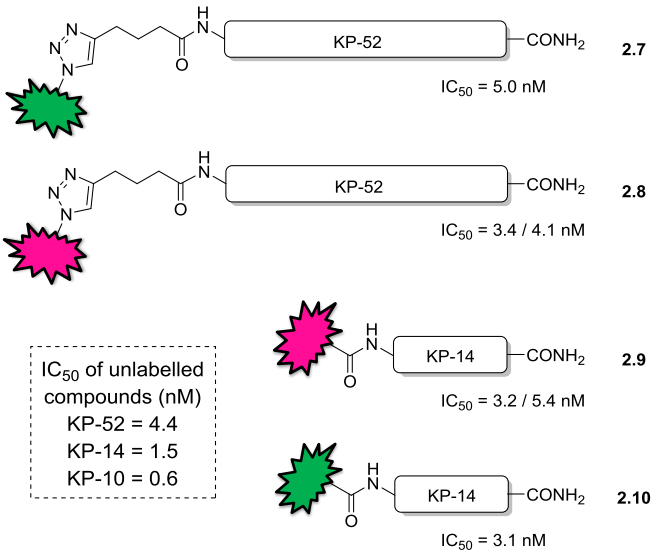


Figure 2-10: Activities of fluorescently labelled kisspeptin compounds (pink = tetramethylrhodamine, green = rhodamine green). Two values are shown for TMR probes representing the two regioisomers of the dye.

After incubation of the probes with GPCR54 expressing CHO cells, fluorescent signals could be seen in the intracellular compartments, although interestingly the localisation was different depending on the dye used, as a portion of the TMR probes were seen in a specific region. The staining was inhibited by the presence of unlabelled KP-10. Fluorescence signals were detected when GPCR54 expressing cells were incubated with the probes then analysed by FACS. Again, fluorescence was not detected in the cells when unlabelled KP-10 was present. These finding suggest that the probes bind to GPCR54 in a specific manner. Intravenous injection of TMR-KP-14 elicited a response expected from the natural ligand (increased plasma LH level), suggesting that the dyes do not interfere with the compounds mode of action *in vivo*.

Recently, modifications to KP-10 have also been carried out by Asami *et al.* at Takeda Pharmaceutical Company Ltd. in order improve its stability following the observation that over half of the peptide was metabolised in mouse serum within 1 minute at 37 °C.[96, 97] Positions vulnerable to protease cleavage were identified at Tyr^{1}-Asn^{2} and Arg^{9}-Phe^{10} (numbered from KP-10). It was thought that replacement of Arg^{9} with a non-natural amino acid would prevent degradation at this site, therefore a range of KP-10 analogues were prepared and tested for agonist activity and stability. Indeed, N^{ω}-methylarginine and $N^{\omega\omega}$-dimethylarginine (asymmetric) showed higher agonistic activity compared to KP-10 (by Ca^{2+} mobilisation). The stability of the N^{ω}-methylarginine analogue was studied, which showed that cleavage at the $Arg(Me)^{9}$-Phe^{10} peptide bond was inhibited.

In addition to this substitution, it was found that replacement of Tyr[1] with D-Tyr allowed agonistic activity to be maintained while avoiding *N*-terminal degradation (Table 2-5).

H-AA[1]-Asn-AA[3]-Asn-Ser-Phe-AA[7]-Leu-AA[9]-Phe-NH_2								
Entry	**AA[1]**	**AA[3]**	**AA[7]**	**AA[9]**	**Agonist activity EC_{50} (nM)**	**Binding affinity IC_{50} (nM)**		**Stability in mouse serum Residual ratio (%)**
						Human	**Rat**	
1	Tyr	Trp	Gly	Arg	0.049	0.08	0.11	0
2	Tyr	Trp	Gly	Arg(Me)	0.023	-	-	-
3	D-Tyr	Trp	Gly	Arg(Me)	0.065	0.11	0.26	18.1
4	D-Tyr	D-Trp	azaGly	Arg(Me)	0.072	0.19	0.22	56.4

Table 2-5: Biological activities and stabilities of KP-10 analogues

However, degradation products were observed from cleavage at other peptide bonds, therefore Phe[6] and Gly[7] were assessed for possible replacements. It is known that substitutions of Phe[6] are not well tolerated due to its importance for receptor binding, therefore isosteres of the peptide bond were focused on, rather than side chain modification. Substitution of Gly with azaGly (Figure 2-11) allowed good activity to be maintained and showed high stability, with the authors commenting that this β-turn inducing motif may be favourable for receptor binding as it allows the peptide to adopt a tight turn conformation.

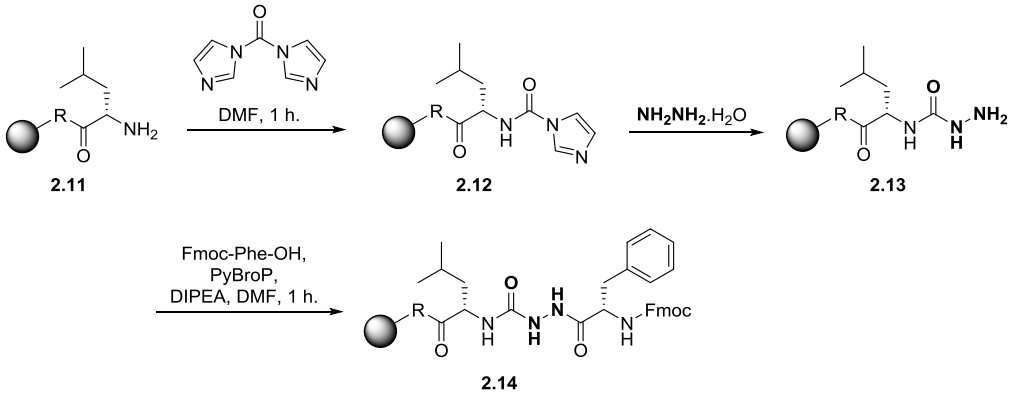
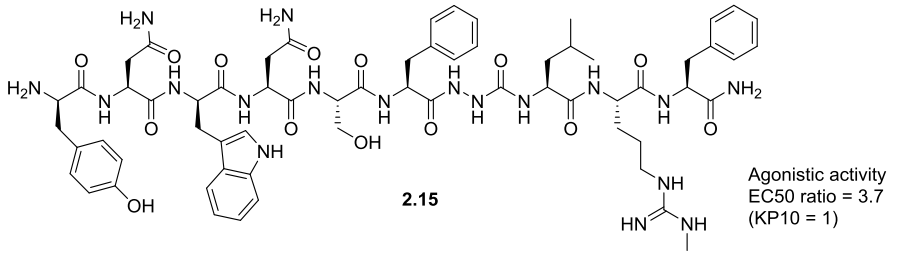


Figure 2-11: Synthesis of azaGly pseudopeptide (azaGly shown in bold)

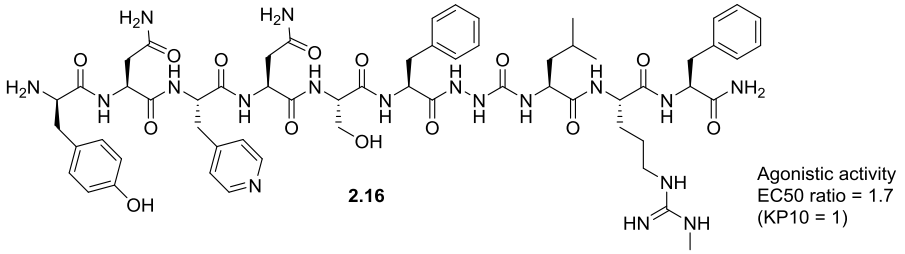
Finally, substitution of Trp[3] for D-Trp gave a peptide with activity comparable to KP-10, with excellent stability in mouse serum with a residual 56% present after incubation for 1 hour at 37 °C (Figure 2-12).



H-D-Tyr[1]-Asn[2]-D-Trp[3]-Asn[4]-Ser[5]-Phe[6]-azaGly[7]-Leu[8]-Arg(Me)[9]-Phe[10]-NH_2

Figure 2-12: KP-10 analogue 2.15 shows high stability in serum and good agonistic activity

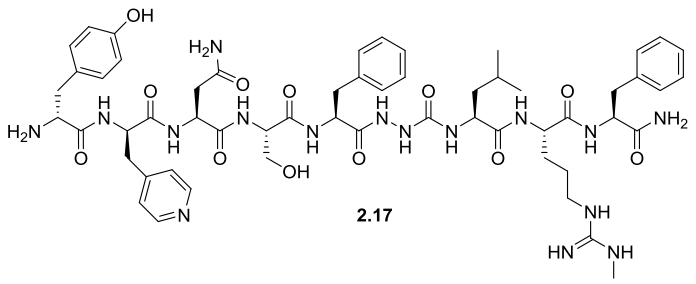
Although good stability had been achieved, the investigation continued in order to achieve improved agonistic activity. Further analogues were prepared by substituting Trp^3. 4-Pyridyl alanine (Pya(4)) was found to be the outstanding replacement, decreasing the agonistic activity ratio from 3.7 to 1.7 (KP-10 = 1) according to Ca^{2+} mobilisation assays (Figure 2-13).



H-D-Tyr[1]-Asn[2]-Pya(4)[3]-Asn[4]-Ser[5]-Phe[6]-azaGly[7]-Leu[8]-Arg(Me)[9]-Phe[10]-NH_2

Figure 2-13: Substitution of D-Trp with Pya(4) increases agonistic activity

Deletion of Asn^2 was also examined, in order to prevent inactivation of the peptide *via* rearrangement of Asn residues sometimes observed in aqueous environments (e.g. conversion to Asp or β-Asp). This resulted in lower activity, however after some exploration, it was found that replacement of Trp^3 with D-Pya(4) resulted in a compound with high binding affinity to the receptor (Figure 2-14). This compound was proven to be long-acting *in vivo*, resulting in it being the first reported kisspeptin analogue to show testosterone-suppressive activity.



2.17

H-D-Tyr[1]-D-Pya(4)[2]-Asn[3]-Ser[4]-Phe[5]-azaGly[6]-Leu[7]-Arg(Me)[8]-Phe[9]-NH_2

Figure 2-14: **Nonamer which shows testosterone-suppressive activity**

These results indicate that peptidic therapeutic compounds must not only be highly potent, but also long-acting. Indeed, Curtis *et al.* reported a decapeptide identical to KP-10 only with the replacement of the tyrosine group with D-tyrosine. This compound proved to have a more potent stimulatory effect on the HPG axis *in vivo* than the endogenous peptide,[98] presumably due to higher stability.

The peptides developed by Asami *et al.* are currently being put forward as candidates for the therapy of diseases such as prostate cancer, which are sex-hormone-dependent. Included in this project are TAK-683 and TAK-448 (Figure 2-15), which have undergone clinical trials[99] where they exhibited anti-tumour activity in several prostate cancer models.[100]

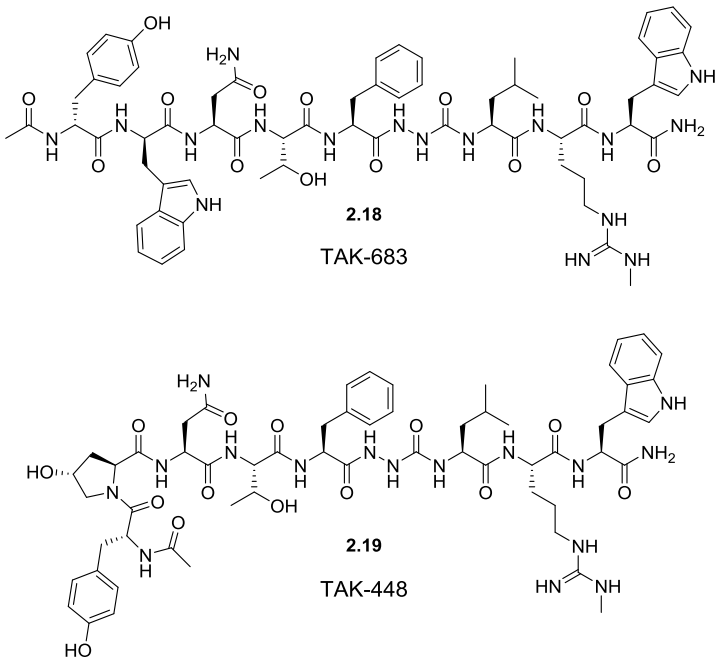


Figure 2-15: TAK-683 and TAK-448 - GPCR54 agonists

2.2.8.2 GPCR54 Antagonists

Roseweir *et al.* reported the first peptidic antagonist for GPCR54. Based on KP-10, intuitive substitutions were made based on SAR investigations. By substituting Leu^8 with D-Trp and Ser^5 with Gly, a potent antagonist was discovered capable of inhibiting inositol phosphate accumulation (Figure 2-16).[72] This compound was instrumental in elucidating crucial information on modes of action of the kisspeptin/GPCR54 pathway *in vitro* and also in some animal models *in vivo*.

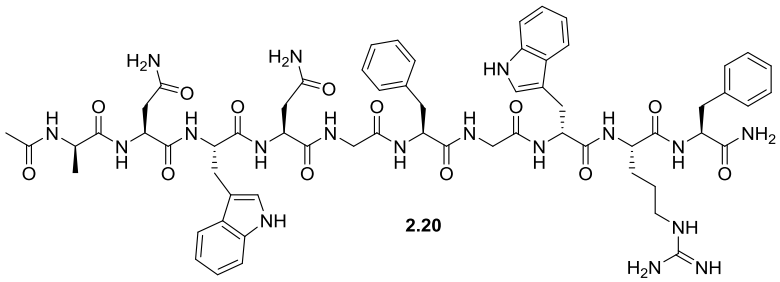
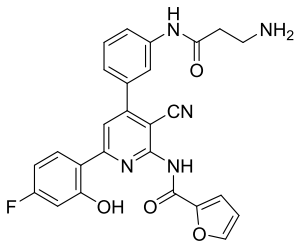


Figure 2-16: A peptidic GPCR54 antagonist developed by Roseweir *et al.*

A series of 2-acylamino-4,6-diphenylpyridines were reported by Kobayashi *et al.* as the first small molecule antagonists of GPCR54. A screen of their compound archive in a receptor binding assay identified compound **2.21**, which has an IC_{50} value of 3.7 nM and shows antagonism in cellular assays (Figure 2-17).



2.21

IC50 = 3.7 nM

Figure 2-17: A small molecule GPCR54 antagonist reported by Kobayashi et al

2.3 Aims

One of the aims of this thesis project was to extend the current TOBOC screening platform established in the Auer lab in order that it can be used to screen on-bead libraries against GPCRs in more efficient ways. Currently, as solubilised proteins are required for CONA screening (see chapter 1), membrane bound proteins cannot be included in the list of screenable targets.

GPCR54 and its endogenous ligands, the kisspeptins, were chosen as the model system for the investigation into possible solutions. Highly potent and stable analogues of this interesting ligand/receptor system could be used in therapeutic areas such as reproductive diseases and sex-hormone related cancers such as ovarian, breast and prostate. Therefore, the two major points of focus for this project were:

1. Investigate the expression of GPCR54 in formats which may allow its use for on-bead screening using CONA

   Or

   Investigate the possibility of a high throughput cellular screening assay against GPCR54 utilising fluorescently tagged OBOC library compounds in-solution
2. Design and synthesise OBOC libraries of kisspeptin analogues using knowledge from reports in the literature

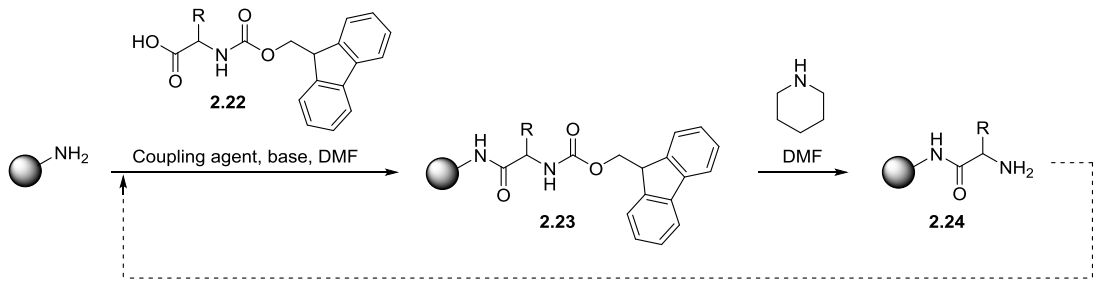
2.4 Results and Discussion

The first step in the development of an assay involving GPCR54 and the kisspeptins was the generation of chemical and biological materials. The synthesis of kisspeptin was the initial focus.

2.4.1 Synthesis of KP-10

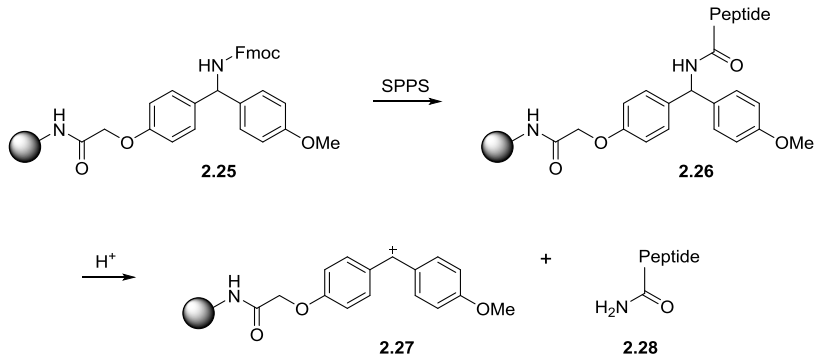
As KP-10 is the smallest of the kisspeptins which retains affinity for GPCR54, it was decided that tool compounds should be based on this decamer. KP-10 was first synthesised in its non-labelled, wild type form to ensure that no 'difficult sequences'[101] exist in the synthesis of the peptide, and also to use as a tool compound in competition assays later in the project. KP-10 was generated using Fmoc based SPPS.[102-104]

Here, amino acids protected at the *N*-terminus by Fmoc protecting group are sequentially added onto a solid support, followed by piperidine mediated deprotection. The sequence is repeated until the full length peptide is synthesised, at which point acid labile peptide side chains can be deprotected and the peptide liberated from the resin (Scheme 2-1).



Scheme 2-1: SPPS by Fmoc strategy. An Fmoc protected amino acid is coupled onto a resin support then the Fmoc group is removed, revealing an amine ready for the next coupling.

For KP-10 to be active, the *C*-terminus must be amidated, therefore the acid labile Rink amide linker was employed as this generates the peptide amide upon cleavage (Scheme 2-2).[105]



Scheme 2-2: RINK amide linker loading and cleavage

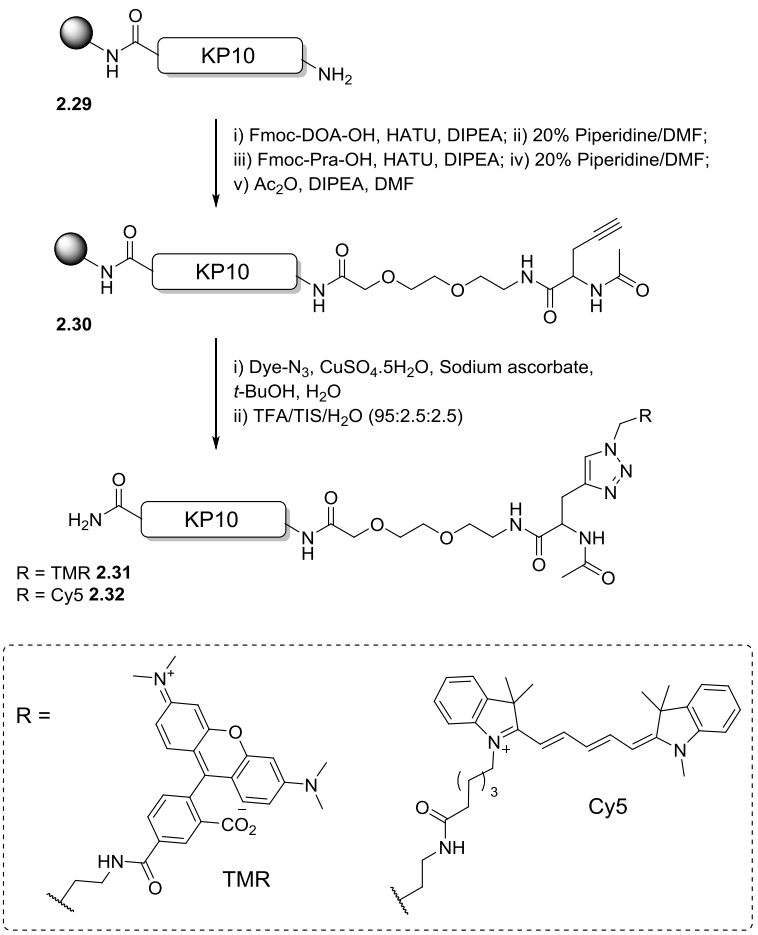
Synthesis of KP-10 proceeded smoothly using this system to give the product in high purity and yield.

2.4.2 Towards a Solution Based Assay for GPCR54/OBOC Libraries

The development of a solution based assay was targeted. This approach would take advantage of the labelling site incorporated into library compounds, which allows their transformation into fluorescent probes. Such an assay would be too expensive for a normal compound archive (high cost of fluorophores), however in this miniaturised format the labelling of thousands of compounds costs relatively little. To assay the labelled compounds, they would be incubated with cell lines expressing GPCR54 and fluorescent imaging used to identify any binding or internalisation. In order to develop such a screen, fluorescent kisspeptin based probes had to be developed.

2.4.2.1 Fluorescently Labelled KP-10

Fluorescent analogues of KP-10 were synthesised for the development of solution based cellular assays. Usually, for SPPS labelling, the alkyne containing Pra amino acid and PEG spacer are introduced at the start of synthesis at the *C*-terminus of library compounds (see chapter 1). In the case of KP-10, however, this would render the compound inactive, as the free *C*-terminal amide is essential for binding. Therefore, the *N*-terminus of KP-10 was used for attachment of the Pra building block required for dye conjugation (Scheme 2-3). Azide functionalised TMR and Cy5 were employed in a copper catalysed 1,3-dipolar cycloaddition reaction to generate KP-10-TMR and KP-10-Cy5.



Scheme 2-3: Synthesis of fluorescently labelled KP-10 by copper catalysed 1,3-dipolar cycloaddition

In order to ensure that the attachment of the fluorophores does not disrupt the activity of the KP-10 analogues, KP-10-Cy5 was biologically assessed in an IP_3 stimulation assay against cell lines stably expressing GPCR54 (Figure 2-18).* This assay involved loading of GPCR54 expressing cells with tritiated *myo*-inotisol, which is incorporated into phosphatidylinositol 4,5-bisphosphate (PIP_2) in the cell.[106] Upon incubation with a ligand with agonistic activity, phospholipase C (PLC) hydrolysis PIP_2 to generate tritiated 1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG) (see section 2.2.6). The use of lithium chloride prevents further degradation of IP_3, which, following extraction of cell contents, is separated from free tritiated inositol using anion exchange chromatography.[107] Liquid scintillation analysis was used to quantify the amount of IP_3 produced (For experimental details see section 8.3.3)

* Biological evaluation of KP-10-Cy5 carried out by Megan Gallant at the Univerity of Capetown, South Africa.
† Western Blot analysis and cloning carried out by Rachel Milne

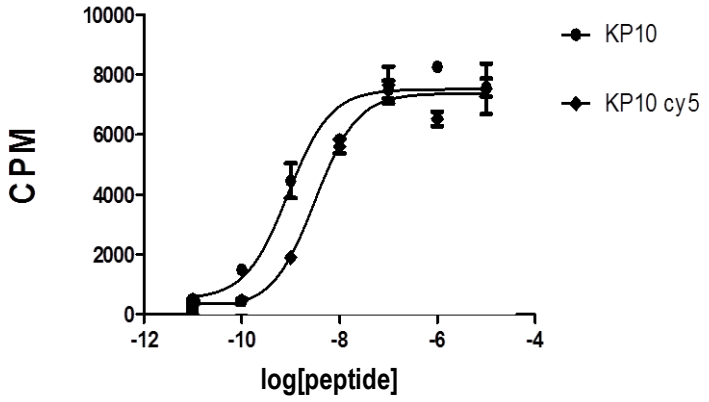


Figure 2-18: Cy5 labelled KP-10 is able to activate GPCR54 signalling in an IP_3 stimulation assay

The EC_{50} measured for KP-10-Cy5 was determined to be 6.6 nM, indicating only a small reduction in activity compared to KP-10 which showed an EC_{50} value of 1.7 nM (Table 2-6).

Entry	Compound No	Compound Name	EC_{50}
1	**2.3**	KP-10	1.7 nM
2	**2.32**	KP-10-Cy5	6.6 nM

Table 2-6: EC_{50} values of KP-10 and labelled analogue KP-10-Cy5

These results indicate that the method of *N*-terminal labelling employed allowed the generation of a fluorescently tagged KP-10 compound which exhibits similar activity to that of the natural ligand. Therefore, this probe was deemed suitable for assay development and the same labelling method utilised in library synthesis.

2.4.2.2 Synthesis of a Non-Natural KP-10 Analogue Library

Building Block Selection

Satisfied that *N*-terminal labelling of KP-10 was not detrimental to its ability to activate GPCR54, this strategy was used to generate a tagged OBOC library. The pentamer library was based around the five critical *C*-terminal amino acids of the kisspeptins, with substitutions including natural and non-natural L-α-amino acids (Figure 2-19).

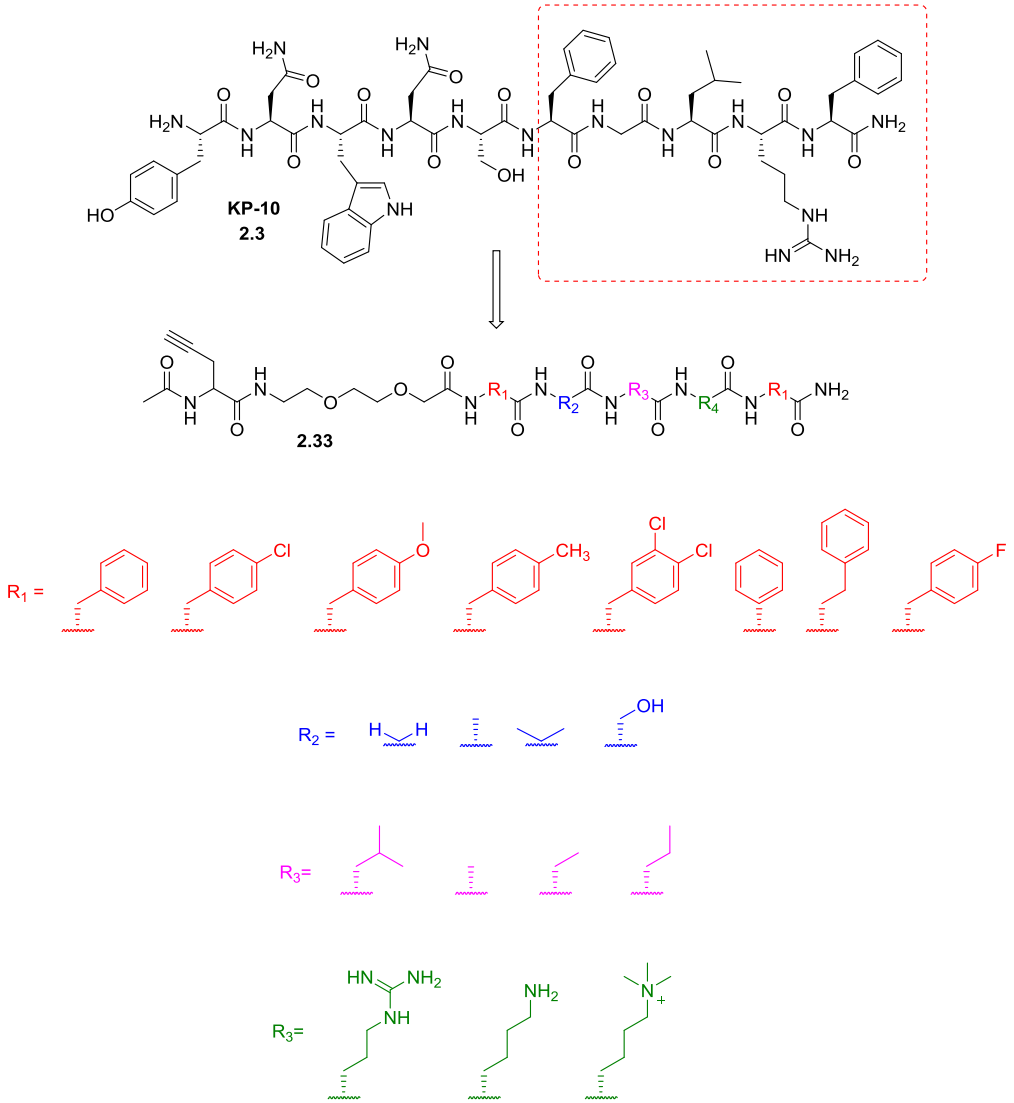


Figure 2-19: An OBOC library was designed based on the five *C*-terminal amino acids of the kisspeptins

Building blocks were selected which showed only modest changes from the natural amino acids, as reports in the literature demonstrated that large changes in the side chains of these amino acids result in depletion of activity. The two phenylalanine (R_1) residues were substituted with only aromatic amino acids, as SAR derived from the literature suggests that

this feature is required for high activity (see section 2.2.8). Halogenated aromatic amino acids 4-chloro-L-phenylalanine, 4-fluoro-L-phenylalanine and 3,4- dichloro-L-phenylalanine were incorporated in the library in order to explore whether halogen bonds could be made with the receptor and whether lipophilic interactions would take place. Methyl and methoxy groups were introduced into the aromatic ring *via* L-Tyrosine and 4-methyl-L-phenylalanine residues to ascertain whether electron donating substituents increased binding affinity. Side chain lengths were also varied using L-phenylglycine and L-homophenylalanine.

To replace glycine (R_2), L-alanine and 2-aminoisobutyric acid were chosen as small lipophilic alternatives. L-Serine was also introduced to assess whether a hydrophilic group was tolerated in this position. Leucine (R_3) was replaced with amino acids with smaller alkyl chains, as it has been shown that larger groups reduce affinity. Arginine substitution (R_4) included L-lysine as a basic alternative. Similar nitrogen containing amino acid trimethyl-L-lysine was also utilised.

Following split and mix synthesis of the library, Doa spacer and Pra were introduced onto the *N*-terminus, which was then acetylated. This library was synthesised on Rink resin so that following fluorescence labelling, compounds could be cleaved to give the *C*-terminally amidated peptides for screening.

Statistical Considerations

The use of these 29 amino acids in 5 combinatorial positions gives a theoretical diversity of (8x4x5x3x8) = 3840 individual compounds. Sublibraries were generated by splitting the resin into 3 portions for coupling at position 4 (Lys) without re-combining. Instead, following coupling, each of the 3 pools of resin was split again into 7 and position 5 (Phe) amino acids coupled. This resulted in 32 sub libraries of 120 compounds each. A total of 640 mg resin was used for synthesis of the library, which yielded 20 mg of resin in each final sub-library.

2.4.2.3 Generation of Cell Lines Stably Expressing GPCR54

As well as fluorescent probes, cell lines expressing GPCR54 were generated for assay development (For experimental details see section 8.2.7)

Hek293 Cell Line Stably Expressing *N*-Terminally FLAG Tagged GPCR54

A cell line stably expressing *N*-terminally FLAG-tagged GPCR54 receptors was produced in HEK293 cells, as this construct has been proven to maintain the characteristics of the

untagged receptor in several assays.[108] The plasmid used was a kind gift from Prof. Bob Millar (University of Capetown, SA) and comprises hGPCR54 cDNA subcloned into a pcDNA3.1 expression vector (Life Technologies) containing an *N*-terminal FLAG epitope (See section 8.2.7.5 for details).

Stable cell lines were produced by transfection of wild type HEK293 cells using transfection reagent Lipofectamine 2000 (Life Technologies), then selecting for stable transfectants by addition of antibiotics (G418) to media. Expression of the tagged receptor was verified by Western blot for FLAG, which revealed a band at ~ 40 kDa (GPCR54 = 43 kDa) in all colonies isolated (Figure 2-20).

In order to investigate the localisation of FLAG-GPCR54 in the cell, immunostaining was carried out (For experimental details see section 8.2.10). Cells were plated on a glass bottomed microtiter plate, suitable for confocal imaging. They were fixed using formaldehyde to freeze all cellular activity and interactions, then permeabilised using Triton to allow molecules to pass through the membrane. Subsequent incubation with anti-FLAG antibody, then an Alexa-555 labelled secondary antibody allowed detection of the FLAG-tagged GPCR54 receptors by fluorescence imaging. DAPI stain was also utilised to allow visualisation of cell nuclei. The cells were assessed on PerkinElmer's Opera® High Content Screening System, with the resulting images exhibiting fluorescence signals throughout the cell upon excitation at 555 nm. This suggests that the cells indeed express FLAG-tagged GPCR54 protein and that it is localised ubiquitously throughout the cell. These findings are consistent with previous reports.[108]

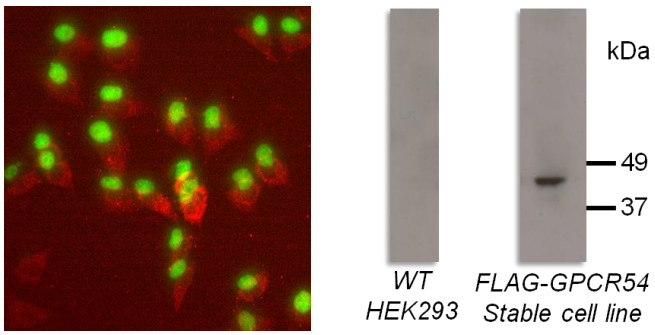


Figure 2-20: Left: Immunostaining of *N*-terminally FLAG tagged GPCR54 expressing Hek293 cell line reveals fluorescence throughput the cell (Cells fixed and permeabilised. Green = DAPI nuclear stain, Red = mouse α-FLAG primary antibody, visualised with Alexa 555 labelled α-mouse). Right: Western blot for FLAG indicates presence of FLAG-GPCR54 in cells.

Hek293 Cell Line Stably Expressing *C*-Terminally eGFP Tagged GPCR54

A second cell line stably expressing *C*-terminally GFP tagged GPCR54 was produced, also in HEK293 cells, using the same methods as above. This construct was also a gift from Prof. Bob Millar, and comprises hGPCR54 cDNA subcloned into a pEGFP-N3 expression vector (Clontech) containing a *C*-terminal eGFP epitope. (See section 8.2.7.5 for details).

Expression of the GFP tagged receptor was verified by fluorescence confocal imaging of the cells using PerkinElmer's Opera® High Content Screening System. Upon excitation at 488 nm, fluorescence was detected throughout the cell, indicating the presence of GFP and therefore presumably GFP tagged GPCR54 (Figure 2-21)

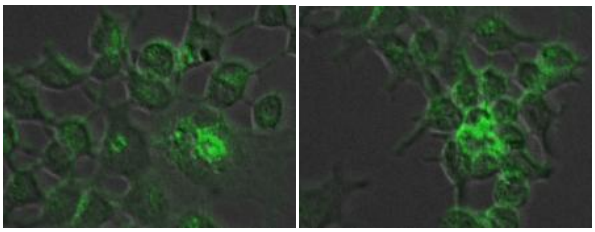
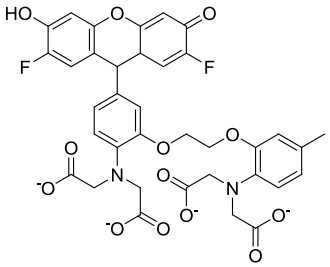


Figure 2-21: Confocal fluorescence image of GPCR54-GFP stable cell line (excitation at 488 nM)

As GPCR54 is a Gq coupled receptor (see section 2.2.6), its activation causes an increase in intracellular calcium as part of its downstream signalling pathway. Therefore, when a ligand is introduced to the cells, the increase in concentration of free calcium is directly linked to the extent of agonism and quantification of this can be used as a functional assay. Calcium sensitive dyes are utilised to allow fluorescence readout of calcium mobilisation. Fluo-4 is a commonly used calcium sensitive dye (Figure 2-22), which responds to Ca^{2+} with an increase in fluorescence intensity when excited at 488 nm.[109]



2.34
Fluo-4
Non-fluorescent, calcium sensitive

Figure 2-22: Calcium sensitive Fluo-4 dye

In order to ascertain whether the GFP-GPCR54 expressing HEK293 cell line was producing functional receptor, a calcium mobilisation experiment was carried out using Fluo-4 Direct

(For experimental details see section 8.2.11). This is a Fluo-4 derivative which is cell permeable and formulated such that, following its addition to cells, does not require to be washed out which minimises the risk of cell detachment. It is also designed to supress background fluorescence. Cells were seeded in a glass bottomed 96 well plate suitable for confocal imaging and loaded with Fluo-4 Direct. Upon addition of KP-10 to a final concentration of 1 μM, wells were imaged using PerkinElmer's Opera® High Content Screening System using 488 nm excitation and fluorescence emission was monitored.

Wild type HEK293 cells showed no increase in fluorescence upon addition of KP-10, however the fluorescence intensity of the GFP-GPCR54 cells increased substantially. This indicates that functional GPCR54 was present in these cells, and therefore upon KP-10 binding, was able to activate its downstream signalling cascade, resulting in the increase in calcium concentration and therefore fluorescence intensity (Figure 2-23). On visual inspection, however, it appeared the viability of the cells had been affected, raising concerns as to whether the over-expression of the receptor was having deleterious effects.

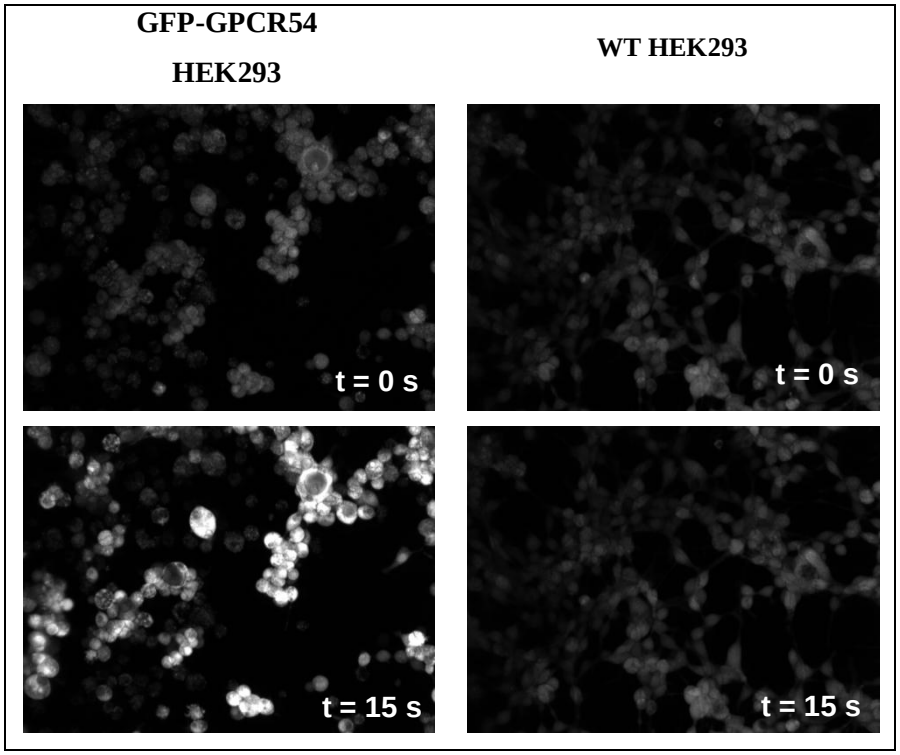


Figure 2-23: Fluorescence intensity increase of GFP-GPCR54 cell line following KP-10 incubation - assayed using Fluo-4AM calcium sensitive dye (WT HEK293 cells shown on right as negative control)

Further assessment of GPCR54 expressing cell lines was required to investigate their interaction with the synthesised fluorescent kisspeptin probe KP-10-Cy5.

2.4.2.4 Fluorescence Activated Cell Sorting

Fluorescence activated cell sorting (FACS) presents a method with which to assess this interaction. This allows the analysis and sorting of individual cells based on their fluorescence, size and granular properties.[110] During FACS, cells are carried in a stream of sheath fluid through a small nozzle, which aligns and separates them in such a way that they pass individually through a detector (Figure 2-24). This detects multiple parameters, such as;

- Forward scatter, which gives information on cell size. This can determine whether cells are large and healthy or small parts of cell debris and allows counting of cells
- Side scatter, which detects light scattered at a 90° angle from the laser beam due to reflection, giving information on the granular properties of cells
- Fluorescence can also be detected, making this method ideal for sorting fluorescent cells from those which don't contain fluorophores, or fluorophores of differing wavelengths.

Sorting of cells is performed by electrical charge. Gaits can be determined which dictate a certain population of cells to be sorted. When a cell which sits within the determined gait passes the detector, its properties are recorded and a charge is applied to the end of the stream as a vibrating nozzle breaks the stream into droplets.[111] This results in a charged fluid droplet containing a single cell, which can be deflected by charged electrodes into a waiting sample tube.

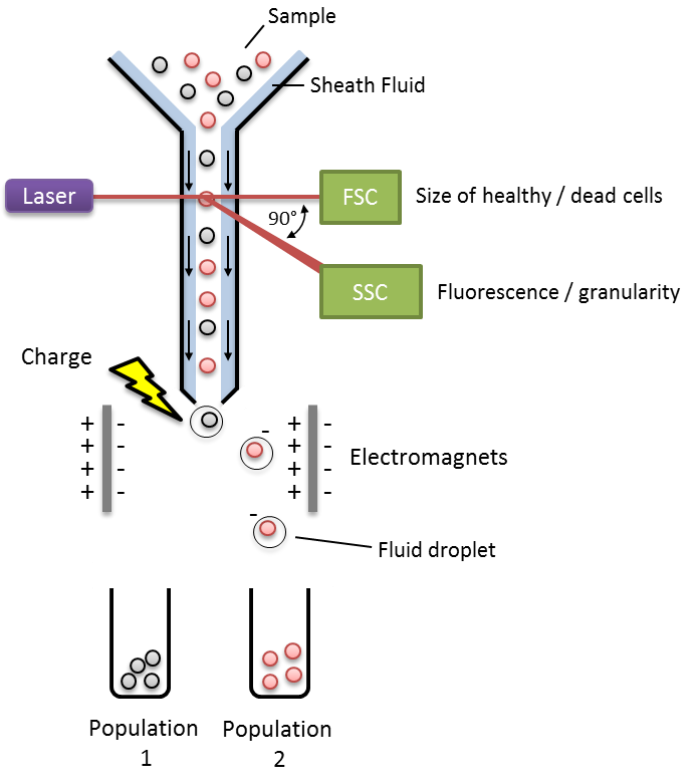


Figure 2-24: Fluorescence activated cell sorting (FACS)

This method was chosen to analyse the result of incubation of fluorescently labelled KP-10 with GPCR54 expressing cells (For experimental details see section 8.2.12). GFP-GPCR54 expressing HEK cells and WT HEK cells were incubated with KP-10-Cy5 (20 uM) in PBS for 2 hours, then washed and analysed by FACS (Figure 2-25). As expected, the negative control WT cells show very little fluorescence in the red (Cy5) or green (GFP) channel, and following incubation with KP-10-Cy5 there is no change. The GFP-GPCR54 cells have a population that shows high fluorescence in the green channel, indicating the presence of GFP. However, upon incubation with KP-10-Cy5, an increase in the fluorescence signal in the red channel is expected corresponding to binding and internalisation of the labelled KP-10. In this case, no increase in fluorescence is observed. As it has been proven that the KP-10-Cy5 ligand used in this experiment is able to activate GPCR54, we can conclude that the GFP-GPCR54 material present in the cell is either not transported to the cell surface, or is misfolded and therefore inactive.

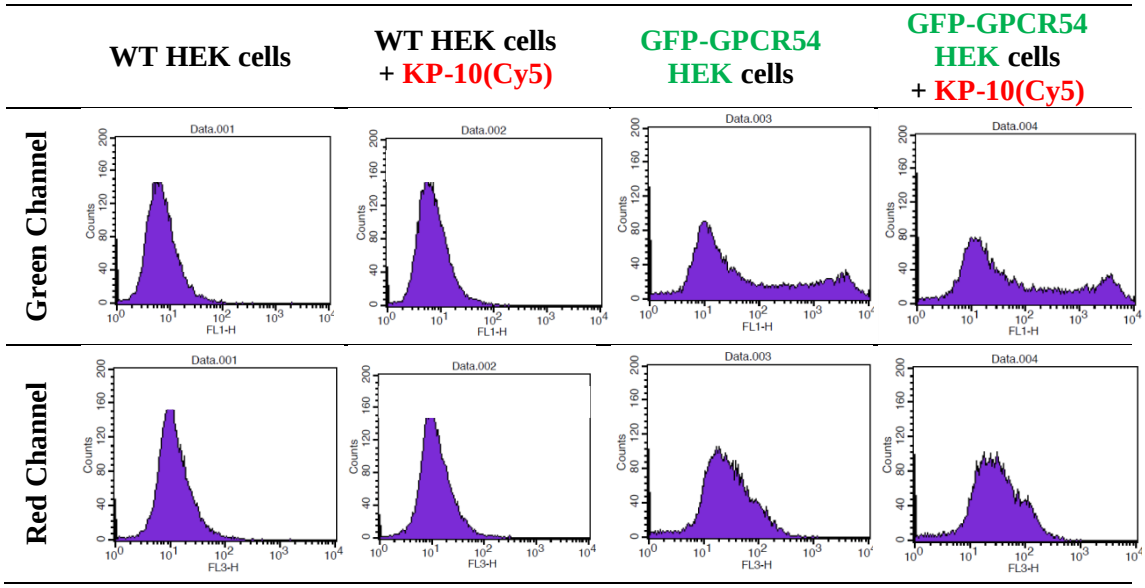


Figure 2-25: FACS analysis of WT and GPCR54 expressing HEK cells before and after incubation with KP-10-Cy5. No binding of the ligand is observed.

In order to progress the project, we focused our efforts on the development of a bead-based imaging screen instead.

2.4.3 Towards an On-Bead Assay for GPCR54/OBOC libraries

2.4.3.1 Reversed KP-10 on Bead

In order to develop an on bead screening method for GPCR54, KP-10 bearing resin was required as a tool compound. However, to present KP-10 on solid support in a manner that it would be able to bind GPCR54, it is essential that the *C*-terminal amide is free and orientated away from the bead. This is the opposite orientation to which peptides are synthesised during SPPS, where the *C*-terminus is attached to the resin and the *N*-terminus is presented outwards (Figure 2-26).

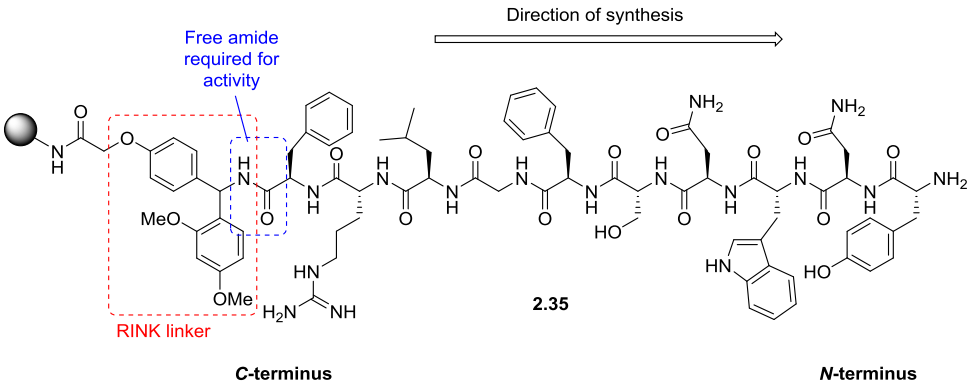


Figure 2-26: Kisspeptin-10 synthesised from *C*→*N*-terminus on solid support with C-terminal amide attached to resin

We therefore required a 'reversed' KP-10 on bead, and anticipated that one possible method to achieve this would be to synthesise KP-10 as shown above, then label the *N*-terminus with a functionality which would allow orthogonal linkage to another surface. Following deprotection of side chains and cleavage from the resin, the labelled KP-10 could be reattached to another resin, this time by the *N*-terminus, allowing the KP-10 compound to be reversed with the *C*-terminus presented outwards (Figure 2-27).

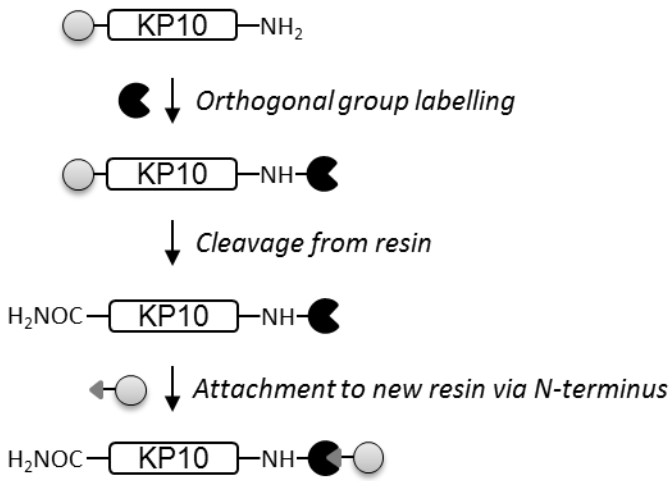
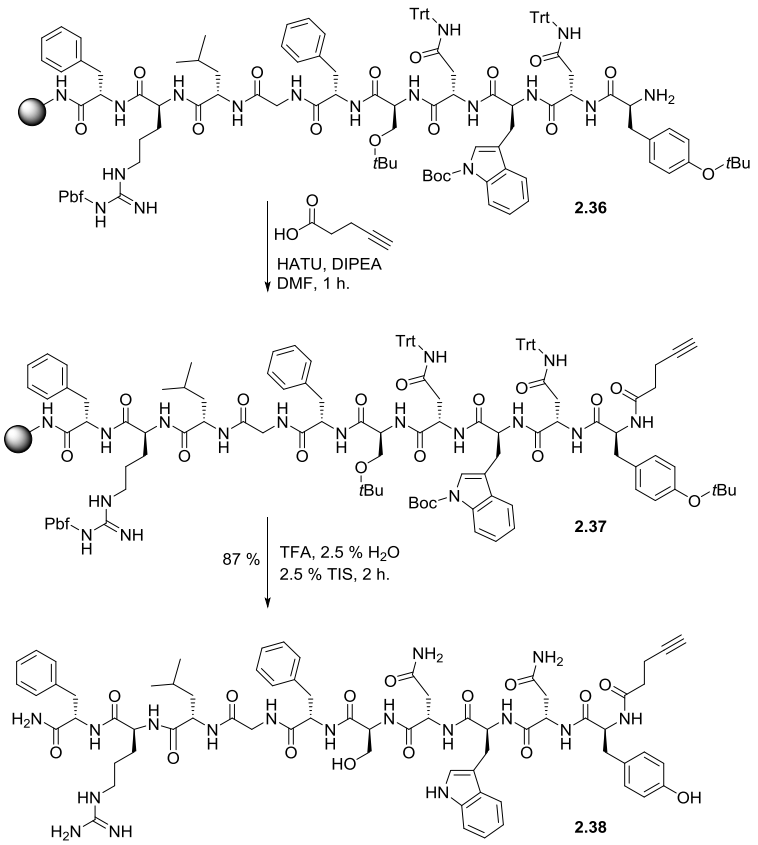


Figure 2-27: Reversing KP-10 to allow *C*-terminal amide to be presented away from resin for binding

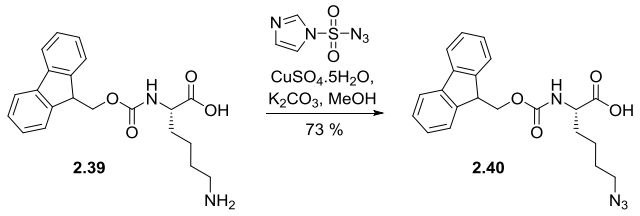
In order to achieve this, Huisgen "click" 1,3-dipolar cycloaddition was employed, as it has proven to be high yielding and orthogonal when used for fluorescent labelling on library compounds. Synthesis of KP-10 was again carried out on TentaGel RINK resin, then 4-pentynoic acid was coupled to the free *N*-terminus to provide alkyne functionalised KP-10.

The peptide was cleaved from the resin using acidic conditions, which simultaneously removed side chain protecting groups (Scheme 2-4).



Scheme 2-4: Synthesis of KP-10-alkyne

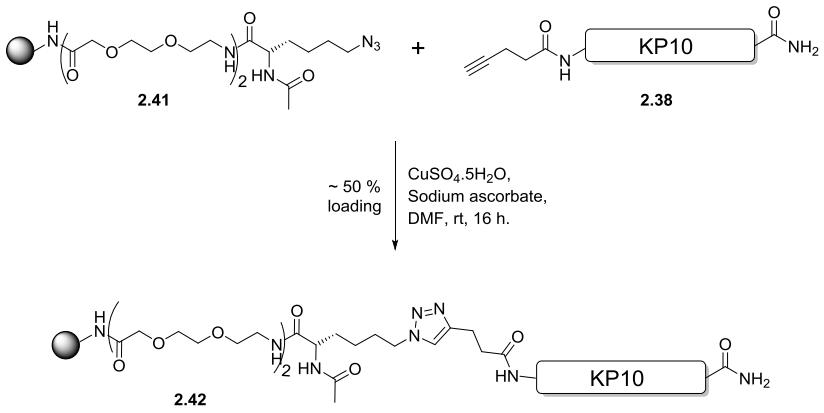
Azide functionalised resin was also prepared. For this purpose, Fmoc-Lys-OH was converted into Fmoc-Lys(N_3)-OH in good yield using diazotransfer reagent imidazole-1-sulfonyl azide hydrochloride (Scheme 2-5).[112]



Scheme 2-5: Synthesis of Fmoc-Lys(N_3)-OH

The azide functionalised amino acid was coupled onto resin following two monomers of PEG spacer 8-amino-3,6-dioxaoctanoic acid (Scheme 2-6). Acetylation of the deprotected amino acid gave resin **2.41**. This was then treated with 3 equivalents KP-10 alkyne **2.38**, along with copper sulfate and sodium ascorbate in DMF overnight to give the final reverse

KP-10 presenting resin **2.42**. Ten equivalents of copper catalyst were used for this reaction, as using stoichiometric amounts resulted in no product formation. DMF was employed as the solvent of choice as KP-10 proved insoluble in solvents such as *tert*-butanol, water and acetonitrile. Analysis of a small sample of the resin indicated that only around half of azide had been loaded with KP-10-alkyne, however this was considered an adequate amount for binding of GPCR54.



Scheme 2-6: Generation of reversed KP-10 using click chemistry

With resin bound KP-10 tool compounds in hand, we focused on generating GPCR54 in a format suitable for on-bead screening.

2.4.3.2 Whole Cells on KP-10 Beads

We envisioned that the simplest on-bead experiment would consist of whole cells binding to KP-10 functionalised resin. Therefore, cells expressing GFP-GPCR54 were incubated with the kisspeptin beads and the results imaged using PerkinElmer's Opera® High Content Screening System at 488 nm (Figure 2-28) (For experimental details, see section 8.2.13.1).

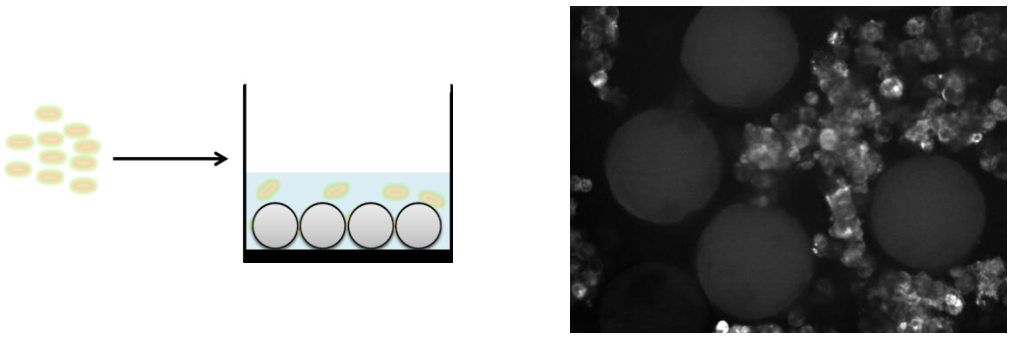


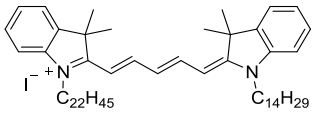
Figure 2-28: GFP-GPCR54 cells incubated with KP-10 resin

Although existing reports document whole cells binding resin presenting ligands,[43] in our experiments the GPCR54 expressing cells appeared to sediment and showed no binding

behaviour towards the beads. The confocal images clearly show fluorescent material throughout the cells, indicating that GPCR54-GFP is indeed being expressed. This led us to conclude that not enough receptors were presented at the cell surface for a strong enough interaction to take place to hold the weight of the rest of the cell in place. Therefore, a different approach was sought.

2.4.3.3 Membrane Fractions

An alternative approach taken was to use fractions of membranes from cells expressing GPCR54 (Figure 2 30) (For experimental details, see section 8.2.13.2). To be sure that receptor was present on the membrane fractions, human GPCR54 membrane preparations, in CHO-K1 cells, were purchased (Perkin Elmer). In order to be able to visualise interactions, the membranes were stained with CellVue Maroon (Tebu Bio). This is a far-red fluorescent cell labelling agent, excitable at 647 nm, which contains hydrophobic tail groups that selectively intercalate with membrane lipids (Figure 2 29).[113] This functionality allows the stain to localise in cell membranes, which is why it was utilised in this experiment to label membrane fractions.



2.43

Figure 2-29: Structure of CellVue Maroon membrane stain

Following staining, the membrane fractions were spun down and the resultant pellet washed with PBS three times. The material was resuspended in 1 x BSA in PBS and added to wells containing KP-10 resin and a negative control well containing blank (NH_2) beads and the samples incubated overnight at 4 °C with shaking.

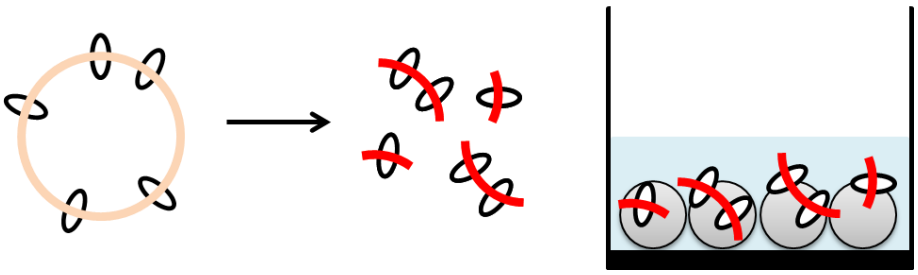


Figure 2-30: Cell membranes containing GPCR54 were stained and incubated with KP-10 resin

Although faint rings were detected around the KP-10 resin upon confocal imaging, indicating binding of a fluorescent substance, the same pattern was seen on the negative control resin which contained amino groups (Figure 2-31). Bright dots are also visible in the background on the images, suggesting that the membrane fractions formed aggregates.

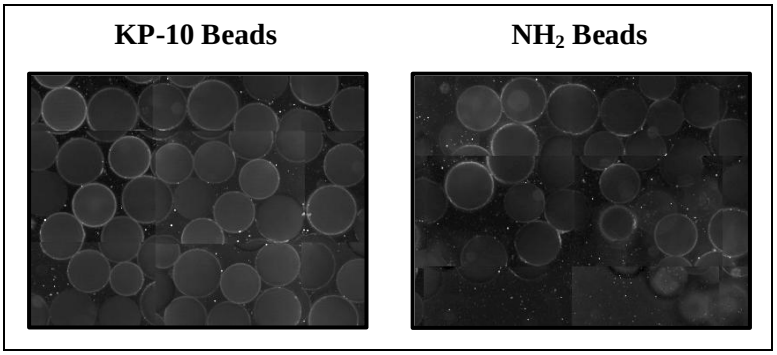


Figure 2-31: KP-10 beads (left) and blank beads (right) following incubation with GPCR54 membrane fractions stained with CellVue Maroon (Tebu Bio) – imaged in 640 nm channel.

This non specific binding could be due to free dye which was not properly washed out. In this case, the problem could be circumvented by detecting the membrane fractions using a fluorescently labelled anti-GPCR54 antibody. This would allow non-labelled membrane fractions to be incubated with the KP-10 resin then non-bound fractions washed away. An antibody against the *C*-terminus of GPCR54 could then be used to detect membranes bound to the resin either in a labelled form, or by following this step with a labelled secondary antibody. This approach would circumvent non-specific interaction of membrane staining dyes with KP-10 resin.

2.4.3.4 MembranePro™ - Lipoparticles

The MembranePro™ kit (LifeTechnologies) is an expression system which allows the production of membrane proteins (e.g. GPCRs) in an aqueous soluble format by taking advantage of virus like particles (VLPs). These are subviral particles which self-assemble from virus-derived core structural proteins (e.g. HIV-1 gag), although as there is no viral genome present, the resultant particles are non-transducing, they are not infectious and do not replicate.[114]

In the MembranePro™ expression system, the lentiviral gag protein is utilised, which is co-transfected into cells along with cDNA of the membrane protein of interest (Figure 2-32). The gag protein travels to the cell membrane and generates buds underneath lipid rafts, where overexpressed GPCRs are likely to be present as lipid rafts play an active role in regulating the sorting and conformational state of membrane proteins. The VLP buds from

the cell and in doing so, captures receptors localised in these microdomains in their natural environment. The lipoparticles separate from the cell and are released into the cell culture medium, which can then be decanted off and spun down to retrieve the lipoparticles.[114]

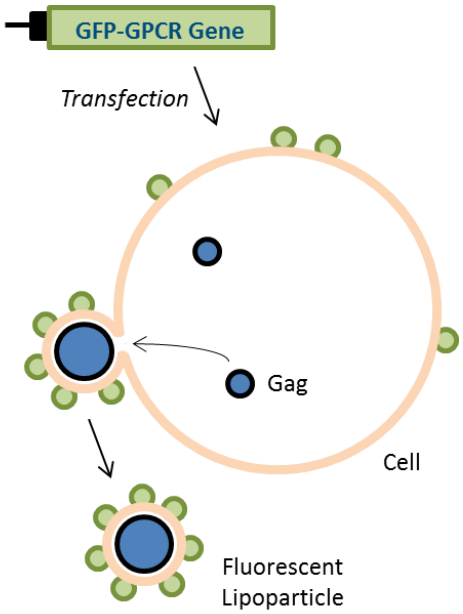


Figure 2-32: MembraneMax™ utilises gag protein to allow the production of lipoparticles displaying membrane proteins of choice.[114]

This expression method appeared viable for the preparation of GPCR54 in a format suitable for CONA screening, as the lipoparticles would be soluble and disperse through buffer, unlike whole cells. Non-fluorescent lipoparticles have also been proven to be compatible with fluorescence polarisation, and therefore could be used for solution based confirmation of hits as well. In a report by Jones *et al.*,[115] CXCR4 lipoparticles were generated and assessed in FP experiments with oregon green labelled T-22, a known peptidic ligand with a reported IC_{50} of 48 – 71 nM.[116] Assessment of the interaction between the CXCR4 containing lipoparticles and labelled T-22 ligand using FP gave an EC_{50} of 15 nM, consistent with reported values.

Co-transfection of 293FT cells with GFP-GPCR54 DNA and the MembraneMax™ reagent was carried out. After two days, the cell culture medium was harvested and a precipitation mix was added. Centrifugation of the sample was carried out and the medium discarded. The resultant pellet was resuspended in buffer and analysed by incubation with KP-10 resin (Figure 2-33) (For experimental details, see section 8.2.13.3).

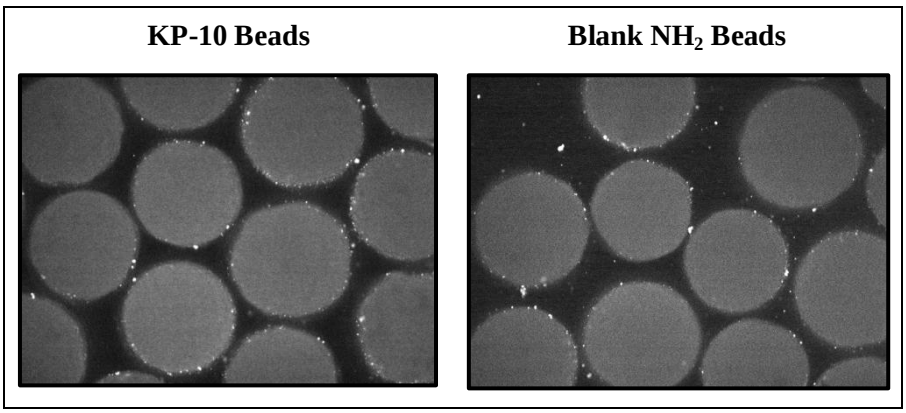


Figure 2-33: KP-10 resin incubated with GFP-GPCR54 lipoparticles generated using MembraneMax™ kit. Although some fluorescent species can be seen, binding is apparent in negative control and can therefore be considered to be non-specific.

A CONA screen shows some fluorescent material bound to the resin, however the material is also seen binding to blank NH_2 resin. There is also very little of it, suggesting production is not very efficient. This procedure was repeated multiple times with similar results seen, therefore a new approach was investigated.

2.4.3.5 MembraneMax™ – Nanolipoprotein Particles

The cell-free production of functional GPCRs is a rapidly advancing field. Such an approach circumvents issues associated with high-level production of GPCRs in cells, such as cytotoxicity and ineffective transportation and insertion into the plasma membrane.[117] GPCRs can be generated in a cell free manner and subsequently stabilised with detergents, however detergents do not mimic the cell membrane environment very well. They are also destabilising and can inactivate proteins over time.[118] The insertion of membrane proteins into unilamelar liposomes has been demonstrated, which provides a more membrane-like environment,[119] however samples tend to be very heterogeneous and insertion conditions must be established for each individual protein.[120]

An alternative method was developed by Kudlicki *et al.*, who aimed to develop a cost effective, heterologous cell-free expression system for membrane proteins.[120-122] This approach involves translation of membrane proteins in the presence of nanometer-sized planar membranes made of nanolipoprotein (NLP), a phospholipid membrane bilayer surrounded by scaffold protein. The GPCRs insert into the nanolipoproteins to form nanolipoprotein particles, which are soluble and contain the folded and functional receptors in a near-native environment. This method has been used to generate soluble Adrenergic Receptor α_2 (ADRB2), Neurokinin 1 Receptor (NK1R) and the Dopamine Receptor D_1 (DRD1), all of which showed affinity for their specific ligands.[123]

It was envisaged that the use of NLPs would also allow us to circumvent problems of fast internalisation and lack of presentation of the receptor on the cell surface. We therefore investigated the commercial MembraneMax™ kit (Life Technologies), which provides reagents required for the production NLPs using *E. coli* extracts (Figure 2-34). Importantly, NLPs have been shown to be compatible with FCS analysis,[124] which could potentially allow them to be used not only for on-bead primary screens, but also in solution based binding experiments, in the same manner that soluble proteins are. NLPs therefore represented the ideal solution to the problem of how to screen TOBOC libraries against GPCRs.

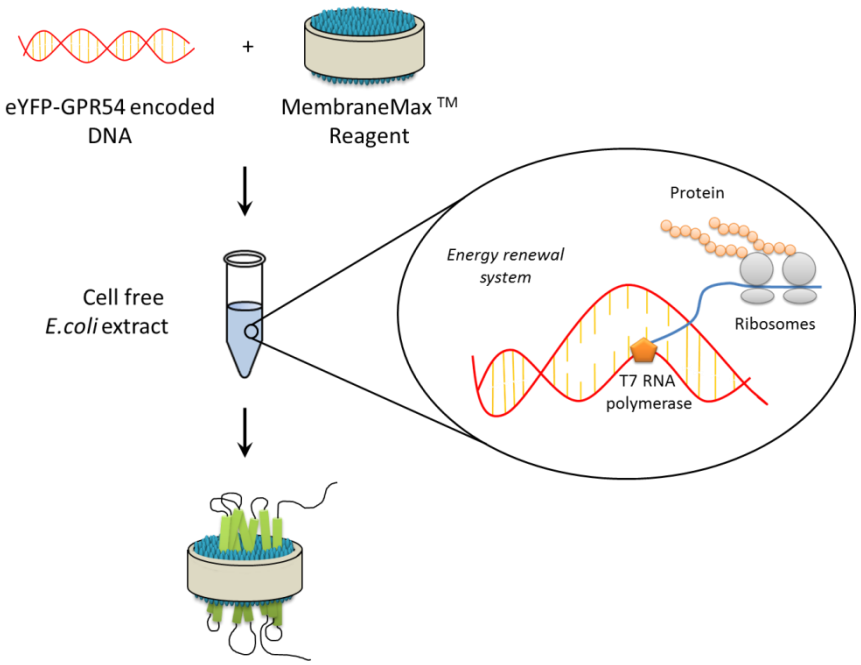


Figure 2-34: Production of soluble receptors in NLPs using MembraneMax™ [114]

For protein expression using MembraneMax™, GPCR54-eYFP was cloned into a vector containing a *C*-terminal 6xHis tag which could be used for NLP purification. The expression was carried out as per kit instructions (For experimental details, see section 8.2.13.4). Following expression, western blot analysis of the mixture was carried out, with bands being visualised using an anti-His antibody (Figure 2-35).† The presence of a band at ~60 kDa was encouraging as this correlates well with the combined mass of GPCR54 plus eYFP (43 kDa + 26 = 67 kDa). Therefore, the crude material was analysed by FCS‡ in order to establish whether any fluorescent species were present.

† Western Blot analysis and cloning carried out by Rachel Milne
‡ FCS analysis carried out by Dr. Nhan Pham

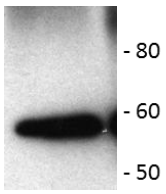


Figure 2-35: Western blot of crude reaction mixture following MembraneMax™ reaction with His-GPCR54-eYFP cDNA. Detection using anti-his antibody.

FCS analysis indicated the presence of fluorescent species in the reaction mixture, but multiple species were present. In order to ascertain whether any functional GPCR54 was present, the crude reaction mixture was incubated with the kisspeptin resin. Following 3 hours incubation, the mixture was imaged on PerkinElmer's Opera® High Content Screening System using 488 nm excitation, however no fluorescent rings were evident on the KP-10 resin (Figure 2-36). Although some debris from the reaction was obscuring the view of the resin, it was clear that no binding to the resin had occurred.

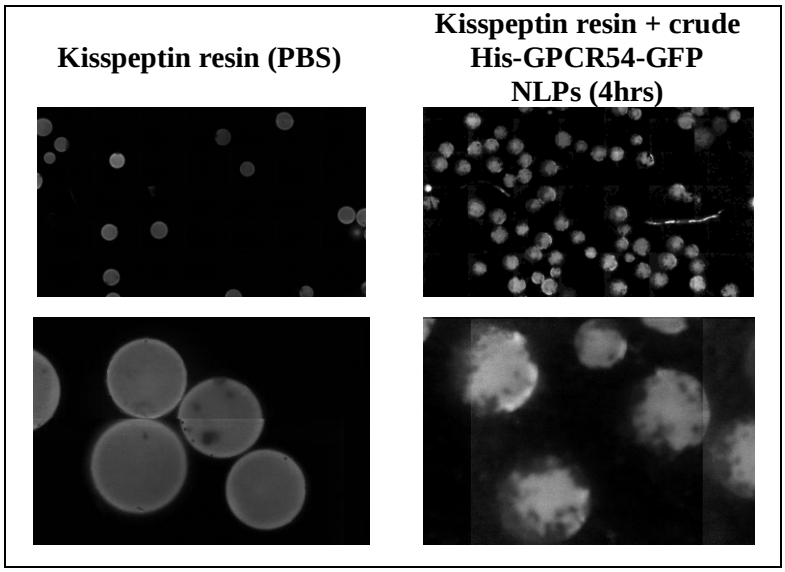


Figure 2-36: Kisspeptin resin in PBS (negative control) and incubated with crude MembraneMax™ NLP reaction mixture. No binding was detected.

Due to time constraints, a deeper investigation into the products of this reaction was not possible. Despite the lack of functional GPCR material generated in this case, in my opinion the use of nanolipoparticles presents an elegant way to screen GPCRs while still taking advantage of the CONA and PS/PS labelling technology developed in the Auer lab. Fluorescently labelled NLPs could allow an initial on-bead screen to be carried out, with single molecule confirmation of hits being carried out using non-labelled NLPs and fluorescently labelled ligands.

2.4.4 GPCR54 *In-Silico* Homology Model

GPCRs are integral membrane proteins, which makes experimental determination of their tertiary structure highly challenging.[125] The first solved crystal structure of a GPCR was that of rhodopsin in 2000 by Palczewski and co-workers.[126] Since then, a handful of other GPCR structures have been determined experimentally,[127] including the β2 adrenergic receptor,[128] the A_{2A} adenosine receptor and the CXCR4 chemokine receptor,[129] amongst others. This collection, however, represents only a small proportion of all GPCRs.

Homology modelling consists of the prediction of the atomistic three-dimensional (3D) structure of a protein sequence based in its alignment to other proteins of which structure is known[130]. This technique has been used extensively for the prediction of GPCR structures. A prerequisite of this process is that in order to generate a useful model, the target protein must share a detectable amount of sequence homology with the template protein. Within a family of proteins, the 3D structure is more conserved than the sequence, therefore detectable sequence similarity allows the assumption of structural similarity. Indeed, the use of multiple templates from related protein structures can lead to a more accurate hybrid model.[131] The process of generating a homology model generally consists of four steps:[132]

- Identify related proteins of known 3D structure to act as template
- Align target sequence with template structure
- Build model for target by substituting amino acids of template protein for those of target protein according to alignment. Energy minimisation is then carried out and conformational aspects checked.
- Evaluate the model with e.g. protein mutagenesis data, ligand binding data.

Homology models have been successfully utilised for the prediction of binding modes of receptor ligands and to rationalise experimentally derived SAR data.[133]

2.4.4.1 Aims

A collaboration was initiated between Steven Shave (University of Edinburgh), Xavier Deupi (PSI, Switzerland) and Agnieszka Bronowska (Heidelberg Institute for Theoretical Studies, Germany) with the aim of generating a homology model of GPCR54, which would allow the prediction of the binding mode of KP-10. Several techniques including molecular modelling, molecular dynamics (MD) simulations and molecular docking calculations were combined in order to achieve three main goals:

1. Obtain an atomistic 3D model of GPCR54
2. Deduce the binding site of GPCR54
3. Predict key features of the kisspeptin-GPCR54 interaction

With this information, it was hoped that insights could be gained into the kisspeptin-GPCR54 interactions which would allow assistance in the design of focused libraries containing high affinity ligands targeting the kisspeptin receptor. Model generation based on CXCR4 was performed by Xavier Deupi (PSI, Switzerland) and Steven Shave (University of Edinburgh). Methods used are detailed in experimental section 8.2.14. The model was refined by Agnieszka Bronowska (HITS, Germany) based on multiple receptors, described below. Results reported below were obtained by Dr. Agnieszka Bronowska (HITS, Germany). Compounds used for docking studies were suggested by myself and provided to Dr. Agnieszka Bronowska as a blind study.

2.4.4.2 Homology Modelling of GPCR54

Since the atomistic structure of GPCR54 receptor remains unknown, a model had to be constructed using the homology modelling technique. A multiple template strategy was applied in order to refine the GPCR54 model based on CXCR4. The delta (PDB codes 4EJ4 and 4N6H) and kappa (4DJH) opioid receptors and the nociceptin/FQ receptor (4EA3) were selected as templates. These receptors share the highest sequence homology with GPCR54 (36% with kappa-opioid, 40% with delta-opioid, and 39% with FQ receptors, respectively) among those family A GPCRs whose structures are available (for instance, bovine rhodopsin and GPCR54 share 21% sequence homology). This degree of homology allows the building of a reliable model. It should be noted that 30% homology or more is considered sufficient. Since structure is more conserved than sequence, it is possible to carry out sequence alignment through a structure-based alignment method. The sequence alignment of three templates and GPCR54 is shown in Figure 2-37.

Figure 2-37: Sequence alignment of delta, kappa and opiod receptors and GPCR54 (KiSS1R)

The overall 3D model of the kisspeptin10-GPCR54 complex is shown in Figure 2-38. The KP-10 peptide occupies the central cavity of the receptor, interacting mostly with residues from transmembrane helices TMH3, TMH5, TMH6, and TMH7, and to some extent, TMH2 and TMH1. This binding mode is consistent with those observed for many family A receptors, including opioid receptors, which were the template for this model. The deepest located fragment is the *C*-terminus of KP-10, with the terminal amide group interacting with critical Gln122 and Gln123 residues in TMH3, which can be regarded as a "switch" between

inactive and active conformation. Deletion of the *C*-terminal amide group leads to dramatic change in electrostatic interactions and a rearrangement of H-bonding network in this area, therefore our results suggest that the terminal amide group is required for maintaining native-like contacts and to bind ligands with high affinity.

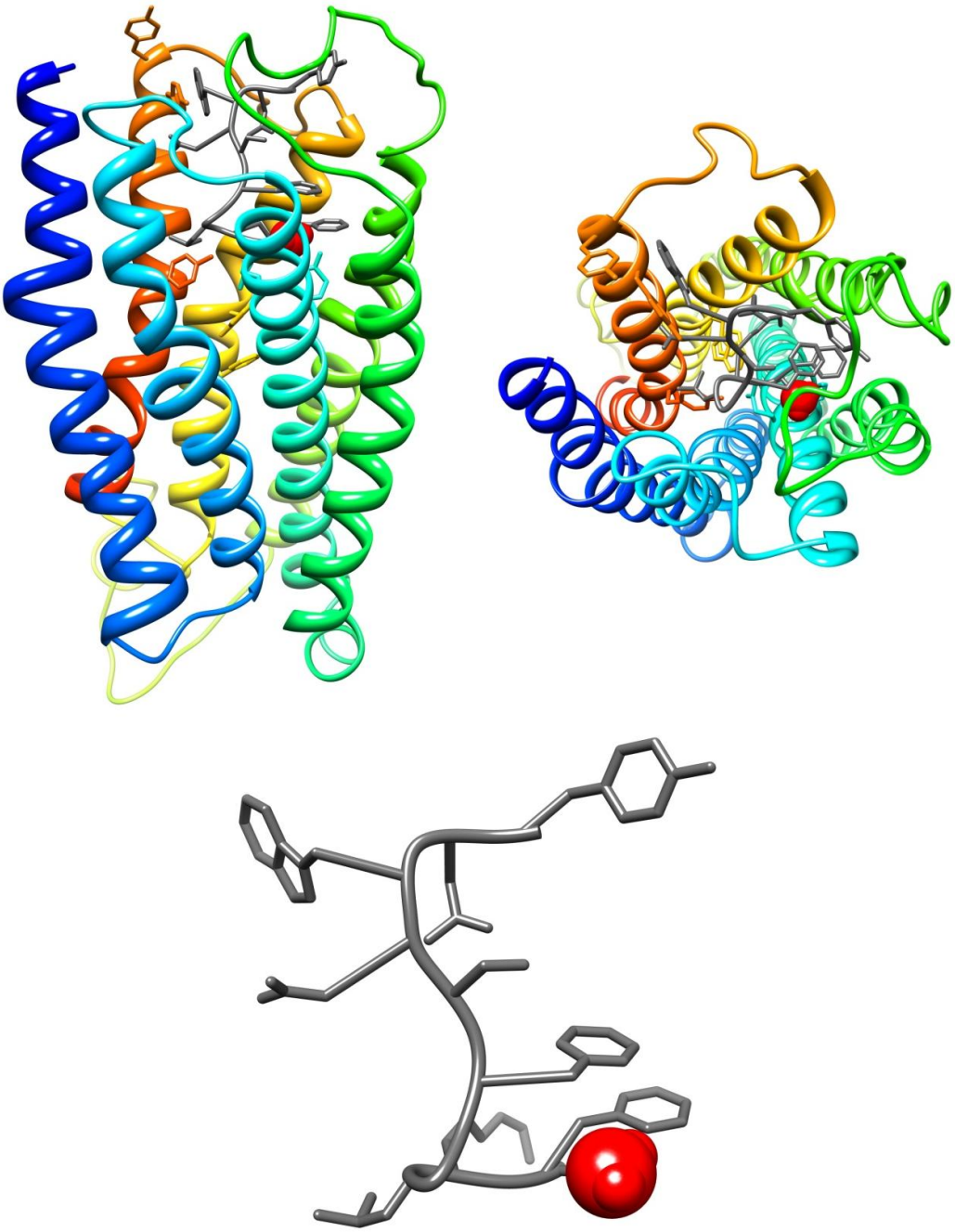


Figure 2-38: 3D model of kisspeptin-10-GPCR54 complex (KP-10 shown in grey, amidated *C*-terminus shown as red sphere)

The key residues predicted to be involved in interactions with kisspeptin-10, including the Gln122 and Gln123 "switch", are highlighted in Figure 2-38. The full list of residues involved in ligand-receptor interactions is shown in Table 2-7.

Residue	TMH	Ligand Residue in Contact	Residue	TMH	Ligand Residue in Contact
Phe48	1	Leu8	Glu201	ECL2	Tyr1
Tyr103	2	unspecific	Arg188	ECL2	Asn2, Trp3
Thr99	2	Leu8	Trp276	6	Arg9
Leu102	2	unspecific	Gln280	6	Arg9
Trp108	ECL1	Ser5	Ile279	6	Arg9
Gln122	3	Tyr10, NH2	Leu283	6	Arg9
Gln123	3	Tyr10, NH2	Gln286	6	Asn2
Asn119	3	Phe6, NH2	Arg202	5	Asn2
Phe204	5	Tyr10	Arg297	ECL3	Asn2, Trp3
Val177	4	Tyr10	Tyr299	7	Trp3
Asn208	5	Arg9, Tyr10	Tyr302	7	Trp3
Tyr190	ECL2	Tyr1	Thr306	7	Asn4
Ser192	ECL2	Tyr1	His309	7	Asn4
Glu193	ECL2	Tyr1	Cys310	7	Leu8
Arg198	ECL2	Tyr1	Tyr313	7	Leu8

Table 2-7: Description of interactions/contacts observed in KP-10–GPCR54 complex

The ultimate validation of an obtained GPCR54 model can only be derived from a direct comparison with its experimentally solved crystal or NMR structure. However, such a comparison is only possible when the experimental structure of the receptor becomes available, which for GPCR54 it currently is not. Thus, the only possible way to validate the model is to test the correlation between predictions generated on its basis and experimental results available. Structure-activity relationship data is available for KP-10 analogues[85, 88] and this was used for the validation of the receptor 3D model. Mutagenesis data is also

available for GPCR54 and receptor model validation based on this is detailed in the experimental section.

2.4.4.3 Reproduction of Experimental Energies of Kisspeptin-10 Mutants and Kisspeptin-Like Pentapeptides

Reported KP-10 derivatives for which experimental binding data is published were chosen for docking in the GPCR54 model. The derivatives chosen were F10H, F10W, S5A, G7A, G7P and N2A (Figure 2-39). For ligand simulation and docking procedures, see materials and methods.

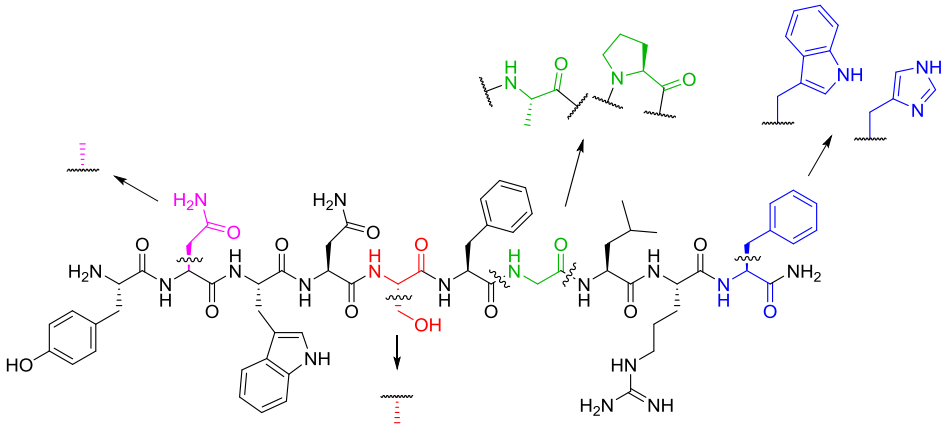


Figure 2-39: KP-10 and mutants were chosen for docking in GPCR54 homology model

The results are shown in Table 2-8 for the energetics evaluation for wild-type kisspeptin-10 and the six derivatives shown above. The interaction energy (Eint) and the ligand's internal energy (EL) are shown separately. These values are combined to give the total theoretical energy for ligand binding (Etotal). It should be noted that the G7P variant failed to adapt the bioactive-like conformation, which causes steric hindrance and makes binding unfavourable. Experimentally determined K_d values for the same ligands are also shown in Table 2-8,[85] which have also been converted into kJ/mol using the equation $\Delta G° = -RT\ln K_a$ (where $K_a = 1/K_i$, T = 298 K and R = 8.314 J/mol/K), to allow comparison between the theoretical and experimental total energy of binding.

Entry	KP-10 mutant	Theoretical			Experimental	
		EintR	EL	Etotal	Binding Energy[A]	K_i[85]
			(kJ/mol)		*(kJ/mol)*	*(nM)*
1	WT KP-10	-2252.3	2193.5	**-58.8**	**-50.7**	1.3
2	S5A	-2376.3	2304.4	**-67.7**	**-51.9**	0.8
3	F10W	-2404.6	2339.2	**-65.3**	**-50.2**	1.6
4	G7A	-2366.9	2319.5	**-48.2**	**-47.4**	5.0
5	N2A	-1758.0	1717.9	**-40.1**	**-43.8**	21.1
6	F10H	-2364.2	2333.3	**-30.9**	**-38.7**	161.8
7	G7P	-2092.2	2245.4	**+153.1**	-	-

Table 2-8: Binding energies from GPCR54 homology model and experimental data for wild-type kisspeptin-10 and six mutants. [A]Experimenally determined binding energy calculated using the equation $\Delta G° = -RT\ln k_a$ (where $k_a = 1/k_i$, T = 298 K and R = 8.314 J/mol/K)

It can be seen from the data shown in Table 2-8 that the total energy values from docking experiments correlate with those derived experimentally, with mostly successful ranking of the mutants in order of binding affinities. Experimentally, mutants S5A and F10W exhibit binding similar to that of KP-10 (entries 1, 2, and 3). The docking studies closely match that observation, with the binding energy for WT KP-10 (-58.8 kJ/mol) very close to that of the mutants (-67.7 kJ/mol and -65.3 kJ/mol for S5A and F10W respectively). Mutants G7A and N2A (entries 4 and 5) show slightly lower affinities for GPCR54 experimentally, which correlates well with the total binding energies of -48.2 kJ/mol and -40.1 kJ/mol observed *in silico*. The F10H mutant shows a binding affinity of 161.8 nM experimentally, which is a large drop in potency compared to KP-10 (1.3 nM). The docking studies replicated that finding, with a high total binding energy of -30.9 kJ/mol for this mutant compared to the wild type peptide. Finally, the G7P mutation depletes all binding in an experimental setup. After docking in the GPCR54 model, a total energy of +153.1 kJ/mol is obtained, once again mimicking the experimental data and providing confidence in the binding modes and interactions proposed. The graph below (Figure 2-40) shows these experimentally derived energy values versus binding energies calculated *in silico*. It can be seen that the values correlate nicely, although the experimental value for entry 7 is omitted in order to maintain a reasonable scale.

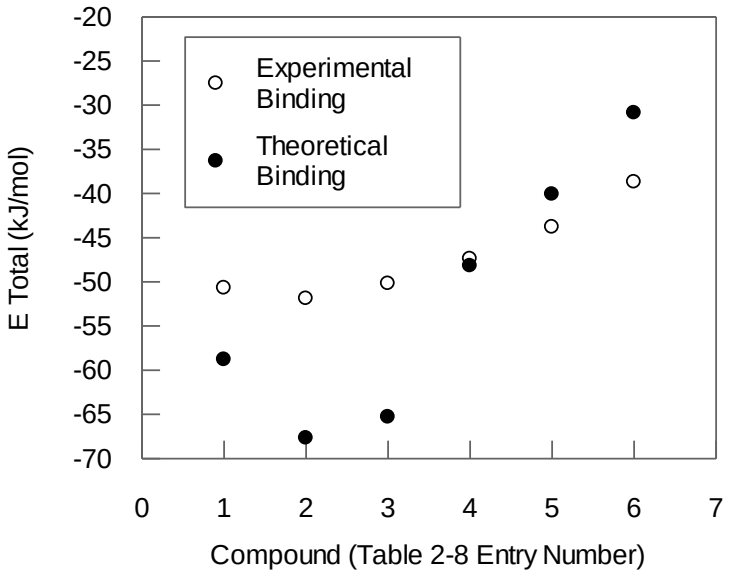


Figure 2-40: Theoretical binding energies vs. experimentally derived binding energies for KP-10 derivatives

A selection of downsized kisspeptin analogues, reported by Tomita *et al.*[90] were also used in docking studies (Figure 2-41).

