LOCATING BINDING SITES OF THE GENERAL ANAESTHETIC PROPOFOL USING PHOTOAFFINITY LABELLING

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

and the Diploma of the Imperial College

by

Grace Mei Shan Yip

January 2013

Biophysics Section
Division of Cell and Molecular Biology
Department of Life Sciences
Sir Ernst Chain Building
Imperial College London
South Kensington
London
SW7 2AZ

Abstract

3-(5′-isopropyl-6′-hydroxyphenyl)-3-trifluoromethyl-diazirine (photofol), a new drug developed in our laboratory, is structurally homologous to the general anaesthetic propofol with one of its isopropyl groups being substituted for a trifluoromethyl-diazirine group. Photofol was shown to act in a similar manner to propofol, albeit with a three-fold reduction in potency on inducing a loss of righting reflex in rats and with reduced modulation of GABA-evoked currents on $\alpha_1\beta_2\gamma_{2S}$ $GABA_A$ receptors for the concentration tested.

The soluble protein human serum albumin (HSA) and the pentameric ligand-gated ion channel GLIC from the prokaryote *Gloeobacter violaceus* were used to test the utility of photofol as a photoaffinity label. Both proteins have been crystallised with propofol to produce high-resolution X-ray structures (Bhattacharya *et al.*, 2000; Nury *et al.*, 2011). Two samples of the photolabel—one containing naturally occurring hydrogen atoms and the other a bespoke proportion of different isotopes with between zero and two deuterium atoms—were used and this novel approach was successfully demonstrated to aid the analysis of mass spectra through their differences in isotopic distribution.

Photolabelling HSA and GLIC followed by enzyme cleavage, liquid chromatography-mass spectrometry and tandem mass spectrometry identified the binding of photofol to six different amino acids. Photofol labelled all three propofol-binding sites of HSA identified by X-ray crystallography. Peptides that are part of the propofol-binding site in the GLIC transmembrane domain reduced in ionisation intensity upon photolabelling, suggesting that photofol had bound to this vicinity. Furthermore, photolabelling of binding sites at the interfaces between subunits in the extracellular domain was reproducibly observed, particularly at lysine-183. Electrophysiology-coupled mutagenesis studies are currently underway to confirm propofol binding to this site, where binding to another general anaesthetic, ketamine, has recently been shown by high-resolution structures (Pan *et al.*, 2012a). In conclusion, the work presented herein indicates that photofol is a suitable photolabel for locating binding sites of propofol.

Acknowledgements

I would first like to thank Nick Franks for providing me with the opportunity to work in his lab during my MRes degree and subsequently to pursue a PhD under his supervision. I must also thank Ed Smith and Chris Edge for their hard work over many years in perfecting a photolabel and developing photofol, thus making this project possible. I am most grateful to Nick, Ed and Chris for their guidance and support throughout my degree.

Many thanks also go to Mark Bennett for teaching me tandem mass spectrometry and assisting me with running samples; to Robert Dickinson, Chris Edge and Thom Gent for carrying out the experiments on the biological characterisation of photofol; to Damian Droste for his company in the lab whilst he worked with a different photolabel; and to crystallographers Stephen Curry, Erhard Hohenester and Peter Brick for their assistance with using structural model software programs.

I wish to also thank those who have collaborated on this project: Pierre-Jean Corringer and Ludovic Sauguet for providing the GLIC proteins; Andrew Jenkins for carrying out mutagenesis studies on GLIC; and Alex Evers' group for extending the use of photofol to label $GABA_A$ receptors.

Thanks are also due to all members of the Biophysics group for their friendship, advice and encouragement, and to the Biotechnology and Biological Sciences Research Council for providing me with funding through my postgraduate studies.

Finally, I would like to express my deep gratitude to my parents, my brother and James for their continued love, support and patience, which have been immensely appreciated.

Declaration of originality

The contents of this thesis are entirely my own work and have not been submitted in any form for another degree. All information derived from the work of others has been acknowledged in the text and a list of references is given.

Table of contents

Abstract **2**

Acknowledgements **3**

Declaration of originality **3**

Table of contents **4**

Table of figures **7**

Acronyms and abbreviations **13**

1 Introduction **15**

1.1 General anaesthetics and their molecular targets 15

1.2 Photoaffinity labelling 23

1.3 Analysis of proteins and peptides using mass spectrometry 28

1.4 Protein targets in this photoaffinity labelling study 34

1.4.1 Human serum albumin 34

1.4.2 *Gloeobacter violaceus* ligand-gated ion channel 38

1.5 Aims of this project 44

2 Materials and methods **45**

2.1 Physical characterisation of the photolabels 45

2.1.1 UV absorbance 45

2.1.2 Hydrolysis and photolysis 45

2.2 Biological characterisation of the photolabels 46

2.2.1 Loss of righting reflex in rats 46

2.2.2 Whole-cell patch clamping 47

2.3 Photolabelling experiments with human serum albumin 49

2.3.1 Light exposure of HSA 49

2.3.2 Denaturation of proteins and modification of cysteine residues 49

2.3.3 Trypsin-digestion of proteins 49

2.4 Photolabelling experiments with *Gloeobacter violaceus* ligand-gated ion channel 50

2.4.1 Light exposure of GLIC 50

2.4.2 Denaturing and digesting GLIC 51

2.4.2.1 Treatment A: denatured in guanidine hydrochloride and digested with chymotrypsin 52

2.4.2.2 Treatment B: photolabelled at pH 8, denatured in SDC and digested with chymotrypsin 52
2.4.2.3 Treatment C: photolabelled at pH 4, denatured in SDC and digested with chymotrypsin 53
2.4.2.4 Treatment D: photolabelled at pH 8, denatured in urea and RapiGest and digested with Lys-C followed by chymotrypsin 54
2.5 High pressure liquid chromatography-mass spectrometry 55
2.5.1 Liquid chromatography-mass spectrometry 55
2.5.2 Tandem mass spectrometry 58

3 Characterisation of the photolabel 60
3.1 Physical characterisation of propofol and photofol 60
3.1.1 UV-Vis absorption profiles 60
3.1.2 Aqueous stability and photolysis in ethanol 62
3.2 Biological characterisation of the photolabels 65
3.2.1 Loss of righting reflex in rats 66
3.2.2 Whole-cell patch clamping in $GABA_A$ receptor-expressing HEK cells 67
3.3 Concluding remarks for the characterisation of photofol 68

4 Photoaffinity labelling of human serum albumin 70
4.1 Identification of tryptic digests of HSA from HPLC-MS 71
4.2 The photolabelling experiment samples 75
4.3 Identifying labelled HSA peptides 77
4.3.1 Identifying the retention times of labelled peptides 78
4.3.2 Determining the masses of labelled peptides 81
4.4 Selected ion monitoring scans and fraction collections 86
4.5 Predicting labelled peptide sequences based on molecular weight 88
4.6 Sequencing labelled peptides by tandem mass spectrometry 90
4.6.1 Labelled peptide LKCASLQK 91
4.6.2 Lysine-labelled peptides 94
4.6.3 Labelled peptide YTK 97
4.6.4 Labelled peptide FPK 99
4.6.5 Labelled peptide HKPK 101
4.6.6 Labelled peptide AWAVAR 104
4.7 Isotopic peak distributions of labelled peptides 107
4.8 Labelled amino acids in the HSA structure 109

4.9 Concluding remarks for the photoaffinity labelling of HSA....114

5 Photoaffinity labelling of *Gloeobacter violaceus* ligand-gated ion channel 118

5.1 Identification of peptides from digested GLIC by HPLC-MS 119

5.1.1 Treatment A with guanidine hydrochloride and chymotrypsin 121

5.1.2 Treatment B with sodium deoxycholate and chymotrypsin 121

5.2 Photolabelling GLIC....124

5.2.1 Treatment B with sodium deoxycholate and chymotrypsin post- photolabelling at pH 8 125

5.2.2 Treatment C with sodium deoxycholate and chymotrypsin post-photolabelling at pH 4 141

5.2.3 Treatment D with urea, RapiGest, Lys-C and chymotrypsin post-photolabelling at pH 8 147

5.3 The labelled amino acids of GLIC 152

5.4 The potential labelled peptides of GLIC 158

5.5 Concluding remarks for the photoaffinity labelling of GLIC 160

6 Overall conclusion 164

References 166

Appendix 180

Table of figures

Figure 1.1: Schematic depicting the architecture of a eukaryote Cys-loop ligand-gated ion channel 17

Figure 1.2: Chemical structures, molecular volumes and anaesthetic potencies of ten propofol analogues 19

Figure 1.3: Behavioural responses to etomidate and propofol in wildtype and $GABA_A$ β_3N265M knock-in mice 21

Figure 1.4: Schematic showing the principal stages of photoaffinity labelling 24

Figure 1.5: Chemical structures of the native anaesthetic propofol and its diazirine photolabel analogue photofol 25

Figure 1.6: Isotopic distributions of molecules with one, ten and 100 carbon atoms 30

Figure 1.7: Schematic showing the principle of tandem mass spectrometry 31

Figure 1.8: Schematic diagram showing cleavages of a peptide at various bonds along the peptide backbone by collision-induced dissociation 32

Figure 1.9: Isotopic distributions of the pPD and dPD samples of photofol 33

Figure 1.10: The binding of endogenous and exogenous ligands to different sites on HSA as defined by crystallographic studies 35

Figure 1.11: X-ray structure of propofol-bound HSA 36

Figure 1.12: Schematic topology of the subunits of prokaryote and eukaryote pentameric ligand-gated ion channels 39

Figure 1.13: X-ray structure of propofol-bound GLIC and functional relevance of the binding site 42

Figure 1.14: Predicted propofol-binding sites on GLIC by computational methods 43

Figure 2.1: Elution profiles for eluents containing either triflouroacetic acid or formic acid 56

Figure 3.1: UV absorbance spectra of propofol 61

Figure 3.2: UV absorbance spectra of photofol 61

Figure 3.3: UV absorbance spectra overlay showing the aqueous stability of photofol 63

Figure 3.4: Reaction scheme showing the photolysis of photofol in ethanol 63

Figure 3.5: UV absorbance spectra overlay showing the photolysis of photofol 64

Figure 3.6: Absorbance at 350 nm of photofol and propofol as a function of irradiation duration 65

Figure 3.7: Quantal dose-response curves for the loss of righting reflex of rats administered with propofol or photofol via tail-vein injection 66

Figure 3.8: Typical current traces showing the EC_7 GABA concentration being potentiated by approximately 100 % by the co-application of 5 μM photofol and EC_7 GABA..... 67

Figure 4.1: Principal steps in photoaffinity labelling experiments of proteins..... 70

Figure 4.2: Typical absorbance chromatograms at 214 and 280 nm and the total ion current chromatogram monitoring the separation of HSA tryptic peptides..... 72

Figure 4.3: Extracted ion current chromatograms for the singly, doubly and triply charged ions of the peptide with a MW of 1370.6 Da..... 73

Figure 4.4: Identification of the elution time of $_1DAHK_4$ with a MW of 469.2 Da by extracting ions of its singly and doubly charged ions followed by investigating the charged state of its *m*/*z* peaks..... 74

Figure 4.5: Chromatogram overlays of absorbance at 214 and 280 nm and total ion current for the separation of tryptic peptides of HSA that was and was not exposed to UV light..... 76

Figure 4.6: Chemical structures of the carbene and quinone methide reactive intermediates of photofol..... 77

Figure 4.7: Chromatogram overlays of absorbance at 214 and 280 nm and total ion current for the separation of tryptic peptides of HSA from protein that was exposed to UV light in the presence of ethanol, pPD and dPD..... 79

Figure 4.8: Absorbance spectra at 214 and 280 nm of tryptic peptides of HSA from protein that was exposed to UV light in the presence of ethanol, pPD and dPD 80

Figure 4.9: Averaged mass spectra overlay at 54.3–55.0 minutes of the HSA samples 82

Figure 4.10: Averaged mass spectra overlay at 68.4–69.4 minutes of the HSA samples 83

Figure 4.11: Averaged mass spectra overlay at 52.0–53.0 minutes of the HSA samples 84

Figure 4.12: Averaged mass spectra overlays at various retention times when labelled HSA peptides were eluted 86

Figure 4.13: Overlays of the SIM scans ion current chromatograms for the unlabelled and the pPD-labelled sample..... 87

Figure 4.14: Summary of the *in silico* sequence predictions of labelled HSA peptides 89

Figure 4.15: Enhanced product ion mass spectra for the fragmentation of the parent ions 582.3 and 582.8 *m/z* for the HSA samples exposed to light in the presence of pPD and dPD....92

Figure 4.16: The expected *m/z* of the singly charged CID ions of the peptide LKCASLQK....93

Figure 4.17: Schematic of the peptide $_{198}$LKCASLQK$_{205}$ with annotations showing the labelled CID fragments identified in the mass spectra....93

Figure 4.18: Chemical structures of labelled-bound internal ions of lysine....95

Figure 4.19: Enhanced product ion spectra for the fragmentation of the of the parent ions 517.4 and 517.9 *m/z* for the pPD- and dPD-labelled samples....96

Figure 4.20: The expected *m/z* of the singly charged CID ions of the peptide YTK....97

Figure 4.21: Enhanced product ion mass spectra for the fragmentation of the parent ions 314.2 and 314.7 *m/z* for the pPD- and dPD-labelled samples....98

Figure 4.23: The expected *m/z* of the singly charged CID ions of the peptide FPK....99

Figure 4.22: Schematic of the peptide $_{411}$YTK$_{413}$ with annotations showing the labelled CID fragments identified in the mass spectra....99

Figure 4.24: Enhanced product ion mass spectra for the fragmentation of the parent ions 304.2 and 304.7 *m/z* for the pPD- and dPD-labelled samples....100

Figure 4.25: Schematic of the peptide $_{223}$FPK$_{225}$ highlighting the only labelled fragment, a_1^, that can be observed on the mass spectra....101

Figure 4.26: The expected *m/z* of the singly charged CID ions of the peptide HKPK....102

Figure 4.27: Enhanced product ion mass spectra for the collision-induced dissociation of the labelled-HKPK peptide....103

Figure 4.28: Schematic of the peptide $_{535}$HKPK$_{538}$ with annotations showing the labelled CID fragments observed in the mass spectra....104

Figure 4.29: Enhanced product ion mass spectra for the fragmentation of the parent ions 445.1 and 445.6 *m/z* for the pPD- and dPD-labelled samples....105

Figure 4.30: The expected *m/z* of the singly charged CID ions of the peptide AWAVAR....106

Figure 4.31: Schematic of the peptide $_{213}$AWAVAR$_{218}$ with annotations showing the labelled CID fragments identified in the mass spectra....107

Figure 4.32: Calculated and observed isotopic distributions for the pPD- and dPD-labelled peptides, YTK, KQTALVELVK and LKCASLQK....108

Figure 4.33: Summary of the identified photofol-bound amino acids within their labelled peptides and their locations in known binding sites of HSA 110

Figure 4.34: Propofol site 1 of HSA showing the proximity of the amino acid tyrosine-411 to the bound propofol in the X-ray structure 111

Figure 4.35: Propofol site 2 of HSA showing the proximities of the amino acids lysine-525, histidine-535, lysine-536 and lysine-541 to the bound propofol in the X-ray structure 112

Figure 4.36: Propofol site 3 of HSA showing the proximities of the amino acids lysine-195, lysine-199, lysine-212, tryptophan-214 and phenylalanine-223 to the bound propofol in the X-ray structure 113

Figure 5.1: Bar charts showing the frequencies of peptides of various amino acid lengths from the hypothetical cleavage of GLIC with chymotrypsin 120

Figure 5.2: Schematic showing the amino acid coverage of GLIC from treatment A 122

Figure 5.3: 214-nm absorbance spectra and total ion current chromatograms of GLIC peptides from chymotrypsin-digestion after denaturation of the protein with either guanidine hydrochloride or sodium deoxycholate 123

Figure 5.4: Schematic showing the amino acid coverage of GLIC from treatment B 124

Figure 5.5: Total ion current, 214-nm and 280-nm absorbance chromatograms of GLIC that was exposed to UV light in the presence of ethano, pPD and dPD prior to post-labelling treatment B 126

Figure 5.6: Averaged mass spectra overlays at three retention time ranges, when labelled GLIC peptides from the photolabelling treatment B experiment were eluted 128

Figure 5.7: Enhanced product ion mass spectra for the collision-induced dissociation of parent ions of 582.3 and 582.8 m/z confirming the sequence of this labelled peptide as being $_{166}$TAVVKPANF$_{174}$ 129

Figure 5.8: Enhanced product ion mass spectra for the fragmentation of the parent ions 834.4 and 834.7 m/z confirming the sequence of this labelled peptide as being $_{175}$ALEDRLESKLDY$_{186}$ 130

Figure 5.9: Enhanced product ion mass spectra of the unlabelled peptides $_{134}$SVDTRNIVL$_{142}$ and $_{175}$ALEDRLESKLDY$_{186}$ 132

Figure 5.10: Total ion current, and 214- and 280-nm absorbance chromatograms of GLIC from the second treatment B photolabelling experiment with the peptide samples separated by eluents containing FA instead of TFA 134

Figure 5.11: Averaged mass spectra overlays at 96.0–97.0 minutes and 97.8–98.8 minutes for peaks C and D of Figure 5.10 respectively 135

Figure 5.12: Bar chart showing the percentage changes in ionisation of twelve peptides from the second GLIC treatment B photolabelling experiment 137

Figure 5.13: Total ion current of GLIC peptide samples from the second and third GLIC treatment B photolabelling experiments 138

Figure 5.14: Bar chart showing percentage changes in ionisation of twelve peptides from the third GLIC treatment B photolabelling experiment 140

Figure 5.15: Chromatograms overlay of the TIC and absorbance at 214 and 280 nm of GLIC peptide samples from the GLIC treatment C photolabelling experiment 142

Figure 5.16: Averaged mass spectra overlay at 90–93 minutes for GLIC samples from the photolabelling treatment C experiment 143

Figure 5.17: Enhanced product ion mass spectra overlays for the fragmentation of the parent ions 931.2 and 931.7 *m/z,* and 854.2 and 855.2 *m/z*of GLIC treatment C samples labelled with pPD and dPD 144

Figure 5.18: Schematic showing the amino acids of GLIC that are part of peptides sequenced from EPI spectra with a confidence score of greater than 50 147

Figure 5.19: Chromatograms overlay of the absorbance at 214 and 280 nm and TIC of GLIC that were exposed to UV light in the presence of ethanol, pPD and dPD at pH 7.8 prior to post-labelling treatment D 149

Figure 5.20: Overlays of averaged mass spectra for samples from GLIC treatment D experiments being eluted at 92.5–96.0 minutes and at 120–145 minutes 150

Figure 5.21: Table summarising the *m/z* peaks that are present in the labelled GLIC samples but absent in the unlabelled sample from the analysis of averaged mass spectra at various retention time ranges 151

Figure 5.22: Labelled amino acids cysteine-27, lysine-170 and lysine-183 in the extracellular domain of GLIC superimposed onto the coordinates of a propofol molecule placed at a potential extracellular propofol-binding cavity 154

Figure 5.23: The potential extracellular binding cavities at the interface between subunits for propofol-bound GLIC, native GLIC and native ELIC 156

Figure 5.24: Alignment of two extracellular domains of GLIC subunits with acetylcholine binding protein complexed with two agonists and two antagonists 156

Figure 5.25: Crystal structure of ketamine-bound GLIC 157

Figure 5.26: Molecular models showing the three photolabelled amino acids K170, K183 and C27 on the surface of the ketamine binding site....158

Figure 5.27: Potential labelled peptides of GLIC that have a reduction in ionisation intensity upon photolabelling with photofol....159

Figure A.1: Enhanced product ion spectra for two peptides with molecular weights of 516.3 Da eluted at 16.8 and 18.3 minutes respectively....180

Figure A.2: Retention times of HSA tryptic peptides eluted from the HPLC column....181

Figure A.3: Enhanced product ion spectra for labelled peptide $_{525}$KQTALVELVK$_{534}$....182

Figure A.4: Enhanced product ion spectra for labelled peptide $_{199}$ASSAKQR$_{205}$....183

Figure A.5: Enhanced product ion spectra for labelled peptide $_{210}$AFKAWAVAR$_{218}$....184

Figure A.6: Table showing the extract ions analysis of GLIC extracellular domain peptides from treatment A at various durations of chymotrypsin digestion....185

Figure A.7: Enhanced product ion spectra for the peptide with a MW of 1168.7 Da eluted at 31.8 minutes....186

Figure A.8: Enhanced product ion spectra for the peptide with a MW of 517.2 Da eluted at 36.2 minutes....186

Figure A.9: Enhanced product ion spectra for the peptide with a MW of 1375.9 Da eluted at 37.6 minutes....187

Figure A.10: Enhanced product ion spectra for the peptide with a MW of 1062.7 Da eluted at 39.7 minutes....187

Figure A.11: Enhanced product ion spectra for the peptide with a MW of 1594.1 Da eluted at 49.8 minutes....188

Figure A.12: Enhanced product ion spectra for the peptide with a MW of 670.3 Da eluted at 52.0 minutes....188

Figure A.13: Enhanced product ion spectra for the peptide with a MW of 1172.4 Da eluted at 52.8 minutes....189

Figure A.14: Enhanced product ion spectra for the peptide with a MW of 1450.9 Da eluted at 56.8 minutes....190

Figure A.15: Enhanced product ion spectra for the peptide with a MW of 1500.8 Da eluted at 63.4 minutes....190

Figure A.16: A list of peptides that match the molecular weights within 1.0 Da of the peaks that correspond to ions of labelled peptides....191

Acronyms and abbreviations

aa	amino acid position(s)
AChBP	acetylcholine binding protein
ACN	acetonitrile
BSA	bovine serum albumin
CID	collision-induced dissociation
D_n	photofol with *n* deuterium atoms substituted for protium atoms
dPD	partially deuterated propofol diazirine
DDM	dodecylmaltoside
DTT	dithiothreitol
ED_{50}	equilibrium ligand dosage at half-maximum effect
ELIC	*Erwinia chrysanthemi* ligand-gated ion channel
EPI	enhanced product ion
ESI	electrospray ionisation
EtOH	ethanol
etomidate	ethyl 1-(1-phenylethyl)-1*H*-imidazole-5-carboxylate
FA	formic acid
GABA	γ-aminobutyric acid
$GABA_A$	γ-aminobutyric acid type A
GLIC	*Gloeobacter violaceus* ligand-gated ion channel
GnHCl	guanidine hydrochloride
HEK	human embryonic kidney
HPLC	high pressure liquid chromatography
HPLC-MS	high pressure liquid chromatography-mass spectrometry
HSA	human serum albumin
hν	light energy
IDA	information-dependent acquisition
LORR	loss of righting reflex
m/*z*	mass-to-charge ratio(s)
MS-MS	tandem mass spectrometry

MW	molecular weight
nACh	nicotinic acetylcholine
NMDA	N-methyl-D-aspartic
OD	optical density
PBS	phosphate-buffered saline
photofol	3-(5′-isopropyl-6′-hydroxyphenyl)-3-trifluoromethyl-diazirine
pPD	protiated propofol diazirine
propofol	2,6-diisopropylphenol
SDC	sodium deoxycholate
SE	standard error
SIM	selected ion monitoring
TIC	total ion current
TFA	trifluoroacetic acid
UV	ultraviolet
UV-Vis	ultraviolet-to-visible
XIC	extracted ion current

1 Introduction

1.1 General anaesthetics and their molecular targets

General anaesthetics are drugs that bring about a reversible loss of consciousness. The first successful public demonstration of the use of anaesthesia in surgery was on 15 October 1846, when dentist Dr William Thomas Green Morton and surgeon Dr John Collins Warren performed a five-minute procedure to remove a tumour from Mr Gilbert Abbott at the Massachusetts General Hospital in Boston, USA (Robinson and Toledo, 2012). Since then, general anaesthetics have been commonly used to induce anaesthesia. Anaesthesia has revolutionised the field of surgery by eliminating the element of pain, transforming the typically quick and brutal procedures prior to its use into the more composed and precise procedures we associate with modern surgery today.

Despite the frequent usage of general anaesthetics for more than 160 years, we are only just beginning to elucidate the molecular mechanisms that underlie their actions. As early as 1952, the notion that general anaesthetics affect neuronal signal transmissions came to light with studies on mammalian sympathetic ganglia showing that anaesthetics at clinically relevant concentrations depressed synaptic transmissions with little effects on action potential propagation (Larrabee and Posternak, 1952). Subsequently in 1964, Matthews and Quilliam (1964) provided the first direct evidence of general anaesthetics suppressing the secretion of acetylcholine in response to pre-ganglionic sympathetic nerves stimulation.

Prior to the 1980s, anaesthetics were thought to act by disrupting the lipid bilayer, which in turn caused configurational changes to their target proteins embedded in the membrane. At the time, this theory was convincing due to the observation that the potency of general anaesthetic agents correlated well with their lipid solubility over several orders of magnitude (Meyer, 1899 and Overton, 1901 cited in Antkowiak, 2001 and Kopp Lugli *et al.*, 2009), a phenomenon termed the Meyer-Overton correlation. However, evidence opposing this lipid theory began to accumulate in the 1980s. Compelling arguments include the inhibition of the soluble protein firefly luciferase by clinically relevant concentrations of general anaesthetics due to competition for the binding site with the natural ligand (Franks and Lieb, 1984), and the discovery of many compounds that were expected to be anaesthetics due to their lipid solubility but did not induce anaesthesia or induced

at a much higher concentration than those used clinically (Koblin *et al.*, 1994). Other evidence came from the abrupt decrease in anaesthetic potency within a series of homologous anaesthetics with increasing lipid solubility once a certain molecular dimension is exceeded, known as 'cut-off' effects (Franks and Lieb, 1985), as well as the discovery that stereoisomers of anaesthetic compounds have different potencies despite being equally soluble in lipids (Franks and Lieb, 1991).

Today, ion channels are thought to be the most important targets of general anaesthetics due to their role in neuronal signal transmissions. Three likely targets, all with evidence from *in vivo* models correlating with data from *in vitro* studies, are γ-aminobutyric acid type A ($GABA_A$) receptors, the N-methyl-D-aspartic acid (NMDA) receptors and the two-pore potassium (2PK) channels (Franks, 2008). Other ion channels that have also been shown to be modulated by general anaesthetics include glycine receptors (Downie *et al.*, 1996), nicotinic acetylcholine (nACh) receptors (Violet *et al.*, 1997), 5-HT_3 receptors (Machu and Harris, 1994) and the *Gloeobacter violaceus* ligand-gated ion channel (GLIC) (Weng *et al.*, 2010), to name a few. $GABA_A$ receptors will be discussed in this section and an overview of the GLIC protein—one of the target proteins for the photoaffinity labelling experiments with a photoactivatable propofol analogue in this project—is presented in section 1.4.2.

$GABA_A$ receptors are part of the Cys-loop family member of ligand-gated ion channels. They are so named due to their subunits each comprising a characteristic loop that contains a disulfide bond linking two cysteine residues. Five subunits, from a selection of at least 19 that have been cloned so far (α_1–α_6, β_1–β_3, γ_1–γ_3, δ, ε, θ, π, ρ_1–ρ_3) (Luscher *et al.*, 2011), are arranged around a central ion channel. Each subunit consists of a large N-terminal extracellular domain, four α-helix transmembrane domains (TM1–4) with a cytoplasmic loop of variable length linking TM3 and TM4, and an extracellular C-terminal that is usually short (Figure 1.1). Despite the wide range of combinations of subunits possible, the vast majority of neuronal $GABA_A$ receptors are $\alpha_1\beta_2\gamma_2$, $\alpha_2\beta_3\gamma_2$ or $\alpha_3\beta_{1-3}\gamma_2$ (Farrant and Nusser, 2005; Franks, 2008), with $\alpha_1\beta_2\gamma_2$ being the most abundant, contributing to 43 % of $GABA_A$ receptors in the rat brain (McKernan and Whiting, 1996). This subfamily of ion channels also includes the nACh receptors, glycine receptors and 5-HT_3 receptors. $GABA_A$ receptors are opened by γ-aminobutyric acid (GABA) binding to the extracellular domain

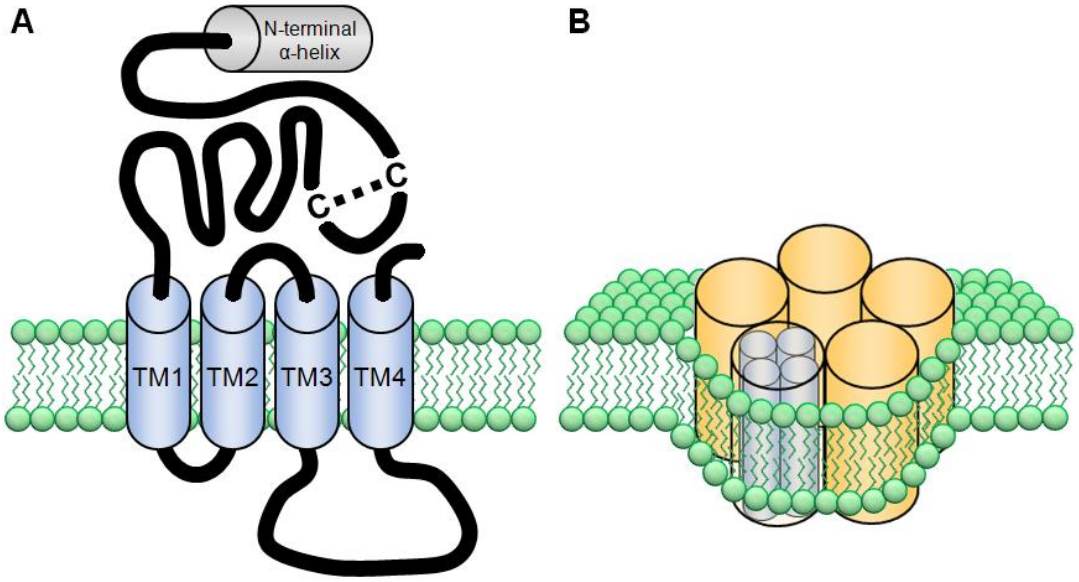


Figure 1.1: Schematic depicting the architecture of a eukaryote Cys-loop ligand-gated ion channel. A typical subunit consists of a large extracellular domain, four transmembrane helices (TM1–4) and a long cytoplasmic loop between TM3 and TM4 (A). The transmembrane domains of five such subunits are arranged around a central core with TM2 lining the ion pore (B).

between the β/α subunit interface. The opened pore passes chloride ions, the entry of which into the cell leads to a hyperpolarisation of the cell membrane, reducing the likelihood of action potential in the post-synaptic neurone, thus impeding excitatory transmission.

Psychotherapeutic drugs such as benzodiazepines are in wide clinical use to enhance the function of $GABA_A$ receptors so as to inhibit neuronal signals (Möhler *et al.*, 2002). These drugs modulate $GABA_A$ receptors by binding to an allosteric site at the α/γ subunit interface of the extracellular domain. The location of binding was deduced from the photoaffinity labelling of histidine-102 at the α/γ subunit interface with [^{3}H]-flunitrazepam (Duncalfe *et al.*, 1996; McKernan *et al.*, 1998), and by knock-out mice that lacked the gene encoding the $GABA_A$ γ_2 subunit being insensitive to the sedative and hypnotic effects of benzodiazepines (Günther *et al.*, 1995), highlighting the importance of subunit requirements of $GABA_A$ receptors for the binding of certain drugs.

$GABA_A$ receptors are thought to be major anaesthetic targets due to being positively modulated by most general anaesthetics, with the exception of gaseous anaesthetics, which primarily modulate NMDA receptors (Yamakura and Harris, 2000). The potentiation of the inhibitory ligand-gated ion channel by anaesthetics results in neuronal signal depression, in line with the general mechanism of general anaesthesia. Further evidence that supports $GABA_A$ channels being an

anaesthetic target includes the correlation of the stereoselectivity of anaesthetics in *in vivo* and *in vitro* studies (Tomlin *et al.*, 1999; Dickinson *et al.*, 2002), and the observations that the anaesthetic effects can be reversed by $GABA_A$ receptor antagonists and enhanced by $GABA_A$ receptor agonists (Tanelian *et al.*, 1993).

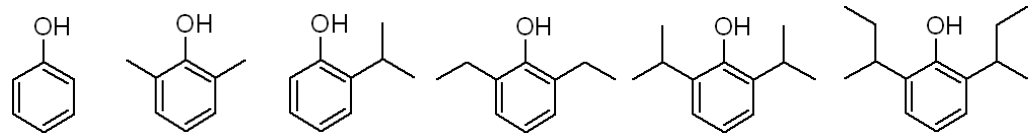
The intravenous anaesthetics 2,6-diisopropylphenol (propofol) and ethyl 1-(1-phenylethyl)-1*H*-imidazole-5-carboxylate (etomidate) are relatively selective for $GABA_A$ receptors (Vanlersberghe and Camu, 2008). Propofol has been shown to modulate the voltage-gated potassium channels (Friederich and Urban, 1999), voltage-gated sodium channels (Lingamaneni *et al.*, 2001) and L-type calcium channels (Guertin and Hounsgaard, 1999). However, Jurd *et al.* (2003) speculated that these channels only play a minor role, if any, towards inducing anaesthesia by propofol. Furthermore, Franks and Lieb (1994) disregarded the voltage-gated ion channels being likely anaesthetic targets due to their minimal modulation by clinically relevant concentrations of general anaesthetics.

Propofol has been shown to enhance the opening of $GABA_A$ receptors via two mechanisms: the potentiation of the GABA-evoked current at clinically relevant concentrations and a direct activation of $GABA_A$ receptors with higher concentrations (Hales and Lambert, 1991; Jones *et al.*, 1995). Propofol-binding was shown to not interfere with the benzodiazepine binding site as it modulated channels consisting of only α_2 and/or β_1 subunits (Jones *et al.*, 1995) and directly activated homomeric β_3 $GABA_A$ channels (Davies *et al.*, 1997). Other studies investigating the subunit requirements of the action of propofol identified the α subunit to be necessary for the GABA-modulatory action of this anaesthetic (Uchida *et al.*, 1997) and that α_1-containing receptors potentiate the GABA-evoked current more than α_6-containing receptors, whilst the reverse is true for the direct gating of the channel by propofol (Krasowski *et al.*, 1997).

Many researchers have suggested that general anesthetics exert their effects by binding to pre-existing cavities of the target proteins and, in so doing, stabilise certain conformational states such as the open or closed states of ion channel proteins (LaBella, 1981; Franks and Lieb, 1984). The binding sites of propofol have been demonstrated by a number of studies to be pre-existing amphiphilic cavities, for example, from the crystal structures of propofol-bound human serum albumin (HSA) (Bhattacharya *et al.*, 2000) and propofol-bound GLIC (Nury *et al.*, 2011), which

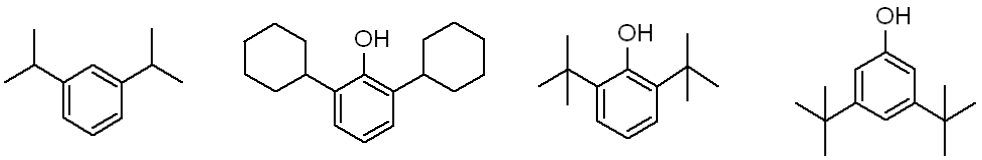
are discussed in more detail in section 1.4, and from the examination of a number of propofol analogues binding to $GABA_A$ receptors (Krasowski *et al.*, 2001a). This thorough study of 27 propofol analogues demonstrated that the binding cavity has limited dimensions favouring hydrophobic groups as shown by an almost linear increase in the potencies of phenol, 2,6-dimethylphenol, 2-isopropylphenol, 2,6-diethylphenol, propofol, and 2,6-di-*sec*-butylphenol in relation to their molecular volumes (Figure 1.2A) up to a 'cut-off' point, whereby analogues with larger alkyl group substitutions at the 2- and 6- positions, such as 2,6-dicyclohexylphenol and 2,6-di-*tert*-butylphenol, are inactive. Krasowski *et al.* (2001a) also highlighted the critical role that hydrogen bonding plays towards propofol-binding as the elimination of the phenolic hydroxyl group of propofol (1,3-diisopropylbenzene), the substitution of non-planar cyclohexyl groups or bulky tertiary butyl groups at the 2- and 6-positions of the phenol adjacent to the hydroxyl group render the compound anaesthetically inactive, whilst the analogue with tertiary butyl groups at the 3- and 5- positions remain active (Figure 1.2B).

A



	phenol	2,6-dimethyl-phenol	2-isopropyl-phenol	2,6-diethyl-phenol	2,6-diisopropyl-phenol	2,6-di-*sec*-butylphenol
Molecular volume	349 Å^3	447 Å^3	495 Å^3	548 Å^3	641 Å^3	740 Å^3
EC_{50} (*in vivo*)	1642 ± 104	68.6 ± 4.6	133 ± 11.4	20.3 ± 2.2	1.9 ± 0.2	0.39 ± 0.1
EC_{50} (*in vitro*)	920 ± 217	135 ± 23	39.7 ± 7.8	13.6 ± 1.1	1.9 ± 0.4	2.8 ± 0.4

B



	1,3-diisopropylbenzene	2,6-dicyclohexylphenol	2,6-di-*tert*-butylphenol	3,5-di-*tert*-butylphenol
Molecular volume	632 Å^3	843 Å^3	711 Å^3	706 Å^3
EC_{50} (*in vivo*)	> 500	> 500	> 500	102 ± 13.3
EC_{50} (*in vitro*)	> 500	> 500	> 500	94.4 ± 26.5

Figure 1.2: Chemical structures, molecular volumes and anaesthetic potencies of ten propofol analogues for the loss of righting reflex assays in *Xenopus laevis* and the potentiation of the GABA response at $GABA_A$ $\alpha_1\beta_2\gamma_{2S}$ receptors expressed in HEK cells. The EC_{50} values of the *in vivo* and *in vitro* responses are expressed in mean μM ± SE. (A) The potency of six propofol analogues generally increases as the molecular volume of the drug increases. (B) The lack of a hydroxyl group in the compound or the presence of non-planar or bulky groups adjacent to the hydroxyl group renders the compound anaesthetically inactive, whilst the propofol analogue with bulky groups at the 3- and 5- positions remains active. (Krasowski *et al.*, 2001a)

The binding sites for the intravenous anaesthetic etomidate had been investigated by chimeric channels of β_1 or β_2 subunits coexpressed in *Xenopus laevis* oocytes with α_6 and γ_2 subunits (Belelli *et al.*, 1997). The insensitivity of the β_1 subunit to etomidate and the highly etomidate-sensitive β_2 subunit (Hill-Venning *et al.*, 1997; Sanna *et al.*, 1997) allowed the idenfication of a region in the N-terminal extracellular domain to be a determinant of etomidate sensitivity through their chimera. Subsequent mutations of this region identified asparagine-289 of β_2 to be involved with the enhancement of GABA-evoked current responses of etomidate as mutation of this amino acid to a serine residue that is at the analogous position of the β_1 subunit strongly suppressed the potentiation response, whereas the mutation of serine-290 to asparagine in α_1 resulted in the gain of etomidate sensitivity of these mutant channels (Belelli *et al.*, 1997). The importance in etomidate sensitivity of this asparagine-289 residue was reinforced by the fact that its mutation in the β_3 subunits of $\alpha_6\beta_3\gamma_2$ receptors resulted in etomidate-insensitive channels (Belelli *et al.*, 1997) when mutated to the amino acid located in the analogous position (methionine-314) of the GABA-gated receptor in *Drosophila melanogaster* (Chen *et al.*, 1994).

Propofol and etomidate were shown to have differential effects on $GABA_A$ receptors (Belelli *et al.*, 1996): etomidate mediates only a modulatory effect on GABA in mammalian $\alpha_3\beta_2\gamma_{2L}$ receptors and has no effects on the invertebrate homologue in *Drosophila melanogaster* (Chen *et al.*, 1994), whilst propofol has GABA-mimetic and -modulatory effects on the mammalian channels and exerts GABA-modulatory effects on the *Drosophila melanogaster* homologue. The amino acid asparagine-289 or, using the common nomenclature that numbers amino acid residues within a transmembrane helix from the N-terminal of that helix in a particular subunit (Ernst *et al.*, 2005), βM2-15, was also shown to be a determinant of propofol sensitivity (Pistis *et al.*, 1999; Siegwart *et al.*, 2002). In addition, the most definitive evidence for $GABA_A$ receptors being a general anaesthesia target was provided from studies of knock-in mice harbouring a single amino acid mutation at the βM2-15 position (Jurd *et al.*, 2003). This study with β_3N265M knock-in mice showed the abolishment of the loss of hindlimb withdrawal reflex that is associated with adminstrations of propofol or etomidate, as well as a reduction in the anaesthestised duration by these intravenous anaesthetics compared to that seen in wildtype mice (Figure 1.3).

