

University of Nottingham
UK | CHINA | MALAYSIA

Site-Selective Bioconjugation via Interception of Free-Radical-Induced Dechalcogenation

Rhys Griffiths

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

Supervisor: Dr Nicholas Mitchell

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 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. I confirm that the work submitted is my own, except work which has formed part of joint-authored publications. The contributions of myself and other authors to this work has been specifically indicated where relevant. I confirm that appropriate credit has been given with the thesis where references has been made to work of others.

The following thesis contains results reported in the following publications:

Site-Selective Modification of Peptides and Proteins via Interception of Free-Radical-Mediated Dechalcogenation; Rhys C. Griffiths, Dr Frances R. Smith, Dr Jed E. Long, Dr Huw E. L. Williams, Prof. Robert Layfield and Dr Nicholas J. Mitchell; *Angew. Chem. Int. Ed.*; **2020**; 59; 23659-23667.

Acknowledgements

Firstly, I would like to thank my supervisor, Dr Nicholas Mitchell, for all the help and guidance he has given me during my PhD. I would also like to thank the other members of the Mitchell group, Andrea, Jo, Heather, Salwa and Lola. I would like to thank Fran for all the help she has given me during her time in the group and for the contributions she has made to my various projects and our published work. I would also like to thank my colleagues from other groups in C17 and B11, particularly Freya and Chris, for the help they have given me during my PhD. Also thanks to the various DTP rotation students (Emma and David), summer students (Nicola and Gigi) and masters students (Carrie, Diana, Yiota, Freddie, Dan, Sam and Sophie).

I would like to thank all the technical staff who have helped me set up the labs each time and with general lab/analytical issues; Zoe and Mark for their help with many lab problems. Ben, Tong and Lola for mass spec and HPLC assistance, and Shaz and Kev for NMR help, particularly when using the 500 MHz instrument. Also thanks to the stores staff, Jim, Eric, John and Chris for sorting a range of problems for me. I would also like to thank my fellow authors on my first publication, Fran for the in lab assistance getting the final bits of data before lockdown, Jed and Huw for NMR analysis of the modified proteins and Rob for providing the modified proteins.

I would like to thank my friends and family for support throughout my studies and life in general. Finally, I owe a massive thanks to my partner, Caitlin, for the constant support she has provided me during my degree and PhD; having to listen to my complaints, being there for me through all the highs and lows, and, very importantly, reminding me to occasionally take breaks and have time away from work.

Abstract

Peptides and proteins play numerous vital roles within cellular biochemistry, as a result they are central to our understanding of the molecular pathways essential to life and the development and progression of disease. The native biological influence of this class of biomolecule has led to the application of polypeptides as effective therapeutics, interest in which has steadily grown over the last four decades. This class of *'biologic'* drug is now at the forefront of modern medicine. In order to exploit the biological properties of peptides and proteins to develop tools to investigate fundamental science and as therapeutic/diagnostic agents, synthetic methods that enable the site-selective installation of biologically relevant groups are required. There are a wide range of established techniques available which selectively target the canonical amino acids. The limitations of these methods include efficiency, operational simplicity, chemoselectivity, versatility and conjugate stability. There have been many recent advances that attempt to address these issues, but novel methods are still in high demand.

A commonly targeted residue is cysteine (Cys), due to its enhanced nucleophilicity and low abundance across the proteome (approximately 2%). An under-utilised residue is the 21st amino acid, selenocysteine (Sec). This residue has a lower abundance than Cys (< 0.1%) and further enhanced nucleophilicity. This thesis describes the development and application of a novel method for the site-selective modification of peptide and proteins by exploiting the process of free-radical-induced dechalcogenation of these two residues. Chapter 1 gives an introduction to site-selective chemistry, focussing mainly on chemistry that targets Cys. It also describes other uses of peptides, including peptide for radiotherapy and the importance of macrocyclic peptides. Chapter 2 describes the generation of a peptide radical (alanyl radical) from Sec and trapping of this intermediate with a persistent radical (TEMPO) bearing biologically relevant functionality. This method is high yielding and operationally simple, allowing the installation of a range of groups. A one-pot procedure was developed, whereby two peptide fragments were ligated via diselenide-selenoester ligation (DSL), followed by direct conjugation of a range of TEMPO probes in high to excellent yields.

Chapter 3 focusses on the application of the developed TEMPO conjugation method to Cys. The method was optimised and applied to a range of Cys containing peptides. This was further extended to the conjugation of a Cys mutant (K48C) of recombinantly expressed ubiquitin (Ub) in high yield. The method was also applied to the dual-functionalisation of a peptide bearing both Sec and Cys residues, thus allowing the site-selective installation of two different biologically useful moieties.

Chapter 4 demonstrates the utilisation of free-radical-induced dechalcogenation to install native and non-proteinogenic functionality via C(sp^3)-C(sp^3) bond formation. Lysine post translational modifications (PTMs) are a widely studied and highly relevant area of science. This chapter describes the development of a novel method to site-selectively install known Lys PTMs. The method is operationally simple and widely applicable to a variety of PTMs. The developed method uses allyl groups to react with the alanyl radical. Using this chemistry, native Lys PTMs were successfully installed into a peptide model in moderate to good yield. The chemistry was then used to modify ubiquitin, installing an acetylated lysine into its native position.

The scope of Chapter 4 was also expanded to an alternate, more stable alkene, for the installation of Lys PTM mimetics and biologically relevant functionality in higher conversions. The alkenes used are structurally similar to the allyl group, but a methyl group on the alkene results in the formation of a stable tertiary radical intermediate. A range of functionality was installed into an 11-mer peptide in good to excellent yield. Using this chemistry, two potential fluorination methods were also developed. A high-yielding, two-step process was developed using known and novel chemistry. A second, one-step method using a novel approach was also developed, albeit with lower isolated yields. This chemistry was also applied to the protein model Ub, installing a trimethylated lysine mimetic and biotin group in excellent yield for a protein model. Additionally, a novel desulfurative arylation reaction is described; while unoptimised, this method has the potential to be developed into a powerful technique for the ate-stage functionalisation of peptides. The data from this chapter shows the scope of this chemistry and the potential for wide-ranging application across numerous fields.

Chapter 5 applies the chemistry developed in Chapter 4 to the macrocyclisation of peptides. The commercially available Fmoc/Boc-allylGlycine-OH was used as the alkene bearing amino acid. The conditions required minor optimisation but were successfully applied to the cyclisation of peptides of various lengths and amino acid sequences. The data gathered in this chapter shows the successful development of a peptide macrocyclisation technique and a novel method of C(sp^3)-C(sp^3) bond formation.

Table of Contents

Declaration ..... **i**

Acknowledgements ..... **ii**

Abstract ..... **iii**

Table of Contents ..... **v**

List of Abbreviations ..... **ix**

List of Tables ..... **i**

List of Figures ..... **iv**

List of Schemes **1**

1. Introduction **7**

1.1 Chemical Modification of Peptides and Proteins **7**

1.2 Chemical Synthesis of Peptides and Proteins **8**

1.2.1 Ligation Techniques 12

1.2.2 Diselenide-Selenoester Ligation 14

1.3 Preparation of Proteins via Recombinant Expression **16**

1.4 Site-Selective Peptide and Protein Modification **17**

1.5 Cysteine-Selective Peptide and Protein Modification **18**

1.5.1 Michael Acceptors 18

1.5.2 Alternative methods for thioether synthesis 20

1.5.3 Alkyl halides – α-halocarbonyls, aminoethylation (alkyl chloride) 21

1.5.4 Transition Metal Mediated Arylation 22

1.5.5 Photoredox Catalysed Methods 24

1.5.6 Oxidative desulfurisation to dehydroalanine 27

1.6 Selenocystine-Selective Conjugation Methods **29**

1.7 Peptide Macrocyclisation **30**

1.8 Protein Radiolabelling **35**

1.9 Aims and Objectives **36**

1) Site-selective bioconjugation via interception of free-radical-mediated dechalcogenation 36

2) Site-selective installation of common Lys post-translational modification (PTMs) 37

3) Macrocyclisation of unprotected peptides 38

2. Selenocystine-selective Bioconjugation via Interception of Free-Radical-Induced Deselenisation with TEMPO traps 39

2.1 Chapter Introduction 39

2.2 Reaction exploration using model peptides 41

2.3 Reaction optimisation using H-UAF-OMe 47

2.4 NMR studies to assess product stereochemistry 71

2.5 Application to one-pot ligation-bioconjugation 72

2.6 Application of one-pot ligation-conjugation to a 58 mer peptide 77

2.7 Chapter Conclusion 79

3. Cysteine-Selective Bioconjugation via Interception of Free-Radical-Induced Desulfurisation with TEMPO traps 80

3.1 Chapter Abstract 80

3.2 Initial Optimisation of Cysteine-Selective Bioconjugation 81

3.3 Reaction of functionalised probes with model peptide 84

3.4 Application to ligation-bioconjugation 86

3.5 Cys/Sec "Dual-functionalisation" 90

3.6 Application of the Cys-selective method to protein model Ub K48C 92

3.7 Conclusion 94

4. Site-selective bioconjugation via desulfurative carbon-carbon bond formation 96

4.1 Chapter Abstract 96

4.2 Initial reactions performed on selenocystine 97

4.3 Reaction optimisation using cysteine-containing peptides 102

4.4 Substrate scope development on larger model systems 118

4.5 Cysteine bioconjugation through C-C bond formation via a stabilised radical intermediate 120

4.6 Cysteine arylation through desulfurative cross-coupling chemistry 131

4.7 Conclusion.....132

5. Application of bioconjugation chemistry to peptide stapling and macrocyclisation 133

5.1 Chapter Abstract 133

5.2 Initial macrocyclisation trials 133

5.3 Further examples for macrocyclisation.....139

5.4 Conclusion.....145

6. Conclusion and Future Work 147

7. Experimental 150

7.1 General Methods 150

Materials 151

Solid Phase Peptide Synthesis (SPPS) 151

Manual Iterative peptide assembly (Fmoc-SPPS)..... 151

Automated solid-phase peptide synthesis 152

Selenoester synthesis general procedure..... 152

General additive-free selenocystine-selenoester ligation procedure 153

General additive-free cysteine-selenoester ligation procedure..... 153

7.2 General Protocols 154

General Protocols for Chapters 2 and 3 154

Initial Optimization of TEMPO Conjugation Method 154

General Protocols for Chapter 4 156

7.3 Compound synthesis.....160

7.4 Peptides synthesised.....172

7.5 TEMPO Bioconjugation Products 175

Ac-YEPLA-SePh (32) and H-UHISKY-NH_2 (33) One-pot Ligation and Conjugation (34-38)..... 187

Ac-YEPLA-SePh (32) and H-UHISCY-NH_2 (39) One-pot Ligation and Conjugation (40 – 42)..... 194

One-pot ligation-conjugation-reductive cleavage – ligation at serine (43) 197
One-pot ligation-oxidation deselenization (Sec to Ser) using Oxone-TCEP method (43b) 198
ZEGFR:1907 - Ligation and Conjugation (46)..... 199
Ac-CWHISKEY-NH_2 Conjugation Reactions (50-54) 202
Reactions Showing Sec/Cys Selectivity (58 and 59)..... 210
K48C Ubiquitin Bioconjugation with Propargyl-TEMPO (61)..... 212
7.6 Allyl Conjugation Products 216
Ac-CWHISKEY-NH_2 (49) Conjugated with N-allyl acetamide (72) 217
Small model Lys PTM installation example for NMR analysis (77 and 78)..... 218
Substrate scope – Application to small histone model (80, 86-91)..... 220
Native Lys PTM installation into proteins..... 227
7.7 Stabilised Alkene Conjugation Products.....228
Ac-CWHISKEY-NH_2 Modification 228
Ac-YEPLACHISKY-NH_2 Modification..... 229
Small model examples for NMR analysis 230
Ac-CAY-NH_2 with 2-methylprop-2-ene-2-sulfonic acid followed by hydrofluorination . 240
Novel fluorination method (116 and 95)..... 243
Protein bioconjugation using stabilised alkene probes (117 and 118)..... 245
7.8 Desulfurative Cross-coupling.....247
7.9 Peptide Macrocyclisation 249
References 270

List of Abbreviations

A/Ala	Alanine
Ac_2O	Acetic anhydride
ADC	Antibody drug conjugate
alG	Allyl glycine
Boc	Tert-Butoxycarbonyl
BSA	Bovine serum albumin
C/Cys	Cysteine
CTC	2-Chlorotrityl chloride
D/Asp	Aspartic acid
d	Doublet
dd	Doublet of doublets
ddd	Doublet of doublet of doublets
DCM	Dichloromethane
Dha	Dehydroalanine
DIC	Diisopropylcarbodiimide
DMF	N,N-dimethylformamide
DMSO	Dimethyl sulfoxide
DSL	Diselenide-selenoester ligation
DTT	1,4-Dithiothreitol
E/Glu	Glutamic acid
EBX	Ethynylbenziodoxolones
Eq.	Equivalents
ESI-MS	Electrospray ionisation mass spectrometry
EY	Eosin Y
F/Phe	Phenylalanine

FIT	Flexible in vitro translation
Fmoc	Fluorenylmethyloxycarbonyl
G/Gly	Glycine
GPx	Glutathione perozidase
H/His	Histidine
HBTU	N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate
HF	Hydrofluoric acid
HMPT	Hexamethylphosphorous triamide
HOBt	1-Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
I/Ile	Isoleucine
Ir(III)	$(Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$
K/Lys	Lysine
L/Leu	Leucine
M/Met	Methionine
m	Multiplet
MeCN	Acetonitrile
N/Asn	Asparagine
NCL	Native chemical ligation
NHS	N-hydroxysuccinimide
Oxone	Potassium peroxymonosulfate
Oxyma	Ethyl cyano(hydroxyimino)acetate
P/Pro	Proline
PB	Phosphate buffer

PPI	Protein-protein interaction
PEG	Polyethylene glycol
PET	Positron emission tomography
Q/Gln	Glutamine
q	Quartet
R/Arg	Arginine
RaPID	Random nonstandard peptide integrated discovery
RCM	Ring-closing metathesis
S/Ser	Serine
s	Singlet
SICLOPPS	Split-intein circular ligation of peptides and proteins
SET	Single-electron transfer
SPECT	Single-photon emission computed tomography
SPPS	Solid-phase peptide synthesis
T/Thr	Threonine
t	Triplet
tBu	Tert-butyl
tBuSH	*tert*-Butyl mercaptan
TCEP	Tris(2-carboxyethyl)phosphine
TIS	Triisopropylsilane
TMAA-Cl	Trimethylallylammonium chloride
Trt	Trityl

TFA	Trifluoroacetic acid
U/Sec	Selenocystine
V/Val	Valine
VA-044	2,2'-Azobis[2-(2-imidazolin-2-yl)propane] Dihydrochloride
W/Trp	Tryptophan
Y/Tyr	Tyrosine

List of Tables

Table 1 - TCEP and TEMPO equivalents optimisation for Sec TEMPO trapping. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**3**). 42

Table 2 - Results from TEMPO trapping optimisation at 37 ^{o}C. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**3**). 43

Table 3 - Data from the stability tests of the H-A(TEMPO)AF-OMe (**11**) conjugate. 63

Table 4 - Data from the stability studies of the TEMPO conjugate H-A(TEMPO)HISKY-NH_2 (**19**)..... 65

Table 5 - Table summarising the yields from the bioconjugation (Scheme 48) of H-UAF-OMe (**10**) with a range of TEMPO probes. 70

Table 6 - Isolated yields for the bioconjugation of Ac-CWHISKEY-NH_2 with a range of TEMPO based probes. 85

Table 7 - Results from the optimisation of Ac-YEPLACHISKY-NH_2 conjugation with propargyl TEMPO (Scheme 57) forming the conjugated product (**35b**). Percentage conversion shows the conversion of starting material to the desired product (**35b**). The time-point conversions are shown in the two graphs below. 88

Table 8 - Initial allyl alcohol trapping experiments (Scheme 62). 98

Table 9 - Allyl alcohol trapping experiments forming allyl-trapped product (**55**). 99

Table 10 - Allyl alcohol trapping experiments forming allyl-trapped product (**54**). 99

Table 11 - Example Sec allyl trapping reactions using eosin Y and photoredox initiator forming conjugated product (**67**)..... 101

Table 12 - Initiator screening for cysteine desulfurisation and DMAA trapping to form trapped product (**70**). Other includes a range of multiple addition (**71**) products, as well as unidentified by-products. 104

Table 13 - Results from pH optimisation trials to form trapped product (**70**), the values in brackets are the corrected values, assuming all values were 2 pH units lower than expected. 105

Table 14 - Results from altering the equivalents of TCEP and the effect this had on the allyl trapping reaction to form trapped product (**70**)..... 107

Table 15 - Results from altering the DMAA (**69**) concentration on the percentage conversion of starting material (**68**) to product (**70**)..... 108

Table 16 - Results from performing the allyl trapping reaction at various peptide concentrations to form trapped product (**60**)..... 109

Table 17 - Initial allyl conjugation reactions using the model peptide Ac-CWHISKEY-NH_2 (**41**) forming the conjugated product (**66**). Peptide – 1 mM, TCEP – 50 mM, 20% DMSO in LB at pH 7 for 2 hours..... 110

Table 18 – Peptide concentration optimisation using 50 eq. TCEP and 500 eq. N-allyl acetamide in 20% DMSO in LB at pH 7 for 2 hours, forming the conjugated product (**72**). 111

Table 19 - TCEP optimisation using 0.5 mM peptide, 500 eq. N-allyl acetamide and either 0.1 eq. eosin Y or 0.01 eq. Ir(III) in 20% DMSO in LB at pH 7 for 2 hours, forming the conjugated product (**72**)..... 112

Table 20 - pH optimisation using 0.5 mM peptide (**49**), 5 eq. TCEP, 500 eq. N-allyl acetamide (**66**) and either 0.1 eq. eosin Y or 0.01 eq. Ir(III) in 20% DMSO in LB at pH 7 for 4 hours. Entry 9 reached completion within 2 hours, forming the conjugated product (**72**). 113

Table 21 - Initiator equivalents optimisation. Peptide 0.5 mM, 5 eq. TCEP, 500 eq. N-allyl acetamide. Eosin Y reactions at pH 7 and Ir(III) reactions at pH 8, forming the conjugated product (**66**). Solvent used was 20% DMSO in LB. Reactions analysed after 1 and 4 hours..... 114

Table 22 - N-allyl acetamide eq. optimisation. 0.5 mM peptide, 5 eq. TCEP. Either 0.5 eq. eosin Y at pH 7 or 0.05 eq. Ir(III) at pH 8, forming the conjugated product (**66**). Reactions analysed after 1 hour..... 115

Table 23 - Low concentration studies using Ac-YEPLACHISKY-NH_2 (**56**) with 100 eq. TCEP, 1000 eq. 3-methyl-3-buten-1-ol (**93**) and 10 mol% Ir(III) in 10% DMSO in LB at pH 8, forming product (**95**). Values given in brackets is the percentage of starting material remaining. 124

Table 24 - Optimisation of the macrocyclisation reaction using the H-S(allyl)AFAC-NH_2 (**125**) model peptide. The solvent system used was ligation buffer at pH 8. 136

Table 25 – Summary of the isolated yields obtained from the desulfurative macrocyclisation of the peptides specified..... 145

Table 26 - Summary of the conditions used in the general protocols above for chapters 2 and 3..... 156

Table 27 - Summary of the conditions used in the general protocols for chapter 4. Ir(III) = ($Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$. All the below reactions use blue LEDs as the initiator, with LB/MeCN at pH 7.8-8.2. 159

List of Figures

Figure 1 - Example protected amino acid (Asp) showing the N-terminal protecting group (Boc, blue) and the side-chain protecting group (Benzyl, red). 9

Figure 2 - Structure of Fmoc-Glu(OtBu)-OH. 10

Figure 3 - Chemical structures of HBTU (left), DIC (middle) and Oxyma (right) 11

Figure 4 - Structure of the cyclic peptide Murepavadin. 32

Figure 5 – Stabilisation of TEMPO radical through formation of a two-center 3-electron bond. 39

Figure 6 - Expected products from TEMPO conjugation reaction, including desired TEMPO conjugated product (**3**), deselenisation to Ala (**4**), oxidative deselenisation to Ser (**5**) and diselenide reduction to selenol (**6**). 42

Figure 7 - HPLC trace from TEMPO trapping experiment using 2.5 mM peptide in LB at 37 ^{o}C for 5 hours with 40 eq. TCEP and 4 eq. TEMPO, showing a significant by-product (**9**) at 4.5 minutes. HPLC gradient 10-70% B over 5 mins, 210 nm. 45

Figure 8 - HPLC trace showing the pure 4.5 minute by-product (**9**). HPLC gradient 10-70% B over 5 mins, 210 nm. 45

Figure 9 - MS analysis of the 4.5 minute by-product product (**9**), showing a single molecular weight ion of 374.2. 46

Figure 10 - Proposed structure of the 4.5 minute by-product (**9**), showing the 6-membered diketopiperazine formed by the N-terminal amine displacing the C-terminal methyl ester. 46

Figure 11 – Reaction scheme and time point conversion progression plot from the bioconjugation of H-UAF-OMe (**7**) with TEMPO at a range of peptide concentrations. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**8**). 48

Figure 12 – H-UAF-OMe (**10**) TEMPO trapping optimization – TCEP concentration. The reaction was doped three times with additional TCEP and TEMPO. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**11**). 49

Figure 13 – Reaction optimisation by altering the number of times that the reaction mixture is doped with addition quantities of TCEP and TEMPO. Percentage conversion shows the

conversion of starting material to products, unless specified the sole product formed was (**11**)..... 50

Figure 14 – Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the equivalents of TEMPO (**2**) in the reaction mixture. All entries were doped twice with additional TCEP and TEMPO (**2**). For all entries the sole product formed was (**11**). 51

Figure 15 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the pH of the reaction mixture. All entries were doped twice with additional TCEP and TEMPO..... 52

Figure 16 – Product (**12**) formed from the radical deselenisation of H-UAF-OMe (**10**) when TEMPO conjugation does not occur. 52

Figure 17 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the percentage of organic solvent in the reaction mixture. All entries were doped twice with additional TCEP and TEMPO. The percentage conversion shows how much desired product (**11**) was formed compared to any by-products or starting material. 53

Figure 18 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the organic co-solvent in the reaction mixture. All entries were doped twice with additional TCEP and TEMPO. For all entries the sole product formed was (**11**)..... 54

Figure 19 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the temperature of the reaction mixture. All entries were doped twice with additional TCEP and TEMPO. For all entries the sole product formed was (**11**). 55

Figure 20 – Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the initiator in the reaction mixture. All entries were doped twice with additional TCEP and TEMPO. For all entries the sole product formed was (**11**). 57

Figure 21 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the equivalents of $Mn(OAc)_3$ in the reaction mixture. All reactions doped twice unless specified. For all entries the sole product formed was (**11**)..... 58

Figure 22 – Reaction scheme and the associated crude HPLC trace compared with the HPLC trace of the pure starting H-UAF-OMe (**10**, red). Analytical HPLC gradient 10-70% B over 5 mins, 210 nm. 59

Figure 23 - Analytical HPLC trace of Ac-UHISKY-OH (**13**) deselenisation to Ala product (**14**). Analytical gradient 10-25% B over 5 min, 210 nm. Blue – Reaction mixture, Black – Starting material (**13**), Green dash – Ac-AHISKY-OH standard (**14**). 60

Figure 24 - Resultant HPLC traces from applying the Mn(III) conditions on the Cys containing peptide Ac-CHISKY-OH (**15**). Three traces are overlaid, reaction mixture (blue), starting material (**15**, black) and Ac-AHISKY-OH standard (**13**, green dash). Analytical gradient 10-25% B over 5 mins, 210 nm. 61

Figure 25 – Chemical structure and analytical HPLC trace of pure TEMPO-WIMKY-NH_2 (**17**). Analytical gradient 10-75% B over 10 min, 210 nm. HRMS Calculated Mas: $[M+H]^+$ = 965.5641, $[M+2H]^{2+}$: 483.2860. Actual Mass: $[M+H]^+$ = 965.5658, $[M+2H]^{2+}$: 483.2879. 62

Figure 26 –HPLC trace comparison, showing the starting material H-A(TEMPO)AF-OMe (**11**, blue) and the reductively cleaved product from Entry 6 (**18**, red). Analytical gradient 10-70% B over 5 mins, 210 nm. 64

Figure 27 – H-A(TEMPO)HISKY-NH_2 (**19**) model used to prove conjugate stability and the ability to selectively cleave the conjugate. 65

Figure 28 – Product (**20**) from reductive cleavage of starting TEMPO conjugate (**19**) and overlaid HPLC traces of the starting conjugate (**19**, green) and the resultant crude trace from Entry 5 (**20**, red). Analytical HPLC gradient 10-70% B over 5 mins, 210 nm. 66

Figure 29 - A range of TEMPO probes synthesised showing varying functionality and biological importance. 67

Figure 30 - A comparison showing the H-UAF-OMe (**10**) starting material (red dash) and the crude propargyl TEMPO conjugated product (**28**) after 1 hour (blue)..... 70

Figure 31 - Structures of the products from H-UAF-OMe conjugation with TEMPO using L-Sec (**11**, left) and D-Sec (**11b**, right). 71

Figure 32 - NMR comparison of D-Sec (**11b**, top) and L-Sec (**11**, bottom) TEMPO conjugated products showing key differences in the NMRs of each product. The absence of highlighted peaks in the L product compared to the D shows racemization does not occur during conjugation. 71

Figure 33 - Summary of the one-pot ligation and conjugation reactions using Ac-YEPLA-SePh (**32**) and H-UHISKY-NH_2 (**34**) fragments. The yields reported are isolated yields based on the amount of starting H-UHISKY-NH_2 (**33**). 74

Figure 34 - Example crude HPLC trace from the one-pot ligation and conjugation using propargyl TEMPO (**23**, blue) and the pure product overlaid (**35**, red dash). Analytical HPLC gradient 10-100% B over 5 mins, 210 nm. 74

Figure 35 - Isolated yields for the one-pot ligation and conjugation of Ac-YEPLA-SePh (**32**) and H-UHISCY-NH_2 (**39**) with selected TEMPO probes. 75

Figure 36 - A comparison between the crude HPLC trace (red) of one-pot ligation and conjugation of Ac-YEPLA-SePh (**32**) and H-UHISCY-NH_2 (**39**) with propargyl TEMPO (**23**) and the purified product (**41**, blue dash). Analytical HPLC gradient 10-100% B over 5 mins, 210 nm. Peak at 5.1 minutes is propargyl TEMPO. 76

Figure 37 - Comparison of the crude trace from the one-pot ligation and conjugation of the ZEGFR protein with propargyl TEMPO (**46**, red) and the pure conjugate (**46**, blue dash). Analytical gradient 10-100% B over 5 mins, 210 nm. 78

Figure 38 - Crude HPLC trace from the reaction of H-CAF-OMe (**48**) with TEMPO using general protocol A (red). The pure TEMPO conjugated product (**11c**) is overlaid (blue dash). Analytical gradient 10-70% B over 5 mins, 210 nm. 82

Figure 39 – Summary of the optimisation reactions performed to increase the conversion observed of **48** to **11c**. All reactions were performed for 2 hours, with time-point monitoring at one and two hours. 82

Figure 40 - Comparison of ^{1}H-NMR of the TEMPO conjugated products using H-UAF-OMe (**11**, top) and H-CAF-OMe (**11c**, bottom). 83

Figure 41 - Overlaid pure HPLC traces for the TEMPO conjugated products using H-UAF-OMe (**11**, red) and H-CAF-OMe (**11c**, blue). Analytical gradient 10-70% B over 5 mins, 210 nm. 84

Figure 42 - Analytical HPLC traces for the bioconjugation of Ac-CWHISKEY-NH_2 (**49**) with propargyl TEMPO (**23**). Crude (red) and pure (**51**, blue dash). 85

Figure 43 - Crude analytical HPLC trace of the ligation of Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**). Analytical gradient 10-100% B over 5 mins, 280 nm. 86

Figure 44 - Crude analytical HPLC trace after addition of NH_2OH and TCEP to the ligation of Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**). Analytical gradient 10-100% B over 5 mins, 280 nm. 87

Figure 45 – Plots showing the percentage conversion against time from the results in Table 7. 89

Figure 46 - Isolated yields from the two-step ligation and conjugation of Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**)..... 89

Figure 47 - Crude HPLC trace from the reductive cleavage of the Ac-YEPLACHISKY-NH_2 propargyl TEMPO product (**57**, blue) compared to the starting conjugate (**35b**, red). Analytical gradient 10-100% B over 5 mins, 280 nm..... 90

Figure 48 - HPLC traces of starting H-UHISCY-NH_2 (**39**, grey), propargyl TEMPO modified product (**58**, orange) and tetraEG TEMPO modified product (**59**, blue). 91

Figure 49 - Structure of the final product (**59**) from the dual-modification of H-UHISCY-NH_2. 91

Figure 50 - Δ chemical shift (ppm) plot comparing NH signals for K48C Ub (**60**) and Ub conjugate (**61**)**.** 93

Figure 51 - Space filling and ribbon models of Ub; modified residue at position 48 highlighted in green, residues that experience significant deviation in chemical shifts for NH signal highlighted in red. 94

Figure 52 - Crude HPLC trace from Entry 2 (Table 11). The peak at 3.8 minutes is Ala (**14**); the peak at 4.2 minutes is desired product (**67**). Analytical HPLC gradient 10-25% B over 5 mins, 280 nm..... 101

Figure 53 - Crude HPLC trace of the reaction shown in Scheme 64 (red) and the pure product (**67b**, blue dash). The peak at 3.8 mins is Ala (**14**) and the peak at 4.6 mins is starting material (**15**). Analytical HPLC gradient 10-25% B over 5 min, 280 nm. 102

Figure 54 - HPLC trace from Entry 7 (Table 13) showing the multiple addition peak at 3.6 mins. The desired product (**70**) is at 3.8 mins and the Ala by-product (**4**) at 4.1 mins. 106

Figure 55 - Rate analysis for the bioconjugation of Ac-CWHISKEY-NH_2 (**49**) with N-allyl acetamide (**66**) to form product (**72**), with time points every 15 minutes. 116

Figure 56 – Chemical structures of the isolated products **77**, L and **78**, D. 117

Figure 57 – NMR comparison of conjugated L (**77**, bottom) and D (**78**, top) Ac-CAY-NH_2 model system with N-allyl acetamide (**66**). 117

Figure 58 - Time point graph for the initial trials on the H-HRLLRCGNYA-NH_2 model, varying the equivalents of TCEP and Ir(III), showing the conversion of (**79**) to (**80**). 118

Figure 59 - Crude HPLC trace for the modification of Ac-CWHISKEY-NH_2 (**49**) with 3-methyl-3-buten-1-ol (**93**). Pure SM (**49**) in orange and crude reaction mixture trace forming product (**94**) in blue. Peak at 5.5 minutes is Ala by-product (**73**), and the peak at 8.2 is desired product (**94**). Analytical HPLC gradient 15-25 over 10 mins, 280 nm. 122

Figure 60 - Low concentration trials on Ac-CWHISKEY-NH_2 (**49**) to assess protein applicability..... 123

Figure 61 - NMR comparison between modified Ac-L-CAY-NH_2 (**96**, top) and Ac-D-CAY-NH_2 (**97**, bottom) conjugated with 3-methyl-3-butene-1-ol (**93**). 125

Figure 62 - NMR overlay of the starting alkene (**114**, top) and the hydrofluorinated product (**115**, bottom). The left expansion shows the disappearance of the alkene protons at 4.75. The right expansion shows the appearance of two methyl groups, showing fluorination occurred on the internal carbon of the alkene. 128

Figure 63 – Proposed structure from the reaction of Ac-CAY-NH_2 with 1-fluoro-2,4,6-trimethylpyridinium tetrafluoroborate..... 129

Figure 64 – NMR comparison of starting peptide (**122**, bottom) and cyclised peptide (**123**, top, majority diastereomer) showing the disappearance of the alkene protons at 4.9 ppm. 134

Figure 65 - Allyl modified Ser (**124**) used in the initial macrocyclisation trials. 134

Figure 66 - The two potential products matching the observed molecular weight from the MS analysis. Both have a calculated molecular weight of 505.2775. The two structures show the cyclised (**126**, left) and linear desulfurised (**127**, right) products..... 135

Figure 67 - Crude analytical HPLC trace from the cyclisation optimisation Entry 5 (Table 24). 137

Figure 68 – NMR overlay showing starting peptide H-S(Allyl)AFAC-NH_2 (**125**, bottom) and cyclised product (**126**, top). 137

Figure 69 - NMR comparison of desulfurised and cyclised H-alGAFAC-NH_2 (**129**, top) and the starting peptide (**130**, bottom)..... 139

Figure 70 – Structure of the cyclised 23-mer peptide H-alGMPWLRC-NH_2 (**133**). 140

Figure 71 – The isolated product (**134**) from the trial macrocyclisation of H-alGYEPLAWHISKEYC-NH_2, showing nucleophilic addition of TCEP into allyl glycine. 140

Figure 72 – Desired product (**135**) from the attempted desulfurative cyclisation of H-alGYEPLAWHISKEYC-NH_2 using 1 mol% catalyst producing a 44-mer cyclic peptide. 141

Figure 73 - Diagram showing the difference between an *i, i*+4 (left) and an *i, i*+7 (right) staple. 141

Figure 74 – Structures of the cyclic peptides from the desulfurative cyclisation of H-alGEDLFYQC-NH_2 (left, **138**) and H-alGDKFNHEC-NH_2 (right, **139**), producing 26-mer cyclic peptides. 142

Figure 75 – Structure of the product from the desulfurative cyclisation of H-DalGFYQC-NH_2 (**141**). 142

Figure 76 – Structure of the native oxytocin, bridged via a disulfide bond between the two Cys residues – H-CYIQNCPLG-NH_2 143

Figure 77 – Structure of the cyclic peptide (**145**) obtained from the desulfurative cyclisation of H-CYIQNalGPLG-NH_2, producing a 20-mer cyclic peptide. 144

Figure 78 - NMR overlay comparing oxytocin model starting material (**144**, top), linear desulfurised (**146**, middle) and cyclised material (**145**, bottom). The dashed red lines highlight the allylic protons, not seen in the bottom trace. The dashed green line shows the presence of a doublet and 1.32 ppm, indicative of the Ala CH_3. 144

Figure 79 - Analytical HPLC trace of pure TEMPO modified H-UF-OMe (**3**). Analytical gradient 10-70% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 406.2706. Observed Mass $[M+H]^+$: 406.2732. 175

Figure 80 - Analytical HPLC trace of pure FITC-TEMPO modified H-UF-OMe (**8**). Analytical gradient 30-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 810.3173, $[M+2H]^{2+}$: 405.6587. Observed Mass $[M+H]^+$: 406.2732, $[M+2H]^{2+}$: 405.6646. 176

Figure 81 - Analytical HPLC trace of pure TEMPO modified H-UAF-OMe (**11**). Analytical gradient 10-70% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 477.3104. Observed Mass $[M+H]^+$: 477.3077. 177

Figure 82 - Analytical HPLC trace of pure conjugated H-CAF-OMe using TEMPO (**11b**). Analytical gradient 10-70% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 477.3077. Observed Mass $[M+H]^+$: 477.3095. 178

Figure 83 - Analytical HPLC trace of pure TEMPO-WIMKY-NH_2 (**17**). Analytical gradient 10-75% B over 10 min, 210 nm. HRMS Calculated Mas: $[M+H]^+$ = 965.5641, $[M+2H]^{2+}$: 483.2860. Observed Mass: $[M+H]^+$ = 965.5658, $[M+2H]^{2+}$: 483.2879. 179

Figure 84 - Analytical HPLC trace purified reductively cleaved product (**18**). Analytical gradient 10-70% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 338.1716. Observed Mass $[M+H]^+$: 338.1731. 180

Figure 85 - Analytical HPLC trace of pure TEMPO modified H-UHISKY-NH_2 (**17**). Analytical gradient 1070% B over 5 min, 280 nm. HRMS Calculated Mass $[M+H]^+$: 872.5358, $[M+2H]^{2+}$: 436.7718. Observed Mass $[M+H]^+$: 872.5387, $[M+2H]^{2+}$: 436.7748. 181

Figure 86 - Analytical HPLC trace of pure tetra-EG-TEMPO modified H-UAF-OMe (**27**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 682.4391. Observed Mass $[M+H]^+$: 682.4426..... 182

Figure 87 - Analytical HPLC trace of pure propargyl-TEMPO modified H-UAF-OMe (**28**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 531.3183. Observed Mass $[M+H]^+$: 531.3218..... 183

Figure 88 - Analytical HPLC trace of pure FITC-TEMPO modified H-UAF-OMe (**29**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 881.3544, $[M+2H]^{2+}$: 441.1811. Observed Mass $[M+H]^+$: 881.3552, $[M+2H]^{2+}$: 441.1834. 184

Figure 89 - Analytical HPLC trace of pure biotin-EDA-TEMPO modified H-UAF-OMe (**30**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 761.4384, $[M+2H]^{2+}$: 381.2231. Observed Mass $[M+H]^+$: 761.4414, $[M+2H]^{2+}$: 381.2250. 185

Figure 90 - Analytical HPLC trace of pure gemcitabine-TEMPO modified H-UAF-OMe (**31**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 766.3587. Observed Mass $[M+H]^+$: 766.3614..... 186

Figure 91 - Analytical HPLC trace of pure tetra-EG-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**34**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 1692.9654, $[M+2H]^{2+}$: 846.9827. Observed Mass $[M+H]^+$: 1693.9599, $[M+2H]^{2+}$: 846.9861. 187

Figure 92 - Analytical HPLC trace of pure propargyl-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**35**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 771.4223. Observed Mass $[M+2H]^{2+}$: 771.4237..... 188

Figure 93 - Analytical HPLC trace of pure conjugated Ac-YEPLACHISKY-NH_2 using propargyl-TEMPO (**35b**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^+$: 1541.8367, $[M+2H]^{2+}$: 771.4223. Observed Mass $[M+H]^+$: 1541.8331, $[M+2H]^{2+}$: 771.4225. 189

Figure 94 - Analytical HPLC trace of pure biotin-EDA-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**36**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 886.4824. Observed Mass $[M+2H]^{2+}$: 886.9884.190

Figure 95 - Analytical HPLC trace of pure FITC-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**37**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+2H]^{2+}$: 946.4404, $[M+3H]^{3+}$: 631.2962. Observed Mass $[M+2H]^{2+}$: 946.9429, $[M+3H]^{3+}$: 631.6318. 191

Figure 96 - Analytical HPLC trace of pure gemcitabine-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**38**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 888.9425. Observed Mass $[M+2H]^{2+}$: 888.9440.192

Figure 97 - Analytical HPLC trace of pure conjugated of Ac-YEPLACHISKY-NH_2 using gemcitabine-TEMPO (**38b**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+2H]^{2+}$: 888.9426. Observed Mass $[M+2H]^{2+}$: 888.9442..... 193

Figure 98 - Analytical HPLC trace of pure TEMPO modified Ac-YEPLAUHISCY-NH_2 from one-pot ligation-conjugation (**40**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^{+}$: 1462.7405, $[M+2H]^{2+}$: 731.8740. Observed Mass $[M+H]^{+}$: 1462.7338, $[M+2H]^{2+}$: 731.8742. 194

Figure 99 - Analytical HPLC trace of pure propargyl-TEMPO modified Ac-YEPLAUHISCY-NH_2 from one-pot ligation-conjugation (**41**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^{+}$: 1517.7588, $[M+2H]^{2+}$: 759.3833. Observed Mass $[M+H]^{+}$: 1517.7446, $[M+2H]^{2+}$: 758.8793. 195

Figure 100 - Analytical HPLC trace of pure biotin-EDA-TEMPO modified Ac-YEPLAUHISCY-NH_2 from one-pot ligation-conjugation (**42**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^{+}$: 1746.8711, $[M+2H]^{2+}$: 873.9395. Observed Mass $[M+H]^{+}$: 1746.8572, $[M+2H]^{2+}$: 873.9427..... 196

Figure 101 - Analytical HPLC trace of pure reductively cleaved product (**43**) after semi-preparative HPLC purification. Analytical gradient 10-100% B over 5 minutes, 280 nm. Calculated Mass $[M+H]^{+}$: 1441.8189, $[M+2H]^{2+}$: 721.3634. Observed Mass $[M+H]^{+}$: 1441.7150, $[M+2H]^{2+}$: 721.3622. 197

Figure 102 - Analytical HPLC trace of pure ligated/oxidatively cleaved product (**43b**). Analytical gradient 10-100% B over 5 minutes, 280 nm. Calculated Mass $[M+H]^{+}$: 1441.8189, $[M+2H]^{2+}$: 721.3634. Observed Mass $[M+H]^{+}$: 1441.7150, $[M+2H]^{2+}$: 721.3622. 198

Figure 103 - Analytical HPLC trace of pure ligated ZEGFR:1907 peptide (**47**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+4H]^{4+}$: 1656.0528,

[M+5H]$^{5+}$: 1325.0438, [M+6H]$^{6+}$: 1104.3711, [M+7H]$^{7+}$: 946.7478, [M+8H]$^{8+}$: 828.5303. Observed Mass [M+4H]$^{4+}$: 1656.8046, [M+5H]$^{5+}$: 1325.6471, [M+6H]$^{6+}$: 1104.7167, [M+7H]$^{7+}$: 947.1764, [M+8H]$^{8+}$: 828.9182. 199

Figure 104 - Analytical HPLC trace of pure conjugated ZEGFR:1907 peptide (**46b**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass [M+5H]$^{5+}$: 1350.6898, [M+6H]$^{6+}$: 1125.7428, [M+7H]$^{7+}$: 965.0664. Observed Mass [M+5H]$^{5+}$: 1351.6782, [M+6H]$^{6+}$: 1126.4290, [M+7H]$^{7+}$: 965.6493. 200

Figure 105 - Analytical HPLC trace of pure one-pot ligated and conjugated ZEGFR:1907 peptide using propargyl-TEMPO (**46**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass [M+4H]^{4}+: 1689.3458, [M+5H]^{5}+: 1351.6782, [M+6H]^{6}+: 1126.5665, [M+7H]^{7}+: 965.7724, [M+8H]^{8}+: 845.1768. Observed Mass [M+4H]^{4}+: 1689.3525, [M+5H]^{5}+: 1351.6898, [M+6H]^{6}+: 1126.5817, [M+7H]^{7}+: 965.7811, [M+8H]^{8}+: 845.1737. 201

Figure 106 - Analytical HPLC trace of pure conjugated Ac-CWHISKEY-NH_2 using tetra-EG-TEMPO (**50**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass [M+H]$^{+}$: 1434.7997, [M+2H]$^{2+}$: 717.9038. Observed Mass [M+H]$^{+}$: 1434.7935, [M+2H]$^{2+}$: 717.9047. 202

Figure 107 - Analytical HPLC trace of pure conjugated Ac-CWHISKEY-NH_2 using propargyl-TEMPO (**51**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass [M+H]$^{+}$: 1283.6788, [M+2H]$^{2+}$: 642.3433. Observed Mass [M+H]$^{+}$: 1283.6829, [M+2H]$^{2+}$: 642.3466. 203

Figure 108 - Analytical HPLC trace of pure conjugated of Ac-CWHISKEY-NH_2 using FITC-TEMPO (**52**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass [M+H]$^{+}$: 1634.7228, [M+2H]$^{2+}$: 817.3614. Observed Mass [M+H]$^{+}$: 1634.7148, [M+2H]$^{2+}$: 817.3640. 204

Figure 109 - Analytical HPLC trace of pure conjugated of Ac-CWHISKEY-NH_2 using Biotin-EDA-TEMPO (**53**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass [M+H]$^{+}$: 1513.7990, [M+2H]$^{2+}$: 757.3995. Observed Mass [M+H]$^{+}$: 1514.7992, [M+2H]$^{2+}$: 757.4059. 205

Figure 110 - Analytical HPLC trace of pure conjugated Ac-CWHISKEY-NH_2 using gemcitabine-TEMPO (**54**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass [M+H]$^{+}$: 1519.7272, [M+2H]$^{2+}$: 759.8636. Observed Mass [M+H]$^{+}$: 1519.7179, [M+2H]$^{2+}$: 759.8660. 206

Figure 111 - Analytical HPLC trace of purified ligated product (**47**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^{+}$: 1364.6672, $[M+2H]^{2+}$: 682.8375. Observed Mass $[M+H]^{+}$: 1364.6685, $[M+2H]^{2+}$: 682.8408. 207

Figure 112 - Analytical HPLC trace of pure reductively cleaved product (**57**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^{+}$: 1348.6901, $[M+2H]^{2+}$: 674.8490. Observed Mass $[M+H]^{+}$: 1348.6926, $[M+2H]^{2+}$: 674.8508. 208

Figure 113 - Analytical HPLC trace of 16 hours crude reaction mixture with no Zn added. Analytical gradient 10-100% B over 5 min. 209

Figure 114 - Analytical HPLC trace of pure conjugated H-UHISCY-NH_2 using propargyl-TEMPO (**58**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^{+}$: 901.4606, $[M+2H]^{2+}$: 451.2342. Observed Mass $[M+H]^{+}$: 901.4622, $[M+2H]^{2+}$: 451.2359. 210

Figure 115 - Analytical HPLC trace of pure conjugated (**59**) using tetra-EG-TEMPO. Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 615.3794. Observed Mass $[M+2H]^{2+}$: 615.3819. 211

Figure 116 - Analytical HPLC trace of pure Ub K48C modified with propargyl-TEMPO (**61**). Analytical gradient 25-60% B over 5 min, 210 nm. Calculated Mass $[M+5H]^{5+}$: 1743.1397, $[M+6H]^{6+}$: 1452.7844, $[M+7H]^{7+}$: 1245.3877, $[M+8H]^{8+}$: 1089.8402, $[M+9H]^{9+}$: 968.8588, $[M+10H]^{10+}$: 872.0737, $[M+11H]^{11+}$: 792.8859. Observed Mass $[M+5H]^{5+}$: 1742.5410, $[M+6H]^{6+}$: 1452.4417, $[M+7H]^{7+}$: 1245.3877, $[M+8H]^{8+}$: 1089.4601, $[M+9H]^{9+}$: 968.6551, $[M+10H]^{10+}$: 871.9165, $[M+11H]^{11+}$: 792.6438. 212

Figure 117 - Overlay of TOSCY spectra of K48C Ub mutant (**601**) and Ub conjugate (**61**).213

Figure 118 - Expansion of NH region of the overlaid TOCSY spectra for K48C Ub mutant (**60**) and Ub conjugate (**61**); signal shifts around modified residue 48 highlighted. 214

Figure 119 - Overlay for 1H NMR spectra of K48C Ub mutant (**60**) and Ub conjugate (**61**); NH region extended across 6.5-9.5 ppm indicating a tertiary fold. 215

Figure 120 - Analytical HPLC trace of pure N-allyl acetamide modified Ac-CHISKY-OH (**67**). Analytical gradient 10-25% B over 5 min, 210 nm. Calculated Mass $[M+H]^{+}$: 859.4678, $[M+2H]^{2+}$: 430.2378. Observed Mass $[M+H]^{+}$: 859.4654, $[M+2H]^{2+}$: 430.2370. 216

Figure 121 - Analytical HPLC trace of pure Ac-CWHISKEY-NH_2 conjugated with N-allyl acetamide (**72**). Analytical gradient 15-25% B over 10 min, 280 nm. Calculated Mass $[M+H]^{+}$: 1173.6057, $[M+2H]^{2+}$: 587.3068. Observed Mass $[M+H]^{+}$: 1173.6064, $[M+2H]^{2+}$: 587.3046. 217

Figure 122 - Analytical HPLC trace of pure Ac-CAY-NH_2 conjugated with N-allyl acetamide (**77**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 464.2509. Observed Mass $[M+H]^+$: 464.2512. 218

Figure 123 - Analytical HPLC trace of pure Ac-D-CAY-NH_2 conjugated with N-allyl acetamide (**78**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 464.2509. Observed Mass $[M+H]^+$: 464.2507. 219

Figure 124 - Analytical HPLC trace of pure H-HRLLRCGNYA-NH_2 conjugated with N-allyl acetamide (**80**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1382.7269, $[M+2H]^{2+}$: 634.8709. Observed Mass $[M+H+TFA]^+$: 1382.7260, $[M+2H]^{2+}$: 634.8702. 220

Figure 125 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N-allyl succinimide (**86**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1326.7395, $[M+2H]^{2+}$: 663.8737. Observed Mass $[M+H]^+$: 1326.7409, $[M+2H]^{2+}$: 663.8742. 221

Figure 126 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N-allyl formamide (**87**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1368.7112, $[M+2H]^{2+}$: 627.8631. Observed Mass $[M+H+TFA]^+$: 1368.7123, $[M+2H]^{2+}$: 627.8653. 222

Figure 127 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N-methylallylamine (**88**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1354.7319, $[M+2H]^{2+}$: 620.8735. Observed Mass $[M+H+TFA]^+$: 1354.7262, $[M+2H]^{2+}$: 620.8728. 223

Figure 128 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N,N-dimethylallylamine (**89**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1368.7476, $[M+2H]^{2+}$: 627.8813. Observed Mass $[M+H+TFA]^+$: 1368.7414, $[M+2H]^{2+}$: 627.8783. 224

Figure 129 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with trimethylallylammonium chloride (**90**). Analytical gradient 5-30% B over 10 min, 210 nm. Calculated Mass $[M+H]^{2+}$: 634.8891. Observed Mass $[M+H]^{2+}$: 634.8786. 225

Figure 130 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N-allyl benzenamide (**91**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1444.7425, $[M+2H]^{2+}$: 665.8788. Observed Mass $[M+H+TFA]^+$: 1444.7340, $[M+2H]^{2+}$: 665.8773. 226

Figure 131 - Analytical HPLC trace of pure K48C ubiquitin conjugated with N-allyl acetamide (**92**). Analytical gradient 25-35% B over 10 min, 210 nm. Calculated Mass $[M+6H]^{6+}$: 1433.2721, $[M+7H]^{7+}$: 1228.6629, $[M+8H]^{8+}$: 1075.2061, $[M+9H]^{9+}$: 955.8507, $[M+10H]^{10+}$: 860.3664, $[M+11H]^{11+}$: 782.2429. Observed Mass $[M+6H]^{6+}$: 1433.7654, $[M+7H]^{7+}$: 1229.3814, $[M+8H]^{8+}$: 1075.6395, $[M+9H]^{9+}$: 956.2682, $[M+10H]^{10+}$: 860.7662, $[M+11H]^{11+}$: 782.5936. 227

Figure 132 - Analytical HPLC trace of pure Ac-CWHISKEY-NH_2 conjugated with 3-methyl-3-buten-1-ol (**94**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 580.8091. Observed Mass $[M+2H]^{2+}$: 580.8094. 228

Figure 133 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugated with 3-methyl-3-buten-1-ol (**95**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^{+}$: 1418.7684, $[M+2H]^{2+}$: 709.8881. Observed Mass $[M+H]^{+}$: 1418.7758, $[M+2H]^{2+}$: 709.8739. 229

Figure 134 - Analytical HPLC trace of pure Ac-CAY-NH_2 conjugated with 3-methyl-3-buten-1-ol (**96**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^{+}$: 451.2557. Observed Mass $[M+H]^{+}$: 451.2555. 230

Figure 135 - Analytical HPLC trace of pure Ac-D-CAY-NH2 conjugated with 3-methyl-3-buten-1-ol (**97**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass [M+H]+: 451.2557. Observed Mass [M+H]+: 451.2562. 231

Figure 136 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugated with prop-biotin (**103**). Analytical gradient 15-45% B over 10 min, 210 nm. Calculated Mass $[M+H]^{+}$: 1629.8463, $[M+2H]^{2+}$: 815.4271. Observed Mass $[M+H]^{+}$: 1629.8422, $[M+2H]^{2+}$: 815.4277. 232

Figure 137 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugated with bute-biotin (**104**). Analytical gradient 15-45% B over 10 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 822.4349. Observed Mass $[M+2H]^{2+}$: 822.4332. 233

Figure 138 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 modified with N,N,2-trimethylprop-2-en-1-amine (**105**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^{+}$: 1431.8000, $[M+2H]^{2+}$: 716.4039. Observed Mass $[M+H]^{+}$: 1431.7973, $[M+2H]^{2+}$: 716.4032. 234

Figure 139 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 modified with N,N-3-trimethylbut-3-en-1-amine (**106**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^{+}$: 1445.8157, $[M+2H]^{2+}$: 723.4118. Observed Mass $[M+H]^{+}$: 1445.8164, $[M+2H]^{2+}$: 723.4015. 235

Figure 140 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 modified with N,N,N-3-tetramethylbut-3-en-1-aminium iodide (**107**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M]^+$: 1459.8313, $[M+H]^{2+}$: 730.4196. Observed Mass $[M]^+$: 1459.8308, $[M+H]^{2+}$: 730.4202. 236

Figure 141 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 modified with diethyl(2-methylallyl)phosphonate (**108**). Analytical gradient 15-45% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1524.7867, $[M+2H]^{2+}$: 762.8973. Observed Mass $[M+H]^+$: 1524.7744, $[M+2H]^{2+}$: 762.8875. 237

Figure 142 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugated with 2-methyl-2-propene-1-sulfonic acid (**112**). Analytical gradient 15-45% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1386.7422, $[M+2H]^{2+}$: 693.8750. Observed Mass $[M+H]^+$: 1386.7421, $[M+2H]^{2+}$: 693.8765. 238

Figure 143 – Analytical HPLC trace of pure Ac-YEPLAXHISKY-NH_2 hydrofluorinated (**113**). Analytical gradient 20-30% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1406.7484, $[M+2H]^{2+}$:703.8781. Observed Mass $[M+H]^+$: 1386.7421, $[M+2H]^{2+}$: 703.8770. 239

Figure 144 - Analytical HPLC trace of pure Ac-CAY-NH_2 conjugated with 2-Methyl-2-propene-1-sulfonic acid sodium salt (**114**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 419.2294. Observed Mass $[M+H]^+$: 419.2288. 240

Figure 145 - Analytical HPLC trace of pure Ac-CAY-NH_2 modified with 2-Methyl-2-propene-1-sulfonic acid sodium salt, then hydrofluorinated (**115**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 439.2356. Observed Mass $[M+H]^+$: 439.2349. 242

Figure 146 – Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugation/fluorination, non-fluorinated product (**95**). Peak at 1019 is unrelated to this sample – MS impurity. Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1418.7684, $[M+2H]^{2+}$: 709.8881. Observed Mass $[M+H]^+$: 1418.7609, $[M+2H]^{2+}$: 709.8862. 244

Figure 147 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugation/fluorination, fluorinated product (**116**). Peak at 1019 is unrelated to this sample – MS impurity. Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1436.7503, $[M+2H]^{2+}$: 718.8794. Observed Mass $[M+H]^+$: 1418.7609, $[M+2H]^{2+}$: 718.8834. 244

Figure 148 - Analytical HPLC trace of pure K48C ubiquitin conjugated with N,N,N-3-tetramethylbut-3-en-1-aminium iodide (**118**). Analytical gradient 25-35% B over 10 min, 210 nm. Calculated Mass $[M+4H]^{5+}$: 1725.7417, $[M+5H]^{6+}$: 1438.2860, $[M+6H]^{7+}$: 1232.9606, $[M+7H]^{8+}$: 1078.9665, $[M+8H]^{9+}$: 959.1933, $[M+9H]^{10+}$: 863.2740. Observed

Mass $[M+4H]^{5+}$: 1726.3903, $[M+5H]^{6+}$: 1438.6706, $[M+6H]^{7+}$: 1233.1971, $[M+7H]^{8+}$: 1079.2648, $[M+8H]^{9+}$: 959.5068, $[M+9H]^{10+}$: 860.8329..... 245

Figure 149 - Analytical HPLC trace of pure K48C ubiquitin conjugated with biotin-butene (**117**). Analytical gradient 25-35% B over 10 min, 210 nm. Calculated Mass $[M+6H]^{6+}$: 1468.6219, $[M+7H]^{7+}$: 1258.9627, $[M+8H]^{8+}$: 1101.7184, $[M+9H]^{9+}$: 979.4172, $[M+10H]^{10+}$: 881.5762, $[M+11H]^{11+}$: 801.5246. Observed Mass $[M+6H]^{6+}$: 1469.4472, $[M+7H]^{7+}$: 1259.7212, $[M+8H]^{8+}$: 1102.2764, $[M+9H]^{9+}$: 980.0792, $[M+10H]^{10+}$: 882.1485, $[M+11H]^{11+}$: 801.9057. 246

Figure 150 - Analytical HPLC trace of pure Ac-CAY-NH_2 modified with 2-nitrophenylboronic acid through desulfurative cross-coupling (**120**). Analytical gradient 5-60% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 486.1989. Observed Mass $[M+H]^+$: 486.1805. 248

Figure 151 – HPLC analysis of the crude cyclisation of 100 after 10 minutes. Analytical gradient 10-30% B over 10 mins, 210 nm. 249

Figure 152 - Pure analytical HPLC trace and MS of cyclised H-S(2-methylallyl)AFAC-NH_2 (**123**), diastereomer 1. Analytical gradient 10-30% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 519.2931. Observed Mass $[M+H]^+$: 519.2927. 250

Figure 153 - Pure analytical HPLC trace and MS of cyclised H-S(2-methylallyl)AFAC-NH_2 (**123**), diastereomer 2. Analytical gradient 10-30% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 519.2931. Observed Mass $[M+H]^+$: 519.2935. 250

Figure 154 - Pure analytical HPLC trace and MS of cyclised H-S(Allyl)AFAC-NH_2 (**126**). Analytical gradient 2-25% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 505.2775. Observed Mass $[M+H]^+$: 505.2785. 251

Figure 155 - Pure analytical HPLC trace and MS of desulfurised and cyclised H-alGAFAC-NH_2 (**130**). Analytical gradient 2-25% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 475.2669. Observed Mass $[M+H]^+$: 465.2668. 252

Figure 156 – NMR comparison of pure desulfurised and cyclised H-alGAFAC-NH_2 (**130**, top) and the starting peptide (**129**, bottom), showing disappearance of allylic protons at 5.71 and 5.15 ppm. 253

Figure 157 - Pure analytical HPLC trace and MS of desulfurised and cyclised H-alGMPWLRX-NH_2 (**133**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass [M+H]+: 869.4820, [M+2H]2+: 435.2449. Observed Mass [M+H]+: 869.4786, [M+2H]2+: 435.2445. 254

Figure 158 - NMR overlay comparing oxytocin model starting peptide H-alGMPWLRC-NH_2 (**131**, bottom) and the cyclised product (**133**, top), showing disappearance of allylic proton peaks at 5.75 and 5.20 ppm. 255

Figure 159 - Pure analytical HPLC trace and MS of desulfurised and TCEP nucleophilic addition product H-A(TCEP)YEPLAWHISKEYA-NH2 (**134**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass $[M+2H]^{2+}$: 977.4678, $[M+3H]^{3+}$: 651.9811. Observed Mass $[M+2H]^{2+}$: 977.4799, $[M+3H]^{3+}$: 651.9608. 256

Figure 160 - Pure starting material H-alGYEPLAWHISKEYC-NH_2 (**132**, bottom) and purified product (**134**, top) NMR comparison, showing disappearance of allylic protons. 257

Figure 161 - ^{31}P NMR ((162 MHz, DMSO) δ 36.62.) of the TCEP-nucleophilic addition product (**134**) with the desulfurised model peptide H-alGYEPLAWHISKEYC-NH_2. 257

Figure 162 - Analytical HPLC trace and MS of the purified inseparable mixture of cyclised and uncyclised product H-alGYEPLAWHISKEYX-NH_2 (**135**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^{+}$: 1703.8671, $[M+2H]^{2+}$: 852.4375. Observed Mass $[M+H]^{+}$: 1703.8773, $[M+2H]^{2+}$: 852.4380. 258

Figure 163 – Pure starting material H-alGYEPLAWHISKEYC-NH_2 (**132**, bottom) and purified isolated mixture of desulfurised and cyclised/uncyclised material (**135**, top) NMR comparison, showing reduction in the integration of allylic protons. 259

Figure 164 - Pure analytical HPLC trace and MS of desulfurised and cyclised H-alGEDLFYQC-NH_2 (**138**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^{+}$: 961.4682. Observed Mass $[M+H]^{+}$: 981.4686. 260

Figure 165 - NMR overlay of starting peptide H-alGEDLFYQC-NH_2 (**136**, bottom) and cyclised product (**138**, top), showing disappearance of allylic proton peaks at 5.75 and 5.20 ppm. 261

Figure 166 – Pure analytical HPLC trace and MS of desulfurised and cyclised H-alGDKFNHEC-NH_2 (**139**). Analytical gradient 2-25% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^{+}$: 956.4590, $[M+2H]^{2+}$: 478.7334. Observed Mass $[M+H]^{+}$: 956.4592, $[M+2H]^{2+}$: 478.7347. 262

Figure 167 - NMR comparison of pure desulfurised and cyclised H-alGDKFNHEC-NH_2 (**139**, top) and the starting peptide (**137**, bottom), showing disappearance of allylic protons at 5.71 and 5.15 ppm. 263

Figure 168 - Pure analytical HPLC trace and MS of desulfurised and cyclised H-DalGFYQC-NH_2 (**141**). Analytical gradient 2-40% B over 10 minutes, 210 nm. Calculated Mass

[M+H]$^{+}$: 739.3415, [M+2H]$^{2+}$: 370.1748. Observed Mass [M+H]$^{+}$: 739.3438, [M+2H]$^{2+}$: 370.1750. 264

Figure 169 - NMR overlay of starting peptide H-DalGFYQC-NH_2 (**140**, bottom) and the cyclised product (**141**, top), showing disappearance of allylic protons at 5.7 and 5.0 ppm. This also shows a lack of characteristic 3H doublet for Ala (usually around 1.3 ppm), also indicating cyclisation has occurred. 265

Figure 170 - Pure analytical HPLC trace and MS of desulfurised H-alGYIQNXPLG-NH_2 (**143**). Analytical gradient 10-30% B over 10 minutes, 210 nm. Calculated Mass [M+H]$^{+}$: 971.5314. Observed Mass [M+H]$^{+}$: 971.5343. 266

Figure 171 - NMR overlay of starting peptide H-alGYIQNXPLG-NH_2 (**142**, bottom) and product (**143**, top). This shows the continuing presence of the allylic protons at 5.7 and 5.15 ppm, indicating cyclisation was not successful. 267

Figure 172 - Pure analytical HPLC trace and MS of cyclised H-XYIQNalGPLG-NH_2 (**145**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass [M+H]$^{+}$: 971.5314, [M+2H]$^{2+}$: 486.2697. Observed Mass [M+H]$^{+}$: 971.5319, [M+2H]$^{2+}$: 486.2710. 268

Figure 173 - Pure analytical HPLC trace and MS of linear desulfurised H-XYIQNalGPLG-NH_2 (**146**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass [M+H]$^{+}$: 971.5314, [M+2H]$^{2+}$: 486.2697. Observed Mass [M+H]$^{+}$: 971.5324, [M+2H]$^{2+}$: 486.2706. 269

List of Schemes

Scheme 1 - General scheme for peptide synthesis using SPPS. 9

Scheme 2 – Generic mechanism of native chemical ligation. 13

Scheme 3 - Reaction scheme for the radical induced desulfurisation of Cys to Ala. 13

Scheme 4 - General scheme showing the ligation of two peptide fragments using diselenide-selenoester ligation (DSL). 15

Scheme 5 - Proposed reaction mechanism for the generation of the selenoate which participates in DSL. 15

Scheme 6 - Reaction scheme for the oxidative deselenisation of selenocysteine, converting Sec into serine. 16

Scheme 7 - Generic modification using maleimides in buffered (phosphate, tris, etc.) solution at room temperature. A reducing agent, such as TCEP, is often included to ensure Cys remains in the reduced form. 19

Scheme 8 - Cys conjugation with phosphonamidate using a high excess (1000 eq.) of 1,4-dithiothreitol (DTT) in aqueous conditions with an organic co-solvent (dimethyl sulfoxide (DMSO)) for 16 hours. 19

Scheme 9 - Modification of Cys residue with quaternary vinyl pyridines at 37 ^{o}C for 1 hour. 20

Scheme 10 - Site-selective modification using EBX reagents for dual orthogonal protein labelling. 20

Scheme 11 - Strain-release modification of cysteine using bicyclic compounds. 21

Scheme 12 - Example reaction for the glycosylation of a protein followed by conversion to the stable thioether by reductive elimination followed by Michael addition. 21

Scheme 13 - General scheme for the modification of Cys residues with iodoacetamide. 22

Scheme 14 - Reaction scheme showing the conversion of Cys to a dimethyl lysine mimetic. 22

Scheme 15 - General reaction scheme for the arylation of Cys residues using organometallic palladium reagents. 23

Scheme 16 - Ni(II) catalysed cross coupling between Cys and 2-nitrophenylboronic acid. .. 23

Scheme 17 -Photocatalysed radical thiol-ene modification of Cys. 24

Scheme 18 - Proposed mechanism for the photocatalysed thiol-ene reaction. 25

Scheme 19 - Photocatalytic S-arylation of Cys containing peptides using diazonium salts of arenes. 25

Scheme 20 - General scheme for Ru/Ni catalysed cross coupling (Y = C or N). The mechanism shows how the thiyl radical is formed using the alkyl radical generated by the single electron transfer of an excited ruthenium catalyst to a silicate. 26

Scheme 21 - General scheme for the desulfurative alkylation method developed by Yingwei Wang et al. [81] 27

Scheme 22 - General scheme for the conversion of Cys to Dha, followed by radical addition into the double bond. 27

Scheme 23 - General scheme for the formation of allyl Cys and cross-metathesis for selective protein modification. 28

Scheme 24 - Alkylation of Sec with phenyloxadiazolyl sulfone. 29

Scheme 25 - Copper catalysed arylation of Sec.[90] 29

Scheme 26 - Arylation of Sec without the requirement of a metal catalyst.[91] 30

Scheme 27 - Modification of Sec-allyl by cross-metathesis to install biologically relevant molecules.[92] 30

Scheme 28 - General scheme for the spontaneous cyclisation of peptides baring an internal Cys residue and a terminal N-chloroacetamide group. 32

Scheme 29 - An example of on-resin macrocyclisation using RCM. 33

Scheme 30 - General scheme for peptide macrocyclisation using modified Cys and Lys residues for diels-alder cyclisation. 33

Scheme 31 - Peptide macrocyclisation using perfluoroaryl-cysteine S_NAr chemistry. 34

Scheme 32 - Reaction scheme for the macrocyclisation of peptide via photoredox-catalysed C-terminal decarboxylation. 34

Scheme 33 - A - direct labelling, B - indirect labelling, C - indirect labelling via complexation. 35

Scheme 34 - General scheme for the TEMPO conjugation of Cys. 37

Scheme 35 - Example of proposed reaction with allyl acetamide. The resultant product is an acetylated lysine. This is a common post-translational modification installed on Lys. .. 38

Scheme 36 - Proposed scheme for peptide macrocyclisation using desulfurative allyl trapping chemistry. 38

Scheme 37 - General scheme for the reaction to be optimised during this chapter. TCEP will be used to induce deselenisation, the persistent radical on TEMPO will trap the resultant alanyl radical to form the desired conjugate..... 40

Scheme 38 - General scheme for the proposed one-pot ligation and conjugation procedure. 40

Scheme 39 - Reaction scheme for TEMPO trapping reactions using H-UF-OMe dipeptide model. 41

Scheme 40 - General reaction scheme for entries 1 to 3. The equivalents of TCEP and TEMPO used in the reaction are shown in Table 1. 42

Scheme 41 - Isolated percentage yields from modified TEMPO trapping reactions. 44

Scheme 42 - General scheme for the conjugation of TEMPO onto H-UAF-OMe (**10**), forming conjugated product **11**. 47

Scheme 43 - Proposed reaction mechanism for the photoredox catalysed deselenisation of selenocysteine. 56

Scheme 44 - Expected deselenisation product (**14**) from the treatment of the Sec containing peptide Ac-UHISKY-OH (**13**) with TCEP, accelerated by Mn(OAc)3..... 60

Scheme 45 - Reductive cleavage of TEMPO conjugate (**11**) to form Ser-containing peptide (**18**)..... 64

Scheme 46 - Reaction scheme showing the conjugation of H-UAF-OMe (**10**) tripeptide model with TEMPO (**2**) to give desired product (**11**). 67

Scheme 47 - Reaction scheme for the synthesis of biotin-TEMPO probe..... 68

Scheme 48 - Reaction scheme for the conjugation of H-UAF-OMe (10) with a variety of TEMPO-based probes. 69

Scheme 49 - General reaction scheme for the one-pot ligation and conjugation of a model peptide sequence. 73

Scheme 50 - Reaction scheme for the one-pot ligation and conjugation of Ac-YEPLA-SePh (**32**) and H-UHISCY-NH^2 (**39**). 75

Scheme 51 - Reaction conditions for the one-pot ligation-conjugation-reductive cleavage of Ac-YEPLA-SePh (**32**) and H-UWIMKY-NH2 (**16**)..... 77

Scheme 52 - Reaction scheme for the one-pot ligation and conjugation of the two fragments of the ZEGFR protein. 78

Scheme 53 - Proposed mechanism for the TEMPO bioconjugation of Cys residues. 80

Scheme 54 - The general reaction scheme for the conjugation of H-CAF-OMe (**48**) with TEMPO (**2**) to form conjugated product (**11c**). 81

Scheme 55 - General reaction scheme for the bioconjugation of Ac-CWHISKEY-NH2 (41) with TEMPO based probes. 84

Scheme 56 – General reaction scheme for the 2-step ligation and conjugation between Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**) to produce TEMPO conjugated products (**35b** and **38b**). 86

Scheme 57 - General reaction scheme for the conjugation of Ac-YEPLACHISKY-NH2 (**56**) with propargyl TEMPO (**23**) forming the conjugated product (**35b**). 87

Scheme 58 - Reaction scheme for the reductive cleavage of TEMPO conjugation (**35b**) to form the native Ser product (**57**). 90

Scheme 59 - Reaction scheme for the dual functionalisation of a peptide with a Sec and Cys residue (H-UHISCY-NH_2 (**39**)). 91

Scheme 60 - Reaction scheme for the bioconjugation of K48C ubiquitin with propargyl TEMPO (**53**)..... 92

Scheme 61 - Reaction mechanism for the conversion of Sec to acetyl lysine..... 97

Scheme 62 - Reaction scheme for the trapping of allyl alcohol by Sec during radical-induced deselenisation, showing the desired product (**63**) and the two potential by-products Ala (**64**) and Ser (**65**). 98

Scheme 63 - Reaction scheme for Sec allyl trapping initiated by eosin Y using the Ac-UHISKY-OH (**13**) starting peptide forming conjugated product (**67**). 100

Scheme 64 - Reaction scheme for the conjugation of Ac-CHISKY-OH (**15**) with N-allyl acetamide to give product (**67b**). 102

Scheme 65 - Reaction scheme for the modification of H-CF-OMe (**68**) with N,N-dimethylallylamine (**69**) using a range of initiators to form trapped product (**70**). Observed by-products include H-atom trapped Ala product (**4**) and multiple addition product (**71**), where the secondary radical formed reacts with another equivalent of DMAA (**69**).... 103

Scheme 66 - General reaction mechanism for the photoredox cycle for the photocatalysts used in Table 12. 104

Scheme 67 - Reaction scheme for the bioconjugation of Ac-CWHISKEY-NH_2 (**49**) with N-allyl acetamide (**66**) forming the conjugated product (**72**), using the altered conditions shown in Table 17. 110

Scheme 68 - Bioconjugation of Ac-CWHISKEY-NH2 (**49**) with N-allyl acetamide (**66**) using fully optimised conditions to form product (**72**)..... 116

Scheme 69 - Reaction scheme for the conversion of H-HRLLRCGNYA-NH_2 (**79**) to the corresponding acetyl Lys baring peptide (**80**), varying the equivalents of TCEP and Ir(III). 118

Scheme 70 - General scheme for the conjugation of H-HRLLRCGNYA-NH_2 (**79**) with a range of allyl compounds. The yields shown are isolated yields. 119

Scheme 71 - Reaction scheme for the installation of acetyl lysine into K48C ubiquitin (**92**). 120

Scheme 72 - Mechanism for the reaction between the alanyl radical and 3-methyl-3-buten-1-ol (**93**), producing a stabilised tertiary radical intermediate. 121

Scheme 73 - Reaction scheme for the bioconjugation of the model peptide Ac-CWHISKEY-NH_2 (**49**) with 3-methyl-3-buten-1-ol (**93**) forming the conjugated product (**94**). 121

Scheme 74 – Conjugation of peptide (**56**) with MBO (**93**) to form desired product (**95**). 122

Scheme 75 - Conjugation of Ac-YEPLACHISKY-NH_2 (**56**) with 3-methyl-3-buten-1-ol (**93**) to form conjugated product (**95**)..... 124

Scheme 76 - Overview of the reaction of Ac-YEPLACHISKY-NH_2 (**56**) with a range of stabilised alkene substrates (**98-104**) to form desired conjugates (**105-110**). The yields specified are isolated yields..... 126

Scheme 77 - The reaction scheme showing the unexpected product when Ac-YEPLACHISKY-NH_2 (**56**) was conjugated with 2-methyl-2-propen-1-sulfonic acid sodium salt (**111**), forming the elimination product (**112**). 126

Scheme 78 – Reaction scheme for the hydrofluorination of peptide **112** using literature conditions as shown, forming the hydrofluorinated product (**113**) in a 69% isolated yield. 127

Scheme 79 - Reaction scheme with conditions and isolated yields for the conjugation (**114**) and hydrofluorination (**115**) of Ac-CAY-NH_2 (**75**)..... 127

Scheme 80 - Reaction scheme showing the attempted fluorination reaction by trapping the intermediate tertiary radical formed after alanyl radical addition into the stabilised alkene. 129

Scheme 81 - Summary of the bioconjugation of K48C ubiquitin with stabilised alkene probes containing biotin and a trimethyl-lysine mimetic..... 130

Scheme 82 - First trial desulfurative cross-coupling combining the conditions found in literature with the novel desulfurisation conditions used previously in this chapter to form product (**120**)..... 131

Scheme 83 - Reaction scheme for the macrocyclisation of (**122**) to form cyclic peptide (**123**). 133

Scheme 84 - Reaction scheme for the macrocyclisation of an allyl glycine containing peptide (**129**) through desulfurative radical addition to form cyclic product **130**..... 138

Scheme 85 – Reaction scheme showing the desulfurisation of starting peptide (**142**) to produce Ala containing peptide (**143**). 143

1. Introduction

1.1 Chemical Modification of Peptides and Proteins

Over the last four decades, there has been a steady increase in the number of polypeptide-based *biologics* approved for use as therapeutics; to date, over 140 proteins and peptides have been approved for clinical use.[1] Currently, 7/10 highest-grossing drugs are proteins (specifically, antibodies); this class of therapeutic now represents the front-line treatment for many diseases.[2] Such therapeutics are identical or very similar to naturally occurring proteins, resulting in minimal chance of an adverse immune response occurring during treatment. Furthermore, native polypeptides exhibit extraordinary target-specificity due to the process of evolution, reducing instances of off-target binding and associated side effects.[3] This specificity has also been exploited as a vehicle for drug delivery via conjugation of a therapeutic small molecule to a protein that targets the cell-line of interest, offering a cell-specific therapy.[3]

The modification of proteins can serve many different functions, including monitoring natural systems, developing unnatural proteins with enhanced/engineered activity, and the preparation of therapeutic bioconjugates. Protein drug conjugates are an emerging class of therapeutics; they involve the selective, covalent attachment of a drug compound to the protein. Since the protein sequence can be designed/selected to target a particular receptor, these conjugates have great potential to be highly selective to a specific target.[4,5] For example, antibody drug conjugates (ADCs) have received significant attention as selective therapeutics, particularly against cancers.[6] Currently, 9 ADCs have been approved for clinical use[7]; the global ADC market is expected to reach approximately $10 bn by 2025 and continue to grow beyond this value.

An issue with the use of polypeptides as drugs concerns their short half-life *in vivo* due to enzymatic degradation; this challenge is especially relevant for peptide therapeutics. Chemical modification can be tailored to enhance the longevity of a peptide or protein *in vivo*, thus increasing its potential as a drug. A common way to achieve this is by covalently attaching a PEG (polyethylene glycol) group to the polypeptide, known as PEGylation. These groups are non-toxic and non-immunogenic, making them suitable for use with therapeutics. Typically, using a longer PEG chain leads to increased activity, higher proteolysis resistance and larger intestinal uptake.[8] Alternative strategies to improve the poor pharmacological properties of peptides and proteins include glycosylation. The position of glycosylation within the polypeptide is crucial, affecting peptide-receptor interactions, biodistribution and pharmacological activity. Glycosylation has been demonstrated as a promising modification to

improve the biodistribution and poor pharmacokinetic profile of radiopharmaceuticals for diagnostics and therapy.[9]

Beyond therapeutics, the chemical modification of peptides and proteins enables the preparation of tools to investigate fundamental biochemical processes, imaging and diagnostic agents, and biomaterials.[10] Due to the broad applications of chemically modified polypeptides, versatile methods to enable the site-selective modification of these biomolecules are in high demand.

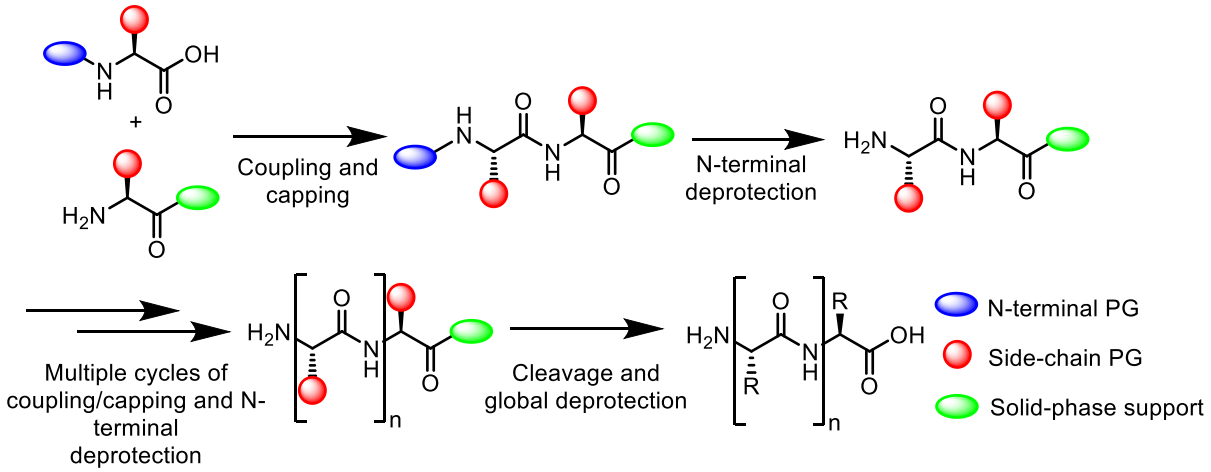
The 21 canonical amino acids which make up the eukaryotic and prokaryotic proteomes present a plethora of chemical functionality which can be utilised to covalently attach a desired small molecule. A central issue of protein modification, therefore, concerns site and residue selectivity. Poor selectivity results in the production of a heterogeneous mixture of modified isoforms. This can be due to the method not being residue specific, leading to multiple residues in a sequence being modified. Alternatively, a method may be specific to a certain residue but not site-specific. This will result in modification of all instances of the target residue, at random, leading to installation of the modified group in varying ratios across the batch. Such isoforms can be challenging, and in some cases impossible, to separate. In the context of protein-drug conjugates, such issues will affect the toxicity of the batch that may vary between batches depending on the degree of modification.

With the increased demand for biologics in medicine, site-selective methods for peptide and protein modification are desired. Thus, in recent years, methods that enable selective modification are becoming more prominent in modern research.

1.2 Chemical Synthesis of Peptides and Proteins

Peptides and proteins used for therapeutics can be produced either chemically or biologically. Chemical peptide synthesis was originally performed using standard solution-phase organic chemistry. This approach required phase extraction and extensive purification at each stage of synthesis. In the 1960's Bruce Merrifield developed solid-phase peptide synthesis (SPPS).[11,12] SPPS involves employing an insoluble solid support functionalised with an appropriate linker. A variety of linkers have been developed that allow the peptide to be cleaved under a range of different conditions. An example is the 2-chlorotrityl chloride (CTC) group that allows cleavage under mildly acidic conditions, as the resulting cleavage product is a stabilised tertiary cation. The insoluble solid support is made from polystyrene which is 1-2% cross-linked with divinylbenzene.[13,14] The synthesis of a peptide on a solid support allows all solvent and reagents to be washed away from the resin using a porous filter. An excess of reagents can therefore be used, and extensive purification steps are unnecessary, reducing the time required for synthesis. SPPS is a stepwise method of synthesis; the general scheme

for this method is shown in Scheme 1. The first amino acid in the sequence is attached to this solid support via the linker. Once attached, the peptide can be synthesised via iterative coupling steps of each incoming amino acid with those on the resin.



Scheme 1 - General scheme for peptide synthesis using SPPS.

The N-terminus of each amino acid being coupled must be protected to prevent multiple addition and polymerisation from occurring. Additionally, the side-chain of the amino acids, both the incoming residue and those already incorporated in the chain, must also be protected. This prevents unwanted side-reactions from occurring. An example of a protected amino acid is shown in Figure 1.

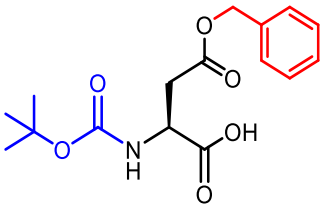


Figure 1 - Example protected amino acid (Asp) showing the N-terminal protecting group (Boc, blue) and the side-chain protecting group (Benzyl, red).

After coupling, the N-terminal protecting group must be removed, in a stage called deprotection. To prevent side-reactions, the N-terminal protecting group must be able to be removed selectively in the presence of the side-chain protecting groups. The original N-terminal protecting group was the Boc group, requiring 50% trifluoroacetic acid (TFA) in dichloromethane (DCM) to be removed. The Merrifield resin used therefore had to be stable to these conditions. The side-chain protecting groups also had to be stable to these conditions, resulting in groups such as benzyl, cyclohexyl or 2,6-dichlorobenzyl being used. Once the peptide was completed, cleavage was performed. Due to the high stability of the side chain protecting groups and the resin linker, acids including hydrofluoric acid (HF) were required.

This made the procedure unsafe, requiring specialised reaction vessels and protective equipment for handling HF. If benzyl-protecting groups were used to protect aspartic acid (Asp) or glutamic acid (Glu), a second deprotection step using hydrogenation was required.

The drawbacks of HF led to the development of alternative N-terminal protecting groups. Sheppard developed the Fmoc group in 1978, which is now the most commonly used N-terminal protecting group.[15] The Fmoc group is base-labile, typically removed using 20% piperidine in *N,N*-dimethylformamide (DMF).[16] This is a mild base, and thus removal conditions are simple and safe. This allowed the side-chains of amino acids to be protected using less stable protecting groups, such as trityl (Trt), tert-butyl (*t*Bu) and *tert*-Butyloxycarbonyl (Boc), which are removed using TFA. In comparison to HF, this is much safer and does not require specialised equipment or safety protocols. The use of a base-labile N-terminal protecting group and acid labile side chain-protecting group shows the orthogonality of the modern protecting group strategy. The structure of protected aspartic acid is shown in Figure 2 below with orthogonal protection on the N terminus and side chain, base labile Fmoc and acid-labile *t*Bu ester respectively.

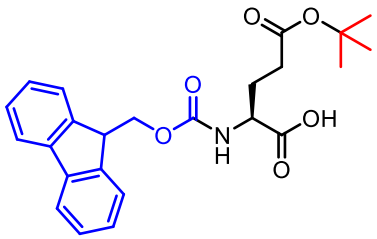


Figure 2 - Structure of Fmoc-Glu(OtBu)-OH.

The introduction of Fmoc chemistry also led to the development of resins that could be cleaved using the same conditions as the side-chain protecting groups, such as Rink Amide, 2-chlorotrityl chloride and Wang. These resins allow for global deprotection of the peptide and cleavage from the resin in a single step. The Wang and chlorotrityl resins are used when a C-terminal carboxylic acid is required post-cleavage. The advantage of the chlorotrityl resin over Wang is the lability of the former in weak acid (acetic acid). The peptide can be cleaved using mildly acidic conditions, resulting in the side-chain protecting groups remaining intact. The C-terminus of the peptide can then be selectively modified, followed by global protecting group removal. The Rink Amide resin affords a peptide bearing a C-terminal primary amide, providing a simple and efficient method for C-terminal amide synthesis.

The peptide chain is extended by activating the amino acid to be coupled to the peptide, forming an amide bond between the resin-bound amino acid and the next amino acid in the sequence. This is achieved by activating the carboxylic acid of the amino acid to be coupled; making it susceptible to nucleophilic attack from the resin bound peptide. The activation is

achieved by the inclusion of coupling reagents. Two of the main classes of coupling reagents used are uronium salts and carbodiimides. A common example of each is *N,N,N',N'*-tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU, uronium) and diisopropylcarbodiimide (DIC, carbodiimides); the structures of each are shown in Figure 3.

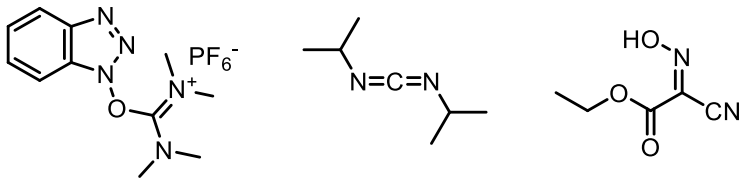


Figure 3 - Chemical structures of HBTU (left), DIC (middle) and Oxyma (right)

When the coupling reagent reacts with the carboxylic acid, it forms an active ester. This is significantly more reactive towards nucleophiles than the original carboxylic acid, allowing reactions to go to completion in less than one hour.[17,18] Using HBTU alone leads to racemisation, so additives such as 1-hydroxybenzotriazole (HOBt) are used to supress this.[19,20] Alternative uronium-based coupling reagents have been developed, such as HATU, which have enhanced coupling efficiency. Drawbacks to these reagents include the potential for guanidinylation.[20,21] This occurs when the coupling reagent reacts with the unprotected primary amine of the peptide, resulting in irreversible capping of the peptide chain. Repeated exposure of the user to HBTU, HATU and HCTU have also been reported to cause anaphylaxis.[22] The highlighted issues have led to the development of more efficient reagents that do not carry potentially fatal health risks. The combination of DIC and ethyl cyano(hydroxyimino)acetate (Oxyma, Figure 3) has been shown to be safer, result in higher percentage conversions and have low associated levels of racemisation.[23]

During peptide synthesis, it is common for couplings to not reach completion due to the coupling reaction not being 100% efficient. This leads to a proportion of peptide missing a single amino acid, known as a deletion sequence. In larger sequences multiple amino acids could be missing from a peptide sequence, these shorter peptides are known as truncated sequences. The presence of these incomplete sequences results in a difficult purification as they often have a very similar retention time to the desired peptide, especially if the missing residue does not affect the peptide polarity, e.g. glycine (Gly). To overcome this, peptide chemists utilise a capping step during synthesis. During this step, a large excess of acetic anhydride (Ac_2O) and base is added to the peptide resin, acetylating any free amines remaining after the coupling step. This significantly reduces the prevalence of closely eluting impurities resulting from deletions/truncations, and the addition of an acetyl group increases the hydrophobicity of these impurities. This step is very useful in enabling effective peptide

purification by high-performance liquid chromatography (HPLC) and can be easily incorporated into automated peptide synthesis.

The use of SPPS allows the simple introduction of unnatural amino acids; these can be introduced as pre-synthesised amino acids. Most unnatural amino acids could be incorporated using standard coupling conditions. If the desired unnatural amino acid is a modification of a natural amino acid, this can be performed on-resin. This technique utilises the orthogonality of SPPS. An example of this is the readily available building block Fmoc-Lys(Mmt)-OH. The side chain-protecting group, 4-methoxytrityl (Mmt), can be removed using 1% TFA in DCM. This allows removal without affecting the N-terminal Fmoc group or the other side-chain protecting groups. The resultant free amine can then be modified as desired.[24]

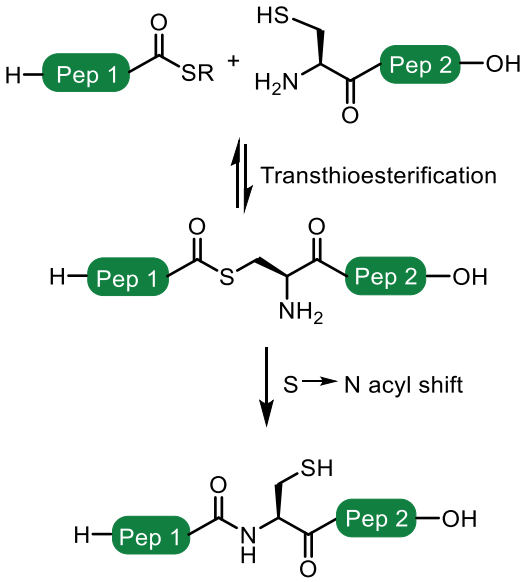
The ease of producing peptides by using SPPS techniques has led to the development of automated systems. Automated peptide synthesisers allow SPPS to be performed without the requirement of manual intervention. Further developments to these systems include the incorporation of microwave units. These reduce the time required for the coupling step in particular, with reaction times from as little as five minutes required. Analysis of peptides can also be automated, using analytical HPLC and electrospray ionisation MS (ESI-MS) to ascertain purity and identify desired products. Peptides can be easily purified using preparative HPLC, with fractions containing the desired products identified by ESI-MS and lyophilised to yield the desired pure peptide as a white fluffy solid.

A limitation of SPPS concerns the length of the peptide; a chain of up to around 70 amino acids can be synthesise on resin, (depending on the sequence).[15,16] This limitation led to the development of peptide ligation. Ligation involves chemoselectively joining two or more peptide fragments together to afford a larger peptide or small protein. This enables the chemical synthesis of larger peptides by synthesising separate fragments and then ligating them together.[13,15,16,25–27] Ligation can also be used to join a large protein produced recombinantly to a shorter, chemically synthesised fragment. This is useful as some proteins can be difficult or impossible to produce entirely via recombinant expression. Ligation of chemically synthesised fragments also allows the incorporation of unnatural modifications, widening the scope of use for these proteins.[28,29]

1.2.1 Ligation Techniques

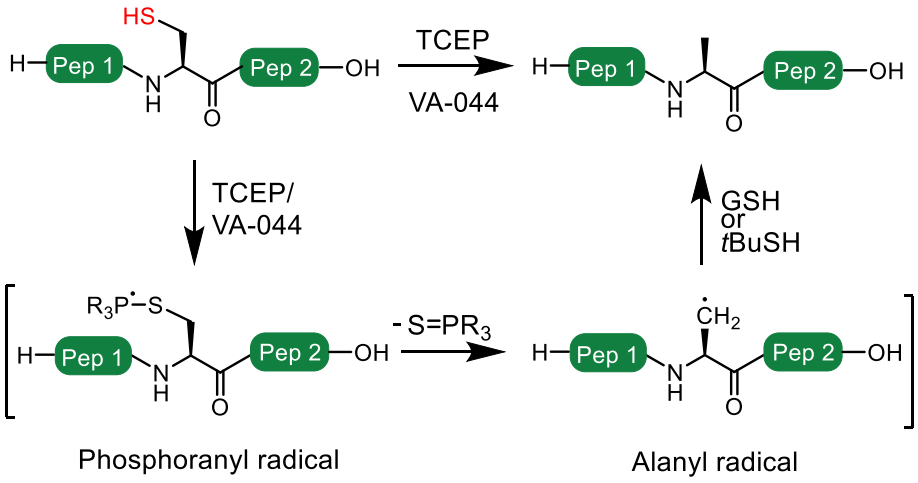
Ligation of peptide fragments has been achieved through a wide range of methods, only some of which result in the desired native amide bond being formed.[30] A successful and widely utilised method of ligation is native chemical ligation (NCL). This method enables the assembly of a peptide carrying an N-terminal Cys residue with a second sequence bearing a C-terminal thioester (Scheme 2). The reaction proceeds via an intermediate thioester followed by an S-

N acyl shift to link the two fragments via an amide bond. This technique has enabled the total chemical synthesis of proteins including HIV-1 protease, RNase A, EPO, etc.[31]



Scheme 2 – Generic mechanism of native chemical ligation.

As this technique can only be used with peptides containing a Cys residue, methods have been developed to modify the Cys residue post ligation. This is useful as not all native proteins contain a Cys or an appropriately placed Cys. The first reported method involved reduction of Cys to alanine (Ala) using either Raney Ni or Pd/Al_2O_3 with H_2.[32] Radical-induced desulfurisation of Cys to Ala has also been achieved by exposure to Tris(2-carboxyethyl)phosphine (TCEP), tert-butyl mercaptan (*t*BuSH) and 2,2'-Azobis[2-(2-imidazolin-2-yl)propane] Dihydrochloride (VA-044) at room temperature.[33] Both of these methods result in the formation of the naturally occurring and highly abundant amino acid Ala (Scheme 3).



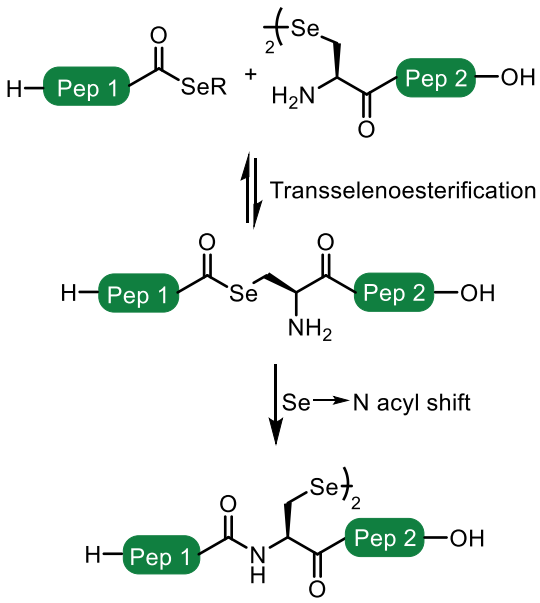
Scheme 3 - Reaction scheme for the radical induced desulfurisation of Cys to Ala.

Despite the highly useful conversion of Cys to Ala expanding the scope of NCL, as well as the development of several thiolated amino acids that enable ligation at multiple native junctions there are still many drawbacks with the technique: ligation reactions can take multiple days to reach completion (or may not reach completion at all) at more hindered residues, such as valine, isoleucine and proline.[31,34,35] Furthermore, the reaction has a narrow pH window and requires the addition of thiol additives to activate the stable alkyl thioester.[31] These issues have led chemists to find alternate methods for the ligation of peptide fragments to form chemically synthesised proteins.

1.2.2 Diselenide-Selenoester Ligation

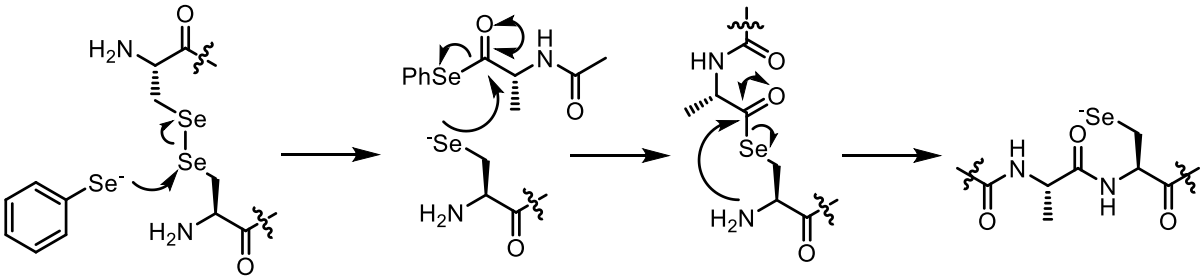
Selenium is known to play an important role in biological systems. More than 35 eukaryotic and prokaryotic proteins contain the residue selenocysteine (Sec). These are often involved in redox processes, with the Sec residue being essential to the catalytic cycle. An example of this is in the enzyme glutathione peroxidase (GPx), which is essential for hydroperoxide metabolism.[36,37] The Sec residue in GPx was shown to be responsible for the efficiency of the catalyst. Enzymes responsible for protecting the cell from oxidative damage, such as reductase enzymes, rely on the low redox potential of Sec (-381 mV).[38] This low redox potential results in Sec existing primarily in the oxidised, diselenide form. This diselenide dimer is referred to as selenocystine. The natural incorporation of Sec in proteins by biological systems led to it being labelled as the 21st proteinogenic amino acid.[36]

A recent advance in peptide ligation utilises the Sec residue, instead of Cys. The rate-limiting step in NCL is the nucleophilic attack of the thiol group to the thioester. The selenol group is more nucleophilic than the thiol group due to the enhanced polarizability of the selenium atom, resulting in accelerated ligation when using Sec.[27] Furthermore, when employing a peptide selenoester in place of the thioester, Sec ligations reach completion within minutes, without the need of additives. This technique, referred to as diselenide-selenoester ligation (DSL), (Scheme 4), also allows ligation to occur at the hindered residues valine (Val), isoleucine (Ile) and proline (Pro) across a wide pH range.[27,39,40]



Scheme 4 - General scheme showing the ligation of two peptide fragments using diselenide-selenoester ligation (DSL).

The exact reaction mechanism of the transselenoesterification step is unknown. One proposed mechanism (Scheme 5) shows the reduction of the Sec diselenide by the phenyl selenoate generated by the hydrolysis of the peptide selenoester.[41] The selenoate generated from reduction then performs a nucleophilic attack on the peptide selenoester, which then N-Se acyl shifts to form the desired product. The selenoate can then perform the initial reduction step to generate another selenoate and produce the expected diselenide final product.

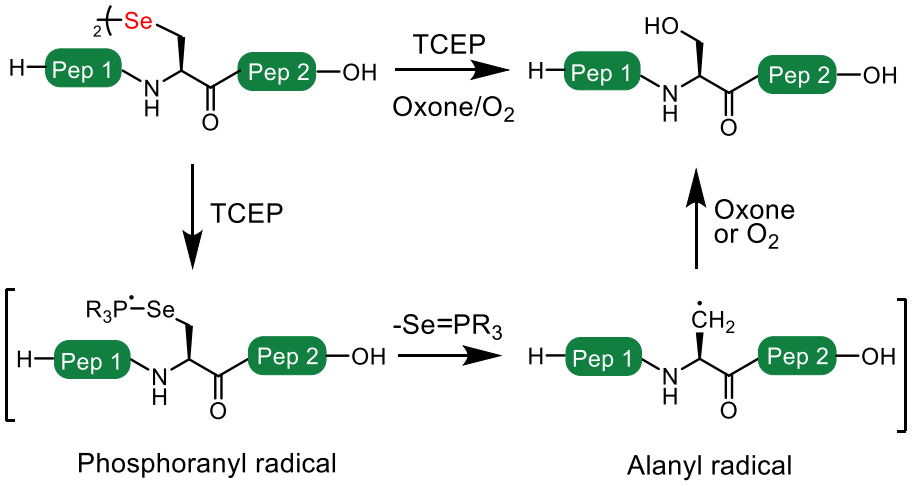


Scheme 5 - Proposed reaction mechanism for the generation of the selenoate which participates in DSL.

DSL can also be performed at high dilution, which increases operational simplicity when ligating larger peptide fragments.[41] Combining NCL and DSL has been used to successfully synthesise proteins, including the 94 amino acid *Mtb* ESAT-6 protein via a three-component one-pot method. The synthesis of small proteins via DSL was further expanded by applying the method to selenolated versions of naturally occurring amino acids. These include lysine (Lys),[42] Pro,[43] leucine (Leu) and phenylalanine (Phe).[44] Post ligation, the amino acids can be converted to a native residue via radical-induced deselenisation permitting DSL at several

native junctions. Furthermore, the use of selenolated proline enables ligation at synthetically challenging proline-proline junctions, not accessible by NCL.[40]

Sec can be converted to native Ala or Ser residue by deselenisation and oxidative deselenisation respectively.[45] These protocols are highly useful as nearly all native proteins contain either an Ala or Ser residue. The C-Se bond is weaker than the C-S bond, allowing formation of the selenyl radical without the inclusion of an initiator. When Sec is exposed to a phosphine-based reducing agent, such as TCEP, deselenisation occurs following the same radical-based mechanism as is seen with Cys to form Ala. The presence of Potassium peroxymonosulfate (oxone) during this radical deselenisation leads to hydroxylation at the β-position leading to the formation of serine, as shown in Scheme 6.[45,46] By employing Sec instead of Cys, significantly faster reaction rates are achieved, with ligation able to be performed at Ala and Ser sites.



Scheme 6 - Reaction scheme for the oxidative deselenisation of selenocysteine, converting Sec into serine.

1.3 Preparation of Proteins via Recombinant Expression

Proteins are readily available and simple to produce using recombinant expression. This technique involved cloning the gene of interest in a vector, transforming this into a host, induce and purify the resulting protein. A commonly used host organism is *E. coli,* due to ready availability and fast growth kinetics. This method has provided simple and quick access to a wide range of proteins at gram scale. There have been many purification techniques developed for proteins produced by expression. A common method involves incorporation of a poly-histidine (His) tag, followed by isolation using metal-affinity columns. The pure protein can then be eluted by disrupting the metal-His interaction.[47] Other methods involving affinity-based tags utilise the strong interaction between biotin and streptavidin.[48]

Producing proteins through recombinant expression provided the ability to alter the protein sequence chosen. This technique, called site-directed mutagenesis, allows one canonical amino acid to be exchanged for another. As certain residues, such as Cys, have a low prevalence across the proteome, selectively installing one into a protein which did not previously have a Cys residue provides a useful handle for selective labelling.[2]

Non-canonical amino acids can also be installed into protein sequences. Incorporation of Sec into proteins by ribosomes is controlled by the opal stop codon UGA. Ribosomes insert a Sec residue instead of stopping terminating protein synthesis when the Sec insertion sequence RNA motif is present.[36–38] This means that incorporation of Sec in recombinantly expressed proteins is possible when using a correctly modified tRNA synthetase. This has been shown in the expression of mammalian thioredoxin reductases using a strain of *E. coli*.[37] Antibodies containing Sec have also been expressed using codon suppression.[49] This method was applied to the expression of the anti HER2 antibody trastuzumab.[50] The amber stop codon and mutated aminoacyl tRNA synthetase enzyme have been repurposed in similar ways to incorporate a range of non-canonical amino acids into proteins.[51,52] This is a powerful technique but challenging due to the extensive protein mutation required for each different amino acid. This multi-step requirement shows why modification of native amino acids is a simpler approach.

Proteins produced by recombinant expression are folded to their native form. The harsh conditions associated with many bioconjugation methods result in protein becoming denatured and requiring refolding. This is not ideal and can often be time consuming. Bioconjugation onto proteins should ideally involve the use of mild conditions, including ambient temperature, aqueous solutions and neutral pH.[2] As many methods fail to meet one or more of these requirements, it shows a clear requirement for new techniques to be developed.

1.4 Site-Selective Peptide and Protein Modification

The site-selective modification of peptides and proteins is an on-going challenge. There have been many methods developed to functionalise proteins; however, these methods usually have significant drawbacks. A commonly targeted residue, highly abundant (5.9%) across the proteome, is Lys.[53] Other reasons for targeting Lys include the facile chemistry/protocols that can be used, the stable products that are formed and the high probability of a protein having a solvent accessible Lys. Modification has been achieved through standard amidation using pre-activated esters of carboxylic acids, such as N-hydroxysuccinimide (NHS) esters. The issue with this method tends to be selectivity. Depending on pH, the NHS esters can react with a range of functional groups, including thiols, imidazole, indole and alcohols. This can lead to multiple sites of modification, resulting in a heterogeneous protein mixture.[54]

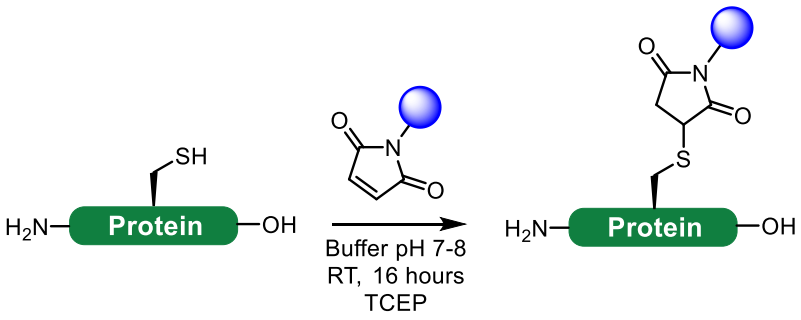
In recent years, reactions that target other canonical residues have been reported with growing prevalence in the literature, with selective modification methods being developed. Chemistry targeting the residues, tryptophan (Trp),[55–57] methionine (Met),[58] His,[59] arginine (Arg),[60,61] serine (Ser),[62,63] as well as the N and C-terminus of proteins, has been reported.[53,64,65] Many of these impressive developments still face issues with selectivity, harsh conditions and ease of operation. Such challenges have led researchers to focus their efforts on Cys. Cys is by far the most commonly targeted residue in the literature, with over 200 currently established methods developed.[62,66]

1.5 Cysteine-Selective Peptide and Protein Modification

The use of Cys as a chemical handle is common due to the thiol group's highly nucleophilic nature. The advantage of modification at Cys is that this residue is one of the least abundant amino acids in the proteome (2%). Therefore, fewer Cys residues are likely to be present in a target sequence, thus site-selective modification is more likely. There is also a high chance that Cys residues will be absent in the protein of interest. This allows selective introduction of a single Cys residue, either chemically or through site-specific mutagenesis during protein expression. The presence of one Cys residue combined with residue-specific chemistry ensures site-selective introduction of the group of interest. This, along with the unique chemistry of Cys, make it an attractive target for selective modification.

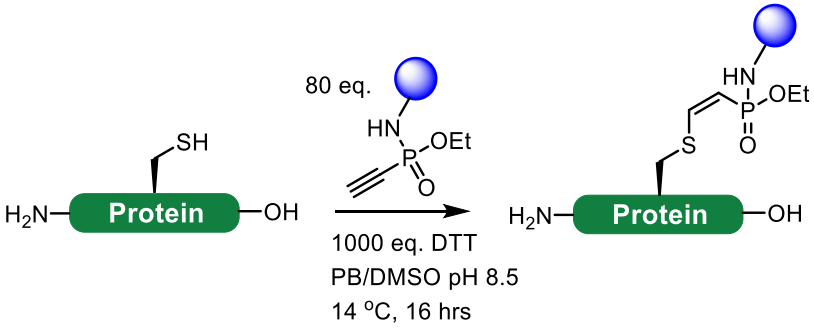
1.5.1 Michael Acceptors

Michael acceptors are alkenes positioned next to electron withdrawing groups. Maleimides are commonly utilised, as shown in Scheme 7. Alternative Michael acceptors include acrylates and vinyl sulfones.[2] The reaction proceeds via nucleophilic attack of the thiol into the alkene, followed by quenching of the resulting carbanion. Examples of Cys being modified using this method date back to 1955 when a bis-maleimide was used to cross-link bovine plasma albumin and wool keratin.[67]



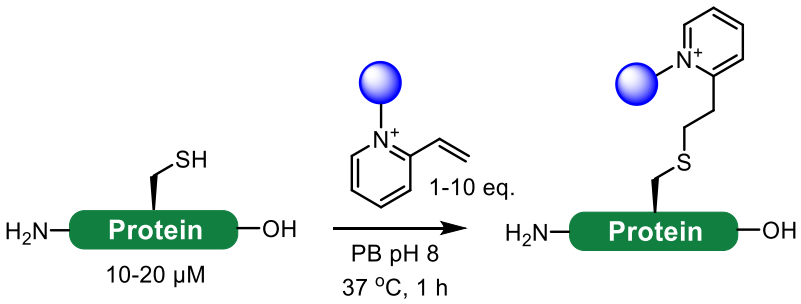
Scheme 7 - Generic modification using maleimides in buffered (phosphate, tris, etc.) solution at room temperature. A reducing agent, such as TCEP, is often included to ensure Cys remains in the reduced form.

A common issue with maleimides is the potential for non-specific reactions with other nucleophilic residues, such as Lys. A further issue is the reversible nature of the reaction, with the retro-Michael addition reaction resulting in removal of the conjugate from the protein. The desired product can also undergo hydrolysis, resulting in ring-opening of the maleimide to produce two isoforms of the formed succinamic acid thioether.[68] An alternative Michael acceptor that reported no issue with selectivity are the phosphonamidates, shown in Scheme 8.[68] These were reported to react only with Cys residues even when 10 equivalents of the reagent per Cys residue was used. They enabled the modification of trastuzumab with a Cy5 dye, generating an antibody-fluorophore conjugate.[69]



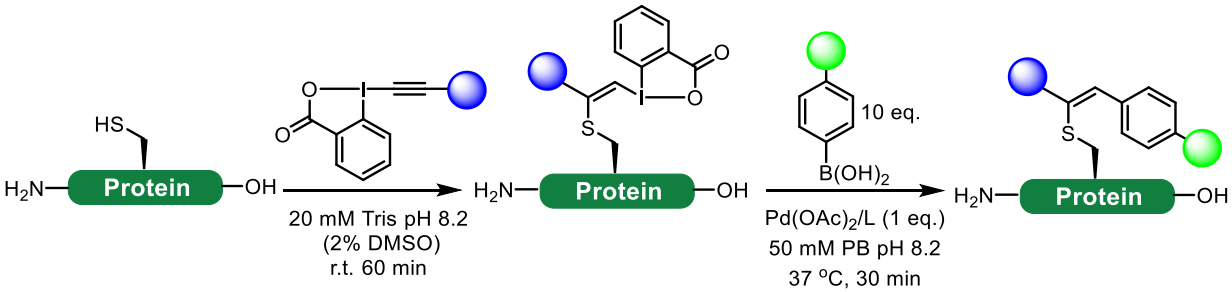
Scheme 8 - Cys conjugation with phosphonamidate using a high excess (1000 eq.) of 1,4-dithiothreitol (DTT) in aqueous conditions with an organic co-solvent (dimethyl sulfoxide (DMSO)) for 16 hours.

Another Michael addition method to selectively modify Cys residues involves the use of quaternary vinyl pyridines, shown in Scheme 9.[70] This method has been applied to proteins and antibodies at low concentration in aqueous conditions to afford quantitative conversion to modified product using 1-10 equivalents of reagent in under 1 hour.



Scheme 9 - Modification of Cys residue with quaternary vinyl pyridines at 37 °C for 1 hour.

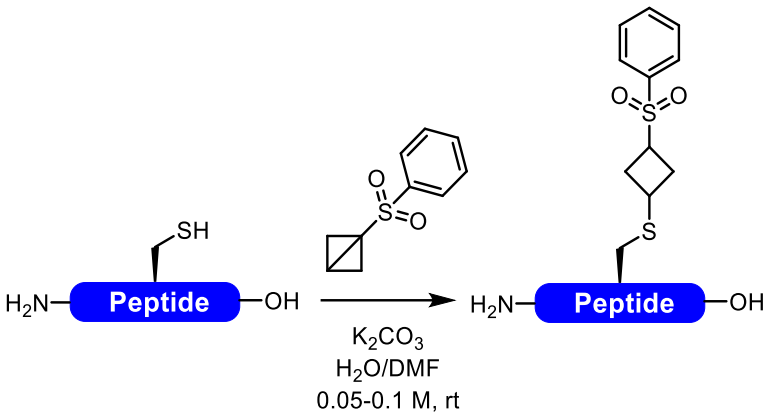
Recently, a method has been developed which utilises hypervalent iodine compounds to enable the introduction of doubly orthogonal reagents. These reagents, ethynylbenziodoxolones (EBXs), consist of an alkyne for Michael addition of the thiol. The hypervalent iodine bond is retained after the initial addition, allowing further cross coupling. The initial addition is performed at room temperature, open to air for 5 to 60 minutes. The reaction is tolerant to a range of groups on the alkyne, including azides, alkynes, alkyl chlorides and cyclopropyl rings. The reaction has been performed successfully in excellent yield on both peptides and proteins. The hypervalent iodine group takes part in Suzuki-Miyaura Pd assisted cross-coupling with aryl boronic acids, allowing the introduction of a range of aromatic groups. To show dual orthogonality, the initial reagent contained an azide, which was reacted with an alkyne using strain-promoted azide alkyne click chemistry, followed by Suzuki-Miyaura cross coupling in a one-pot method on a protein model, shown in Scheme 10.



Scheme 10 - Site-selective modification using EBX reagents for dual orthogonal protein labelling.

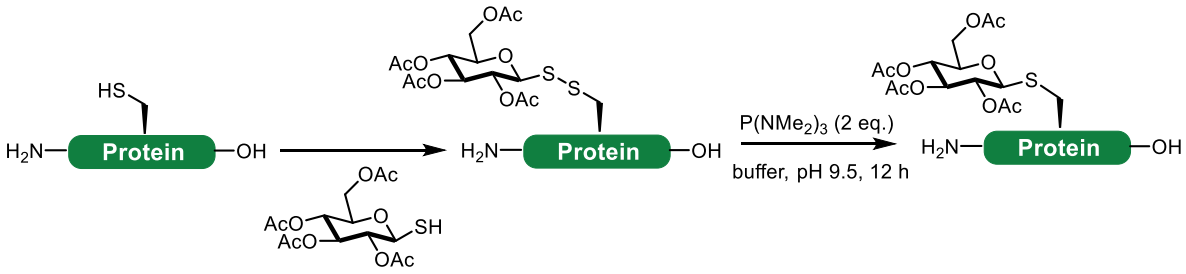
1.5.2 Alternative methods for thioether synthesis

An alternative to the thiol-ene click reaction for generating thioethers is the use of strained 'bicyclo' molecules, shown in Scheme 11. These strained structures undergo nucleophilic attack from the thiol group, opening one of the rings to form a new S-C bond. The reaction requires DMF as a co-solvent to solubilise the strain-release agent. The basic conditions (e.g. pH 9) enhance the nucleophilicity of the thiol. The reaction is selective in the presence of other nucleophilic residues, such as Lys.



Scheme 11 - Strain-release modification of cysteine using bicyclic compounds.

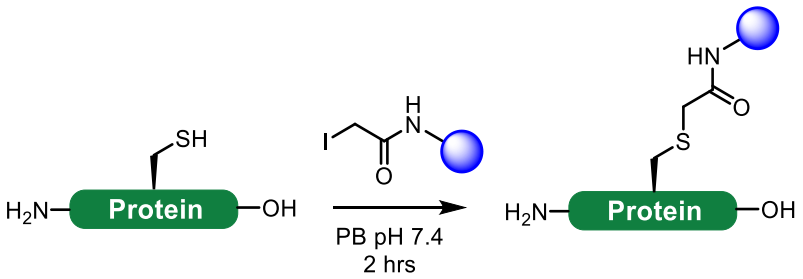
Peptides and proteins have also been modified via disulfides. The thiol of Cys will readily oxidise with another thiol to form a disulfide bridge. The issue with this form of modification is the instability of the disulfide bond, meaning this method cannot be used for permanent modifications. A method to convert the disulfide bond into a more stable thioether has been developed.[71] In the referenced work by Davis and co-workers, a protein modified with a glycosyl thiol is treated with hexamethylphosphorous triamide (HMPT). This reduces the disulfide and forms a phosphonium ion that undergoes β-elimination to form dehydroalanine (Dha). The previously released glycosyl thiolate then nucleophilically attacks the newly formed alkene, forming a thioether, as shown in Scheme 12.



Scheme 12 - Example reaction for the glycosylation of a protein followed by conversion to the stable thioether by reductive elimination followed by Michael addition.

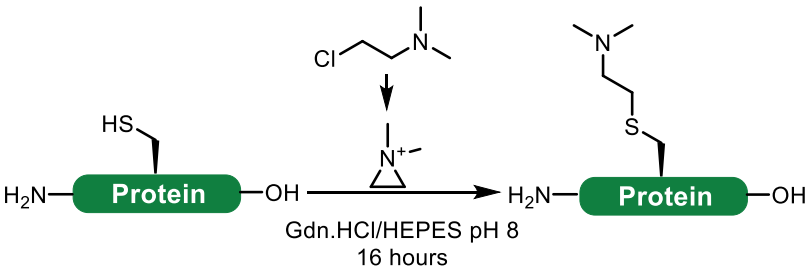
1.5.3 Alkyl halides – α-halocarbonyls, aminoethylation (alkyl chloride)

The use of α-halocarbonyls for protein modification can be dated as early as 1935[72] where iodoacetamide was used to modify and study the Cys residues of keratin, as shown in Scheme 13.



Scheme 13 - General scheme for the modification of Cys residues with iodoacetamide.

Aminoethylation has proven to be a useful tool for generating Lys mimetics; a range of known Lys PTMs can be installed, site-selectively, using this approach. The reported method, as shown in Scheme 14, used the alkyl chloride of the desired aminoethyl group.[73] This molecule was treated with base, forming the aziridinium intermediate. The cyclic intermediate undergoes nucleophilic attack from the thiol, forming the aminoethylated product.

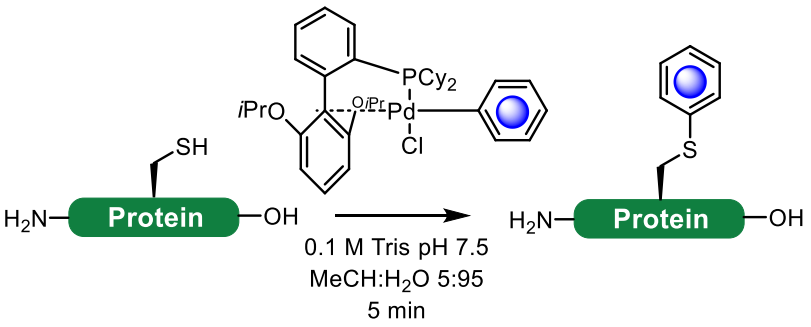


Scheme 14 - Reaction scheme showing the conversion of Cys to a dimethyl lysine mimetic.

1.5.4 Transition Metal Mediated Arylation

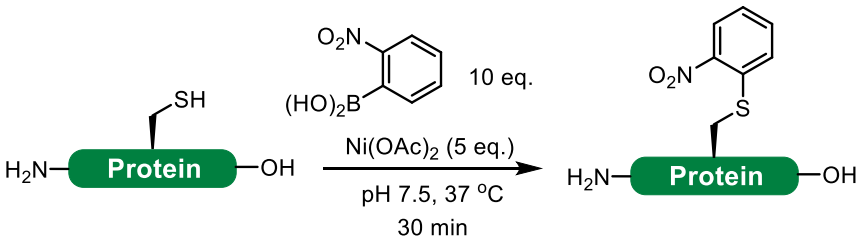
Another area of increasing attention is the modification of peptides and proteins through transition-metal-mediated arylation. Such reactions proceed under mild conditions and the resulting conjugates are stable to a range of conditions; thus, these methods are applicable to the preparation of protein drug conjugates.

A versatile and rapid method involving the use of organometallic palladium reagents was developed in 2015, shown in Scheme 15.[74] The reaction was successfully applied to the arylation of peptides, proteins, and antibodies. [74] The reaction was successfully applied to the arylation of peptides, proteins and antibodies. Peptides were quantitatively modified in 5 minutes using 2-6 equivalents of the palladium reagents in aqueous conditions. The low peptide concentration in these reactions, 10 µM, make transferring the method from peptides to proteins and antibodies simple. The antibody trastuzumab was modified with the drug molecule Vandetanib in 30 minutes with a drug to antibody ratio of 4.4:1. This method was also used to arylate two Cys residues of a peptide, forming a stapled peptide.



Scheme 15 - General reaction scheme for the arylation of Cys residues using organometallic palladium reagents.

Another cross-coupling example was developed in early 2019.[75] This method involves the use of a simple Ni(II) salt, $Ni(OAc)_2$, at high dilution (0.2 mM) in aqueous conditions (Scheme 16). The reaction proceeds at 37 ^oC and is complete within 30 minutes. This method has been used on a range of peptides containing multiple different residues. It can also tolerate the presence of TCEP (4 eq.), which is useful as it allows the functionalisation of oxidised Cys residues. The advantage of this method using a Ni(II) catalyst is that it is unaffected by the presence of water and oxygen, useful for protein modification. The issue with this reaction compared with the previously discussed cross coupling method is the substrate scope. The Ni(II)-catalysed technique only works with 2-nitrophenylboronic acids, which is a significant limitation. The reaction is tolerant to a range of groups in the para position, which does widen the potential scope of this method. The reaction does not require a ligand for peptides, but for protein models a 6,6'-dimethyl-2,2'-dipyridyl ligand in 1:1 equivalence with the Ni catalyst is required. The protein bovine serum albumin (BSA) was successfully modified with 2-nitrophenylboronic acid modified with an alkyne group at the para position to a conversion of >90% within 10 minutes.[76]



Scheme 16 - Ni(II) catalysed cross coupling between Cys and 2-nitrophenylboronic acid.

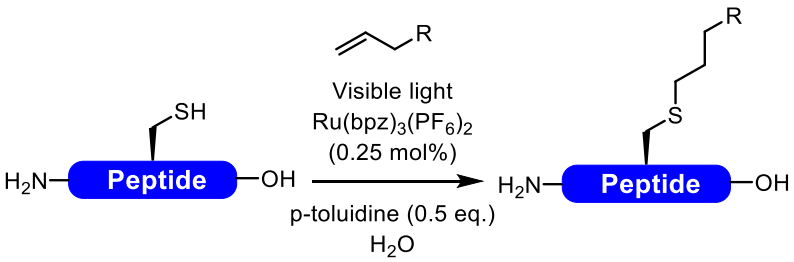
Another reported method uses preformed gold complexes for selective arylation.[77] The solvent system is aqueous, with a pH range of 0.5-14, 3 equivalents of the gold complex and reaction times of less than 5 minutes. This method has been used to install biologically useful moieties, including fluorescent dyes, affinity tags and drug molecules. It has been applied to peptide and protein models. Protein models require higher equivalents of the gold complex (20 eq.)

and longer reaction times (30 minutes). The arylation reaction scheme is similar to that shown previously for the Pd mediated arylation in Scheme 15.