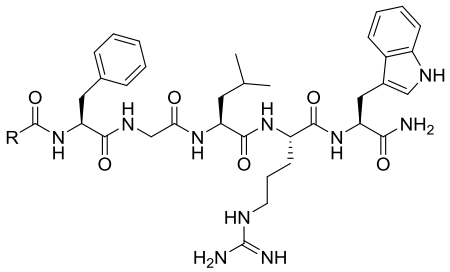


Figure 2-41: Downsized kisspeptin analogues chosen for binding studies in GPCR54 model

Again, contributions of interaction energy (Eint) and ligands internal energy (EL) were combined in order to give a total energy score for each compound. The results are shown in Table 2-9 along with experimentally determined EC_{50} values for the same ligands.[90] The experimentally determined values have been converted into kJ/mol using the equation $\Delta G° = -RTlnK_a$ (where $K_a = 1/K_i$, T = 298 K and R = 8.314 J/mol/K), to allow comparison between the theoretical and experimental total energy of binding. Although this is only an approximation to the actual affinity of the ligands, as EC_{50} values are not equal to K_i values in general, the conversion was made in order to allow a useful comparison to be seen between the two data sets.

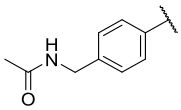
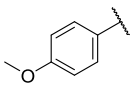
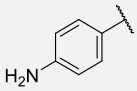
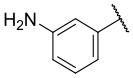
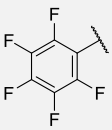
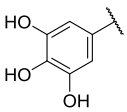
Entry	R	Theoretical			Experimental	
		EintR	EL	Etotal	Binding Energy [A]	EC_{50} [90]
			(kJ/mol)		*(kJ/mol)*	*(nM)*
1	KP-10 WT	-2252.33	2193.55	-58.77	-52.78	0.56
2		-1378.71	1249.76	-128.7	-51.90	0.80
3		-1423.55	1302.22	-121.31	-51.37	0.99
4		-1160.58	1068.10	-92.42	-51.37	0.99
5		-1188.55	1128.18	-60.36	-50.89	1.20
6		-1163.93	1119.13	-44.9	-50.18	1.60
7		-1173.23	1130.65	-42.55	-50.03	1.70
8		-1130.02	1084.80	-45.0	-49.63	2.00
9		-1191.98	1150.53	-41.6	-48.62	3.00
10		-1211.24	1166.02	-45.2	-48.54	3.10
11		-1114.53	1081.87	-32.5	-47.51	4.70
12		-1148.86	1137.81	-11.05	-52.78	>100
13		-1213.17	1234.73	+21.55	-51.90	>100

Table 2-9 (previous page): Binding energetics from GPCR54 homology model and experimental data for downsized kisspeptin analogues. [A]Experimenally determined binding energy calculated as in Table 2-8.

This data shows again that although the range of total energies is larger for the theoretical data, the general trend obtained for docking each ligand in the GPCR54 model correlate to experimentally acquired binding data. Compounds which activate GPCR54 with potency similar to that of KP-10 (e.g. entries 2, 3 and 4) show very low total energy values in the docking studies. Importantly, compounds which were not able to active GPCR54 (e.g. entries 12 and 13) gave higher total energy scores *in silico*. These results further validate the model and its ability to distinguish high affinity compounds from non-binders even when their difference in structure is as discrete as the introduction of a nitrogen into an aryl ring (entry 4 vs. entry 9).

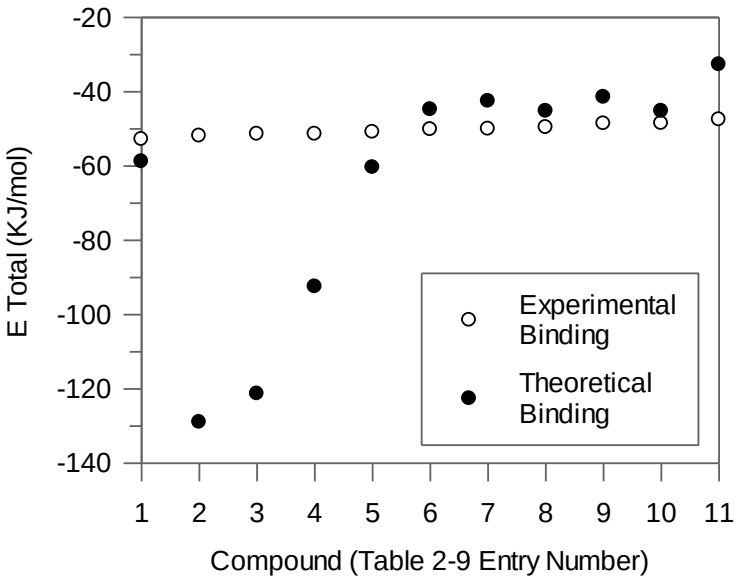


Figure 2-42: Theoretical binding energies vs. Experimentally derived Ki's for downsized KP-10 analogues

Figure 2-42 shows a plot of theoretically calculated and experimentally obtained affinities of the downsized KP-10 analogues shown in Table 2-9. It can be seen that overall, the trend of the ligand affinities correlate nicely, even though the scales are slightly different. Experimental values for entries 12 and 13 were omitted due to an undefined number being reported (>100).

2.5 Conclusions and Future Work

Biologically active fluorescent probes for GPCR54 were designed and synthesised and a 'reversed' kisspeptin-10 resin was generated to aid the development of an assay to screen

GPCR54 against OBOC libraries. Several approaches have been investigated for establishment of an on-bead screen of KP-10 against GPCR54, including the use of whole cells, membrane fractions, lipoparticle and nanolipoprotein particles. No interactions were seen between whole cells and the resin, possibly due to a lack of receptor presented on the surface of the HEK cell line used (indicated by lack of fluorescent ligand binding in a FACS assay). Membrane fractions were stained and analysed against KP-10 resin, however the resultant fluorescent rings observed were also present in the negative control, indicating that a non-specific interaction was taking place. Unfortunately, the use of commercial kits did not allow the successful synthesis of active lipoparticles or nanolipoprotein particles. Further efforts should focus on the investigation of these materials to ascertain whether they would indeed be suitable for the screening of GPCRs on-bead.

An *in silico* homology model of GPCR54 was generated based on the structure of related opiod receptors, with which it shares good enough sequence homology. It has been demonstrated that the sequence homology modelling-molecular-docking-MD simulation procedure used is able to generate the 3D structures of peptide-GPCR54 complexes, of which predicted binding energies correlates very well with experimental binding data, even in the absence of experimentally obtained crystal structure of the receptor and without prior knowledge of the binding site. The most important requirement of the molecular docking calculation is its ability to distinguish the real binding modes from ‘decoys’ (nonspecific or energetically unfavourable binding modes). In the absence of structural data for GPCR54 and its ligands, binding energies were used to validate the procedure. In all tested cases, the sequential docking procedure resulted in the correct estimation of the binding strength of the tested peptides and distinguished binders from non-binders. This raises the confidence in the applied methodology and enables the system for efficient design of new ligands for GPCR54 receptor with great therapeutic potential.

CHAPTER 3

INVESTIGATION OF KISSPEPTIN *C*-TERMINUS

&

DEVELOPMENT OF THE ALBA LINKER

3.1 Abstract

During the development of a library biased towards GPCR54, an investigation into the structure activity relationship of the kisspeptins drew attention to the highly conserved RFamide motif. In the literature, all chemical manipulations of the amide moiety carried out on the kisspeptins and other RFamide peptides have resulted in either inactive or highly reduced activity compounds. This chapter describes the synthesis of a KP-10 analogue where the RFamide has been replaced with an RFamine motif. The KP-10-amine was biologically evaluated but showed low activity against GPCR54, indicating that the amide carbonyl is essential for KP-10 binding. Comparison of the orientations of KP-10 and KP-10-amine following *in-silico* docking in a GPCR54 homology model suggests reasons why this is the case. During the synthesis of KP-10-amine, a novel acid labile linker for the attachments of amines to a solid support was developed.

3.2 Introduction

3.2.1 RFamide Peptides

The kisspeptins, introduced in chapter 1, belong to the RFamide family of peptides. These are neuropeptides which act on GPCRs and all of which possess a highly conserved *C*-terminal motif of Arginine, amidated Phenylalanine (Figure 3-1).

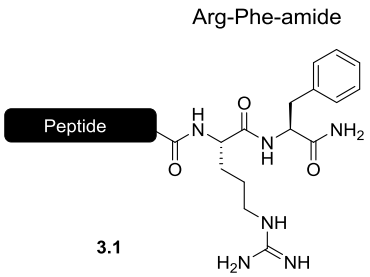


Figure 3-1: *C*-terminal RFamide motif

Several groups of RFamide peptides have been discovered in mammals (Table 3-1) with much more known in invertebrates.[134] They are split into five groups which, as well as the kisspeptins, include:

Neuropeptide FF (NPFF) group

Octapeptide neuropeptide FF (NPFF) and octadecapeptide neuropeptide AF (NPAF) make up this group of RFamide peptides. They bind neuropeptide FF receptor 1 (NPFF1R or GPCR147) and neuropeptide FF receptor 2 (NPFF2R or GPCR74).[135] Their names are

generated from their *N*- and *C*-terminal amino acids and they are primarily involved in the modulation of pain stimuli.[136]

Gonadotropin-inhibitory (GnIH) group

The gonadatrophin inhibitory group includes RFRP-1 (NPSF) and RFRP-3 (NPVF), which have inhibitory effects on GnRH neurons and bind GPCR147.[137]

26RFa group

Pyroglutamylated RFamide peptide receptor (QRFPR or GPCR103) is the endogenous receptor for the 26RFa group of RFamide peptides, which includes 43RFa (QRFP) and truncated version 26RFa. These peptides have been implicated in various biological processes such as activation of the gonadatropic axis,[138] regulation of appetite and energy expenditure[139] and regulation of the cardiovascular system,[140] among others.

Prolactin-releasing peptide (PrRP) group

PrRP31 and its truncated analogue PrRP20 make up the PrRP group, which endogenously bind prolactin-releasing peptide receptor (PrRPR or GPCR10). These peptides were originally identified as prolactin releasing factors in a study from which their names were derived.[141] They have since been implicated with processes such as control of body weight[142] and stress response.[143]

Group	Peptide	Sequence	Receptor
Kisspeptin	kisspeptin-54	GTSLSPPPESSGSRQQPGLSAPHSRQI PAPQ-GAVLVQREKDLPNYNWNSFGL**RF-NH$_2$**	GPCR54
	kisspeptin-14	DLPNYNWNSFGL**RF-NH$_2$**	
	kisspeptin-13	LPNYNWNSFGL**RF-NH$_2$**	
	kisspeptin-10	YNWNSFGL**RF-NH$_2$**	
NPFF	NPFF	SQAFLFQPQ**RF-NH$_2$**	GPCR74
	NPAF	AGEGLNSQFWSLAAPQ**RF-NH$_2$**	
GnIH	RFRP-1 (NPSF)	MPHSFANLPL**RF-NH$_2$**	GPCR147
	RFRP-3 (NPVF)	VPNLPQ**RF-NH$_2$**	
26RFa	43RFa (QRFP)	<EDEGSEATGFLPAAGEK-TSGPLGNLAEELNGYSRKKGGFSF**RF-NH$_2$**	GPCR103
	26RFa	TSGPLGNLAEELNGYSRKKGGFSF**RF-NH$_2$**	
PrRP	PrRP31	SRTHR-HSMEIRTPDINPAWYASRGIRPVG**RF-NH$_2$**	GPCR10
	PrRP20	TPDINPAWYASRGIRPVG**RF-NH$_2$**	

Table 3-1: Human RFamide peptides

Existing RFamide peptides are known to be involved in a vast array of biological processes. These include metastasis suppression,[48] regulation of reproduction,[144, 145] reduction of food consumption[146] and elevation of blood pressure.[147] Due to the vast range of disease states these peptides are implicated in and the fact that novel examples are predicted to be discovered in the future, any information on the structure activity relationship (SAR) is valuable in understanding binding modes and for generating active analogues.

Chemical manipulations on various RFamide peptides have been reported, with the aim of elucidating the importance of the amide functionality of the RFamide. These will be summarised *vide infra*. The *C*-terminus of Substance P, a non RFamide family member which also possesses a *C*-terminal amide, has also been examined.

3.2.2 *C*-Terminal Investigations of Kisspeptin

Ohtaki *et al.* reported an analogue of kisspeptin-54 whereby the *C*-terminal carboxamide was replaced with a free acid.[50] This compound showed only very weak activity on GPCR54 in a calcium mobilisation assay, suggesting that the amide moiety is very important for receptor activation.

As part of a large SAR investigation into kisspeptin, Orsini *et al.* generated a KP-13 analogue where the *C*-terminal amide was replaced with a free acid group (Figure 3-2).[85] Biological evaluation in a radioligand binding assay showed that this modification significantly inhibited receptor binding with an increase in K_i from 1.1 ± 0.4 nM for the amidated KP-13 to 551.8 ± 54.0 nM for the free acid form. Similar results were seen for agonist function in a calcium mobilisation assay, with an increase in EC_{50} from 22.8 ± 10.9 nM to 1068.5 ± 116.1 nM.

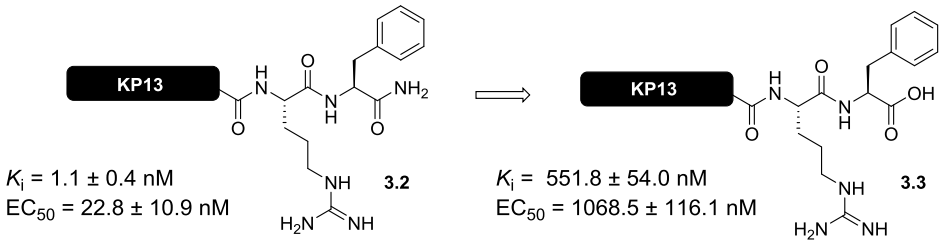


Figure 3-2: Replacement of *C*-terminal amide with free acid in KP13 severely depletes binding affinity to GPCR54

Tomita and co-workers developed potent downsized peptidic analogues of kisspeptin, which were used to probe the effects of modification of the *C*-terminus (Figure 3-3).[89] The analogue used for *C*-terminal investigation was a pentamer, composed of the five most *C*-terminal residues of kisspeptin (but with Phe^{10} replaced by Trp) and the *N*-terminus flanked with 4-aminobenzoic acid. This analogue showed 96.5 ± 0.3% agonistic activity compared to KP-10 (an EC_{50} of 1.4 nM) in a calcium mobilisation assay.

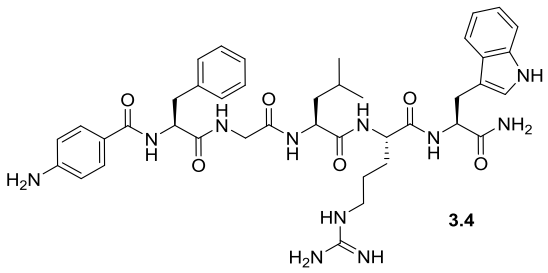


Figure 3-3: Downsized kisspeptin analogue reported by Tomita *et al*.

A series of analogues of this compound were generated whereby the *C*-terminal amide was replaced by various groups, including an acid (**3.5**), *N*-methylamide (**3.6**), *N*,*N*-dimethylamide (**3.7**) and hydrazide (**3.8**) (Figure 3-4). Agonistic activities of this series

were determined using a calcium flux assay, which showed a remarkable loss of activity for all compounds (Table 3-2).

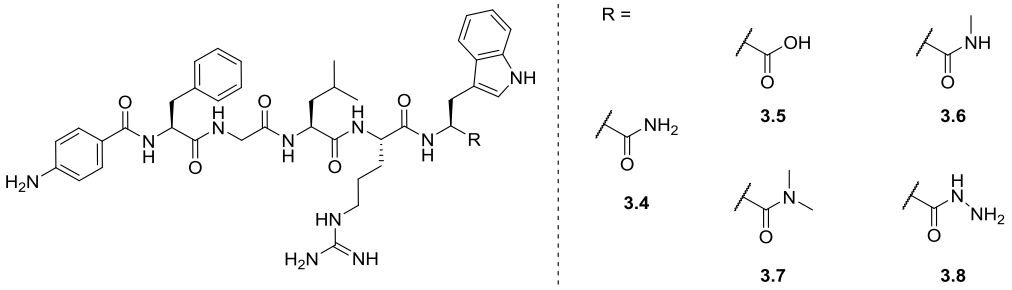


Figure 3-4: *C*-terminal modifications carried out by Tomita *et al.*

Entry	R	Cpd. No	% activity	EC_{50} (nM)
1	KP-10	-	-	0.18 – 1.1
2	-Tryp-NH_2	**3.4**	96.5 ± 0.3	1.4
3	–Tryp-CO_2H	**3.5**	3.8 ± 0.1	-
4	–Trp-NHMe	**3.6**	0.2 ± 0.1	>100
5	–Trp-NMe_2	**3.7**	0.1 ± 0.1	>100
6	–Trp-$NHNH_2$	**3.8**	0.0 ± 0.0	>100

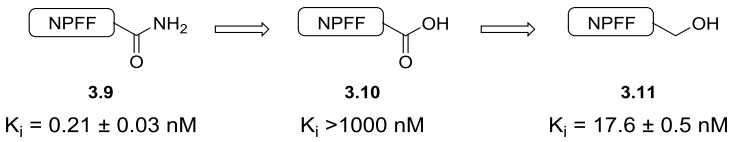
Table 3-2: Bioactivities of *C*-terminally-modified pentapeptide analogues, H-Amb-Phe-Gly-Leu-Arg-Trp-R, as reported by Tomita *et al.*

The authors note their surprise that even the addition of a methyl group (in the case of the *N*-methylamide analogue, entry 4) causes such severe activity loss and hypothesise that the amide hydrogens may be involved in hydrogen bonding with GPCR54.

3.2.3 *C*-Terminal Investigations of Additional RFamide Peptides

3.2.3.1 Neuropeptide FF

While investigating the SAR of neuropeptide FF, Mazarguil and co-workers generated and tested some *C*-terminally modified derivatives (Figure 3-5).[148] The first manipulation was synthesis of the peptide with a *C*-terminal free acid (**3.10**), which resulted in a loss of binding affinity from 0.21 ± 0.03 nM for the natural peptide to > 1000 nM in a radioligand binding assay. This result supports similar findings by Payza *et al.*,[149] who also observed loss of binding upon replacement of the amide with a free acid.



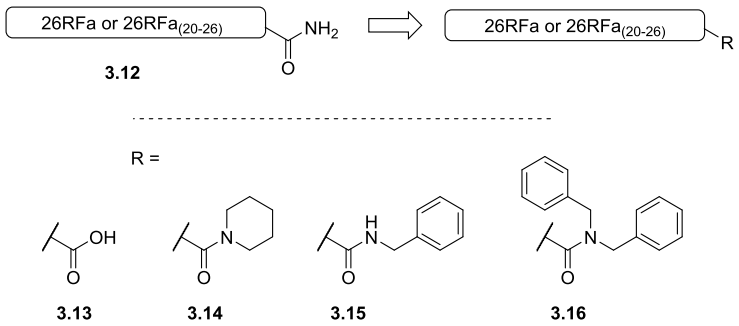
3.9 K_i = 0.21 ± 0.03 nM **3.10** K_i >1000 nM **3.11** K_i = 17.6 ± 0.5 nM

Figure 3-5: *C*-terminal modifications of neuropeptide FF

The *C*-terminal acid group was subsequently reduced to give the corresponding peptide alcohol (**3.11**), which was also biologically assessed. Interestingly, this showed a binding affinity of 17.6 ± 0.5 nM, suggesting that some binding features could have been redeemed upon reduction. It could be that the negative charge of the carboxylate at physiological pH is unfavoured in the binding site and changing this to the alcohol removes the charge.

3.2.3.2 26RFa/43RFa

Fukusumi and co-workers reported that exchange of the *C*-terminal amide of 26RFa to the corresponding carboxylic acid leads to a 400 fold lower affinity of the peptide for GPCR103.[150] Modifications to the *C*-terminal amide of 26RFa peptide were also carried out by Le Marec *et al.* (Table 3-3).[151] Mono- and di- alkylated amines were coupled to the free *C*-terminus of human 26RFa and active heptapeptide analogue $26RFa_{(20-26)}$ to give a series of compounds which were then evaluated against GPCR103 in a calcium flux assay.



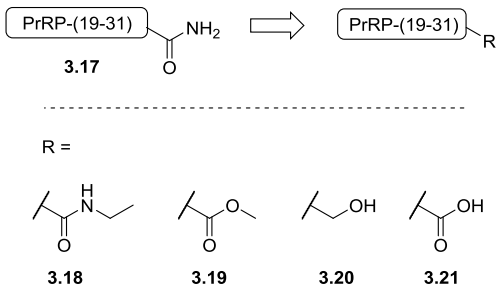
Entry	R	Cpd. No	EC_{50} (nM) 26RFa	EC_{50} (nM) 26RFa$_{(20-26)}$
1	$CONH_2$	**3.12**	10.4 ± 1.5	739 ± 149
2	CO_2H	**3.13**	531 ± 319	-
3	*N*,*N*-piperidinyl	**3.14**	282 ± 134	1198 ± 286
4	*N*-benzyl	**3.15**	39 ± 5.5	-
5	*N*,*N*-dibenzyl	**3.16**	166 ± 61	-

Table 3-3: *C*-terminal modifications made to 26RFa

These results agree with those of Fukusumi,[150] with the free acid peptide **3.13** showing very poor activity. Alkylation of the *C*-terminal amide also decreases Ca^{2+}-mobilising activity compared to the native peptide, leading the authors to suggest that the hydrogens present on the amide are involved in hydrogen bonding either to the receptor, or more likely, to another part of the peptide.

3.2.3.3 PrRP31/PrRP20

Like its family members, the free acid variant of PrRP31 was shown by Hinuma and co-workers to have significantly reduced receptor binding activity compared to the amidated form.[152] Boyle *et al*. performed SAR studies around tridecapeptide analogues of PrRP-(19–31)-peptide.[153] The approach included both a calcium mobilisation based functional assay, and a radioligand binding assay. With respect to the *C*-terminus, their findings agreed with those of Hinuma, as the free acid **3.21** was not functionally active (Table 3-4). However, alkylation of the amide (**3.18**) still showed high binding affinity. Methyl ester variant **3.19** also surprisingly showed good binding affinity but analogue **3.20**, bearing a terminal alcohol showed lower affinity binding.



Entry	R	Cpd No	FLIPR EC_{50} (μM)	Binding K_i (nM)
1	$CONH_2$	**3.17**	0.02	5.3
2	CONHEt	**3.18**	not tested	4.4
3	CO_2Me	**3.19**	not tested	35.3
4	CH_2OH	**3.20**	0.89	23.5
5	CO_2H	**3.21**	-	5000

Table 3-4: *C*-terminal modifications made to PrRP-(19–31)

3.2.4 *C*-Terminal Investigations of Other *C*-Terminally Amidated Peptides

3.2.4.1 Substance P

Substance P is a neuropeptide with an amidated *C*-terminus which acts upon neurokinin 1 receptor. Although not a member of the RFamide family of peptides, investigations into its amidated *C*-terminus led to some interesting conclusions being drawn, which could be applicable to the case of kisspeptin and the RFamide peptides.

Escher and co-workers undertook a detailed study of the *C*-terminus of **3.22** (Figure 3-6), an active seven amino acid length truncation of substance P, with the *C*-terminal methionine having been replaced with norleucine. Replacement of its *C*-terminal carboxamide with an acid group (**3.23**) resulted in a compound that was almost inactive.[154]

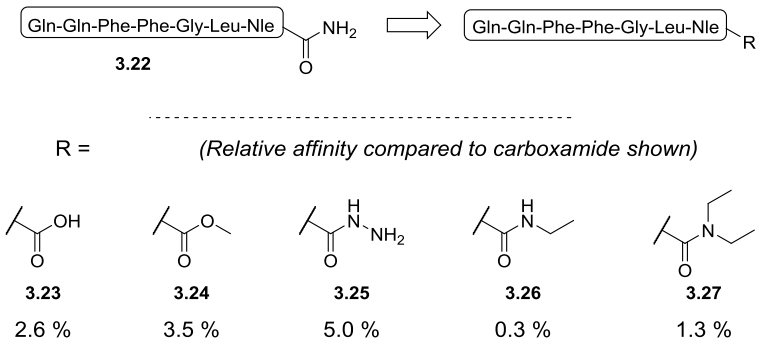


Figure 3-6: Compounds made in the *C*-terminal investigation into substance P analogue

The group also synthesised peptides with the carboxamide replaced with a methyl ester (**3.24**), hydrazide (**3.25**), *N*-ethylamide (**3.26**) and *N*,*N*-diethylamide (**3.27**). These modifications rendered the compounds much less active, with the authors noting that affinity being at least 20 fold less than the amide in each case.

The authors go on to deliberate the possible cause of this inactivity, reasoning that all modifications carried out either increase the size of the group relative to the amide or introduce a charge and prevent the amide protons from forming hydrogen bonds. Therefore two possible explanations are offered:

A. The binding site accommodating the peptides is not large enough to accept any group bigger than the amide. In the case of the free acid, this would be presented in the form of a charged carboxylate at physiological pH and therefore be surrounded in a hydrating shell.
B. The amide hydrogens are participating in hydrogen bonding either to the receptor, or in an intramolecular fashion to another point on substance P. If both hydrogens are involved in such an interaction, even mono alkylation would disrupt binding.

3.2.5 Summary

From all of these studies, it can be concluded that changes to *C*-terminal amide moieties in RFamide peptides and others are not well tolerated and generally lead to drastically depleted binding affinity and receptor activation. The possibility that the amide hydrogens are implicated in hydrogen bonding is offered as an explanation more than once, which would explain why activities are ablated even upon single methylation of the amide. However, Mazarguil showed that a hydroxyl functionalised neuropeptide FF analogue showed some activity in binding and activation of $NPFF_2$ receptors, which calls this argument into question.[148]

The hydroxyl modification is the smallest of all examples shown *vide supra*, which supports the argument that perhaps the receptors with which the peptides are interacting do not have a binding pocket large enough to accommodate larger substituents.

Similar results are seen across the entire family of RFamide peptides when modifications are made to the *C*-terminus. It is therefore possible that a modification on the *C*-terminus of kisspeptin that results in higher activity could also be applicable across the whole RFamide family of peptides, and maybe even further to other amide modified peptides such as substance P.

It was noticed that with the exception of the hydroxyl analogue of Neuropeptide FF and PrRP31, all modifications were taking place on the amide nitrogen (Figure 3-7). As the hydroxyl peptide completely removes the carbonyl of the amide with less drastic effects than the other substituents, the importance of the role played by the carbonyl group could be questioned.

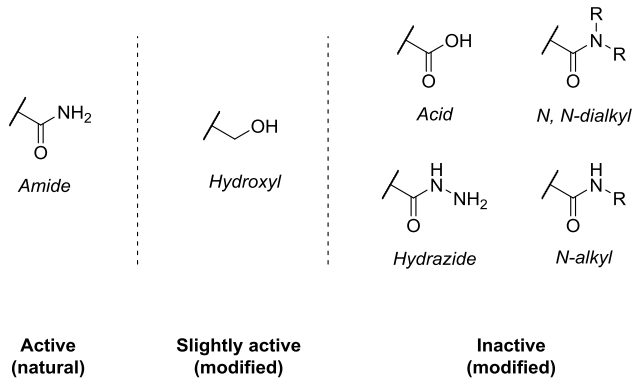


Figure 3-7: Reported changes to *C*-terminal amide peptides

A *C*-terminal amine peptide would be an interesting addition to this field of study (Figure 3-8), as it would likely be a small enough analogue fit into a binding site that is known to accommodate the larger amide group. An amine group would also leave the amine hydrogens in place, and therefore retain the ability of the compound to make hydrogen bonds to the receptor or in an intramolecular fashion.

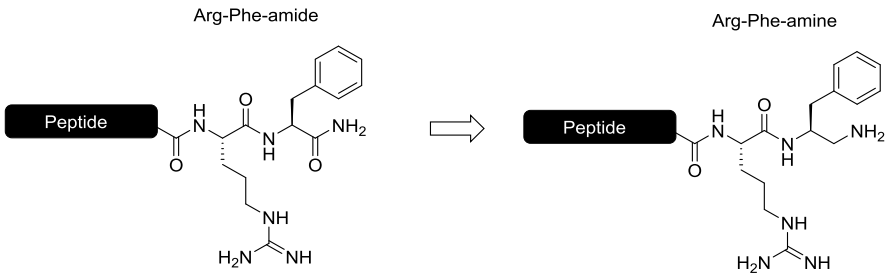
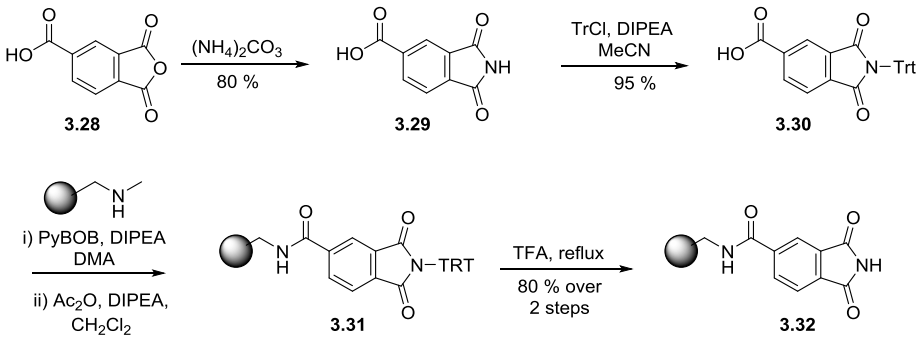


Figure 3-8: Structure of proposed RFamine peptide

Such a peptide would best be prepared *via* solid supported synthesis, therefore options for anchoring amines onto resins for SPPS were evaluated.

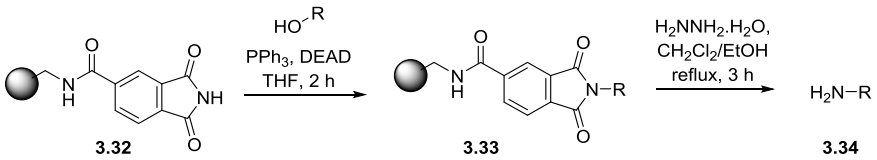
3.2.6 Amine Linkers for Solid Phase Synthesis

Only a few linkers exist in the literature that allow capture and release of primary amines on solid support. One such linker was reported by Aronov and co-workers (Scheme 3-1).[155] Here, a phthalimide reagent was anchored on resin and used to generate primary amines from alcohols using Mitsunobu chemistry (Scheme 3-2), in order to circumvent the laborious work up procedures experienced in solution phase. Trimellitic anhydride **3.28** was reacted with ammonium carbonate in a melt reaction at 280 °C to yield imide **3.29** in 80% yield. This was tritylated in 95% yield using trityl chloride to give **3.30**. The tritylated imide was then loaded onto aminomethyl resin, after which any free amines were capped using acetic anhydride. Removal of the trityl group gave **3.32** and allowed quantification of loading, which was determined to be 80%.



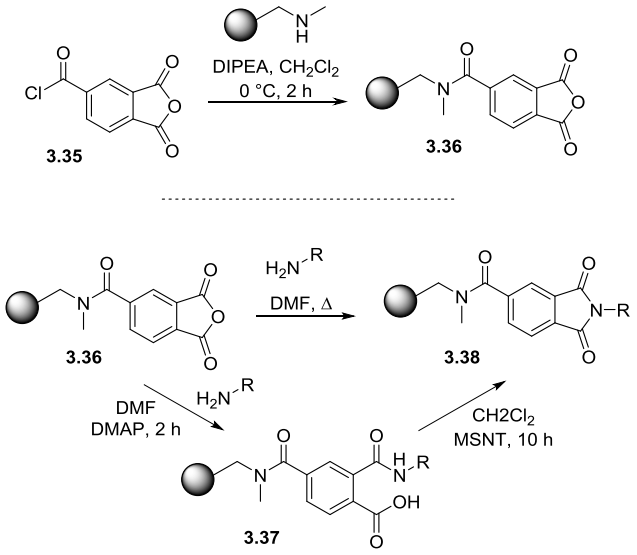
Scheme 3-1: Generation of activated resin for Mitsunobu reaction

Following resin preparation, a number of alcohols were loaded under Mitsunobu conditions and subsequently cleaved using hydrazine, affording the desired primary amines in high purity and 85 – 93% yields (Scheme 3-2).



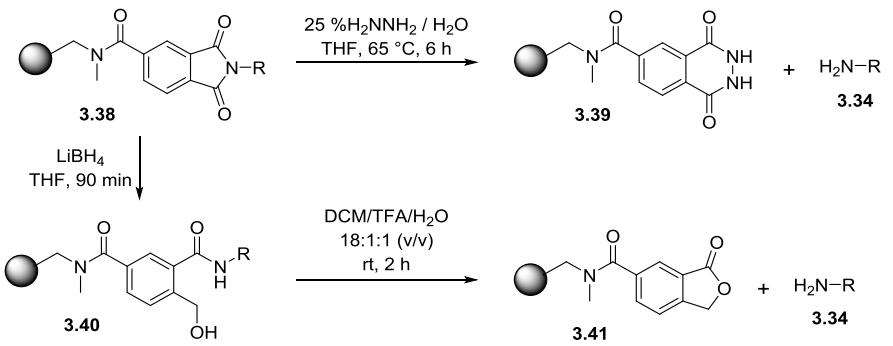
Scheme 3-2: Loading of alcohols to phthalimide resin *via* Mitsonubu reaction and cleavage to afford primary amines

Another amine linker, also based around a phthalimide structure, was described by Bauer *et al*. (Scheme 3-3).[156] The Trimellitic anhydride linker (TAL) was generated from trimellitic anhydride chloride **3.35**, which was loaded directly onto amine resin to give **3.36**. Primary amines were then introduced to generate phthalimides **3.38**. This was achieved by one of two methods, either by heating, or in a milder two-step procedure by firstly generating the secondary amide **3.37**, then encouraging condensation and formation of the phthalimide using MSNT.



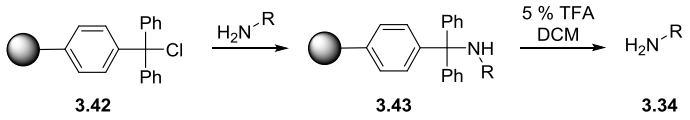
Scheme 3-3: Loading of amines onto TAL

TAL was used to generate labelled amine bearing carbohydrate derivatives. Once labelled, the compounds could be cleaved from the resin by a choice of methods (Scheme 3-4). A nucleophile mediated approach using hydrazine at elevated temperatures could be employed, or if the products were base sensitive, *via* a safety catch method. In this case, the phthalimide was reduced using lithium borohydride to give secondary amide **3.40**. Mild acidic treatment of this species encouraged intramolecular cyclisation of the linker by nucleophilic attack of the hydroxyl group on the amide carbonyl to form phthalide **3.41**, causing the product amine **3.34** to be eliminated and therefore released from the resin.



Scheme 3-4: Two possible cleavage methods of amines from TAL

Chlorotrityl resin **3.42** can also be used for the attachment of amines to a solid support (Scheme 3-5).[157-159] Amines can be easily attached to the resin *via* displacement of the chloride and cleavage can be achieved in very weak acidic conditions. However, the resin is extremely water sensitive and loaded compounds can be prematurely cleaved, even with coupling reagents such as HOBT.



Scheme 3-5: Immobilisation and cleavage of amines on chlorotrityl resin

3.3 Aims

The objective of this work was to synthesise and biologically evaluate an analogue of kisspeptin-10 whereby the carbonyl group of the amide was removed, effectively generating an RFamine motif.

3.4 Results and Discussion

In order to achieve the above aim, a method had to be established by which the required amine could be tethered to a solid support and SPPS carried out as normal. For this purpose, a phthalimide based linker was chosen due to its stability and choice of cleavage conditions. Figure 3-9 shows the setup targeted in order to allow synthesis of a KP-10-amine analogue.

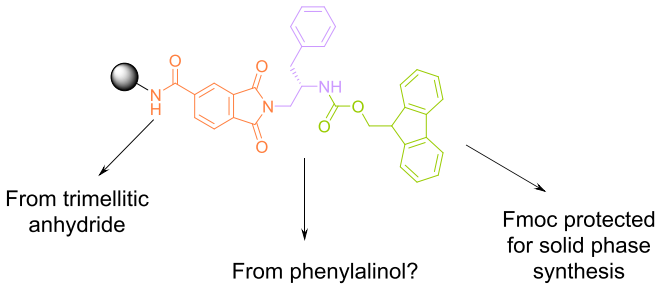
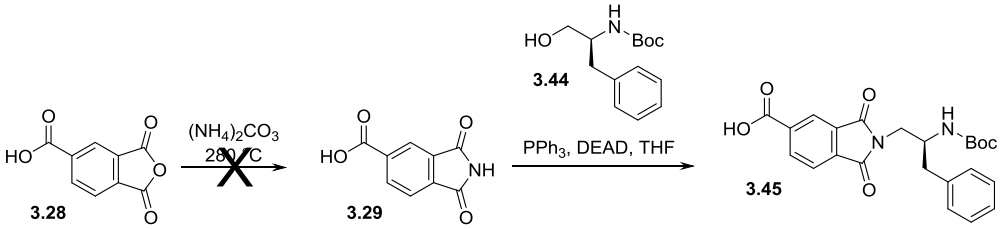


Figure 3-9: Proposed structure of phthalimide linker Phe analogue

Protected phenylalinol **3.44** was identified as a starting material, with the intention to use Mitsunobu reaction conditions to transform the alcohol into the desired amine, as demonstrated by Aronov and co-workers in Scheme 3-2.[155] This could allow production of building block **3.45** containing both the required diamine and linker, which could be loaded onto resin for subsequent functionalisation (Scheme 3-6).



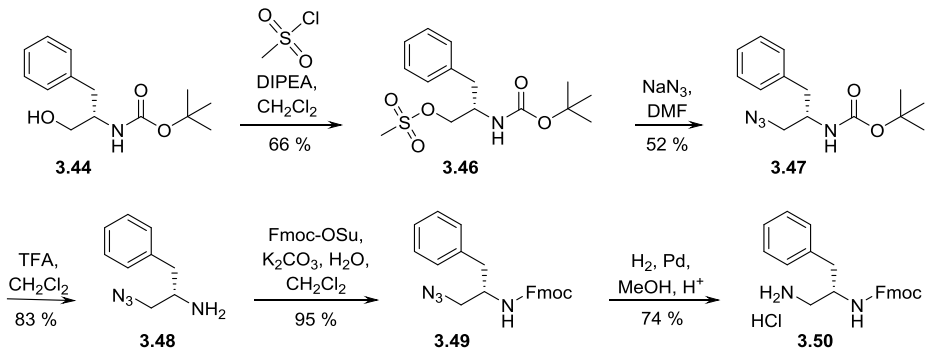
Scheme 3-6: First proposed synthesis of KPamine building block

The melt reaction described in the report by Aronov *et al*. between trimellitic anhydride and ammonium carbonate was difficult to reproduce due to the high temperatures required, which are not easily achievable using standard laboratory equipment. Therefore, an alternative synthetic strategy was explored. This involved synthesis of the required diamine building block first. This would then be Fmoc protected in order to make the compound compatible with SPPS. As a final step, this building block would be loaded into a phthalimide linker for bead attachment.

3.4.1 Synthesis of Diamine Building Block

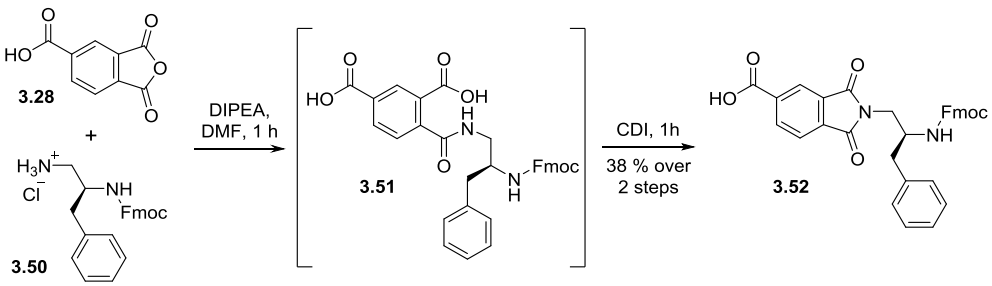
Conversion of the phenylalinol alcohol group into an amine was targeted (Scheme 3-7). This was achieved by reacting Boc protected phenylalinol **3.44** with mesyl chloride to generate the corresponding mesylate **3.46** in good yield. The mesylate was displaced using sodium azide at elevated temperatures to generate the corresponding azide **3.47**. Boc deprotection of the amine was then carried out and replaced with an Fmoc group in good yield to give **3.49**.

Finally, the azide was reduced in the presence of HCl to avoid Fmoc deprotection to give the desired diamine product **3.50** as the hydrochloride salt.



Scheme 3-7: Synthesis of KP-10-amine building block

The diamine was incorporated into carboxylic acid functionalised phthalic anhydride to form the desired phthalimide for bead linkage. Initial attempts involved heating the reactants together, however upon LCMS analysis, no required product was observed. Instead, multiple side products were observed from degradation of the Fmoc protecting group. Mild conditions were sought, and reaction of the anhydride and amine was carried out at room temperature with pyridine in DMF (Scheme 3-8). After an hour, formation of the pththalylic acid intermediate **3.51** was evident by LCMS analysis. Addition of 1,1'-carbonyl diimidazole to the crude mixture encouraged the ring closing step to take place and afforded desired phthalimide **3.52** within 1 hour at room temperature.

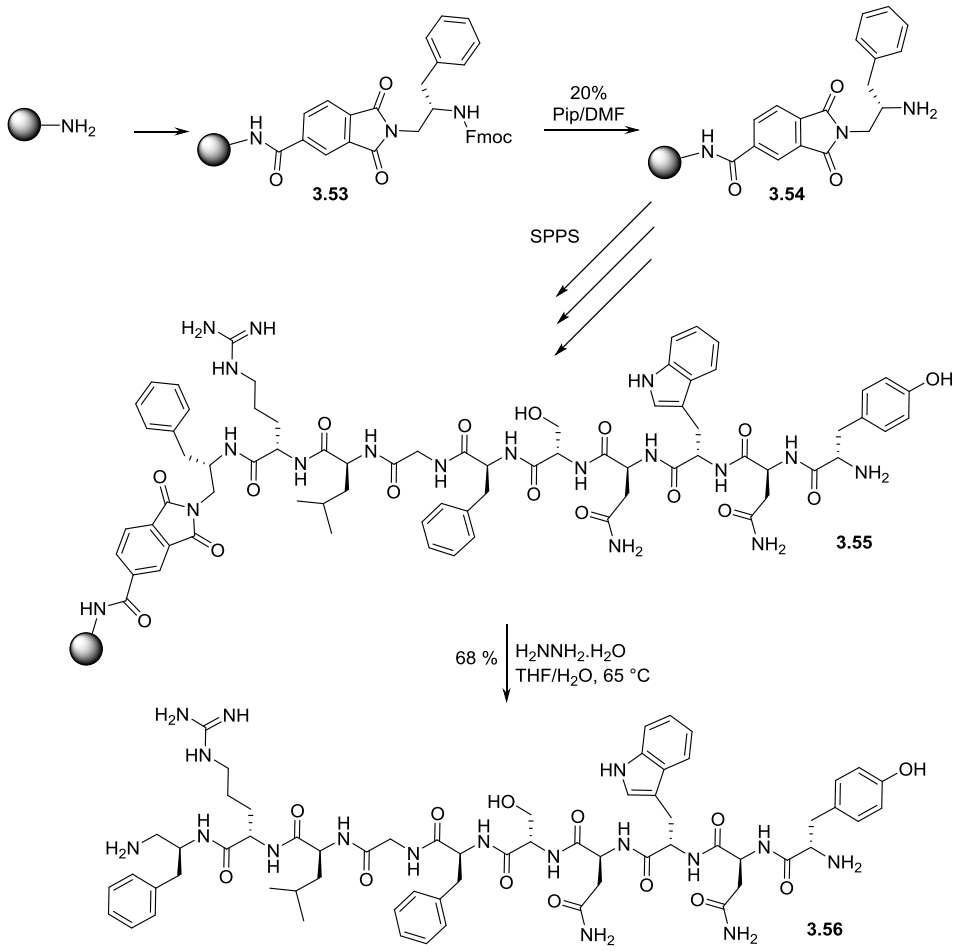


Scheme 3-8: Loading KP-10-amine building block into trimellitic anhydride

3.4.2 Synthesis of KP-10-Amine

Phenylalanine analogue loaded phthalide **3.52** was coupled to TentaGel NH_2 beads and Fmoc deprotection carried out. The compound was elaborated by SPPS until the full kisspeptin decamer analogue **3.55** was made (Scheme 3-9). Deprotection of side chains was carried out using TFA:H_2O:TIS (95:2.5:2.5) and then cleavage of the phthalimide linker

performed using 10% hydrazine in THF at 65 °C for 6 hours to give the desired KP-10-amine **3.56** in 68% yield.



Scheme 3-9: Synthesis of KP-10-amine using SPPS

3.4.3 Biological Evaluation of KP-10-amine

KP-10-amine was tested in a cellular functional assay against GPCR54 (for details see sections 2.4.2.1 and 8.3.3).[§] Unfortunately, the results suggest that the KP-10-amine compound has severely reduced activity against GPCR54 compared to wild type KP-10 (Figure 3-10).

[§] Biological evaluation of KP-10 analogue carried out by Megan Gallant at the Univerity of Capetown, South Africa.

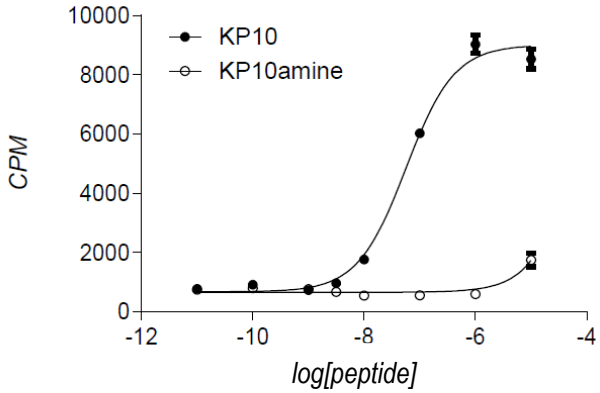


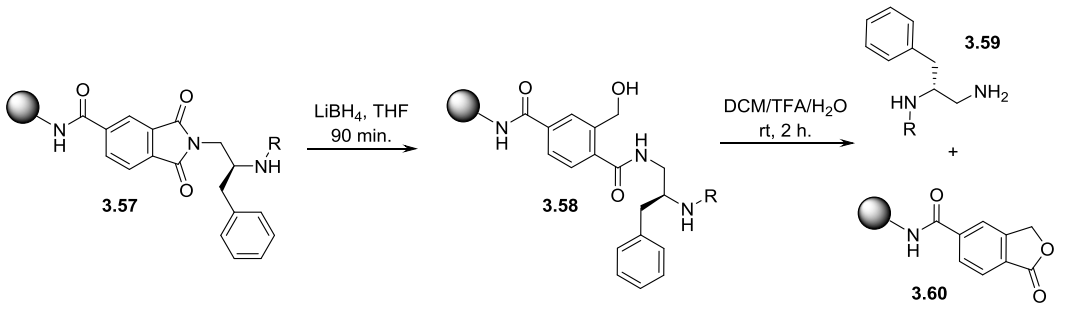
Figure 3-10: KP-10-amine was not able to stimulate signalling in GPCR54

It can therefore be concluded that replacement of the *C*-terminal amide of KP-10 with an amine group renders kisspeptin unable to activate GPCR54. From this we can conclude that the presence of the carbonyl group of the *C*-terminal amide is very important in the binding of KP-10 to GPCR54. This demonstrates the importance of conservation of the carbonyl moiety in the *C*-terminal RFamide motif.

3.4.4 Development of an Acid Labile Benzofuranone Based Amine (ALBA) Linker for Solid Phase Synthesis

During planning of the KP-10-amine synthesis, previously reported linkers were considered with which to attach the required amine. However, it was noticed that only the trityl resin was acid cleavable, and that in fact this was so sensitive to acidic conditions that premature cleavage becomes an issue. An acid cleavable linker for attachment of amines to a solid support would therefore be highly beneficial, as most building blocks used for solid phase peptide synthesis are provided with acid labile protecting groups.

The work of Bauer *et al.* was revisited.[156] Here, an alternative method for cleavage of compounds from a phthalimide based resin was reported, which first involves reduction of the phthalimide to give the corresponding hydroxymethyl benzoic amide (Scheme 3-10). Acidic conditions are then employed to encourage the formation of a stable benzolactone through nucleophilic attack of the hydroxyl group on the amide carbonyl, cleaving the amide bond and liberating the free amine.



Scheme 3-10: Reductive/acidic cleavage of TAL

The development of a linker based on the 2-hydroxymethyl benzoic acid (2-HMBA) intermediate **3.58** was investigated (Figure 3-11). This would allow a one-step acid mediated cleavage mechanism for amines and also allow loading of the linker using highly efficient amide bond formation. Masking of the hydroxymethyl functionality using a protecting group would prevent premature cyclisation and allow synthesis from the carboxylic acid.

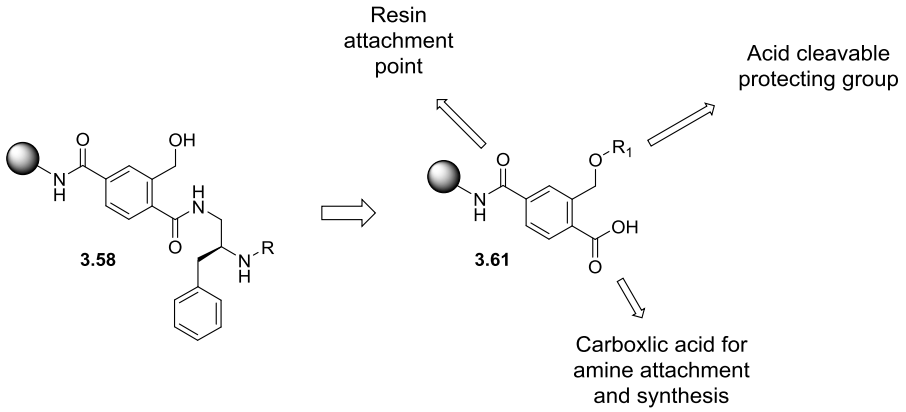
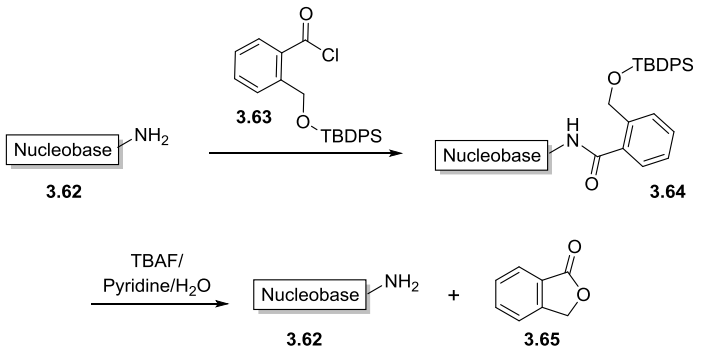


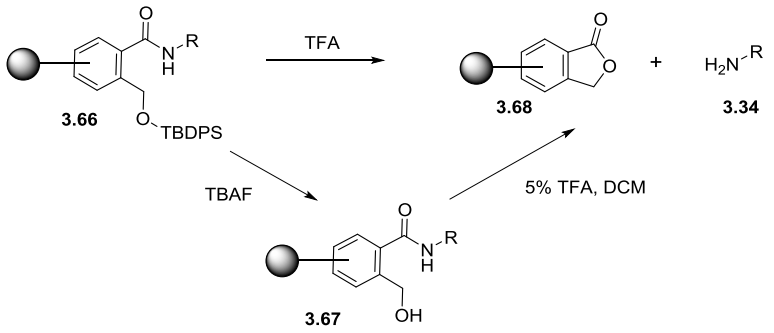
Figure 3-11: Requirements for development of new linker based on 2-hydroxymethylbenzoic acid

In order to cleave compounds from the resin following synthesis, the hydroxyl protecting group would be removed and linker cyclisation could take place upon treatment with acid. Dreef-Tromp and co-workers reported 2-(*tert*-butyldiphenylsilyloxymethyl)benzoyl chloride (SiOMB-Cl, **3.63**) for the protection of nucleobase amines (Scheme 3-11).[160] This uses the same benzolactone forming mechanism described above. Here, deprotection is achieved by subjection to TBAF, pyridine is then used to encourage cyclisation of the 2-HMBA and release the unprotected nucleobase.



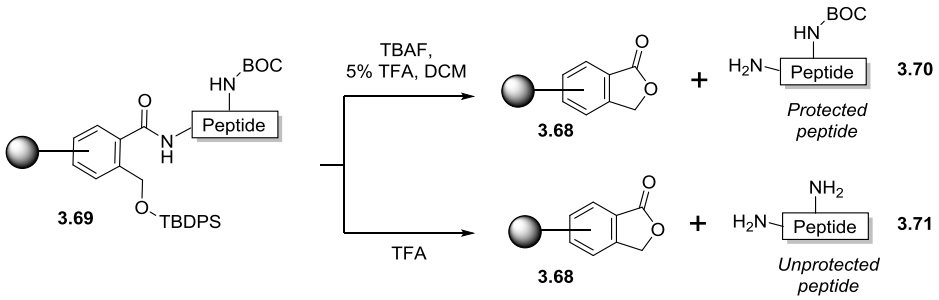
Scheme 3-11: 2-HMBA based protecting group reported by Dreef-Tromp *et al.*

The use of TBDPS presents two possible modes of compound cleavage (Scheme 3-12). Firstly by deprotection of the silyl protecting group using fluoride, followed by cyclisation induced using dilute acid. A second option would be to choose conditions that would deprotect the TBDPS group and encourage cyclisation simultaneously, which could be carried out using neat TFA.



Scheme 3-12: Two proposed possible cleavage modes of ALBA linker

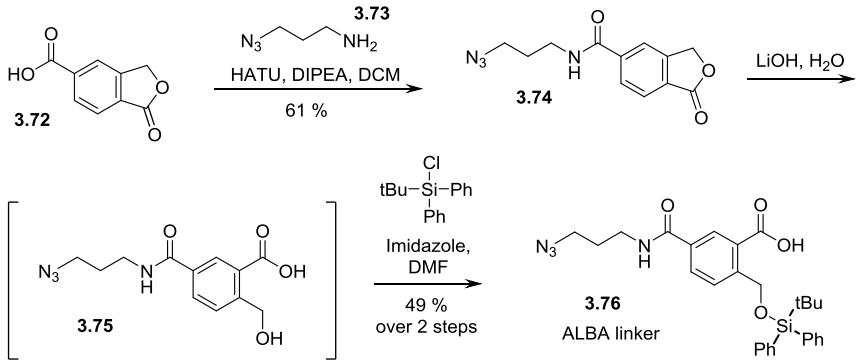
As the cleavage conditions using TBAF/dilute acid are mild, we reasoned it could also be possible to cleave the linker whilst leaving acid labile protecting groups of amino acids intact, therefore cleaving the fully protected peptide (Scheme 3-13). This could be useful for fragment based peptide synthesis or modification of the *C*-terminus following cleavage.



Scheme 3-13: ALBA linker could potentially be used to cleave either protected or deprotected peptides

3.4.4.1 Synthesis of the ALBA Linker

Synthesis of the ALBA (acid labile benzofuranone amine) linker was undertaken. The first step in linker design introduced a point in the 2-HMBA scaffold to allow anchoring to a solid support (Scheme 3-14). 5-Carboxyphthalide **3.72** provided the required functionality for coupling of azidopropylamine **3.73**, which was employed as the azide allows attachment of the linker to resin either *via* click chemistry, or by reducing to the amine and using amide coupling. These options are preferable to ethers, as these can lead to unwanted side reactions during acidic cleavage conditions.[161] Subsequent opening of the phthalide ring with hydroxide gave intermediate **3.75**, and *in situ* protection of the resulting alcohol using *tert*-butyldiphenyl silyl chloride[162] furnished the completed ALBA linker **3.76**.

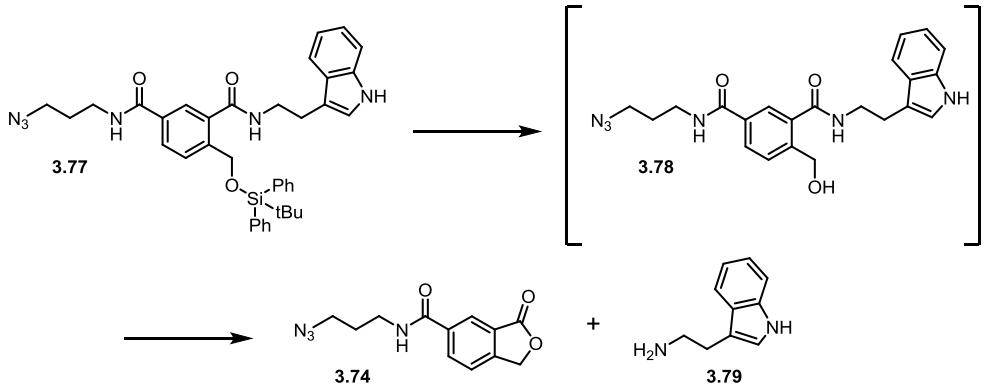


Scheme 3-14: Synthesis of ALBA Linker

3.4.4.2 ALBA Linker Stability Studies

The ALBA linker was coupled with tryptamine to give **3.77** and this compound assessed in stability studies in solution (Table 3-5). The linker proved stable to Fmoc based SPPS coupling solvents (10% DIPEA in DMF) and deprotection conditions (20% piperidine in DMF). Formation of the phthalide and liberation of tryptamine by subjection to TFA was

shown to be complete within 4 hours, confirming that this mode of cleavage is effective. TBAF deprotection of the TBDPS group was complete after 30 minutes, however subsequent cyclisation and elimination of tryptamine in basic conditions (pyridine/H_2O) as described by Dreef-Tromp *et al.*[160] was not observed after 24 hours. 2-HMBA can be cyclised using mild acidic conditions (5% TFA in CH_2Cl_2),[156] therefore these conditions were assessed with the tryptamine-loaded ALBA linker and was shown to yield the released product within 4 hours (For experimental details, see section 8.3.5).

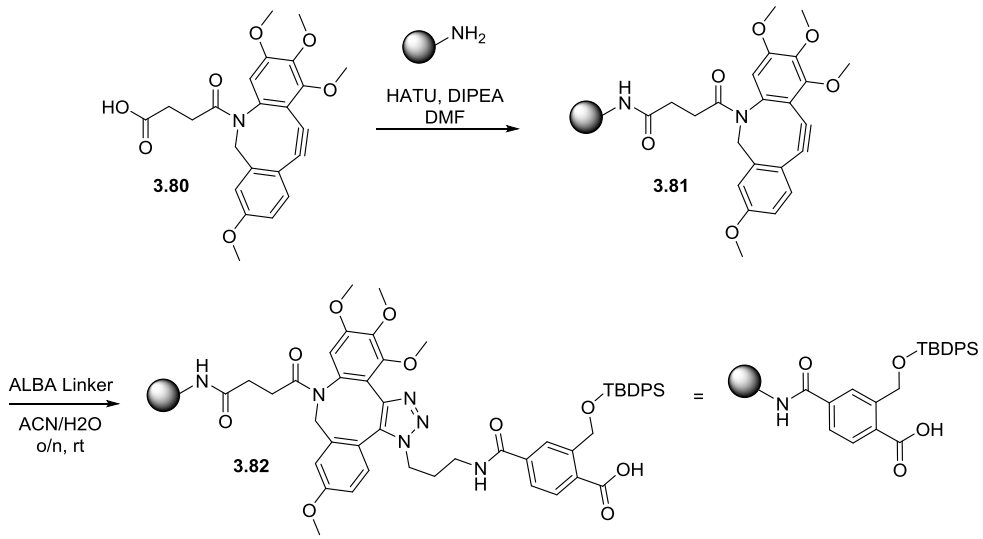


Entry	Conditions	Species Observed
1	20% Piperidine/DMF	**3.77**
2	10% DIPEA/DMF	**3.77**
3	1 mM TBAF	**3.78**
4	TFA	**3.74** + **3.79**
5	1 mM TBAF then Pyridine/H_2O	**3.78**
6	1 mM TBAF then 5% TFA in CH_2Cl_2	**3.74** + **3.79**

Table 3-5: Determining Linker Stability/Cleavage Conditions - species present were determined using LC-MS analysis

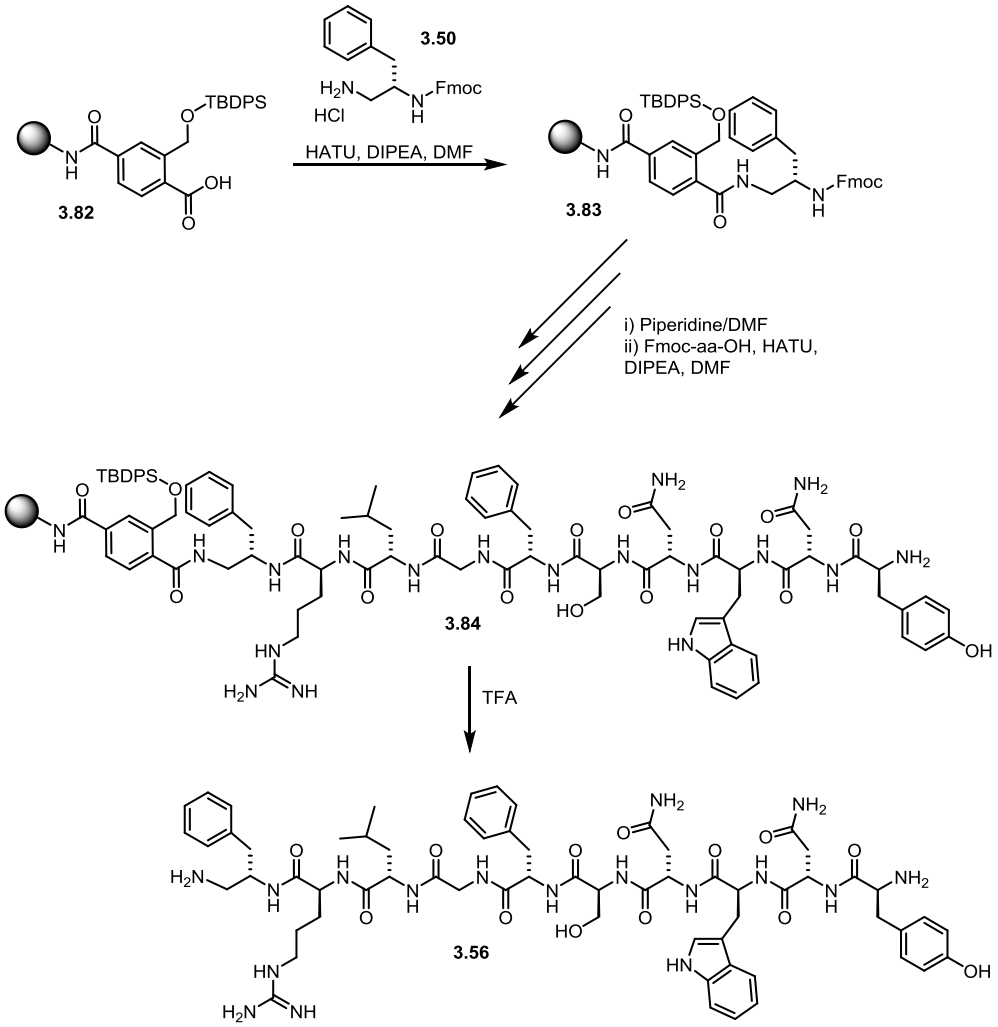
3.4.4.3 Synthesis of KP-10-amine Using ALBA Linker

With the ALBA linker and amine building block in hand, synthesis of the kisspeptin analogue was undertaken again to prove utility of ALBA (Scheme 3-15). TentaGel NH_2 beads were loaded with strained cyclooctyne **3.80**[163] to give alkyne functionalised resin **3.81**. This allowed the ALBA linker to be loaded using copper free 1,3-dipolar cycloaddition,[164] generating ALBA linker resin **3.82**.



Scheme 3-15: Loading ALBA onto resin *via* copper free click chemistry

An advantage of using copper free 1,3-dipolar cycloaddition is that no other reagents are required, meaning unused linker material can be recovered, allowing larger excesses of the linker to be used to ensure high reaction yields. Amine building block **3.50** was coupled to the linker using standard coupling conditions and then SPPS allowed the build-up of the full KP-10-amine decamer **3.84** (Scheme 3-16).



Scheme 3-16: Synthesis of KP-10-amine on ALBA linker

Following synthesis, release of the final product from the solid support was achieved by incubating the resin in TFA. Isolation of KP-10-amine peptide **3.56** demonstrates the utility of the ALBA linker in the synthesis of *C*-terminally amine functionalised peptides.

3.5 Conclusions and Future Work

A kisspeptin analogue was synthesised whereby the highly conserved RFamide motif was replaced with an RFamine. The synthesis of this novel compound aimed to probe the importance of the carbonyl moiety in the highly conserved RFamide motif. Indeed, many modifications to members of the RFamide family of peptides involve removal or substitution on the nitrogen of the amide, whereas the carbonyl remained unchanged. Synthesis of the peptide was achieved by the generation of a phenylalanine analogue whereby the carboxylic acid was replaced with an amine. Solid phase synthesis of the peptide was performed using a phthalimide based linker. Upon biological evaluation, KP-10-amine showed very little

activity against GPCR54, allowing us to conclude that the carbonyl portion of the *C*-terminal amide is essential and that the amine alone is not enough for receptor activation.

Subsequent to this work being carried out, an *in silico* model of GPCR54 was created[**] and docking of KP-10[††] carried out (see Chapter 2 for details). This work suggests a possible explanation for the inability of KP-10-amine to activate GPCR54, as KP-10 and KP-10-amine are believed to be orientated differently within the receptor (Figure 3-12).

In Figure 3-12, ligands are coloured green, critical atoms (involved in H-bonding) are represented as balls and coloured by type (red is oxygen, blue is nitrogen). Important H-bonds are showed as black lines. Image A shows the *C*-terminus of KP-10 complexed with GPCR54, where two adjacent Glutamine residues Q122 and Q123 are involved in hydrogen bonding with the *C*-terminal carbonyl, as well as tyrosine Y213 with the nitrogen.

Images B and D show the positioning of KP-10 (B) and KP-10-amine (C) bound to GPCR54 from the top of the receptor. It can clearly be seen that KP-10 is oriented differently than KP-10-amine, and that the hydrogen bonding network is very different. None of the residues involved in hydrogen bonding to KP-10 is maintained in the KP-10-amine complex. Here, the *C*-terminal amine group in KP-10-amine instead makes H-bonds with glutamic acid E193. It is therefore predicted to bind to GPCR54 with much less affinity than KP-10. The experimental work carried out in this chapter therefore further validates this *in silico* model.

A novel linker was developed for the attachment of amines to a support during solid phase synthesis. The aim was to develop an amine linker which would be compatible with SPPS and therefore stable to basic conditions and be acid cleavable. This aim was fulfilled by the synthesis of the ALBA linker, which was shown to be stable to all conditions required. It was also shown that cleavage could be carried out *via* two modes, either with neat TFA, or in more mild conditions using TBAF and 5% TFA in CH_2Cl_2. This linker could serve as a useful tool for amine linkage to solid support, particularly for Fmoc based peptide synthesis.

Use of phthalide compounds in the synthesis of the ALBA linker inspired the work in the following chapter. Here, phthalides were identified as starting materials in a novel synthetic route to generate rhodamine dyes in an isomerically pure form.

[**] GPCR54 model generated by Xavier Deupi, PSI, Switzerland
[††] Molecular dynamics, ligand docking and image generation carried out by Agnieszka Bronowska, HITS, Germany

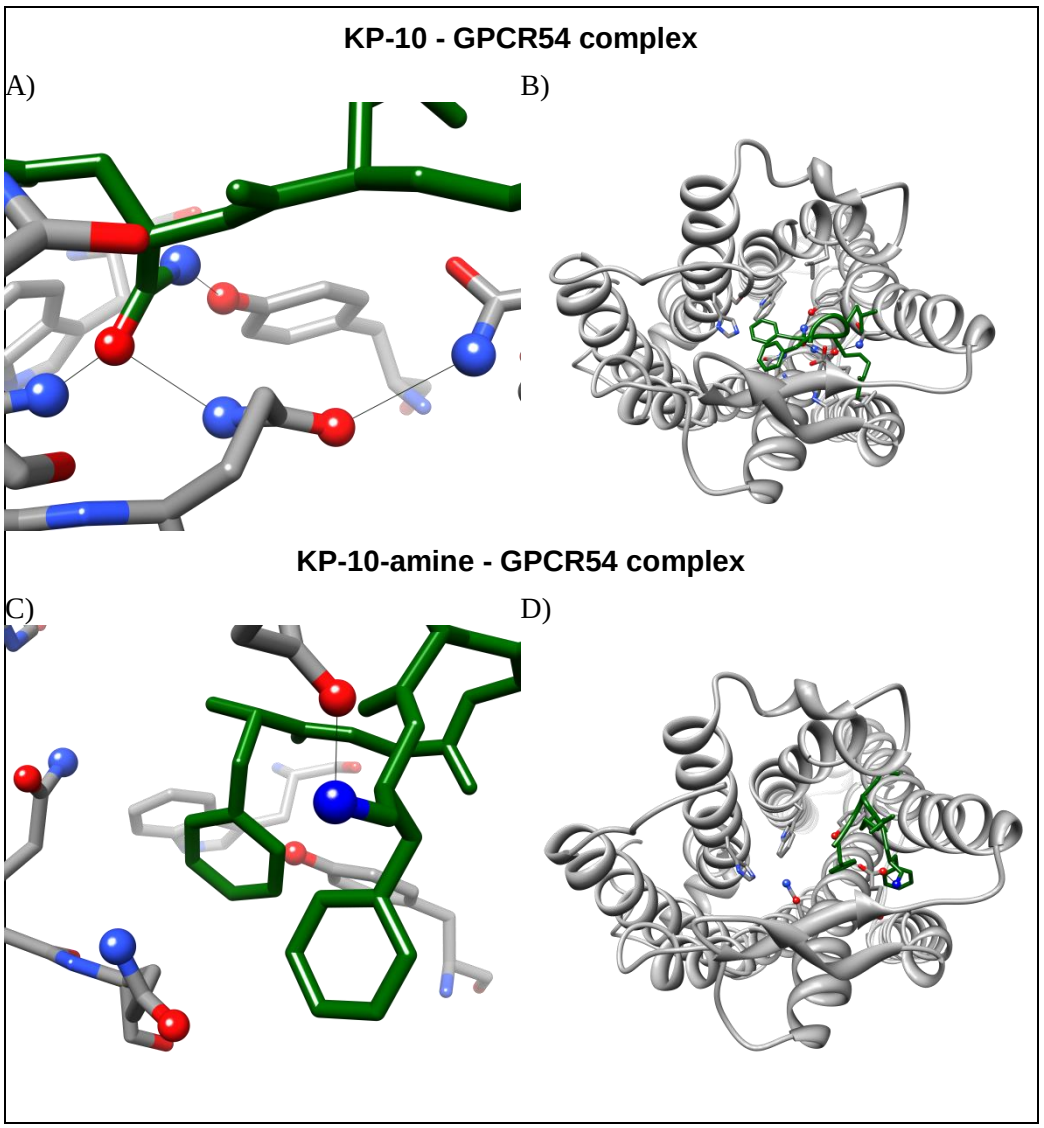


Figure 3-12: Orientation and key binding interactions of KP-10 and KP-10-amine in GPCR54 homology model

CHAPTER 4

DEVELOPMENT OF A NOVEL SYNTHETIC ROUTE TO SINGLE ISOMER FUNCTIONALISED RHODAMINE DYES

4.1 Abstract

The use of phthalide containing compounds in the production of the ALBA linker inspired a novel synthetic route to isomerically pure rhodamine dyes. The generally accepted synthetic method which uses anhydrides as starting materials generates regioisomeric products when a functional handle for attachment to chemical compounds or biomolecules is introduced. These isomers are difficult to separate and use of mixed isomers for labelling can lead to complicated analyses of final products. The removal of a carbonyl group when changing from an anhydride starting material to a phthalide also removes the possibility of isomer production. This allows simple and direct generation of a range of isomerically pure functionalised rhodamine dyes. Commercial interest in this work resulted in the submission of a patent application.

4.2 Introduction

4.2.1 Tetramethylrhodamine

Tetramethylrhodamine (TMR) is a fluorescent dye belonging to the rhodamine family, which is built around a similar xanthene scaffold (Figure 4-1). Rhodamines are highly photostable, have high absorption coefficients and high fluorescence quantum yields.[21] TMR has been utilised in a wide range of research areas, including protein labelling,[165] oligonucleotide labelling[166] and DNA sequencing.[167] An azide derivatised analogue (TMR-N_3) is used for PS/PS labelling of compound libraries to allow highly sensitive quality control (QC) and fluorescence based single molecule binding experiments following CONA screening (see Chapter 1).[9, 11]

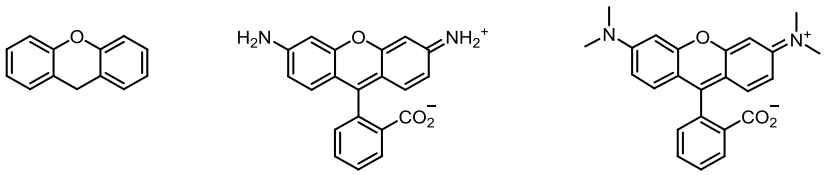
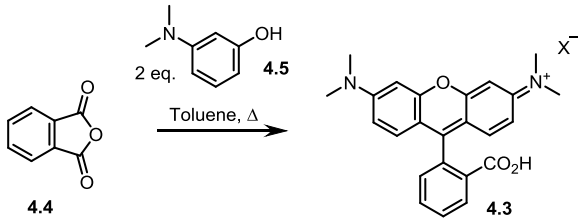


Figure 4-1: Structures of xanthene, rhodamine and tetramethylrhodamine (L to R)

4.2.2 Synthesis of TMR

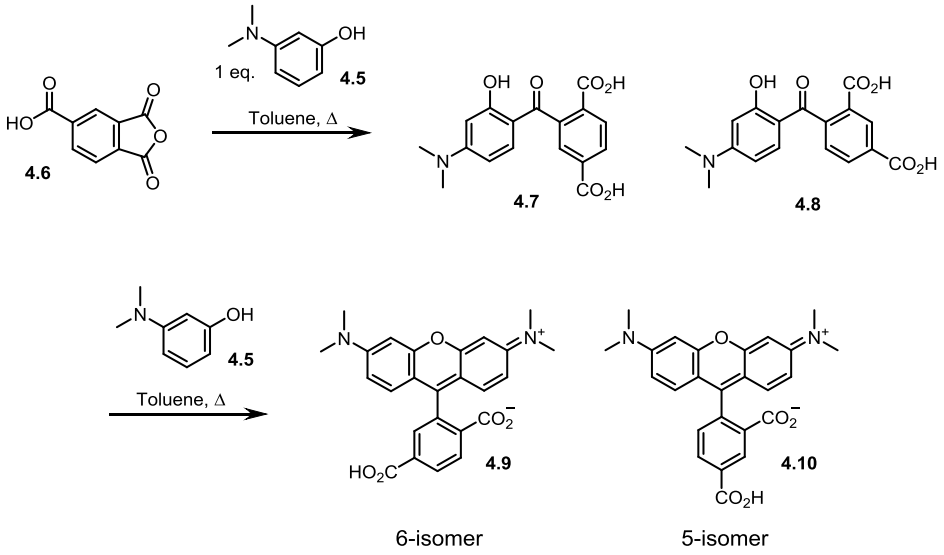
Tetramethylrhodamine is generally prepared by reacting phthalic anhydride **4.4** with two equivalents of 3-dimethylaminophenol **4.5** (Scheme 4-1). The reaction involves two

sequential Friedel-Crafts electrophilic aromatic substitutions on the anyhdride to form the xanthene core.



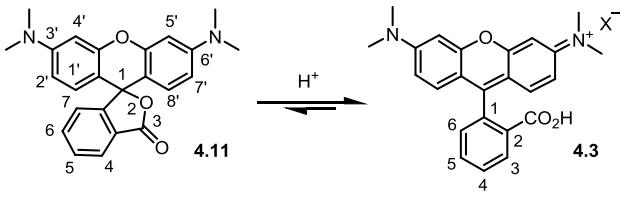
Scheme 4-1: General synthesis of tetramethylrhodamine

In order to utilise such fluorophores for labelling of biological or chemical entities, a linkage point must be introduced into the benzoic acid moiety. While this makes the compounds more useful, it also introduces complexity into the synthesis. When a functionalised anhydride is used, the aminophenol employed can attack the anhydride at either of the carbonyl carbons (Scheme 4-2). This leads to two regioisomeric isomers, **4.7** and **4.8**. Following reaction with a second equivalent of aminophenol, these isomers are converted into the two regioisomers of TAMRA, **4.9** and **4.10**. These isomers can be complicated to separate due to their structural similarity and high polarity. The scheme below exemplifies this reaction for the case of carboxytetramethylrhodamine (TAMRA).



Scheme 4-2: Introduction of regioisomers in functionalised TMR synthesis

The isomers are numbered historically from rhodamines having been drawn in a tautomeric spirolactone form in which they occur when in basic conditions, although in acidic conditions they exist in their open chain form (Scheme 4 3).[168, 169]



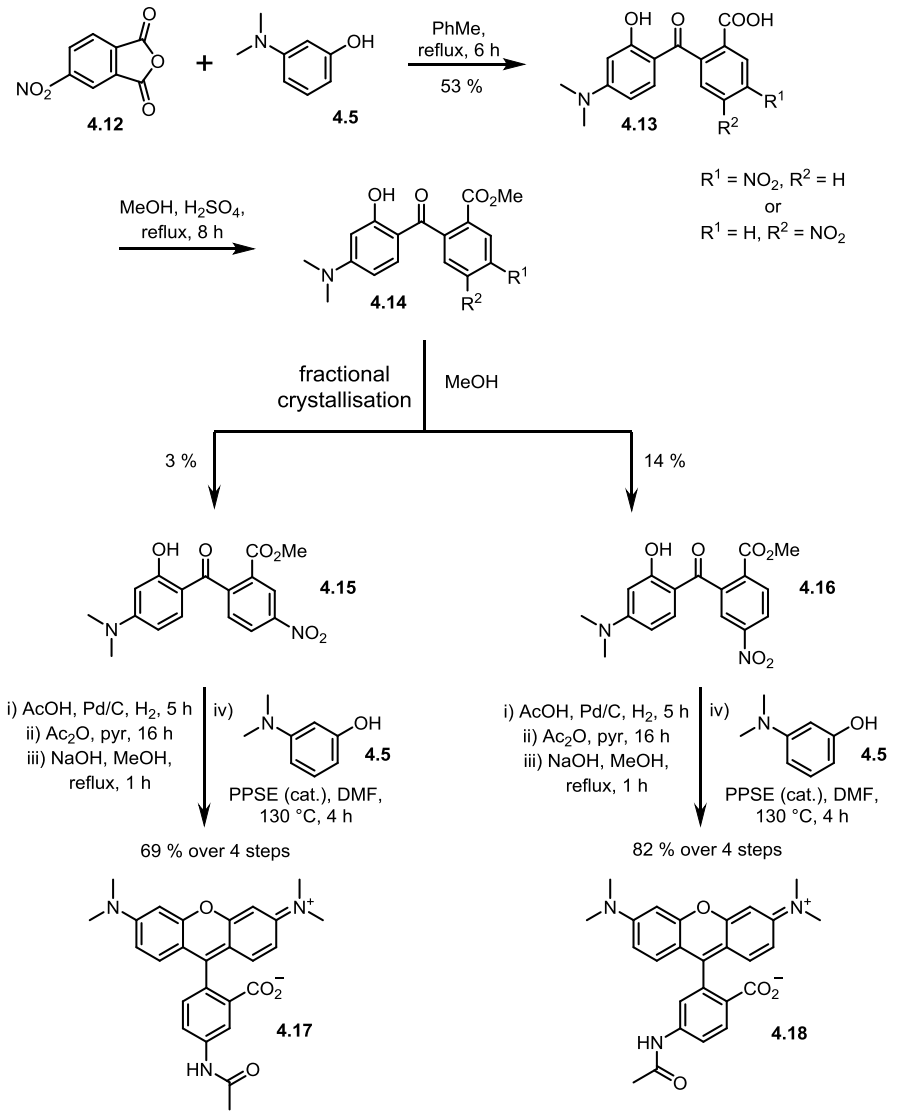
Scheme 4-3: Numbering of lactone (left) and open chain (right) form of tetramethylrhodamines

For some purposes, working with a mixture of 5- and 6- isomers is not a problem. However in some applications their difference in physical properties can cause issues. Particularly problematic is the difference in electrophoretic mobility, which can induce errors when used for sequencing applications.[170] Single isomer material is also preferable for the HPLC analysis of TMR tagged compounds, as the different retention times of the regioisomers lead to complicated traces.

4.2.3 Reported Procedures for Separation of TMR Isomers

Fractional crystallisation has been used to separate isomeric benzoylbenzoic acid intermediates (such as **4.7** and **4.8**) formed when only one equivalent of aminophenol is reacted with a substituted anhydride. Following separation, these can be converted into isomerically pure tetramethylrhodamines *via* reaction with another equivalent of aminophenol.

Corrie *et al.* used this method to prepare isomerically pure 5- and 6-aminotetramethylrhodamine (Scheme 4-4).[168] Nitrophthalic anhydride was refluxed with equimolar 3-dimethylaminophenol in toluene for 6 hours to generate the corresponding nitro substituted benzoylbenzoic acid intermediates **4.13** in 53% yield. Refluxing these in methanol with catalytic sulphuric acid gave esters **4.14**. Fractional crystallisation using methanol was employed to separate the isomers. However the author described this process as ‘laborious and inefficient’, with yields of 14% of the 5- and 3% of the 6- isomer isolated.

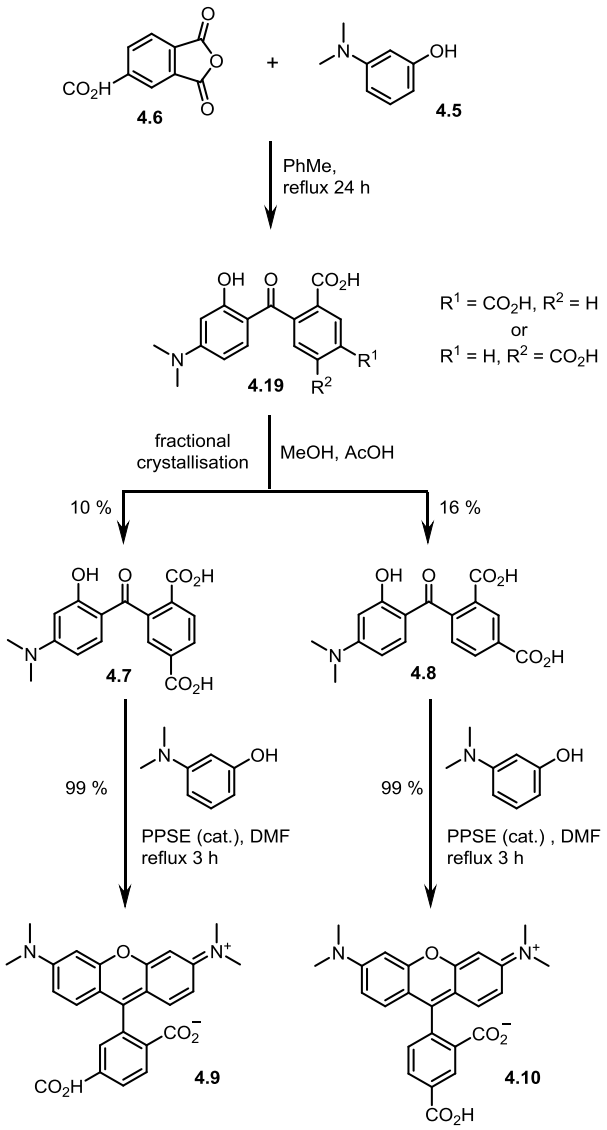


Scheme 4-4: TMR isomer separation as described by Corrie *et al.*[168]

Following crystallisation, reduction of each nitro intermediate using Pd/H_2, followed by acetylation of the resulting amine groups and saponification of the ester groups were carried out. This allowed the rhodamine forming Friedel Crafts reaction to be completed by refluxing in DMF with catalytic trimethylsilyl polyphosphate to give isomerically pure acetamido tetramethylrodamine compound **4.17** in 69% yield and **4.18** in 82% yield. These compounds were derivatised further to generate thiol reactive iodoacetamido tetramethyl rhodamine derivatives.

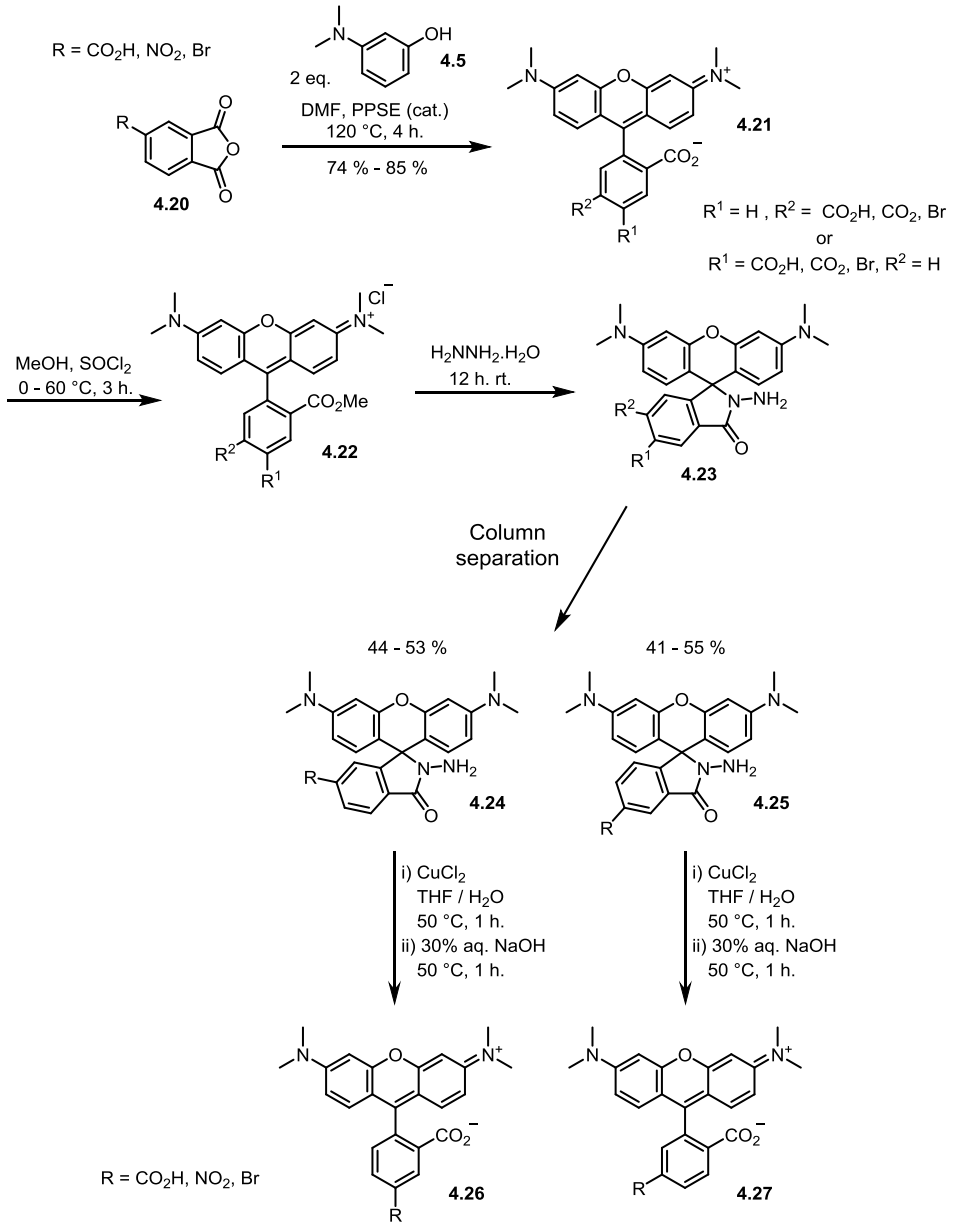
Kvach *et al.* used a similar method to generate isomerically pure TAMRA through fractional crystallisation of its benzoylbenzoic acid intermediates (Scheme 4-5).[171] They reported that their crystallisation method enabled gram-scale amounts of TAMRA to be prepared. In order

to achieve this they heated equimolar amounts of phthalic anhydride with 3-dimethylaminophenol in toluene overnight to generate the intermediates **4.19** as mixed isomers. The isomers were then separated by fractional crystallisation from methanol and acetic acid, yielding 10% and 16% for the 5- and 6- isomers respectively, although the authors do note that a 20% yield of mixed isomer material was recovered and could be used for further crystallisation. Following separation, the benzoyl benzoic acids were refluxed with another equivalent of 3-dimethylaminophenol and catalytic trimethylsilylpolyphosphate in DMF to afford isomerically pure carboxytetramethylrhodamines **4.9** and **4.10**, which were precipitated in acidic aqueous conditions. The lack of column chromatography in this synthesis and its scalability makes it appear appealing, however low yields and protracted crystallisation methods detract from these advantages.