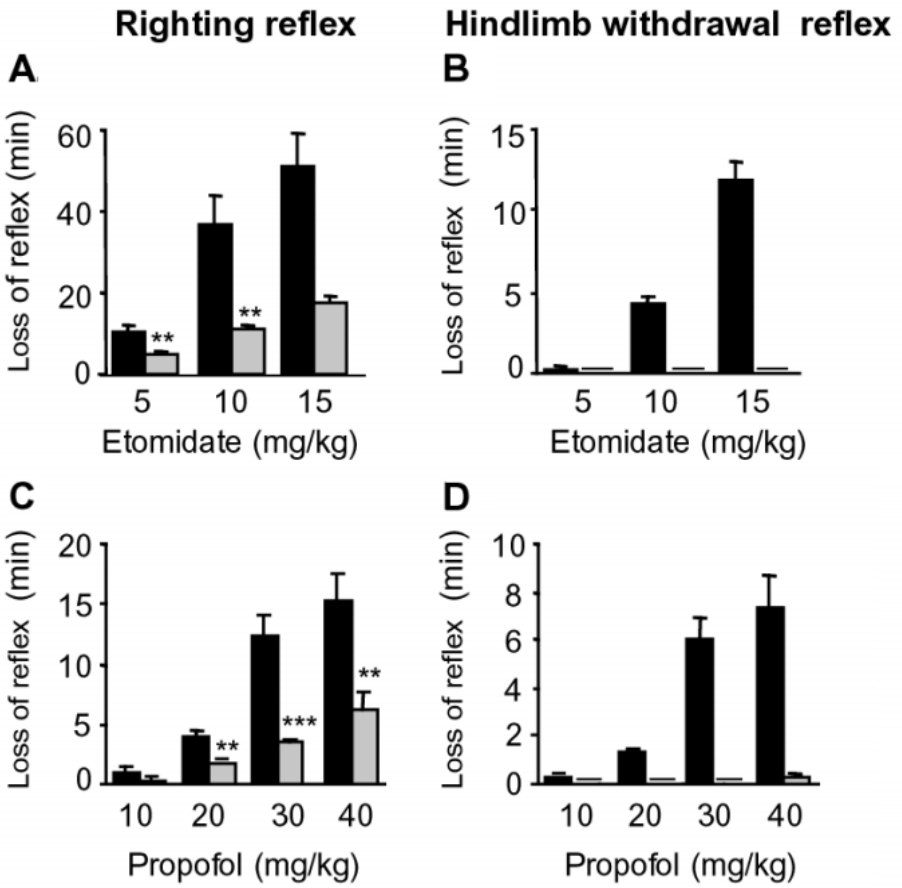


Figure 1.3: Behavioural responses to etomidate and propofol in wildtype and $GABA_A$ β_3N265M knock-in mice. Reductions in the durations of the loss of righting reflex (A and C) and failure to induce the loss of hindlimb withdrawal reflex (B and D) are seen for the knock-in mice (grey) induced by both anaesthetics compared to wildtype mice (black). Durations of reflex are expressed as mean ± SE. Two and three asterisks denote $P < 0.01$ and $P < 0.001$ respectively. Figure taken from Jurd *et al.* (2003).

Another determinant of propofol sensitivity on $GABA_A$ receptors was identified to be the methionine-286 residue of the β_1 subunits (Krasowski *et al.*, 1998). This fourth residue of the transmembrane helix 3 of β subunits (βM3–4) abolished the GABA-modulatory action of propofol when it was mutated from a methionine to a tryptophan residue in $\alpha_2\beta_1$ channels but did not affect the direct modulation of $GABA_A$ receptors (Krasowski *et al.*, 1998). The same results have also been observed on $\alpha_2\beta_3$M286Wγ_2 channels (Siegwart *et al.*, 2002). Methionine-286 has been implicated in propofol-binding in the following two studies. Firstly, Krasowski *et al.* (2001b) mutated methionine to various amino acids and concluded that the mutation alters the dimensions of the binding pocket for propofol. Secondly, Bali and Akabas (2004) showed the protection from labelling of a cysteine residue substituted at that amino acid position. The same study also showed that the cysteine-substituted equilvalent amino acid on the α_1 subunit

alanine-291, the substitution of the aforementioned determinant of propofol sensitivity βM2-15, nor its equivalent αM2-15 residue were protected from the sulfhydryl-specific reagent by the binding of propofol. This suggests firstly that the M3-4 of the β subunit and not of the α subunit is critical for mediating propofol actions, and secondly that propofol binds closer to βM3-4 than to βM2-15, which are about 10 Å apart (Bali and Akabas, 2004) based on modelling to the 4-Å resolution structure of the nACh receptor (Miyazawa *et al.*, 2003), if it binds near the latter amino acid at all. However, modelling $GABA_A$ receptors with nACh receptors should be treated with caution as the amino acid alignments and structural alignments of the M2 and M3 helices of nACh receptors with other ligand-gated ion channels, namely GLIC (Baenziger and Corringer, 2011) and the *Caenorhabditis elegans* glutamate-gated chloride channel (Hibbs and Gouaux, 2011), indicate an off-alignment by one α-helix turn.

Photoaffinity labelling experiments with azietomidate photolabelled the β subunit methionine-286 residue as well as the α_1 subunit methionine-236 in TM1, three residues after a conserved proline residue (Li *et al.*, 2006). Based on cysteine cross-linking studies (Bali *et al.*, 2009) and alignment with the known structure of nACh receptors (Miyazawa *et al.*, 2003), the etomidate-binding site was predicted to be in an intersubunit cavity spanning at least from methionine-236 in TM1 of the α_1 subunit to the methionine-286 in TM3 of the β subunit, which is one helix turn more extracellular. It is interesting to note that methionine-283, another helix turn further up towards the extracellular side of the lipid bilayer compared to methionine-236, was not labelled, suggesting that the binding cavity does not extend to this amino acid.

A high concentration of propofol was shown to partially inhibit the azietomidate photolabel incorporation into these two methionines at positions 236 and 286 (Li *et al.*, 2010). One explanation for this is that propofol competes for binding to the exact same site as etomidate and that some $GABA_A$ receptors with certain subunit combinations only bind etomidate and not propofol. However, no subunit combinations have been identified so far that fit such a description, whilst subunit combinations most sensitive to etomidate are also sensitive to propofol (Hill-Venning *et al.*, 1997; Siegwart *et al.*, 2002). Alternative theories for the partial inhibition of azietomidate-photolabelling by propofol suggested by the authors include the binding of propofol allosterically inhibiting photolabel-binding, as well as the binding of the photolabel and propofol

to the β/α subunit interface within a binding pocket with the two drugs in close proximity (Li *et al.*, 2010).

Other determinants of propofol sensitivity on $GABA_A$ receptors include the β_2 subunit tyrosine-157 at the extracellular domain region that reduced the direct activation of propofol when mutated to a phenylalanine (Fukami *et al.*, 1999 cited in Bali and Akabas, 2004), and the α_1 subunit lysine-278 residue that affected the direct response of propofol on $GABA_A$ receptors when it was mutated to methionine (Hales *et al.*, 2006). However, both of these mutations were also shown to affect the channel response to GABA (Fukami *et al.*, 1999 cited in Bali and Akabas, 2004; Hales *et al.* 2006), suggesting that it is likely that these determinants of propofol sensitivity have affected the transducion mechanism of binding to gating, rather than the binding of the anaesthetic alone.

To determine whether the determinants of propofol sensitivity identified from mutagenesis-coupled electrophysiology studies are involved with their interaction with amino acids at the binding site or with the gating of the channel, high-resolution crystal structures of the anaesthetic-bound target proteins are ideal. However, membrane proteins are hard to crystallise due to their hydrophobicity. A number of techniques have been used to locate binding sites of proteins including a range of spectroscopic techniques and affinity labelling. Photoaffinity labelling with a novel propofol photolabel analogue forms the basis of this project and this technique for studying the interactions of macromolecules is outlined in the next section.

1.2 Photoaffinity labelling

Photoaffinity labelling as a technique for studying the interactions of biological macromolecules was introduced fifty years ago with the photolysis of *p*-nitrophenyl diazoacetate acylating chymotrypsin (Singh *et al.*, 1962). Photoaffinity labelling has the advantage over affinity labelling in that the inertness of the ligand prior to photoactivation allows assays for biological activity to be carried out, allowing the similarities in the activities between the precursor ligand and the photolabel analogue to be tested.

The basic principle of photoaffinity labelling is to use a ligand that, upon irradiation, converts to a highly reactive intermediate that in turn forms an irreversible covalent bond with an amino acid

present at the ligand binding site. The photolabelled protein is then isolated and fragmented, allowing the fragments to be analysed in order to locate any labelled amino acids (Figure 1.4).

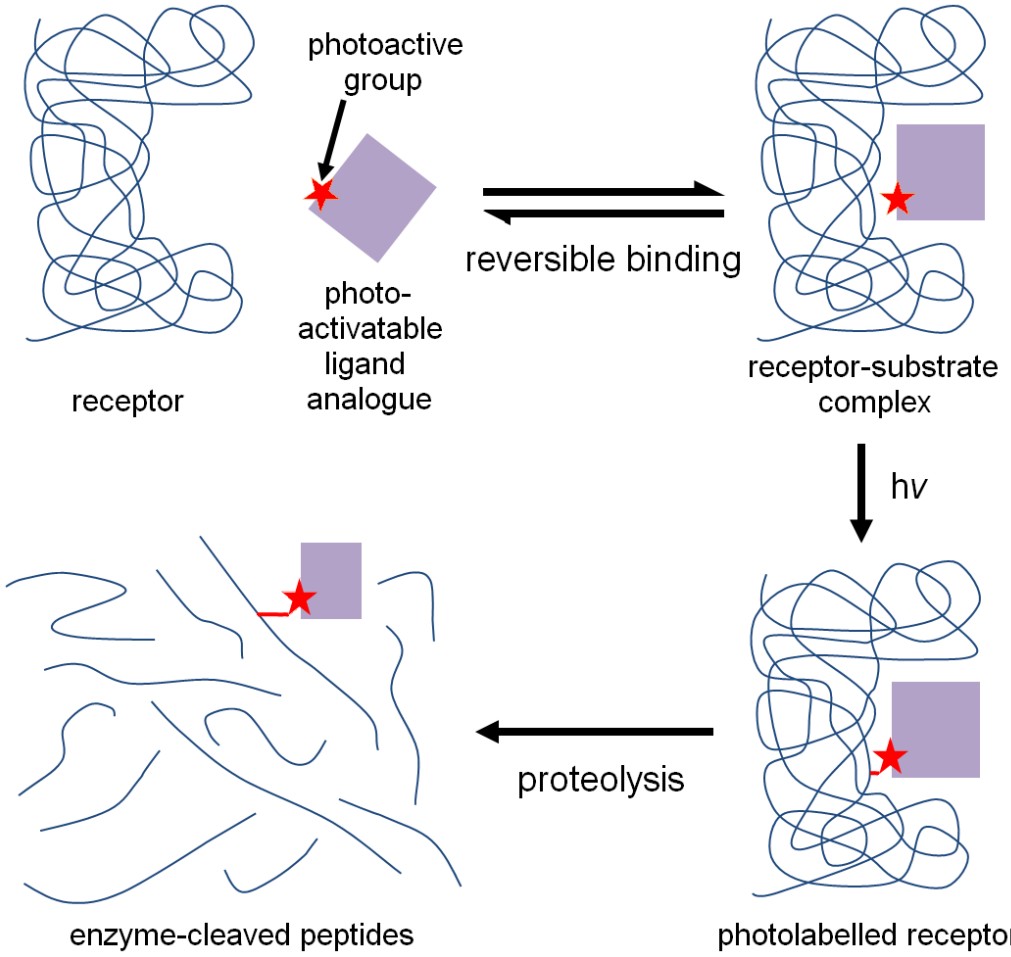
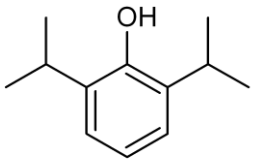


Figure 1.4: Schematic showing the principal stages of photoaffinity labelling. A chemically inert photoactivable ligand analogue binds reversibly to the same binding site as its parent ligand, forming a receptor-substrate complex. When in complex with the receptor at the binding site, the photolabel becomes activated by light (*hv*), producing a highly reactive intermediate that binds irreversibly to the receptor at the ligand-binding site. The labelled protein then undergoes proteolysis to produce peptides, some of which have the photolabel covalently bound.

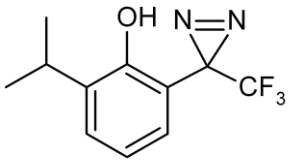
The main properties of an ideal photolabel are as follows. The photolabel should be structurally and chemically not too dissimilar to its parent ligand, such that it is likely to bind to the same binding sites as the parent compound, thus allowing the binding site of the native ligand to be identified. The photolabel should be excited by light outside the range of wavelengths that interact with proteins, particularly avoiding wavelengths of 214 nm and 280 nm, where peptide bonds and aromatic residues absorb strongly, so that the protein is minimally affected during light exposure. The reactive intermediate formed upon irradiation should have a shorter lifetime than the

dissociation time of the receptor-ligand complex, such that the labelling occurs whilst the photolabel is bound within the ligand binding site of the target protein. Furthermore, the reactive intermediate should not have preference over certain amino acids and should react with aliphatic C–H groups as well as nucleophilic X–H bonds.

Many different photoactivatable groups have been employed in photoaffinity experiments, which generally produce three classes of reactive intermediates: nitrenes, carbenes and free radicals. The photoaffinity label used in this project is a diazirine derivative of propofol, 3-(5′-isopropyl-6′-hydroxyphenyl)-3-trifluoromethyl-diazirine, hereafter called photofol (Figure 1.5). As with the photolabel used in the seminal photoaffinity labelling experiment with a diazo compound in 1962, upon irradiation of these chemicals, carbene intermediates are produced that can label any of the twenty amino acids. The demonstration of the insertion of the photolabel carbene intermediate from the photolysis of diazomalonyltrypsin into a C–H bond showcased the ability of photoaffinity labelling to incorporate into these ubiquitous and unreactive bonds (Vaughan and Westheimer, 1969 cited in Chowdhry and Westheimer, 1979).



2,6-diisopropylphenol
(propofol)



3-(5'-isopropyl-6'-hydroxyphenyl)-
3-trifluoromethyl-diazirine (photofol)

Figure 1.5: Chemical structures of the native anaesthetic propofol and its diazirine photolabel analogue photofol.

The photolabelling of diazoacetylchymotrypsin study by Singh *et al.* (1962) was shown to be only about 20–30 % efficient (Plate, 1997). This has been partly attributed to the susceptibility to Wolff-type rearrangement of α-keto carbenes produced to form ketenes, which are more long-lived and can migrate out of the binding site to label more nucleophilic residues or react with water. The diazirine photoactive group that was introduced by Smith and Knowles (1973) was suggested to have superior photochemistry due to being more stable and processing an absorbance band at around 350 nm, well away from light of wavelengths that interact with proteins. Diazirines are cyclic isomers of diazo compounds and are photolysed to produce carbenes as well as being

rearranged to form diazo compounds. At the wavelength used to photolyse diazirine, the diazo compounds have long half-lives. For this reason, many diazirine compounds synthesised for photolabelling purposes, including photofol (Figure 1.5), consist of a trifluoromethyl group adjacent to the diazirine group as this poses strong inductive electron-withdrawing effects to reduce the reactivity of the diazo group, thus minimising labelling via the diazo route, which may involve more unspecific labelling away from the binding sites.

Photoaffinity labelling has been particularly useful for the study of ligand interactions with transmembrane proteins because the conventional method of obtaining high-resolution crystallographic structures is difficult for these large hydrophobic proteins. As mentioned in section 1.1, the binding site of benzodiazepine was identified through photoaffinity labelling with the benzodiazepan agonist, [^{3}H]-flunitrazepam, which labelled histidine-102 of the α_1 subunit of mammalian $GABA_A$ receptors (Duncalfe *et al.*, 1996; McKernan *et al.*, 1998). The binding site for the potent agonist of $GABA_A$ receptors was also identified from photoaffinity labelling with [^{3}H]-muscimol incorporating into the α_1 subunit phenylalanine-65.

This technique for probing ligand-protein interactions has also been carried out with anaesthetic photolabels. The first of these studies was carried out on bovine and human serum albumin (BSA and HSA) with halothane, a native photoactivatable drug. Upon photolysis at 254 nm, [C_{14}]-halothane decomposes to bromine and trifluorochloroethane free radicals. However, the labelling results identified only one of the seven binding sites of HSA (Eckenhoff and Shuman, 1993; Eckenhoff, 1996a) that had been confirmed by crystallographic data (Bhattacharya *et al.*, 2000) and subsequent experiments of photolabelling nACh receptors with [C_{14}]-halothane appeared to be non-specific (Eckenhoff, 1996b). A more successful example of using photoaffinity labelling to locate anaesthetic-binding sites was with an etomidate diazirine analogue. Li *et al.* (2006) used [^{3}H]-azietomidate—which replaces the ethyl group of etomidate with a 2-(3-methyl-3*H*-diaziren3-yl)ethyl group—to photolabel $GABA_A$ receptors, identifying two amino acids, α subunit methionine-236 and β subunit methionine-286, labelled with 3 % and 1 % efficiency respectively. Azietomidate photolabelling is inhibited by etomidate, suggesting that the labelling had occurred at the etomidate binding site. Due to the labelling efficiency of the photolabelling examples given so far being low, radioactive isotopes were deliberately incorporated in these

photolabels so that scintillation count could be used to identify labelled peptides and amino acids through electrophoresis and microsequencing. In photolabelling experiments where the efficiency of labelling is greater—for example, the photolabelling of HSA with an acyl azide etomidate analogue reported an average photolabelling efficiency of 32 % (Bright *et al.*, 2007)—mass spectrometry techniques coupled with high pressure liquid chromatography (HPLC) is used to identify the labelled peptides and amino acids, a process which is discussed in section 1.3.

The synthesis and characterisation of two propofol diazirine analogues have been reported recently: *m*-azipropofol, which has a trifluoromethyl-diazirinyl group added to the *meta* position of the phenol, adjacent to where an isopropyl group at the *ortho* position had been eliminated (Hall *et al.*, 2010) and *p*-(4-azipentyl)propofol, where a 3-(3-methyl-3*H*-diazirinyl)propyl group was added to the *para* position from the hydroxyl group of propofol (Stewart *et al.*, 2011). Both photolabels showed comparable potency to propofol *in vivo*, whilst some pharmacological differences from the native anaesthetics were seen. For example, at 30 μM concentration, where both GABA-mimetic and GABA-modulatory responses are seen for propofol, the *m*-azipropofol blocks the $GABA_A$ channel (Hall *et al.*, 2010); whilst *p*-(4-azipentyl)propofol showed reduced efficacy in enhancing the binding of muscimol and flunitrazepem compared to propofol and did not cause a decrease in the binding of these agonists at a high anaesthetic concentration as exhibited by propofol (Stewart *et al.*, 2011). Hall *et al.* (2010) photolabelled horse spleen apoferritin with *m*-azipropofol and claimed from their mass spectrometry sequencing data that two leucine residues—which are shown to line the propofol binding site of the target protein in the crystallographic structure (Vedula *et al.*, 2009)—had been labelled. Photolabelling was also carried out with *p*-(4-azipentyl)propofol: the labelling of nACh receptor-rich membranes showed photoincorporation occurred most efficiently into α subunits, but still only with an efficiency of 0.2 % (Stewart *et al.*, 2011).

Together, these propofol photolabel analogues, including photofol, could aid the exploration of propofol-binding sites in detail. Photofol, which has a diazirine moiety at the *ortho* position, can potentially identify amino acids that are close to the functional part (hydroxyl group) of propofol within the binding site, whilst amino acids slightly further away from the functional region of the active site could be identified by *m*-azipropofol, and amino acids further still could be identified

with *p*-(4-azipentyl)propofol, which has its diazirine moiety on the fourth position of a pentyl chain. The anaesthetic potency and photoaffinity labelling ability of photofol has been investigated during this project with the aim of confirming this hypothesis and ascertaining its suitability as a photolabel, with the results discussed herein.

1.3 Analysis of proteins and peptides using mass spectrometry

Mass spectrometry now plays an important role in the study of proteins and peptides but a breakthrough for this tool in the field of biomolecular analysis only came about in the 1980s with the development of soft ionisation techniques such as fast atom bombardment (Barber *et al.*, 1981), electrospray ionisation (ESI) (Fenn *et al.*, 1989) and matrix-assisted laser desorption ionisation (MALDI) (Tanaka *et al.*, 1988). Prior to these discoveries, mass spectrometry was restricted to analytes present in the gas phase and thus was only suitable for volatile compounds. These soft ionisation techniques, of which ESI and MALDI are now most commonly used, allow stable ions to be released from the protein or peptide source with no fragmentation, thus enabling the molecular weights of intact molecules to be measured. Due to the lack of fragmentation, complex mixtures can be analysed without prior separation. However, competitive ionisation—whereby the ionisation of some peptides is suppressed when they are in a mixture with other peptides (Cohen and Chait, 1996)—is observed and mass spectra containing fewer different peptides are more easily interpretable. Mass spectrometry is therefore usually coupled with separation techniques such as HPLC.

One of the early uses of mass spectrometers for the study of proteins was simply to measure the molecular weights of polypeptides. This was extended to monitoring non-covalent protein complexes (Ganem *et al.*, 1991): the stoichiometry of binding can be determined from measuring molecular weights of the intact complex and of the individual binding partners. This type of analysis can be carried out with single-stage mass spectrometers.

A typical mass spectrum for a protein or peptide in an ESI-mass spectrometry experiment shows its intact molecules, with molecular weights of M, being protonated by a number (n) of protons (H^+), thus carrying n charges that are separated by their mass-to-charge ratios (m/z), which are calculated as $\frac{M+n}{n}$. It was reported that the maximum number of charges of proteins or peptides

correlates well with their total number of basic amino acids plus one for an N-terminal group that is not acylated (Smith *et al.*, 1990 cited in Biemann, 1992), therefore suggesting that all tryptic peptides—those cleaved after arginines and lysines—without N-terminal acylation have the potential to be at least doubly charged. This ability of ESI to produce multiply charged ions allows the molecular weights of large proteins to be determined by a mass analyser with a relatively small *m*/*z* range. To determine the molecular weights of peptides, or indeed any other species, from peaks on mass spectra, the charge states of the species need to be determined. To address this, a brief overview of isotopes will first be presented.

An atom is comprised of a small dense nucleus, consisting of protons and neutrons surrounded by electrons. The number of protons in the atom defines the element, and the total number of protons and neutrons defines the mass of the atom. The number of neutrons in atoms of a particular element is not fixed for most elements, with the exception of phosphorus and rhodium (Herbert and Johnstone, 2003, p339). For example, all phosphorus atoms are comprised of 15 protons and 16 neutrons and therefore have a mass of 31 Da, denoted as ^{31}P. On the other hand, hydrogen has three naturally occurring isotopes, all with one proton but a varying number of neutrons. The most abundant, consisting of more than 99.985 % of all hydrogen atoms (Herbert and Johnstone, 2003, p346) has no neutrons and is named protium (^{1}H), whereas the isotope with one neutron is called deuterium (^{2}H) and that with two neutrons is tritium (^{3}H). An important element whose isotopes need to be considered in the analysis of peptide mass spectra is carbon. ^{12}C contributes to 98.892 % of the natural abundance of carbon, whilst ^{13}C contributes to 1.108 % (Herbert and Johnstone, 2003, p346). Isotopes of the same element have almost identical chemistry but their presence can be observed using a mass spectrometer. A simple organic compound with one carbon, such as CH_4, exhibits a peak with a relative abundance of 1.1 % at 18 *m*/*z* compared to its base peak at 17 *m*/*z*: For a compound with ten carbons, such as $C_{10}H_{22}$, the peak 1 Da greater than the base peak of 143.2 *m*/*z* has a relative abundance of about 10 %, and for a compound with 100 carbons, such as the alkane $C_{100}H_{202}$, the peaks of the monoisotopic mass at 1405.6 *m*/*z* and that which is 1 Da greater have similar abundances (Figure 1.6).

The isotopic distributions of peaks on the mass spectra aid the determination of charge states by the spacing between the peaks of the isotopes. A spacing of one unit between isotopes indicates a

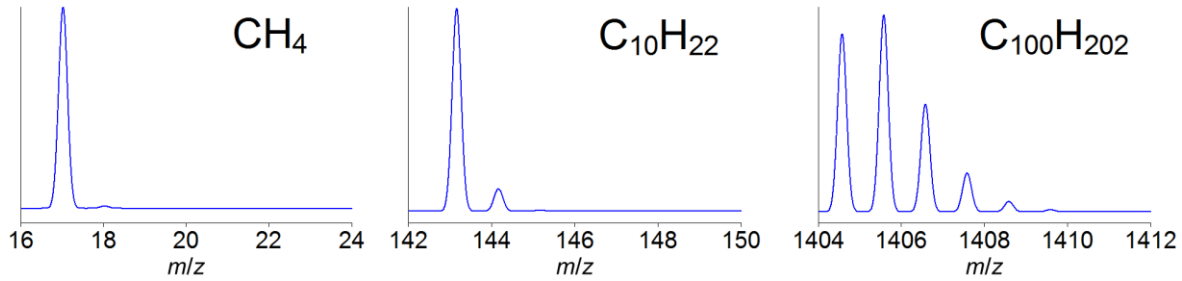


Figure 1.6: Isotopic distributions of molecules with one, ten and 100 carbon atoms.

charge of one, spacing of ½ unit denotes a charge of two and in general, spacing of $\frac{1}{z}$ units denote a charge of *z*. With this information in hand, the molecular weights of the peptides or other species can be calculated from the *m*/*z* of most predominant isotope peak, or the average mass peak if the isotope peaks are not resolved, multiplied by the charge state followed by subtraction of the charge state. For peptides which multiple charged ions are observed, the calculated molecular weights from each ion with a different number of charges are averaged.

One of the major advances in mass spectrometry analysis of proteins and peptides is with the development of tandem mass spectrometry (MS-MS). This has provided additional details in terms of the amino acid sequences and the identification and location of post-translation modifications. So far, the discussion of soft ionisation methods have emphasised the formation of ions that are stable and do not yield any fragments, but this property also leads to a lack of information about the structure of the protein or peptide. In MS-MS, sufficient energy for the peptide ions to fragment is obtained from converting some of their kinetic energy to vibrational energy (Biemann, 1992). In the most common experiment carried out on a triple quadrupole mass spectrometer, the ions of a user-selected peptide, known as precursor or parent ions, are transmitted to the second quadrupole analyser, where they are fragmented by colliding with neutral molecules in a process known as collision-induced dissociation (CID) to allow fragmentation to occur along the peptide backbone, yielding product ions and neutral fragments. The product ions are then analysed by the third quadrupole mass analyser to produce enhanced product ion (EPI) mass spectra (Figure 1.7). MS-MS can also be performed in a single mass analyser, such as the ion trap mass analyser, where the instrument ejects all except for the single desired ion species, followed by fragmentation of the remaining ions to produce product ion mass spectra. These ion trap instruments have the potential to perform several sequential steps in MS-MS, allowing detailed fragmentation studies.

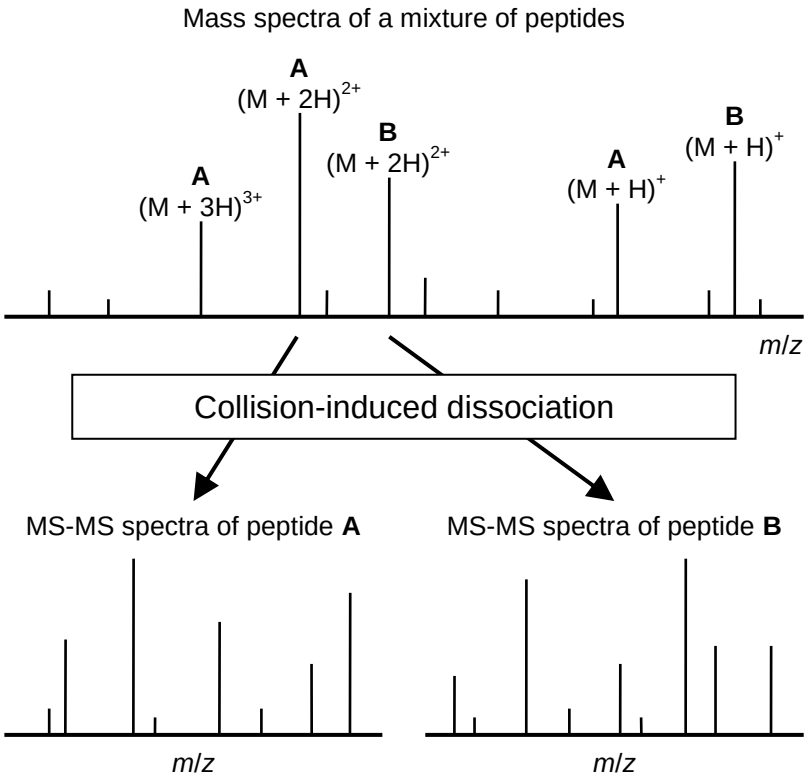


Figure 1.7: Schematic showing the principle of tandem mass spectrometry in which precursor ions are selected to undergo collision-induced dissociation to produce product ions spectra that contain more structural information about the ions.

Three different types of bonds occur along the peptide backbone: the NH–CH, CH–CO and the peptide bond, CO–NH. The cleavages at each of these positions (Figure 1.8A) give rise to one charged and one neutral species, of which only the charged one is monitored by the mass spectrometer. These cleavages result in six common types of ions, a_n, b_n, c_n, x_n, y_n and z_n ions (Figure 1.8B), of which the b_n and y_n types are the most prominent. In the former three types of ions, the charge is retained on the N-terminal fragment, whereas in the latter three, the charge is retained on the C-terminal fragment. Some ions lose fragments from both the N- and C-terminals, producing internal ions. The most common of these are the immonium ions from b_n- and x_n-type cleavages and the amino acylium ions from cleavages of the b_n and y_n types (Figure 1.8C). Many other related ions associated with different amino acids have been documented by Falick *et al.* (1993) and Papayannopoulos (1995) from detailed analysis of thousands of mass spectra. Other fragments that are often observed involve the loss of water, denoted with a degree symbol (°) after the common CID ion nomenclature, or the loss of ammonia, denoted with an asterisk (*) after the CID ion nomenclature, from these ions.

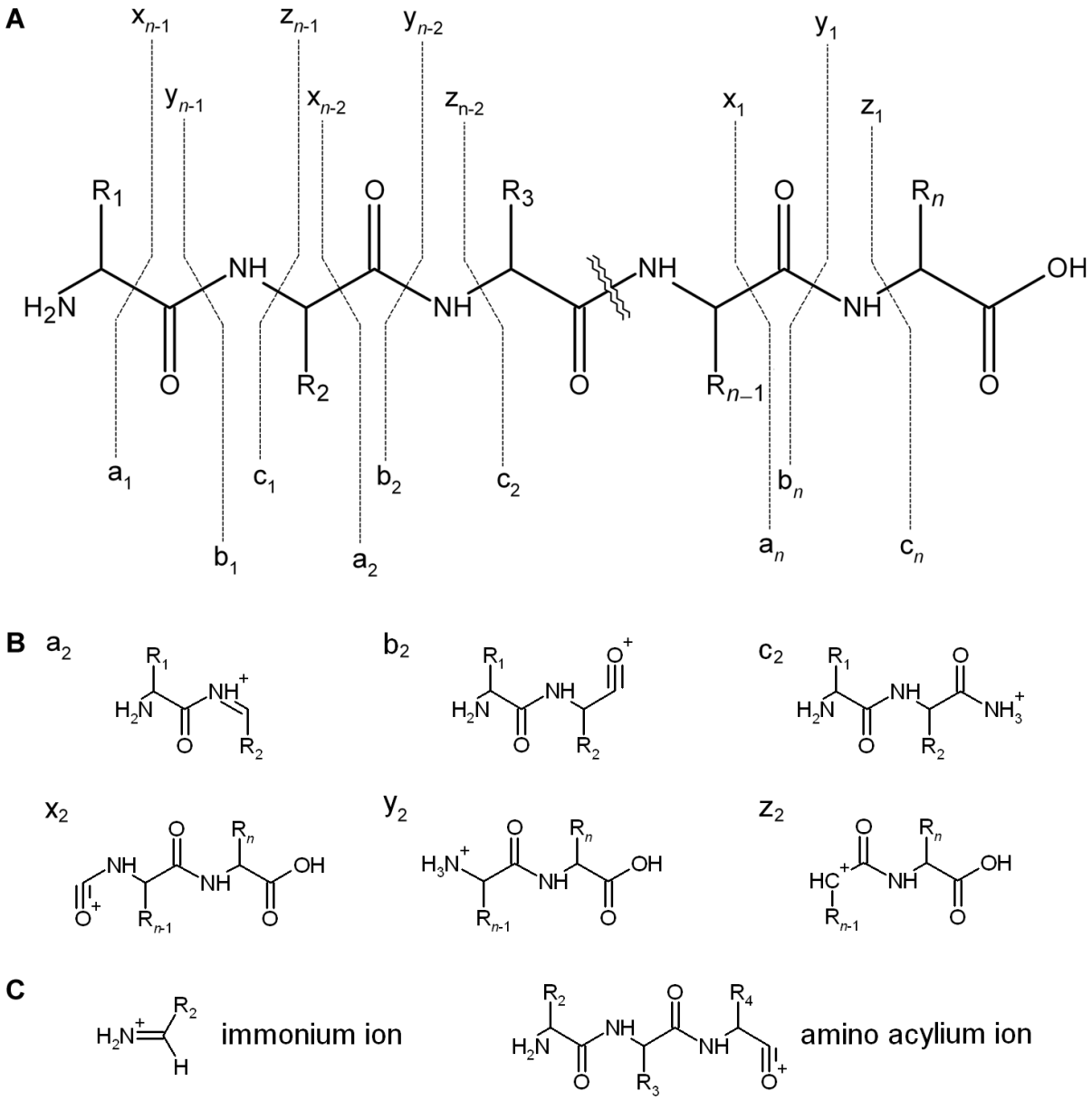


Figure 1.8: Schematic diagram showing cleavages of a peptide at various bonds along the peptide backbone by collision-induced dissociation to produce characteristic ions (A). The structures of the a_n, b_n and c_n ions, with the proton retained on the N-terminal of the peptide, and x_n, y_n and z_n ions, with the charge retained on the C-terminal of the peptide (B), along with two common internal ions (C) are shown. Dotted lines highlight the positions of cleavages along the peptide: a_n and x_n ions arise from CH–CO bond cleavage, b_n and y_n from the cleavage of the peptide bond CO–NH, and c_n and z_n arise from cleavage between the NH–CH bond. R_n represents the side-chain group of the n^{th} amino acid and n denotes the number of amino acids in the parent peptide.

Ions of the same a_n, b_n, c_n, x_n, y_n or z_n series and incremented n differ by the mass of an amino acid less water, also known as the residue mass. This means that the presence of a complete series allows the amino acid sequence of the peptide to be deduced with the exception of not distinguishing between amino acids from the isobaric pairs isoleucine and leucine, and lysine and

glutamine. However, a full ion series is generally not observed and overlapping N- and C-terminal ions are used to arrive at the sequence. Overall, the interpretation of product ion mass spectra relies on assigning mass spectra peaks to ions produced mainly from reasonable cleavages along the peptide backbone and, in some cases, cleavages within the amino acid side chains as well.

In this project, HPLC coupled with ESI-mass spectrometry is used for the separation and detection of labelled peptides, and subsequently MS-MS is used to sequence the labelled peptides and identify the labelled amino acids. One of the advantages of using a mass spectrometry-based approach in the analysis of peptides from the photoaffinity labelling experiments is the direct nature of the measurement. Generally, a mass shift occurs when an unlabelled peptide becomes heavier due to the attachment of a photolabel, but it could also result from different cleavage patterns when the bound photolabel interferes with enzyme cleavage sites. To allow unambiguous identification of labelled peptides, the ability of mass spectrometers to distinguish between different isotopes with the same chemical reactivity is exploited. Different isotopes of the propofol diazirine analogue photoaffinity label photofol with different degrees of deuteration on the benzene ring were synthesised (Figure 1.9). These were grouped into two samples: one containing hydrogen atoms that are of the type that is most abundant naturally, ^{1}H, on the three unsubstituted positions on the benzene ring, named protiated propofol diazirine (pPD), and the other consisting of a mixture of different isotopic forms of photofol, called partially deuterated propofol diazirine (dPD). The dPD sample has a specific proportion of the different isotopes of photofol with up to two deuterium atoms (D_0: 34%; D_1: 43 %; and D_2: 23 %) replacing the protium atoms at the *meta* positions on the benzene ring, which provides a distinct isotopic

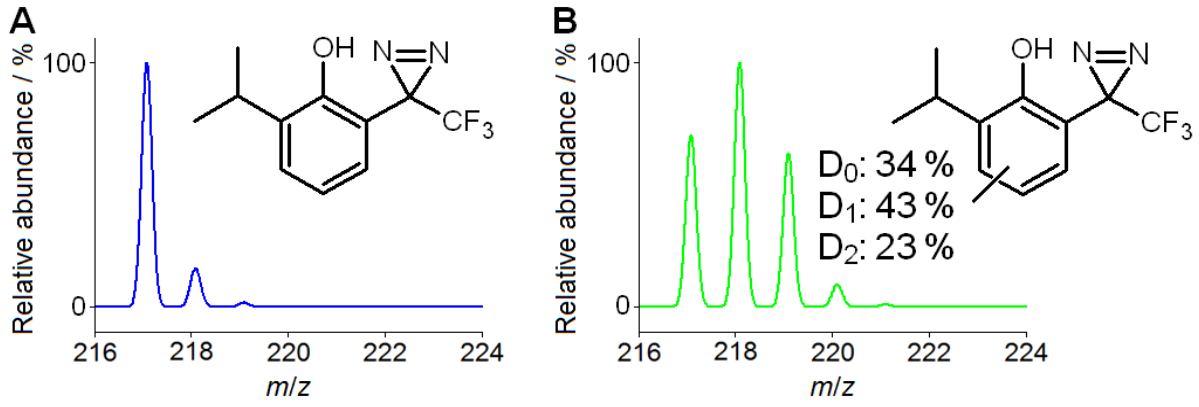


Figure 1.9: Isotopic distributions of the pPD and dPD samples of photofol: (A) pPD, consisting of just D_0 photolabels; and (B) dPD, consisting of 34 % D_0, 43 % D_1 and 23 % D_2 photofol.

distribution pattern for, and subsequently allows easier identification of, labelled-peptide peaks on mass spectra.

Many examples have been documented where the use of stable isotope affinity labels have aided the mass spectra analysis of proteomics studies (Gygi *et al.*, 1999; Schmidt *et al.*, 2005). However, the use of a mixture consisting of a deliberately-chosen proportion of multiple isotopes to achieve a recognisable isotopic distribution pattern in mass spectra appears to be novel.

1.4 Protein targets in this photoaffinity labelling study

The soluble protein HSA and the *Gloeobacter violaceus* ligand-gated ion channel (GLIC) were used to validate the newly synthesised *ortho*-propofol diazirine, photofol (Figure 1.5) as a photolabel. In the remaining part of this chapter, brief introductions to HSA and the bacterial ion channel protein GLIC are given.

1.4.1 Human serum albumin

HSA is the most abundant protein in the plasma, with a concentration of about 0.6 mM. The principal function of HSA is to transport fatty acids (Curry *et al.*, 1998) but it also binds a wide range of administered drugs, contributing significantly to the free concentration available for these drugs to bind to their physiological targets, based on their binding affinity to HSA. The structure of this 66-kDa soluble protein was solved by X-ray crystallography at a resolution of 2.8 Å in 1992, showing three homologous helical domains (I–III), each of which are divided into subdomains A and B (He and Carter, 1992). The three domains assemble with each other asymmetrically, forming a heart-shaped protein with domains I and II almost perpendicular in a T-shape assembly and domain III protruding from domain IIB at 45 ° and forming a Y-shape assembly (Sugio *et al.*, 1999). Despite the similarities in secondary structures of the three domains, they consist of binding sites that differ in their binding cavity environment.

Two binding sites within HSA, termed sites I and II, were described by Sudlow's group in the mid-1970s from their differential binding to fluorescent probes and a variety of drugs such as phenylbutazone, warfanin, ibuprofen and flufenamic acid (Sudlow *et al.*, 1975; Sudlow *et al.*, 1976). With the aid of almost eighty crystal structures of HSA that had been deposited in the Protein Data Bank over the last two decades, most of them co-crystallising with a wide range of

endogenous and exogenous small molecules, many binding sites of HSA including sites I and II have been characterised. Although both sites I and II are in subdomain A, IIA and IIIA respectively, their similarity only extends to both cavities being mainly apolar and their entrances being framed by helices 1, 2 and 6 of their respective subdomains (Curry, 2009). One of the major differences is that site I is bigger in volume than site II. This is due to an extension of its cavity being stabilised by interactions with subdomains IIB and IIIA, where drugs such as indomethacin and iodipamide bind (Ghuman *et al.*, 2005). Due to the sheer size of this site, the binding cavity was further categorised into three sub-sites by their independent binding of three drugs to discrete regions within this site (Yamasaki *et al.*, 1996).