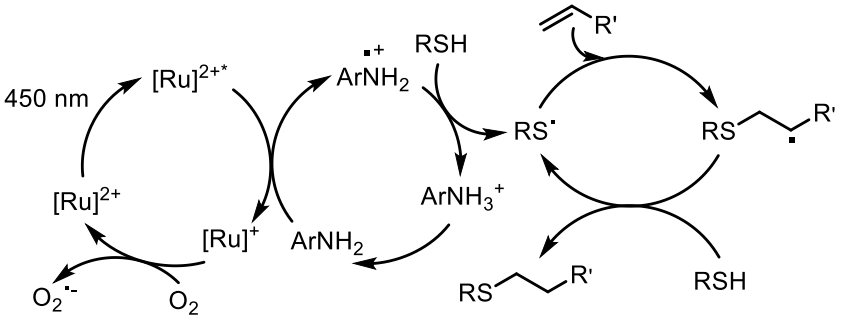
1.5.5 Photoredox Catalysed Methods

Photoredox catalysis has received increasing attention in recent years as a more sustainable method of performing useful chemical transformations. The catalysts used in these reactions become excited after absorbing visible light, allowing them to perform single-electron transfer (SET). This has been shown to be particularly useful with Cys, transferring an electron onto the thiol group to form a radical cation. The radical cation is readily deprotonated, forming a thiyl radical. This is a very mild method to form the thiyl radical when compared to other methods which use excesses of radical initiator (VA-044) or UV light (254 nm). These mild conditions therefore allow the formation of the thiyl without disrupting the protein's tertiary structure.[78]

Techniques that utilise photocatalysis to modified Cys are increasing in prominence. An example is the radical thiol-ene reaction. An example, shown in Scheme 17, uses the ruthenium based catalyst $Ru(bpz)_3(PF_6)_2$ at 0.25 mol%, with p-toluidine used as a redox mediator. This is a very mild reaction, performed in aqueous conditions, allowing addition of unactivated alkenes such as 6-azido-1-hexene in yields >70%.[79] The mechanism proposed for this reaction is shown in Scheme 18. It shows the initial excitation of the ruthenium catalyst by blue light. The excited catalyst performs a single electron transfer, removing an electron from the aniline additive to form a radical cation. This is then removes an H atom from the thiol, forming the thiyl radical, which reacts with the alkene forming a thioether bond. They then propose the resulting carbon centred radical is quenched by abstracting an H atom from another thiol. The proposed mechanism also shows the use of oxygen to regenerate the photocatalyst. This adds a degree of operational simplicity to the reactions as degassing a reaction mixture can be time consuming and very difficult when using small reaction volumes, typical when performing peptide modifications. A more recent development uses an organic-based photocatalyst, allowing metal-free thiol-ene protein conjugation.[47]

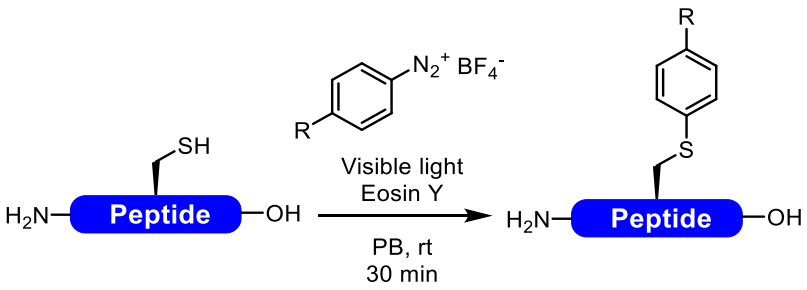


Scheme 17 -Photocatalysed radical thiol-ene modification of Cys.



Scheme 18 - Proposed mechanism for the photocatalysed thiol-ene reaction.

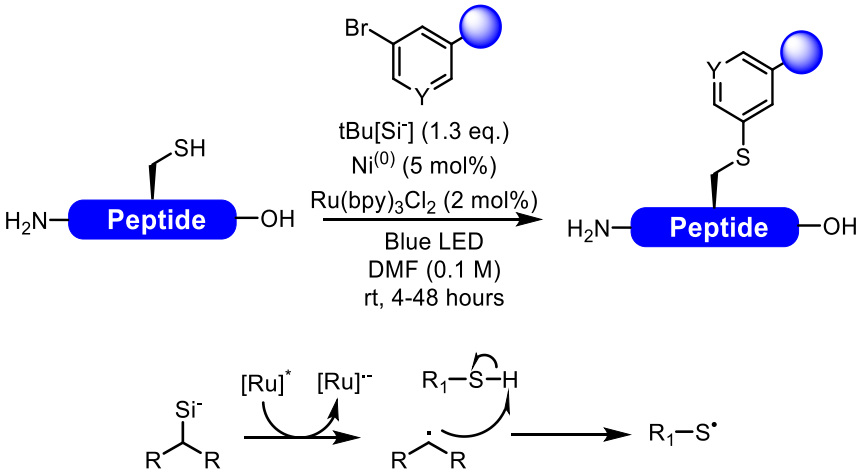
Aside from thiol-ene chemistry, photoredox catalysis has also been used to enable Cys-selective arylations. The example shown in Scheme 19 uses eosin Y (EY) to form the thiyl radical. It was proposed that eosin Y catalyses the generation of the thiyl radical following the same mechanism as shown in Scheme 18. The thiyl radical displaces a diazonium salt, either pre-formed or formed *in situ.* The reaction can be performed on peptides and is tolerant to multiple functional groups. Aqueous conditions can be used, and the reaction can be performed in air. The reaction can be performed in batch, taking 30 minutes, or in flow, taking as little as 30 seconds. Conditions such as these are very useful to chemical biologists as they often do not have the facilities available for inert atmospheres or completely dry reaction medium.[76] A recent publication shows the use of an organic photoredox catalyst, allowing a metal-free thiol-ene reaction.[80]



Scheme 19 - Photocatalytic S-arylation of Cys containing peptides using diazonium salts of arenes.

Another example of photoredox-catalysed arylation uses a combination of ammonium bis(catechol) silicate (*t*Bu[Si$^-$]) and a Ru catalyst to form the thiyl radical (Scheme 20). The Ru catalyst is excited by blue LEDs, this removes an electron from the silicate to generate an alkyl radical. The alkyl radical then abstracts an H-atom from the thiol to form the thiyl radical. The thiyl radical then donates into a Ni(0) catalyst, which cross couples with an aryl bromide to form the desired product. The reaction has been used to install a range of functionalised aryl bromides, including drug molecules such as benzodiazepine. Yields are generally high, with complete conversion observed on small peptide models. The main issues with this method

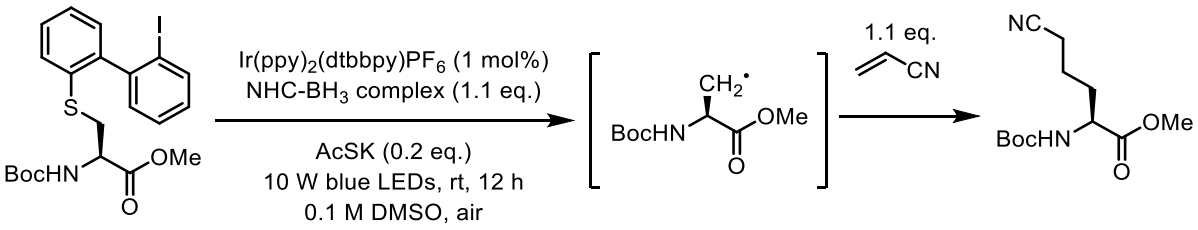
are that the reaction cannot tolerate more than 20 equivalents of water, and high dilution (< 10 mM) leads to a significant decrease in conversion.[76]



Scheme 20 - General scheme for Ru/Ni catalysed cross coupling (Y = C or N). The mechanism shows how the thiyl radical is formed using the alkyl radical generated by the single electron transfer of an excited ruthenium catalyst to a silicate.

The highlighted reactions show the potential for photocatalysis to be used for Cys-selective bioconjugation. Photocatalysis has significant advantages over traditional chemistry, including sustainability, mild reaction conditions, and high selectivity. This is an area of protein bioconjugation that requires further exploration and reaction development.

A recent publication (Oct 2020) showed the development of a technique that utilised the alanyl radical generated from a modified Cys residue.[81] This radical donates into mostly activated alkenes to form the corresponding alkylated product (Scheme 21). This was mostly applied to single, fully protected amino acids or fully protected tripeptides. They successfully synthesised a very wide range of unnatural and native amino acid residues, with high enantiomeric excess values. The mechanism proposed involves a single electron transfer from the thiolate to the excited iridium, forming a thiyl radical. The thiyl radical then abstracts an H-atom from the borane complex, forming a C-BH_2 radical which removed the iodo from the aromatic ring, forming an aryl radical. This undergoes an intramolecular radical substitution reaction with the sulfur atom, eliminating the dibenzothiophene and generating the alanyl radical. The alanyl radical then donates into the activated alkenes, forming the alkylated product. They then proposed that the secondary radical generated from this radical addition reaction is reduced to the corresponding carbanion by the iridium catalyst in the +2 oxidation state. This then regenerates the Ir(III) catalyst required for further reactions.

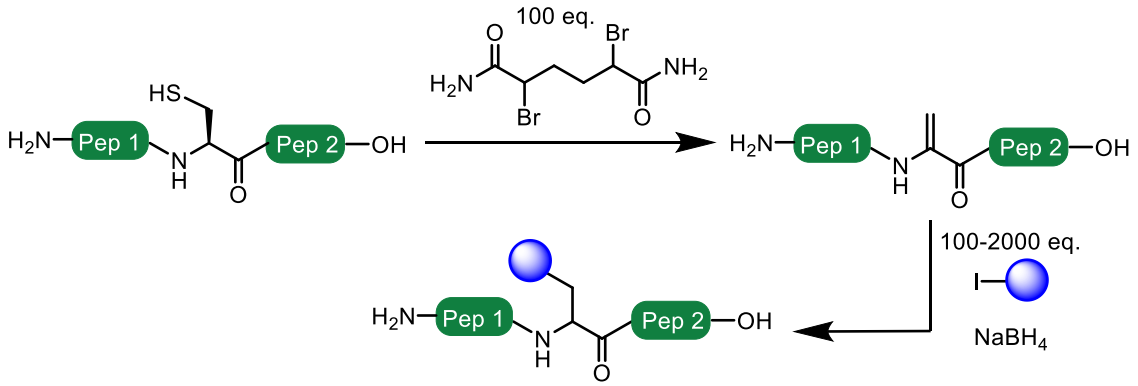


Scheme 21 - General scheme for the desulfurative alkylation method developed by Yingwei Wang et al. [81]

1.5.6 Oxidative desulfurisation to dehydroalanine

One of the main concerns regarding the use of Cys as a chemical handle to install biologically relevant groups is that the resulting C-S bond can be unstable. This instability leads to the modification being lost, resulting in mitigation of activity. To overcome this issue, methods have been developed which allow Cys to be converted to Dha, achieved via reductive elimination. A range of methods have been developed to allow access to Dha, a common example being dialkylation of the thiol, using a dibromo reagent such as 2,5-dibromohexanediamide, which eliminates to form Dha. [82–84]

There have been a range of methods developed over recent years to install functional groups onto Dha. A powerful method involves the generation of alkyl radicals from alkyl halides using sodium borohydride. The alkyl radicals selectively donate into the double bond of Dha, forming a carbon-carbon bond. This method is high yielding and has been used to install a wide range of groups, including modified lysine and arginine residues, glycans and phosphorylations.[85] The general reaction scheme for this method is shown in Scheme 22.



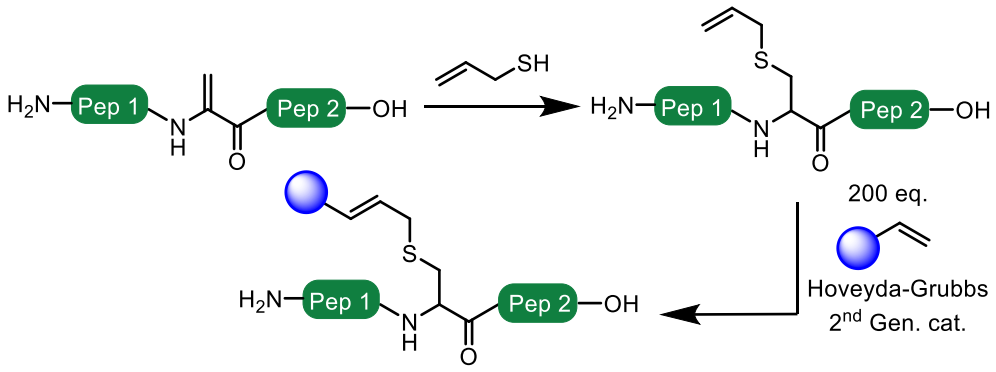
Scheme 22 - General scheme for the conversion of Cys to Dha, followed by radical addition into the double bond.

A similar method, also published in 2016, formed Dha from phosphoserine.[86] The phosphoserine containing protein was treated with base to eliminate the phospho group, producing Dha. The alkyl radicals used for addition into the double bond were generated from alkyl iodides using Zn powder and $Cu(OAc)_2$.

A recent advance to the modification to Dha has been made by the Davis group.[87] Instead of using $NaBH_4$ to generate alkyl radicals from alkyl halides, they use a ruthenium-based photoredox catalyst (either in excess or sub-stoichiometric) to generate alkyl radicals from either alkyl boronic acids or pyridyl-sulfone-fluoride reagents. These then react with Dha in the same manner. This method was again applicable to proteins, allowing installation of native PTMs, including acetyl and benzyl lysine. They also used this method to install an alkyl halide onto a protein. This was used as a complementary handle to undergo nucleophilic substitution reactions with a range of groups including proteins, phosphines, sodium azide and thiols.

These methods are undeniably powerful, providing access to a wide range of modified proteins linked to their specific modification via a stable C-C bond. However, the reaction results in a loss of the stereochemical information at the target residue due to the formation of Dha, which can affect the tertiary structure and therefore its natural activity.

As well as installing a range of groups generated from alkyl halides, this method has also been used to install cysteine mimetics. The alkene readily undergoes Michael addition with a thiol to reform cysteine, albeit with a loss of stereochemistry. This has been exploited on a protein model to produce allyl cysteine. This was further modified using cross-metathesis to install a range of groups, including sugar and PEG molecules, into a protein. The general scheme for this is shown Scheme 23.



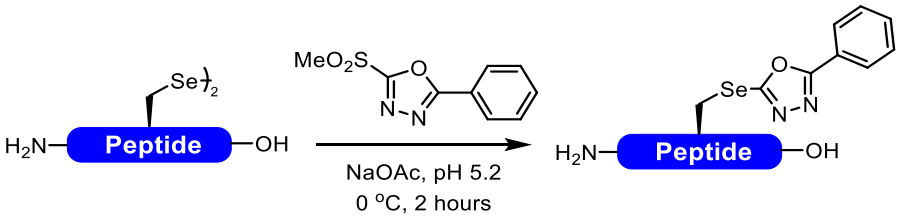
Scheme 23 - General scheme for the formation of allyl Cys and cross-metathesis for selective protein modification.

While there are clear advantages for choosing Cys as a chemical handle, resulting in the numerous examples discussed here, the lability of the C-S bond linking the small molecule to the protein can result in loss of the modification under certain conditions.[77] The most common method of bioconjugation uses the versatile maleimide group. The issue of reversibility is common with maleimides, resulting in the maleimide and free Cys being regenerated.[68,88] This shows the need for novel methods to be developed which match the efficiency of the maleimide conjugation, whilst producing a stable product.

1.6 Selenocystine-Selective Conjugation Methods

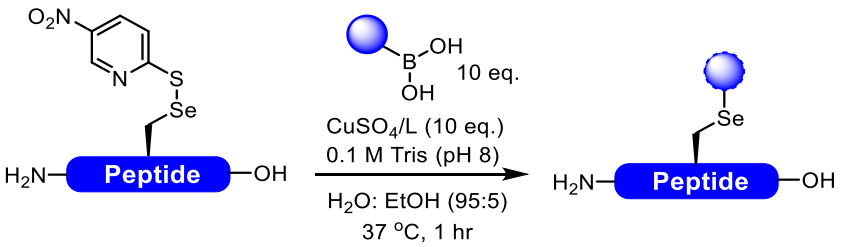
The methods discussed thus far have been sequence specific, relying on the presence or absence of specific residues. They have allowed the introduction of a range of functional groups, often at the cost of expensive metal catalysts, harsh conditions or loss of chirality. Using Sec as the modification handle could provide a universal modification method and offer site-specific modification whilst maintaining the stereochemistry of the protein backbone and the protein's activity.

The use of Sec for ligation and subsequent site-selective modification has yet to be investigated in depth. The most common methods used to modify Sec mimic those of Cys. These include exploiting the enhanced nucleophilicity of Sec to alkylate this residue with a range of useful molecules. An example of this is shown in Scheme 24. This uses alkyl sulfones to install phenyloxadiazolyl.[89]



Scheme 24 - Alkylation of Sec with phenyloxadiazolyl sulfone.

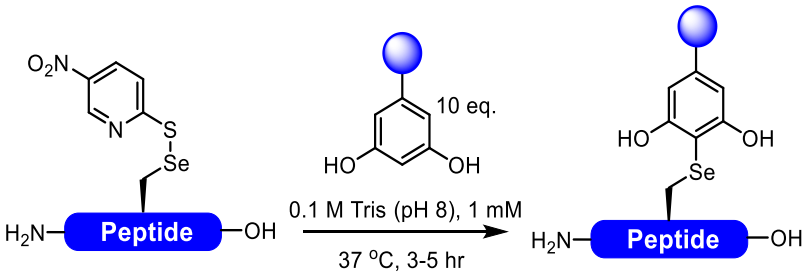
Other uses of Sec include chemoselective arylation using an umpolung approach. The arylation is copper catalysed, using aryl boronic acids, as shown in Scheme 25. The method enables the attachment of a range of aryl groups to proteins that contain Sec, including coumarin and L-phenylalanine. The method is also compatible with a range of *ortho, meta* and *para* substituted aromatics, bearing both electron withdrawing and donating substituents.[90]



Scheme 25 - Copper catalysed arylation of Sec.[90]

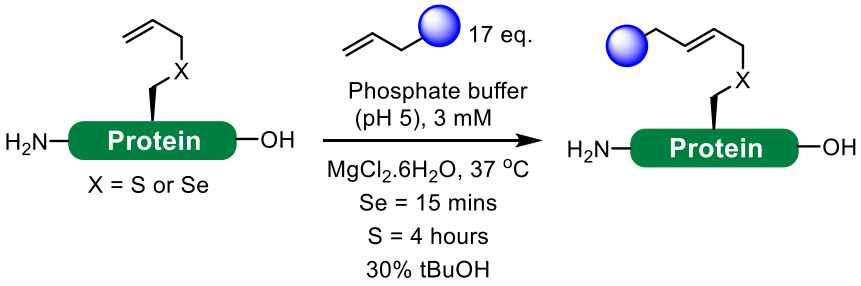
This approach has been further developed to allow the arylation of Sec without the need of a metal catalyst. The example in Scheme 26 shows the arylation of Sec with the antibiotic vancomycin occurring in at pH 8 at 37 °C. This has been performed on a range of peptides, with yields varying from 40 to 90%. The researchers installed a variety of activated aromatics, allowing access to peptides with a range of useful functional groups. They also installed

interesting natural products and pharmaceuticals with a yield >60% including fenoterol, resveratrol and phlorizin. The ease of arylation in metal-free conditions and high yields shows how this is a very useful method for installing biologically relevant modifications.[91]



Scheme 26 - Arylation of Sec without the requirement of a metal catalyst.[91]

Another approach explored involves the utilisation of allyl Sec as a handle for cross-metathesis, using the same methodology as shown in Scheme 23. The seleno-allyl group demonstrates enhanced reactivity when compared to thio-allyl. This is illustrated by the rapid reaction rate with Sec allyl (15 minutes), compared to 4 hours for Cys allyl. This method has allowed protein glycosylation with a yield of 75%, while the thiol equivalent did not produce desired product. The conjugate can be cleaved when using the seleno form by addition of aqueous H_2O_2, forming Dha. This has allowed a "write-read-erase" cycle using an acetylated lysine mimic. This method has provided chemical access to rapid, reversible protein modifications.[92]



Scheme 27 - Modification of Sec-allyl by cross-metathesis to install biologically relevant molecules.[92]

1.7 Peptide Macrocyclisation

Cyclic peptides have received increasing attention for their use as pharmaceuticals. The constrained secondary structure of these sequences, covalently locked into place via a 'staple', affords several advantages over linear peptides. If this orientation has been optimised to maximise the interaction of the cyclic peptide with the target protein, then this restricted movement will afford a high binding affinity. The use of peptides for this design is advantageous as a wide variety of side-chain functionality can be introduced synthetically in an attempt to increase this binding. These properties allow cyclic peptides to be used to

effectively interrupt protein-protein interaction (PPI), a complex network of interactions across the proteome which mediate the majority of biochemical interactions. It has been proven that PPIs play an important role in a wide range of diseases including various cancers and those caused by bacteria and viruses.[93] This can either result from the disruption to native PPIs, or the formation of new PPIs between human and bacterial/viral proteins. The interactions between proteins involves surface-to-surface interactions rather than an active site 'pocket', making them difficult to target using standard small molecules drugs. The perception of these interactions as "undruggable" has led to the development of alternative approaches, including cyclic peptides.[94]

Cyclic peptides have the same associated advantages of linear peptides over small molecules, including higher specificity and lower toxicity. While linear peptides suffer from a high rate of degradation in the blood due to enzymatic activity, the act of cyclising the peptide enhances the stability to proteases in the blood as the typical routes of degradation, from the N or C terminus, cannot be performed.[95–98] This approach affords the peptides a longer serum lifetime. The cyclised forms also have a larger surface area, allowing binding to large areas of shallow protein surfaces. The cyclised forms also tend to have a higher cell-membrane permeability, allowing them to be used for both extracellular and intracellular targets.[99] Thus cyclic peptides are able to bridge the gap between the advantageous physicochemical properties displayed by small molecules, and the exceptional activity, specificity, and bioavailability inherent to protein-based biologics.

Cyclic peptides used in pharmaceutical are relatively short (5 to 14 amino acids) allowing simple synthesis using SPPS.[99] The use of SPPS also allows the introduction of unnatural amino acids, which often enables the development of more target-specific peptides, producing a drug with a higher activity at a lower concentration.[100,101] There is currently a wide range of cyclic peptides approved for the treatment of diseases. These include macrocyclic peptides derived from nature, such as cyclosporine A, used to prevent organ rejection,[102] as well as synthetic examples such as the cyclic peptide Murepavadin (Figure 4), used as a pseudomonas specific antibiotic.[102]

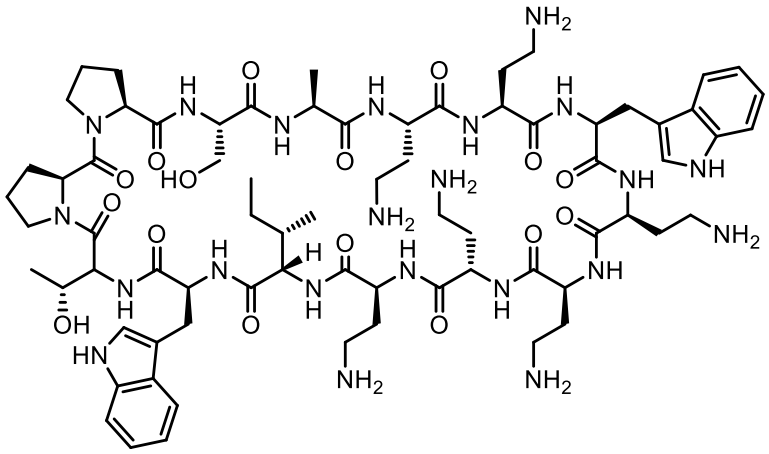
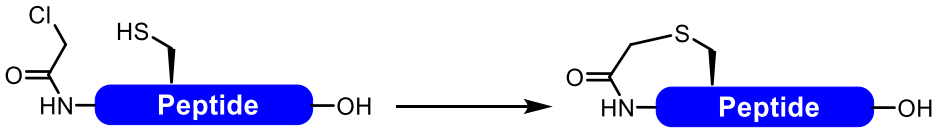


Figure 4 - Structure of the cyclic peptide Murepavadin.

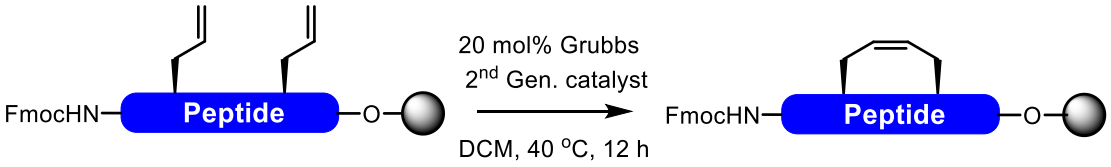
There are a range of methods available to generate libraries of cyclic peptides for screening against specific PPIs. For example, phage display has been utilised to express libraries of peptides which can be cyclised via conjugation to a core scaffold.[103] Another example is the split-intein circular ligation of peptides and proteins (SICLOPPS) technique. This method utilises N and C-terminal intein fragments to generate a cyclic target peptide. The target peptide can contain any canonical amino acid and be any length, providing access to peptide with a wide range of functionality and ring sizes.[104] Another technique is referred to as random nonstandard peptide integrated discovery (RaPID).[105] This method utilises flexible *in vitro* translation (FIT). The technique is facilitated by RNA enzymes (flexizymes) that catalyse tRNA acylation with the target non-canonical amino acids. This integration of FIT with mRNA display has been used to develop libraries of macrocyclic peptides against both extracellular and intracellular target proteins. The FIT method has also been used independently to develop cyclic peptide libraries.[106] This method relied on the peptides having a Cys residue and an N-terminal chloroacetamide group. The free thiol spontaneously reacts with the chloroacetamide group to form the desired cyclic peptides (Scheme 28). Although the incorporation of non-native residues is possible using the RaPID/FIT technique, synthetic methods that enable the cyclisation of engineered peptides are still in demand.



Scheme 28 - General scheme for the spontaneous cyclisation of peptides baring an internal Cys residue and a terminal N-chloroacetamide group.

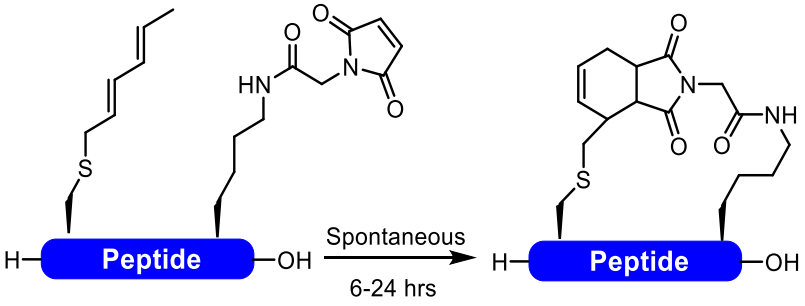
A wide range of methods have been employed for the synthesis of cyclised peptides; these include click chemistry with azides, ring-closing metathesis (RCM, Scheme 29) with installed alkenes, and thiol-ene chemistry.[95] These techniques require the installation of unnatural

amino acids, bearing alkynes/azides or alkenes, and are commonly performed on-resin, reducing the number of steps required to produce cyclic peptides.[107–109]



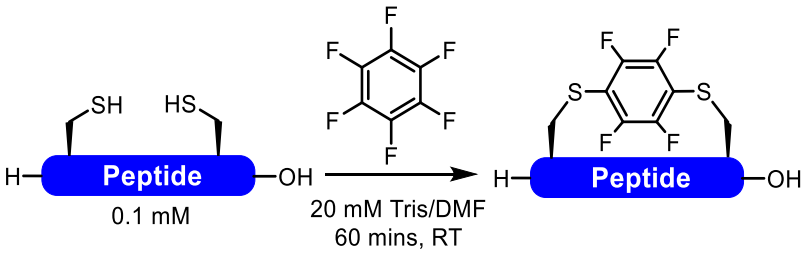
Scheme 29 - An example of on-resin macrocyclisation using RCM.

An interesting method that employs Diels-Alder cycloaddition chemistry has recently been developed.[98] This involves the synthesis and incorporation during SPPS of two unnatural amino acids. It utilises orthogonal protecting groups, allowing the installation of these unnatural groups on-resin. This gives the options of either synthesising the amino acid for installation, or just installing the actual group for installation onto the amino acid on resin. A hexadiene group installed on the Cys undergoes spontaneous cycloaddition with the maleimide installed onto a Lys residue, forming a cyclic peptide (Scheme 30).



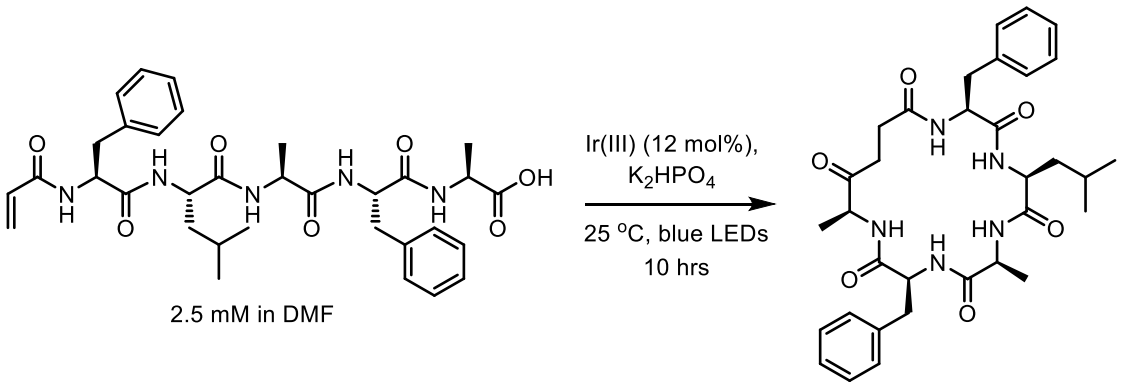
Scheme 30 - General scheme for peptide macrocyclisation using modified Cys and Lys residues for diels-alder cyclisation.

Macrocyclisation techniques that can be performed on fully unprotected peptides using only canonical amino acids are relatively uncommon. An example of this uses the previously mentioned NCL chemistry. Having the N-terminal Cys residue and activating group (hydrazide) on the same molecule, head-to-tail cyclisation occurs.[110] This same paper also describes a method of forming macrocycles on unprotected peptide that contained a single Cys and Lys residue, with the N-terminus acetylated. This method uses *ortho*-phthalaldehyde. The amine of the lysine reacts through reductive amination forming an imine, the thiol adds into the imine and consequently the amine attacks the other aldehyde. This then rearranges, resulting in the elimination of the alcohol, forming a stable, cyclised peptide. As with bioconjugations, Cys is a commonly targeted residue for cyclisations due to its unique reactivity. A method developed in 2014 utilised the specificity with which Cys reacts with perfluorinated aromatics to selectively bridge two Cys residues (Scheme 31).[111]



Scheme 31 - Peptide macrocyclisation using perfluoroaryl-cysteine S_NAr chemistry.

There are relatively few examples of peptide macrocyclisation utilising visible light photoredox catalysis. A recent review of this area, by L. Raynal and co-workers, highlights the current extent of these methods developed.[112] The majority of methods that use Cys utilise UV light (365 nm) and UV-activated catalysts to perform an intramolecular thiol-ene reaction on unactivated alkenes and alkynes. Aside from Cys, an interesting method was developed by the Macmillan group in 2017 (Scheme 32).[113] They used a fully side-chain protected peptide baring an N-terminal acroyl group and a glycine residue with a free carboxylic acid on the C-terminus. An iridium-based catalyst was used to afford C-terminal decarboxylation, generating a primary alkyl radical that donated into the activated alkene, forming the cyclised peptide. A variety of ring sizes were formed, ranging from 11 to 47 with isolated yields of 22 to 41%.



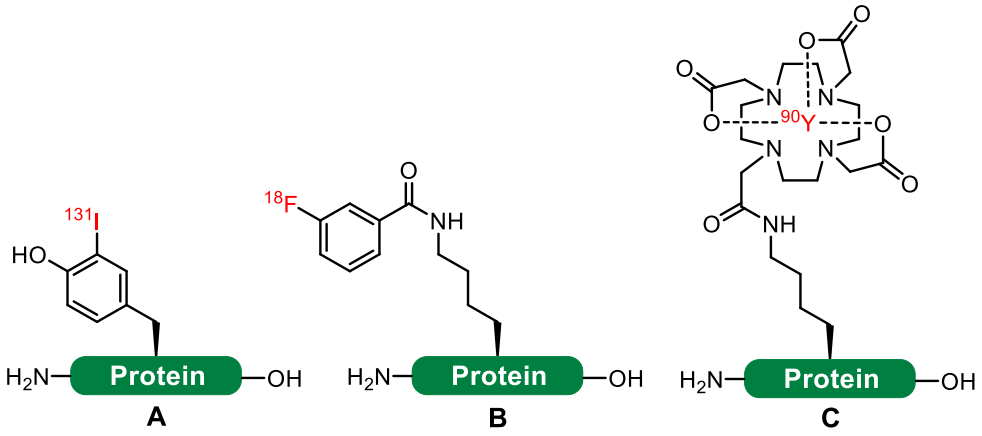
Scheme 32 - Reaction scheme for the macrocyclisation of peptide via photoredox-catalysed C-terminal decarboxylation.

The relatively few macrocyclisation methods that utilise photoredox catalysis highlights a possible area for expansion. Another issue with methods currently available is the use or formation of bulky groups to form the macrocycle. Such moieties have the potential to significantly distort the peptide's structure, which could in turn reduce the activity of the cyclic peptide. Many of these reactions also require the peptide to be fully protected; a method that utilises photoredox catalysis on fully unprotected peptides and sterically minimal groups would be a welcome introduction into this area of chemistry.

1.8 Protein Radiolabelling

Another area of great interest and application in bioconjugation chemistry is radiolabelling. A radiolabelled protein can be tracked, *in vivo*, to enable elucidation of its mechanism of function or used as a diagnostic agent. This can be achieved using imaging methods such as single-photon emission computed tomography (SPECT) or positron emission tomography (PET). These are highly sensitive methods and only require micro to pico molar concentrations of imaging agent. Radiolabelled proteins are also used as therapeutics, a field known as radioimmunotherapy (RIT). The majority of examples in this area are antibody-based labelled drugs, used in the treatment of cancer.[114]

A wide range of methods, using various isotopes, including ^{18}F, ^{68}Ga, ^{123}I and ^{124}I, have been developed.[115–117] Three main strategies of labelling are utilised; direct labelling, indirect labelling using a prosthetic group, and indirect labelling via complexation, as shown in Scheme 33.



Scheme 33 - A - direct labelling, B - indirect labelling, C - indirect labelling via complexation.

A key advantage of radioactive fluorine is the short half-life of 109.7 minutes.[116] Medically, this is useful; however, the chemical methods used for the incorporation of ^{18}F must be rapid. This time limitation means that ideally the ^{18}F will be installed as the final step to reduce time between labelling and use. Fluorine is also preferred over other radioactive elements as the small size results in a relatively insignificant effect of the overall structure and activity of the protein.[115]

Recent developments focussing on direct radiolabelling of peptides and proteins includes site-selective chemistry for trifluoromethylation. This has included focus on Cys, Trp and Tyr residues. Cys was trifluoromethylated in fully unprotected peptides using the Umemoto reagent in under 20 minutes. More recently, a technique for the trifluoromethylation of Trp in unprotected peptides and proteins has been developed. This method uses readily available

sodium trifluoromethansulfinate and *tert*-butylhydroperoxide to label Trp residues in peptides and proteins.[57] Minor side-products where the method targeted His, Cys, Phe and Tyr were also observed using this method. Another group, following a similar method but using the zinc trifluoromethanesulfinate were able to selectively trifluormethylate Tyr in moderate to good yields.[118]

Other methods first attach a prosthetic group to the peptide or protein, then radiolabel this moiety through selective chemistry. There have been a wide range of publications reporting this technique, using chemistry including copper-catalysed and copper free azide-alkyne click chemistry, and nucleophilic displacement of aromatic leaving groups including stannanes[119] and quaternary amines.[120] Such methods typically require the installation of non-canonical amino acids that make them difficult to apply to biologically expressed proteins.

Alternatively, the group being attached to the protein can already bare the radioactive fluorine. An example of this uses the maleimide active ester of radiolabelled 3-fluorobenzoic acid. This undergoes reaction with either the amine group on the side chain of Lys, or the amine group on the N-terminus of the protein. There is also chance of other nucleophilic residues, such as Cys, Ser or Thr reacting with this activated ester. This leads to a heterogeneous mixture being formed due to the poor selectivity of this method. The procedure also involves multiple steps, resulting in a complex method and low yields, especially an issue when using radionuclides with a short half-life.

Direct labelling minimally affects a peptide/protein's structure as only a single element has changed. This is advantageous when the aim is to assess how the native protein interacts internally. Site-specificity ensures repeatability and guarantees the same product each time, important for therapeutic consideration as non-specific labelling would produce a mixture of modified isoforms. As direct labelling methods are often non-specific, require certain residues solvent accessible or use denaturing conditions they are not always the most suitable. Installing a prosthetic group can often be simpler, higher yielding and allows the protein to be refolded before the often simpler and milder radiolabelling step. Installing the radioactive group in the final step ensures a greater radiochemical yield.

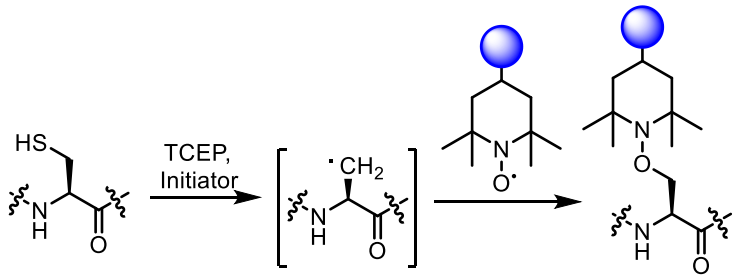
1.9 Aims and Objectives

1) Site-selective bioconjugation via interception of free-radical-mediated dechalcogenation

There is a wide range of chemical functionality displayed across the proteome, presenting the challenge to develop chemistry that is able to site-selectively modify peptides or proteins. Among the numerous methods developed to chemoselectively target the canonical amino acids, Cys features heavily. Many of these methods suffer from a range of drawbacks; these

include residue incompatibility, harsh reaction conditions, limited substrate scope, bioconjugate linker instability, and sequence dependent reactions.

The aim of Chapters 2 and 3 is to develop novel methods that address as many of these drawbacks as possible. Radical dechalcogenation is yet to be explored as a method of installing biologically relevant groups. We propose to install a functionalised 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) group, as shown in Scheme 34 via interception of the alanyl radical intermediate produced during dechalcogenation. This reaction should be rapid, chemoselective, and tolerant to a variety of groups appended to the persistent radical trap. As Cys is of low abundance and easily installed into recombinantly expressed proteins through site-specific mutagenesis, this would be a widely applicable method.

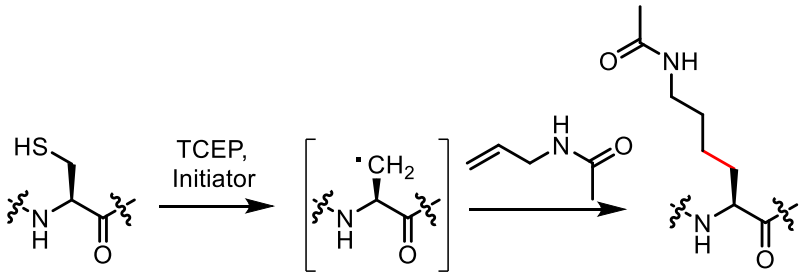


Scheme 34 - General scheme for the TEMPO conjugation of Cys.

Both Cys and Sec have proved vital in the development of methods to synthesise both difficult peptides and small proteins. Cys and Sec are exploited in NCL or DSL respectively to produce a ligated peptide or small protein with a native amide bond. These techniques will produce ligated sequences with either a single Cys or Sec residue. So far, these sites have only been used to convert Cys to Ala and Sec to Ala or Ser. This shows great potential for a method to be developed that exploits the produced alanyl radical for site-specific chemistry, post-ligation. If the modification of larger peptides is successful, then the method will be applied to biologically expressed proteins carrying a single Cys residue.

2) Site-selective installation of common Lys post-translational modification (PTMs)

Based on the approach described in Chapter 3, Chapter 4 will explore the development of a novel method to install functional groups into polypeptides via a stable $C(sp^3)$-$C(sp^3)$ bond. Any group installed using this method would be stable and, using an appropriate trapping agent, enable the incorporation of native PTMs common to Lys residues. The proposed method utilises allylamine groups; radical addition onto the unactivated alkene would form a four-carbon linker to the amine group attached via a $C(sp^3)$-$C(sp^3)$ bond. This method would allow the preparation of PTM'd proteins to use as tools to elucidate the function of Lys PTMs.

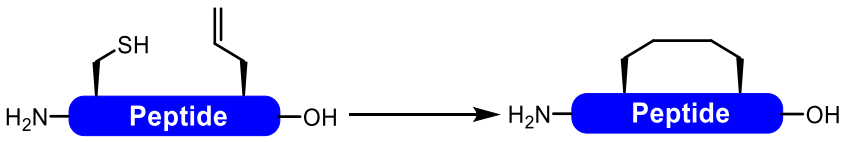


Scheme 35 - Example of proposed reaction with allyl acetamide. The resultant product is an acetylated lysine. This is a common post-translational modification installed on Lys.

The example in Scheme 35 shows the production of acetylated Lys. This is a common PTM of Lys seen in nature. It is important as it regulates the binding of histone proteins to DNA, thus controlled the expressions of genes on DNA. This is just one example of a Lys PTM, other include methylation, succinlyation, ubiquitination and formylation. All of these moieties could theoretically be installed in a site-selective manner using the proposed reaction to demonstrate the potential impact of this method.

3) Macrocyclisation of unprotected peptides

Macrocyclisation is a vital tool for the peptide pharmaceutical industry. Cyclic peptides have been shown to have higher activity and proteolytic stability than their linear counterparts. Chapter 5 describes the application of our radical chemistry to the production of a stable, unprotected macrocyclic peptide. The TEMPO conjugation chemistry could be developed to install an unnatural amino acid baring a TEMPO group, followed by dechalcogenation and intramolecular trapping. The allyl trapping method could also be applied in a similar manner, by synthesising or purchasing an allyl-modified amino acid for incorporating during SPPS. This method would enable cyclisation of a peptide sequence via a C-C bond, to produce a stable macrocycle (Scheme 36).



Scheme 36 - Proposed scheme for peptide macrocyclisation using desulfurative allyl trapping chemistry.

2. Selenocystine-selective Bioconjugation via Interception of Free-Radical-Induced Deselenisation with TEMPO traps

2.1 Chapter Introduction

Sec, the 21st proteinogenic amino acid, has become a utilised handle in the chemical synthesis of peptides and small proteins in recent years. Selenocystine, the oxidised dimer of Sec, is used in the rapid ligation method, DSL, to permit ligation in under 60 seconds.[27] Furthermore, internal Sec residues incorporated via SPPS or expression techniques have been utilised for Cys-like chemistry with electrophiles[92], via metal-assisted[90] and metal-free arylation[91], and allenamidation.[121] Sec can also be simply converted to other residues including Ala[27], Ser[45,46,122] and Dha.[123,124] The increasing prevalence and usefulness of Sec in ligation techniques and the increasing ease of incorporation of Sec into biologically expressed proteins demonstrates the potential of this highly useful reactive handle.

TEMPO is an example of a nitroxide and is characterised by the presence of a persistent organic radical. The radical is stabilised by electron donation from the nitrogen lone pair. This forms a two-center 3-electron bond as shown in Figure 5. This stabilisation combined with the steric hindrance from the gem-dimethyl groups prevent TEMPO from combining with itself. TEMPO has found a wide variety of uses, including as spin labels for biomolecules[125], nitroxide-mediated polymerisation of styrene[126] and as a catalyst for the metal-free oxidation of primary and secondary alcohols[127].

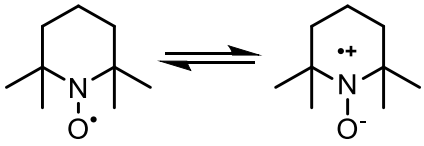
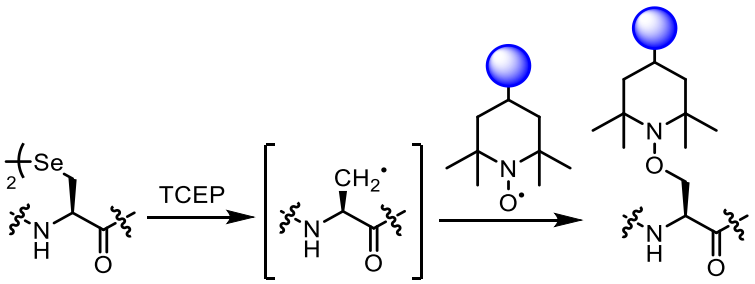


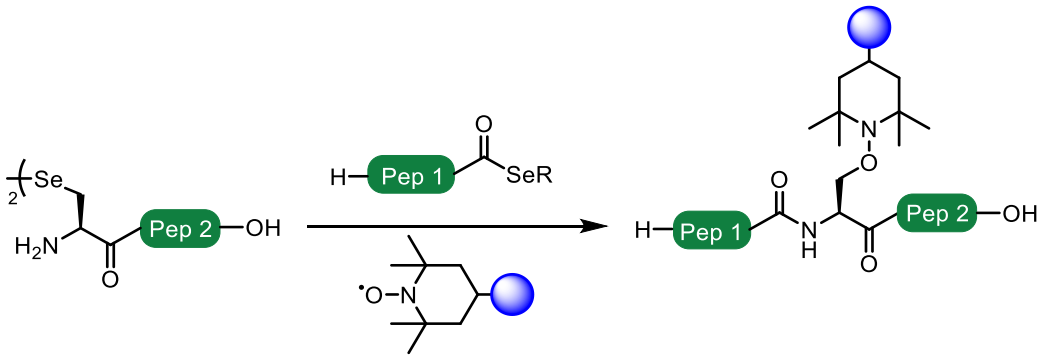
Figure 5 – Stabilisation of TEMPO radical through formation of a two-center 3-electron bond.

This chapter describes the development of a reaction that allows the installation of TEMPO into peptide models (Scheme 37). The method was first developed on a small model system (H-UAF-OMe), using rapid initiator inclusive conditions or milder initiator-free conditions. Both conditions result in complete conversion of the starting peptide to the desired conjugated product.



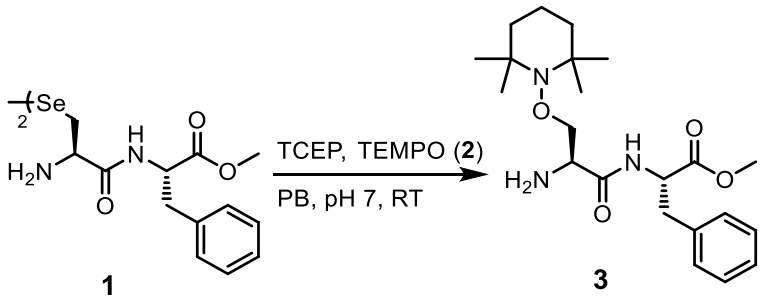
Scheme 37 - General scheme for the reaction to be optimised during this chapter. TCEP will be used to induce deselenisation, the persistent radical on TEMPO will trap the resultant alanyl radical to form the desired conjugate.

After the initial optimisation, a range of biologically relevant functionality was attached to the 4-position of TEMPO. The substrate scope was then assessed on the small model system, showing full compatibility with a range of biologically relevant functional groups. Next, the selectivity of the reaction was assessed, including a range of canonical amino acids that represent all functionality that could be present in a peptide. Finally, the method was applied to the development of a one-pot ligation-conjugation protocol (Scheme 38).



Scheme 38 - General scheme for the proposed one-pot ligation and conjugation procedure.

2.2 Reaction exploration using model peptides



Scheme 39 - Reaction scheme for TEMPO trapping reactions using H-UF-OMe dipeptide model.

The chosen model for initial reactions was H-UF-OMe (**1**), as shown in Scheme 39. This sequence was selected to balance the ease of synthesis whilst still representing a peptide system. The presence of an aromatic residue also aids in detection during HPLC analysis and purification. For these reactions a tertiary phosphine was required. As the reaction was to be performed in water, the water-soluble phosphine TCEP was chosen. The unsubstituted, commercially available TEMPO (**2**) was used as the radical trap. Due to TEMPO not being completely soluble in water, a co-solvent was included. Bioconjugations performed in aqueous conditions are often pH sensitive, so phosphate buffer was used to maintain the pH at 7.

All optimisation reactions were monitored by analytical HPLC to ascertain the reaction progress. The main peaks in the spectrum were integrated to show the percentage of each peak in the reaction mixture. For all tables of data the reported conversion shows the amount of starting material converted to products. Unless specified, the only product formed was the desired TEMPO conjugated product (**3**). There were a range of potential products that could be formed in these reactions, these are shown in Figure 6. For the initial reactions two main by-products were observed, the deselelnisation to Ala (**4**) and oxidative deselenisation to Ser (**5**). The initial role of TCEP in the reduction of the diselenide bond, forming the selenol product (**6**). When calculating conversion, this was treated as starting material.

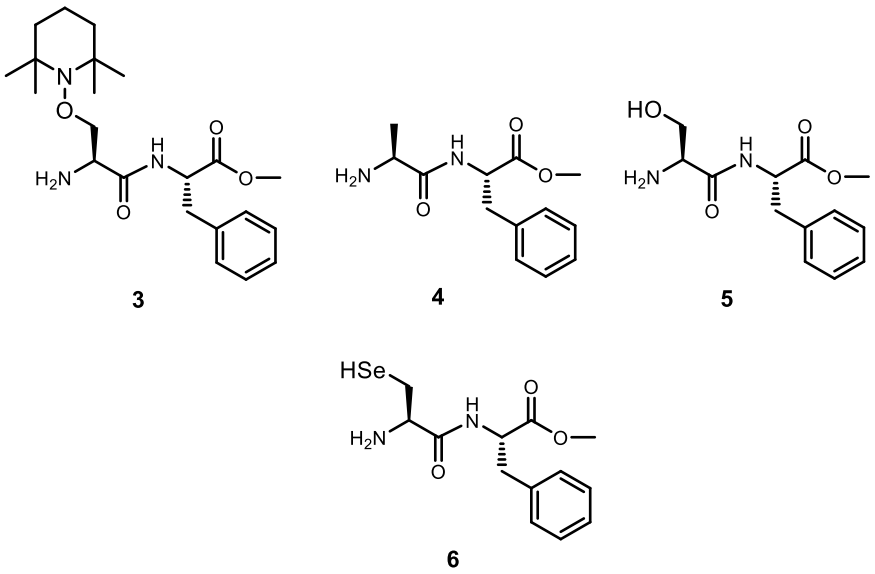
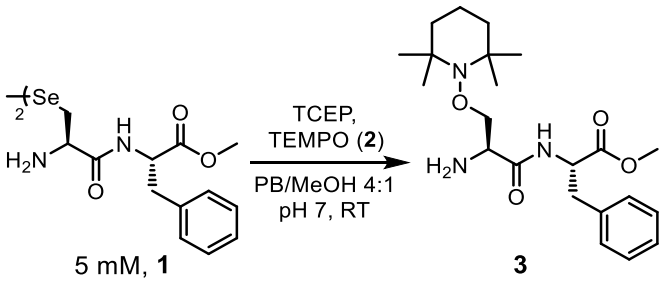


Figure 6 - Expected products from TEMPO conjugation reaction, including desired TEMPO conjugated product (**3**), deselenisation to Ala (**4**), oxidative deselenisation to Ser (**5**) and diselenide reduction to selenol (**6**).

The data in Table 1 shows the results of the small-scale initial reactions. All reactions were performed by first preparing stock solutions of reagents in either phosphate buffer (PB, pH 7) (peptide (**1**) and TCEP) or methanol (TEMPO) These initial reactions showed the effect of increasing the equivalents of both TCEP and TEMPO on the conversion of starting material to product.



Scheme 40 - General reaction scheme for entries 1 to 3. The equivalents of TCEP and TEMPO used in the reaction are shown in Table 1.

Table 1 - TCEP and TEMPO equivalents optimisation for Sec TEMPO trapping. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**3**).

Entry	TCEP/ eq.	TEMPO/ eq.	1 hour %conv.	2 hours %conv.	6 hours %conv.	16 hours %conv.
1	2	1	0	0	0	8
2	5	4	0	6	67	74
3	10	8	25	27	85	86

The data in Table 1 showed increasing conversion of starting material to product as the equivalents of TCEP and TEMPO were increased. At this point, the reaction was taking more than 6 hours to reach a reasonable level of conversion. Ideally, the reaction time would be significantly shorter than this, but these results represented a promising starting point. These trial reactions were performed with 20% methanol as the co-solvent, at ambient temperature (< 25 ^{o}C). An organic co-solvent was necessary to solubilise TEMPO. The reactions were doped with the same quantity of TCEP and TEMPO as the starting equivalents at 15, 30 and 90 minutes. This was to ensure a constant presence of TCEP and TEMPO, as it was observed that TCEP and TEMPO react with each other over time, forming oxidised TCEP and the corresponding piperidine.

This data shows that a higher equivalence of TCEP and TEMPO leads to a higher conversion of starting material to desired product (**3**) and faster reaction rate. Ideally, the TEMPO conjugate will be used in as low equivalence as possible as it could be modified with an expensive or difficult to synthesise group. TCEP is relatively cheap so using it in high equivalents is not a significant issue. In an attempt to increase the reaction rate, alternative conditions were trialled. The temperature was changed from ambient temperature to 37 ^{o}C, and the radical initiator VA-044 was included. It was theorised that the rate-limiting step could be the formation of the selenyl radical. As VA-044 is usually used to abstract the H-atom from a Cys during radical desulfurisation, this seemed a logical additive to trial. As the main by-product in the previous reactions was the formation of Ala, altering the concentration from 5 mM to 2.5 mM was also trialled to observe the affects. The results from these experiments are shown in Table 2.

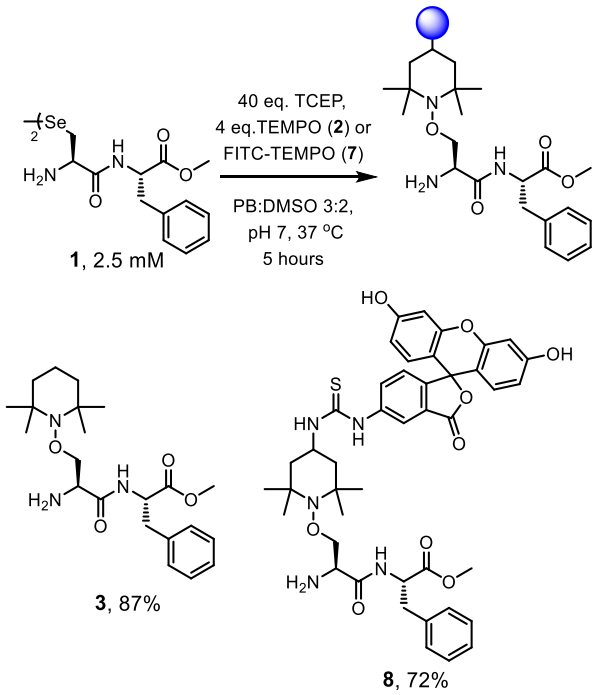
Table 2 - Results from TEMPO trapping optimisation at 37 ^{o}C. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**3**).

Entry	Peptide (**1**)/mM	TCEP/ eq.	TEMPO/ eq.	VA-044/ eq.	1 hour %conv. (**3**)	3 hours %conv. (**3**)	5 hours %conv. (**3**)
1	5	20	2	0	34	76	88
2	2.5	20	2	0	33	80	94
3	2.5	20	4	0	31	85	91
4	2.5	40	4	0	44	90	97
5	2.5	20	4	8	33	64	75

This data demonstrates that increasing the temperature from ambient (< 25 ^{o}C) to 37 ^{o}C causes a significant increase in reaction rate and decreasing the concentration from 5 mM to

2.5 mM results in a slight increase in final conversion. By the 5-hour injection, the conversion is > 90% for Entries 3 to 5 (Table 2), whereas at ambient temperature these levels of conversion were not reached, even after 16 hours. Entry 5 (Table 2) also showed that the incorporation of a radical initiator (VA-044), which increases the rate for standard deselenisation to Ala, has limited effect on the reaction rate. This could have been due to the low equivalents of VA-044 used.

The reaction was then performed on a preparative scale, using conditions from Entry 4 in Table 2, giving the desired conjugated product (**3**) in an 87% isolated yield. These conditions were also applied to an alternative TEMPO-based probe. This was to ensure the reaction could be applied to larger probes with more functionality. The chosen probe was functionalised with a fluorescein moiety using established chemistry (see the Experimental section 7.2 for details). An advantage of the TEMPO probes is the range of functionalised TEMPO compounds commercially available. These are typically functionalised in the 4-position, with functional groups including amino, oxo, carboxyl and hydroxyl all commercially available. This allowed a wide variety of simple chemistry to be considered when deciding which biologically relevant groups to append onto TEMPO. The synthesis of TEMPO probes will be discussed later in this chapter. Due to the hydrophobic nature of the fluorescein group, 40% DMSO was used to aid solvation. The products and resulting isolated yields from the fluorescein TEMPO probe (**7**) compared with standard TEMPO are shown in Scheme 41.



Scheme 41 - Isolated percentage yields from modified TEMPO trapping reactions.

Since the use of Sec over Cys enables the rapid ligation of peptide fragments, it was hoped to apply this reaction to a one-pot ligation and conjugation procedure. These ligation reactions require a solvent system of 6 M Gdn•HCl and 0.1 M Na_2HPO_4 (PB) in water, known as ligation buffer (LB). The inclusion of the guanidine in the solvent system denatures the peptide and aids dissolution of the peptide fragments. Currently, all reactions described have only used 0.1 M PB as the solvent, so from this point onwards ligation buffer was used. Interestingly, this solvent change resulted in an impurity peak at 4.5 minutes, with the desired product (**3**) showing at 5 minutes. The trace for this reaction is shown in Figure 7.

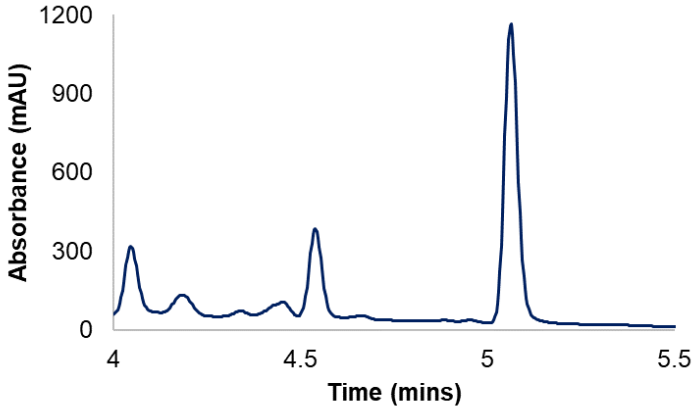


Figure 7 - HPLC trace from TEMPO trapping experiment using 2.5 mM peptide in LB at 37 °C for 5 hours with 40 eq. TCEP and 4 eq. TEMPO, showing a significant by-product (**9**) at 4.5 minutes. HPLC gradient 10-70% B over 5 mins, 210 nm.

The reaction from Scheme 41 was performed on a larger scale to allow isolation of the side-product (**9**) eluting at 4.5 minutes on the analytical HPLC trace shown in Figure 7. The HPLC trace from the pure 4.5 minute product is shown in Figure 8. The product (**9**) was also analysed by MS, the results of this are shown in Figure 9.

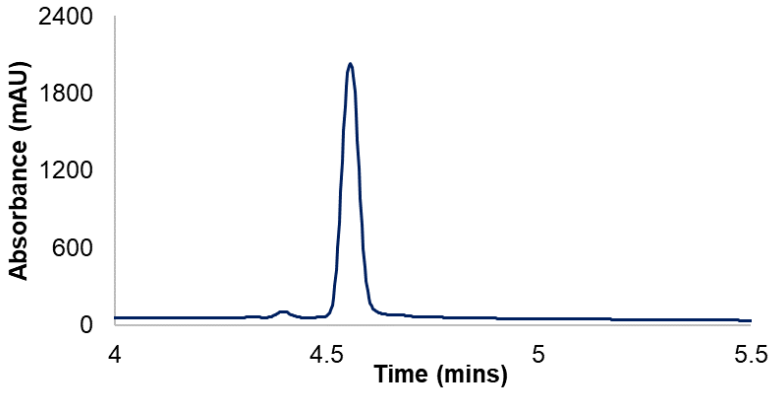


Figure 8 - HPLC trace showing the pure 4.5 minute by-product (**9**). HPLC gradient 10-70% B over 5 mins, 210 nm.

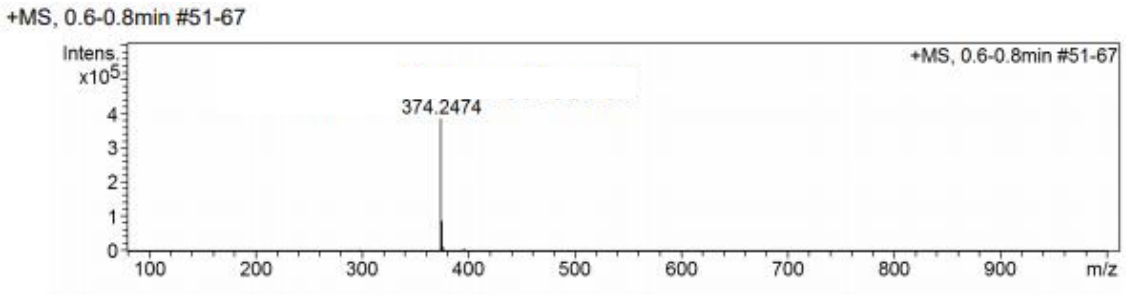


Figure 9 - MS analysis of the 4.5 minute by-product product (**9**), showing a single molecular weight ion of 374.2.

Based on the mass of the product it would appear that during the reaction a known base-induced peptide side-reaction was occurring. This involves the N-terminus of a dipeptide attacking the C-terminus to form a diketopiperazine by-product. In this case, the methyl ester acts as a leaving group, allowing the cyclisation to occur. The proposed product is shown in Figure 10. This is a proposed structure and could not be confirmed by NMR as the quantity isolated too little to obtain an NMR. To prevent this from occurring in future studies, a tripeptide H-UAF-OMe (**10**) was used. This was synthesised as described in section 7.3. This model will prevent DKP formation from occurring, allowing simple analysis of the reaction mixtures.

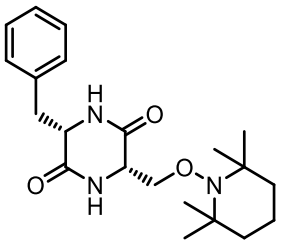
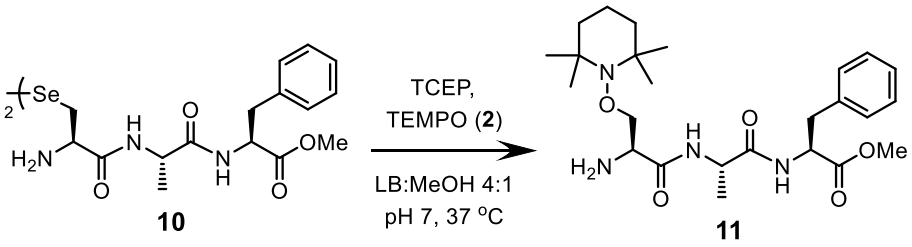


Figure 10 - Proposed structure of the 4.5 minute by-product (**9**), showing the 6-membered diketopiperazine formed by the N-terminal amine displacing the C-terminal methyl ester.

2.3 Reaction optimisation using H-UAF-OMe

The data from section 2.2 using the H-UF-OMe (**1**) model showed that the desired TEMPO conjugated product was being formed with high associated yields. During this section further reaction optimisation was performed on the H-UAF-OMe model (**10**). The aim was to develop a selective method of rapid conjugation, using the lowest equivalents of the TEMPO reagents as possible. As previous reactions have shown the reaction rate is fastest with higher equivalents of the reagents, 50 eq. of TCEP and 25 eq. of TEMPO were initially chosen. The previous data also showed that performing the reactions at 37 ^{o}C instead of ambient temperature resulted in a higher reaction rate, so the following optimisation was performed at 37 ^{o}C.



Scheme 42 - General scheme for the conjugation of TEMPO onto H-UAF-OMe (**10**), forming conjugated product **11**.

For the initial optimisation the protocol followed is outlined in section 7.2. Briefly, the model peptide (**10**) was dissolved in LB/MeOH, TCEP and TEMPO (**2**) were then added and the pH of the reaction was ensured to be at 7 (unless otherwise specified). Reactions which were doped had additional TCEP and TEMPO (**2**) added at specific time points. If doped once at 15 minutes, if doped twice at 15 and 45 minutes, and if doped three times at 15, 30 and 45 minutes. If heated, reactions were placed in a Thermo heater/shaker at the specified temperatures. All reactions were monitored by analytical HPLC, the conversion was assessed by integrating the main peaks. For reported reactions that do not reach completion, the remainder of the material present was either starting material or deselenisation to Ala.

The initial factor optimised was the concentration of the peptide (**10**). Previous data (Table 2) showed that decreasing the concentration from 5 mM to 2.5 mM resulted in an increase in reaction rate. The concentration of the peptide was further decreased to ascertain if the rate would continue to increase. In addition, as our intention is to perform this reaction on large peptides and small proteins, a low concentration is more appropriate to complex polypeptide systems in terms of operationally simplicity and analysis. The results from this are shown in Figure 11.

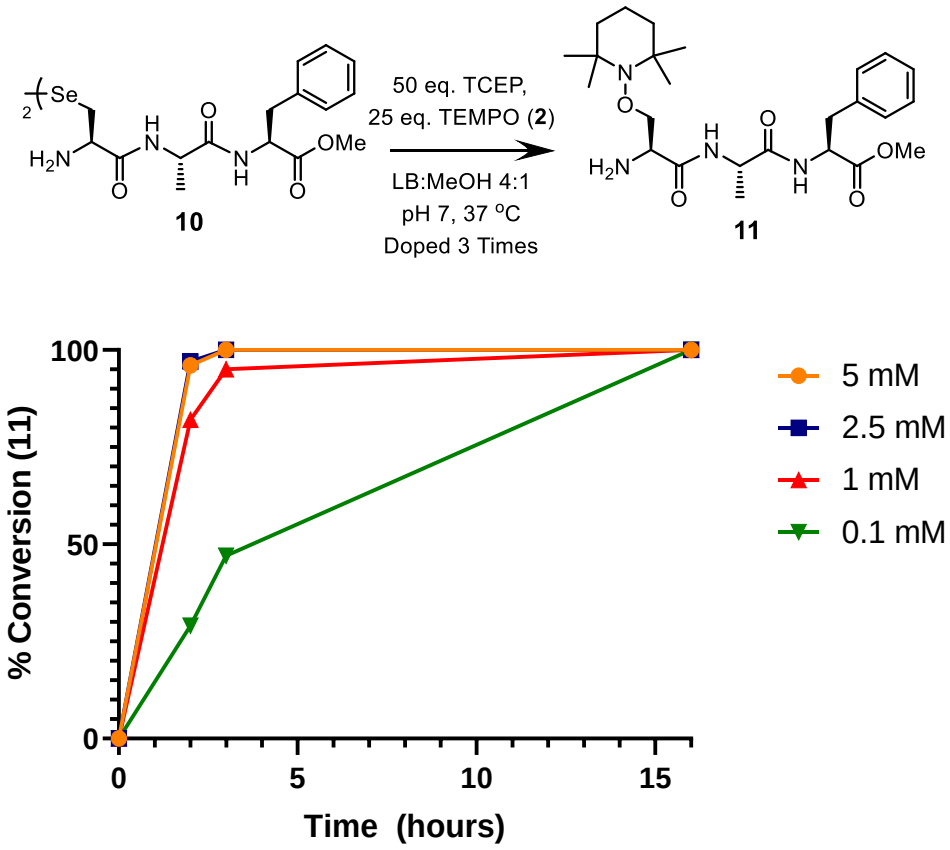


Figure 11 – Reaction scheme and time point conversion progression plot from the bioconjugation of H-UAF-OMe (**7**) with TEMPO at a range of peptide concentrations. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**8**).

The data from Figure 11 shows that performing the reaction at lower concentrations results in a lower rate of reaction, with the optimal concentration being 2.5 mM. Although slower, the reaction does still reach completion after 16 hours, which is promising for the potential application to protein models. These reactions were all doped with additional TCEP and TEMPO three times. For all further optimisation, the reaction was performed at the optimal concentration, 2.5 mM. As the equivalents of TCEP has so far shown to have a significant effect on the reaction, this was the next variable to be optimised. The data from this optimisation is shown in Figure 12.

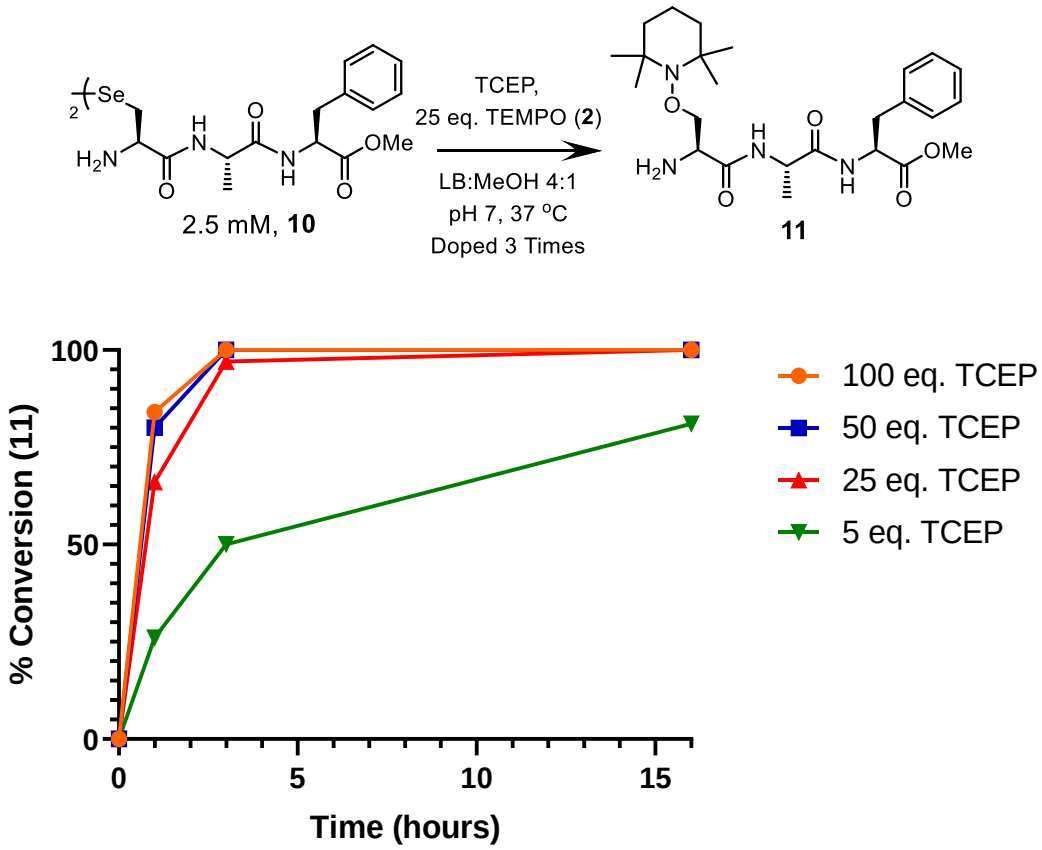


Figure 12 – H-UAF-OMe (**10**) TEMPO trapping optimization – TCEP concentration. The reaction was doped three times with additional TCEP and TEMPO. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**11**).

The optimisation of TCEP equivalence, as shown in Figure 12, shows that the reaction rate increases as the equivalence of TCEP is increased. This was expected, as TCEP is involved firstly in the reduction of the diselenide bond, then in the deselenisation stage.

Based on this data it is clear that an excess of TCEP is vital to reaction rate, so the TCEP equivalents remained at 50 for future reactions. In the studies up to this stage, all reactions were doped with extra TCEP and TEMPO (**2**) three times in the first hour. As this results in a very large excess of each reagent being used, the next set of optimisation data focussed on ascertaining whether doping was necessary. This data is shown in Figure 13.

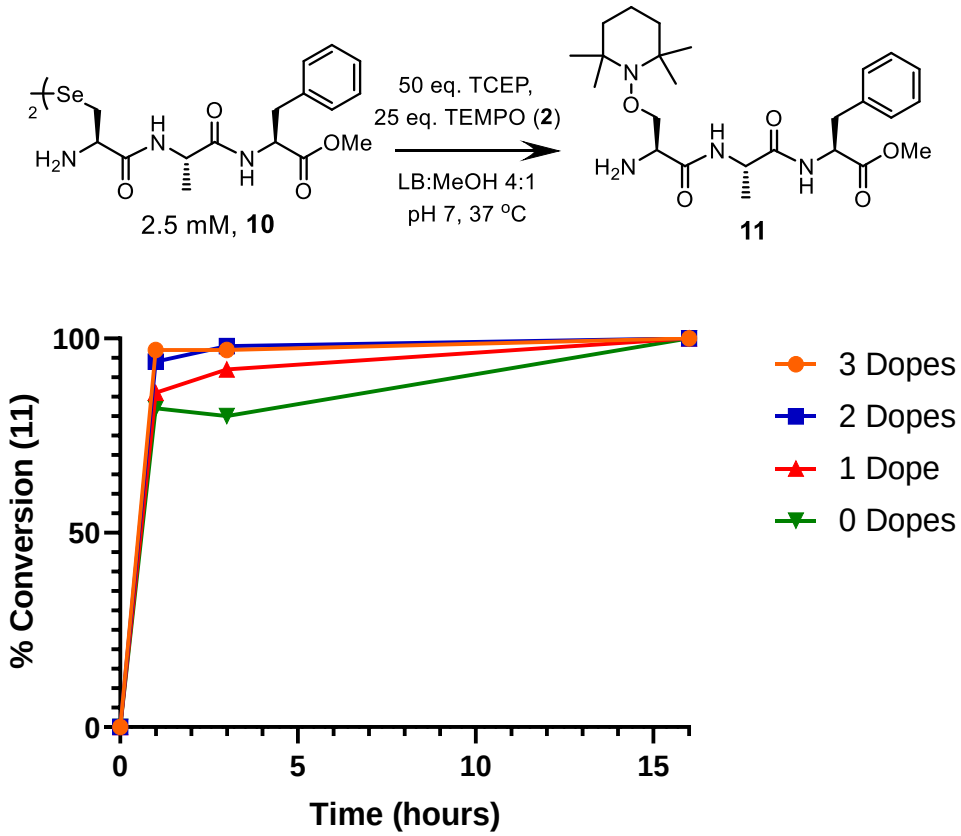


Figure 13 – Reaction optimisation by altering the number of times that the reaction mixture is doped with addition quantities of TCEP and TEMPO. Percentage conversion shows the conversion of starting material to products, unless specified the sole product formed was (**11**).

The doping optimisation data showed that the reaction rate was not significantly affected by reducing the number of dopes from three to two. Reducing the number of dopes even further still resulted in quantitative conversion, albeit with a slightly longer reaction time. In the interest of finding the optimal conditions for rapid bioconjugation, two dopes were used as the optimised conditions in further reactions, unless otherwise stated. The next variable focussed on was the equivalents of TEMPO (**2**). As TEMPO (**2**) reacts directly with the alanyl radical, it is clear that the TEMPO (**2**) equivalents should have a significant effect on the reaction rate. As the aim for this project is for the TEMPO (**2**) to be further functionalised with biologically relevant molecules, such as drug molecules, having the lowest equivalents possible is necessary to reduce the cost of the reaction. Reducing the cost of the reaction is vital for the method to be considered for the pharmaceutical industry. The data from TEMPO (**2**) optimisation studies is shown in Figure 14.

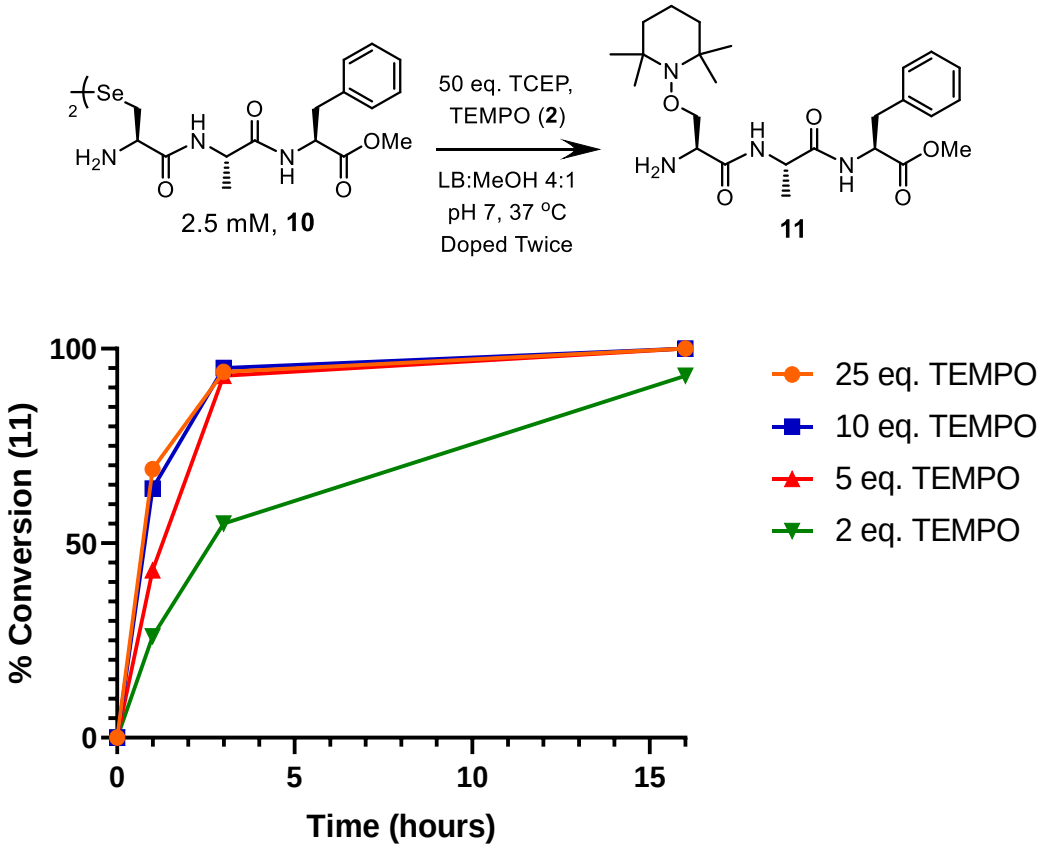


Figure 14 – Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the equivalents of TEMPO (**2**) in the reaction mixture. All entries were doped twice with additional TCEP and TEMPO (**2**). For all entries the sole product formed was (**11**).

The data obtained for the optimisation of TEMPO (**2**) equivalents was very promising as it showed a high excess of the TEMPO (**2**) probe was not necessary. Reducing the equivalents of TEMPO (**2**) from 25 to 5 did not have a significant effect on the reaction rate, therefore, 5 equivalents was used for all further reactions. In the interest of reducing the equivalents, the TCEP concentration was also reduced to 25. For a bioconjugation method to be applied to proteins it is important that a near-neutral pH can be used. As all the reactions to this point were performed at pH 7 and resulted in quantitative conversion this was very promising. As it is shown in the literature that pH can have a large effect on reaction rate, it was of interest to see whether that would be the case for this bioconjugation method. The data from performing this reaction at a range of pH values is shown in Figure 15.

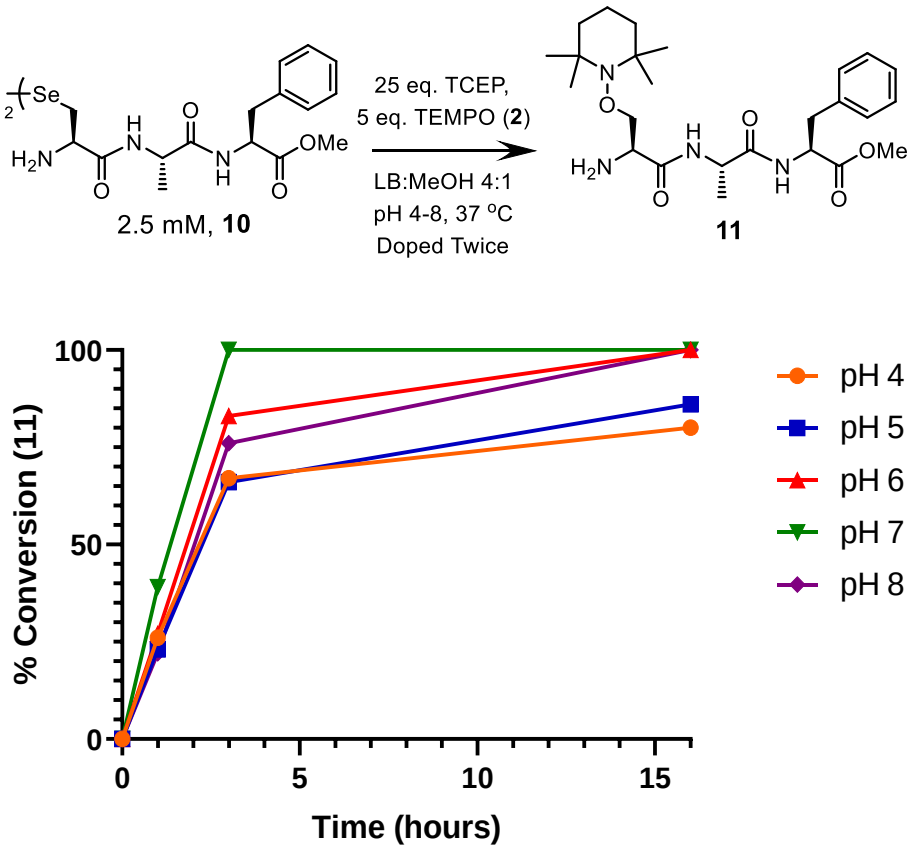


Figure 15 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the pH of the reaction mixture. All entries were doped twice with additional TCEP and TEMPO.

The data from varying the pH of the bioconjugation reaction shows that a pH value of 6 to 8 is optimal for the reaction, forming only desired product (**11**). Reducing the pH below six results in the formation of the Ala by-product (**12,** Figure 16) associated with H-atom abstraction and quenching of standard radical deselenisation. Above pH 8, there is an increased risk of hydrolytic degradation of the peptide material. This data is useful, showing that further optimisation can continue to be performed at pH 7.

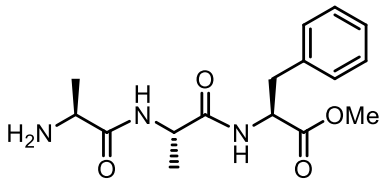


Figure 16 – Product (**12**) formed from the radical deselenisation of H-UAF-OMe (**10**) when TEMPO conjugation does not occur.

With these conditions in hand, the organic co-solvent composition was the next factor to be probed. Ideally, the conjugation reaction would be performed in pure aqueous conditions. This, unfortunately, will not always be possible depending on the solubility of the TEMPO (**2**) probe

being used. To allow for a range of probes to be used, it will be important for the reaction to be unaffected by the percentage of organic solvent used in the reaction. The data from this study is shown in Figure 17.

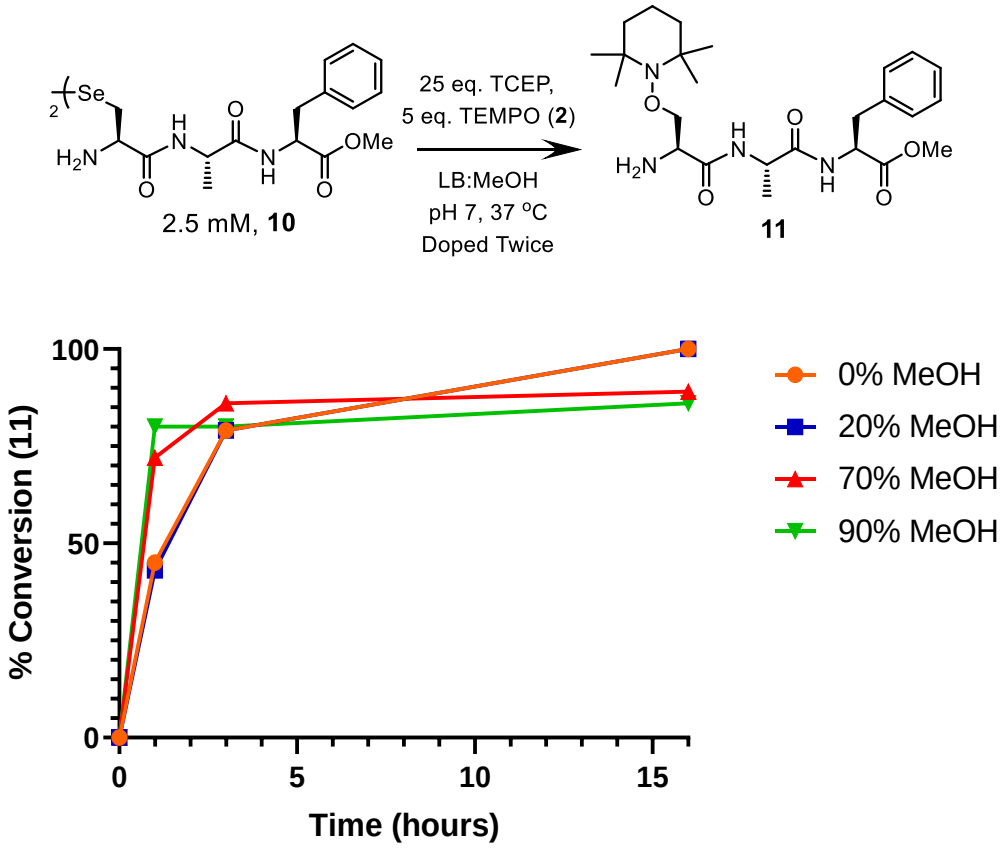


Figure 17 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the percentage of organic solvent in the reaction mixture. All entries were doped twice with additional TCEP and TEMPO. The percentage conversion shows how much desired product (**11**) was formed compared to any by-products or starting material.

The data in Figure 17 illustrates that the reaction is affected by the percentage of organic solvent used. The higher percentages (70% and 90%) of organic solvent still resulted in good conversion to desired product (**11**), but Ala by-product (**12**) was observed. Gratifyingly, this data does show that the reaction can be performed in 100% aqueous solvent conditions, important for the application to proteins. With the same reasoning in mind as for the previous study, the next factor focussed on was which organic solvents are suitable to be used as co-solvents for the reaction. The results from this data are shown in Figure 18 .

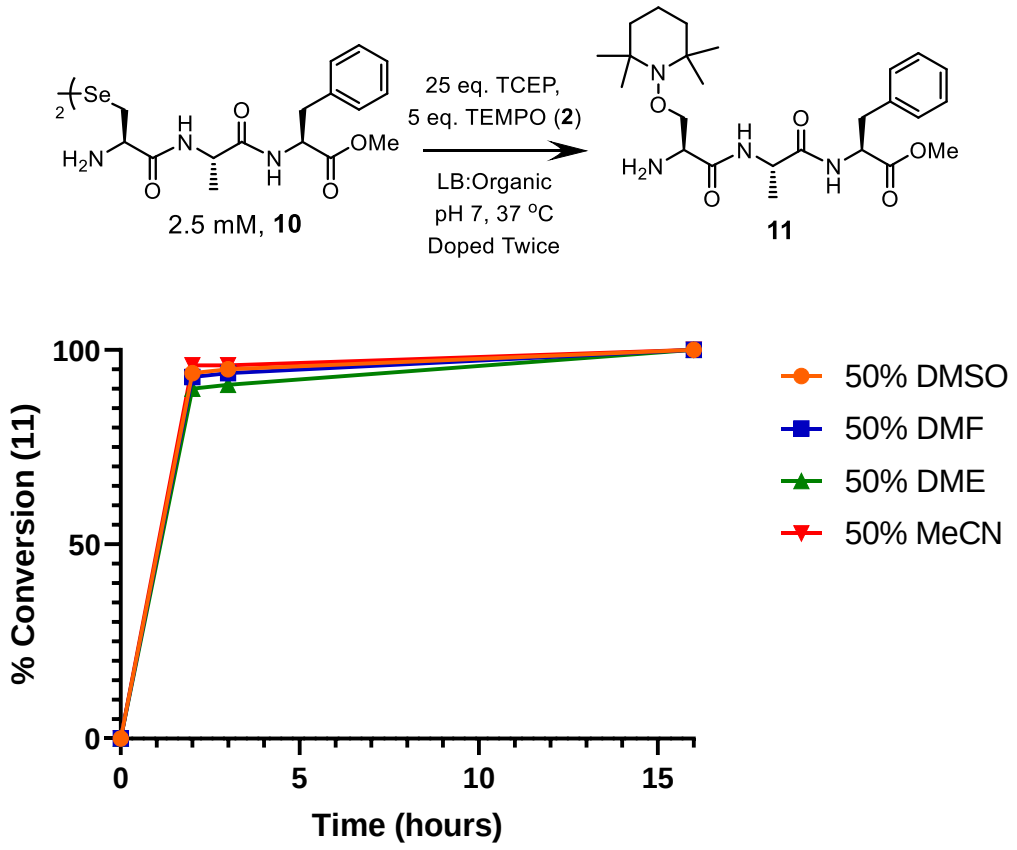


Figure 18 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the organic co-solvent in the reaction mixture. All entries were doped twice with additional TCEP and TEMPO. For all entries the sole product formed was (**11**).

The data in Figure 18 demonstrates that a range of organic co-solvents can be used in this reaction. This is useful as it will allow for a wide range of TEMPO probes to be used. The aim from this point onwards was to develop conditions that result in complete conversion to desired product (**10**) in under 1 hour, whilst only requiring low equivalents of TEMPO. The results in Figure 19 show how further increasing the temperature affects the reaction rate.

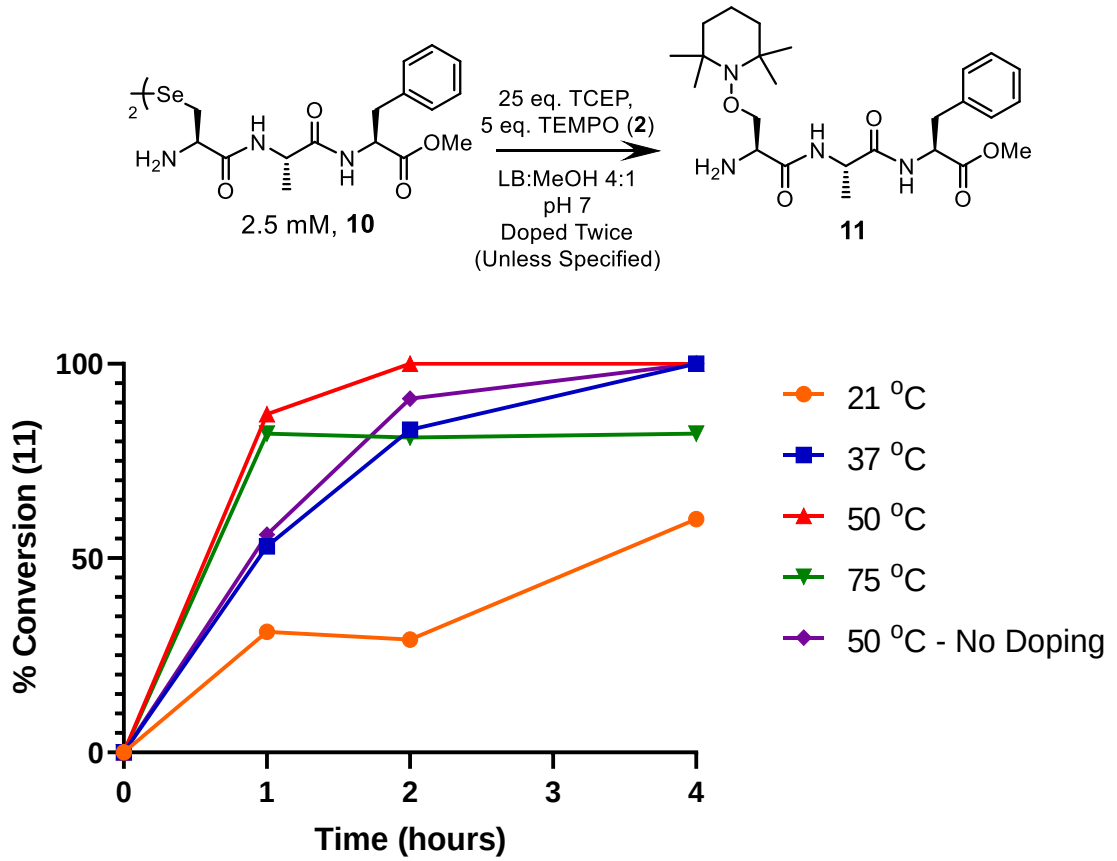
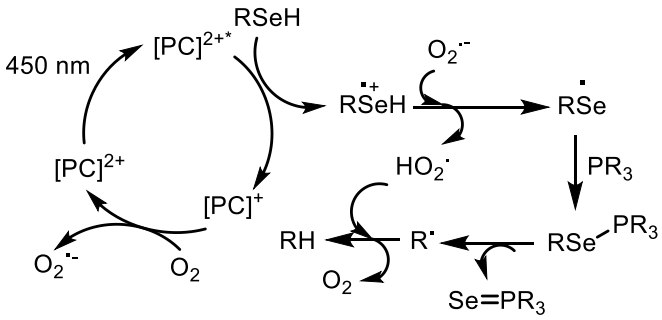


Figure 19 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the temperature of the reaction mixture. All entries were doped twice with additional TCEP and TEMPO. For all entries the sole product formed was (**11**).

As previously shown, increasing the temperature from ambient to 37 °C resulted in a significant increase in reaction rate. The data in Figure 19 show that increasing the temperature to 50 °C results in a further increase in reaction rate. Interestingly, increasing the temperature to 75 °C resulted in a lower conversion due to the production of the Ala by-product. As increasing the temperature resulted in a significant increase in reaction rate, it was tested whether doping was still beneficial to the reaction. This showed the rate is slightly reduced, but as the equivalents of TEMPO used was significantly lower these conditions were preferred. Using less TEMPO is ideal as it means this technique could be applied to the installation of expensive or hard to synthesise conjugates.

The previous data showed that doping the reaction mixture was not necessary to achieve complete conversion to TEMPO conjugated product. To try and further increase the rate of the reaction, a range of initiators were tried. The role of the initiators was to increase the rate of selenyl radical formation. The metals/photocatalysts typically achieve this by a single electron transfer (SET) mechanism (Scheme 43). A single electron is removed from the selenol, forming a radical cation, which is rapidly deprotonated to form the selenyl radical. It

was expected that the chosen metals would follow a similar SET mechanism, without being regenerated by oxygen so 2 equivalents of these were used.



Scheme 43 - Proposed reaction mechanism for the photoredox catalysed deselenisation of selenocysteine.

Diazo initiators, such as VA-044, play a slightly different role. These initiators directly abstract the H-atom from the selenol to form the selenyl radical. The following optimisation reactions were performed at 50 ^{o}C with the initiator being added after the TCEP and TEMPO (**2**). The results from this are shown in Figure 20.

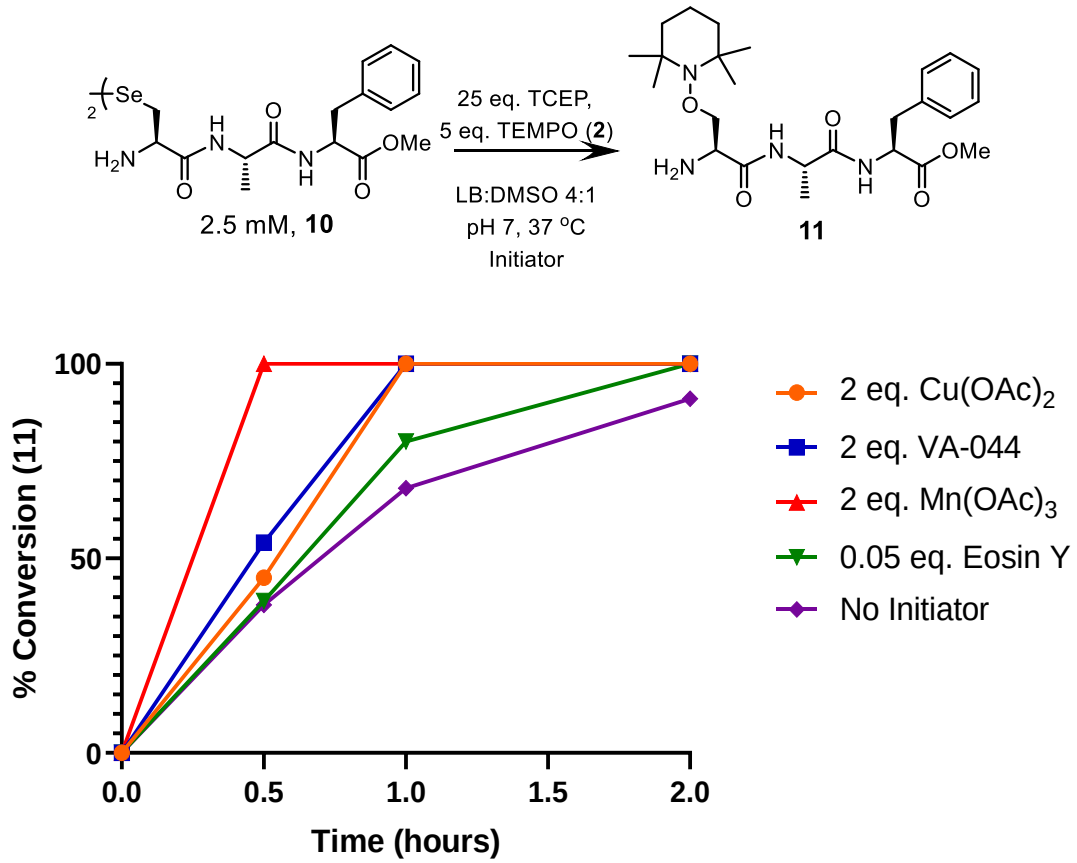


Figure 20 – Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the initiator in the reaction mixture. All entries were doped twice with additional TCEP and TEMPO. For all entries the sole product formed was (**11**).

The results in Figure 20 were very promising as they showed that a range of initiators are able to increase the rate of the reaction. Using 2 eq. $Mn(OAc)_3$ was particularly successful as it showed complete conversion to the desired product (**11**) in only 30 minutes. In the literature Mn(III) reagents are known to be single electron oxidants of thiols to form thiyl radicals.[128] The increase in reaction rate seen here could show that, previously, the rate-limiting step was the generation of the selenyl radical after reduction of the diselenide. The addition of the Mn(III) reagent rapidly forms the selenyl radical, increasing the overall reaction rate. To further probe the Mn(III) accelerated conditions, a range of reactions were performed to see if the amount of TEMPO (**2**) probe could be further reduced without significantly reducing the reaction rate. The results from these experiments are shown in Figure 21.

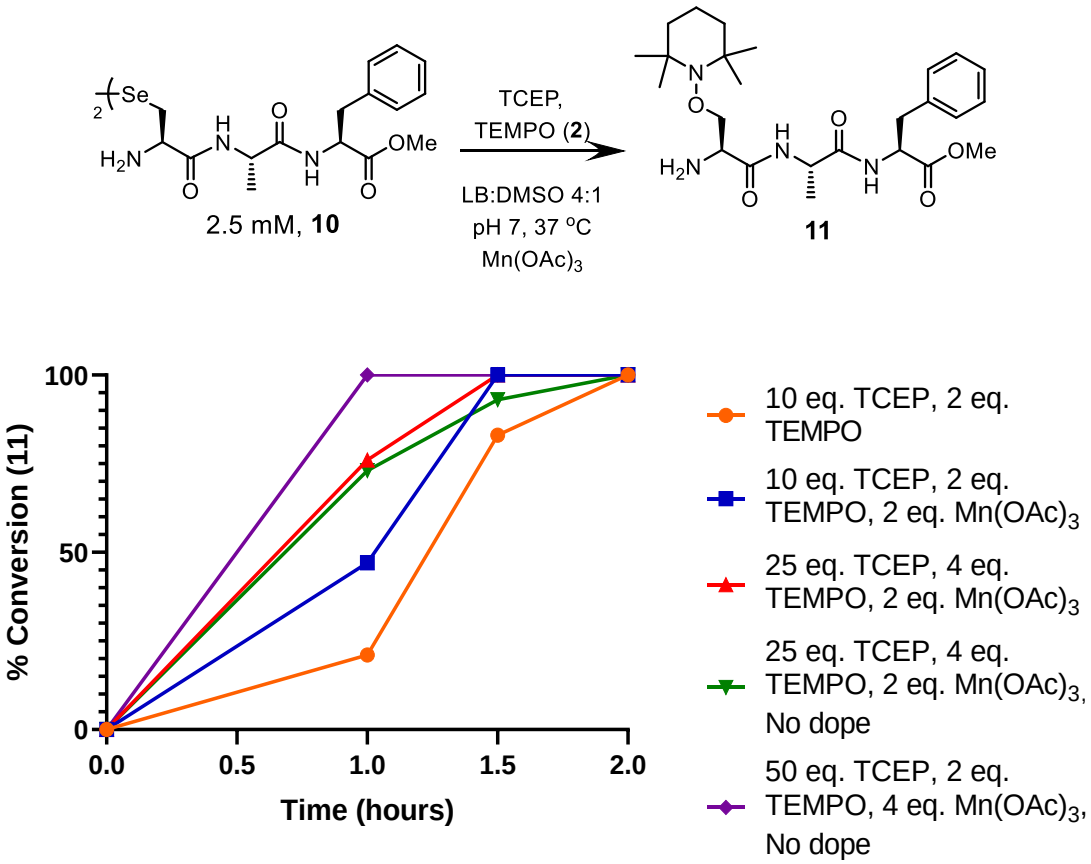


Figure 21 - Reaction scheme and conversion plot showing the conversion of starting material (**10**) to products over time, varying the equivalents of $Mn(OAc)_3$ in the reaction mixture. All reactions doped twice unless specified. For all entries the sole product formed was (**11**).

The data in Figure 21 showed that the addition of the Mn(III) additive enabled the reduction of TEMPO (**2**) equivalents to two. This is a significant decrease from where the optimisation of this reaction started. The decrease to two equivalents allows the attachment of expensive groups that are challenging to synthesise without significant waste. The reaction still reached completion in under 1 hour, showing the ability of the developed method to rapidly label peptides. The crude HPLC trace using the final optimised conditions (2.5 mM peptide (**10**), 50 eq. TCEP, 2 eq. TEMPO, 4 eq. $Mn(OAc)_3$ in 4:1 LB:DMSO pH 7 at 37 °C for 1 hour) is shown in Figure 22.

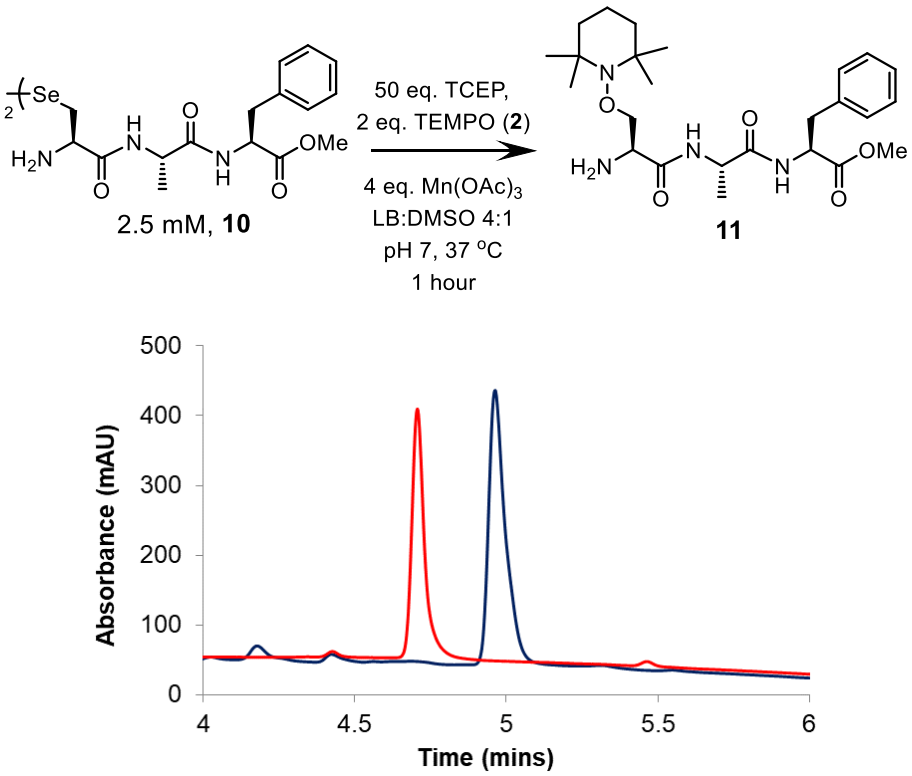
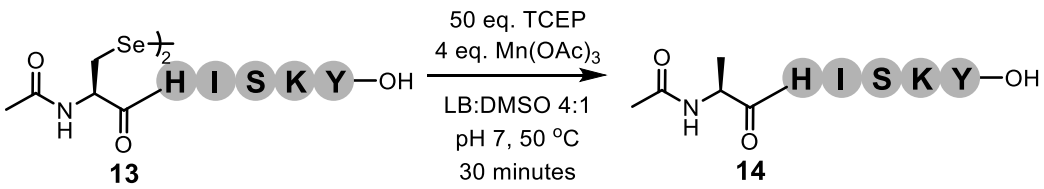


Figure 22 – Reaction scheme and the associated crude HPLC trace compared with the HPLC trace of the pure starting H-UAF-OMe (**10**, red). Analytical HPLC gradient 10-70% B over 5 mins, 210 nm.

The crude HPLC trace in Figure 22 shows complete conversion from the starting H-UAF-OMe (**10**, Red) to the desired TEMPO conjugated product (**11**, blue) by the time the reaction was analysed at 1 hour. The final optimised conditions were 2.5 mM peptide, 50 eq. TCEP, 2 eq. TEMPO (**2**) and 4 eq. $Mn(OAc)_3$ in LB at pH 7 for 1 hour, referred to as general protocol A. The next step was to attempt modification of larger models, and the first attempt of developing a one-pot method for the ligation and bioconjugation of peptides and small proteins. Firstly, as Mn(III) reagents are not commonly used for the modification of peptides, it was important to ensure that only the expected reaction at Sec occurred in the presence of other canonical amino acids.

This control reaction was performed on the Sec containing peptide model Ac-UHISKY-OH (**13**). This peptide model was synthesised using SPPS, as described in section 7.4. This model was chosen as it contained some diverse and reactive functionality commonly found in peptides and proteins. The expected reaction scheme for this control reaction is shown in Scheme 44.



Scheme 44 - Expected deselenisation product (**14**) from the treatment of the Sec containing peptide Ac-UHISKY-OH (**13**) with TCEP, accelerated by Mn(OAc)3.

The expected by-product was also synthesised by SPPS (Ac-AHISKY-OH, **14**) as described in section 7.4 and used as a reference so the analytical scale reaction could be compared to the expected product trace. Analytical scale reaction performed using Ac-UHISKY-OH (**13**) using 2.5 mM peptide with 50 eq. TCEP and 4 eq. $Mn(OAc)_3$ at 50 °C for 30 minutes. (General protocol A without the addition of TEMPO). Reactions were monitored after being at 50 °C for 30 minutes. The resultant crude HPLC trace for this reaction is shown in Figure 23.

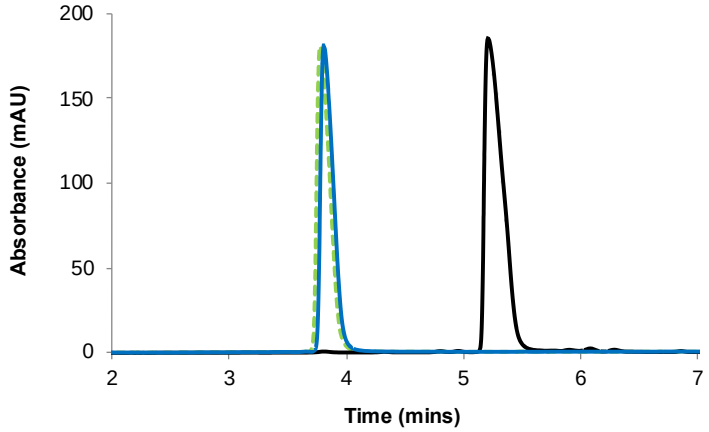


Figure 23 - Analytical HPLC trace of Ac-UHISKY-OH (**13**) deselenisation to Ala product (**14**). Analytical gradient 10-25% B over 5 min, 210 nm. Blue – Reaction mixture, Black – Starting material (**13**), Green dash – Ac-AHISKY-OH standard (**14**).

The HPLC trace in Figure 23 shows that the expected reaction did occur. There were no by-products present, only clean deselenisation to the corresponding Ala product. This shows that the Mn(III) chemistry developed in this project can also be applied to the rapid and clean deselenisation of Sec (a conversion that usually takes a few hours with the VA-044 reagent which is commonly employed). After observing the expected outcome, it was important to see whether these reaction conditions could be applied to a peptide that also carries a Cys residue. The above conditions were applied to the model peptide Ac-CHISKY-OH (**15**). The analytical scale reaction showed that after 30 minutes the same product was formed as previously seen on the Ac-UHISKY-OH (**13**) model. This shows that the Mn(III) conditions can also be used to

afford rapid and clean desulfurisation of Cys to Ala, therefore, the optimised conditions are not applicable to Cys-containing peptides. The results from this reaction are shown in Figure 24.

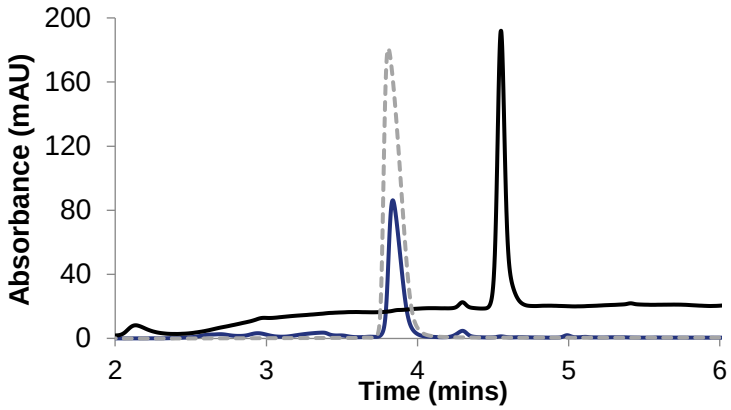
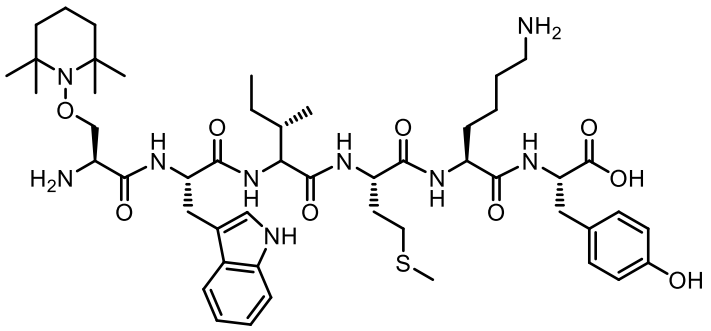
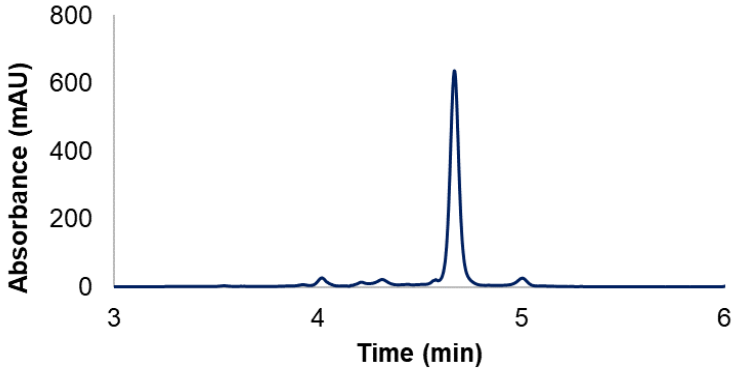


Figure 24 - Resultant HPLC traces from applying the Mn(III) conditions on the Cys containing peptide Ac-CHISKY-OH (**15**). Three traces are overlaid, reaction mixture (blue), starting material (**15**, black) and Ac-AHISKY-OH standard (**13**, green dash). Analytical gradient 10-25% B over 5 mins, 210 nm.

Next, the model peptide H-UWIMKY-NH_2 (**16**) was synthesised (section 7.4) and the optimised Mn(III) accelerated TEMPO trapping conditions (general protocol A, section 7.2) applied. This sequence was chosen as it includes the oxidation sensitive Met and Trp residues and the Mn(III) reagent is known to be oxidising. The following work was performed by Dr. Frances Smith (PDRA in the Mitchell Group) following the methods developed during this project. H-UWIMKY-NH_2 (**16**) was synthesised and subjected to general protocol A (section 7.2).



17



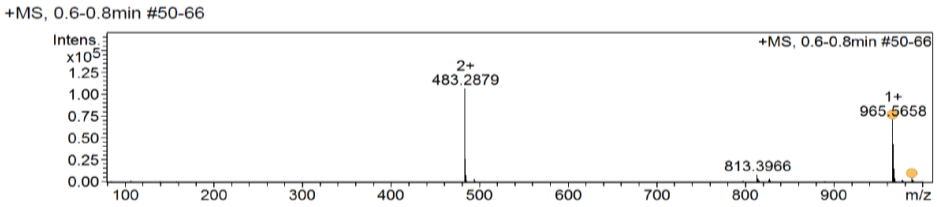
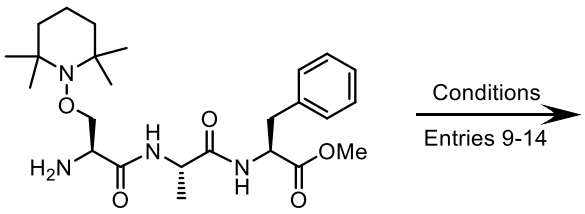


Figure 25 – Chemical structure and analytical HPLC trace of pure TEMPO-WIMKY-NH_2 (**17**). Analytical gradient 10-75% B over 10 min, 210 nm. HRMS Calculated Mas: $[M+H]^+$ = 965.5641, $[M+2H]^{2+}$: 483.2860. Actual Mass: $[M+H]^+$ = 965.5658, $[M+2H]^{2+}$: 483.2879.

The resultant pure HPLC trace and product MS is shown in Figure 25. The MS shows only the molecular weight for the desired product (**17**) and no sign of Met oxidation. The desired product was isolated in a 68% yield. This was an important validation reaction as it showed the compatibility of the bioconjugation method with sensitive canonical amino acids.

The data from this chapter so far shows the successful development of a residue-specific method of bioconjugation. The method uses only 2 equivalents of the conjugate probe, in aqueous conditions at neutral pH with a short reaction time. These results showed excellent promise that warranted further study. As the reaction can be performed using reasonable stoichiometric amounts of TEMPO in a rapid manner, the method development was continued. For bioconjugates, it is important that the conjugate formed is stable. The data in Table 3 shows the results from subjecting the probe to a range of conditions.

Table 3 - Data from the stability tests of the H-A(TEMPO)AF-OMe (**11**) conjugate.

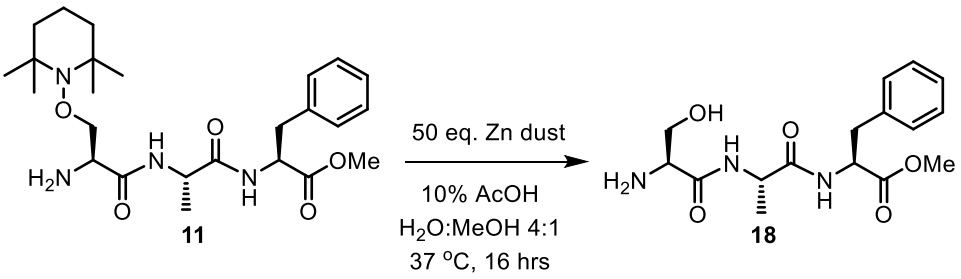


11

Entry	Conditions	Time/hr	% Degradation
1	pH 2	16	0
2	pH 9	16	0
3	5 eq. GSH/DTT	16	0
4	50 ^{o}C	16	0
5	5 eq. V-50	16	0
6	50 eq. Zn/ 10% AcOH - 37 ^{o}C	16	100

The results in Table 3 show a range of conditions employed to test the stability of the TEMPO conjugate (**11**). Pleasingly, the conjugate showed complete stability to all conditions typically employed to new conjugates. For the potential of this method to be employed for the synthesis of peptide-drug conjugates, the stability of the conjugate to acidic conditions and thiol additives are particularly important. Furthermore, the conjugate showed excellent stability to high temperatures, basic pH and radical initiators.

The conditions in Scheme 45 were attempted to explore reductive cleavage of the probe. As mentioned in the introduction, the N-O bond can be cleaved in the presence of Zn in acetic acid. This result was interesting, as after 16 hours the conjugate had fully degraded, showing complete and clean conversion to a new product (**18**). The HPLC trace in Figure 26 shows the starting material trace (**11**, blue) and the product (**18**) from the conditions in Entry 6 (Table 3, red). The conditions in Entry 6 (Table 3, red) were repeated on an isolatable scale to characterise the degradation product. Interestingly, the molecular weight of the new product, as shown by ESI-MS, matched that of N-O homolysis forming the native residue serine (Ser, **18**). The product was obtained in a 91% isolated yield.



Scheme 45 - Reductive cleavage of TEMPO conjugate (**11**) to form Ser-containing peptide (**18**).

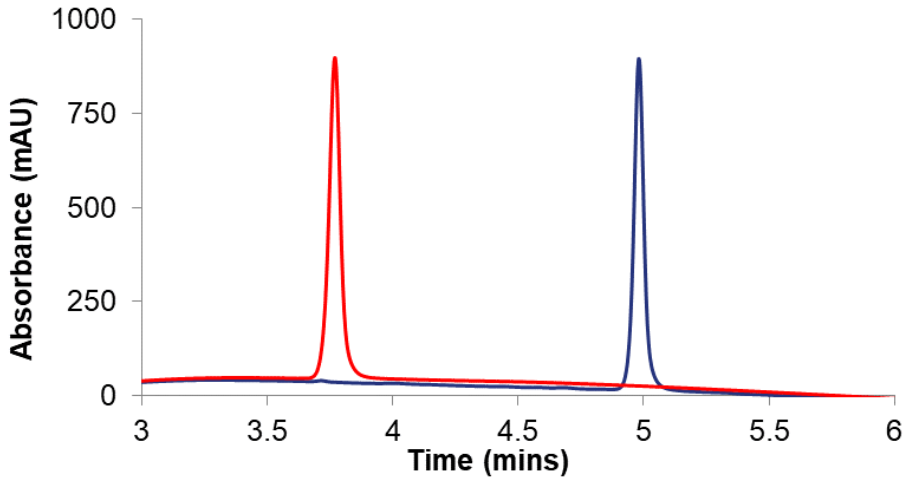
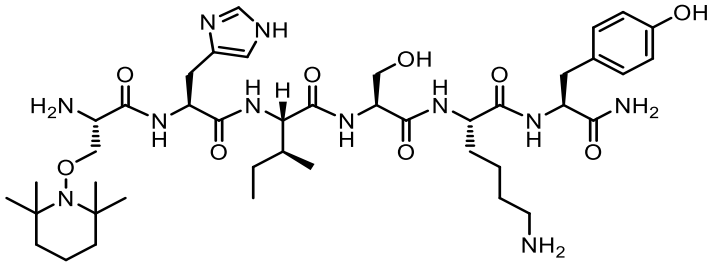


Figure 26 –HPLC trace comparison, showing the starting material H-A(TEMPO)AF-OMe (**11**, blue) and the reductively cleaved product from Entry 6 (**18**, red). Analytical gradient 10-70% B over 5 mins, 210 nm.

This was an interesting result as it shows a novel, simple and high yielding method to convert Sec into Ser. Furthermore, the ability to remove the functionalisation shows a potential method of delivering a conjugate to a certain location. The ability to trap and cleave using simple chemistry also shows the potential for this method to be applied to the isolation and purification of peptides and proteins.

To ensure the stability and cleavability of the conjugate was not affected by extra functionality, these reactions were performed on a larger TEMPO conjugated model, H-A(TEMPO)HISKY-NH_2 (**19**, Figure 27). The data from these reactions is shown in Table 4.



19

Figure 27 – H-A(TEMPO)HISKY-NH_2 (**19**) model used to prove conjugate stability and the ability to selectively cleave the conjugate.

Table 4 - Data from the stability studies of the TEMPO conjugate H-A(TEMPO)HISKY-NH_2 (**19**).

Entry	Conditions	Time/hr	% Degradation
1	pH 2	16	0
2	pH 9	16	0
3	5 eq GSH	16	0
4	50 ^{o}C	16	0
5	50 eq Zn/ 10% AcOH - 37 ^{o}C	16	100

The stability data again showed full conjugate stability to a range of conditions. The conditions in Entry 5 (Table 4) showed complete degradation of the conjugate, with a clear and clean shift in the HPLC trace being observed. The HPLC trace from this reaction compared with the starting material is shown in Figure 28.