Scheme 4-5: TMR isomer separation using fractional crystallisation as described by Kvach *et al.*[171]

Yu and co-workers took a different approach to the preparation of single isomer substituted TMR analogues. Their method involves the synthesis of mixed isomer TMR bearing a variety of substitution groups, followed by manipulation of the products to generate hydrazine spirolactam derivatives **4.25**, which they report can be separated using column chromatography (Scheme 4-6).[172]

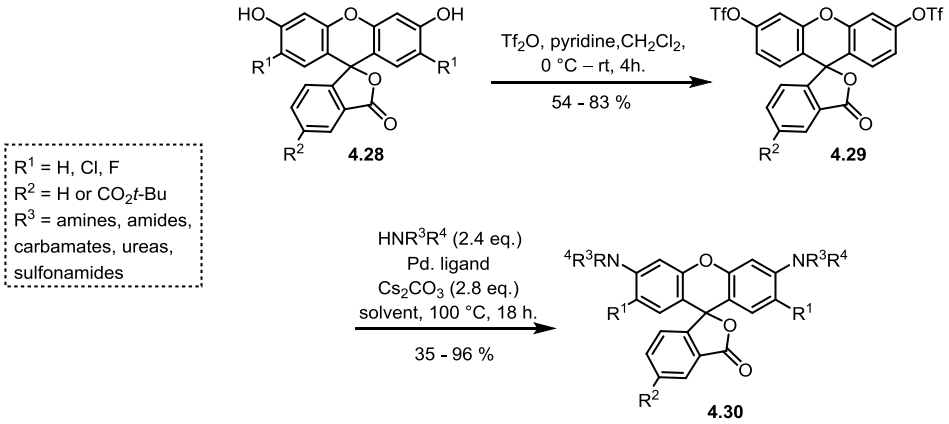


Scheme 4-6: TMR isomer separation *via* column chromatography of spirolactams as described by Yu *et al.* [172]

Following rhodamine synthesis, the corresponding spirolactams were prepared by conversion of the 1-carboxylic acid into methyl ester **4.22** by heating in methanol in the presence of thionyl chloride. Subsequent displacement of the ester using hydrazine hydrate encouraged formation of the cyclic lactam product **4.23**. The authors claim that these isomeric spirolactams are easier to separate than their open chain counterparts by column chromatography, due to their neutral state which results in lower polarity and better solubility in organic solvents. Following isomer separation, copper chloride allowed ring opening to occur and subsequent hydrolysis using sodium hydroxide gave the required

tetramethylrhodamine products. This method was demonstrated as being more universally applicable than crystallisation methods by the generation of isomerically pure bromo-, nitro- and carboxytetramethylrhodamine, with overall yields reported between 31 and 45%.

Recently, the Lavis group reported an elegant synthetic strategy where a variety of single isomer rhodamines bearing a 5-carboxylic acid linkage group were synthesised directly *via* conversion of isomerically pure trifluoromethane sulfonate functionalised fluoresceins **4.29** into rhodamines using palladium mediated C-N cross coupling (Scheme 4-7).[173] Known and novel rhodamines were synthesised using this method, which reports yields ranging between 54% and 83%.



Scheme 4-7: Rhodamine synthesis by palladium mediated cross coupling as reported by the Lavis group.[173]

A wide array of amines were used to demonstrate this method, including symmetric and asymmetric secondary linear amines, aliphatic amine containing cycles (e.g. pyrrolidine, piperidine) and nitrogen containing aromatic ring systems (e.g. pyrrole and indole). Also demonstrated to couple well with the trifluoromethane sulfonate functionalised fluoresceins were amides, carbamates and other nitrogen nucleophiles such as hydroxylamines. However, only six of the thirty two examples shown were exemplified with the *tert*-butyl protected carboxylic acid functionality installed in the pendent ring.

Whilst useful, this method does not allow the generation of X-rhodamine and Q-rhodamine, where the xanthene anilines are substituted with fused cyclic rings (Figure 4-2). The use of palladium cross coupling also precludes the use of some halogen functionality on the pendent aromatic ring. The elevated temperatures employed in this reaction also render volatile amines such as dimethylamine (used to generate tetramethylrhodamine) impractical to use on a large scale.

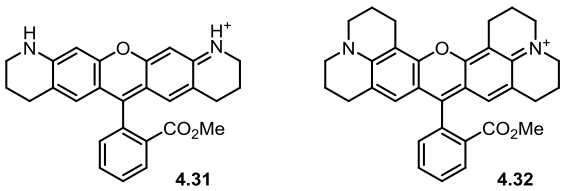
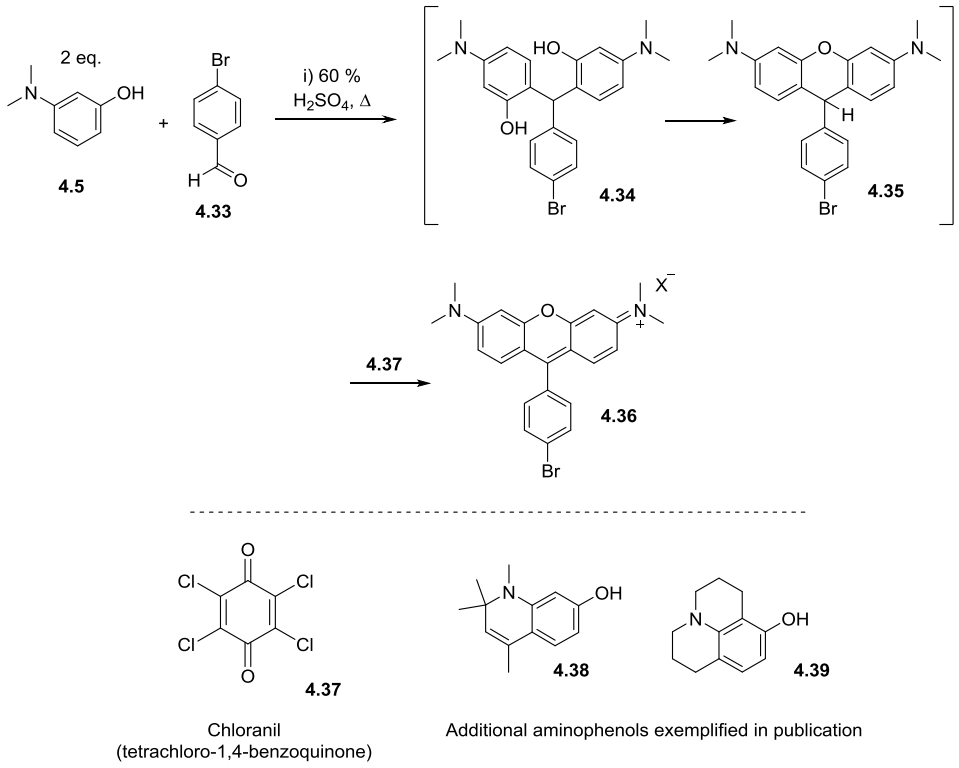


Figure 4-2: Structures of Q-rhodamine and X-rhodamine

4.2.4 Synthesis of Rosamines

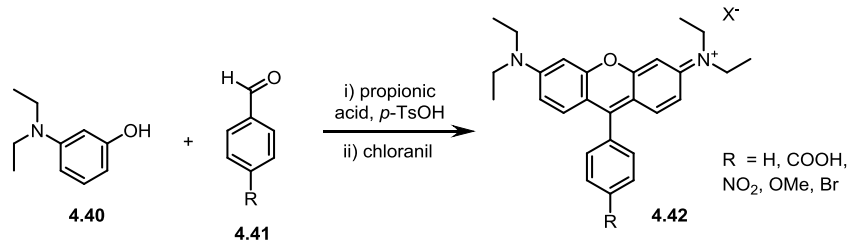
Rosamines are dyes which have a similar structure to rhodamines, but the acid group on the pendent ring is absent. They possess slightly less favourable properties than rhodamines, such as lower solubility and fluorescence in aqueous solvents. Jiao and co-workers describe the synthesis of three examples *via* reaction of a *para*-brominated benzaldehyde with various aminophenols in 60% sulphuric acid with heating. During this reaction, two aminophenol compounds perform Friedel Crafts electrophilic aromatic substitution at the aldehyde carbon, generating diphenol intermediate **4.34**, which can react in an intramolecular fashion to generate **4.35** (Scheme 4 8).[174] Chloranil **4.37** (or tetrachloro-1,4-benzoquinone), a mild oxidant and hydride acceptor which is often used in aromatisation reactions,[175] is employed to generate the desired aromatic rosamine product **4.36**.



Scheme 4-8: Bromo-rosamine synthesis according to Jiao *et al.* [175]

The use of an aldehyde starting material in this work means that there is only one point available for aminophenols to react (rather than two in the case of an anhydride). As such, only a single isomer is formed during synthesis of these dyes. Both conventional heating and microwave irradiation were used to generate three examples in 5 – 73% yield, with microwave irradiation giving the best results.

Cardoso *et al.* extended this work to generate rosamines by bearing various linkage points, including carboxyl-, nitro- and bromo- by reacting 3-diethylaminophenol with various benzylic aldehydes (Scheme 4-9).[176] The nitro group was subsequently reduced using Pd/H_2 to give the corresponding amine.



Scheme 4-9: Synthesis of rosamines according to Cardoso *et al.*

Here, the reactions were carried out at high temperatures in propionic acid with catalytic *p*-toluenesulfonic acid, followed by addition of chloranil. Yields reported varied from 26 – 51%, again with higher yields observed when microwave heating was employed.

4.2.5 Summary

Although methods for the separation of functionalised TMR compounds have been reported in the literature, they possess disadvantages. Fractional crystallisation conditions are substrate specific and cannot be applied directly to other analogues or species, therefore the applicability of each reported method is limited to the molecule around which it was optimised. Establishing conditions for a number of different fluorophores or tetramethylrhodamine derivatives would likely be a time consuming and arduous process.

While generating spirolactam derivatives of functionalised TMR compounds allows separation of isomers by column chromatography, this introduces multiple synthetic steps, which lowers yield and increases the number of purification steps.

Rosamine synthesis removes the problem of generating regioisomers altogether, however the properties of these dyes are inferior compared to rhodamines. Nevertheless, inspiration could

be taken from their synthesis, as the aldehydes used may be applicable to rhodamine synthesis.

4.3 Aims

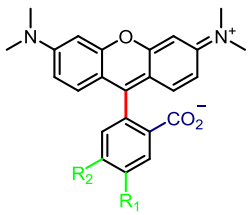
The lack of a reliable and simple method for rhodamine isomer separation results in a very high commercial value for single isomer TMR derivatives (5-Carboxytetramethylrhodamine = £250.50/10 mg[‡‡]), meaning that purchasing in large amounts is very costly. As these reagents are crucial components in the PS/PS screening platform, a robust and reproducible method for generating single isomer substituted tetramethylrhodamines and other similar dyes would be highly beneficial. An examination into a whether a synthetic route could be established which would produce single isomer rhodamine material was undertaken. Such a development would eliminate the need for arduous separation of mixed isomers.

4.4 Results and Discussion

4.4.1 Synthetic Route Planning

As the use of anhydrides is the cause of regioisomer formation in classical rhodamine synthesis, possible alternative starting materials were assessed. Such a compound would be required to fulfil the following criteria:

- Central aromatic ring
- Point of regioselective (or chemo selective) reactivity to attach amino phenols (red)
- Carboxylic acid ortho to xanthene core attachment point (blue)
- Functionality to link final compound to biomolecules/other chemical species in either the 5- or 6- position (green)



Having identified the features required, it was decided that attachment of aminophenol moieties would be achieved *via* reaction with an aldehyde, as in the synthesis of rosamines. A retrosynthetic analysis was carried out, in which substituted phthalides were identified as possible reagents for this synthesis (Figure 4 3).

[‡‡] Price taken from Sigma Aldrich website December 2014

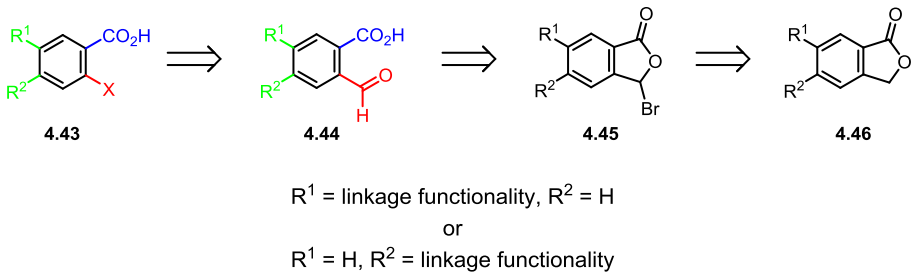
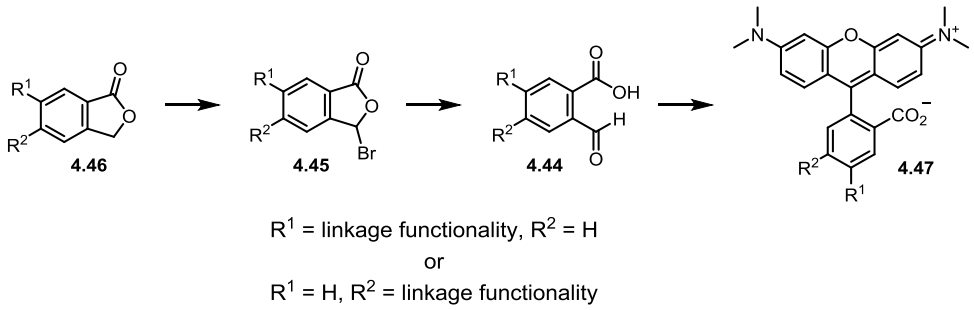


Figure 4-3: Retrosynthetic analysis of required building block for TMR synthesis

The transformation of phthalides into phthalaldehydic acids is well documented, and involves bromination of the phthalide in the 3-position then subsequent hydrolysis of the brominated species to give the phthalaldehydic acid (Scheme 4-10).[177, 178] This species was predicted to contain the necessary functionalities to enable a double Friedel-Crafts reaction to be carried out in order to obtain the desired rhodamine compounds.



Scheme 4-10: Planned synthesis of TMR compounds

Having identified a promising starting point, the commercial availability of functionalised phthalides was explored. Gratifyingly, a range of inexpensive phthalides with functional groups pre-installed were commercially available as both the 5- and 6- isomers (Figure 4-4), allowing rapid progression onto testing of the planned synthetic route.

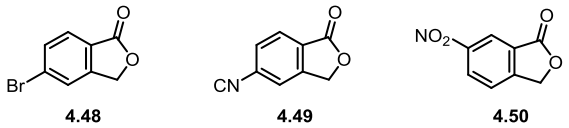
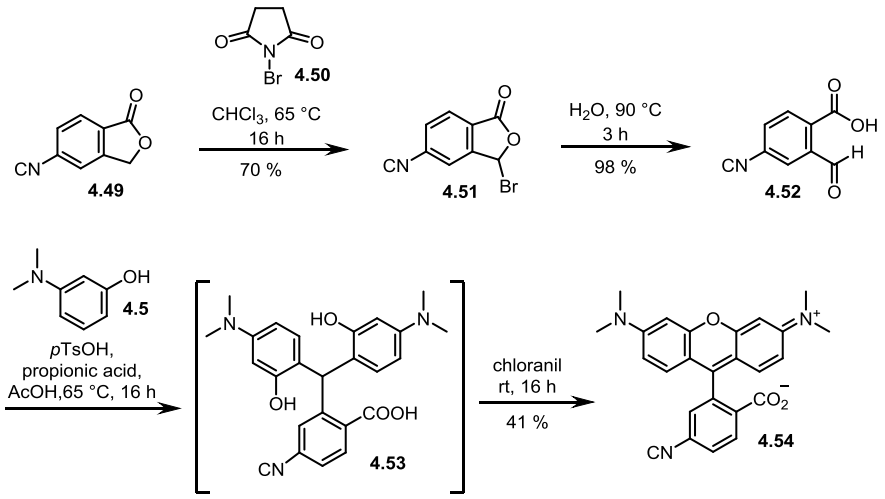


Figure 4-4: Some commercially phthalides, from the left: 5-bromomophthalide; 5-cyanophthalide and 6-nitrophthalide

4.4.2 Proving the Synthetic Route – Synthesis of 6-CN-TMR

5-Cyanophthalide was chosen as the first example to be assessed in this synthesis, as following rhodamine generation, the pendent nitrile group could be hydrolysed, providing access to TAMRA. The synthesis carried out is shown in Scheme 4-11.

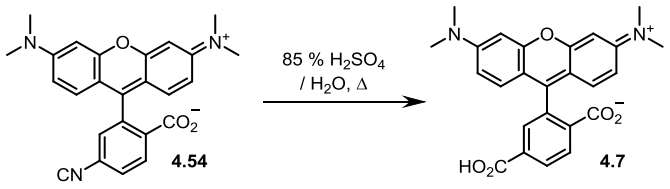


Scheme 4-11: Synthesis of 6-CN TMR from 5-cyanophthalide

Bromination of 5-cyanophthalide was carried out in good yield by heating with *N*-bromosuccinimide (NBS) in chloroform overnight to give **4.51**. The bromide was hydrolysed by heating in water to give the corresponding phthalaldehydic acid **4.52** in excellent yield. This was then reacted with two equivalents of 3-dimethylaminophenol in the presence of catalytic *p*-TsOH in propionic acid at 65 °C overnight, conditions similar to those used for rosamine synthesis.[176] LC-MS analysis of the reaction mixture indicated that the formation of intermediate **4.53** had occurred, which was oxidised *in situ* using chloranil to give final product 6-cyanotetramethylrhodamine (6-CN-TMR, **4.54**) in 41% yield. The isolation of this compound proved that the planned synthesis worked and that phthalides could indeed be used in order to access a single isomer functionalised rhodamine compound. Assessment of the route to generate useful rhodamines, such as carboxytetramethylrhodamine (TAMRA), was next investigated.

4.4.3 Generation of TAMRA from 6-CN-TMR

6-CN-TMR was refluxed in aqueous sulfuric acid with the aim of generating 6-TAMRA **4.7** *via* nitrile hydrolysis (Scheme 4-12). Completion of the reaction took two days, during which time some decomposition of the compounds was observed. Isolation of clean material therefore required another chromatographic purification, which proved difficult due to the charged and polar nature of TAMRA. The extended reaction time required in these relatively harsh conditions raised concerns about possible damage to the integrity of the dye. Due to these reasons, an alternative approach to the generation of TAMRA was therefore taken.

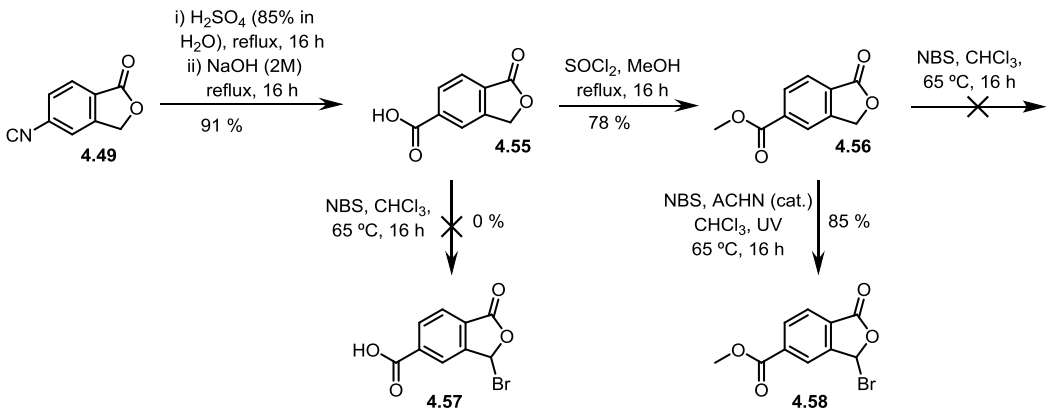


Scheme 4-12: Hydrolysis of 6-CN-TMR to generate 6-TAMRA

4.4.4 Synthesis of 6-TAMRA from 5-Carbomethoxyphthalide

In order to alleviate the problems faced in the hydrolysis of 6-CN-TMR, the desired carboxylic acid linkage group was instead introduced through hydrolysis of the phthalide starting material, 5-cyanophthalide. This conversion was carried out over 2 steps in excellent yield (Scheme 4-13). Bromination of the resulting 5-carboxyphthalide was attempted, however insolubility of this compound prevented the reaction from occurring. Solubility was increased by conversion of the free acid to the corresponding methyl ester to give **4.56** by heating to reflux in the presence of thionyl chloride. Disappointingly, no product was formed when bromination of the methyl ester was attempted under the conditions used previously.

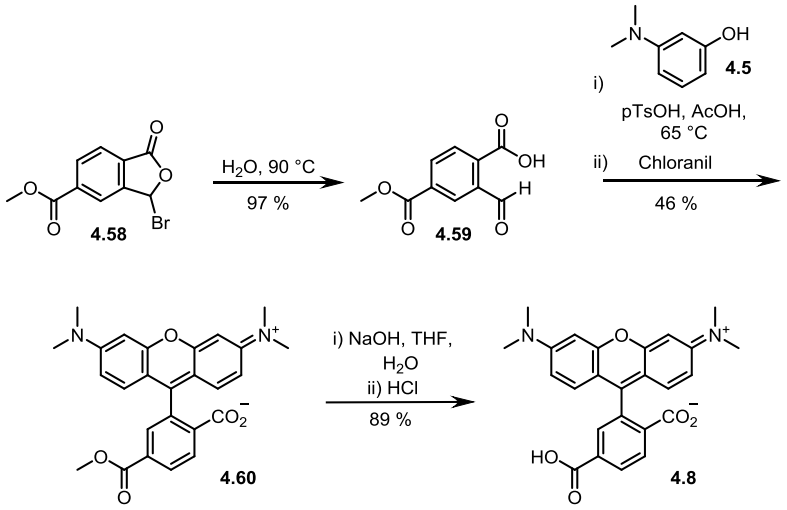
The use of radical initiators has been reported to aid the formation of radical species in allylic bromination reactions such as this one.[178] The reaction was repeated, this time with the addition of radical initiator 1,1′-azobis(cyclohexanecarbonitrile) (ACHN) and irradiation with UV light (254 nm, 4 Watt) throughout. Under the new conditions, brominated 5-carbomethoxyphthalide **4.58** was isolated in 85% yield.



Scheme 4-13: Bromination of 5-carboxyphthalide was unsuccessful, but worked well with the 5-carbomethoxyphthalide

With the brominated methyl ester bearing compound in hand, the rest of the synthesis was carried out (Scheme 4-14). Formation of phthalaldehydic acid **4.59** was achieved in excellent

yield. Subsequent formation of the rhodamine structure by reaction with aminophenol **4.5** yielded 6-CO_2Me-TMR **4.60** in 46% yield.



Scheme 4-14: Synthesis of 6-TAMRA from 5-methylcarboxyphthalide

Generation of 6-TAMRA *via* hydrolysis of the ester group proceeded smoothly within three hours using sodium hydroxide. Subsequent acidification of the reaction mixture with hydrochloric acid caused the product to precipitate. Filtration of the solid gave 6-TAMRA **4.8** in 89% yield, with no sign of degradation and therefore no extra chromatography required.

4.4.5 Analysis of 6-TAMRA

Analysis of the synthesised 6-TAMRA was carried out to ensure isomeric purity of the prepared material. The ^{1}H NMR spectrum of synthesised 6-TAMRA (Figure 4-5) indicates the presence of only one isomer. In the aromatic region of the spectrum, proton integrations correlate exactly to the desired compound, with no evidence of another species present.

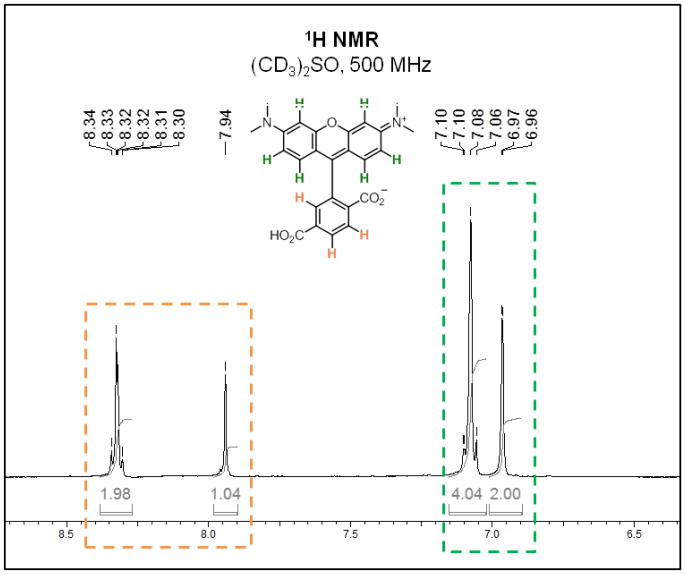


Figure 4-5: Aromatic region of ^{1}H NMR spectrum of synthesised 6-TAMRA

A sample of the synthesised 6-TAMRA was spiked with previously prepared 5-TAMRA and analysed by RP-HPLC (Figure 4 6).[§§] The trace shows clean material, with the two isomers eluting at different retention times, indicating that they are different chemical species. These analyses confirm the isomeric purity of the synthesised compound and its identity as the 6-isomer.

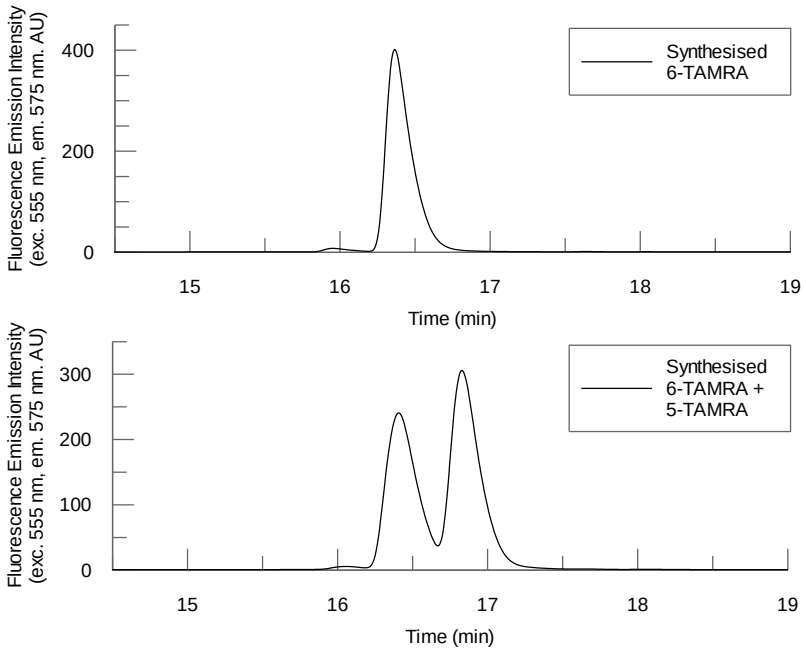


Figure 4-6: RP HPLC analysis of synthesised 6-TAMRA (above) and spiked with 5-TAMRA (below)

[§§] 5-TAMRA synthesised by Dr. Angus Brown

With a method to generate isomerically pure 6-TAMRA clearly defined, the scope of this reaction was next examined.

4.4.6 Production of additional TMR Analogues

To ascertain whether the developed reaction conditions were amenable to other useful functional groups on the TMR pendent ring, two additional commercially available phthalides were examined. 5-Bromophthalide and 6-nitrophthalide were assessed in the synthesis. Both enabled their corresponding tetramethylrhodamine analogues to be generated. Both 5-Br-TMR (**4.61**) and 6-NO_2-TMR (**4.62**) were obtained in similar yields to those of earlier examples, without the need for optimisation of any reaction conditions (Figure 4-7, Table 4-1).

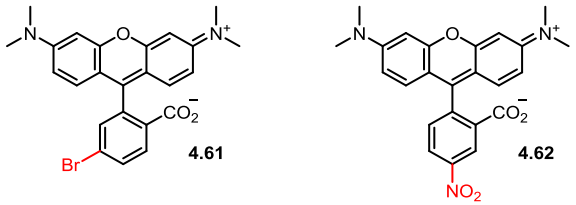
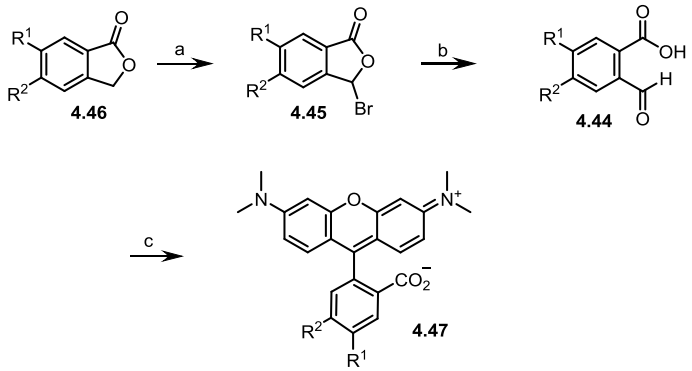


Figure 4-7: 6-Br-TMR and 5-NO_2-TMR were produced using the developed synthetic route

A review of reactions carried out using this synthetic route shows that each step demonstrates consistency in reproducibility and yield, despite variation in linkage localisation (5- or 6- position) and their chemical composition (Table 4-1).

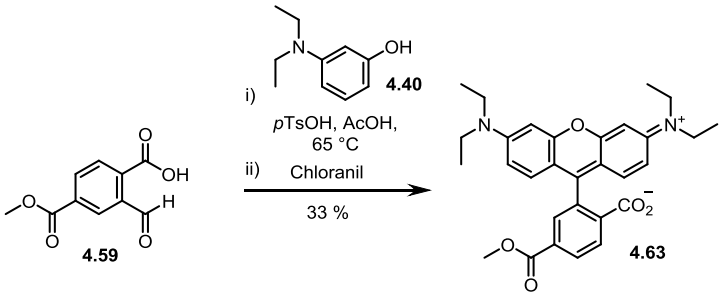


Entry	R^1	R^2	Step Yield (%) a	b	C
1	H	CN	70	98	41
2	H	Br	97	90	37
3	H	CO_2Me	85	97	46
4	NO_2	H	61	85	36

Table 4-1: Yields of TMR analogues synthesised

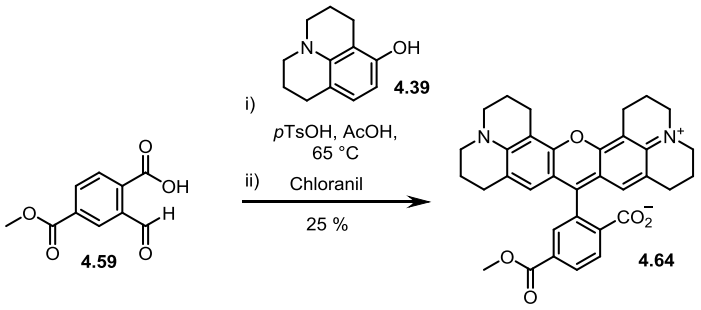
4.4.7 Production of Additional Rhodamines

To establish whether rhodamines other than tetramethylrhodamine analogues could be produced using this method, two different amino phenols were employed. In the first example, previously prepared intermediate 2-formyl-4-methoxycarbonyl-benzoic acid **4.59** was reacted with 3-diethylaminophenol with the view to generating 6-CO_2Me functionalised rhodamine B (Scheme 4-15).



Scheme 4-15: Synthesis of 6-CO_2Me-Rhodamine B

This reaction proceeded smoothly to give compound **4.63**, although in slightly lower yield than that of the TMR derivative. Encouraged by this, 8-hydroxyjulolidine **4.39** was reacted with 2-formyl-4-methoxycarbonyl-benzoic acid **4.59** under the same conditions to give 6-CO_2Me functionalised X-rhodamine **4.64**, albeit in low yield (Scheme 4-16).

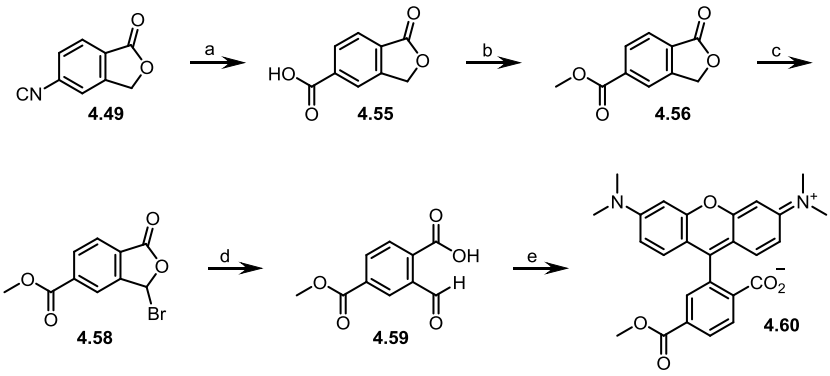


Scheme 4-16: Synthesis of 6-CO_2Me-X-Rhodamine

These successful reactions demonstrate the generality of this synthetic route, not only for the production of isomerically pure tetramethylrhodamines, but also for other members of the rhodamine family.

4.4.8 Large Scale Production of 6-TAMRA

In the development of this synthetic route, syntheses were generally carried out on a 0.5 mmol scale. In order to assess the utility of this route, a larger scale synthesis was undertaken (Table 4-2).[***]



a) i) H_2SO_4 (85% in H_2O), reflux, 16 h; ii) NaOH (2M), reflux, 16 h; b) $SOCl_2$, MeOH, reflux, 16 h;
c) NBS, ACHN (cat.), $CHCl_3$, UV, 65 ºC, 16 h; d) H_2O, 90 ºC, 3 h.;
e) i) 2 eq. 3-dimethylaminophenol, pTsOH (cat.), propionic acid, 65 ºC, 16 h; ii) chloranil, rt, 16 h

[***] All large scale synthesis work carried out in collaboration with Irene Pérez Pi

Entry	Step	Scale (mmol)	Yield (%)	Yield (0.5 mmol scale)
1	a	125	96	91
2	b	125	90	78
3	c	62	83	85
4	d	62	89	97
5	e	15	38	46
6	**Overall**		**24**	**27**

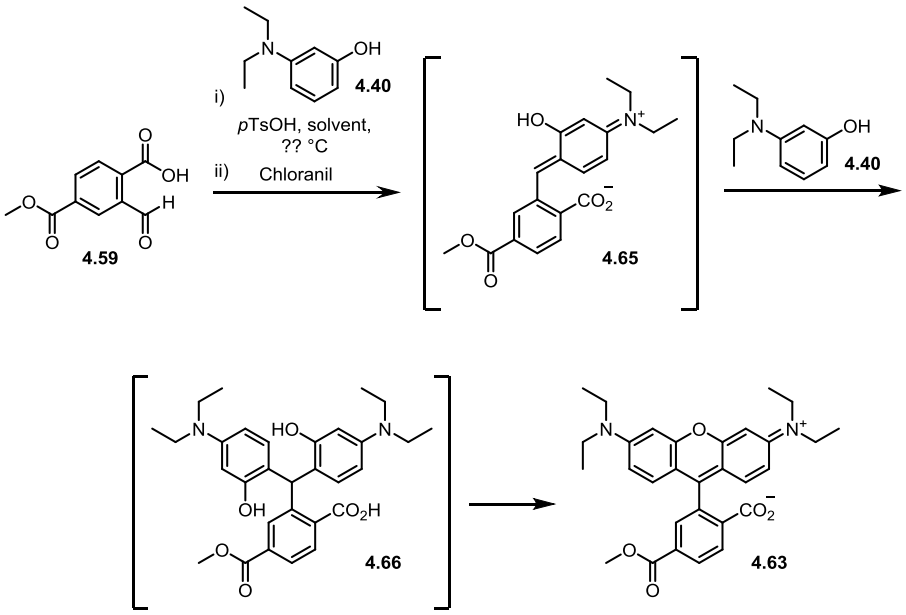
Table 4-2: Yield comparison for development and scale up syntheses of 6-TMR-CO_2Me

The comparable yields of large scale synthesis with the small scale conversions shown in Table 4-2 demonstrates the suitability of this process to gram scale reactions. Generation of 5-carbomethoxyphthalide from 5-cyanophthalide was executed in excellent yield with no purification steps required. Column chromatography was required following bromination, however separation was easily achieved. Subsequent hydrolysis of brominated phthalide **4.58** occurred in excellent yield. The final rhodamine forming step proceeded with slightly lower yield compared to the small scale reaction, however following a final chromatography purification, almost a gram of analytically pure 6-CO_2Me-TMR was afforded, which corresponds to a commercial value of over £12, 000.[†††]

4.4.9 Synthesis Optimisation

Although the established reaction conditions allow access to various single isomer rhodamines in up to gram scale amounts, it was thought that the moderate yields of the final rhodamine forming step could be improved upon. A screen of conditions was conducted using phthalaldehydic acid intermediate **4.59** and 3-diethylaminophenol **4.40**, with solvent, temperature and reagent amounts being varied (Table 4-3).

[†††] Price taken from Sigma Aldrich website December 2014



Reaction no.	Solvent	Equiv. aminophenol	Temperature (°C)
1	Propionic acid†	2	60
2	Propionic acid†	3	60
3	Propionic acid†	5	60
4	Propionic acid†	10	60
5	Acetic acid†	3	60
6	Methanesulfonic acid	3	60
7	60% HCl	3	60
8	60% H_2SO_4	3	60
9	Propionic acid†	3	80
10	Propionic acid†	3	100

Table 4-3: Conditions used in rhodamine synthesis optimisation array.
Reactions carried out on a 0.24 mmol scale. †0.2 eq. *p*-TsOH added to these reactions

Reaction mixtures were analysed by LC-MS following 16 hours of heating.

4.4.9.1 Solvent Screen

Methanesulfonic acid, 60% HCl and 60% H_2SO_4 proved ineffective as solvents due to insolubility of reagents (conditions 6, 7 and 8). Acetic acid showed less conversion of starting material into the diphenol intermediate **4.66** within the time given. We therefore concluded that the use of propionic acid as a solvent was optimal.

4.4.9.2 Quantity of Reagents

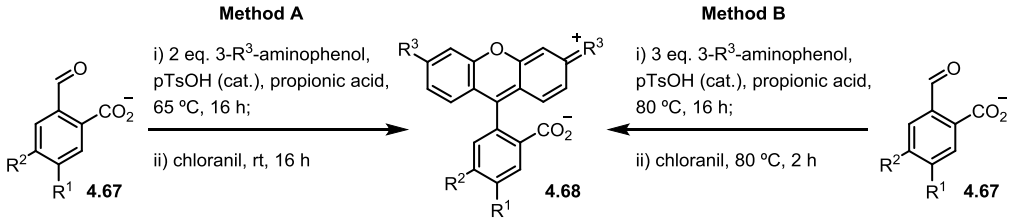
Increasing the equivalents of 3-diethylaminophenol **4.40** from two to three increased the amount of diphenol intermediate **4.66**. However, any increase past this was not of any significant benefit.

4.4.9.3 Temperature

Increasing the temperature of the reaction to 80 °C resulted in a higher amount of diphenol intermediate **4.66** being formed. This also appeared to produce a significant amount of final rhodamine product during the reaction, suggesting that conversion of the diphenol intermediate into rhodamine may be more favourable at higher temperatures. Heating to 100 °C did not show any beneficial effects beyond this.

4.4.9.4 Chloranil Addition

A final reaction was performed which included the use of 3 equivalents of 3-diethylaminophenol and heating at 80 °C. When the addition of chloranil was also carried out at 80 °C, Rhodamine B compound **4.63** was isolated in 70% yield. This represented a great increase in the efficiency of the reaction compared to the original conditions used, which gave 33% yield. These new conditions were therefore employed with all previously synthesised rhodamines. The results are summarised in Table 4-4.



Entry	R^1	R^2	R^3	Cpd no	Yield Method A (%)	Yield Method B (%)
1	H	CO_2Me	NEt_2	**4.63**	33	70
2	H	CO_2Me	Piperidine	**4.64**	25	47
3	H	CO_2Me	NMe_2	**4.60**	46	51
4	H	CN	NMe_2	**4.54**	41	50
5	H	Br	NMe_2	**4.61**	37	54
6	NO_2	H	NMe_2	**4.62**	36	57

Table 4-4: Yields of rhodamine forming step from original conditions and newly developed conditions

A large increase in yield was also seen for the X-rhodamine analogue (entry 2) which went from 25% to 47% under the new conditions. All tetramethylrhodamine derivatives also saw an increase in yield, though these were more modest.

4.5 Conclusions and Future Work

The aim of this work was to generate a robust and repeatable synthesis for the generation of single isomer functionalised rhodamine derivatives, which has been achieved. The developed

synthesis has been proven suitable for the generation of tetramethylrhodamines bearing a diverse range of useful functional handles in either the 5- or 6- position. It has also been shown that various rhodamines can be generated from the phthalaldehydic acid intermediates prepared and that the process is amenable to large scale syntheses.

In comparison to existing methods of generating single isomer TMR compounds, this method holds some advantages and some disadvantages. One drawback is the number of transformations required to generate the phthalaldehydic acid intermediate. From the starting material, four steps are undertaken before TAMRA synthesis can be carried out, after which there is a further deprotection step. A redeeming factor is that only one of these steps requires purification.

The moderate yield of the TMR forming step was addressed during this work, however additional development work could increase this yield further. One option to investigate could be the use of use microwave heating, which was reported to give higher yields than conventional heating in the synthesis of rosamines.[174, 176]

The main advantage of this method is that the need for separation of isomers is circumvented, which is an arduous and time consuming task. This is reflected in the commercial pricing of TAMRA, where single isomer derivatives cost up to ten times that of mixed isomers (Table 4-5).

Compound	Price/10 mg
Mixed 5/6-Carboxytetramethylrhodamine	£14.95
5-Carboxytetramethylrhodamine	£250.50
6-Carboxytetramethylrhodamine	£127.50

Table 4-5: Prices of mixed isomer and isomerically pure TAMRA‡‡‡

This method allows fine tuning of final dye compounds. Phthalides are inert to many chemical conditions, therefore transformations can be carried out on pre-installed groups of commercially available materials, as demonstrated in the case of the conversion of 5-cyanophthalide to 5-carbomethoxyphthalide for TAMRA generation. This approach also allows flexibility in the localisation of substituents on the final TMR compound, which can be chosen from the 5- or 6- position. Also, by employing different aminophenols, various rhodamine types can be synthesised from the same phthalaldehydic acid intermediates.

All of this, in conjunction with consistent yields and high reproducibility, renders this method a highly useful and valuable tool in the synthesis of single isomer functionalised

‡‡‡ Prices taken from Sigma Aldrich website December 2014

rhodamine dyes. Future work should focus on the generation of novel rhodamine compounds which would be inaccessible by other synthetic routes. For example, compounds containing bioorthogonal groups, ‘switch on’ dyes, which only become fluorescent upon reaction with another species or novel dyes for applications such as super resolution microscopy.

The production of gram amounts of isomerically pure rhodamine dyes allows its use in myriad projects within the Auer lab. Particularly useful is its use as a reporter group in fluorescently labelled ligands and TOBOC libraries, detailed in chapter 1. It was also used in the development of various multifunctional reagents synthesised using a solid supported trifunctional orthogonal scaffold, which will be described in the following chapter.

CHAPTER 5

TOBA RESIN FOR GENERATION OF TRIFUNCTIONAL REAGENTS

5.1 Abstract

The development of a novel synthetic route to access large amounts of isomerically pure rhodamine dyes allowed its use as a reporter group in molecules for multiple purposes. For example, in this chapter it is used in a trifunctional reagent to aid the development of a bead-based method to probe protein-protein interactions. Here, there was a need for a molecule to enable simultaneous fluorescent labelling and immobilisation of target proteins. In this chapter, an orthogonally protected trifunctional backbone is developed on solid-phase to allow the generation of trifunctional reagents. The use of this functionalised resin is exemplified by the synthesis of a TMR-maleimide-N_3 tag, which allowed simultaneous fluorescent labelling and immobilisation of an alpha-synuclein mutant containing only one cysteine.

5.2 Introduction

5.2.1 Trifunctional Reagents

Trifunctional reagents are molecules which possess three different reactive groups, in most cases one being a conjugating group and one being a reporter group (Figure 5-1).

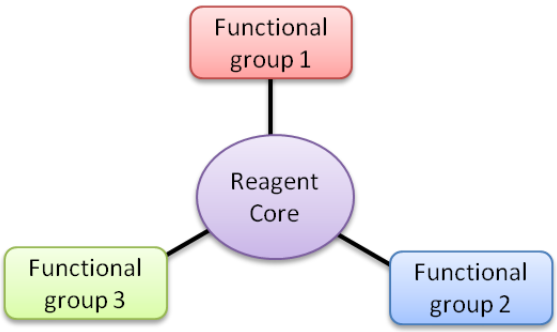


Figure 5-1: Trifunctional reagent

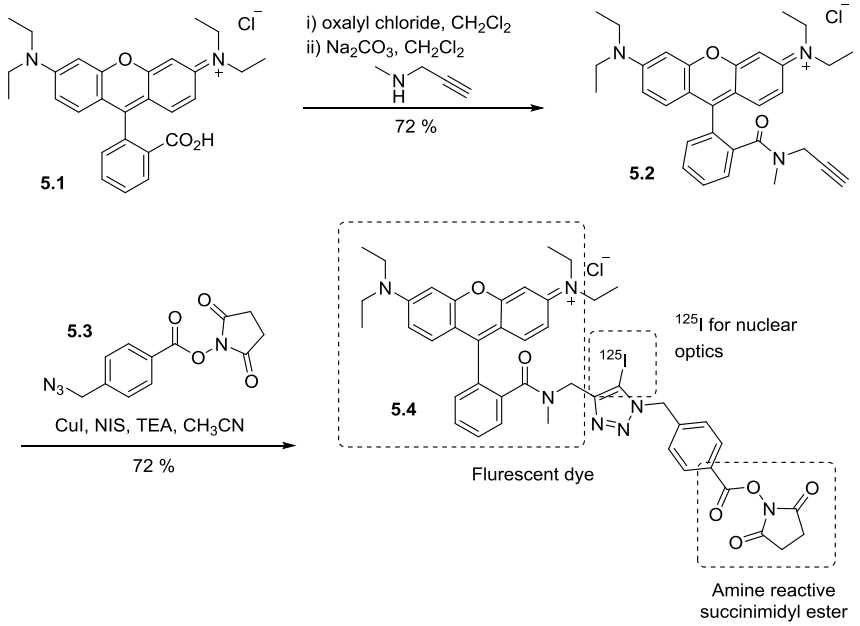
Their extremely broad utility has been demonstrated in many areas of scientific research such as clinical applications,[179],[180] chemoproteomics,[181] and assay development.[182],[183] Most trifunctional reagents reported in the literature were designed and made for one specific purpose relevant to their intended application. To illustrate the broad utility and usefulness of trifunctional probes, a selection of these reagents and their application will be reviewed in section 5.2.1.1.

A number of trifunctional backbones have been developed which have a more generic use, allowing a range of trifunctional tags to be generated quickly and simply. This is generally

done in one of two ways: 1) by synthesising a backbone bearing masked or protected reactive groups that can be revealed and decorated sequentially using just one or two chemical reactions, or 2) by employing a different reactive group on each arm, so that functional groups bearing a complimentary reactive species can be added regioselectively, sometimes in a one-pot manner. These approaches will be reviewed *vide infra*, in section 5.2.1.2 and 5.2.1.3.

5.2.1.1 Applications of Trifunctional Reagents

In 2011, Yan and co-workers developed a one pot synthesis of a trifunctional reagent for imaging with nuclear and optical techniques.[180] Here, a rhodamine dye is functionalised with a terminal alkyne, which is then conjugated to a succinimidyl ester bearing a pendent azide using click chemistry in the presence of electrophilic iodine to give a 5-iodo-1,2,3-triazole **5.4** (Scheme 5-1).



Scheme 5-1: Trifunctional tag for imaging reported by Yan *et al.*

By incorporating ^{125}I, a reagent is produced with three functionalities: a reactive succinimidyl ester group for ligation to proteins of interest; a fluorophore for optical imaging and a nuclear label for nuclear imaging. All of these functionalities combine to give a reagent for multiscale imaging, which allows the shortcomings of individual techniques to be overcome. The reagent was used to label carcinoembryonic antigen (CEA) specific antibody A5B7 and evaluated in mice bearing human colorectal xenografts. Uptake was observed across tumour, where fluorescence signals were in good agreement with the distribution of

radioactivity shown in a radioluminograph. Although this synthesis is short and simple, it is very specific for its intended use, as the iodine functionality in this setup could not be replaced with any other species. Therefore this approach could not be easily adapted for a wide range of tags.

Frei *et al.* reported the 'TRICEPS' (trifunctional chemoproteomic) reagent (Figure 5-2), which was used on living cells and tissue to identify receptors with which ligands were interacting.[181]

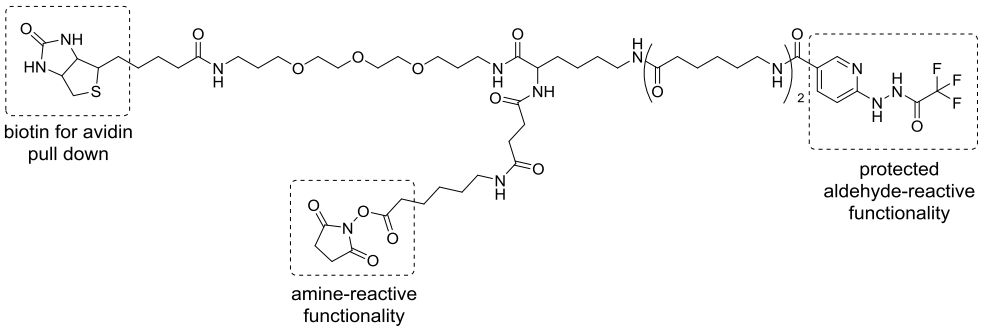
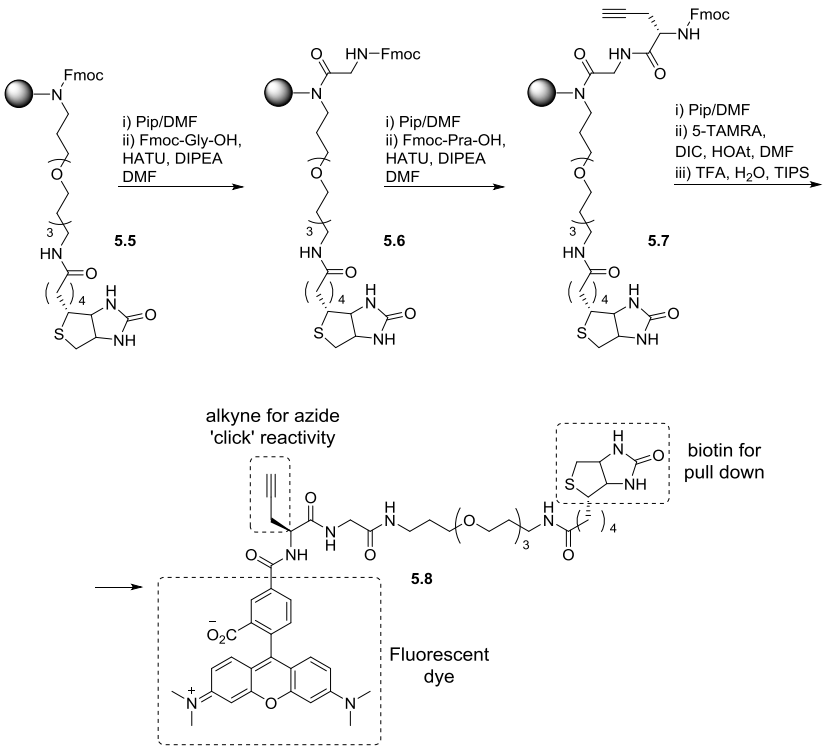


Figure 5-2: The TRICEPS tag, reported by Frei *et al.*

The succinimidyl ester reacted with free amines on peptidic ligands. Cells were then subjected to periodate oxidation, in order to transform hydroxyl groups of carbohydrates into aldehydes. The ligand-TRICEPS conjugate was then added to the cells, allowing the protected hydrazine to bind to the aldehydes on the glycosolated receptors. Cells were then digested and subjected to MS analysis to elucidate the receptor proteins. This approach was demonstrated by the identification of, for example, interactions of peptide ligands with their endogenous GPCRs on living cells, despite the low abundance and integral membrane nature of these receptors.

This application demonstrates nicely the advantages of using a compound with three functionalities, rather than two. The authors state that before TRICEPS, a bifunctional ligand was used incorporating only the biotin and hydrazide groups. This was useful for identifying receptors which sit on the cell surface, but by adding an additional functionality, the same experimental setup can be used for more specific and focused studies. The synthesis of TRICEPS was carried out in solution and consisted mainly of peptide bond formations and deprotections. The backbone of the reagent could be used for various purposes, by simply switching the functional groups used in this example for other reagents.

A trifunctional tag was synthesised for protein labelling by Heal *et al*. to use alongside an enzymatic reaction which introduces an azide onto the *N*-terminus of proteins (Scheme 5-2).[184] The reaction takes advantage of the post translational modification myristoylation, which uses enzyme myristoyl-CoA:protein *N*-myristoyltransferase (NMT) to transfer myristic acid onto an *N*-terminal glycine residue. By expressing proteins in Escherichia coli, which have no endogenous NMT, azide bearing analogues of myristic acid can be incorporated. Once the azide tagged protein is generated, clickable tag **5.8** can be bioorthogonally attached *via* the terminal alkyne it presents.



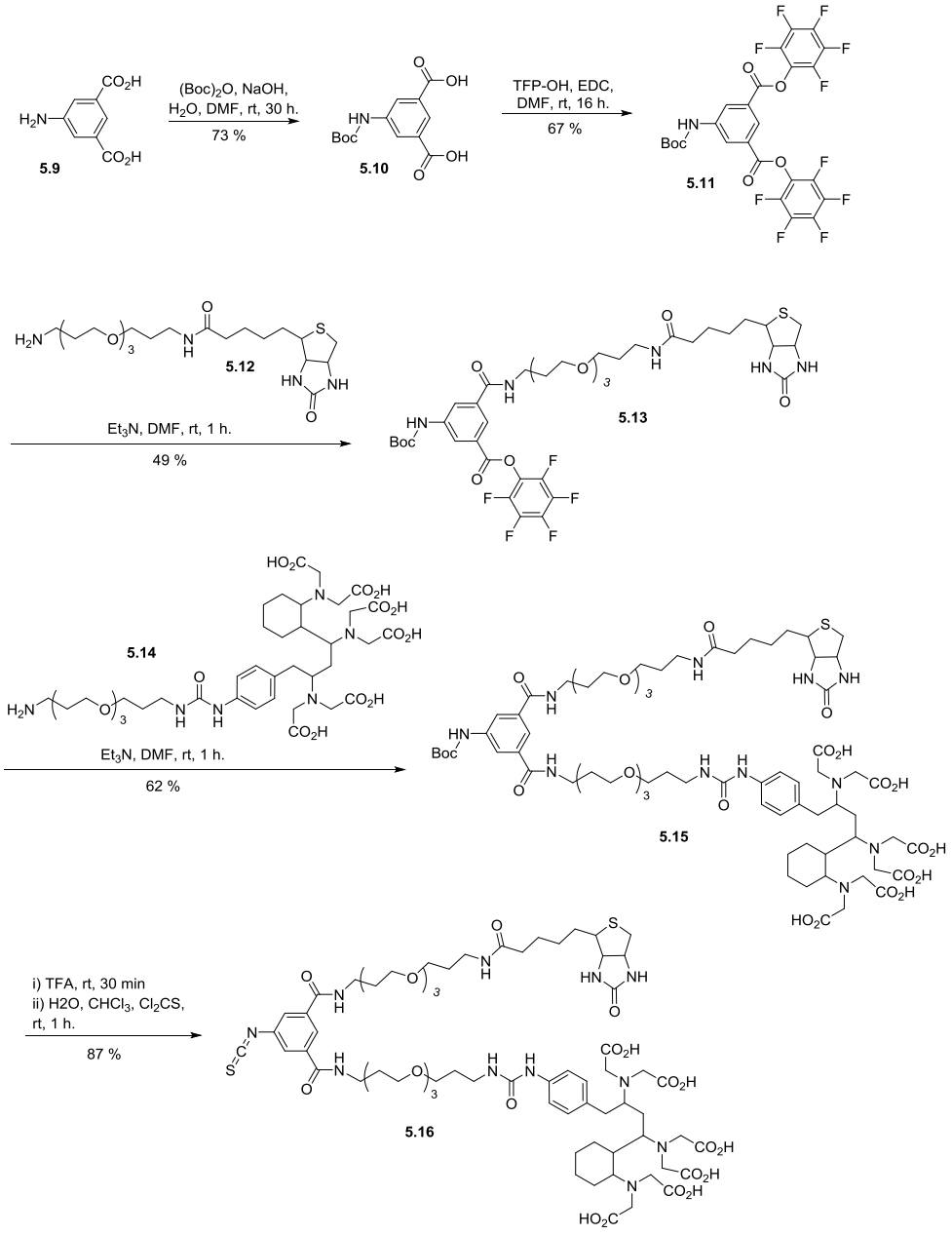
Scheme 5-2: 'Clickable' labelling reagent synthesised by Heal *et al.*

The tag allows for fluorescence detection of the protein *via* the incorporated teramethylrhodamine dye and affinity purification, through the biotin moiety. The tag is synthesised on Biotin-PEG NovaTag resin, where linkage onto resin *via* an amide backbone allows the convenience of solid phase synthesis while also allowing it to be cleaved in a traceless manner. Again, the peptidic nature of the tag could allow different functional groups to be incorporated. The solid phase approach to the synthesis circumvents many purification steps in the build-up of the final product.

Wilbur *et al*. reported trifunctional conjugation reagents which contain biotin, a radiometal chelation moiety and a reactive isocyanate for linkage to biological entities (Scheme 5-3).[179]

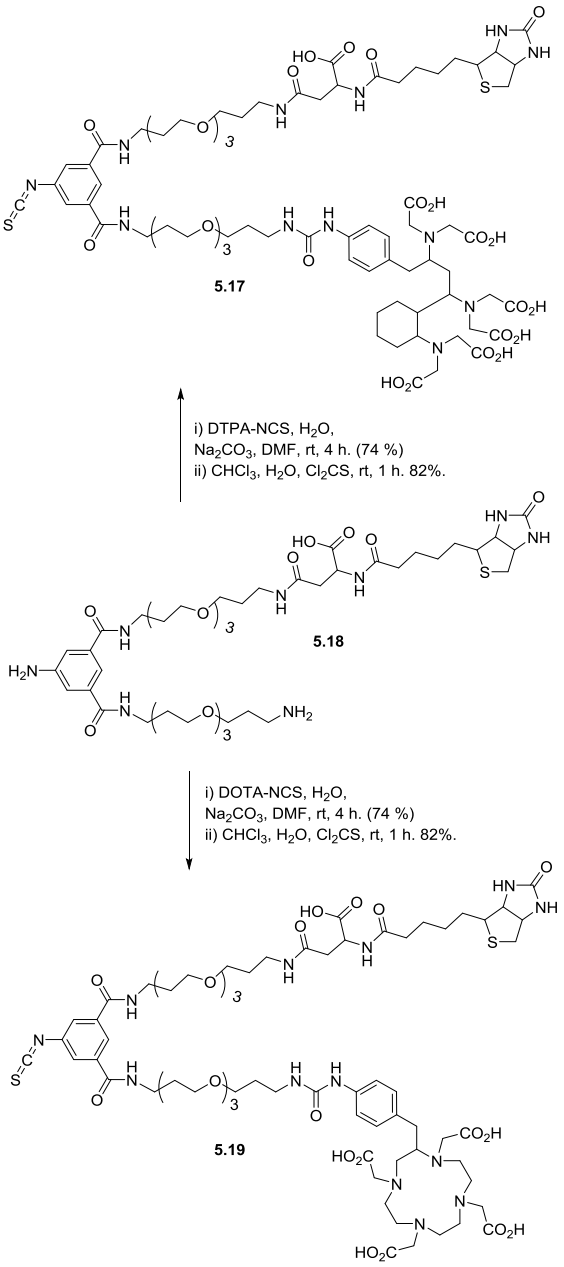
These reagents were produced for the removal of radiolabelled antibodies from the blood of cancer patients undergoing radiation therapy, by biotinylation of the antibodies and subjection of these to an avidin column. The ability to completely remove these antibodies from the blood could allow higher doses to be administered to patients. In the preparation of these antibodies, it was found that successive radiolabelling then biotinylation led to a mixture of products. Also double labelling of antibodies can significantly reduce binding to an antigen. Therefore, a trifunctional reagent was required which could simultaneously incorporate a chelate for radiolabelling and a biotin moiety in equal ratios, while minimising the number of conjugation reactions.

The use of aminoisophthalic acid as a starting material allows each arm of the tag to be orientated in an equally spaced manner. However as there are two acid groups present on the compound, reactivity between these is not easily distinguished, leading to selectivity issues. The linear nature of the synthesis meant that for each new reagent, synthesis had to begin again from starting materials. A more modular or convergent approach would have allowed the replacement of each arm to be carried out from intermediates.



Scheme 5-3: Synthesis of trifunctional reagent described by Wilbur *et al.*

This point was noted by the authors, who then generated biotinylated intermediate **5.18** (Scheme 5-4), which can be functionalised further at two amino groups with differing reactivities, which they used to generate two further trifunctional reagents incorporating a biotinidase blocking group and two different chelation groups The authors also mention that different reporter groups could be incorporated too.

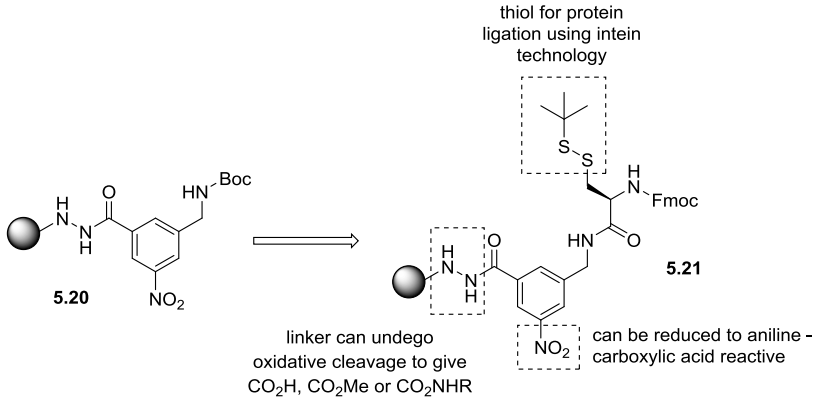


Scheme 5-4: Intermediate allows distinction of reactivity between aromatic and aliphatic amine for regioselective functionalisation

The reagents were conjugated to antibodies and it was found that up to three conjugates could be added before loss of activity of the antibody. Radiolabels ^{111}In and ^{90}Y were chelated easily to the reagents and it was found that most of the biotinylated antibodies bound to an avidin column in one pass.

5.2.1.2 Generic Trifunctional Reagent Scaffolds Synthesised on Solid Support

Trifunctional building block **5.21** was described by Watzke and co-workers was designed as a flexible method to simultaneously allow linkage to a protein of interest, immobilisation to a surface and attachment of another tag such as a reporter group (Scheme 5-5).[185] A solid supported synthesis approach was taken, whereby three functional groups were orthogonally masked to allow chemoselective functionalisation of each. These included an aromatic nitro group, which could be reduced to give an aniline, a Boc protected amine and a carboxylic acid masked by attachment to the solid support by means of a traceless hydrazine linker.[182]



Scheme 5-5: Trifunctional building block for synthesis of trifunctional protein labelling reagents reported by Watzke and co-workers

Removal of Boc and subsequent introduction of a cysteine monomer was carried out to allow linkage to proteins that, through the use of intein technology, contain a thioester group. The following compounds were generated using the building block (Figure 5-3):

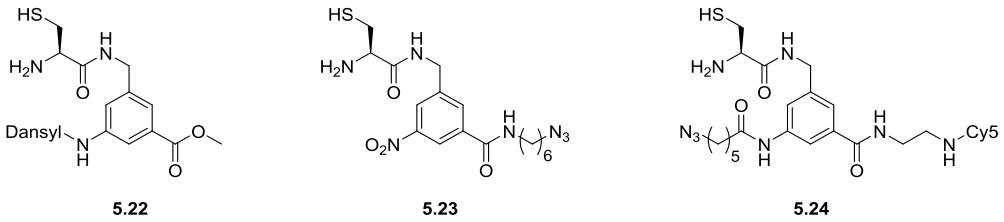
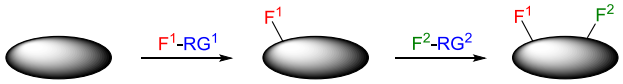


Figure 5-3: Compounds synthesised for protein labelling by Watzke and co-workers

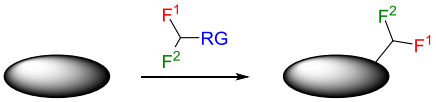
Dansyl bearing compound **5.22** was ligated to the mercaptoethanesulfonic acid (MESNA) thioester of Ypt1D3, the yeast homologue of the human Rab7. Analysis by SDS page revealed that the ligation was successful and the labelled protein was obtained quantitatively. Compound **5.23** was also ligated to Ypt1D3 and the azido group used to immobilise the tagged protein on a phosphane functionalised solid support *via* the azide using Staudinger

ligation.[186] Cy5 labelled anti RAS antibody was then used to detect the protein on the functionalised surface, confirming immobilisation of the protein had been successful. Although this synthetic approach was used to generate a fully trifunctional compound **5.24**, its use for protein functionalisation, fluorescent labelling and immobilisation was not reported.

In 2008, Garanger *et al*. coined the phrase multifunctional single-attachment-point (MSAP) reagents (Figure 5-4).[187] The group were aiming to solve the problem of stoichiometric ratios which were encountered when sequential addition of reporter groups bearing different reactive functionalities were added to nanoparticles or polymers. As the functional groups are ligated one after the other using different chemistries, any incomplete reactions invariably affect the ratio of one reporter group against the other, a problem which escalates with increasing numbers of different reporter groups, or decreasing the size of the material being functionalised. They believed this problem could be overcome by generating a multifunctional probe with one single reactive group, with which the probe could be attached to a substrate in one single step.



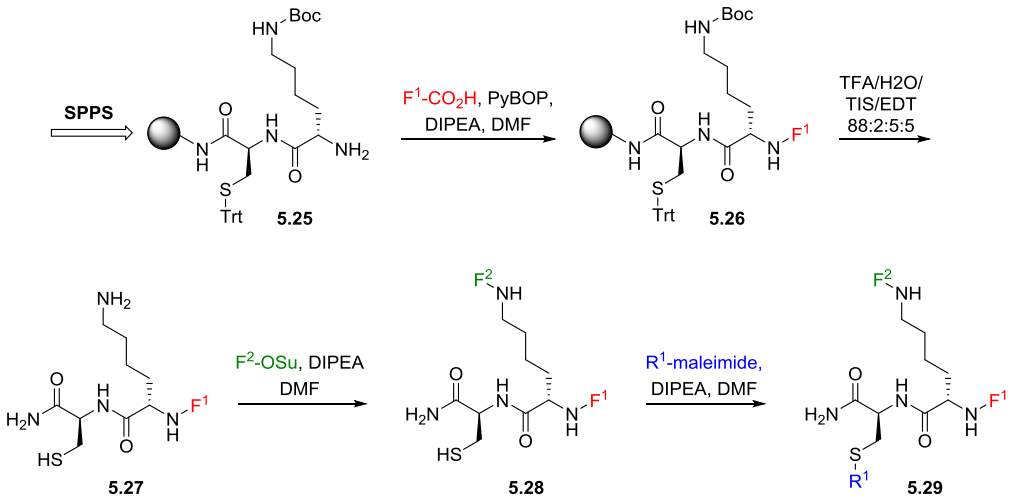
A) Sequential addition can lead to varying ratios of functional groups



B) Use of a multifunctional single-attachment-point (MSAP) reagents allows introduction of multiple functional groups in a single step

Figure 5-4: Sequential v's MSAP labelling

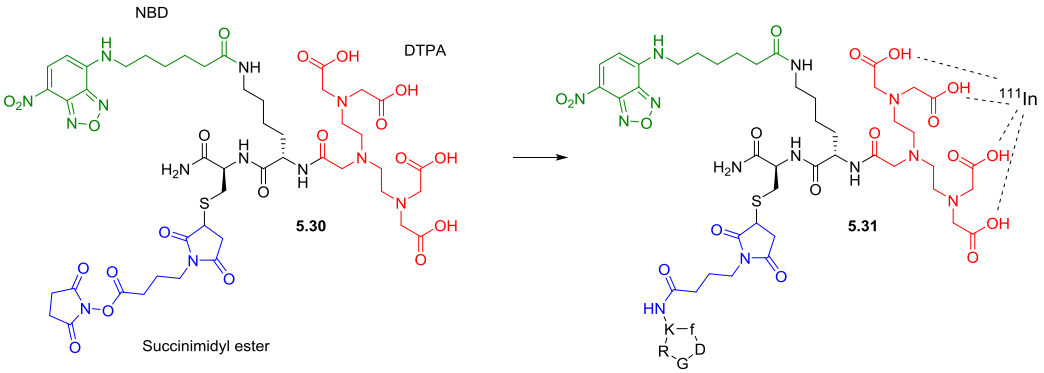
A peptidic scaffold based around a Lys-Cys dimer was utilised to generate examples of MSAP compounds bearing two functional groups and one reactive group (Scheme 5-6).



Scheme 5-6: General synthesis of bifunctional MSAP as reported by Garanger *et al.*

The backbone of the reagent **5.25** was built up using standard SPPS using lysine and cysteine residues incorporating acid labile protecting groups, the three points of functionalisation being the lysine ε-amine and the peptide backbone amine, along with the cysteine thiol. The amino groups were differentiated from each other by adding the first functional group on to the peptide *N*-terminus before deprotection of the side chains (and simultaneous cleavage from the resin). The differential reactivity of the thiol and lysine side chain amine was exploited by use of an activated ester to ligate the second functional group to the Lysine ε-amine, and finally by conjugation of a maleimide with the cysteine thiol introduced the MSAP reagent's reactive group.

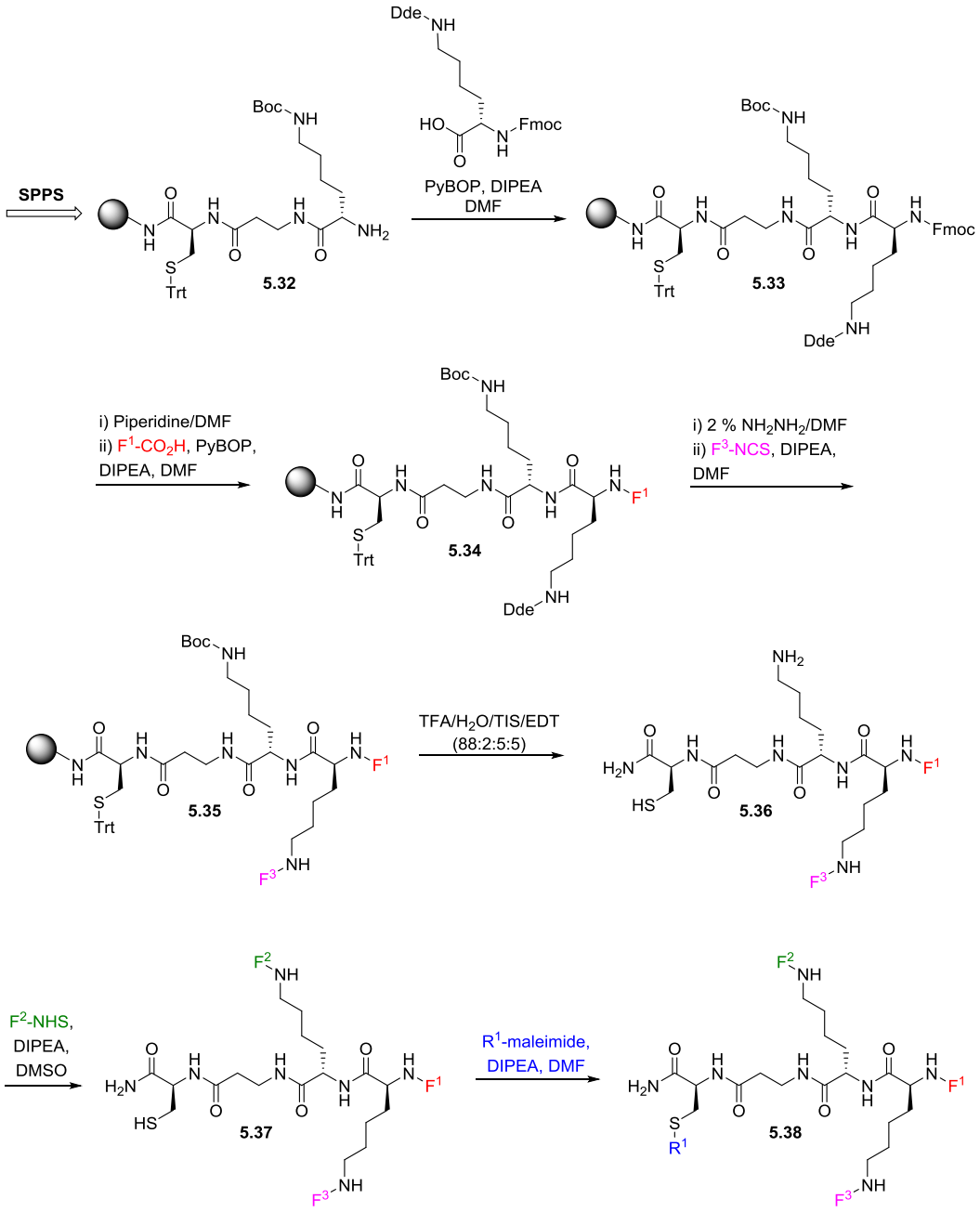
Bifunctional compound shown in Scheme 5-7 was synthesised using this method. Incorporated onto the backbone is an NBD fluorophore and a metal chelating DTPA group, along with a succinimidyl ester group for amine conjugation. The succinimidyl ester group was used to link the tag to a the ε-amino group of a lysine residue in a cyclo[-RGDfK-] peptide, known to have high affinity for integrins. The NBD fluorophore allowed clear monitoring of the reaction by HPLC. The DTPA group was then complexed with ^{111}In, to give a bifunctional probe which was administered intravenously to a mouse bearing an integrin-expressing B16F0 melanoma. Distribution of the probe was traced by SPECT-CT and immunohistochemistry was employed to determine distribution of the probe within the tumour by employing an NBD antibody. The MSAP approach had been successful in allowing easy construction of a sophisticated, dual readout tracer.



Scheme 5-7: Bifunctional reagent synthesised by Garanger *et al.*

This scaffold was also used to generate a MSAP reagent for the controlled biotinylation of proteins.[188] Here, a biotin, fluorescein and reactive succinimidyl ester were combined in order to biotinylate protein substrates. Determination of the degree of biotinylation by absorbance or fluorescence using the fluorescein dye was carried out.

The authors next required a construct that would allow the introduction of three functional groups onto an MSAP reagent. Another scaffold, **5.32**, generated from a Lys-Lys-βAla-Cys motif was synthesised and used to generate MSAP reagents presenting three functional groups (Scheme 5-8). A Dde protecting group was chosen for the extra lysine incorporated into the backbone, which could be removed selectively in the presence of the Fmoc and acid labile protecting groups already present, allowing a third functional group to be introduced selectively.



Scheme 5-8: General synthesis of trifunctional MSAP as reported by Garanger *et al.*

This trifunctional MSAP was decorated with two functional groups (fluorescein for detection and DTPA for metal complexation) and one for stability (PEG) to give **5.39** (Figure 5-5). The reactive thiol group was used to conjugate the tag to gold nanoparticles.

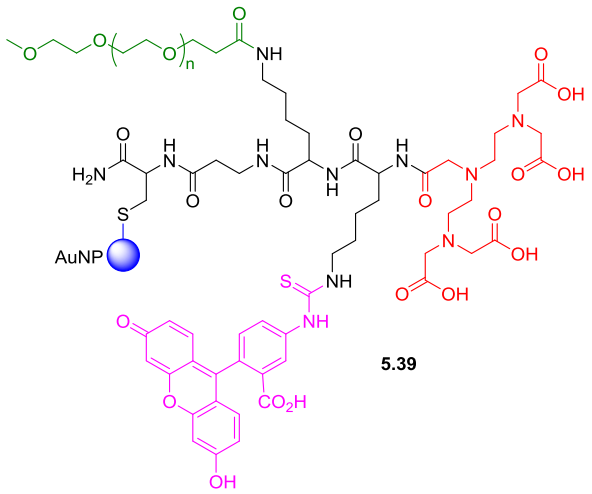
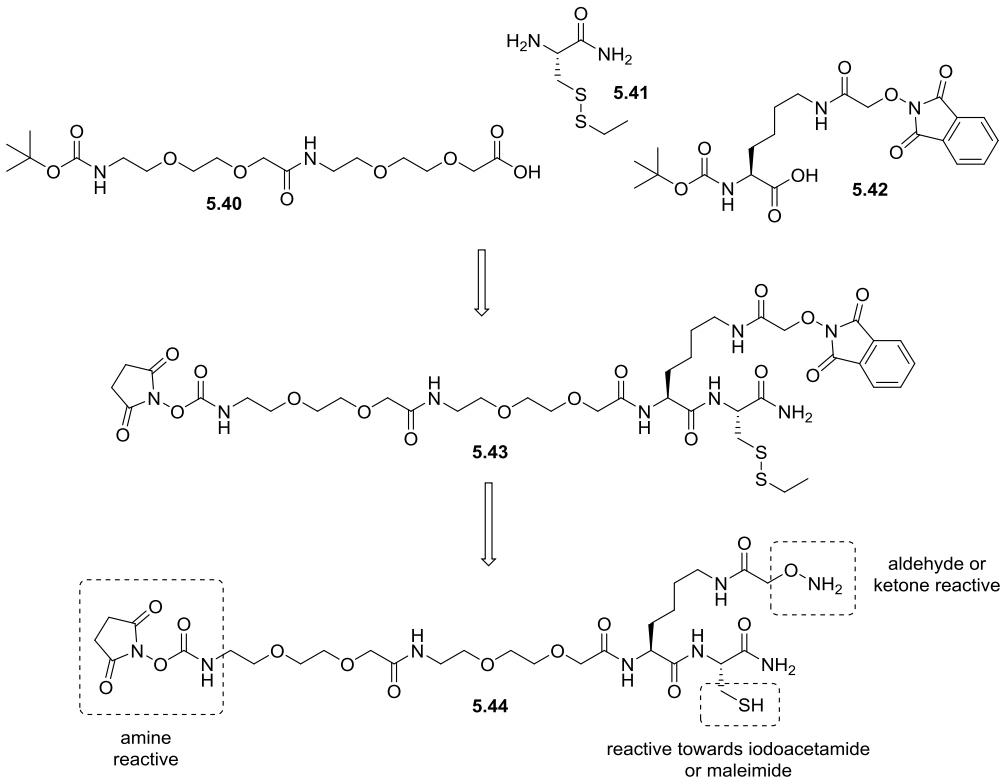


Figure 5-5: Trifunctional reagent synthesised by Garanger *et al.*

This approach in synthesising MSAP reagents is highly flexible and could allow easy generation of a wide range of trifunctional entities. Although highly useful, this method does have a small drawback, in that at least one of the functional groups needs to be introduced during solid supported synthesis. It therefore would not be feasible to synthesise a tag in which every functional group is sensitive to the relatively harsh conditions required for cleavage of the compound from the resin. This issue could be overcome by the introduction of an additional orthogonal protecting group, which would allow deprotection of each amine individually in solution, therefore giving complete flexibility over placement of each functional group.

5.2.1.3 Generic Trifunctional Reagent Scaffolds Synthesised in Solution

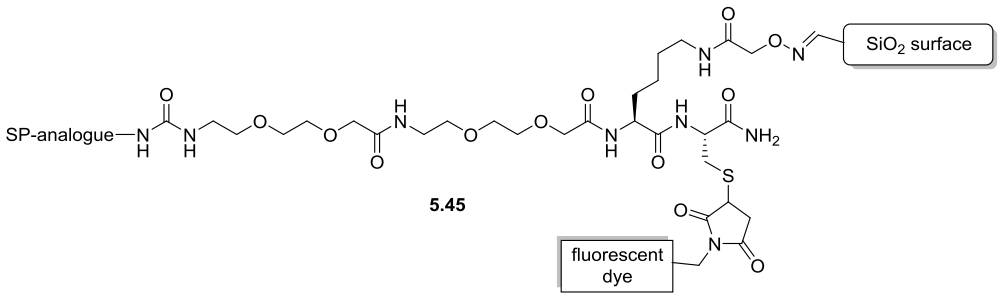
In 2008, the Romieu group reported generic heterotrifunctional peptide-based reagent **5.44** named tripod, bearing three different reactive functionalities (Scheme 5-9).[183] An *N*-succinimidyl carbamate moiety, an aminooxy and a thiol group on a peptidic backbone presented the opportunity to introduce amines, aldehyde/ketone groups and thiol reactive compounds, respectively. This heterotrifunctional probe is based on a dipeptide (Lys-Cys) architecture with a PEG spacer incorporated to increase water solubility. Fully protected building blocks were selected for the backbone build-up to ensure the solution phase synthesis was convergent and simple.



Scheme 5-9: Synthesis of heterotrifunctional peptide-based reagent (tripod) as reported by Romieu *et al.*

The aminooxy and thiol groups were protected until required using a phthalide and ethyl disulfide group respectively, to prevent degradation of the tag through attack on the nucleophile-sensitive active carbamate.

Once deprotected, the trifunctional reagent was sequentially and chemoselectively derivatised with three molecular partners (Figure 5-6). Firstly, the *N*-succinimidyl carbamate was attached to the *N*-terminus of an analogue of SP, an eleven amino acid neuropeptide, in a fast and efficient manner. Subsequent removal of the phthalide and ethyl disulfide groups allowed chemoselective reaction of the thiol group to a maleimide functionalised fluorescent dye *via* Michaeal addition. Importantly, no reactivity was observed between the aminooxy group and the dye. Finally, the aminooxy group was utilised to link the compound to an aldehyde-modified silica surface.

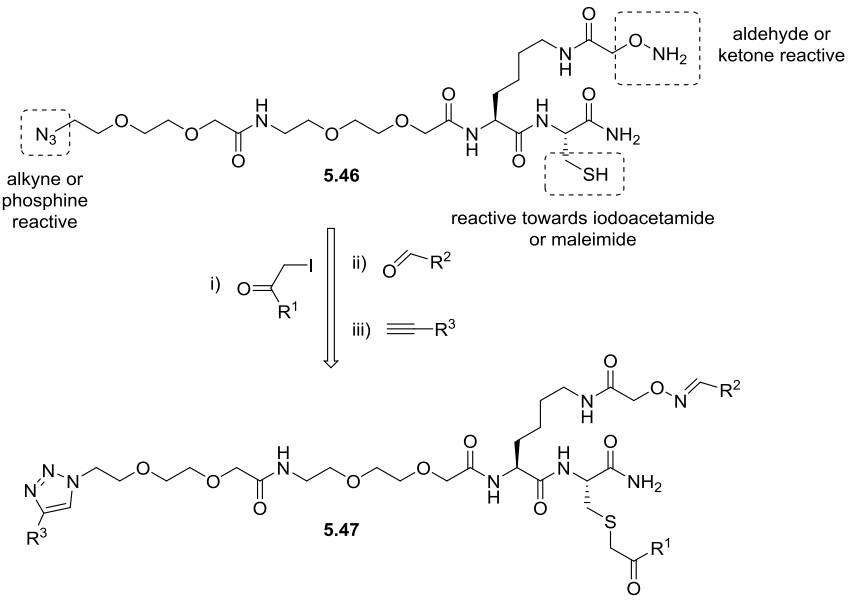


5.45

Figure 5-6: Trifunctional reagent for peptide labelling and pull down to solid surface

Although elegant, this approach presents a few disadvantages. Firstly, the solution based synthesis requires multiple purification steps, most of which need to be carried out using HPLC which can lead to significant loss of product. Secondly, only ligands with a single free amine could be used in this system as they would react with the succinimidyl ester of the tripod reagent, which vastly limits the ligands that can be used in this system. Indeed, the SP analogue utilised had all of the NHS-ester reactive sites capped for this purpose. Also, the instability of the compound when all three protecting groups are removed is an issue, as it limits the sequence in which conjugation to the tag can occur.

The second and third pitfalls of this system were addressed by the group when in 2010 they presented the second generation tripod reagent (Scheme 5-10).[189] Here, the nucleophile sensitive *N*-succinimidyl carbamate was replaced with an azide, therefore allowing biorthogonal linkage to alkyne tagged chemical entities *via* copper-mediated 1,3-dipolar cycloaddition and also negating the need for protecting groups on the aminooxy- and sulfhydryl groups.



Scheme 5-10: Second generation tripod reagent

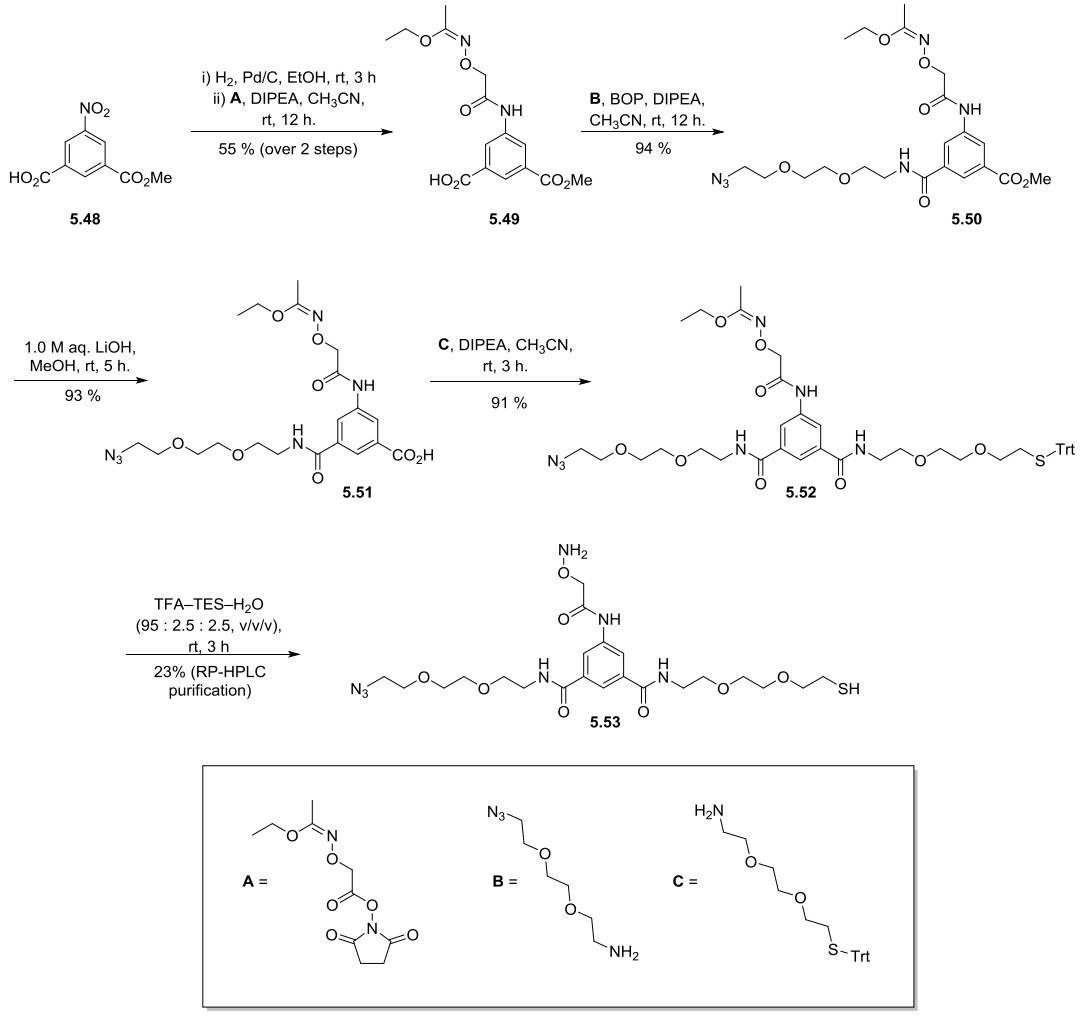
However, this approach still limited the sequence in which these reactive functionalities could be decorated. Thiol alkylation and oxime ligation had to be carried out before copper-catalysed azide-alkyne 1,3-dipolar cycloaddition (CuAAC) as it was feared that the thiol groups could be oxidatively damaged and N-O bonds cleaved by the reagents and side products produced during the cycloaddition reaction. However, as conjugation to each reactive point of the trifunctional reagent can be carried out under very mild conditions, this method does allow attachment of chemically sensitive groups, demonstrated by the attachment of DNA fragments and nuceophile sensitive lactone-containing mycotoxin AFB2. The authors comment that this synthesis has the problem that there is a requirement for multiple chromatography steps during tag derivitisation, which leads to substantial loss of material.

Therefore, in 2013, a one pot procedure for functionalisation of a new tripod reagent was reported (Scheme 5-11).[190] This version of the tripod compound saw the peptidic backbone replaced with a benzene based core, in an attempt to generate a structurally simple compound with a more convergent synthetic route. The three aforementioned orthogonal bioconjugatable groups were grafted onto the aromatic core in a 6 step synthesis starting from commercially available 5-nitroisophthalic acid monomethyl ester **5.48**.

The five step synthesis relies on standard and easily implemented reactions (e.g. reduction of nitro group, saponification of methyl ester) to systematically decorate the benzene core with the three bioreactive functionalities. PEG spacers were used to ensure solubility of the tripod

in aqueous solvent. It was found that upon storage, the deprotected tripod compound **5.53** degraded *via* dimerisation of the thiol groups and reaction of the aminooxy- group with acetone in the lab atmosphere. Therefore the authors suggest deprotection immediately prior to use.

Decoration of the tripod was carried out in a one-pot fashion under mild aqueous conditions in the following order: 1) Oxime ligation; 2) thio alkylation; 3) CuAAC reaction. This allowed isolation of a tri-substituted product in 28% yield, compared to the 6% and 9% yields achieved when isolation of each intermediate was carried out.



Scheme 5-11: Synthesis of benzene based tripod reagent

Despite the elegant reactivity of this tripod reagent, and the practical problems overcome in its development, another drawback to this approach becomes evident, in that each entity which is tethered to the tripod reagent must be functionalised with a reactive partner of one of the tripod arms. In one use of the tripod, three fluorophores are attached to the scaffold in

order to generate a triple FRET cascade. In this application, three dyes must be functionalised with three different functional groups. Not only is this labour intensive, but during each reaction it can be assumed that some precious dye material is lost.