Although many research groups have focused mainly on drug sites I and II, the existance of more than two different specificity sites on HSA was already known from the studies of Sudlow's group. This was demonstrated by probenecid and amitriptyline not displacing fluorenscent probes bound to these sites despite them being able to strongly bind HSA (Sudlow *et al.*, 1975). Figure 1.10,

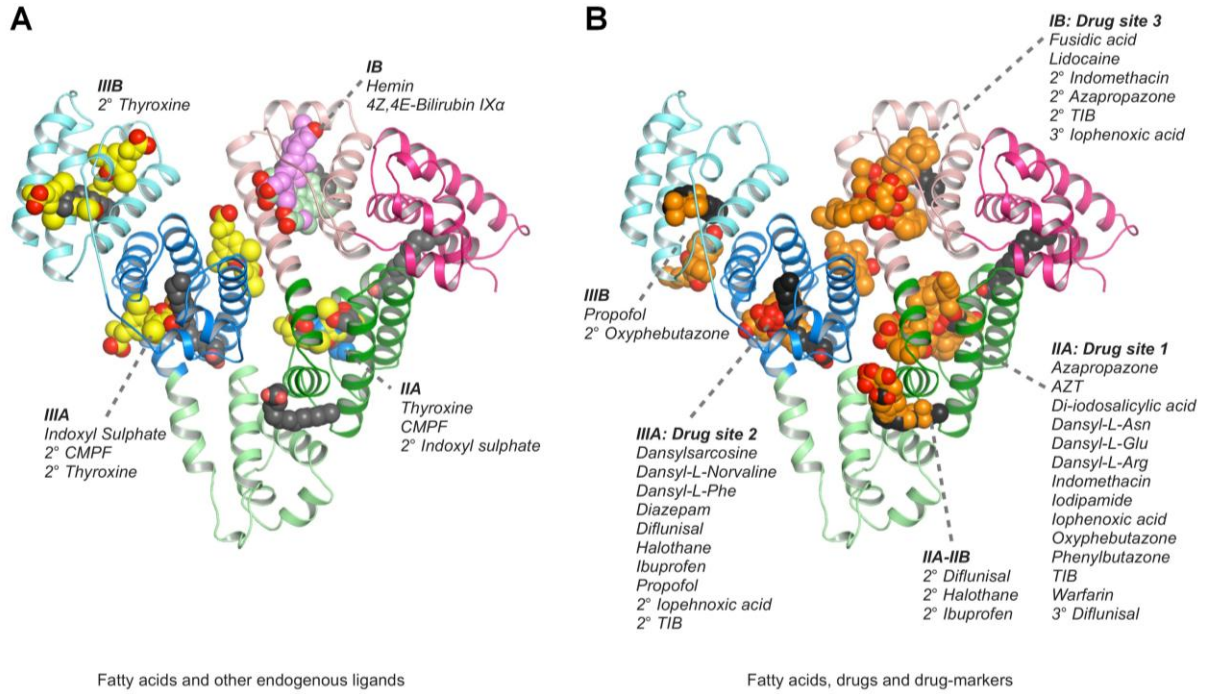


Figure 1.10: The binding of endogenous (A) and exogenous ligands (B) to different sites on HSA as defined by crystallographic studies. All ligands are depicted in space-filling representation with oxygen atoms coloured red. Non-oxygen atoms of fatty acids, bilirubin, hemin, CMPF, thyroxine and all exogenous drugs are coloured grey, pink, green, blue, yellow and orange respectively. For drugs that have differential affinity to more than one site, the binding to secondary and tertiary sites are denoted with 2° and 3° respectively. Figure taken from Curry (2011).

taken from Curry (2011), shows a selection of small molecules that have been co-crystallised with HSA, illustrating its many binding sites where a wide variety of subtances bind.

One of the high-resolution structures of ligand-bound HSA is with the anaesthetic propofol (Bhattacharya *et al.*, 2000). As shown in both Figure 1.10 and Figure 1.11, propofol binds to two discrete sites in subdomains IIIA and IIIB that also bind fatty acids and other drugs. Both of these

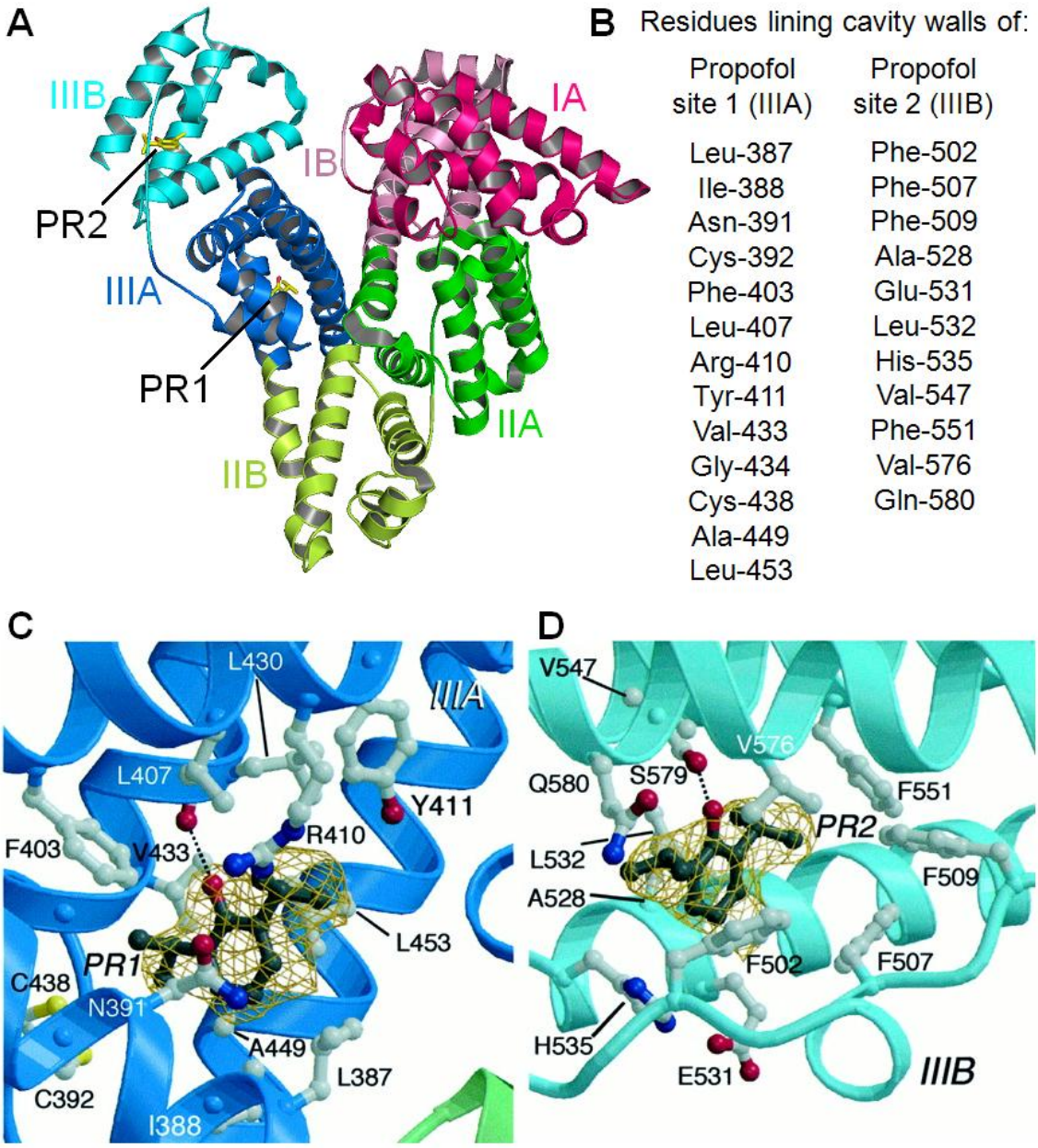


Figure 1.11: X-ray structure of propofol-bound HSA. Two distinct binding sites of propofol at subdomains IIIA and IIIB are shown in (A). The residues that are close to the bound propofol molecules are listed in (B), and shown as ball and stick models in enlarged views of propofol site 1 (C) and propofol site 2 (D). The propofol molecules are named PR1 and PR2 based on their binding to propofol sites 1 or 2 respectively. Figure adapted from Bhattacharya *et al.* (2000).

binding sites comprise of preformed apolar pockets. The binding of a propofol molecule to propofol site 1 with its aromatic group sandwiched between leucine-453 and asparagine-391 involves a hydrogen bond interaction between the phenolic hydroxyl group of propofol and the carbonyl of leucine-430, which are 3.1 Å apart. One of the isopropyl groups of propofol interacts with amino acid residues deep inside the binding pocket via van der Waals interactions upon a 120 ° rotation of the side chain of valine-433 to accommodate this bulky apolar group binding there (Bhattacharya *et al.*, 2000). The outer isopropyl arm is exposed to solvent in close contact with asparagine-391, leucine-407, arginine-410 and tyrosine-411. This arm is very rigid as indicated by its well-defined electron density. The binding of propofol to propofol site 2 involves interaction with phenylalanine-502 and leucine-532, where its aromatic ring is sandwiched between, and the binding is strengthened via a hydrogen bond interaction with serine-579 that is 2.9 Å away. The entrance to propofol site 2 is relatively polar, with glutamic acid-531 and histidine-535 both within about 4 Å to the base of the bound propofol molecule and lining the entrance. As with binding to propofol site 1, only minor conformational changes have been noted when compared to the apo HSA structure, with the major change being a 90 ° rotation of the C_α–C_β bond of phenylalanine-507 (Bhattacharya *et al.*, 2000). Propofol site 1 was shown to have a higher affinity for binding propofol than propofol site 2 as demonstrated by the disappearance of electron density for propofol in subdomain IIIB when a lower propofol concentration was used in the crystallisation experiments. At clinically relevant blood propofol concentration, only one propofol site should be occupied (Bhattacharya *et al.*, 2000). A much higher concentration than that of clinically relevant blood propofol concentration was used in the crystallisation study carried out by Bhattacharya *et al.* (2000) to investigate the claim that there are many binding sites of propofol on HSA.

Before the availability of crystal structures of propofol-bound HSA, the number of binding sites for propofol was predicted to be between nine and twenty (Mazoit and Samii, 1999). Since the availability of the high-resolution structure, the binding of propofol to HSA had also been studied by non-crystallographic methods such as elution chromatography, isothermal titration calorimetry and hydrogen-tritium exchange experiments (Liu *et al.*, 2004; Liu *et al.*, 2005). These studies generally showed two binding sites of propofol on HSA, but also suggested the presence of additional low-affinity sites that are not within domain III.

Besides propofol, the binding sites of HSA for other anaesthetics have been identified. Seven or eight sites of halothane have been reported to be distributed unevenly across the three domains of HSA with most binding to domain II (Bhattacharya *et al.*, 2000; Liu *et al.*, 2004). However, photolabelling of HSA with [C_{14}]-halothane prior to these studies only labelled a 5-kDa fragment starting at aspartic acid-187, which contained the conserved tryptophan-214 amino acid within drug site I. This residue was predicted to be photolabelled due to it being analogous to the tryptophan-212 of BSA that has been labelled and identified from microsequencing (Eckenhoff, 1996a). Another anaesthetic, etomidate, was predicted to bind to drug site I in subdomain IIA and in propofol-binding site II in subdomain IIIB by photoaffinity labelling studies (Bright *et al.*, 2007). The labelling of lysine-199 and lysine-525 with an azyl azide analogue of etomidate was identified through MS-MS sequencing. Other photolabelling studies have been performed using HSA as a target, for example the binding site of semotiadil, an antagonist for L-type calcium channels, was identified to be in subdomains IIIA and IIIB through the labelling of lysine-414 and lysine-541 respectively with an azidophenyl derivative of the drug (Kawahara *et al.*, 2002). These photolabelled amino acids, identified by MS-MS, suggest that semotiadil and propofol would compete for binding to the same sites on HSA.

1.4.2 *Gloeobacter violaceus* ligand-gated ion channel

The cyanobacteria genus *Gloeobacter* is unique amongst other cyanobacteria in that they lack thylakoid membranes, thus photosynthesis occurs within the plasmic membrane and protons fluxed out of the cell instead of from the thylakoid compartments into the cytoplasm. The decrease in extracellular pH activates the proton-gated pentameric ion channel GLIC, which depolarises the cell and may play a role in the phototaxis of the organism (Bocquet *et al.*, 2007; Wang *et al.*, 2012). GLIC was first cloned and expressed in human embryonic kidney (HEK)-293 cells and *Xenopus* oocytes (Bocquet *et al.*, 2007). Upon pore opening, GLIC permeates sodium and potassium ions but not chloride ions, much like the nACh receptor, another pentameric ligand-gated ion channel. The bacterial origin of GLIC and its lack of post-translational modification in the extracellular domain that is typical of its eukaryote homologues make it a great candidate for expression in *Escherichia coli*, allowing an abundance of the protein to be produced.

The first high-resolution X-ray structure of a pentameric ligand-gated ion channel appeared in 2008 (Hilf and Dutzler, 2008). The ion channel, named ELIC from the prokaryote *Erwinia chrysanthemi*, was crystallised at pH 6.5 in a closed-pore conformation. A year later, two groups published the X-ray structure of GLIC in an open conformation (Bocquet *et al.*, 2009; Hilf and Dutzler, 2009) at 2.9- and 3.1-Å resolution. Despite GLIC and ELIC only sharing 18 % sequence homology, their topological organisation is very similar, both composing of five homomeric subunits each consisting of an α-helix and ten β-strands that are organised into two β sheets in their extracellular domains. They both differ from eukaryotic Cys-loop ligand-gated ion channels, such as $GABA_A$ receptors, by their absence of an N-terminal α-helix domain, a cysteine-loop that joins the β6 and β7 strands and a large cytoplasmic domain between TM3 and TM4 (Figure 1.12).

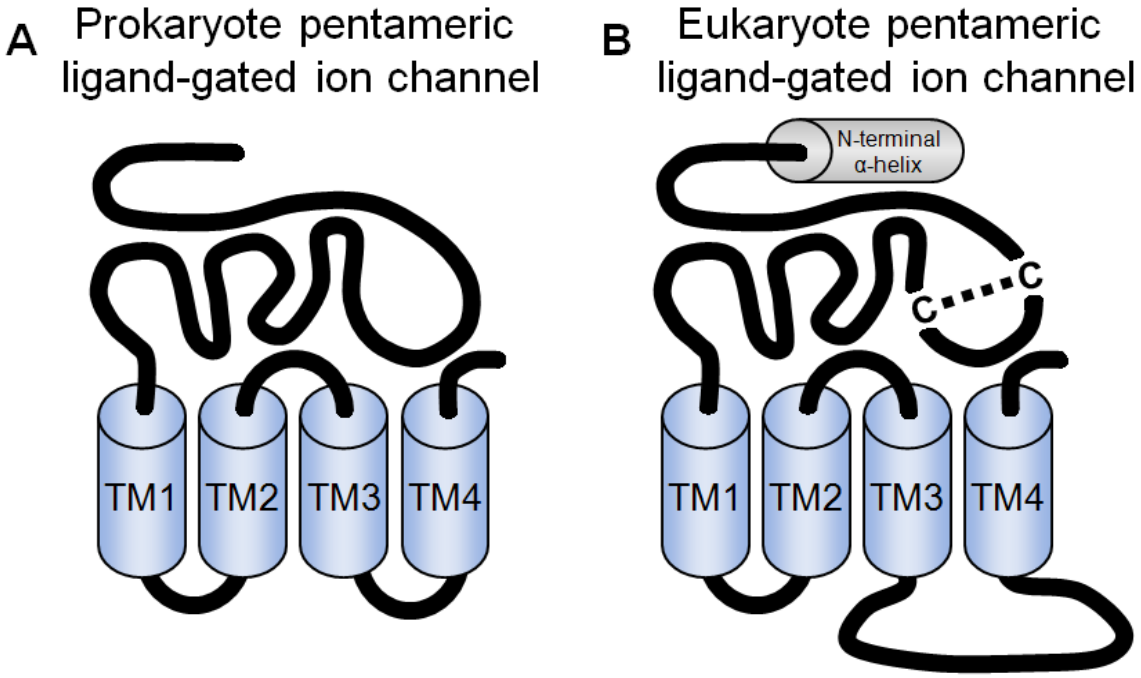


Figure 1.12: Schematic topology of the subunits of prokaryote (A) and eukaryote pentameric ligand-gated ion channels (B) showing the absence of an N-terminal α helix, a cysteine-loop and a large cytoplasmic domain between TM3 and TM4 in the bacterial origin channel.

Since activating ligands of ELIC were not identified to be a class of primary amines until more recently (Zimmermann and Dutzler, 2011) and subsequently followed by the co-crystallisation of ligand and protein, most of the work on understanding the gating of ligand-gated channels were from comparing the apparent open-pore structure of GLIC and the ELIC closed-pore structure. Such a comparison showed a wider opening of the channel pore in GLIC than in ELIC as TM2 and TM3 tilt away from the axis of the channel. This reorganisation of the transmembrane

domain is likely to have resulted from the interactions of the β6–β7 loop and/or the β1–β2 loop with the short loop linking TM2 and TM3 (Bocquet *et al.*, 2009; Corringer *et al.*, 2010; Maksay, 2009), in which all three of these elements show conserved lengths across the pentameric ligand-gated ion channel family. Another notable difference in their crystal structures is seen between their subunit interfaces at the extracellular domain where orthosteric ligands bind in their eukaryote homologues (Maksay, 2009; Zimmermann and Dutzler, 2011; Pan *et al.*, 2012b). The binding interactions of agonists and antagonists in this region—studied on nACh receptors and acetylcholine binding protein (AChBP)—has been shown to affect the mobility of loop C that links β9 and β10, and loop F that links β8 and β9 in that the binding cavity is more constricted when bound to agonists, and more open when bound to antagonists, compared to the apo AChBP structure (Hansen *et al.*, 2005; Cheng *et al.*, 2006). These movements near the orthosteric binding site are transmitted downwards via the β strands, particularly β10. The resultant outward and upward motions of the lower part of the β strands are then conveyed onto the transmembrane domain.

GLIC is activated by protons with a pH_{50} (the pH at which half maximal current is achieved) of pH 5.1 ± 0.1 with a Hill coefficient n_H of 1.9 ± 0.4 in HEK cells, and a pH_{50} of 4.9 ± 0.5 and an n_H of 1.5 ± 0.2 in *Xenopus* oocytes with no desensitisation observed (Bocquet *et al.*, 2007), although contradicting results have since come to light indicating that desensitisation of proton-activation does occur (Gonzalez-Gutierrez and Grosman, 2010; Parikh *et al.*, 2011; Wang *et al.*, 2012). This raises the question of whether the crystal structure of GLIC obtained at pH 4 indeed represents the channel in an open-pore conformation or in a desensitised state.

There are also contradicting reports about the site on GLIC that is responsible for proton-sensing. Wang *et al.* (2012) suggested that histidine-235 in TM2 is a proton-binding site linked to the opening of the channel. The crystal structure of GLIC (Bocquet *et al.*, 2009) shows histidine-235 hydrogen bonds to the carbonyl of isoleucine-259 situated on TM3. This inter-helical hydrogen bond interaction is believed to stabilise the open-channel conformation as suggested by the abolishment of proton-activation in the H235F GLIC mutant (Wang *et al.*, 2012). However, in another study, Duret *et al.* (2011) suggested that the proton sensor of the channel is located in the extracellular domain through functional chimera channels—comprising the extracellular domain

of GLIC and the transmembrane domain of α_1 glycine receptors ($GLIC_{ECD}$-α_1GlyR_{TMD})—being activated by protons, whilst α_1GlyR are inhibited by protons.

The proton-induced current of GLIC is modulated by a number of volatile anaesthetics such as desflurane, halothane, isoflurane, sevoflurane and the intravenous anaesthetic propofol at clinical relevant concentrations (Weng *et al.*, 2010). All of these anaesthetics were shown to inhibit the proton-activation current with propofol having a half maximal inhibitory concentration, IC_{50} of 0.51 μM. The Hill coefficients for these anaesthetics were all below one with propofol having an n_H of 0.42. This indicates either negative cooperativity in binding to identical sites or the presence of multiple binding sites on GLIC.

The high-resolution structures of propofol-bound GLIC and desflurane-bound GLIC became available at the beginning of last year (Nury *et al.*, 2011). The structure shows five anaesthetic molecules binding to a pre-existing binding site in the transmembrane domain of each of the five subunits (Figure 1.13A). Desflurane was shown to bind deeper inside the binding pocket, whilst propofol was bound at the entrance, displacing a lipid molecule at TM1 and TM4 that is observed in the apo GLIC and desflurane-bound GLIC structures. The displacement of the tight binding lipids at the entrance of the binding cavity has been suggested to cause the functional inhibition of this channel (Nury *et al.*, 2011) but this does not explain the mechanism of action of desflurane, which did not compete for binding with the lipid.

The propofol molecule in the orientation that best fits the electron density is seen sandwiched between TM1 and TM3 having van der Waals interactions with mainly threonine-255 and possibly hydrogen bond interactions with tyrosine-254 (Figure 1.13B). Some key residues within the binding site were mutated to alanine or more bulky residues. Isoleucine-202 in TM1 that is in contact with propofol was mutated to alanine, tryptophan and tyrosine, and all three mutants showed no significant differences in propofol-inhibition compared to the wildtype channel. However, valine-242 in TM2, deep in the binding pocket and 6 Å from propofol, and threonine-255 in TM3, another residue in contact with propofol, both increased the propofol response of the proton-activation (Figure 1.13C). These electrophysiology data confirmed the contribution of propofol binding to this site in causing inhibition of the channel and together with

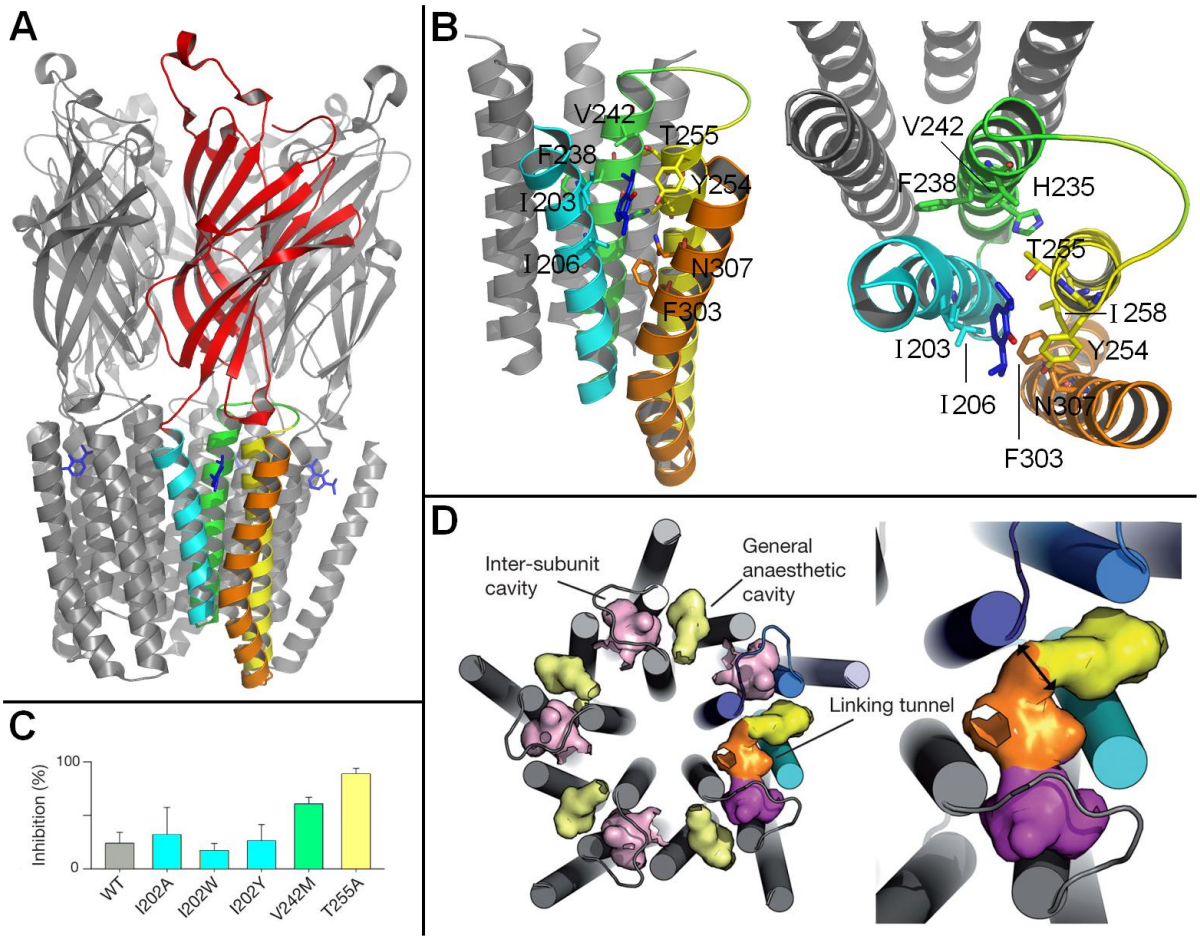


Figure 1.13: X-ray structure of propofol-bound GLIC and functional relevance of the binding site. (A) Cartoon representation showing propofol molecules (blue sticks) bound to the intra-subunit binding cavities at the top of the transmembrane domain, with the extracellular domain (red), TM1 (cyan), TM2 (green), TM3 (yellow) and TM4 (orange) of one subunit coloured. (B) Enlarged side (left) and top (right) views of the propofol-binding site with residues that are on the side of the α helices lining the binding cavity and propofol shown in stick models. The TM2 of the other four subunits are shown in grey and the extracellular domain is omitted from the top view picture. (C) Bar chart showing the inhibition of EC_{20} currents by 10 µM propofol on wildtype and mutant channels. (D) Surfaces of the intra-subunit cavities (yellow), inter-subunit cavities (pink) and the connection of one of the intra- and inter-subunit cavities by a narrow tunnel (orange) of the pentameric channel, with its 20 transmembrane helices depicted as cylinders, as viewed from the extracellular domain, which has been removed for clarity are shown. Figures A and B are from the PDB file with accession code 3P50 and Figures C and D are taken from Nury *et al.* (2011).

the molecular dynamics simulation data also presented in the same paper suggest propofol has significant mobility within the binding pocket (Nury *et al.*, 2011).

The intra-subunit anaesthetic-binding site described so far is connected to an inter-subunit cavity of comparable volume via a narrow tunnel that is less than 3 Å wide (Figure 1.13D). It is

interesting to note that the two residues that have been photolabelled by an etomidate analogue on $GABA_A$ receptors, α_1M236 and β_3M286 (Li *et al.*, 2006), point towards the inter-subunit cavity when the GLIC and $GABA_A$ receptor sequences were aligned.

Other research groups have predicted different binding sites of propofol on GLIC by molecular simulations. Liu *et al.* (2012) presented data that showed the most probable binding site of propofol using docking was predicted to be at the same site as that identified crystallographically. This was however not the case for desflurane, where although its crystallographic site remained among the top three ranked potential binding sites, the most probable binding site was predicted to be in the extracellular domain near the loop linking β6 and β7 at the interface between subunits (Figure 1.14A). Propofol was also predicted to dock to this extracellular site, but with less probability. Another group, LeBard *et al.* (2012) focused on locating the possibility of propofol binding to the pore of GLIC as pore-blocking by propofol was not addressed in the high-

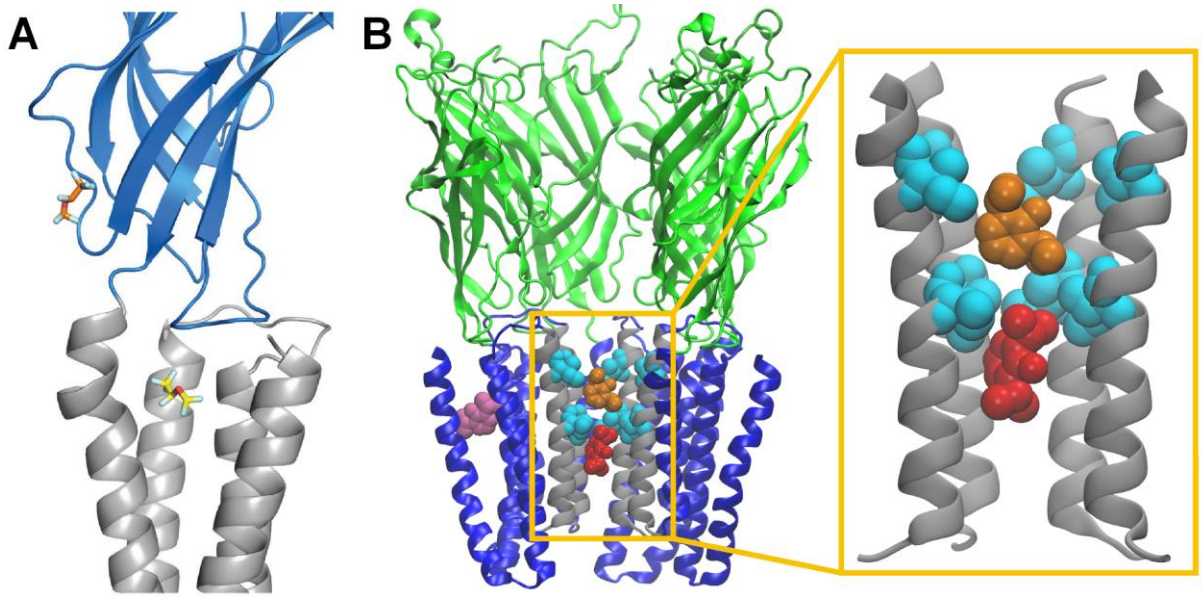


Figure 1.14: Predicted propofol-binding sites on GLIC by computational methods. (A) The most probable binding site of desflurane, predicted to be at an extracellular site (orange) and the less probable binding to the crystallographically-confirmed transmembrane site (yellow) illustrating the rough position of binding for propofol, as both these drugs have been predicted to bind to the same sites using docking studies (taken from Liu *et al.* (2012)). (B) Two propofol molecules (orange and red) are shown to block the pore of GLIC, whilst another molecule (purple) is predicted to bind to the crystallographically-confirmed transmembrane site from free energy perturbation studies. Four of the five TM2 lining the pore are coloured grey with isoleucine-233 (9′) and isoleucine-240 (16′) depicted in cyan space filling models (taken from LeBard *et al.* (2012)).

resolution structural data due to detergents binding there during crystallisation. By investigating the interaction strength of propofol with the pore, two propofol molecules were predicted to bind between the ninth amino acid of TM2, denoted 9′ (isoleucine-233), and 16′ (isoleucine-240), and between 6′ (serine-230) and 9′ (Figure 1.14B). The pore region between 9′ and 16′ harbour the bromine atom of bromolidocaine (Hilf *et al.*, 2010) and the hydrophobic tail of dodecylmaltoside (DDM) (Bocquet *et al.*, 2009; Nury *et al.*, 2011) that bind to GLIC, whilst in the region between 6' and 9', the hydrophilic head of one of the six DDM molecules interacts with the pore-lining residues (Bocquet *et al.*, 2009; Nury *et al.*, 2011). Note that the method used in this molecular simulation study (LeBard *et al.*, 2012) also identified the intra-subunit transmembrane domain site of the crystallographic data.

1.5 Aims of this project

The aims of this project are two-fold: first is to ascertain whether the anaesthetic properties of the novel drug photofol are similar to those of propofol through *in vivo* experiments to test the loss of righting reflex in rats and *in vitro* experiments to assess the modulation of the GABA-evoked response in $GABA_A$ receptors. Secondly, the project intends to validate that the photolabel binds to the same molecular sites as its parent drug in proteins, namely HSA and GLIC, for which X-ray crystallographic structures of propofol bound to these proteins are available.

2 Materials and methods

2.1 Physical characterisation of the photolabels

2.1.1 UV absorbance

The UV absorbance spectra of propofol and the photolabels up to concentrations of 750 µM dissolved in ethanol were measured with a spectrophotometer (Varian, Palo Alto, California, USA, Cary 50 Probe) after zeroing with ethanol. Data files were converted into ASCII format and exported to the graphing software *SigmaPlot 10.0* for analysis. The extinction coefficient at the peak absorbance wavelength of each drug was calculated using Beer's Law:

$$A = \varepsilon c l$$

where A is the absorbance in optical density (OD) units, ε is the extinction coefficient in mol dm^{-3} cm^{-1}, c is the concentration in mol dm^{-3} and l is the length that the light had to travel through the solution in centimetres.

2.1.2 Hydrolysis and photolysis

The aqueous stability of propofol and photofol in ethanol were investigated over periods of two and seven hours respectively. The photofol sample that was stored at 250 mM in ethanol at −20 °C was diluted to 500 µM in ethanol and left at room temperature inside a 4-ml quartz cuvette. The sample was covered by Parafilm and exposed to diminished ambient light during the experiment. At 30-minute intervals, ultraviolet-to-visible (UV-Vis) light spectra were obtained with the spectrophotometer and the data were analysed with *SigmaPlot*. Additional spectra for the photofol sample were also obtained after 24 and 48 hours.

Propofol and photofol at 0.2 mg ml^{-1} were both dissolved in ethanol and subjected to light exposure while in a 1-cm-depth quartz cuvette with a magnetic stirrer. Light with a wavelength range of 280–400 nm from a 500 W mercury (xenon) light source was condensed onto a beam-turning dichroic mirror (LOT-Oriel, Darmstadt, Germany, model #66216). Most of the infrared light was transmitted and excess heat was dissipated via a heat sink on the reverse side of the mirror housing. The reflected ultraviolet (UV) light from the mirror then passed through a manually controlled shutter followed by a 320-nm long-pass filter (LOT-Oriel, part #320FG01-50)

before being focused by a 7.6-cm-focal length secondary lens (LOT-Oriel, part #6197) onto the cuvette that was held in place by a stainless steel holder in a polished chamber. An Ozone Eater (LOT-Oriel, part #66087) was used to remove the ozone emitted. The samples were repeatedly exposed to the filtered light source and their absorbance at 250–500 nm was monitored by the spectrophotometer in between each exposure. The absorbance change at 350 nm was fitted to a single exponential decay equation to calculate the time constant using *SigmaPlot*.

2.2 Biological characterisation of the photolabels

2.2.1 Loss of righting reflex in rats

Experiments assessing the potency of photofol compared to propofol in causing anaesthesia through loss of righting reflex (LORR) assays were conducted by Dr Christopher Edge and Dr Thomas Gent. The righting reflex is the ability of an animal to orient itself and put its paws on the ground when placed on its back. An animal that is unable to continuously keep itself upright whilst being rotated inside a rotation chamber after injection is considered to be anaesthetised if it is later able to regain its righting reflex. Propofol and the protiated form of photofol (pPD) were used to anaesthetise male Sprague-Dawley rats (B & K Universal, London, UK) of weight 498 ± 7.7 g (mean ± standard error (SE)) via tail vein injections and the righting reflexes of the rats one minute post-injection were tested for various doses.

Photofol from a 100-mg ml^{-1} ethanol stock was mixed with up to 500 µl of 0.2 % intralipid prior to injection, whilst the required amount of propofol from commercially available emulsion at 10 mg ml^{-1} (Fresenius Kabi, Germany) was mixed with 100 µl of ethanol before injection to ensure that the presence of ethanol in the photolabel solution did not affect this assay.

The tail vein of each rat was cannulated with a 26-gauge cannula while restrained in a specially made acrylic glass tube. The line was first checked for patency by flushing it with 200 µl of 0.9 % w/v NaCl (Sigma-Aldrich, Poole, UK, S9625), which is physiological saline. The 500-µl pre-prepared drug solution was then injected, followed by another 200 µl of 0.9 % saline. Immediately post-injection, each rat was carefully moved into the rotation chamber, rotating at three revolutions per minute, and the LORR of the rats within one minute post-injection were recorded.

The quantal response of the LORR at one minute post-injection as a function of dose was fitted using the logistic equation of Waud (Waud, 1972):

$$R = \frac{D^n}{D^n + \mathrm{ED}_{50}{}^n} = [0, 1]$$

where R is the LORR response of the rat, D is the anaesthetic dose injected, n is the slope factor and ED_{50} is the dosage for a half-maximal anaesthetic effect.

2.2.2 Whole-cell patch clamping

Whole-cell patch clamping experiments to assess the electrophysiology of HEK-293 cells expressing $\alpha_1\beta_2\gamma_{2S}$ channels were carried out by Dr Robert Dickinson. This technique is implemented by creating a high-resistance seal between a pipette and the cell membrane surface followed by rupturing the patch of membrane enclosed within the tip of the pipette, thus resulting in electrical continuity between the electrolyte solution in the pipette and the interior of the cell. The voltage at the top of the pipette is controlled by a voltage-clamp circuit that prevents the influx of ions from altering the membrane potential set by the experimenter. A single electrode is used simultaneously for both voltage-recording and current-passing.

HEK cells were maintained in Eagle's minimum essential media (Sigma-Aldrich, M4655) at 37 °C, 5 % CO_2 in humidified incubators (Kendro Laboratory Products, Bishop's Stortford, UK, BB6220 CU incubator). The cells adhered to poly-D-lysine (1 mg ml^{-1})-treated coverslips that were co-transfected with pcDNA3.1/GFP, pcDNA3.1/α_1, pcDNA3.1/β_2 and pcDNA3.1/γ_{2S} using a modified calcium phosphate method (Chen and Okayama, 1987). For each transfection, two solutions were prepared in separate eppendorf tubes. Solution A (0.5 µg pcDNA3.1/GFP, 0.33 µg of each pcDNA3.1 subunit, 16.9 µl 2 M $CaCl_2$ and up to 73.75 µl H_2O) was then added drop-wise to solution B (75 µl phosphate-free buffer (280 mM NaCl and 50 mM HEPES, pH 6.96) and 1.35 µl Na_2HPO_4) for each well of cells to be transfected. This solution was left at room temperature for 15 minutes for $CaPO_4$/DNA precipitate to form before 150 µl was added to each well that had been lined with coverslips. The cells were incubated in a 37 °C, 3 % CO_2 humidified incubator for 7–8 hours then washed three times with phosphate-buffered saline (PBS; Sigma-Aldrich, P4417) before 2 ml of the fresh growth media was added to each well. The cells were then incubated overnight before being used for electrophysiological studies.

During the experiment, a cell-attached cover slip, placed in a purpose-built Perspex recording chamber, was observed with an inverted microscope (Nikon Corporation, Tokyo, Japan, Diaphot 200) fitted with a fluorescent attachment for exposing UV light. A gravity-driven perfusion system enabled external solution (140 mM NaCl, 2.5 mM KCl, 2 mM $MgCl_2$, 10 mM HEPES, 10 mM D-glucose and 1 mM $CaCl_2$, pH 7.4) to continuously fill the chamber directly via PTEE tubing connected to the chamber, and test and control solutions were directed to the bath via a double-barrelled perfusion arm to allow fast switching between the two solutions. The flow of solution into the two lines of the perfusion arm was controlled by Teflon-lined valves and solution was drained from the opposite end of the bath using a water-powered vacuum pump.

Cells expressing green fluorescent protein (GFP), and subsequently the $\alpha_1\beta_2\gamma_{2S}$ channels, were selected by their fluorescence under UV light. Whole-cell configuration was achieved using a pipette electrode—back-filled with intracellular solution (140 mM CsCl, 4 mM NaCl, 10 mM HEPES, 5 mM EGTA, 0.5 mM $CaCl_2$ and 2 mM Mg-ATP, pH 7.30)—that was mounted onto a headstage amplifier via a Perspex holder enclosing a silver chloride-coated wire. Switches between the control and test solutions were carried out by the movement of a macromanipulator (Universal Imaging Corporation, Marlow, UK, Sutter MP-225), upon which the double-barrelled perfusion arm was mounted. For the recording of cells, the −60-mV-clamped current of the GABA response was first measured by switching between the external solution and a concentration of GABA required to bring about 7 % of the maximum response (EC_7) that was identified to be 3 μM in preliminary experiments. Upon switching back to the external solution, at least 30 seconds elapsed to allow the current to return to the baseline before testing the GABA-modulatory response of propofol at 3 μM or photofol at 5 μM. These drugs were pre-applied for twenty seconds before the co-application of the aromatic drug with 3 μM GABA for ten seconds. The potentiation of the GABA-induced current by propofol or photofol was calculated by the following equation:

$$P(\%) = 100 \times \frac{I - I_0}{I_0}$$

where $P(\%)$ is the percentage potentiation by the drug, I is the peak of the GABA-induced current in the presence of the drug and I_0 is the peak of the control GABA-induced current.

2.3 Photolabelling experiments with human serum albumin

The validity of photofol as a photoaffinity label to probe propofol-binding sites was first investigated with the soluble protein human serum albumin (HSA). Essentially fatty acid-free HSA (Sigma-Aldrich, A3782) was dissolved in PBS at 10 mg ml^{-1} concentration.

2.3.1 Light exposure of HSA

Samples for light exposure were prepared by mixing 6 µl of the protiated (pPD) or partially deuterated (dPD) photolabel sample, both at 250 mM in ethanol, with 1 ml of HSA stock solution in a 4-ml quartz cuvette. For the control experiments, 1 ml of HSA stock solution was mixed with 6 µl of ethanol, or 6 µl of photolabel was mixed with 1 ml of PBS. After briefly mixing the samples with a magnetic stirrer, they were rested for five minutes in the dark at room temperature to allow time for equilibration. The same mercury lamp setup as was used in the photolysis experiment (section 2.1.2) was then used to irradiate the samples for two minutes. During the light exposure, the samples were continuously stirred with a magnetic stirrer. An additional control sample consisting of a mixture of 1 ml of HSA with 6 µl of ethanol was not exposed to light in order to investigate the effects of light exposure on HSA.

2.3.2 Denaturation of proteins and modification of cysteine residues

After irradiation, the 1-ml light-exposed sample was transferred to a 3.75-ml glass vial containing 0.8 g of guanidine hydrochloride (GnHCl; Sigma-Aldrich, G3272) and 20 µl of an EDTA stock to achieve a final concentration of 6 M GnHCl and 2.0 mM EDTA. The samples were then incubated at 37 °C for two hours to denature the protein. 10 µl of dithiothreitol (DTT; 50 mg ml^{-1} in PBS; Sigma-Aldrich, D0636) was added to reduce disulphide bonds at 37 °C for an extra hour before 80 µl of iodoacetamide (Sigma-Aldrich, I1149) was added to block the reformation of disulphide bonds by forming carbamidomethyl-cysteines for 30 minutes at room temperature.

2.3.3 Trypsin-digestion of proteins

The denatured cysteine-modified protein samples were transferred into 1-cm-wide dialysis tubing (Medicell International Ltd, London, UK) using fire-polished Pasteur pipettes. The ends of the tubes were closed with plastic clips and the samples were dialysed separately in 1 L of freshly-

made 100 mM ammonium bicarbonate (Sigma-Aldrich, A6141), pH 8.0 three times over the course of about 24 hours at room temperature with gentle stirring.