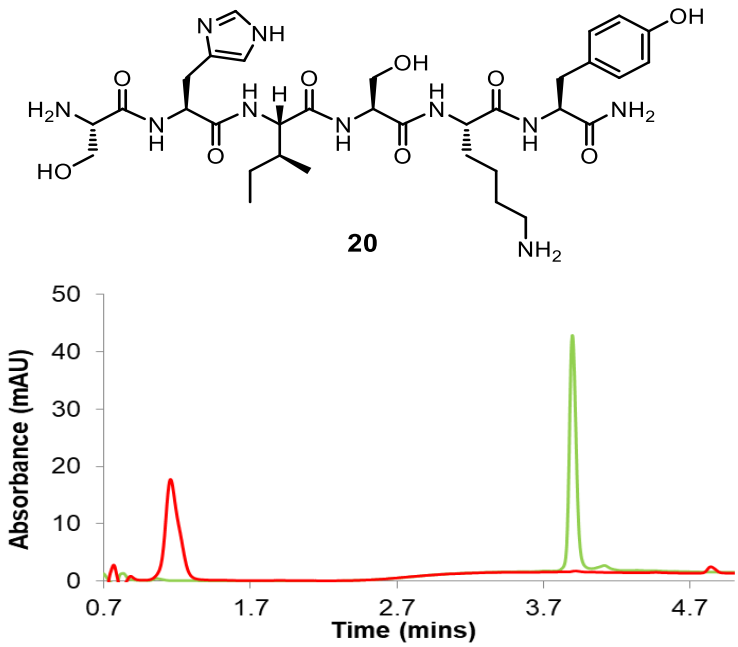
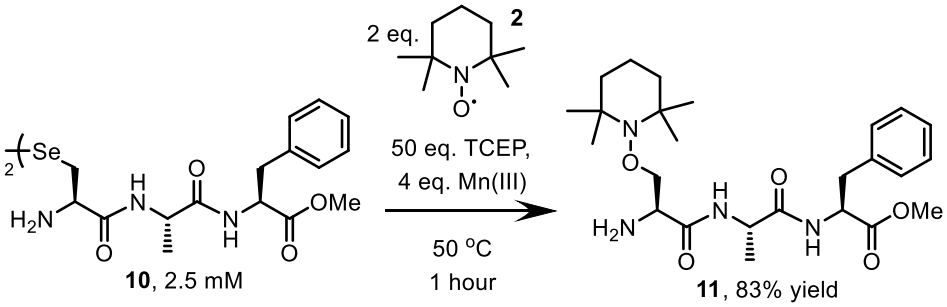


Figure 28 – Product (**20**) from reductive cleavage of starting TEMPO conjugate (**19**) and overlaid HPLC traces of the starting conjugate (**19**, green) and the resultant crude trace from Entry 5 (**20**, red). Analytical HPLC gradient 10-70% B over 5 mins, 210 nm.

After demonstrating the stability of the TEMPO conjugate (**19**), the next important factor to study was the tolerance of the method to different functional groups on the TEMPO moiety. Previously in this chapter the reaction was shown to tolerate propargyl and FITC groups. The aim of the following data was to further explore the suitability of this method to a range of biologically relevant groups. It was also important to perform NMR studies to ensure the assumed product was actually what was expected. These NMR studies were also applied to assess whether the method was affecting the stereocentre of the Sec residue.

Before going any further in this project, it was important to isolate and fully characterise the product to ensure it was as expected. The H-UAF-OMe (**10**) tripeptide model was subjected to conjugation with TEMPO using general protocol A. Briefly, to peptide dissolved in 20% DMSO in ligation buffer to a concentration of 5 mM was added a solution of TCEP (pH adjusted to 7, 50 eq), TEMPO (2 eq) and $Mn(OAc)_3$ (4 eq). The pH of the reaction mixture was checked to be 7-7.5 then the reaction mixture diluted to a final peptide concentration of 2.5 mM. The reaction was then heated to 50 ^{o}C and shaken for 1 hour. Once the starting material was shown by HPLC to be fully consumed, the reaction mixture was purified by semi-preparative HPLC, producing compound **11** with an 83% yield. The scheme for this reaction is shown in Scheme 46.



Scheme 46 - Reaction scheme showing the conjugation of H-UAF-OMe (**10**) tripeptide model with TEMPO (**2**) to give desired product (**11**).

The isolated product was analysed by ^{1}H-NMR and showed the expected peaks for the desired product, with no side products observed (NMR details in section 7.4, compound **11**). The ESI-MS analysis (section 7.4, compound **11**) showed only single addition of the TEMPO molecule to the peptide, with loss of the splitting pattern characteristic of Se molecules. This showed the loss of Se from Sec.

After proving that the desired product (**11**) was being produced, a range of TEMPO probes with varying biologically relevant functional groups were synthesised. The desire is for the developed method to be operationally simple and generally applied, so the TEMPO probes must also be easy to synthesise. The chosen probes are shown in Figure 29.

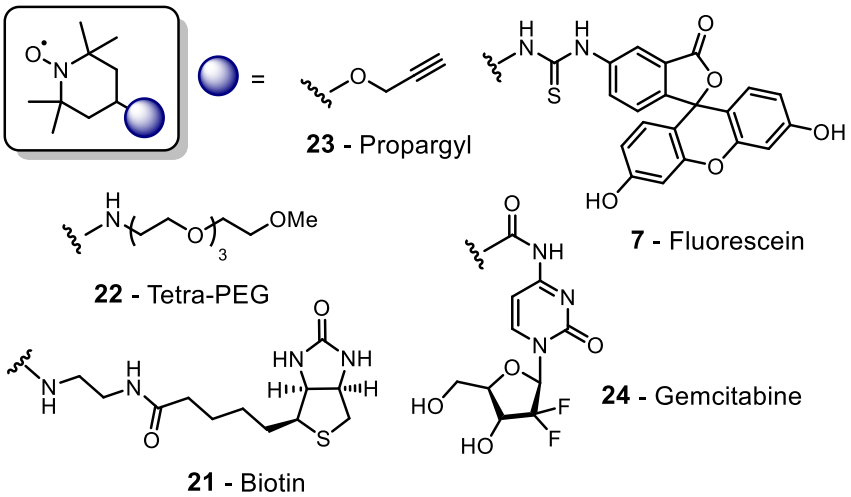
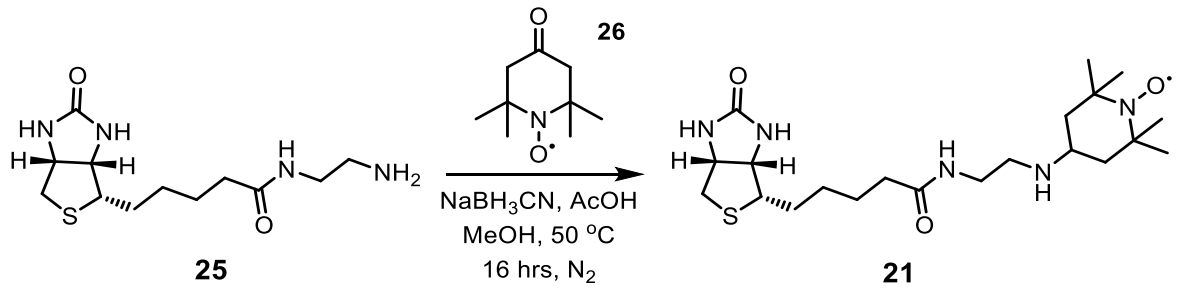


Figure 29 - A range of TEMPO probes synthesised showing varying functionality and biological importance.

The probes shown in Figure 29 were all synthesised from commercially available starting materials; either 4-amino, oxo, hydroxy or carboxy TEMPO. An example reaction scheme for the synthesis of biotin probe (**21**) is shown in Scheme 47. This was synthesised by a reductive

amination reaction between ethylenediamine modified biotin (**25**) and commercially available 4-oxoTEMPO (**26**), as shown in Scheme 47.

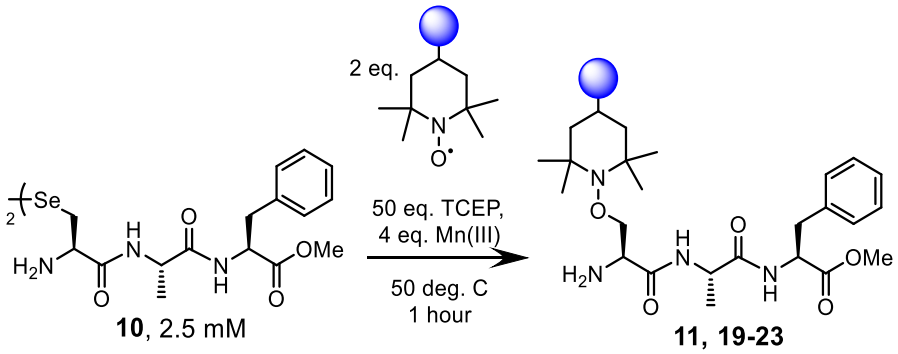


Scheme 47 - Reaction scheme for the synthesis of biotin-TEMPO probe.

All other probes were synthesised using simple, well-established chemistry. Tetra-EG probe **22** was also synthesised by reductive amination, with Tetra-EG-NH_2 and 4-oxoTEMPO (**26**). Propargyl probe **23** was synthesised by alkylating commercially available 4-hydroxyTEMPO with propargyl bromide. Fluorescein probe **7** was synthesised by reacting commercially available 4-aminoTEMPO with fluorescein isothiocyanate. Finally, gemcitabine probe **24** was synthesised by amidating commercially available 4-carboxyTEMPO with gemcitabine. The details of the syntheses performed can be found in section 7.3. The range of conditions used in the TEMPO probe syntheses show the stability of the TEMPO radical trap to a range of conditions, including acidic/basic pH, high temperatures and reducing conditions for extended periods. Other conditions used include mildly (Et_3N, DIPEA) and strongly basic (NaH), with a range of coupling reagents (DMTMM, EDC/HOBt), solvents (DMF, MeOH, THF) and additives. This range of conditions allows the potential synthesis of numerous biologically relevant probes. The probes are column and bench stable; they do not require any stringent storage conditions and are effective after over 6 months of storage.

The probes used in the following reactions to test the TEMPO conjugation method were chosen due to their biological relevance. The PEGylated (**22**) example was chosen because of its relevance in the pharmaceutical industry. The addition of a PEG group to peptide or protein-based drugs has been shown to increase the drug's longevity and activity. The hydrophilicity of PEG is also useful in allowing the use of hydrophobic groups or peptides.[129,130] The alkyne (**23**) probe was chosen to prove the method could be used to convert the Sec residue into a highly utilised bioorthogonal handle. The alkyne group is commonly used in copper-catalysed azide-alkyne click chemistry. A range of functionality, including fluorescent and PEG groups, and drug molecules have been attached using this chemistry.[131,132] The click reaction has also been used to label proteins in live cells, allowing cell protein charecterisation.[133] Gemcitabine (**24**) was chosen to represent the potential of this method for the synthesis of peptide or protein drug conjugates. Previous research has shown that

gemcitabine peptide conjugates have improved metabolic stability, safety and specificity.[134–136] Biotin (**21**) was chosen as a highly useful handle for protein purification. The biotin groups strongly bind to streptavidin coated magnetic beads. This allows the isolation of biotinylated protein, with non-biotinylated protein simply washed away. The biotinylated protein is then eluted by changing elution buffer and increasing temperature.[48,137,138] This purification method has also been applied to the isolation of the most reactive/abundant proteins from cell lysate.[137] The fluorescent (**7**) probe was chosen as peptide/protein fluorescent probe conjugates feature prominently in literature. They are attached using a range of methods and used for the visualisation and tracking of proteins in cell lines.[139–142] Such fluorophores have also been used to label the pharmaceutically prominent antibodies bevacizumab and trastuzumab. In these cases, the labelling enabled monitoring of the rate of antibody clearance from plasma.[143] Aside from the broad range of biological relevance, the chosen probes offer a variety of reactive functional groups, including phenols, alkynes, primary alcohols, and secondary alcohols to test the bioconjugation method compatibility.



Scheme 48 - Reaction scheme for the conjugation of H-UAF-OMe (10) with a variety of TEMPO-based probes.

After synthesising the chosen probes, as outlined in section 7.3, the next step was to attempt bioconjugation onto the tripeptide H-UAF-OMe (**10**). The general scheme for this reaction is shown in Scheme 48. For the following bioconjugations general protocol A was utilised. These reactions were performed using the probes shown in Figure 29, with the resulting isolated yields shown below in Table 5.

Table 5 - Table summarising the yields from the bioconjugation (Scheme 48) of H-UAF-OMe (**10**) with a range of TEMPO probes.

Entry	Product	Probe	Isolated Yield
1	**11**	TEMPO (**2**)	83%
2	**27**	Tetra-EG (**22**)	85%
3	**28**	Propargyl (**23**)	82%
4	**29**	Fluorescein (**7**)	80%
5	**30**	Biotin (**21**)	79%
6	**31**	Gemcitabine (**24**)	67%

The data in Table 5 shows that the isolated yields of the bioconjugation reaction were relatively unaffected by the functionality installed on the TEMPO molecule. This is gratifying as it demonstrates that this method can be used to install a wide range of functionality onto a peptide. Shown in Figure 30 is an example of the crude reaction mixture of H-UAF-OMe (**10**) with propargyl TEMPO (**23**) compared with the starting material. The crude trace shows clean conversion, as was seen when using the standard TEMPO probe.

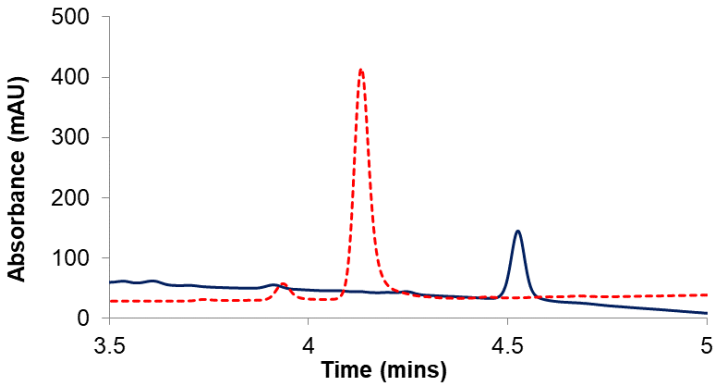


Figure 30 - A comparison showing the H-UAF-OMe (**10**) starting material (red dash) and the crude propargyl TEMPO conjugated product (**28**) after 1 hour (blue).

2.4 NMR studies to assess product stereochemistry

After the success of the above reactions, an important aspect of the reaction to investigate was whether the stereochemistry of the conjugate amino acid alpha centre was affected. For this, the D-Sec was synthesised by Dr. Frances Smith, incorporated into the H-D-UAF-OMe (**10b**) model, and the TEMPO conjugation reaction performed using general protocol A (this work was performed by Dr. Frances Smith). The two different isomers produced from these reactions (**10** and **10b**) are shown in Figure 31. The NMR spectra from the two products were stacked for comparison, as shown in Figure 32.

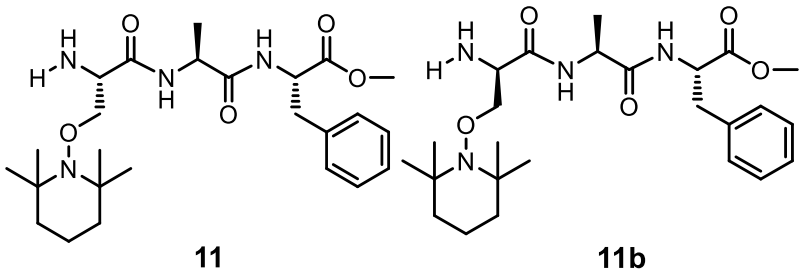


Figure 31 - Structures of the products from H-UAF-OMe conjugation with TEMPO using L-Sec (**11**, left) and D-Sec (**11b**, right).

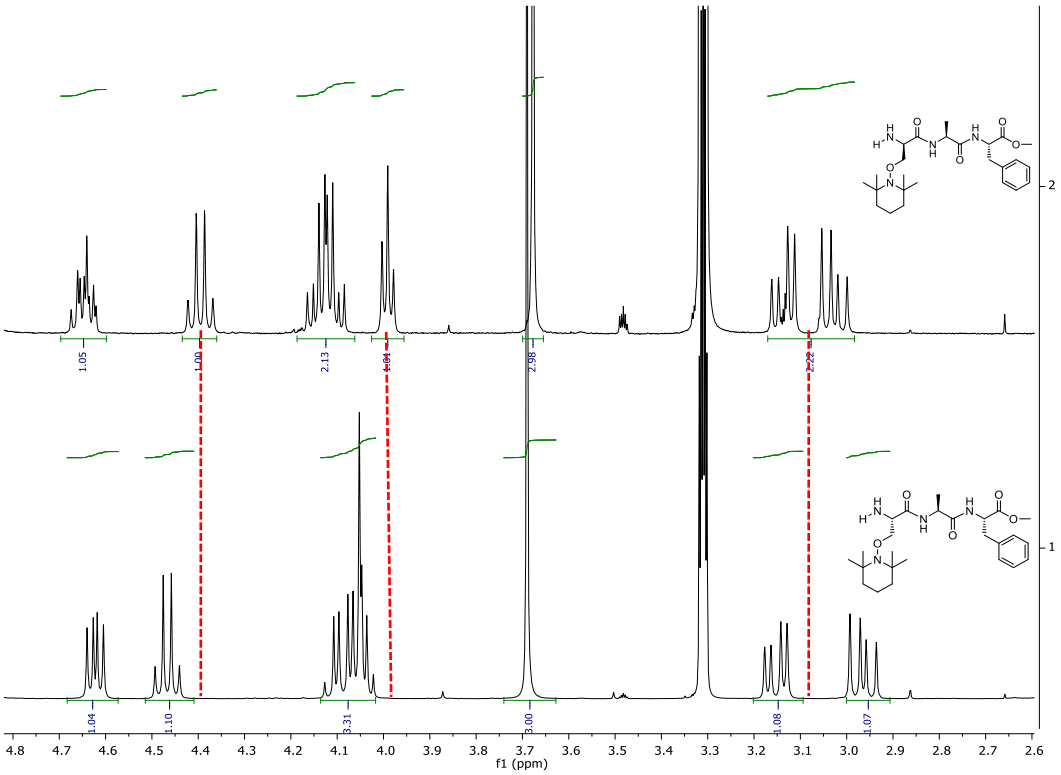


Figure 32 - NMR comparison of D-Sec (**11b**, top) and L-Sec (**11**, bottom) TEMPO conjugated products showing key differences in the NMRs of each product. The absence of highlighted peaks in the L product compared to the D shows racemization does not occur during conjugation.

This study was performed to explore the integrity of the alpha centre of the Sec residue during conjugation. If racemisation was occurring during the reaction, the NMR spectra would be identical as both reactions would produce diastereomers. This would not be ideal as the aim for this reaction is to retain stereochemistry, giving it an advantage over other currently available methods.[85,144] The spectra in Figure 32 clearly show a difference in the peaks present, as shown by the dotted red lines. The difference in peaks present shows that each reaction is producing a different product (**11** or **11b**) from the other. Another clear difference between the two spectra is the larger split in the doublet of doublets peak at 3.15 and 2.98. This also supports the idea that the two products (**11** or **11b**) are clearly different. Based on the NMR data it is clear that two different products are formed from the conjugation of D- (**10b**) and L-Sec (**10**), showing a retention in the stereochemistry. This result was very important for the validity of the developed method.

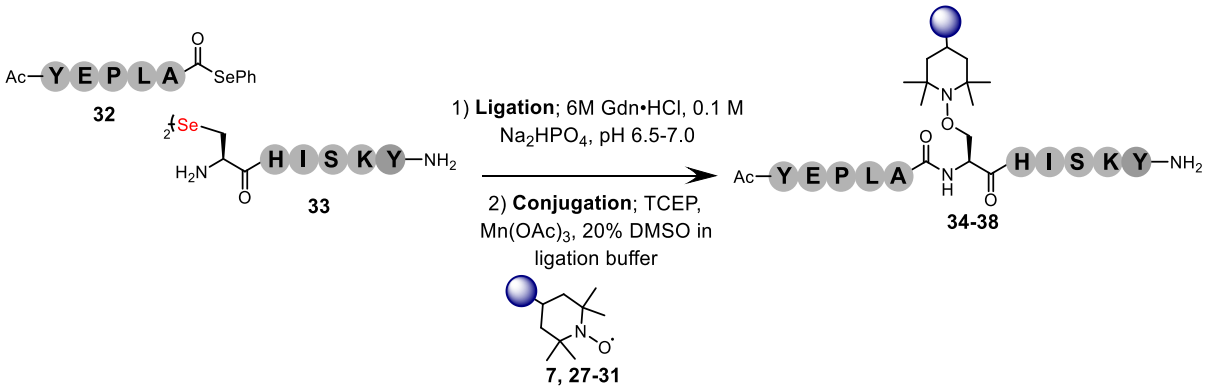
At this stage the method had been fully optimised, tested on a range of substrates, the products fully characterised by NMR, and the retention of stereochemistry confirmed. The next development shows the method being applied to a one-pot ligation and conjugation reaction. This would be very impactful for the synthesis and modification of small proteins as current methods are limited.

2.5 Application to one-pot ligation-bioconjugation

The aim of this chapter was to apply the conjugation to larger model peptides generated using the DSL technique. Ideally, the ligation and conjugation would be suitable to be performed as a one-pot two-step reaction. This would increase time efficiency and yield as only one preparative purification step would be required. The second aim of this chapter was to increase the complexity of the peptide sequences used. This would provide insights as to whether any particular amino acid functionality would cause issues and reduce yield from side-reactions occurring.

For this chapter two model peptides were chosen as the two peptide fragments for ligation. The N-terminal fragment was Ac-YEPLA-SePh (**32**) and the C-terminal fragment H-UHISKY-NH_2 (**33**). The peptide fragments were synthesised as described in section 7.1. These sequences were chosen as they contain the majority of the functionality displayed across the proteome. This means that if the conjugation reactions were successful, then there should be no limit regarding which peptides or proteins could be modified using this chemistry. So far, the only Sec residue that had been modified was an N-terminal Sec residue. These reactions allowed for the exploration of conjugation to an internal Sec residue. The internalisation of the conjugation site increases the steric bulk that could prove problematic. This issue had the

potential to significantly reduce reaction rate and product yield. The general scheme for these reactions is shown in Scheme 49.



Scheme 49 - General reaction scheme for the one-pot ligation and conjugation of a model peptide sequence.

The one-pot ligation and conjugation reactions were performed using the following method – Standard DSL was performed using 1.05 eq. of the selenoester fragment (**32**). Ligations were performed at a concentration of 5 mM with respect to the N-terminal selenopeptide monomer (**33**). Ligations were performed for 10 minutes before extracting the precipitated diphenyl diselenide (DPDS) using hexane. Conjugation was immediately performed by adding TCEP (50 eq.), TEMPO probe (2 eq.) and $Mn(OAc)_3$ (4 eq.). The reactions were diluted so the final peptide concentration was 2.5 mM and the final solvent composition was 20% DMSO in LB. The pH of the reaction was adjusted to 7-7.5 using 10 M NaOH and the reaction shaken at 50 ^{o}C for one hour. The reactions were immediately purified using preparative RP-HPLC. The results from these reactions are summarised in Figure 33, with the percentages reported being the isolated yields. An example HPLC trace showing the crude reaction mixture and the pure product when using propargyl TEMPO (**23**) is shown in Figure 34.

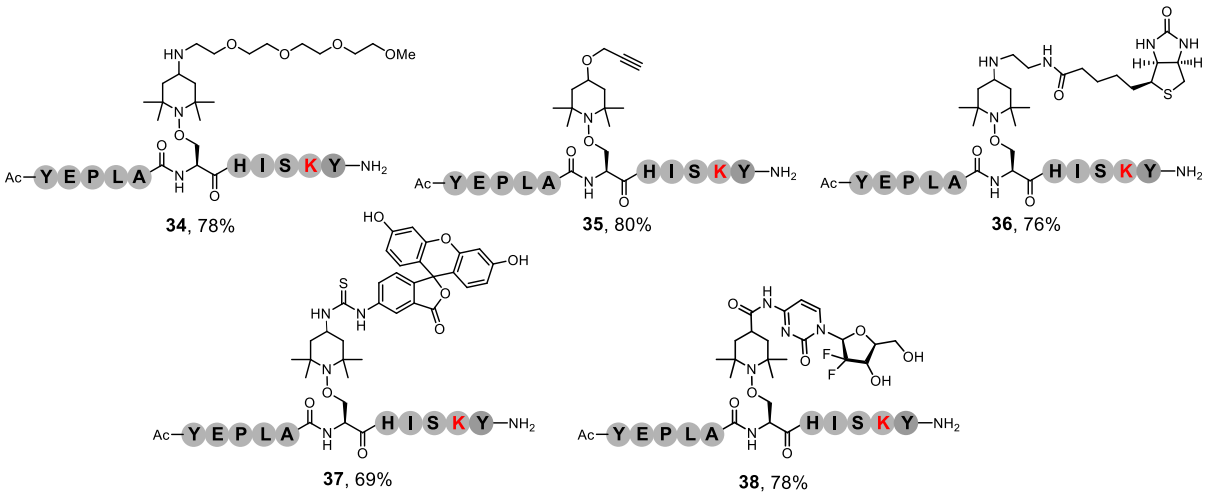


Figure 33 - Summary of the one-pot ligation and conjugation reactions using Ac-YEPLA-SePh (**32**) and H-UHISKY-NH_2 (**34**) fragments. The yields reported are isolated yields based on the amount of starting H-UHISKY-NH_2 (**33**).

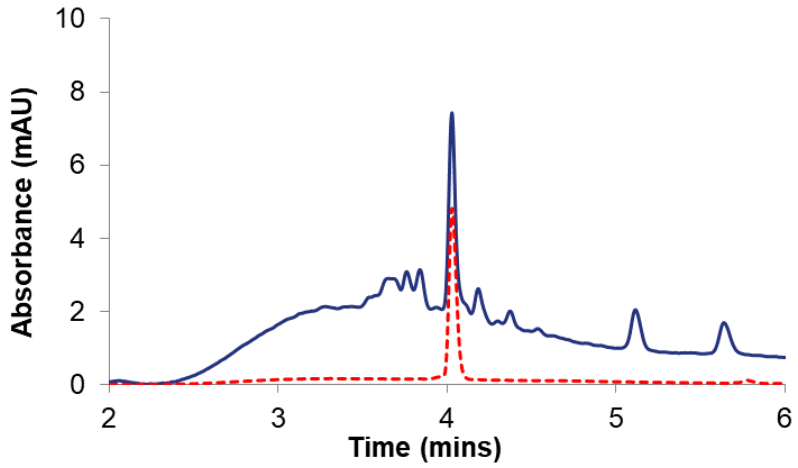
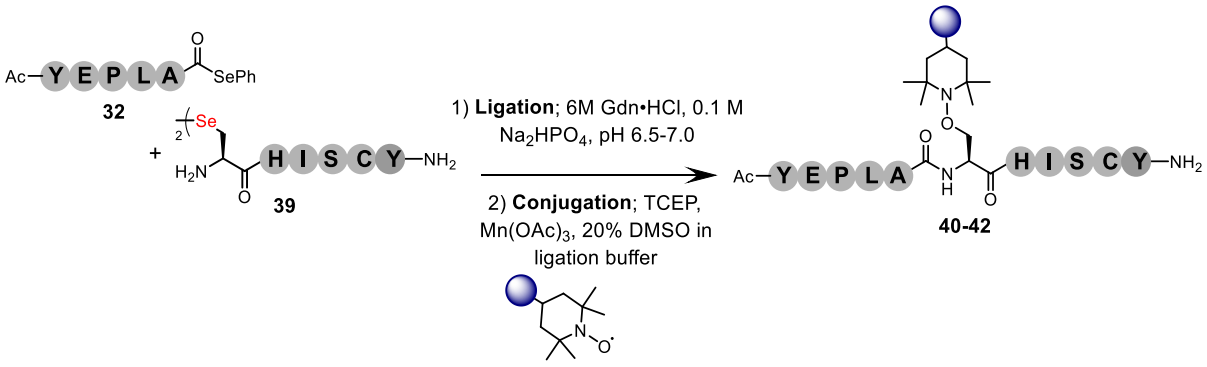


Figure 34 - Example crude HPLC trace from the one-pot ligation and conjugation using propargyl TEMPO (**23**, blue) and the pure product overlaid (**35**, red dash). Analytical HPLC gradient 10-100% B over 5 mins, 210 nm.

The method was successfully applied, allowing the ligation and subsequent conjugation in a one-pot procedure in excellent yields. The crude HPLC trace in Figure 34 shows the reaction is relatively clean, with minimal impurities present. This made the purification simple, achieving completely pure product from a single purification run. This method is very useful as it provides an operationally simple method to further functionalise the Sec residue required for DSL. This greatly expands the current post-ligation reactions available to peptide chemists.



Scheme 50 - Reaction scheme for the one-pot ligation and conjugation of Ac-YEPLA-SePh (**32**) and H-UHISCY-NH^2 (**39**).

Next, the C-terminal fragment was adjusted, replacing the Lys residue for Cys, H-UHISCY-NH_2 (**39**) (Scheme 50). This was an important study, as it will show that the reaction can be successfully applied to conjugate an internal Sec residue over an internal Cys residue. As shown previously, the Mn(III) accelerated conditions result in radical-induced desulfurisation of Cys residues. Therefore, mild conditions were explored. The ligation conditions were the same as previously described, but during conjugation Mn(III) was omitted, instead the TEMPO equivalents were increased to 5 and reaction times extended to 4 hours (General protocol C, details in section 7.2). The conjugation was performed on selected probes to prove the reaction was possible and that the yield was not significantly affected. The results from these reactions is shown in Figure 35, where the reported yields are isolated yields.

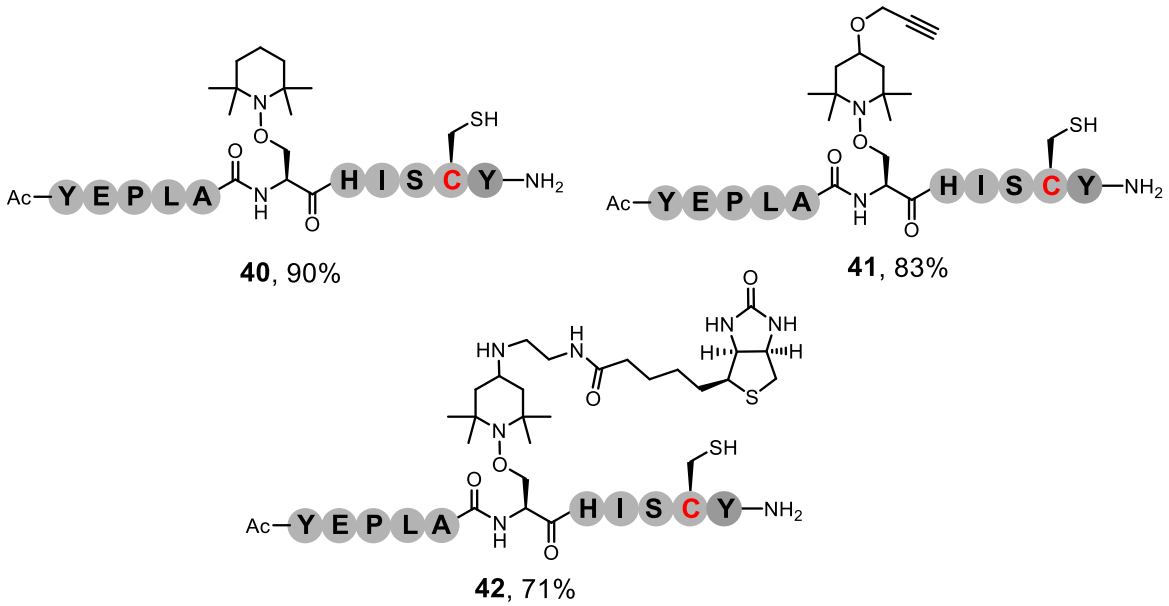


Figure 35 - Isolated yields for the one-pot ligation and conjugation of Ac-YEPLA-SePh (**32**) and H-UHISCY-NH_2 (**39**) with selected TEMPO probes.

Pleasingly, the results in Figure 35 show that the yields of reactions with the H-UHISCY-NH_2 (**39**) peptide are comparable to those with the H-UHISKY-NH_2 (**33**) peptide. The crude and pure HPLC traces for the propargyl TEMPO model (**41**) are shown in Figure 36. These traces show similar results to those seen with the propargylTEMPO modified H-UHISKY-NH_2 (**35**) model. The crude trace shows relatively clean conversion to desired product. The pure trace shows that the desired product was obtained in a high purity after a single purification. These results show that the developed method is completely chemoselective to Sec residues.

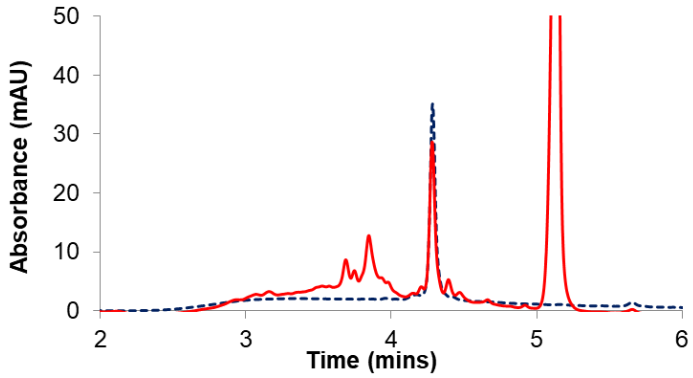
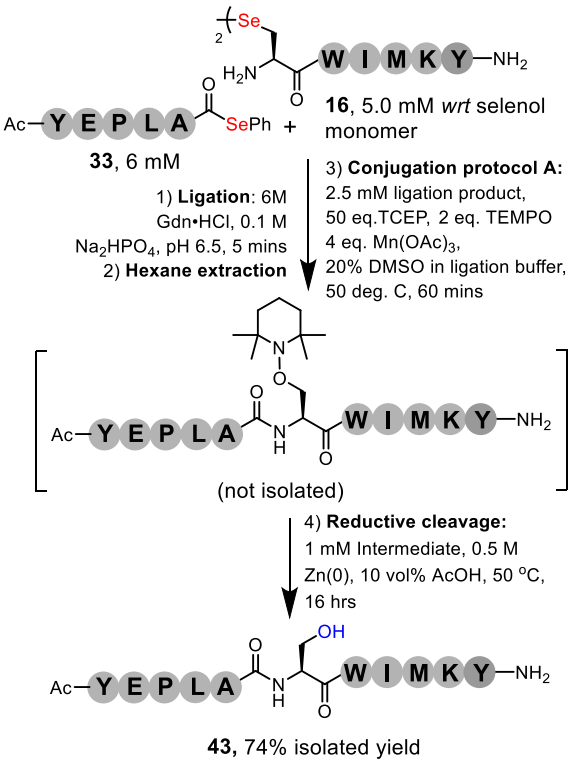


Figure 36 - A comparison between the crude HPLC trace (red) of one-pot ligation and conjugation of Ac-YEPLA-SePh (**32**) and H-UHISCY-NH_2 (**39**) with propargyl TEMPO (**23**) and the purified product (**41**, blue dash). Analytical HPLC gradient 10-100% B over 5 mins, 210 nm. Peak at 5.1 minutes is propargyl TEMPO.

The results from the one-pot ligation and conjugation reactions so far have large implications. They show that the developed method can be applied to the synthesis of small proteins using the DSL method, followed by immediate functionalisation to install a wide range of biologically relevant groups.

As the one-pot ligation and conjugation method was successful, combining this with the reductive cleavage conditions was trialled (Scheme 51). This would provide a complementary method to the oxidative deselenisation method already published, allowing conversion of Sec to Ser post-ligation.[122] Briefly, this method uses 50 eq. of TCEP and oxone at a low peptide concentration (0.1 mM) to afford the oxidatively deselenised product. The presence of oxone has been shown to cause Met oxidation, so the method developed in this chapter would provide an alternative for oxidation sensitive peptides. The initial trial, using 50 eq. Zn(0) was unsuccessful. Simply by increasing the amount of zinc in the reaction mixture the desired one-pot ligation-conjugation-reductive cleavage reaction was achieved in a 74% isolated yield (**43**). This yield was comparable to the already established oxidative deselenisation method, where the desired Ser peptide was isolated in a 68% yield using the same model peptides.

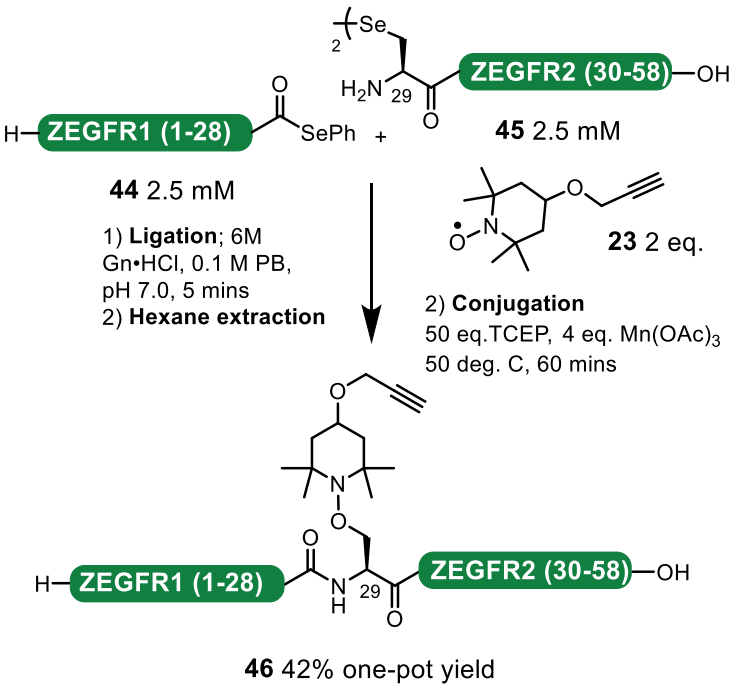


Scheme 51 - Reaction conditions for the one-pot ligation-conjugation-reductive cleavage of Ac-YEPLA-SePh (**32**) and H-UWIMKY-NH2 (**16**).

2.6 Application of one-pot ligation-conjugation to a 58 mer peptide

As the overall aim for this method is to apply it to biologically relevant peptides and proteins, for use in research and the pharmaceutical industry, a relevant small protein model was chosen to demonstrate our technique on a more complex system. The small protein model chosen was ZEGFR-1907,[145] a sequence derived from immunoglobin-binding protein A.

The fragments used for the ligation and conjugation reaction were synthesised using microwave assisted automated peptide synthesis. For ZEGFR1 (**44**), the peptide was removed from the 2-chlorotrityl chloride resin using mild conditions (30% HFIP in DCM), leaving the side-chain protecting groups intact. The C-terminal selenoester was then synthesised by addition of diphenyl diselenide (DPDS) and tributylphosphine. For ZEGFR2 (**45**), the peptide was synthesised on 2-chlorotrityl chloride resin and Boc-Sec-OH coupled as described in the experimental section. Both peptides were purified by preparative reversed-phase HPLC. The ligation-conjugation reaction was performed using the general one-pot ligation and conjugation protocol as described in the experimental section 7.2. As Cys was not present in either fragment of this protein, the Mn(III) accelerated conditions were applied. The scheme for this reaction is shown in Scheme 52.



46 42% one-pot yield

Scheme 52 - Reaction scheme for the one-pot ligation and conjugation of the two fragments of the ZEGFR protein.

Pleasingly, the one-pot ligation and conjugation of the ZEGFR protein was successful. The desired product (**46**) was isolated in a 42% yield based on the starting selenopeptide. The crude and pure HPLC traces are shown in Figure 37. The product was obtained with a purity of > 90% after a single purification. The crude trace shows relatively clean conversion to the desired product. As this yield was lower than seen in the previous one-pot reactions, the cause to this was further explored. The reaction was performed as a two-pot reaction, isolating after the ligation to ascertain whether this was causing problems.

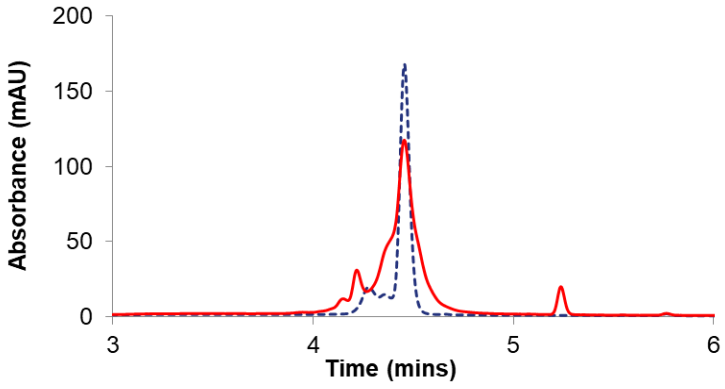


Figure 37 - Comparison of the crude trace from the one-pot ligation and conjugation of the ZEGFR protein with propargyl TEMPO (**46**, red) and the pure conjugate (**46**, blue dash). Analytical gradient 10-100% B over 5 mins, 210 nm.

The ligation was performed using the standard DSL procedure, using 1.2 eq. of the selenoester fragment. The yield isolated from this reaction was only 32%, which suggests that the ligation step for this sequence is responsible for the low one-pot yield (**47**). This product was taken forward to the conjugation step, and, gratifyingly, the conjugated product was obtained in an 88% isolated yield (**46b**). This confirms that the lower-than-expected yield was caused by the ligation step, rather than the conjugation step. Nonetheless, synthesising a small, functionalised protein in a 42% isolated yield can still be considered an excellent result.

2.7 Chapter Conclusion

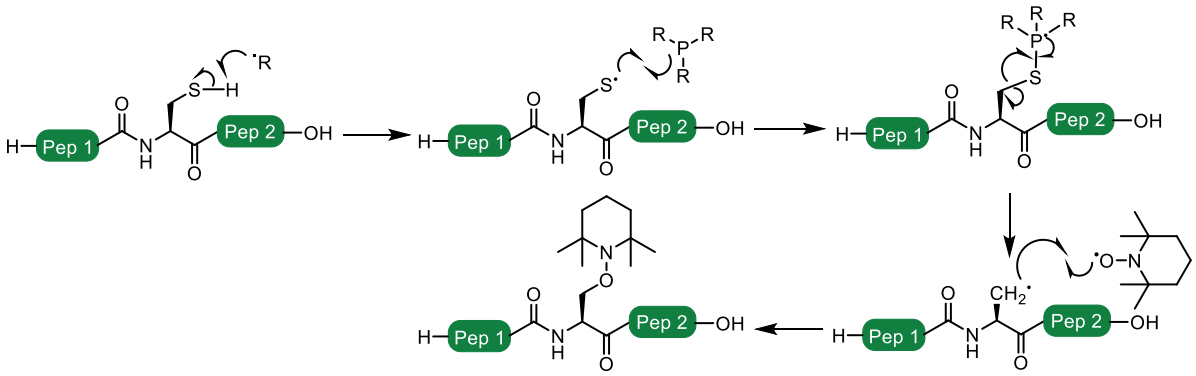
To summarise this chapter, a novel method to selectively modify Sec residues has been developed. This method uses low equivalents of the conjugate and is performed in aqueous solution at neutral pH. The reaction itself is rapid, reaching completion in under one hour, and high yielding. Furthermore, this reaction was applied to the development of a one-pot procedure involving ligation of the peptide fragments, followed by further functionalisation at the Sec ligation site using the described developments. The technique was further applied to a small protein model to yield a functionalised biologically relevant small protein in a reasonable yield. This protein model and previous peptide models show that the reaction is fully chemoselective, with reaction occurring at Sec alone, with no racemisation at the target residue. This method shows a great deal of promise in the desired application of diagnostics and in the synthesis of peptide/protein drug conjugates.

3. Cysteine-Selective Bioconjugation via Interception of Free-Radical-Induced Desulfurisation with TEMPO traps

3.1 Chapter Abstract

Methods that enable the site-selective labelling of peptides and proteins via the canonical amino acids, to form stable conjugates using mild conditions and low equivalents of the conjugate, are in high demand. Many of the currently available methods fail when comparing the procedure to the above criteria. As previously discussed in Chapter 1, many methods have been developed to target Cys residues due to the low natural abundance of Cys across the proteome and high nucleophilicity of the thiol group. As Cys is a naturally occurring amino acid, it can be installed into a protein during expression through side-specific mutagenesis. This simple and usually high yielding procedure enables the exchange of the native residue for another canonical residue.[128] Once the protein with the desired mutant is produced, it can then be subjected to conjugation.

This chapter describes the application of the bioconjugation method developed in Chapter 2 to the interception of a peptide alanyl radical formed via free-radical induced dechalcogenation of Cys. The proposed mechanism for this method of modification is shown in Scheme 53.



Scheme 53 - Proposed mechanism for the TEMPO bioconjugation of Cys residues.

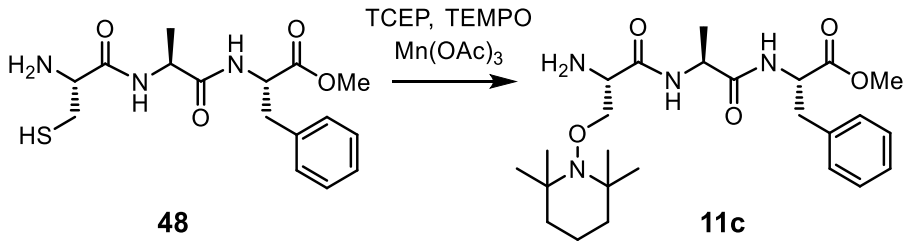
This mechanism is very similar to the deselenisation of Sec; the main difference is the requirement of a radical initiator to abstract the H-atom from the thiol of the Cys, forming a thiyl radical. This radical can then donate into a phosphine and follows the same mechanism as deselenisation of Sec. The thiyl radical can be generated through a range of techniques, including H-atom abstraction using a radical initiator, such as VA-044,[32] the application of UV light (254 nm).[146,147] Alternative methods involve the single-electron oxidation of the thiol, forming a radical cation, which is readily deprotonated to form the thiyl radical. This has been

achieved using both photoredox catalysis[148–150] and readily-reducible metal ions.[128] All of these methods require the presence of a tertiary phosphine or phosphite to induce desulfurisation.

Successful exploration and optimisation of this reaction again enables us to develop a one-pot ligation and conjugation using an N-terminal Cys containing peptide. Furthermore, the method is applied to the site-selective installation of a desired group using a Cys mutant of a recombinantly produced protein to demonstrate the application to pharmaceuticals and general protein chemistry.

3.2 Initial Optimisation of Cysteine-Selective Bioconjugation

The first stage of this investigation involved applying the conditions optimised in the previous chapter (general protocol A) to model peptide H-CAF-OMe (**48**). The general scheme for this reaction is shown in Scheme 54.



Scheme 54 - The general reaction scheme for the conjugation of H-CAF-OMe (**48**) with TEMPO (**2**) to form conjugated product (**11c**).

Using general protocol A resulted in only 70% conversion to the desired product after 2 hours. There remained 15% starting material and 15% conversion to the undesired Ala product at this time point. The HPLC trace from this reaction is shown in Figure 38. When using Sec, the reaction had reached completion by 1 hour, the Cys-model had only 43% conversion to product by this point. The slower reaction rate can be attributed to the stronger C-S bond relative to C-Se. This will result in the homolysis step taking longer, increasing the overall reaction time.

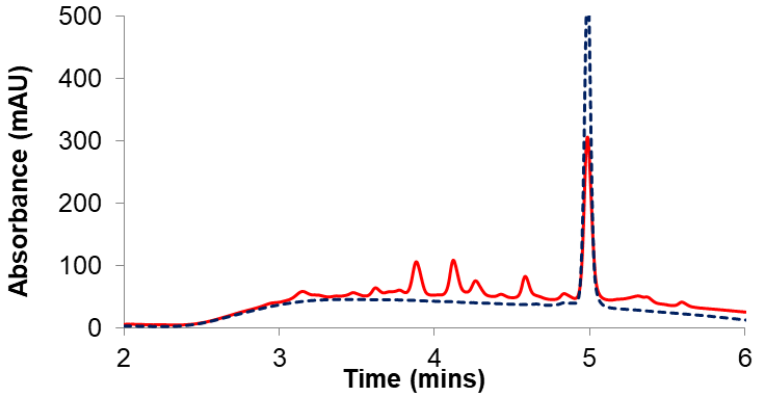


Figure 38 - Crude HPLC trace from the reaction of H-CAF-OMe (**48**) with TEMPO using general protocol A (red). The pure TEMPO conjugated product (**11c**) is overlaid (blue dash). Analytical gradient 10-70% B over 5 mins, 210 nm.

After this initial promising result, alternative reaction conditions were trialled. As the TEMPO probe would be carrying the potentially expensive functionality, the desire was to keep the equivalents of this as low as possible. As both TCEP and $Mn(OAc)_3$ are both relatively cheap and readily available, these were increased in concentration. The reaction was also diluted to try to prevent alanyl radical quenching from the thiol present. The results from these trials are shown in Figure 39.

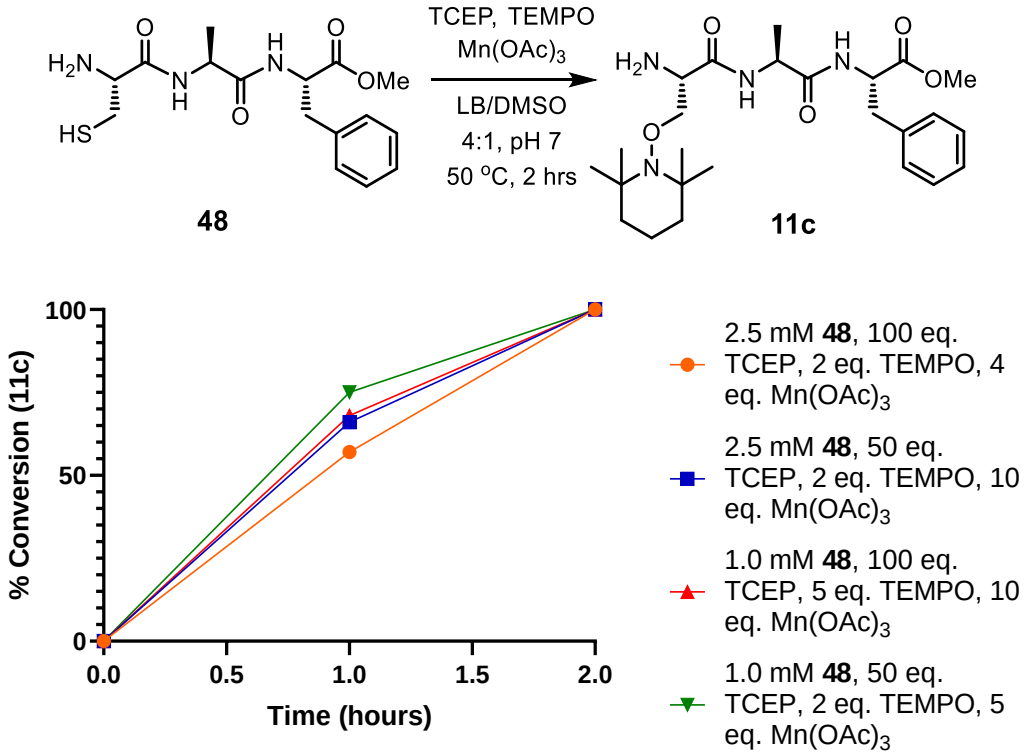


Figure 39 – Summary of the optimisation reactions performed to increase the conversion observed of **48** to **11c**. All reactions were performed for 2 hours, with time-point monitoring at one and two hours.

As Figure 39 shows, the reaction was successfully optimised, leading to an excellent conversion of starting material (**48**) to product (**11c**). Importantly, no Ala by-product (**12**) was detected when applying the conditions as shown in Figure 39. This is important as when moving to a larger peptide or protein model, the Ala product would be difficult to isolate from the desired product. A reaction time of 2 hours can still be considered fast for a bioconjugation reaction. A higher dilution is operationally useful, especially when performing bioconjugation reactions on proteins. Due to the size of the protein the volume of liquid used will be low, so a dilute reaction allows for a more appropriate reaction volume. The results in Figure 39 show that the optimal conditions to be used in future reactions were 1 mM peptide, 50 eq. TCEP, 2 eq. TEMPO and 5 eq. $Mn(OAc)_3$ at 50 ^oC for 2 hours in 4:1 LB:DMSO pH 7 (general protocol B, section 7.2). These conditions result in high conversion to desired product (**11c**), whilst only requiring 2 eq. of TEMPO conjugate, which is ideal for the TEMPO group to bare precious/expensive functionality.

To ensure the correct product was being synthesised, the product from H-UAF-OMe TEMPO conjugation (**11**) was compared to the product from H-CAF-OMe TEMPO conjugation (**11c**). As these products should be identical, the NMR data should also be identical. As shown previously, if epimerisation were occurring then a clear shift in NMR peaks would also be seen. The comparison of these spectra is shown in Figure 40.

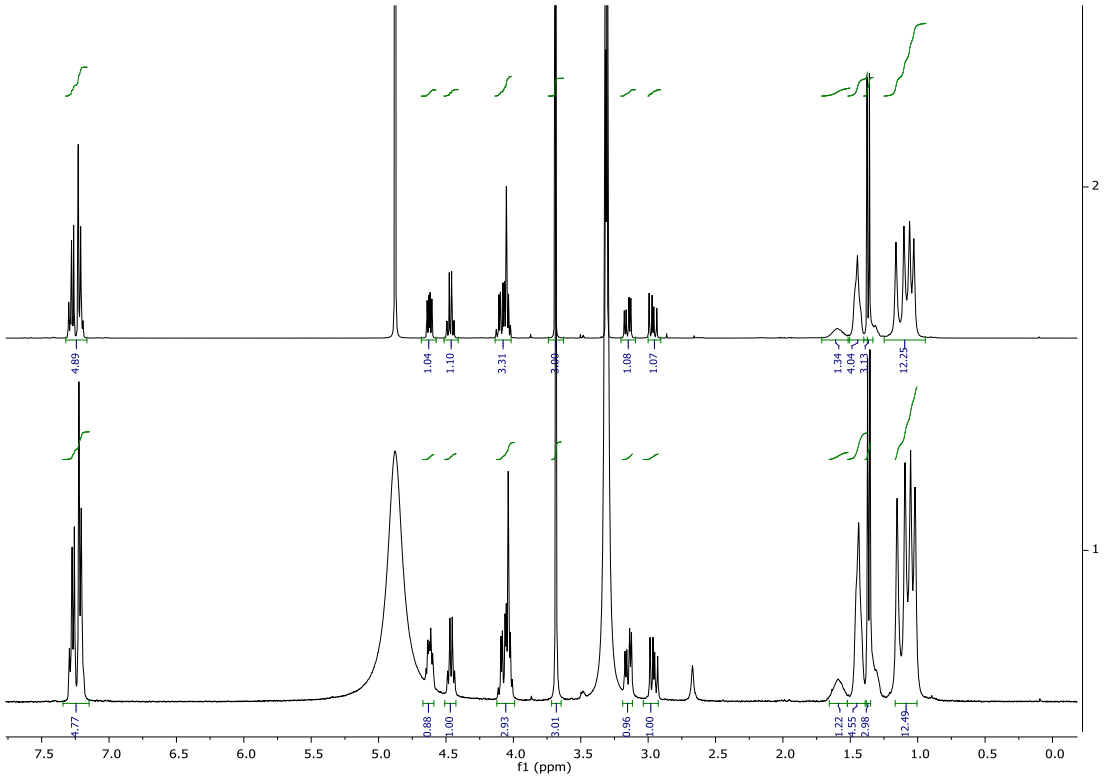


Figure 40 - Comparison of ^{1}H-NMR of the TEMPO conjugated products using H-UAF-OMe (**11**, top) and H-CAF-OMe (**11c**, bottom).

The same splitting patterns are seen in the peaks relating to the α and β protons of the Ala and Phe amino acids. The overlapping α and β protons at 4.05 ppm look identical, which is important as this corresponds to the conjugated residue. The MS and HPLC data also proved that the same product was synthesised. The overlapping HPLC data is shown in Figure 41. This, combined with the NMR data, was sufficient proof that the expected product was made.

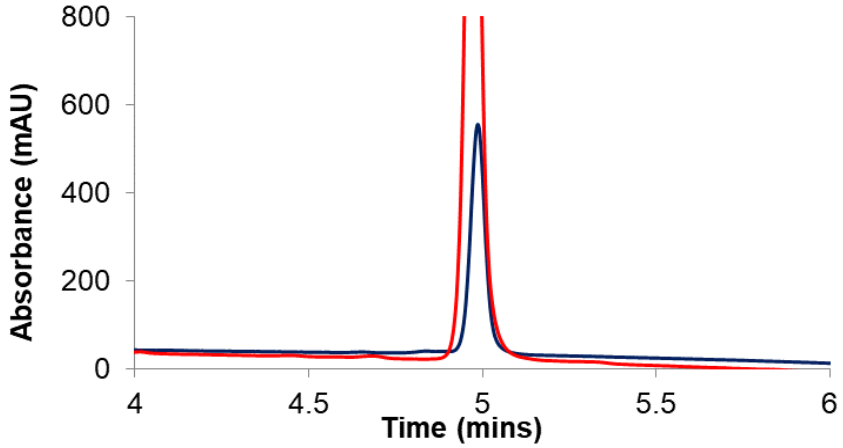
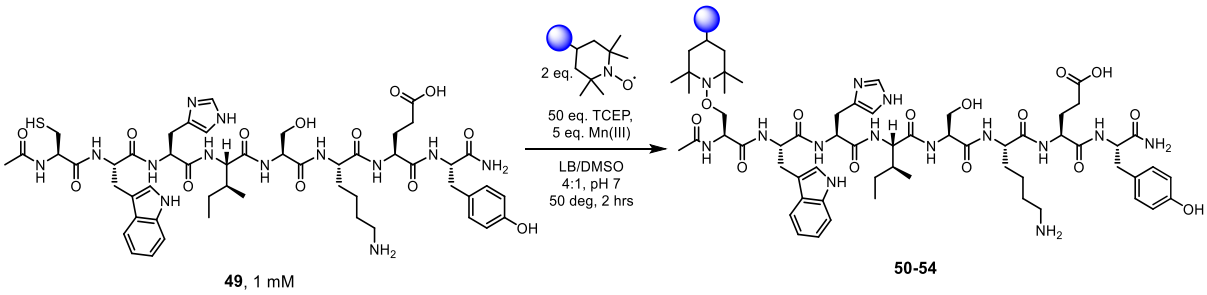


Figure 41 - Overlaid pure HPLC traces for the TEMPO conjugated products using H-UAF-OMe (**11**, red) and H-CAF-OMe (**11c**, blue). Analytical gradient 10-70% B over 5 mins, 210 nm.

3.3 Reaction of functionalised probes with model peptide

The next step was to assess whether adding extra functionality would affect the reaction. This was not previously an issue with Sec, however, since the conditions have been altered, this was an important aspect to assess. The compatibility with the other TEMPO probes used previously was also to be assessed. The peptide sequence used was Ac-CWHISKEY-NH_2 (**49**), and the general reaction outline is shown in Scheme 55.



Scheme 55 - General reaction scheme for the bioconjugation of Ac-CWHISKEY-NH2 (41) with TEMPO based probes.

These reactions were performed using general protocol B, as specified in the experimental section 7.2. Briefly, the peptide was dissolved in 20% DMSO in ligation buffer to a concentration of 2 mM. TCEP (pH adjusted to 7, 50 eq), TEMPO (2 eq) and $Mn(OAc)_3$ (5 eq) were then added. The pH was checked to be 7-7.5 then the reaction mixture diluted to a final

peptide concentration of 1.0 mM. The reaction was heated to 50 °C and shaken for 2 hours. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC. The isolated yields from these reactions are shown in Table 6.

Table 6 - Isolated yields for the bioconjugation of Ac-CWHISKEY-NH_2 with a range of TEMPO based probes.

Entry	Product	Probe	Isolated Yield
1	**50**	Tetra-EG (**22**)	85%
2	**51**	Propargyl (**23**)	82%
3	**52**	Fluorescein (**7**)	71%
4	**53**	Biotin (**21**)	73%
5	**54**	Gemcitabine (**24**)	80%

The isolated yields outlined in Table 6 show that the reaction was as efficient on this model (**49**) as it was with the H-UAF-OMe (**10**) model. The reaction proceeded efficiently and with a similarly clean crude reaction mixture, showing mostly desired product, as shown in Figure 42.

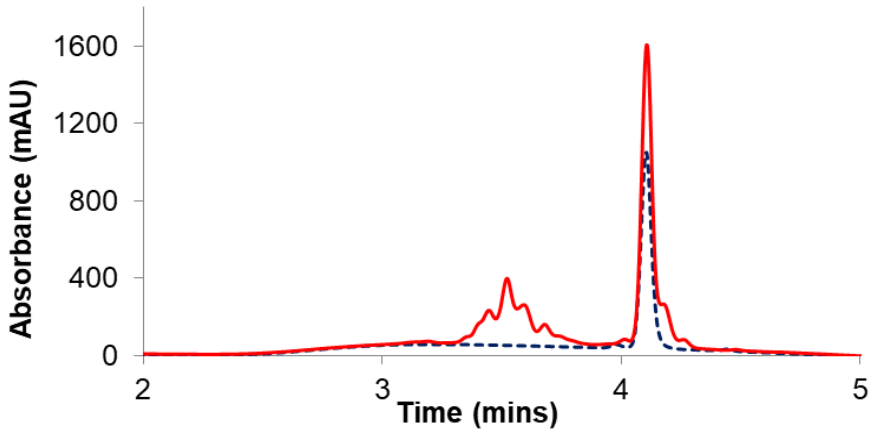
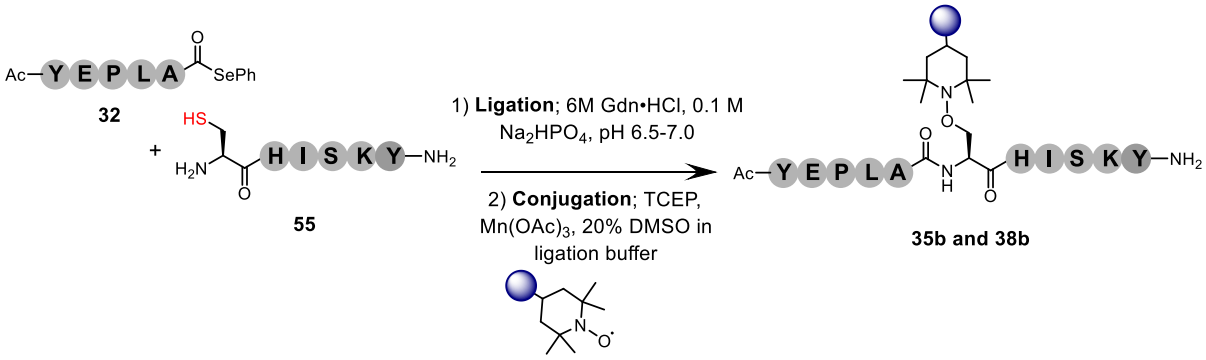


Figure 42 - Analytical HPLC traces for the bioconjugation of Ac-CWHISKEY-NH_2 (**49**) with propargyl TEMPO (**23**). Crude (red) and pure (**51**, blue dash).

After proving that the reaction can be applied to Cys peptides with a range of functionality, the next step was to see whether the one-pot conditions described in Chapter 2 could be adapted and applied. Although slower, NCL still features prominently in the literature as a method for chemically synthesising small proteins.

3.4 Application to ligation-bioconjugation



Scheme 56 – General reaction scheme for the 2-step ligation and conjugation between Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**) to produce TEMPO conjugated products (**35b** and **38b**).

A one-pot ligation and conjugation reaction was performed using the peptide fragments Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**) (Scheme 56). The ligation was performed in the same manner as for Sec, with reactions being left for 10 minutes. The ligation solution was then subjected to the conditions previously used for the modification of Cys containing peptides. The ligation itself proved slightly more problematic than the Sec ligations. The HPLC trace of the crude ligation showed a number of peaks, as shown in Figure 43.

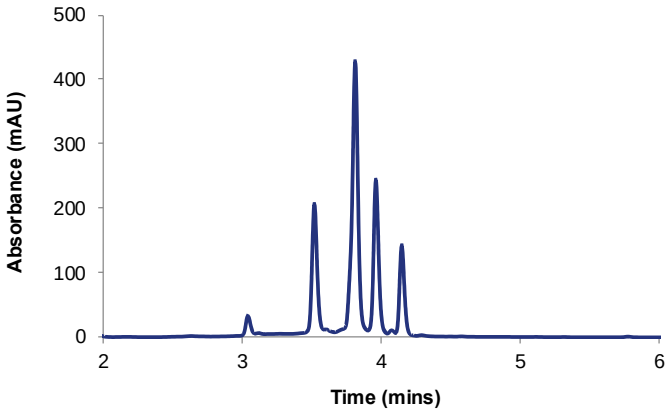


Figure 43 - Crude analytical HPLC trace of the ligation of Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**). Analytical gradient 10-100% B over 5 mins, 280 nm.

The presence of multiple peaks was intriguing, as NCL is usually a relatively clean procedure. It was theorised that the presence of impurity peaks could be due to a mixture of side reactions. For example, after the ligation is complete, the resultant free thiol can further react with another selenoester. A second possibility is that a sulfide-selenide could formed with the phenyl selenide produced as a by-product of the ligation. To test these theories, NH_2OH and TCEP were added to the ligation solution. This would result in the hydrolysis of the pendant thioester and the reduction of the sulfide-selenide and should result in a single peak being formed. The resultant HPLC trace from this is shown in Figure 44.

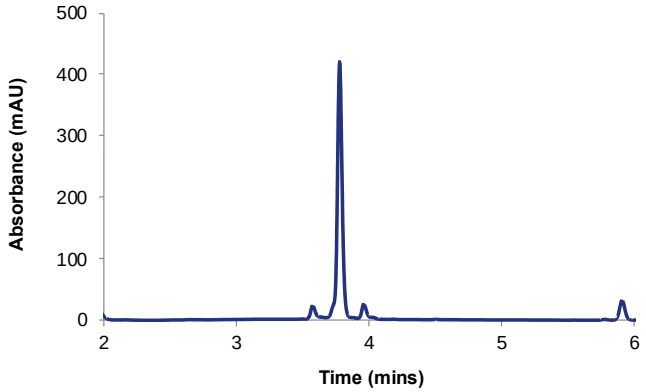
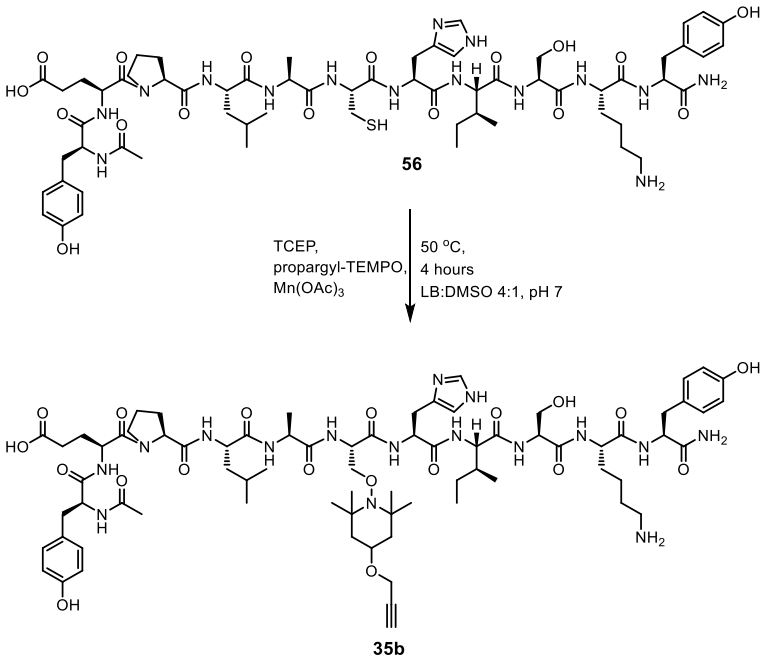


Figure 44 - Crude analytical HPLC trace after addition of NH_2OH and TCEP to the ligation of Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**). Analytical gradient 10-100% B over 5 mins, 280 nm.

This proved the theory was correct, as the extra peaks are no longer present. The reaction mixture was extracted with hexane to remove the precipitated DPDS and taken forward to the conjugation step. Unfortunately, even after 16 hours in the reaction conditions only ligated peptide was obtained from purification. This was interesting, as it would be expected that desulfurisation would have at least occurred. This result could suggest one of two problems, a product from the ligation itself was preventing further reaction from occurring, or the reaction does not work on internal Cys residues. To test this, the two fragments were ligated then purified, giving the desired ligated peptide (Ac-YEPLACHISKY-NH_2, **56**). Next, the ligated peptide was subjected to the conjugation conditions, the general scheme for this reaction is shown in Scheme 57.

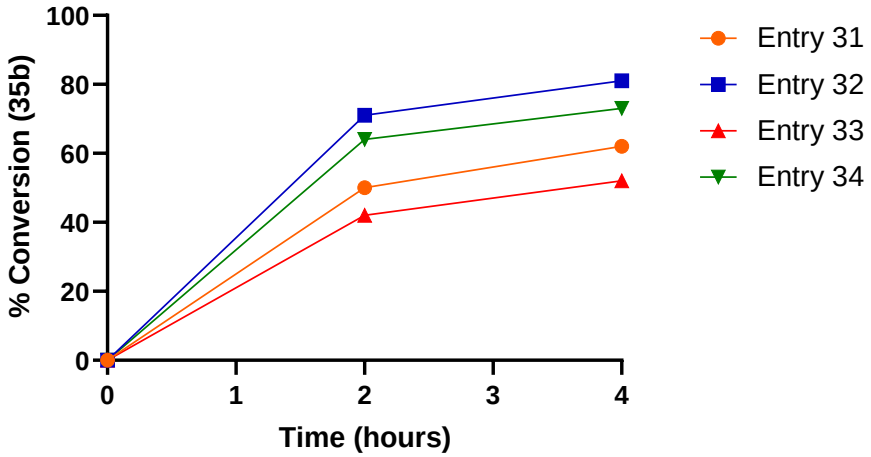


Scheme 57 - General reaction scheme for the conjugation of Ac-YEPLACHISKY-NH2 (**56**) with propargyl TEMPO (**23**) forming the conjugated product (**35b**).

For initial trials, propargyl TEMPO (**23**) was used as the TEMPO probe. The conjugation reaction was slower than previously seen for Cys conjugation. After 4 hours the reaction had only reached 45% conversion to desired product (**35b**). This at least proved that conjugation at an internal Cys residue was possible, and it was likely a by-product from ligation that was preventing the reaction from occurring. This reaction was further optimised to see whether it was possible to reach 100% conversion. The results from these reactions are shown in Table 7. All the reactions stated were performed at 50 °C and analysed at the specified time points.

Table 7 - Results from the optimisation of Ac-YEPLACHISKY-NH_2 conjugation with propargyl TEMPO (Scheme 57) forming the conjugated product (**35b**). Percentage conversion shows the conversion of starting material to the desired product (**35b**). The time-point conversions are shown in the two graphs below.

Entry	Peptide (**56**)/ mM	TCEP eq.	Propargyl TEMPO (**23**) eq.	$Mn(OAc)_3$ eq.	2 hrs %conv. (**35b**)	4 hrs %conv. (**35b**)
1	1	50	2	5	39	45
2	1	50	2	20	55	73
3	1	100	2	20	52	82
4	1	100	5	20	53	83
5	2.5	50	5	10	50	62
6	2.5	50	2	10	71	81
7	2.5	100	2	5	42	52
8	2.5	100	5	10	64	73



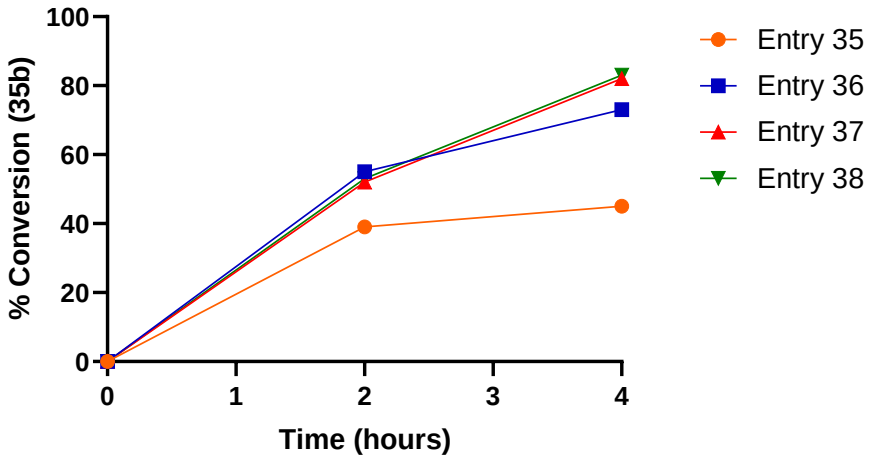
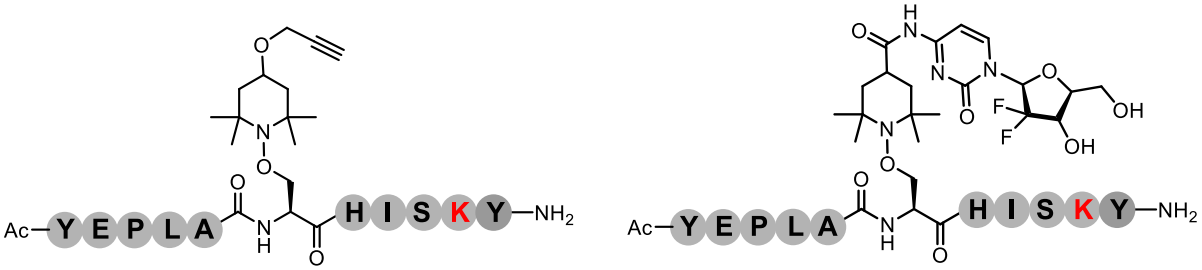


Figure 45 – Plots showing the percentage conversion against time from the results in Table 7.

The results from the reaction optimisation shown in Table 7 show the reaction could be optimised to a 4-hour conversion of 83%. This was a significantly greater conversion than seen previously using the original conditions. These conditions were applied to the propargyl and gemcitabine probes to obtain an isolated yield, forming products **35b** and **38b** respectively. These reactions were successful, with the results shown in Figure 46. This shows the successful application of the TEMPO conjugation method to ligated peptides and internal Cys residues.



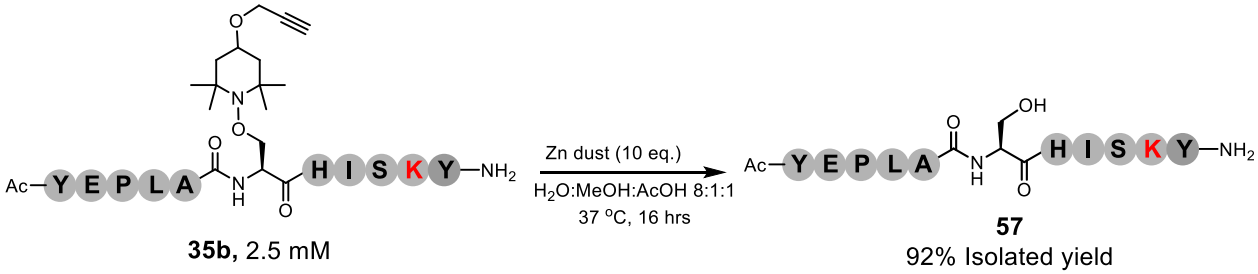
35b
67% (Conjugation yield)
52% (2-step yield)

38b
62% (Conjugation yield)
48% (2-step yield)

Figure 46 - Isolated yields from the two-step ligation and conjugation of Ac-YEPLA-SePh (**32**) and H-CHISKY-NH_2 (**55**).

As mentioned previously, the resulting TEMPO conjugate is very stable. This research has shown that the conjugate can be selectively cleaved using Zn/AcOH. This would be a very interesting addition to NCL chemistry as it would allow the ligation of peptide fragments, followed by conjugation and reductive cleavage to Ser. Up to now, the only post-ligation conversion of Cys has been to Ala. This development would allow the NCL procedure to be applied to Ser ligation junctions as well as Cys and Ala, opening the potential to the synthesis

of a wide range of small proteins. To test this, the propargyl TEMPO conjugated product (**35b**) was subjected to the Zn(0)/AcOH reductive cleavage conditions (Scheme 58). The resultant crude HPLC trace from this reaction is shown in Figure 47.



35b, 2.5 mM

57
92% Isolated yield

Scheme 58 - Reaction scheme for the reductive cleavage of TEMPO conjugation (**35b**) to form the native Ser product (**57**).

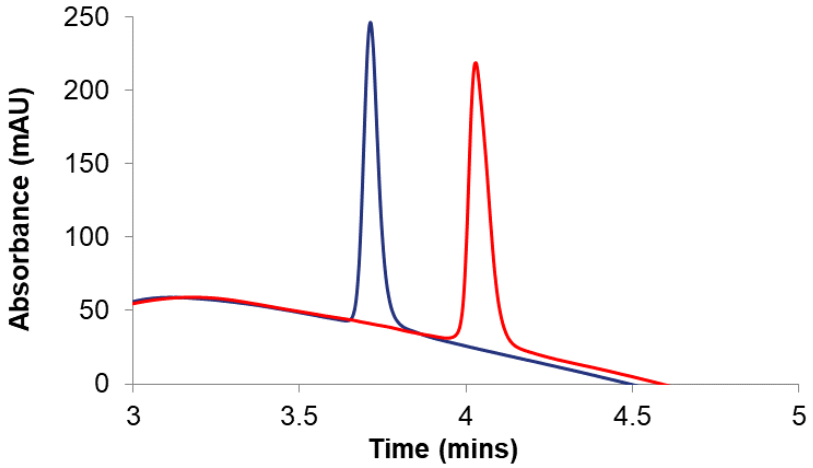


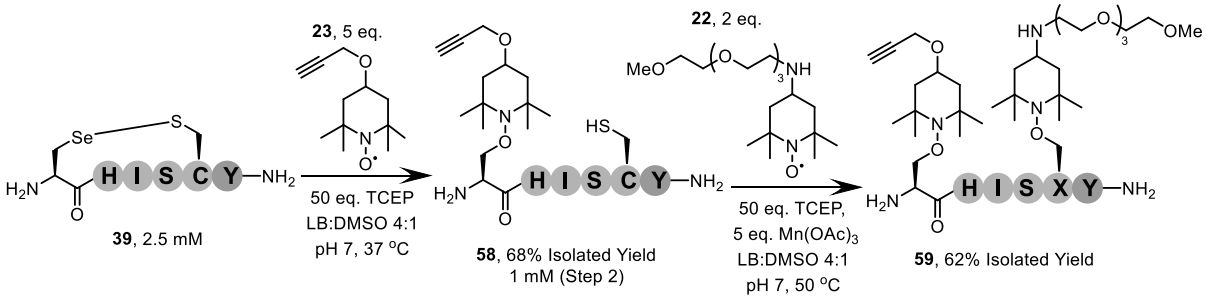
Figure 47 - Crude HPLC trace from the reductive cleavage of the Ac-YEPLACHISKY-NH_2 propargyl TEMPO product (**57**, blue) compared to the starting conjugate (**35b**, red). Analytical gradient 10-100% B over 5 mins, 280 nm.

The crude HPLC trace shown in Figure 47 shows that the TEMPO conjugate from Cys ligation and conjugation can be very cleanly cleaved to form Ser (**57**). The isolated yield from this reaction was 92%, showing very efficient removal of the conjugate. This is a novel method, allowing the NCL procedure to be performed at Ser residues.

3.5 Cys/Sec "Dual-functionalisation"

Another challenge faced in the modification of peptides and proteins is the ability to dual label at specific sites. Due to the lack of reactive, orthogonal and low abundant amino acid residues, selectively labelling two different positions is difficult. As shown in the Sec chapter of this thesis, this method was successfully applied to Sec in the presence of Cys. The Cys residue remained unaffected by the conditions used. The aim of the following reactions was to explore whether the Cys residue could be functionalised after the Sec residue. This would allow the

introduction of two differently functionalised TEMPO groups at specific positions in the peptide. The general scheme for this reaction is shown in Scheme 59.



Scheme 59 - Reaction scheme for the dual functionalisation of a peptide with a Sec and Cys residue (H-UHISCY-NH_2 (**39**)).

The model peptide chosen for this study was the same as used previously in the one-pot studies, H-UHISCY-NH_2 (**39**). The first step of this involved functionalisation of the Sec residue using propargyl TEMPO (**23**). This was achieved in a 68% yield (**58**) using general protocol C (Mn free conditions). The Cys residue was then targeted, using Mn accelerated conditions, general protocol B. The tetraEG TEMPO (**22**) probes was chosen for this reaction, and the desired product isolated in a 64% yield (**59**). The HPLC traces from these reactions are shown in Figure 48 and the final product in Figure 49.

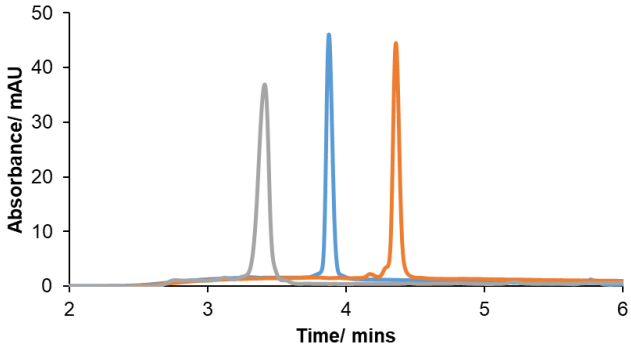
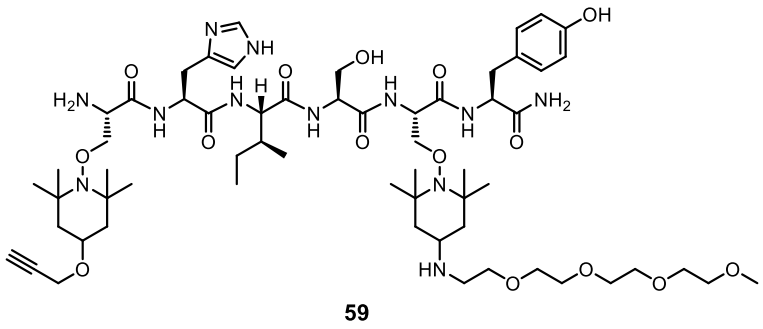


Figure 48 - HPLC traces of starting H-UHISCY-NH_2 (**39**, grey), propargyl TEMPO modified product (**58**, orange) and tetraEG TEMPO modified product (**59**, blue).



59

Figure 49 - Structure of the final product (**59**) from the dual-modification of H-UHISCY-NH_2.

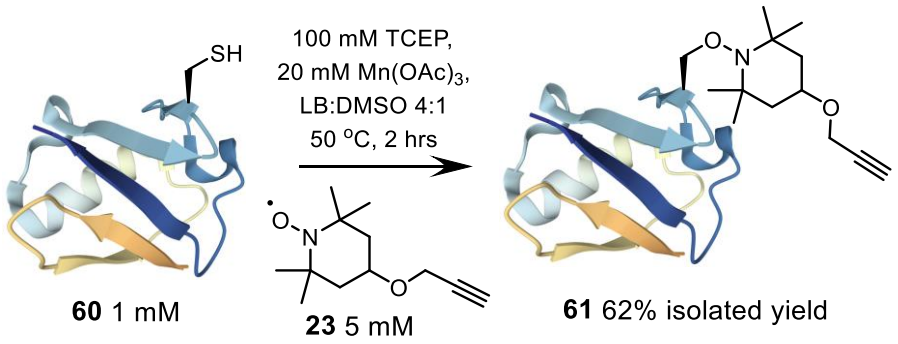
This data proved that a peptide could be functionalised with two different TEMPO probes with peptides having a Sec and Cys residue. The HPLC traces shown in Figure 48 show clean conversions from starting material to dual functionalised peptide. This would be a very useful technique, with potential applications in imaging and pharmaceuticals.

3.6 Application of the Cys-selective method to protein model Ub K48C

The final test for this procedure was to apply the conjugation method to a recombinantly expressed protein. This work used a K48C mutatant of ubiquitin (**60**, Ub) provided by Prof. Robert Layfield (School of Life Sciences, University of Nottingham). Trialling the conjugation procedure on a biologically expressed protein would demonstrate that this method could be applied to the preparation of protein drug conjugates or labelling for imaging purposes. The protein sequence used was as follows –

H-MQIFVKTLTGKTITLEVESSDTIDNVKSKIQDKEGIPPDQQRLIFAG**C**QLEDGRTLSDYNIQ KESTLHLVLRLRGG-OH

The protein contained a vast range of functionality, with the Cys residue close to the middle of the sequence. This would show whether the issues of Cys internalisation seen when using the Ac-YEPLACHISKY-NH_2 model would be further exacerbated by the larger number of residues on either side. The scheme for this reaction is shown in Scheme 60.



Scheme 60 - Reaction scheme for the bioconjugation of K48C ubiquitin with propargyl TEMPO (**53**).

Due to the small amount of protein material available (< 10 mgs) and the high molecular weight, it would be operationally difficult to apply the conditions used above for the Ac-YEPLACHISKY-NH_2 model. Therefore, for the conjugation of the protein model, lower concentration conditions, referred to as general protocol E (details in section 7.2), were applied. The lower concentration increases the operational simplicity of the reaction. As the protein has a high molecular weight, increasing the TEMPO equivalents from 2 to 5 does not represent a significant increase. The amount of propargyl-TEMPO (**23**) used in the reaction

was less than 1 mg. The conjugation was performed using general protocol E, and the desired conjugated product was obtained in a 62% isolated yield (**61**). This was pleasing as this isolated yield was comparable to the yield when conjugating onto the Ac-YEPLACHISKY-NH_2 model.

To ensure that the modified protein could be re-folded after conjugation under denaturing conditions (6M Gdn•HCl) and HPLC purification, the purified Ub-conjugate (**61**) and the starting K48C protein (**60**) were dissolved in ligation buffer and dialysed into folding buffer (25 mM Na_2PO_4, 100 mM NaCl, pH 6.9) overnight. Each sample was then analysed by NMR using an Avance III 800 MHz NMR system. The NMR analysis and data processing was performed by Dr Huw Williams and Dr Jed Long, respectively. The NMR data is outlined in the experimental and the comparison data is shown in Figure 50 and Figure 51.

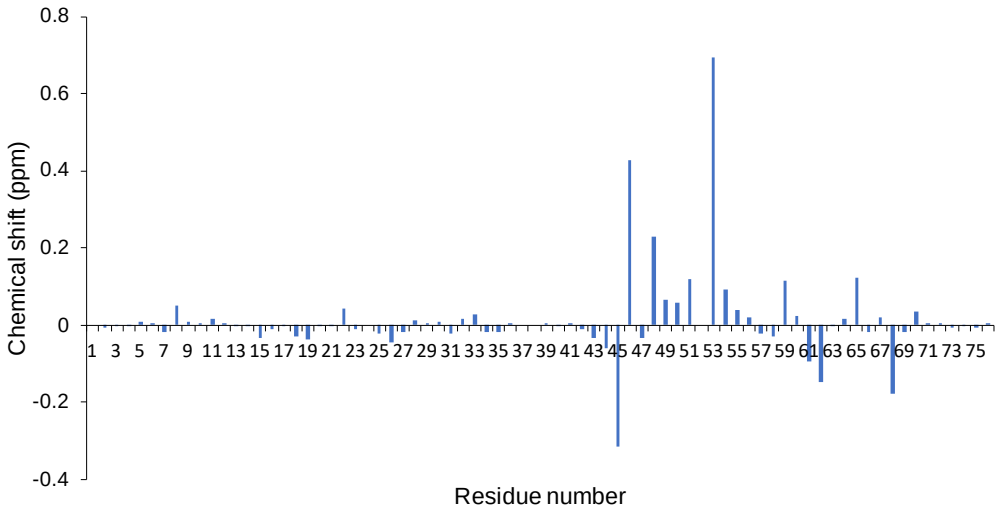


Figure 50 - Δ chemical shift (ppm) plot comparing NH signals for K48C Ub (**60**) and Ub conjugate (**61**)**.**

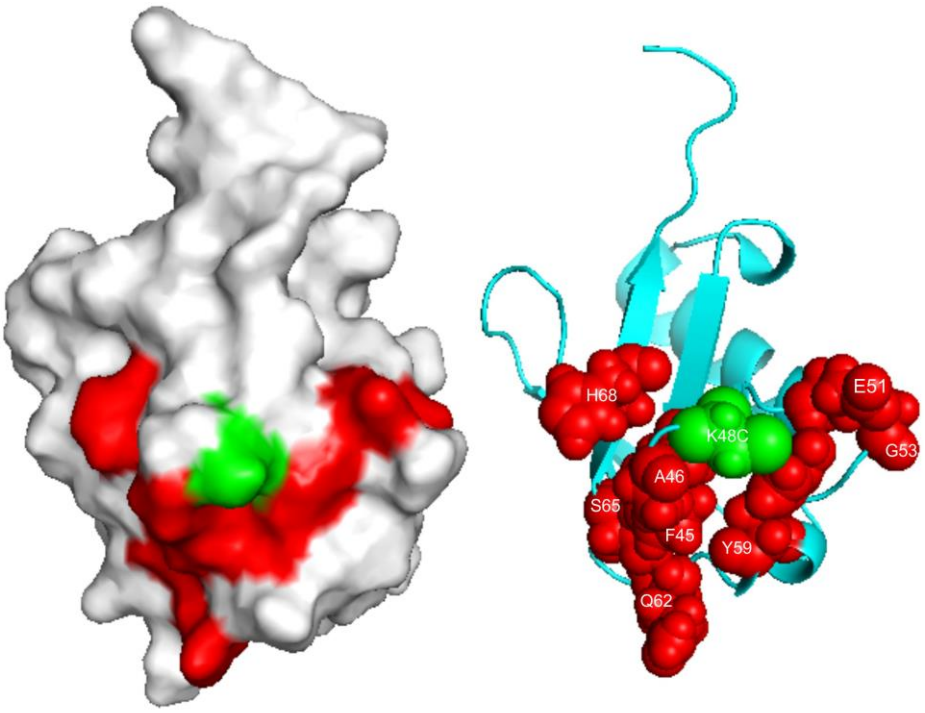


Figure 51 - Space filling and ribbon models of Ub; modified residue at position 48 highlighted in green, residues that experience significant deviation in chemical shifts for NH signal highlighted in red.

The TOCSY data of both the K48C Ub mutant starting protein (**60**) and the conjugated Ub (**61**) were recorded and assigned by comparison to the published ubiquitin data (BMRB 68 and 4769).[151] These spectra show that the fold of ubiquitin is intact, both the conjugate and the unmodified sample showed extended NH regions (6.5 – 9.5 ppm) and upshifted methyl groups indicating formation of the native tertiary structure. A comparison of the chemical shift values shows very little deviation across the majority of the sequence. While the conjugated propargyl TEMPO could not be observed directly, the significant deviation around Gly53 suggests it is there and the alkyne group is close to the glycine.

3.7 Conclusion

This chapter describes the adaptation of the TEMPO bioconjugation method developed in Chapter 2 to enable Cys-selective bioconjugation. The optimised method was used to conjugate a range of TEMPO probes onto the model peptide Ac-CWHISKEY-NH_2. This showed that a range of functionality could be installed onto the Cys residue, without interfering with other reactive residues. The method was then altered slightly to allow conjugation onto internal Cys residues formed from peptide fragment NCL. This modification was achieved in good yields, allowing installation of interesting functionality to peptide synthesised through NCL. The conjugate formed was then reductively cleaved to form a serine residue. This marks a novel method for the conversion of a Cys residue into a Ser residue, expanding the

usefulness of the NCL method to Ser ligation junctions. The method was then successfully applied to the recombinantly produced K48C ubiquitin mutant. The same conditions were used as developed for the ligation model, with similar isolated yields observed. This was a vital result as it demonstrates that significantly increasing the complexity of the target did not have an adverse effect on the reaction. Finally, the method was used to produce “dual functionalised” peptides. A peptide containing a Sec and Cys residue was first modified at the Sec residue, then the Cys, installing two different biologically relevant groups. The tolerance, selectivity, versatility, and operational simplicity of this novel method will enable researchers to employ the described protocols to useful polypeptide-small molecule tools for a range of applications.

4. Site-selective bioconjugation via desulfurative carbon-carbon bond formation

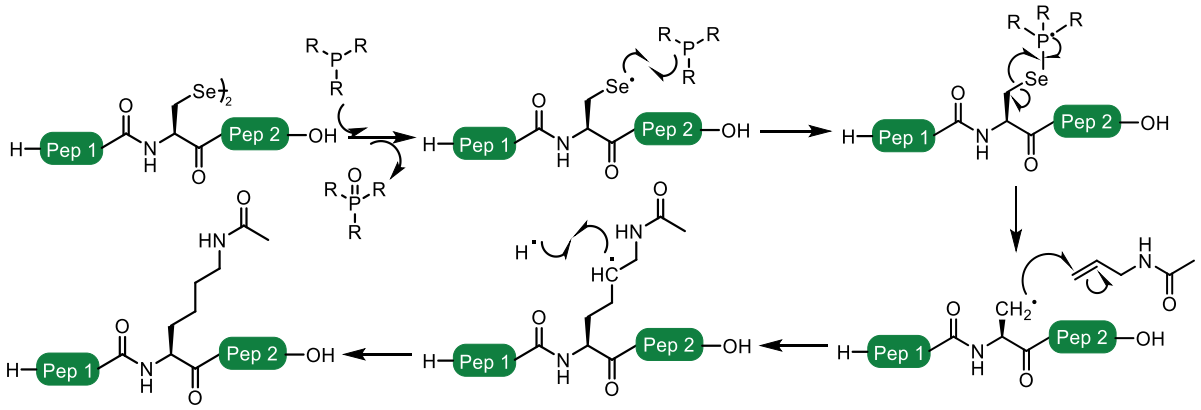
4.1 Chapter Abstract

The post-translational modification (PTM) of Lys is an extensively researched area of biology. A large body of study in this field focuses on the PTMs of histone proteins. These proteins play a key role in the regulation of gene expression. DNA wraps around the histone proteins and this complex further coils around itself to form chromatin. The location and degree of PTMs present on the histone proteins regulates the chromatin structure and access to genetic information. This, in turn, controls which proteins are ultimately expressed.[152] There are a range of Lys PTMs which are known to regulate histone biology, including acetylation,[152,153] methylation,[152,154–156] succinylation,[157] malonylation,[158] propionylation, and butyrylation.[159] Another example of a key Lys PTM that has received extensive attention is the acetylation of Lys residues in the protein ubiquitin (Ub). Ub is a PTM itself, marking proteins for degradation and aiding in DNA repair. Lys acetylation in Ub represses polyubiquitin chain elongation and regulates the activity of Ub.[160]

Due to the obvious importance of Lys PTMs, extensive research has been conducted focussing on the selective installation of these PTMs. This has been achieved both biologically and chemically. Biological methods involve the use of amber-codon suppression to install a PTM Lys residue at a specific site.[158,160–162] Another biological method utilises enzymes, such as acyltransferase. These have been successfully used to install lysine PTMs including acetylation, propionylation and butyrylation.[159]

There have also been a wide range of chemical methods developed for the installation of Lys PTMs. Most of these focus on the installation of Lys mimetics, utilising the unique reactivity of Cys. This has been achieved in two main ways; the first uses Cys directly by reacting it with alkyl halides or activated alkenes.[153,157,163–165] Another common method first converts Cys to Dha followed by addition of a thiol bearing the desired functional group to produce the Lys mimetic.[124,153,155,166] Recently, a method to generate site-specific Lys PTMs has been developed.[86,167] This uses radicals generated from alkyl halides to react with the Dha residue, forming a new C-C bond. These methods allowed the installation of a wide range of groups forming common Lys PTMs including mono, di and tri methylations. Although highly applicable to a wide range of proteins, it is operationally difficult as it requires strictly inert conditions. In addition, as Dha is first formed, the stereochemistry of the residue is lost. The prominence of Lys PTMs in the literature and lack of methods to easily install these site-selectively onto proteins shows why a novel method to accomplish this is required.

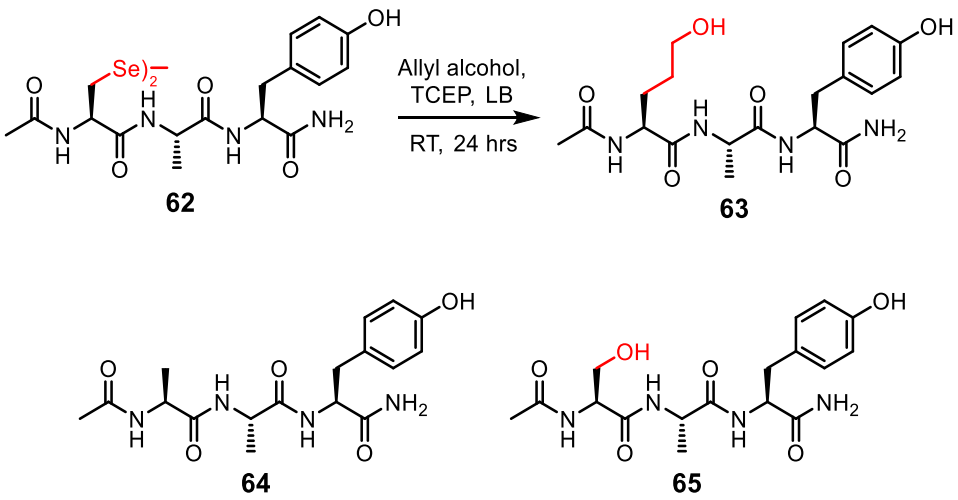
This chapter describes the development of a novel C-C bond forming reaction that enables the installation of native Lys PTMs, Lys PTM mimics, and desirable non-proteinogenic functionality at a Cys residue. The method builds on the chemistry described in the previous chapters, using alkene radical traps to intercept the alanyl radical formed via desulfurisation (Scheme 61). Furthermore, alkene substrates that yield a stabilised radical intermediate were also explored in an attempt to develop a faster, more efficient visible-light-catalysed C-C bond forming bioconjugate reaction. This proved successful, allowing rapid, high yielding conjugation of peptides and proteins with a range of substrates. This method produces highly stable bioconjugates, with the conjugate linked to the peptide/protein via a C-C bond.



Scheme 61 - Reaction mechanism for the conversion of Sec to acetyl lysine.

4.2 Initial reactions performed on selenocystine

After demonstrating that the alanyl radical produced through deselenisation of Sec could be trapped with functionalised TEMPO under mild conditions, we moved on to the exploration of alkene traps that would enable the installation of desired functionality via C-C bond formation. The first reagent explored was allyl alcohol using the Ac-UAY-NH_2 (**62**) model peptide in the presence of the phosphine TCEP to form conjugated product **63** as shown in Scheme 62.



Scheme 62 - Reaction scheme for the trapping of allyl alcohol by Sec during radical-induced deselenisation, showing the desired product (**63**) and the two potential by-products Ala (**64**) and Ser (**65**).

Initial experiments using varying equivalents of allyl alcohol were explored to test the potential of this reaction. All reactions were performed for 24 hours at ambient temperature unless otherwise specified. The results from these are shown in Table 8.

Table 8 - Initial allyl alcohol trapping experiments (Scheme 62).

Entry	Peptide (**62**)/ mM	TCEP eq.	Allyl alcohol eq.	Product (**63**) %conv.	SM (**62**) %rem.	Ala (**64**) %conv.	Ser (**65**) %conv.
1	5.0	25	200	5	69	13	7
2	2.5	100	400	10	58	19	0
3	1.0	10	100	3	0	7	84

These experiments show minimal conversion to the allyl-trapped product (**63**). This could be due to the unactivated alkene being a poor radical trap for the alanyl radical, the radical extracting an H-atom from allyl alcohol, or the resulting radical not being quenched and polymerisation occurring. Polymerisation is, however, unlikely as products from this should have been detectable. To eliminate the possibility of polymerisation, tributyltin hydride was added to the reaction as an H-atom donor. These results are shown in Table 9.

Table 9 - Allyl alcohol trapping experiments forming allyl-trapped product (**55**).

Entry	Peptide (**62**)/ mM	TCEP eq.	Allyl alcohol eq.	Bu_3SnH eq.	Product (**63**) %conv.	SM (**62**) %rem.	Ala (**64**) %conv.
1	2.5	25	200	10	3	84	12
2	2.5	25	200	20	4	87	7

This did not lead to an increase in the conversion to allyl-trapped peptide. All trials so far have shown that starting material remains in significant quantities, so it is possible that higher percentage conversions can be achieved. Further optimisation was attempted by adding a water-soluble radical initiator to the reaction, VA-044. This has been shown in literature to increase the rate of deselenisation.[33] It was also considered that tributyltin hydride could be quenching the alanyl/selenyl radical, as it is an established H-atom donor, so the milder sodium cyanoborohydride was trialled. These attempts followed the general Scheme 62 and produced the results shown in Table 10.

Table 10 - Allyl alcohol trapping experiments forming allyl-trapped product (**54**).

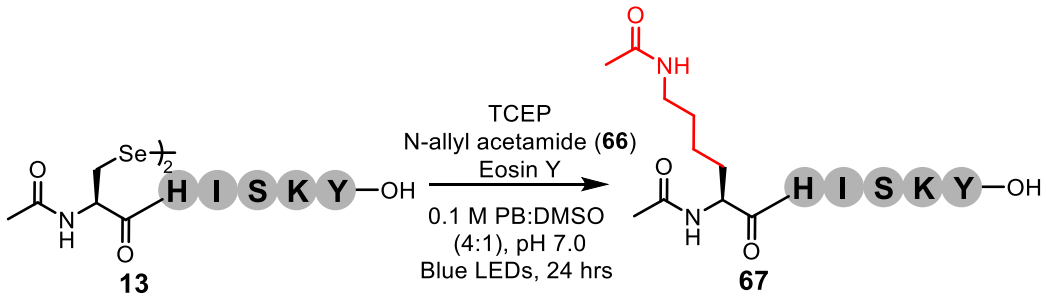
Entry	Peptide (**62**)/ mM	TCEP eq.	Allyl alcohol eq.	Bu_3SnH eq.	$NaCNBH_4$ eq.	VA-044 eq.	Product (**63**) %conv.	SM (**62**) %rem.	Ala (**64**) %conv.	Ser (**65**) %conv.
1	5	25	200	/	/	10	41	20	38	0
2	5	25	50	/	/	10	24	3	66	5
3	5	25	50	10	/	10	24	4	66	4
4	2.5	100	400	/	/	40	36	15	40	0
5	2.5	100	400	/	/	20	39	17	33	4
6	2.5	100	400	20	/	20	35	12	37	2
7	2.5	100	400	/	20	20	24	25	29	10
8	2.5	100	400	20	/	100	30	7	61	0
9	1	50	1000	/	/	10	30	9	16	23
10	1	50	1000	/	/	100	34	17	32	4
11	1	10	100	/	/	20	28	1	28	4
12	2.5	20	400	/	/	8	31	0	62	6

The overall conversion to the desired product using the conditions in Entry 1 (Table 10) increased to 41%, showing significant improvement from the initial experiments. A 41% conversion is not as high as required for this method to be considered viable for protein modification, so further modifications to the methodology were made in an attempt to improve the yield of trapped peptide. For the next reaction the allyl probe being used was switched to N-allyl acetamide (**66**), as this produced a more biologically relevant final product. As previously mentioned, this is an important and functional lysine PTM. As well as this, the model was also changed to add in extra functionality. The model peptide Ac-UHISKY-OH (**13**) was chosen for the following studies.

The first reaction attempted on this model used the conditions from Entry 1 (Table 10). Unfortunately, the conversions achieved were <10%. It was unclear as to what was occurring to result in such low conversions, but multiple peaks were present by HPLC.

The low conversions could indicate that the bulkier neighbouring His residue was the issue. This led to alternative conditions being explored. It could also indicate that the large excesses of the radical initiator VA-044 present was causing side reactions to occur. With these thoughts in mind, a milder method to assist the reaction was sought.

Recent advances in the use of photoredox catalysis to induce radical desulfurisation led to the decision to use eosin Y as the photoredox catalyst. Eosin Y is a cheap organic-based photocatalyst, which when excited by visible light will induce single electron oxidation of the selenol or thiol. This results in a radical cation being formed on the selenium, which is readily deprotonated to form the selenyl radical. This can then donate into the tertiary phosphine as explained before in the TEMPO trapping section in Chapter 2.3 (Scheme 43). The general scheme for this reaction is shown in Scheme 63.



Scheme 63 - Reaction scheme for Sec allyl trapping initiated by eosin Y using the Ac-UHISKY-OH (**13**) starting peptide forming conjugated product (**67**).

The initiator used, eosin Y, should be catalytic so this was used in sub-stoichiometric quantities. As the alanyl radical formed will be very unstable and attempt to immediately quench, it was theorised that a very high excess of N-allyl acetamide (**66**) would be necessary.

The high excess would give the radical a greater chance to react with the allyl group, initiating the reaction. The previous results showed that 25 eq. of TCEP was enough to successfully induce deselenisation, so this amount was used. The results from these reactions are shown in Table 11.

Table 11 - Example Sec allyl trapping reactions using eosin Y and photoredox initiator forming conjugated product (**67**).

Entry	Peptide (**13**)/ mM	TCEP eq.	N-Allyl acetamide (**66**) eq.	Eosin Y eq.	Bu_3N eq.	Product (**67**) %conv.	Ala (**14**) %conv.
1	1	25	1000	0.01	2	49	51
2	0.5	25	1000	0.001	2	52	48
3	0.5	25	2000	0.01	2	49	51

The results shown in Table 11 were very promising. They show in increase in percentage conversion from 10% using VA-044 to 52% using eosin Y of the allyl trapped product (**67**). The crude trace from Entry 2 (Table 11) is shown in Figure 52.

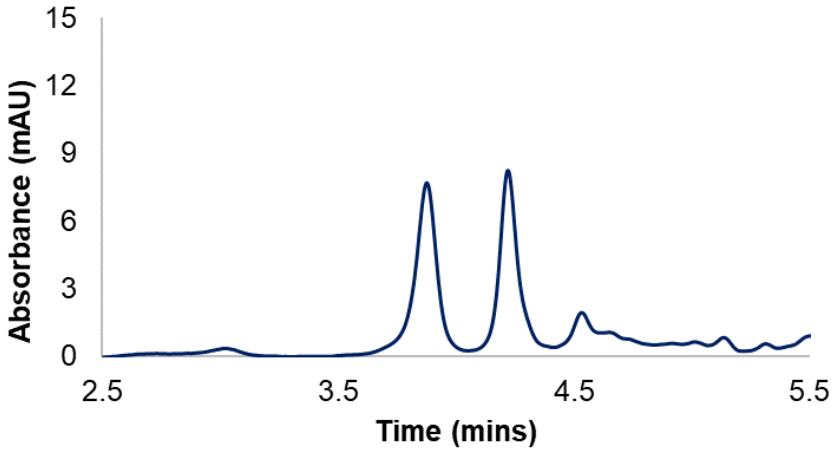
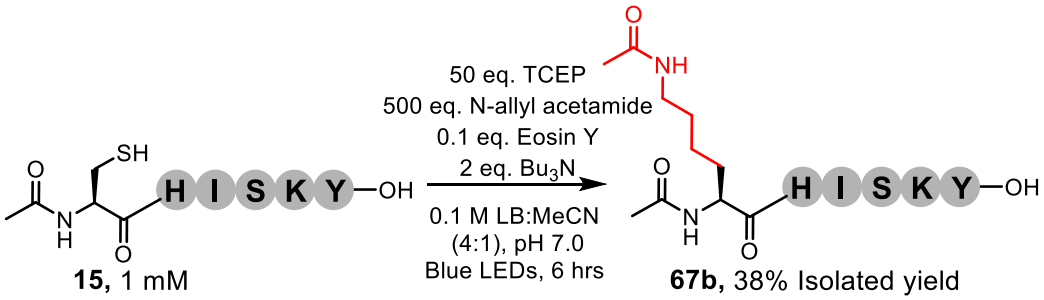


Figure 52 - Crude HPLC trace from Entry 2 (Table 11). The peak at 3.8 minutes is Ala (**14**); the peak at 4.2 minutes is desired product (**67**). Analytical HPLC gradient 10-25% B over 5 mins, 280 nm.

The crude HPLC trace shown in Figure 52 shows that the reaction is relatively clean, resulting predominantly in two clear products. This reaction would have to be further developed before being deemed applicable to peptide and protein models. It would be interesting to apply to a larger model, as previous reactions showed that the size of the model did not affect the isolated yield. If this could be applied to larger peptides or proteins and achieve similar yields, this would be a very useful method for installing lysine PTMs.

4.3 Reaction optimisation using cysteine-containing peptides

After some success with Sec, attempts were made to see whether it was possible to perform the ally trapping reactions with Cys. The first reaction performed mimicked conditions developed for the modification of Sec. As eosin Y has been used for radical desulfurisation of Cys to Ala, it was thought these conditions could be successful. The general scheme for the reaction performed is shown in Scheme 64.



Scheme 64 - Reaction scheme for the conjugation of Ac-CHISKY-OH (**15**) with N-allyl acetamide to give product (**67b**).

The reaction shown in Scheme 64 resulted in an isolated yield of 38%. The crude and pure HPLC traces are shown superimposed in Figure 53. The crude trace shows that the reaction is relatively clean, with only three clear peaks present. The product HPLC trace shows the desired product was obtained to a purity of > 95% after a single purification. This shows that, on this model, the desired conjugated product is easily separated from the Ala by-product.

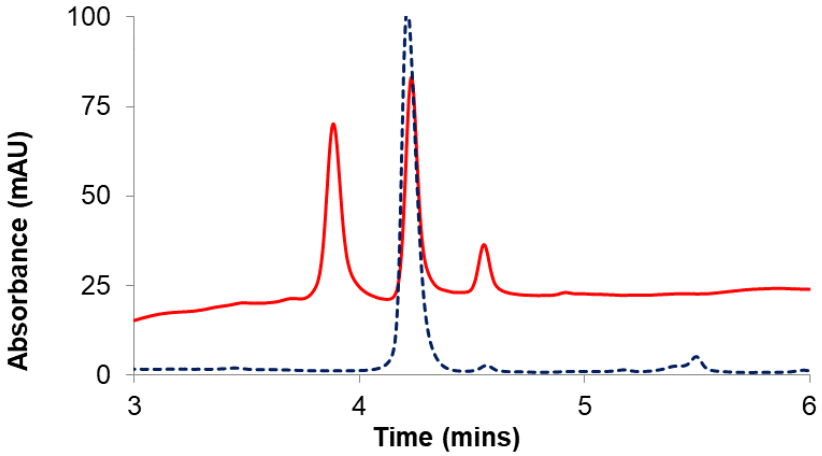
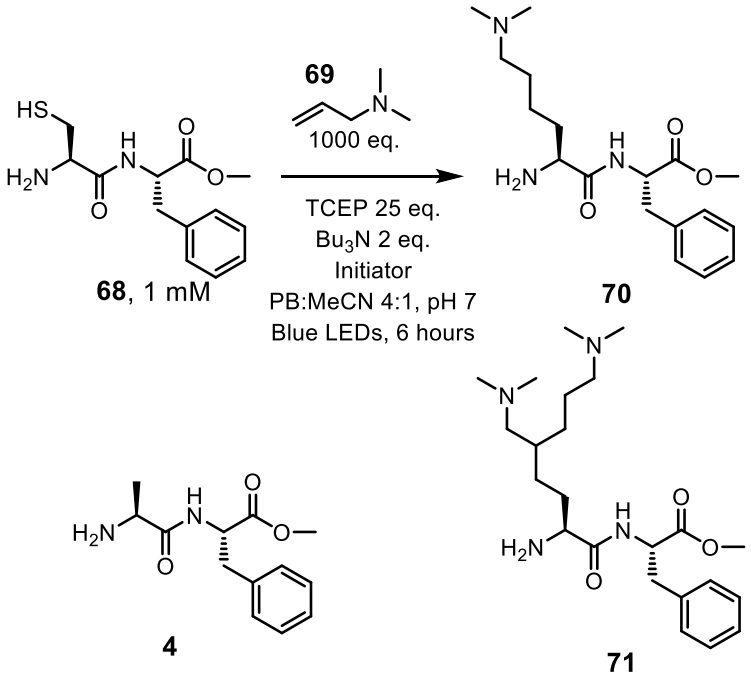


Figure 53 - Crude HPLC trace of the reaction shown in Scheme 64 (red) and the pure product (**67b**, blue dash). The peak at 3.8 mins is Ala (**14**) and the peak at 4.6 mins is starting material (**15**). Analytical HPLC gradient 10-25% B over 5 min, 280 nm.

This was a promising starting point, and, if transferrable to larger models, would also be considered a moderately good yield. As altering the radical initiator from VA-044 to eosin Y

had such a large affect previously, the first factor assessed was the initiator. The reaction was performed on a small Cys dipeptide (H-CF-OMe, **68**) to screen a range of initiators. The scheme for this reaction is shown in Scheme 65. The results from the initiator screening are shown in Table 12. All reactions were performed with a peptide concentration of 1 mM with 1000 eq. N,N-dimethylallylamine (**69**, DMAA), 2 eq. Bu_3N and 25 eq. TCEP. The Ir(III) catalysts refers to $(Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$. It should be noted here that after running these experiments it was made apparent that the pH determination method being used for all H-CF-OMe (**68**) reactions was inaccurate. The Laquatin pH probe was damaged and the pH paper used for pH determination proved to be under-reading. This resulted in the actual pH being around two units lower than expected. This led to the requirement of the initial reactions and optimisation being repeated. The repeats of the optimisation data are later in this chapter on the larger Ac-CWHISKEY-NH_2 (**49**) model.

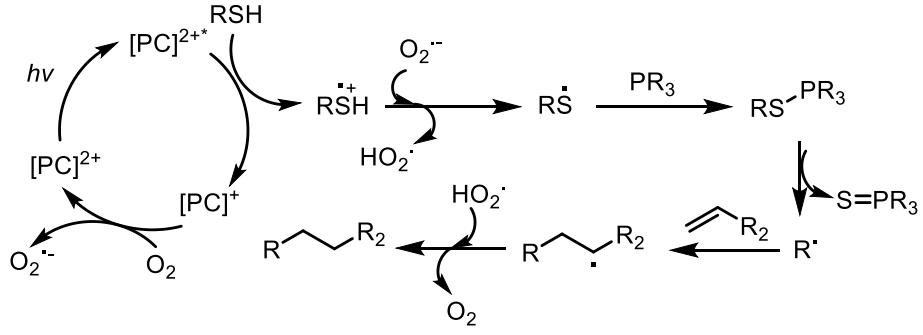


Scheme 65 - Reaction scheme for the modification of H-CF-OMe (**68**) with N,N-dimethylallylamine (**69**) using a range of initiators to form trapped product (**70**). Observed by-products include H-atom trapped Ala product (**4**) and multiple addition product (**71**), where the secondary radical formed reacts with another equivalent of DMAA (**69**).

Table 12 - Initiator screening for cysteine desulfurisation and DMAA trapping to form trapped product (**70**). Other includes a range of multiple addition (**71**) products, as well as unidentified by-products.

Entry	Initiator	Eq.	Product (**70**) %conv.	Ala (**4**) %conv.	SM (**68**) %rem.	Other %conv.
1	Eosin Y	0.01	0	0	0	100
2	Rose Bengal	0.01	10	6	0	84
3	4-CzIPN	0.01	25	11	0	64
4	TAP	0.01	0	0	5	95
5	Ir(III)	0.01	35	28	0	34
6	ZnTPP	0.1	0	19	0	81
7	Eosin Y	0.1	27	19	22	26
8	4-CzIPN	0.1	33	29	14	16

The results in Table 12 show that the most successful initiators at this stage were Ir(III), eosin Y and 4-CzIPN. As shown in Scheme 66, the role of the initiator involves a single electron oxidation of the thiol, forming a radical cation. This is deprotonated, forming a thiyl radical, which can be desulfurized following the standard mechanism. It is interesting that different catalysts show different results. Looking at the oxidation potential for each catalyst does not offer much insight, with oxidation potentials ranging from +0.31 to +1.4 V.[168]



Scheme 66 - General reaction mechanism for the photoredox cycle for the photocatalysts used in Table 12.

Unfortunately, the conversions are still relatively low, with the majority of peptide being converted to Ala. This is potentially due to the thiol being a good H-atom donor; when the alanyl radical is formed it preferentially abstracts an H-atom from a thiol instead of donating into the non-activated allyl group. Most reactions involving a radical donating into an alkene

involve an electron-withdrawing group located next to the alkene.[2,68,169] This draws electron density away from the double bond making it more reactive towards radicals. We therefore expect the reaction we are exploring to have limitations, however, the promise of the installation of native PTM Lys side chains motivated us to continue exploring the allyl amine series.

The Ir(III) catalyst was chosen to pursue further, with hope that fully optimising the conditions would lead to a higher conversion. The next factor to optimise was the pH at which to perform the reaction. The results from these trials are in Table 13. The conditions for these reactions were as follows – 1 mM H-CF-OMe (**68**), 10 eq. TCEP, 500 eq. DMAA (**69**), 2 eq. Bu_3N and 0.01 eq. Ir(III) catalyst.

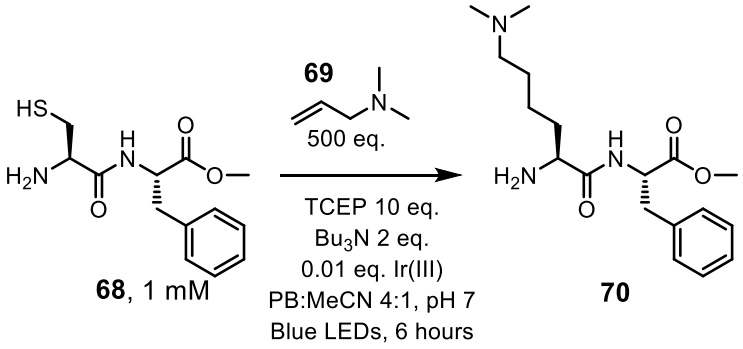


Table 13 - Results from pH optimisation trials to form trapped product (**70**), the values in brackets are the corrected values, assuming all values were 2 pH units lower than expected.

Entry	pH	Product (**70**) %conv.	Ala (**4**) %conv.	SM (**68**) %rem.	MA (**71**) %conv.
1	4 (2)	0	0	100	0
2	5 (3)	0	0	100	0
3	6 (4)	0	0	100	0
4	7 (5)	3	12	75	5
5	8 (6)	10	31	11	13
6	9 (7)	25	31	0	44
7	10 (8)	39	45	0	16

The pH values show that the more basic the reaction mixture the higher conversion to desired product was observed. An issue that began to occur was that of multiple addition (**71**, MA) (Figure 54). This would have been a result of the secondary radical that is formed after initial trapping reacting with another DMAA molecule. Other unidentified products account for the rest of the conversion, where numbers do not add to 100%.

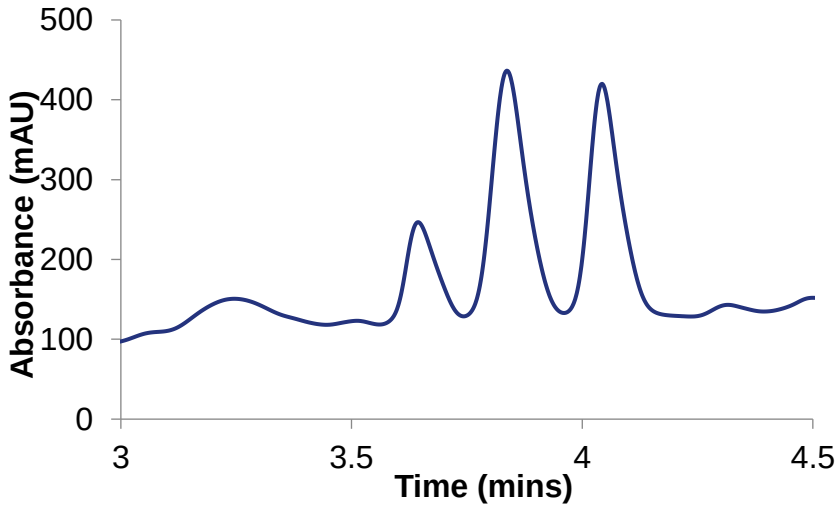


Figure 54 - HPLC trace from Entry 7 (Table 13) showing the multiple addition peak at 3.6 mins. The desired product (**70**) is at 3.8 mins and the Ala by-product (**4**) at 4.1 mins.

These trials showed that a more basic pH leads to the higher conversion. They also showed that outcompeting multiple addition was likely to become a difficulty, along with the previously mentioned H-atom abstraction. With this in mind, it was still decided to proceed with optimisation, in hope that optimising peptide, TCEP, DMAA, base and initiator concentrations would lead to the desired outcome. The concentration of TCEP was altered to assess how this would affect the outcome. For these reactions, the peptide concentration was 1 mM, and 500 eq. of DMAA, 0.01 eq. Ir(III) and 2 eq. 2,6-lutidine was used. The reactions were performed in LB at pH 7. The results from these reactions are shown in Table 14.

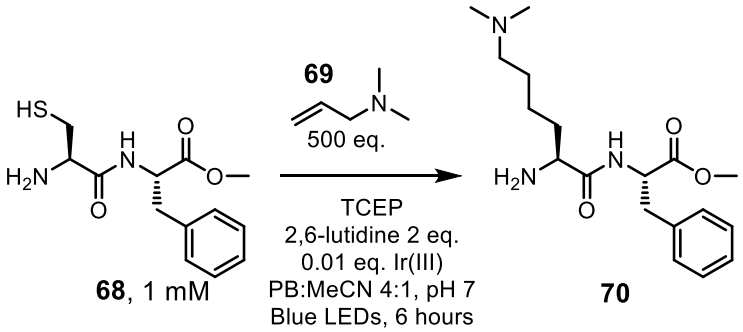


Table 14 - Results from altering the equivalents of TCEP and the effect this had on the allyl trapping reaction to form trapped product (**70**).