5.3 Aims

Various generic scaffolds have been presented in the literature from which trifunctional reagents are constructed. However these constructs have various drawbacks as discussed *vide supra*. Within our lab, work is oriented towards high throughput techniques and parallelisation, therefore we require a highly versatile tool which would allow rapid and clean synthesis of a variety of trifunctional reagents, which could be employed easily in a broad range of applications, including the tagging of PS/PS library members. Therefore we aimed to design and synthesise a generic trifunctional reagent backbone which is simple and fast to synthesise and can be functionalised in a regioselective and time efficient manner.

5.4 Results and Discussion

5.4.1 Trifunctional Reagent Backbone Design

Previous approaches in the synthesis of these trifunctional tool compounds have been to utilise three different reactive groups on a backbone in order to obtain regioselectivity. However, this approach has a drawback, in that each group added to the trifunctional backbone has to be modified or purchased to incorporate the desired reactive partner. Also, as exemplified above, the use of different reactive species on a backbone can lead to intramolecular reactions and therefore degradation. Rather than this, an approach to utilise a backbone which presents three similarly reactive groups, so that no cross reactivity could take place was decided upon (Figure 5-7). Orthogonal protecting groups should allow the addition of functional groups to any location on the backbone in any order.

It was decided that amines would be the reactive species used on the tag, as they present two major benefits. Firstly the high yields and chemical reliability associated with peptide bond formation on solid support could be exploited to ensure complete conversion and clean products. Secondly, most moieties we envisage being employed to decorate the tags (e.g. fluorophores, biotin etc.) are available commercially in either carboxylic acid or activated ester form, therefore no modification of reagents or purchasing of expensive adapted reagents would be required for positioning at any point on the resin. Finally, a wide range of orthogonal amine protecting groups have been developed for use in peptide synthesis, which could be exploited here.

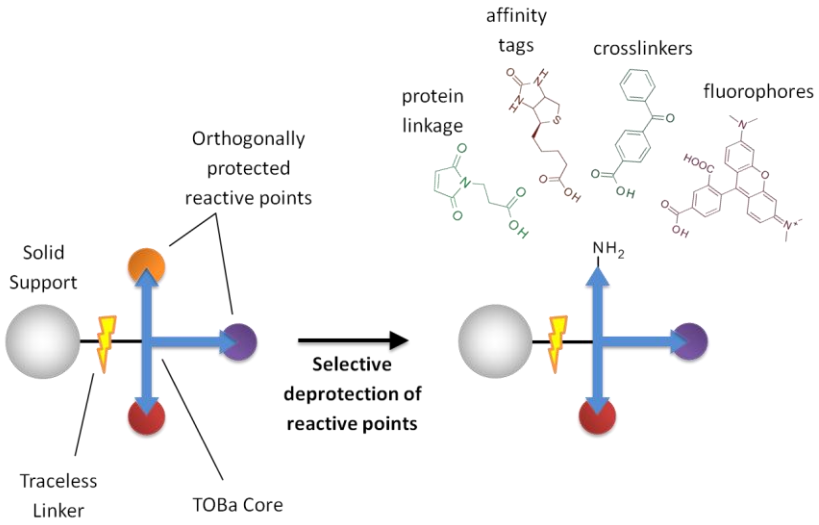
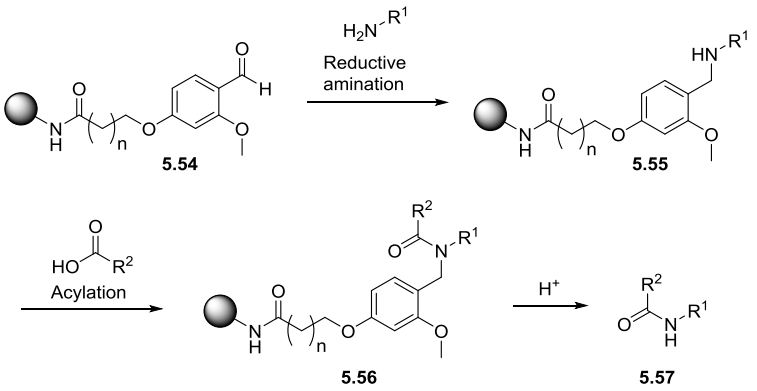


Figure 5-7: Setup of planned trifunctional backbone resin

We envisaged using a combination of solution and solid-phase methods for the preparation of the generic trifunctional reagent. The backbone would be prepared to high purity in solution and then applied to solid phase resin for rapid and simple functionalisation, avoiding the need for multiple purification steps of complex molecules which could lead to substantial loss of material. The fully orthogonal approach also means that sensitive functional groups which could not withstand resin cleavage conditions, the tag could be cleaved at any point and the synthesis continued in solution.

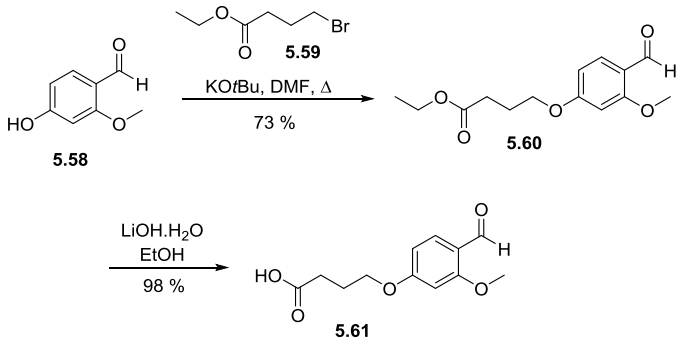
5.4.2 Solution Synthesis of TOBa

The backbone amide linker **5.54** (BAL) was chosen to attach the backbone to a solid support (Scheme 5-12).[191] This linker is based around an electron rich aromatic aldehyde, onto which primary amines can be reductively aminated. At this point (**5.55**), when loaded with a secondary amine, the linker cannot be cleaved. Subsequent acylation of the amine gives an amide bond, with the linker being attached to the peptidic compound *via* the nitrogen of the peptide bond. Acid mediated cleavage of the linker is now possible, with electron donating methyloxy groups on the ring assisting the liberation of synthesised compounds in a traceless manner.



Scheme 5-12: BAL Linker loading and cleavage

As the linker is attached onto the amide backbone, elongation of anchored molecules can be carried out in two directions. The acid labile character means that the use of plethora protecting groups is possible, such as nucleophile, base or catalytically labile ones. For the preparation of a trifunctional backbone, BAL linker was synthesised in two steps from 4-hydroxy-2-methoxybenzaldehyde **5.58** (Scheme 5-13). The phenol was deprotonated using potassium *tert*-butoxide and alkylated with ethyl 4-bromobutanoate to introduce a protected carboxlic acid bead linkage point in good yield. The ethyl ester was hydrolysed in 98% yield using lithium hydroxide to give the desired BAL linker **5.61**.



Scheme 5-13: Synthesis of BAL linker

Next, an easily accessible tag backbone was designed to present the three masked amines in a spatially separated orientation. Synthesis of the backbone with the BAL linker was carried out in solution, so that the two components could be loaded onto resin in one step. A mono-protected diamine was used for reductive amination onto the BAL linker to give the first masked amine. The resulting secondary amine was then acylated with a lysine amino acid bearing two other protecting groups orthogonal to the first and each other, generating a

backbone with three orthogonally protected amines. Fmoc, Alloc and Dde protecting groups are frequently used for SPPS and can all be cleaved in an orthogonal manner.[192] They were therefore employed in the trifunctional tag backbone (Figure 5-8), which was named TOBa resin (**T**rifunctional **O**rthogonal **Ba**ckbone resin).

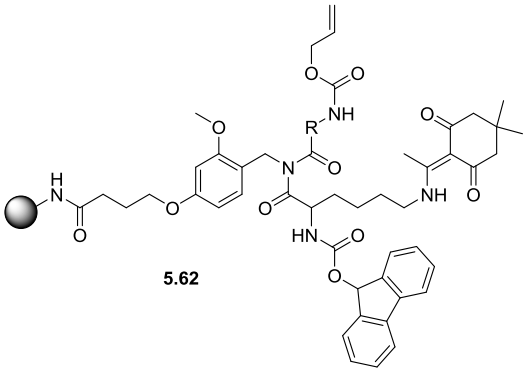
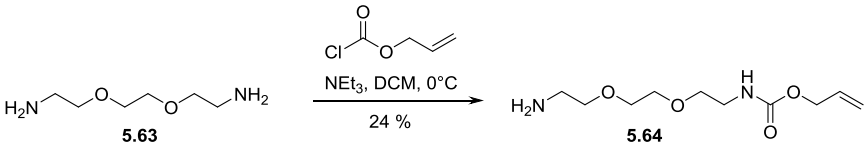


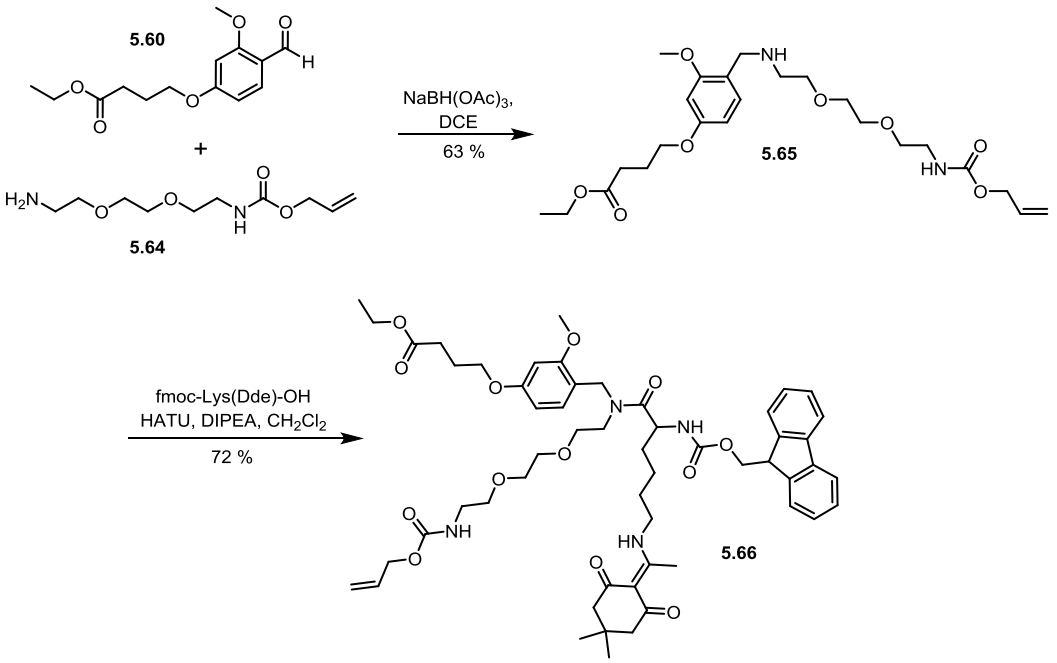
Figure 5-8: Planned structure of TOBa backbone

The backbone was synthesised in solution, starting with ethyl ester protected BAL linker compound **5.60**. For the reductive amination step, mono-allyl protected PEG diamine **5.64** was synthesised (Scheme 5-14).



Scheme 5-14: Synthesis of mono allyl protected PEG diamine

This amine was combined with the protected BAL linker and reduction of the resulting imine performed using sodium triacetoxyborohydride to give secondary amine **5.65** (Scheme 5-15). This was acylated with a lysine residue incorporating an Fmoc protected α-amine and a Dde protected ε amino group to give **5.66** in good yield.

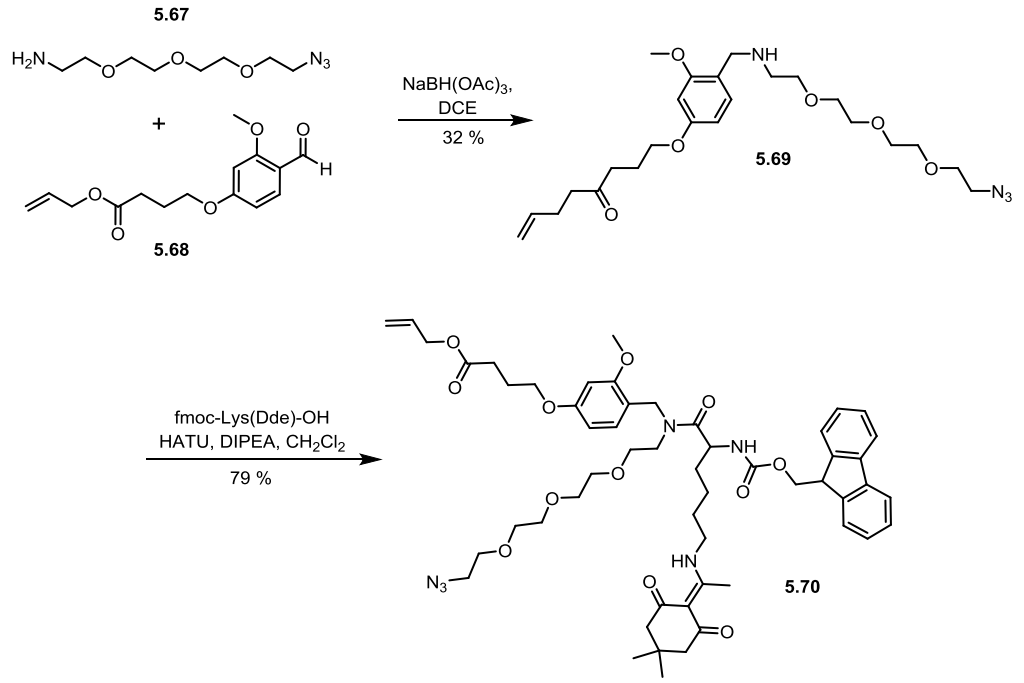


Scheme 5-15: Synthesis of ethyl ester protected TOBa backbone

With the backbone fully established, deprotection of the ethyl ester was attempted to reveal the free acid for bead attachment. Using sodium and lithium hydroxide, significant deprotection of the Fmoc group was observed. Alkyl esters can be cleaved in acidic conditions, however this would cleave the backbone from the BAL linker. Therefore, another approach was sought.

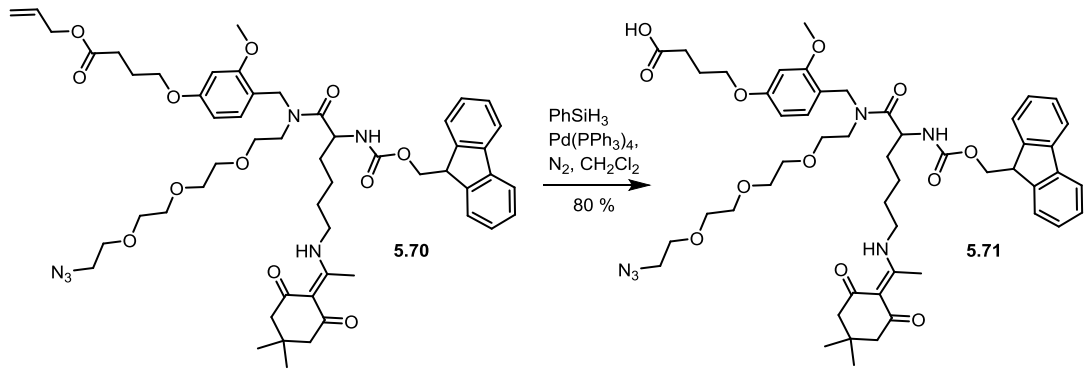
An additional chemical group was needed which would provide orthogonality over those already present on the compound. An azide was chosen to fulfil this requirement, as it can be orthogonally reduced to the corresponding amine using phosphine mediated Staudinger reduction in the presence of the other protecting groups. Use of an azide also permits an allyl group to be employed for protection of the BAL acid. Finally, this approach also presents the opportunity to use Huisgen 1,3-dipolar cycloaddition 'click' chemistry with the tags, which could be useful if the tags are to have bioorthogonal utility. The backbone was therefore resynthesised, incorporating these modifications.

Allyl protected BAL linker **5.68** was prepared, onto which commercially available 11-Azido-3,6,9-trioxaundecan-1-amine **5.67** was reductively aminated (Scheme 5-16). Following formation of the secondary amine, acylation was carried out with Fmoc-Lys(Dde)-OH in the presence of HATU coupling reagent and base to give allyl protected TOBa compound **5.70**.



Scheme 5-16: Synthesis of allyl protected TOBa backbone

Initial attempts at removal of the allyl protecting group resulted in reduction of the azide. Therefore, several conditions were assessed with variation in hydride source and palladium catalyst loading. Finally, allyl deprotection was carried out in high yield using $PhSiH_3$ and $Pd(PPh_3)_4$ in CH_2Cl_2 with no sign of azide reduction to give the free acid, TOBa compound **5.71** in excellent yield (Scheme 5-17).

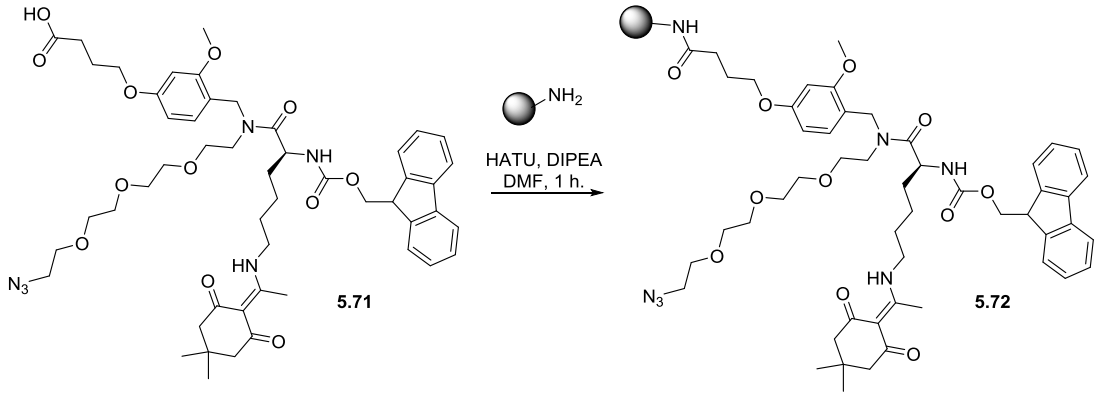


Scheme 5-17: Allyl deprotection of TOBa

5.4.3 Loading of TOBa onto Solid Support and Analysis

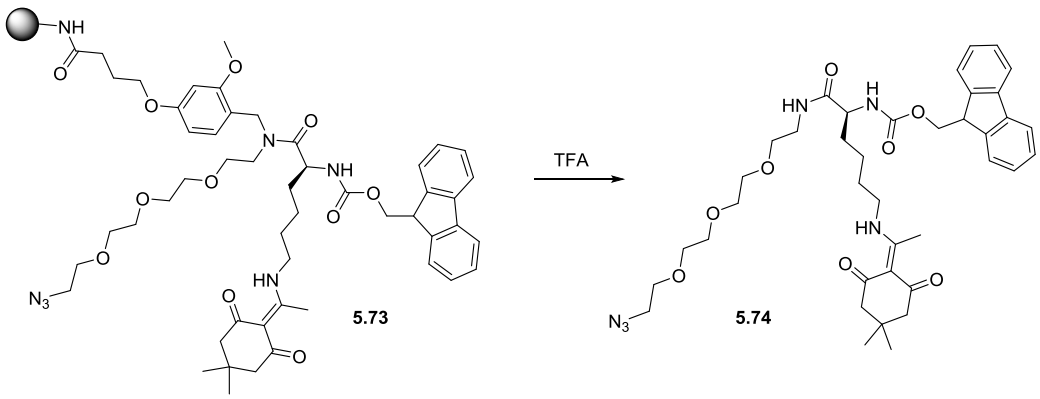
TentaGel beads were chosen as the solid support onto which to load TOBa, as their flexible nature would minimise steric hindrance during synthesis and therefore aid high yields and purity. TOBa compound **5.71** was loaded onto TentaGel NH_2 resin using HATU mediated

peptide coupling (Scheme 5-18). As the TOBa compound is quite large, we were concerned that incomplete coupling may be an issue. Indeed, a TNBS test was performed following one hour of using 5 equivalents of the TOBa compound and HATU which indicated free amines remained on the resin. Therefore a second coupling was performed, after which no free amines were detected. Therefore completion of the reaction could be assumed.



Scheme 5-18: Loading TOBa onto TentaGel NH_2 resin

As all three protecting groups chosen are stable to TFA treatment, cleavage the fully protected backbone **4.74** from solid support was carried out (Scheme 5-19) and was found to have a purity of 96% by RP-HPLC.



Scheme 5-19: Cleavage of protected TOBa tag

To prove full orthogonality of the resin, selective removal of Fmoc and Dde protecting groups was undertaken and the resulting mono-deprotected product cleaved from the resin and assessed by RP-HPLC (Figure 5-9, results summarised in Table 5-1).

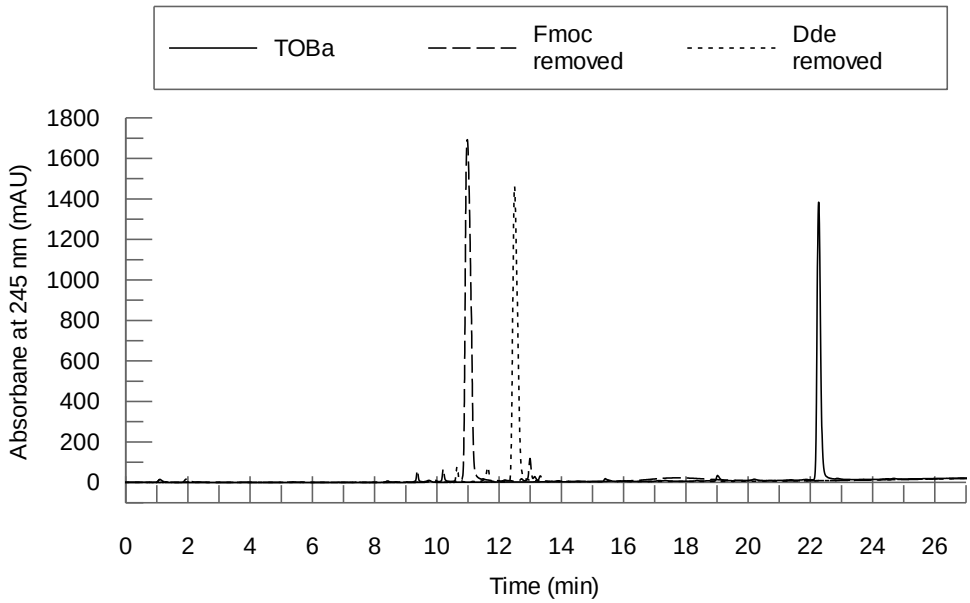


Figure 5-9: RP-HPLC trace of cleaved TOBa backbone, TOBA backbone with Fmoc removed and TOBa backbone with Dde removed.

Entry	Group removed[a]	Expected Mass for $[M+H]^+$	Observed Mass	Purity[b]
1	None (**4.74**)	733.39	733.46	96%
2	Fmoc	511.32	511.40	93%
3	Dde	569.31	569.47	95%

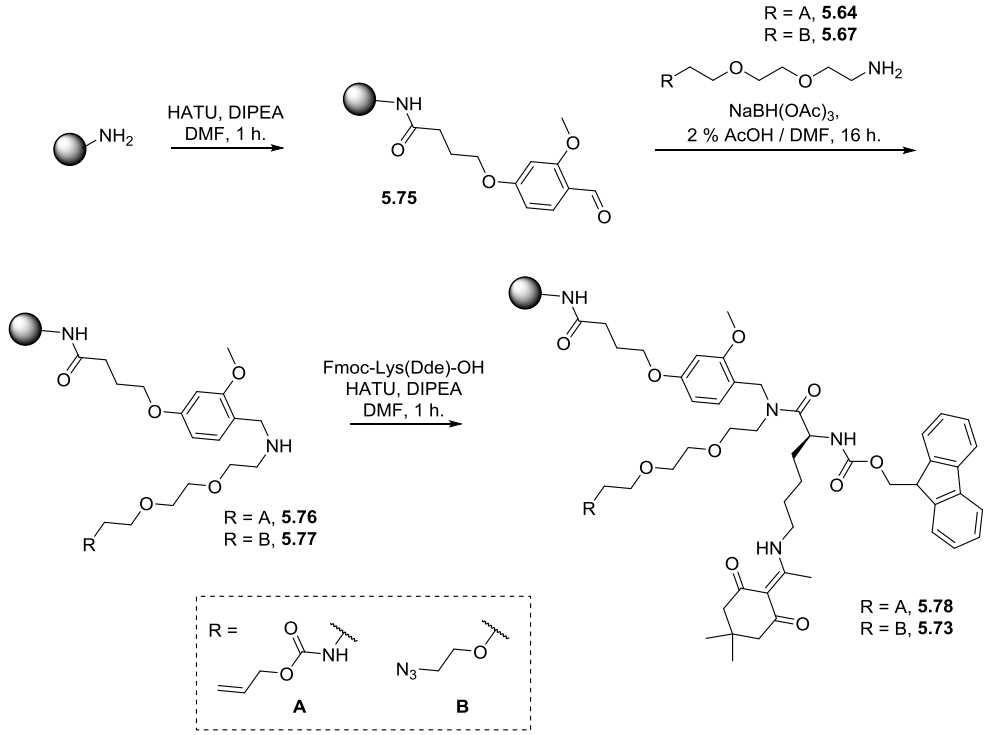
Table 5-1: Investigation of orthogonality of protecting groups.
[b]Determined by RP-HPLC at 254nm.

Each deprotection proved to be completely orthogonal, with no evidence of side products resulting from the removal of the other protecting groups.

5.4.4 Solid Supported Synthesis of TOBa

A solid phase synthesis of TOBa was also devised (Scheme 5-20). This allows rapid synthesis of tags but does not allow for purification of the TOBa backbone before functionalisation as the solution phase synthesis does. BAL linker **5.61** was loaded onto TentaGel NH_2 resin using standard coupling conditions. Next, either protected PEG based diamine **5.64** or **5.67** was reductively aminated onto the aldehyde using 5 equivalents of amine and 7 equivalents of sodium triacetoxyborohydride and the resins shaken overnight. A test for secondary amines gave confirmation of the presence of this functionality, so Fmoc-Lys(Dde)-OH was coupled, again using standard conditions. This yielded the two TOBa compounds **5.73** and **5.78** in high purity, and can therefore be regarded as an acceptable alternative route for the synthesis of the TOBa backbone. The solid phase route boasts the advantage that the azide on the backbone can be replaced with the allyl protected

amine which was found to be incompatible with the solution phase route, and therefore allows more flexibility in the system.



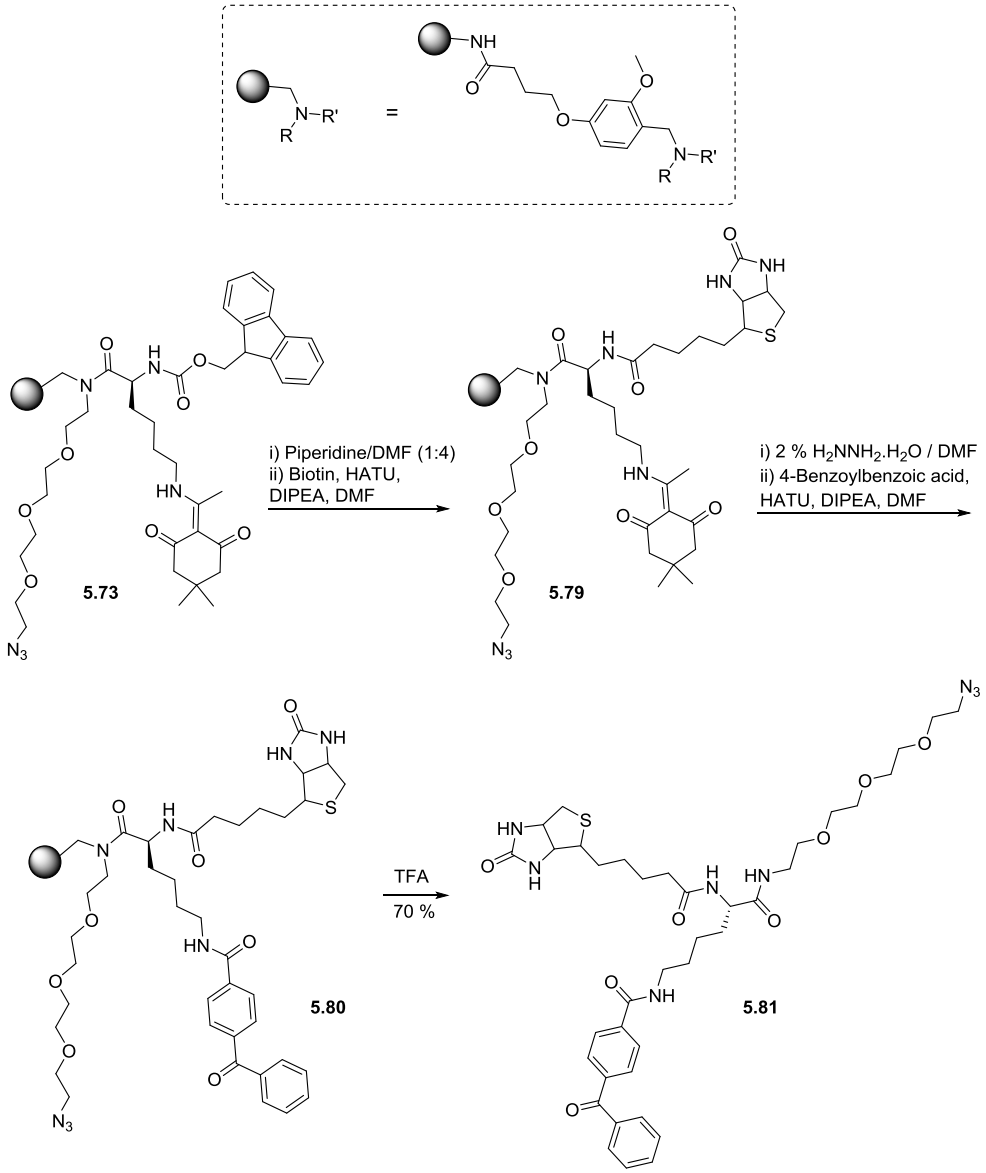
Scheme 5-20: Solid phase synthesis of TOBa

The application of TOBa functionalised resin was investigated in the synthesis of some proof-of-concept trifunctional reagents.

5.4.5 Application of TOBa: Synthesis of Tags

5.4.5.1 Pull Down Tag

Following backbone synthesis, a tag incorporating biotin and a benzophenone cross-linker was synthesised using commercially available biotin and 4-benzoylbenzoic acid (Scheme 5-21). The final product was isolated in 70% yield and 72% crude purity.



Scheme 5-21: Synthesis of azide-biotin-benzophenone trifunctional reagent

This tag is to be utilised in experiments to identify binding partners for chemical compounds and combinatorial library members. In an experiment using this tag, ligands incorporating Pra amino acid will be attached to the trifunctional tag using 1,3-dipolar cycloaddition chemistry, then incubated with a lysate. Following incubation, exposure of the mixture to UV light would cause the photoactivatable benzophenone to cross-link to any protein with which the ligand has bound. Purification of the samples using biotin/avidin affinity will then allow identification of the peptides by mass spectrometry.

5.4.5.2 Tags for Protein Labelling and Resin Conjugation

Within the Auer lab, a variation of the CONA screening technology is under development which aims at a systematic and systems based approach to the identification of modulators of protein-protein interactions. This involves the immobilisation of fluorescently labelled proteins *via* an affinity tag onto resin functionalised with the tag's reactive partner. For example, the bead immobilisation of fluorescently labelled His tagged proteins has been demonstrated using Nickel-NTA agarose resin. In order to screen compounds, a second protein which is known to interact with the immobilised one, labelled in a different colour, is then incubated. Binding can be visualised by co-localisation of the two fluorophores around the resin by confocal fluorescence imaging (Figure 5-10). Compounds can be added either before or after protein binding partner in colour 2, and the fluorescent image monitored. If dissociation of the protein-protein interaction is observed, the compound can be considered an inhibitor of that interaction.

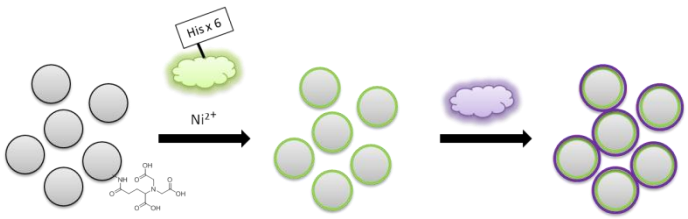


Figure 5-10: Principle of visualising protein-protein interactions on resin

For the analysis of data in this experiment, it is required that the resin material is changed from agarose to a more uniformly sized resin, such as TentaGel. It is also desired that proteins which are not fluorescently labelled but do contain an affinity tag can be used in the assay. Therefore, a trifunctional reagent was required that could simultaneously fluorescently label proteins and incorporate a functional group which would allow immobilisation of the labelled protein onto functionalised TentaGel resin (Figure 5-11).

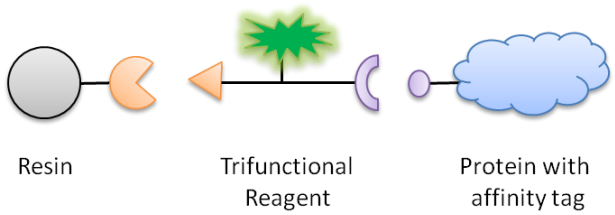
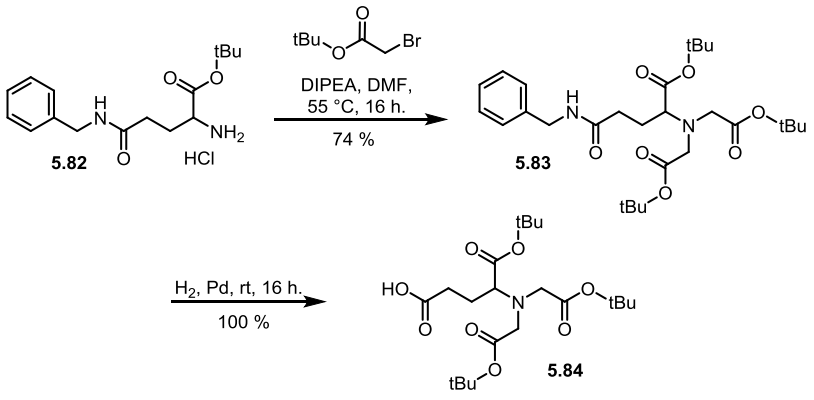


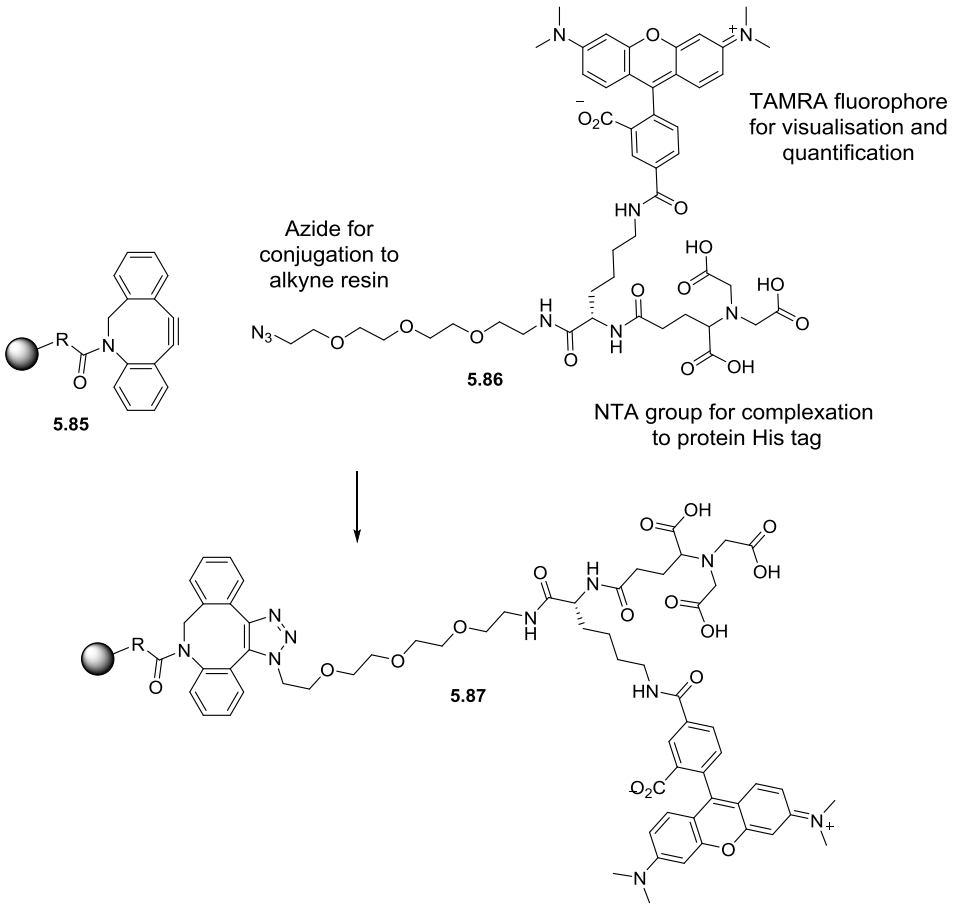
Figure 5-11: A trifunctional reagent is required to simultaneously fluorescently label proteins and immobilise them on resin *via* the protein affinity tag

Strain-promoted 1,3-dipolar cycloaddition chemistry was the chosen method for immobilisation of proteins onto resin, as it is completely bioorthogonal and can be carried out under a vast range of conditions. In order to test the feasibility of this approach, an azide-TMR-NTA tag was produced using TOBa, with the intention of immobilising His tagged proteins onto cyclooctyne functionalised resin (Scheme 5-23). Protected NTA monomer **5.84** was prepared for this purpose (Scheme 5-22).



Scheme 5-22: Synthesis of protected NTA monomer

This NTA reagent was applied to the TOBa backbone in replacement of the Fmoc protecting group. Subsequent removal of Dde and coupling of 5-TAMRA gave the required trifunctional reagent, which was cleaved from the resin using TFA to give **5.86** (Scheme 5-23).



Scheme 5-23: Trifunctional reagent prepared for fluorescent labelling and immobilisation of His tagged proteins onto DBCO resin.

Incubation of **5.86** with dibenzyl cyclooctyne (DBCO) functionalised resin **5.85** resulted in strongly fluorescent rings around each bead when analysed on PerkinElmer's Opera® High Content Screening System[§§§], suggesting that resin-tag conjugate **5.87** had been formed and therefore that immobilisation using this reaction is achievable (Figure 5-12 – More details in experimental section 8.4.5).

Next, the immobilisation of a protein using this compound was investigated. For this purpose, human CASK-interacting nucleosome assembly protein (CINAP), a 79 kDa protein that is involved in the modulation of gene expression,[193] was used. The CINAP protein contained a His-tag to enable complexation to the NTA functionality of **5.86** and was also Cy5 labelled, which would allow detection of the protein during experiments. Upon the addition of his-hCINAP-Cy5 in nickel containing buffer, only TMR fluorescence could be detected on the resin, whereas a fluorescent signal from Cy5 was absent, suggesting that

[§§§] Incubations and imaging carried out by Nhan Pham

complexation of the NTA portion of the trifunctional reagent with the protein His tag did not occur under these conditions.

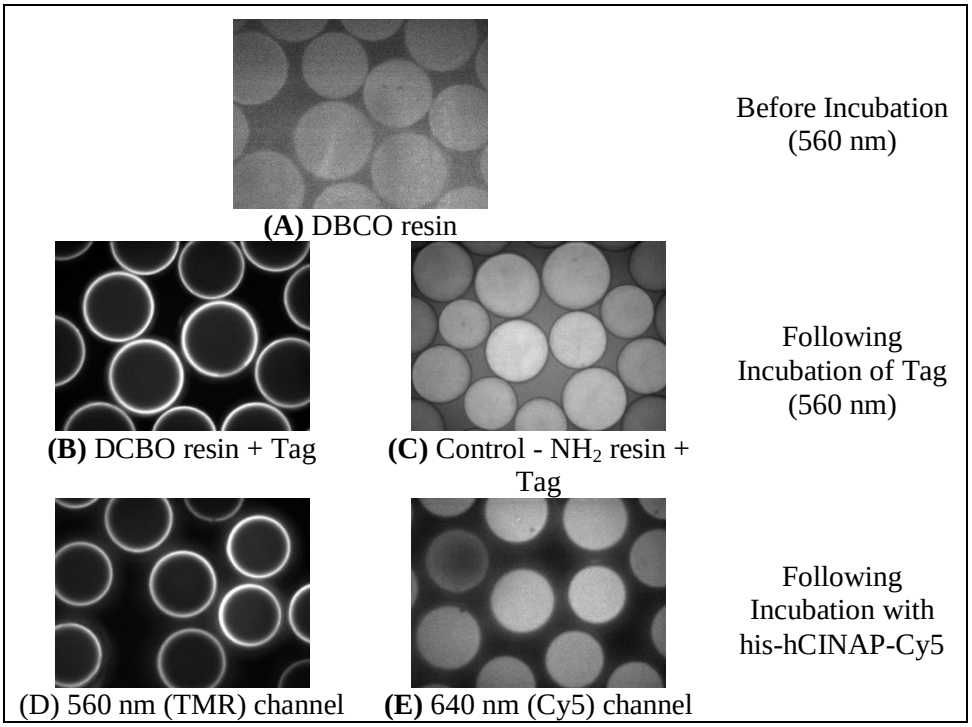


Figure 5-12: (A) DBCO resin (B) DBCO following 2 hour incubation with TMR-NTA-N_3 tag and washing – fluorescent rings can be seen indicating that the reaction of the tag with the resin was successful; (C) Control – NH_2 resin following 2 hour incubation with TMR-NTA-N_3 tag. No rings are seen, suggesting that the rings seen on the DBCO resin are from click reaction; (D) DBCO-tag resin following incubation with his-hCINAP-Cy5, TMR is still visible; (E) in the Cy5 channel, no his-hCINAP-Cy5 can be detected suggesting that the complexation between the tag and protein was unsuccessful.

In a second experiment, protein incubation with reagent **5.86** was carried out prior to pull-down onto resin, in the hope that complexation would occur when the reagent could freely diffuse throughout the mixture. Reagent **5.86** was incubated with His-hCINAP-Cy5 prior to addition of the DBCO resin. Unfortunately, rings in the 640 nm channel were again absent, whereas in a positive control using agarose-NTA resin His-hCINAP-Cy5 was clearly detected (Figure 5-13).

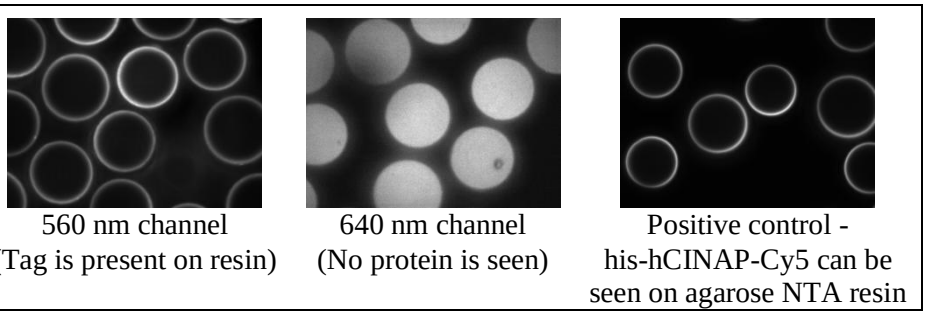
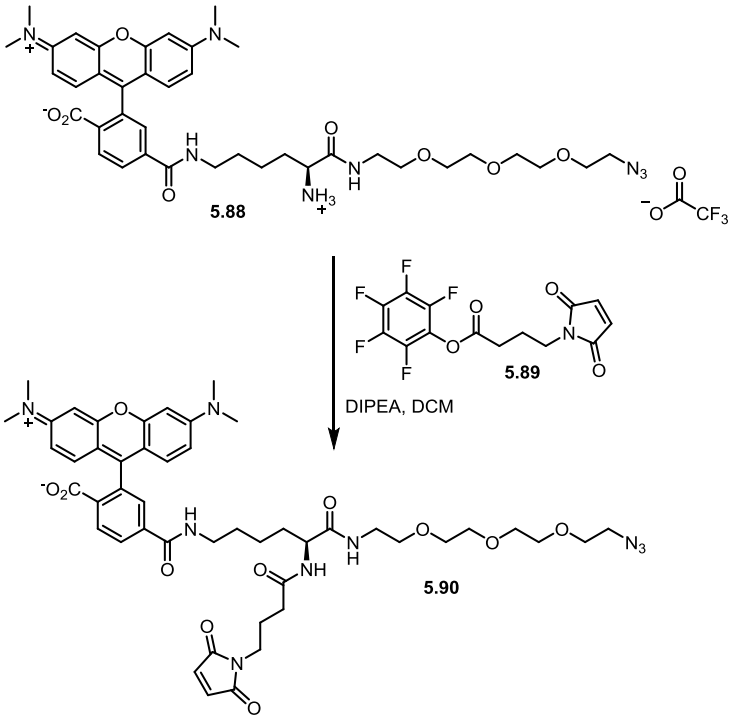


Figure 5-13: Images of resin following pre-incubation of tag with his-hCINAP-Cy5. TMR labelled tag was observed on the resin in 560 nm channel, but his-hCINAP-Cy5 was not observed in the 640 nm channel. In the positive control using agarose beads, the protein can clearly be seen in the 640 nm channel.

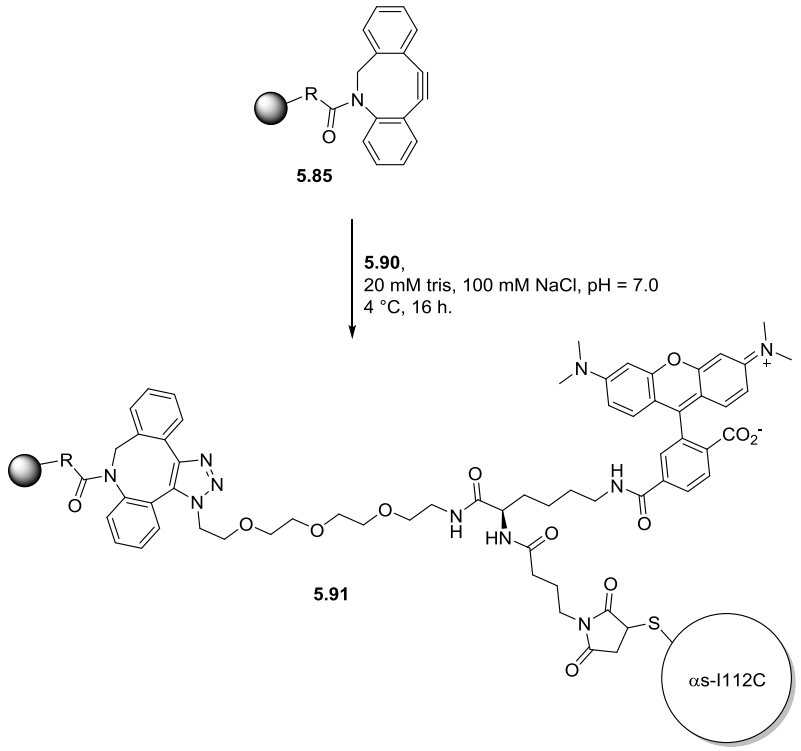
It was decided that a covalent bond between the trifunctional reagent and protein would better enable analysis. Therefore a new construct was synthesised incorporating a maleimide in place of the NTA (Scheme 5 24). Initial attempts to synthesise the complete tag on the solid support were impeded by hydrolysis of the maleimide group during cleavage from the resin. Instead, intermediate **5.88** was synthesised using TOBa and cleaved from the resin, allowing coupling of the pentafluorophenol ester functionalised maleimide **5.89** in solution, thereby avoiding subjection of this sensitive functionality to harsh cleavage conditions.



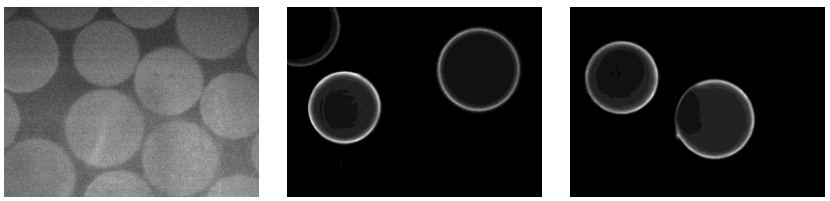
Scheme 5-24: Synthesis of TMR-meleimide-N_3 trifunctional reagent

Trifunctional compound **5.90** was used to label the alpha-synuclein mutant I112C at this single sulfhydryl site as a first evaluation experiment with a protein accessible in the lab

(Scheme 5-25).[****] Protein labelling was confirmed by HPLC and excess tag removed using a NAP™ column. The labelled protein was incubated with DBCO resin and following washing, confocal images taken (Figure 5-14 - Experimental details in 8.4.5).[††††]



Scheme 5-25: Loading of TMR-N_3-maleimide labelled alpha synuclein mutant I112C onto DBCO resin



DBCO Resin DBCO resin + labelled α-synuclein

Figure 5-14: Confocal images of DBCO resin before and after incubation with TMR-N_3-maleimide labelled alpha synuclein mutant I112C

The images show fluorescent halos around the beads, indicative of successful binding of the tagged protein to the resin *via* strain promoted 1,3-dipolar cycloaddition. This experiment demonstrates the use of the TOBa resin for the generation of trifunctional reagents which can be used to add multiple functionalities onto a single reactive site. Here, a protein bearing a single cysteine was fluorescently labelled and immobilised on a solid support using one

[****] Protein expressed by and labelling perfomed with Irene Pérez Pi.
[††††] Imaging performed by Nhan Pham

reagent. This construct will allow the development of assays based around immobilised alpha synuclein.

5.5 Conclusions and Future Work

An orthogonally protected backbone, TOBa, has been designed and synthesised on TentaGel resin, which allows rapid access to trifunctional reagents. By using amines as reactive points, high yielding and efficient amide bond formation is utilised to introduce required functionalities. The complete orthogonality between Fmoc, Dde and azide has been demonstrated, allowing total control over the regioselectivity of introduced functional groups. Two methods for the preparation of TOBa were developed, one in solution and the other on-bead. Both of these routes proved effective, however the solution phase route precludes the use of an allyl protecting group in place of the azide, which was shown to be effective on solid phase.

Previous backbones for the preparation of trifunctional reagents (reviewed in section 5.2.1.2 - 5.2.1.3) have been reported, but suffer from some inherent problems. In some cases, intramolecular reactivity was observed, leading to degradation of the tag upon storage.[183] In the case of TOBa, with all protecting groups intact, the compound is inert. It should be noted, though, that as Fmoc and Dde groups are sensitive to primary amines, reactions on TOBa should be carried out immediately following deprotection. Some backbone syntheses suffer from a lack of divergence, which means new compounds must be generated from early intermediates over numerous steps. TOBa allows a completely divergent synthesis, allowing the build-up of almost any trifunctional reagent in only three steps. As the backbone is immobilised on-bead, the need for purification between functionalisation reactions is circumvented. Also, the option to cleave the backbone at any stage of synthesis allows the use of sensitive functional groups which would not withstand harsh acidic resin cleavage conditions. The ease of synthesis of TOBa is also an advantage, with only three synthetic steps required for its solid supported synthesis or 4 steps in solution, the reagent can be easily accessed using simple chemical transformations. Finally, TOBa represents, to our knowledge, the first example of a triply orthogonally protected backbone for trifunctional reagent synthesis. This unique property allows the use of acid-containing building blocks for each point of backbone decoration. As many useful reporter and linkage reagents are available commercially in carboxylic acid functionalised form, this means that no modification is required for them to be utilised with TOBa.

TOBa functionalised resin is a useful tool for the general synthesis of multifunctional reagents and will be utilised in chemical biology applications and for the development of novel screening methods. This was demonstrated with the synthesis of a tag bearing an azide, photoactivatable cross-linker and biotin, which will be used in the future for target identification and validation.

The TOBa functionalised resin was also used to generate a series of tags for simultaneous labelling and immobilisation of proteins onto TentaGel resin for the development of assays to probe protein-protein interactions. A TMR-NTA-N_3 reagent was synthesised for immobilisation of His-tagged proteins onto alkyne functionalised TentaGel resin. Although the pull-down reaction proved to work well, formation of a complex between the NTA and His tag did not take place, preventing the success of this experiment. Future development will involve optimisation of complexation conditions.

Sensitive groups such as maleimides were found to degrade upon acid cleavage of the resin bound backbone. This problem was circumvented by release of tag compounds at intermediate stages of synthesis, allowing completion of syntheses in milder conditions in solution. Synthesis of a TMR-N_3-maleimide tag was performed using this approach. The tag was used for cysteine labelling and immobilisation of α-synuclein mutant I112C onto alkyne functionalised TentaGel resin. This construct will be used for the development of assays probing α-synuclein aggregation and ligand binding.

Future applications of the TOBa resin are potentially wide ranging, however a particular focusa of the group would be in the tagging of individual compounds, whole libraries or proteins. Single compounds could be tagged for localisation studies, for example using a cell penetrating peptide and a fluorescent dye to assess where compounds which are cell impermeable would travel when carried inside cells. Another example would be the use of two different tracing agents such as a chelating group for a radioactive element and a fluorescent dye for simultaneous fluorescence and PET imaging of compounds. Alternatively, two different compounds could be attached to the scaffold with one reporter group. For example, if a target protein were known to form a homodimer, then two ligands could be tethered onto the backbone in an attempt to increase potency of its interaction. A reporter group such as a fluorophore could also be included on the third arm for the purpose of assay readout using, for example, fluorescence polarisation.

Proteins could also be tagged using trifunctional reagents generated through TOBa. As demonstrated already, the simultaneous immobilisation and fluorescent labelling of preoteins

will be used for the development of assays to identify protein-protein interactions. The rapid synthesis of trifunctional reagents using TOBa will allow a combinatorial approach to the synthesis of these reagents, allowing functional groups to be easily substituted to simplify optimisation of assay conditions.

The future development of a trufunctional crosslinking reagent which would allow whole libraries to be tagged and incubated with a cell lysate to generate a multi-ligand, multi-target screen is a goal of the Auer lab. In such an assay, ligands with affinity for a taget would be covalently bound using a photoreactive crosslinker and then complexes formed would be purified for subsequent MS identification through use of a biotin group or another suitable purification moiety. For this goal to become reality, a suitable purification method must be established. In the following chapter, one such method is investigated in a simpler setup. Here, the development of a trifunctional reagent is described which incorporates two cross-linking groups and a salicylhrdroxamic acid for purification of cross-linked peptides. Its use in the cross-linking of a protein in order to elucidate structural information is also detailed.

CHAPTER 6

NOVEL PROTEIN CROSS-LINKING REAGENTS

6.1 Abstract

The usefulness of trifunctional reagents has been demonstrated in the field of chemical cross-linking of proteins for analysis with mass spectrometry. In this rapidly developing method for protein structure determination, the third arm of the trifunctional reagent is used for purification. Cross-linking and protein digestion leads to a large number of non-modified peptides and only a small proportion of fragments which have been chemically cross-linked, leading to a low signal upon MS analysis. Enrichment of cross-linked fragments can be carried out by incorporating a purification group into a cross-linker. In this chapter, the development of an enrichable cross-linker is described, the mode of purification of which relies on the complex formed between salicylhydroxamic acids and boronates. A heterobifunctional cross-linker was also synthesised and evaluated by performing intramolecular cross-linking with human serum albumin (HSA). Cross-links using this reagent were found to be almost exclusively complimentary to those seen when using frequently employed sulfodiazirine (SDA) cross-linker. This bezophenone cross-linker will be used in combination with other cross-linkers in future studies to allow the extraction of more cross-linking data from protein samples, leading to higher resolution structural information being obtained.

6.2 Introduction

6.2.1 Protein and Peptide Cross-linking

Chemical cross-linkers are tools used to generate covalent bonds between interacting proteins and protein domains that are in close proximity. This introduces distance constraints, which essentially freezes interactions in time, generating a tether between the proteins which can withstand denaturing conditions. Proteins can be digested into peptides and analysed by mass spectrometry, with cross-links keeping the three-dimensional structural information intact. Such experiments allow the probing of multi-protein complexes and provide structural data.

Chemical cross-linking and mass spectrometry can be considered as a complimentary approach to other protein structure elucidating methods such as X-ray crystallography, nuclear magnetic resonance (NMR), electron paramagnetic resonance (EPR) and Förster resonance energy transfer (FRET).

X-ray crystallography involves the generation of crystallised proteins which are subjected to intense x-ray beams.[194] The beams are diffracted and generate patterns, which are then

analysed and a detailed map of the electron density in the protein can be generated. Very ordered crystals allow highly detailed atomic information to be gained about a protein, however this technique does have some drawbacks, with the main one being the difficulty in producing protiein crystals (extremely difficult for flexible proteins). This method also only gives a snapshot of the protein when packed in a crystal and does not allow the sudy of these dynamic entities in their natural environment.

NMR spectroscopy overcomes this issue, by allowing proteins to be analysed in solution.[195] Here, proteins are purified and subjected to radio waves in a strong magnetic field. Observed resonances allow the build up of information about nuclei which are close in space, which can eventually be used to generate a model of the protein showing the locations of all atoms. Disadvantages of this technique, however, are the restrictions on the size of protein which can be analysed (small to medium only, otherwise signal overlap is observed), the relatively large amount of protein required for analysis and the high cost of running the experiments. Proteins are also required to be isotopically labelled in order for a signal to be observed.

FRET and EPR are both techniques from which information about protein dynamics can be gained as well as protein structure. In FRET, the energy transfer efficiency between a suitable donor and acceptor chromophore on a biomolecule can be used to measure the distance between the two.[196] This technique is most sensitive over a few nanometres, making it ideal for probing protein structures. It can also be carried out using very small quantities of material and is highly sensitive. Disadvantages of this method include the fact that in a single experiment only the distance between two sites can be measured, and that it requires the protein to be fluorescently labelled specifically at the site that a measurement is to take place, which often requires mutagenesis. This can also affect the characteristics of the protein being studied.

EPR spectroscopy involves the labelling of proteins with labels which contain unpaired electrons (such as the nitroxyl radical). The tagged proteins are then subjected to a magnetic field, with which the unpaired electrons interact. Spin-spin interactions between two labels on a protein can be quantified to give distances between them.[197] Although proteins of any size can be studied in solution using this technique, it does not offer the high resolution data that NMR or crystallography canand requires the labelling of proteins with single electron tags.

Chemical cross-linking and mass spectrometry methods have also not yet matured enough to be able to offer the kind of atomistic protein models obtained from NMR or x-ray

crystallography, however as a technique it holds some advantages over its higher resolution counterparts, such as;

- Only a small amount of sample is required compared to NMR
- Proteins are analysed in solution, so crystals are not required
- Short lived interactions can be captured
- Proteins can be cross-linked in their native environment

The cross-linking/MS process can be roughly divided into four sequential steps,[198] shown in Figure 6-1.

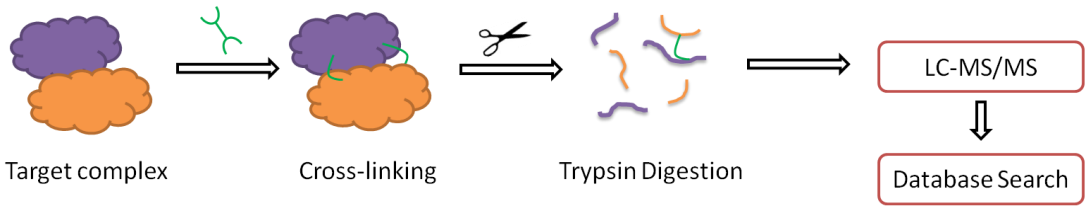


Figure 6-1: The cross-linking/MS process

1. **Cross-linking** – Chemical compounds are used to enable inter or intramolecular cross-links in proteins *via* functional groups on amino acid side chains, usually amines or sulfhydryl groups (discussed *vide infra*). The extent of this can be approximated using gel electrophoresis or MALDI-TOF.
2. **Trypsin Digestion** – Proteins are enzymatically cleaved into peptide fragments. Cross-linked peptides stay covalently bound to one another.
3. **LC-MS/MS** – Techniques such as MALDI and ESI allow the transfer of large, intact biological compounds into the gas phase. Molecular weights of peptides and their fragmentation patterns can be determined.
4. **Database Mining** – Large amounts of complex data detailing peptide fragments are subjected to searches in databases in order to identify cross-linked peptides.

The analysis of such a process can be complicated due to the number of peptide species present in a sample following cross-linking and digestion (Figure 6-2). These include: peptides which have not been cross-linked; species in which the cross-linker has reacted at one site only, with the other end having been hydrolysed (type 0); cyclic peptides, where the cross-linker has reacted twice along the same fragment (type 1); cross-linked peptides where one end of the cross-linker has reacted with a fragment different to the other end (type 2); and finally higher order peptides, where many different cross-links have occurred (type 3).

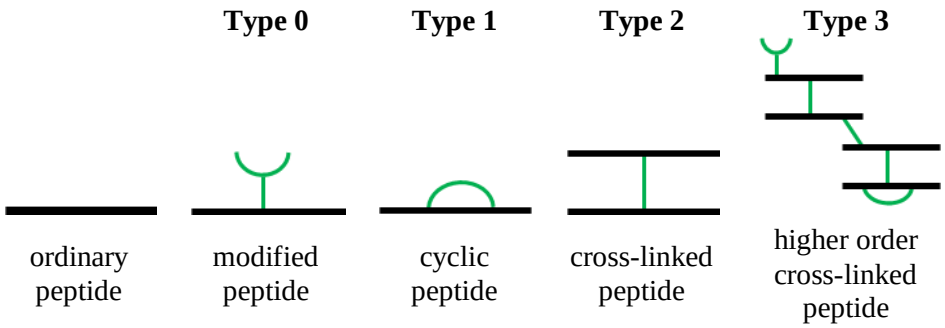


Figure 6-2: Peptide types observed following cross-linking and digestion

6.2.2 Mass Spectrometric Analysis of Cross-Linking Products

There are two methods for the analysis of proteins by mass spectrometry; MALDI (matrix assisted laser desorption/ionisation) and ESI (electron spray ionisation). These soft ionisation techniques allow large biopolymers to enter the gas phase intact, so that molecular ions can be detected. Peptidic fragments of proteins can also be detected to a high degree of accuracy, following digestion.

6.2.2.1 Matrix Assisted Laser Desorption Ionisation (MALDI)

The term MALDI was first coined in 1985 when it was noticed that the ionisation of alanine by laser irradiation could be carried out more easily if mixed with tryptophan.[199] The tryptophan acts as an absorbing matrix for the laser energy and helps to ionise the non-absorbing alanine.

MALDI allows the ionisation of non-volatile bioorganic compounds using very small amounts of sample that can be easily prepared and measured in a short time.[200] This very sensitive approach is robust and tolerant to a range of buffers. In MALDI, sample is mixed with a matrix, typically made up of aromatic acids (e.g. cinnamic acid derivatives such as 4-hydroxy-3-methoxycinnamic acid),[201] then introduced to a vacuum chamber. Short pulses of laser light are then fired at the sample, inducing ionisation of the matrix which then transfers protons to the sample molecules, generating charged species. The coupling of MALDI to a TOF (time of flight) analyser gives very high sensitivity and detection of high mass ions with very good mass accuracy.

6.2.2.2 Electrospray Ionisation (ESI)

Electrospray ionisation was first described in 1984 as a technique to enable mass spectrometric analysis of large molecules which are too complex, large, fragile or non-volatile for ionisation by other existing methods.[202] Here, high voltage is applied to a

liquid sample to generate an aerosol of charged species. As this is a soft ionisation technique, very little fragmentation is seen, so ESI-MS does not offer much structural information. This can be overcome by coupling ESI with tandem MS/MS.

6.2.3 Cross-linking Techniques

Cross-linkers generally contain chemical groups which are specifically reactive to certain amino acid side chains in proteins, in most instances primary amines of lysine amino acids and sulfhydryl groups present in cysteine residues. These nucleophilic residues can be trapped using suitably designed electrophilic cross-linking reagents.

6.2.3.1 Amine Reactive Cross-linking Groups

Groups that are reactive towards amines include *N*-hydroxysuccinimidyl esters (NHS esters) **6.1**, sulfo NHS esters (NHSS esters) **6.2** or imidoesters **6.3** (Figure 6-3).

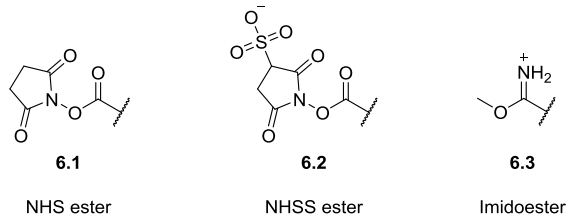
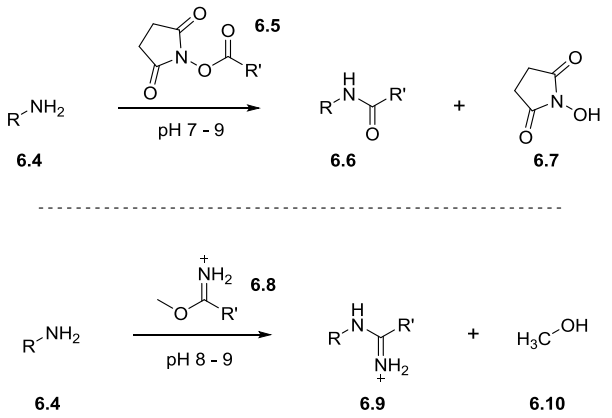


Figure 6-3: Cross-linking groups which are reactive towards amines

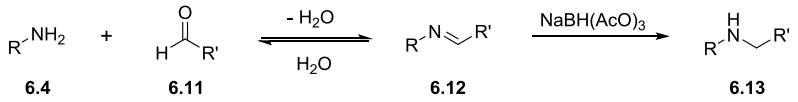
Under basic conditions, these groups specifically react with the free primary amines on a protein, such as those on lysine side chains or the protein *N*-terminus (Scheme 6 1). As amines are charged at physiological pH, they tend to be over-represented in solvent accessible (surface) areas of protein structures, meaning they are readily accessible for cross-linking.[203]

N-hydroxysuccinimidyl (NHS) esters have been used as amine reactive probes for decades, and generate stable amide bonds upon reaction with amines. Although NHS react very quickly, hydrolysis of the ester group can also occur, which leaves a non-reactive carboxylate group and NHS. Highly soluble sulfo-NHS esters allow dissolution directly into aqueous media, whereas NHS esters often require to be dissolved first in organic solvents.[204] Iminoesters react with amines to form an amidine bond, however this has a relatively short half-life and side reactions can occur.



Scheme 6-1: Amine-reactive groups. NHS ester (above) and imidoester (below)

Reductive amination is another method by which protein amines can be modified. Here, an aldehyde is used as the reactive partner, with which the nucleophilic amine reacts to generate an imine (also called a Schiff base, Scheme 6-2). Unlike with hydrazine or hydroxylamine based reagents, where this type of reaction results in a relatively stable hydrazone, the reaction between amines in a protein and an aldehyde is reversible due to hydrolysis. Therefore, a mild reducing reagent (sodium triacetoxyborohydride) is employed which selectively reduces the imine, resulting in a stable secondary amine.

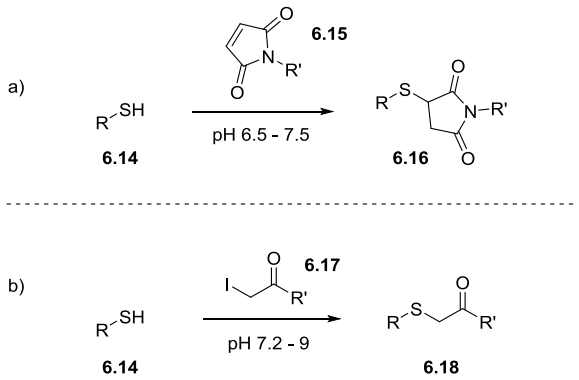


Scheme 6-2: Modification of protein amines by reductive amination using aldehydes. Imine (or Schiff base) formation is reversible until reduced to a secondary amine.

Although many other groups can be used to modify amines, such as isocyanates, sulfonyl chlorides, anhydrides, epoxides and pentafluorophenyl esters, the most prevalently used in protein chemistry are NHS esters and iminoesters.

6.2.3.2 Sulfhydryl Reactive Cross-linking Groups

Maleimide and iodoacetyl groups are often used in the cross-linking of sulfhydryl groups present on cysteine side chains (Scheme 6-3). The cross-linking occurs *via* a conjugate addition into the maleimide or alkylation of the iodoacetyl by the thiol.



Scheme 6-3: Reaction of a sulfhydryl group with a maleimide (a) and iodoacetyl (b) group to form a stable thioether

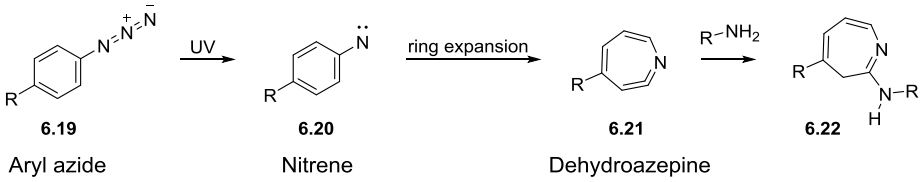
These reactions result in the formation of a stable thioether linkage, however disulfide bonds in a protein must be reduced in order to free up thiol groups for reaction. This reaction also precludes the use of other thiol groups (e.g. reducing agents) being used in buffers as they will compete for reactivity.

6.2.3.3 Photoreactive Cross-linking Groups

Photoreactive compounds are reasonably chemically inert and only become reactive following exposure to UV or visible light.[204] Although this feature allows their application under highly controlled conditions, their reactivity is not as specific as the groups mentioned in section 6.2.3.2, therefore predicting fragments from cross-links made with these species can be difficult.

6.2.3.3.1 Aryl Azides

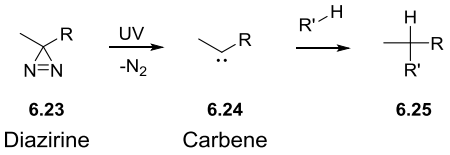
Aryl azides are useful photoreactive cross-linking reagents. Upon UV irradiation at wavelengths around 360 nm, aryl azides form nitrenes which are short lived and react rapidly with the surrounding environment (Scheme 6-4).[205] This group inserts non-specifically into double bonds, active hydrogen bonds and C-H and N-H sites. Ring expansion can also occur, to create a seven membered dehydroazepine, which preferentially react with nucleophiles, specifically amines.



Scheme 6-4: Aryl azide cross-linking mode of action

6.2.3.3.2 Diazirines

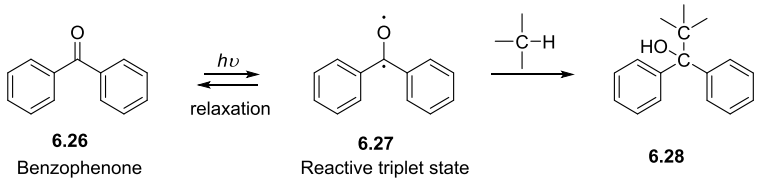
Diazirines consist of a three membered ring containing two nitrogen atoms connected by a double bond (Scheme 6-5). A highly reactive carbene species is formed upon irradiation of diazirines with UV light, which can insert into C-H or N-H bonds, or into points of unsaturation.



Scheme 6-5: Diazirine mode of action

6.2.3.3.3 Benzophenones

Benzophenones are stable to normal light conditions and are inexpensive and commercially available with a variety of functional groups attached. Irradiation of these groups forms a reactive triplet state, which can insert non-specifically into e.g. amide bonds. They have a unique advantage as photoreactive cross-linkers, in that following formation of the triplet state, if reaction does not occur the benzophenone can relax and return to its non-reactive ground state (Scheme 6-6).[206]



Scheme 6-6: Benzophenone cross-linking mode of action

This means that irradiation can be repeated, which can lead to a higher number of cross-links being achieved.

6.2.4 Bifunctional cross-linkers

Bifunctional cross-linkers allow the covalent linkage of two moieties *via* two reactive chemical groups (Figure 6-4) separated by a spacer. Many bifunctional cross-linkers are commercially available with different properties and reactive groups to target different amino acids on proteins.

Figure 6-4: Homo- and heterobifunctional cross-linkers

A particularly important consideration in choice of cross-linker is the distance between the two reactive groups, as this determines the length over which cross-links can be made. Other considerations during cross-linker design/purchase are hydrophobicity/hydrophilicity (or solubility) and whether the spacer should incorporate a cleavable moiety, so that cross-linked products can be separated.

Homobifuctional cross-linkers possess the same reactive groups at each end of the spacer, therefore both ends react with similar groups on the protein (e.g. amines, sulfhydryl). Although widely used, these cross-linkers do have drawbacks, as products from cross-linking can be poorly defined. For example, when one end of the reagent reacts with a protein, it becomes an active intermediate. This species can then cross-link with a different protein, another molecule of the same protein, or in an intramolecular fashion. In some cases they may even react with proteins which have already been cross-linked, generating various sized oligomers which may polymerise and precipitate.[207] Some examples of commercially available homobifunctional cross-linkers are given below (Table 6-1).

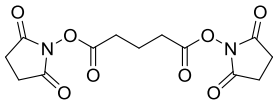
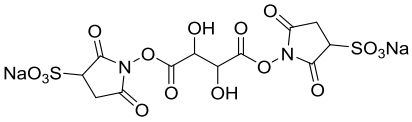
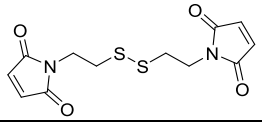
Entry	Structure	Name	Reactive to	Spacer length (Å)	Cleaved with
1		DSG	Amino	7.7	-
2		Sulfo-DST	Amino	6.4	Periodate
3		DTME	Sulfhydryl	13.3	Reducing agents

Table 6-1: Examples of homobifunctional cross-linkers Disuccinimidyl glutarate (DSG), Disulfosuccinimidyl tartarate (Sulfo-DST) and Dithiobismaleimidoethane (DTME) and their properties

Heterobifunctional cross-linkers have two different reactive groups present to join two separate functional groups. In the case of NHS-Diazirine (SDA, Table 6-2, entry 1), an NHS ester allows reactivity towards an amine and a diazirine allows photo-inducible non-specific reactivity. This type of reagent can offer more control over cross-linking reactions, as the NHS ester can be reacted first, excess reagent removed, and then the diazirine irradiated for cross-linking. Some examples of homobifunctional cross-linkers are given below (Table 6-2).

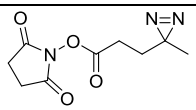
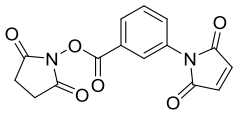
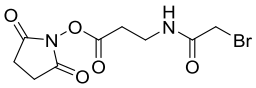
Entry	Structure	Name	Reactive to	Spacer length (Å)
1		SDA	Amino, anything	3.9
2		MBS	Amino, sulfhydryl	7.3
3		SBAP	Amino, sulfhydryl	6.2

Table 6-2: Examples of heterobifunctional cross-linkers NHS-diazirine (SDA), *m*-maleimideobenzoyl-*N*-hydroxysuccinimide ester (MBS) and succinimidyl-3-(bromoacetamido)propionate (SBAP) and their properties

6.2.5 Enrichable Cross-linkers

Despite the benefits that cross-linking and MS offer as a technique, analysis of data is not straightforward. The number of non-modified peptides in a mixture far outweighs the

number of peptides bearing cross-links, therefore their analysis becomes akin to looking for a needle in a haystack.[204] In order to aid the detection of these peptides, a number of strategies have been developed. These include the incorporation isotopes or fluorescent groups into cross-linking reagents, the generation of enrichment tags and the use of MS techniques which provide high mass accuracy (FTICR-MS).[204] Enrichment tags will be the focus of subsequent discussion.

6.2.5.1 Biotin Enrichment

The synthesis of trifunctional linkers containing two cross-linking groups and a biotin moiety has been reported for enrichment of cross-linked proteins from a protein mixture.[208] Sulfo-SBED is one such construct which is now commercially available (Thermo Scientific Pierce - Figure 6-5). This trifunctional reagent consists of two cross-linking groups; an amine reactive NHSS ester and a photoactive aryl azide. It also possesses a biotin moiety for affinity purification of cross-linked species using monomeric avidin.

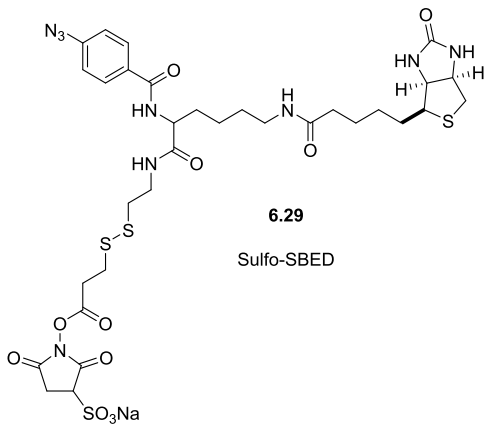


Figure 6-5: Photo-cross-linking and biotin enrichment reagent

In 2004, Sinz *et al*. reported that this construct could be used to efficiently identify interacting amino acid sequences in the complex between calmodulin (CaM) and its target peptide M13, which are part of the skeletal muscle myosin light chain kinase.[209] Following purification, the peptides observed by MS were almost exclusively ones which had been modified with the sulfo-SBED tag. Before avidin enrichment, these had only been observed as weak signals in the original reaction mixture following cross-linking and digestion. In this report, however, some non-specific binding of peptides to avidin beads was observed.

Sulfo-SBED has also been used to demonstrate enrichment of an intramolecular cross-link using the peptide neurotensin as a model system.[210] The results of this study were promising

for the enrichment, however the authors commented on non-specific binding of peptides to avidin resin.

6.2.5.2 1,3-Dipolar Clycloaddition (Click) Chemistry Enrichment

In 2007, Kasper *et al*. introduced bis(succinimidyl)-3-azidomethyl glutarate (BAMG), a homobifuctional cross-linker which incorporates an azide, a functional group not present in biological systems, on the linker backbone (Figure 6-6)[211].

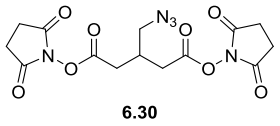
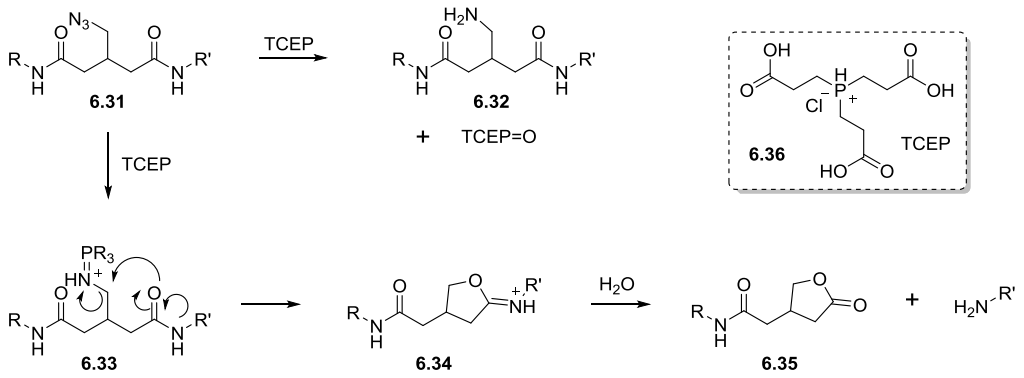


Figure 6-6: Bis(succinimidyl)-3-azidomethyl glutarate (BAMG)

BAMG was developed in order to separate cross-linked peptides from non-modified peptides following digestion of proteins. Upon Staudinger reduction (e.g. with TCEP), the azide group undergoes two competing reactions. It can either be reduced to the corresponding amine, or conversion to a leaving group that is displaced by the neighbouring carbonyl oxygen of one of the two amide groups by an intramolecular S_N2 reaction to form an imido ester, which is then hydrolysed (Scheme 6-7). As the compound is symmetrical, this can occur on either side.

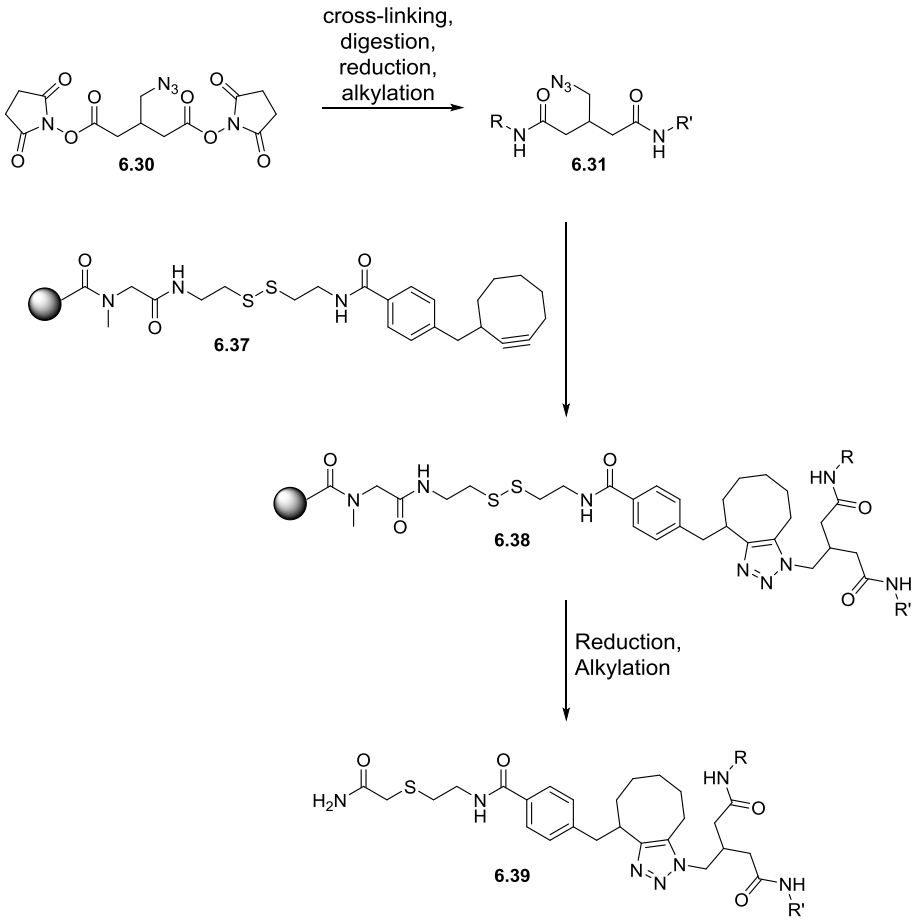


Scheme 6-7: Two reactions which can occur upon subjection of cross-linked BAMG to TCEP

This allows separation of cross-linked peptides from non-modified peptides *via* reversed-phase (RP) diagonal chromatography. Here, digested peptides are subjected to fractionation using RP chromatography. The fractions are then modified (in this case, TCEP reduction) and another chromatography step is performed, whereby modified peptides will have a significantly different retention time to the non-modified peptides due to the structural changes they have undergone. The two reaction methods also allow for the isolation and

detection of the parent cross-linked peptide fragment and the two cleaved fragments from the unmodified peptide bulk, which will increase the sensitivity of the cross-linked peptide signals and allow multiple masses of each cross-link, depending on the reaction of the azide. Although this method allowed enrichment of a cross-linked sample, it may have some drawbacks, including whether reducing agents in protein samples would prematurely reduce the azide group, and the possible loss of material through running two chromatography separations on each sample.

Later, in 2009, BAMG was used in combination with bioorthogonal copper-free strain promoted click chemistry in order to purify cross-linked peptides from non-modified peptides.[212] Here, azide reactive cyclooctyne resin (ARCO) was used to purify azide containing peptides (Scheme 6-8).

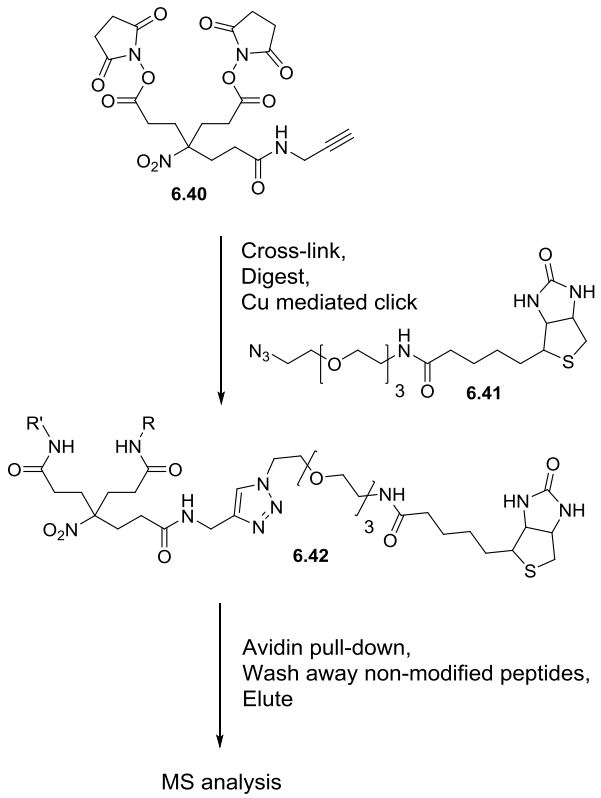


Scheme 6-8: Enrichment of azide containing peptides using ARCO resin

Cytochrome c was cross-linked *in vitro* by BAMG and the protein reduced (non-TCEP conditions) digested and analysed by MS, which showed a mixture of modified and non-modified peptides. Following enrichment, the MS data revealed only a single unmodified

peptide, amongst numerous examples with the expected modification. Although this data shows the ARCO/BAMG enrichment strategy to work effectively, the hydrophobicity of the cyclooctyne resulted in modified peptides being highly insoluble in aqueous buffers. It was also found that unreacted cyclooctyne dominated MS spectra and a chromatography step had to be introduced to remove it.

Copper mediated Huisgen 1,3-Dipolar Cycloaddition 'click' chemistry was also utilised for the enrichment of cross-linked peptides by Chowdhury and co-workers in 2009.[213] A compact cross-linker named CLIP (click-enabled linker for interacting proteins) was synthesised which incorporated a terminal alkyne for click conjugation and a nitro group for solubility and to act as a neutral fragment group (Scheme 6-9).



Scheme 6-9: Click-enabled linker for interacting proteins (CLIP)

Following CLIP cross-linking and digestion, click chemistry was used to react an azide functionalised biotin moiety onto the terminal alkyne now present on cross-linked peptides. Purification of these peptide fragments using monomeric streptavidin was then carried out. The procedure was successfully applied to the enrichment of cross-linked ubiquitin from a complex biological mixture, with results found to be identical to those seen from cross-linking purified ubiquitin. This method has several drawbacks, however, including the harsh

conditions required for elution of biotin from avidin resin, and the fact that samples may be contaminated with natural biotin containing peptides present in *E.coli*.

Very recently, a new enrichable strategy for *in vivo* cross-linking and mass spectrometry was described by Kaake *et al.*, which allows the definition of protein-protein interactions in living cells.[214] This approach utilises an enrichable cross-linker named Azide-A-DSBSO (Azide-tagged, Acid-cleavable DiSuccinimidyl BisSulfOxide) which boasts many features (Figure 6-7).