After dialysis, the samples were transferred to HPLC-sample vials with fire-polished Pasteur pipettes and 50 µg (1:200, w/w) of sequencing grade trypsin (Sigma-Aldrich, T8658) dissolved in 25 µl of 100 mM ammonium bicarbonate was added to each vial. The samples were then incubated at 37 °C overnight before being stored at 4 °C ready for analysing by HPLC-MS.

2.4 Photolabelling experiments with *Gloeobacter violaceus* ligand-gated ion channel

The validity of photofol as a photoaffinity label to probe propofol-binding sites was then investigated with the pentameric ligand-gated ion channel from *Gloeobacter violaceus* (GLIC). GLIC was kindly supplied by Professor Pierre-Jean Corringer and Dr Ludovic Sauguet having been expressed and purified as described by Bocquet *et al.* (2007; 2009) and Nury *et al.* (2011), with the main contaminant being *Escherichia coli* maltoporin that co-purifies together with GLIC. The protein was supplied at 8.5 mg ml^{-1} or 10.0 mg ml^{-1} in a GLIC buffer (300 mM NaCl, 20 mM Tris-HCl and 0.02 % DDM, pH 7.6) and was stored in 50-µl aliquots at −80 °C.

2.4.1 Light exposure of GLIC

A bespoke cuvette was constructed by Professor Nicholas Franks by tightly fitting a piece of polytetrafluoroethylene (PTFE) Teflon inside a semi-micro quartz cuvette such that the inner width and depth were 2 × 5 mm, reducing the depth of the solution it held from 10 to 5 mm, thereby increasing the height of the liquid.

Just prior to each experiment, GLIC solution—from one aliquot or two aliquots combined—was thawed and mixed with either 100 mM pH 7.8 Tris buffer (described in more detail in sections 2.4.2.2 and 2.4.2.4) or with a mixture of Na_2HPO_4 and citric acid to achieve a pH of 4 (see section 2.4.2.3). The diluted GLIC solution was allowed to equilibrate with either the Tris or Na_2HPO_4-citric acid buffer for a few minutes before being aliquoted into three 500-µl eppendorfs. The three separate GLIC samples were for the addition of ethanol, pPD or dPD: for every 1 mg of protein, 2 µl of 250-mM photolabel or 2 µl of 100 % ethanol was added. Upon adding ethanol or a photolabel, the samples were mixed then subsequently transferred using a 100-µl Hamilton

syringe to the specially made cuvette, where the samples were allowed to further equilibrate for two minutes in the dark before UV light exposure.

The same mercury lamp setup as described in sections 2.1.2 and 2.3.1 was used to expose GLIC to light of wavelengths ranging between 320 and 400 nm for a total duration of two minutes. A magnetic stirrer was not used to stir the GLIC samples continuously during exposure to light as no stirrers were small enough to move freely while being totally submerged in the small volumes of GLIC solution being irradiated. Instead, the needle of the Hamilton syringe was used to stir the solutions in between four 30-second intervals of exposure.

2.4.2 Denaturing and digesting GLIC

Four different methods were used to process the GLIC protein into peptides. The major differences in all of these methods compared to processing HSA into peptides, discussed in sections 2.3.2 and 2.3.3, are due to the relatively low abundance of GLIC sample available, the presence of only one cysteine residue in a GLIC subunit and the nature of the GLIC sequence, which has far fewer lysine and arginine residues than HSA and therefore renders the commonly used trypsin ineffective as the sole digestive enzyme.

As low as 100 μg of GLIC was used for each sample in an experiment compared to using 10 mg of HSA for each sample in all of the HSA experiments. The total reaction volume was subsequently downscaled by approximately ten-fold compared to that in the HSA experiments, where it was about 1.7 ml, so that peptides from the digestion of similar molar amounts of protein could be analysed by HPLC-MS via the injection of a comparable volume of sample without the need to concentrate them just prior to injection. With their small reaction volumes, these samples were not subjected to dialysis before adding protease but the denaturing agents were diluted with 100 mM pH 7.8 Tris solution before digestion.

The presence of only one cysteine residue in the GLIC peptide sequence renders the cysteine-modification steps involving DTT and iodoacetamide that were used for the HSA experiments almost redundant. There is however the possibility of cysteine residues from different subunits cross-linking but this should not make the analysis of cysteine-containing peptides overly complicated. The use of cysteine-reduction and -modification agents without following up with

dialysis prior to protease-digestion would result in reduced enzymatic activity of the enzyme, thus cysteine-modification steps were avoided.

From the *in silico* digestion analysis of GLIC, chymotrypsin was chosen as the enzyme to cleave GLIC into predominantly manageable-sized peptides for HPLC-MS analysis. Chymotrypsin was used at a concentration of 1:100 w/w. In one of the experiments, the endoproteinase Lys-C was also used prior to digestion with chymotrypsin (see section 2.4.2.4).

The following subsections describe each of the different GLIC denaturation and digestion treatments used in detail.

2.4.2.1 Treatment A: denatured in guanidine hydrochloride and digested with chymotrypsin

The first treatment involved the denaturation of GLIC in GnHCl, similar to the initial treatment to light-exposed HSA. This method was used in the initial control experiment to identify chymotryptic peptides of GLIC from HPLC-MS at different durations of digestion, ranging from four hours and 20 minutes to 34 hours and 40 minutes at 140-minute intervals.

500 µg of GLIC in 50 µl of GLIC buffer was exposed to light for two minutes with manual stirring using a Hamilton syringe needle at 30-second intervals (section 2.4.1). The light-exposed sample was then transferred to a HPLC-glass vial containing 52.7 mg of GnHCl and 2.7 µl of a 25-mg ml^{-1} EDTA stock solution to achieve a concentration of 6 M GnHCl and 2.02 mM EDTA in 89.9 µl. The reaction mixture was then incubated for two hours at 37 °C to denature the protein, before 410.2 µl of proteolysis buffer (100 mM Tris-HCl and 10 mM $CaCl_2$, pH 7.8) was added to dilute the GnHCl by 5.6-fold to 1.08 M. Upon dilution, 25 µl of chymotrypsin (Roche, Welwyn Garden City, UK, sequencing grade) resuspended in 100 mM HCl at 0.2 mg ml^{-1} was added to begin cleavage of GLIC at 25 °C. 30 µl of GLIC sample was directly injected from the reaction vial into the HPLC-coupled mass spectrometer for analysis at 140-minute intervals.

2.4.2.2 Treatment B: photolabelled at pH 8, denatured in SDC and digested with chymotrypsin

A protocol that claims to improve the identification of both hydrophilic and hydrophobic peptides—involving the use of the bile salt sodium deoxycholate (SDC) (Masuda *et al.*, 2008)—

was adapted for the treatment of GLIC after exposure to UV light in the absence or presence of a photolabel.

Two aliquots of 8.5 mg ml^{-1} GLIC totalling 850 µg were combined. 33 µl of GLIC sample was mixed with 14.4 µl of 100 mM pH 7.8 Tris buffer. After equilibration of the mixture for two minutes, 0.6 µl of ethanol or the same quantity of a photolabel from a 250 mM stock was added and transferred to the bespoke cuvette for light exposure as described in section 2.4.1. Upon light exposure, the sample was transferred to a 500-µl eppendorf, where 12 µl of a 10 % SDC (Sigma-Aldrich, D6750) stock was added, achieving a concentration of 2 % SDC. The solution was incubated for one hour at 37 °C to allow the protein to denature in the presence of SDC, after which the reaction mixture was diluted with Tris buffer and 2.8 µg of chymotrypsin was added to allow digestion in the presence of 1 % SDC.

Each of the samples was then aliquoted between two HPLC-vial glass inserts (Agilent Technologies, Yarnton, UK, 250-µl pulled point glass) for digesting at two different durations. Four and ten hours after initiating chymotrypsin-digestion, the reactions were stopped by adding 1 % trifluoroacetic acid (TFA; Sigma-Aldrich, 302031), which caused the formation of a white precipitate. This was a breakdown product of SDC and was removed by the phase transfer surfactant method (Masuda *et al.*, 2008). A volume of ethyl acetate equal to that of the sample was added. The samples were mixed well by pipetting and inverting the glass inserts once covered with Parafilm. The organic and inorganic layers were separated by centrifuging the sample at 13,200 rpm for five minutes upon sitting the glass insert inside a 500-µl eppendorf enclosed in a 1.5-ml eppendorf. The top layer containing the organic solvent and the breakdown products of SDC was pipetted off. The clear bottom layer solution was transferred to a new glass insert—avoiding the white precipitant, which in some instances caused the subsequent blockage of the chromatography column—ready for loading onto the HPLC column.

2.4.2.3 Treatment C: photolabelled at pH 4, denatured in SDC and digested with chymotrypsin

The Na_2HPO_4-citric acid buffer has a pH range of between approximately 2.6 and 7.6 (Sigma-Aldrich, n.d.), which makes it suitable for allowing GLIC to be photolabelled at pH 4 followed by the denaturation and cleavage of the labelled protein at neutral or a slightly basic pH. Preliminary

experiments were carried out by mixing different ratios of citric acid monohydrate ($C_6H_8O_7.H_2O$; Sigma C7129), sodium hydrophosphate (Na_2HPO_4; Sigma-Aldrich, S3264) and the GLIC buffer without DDM (300 mM NaCl and 20 mM Tris-HCl, pH 7.6) and measuring the acidity of the mixture with a pH meter (Radiometer, Copenhagen, Denmark, PHM 93). It was found that for a volume, *x*, of GLIC buffer, adding 0.6*x* of 100 mM citric acid and 0.36*x* of 500 mM Na_2HPO_4 resulted in a mixture with a pH of 4. This could then be neutralised to pH 7 by adding *x* of 1 M Na_2HPO_4, whilst the addition of 2*x* instead raised the pH to approximately 7.8. At neutral pH, the acid-cleaved detergent SDC could be added without precipitation.

For the photolabelling experiment, 135 µg (17.5 µl) of GLIC was mixed with 6.3 µl of 500 mM Na_2HPO_4 and 10.5 µl of 100 mM citric acid in a 500-µl eppendorf. After equilibration of the mixture, 0.3 µl of ethanol or a photolabel sample from a 250 mM stock dissolved in ethanol was added and transferred to the bespoke cuvette for light exposure as described in section 2.4.1. Upon light exposure, the sample was transferred back to the small eppendorf, where 18.6 µl of 1 M Na_2HPO_4 and 8.4 µl of 10 % SDC stock solution were added to achieve a pH of 7 and SDC content of 1.4 %. The solution was incubated for one hour at 37 °C to allow the protein to denature in the presence of SDC. Another 15 µl of 1 M Na_2HPO_4 was then added along with 1.35 µg of chymotrypsin (from a stock of 0.2 mg ml^{-1} in 100 mM pH 7.8 Tris buffer) for digestion of GLIC to occur in the presence of 1 % SDC. After four hours of chymotrypsin-digestion, the reaction was stopped by 1 % TFA and the samples were prepared for injection into the HPLC column as described in section 2.4.2.2.

2.4.2.4 Treatment D: photolabelled at pH 8, denatured in urea and RapiGest and digested with Lys-C followed by chymotrypsin

In this protocol, the post-light exposure treatment of GLIC was designed to mimic that from Professor Alex Evers' group on the complete sequencing of the $GABA_A$ α_1 subunit (Chen *et al.*, 2012a). The main changes from the previous treatments described herein are that an extra enzyme, the endoproteinase Lys-C (Roche, sequencing grade), was used before chymotrypsin; urea (Sigma-Aldrich, U5376) was used as it has been reported that at a high concentration of this denaturing agent, Lys-C still remains active (Klammer and MacCoss, 2006); and SDC was swapped for another acid-cleavable detergent, RapiGest (Waters Corporation, Milford, Massachusetts, USA).

GLIC dissolved in the GLIC buffer at a concentration of 8.7 mg ml^{-1} was aliquoted into 32.2-µl batches in LoBind 500-µl eppendorfs. 0.6 µl of ethanol or the same volume of 250 mM pPD or dPD photolabel was added to separate aliquots to give three sets of samples and the solutions were exposed to light as described in section 2.4.1. The light-exposed samples were then transferred to 1.5-ml LoBind eppendorfs with 28.8 mg of urea along with 2.5 µl of a 2.4 % RapiGest stock solution made from reconstituting 1 mg in 60 µl of 100 mM pH 7.8 Tris buffer. These 60-µl samples were incubated at 37 °C for three hours before 0.7 µg of Lys-C was added from a stock solution of 0.1 mg ml^{-1} in water. Lys-C digestion took place at 37 °C for two days to ensure complete digestion by the enzyme. 402.5 µl of 100 mM pH 7.8 Tris buffer was added to dilute the urea concentration to 1 M and $CaCl_2$ was added to achieve a final concentration of 10 mM. 2.8 µg of chymotrypsin was added and half of each of the three sets of samples was digested for five hours at 25 °C, whilst the remaining samples were digested for 20 hours. Formic acid (FA; J.T.Baker, Deventer, Netherlands) to a final concentration of 1 % was then added to inactivate the chymotrypsin. The reaction mixtures were subsequently incubated at 37 °C for 45 minutes to allow the breakdown of RapiGest as per the manufacturer's instructions. The phase transfer surfactant method for removing the breakdown product of the acid-cleaved detergent SDC described in sections 2.4.2.2 and 2.4 was not used. The precipitation that formed was separated by transferring each sample into a 250-µl HPLC-vial glass insert and centrifuging the sample by sitting the glass insert inside a 500-µl eppendorf enclosed in a 1.5-ml eppendorf at 13,200 rpm for 15 minutes. The supernatant peptide solution was pipetted to a clean glass insert ready for injection into the HPLC column.

2.5 High pressure liquid chromatography-mass spectrometry

2.5.1 Liquid chromatography-mass spectrometry

Peptides samples were analysed by HPLC-MS using an Agilent 1100 Series instrument coupled with an electrospray ionisation (ESI)-mass spectrometer (Agilent Technologies, G1946C). The HPLC components included an autosampler module that rested on top of a thermostat module, allowing samples to be kept at 4 °C as they sat inside the autosampler. A quad pump delivered eluents (solvent A: 99.94 % water and 0.06 % TFA or FA; solvent B: 19.944 % water, 80 % acetonitrile and 0.056 % TFA or FA) through a reverse phase 5-µm C18 column with a 2.1-mm

internal diameter and 250-mm length (Grace Vydac, Hesperia, California, USA), where peptides were separated by their size and polarity. The eluate passing out of the column was monitored using a photodiode detector array housed in a multiple wavelength detector module (Agilent Technologies, G1365B), where its absorbance at 214 and 280 nm was recorded. Solutions leaving that module were then directed into either a fraction collector module (Agilent Technologies, G1364A) or the mass spectrometer.

Two elution profiles were used throughout this project: a 140-minute run for the eluents containing TFA and a 170-minute run for the eluents containing FA. HPLC-grade water and acetonitrile (ACN) were purchased from J.T.Baker. The flow rate of the eluents was 200 µl min^{-1}. The elution profile used for eluting solvents A and B containing TFA was 100 % solvent A / 0 % solvent B to 63 % solvent A / 37 % solvent B in the first 63 minutes, then to 25 % solvent A / 75 % solvent B between 63 and 95 minutes, and to 2 % solvent A / 98 % solvent B over the following ten minutes. The solvent gradient was then reversed rapidly over one minute back to 100 % solvent A, at which point it was held for 34 minutes (Figure 2.1; solid line). The 170-minute profile for eluting solvents A and B containing FA was 100 % solvent A / 0 % solvent B for the first eight

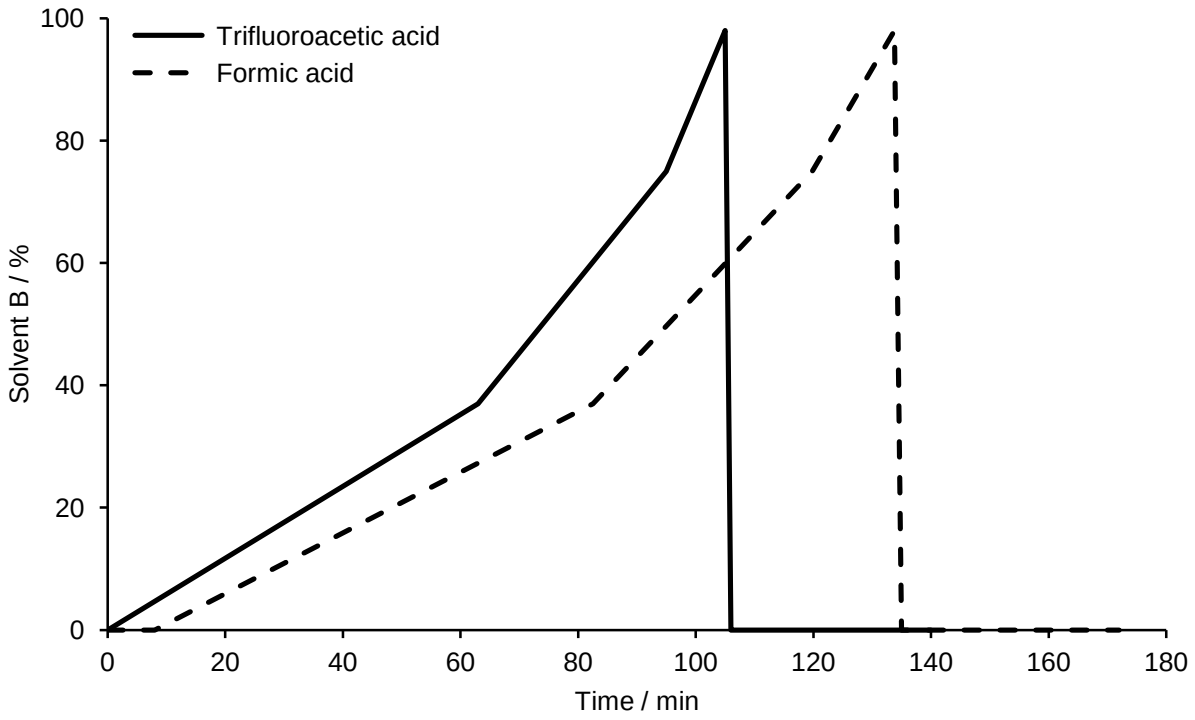


Figure 2.1: Elution profiles for eluents containing either triflouroacetic acid (solid line) or formic acid (dashed line).

minutes, then to 63 % solvent A / 37 % solvent B between eight and 82.5 minutes, then to 25 % solvent A / 75 % solvent B up to 119.8 minutes, and to 2 % solvent A / 98 % solvent B over the following 14 minutes. The solvent gradient was then reversed rapidly over one minute back to 100 % solvent A, at which point it was held for 37 minutes (Figure 2.1; dashed line).

In the early experiments of the project, eluates were directed to the mass spectrometer seven minutes after the start of the chromatography run, instead of just 2.5 minutes in the later experiments, to allow salts in the flowthrough to divert away from the ionisation chamber. In the chamber, the drying gas, nitrogen, was passed at a flow rate of 10 L min^{-1} at 350 °C, and the nebuliser pressure of 30 psi and fragmentation voltage of 70 V were set. Eluates were subjected to ESI and their positive ions were directed to the single quadrupole for their mass-to-charge ratios (m/z) in the range of 100–1500 to be detected. The mass spectrometer was configured to run in 'scan' mode normally, in which the m/z range was scanned continuously. The selected ion monitoring (SIM) mode, in which a specific set of m/z values was monitored, was also used to increase the signal-to-noise ratio of selected ions.

When the solution flow was directed into the fraction collector instead of the mass spectrometer, the collection criteria used was time-based. The collection times were set upon analysis of HPLC-MS runs of the same injection amount of the same sample.

Data analysis was performed using the *ChemStation B.04.03* software (Agilent Technologies) as well as in *SigmaPlot* upon exportation of the data.

The typical volume of HSA sample injected for HPLC-MS analysis was 20 µl, which contained peptides from the digestion of about 110 µg or 1.68 nmole of HSA. For the purification of labelled HSA peptides for sequencing, up to 100 µl of peptide sample was injected in each run and fractions collected from several runs were pooled. However, for the GLIC samples, peptides from the digestion of about 50 µg or 1.45 nmole of a GLIC subunit were typically injected for each sample run for the analysis of the elution times of peptides and the investigation of the presence of labelled peptides. Similar amounts of GLIC samples were injected for the purification runs of peptides for sequencing without pooling fractions from multiple HPLC runs.

2.5.2 Tandem mass spectrometry

Peptides samples collected from the fraction collector were concentrated by evaporating the ACN under a stream of nitrogen gas. The concentrated samples were acidified by the addition of 0.2 % FA and 0.02 % TFA and analysed by a triple quadrupole mass spectrometer (Applied Biosystems, Warrington, UK) coupled to an Agilent 1100 Series HPLC stack consisting of a binary pump (Agilent Technologies, G1312A), a capillary pump (Agilent Technologies, G1376A), an autosampler (Agilent Technologies, G1367A), a column oven (Agilent Technologies, 1316A) and an integrated six-port valve.

Throughout the duration of this project, a few different configuration settings were used. Typically, samples were loaded onto a trap column (Agilent Technologies, Zorbax SB, 5 µm × 0.3 mm × 35 mm) using the binary pump. The binary pump delivered the loading/washing solvent (99.78 % H_2O, 0.2 % FA and 0.02 % TFA) at a flow rate of 150 µl min^{-1} for five minutes to remove salt impurities before the valve was switched into the capillary flow to direct the peptides into the capillary column (Agilent Technologies, SB, 5 µm × 0.5 mm × 150 mm) to be separated. The column temperature was set to 40 °C. The flow rate of the capillary pump was 10 µl min^{-1} and the elution profile was 100 % solvent C (94.9 % water, 5 % ACN and 0.1 % FA) / 0 % solvent D (5 % water, 94.9 % ACN and 0.1 % FA) for the first five minutes, then ramped to 40 % solvent C / 60 % solvent D in the next 45 minutes, then to 20 % solvent C / 80 % solvent D over the next two minutes and held at that composition for one minute. The gradient was subsequently reversed to 100 % solvent C / 0 % solvent D over the following minute, where it remained for the final seven minutes.

For sequencing polar peptides, a HPLC configuration setting was used that bypassed the capillary column. This was to avoid sample loss in the flowthrough of the trap column in the previously described HPLC configuration because polar peptides do not retain well on the column.

Eluted peptides were directed to the QTRAP mass spectrometer with a TurboIon Spray Source at 150 °C. The mass spectrometer was used either to directly acquire enhanced product ion (EPI) scans by selecting parent ion *m*/*z* to fragment and manually deciding the collision energy to use, or with an information-dependent acquisition (IDA) method, which consisted of a survey scan from 350 to 1600 *m*/*z*, an enhanced resolution scan and EPI scans for the four most abundant ions.

The mass spectra obtained from EPI scans were analysed with the software program *Analyst 1.5.1* (AB SCIEX, Warrington, UK) through manually sequencing the peptide with the aid of the Peptide Tools function in *ChemStation* to generate the *m*/*z* of the singly charged collision-induced dissociation (CID) fragments of specific peptide sequences. For spectra in which the peptide sequence could not be predicted based on its parent ion *m*/*z* with an unknown charged state, a list of CID fragment ion *m*/*z* could not be obtained from the Peptide Tools function for matching with peaks on the EPI mass spectra. In such cases, sequencing involved identifying high-intensity mass peaks on the EPI spectra that differed by an amino acid residue mass. Adjacent mass differences corresponding to different amino acids were then used as sequence tags to search the peptide sequence of the proteins used, much similar to the algorithm used by proteomics software such as *ProteinPilot* (AB SCIEX). *ProteinPilot 4.0* was used for the analysis of data from IDA methods (although individual EPI scan data could not be analysed with this software). User-defined modifications were incorporated into the sequencing algorithm to include the possibility of the addition of $C_{11}H_{11}F_3O$, the molecular formula of the D_0 photolabel, or of $C_{11}H_{10}DF_3O$, the molecular formula of the D_1 photolabel, to all amino acids. Restricted search databases were used that included the target proteins HSA and GLIC.

3 Characterisation of the photolabel

A novel diazirine analogue of the general anaesthetic propofol, 3-(5′-isopropyl-6′-hydroxyphenyl)-3-trifluoromethyl-diazirine (photofol), was synthesised by Dr Edward Smith and Dr Christopher Edge in our laboratory for use as a photolabel to probe the propofol-binding site of anaesthetic protein targets, such as $GABA_A$ receptors.

Before proceeding with photolabelling experiments, photofol was characterised both physically, in terms of its absorbance profile, stability and photolysis; and biologically, in terms of its anaesthetic activity in rats and action on $GABA_A$ receptors, a well-regarded anaesthetic ion channel target, expressed in HEK cells.

Two samples of the photolabel photofol were produced: pPD, with isotopes that had no deuterium atoms substituted (D_0), and dPD, consisting of a mixture of 34 % D_0, 43 % with one deuterium substituted (D_1) and 23 % with two deuteriums substituted (D_2). Only the pPD photofol sample was used in the experiments presented in this chapter. It is believed that the dPD sample would behave in the same manner in these physical and biological characterisations.

3.1 Physical characterisation of propofol and photofol

3.1.1 UV-Vis absorption profiles

The UV absorbance spectrum of photofol was investigated by comparing it with the absorption profile of propofol. The UV absorption spectra of propofol at 250 μM, 500 μM and 750 μM were obtained (Figure 3.1). The extinction coefficient (ε) at the peak wavelength of absorbance (275 nm) was calculated from the inset graph of Figure 3.1. The gradient of the line of best fit—fitted with the linear equation $y = mx + c$ using *SigmaPlot*—of the concentration against optical density (OD) unit graph is $2.487 \times 10^{-3} \pm 0.000088$, which is equlvalent to an ε value of $2490 \pm 90\ M^{-1}\ cm^{-1}$. This is comparable to $1900 \pm 200\ M^{-1}\ cm^{-1}$ for propofol dissolved in methanol (Heyne *et al.*, 2006). When the absorption profile of photofol was measured at concentrations of up to 700 μM at 100-μM intervals, the peak wavelength of absorbance was identified to have been slightly right-shifted to 283 nm (Figure 3.2). Through the same approach as with analysing the propofol data, the extinction coefficient of photofol at the peak wavelength was calculated to be $1970 \pm 20\ M^{-1}\ cm^{-1}$ (Figure 3.2, top inset).

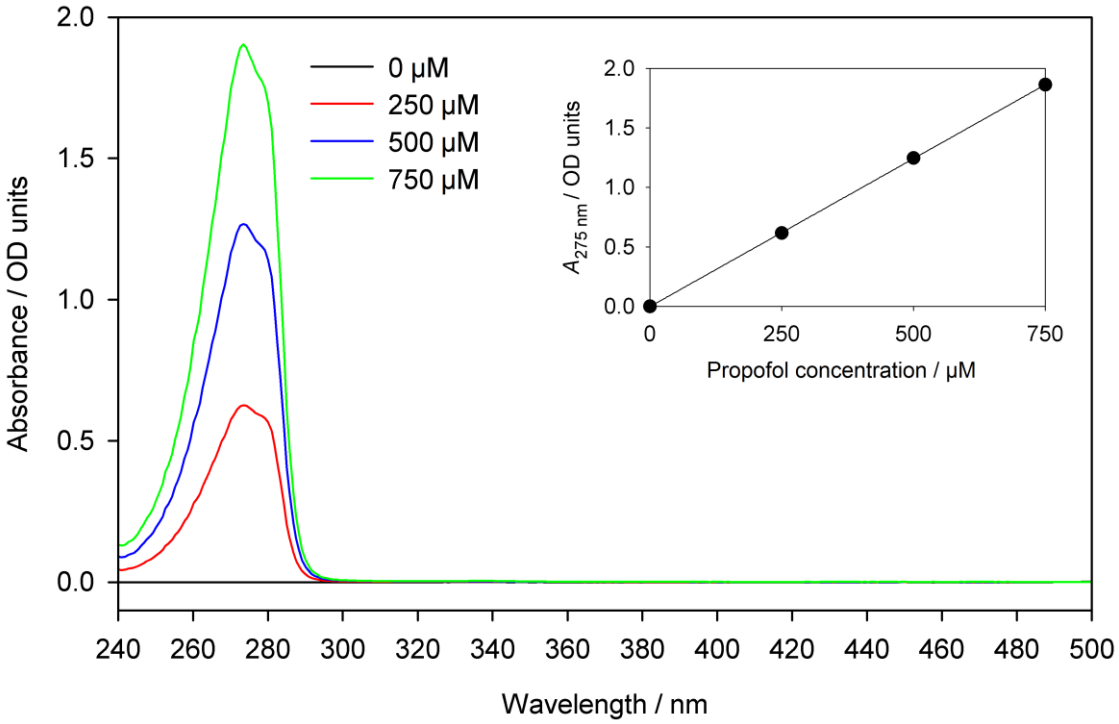


Figure 3.1: UV absorbance spectra of propofol shows its peak absorbance at 275.1 nm varies proportionally to the concentration of propofol. The extinction coefficient of propofol is calculated to be 2490 ± 90 M^{-1} cm^{-1} from the gradient of the line of best fit (inset).

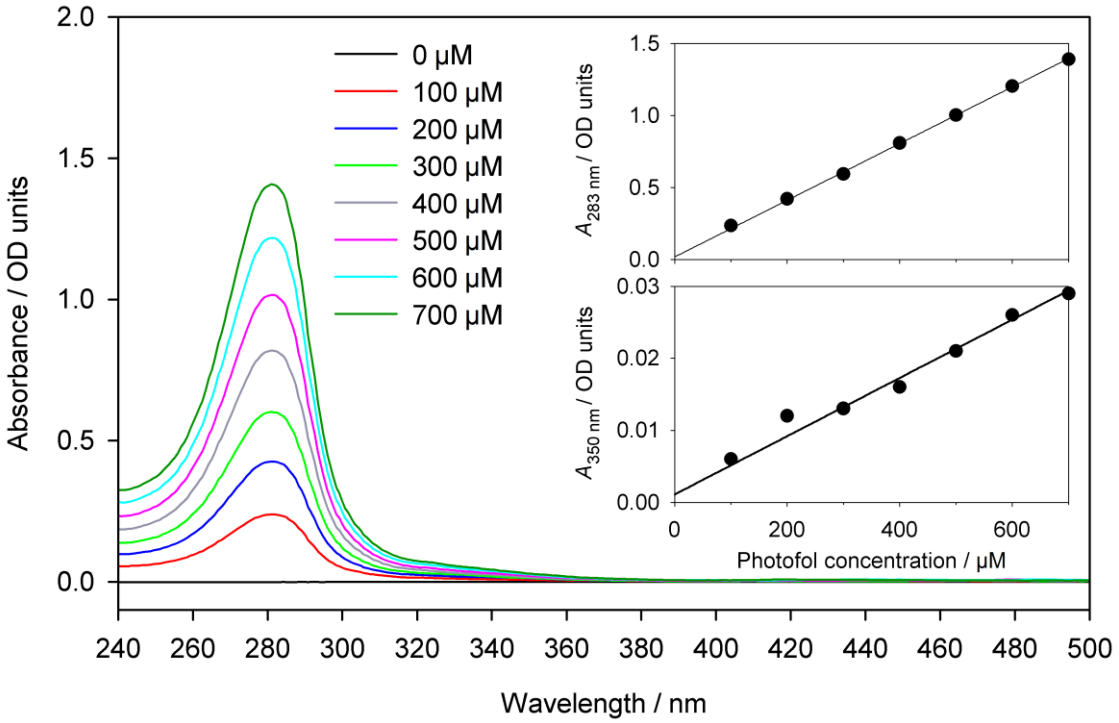


Figure 3.2: UV absorbance spectra of photofol shows its peak absorbance at 283.3 nm varies proportionally to the concentration of photofol. The extinction coefficient at λ_{max} is calculated to be 1970 ± 20 M^{-1} cm^{-1} (top inset) and at 350 nm to be 40.5 ± 2.5 M^{-1} cm^{-1} (bottom inset) from the gradients of fitted linear graphs.

Absorbance in the range of 300 nm and 370 nm that is exhibited by the photolabel is not seen by propofol. This is as expected as photofol has a diazirine group, which typically absorbs at 350 nm. The extinction coefficient of photofol at 350 nm was calculated to be 40.5 ± 2.5 M^{-1} cm^{-1} (Figure 3.2, bottom inset). This value is much lower than at $\lambda_{283\,nm}$. The absorbance by the diazirine moiety at this longer wavelength of 350 nm is likely to be due to the energy transitions of electrons from non-bonded (n) orbitals of the nitrogens into mainly the antibonding π^* orbitals. However, electrons in the n orbital are spatially quite far from those in the π^* orbitals, thus the probability of such n transition is low. Since ε depends on the number of electrons undergoing transitions, $\varepsilon_{350\,nm}$ is relatively low compared to that of $\varepsilon_{283\,nm}$, which mainly results from transitions of π to π^* orbitals.

3.1.2 Aqueous stability and photolysis in ethanol

The aqueous stability and the photolysis—by light with a wavelength of 320–400 nm—of propofol in ethanol were investigated as a control for the characterisation of photofol with the same parameters. Propofol was found to be stable when dissolved in ethanol; there was no change in its UV absorption spectrum over a period of two hours left at room temperature. Exposing propofol (750 μM) dissolved in ethanol to light of wavelengths in the range of 320–400 nm also had no effect on the absorbance profile of propofol when exposed for a period of 60 seconds. UV absorption spectra were taken after successive ten-second exposures and the spectra all overlap each other. The aqueous and photolytic stabilities of propofol over a longer duration were not tested.

The stability of photofol in ethanol at 500 μM at 25 °C was investigated. The photofol sample was already dissolved in ethanol at 250 mM while stored at –20 °C. The photofol stock was diluted to 500 μM and left at room temperature inside a 4-ml quartz cuvette while covered by Parafilm but exposed to diminished ambient light. UV-Vis spectra were obtained at 30-minute intervals. The absorbance at λ_{max} was not altered by more than 5 % of the initial absorbance (A_0) for spectra obtained from the first seven hours. Upon leaving the sample for 24 hours and 48 hours, two more UV-Vis spectra were obtained. An overlay of these spectra and those from the initial seven hours of storing photofol at room temperature is shown in Figure 3.3. After one day of storage at room temperature at 500 μM, there was a reduction in absorbance of about 16 % from A_0 and after two

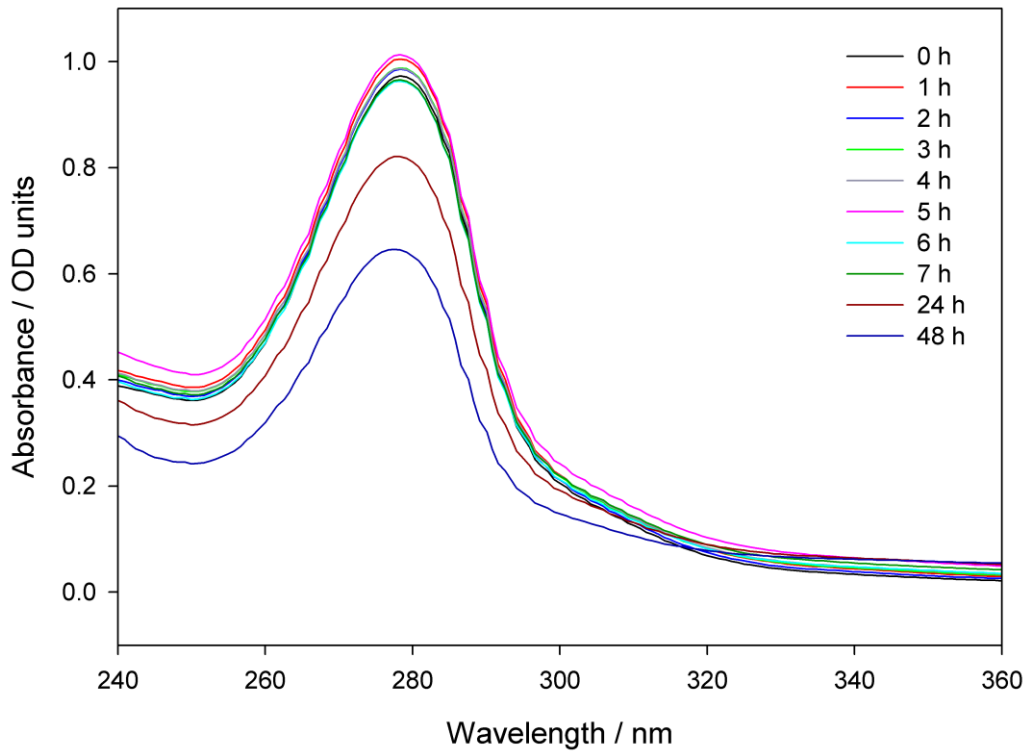


Figure 3.3: UV absorbance spectra overlay showing the aqueous stability of photofol of a 800 µM sample dissolved in ethanol and left for increasing durations at 25 °C under diminished ambient light.

days, a 34 % reduction from A_0 was seen. Over the duration of a typical photolabelling experiment with proteins, there should be negligible hydrolytic decomposition of the photolabel during the preparation stage before the drug undergoes photolysis in the presence of their target proteins.

Photofol was diluted to 0.2 mg ml^{-1} in ethanol for photolysis with a 500 W mercury (xenon) light source at a wavelength range of 320–400 nm. Over the duration of five minutes, the sample was progressively exposed to increasing durations of this light to allow the reaction scheme described in Figure 3.4 to occur and the UV absorption profile of the reaction mixture was monitored regularly.

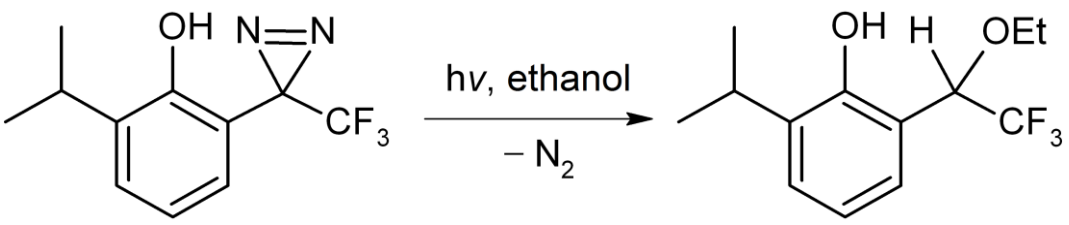


Figure 3.4: Reaction scheme showing the photolysis of photofol in ethanol.

The overlay of some of the spectra obtained in the initial two minutes of light exposure is shown in Figure 3.5. Increases in absorbance at the absorption maxima and at 350 nm—the characteristic wavelength of absorbance of diazirines—are seen upon increasing the duration of light exposure. There is also a slight shift of the absorption maxima to a lower wavelength upon photolysis (Figure 3.5, inset).

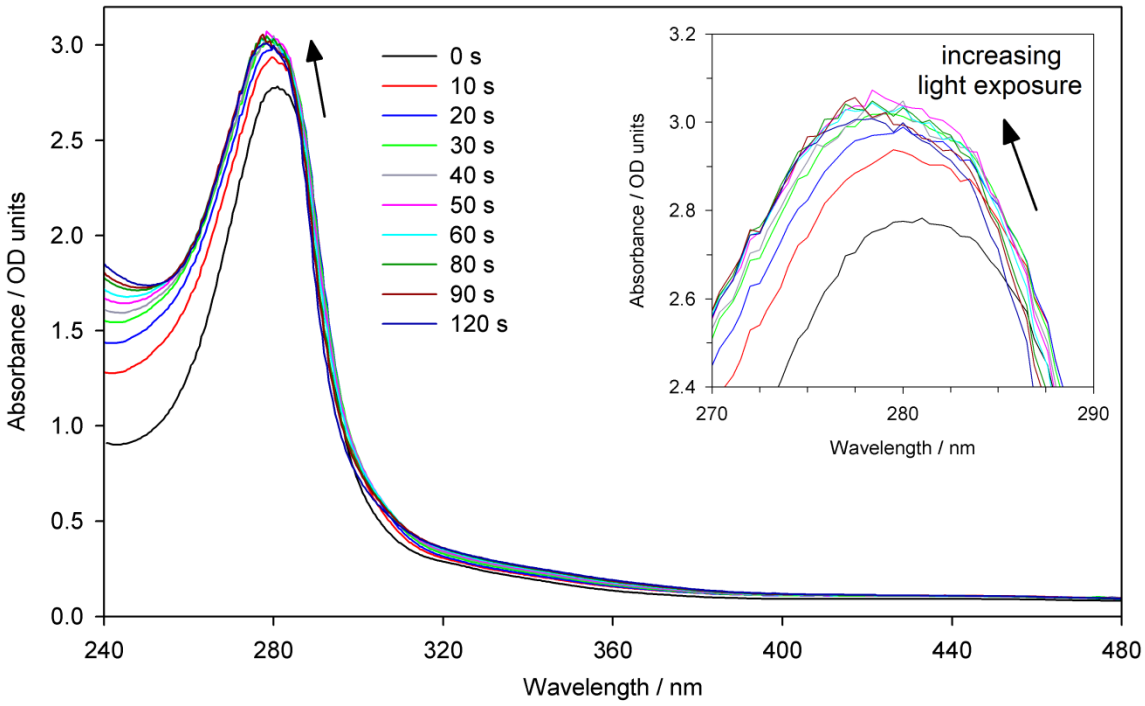


Figure 3.5: UV absorbance spectra overlay showing the photolysis of photofol by increasing durations of light exposure on a 0.2 mg ml^{-1} sample dissolved in ethanol. Arrows mark the shift in the wavelength of maximum absorbance around 280 nm, for which an enlarged spectra is shown in the inset.