Entry	TCEP eq.	Product (**70**) %conv.	Ala (**4**) %conv.	SM (**68**) %rem.	MA (**71**) %conv.
1	2.5	6	16	21	57
2	5	29	54	0	17
3	10	32	45	0	23
4	25	35	45	0	19
5	50	35	43	0	21
6	100	32	50	0	18
7	15	30	41	0	28
8	35	31	39	0	30

The results shown in Table 14 show that the optimal equivalents of TCEP was either 25 or 50 equivalents. It was observed, however, that the production of Ala was still a problem. To try to reduce the amount of Ala produced, the equivalents of DMAA used was altered to see how this affected the reaction. The results from these trials are shown in Table 15. For these reactions the peptide (**68**) concentration was 1 mM. The equivalents of TCEP (25), Ir(III) (0.01) an 2,6-lutidine (2) were maintained for all reactions. The reactions were performed in LB at pH 7.

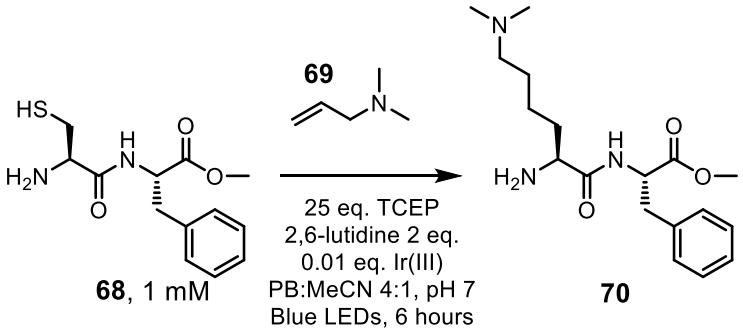


Table 15 - Results from altering the DMAA (**69**) concentration on the percentage conversion of starting material (**68**) to product (**70**).

Entry	DMAA eq.	Product (**70**) %conv.	Ala (**4**) %conv.	SM (**68**) %rem.	MA (**71**) %conv.
1	10	0	77	0	23
2	50	5	57	0	38
3	100	10	58	0	32
4	250	23	46	0	31
5	500	30	33	0	37
6	1000	36	19	0	44

The results in Table 15 show that the more DMAA (**69**) that is added, the higher conversion of SM to product (**70**). Unfortunately, the amount of multiple addition product (**71**) also increases. In attempt to overcome the issue of multiple addition, trialling the reaction at higher dilution was attempted. The results from these reactions are shown in Table 16. All reactions used 25 eq. TCEP, 1000 eq. DMAA, 0.01 eq. Ir(III) and 2 eq. 2,6-lutidine. The reactions were performed in LB at pH 7.

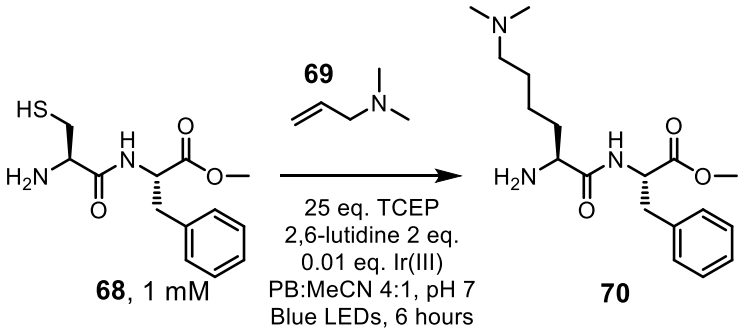
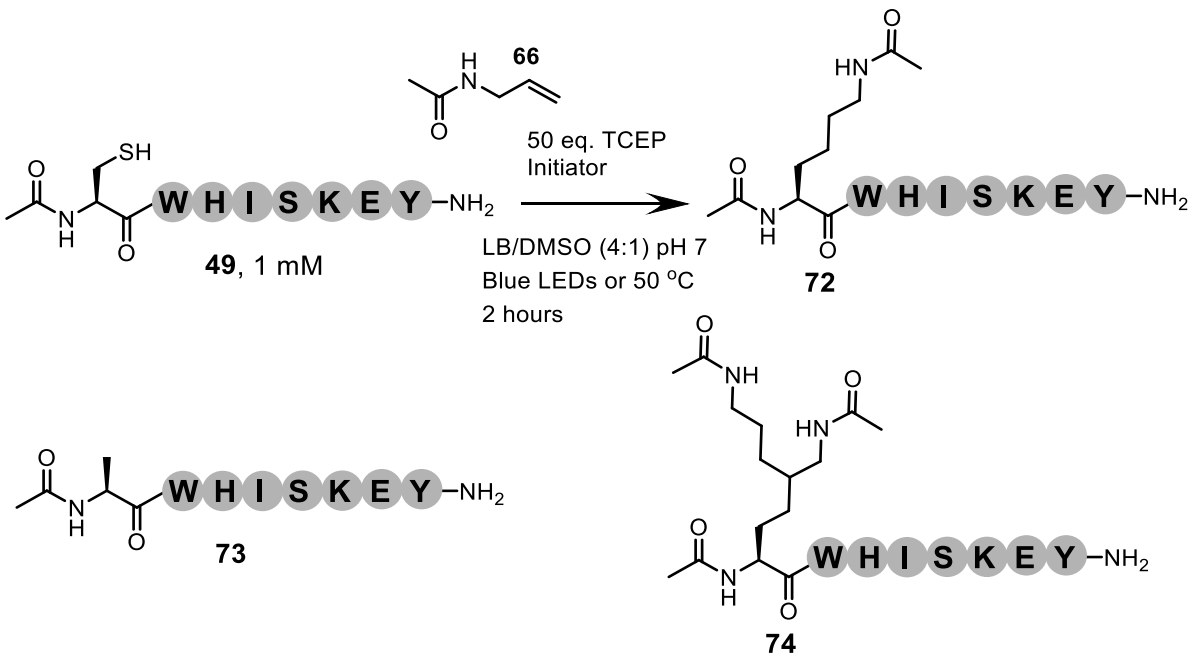


Table 16 - Results from performing the allyl trapping reaction at various peptide concentrations to form trapped product (**60**).

Entry	Peptide (**68**)/ mM	Product (**70**) %conv.	Ala (**4**) %conv.	SM (**68**) %rem.	MA (**71**) %conv.	Other(Unk)
1	1	30	18	18	40	12
2	2.5	27	27	0	52	14
3	0.5	18	12	0	66	4

The results in Table 16 show that higher dilution resulted in a lower amount of desired product (**70**) and a higher amount of multiple addition product (**71**). Upon MS analysis of the peaks, it was seen that the peak for the multiple addition product co-elutes with other compounds. These include the methyl ester hydrolysis product of both the desired product and the multiple addition product. As the methyl ester hydrolysis was complicating optimisation, it was decided to change the model. These reactions showed that multiple addition was an issue. As the desire is to apply this method to large peptides and proteins, a more representative model of these was to be used. An increase in bulk around the modified residue could also reduce the chance for multiple addition due to steric hindrance.

Due to the issue stated above with pH determination the optimisation reactions were repeated. In addition to the pH issue, as seen in Chapter 2.3.1, the use of small dipeptides can lead to DKP formation. This, along with the previously mentioned issue of methyl ester hydrolysis led to the decision to use the model peptide Ac-CWHISKEY-NH_2 (**49**) for all subsequent optimisation. This peptide was synthesised by SPPS, as described in section 7.1. After the success of the TEMPO conjugation work using $Mn(OAc)_3$, conditions based on these reactions were also trialled. The results from these trials are shown in Table 17. All reactions were performed at pH 7, using 20% DMSO in LB as the reaction solvent (Scheme 67). Ir(III) refers to $(Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$, EY to eosin Y and Mn to $Mn(OAc)_3$. Reactions using Ir(III) or EY were performed using blue LEDs, and reactions using Mn were performed at 50 ^oC.



Scheme 67 - Reaction scheme for the bioconjugation of Ac-CWHISKEY-NH_2 (**49**) with N-allyl acetamide (**66**) forming the conjugated product (**72**), using the altered conditions shown in Table 17.

Table 17 - Initial allyl conjugation reactions using the model peptide Ac-CWHISKEY-NH_2 (**41**) forming the conjugated product (**66**). Peptide – 1 mM, TCEP – 50 mM, 20% DMSO in LB at pH 7 for 2 hours.

Entry	N-allyl acetamide (**66**) eq.	Initiator eq.	Product (**72**) %conv.	Ala (**73**) %conv.	SM (**49**) %rem.	MA (**74**) %conv.
1	500	Ir(III) (0.01)	36	54	0	10
2	500	EY (0.1)	37	49	0	13
3	10	EY (0.1)	0	100	0	0
4	500	Mn (5)	28	72	0	0
5	10	Mn (5)	0	100	0	0
6	2	Mn (5)	0	100	0	0

The results in Table 17 show that the Ir(III) and EY catalysts are more suitable to this reaction than the Mn initiator. As the Ir(III) and EY catalysts provide a similar level of conversion, they will both be used for all future optimisation reactions. Factors left to optimise include peptide concentration, TCEP eq., initiator eq. and allyl eq. The first of the remaining factors varied was the peptide concentration. This was chosen as the most likely cause of Ala formation would be H-atom abstraction, which could occur due to the presence of thiol. The results from these trials are shown in Table 18.

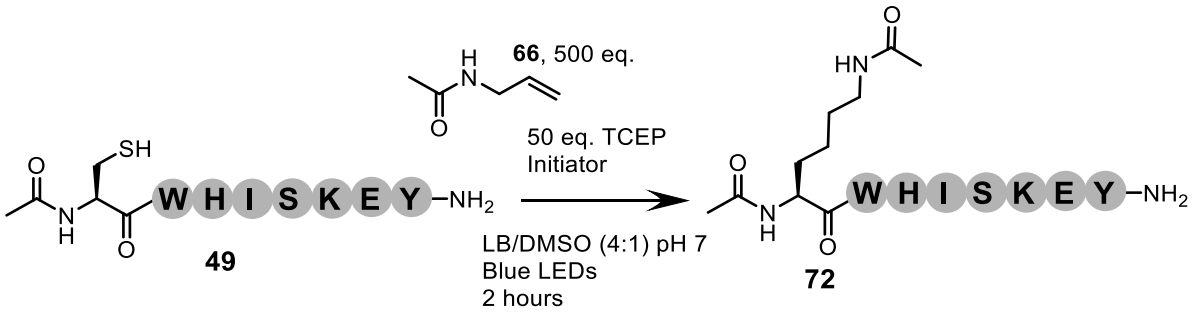


Table 18 – Peptide concentration optimisation using 50 eq. TCEP and 500 eq. N-allyl acetamide in 20% DMSO in LB at pH 7 for 2 hours, forming the conjugated product (**72**).

Entry	Peptide (**49**) conc. mM	Initiator eq.	Product (**72**) %conv.	Ala (**73**) %conv.	SM (**49**) %rem.	MA (**74**) %conv.
1	5	EY (0.1)	7	23	0	70
2	2.5	EY (0.1)	28	48	0	24
3	1	EY (0.1)	37	49	0	13
4	0.5	EY (0.1)	38	45	0	16
5	0.2	EY (0.1)	35	56	0	9
6	5	Ir(III) (0.01)	17	30	0	53
7	2.5	Ir(III) (0.01)	27	38	0	35
8	1	Ir(III) (0.01)	30	40	0	30
9	0.5	Ir(III) (0.01)	43	49	0	9
10	0.2	Ir(III) (0.01)	39	51	0	10

The results from the peptide concentration optimisation show that a more dilute reaction is required to increase the conversion of starting material to desired product. They show that decreasing the concentration from 5 mM to 0.5 mM results in an increase in the product (**72**) conversion of 7% to 38%. Diluting further from 0.5 mM to 0.2 mM reduced the amount of desired product, showing that the optimal peptide concentration of these reactions was 0.5 mM. The next factor optimised was the equivalents of TCEP present. This plays an important role, reacting with the thiyl radical and forming the alanyl radical. The results from this are shown in Table 19.

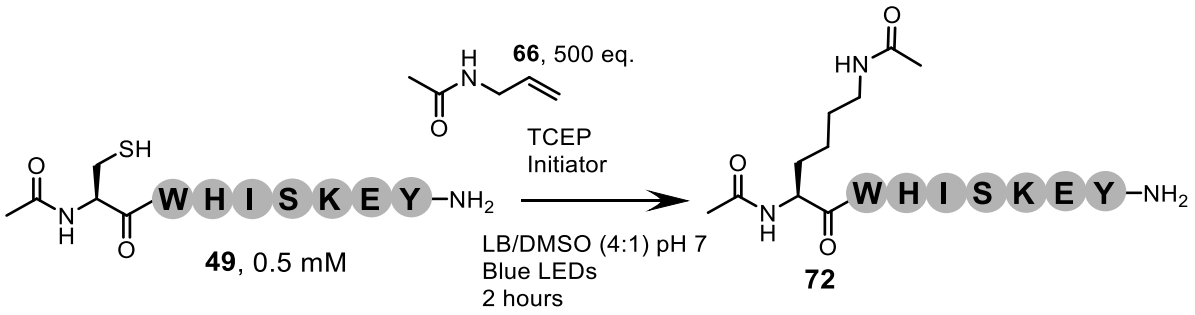


Table 19 - TCEP optimisation using 0.5 mM peptide, 500 eq. N-allyl acetamide and either 0.1 eq. eosin Y or 0.01 eq. Ir(III) in 20% DMSO in LB at pH 7 for 2 hours, forming the conjugated product (**72**).

Entry	TCEP eq.	Initiator eq.	Product (**72**) %conv.	Ala (**73**) %conv.	SM (**49**) %rem.	MA (**74**) %conv.
1	100	EY (0.1)	33	58	0	9
2	50	EY (0.1)	38	45	0	16
3	10	EY (0.1)	42	37	0	21
4	5	EY (0.1)	44	35	0	21
5	100	Ir(III) (0.01)	0	14	0	86
6	50	Ir(III) (0.01)	30	40	0	30
7	10	Ir(III) (0.01)	42	39	0	19
8	5	Ir(III) (0.01)	42	40	0	17

It was expected that, as TCEP plays a key role in desulfurisation, that increasing the equivalents of TCEP would result in a greater conversion to product. However, when these reactions were performed it was apparent that more TCEP resulted in a lower product conversion. For both eosin Y and Ir(III), 5 and 10 equivalents resulted in similar conversion, but increasing to 50 and 100 equivalents resulted in a large reduction of product. The optimal TCEP equivalents found was taken forward to the next optimisation reaction, which focussed on the pH. In aqueous conditions, pH plays a key role in reactions. Focussing on Cys alone, depending on pH this can exist either as the thiol or as the thiolate anion. The pH optimisation data is shown in Table 20.

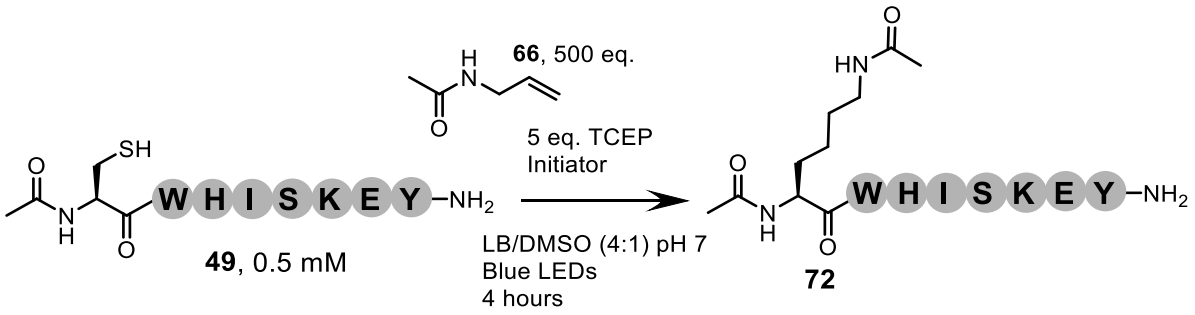


Table 20 - pH optimisation using 0.5 mM peptide (**49**), 5 eq. TCEP, 500 eq. N-allyl acetamide (**66**) and either 0.1 eq. eosin Y or 0.01 eq. Ir(III) in 20% DMSO in LB at pH 7 for 4 hours. Entry 9 reached completion within 2 hours, forming the conjugated product (**72**).

Entry	pH	Initiator eq.	Product (**72**) %conv.	Ala (**73**) %conv.	SM (**49**) %rem.	MA (**74**) %conv.
1	5	EY (0.1)	27	40	0	33
2	6	EY (0.1)	33	47	0	20
3	7	EY (0.1)	43	41	0	17
4	8	EY (0.1)	42	41	0	17
5	9	EY (0.1)	35	48	0	17
6	5	Ir(III) (0.01)	0	0	100	0
7	6	Ir(III) (0.01)	15	38	46	0
8	7	Ir(III) (0.01)	42	40	0	17
9	8	Ir(III) (0.01)	52	35	0	13
10	9	Ir(III) (0.01)	40	27	0	33

For reactions on proteins, it is useful if near-neutral pH can be used as this will not result in the protein tertiary structure being affected. Both peptides and proteins are also more stable at neutral pH, as if the pH is too acidic or basic amide bond hydrolysis can occur. The optimisation data showed that for Ir(III) a pH of 8 was optimal, whilst for EY a pH of 7 was optimal. This also was the first time that the percentage conversion reached over 50%, which was a significant result for the usefulness of this reaction. In Table 21 the equivalents of the catalysts were altered to see whether this would affect the conversion. As the catalyst plays a key role in the generation of the thiyl radical, the rate at which this is formed could play a key role.

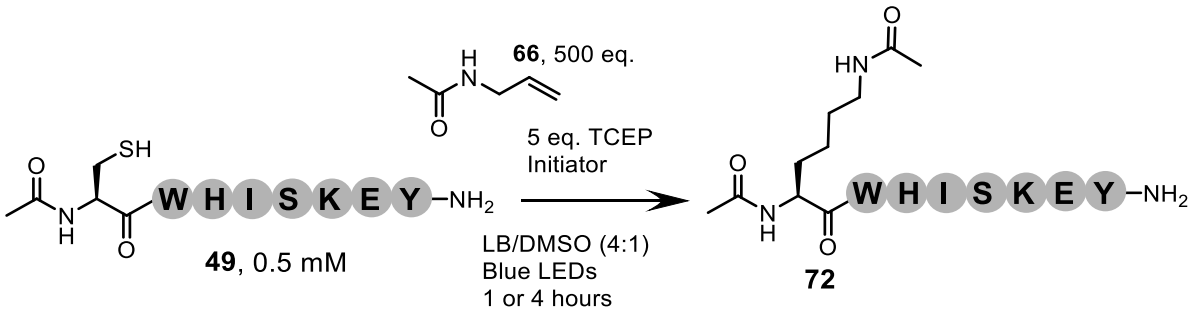


Table 21 - Initiator equivalents optimisation. Peptide 0.5 mM, 5 eq. TCEP, 500 eq. N-allyl acetamide. Eosin Y reactions at pH 7 and Ir(III) reactions at pH 8, forming the conjugated product (**66**). Solvent used was 20% DMSO in LB. Reactions analysed after 1 and 4 hours.

Entry	Initiator eq.	Product (**72**) %conv.	Ala (**73**) %conv.	SM (**49**) %rem.	MA (**74**) %conv.	Completion time
1	EY (0.1)	48	36	0	17	4 hours
2	EY (0.5)	50	32	0	18	1 hour
3	EY (0.02)	38	36	0	26	4 hours
4	Ir(III) (0.01)	48	33	0	19	4 hours
5	Ir(III) (0.05)	51	29	0	20	1 hour
6	Ir(III) (0.001)	20	21	59	0	>4 hours

This data showed that for both catalysts higher equivalents results in a higher degree of conversion. However, this was not significant as increasing the mol% of Ir(III) from 1 to 5 only increased the conversion by 3% i.e. within the error of the repetitions. Similar results were seen with eosin Y. The final factor left to optimise was the equivalents of alkene present. This has to be high due to the high reactivity and instability of the alanyl radical, to ensure reaction occurs. However, the presence of a high excess of alkene was resulting in the multiple addition side-product that was resulting in lower conversion to the desired product (**72**). The results from the alkene equivalents optimisation is shown in Table 22.

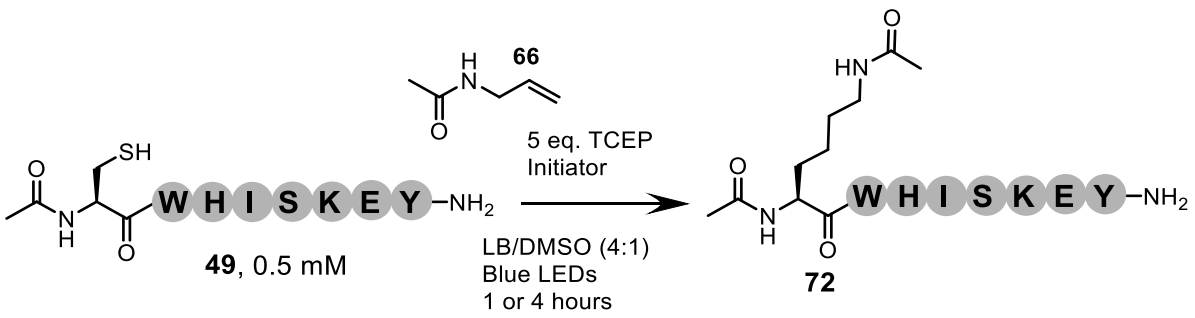
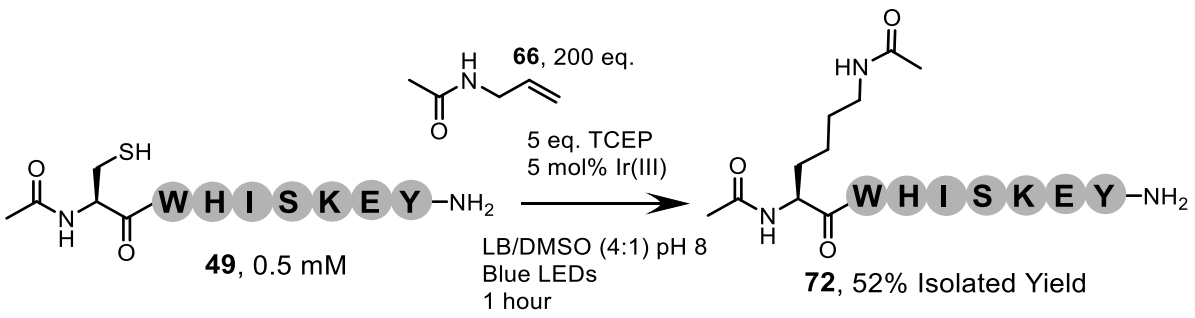


Table 22 - N-allyl acetamide eq. optimisation. 0.5 mM peptide, 5 eq. TCEP. Either 0.5 eq. eosin Y at pH 7 or 0.05 eq. Ir(III) at pH 8, forming the conjugated product (**66**). Reactions analysed after 1 hour.

Entry	N-allyl acetamide (**66**) eq.	Initiator eq.	Product (**72**) %conv.	Ala (**73**) %conv.	SM (**49**) %rem.	MA (**74**) %conv.
1	1000	EY (0.5)	38	24	14	24
2	500	EY (0.5)	48	32	0	20
3	200	EY (0.5)	45	44	0	12
4	100	EY (0.5)	38	55	0	7
5	1000	Ir(III) (0.05)	41	28	8	24
6	500	Ir(III) (0.05)	51	30	0	19
7	200	Ir(III) (0.05)	49	38	0	12
8	100	Ir(III) (0.05)	43	51	0	6

The alkene equivalents optimisation showed that a high equivalent was required for a reasonable product conversion to occur. The amount of multiple addition did correlate with the amount of alkene, with more alkene resulting in more multiple addition. There was not a significant decrease in product conversion when decreasing from 500 to 200 eq., however decreasing further to 100 did produce less desired product. Based on these optimisation reactions the optimal reaction conditions, taking into account percentage conversion and reagent equivalents, were designated – 0.5 mM peptide, 5 eq. TCEP, 200 eq. alkene, 0.05 eq. Ir(III) at pH 8 for 1 hour.



Scheme 68 - Bioconjugation of Ac-CWHISKEY-NH2 (**49**) with N-allyl acetamide (**66**) using fully optimised conditions to form product (**72**).

These conditions were applied to a large-scale reaction on Ac-CWHISKEY-NH_2 (**49**) to ascertain the overall isolated yield and reaction rate. The reaction was monitored at 15 minute intervals and was shown to reach completion after 45 minutes. At this time, the conversion of starting material to product (**72**) was 55%. The remainder was 37% Ala (**73**) and 7% multiple addition product (**74**). The yield of desired product (**72**) isolated was 52%.

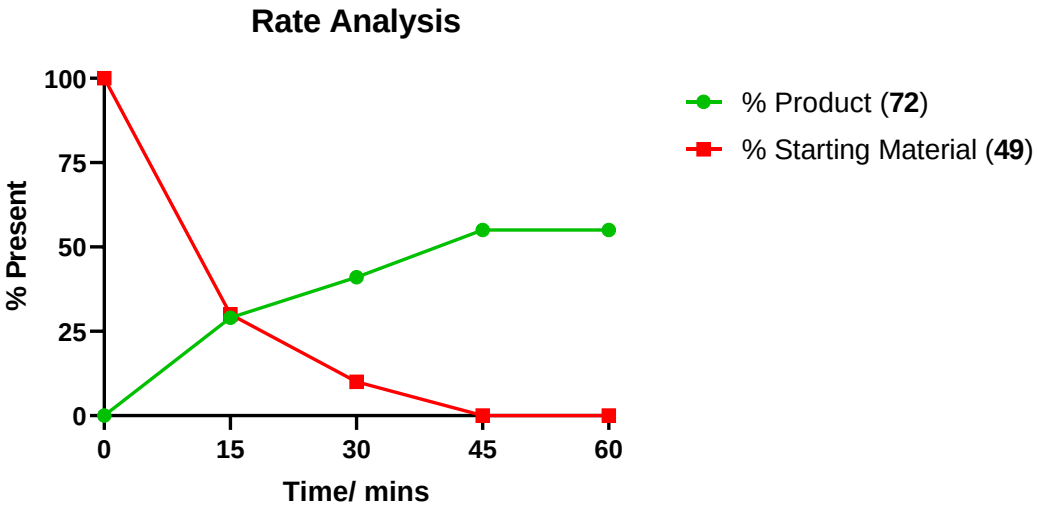


Figure 55 - Rate analysis for the bioconjugation of Ac-CWHISKEY-NH_2 (**49**) with N-allyl acetamide (**66**) to form product (**72**), with time points every 15 minutes.

The data so far shows the reaction has been significantly optimised in comparison with the starting conversions seen. Due to the employment of an non-activated alkene, the interception of the alanyl radical with these traps is not able to fully outcompete H-atom abstraction. However, considering native modifications can be installed using this method at moderate yield, we were content with this level of optimisation. Therefore, moving forward, it was important to ensure that racemisation of the conjugated alpha centre was not occurring. To do this, two small model systems were synthesised (as described in section 7.1), Ac-CAY-NH_2 (**75**) and Ac-D-CAY-NH_2 (**76**). The conjugation was performed using the optimised conditions

(0.5 mM peptide (**75** or **76**), 2.5 mM TCEP, 200 mM N-allyl acetamide (**66**), 0.025 mM Ir(III), in LB at pH 8 for 1 hour). The resulting structures (Figure 56) and NMR spectra (Figure 57) are shown below.

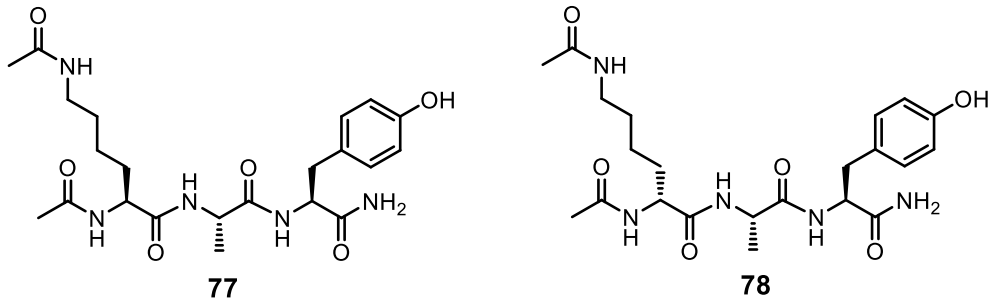


Figure 56 – Chemical structures of the isolated products **77**, L and **78**, D.

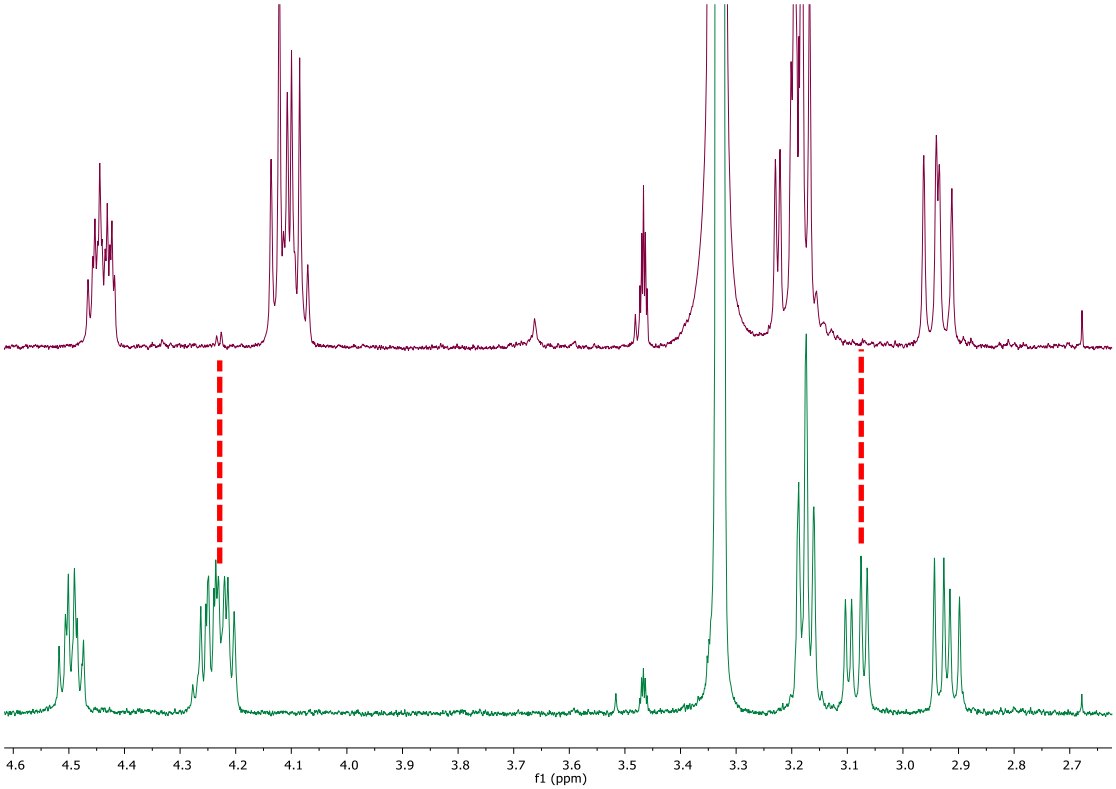
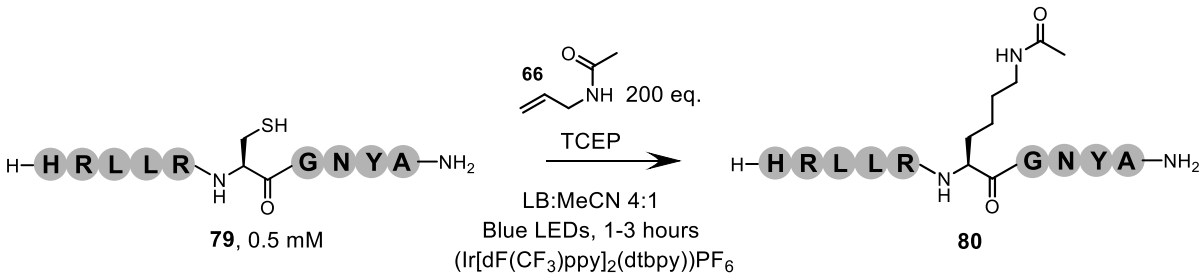


Figure 57 – NMR comparison of conjugated L (**77**, bottom) and D (**78**, top) Ac-CAY-NH_2 model system with N-allyl acetamide (**66**).

The NMR comparison clearly shows key differences between the spectrum for the L (**77**) and D (**78**) conjugated peptides. The absence of certain peaks without peak broadening shows that racemisation of the conjugated amino acid is not occurring. This was vital as it shows that the native Lys PTM is being formed.

4.4 Substrate scope development on larger model systems

After confirming that racemisation was not occurring, the optimised conditions were taken forward to a more relevant model. The chosen model for substrate scope analysis was H-HRLLRCGNYA-NH_2 (**79**), synthesised as described in section 7.1. This model represents a small section from a histone H2A protein; histones undergo extensive Lys PTM which mediates the structure of chromatin and therefore access to genetic information. This model also allowed us to investigate the conjugation of an internal Cys residue, a move which created issues with the TEMPO trapping reaction in Chapter 2. Initial reactions were performed on this model system, including the previously optimised conditions, as shown in Figure 58, to form conjugated product (**80**).



Scheme 69 - Reaction scheme for the conversion of H-HRLLRCGNYA-NH_2 (**79**) to the corresponding acetyl Lys baring peptide (**80**), varying the equivalents of TCEP and Ir(III).

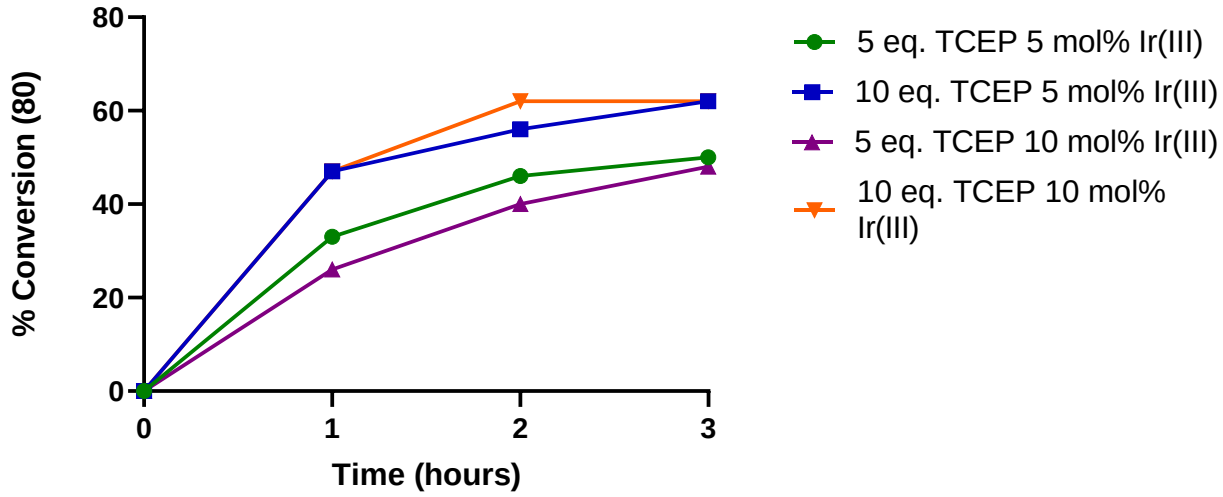
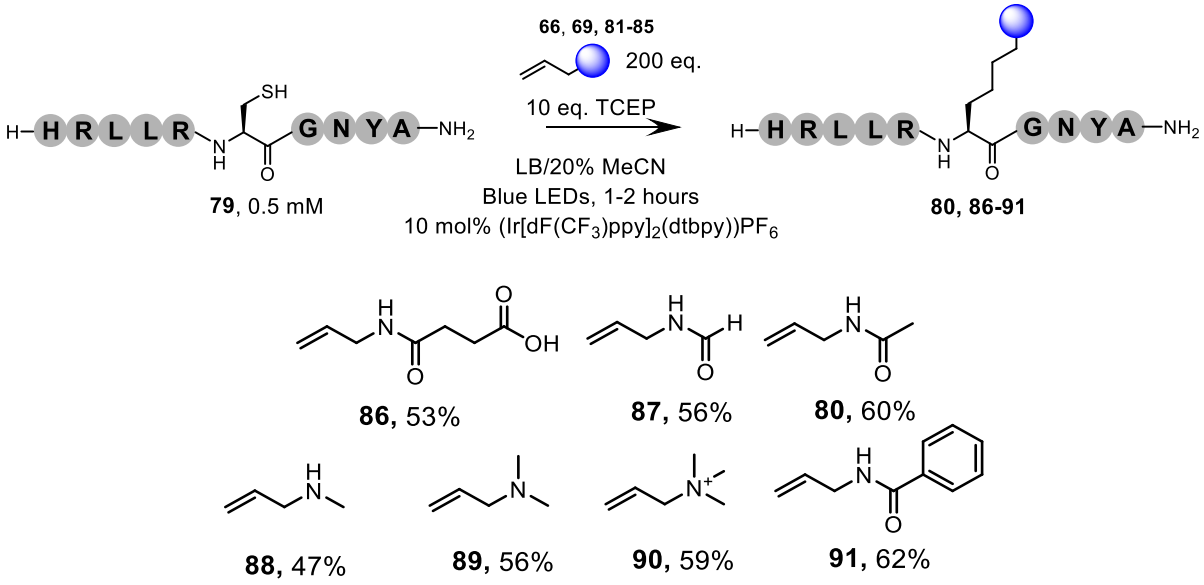


Figure 58 - Time point graph for the initial trials on the H-HRLLRCGNYA-NH_2 model, varying the equivalents of TCEP and Ir(III), showing the conversion of (**79**) to (**80**).

The initial trials on the small histone model show that, as with the TEMPO reaction, internalising the Cys residue required a slight modification to the reaction conditions. In this instance, increasing the TCEP eq. from 5 to 10 eq. and the Ir(III) eq. from 5 mol% to 10 mol%

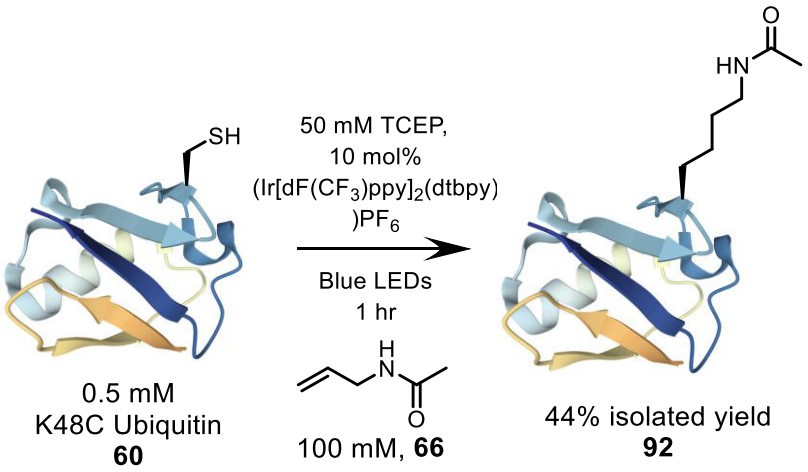
led to similar conversion as shown with the Ac-CWHISKEY-NH_2 model in 2 hours. Using these conditions, the organic co-solvent was also evaluated. The DMSO could be exchanged for methanol or acetonitrile with no decrease in conversion observed.

Using the most optimal conditions (General protocol G, section 7.2, 0.5 mM peptide, 10 eq. TCEP, 200 eq. allyl substrate, 10 mol% Ir(III), LB:MeCN 4:1, pH 8 for 1-2 hours), the substrate scope of the reaction was tested. A range of substrates were either purchased and used as received, or synthesised from commercially available allyl amine or allyl chloride. The chosen substrates were N-allylsuccinimide (synthesised by Dr Frances Smith, NASA, **81**), N-allyl formamide (**82**), N-allyl acetamide (**66**), N-methylallylamine (purchased, MMAA, **83**), N,N-dimethylallylamine (purchased, DMAA, **69**), allyltrimethylammonium chloride (**84**) and N-allylbenzamide (**85**). These were either purchased or synthesised using common, well-established chemistry. Common reaction used were amidation of allyl amine (**66**, **81**, **82**, **85**) and nucleophilic displacement (**84**). Details of the reactions performed are in section 7.3. The general reaction scheme, allyl compounds synthesised and resulting isolated conjugation yields are shown in Scheme 70.



Scheme 70 - General scheme for the conjugation of H-HRLLRCGNYA-NH_2 (**79**) with a range of allyl compounds. The yields shown are isolated yields.

The data shows that the developed method can successfully install a wide range of native Lys PTMs onto a peptide model in moderate-to-good isolated yield. The final challenge for this method was to apply it to a protein model. For this, the K48C mutant of Ub (**60**) was again chosen. The Lys at position 48 in endogenous Ub is acetylated, this PTM mediates the formation of poly-Ub chains and therefore controls the activity of this protein to label other protein substrates for degradation.[160] The scheme for this reaction is shown in Scheme 71.



Scheme 71 - Reaction scheme for the installation of acetyl lysine into K48C ubiquitin (**92**).

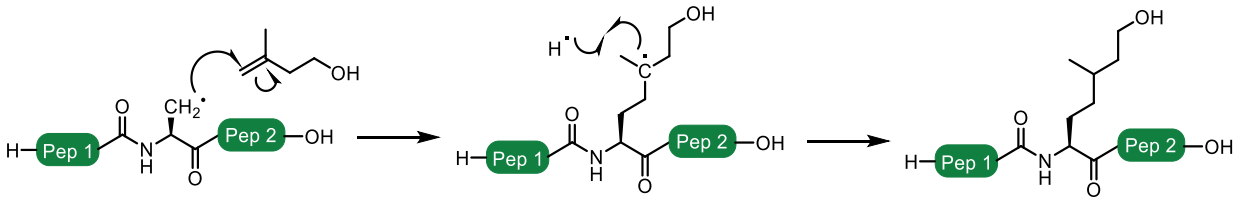
The method developed in this chapter was successfully applied to the installation of a native Lys PTM into a protein model (HPLC and MS data in section 7.6). In this case, the Cys at position 48 of the K48C mutant was exchanged for an acetylated Lys residue. To the best of our knowledge, this is the first chemical method that can successfully install fully native Lys PTMs into proteins in a site-selective manner without compromising the stereochemistry at the target residue. Full NMR characterisation of the folded product, and associated biological assays to demonstrate native activity, are currently underway. This result shows that the method can be applied to protein models, making this a useful technique for chemical biologists probing the roles of native PTMs in proteins.

4.5 Cysteine bioconjugation through C-C bond formation via a stabilised radical intermediate

The chemistry described in sections 4.2 and 4.3 not only shows a novel method for installing common Lys PTMs into peptides and proteins, but also provides a novel procedure for C(sp^3)-C(sp^3) bond formation. An advantage of bioconjugation through C-C bond formation is the unparalleled stability of the resulting conjugate to biological conditions. The aim of this section was to further extend this chemistry.

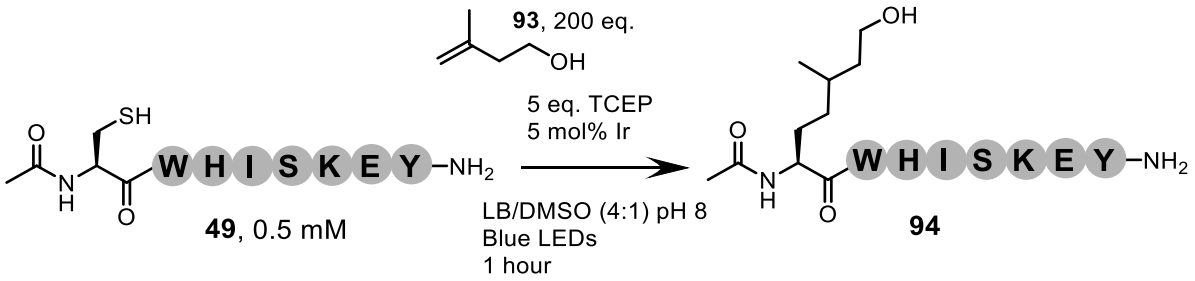
The main issue with the chemistry developed so far in this chapter is the moderate yields observed. There were two key reasons for the loss of yield, the first being the multiple addition observed forming branched products. The second involves the rapid H-atom abstraction of the alanyl radical, competing with the radical addition into the unactivated alkenes, forming Ala. This occurred so readily because once the alanyl radical donates into the allyl group, a non-sterically hindered, unstabilised secondary radical is formed. This unstable radical is able to react with the excess of allyl compound in solution. This led to alternate alkene groups being studied. We decided to trial the methyl-alkene traps, as the methyl group would provide steric

bulk to limit multiple addition, and also result in a stable tertiary radical that should enable the desired route to out-complete H-atom abstraction (Scheme 72).



Scheme 72 - Mechanism for the reaction between the alanyl radical and 3-methyl-3-buten-1-ol (**93**), producing a stabilised tertiary radical intermediate.

The commercially available compound 3-methyl-3-buten-1-ol (**93**) was therefore initially trialled as a potential alternative to the allyl motif as the radical formed after addition would be a stabilised, sterically hindered tertiary radical; the scheme for this reaction is shown in Scheme 73.



Scheme 73 - Reaction scheme for the bioconjugation of the model peptide Ac-CWHISKEY-NH_2 (**49**) with 3-methyl-3-buten-1-ol (**93**) forming the conjugated product (**94**).

The reaction was performed in the same manner as previously shown on the Ac-CWHISKEY-NH_2 (**49**) model peptide when using N-allyl acetamide. Instead of multiple impurity peaks previously observed by HPLC of the reaction mixture, the analysis of this reaction mixture showed only a major and minor product. The two products were isolated; the minor product was shown to be the desulfurisation product (Ala, **73**), and the major product was the desired conjugate (**94**). Based on HPLC conversion of the crude trace shown in Figure 59, the amount of product (**94**) present was 83% compared to only 17% of Ala product (**73**). The reaction was repeated on an isolatable scale, with the reaction resulting in a 79% isolated yield (**94**). This is significantly higher than the isolated yield when using allyl compounds, so this motif will be taken forward for further experimentation.

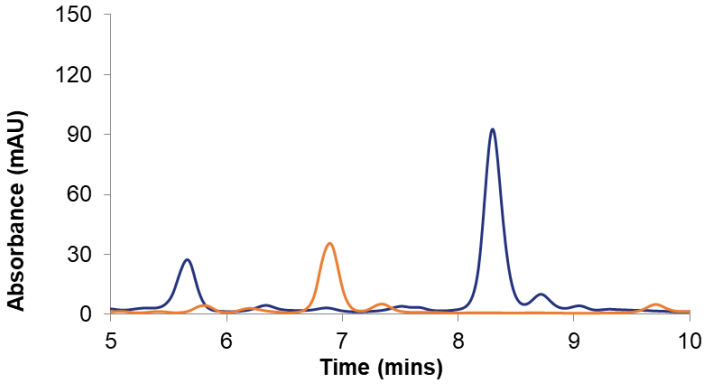
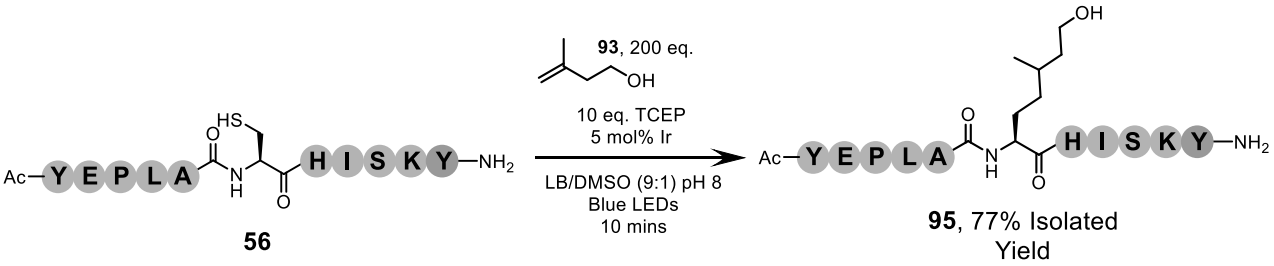


Figure 59 - Crude HPLC trace for the modification of Ac-CWHISKEY-NH_2 (**49**) with 3-methyl-3-buten-1-ol (**93**). Pure SM (**49**) in orange and crude reaction mixture trace forming product (**94**) in blue. Peak at 5.5 minutes is Ala by-product (**73**), and the peak at 8.2 is desired product (**94**). Analytical HPLC gradient 15-25 over 10 mins, 280 nm.

The greater conversion to desired product (**94**) and cleaner reaction mixture ensures that this reaction will be highly suitable for peptide and protein bioconjugation. To ensure full optimisation of this method, the equivalents of alkene required should ideally be lower than 200. However, attempts to reduce the alkene equivalents were, unfortunately, unsuccessful. Even a reduction from 200 eq. to 100 eq. results in a decrease in conversion from 83% to 61%. Reducing further to 50 eq. and 20 eq. results in a 51% and 30% conversion to product (**94**) respectively. As issues with modifying internal Cys residues have occurred with the previously developed reactions, the current optimal conditions (Scheme 74) were applied to Ac-YEPLACHISKY-NH_2 (**56**). The TCEP equivalents were increased to 10, as increasing the amount of TCEP for internalised Cys residues was previously found to be necessary.



Scheme 74 – Conjugation of peptide (**56**) with MBO (**93**) to form desired product (**95**).

The reaction was analysed after only 10 minutes and was shown to have formed 82% product, with no starting peptide (**56**) remaining. The desired product (**95**) was isolated in a 77% yield. These conditions will be used for substrate scope screening later in this chapter, being referred to as general protocol I (details in section 7.2). For protein bioconjugation, the lower the concentration the simpler the reaction is to perform due to the small amount of material available. The results in Figure 60 show that lowering the concentration to 0.05 mM required both the alkene and TCEP equivalents to be increased.

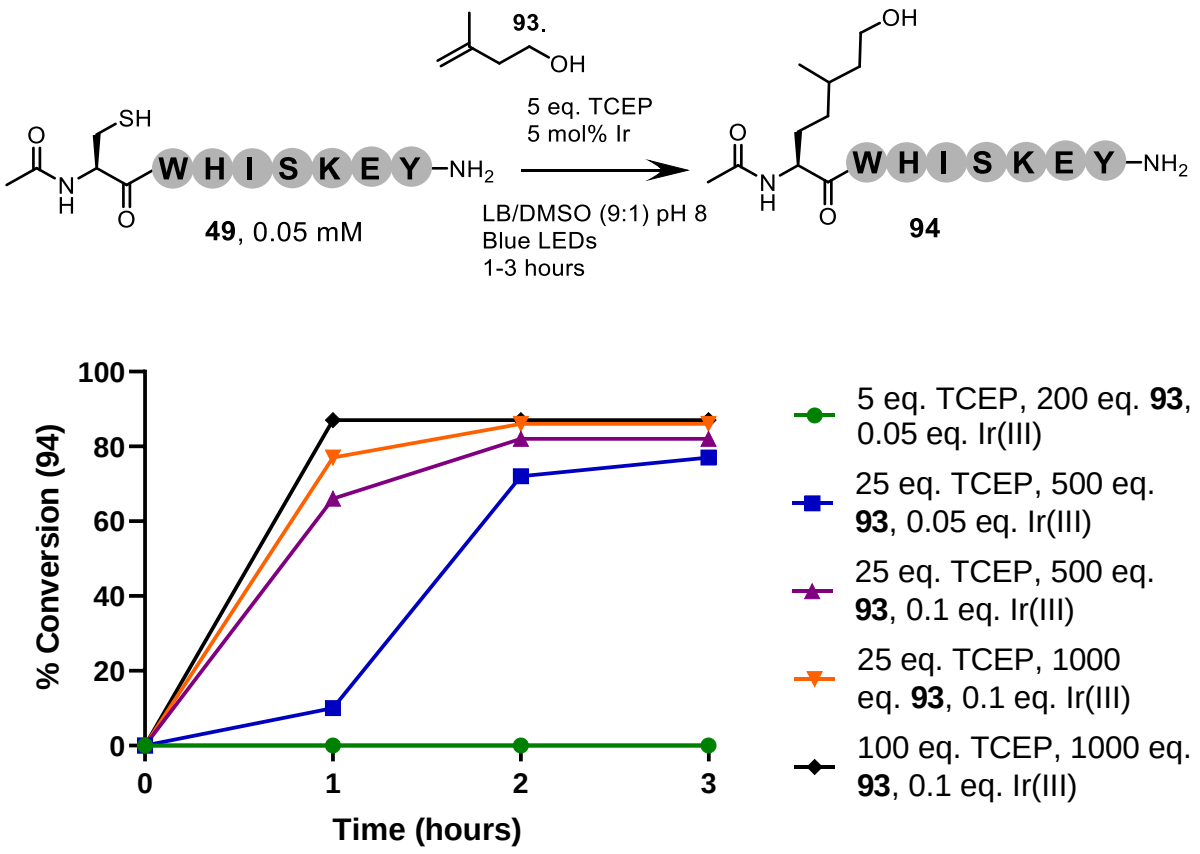
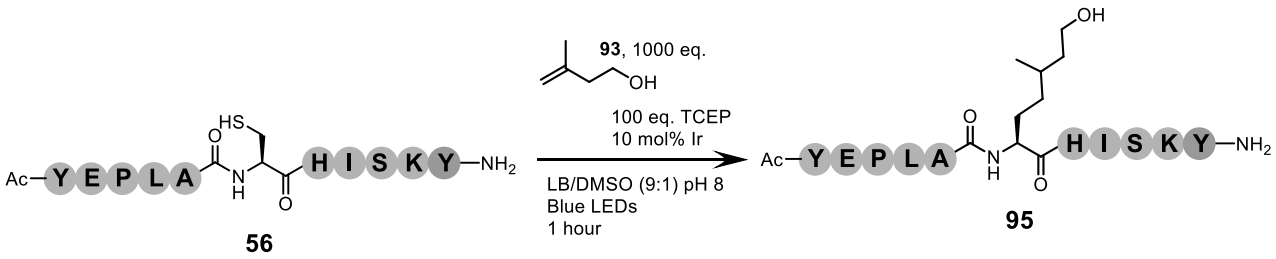


Figure 60 - Low concentration trials on Ac-CWHISKEY-NH_2 (**49**) to assess protein applicability.

This data demonstrates that although lower concentrations are possible, this does mean the equivalents of alkene and TCEP do need to be increased. When using large molecular weight proteins on small scale, this is not a substantive drawback, as the amount of comparatively low molecular weight alkene required would be minimal. The optimal conditions (100 eq. TCEP, 1000 eq, **93**, 0.1 eq. Ir(III)) were repeated on the Ac-YEPLACHISKY-NH_2 (**56**) model to test the effect of Cys internalisation, giving conjugated peptide (**95**). The potential of these conditions for lower concentration reactions were also trialled. This would be an important result for the applicability of this reaction to protein substrates. The data from these trials (Scheme 75) are shown in Table 23.



Scheme 75 - Conjugation of Ac-YEPLACHISKY-NH_2 (**56**) with 3-methyl-3-buten-1-ol (**93**) to form conjugated product (**95**).

Table 23 - Low concentration studies using Ac-YEPLACHISKY-NH_2 (**56**) with 100 eq. TCEP, 1000 eq. 3-methyl-3-buten-1-ol (**93**) and 10 mol% Ir(III) in 10% DMSO in LB at pH 8, forming product (**95**). Values given in brackets is the percentage of starting material remaining.

Entry	Peptide (**56**) µM	1 hour %conv. (**95**)
1	50	82 (0)
2	25	75 (8)
3	10	25(57)
4	5	0 (100)
5	1	0 (100)

The data shows the reaction could be performed at 0.025 mM, but lower concentrations were not possible. The conditions from Entry 1 (Table 23) were also repeated using PB instead of LB. As the guanidine present in LB results in protein denaturing, it would be useful if this were not required as proteins could be modified without the need for refolding. This was successful, with a conversion of 80% observed after 1 hour. The reaction was repeated on an isolatable scale using the optimised conditions from general protocol I (details in section 7.2) to give the desired conjugated peptide (**95**) in a 77% isolated yield.

As with the previously developed conjugation techniques, it was important to ensure that the stereocentre of the modified amino acid remained unchanged. This was again proven through NMR analysis of the modified L (**96**) and D (**97**) versions of Ac-CAY-NH_2 (**75** (L) and **76** (D)).

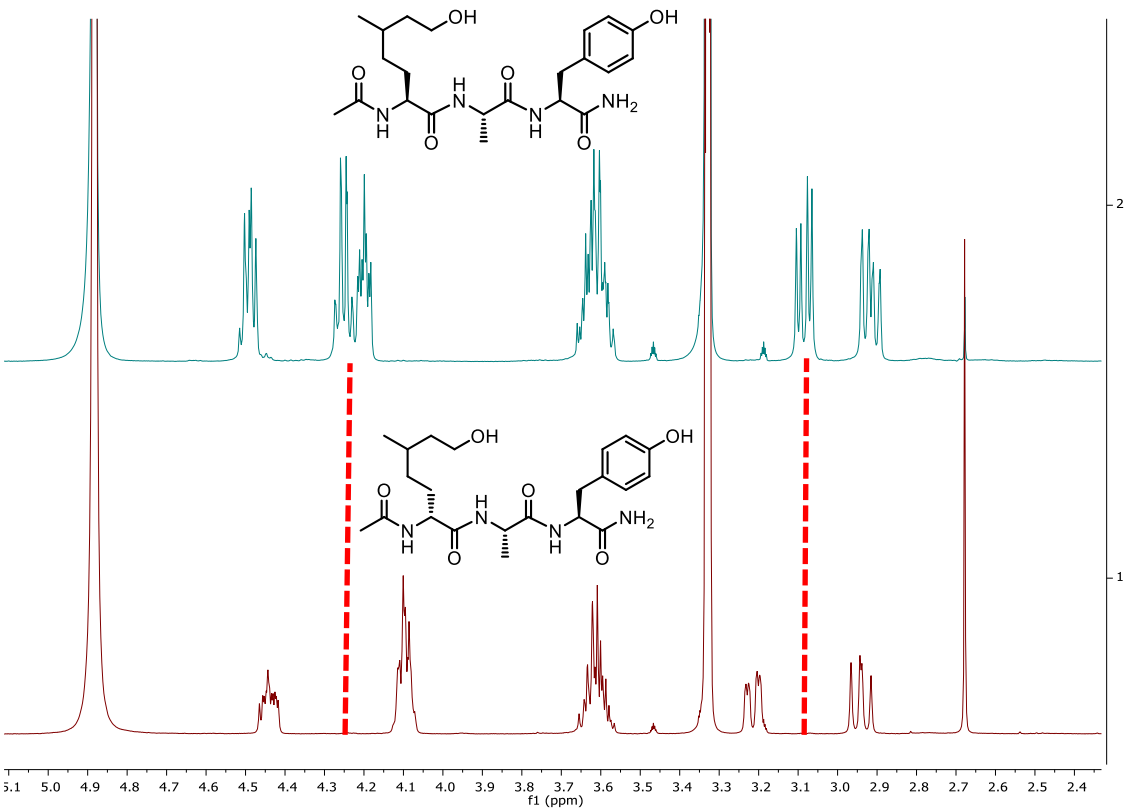
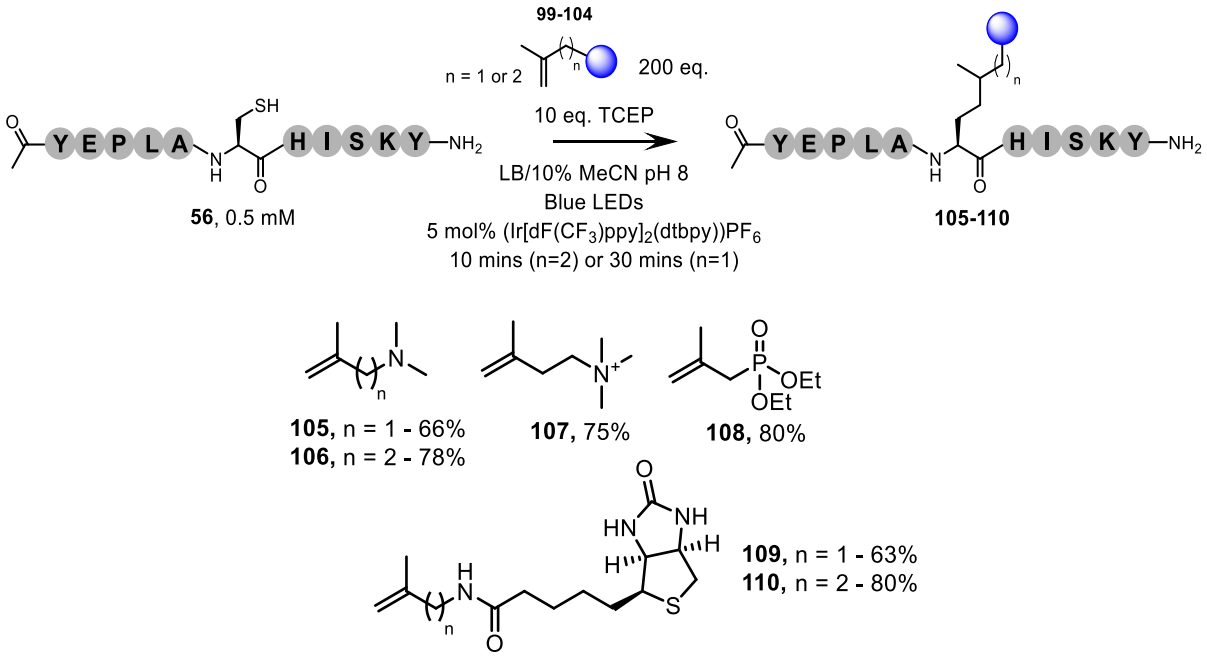


Figure 61 - NMR comparison between modified Ac-L-CAY-NH_2 (**96**, top) and Ac-D-CAY-NH_2 (**97**, bottom) conjugated with 3-methyl-3-butene-1-ol (**93**).

The next test for this conjugation technique was substrate scope. As only the 3-methyl-3-butene-1-ol is commercially available, this had to be converted to the amine (**98**) for ease of functionalisation (experimental details outlined in section 7.3). Once converted it was utilised in either an amidation or nucleophilic displacement reaction (experimental details in section 7.3). The model system Ac-YEPLACHISKY-NH_2 (**56**) was used to assess the substrate scope of this reaction. During this assessment, the one carbon shorter 2-methylallyl conjugates were also trialled. These conjugates required a longer reaction time (30 mins). The advantage of these reagents is that the amine and aldehyde are commercially available. Another advantage is that they are closer mimetics to the native Lys, only varying by a single methyl group. This means that these conjugates have a greater likelihood of being successful mimetics for Lys PTMs. The optimised conditions with a peptide concentration of 0.5 mM were chosen, using 10 eq. TCEP, 200 eq. alkene and 5 mol% Ir(III) catalyst. These were chosen over the 0.05 mM conditions as a lower amount of alkene and TCEP is preferable. The probes trialled in these reactions were the dimethylated 2-methylpropene (**99**) and 3-methylbutene (**100**), trimethylated 3-methylbutene (**101**), phosphorylated 2-methylpropene **102**, commercially available), and biotin modified 2-methylpropene (**103**) and 3-methylbutene (**104**) (synthesised

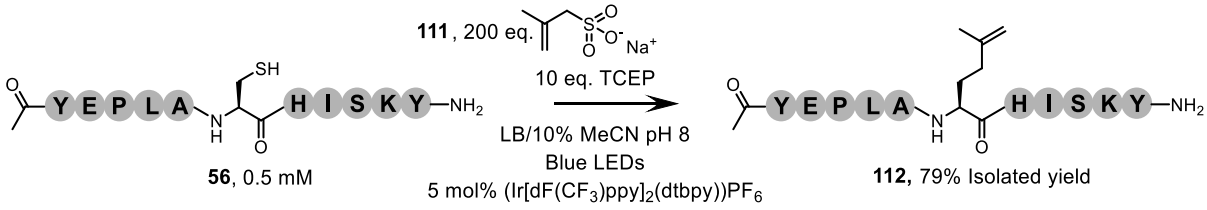
by Dr Frances Smith). The reaction scheme and substrate isolated yields for the performed conjugations are shown in Scheme 76.



Scheme 76 - Overview of the reaction of Ac-YEPLACHISKY-NH_2 (**56**) with a range of stabilised alkene substrates (**98-104**) to form desired conjugates (**105-110**). The yields specified are isolated yields.

These yields show that the developed method could be used to install biologically relevant functionality into the model peptide in moderate to high isolated yields. This data is very promising and shows the potential for the method to be further developed to the conjugation of a wide range of functionality into peptides via a highly stable C(sp^3)-C(sp^3) bond.

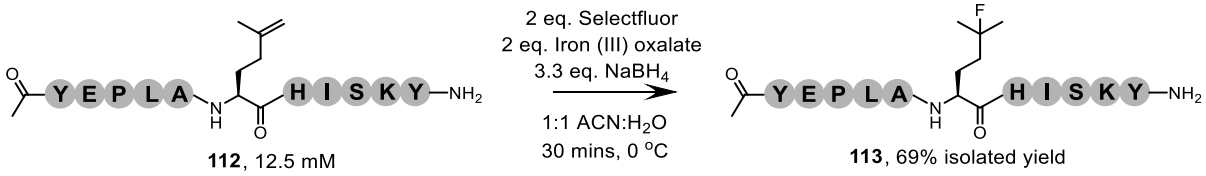
Interestingly, when the commercially available 2-methyl-2-propene-1-sulfonic acid sodium salt (**111**) was trialled the desired sulfonic acid conjugated peptide was not observed. Based on the molecular weight observed in the MS trace (shown in section 7.7), it was most likely the elimination product (**112**) which was isolated in a 79% yield (Scheme 77).



Scheme 77 - The reaction scheme showing the unexpected product when Ac-YEPLACHISKY-NH_2 (**56**) was conjugated with 2-methyl-2-propen-1-sulfonic acid sodium salt (**111**), forming the elimination product (**112**).

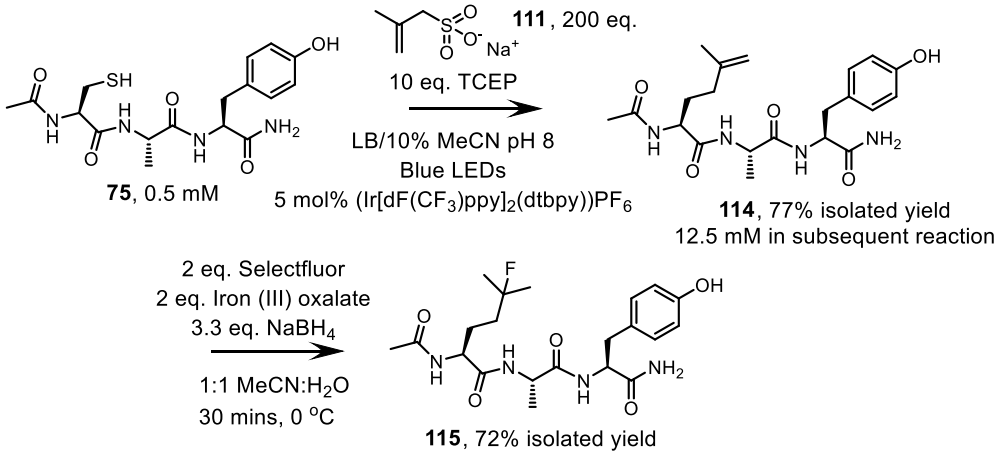
As an original aim of this project was to develop a novel method for the site-specific radiolabelling of peptides and proteins, this offered the potential for hydrofluorination. The first

method trialled used conditions reported in literature.[170] This was performed on the above product, using conditions as stated in the publication (Scheme 78, to form product (**113**). The results are shown in Scheme 78.



Scheme 78 – Reaction scheme for the hydrofluorination of peptide **112** using literature conditions as shown, forming the hydrofluorinated product (**113**) in a 69% isolated yield.

This was a very promising isolated yield for a fluorination reaction on a peptide model, and, importantly, with a short reaction time. The speed of the reaction is vital when using radioactive fluorine (^{18}F) due to its short half-life (109 minutes). To ensure the reaction was proceeding as predicted, it was repeated on the small model Ac-CAY-NH_2 system for NMR analysis (Scheme 79), with the resulting NMR data shown in Figure 62.



Scheme 79 - Reaction scheme with conditions and isolated yields for the conjugation (**114**) and hydrofluorination (**115**) of Ac-CAY-NH_2 (**75**).

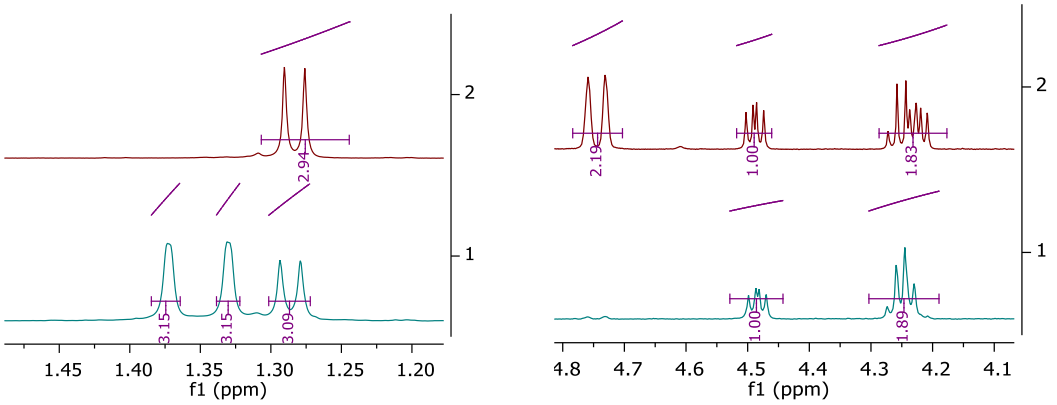


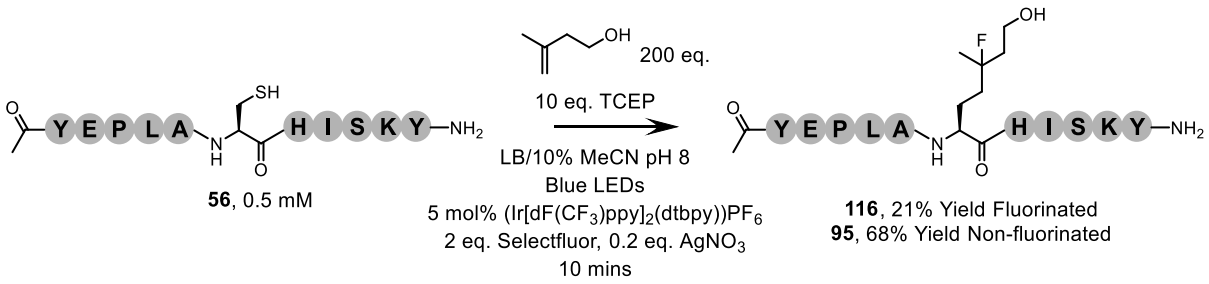
Figure 62 - NMR overlay of the starting alkene (**114**, top) and the hydrofluorinated product (**115**, bottom). The left expansion shows the disappearance of the alkene protons at 4.75. The right expansion shows the appearance of two methyl groups, showing fluorination occurred on the internal carbon of the alkene.

The NMR data in Figure 62 showed that the sulfonic acid elimination was occurring, and that the resultant alkene (**114**) could be utilised for hydrofluorination. The ease of site-selective alkene installation, combined with the published rapid hydrofluorination conditions, provide an efficient and high yielding route to fluorinated peptides. This is a very useful method of installing a terminal alkene site-selectively, as the sulfonic acid used is highly water-soluble. Compared with other radiolabelling methods that use a prosthetic group, this method utilises a minimal biogenic homo-leucine mimetic group, meaning the effect of this on a peptide or protein's tertiary structure is likely to be minimal.

An alternative novel fluorination method was also briefly explored; an advantage of using these alkene substrates that form stabilised tertiary radicals after alanyl radical addition is the prevention of multiple addition products. A publication from 2017 showed the use of Selectfluor and silver nitrate to fluorinate alkyl radicals generated from aliphatic decarboxylation.[171] As the method developed here also produces an alkyl radical, the conditions should be applicable. Another paper, released in 2020, generates an alanyl radical from modified Cys, followed by alkylation with activated alkenes (Scheme 21).[81] This chemistry is performed on a fully protected, single amino acid scale. The mechanism proposed in this paper suggests that once the alanyl radical donates into the alkene, a secondary radical is formed as expected. They then propose that this radical is reduced by the Ir(II) to form a carbanion and Ir(III), thus regenerating the catalyst. This carbanion is then quenched, either by gaining an H^+ or by losing an F^-. Based on this mechanism, fluorination could occur through an electrophilic fluorine donor.

Due to time constraints, limited work was undertaken in this area. The first reaction performed used the Ac-YEPLACHISKY-NH_2 model, as shown in Scheme 80. The standard alkene

trapping conditions were combined with the alkyl radical fluorination conditions and the reaction mixture irradiated for 10 minutes.



Scheme 80 - Reaction scheme showing the attempted fluorination reaction by trapping the intermediate tertiary radical formed after alanyl radical addition into the stabilised alkene.

These results show that the method has the potential to be used as a novel site-selective fluorination technique if further optimisation is possible. The reaction was repeated on the smaller model system Ac-CAY-NH_2 using 10 equivalents of Selectfluor instead of 2. This time the only product observed was oxidised starting material. This indicated that increasing the amount of Selectfluor might not be a useful option, as Selectfluor is known to be oxidising, out-competing the reducing properties of TCEP. In an attempt to overcome this issue, an alternative fluorinating reagent was sought. A literature search showed Selectfluor has been proven to possess a high rate constant when compared to other fluorinating agents. For comparison, a fluorinating reagent with a low rate constant was also trialled. The fluorinating reagent used, 1-fluoro-2,4,6-trimethylpyridinium tetrafluoroborate, unfortunately was also not successful. This led to clean conversion, but MS analysis of the major peak showed a product with an observed mass of 516.2264 $[M]^+$. This did not match any expected products, however the only logical outcome with that molecular weight was the structure shown in Figure 63, with a calculated mass of 516.2280 $[M]^+$.

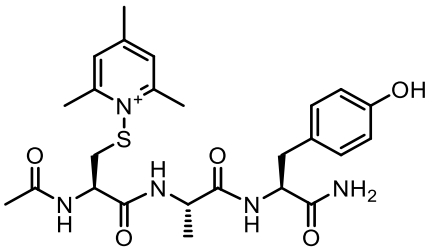
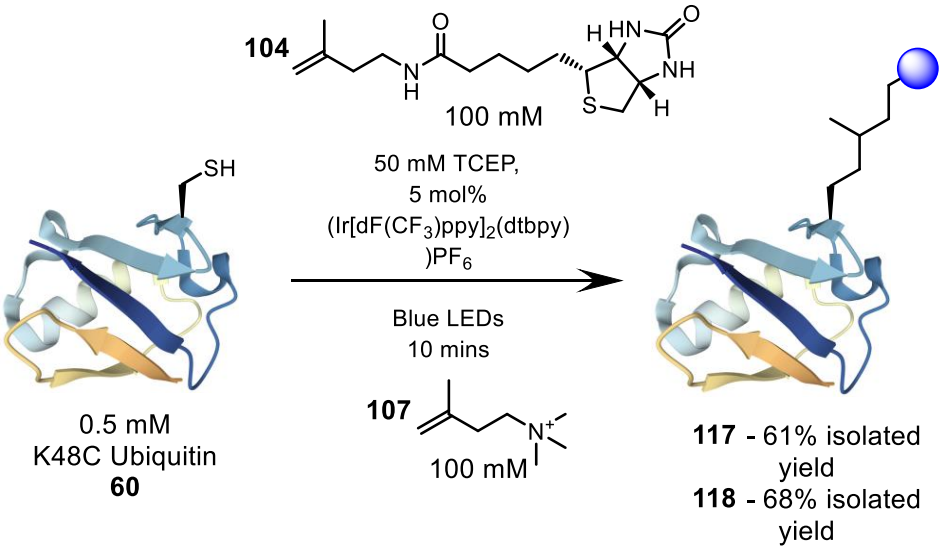


Figure 63 – Proposed structure from the reaction of Ac-CAY-NH_2 with 1-fluoro-2,4,6-trimethylpyridinium tetrafluoroborate.

This reagent could potentially provide an interesting arylation method for Cys, but unfortunately was not useful for the desired outcome of fluorination. This method therefore requires a great deal of optimisation before it can be considered a useful method for

fluorination and radiolabelling. The results from Scheme 80 do, however, show the potential for utilising the tertiary radical/carbanion produced from alanyl radical addition into the alkenes, and the potential high impact of this site-selective fluorination method makes further optimisation a priority.

The final area explored for this conjugation method was the installation of desired functionality into protein systems. The ability of this technique to install both Lys PTM mimetics and other biologically useful groups was tested. In Chapter 2, when applying the TEMPO conjugation method to the protein model, higher equivalents of TCEP was required when compared to the peptide models. Due to time constrains and limited protein material, this was assumed the case for this method. As TCEP is cheap and readily available, having a large excess was not considered an issue. All other conditions, including reaction time, were kept consistent with those seen for the peptide conjugation reactions. The outcome of these protein bioconjugations are shown in Scheme 81.



Scheme 81 - Summary of the bioconjugation of K48C ubiquitin with stabilised alkene probes containing biotin and a trimethyl-lysine mimetic.

The application of this technique to the bioconjugation of proteins was successful, with good isolated yields observed for the substrates trialled. The operational simplicity, short reaction times and ease of substrate synthesis makes this a very useful technique for installing a range of biologically relevant functionality and Lys PTM mimetics into proteins. After this success,

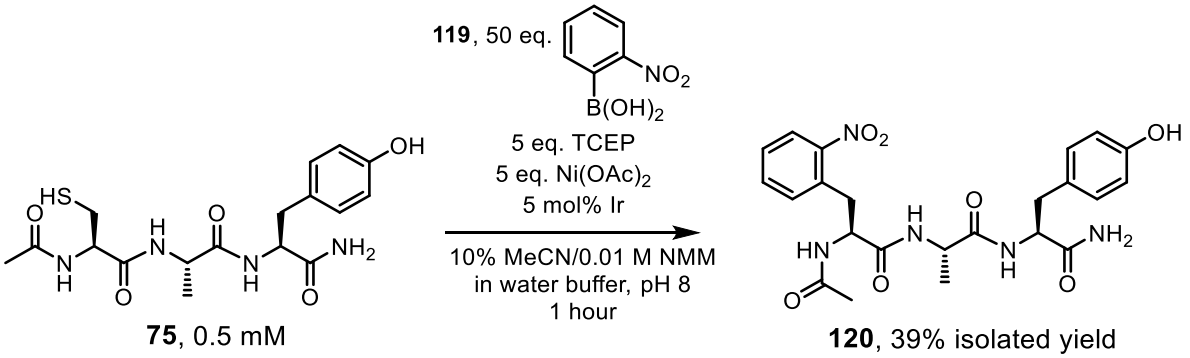
future work for this chapter will involve testing the method against alternative proteins and installing other functionality into these proteins.

4.6 Cysteine arylation through desulfurative cross-coupling chemistry

Chapter 1, section 1.4.4 showed how a range of metal-mediated arylation and cross-coupling techniques have been developed in recent years. An interesting reaction already discussed utilises nickel(II) compounds ($Ni(OAc)_2$) and aromatic boronic acids (Scheme 16). The advantage of this over other techniques is that the nickel compounds used are not air sensitive, as they are already in the oxidised +2 oxidation state. This allows reactions to be performed open to air.

In this section, the aim was to explore whether this chemistry could be applied to the alanyl radical formed during radical-induced desulfurisation. A potential issue when compared to the reaction using a thiol included having to outcompete the thiol cross-coupling reaction. In addition, H-atom abstraction by the alanyl radical to form Ala has been an unavoidable issue with the chemistry developed in this chapter so far.

The initial reaction was performed on the Ac-CAY-NH_2 (**75**) model peptide using conditions similar to those used in this chapter. The reaction was performed in 0.01 M NMM in H_2O/MeCN (9:1, pH 8) with a peptide concentration of 0.5 mM, using 5 eq. TCEP, 5 eq. $Ni(OAc)_2$, 50 eq. 2-nitrophenylboronic acid (**119**, commercially available) and 5 mol% Ir(III) catalyst for 1 hour. The traditional ligation buffer could not be used as the nickel salt was not soluble. The results from this reaction are shown in Scheme 82.



Scheme 82 - First trial desulfurative cross-coupling combining the conditions found in literature with the novel desulfurisation conditions used previously in this chapter to form product (**120**).

Upon reaction completion and purification, the desired product was isolated with a 39% yield. The product was analysed by MS and NMR to ensure the desired product was being formed. This moderate conversion to the desired product was interesting. When performed without

$Ni(OAc)_2$, only starting material was observed. Interestingly, the main peak in the crude reaction mixture was not peptide related. NMR analysis showed only aromatic protons and analysis by MS showed an $[M+H]^+$ of 231.0763. This product is formed simply by combining TCEP and 2-nitrophenylboronic acid; no catalyst or light is required. To understand this reaction fully, further studies are required to ascertain the nature of this product. Unfortunately, due to time constraints no further development of this chemistry could be performed during this project. The development of this Cys arylation reaction will make for very interesting future work and optimisation will continue within the group.

4.7 Conclusion

The data in this chapter demonstrates that the alanyl radical generated from radical desulfurisation of Cys can be trapped using alkenes, forming a novel carbon-carbon bond. In the first half of this chapter, this reaction is used to install native Lys PTMs into peptides and proteins. The reaction is moderately yielding and has a wide substrate scope, installing seven different native PTMs into a model peptide. An advantage of this method over previously published methods is that the stereochemistry of the alpha centre is retained.[87,166,167] To the best of our knowledge, this is the only known chemical method to facilitate the installation of fully native Lys PTMs site-selectively into unprotected peptides and proteins without epimerisation of the target residue.

The later work in this chapter uses a stabilised alkene to provide high yielding and rapid installation of a range of biologically relevant functionality and Lys PTM mimetics into peptides and proteins. The alanyl radical donates into the methyl-substituted alkenes and a stabilised radical is formed. This prevents multiple addition from occurring and increases the yield compared to the unsubstituted allyl counterpart.

The chemistry developed in this chapter is operationally simple; the protocols do not require any degassing, which is often a timely task requiring high quality pumps and liquid nitrogen for freeze-thaw drying. The tolerance of this reaction to ambient atmosphere means an expensive glove box is not required. The reaction is catalysed by a photoredox catalyst, requiring cheap blue LEDs for initiation. Overall, this reaction is cheap and simple to perform, requiring commercially available materials to install a range of native and non-native functionality into peptides and proteins in moderate to excellent isolated yields.

5. Application of bioconjugation chemistry to peptide stapling and macrocyclisation

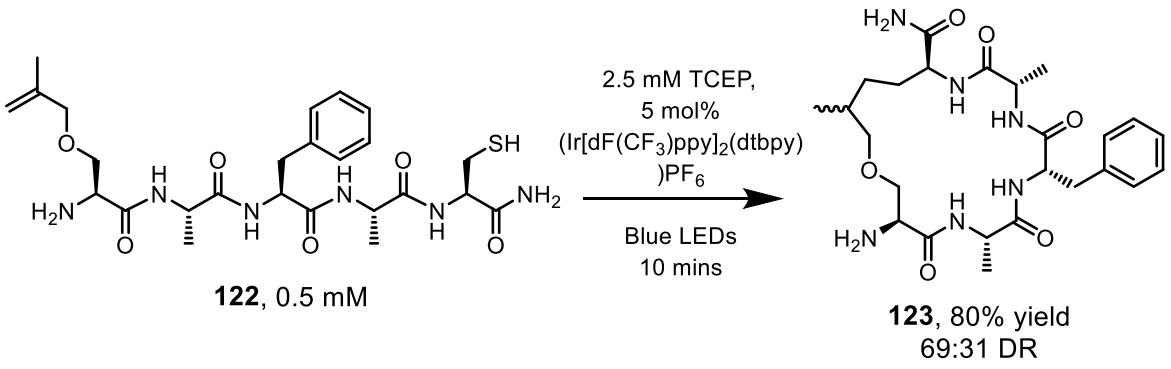
5.1 Chapter Abstract

The utilisation of macrocyclic peptides as therapeutic agents has gained increasing attention in recent years.[99] Aside from pharmaceuticals, antimicrobial cyclic peptides have found use in agriculture as fungicides. Due to their high stability to proteases and favourable degradability into non-toxic by-products they are attractive new targets[172].

As discussed, there are a wide range of techniques available for producing cyclic peptides, each with their own positives and drawbacks. This chapter describes the utilisation of the previously discussed chemistry to create a novel method of synthesising peptide macrocycles. The method is compatible with fully unprotected amino acids and is fully chemoselective. The technique involves the intramolecular trapping of a Cys-derived alanyl radical with alkene-containing amino acids, cyclising the peptide via a stable $C(sp^3)$-$C(sp^3)$ bond.

5.2 Initial macrocyclisation trials

Initial trials focussed on the use of 2-methylallyl modified serine (**121**), synthesised by reacting Boc-Ser-OH with NaH, followed by the addition of 3-bromo-2-methylpropene. This crude product was then incorporated into the model peptide H-S(2-methylallyl)AFAC-NH_2 (**122**) using standard coupling conditions. This was chosen as the 2-methylallyl substrate conjugation in chapter 4 proceeded with moderate to high isolated yield. The first trial of this reaction was performed using the conditions from general protocol G (details in section 7.2), without the addition of alkene. The results from this reaction are shown in Scheme 83 and NMR spectra comparison between starting material and product in Figure 64.



Scheme 83 - Reaction scheme for the macrocyclisation of (**122**) to form cyclic peptide (**123**).

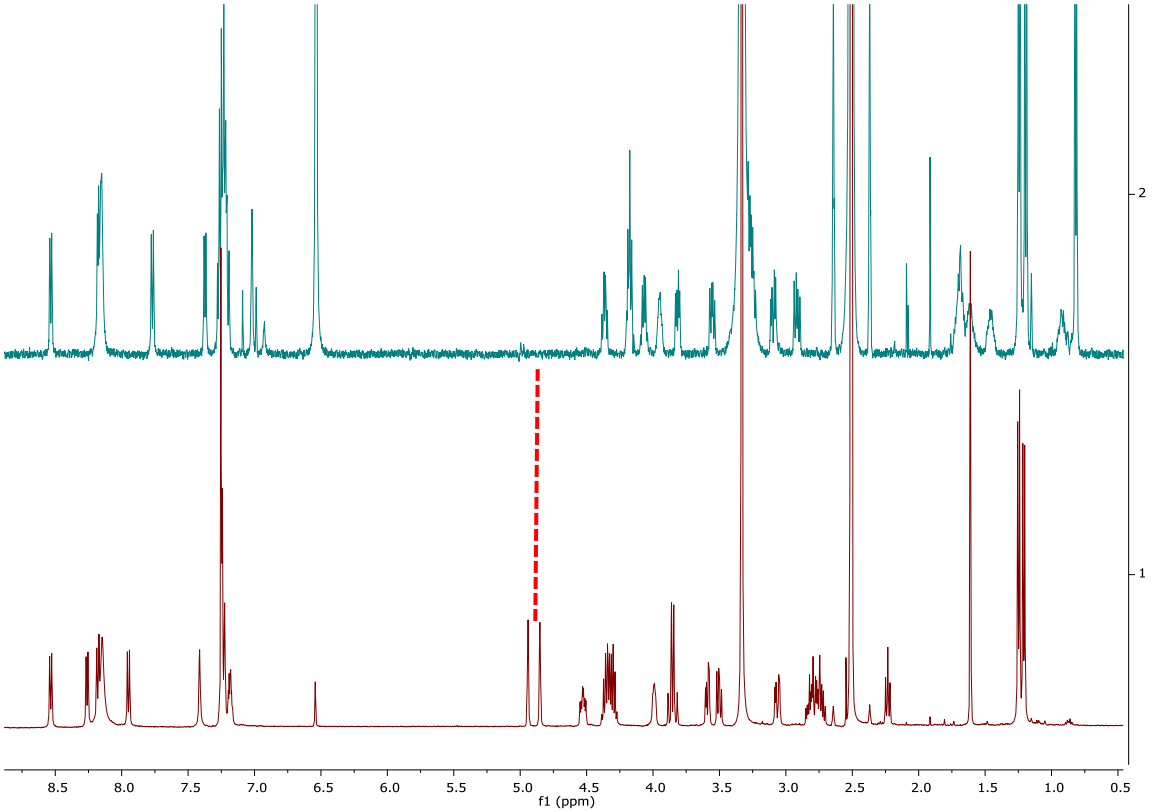


Figure 64 – NMR comparison of starting peptide (**122**, bottom) and cyclised peptide (**123**, top, majority diastereomer) showing the disappearance of the alkene protons at 4.9 ppm.

As indicated in Scheme 83, a mixture of diastereomers was formed from this reaction. Fortunately, the diastereomers could be separated by preparative HPLC and NMR analysis of each diastereomer showed that the desired cyclisation had occurred. The correct molecular weight present by MS and the disappearance of the alkene protons show that desulfurisation and cyclisation had occurred.

As the results from the use of the 2-methylallyl modified amino acid resulted in unwanted diastereomers, but higher yields than expected, the use of allyl modified serine was trialled (Figure 65). This was synthesised by reacting Boc-Ser-OH with NaH and allyl bromide, again using the crude product for installation into peptides. The allyl substrate conjugation in chapter 4 proceeds with moderate yields, so this model was also expected to result in moderate yields.

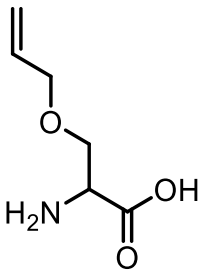
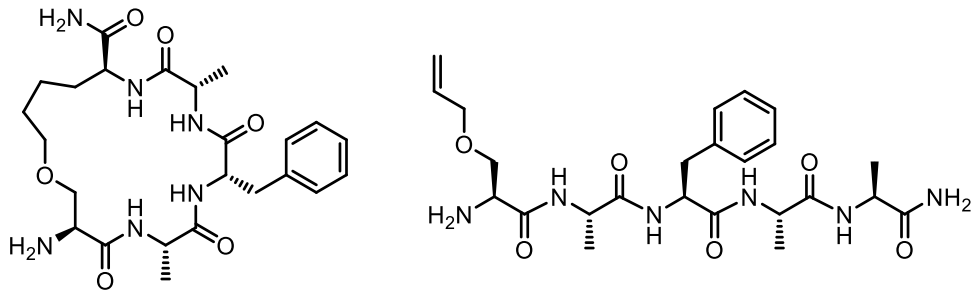


Figure 65 - Allyl modified Ser (**124**) used in the initial macrocyclisation trials.

A simple model peptide, H-S(Allyl)AFAC-NH_2 (**125**) was synthesised and subjected to the optimised conditions from Chapter 4.3. The resultant HPLC trace showed two major products being formed. The HPLC purification was able to separate these two compounds. Analysis by MS showed the same molecular weight for each product. This was indicative of a mixture of cyclised material (**126**) and linear desulfurised material (**127**) (Figure 66).



126 **127**

Figure 66 - The two potential products matching the observed molecular weight from the MS analysis. Both have a calculated molecular weight of 505.2775. The two structures show the cyclised (**126**, left) and linear desulfurised (**127**, right) products.

To confirm which peak was the linear desulfurised product, the peptide S(allyl)AFAA-NH_2 (**127**) was synthesised by SPPS, purified and analysed by analytical HPLC. This showed that the minor product peak was the linear desulfurised product. This was promising as it showed the reaction produced the cyclised product as the majority of the material. After this initial result, the reaction conditions were altered in an attempt to produce a single product. In these reactions the equivalents of TCEP and $(Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$ were altered and the analytical scale reactions monitored by HPLC. The data from these trials is shown in Table 24.

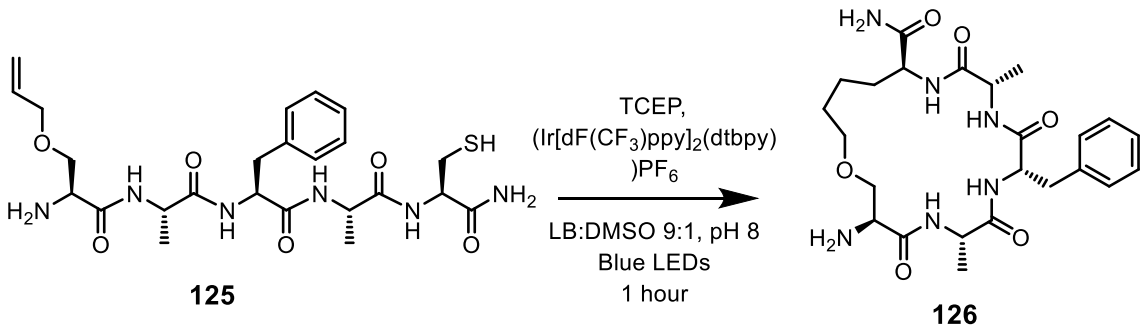


Table 24 - Optimisation of the macrocyclisation reaction using the H-S(allyl)AFAC-NH_2 (**125**) model peptide. The solvent system used was ligation buffer at pH 8.

Entry	Peptide/ mM (**125**)	TCEP eq.	Ir(III) eq.	1 hour %conv. (**126**)
1	0.5	10	0.05	72
2	0.5	5	0.05	38
3	0.5	10	0.01	36
4	0.1	10	0.05	0
5	0.5	50	0.05	100
6	1	10	0.05	90
7	1	50	0.05	100

The optimisation results show that by increasing the equivalents of TCEP from 10 to 50 the cyclised product was obtained as the majority product. There was no trace of starting material or linear desulfurised product observed. The HPLC trace from the successful Entry 5 (Table 24) is shown in Figure 67, showing that the crude reaction proceeds with very clean conversion. Entry 7 (Table 24)also shows an absence of starting material and uncyclised product, showing the reaction can successfully be performed at a higher concentration. The conditions from Entry 5 (Table 24) were repeated without the presence of Gdn•HCl to ensure this does not affect the reaction, resulting in the same conversion. The conditions were also repeated in the dark, resulting in only starting material being observed after 1 hour, showing the reaction is catalysed by light.

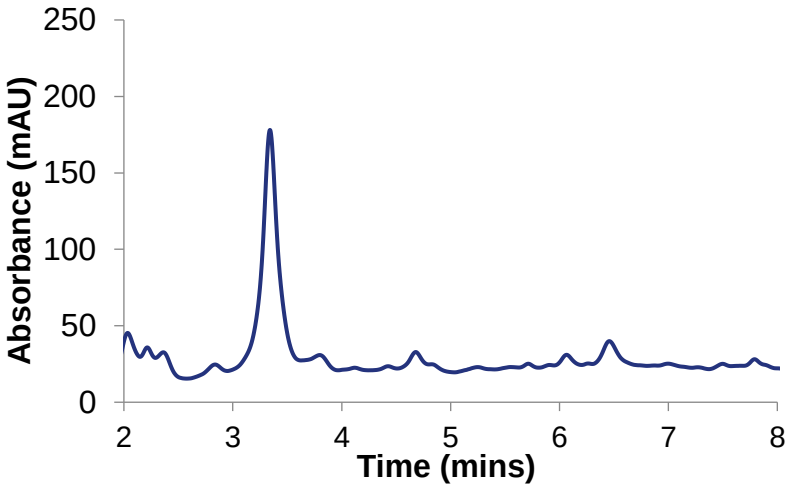


Figure 67 - Crude analytical HPLC trace from the cyclisation optimisation Entry 5 (Table 24).

Using the conditions from Entry 5 (Table 24), the reaction was repeated on an isolatable scale to allow product analysis by NMR. As this cyclisation reaction would always have the potential of producing the cyclised and uncyclised products with the same molecular weight, it was important to have a definitive method of analysis. For this, NMR analysis could be used to differentiate between the cyclised and uncyclised forms via the absence or presence of allylic protons. These signals appear in a region typically uninhabited by peptide functional groups (5-6 ppm), therefore providing a useful indicator of cyclisation.

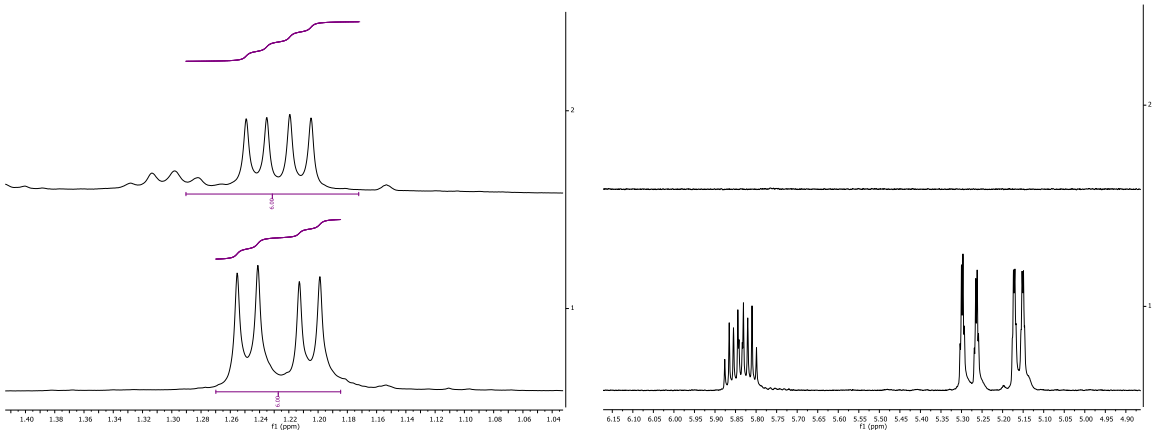
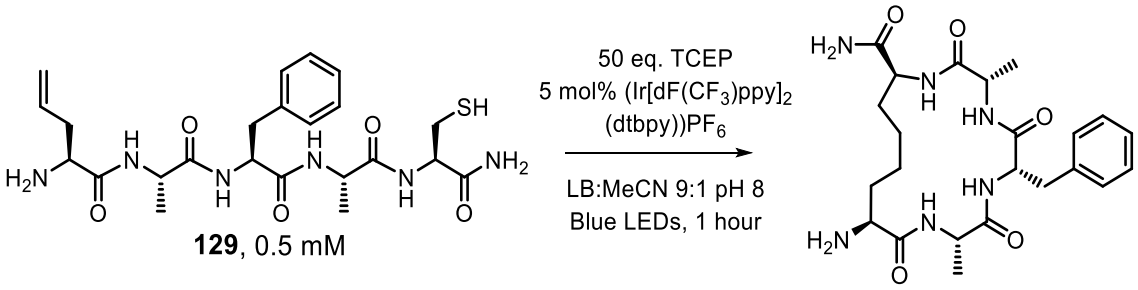


Figure 68 – NMR overlay showing starting peptide H-S(Allyl)AFAC-NH_2 (**125**, bottom) and cyclised product (**126**, top).

The overlay in Figure 68 shows the Ala CH_3 peaks (left) and allylic proton region (right) of each NMR. They show that both starting material and product have 6x CH_3 protons, indicating the presence of only two Ala residues. This shows that desulfurisation without cyclisation has not occurred as this would result in three Ala residues. The absence of allyl peaks, presence of only 2 Ala residues and correct molecular weight by MS indicates successful macrocyclisation, isolated in a 77% yield.

As the Boc-Ser(Allyl)-OH (**124**) is not commercially available, this procedure would not be simple to repeat for researchers without training in synthetic chemistry. We would ideally wish to develop a macrocyclisation technique which utilises only commercially available materials. With this in mind, the Boc and Fmoc protected versions of allyl glycine (**128**, alG) were purchased. This was then incorporated by SPPS into the model peptide H-alGAFAC-NH_2 (**129**) and cyclised (**130**), as shown in Scheme 84.



Scheme 84 - Reaction scheme for the macrocyclisation of an allyl glycine containing peptide (**129**) through desulfurative radical addition to form cyclic product **130**.

The previously optimised conditions from Entry 5 (Table 24) were repeated on this model peptide, resulting in successful cyclisation to give product (**130**) with a 79% isolated yield. The NMR comparison (Figure 69) of product and starting peptide again shows the disappearance of the allylic protons at around 5.7 and 5.1 ppm, indicating cyclisation has occurred successfully.

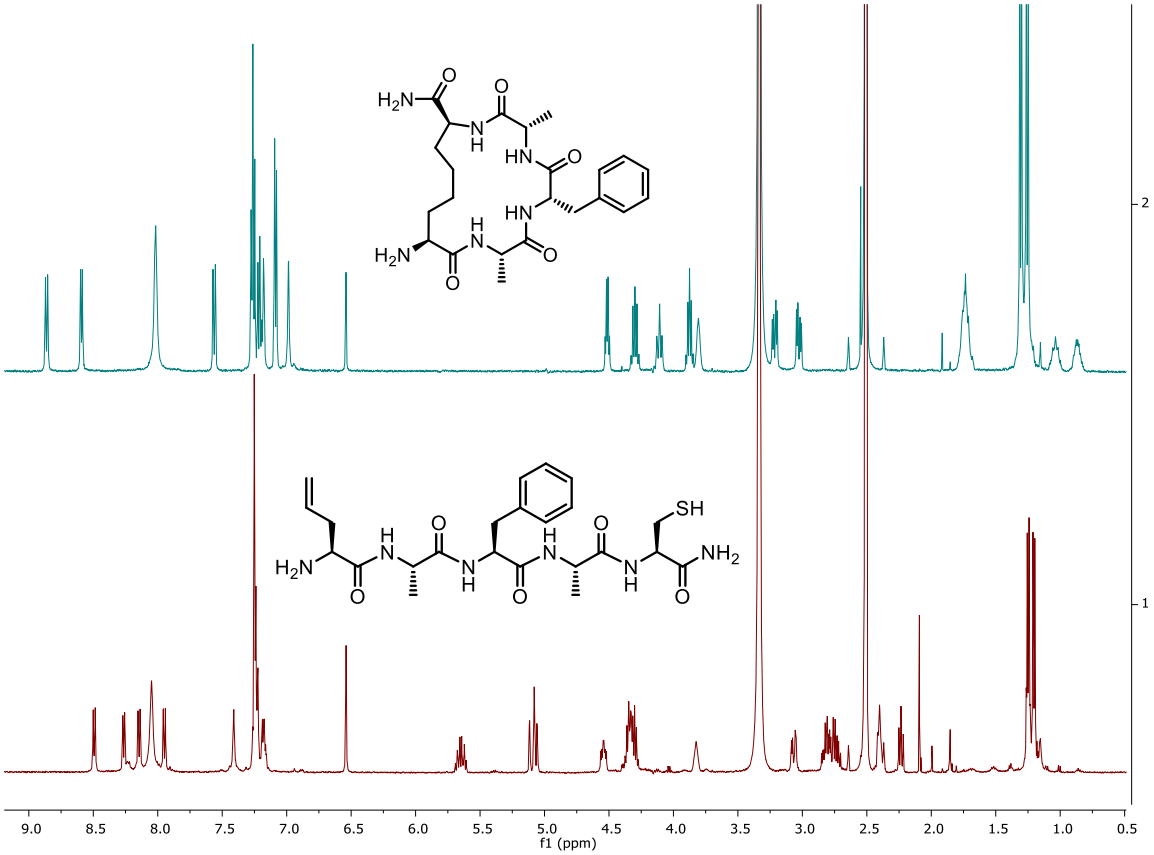
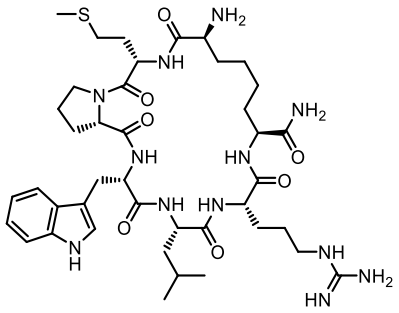


Figure 69 - NMR comparison of desulfurised and cyclised H-alGAFAC-NH_2 (**129**, top) and the starting peptide (**130**, bottom).

The data showed that allyl glycine could be successfully used as the alkene source for the intramolecular reaction between this and the alanyl radical produced from free-radical-induced desulfurisation. Due to the cyclisation success when using allyl glycine and its commercial availability for direct incorporation into peptides through SPPS, it was decided that this would be used for all future model peptides.

5.3 Further examples for macrocyclisation

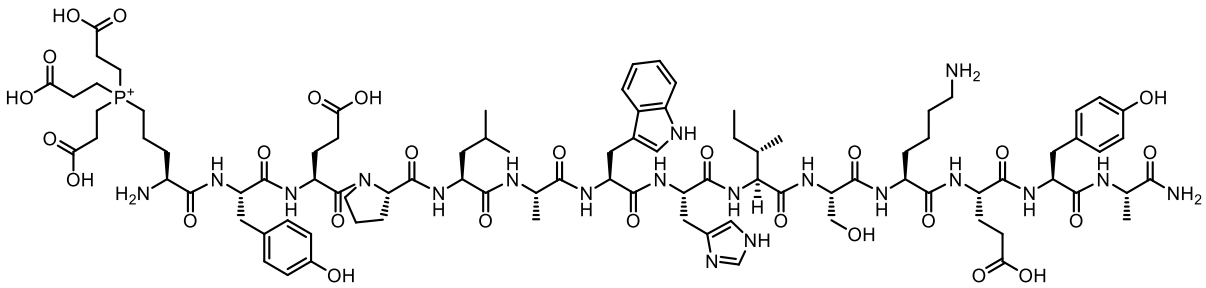
After the successful initial macrocyclisation reactions a range of model peptides were synthesised incorporating allyl glycine either at the N-terminus or internally. Firstly, two larger peptides were trialled. The sequences were chosen to test the method against bulky residues to ensure these do not affect the reactions. The two sequences were H-alGMPQLRC-NH_2 (**131**) and H-alGYEPLAWHISKEYC-NH_2 (**132**). The model peptide H-alGMPWLRC-NH_2 (**131**, Figure 70) was successfully cyclised in a 65% isolated yield (**133**). The NMR data showed no presence of allylic protons.



133, 65% Isolated Yield

Figure 70 – Structure of the cyclised 23-mer peptide H-alGMPWLRC-NH_2 (**133**).

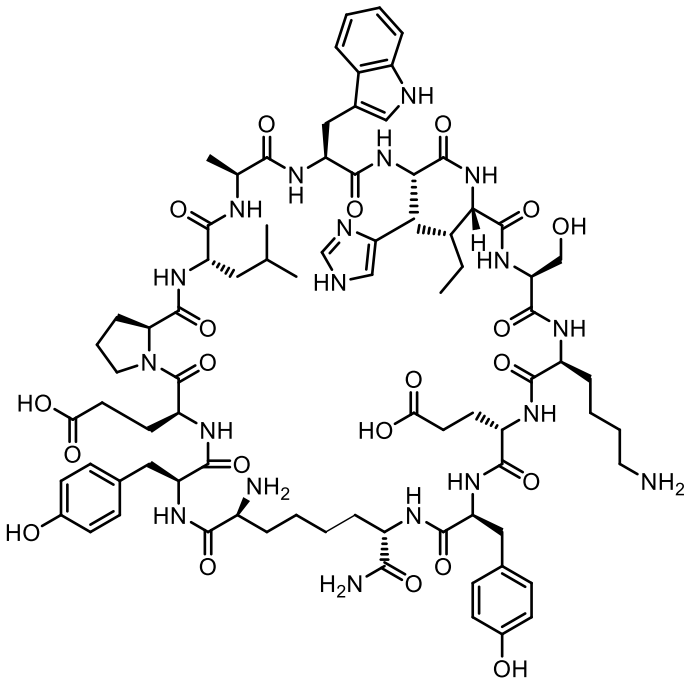
When the same conditions were applied to the larger model H-alGYEPLAWHISKEYC-NH_2 (**132**) an unknown majority product was formed. After MS analysis, this was shown to be the nucleophilic addition product of TCEP into the double bond of allyl glycine (**134**, Figure 71). This was further supported by the disappearance of the allylic protons in the ^{1}H-NMR and the appearance of a peak in the ^{31}P-NMR.



134

Figure 71 – The isolated product (**134**) from the trial macrocyclisation of H-alGYEPLAWHISKEYC-NH_2, showing nucleophilic addition of TCEP into allyl glycine.

To avoid this nucleophilic addition by-product, the reaction was attempted with a lower equivalent of the radical initiator Ir(III) (1 mol%). This reaction did not produce the nucleophilic addition product; however, it did produce an inseparable mixture of cyclised and linear desulfurised products (**135**, Figure 72).



135, 72%
Mix of Cyc/Uncyc

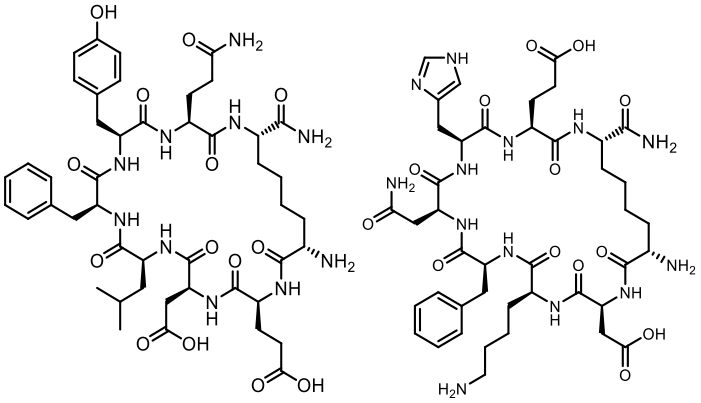
Figure 72 – Desired product (**135**) from the attempted desulfurative cyclisation of H-alGYEPLAWHISKEYC-NH_2 using 1 mol% catalyst producing a 44-mer cyclic peptide.

As the application of this method to a larger peptide (**132**) resulted in an inseparable mixture of cyclised and linear desulfurised material, it is clear that the developed method is more applicable to the cyclisation of shorter peptides. Based on this conclusion, a range of smaller models were synthesised and the cyclisation attempted. These model systems are based on the human ACE2 protein which interacts with the SARS-CoV-2 spike protein of COVID-19, mediating cell-entry during infection (amino acid residues 23-42; H-EQAKTFLDKFNHEAEDLFYQ-OH) and are thus pharmaceutically relevant. The following experiments were performed to assess whether this method could be used to form *i,i+4* and *i,i+7* stapled peptides (Figure 73). The reason stapled peptides have received interest is that the staple of choice locks the peptide into its alpha helical conformation.[173,174] The overall aim of this is to stabilise peptides and increase their longevity in vivo.



Figure 73 - Diagram showing the difference between an *i, i*+4 (left) and an *i, i*+7 (right) staple.

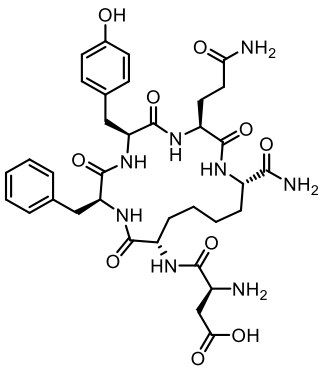
The first two sequences trialled to form *i,i+7* cyclic peptides were H-alGEDLFYQC-NH_2 (**136**) and H-alGDKFNHEC-NH_2 (**137**). Pleasingly, using only 1 mol% of the catalyst, the desired cyclised products were isolated in a 65% (**138**) and 70% (**139**) yield respectively.



138, 65% Isolated Yield **139,** 70% Isolated Yield

Figure 74 – Structures of the cyclic peptides from the desulfurative cyclisation of H-alGEDLFYQC-NH_2 (left, **138**) and H-alGDKFNHEC-NH_2 (right, **139**), producing 26-mer cyclic peptides.

Next, the method was applied to the synthesis of stapled peptides. The *i,i+4* model peptide, based on the ACE2 protein, was H-DalGFYQC-NH_2 (**140**). This cyclised peptide (**141**) was isolated in an 82% isolated yield, showing that internalising the allyl glycine residue did not affect the performance of the reaction (Figure 75).



141, 82% Isolated Yield

Figure 75 – Structure of the product from the desulfurative cyclisation of H-DalGFYQC-NH_2 (**141**).

The final model used to explore internal cyclisations was the hormone oxytocin (Figure 76). In nature, this is bridged by a disulfide bond. This method would form a permanent bridge if successful.

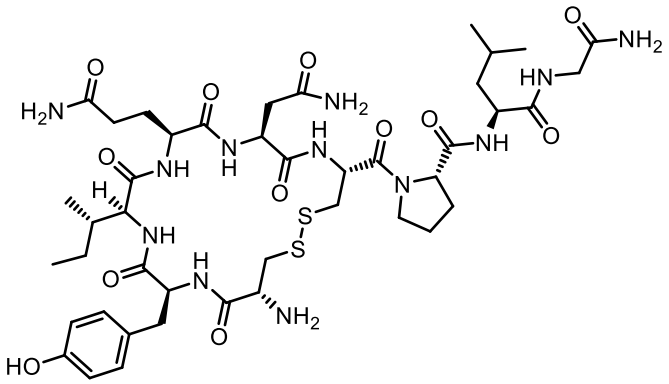
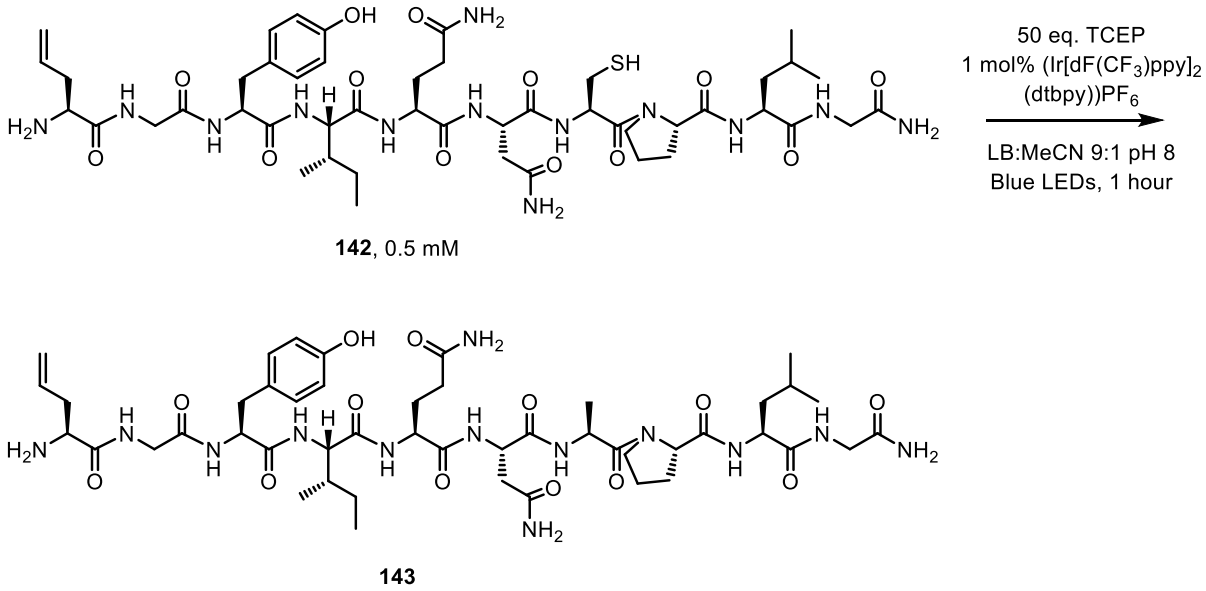


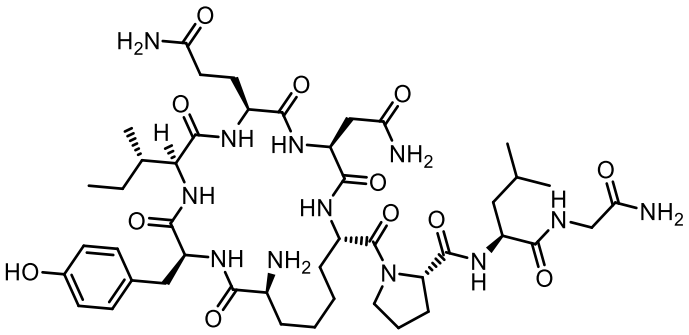
Figure 76 – Structure of the native oxytocin, bridged via a disulfide bond between the two Cys residues – H-CYIQNCPLG-NH_2.

The first trial used the model peptide H-alGYIQNCPLG-NH_2 (**142**). The crude cyclisation showed a single majority peak, however, upon NMR analysis this was shown to contain allylic protons, indicating the product was linear desulfurised (**143**, Scheme 85).



Scheme 85 – Reaction scheme showing the desulfurisation of starting peptide (**142**) to produce Ala containing peptide (**143**).

The unsuccessful cyclisation observed was theorised to be due to the unfavourable conformation that would be required for the alanyl radical to react with the alkene. To test this, the peptide was synthesised exchanging the Cys for the alG, H-CYIQNalGPLG-NH_2 (**144**). As the N-terminus is a more flexible position, this should have been able to react with the internal alkene. The result of this cyclisation was a relatively even mixture of two products. NMR analysis (Figure 78) showed both cyclised (**145,** Figure 77) and linear desulfurised (**146**) products, with an isolated yield of 52% and 35% respectively. This data shows that this method is not compatible with an internal Cys residue; however it is compatible with an internal alG.



145, 52% Isolated Yield

Figure 77 – Structure of the cyclic peptide (**145**) obtained from the desulfurative cyclisation of H-CYIQNalGPLG-NH_2, producing a 20-mer cyclic peptide.

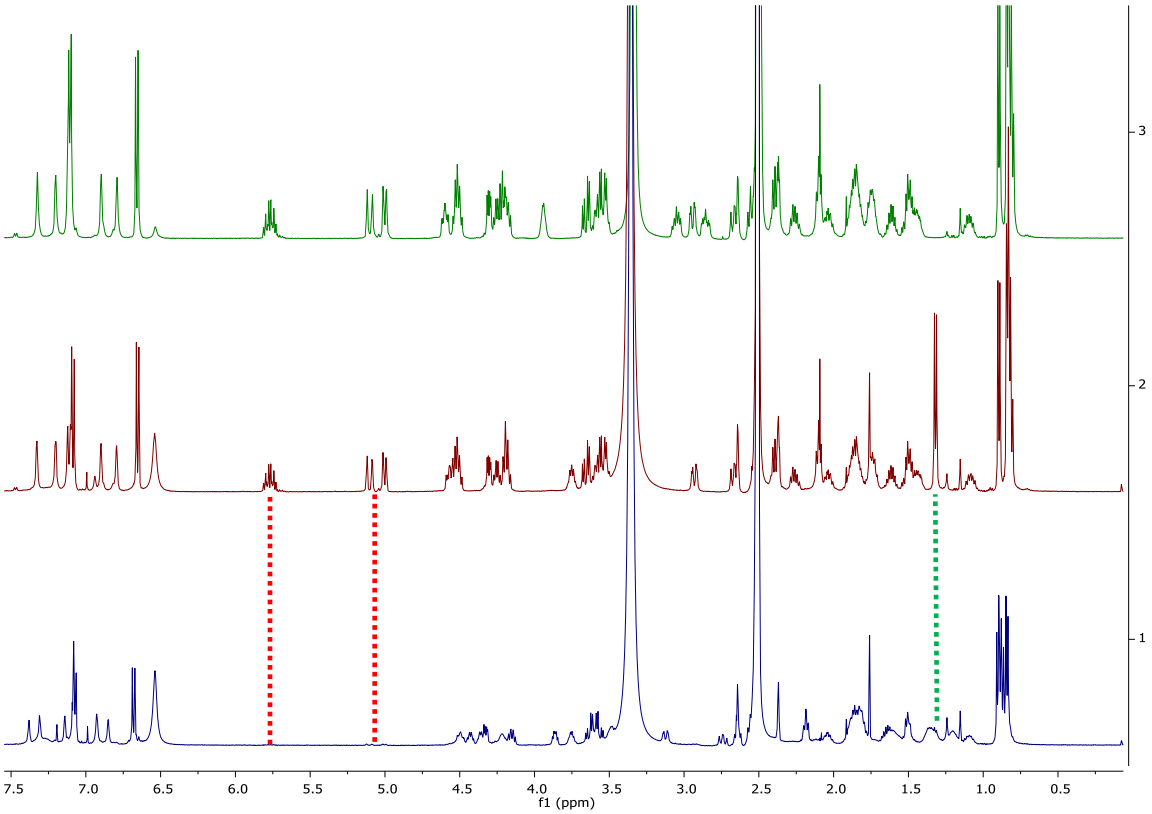


Figure 78 - NMR overlay comparing oxytocin model starting material (**144**, top), linear desulfurised (**146**, middle) and cyclised material (**145**, bottom). The dashed red lines highlight the allylic protons, not seen in the bottom trace. The dashed green line shows the presence of a doublet and 1.32 ppm, indicative of the Ala CH_3.

Overall, this chapter demonstrates the successful development of a novel macrocyclisation technique based on the chemistry developed in Chapter 4. The cyclised yields obtained are summarised in Table 25. The data shows a moderate to excellent isolated yield for a range of peptide sequences, showing various unprotected functional groups and steric bulk can be

tolerated. The method is fast, operationally simple and uses commercially available chemicals, 1 mol% of catalyst and a cheap blue LED setup. The technique is a novel example of a method to product macrocyclised peptides, linked with a carbon-carbon bond. The advantages of this technique include the use of purely commercially available materials, the operational simplicity of the reaction, the short reaction times and high isolated yields.

Table 25 – Summary of the isolated yields obtained from the desulfurative macrocyclisation of the peptides specified.

Compound Number	Peptide Sequence	Macrocycle Size	Isolated Yield (%)
123	H-S(2-methylallyl)AFAC-NH_2	19-mer	80
126	H-S(allyl)AFAC-NH_2	19-mer	77
130	H-alGAFAC-NH_2	17-mer	79
133	H-alGMPWLRC-NH_2	23-mer	65
134	H-alGYEPLAWHISKEYC-NH_2	44-mer	72 (Mixture)
138	H-alGEDLFYQC-NH_2	26-mer	65
139	H-alGDKFNHEC-NH_2	26-mer	70
141	H-DalGFYQC-NH_2	17-mer	82
145	H-CYIQNalGPLG-NH_2	20-mer	52

5.4 Conclusion

The data from this chapter shows that a novel method for peptide macrocyclisation has been developed. The method utilises the commercially available allyl glycine and Cys residues. The desulfurisation conditions previously developed in this work (albeit slightly altered) permit the intramolecular-addition of the alanyl radical into the allyl glycine, forming the cyclised product in moderate to excellent yields. The method does have limitations, the first being macrocycle size, with 26-mers being tolerated but 44-mers producing an inseparable mixture of cyclised and linear desulfurised products. The second limitation is the positioning of the Cys residue. The attempted cyclisation of the oxytocin mimetic with the internal Cys residue produced only linear desulfurised product, likely due to steric hindrance preventing the alanyl radical reacting with the alkene. This can be overcome by having the Cys residue on either termini, which is

an acceptable solution for the majority of desired cyclic products. The method tolerates internalisation of allyl glycine, therefore, this residue can be used to produce both cyclic and stapled peptides. This method represents a powerful addition into the field of cyclic peptide synthesis, providing a novel technique to produce highly stable cyclic peptides, forming a carbon-carbon bond.