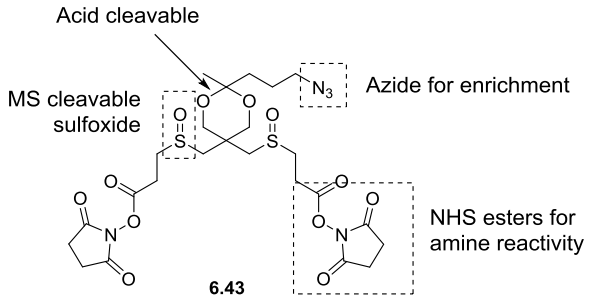
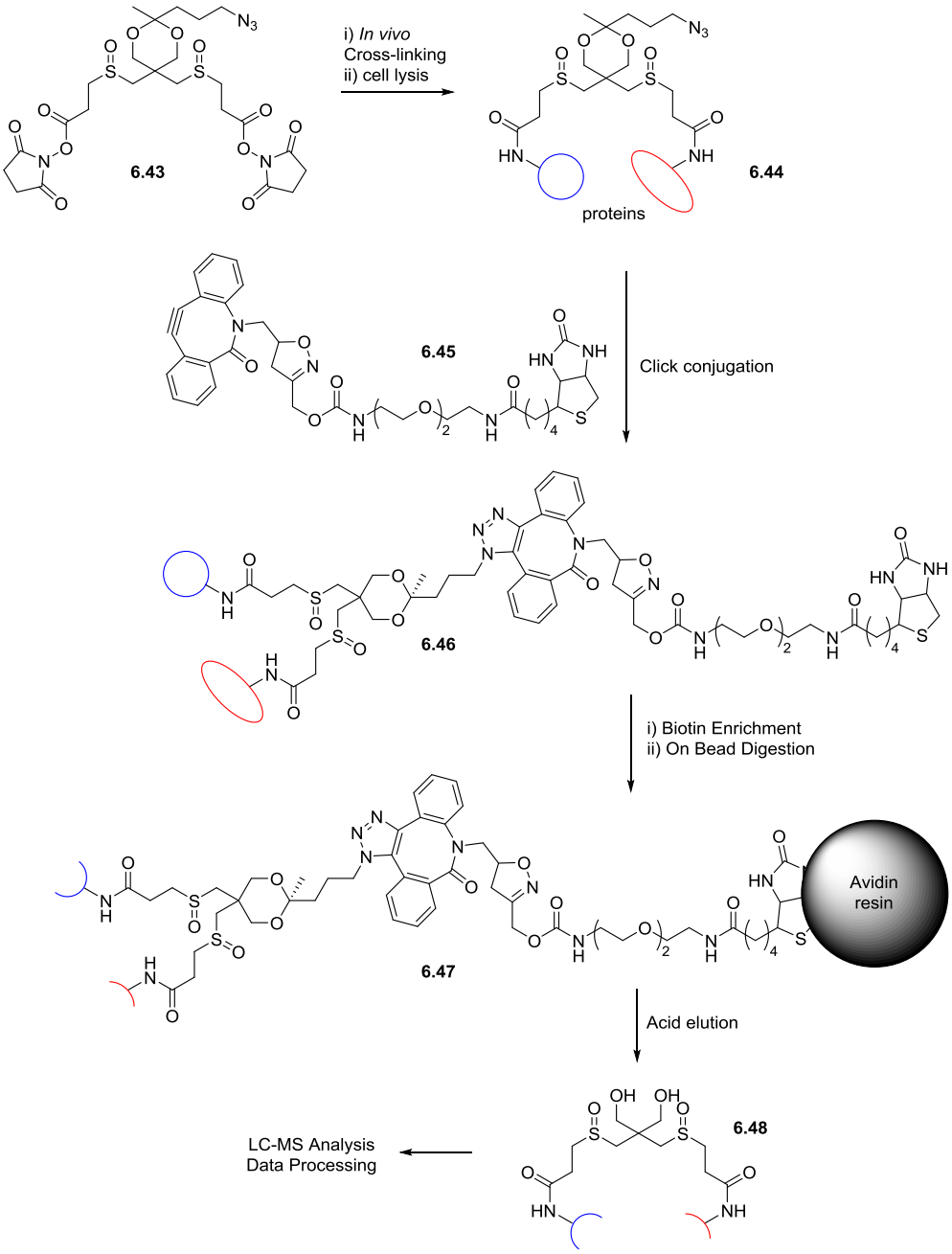


Figure 6-7: Azide-A-DSBSO (Azide-tagged, Acid-cleavable DiSuccinimidyl BisSulfoxide)

The cross-linker bears an azide for bioorthogonal enrichment of modified peptides, an acid cleavable site to enable liberation of cross-linked peptides from a solid support following enrichment and two sulfoxide groups which are MS cleavable, to allow fast and unambiguous identification of modified peptides during MS analysis.

The aim of this work was to map protein-protein interactions in living cells, which was achieved by the following workflow (Scheme 6-10): cross-linking was carried out in live HEK293 cells which were then lysed. The azide was tagged with a heterofunctional linker consisting of a cyclooctyne (for copper-free click to the azide group of Azide-A-DSBSO) and a biotin, which was used to pull the modified proteins down onto avidin beads. The resin is then washed, removing non-cross-linked proteins and digestion of modified proteins carried out on the resin, leaving only cross-linked peptides attached. The acid cleavable site on the Azide-A-DSBSO can then be utilised to elute the modified fragments from the resin. Incorporation of a cleavable site in the tag means that the bulky groups required for copper-free click are removed from the cross-linked peptides, which circumvents associated hydrophobicity problems. The enriched fragments are then analysed by MS to generate an in-vivo protein-protein interaction map.



Scheme 6-10: Azide-A-DSBSO cross-linking workflow

This approach allowed the identification of numerous novel cross-links across a wide range of protein-protein interactions implicated in various signalling pathways. The possibility to focus on a certain protein complex was also demonstrated by carrying out affinity purification of a certain protein prior to enrichment and analysis.

6.2.5.3 Summary

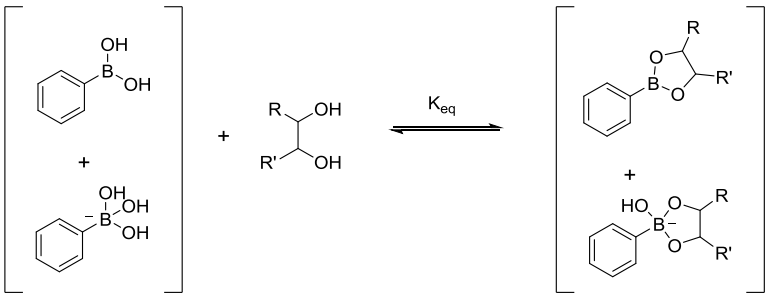
The enrichment methods summarised above all rely on either click chemistry or biotin-avidin interactions for the purification of cross-linked samples. Indeed, some of the methods which use click chemistry still utilise the biotin-avidin interaction as part of their workflow, which is a testament to the power of the biotin/avidin interaction and their ability to 'find' each other in complex mixtures.

However, both of these methods of purification hold some disadvantages, including non-specific interactions between non cross-linked proteins and avidin, the difficulty in elution of biotinylated material from avidin and the inherent inseparability of click reaction products. In order to circumvent these problems, cleavable cross-linkers have been developed which risk being prematurely dissociated and involve complex syntheses. Multistep purification methods also alleviate these problems but complicate workflows, are time consuming and risk the loss of precious material.

Another way to circumvent these problems would be to use a different method of purification. A bioorthogonal, reversible interaction between a reactive species on a cross-linker and a purification medium would allow simple purification without the need to compromise the stability of the cross-linking reagent. Work towards such an approach by the Rappsilber group (University of Edinburgh) is discussed below.

6.2.6 Towards an Enrichable Cross-linker using Reversible Boronate Ester Conjugation

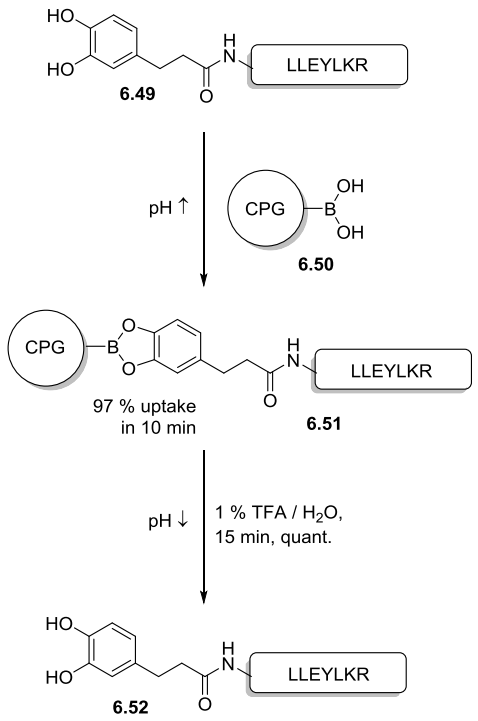
The Rappsilber lab have made promising progress on the development of an efficient tag for protein cross-linking and enrichment.[215] Their aim was to develop a compound which takes advantage of the reversible complex formed between phenylboronic acids and a variety of polar groups, such as diacids, α-hydroxy carboxylic acids and diols in aqueous media to form boronate esters.[216-222] The most common form of the complex is with polyols (mostly 1,2- and 1,3- diols) to form 5- or 6- membered heterocyles (Scheme 6-11).



Scheme 6-11: Reversible boronate ester formed between phenyl boronic acid and a 1,2- diol

This interaction had already been exploited by the immobilisation of phenylboronates for affinity purification of a range of compounds, including 2-hydroxycarboxylic acids,[223] carbohydrates and RNA.[224] Conversely, a range of solid supported reagents for the removal of boronic acids from organic solvents using this interaction have been reported.[225, 226]

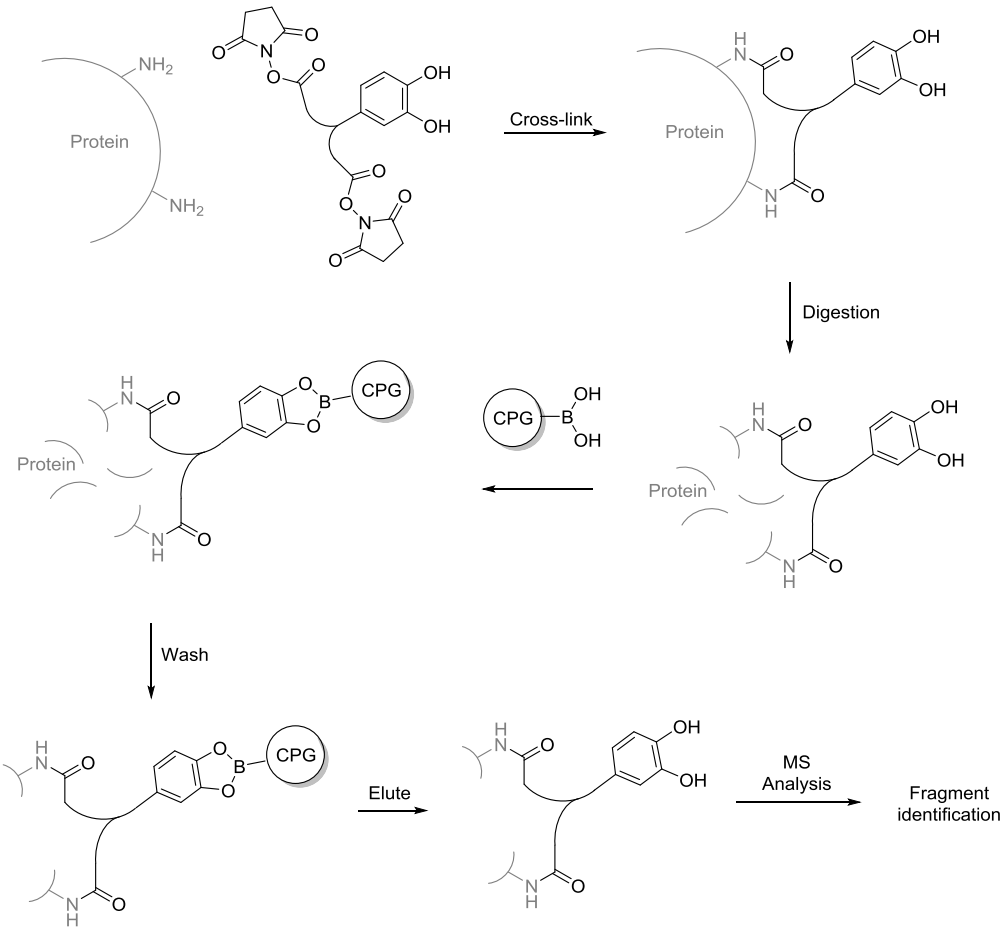
In order to test the applicability of this interaction for the enrichment of cross-linked peptides, Dr. Adam Belsom (Univeristy of Edinburgh) synthesised catechol tagged peptide LLEYLKR and assessed the extent of enrichment when subjected to controlled pore glass (CPG) functionalised with arylboronic acid (Scheme 6-12).



Scheme 6-12: Capture and release of catechol tagged peptide LLEYLKR using boronic acid functionalised controlled pore glass (CPG) *via* reversible boronate ester complexation

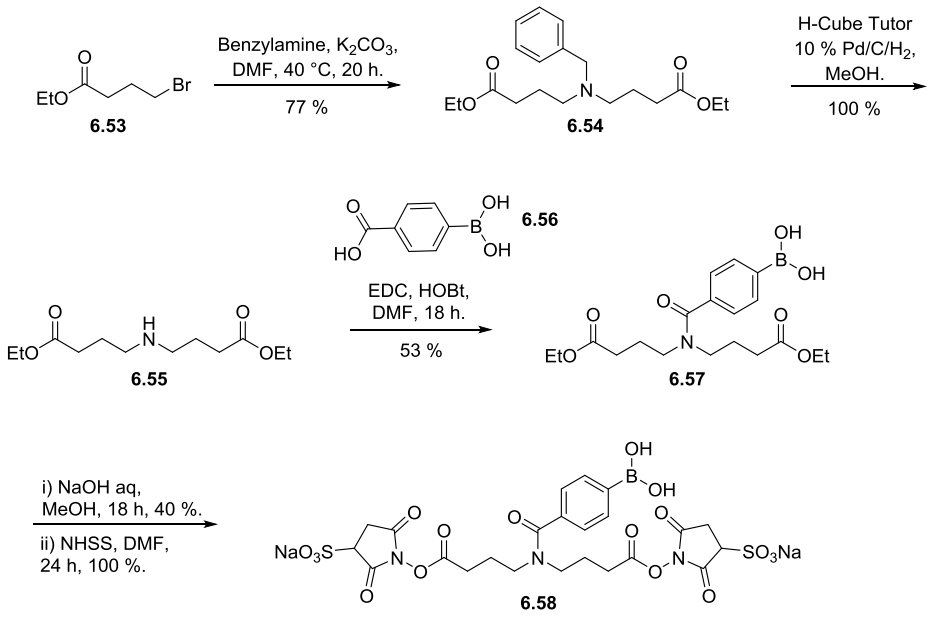
A high degree of degradation of this peptide was observed at high pH due to the catechol group, which was controlled by addition of 5 mM TCEP and working with a peptide concentration of 0.6 mM, with the pH of the solution at 7.6. It was shown that within 10 minutes, using 20 equivalents of CPG, near-quantitative amounts of the tagged peptide were taken up onto the solid support. This material was eluted cleanly within 15 minutes using 1% TFA in water.

This tagged peptide was then spiked into *E.coli* extract to assess whether the peptide could be enriched from a complex mixture. Following mixing, digestion of the sample was carried out with trypsin followed by purification of catechol containing fragments using arylboronic acid CPG. Eluted material was analysed by MS and compared with the crude mixture. In the enriched sample, 35 peptide fragment sequences were identified, whereas 1641 peptide sequences were found in the crude mixture, thus proving that enrichment *via* boronate ester formation can be successfully carried out. Having proven boronate ester complexation suitable for sample enrichment, the group targeted a cross-linking compound with built in enrichment functionality using this system (Scheme 6-13).



Scheme 6-13: Enrichment strategy using a catechol functionalised cross-linker. Following cross-linking, modified peptides are immobilised on arylboronic acid functionalised CPG and non-modified peptides washed away. Elution of only modified peptides simplifies MS analysis.

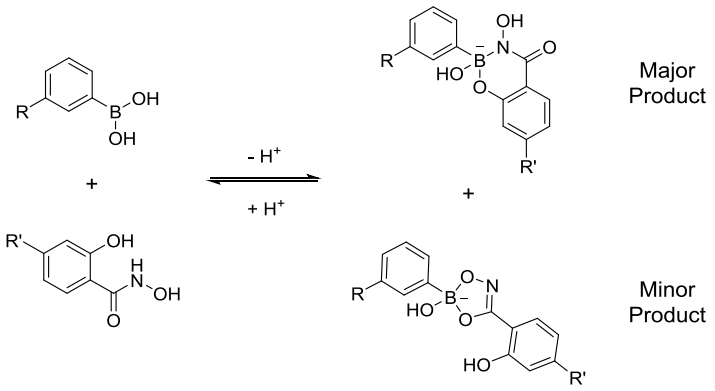
Attempts to synthesise such a linker bearing a catechol enrichment group were impeded by the inherent instability of the group, therefore a new approach was needed. Functionalities were reversed and instead an arylboronic acid was incorporated into the tag and a catechol functionalised solid support was generated. Scheme 6-14 shows the arylboronic acid bearing cross-linker synthesised.



Scheme 6-14: Boronic acid functionalised cross-linker

This compound proved successful in the ability to cross-link proteins, as indicated by gel electrophoresis analysis. Unfortunately, following tryptic digestion of excised gel bands, no detectable modified peptides were observed. It was found that no digestion was taking place due to the fact that the boronic acid group inhibits trypsin.

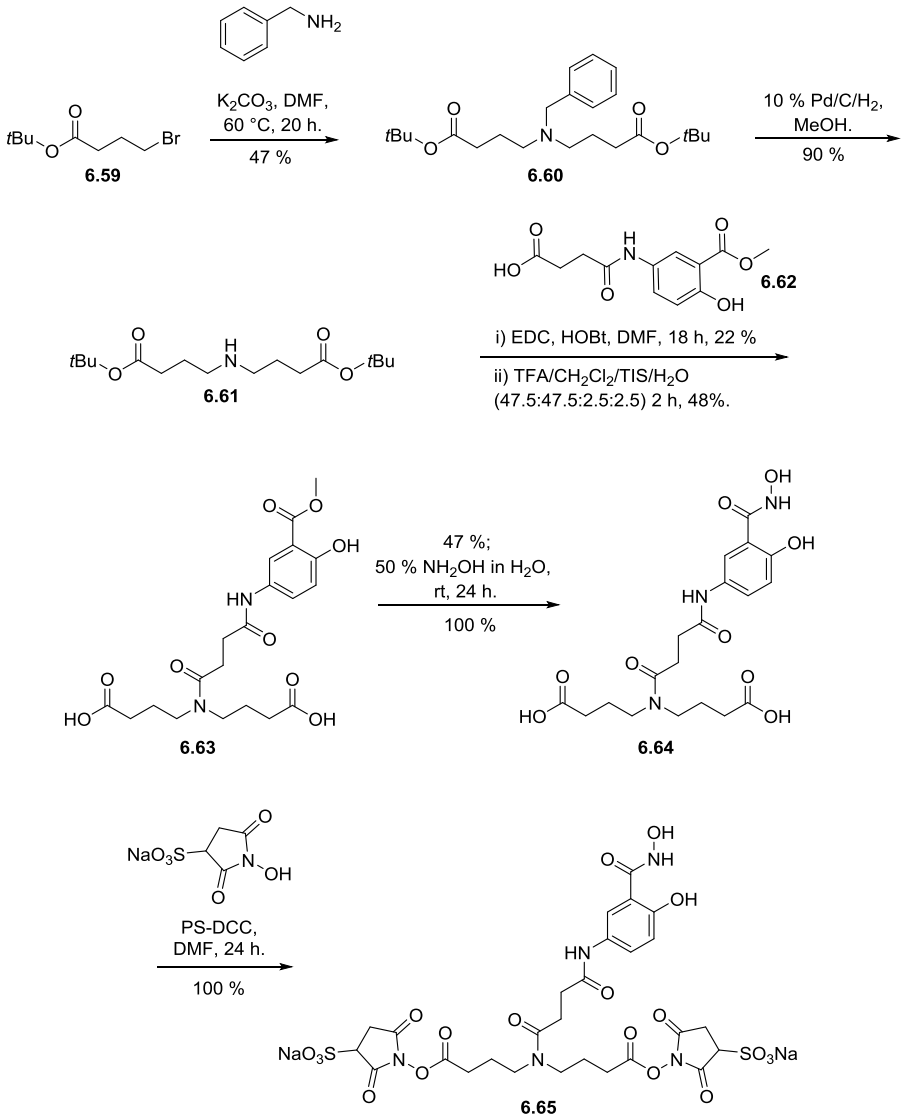
In addition to the above mentioned partners for boronic acids, salicylhydroxamic acid (SHA) has also been shown to form reversible complexes at around pH 8-9, which are typically dissociated by lowering the pH. Stolowitz identified this interaction as a new bioorthogonal reaction which could be used to conjugate two moieties, each labelled with either a phenylboronic acid (PBA) or salicylhydroxamic acid (SHA).[227] This interaction forms two products, predominantly a six-membered ring but with a small amount of five-membered ring also formed, as shown in Scheme 6-15.



Scheme 6-15: Major and minor product complexes formed between phenylboronic acids and salicylhydroxamic acids

Salicylhydroxamic acid derivatives were therefore employed by the Rappsilber group in the development of subsequent enrichable cross-linkers. A SHA modified peptide was synthesised and assessed with arylboronic acid functionalised CPG. It was found that by using 20 equivalents of functionalised CPG, only 4% of peptide remained in solution after 1 minute, and that elution could be achieved using 1% TFA in water for 25 minutes.

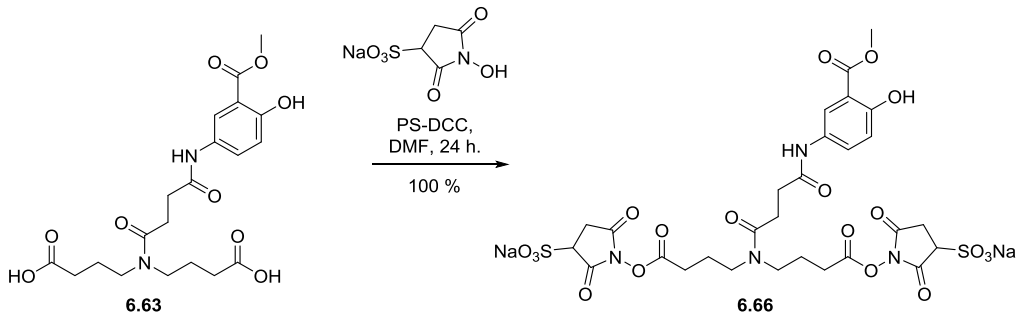
Cross-linker **6.65** (Scheme 6-16) was synthesised, however a certain amount of decomposition of the penultimate hydroxamic acid compound **6.64** was observed upon HPLC purification.



Scheme 6-16: Synthesis of SHsA functionalised cross-linker

The activated cross-linker was applied to a peptide (H_2N-LLEYLKR-OH) under basic conditions. Analysis of the mixture showed the main product to be the internally cross-linked peptide (from K side chain to *N*-terminus), but with the hydroxamic acid having been hydrolysed to the corresponding carboxylic acid. Other concerns were also raised about this cross-linker, such as whether the hydroxamic acid could react with the carbodiimide reagent used to introduce the NHSS groups, and whether the hydroxamic acid undergoes intramolecular nucleophilic attack on the activated esters once they are formed.

Therefore, the intermediate methyl ester **6.63** was activated with NHSS groups and used in cross-linking experiments as an activatable precursor to the hydroxamic acid compound (Scheme 6-17).



Scheme 6-17: Activation of methyl ester 'hydroxamic acid precursor' cross-linker

Compound **6.66** was mixed with beta-lactoglobulin and after 2 hours, cross-linked protein was identified by gel electrophoresis. This was trypsinised and subjected to 50% aqueous hydroxylamine in order to convert the methyl ester group into the desired hydroxamic acid. Digestion and analysis using MS indicated that the conversion was successful and an intermolecular cross-link was identified.

Although these results were encouraging, lack of solubility of the tag was an issue. Therefore, a next generation enrichment tag **6.67** was prepared, which featured a PEG spacer in an attempt to resolve this problem (Figure 6-8).

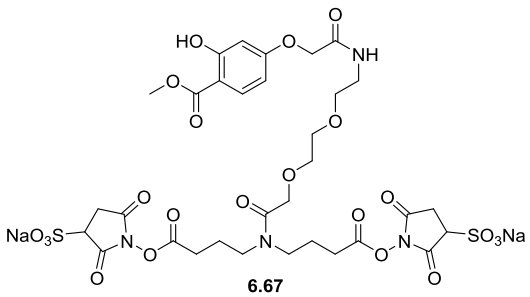


Figure 6-8: Next Generation methyl ester enrichable tag with increased solubility

Beta-lactoglobulin was cross-linked with the hydrophilic cross-linker and again, cross-linked products were observed by gel electrophoresis. Subsequent digestion and MS analysis showed the same cross-link as observed in the earlier experiment, suggesting consistency in the technique. Five additional cross-links were also identified. The mixture was subjected to hydroxylamine solution in order to generate the hydroxamic acid required. The solid supported phenylboronic acid was introduced, then subsequent washing and elution resulted in an enriched sample of cross-linker modified peptides.

These results show the boronate ester mediated enrichment of peptide mixtures to be very promising. However, an observation was made which impairs the use of the methyl ester

precursor tag. Following hydroxylamine treatment, several amino acids such as glutamine and asparagine were also modified, causing them to also form a complex with the boronic acid CPG, contaminating enriched samples Therefore, a new approach whereby treatment with hydroxylamine is not required subsequent to cross-linking would be highly beneficial.

6.3 Aims

In the development of a boronate based enrichable tag, the Rappsilber group encountered a number of issues (See section 6.2.4.3). These included:

- Reactivity of the unprotected hydroxamic acid towards activated esters of cross-linker and coupling agents (compound **6.65**, Figure 6-9)
- Hydrolysis of free hydroxamic acid to carboxylic acid during cross-linking (compound **6.65**)
- Modification of protein sample while carrying out conversion of methyl ester to hydroxamic acid *in situ* using hydroxylamine (compound **6.66**)

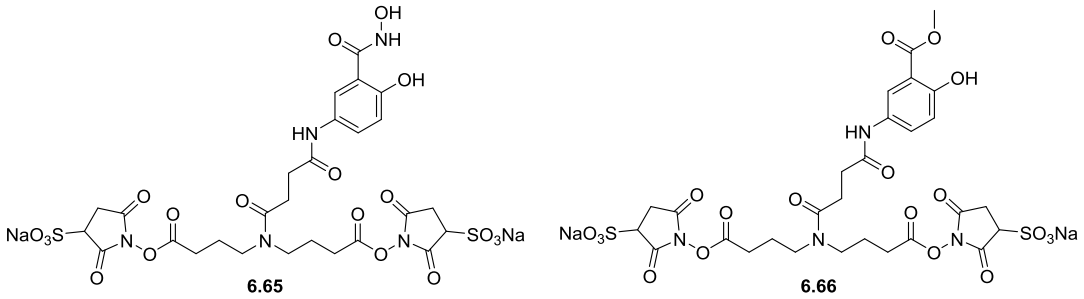


Figure 6-9: Enrichable cross-linkers synthesised by Dr. Adam Belsom, Rappsilber lab.

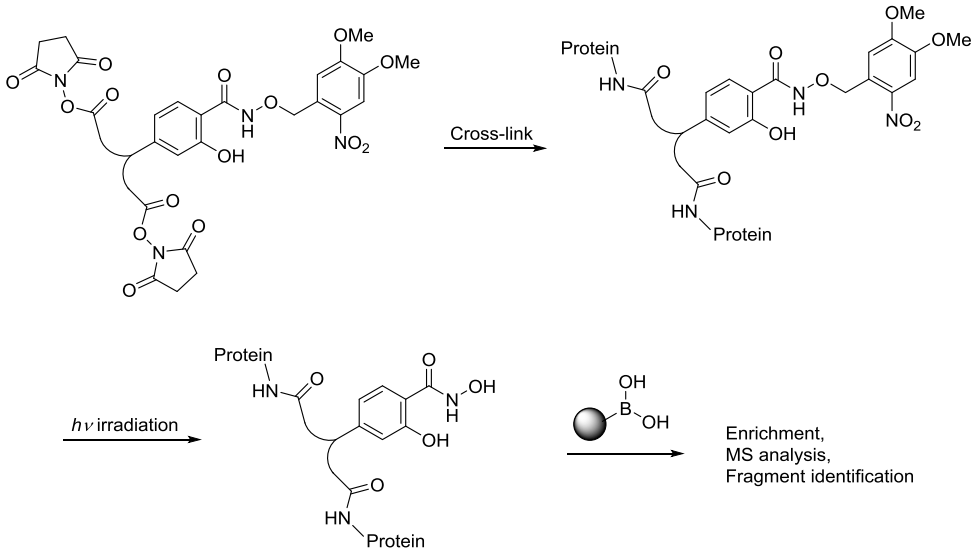
In a collaborative effort, we sought to overcome these problems and generate a useful enrichable tag for protein cross-linking and mass spectrometry analysis. Based on the above observations made by the Rappsilber group, our initial strategy was to generate a tag with a protected hydroxamic acid moiety in place. The chosen group should be labile to mild conditions, so that following cross-linking of the protein, the free hydroxamic acid could be revealed without modification or degradation of the protein. This would allow a controlled way of generating the 'active' salicylhydroxamic acid moiety required for complexation to solid supported boronate to allow immobilisation of cross-linked peptides. Washing away of non-modified peptide fragments would then yield an enriched sample for analysis by mass spectrometry. In order to generate such a cross-linker, several protecting groups were assessed.

6.4 Results and Discussion

6.4.1 Development of a Boronate Ester based Enrichable Tag for Cross-linking

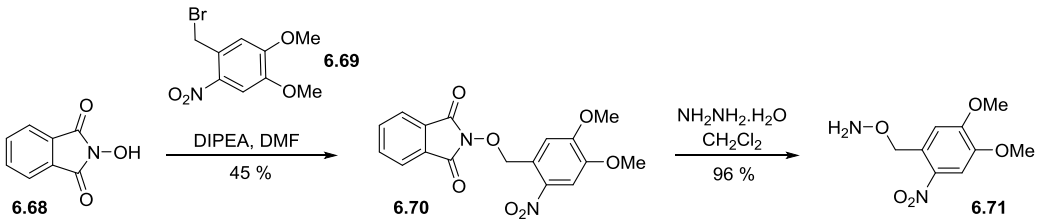
6.4.1.1 Assessment of a Caged Hydroxamic Acid

A caging group is designed to maximally interfere with the activity of a molecule until photolysis, at which point it rapidly dissociates from the compound, revealing the active product. We envisaged using a caged hydroxamic acid in this instance, which would be masked using a photolabile protecting group. 4,5-Dimethoxy-2-nitrobenzyl groups are well known photoreactive protecting groups for alcohols which, upon UV illumination at 355 nm, rapidly dissociate to reveal the underlying hydroxyl functionality (Scheme 6-18).[228, 229] As UV illumination is known to be compatible for use with peptides, we reasoned that this would be a good approach for our enrichable cross-linking tag.



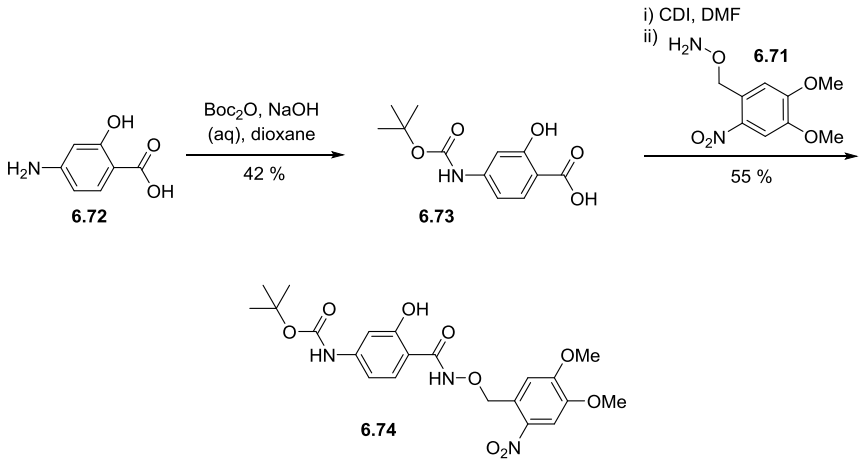
Scheme 6-18: Envisaged mode of action for caged enrichable tag

To achieve this, *O*-4,5-Dimethoxy-2-nitrobenzyl hydroxylamine **6.71** was synthesised by reaction of 4,5-dimethoxy-2-nitrobenzyl bromide **6.69** with *N*-hydroxyphthalimide (Scheme 6-19). The phthalimide group was then removed using hydrazine to give the desired caged hydroxylamine product in excellent yield.



Scheme 6-19: Synthesis of *O*-4,5-dimethoxy-2-nitrobenzyl hydroxylamine

The protected hydroxylamine was then coupled with *N*-boc-4-aminosalicylic acid **6.72**, which was synthesised from 4-aminosalicylic acid (Scheme 6-20).



Scheme 6-20: Coupling of *O*-4,5-dimethoxy-2-nitrobenzyl hydroxylamine with *N*-boc-4-aminosalicylic acid tag backbone

Following Boc deprotection, synthesis of the rest of the tag was planned *via* functionalisation of the aniline group. However, before this synthesis was undertaken, a test deprotection of the caged hydroxamic acid was carried out. Compound **6.73** (1 mg) was dissolved in 1 mL 25% MeCN in H_2O in a clear glass vial and subjected to UV irradiation using a Stratalinker 2400 (365 nm, 15 Watts). Samples were taken at various time intervals and subjected to HPLC and LC-MS analysis (Figure 6-10).

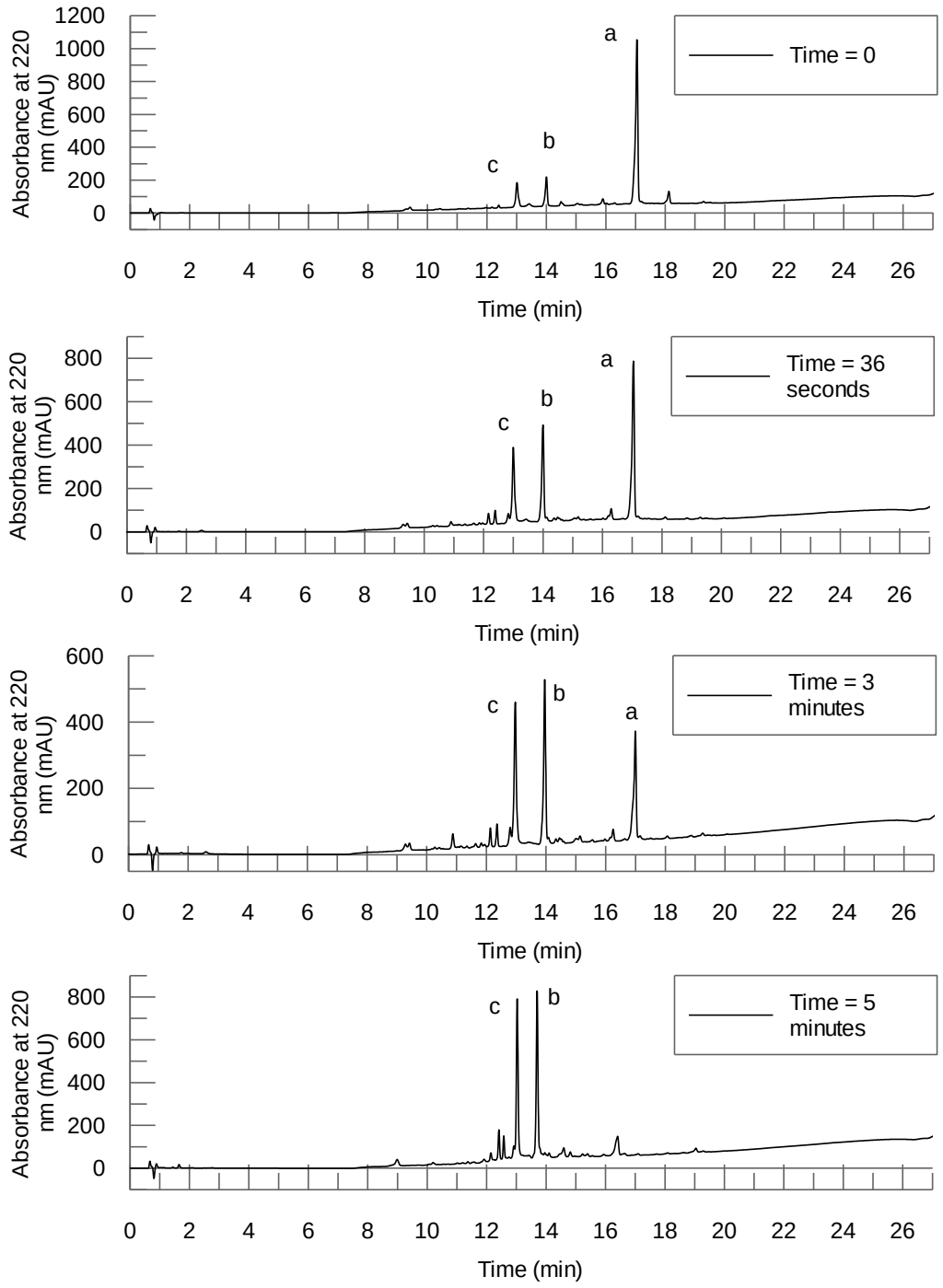


Figure 6-10: HPLC traces of caged hydroxamic acid following UV irradiation. Peak a = starting material 5.74. Peak b = side product 5.76. Peak c = Uncaged hydroxamic acid 5.75.

From the spectra above, it can clearly be seen that the amount of starting material (peak a) decreased over the course of irradiation until, after 5 minutes, there was only trace amounts left. However, the appearance of two product peaks of roughly equal size, suggested that a side reaction was occurring. Upon LC-MS analysis, it became clear that peak c represents the desired uncaged hydroxamic acid **6.75**, whereas peak b represents the corresponding

carbamate **6.76**, indicating that the N-O bond of the hydroxamic acid had cleaved (Table 6-3). Therefore around 50% of the uncaged material is unwanted side product.

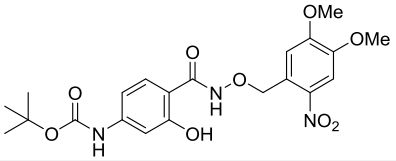
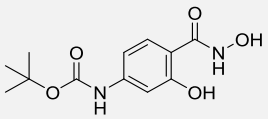
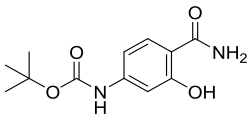
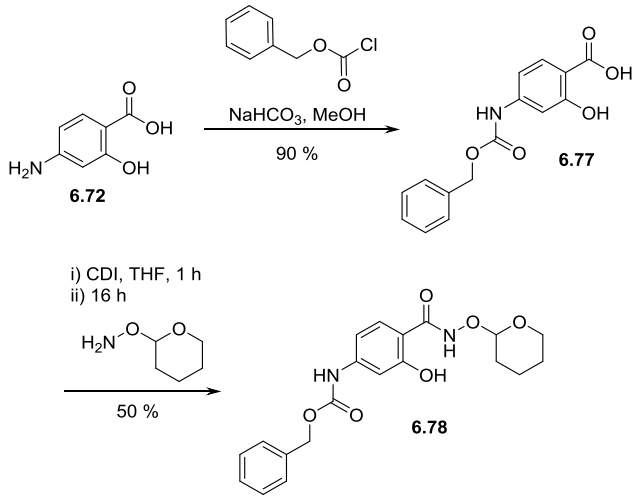
Structure		Peak t_R (min)	Mass Expected $[M+H]^+$	Mass Found
	**6.74**	17	464.17	464.30
	**6.75**	13	269.11	269.12
	**6.76**	14	253.12	253.13

Table 6-3: Species identified from caged hydroxamic acid irradiation

Another observation made was that without irradiation, a small amount of the uncaged material and side product could be seen, meaning that the starting material degraded to some extent before UV irradiation. These results compelled us to conclude that an alternative strategy was required.

6.4.1.2 Assessment of THP Protected Hydroxamic Acid

Substitution of the UV reacting caging group with another protecting group which could be removed under mild conditions was decided upon. Tetrahydropyran has been used previously as a protective group for hydroxamic acids, with deprotection in some cases carried out using amberlyst.[230] This appeared to be an attractive option for the enrichable tag, as following cross-linking, amberlyst could be added to the peptide mixture, then simply filtered off when deprotection had occurred. Synthesis of an enrichable tag which incorporated a THP protected hyroxamic acid was therefore targeted. In this case, 4-aminosalicyclic acid was protected using a benzyl group, as Boc removal would not be compatible with THP (Scheme 6-21). *N*-Carboxybenzyl protection of 4-aminosalicylic acid was carried out in high yield using benzyl chloroformate, then commercially available *O*-(tetrahydro-2*H*-pyran-2-yl)hydroxylamine was coupled to the acid group by first activating it with carbonyldiimidazole.



Scheme 6-21: synthesis of THP protected salicyl hydroxamic acid

Once again, before further synthesis was carried out, deprotection studies to ensure compatibility were carried out. Solutions of THP protected hydroxamic acid **6.78** (1 mg) in 1 mL 25% MeCN in H_2O, then 1 mg amberlyst was added to one vial. Samples were taken at various time points and subjected to LC-MS analysis Figure 6-11**.**

Pleasingly, the chromatograms suggest that amberlyst cleanly removes the THP protecting group without degradation of the compound or the production of side products. Using MS analysis it was found that that expected mass of the desired product matched that of the product seen in the trace, giving confidence that the expected THP deprotection reaction had taken place. In the sample containing no amberlyst, only a very small amount of deprotection was observed following 48 hours in aqueous medium. Satisfied with this data, we proceeded to the synthesis of the rest of the tag.

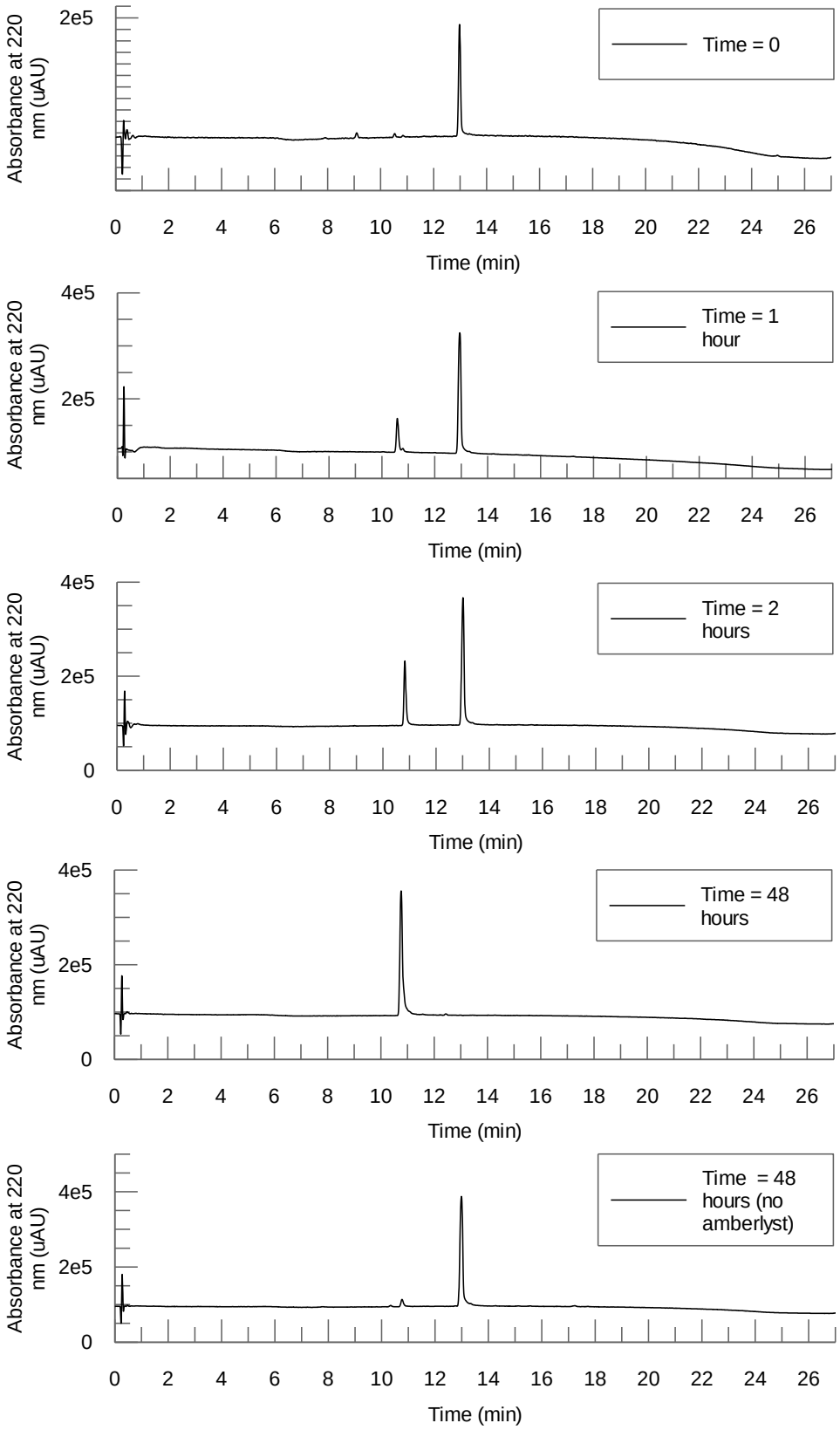
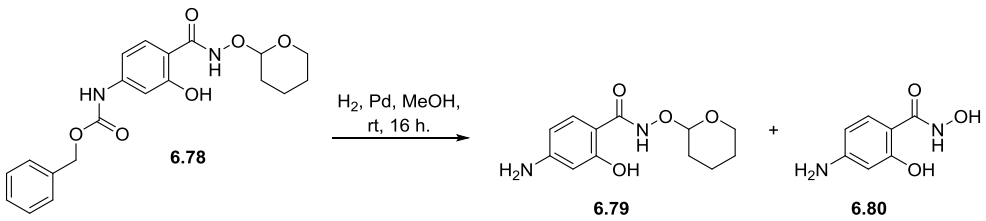


Figure 6-11: LC traces showing deprotection of THP protected salicyl hydroxamic acid over time using amberlyst.

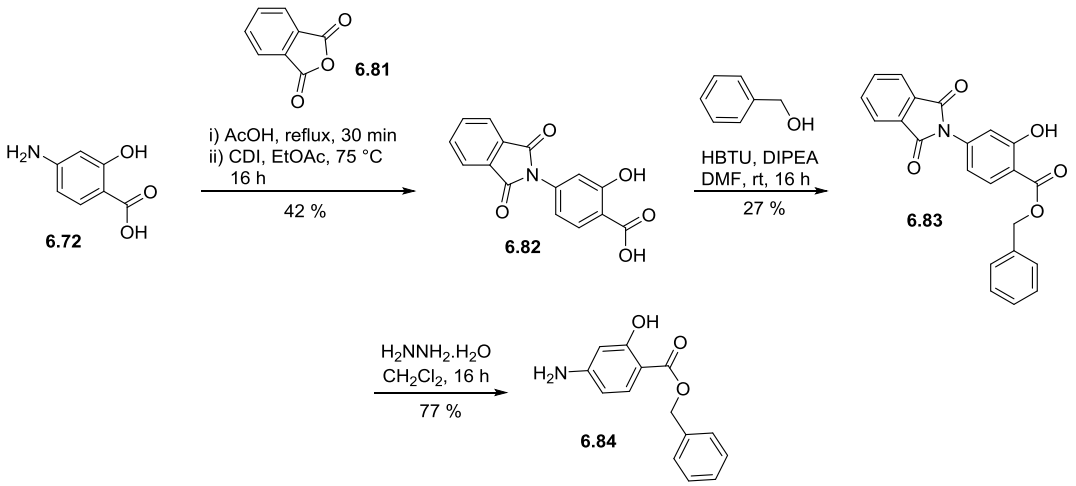
6.4.1.3 Cross-linker Synthesis

Surprisingly, removal of carboxybenzyl group from the aniline by hydrogenolysis resulted in a mixture of products, corresponding to the desired product plus that of the THP group also having being removed (Scheme 6-22).



Scheme 6-22: Hydrogenolysis removed *N*-carboxybenzyl group but also caused unwanted liberation of THP

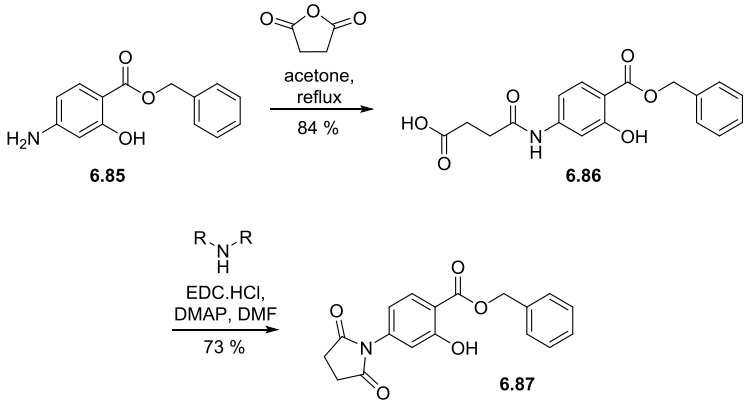
The benzyl protecting group was therefore replaced with a phthalimide, as hydrazine mediated cleavage of this should be compatible with the THP functionality (Scheme 6-23). This intermediate was also used to generate test compound **6.84**, which was used to investigate conditions for aniline functionalisation.



Scheme 6-23: Synthesis of phthalimide protected 4-aminosalicylic acid and test compound benzyl 4-amino-2-hydroxybenzoate

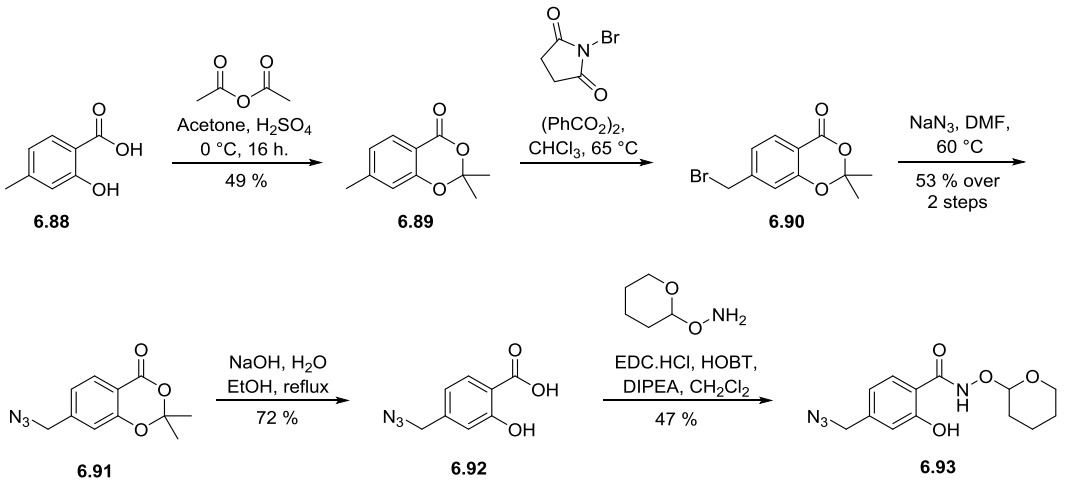
Functionalisation of the aniline group with various carboxylic acids was attempted using a range coupling reagents, without success. Carbodiimide based coupling reagents, HATU and acyl chloride derivatives were all used in an attempt to acylate the aniline, however no product was detected in any of these reactions. Indeed, other groups have reported issues with the acylation of similar anilines due to their low nucleophilicity, and note that only a few very carefully selected reagents and conditions yielded the desired amide product.[231]

Finally, functionalisation of the aniline was achieved in high yield by refluxing in acetone in the presence of succinic anhydride (Scheme 6-24). However, when coupling was attempted on the resulting free carboxylic acid, cyclisation occurred and a stable 5 membered ring was formed rather than the desired product.



Scheme 6-24: Attempts to functionalise acid 6.86 resulted in unwanted cyclised product 6.87

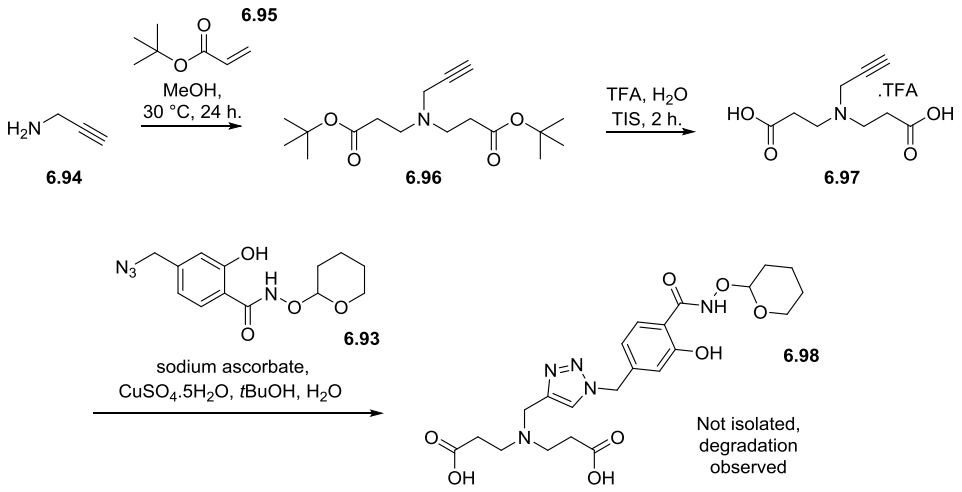
Due to the difficulty in acylating the aniline of 4-aminosalicylic acid, an alternative approach was sought. It was reasoned that incorporating a methylene group between the aromatic ring and the nitrogen would give an amine that was nucleophilic enough to undergo coupling reactions. Therefore, 4-methylsalicylic acid was chosen as an alternative starting material for tag synthesis. In this approach, the acid and phenol functionalities were masked using an acetonide (Scheme 6-25). The methyl was then brominated using NBS and catalytic benzyl peroxide. The bromide was displaced using sodium azide *in situ* with a view that this could later be reduced to give an amine for functionalisation. The acetonide group was then removed to reveal the original phenol and acid group, which was converted into the THP protected hydroxamic acid by EDC.HCl mediated coupling with *O*-THP-hydroxylamine in moderate yield.



Scheme 6-25: Synthesis of azide functionalised THP protected salicylhydroxamic acid

Although reduction of the azide to the corresponding amine had been planned, we saw an opportunity to functionalise the tag without introduction of another peptide bond, which is unfavourable due to fragmentation during MS analysis. Instead, it was decided to use Huisgen 1,3-dipolar cycloaddition to join this compound to an alkyne functionalised section for cross-linking.

With this in mind, we focused on the synthesis of cross-linking portion of the tag. Propargylamine was reacted with excess *tert*-butyl acrylate to give an alkyne functionalised di-*tert*-butyl ester (Scheme 6-26). This compound also contains a tertiary amine, designed to aid solubility and ionisation in the mass spectrometer. Subsequent treatment of the di-ester with TFA yielded the corresponding diacid **6.97**, which was conjugated to azide **6.93** using copper catalysed 1,3-dipolar cycloaddition.



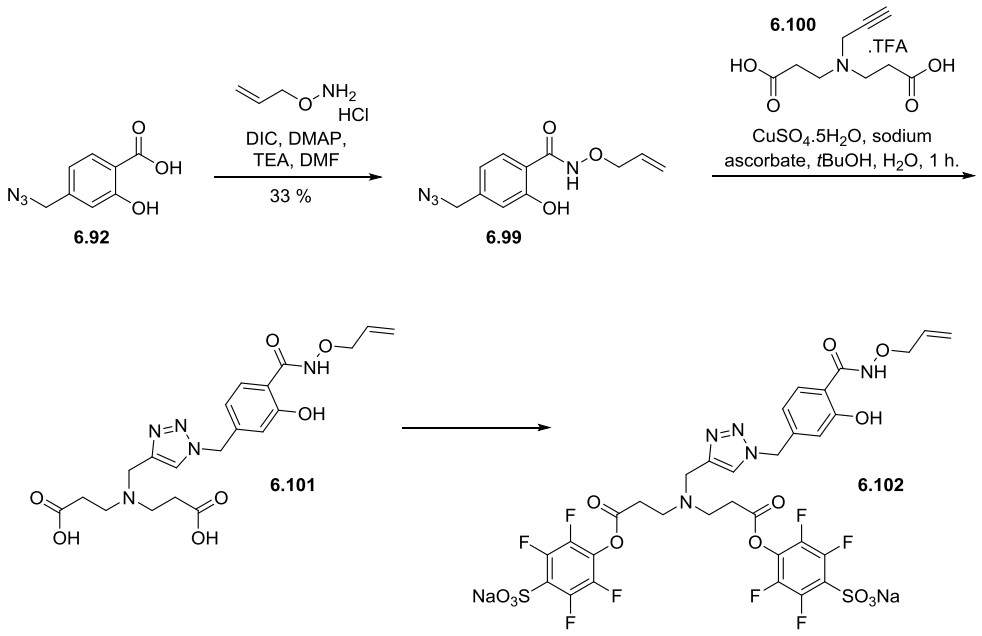
Scheme 6-26: Direct synthesis of free acid enrichable tag which degraded upon purification

This product proved difficult to isolate, as decomposition was observed during silica chromatography. Preparative HPLC also caused degradation, with the main product being the THP deprotected compound. Hence it was decided to exchange the THP protecting group for one more stable to acidic conditions, but that could be removed in a bioorthogonal manner following protein cross-linking.

6.4.1.4 Allyl Protected Hydroxamic Acid

Weiss *et al.* have described the use of palladium in bioorthogonal organometallic (BOOM) reactions.[232] In these reports, palladium resins[233] were utilised for the deptrotection of palladium labile groups under biocompatible conditions *in vitro* and also in a zebrafish embryo *in vivo*.[234]

In order to utilise this method in the synthesis of an enrichment tag, *O*-allylhydroxylamine was introduced onto acid **6.92** to generate the corresponding allyl protected hydroxamic acid **6.99** (Scheme 6-27). This was reacted with alkyne bearing diacid **6.97** in a copper catalysed 1,3-dipolar cycloaddition to give diacid **6.101** which was purified using preparative RP-HPLC. Finally, DCC mediated introduction of 4-sulfotetrafluorophenyl[235] groups was performed to give the final product, enrichable tag **6.102**, which was used directly from the crude mixture.



Scheme 6-27: Synthesis of allyl-protected hydroxamic acid based enrichable tag

Glutathione S-transferase (GST) is a homodimer which, when analysed using SDS-PAGE in denaturing conditions, appears as a monomer. This protein is used to determine cross-linking efficiency of compounds, as cross-links will covalently bind the dimers which can then be visualised on the gel. Compound **6.102** was assessed in this assay to determine whether it was capable of forming cross-links.[‡‡‡‡] Figure 6-12 shows that cross-linking can be detected even at low ratios of **6.102** (0.14:1 tag/protein). Increasing the ratio of tag to protein increases the prevalence of the dimer band as expected, suggesting that cross-linking efficiency is dependent on the amount of cross-linker present. However, the dimer bands appear faint, even with an 18 fold excess of cross-linker. At this ratio, it would be expected that no monomer band be visible.

[‡‡‡‡] Cross-linking and SDS-PAGE analysis performed by Dr. Adam Belsom

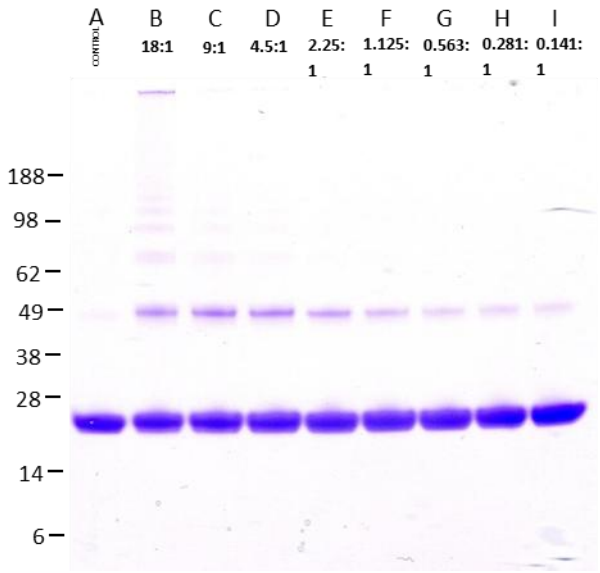
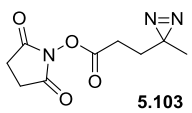


Figure 6-12: Cross-linking of GST using enrichment tag 5.102[§§§§]

Optimisation of the tag activation and cross-linking procedure is required. For example, rather than using the crude reaction mixture, purification and isolation of final cross-linker **6.102** would allow more precise ratios of cross-linker to protein to be established. This has so far proven difficult, due to its high polarity and susceptibility to ester hydrolysis.

6.4.2 Synthesis and Evaluation of a Novel Benzophenone Heterobifunctional Cross-linker

The Rappsilber group use heterobifunctional cross-linker SDA in cross-linking experiments to ascertain structural information about proteins (Figure 6-13). In cross-linking experiments, the succinimidyl ester group reacts with amino groups on proteins, then the sample is irradiated to activate the photoreactive diazirine which cross-links in an non-specific fashion *via* an intermediate carbene. The cross-linked protein is then digested and the fragments analysed by mass spectrometry. Cross-links are assigned by Xi (software developed in-house by the Rappsilber lab) and from this analysis, structural information can be obtained.



5.103

Figure 6-13: SDA cross-linker

[§§§§] Cross-linking and SDS-PAGE analysis performed by Dr. Adam Belsom

Although SDA has proven a useful tool for the Rappsilber group, it does possess some disadvantages. Diazirine synthesis is non-trivial, which leads to high costs if bought commercially. They are also unstable under normal light conditions meaning handling and storage can be troublesome. Therefore, a cheaper and more stable alternative would be advantageous.

4-(*N*-Succinimidylcarboxy)benzophenone (SBP, Figure 6-14) is commercially available, and has been previously incorporated into biological entities to generate photoreactive peptides and nucleotides *via* the benzophenone moiety,[236, 237] but to our knowledge, has never been used as a standalone cross-linker for protein cross-linking experiments. Due to the benefits of benzophenone use in cross-linking discussed *vide supra*, we were curious as to how a benzophenone based bifunctional cross-linker would perform in structure determination experiments compared to its diazirine counterpart.

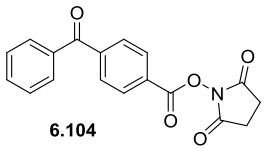


Figure 6-14: 4-(*N*-Succinimidylcarboxy)benzophenone (SBP)

In such experiments, the length of the linker separating the two reactive groups is important, as it determines the distance through space between which residues can be linked. The spacer arm of SDA is 3.9 Å, whereas the spacer arm of the succinimidyl ester benzophenone (SBP) was predicted to 5.7 Å using the minimiser in Omega2 (OpenEye Scientific Software)[238] by generating a low energy 3D conformation (Figure 6-15).***** Although this is slightly longer than SDA, it is acceptable for this type of experiment.

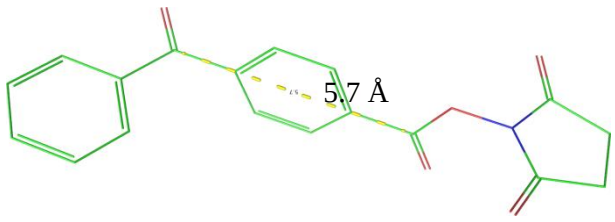
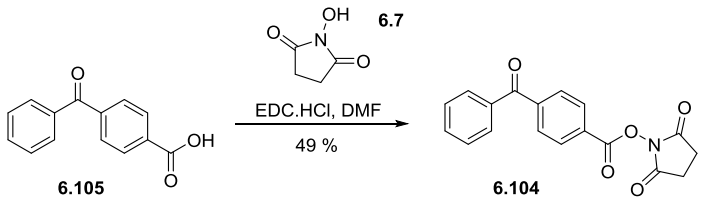


Figure 6-15: Distance estimation between SBP's two reactive centres

***** Distance predicted by Dr. Steven Shave using the minimiser in Omega2 (OpenEye Scientific Software) by generating a low energy 3D conformation.

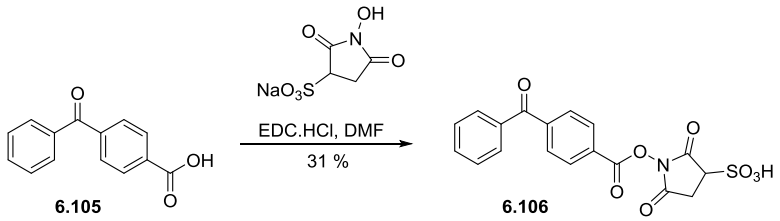
6.4.2.1 Benzophenone Cross-linker Synthesis

The cross-linker was synthesised from 4-benzoylbenzoic acid by reaction with *N*-hydroxysuccinimide in the presence of coupling reagent EDC.HCl to give the corresponding NHS ester (Scheme 6-28).



Scheme 6-28: Synthesis of SBP

However, the very low solubility of this compound in aqueous solutions meant that SBP could not be used in cross-linking experiments. In order to increase solubility, 4-benzoylbenzoic acid was instead reacted with *N*-hydroxysulfosuccinimide using a diimide based coupling reagent to generate the sulfosuccinimidyl ester benzophenone (SSBP) **6.106**, which was purified using preparative RP-HPLC (Scheme 6-29).



Scheme 6-29: Synthesis of sulfosuccinimidyl benzophenone (SSBP)

6.4.2.2 Analysis of Benzophenone Cross-linker

SSBP was found to be soluble in aqueous media and so was used in intramolecular cross-linking experiments with human serum albumin (HSA). All cross-linking experimentation and data analysis reported subsequently was carried out by Dr. Adam Belsom (Rappsilber Lab, University of Edinburgh). HSA is a 66.5 kDa soluble monomeric protein which is produced in the liver and is abundant in blood plasma. Cross-linking was carried out under various conditions (irradiation time, cross-linker to protein ratio) to establish those which showed the most promising results, determined by gel electrophoresis (Figure 6-16).

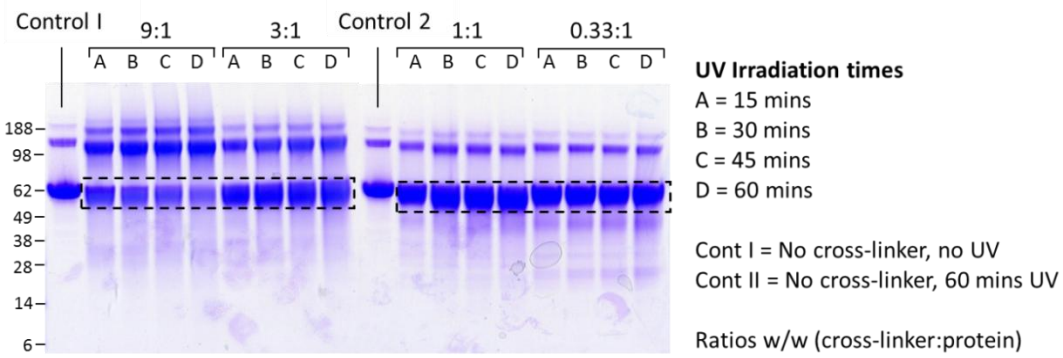


Figure 6-16: Gels showing extent of cross-linking under different conditions using 5.106

In Figure 6-16, control 1 shows HSA with no cross-linker present and no irradiation. The large band (~ 62 kDa) represents HSA monomer, whereas the band above shows a small amount of dimer HSA present in the sample. These high order aggregates are always seen when HSA is analysed. For the cross-linking experiments, various ratios of cross-linker to HSA (0.33:1, 1:1, 3:1 and 9:1 w/w) were mixed and irradiated for different lengths of time (15, 30, 45 and 60 minutes). The results show that as the relative amount of cross-linker is increased, more of the dimer and even higher mass material is seen on the gel, particularly with the 3:1 and 9:1 ratios. This indicates that intermolecular cross-linking has occurred, therefore the ratio is too high. At lower ratios, these multimers are not seen, therefore it can be assumed that cross-linking here is restricted to intramolecular reactions. The bands likely to contain intramolecular cross-links were excised for digestion and MS analysis[239] (See experimental section 8.5.4 for details).

Mass spectral data is processed in Xi software (unpublished) which was developed by the Rappsilber group. This analyses MS data and compares it to a database of possible cross-links for that experiment. The software generates a list of cross-links, ranked by score, which is based on a number of factors with the main one being peptide fragmentation. To introduce confidence limits, the search incorporates decoy sequences not present in the protein being analysed, which produce decoy cross-link matches. The final output therefore consists of true positive cross-links as well as false decoy matches and false target matches. An FDR (false discovery rate) cut-off is applied to the list at the point where the number of decoy matches exceeds the allowed limit for the stated FDR. The decoy matches are removed from the dataset and the remaining list contains true cross-link matches with an error rate equal to the FDR reported.

Analysis of the SDS-PAGE samples from SSBP cross-linking of HSA revealed that 151 unique cross-links had been achieved at 5% FDR, indicating that around 7 of these are likely mis-matches. The identified cross-links are shown pictorially in Figure 6-17. The linkage map shows red lines indicating the cross-link pairs identified in the protein. The chart underneath shows a linear version of this information, with residue number given on the x axis and cross-links made shown by curved lines. Here, green lines represent cross-links of length 0 – 15 Å, orange lines represent cross-links 15 – 25 Å long and purple lines show over length links (> 25 Å).

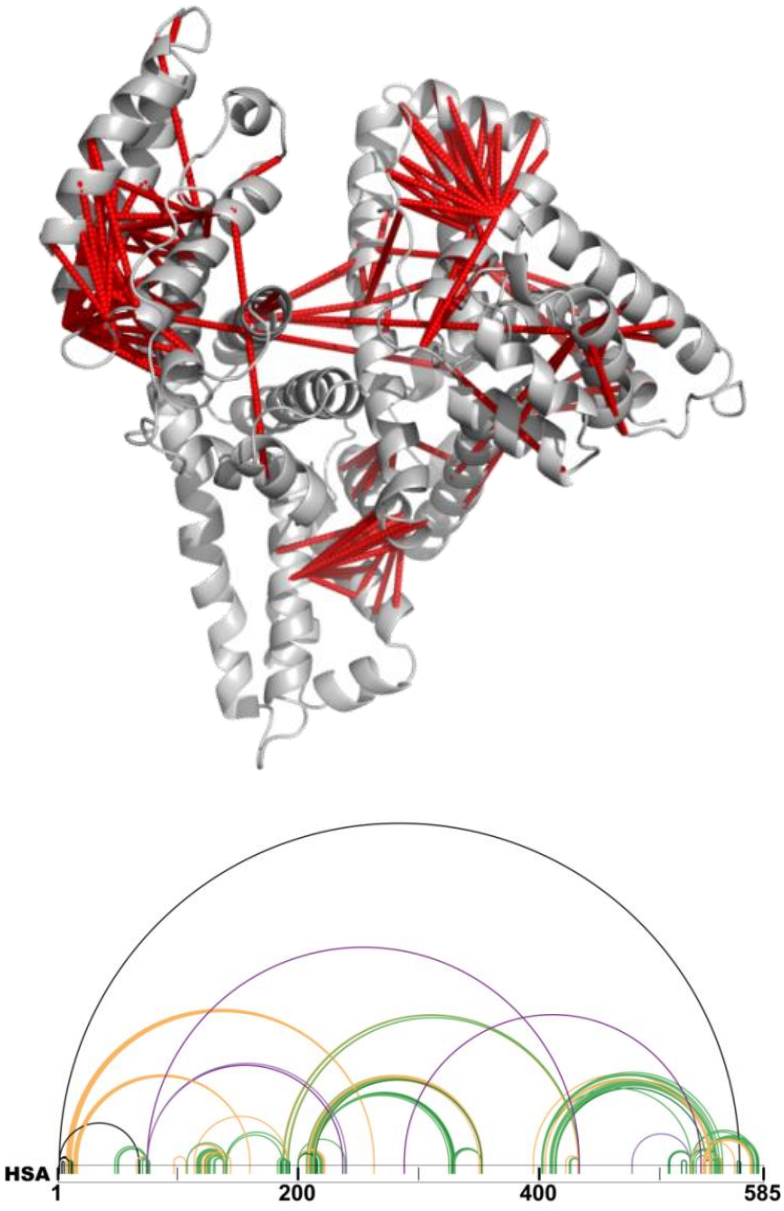


Figure 6-17: Linkage map of observed cross-link pairs on HSA using SSBP based on crystal structure (left) and linear linkage map (right)

Figure 6-18, below, shows a random distribution of alpha-carbon distances between lysine residues and other residues in HSA (red). Also shown is the distribution of lengths of actual cross-links made during the experiment (blue). A prediction of the upper limit of cross-linkable distance between lysine pairs can be calculated[198] from the length of the lysine chain (6.5 Å), the length of the cross-linker spacer arm (5.7 Å, Figure 6-15) and an estimation of the length of the side chain of the residue to which the benzophenone reacted (~ 6.5 Å). Also taken into account is the error in the crystal structure (~ 1 Å). Therefore, it is expected that a majority of the cross-links would be between lysine pairs closer than 19.7 Å. Indeed, the data shows that a high percentage of the cross-links observed sit within this limit. However, a few data points show cross-links to lysines much further apart (up to 47 Å). Although this could be rationalised as capturing possible conformations of HSA in solution that do not exist in the crystal structure, it is more likely that these data are false positives and should be disregarded.

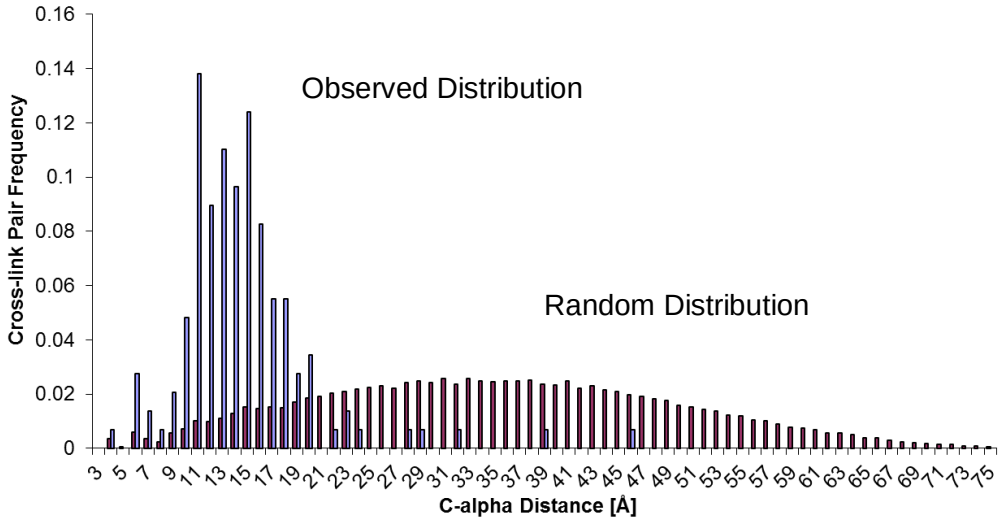


Figure 6-18: Distribution of cross-link distances using sulfo-SBP

If the linkage map of SSBP is compared to that of SDA, for which 1066 cross-links have been identified (Figure 6-19), it can clearly be seen that far fewer cross-links have been made with SSBP. This may be explained by the fact that the cross-linking data shown for SDA represents the combined reults from many cross-linking experiments over a three month period where cross-linking conditions had been optimised, whereas only initial results from trial experiments with SSBP are shown.

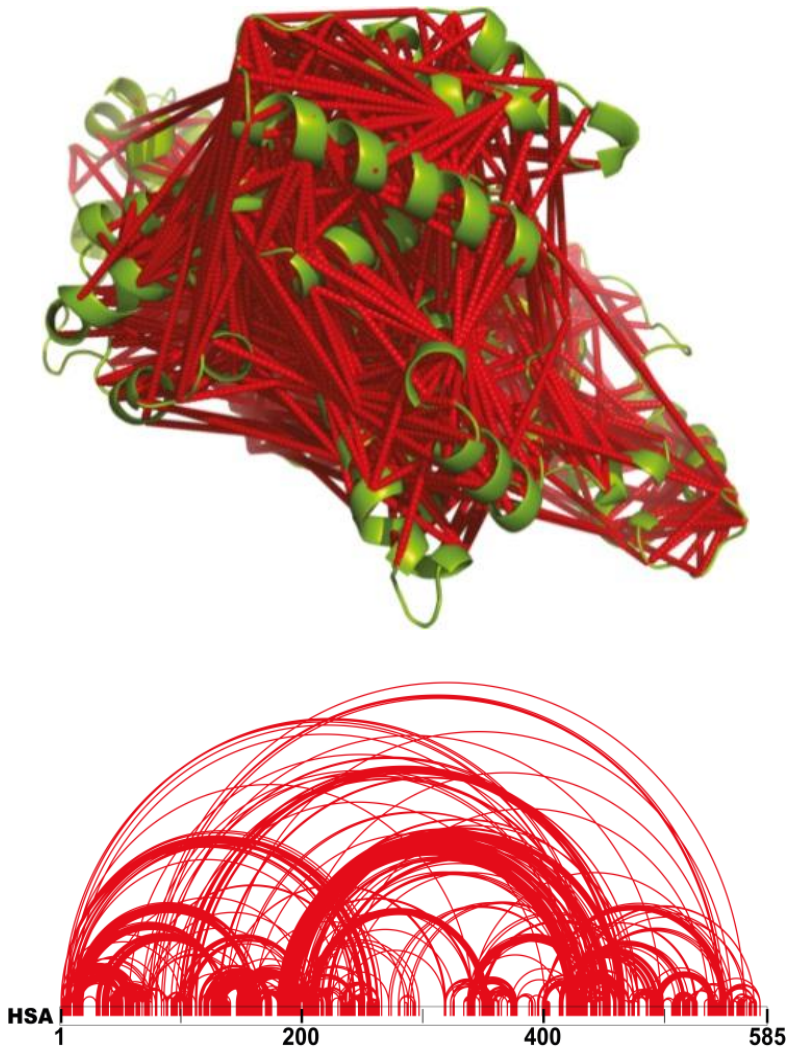


Figure 6-19: Linkage maps of observed cross-link pairs on HSA using SDA

Interestingly, when individual cross-links of HSA made by SDA and SSBP are compared, it becomes evident that despite the vast number of cross-links identified using SDA, the benzophenone reagent is still capable of making novel links from lysine residues (Figure 6-20). Here we see all of the cross-links made by SSBP, with ones already identified by SDA in red and the new, complimentary cross-links shown in blue. From this we can conclude that the benzophenone cross-linker could be used as a complimentary tool to SDA, to generate cross-links in areas that the diazirine does not, in order to generate extra structural information. This data alludes to the fact that different photoactivatable cross-linkers may show preference and specificity to certain regions of proteins. It also demonstrates the fact that cross-linkers themselves are chemicals, which have individual specificities and interactions with proteins, depending on their chemical makeup. This may represent a new consideration to be taken into account when performing cross-linking experiments. Perhaps

for a more complete dataset of cross-links on a protein, a series of chemically different cross-linkers are required to cover every different surface or envirmonment present.

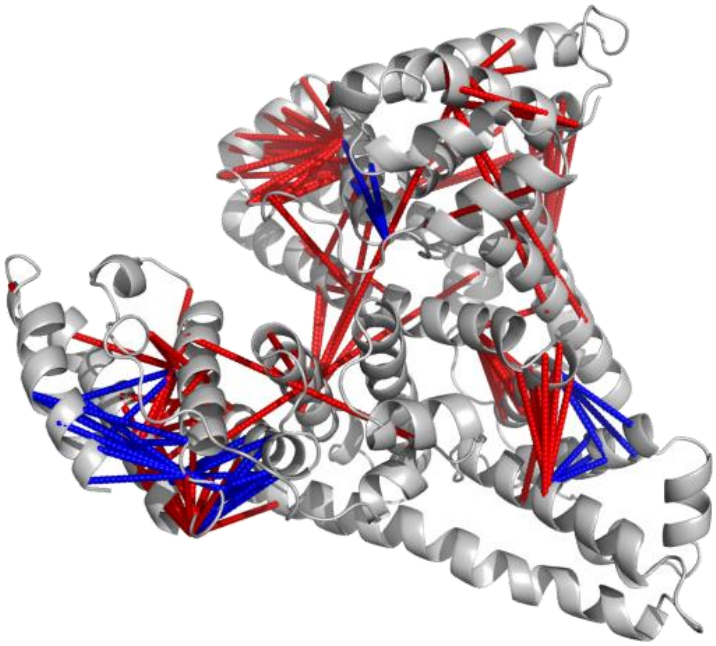


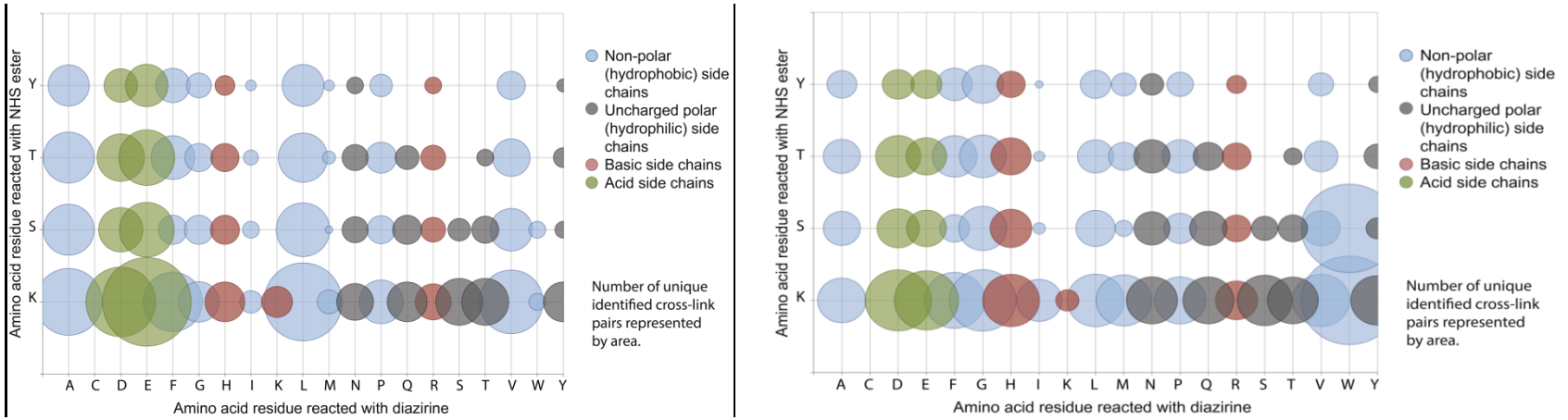
Figure 6-20: New (complimentary) cross-links made by SSBP but not SDA shown in blue. Common cross-links between SDA and SSBP are shown in red.

An analysis of the distribution of cross-links made by SSBP to different amino acids present in HSA was conducted and compared to that of SDA (Figure 6 21). Here, the y axis shows the different amino acids which reacted with the activated ester side of the tags, whereas the x axis shows the various amino acids with which the photoactivatable part of the tag reacted. The two diagrams on the left show non normalised data, which means the absolute number of cross-links made to each amino acid are shown. In the right hand diagrams, the data is normalised, which means the relative number of cross-links to each amino acid is shown by dividing the number of cross-links made to each amino acid by the number of times that particular residue is present.

The non-normalised data would suggest that both cross-linkers have a preference to react with leucine, glutamic acid, alanine and valine residues. However, this may be because these residues are most abundant in the protein (Glu makes up 78 of the 608 residues of HSA, while there are 60 residues of Leu, 62 of Ala and 41 of Val).[240] Tryptophan appears to be involved in very few cross-links, however this can be explained by the fact that there is only a single tryptophan residue present in HSA.

The distribution of normalised HSA cross-links made by the diazirine, which gives a better indication of preference, shows roughly equal numbers of cross-links occurring on each amino acid type (non-polar, polar, basic and acidic), implying that it reacts in a rather non specific way across the whole of the protein. However, when the normalised SSBP cross-links are examined, there appears to be a preference for non-polar, hydrophobic amino acids, particularly methionine and tryptophan residues. This can be explained by the relatively hydrophobic character of the SSBP cross-linker itself, which most probably forming hydrophobic interactions with these residues.

SDA (left = non-normalised, right = normalised)



SSBP (left = non-normalised, right = normalised)

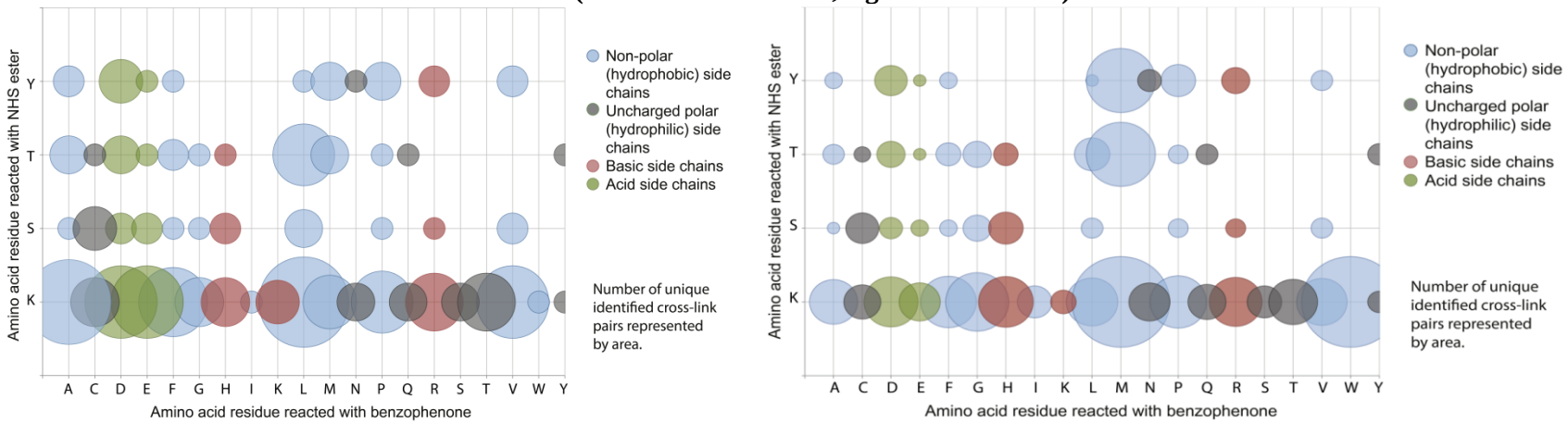


Figure 6-21: Distribution of cross-links to different amino acids using SDA (above) and sulfo-SBP (below)

6.5 Conclusions and Further Work

Development of an enrichable cross-linker based on reversible boronate ester conjugation was undertaken. Previous hydroxamic acid bearing cross-linkers synthesised by Dr. Adam Belsom (Rappsilber lab, University of Edinburgh) underwent degradation when subjected to buffer. In order to prevent this, a caged hydroxamic acid was synthesised and assessed. It was found that upon irradiation, an impurity formed in equal amounts to the desired product. Protecting group THP was incorporated and assessed which showed clean deprotection to give the desired product, however degradation of this group meant that isolation of the final compound was not achieved. This group was therefore replaced with an allyl protecting group, which allowed synthesis of the tag to be completed. Future work will focus on the bioorthogonal organometallic (BOOM) removal of this protecting group to enable enrichment of cross-linked peptides.

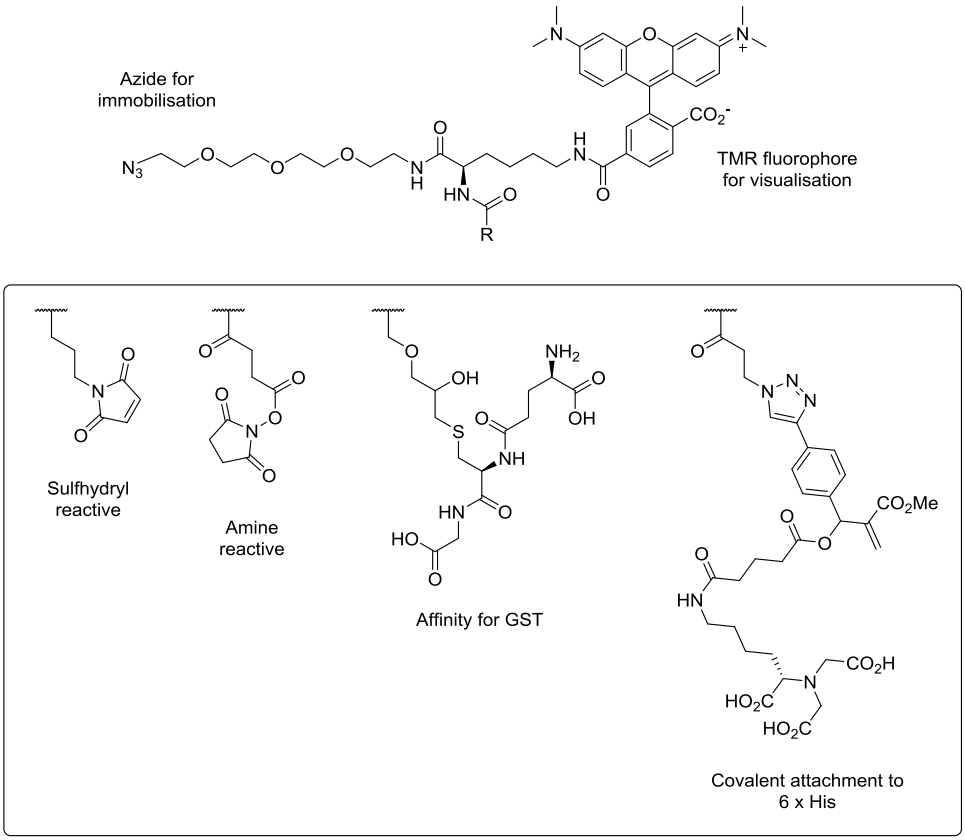
A novel benzophenone based cross-linker was synthesised which appears to have the potential to elucidate cross-linking data about HSA which SDA does not. Despite the high degree of cross-linking of HSA achieved by SDA, identification of novel cross-links was still achieved using our new benzophenone reagent. Further optimisation experiments are required in order to establish whether the number of cross-links made using this reagent can be improved. What can be concluded, however, is that the SSBP allows cross-links complimentary to the SDA, which gives additional structural information for each protein examined. In this respect, SSBP is a promosing new tool for chemical cross-linking. Rather than being used as a replacement for SDA as first intended, the reults of experiments carried out using HSA suggest that it should instead be used alongside the diazirine based reagent to extract maximum structural information. The cross-linking and analysis of additional proteins should be carried out in order to further demonstrate this theory. An interesting future investigation could involve the analysis of the protein surface where the SSBP cross-linker suceeded in generating novel cross-links to establish whether these areas are particularly hydrophobic compared to those where only the SDA made cross-links. Thios would show that a hydrophobic compound such as SSBP would be more likely to generate cross-links in areas where it can form hydrophobic interactions with the protein. If this were true, then it may be possible to specifically target areas of proteins which show few cross-links simply by chosing a cross-linking reagent which shows affinity for the properties of the protein surface in that area (e.g. SSBP for hydrophobic areas). This could potentially allow more complete coverage of a protein in cross-links.