The rate of photoconversion of photofol in ethanol was determined by plotting the time-dependent increase in absorbance at 350 nm and fitting the data to the exponential decay equation $A_{350\,nm} = A_0 + ae^{-bt}$, where $A_{350\,nm}$ is the absorbance at 350 nm, A_0 is the initial absorbance, a is the scaling constant, b is the rate constant and t is time. The time constant (τ) of the reaction can then be calculated from the reciprocal of the rate constant (b). Figure 3.6 shows the change in absorbance at 350 nm for both photofol and propofol. It is clear that upon irradiation, there is no absorbance at 350 nm by propofol, as with before irradiation. However, the initial absorbance of 0.164 OD units at this wavelength by 0.2 mg ml^{-1} photofol was increased progressively with a time constant, τ of 42.0 seconds over the initial 100 seconds of light exposure. The fitted graph then plateaus off at 0.229 OD units.

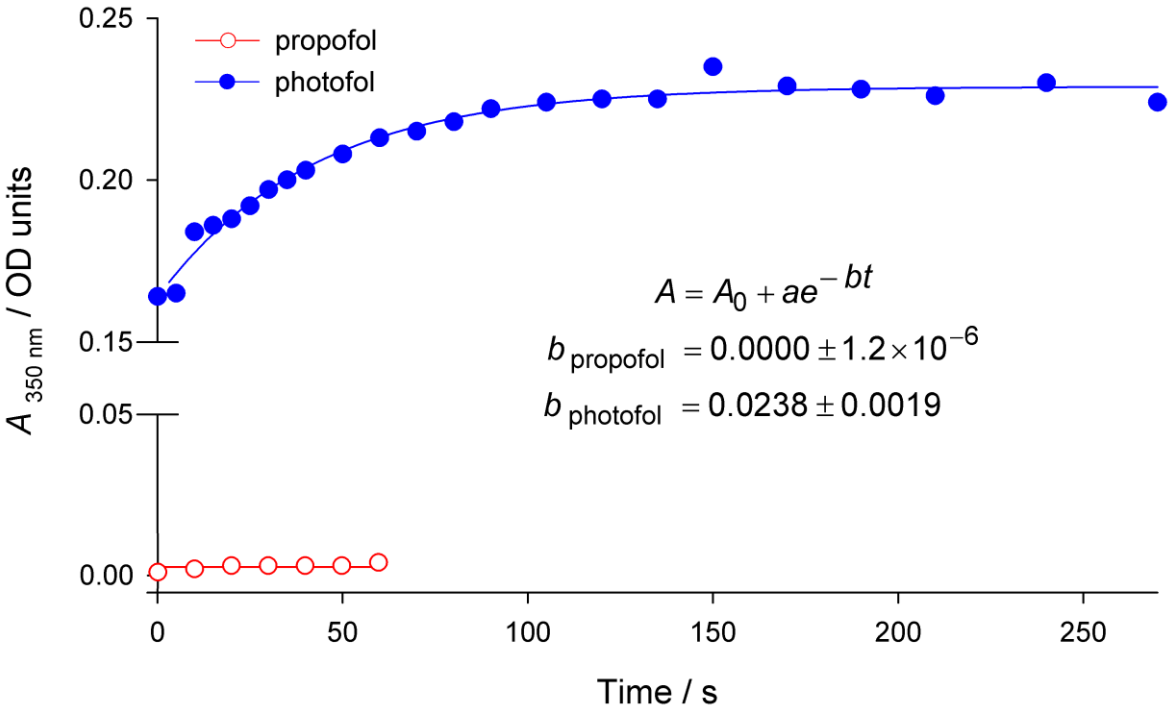


Figure 3.6: Absorbance at 350 nm of photofol and propofol as a function of irradiation duration. The photofol (red) and the propofol (black) data were both fitted with the exponential decay equation shown and the rate constant (*b*) for each set of data is displayed.

The absorbance at 350 nm is expected to decrease upon increasing the duration of light exposure of photofol due to the conversion of the diazirine into carbenes, which react with ethanol to form an ether. Despite the unexpected increase in absorbance at 350 nm, the changes at this wavelength indicate that the photolysis reaction reaches near completion after about 120 seconds of light exposure, thus this was the chosen exposure duration for the photolabelling experiments. The unexpected increase could be contributed to by-products from the photolysis reaction shown in Figure 3.4, of which one of these end products may absorb strongly at this wavelength band, due possibly to containing an extensive conjugated π system.

3.2 Biological characterisation of the photolabels

The feasibility of photofol as an anaesthetic, acting in a similar fashion to propofol, was investigated both *in vivo* and *in vitro*. *In vivo* experiments were carried out by examining the loss of righting reflex (LORR) in rats post-tail-vein injection of the drug and *in vitro* experiments were conducted by whole-cell patch clamp experiments on HEK cells expressing $GABA_A$ $\alpha_1\beta_2\gamma_{2S}$ receptors.

3.2.1 Loss of righting reflex in rats

The following experiments were carried out by Dr Christopher Edge and Dr Thomas Gent. LORR is a common assay for measuring unconsciousness in rodents. The righting reflex is the ability of an animal to orient itself and put its paws on the ground when placed on its back. To determine the potency of propofol and photofol in male Sprague-Dawley rats weighing 498 ± 7.7 g, these drugs were administered by tail-vein injection via a cannula and the rat was immediately placed in a rotating chamber to assess its righting reflex. The rat was judged as anaesthetised if it failed to right itself by placing all four paws on the ground one minute post-injection. The quantal dose-response curve, where a response of zero represents being awake and one represents being anaesthetised, was fitted to the logistic equation of Waud (1972).

Twelve rats were injected with propofol and eleven with photofol for the LORR assay. Both drugs were able to cause reversible LORR in rats. Figure 3.7 shows the dose-response data for both anaesthetics. When the data was fitted using the equation of Waud. The dosage of photofol with which 50 % of the population were anesthetised, ED_{50}, was determined to be 4.36 ± 0.54 mg kg^{-1},

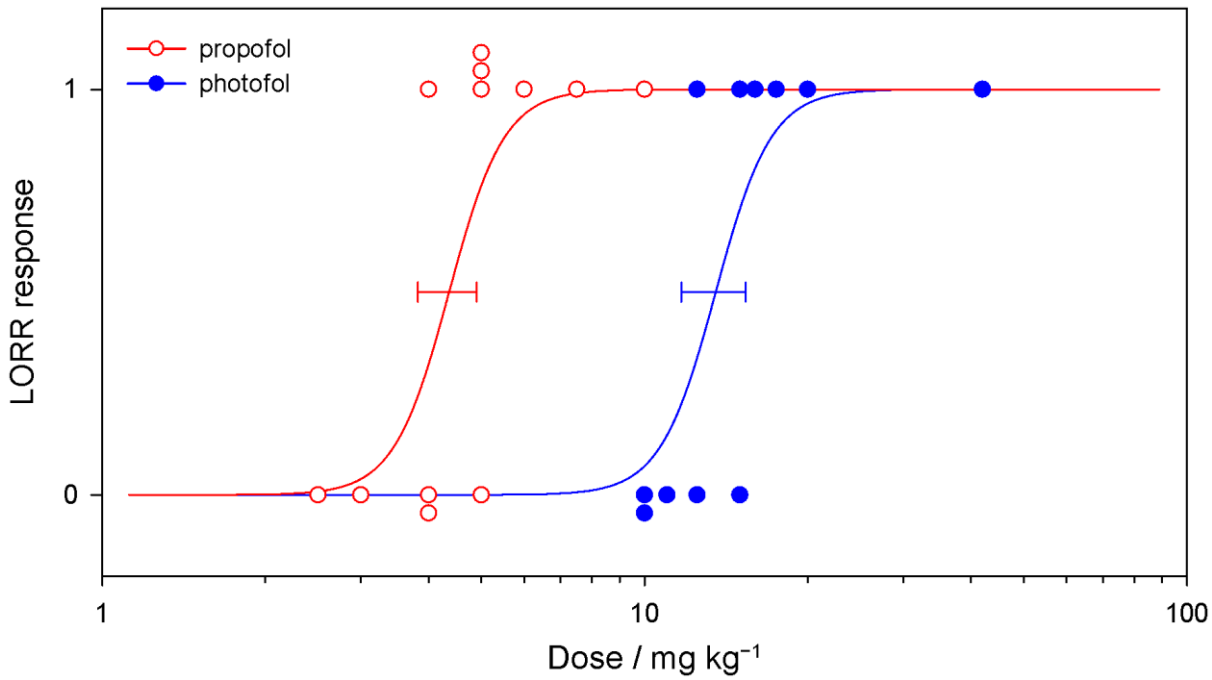


Figure 3.7: Quantal dose-response curves for the loss of righting reflex of rats administered with propofol (red) or photofol (blue) via tail-vein injection. Each circle represents an individual animal administered with one of the drugs. A response of zero denotes an awake rat and one denotes an anaesthetised rat. The data was fitted to the Waud equation (Waud, 1972). Error bars intercepting the fitted curves mark the ED_{50} dose ± standard error.

whilst for photofol, the ED_{50} was 13.5 ± 1.8 mg kg^{-1}. The slope factors, *n*, of both anaesthetics were 8.5 with standard errors of ±7.6 and ±7.2 for propofol and photofol respectively.

3.2.2 Whole-cell patch clamping in $GABA_A$ receptor-expressing HEK cells

Whole-cell patch clamp experiments testing the effects of photofol on the GABA-evoked response of $GABA_A$ $\alpha_1\beta_2\gamma_{2S}$ receptors expressed in HEK cells were conducted by Dr Robert Dickinson. The $\alpha_1\beta_2\gamma_{2S}$ $GABA_A$ receptor was selected due to it being the most abundant subunit combination observed in the mammalian central nervous system (McKernan and Whiting, 1996). An EC_7 GABA concentration of 3 µM was chosen for investigating the modulatory response of propofol and photofol. This low GABA concentration induces currents of approximately 1 nA with essentially no desensitisation and is potentiated significantly by propofol and photofol compared to higher GABA concentrations. The anaesthetic propofol at 3 µM or photofol at 5 µM was pre-applied to cells using a local perfusion system, which allowed for fast solution exchange into the test solution, containing the appropriate concentration of anaesthetics and 3 µM GABA. A typical current trace showing the potentiation of the GABA-induced current by photofol is shown in Figure 3.8. Control GABA responses were measured before and after the co-application of

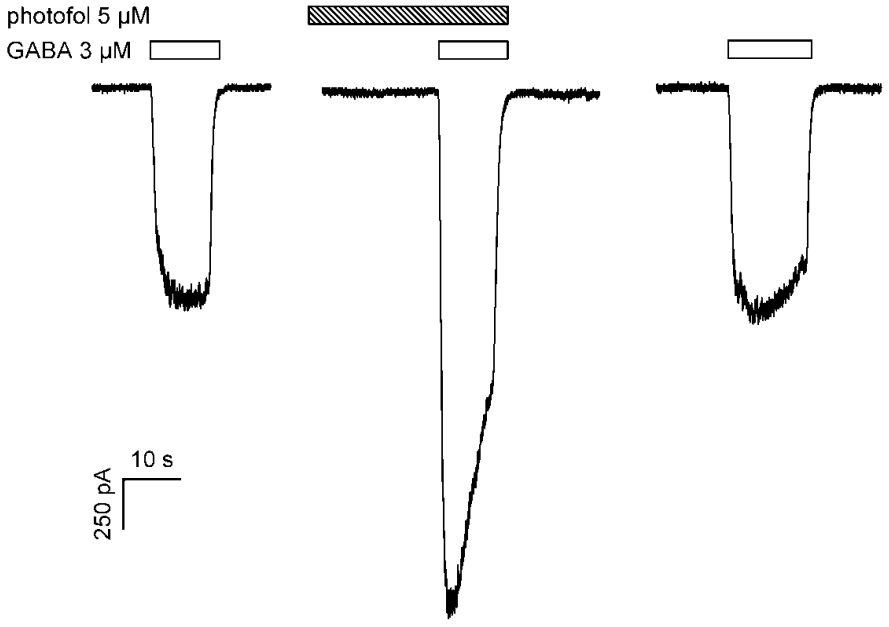


Figure 3.8: Typical current traces showing the EC_7 GABA concentration being potentiated by approximately 100 % by the co-application of 5 µM photofol and EC_7 GABA. The voltage-clamped currents induced by an EC_7 GABA concentration before and after the co-application of GABA and the anaesthetics are about 1 nA.

anaesthetics and GABA. The mean control current was used in the calculation of the percentage potentiation of the current response in the presence of the anaesthetic with the following equation:

$$P(\%) = 100 \times \frac{I - I_0}{I_0}$$

where $P(\%)$ is the percentage potentiation by the drug, I is the peak of the GABA-induced current in the presence of the drug and I_0 is the peak of the control GABA-induced current.

The percentage potentiation of EC_7 GABA-evoked current by 3 μM propofol was calculated to be 188 ± 34 % (n = 12) and for 5 μM photofol to be 164 ± 45 % (n = 12), indicating reduced modulation of the GABA-evoked currents by photofol compared to its parent analogue, propofol for the drug-concentration tested.

3.3 Concluding remarks for the characterisation of photofol

Physical and biological characterisations were carried out on the novel photolabel and compared with propofol. The absorbance peak at the maximum wavelength of propofol was shifted by approximately 10 nm to a longer wavelength upon replacing one of its isopropyl groups with an alkyldiazirinyl group. The UV spectra of photofol also showed absorbance intensity between 300 and 380 nm which is absent in that of propofol. The extinction coefficients at the maximum wavelengths for propofol and photofol were 2490 ± 90 $M^{-1}\,cm^{-1}$ and 1970 ± 20 $M^{-1}\,cm^{-1}$ respectively, whilst the extinction coefficient for photofol at 350 nm—the typical absorbance of diazirine moiety—is much less at 40.5 ± 2.5 $M^{-1}\,cm^{-1}$, owing to the less frequent energy transitions of electrons in n orbitals to π* orbitals. The novel drug was shown to be stable in solution at room temperature under diminished light—much like the conditions during the sample preparation stage prior to UV exposure in a photolabelling experiment—when it was left for seven hours but did not show more than 5 % change in its absorbance at λ_{max} (Figure 3.3).

When photofol was photolysed by a filtered mercury light source with a wavelength range of 320–400 nm, a small left shift of the maximum peak wavelength, as with an increase in absorbance intensities at around 350 nm, were seen with increasing durations of photolysis (Figure 3.5). The increase in absorbance had a time constant of 42 seconds and reached equilibrium after about two

minutes of light exposure (Figure 3.6), thus this duration was chosen for photolabelling experiments.

The photolabel was tested for its anaesthetic actions *in vivo* through LORR assays in rats and *in vitro* through whole-cell patch clamp experiments on HEK cells expressing $GABA_A$ receptors. The potency of photofol was identified to be 3.1- and 1.9-fold less than propofol in the *in vivo* and *in vitro* experiments respectively. It is promising that both of these experiments showed that photofol exhibits anaesthetic actions, be it with less potency. A reason for the reduced potency could be due to the presence of a more bulky alkyldiazirinyl group that substituted one of the isopropyl groups at the *ortho* position to the hydroxyl group of propofol. These isopropyl groups already hinder hydrogen bond formations of propofol (Leon *et al.*, 2012), thus the bulkier replacement photoactivatable group is likely to cause more hindrance to the interaction of the hydroxyl group with its target. These hydrogen-bonding interactions of propofol have been shown to play a crucial role in its binding to $GABA_A$ receptors via electrophysiology studies (Richardson *et al.*, 2007), and to HSA (Bhattacharya *et al.*, 2000) and GLIC (Nury *et al.*, 2011) from clear evidence of hydrogen bond interactions in their propofol-bound X-ray structures. This explains the potential weaker binding of photofol to its anaesthetic targets compared to propofol.

4 Photoaffinity labelling of human serum albumin

Photoaffinity labelling has served as a useful tool for studying molecular interactions between proteins and their ligands, following the steps that are briefly outlined in Figure 4.1. This technique is especially useful for elucidating the binding sites of proteins for which crystallographic data is hard to obtain, such as ion channel proteins. With ion channels being the accepted target of general anaesthetics, it is therefore appropriate to use anaesthetic photolabels to probe the binding sites of these drugs. The novel anaesthetic photofol, synthesised in our laboratory, has been tested for its anaesthetic actions on rats and in HEK cells expressing $GABA_A$ receptors and was found to have a 3.1- and 1.9-fold reduction in potency respectively compared to propofol (section 3.2). Its validity as a photoaffinity label was tested by labelling the soluble protein human serum albumin (HSA)—the results of which are presented in this chapter—and the pentameric ligand-gated ion channel GLIC from the prokaryote *Gloeobacter violaceus*—presented in chapter 5—for which X-ray crystallographic data of propofol bound to these proteins are already available (Bhattacharya *et al.*, 2000; Nury *et al.*, 2011).

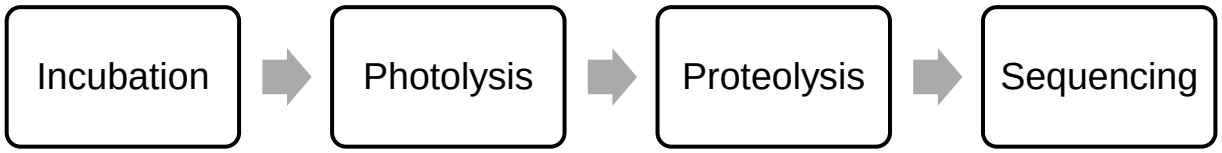


Figure 4.1: Principal steps in photoaffinity labelling experiments of proteins.

The X-ray structure of HSA with propofol bound was solved at a resolution of 2.4 Å (Bhattacharya *et al.*, 2000), presenting two propofol binding sites: propofol site 1 in subdomain IIIA of HSA and propofol site 2 in subdomain IIIB. Our novel anaesthetic photofol was used to label HSA to replicate the observed binding sites in the crystallographic data. Photoaffinity labelling of HSA with an etomidate diazirine analogue has been carried out by our group (Bright *et al.*, 2007). The etomidate analogue labelled two lysine residues: one in propofol site 2 (Bhattacharya *et al.*, 2000) and one in drug binding site 1 (Sudlow *et al.*, 1975; Sudlow *et al.*, 1976) of HSA. A similar approach was taken to label HSA with the novel photolabel, photofol, beginning with the characterisation of tryptic digests of the protein from HPLC-MS.

4.1 Identification of tryptic digests of HSA from HPLC-MS

The *in silico* digestion of HSA with trypsin yields 81 peptides. The mass spectrometer that was used in this project detects ions with a mass-to-charge ratio (*m*/*z*) of between 100 and 1500. Assuming that it can detect peptides that are at least three amino acids long, 55 tryptic peptides, corresponding to 59 % amino acid coverage of HSA and having a molecular weight (MW) of between approximately 300 and 1500 Da, are potentially identifiable by their singly charged ions. An additional thirteen peptides are theoretically detectable by their ions with multiple charges as their masses range from 1500 to 3000 Da. Trypsin digestion of HSA followed by analysis with our mass spectrometer set up could lead to 96.6 % amino acid coverage of HSA.

HSA was denatured and cysteine-modified as outlined in section 2.3.2, then dialysed with ammonium bicarbonate buffer before being digested with trypsin as outlined in section 2.3.3. The peptide sample was then separated by HPLC on a reverse phase column with an increasing gradient of acetonitrile (ACN). The separated peptides were monitored at 214 and 280 nm by a multiple wavelength detector, and loaded directly into an ESI-mass spectrometer.

The absorbance chromatograms at 214 nm, the wavelength at which peptide bonds absorb strongly, and at 280 nm, which indicates the presence of peptides containing aromatic amino acids, are shown in Figure 4.2A. The total ion current (TIC) profile of the HSA tryptic digest is shown in Figure 4.2B and shows the sum of the current from all ion species passing through the mass spectrometer in the range of the *m*/*z* being detected at any given time. For a given peptide with a molecular weight of M, ESI produces charged species of $(M + nH)^{n+}$, so for example, a peptide with a molecular weight of 1200 Da might generate ions of 1201, 601, 401 and 301 *m*/*z* for $(M + H)^{+}$, $(M + 2H)^{2+}$, $(M + 3H)^{3+}$ and $(M + 4H)^{4+}$ respectively.

To determine the elution times of the tryptic peptides of HSA predicted by *in silico* digestion, the Extract Ions tool in the software program *ChemStation* was used. The extracted ion current (XIC) chromatogram that is outputted shows the current intensity for a selected *m*/*z* value or window as a function of time.

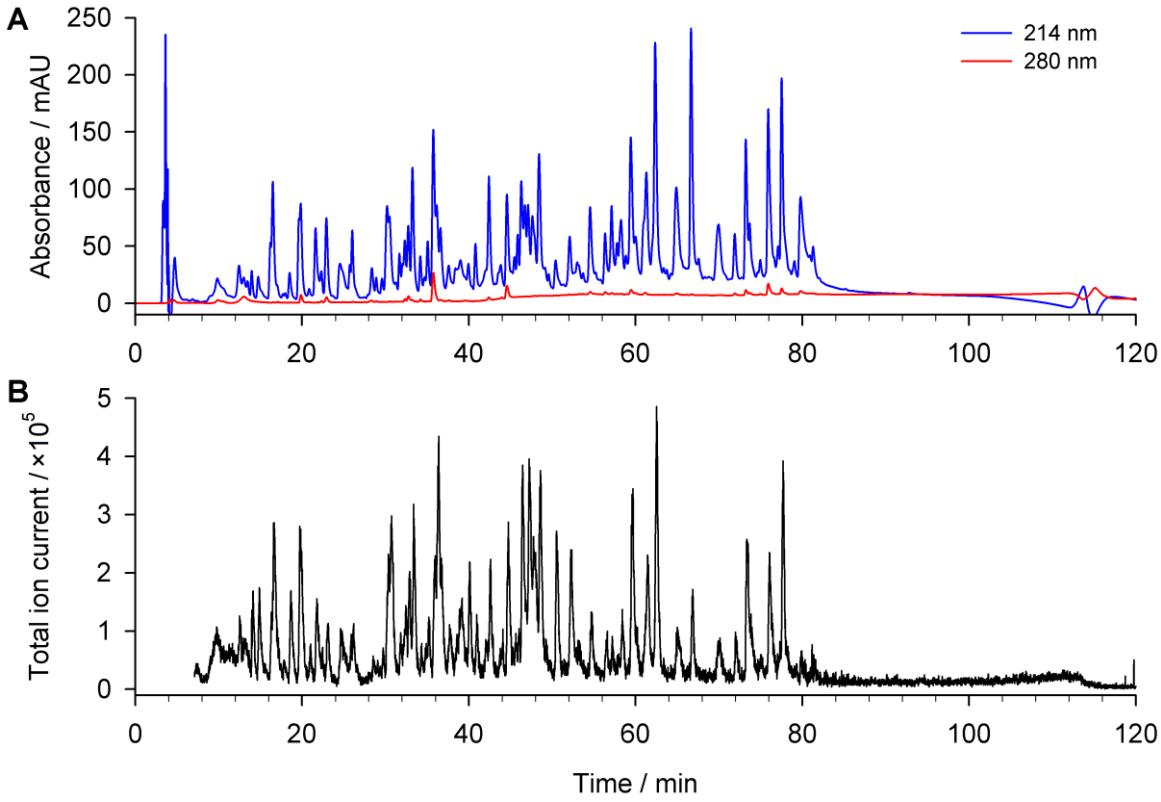


Figure 4.2: Typical absorbance chromatograms at 214 and 280 nm (A) and the total ion current chromatogram (B) monitoring the separation of HSA tryptic peptides. HSA was denatured, reduced, alkylated by iodoacetamide and digested by trypsin as outlined in sections 2.3.2 and 2.3.3. 20 µl of the final reaction mixture was injected into the HPLC-coupled mass spectrometer to be separated by a reverse phase column using a gradient elution of increasing ACN.

By extracting ions for three different charged species—ions with either one to three charges or two to four charges—for each of the predicted peptides in turn and recording the time points in which at least two peaks in the XIC traces aligned (Figure 4.3), it was possible to identify all of the expected peptides having a MW of more than 500 Da with certainty. Figure 4.3 illustrates the search for the elution time of the tryptic peptide $_{163}$AAFTE<u>CC</u>QAADK$_{174}$ with a MW of 1370.6 Da, where <u>C</u> denotes a cysteine residue modified by iodoacetamide alkylation. A peak from each of the chromatograms extracting the singly (1371.6 *m*/*z*), doubly (686.3 *m*/*z*) and triply (457.9 *m*/*z*) charged ions of the peptide aligns at 32.0 minutes, which suggests the elution of this peptide at this time when the concentration of ACN in the eluent is 19 %.

For smaller peptides or those that did not exhibit an alignment of peaks from the XIC chromatograms of different charged species, notable peaks in the XIC graph for the $(M + H)^+$ ion were further investigated as follows: a mass spectrum was obtained by averaging all the spectra

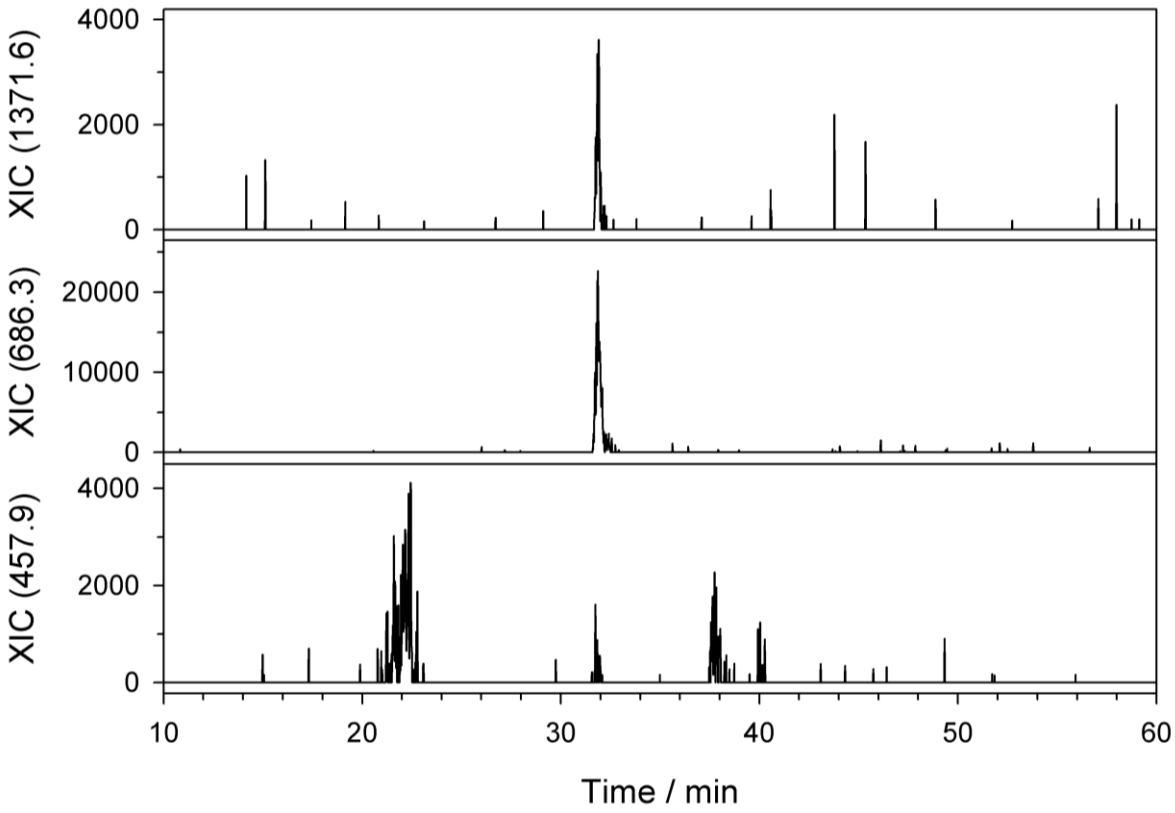


Figure 4.3: Extracted ion current chromatograms for the singly, doubly and triply charged ions of the peptide with a MW of 1370.6 Da that has been cysteine-modified, $_{163}$AAFTECCQAADK$_{174}$. The elution time of this peptide was identified to be at 32.0 minutes.

spanning the duration of each peak and the isotopic distribution pattern of the extracted-ion mass was checked to identify the charge state of the ion that gave rise to the XIC signal (Figure 4.4). Figure 4.4A shows that extracting ions of 470.2 and 235.6 *m*/*z* to identify the elution time of the first predicted tryptic peptide of HSA, $_{1}$DAHK$_{4}$, resulted in no alignment of peaks. The mass spectra of the two peaks present in the XIC spectra for the singly charged ion of $_{1}$DAHK$_{4}$ (top panel of Figure 4.4A) at around 29.6 minutes is shown in Figure 4.4B and at around 30.4 minutes is shown in Figure 4.4C. The mass spectra show that the first of these two elution times for the 470.2 *m*/*z* ion arose from a singly charged species whilst the XIC (470.2) peak at the latter elution time is due to the presence of a peptide with a MW of 939.4 Da, where 470.2 *m*/*z* is the *m*/*z* of its doubly charged ion.

However, overall, there were still some ambiguities over a few peptides where XIC signals of their charged species aligned at more than one time point, or where more than one peak in the $(M + H)^+$-value XIC chromatogram had arisen due to singly charged ions as confirmed by their

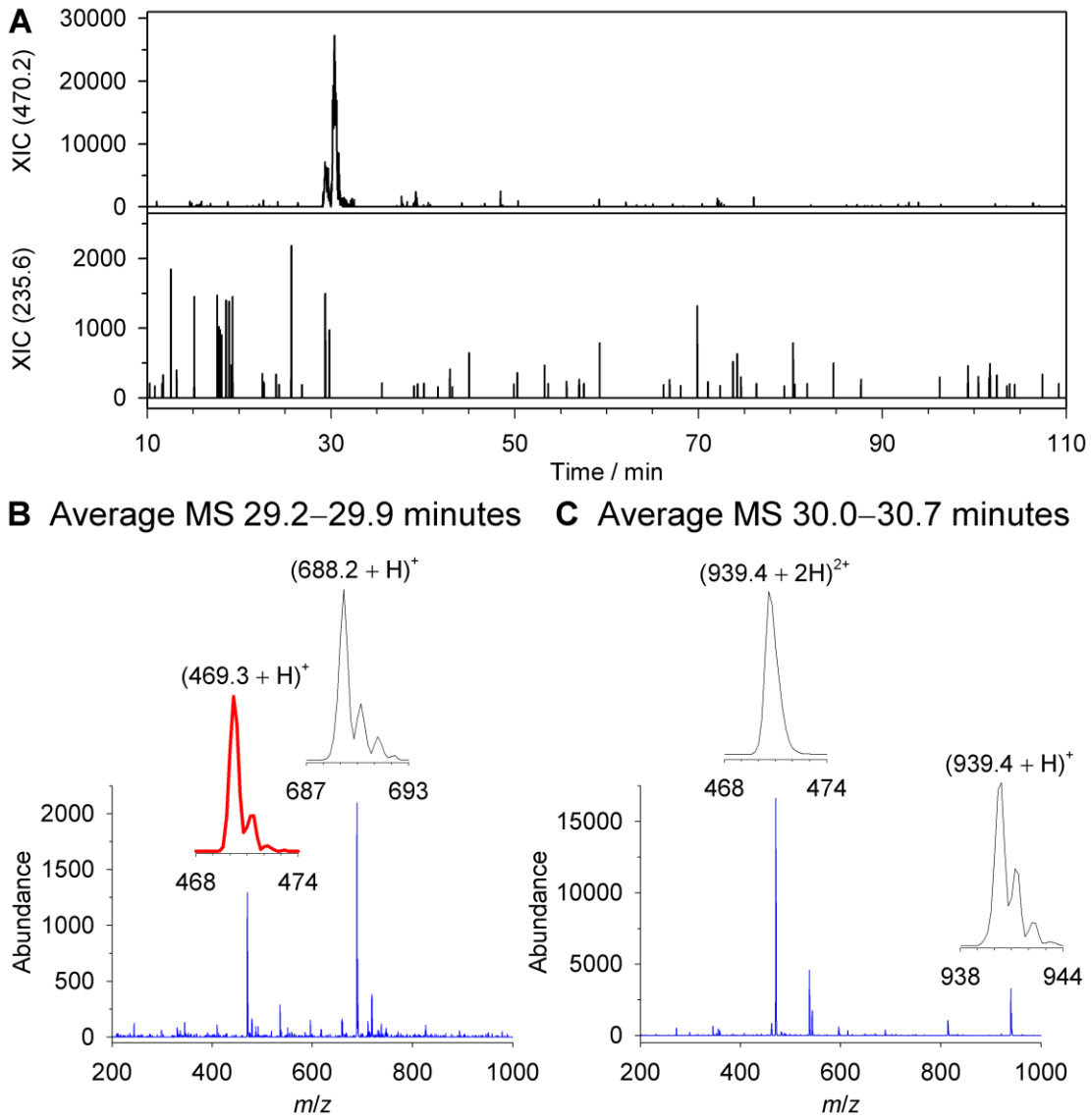


Figure 4.4: Identification of the elution time of $_{1}$DAHK$_{4}$ with a MW of 469.2 Da by extracting ions of its singly and doubly charged ions (A) followed by investigating the charged state of its *m/z* peaks in the mass spectra spanning the two XIC peaks at 29.2–29.9 minutes (B) and at 30.0–30.7 minutes (C). Insets in B and C show the enlarged mass spectra of the significant *m/z* peaks with the molecular weight of the peptide and the charge state of its ions noted above the maximum intensity peak. The singly charged ions of the peptide $_{1}$DAHK$_{4}$ in the enlarged mass spectra (red) show a separation of 1.0 *m/z* unit between its isotopic peaks.

isotopic distribution patterns. These occur when there are two or more peptides with similar masses that are not distinguishable by the resolution and sensitivity of the mass spectrometer used. Some of these ambiguities were resolved by sequencing with MS-MS (explained in more detail in section 4.6). For example, the peptides $_{258}$ADLAK$_{262}$ and $_{542}$EQLK$_{545}$ have a monoisotopic masses of 516.2908 and 516.2909 Da respectively, so their elution times were determined by sequencing the peptides of mass 516.3 Da eluted at 16.8 and 18.3 minutes (see Figure A.1 in the appendix for

the two EPI mass spectra). A list of all the identified HSA tryptic peptides resulting in amino acid sequence coverage totalling 97.3 %, along with their elution times is presented in Figure A.2 of the appendix. It should be noted that this percentage is greater than originally expected as it was possible to identify the elution times of some dipeptides.

4.2 The photolabelling experiment samples

Photolabelling was carried out with 10 mg/ml HSA in PBS with 1.5 mM of photofol. Two samples of photofol were used in the labelling experiments: the protiated form of the diazirine (pPD) and the photofol mixture with a specific ratio of its isotopes with various degrees of deuteration (dPD), as detailed in section 1.3. Photolysis experiments of pPD were conducted in ethanol; the results indicated that upon two minutes of irradiation with a wavelength range of 320–400 nm, all of the photoactivatable moiety had converted into their final products (section 3.1.2). It was assumed that the dPD photofol sample has the same property as pPD in terms of its reactivity and would also be completely activated upon light exposure for two minutes. Consequently, the samples were exposed to light (hν) for two minutes as necessary, resulting in the following sample conditions:

1. HSA + ethanol (EtOH)
2. HSA + ethanol (EtOH) + hν
3. HSA + pPD + hν
4. HSA + dPD + hν

These samples were then denatured, cysteine-modified and digested with trypsin as outlined in sections 2.3.2 and 2.3.3 before being analysed using HPLC-MS.

Comparing the two ethanol control experiments with and without exposure to UV light shows the effect of exposing HSA to light (Figure 4.5). The UV and TIC chromatograms of the light-treated sample generally overlap well with the traces for the unexposed sample. Small decreases in TIC are seen at 45.8 and 65.2 minutes for the light-exposed sample compared to the unexposed sample. Besides the small reductions in absorbance seen at the same time points for the two monitored wavelengths upon irradiation, there are peaks at 39.0, 39.5, 41.2 and 66.2 minutes where a minor increase is seen in the light-exposed sample compared to the unexposed sample. However, these changes between the UV fingerprints of the two HSA + ethanol control samples are only minute,

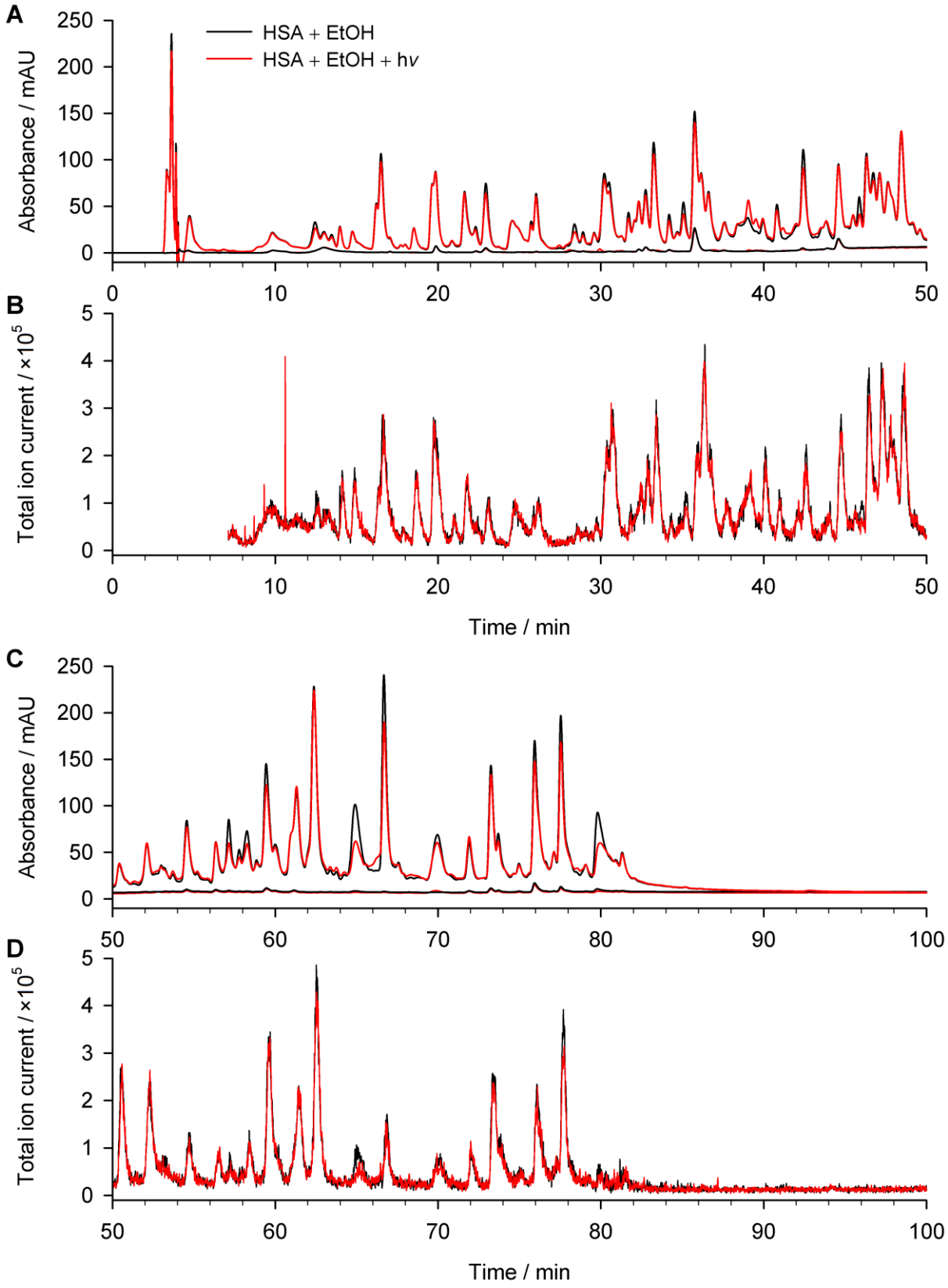


Figure 4.5: Chromatogram overlays of absorbance at 214 and 280 nm (A and C) and total ion current (B and D) for the separation of tryptic peptides of HSA that was (red) and was not exposed (black) to UV light prior to protein denaturation and trypsin digestion.

suggesting that the two-minute exposure to filtered light having a wavelength range of 320–400 nm has a minimal effect on the gross structure of the protein.

Comparing the HPLC-MS data of these control samples with that of the labelled samples allows the identification of peptides that have a photolabel attached, which is examined in the next section.

4.3 Identifying labelled HSA peptides

When HSA was exposed to light in the presence of photofol, the photolabel covalently bound to an amino acid and remained bound during tryptic digestion of the protein into peptides. The retention times of the labelled peptides, along with their molecular weights—which are greater by the mass of the photolabel compared to their unlabelled parent peptides—are identified and presented in this section.

Photofol, with the chemical formula $C_{11}H_{11}ON_2F_3$, has a MW of 230.1 Da. Upon photolysis, its two nitrogen atoms are eliminated to produce a carbene intermediate (Figure 4.6A). The carbene moiety is capable of inserting even into C–H bonds (Doering and Knox, 1956) or it can rearrange to form a quinone methide (Figure 4.6B), which can attach to hydroxy, amino and mercapto groups by nucleophilic reaction. Both the carbene and quinone methide have a MW of 216.1 Da. When these reactive intermediates covalently bind to amino acids, the tryptic peptides containing the labelled amino acid increase in molecular mass by 216.1 Da. The dPD samples primarily consist of an abundance of the photofol isotope D_1, which has a MW of 231.1 Da due to one deuterium replacing a protium on the benzene ring. Most of the dPD-labelled peptides therefore have a 217.1-Da addition to the native peptide MW, whilst isotopic peaks corresponding to

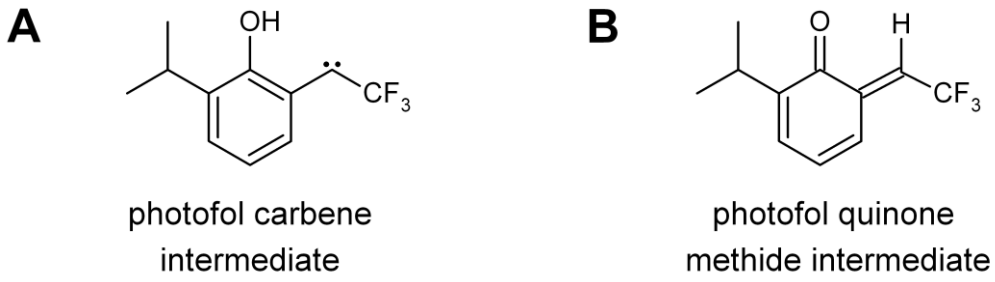


Figure 4.6: Chemical structures of the carbene (A) and quinone methide (B) reactive intermediates of photofol, both of which have the chemical formula $C_{11}H_{11}OF_3$ and a molecular weight of 216.1 Da.