6. Conclusion and Future Work

There is an ever-increasing prominence of peptides and proteins in general research and within the pharmaceutical industry. Methods to site-selectively install biologically relevant groups are, therefore, in high demand. There are a wide range of powerful methods currently available, each with their advantages and disadvantages. The aim of this thesis was to address the disadvantages and develop a novel method for site-selective bioconjugation.

Chapters 2 and 3 demonstrate a novel approach to site-selective bioconjugation. The method involves the installation of a broad variety of interesting groups via trapping of the alanyl radical produced from free-radical induced dechalcogenation. The chemistry developed targets Cys and Sec residues and can be applied to Sec selectively in the presence of Cys. The method is operationally simple, uses low equivalents of the conjugate, and is highly tolerant to a range of moieties on the probe. The conjugated products can be isolated in good to excellent yields, with reaction times as little as 30 minutes on Sec and 2 hours on Cys. A model peptide was "dual-functionalised", selectively targeting Sec and Cys with different functionalised TEMPO probes. The method was successfully applied to a Cys containing protein, isolating conjugated product in good yield. A technique has also been developed to allow the one-pot ligation and conjugation of peptides and small proteins.

Chapter 4 describes the development of a novel method for the site-selective installation of common Lys PTMs via visible-light-catalysed desulfurative $C(sp^3)$-$C(sp^3)$ bond formation. Current chemical methods can be used to either install mimetics, or install racemic versions of Lys PTMs, however, this method is, to our knowledge, the first to directly enable the installation of PTM'd Lys sidechains without eliminating the stereochemistry at the modified residue.

The second part of Chapter 4, utilising stabilised alkene substrates, demonstrates the application of this method as a high-yielding and versatile general bioconjugation technique, again via visible-light-catalysed desulfurative $C(sp^3)$-$C(sp^3)$ bond formation. A final method briefly explored in this chapter involves a novel $C(sp^3)$-$C(sp^2)$ cross-coupling reaction at Cys via the same approach using aryl boronic acid. With further optimisation, this Cys-arylation reaction could be a highly useful technique for the late-stage functionalisation of peptides.

Chapter 5 focusses on the application of the alkene conjugation method, developed in Chapter 4, to the intramolecular trapping of an alkene to form a cyclic peptide. After initial optimisation on a five amino acid model peptide, the cyclic product was obtained in an excellent isolated yield. Increasing the peptide length was problematic, with cyclised and uncyclised mixtures being produced. In addition, Michael addition of TCEP into the allyl glycine also occurred. The latter issue could be avoided by decreasing the amount of catalyst from 5 to 1 mol%, which is

advantageous as the catalyst is the expensive reagent in the reaction. The optimised conditions were applied to a range of peptides (5 to 9 residues), with cyclised products being isolated with moderate to excellent yield. The cyclisation of the biologically relevant oxytocin was also successful, showing that the allyl glycine could be internalised. This opens the method to the potential of producing stapled peptides as well as cyclic peptides. This method showed to be very successful, allowing cyclisation of a range of unprotected peptides with any of the canonical amino acids present.

Future work based on Chapters 2 and 3 may involve the further application of the method in catch-and-release proteomics to profile reactive Cys residues. In this case, cell lysate would be isolated and the conditions applied using varying concentrations of the biotin-TEMPO probe. Proteins carrying reactive Cys would be modified, isolated using magnetic streptavidin beads, and released using the Zn/AcOH conditions. Furthermore, this catch-and-release method could be used to purify peptides; using a solid support functionalised with TEMPO, Cys containing peptides could be 'fished' out of a crude solution and cleaved to afford a Ser at the site of capture.

Future work on Chapter 4 will involve applying the method to additional protein models that are known to have PTM Lys residues to develop a library of modified proteins to further probe fundamental biochemical pathways. Future work on this area will focus on developing this chemistry further, assessing the substrate scope and applying to larger peptides and proteins. The native Lys PTM reaction will be applied to alternative protein examples and tested to show activity between the native and chemically produced modified proteins. This can then be used to generate large libraries of proteins baring native Lys PTMs, giving the ability to screen the effects different PTMs have on the protein's activity.

The methyl-alkene trapping described in this chapter will also be taken further, using larger protein systems to demonstrate the applicability of this versatile, operationally simple, and high-yielding bioconjugation method. This method will also be tested to assess whether the Lys PTM mimetics have comparable activity to the native PTMs. This method could also be used as a simple technique to generate other modified peptides, including stable peptide-drug conjugates for potential pharmaceuticals. The potential of this method to be used for radiolabelling will also be further developed. An alkene can be simply and site-selectively installed into a peptide/protein substrate, a fluorinated by hydrofluorination. This method requires further development to show it can be applied to proteins, and the reaction conditions optimised to ensure protein compatibility. Additionally the Cys-arylation work will be continued by other members of the group, in an attempt to optimise this chemistry for complex peptides. This method has the potential of being a very powerful technique, forming highly stable conjguates mimicking the native amino acid phenylalanine.

Future work for Chapter 5 will include fully assessing the ability of the method to produce stapled peptides. It would be interesting to see whether a dual-stapled peptide could be produced. It would also be interesting to explore the position of the Cys residue further, as internalisation closer to the termini may well be tolerated. The ability of this method will also be tested further, attempting to produce dual-stapled peptides in a selective manner. Finally, the macrocyclic ACE2 peptides, generated via intramolecular interception of the alanyl radical, will be assayed against the SARS-CoV-2 spike protein to determine their binding kinetics, and evaluate their potential as competitive inhibitors against COVID-19.

7. Experimental

7.1 General Methods

NMR: Samples were analysed on a Bruker AVIII 400 NMR system (^{1}H-NMR frequency 400 MHz; ^{13}C-NMR frequency 101 MHz). Chemical shifts are reported in parts per million (ppm) and are referenced to solvent residual signals: $CDCl_3$ (δ 7.26 [^{1}H]) (δ 77.16 [^{13}C]), DMSO-d_6 (δ 2.50 [^{1}H]) (δ 39.52 [^{13}C]), CD_3OD (δ 3.31 [^{1}H]) (δ 49.00 [^{13}C]), D_2O (δ 4.79 [^{1}H]). ^{1}H NMR data is reported as chemical shift (δ), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, dd = doublet of doublets, ddd = doublet of doublet of doublets or combinations of these multiples; m = unassigned multiplet), relative integral, coupling constant (J Hz) and assignment where possible. TEMPO probes containing persistent radicals were analysed using either one of two methods. NMR method A involved direct addition of phenyl hydrazine to the NMR tube prior to analysis. Method B involved injection into an RP-HPLC column using 0.1% TFA in water/acetonitrile, product collection and lyophilisation. This method has proven useful in quenching the TEMPO radical without masking the aromatic region of the NMR with phenyl hydrazine peaks.

High-resolution mass spectrometry (HRMS): Recorded on a Bruker MicrOTOF Focus II MS (ESI) operating in positive or negative ionisation mode.

Analytical HPLC: Performed on a Thermo Ultimate 3000 uHPLC system equipped with PDA eλ detector (λ = 210 – 400 nm). Peptides were analysed using a Waters SunFire 5 µm, 2.1 x 150 mm column (C-18) at a flow rate of 0.6 mL min^{-1}. The mobile phase composed of 0.1% trifluoroacetic acid in H_2O (Solvent A) and 0.1% trifluoroacetic acid in acetonitrile (Solvent B). The analysis of the chromatograms was conducted using Chromeleon 7 software.

Preparative reverse-phase HPLC: Performed using a Waters 1525 binary pump HPLC equipped with a dual wavelength UV detector set to 210 nm and 280 nm. Peptides were purified on a Waters SunFire 5 µm (C-18) preparative column with 5-µm particle size, 19 x 150 mm, operating at a flow rate of 6 mL min^{-1} using a mobile phase of 0.1% trifluoroacetic acid in water (Solvent A) and 0.1% trifluoroacetic acid in acetonitrile (Solvent B) using the gradient specified for each specific compound. Semi-preparative reverse-phase HPLC was performed using the same HPLC and solvent system. The column used was a Waters SunFire 5 µm (C-18) preparative column, 10 x 250 mm, operating at a flow rate of 5 mL min^{-1} using the gradient specified for each specific compound.

Materials

Commercial materials were used as received unless otherwise noted. Amino acids, coupling reagents and resins were obtained from Novabiochem, Fluorochem or GL Biochem. Reagents that were not commercially available were synthesised as outlined for each specific compound. Solvents were obtained as reagent grade from Merck or Fisher. $(Boc\text{-}Sec\text{-}OH)_2$ was synthesised from commercially available $(H_2N\text{-}Sec\text{-}OH)_2$ (Acros organic).

A range of compounds used in this thesis were synthesised by Dr Frances Smith. These include Boc-D-Sec-OH, H-D-UAF-OMe, H-D-A(TEMPO)AF-OMe, biotin modified with 3-methyl-3-buten-1-amine and 2-methylallylamine, and H-A(TEMPO)WIMKY-NH_2.

Solid Phase Peptide Synthesis (SPPS)

Preloading Rink Amide resin*:* Rink amide resin was initially washed with DCM (5 × 3 mL) followed by removal of the Fmoc group by treatment with 20% piperidine/DMF (2 × 5 min). The resin was washed with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL). Oxyma (4 eq.) and DIC (4 eq.) were added to a solution of Fmoc-AA-OH (4 eq.) in DMF. After 5 min of pre-activation, the mixture was added to the resin. After 2 h the resin was washed with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL), capped with acetic anhydride/pyridine (1:9 v/v) (2 × 3 min) and washed with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL).

Preloading 2-chlorotrityl chloride resin*:* 2-Chlorotrityl chloride resin was swollen in DCM for 30 min then washed with DCM (2 × 3 mL). A solution of Fmoc-AA-OH (0.5 eq. relative to resin functionalization) and *i*Pr_2NEt (2 eq. relative to resin functionalization) in DCM (final concentration 0.1 M of amino acid) was added and the resin shaken at rt for 16 h. The resin was washed with DMF (5 × 3 mL) and DCM (5 × 3 mL). The resin was treated with a solution of DCM/CH_3OH/*i*Pr_2NEt (17:2:1 v/v/v, 3 mL) for 1 h and washed with DMF (5 × 3 mL), DCM (5 × 3 mL), and DMF (5 × 3 mL). The resin was subsequently submitted to iterative peptide assembly (Fmoc-SPPS).

Estimation of amino acid loading*:* The resin was treated with 20% piperidine/DMF (2 x 3 mL, 3 min) and 20 µL of the combined deprotection solution was diluted to 10 mL using 20% piperidine/DMF in a volumetric flask. The UV absorbance of the resulting piperidine-fulvene adduct was measured (λ = 301 nm, ε = 7800 M^{-1} cm^{-1}) to determine the amount of amino acid loaded onto the resin.

Manual Iterative peptide assembly (Fmoc-SPPS)

General amino acid coupling: A solution of protected amino acid (4 eq.), DIC (4 eq.) and oxyma pure (4 eq.) in DMF (final concentration 0.1 M) was added to the resin. After 1 h, the

resin was washed with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL).

Deprotection*:* The resin was treated with 20% piperidine/DMF (2 x 3 mL, 3 min) and washed with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL).

Capping: Acetic anhydride/pyridine (1:9 v/v) was added to the resin (3 mL). After 3 min the resin was washed with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL).

Coupling conditions for (Boc-Sec-OH)$_2$: A solution of compound (Boc-Sec-OH)$_2$ (1 eq.), DIC (2 eq.), and Oxyma pure (2 eq.) in DMF was added to the resin (1 eq.) and shaken at rt for 16 h. The resin was then washed with DMF (5 × 3 mL), and DCM (10 × 3 mL).

Coupling conditions for Boc or Fmoc protected allyl glycine*:* A solution of compound Boc/Fmoc-allylglycine (1.2 eq.), DIC (1.5 eq.), and Oxyma pure (1.5 eq.) in DMF was added to the resin (1 eq.) and shaken at rt for 16 h. The resin was then washed with DMF (5 × 3 mL), and DCM (10 × 3 mL).

Cleavage: For peptides containing Cys/Met: A mixture of TFA, thioanisole, tri*iso*propylsilane (TIS) and water (90:4:4:2 v/v/v/v) was added to the resin. After 3 h, the resin was washed with TFA (3 × 2 mL). For all other peptides: A mixture of TFA, tri*iso*propylsilane (TIS) and water (95:2.5:2.5 v/v/v) was added to the resin. After 3 h, the resin was washed with TFA (3 × 2 mL).

Work-up: The combined solutions were concentrated under a stream of nitrogen to < 5 mL. 40 mL of diethyl ether was added to precipitate the peptide and the suspension centrifuged. The pellet was then dissolved in water containing 0.1% TFA, filtered and purified by preparative HPLC and analysed by LC-MS and ESI mass spectrometry.

Automated solid-phase peptide synthesis

Automated Fmoc-SPPS was carried out on a Biotage Initiator$^+$ Alstra microwave peptide synthesizer. General synthetic procedures for Fmoc-deprotection and capping were carried out in accordance with the manufacturer's specifications. Standardized amino acid couplings were performed for 15 min at 50 ^{o}C under microwave irradiation in the presence of amino acid (0.5 M in DMF, 4 eq.), Oxyma (0.5 M in DMF, 4 eq.) and di*iso*propylcarbodiimide (0.5 M in DMF, 4 eq.). Peptide cleavage and work-up were carried out as described above for manual SPPS.

Selenoester synthesis general procedure

Model peptide selenoesters were prepared on 2-chlorotrityl chloride resin using Fmoc-SPPS as described in the general methods. Cleavage of the peptides from the resin was achieved

by treating with 30 vol.% HFIP in DCM (5 ml) for 2 hours before concentrating *in vacuo*. The resulting residue was dissolved in DCM and cooled to 0 °C. Diphenyl diselenide (DPDS) (30 eq. in DCM) was added to the solution followed by Bu_3P (30 eq.). The reaction was allowed to proceed at 0 °C for 3 h, after which time the solvent was removed *in vacuo*. The crude material was put on ice and the protecting groups removed via treatment with TFA:TIS:thioanisole:H_2O (90:4:4:2 v/v/v/v). After 3 h at room temperature, the cleavage cocktail was removed under a stream of N_2 and the crude residue suspended in diethyl ether and cooled to -50 °C. The precipitate was pelleted by centrifugation at 4000 rpm for 5 min, the supernatant decanted and the pellet dissolved in 0.1% TFA/H_2O with addition of 0.1% TFA/MeCN and purified via preparative HPLC.

General additive-free selenocystine-selenoester ligation procedure

Peptide selenoesters (1.05-1.20 eq.) and peptide dimers bearing N-terminal selenocystine (0.5 eq) were dissolved separately in ligation buffer (6 M guanidine hydrochloride, 100 mM Na_2HPO_4, pH 7.2) to a concentration of 10.5-12 mM with respect to the peptide selenoester and 5 mM with respect to the peptide diselenide dimer. The solutions were combined to give an overall concentration of 2.5 mM with respect to the selenopeptide dimer, 5.25-6 mM with respect to the peptide selenoester and a final pH of 6.2-6.5. The reaction was monitored by RP-HPLC and once complete the DPDS was extracted with hexane. Immediately before purification by preparative RP-HPLC, a small amount of TCEP was added to the ligation mixture. Fractions containing the desired ligated product were lyophilized.

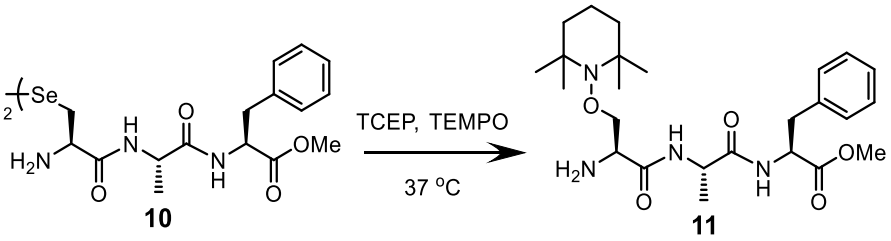
General additive-free cysteine-selenoester ligation procedure

The peptide selenoester (1.05 eq.) and peptide bearing N-terminal cysteine (1.0 eq) were dissolved separately in ligation buffer (6 M guanidine hydrochloride, 100 mM Na_2HPO_4, pH 7.2) to a concentration of 10.5 mM and 10 mM respectively. The solutions were combined to give an overall concentration of 5 mM with respect to the peptide bearing an N-terminal cysteine and 5.25 mM with respect to the peptide selenoester, and a final pH of the reaction of 6.5-7.0. If the pH was below 6.5 the pH was carefully adjusted with 2 M NaOH, ensuring to not exceed pH 7. The reaction was monitored by RP-HPLC and once complete the DPDS was extracted with hexane. Hydroxylamine (10 eq) was added to the ligation solution and the pendant thioester hydrolysis monitored by RP-HPLC. Once hydrolysed, TCEP was added and purification by preparative HPLC was performed. Fractions containing the desired ligated product were lyophilized.

7.2 General Protocols

General Protocols for Chapters 2 and 3

Initial Optimization of TEMPO Conjugation Method



Unless otherwise specified the following method was followed:

To H-UAF-OMe (**10**, 2.5 mM final concentration) in ligation buffer (LB)/MeOH (20% MeOH) at pH 7-7.5 was added TCEP (0.625 M in LB pH adjusted to 7) followed by TEMPO (0.5 M in MeOH). The reaction was shaken at 37 °C and analysed by RP-HPLC (10-70% B over 5 minutes) at the specified time points. When doping, the same quantities of TCEP and TEMPO were added at 15, 45 and 60 minutes if doping three times. If doping twice, additional reagents added at 15 and 45 minutes. If doping once, additional reagents were added at 15 minutes. All reactions were performed in a Thermo heater/shaker set to the specified temperature and shaken at 800 rpm. Conversion to product reported was measured based on integration of the product peak from the HPLC trace compared to integration of starting material and by-product peaks.

General one-pot ligation-conjugation protocol

Peptide selenoesters (1.05-1.20 eq.) and peptide dimers bearing N-terminal selenocystine (0.5 eq) were dissolved separately in LB (pH 7.2) to a concentration of 10.5-12 mM with respect to the peptide selenoester and 10 mM with respect to the selenopeptide monomer. The solutions were combined to give an overall concentration of 5.0 mM with respect to the selenol monomer, 5.25-6 mM with respect to the peptide selenoester and a final pH of the reaction of 6.0-7.0. After 10 minutes, hexane (equal volume to current solution) was added to the ligation mixture to extract away the DPDS. The ligation solution was then subjected to general conjugation protocol A: to the ligation mixture was added a solution of TCEP (0.625 M stock solution adjusted to pH 7, 50 eq. (eq. based on moles), 125 mM), TEMPO probe (0.1 M solution in DMSO, 2 eq., 5 mM) and $Mn(OAc)_3$ (0.1 M stock in DMSO, 4 eq., 10 mM). The solution was diluted to a final selenol monomer concentration of 2.5 mM. The final solution composition was 20% DMSO in LB. The pH was checked to be 7-7.5 and the reaction shaken at 50 °C for 1 hour. Reaction completion was checked by RP-HPLC and purified by semi-

preparative HPLC.

For selenopeptide fragments containing Cys the same ligation protocol was followed, using general protocol C for the conjugation;

General TEMPO conjugation protocol A – Accelerated modification of Sec

To peptide dissolved in 20% DMSO in LB to a concentration of 5 mM was added a solution of TCEP (0.625 M stock solution in LB, pH adjusted to 7, 50 eq.), TEMPO probe (0.1 M stock solution in DMSO, 2 eq.) and $Mn(OAc)_3$ (0.1 M stock solution in DMSO, 4 eq.). The pH of the reaction mixture was checked to be 7-7.5 then the reaction mixture diluted so the final peptide concentration was 2.5 mM. The reaction was then heated to 50 ^oC and shaken for 1 hour. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

General TEMPO conjugation protocol B – Accelerated modification of Cys

To peptide dissolved in 20% DMSO in LB to a concentration of 2 mM was added a solution of TCEP (0.625 M stock solution in LB, pH adjusted to 7, 50 eq.), TEMPO (0.1 M stock solution in DMSO, 2 eq.) and $Mn(OAc)_3$ (0.1 M stock solution in DMSO, 5 eq.). The pH of the reaction mixture was checked to be 7-7.5 then the reaction mixture diluted so the final peptide concentration was 1.0 mM. The reaction was then heated to 50 ^oC and shaken for 2 hours. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

General TEMPO conjugation protocol C – Sec-selective TEMPO conjugation in presence of Cys

To peptide dissolved in 20% DMSO in LB to a concentration of 5 mM was added a solution of TCEP (0.625 M stock solution in LB, pH adjusted to 7, 50 eq.) and TEMPO (0.1 M stock solution in DMSO, 5 eq.). The pH of the reaction mixture was checked to be 7-7.5 then the reaction mixture diluted so the final peptide concentration was 2.5 mM. The reaction was then heated to 37 ^oC and shaken for 4 hours. Analytical RP-HPLC was used to show full consumption of starting peptide and the reaction mixture was purified by semi-preparative HPLC.

General TEMPO conjugation protocol D – Internal Cys residue conjugation

To peptide dissolved in 20% DMSO in LB to a concentration of 5 mM was added a solution of TCEP (0.625 M stock solution in LB, pH adjusted to 7, 50 eq.), TEMPO (0.1 M stock solution in DMSO, 2 eq.) and $Mn(OAc)_3$ (0.1 M stock solution, 4 eq.). The pH of the reaction mixture was checked to be 7-7.5 then the reaction mixture diluted so the final peptide concentration was 2.5 mM. The reaction was then heated to 50 ^oC and shaken for 1 hour. Once the starting

material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

General TEMPO conjugation protocol E – Internal Cys residue protein conjugation

To peptide dissolved in 20% DMSO in ligation buffer to a concentration of 2 mM was added a solution of TCEP (0.625 M stock solution in LB, pH adjusted to 7, 100 eq.), TEMPO (0.1 M stock solution in DMSO, 5 eq.) and $Mn(OAc)_3$ (0.25 M stock solution in DMSO, 20 eq.). The pH of the reaction mixture was checked to be 7-7.5 then the reaction mixture diluted so the final peptide concentration was 1.0 mM. The reaction was then heated to 50 ^oC and shaken for 2 hours. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

TEMPO Conjugation General Protocol Summary

Table 26 - Summary of the conditions used in the general protocols above for chapters 2 and 3.

General Protocol	Peptide/ mM	TCEP eq.	TEMPO eq.	$Mn(OAc)_3$ eq.	Temp/ ^oC	Time/ hours
A	2.5	50	2	4	50	1
B	1.0	50	2	5	50	2
C	2.5	50	5	0	37	4
D	2.5	50	2	10	50	4
E	1.0	100	5	20	50	2

General Protocols for Chapter 4

Initial protocol for allyl trap on Sec containing peptides

The following stock solutions were first prepared in ligation buffer (6 M Gdn.HCl, 0.1 M PBS, pH 7.0); 0.0125 M Ac-UAY-NH_2, 5 M allyl compound, 0.5 M VA-044 and 1 M TCEP. The pH of the TCEP solution was adjusted to 7.0 using 10 M NaOH. The desired final concentrations were 5 mM peptide, 1 M allyl compound, 0.05 M VA-044 and 0.25 M TCEP. To a 2 ml Eppendorf was added 10 µL of peptide and 2.5 µL of ligation buffer. This was followed by 10 µL allyl compound, 5 µL VA-044 and 12.5 µL TCEP. This was left shaking for 24 hours.

Stock solutions of peptide and TCEP were prepared as stated for radical traps. Again, all other reagents were prepared in either LB or DMSO. When doping reactions, a new stock solution of VA-044 was prepared for each dope. The RM was prepared and left to react as previously stated for radical traps.

Photoredox reaction setup



The photoredox setup used standard iNextStation blue LED strip lights wrapped inside a Pyrex dish. Reaction vessels were chosen based on the reaction volume, using either 2 ml, 8 ml or 25 ml screw top vials. The temperature inside the dish was within the range of 30-37 ^{o}C. The reactions were set up inside a standard fume hood, then transferred into the LED dish once pH adjusted, no degassing was necessary. The reactions stopped when not exposed to light, so time points were monitored by removing 0.1 ml of reaction mixture and analysing by analytical HPLC.

General optimisation protocol using Ac-CWHISKEY-NH_2

Peptide stock solutions were prepared in either LB (peptide, TCEP and eosin Y) or DMSO (N-allyl acetamide, $(Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$ (Ir), $Mn(OAc)_3$).

Peptide stock solution was diluted with LB/DMSO, followed by addition of TCEP, N-allyl acetamide and initiator. The pH of the solution was adjusted to the specified pH and the reaction mixture placed in a Pyrex dish wrapped with blue LEDs (Ir(III)/eosin Y) or in a Thermo shaker set to the specified temperature ($Mn(OAc)_3$). Samples were removed and analysed at the specified time point using analytical HPLC (15 to 25% B over 10 minutes, 0.6 ml/min).

Final Ac-CWHISKEY-NH_2 optimised protocol F

To peptide dissolved in 20% acetonitrile (MeCN) in LB to a concentration of 1 mM was added a solution of TCEP (0.5 M stock solution in LB, pH adjusted to 7, 5 eq.), allyl compound (pure, 200 eq.) and $(Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$ (1 mM stock solution in MeCN, 0.05 eq.). The reaction mixture was diluted to a final peptide concentration of 0.5 mM and the pH of the reaction mixture was checked to be around 8 (7.8-8.2), this was adjusted as required by addition of either 10 M NaOH or 6 M HCl (care must be taken to not allow the pH to rise above 9). The reaction vessel was then placed into blue LEDs and left for 1 hour. Once the starting

material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

Optimised General Protocol G for Installation of Native Lys PTMs onto model H-HRLLRCGNYA-NH_2

To peptide dissolved in 20% acetonitrile (MeCN) in LB to a concentration of 1 mM was added a solution of TCEP (0.5 M stock solution in LB, pH adjusted to 7, 10 eq.), allyl compound (pure, 200 eq.) and (Ir[dF(CF_3)ppy]$_2$(dtbpy))PF_6 (1 mM stock solution in MeCN, 0.1 eq.). The reaction mixture was diluted to a final peptide concentration of 0.5 mM and the pH of the reaction mixture was checked to be around 8 (7.8-8.2), this was adjused as required by addition of either 10 M NaOH or 6 M HCl (Care must be taken to not allow the pH to rise above 9). The reaction vessel was then placed into blue LEDs and left for 1 to 2 hours. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

General Protocol H for Installation of Native Lys PTMs onto proteins

To protein dissolved in 20% acetonitrile (MeCN) in LB to a concentration of 1 mM was added a solution of TCEP (0.5 M stock solution in LB, pH adjusted to 7, 100 eq.), allyl compound (pure, 200 eq.) and (Ir[dF(CF_3)ppy]$_2$(dtbpy))PF_6 (1 mM stock solution in MeCN, 0.1 eq.). The reaction mixture was diluted to a final peptide concentration of 0.5 mM and the pH of the reaction mixture was checked to be around 8 (7.8-8.2), this was adjusted as required by addition of either 10 M NaOH or 6 M HCl (care must be taken to not allow the pH to rise above 9). The reaction vessel was then placed into blue LEDs and left for 1 hour. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

General Protocol I for peptide bioconjugation via C-C bond formation

To peptide dissolved in 10% acetonitrile (MeCN) in LB to a concentration of 1 mM was added a solution of TCEP (0.5 M stock solution in LB, pH adjusted to 7, 10 eq.), stabilised alkene compound (pure, 200 eq.) and (Ir[dF(CF_3)ppy]$_2$(dtbpy))PF_6 (1 mM stock solution in MeCN, 0.05 eq.). The reaction mixture was diluted to a final peptide concentration of 0.5 mM and the pH of the reaction mixture was checked to be around 8 (7.8-8.2), this was adjusted as required by addition of either 10 M NaOH or 6 M HCl (care must be taken to not allow the pH to rise above 9). The reaction vessel was then placed into blue LEDs and left for 10 minutes (butene probes) or 30 minutes (propene probes). Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

General Protocol J for protein bioconjugation via C-C bond formation

To protein dissolved in 10% acetonitrile (MeCN) in LB to a concentration of 1 mM was added a solution of TCEP (0.5 M stock solution in LB, pH adjusted to 7, 100 eq.), stabilised alkene

compound (pure, 200 eq.) and (Ir[dF(CF_3)ppy]$_2$(dtbpy))PF_6 (1 mM stock solution in MeCN, 0.05 eq.). The reaction mixture was diluted to a final peptide concentration of 0.5 mM and the pH of the reaction mixture was checked to be around 8 (7.8-8.2), this was adjusted as required by addition of either 10 M NaOH or 6 M HCl (care must be taken to not allow the pH to rise above 9). The reaction vessel was then placed into blue LEDs and left for 10 minutes. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC.

Chapter 4 General Protocol Summary

Table 27 - Summary of the conditions used in the general protocols for chapter 4. Ir(III) = (Ir[dF(CF_3)ppy]$_2$(dtbpy))PF_6. All the below reactions use blue LEDs as the initiator, with LB/MeCN at pH 7.8-8.2.

General Protocol	Peptide/ mM	TCEP eq.	Alkene eq.	Ir(III) eq.	Time/ mins
F	0.5	5	200	0.05	60
G	0.5	10	200	0.1	60-120
H	0.5	100	200	0.1	60
I	0.5	5	200	0.05	10-30
J	0.5	100	200	0.05	10

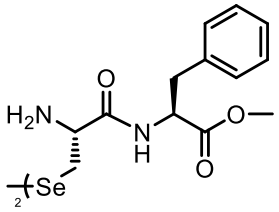
General Protocol K for Chapter 5

For peptide sequences containing allylglycine, the shortened code used was alG.

To linear peptide dissolved in LB/MeCN 9:1 to a concentration of 1 mM was added TCEP (50 eq., 0.5 M stock solution in water, pH adjusted to 7) and (Ir[dF(CF_3)ppy]$_2$(dtbpy))PF_6 (0.01 eq., 1 mM stock solution in MeCN). The reaction mixture was diluted to a final peptide concentration of 0.5 mM and the pH of the reaction mixture was checked to be around 8 (7.5-8.2), this was adjusted as required by addition of either 10 M NaOH or 6 M HCl (care must be taken to not allow the pH to rise above 9). The reaction mixture was placed into the general LED setup and left for 1 hour. At this time, the reaction mixture was purified by semi-preparative HPLC and fractions containing the desired product lyophilised to give the desired cyclic peptide as a fluffy white solid. The pure peptide was analysed by NMR (500 MHz) to ensure disappearance of the allylic protons.

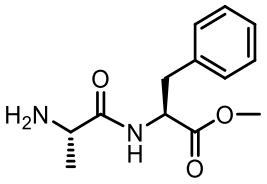
7.3 Compound synthesis

H-UF-OMe (1)



To a solution of Boc-Sec-OH (0.25 g, 0.468 mmol), H-Phe-OMe (0.22 g, 1.03 mmol) and NMM (0.17 ml, 1.40 mmol) in DMF (10 ml) at room temperature was added DMTMM (0.28 g, 1.03 mmol) in one portion. This was left to stir overnight before being diluted with water (50 ml) and extracted with EtOAc (3 x 50 ml). The combined organic extracts were washed with 1 M HCl (3 x 50 ml), 1 M $NaHCO_3$ (3 x 50 ml), water (1 x 50 ml) and brine (1 x 50 ml), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The resultant white foam was redissolved in 50% TFA in DCM (50 ml) and stirred for 1 hour before being concentrated *in vacuo*. The clear oil was co-evaporated with toluene (3 x 25 ml) to fully evaporate TFA, resulting in an off-white solid. This was washed with diethyl ether and dried under high vacuum, yielding a free-flowing, off-white solid (0.27 g, 88% yield, 0.41 mmol). HRMS Calc. = 659.0887 $[M+H]^+$, Obs. = 659.0935 $[M+H]^+$. ^{1}H NMR (400 MHz, CD_3OD) δ 7.40 – 7.17 (m, 5H), 4.81 (dd, J = 8.5, 5.8 Hz, 1H), 4.20 (dd, J = 7.9, 4.8 Hz, 1H), 3.72 (s, 3H), 3.54 (dd, J = 13.7, 4.9 Hz, 1H), 3.30 (dd, J = 13.7, 4.9 Hz, 1H), 3.23 (dd, J = 14.0, 5.8 Hz, 1H), 3.05 (dd, J = 14.0, 5.8 Hz, 1H). ^{13}C NMR (101 MHz, CD_3OD) δ 171.6 (C), 167.5 (C), 136.39 (Ar), 128.8 (Ar-H), 128.3 (Ar-H), 126.7 (Ar-H), 54.3 (CH), 52.6 (CH_3), 51.6 (CH), 36.9 (CH_2), 28.7 (CH_2).

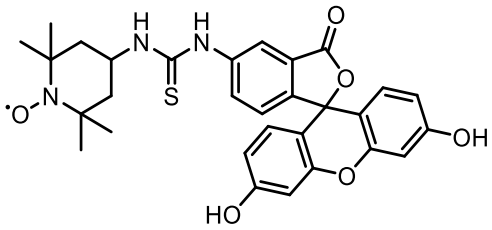
H-AF-OMe (4)



To a solution of Boc-Ala-OH (1.03 g, 5.48 mmol) and H-Phe-OMe.HCl (1.30 g, 6.03 mmol) in DMF (10 ml) was added DMTMM (1.70 g, 6.03 mmol) portion-wise. This was left to stir at room temperature overnight before being diluted with water (50 ml) and extracted with ethyl acetate (3 x 50 ml). The combined organic layers were then washed with 1 M HCl (3 x 20 ml), 1 M $NaHCO_3$ (3 x 20 ml), water (1 x 20 ml) and brine (1 x 20 ml), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The resulting white foam was redissolved in 50% TFA in DCM (100 ml) and stirred for 1 hour. This was concentrated *in vacuo,* with excess TFA co-evaporated

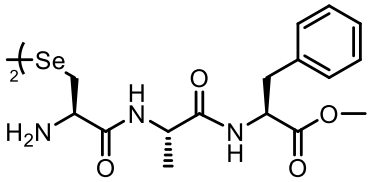
with toluene (3 x 25 ml). The resulting white solid was washed with diethyl ether (3 x 50 ml) to give the title compound as a free-flowing white powder (1.26 g, 5.04 mmol, 92% yield). HRMS Calc.: 251.1396 $[M+H]^+$. Obs.: 251.1387 $[M+H]^+$. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.86 (d, J = 7.5 Hz, 1H), 8.12 (s, 3H), 7.47 – 7.16 (m, 5H), 4.54 (ddd, J = 9.2, 7.5, 5.3 Hz, 1H), 3.84 (q, J = 7.0 Hz, 1H), 3.63 (s, 3H), 3.10 (dd, J = 14.0, 5.3 Hz, 1H), 2.96 (dd, J = 14.0, 9.3 Hz, 1H), 1.34 (d, J = 6.9 Hz, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 171.9 (C), 170.3 (C), 137.4 (Ar), 129.5 (Ar-H), 128.8 (Ar-H), 127.2 (Ar-H), 54.4 (CH), 52.5 (CH_3), 48.4 (CH), 36.7 (CH_2), 17.5 (CH_3).

Fluorescein isothiocyanate-TEMPO (7)



To a solution of 4-amino TEMPO (0.146 g, 0.852 mmol) in THF (6 ml) was added FITC (0.316 g, 0.811 mmol) in one portion. The suspension was left to stir at room temperature overnight. The reaction mixture was poured into water (25 ml), and the aqueous layer extracted with EtOAc (3 x 20 ml). The combined organic layers were washed with 1 M HCl (2 x 10 ml), water (1 x 10 ml) and brine (1 x 10 ml), before being dried over $MgSO_4$, filtered and concentrated *in vacuo*. The resulting orange solid was purified by column chromatography (1:4 Hexane: Ethyl acetate) to give the title compound as an orange solid (0.364 g, 80%, 0.649 mmol). HRMS Calc.: 561.1934 $[M+H]^+$. Obs: 561.1951 $[M+H]^+$. NMR Method B used for NMR analysis. ^{1}H NMR (400 MHz, CD_3OD) δ 8.31 (d, *J* = 2.1 Hz, 1H), 7.85 (d, *J* = 2.1 Hz, 1H), 7.21 (d, J=2.1 Hz, 1H), 6.91 (m, 4H), 6.71 (d, *J* = 2.4 Hz, 2H), 5.03 (m, 1H), 2.43 (m, 2H), 1.91 (t, *J* = 13.3 Hz, 2H), 1.53 (d, *J* = 35.9 Hz, 12H). ^{13}C NMR (101 MHz, CD_3OD) δ 181.3 (C), 168.9 (C), 162.8 (Ar-H), 161.1 (Ar), 154.4 (Ar), 141.4 (Ar), 129.9 (Ar-H), 128.4 (Ar-H), 125.6 (Ar-H), 120.2 (Ar-H), 117.7 (Ar-H), 114.8 (Ar-H), 114.1 (Ar-H), 111.8 (Ar-H), 102.1 (Ar-H), 68.5 (C), 44.0 (CH), 41.3 (CH_2), 26.8 (CH_3), 19.0 (CH_3). Melting point – decomposition at 190 ^{o}C. IR (ATR, cm^{-1}) - 3202 brw, 3053w, 2973w, 2934w, 1732w, 1605m, 1530m, 1502m, 1443m, 1381m, 1314m, 1238m, 1206s, 1169s, 1108s, 1047m.

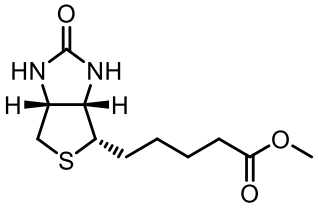
H-UAF-OMe (10)



To Boc-Sec-OH (0.1 g, 0.187 mmol) and H-AF-OMe (**4**, 0.094 g, 0.374 mmol) in DMF (5 ml) was added NMM (0.062 ml, 0.561 mmol) and DMTMM (0.103 g, 0.374 mmol). The solution was stirred at room temperature for 16 hours before being diluted with 1 M HCl (5 ml) and extracted with ethyl acetate (3 x 20 ml). The combined organic layers were washed with 1 M HCl (3 x 5 ml), saturated $NaHCO_3$ (3 x 5 ml), water (5 ml) and brine (5 ml), then dried over $MgSO_4$, filtered and concentrated in vacuo. The resultant residue was dissolved in a 1:1 solution of DCM: TFA (10 ml) and stirred at room temperature for 30 minutes before being concentrated in vacuo. The remaining TFA was co-evaporated using toluene (2 x 5 ml). The resultant solid was washed with diethyl ether (3 x 20 ml), then dried under high vacuum to give the title compound as a pale-yellow solid (0.125 g, 0.157 mmol, 84% yield). HRMS Calc. $[M+H]^+$ = 801.1650. Obs. $[M+H]^+$ = 801.1670. ^{1}H NMR (400 MHz, CD_3OD) δ 8.53 (d, *J* = 7.7 Hz, 1H), 7.35 – 7.11 (m, 5H), 4.64 (m, 1H), 4.58 (m, 1H), 4.19 (dd, *J* = 8.0, 4.0 Hz, 1H), 3.67 (s, 3H), 3.47 (dd, *J* = 13.6, 5.0 Hz, 1H), 3.18 – 3.10 (m, 2H), 3.01 (dd, *J* = 13.9, 8.4 Hz, 1H), 1.37 (d, *J* = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, CD_3OD) δ 174.6 (C), 173.3 (C), 168.5 (C), 138.2 (Ar), 130.4 (Ar-H), 129.8 (Ar-H), 128.1 (Ar-H), 55.7 (CH), 54.0 (CH), 52.9 (CH_3), 50.4 (CH), 38.4 (CH_2), 30.3 (CH_2), 18.7 (CH_3).

Biotin-EDA-TEMPO (21)

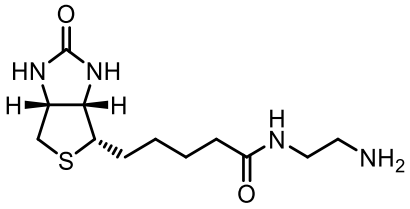
Biotin-OMe (124)



Biotin-OMe (**124**) was synthesised following a literature procedure.[175] To methanol (10 ml) at 0 °C was added acetyl chloride (1.46 ml, 20.5 mmol) dropwise. This was stirred for 5 minutes before adding dropwise to a suspension of Biotin (1.0 g, 4.09 mmol) in methanol (10 ml) at 0 °C. The resulting solution was allowed to warm to room temperature and stirred for 1 hour. Over this time the solution became clear. The reaction mixture was concentrated *in vacuo*, then redissolved in 2% methanol in DCM (50 ml). This solution was extracted with saturated $NaHCO_3$ (3 x 10 ml), water (10 ml) and brine (10 ml), then dried over $MgSO_4$, filtered and

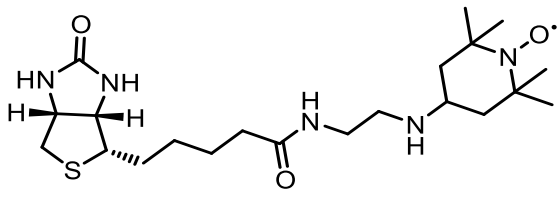
concentrated *in vacuo* to give the title compound as an off-white solid (1.04 g, 4.01 mmol, 98% yield). ^{1}H NMR (400 MHz, DMSO-d_6) δ 6.50 – 6.43 (s, 1H), 6.40 – 6.33 (s, 1H), 4.39 – 4.24 (m, 1H), 4.14 (dd, J = 7.7, 4.4 Hz, 1H), 3.59 (s, 3H), 3.10 (ddd, J = 8.6, 6.2, 4.4 Hz, 1H), 2.83 (dd, J = 12.5, 5.1 Hz, 1H), 2.58 (d, J = 12.4 Hz, 1H), 2.51 (p, J = 1.8 Hz, 1H), 2.30 (t, J = 7.4 Hz, 2H), 1.70 – 1.26 (m, 6H).^{13}C NMR (101 MHz, DMSO-d_6) δ 173.8 (C=O), 163.2 (C=O), 61.5 (CH), 61.4 (CH), 59.6 (CH), 59.5 (CH), 55.8 (CH), 51.7 (CH_3), 33.6 (CH_2), 28.5 (CH_2), 24.9 (CH_2).

Biotin-ethylenediamine (25)



Biotin-ethylenediamine (**25**) was synthesised following a literature procedure.[175] To Biotin-OMe (**124**, 1.0 g, 3.87 mmol) in methanol (25 ml) was added ethylenediamine (6.5 ml, 96.8 mmol). The resulting solution was stirred at 60 ^{o}C for 16 hours. TLC was used to confirm complete consumption of the starting material at this point. The reaction mixture was concentrated *in vacuo*, with excess ethylenediamine being co-evaporated with toluene (3 x 20 ml). The resulting solid was washed with ethyl acetate (3 x 25 ml) and dried under high vacuum to give the title compound as an off-white solid (1.0 g, 3.49 mmol, 90% yield). HRMS Calc.: 287.1542 $[M+H]^+$. Obs.: 287.1563 $[M+H]^+$. ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.74 (t, J = 5.6 Hz, 1H), 6.44 (s, 1H), 6.37 (s, 1H), 4.35 – 4.28 (m, 1H), 4.13 (ddd, J = 7.8, 4.5, 1.9 Hz, 1H), 3.10 (ddd, J = 8.5, 6.1, 4.4 Hz, 1H), 3.02 (q, J = 6.2 Hz, 2H), 2.83 (dd, J = 12.4, 5.1 Hz, 1H), 2.58 (d, J = 15.4 Hz, 1H), 2.53 (t, J = 6.4 Hz, 2H), 2.06 (t, J = 7.4 Hz, 2H), 1.69 – 1.22 (m, 6H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 172.5 (C=O), 163.2 (C=O), 61.5 (CH), 59.7 (CH), 55.9 (CH), 42.7 (CH_2), 41.9 (CH_2), 40.6 (CH_2), 35.7 (CH_2), 28.7 (CH_2), 28.5 (CH_2), 25.8 (CH_2).

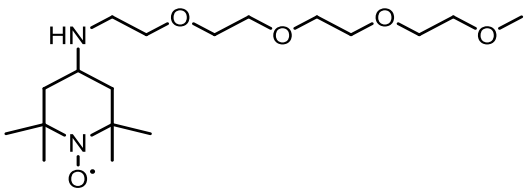
Biotin-EDA-TEMPO (21)



To 4-oxoTEMPO (0.5 g, 2.94 mmol) and biotin-EDA (**25**, 1.01 g, 3.52 mmol) in dry methanol (5 ml) under nitrogen was added glacial acetic acid (0.15 ml, 2.65 mmol). The solution was heated to 50 ^{o}C and stirred under nitrogen for 3 hours. At this time the solution was allowed to cool to room temperature and sodium cyanoborohydride (0.20 g, 3.23 mmol) was added in a single portion. The reaction mixture was again heated to 50 ^{o}C and stirred under nitrogen

for 16 hours. The reaction was quenched by addition of saturated sodium bicarbonate (5 ml) and the aqueous layer extracted with DCM (5 x 10 ml). The combined organics were washed with brine (5 ml), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The resultant orange oil was dried under high vacuum to give the title compound as a pale orange crystalline solid (1.11 g, 2.53 mmol, 86% yield). HRMS – Calc. $[M+H]^+$ - 441.2783. Obs. $[M+H]^+$ - 441.2774. NMR Method B used for NMR analysis. ^1H NMR (400 MHz, CD_3OD) δ 4.53 (dd, *J* = 7.9, 4.4 Hz, 1H), 4.34 (dd, *J* = 7.9, 4.4 Hz, 1H), 3.90 (m, 1H), 3.58 – 3.52 (m, 2H), 3.30 – 3.20 (m, 3H), 2.95 (dd, *J* = 12.8, 5.0 Hz, 1H), 2.73 (d, *J* = 12.8 Hz, 1H), 2.44 (m, 2H), 2.29 (t, *J* = 7.4 Hz, 2H), 2.06 (m, 2H), 1.83 – 1.56 (m, 4H), 1.52 (d, *J* = 4.5 Hz, 12H). ^{13}C NMR (101 MHz, CD_3OD) δ 177.6 (C=O), 166.1 (C=O), 162.8 (C=O), 119.4 (CH), 116.5 (CH), 68.5 (C), 63.3 (CH), 61.7 (CH), 57.0 (CH), 46.2 (CH_2), 41.0 (CH_2), 37.2 (CH_2), 36.4 (CH_2), 29.7 (CH_2), 29.4 (CH_2), 28.2 (CH_2), 26.4 (CH_3), 20.2 (CH_3). Melting Point (˚C) - 54-56. FTIR (ATR) νmax/cm^{-1} 3271brm, 3080w, 2971w, 2929m, 2859w, 1692s, 1648s, 1546m, 1458m, 1376w, 1361w, 1331w, 1308w, 1243w, 1179w, 1120w.

Tetra-EG-TEMPO (22)



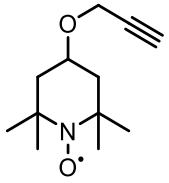
To tetraethylene glycol monomethyl ether (1.0 g, 4.80 mmol) in DCM at 0 ºC was added triethylamine (3.35 ml, 24.0 mmol) followed by mesyl chloride (1.11 ml, 14.4 mmol) dropwise. The reaction mixture was allowed to warm to room temperature and stir for 2 hours. The reaction mixture was washed with 1 M HCl (3 x 10 ml), saturated $NaHCO_3$ (3 x 10 ml), brine (10 ml), dried over $MgSO_4$ and concentrated in vacuo.

To the crude mesylate (1.30 g, 4.54 mmol) in 35% ammonia in water (50 ml) was added ammonium acetate (4.86 g, 90.8 mmol). This was stirred at room temperature for 48 hours. The reaction mixture was saturated with NaCl and extracted with DCM (5 x 20 ml). The combined organic layers were washed with brine (10 ml), dried over $MgSO_4$, filtered and concentrated in vacuo.

The resulting colourless oil (0.85 g, 4.10 mmol) was redissolved in dry MeOH (20 ml) and, under N_2, 4-oxoTEMPO (0.63 g, 3.73 mmol) and glacial acetic acid (0.19 ml, 3.36 mmol) were added. The reaction mixture was heated to 50 ºC and stirred for 2 hours. The reaction mixture was cooled to room temperature and sodium cyanoborohydride (0.70 g, 11.2 mmol) was added and the reaction mixture again heated to 50 ºC and stirred under N_2 for 16 hours. The reaction mixture was cooled, then poured into saturated $NaHCO_3$ (10 ml) and the aqueous

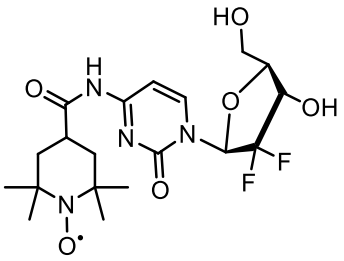
layer extracted with DCM (5 x 20 ml). The organic layers were washed with brine (20 ml), dried over $MgSO_4$, filtered and concentrated in vacuo. The resultant orange oil was purified by column chromatography (DCM: MeOH 9:1) to give the title compound as an orange oil (1.13 g, 3.13 mmol, 84% yield). HRMS Calc.: 362.2781 $[M+H]^+$. Obs.: 362.2797 $[M+H]^+$. NMR Method A used for NMR analysis. ^1H NMR (400 MHz, CD_3OD) δ 3.69 – 3.59 (m, 14H), 3.58 – 3.52 (m, 2H), 3.37 (s, 2H), 2.97 – 2.86 (m, 1H), 2.80 (t, J = 5.3 Hz, 2H), 1.88 – 1.80 (m, 2H), 1.35 – 1.27 (m, 3H), 1.17 (d, J = 9.8 Hz, 12H), 1.05 (s, 1H). ^{13}C NMR (101 MHz, CD_3OD) δ 71.6 (CH), 70.1 (CH), 70.0 (C), 69.5 (CH_2), 58.3 (CH_3), 57.7 (CH_2), 45.5 (CH_2), 44.7 (CH_2), 31.6 (CH_3), 19.2 (CH_3). IR (ATR, cm^{-1}) 3276m, 3125 brw, 2989w, 2967m, 2929m, 2887m, 2854m, 1487s, 1391m, 1378m, 1179w, 1137s, 1126s, 1116m, 1098w.

Propargyl TEMPO (23)



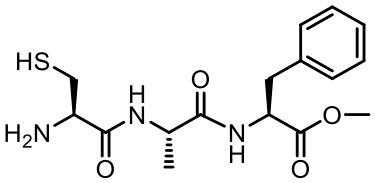
To a solution of NaH (60% in mineral oil, 0.70 g, 17.4 mmol) in anhydrous DMF (5 ml) under N_2 was added 4-hydroxy TEMPO (1.0 g, 5.81 mmol) in anhydrous DMF (5 ml). The solution was stirred for 30 minutes before being cooled to 0 ^oC and propargyl bromide (80% in toluene, 2.07 g, 17.4 mmol) was added dropwise over 15 minutes. The solution was then left to stir at room temperature under N_2 overnight. The solution was poured into water (25 ml), and the solution extracted with ethyl acetate (3 x 25 ml). The combined organics were washed with brine (25 ml), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The resulting brown oil was purified by column chromatography (95:5 hexane: ethyl acetate) to give the title compound as an orange solid (0.81 g, 3.86 mmol, 66% yield). HRMS Calc.: 211.1572 $[M+H]^+$. Obs.: 211.1574 $[M+H]^+$. NMR Method A used for NMR analysis. ^1H NMR (400 MHz, $CDCl_3$) δ 4.20 (d, J = 2.4 Hz, 2H), 3.88 (tt, J = 11.2, 4.1 Hz, 1H), 2.46 (t, J = 2.4 Hz, 1H), 1.99 (m, 2H), 1.58 – 1.42 (m, 2H), 1.24 (d, J = 18.8 Hz, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 74.2 (C), 69.8 (CH), 59.6 (CH), 55.3 (CH), 44.3 (CH_2), 32.0 (CH_3), 20.7 (CH_3). Melting point – 58-59 ^oC. FTIR (ATR) νmax/cm^{-1} 3229m, 2996w, 2973w, 2953w, 2936w, 2870w, 2111w, 1466w, 1445w, 1379w, 1365m, 1348m, 1312m, 1020s, 1000m.

Gemcitabine-TEMPO (24)



To gemcitabine hydrochloride (0.25 g, 0.834 mmol) and 4-carboxyTEMPO (0.18 g, 0.918 mmol) in DMF: DMSO 3:1 (10 ml) was added NMM (0.10 ml, 0.918 mmol), HOBt (0.12 g, 0.918 mmol) and EDC.HCl (0.21 g, 1.08 mmol). The reaction mixture was stirred at room temperature for 16 hours before adding brine (10 ml) and extracting with EtOAc (3 x 20 ml). The combined organic layers were washed with 20% LiCl (10 ml), sat. $NaHCO_3$ (10 ml), brine (10 ml), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The resultant oil was purified by column chromatography ($CHCl_3$: MeOH 9.5:0.5) to give the title compound as an orange solid (0.15 g, 0.336 mmol, 40%). HRMS Calc.: 446.1977 $[M+H]^+$. Obs.: 446.1978 $[M+H]^+$. NMR Method B used for NMR analysis.^{1}H NMR (400 MHz, CD_3OD) δ 8.40 (d, J = 8.0 Hz, 1H), 7.46 (d, J = 8.0 Hz, 1H), 6.29 (t, *J* = 8 Hz, 1H), 4.31 (m, 1H), 4.01 (m, 2H), 3.89 – 3.80 (m, 1H), 2.21 (dd, J = 14.8, 3.3 Hz, 2H), 2.04 (t, J = 13.6 Hz, 2H), 1.51 (d, J = 13.6 Hz, 12H). ^{13}C NMR (101 MHz, CD_3OD) δ 173.8 (C), 163.4 (C), 156.1 (C), 144.8 (CH), 125.1 (C), 122.5 (CH), 96.9 (CH), 81.5 (CH), 69.0 (CH), 67.6 (C), 58.8 (CH_2), 38.6 (CH_2), 26.8 (CH_3), 18.7 (CH_3). ^{19}F NMR (376 MHz, CD_3OD) δ -119.15 (d, J = 248 Hz, 1F), -119.82 (dd, J = 236 Hz, 1F). Melting point – 212-214 ^{o}C. IR (ATR, cm^{-1}) - 3172m, 2975w, 2933m, 1720 m, 1648s, 1617m, 1559s, 1491s, 1467s, 1437m, 1384s, 1314m, 1275s, 1241m, 1100s, 1083s, 1071s.

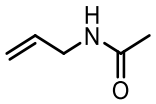
H-CAF-OMe (48)



To H-AF-OMe (**4**, 0.15 g, 0.59 mmol) and Boc-Cys(Trt)-OH (0.25 g, 0.54 mmol) in DMF (10 ml) was added NMM (0.18 ml, 1.62 mmol) and DMTMM (0.16 g, 0.59 mmol). The reaction mixture was stirred at room temperature for 16 hours before being diluted with H_2O (10 ml) and extracted with EtOAc (3 x 20 ml). The combined organic layers were washed with 1 M HCl (3 x 10 ml), saturated $NaHCO_3$ (3 x 10 ml) and brine (10 ml), then dried with $MgSO_4$, filtered and concentrated *in vacuo* to yield a colourless residue. The residue was redissolved in a solution of 1:1:0.05 DCM:TFA:TIS (25 ml) and stirred at room temperature for 30 minutes before being concentrated *in vacuo*. The excess TFA was co-evaporated with toluene (3 x 10 ml). The resulting solid was washed with Et_2O (3 x 20 ml) and dried under high vacuum to

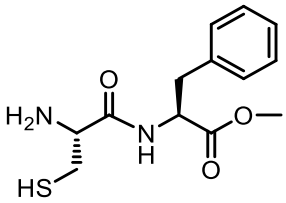
yield an off-white solid. This was purified by reversed-phase preparative HPLC (10-80%B over 30 minutes, 6 ml/min) to give the title compound as a white solid after lyophilisation (0.12 g, 0.34 mmol, 63% yield). HRMS – Calc. $[M+H]^+$ = 354.1488. Obs. $[M+H]^+$ = 354.1500. ^{1}H NMR (400 MHz, CD_3OD) δ 7.36 – 7.19 (m, 5H), 4.71 – 4.64 (m, 1H), 4.63 – 4.56 (m, 1H), 4.21 (q, J = 8.0, 4.9 Hz, 1H), 3.69 (s, 3H), 3.49 (dd, J = 13.6, 5.0 Hz, 1H), 3.15 (m, 2H), 3.03 (dd, J = 13.9, 8.4 Hz, 1H), 1.39 (d, J = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, CD_3OD) δ 173.1 (C), 171.8 (C), 167.0 (C), 136.6 (Ar), 128.9 (Ar-H), 128.2 (Ar-H), 126.6 (Ar-H), 54.2 (CH), 52.5 (CH), 51.4 (CH_3), 48.9 (CH), 36.9 (CH_2), 28.8 (CH_2), 17.1 (CH_3).

N-allylacetamide (66)



To a solution of allyl amine (10.0 ml, 133.2 mmol) and triethylamine (11.1 ml, 80.0 mmol) in THF (100 ml) at 0 °C was added acetyl chloride (9.48 ml, 133.2 mmol) dropwise. The reaction was left to stir at room temperature for 2 hours before being purified by vacuum distillation (15 mbar). The desired product eluted at 90 °C as a colourless oil (12.59 g, 127.2 mmol, 95% yield). HRMS Calc. = 100.0762 $[M+H]^+$. Obs. = 100.0764 $[M+H]^+$. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.11 (s, 1H), 5.82 (ddt, *J* = 17.2, 10.2, 5.7 Hz, 1H), 5.30 – 5.02 (m, 2H), 3.85 (ddd, *J* = 5.7, 4.1, 1.6 Hz, 2H), 1.99 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.2 (C), 134.2 (CH_2), 116.3 (CH), 42.0 (CH_3), 23.1 (CH_2). IR (ATR, cm^{-1}) - 3277brm, 3081m, 2986m, 1634s, 1544s, 1422m, 1372m, 1281m, 1037m, 991m, 918m.

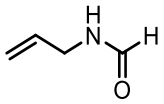
H-CF-OMe (68)



To a solution of Boc-Cys(Trt)-OH (2.0 g, 4.76 mmol), H-Phe-OMe. HCl (1.03 g, 4.76 mmol) and NMM (0.95 ml, 8.66 mmol) in DMF (10 ml) at room temperature was added DMTMM (1.17 g, 4.76 mmol) in one portion. This was left to stir overnight before being diluted with water (50 ml) and extracted with EtOAc (3 x 50 ml). The combined organic extracts were washed with 1 M HCl (3 x 50 ml), 1 M $NaHCO_3$ (3 x 50 ml), water (1 x 50 ml) and brine (1 x 50 ml), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The resultant white foam was redissolved in 50% TFA in DCM (50 ml) and stirred for 1 hour before being concentrated *in vacuo*. The clear oil was co-evaporated with toluene (3 x 25 ml) to fully evaporate TFA, resulting in an off-white solid. This was washed with diethyl ether and dried under high vacuum, yielding a free-flowing, white solid (1.63 g, 4.12 mmol, 95% yield, TFA salt). HRMS Calc. = 283.1116 $[M+H]^+$. Obs. =

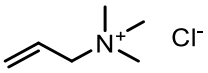
283.1126 $[M+H]^+$. ^1H NMR (400 MHz, DMSO-d_6) δ 8.98 (d, *J* = 7.4 Hz, 1H), 8.29 (s, 3H), 7.57 – 6.86 (m, 5H), 4.58 (ddd, *J* = 9.0, 7.4, 5.4 Hz, 1H), 4.03 (dd, *J* = 6.0, 4.5 Hz, 1H), 3.63 (s, 3H), 3.16 – 2.88 (m, 4H). ^{13}C NMR (101 MHz, DMSO) δ 171.9 (C), 167.6 (C), 137.3 (Ar), 129.5 (Ar-H), 128.9 (Ar-H), 127.2 (Ar-H), 54.5 (CH), 54.0 (CH_3), 52.6 (CH), 36.7 (CH_2), 25.8 (CH_2).

N-allylformamide (82)



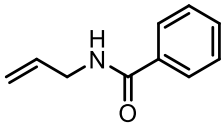
To allylamine (9.86 ml, 131.4 mmol) at 0 °C was added ethyl formate (21.2 ml, 262.8 mmol) dropwise. The reaction mixture was allowed to warm to room temperature and stirred overnight. The solution was concentrated *in vacuo* and purified by vacuum distillation (15 mbar). The product eluted at 85-90 °C as a colourless oil (9.17 g, 107.9 mmol, 82% yield). ^1H NMR (400 MHz, $CDCl_3$) δ 8.17 (s, 1H), 6.46 (s, 1H), 5.81 (ddt, J = 17.1, 10.3, 5.6 Hz, 2H), 5.40 – 4.89 (m, 4H), 3.93 – 3.84 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.0 +161.4 (C), 134.4 + 133.6 (CH_2), 116.8 + 116.5 (CH), 44.2 + 40.4 (CH_2). NMR showed a mixture of rotamers in a 1:0.2 ratio. IR (ATR, cm^{-1}) - 3285 brm, 3051w, 2986w, 2865w, 1656s, 1650s, 1526s, 1382s, 1231m, 991m, 918m.

Allyltrimethylammonium chloride (84)



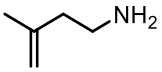
TMA (30% in EtOH, 2.00 ml, 6.43 mmol) was diluted to 10 ml with ethanol and cooled to 0 °C. A 1 M NaOH in water solution (6.43 ml, 6.43 mmol) was cooled and added dropwise to the TMA solution, then left to stir for 10 minutes. Allyl chloride (0.52 ml, 6.43 mmol) was then added dropwise, the RM warmed to RT and left to stir for 16 hours. The RM was concentrated *in vacuo*, the residue dissolved in dry ethanol, filtered and concentrated *in vacuo*. The residue was dissolved in chloroform, filtered and acetone slowly added until crystals formed. The resulting white crystals were dried under high vacuum for 2 hours, (0.49 g, 2.72 mmol, 42% yield). HRMS Calc. = 100.1126 $[M]^+$. Obs. = 100.1146 $[M]^+$. ^1H NMR (400 MHz, DMSO-d_6) δ 6.06 (m, 1H), 5.66 – 5.55 (m, 2H), 4.06 (d, 2H), 3.08 (s, 9H). ^{13}C NMR (101 MHz, DMSO) δ 127.9 (CH_2), 126.8 (CH), 67.4 (CH_3), 52.2 (CH_2). Melting point – 184-186 °C. IR (ATR, cm^{-1}) - 3464brw, 3008m, 2943m, 1640w, 1473s, 1404m, 919s, 889s.

N-allylbenzamide (85)



To N-allylamine (0.64 ml, 8.54 mmol) and triethylamine (2.97 ml, 21.3 mmol) in DCM (25 ml) at 0 ^{o}C was added benzoyl chloride (0.83 ml, 7.11 mmol) dropwise. The reaction mixture was allowed to warm to room temperature and stir for 16 hours before being diluted with DCM (25 ml) and washed with 1 M HCl (3 x 10 ml), sat. $NaHCO_3$ (3 x 10 ml) and brine (10 ml). The organic layer was then dried over $MgSO_4$, filtered and concentrated to produce a pale yellow oil. This was purified by column chromatography (DCM) to yield the title compound as a colourless oil, which formed colourless crystals when stored at 4 ^{o}C (0.96 g, 5.96 mmol, 84% yield). HRMS Calc. = 162.0919 $[M+H]^+$. Obs. = 162.0920 $[M+H]^+$. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.67 (s, 1H), 7.88 (d, J = 7.0 Hz, 2H), 7.59 – 7.39 (m, 3H), 5.91 (ddt, J = 17.3, 10.3, 5.2 Hz, 1H), 5.23 – 5.03 (m, 2H), 4.00 – 3.83 (m, 2H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 166.8 (C), 135.9 (Ar), 134.9 (Ar-H), 131.6 (Ar-H), 128.7 (Ar-H), 127.7 (CH_2), 115.6 (CH), 42.0 (CH_2). IR (ATR, cm^{-1}) - 3305 brm, 3065w, 2985w, 2916w, 1636s, 1532s, 1420w, 1294s, 1028w, 992s.

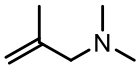
3-methyl-3-buten-1-amine (98)



To a solution of 3-methyl-3-buten-1-ol (**93**, 5.86 ml, 58.1 mmol) and mesyl chloride (13.5 ml, 174 mmol) in DCM (50 ml) at 0 ^{o}C was added triethylamine (24.3 ml, 174 mmol) dropwise. After addition the reaction mixture was allowed to warm to room temperature and stirred for 4 hours. The reaction mixture was then diluted with DCM (50 ml) and washed with 1 M HCl (3 x 20 ml), sat. $NaHCO_3$ (3 x 20 ml) and brine (20 ml), dried over $MgSO_4$, filtered and concentrated to give the crude mesylate (9.4 g, 56.9 mmol, 98% yield).

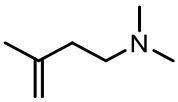
To a portion of the crude mesylate (2.0 g, 12.2 mmol) was added ammonia solution (30 ml, 304 mmol, 35% solution in water) followed by ammonium acetate (18.8 g, 244 mmol). The solution was stirred at room temperature for 48 hours before being saturated with potassium carbonate and extracted with diethyl ether (5 x 20 ml). The combined organic extracts were dried over $MgSO_4$, filtered and purified by distillation (85 ^{o}C elution) to give the title compound as a colourless liquid (0.8 ml, 9.4 mmol, 77% yield). ^{1}H NMR (400 MHz, D_2O) δ 4.77 (t, J = 1.9 Hz, 1H), 2.65 (t, J = 6.8 Hz, 2H), 2.10 (t, J = 6.8 Hz, 2H), 1.65 (d, J = 1.4 Hz, 3H). Second alkene proton singlet overlaps with D_2O peak. ^{13}C NMR (101 MHz, D_2O) δ 144.7 (C), 111.3 (CH_2), 39.8 (CH_2), 38.1 (CH_2), 21.2 (CH_3). IR (ATR, cm^{-1}) - 3364brw, 3291brw, 3074w, 2986m, 2934w, 1648m, 1598w, 1442m, 1375m, 1056w, 884s.

N,N,2-trimethylprop-2-en-1-amine (99)



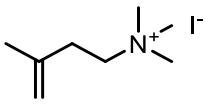
To dimethylamine (46.0 ml, 408 mmol, 40% solution in H_2O) at 0 ^oC was added 3-chloro-2-methylprop-1-ene (10.0 ml, 102 mmol) dropwise. After complete addition the reaction mixture was allowed to warm to room temperature and stir for 16 hours. The reaction mixture was saturated with potassium carbonate, then extracted with diethyl ether (5 x 20 ml). The combined organics were dried over $MgSO_4$, filtered and purified by distillation. The title compound eluted as a colourless liquid at 82 ^oC (6.21 g, 62.6 mmol, 61% yield). HRMS Calc. = 100.1126 $[M+H]^+$. Obs. = 100.1128 $[M+H]^+$. ^1H NMR (400 MHz, $CDCl_3$) δ 4.91 – 4.84 (m, 2H), 2.84 (s, 2H), 2.22 (s, 6H), 1.78 (s, 2H). 13C NMR (101 MHz, $CDCl_3$) δ 143.1 (C), 113.3 (CH_2), 66.9 (CH_2), 45.3 (CH_3), 20.8 (CH_3). IR (ATR, cm^{-1}) – 3371brw, 3075w, 2969m, 2942m, 2858m, 2815m, 2765m, 1649m, 1537s, 1374m, 1097m, 884s.

N,N-3-trimethylbut-3-en-1-amine (100)



The mesylate of 3-methyl-3-buten-1-ol was prepared as previously described for compound **98**. To the crude mesylate (2.5 g, 15.2 mmol) at 0 ^oC was added dimethylamine (8.6 ml, 76.1 mmol, 40% solution in H_2O) dropwise. The solution was allowed to warm to room temperature and stirred for 3 hours before being saturated with potassium carbonate and extracted with diethyl ether (5 x 20 ml). The combined organic extracts were dried over $MgSO_4$, filtered and purified by distillation (115 ^oC elution) to yield the title compound as a colourless liquid (1.5 g, 13.3 mmol, 88% yield). HRMS Calc. = 114.1283 $[M+H]^+$. Obs. = 114.1288 $[M+H]^+$. ^1H NMR (400 MHz, $CDCl_3$) δ 4.77 (dd, J = 2.6, 1.2 Hz, 1H), 4.72 (dt, J = 2.1, 1.1 Hz, 1H), 2.43 (dd, J = 8.8, 6.7 Hz, 2H), 2.27 (s, 6H), 2.22 (t, J = 7.7 Hz, 2H), 1.76 (t, J = 1.1 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.9 (C), 110.9 (CH_2), 58.0 (CH_2), 45.4 (CH_3), 35.8 (CH_2), 22.5 (CH_3). IR (ATR, cm^{-1}) - 3380brw, 3074w, 2970m, 2941m, 2815m, 2765m, 1649w, 1461s, 1306w, 1040m, 884s.

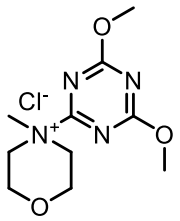
N,N,N-3-tetramethylbut-3-en-1-aminium iodide (101)



To N,N-3-trimethylbut-3-en-1-amine (**98**, 0.15 g, 1.33 mmol) in diethyl ether (25 ml) at 0 ^oC was added methyl iodide (0.25 ml, 3.98 mmol) dropwise. The resulting solution was allowed to warm to room temperature and stir for 2 hours, where a white precipitate formed over time. The solution was filtered and the solid washed with diethyl ether. The white powder was then

recrystallized from chloroform to produce the title compound as off-white needles (0.24 g, 0.941 mmol, 70% yield). HRMS Calc. = 128.1439 $[M]^+$. Obs. = 128.1452 $[M]^+$. ^{1}H NMR (400 MHz, CD_3OD) δ 4.94 (t, J = 1.6 Hz, 1H), 4.90 (q, J = 1.1 Hz, 1H), 3.54 – 3.46 (m, 2H), 3.18 (s, 9H), 2.60 – 2.51 (m, 2H), 1.83 (t, J = 1.2 Hz, 3H). ^{13}C NMR (101 MHz, CD_3OD) δ 120.1 (C), 114.2 (CH_2), 66.2 (CH_2), 53.7 (CH_3), 53.7 (CH_3) , 53.6 (CH_3), 31.8 (CH_2), 22.6 (CH_3). Melting point – 185-187 ^{o}C. IR (ATR, cm^{-1}) - 3081w, 3005w, 2979w, 2947w, 1650w, 1477m, 1444w, 897s.

DMTMM (147)



DMTMM (**147**) was synthesised following a literature procedure.[176] To a solution of CDMT (2.5 g, 14.2 mmol) in THF (25 ml) was added NMM (1.49 ml, 13.5 mmol) dropwise. Almost immediately a white precipitate began to form. The reaction was left to stir for a further 15 minutes before being filtered, and the white crystals washed with THF. The crystals were then dried under high vacuum to give the title compound as free-flowing white crystals (3.51 g, 12.69 mmol, 94%). HRMS Calc.: 241.1301 $[M]^+$ Obs.: 241.1321 $[M]^+$. ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.36 (d, 2H), 4.11 (s, 6H), 4.05 – 3.89 (m, 4H), 3.82 – 3.71 (m, 2H), 3.50 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 173.8 (C), 170.6 (C), 149.9 (CH_3), 61.8 (CH_3), 59.9 (CH_2), 57.2 (CH_2).

7.4 Peptides synthesised

The following is a list of peptides that have been synthesised, along with their respective expected and observed mass values. Pure HPLC traces with analytical/preparative solvent gradients can be found in the appendix. U = Sec. alG = allyl glycine.

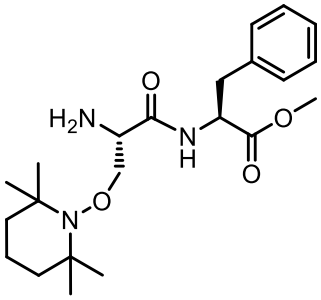
Compound number	Sequence	Calculated Mass	Observed Mass	Yield (%)
13	Ac-UHISKY-OH	$[M+2H]^{2+}$: 839.3081, $[M+3H]^{3+}$: 559.8746	$[M+2H]^{2+}$: 839.3119, $[M+3H]^{3+}$: 559.8803	41 mg (46%)
14	Ac-AHISKY-OH	$[M+H]^{+}$: 760.3994, $[M+2H]^{2+}$: 380.7036	$[M+H]^{+}$: 760.3999, $[M+2H]^{2+}$: 380.7045	29 mg (54%)
15	Ac-CHISKY-OH	$[M+H]^{+}$: 792.3714, $[M+2H]^{2+}$: 396.6896	$[M+H]^{+}$: 792.3714, $[M+2H]^{2+}$: 396.6896	48 mg (58%)
16	H-UWIMKY-NH_2	$[M+2H]^{2+}$ = 889.3423, $[M+3H]^{3+}$: 593.2308	$[M+2H]^{2+}$ = 889.3464, $[M+3H]^{3+}$: 593.2349	31 mg (17%)
32	Ac-YEPLA-SePh	$[M+H]^{+}$: 774.2617	$[M+H]^{+}$: 774.2634	25 mg (18%)
33	H-UHISKY-NH_2	$[M+2H]^{2+}$: 796.3135, $[M+3H]^{3+}$: 531.2137, $[M+4H]^{4+}$: 398.6622.	$[M+2H]^{2+}$: 796.3166, $[M+3H]^{3+}$: 531.2162, $[M+4H]^{4+}$: 398.6650.	26 mg (17%)
39	H-UHISCY-NH_2	$[M+H]^{+}$: 770.2199	$[M+H]^{+}$: 770.2297.	26 mg (17%)

44	ZEGFR-1 Selenoester	$[M+2H]^{2+}$: 1787.8010, $[M+3H]^{3+}$: 1192.2033, $[M+4H]^{4+}$: 894.4044	$[M+2H]^{2+}$: 1788.2217, $[M+3H]^{3+}$: 1192.5592, $[M+4H]^{4+}$: 894.6604.	40 mg (6%)
45	ZEGFR-2	$[M+2H]^{2+}$: 1603.2863, $[M+3H]^{3+}$: 1069.1935, $[M+4H]^{4+}$: 802.1471	$[M+2H]^{2+}$: 1603.2895, $[M+3H]^{3+}$: 1069.5331, $[M+4H]^{4+}$: 802.4000	78 mg (12%)
49	Ac-CWHISKEY-NH_2	$[M+2H]^{2+}$: 553.7586	$[M+2H]^{2+}$: 553.7610.	89 mg (40%)
55	H-CHISKY-NH_2	$[M+H]^{+}$: 749.3769, $[M+2H]^{2+}$: 375.1924	$[M+H]^{+}$: 749.3782, $[M+2H]^{2+}$: 375.1939	45 mg (30%)
56	Ac-YEPLACHISKY-NH_2	$[M+H]^{+}$: 1364.6673, $[M+2H]^{2+}$: 682.8376	$[M+H]^{+}$: 1315.6316, $[M+2H]^{2+}$: 682.8393	44 mg (16%)
75	Ac-CAY-NH_2	$[M+H]^{+}$: 397.1546	$[M+H]^{+}$: 397.1542	38 mg (48%)
76	Ac-D-CAY-NH_2	$[M+H]^{+}$: 397.1546	$[M+H]^{+}$: 397.1545	32 mg (41%)
79	H-HRLLRCGNYA-NH_2	$[M+H+TFA]^{+}$: 1315.6305, $[M+2H]^{2+}$: 601.3228	$[M+H+TFA]^{+}$: 1315.6316, $[M+2H]^{2+}$: 601.2185.	50 mg (21%)

122	H-S(2-methylallyl)AFAC-NH_2	$[M+H]^+$: 551.2652	$[M+H]^+$: 551.2649	12 mg (44%)
129	H-alGAFAC-NH_2	$[M+H]^+$: 507.2390	$[M+H]^+$: 507.2403	14 mg (55%)
131	H-alGMPQLRC-NH_2	$[M+H]^+$: 901.4541, $[M+2H]^{2+}$: 451.2310	901.4544, $[M+2H]^{2+}$: 451.2319.	12 mg (27%)
132	H-alGYEPLAWHISKEYC-NH_2	$[M+H]^+$: 1735.8392, $[M+2H]^{2+}$: 868,4235	$[M+H]^+$: 1735.7704, $[M+2H]^{2+}$: 868.4207.	20 mg (23%)
136	H-alGEDLFYQC-NH_2	$[M+2H]^{2+}$: 507.2390	$[M+2H]^{2+}$: 507.2403	19 mg (38%)
137	H-alGDKFNHEC-NH_2	$[M+2H]^{2+}$: 494.7195	$[M+2H]^{2+}$: 494.7206	14 mg (28%)
140	H-DalGFYQC-NH_2	$[M+H]^+$: 771.3136	$[M+H]^+$: 771.3160	16 mg (42%)
142	H-alGYIQNCPLG-NH_2	$[M+2H]^{2+}$: 502.2557	$[M+2H]^{2+}$: 502.2574	18 mg (36%)
144	H-CYIQNalGPLG-NH_2	$[M+H]^+$: 1003.5035	$[M+H]^+$: 1003.5076	20 mg (40%)

7.5 TEMPO Bioconjugation Products

H-UF-OMe TEMPO Trap Product (3)



To a solution of H-UF-OMe (**1**, 19.8 mg, 0.06 mmol monomer) in 0.01 M NMM (12 ml, pH 7) and methanol (7.2 ml) was added a solution of TCEP (0.625 M, 2.4 ml, 25 eq.) and TEMPO (0.5 M, 2.4 ml, 20 eq.). The solution was left at room temperature, with doping at 60 and 120 minutes. The reaction was left for 4 hours before being lyophilised. The resulting residue was purified using a 15-80% B over 30 minutes 6 ml/min method, yielding the desired product after lyophilisation of the correct fractions (21.6 mg, 89% yield, 0.053 mmol). ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.98 (d, J = 7.7 Hz, 1H, N*H*), 8.26 (d, J = 5.1 Hz, 3H, NH_3), 7.69 – 6.92 (m, 5H, Ar*H*), 4.64 (td, J = 7.5, 6.4 Hz, 1H, CH_α), 4.06 – 3.85 (m, 3H, CH_α, C$H_{2\beta}$), 3.59 (s, 3H, OCH_3), 3.17 – 2.88 (m, 2H, C$H_{2\beta}$), 1.74 – 1.20 (m, 6H, TEMPO CH_2), 1.19 – 0.88 (m, 12H, TEMPO CH_3). ^{13}C NMR (101 MHz, DMSO-d_6) δ 171.4, 166.8, 158.8, 158.4, 136.9, 129.6, 128.9, 127.3, 74.0, 60.2, 54.2, 52.5, 52.0, 37.5, 32.9, 20.3, 16.9.

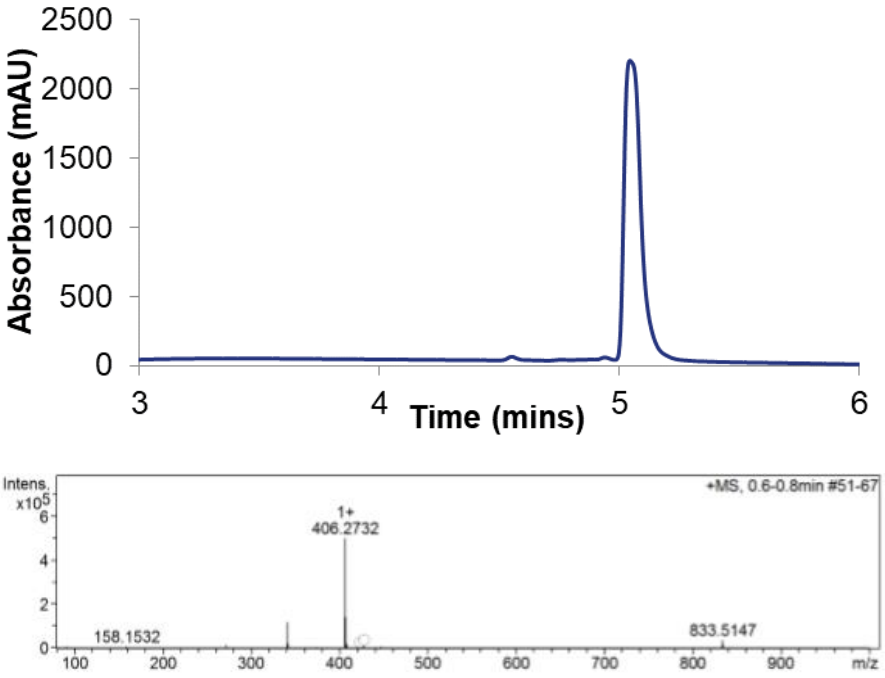
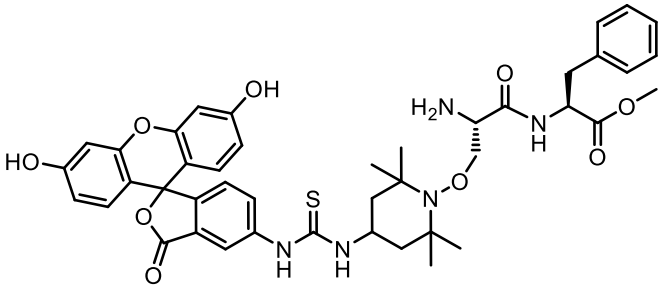


Figure 79 - Analytical HPLC trace of pure TEMPO modified H-UF-OMe (**3**). Analytical gradient 10-70% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 406.2706. Observed Mass $[M+H]^+$: 406.2732.

H-UF-OMe FITC-TEMPO Trap Product (8)



To a solution of H-UF-OMe (**1**, 1.25 mg, 3.79 µmol) in 60% DMSO: 40% 0.01 M NMM (pH 7.5) (1.349 ml) was added TCEP (0.625 M, 0.12 ml, 0.15 mmol, 20 eq.) and FITC-TEMPO (**7**, 0.24 M, 0.031 ml, 2 eq.). The reaction was shaken at room temperature and doped with more TCEP (10 eq) and FITC TEMPO (2 eq) at 15, 60 and 120 minutes. The reaction was left at room temperature for 16 hours before being purified by preparative HPLC (30-80% B over 30 minutes, 6 ml/min). Impure fractions were re-purified by semi-preparative HPLC (30-60% B over 30 minutes, 6 ml/min). The pure fractions were lyophilised to produce an orange solid (2.2 mg, 2.72 µmol, 72% yield). ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.13 (s, 2H), 9.86 (s, 1H), 8.98 (d, J = 7.8 Hz, 1H, TEMPO-N*H*), 8.33 – 8.12 (m, 4H), 8.00 (d, J = 7.4 Hz, 1H), 7.73 (d, J = 8.6 Hz, 1H), 7.24 (ddd, J = 38.1, 18.4, 7.7 Hz, 6H), 6.71 – 6.53 (m, 6H), 4.63 (dd, 2H, TEMPO C*H,* CH_{α}), 3.61 (s, 3H, OCH_3), 3.04 (m, 3H, C$H_{2\beta}$, CH_{α}), 1.90 – 1.78 (m, J = 12.5 Hz, 2H, TEMPO CH_2), 1.43 (t, J = 12.5 Hz, 2H, TEMPO CH_2), 1.22 – 1.10 (m, 12H, TEMPO CH_3).

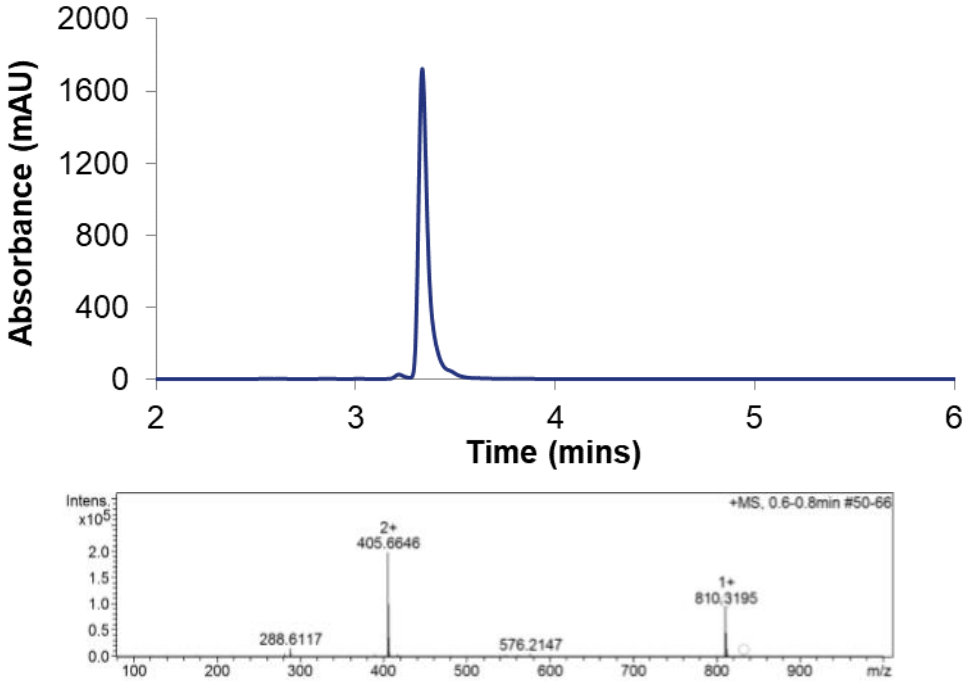
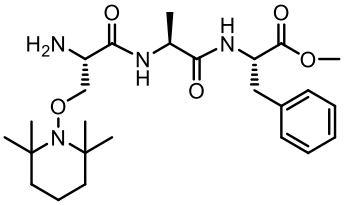


Figure 80 - Analytical HPLC trace of pure FITC-TEMPO modified H-UF-OMe (**8**). Analytical gradient 30-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 810.3173, $[M+2H]^{2+}$: 405.6587. Observed Mass $[M+H]^+$: 406.2732, $[M+2H]^{2+}$: 405.6646.

H-UAF-OMe modified with TEMPO (11)



Synthesised using general protocol A using H-UAF-OMe (**10**) (8 mg, 0.02 mmol). Purified by preparative RP-HPLC (15 to 80% B over 30 minutes) and lyophilized to produce the desired modified peptide (7.9 mg, 0.0166 mmol, 83% yield). ^{1}H NMR (400 MHz, CD_3OD) δ 7.34 – 7.18 (m, 5H, ArH), 4.64 (dd, *J* = 8.8, 5.4 Hz, 1H, CHα), 4.49 (q, *J* = 7.0 Hz, 1H, CHα), 4.16 – 4.04 (m, 3H, CHα and $CH_{2\beta}$), 3.71 (s, 3H, OCH_3), 3.08 (qd, *J* = 56, 13.9, 8 Hz, 2H, CH_β), 1.73 – 1.52 (m, 2H, CH_2), 1.54 – 1.39 (m, 4H, CH_2), 1.39 (d, *J* = 7.0 Hz, 3H), 1.27 – 0.96 (m, 12H, TEMPO CH_3).^{13}C NMR (101 MHz, CD_3OD) δ 172.6, 171.9, 165.7, 136.7, 128.8, 128.2, 126.6, 73.6, 60.2, 60.1, 54.1, 52.3, 51.3, 48.8, 39.4, 36.8, 31.9, 19.2, 17.6, 16.5.

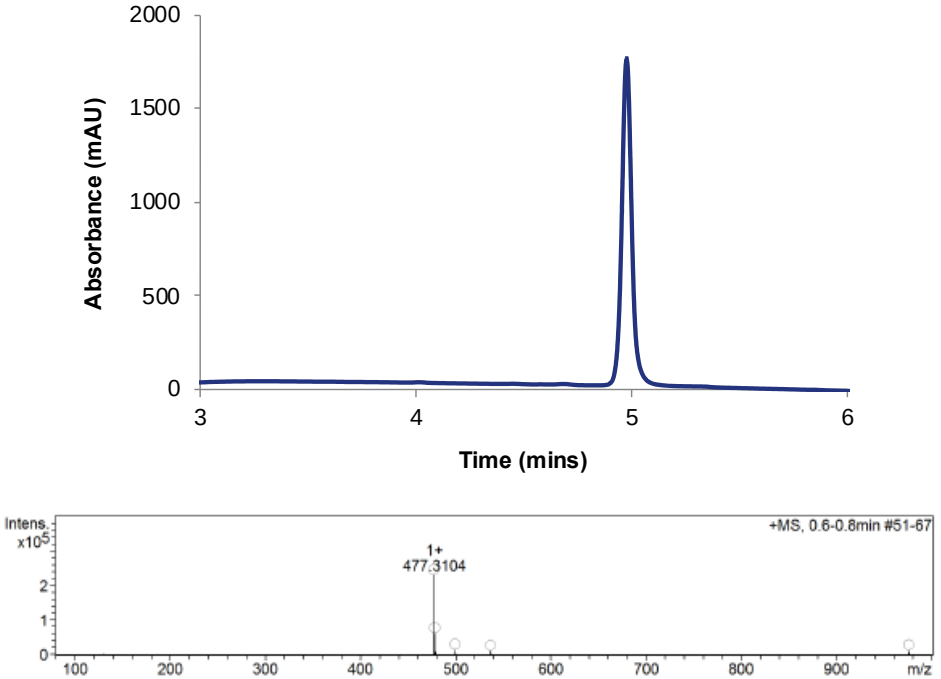
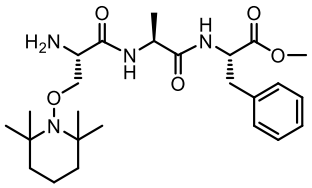
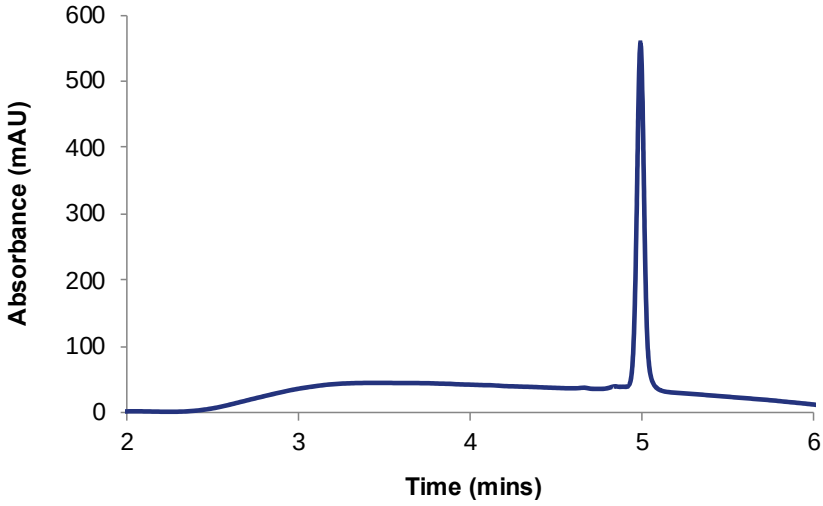


Figure 81 - Analytical HPLC trace of pure TEMPO modified H-UAF-OMe (**11**). Analytical gradient 10-70% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 477.3104. Observed Mass $[M+H]^+$: 477.3077.

H-CAF-OMe conjugation with TEMPO (11b)



Synthesised following general protocol B using H-CAF-OMe (**48**) (4 mg, 11.3 µmol) followed by purification using preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (3.8 mg, 7.97 µmol, 71% yield). ^{1}H NMR (400 MHz, CD_3OD) δ 7.33 – 7.20 (m, 5H), 4.64 (m, 1H), 4.48 (q, J = 7.0 Hz, 1H), 4.14 – 4.01 (m, 3H), 3.70 (s, 3H), 3.19 – 3.11 (dd, J = 14.0, 8.8 Hz, 1H), 2.98 (dd, J = 14.0, 8.8 Hz, 1H), 1.63 (m, 2H), 1.51 – 1.42 (m, 4H), 1.38 (d, J = 7.0 Hz, 3H), 1.19 – 1.00 (m, 12H).^{13}C NMR (101 MHz, CD_3OD) δ 172.6, 171.9, 165.6, 136.7, 128.8, 128.2, 126.6, 73.6, 60.2, 60.1, 54.1, 52.3, 51.3, 48.8, 39.4, 36.8, 32.0, 19.2, 17.6, 16.5.



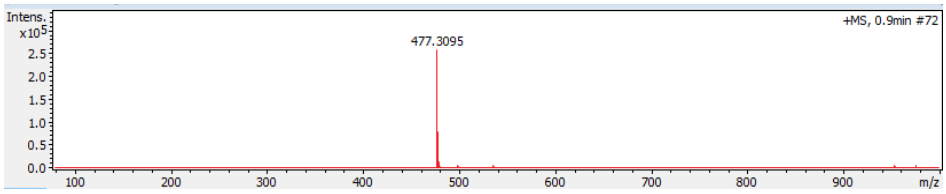
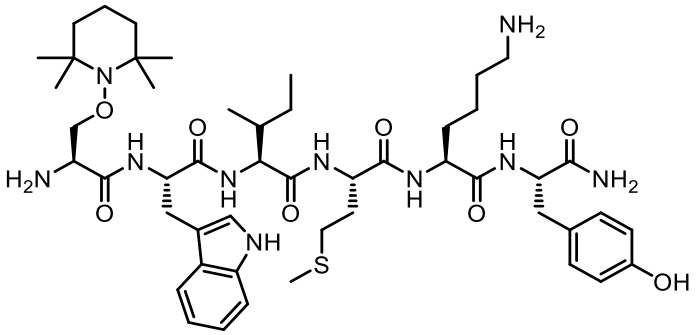


Figure 82 - Analytical HPLC trace of pure conjugated H-CAF-OMe using TEMPO (**11b**). Analytical gradient 10-70% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 477.3077. Observed Mass $[M+H]^+$: 477.3095.

H-A(TEMPO)WIMKY-NH_2 (17) – Met stability to conditions



Synthesised using general protocol A using H-UWIMKY-NH_2 (**16**) (4 mg, 0.0044 mmol). Purified by preparative RP-HPLC (15 to 80% B over 30 minutes) and lyophilized to produce the desired modified peptide (2.8 mg, 0.003 mmol, 68%).

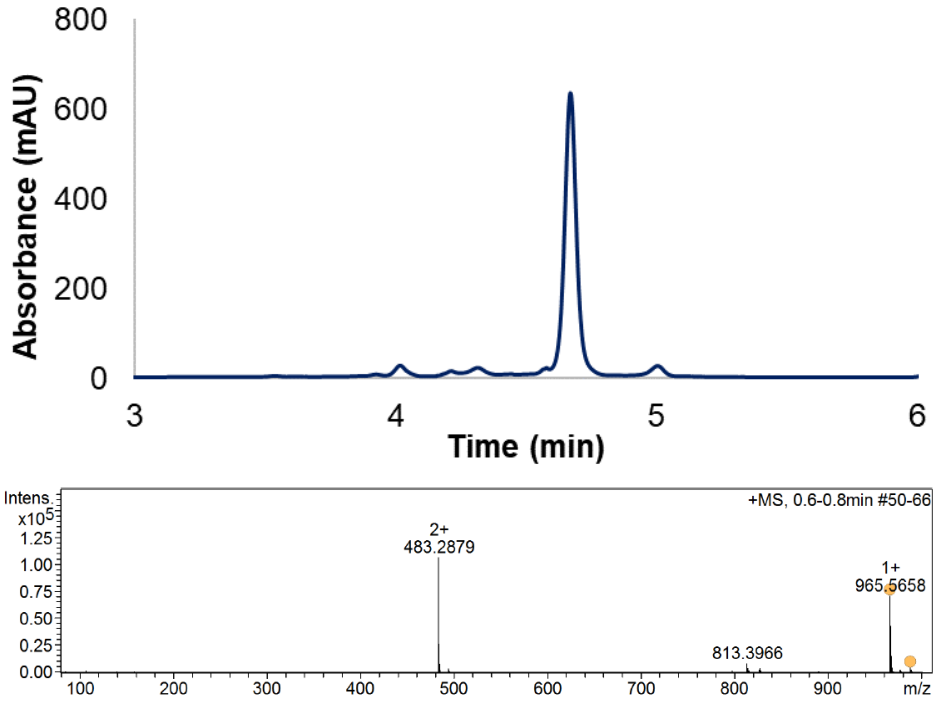
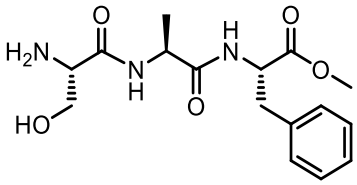


Figure 83 - Analytical HPLC trace of pure TEMPO-WIMKY-NH_2 (**17**). Analytical gradient 10-75% B over 10 min, 210 nm. HRMS Calculated Mas: $[M+H]^+$ = 965.5641, $[M+2H]^{2+}$: 483.2860. Observed Mass: $[M+H]^+$ = 965.5658, $[M+2H]^{2+}$: 483.2879.

TEMPO conjugate reductive cleavage to Ser (18)



TEMPO conjugate H-(A(TEMPO)AF-OMe (**11**) (5 mg, 0.0105 mmol) was dissolved in H_2O: MeOH: AcOH 8:1:1 (4.2 ml) to a concentration of 2.5 mM. Zinc dust (0.034 mg, 0.525 mmol) was added. The solution was stirred at 37 °C for 16 hours then purified by semi-preparative RP-HPLC (5-60% B over 30 minutes) and lyophilized to produce the desired peptide (3.4 mg, 0.010 mmol, 91% yield).^{1}H NMR (400 MHz, CD_3OD) δ 8.34 (d, *J* = 7.8 Hz, 1H), 7.34 – 7.16 (m, 5H), 4.71 – 4.59 (m, 1H), 4.41 (q, *J* = 7.2 Hz, 1H), 3.98 – 3.79 (m, 3H), 3.68 (s, 3H), 3.14 (dd, 1H), 3.00 (dd, *J* = 13.9, 8.3 Hz, 1H), 1.34 (d, *J* = 7.2 Hz, 3H).^{13}C NMR (101 MHz, CD_3OD) δ 173.1, 171.8, 166.6, 136.6, 128.9, 128.1, 126.5, 60.4, 54.7, 54.0, 51.3, 49.0, 36.9, 16.7.

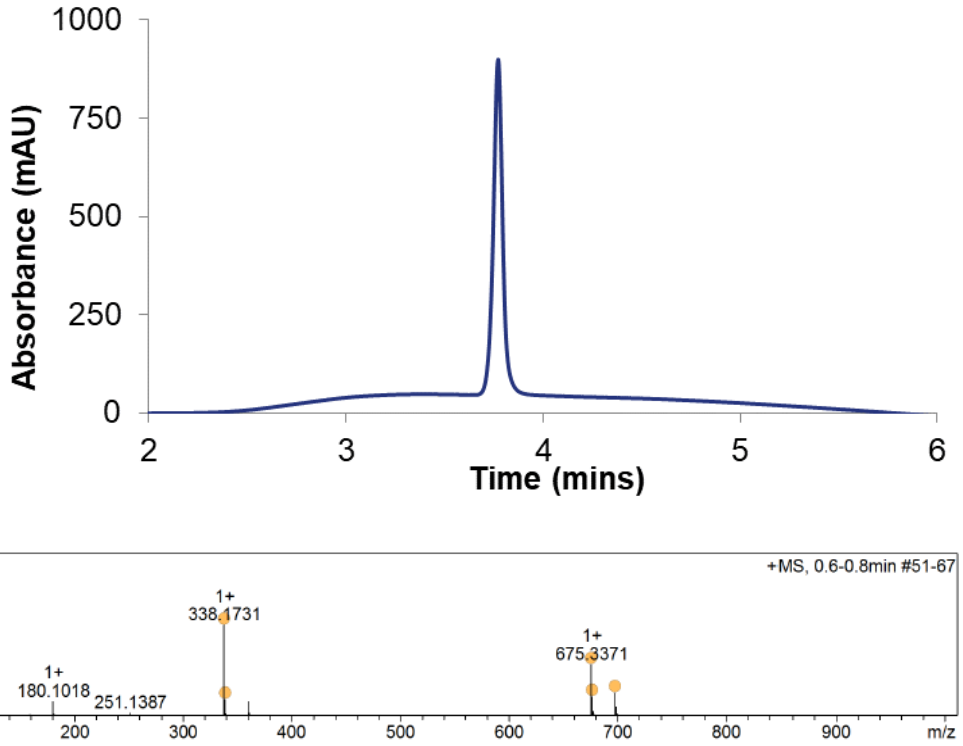
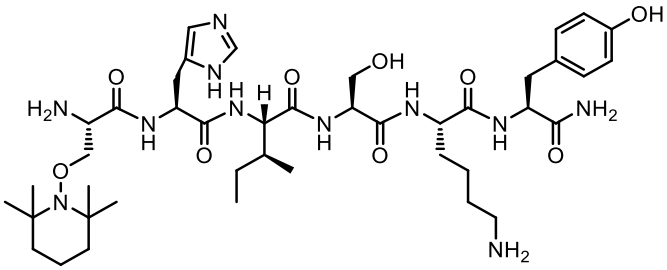


Figure 84 - Analytical HPLC trace purified reductively cleaved product (**18**). Analytical gradient 10-70% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 338.1716. Observed Mass $[M+H]^+$: 338.1731.

H-A(TEMPO)HISKY-NH_2 (19)



Synthesised using general protocol A using H-UHISKY-NH_2 (**33**) (5 mg, 6.28 µmol). Purified by preparative RP-HPLC (10 to 80% B over 30 minutes) and lyophilized to produce the desired modified peptide (4.6 mg, 5.28 µmol, 84% yield).

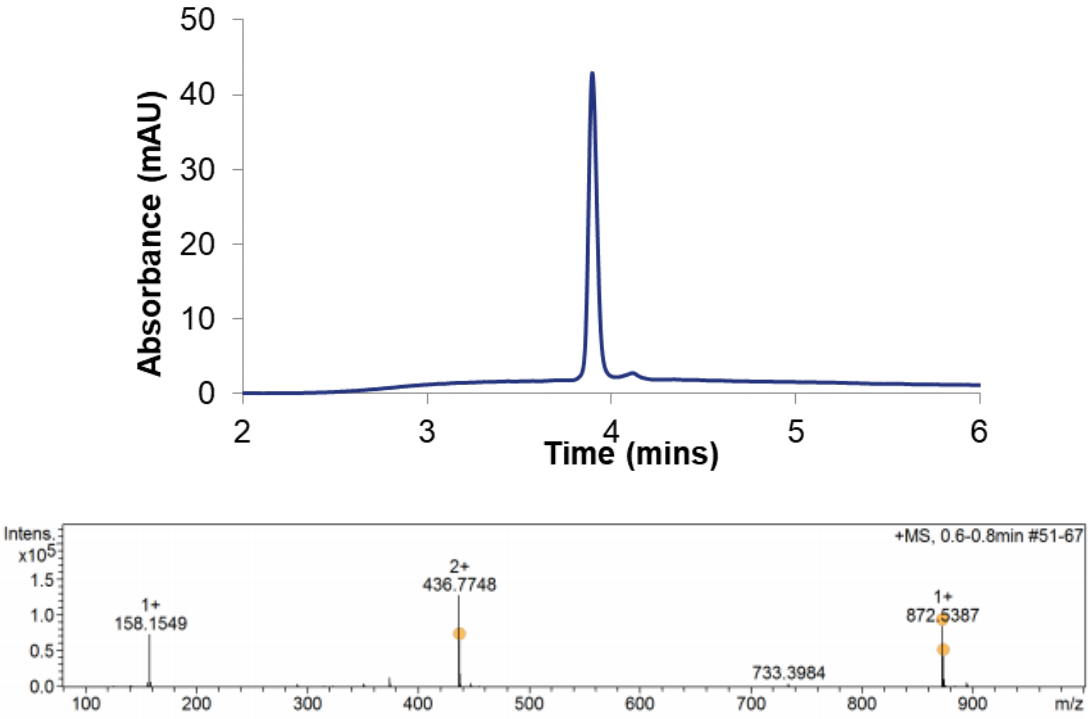
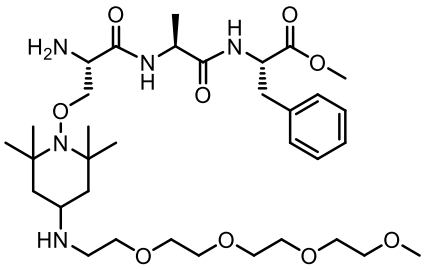


Figure 85 - Analytical HPLC trace of pure TEMPO modified H-UHISKY-NH_2 (**17**). Analytical gradient 1070% B over 5 min, 280 nm. HRMS Calculated Mass $[M+H]^+$: 872.5358, $[M+2H]^{2+}$: 436.7718. Observed Mass $[M+H]^+$: 872.5387, $[M+2H]^{2+}$: 436.7748.

Tetra-EG-TEMPO Conjugate (27)



Synthesised using general protocol A using H-UAF-OMe (**10**, 5 mg, 0.0125 mmol). Purified by preparative RP-HPLC (15 to 80% B over 30 minutes) and lyophilised to produce the desired modified peptide (6.1 mg, 8.95 mmol, 72% yield). ^{1}H NMR (400 MHz, CD_3OD) δ 7.35 – 7.19 (m, 5H), 4.63 (dd, *J* = 8.7, 5.4 Hz, 1H), 4.50 (q, *J* = 7.0 Hz, 1H), 4.17 – 4.05 (m, 3H), 3.77 (m, 2H), 3.70 (m, 7H), 3.68 – 3.62 (m, 6H), 3.58 – 3.54 (m, 2H), 3.53 – 3.44 (m, 1H), 3.37 (s, 3H), 3.27 (m, 2H), 3.17 (dd, *J* = 14.0, 5.4 Hz, 1H), 2.98 (dd, *J* = 14.0, 4.0 Hz, 1H), 1.96 (m, 2H), 1.56 (q, *J* = 12.0 Hz, 2H), 1.39 (d, *J* = 7.0 Hz, 3H), 1.30 – 1.08 (m, 12H). ^{13}C NMR (100 MHz, CD_3OD) δ 172.7, 171.8, 165.5, 136.8, 128.9, 128.2, 126.6, 73.9, 71.5, 70.0, 65.7, 59.8, 57.7, 54.1, 52.2, 51.3, 49.8, 44.6, 41.1, 36.8, 31.8, 19.5, 17.6.

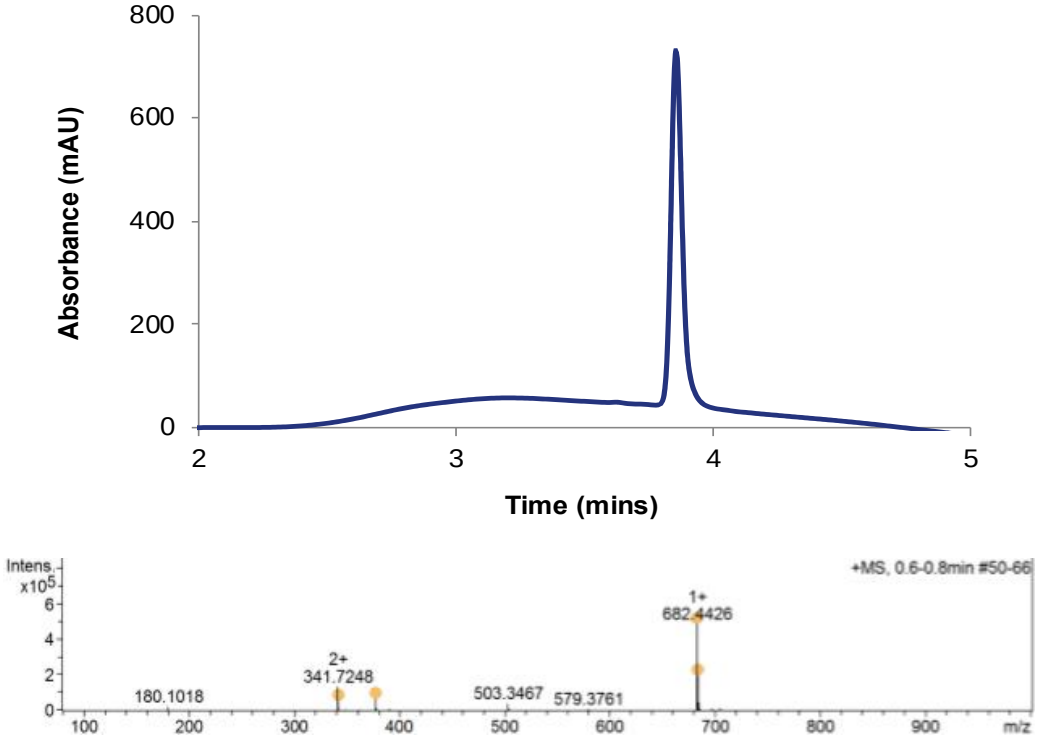
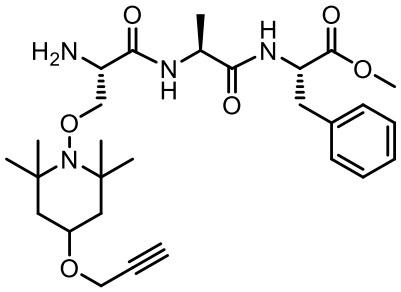
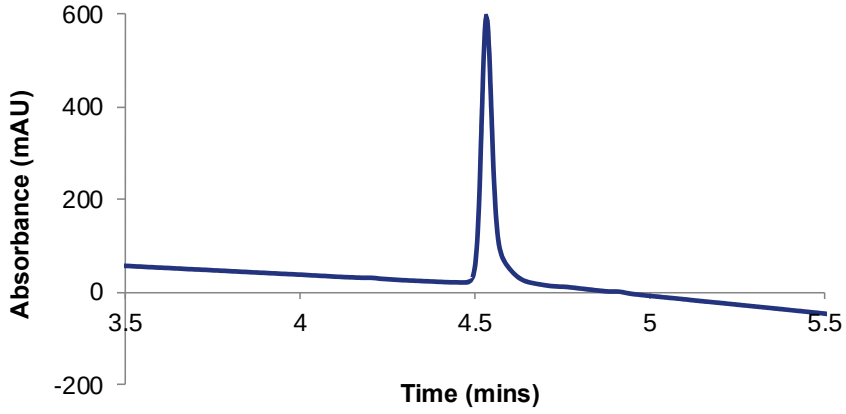


Figure 86 - Analytical HPLC trace of pure tetra-EG-TEMPO modified H-UAF-OMe (**27**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 682.4391. Observed Mass $[M+H]^+$: 682.4426.

Propargyl TEMPO Conjugate (28)



Synthesised using general protocol A using H-UAF-OMe (**10**) (5 mg, 0.0125 mmol). Purified by preparative RP-HPLC (15 to 80% B over 30 minutes) and lyophilised to produce the desired modified peptide (5.4 mg, 0.01 mmol, 80% yield). ^{1}H NMR (400 MHz, CD_3OD) δ 7.35 – 7.22 (m, 5H, Ar-H), 4.64 (dd, *J* = 9.0, 5.2 Hz, 1H, CHα), 4.49 (q, *J* = 7.0 Hz, 1H, CHα), 4.20 (d, *J* = 2.4 Hz, 2H, OCH_2), 4.14 – 4.03 (m, 3H, CHα and $CH_{2\beta}$), 3.86 (m, 1H, TEMPO CH), 3.72 (s, 3H, OCH_3), 3.18 (dd, *J* = 14.0, 5.2 Hz, 1H, CH_β), 2.97 (dd, *J* = 14.0, 9.0 Hz, 1H, CH_β), 2.88 (t, *J* = 2.4 Hz, 1H, alkyne CH), 1.87 (m, 2H, CH_2), 1.39 (d, *J* = 7.0 Hz, 3H, Ala CH_3), 1.37 – 1.29 (m, 2H, CH_2), 1.13 (m, 12H, TEMPO CH_3). ^{13}C NMR (101 MHz, CD_3OD) δ 172.6, 171.9, 165.6, 136.8, 128.9, 128.2, 126.6, 79.6, 74.2, 73.6, 69.0, 60.3, 54.7, 54.1, 52.2, 51.3, 48.8, 44.4, 36.8, 32.1, 20.0, 17.7.



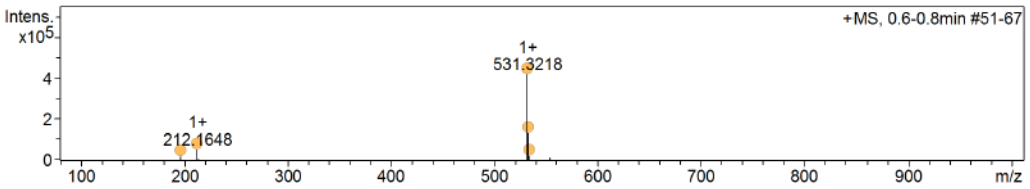
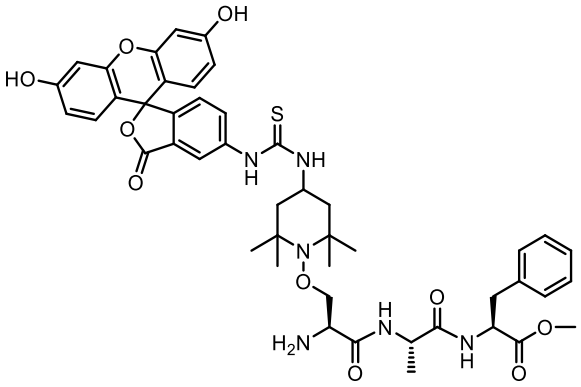


Figure 87 - Analytical HPLC trace of pure propargyl-TEMPO modified H-UAF-OMe (**28**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 531.3183. Observed Mass $[M+H]^+$: 531.3218.

FITC-TEMPO Conjugate (29)



Synthesised using general protocol A using H-UAF-OMe (**10**) (5 mg, 0.0125 mmol). Purified by preparative RP-HPLC (15 to 80% B over 30 minutes) and lyophilized to produce the desired modified peptide (7.6 mg, 69% yield, 8.62 mmol). ^{1}H NMR (400 MHz, CD_3OD) δ 8.27 (d, J = 1.6 Hz, 1H), 7.84 (d, J = 4.0 Hz, 1H), 7.38 – 7.17 (m, 6H), 6.88 (m, 4H), 6.70 (m, 2H), 4.65 (m, 1H), 4.51 (q, J = 12.0, 4.0 Hz, 1H), 4.09 (d, J = 5.4 Hz, 3H), 3.72 (s, 3H), 3.19 (dd, J = 14.0, 5.2 Hz, 1H), 2.96 (dd, J = 14.0, 5.1 Hz 1H), 2.68 (s, 1H), 2.01 – 1.90 (m, 2H), 1.66 – 1.44 (m, 2H), 1.40 (d, J = 7.0 Hz, 3H), 1.26 – 1.03 (m, 12H). ^{13}C NMR (101 MHz, CD_3OD) δ 180.7, 172.7, 171.9, 165.6, 154.1, 141.5, 136.8, 129.7, 128.8, 128.3, 126.8, 125.3, 113.6, 102.1, 73.7, 60.5, 54.2, 52.3, 51.4, 48.8, 45.6, 44.3, 36.8, 31.8, 19.8, 17.7, 15.9.

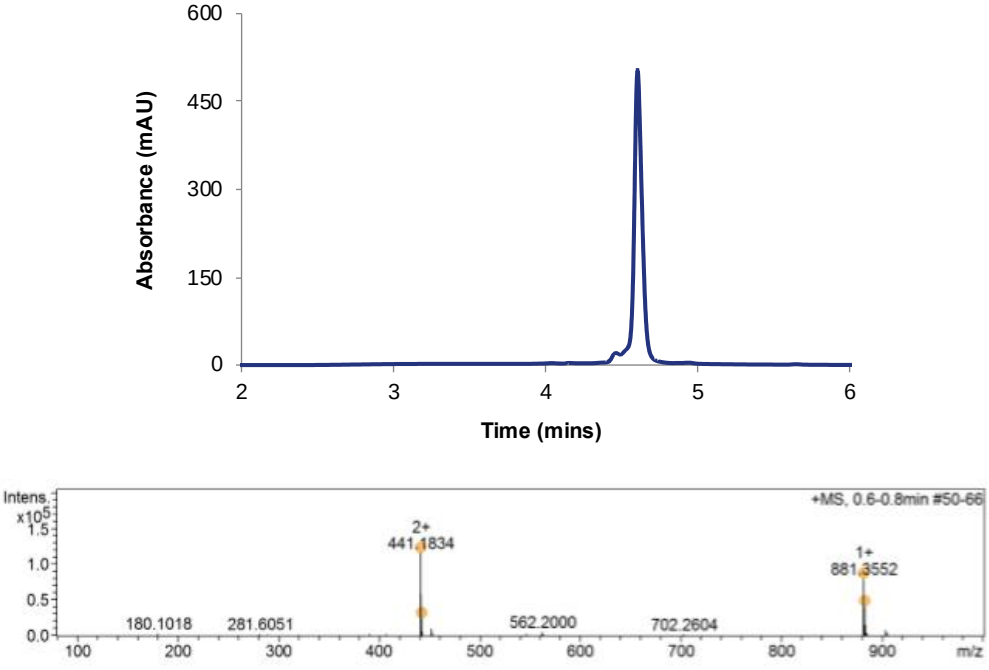
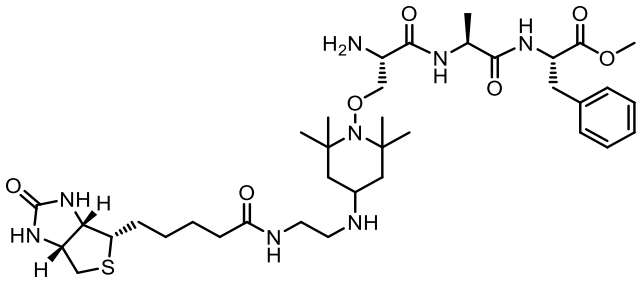


Figure 88 - Analytical HPLC trace of pure FITC-TEMPO modified H-UAF-OMe (**29**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 881.3544, $[M+2H]^{2+}$: 441.1811. Observed Mass $[M+H]^+$: 881.3552, $[M+2H]^{2+}$: 441.1834.

Biotin-EDA-TEMPO Conjugate (30)



Synthesised using general protocol A using H-UAF-OMe (**10**) (5 mg, 0.0125 mmol). Purified by preparative HPLC (15 to 80% B over 30 minutes) and lyophilized to produce the desired modified peptide (7.4 mg, 0.0097 mmol, 78% yield).^{1}H NMR (400 MHz, CD_3OD) δ 8.53 (d, *J* = 7.7 Hz, 1H), 7.39 – 7.18 (m, 5H), 4.63 (m, 1H), 4.52 (m, 2H), 4.32 (m, 1H), 4.19 – 4.02 (m, 3H), 3.71 (s, 3H), 3.55 – 3.44 (m, 3H), 3.28 – 3.13 (m, 4H), 3.03 – 2.90 (m, 2H), 2.78 – 2.65 (m, 1H), 2.30 (t, *J* = 7.3 Hz, 2H), 1.93 (m, 2H), 1.85 – 1.43 (m, 8H), 1.39 (d, *J* = 7.0 Hz, 3H), 1.32 – 1.05 (m, 12H). ^{13}C NMR (101 MHz, CD_3OD) δ 176.4, 172.7, 171.8, 165.5, 164.7, 136.8, 128.9, 128.2, 126.6, 73.9, 62.0, 60.3, 59.7, 55.6, 54.1, 52.2, 51.3, 49.6, 44.7, 41.1, 39.6, 36.8, 36.1, 35.0, 31.8, 28.4, 28.1, 25.1, 19.4, 17.6.

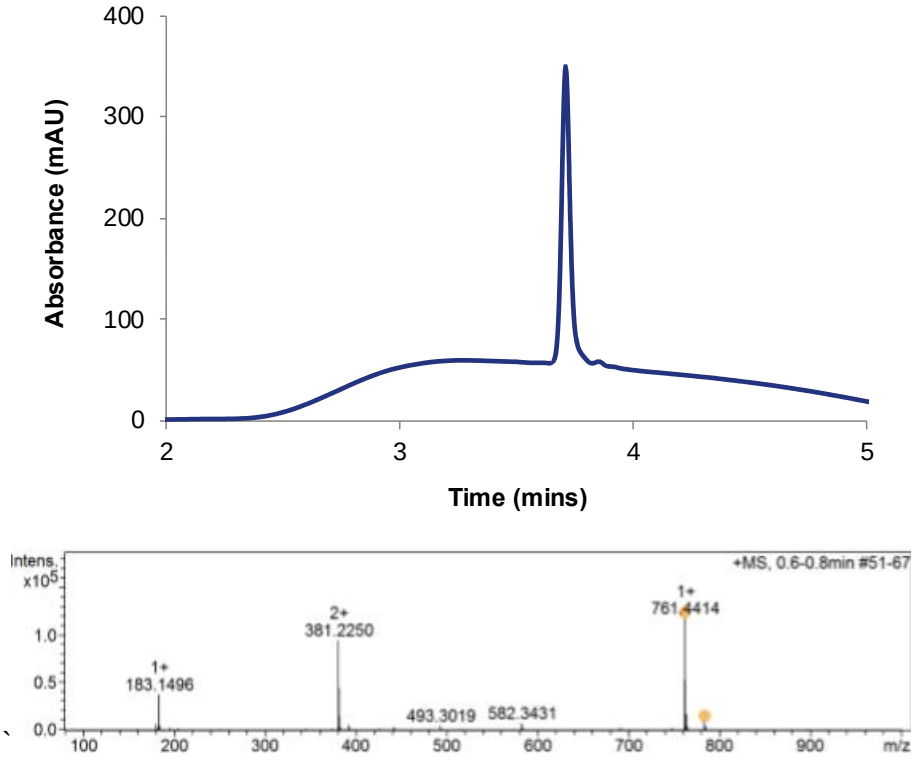
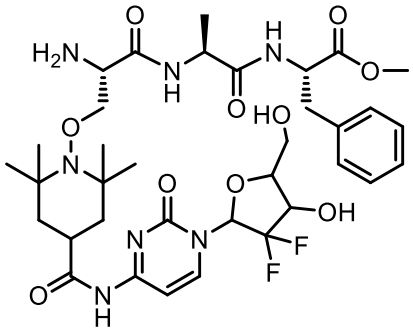


Figure 89 - Analytical HPLC trace of pure biotin-EDA-TEMPO modified H-UAF-OMe (**30**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 761.4384, $[M+2H]^{2+}$: 381.2231. Observed Mass $[M+H]^+$: 761.4414, $[M+2H]^{2+}$: 381.2250.