CHAPTER 7

The CONA PS/PS platform developed in the Auer lab has been fully established and optimised for the screening of soluble proteins against peptides and small molecules. Here, a However, this elegant approach to high-throughput screening - where OBOC libraries are used for an initial on-bead screen, with hit compounds taken directly into single molecule solution phase binding confirmation experiments using only the material from a single bead - is still not realising its potential. Work has been initiated towards the inclusion of membrane proteins into the list of screenable targets. Expression of GPCRs in lipoparticles or nanodicscs represent encouraging approaches for on-bead to solution phase screening of compound libraries at the single bead level. Receptors in either of these formats could be implemented in on-bead screening, and also in solution phase fluorescence based assays for binding confirmation using only 50 picomoles of substance.

As a result of the development of the new ALBA linker we now have a new tool for the manipulation of amine bearing compounds on solid support. Future applications of this could include the tethering of other amine containing compounds such as nucleotides, or proteinogenic reagents *via* lysine side chains to a solid support. The anchoring of lysine side chains would allow a novel method for the synthesis of cyclic peptides, with the *N*- and *C*- termini of the amino acid free to partake in the cyclisation reaction. The linker may also find broader application, with the potential to immobilise other functional groups such as alcohols (as an ester linkage) and thiols (as a thio-ester). This would allow the ALBA linker to be used, for example, in carbohydrate chemistry.

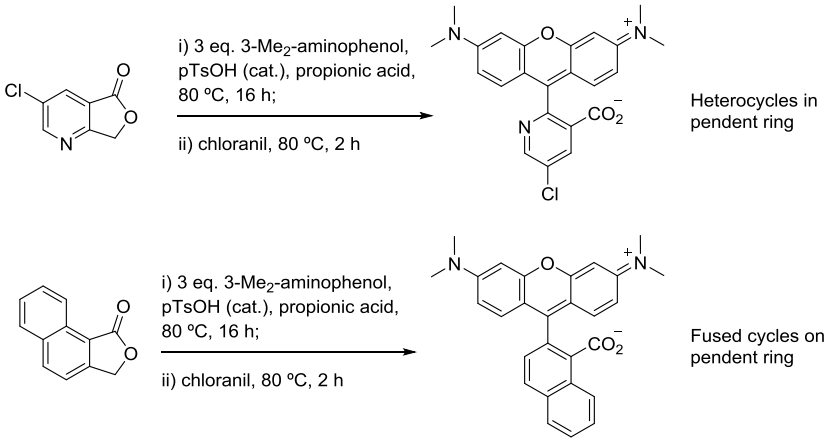
The development of TOBa functionalised resin represents an opportunity to take the CONA and PS/PS platform beyond classical on bead screening. This solid supported triply orthogonally protected backbone will allow rapid and simple regioselective synthesis of almost limitless combinations of multifunctional reagents for novel screening and chemical biology purposes. Exemplified in this thesis is the application of TOBa towards the simultaneous fluorescent labelling and immobilisation of a sulfhydryl containing protein in order to probe protein-protein interactions. An expansion of this proof-of-concept tag format will be carried out to allow labelling and immobilisation of proteins through reaction of lysine amino groups (using an activated ester), His-tagged proteins (utilising an NTA bearing compound with the possibility of covalent linkage)[241] and glutathione S-transferase (GST) tagged proteins (through use of glutathione bearing tags) among others. This will allow proteins containing almost any affinity tag to be incorporated into the protein-protein interaction assays.



Beyond this, novel screening methods which take advantage of the developments made during this thesis in the area of chemical cross-linking and mass spectrometry can be foreseen. For example, the trifunctional reagent synthesised which bears an azide, a benzophenone and biotin could be employed in a cross-linking and pull-down based screening approach for the identification of targets for libraries. OBOC libraries incorporating Pra would be labelled with the azide functionality using click chemistry. Following this, libraries would be incubated with cell lysate then irradiated with UV light. Ligands would then be covalently bound to targets *via* the benzophonone cross-linker. Biotin could be used to purify from the cell lysate and MS analysis used for the identification of bound proteins. The biotin moiety could also be replaced with the salicylhydroxamic acid used for enrichment in Chapter 5.

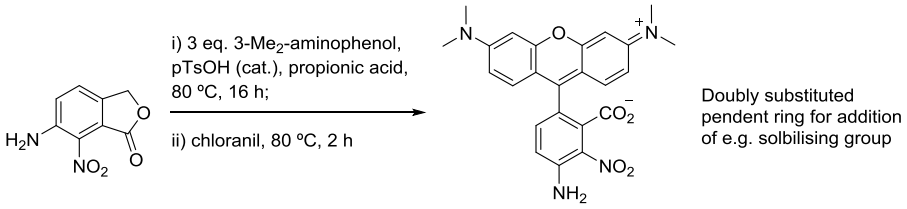
For the production of labelled libraries and tags, fluorescent material must be easily available. Single isomer rhodamine derivatives are relatively unattainable in large quantities due to high commercial cost and the difficulty faced in the separation of structurally similar isomers formed during synthesis. The novel synthetic route reported herein allows simple and direct generation of a range of isomerically pure functionalised rhodamine dyes. A further advantage of this method is the use of stable phthalides as starting materials, which

allows installation of various linkage functionalities before rhodamine construction. We envisage that this new approach will allow access to novel isomerically pure functionalised rhodamine compounds with altered spectroscopic and physical properties. The few suggestions below aim to highlight the possibilities available using our novel synthetic method. The modulation of spectroscopic properties such as absorption wavelength may be achieved, for example, by the use of heterocyclic phthalides or phthalides with fused ring systems.



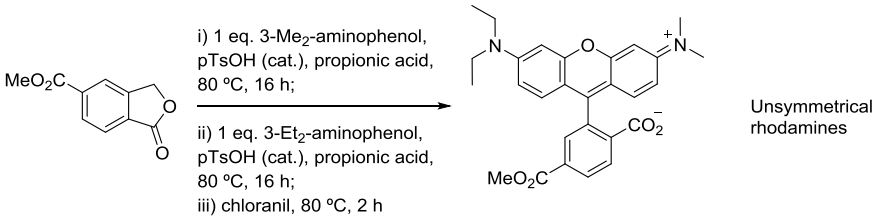
Scheme 7-1: Heterocyclic phthalides or phthalides with fused ring systems could be used for the generation of novel rhodamine analogues

Doubly functionalised rhodamines may be synthesised using this method, whereby phthalides with two pre-installed functional groups could be utilised to introduce e.g. a linkage point and a solubilising agent such as a PEG chain.



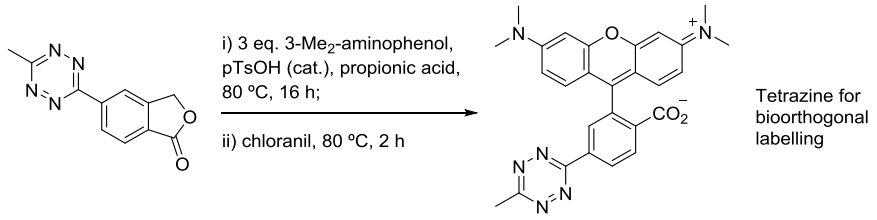
Scheme 7-2: Doubly substituted phthalides could allow rhodamines with a linkage point and a site to incorporate other functionalities, such as solubilising agents

Unsymmetrical rhodamines have been previously unattainable, however using this method we envisage that two different amino phenols could be utilised to generate this unusual compound class.



Scheme 7-3: Unsymmetrical rhodamines may be accessible using our novel method

Finally, the introduction of bioorthogonal reactive groups into dyes can generate highly useful imaging agent, capable of site specific labelling of compounds *in vivo / vitro* which bear the reactive partner. A tetrazine functionalised phthalide would allow construction of a tetrazine containing rhodamine dye, capable of selectively labelling strained alkenes.[242]



Scheme 7-4: Rhodamines with groups installed for use in bioorthogonal reactions will be targeted in the future

In summary, the dyes, linkers, tags and libraries presented in this thesis have been utilised for specific applications, however their full potential and usage has not yet been realised given the time frame of development within one PhD thesis. They may therefore be regarded as new, useful tools for application towards future systems chemical biology investigations.

EXPERIMENTAL

8.1 General Information

All commercially available reagents were purchased from Sigma Aldrich or VWR unless otherwise stated and used as received. Thin layer chromatography (TLC) was performed on Merck DF-Alufoilien 60F254 0.2 mm precoated plates. Product spots were visualised by UV light at 254 nm, and subsequently developed using vanillin, iodine, or ninhydrin solution as appropriate. Flash column chromatography was carried out using silica gel (Fisher Scientific 60 Å particle size 35-70 micron) or using Biotage Zip cartridges (5 – 80 g). ^{1}H NMR spectra were recorded on a Bruker AVA500 (500 MHz), Bruker Pro500 (500 MHz), Bruker AVA600 (600 MHz) or a Bruker AVA400 (400 MHz) spectrometer. Chemical shifts (δ) are quoted in parts per million (ppm) downfield of tetramethylsilane, using residual protonated solvent as internal standard ($CDCl_3$ at 7.27 ppm). Abbreviations used in the description of resonances are: s (singlet), d (doublet), t (triplet), q, (quartet), app (apparent), br (broad). Coupling constants (*J*) are quoted to the nearest 0.1 Hz. Proton-decoupled ^{13}C NMR spectra were recorded on a Bruker AV500 (125.8 MHz) spectrometer or a Bruker AVA400 (100.6 MHz) spectrometer. Chemical shifts (δ) are quoted in parts per million (ppm) downfield of tetramethylsilane, using deuterated solvent as internal standard ($CDCl_3$ at 77.0 ppm). Assignments were made using the DEPT sequence with secondary pulses at 90° and 135°.

LC-MS

Samples were analysed on an HPLC-ESI-MS system consisting of a Finnigan Deka XP Plus ESI-MS detector connected to an Agilent 1100 capillary LC with a G1376A capillary pump, a G1377A Micro-autosampler and a G1325B DAD detector with micro flow cell. The system was run with a spray voltage of 5 kV, capillary temperature of 275 °C, a capillary voltage of 15 V in single MS mode. The LC-system contained a Zorbax-SB-C18 0.5x35mm 3.5 μm particle size column and was run with a gradient of solvent A: H_2O containing 0.1% TFA and B: MeCN (LCMS grade) containing 0.1% TFA. Standard Gradient: 2.5 min 0% B, 22.5 min 0 - 100% B, 25 min 100 - 0% B, 28 min 0% B at a flow rate of 200 μl/min. HPLC-ESI-MS data were processed using the Xcalibur software package (version 2.0, Thermo Electron Corporation, MA, USA).

RP-HPLC

HPLC analysis was performed on an Agilent 1100 series HPLC system, consisting of a quaternary pump (G1311A), a degasser (G1322A), an FLD detector (G1321A) and a DAD

detector; column: Vydac peptide C8, 4.6 mm x 150 mm, 5 μm particle diameter size. Analyses were performed using a linear gradient of A: H_2O containing 0.1% TFA and B: MeCN (HPLC grade) containing 0.1% TFA with a flow rate of 0.8 mL/min, Retention times (t_R) are denoted in minutes. All analyses use standard gradient unless stated otherwise. Standard gradient: 5 min 0% B, 25 min 0 – 100% B, 27.5 min 100 – 0% B, 30 min 0% B.

Semi-Preparative RP-HPLC

Purification of mg quantities of peptides and dyes was carried out using a preparative HPLC system (Agilent 1100 prep-HPLC system), equipped with a preparative autosampler (G2260A), preparative scale pumps (G1361A), a fraction collector (G1364B-prep) and a multiwavelength UV detector (G13658 MWD with preparative flow cell). Material was separated at a flow rate of 20 mL/min on an Agilent RP-C-18 column (21.2 x 150 mm, 10 μm particle size), using a H_2O/MeCN gradient and a detection wavelength of 210 nm. Solvent A: H_2O, 0.1% TFA; solvent B: MeCN, 0.1% TFA.

Method A: 15% B for 5 minutes, the increase to 55% B over 20 minutes, then 100% B over 10 minutes.

Method B: 5% to 40% B in 20 minutes, 40% B to 95% B in 5 minutes.

Method C: 0% B for 5 minutes, 0% to 80% B in 40 minutes, 80% B for 5 minutes, 80% to 0% B in 5 minutes, 0% B for 5 minutes.

8.2 Chapter 2 Experimental

8.2.1 Standard Solid Phase Synthesis Methods

Amino Acid Coupling

To a solution of amino acid (5 eq.) and HATU (4.9 eq.) in DMF (1 mL/100 mg resin) is added DIPEA (10 eq.). The solution is transferred into an SPE fitted with a frit containing resin. The SPE is capped and shaken for 55 minutes then the solution drained and the resin washed extensively with DMF and CH_2Cl_2.

Fmoc Deprotection

A solution of 20% piperidine in DMF is added to resin in an SPE fitted with a frit. The mixture is shaken for 30 minutes then the solution drained and the resin washed extensively with DMF then CH_2Cl_2.

TNBS Test

A few representative beads are transferred by pipette to a vial and a few drops of 10% DIPEA/DMF solution are added. One drop of 5% (w/v) in H_2O picrylsulfonic acid solution is added and the colour of the beads determined under a microscope. Orange beads indicate free primary amines presented on the resin. Clear beads indicate no primary amines.

***N*-Terminus Acetylation**

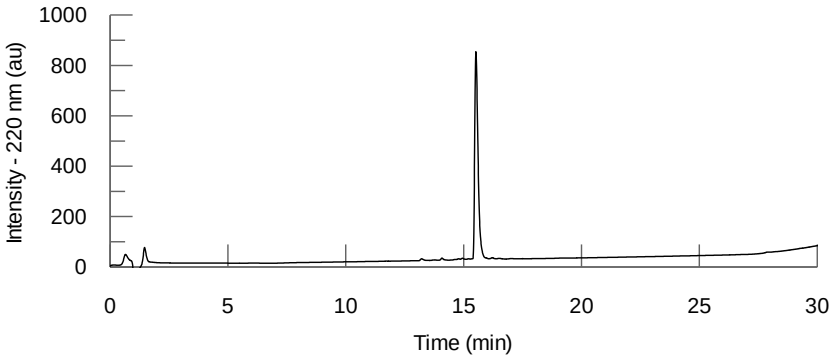
Resin is subjected to a solution of acetic anhydride (10 eq.) and DIPEA (20 eq.) in DMF for 1 hour. The resin is then washed extensively with DMF and CH_2Cl_2.

Standard Cleavage from Rink Resin

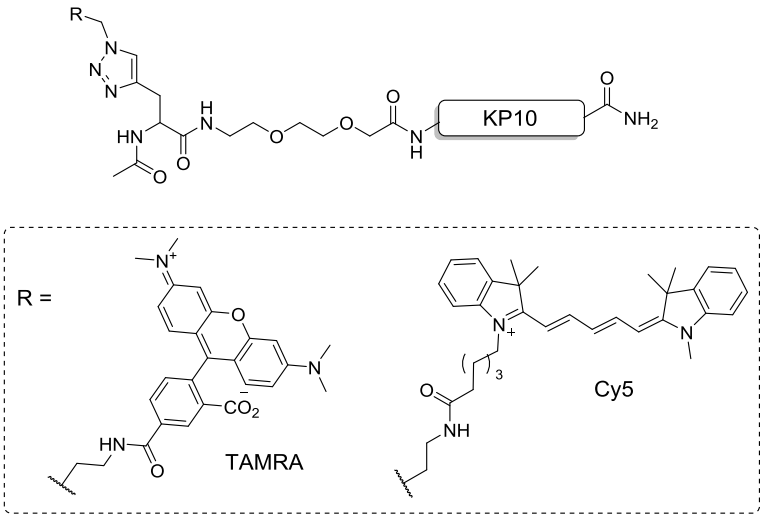
Resin is subjected to a solution of 2.5% H_2O and 2.5% TIS in TFA for 2 hours. Following this, the solution is drained into a round bottomed flask and the resin washed with CH_2Cl_2 three times. Volatiles are removed under reduced pressure and the residue subjected to diethyl ether to cause precipitation of peptides. The solid is filtered off and dried.

8.2.2 KP-10

KP-10 (**2.3**) was synthesised on TentaGel S RAM resin (Rapp Polymere) using standard procedures. MS (ESI) Exact mass calcd for $C_{63}H_{84}N_{17}O_{14}$ $[M+H]^+$: 1302.64, found: 1303.02; Crude purity (RP-HPLC): 92%. The peptide was purified using semi preparative RP-HPLC method A. >95% purity.



8.2.3 Fluorescently Labelled KP-10



Ac-Pra-DOA-KP-10-NH_2 (2.30)

Ac-Pra-DOA-KP-10-NH_2 was synthesised on TentaGel S RAM resin (Rapp Polymere) using standard procedures. MS (ESI) Exact mass calcd for $C_{76}H_{102}N_{19}O_{19}$ $[M+H]^+$: 1584.76, found: 1585.07. Crude purity (RP-HPLC): 74% (220 nm).

Ac-Pra(TMR)-DOA-KP-10-NH_2 (KP-10-TMR, 2.31)

Ac-KP-10-DOA-Pra resin **2.30** (100 mg, 0.023 mmol) was dried and to this was added a solution of TMR-N_3 (11.8 mg, 0.023 mmol) in MeOH (0.5 mL). A solution of $CuSO_4.5H_2O$ (16.8 mg, 0.069 mmol), Sodium ascorbate (13.7 mg, 0.069 mmol) and TBTA (37.0 mg, 0.069 mmol) was made up in 0.5 mL 1:1 $^tBuOH/H_2O$ and this was added to the resin also. The mixture was covered with foil and shaken for 48 hours. The resin was washed thoroughly with CH_2Cl_2, DMF, MeOH then dried. Cleavage from the resin was carried out using standard procedures to yield a purple solid. MS (ESI) Exact mass calcd for $C_{104}H_{130}KN_{25}O_{23}$ $[M+H+K]^{2+}$: 1067.99, found: 1067.51; The peptide was purified using semi preparative RP-HPLC method A. Purity >95% by HPLC (220 nm).

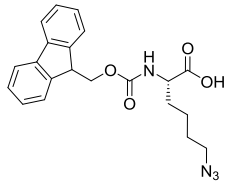
Ac-Pra(Cy5)-DOA-KP-10-NH_2 (KP-10-Cy5, 2.32)

Ac-KP-10-DOA-Pra resin **2.30** (50 mg, 0.012 mmol) was dried and to this was added a solution of Cy5-N_3 (8.9 mg, 0.012 mmol) in MeOH (0.5 mL). A solution of $CuSO_4.5H_2O$ (8.5 mg, 0.035 mmol), Sodium ascorbate (6.8 mg, 0.035 mmol) and TBTA (18.0 mg, 0.035 mmol) was made up in 0.5 mL 1:1 $^tBuOH/H_2O$ and this was added to the resin also. The mixture was covered with foil and shaken for 48 hours. The resin was washed thoroughly with CH_2Cl_2, DMF, MeOH then dried. Cleavage from the resin was carried out using standard procedures to yield a blue solid. MS (ESI) Exact mass calcd for $C_{112}H_{148}KN_{25}O_{26}S_2$ $[M+H+K]^{2+}$: 1181.01, found: 1180.98; The peptide was purified using semi preparative RP-HPLC method A. Purity >95% by HPLC (220 nm).

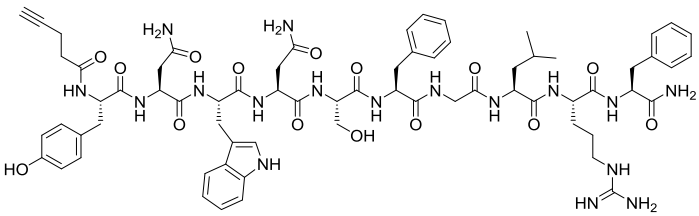
Ac-Pra(Cy5)-DOA-KP-10-NH_2 Biological Testing

Protocol is detailed in section 8.3.7

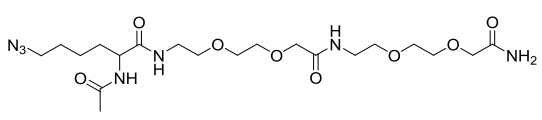
8.2.4 Reversed KP-10 Resin



Fmoc-Lys(N_3)-OH (2.40): This compound was prepared as described in the literature.[243]



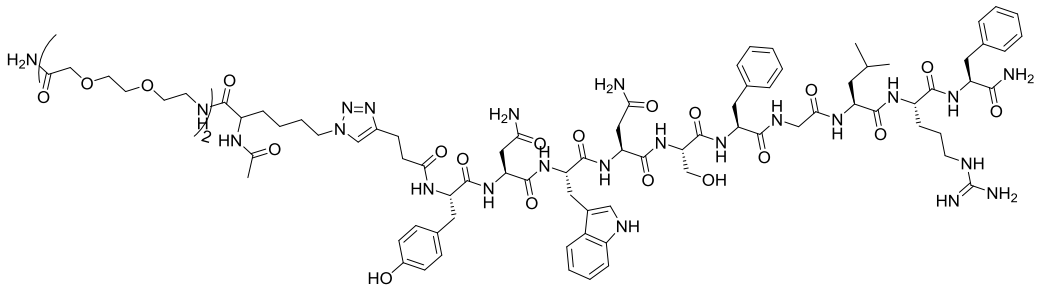
4-Pentynoic acid tagged KP-10 (2.38): Kisspeptin-10 was synthesised on 750 mg (0.26 mmol/g loading, 0.195 mmol) TentaGel rink amide resin using standard conditions. Following this, 4-pentynoic acid was coupled onto the *N*-terminus and the compound cleaved from the resin using standard conditions. The alkyne tagged KP-10 was isolated as a white solid (233 mg, 0.17 mmol, 86%). MS (ESI) Exact mass calcd for $C_{68}H_{88}N_{17}O_{15}$ $[M+H]^+$: 1382.66, found: 1383.13.



Ac-DOA-DOA-Lys(N_3) (2.41): Fmoc-8-amino-3,6-dioxaoctanoic acid (DOA) was loaded onto 100 mg (0.26 mmol/g loading, 0.026 mmol) TentaGel rink amide resin using standard conditions. Another DOA unit was added, then Fmoc-Lys(N_3)-OH was coupled and the *N*-terminus deprotected and acetylated.

A small amount of resin was subjected to standard cleavage conditions in order to determine the purity of the compound. MS (ESI) Exact mass calcd for $C_{20}H_{38}N_7O_8$ $[M+H]^+$: 504.28, found: 504.11.

DOA-DOA-Reversed KP-10 (2.42):



Ac-DOA-DOA-Lys(N_3) Resin **2.41** was swelled in CH_2Cl_2 then drained. 4-Pentynoic acid tagged KP-10 **2.38** (10 eq.) was dissolved in DMF and added to the rein. A solution of $CuSO_4.5H_2O$ (10 eq.), and sodium ascorbate (10 eq.) was added to the resin also and the mixture shaken for 16 hours. The resin was washed with DMF and CH_2Cl_2 then dried. Compound was cleaved from the resin under standard conditions. LC-MS analysis shows around 50 % loading of KP-10. MS (ESI) Exact mass calcd for $C_{88}H_{125}N_{24}O_{23}$ $[M+H]^+$: 1885.93, found: 1886.88.

8.2.5 Bacterial culture and handling

8.2.5.1 Bacterial growth media (prepared by Joanna Koszela)

Liquid lysogeny broth (LB): 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl were dissolved in 950 mL of dH_2O water. The pH was adjusted to 7.0 using 1 M NaOH, the volume was brought up to 1 L, and the medium was sterilised by autoclaving at 120 °C.

Agar (solid) LB: Prepared as liquid LB, but 15 g/L agar was added prior to autoclaving. After cooling down the media to around 55 °C, appropriate amounts of antibiotics were added and the solution was poured onto sterile 10 cm petri dishes and left to solidify at ambient temperature for 2-3 days.

8.2.5.2 Bacterial transformation (performed with the aid of Joanna Koszela):

50 µL of chemically competent E. coli DH5α in Eppendorf tubes were thawed on ice. 5 µL of plasmid DNA was added, mixed gently without pipetting and incubated for 30 minutes on ice. Then, the bacteria were heat-shocked at 42 °C for 45 seconds and cooled down on ice for 2 minutes. 250 µL of liquid LB was added and the solution was incubated on a shaker for

1 hour at 37 °C. Finally, 50 µL of bacteria were spread with sterile glass beads onto prewarmed agar LB plates with appropriate antibiotic selection and grown overnight at 37 °C.

8.2.6 DNA amplification and isolation

Selected resistant colonies were picked up and grown overnight at 37 °C in 5 mL of liquid LB with appropriate antibiotics. 1 mL of the overnight culture was placed into Eppendorf tubes, span down for 1 minute at 13,000 rpm and the resulting pellets were subjected to plasmid DNA extraction with the QIAprep Spin Miniprep Kit (Qiagen, Cat. n° 27106) according to manufacturer's indications. The DNA purity and concentration was measured using NanoDrop (ThermoScientific).

8.2.7 General Cell Culture Methods

The following cell lines were used in this work: HEK293, CHO.

8.2.7.1 Cell Culture

Cells were cultured at 37 °C in humidified atmosphere supplemented with 5% CO_2, in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, Life Technologies, cat. n° 11960-044) supplemented with 10% defined Fetal Bovine Serum (HyClone, cat. n° SH30070.03), 2 mM *L*-Glutamine (Gibco, Life Technologies, cat. n° 25030) and 1% Penicillin-Streptavidin (Gibco, Life Technologies, cat. n° 18140). Cells were split 1:10 every 2-3 days.

8.2.7.2 Thawing cells

Frozen cryovials with cells in 20% DMSO were thawed by dipping into 37 °C water bath, sprayed with 70% ethanol to sterilise the tubes and the cells were immediately pipetted into 10 mL of pre-warmed medium in T25 flasks (Nunclon). The flasks were kept overnight in incubators at 37 °C and 5% CO_2. The medium was exchanged the following morning.

8.2.7.3 Freezing Cells

80 – 90% confluent cells in T25 flasks were washed twice with PBS and detached using 2 mL 0.05% trypsin-EDTA (Gibco, Life Technologies, cat. n° 25300-054) for 5-10 minutes at 37 °C. Cells were collected in 8 mL of DMEM without antibiotics and centrifuged for 5 min at 300 x *g* at 4 °C in 50 mL Falcon tubes. Cells were carefully resuspended in 2 mL per flask of freezing medium (DMEM + 10% DMSO), distributed into 2 cryovials (1 mL per vial) and immediately placed on dry ice. Frozen cells were kept at -80 °C before being

transferred into liquid nitrogen for longer storage.

8.2.7.4 Cell Transfection

A day before transfection, cells were plated in 10 cm dishes to reach 70-90% confluence at the end point of the experiment, in DMEM without antibiotics. On the transfection day, 16 µg DNA was mixed with 800 µL OptiMEM transfection medium (Gibco, Life Technologies, cat. n° 31985). 90 µL Lipofectamine (Life Technologies, cat. n° 18324) was diluted in 710 µL OptiMEM, mixed with the DNA solution and incubated for 20 minutes at rt. In the meantime, the medium on the cells was exchanged to 2/3 of required volume of fresh DMEM with no serum and no antibiotics. After incubation, the DNA-Lipofectamine mix was further diluted in OptiMEM and pipetted onto the cells in drop wise manner, while shaking the plate with cells to allow immediate mixing. Transfected cells were incubated for 4 hours to overnight before the medium was changed to DMEM with serum and antibiotics if required.

8.2.7.5 Stable Cell Line Generation

Antibiotic Concentration Determinations

Cells (0.5×10^5) were plated in 0.5 ml complete growth medium per well in a 24-well tissue culture plate and incubated overnight at 37 °C, 5% CO_2. Increasing amounts (0, 400, 600, 800, 1000, 1200, 1400, 16000, 1800 and 2000 µg/ml) of G418 was added to wells of cells and media plus antibiotic was replaced every 2 – 3 days and the cells inspected for toxicity. The lowest concentration of antibiotic to kill all cells in around a week was selected for stable cell line generation.

Harvesting Cells Stably Expressing Gene of Interest

Cell transfection was carried out as per section 8.2.7.4. The day following transfection, medium was changed and the cells incubated overnight. 48 Hours post transfection, antibiotics were added to the medium (amount determined from antibiotic concentration determination experiment) and the cells incubated at 37 °C, 5% CO_2. Media plus antibiotics were replaced every 2 – 3 days for around 10 days, during which time cells which did not undergo plasmid integration died and were washed away. Surviving cell colonies were picked using a pipette tip and transferred into a 24 well cell culture plate and grown up to confluency, at which point they were transferred into T-25 flasks. Cell lines were analysed and samples frozen down.

FLAG-GPCR54 Construct

This construct was a gift from Prof. Bob Millar. It was prepared by amplification of hGPCR54 by PCR to add 5' EcoRI and 3' BamHI sites to allow subcloning into pcDNA3.1 mammalian expression vector containing an *N*-terminal FLAG epitope tag.[244]

eGFP-GPCR54 Construct

This construct was a gift from Prof. Bob Millar. It was prepared by amplification of hGPCR54 by PCR to add 5' BamHI and 3' XhoI sites using PCR to allow subcloning pEGFP-N3 containing a C-terminal GFP epitope tag.[244]

8.2.8 Protein Extraction for Western Blotting

Mammalian cells were grown on sterile 10 cm petri dishes and collected by washing twice with ice-cold PBS (Gibco, Life Technologies), then collected in Eppendorf tubes, pelleted and lysed for 30 minutes on ice in lysis buffer (50 mM Tris-HCl, pH 7.4, 100 mM NaCl, 2% Triton (Pierce), 10% glycerol, benzonase HC (Novagen), Protease inhibitor cocktail (Roche), 1 mM DTT (Melford)). The lysates were cleared by centrifugation at 4 °C at 12,000 rpm for 10 minutes.

8.2.9 Western Blotting (Carried out with Joanna Koszela)

Protein concentration was measured using Bradford solution (BioRad). Samples in SDS loading buffer were boiled at 95 °C for 5 minutes and run on self-made 12% gel in running buffer (recipes below) or 4-12% NuPAGE SDS-PAGE Bis-Tris precast mini-gels (Novex, Life Technologies cat. n° NP0322BOX) in NuPAGE SDS MOPS running buffer (cat. n° NP0001). Proteins were transferred to a 0.2 μm nitrocellulose membrane with a semi-dry transfer system (Trans Blot SD, BioRad). The membrane was briefly washed in Trisbuffered saline-Tween (TBS-T) and blocked with 5% milk in TBS-T for 1 hour, then incubated with primary antibody in 5% milk in TBS-T overnight at 4 °C. The following day, the membrane was washed 3 times for 5 minutes in TBS-T and incubated with a secondary antibody at 1:10,000 dilution in 5% milk in TBS-T for 1 hour at rt, then washed 3 times for 15 minutes in TBS-T and finally revealed using ECL reagent (Pierce, ThermoScientific) and developed using automated film processor (SRX-101A, Konica).

10x Running buffer (1 L):

30.28 g Tris (BioRad, cat. n° 161-0719), 144.14 g glycine (BioRad, cat. n° 161-0718), 10 g SDS

Transfer buffer (2 L):

5.98 g Glycine, 11.78 g Tris, 3.7 mL 20% SDS, 400 mL methanol

TBS-T:

50 mM Tris pH 7.5, 150 mM NaCl, 0.02% Tween 20 (BioRad, cat. n° 161-6531)

12% resolving gel (quantities for 2):

40% Acrylamide solution (BioRad, cat. n° 161-0140) 2.93 mL, 2% bisacrylamide solution (BioRad, cat. n° 161-0142) 1.6 mL, 20% SDS solution 50 µL, 1 M Tris, pH 8.8 3.75 mL, H_2O 1.5 mL, 10% ammonium persulfate (APS) 100 µL, tetramethylethylenediamine (TEMED) 5 µL (BioRad, cat. n° 161-0801)

5% stacking gel:

40% Acrylamide solution 0.6 mL, 2% Bisacrylamide solution 0.375 mL, 20% SDS solution 25 µL, 1 M Tris, pH 6.8 0.65 mL, H_2O 3.4 mL, 10% APS 16.67 µL, TEMED 6.67 µL

8.2.10 Immunofluorescence (Carried out with Joanna Koszela)

Cells were plated to allow 70-90% confluence at the time of fixation and allowed to adhere before the treatment. After the treatment and required incubation time, cells were washed twice with pre-warmed PBS and fixed with 4% paraformaldehyde in PBS for 20 minutes at rt. After two washes with PBS at rt, cells were permeabilised with 1 x permeabilisation solution (0.2% Triton® X-100 in PBS) and washed twice. Blocking solution (10% Normal Goat Serum in PBS) was applied for 1 hour at rt. The cells were then incubated in primary antibody diluted 1:500-1000 in blocking solution overnight at 4 °C. The next day, cells were washed 4 times with blocking solution and incubated with fluorophore-labelled secondary antibody diluted 1:1000 in PBS for 1 hour at rt protected from light. Finally, the cells were washed 3 times with PBS, counterstained with DAPI if required (20 min incubation in 1 µg/mL DAPI in PBS), washed again and kept in PBS for imaging on PerkinElmer's Opera® High Content Screening System. Cells were visualised at 20 x magnification with air lenses.

8.2.11 Calcium Mobilisation Assay Using Fluo-4 Direct

1.0×10^3 Cells were plated in a glass bottomed cell culture well plate and incubated at 37 °C, 5% CO_2 overnight. 1 mL of Fluo-4 Direct™ calcium assay buffer was added to probenecid to generate a 250 mM stock solution. This was used to make up a 2 X Fluo-4 Direct™ calcium reagent loading solution with a final probenecid concentration of 5 mM. An equal volume of 2 X Fluo-4 Direct™ calcium reagent loading solution to cell culture medium was added to each well and the plate incubated at 37 °C for 1 hour. Ligand dissolved in Calcium free medium was added to the cells, which were imaged using PerkinElmer's Opera® High Content Screening System at 488 nm excitation using 20 x magnification with air lenses.

8.2.12 FACS Analysis

Cells were passaged on to a 10 cm dish and incubated in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, Life Technologies, cat. n° 11960-044) supplemented with 10% defined Fetal Bovine Serum (HyClone, cat. n° SH30070.03). The next day, KP-10-Cy5 was diluted in medium to a final concentration of 20 µM and added to cells and the mixture was incubated for four 2 hours. Cells were gently pipetted off the plate and centrifuged at 1000 rpm for five minutes to pellet. FACS buffer (PBS + 2% FBS) was used to wash cells, and they were centrifuged as above. Re-suspension of cells in 1ml FACS buffer was carried out and they were transferred to FACS tubes, which was wrapped in foil and kept on ice until flow cytometry analysis. Cells were analysed on Becton Dickenson FACS Analyser equipped with 15 mW 488 nm argon laser. Fluorescence data was collected using photomultipliers with bandpass filters: 530 nm, and 661 nm.

8.2.13 Bead Based Assays

8.2.13.1 Whole Cells

This assay was carried out as described by Udugamasooriya *et al.*[42] with some modifications. Around 2000 beads (1 mg) were equilibrated in 3% BSA containing DMEM for 1 hour. Using an enzyme free cell dissociation buffer, GFP-GPCR54 expressing cells were dissociated from plates, washed and resuspended in DMEM medium. 20 x 10^3 cells were added to the beads and the mixture incubated at 37 °C with gentle shaking for an hour. The beads were gently washed with PBS then imaged on PerkinElmer's Opera® High Content Screening System using 488 nm laser.

8.2.13.2 Membrane Fractions

The membrane stain (CellVue Maroon Kit - Tebu Bio) was prepared by diluting down to suggested concentration (2 µM) with PBS. Membranes (10 µL, Perkin Elmer Metastin Receptor GPCR54 human membrane preparation, in CHO-K1 cells) were diluted with 15 µL PBS then 25 µL dye solution added and incubated for 5 minutes. The mixture was spun down at 20,000 rpm for 30 minutes (a blue pellet was visible). The supernatant was aspirated off and the pellet resuspended in 50 µL 1 x BSA in PBS, and this was spun down again for 30 min and the wash step repeated once more. The pellet was resuspended in 60 µL 1 x BSA in PBS and 15 µL added to a well of a 96 well plate which contained beads that had been pre incubated with 50 µL 1 x BSA in PBS. The plate was incubated at 4 °C overnight with

shaking then imaged on PerkinElmer's Opera® High Content Screening System using 640 nm laser.

8.2.13.3 Lipoparticle Generation using MembranePro™ Kit (Life Technologies)

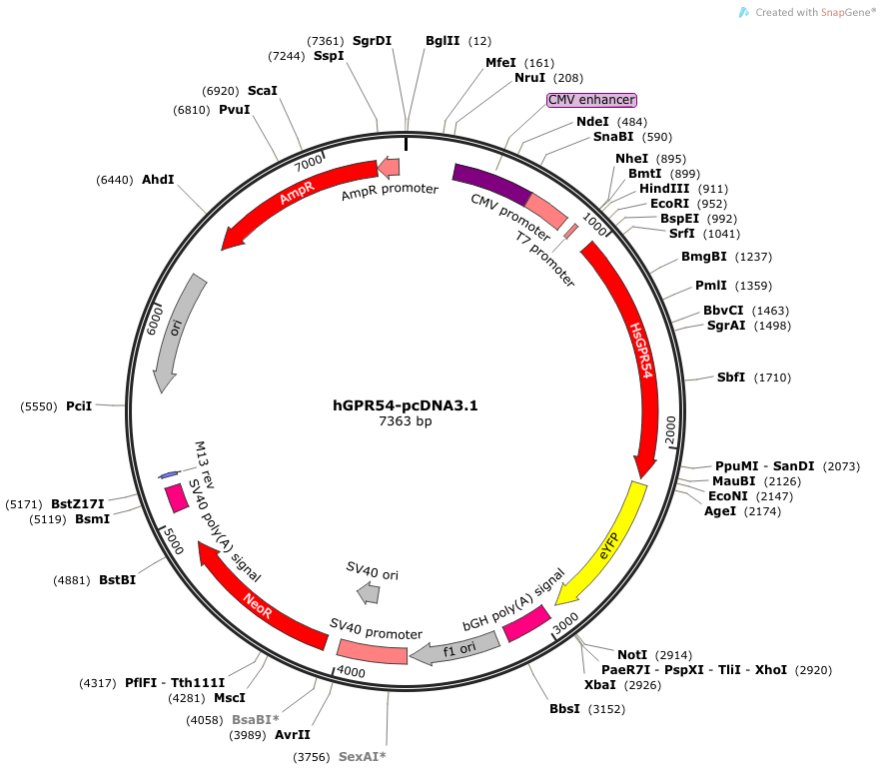
Lipoparticle generation was carried out as per kit instructions (Life Technologies, cat. No. A11667). Briefly, $1x10^7$ FT cells were plated in a T-175 tissue culture flask in 25 mL complete growth medium and incubated overnight at 37 °C, 5% CO_2. The next day, cells were co-transfected with MembranePro™ reagent and plasmid containing GPCR54 using lipofectamine. 9 µg of purified expression construct, 27 µg of MembranePro™ Reagent, and 4 mL of Opti-MEM reduced serum medium were mixed gently in a 15 mL falcon tube. In another tube, 180 µL of Lipofectamine 2000 and 4 mL of Opti-MEM reduced serum medium were mixed. After 5 minutes, both solutions were mixed and incubated again for 20 minutes. The solution was added dropwise to the tissue culture flask while it was rocked gently. The cells were incubated overnight at 37 °C, 5% CO_2.

The medium was discarded and replaced with 32 mL complete culture medium without antibiotics. The cells were incubated again overnight at 37 °C, 5% CO_2. At this point the VLP's should bud from the cell membrane into the culture medium.

After 24 hours, the growth medium was decanted into a 50 mL centrifuge tube and the mixture centrifuged at 2,000 x *g* for 10 minutes to pellet any cell debris. The supernatant was carefully aspirated off using a pipette, making sure not to disturb the pellet. 1/5 Volume of MembranePro™ precipitation mix was added to the medium and the sample mixed by inversion. The sample was incubated for at least 18 hours at 4 °C then VLP's recovered by centrifuging at 5,500 x *g* for 30 minutes at 4 °C. Medium was decanted off and the pellet resuspended in 500 µL PBS. VLPs were stored at 4 °C for 2 days and -80 °C for longer term storage.

8.2.13.4 Generation of His-GPCR54-GFP Nanodiscs using MembraneMax Kit (Life Sciences)

His tagged GPCR54-eYFP plasmid was generated by Rachel Milne (University of Edinburgh) according to the plasmid map below.



Production of recombinant membrane protein from the DNA template using the MembraneMax™ kit (Life Technologies, cat. No. A10632) was carried out as per kit instructions. Briefly, for a typical 100 µL reaction, the following reagents were added to a sterile eppendorf tube (all supplied in kit except DNA template):

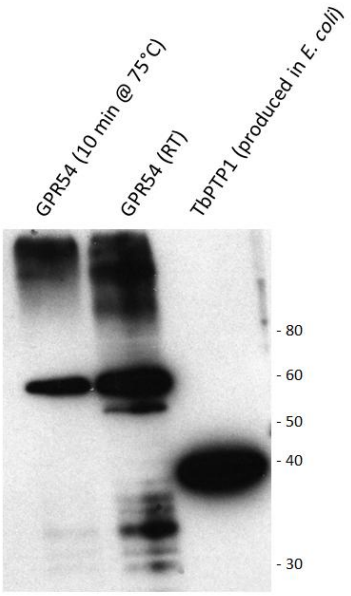
Reagent	Amount
E. coli slyD- Extract	20 µL
2.5X IVPS Reaction Buffer (- amino acids)	20 µL
50 mM Amino Acids (–Met)	1.25 µL
75 mM Methionine	1 µL
MembraneMax™ reagent	2 µL
T7 Enzyme Mix	1 µL
DNA Template	1 µg
DNase/RNase-Free Water	Make up to 50 µL

Tubes were sealed and incubated in a thermomixer at 1,200 rpm, 37 °C for 30 minutes. Feed buffer was prepared in a sterile micro centrifuge tube according to the following amounts:

Reagent	Amount
2X IVPS Feed Buffer	25 µL
50 mM Amino Acids (–Met)	1.25 µL
75 mM Methionine	1 µL
DNase/RNase-Free Water	Make up to 50 µL

Following 30 minutes sample incubation, feed buffer was added and the tube capped and incubated for a further 1.5 - 2 hours at 1,200 rpm, 37 °C. Samples were placed on ice and analysed. Samples were stored long term at -20 °C.

Analysis of reaction samples by Western blot was carried out by Rachel Milne (UoE). Visualisation was carried out using α-His antibody according to the procedure in 7.2.9. Samples were run both at rt. and following incubation at 75 °C for 10 minutes. This did not affect the outcome.



8.2.14 GPCR54 Homology Model

8.2.14.1 Homology Model Generation Based on CXCR4 (Xavier Deupi, PSI and Steven Shave, University of Edinburgh)

GPCR54 belongs to the gamma group of rhodopsin-like GPCRs[245]. This group also contains the chemokine receptor cluster, including CXCR4, whose structure has been recently solved[129]. The sequence identity between GPCR54 and CXCR4 is 20% overall and 23% in the transmembrane region, which allows to use CXCR4 as a template for homology modelling of GPCR54[246]

The sequences of human GPCR54 (UniProt accession number Q969F8) and CXCR4 were aligned using Clustal Omega[247]. This initial alignment was manually refined using Chimera[248] to adjust some of the gaps in the loop regions. Using this alignment, a 3D model of GPCR54 was built using the structures of CXCR4 and bovine rhodopsin[249] as templates using Modeller[250] as follows: (i) the lysozime fused between Ser229 and Lys330 of CXCR4 was deleted; (ii) the 20-residue insertion (between residues 229 and 230 of CXCR4) in the third intracellular loop (ICL3) of GPCR54, which is too long to be modelled reliably using ab-initio methods, was modelled as a short 7-residue loop using ICL3 of rhodopsin as a template; (iii) the C-terminal part of transmembrane helix 7 (TM7) and helix 8, which are partially unfolded in the crystal structure of CXCR4, were modelled instead using the corresponding regions of rhodopsin. As the crystal structure of CXCR4 lacks residues 1-26 in the *N*-terminus, which do not feature clear electron density, and residues 326-352 in the C-terminus, which were truncated in the construct, these regions were not modelled. The cysteine bridge between Cys115 in TM3 and Cys191 in ECL2 (Cys109 and Cys186 in the CXCR4 template) was explicitly defined during model building. Residues missing in the template were refined using the loop optimization method. All models were subjected to 300 iterations of VTFM optimization and thorough MD and SA optimization, and scored using the DOPE potential. The 20 best scoring models were analysed visually, and a suitable model (in term of low score and structure of the loops) was selected for the next step[251]. Eight intramolecular waters, expected to be conserved in GPCRs[252, 253], were added to the modelled GPCR54, and PDB2PQR[251] and PropKa[254] were used to determine protonation states of titratable groups and to add hydrogens to the structure. This model was embedded in a pre-equilibrated lipid bilayer consisting of 279 molecules of 1-palmitoyl-2-oleoyl-sn-glycerol-3-phosphatidylcholine (POPC), and the geometry of the system was optimized by 1000 steps of energy minimization. The temperature was then raised to 300K in 30.000 steps by the temperature reassignment method, followed by 10 ns of equilibration. The final structure of the equilibrated GPCR54 model was used for further analysis. Simulations were performed using the generalized Born implicit solvent method with an ion concentration of 0.2 M, Langevin dynamics for temperature control and a time step of 2 fs, with NAMD 2.8[255] and the CHARMM27 all-hydrogen force field[256].

8.2.14.2 Ligand simulations and docking (Carried out and written up by Agnieszka Bronowska, HITS, Germany)

Prior to molecular docking, the wild-type human kisspeptin decamer (sequence YNWNSFGLRF-NH_2) and all mutants were simulated in explicit water, starting from the fully stretched conformation. The simulation time was 50 ns. Cluster analysis was performed

on the resulting trajectory and structures representative for 3 most populated clusters were selected for peptide-receptor docking calculations. Since kisspeptin is very flexible in water solution and lacking well-defined structure, prediction of the binding modes is very challenging. Thus, the peptide length was initially reduced to terminal 5 residues - the shortest sequence which retains GPCR54 binding and biological activity. This pentamer was docked into GPCR54 receptor using two independent docking methods: UCSF DOCK6.5[257] and Rosetta FlexPepDock.[258, 259] During docking procedure using DOCK6.5 for an initial prediction of the peptide-GPCR54 complexes, in the first step, the last five residues of kisspeptin with amidated *C*-terminus were used to generate models of bioactive complexes. The remaining five residues were added subsequently. The flexible docking approach was used with energy grid scoring in an implicit solvent. The internal degrees of freedom of the peptide were sampled with the anchor-and-grow incremental construction approach.[260] The number of maximum anchor iterations was 500, and the number of maximum grow iterations was 250. Then the procedure was repeated, and the new residue was added, until the whole 10-mer has been docked successfully. The grid space was 0.25 Å. The grid included 20 Å beyond the COM of the putative binding site. The energy score was the sum of electrostatic and Van der Waals contributions. Obtained final structure of peptide-GPCR54 complex was then subjected to 15000 cycles of molecular-mechanical energy minimisation using Amber99SB force field (Amber11 package) with the Generalised Born implicit water model.[261] The cut-off for non-bonded interactions was 12 Å.

In order to refine the obtained binding poses, Rosetta FlexPepDock was employed, which is a high-resolution peptide docking protocol, implemented within the Rosetta framework. It is able to successfully refine a starting structure of a peptide-receptor complex to a near-native model of the interaction. Rosetta FlexPepDock consisted of two alternating modules that optimise the peptide backbone and rigid body orientation, respectively, using the Monte-Carlo with energy minimisation approach. The starting structure was refined in 200 independent simulations. 100 of these simulations included a low-resolution pre-optimisation step, followed by the high-resolution refinement, and 100 of the simulations were carried out strictly in high-resolution mode. A total of 200 models were thus created and then ranked based on their all-atom energy score. For more details on the method, see Raveh et al., 2011.[259] Using this approach, the whole KP-10 was docked. The best-scoring poses obtained were consistent with those obtained for the pentamers and the poses obtained by DOCK, which adds confidence to the predicted binding modes.

Downsized analogues of kisspeptin were simulated independently, starting from the fully stretched conformation. The simulation time was 20 ns in each case. For the head groups of studied pentapeptides, the atomic charges were assigned using AM1-BCC procedure.[262] The head groups were parametrised using Amber99SB force field.[261] All simulations were carried out using the Amber software package (Case et al., 2014).[263] The Amber99SB all-atom force field[261] with the Generalised Born implicit water model were used. The temperature was kept constant at 300 K with using velocity rescaling with a coupling time of 2.5 ps. An integration time step of 2 fs was used. The cut-off for non-bonded interactions was 12 A.

Prior to MD simulations, all ligands were energy-minimised using the steepest-descent method for 500 steps and conjugated gradients for 4500 steps, followed by 20 ps of MD simulations during which positional restraints were used on all bonds and gradually reduced to zero. Finally, ligands were equilibrated for 5 ns. The production period of each simulation was 15 ns, and this was used for the cluster analysis. The lowest-energy clusters were used for molecular docking calculations. The assumption is that the ligands available for binding would belong to the lowest-energy basin.

For molecular docking, the initial ligand conformation was taken from the MD simulation of free (unbound) ligands, as described in previous sections. The grid space was 0.25 A. The grid included 20 A beyond the COM of the putative binding site. The energy score was the sum of electrostatic and Van der Waals contributions. During the docking procedure, the ligands were subjected to 2500 cycles of molecular-mechanical energy minimisation. The number of maximum ligand orientations was 2500. The best-scoring 10 peptide-GPCR54 complexes were further analysed by means of molecular mechanical energy calculations, followed by short (250 ps) MD simulation. The best-scoring poses were subjected to the further refinement. The best-scoring kisspeptin-10-GPCR54 complex resulting from the combination of DOCK and Rosetta FlexPepDock was subjected to 1.5 ns of MD simulations. The assumption was that the positions of the peptides at the binding site of GPCR54 was overall similar, with some differences arising from the *N*-terminal moieties of peptides (head groups in case of pentapeptides, or 5 residues in case of decapeptides, respectively). The simulation protocol was similar to the one described for ligands: molecular-mechanical energy minimisation followed by equilibration and production periods of MD simulation at 300 K. Implicit solvent and positional restraints on backbone atoms of transmembrane helices of GPCR54 were used. The coordinates were saved each 1 ps. The binding energies were calculated using the procedure similar to the standard MM-GBSA approach.[263]

8.2.14.3 Validation of the GPCR54 model by mutagenetic data available (Carried out by Agnieszka Bronowska, HITS, Germany)

The following mutations were specifically investigated: R297L/C223R (R297 is at extracellular loop close to the ligand binding site, C223 at buried part of the transmembrane helix), L102P (putative ligand binding site, at the end of the transmembrane helix), F272S (within the transmembrane helical bundle, involved in interactions with residues that interact with ligand). All of these mutants exhibit impaired GPCR54 signalling.[60, 62]

The *N*-terminal section of docked kisspeptin-10 peptide (residues YNWNSF) seems to form favourable, albeit non-specific hydrophobic contacts with Leu102. Mutation to proline results in a kink in the transmembrane helix and allows for the interactions between the ligand and leucine Leu101 residue. This mutation may affect receptor function due to induced conformational changes, but should not affect ligand binding affinity, since there is another leucine residue in the close proximity. Observed structural information agree with the experimental results, which show that upon Leu102P mutation, an absence of inositol triphosphate accumulation under kisspeptin challenge can be observed, but affinity for kisspeptin remains unaffected.[62] This indirectly supports the GPCR54 model we have and also the proposed model of ligand-receptor interactions.

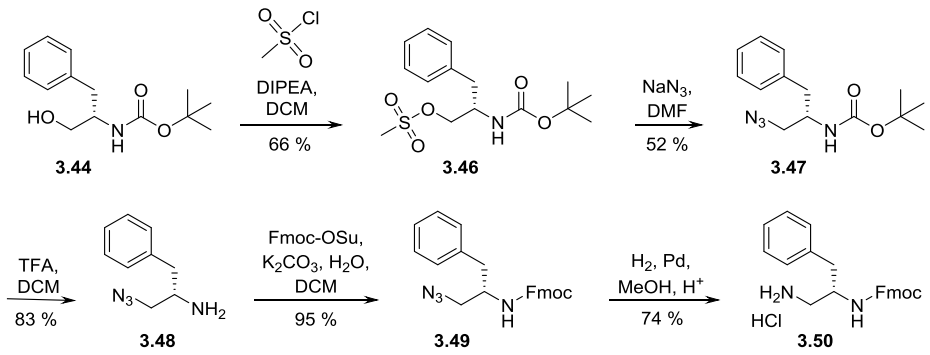
Phenylalanine Phe272 residue is involved in hydrophobic and aromatic network of interactions, which involve residues directly interacting with ligands: most notably with Trp276, also with Val126, Leu212 and Gln280, which interact with Asn208 and Gln122 - key residues predicted to interact with *C*-terminal part of KP-10 peptide, particularly with the side chain of Arg9. Thus, Phe272 is likely to be a transducer of signal from the ligand binding site to G-protein binding area. In experiments, F272S mutation resulted in almost complete inhibition of kisspeptin-induced GPCR54 signalling. This may be rationalised in terms of impaired signal transduction upon mutation (aromatic moiety is crucial for signal transduction, and, likely, for the folding/stability). It validates the GPCR54 model and the proposed model of the binding site. It also indirectly supports the model of the ligand-receptor complex (the orientation of peptide at the binding site).

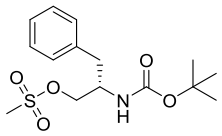
Arginine Arg297 is proximal to ligand binding site and it may directly interact with *N*-terminal part of bound KP-10 peptide (Table 2-7). The interactions are mostly electrostatic in nature (charge-charge) and involve the side-chain of KP-10 and thus are not very specific. Experimental data show that mutation of this arginine to leucine (bulky but lacking positive

charge) results in mild reduction in ligand-stimulated activity.[60] This supports the suggested model of ligand-receptor interactions.

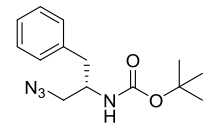
8.3 Chapter 3 Experimental

8.3.1 Preparation of KP-10amine Building Block



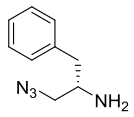


***N*-[(1S)-1-[[(Methylsulfonyloxy)]methyl]-2-phenylethyl]carbamic acid tert-butyl ester (3.46):** *N*-Boc-L-Phenylalaninol (2.0 g, 7.96 mmol) was added to CH_2Cl_2 (100 mL) and to this was added DIPEA (1.7 mL, 9.55 mmol). The mixture was cooled to 0 °C and mesyl chloride (0.7 mL, 9.15 mmol) was added slowly and the mixture allowed to warm to room temperature and stir for 2 hours. The organic layers were washed with water twice then dried and concentrated to dryness to give a white solid, which was purified using flash chromatography (CH_2Cl_2) to give the mesylate (1.74 g, 5.29 mmol, 66%) as a white solid. R_f = 0.47 (50% CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.34 – 7.22 (5H, m, Ar**H**), 4.74 (1H, br s, N**H**), 4.25 (1H, m, NHC**H**), 4.12 (2H, m, CHC**H**$_2$O), 3.02 (3H, s, SC**H**$_3$), 2.96 – 2.84 (2H, m, CHC**H**$_2$Ar), 1.43 (9H, s, C(C**H**$_3$)$_3$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 155.1 (C), 136.6 (C), 129.2 (CH), 128.8 (CH), 127.0 (CH), 80.0 (C), 69.8 (CH_2), 52.5 (CH), 50.8 (CH_2), 37.2 (CH_3), 28.3 (CH_3); MS (ESI) Exact mass calcd for $C_{15}H_{23}NO_5SNa$ $[M+Na]^+$: 352.12, found: 352.05. Spectroscopic data is consistent with that reported previously[264]

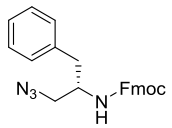


***tert*-Butyl-*N*-[(1*S*)-1-(azidomethyl)-2-phenyl-ethyl] carbamate (3.47):** Methane sulfonate **3.46** (1.7 g, 5.15 mmol) was stirred in DMF (55 mL) and to this was added sodium azide (768 mg, 5.66 mmol) and the mixture stirred at 55 °C overnight. Solvents were removed under partial pressure and the residue taken up in CH_2Cl_2 and washed with water. The organics were dried over $MgSO_4$ then concentrated to dryness. The residue was purified using flash chromatography (1:1 CyHex/EtOAc) to give the title compound (738 mg, 2.67 mmol, 52%) as a white solid. R_f = 0.80 (1:1 CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38 – 7.19 (5H, m, Ar**H**), 4.68 (1H, br s, N**H**), 4.05 – 3.93 (1H, br m, NHC**H**), 3.45 (1H, dd, *J* = 12.1, 3.9 Hz, C**H**$_2$N$_3$), 3.34 (1H, dd, *J* = 12.3, 4.4 Hz, C**H**$_2$N$_3$), 2.94 – 2.86 (1H, m, CHC**H**$_2$Ar), 2.84 – 2.79 (1H, m,

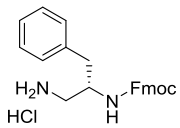
CHC**H**$_2$Ar), 1.45 (9H, s, C(C**H**$_3$)$_3$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 155.1 (C), 137.1 (C), 129.3 (CH), 128.7 (CH), 126.7 (CH), 79.8 (C), 53.1 (CH_2), 51.3 (CH), 38.1 (CH_2), 28.3 (CH_3) MS (ESI) Exact mass calcd for $C_{14}H_{21}N_4O_2$ $[M+H]^+$: 277.17, found: 276.68. The data is in agreement with the literature[265]



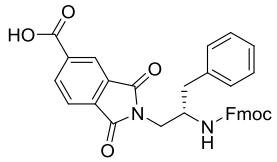
1-Azido-3-phenyl-propan-2-amine (3.48): Azide **3.47** (731 mg, 2.65 mmol) was stirred in 10 mL 50% TFA in CH_2Cl_2 for 5 hours. Solvents were removed under vacuo and the residue taken up in 10 mL CH_2Cl_2 and washed with saturated $NaHCO_3$ (3 x 10 mL). The organics were dried and concentrated to dryness and the residue purified by flash chromatography (9.5:0.5 CH_2Cl_2/MeOH) to give the amine (387 mg, 2.20 mmol, 83%) as a colourless oil. R_f = 0.33 (9.5:0.5 CH_2Cl_2/MeOH); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40 – 7.20 (5H, m, Ar**H**), 3.45 – 3.40 (1H, m, NH_2C**H**), 3.29 – 3.20 (2H, m, C**H**$_2$N$_3$), 2.82 (1H, dd, *J* = 13.4, 5.5 Hz, C**H**$_2$Ar), 2.65 (1H, dd, *J* = 13.4, 5.8 Hz, C**H**$_2$Ar), 1.70 (2H, s, N**H**$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 138.0 (C), 129.2 (CH), 128.7 (CH), 126.7 (CH), 57.2 (CH_2), 52.4 (CH), 41.3 (CH_2). MS (ESI) Exact mass calcd for $C_9H_{13}N_4$ $[M+H]^+$: 177.11, found: 177.00. The data is in agreement with the literature[266]



9H-Fluoren-9-ylmethyl *N*-[(1*S*)-1-(azidomethyl)-2 phenyl-ethyl] carbamate (3.49): (2*S*)-1-Azido-3-phenyl-propan-2-amine **3.48** (350 mg, 1.99 mmol) was stirred in water (15 mL) and to this was added sodium carbonate (421 mg, 3.98 mmol). The mixture was cooled in an ice bath and Fmoc succinimidyl ester (737 mg, 2.19 mmol) in 15 mL CH_2Cl_2 was added with stirring. This was left to warm to room temperature overnight. The organics were run off and washed with water (3x) then dried over $MgSO_4$ and concentrated to dryness. The residue was purified using flash chromatography (1:4 EtOAc/CyHex) to give the title compound (753 mg, 1.89 mmol, 95%) as a white solid. R_f = 0.27 (1:4 EtOAc/CyHex); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 (2H, d, *J* = 7.6 Hz, Ar**H**), 7.58 (2H, t, *J* = 6.5 Hz, Ar**H**), 7.43 (2H, t, *J* = 7.5 Hz , Ar**H**), 7.34 (4H, t, *J* = 7.3 Hz, Ar**H**), 7.30 – 7.25 (1H, m, Ar**H**), 7.22 (2H, d, *J* = 6.8 Hz, Ar**H**), 7.20 (1H, br d, *J* = 7.7 Hz, N**H**), 4.40 (2H, m, CHC**H**$_2$O), 4.20 (1H, t, *J* = 6.6 Hz, C**H**CH$_2$O), 4.05 (1H, m, NHC**H**), 3.40 (2H, m, C**H**$_2$N$_3$), 2.87 (2H, m, C**H**$_2$Ar); ^{13}C NMR (125 MHz, $CDCl_3$) δ 155.6 (C), 143.8 (C), 141.4 (C), 136.8 (C), 129.3 (CH), 128.8 (CH), 127.4 (CH), 127.1 (CH), 126.9 (CH), 125.1 (CH), 120.0 (CH), 66.8 (CH_2), 53.2 (CH_2), 51.8 (CH), 47.3 (CH), 38.1 (CH_2); MS (ESI) Exact mass calcd for $C_{24}H_{23}N_4O_2$ $[M+H]^+$: 399.18, found: 398.96. The data is in agreement with the literature[267]



9H-Fluoren-9-ylmethyl-*N*-[(1*S*)-1-(aminometh-yl)-2-phenyl-ethyl] carbamate hydrochloride (3.50): Azide **3.49** (350 mg, 0.88 mmol) was stirred in 25 mL methanol and to this was added 2M HCl in dioxane (1.36 mL, 2.72 mmol). The reaction was put under an argon atmosphere and palladium on carbon 5% loading (25 mg) was added. The flask was then evacuated and filled with a hydrogen atmosphere and stirred overnight. The mixture was filtered through celite and concentrated to dryness. The resulting gum was triturated with diethyl ether to give the amine hydrochloride salt (275 mg, 0.67 mmol, 74%) as an off white solid. R_f = 0.68 (1:9 MeOH/CH_2Cl_2); ^{1}H NMR (400 MHz, DMF-d_7) δ 8.93 (2H, br s, N**H**$_2$), 7.94 (2H, d, *J* = 7.6 Hz, Ar**H**), 7.75 (2H, t, *J* = 8.4 Hz, Ar**H**), 7.45 (2H, t, *J* = 7.5 Hz, Ar**H**), 7.4 - 7.26 (6H, m, Ar**H**), 7.25 - 7.18 (1H, m, Ar**H**), 4.29 – 4.13 (4H, m, CHC**H**$_2$O, C**H**CH$_2$O, NHC**H**), 3.22 (2H, br s, C**H**$_2$NH$_2$), 3.06 (2H, m, C**H**$_2$Ar); ^{13}C NMR (200 MHz, DMF-d_7) δ 157.3 (C), 145.3 (C), 142.2 (C), 139.4 (C), 130.8 (C), 130.5 (C), 129.8 (CH), 129.4 (CH), 128.8 (CH), 128.2 (CH), 128.2 (CH), 127.5 (CH), 126.8 (CH), 126.6 (CH), 121.1 (CH), 67.3 (CH), 52.5 (CH), 48.1(CH$_2$), 39.1 (CH$_2$), 37.7 (CH$_2$); MS (ESI) Exact mass calcd for $C_{24}H_{25}N_2O_2$ $[M+H]^+$: 373.19, found: 373.17.



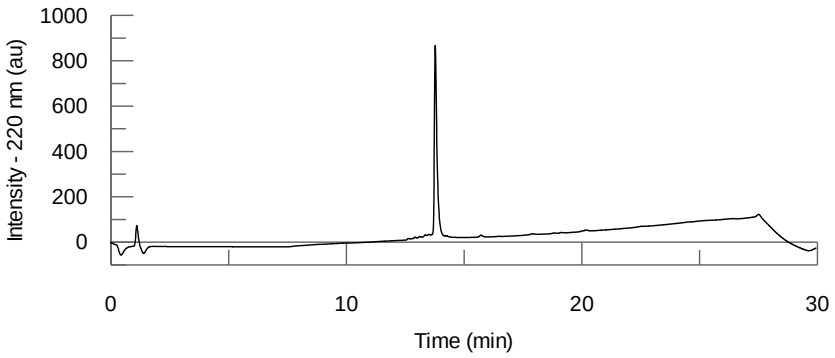
2-[(2S)-2-(9H-Fluoren-9-ylmethoxycarbonyl -amino)-3-phenyl-propyl]-1,3-dioxo-isoindo line-5-carboxylic acid (3.52): Amine hydrochloride **3.50** (100 mg, 0.93 mmol) and trimellitic anhydride (267 mg, 1.39 mmol) were stirred in 25ml DMF then pyridine (186 μL, 2.32 mmol) was added and the mixture stirred at rt for 1.5 hours. LCMS of the crude mixture showed consumption of starting material and formation of the phthlamic acid intermediate. 1,1'-carbonyldiimidazole (450 mg, 2.78 mmol) was added to the solution and this was stirred at room temperature for another hour. Solvents were removed under vacuo and the residue taken up in water which was acidified to pH = 1 using 1M HCl. The precipitated solid was taken up in EtOAc then the organic layers washed twice more with water. Organics were dried over $MgSO_4$ then concentrated to dryness. The resulting white gum was recrystallised from H_2O/EtOH to give the phthalimide (47 mg, 0.09 mmol, 38%) as a white solid. R_f = 0.30 (1:9 MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, DMSO d_6) δ 13.65 (1H, br s, COO**H**), 8.29 (1H, d, *J* = 7.4 Hz, Ar**H**), 8.15 (1H, s, Ar**H**), 7.90 (1H, d, *J* = 7.6 Hz, Ar**H**), 7.85 (2H, d, *J* = 7.6 Hz, Ar**H**), 7.55 (1H, d, *J* = 7.6 Hz, Ar**H**), 7.49 (1H, d, *J* = 7.4 Hz, Ar**H**), 7.45 – 7.34 (3H, m, Ar**H**), 7.32 – 7.20 (5H, m, Ar**H**), 7.15 (1H, d, *J* = 5.8 Hz, Ar**H**), 4.12 – 4.05 (1H, m, COOC**H**$_2$CH), 4.05 – 3.99 (1H, m, NHC**H**$_2$), 3.98 - 3.91 (2H, m, COOC**H**$_2$), 3.75 (2H, d, *J* = 6.0 Hz, NC**H**$_2$), 2.93 – 2.72 (2H, m, C**H**$_2$Ar); ^{13}C NMR (125 MHz, $CDCl_3$) δ 167.4 (C), 167.4 (C), 165.9 (C), 156.0 (C), 144.1 (C), 143.7 (C), 140.7 (C),

140.7 (C), 138.8 (C), 136.2 (C), 135.3 (C), 135.1 (C), 132.2 (CH), 129.2 (CH), 128.3 (CH), 127.7 (CH), 127.2 (CH), 127.1 (CH), 126.3 (CH), 125.4 (CH), 125.2 (CH), 123.5 (CH), 123.1 (CH), 120.2 (CH), 65.6 (CH_2), 51.5 (CH), 46.7 (CH), 42.2 (CH_2), 37.6 (CH_2). MS (ESI) Exact mass calcd for $C_{26}H_{22}N_2O_6$ $[M+H]^+$: 547.14, found: 547.04.

8.3.2 Synthesis of KP-10amine Peptide

KP-10-amine (3.56)

The peptide was synthesised according to the general procedures outlined. MS (ESI) Exact mass calcd for $C_{63}H_{86}N_{17}O_{13}$ $[M+H]^+$: 1288.66, found: 1288.66; Crude purity (RP-HPLC): 68%. The peptide was purified using semi preparative RP-HPLC method B. >95% purity.



8.3.3 Inositol-(1, 4, 5)-triphosphates accumulation assay

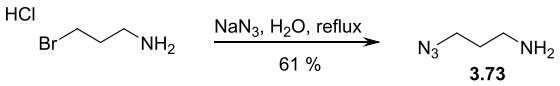
This assay was performed by Megan Gallant, University of Capetown, SA

COS-7 cells electroporated with GPCR54 were seeded in 12 well plates (100 000 cells/1ml) and incubated for 24 hours in a 37 °C 5% CO_2 incubator. Thereafter, the medium was aspirated and cells were rinsed with 0.5 mL of Med199 medium containing 2% FCS for 15 minutes. Cells were radiolabelled with 0.5 mL 2uCi/ mL myo [2-3H (N)]-inositol (Perkin-Elmer) in Med199/2%FCS and incubated for 20 hours. After which, cells were incubated with Buffer-I (140mM NaCl, 4 mM KCl, 20 mM Hepes, 8 mM glucose, 0.1% bovine serum albumin (BSA), 1 mM $MgCl_2$, 1 mM $CaCl_2$, and 10 mM LiCl) for 15 minutes at 37 ˚C. Peptides were diluted (0-1 µM) in Buffer- I and cells were stimulated for 90 minutes at 37 ˚C. The media was aspirated and 1 mL 10 mM formic acid was added to each well and the cells were extracted by incubation at 4 ˚C for 30 minutes.

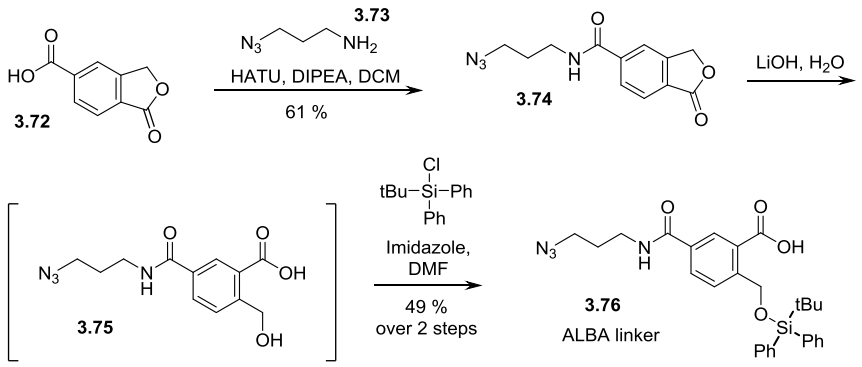
Columns packed with Dowex 1x 8-200 ion-exchange resin (Sigma) were washed sequentially with 3 mL 3M ammonium formate/0.1 M formic acid and 10 mL distilled water. The cell extracts were added to the columns, followed by 10 mL distilled water and then 5 mL 0.1 M ammonium formate/0.1 M formic acid. Samples were then eluted in 1 mL 1M ammonium formate/0.1M formic acid. Scintillation vials were prepared with 2.6 mL scintillation fluid and 1 mL of elute was added to each vial and mixed well. Samples were counted in a Tri-carb 2100TR liquid scintillation analyser (Packard).

8.3.4 Preparation of ALBA Linker

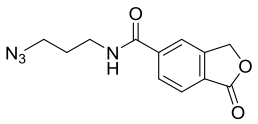
3-Azidopropylamine (3.73):



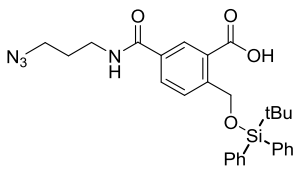
1-Bromo-3-aminopropane hydrochloride (1.0 g, 5.73 mmol) was stirred in water (5 mL) followed by the addition of NaN_3 (1.1 g, 17.0 mmol) in water (5 mL). The mixture was heated to reflux overnight then concentrated to a third of the volume under partial pressure. The remaining mixture was cooled to 0ºC and Et_2O (10 mL) and KOH pellets (1.2 g, 21.43 mmol) added. The organic layer was separated and the aqueous washed twice more with Et_2O and the organic layers pooled and dried over $MgSO_4$ and concentrated to dryness. This gave a yellow oil (380 mg, 3.8 mmol, 67%) which was used without further purification. Rf = 0.17 (10:1:0.1 CH_2Cl_2/MeOH/7M NH_3 in MeOH); ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.38 (2H, t, J = 6.8 Hz, $N_3C\mathbf{H}_2$), 2.82 (2H, t, J = 6.8 Hz, $NH_2C\mathbf{H}_2$), 1.74 (2H, quint, J = 6.8 Hz, $CH_2C\mathbf{H}_2CH_2$); ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 49.3 (CH_2), 39.4 (CH_2), 32.4 (CH_2). MS (ESI) Exact mass calcd for $C_3H_9N_4$ $[M+H]^+$: 101.08 found: 100.95. Data is in agreement with the literature[268]



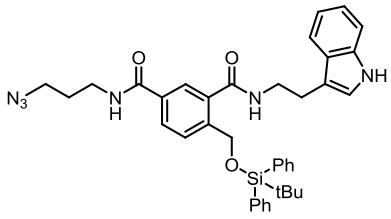
For the synthesis of 5-carboxyphthalide see section 7.6.1



***N*-(3-Azidopropyl)-1-oxo-3H-isobenzofuran-5-carboxamide (3.74):** 3-3-Azidopropylamine **3.73** (380 mg, 3.8 mmol), 5-carboxyphthalide (676 mg, 3.8 mmol) and HBTU (1729 mg, 4.56 mmol) were stirred in dry CH_2Cl_2 (40 mL) at rt and DIPEA (1580 mL, 9.12 mmol) was added. The solution was stirred overnight then washed with sat. $NaHCO_3$ (2 x 30 mL). The organics were dried over $MgSO_4$ then concentrated to dryness. The residue was purified by flash chromatography (20% EtOAc/CyHex) to give the title product as a white solid (602 mg, 2.3 mmol, 61%). R_f = 0.28 (20% EtOAc/CyHex); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.79 (1H, t, *J* = 5.3 Hz, N**H**), 8.08 (1H, s, Ar**H**), 8.00 (1H, d, *J* = 8.0 Hz, Ar**H**), 7.93 (1H, d, *J* = 8.0 Hz, Ar**H**), 5.47 (2H, s, C=OOC**H**$_2$), 3.43 (2H, t, *J* = 6.7 Hz, $N_3C\mathbf{H}_2$), 3.36 (2H, m, NHC**H**$_2$), 1.80 (2H, quin, *J* = 6.7 Hz, $CH_2C\mathbf{H}_2CH_2$); ^{13}C NMR (125.8 MHz, $(CD_3)_2SO$) δ 170.4 (C), 165.8 (C), 147.8 (C), 140.2 (C), 128.2 (C), 127.3 (CH), 125.2 (CH), 122.3 (CH), 70.3 (CH_2), 48.8 (CH_2), 37.2 (CH_2), 28.6 (CH_2); MS (ESI) Exact mass calcd for $C_{12}H_{13}N_4O_3$ $[M+H]^+$: 261.26, found: 260.94.



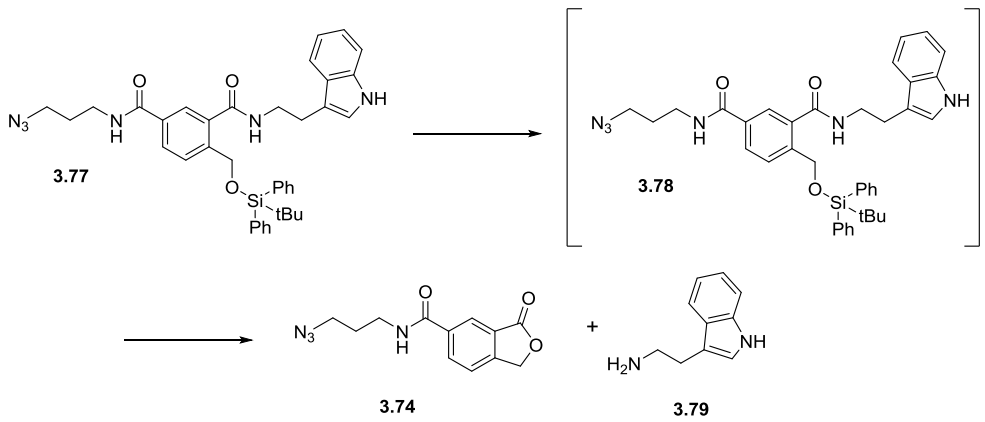
ALBA Linker (3.76): Phthalide **3.74** (560 mg, 2.15 mmol) was stirred in 5 mL of an 85% solution of methanol in H_2O solution and $LiOH.H_2O$ (90 mg, 2.15 mmol) was added. The mixture was heated to 50 °C overnight. Solvents were removed under reduced pressure then the residue taken up in pyridine (5 mL) and TBDPSiCl (7.1 mmol, 1.8 mL). The mixture was stirred overnight then saturated $NaHCO_3$ was added and extracted with CH_2Cl_2. The organics were run off and dried then concentrated to dryness, then the resulting oil was dissolved in methanol (20 mL) and THF (7 mL) and treated with an aqueous solution of K_2CO_3 (700 mg, 7 mL). This was stirred for 30 minutes then the solution concentrated to around one quarter. The mixture was cooled on ice and adjusted to pH 1 with 1M potassium hydrogen sulphate, extracted with diethyl ether and dried over $MgSO_4$. The residue was purified using flash chromatography (9:1 CyHex/EtOAc, then EtOAC) to give the title compound as a white solid (548 mg, 1.06 mmol, 49%). R_f = 0.33 (EtOAc); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.68 (1H, t, *J* = 5.6 Hz, N**H**), 8.40 (1H, d, *J* = 1.7 Hz, Ar**H**), 7.93 (1H, d, *J* = 8.0 Hz, Ar**H**), 7.80 (1H, dd, *J* = 8.0, 1.7 Hz, Ar**H**), 7.65 (4H, m, Ar**H**), 7.44 (6H, m, Ar**H**), 5.13 (2H, s, CC**H**$_2$O), 3.44 (2H, t, *J* = 6.8 Hz, $N_3C\mathbf{H}_2$), 3.36 (2H, m, NHC**H**$_2$), 1.8 (2H, quin, *J* = 6.8 Hz, $CH_2C\mathbf{H}_2CH_2$), 1.06 (9H, s, $C(C\mathbf{H}_3)_3$); ^{13}C NMR (125.8 MHz, $(CD_3)_2SO$) δ 167.5 (C), 166.0 (C), 142.4 (C), 137.9 (C), 135.0 (CH), 132.7 (C), 130.2 (CH), 130.1 (C), 130.0 (CH), 128.0 (CH), 125.6 (CH), 125.1 (CH), 63.4 (CH_2), 48.5 (CH_2), 36.7 (CH_2), 28.3 (CH_2), 26.6 (CH), 18.9 (C), 14.1 (CH_3); MS (ESI) Exact mass calcd for $C_{28}H_{33}N_4O_4Si$ $[M+H]^+$: 517.23, found: 517.12.



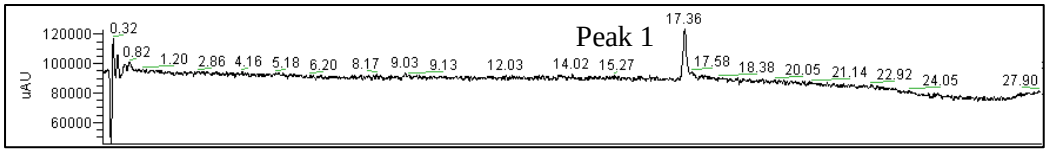
ALBA-Tryptamine (3.77): ALBA linker **3.76** (50 mg, 0.097 mmol) was stirred with tryptamine (19 mg, 0.116 mmol) in 5 mL CH_2Cl_2. To this was added HATU (44 mg, 0.116 mmol) and DIPEA (34 µL, 0.194 mmol) at room temperature with stirring. After 3 hours, the mixture was washed with water x 3 and the organic layers run off and dried over $MgSO_4$. Solvents were removed in vacuo and the residue purified using silica gel chromatography (8:2 EtOAc/CyHex) to give the amide as a white solid (52 mg, 0.080 mmol, 82%). R_f = 0.49 (EtOAc); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 7.96 (1H, br s, N**H**), 7.70 (1H, dd, *J* = 8.0, 1.8 Hz, Ar**H**), 7.65 - 7.55 (6H, m, Ar**H**), 7.46 - 7.42 (2H, m, Ar**H**), 7.39 - 7.32 (5H, m, Ar**H**), 7.23 – 7.08 (3H, m, Ar**H**), 6.96 (1H, d, *J* = 2.3 Hz, Ar**H**), 6.20 (1H, br t, *J* = 5.8 Hz, N**H**), 4.84 (2H, s, C**H**$_2$Si), 3.79 - 3.72 (2H, m, tryptamine C**H**$_2$CH$_2$NH), 3.51 (2H, q, *J* = 6.6 Hz, N_3C**H**$_2$CH$_2$CH$_2$), 3.42 (2H, t, *J* = 6.6 Hz, $N_3CH_2CH_2C$**H**$_2$), 3.06 (2H, t, *J* = 7.0 Hz, tryptamine CH_2C**H**$_2$NH), 1.89 (2H, quin, *J* = 6.6 Hz, N_3CH_2C**H**$_2CH_2$), 1.03 (9H, s, C(C**H**$_3$)$_3$); ^{13}C NMR (125.8 MHz, $(CD_3)_2SO$) δ 167.8 (C), , 166.7 (C), 138.1 (C), 137.9 (C), 136.4 (C), 135.8 (C), 135.7 (CH), 132.8 (C), 130.0 (CH), 128.9 (CH), 127.9 (CH), 127.3 (C), 127.2 (CH), 126.6 (CH), 122.2 (CH), 121.9 (CH), 119.5 (CH), 118.7 (CH), 112.9 (C), 111.2 (CH), 64.3 (CH_2), 49.4 (CH_2), 40.4 (CH_2), 37.7 (CH_2), 28.8 (CH_2), 26.8 (CH_3), 25.4 (CH_2), 19.2 (C); MS (ESI) Exact mass calcd for $C_{38}H_{42}N_5O_4Si$ $[M+H]^+$: 659.31, found: 659.22.

8.3.5 Solution Based ALBA Stability Studies

In order to test the stability of the ALBA linker to certain conditions, it was loaded with tryptamine (see above for experimental procedure) in solution. This compound was then weighed into a vial (1 mg/mL) and stirred in different solutions for a given amount of time. LCMS analysis of the product showed how the linker behaved in each. LCMS photodiode array traces shown.

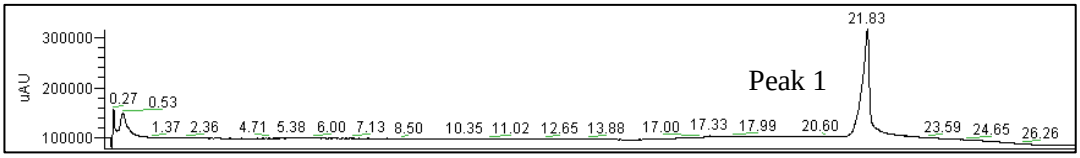


Reference - Tryptamine loaded ALBA linker (3.77)



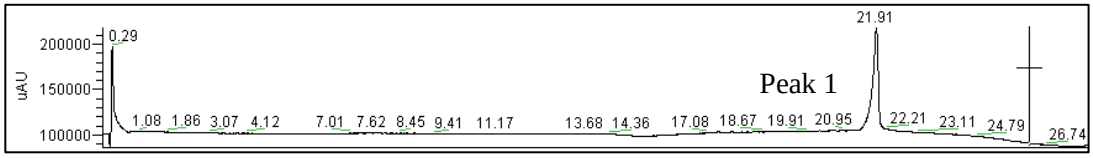
Peak	t_R	m/z observed	Species Observed
1	17.36	658.8	**3.77**

1 hour 20% piperidine in DMF - Linker stable to these conditions



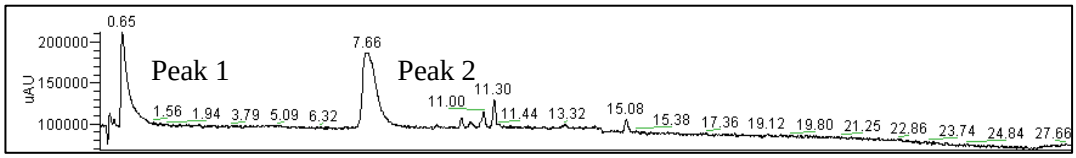
Peak	t_R	m/z observed	Species Observed
1	21.83	659.0	**3.77**

1 hour 10% DIPEA in DMF - Linker stable to these conditions



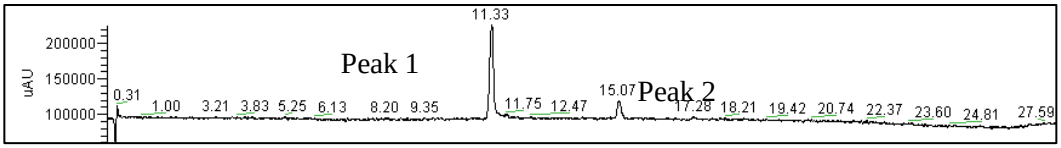
Peak	t_R	m/z observed	Species Observed
1	21.91	659.2	**3.77**

4 hours in TFA - Linker cleavage appears complete after 4 hours due to disappearance of starting material and formation of benzofuranone compound **3.74** (peak 2) and release of tryptamine **3.79** (peak 1).



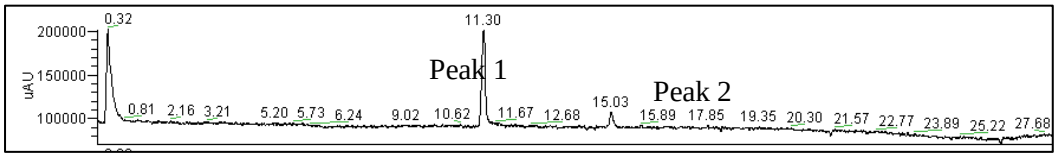
Peak	t_R	m/z observed	Species Observed
1	0.65	161.0	**3.79**
2	7.66	260.9	**3.74**
3	15.08	None	TBDPS group

30 min 1mM TBAF in THF - TBDPS group removed after 30 minutes due to disappearance of starting material and appearance of deprotected linker compound **3.78**.



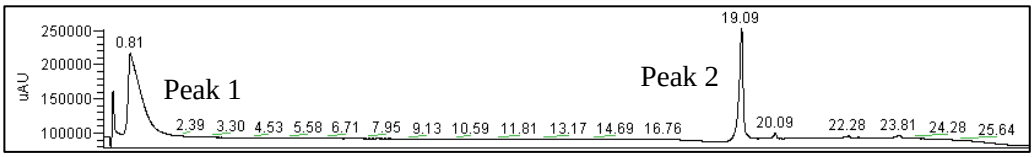
Peak	t_R	m/z observed	Species Observed
1	11.33	421.1	**3.78**
2	15.07	none	TBDPS group

30 min 1mM TBAF in THF then overnight pyridine/H_2O - Linker cleavage with pyridine appears unsuccessful, even overnight with no evidence of benzofuranone, only deprotected linker compound **3.78**.



Peak	t_R	m/z observed	Species Observed
1	11.30	421.1	**3.78**
2	15.03	none	TBDPS group

4 hours 0.1 M TBAF, 5% TFA, 5% H_2O in CH_2Cl_2 - Linker cleavage appears complete after 4 hours due to disappearance of starting material and formation of benzofuranone compound **3.74** (peak 2) and release of tryptamine **3.79** (peak 1).



Peak	t_R	m/z observed	Species Observed
1	0.81	160.97	**3.79**
2	19.09	260.92	**3.74**
3	15.08	None	TBDPS group

8.3.6 Preparation of Strained Cyclooctyne

Oxo(1,2,3,8-tetramethoxy-11,12-didehydrodibenzo[*b*,*f*]- azocin-5(6*H*)-yl)butanoic acid: This was prepared according to the literature.[163]

8.3.7 Solid Phase Protocols

Cleavage of phthalimide Linker

Phthalimide linker was cleaved as described in the literature.[156]

Loading ALBA Linker onto Resin

ALBA linker (10eq) was dissolved in MeCN/H_2O (1:1) then the solution applied to resin loaded with oxo(1,2,3,8-tetramethoxy-11,12-didehydrodibenzo[b,f]- azocin-5(6H)-yl)butanoic acid overnight. The mixture was shaken overnight at room temperature. The resin was drained and the excess linker retained for future use. Beads were washed extensively with DMF, then CH_2Cl_2.

TFA Cleavage of ALBA Linker

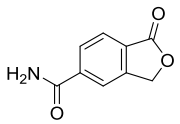
Resin was subjected to a solution of 2.5% TIS, 2.5% H_2O in TFA for 4 hours. Following this, the resin was drained and the liquid collected. The resin was rinsed with TFA again and the filtrate concentrated to dryness. The residue was taken up in cold diethyl ether to precipitate the side chain deprotected peptide, which was filtered off and dried.

Fluoride/Dilute Acid Cleavage of ALBA Linker

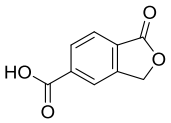
Resin was subjected to a solution of 0.1 M TBAF in THF for 30 minutes. The solution was then drained from the resin and the resin washed with DMF 3 times. A solution of 5% TFA, 5% H_2O in CH_2Cl_2 and this was added to the resin. The mixture was shaken for 4 hours. This solution was collected and concentrated to dryness.

8.4 Chapter 4 Experimental

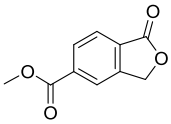
8.4.1 Preparation of 5-CO_2Me-Phthalide



5-Carbomylphthalide: This was prepared as per the literature.[269] 5-Cyanophthalide (4.10 g, 25.8 mmol) was stirred in 75% sulphuric acid (40 mL) and the mixture heated to 80 °C for 3 hours. The solution was allowed to cool to rt then poured over ice, causing a white solid to precipitate which was filtered off and washed with water. This was dried overnight at 120 °C to give 5-carbamoylphthalide (4.25 g, 23.99 mmol, 93%). R_f = 0.52 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.22 (1H, s, N**H**$_2$), 8.11 (1H, s, N**H**$_2$), 8.03 (1H, t, *J* = 8.0 Hz, Ar**H**), 7.92 (1H, d, *J* = 8.0 Hz, Ar**H**), 7.66 (1H, s, Ar**H**), 5.47 (2H, s, C**H**$_2$); ^{13}C NMR (125.8 MHz, $(CD_3)_2SO$) δ 170.1 (C), 167.0 (C), 147.4 (C), 139.6 (C), 128.1 (C), 127.1 (CH), 124.9 (CH), 122.1 (CH), 70.0 (CH_2); MS (ESI) Exact mass calcd for $C_9H_8NO_3$ $[M+H]^+$: 178.05, found: 178.16.