216.1- and 218.1-Da additions can also be seen due to the presence of D_0 and D_2 photolabels at 34 % and 23 % in the dPD samples respectively.

The labelled peptides are likely to have a different retention time compared to the unmodified peptides due to their differences in mass and polarity. This results in the appearance of new peaks in the 214-nm spectra, which monitors mainly absorbance by peptide bonds, and also in the 280-nm spectra due to the presence of an aromatic ring that absorbs strongly at this wavelength on the peptide-attached photolabel. The appearance of new peaks in the 214-nm spectra as a result of a labelled peptide occurs in parallel with the reduction in size or the disappearance of another peak as the concentration of the unmodified peptide decreases.

4.3.1 Identifying the retention times of labelled peptides

Since almost all peaks in the 214-nm spectra of the peptide samples correlates with a peak in the TIC spectra, for the labelled HSA samples, it is anticipated that new peaks—corresponding to the ionisation of the labelled peptides being eluted at a different retention time to their unmodified parent peptides—and reductions in peak sizes at the retention times of the unmodified peptides would be seen on the TIC chromatogram. Figure 4.7 shows the TIC and UV absorbance at 214 and 280 nm for the unlabelled and pPD- and dPD-labelled samples. For almost every peak in the UV absorbance spectra, there is a peak in the TIC chromatogram that corresponds to the ionisation current of peptides eluted at that retention time. Reductions in peak sizes for 214-nm absorbance and TIC can be observed at 21.0 and 45.8 minutes. It should be noted that differences between the light-exposed and unexposed HSA samples were also observed at this latter time point (Figure 4.5), suggesting that the reduction in peak size seen at 45.8 minutes upon photolabelling may be due to the effects of exposure and not photolabelling. From the TIC chromatograms of Figure 4.7, it can also be observed that there are new peaks at 54.6 and 68.9 minutes in the spectra of both the labelled samples. These two peaks are likely to correspond to the elution of labelled peptides, as no differences were seen in the chromatograms between the two control samples at these time points in Figure 4.5. Differences between the unlabelled and labelled samples were seen in their absorbance chromatograms at other time points, but not in their TIC chromatograms. These are investigated further in Figure 4.8, which shows only the absorbance chromatograms. The lack of peak size differences in the TIC graphs simultaneous to differences

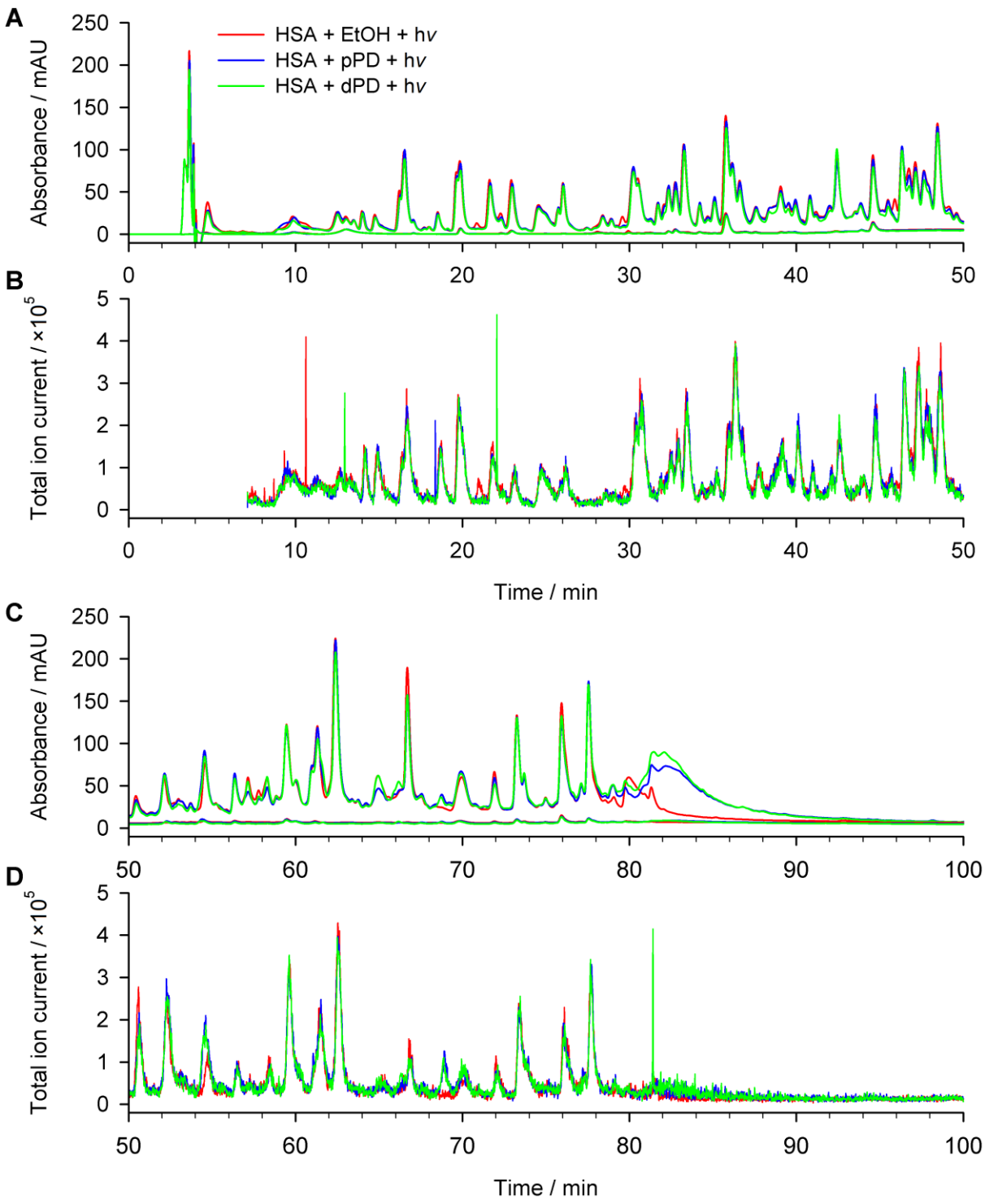


Figure 4.7: Chromatogram overlays of absorbance at 214 and 280 nm (A and C) and total ion current (B and D) for the separation of tryptic peptides of HSA from protein that was exposed to UV light in the presence of ethanol (i.e. in the absence of photolabel; red), pPD (blue) and dPD (green) prior to protein denaturation and trypsin digestion.

seen in the 214- and 280-nm absorbance could be due to more than one peptide eluting at a similar retention time and/or that the labelled peptide does not ionise well on the mass spectrometer.

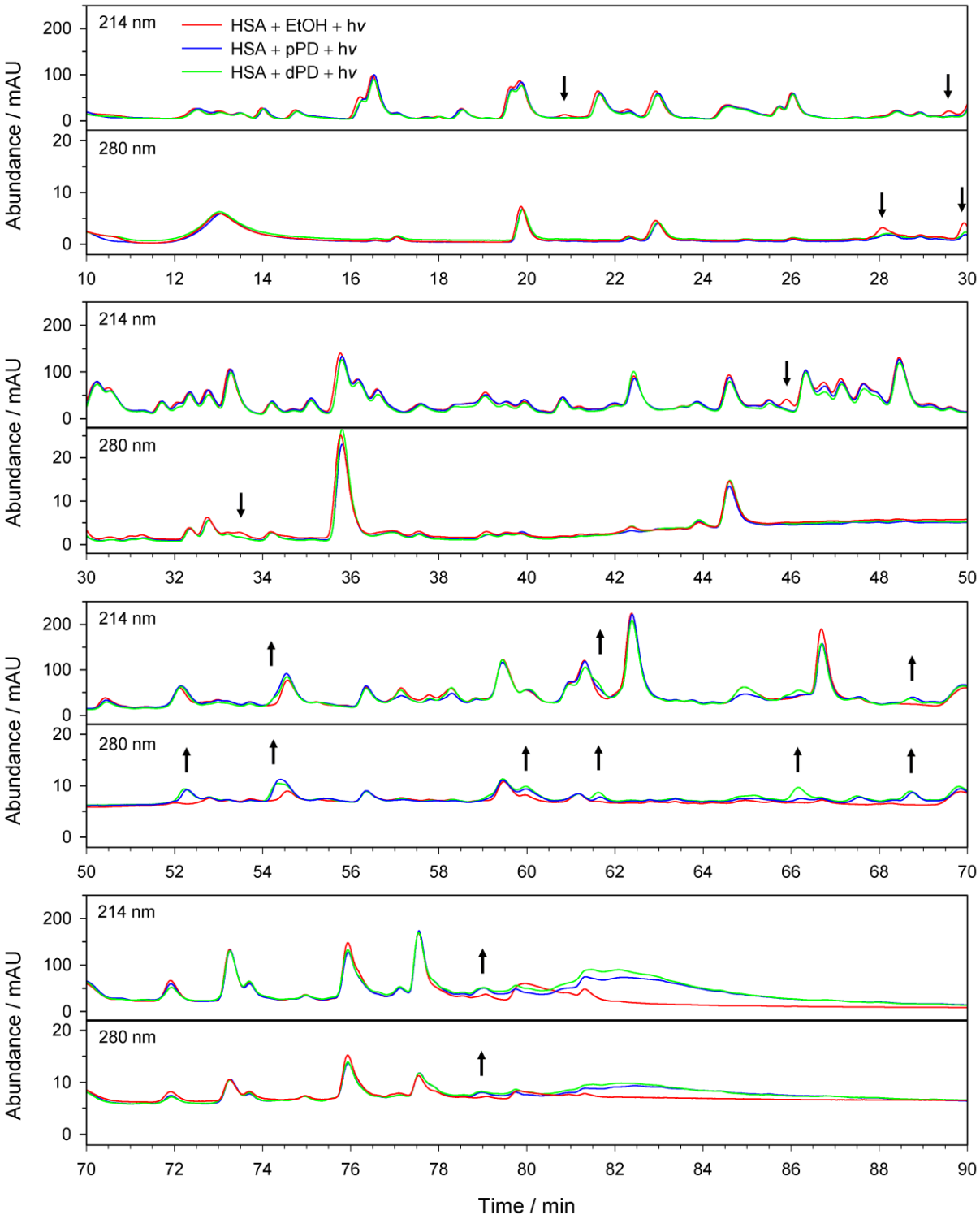


Figure 4.8: Absorbance spectra at 214 (top panels) and 280 nm (bottom panels) of tryptic peptides of HSA from protein that was exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green) prior to protein denaturation and trypsin digestion. Upward arrows indicate peaks that have appeared or increased in intensity, whilst downward arrows indicate peaks that have disappeared or reduced in intensity, when HSA is exposed in the presence of a photolabel.

Figure 4.8 shows that there are reductions in 214-nm absorbance of the labelled samples compared to the unlabelled sample at 21.0, 29.6 and 45.8 minutes, whilst increases in the peak sizes of the labelled samples compared to that unlabelled are seen at 52.3, 54.6, 60.1, 61.7, 67.7, 68.9, 79.1 and between 81 and 88 minutes. These new peaks and reductions in peak sizes are seen in both the pPD- and dPD-labelled samples simultaneously.

4.3.2 Determining the masses of labelled peptides

Analysing the averaged mass spectra spanning the durations of the new peaks for the pPD- and dPD-labelled samples together with those for unlabelled sample between the same time points allows the masses of labelled peptides to be determined. An initial analysis was carried out with the Deconvolution tool of *ChemStation*, which resolves peaks corresponding to multiply charged forms of the same peptide and outputs the MW of that peptide. When this approach did not result in identifying molecular weights that only occurred in the labelled but not the unlabelled samples, then the presence of new species on the mass spectra of the labelled samples was extensively searched by exporting the average mass spectra for all three samples to *SigmaPlot* to overlay them for closer inspection.

Apparent new peaks in the TIC spectra upon labelling were seen at 54.3–55.0 and 68.4–69.4 minutes (Figure 4.7). The averaged mass spectra over these two periods were obtained. Figure 4.9 identifies the labelled peptides with molecular weights of 1162.4 Da (when labelled with pPD) and primarily 1163.4 Da (when labelled with dPD) being eluted at 54.3–55.0 minutes along with an unlabelled peptide with a MW of 1910.3 Da. The singly, doubly and triply charged ions for the labelled peptide can be seen on the overlaid averaged mass spectra for both the photolabelled samples. The mass spectra insets in Figure 4.9 displaying the enlarged peaks illustrate the distinct differences in the isotopic distribution of the pPD- and dPD-labelled $(M + H)^+$ ions, as well as the right shift of the peak apices of the $(M + 2H)^{2+}$ and $(M + 3H)^{3+}$ ions for the dPD-labelled sample compared to those for the pPD-labelled sample. A right shift of 0.5 m/z is expected for a doubly charged dPD-labelled ion whilst a shift of 0.33 m/z is expected for a triply charged dPD-labelled ion. Shifts of 0.6 and 0.4 m/z are seen for the labelled doubly and triply charged ions respectively. These new peaks, which are only present in the mass spectra of the labelled samples, indicate that a peptide with a mass of 946.3 Da had been labelled. This is the most abundant singly charged

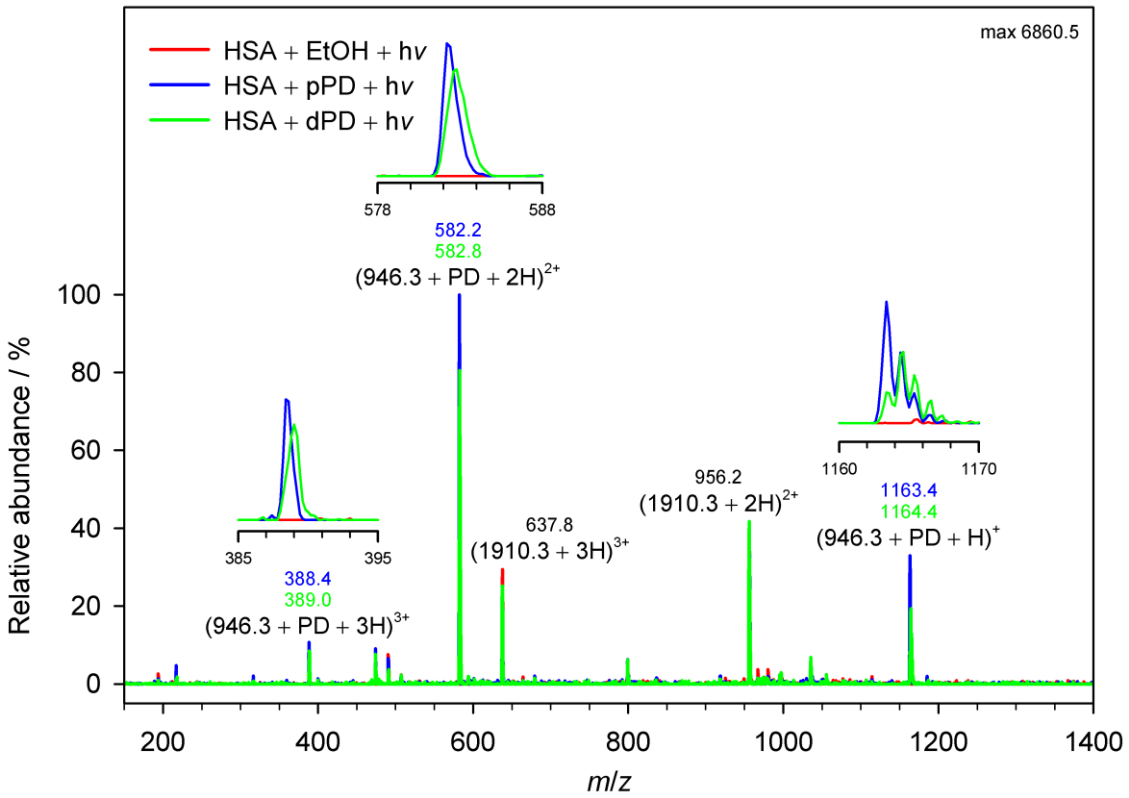


Figure 4.9: Averaged mass spectra overlay at 54.3–55.0 minutes of the HSA samples exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green). The ionisation current is normalised to the most ionising peak from the three traces within the *m/z* range displayed. Insets show enlarged spectra around the peaks for the ions of the parent peptide with a molecular weight of 946.3 Da, which has become labelled. The *m/z* of the peak apices are noted, as are the molecular weights and charged states of their corresponding ions.

species of the labelled peptide, 1163.4 *m/z* for the pPD sample and 1164.4 *m/z* for the dPD sample, less both the charge (1 Da) and the molecular weights of the photolabels attached (216.1 Da for a D_0 photolabel and 217.1 Da for a D_1 photolabel as this is the most abundant isotope in the dPD sample).

Obtaining the averaged mass spectra spanning the new TIC peak at 68.4–69.4 minutes and proceeding with the same workflow identified a peptide with a MW of 1127.5 Da being labelled by both the photolabel samples (Figure 4.10). Once again, the singly, doubly and triply charged ions of the labelled peptides are detected by the mass spectrometer and displayed the characteristic differences in isotopic distribution between the two photolabelled samples.

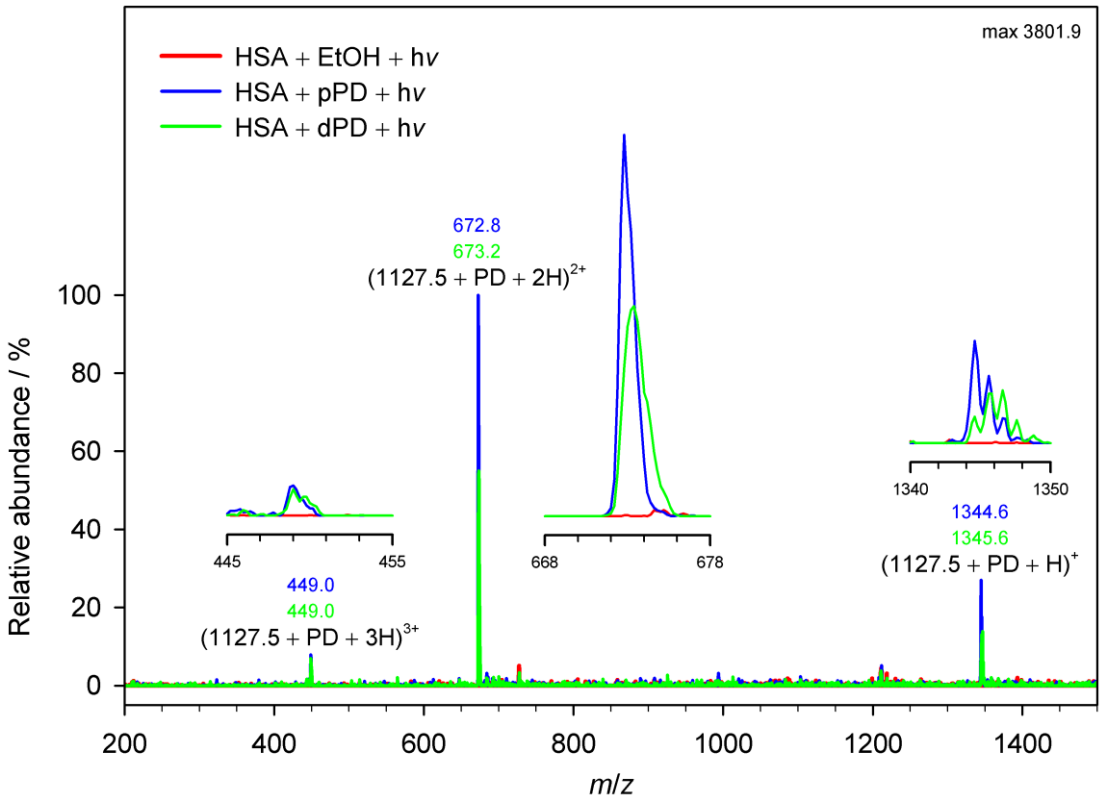


Figure 4.10: Averaged mass spectra overlay at 68.4–69.4 minutes of the HSA samples exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green). The ionisation current is normalised to the most ionising peak from the three traces within the *m/z* range displayed. Insets show enlarged spectra around the peaks for the ions of the parent peptide with a molecular weight of 1127.5 Da, which has become labelled. The *m/z* of the peak apices are noted, as are the molecular weights and charged states of their corresponding ions.

The analysis of these mass spectra has demonstrated the usefulness of labelling with pPD and dPD samples concurrently. The mass spectra of Figure 4.9 and Figure 4.10 show that the new (not present in the unlabelled sample) peak species corresponding to singly charged ions of pPD- and dPD-labelled peptides display different isotopic distributions. The maximum isotopic peak for the singly charged ion of the labelled peptides of the dPD-labelled sample is one unit greater than that for the pPD-labelled sample; the monoisotopic ion peak for the D_0-labelled peptide is more prominent in the pPD than in the dPD-labelled peptide; whilst the isotopic peaks that are two and three units greater than the monoisotopic ion peak are larger for the dPD-labelled sample. A similar distribution of isotopic peaks should theoretically occur for multiply charged ions, but with the separation between the peak apices being $1/z$, where z is the charge state of the ion. The isotopic distribution peak patterns can be calculated for both unlabelled and labelled peptides of known sequences. A few examples comparing the isotopic patterns of labelled peptides from

experimental data with their predicted isotopic distributions are presented in section 4.7 following a discussion on sequencing these labelled peptides in section 4.6.

Increases in UV absorbance at 214 or 280 nm are seen at elution times of 52.3, 60.1, 61.7, 67.7 and 79.1 minutes but increases in ionisation are not seen at these time points in the TIC chromatogram (Figure 4.7). Averaged mass spectra were obtained between −0.3 and +0.7 minutes of the apex of each of these new absorbance peaks. Out of all of these time ranges, only the analysis of the averaged mass spectra around 52.3 minutes resulted in the identification of the mass of a labelled peptide.

Despite no new peaks being seen in the TIC chromatogram at 52.3 minutes, the averaged mass spectra for the three samples were acquired. Figure 4.11 of the averaged mass spectra at 52.0–53.0 minutes shows that unlabelled peptides with molecular weights of 1012.7 Da and 1034.4 Da are eluted at this time range. The peaks at 507.2, 1013.7 and 1035.4 *m/z* for all three samples overlay

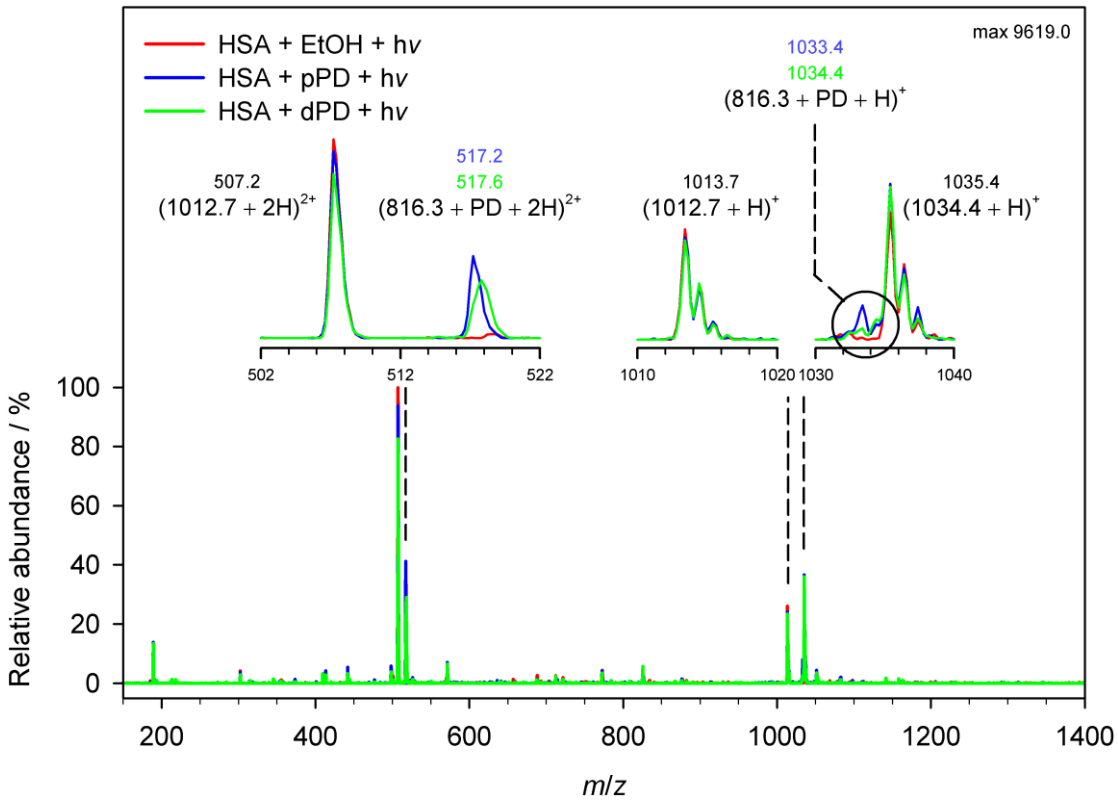


Figure 4.11: Averaged mass spectra overlay at 52.0–53.0 minutes of the HSA samples exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green). The ionisation current is normalised to the most ionising peak from the three traces within the *m/z* range displayed. Insets show enlarged spectra for *m/z* ranges of 502–522, 1010–1020 and 1030–1040. The *m/z* of the peak apices are noted, as are the molecular weights and charged states of their corresponding ions.

each other exactly, whilst there are peaks at 517.2 and 517.6 *m*/*z* in the mass spectra of the pPD- and dPD-labelled samples respectively that are absent in that of the unlabelled sample. New peaks at 1033.4 *m*/*z* at a higher intensity for the pPD-labelled sample and a lower intensity for the dPD-labelled sample, as well as vice versa for the peaks at 1034.4 *m*/*z*, can also be seen (circled in Figure 4.11). These peaks around 517, 1033 and 1034 *m*/*z* indicate that a peptide of 816.3 Da had been labelled.

To find more labelled peptides and their masses, a function to obtain averaged mass spectra in *ChemStation* was used over the total duration between seven and 140 minutes of the HPLC-MS run, as well as over epochs of between one and five minutes across the whole duration of the run. The averaged mass spectra of the unlabelled and pPD- and dPD-labelled samples for each period were overlaid. Any new peaks that occurred upon labelling were noted and their *m*/*z* values were then assumed to have arisen from ions that had up to four charges to predict the potential MW of their ionised peptides. When two or more new peaks on the same set of averaged mass spectra suggested the presence of a peptide, its MW was further investigated using the Extract Ions function in *ChemStation* to confirm its presence in only the labelled samples.

Through this approach, six more labelled peptides along with their masses were identified, bringing the total number of labelled peptides identified to nine. Figure 4.12 shows the mass spectra overlay over six epochs when labelled peptides were eluted from the chromatography column. Peaks corresponding to the singly and doubly charged ions of the labelled peptides are displayed, demonstrating the expected differences in isotopic distribution between the pPD- and dPD-labelled samples. The averaged mass spectra at 80.0–83.5 minutes (Figure 4.12F) gives weak evidence that labelled peptides with molecular weights of 606.0 *m*/*z* (when labelled with pPD) and 607.0 *m*/*z* (when labelled with dPD) are present based on the difference in isotopic distribution pattern of peaks between 606 and 610 *m*/*z* with ionisation abundance that is close to background noise. More evidence was provided by using the Extract Ions tool in *ChemStation* on the *m*/*z* of the labelled singly charged ions (607.0 and 608.0 *m*/*z*), which showed no ionisation current between 80 and 85 minutes for the unlabelled sample, but ionisation caused by ions of 607.0 and 608.0 *m*/*z* can be seen in the pPD- and dPD-labelled samples respectively (data not shown).

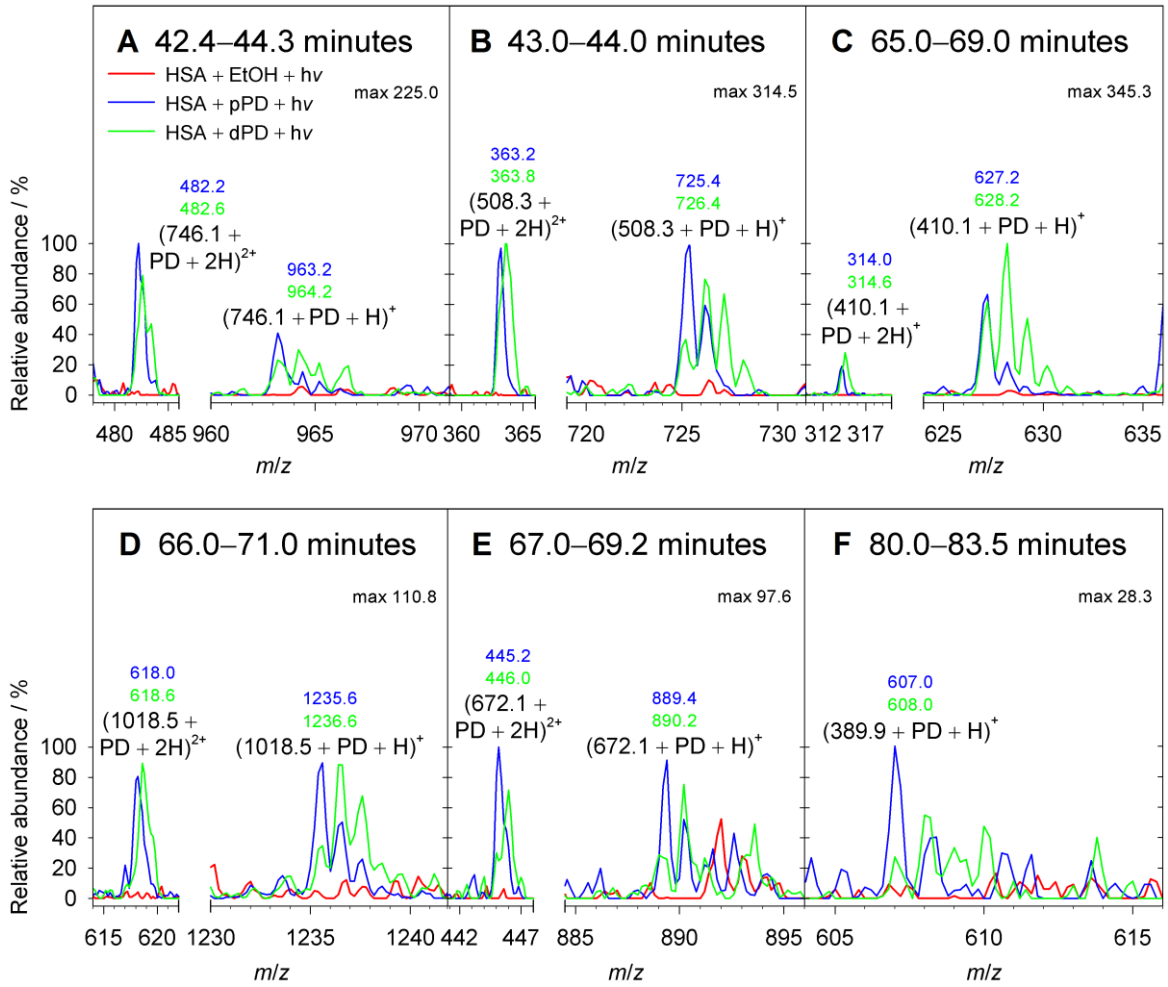


Figure 4.12: Averaged mass spectra overlays at various retention times when labelled HSA peptides were eluted for the protein samples exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green). The labelled peptides that were eluted between 42.4 and 44.3 minutes (A), 43.0 and 44.0 minutes (B), 65.0 and 69.0 minutes (C), 66.0 and 71.0 minutes (D), 67.0 and 69.2 minutes (E), and 80.0 and 83.5 minutes (F) have unlabelled parent peptides with molecular weights of 746.1, 508.3, 410.1, 1018.5, 672.1 and 389.9 Da respectively. The ionisation current is normalised to the most ionising peak from the three traces within the *m/z* ranges displayed in each of these time ranges. The *m/z* of the peak apices are noted, as are the molecular weights and charged states of their corresponding ions.

Evidence similar to the extract ion analysis for this labelled peptide in the form of selected ion monitoring scan is presented in Figure 4.13B in the next section.

4.4 Selected ion monitoring scans and fraction collections

Quadrupole analysers of mass spectrometers can be run in either scan mode or selected ion monitoring (SIM) mode. In SIM mode, only a few ions are monitored by the mass analyser rather than the continuous surveying of a range of *m/z* by scanning. SIM mode runs have the advantage

of being significantly more sensitive than scan mode because more time is spent focusing on only a few ions. SIM mode experiments were carried out on the unlabelled and pPD-labelled HSA samples to monitor groups consisting of the $(M + H)^{+}$, $(M + 2H)^{2+}$ and $(M + 3H)^{3+}$ ions of each of the nine labelled peptides for which the molecular weights of the labelled peptides had been identified (section 4.3.2). These HPLC-MS runs provide a more accurate indication of the elution times of the labelled peptides than performing extract ions analysis on the full scan mode data.

The overlays of the SIM ion current chromatograms for the unlabelled and pPD-labelled samples for five of the nine identified labelled peptides are shown in Figure 4.13. Figures A and B show pPD-labelled peptides with molecular weights of 626.2 and 606.2 Da respectively. These peptides have relatively low ionisation currents, compared to the ionisation caused by the 1234.6-Da labelled peptide represented by the SIM data in figure C. The ion current chromatogram for these three labelled peptides indicate that the fraction collection times of these peptides should be 66.1–

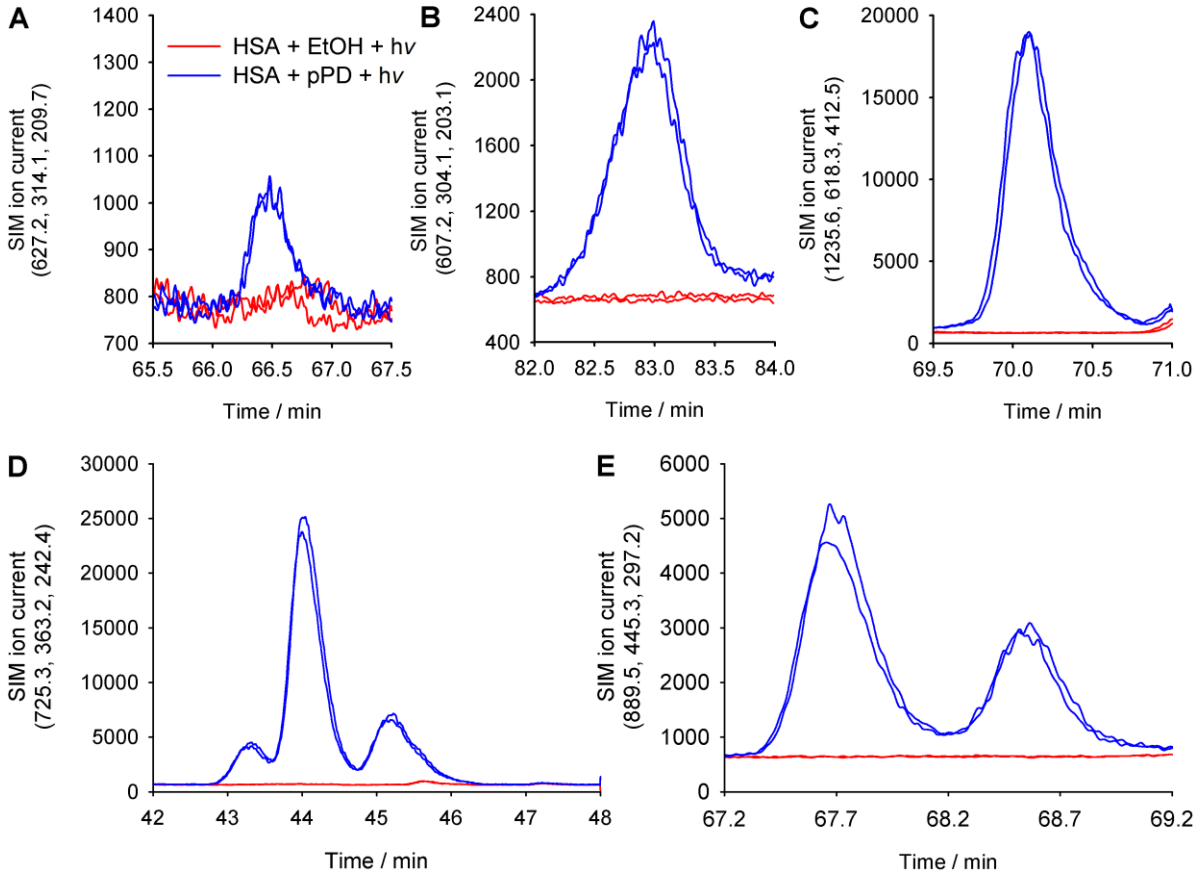


Figure 4.13: Overlays of the SIM scans ion current chromatograms for the unlabelled (red) and the pPD-labelled sample (blue) monitoring the combined ion current from the singly, doubly and triply charged ions of the labelled peptides with molecular weights of 626.2 (A), 606.2 (B), 1234.6 (C), 724.3 (D) and 888.5 Da (E).

67.0, 82.0–84.0 and 69.5–70.8 minutes. These retention ranges are more accurate than when they are determined by the Extract Ions function due to greater signal-to-noise ratio. The use of SIM scans is better demonstrated by Figures D and E. Figure 4.13D shows that the pPD-labelled peptides with MW of 724.3 Da elute off the column between 42.8–46.4 minutes, similar to the results from extracting the singly and doubly charged ions on a full scan mode TIC chromatogram (data not shown). However, with the greater sensitivity of the SIM mode data, it is apparent that not all the 724.3-Da pPD-labelled peptides have the same polarity—a mixture of three different labelled peptides is present—resulting in their elution at slightly different ACN concentration and thus three peaks appear in the labelled sample SIM chromatogram that are absent in the unlabelled sample chromatogram. This could be due to more than one 508.2-Da HSA peptide being photolabelled and coincidently eluting at similar times, or that one peptide with MW of 508.2 Da being labelled along different positions along its sequence. Figure 4.13E also shows the elution of labelled peptides with the same MW—888.5 Da—at more than one retention time range. These two peaks that are only present in the chromatogram of the labelled sample correspond to either two different peptides with MW of 672.4 Da being labelled or that the photolabel can bind to one of two amino acid positions of the same peptide.

The labelled peptides were collected into separate fractions using an automated fraction collector module by pooling a handful of liquid chromatography runs of each of the two photofol-labelled HSA samples. These fractions were then concentrated by evaporating ACN under a stream of nitrogen gas and adding 0.2 % FA and 0.02 % TFA to acidify these concentrated peptide samples.

4.5 Predicting labelled peptide sequences based on molecular weight

Upon identifying the masses of the labelled peptides as described in section 4.3.2, the sequences of the labelled peptides can be predicted using *in silico* methods. The molecular weights of the parent HSA peptides that were labelled can be calculated by subtracting the mass of the photolabel (216.1 Da for D_0-labelled and 217.1 Da for D_1-labelled peptides, which are the most abundant isotopes in the pPD- and dPD-labelled samples respectively) from the labelled peptide mass.

Through searching the masses of the expected tryptic peptides of HSA for the calculated molecular weights of the native peptides that became labelled, only four of the nine identified

labelled peptides can be assigned potential sequences. These are $_{213}$AWAVAR$_{218}$ (672.3 Da), $_{223}$FPK$_{225}$ (390.2 Da), $_{411}$YTK$_{413}$ (410.2 Da) and $_{535}$HKPK$_{538}$ (508.3 Da). The Peptide Tools function of *ChemStation* allows the searching of any sequence within HSA for peptide fragments of a particular mass. Sieving through the shortlisted peptide candidates for peptides that begin after, as well as end with, lysine or arginine residues identified six tryptic peptides each with a single missed cleavage matching the required native peptide masses of the labelled peptides. These peptides, along with the aforementioned four tryptic peptides containing no missed cleavages, are summarised in Figure 4.14.

Retention time (minutes)	Molecular weight (Da)		Sequence
	Labelled	Unlabelled	
42.9–46.0	724.3	508.2	$_{535}$HKPK$_{538}$
43.0–44.0	962.2	746.1	$_{191}$ASSAK↓QR$_{197}$
52.1–53.4	1032.4	816.3	$_{539}$ATK↓EQLK$_{545}$
54.3–55.3	1162.4	946.3	$_{198}$LK↓CASLQK$_{205}$
66.1–66.9	626.2	410.2	$_{411}$YTK$_{413}$
67.4–69.0	888.4	672.3	$_{213}$AWAVAR$_{218}$ $_{520}$ER↓QIK$_{524}$
68.6–69.6	1343.6	1127.5	$_{525}$K↓QTALVELVK$_{534}$
69.6–70.6	1234.6	1018.5	$_{210}$AFK↓AWAVAR$_{218}$
82.2–83.8	606.2	390.2	$_{223}$FPK$_{225}$

Figure 4.14: Summary of the *in silico* sequence predictions of labelled HSA peptides with their corresponding retention times and their molecular weights with and without a 216.1-Da photofol photolabel attached. Arrows denote missed trypsin-cleavage sites. Note that both $_{213}$AWAVAR$_{218}$ and $_{520}$ERQIK$_{524}$ are 672.1 Da.