Gemcitabine-TEMPO Conjugate (31)



Synthesised using general protocol A using H-UAF-OMe (**10**) (5 mg, 0.0125 mmol). Purified by preparative RP-HPLC (15 to 80% B over 30 minutes) and lyophilized to produce the desired modified peptide (6.4 mg, 67% yield, 8.36 mmol). ^{1}H NMR (400 MHz, CD_3OD) δ 8.48 (d, J = 7.7 Hz, 0.5H), 8.37 (d, J = 7.6 Hz, 1H), 8.28 (d, J = 7.7 Hz, 0.5H), 7.50 (d, J = 7.6 Hz, 1H), 7.37 – 7.16 (m, 5H), 6.29-6.15 (m, 1H), 4.64 (m, 1H), 4.49 (m, 1H), 4.32 (m, 1H), 4.18 – 4.06 (m, 3H), 4.04 – 3.94 (m, 2H), 3.83 (m, 1H), 3.71 (s, 3H), 3.18 (ddd, J = 13.9, 5.4, 2.9 Hz, 1H), 3.06 – 2.94 (m, 1H), 1.78 – 1.59 (m, 3H), 1.57 – 1.46 (m, 1H), 1.39 (d, J = 7.0, 1.0 Hz, 3H), 1.28 – 1.05 (m, 12H). ^{13}C NMR (101 MHz, CD_3OD) δ 185.1, 183.2, 171.9, 165.6, 162.8, 154.4, 150.5, 136.7, 128.2, 126.6, 118.1, 73.73, 71.2, 59.7, 59.6, 58.8, 54.1, 52.2, 51.3, 48.8, 44.4, 41.7, 36.8, 31.8, 19.6, 17.6, 16.1. ^{19}F NMR (376 MHz, CD_3OD) δ -119.22 (d, J = 248 Hz, 1F), -120.14 (dd, J = 256 Hz, 1F).

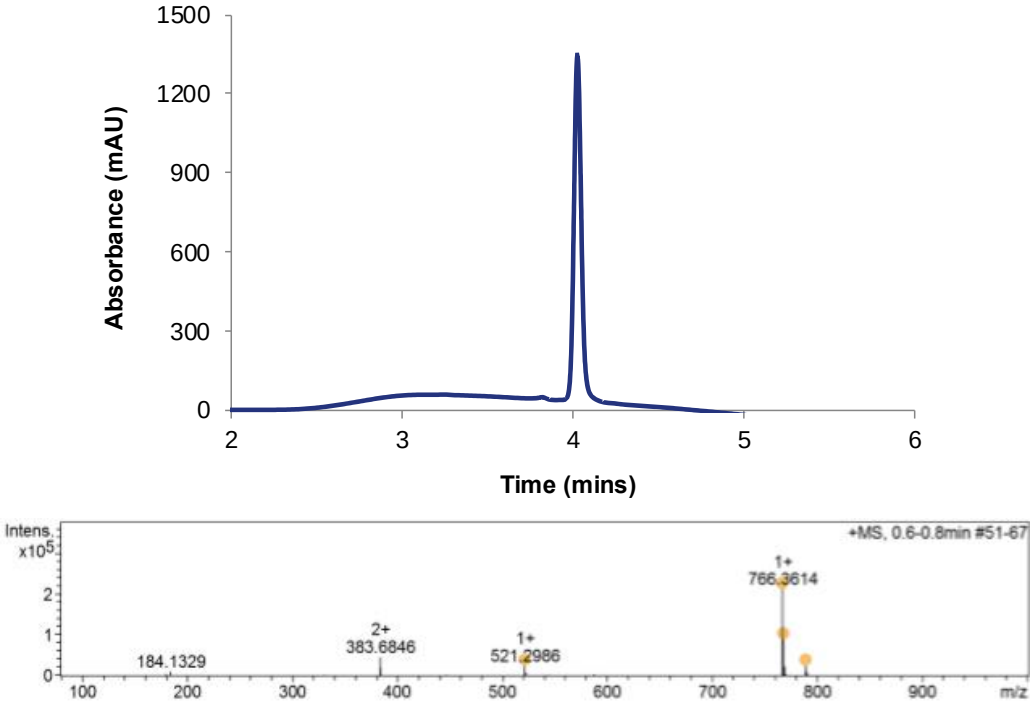
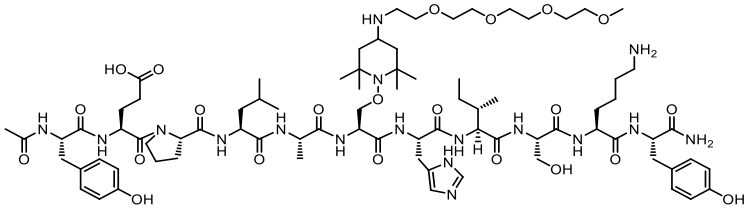


Figure 90 - Analytical HPLC trace of pure gemcitabine-TEMPO modified H-UAF-OMe (**31**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 766.3587. Observed Mass $[M+H]^+$: 766.3614.

Ac-YEPLA-SePh (32) and H-UHISKY-NH_2 (33) One-pot Ligation and Conjugation (34-38)

Tetra-EG-TEMPO Conjugate (34)



Synthesised following general one-pot ligation-conjugation protocol using H-UHISKY-NH_2 (**33**) (2 mg, 2.51 µmol) and Ac-YEPLA-SePh (**32**) (2.04 mg, 2.64 µmol). Purified by semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilised to yield the desired modified peptide (3.3 mg, 1.95 µmol, 78% yield).

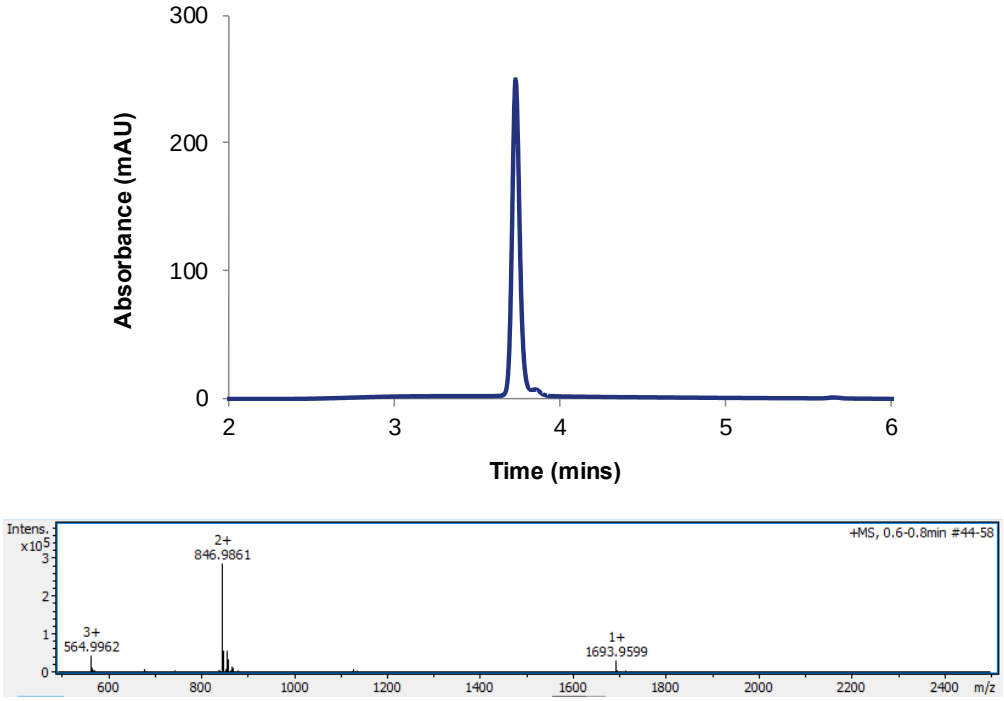
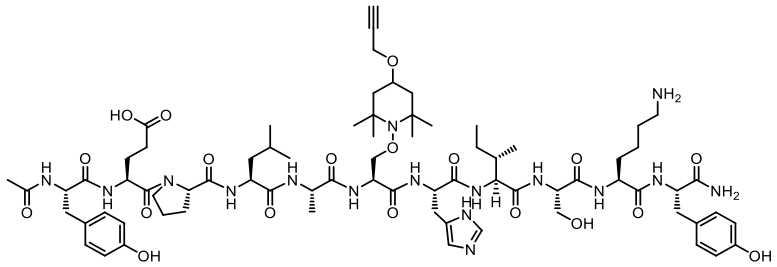


Figure 91 - Analytical HPLC trace of pure tetra-EG-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**34**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 1692.9654, $[M+2H]^{2+}$: 846.9827. Observed Mass $[M+H]^+$: 1693.9599, $[M+2H]^{2+}$: 846.9861.

Propargyl-TEMPO Conjugate (35)



Synthesised following general one-pot ligation-conjugation protocol using H-UHISKY-NH_2 (**33**) (2 mg, 2.51 µmol) and Ac-YEPLA-SePh (**32**) (2.04 mg, 2.64 µmol). Purified by semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilised to yield the desired modified peptide (3.1 mg, 2.01 µmol, 80% yield).

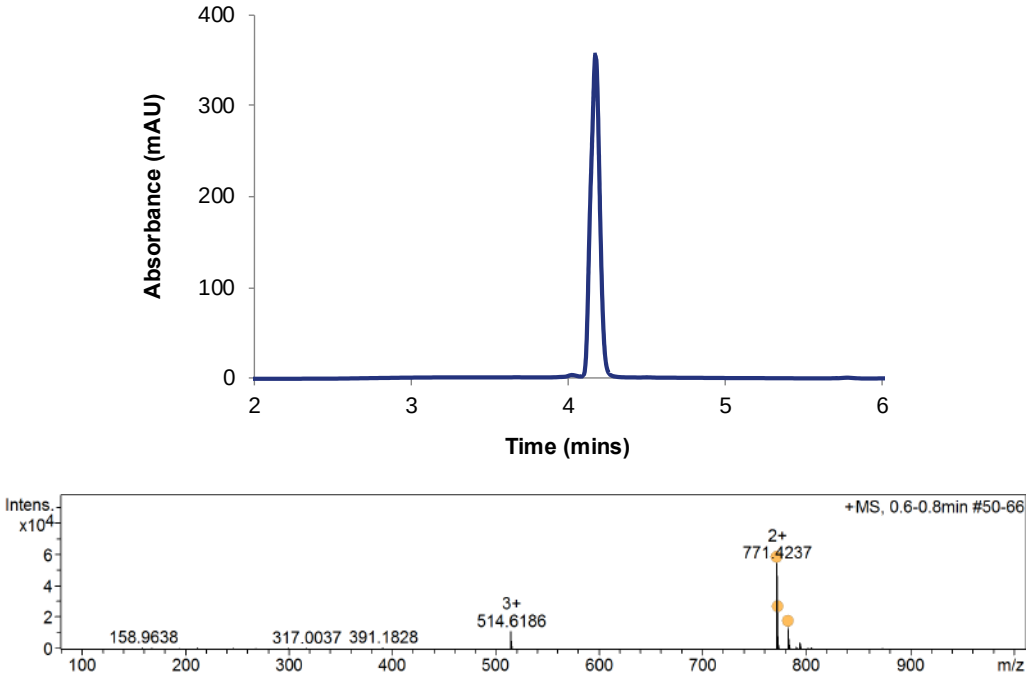
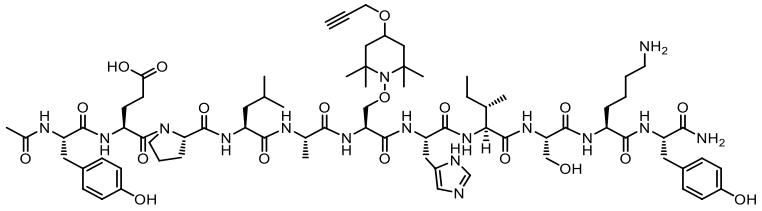


Figure 92 - Analytical HPLC trace of pure propargyl-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**35**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 771.4223. Observed Mass $[M+2H]^{2+}$: 771.4237.

Propargyl-TEMPO Conjugate from Ac-YEPLACHISKY-NH_2 (35b)



Synthesised following general protocol D using Ac-YEPLACHISKY-NH_2 (**56**) (4.0 mg, 2.93 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (3.0 mg, 1.95 µmol, 67% yield).

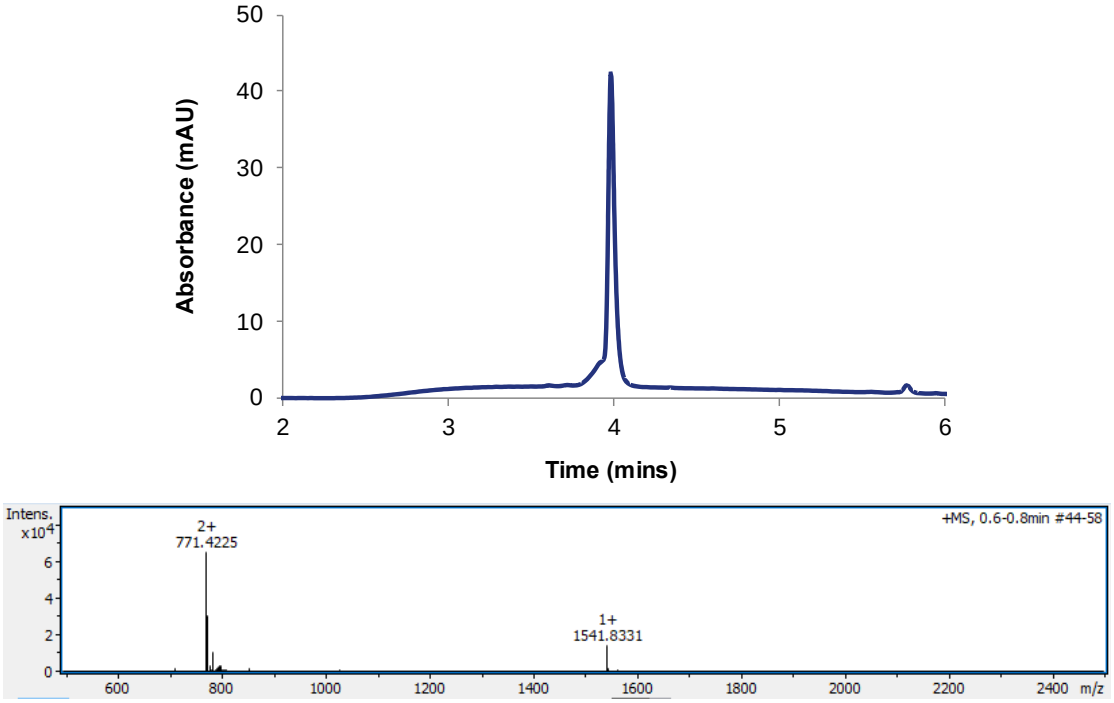
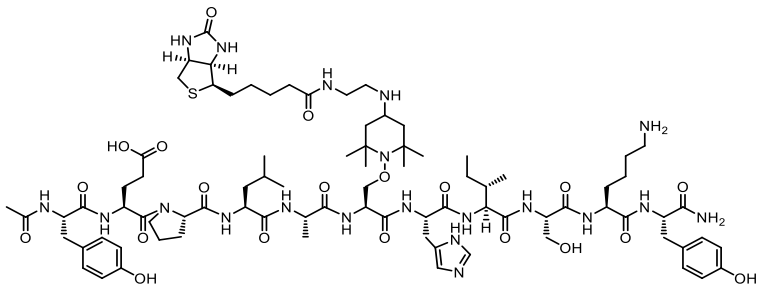


Figure 93 - Analytical HPLC trace of pure conjugated Ac-YEPLACHISKY-NH_2 using propargyl-TEMPO (**35b**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^+$: 1541.8367, $[M+2H]^{2+}$: 771.4223. Observed Mass $[M+H]^+$: 1541.8331, $[M+2H]^{2+}$: 771.4225.

Biotin-TEMPO Conjugate (36)



Synthesised following general one-pot ligation-conjugation protocol using H-UHISKY-NH_2 (**33**) (2 mg, 2.51 µmol) and Ac-YEPLA-SePh (**32)** (2.04 mg, 2.64 µmol). Purified by semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilised to yield the desired modified peptide (3.4 mg, 1.92 µmol, 76% yield).

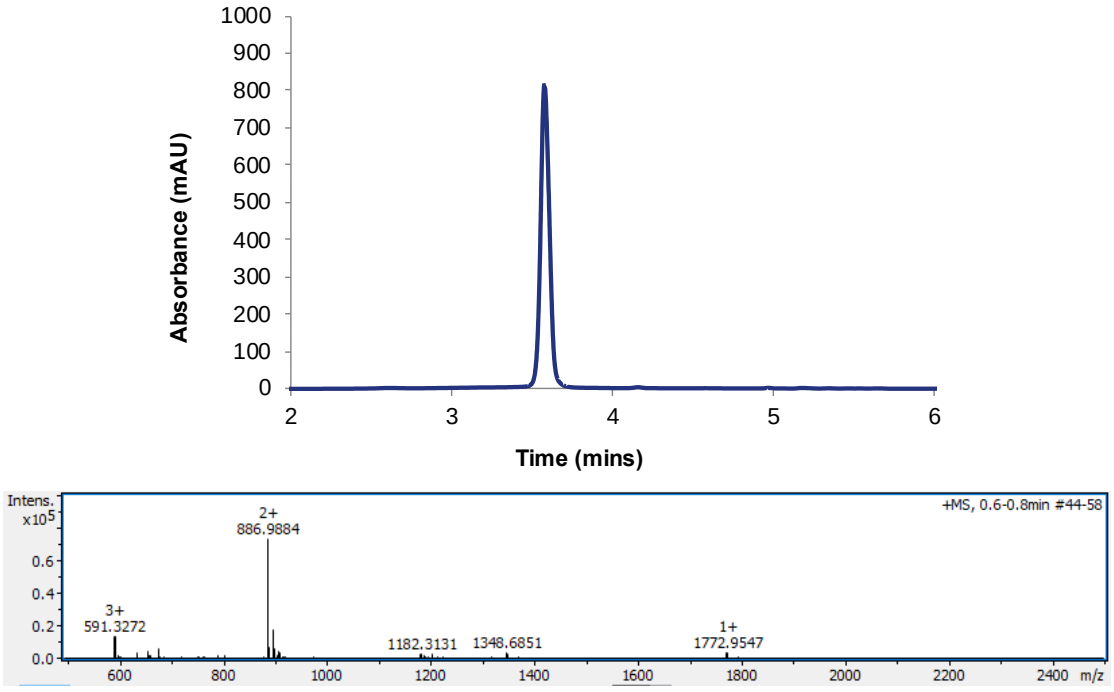
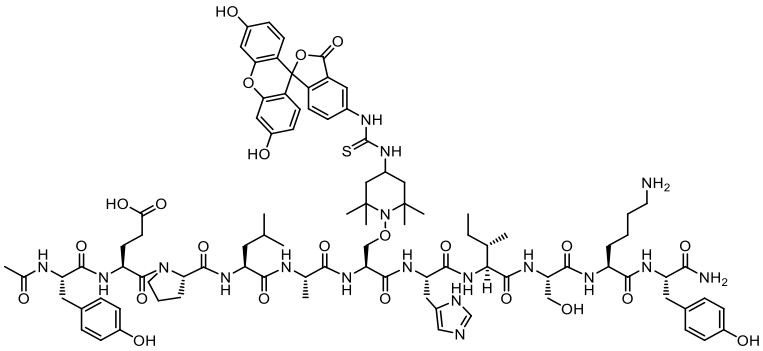


Figure 94 - Analytical HPLC trace of pure biotin-EDA-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**36**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 886.4824. Observed Mass $[M+2H]^{2+}$: 886.9884.

FITC-TEMPO Conjugate (37)



Synthesised following general one-pot ligation-conjugation protocol using H-UHISKY-NH_2 (**33**) (2 mg, 2.51 µmol) and Ac-YEPLA-SePh (**32**) (2.04 mg, 2.64 µmol). Purified by semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilised to yield the desired modified peptide (3.3 mg, 1.74 µmol, 69% yield).

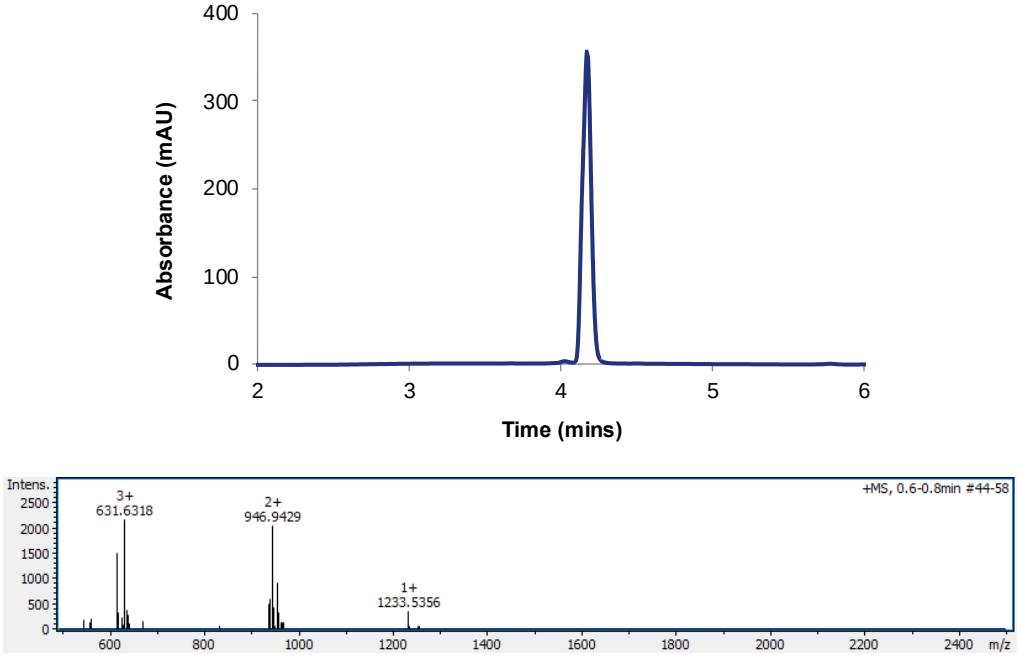
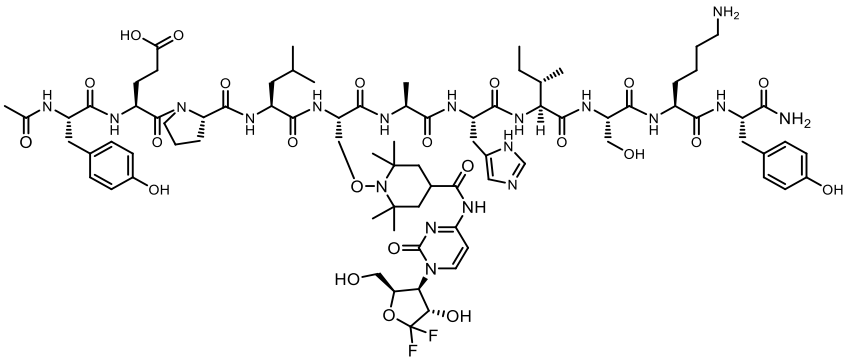


Figure 95 - Analytical HPLC trace of pure FITC-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**37**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+2H]^{2+}$: 946.4404, $[M+3H]^{3+}$: 631.2962. Observed Mass $[M+2H]^{2+}$: 946.9429, $[M+3H]^{3+}$: 631.6318.

Gemcitabine-TEMPO Conjugate (38)



Synthesised following general one-pot ligation-conjugation protocol using H-UHISKY-NH_2 (**33**) (2 mg, 2.51 µmol) and Ac-YEPLA-SePh (**32**) (2.04 mg, 2.64 µmol). Purified by semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilised to yield the desired modified peptide (3.5 mg, 1.97 µmol, 78% yield).

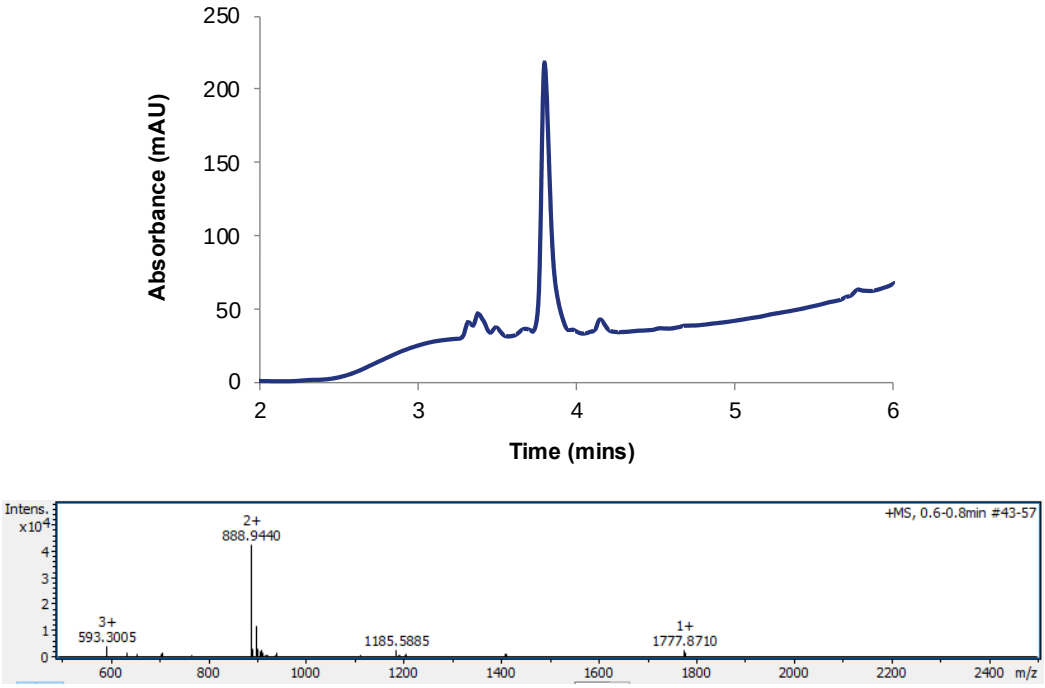
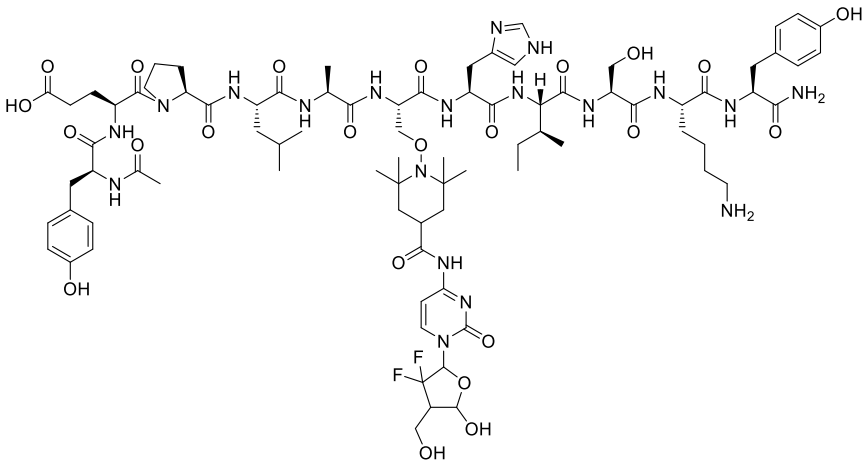


Figure 96 - Analytical HPLC trace of pure gemcitabine-TEMPO modified Ac-YEPLAUHISKY-NH_2 from one-pot ligation-conjugation (**38**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 888.9425. Observed Mass $[M+2H]^{2+}$: 888.9440.

Gemcitabine-TEMPO Conjugate with Ac-YEPLACHISKY-NH_2 (38b)



Synthesised following general protocol D using Ac-YEPLACHISKY-NH_2 (**56**) (4.0 mg, 2.93 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (3.2 mg, 1.81 µmol, 62% yield).

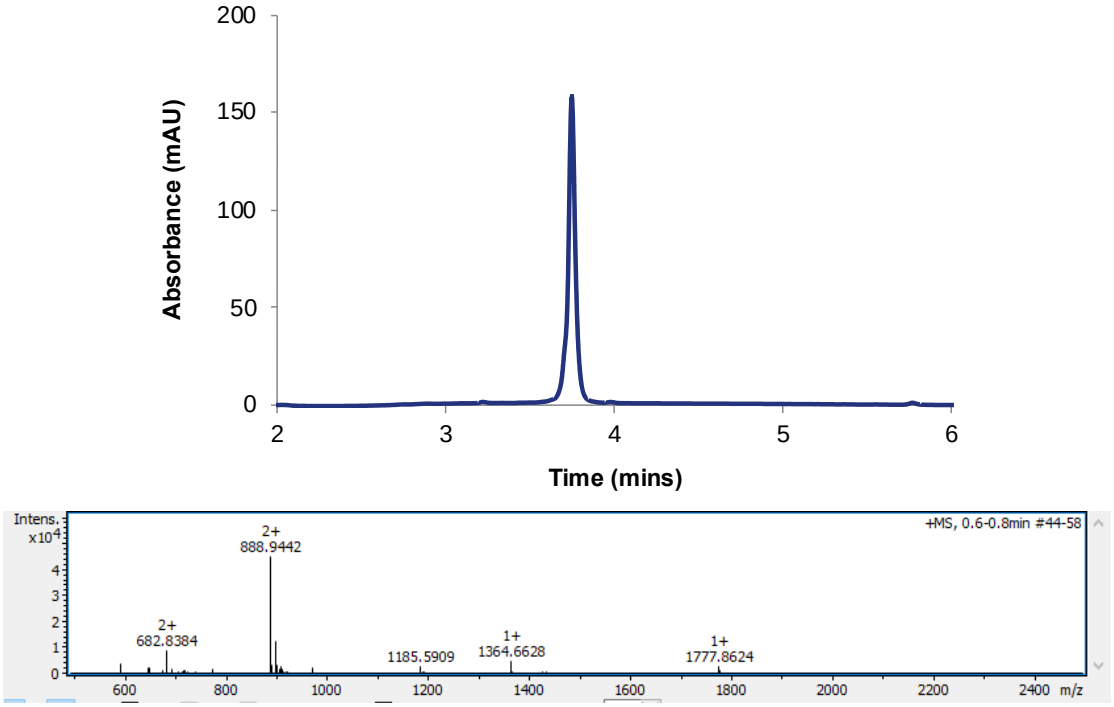
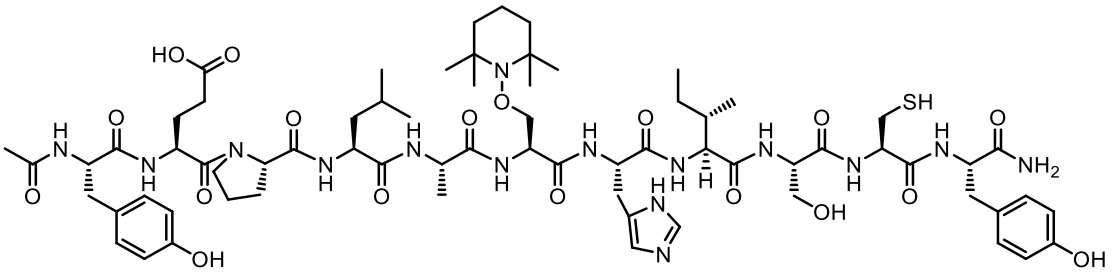


Figure 97 - Analytical HPLC trace of pure conjugated of Ac-YEPLACHISKY-NH_2 using gemcitabine-TEMPO (**38b**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+2H]^{2+}$: 888.9426. Observed Mass $[M+2H]^{2+}$: 888.9442.

Ac-YEPLA-SePh (32) and H-UHISCY-NH_2 (39) One-pot Ligation and Conjugation (40 – 42)

TEMPO Conjugate (40)



Synthesised following general one-pot ligation-conjugation protocol using H-UHISCY-NH_2 (**39**) (2.05 mg, 2.66 µmol) and Ac-YEPLA-SePh (**32**) (2.16 mg, 2.79 µmol). Purified by semi-preparative RP-HPLC (10-80% B over 30 minutes) and lyophilised to yield the desired modified peptide (3.5 mg, 2.39 µmol, 90% yield).

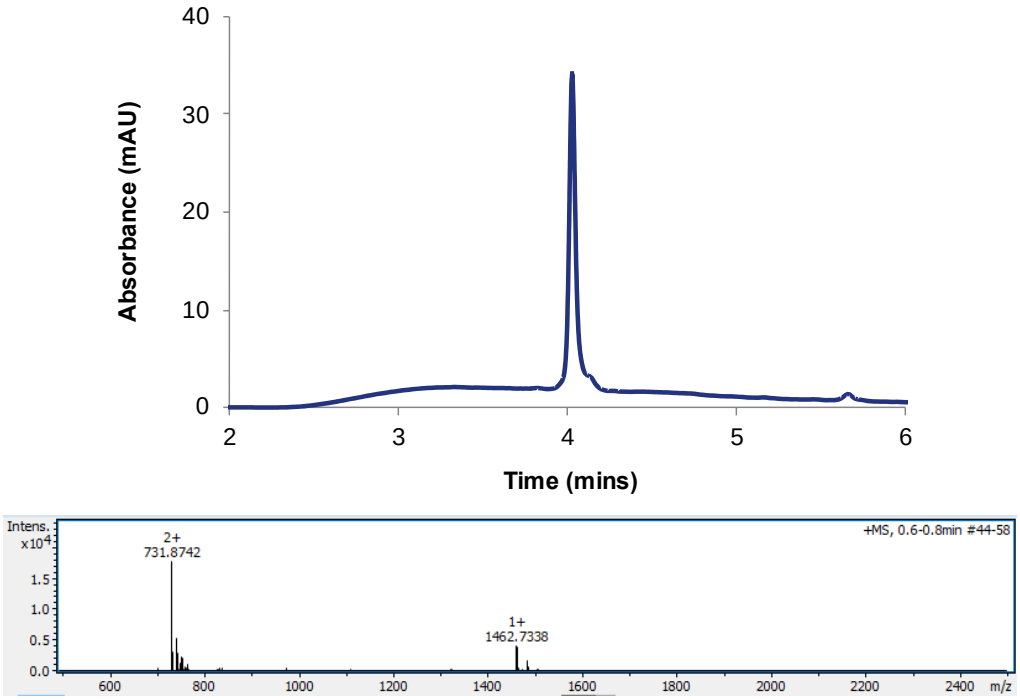
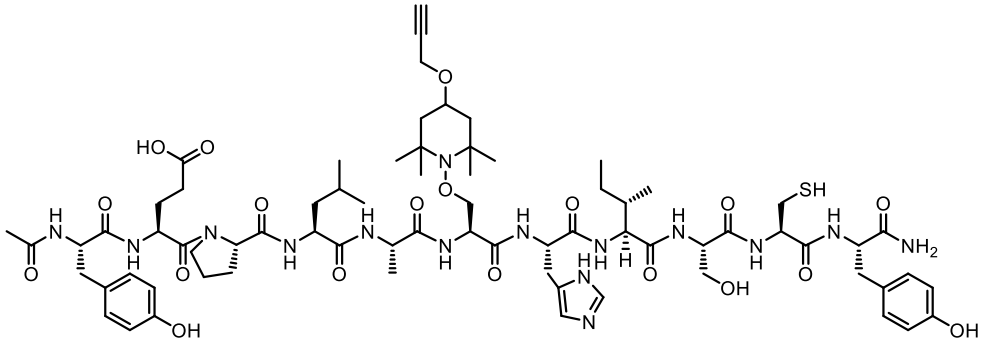


Figure 98 - Analytical HPLC trace of pure TEMPO modified Ac-YEPLAUHISCY-NH_2 from one-pot ligation-conjugation (**40**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 1462.7405, $[M+2H]^{2+}$: 731.8740. Observed Mass $[M+H]^+$: 1462.7338, $[M+2H]^{2+}$: 731.8742.

Propargyl-TEMPO Conjugate (41)



Synthesised following general one-pot ligation-conjugation protocol using H-UHISCY-NH_2 (**39**) (2.05 mg, 2.66 µmol) and Ac-YEPLA-SePh (**32**) (2.16 mg, 2.79 µmol). Purified by semi-preparative RP-HPLC (10-80% B over 30 minutes) and lyophilised to yield the desired modified peptide (3.3 mg, 2.20 µmol, 83% yield).

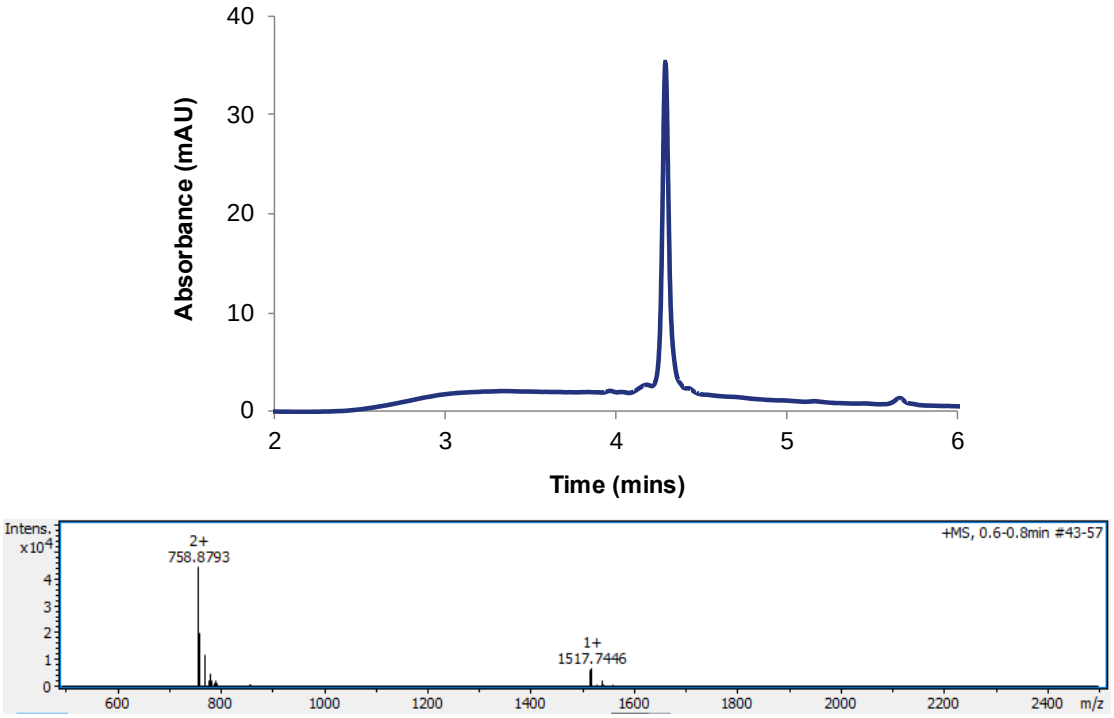
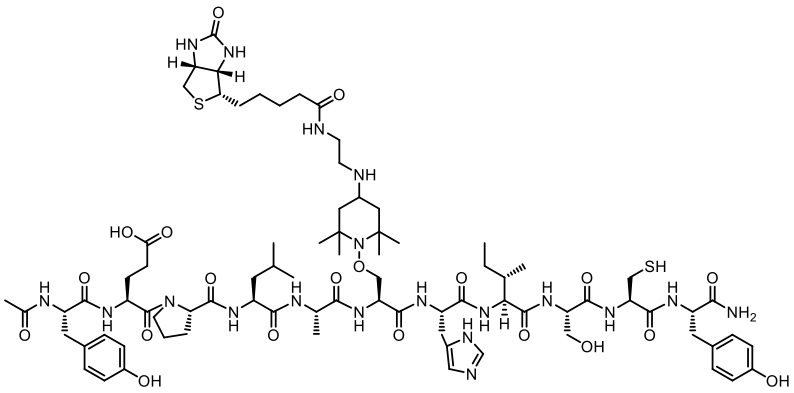


Figure 99 - Analytical HPLC trace of pure propargyl-TEMPO modified Ac-YEPLAUHISCY-NH_2 from one-pot ligation-conjugation (**41**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 1517.7588, $[M+2H]^{2+}$: 759.3833. Observed Mass $[M+H]^+$: 1517.7446, $[M+2H]^{2+}$: 758.8793.

Biotin-EDA-TEMPO (42)



Synthesised following general one-pot ligation-conjugation protocol using H-UHISCY-NH_2 (**39**) (2.05 mg, 2.66 µmol) and Ac-YEPLA-SePh (**32**) (2.16 mg, 2.79 µmol). Purified by semi-preparative RP-HPLC (2-70% B over 30 minutes) and lyophilised to yield the desired modified peptide (3.3 mg, 1.89 µmol, 71% yield).

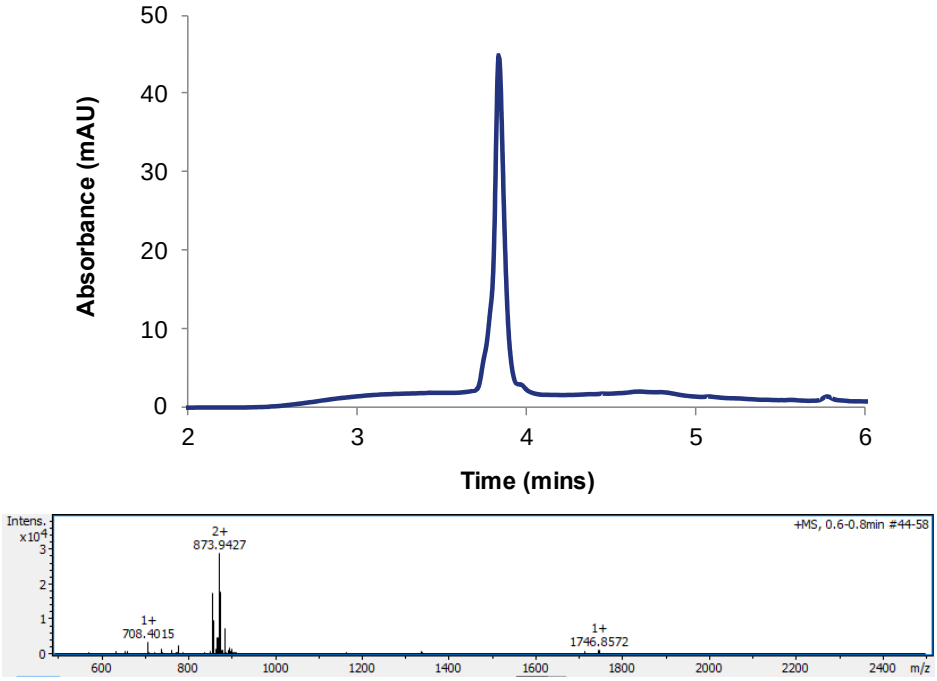
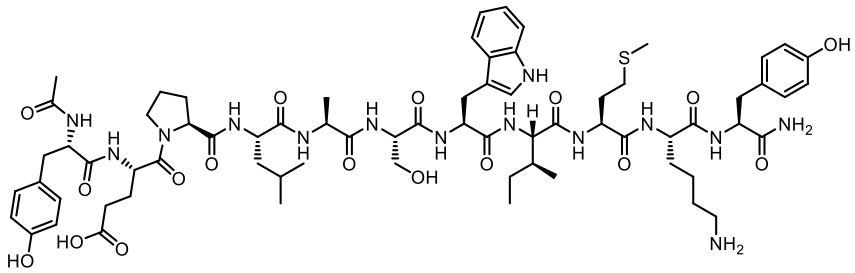
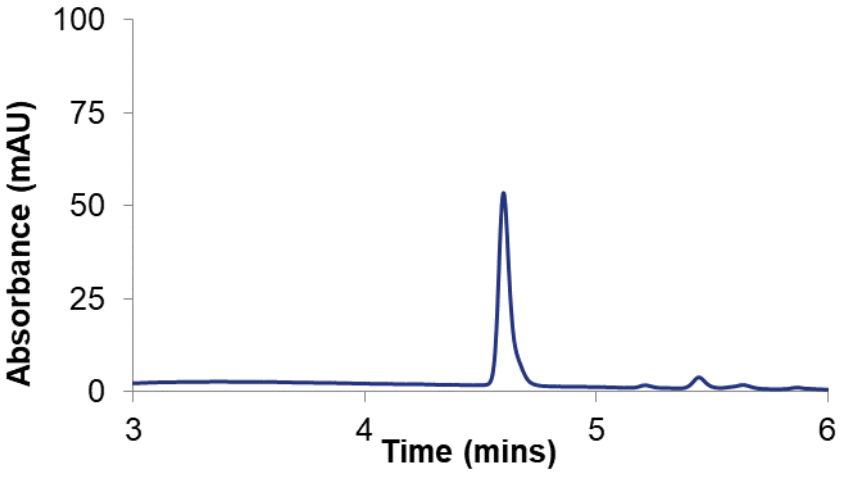


Figure 100 - Analytical HPLC trace of pure biotin-EDA-TEMPO modified Ac-YEPLAUHISCY-NH_2 from one-pot ligation-conjugation (**42**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 1746.8711, $[M+2H]^{2+}$: 873.9395. Observed Mass $[M+H]^+$: 1746.8572, $[M+2H]^{2+}$: 873.9427.

One-pot ligation-conjugation-reductive cleavage – ligation at serine (43)



Synthesised following general one-pot ligation-conjugation protocol using H-UWIMKY-NH_2 (**16**) (1 mg, 1.12 µmol) and Ac-YEPLA-SePh (**32**) (1.04 mg, 1.35 µmol). To conjugation solution was added Zn dust (36.6 mg, 0.56 mmol) and glacial acetic acid and ligation buffer to a final composition of 10 vol% acetic acid, 1 mM wrt the conjugated peptide . The resulting solution was shaken at 50 °C for 16 hours before being purified by semi-preparative RP-HPLC (10-100% B over 30 minutes) and lyophilised to yield the desired peptide (1.2 mg, 0.83 µmol, 74% yield).



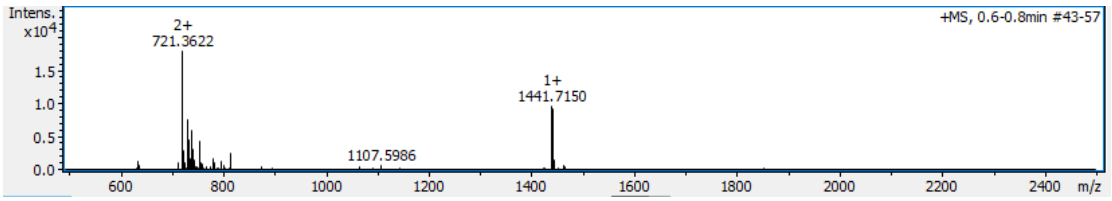


Figure 101 - Analytical HPLC trace of pure reductively cleaved product (**43**) after semi-preparative HPLC purification. Analytical gradient 10-100% B over 5 minutes, 280 nm. Calculated Mass $[M+H]^+$: 1441.8189, $[M+2H]^{2+}$: 721.3634. Observed Mass $[M+H]^+$: 1441.7150, $[M+2H]^{2+}$: 721.3622.

One-pot ligation-oxidation deselenization (Sec to Ser) using Oxone-TCEP method (43b)

Synthesised following general ligation protocol using H-UWIMKY-NH_2 (**16**) (1 mg, 1.12 µmol) and Ac-YEPLA-SePh (**32**) (1.04 mg, 1.35 µmol). The ligation solution was diluted to 0.3 mM with ligation buffer. Next, equal volumes of TCEP stock (15 mM stock, 50 eq.) and oxone stock (15 mM, 50 eq.) both in distilled water were added simultaneously, giving a final peptide concentration of 0.1 mM. This solution was agitated at room temperature for 16 hours before being purified by semi-preparative RP-HPLC (10-100% B over 30 minutes) and lyophilised to yield the desired peptide (1.1 mg, 0.76 µmol, 68% yield).

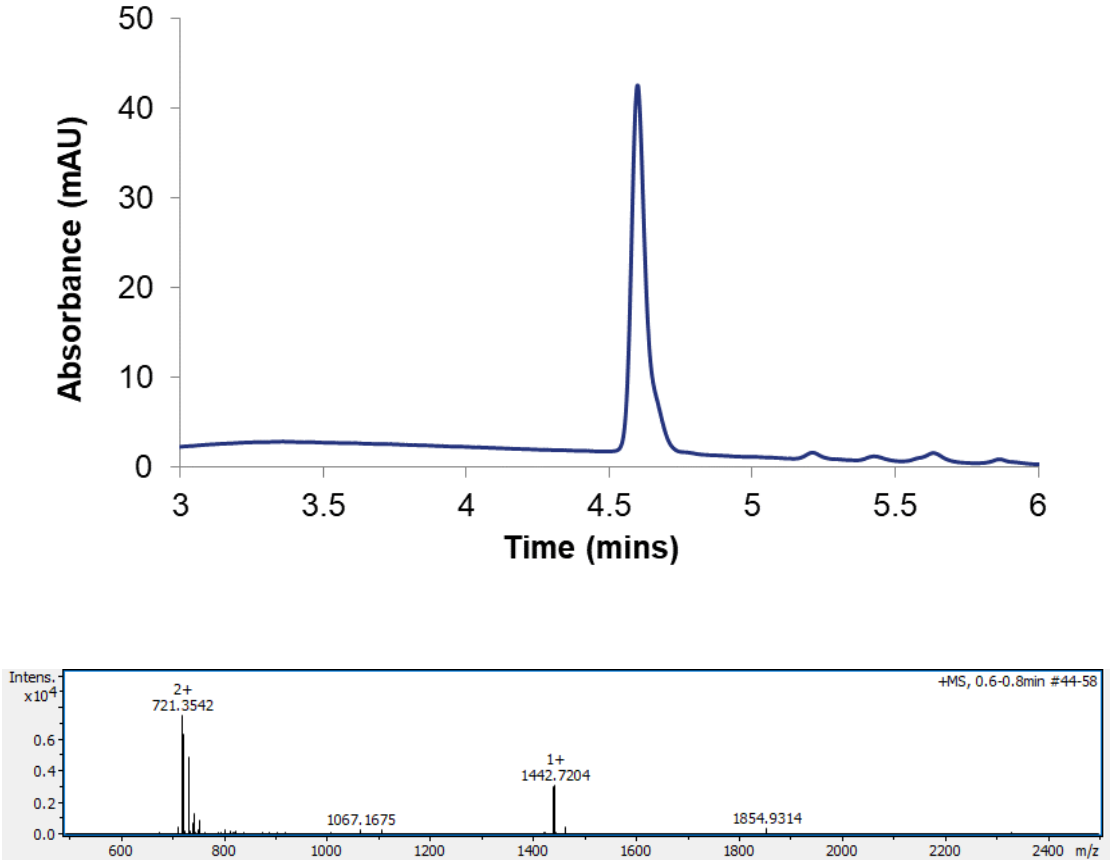


Figure 102 - Analytical HPLC trace of pure ligated/oxidatively cleaved product (**43b**). Analytical gradient 10-100% B over 5 minutes, 280 nm. Calculated Mass $[M+H]^+$: 1441.8189, $[M+2H]^{2+}$: 721.3634. Observed Mass $[M+H]^+$: 1441.7150, $[M+2H]^{2+}$: 721.3622.

ZEGFR:1907 - Ligation and Conjugation (46)

Two step method - ligation followed by conjugation

Step 1- Ligation (47)

ZEGFR-1 Selenoester (**44**) –
H-VDNKFNKEMWAAWEEIRNLPNLNGWQMT-SePh
ZEGFR-2 Selenopeptide (**45**) –
H-UFIASLVDDPSQSANLLAEAKKLNDAQAPK-OH
ZEGFR:1907 Ligated (**47**)–
H-VDNKFNKEMWAAWEEIRNLPNLNGWQMTUFIASLVDDPSQSANLLAEAKKLNDAQAPK-OH

Synthesised using a two-step protocol, using the general additive-free ligation protocol followed by general protocol A. Firstly, ligation was performed using ZEGFR-1 selenoester (**44**) (3.9 mg, 1.09 µmol) and ZEGFR-2 selenopeptide (**45**) (2.92 mg, 0.910 µmol), followed by purification using semi-preparative RP-HPLC (25-60%B over 30 minutes) and lyophilisation to produce the desired ligated peptide (1.90 mg, 0.287 µmol, 32% yield).

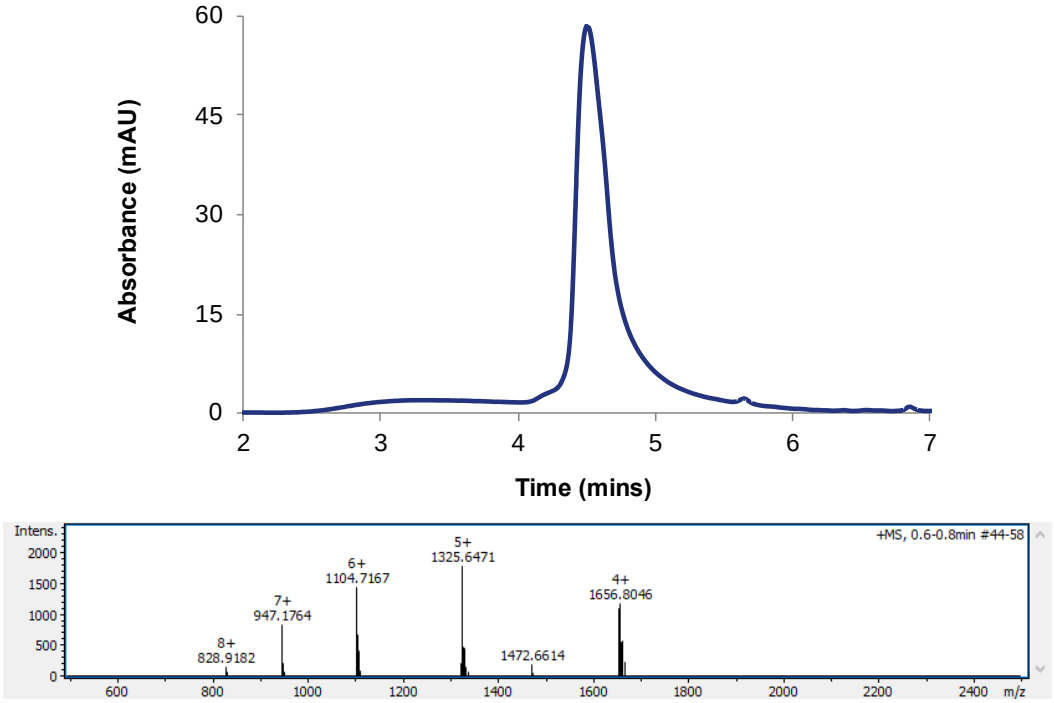
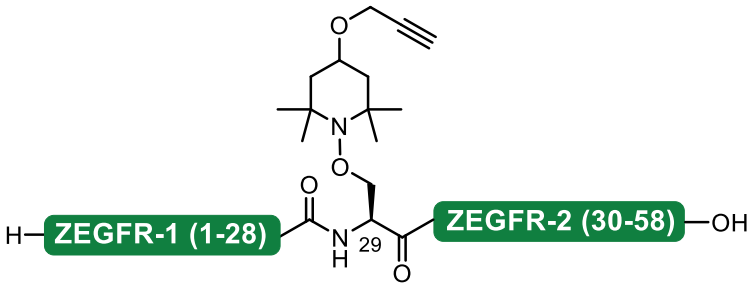


Figure 103 - Analytical HPLC trace of pure ligated ZEGFR:1907 peptide (**47**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+4H]^{4+}$: 1656.0528, $[M+5H]^{5+}$: 1325.0438, $[M+6H]^{6+}$: 1104.3711, $[M+7H]^{7+}$: 946.7478, $[M+8H]^{8+}$: 828.5303. Observed Mass $[M+4H]^{4+}$: 1656.8046, $[M+5H]^{5+}$: 1325.6471, $[M+6H]^{6+}$: 1104.7167, $[M+7H]^{7+}$: 947.1764, $[M+8H]^{8+}$: 828.9182.

Step 2 - Conjugation (46b)



Synthesised following general protocol A using ZEGFR:1907 (**47**, 1.9 mg, 0.287 µmol), followed by purification using semi-preparative RP-HPLC (25-60%B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (1.7 mg, 0.252 µmol, 88% yield).

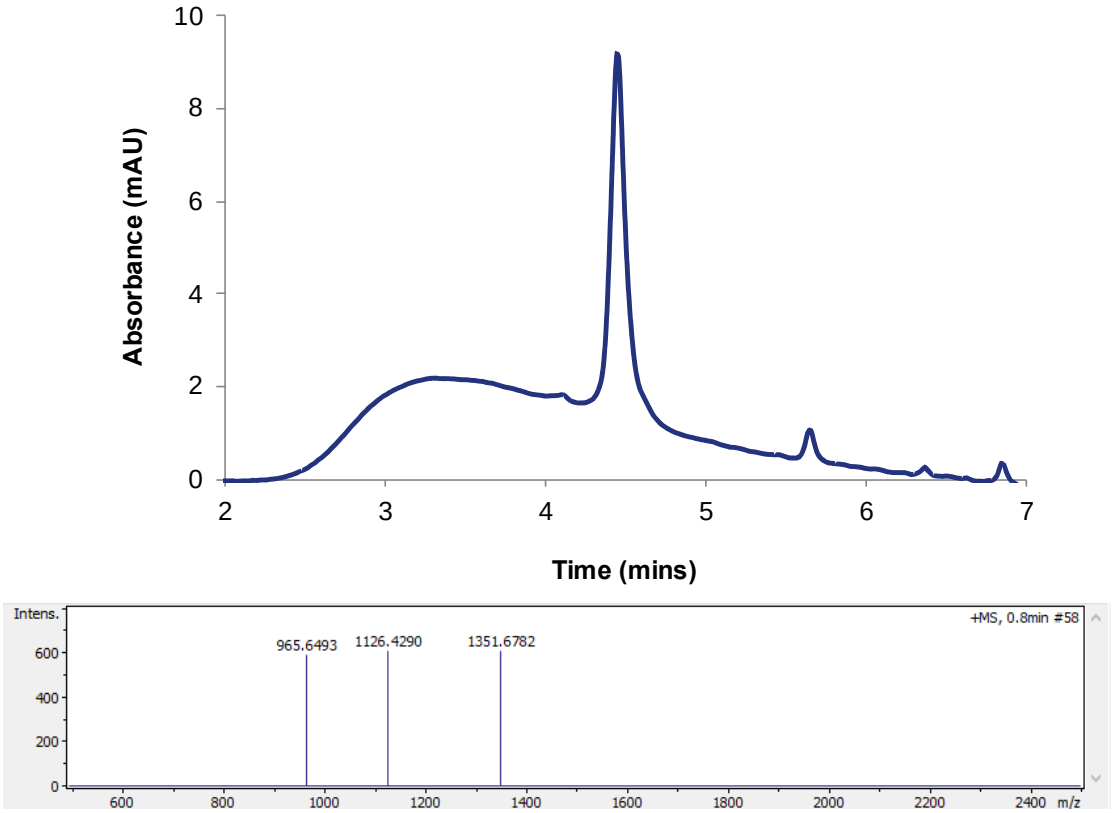
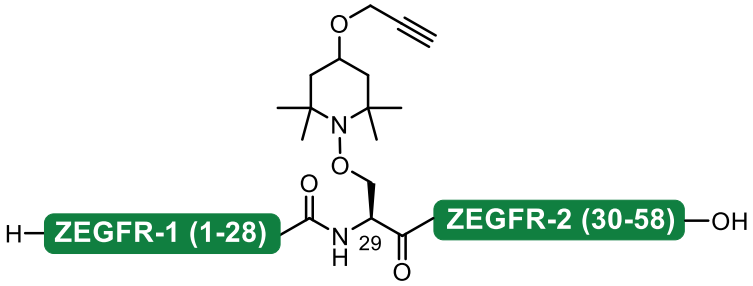


Figure 104 - Analytical HPLC trace of pure conjugated ZEGFR:1907 peptide (**46b**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+5H]^{5+}$: 1350.6898, $[M+6H]^{6+}$: 1125.7428, $[M+7H]^{7+}$: 965.0664. Observed Mass $[M+5H]^{5+}$: 1351.6782, $[M+6H]^{6+}$: 1126.4290, $[M+7H]^{7+}$: 965.6493.

One-pot method to form ZEGFR:1907 conjugated product (46)



Synthesised following general protocol A using ZEGFR-1 (**44**) (4.37 mg, 1.22 µmol) and ZEGFR-2 (**45**) (3.27 mg, 1.02 µmol) followed by purification using semi-preparative RP-HPLC (25-60%B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (2.9 mg, 0.429 µmol, 42% yield).

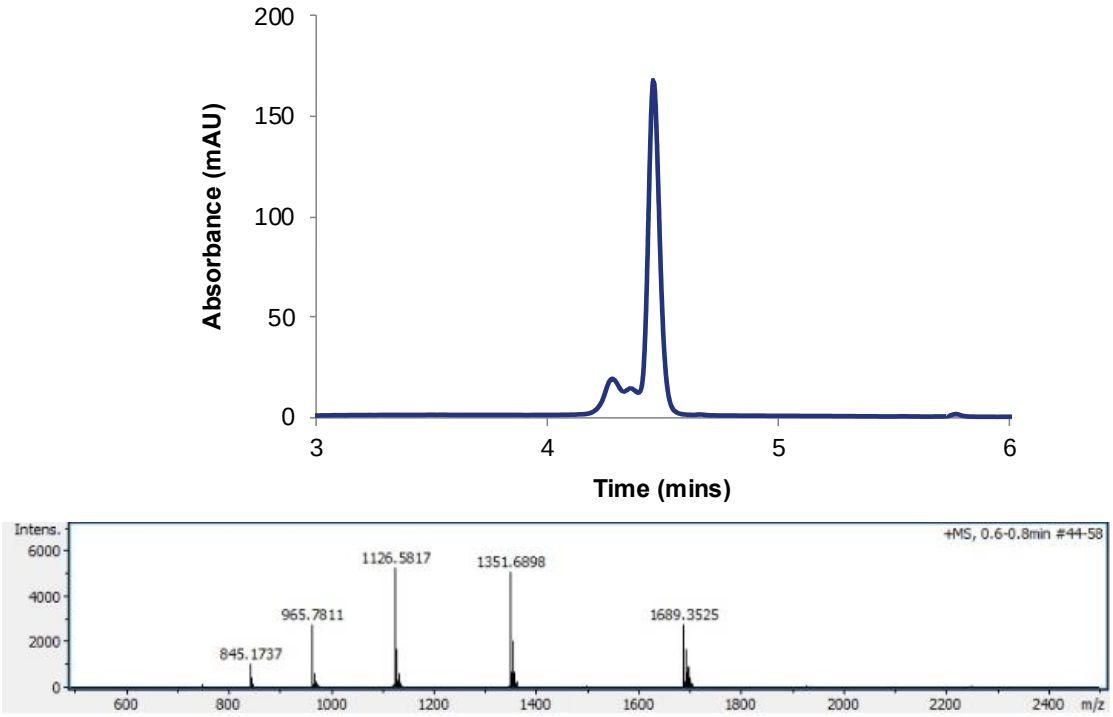
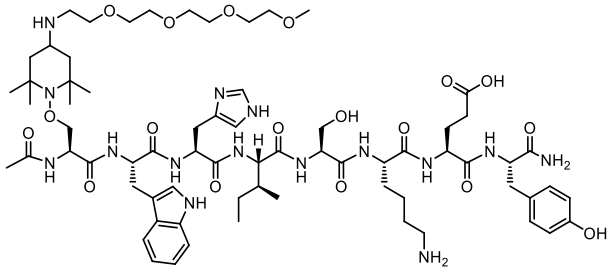


Figure 105 - Analytical HPLC trace of pure one-pot ligated and conjugated ZEGFR:1907 peptide using propargyl-TEMPO (**46**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+4H]^{4}+$: 1689.3458, $[M+5H]^{5}+$: 1351.6782, $[M+6H]^{6}+$: 1126.5665, $[M+7H]^{7}+$: 965.7724, $[M+8H]^{8}+$: 845.1768. Observed Mass $[M+4H]^{4}+$: 1689.3525, $[M+5H]^{5}+$: 1351.6898, $[M+6H]^{6}+$: 1126.5817, $[M+7H]^{7}+$: 965.7811, $[M+8H]^{8}+$: 845.1737.

Ac-CWHISKEY-NH_2 Conjugation Reactions (50-54)

Ac-CWHISKEY-NH_2 with tetra-EG-TEMPO 50)



Synthesised following general protocol B using Ac-CWHISKEY-NH_2 (**49**) (2 mg, 1.81 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (2.7 mg, 3.47 µmol, 80% yield).

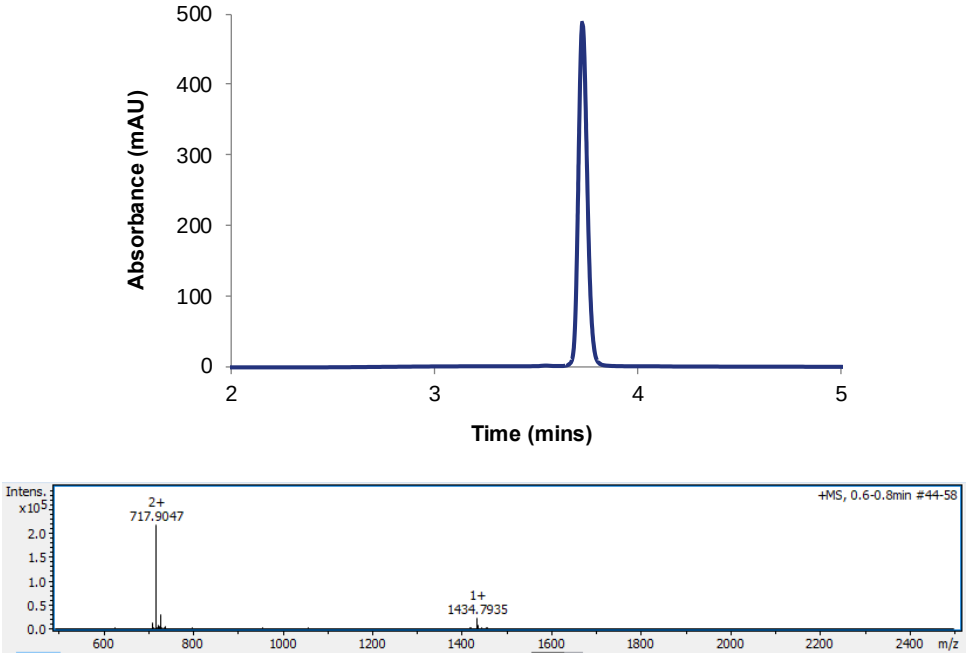
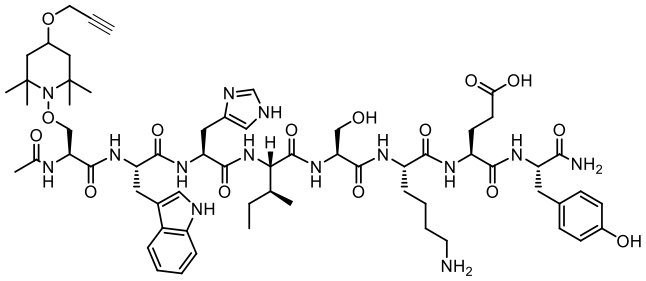


Figure 106 - Analytical HPLC trace of pure conjugated Ac-CWHISKEY-NH_2 using tetra-EG-TEMPO (**50**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 1434.7997, $[M+2H]^{2+}$: 717.9038. Observed Mass $[M+H]^+$: 1434.7935, $[M+2H]^{2+}$: 717.9047.

Ac-CWHISKEY-NH_2 with Propargyl-TEMPO (51)



Synthesised following general protocol B using Ac-CWHISKEY-NH_2 (**49**) (2 mg, 1.81 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (1.9 mg, 1.48 µmol, 82% yield).

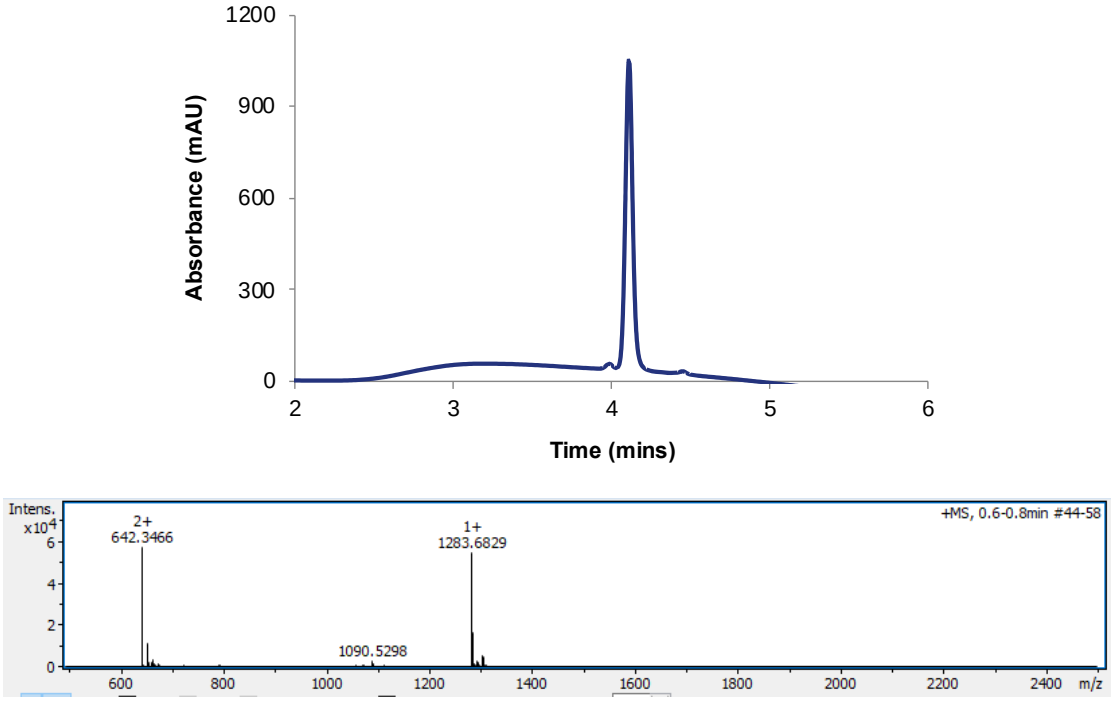
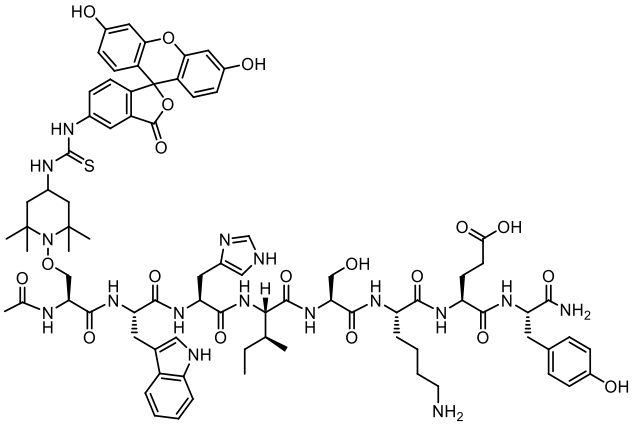


Figure 107 - Analytical HPLC trace of pure conjugated Ac-CWHISKEY-NH_2 using propargyl-TEMPO (**51**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 1283.6788, $[M+2H]^{2+}$: 642.3433. Observed Mass $[M+H]^+$: 1283.6829, $[M+2H]^{2+}$: 642.3466.

Ac-CWHISKEY-NH_2 with FITC-TEMPO (52)



Synthesised following general protocol B using Ac-CWHISKEY-NH_2 (**49**) (2 mg, 1.81 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (2.1 mg, 1.29 µmol, 71% yield).

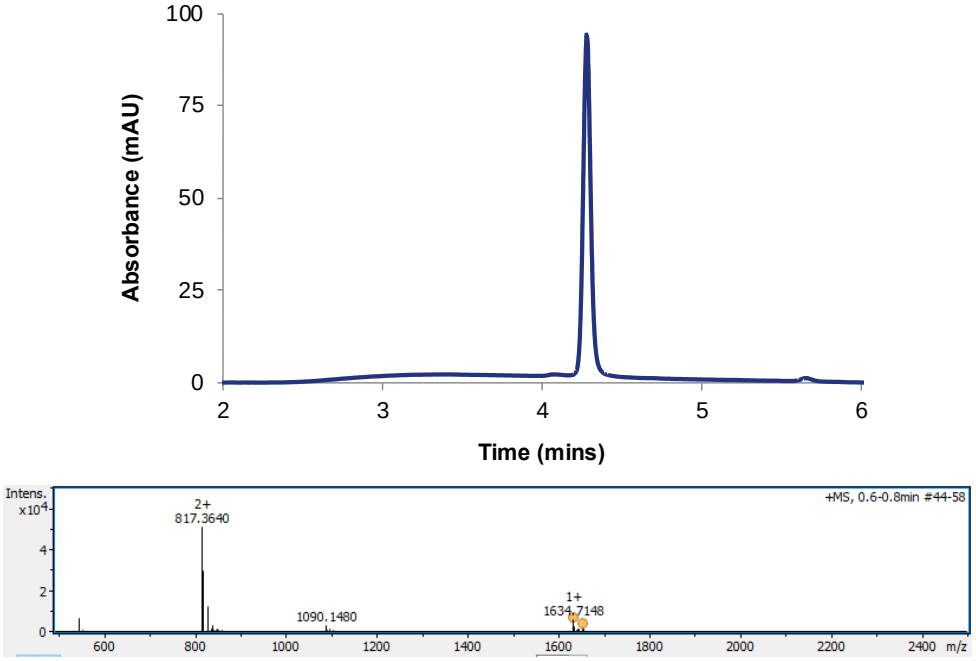
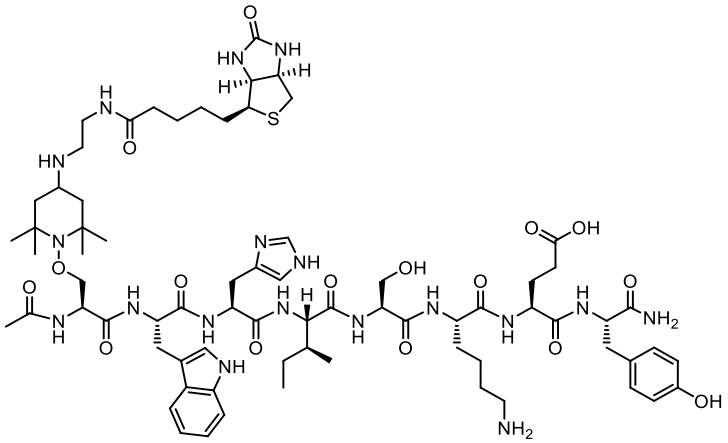


Figure 108 - Analytical HPLC trace of pure conjugated of Ac-CWHISKEY-NH_2 using FITC-TEMPO (**52**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^+$: 1634.7228, $[M+2H]^{2+}$: 817.3614. Observed Mass $[M+H]^+$: 1634.7148, $[M+2H]^{2+}$: 817.3640.

Ac-CWHISKEY-NH_2 with Biotin-EDA-TEMPO (53)



Synthesised following general protocol B using Ac-CWHISKEY-NH_2 (**49**) (2 mg, 1.81 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (2.0 mg, 1.32 µmol, 73% yield).

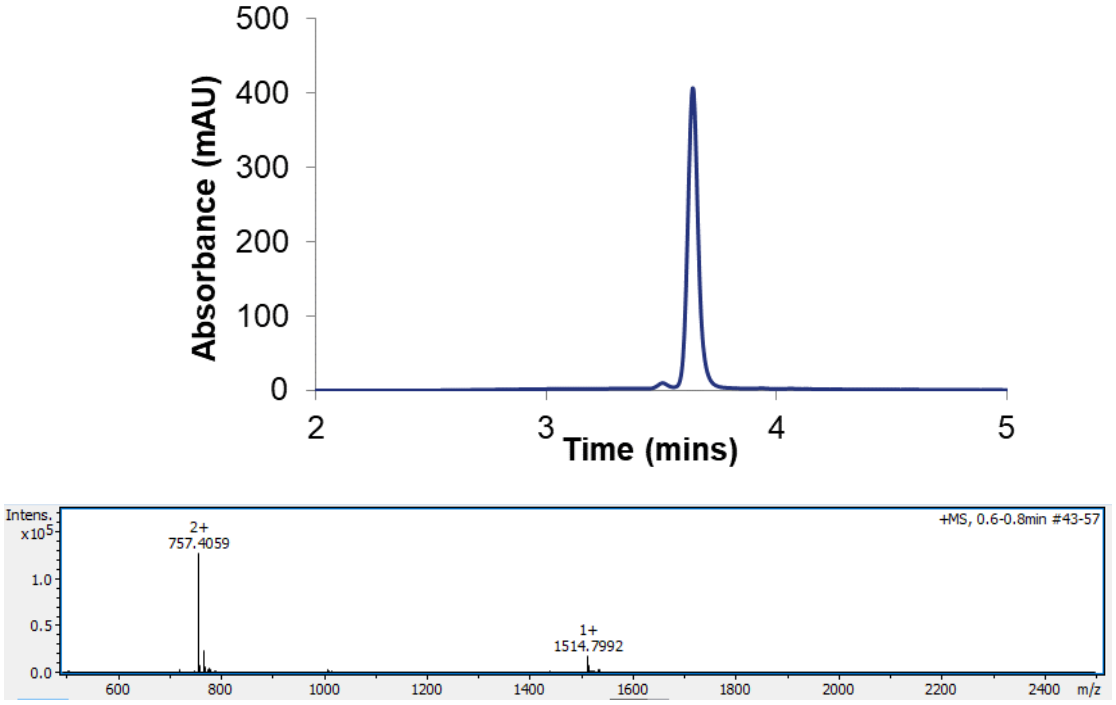
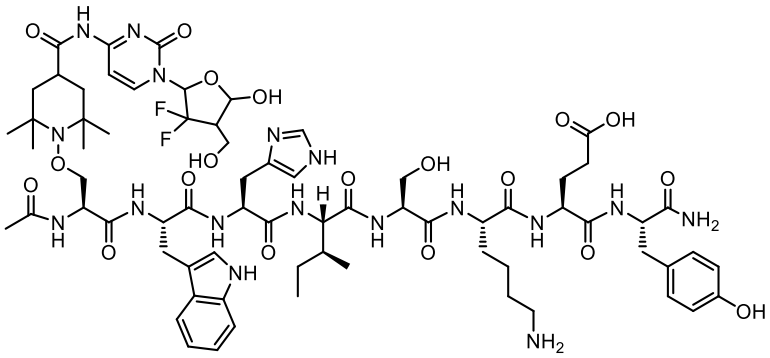


Figure 109 - Analytical HPLC trace of pure conjugated of Ac-CWHISKEY-NH_2 using Biotin-EDA-TEMPO (**53**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^+$: 1513.7990, $[M+2H]^{2+}$: 757.3995. Observed Mass $[M+H]^+$: 1514.7992, $[M+2H]^{2+}$: 757.4059.

Ac-CWHISKEY-NH2 with Gemcitabine-TEMPO (54)



Synthesised following general protocol B using Ac-CWHISKEY-NH_2 (**49**) (5 mg, 4.52 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (5.2 mg, 3.47 µmol, 80% yield).

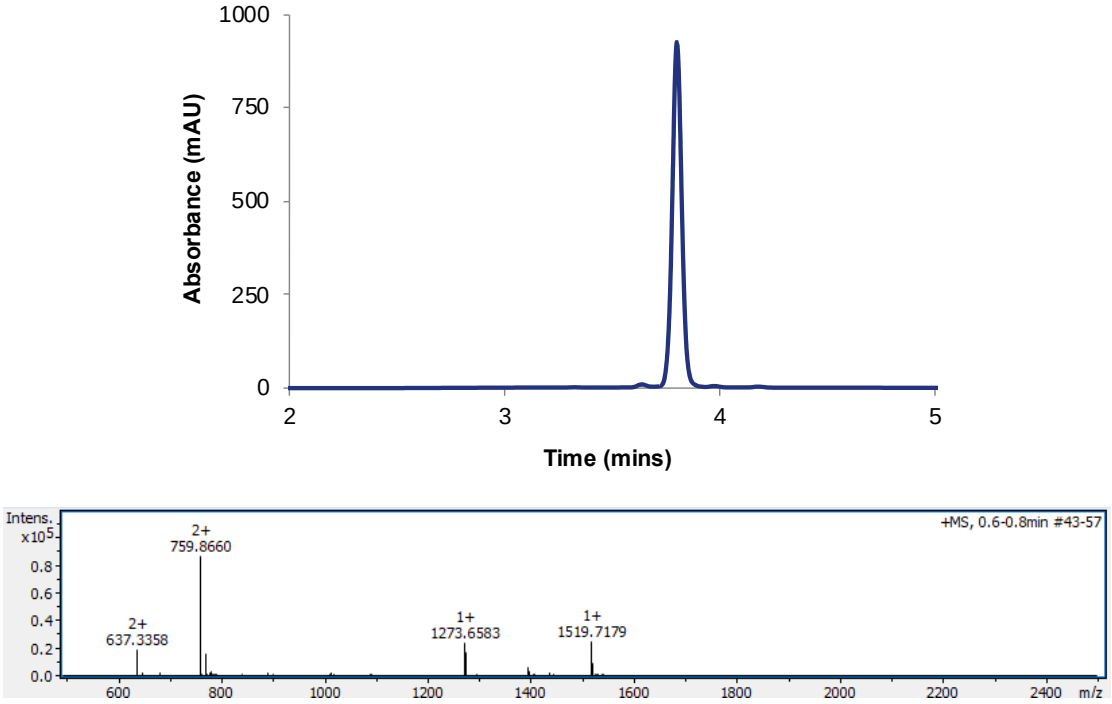
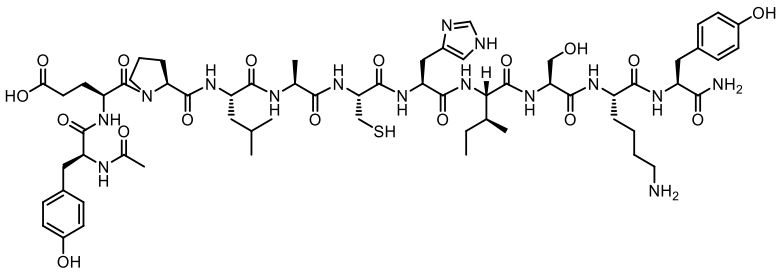


Figure 110 - Analytical HPLC trace of pure conjugated Ac-CWHISKEY-NH_2 using gemcitabine-TEMPO (**54**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^+$: 1519.7272, $[M+2H]^{2+}$: 759.8636. Observed Mass $[M+H]^+$: 1519.7179, $[M+2H]^{2+}$: 759.8660.

Ac-YEPLACHISKY-NH_2 Synthesis by Native Chemical Ligation (56)



Synthesised following the general ligation protocol for N-terminal Cys peptides using H-CHISKY-NH_2 (**46**) (2.3 mg, 3.07 µmol) and Ac-YEPLA-SePh (**32**) (2.49 mg, 3.22 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired ligated peptide (3.2 mg, 2.35 µmol, 77% yield).

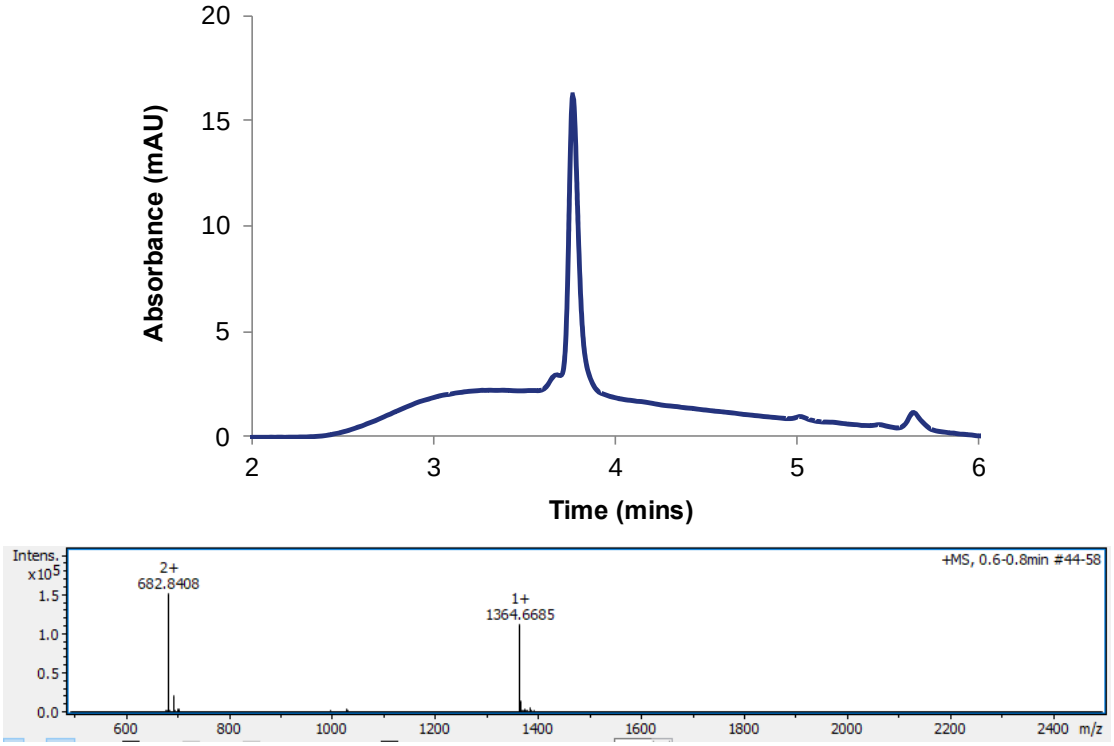
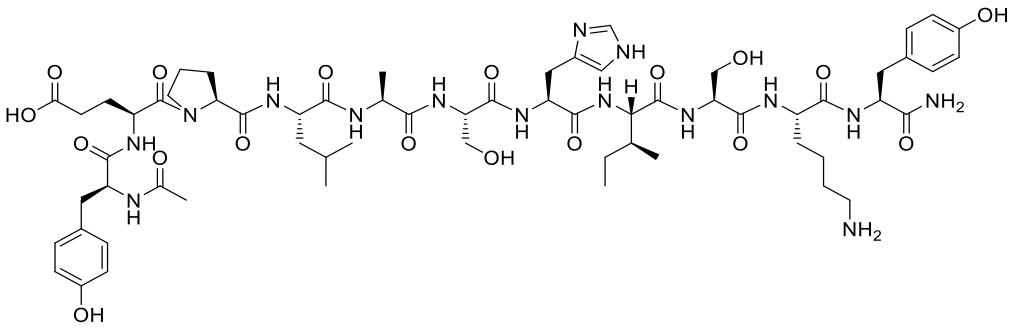


Figure 111 - Analytical HPLC trace of purified ligated product (**47**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^+$: 1364.6672, $[M+2H]^{2+}$: 682.8375. Observed Mass $[M+H]^+$: 1364.6685, $[M+2H]^{2+}$: 682.8408.

Ac-YEPLAA(propargyl-TEMPO)HISKY reductive cleavage (57)



To conjugated product (**35b**, 2.0 mg, 1.30 µmol) dissolved in H_2O: MeOH: AcOH 8:1:1 (0.52 ml, 2.5 mM peptide concentration) was added zinc powder (0.85 mg, 13.0 µmol, 10 eq). The reaction mixture was shaken at 37 °C for 16 hours before being purified by semi-preparative HPLC (5-60% B over 30 minutes) and lyophilized to produce the desired reductively cleaved peptide (1.6 mg, 1.19 µmol, 92% yield).

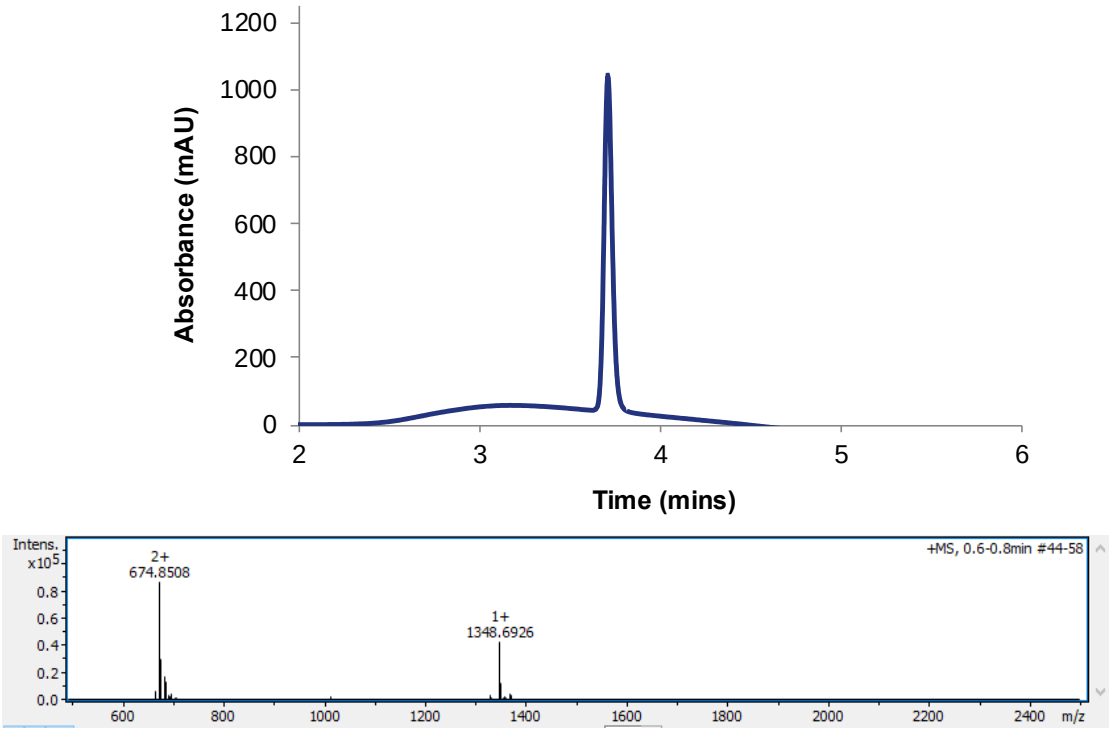


Figure 112 - Analytical HPLC trace of pure reductively cleaved product (**57**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass $[M+H]^+$: 1348.6901, $[M+2H]^{2+}$: 674.8490. Observed Mass $[M+H]^+$: 1348.6926, $[M+2H]^{2+}$: 674.8508.

Ac-YEPLAA(propargyl-TEMPO)HISKY-NH_2 control reductive cleavage – No Zinc

The same reaction conditions were applied to starting peptide (**35b**) as to afford reductively cleaved product (**57**), without the addition of zinc. The reaction mixture was shaken at 37 ºC for 16 hours then analyzed by analytical RP-HPLC.

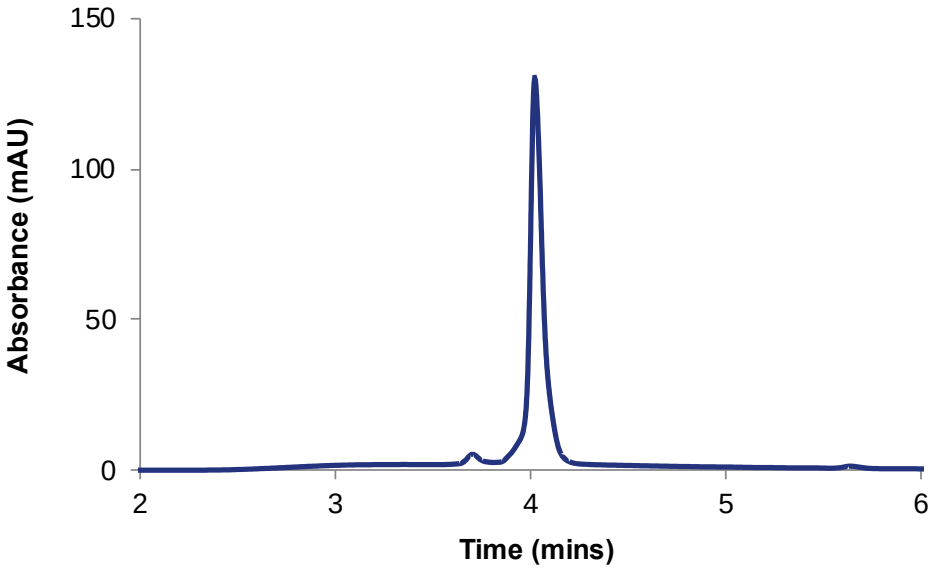
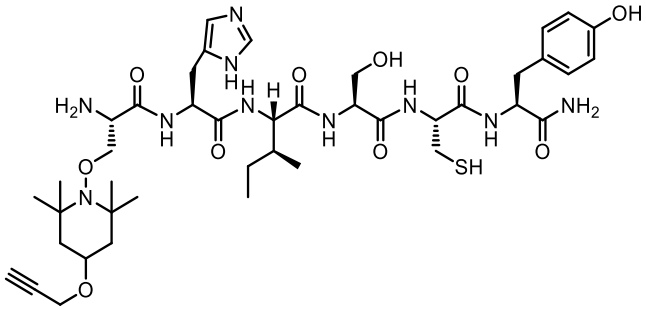


Figure 113 - Analytical HPLC trace of 16 hours crude reaction mixture with no Zn added. Analytical gradient 10-100% B over 5 min.

Reactions Showing Sec/Cys Selectivity (58 and 59)

H-UHISCY-NH_2 Selective modification – Part 1 (58)



Synthesised following general protocol C using H-UHISCY-NH_2 (**39**) (3.0 mg, 3.89 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (2.4 mg, 2.66 µmol, 68% yield).

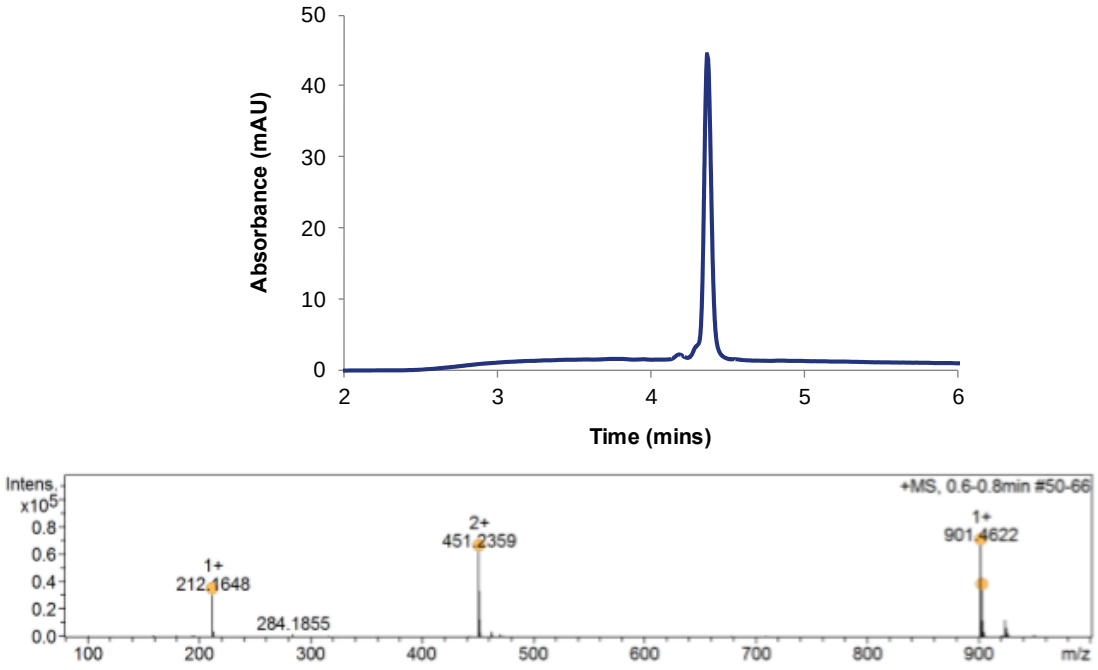
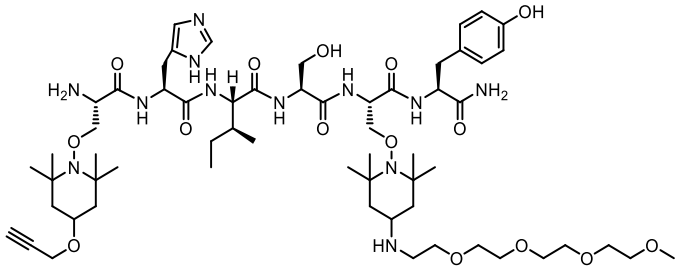


Figure 114 - Analytical HPLC trace of pure conjugated H-UHISCY-NH_2 using propargyl-TEMPO (**58**). Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 901.4606, $[M+2H]^{2+}$: 451.2342. Observed Mass $[M+H]^+$: 901.4622, $[M+2H]^{2+}$: 451.2359.

H-UHISCY-NH_2 Selective modification – Part 2 (59)



Synthesised following general protocol A using H-A(propargylTEMPO-HISCY-NH_2 (**58**) (2.4 mg, 2.66 µmol) followed by purification using semi-preparative RP-HPLC (10-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (2.1 mg, 1.71 µmol, 64% yield).

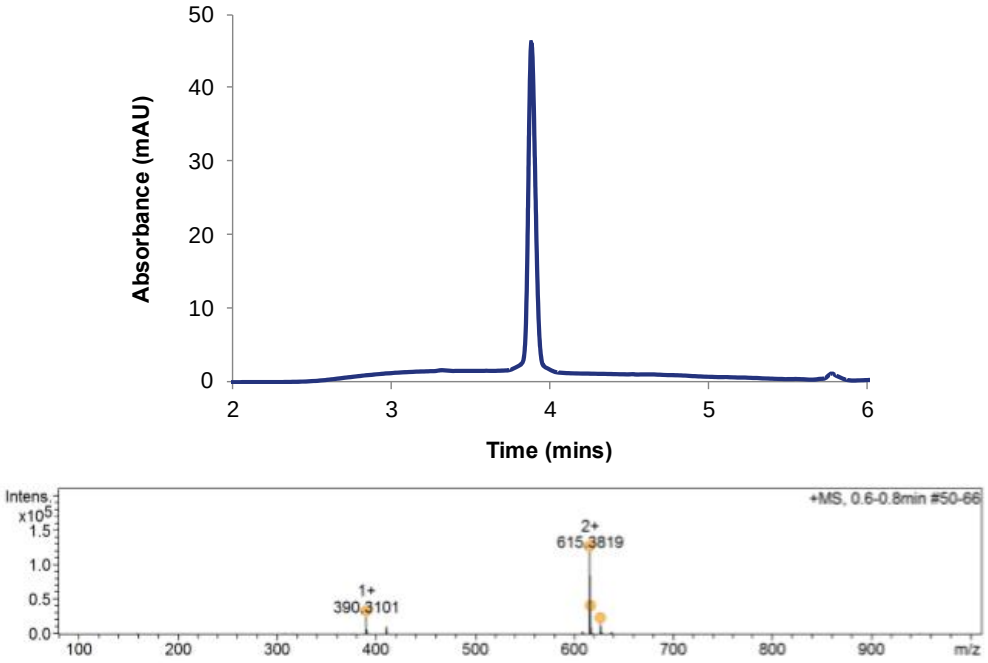
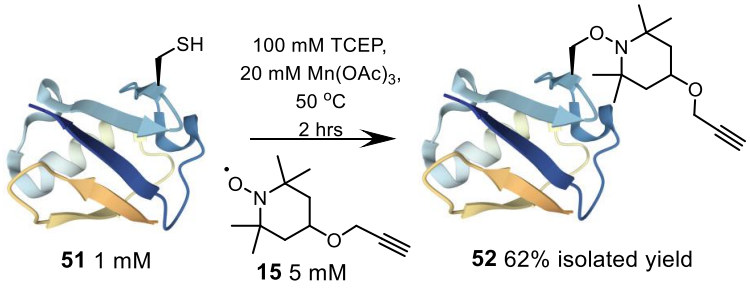


Figure 115 - Analytical HPLC trace of pure conjugated (**59**) using tetra-EG-TEMPO. Analytical gradient 10-100% B over 5 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 615.3794. Observed Mass $[M+2H]^{2+}$: 615.3819.

K48C Ubiquitin Bioconjugation with Propargyl-TEMPO (61)

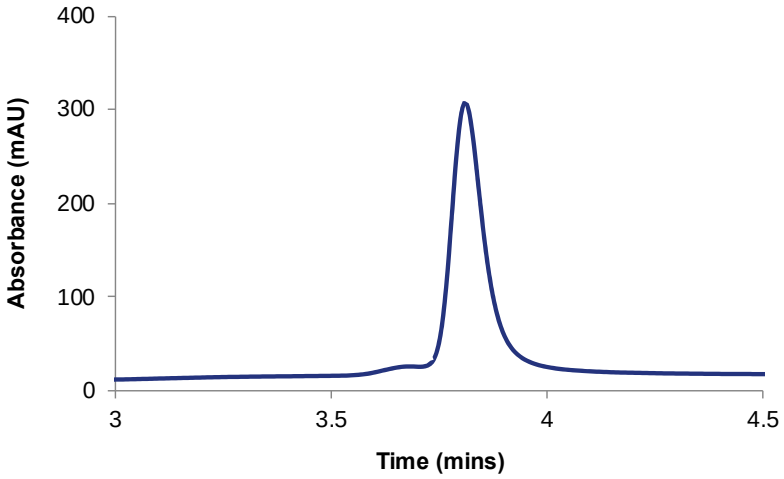
Ubiquitin K48C mutant protein (**51**) was recombinantly expressed and purified as described in Garner *et al.* 2011 Biochemistry with SDS PAGE analysis confirming ~99% purity (not shown).



Ub K48C (43) Sequence – H-MQIFVKTLTGKTITLEVESSDTIDNVKSKI
QDKEGIPPDQQRLIFAGXQLEDGRTLSDYNIQKESTLHLVLRLRGG-OH

X = Modification site

Synthesised following general protocol E using K48C Ubiquitin mutant (**60**, 6.6 mg, 0.774 µmol) followed by purification using semi-preparative RP-HPLC (25-45% B over 30 minutes) and lyophilisation to produce the desired conjugated protein (4.2 mg, 0.482 µmol, 62% yield).



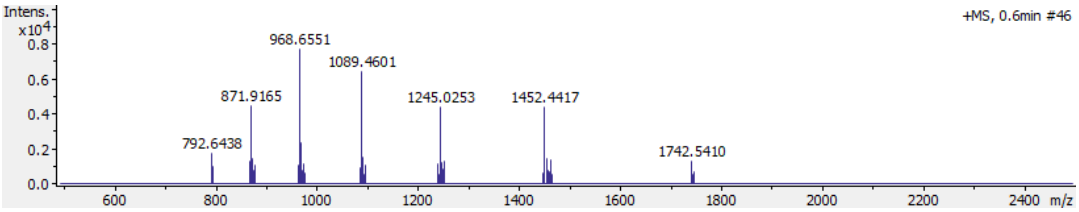


Figure 116 - Analytical HPLC trace of pure Ub K48C modified with propargyl-TEMPO (**61**). Analytical gradient 25-60% B over 5 min, 210 nm. Calculated Mass $[M+5H]^{5+}$: 1743.1397, $[M+6H]^{6+}$: 1452.7844, $[M+7H]^{7+}$: 1245.3877, $[M+8H]^{8+}$: 1089.8402, $[M+9H]^{9+}$: 968.8588, $[M+10H]^{10+}$: 872.0737, $[M+11H]^{11+}$: 792.8859. Observed Mass $[M+5H]^{5+}$: 1742.5410, $[M+6H]^{6+}$: 1452.4417, $[M+7H]^{7+}$: 1245.3877, $[M+8H]^{8+}$: 1089.4601, $[M+9H]^{9+}$: 968.6551, $[M+10H]^{10+}$: 871.9165, $[M+11H]^{11+}$: 792.6438.

NMR analysis of folded Ub conjugate (61)

Purified Ub-conjugate (**61**) and the starting K48C protein (**60**) were dissolved in ligation buffer and dialysed into folding buffer (25 mM Na_2HPO_4/NaH_2PO_4, 100 mM NaCl, pH 6.9) overnight. To each sample was added 10 vol% D_2O and the samples were analysed on an Avance III 800 MHz NMR system. The NMR analysis was performed by Dr Huw Williams and data analysis was performed by Dr Jed Long.

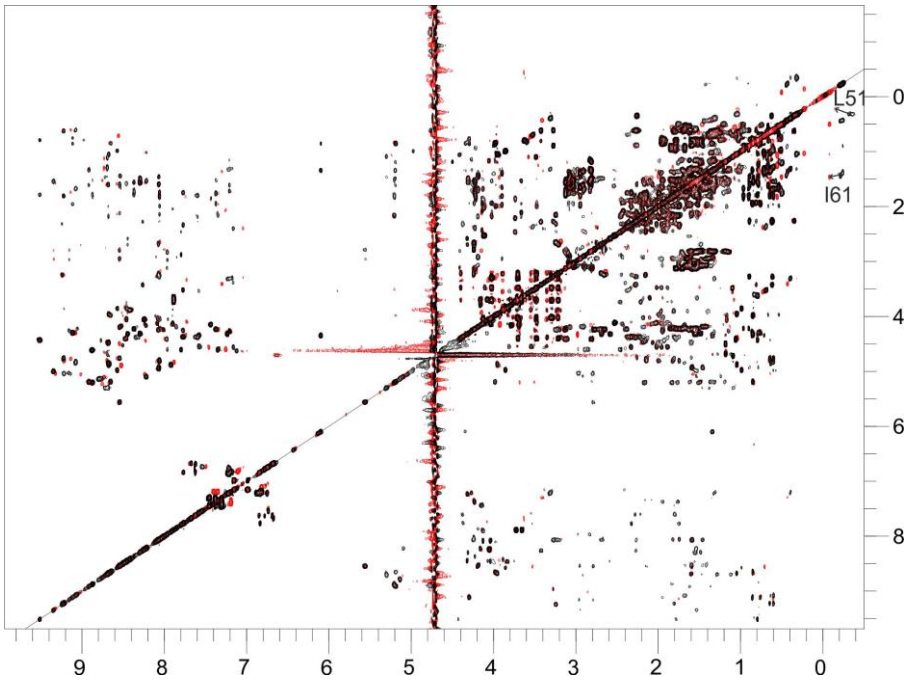


Figure 117 - Overlay of TOSCY spectra of K48C Ub mutant (**601**) and Ub conjugate (**61**).

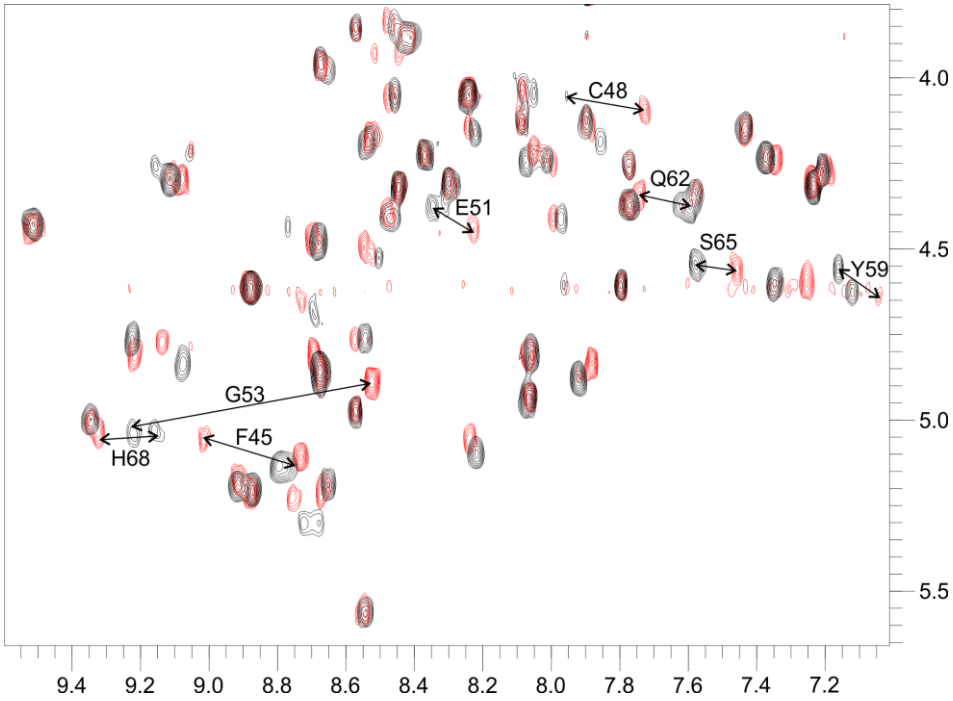


Figure 118 - Expansion of NH region of the overlaid TOCSY spectra for K48C Ub mutant (**60**) and Ub conjugate (**61**); signal shifts around modified residue 48 highlighted.

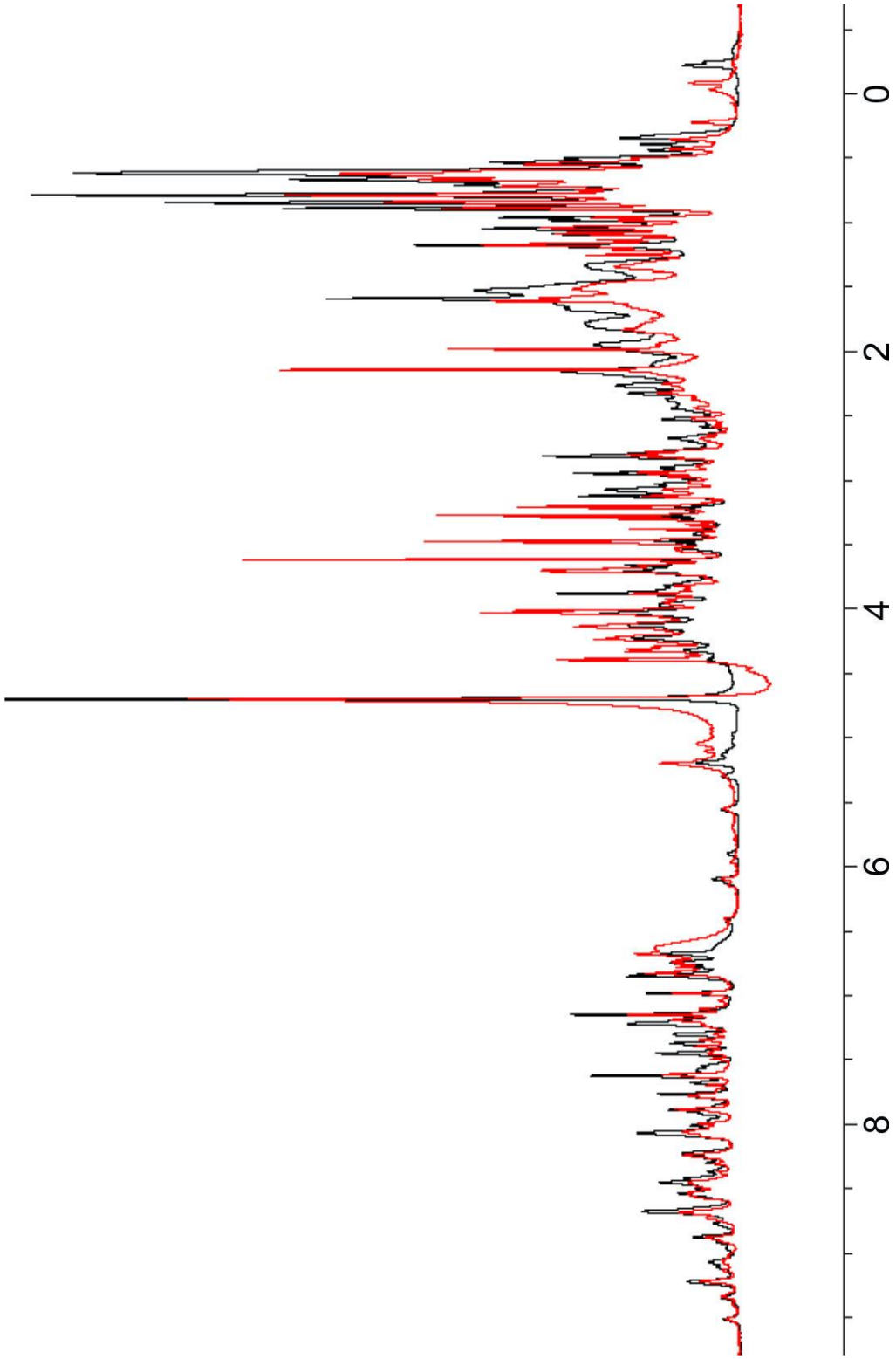
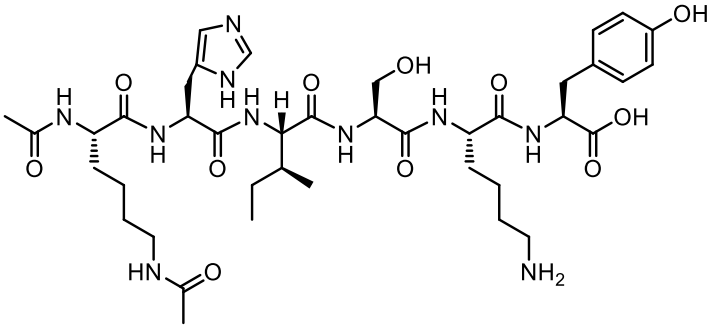


Figure 119 - Overlay for 1H NMR spectra of K48C Ub mutant (**60**) and Ub conjugate (**61**); NH region extended across 6.5-9.5 ppm indicating a tertiary fold.

7.6 Allyl Conjugation Products

Ac-CHISKY-NH_2 with N-allyl acetamide (67)



To Ac-CHISKY-OH (**15**, 1.5 mg, 1.89 µmol) dissolved in 20% MeCN in 0.1 M PB (1.07 ml) was added stock solutions of TCEP (50 eq.), N-allyl acetamide (500 eq.) and Eosin Y (0.1 eq.). The reaction mixture was diluted to give a final peptide concentration of 1 mM and the pH checked to be 7, then placed in blue LEDs for 6 hours. The reaction mixture was purified by preparative RP-HPLC (5-40% B over 30 mins) and lyophilised to produce the desired modified peptide (0.62 mg, 0.722 µmol, 38% yield).

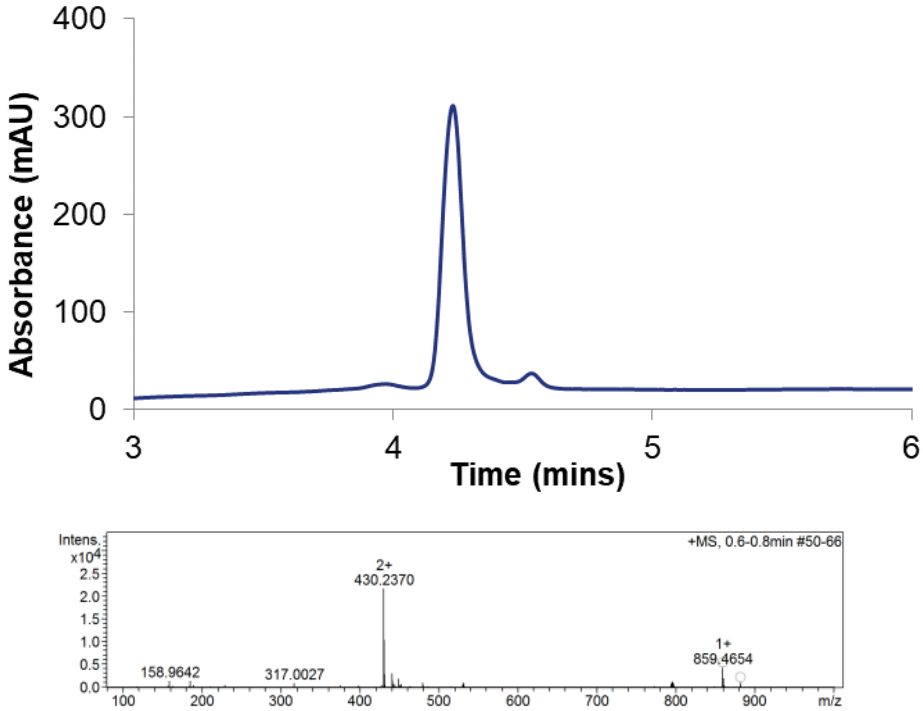
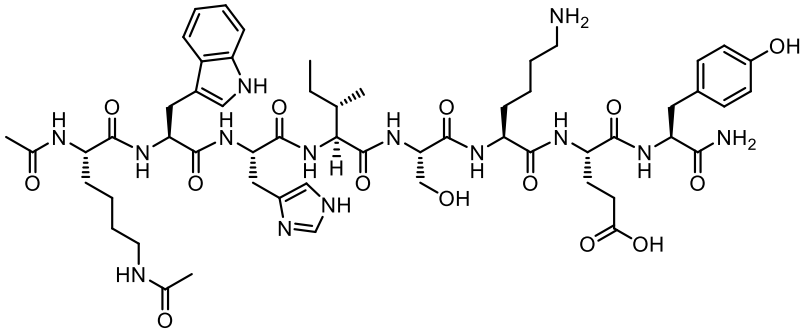


Figure 120 - Analytical HPLC trace of pure N-allyl acetamide modified Ac-CHISKY-OH (**67**). Analytical gradient 10-25% B over 5 min, 210 nm. Calculated Mass $[M+H]^+$: 859.4678, $[M+2H]^{2+}$: 430.2378. Observed Mass $[M+H]^+$: 859.4654, $[M+2H]^{2+}$: 430.2370.

Ac-CWHISKEY-NH_2 (49) Conjugated with N-allyl acetamide (72)



Synthesised following general conjugation procedure 1 using Ac-CWHISKEY-NH_2 (**49**, 4 mg, 3.62 µmol) and purified by semi preparative HPLC (15-80%B over 30 minutes) to produce the desired conjugated peptide (2.2 mg, 1.88 µmol, 52% yield).

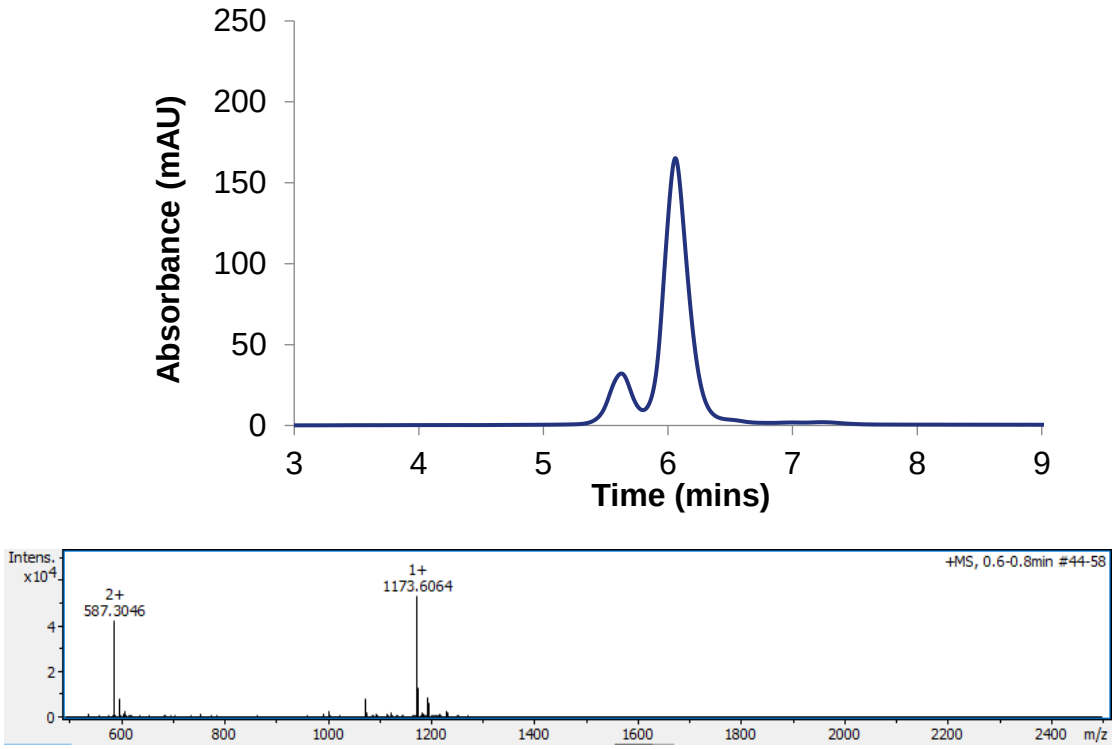
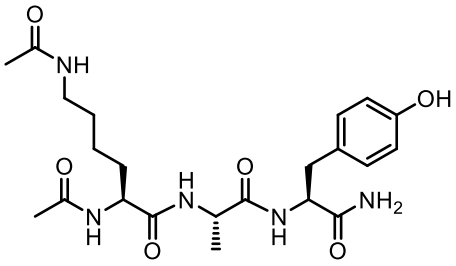


Figure 121 - Analytical HPLC trace of pure Ac-CWHISKEY-NH_2 conjugated with N-allyl acetamide (**72**). Analytical gradient 15-25% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 1173.6057, $[M+2H]^{2+}$: 587.3068. Observed Mass $[M+H]^+$: 1173.6064, $[M+2H]^{2+}$: 587.3046.