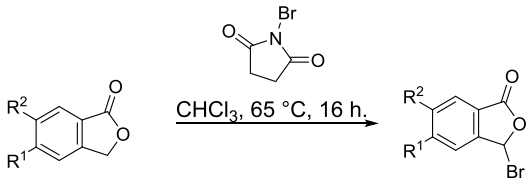
5-Carboxyphthalide (4.55): This was prepared as per the literature.[269] 5-Carbamoylphthalide (4.10 g, 23.02 mmol) was stirred in 2M NaOH solution and heated to reflux overnight. The solution was cooled to rt and acidified to pH1 using HCl which caused a white solid to precipitate. This was filtered off and washed with water. This was dried at 120 °C overnight to give 5-carboxyphthalide (4.01 g, 22.51 mmol, 98%). R_f = 0.33 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.22 (1H, s, Ar**H**), 8.10 (1H, dd, *J* = 8.0, 0.6 Hz, Ar**H**), 7.95, (1H, d, *J* = 8.0 Hz, Ar**H**), 5.47 (2H, s, C**H**$_2$); ^{13}C NMR (125.8 MHz, $(CD_3)_2SO$) δ 170.4 (C), 167.0 (C), 148.1 (C), 136.3 (C), 130.3 (C), 129.0 (CH), 125.6 (CH), 124.5 (CH), 70.6 (CH_2); MS (ESI) Exact mass calcd for $C_9H_7O_4$ $[M+H]^+$: 179.03, found: 179.24.



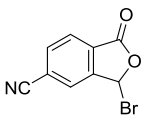
Methyl 1-oxo-3H-isobenzofuran-5-carboxylate (4.56): 5-Carboxyphthalide **4.55** (92 mg, 0.52 mmol) was stirred in MeOH (0.5 mL) and to this was added thionyl chloride (190 µL, 2.58 mmol) and the mixture was heated to reflux for 4 hours. Solvents were removed under reduced pressure and the residue triturated with diethyl ether to give a white solid (52 mg, 2.69 mmol, 52%). R_f = 0.57 (40% CyHex/EtOAc); ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.22 (1H, d, *J* = 7.9 Hz, Ar**H**), 8.19 (1H, s, Ar**H**), 8.0 (1H, d, *J* = 7.9 Hz, Ar**H**), 5.39 (2H, s, C**H**$_2$), 3.99 (3H, s, C**H**$_3$); ^{13}C NMR (150 MHz, $CDCl_3$) δ 169.9 (C), 165.8 (C), 146.4 (C), 135.4 (C), 130.4 (CH), 129.4 (C), 125.8 (CH), 123.5 (CH), 69.6 (CH_2), 52.7 (CH_3); MS (ESI) Exact mass calcd for $C_{10}H_9O_4$ $[M+H]^+$: 193.05, found: 193.22.

8.4.2 Preparation of 3-Bromophthalides

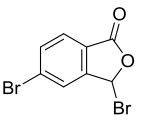
General procedure for bromination of phthalides (no radical initiator): Phthalide (1eq) was stirred in $CHCl_3$ (100 mg/mL) in a vessel and to this was added *N*-bromosuccinimide (1.6 eq). The mixture was stirred and heated at 65 °C overnight. Precipitated solid was filtered off and the resulting solution used directly for flash chromatography.



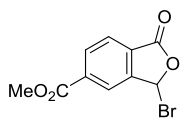
General procedure for bromination of phthalides (with radical initiator): The procedure above was followed, however ACHN (0.2 eq) was added to the reaction mixture before heating and the reaction was irradiated with UV light (245 nm, 6 Watt, Spectroline EBF-260C lamp) overnight.



3-Bromo-1-oxo-3H-isobenzofuran-5-carbonitrile (4.51): 5-Cyanophthalide **4.49** (400 mg, 2.51 mmol) was used as starting material and the title compound was obtained as a white solid (420 mg, 1.76 mmol, 70%). Eluent for chromatography: CH_2Cl_2; R_f = 0.39; ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.07 (1H, d, *J* = 7.9 Hz, Ar**H**), 7.96 (1H, s, Ar**H**), 7.92 (1H, dd, *J* = 7.9, 0.9, Ar**H**), 7.44 (1H, s, C**H**Br); ^{13}C NMR (125 MHz, $CDCl_3$) δ 165.3 (C), 149.2 (C), 134.5 (CH), 127.6 (C), 127.6 (CH), 126.9 (CH), 118.9 (C), 116.8 (C), 73.0 (CH); MS (ESI) Exact mass calcd for $C_9H_4BrNNaO_2$ $[M+Na]^+$: 259.93, found: 259.93.

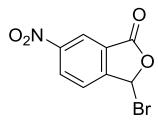


3,5-Dibromo-3H-isobenzofuran-1-one: 5-Bromophthalide **4.48** (500 mg, 2.35 mmol) was used as starting material and the title compound was obtained as a white solid (428 mg, 1.48 mmol, 63%). Eluent for chromatography: $CHCl_3$; R_f = 0.59; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.82 – 7.78 (2H, m, Ar**H**), 7.78 – 7.77 (1H, m, Ar**H**), 7.35 (1H, s, C**H**Br); ^{13}C NMR (125 MHz, $CDCl_3$) δ 166.3 (C), 150.4 (C), 134.6 (CH), 130.5 (C), 127.1 (CH), 127.0 (CH), 123.0 (C), 73.2 (CH); MS (EI) Exact mass calcd for $C_8H_4Br_2O_2$ $[M+H]^+$: 290.8, found: 290.8.



Methyl 3-bromo-1-oxo-3H-isobenzofuran-5-carboxylate (4.58): Methyl 1-oxo-3H-isobenzofuran-5-carboxylate **4.58** (50 mg, 0.26 mmol) was used as starting material and the title compound was obtained as a white solid (59 mg, 0.22 mmol, 85%). Eluent for chromatography: $CHCl_3$; R_f = 0.58; ^{1}H NMR (500

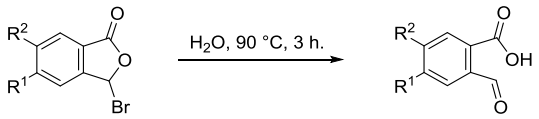
MHz, $CDCl_3$) δ 8.32 – 8.27 (2H, m, Ar**H**), 8.01 (1H, d, *J* = 8.2 Hz, Ar**H**), 7.44 (1H, s, C**H**Br), 4.01 (3H, s, C**H**$_3$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 166.3 (C), 165.1 (C), 148.9 (C), 136.6 (C), 132.1 (CH), 127.5 (C), 126.0 (CH), 124.9 (CH), 74.0 (CH), 53.0 (CH_3); MS (ESI) Exact mass calcd for $C_{10}H_7BrNaO_4$ $[M+Na]^+$: 292.94, found: 292.94.



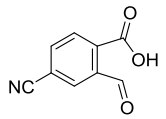
3-Bromo-6-nitro-3H-isobenzofuran-1-one: 5-Nitrophthalide **4.50** (200 mg, 1.12 mmol) was used as starting material and the title compound was obtained as a colourless oil (174 mg, 0.68 mmol, 61%). Eluent for chromatography: $CHCl_3$; R_f = 0.20; ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.76 (1H, d, *J* = 2.0 Hz, Ar**H**), 8.66 (1H, dd, *J* = 8.4, 2.0 Hz, Ar**H**), 7.86 (1H, d, *J* = 8.4 Hz, Ar**H**), 7.49 (1H, s, C**H**Br); ^{13}C NMR (125 MHz, $CDCl_3$) δ 164.9 (C), 153.6 (C), 149.9 (C), 130.1 (CH), 125.8 (C), 125.0 (CH), 121.6 (CH), 73.0 (CH).

8.4.3 Preparation of Phthalaldehydic Acids

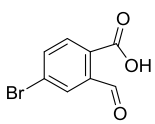
General procedure for phthalaldehydic acid preparation:



Bromophthalide was stirred in H_2O (50 mg/mL) and the mixture heated to 90 °C for 3 hours then allowed to cool to rt. The aqueous solution was extracted with Et_2O (3x) then the organic layers were combined, dried over $MgSO_4$ and concentrated to dryness to give the required product.

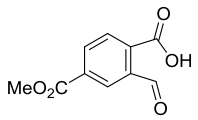


4-Cyano-2-formyl-benzoic acid (4.52): 3-Bromo-1-oxo-3H-isobenzofuran-5-carbonitrile **4.51** (400 mg, 1.69 mmol) was used as starting material and the title compound was obtained as a white solid (290 mg, 1.66 mmol, 98%). R_f = 0.19 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.42 (1H, br s, CO_2**H**), 8.26 (1H, s, Ar**H**), 8.13 (1H, d, *J* = 7.8 Hz, Ar**H**), 8.02 (1H, d, *J* = 7.8 Hz, Ar**H**), 6.73 (1H, br s, CO**H**); ^{13}C NMR (150 MHz, $(CD_3)_2SO$) δ 167.0 (C), 147.7 (C), 134.6 (CH), 130.4 (C), 128.2 (CH), 125.7 (CH), 117.8 (C), 116.5 (C), 98.3 (CH); MS (EI) Exact mass calcd for $C_9H_5NO_3$ M^+: 175.0, found: 175.0.

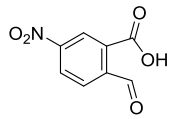


4-Bromo-2-formyl-benzoic acid: 3,5-Dibromo-3H-isobenzofuran-1-one (400 mg, 1.38 mmol) was used as starting material and the title compound was obtained as a white solid (282 mg, 1.23 mmol, 90%). R_f = 0.56 (10%

MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.28 (1H, br s, CO_2**H**), 7.93 (1H, s, ArH), 7.86 (1H, d, *J* = 7.7 Hz, Ar**H**), 7.77 (1H, d, *J* = 7.8 Hz, Ar**H**), 6.65 (1H, br s, CO**H**); ^{13}C NMR (150 MHz, $(CD_3)_2SO$) δ 167.5 (C), 149.4 (C), 133.8 (CH), 128.5 (C), 127.0 (CH), 126.4 (CH), 125.8 (C), 97.7 (CH); MS (EI) Exact mass calcd for $C_8H_6BrO_3$ M^+: 227.9, found: 227.9.



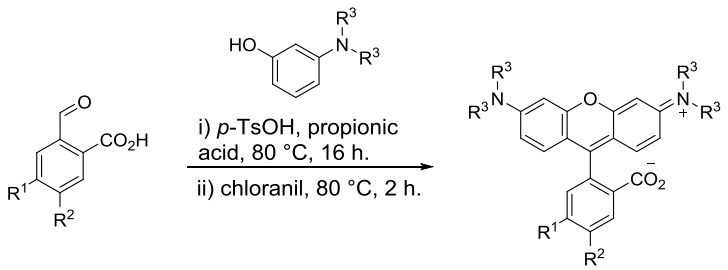
2-Formyl-4-methoxycarbonyl-benzoic acid (4.59): Methyl 3-bromo-1-oxo-3H-isobenzofuran-5-carboxylate **4.58** (1066 mg, 3.95 mmol) was used as starting material and the title compound was obtained as a white solid (800 mg, 3.85 mmol, 97%). R_f = 0.13 (9:1 CH_2Cl_2/MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.34 (1H, br s, CO_2**H**), 8.19 (1H, d, *J* = 7.9 Hz, ArH), 8.14 (1H, s, Ar**H**), 7.96 (1H, d, *J* = 8.0 Hz, Ar**H**), 6.74 (1H, br s, CO**H**), 3.93 (3H, s, C**H**$_3$); ^{13}C NMR (150 MHz, $(CD_3)_2SO$) δ 167.5 (C), 165.2 (C), 147.7 (C), 134.9 (C), 131.3 (CH), 130.4 (C), 125.1 (CH), 124.3 (CH), 98.4 (CH), 52.8 (CH_3); MS (ESI) Exact mass calcd for $C_{10}H_8O_5$ $[M+H]^+$: 209.04, found: 208.93.



2-formyl-5-nitro-benzoic acid: 3-Bromo-6-nitro-3H-isobenzofuran-1-one (143 mg, 0.54 mmol) was used as starting material and the title compound was obtained as a white solid (89 mg, 0.46 mmol, 85%). R_f = 0.11 (9:1 CH_2Cl_2/MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.62 - 8.57 (1H, m, Ar**H**), 8.53 (1H, s, CO_2**H**), 8.50 (1H, s, Ar**H**), 7.96 (1H, d, *J* = 8.4 Hz, Ar**H**), 6.81 (1H, br s, CO**H**); ^{13}C NMR (150 MHz, $(CD_3)_2SO$) δ 166.5 (C), 152.6 (C), 149.5 (C), 129.4 (CH), 128.2 (C), 125.7 (CH), 119.8 (CH), 98.4 (CH).

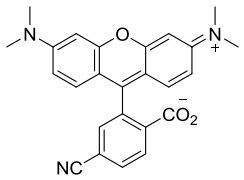
8.4.4 Preparation of Rhodamines

General procedure for preparation of rhodamines:

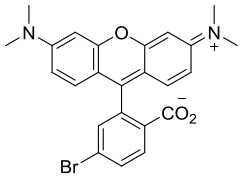


Phthalaldehydic acid (1 eq) was stirred in propionic acid (50 mg/mL) and aminophenol (3 eq) was added, along with *p*-TsOH (0.2 eq). The mixture was heated to 80 °C overnight then chloranil (1 eq) was added, maintaining the temperature at 80 °C and the mixture stirred

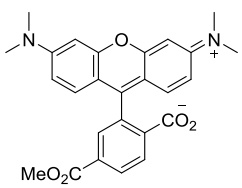
for 2 hours. The mixture was cooled to rt then solvent was removed under reduced pressure and the crude residue purified using flash chromatography, eluting with MeCN/H_2O. Following chromatography, products were dissolved in 10% MeOH/CH_2Cl_2 and filtered through cotton wool to remove any eluted silica.



6-CN-TMR (4.54): 4-Cyano-2-formyl-benzoic acid **4.52** (50 mg, 0.24 mmol) was used as the starting material which was reacted with 3-dimethylaminophenol. The title compound was obtained as a dark purple solid (49 mg, 0.12 mmol, 50%). Eluent for chromatography: gradient from 5% H_2O/MeCN to 20% H_2O/MeCN; Rf = 0.32 (15% H_2O/MeCN); ^{1}H NMR (500 MHz, (CD_3OD) δ 8.22 (1H, d, *J* = 8.1 Hz, Ar**H** 2), 8.05 (1H, d, *J* = 8.8 Hz, Ar**H**), 7.72 (1H, s, Ar**H**), 7.26 (2H, d, *J* = 9.5 Hz, Ar**H**), 7.07 (2H, dd, *J* = 9.5, 1.7 Hz, Ar**H**), 6.95 (2H, d, *J* = 1.5 Hz, Ar**H**), 3.31 (12H, s, NC**H**$_3$); ^{13}C NMR (125 MHz, CD_3OD) δ 171.7 (C), 160.8 (C), 159.3 (C), 159.0 (C), 146.5 (C), 134.8 (CH), 134.7 (C), 134.3 (CH), 132.6 (CH), 131.8 (CH), 119.0 (C), 115.4 (CH), 115.1 (C), 114.4 (C), 97.5 (CH), 41.0 (CH_3); MS (ESI) Exact mass calcd for $C_{25}H_{22}N_3O_3$ $[M+H]^+$: 412.17, found: 412.35.

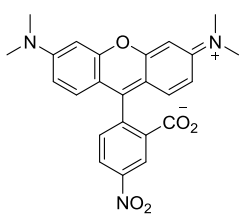


6-Br-TMR (4.61): 4-Bromo-2-formyl-benzoic acid (50 mg, 0.22 mmol) was used as the starting material which was reacted with 3-dimethylaminophenol. The title compound was obtained as a dark purple solid (55 mg, 0.12 mmol, 54%). Eluent for chromatography: gradient from 5% H_2O/MeCN to 15% H_2O/MeCN R_f = 0.32 (15% H_2O/MeCN); ^{1}H NMR (500 MHz, CD_3OD) δ 8.00 (1H, d, *J* = 8.4 Hz, Ar**H**), 7.81 (1H, dd, *J* = 8.4, 2.0 Hz, Ar**H**), 7.43 (1H, d, *J* = 2.0 Hz, Ar**H**), 7.25 (2H, d, *J* = 9.5 Hz, Ar**H**), 7.01 (2H, dd, *J* = 9.5, 2.5 Hz, Ar**H**), 6.88 (2H, d, *J* = 2.4 Hz, Ar**H**), 3.26 (12H, s, C**H**$_3$); ^{13}C NMR (125 MHz, CD_3OD) δ 172.3 (C), 160.6 (C), 159.2 (C), 158.8 (C), 140.8 (C), 136.4 (C), 134.0 (CH), 133.0 (CH), 133.0 (CH), 132.6 (CH), 124.8 (C), 115.2 (CH), 115.0 (C), 97.5 (CH), 41.0 (CH_3); MS (ESI) Exact mass calcd for $C_{24}H_{23}BrN_2O_3$ $[M+H]^+$: 466.09, found: 465.35.

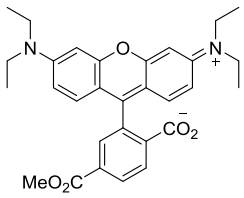


6-CO_2Me-TMR (4.60): 2-Formyl-4-methoxycarbonyl-benzoic acid **4.59** (100 mg, 0.48 mmol) was used as the starting material which was reacted with 3-dimethylaminophenol. The title compound was obtained as a dark purple solid (108 mg, 024. mmol, 51%). Eluent for chromatography: gradient from 5% H_2O/MeCN to 20% H_2O/MeCN R_f = 0.27 (15% H_2O/MeCN); ^{1}H NMR (500 MHz, CD_3OD) δ 8.25 (1H, dd, *J* = 8.2, 1.7 Hz, Ar**H**), 8.14 (1H,

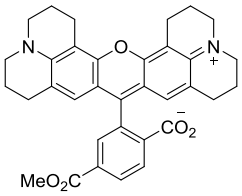
d, *J* = 8.2 Hz, Ar**H**), 7.85 (1H, d, *J* = 1.4 Hz, Ar**H**), 7.23 (2H, d, *J* = 9.5 Hz, Ar**H**), 6.98 (2H, dd, *J* = 9.5, 2.5 Hz, Ar**H**), 6.87 (2H, d, *J* = 2.5 Hz, Ar**H**), 3.90 (1H, s, OC**H**$_3$), 3.25 (12H, s, 3 x NC**H**$_3$); ^{13}C NMR (125 MHz, CD_3OD) δ 172.5 (C), 167.5 (C), 161.9 (C), 159.2 (C), 158.8 (C), 146.3 (C), 133.9 (C), 132.8 (CH), 132.2 (C), 131.9 (CH), 131.6 (CH), 131.2 (CH), 115.2 (CH), 115.1 (C), 97.5 (CH), 53.1 (CH_3), 41.0 (CH_3); MS (ESI) Exact mass calcd for $C_{26}H_{25}N_2O_5$ $[M+H]^+$: 445.18, found: 445.38.



5-NO$_2$-TMR (4.62): 2-formyl-5-nitro-benzoic acid (50 mg, 0.26 mmol) was used as the starting material which was reacted with 3-dimethylaminophenol. The title compound was obtained as a dark purple solid (63 mg, 0.15 mmol, 58%). Eluent for chromatography: gradient from 5% H_2O/MeCN to 15% H_2O/MeCN R_f = 0.38 (15% H_2O/MeCN); ^{1}H NMR (500 MHz, CD_3OD) δ 8.91 (1H, d, *J* = 2.4 Hz, ArH), 8.44 (1H, dd, *J* = 8.3, 2.4 Hz, ArH), 7.53 (1H, d, *J* = 8.4 Hz, ArH), 7.21 (2H, d, *J* = 9.5 Hz, ArH), 7.03 (2H, dd, *J* = 9.5, 2.4 Hz, ArH), 6.94 (2H, d, *J* = 2.4 Hz, ArH), 3.29 (12H, s, CH_3); ^{13}C NMR (125 MHz, CD_3OD) δ 170.5 (C), 161.5 (C), 159.2 (C), 159.0 (C), 150.4 (C), 143.6 (C), 140.1 (C), 132.5 (CH), 132.3 (CH), 126.0 (CH), 125.2 (CH), 115.4 (CH), 114.8 (C), 97.5 (CH), 41.0 (CH_3); MS (ESI) Exact mass calcd for $C_{24}H_{22}N_3O_5$ $[M+H]^+$: 432.16, found: 432.25.

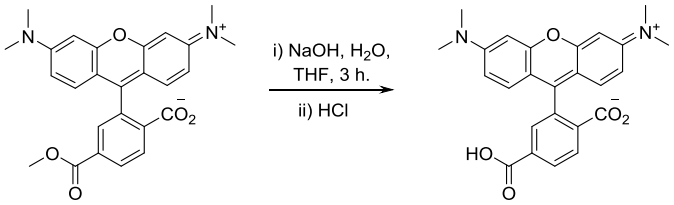


5-CO$_2$Me-Rhodamine-B (4.63): 2-Formyl-4-methoxycarbonyl-benzoic acid **4.59** (50 mg, 0.24 mmol) was used as the starting material which was reacted with 3-diethylaminophenol **4.40**. The title compound was obtained as a dark purple solid (83 mg, 0.17 mmol, 70%). Eluent for chromatography: gradient from 5% H_2O/MeCN to 15% H_2O/MeCN R_f = 0.49 (15% H_2O/MeCN); ^{1}H NMR (500 MHz, CD_3OD) δ 8.24 (1H, dd, *J* = 8.1, 1.7 Hz, Ar**H**), 8.13 (1H, d, *J* = 8.1 Hz, Ar**H**), 7.87 (1H, d, *J* = 1.7 Hz, Ar**H**), 7.23 (2H, d, *J* = 9.5 Hz, Ar**H**), 6.99 (2H, dd, *J* = 9.5, 2.4 Hz, Ar**H**), 6.92 (1H, d, *J* = 2.4 Hz, Ar**H**), 3.90 (3H, s, C**H**$_3$), 3.66 (8H, q, *J* = 7.1 Hz, C**H**$_2$), 1.30 (12H, t, *J* = 7.1 Hz, Ar**H**); ^{13}C NMR (125 MHz, CD_3OD) δ 172.6 (C), 167.5 (C), 161.9 (C), 159.6 (C), 157.1 (C), 146.5 (C), 133.7 (C), 133.1 (CH), 132.1 (C), 131.8 (CH), 131.6 (CH), 131.2 (CH), 115.2 (CH), 115.1 (C), 97.3 (CH), 53.1 (CH_3), 46.9 (CH_2), 13.0 (CH_3); MS (ESI) Exact mass calcd for $C_{24}H_{22}N_3O_5$ $[M+H]^+$: 501.24, found: 501.39.



5-CO_2Me-X-Rhodamine (4.64): 2-Formyl-4-methoxycarbonyl-benzoic acid **4.59** (50 mg, 0.24 mmol) was used as the starting material which was reacted with 8-hydroxyjulolidine **4.39**. Following chromatography, the residue was taken up in CH_2Cl_2 and washed with water instead of filtering through cotton wool. The organics were then dried ($MgSO_4$) and concentrated to dryness to give the title compound as a dark purple solid (62 mg, 0.11 mmol, 47%). Eluent for chromatography: gradient from 5% H_2O/MeCN to 20% H_2O/MeCN. R_f = 0.39 (15% H_2O/MeCN); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.32 (1H, d, *J* = 8.1 Hz, Ar**H**), 8.19 (1H, dd, *J* = 8.1, 1.7 Hz, Ar**H**), 7.74 (1H, d, *J* = 1.4 Hz, Ar**H**), 6.71 (2H, s, Ar**H**), 3.89 (3H, s, C**H**$_3$), 3.42 – 3.31 (8H, m, 4 x C**H**$_2$), 3.00 - 2.94 (4H, m, 2 x C**H**$_2$), 2.73 – 2.53 (4H, m, 2 x C**H**$_2$), 2.09 – 2.01 (4H, m, 2 x C**H**$_2$), 1.95 – 1.84 (4H, m, 2 x C**H**$_2$); ^{13}C NMR (125 MHz, $CDCl_3$); δ 168.7 (C), 166.7 (C), 151.6 (C), 149.6 (C), 144.4 (C), 136.3 (C), 130.3 (C), 130.2 (CH), 129.9 (CH), 129.0 (CH), 127.0 (CH), 121.9 (C), 112.1 (C), 105.0 (C), 52.2 (CH_3), 50.6 (CH_2), 50.1 (CH_2), 27.4 (CH_2), 20.9 (CH_2), 20.2 (CH_2), 20.2 (CH_2); MS (ESI) Exact mass calcd for $C_{34}H_{33}N_2O_5$ $[M+H]^+$: 549.24, found: 549.48.

5-CO_2H-TMR (4.8):



6-CO_2Me-TMR **4.60** (1.0 g, 2.25 mmol) was stirred in H_2O/THF (1:1, 50 mL) and to this was added NaOH (225 mg, 5.63 mmol). The mixture was shielded from light and stirred for 3 hours. The solution was acidified with conc. HCl causing a solid to precipitate, which was filtered and dried to give the title product as a purple solid (863 mg, 2.01 mmol, 89%); ^{1}H NMR (500 MHz, MeOD) δ 8.49 - 8.38 (2H, m, Ar**H**), 8.01 (1H, s, Ar**H**), 7.17 (2H, d, *J* = 9.5 Hz, Ar**H**), 7.08 (2H, dd, *J* = 9.4, 1.7 Hz, Ar**H**), 6.99 (2H, s, Ar**H**), 3.33 (12H, s, 3 x C**H**$_3$); ^{13}C NMR (125 MHz, CD_3OD) δ 167.8 (C), 167.5 (C), 160.5 (C), 159.2 (C), 159.1 (C), 136.2 (C), 136.1 (C), 135.5 (C), 133.0 (CH), 132.5 (CH), 132.5 (CH), 132.1 (CH), 115.8 (CH), 115.1 (C), 97.6 (CH), 41.1 (CH_3); MS (ESI) Exact mass calcd for $C_{25}H_{23}N_2O_5$ $[M+H]^+$: 431.16, found: 431.33.

8.5 Chapter 5 Experimental

8.5.1 Preparation of TOBa on resin

Loading of aldehyde linker onto beads

TentaGel S NH_2 beads (0.26 mmol/g) were weighed into a disposable syringe fitted with a frit and swelled in CH_2Cl_2 then drained. Coupling of linker **5.61** was carried out using 5 equiv. linker (relative to resin), 5 equiv. HATU and 10 equiv. DIPEA in DMF (10 ml/g resin). Beads were shaken in this mixture for 1 hour then washed extensively with DMF and CH_2Cl_2

Reductive amination of amine

Beads were swollen in CH_2Cl_2 then drained. 5 equiv. diamine HCl were dissolved in 2% AcOH in DMF (10 ml/mg resin) and added to the resin. This was shaken for 1 hour then 7 equiv. $Na(AcO)_3BH$ was added and the mixture shaken at room temperature overnight. Beads were washed extensively with DMF, MeOH then CH_2Cl_2.

Chloranil test for secondary amines[270]

A few representative beads were placed in a vial and to this was added a few drops of 2% chloranil in DMF, then a few drops of 2% acetaldehyde in DMF. A deep blue colour indicates the presence of secondary amines, whereas pale yellow beads represent a negative result.

General coupling conditions

Beads were swelled in CH_2Cl_2 then drained. Couplings were carried out using 5 equiv. acid (relative to resin), 4.9 equiv. HATU and 10 equiv. DIPEA in DMF (10 ml/g resin). Beads were shaken in this mixture for 1 hour then washed extensively with DMF and CH_2Cl_2.

Removal of Fmoc protecting group

Beads were swollen in CH_2Cl_2 then drained and 20% Piperidine in DMF was added (10 ml/g resin) and the mixture shaken for 30 minutes, then washed with DMF and CH_2Cl_2.

Removal of Alloc protecting group

Beads were swollen in CH_2Cl_2 then drained and to them was added a freshly made solution of 0.2 equiv. Tetrakis(triphenylphosphine)palladium(0) and 20 equiv. Phenylsilane in CH_2Cl_2. The beads were shaken for 30 minutes then washed with CH_2Cl_2 and the procedure repeated. Washing with CH_2Cl_2 and methanol was carried out.

Removal of Dde protecting group

Dde protecting group was carried out according to Díaz-Mochón and co-workers[192]. Briefly, a solution of 1.25 g $NH_2OH.HCl$ and 0.918 g imidazole were suspended in 5 mL NMP. Sonication was carried out until the solids dissolved completely. The required amount of solution was diluted 5:1 with CH_2Cl_2 then added to the beads for 3 hours.

Non-orthogonal removal of Dde protecting group

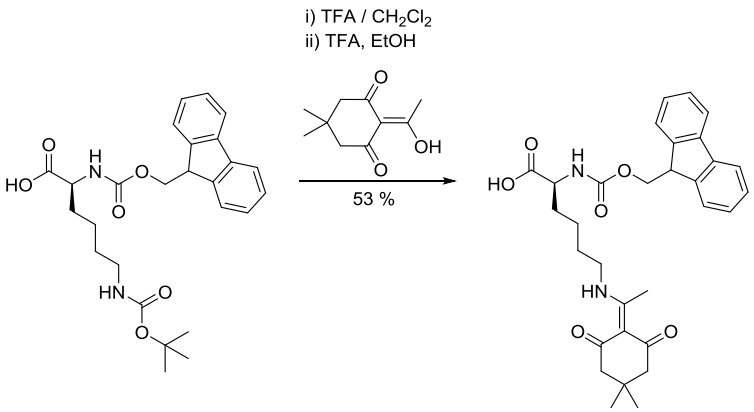
(This procedure was used only when Dde was the final protecting group left on the resin as it will remove Fmoc and reduce Alloc groups) Bead were swollen in CH_2Cl_2 then drained. A solution of 2% hydrazine hydrate in DMF was added to the beads and shaken for 3 minutes. This was repeated 3 times then the beads were washed with DMF and CH_2Cl_2.

Cleavage from bead

Beads were dried then subjected to TFA 2 hours with shaking. If groups with acid labile protecting groups e.g. ^tBu were used during synthesis, then 2.5% H_2O and 2.5% TIS was added to the mixture.

8.5.2 Preparation of TOBa in Solution

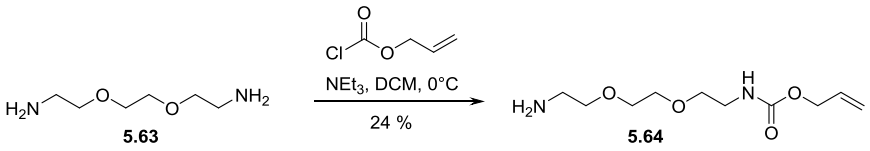
Fmoc-Lys(Dde)-OH



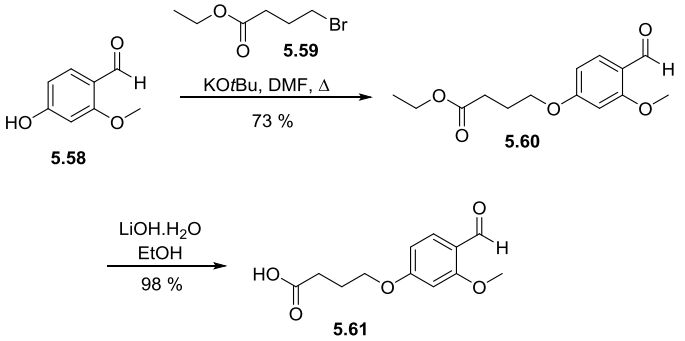
Fmoc-Lys(Boc)-OH (1.00 g, 2.14 mmol) was stirred in CH_2Cl_2 (15 mL) and to this was added 15 mL TFA. The mixture was stirred at rt for 1 hour, until TLC analysis showed consumption of the starting material. Volatiles were removed under reduced pressure to give an orange oil. This was taken up in 30 mL ethanol and to this was added Dde-OH (2.27 g, 4.27 mmol) and TFA (16 µL, 0.21 mmol). The solution was heated to reflux for 48 hours. Solvents were removed under partial pressure and the residue taken up in EtOAc and washed with 1M $KHSO_4$ twice. The organics were run off and dried over $MgSO_4$ then concentrated to dryness. The residue was triturated with CyHex then the remaining sticky oil was taken up in a minimum amount of EtOAc and added dropwise to a stirring beaker of petroleum ether

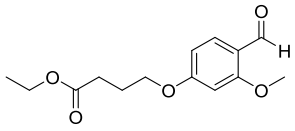
to crash out an off white solid, which was filtered off and dried to give the title compound (0.61 g, 1.13 mmol, 53%). R_f = 0.26 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $CDCl_3$) δ 13.33 (1H, br s, N$^{\varepsilon}$**H**), 7.76 (2H, d, *J* = 7.6 Hz, Fmoc Ar**H**), 7.60 (2H, t, *J* = 8.2 Hz, Fmoc Ar**H**), 7.39 (2H, t, *J* =7.4 Hz, Fmoc Ar**H**), 7.29 (2H, m, Fmoc Ar**H**), 5.74 (1H, d, *J* =7.8 Hz, N$^{\alpha}$**H**), 4.47 (1H, m, C$^{\alpha}$**H**), 4.38 (2H, d, *J* =7.0 Hz, Fmoc C**H**$_2$), 4.21 (1H, m, Fmoc C**H**), 3.43 (2H, d, *J* =4.6 Hz, C$^{\varepsilon}$**H**$_2$), 2.56 (3H, s, C=CC**H**$_3$), 2.40 (4H, s, Dde 2 x C**H**$_2$), 1.99 (1H, m, C^{ß}**H**$_2$), 1.82 (1H, m, C$^{\beta}$**H**$_2$), 1.74 (2H, m, C$^{\delta}$**H**$_2$), 1.54 (2H, m, C$^{\gamma}$**H**$_2$), 1.02 (6H, s, Dde (C**H**$_3$)$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 198.2 (C), 174.3 (C), 174.2 (C), 156.1 (C), 143.9 (C), 143.7 (C), 141.3 (C), 127.7 (CH), 127.1 (CH), 125.1 (CH), 120.0 (CH), 107.8 (C), 67.1 (CH_2), 53.4 CH, 52.2 (CH_2), 47.2 CH, 43.4 (CH_2), 31.9 (CH_2), 30.2 (C), 28.2 (CH_2), 28.2 (CH_3), 22.3 (CH_2), 18.3 (CH_3); Exact mass calcd for $C_{31}H_{37}N_2O_6$ $[M+H]^+$: 533.26, found: 533.35. Data consistent with the literature.[271]

Allyl *N*-[2-[2-(2-aminoethoxy)ethoxy]ethyl]carbamate (5.64):

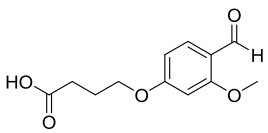


The allyl protected amine was prepared according to the literature[272] with some slight modifications. 2-[2-(2-Aminoethoxy)ethoxy]ethanamine **5.63** (13.3 g, 90 mmol) was stirred in 200 mL CH_2Cl_2 at 0ºC and to this was added dropwise a solution of allylchloroformate (0.96 mL, 9.0 mmol) in 100 mL CH_2Cl_2. The mixture was left in ice and allowed to stir overnight then volatiles were removed in vacuo until ~100 mL was left. This was washed with water (3 x 30 mL) then the organic layers dried ($MgSO_4$) and concentrated to dryness and the residue purified by flash chromatography (10:1:0.1 CH_2Cl_2/MeOH/NEt_3) to give the amine as a light yellow oil (500 mg, 24%). R_f = 0.45 (CH_2Cl_2/MeOH/$NEt_3$10:1:0.1); ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.96 – 5.88 (1H, m, C**H**=CH_2), 5.55 (1H, br s, N**H**), 5.32 – 5.19 (2H, m, CH=C**H**$_2$), 4.56 (2H, d, *J* = 5.3 Hz, COOC**H**$_2$), 3.63 – 3.60 (4H, m, 2 x C**H**$_2$O,), 3.58 – 3.53 (4H, m, 2 x C**H**$_2$O), 3.38 (2H, m, C**H**$_2$NH), 2.90 (2H, t, *J* = 5.0 Hz, C**H**$_2$NH$_2$), 2.37 (2H, s, N**H**$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 156.4 (C), 133.0 (CH), 117.5 (CH_2), 72.7 (CH_2), 70.3 (CH_2), 70.1 (CH_2), 70.0 (CH_2), 65.5 (CH_2), 41.4 (CH_2), 40.8 (CH_2); Exact mass calcd for $C_{10}H_{21}N_2O_4$ $[M+H]^+$: 233.15, found: 233.12.



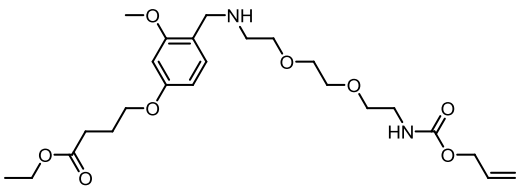


Ethyl 4-(4-formyl-3-methoxy-phenoxy)butanoate (5.60): 4-Hydroxy-2-methoxybenzaldehyde **5.58** (100 mg, 0.66 mmol) was stirred in 5 mL dry DMF and to this was added Potassium tert-butoxide (81 mg, 0.72 mmol). Ethyl-4-bromobutyrate (103 µl, 0.72 mmol) was then added and the mixture stirred at 60 °C overnight. DMF was removed under partial pressure and residue taken up in CH_2Cl_2 and washed with water. The organics were dried and concentrated to dryness and the residue purified by flash chromatography (30% EtOAC/CyHex). The product was afforded as white crystals (129 mg, 0.48 mmol, 73%). R_f = 0.30 (70% CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.28 (1H, s, C**H**O), 7.79 (1H, d, *J* = 8.7 Hz, Ar**H**), 6.52 (1H, dd, *J* = 8.7, 2.1 Hz, Ar**H**), 6.45 (1H, d, *J* = 2.1 Hz, Ar**H**), 4.15 (2H, q, 7.1 Hz, COOC**H**$_2$), 4.08 (2H, t, *J* = 6.2 Hz, C**H**$_2$CH$_2$CH$_2$), 3.90 (3H, s, ArOC**H**$_3$), 2.52 (2H, t, *J* = 7.3 Hz, CH$_2$CH$_2$C**H**$_2$), 2.13 (2H, tt, *J* = 7.1, 6.2 Hz, CH$_2$C**H**$_2$CH$_2$), 1.26 (3H, t, *J* = 7.1 Hz, COOHCH$_2$C**H**$_3$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 188.4 (C), 173.0 (C), 165.4 (C), 163.6 (C), 130.8 (CH), 119.1 (C), 106.3 (CH), 98.3 (CH), 67.2 (CH_2), 60.6 (CH_2), 55.6 (CH_3), 30.6 (CH_2), 24.4 (CH_2), 14.2 (CH_3). Exact mass calcd for $C_{14}H_{19}O_5$ $[M+H]^+$: 267.12, found: 267.00.

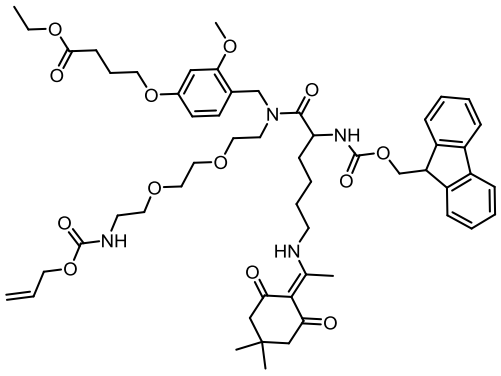


4-(4-Formyl-3-methoxy-phenoxy)butanoic acid (5.61): Ethyl 4-(4-formyl-3-methoxy-phenoxy)butanoate **5.60** (120 mg, 0.45 mmol) was stirred in 20 mL 1M sodium hydroxide and 20 mL ethanol overnight. Ethanol was removed under partial pressure and the aqueous extracted with EtOAc then acidified to pH 4 with 1M HCl. A white solid crashed out which was extracted with EtOAc. The organics were dried and concentrated to dryness to give a white solid (105 mg, 0.44 mmol, 98%). R_f 0.38 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.28 (1H, s, C**H**O), 7.79 (2H, d, *J* = 8.7 Hz, Ar**H**), 6.52 (1H, dd, *J* = 8.7, 2.1 Hz, Ar**H**), 6.45 (1H, d, *J* = 2.1 Hz, Ar**H**), 4.11 (2H, t, *J* = 6.2 Hz, C**H**$_2$CH$_2$CH$_2$), 3.90 (3H, s, ArOC**H**$_3$), 2.60 (2H, t, *J* = 7.1 Hz, CH$_2$CH$_2$C**H**$_2$), 2.15 (2H, tt, *J* = 7.1, 6.2 Hz, CH$_2$C**H**$_2$CH$_2$). ^{13}C NMR (125 MHz, $CDCl_3$) δ 188.5 (C), 178.0 (C), 165.4 (C), 163.7 (C),

130.8 (CH), 119.1 (C), 106.2 (CH), 98.3 (CH), 66.9 (CH_2), 55.6 (CH_3), 30.1 (CH_2), 24.2 (CH_2); Exact mass calcd for $C_{12}H_{15}O_5$ $[M+H]^+$: 239.09, found: 239.03. The data is in agreement with the literature[273]

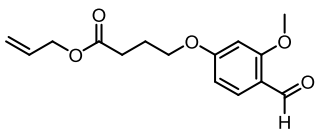


Ethyl-4-[4-[[2-[2-[2-(allyloxycarbonylamino)ethoxy]ethoxy]ethylamino]methyl]-3-methoxy-phenoxy]butanoate (5.65): Allyl *N*-[2-[2-(2-aminoethoxy)ethoxy]ethyl]carbamate **5.64** (139 mg, 0.60 mmol) and ethyl 4-(4-formyl-3-methoxy-phenoxy)butanoate **5.60** (100 mg, 0.38 mmol) were stirred in DCE (7.5 mL) for 10 minutes. To this was added $NaBH(AcO)_3$ (126 mg, 0.59 mmol) and the mixture stirred for 3 hours. Water (15 mL) was added and the mixture stirred for a further 5 minutes 10 mL CH_2Cl_2 was added to the organic layer which was washed with water, dried over $MgSO_4$ and concentrated to dryness. The residue was purified using flash chromatography (10:1 CH_2Cl_2/MeOH) to give a colourless oil (115 mg, 0.24 mmol, 63%). R_f = 0.40 (10:1 CH_2Cl_2/MeOH); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.14 (1H, d, *J* = 8.1 Hz, Ar**H**), 6.45 (1H, d, *J* = 2.2 Hz, Ar**H**), 6.42 (1H, d, *J* = 8.1, 2.2 Hz, Ar**H**), 5.9 (1H, ddt, *J* = 16.7, 10.9, 5.5 Hz, C**H**=CH_2), 5.80 (1H, br s, N**H**), 5.23 (2H, m, CH=C**H**$_2$), 4.54 (2H, d, *J* = 5.5 Hz, C**H**$_2$CH=CH_2), 4.15 (2H, q, *J* = 5.0 Hz, C**H**$_2$$CH_3$), 4.0 (2H, t, *J* = 6.1 Hz, C=OC**H**$_2$$CH_2CH_2$), 3.87 (2H, s, NHC**H**$_2$Ar), 3.82 (3H, s, OC**H**$_3$), 3.66 (2H, t, *J* = 5.2 Hz, C**H**$_2$O), 3.60 (4H, s, 4 x OC**H**$_2$), 3.55 (2H, t, *J* = 5.0 Hz, C**H**$_2$O), 3.34 (2H, q, *J* = 5.3 Hz, C**H**$_2$NHC=O), 2.86 (2H, t, *J* = 5.2 Hz, C**H**$_2$NHCH$_2$Ar), 2.51 (2H, t, *J* = 7.3 Hz, C=OCH_2CH_2C**H**$_2$), 2.10 (2H, quin, *J* = 6.7 Hz, C=OCH_2C**H**$_2$$CH_2$), 1.99 (1H, s, CH_2N**H**CH_2); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.2 (C), 160.1 (C), 158.7 (C), 156.5 (C), 133.0 (CH), 131.2 (CH), 117.5 (CH_2), 117.3 (C), 104.8 (CH), 99.0 (CH), 70.2 (CH_2), 68.9 (CH_2), 66.8 (CH_2), 65.4 (CH_2), 60.4 (CH_2), 55.3 (CH_3), 47.7 (CH_2), 47.1 (CH_2), 40.8 (CH_2), 30.7 (CH_2), 24.6 (CH_3), 14.20 (CH_2); MS (ESI) Exact mass calcd for $C_{24}H_{39}N_2O_8$ $[M+H]^+$: 483.27, found: 483.25.

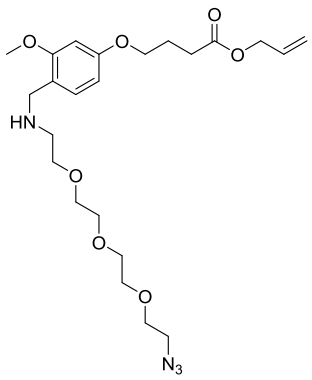


Ethyl TOBa 1 (5.66): Amine **5.65** (100 mg, 0.21 mmol) and Fmoc-Lys(Dde)-OH (108 mg, 0.25 mmol) were stirred in 5 mL CH_2Cl_2 and to this was added HATU (95 mg, 0.25 mmol) then DIPEA (72 µL, 0.41 mmol). The mixture was stirred for 2 hours at rt then the organic layers washed with water twice, dried over $MgSO_4$ then concentrated to dryness. The resulting yellow oil was purified using silica gel chromatography (EtOAc) to give a colourless oil

(130 mg, 0.13 mmol, 63%). R_f = 0.29 (EtOAc); Exact mass calcd for $C_{55}H_{73}N_4O_{13}$ $[M+H]^+$: 997.52, found: 997.60. Purity 81% by µHPLC (220 nm).

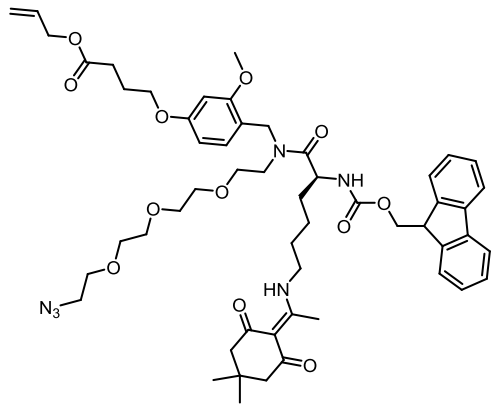


Allyl 4-(4-formyl-3-methoxy-phenoxy)butanoate: 4-(4-formyl-3-methoxy-phenoxy)butanoic acid (**5.61**) (500 mg, 2.12 mmol) was stirred in CH_2Cl_2 (25 mL) and to this was added allyl alcohol (180 µL, 2.65 mmol), EDC.HCl (508 mg, 2.65 mmol) and finally DMAP (13 mg, 0.11 mmol). The mixture was stirred at rt under N_2 for 4 hours, then water was added and the organic layers ran off and dried over $MgSO_4$ and concentrated to dryness. The residue was purified using flash chromatography (20% EtOAc/CyHex) to give the allyl ester as a colourless oil, which solidified upon standing to give a white solid (394 mg, 1.41 mmol, 67%). R_f = 0.45 (20% EtOAc/CyHex); ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.29 (1H, s, C**H**=O), 7.80 (1H, d, *J* = 8.6 Hz, Ar**H**), 6.53 (1H, dd, *J* = 8.6, 2.1 Hz, Ar**H**), 6.45 (1H, d, J = 2.1 Hz, Ar**H**), 5.93 (1H, ddt, *J* = 16.9, 10.8, 5.8 Hz, CH_2=C**H**), 5.33 (1H, dd, J = 16.9, 1.5 Hz, C**H**$_2$=CH trans), 5.25 (1H, dd, *J* = 10.8, 1.1 Hz, C**H**$_2$=CH cis), 4.61 (2H, d, *J* = 5.8 Hz, CH_2=CHC**H**$_2$), 4.10 (2H, d, *J* = 6.1 Hz, C=OC**H**$_2$), 3.91 (3H, s, OC**H**$_3$), 2.52 (2H, t, *J* = 7.2 Hz, C=OCH_2CH_2C**H**$_2$), 2.16 (2H, quin, *J* = 6.7 Hz, C=OCH_2C**H**$_2CH_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 188.6 (CH), 172.8 (C), 165.6 (C), 163.9 (C), 132.3 (CH), 131.0 (CH), 119.3 (C), 118.7 (CH_2), 106.5 (CH), 98.6 (CH), 67.3 (CH_2), 65.5 (CH_2), 55.9 (CH_3), 30.7 (CH_2), 24.6 (CH_2); MS (ESI) Exact mass calcd for $C_{15}H_{19}O_5$ $[M+H]^+$: 279.12, found: 279.00.

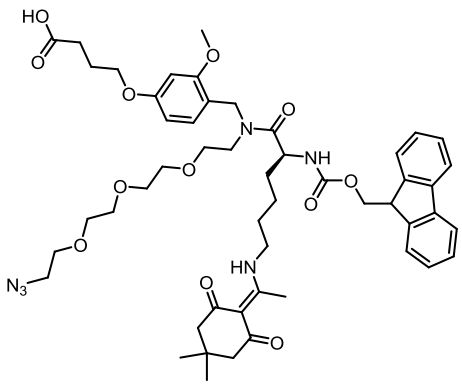


Allyl 4-[4-[[2-[2-[2-(2-azidoethoxy)ethoxy]ethoxy]ethylamino]methyl]-3-methoxy -phenoxy] butanoate (5.69): Amine **5.67** (310 µL, 1.56 mmol) and aldehyde **5.68** (290 mg, 1.04 mmol) were stirred in 15 mL DCE at room temperature. To this was added sodium triacetoxyborohydride (330 mg, 1.56 mmol) in one portion. The mixture was stirred overnight then diluted with CH_2Cl_2. The organics were washed with water three times then dried over $MgSO_4$ and concentrated to dryness. The residue was purified using silica gel chromatography (10% MeOH/CH_2Cl_2) to give the title compound as a yellow oil (241 mg, 0.50 mmol, 32%). R_f = 0.40 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.28 (1H, m, Ar**H**), 6.46 (2H, m, Ar**H**), 5.92 (1H, ddt, *J* = 17.2, 10.4, 5.7 Hz, CH_2=C**H**), 5.28 (2H, m, C**H**$_2$=CH), 4.60 (2H, dt, *J* = 5.8, 1.3 Hz, CH_2=CHC**H**$_2$), 4.10 (2H, s, ArC**H**$_2$), 4.00 (2H, t, *J* = 6.1 Hz, C=OC**H**$_2$), 3.86 (3H, s, OC**H**$_3$), 3.81 (2H, m, N_3C**H**$_2$), 3.65 (10H, d, *J* = 2.6 Hz, 5 x OC**H**$_2$), 3.38 (2H, m,

$NHCH_2C\mathbf{H}_2$), 3.02 (2H, m, $NHC\mathbf{H}_2$), 2.55 (2H, t, *J* = 7.3 Hz, $C{=}OCH_2CH_2C\mathbf{H}_2$), 2.12 (2H, m, $C{=}OCH_2C\mathbf{H}_2CH_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.7 (C), 161.2 (C), 158.8 (C), 132.5 (CH), 132.1 (CH), 118.3 (CH_2), 112.3 (C), 105.3 (CH), 99.0 (CH), 70.5 (CH_2), 70.5 (CH_2), 70.3 (CH_2), 70.2 (CH_2), 70.0 (CH_2), 66.9 (CH_2), 66.2 (CH_2), 65.2 (CH_2), 55.5 (CH_3), 50.7 (CH_2), 47.1 (CH_2), 46.1 (CH_2), 30.6 (CH_2), 24.5 (CH_2); MS (ESI) Exact mass calcd for $C_{23}H_{37}N_4O_7$ $[M+H]^+$: 481.27, found: 481.13.

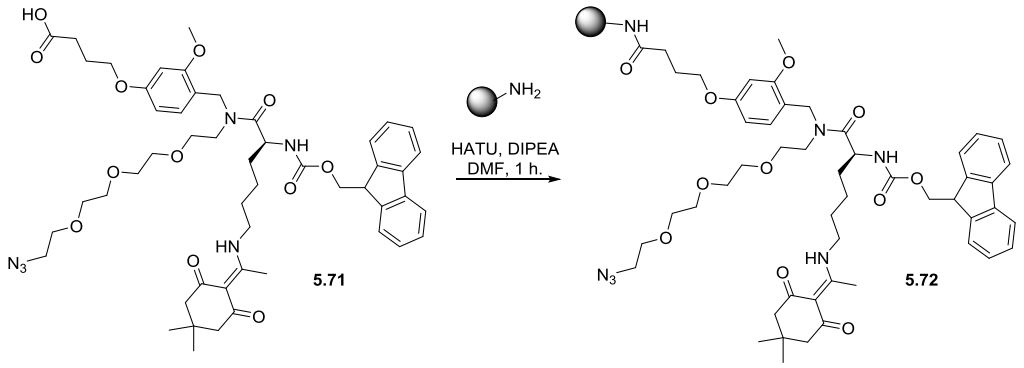


Allyl-TOBa (5.70): Amine **5.69** (215 mg, 0.45 mmol) was stirred in 5 mL CH_2Cl_2 and to this was added Fmoc-lys(Dde)-OH (286 mg, 0.54 mmol), HATU (205 mg, 0.54 mmol) then DIPEA (156 µL, 0.90 mmol). The mixture was stirred at rt for 16 hours then diluted with more CH_2Cl_2 and the organic layers washed with water. The organics were run off and dried ($MgSO_4$) then concentrated to dryness. The residue was purified using flash chromatography (7:3 EtOAc/CyHex, then EtOAc) to give a colourless oil (355 mg, 0.35 mmol, 79%). R_f = 0.27 (EtOAc); MS (ESI) Exact mass calcd for $C_{54}H_{71}N_6O_{12}$ $[M+H]^+$: 995.51, found: 995.61. Purity 60% by µHPLC (220 nm).



TOBa(N_3) (5.71): Allyl ester **5.70** (150 mg, 0.15 mmol) was stirred in 2 mL CH_2Cl_2 at 0 °C for 10 minutes. To this was added phenylsilane (9.2 µL, 0.075 mmol) then $Pd(PPh_3)_4$ (17 mg, 0.015 mmol) and the mixture was stirred under N_2 and allowed to warm to room temperature for 20 minutes. Next, CH_2Cl_2 and H_2O (2 x 10 mL) were added to the solution. The organics were run off and washed twice more with H_2O before being dried ($MgSO_4$) and concentrated to dryness. The residue was purified using flash chromatography (10% MeOH/CH_2Cl_2) to give the title compound as a yellow oil (115 mg, 0.12 mmol, 80%). R_f = 0.22 (10% MeOH/CH_2Cl_2); MS (ESI) Exact mass calcd for $C_{51}H_{67}N_6O_{12}$ $[M+H]^+$: 955.48, found: 955.58. Purity 79% by µHPLC (220 nm).

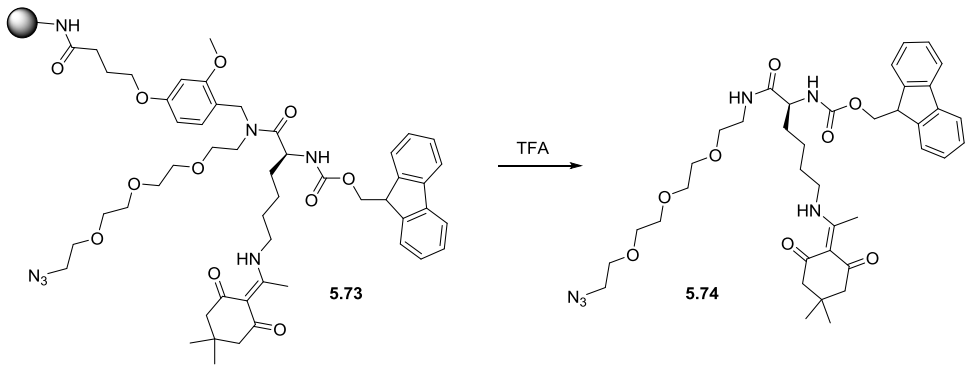
TOBa resin (5.72):



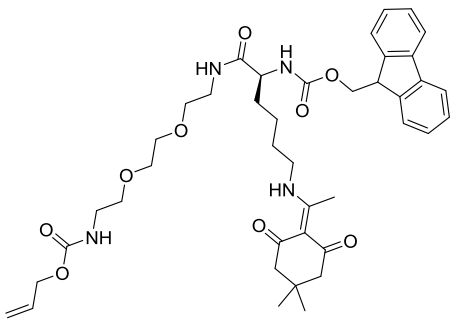
TentaGel S NH_2 beads (0.26 mmol/g) were weighed into a disposable syringe fitted with a frit and swelled in CH_2Cl_2 then drained. Coupling of TOBa compound **5.71** was carried out using 5 equiv. TOBa **5.71** (relative to resin), 5 equiv. HATU and 10 equiv. DIPEA in DMF (10 ml/g resin). Beads were shaken in this mixture for 1 hour then washed extensively with DMF and CH_2Cl_2 and dried. A TNBS test indicated complete reaction of amines on resin.

8.5.3 TOBa Analysis

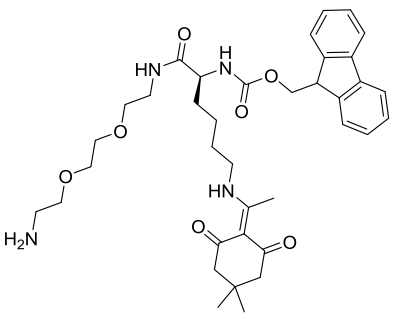
TOBa-N_3 cleaved from resin (5.74):



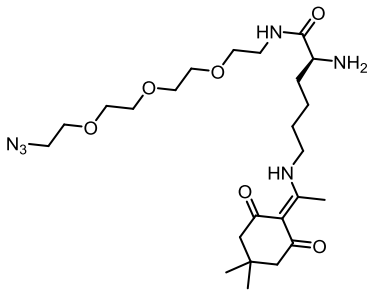
TOBa resin **5.73** was dried then subjected to neat TFA for 2 hours. The liquids were filtered off and collected and the resin washed with CH_2Cl_2, which was also collected. The filtrates were pooled and concentrated to dryness to give the triply protected backbone **5.74**. Exact mass calcd for $C_{39}H_{53}N_6O_8$ $[M+H]^+$: 733.39, found: 733.46. Purity 96% by HPLC (245 nm).



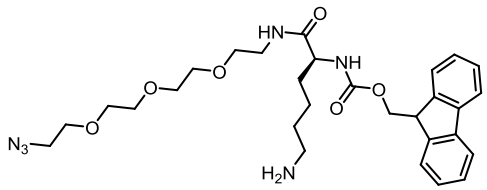
TOBa(Alloc) cleaved from resin: Alloc protected TOBa resin **5.78** was dried, then subjected to neat TFA for 2 hours. The liquids were filtered off and collected and the resin washed with CH_2Cl_2, which was also collected. The filtrates were pooled and concentrated to dryness to give Alloc bearing triply protected backbone. Exact mass calcd for $C_{41}H_{55}N_4O_9$ $[M+H]^+$: 747.40, found: 747.50. Purity 88% by µHPLC (220 nm).



TOBa – Alloc removed: Alloc protected TOBa resin **5.78** was subjected to Alloc removal conditions as outlined in general procedures. The resin was washed, dried and then subjected to neat TFA for 2 hours. The liquids were filtered off and collected and the resin washed with CH_2Cl_2, which was also collected. The filtrates were pooled and concentrated to dryness to give the doubly protected backbone. Exact mass calcd for $C_{37}H_{51}N_4O_7$ $[M+H]^+$: 663.38, found: 663.51. Purity 80% by µHPLC (220 nm).



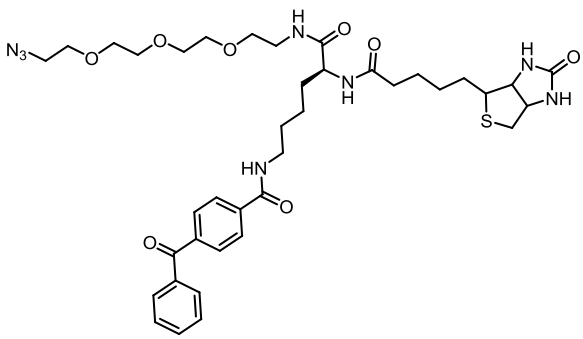
TOBa-N_3-NH_2-Dde: A sample of TOBa resin **5.73** was shaken in 20% Piperidine/DMF for 30 minutes. The solution was drained and the resin washed extensively. The resin was dried then subjected to neat TFA for 2 hours. The liquids were filtered off and collected and the resin washed with CH_2Cl_2, which was also collected. The filtrates were pooled and concentrated to dryness to give the title compound as a white solid. MS (ESI) Exact mass calcd for $C_{24}H_{43}N_6O_6$ $[M+H]^+$ 511.32, found: 511.40. Purity 93% by HPLC (245 nm).



TOBa-N_3- Fmoc-NH_2: a solution of 1.25 g $NH_2OH.HCl$ and 0.918 g imidazole were suspended in 5 mL NMP. Sonication was carried out until the solids dissolved completely. The required amount of solution was diluted 5:1 with CH_2Cl_2 then added to the TOBa resin **5.73** for 3 hours then drained and the resin washed extensively. The resin was dried then subjected to neat TFA for 2 hours. The liquids were filtered off and collected and the resin washed with CH_2Cl_2, which was also collected. The filtrates were pooled and concentrated to dryness to give the title compound as a white solid. MS (ESI) Exact mass calcd for $C_{29}H_{41}N_6O_6$ $[M+H]^+$ 569.31, found: 569.38. Purity 95% by HPLC (245 nm).

8.5.4 TOBa Tags

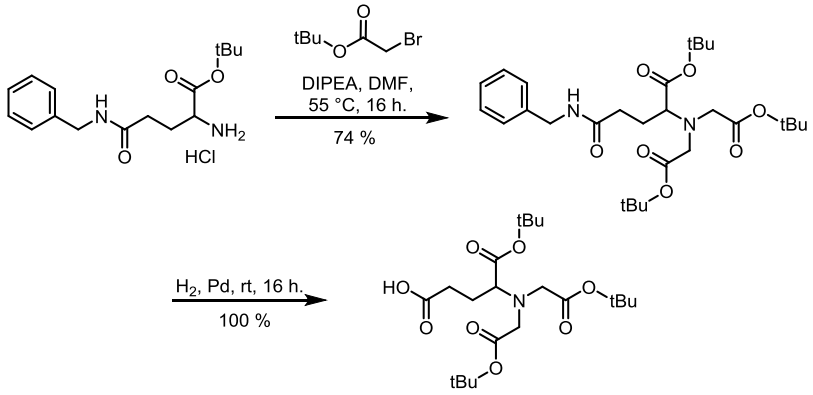
8.5.4.1 N_3, Biotin, Benzophenone

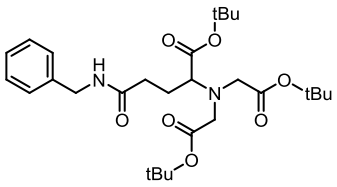


TOBa- N_3, Biotin, Benzophenone: TOBa resin **5.73** (100 mg, 0.026 mmol) was swelled in CH_2Cl_2 and then the Fmoc group was removed using conditions outlined above. Biotin (31.7 mg, 0.13 mmol) was coupled using standard conditions, then Dde removed using the non-orthogonal standard method. Benzophenone carboxylic acid (29.4 mg, 0.13 mmol) was coupled and the compound cleaved from the resin using TFA. The residue was taken up in H_2O and lyophilised to give the title compound as a white solid (15 mg, 0.019 mmol). MS (ESI) Exact mass calcd for $C_{38}H_{53}N_8O_8S$ $[M+H]^+$ 781.37, found: 781.30. Purity 74% by HPLC (220 nm).

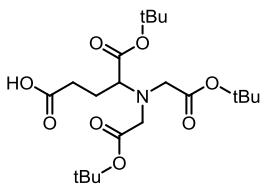
8.5.4.2 N_3, NTA,TMR

NTA Building Block



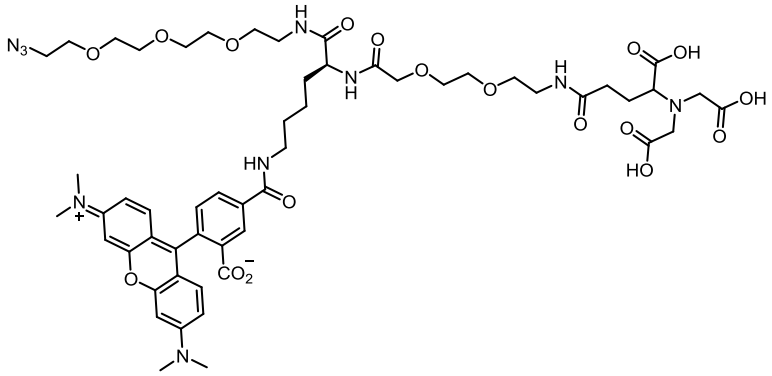


(^{t}Bu)$_2$-Glu(OBzl)-O^{t}Bu (5.83): H-Glu(OBzl)-OtBu.HCl **5.82** (2.00 g, 6.06 mmol) was stirred in 50 mL DMF and to this was added *tert*-butyl bromoacetate (2.70 g, 8.18 mmol) and DIPEA (4.20 mL, 24.24 mmol) slowly. The mixture was heated to 55 °C and left to stir overnight. Solvents were removed under vacuo and the residue taken up in EtOAc. A solid was precipitated which was filtered off. The filtrates were reduced to dryness then purified using flash chromatography (20:1 Cyhex/EtOAc) to give a pale yellow oil (2.30 g, 4.47 mmol, 74%). R_f = 0.27 (20:1 Cyhex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.27 – 7.4 (m, 5H, Ar**H**), 5.14 (s, 2H, ArC**H**$_2$), 3.46 (s, 4H, NC**H**$_2$), 3.41 (q, *J* = 5.5 Hz, 1H, NC**H**), 2.58 – 2.77 (m, 2H, NCHCH$_2$C**H**$_2$), 1.89 – 2.08 (m, 2H, NCHC**H**$_2$CH$_2$), 1.48 (s, 9H, CC(O)OC(C**H**$_3$)$_3$), 1.45 (s, 18H N(CH$_2$C(O)OC(C**H**$_3$)$_3$)$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.7 (C), 172.0 (C), 170.7 (C), 136.4 (C), 128.6 (CH), 128.3 (CH), 128.2 (CH), 81.5 (C), 80.9 (C), 66.3 (CH_2), 64.5 (CH), 54.0 (CH_2), 30.7 (CH_2), 28.4 (CH_3), 28.3 (CH_3), 25.4 (CH_2); MS (ESI) Exact mass calcd for $C_{28}H_{44}NO_8$ $[M+H]^+$: 522.48, found: 522.18. Data consistent with the literature.[274]



(^{t}Bu)$_2$-Glu-O^{t}Bu (5.84): (tBu)$_2$-Glu(OBzl)-OtBu **5.83** (3.25 g, 6.24 mmol) was stirred in 50 mL dry methanol under and argon atmosphere. To this was added 10%palladium on charcoal (500 mg). The flask was fitted with a septum and evacuated, then the atmosphere replaced with hydrogen and the mixture was stirred overnight. Filtration through celite and removal of solvent gave the title compound as a colourless oil (2.74 g, 6.33 mmol, 100%). R_f = 0.43 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.45 (s, 4H, N(C**H**$_2$CO)$_2$), 3.37 (dd, J = 5.0, 10.6 Hz, 1H, NC**H**), 2.73 – 2.59 (m, 2H, C**H**$_2$COOH),

2.05 – 1.87 (m, 2H, NCHC**H**$_2$), 1.46 (s, 8H, NCHCOO(C**H**$_3$)$_3$), 1.45 (s, 18H, NCH$_2$COO(C**H**$_3$)$_3$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 177.5 (C), 171.5 (C), 170.6 (C), 81.6 (C), 81.1 (C), 64.6 (CH), 54.0 (CH_2), 30.9 (CH_2), 28.2 (CH_3), 28.1 (CH_3), 25.3 (CH_2); MS (ESI) Exact mass calcd for $C_{21}H_{47}N_2O_{10}$ $[M+H]^+$: 432.69, found: 432.05. Data consistent with the literature.[274]



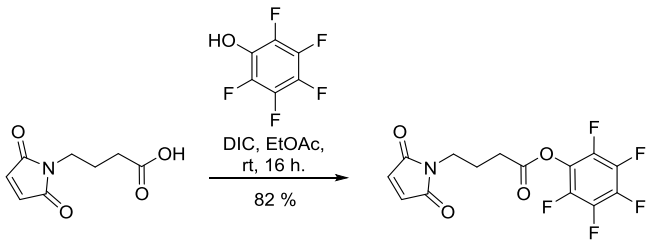
TOBa- N_3, TMR, PEG-NTA (5.87): TOBa resin **5.73** (100 mg, 0.026 mmol) was swelled in CH_2Cl_2 and then the Fmoc group removed using conditions outlined above. Fmoc-DOA-OH (50.1 mg, 0.13 mmol) was coupled using standard conditions, then (^{t}Bu)$_2$-Glu-O^{t}Bu (56.2 mg, 0.13 mol). Dde was removed using the non-orthogonal standard method. 5-TAMRA (55.9 mg, 0.13 mmol) was coupled and the compound cleaved from the resin using TFA. The residue was taken up in H_2O and lyophilised to give the title compound as a purple solid. MS (ESI) Exact mass calcd for $C_{54}H_{73}N_{10}O_{18}$ $[M+H]^+$ 1149.51, found: 1149.16. Purity 83% by µHPLC (220 nm).

8.5.4.3 DBCO Resin

TentaGel NH_2 resin was coupled with DBCO-acid (Click Chemistry Tools) using standard conditions (see section 7.4.1).

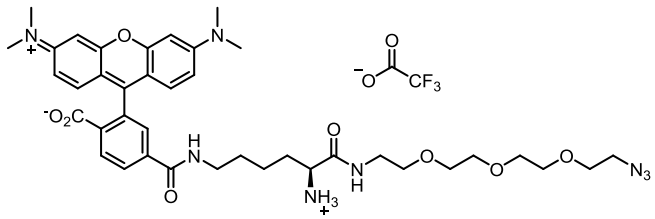
8.5.4.4 N_3, TMR, Maleimide:

Pentafluorophenyl 4-(2,5-dioxopyrrol-1-yl)butanoate (5.89):

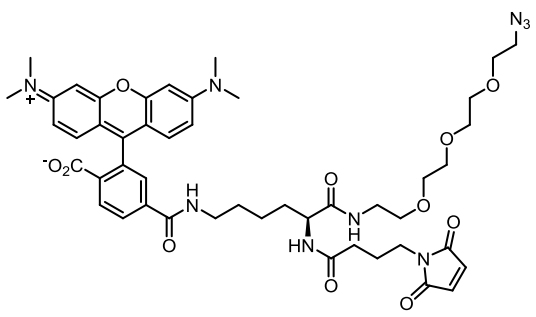


N-(3-Carboxypropyl)maleimide (100 mg, 0.55 mmol, ChemImpex) and pentafluorophenol (110 g, 0.60 mmol) were dissolved in 1.5 mL EtOAc and to this was added DIC (93 µL, 0.6 mmol). The mixture was stirred for 16 hours then solvents removed under reduced

pressure and the residue purified using flash chromatography (7:3 CyHex/EtOAc) to give the title compound as a white solid (158 mg, 0.45 mmol, 82%). R_f = 0.6 (7:3 CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.74 (2H, s, C**H**=C**H**), 3.67 (2H, t, *J* = 6.8 Hz, NC**H**$_2$CH$_2$CH$_2$), 2.71 (2H, t, *J* = 7.4 Hz, NCH$_2$CH$_2$C**H**$_2$), 2.08 (2H, t, *J* = 7.1 Hz, NCH$_2$C**H**$_2$CH$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.7 (C), 168.6 (C), 142.1 (m, CF), 140.5 (m, CF), 140.1 (m, CF), 138.9 (m, CF), 138.5 (m, CF), 136.9 (m, CF), 134.19 (CH), 36.7 (CH_2), 30.6 (CH_2), 23.7 (CH_2). ^{19}F NMR (376.5 MHz, $CDCl_3$) δ -152.6 (2F, m), -157.9 (1F, m), -162.3 (2F, t, *J* = 21.6).



TOBa-N_3, NH_2.TFA, TMR (5.88): TOBa resin 4.73 (100 mg, 0.026 mmol) was swelled in CH_2Cl_2 and then the Fmoc group removed using conditions outlined above. 6-TAMRA (55.9 mg, 0.13 mmol) was coupled using standard conditions, then Dde was removed using the non-orthogonal standard method. The compound cleaved from the resin using TFA to give a purple solid and volatiles removed under reduced pressure. The residue was taken up in H_2O and purified using semi preparative RP-HPLC method C to give the product as the TFA salt. MS (ESI) Exact mass calcd for $C_{39}H_{51}N_8O_8$ [M+H]+ 759.38, found: 759.36. Purity >95% by HPLC (220 nm).



N_3, Maleimide, TMR (5.90): TOBa-N_3, NH_2.TFA, TMR 4.88 (1 mg, 1.15 µmol) was stirred in 0.5 mL CH_2Cl_2 with DIPEA (1.0 µL, 5.7 µmol) and to this was added 4.89 (0.4 mg, 1.15 µmol). The mixture was stirred for 5 hours then volatiles removed under reduced pressure. The residue was taken up in buffer (depending on experiment) and used crude. MS (ESI) Exact mass calcd for $C_{47}H_{58}N_9O_{11}$ $[M+H]^+$ 924.43, found: 924.41. Purity 57% by µHPLC (220 nm).

8.5.5 Trifunctional Tags Proof of Concept Studies

8.5.5.1 Pull-down of His-hCINAP-Cy5 using 5.87 (Experiment performed by Dr. Nhan Pham):

Experiment 1 – Reaction of Tag 5.87 with resin 5.85, then addition of His-hCINAP-Cy5

DCBO resin **5.85** (0.5 mg) was suspended in PBS pH 7.4 with 0.1% Pluronic. NTA-N_3-TMR tag **5.87** was added to a final concentration of 500 nM and left to incubate for 2 hours. The resin was washed with PBS buffer 3 times, then 500 nM $NiCl_2$ (final concentration) and 500 nM of His-hCINAP-Cy5 was added and the mixture incubated for 2 hours. Resin was again washed with PBS buffer 3 times. Imaging on PerkinElmer's Opera® High Content Screening System was carried out. Repeat experiments were carried out with concentrations of $NiCl_2$ up to 2uM.

Imaging conditions: 560nm laser 1000uW, expo. Time 120ms, 640nm laser 2000uW, exp. Time 200 ms

Experiment 2 – Reaction of tag 5.87 with His-hCINAP-Cy5, then addition of resin 5.85

500 nM His-hCINAP-Cy5, 500 nM NTA-N_3-TMR tag **5.87** and 500 nM $NiCl_2$ were incubated together for 2 hours. DCBO resin **5.85** (0.5 mg) was added to the solution and this was incubated for 2 hours at rt. Imaging on PerkinElmer's Opera® High Content Screening System was carried out.

Imaging conditions: 560nm laser 1000uW, expo. Time 120ms, 640nm laser 2000uW, exp. Time 200 ms

8.5.5.2 Pull-down of alpha-synuclein I112C using 5.90 (Carried out with Irene Pérez Pi and Nhan Pham):

Labelling of alpha-synuclein I112C

Alpha-synuclein I112C (27.1 nmol, expressed by Irene Pérez Pi) was incubated with 20 equivalents **5.90** in 200 µL buffer (20 mM tris, 100 mM NaCl, pH = 7.0). The mixture was incubated overnight at 4 ºC then purified by NAP5^{TM} column (GE healthcare). Fractions were analysed by SDS-PAGE and HPLC. Fractions containing labelled protein and no free tag were used for pull down experiments.

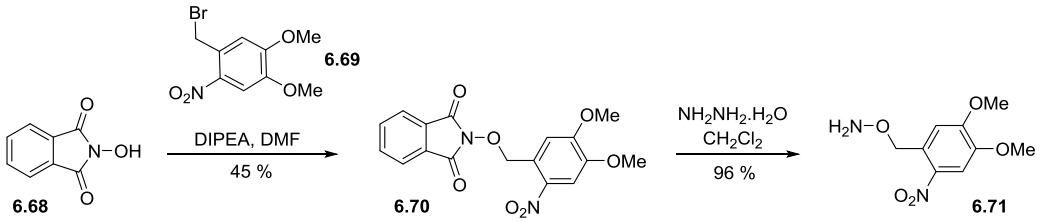
Pull-down onto DBCO resin and Imaging

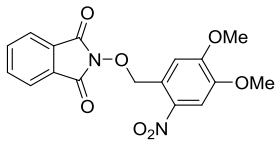
DBCO resin **5.85** (0.5 mg) was incubated with Alpha-synuclein I112C labelled with **5.90** for 3 hours. The resin was washed with buffer (20 mM tris, 100 mM NaCl, pH = 7.0) three times then the resin imaged on PerkinElmer's Opera® High Content Screening System.

Imaging conditions: 560nm laser 1000uW, expo. Time 120ms.

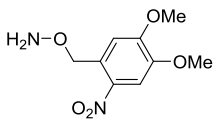
8.6 Chapter 6 Experimental

8.6.1 Caged Hydroxamic Acid



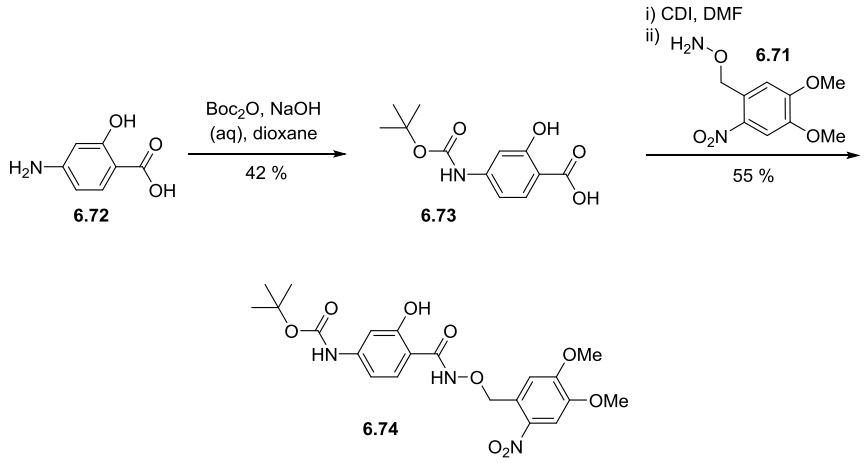


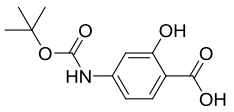
2-[(4,5-Dimethoxy-2-nitro-phenyl)methoxy]isoindoline-1,3-dione (6.70): To a solution of *N*-hydroxyphthalimide **6.68** (100 mg, 0.73 mmol) and DIPEA (190 µL) in 2 mL DMF was added 4,5-dimethoxy-2-nitrobenzylbromide **6.69** (101 mg, 0.37 mmol) and the reaction heated to 70 °C for 2 hours. A precipitate formed which was filtered off and washed with MeOH. The filtrate was concentrated to dryness then taken up in EtOAc and washed with $NaHCO_3$ three times, then the organic layers dried ($MgSO_4$) and concentrated to dryness. The crude material was combined with the filtered solid and purified using flash chromatography (CH_2Cl_2) to give the title compound as a yellow solid (60 mg, 0.17 mmol, 45%). R_f = 0.38 (7:3 CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.86 (2H, dd, *J* = 5.5, 3.1 Hz, 2 x phthalimide C**H**), 7.78 (2H, dd, *J* = 5.5, 3.0 Hz, 2 x Phthalimide C**H**), 7.72 (1H, s, Ar**H**), 7.62 (1H, s, Ar**H**), 5.69 (2H, s, OC**H**$_2$), 4.09 (3H, s, C**H**$_3$), 3.97 (3H, s, C**H**$_3$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 163.4 (C), 153.7 (C), 148.3 (C), 139.6 (C), 134.7 (CH), 128.8 (C), 126.3 (C), 123.7 (CH), 110.4 (CH), 107.9 (CH), 76.2 (CH_2), 56.7 (CH_3), 56.4 (CH_3); Exact mass calcd for $C_{17}H_{15}N_2O_7$ $[M+H]^+$: 359.09, found: 359.04.



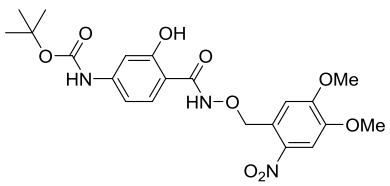
O-[(4,5-Dimethoxy-2-nitro-phenyl)methyl]hydroxyl amine (6.71): 2-[(4,5-Dimethoxy-2-nitro-phenyl)methoxy]isoindoline-1,3-dione **6.70** (286 mg, 0.80 mmol) was stirred in 3 mL CH_2Cl_2 and to this was added hydrazine hydrate (79 mg, 1.60 mmol). The mixture was stirred overnight during which time a precipitate formed, which was removed by filtration. The filtrate was washed with sat. $NaHCO_3$ solution (3 x 3 mL) then dried ($MgSO_4$) and concentrated to dryness to give the title compound as a yellow solid (175 mg, 0.76 mmol, 96%). R_f = 0.54 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.69 (1H, s, Ar**H**), 7.13 (1H, s, Ar**H**), 5.58 (2H, s, N**H**$_2$), 5.10 (2H, s, C**H**$_2$), 4.01 (3H, s, OC**H**$_3$), 3.96 (3H, s, OC**H**$_3$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 153.6 (C), 147.8 (C), 140.1 (C), 129.6 (C), 109.9 (CH), 108.0 (CH), 74.5

(CH_2), 56.4 (CH_3), 56.4 (CH_3); Exact mass calcd for $C_9H_{13}N_2O_5$ $[M+H]^+$: 229.08, found: 228.82.





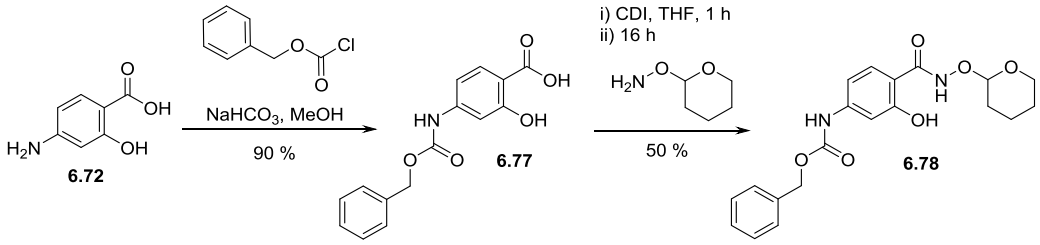
4-(*tert*-Butoxycarbonylamino)-2-hydroxy-benzoic acid (6.73): 4-Aminosalicylic acid **6.72** (1.53 g, 10.00 mmol) was stirred in 30 mL dioxane/H_2O (2:1) and to this was added 10 mL 1M NaOH. The mixture was cooled to 0 °C and a solution of Boc_2O (2.40 g, 10.10 mmol) in 10 mL THF added dropwise. The solution was then allowed to warm to rt and was stirred for 16 hours. A further 10 mL 1M NaOH and 1.1 eq. Boc_2O (2.40 g) was added as the reaction had not gone to completion. After another 16 h, 10 mL EtOAc was added and the mixture acidified with $KHSO_4$. The aqueous phase was extracted with EtOAc (3 x 10 mL) and the organic layers pooled, dried ($MgSO_4$) and concentrated to dryness. The residue was triturated using MeOH to give the title compound as a white solid (1.06 g, 4.19 mmol, 42%); R_f = 0.20 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $(CD_3)_2SO$)) δ 9.68 (1H, s, COO**H**), 7.65 (1H, d, *J* = 8.7 Hz, Ar**H**), 7.12 (1H, d, *J* = 2.0 Hz, Ar**H**), 6.99 (1H, dd, *J* = 8.7, 2.0 Hz, Ar**H**), 1.48 (9H, s, C(C**H**$_3$)$_3$); ^{13}C NMR (125 MHz, $(CD_3)_2SO$)) δ 171.7 (C), 162.2 (C), 162.4 (C), 146.2 (C), 130.9 (CH), 109.2 (CH), 106.6 (CH), 104.5 (C), 79.8 (C), 28.0 (CH_3); Exact mass calcd for $C_{12}H_{16}NO_5$ $[M+H]^+$: 254.10, found: 254.19.

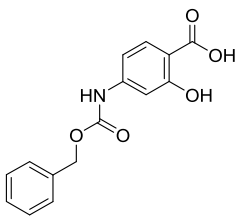


***tert*-Butyl *N*-(4-{[(4,5-dimethoxy-2- nitrophenyl) methoxy]carbamoyl}-3-hydroxyphenyl)carbamate (6.74):** 4-(*tert*-Butoxycarbonylamino)-2-hydroxy-benzoic acid **6.73** (14 mg, 0.06 mmol) was stirred in 0.5 mL THF and to this was added CDI (13 mg, 0.083 mmol). This was stirred for 4 hours then *O*-[(4,5-dimethoxy-2-nitro-phenyl)methyl]hydroxyl amine (20 mg, 0.09 mmol) was added and the mixture stirred overnight, during which time a precipitate was formed. The

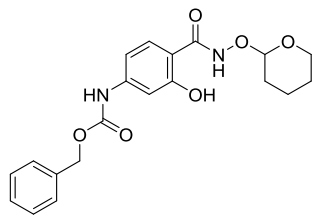
mixture was filtered and the filtrate diluted with Et_2O and washed with $NaHCO_3$. The organics were pooled and dried ($MgSO_4$) then concentrated to dryness. The resultant yellow residue was purified using flash chromatography (40% EtOAc/CyHex) to give the title product as an orange semisolid (14 mg, 0.03 mmol, 55%). R_f = 0.27 (40% EtOAc/CyHex); ^{1}H NMR (500 MHz, ($CDCl_3$) δ 11.52 (1H, s, C=ON**H**O), 9.23 (1H, s, ArO**H**), 7.68 (1H, s, Ar**H**), 7.41 (1H, s, Ar**H**), 6.97 – 6.90 (2H, m, Ar**H**), 6.66 (1H, s, Ar**H**), 5.43 (2H, s, ArC$\mathbf{H}_2$), 4.01 (3H, s, OC$\mathbf{H}_3$), 3.95 (3H, s, OC$\mathbf{H}_3$), 1.52 (9H, s, C(C$\mathbf{H}_3$)$_3$); MS (ESI) Exact mass calcd for $C_{21}H_{26}N_3O_9$ $[M+H]^+$: 464.17, found: 464.15.

8.6.2 THP-Protected Hydroxamic Acid

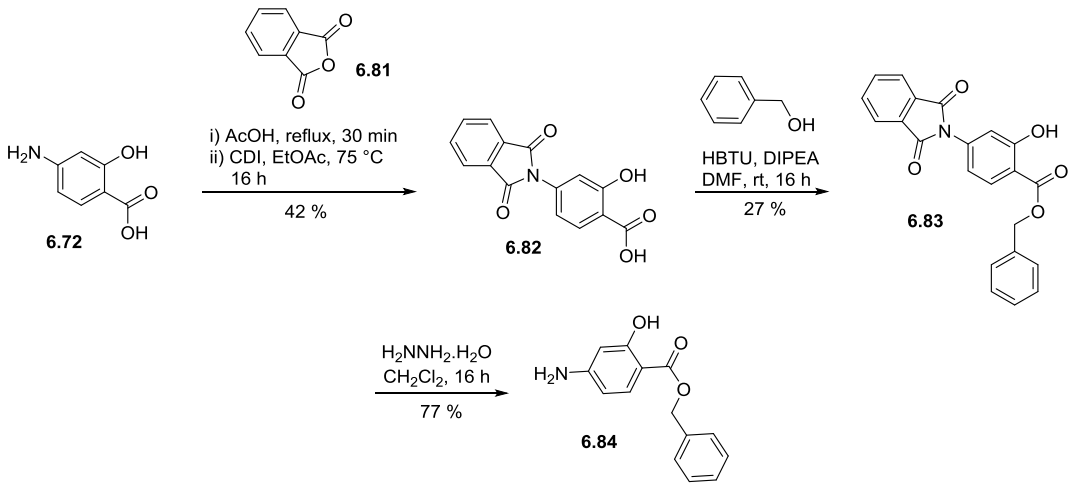


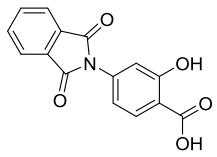


4-{[(Benzyloxy)carbonyl]amino}-2-hydroxybenzoic acid (6.77): This compound was prepared according to the reported procedure.[275] 4-Aminosalicylic acid **6.72** (1.50 g, 10.00 mmol) and $NaHCO_3$ (1.0 g, 12 mmol) were stirred in 8 mL methanol and to this was added benzyl chloroformate (1.70 ml, 12.00 mmol). The mixture was stirred at room temperature overnight. Solvents were removed under partial pressure and the residue taken up in EtOAc and washed with 1 M HCl. The aqueous layer was extracted with EtOAc and the combined organics were washed with brine and dried ($MgSO_4$). The solvent was evaporated to obtain the title compound as a light pink solid (2.57 g, 9.00 mmol, 90%). R_f = 0.30 (10% H_2O/MeCN); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 11.35 (1H, br s, N$\mathbf{H}_2$), 10.12 (1H, s, ArO**H**), 7.69 (1H, d, *J* = 8.7, Ar**H**), 7.48 – 7.30 (5H, m, $CH_2C_6\mathbf{H}_5$), 7.16 (1H, d, *J* = 2.0 Hz, Ar**H**), 6.99 (1H, dd, *J* = 8.7, 2.1 Hz, Ar**H**), 5.17 (2H, s, ArC$\mathbf{H}_2$); ^{13}C NMR (125 MHz, $(CD_3)_2SO$) δ 171.6 (C), 162.2 (C), 153.0 (C), 145.7 (C), 136.2 (C), 131.1 (CH), 128.5 (CH), 128.2 (CH), 128.2 (CH), 109.3 (CH), 106.9 (C), 104.7 (CH), 66.1 (CH_2); MS (ESI) Exact mass calcd for $C_{15}H_{14}NO_5$ $[M+H]^+$: 288.09, found: 288.14. Data is in agreement with the literature.[275]



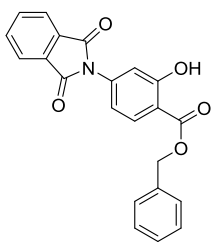
Benzyl *N*-{3-hydroxy-4-[(oxan-2-yloxy)carbamoyl]phenyl} carbamate (6.78): 4-{[(Benzyloxy)carbonyl]amino}-2-hydroxy benzoic acid **6.77** (204 mg, 0.71 mmol) was stirred in 2 mL THF with CDI (126 mg, 0.78 mmol) for 1 hour. *O*-(Tetrahydro-2*H*-pyran-2-yl)hydroxylamine (100 mg, 0.85 mmol) was added and the mixture stirred for 3 hours, after which it was washed 3x with sat. $NaHCO_3$ solution. The organics were dried ($MgSO_4$) and concentrated to dryness. The residue was purified using flash chromatography (1:1 CyHex/EtOAc) to give the title compound as a white solid (137 mg, 0.35 mmol, 50%). R_f = 0.49 (1:1 CyHex/EtOAc); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 11.68 (1H, s, ArO**H**), 9.03 (1H, s, C=ON**H**O), 7.43 – 7.34 (4H, m, Ar**H**), 7.31 (1H, d, *J* = 8.7 Hz, Ar**H**), 7.0 (1H, dd, *J* = 8.7, 1.9 Hz, Ar**H**), 6.95 (1H, d, *J* = 2.1 Hz, Ar**H**), 6.87 (1H, s, Ar**H**), 5.21 (2H, s, ArC**H**$_2$), 5.05 (1H, t, *J* = 2.9 Hz, THP 1**H**), 4.06 – 3.98 (1H, m, THP OCHC**H**$_2$), 3.72 – 3.65 (1H, m, THP OCHC**H**$_2$), 1.93 – 1.82 (2H, m, THP C**H**$_2$), 1.72 – 1.56 (4H, m, THP 2 x C**H**$_2$); MS (ESI) Exact mass calcd for $C_{20}H_{23}N_2O_6$ $[M+H]^+$: 387.16, found: 387.05.