For the labelled peptide with a monoisotopic mass of 888.4 Da, the unlabelled peptide could be $_{213}$AWAVAR$_{218}$ or $_{520}$ERQIK$_{524}$ assuming a missed cleavage after the arginine in the latter sequence. There is also a possibility that both of these 672.3-Da peptides were labelled. That would complement with the SIM scan data of the 888.5-Da labelled peptide (Figure 4.13E) that showed two elution times for labelled peptides of this MW.

The tryptic peptides $_{191}$ASSAKQR$_{197}$, $_{198}$LKCASLQK$_{205}$, $_{210}$AFKAWAVAR$_{218}$ and $_{539}$ATKEQLK$_{545}$ were not found in the unlabelled or labelled samples by extracting their $(M + H)^+$ and $(M + 2H)^{2+}$

ions. This is not surprising as upon incubating the denatured HSA sample with trypsin overnight, trypsin digestion should have reached completion. These missed-cleavage peptides are only present in the labelled sample when a photolabel is bound, adding 216.1 Da or 217.1 Da of mass to them depending on which photolabel isotope has been attached. This suggests that the presence of an attached photolabel might have blocked the trypsin-cleavage site.

Only one retention time, at 35.9 minutes, was identified by extracting ions of 673.3 and 337.2 *m*/*z* for the peptide with a MW of 672.3 Da in the unlabelled sample. This is likely to correspond to the elution of the peptide $_{213}$AWAVAR$_{218}$ rather than $_{520}$ERQIK$_{524}$, which contains a missed cleavage site. However, extracting the singly and doubly charged ions for the peptide with a MW of 1127.5 Da in the unlabelled sample suggests that such a peptide exists and elutes from the column at 47.9 minutes. The missed-cleavage peptide $_{525}$KQTALVELVK$_{534}$ is the only tryptic peptide in HSA that has a MW of 1127.5 Da. The presence of this missed-cleavage peptide can potentially be explained by the occurrence of another trypsin-cleavage site just preceding this peptide. Amino acid position 524 of HSA is also a lysine residue and the two adjacent cleavage sites might hinder trypsin-binding and thus reduce the rate of hydrolysis at the C-terminal of lysine-525 and yield $_{525}$KQTALVELVK$_{534}$.

4.6 Sequencing labelled peptides by tandem mass spectrometry

The isolated labelled peptides obtained via the automated fraction collector of the HPLC module, described in section 4.4, were concentrated and acidified by 0.2 % FA and 0.02 % TFA prior to injection into the HPLC-coupled triple quadrupole mass spectrometer.

MS-MS is a powerful technique for gaining structural information on analytes. The mass spectra shown so far in this chapter were obtained by a single-stage mass spectrometer with ESI. This relatively ‘soft’ ionisation is ideal for analysing peptide mixtures as it leaves peptides intact. However, to acquire more structural information, the ions of the peptides are fragmented by colliding with neutral molecules in a process known as collision-induced dissociation (CID) in order to identify their constituent amino acids. In a triple quadrupole mass spectrometer setup, the first of the three mass analysers was configured to allow passage of the doubly charged ions of the labelled peptides, known as parent ions, into the second analyser, which acts as a CID chamber.

This is where ions are fragmented by argon mainly along the peptide backbone, particularly across the peptide bond, producing characteristic b_n and y_n ions, as well as a_n, c_n, x_n and z_n ions (Figure 1.8). The third mass analyser then generates the mass spectrum of the resulting product ions.

4.6.1 Labelled peptide LKCASLQK

At the retention time of 54.6 minutes, there was clear evidence of increases in TIC (Figure 4.7) and absorbance at 214 and 280 nm (Figure 4.8) in the labelled samples compared to the unlabelled sample. The peptide $_{198}$LKCASLQK$_{205}$ was predicted to be labelled through analysis of the averaged mass spectra spanning 54.3–55.0 minutes (Figure 4.9) coupled with the analysis of *in silico* digestion of HSA.

Due to its doubly charged ions at 582.3 and 582.8 *m/z* for the pPD- and dPD-labelled samples respectively being the most abundant charged ions, these were selected as the parent ions for analysis by MS-MS. The EPI mass spectra showing the daughter ions from these pPD- and dPD-labelled peptides are shown in Figure 4.15.

The characteristic CID ions for the predicted peptide $_{198}$LKCASLQK$_{205}$ (section 4.5) are shown in Figure 4.16. Most of the singly charged b_n and y_n ions for this peptide were identifiable on the mass spectra (indicated by black labels in Figure 4.15 and highlighted in bold in Figure 4.16), confirming that the peptide $_{198}$LKCASLQK$_{205}$ had been labelled.

It was hypothesised that the photolabel had attached to an amino acid near the missed trypsin-cleavage site at the carboxyl end of lysine-199 (section 4.5). The EPI spectra confirmed that the photolabel had bound to lysine-199. Peptide fragments with the photolabel bound are identifiable on the spectra (indicated by blue and green labels for the D_0- and D_1-labelled fragments respectively in Figure 4.15). There is a 1 ± 0.1 *m/z* unit difference between the most abundant isotopes of the singly charged labelled fragments of the two spectra. For example, the peaks for the label-bound fragment consisting of the first seven amino acids cleaved at the peptide bond between the seventh and last amino acids, b_7^, are at 1017.4 and 1018.5 *m/z* for the two samples. For a doubly charged labelled fragment, for example, y_8^{2+}^, there is only a 0.5 *m/z* difference. The presence of peaks corresponding to labelled b_2 (b_2^) and labelled y_7 (y_7^) fragments on both spectra suggests that the label must be bound to the first lysine at amino acid position two of this

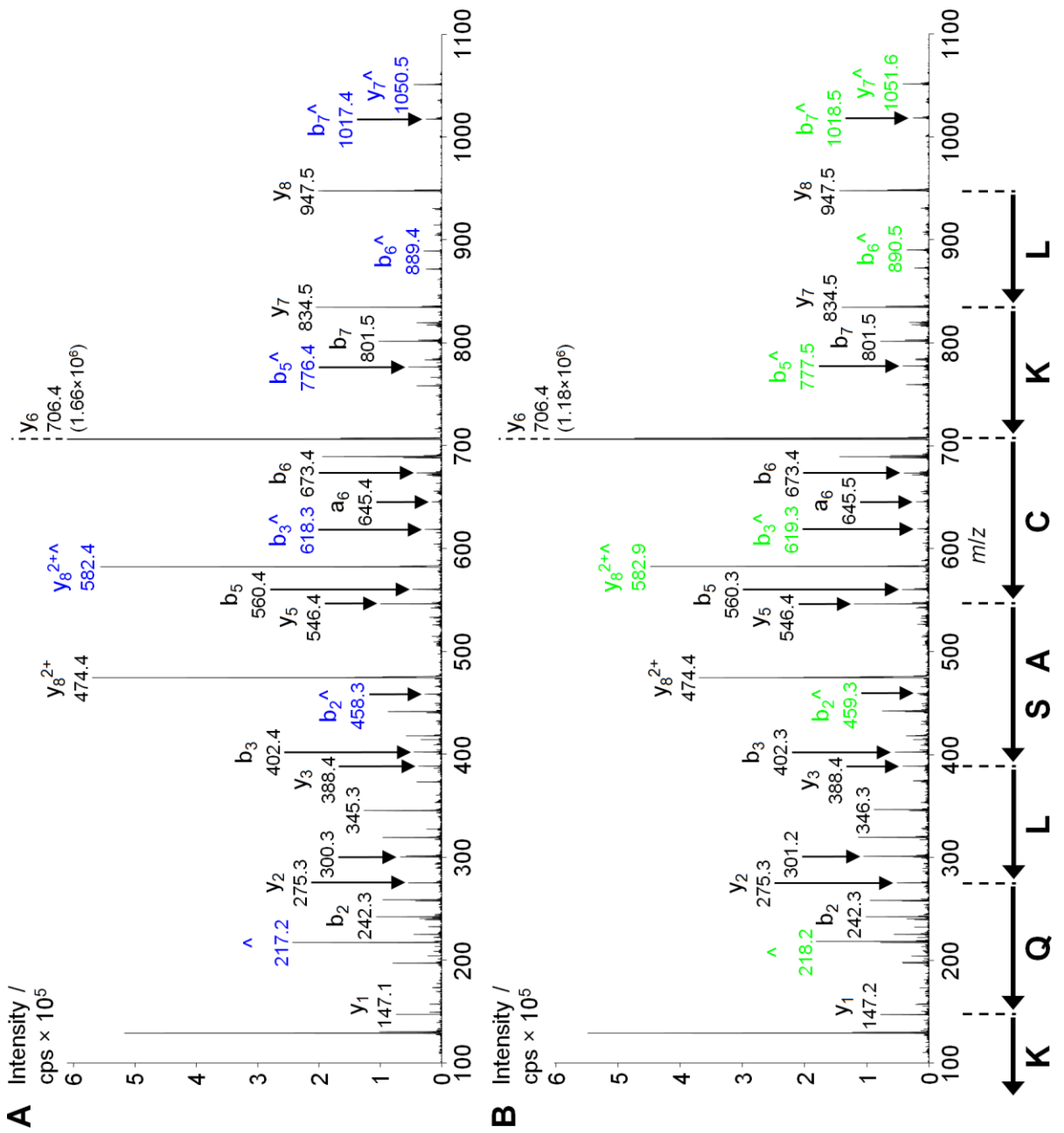


Figure 4.15: Enhanced product ion mass spectra for the fragmentation of the parent ions 582.3 and 582.8 *m/z* for the HSA samples exposed to light in the presence of pPD (A) and dPD (B) respectively. The labelled peptide is identified as LKCASLQK based on the presence of many b_n and y_n ions (*m/z* values and CID ion notations in black) that are representative of this peptide. Peaks that are 216.1 (pPD-labelled) or 217.1 (dPD-labelled) *m/z* units greater than the b_n and y_n ions are marked with an additional caret (*m/z* values and CID ion notations in blue (pPD-labelled) and green (dPD-labelled). The presence of peaks corresponding to a labelled-b_2 ion, b_2^ and a labelled-y_7 ion, y_7^, indicates that photofol is bound to the first lysine of this peptide. The singly charged photolabel products with 217.2 and 218.2 *m/z* detached from the pPD- and dPD-labelled peptides respectively were also marked by a caret. Peaks which extend beyond the *y*-axis scale were marked with a dashed line at the top of the spectra and their intensities in cps were noted in brackets.

a_n ions (m/z) →	86.1	214.2	374.2	445.3	532.3	**645.4**	773.4	901.5
b_n ions (m/z) →	114.1	**242.2**	**402.2**	473.3	**560.3**	**673.4**	**801.5**	929.5
c_n ions (m/z) →	131.1	259.2	419.2	490.3	577.3	690.4	818.5	946.6

n =	1	2	3	4	5	6	7	8	
aa(198–205)	**L**	**K**	**C**	**A**	**S**	**L**	**Q**	**K**	
	8	7	6	5	4	3	2	1	= n
	973.5	860.4	732.3	572.3	501.3	414.2	301.2	173.1	← x_n ions (m/z)
	947.5	**834.5**	**706.4**	**546.3**	475.3	**388.3**	**275.2**	**147.1**	← y_n ions (m/z)
	930.5	817.4	689.3	529.3	458.3	371.2	258.2	130.1	← z_n ions (m/z)

Figure 4.16: The expected m/z of the singly charged CID ions of the peptide LKCASLQK. The m/z that are observed in the EPI mass spectra (Figure 4.16) are in bold. The amino acid cysteine, which was alkylated with iodoacetamide (C), has an additional mass of 57 Da compared to an unmodified cysteine residue.

peptide. Figure 4.17 illustrates the labelled fragments that have been observed and marks the position where the photolabel is bound. The peaks at 217.2 and 218.2 m/z in Figure 4.15 correspond to the photolabels with zero or one deuterium substitutions being detached from the

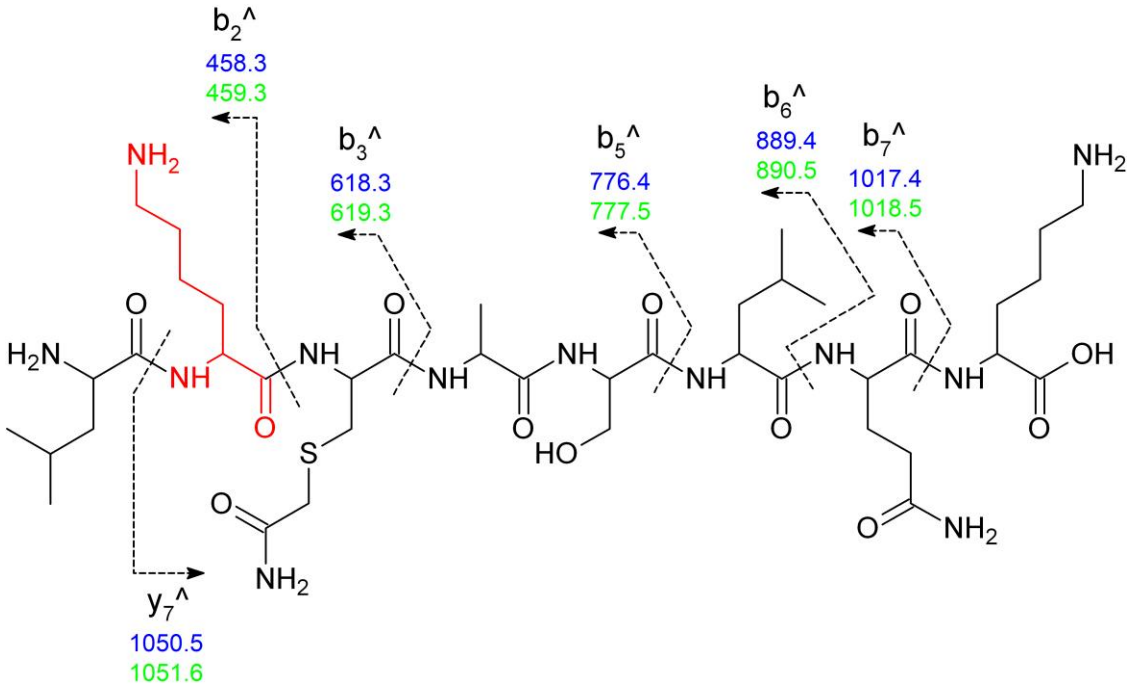


Figure 4.17: Schematic of the peptide $_{198}$LKCASLQK$_{205}$ with annotations showing the labelled CID fragments identified in the mass spectra of Figure 4.15. The observed m/z for pPD-labelled (blue) and dPD-labelled (green) fragments are listed below their CID notations. A caret denotes the presence of a label bound to the fragment. Dashed lines intersecting with the peptide indicate the bonds that become cleaved, whilst arrows point towards the amino acids that are included in the peptide fragment. The potential binding site of the photolabel is deduced to be lysine-199, shown in red.

labelled peptide. The ionisation intensity of these two peaks increases for spectra obtained from MS-MS runs with a higher collision energy setting. Concurrent to the increases in the collision energy setting and the peak intensities of the detached photolabels, the peaks for the parent ions (582.4 and 582.9 *m/z*) and for peptide fragments that have a photolabel bound reduce in ionisation intensity. It is critical to set the collision energy parameter not too high so that the photolabel remains on the CID fragments of the labelled peptide in order to allow the identification of the labelled amino acid. However, too low a collision energy setting results in insufficient fragmentation of the parent ion, leading to very low relative intensity peaks when normalised to the maximum-intensity peak of the parent ion, making sequencing difficult.

4.6.2 Lysine-labelled peptides

The previous subsection presented a good example of a lysine-labelled peptide, $_{198}$LKCASLQK$_{205}$. The EPI mass spectra showed clear peaks for many characteristic b_n- and y_n-fragment ions to confirm the peptide sequence, and the presence of labelled-fragment ions from the b_n- and y_n-series, particularly b_2^ and y_7^, indicated that the location of the bound photolabel was on the lysine prior to the missed trypsin-cleavage site.

The mass spectra for the pPD- and dPD-labelled $_{198}$LKCASLQK$_{205}$ also showed four significant peaks at 300.3 and 345.3, and 301.2 and 346.3 *m/z* respectively for the two samples (Figure 4.15). These peaks correspond to the immonium-less-ammonia (K_{im}*) and amino acylium ions of lysine residues (K_{aa}) with a photolabel attached (Figure 4.18). These K_{im}* and K_{aa} ions of 84 and 130 *m/z* respectively are common internal lysine ions that occur in MS-MS (Falick *et al.*, 1993; Papayannopoulos, 1995). K_{im}* ions occur from c_n- and x_n-type cleavages at the bonds around the R group of lysine, whilst K_{aa} ions occur from b_n- and y_n-type cleavages at the peptide bonds of the lysine.

In section 4.5, four other labelled peptides were predicted to have a missed tryptic cleavage site at the C-terminal to a lysine residue. The sequences of these peptides—$_{191}$ASSAKQR$_{197}$, $_{210}$AFKAWAVAR$_{218}$, $_{525}$KQTALVELVK$_{534}$ and $_{539}$ATKEQLK$_{545}$—were all confirmed by MS-MS. The presence of some labelled-fragment ions of the b_n and y_n series confirms the location of the bound-photolabel to a single amino acid, the missed trypsin-cleavage lysine, or to a sequence of

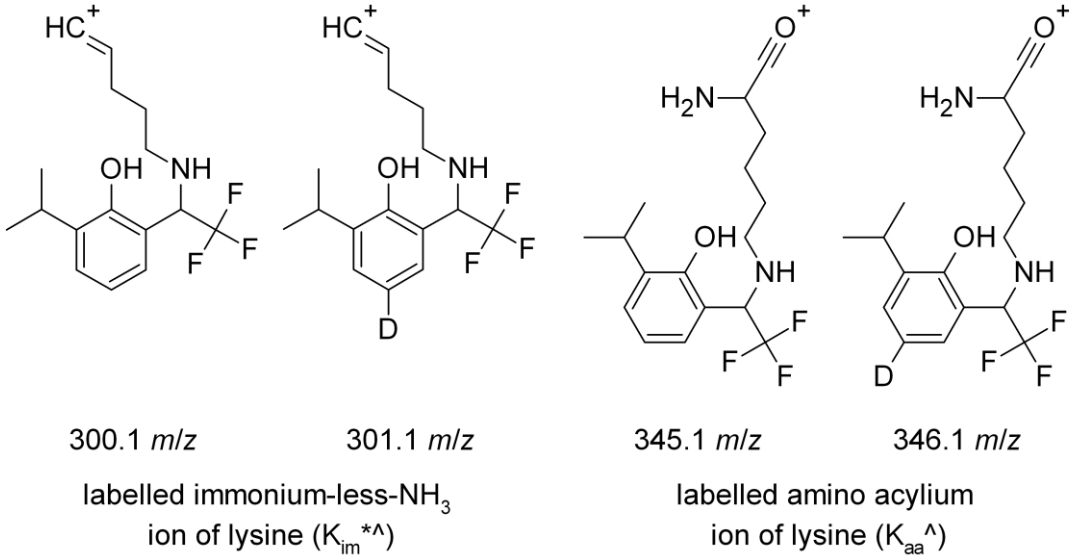


Figure 4.18: Chemical structures of labelled-bound internal ions of lysine. 300.1, 345.1 *m/z* (for the D_0-labelled from the pPD-labelled sample) and 301.1 and 346.1 *m/z* (for the D_1-labelled from the dPD-labelled sample) are the calculated *m/z* for the label-bound immonium-less-ammonia (K_{im}*^) and amino acylium (K_{aa}) ions of lysine. D denotes a deuterium atom which has replaced a protium atom on the benzene ring of photofol.

up to two amino acids, of which one is the missed-cleavage lysine. The characteristic pPD- and dPD-labelled K_{im}* and K_{aa} ions were also observed on these spectra, providing further evidence that a lysine residue had been labelled in these peptides. The annotated EPI mass spectra of pPD- and dPD-labelled $_{539}$ATKEQLK$_{545}$ are shown in Figure 4.19 and those for the other three lysine-labelled peptides are shown in Figures A.3–A.5 in the appendix.

The EPI spectra for the labelled-$_{539}$ATKEQLK$_{545}$ in Figure 4.19 provides a good demonstration of the usefulness of labelling with the pair of photolabel samples, pPD and dPD, in parallel. The expected *m/z* for the singly charged y_4 ions, labelled-y_7^{2+} ions and the potentially labelled-b_3 ions are isobaric, sharing an *m/z* of 517.3, whilst y_2 ions and the potentially labelled a_1 ions form an isobaric pair with an *m/z* of 260.3. The presence of a maximum peak at 517.9 *m/z* as well as a peak with 30 % maximal ionisation intensity at 518.3 *m/z* in the dPD-labelled sample of the EPI mass spectra, compared to a maximum peak at 517.4 *m/z* with no adjacent peaks at 517.9 or 518.3 *m/z* in that of the pPD-labelled sample, indicates the presence of both the y_7^{2+}^ and b_3^ ions from the fragmentation of the pPD- and dPD-labelled $_{539}$ATKEQLK$_{545}$. The deduction of the presence of b_3^ ions contributes to the conclusion that the internal lysine of this peptide had been labelled.

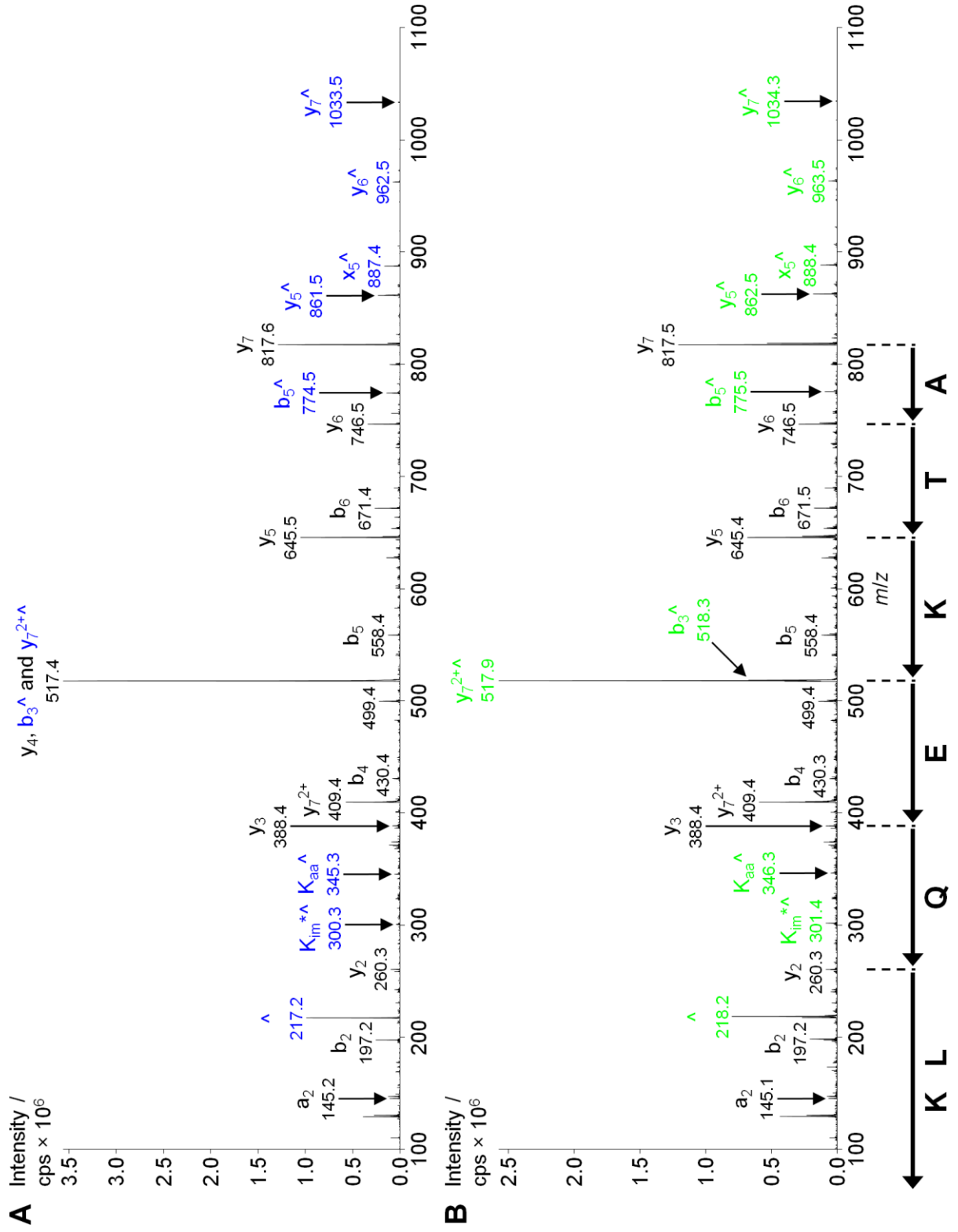


Figure 4.19: Enhanced product ion spectra for the fragmentation of the of the parent ions 517.4 and 517.9 *m/z* for the pPD- (A) and dPD-labelled (B) samples respectively showing that the label is attached to the third amino acid of the peptide $_{539}$ATKEQLK$_{545}$ based on the presence of b_3^, x_5^ and y_5^ ions. Carets indicate fragments containing the photolabel. K_{im}* and K_{aa} are internal lysine CID ions of 84 and 129 *m/z* respectively that were acknowledged by Falick *et al.*, 1993 and Papayannopoulos, 1995. The presence of these internal lysine ions being labelled further confirms the labelling of lysine-541.

Conversely, both EPI spectra exhibit a peak at 260.3 *m/z* and lack a peak at 261.3 *m/z*, indicating that the peaks at 260.3 *m/z* correspond to y_2 and not a_1^ ions.

4.6.3 Labelled peptide YTK

The short labelled peptide with a MW of 626.2 Da, which was eluted at the retention time of 66.5 minutes, was predicted to be the peptide $_{411}$YTK$_{413}$ (section 4.5). Figure 4.20 lists the calculated expected *m/z* of its singly charged ion fragments and those that are observed on the EPI mass spectra of the fragmentation of the predominantly doubly charged ions of 314.2 and 314.7 *m/z* for the pPD- and dPD-labelled peptide respectively (Figure 4.24) are highlighted in bold. The detection of the full y_n ion series confirmed $_{411}$YTK$_{413}$ as being the 410.2-Da peptide labelled.

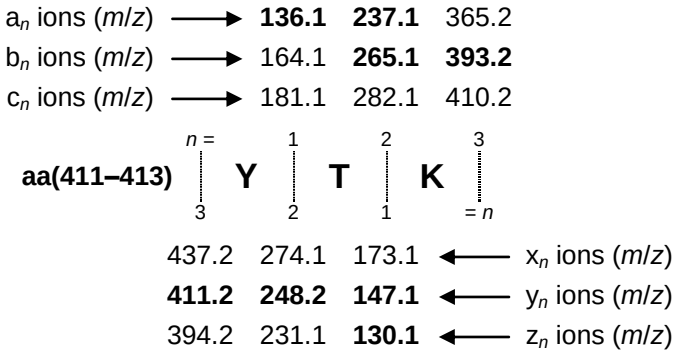


Figure 4.20: The expected *m/z* of the singly charged CID ions of the peptide YTK. The *m/z* that are observed in the EPI mass spectra (Figure 4.21) are in bold.

The EPI spectra of Figure 4.21 show that the photolabels have attached to the N-terminal tyrosine of $_{411}$YTK$_{413}$ based on the presence of a_1^ peaks at 352.2 and 353.2 *m/z* for the pPD- and dPD-labelled peptide respectively. These peaks differ by one unit between the two samples due to the photolabel isotopes with one deuterium substitution being the most abundant in the dPD-photolabel sample. The prominent peaks at 585.3 *m/z* for a pPD-labelled fragment and 586.3 *m/z* for a dPD-labelled fragment do not correspond to any common CID ion fragments of $_{411}$YTK$_{413}$ with the addition of the mass of a photolabel (216.1 or 217.1 Da). These peaks are likely to be y_3 ions, that is to say the non-fragmented singly charged peptide, with a partial photolabel molecule—one that lacks its isopropyl group—attached to the peptide.

The labelled fragments of the labelled $_{411}$YTK$_{413}$ peptide that were observed in the EPI mass spectra are summarised in Figure 4.22.

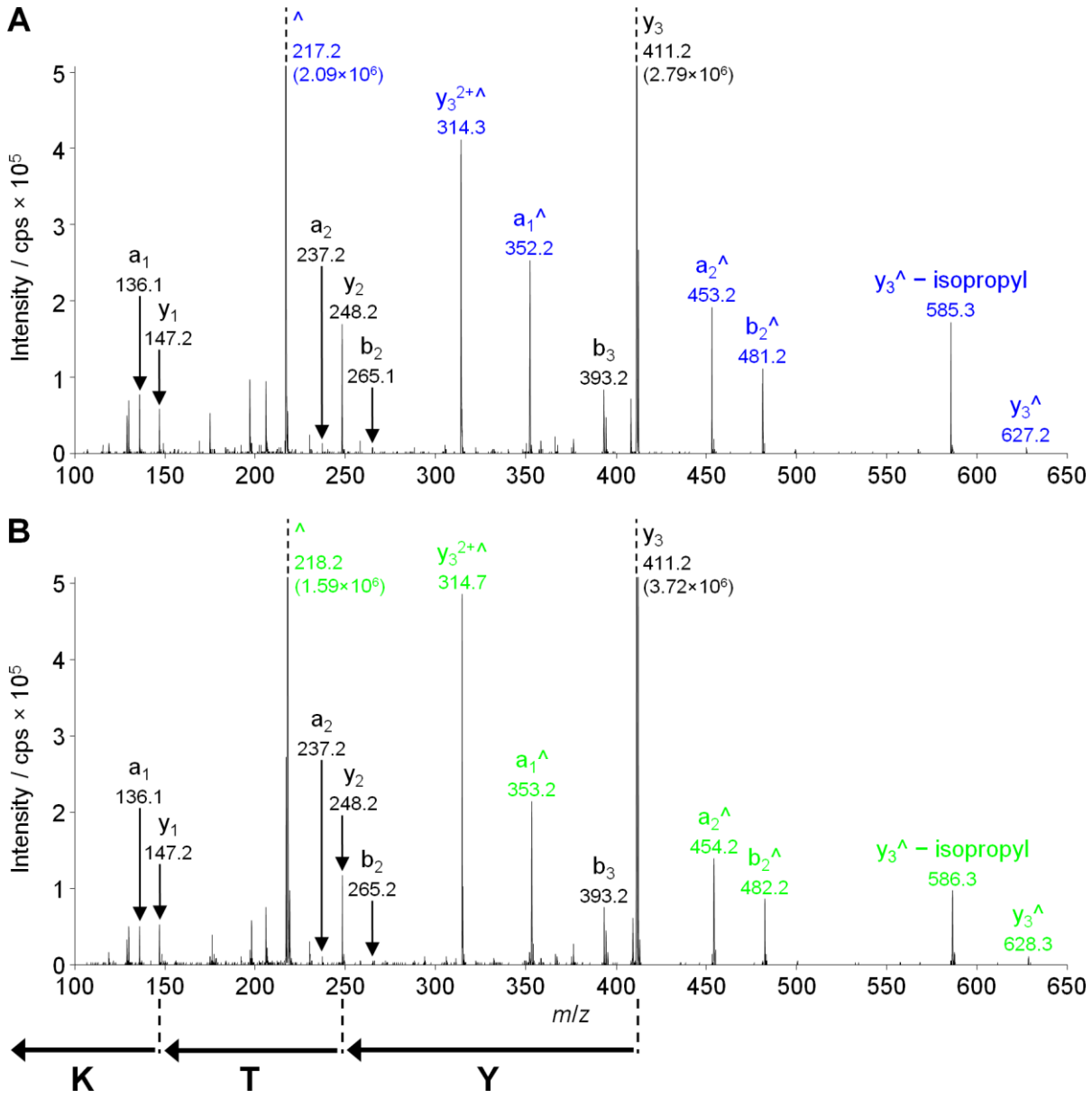


Figure 4.21: Enhanced product ion mass spectra for the fragmentation of the parent ions 314.2 and 314.7 *m/z* for the pPD- (A) and dPD-labelled (B) samples respectively. The labelled peptide is confirmed to be $_{411}$YTK$_{413}$ based on the full y_n-ion series for this peptide being observed (*m/z* values and CID ion notations in black). Peaks that correspond to labelled fragment ions are marked with an additional caret (*m/z* values and CID ion notations in blue (pPD-labelled) and green (dPD-labelled)). The presence of peaks corresponding to a labelled-a_1 ion, a_1^, indicates that photofol is bound to tyrosine-411. Standalone carets at 217.2 and 218.2 *m/z* mark the photolabels that have been cleaved from the labelled peptides. Peaks which extend beyond the *y*-axis scale were marked with a dashed line at the top of the spectra and their intensities in cps were noted in brackets.

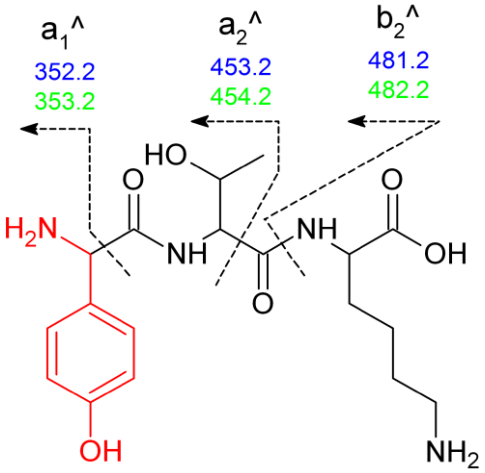


Figure 4.22: Schematic of the peptide $_{411}$YTK$_{413}$ with annotations showing the labelled CID fragments identified in the mass spectra of Figure 4.21. The observed *m/z* for pPD-labelled (blue) and dPD-labelled (green) fragments are listed below their CID notations. Caret symbol denotes the presence of a label bound to the fragment. Dashed lines intersecting with the peptide marks the bond being cleaved, whilst arrows point towards the amino acids that are included in the peptide fragment. The photolabel is bound to the a_1 fragment, which consists of tyrosine-411 and is shown in red.

4.6.4 Labelled peptide FPK

As with sequencing and identifying the amino acid labelled in the small peptide $_{411}$YTK$_{413}$, the confirmation that the predicted labelled peptide $_{223}$FPK$_{225}$ (section 4.5) from the single-stage mass spectrometer data was correct and the identification of the location of the labelled amino acid was straightforward. Figure 4.23 lists the expected CID singly charged ions of peptide FPK and those that are present on the EPI mass spectra of the fragmentation of the labelled doubly charged *m/z* of 304.2 and 304.7 (Figure 4.24) are in bold.

a_n ions (*m/z*) ⟶	**120.1**	**217.1**	345.2	
b_n ions (*m/z*) ⟶	148.1	**245.1**	**373.2**	
c_n ions (*m/z*) ⟶	165.1	262.2	390.3	
	n = 3	1 2	2 1	3 = *n*
aa(223–225)	**F**	**P**	**K**	
	417.2	270.1	173.1	⟵ x_n ions (*m/z*)
	391.2	**244.2**	**147.1**	⟵ y_n ions (*m/z*)
	374.2	227.1	130.1	⟵ z_n ions (*m/z*)

Figure 4.23: The expected *m/z* of the singly charged CID ions of the peptide FPK. The *m/z* that are observed in the EPI mass spectra (Figure 4.24) are in bold.

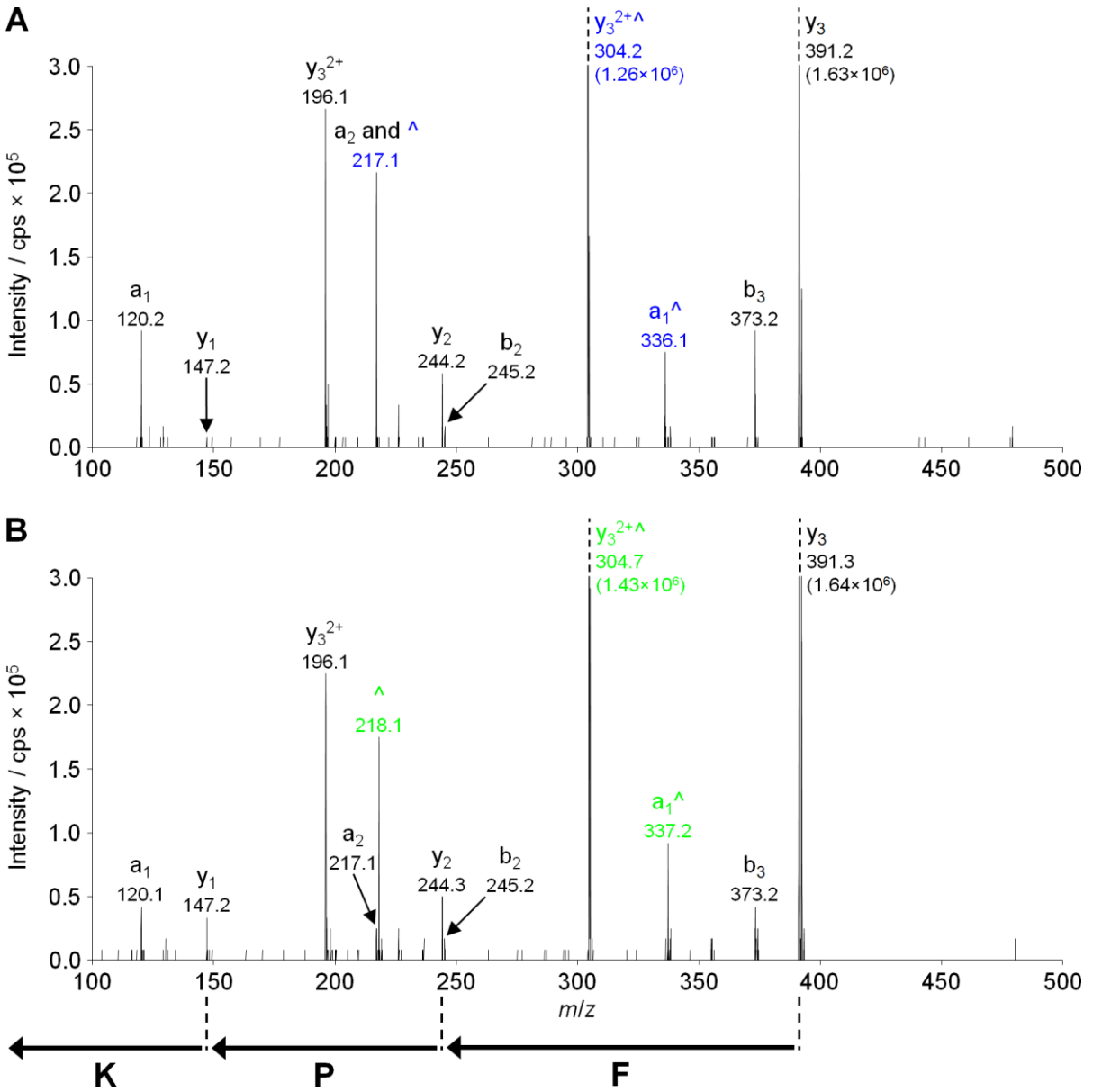


Figure 4.24: Enhanced product ion mass spectra for the fragmentation of the parent ions 304.2 and 304.7 *m/z* for the pPD- (A) and dPD-labelled (B) samples respectively. The labelled peptide is confirmed to be $_{223}$FPK$_{225}$ based on the full y_n-ion series for this peptide being observed (*m/z* values and CID ion notations in black). Peaks which correspond to labelled (^) fragment ions are marked in blue (pPD-labelled) and green (dPD-labelled). The presence of peaks corresponding to a labelled-a_1 ion, a_1^ indicates that photofol is bound to phenylalanine-223. Standalone carets at 217.2 and 218.2 *m/z* mark the photolabels that have been cleaved from the labelled peptides. Peaks which extend beyond the *y*-axis scale were marked with a dashed line at the top of the spectra and their intensities in cps were noted in brackets.

The EPI mass spectra in Figure 4.24 show that only two common CID fragments, a_1 and y_3, have been observed with a photolabel attached to them, a_1^ and y_3^{2+}^. The latter of these ions did not result from fragmentation by collision and is simply the doubly charged parent ion. The presence

of a_1^ is sufficient to deduce that the label must be bound to phenylalanine-223 (Figure 4.25), either attached to its amine group or inserted into a C–H or C=C bond of the side chain of phenylalanine.

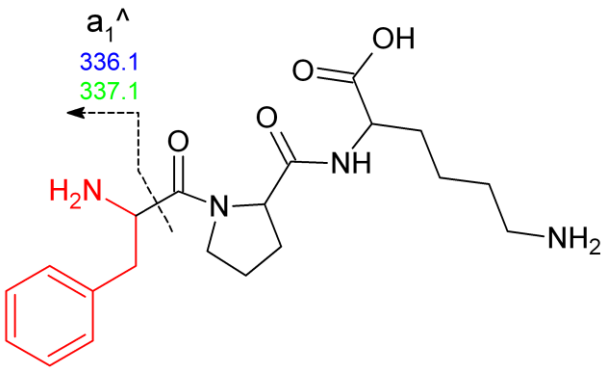


Figure 4.25: Schematic of the peptide $_{223}$FPK$_{225}$ highlighting the only labelled fragment, a_1^, that can be observed on the mass spectra of Figure 4.24 from the CID of its doubly charged parent ion. The observed *m/z* for pPD-labelled (blue) and for dPD-labelled (green) a_1 ions are listed over dashed line that intersects with the peptide, marking the bond being cleaved to generate the a_1 ion. The arrow direction point towards the amino acid that is included in the a_1 fragment, which is shown in red, as phenylalanine-223 is where photofol binds.