Small model Lys PTM installation example for NMR analysis (77 and 78)

Ac-L-CAY-NH_2 modified with N-allyl acetamide (77)



Synthesised following the general protocol F Ac-CAY-NH_2 (**75**, 4 mg, 10.1 µmol) and purified using semi-preparative HPLC (2-60% B over 30 minutes) to yield the desired conjugated peptide (2.6 mg, 5.61 µmol, 56% yield). ^{1}H NMR (500 MHz, CD_3OD) δ 8.25 (d, J = 6.1 Hz, 1H), 7.78 (d, J = 8.0 Hz, 1H), 7.07 (d, J = 8.5 Hz, 2H), 6.71 (d, J = 8.5 Hz, 2H), 4.50 (td, J = 8.2, 5.6 Hz, 1H), 4.28 – 4.20 (m, 2H), 3.17 (dd, J = 7.5, 6.2 Hz, 2H), 3.08 (dd, J = 14.0, 5.6 Hz, 1H), 2.92 (dd, J = 14.0, 8.5 Hz, 1H), 2.02 (s, 3H), 1.95 (s, 3H), 1.79 – 1.60 (m, 2H), 1.53 (p, J = 7.2 Hz, 2H), 1.39 (dq, J = 16.4, 8.2, 7.5 Hz, 2H), 1.29 (d, J = 7.2 Hz, 3H).^{13}C NMR (126 MHz, CD_3OD) δ 174.7 (C), 173.4 (C), 173.3 (C), 172.3 (C), 171.9 (C), 155.9 (Ar), 129.9 (Ar), 127.7 (Ar-H), 114.8 (Ar-H), 54.5 (CH), 53.8 (CH), 49.5 (CH), 38.8 (CH_2), 36.4 (CH_2), 31.0 (CH_3), 28.5 (CH_3), 22.8 (CH_2), 21.2 (CH_2), 21.1 (CH_2), 16.2 (CH_3).

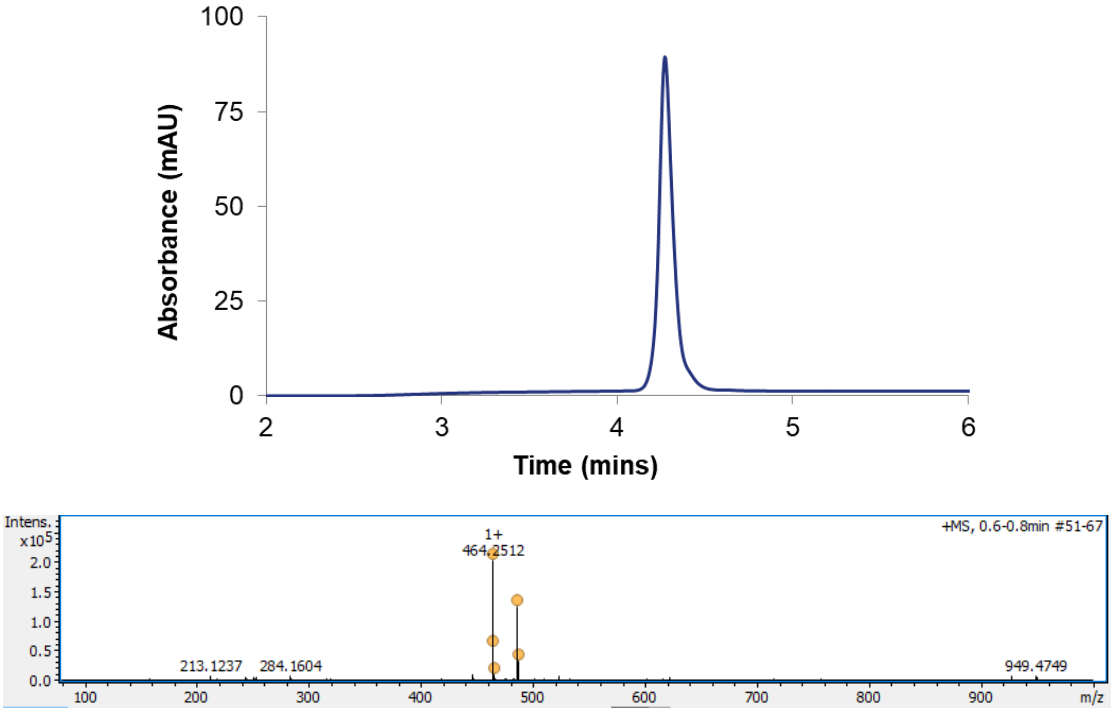
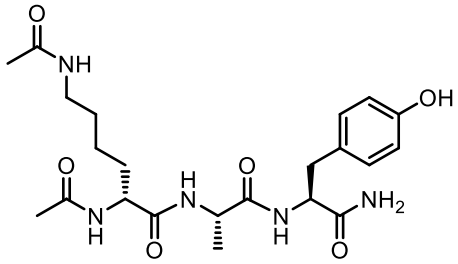


Figure 122 - Analytical HPLC trace of pure Ac-CAY-NH_2 conjugated with N-allyl acetamide (**77**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 464.2509. Observed Mass $[M+H]^+$: 464.2512.

Ac-D-CAY-NH_2 modified with N-allyl acetamide (78)



Synthesised following the general protocol F Ac-D-CAY-NH_2 (**76**, 4 mg, 10.1 µmol) and purified using semi-preparative HPLC (2-60% B over 30 minutes) to yield the desired conjugated peptide (2.4 mg, 5.18 µmol, 51% yield). ^{1}H NMR (500 MHz, CD_3OD) δ 7.13 (d, J = 8.5 Hz, 2H), 6.71 (d, J = 8.5 Hz, 2H), 4.48 – 4.40 (m, 1H), 4.16 – 4.05 (m, 2H), 3.20 – 3.15 (m, 3H), 2.94 (dd, J = 13.9, 11.3 Hz, 1H), 2.00 (s, 3H), 1.94 (s, 3H), 1.72 (dd, J = 9.6, 6.4 Hz, 1H), 1.55 (dd, J = 8.2, 6.2 Hz, 2H), 1.49 – 1.28 (m, 3H), 1.20 (d, J = 7.3 Hz, 3H). ^{13}C NMR (126 MHz, CD_3OD) δ 175.4 (C), 175.0 (C), 173.6 (C), 172.5 (C), 171.9 (C), 155.8 (Ar), 129.8 (Ar), 128.4 (Ar-H), 114.7 (Ar-H), 55.0 (CH), 54.5 (CH), 50.3 (CH), 38.7 (CH_2), 35.8 (CH_2), 30.2 (CH_3), 28.7 (CH_3), 22.8 (CH_2), 21.1 (CH_2), 20.9 (CH_2), 15.6 (CH_3).

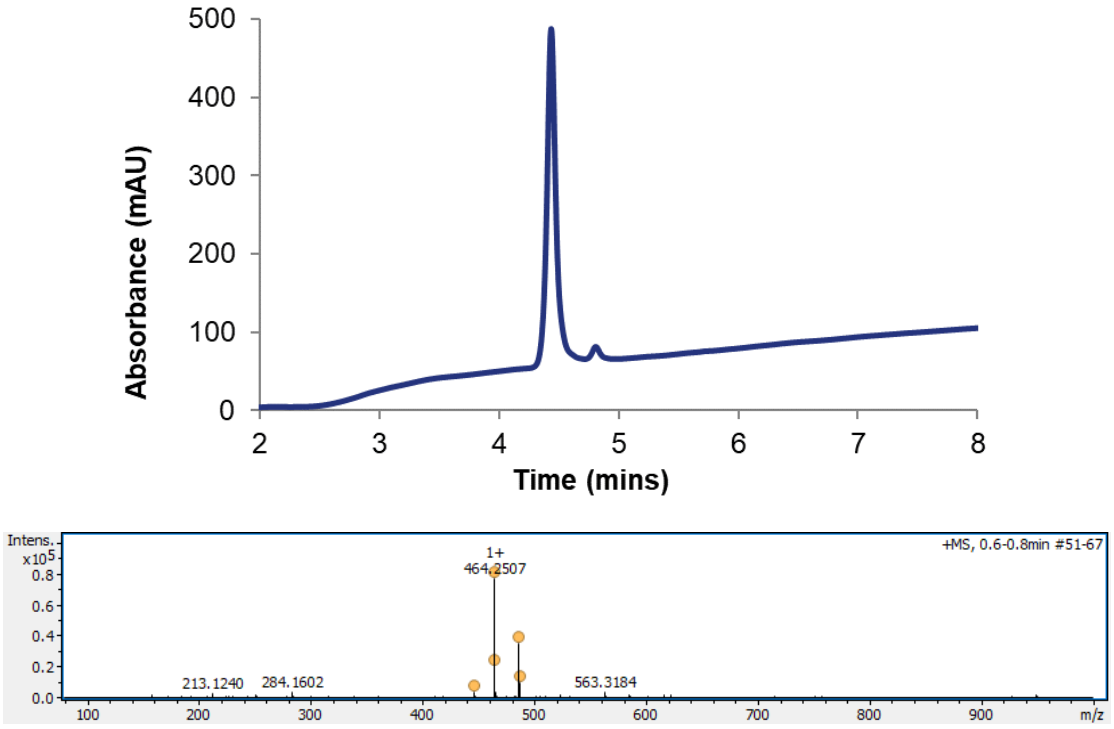
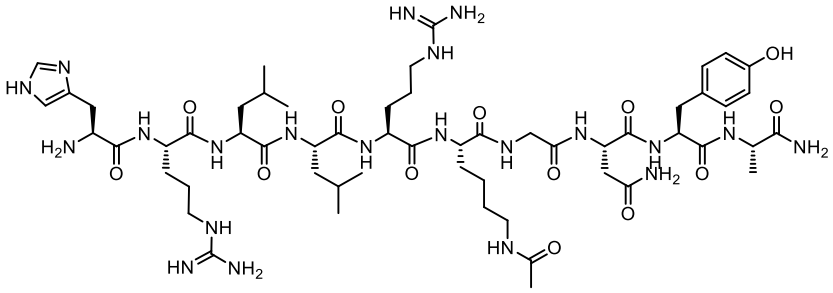


Figure 123 - Analytical HPLC trace of pure Ac-D-CAY-NH_2 conjugated with N-allyl acetamide (**78**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 464.2509. Observed Mass $[M+H]^+$: 464.2507.

Substrate scope – Application to small histone model (80, 86-91)

H-HRLLRCGNYA-NH_2 modified with N-Acetyl lysine (80)



Synthesised following the general conjugation procedure 2 for 1 hour using H-HRLLRCGNYA-NH_2 (**79**, 2 mg, 1.66 µmol) and purified using semi-preparative HPLC (5-40% B over 30 minutes) to yield the desired conjugated peptide (1.2 mg, 0.946 µmol, 57% yield).

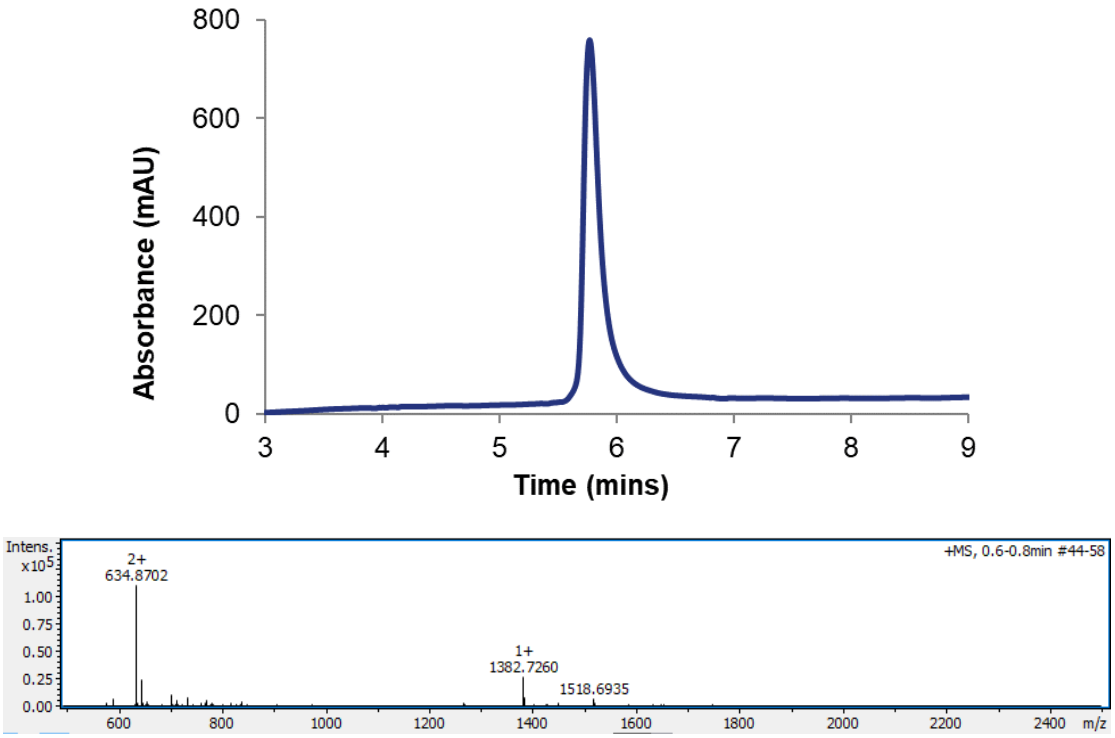
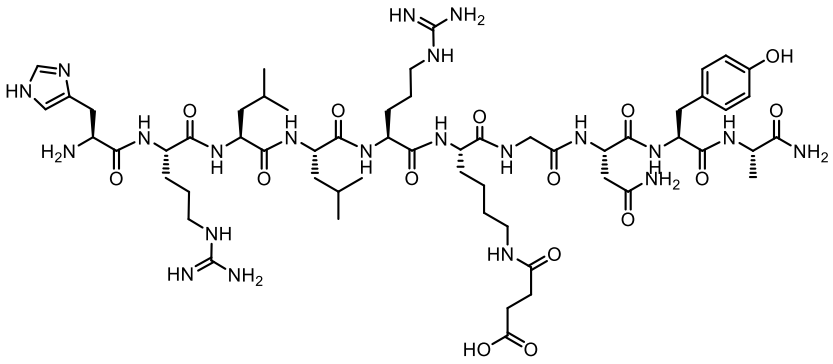


Figure 124 - Analytical HPLC trace of pure H-HRLLRCGNYA-NH_2 conjugated with N-allyl acetamide (**80**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1382.7269, $[M+2H]^{2+}$: 634.8709. Observed Mass $[M+H+TFA]^+$: 1382.7260, $[M+2H]^{2+}$: 634.8702.

H-HRLLRCGNYA-NH_2 modified with N-allylsuccinimide (86)



Synthesised following the general conjugation procedure for 1 hour using H-HRLLRCGNYA-NH_2 (**79**, 2 mg, 1.66 µmol). After analysis the remainder of material (1.565 µmol) was purified using purified using semi-preparative HPLC (5-40% B over 30 minutes) to yield the desired conjugated peptide (1.1 mg, 0.829 µmol, 53% yield).

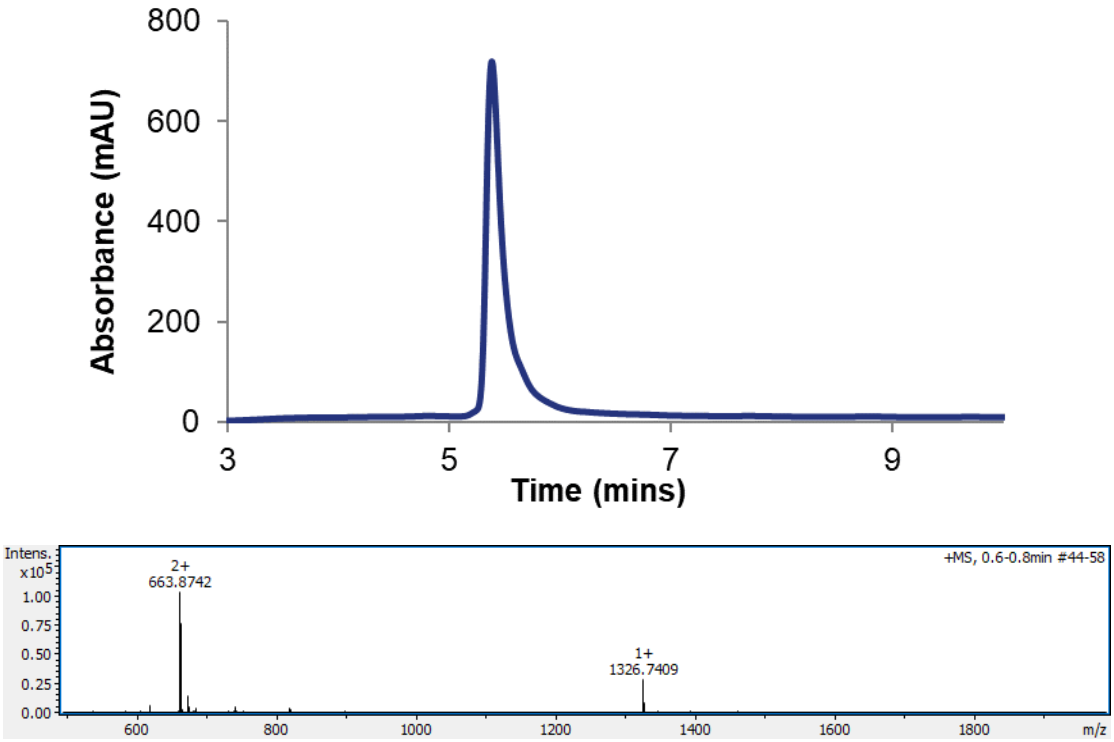
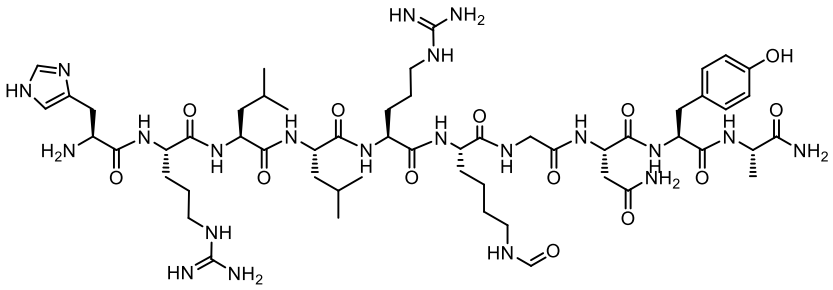


Figure 125 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N-allyl succinimide (**86**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1326.7395, $[M+2H]^{2+}$: 663.8737. Observed Mass $[M+H]^+$: 1326.7409, $[M+2H]^{2+}$: 663.8742.

H-HRLLRCGNYA-NH_2 modified with N-allyl formamide (87)



Synthesised following the general conjugation procedure for 1 hour using H-HRLLRCGNYA-NH_2 (**79**, 2 mg, 1.66 µmol). After analysis the remainder of material (1.565 µmol) was purified using semi-preparative HPLC (5-40% B over 30 minutes) to yield the desired conjugated peptide (1.1 mg, 0.877 µmol, 56% yield).

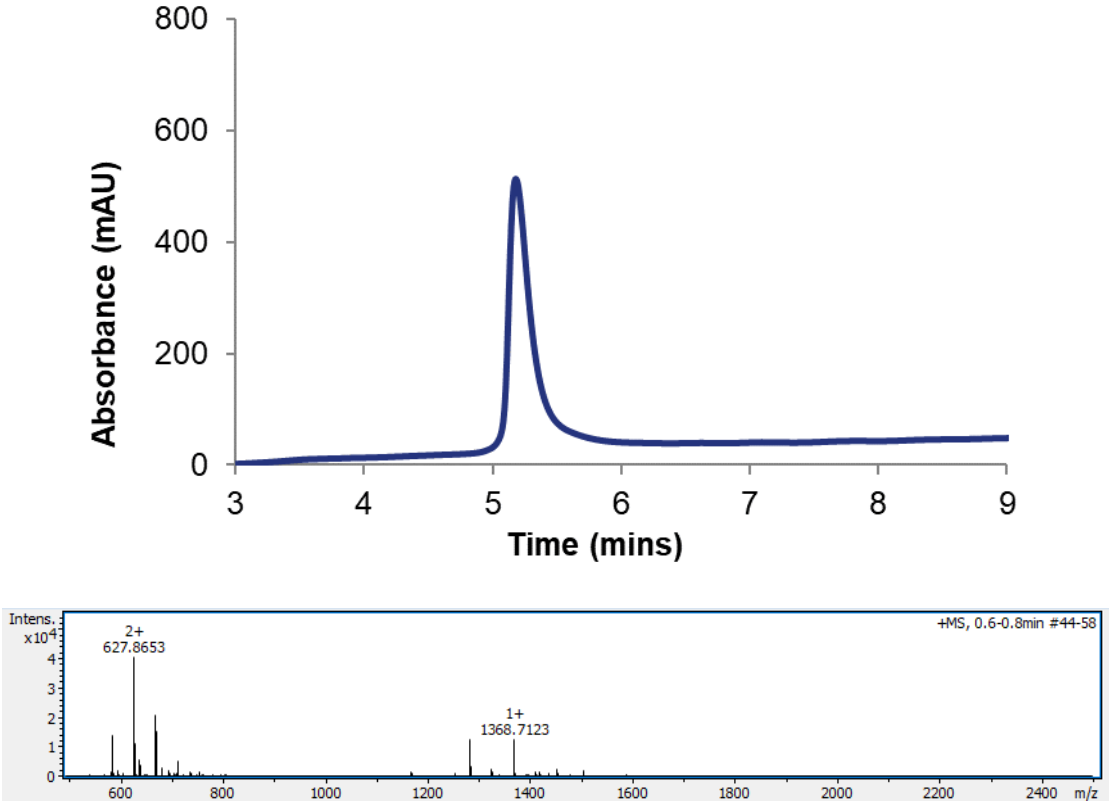
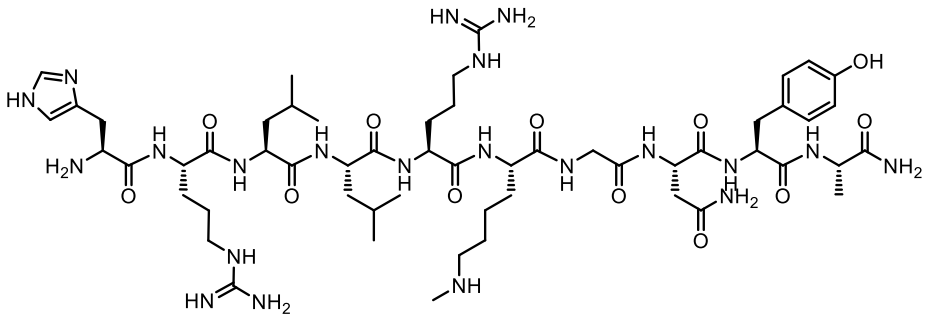


Figure 126 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N-allyl formamide (**87**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1368.7112, $[M+2H]^{2+}$: 627.8631. Observed Mass $[M+H+TFA]^+$: 1368.7123, $[M+2H]^{2+}$: 627.8653.

H-HRLLRCGNYA-NH_2 modified with N-methylallylamine (88)



Synthesised following the general conjugation procedure for 1 hour using H-HRLLRCGNYA-NH_2 (**79**, 4 mg, 3.32 µmol). After analysis the remainder of material (3.23 µmol) was purified using purified using semi-preparative HPLC (5-40% B over 30 minutes) to yield the desired conjugated peptide (1.9 mg, 1.53 µmol, 47% yield).

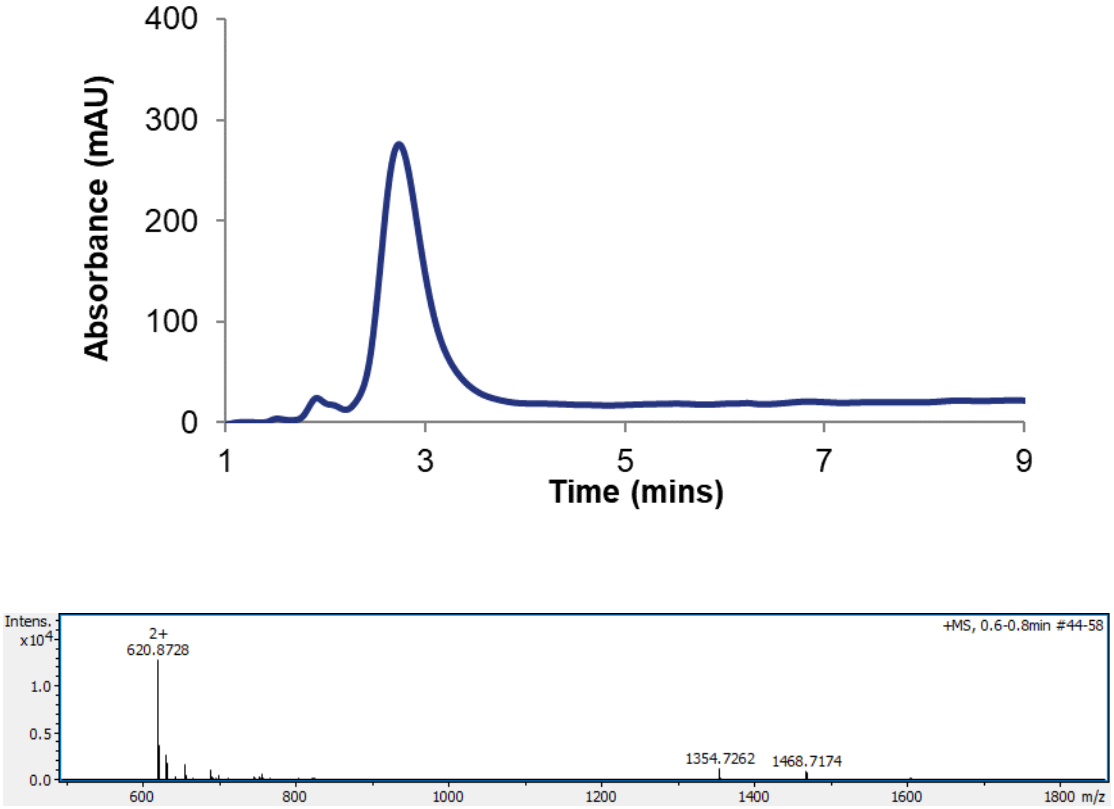
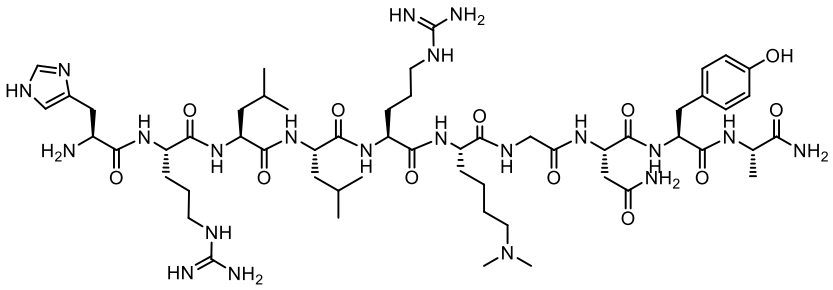


Figure 127 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N-methylallylamine (**88**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1354.7319, $[M+2H]^{2+}$: 620.8735. Observed Mass $[M+H+TFA]^+$: 1354.7262, $[M+2H]^{2+}$: 620.8728.

H-HRLLRCGNYA-NH_2 modified with N,N-dimethylallylamine (89)



Synthesised following the general conjugation procedure for 1 hour using H-HRLLRCGNYA-NH_2 (**79**, 2 mg, 1.66 µmol). After analysis the remainder of material (1.565 µmol) was purified using semi-preparative HPLC (5-40% B over 30 minutes) to yield the desired conjugated peptide (1.1 mg, 0.877 µmol, 56% yield).

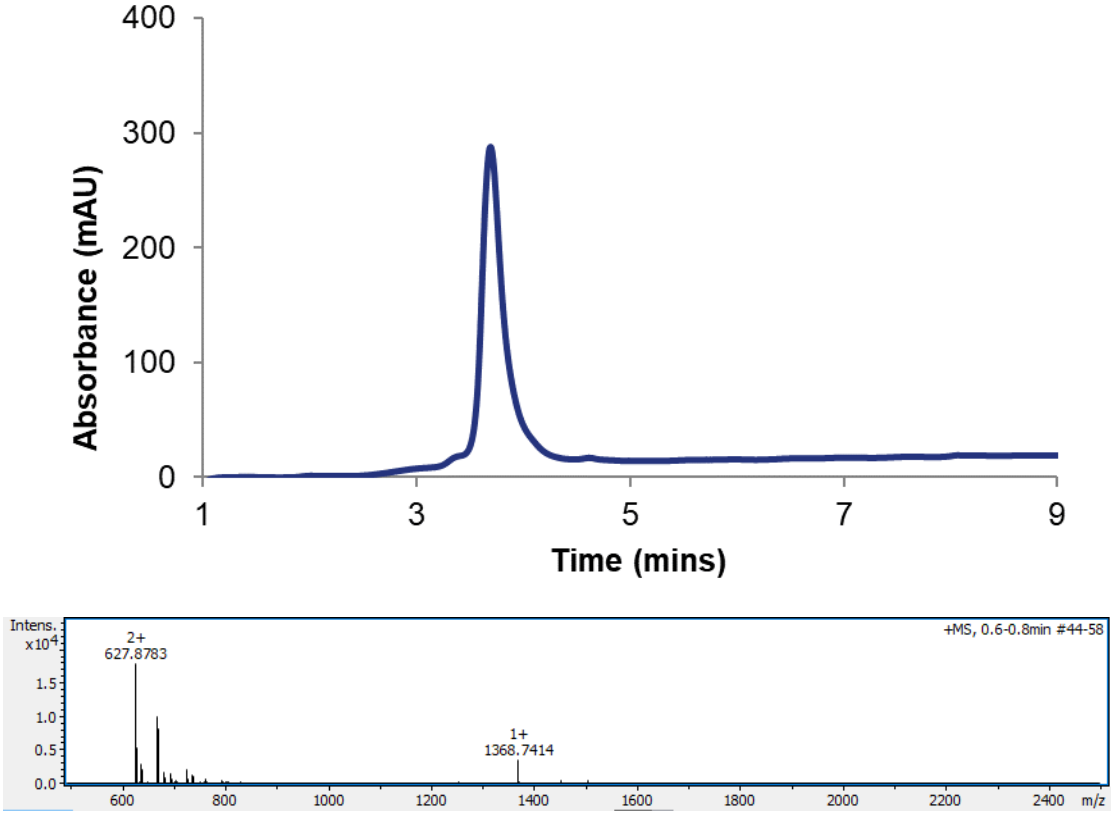
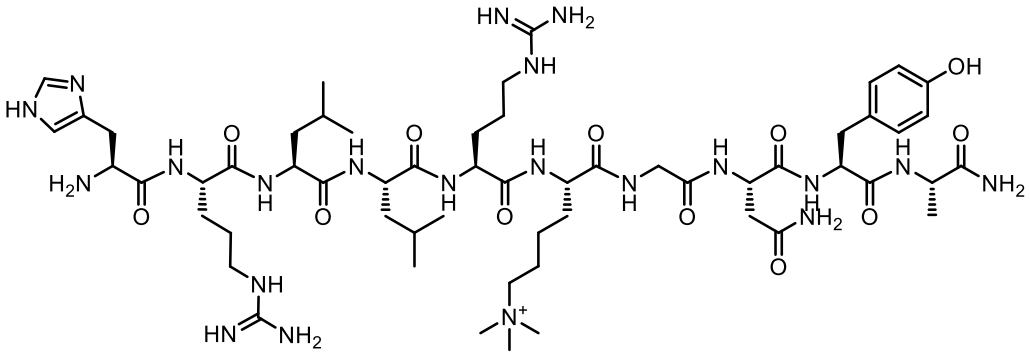
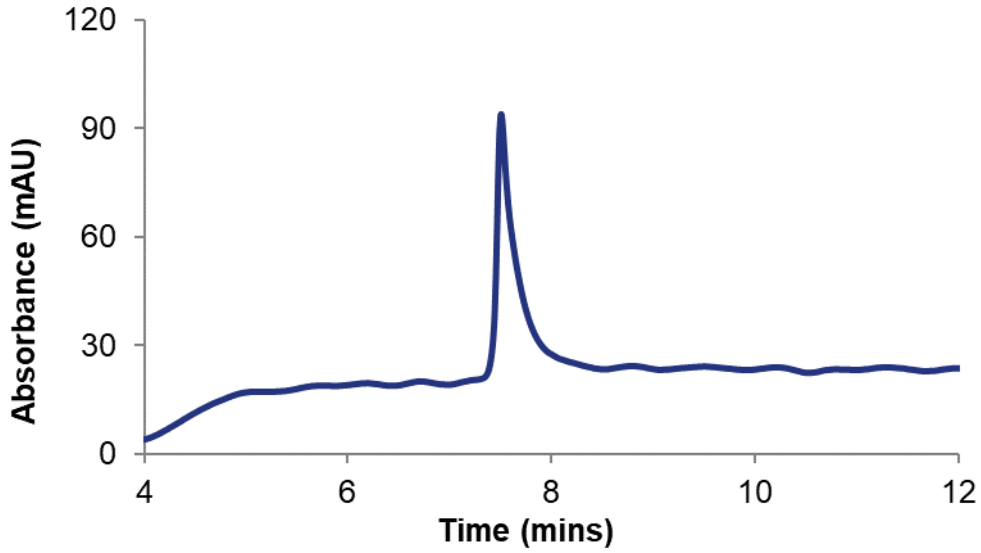


Figure 128 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N,N-dimethylallylamine (**89**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1368.7476, $[M+2H]^{2+}$: 627.8813. Observed Mass $[M+H+TFA]^+$: 1368.7414, $[M+2H]^{2+}$: 627.8783.

H-HRLLRCGNYA-NH_2 modified with allyltrimethylammonium chloride (90)



Synthesised following the general conjugation procedure for 1 hour using H-HRLLRCGNYA-NH_2 (**79**, 2 mg, 1.66 µmol). After analysis the remainder of material (1.565 µmol) was purified using semi-preparative HPLC (5-40% B over 30 minutes) to yield the desired conjugated peptide (1.2 mg, 0.920 µmol, 59% yield).



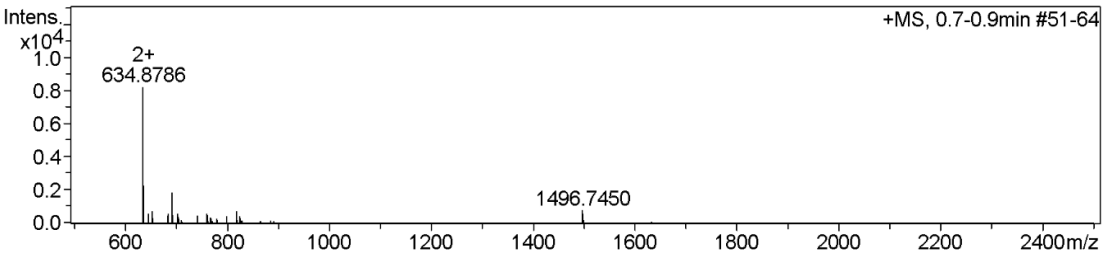
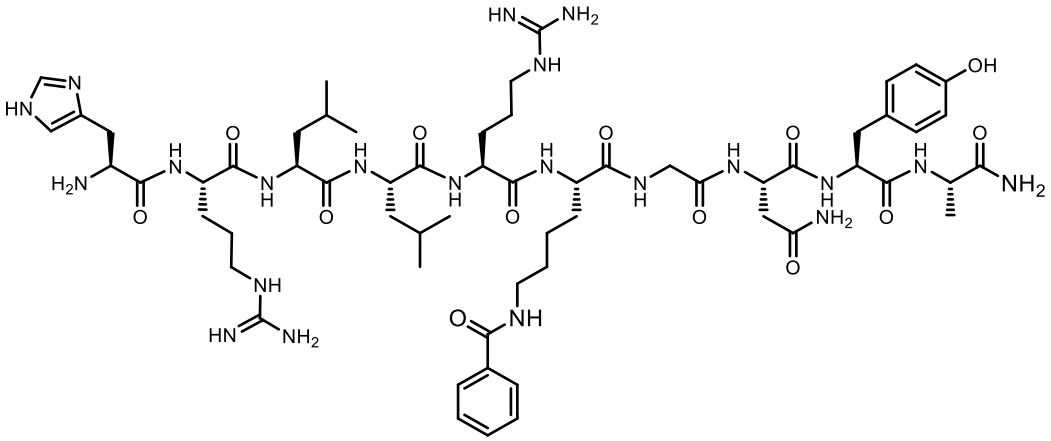


Figure 129 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with trimethylallylammonium chloride (**90**). Analytical gradient 5-30% B over 10 min, 210 nm. Calculated Mass $[M+H]^{2+}$: 634.8891. Observed Mass $[M+H]^{2+}$: 634.8786.

H-HRLLRCGNYA-NH_2 modified with N-allylbenzenamide (91)



Synthesised following the general conjugation procedure for 1 hour using H-HRLLRCGNYA-NH_2 (**79**, 2 mg, 1.66 µmol). After analysis the remainder of material (1.565 µmol) was purified using semi-preparative HPLC (5-40% B over 30 minutes) to yield the desired conjugated peptide (1.3 mg, 0.977 µmol, 62% yield).

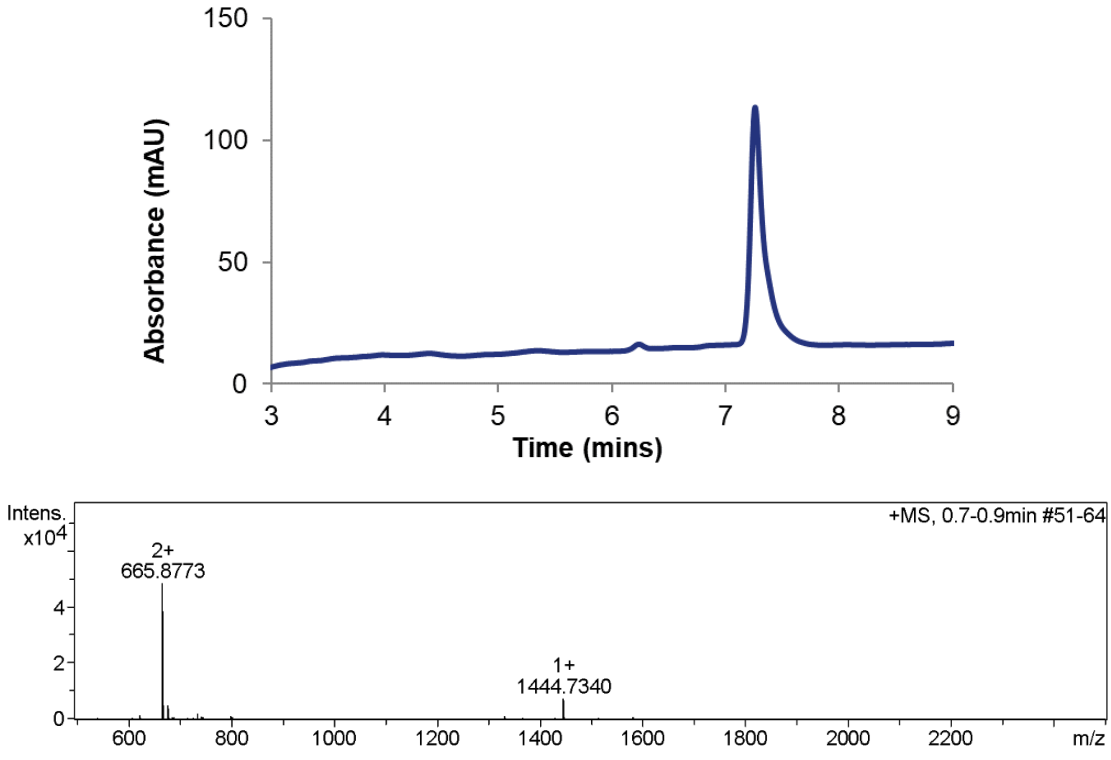
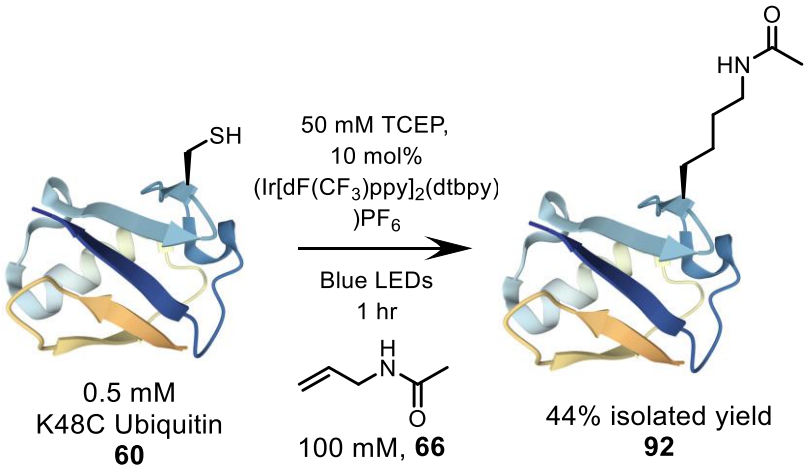


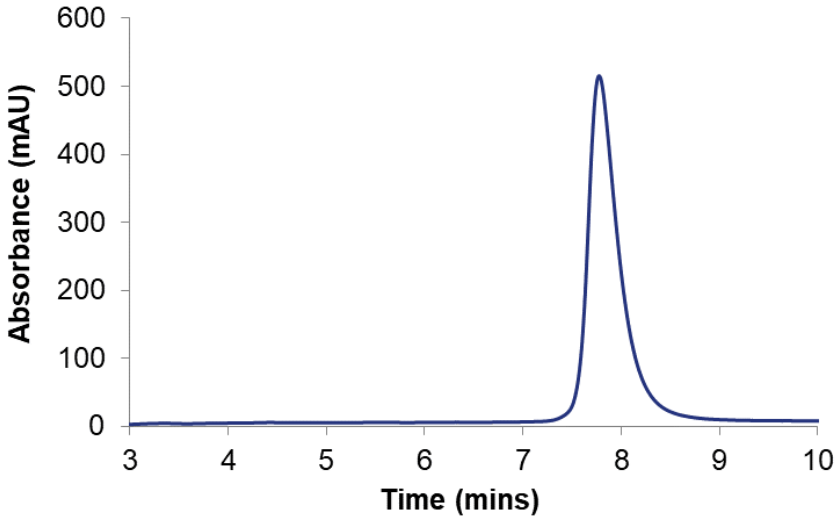
Figure 130 - Analytical HPLC trace of H-HRLLRCGNYA-NH_2 conjugated with N-allyl benzenamide (**91**). Analytical gradient 10-30% B over 10 min, 210 nm. Calculated Mass $[M+H+TFA]^+$: 1444.7425, $[M+2H]^{2+}$: 665.8788. Observed Mass $[M+H+TFA]^+$: 1444.7340, $[M+2H]^{2+}$: 665.8773.

Native Lys PTM installation into proteins

Installation of Lys(Ac) into K48C Ubiquitin (92)



Synthesised following the general conjugation procedure, using 100 eq. TCEP instead of 10 eq., for 1 hour using K48C Ubiquitin (**60**, 4 mg, 0.469 µmol). After analysis the remainder of material (0.444 µmol) was purified using semi-preparative HPLC (25-40% B over 60 minutes) to yield the desired conjugated peptide (1.7 mg, 0.197 µmol, 44% yield).



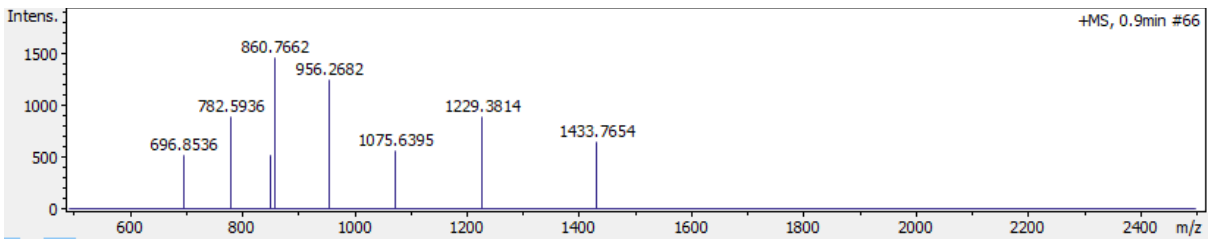
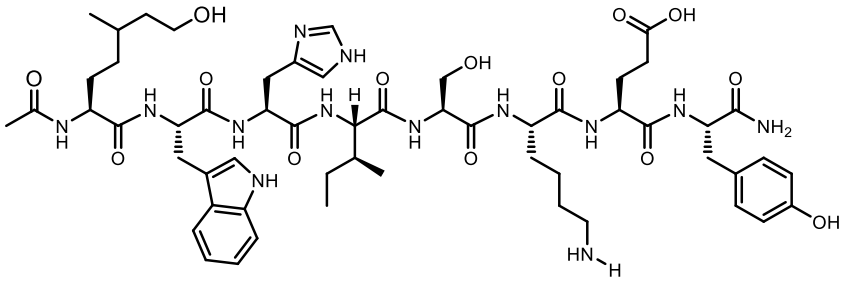


Figure 131 - Analytical HPLC trace of pure K48C ubiquitin conjugated with N-allyl acetamide (**92**). Analytical gradient 25-35% B over 10 min, 210 nm. Calculated Mass $[M+6H]^{6+}$: 1433.2721, $[M+7H]^{7+}$: 1228.6629, $[M+8H]^{8+}$: 1075.2061, $[M+9H]^{9+}$: 955.8507, $[M+10H]^{10+}$: 860.3664, $[M+11H]^{11+}$: 782.2429. Observed Mass $[M+6H]^{6+}$: 1433.7654, $[M+7H]^{7+}$: 1229.3814, $[M+8H]^{8+}$: 1075.6395, $[M+9H]^{9+}$: 956.2682, $[M+10H]^{10+}$: 860.7662, $[M+11H]^{11+}$: 782.5936.

7.7 Stabilised Alkene Conjugation Products

Ac-CWHISKEY-NH_2 Modification

Ac-CWHISKEY-NH_2 modified with 3-methyl-3-buten-1-ol (94)



Synthesised following the general conjugation procedure I using Ac-CWHISKEY-NH_2 (**49**, 2 mg, 1.81 µmol). After analysis the remainder of material (1.745 µmol) was purified using semi-preparative HPLC (10-70% B over 30 minutes) to yield the desired conjugated peptide (1.6 mg, 1.38 µmol, 79% yield).

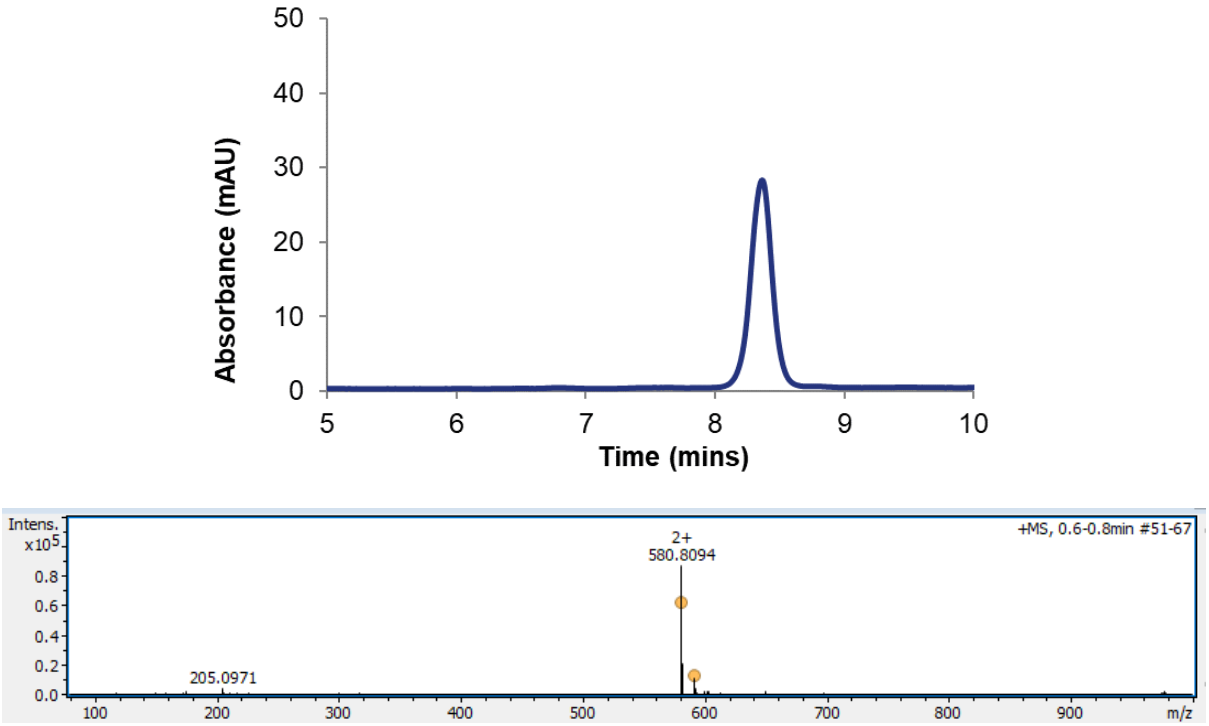
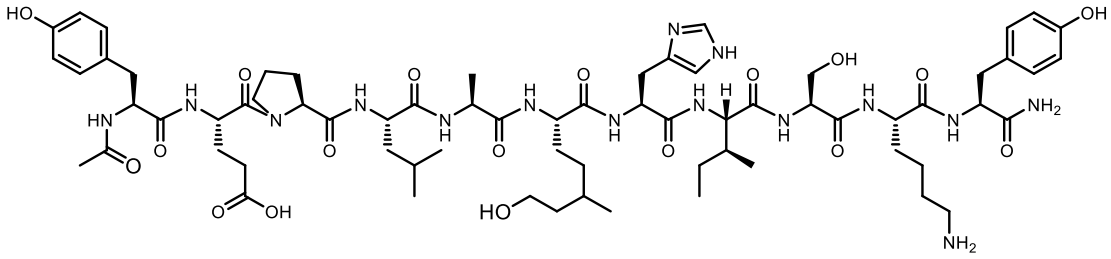


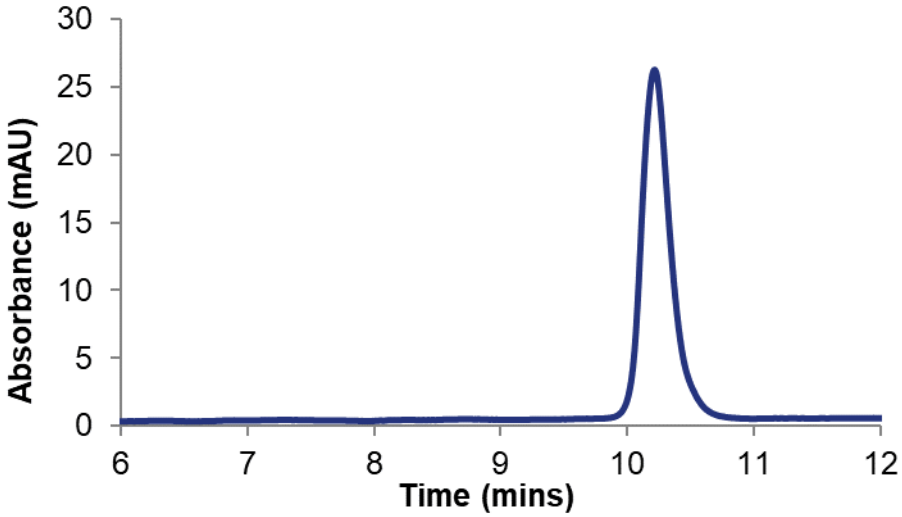
Figure 132 - Analytical HPLC trace of pure Ac-CWHISKEY-NH_2 conjugated with 3-methyl-3-buten-1-ol (**94**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 580.8091. Observed Mass $[M+2H]^{2+}$: 580.8094.

Ac-YEPLACHISKY-NH_2 Modification

Large model example using optimized conditions (95)



Synthesised following the general protocol I Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.47 µmol) and purified using semi-preparative HPLC (15-80% B over 30 minutes) to yield the desired conjugated peptide (1.6 mg, 1.13 µmol, 77% yield).



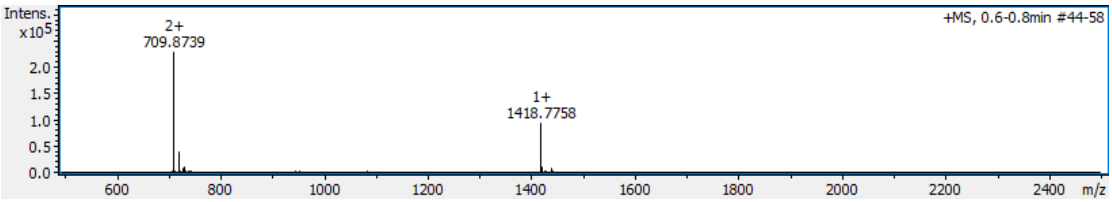
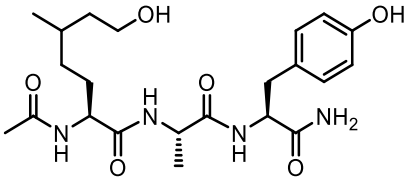


Figure 133 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugated with 3-methyl-3-buten-1-ol (**95**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1418.7684, $[M+2H]^{2+}$: 709.8881. Observed Mass $[M+H]^+$: 1418.7758, $[M+2H]^{2+}$: 709.8739.

Small model examples for NMR analysis

Modification of L amino acid (96)



Synthesised following general protocol I using Ac-CAY-NH_2 (**75**) (4.0 mg, 10.1 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (3.5 mg, 7.80 µmol, 77% yield). ^{1}H NMR (500 MHz, CD_3OD) δ 7.08 (d, J = 8.4 Hz, 2H), 6.71 (d, J = 8.4 Hz, 2H), 4.49 (dd, J = 8.6, 5.8 Hz, 1H), 4.25 (qd, J = 7.2, 1.4 Hz, 1H), 4.20 (ddd, J = 8.3, 5.6, 2.5 Hz, 1H), 3.67 – 3.56 (m, 2H), 3.08 (dd, J = 14.0, 5.7 Hz, 1H), 2.92 (ddd, J = 14.0, 8.6, 1.8 Hz, 1H), 2.03 (d, J = 0.7 Hz, 3H), 1.87 – 1.55 (m, 4H), 1.49 – 1.32 (m, 2H), 1.29 (d, J = 7.2 Hz, 3H), 1.26 – 1.14 (m, 1H), 0.93 (dd, 3H). ^{13}C NMR (126 MHz, CD_3OD) δ 174.7 (C), 173.5 (C), 173.3 (C), 172.3 (C), 155.9 (Ar), 129.9 (Ar), 127.7 (Ar-H), 114.8 (Ar-H), 59.5 (CH), 54.6 (CH), 54.3 (CH), 49.5 (CH), 39.2 (CH_2), 39.0 (CH_2), 36.4 (CH_2), 32.7 (CH_2), 29.1 (CH_3),29.0 (CH_2), 21.1 (CH_2), 18.5 (CH_3), 18.4 (CH_2) 16.2 (CH_3).

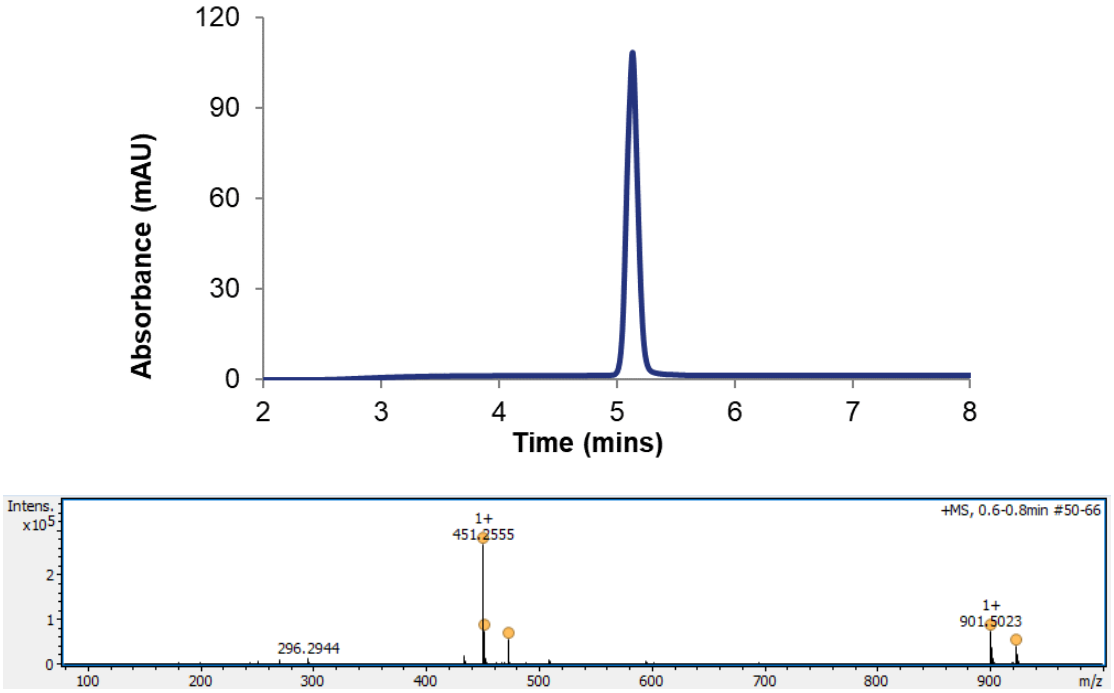
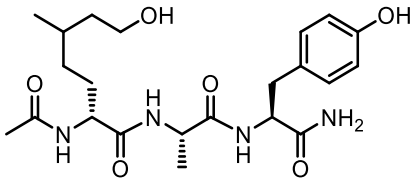


Figure 134 - Analytical HPLC trace of pure Ac-CAY-NH_2 conjugated with 3-methyl-3-buten-1-ol (**96**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 451.2557. Observed Mass $[M+H]^+$: 451.2555.

Modification of Ac-D-CAY-NH_2 with 3-methyl-3-buten-1-ol (97)



Synthesised following general protocol I using Ac-D-CAY-NH_2 (**76**) (4.0 mg, 10.1 µmol) followed by purification using semi-preparative RP-HPLC (15-80% B over 30 minutes) and lyophilisation to produce the desired conjugated peptide (3.3 mg, 7.33 µmol, 73% yield). ^{1}H NMR (500 MHz, CD_3OD) δ 7.97 – 7.89 (m, 1H), 7.17 – 7.10 (m, 2H), 6.71 (d, J = 8.5 Hz, 2H), 4.49 – 4.40 (m, 1H), 4.10 (tt, J = 7.6, 2.3 Hz, 2H), 3.70 – 3.54 (m, 2H), 3.21 (ddd, J = 14.1, 4.3, 1.7 Hz, 1H), 2.94 (dd, J = 14.0, 11.2 Hz, 1H), 2.68 (s, 1H), 2.00 (s, 3H), 1.83 – 1.68 (m, 1H), 1.68 – 1.54 (m, 2H), 1.53 – 1.45 (m, 1H), 1.42 – 1.28 (m, 2H), 1.20 (d, J = 7.3 Hz, 3H), 0.94 (dd, J = 6.5, 3.6 Hz, 3H). ^{13}C NMR (126 MHz, CD_3OD) δ 175.4 (C), 175.4 (C), 173.7 (C), 172.5 (C), 155.8 (Ar), 129.8 (Ar), 128.4 (Ar-H), 114.7 (Ar-H), 59.5 (CH), 59.4 (CH), 55.1 (CH), 54.8 (CH), 50.3 (CH_2), 39.1 (CH_2), 35.8 (CH_2), 32.7 (CH_2), 32.6 (CH_2), 29.1 (CH_2), 29.0 (CH_3), 28.1 (CH_2), 20.9 (CH_2), 18.4 (CH_3), 15.6 (CH_3).

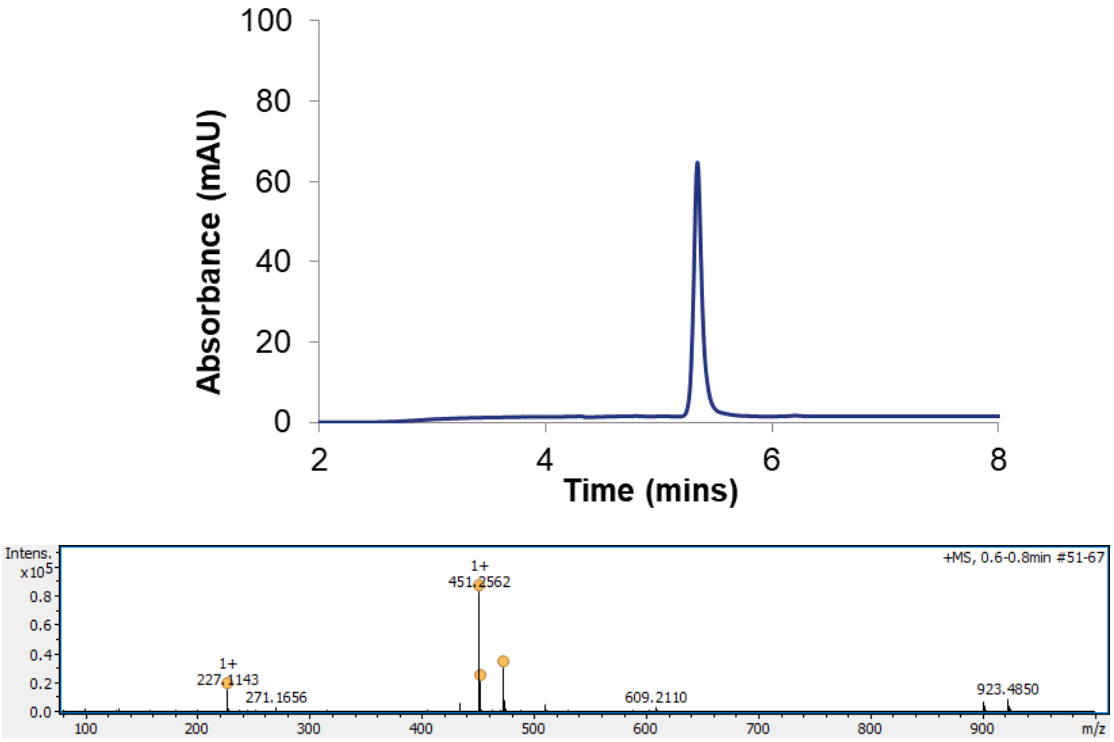
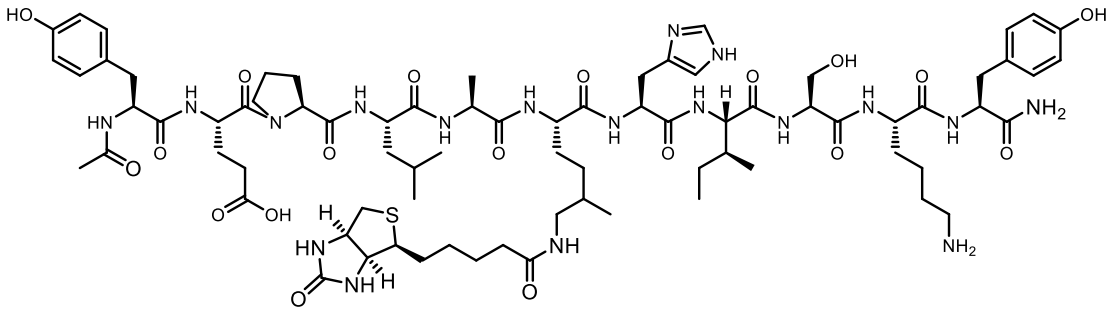


Figure 135 - Analytical HPLC trace of pure Ac-D-CAY-NH2 conjugated with 3-methyl-3-buten-1-ol (**97**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass [M+H]+: 451.2557. Observed Mass [M+H]+: 451.2562.

Biotin-prop (103)



Synthesised following the general protocol I using Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.466 µmol). After analysis the remainder of material (1.365 µmol) was purified using semi-preparative HPLC (15-80% B over 30 minutes) to yield the desired conjugated peptide (1.4 mg, 0.858 µmol, 63% yield).

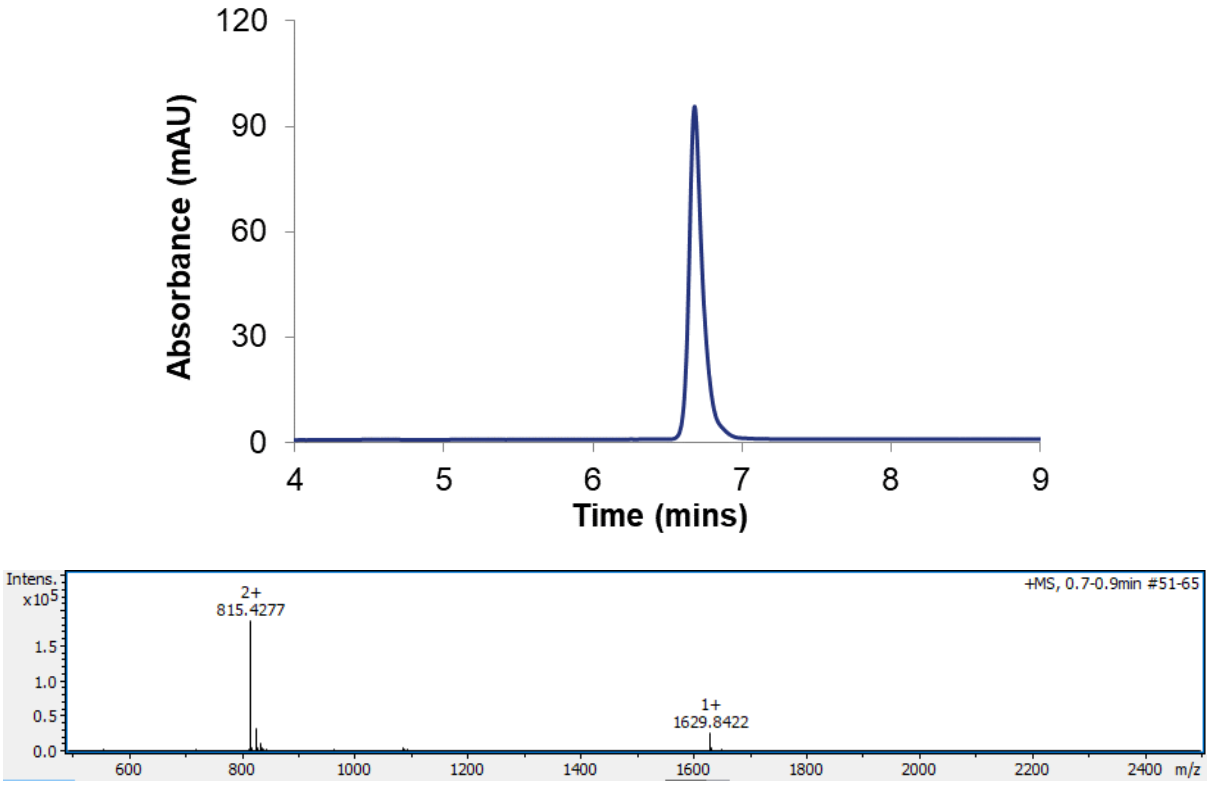
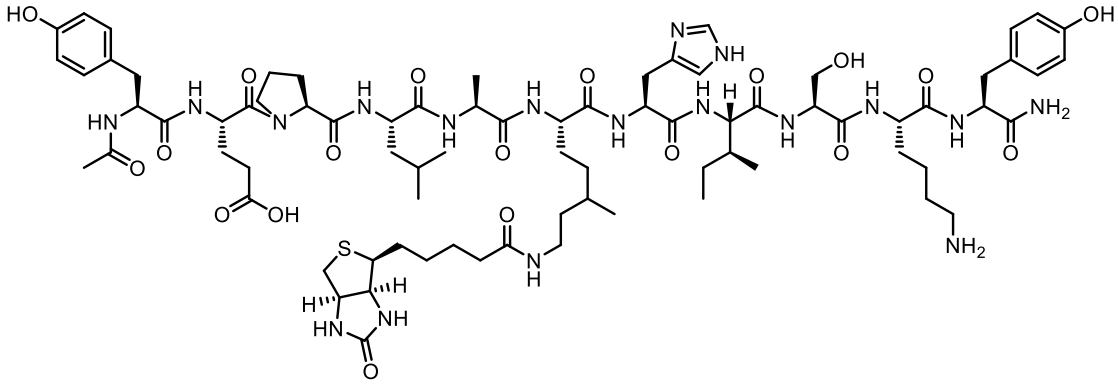


Figure 136 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugated with prop-biotin (**103**). Analytical gradient 15-45% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1629.8463, $[M+2H]^{2+}$: 815.4271. Observed Mass $[M+H]^+$: 1629.8422, $[M+2H]^{2+}$: 815.4277.

Biotin-bute (104)



Synthesised following the general protocol I using Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.466 µmol). After analysis the remainder of material (1.365 µmol) was purified using semi-preparative HPLC (15-80% B over 30 minutes) to yield the desired conjugated peptide (1.8 mg, 1.094 µmol, 80% yield).

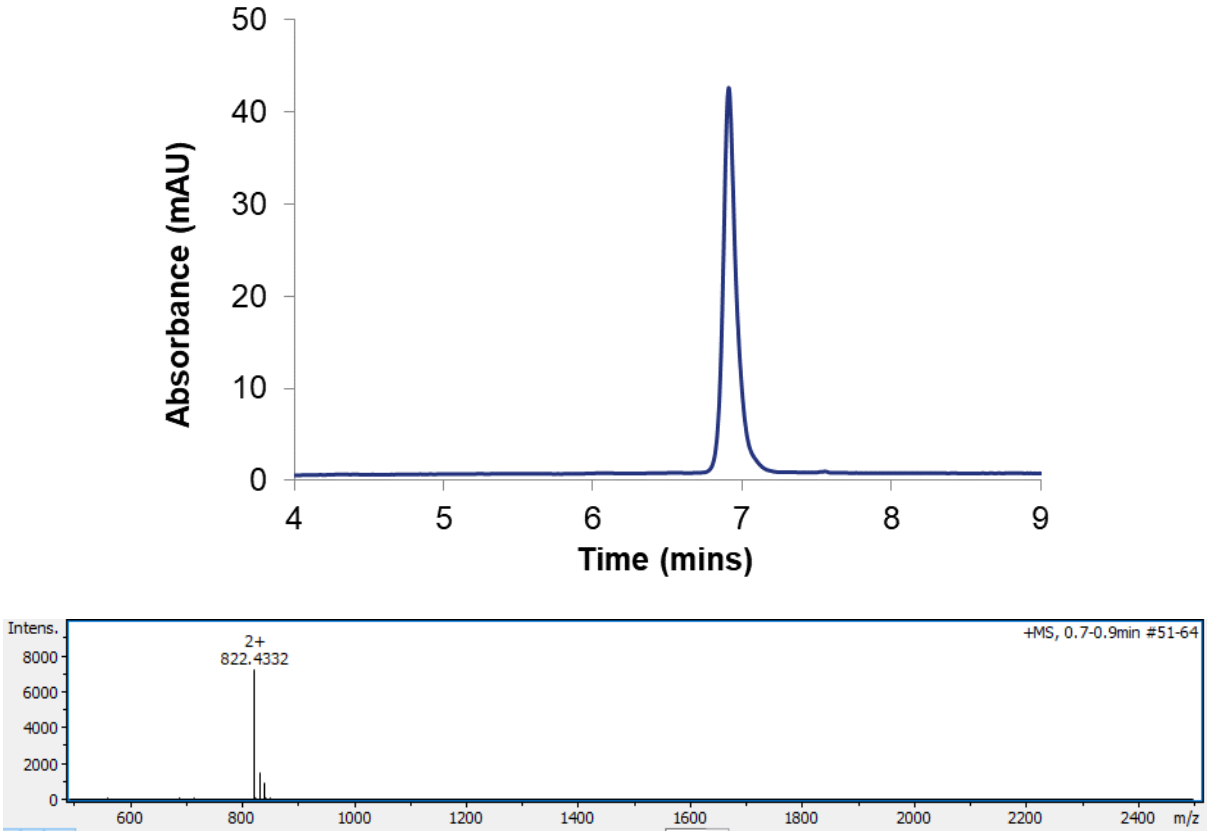
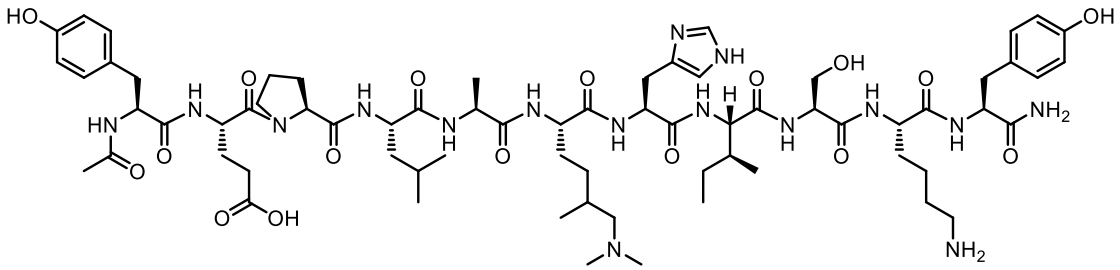
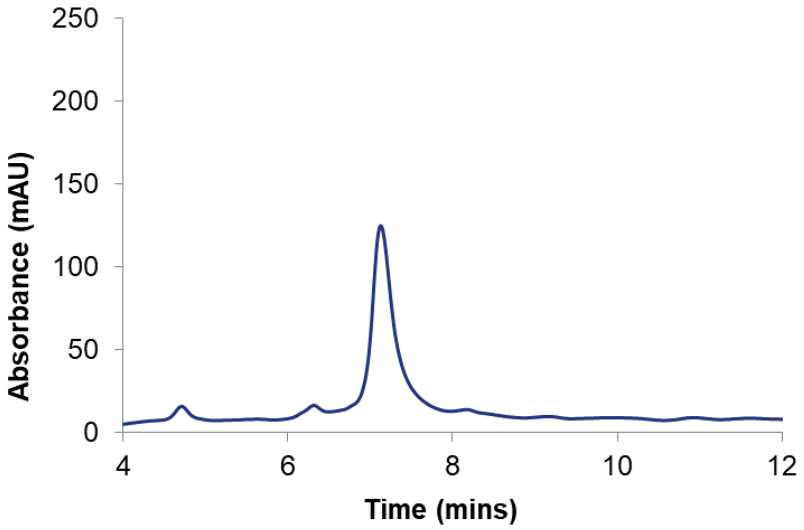


Figure 137 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugated with bute-biotin (**104**). Analytical gradient 15-45% B over 10 min, 210 nm. Calculated Mass $[M+2H]^{2+}$: 822.4349. Observed Mass $[M+2H]^{2+}$: 822.4332.

Ac-YEPLACHISKY-NH_2 modified with N,N,2-trimethylprop-2-en-1-amine (105)



Synthesised following the general protocol I using Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.466 µmol). After analysis, the remainder of material (1.365 µmol) was purified using semi-preparative HPLC (5-60% B over 30 minutes) to yield the desired conjugated peptide (1.3 mg, 0.908 µmol, 67% yield).



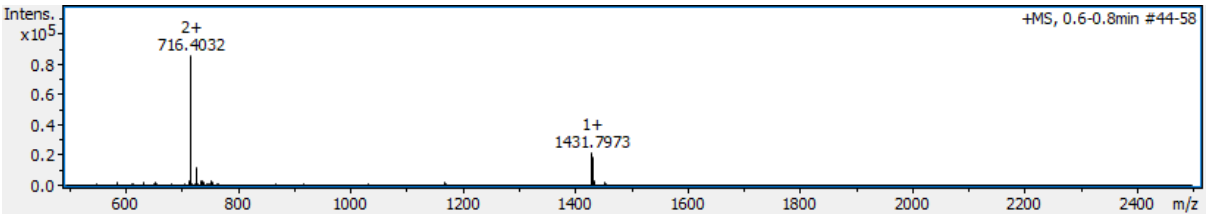
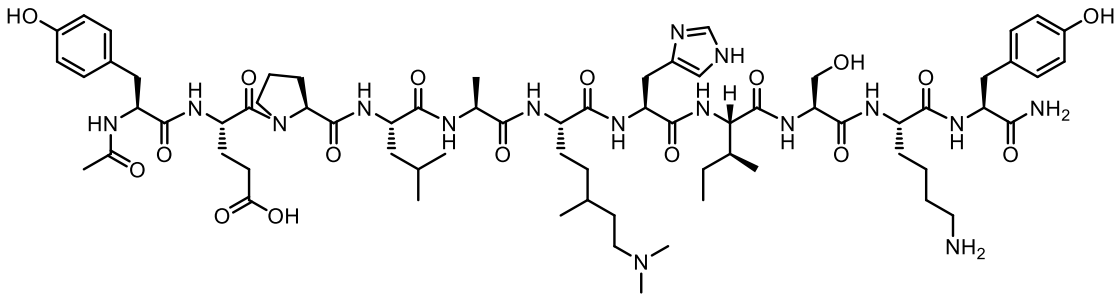


Figure 138 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 modified with N,N,2-trimethylprop-2-en-1-amine (**105**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1431.8000, $[M+2H]^{2+}$: 716.4039. Observed Mass $[M+H]^+$: 1431.7973, $[M+2H]^{2+}$: 716.4032.

Ac-YEPLACHISKY-NH_2 modified with N,N-3-trimethylbut-3-en-1-amine (106)



Synthesised following the general protocol I using Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.466 µmol). After analysis the remainder of material (1.365 µmol) was purified using semi-preparative HPLC (5-60% B over 30 minutes) to yield the desired conjugated peptide (1.6 mg, 1.11 µmol, 81% yield).

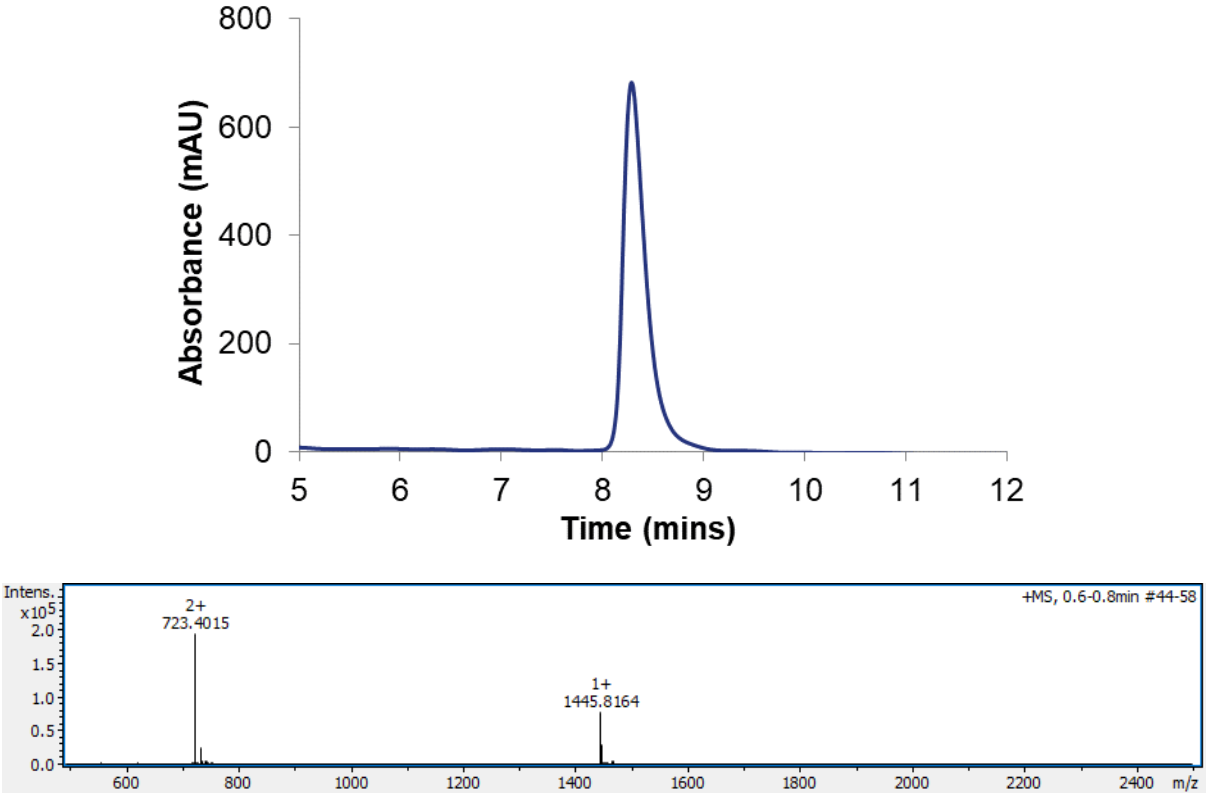
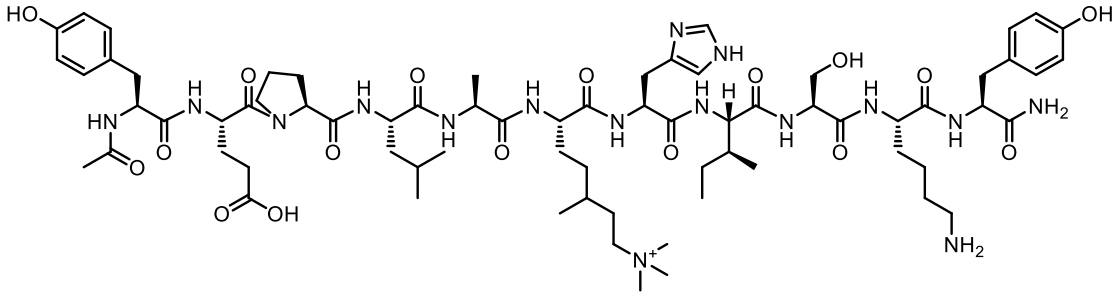


Figure 139 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 modified with N,N-3-trimethylbut-3-en-1-amine (**106**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1445.8157, $[M+2H]^{2+}$: 723.4118. Observed Mass $[M+H]^+$: 1445.8164, $[M+2H]^{2+}$: 723.4015.

Ac-YEPLACHISKY-NH_2 modified with N,N,N-3-tetramethylbut-3-en-1-aminium iodide (107)



Synthesised following the general protocol I using Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.466 µmol). After analysis the remainder of material (1.365 µmol) was purified using semi-preparative HPLC (5-60% B over 30 minutes) to yield the desired conjugated peptide (1.7 mg, 1.07 µmol, 78% yield).

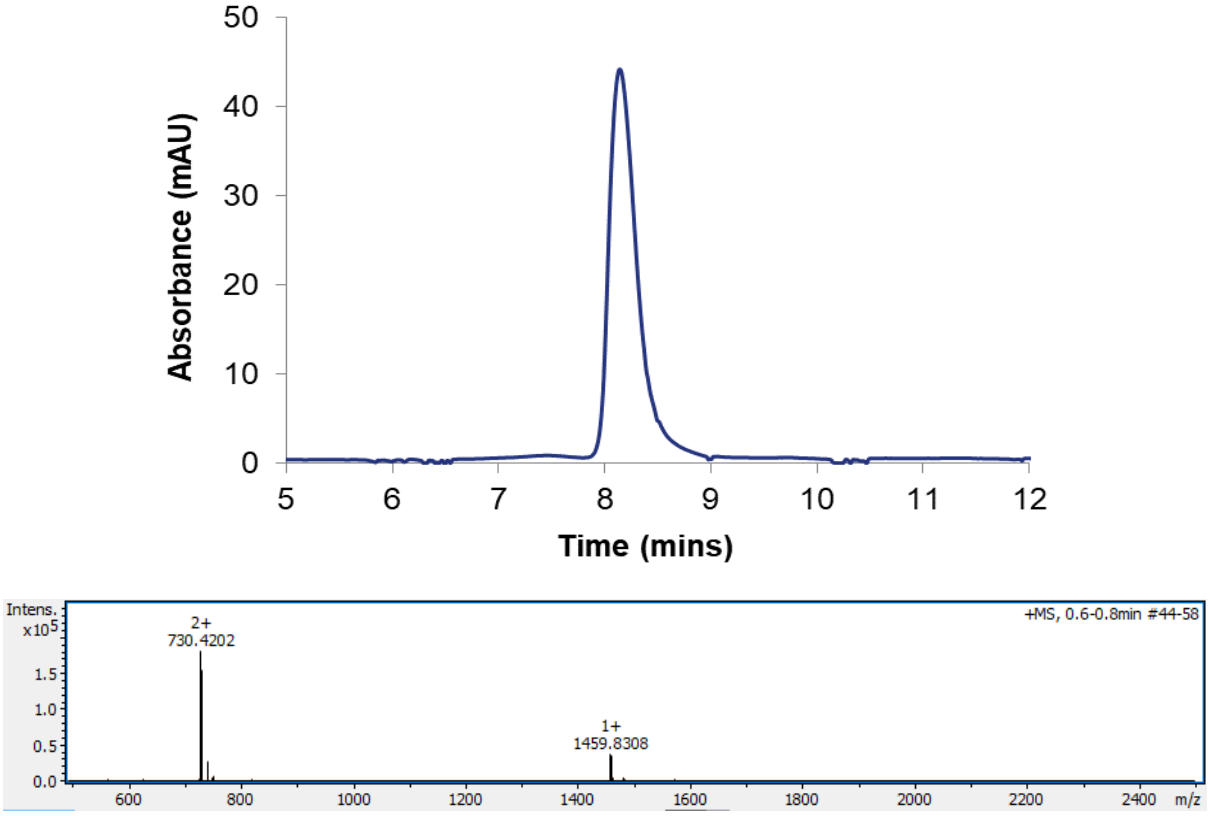
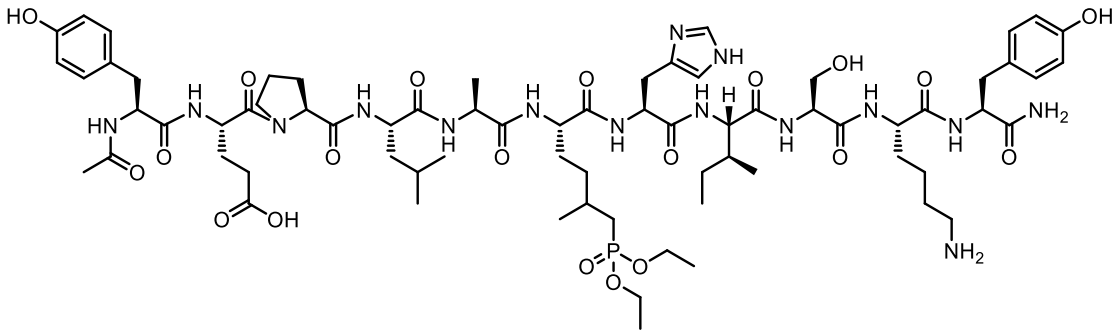
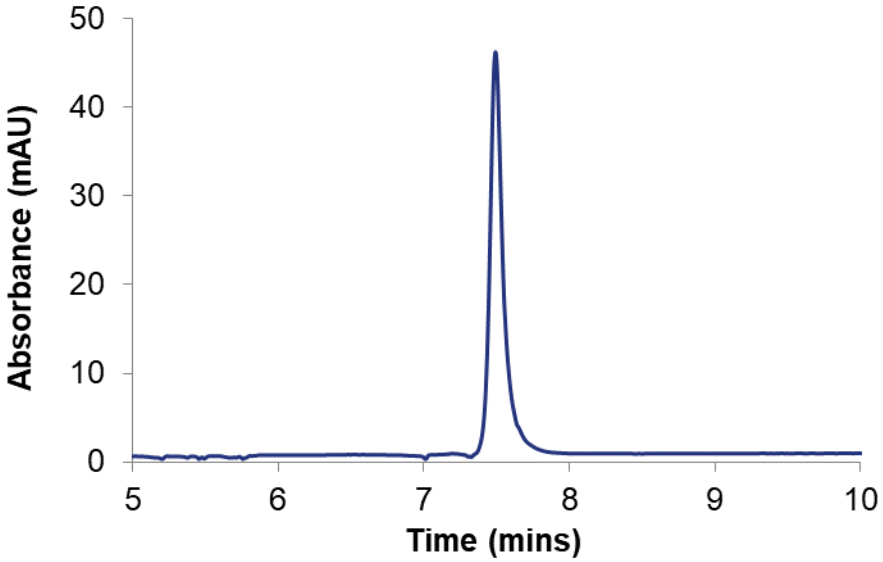


Figure 140 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 modified with N,N,N-3-tetramethylbut-3-en-1-aminium iodide (**107**). Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M]^+$: 1459.8313, $[M+H]^{2+}$: 730.4196. Observed Mass $[M]^+$: 1459.8308, $[M+H]^{2+}$: 730.4202.

Ac-YEPLACHISKY-NH_2 modified with diethyl(2-methylallyl)phosphonate (108)



Synthesised following the general protocol I using Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.466 µmol). After analysis the remainder of material (1.365 µmol) was purified using semi-preparative HPLC (5-60% B over 30 minutes) to yield the desired conjugated peptide (1.7 mg, 1.11 µmol, 81% yield).



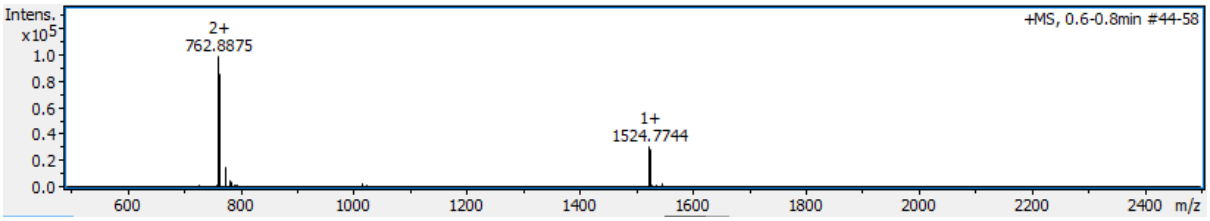
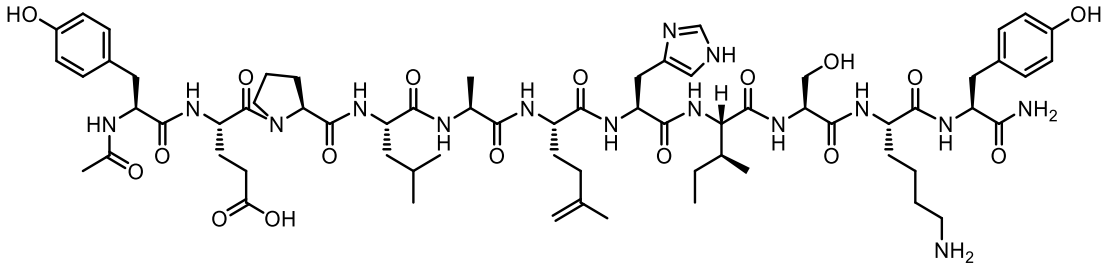


Figure 141 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 modified with diethyl(2-methylallyl)phosphonate (**108**). Analytical gradient 15-45% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1524.7867, $[M+2H]^{2+}$: 762.8973. Observed Mass $[M+H]^+$: 1524.7744, $[M+2H]^{2+}$: 762.8875.

Ac-YEPLACHISKY-NH_2 modified with 2-methyl-2-propene-1-sulfonic acid – elimation product (112)



Synthesised following the general protocol I using Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.466 µmol). After analysis the remainder of material (1.365 µmol) was purified using semi-preparative HPLC (5-60% B over 30 minutes) to yield the desired conjugated peptide (1.5 mg, 1.08 µmol, 79% yield).

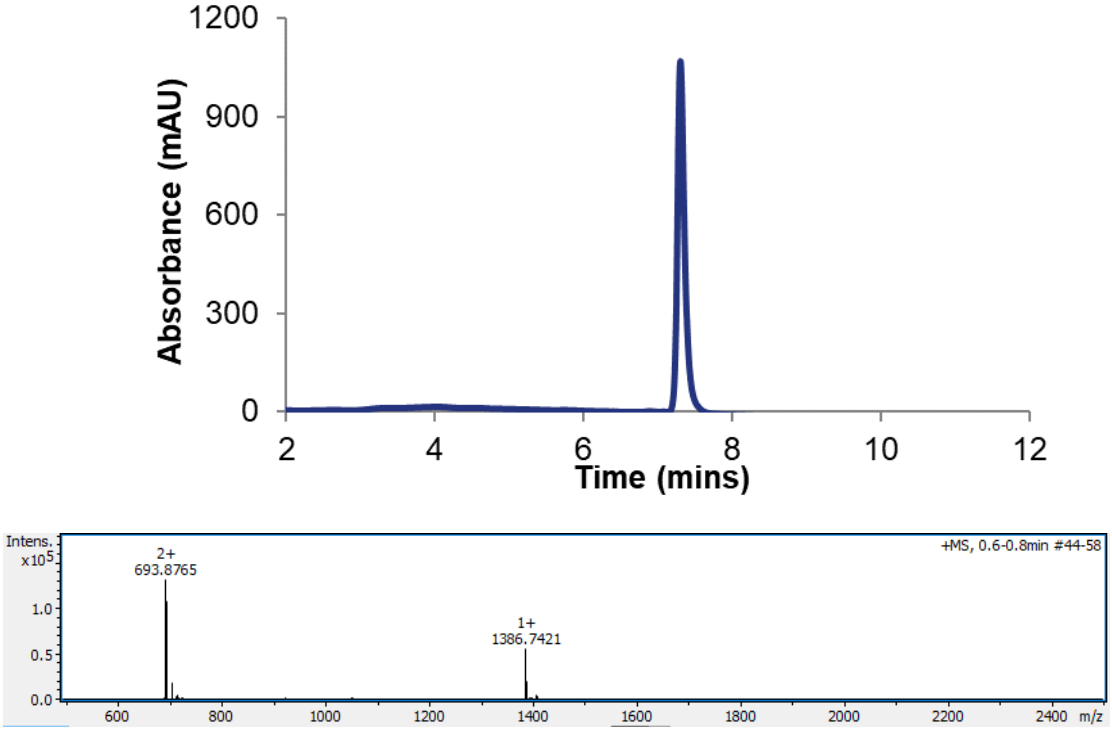
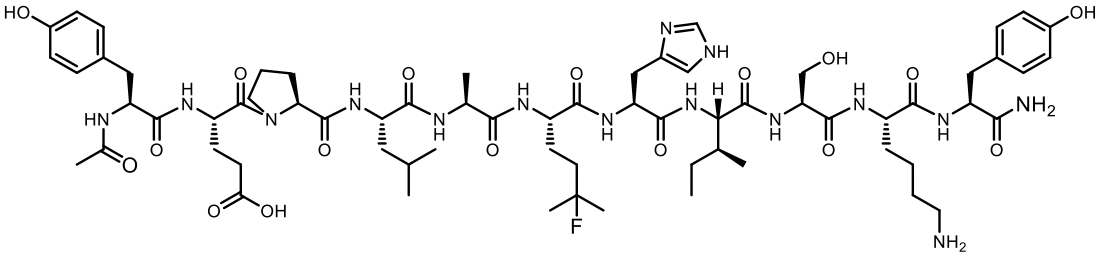


Figure 142 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugated with 2-methyl-2-propene-1-sulfonic acid (**112**). Analytical gradient 15-45% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1386.7422, $[M+2H]^{2+}$: 693.8750. Observed Mass $[M+H]^+$: 1386.7421, $[M+2H]^{2+}$: 693.8765.

Step 2 - Hydrofluorination of elimination product (113)



The following solutions were first prepared: 0.05 M solution of iron (III) oxalate (2.4 mg per 0.1 ml H_2O, degassed with Ar), 0.5 M Selectfluor (17.7 mg per 0.1 ml H_2O) and 0.5 M $NaBH_4$ (1.9 mg per 0.1 ml MeCN). The final concentrations of the initial reaction mixture were 0.0125 M peptide (**112**), 0.025 M iron (III) oxalate, 0.025 M Selectfluor and 0.04 M $NaBH_4$. To modified Ac-YEPLAXHISKY-NH (**112**, 1 mg, 0.722 µmol) in acetonitrile (0.024 ml) was added iron (III) oxalate (0.029 ml, 0.05 M stock, 2 eq.) and Selectfluor (0.003 ml, 0.5 M stock, 2 eq.) and the solution briefly degassed with Ar. The reaction mixture was placed on ice and $NaBH_4$ (0.005 ml, 0.5 M stock, 1.6 eq.) was added. The reaction mixture was shaken at 0 °C for 2 minutes and a further quantity of $NaBH_4$ (0.005 ml, 0.5 M stock, 2 eq.) was added. The reaction mixture was shaken on ice for a further 30 minutes, after which time the reaction was analysed by HPLC and purified by semi-preparative HPLC (5-60% B over 30 minutes). The fractions containing the desired peptide were lyophilised to give the title compound as a fluffy white solid (0.7 mg, 0.498 µmol, 69% yield).

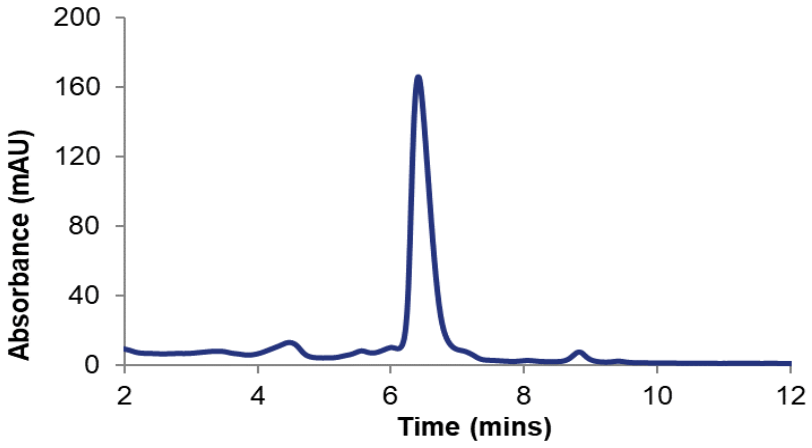
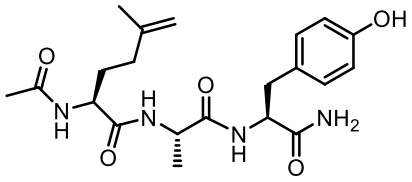


Figure 143 – Analytical HPLC trace of pure Ac-YEPLAXHISKY-NH_2 hydrofluorinated (**113**). Analytical gradient 20-30% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1406.7484, $[M+2H]^{2+}$:703.8781. Observed Mass $[M+H]^+$: 1386.7421, $[M+2H]^{2+}$: 703.8770.

Ac-CAY-NH_2 with 2-methylprop-2-ene-2-sulfonic acid followed by hydrofluorination

Ac-CAY-NH_2 modified for hydrofluorination – Part 1 (114)



Synthesised following the general protocol I Ac-CAY-NH_2 (**75**, 5 mg, 12.6 µmol) and purified using semi-preparative HPLC (5-60% B over 30 minutes) to yield the desired conjugated peptide (4.1 mg, 9.75 µmol, 77% yield). ^{1}H NMR (500 MHz, CD_3OD) δ 7.07 (d, *J* = 8.5 Hz, 2H), 6.71 (d, *J* = 8.5 Hz, 2H), 4.80 – 4.72 (m, 2H), 4.49 (dd, *J* = 8.6, 5.7 Hz, 1H), 4.29 – 4.15 (m, 2H), 3.08 (dd, *J* = 14.1, 5.6 Hz, 1H), 2.91 (dd, *J* = 14.0, 8.6 Hz, 1H), 2.11 (dt, *J* = 9.4, 6.3 Hz, 2H), 2.03 (s, 3H), 1.97 – 1.88 (m, 1H), 1.83 – 1.78 (m, 1H), 1.76 (t, *J* = 1.1 Hz, 3H), 1.28 (d, *J* = 7.2 Hz, 3H). ^{13}C NMR (126 MHz, CD_3OD) δ 174.7, 173.4, 173.3, 172.4, 155.9, 144.3, 129.9, 127.7, 114.8, 109.9, 54.6, 53.7, 49.5, 36.4, 33.4, 29.5, 21.2, 21.1, 16.2.

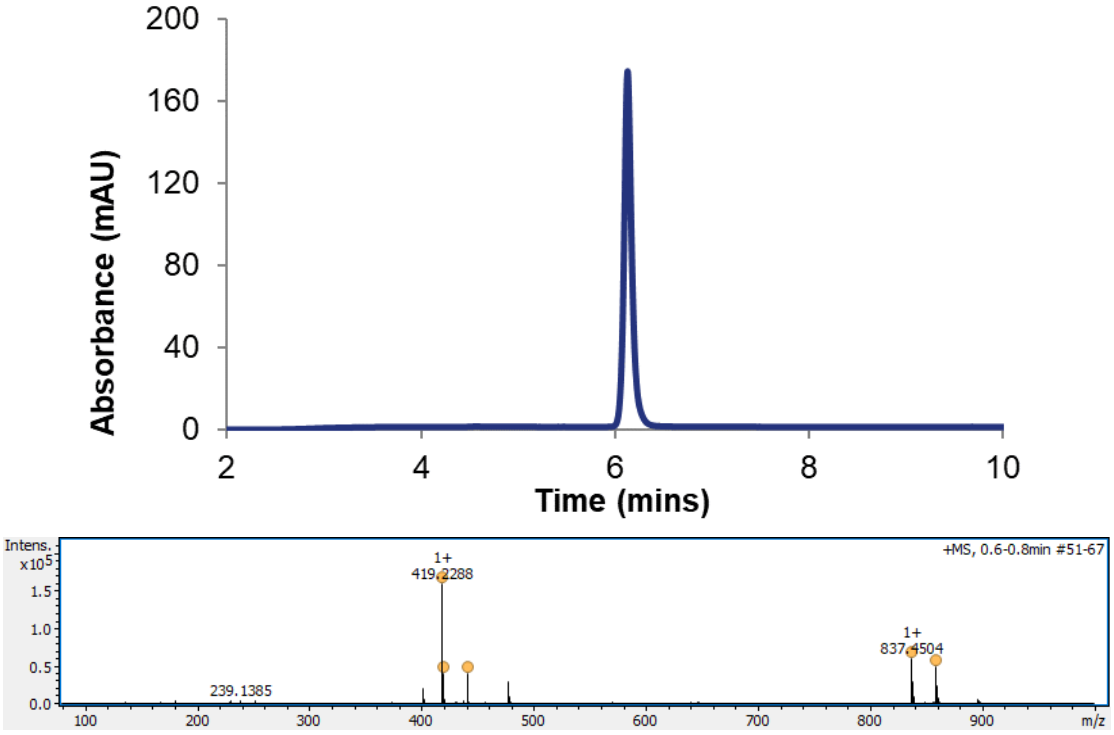
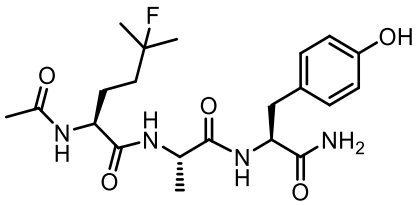


Figure 144 - Analytical HPLC trace of pure Ac-CAY-NH_2 conjugated with 2-Methyl-2-propene-1-sulfonic acid sodium salt (**114**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 419.2294. Observed Mass $[M+H]^+$: 419.2288.

Ac-CAY-NH_2 modified for hydrofluorination – Part 2 (115)



Hydrofluorination was performed following a literature method.[171] The following solutions were first prepared: 0.05 M solution of iron (III) oxalate (2.4 mg per 0.1 ml H_2O, degassed with Ar), 0.5 M Selectfluor (17.7 mg per 0.1 ml H_2O) and 0.5 M $NaBH_4$ (1.9 mg per 0.1 ml MeCN). The final concentrations of the initial reaction mixture were 0.0125 M peptide, 0.025 M iron (III) oxalate, 0.025 M Selectfluor and 0.04 M $NaBH_4$. To previously modified peptide (**114**) (2 mg, 4.78 µmol) was added iron (III) oxalate (2 eq.) and Selectfluor (2 eq.) and the solution briefly degassed with Ar. The reaction mixture was placed on ice and $NaBH_4$ (1.6 eq.) was added. The reaction mixture was shaken at 0 ^oC for 2 minutes and a further quantity of $NaBH_4$ (2 eq.) was added. The reaction mixture was shaken on ice for a further 30 minutes, after which time the reaction was analysed by HPLC and purified by semi-preparative HPLC (5-60% B over 30 minutes) to yield the title compound as a fluffy white solid (1.5 mg, 3.42 µmol, 72% yield). ^1H NMR (500 MHz, CD_3OD) δ 7.08 (d, J = 8.5 Hz, 2H), 6.71 (d, J = 8.5 Hz, 2H), 4.48 (dd, J = 8.5, 5.8 Hz, 1H), 4.24 (dd, J = 8.2, 6.2 Hz, 2H), 3.08 (dd, J = 14.0, 5.8 Hz, 1H), 2.91 (dd, J = 14.0, 8.5 Hz, 1H), 2.03 (s, 3H), 1.95 – 1.84 (m, 1H), 1.81 – 1.61 (m, 3H), 1.37 (d, J = 1.8 Hz, 3H), 1.33 (d, J = 1.8 Hz, 3H), 1.29 (d, J = 7.2 Hz, 3H).^{13}C NMR (126 MHz, CD_3OD) δ 174.7 (C), 173.2 (C), 173.1 (C), 172.3 (C), 155.9 (Ar), 129.9 (Ar), 127.7 (Ar-H), 114.78 (Ar-H), 93.4 (C), 54.6 (CH), 54.0 (CH), 49.5 (CH), 37.1 (CH), 36.9 (CH_2), 36.4 (CH_2), 29.3 (CH_3), 26.0 (CH_2), 25.4 (CH_3), 21.0 (CH_3), 16.2 (CH_3). ^{19}F NMR (376 MHz, CD_3OD) δ -140.15.

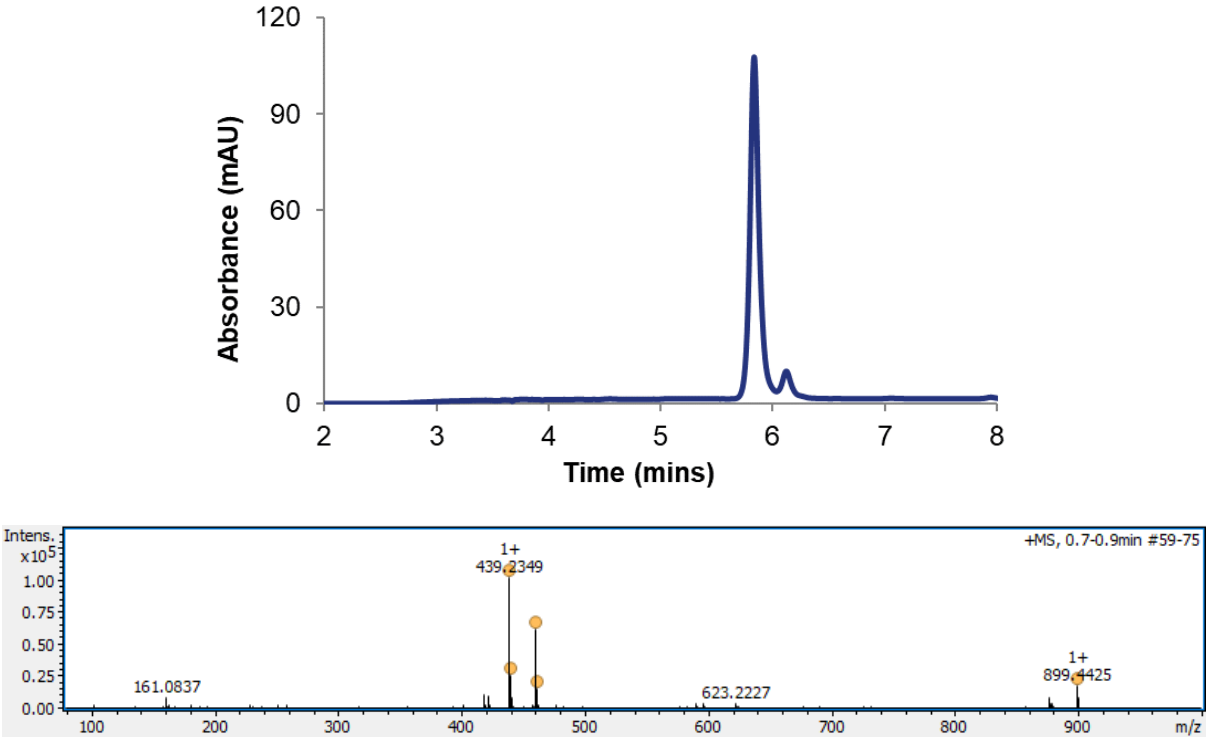
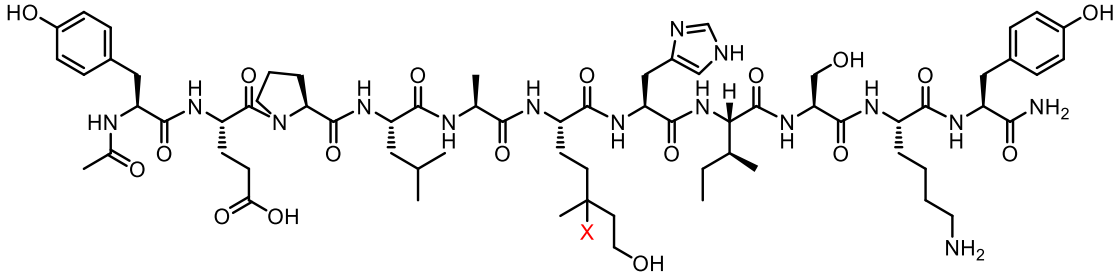


Figure 145 - Analytical HPLC trace of pure Ac-CAY-NH_2 modified with 2-Methyl-2-propene-1-sulfonic acid sodium salt, then hydrofluorinated (**115**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass $[M+H]^+$: 439.2356. Observed Mass $[M+H]^+$: 439.2349.

Novel fluorination method (116 and 95)



X = F - **116**
X = H - **95**

Separate stock solutions of Selectfluor (0.5 M) and $AgNO_3$ (0.05 M) were prepared, then mixed in a 1:1 ratio for pre-activation. To Ac-YEPLACHISKY-NH_2 (**56**, 2 mg, 1.47 µmol) dissolved in 10% acetonitrile (MeCN) in ligation buffer to a concentration of 1 mM was added a solution of TCEP (0.5 M stock solution in LB, pH adjusted to 7, 10 eq.), 3-methyl-3-butene-1-ol (pure, 200 eq.) and (Ir[dF(CF3)ppy]2(dtbpy))PF6 (1 mM stock solution in MeCN, 0.05 eq.). The pH of the reaction mixture was checked to be around 8 (7.8-8.2), this was adjusted as required by addition of either 10 M NaOH or 6 M HCl (care must be taken to not allow the pH to rise above 9). The pre-activated solution of Selectfluor and $AgNO_3$ (0.012 ml, 2 eq. Selectfluor and 0.2 eq. $AgNO_3$) was added to the reaction mixture and the pH checked to still be between 7.8 and 8.2. The reaction vessel was then placed into blue LEDs and left for 10 minutes. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified (1.35 µmol remaining after analysis) by semi-preparative HPLC (15-80% B over 30 mins) and fractions containing the desired peptide lyophilised to yield a mixture of fluorinated (**116**, 0.4 mg, 0.278 µmol, 21% yield) and non-fluorinated (**95**, 1.3 mg, 0.917 µmol, 68% yield) products as fluffy white solids.

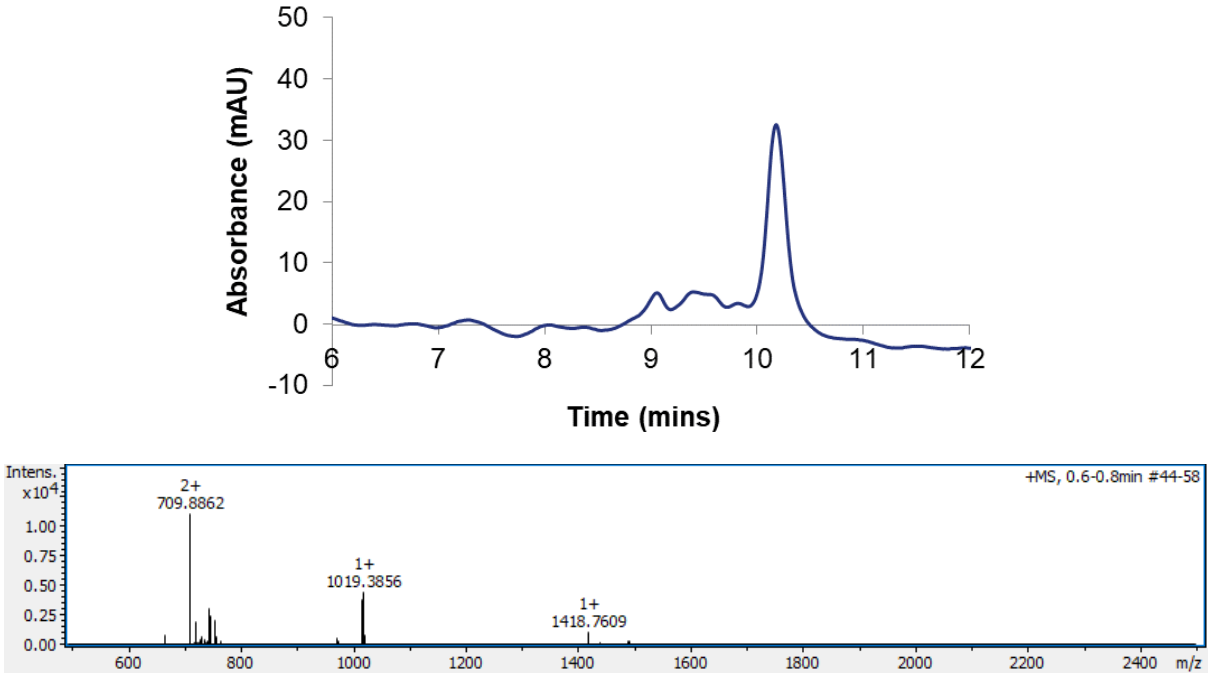


Figure 146 – Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugation/fluorination, non-fluorinated product (**95**). Peak at 1019 is unrelated to this sample – MS impurity. Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1418.7684, $[M+2H]^{2+}$: 709.8881. Observed Mass $[M+H]^+$: 1418.7609, $[M+2H]^{2+}$: 709.8862.

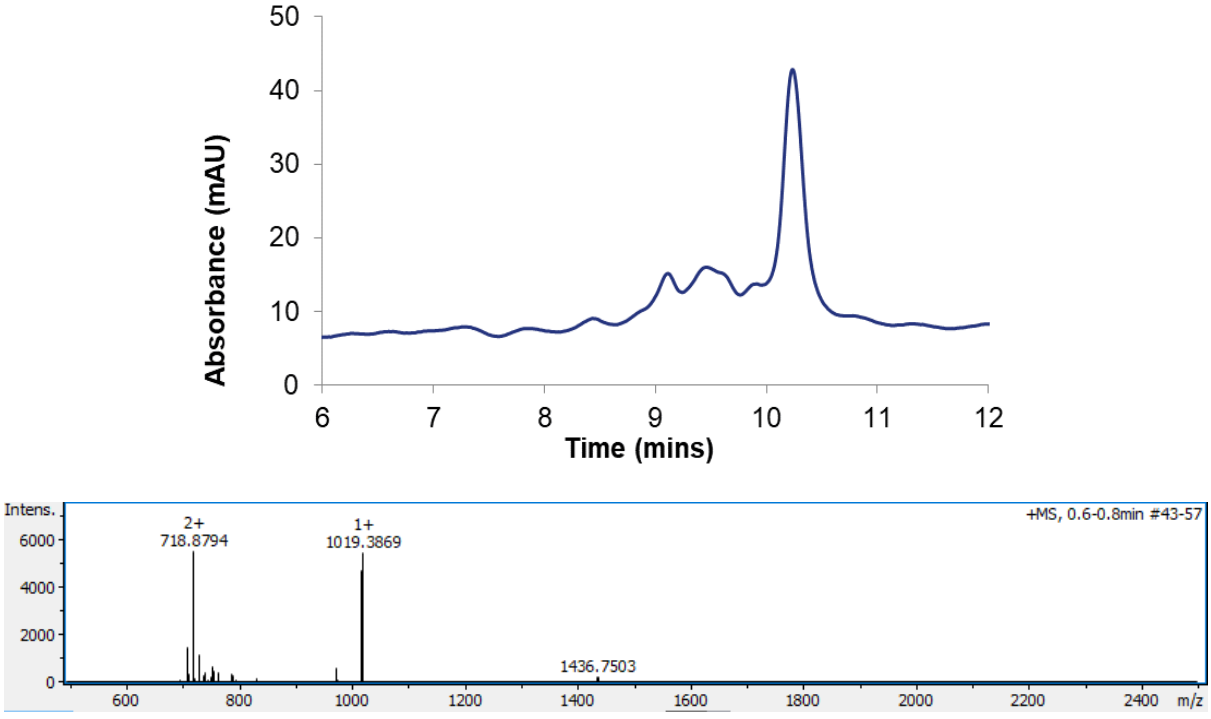
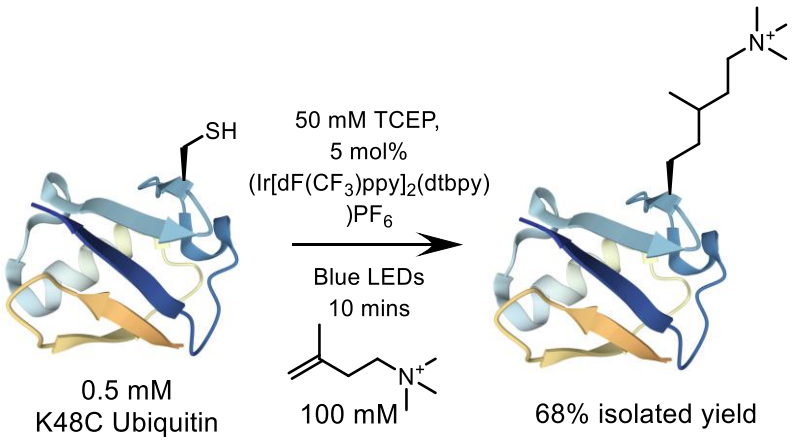


Figure 147 - Analytical HPLC trace of pure Ac-YEPLACHISKY-NH_2 conjugation/fluorination, fluorinated product (**116**). Peak at 1019 is unrelated to this sample – MS impurity. Analytical gradient 15-25% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 1436.7503, $[M+2H]^{2+}$: 718.8794. Observed Mass $[M+H]^+$: 1418.7609, $[M+2H]^{2+}$: 718.8834.

Protein bioconjugation using stabilised alkene probes (117 and 118)

K48C bioconjugation with N,N,N-3-tetramethylbut-3-en-1-aminium iodide (118)



Synthesised following the general protocol J using K48C Ubiquitin (**60**, 2 mg, 0.234 µmol). After analysis the remainder of material (0.222 µmol) was purified using semi-preparative HPLC (25-40% B over 30 minutes) to yield the desired conjugated peptide (1.3 mg, 0.151 µmol, 68% yield).

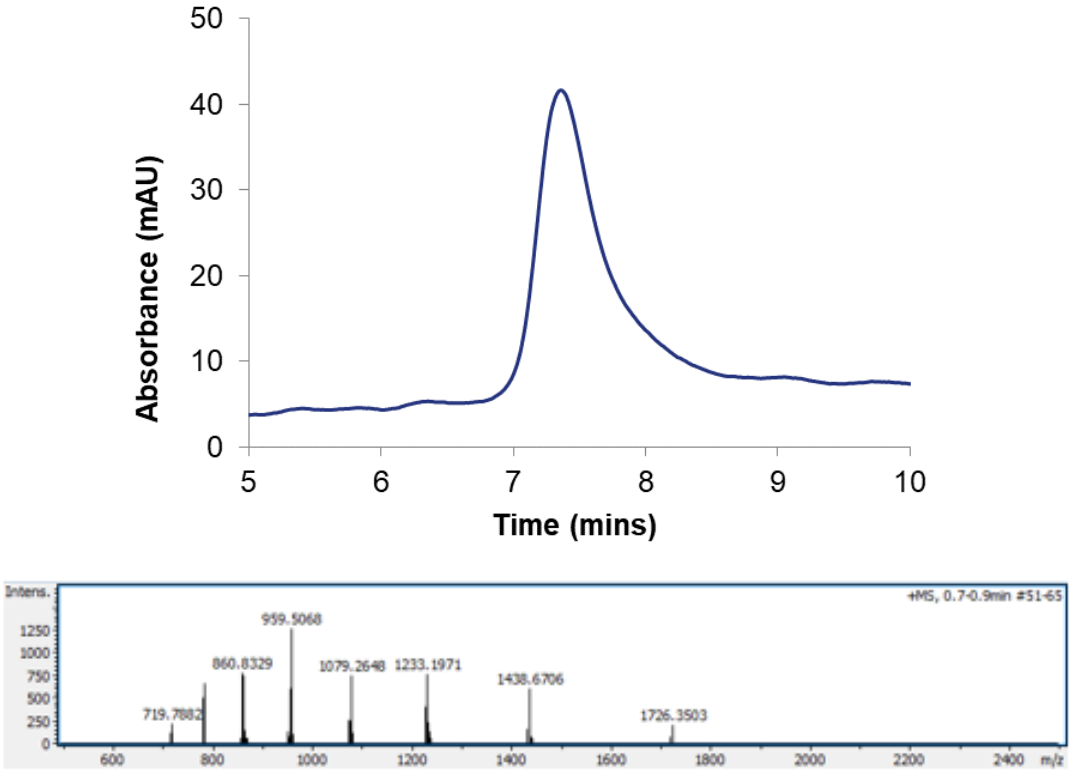
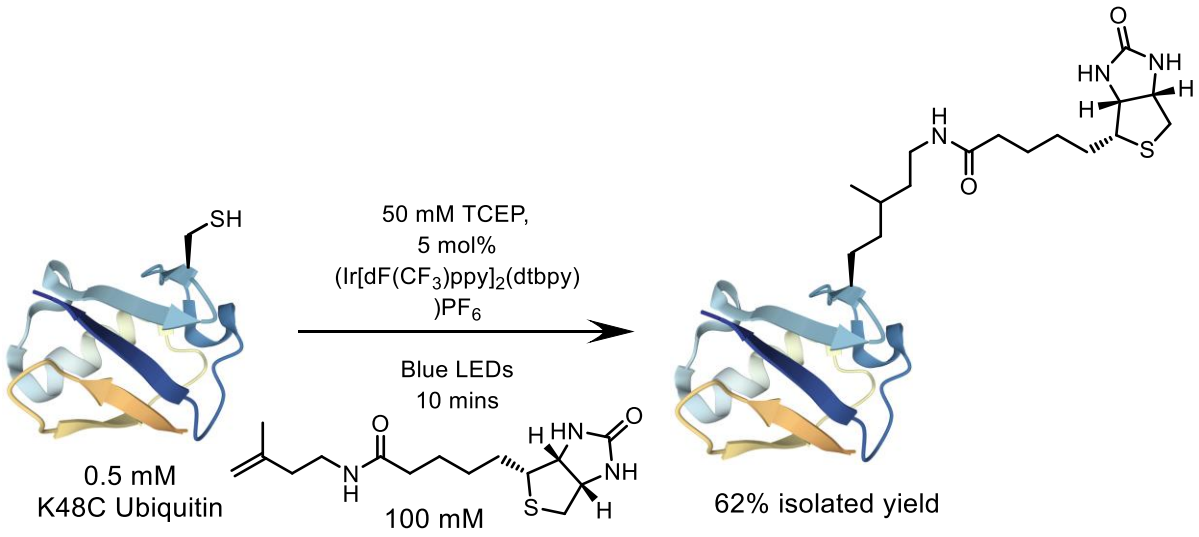


Figure 148 - Analytical HPLC trace of pure K48C ubiquitin conjugated with N,N,N-3-tetramethylbut-3-en-1-aminium iodide (**118**). Analytical gradient 25-35% B over 10 min, 210 nm. Calculated Mass $[M+4H]^{5+}$: 1725.7417, $[M+5H]^{6+}$: 1438.2860, $[M+6H]^{7+}$: 1232.9606, $[M+7H]^{8+}$: 1078.9665, $[M+8H]^{9+}$: 959.1933, $[M+9H]^{10+}$: 863.2740. Observed Mass $[M+4H]^{5+}$: 1726.3903, $[M+5H]^{6+}$: 1438.6706, $[M+6H]^{7+}$: 1233.1971, $[M+7H]^{8+}$: 1079.2648, $[M+8H]^{9+}$: 959.5068, $[M+9H]^{10+}$: 860.8329.