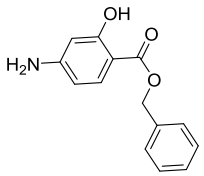


4-(1,3-Dioxo-2,3-dihydro-1H-isoindol-2-yl)-2- hydroxybenzoic acid (6.82): 4-Aminosalicylic acid **6.72** (500 mg, 3.27 mmol) was heated with phthalic anhydride (484 mg, 3.27 mmol) in 20 mL acetic acid. A precipitate formed which was dissolved in EtOAc and the organic layers washed with 1M HCl. The organics were dried ($MgSO_4$) then concentrated to dryness then taken up in EtOAc and CDI (556 mg, 3.43 mmol) added and the heated at 75 °C overnight. Solvents were removed under reduced pressure and the resultant residue suspended in H_2O and filtered. The collected solid was washed again with H_2O and dried to give the title compound (389 mg, 1.37 mmol, 42%). R_f = 0.30 (10% H_2O/MeCN); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 8.17 – 7.77 (5H, m, Ar**H**), 7.10 (1H, s, Ar**H**), 7.06 (1H, d, *J* = 8.4 Hz, Ar**H**); ^{13}C

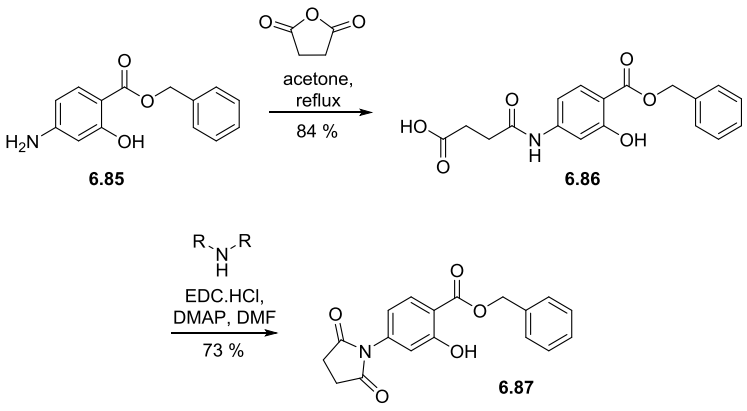
NMR (125 MHz, $(CD_3)_2SO$) δ 171.3 (C), 166.5 (C), 161.1 (C), 137.9 (C), 134.8 (CH), 131.4 (C), 130.6 (CH), 123.6 (CH), 117.7 (CH), 115.3 (CH), 112.4 (C); MS (ESI) Exact mass calcd for $C_{15}H_{10}NO_5$ $[M+H]^+$: 284.06, found: 284.20.

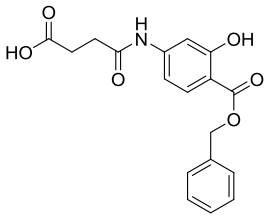


Benzyl 4-(1,3-dioxo-2,3-dihydro-1*H*-isoindol-2-yl)-2- hydroxylbenzo ate (6.83): 4-(1,3-Dioxo-2,3-dihydro-1H-isoindol-2-yl)-2-hydroxybenzoic acid **6.82** (300 mg, 1.48 mmol) was stirred in 20 mL DMF with benzyl alcohol (169 µL, 1.63 mmol) and to this was added HBTU (618 mg, 1.63 mmol), then DIPEA (515 µL, 2.96 mmol). The solution was stirred overnight then taken up in sat. $NaHCO_3$ solution. This was extracted using EtOAc 3x then the organic layers pooled, dried ($MgSO_4$) and concentrated to dryness. The residue was purified using flash chromatography (1:1 EtOAc/CyHex) to give the title compound (148 mg, 0.40 mmol, 27%). MS (ESI) Exact mass calcd for $C_{22}H_{16}NO_5$ $[M+H]^+$: 374.10, found: 374.11. This was taken on to the next stage without further characterisation.

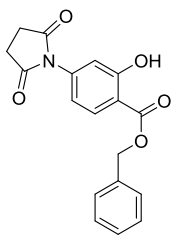


Benzyl 4-amino-2-hydroxybenzoate (6.84): Benzyl 4-(1,3-dioxo-2,3-dihydro-1H-isoindol-2-yl)-2- hydroxylbenzoate **6.83** (148 mg, 0.44 mmol) was stirred in 5mL CH_2Cl_2 and to this was added hydrazine hydrate (40 mg, 0.8 mmol). The mixture was stirred overnight, and the resulting precipitate removed by filtration. The resulting filtrate was concentrated to dryness then triturated using CyHex to give the title compound as a white solid (75 mg, 0.31 mmol, 77%). R_f = 0.60 (7:3 CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.94 (1H, s, COH), 7.68 (1H, d, J = 8.6 Hz, Ar**H**), 7.49 – 7.32 (5H, m, Ph), 6.18 (1H, d, *J* = 2.2 Hz, Ar**H**), 6.15 (1H, dd, *J* = 8.6, 2.2 Hz, Ar**H**), 5.34 (2H, s, C**H**$_2$), 4.09 (2H, br s, N**H**$_2$); ^{13}C NMR (125 MHz, $(CD_3)_2SO$) δ 169.8 (C), 163.7 (C), 153.4 (C), 135.9 (C), 131.7 (CH), 128.6 (CH), 128.3 (CH), 128.1 (CH), 106.8 (CH), 103.0 (C), 100.7 (CH), 66.2 (CH_2); MS (ESI) Exact mass calcd for $C_{14}H_{14}NO_3$ $[M+H]^+$: 244.10, found: 244.15. Data in agreement with the literature.[276]



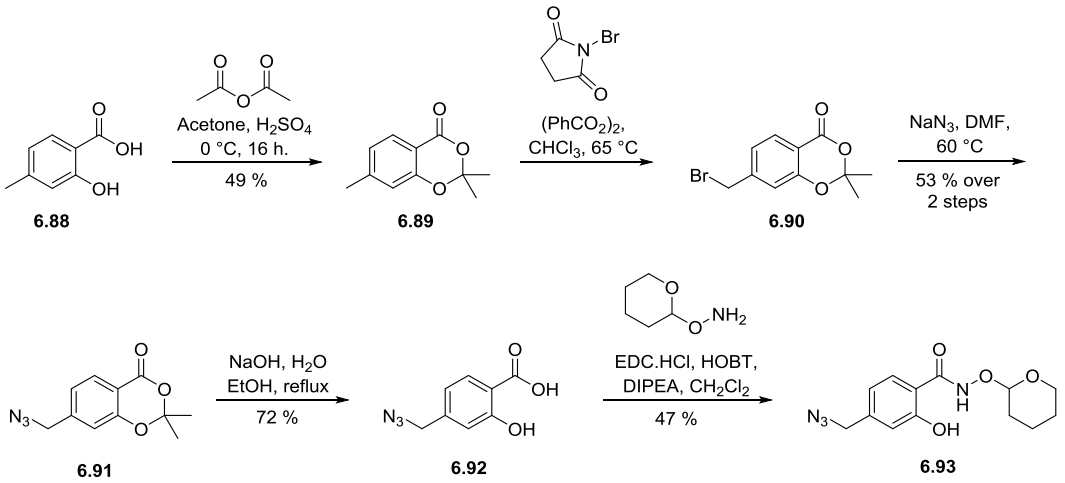


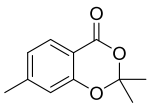
4-({4-[(Benzyloxy)carbonyl]-3-hydroxyphenyl}amino)-4-oxobutanoic acid (6.86): Benzyl 4-amino-2-hydroxybenzoate **6.85** (1 g, 4.12 mmol) was stirred in 20 mL acetone with succinic anhydride (534 mg, 5.35 mmol) and the suspension heated to reflux for 72 hours. Solvents were removed under reduced pressure to give an off white solid. CH_2Cl_2 was added which resulted in precipitation of a fine white solid. This was filtered off and dried to give the title compound as a white solid (1.18 g, 3.44 mmol, 84%). R_f = 0.32 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 10.61 (1H, s, CO_2**H**), 10.28 (1H, s, CO**H**), 7.75 (1H, d, *J* = 8.7 Hz, Ar**H**), 7.48 (2H, d, *J* = 7.2 Hz, Ar**H**), 7.41 (4H, m, Ar**H**), 7.07 (1H, d, *J* = 8.2 Hz, Ar**H**), 5.37 (2H, s, benzyl C**H**$_2$), 2.6 (2H, m, C**H**$_2$CH$_2$), 2.52 (2H, m, succinic CH$_2$C**H**$_2$); ^{13}C NMR (125 MHz, $(CD_3)_2SO$) δ 173.7 (C), 171.0 (C), 168.4 (C), 161.4 (C), 145.8 (C), 135.8 (C), 130.9 (CH), 128.6 (CH), 128.3 (CH), 128.1 (CH), 110.4 (CH), 107.2 (C), 105.9 (CH), 66.3 (CH_2), 31.3 (CH_2), 28.6 (CH_2); MS (ESI) Exact mass calcd for $C_{18}H_{18}NO_6$ $[M+H]^+$: 344.11, found: 344.08.



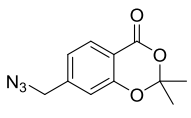
Benzyl 4-(2,5-dioxopyrrolidin-1-yl)-2-hydroxybenzoate (6.87): 4-({4-[(Benzyloxy)carbonyl]-3-hydroxyphenyl}amino)-4-oxobutanoic acid **6.86** (600 mg, 1.75 mmol) and dimethyl 2,2'-iminodiacetate (253 mg) were stirred together in 15 mL DMF and to this was added EDC.HCl (361 mg, 1.88 mmol) and DMAP (38 mg, 0.31 mmol). Consumption of the acid was seen by TLC so solvents were removed under reduced pressure and the residue taken up in CH_2Cl_2. The organics were washed with H_2O and sat. aq. $NaHCO_3$ then dried with $MgSO_4$ and concentrated to dryness. The residue was purified using flash chromatography (1:1 EtOAc/CyHex) to give a white solid. Upon analysis, this appeared to be the title compound (416 mg, 1.28 mmol, 73%). R_f = 0.21 (1:1 EtOAc/CyHex); ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.84 (1H, s, CO**H**), 7.99 (1H, d, *J* = 8.6 Hz, Ar**H**), 7.47 – 7.35 (5H, m, Ar**H**), 7.01 (1H, d, *J*

= 2.0 Hz, Ar**H**), 6.87 (1H, dd, *J* = 8.6, 2.0 Hz, Ar**H**), 5.41 (2H, s, Benzyl C**H**$_2$), 2.90 (4H, s, 2 x C**H**$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 175.4 (C), 169.3 (C), 162.1 (C), 138.2 (C), 135.1 (C), 130.7 (CH), 128.7 (CH), 128.6 (CH), 128.2 (CH), 117.0 (CH), 115.6 (CH), 112.3 (C), 67.2 (CH_2), 28.4 (CH_2); MS (ESI) Exact mass calcd for $C_{18}H_{16}NO_5$ $[M+H]^+$: 326.10, found: 326.16.



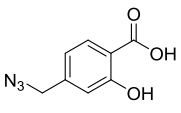


2,2,7-Trimethyl-1,3-benzodioxin-4-one (6.89): The procedure was carried out as described in the literature.[277] A stirred mixture of 4-methylsalicylic acid **6.88** (8.00 g, 52.63 mmol), acetic anhydride (7.50 mL), and acetone (25 mL) was cooled in an ice/acetone bath to -8 °C. To the cooled mixture was added 0.1 mL conc. H_2SO_4 dropwise. The flask was kept in the ice bath overnight, slowly reaching room temperature. Volatiles were removed under reduced pressure are the resulting brown solid purified using flash chromatography (1:1 EtOAc/CyHex) to give the title compound as a white solid (4.96 g, 25.68 mmol, 49%). R_f = 0.42 (9:1 CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.84 (1H, d, *J* = 8.0 Hz, Ar**H**), 6.93 (1H, m, Ar**H**), 6.77 (1H, m, Ar**H**), 2.39 (3H, s, ArC**H**$_3$), 1.73 (6H, s, $C(CH_3)_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 161.2 (C), 156.1 (C), 148.0 (C), 129.5 (CH), 123.8 (CH), 117.3 (CH), 111.0 (C), 106.2 (C), 25.8 (CH_3), 22.0 (CH_3); MS (ESI) Exact mass calcd for $C_{11}H_{13}O_3$ $[M+H]^+$: 193.09, found: 193.01.

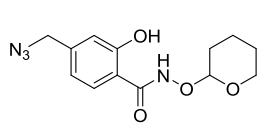


7-(Azidomethyl)-2,2-dimethyl-1,3-benzodioxin-4-one (6.91): 2,2,7-Trimethyl-1,3-benzodioxin-4-one **6.89** (5.40 g. 28.00 mmol) was dissolved in $CHCl_3$ (20 mL) and to the solution was added *N*-bromo succinimide (5.78 g, 32.48 mmol) and benzoyl peroxide (1.49 g, 6.16 mmol) and the mixture heated to 65 °C overnight. The solution was cooled on ice and the resulting precipitate filtered off and washed with CyHex. The filtrate was concentrated to dryness and purified

using flash chromatography (CH_2Cl_2) to give a yellow oil containing both remaining starting material and the brominated intermediate, which was used directly in the next step. The oil was dissolved in DMF (50 mL) and to this was added NaN_3 (1.82 g, 28 mmol) and the mixture heated to 60 °C for 30 minutes. Solvents were removed under reduced pressure and the brown residue taken up in EtOAc (50 mL) and washed with water (3 x 30 mL). The organics were dried ($MgSO_4$) and concentrated to dryness then purified using flash chromatography (9:1 CyHex/EtOAc) to give a colourless oil (3.44g, 14.76 mmol, 53%). R_f = 0.29 (1:1 CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.97 (1H, d, *J* = 8.0 Hz, Ar**H**), 7.06 (1H, m, Ar**H**), 6.95 (1H, m, Ar**H**), 4.41 (2H, s, C**H**$_2$Br), 1.75 (6H, s, C(CH_3)$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 160.6 (C), 156.3 (C), 144.7 (C), 130.2 (CH), 121.8 (CH), 116.2 (CH), 113.2 (C), 106.6 (C), 54.1 (CH_2), 25.8 (CH_3); MS (ESI) Exact mass calcd for $C_{11}H_{12}N_3O_3$ $[M+H]^+$: 234.09, found: 233.97.

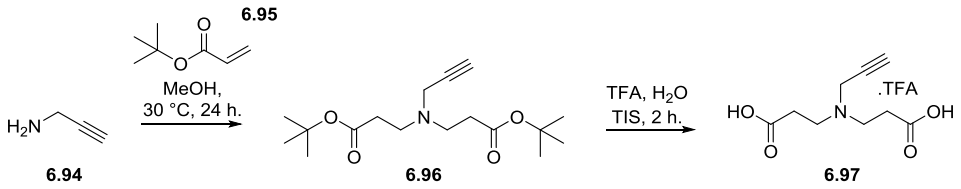


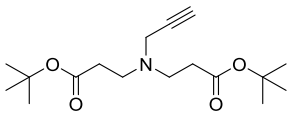
4-(azidomethyl)-2-hydroxybenzoic acid (6.92): 7-(Azidomethyl)-2,2-dimethyl-1,3-benzodioxin-4-one **6.91** (3.42 g. 14.74 mmol) was dissolved in ethanol (100 mL) and to this was added NaOH (2.95 g, 73.8 mmol) in 14 mL H_2O. The solution was heated to reflux for 3 hours then stirred at rt for 24 hours. 1M HCl (80 mL) was added to the mixture then the ethanol was removed under partial pressure. The remaining aqueous fraction was extracted using EtOAc (3 x 50 mL) then the organic layers combined, dried ($MgSO_4$) and concentrated to dryness. The resulting solid was purified using flash chromatography (EtOAc) to give a light red solid (2.05 g, 10.62 mmol, 72%). R_f = 0.45 (9:1:0.1 CH_2Cl_2/MeOH/AcOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 7.80 (1H, d, *J* = 8.0 Hz, Ar**H**), 6.92 (1H, d, *J* = 1.3 Hz, Ar**H**), 6.88 (1H, dd, *J* = 8.1, 1.6 Hz, Ar**H**), 4.47 (2H, s, C**H**$_2$N$_3$); ^{13}C NMR (125 MHz, $(CD_3)_2SO$) δ 171.6 (C), 161.3 (C), 143.3 (C), 130.7 (CH), 118.6 (CH), 116.3 (CH), 113.2 (C), 53.0 (CH_2).



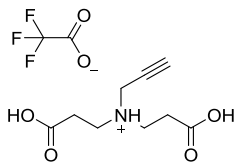
4-(Azidomethyl)-2-hydroxy-*N*-(oxan-2-yloxy)benzamide (6.93): 4-(azidomethyl)-2-hydroxybenzoic acid **6.92** (1.00 g, 5.18 mmol) was stirred in CH_2Cl_2 with *O*-(Tetrahydro-2*H*-pyran-2-yl)hydroxylamine (729 mg, 6.22 mmol) and to the stirring solution was added EDC.HCl (1.49 g, 7.77 mmol), HOBt (1.19 g, 7.77 mmol) and DIPEA (1.80 mL, 10.36 mmol). The mixture was stirred at rt overnight then sat. $NaHCO_3$ solution was added. The organics were separated off dried ($MgSO_4$) then concentrated to dryness. The residue was purified using flash chromatography (7:3 CyHex/EtOAc) to give a white solid (714 mg, 2.44 mmol, 47%). R_f = 0.30 (8:2 CyHex/EtOAc); ^{1}H NMR (500 MHz, $(CDCl_3)$ δ 11.60 (1H, s, ArO**H**), 9.11 (1H, s, Ar**H**), 7.40 (1H, d, *J* = 8.1 Hz, Ar**H**), 6.96 (1H, d, *J* = 1.6 Hz, Ar**H**), 6.81 (1H, dd, *J* =

8.2, 1.6 Hz, Ar**H**), 5.07 (1H, t, *J* = 2.8 Hz, NHOC**H**), 4.33 (2H, s, C**H**$_2$N$_3$), 4.03 (1H, m, THP C**H**$_2$), 3.69 (1H, m, THP C**H**$_2$), 1.89 (3H, m, THP C**H**$_2$), 1.65 (3H, m, THP C**H**$_2$); ^{13}C NMR (125 MHz, ($CDCl_3$) δ 168.0 (C), 161.6 (C), 142.6 (C), 126.0 (CH), 118.2 (CH), 117.9 (CH), 111.9 (C), 103.2 (CH), 63.0 (CH_2), 54.1 (CH_2), 28.0 (CH_2), 24.9 (CH_2), 18.7 (CH_2); MS (ESI) Exact mass calcd for $C_{13}H_{17}N_4O_4$ $[M+H]^+$: 293.12, found: 293.05.

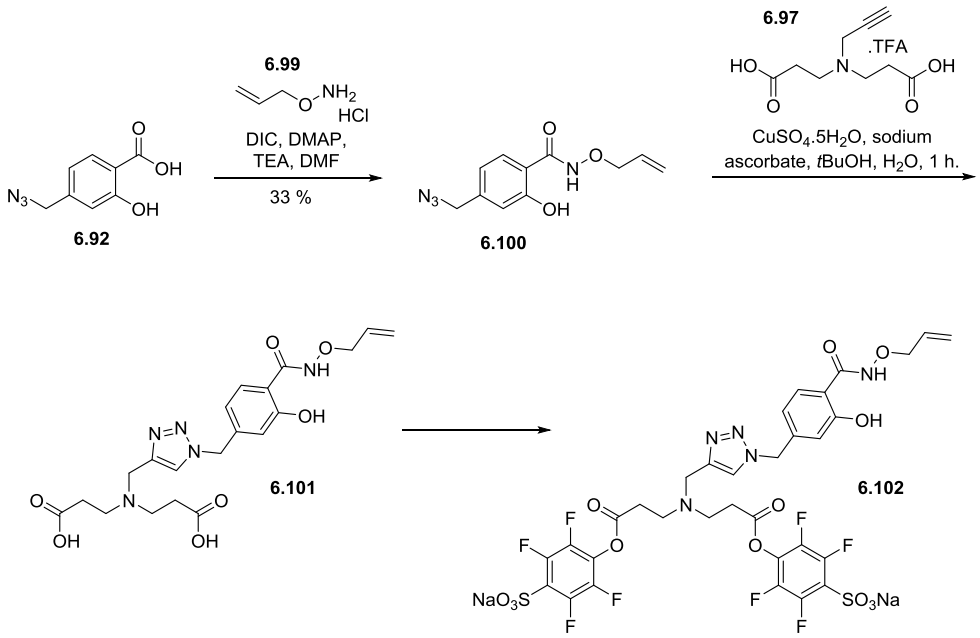


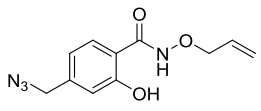


***tert*-Butyl 3-{[3-(tert-butoxy)-3-oxopropyl](prop-2-yn-1-yl)ami no}propanoate (6.96)**: To a solution of propargylamine **6.94** (220 mg, 4.00 mmol) in MeOH (10 mL) was added dropwise *tert*-butyl acrylate **6.95** (2.11 mL, 14.4 mmol). The solution was stirred at 30 °C for 36 hours then concentrated to dryness. The residue was purified using silica gel chromatography (6:4 CyHex/EtOAc) to give a colourless oil (589 mg, 1.89 mmol, 47%). ^{1}H NMR (500 MHz, ($CDCl_3$) δ 3.42 (2H, app d, *J* = 1.7 Hz, C**H**$_2$C≡CH)), 2.80 (4H, t, *J* = 7.3 Hz, 2 x C=OC**H**$_2$), 2.38 (4H, J = 7.3 Hz, 2 x NC**H**$_2$CH$_2$), 2.19 (1H, m, C≡C**H**), 1.45 (18H, s, 2 x C(C**H**$_3$)$_3$); ^{13}C NMR (125 MHz, ($CDCl_3$) δ 171.6 (C), 80.3 (C), 78.4 (C), 73.0 (CH), 49.2 (CH_2), 41.9 (CH_2), 34.4 (CH_2), 28.1 (CH_3); MS (ESI) Exact mass calcd for $C_{17}H_{30}NO_4$ $[M+H]^+$: 312.22, found: 312.09.

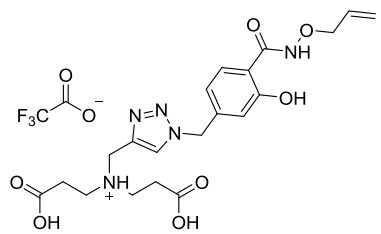


3-[(2-Carboxyethyl)(prop-2-yn-1-yl)amino]propanoic acid TFA salt (6.97): The di-*tert*-butyl ester **6.96** (550 mg, 1.77 mmol) was stirred in 10 mL of a mixture of TFA:H_2O:TIS (90:5:5) for 2 hours, then volatiles were removed under reduced pressure. The resultant yellow gum was triturated with diethyl ether and CyHex (1:1) to give a while solid, which was isolated by filtration and dried to give the diacid as a trifluoroacetate salt (505 mg, 1.61 mmol, 91%). ^{1}H NMR (500 MHz, $SO(CD_3)_2$) δ 4.05 (2H, br s, C**H**$_2$C≡CH), 3.74 (1H, br s, C≡C**H**), 3.24 (4H, t, *J* = 7.3 Hz, 2 x C=OC**H**$_2$), 2.69 (4H, t, *J* = 7.3 Hz, 2 x NC**H**$_2$CH$_2$); ^{13}C NMR (125 MHz, $SO(CD_3)_2$) δ 171.8 (C), 158.2 (q, ^2J = 32.9 Hz, **C**=O**C**F$_3$), 116.7 (q, *J* = 297.7 Hz, C=O**C**F$_3$), 80.4 (C), 74.1 (CH), 48.6 (CH_2), 42.4 (CH_2), 29.3 (CH_2); MS (ESI) Exact mass calcd for $C_9H_{14}NO_4$ $[M+H]^+$: 200.09, found: 200.15.

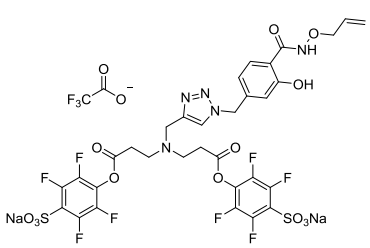




4-(azidomethyl)-2-hydroxy-*N*-(prop-2-en-1-yloxy)benzamide (6.100): To a solution of 4-(azidomethyl)-2-hydroxybenzoic acid **6.92** (100 mg, 0.52 mmol), DIC (97 µL, 0.63 mmol) and DMAP (32 mg, 0.26 mmol) in 5 mL DMF was added a solution of allyl hydroxylamine hydrochloride hydrate **6.99** (76 mg, 1.04 mmol) and triethylamine (75 µL, 0.54 mmol) in DMF. The solution was stirred at rt for 16 hours then solvents removed under reduced pressure. The residue was taken up in CH_2Cl_2 and washed x 3 with sat. $NaHCO_3$ solution. The organics were dried ($MgSO_4$) and concentrated to dryness and the residue purified using flash chromatography (6:4 CyHex/EtOAc) to give an orange oil (43 mg, 0.17 mmol, 33%). R_f = 0.44 (1:1 CyHex/EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.40 (1H, d, J = 7.5 Hz, Ar**H**), 6.96 (1H, m, Ar**H**), 6.81 (1H, dd, *J* = 8.1, 1.7 Hz, Ar**H**), 6.04 (1H, ddt, *J* = 17.1, 10.4, 6.4 Hz, $OCH_2C\mathbf{H}=CH_2$), 5.38 (2H, m, $OCH_2CH=C\mathbf{H}_2$), 4.53 (2H, d, *J* = 6.4, $OC\mathbf{H}_2CH=CH_2$), 4.33 (2H, s, $N_3C\mathbf{H}_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ 167.3 (C), 161.0 (C), 142.3 (C), 131.8 (CH), 126.1 (CH), 121.3 (C), 118.4 (CH), 117.7 (CH), 112.0 (CH_2), 77.6 (CH_2), 54.1 (CH_2); MS (ESI) Exact mass calcd for $C_{11}H_{12}N_4O_3$ $[M+H]^+$: 249.10, found: 249.06.



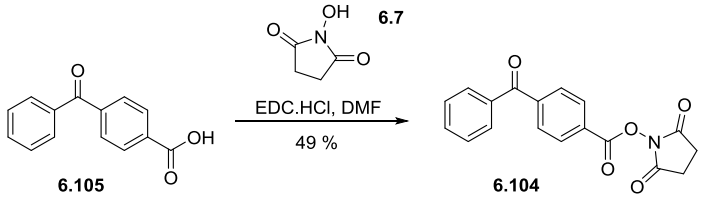
3-[(2-Carboxyethyl)({[1-({3-hydroxy-4-[(prop-2-en-1-yloxy)carbamoyl]phenyl}methyl)-1H-1,2,3-triazol-4-yl]methyl})amino]propanoic acid (Allyl Enrichment Tag – 6.101): 4-(Azidomethyl)-2-hydroxy-*N*-(prop-2-en-1-yloxy)benzamide **6.99** (40 mg, 0.16 mmol) was stirred in ^{t}BuOH with 3-[(2-carboxyethyl)(prop-2-yn-1-yl)amino]propanoic acid TFA salt **6.97** (50 mg, 0.16 mmol). To a vial containing $CuSO_4.5H_2O$ (40 mg, 0.16 mmol) and sodium ascorbate (32 mg, 0.16 mmol) was added 1 mL H_2O and the components mixed until a light yellow colour was seen. This was added to the *t*BuOH solution and the mixture stirred at rt for 30 minutes. The suspension was filtered and the filtrate concentrated to dryness under reduced pressure. The residue was purified using semi preparative RP-HPLC method C to give the title compound as a white solid (16.4 mg, 0.029 mmol, 18%). ^{1}H NMR (500 MHz, $(CD_3)_2SO$) δ 11.81 (1H, s, C=ON**H**O), 11.69 (1H, s, ArO**H**), 8.37 (1H, s, Ar**H**), 7.64 (1H, d, *J* = 8.1 Hz, Ar**H**), 6.83 (1H, d, *J* = 1.5 Hz, Ar**H**), 6.78 (1H, dd, *J* = 8.2, 1.6 Hz, Ar**H**), 5.99 (1H, ddt, *J* = 17.1, 10.6, 6.1 Hz, $OCH_2C\mathbf{H}=CH_2$), 5.65 (2H, s, $ArC\mathbf{H}_2$triazole), 5.35 (1H, dd, J = 17.3, 1.6 Hz, $OCH_2CH=C\mathbf{H}_2$), 5.27 (1H, m, $OCH_2CH=C\mathbf{H}_2$), 4.46 (2H, s, $NC\mathbf{H}_2$triazole), 4.42 (2H, d, J = 6.0, $OC\mathbf{H}_2CH=CH_2$), 3.26 (4H, t, J = 7.3, $NC\mathbf{H}_2CH_2$), 2.79 (4H, t, J = 7.3, $NCH_2C\mathbf{H}_2$); ^{13}C NMR (125 MHz, $(CD_3)_2SO$) δ 171.6 (C), 165.5 (C), 158.8 (C), 157.8 (q, ^{2}J = 31.4 Hz **C**=O**C**F_3), 141.2 (C), 136.7 (C), 132.8 (CH), 128.4 (CH), 127.3 (CH), 119.6 (CH_2), 118.1 (CH), 117.0 (q, *J* = 297.5 Hz, C=O**C**F_3), 116.3 (CH), 114.5 (C), 76.2 (CH_2), 52.4 (CH_2), 48.2 (CH_2), 46.8 (CH_2), 28.7 (CH_2); MS (ESI) Exact mass calcd for $C_{20}H_{26}N_5O_7$ $[M+H]^+$: 448.18, found: 448.19.



Activated Allyl Enrichment Tag (6.102): Allyl Enrichment Tag **6.101** (5.0 mg, 8.90 μmol) was dissolved in anhydrous DMF and to this was added 4-sulfotetrafluorophenol sodium salt (4.8 mg, 0.018 mmol) and DCC (3.7 mg, 0.018 mmol) and the mixture stirred for 16 hours. Solvents were removed under vacuum and the residue taken up in buffer (subject to experiment) and filtered, then used without further purification.

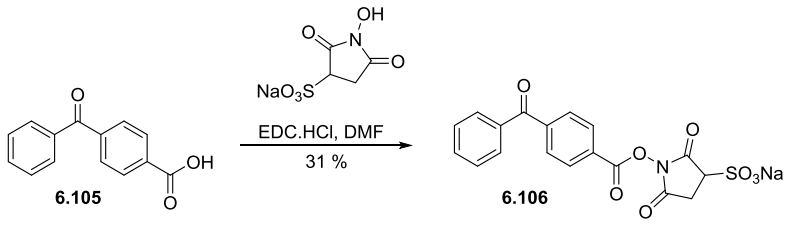
8.6.3 Benzophenone cross-linkers

SBP Cross-linker (6.104):



DMAP (10 mg) was added to a stirring solution of 4-benzoylbenzoic acid **6.105** (1.00 g, 4.42 mmol), *N*,*N*'Disuccinimidyl carbonate (2.3 g, 8.84 mmol) and triethylamine (2.5 mL, 17.68 mmol) in DMF (20 mL). The reaction mixture was stirred at rt for 16 hours. CH_2Cl_2 was added and this was washed with brine and water (x3). The organics were dried over $MgSO_4$ and concentrated to dryness to give a yellow oil. Diethyl ether was used to triturate the mixture which yielded an off white solid. This was filtered through silica, eluting with CH_2Cl_2. The organics were removed under vacuo to give a white solid (700 mg, 2.1 mmol, 49%). R_f = 0.91 (10% MeOH/CH_2Cl_2); ^{1}H NMR (500 MHz, ($CDCl_3$) δ 8.26 (m, 2H, Ar**H**), 7.89 (m, 2H, Ar**H**), 7.80 (m, 2H, Ar**H**), 7.64 (m, 1H, Ar**H**), 7.51 (m, 2H,Ar**H**), 2.93 (s, 4H, C**H**$_2$); ^{13}C NMR (125 MHz, ($CDCl_3$) δ 195.7 (C), 169.2 (C), 161.4 (C), 143.2 (C), 136.7 (C), 133.4 (CH), 130.7 (CH), 130.3 (CH), 130.2 (CH), 128.8 (CH), 128.3 (C), 25.9 (CH_2); MS (ESI) Exact mass calcd for $C_{18}H_{14}NO_5$ $[M+H]^+$: 324.31, found: 323.99. Data in agreement with the literature.[278]

SSBP Cross-linker (6.105):



4-Benzoylbenzoic acid **6.105** (37 mg, 0.16 mmol) and *N*-hydroxysulfosuccinimide (50 mg, 0.23 mmol) were stirred in 1 mL DMF and to this was added DIC (32 µL, 0.21 mmol) and the mixture stirred overnight. Solvents were removed under reduced pressure and the residue purified using semi-preparative RP-HPLC method C to yield the title product as a white solid (20 mg, 0.05 mmol, 31%). ^{1}H NMR (500 MHz, D_2O) δ 8.29 – 8.15 (2H, m, Ar**H**), 7.92 – 7.66 (5H, m, Ar**H**), 7.61 – 7.48 (2H, m, Ar**H**), 4.55 (1H, d, *J* = 7.1 H, SC**H**), 3.46 (1H, dd, *J* = 18.0, 8.7 Hz, C**H**$_2$), 3.26 (1H, m, C**H**$_2$); ^{13}C NMR (125 MHz, D_2O) δ 199.2 (C), 170.0 (C), 166.0 (C), 161.9 (C), 142.6 (C), 135.8 (C), 134.1 (CH), 130.5 (CH), 130.4 (CH), 130.3

(CH), 128.7 (CH), 127.6 (C), 56.6 (CH), 30.0 (CH_2); MS (ESI) Exact mass calcd for $C_{18}H_{14}NO_8S$ $[M+H]^+$: 404.44, found: 404.15.

8.6.4 Cross-linking Procedure (Carried out by Dr. Adam Belsom, University of Edinburgh)

Cross-linking was carried out using three cross-linker-to-protein weight-to-weight ratios of 0.33:1, 1:1 and 3:1. HSA (15 µg, 0.75 mg/mL) was mixed with SSBP (11.2 mM at 3:1 ratio) in cross-linking buffer (50 mM HEPES-OH, 20 mM NaCl, 5 mM $MgCl_2$, pH 7.8) to initiate lysine reaction with the sulfo-NHS ester component of the cross-linker. The benzophenone group was then photo-activated using UV irradiation. A UVP CL-1000L UV Cross linker at 365 nm was utilised for photo-activation. Samples were spread onto the inside of Eppendorf tube lids to form a thin film, placed on ice at a distance of 5 cm from the lamp and irradiated for either 15, 30, 45 or 60 minutes. The resulting cross-linked mixture was separated using an SDS-PAGE gel, bands were excised and the proteins reduced with 20 mM DTT, alkylated using 55 mM IAA and digested using trypsin following standard protocols. The resulting digests were de-salted using C-18 StageTips prior to MS analysis by ES-LCMS using an LTQ-Orbitrap Velos and identification of cross-links was carried out using in-house developed software Xi.

REFERENCES

References

[1] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149.

[2] H. M. Geysen, R. H. Meloen, S. J. Barteling, *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 3998.

[3] S. P. Fodor, J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu, D. Solas, *Science* **1991**, *251*, 767.

[4] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, *354*, 82.

[5] A. Furka, F. Sebestyen, M. Asgedom, G. Dibo, *Int. J. Pept. Protein. Res.* **1991**, *37*, 487.

[6] J. W. Davies, M. Glick, J. L. Jenkins, *Curr. Protoc. Chem. Biol.* **2006**, *10*, 343.

[7] M. Hintersteiner, "*Single Beads, Single Molecules, Single Cells*", PhD. Thesis, University of Salzburg (Salzburg), **2007**.

[8] M. Bagheri, M. Beyermann, M. Dathe, *Antimicrob. Agents Chemother.* **2009**, *53*, 1132.

[9] M. Hintersteiner, T. Kimmerlin, F. Kalthoff, M. Stoeckli, G. Garavel, J.-M. Seifert, N.-C. Meisner, V. Uhl, C. Buehler, T. Weidemann, M. Auer, *Chem. Biol.* **2009**, *16*, 724.

[10] V. Uhl, H. Gstach, J. M. Seifert, T. Jung, R. Falchetto, A. Fetsch, R. Stange, M. Schmied, C. Graf, A. Adam, F. Amberger, D. Monteux, S. Prechelmacher, A. Pruckner, W. Dautz, P. Graf, E. Hermes, P. Van Hooft, A. Berces, J. Müller, M. Auer, *Single Mol.* **2002**, *3*, 174.

[11] M. Hintersteiner, C. Buehler, V. Uhl, M. Schmied, J. MuÌller, K. Kottig, M. Auer, *J. Comb. Chem.* **2009**, *11*, 886.

[12] X. Chen, P. H. Tan, Y. Zhang, D. Pei, *J. Comb. Chem.* **2009**, *11*, 604.

[13] M. Auer, H. Gstach, *Fluorescent dyes (AIDA) for solid phase and solution phase screening*, Novartis AG, US6207831, **2001**.

[14] M. Hintersteiner, G. z. Ambrus, J. Bednenko, M. Schmied, A. J. S. Knox, N.-C. Meisner, H. Gstach, J.-M. Seifert, E. L. Singer, L. Gerace, M. Auer, *ACS Chem. Biol.* **2010**, *5*, 967.

[15] N.-C. Meisner, M. Hintersteiner, J.-M. Seifert, R. Bauer, R. M. Benoit, A. Widmer, T. Schindler, V. Uhl, M. Lang, H. Gstach, M. Auer, *J. Mol. Biol.* **2009**, *386*, 435.

[16] M. Hintersteiner, J. Kallen, M. Schmied, C. Graf, T. Jung, G. Mudd, S. Shave, H. Gstach, M. Auer, *Angew. Chem., Int. Ed.* **2014**, *53*, 4322.

[17] R. Huisgen, *Proc. Chem. Soc.* **1961**, 357.

[18] C. W. Tornoe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057.

[19] H. C. Kolb, K. B. Sharpless, *Drug Discovery Today* **2003**, *8*, 1128.

[20] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem., Int. Ed.* **2002**, *41*, 2596.

[21] M. Beija, C. A. M. Afonso, J. M. G. Martinho, *Chem. Soc. Rev.* **2009**, *38*, 2410.

[22] E. Atherton, C. J. Logan, R. C. Sheppard, *J. Chem. Soc., Perkin Trans. 1* **1981**, 538.

[23] M. Hintersteiner, T. Kimmerlin , G. Garavel, T. Schindler, R. Bauer, N.-C. Meisner, J.-M. Seifert, V. Uhl, M. Auer *ChemBioChem* **2009**, *10*, 994.

[24] M. Hintersteiner, A. J. Knox, G. Mudd, M. Auer, *J. Chem. Biol* **2012**, *5*, 63

[25] J. J. Babcock, M. Li, *Acta Pharmacol. Sin.* **2013**, *35*, 11.

[26] L. Fagerberg, K. Jonasson, G. von Heijne, M. Uhlén, L. Berglund, *Proteomics* **2010**, *10*, 1141.

[27] M. A. Yildirim, K.-I. Goh, M. E. Cusick, A.-L. Barabasi, M. Vidal, *Nat Biotech* **2007**, *25*, 1119.

[28] R. Fredriksson, M. C. Lagerström, L.-G. Lundin, H. B. Schiöth, *Mol. Pharmacol.* **2003**, *63*, 1256.

[29] T. K. Bjarnadottir, D. E. Gloriam, S. H. Hellstrand, H. Kristiansson, R. Fredriksson, H. B. Schioth, *Genomics* **2006**, *88*, 263.

[30] J. r. Drews, *Science* **2000**, *287*, 1960.

[31] N. Tuteja, *Plant Signaling Behav.* **2009**, *4*, 942.

[32] M. J. Marinissen, J. S. Gutkind, *Trends Pharmacol. Sci.* **2001**, *22*, 368.

[33] H. E. Hamm, *J. Biol. Chem.* **1998**, *273*, 669.

[34] R. N. Zuckermann, E. J. Martin, D. C. Spellmeyer, G. B. Stauber, K. R. Shoemaker, J. M. Kerr, G. M. Figliozzi, D. A. Goff, M. A. Siani, *J. Med. Chem.* **1994**, *37*, 2678.

[35] K. C. Appell, T. D. Y. Chung, M. J. H. Ohlmeyer, N. H. Sigal, J. J. Baldwin, D. Chelsky, *J. Biomol. Screening* **1996**, *1*, 27.

[36] G. Heizmann, P. Hildebrand, H. Tanner, S. Ketterer, A. Pansky, S. Froidevaux, C. Beglinger, A. N. Eberle, *J. Recept. Signal Transduction* **1999**, *19*, 449.

[37] B. Evans, A. Pipe, L. Clark, M. Banks, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1297.

[38] C. K. Jayawickreme, J. M. Quillan, G. F. Graminski, M. R. Lerner, *J. Biol. Chem.* **1994**, *269*, 29846.

[39] A. O. Iuga, V. B. Reddy, E. A. Lerner, *Peptides* **2005**, *26*, 2124.

[40] L. Peng, R. Liu, J. Marik, X. Wang, Y. Takada, K. S. Lam, *Nat Chem Biol* **2006**, *2*, 381.

[41] H. J. Olivos, K. Bachhawat-Sikder, T. Kodadek, *ChemBioChem* **2003**, *4*, 1242.

[42] D. G. Udugamasooriya, S. P. Dineen, R. A. Brekken, T. Kodadek, *J. Am. Chem. Soc.* **2008**, *130*, 5744.

[43] D. G. Udugamasooriya, T. Kodadek, *Curr. Protoc. Chem. Biol.* **2012**, *4*, 35.

[44] X. Qi, J. Astle, T. Kodadek, *Mol. BioSyst.* **2009**, *6*, 102.

[45] D. K. Lee, T. Nguyen, G. P. O'Neill, R. Cheng, Y. Liu, A.-T. Tang-Nguyen, S. R. George, B. F. O'Dowd, *FEBS Lett.* **1999**, *446*, 103.

[46] M. K. Clements, T. P. McDonald, R. Wang, G. Xie, B. F. O'Dowd, S. R. George, C. P. Austin, Q. Liu, *Biochem. Biophys. Res. Commun.* **2001**, *284*, 1189.

[47] J. H. Lee, D. R. Welch, *Cancer Res.* **1997**, *57*, 2384.

[48] J.-H. Lee, M. E. Miele, D. J. Hicks, K. K. Phillips, J. M. Trent, B. E. Weissman, D. R. Welch, *J. Natl. Cancer Inst.* **1996**, *88*, 1731.

[49] A. West, P. J. Vojta, D. R. Welch, B. E. Weissman, *Genomics* **1998**, *54*, 145.

[50] T. Ohtaki, Y. Shintani, S. Honda, H. Matsumoto, A. Hori, K. Kanehashi, Y. Terao, S. Kumano, Y. Takatsu, Y. Masuda, Y. Ishibashi, T. Watanabe, M. Asada, T. Yamada, M. Suenaga, C. Kitada, S. Usuki, T. Kurokawa, H. Onda, O. Nishimura, M. Fujino, *Nature* **2001**, *411*, 613.

[51] M. Kotani, M. Detheux, A. Vandenbogaerde, D. Communi, J. Vanderwinden, E. Le Poul, S. Brezillon, R. Tyldesley, N. Suarez-Huerta, F. Vandeput, C. Blanpain, S. N. Schiffmann, G. Vassart, M. Parmentier, *J. Biol. Chem.* **2001**, *276*, 34631.

[52] A. I. Muir, L. Chamberlain, N. A. Elshourbagy, D. Michalovich, D. J. Moore, A. Calamari, P. G. Szekeres, H. M. Sarau, J. K. Chambers, P. Murdock, K. Steplewski, U. Shabon, J. E. Miller, S. E. Middleton, J. G. Darker, C. G. C. Larminie, S. Wilson, D. J. Bergsma, P. Emson, R. Faull, K. L. Philpott, D. C. Harrison, *J. Biol. Chem.* **2001**, *276*, 28969.

[53] J. P. Castano, A. J. Martinez-Fuentes, E. Gutierrez-Pascual, H. Vaudry, M. Tena-Sempere, M. M. Malagon, *Peptides* **2009**, *30*, 10.

[54] A. Hori, S. Honda, M. Asada, T. Ohtaki, K. Oda, T. Watanabe, Y. Shintani, T. Yamada, M. Suenaga, C. Kitada, H. Onda, T. Kurokawa, O. Nishimura, M. Fujino, *Biochem. Biophys. Res. Commun.* **2001**, *286*, 958.

[55] S. Messager, *J. Neuroendocrinol.* **2005**, *17*, 687.

[56] S. B. Seminara, U. B. Kaiser, *Endocrinology* **2005**, *146*, 1686.

[57] N. de Roux, E. Genin, J.-C. Carel, F. Matsuda, J.-L. Chaussain, E. Milgrom, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 10972.

[58] S. B. Seminara, S. Messager, E. E. Chatzidaki, R. R. Thresher, J. S. Acierno, J. K. Shagoury, Y. Bo-Abbas, W. Kuohung, K. M. Schwinof, A. G. Hendrick, D. Zahn, J. Dixon, U. B. Kaiser, S. A. Slaugenhaupt, J. F. Gusella, S. O'Rahilly, M. B. L. Carlton, W. F. Crowley, S. A. J. R. Aparicio, W. H. Colledge, *N. Engl. J. Med.* **2003**, *349*, 1614.

[59] F. Lanfranco, J. Gromoll, E. S. von, E. M. Herding, E. Nieschlag, M. Simoni, *Eur. J. Endocrinol.* **2005**, *153*, 845.

[60] R. K. Semple, J. C. Achermann, J. Ellery, I. S. Farooqi, F. E. Karet, R. G. Stanhope, S. O'Rahilly, S. A. Aparicio, *J. Clin. Endocrinol. Metab.* **2005**, *90*, 1849.

[61] F. Cerrato, J. Shagoury, M. Kralickova, A. Dwyer, J. Falardeau, M. Ozata, V. G. Van, P. Bouloux, J. E. Hall, F. J. Hayes, N. Pitteloud, K. A. Martin, C. Welt, S. B. Seminara, *Eur. J. Endocrinol.* **2006**, *155*, S3.

[62] Y. Tenenbaum-Rakover, M. Commenges-Ducos, A. Iovane, C. Aumas, O. Admoni, R. N. de, *J. Clin. Endocrinol. Metab.* **2007**, *92*, 1137.

[63] R. Nimri, Y. Lebenthal, L. Lazar, L. Chevrier, M. Phillip, M. Bar, E. Hernandez-Mora, R. N. de, G. Gat-Yablonski, *J. Clin. Endocrinol. Metab.* **2011**, *96*, E536.

[64] S. Funes, J. A. Hedrick, G. Vassileva, L. Markowitz, S. Abbondanzo, A. Golovko, S. Yang, F. J. Monsma, E. L. Gustafson, *Biochem. Biophys. Res. Commun.* **2003**, *312*, 1357.

[65] X. d'Anglemont de Tassigny, L. A. Fagg, J. P. C. Dixon, K. Day, H. G. Leitch, A. G. Hendrick, D. Zahn, I. Franceschini, A. Caraty, M. B. L. Carlton,

S. A. J. R. Aparicio, W. H. Colledge, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10714.

[66] A. K. Topaloglu, J. A. Tello, L. D. Kotan, M. N. Ozbek, M. B. Yilmaz, S. Erdogan, F. Gurbuz, F. Temiz, R. P. Millar, B. Yuksel, *N. Engl. J. Med.* **2012**, *366*, 629.

[67] X. Luan, H. Yu, X. Wei, Y. Zhou, W. Wang, P. Li, X. Gan, D. Wei, J. Xiao, *Neuroendocrinology* **2007**, *86*, 77.

[68] M. G. Teles, S. D. C. Bianco, V. N. Brito, E. B. Trabach, W. Kuohung, S. Xu, S. B. Seminara, B. B. Mendonca, U. B. Kaiser, A. C. Latronico, *N. Engl. J. Med.* **2008**, *358*, 709.

[69] L. G. Silveira, S. D. Noel, A. P. Silveira-Neto, A. P. Abreu, V. N. Brito, M. G. Santos, S. D. C. Bianco, W. Kuohung, S. Xu, M. Gryngarten, M. E. Escobar, I. J. P. Arnhold, B. B. Mendonca, U. B. Kaiser, A. C. Latronico, *J. Clin. Endocrinol. Metab.* **2010**, *95*, 2276.

[70] A. Akinci, D. Cetin, N. Ilhan, *J. Clin. Res. Pediatr. Endocrinol.* **2012**, *4*, 60.

[71] S. Messager, E. E. Chatzidaki, D. Ma, A. G. Hendrick, D. Zahn, J. Dixon, R. R. Thresher, I. Malinge, D. Lomet, M. B. L. Carlton, W. H. Colledge, A. Caraty, S. A. J. R. Aparicio, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 1761.

[72] A. K. Roseweir, A. S. Kauffman, J. T. Smith, K. A. Guerriero, K. Morgan, J. Pielecka-Fortuna, R. Pineda, M. L. Gottsch, M. Tena-Sempere, S. M. Moenter, E. Terasawa, I. J. Clarke, R. A. Steiner, R. P. Millar, *J. Neurosci.* **2009**, *29*, 3920.

[73] M. L. Gottsch, M. J. Cunningham, J. T. Smith, S. M. Popa, B. V. Acohido, W. F. Crowley, S. Seminara, D. K. Clifton, R. A. Steiner, *Endocrinology* **2004**, *145*, 4073.

[74] M. S. Irwig, G. S. Fraley, J. T. Smith, B. V. Acohido, S. M. Popa, M. J. Cunningham, M. L. Gottsch, D. K. Clifton, R. A. Steiner, *Neuroendocrinology* **2004**, *80*, 264.

[75] C. N. Jayasena, G. M. K. Nijher, O. B. Chaudhri, K. G. Murphy, A. Ranger, A. Lim, D. Patel, A. Mehta, C. Todd, R. Ramachandran, V. Salem, G. W. Stamp, M. Donaldson, M. A. Ghatei, S. R. Bloom, W. S. Dhillo, *J. Clin. Endocrinol. Metab.* **2009**, *94*, 4315.

[76] W. S. Dhillo, O. B. Chaudhri, E. L. Thompson, K. G. Murphy, M. Patterson, R. Ramachandran, G. K. Nijher, V. Amber, A. Kokkinos, M. Donaldson, M. A. Ghatei, S. R. Bloom, *J. Clin. Endocrinol. Metab.* **2007**, *92*, 3958.

[77] C. N. Jayasena, G. M. K. Nijher, A. Abbara, K. G. Murphy, A. Lim, D. Patel, A. Mehta, C. Todd, M. Donaldson, G. H. Trew, M. A. Ghatei, S. R. Bloom, W. S. Dhillo, *Clin. Pharmacol. Ther.* **2010**, *88*, 840.

[78] C. N. Jayasena, A. Abbara, A. N. Comninos, G. M. K. Nijher, G. Christopoulos, S. Narayanaswamy, C. Izzi-Engbeaya, M. Sridharan, A. J. Mason, J. Warwick, D. Ashby, M. A. Ghatei, S. R. Bloom, A. Carby, G. H. Trew, W. S. Dhillo, *J. Clin. Invest.* **2014**, *124*, 3667.

[79] W. S. Dhillo, O. B. Chaudhri, M. Patterson, E. L. Thompson, K. G. Murphy, M. K. Badman, B. M. McGowan, V. Amber, S. Patel, M. A. Ghatei, S. R. Bloom, *J. Clin. Endocrinol. Metab.* **2005**, *90*, 6609.

[80] J. T. George, J. D. Veldhuis, A. K. Roseweir, C. L. Newton, E. Faccenda, R. P. Millar, R. A. Anderson, *J. Clin. Endocrinol. Metab.* **2011**, *96*, E1228.

[81] Y. Lyubimov, M. Engstrom, S. Wurster, J. M. Savola, E. R. Korpi, P. Panula, *Neuroscience* **2010**, *170*, 117.

[82] K. P. Tolson, C. Garcia, S. Yen, S. Simonds, A. Stefanidis, A. Lawrence, J. T. Smith, A. S. Kauffman, *J. Clin. Invest.* **2014**, *124*, 3075.

[83] J. L. Calley, W. S. Dhillo, *Advances in Biology* **2014**, *2014*, 10.

[84] M. K. Clements, T. P. McDonald, R. Wang, G. Xie, B. F. O'Dowd, S. R. George, C. P. Austin, Q. Liu, *Biochem. Biophys. Res. Commun.* **2001**, *284*, 1189.

[85] M. J. Orsini, M. A. Klein, M. P. Beavers, P. J. Connolly, S. A. Middleton, K. H. Mayo, *J. Med. Chem.* **2007**, *50*, 462.

[86] R. Shin, D. Welch, V. Mishra, K. Nash, D. Hurst, N. Rama Krishna, *Clin. Exp. Metastasis* **2009**, *26*, 7.

[87] E. Gutiérrez-Pascual, J. Leprince, A. J. Martínez-Fuentes, I. Ségalas-Milazzo, R. Pineda, J. Roa, M. Duran-Prado, L. Guilhaudis, E. Desperrois, A. Lebreton, L. Pinilla, M.-C. Tonon, M. M. Malagón, H. Vaudry, M. Tena-Sempere, J. P. Castaño, *Mol. Pharmacol.* **2009**, *76*, 58.

[88] A. Niida, Z. Wang, K. Tomita, S. Oishi, H. Tamamura, A. Otaka, J.-M. Navenot, J. R. Broach, S. C. Peiper, N. Fujii, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 134.

[89] K. Tomita, A. Niida, S. Oishi, H. Ohno, J. Cluzeau, J.-M. Navenot, Z.-x. Wang, S. C. Peiper, N. Fujii, *Bioorg. Med. Chem.* **2006**, *14*, 7595.

[90] K. Tomita, S. Oishi, J. r. m. Cluzeau, H. Ohno, J.-M. Navenot, Z.-x. Wang, S. C. Peiper, M. Akamatsu, N. Fujii, *J. Med. Chem.* **2007**, *50*, 3222.

[91] K. Tomita, S. Oishi, H. Ohno, N. Fujii, *Pept. Sci.* **2008**, *90*, 503.

[92] K. Tomita, T. Narumi, A. Niida, y. Oishi, H. Ohno, N. Fujii, *Synthesis and Application of (Z)-Alkene- and (E)-Fluoroalkene-Dipeptide Isosteres as cis-Amide Equivalents, Vol. 611*, Springer New York, **2009**.

[93] K. Tomita, S. Oishi, H. Ohno, S. C. Peiper, N. Fujii, *J. Med. Chem.* **2008**, *51*, 7645.

[94] S. Oishi, R. Misu, K. Tomita, S. Setsuda, R. Masuda, H. Ohno, Y. Naniwa, N. Ieda, N. Inoue, S. Ohkura, Y. Uenoyama, H. Tsukamura, K.-i. Maeda, A. Hirasawa, G. Tsujimoto, N. Fujii, *ACS Med. Chem. Lett.* **2010**, *2*, 53.

[95] M. Kaneda, R. Misu, H. Ohno, A. Hirasawa, N. Ieda, Y. Uenoyama, H. Tsukamura, K.-i. Maeda, S. Oishi, N. Fujii, *Bioorg. Med. Chem.* **2014**.

[96] T. Asami, N. Nishizawa, Y. Ishibashi, K. Nishibori, M. Nakayama, Y. Horikoshi, S.-i. Matsumoto, M. Yamaguchi, H. Matsumoto, N. Tarui, T. Ohtaki, C. Kitada, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6391.

[97] T. Asami, N. Nishizawa, Y. Ishibashi, K. Nishibori, Y. Horikoshi, H. Matsumoto, T. Ohtaki, C. Kitada, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6328.

[98] A. E. Curtis, J. H. Cooke, J. E. Baxter, J. R. C. Parkinson, A. Bataveljic, M. A. Ghatei, S. R. Bloom, K. G. Murphy, *Endocrinol. Metab.* **2010**, *298*, E296.

[99] G. Scott, I. Ahmad, K. Howard, D. MacLean, C. Oliva, S. Warrington, D. Wilbraham, P. Worthington, *Br. J. Clin. Pharmacol.* **2013**, *75*, 381.

[100] H. Matsui, T. Masaki, Y. Akinaga, A. Kiba, D. Nakata, A. Tanaka, T. Watanabe, T. Ohtaki, M. Kusaka, *EJC Supplements* **2010**, *8*, 205.

[101] A. K. Tickler, J. D. Wade, *Overview of Solid Phase Synthesis of "Difficult Peptide" Sequences*, John Wiley & Sons, Inc., **2001**.

[102] L. A. Carpino, G. Y. Han, *J. Org. Chem.* **1972**, *37*, 3404.

[103] E. Atherton, H. Fox, D. Harkiss, C. J. Logan, R. C. Sheppard, B. J. Williams, *J. Chem. Soc., Chem. Commun.* **1978**, 537.

[104] C. D. Chang, J. Meienhofer, *Int. J. Pept. Protein. Res.* **1978**, *11*, 246.

[105] H. Rink, *Tetrahedron Lett.* **1987**, *28*, 3787.

[106] P. E. Brandish, L. A. Hill, W. Zheng, E. M. Scolnick, *Anal. Biochem.* **2003**, *313*, 311.

[107] W. R. Huckle, P. M. Conn, A. R. M. P. Michael Conn, *Use of lithium ion in measurement of stimulated pituitary inositol phospholipid turnover, Vol. 141*, Academic Press, **1987**.

[108] M. Pampillo, N. Camuso, J. E. Taylor, J. M. Szereszewski, M. R. Ahow, M. Zajac, R. P. Millar, M. Bhattacharya, A. V. Babwah, *Mol. Endocrinol.* **2009**, *23*, 2060.

[109] K. R. Gee, K. A. Brown, W. N. U. Chen, J. Bishop-Stewart, D. Gray, I. Johnson, *Cell Calcium* **2000**, *27*, 97.

[110] W. A. Bonner, H. R. Hulett, R. G. Sweet, L. A. Herzenberg, *Rev. Sci. Instrum.* **1972**, *43*, 404.

[111] M. H. Julius, T. Masuda, L. A. Herzenberg, *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 1934.

[112] E. D. Goddard-Borger, R. V. Stick, *Org. Lett.* **2007**, *9*, 3797.

[113] B. D. Gray, *Fluorescent membrane intercalating probes and methods for their use*. US 2003/0223935 A1, **2003**.

[114] Thermo-Fisher-Scientific, http://www.lifetechnologies.com/, **2014**.

[115] J. W. Jones, T. A. Greene, C. A. Grygon, B. J. Doranz, M. P. Brown, *J. Biomol. Screening* **2008**, *13*, 424.

[116] H. Tamamura, Y. Xu, T. Hattori, X. Zhang, R. Arakaki, K. Kanbara, A. Omagari, A. Otaka, T. Ibuka, N. Yamamoto, H. Nakashima, N. Fujii, *Biochem. Biophys. Res. Commun.* **1998**, *253*, 877.

[117] C. G. Tate, *FEBS Lett.* **2001**, *504*, 94.

[118] A. M. Seddon, P. Curnow, P. J. Booth, *Biochim. Biophys. Acta, Biomembr.* **2004**, *1666*, 105.

[119] R. Kalmbach, I. Chizhov, M. C. Schumacher, T. Friedrich, E. Bamberg, M. Engelhard, *J. Mol. Biol.* **2007**, *371*, 639.

[120] F. Katzen, J. E. Fletcher, J.-P. Yang, D. Kang, T. C. Peterson, J. A. Cappuccio, C. D. Blanchette, T. Sulchek, B. A. Chromy, P. D. Hoeprich, M. A. Coleman, W. Kudlicki, *J. Proteome Res.* **2008**, *7*, 3535.

[121] S. Doyle, J. Cappuccio, A. Hinz, E. Kuhn, J. Fletcher, E. Arroyo, P. Henderson, C. Blanchette, V. Walsworth, M. Corzett, R. Law, J. Pesavento, B. Segelke, T. Sulchek, B. Chromy, F. Katzen, T. Peterson, G. Bench, W. Kudlicki, P. Hoeprich, Jr., M. Coleman, *Cell-Free Expression for Nanolipoprotein Particles: Building a High-Throughput Membrane Protein Solubility Platform, Vol. 498*, Humana Press, **2009**.

[122] J.-P. Yang, T. Cirico, F. Katzen, T. Peterson, W. Kudlicki, *BMC Biotechnol.* **2011**, *11*, 57.

[123] T. Gao, J. Petrlova, W. He, T. Huser, W. Kudlick, J. Voss, M. A. Coleman, *PLoS ONE* **2012**, *7*, e44911.

[124] T. Gao, C. D. Blanchette, W. He, F. Bourguet, S. Ly, F. Katzen, W. A. Kudlicki, P. T. Henderson, T. A. Laurence, T. Huser, M. A. Coleman, *Protein Sci.* **2011**, *20*, 437.

[125] S. Radestock, T. Weil, S. Renner, *J. Chem. Inf. Model.* **2008**, *48*, 1104.

[126] K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. L. Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, *Science* **2000**, *289*, 739.

[127] M. Congreve, C. J. Langmead, J. S. Mason, F. H. Marshall, *J. Med. Chem.* **2011**, *54*, 4283.

[128] S. G. F. Rasmussen, H.-J. Choi, D. M. Rosenbaum, T. S. Kobilka, F. S. Thian, P. C. Edwards, M. Burghammer, V. R. P. Ratnala, R. Sanishvili, R. F. Fischetti, G. F. X. Schertler, W. I. Weis, B. K. Kobilka, *Nature* **2007**, *450*, 383.

[129] B. Wu, E. Y. T. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov, R. C. Stevens, *Science* **2010**, *330*, 1066.

[130] M. Marti-Renom, A. C. Stuart, A. Fiser, R. Sanchez, F. Melo, A. Sali, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291.

[131] J. Peng, J. Xu, *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 1930.

[132] R. Sanchez, A. Sali, *Curr. Opin. Struct. Biol.* **1997**, *7*, 206.

[133] C. N. Cavasotto, S. S. Phatak, *Drug Discovery Today* **2009**, *14*, 676.

[134] M. Findeisen, D. Rathmann, A. G. Beck-Sickinger, *Pharmaceuticals* **2011**, *4*, 1248.

[135] J. A. Bonini, K. A. Jones, N. Adham, C. Forray, R. Artymyshyn, M. M. Durkin, K. E. Smith, J. A. Tamm, L. W. Boteju, P. P. Lakhlani, R. Raddatz, W.-J. Yao, K. L. Ogozalek, N. Boyle, E. V. Kouranova, Y. Quan, P. J. Vaysse, J. M. Wetzel, T. A. Branchek, C. Gerald, B. Borowsky, *J. Biol. Chem.* **2000**, *275*, 39324.

[136] H. Y. Yang, W. Fratta, E. A. Majane, E. Costa, *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 7757.

[137] E. Ducret, G. M. Anderson, A. E. Herbison, *Endocrinology* **2009**, *150*, 2799.

[138] Navarro, F. Fernandez, Nogueiras, Vigo, Tovar, Chartrel, M. Le, Leprince, Aguilar, Pinilla, Dieguez, Vaudry, S. Tena, *J Physiol.* **2006**, *573*, 237

[139] S. Takayasu, T. Sakurai, S. Iwasaki, H. Teranishi, A. Yamanaka, S. C. Williams, H. Iguchi, Y. I. Kawasawa, Y. Ikeda, I. Sakakibara, K. Ohno, R. X.

Ioka, S. Murakami, N. Dohmae, J. Xie, T. Suda, T. Motoike, T. Ohuchi, M. Yanagisawa, J. Sakai, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 7438

[140] Q. Fang, Q. Liu, N. Li, T.-n. Jiang, Y.-l. Li, X. Yan, R. Wang, *Eur. J. Pharmacol.* **2009**, *621*, 61

[141] S. Hinuma, Y. Habata, R. Fujii, KawamataYuji, M. Hosoya, S. Fukusumi, C. Kitada, Y. Masuo, T. Asano, H. Matsumoto, M. Sekiguchi, T. Kurokawa, O. Nishimura, H. Onda, M. Fujino, *Nature* **1998**, *393*, 272.

[142] W. Gu, B. J. Geddes, C. Zhang, K. P. Foley, A. Stricker-Krongrad, *J. Mol. Neurosci.* **2004**, *22*, 93

[143] M. Maruyama, H. Matsumoto, K. Fujiwara, J. Noguchi, C. Kitada, M. Fujino, K. Inoue, *Endocrinology* **2001**, *142*, 2032

[144] J. Roa, E. Aguilar, C. Dieguez, L. Pinilla, M. Tena-Sempere, *Front. Neuroendocrinol.* **2008**, *29*, 48.

[145] A. K. Roseweir, R. P. Millar, *Hum. Reprod. Update* **2009**, *15*, 203.

[146] M. A. Cline, D. C. Godlove, W. Nandar, C. N. Bowden, B. C. Prall, *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.* **2007**, *148*, 657.

[147] S. Takayasu, T. Sakurai, S. Iwasaki, H. Teranishi, A. Yamanaka, S. C. Williams, H. Iguchi, Y. I. Kawasawa, Y. Ikeda, I. Sakakibara, K. Ohno, R. X. Ioka, S. Murakami, N. Dohmae, J. Xie, T. Suda, T. Motoike, T. Ohuchi, M. Yanagisawa, J. Sakai, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 7438.

[148] H. Mazarguil, C. Gouardères, J.-A. M. Tafani, D. Marcus, M. Kotani, C. Mollereau, M. Roumy, J.-M. Zajac, *Peptides* **2001**, *22*, 1471.

[149] K. Payza, C. A. Akar, H. Y. Yang, *J. Pharmacol. Exp. Ther.* **1993**, *267*, 88.

[150] S. Fukusumi, H. Yoshida, R. Fujii, M. Maruyama, H. Komatsu, Y. Habata, Y. Shintani, S. Hinuma, M. Fujino, *J. Biol. Chem.* **2003**, *278*, 46387.

[151] O. Le Marec, C. Neveu, B. Lefranc, C. Dubessy, J. A. Boutin, J.-C. Do-Régo, J. Costentin, M.-C. Tonon, M. Tena-Sempere, H. Vaudry, J. r. m. Leprince, *J. Med. Chem.* **2011**, *54*, 4806.

[152] S. Hinuma, R. Fujii, M. Fujino, S. Fukusumi, Y. Habata, M. Hosoya, Y. Kawamata, C. Kitada, T. Kurokawa, Y. Masuo, H. Matsumoto, H. Onda, T. Asano, O. Nishimura, M. Sekiguchi, *Nature* **1998**, *393*, 272

[153] R. G. Boyle, R. Downham, T. Ganguly, J. Humphries, J. Smith, S. Travers, *J. Pept. Sci.* **2005**, *11*, 161.

[154] E. Escher, R. Couture, C. Poulos, N. Pinas, J. Mizrahi, et al., *J. Med. Chem.* **1982**, *25*, 1317

[155] A. M. Aronov, M. H. Gelb, *Tetrahedron Lett.* **1998**, *39*, 4947.

[156] J. Bauer, J. Rademann, *Tetrahedron Lett.* **2003**, *44*, 5019.

[157] S. Manku, C. Laplante, D. Kopac, T. Chan, D. G. Hall, *J. Org. Chem.* **2001**, *66*, 874.

[158] M. Schuster, J. Pernerstorfer, S. Blechert, *Angew. Chem., Int. Ed.* **1996**, *35*, 1979.

[159] Y. Guan, M. A. Green, D. E. Bergstrom, *J. Comb. Chem.* **2000**, *2*, 297.

[160] C. M. Dreef-Tromp, P. Hoogerhout, G. A. van der Marel, J. H. van Boom, *Tetrahedron Lett.* **1990**, *31*, 427.

[161] V. Castro, H. Rodriguez, F. Albericio, *Org. Lett.* **2012**, *15*, 246.

[162] T. Guerlavais-Dagland, A. Meyer, F. Morvan, *J. Chem. Res., Synop.* **2002**, *2002*, 606.

[163] F. Starke, M. Walther, H.-J. Pietzsch, *ARKIVOC (Gainesville, FL, U. S.)* **2010**, *11*, 350

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

[165] L. N. Chan, C. Hart, L. Guo, T. Nyberg, B. S. J. Davies, L. G. Fong, S. G. Young, B. J. Agnew, F. Tamanoi, *Electrophoresis* **2009**, *30*, 3598.

[166] B. Mullah, A. Andrus, *Tetrahedron Lett.* **1997**, *38*, 5751.

[167] W. Ansorge, B. S. Sproat, J. Stegemann, C. Schwager, *J. Biochem. Biophys. Methods* **1986**, *13*, 315.

[168] J. Corrie, *J. Chem. Soc., Perkin Trans. 1* **1994**, 2967.

[169] A. Brunet, T. Aslam, M. Bradley, *Bioorg. Med. Chem. Lett.* **2014**.

[170] S. M. Menchen, S. Fung, *5- and 6-Succinimidylcarboxylate Isomers of Rhodamine Dyes*, **1991**.

[171] M. V. Kvach, I. A. Stepanova, I. A. Prokhorenko, A. P. Stupak, D. A. Bolibrukh, V. A. Korshun, V. V. Shmanai, *Bioconjugate Chem.* **2009**, *20*, 1673.

[172] H. Yu, Y. Xiao, H. Guo, *Org. Lett.* **2012**, *14*, 2014.

[173] J. B. Grimm, L. D. Lavis, *Org. Lett.* **2011**, *13*, 6354.

[174] G.-S. Jiao, J. C. Castro, L. H. Thoresen, K. Burgess, *Org. Lett.* **2003**, *5*, 3675.

[175] D. R. Buckle, *Chloranil*, John Wiley & Sons, Ltd, **2001**.

[176] I. C. S. Cardoso, A. L. Amorim, C. Queirós, S. C. Lopes, P. Gameiro, B. de Castro, M. Rangel, A. M. G. Silva, *Eur. J. Org. Chem.* **2012**, *2012*, 5810.

[177] Shriner, Wolf, *Org. Synth.* **1943**, *23*, 74.

[178] V. T. Abaev, A. S. Dmitriev, S. A. Podelyakin, A. V. Butin, A. V. Gutnov, *J. Heterocycl. Chem.* **2006**, *43*, 1195.

[179] D. S. Wilbur, M.-K. Chyan, D. K. Hamlin, B. B. Kegley, R. Nilsson, B. E. B. Sandberg, M. Brechbiel, *Bioconjugate Chem.* **2002**, *13*, 1079.

[180] R. Yan, E. El-Emir, V. Rajkumar, M. Robson, A. P. Jathoul, R. B. Pedley, E. Årstad, *Angew. Chem., Int. Ed.* **2011**, *50*, 6793.

[181] A. P. Frei, O. Y. Jeon, S. Kilcher, H. Moest, L. M. Henning, C. Jost, A. Pluckthun, J. Mercer, R. Aebersold, E. M. Carreira, B. Wollscheid, *Nat Biotechnol* **2012**, *30*, 997.

[182] F. Stieber, U. Grether, H. Waldmann, *Chem. - Eur. J.* **2003**, *9*, 3270.

[183] G. Clave, H. Boutal, A. Hoang, F. Perraut, H. Volland, P.-Y. Renard, A. Romieu, *Org. Biomol. Chem.* **2008**, *6*, 3065.

[184] W. P. Heal, M. H. Wright, E. Thinon, E. W. Tate, *Nat. Protoc.* **2012**, *7*, 105.

[185] A. Watzke, M. Köhn, M. Gutierrez-Rodriguez, R. Wacker, H. Schröder, R. Breinbauer, J. Kuhlmann, K. Alexandrov, C. M. Niemeyer, R. S. Goody, H. Waldmann, *Angew. Chem., Int. Ed.* **2006**, *45*, 1408.

[186] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007.

[187] E. Garanger, E. Aikawa, F. Reynolds, R. Weissleder, L. Josephson, *Chem. Commun.* **2008**, *39*, 4792.

[188] E. Garanger, R. Weissleder, L. Josephson, *Bioconjugate Chem.* **2008**, *20*, 170.

[189] G. Clave, H. Volland, M. Flaender, D. Gasparutto, A. Romieu, P.-Y. Renard, *Org. Biomol. Chem.* **2010**, *8*, 4329.

[190] G. Viault, S. Dautrey, N. Maindron, J. Hardouin, P.-Y. Renard, A. Romieu, *Org. Biomol. Chem.* **2013**, *11*, 2693.

[191] K. J. Jensen, J. Alsina, M. F. Songster, J. Vágner, F. Albericio, G. Barany, *J. Am. Chem. Soc.* **1998**, *120*, 5441.

[192] J. J. Díaz-Mochón, L. Bialy, M. Bradley, *Org. Lett.* **2004**, *6*, 1127.

[193] C.-W. Lin, T.-N. Huang, G.-S. Wang, T.-Y. Kuo, T.-Y. Yen, Y.-P. Hsueh, *J. Comp. Neurol.* **2006**, *494*, 606.

[194] J. Keith, A. Ilari, C. Savino, in *Bioinformatics, Vol. 452*, Humana Press, **2008**, pp. 63.

[195] M. Billeter, G. Wagner, K. WÃ¼thrich, *Journal of biomolecular NMR* **2008**, *42*, 155.

[196] B. Schuler, *Journal of Nanobiotechnology* **2013**, *11*, S2.

[197] L. Berliner, G. Eaton, S. Eaton, E. Hustedt, A. Beth, in *Distance Measurements in Biological Systems by EPR, Vol. 19*, Springer US, **2002**, pp. 155.

[198] J. Rappsilber, *J. Struct. Biol.* **2011**, *173*, 530.

[199] M. Karas, D. Bachmann, F. Hillenkamp, *Anal. Chem.* **1985**, *57*, 2935.

[200] M. Karas, F. Hillenkamp, *Anal. Chem.* **1988**, *60*, 2299.

[201] R. C. Beavis, B. T. Chait, H. M. Fales, *Rapid Commun. Mass Spectrom.* **1989**, *3*, 432.

[202] M. Yamashita, J. B. Fenn, *J. Phys. Chem.* **1984**, *88*, 4451.

[203] G. T. Hermanson, *Bioconjugate Techniques*, 2nd ed., Academic Press, New York, **2008**.

[204] A. Sinz, *J. Mass Spectrom.* **2003**, *38*, 1225.

[205] G. T. Hermanson, *Bioconjugate Techniques*, 2nd ed., Academic Press, New York, **2008**.

[206] G. Dorman, G. D. Prestwich, *Biochemistry* **1994**, *33*, 5661.

[207] G. T. Hermanson, *Bioconjugate Techniques*, 2nd ed., Academic Press, New York, **2008**.

[208] S. C. Alley, F. T. Ishmael, A. D. Jones, S. J. Benkovic, *J. Am. Chem. Soc.* **2000**, *122*, 6126.

[209] A. Sinz, S. Kalkhof, C. Ihling, *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1921.

[210] G. B. Hurst, T. K. Lankford, S. J. Kennel, *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 832.

[211] P. T. Kasper, J. W. Back, M. Vitale, A. F. Hartog, W. Roseboom, L. J. de Koning, J. H. van Maarseveen, A. O. Muijsers, C. G. de Koster, L. de Jong, *ChemBioChem* **2007**, *8*, 1281.

[212] M. A. Nessen, G. Kramer, J. Back, J. M. Baskin, L. E. J. Smeenk, L. J. de Koning, J. H. van Maarseveen, L. de Jong, C. R. Bertozzi, H. Hiemstra, C. G. de Koster, *J. Proteome Res.* **2009**, *8*, 3702.

[213] S. M. Chowdhury, X. Du, N. Tolic, S. Wu, R. J. Moore, M. U. Mayer, R. D. Smith, J. N. Adkins, *Anal. Chem.* **2009**, *81*, 5524.

[214] R. M. Kaake, X. Wang, A. Burke, C. Yu, W. Kandur, Y. Yang, E. J. Novtisky, T. Second, J. Duan, A. Kao, S. Guan, D. Vellucci, S. D. Rychnovsky, L. Huang, *Mol. Cell. Proteomics* **2014**.

[215] A. Belsom, "*Exploiting reversible interactions: hydrogels and protein cross-linkers*", PhD. Thesis, University of Edinburgh (Edinburgh), **2010**.

[216] H. Eggert, J. Frederiksen, C. Morin, J. C. Norrild, *J. Org. Chem.* **1999**, *64*, 3846.F

[217] S. Friedman, R. Pizer, *J. Am. Chem. Soc.* **1975**, *97*, 6059.

[218] R. Pizer, R. Selzer, *Inorg. Chem.* **1984**, *23*, 3023.

[219] K. Kustin, R. Pizer, *J. Am. Chem. Soc.* **1969**, *91*, 317.

[220] L. Babcock, R. Pizer, *Inorg. Chem.* **1980**, *19*, 56.

[221] J. P. Lorand, J. O. Edwards, *J. Org. Chem.* **1959**, *24*, 769.

[222] R. Pizer, C. Tihal, *Inorg. Chem.* **1992**, *31*, 3243.

[223] S. Higa, S. Kishimoto, *Anal. Biochem.* **1986**, *154*, 71.

[224] H. L. Weith, J. L. Wiebers, P. T. Gilham, *Biochemistry* **1970**, *9*, 4396.

[225] W. Li, K. Burgess, *Tetrahedron Lett.* **1999**, *40*, 6527.

[226] T. Arnauld, A. G. M. Barrett, R. Seifried, *Tetrahedron Lett.* **2001**, *42*, 7899.

[227] M. L. Stolowitz, C. Ahlem, K. A. Hughes, R. J. Kaiser, E. A. Kesicki, G. Li, K. P. Lund, S. M. Torkelson, J. P. Wiley, *Bioconjugate Chem.* **2001**, *12*, 229.

[228] W. F. Veldhuyzen, Q. Nguyen, G. McMaster, D. S. Lawrence, *J. Am. Chem. Soc.* **2003**, *125*, 13358.

[229] T. Kobayashi, T. Komatsu, M. Kamiya, C. Campos, M. González-Gaitán, T. Terai, K. Hanaoka, T. Nagano, Y. Urano, *J. Am. Chem. Soc.* **2012**, *134*, 11153.

[230] A. Bapna, E. Vickerstaffe, B. H. Warrington, M. Ladlow, T.-P. D. Fan, S. V. Ley, *Org. Biomol. Chem.* **2004**, *2*, 611.

[231] G. M. Burslem, A. J. Wilson, *Synlett* **2014**, *25*, 324.

[232] J. T. Weiss, J. C. Dawson, C. Fraser, W. Rybski, C. Torres-Sánchez, M. Bradley, E. E. Patton, N. O. Carragher, A. Unciti-Broceta, *J. Med. Chem.* **2014**, *57*, 5395.

[233] A. Unciti-Broceta, E. M. V. Johansson, R. M. Yusop, R. M. Sanchez-Martin, M. Bradley, *Nat. Protocols* **2012**, *7*, 1207.

[234] J. T. Weiss, J. C. Dawson, K. G. Macleod, W. Rybski, C. Fraser, C. Torres-Sánchez, E. E. Patton, M. Bradley, N. O. Carragher, A. Unciti-Broceta, *Nat Commun* **2014**, *5*.

[235] K. R. Gee, E. A. Archer, H. C. Kang, *Tetrahedron Lett.* **1999**, *40*, 1471.

[236] J. Gariépy, G. K. Schoolnik, *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 483.

[237] C. Berens, P. J. Courtoy, E. Sonveaux, *Bioconjugate Chem.* **1998**, *10*, 56.

[238] P. C. D. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson, M. T. Stahl, *J. Chem. Inf. Model.* **2010**, *50*, 572.

[239] Z. A. Chen, A. Jawhari, L. Fischer, C. Buchen, S. Tahir, T. Kamenski, M. Rasmussen, L. Lariviere, J. C. Bukowski-Wills, M. Nilges, P. Cramer, J. Rappsilber, *The EMBO Journal* **2010**, *29*, 717.

[240] P. F. Spahr, J. T. Edsall, *Journal of Biological Chemistry* **1964**, *239*, 850.

[241] X. X. Liu, A. Melman, *Chem. Commun.* **2013**, *49*, 9042.

[242] K. Lang, J. W. Chin, *Chemical Reviews* **2014**, *114*, 4764.

[243] C. Byrne, P. A. McEwan, J. Emsley, P. M. Fischer, W. C. Chan, *Chem. Commun.* **2011**, *47*, 2589.

[244] J. L. Wacker, D. B. Feller, X.-B. Tang, M. C. DeFino, Y. Namkung, J. S. Lyssand, A. J. Mhyre, X. Tan, J. B. Jensen, C. Hague, *J. Biol. Chem.* **2008**, *283*, 31068.

[245] R. Fredriksson, M. C. Lagerström, L.-G. Lundin, H. B. Schiöth, *Molecular Pharmacology* **2003**, *63*, 1256.

[246] M. Olivella, A. Gonzalez, L. Pardo, X. Deupi, *Bioinformatics* **2013**, *29*, 1589.

[247] F. Sievers, D. Dineen, D. G. Higgins, A. Wilm, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soeding, J. D. Thompson, *Molecular Systems Biology* **2011**, *7*.

[248] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *Journal of Computational Chemistry* **2004**, *25*, 1605.

[249] J. Li, P. C. Edwards, M. Burghammer, C. Villa, G. F. X. Schertler, *Journal of Molecular Biology* **2004**, *343*, 1409.

[250] N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M.-y. Shen, U. Pieper, A. Sali, in *Current Protocols in Protein Science*, John Wiley & Sons, Inc., **2001**.

[251] T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe, N. A. Baker, *Nucleic acids research* **2007**, *35*, W522.

[252] L. Pardo, X. Deupi, N. Dölker, M. L. López-Rodríguez, M. Campillo, *ChemBioChem* **2007**, *8*, 19.

[253] T. E. Angel, K. Palczewski, M. R. Chance, *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 8555

[254] H. Li, A. D. Robertson, J. H. Jensen, *Proteins: Structure, Function, and Bioinformatics* **2005**, *61*, 704.

[255] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *Journal of Computational Chemistry* **2005**, *26*, 1781.

[256] A. D. MacKerell, D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, *The Journal of Physical Chemistry B* **1998**, *102*, 3586.

[257] P. T. Lang, S. R. Brozell, S. Mukherjee, E. F. Pettersen, E. C. Meng, V. Thomas, R. C. Rizzo, D. A. Case, T. L. James, I. D. Kuntz, *RNA* **2009**, *15*, 1219.

[258] N. London, B. Raveh, E. Cohen, G. Fathi, O. Schueler-Furman, *Nucleic Acids Res.* **2011**, *39*, 249.

[259] B. Raveh, N. London, L. Zimmerman, O. Schueler-Furman, *PLoS ONE* **2011**, *6*, e18934.

[260] T. J. Ewing, S. Makino, A. G. Skillman, I. D. Kuntz, *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411.

[261] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* **2006**, *65*, 712.

[262] K. C. Tsai, S. H. Wang, N. W. Hsiao, M. Li, B. Wang, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3509.

[263] V. B. D.A. Case, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu and P.A. Kollman, 14 ed., University of California, San Francisco, **2014**.

[264] Arwin J. Brouwer, Suzanne J. E. Mulders, Rob M. J. Liskamp, *Eur. J. Org. Chem.* **2001**, *2001*, 1903.

[265] H. Sajiki, K. Y. Ong, *Tetrahedron* **1996**, *52*, 14507.

[266] P.-Y. Yang, H. Wu, M. Y. Lee, A. Xu, R. Srinivasan, S. Q. Yao, *Org. Lett.* **2008**, *10*, 1881.

[267] A. Boeijen, J. van Ameijde, R. M. J. Liskamp, *J. Org. Chem.* **2001**, *66*, 8454.

[268] K. D. Park, P. Morieux, C. Salomé, S. W. Cotten, O. Reamtong, C. Eyers, S. J. Gaskell, J. P. Stables, R. Liu, H. Kohn, *J. Med. Chem.* **2009**, *52*, 6897.

[269] L. F. Levy, H. Stephen, *J. Chem. Soc.* **1931**, *0*, 867.

[270] T. Vojkovsky, *Pept. Res.* **1995**, *8*, 236.

[271] B. W. Bycroft, W. C. Chan, S. R. Chhabra, N. D. Hone, *J. Chem. Soc., Chem. Commun.* **1993**, 778.

[272] V. Kumar, J. V. Aldrich, *Org. Lett.* **2003**, *5*, 613.

[273] M. J. Liley, T. Johnson, S. E. Gibson, *J. Org. Chem.* **2006**, *71*, 1322.

[274] S. Lata, A. Reichel, R. Brock, R. Tampe, J. Piehler, *J. Am. Chem. Soc.* **2005**, *127*, 10205.

[275] T. Terai, H. Ito, K. Kikuchi, T. Nagano, *Chem. - Eur. J.* **2012**, *18*, 7377.

[276] P. Desreumaux, S. Bellinvia, P. Chavatte, S. Baroni, *Multitarget Compounds Active at a PPAR and Cannabinoid Receptor*, **2011**.

[277] S.-W. Kang, C. M. Gothard, S. Maitra, W. Atia tul, J. S. Nowick, *J. Am. Chem. Soc.* **2007**, *129*, 1486.

[278] H. F. Olivo, N. Perez-Hernandez, D. Liu, M. Iruthayanathan, B. O'Leary, L. L. Homan, J. S. Dillon, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1153.

Appendix

Curriculum Vitae

Work Experience

Jun 2009 – Dec 2011	**Research Assistant, *Prof. Manfred Auer,*** The University of Edinburgh (UK)
Jun 2008 – Jun 2009	**Manufacturing Technician,** Piramal Healthcare (Grangemouth, UK)
Sep 2007 – Jun 2008	**Assistant Scientist,** SEPA (East Kilbride, UK)
Jun 2006 – Sep 2007	**Medicinal Chemistry Industrial Placement**, GlaxoSmithKline Psychiatry Centre of Excellence for Drug Discovery (Harlow, UK)

Education

2004 – 2007	**MChem (2:1, Hons),** University of York

Patent Applications

"*A General Synthetic Route to Isomerically Pure Functionalised Rhodamine Dyes*", Mudd, G.; Auer, M. *Patent application No. 1413685.7*, Aug 2014.

Publications

"*A General Synthetic Route to Isomerically Pure Functionalised Rhodamine Dyes*", Mudd, G.; Pérez Pi, I.; Auer, M. *Manuscript in preparation.*

"*Identification and X-ray Co-crystal Structure of a Small-Molecule Activator of LFA-1-ICAM-1 Binding*", Hintersteiner, M.; Kallen, J.; Schmied, M.; Graf, C.; Jung, T.; Mudd, G.; Shave, S.; Gstach, H.; Auer, M. *Angew. Chem. Int. Ed. Engl*. **2014**, *17*, 4322 - 4326.

"*Towards mimicking short linear peptide motifs: identification of new mixed α,β-peptidomimetic ligands for SLAM-Associated Protein (SAP) by confocal on-bead screening*", Hintersteiner, M.; Knox, A. J.; Mudd, G.; Auer, M. *J. Chem. Biol.* **2012**, *2*, 63 - 79.

Scientific Communication

"Novel Route to Single Isomer Functionalised Rhodamines", Oral presentation & poster presentation, FB^3 - Fluorescent Biomolecules and their Building Blocks, La Jolla (CA, USA - Aug 2014)

"*High Throughput Identification of GPCR Modulators using TOBOC Technology: Simultaneous Binding and Functional Readout via 3 Colour Assay*", Oral presentation, SULSA Student Science Conference, Edinburgh (UK - Jun 2011). **Awarded 1st prize** (translational biology section).