K48C bioconjugation with biotin modified 3-methyl-3-butene-1-amine (117)



Synthesised following the general conjugation procedure J using K48C Ubiquitin (**60**, 2 mg, 0.234 µmol). After analysis the remainder of material (0.222 µmol) was purified using semi-preparative HPLC (25-40% B over 30 minutes) to yield the desired conjugated peptide (1.2 mg, 0.136 µmol, 62% yield).

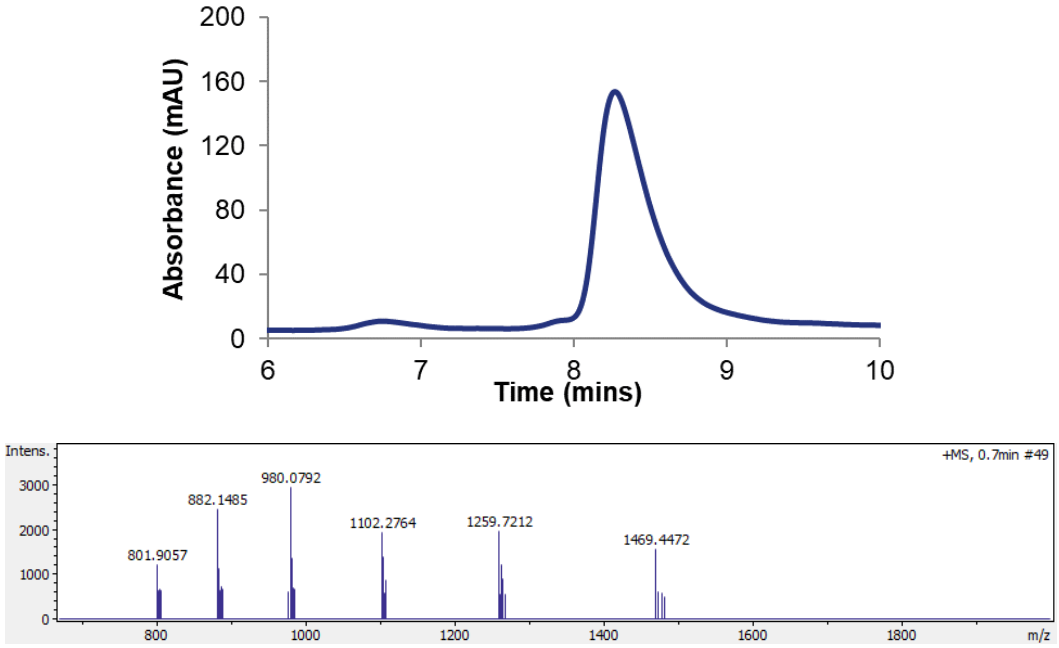
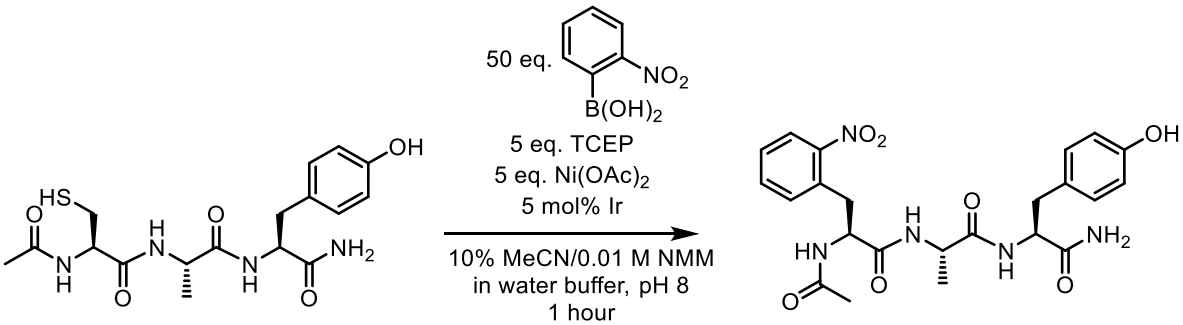


Figure 149 - Analytical HPLC trace of pure K48C ubiquitin conjugated with biotin-butene (**117**). Analytical gradient 25-35% B over 10 min, 210 nm. Calculated Mass $[M+6H]^{6+}$: 1468.6219, $[M+7H]^{7+}$: 1258.9627, $[M+8H]^{8+}$: 1101.7184, $[M+9H]^{9+}$: 979.4172, $[M+10H]^{10+}$: 881.5762, $[M+11H]^{11+}$: 801.5246. Observed Mass $[M+6H]^{6+}$: 1469.4472, $[M+7H]^{7+}$: 1259.7212, $[M+8H]^{8+}$: 1102.2764, $[M+9H]^{9+}$: 980.0792, $[M+10H]^{10+}$: 882.1485, $[M+11H]^{11+}$: 801.9057.

7.8 Desulfurative Cross-coupling

Ac-CAY-NH_2 modified with 2-nitrophenylboronic acid (120)



61, 0.5 mM

100, 39% isolated yield

To Ac-CAY-NH_2 (**75**, 4 mg, 10.1 µmol) dissolved in 10% acetonitrile (MeCN) in 0.01 M NMM in H_2O to a concentration of 1 mM was added a solution of TCEP (0.5 M stock solution in H_2O, pH adjusted to 7, 5 eq.) and $(Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$ (1 mM stock solution in MeCN, 0.05 eq.). The reaction mixture was diluted to a peptide concentration of 0.5 mM and the pH of the solution was adjusted to 7.8-8.2. Quickly, $Ni(OAc)_2$ (0.5 M stock solution in H_2O, 5 eq.) followed by 2-nitrophenylboronic acid (50 eq.) were added. The pH of the reaction mixture was checked to be around 8 (7.8-8.2) and the reaction vessel was then placed into blue LEDs and left for 1 hour. The setup must be quick to avoid the thiol-cross-coupling occurring before desulfurisation occurs. Once the starting material was shown to be fully consumed by HPLC the reaction mixture was purified by semi-preparative HPLC (5-60 over 30 mins) to yield the title compound as an off-white fluffy solid (1.9 mg, 3.91 µmol, 39% yield). ^{1}H NMR (500 MHz, CD_3OD) δ 7.21 (d, J = 8.5 Hz, 2H), 7.11 – 6.96 (m, 3H), 6.70 (d, J = 2.6 Hz, 1H), 4.74 (dd, J = 7.7, 4.5 Hz, 1H), 4.49 – 4.30 (m, 2H), 3.21 – 3.14 (m, 1H), 3.05 (dd, J = 14.2, 7.7 Hz, 1H), 2.94 (dd, J = 14.2, 10.0 Hz, 1H), 1.35 (d, J = 7.1 Hz, 3H). ^{13}C NMR (126 MHz, CD_3OD) δ 174.2, 172.5, 172.1, 170.7, 156.5, 131.8, 130.5, 130.0, 119.7, 118.4, 118.0, 114.8, 52.8, 51.0, 49.1, 36.6, 34.6, 21.1, 15.7.

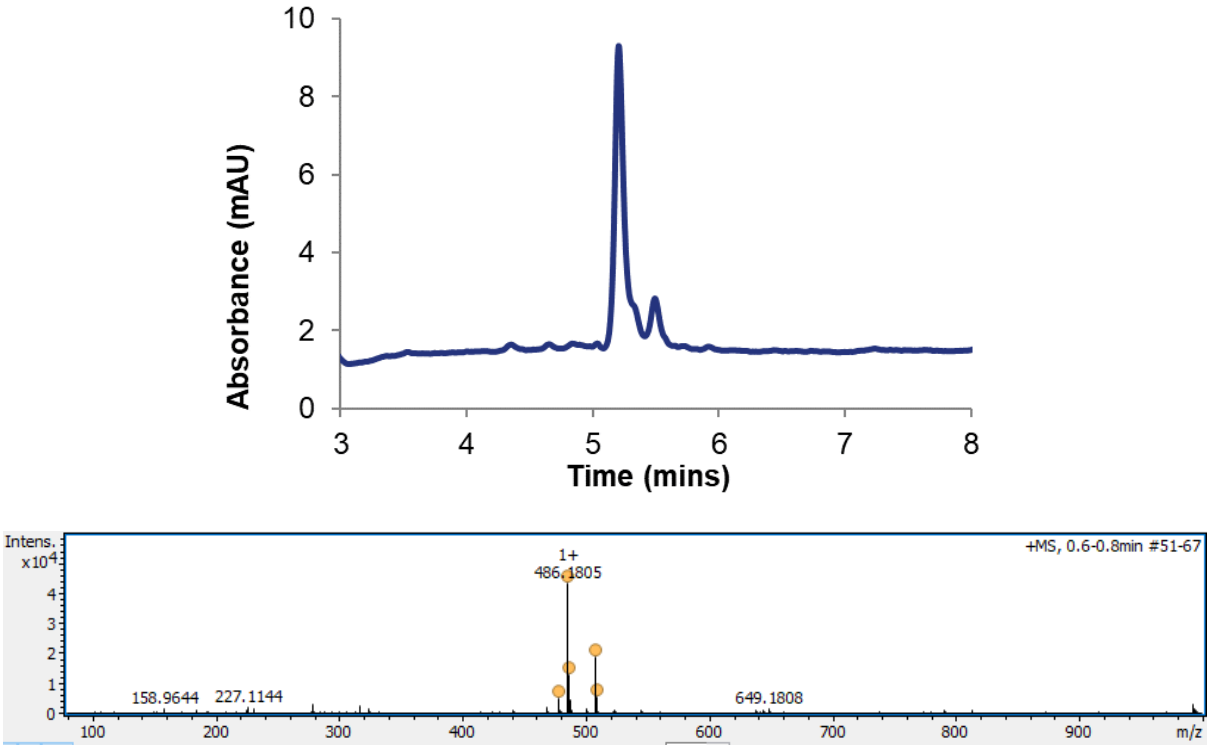
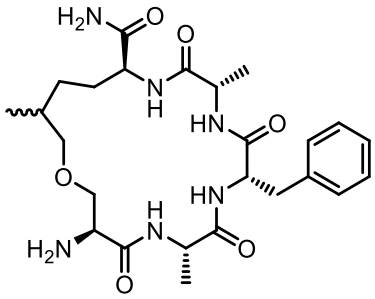


Figure 150 - Analytical HPLC trace of pure Ac-CAY-NH_2 modified with 2-nitrophenylboronic acid through desulfurative cross-coupling (**120**). Analytical gradient 5-60% B over 10 min, 210 nm. Calculated Mass $[M+H]^+$: 486.1989. Observed Mass $[M+H]^+$: 486.1805.

7.9 Peptide Macrocyclisation

H-S(2-methylallyl)AFAX-NH_2 (123)



Synthesised following general protocol I without any alkene addition using H-S(Allyl)AFAC-NH_2 (**122**, 4 mg, 7.26 µmol). After analysis the remaining solution (7.21 µmol) was purified using semi-preparative HPLC (5-60% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid as a mixture of diastereomers - (2.0 mg, 3.88 µmol, 53% yield and 0.9 mg, 1.74 µmol, 24% yield).

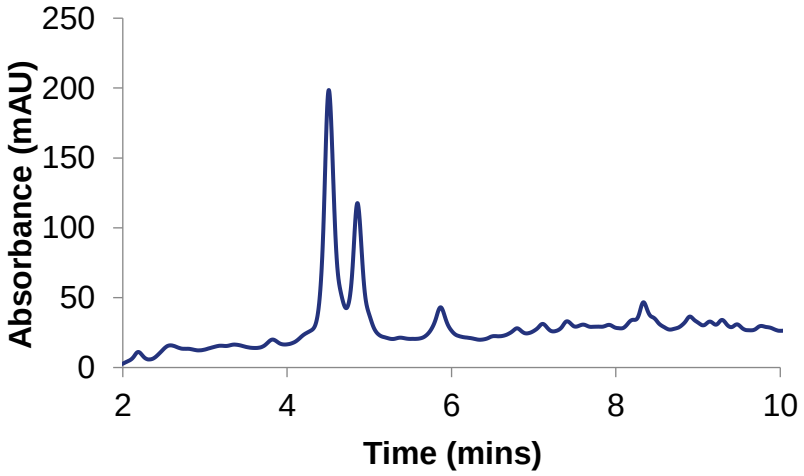
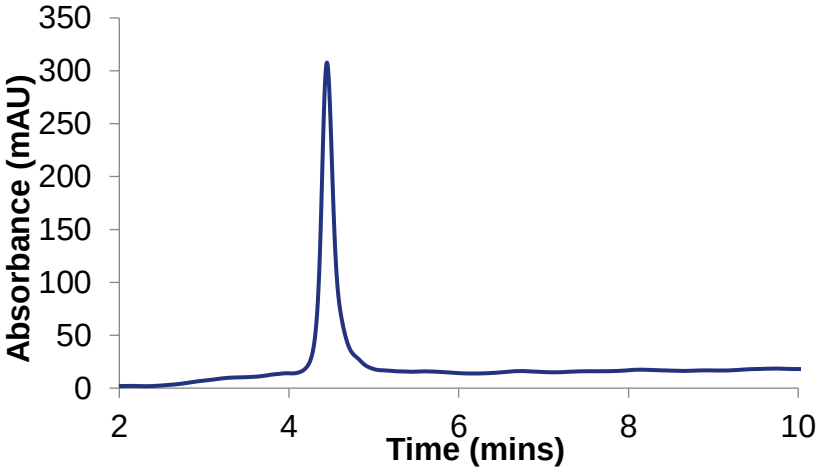


Figure 151 – HPLC analysis of the crude cyclisation of 100 after 10 minutes. Analytical gradient 10-30% B over 10 mins, 210 nm.



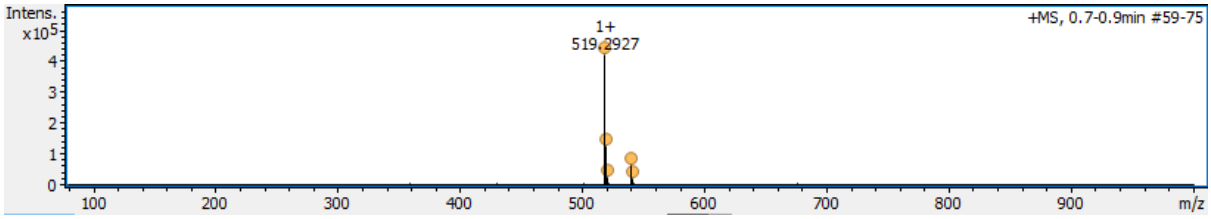
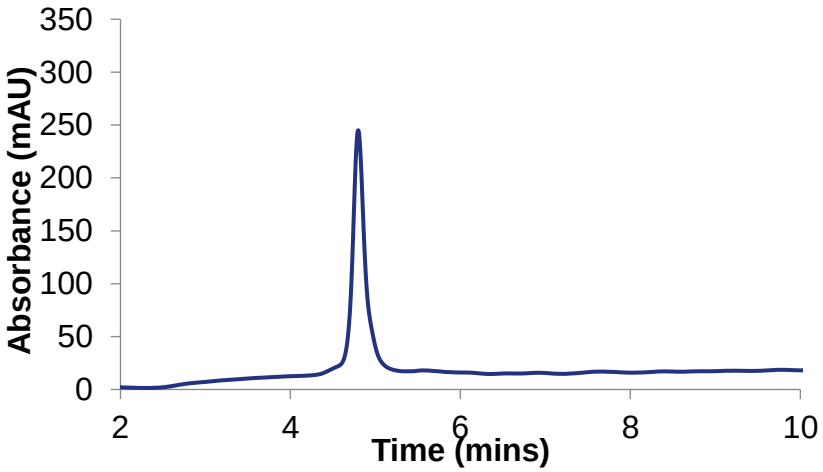


Figure 152 - Pure analytical HPLC trace and MS of cyclised H-S(2-methylallyl)AFAC-NH_2 (**123**), diastereomer 1. Analytical gradient 10-30% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 519.2931. Observed Mass $[M+H]^+$: 519.2927.



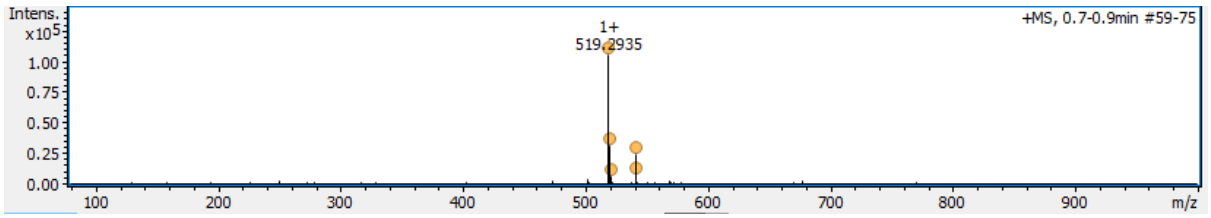
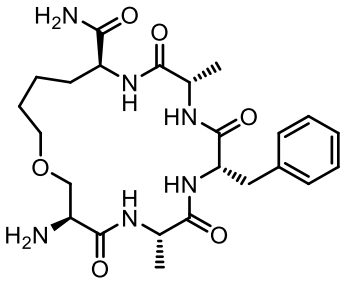


Figure 153 - Pure analytical HPLC trace and MS of cyclised H-S(2-methylallyl)AFAC-NH_2 (**123**), diastereomer 2. Analytical gradient 10-30% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 519.2931. Observed Mass $[M+H]^+$: 519.2935.

H-S(Allyl)AFAX-NH_2 (126)



Synthesised following general protocol K using H-S(Allyl)AFAC-NH_2 (**125**, 4 mg, 7.43 µmol) with 5 mol% Ir catalyst. After analysis the remaining solution (7.35 µmol) was purified using semi-preparative HPLC (5-60% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (2.9 mg, 5.75 µmol, 78% yield).

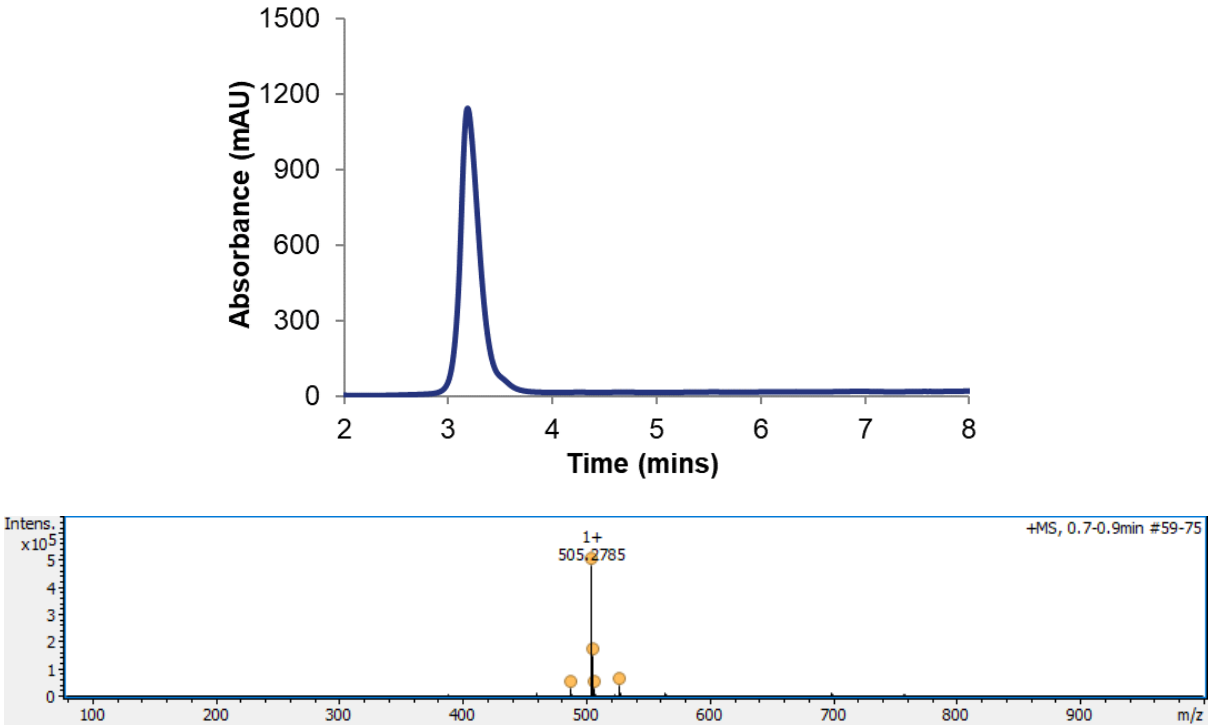
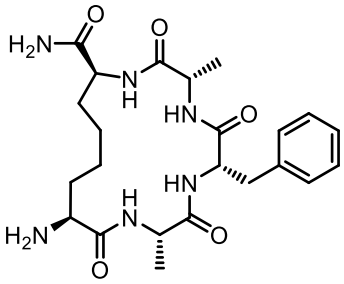


Figure 154 - Pure analytical HPLC trace and MS of cyclised H-S(Allyl)AFAC-NH_2 (**126**). Analytical gradient 2-25% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 505.2775. Observed Mass $[M+H]^+$: 505.2785.

H-alGAFAX-NH_2 (alG-X Cyclised) (130)



Synthesised following general protocol K using H-alGAFAC-NH_2 (**129**, 3 mg, 5.92 µmol). After analysis the remaining solution (5.85 µmol) was purified using semi-preparative HPLC (5-60% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (2.2 mg, 4.64 µmol, 79% yield).

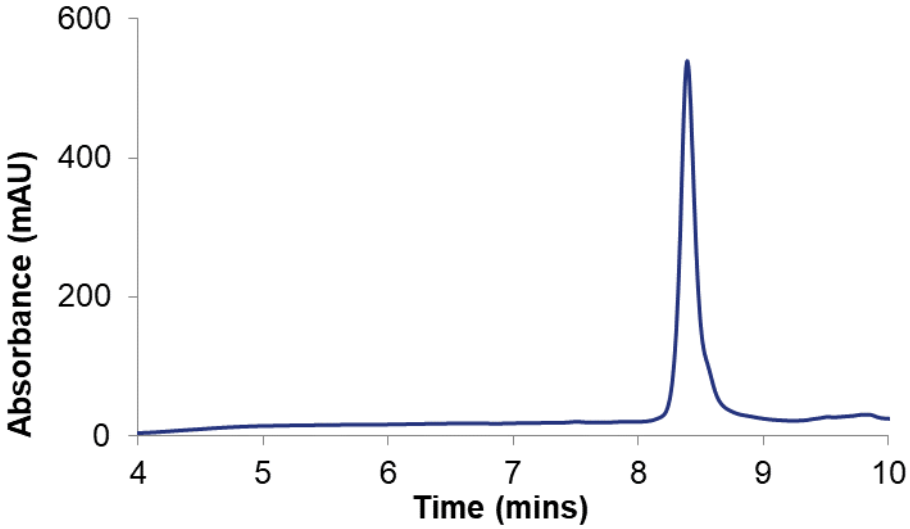


Figure 155 - Pure analytical HPLC trace and MS of desulfurised and cyclised H-alGAFAC-NH_2 (**130**). Analytical gradient 2-25% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 475.2669. Observed Mass $[M+H]^+$: 465.2668.

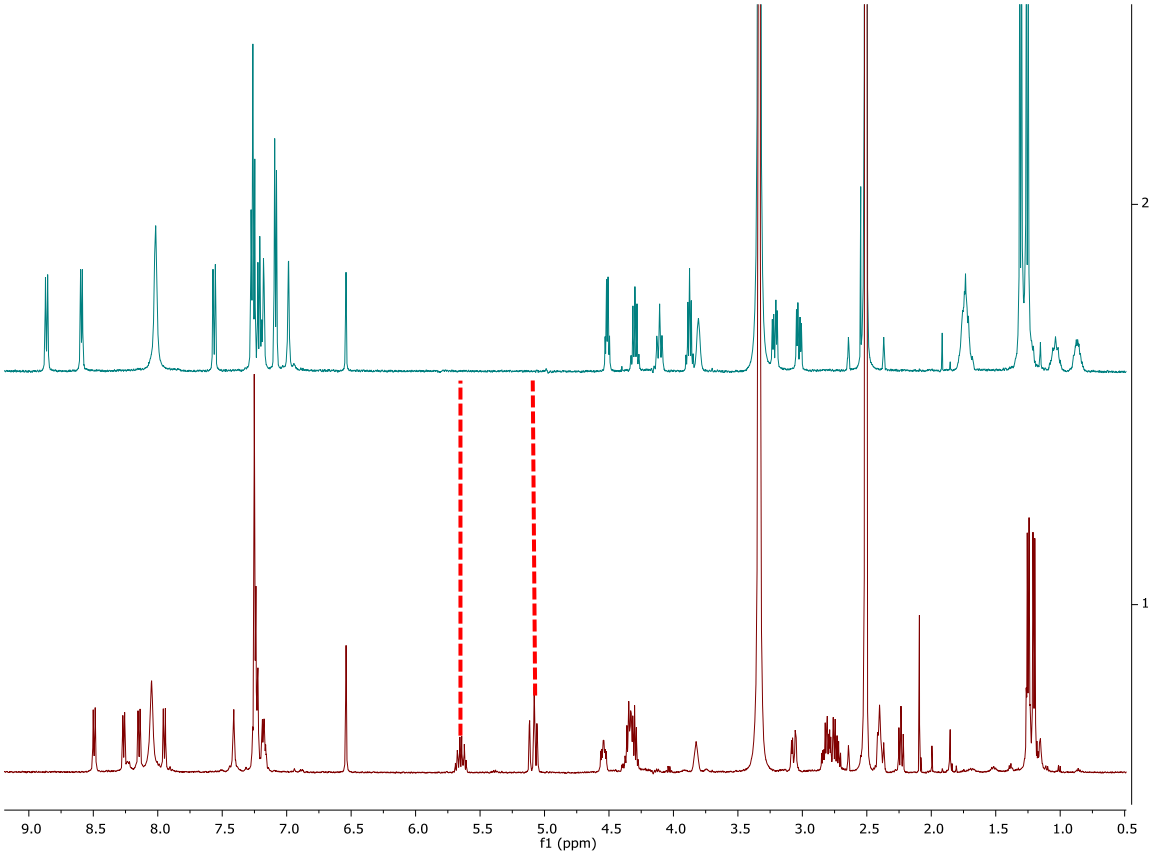
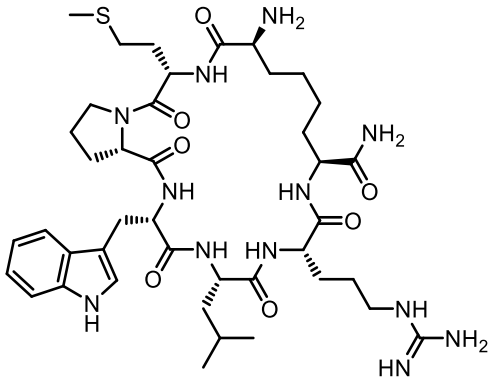
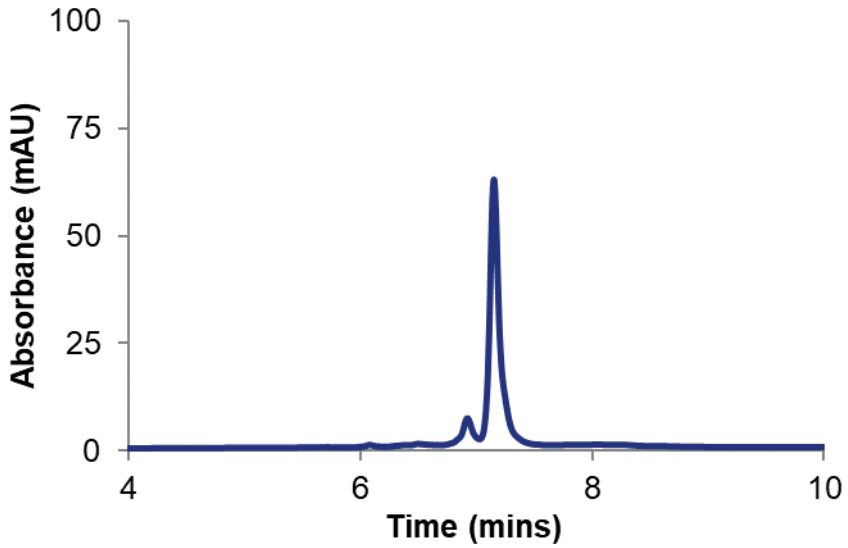


Figure 156 – NMR comparison of pure desulfurised and cyclised H-alGAFAC-NH_2 (**130**, top) and the starting peptide (**129**, bottom), showing disappearance of allylic protons at 5.71 and 5.15 ppm.

H-alGMPWLRX-NH_2 (133)



Synthesised following general protocol K using H-alGMPWLRC-NH_2 (**131**, 2 mg, 2.22 µmol). After analysis the remaining solution (2.12 µmol) was purified using semi-preparative HPLC (10-70% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (1.2 mg, 1.38 µmol, 65% yield).



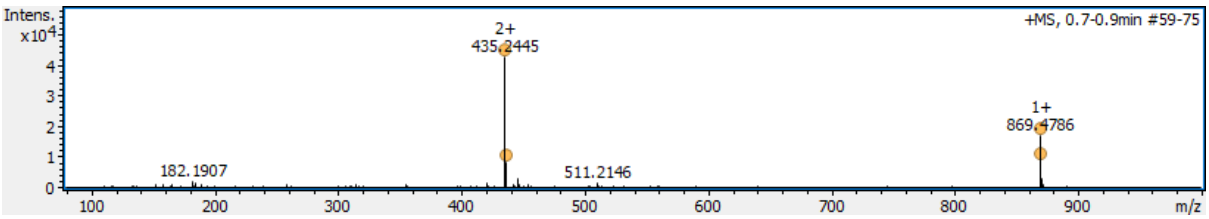


Figure 157 - Pure analytical HPLC trace and MS of desulfurised and cyclised H-alGMPWLRX-NH_2 (**133**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass [M+H]+: 869.4820, [M+2H]2+: 435.2449. Observed Mass [M+H]+: 869.4786, [M+2H]2+: 435.2445.

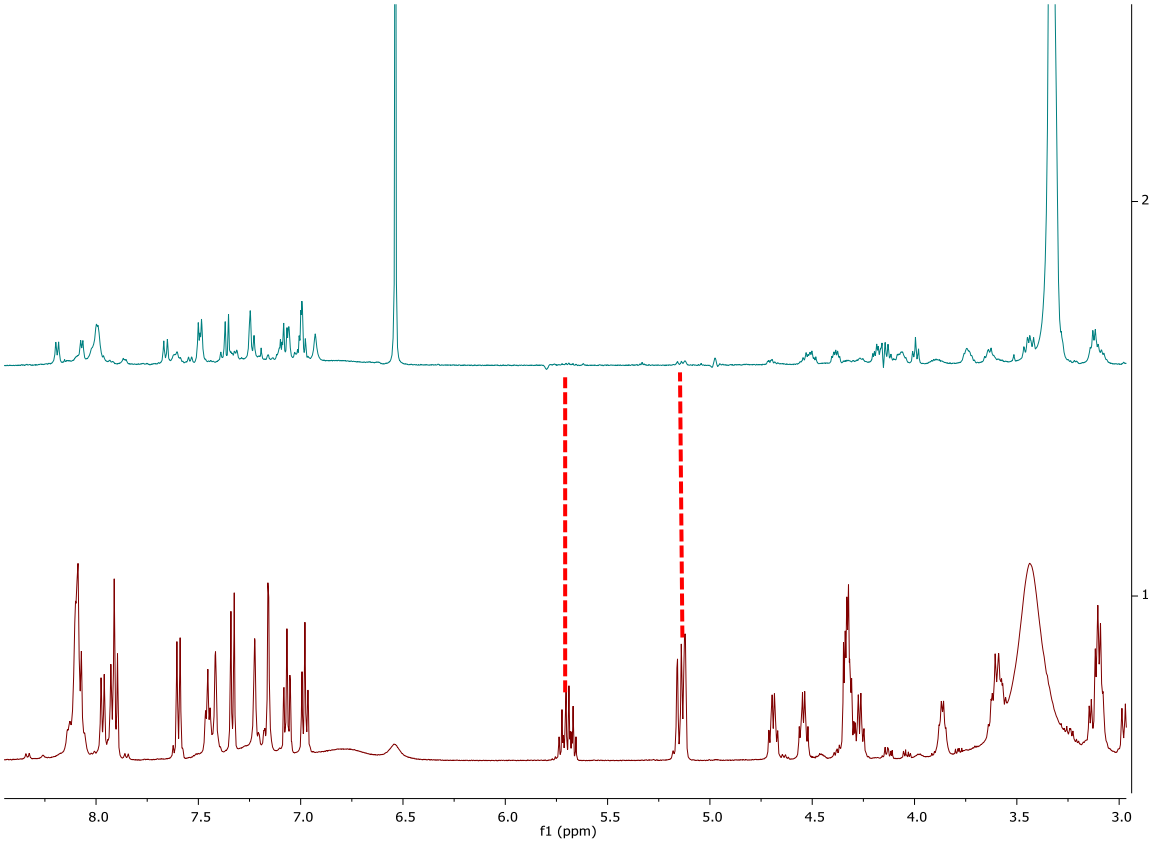
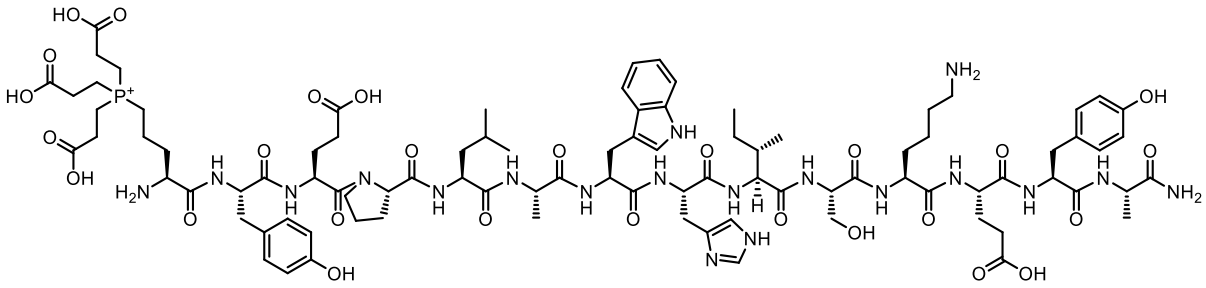
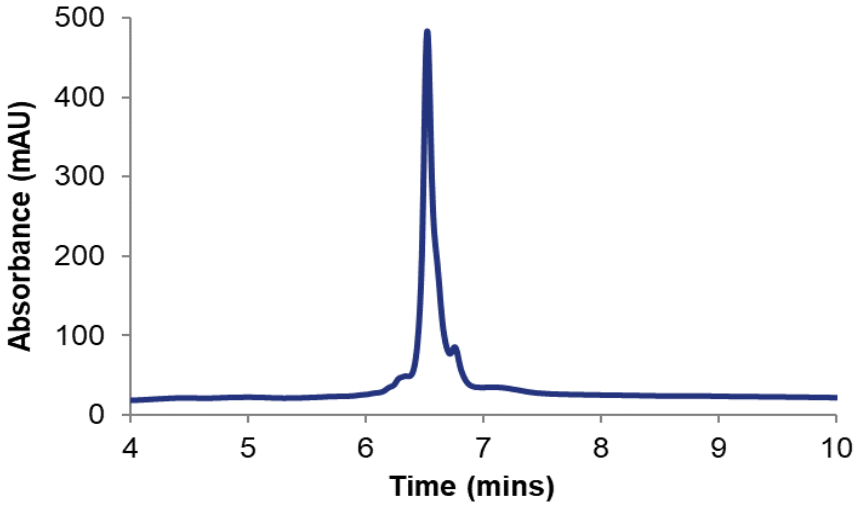


Figure 158 - NMR overlay comparing oxytocin model starting peptide H-alGMPWLRC-NH_2 (**131**, bottom) and the cyclised product (**133**, top), showing disappearance of allylic proton peaks at 5.75 and 5.20 ppm.

H-alGYEPLAWHISKEYC-NH_2 Trial 1 (134)



Synthesised following general protocol K using H-alGYEPLAWHISKEYC-NH_2 (**132**, 2 mg, 1.15 µmol), with 5 mol% Ir(III) catalyst. After analysis the remaining solution (1.06 µmol) was purified using semi-preparative HPLC (10-80% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (1.2 mg, 0.614 µmol, 58% yield).



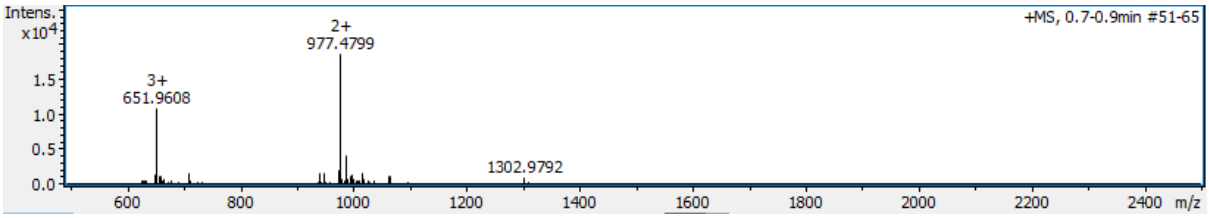


Figure 159 - Pure analytical HPLC trace and MS of desulfurised and TCEP nucleophilic addition product H-A(TCEP)YEPLAWHISKEYA-NH2 (**134**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass $[M+2H]^{2+}$: 977.4678, $[M+3H]^{3+}$: 651.9811. Observed Mass $[M+2H]^{2+}$: 977.4799, $[M+3H]^{3+}$: 651.9608.

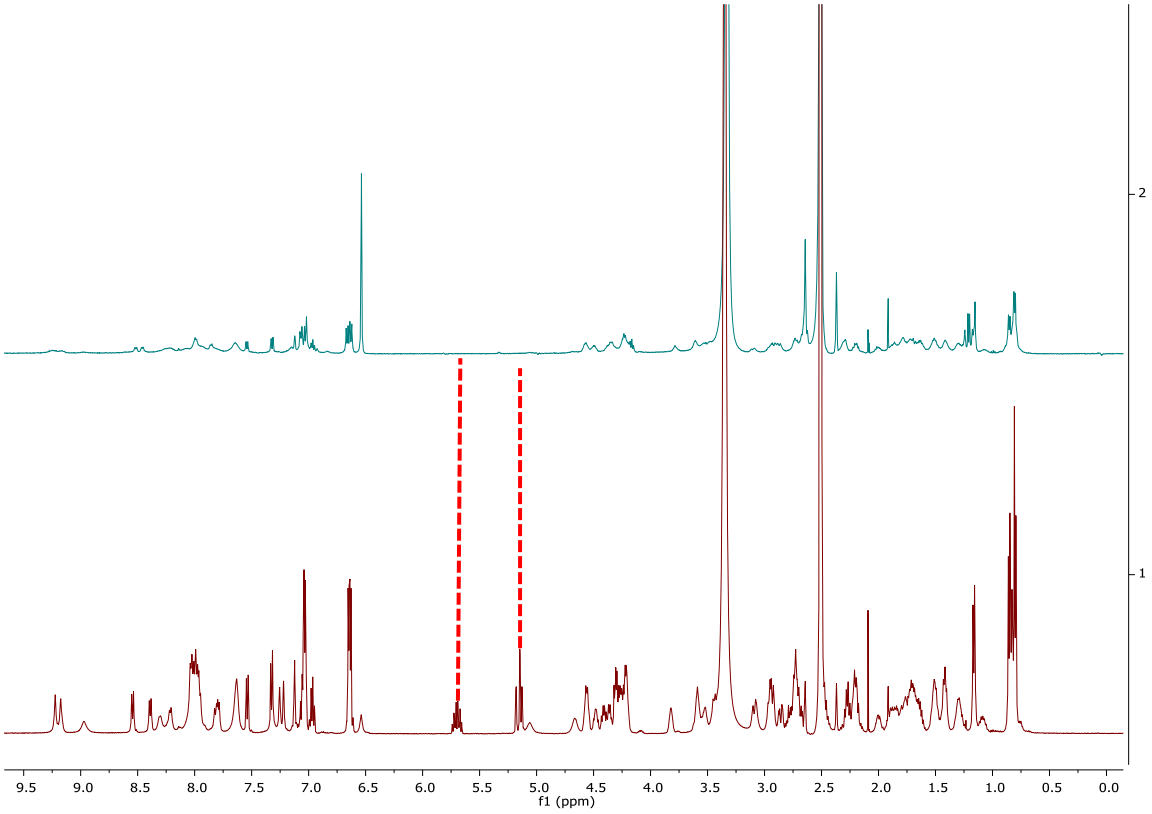


Figure 160 - Pure starting material H-alGYEPLAWHISKEYC-NH_2 (**132**, bottom) and purified product (**134**, top) NMR comparison, showing disappearance of allylic protons.

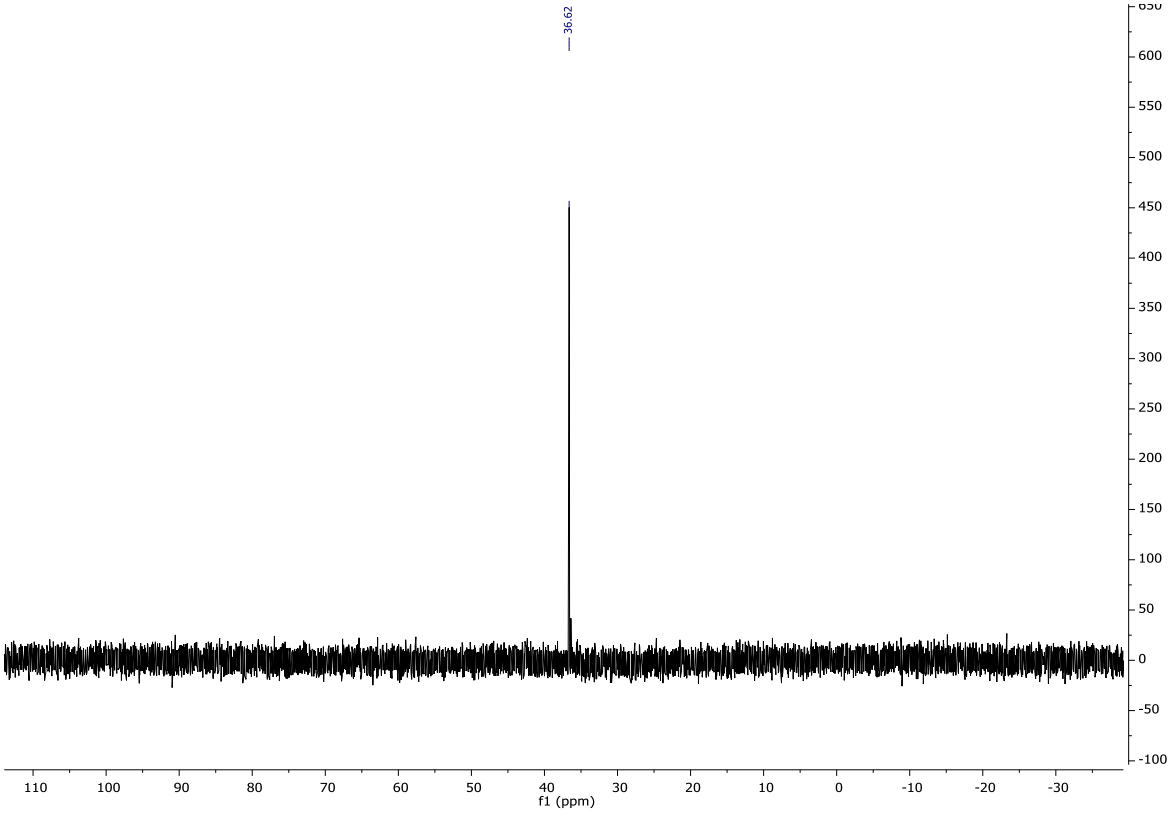
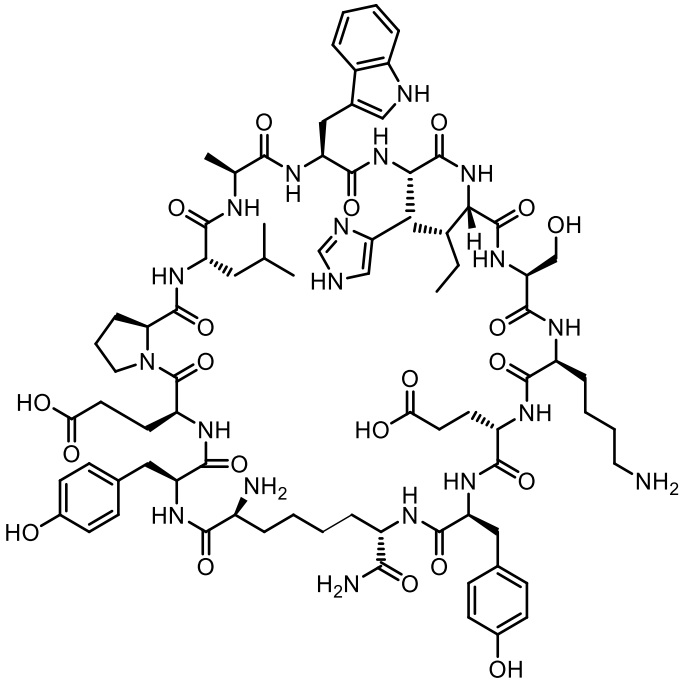


Figure 161 - ^{31}P NMR ((162 MHz, DMSO) δ 36.62.) of the TCEP-nucleophilic addition product (**134**) with the desulfurised model peptide H-alGYEPLAWHISKEYC-NH_2.

H-alGYEPLAWHISKEYX-NH_2 Trial 2 (Mixture of alG-X cyclised (X=bridged residue) and uncyclised (X=A)) (135)



Synthesised following general protocol K using H-alGYEPLAWHISKEYC-NH_2 (**132**, 2 mg, 1.15 µmol). After analysis the remaining solution (1.06 µmol) was purified using semi-preparative HPLC (10-70% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (1.3 mg, 0.763 µmol, 72% yield).

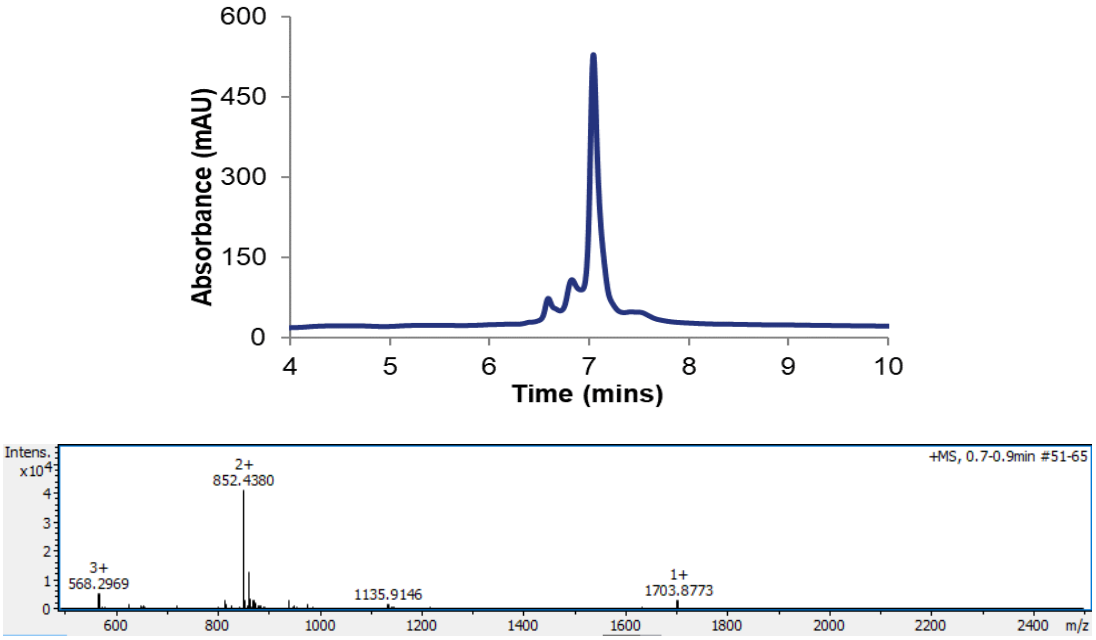


Figure 162 - Analytical HPLC trace and MS of the purified inseparable mixture of cyclised and uncyclised product H-alGYEPLAWHISKEYX-NH_2 (**135**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 1703.8671, $[M+2H]^{2+}$: 852.4375. Observed Mass $[M+H]^+$: 1703.8773, $[M+2H]^{2+}$: 852.4380.

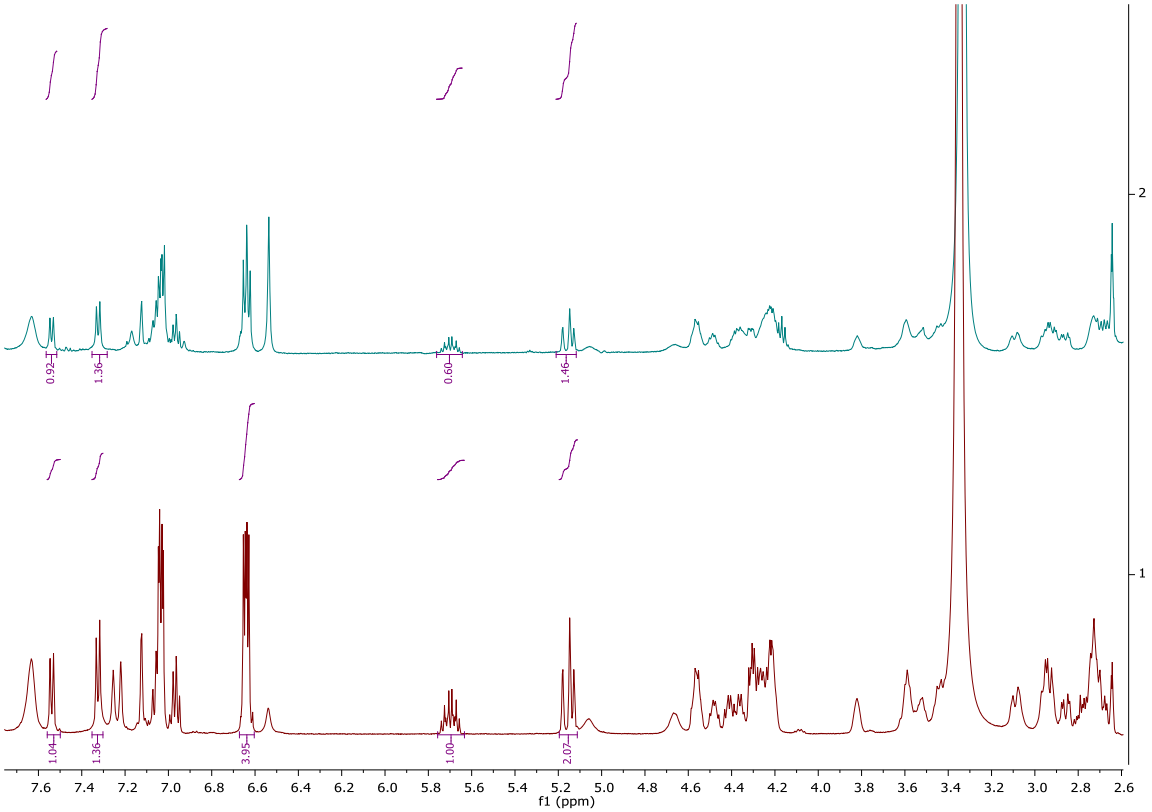
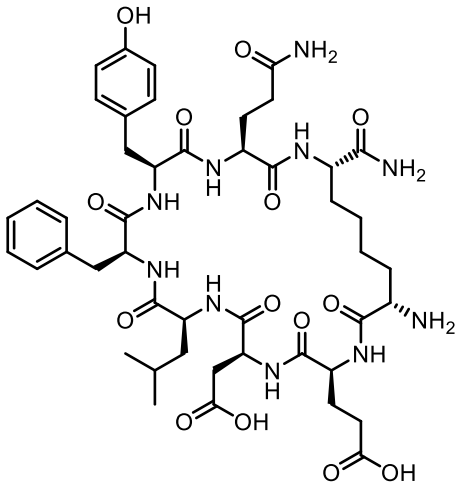


Figure 163 – Pure starting material H-alGYEPLAWHISKEYC-NH_2 (**132**, bottom) and purified isolated mixture of desulfurised and cyclised/uncyclised material (**135**, top) NMR comparison, showing reduction in the integration of allylic protons.

H-alGEDLFYQX-NH_2 (alG-X cyclised) (138)



Synthesised following general protocol K using H-alGEDLFYQX-NH_2 (**136**, 2 mg, 1.97 µmol). After analysis the remaining solution (1.88 µmol) was purified using semi-preparative HPLC (15-80% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (1.2 mg, 1.22 µmol, 65% yield).

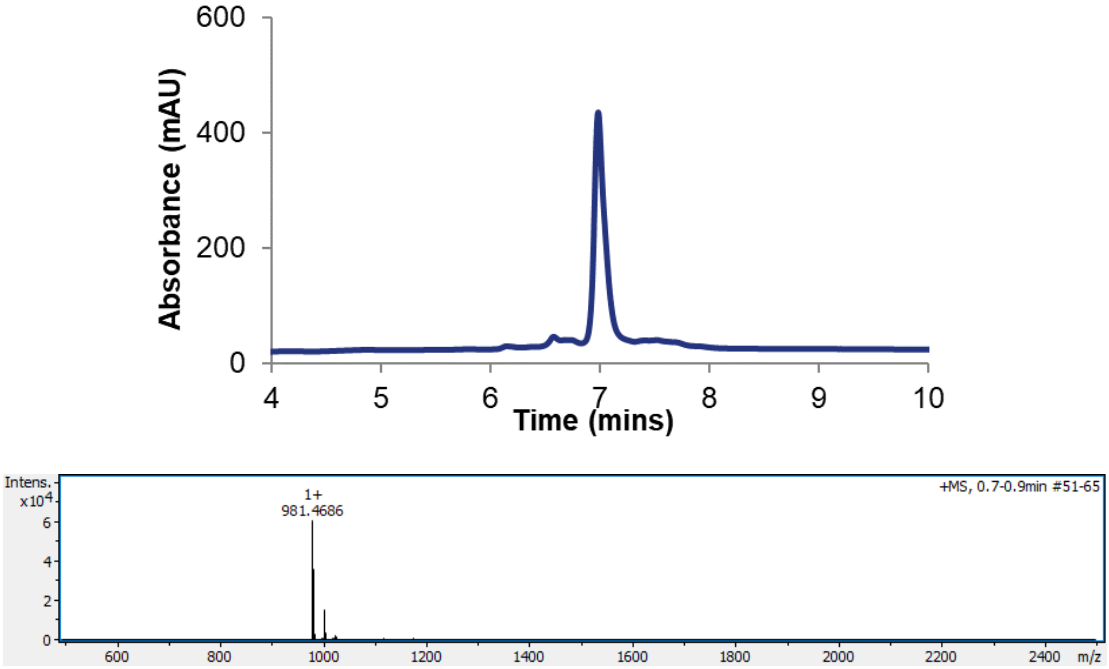


Figure 164 - Pure analytical HPLC trace and MS of desulfurised and cyclised H-alGEDLFYQC-NH_2 (**138**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 961.4682. Observed Mass $[M+H]^+$: 981.4686.

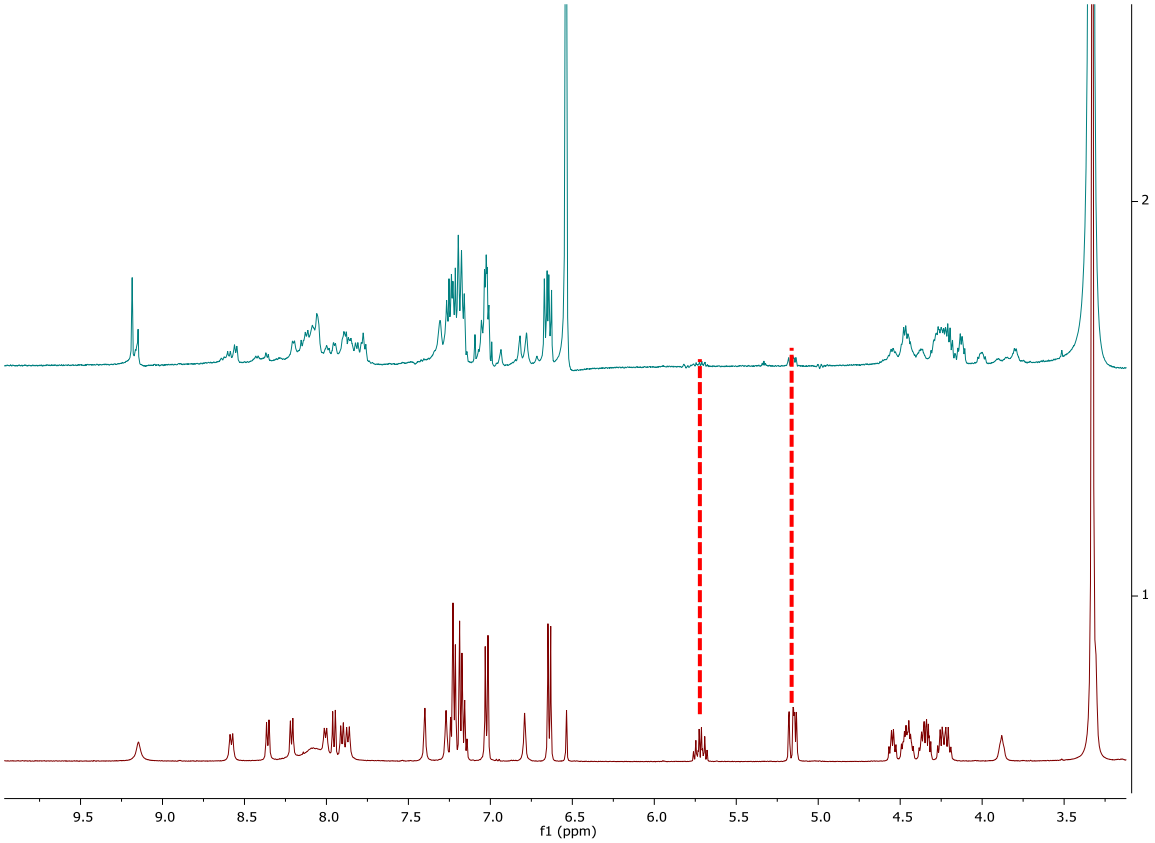
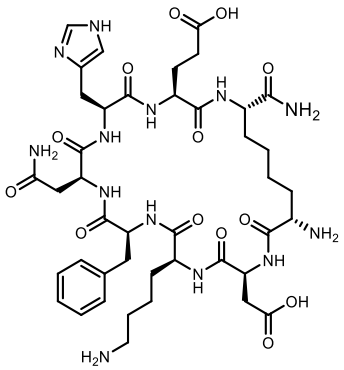


Figure 165 - NMR overlay of starting peptide H-alGEDLFYQC-NH_2 (**136**, bottom) and cyclised product (**138**, top), showing disappearance of allylic proton peaks at 5.75 and 5.20 ppm.

H-alGDKFNHEX-NH2 (alG-X Cyclised) (139)



Synthesised following general protocol K using H-alGDKFNHEC-NH_2 (**137**, 2 mg, 2.02 µmol). After analysis the remaining solution (1.93 µmol) was purified using semi-preparative HPLC (2-60% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (1.3 mg, 1.36 µmol, 70% yield).

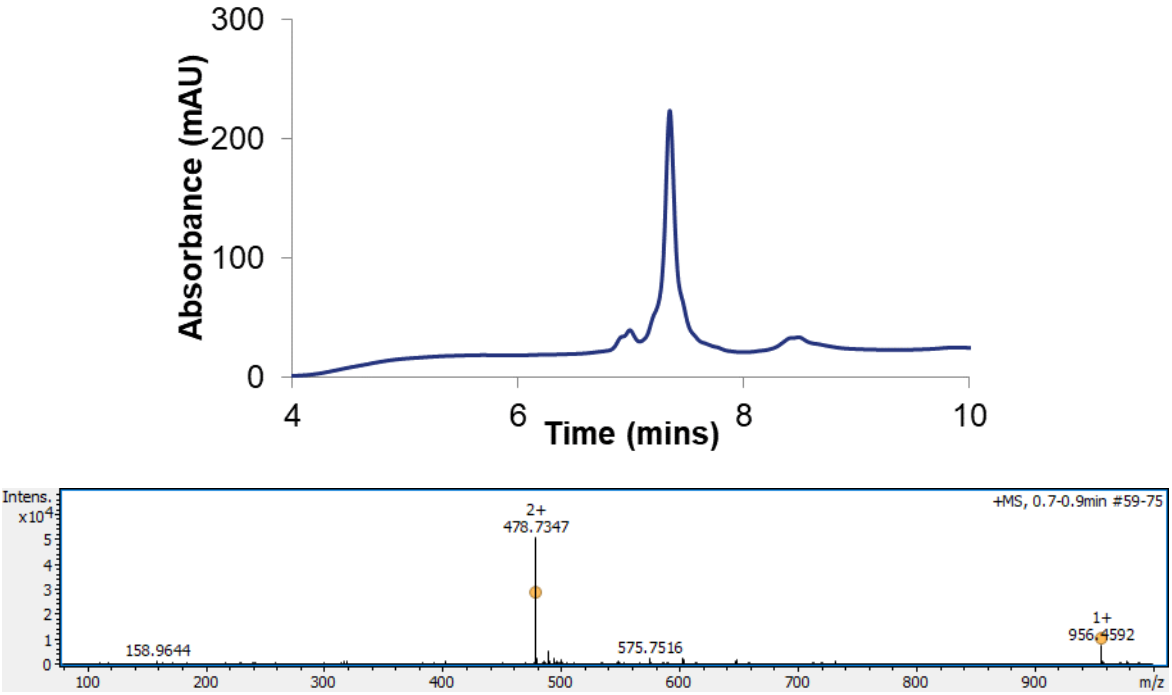


Figure 166 – Pure analytical HPLC trace and MS of desulfurised and cyclised H-alGDKFNHEC-NH_2 (**139**). Analytical gradient 2-25% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 956.4590, $[M+2H]^{2+}$: 478.7334. Observed Mass $[M+H]^+$: 956.4592, $[M+2H]^{2+}$: 478.7347.

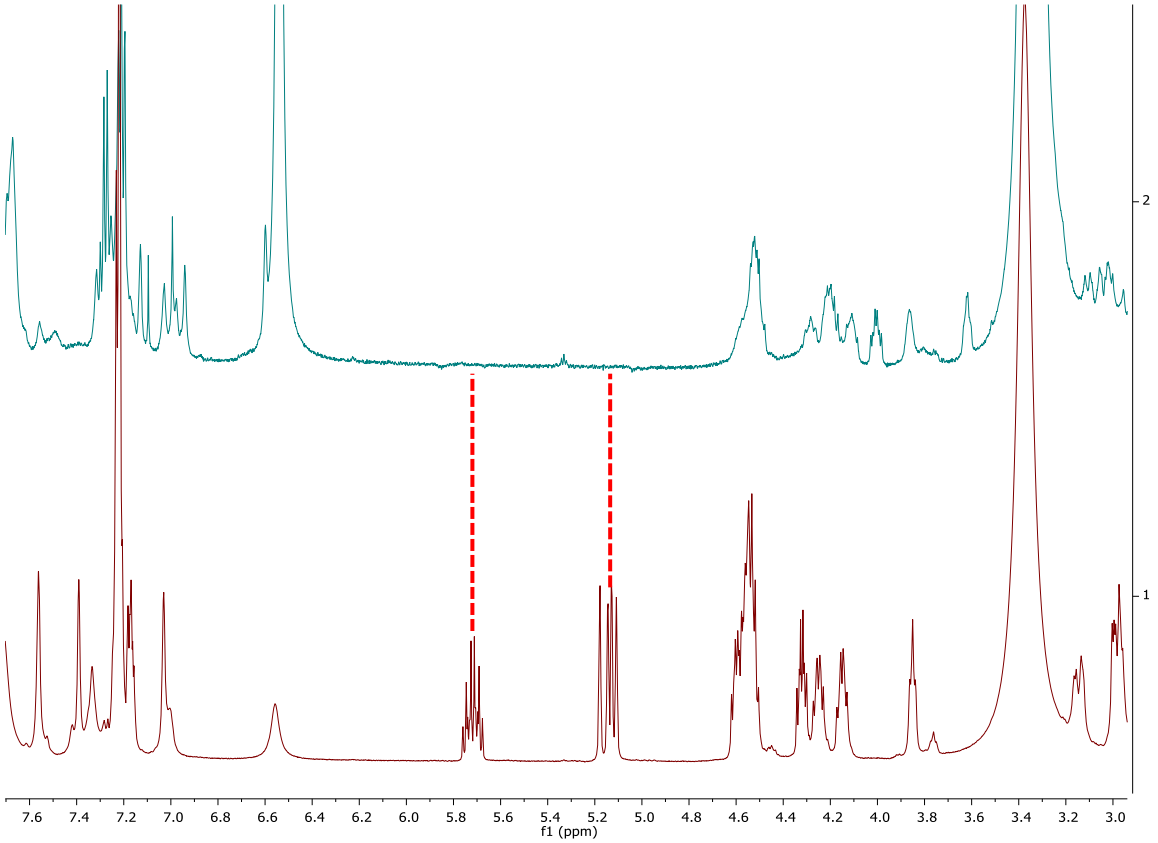
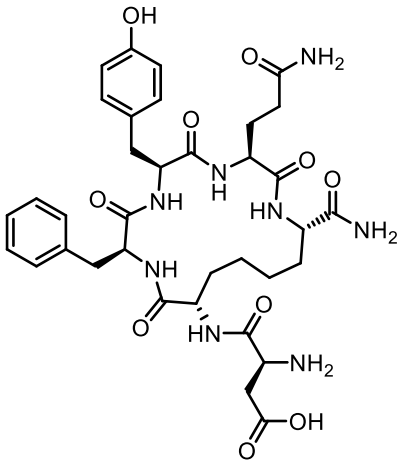
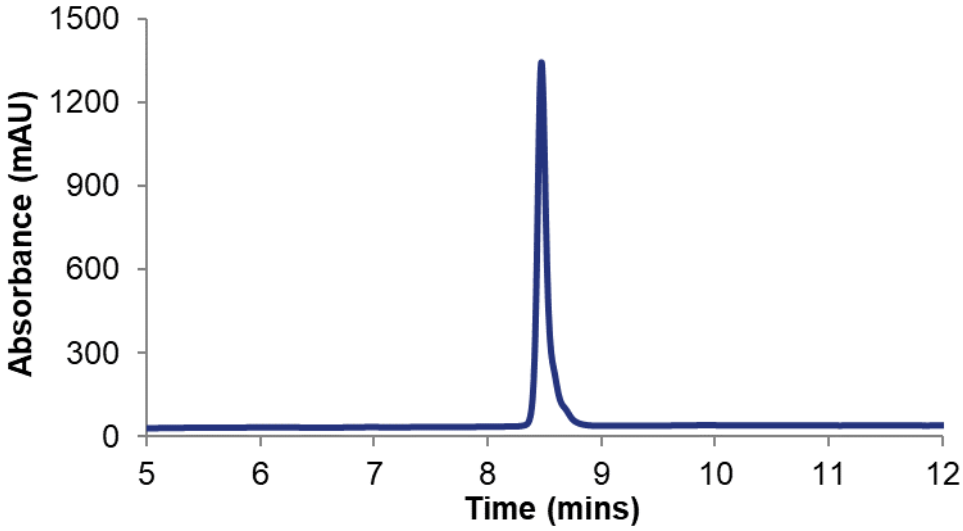


Figure 167 - NMR comparison of pure desulfurised and cyclised H-alGDKFNHEC-NH_2 (**139**, top) and the starting peptide (**137**, bottom), showing disappearance of allylic protons at 5.71 and 5.15 ppm.

H-DalGFYQX-NH_2 (alG-X cyclised) (141)



Synthesised following general protocol K using H-DalGFYQC-NH_2 (**140**, 4 mg, 5.20 µmol). After analysis the remaining solution (5.10 µmol) was purified using semi-preparative HPLC (10-50% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (3.1 mg, 4.20 µmol, 82% yield).



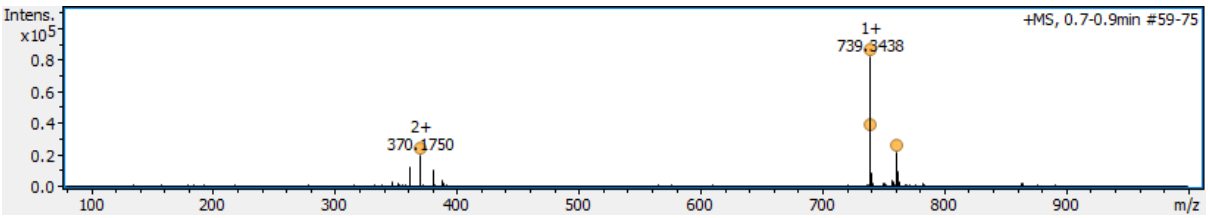


Figure 168 - Pure analytical HPLC trace and MS of desulfurised and cyclised H-DalGFYQC-NH_2 (**141**). Analytical gradient 2-40% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 739.3415, $[M+2H]^{2+}$: 370.1748. Observed Mass $[M+H]^+$: 739.3438, $[M+2H]^{2+}$: 370.1750.

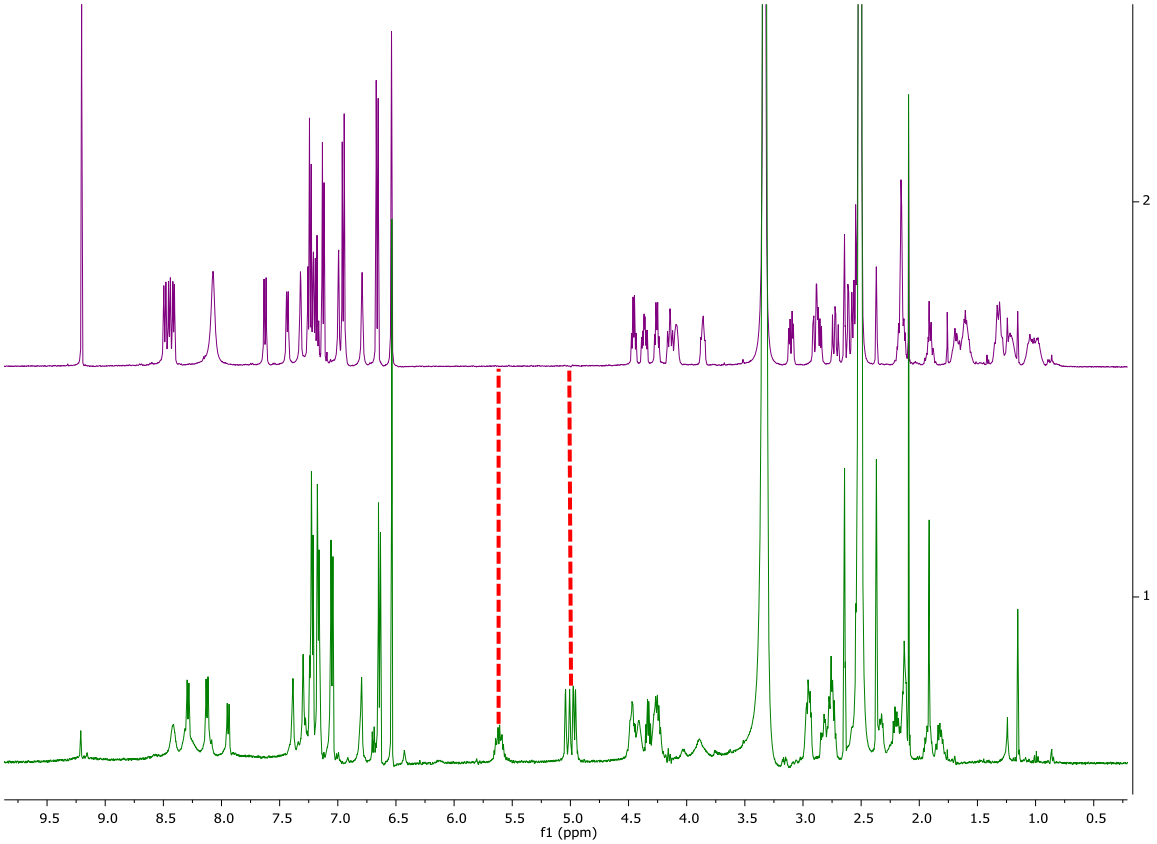
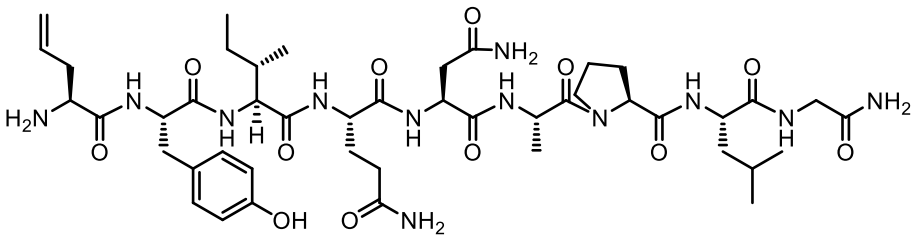


Figure 169 - NMR overlay of starting peptide H-DalGFYQC-NH_2 (**140**, bottom) and the cyclised product (**141**, top), showing disappearance of allylic protons at 5.7 and 5.0 ppm. This also shows a lack of characteristic 3H doublet for Ala (usually around 1.3 ppm), also indicating cyclisation has occurred.

H-alGYIQNAPLG-NH_2 (143)



Synthesised following general protocol K using H-alGYIQNXPLG-NH_2 (**142**, 2 mg, 1.99 µmol). After analysis the remaining solution (1.90 µmol) was purified using semi-preparative HPLC (10-70% B over 30 minutes) and fractions containing the desired product lyophilised to yield the title compound as a fluffy white solid (1.5 mg, 1.55 µmol, 82% yield).

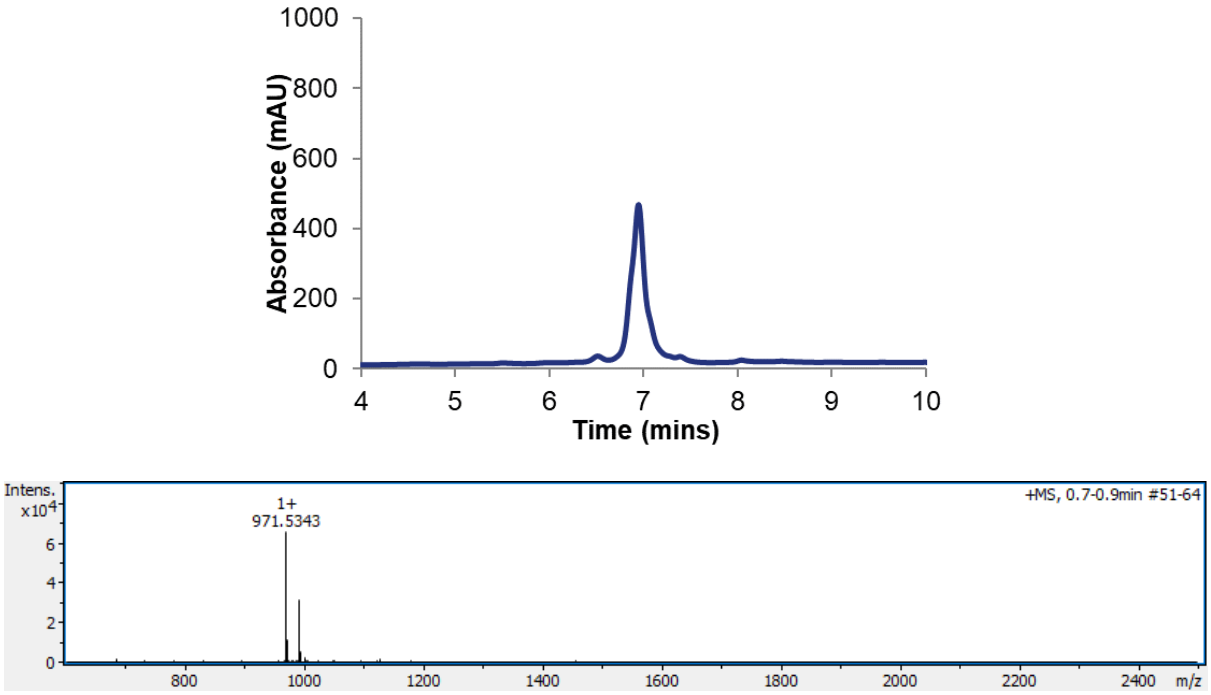


Figure 170 - Pure analytical HPLC trace and MS of desulfurised H-alGYIQNXPLG-NH_2 (**143**). Analytical gradient 10-30% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 971.5314. Observed Mass $[M+H]^+$: 971.5343.

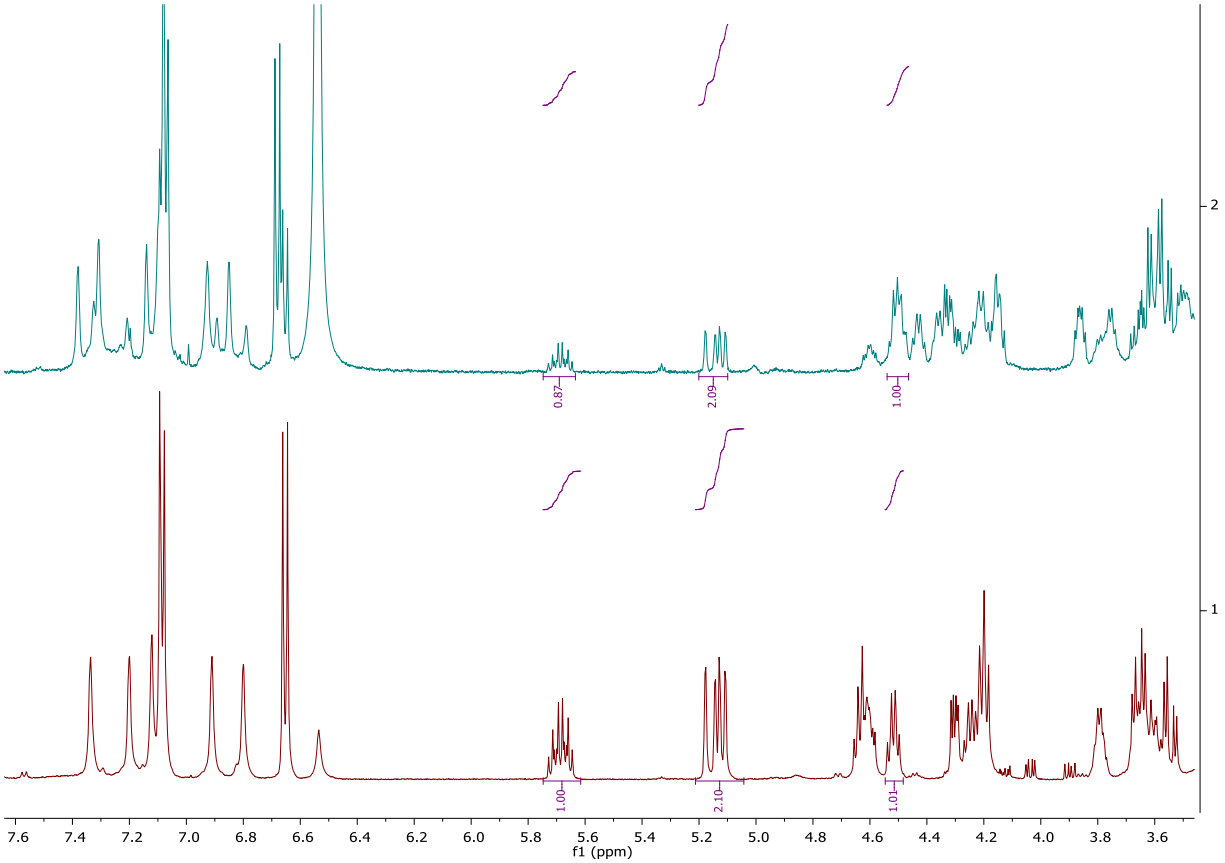
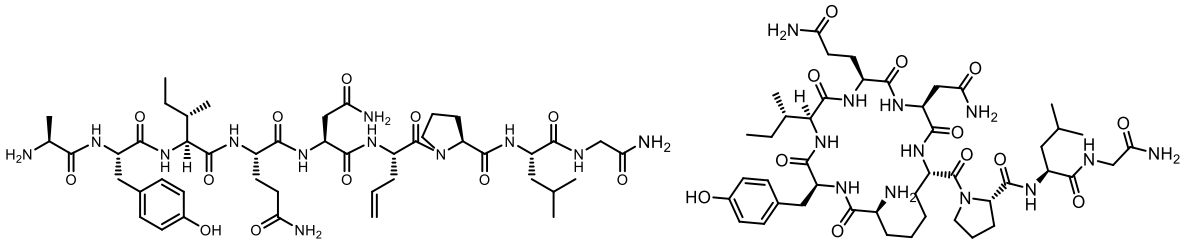


Figure 171 - NMR overlay of starting peptide H-alGYIQNXPLG-NH_2 (**142**, bottom) and product (**143**, top). This shows the continuing presence of the allylic protons at 5.7 and 5.15 ppm, indicating cyclisation was not successful.

H-XYIQNalGPLG-NH_2 (144) (Separated mixture of X-alG cyclised and uncyclised) (145 and 146)



Synthesised following general protocol K using H-CYIQNalGPLG-NH_2 (**144**, 5 mg, 4.98 µmol). After analysis the remaining solution (4.89 µmol) was purified using semi-preparative HPLC (10-70% B over 30 minutes) and fractions containing the desired product lyophilised to yield the uncyclised title compound (**146**, 1.7 mg, 1.73 µmol, 35% yield) and the cyclised title compound (**145**, 2.5 mg, 2.54 µmol, 52% yield) as fluffy white solids.

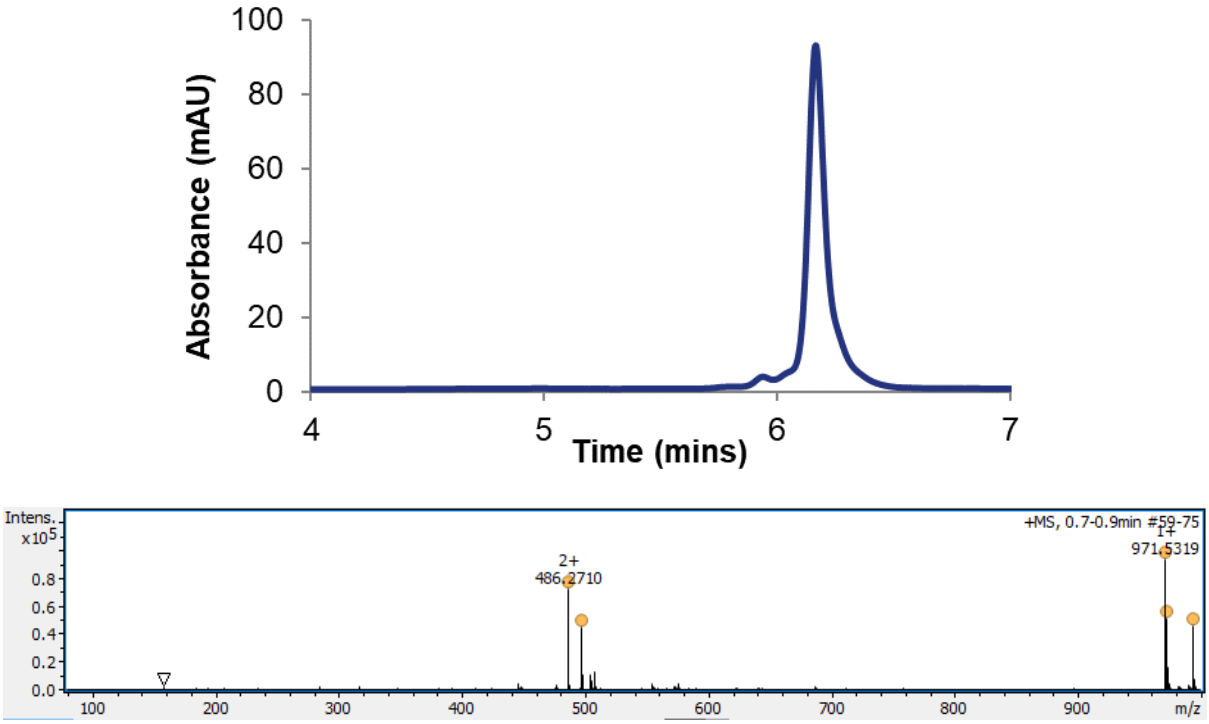


Figure 172 - Pure analytical HPLC trace and MS of cyclised H-XYIQNalGPLG-NH_2 (**145**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 971.5314, $[M+2H]^{2+}$: 486.2697. Observed Mass $[M+H]^+$: 971.5319, $[M+2H]^{2+}$: 486.2710.

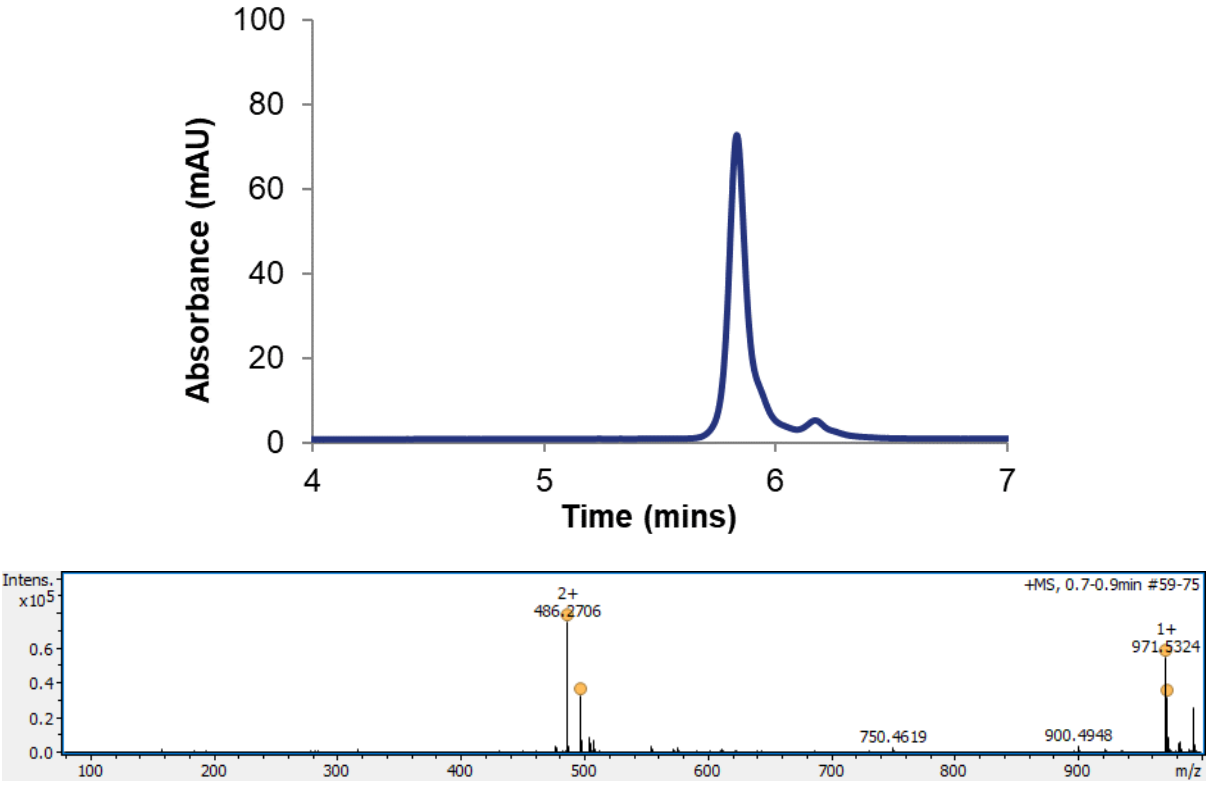


Figure 173 - Pure analytical HPLC trace and MS of linear desulfurised H-XYIQNalGPLG-NH_2 (**146**). Analytical gradient 10-50% B over 10 minutes, 210 nm. Calculated Mass $[M+H]^+$: 971.5314, $[M+2H]^{2+}$: 486.2697. Observed Mass $[M+H]^+$: 971.5324, $[M+2H]^{2+}$: 486.2706.

References

1 K. Fosgerau and T. Hoffmann, *Drug Discov. Today*, 2015, **20**, 122–8.

2 D. G. Rawale, K. Thakur, S. R. Adusumalli and V. Rai, *European J. Org. Chem.*, 2019, **2019**, 6749–6763.

3 Y. Go, J. D. Chandler and D. P. Jones, *Free Radic Biol Med. 2015*, 2015, **84**, 227–245.

4 P. Hoppenz, S. Els-Heindl and A. G. Beck-Sickinger, *Front. Chem.*, 2020, **8**, 1–24.

5 Y. Wang, A. G. Cheetham, G. Angacian, H. Su, L. Xie and H. Cui, *Adv. Drug Deliv. Rev.*, 2017, **110**–**111**, 112–126.

6 P. Zhao, Y. Zhang, W. Li, C. Jeanty, G. Xiang and Y. Dong, *Acta Pharm. Sin. B*, 2020, **10**, 1589–1600.

7 P. Khongorzul, C. J. Ling, F. U. Khan, A. U. Ihsan and J. Zhang, *Mol. Cancer Res.*, 2020, **18**, 3–19.

8 N. Yin, *Med. Chem.*, 2015, **4**, 763–769.

9 S. V. Moradi, W. M. Hussein, P. Varamini, P. Simerska and I. Toth, *Chem. Sci.*, 2016, **7**, 2492–2500.

10 K. Hosoyama, C. Lazurko, M. Muñoz, C. D. McTiernan and E. I. Alarcon, *Front. Bioeng. Biotechnol.*, 2019, **7**, 225

11 T. Kimmerlin and D. Seebach, *J. Pept. Res.*, 2005, **65**, 229–260.

12 V. Mäde, S. Els-Heindl and A. G. Beck-Sickinger, *Beilstein J. Org. Chem*, 2014, **10**, 1197–1212.

13 S. L. Jensen, Knud J.; Tofteng Shelton, Pernille; Pedersen, *Peptide Synthesis and Applications*, 2013.

14 Abdelmoty, F. Albericio, L.A. Carpino, B.M. Foxman, and S.A. Kates, 1994, **1**, 1–55.

15 M. Stawikowski and G. B. Fields, *Curr. Protoc. Protein Sci.*, 2002, **26**, 1–17.

16 C. Petrou and Y. Sarigiannis, *Peptide synthesis: Methods, trends, and challenges*, Elsevier Ltd, 2017.

17 E. Valeur and M. Bradley., *Chem. Soc. Rev.,* 2009, **38**, 606-631

18 T. I. Al-Warhi, H. M. A. Al-Hazimi and A. El-Faham, *J. Saudi Chem. Soc.*, 2012, **16**, 97–116.

19 A. Di Fenza, M. Tancredi, C. Galoppini and P. Rovero, *Tetrahedron Lett.*, 1998, **39**,

8529–8532.

20 T. I. Al-Warhi, H. M. A. Al-Hazimi and A. El-Faham, *J. Saudi Chem. Soc.*, 2012, **16**, 97–116.

21 E. I. Vrettos, N. Sayyad, E. M. Mavrogiannaki, E. Stylos, A. D. Kostagianni, S. Papas, T. Mavromoustakos, V. Theodorou and A. G. Tzakos, *RSC Adv.*, 2017, **7**, 50519–50526.

22 K. J. McKnelly, W. Sokol and J. S. Nowick, *J. Org. Chem.*, 2020, **85**, 1764–1768.

23 R. Subirós-Funosas, R. Prohens, R. Barbas, A. El-Faham and F. Albericio, *Chem. - A Eur. J.*, 2009, **15**, 9394–9403.

24 M. G. Ryadnov and D. N. Woolfson, *Angew. Chemie - Int. Ed.*, 2003, **42**, 3021–3023.

25 L. Liu, *Isr. J. Chem.*, 2019, **59**, 64–70.

26 N. Ollivier, R. Desmet, H. Drobecq, A. Blanpain, E. Boll, B. Leclercq, A. Mougel, J. Vicogne and O. Melnyk, *Chem. Sci.*, 2017, **8**, 5362–5370.

27 N. J. Mitchell, L. R. Malins, X. Liu, R. E. Thompson, B. Chan, L. Radom and R. J. Payne, *J. Am. Chem. Soc.*, 2015, **137**, 14011–14014.

28 D. Macmillan and C. R. Bertozzi, *Angew. Chem. Int. Ed.*, 2004, **43**, 1355–1359.

29 R. J. Payne and C. H. Wong, *Chem. Commun.*, 2010, **46**, 21–43.

30 B. L. Nilsson, L. L. Kiessling and R. T. Raines, *Org. Lett.*, 2000, **2**, 1939–1941.

31 P. Thapa, R. Y. Zhang, V. Menon and J. P. Bingham, *Molecules*, 2014, **19**, 14461–14483.

32 Q. Wan and S. J. Danishefsky, *Angew. Chemie - Int. Ed.*, 2007, **46**, 9248–9252.

33 S. Dery, S. Reddy, L. Dery, R. Mousa, R. N. Dardashti and N. Metanis., *Chem. Sci.,* 2015, 6, **11**, 6207-6212..

34 D. J. Boerema, V. A. Tereshko and S. B. H. Kent, *Biopolym. - Pept. Sci. Sect.*, 2008, **90**, 278–286.

35 C. Araman, G. Laboratories and C. C. Leiden, 2017, **2**, 5–7.

36 J. Köhrl, R. Brigelius-Flohé, A. Böck, R. Gärtner, O. Meyer and L. Flohé, *Biol. Chem.*, 2000, **381**, 849–64.

37 Q. Cheng and E. S. J. Arnér, *J. Biol. Chem.*, 2017, **292**, 5476–5487.

38 R. Hondal., *Protein Pept. Lett.,* 2013, **12**, 757–764.

39 S. D. Townsend, Z. Tan, S. Dong, S. Shang, J. A. Brailsford and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2012, **134**, 3912–3916.

40 J. Sayers, P. M. T. Karpati, N. J. Mitchell, A. M. Goldys, S. M. Kwong, N. Firth, B. Chan and R. J. Payne, *J. Am. Chem. Soc.*, 2018, **140**, 13327–13334.

41 T. S. Chisholm, S. S. Kulkarni, K. R. Hossain, F. Cornelius, R. J. Clarke and R. J. Payne, *J. Am. Chem. Soc.*, 2020, **142**, 1090–1100.

42 R. N. Dardashti, S. Kumar, S. M. Sternisha, P. S. Reddy, B. G. Miller and N. Metanis, *Chem. - A Eur. J.*, 2020, 4952–4957.

43 J. Sayers, P. M. T. Karpati, N. J. Mitchell, A. M. Goldys, S. M. Kwong, N. Firth, B. Chan and R. J. Payne, *J. Am. Chem. Soc.*, 2018, **140**, 13327–13334.

44 X. Wang, L. Corcilius, B. Premdjee and R. J. Payne, *J. Org. Chem.*, 2020, **85**, 1567–1578.

45 N. J. Mitchell, S. S. Kulkarni, L. R. Malins, S. Wang and R. J. Payne, *Chem. Eur. J.*, 2017, **23**, 946–952.

46 S. Dery, P. S. Reddy, L. Dery, R. Mousa, R. N. Dardashti and N. Metanis, *Chem. Sci.*, 2015, **6**, 6207–6212.

47 H. Choi, M. Kim, J. Jang and S. Hong, *Angew. Chemie - Int. Ed.*, 2020, 2–10.

48 J. N. Rybak, S. B. Scheurer, D. Neri and G. Elia, *Proteomics*, 2004, **4**, 2296–2299.

49 X. Li, J. Yang and C. Rader, *Methods*, 2014, **65**, 133–138.

50 X. Li, C. G. Nelson, R. R. Nair, L. Hazlehurst, T. Moroni, P. Martinez-Acedo, A. R. Nanna, D. Hymel, T. R. Burke and C. Rader, *Cell Chem. Biol.*, 2017, **24**, 433-442.e6.

51 I. Nikic, J. H. Kang, G. E. Girona, I. V. Aramburu and E. A. Lemke, *Nat. Protoc.*, 2015, **10**, 780–791.

52 S. Smolskaya and Y. A. Andreev, *Biomolecules*, , DOI:10.3390/biom9070255.

53 C. B. Rosen and M. B. Francis, *Nat. Chem. Biol.*, 2017, **13**, 697–705.

54 S. Mädler, C. Bich, D. Touboul and R. Zenobi, *J. Mass Spectrom.*, 2009, **44**, 694–706.

55 M. Kanai, K. Oisaki, Y. Sohma, T. Ishiyama, D. Sasaki, Y. Seki and J. Abe, *J. Am. Chem. Soc.*, 2016, **138**, 10798–10801.

56 Y. Yu, L. K. Zhang, A. V. Buevich, G. Li, H. Tang, P. Vachal, S. L. Colletti and Z. C. Shi, *J. Am. Chem. Soc.*, 2018, **140**, 6797–6800.

57 M. Imiołek, G. Karunanithy, W. L. Ng, A. J. Baldwin, V. Gouverneur and B. G. Davis, *J. Am. Chem. Soc.*, 2018, **140**, 1568–1571.

58 S. Lin, X. Yang, S. Jia, A. M. Weeks, M. Hornsby, P. S. Lee, R. V. Nichiporuk, A. T. Iavarone, J. A. Wells, F. D. Toste and C. J. Chang, *Science (80-.).*, 2017, **355**, 597–602.

59 X. Chen, F. Ye, X. Luo, X. Liu, J. Zhao, S. Wang, Q. Zhou, G. Chen and P. Wang, *J. Am. Chem. Soc.*, 2019, **141**, 18230–18237.

60 Y. Gong, D. Andina, S. Nahar, J. C. Leroux and M. A. Gauthier, *Chem. Sci.*, 2017, **8**, 4082–4086.

61 I. Dovgan, S. Erb, S. Hessmann, S. Ursuegui, C. Michel, C. Muller, G. Chaubet, S. Cianférani and A. Wagner, *Org. Biomol. Chem.*, 2018, **16**, 1305–1311.

62 J. C. Vantourout, S. R. Adusumalli, K. W. Knouse, D. T. Flood, A. Ramirez, N. M. Padial, A. Istrate, K. Maziarz, J. N. deGruyter, R. R. Merchant, J. X. Qiao, M. A. Schmidt, M. J. Deery, M. D. Eastgate, P. E. Dawson, G. J. L. Bernardes and P. S. Baran, *J. Am. Chem. Soc.*, 2020, **142**, 17236–17242.

63 R. J. Spears, R. L. Brabham, D. Budhadev, T. Keenan, S. McKenna, J. Walton, J. A. Brannigan, A. M. Brzozowski, A. J. Wilkinson, M. Plevin and M. A. Fascione, *Chem. Sci.*, 2018, **9**, 5585–5593.

64 W. Gong, G. Zhang, T. Liu, R. Giri and J. Q. Yu, *J. Am. Chem. Soc.*, 2014, **136**, 16940–16946.

65 S. Bloom, C. Liu, D. K. Kölmel, J. X. Qiao, Y. Zhang, M. A. Poss, W. R. Ewing and D. W. C. Macmillan, *Nat. Chem.*, 2018, **10**, 205–211.

66 G. Z. Elek, K. Koppel, D. M. Zubrytski, N. Konrad, I. Järving, M. Lopp and D. G. Kananovich, *Org. Lett.*, 2019, **21**, 8473–8478.

67 J. E. Moore and W. H. Ward, *J. Am. Chem. Soc.*, 1956, **78**, 2414–2418.

68 J. M. J. M. Ravasco, H. Faustino, A. Trindade and P. M. P. Gois, *Chem. - A Eur. J.*, 2019, **25**, 43–59.

69 M. A. Kasper, M. Glanz, A. Stengl, M. Penkert, S. Klenk, T. Sauer, D. Schumacher, J. Helma, E. Krause, M. C. Cardoso, H. Leonhardt and C. P. R. Hackenberger, *Angew. Chemie - Int. Ed.*, 2019, **58**, 11625–11630.

70 M. J. Matos, C. D. Navo, T. Hakala, X. Ferhati, A. Guerreiro, D. Hartmann, B. Bernardim, K. L. Saar, I. Compañón, F. Corzana, T. P. J. Knowles, G. Jiménez-Osés and G. J. L. Bernardes, *Angew. Chemie - Int. Ed.*, 2019, **58**, 6640–6644.

71 G. J. L. Bernardes, E. J. Grayson, S. Thompson, J. M. Chalker, J. C. Errey, F. El Oualid, T. D. W. Claridge and B. G. Davis, *Angew. Chemie - Int. Ed.*, 2008, **47**, 2244–2247.

72 *Deriv. Keratin*, 1935, **112**, 361–371.

73 M. D. Simon, F. Chu, L. R. Racki, C. C. de la Cruz, A. L. Burlingame, B. Panning, G. J. Narlikar and K. M. Shokat, *Cell*, 2007, **128**, 1003–1012.

74 E. V. Vinogradova, C. Zhang, A. M. Spokoyny, B. L. Pentelute and S. L. Buchwald, *Nature*, 2015, **526**, 687–691.

75 K. Hanaya, J. Ohata, M. K. Miller, A. E. Mangubat-Medina, M. J. Swierczynski, D. C. Yang, R. M. Rosenthal, B. V. Popp and Z. T. Ball, *Chem. Commun.*, 2019, **55**, 2841–2844.

76 B. A. Vara, X. Li, S. Berritt, C. R. Walters, E. J. Petersson and G. A. Molander, *Chem. Sci.*, 2018, **9**, 336–344.

77 M. S. Messina, J. M. Stauber, M. A. Waddington, A. L. Rheingold, H. D. Maynard and A. M. Spokoyny, *J. Am. Chem. Soc.*, 2018, **140**, 7065–7069.

78 C. Bottecchia and T. Noël, *Chem. - A Eur. J.*, 2019, **25**, 26–42.

79 E. L. Tyson, Z. L. Niemeyer and T. P. Yoon, *J. Org. Chem.*, 2014, **79**, 1427–1436.

80 H. Choi, M. Kim, J. Jang and S. Hong, *Angew. Chemie - Int. Ed.*, 2020, **59**, 22514–22522.

81 Y. Wang, L. F. Deng, X. Zhang, Z. D. Mou and D. Niu, *Angew. Chemie - Int. Ed.*, 2020, 60, **4**, 2155-2159.

82 C. D. Spicer and B. G. Davis, *Nat. Commun.*, 2014, **5**, 1–14.

83 O. Boutureira and G. J. L. Bernardes, *Chem. Rev.*, 2015, **115**, 2174–2195.

84 B. J. Bruno, G. D. Miller and C. S. Lim, *Ther. Deliv.*, 2014, **4**, 1443–1467.

85 T. H. Wright, B. J. Bower, J. M. Chalker, G. J. L. Bernardes, R. Wiewiora, W.-L. Ng, R. Raj, S. Faulkner, M. R. J. Vallée and A. Phanumartwiwath, *Science (80-.).*, 2016, **354**, 1465.

86 A. Yang, S. Ha, J. Ahn, R. Kim, S. Kim, Y. Lee, J. Kim, D. Söll, H. Lee and H. Park, *Science (80-.).*, 2016, **354**, 1–5.

87 B. Josephson, C. Fehl, P. G. Isenegger, S. Nadal, T. H. Wright, A. W. J. Poh, B. J. Bower, A. M. Giltrap, L. Chen, C. Batchelor-McAuley, G. Roper, O. Arisa, J. B. I. Sap, A. Kawamura, A. J. Baldwin, S. Mohammed, R. G. Compton, V. Gouverneur and B. G. Davis, *Nature*, 2020, **585**, 530–537.

88 A. D. Baldwin and K. L. Kiick, *Bioconjug. Chem.*, 2011, **22**, 1946–1953.

89 L. Pedzisa, X. Li, C. Rader and W. R. Roush, *Org. Biomol. Chem.*, 2016, **14**, 5141–5147.

90 D. T. Cohen, C. Zhang, B. L. Pentelute and S. L. Buchwald, *J. Am. Chem. Soc.*, 2015, **137**, 9784–9787.

91 D. T. Cohen, C. Zhang, C. M. Fadzen, A. J. Mijalis, L. Hie, K. D. Johnson, Z. Shriver, O. Plante, S. J. Miller, S. L. Buchwald and B. L. Pentelute, *Nat. Chem.*, 2019, **11**, 78–85.

92 Y. A. Lin, O. Boutureira, L. Lercher, B. Bhushan, R. S. Paton and B. G. Davis, 2013, 4–7.

93 X. Yang, K. R. Lennard, C. He, M. C. Walker, A. T. Ball, C. Doigneaux, A. Tavassoli and W. A. Van Der Donk, *Nat. Chem. Biol.*, 2018, **14**, 375–380.

94 L. D. Walensky and G. H. Bird, *J. Med. Chem.*, 2014, **57**, 6275–6288.

95 N. S. Robertson, S. J. Walsh, E. Fowler, M. Yoshida, S. M. Rowe, Y. Wu, H. F. Sore, J. S. Parker and D. R. Spring, *Chem. Commun.*, 2019, **55**, 9499–9502.

96 A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y. S. Lin and B. L. Pentelute, *J. Am. Chem. Soc.*, 2013, **135**, 5946–5949.

97 Y. Wu, Y. H. Li, X. Li, Y. Zou, H. L. Liao, L. Liu, Y. G. Chen, D. Bierer and H. G. Hu, *Chem. Sci.*, 2017, **8**, 7368–7373.

98 J. E. Montgomery, J. A. Donnelly, S. W. Fanning, T. E. Speltz, X. Shangguan, J. S. Coukos, G. L. Greene and R. E. Moellering, *J. Am. Chem. Soc.*, 2019, **141**, 16374–16381.

99 D. Gang, D. W. Kim and H. S. Park, *Genes.*, 2018, 9, **11**, 557.

100 D. Meister, S. M. Taimoory and J. F. Trant, *Pept. Sci.*, 2019, **111**, 24058.

101 S. W. Millward, S. Fiacco, R. J. Austin and R. W. Roberts, *ACS Chem. Biol.*, 2007, **2**, 625–634.

102 H. S. Sader, G. E. Dale, P. R. Rhomberg and R. K. Flamm, *Antimicrob. Agents Chemother.*, 2018, **62**, 1–6.

103 W. Quispe-Tintaya, *Physiol. Behav.*, 2017, **176**, 139–148.

104 K. R. Lennard and A. Tavassoli, *Chem. - A Eur. J.*, 2014, **20**, 10608–10614.

105 T. Passioura and H. Suga, *Chem. Commun.*, 2017, **53**, 1931–1940.

106 Y. Goto, T. Katoh and H. Suga, *Nat. Protoc.*, 2011, **6**, 779–790.

107 A. Stefanucci, W. Lei, S. Pieretti, E. Novellino, M. P. Dimmito, F. Marzoli, J. M. Streicher and A. Mollica, *Sci. Rep.*, 2019, **9**, 1–13.

108 A. A. Aimetti, R. K. Shoemaker, C. C. Lin and K. S. Anseth, *Chem. Commun.*, 2010, **46**, 4061–4063.

109 V. R. Pattabiraman, J. L. Stymiest, D. J. Derksen, N. I. Martin and J. C. Vederas, *Org. Lett.*, 2007, **9**, 699–702.

110 Y. Zhang, Q. Zhang, C. T. T. Wong and X. Li, *J. Am. Chem. Soc.*, 2019, **141**, 12274–12279.

111 Y. Zou, A. M. Spokoyny, C. Zhang, M. D. Simon, H. Yu, Y. S. Lin and B. L. Pentelute, *Org. Biomol. Chem.*, 2014, **12**, 566–573.

112 L. Raynal, N. C. Rose, J. R. Donald and C. D. Spicer, *Chem. - A Eur. J.*, 2020, 1–21.

113 S. J. McCarver, J. X. Qiao, J. Carpenter, R. M. Borzilleri, M. A. Poss, M. D. Eastgate, M. M. Miller and D. W. C. MacMillan, *Angew. Chemie - Int. Ed.*, 2017, **56**, 728–732.

114 R. Jones, *Bone*, 2014, **23**, 1–7.

115 G. Sugiura, H. Kühn, M. Sauter, U. Haberkorn and W. Mier, *Molecules*, 2014, **19**, 2135–2165.

116 J. Ermert., *Bio. Med. Res.,* 2014, 2014, 15.

117 I. M. Jackson, P. J. H. Scott and S. Thompson, *Semin. Nucl. Med.*, 2017, **47**, 493–523.

118 N. Ichiishi, J. P. Caldwell, M. Lin, W. Zhong, X. Zhu, E. Streckfuss, H. Y. Kim, C. A. Parish and S. W. Krska, *Chem. Sci.*, 2018, **9**, 4168–4175.

119 T. Yamasaki, R. Azuma, K. Sano, M. Munekane, Y. Matsuoka, K. I. Yamada and T. Mukai, *ACS Med. Chem. Lett.*, 2020, **11**, 45–48.

120 J. Becaud, L. Mu, M. Karramkam, P. A. Schubiger, S. M. Ametamey, K. Graham, T. Stellfeld, L. Lehmann, S. Borkowski, D. Berndorff, L. Dinkelborg, A. Srinivasan, R. Smits and B. Koksch, *Bioconjug. Chem.*, 2009, **20**, 2254–2261.

121 L. Pedzisa, X. Li, C. Rader and W. R. Roush, *Org. Biomol. Chem.*, 2016, **14**, 5141–5147.

122 L. R. Malins, N. J. Mitchell, S. McGowan and R. J. Payne, *Angew. Chemie - Int. Ed.*, 2015, **54**, 12716–12721.

123 K. M. Reddy and G. Mugesh, *RSC Adv.*, 2019, **9**, 34–43.

124 J. Dadová, S. R. Galan and B. G. Davis, *Curr. Opin. Chem. Biol.*, 2018, **46**, 71–81.

125 M. M. Haugland, J. E. Lovett and E. A. Anderson, *Chem. Soc. Rev.*, 2018, **47**, 668–680.

126 C. Marestin, C. Noël, A. Guyot and J. Claverie, *Macromolecules*, 1998, **31**, 4041–4044.

127 T. Okada, T. Asawa, Y. Sugiyama, M. Kirihara, T. Iwai and Y. Kimura, *Synlett*, 2014, **25**, 596–598.

128 V. H. Nguyen, H. Nishino, S. Kajikawa and K. Kurosawa, *Tetrahedron*, 1998, **54**, 11445–11460.

129 V. Gupta, S. Bhavanasi, M. Quadir, K. Singh, G. Ghosh, K. Vasamreddy, A. Ghosh, T. J. Siahaan, S. Banerjee and S. K. Banerjee, *J. Cell Commun. Signal.*, 2019, **13**, 319–330.

130 L. H. Belén, C. De Oliveira Rangel-Yagui, J. F. Beltrán Lissabet, B. Effer, M. Lee-Estevez, A. Pessoa, R. L. Castillo and J. G. Farías, *Front. Pharmacol.*, 2019, **10**, 1–16.

131 M. Meldal and C. W. Tomøe, *Chem. Rev.*, 2008, **108**, 2952–3015.

132 C. S. McKay and M. G. Finn, *Chem. Biol.*, 2014, **21**, 1075–1101.

133 V. Hong, N. F. Steinmetz, M. Manchester and M. G. Finn, *Bioconjug. Chem.*, 2010, **21**, 1912–1916.

134 R. He, B. Finan, J. P. Mayer and R. D. DiMarchi, *Molecules*, 2019, **24**, 11–14.

135 T. Karampelas, E. Skavatsou, O. Argyros, D. Fokas and C. Tamvakopoulos, *Mol. Pharm.*, 2017, **14**, 674–685.

136 A. Ferreira, R. Lapa and N. Vale, *Biomolecules*, , DOI:10.3390/biom9110693.

137 J. S. Cheah and S. Yamada, *Biochem. Biophys. Res. Commun.*, 2017, **493**, 1522–1527.

138 P. R. Karhemo, S. Ravela, M. Laakso, I. Ritamo, O. Tatti, S. Mäkinen, S. Goodison, U. H. Stenman, E. Hölttä, S. Hautaniemi, L. Valmu, K. Lehti and P. Laakkonen, *J. Proteomics*, 2012, **77**, 87–100.

139 A. N. Kapanidis, Y. W. Ebright and R. H. Ebright, *J. Am. Chem. Soc.*, 2001, **123**, 12123–12125.

140 A. Hawe, M. Sutter and W. Jiskoot, *Pharm. Res.*, 2008, **25**, 1487–1499.

141 M. Sameiro and T. Gonçalves, *Chem. Rev.*, 2009, **109**, 190–212.

142 H. Gong, J. L. Kovar, B. Baker, A. Zhang, L. Cheung, D. R. Draney, I. R. Corrêa, M. Q.

Xu and D. M. Olive, *PLoS One*., 2011, 7, **3**, e34003.

143 C. Cilliers, I. Nessler, N. Christodolu and G. M. Thurber, *Mol. Pharm.*, 2017, **14**, 1623–1633.

144 C. D. Spicer and B. G. Davis, *Nat. Commun.*, 2014, 5, 1-14.

145 M. Friedman, A. Orlova, E. Johansson, T. L. J. Eriksson, I. Höidén-Guthenberg, V. Tolmachev, F. Y. Nilsson and S. Ståhl, *J. Mol. Biol.*, 2008, **376**, 1388–1402.

146 T. S. Chisholm, D. Clayton, L. J. Dowman, J. Sayers and R. J. Payne, *J. Am. Chem. Soc.*, 2018, **140**, 9020–9024.

147 J. T. Ge, Y. Y. Li, J. Tian, R. Z. Liao and H. Dong, *J. Org. Chem.*, 2017, **82**, 7008–7014.

148 X. F. Gao, J. J. Du, Z. Liu and J. Guo, *Org. Lett.*, 2016, **18**, 1166–1169.

149 M. Lee, S. Neukirchen, C. Cabrele and O. Reiser, *J. Pept. Sci.*, 2017, **23**, 556–562.

150 E. H. Discekici, S. L. Shankel, A. Anastasaki, B. Oschmann, I. H. Lee, J. Niu, A. J. McGrath, P. G. Clark, D. S. Laitar, J. R. De Alaniz, C. J. Hawker and D. J. Lunn, *Chem. Commun.*, 2017, **53**, 1888–1891.

151 B. M. R. D. B. (BMRDB) and http://www. bmrb. wisc. edu/-entrie. 68 and 4769., Biological Magnetic Resonance Data Bank (BMRDB) http://www.bmrb.wisc.edu/ - entries 68 and 4769.

152 T. Jenuwein and C. D. Allis, *Science (80-.).*, 2001, **293**, 1074–1080.

153 F. Li, A. Allahverdi, R. Yang, G. B. J. Lua, X. Zhang, Y. Cao, N. Korolev, L. Nordenskiöld and C. Liu, *Angew. Chem. Int. Ed.*, 2011, **123**, 9785–9788.

154 G. D. Bowman and M. G. Poirier, *Chem. Rev.*, 2015, **115**, 2274–2295.

155 J. Guo, J. Wang, J. S. Lee and P. G. Schultz, *Angew. Chemie - Int. Ed.*, 2008, **47**, 6399–6401.

156 M. M. Müller and T. W. Muir, *Chem. Rev.*, 2015, **115**, 2296–2349.

157 Y. Jing, Z. Liu, G. Tian, X. Bao, T. Ishibashi and X. D. Li, *Cell Chem. Biol.*, 2018, **25**, 166-174.e7.

158 T. Ishiguro, K. Tanabe, Y. Kobayashi, S. Mizumoto, M. Kanai and S. A. Kawashima, *Sci. Rep.*, 2018, **8**, 1–10.

159 Y. Chen, R. Sprung, Y. Tang, H. Ball, B. Sangras, S. C. Kim, J. R. Falck, J. Peng, W. Gu and Y. Zhao, *Mol. Cell. Proteomics*, 2007, **6**, 812–819.

160 F. Ohtake, Y. Saeki, K. Sakamoto, K. Ohtake, H. Nishikawa, H. Tsuchiya, T. Ohta, K.

Tanaka and J. Kanno, *EMBO Rep.*, 2015, **16**, 192–201.

161 T. P. Garner, J. Strachan, E. C. Shedden, J. E. Long, J. R. Cavey, B. Shaw, R. Layfield and M. S. Searle, *Biochemistry*, 2011, **50**, 9076–9087.

162 L. Herhaus and I. Dikic, *EMBO Rep.*, 2015, **16**, 1071–1083.

163 R. Huang, M. A. Holbert, M. K. Tarrant, S. Curtet, D. R. Colquhoun, B. M. Dancy, B. C. Dancy, Y. Hwang, Y. Tang, K. Meeth, R. Marmorstein, R. N. Cole, S. Khochbin and P. A. Cole, *J. Am. Chem. Soc.*, 2010, **132**, 9986–9987.

164 S. Bhat, Y. Hwang, M. D. Gibson, M. T. Morgan, S. D. Taverna, Y. Zhao, C. Wolberger, M. G. Poirier and P. A. Cole, *J. Am. Chem. Soc.*, 2018, 140, 9478–9485.

165 B. J. G. E. Pieters, J. C. J. Hintzen, Y. Grobben, A. H. K. Al Temimi, J. J. A. G. Kamps and J. Mecinović, *Bioconjug. Chem.*, 2019, **30**, 952–958.

166 J. M. Chalker, L. Lercher, N. R. Rose, C. J. Schofield and B. G. Davis, *Angew. Chemie - Int. Ed.*, 2012, **51**, 1835–1839.

167 T. H. Wright, B. J. Bower, J. M. Chalker, G. J. L. Bernardes, R. Wiewiora, W. L. Ng, R. Raj, S. Faulkner, M. R. J. Vallée, A. Phanumartwiwath, O. D. Coleman, M. L. Thézénas, M. Khan, S. R. G. Galan, L. Lercher, M. W. Schombs, S. Gerstberger, M. E. Palm-Espling, A. J. Baldwin, B. M. Kessler, T. D. W. Claridge, S. Mohammed and B. G. Davis, *Science.,* 2016, 354, **6312**, 1465.

168 L. Marzo, S. K. Pagire, O. Reiser and B. König, *Angew. Chemie - Int. Ed.*, 2018, **57**, 10034–10072.

169 M. J. Matos, B. L. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. M. S. D. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. J. Deery, J. M. Chalker, F. Corzana, G. Jiménez-Osés and G. J. L. Bernardes, *J. Am. Chem. Soc.*, 2018, **140**, 4004–4017.

170 T. J. Barker and D. L. Boger, *J. Am. Chem. Soc.*, 2012, **134**, 13588–13591.

171 F. Yin, Z. Wang, Z. Li and C. Li, *J. Am. Chem. Soc.*, 2012, **134**, 10401–10404.

172 D. W. Lee and B. S. Kim, *Plant Pathol. J.*, 2015, **31**, 1–11.

173 J. Iegre, N. S. Ahmed, J. S. Gaynord, Y. Wu, K. M. Herlihy, Y. S. Tan, M. E. Lopes-Pires, R. Jha, Y. H. Lau, H. F. Sore, C. Verma, D. H. O'Donovan, N. Pugh and D. R. Spring, *Chem. Sci.*, 2018, **9**, 4638–4643.

174 M. Moiola, M. G. Memeo and P. Quadrelli, *Molecules*., 2019, 24, **20**, 3654.

175 M. B. Sims, J. J. Lessard, L. Bai and B. S. Sumerlin, *Macromolecules*, 2018, **51**, 6380–6386.

176 M. Kunishima, C. Kawachi, J. Morita, K. Terao, F. Iwasaki and S. Tani, *Tetrahedron*, 1999, **55**, 13159–13170.