4.6.5 Labelled peptide HKPK

It was noted in section 4.4 that the pPD-labelled peptides with a MW of 724.4 Da eluted at three different retention times spanning 42.9–46.0 minutes. Peptide $_{535}$HKPK$_{538}$ is the only tryptic fragment in HSA that matches the requirement of having a MW of 508.2 Da (the MW of the labelled peptide minus the MW of the photolabel) and was therefore predicted to be the peptide that was labelled.

MS-MS was carried out on the three different labelled peptides eluted at the different retention times. All of the EPI mass spectra confirmed the peptide sequence to be $_{535}$HKPK$_{538}$. Figure 4.26 lists the expected singly charged CID ions of this peptide. Those *m/z* values that are located on the EPI mass spectra (Figure 4.27) from the fragmentation of the singly and doubly charged pPD-labelled ions at 363.1 and 725.2 *m/z*, and the respective dPD-labelled ions at 363.6 and 726.2 *m/z* are in bold. Figure 4.27A and B show the EPI mass spectra for the fragmentation of the most ionising labelled-$_{535}$HKPK$_{538}$ peptides of the pPD- and dPD-labelled samples (at a retention time of 43.8–44.6 minutes in Figure 4.13D). The labelled $_{535}$HKPK$_{538}$ that eluted at an earlier time point

gave the same set of *m/z* peaks but with reduced intensity. The presence of a_1^ ions in both the pPD- and dPD-labelled samples (Figure 4.27A and B) is sufficient to deduce that photofol must be bound to histidine-535 of the most ionising labelled-$_{535}$HKPK$_{538}$ peptide (Figure 4.28A). It was not possible to establish the location of the bound photofol in the labelled-$_{535}$HKPK$_{538}$ peptide that eluted first as the obtained EPI mass spectra also suggested that the a_1 ion fragment was labelled due to the labelled-histidine peptide being much more ionising than the other, more polar, labelled peptides.

a_n ions (*m/z*)	→	**110.1**	238.2	335.2	463.3
b_n ions (*m/z*)	→	138.1	**266.2**	**363.2**	**491.3**
c_n ions (*m/z*)	→	155.1	283.2	380.2	508.3

n =	1	2	3	4	
aa(535–538)	**H**	**K**	**P**	**K**	
4	3	2	1	= *n*	

535.3	398.2	270.1	173.1	←	x_n ions (*m/z*)
509.3	**372.3**	**244.2**	**147.1**	←	y_n ions (*m/z*)
492.3	355.2	227.1	130.1	←	z_n ions (*m/z*)

Figure 4.26: The expected *m/z* of the singly charged CID ions of the peptide HKPK. The *m/z* that are observed in the EPI mass spectra (Figure 4.27) are in bold.

Figure 4.27C and D show the EPI mass spectra of the labelled-$_{535}$HKPK$_{538}$ peptides that eluted last at around 45.1 minutes on the HPLC column that is coupled to the fraction collector. These two spectra lack the characteristic a_1^ peaks at 326.3 and 327.4 *m/z* for the pPD- and dPD-labelled samples respectively that were seen in Figure 4.27A and B. However, peaks with 1 ± 0.1 *m/z* unit difference between the pPD- and dPD-labelled samples occur at 300.4 and 345.3, and at 301.5 and 346.3 *m/z* respectively. These peaks correspond to the pPD- and dPD-labelled K_{im}* and K_{aa} ions (discussed in section 4.6.2), suggesting that a lysine residue was labelled. The presence of b_2^ ions allows the position of the bound photolabel to be determined as the lysine sandwiched between histidine-535 and proline-537 (Figure 4.28B). It is interesting to note that b_2^ ions were observed in all of the spectra presented in Figure 4.27 but it is believed that they arose from histidine-535 being labelled in the spectra of panels A and B, and lysine-536 being labelled in the spectra of panels C and D. This is due to the co-appearance of a_1^ ions and the absence of labelled-internal lysine ions in the former two spectra, and vice versa for the latter two spectra of Figure 4.27.

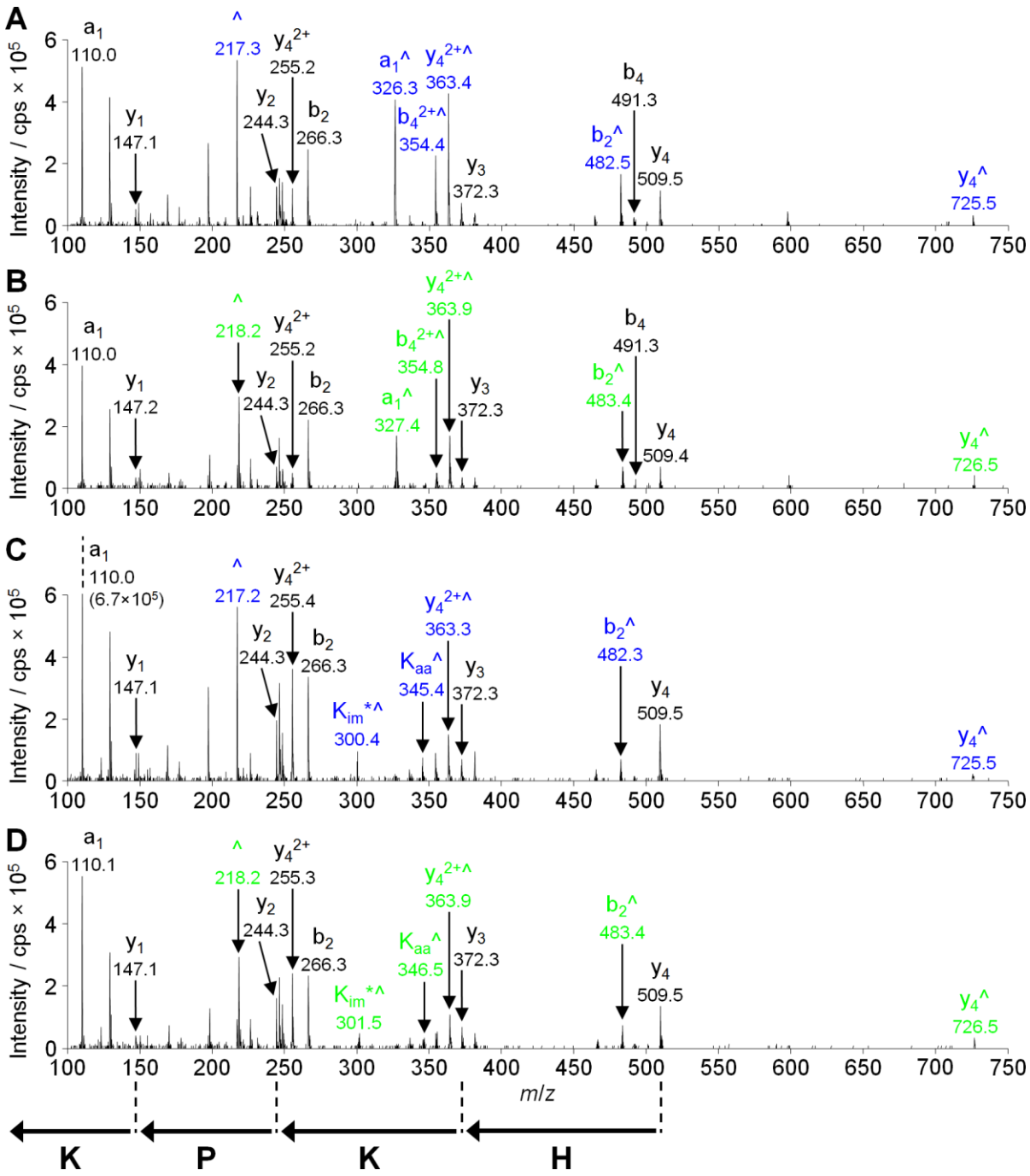


Figure 4.27: Enhanced product ion mass spectra for the collision-induced dissociation of the labelled-HKPK peptide. Data from fragmenting the singly and doubly charged ions were combined and shown on the same spectrum. Spectra for the pPD- (A) and dPD-labelled (B) HKPK eluted at 44.0 minutes (Figure 4.13D) indicate labelling on histidine-535. Spectra for the pPD- (C) and dPD-labelled (D) HKPK eluted at 45.1 minutes indicate labelling on lysine-536. The peptide sequence was confirmed as HKPK by the presence of many b_n and y_n ions (black). Peaks corresponding to labelled fragments are marked with a caret (blue for pPD-labelled and green for dPD-labelled fragments). K_{im}*^ denotes the labelled immonium-less-ammonia ion of lysine. K_{aa}^ denotes the labelled amino acylium ion of lysine.

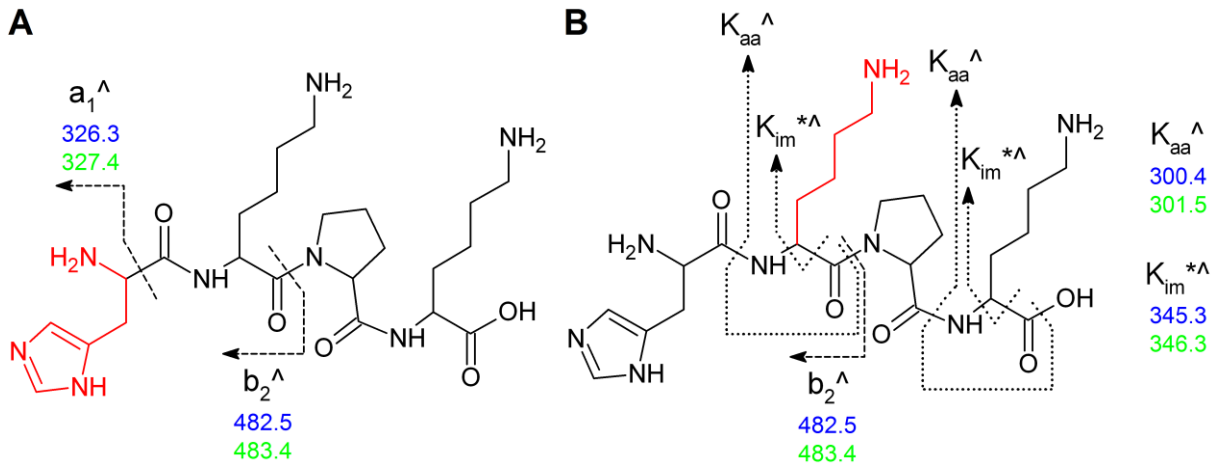


Figure 4.28: Schematic of the peptide $_{535}$HKPK$_{538}$ with annotations showing the labelled CID fragments observed in the mass spectra of Figure 4.27 for the labelled HKPK eluted at 44.0 minutes (A) indicate labelling on histidine-535 (red); and for the labelled HKPK eluted at 45.1 minutes (B), indicating label on lysine-536 (red). The observed *m/z* for pPD-labelled (blue) and dPD-labelled (green) fragments are listed below their CID notations. A caret signifies the presence of a label bound to the fragment. K_{im}*^ denotes the labelled immonium-less-ammonia ion of lysine. K_{aa}^ denotes the labelled amino acylium ion of lysine. Dashed lines intersecting with the peptide indicate the bonds that become cleaved to generate ion fragments which extends towards the N-terminal (left arrow) or C-terminal (right arrow). Dotted lines intersecting with the peptide indicate the pair of bonds that become cleaved to generate internal ion fragments.

4.6.6 Labelled peptide AWAVAR

The labelled peptides with a MW of 888.2 Da eluted from the HPLC column at two retention time ranges (67.3–68.2 and 68.2–69.2 minutes in Figure 4.13E). The peptides were predicted to be $_{213}$AWAVAR$_{218}$ and $_{520}$ERQIK$_{524}$ as both of these peptides have a MW of 672.3 Da (section 4.5). The other hypothesis was that only one of these peptides had been labelled but at two different positions along its sequence, much like the peptide $_{535}$HKPK$_{538}$ (section 4.6.5). Almost identical EPI spectra were obtained for the fragmentation of the peptides eluting at the two slightly different time ranges, with the same set of *m/z* peaks being observed and only differences seen in the intensities of the peaks. Representative EPI spectra for both the pPD- and dPD-labelled samples of the daughter ions of the doubly charged ions of 445.1 and 445.6 *m/z* respectively are shown in Figure 4.29. The presence of *m/z* peaks corresponding to five of the six y_n-series ions of $_{213}$AWAVAR$_{218}$ confirmed that this peptide was labelled. Besides the parent ions, which are equivalent to the y_6^{2+}^ ions, none of the common CID fragments of the peptide $_{520}$ERQIK$_{524}$ were

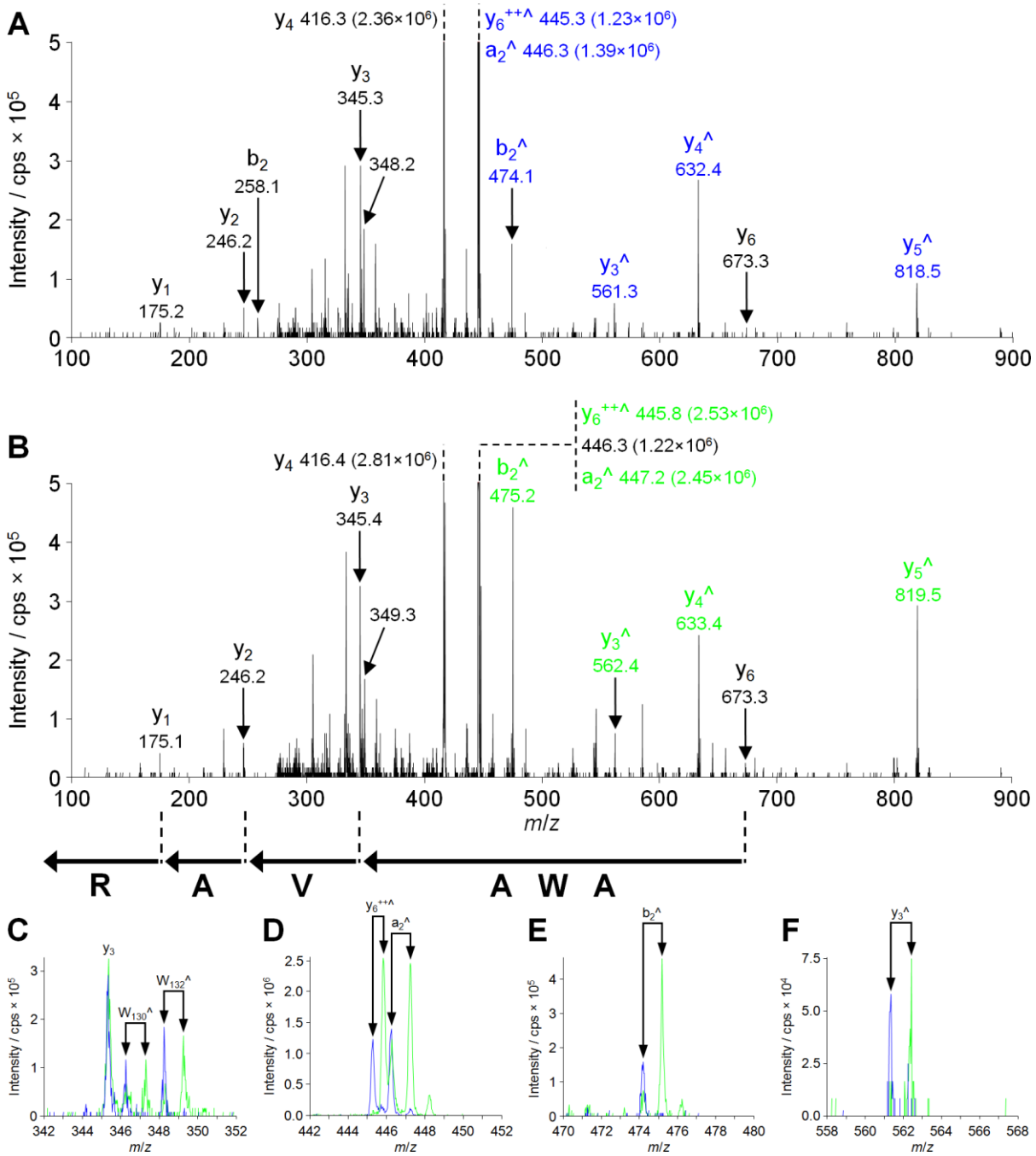


Figure 4.29: Enhanced product ion mass spectra for the fragmentation of the parent ions 445.1 and 445.6 *m/z* for the pPD- (A) and dPD-labelled (B) samples respectively. The labelled peptide is confirmed to be $_{213}$AWAVAR$_{218}$ based on almost all of the y_n-ion series for this peptide being observed. Peaks which correspond to labelled (^) fragment ions are marked in blue (pPD-labelled) and green (dPD-labelled). $W_{m/z}$^ denotes the related ions of tryptophan, with 130 or 132 *m/z* (Falick *et al.*, 1993) (Papayannopoulos, 1995) being labelled. Peaks which extend beyond the y-axis scale were marked with a dashed line at the top of the spectra and their intensities in cps were noted in brackets. Figures C–F show overlay mass spectra for the pPD-(blue) and dPD-labelled samples (green) centred around the y_3, W_{130}^ and W_{132}^ peaks (C); the y_6^{2+}^ and a_2^ peaks (D); the b_2^ peaks (E) and the y_3^ peaks (F).

observed, ruling out this peptide as being labelled. Figure 4.30 shows the common CID fragment ions of $_{213}$AWAVAR$_{218}$ and highlights those that were observed on the EPI mass spectra.

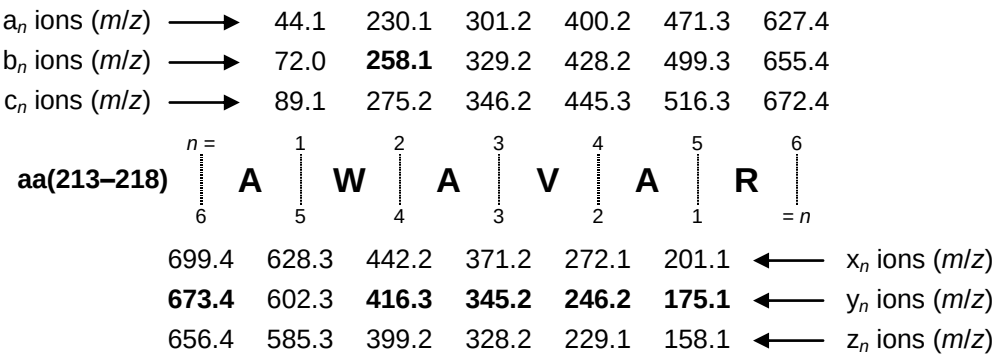


Figure 4.30: The expected *m/z* of the singly charged CID ions of the peptide AWAVAR. The *m/z* that are observed in the EPI mass spectra (Figure 4.29) are in bold.

There is clear evidence that the photolabel attached to $_{213}$AWAVAR$_{218}$ at different positions. Labelled-peptide fragments consisting of the first two amino acids were identified in the form of a_2^ and b_2^ ions. These two ions suggest that the photolabel had been bound to either alanine-213 or tryptophan-214. However, peaks corresponding to the fragments comprising of the three, four and five C-terminal amino acids with the photolabel bound—y_3^, y_4^ and y_5^—were also detected. A y_3^ ion fragment confines the photolabel to the amino acids valine-216, alanine-217 or arginine-218.

A few different collision energy settings were attempted by trial and error to attain better EPI spectra in order to ascertain more precise locations of photofol-binding on $_{213}$AWAVAR$_{218}$ but this was not very successful. Due to the lack of b_n-series ions, both with and without a photolabel bound, it was not possible to pinpoint the C-terminal photolabel site to a single amino acid. However, there was evidence suggesting that tryptophan-214 had been labelled. Fragments at 346.2 and 348.2 *m/z* were seen in the pPD-labelled peptide, whilst fragments at 347.2 and 349.3 *m/z* appeared in the dPD-labelled peptide (Figure 4.29C). Assuming that these are singly charged ions—due to the *m/z* of the D_0- and D_1-labelled isotopes being one *m/z* unit apart—this suggests that the singly charged ions of 130 and 132 *m/z* were labelled. Two related ions of tryptophan from CID were documented to have such monoisotopic *m/z* values (Papayannopoulos, 1995). The presence of these labelled internal tryptophan ions therefore identifies one of the photofol-binding locations as tryptophan-214.

The two labelled regions at tryptophan-214 and aa(216–218), along with the labelled fragments that had been observed in the EPI mass spectra of Figure 4.29 to deduce them, are summarised in Figure 4.31.

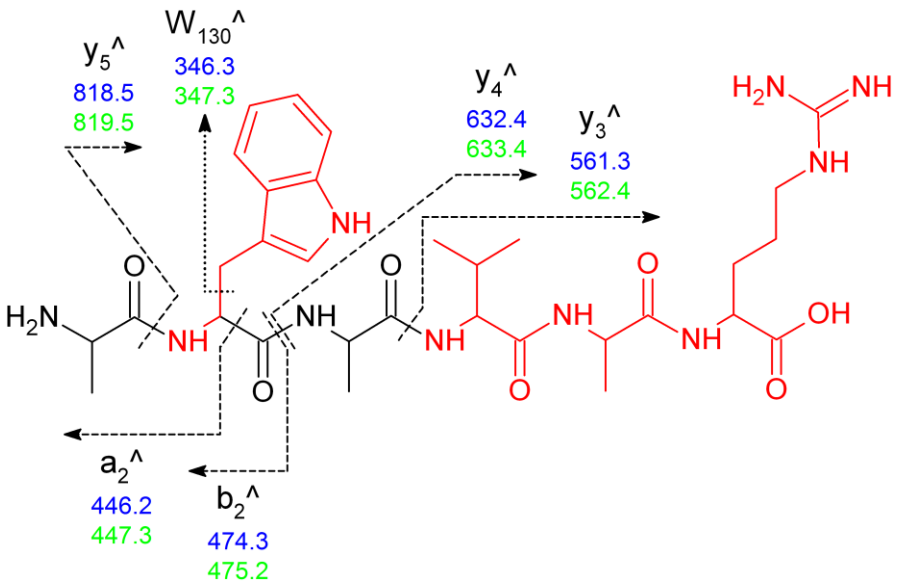


Figure 4.31: Schematic of the peptide $_{213}$AWAVAR$_{218}$ with annotations showing the labelled CID fragments identified in the mass spectra of Figure 4.29. The observed *m/z* for pPD-labelled (blue) and dPD-labelled (green) fragments are listed below their CID notations. A caret signifies the presence of a label bound to the fragment. W_{130}^ denotes a related ion of tryptophan (Falick *et al.*, 1993; Papayannopoulos, 1995) being labelled, with the dotted line intersecting with the peptide marking the cleaved bond to generate this internal ion. Dashed lines intersecting with the peptide indicate the bonds that become cleaved to generate ion fragments which extends towards the N-terminal (left arrow) or C-terminal (right arrow). Two regions of this peptide had been labelled (marked in red), concurring with the SIM scan ion current chromatogram that showed the labelled peptides with MW of 888.2 Da eluting from the HPLC column at two different time points Figure 4.13E.

4.7 Isotopic peak distributions of labelled peptides

Having established the sequences of the labelled HSA peptides in the previous section, it is now possible to predict their isotopic distributions using the known composition ratio of the different photofol isotopes. In this section, the theoretical isotopic distributions of three of these labelled peptides are compared with the experimental data. A function based on the *isotopicdist* function of *MATLAB R2012a* was coded by Dr Christopher Edge incorporating functionalities for allowing the modification of cysteine residues by iodoacetamide and the addition of photofol with zero to two deuterium substitutions to the peptide ($D_{0–2}$). The addition of D_0 photofol to peptides with

this bespoke function gives rise to the isotopic distribution of labelled peptides in the pPD-labelled sample. For the theoretical isotopic distribution of the dPD-labelled peptides, the combined isotopic density from the D_0-, D_1- and D_2-labelled peptides was plotted according to their percentage composition in the dPD sample, which is 43 %, 34 % and 23 % respectively for the three isotopes (Figure 4.32).

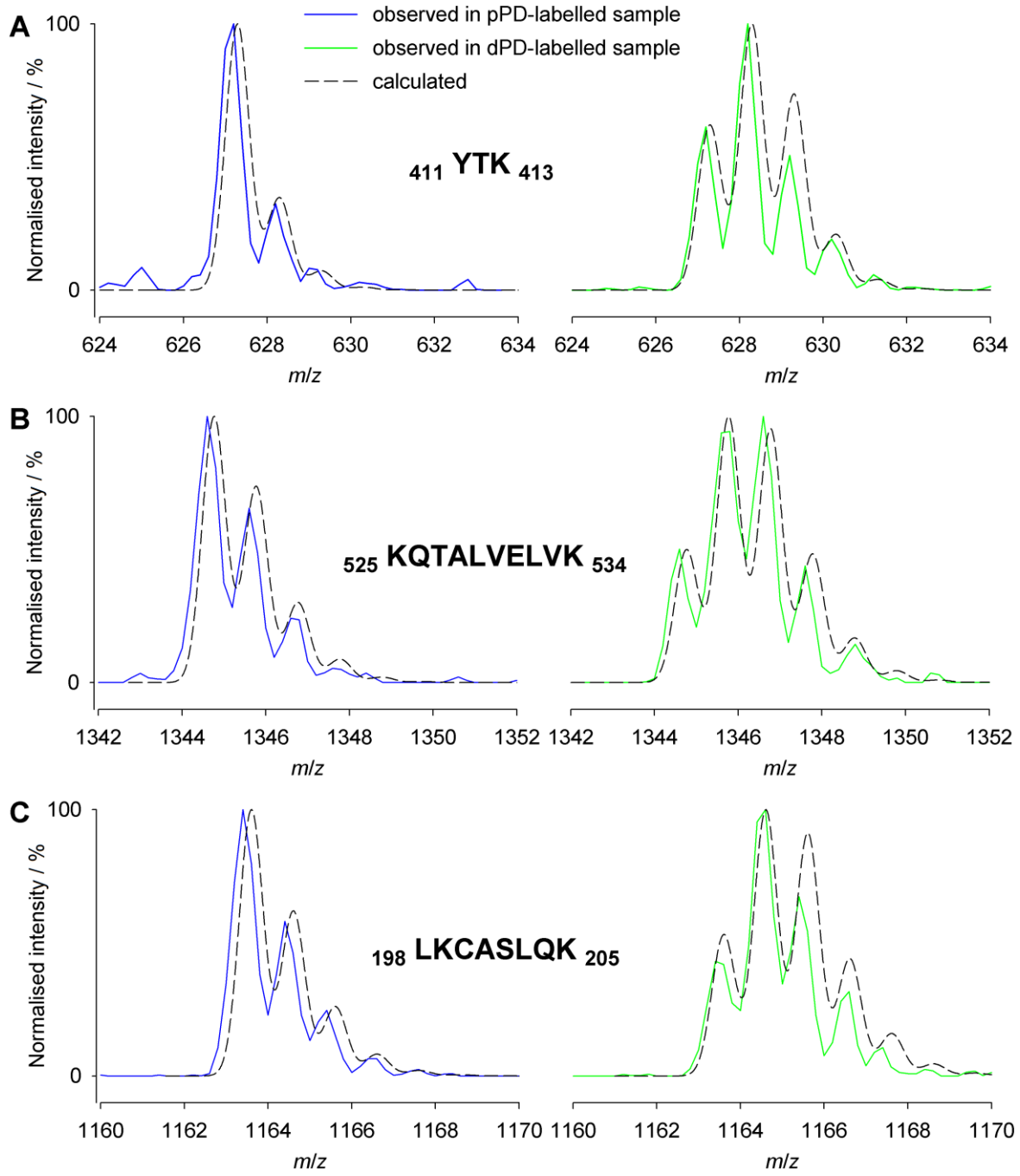


Figure 4.32: Calculated (black) and observed isotopic distributions for the pPD- (blue) and dPD-labelled (green) peptides, YTK (A), KQTALVELVK (B) and LKCASLQK (C). Ionisation intensities are normalised to the maximum intensity of the peaks in view in each of the graphs.

The calculated and observed isotopic distributions for the pPD-labelled samples are almost identical (left-hand side mass spectra of Figure 4.32). For the dPD-labelled samples, the fit between the theoretical and experimental distributions is good for the zero and one deuterium-substitution peaks but beyond that the distributions do not match so well, as demonstrated by the spectra on the right-hand side of Figure 4.32A and C. The observed intensities of the D_2 peaks for the dPD-labelled $_{411}$YTK$_{413}$ and $_{198}$LKCASLQK$_{205}$ are both less than their calculated intensities. However, for the largest labelled peptide, $_{525}$KQTALVELVK$_{534}$, the observed intensity for the dPD-labelled sample is similar to the calculated distribution for all of its isotopic peaks (Figure 4.32B). Despite the fit between the theoretical and experimental data not being perfect, it is still clear that the observed patterns for dPD-labelled peptides are distinct to those of labelled peptides in the pPD sample and are unlikely to occur by chance, thus allowing these deliberate patterns of isotopic peaks to be used as practical markers for distinguishing between peaks arising from unlabelled and labelled peptides.

4.8 Labelled amino acids in the HSA structure

The published crystallographic data of propofol-bound HSA revealed two binding sites of propofol in subdomains IIIA and IIIB, namely propofol sites 1 and 2 respectively (Bhattacharya *et al.*, 2000). In this chapter, as well as labelling amino acids within these two known propofol-binding sites, five amino acids have been identified to be labelled by photofol within a binding cavity known as drug site I (Sudlow *et al.*, 1975; Sudlow *et al.*, 1976). This cavity is formed from the end of subdomain IB helix 4 and all of subdomain IIA, and has been shown to bind to a variety of drugs, such as the endogenous toxin CMPF (3-carboxy-4-methyl-5-propyl-2-furanpropionic acid), oxyphenbutazone and phenylbutazone (Ghuman *et al.*, 2005). The electron density map for propofol-bound HSA was reinvestigated and clear density at the region of drug site I was visible. Professor Stephen Curry, one of the co-authors of the paper that showed the high-resolution structure of propofol-bound HSA (Bhattacharya *et al.*, 2000), provided a modified PDB file—from the original file with accession code 1E7A—which includes a third propofol molecule fitted to its electron density at drug site I, hereafter referred to as propofol site 3. Figure 4.33 summarises the amino acids identified to be labelled and their locations categorised into these three propofol-binding sites from the X-ray structural data.

Labelled peptide	Binding site
191 ASSA**K**QR 197	Propofol site 3
198 L**K**CASLQK 205	Propofol site 3
210 AF**K**AWAVAR 218	Propofol site 3
213 A**W**AVAR 218	Propofol site 3
223 **F**PK 225	Propofol site 3
411 **Y**TK 413	Propofol site 1
525 **K**QTALVELVK 534	Propofol site 2
535 **HK**PK 538	Propofol site 2
539 AT**K**EQLK 545	Propofol site 2

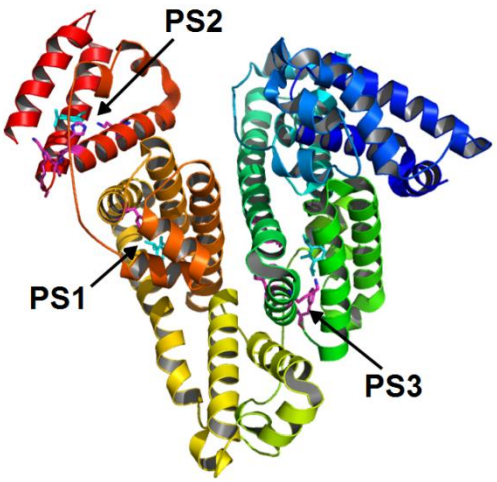


Figure 4.33: Summary of the identified photofol-bound amino acids within their labelled peptides and their locations in known binding sites of HSA. (Left) Red in bold in the table indicates the amino acids that have been photolabelled. (Right) A cartoon representation of the structure of HSA with propofol molecules (cyan sticks) fitted to their electron density maps on the three binding sites, propofol sites 1–3 (PS1, PS2 and PS3), respectively. Labelled amino acids are also represented in stick models in magenta.

Figure 4.34 shows that the sole amino acid that has been labelled in the vicinity of propofol site 1, tyrosine-411, has a nucleophilic hydroxyl group that is 4.9 Å from the middle carbon of one of the isopropyl groups of propofol, where a carbene moiety would be if the propofol molecule is replaced by a photoactive photofol molecule. The tyrosine residue is located at the opening of this binding cavity close to the carbonyl group of leucine-430 that has been shown to interact with the hydroxyl group of propofol in the high resolution structure (Bhattacharya *et al.*, 2000). There is no evidence showing labelling of other amino acids close to the isopropyl arm of propofol that is deep inside the binding pocket. This suggests that the bulkier photoactivatable moiety of photofol compared to an isopropyl arm of propofol is less inclined to enter this relatively small binding cavity compared to the other two binding sites. The internal surface cavity volume has been reported to be around 340 $Å^3$ (Liu *et al.*, 2005) and a 120 ° rotation of the side chain, valine-433, was required to accommodate the inner isopropyl group in the drug-bound HSA structure (Bhattacharya *et al.*, 2000). Consistent with the labelling of only a single amino acid by an alkyldiazirinyl moiety replacing the solvent-exposed isopropyl arm of propofol at this site is the electron density for this isopropyl group being much better defined than that of the isopropyl group deep inside the binding cavity (Bhattacharya *et al.*, 2000).

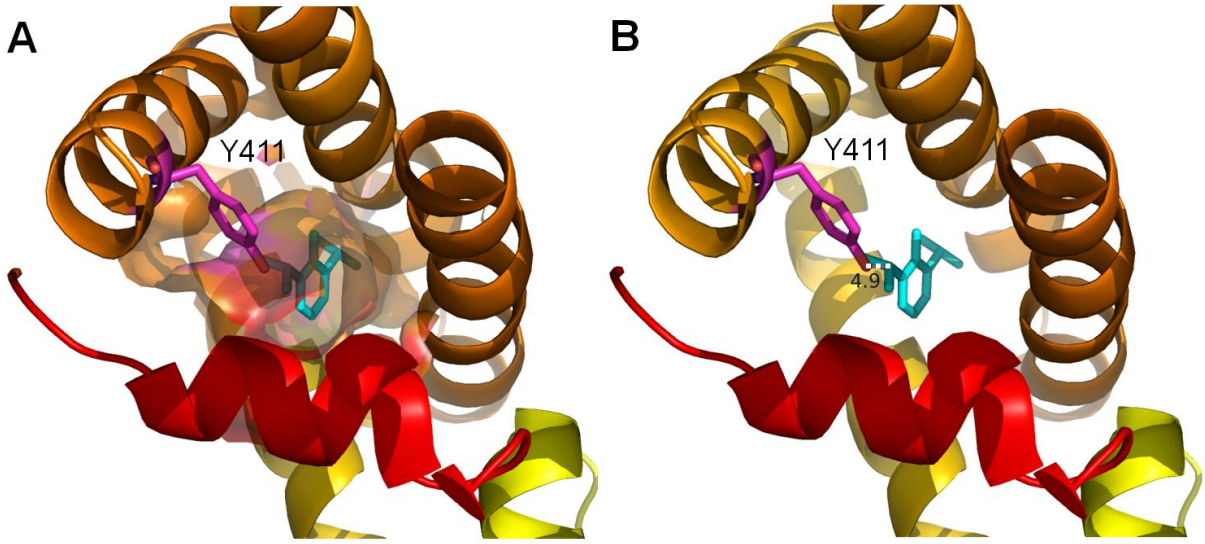


Figure 4.34: Propofol site 1 of HSA showing the proximity of the amino acid tyrosine-411—which has been shown to be labelled by photofol—to the bound propofol in the X-ray structure. (A) The surface of the binding cavity of propofol site 1 superimposed onto a cartoon representation of subdomain IIIA of HSA with the bound propofol (cyan) and tyrosine-411 (magenta) in sticks representation. (B) The distance between the hydroxyl group of tyrosine-411 and the middle carbon of one of the isopropyl arms of propofol is shown to be 4.9 Å.

The internal surface cavity of propofol site 2 was reported to be 810 $Å^3$ (Liu *et al.*, 2005). The crystallographic data of propofol bound to HSA shows a hydrogen bond between the hydroxyl groups of serine-579 and propofol (Bhattacharya *et al.*, 2000). Figure 4.35 shows the proximity of one of the nitrogens of the R group of histidine-535 being 5.0 Å from the propofol isopropyl group that is on the entry side of this binding cavity. The adjacent lysine residue, lysine-536, also lines the surface of the entrance of this pocket, whilst the terminal nitrogen of lysine-525, with its flexible alkyl side chain pointing towards the opening of the pocket, lines the back of the pocket and is 4.9 Å from the middle carbon of the opposite isopropyl arm of propofol. The third lysine that has been labelled in subdomain IIIB of HSA is lysine-541. As with lysine-536, this lysine sits on a flexible loop at the opening of the cavity. It is further away from either of the isopropyl arms of the bound propofol molecule in the X-ray structure than the other labelled amino acids identified in this subdomain despite modelling the R group of this lysine such that its alkyl chain points towards the propofol molecule in the PDB file, even though almost no electron density is seen beyond its beta carbon.

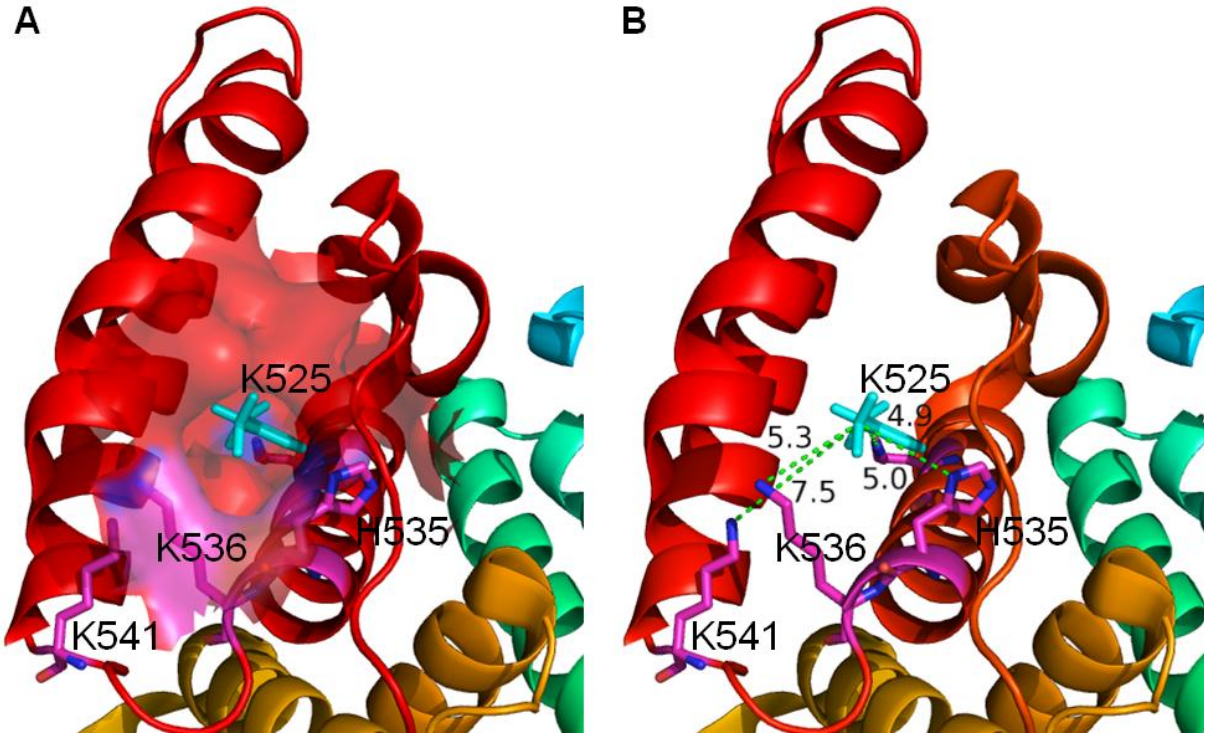


Figure 4.35: Propofol site 2 of HSA showing the proximities of the amino acids lysine-525, histidine-535, lysine-536 and lysine-541—all of which have been shown to be labelled by photofol—to the bound propofol in the X-ray structure. (A) The surface of the binding cavity of propofol site 2 superimposed onto a cartoon representation of subdomain IIIB of HSA with the bound propofol (cyan) and the amino acids identified to bind to photofol (magenta) in sticks representation. (B) The distances between the terminal nitrogen atoms in the R groups of these lysines, or one of the nitrogens of the R group of histidine-535 and the middle carbon of one of the propofol isopropyl arms are marked with dashed lines and recorded in angstroms. Note that the coordinates of the gamma, delta and epsilon carbons and the terminal nitrogens of the three lysine residues were manipulated on the propofol-bound HSA PDB file—accession code 1E7A—to determine the minimal distance between the terminal nitrogen and the isopropyl arm of propofol.

Propofol site 3, previously described as drug site 1 in the field of HSA-ligand interactions (Sudlow *et al.*, 1975; Sudlow *et al.*, 1976), was identified to harbour five amino acids that are labelled by photofol (Figure 4.36). The main chamber of this site, which is where propofol electron density had been located, contains the amino acids leucine-238 and alanine-291. These two amino acids interact with drugs containing flat aromatic moiety (Ghuman *et al.*, 2005), allowing them to be sandwiched between the two apolar residues (Figure 4.36B). The X-ray data also suggests hydrogen bond interaction between the hydroxyl group of propofol and one of the terminal nitrogens of arginine-222, forcing the hydroxyl group of the anaesthetic drug within this spacious binding cavity to point towards subdomain IIIA. Assuming that photofol bears the same

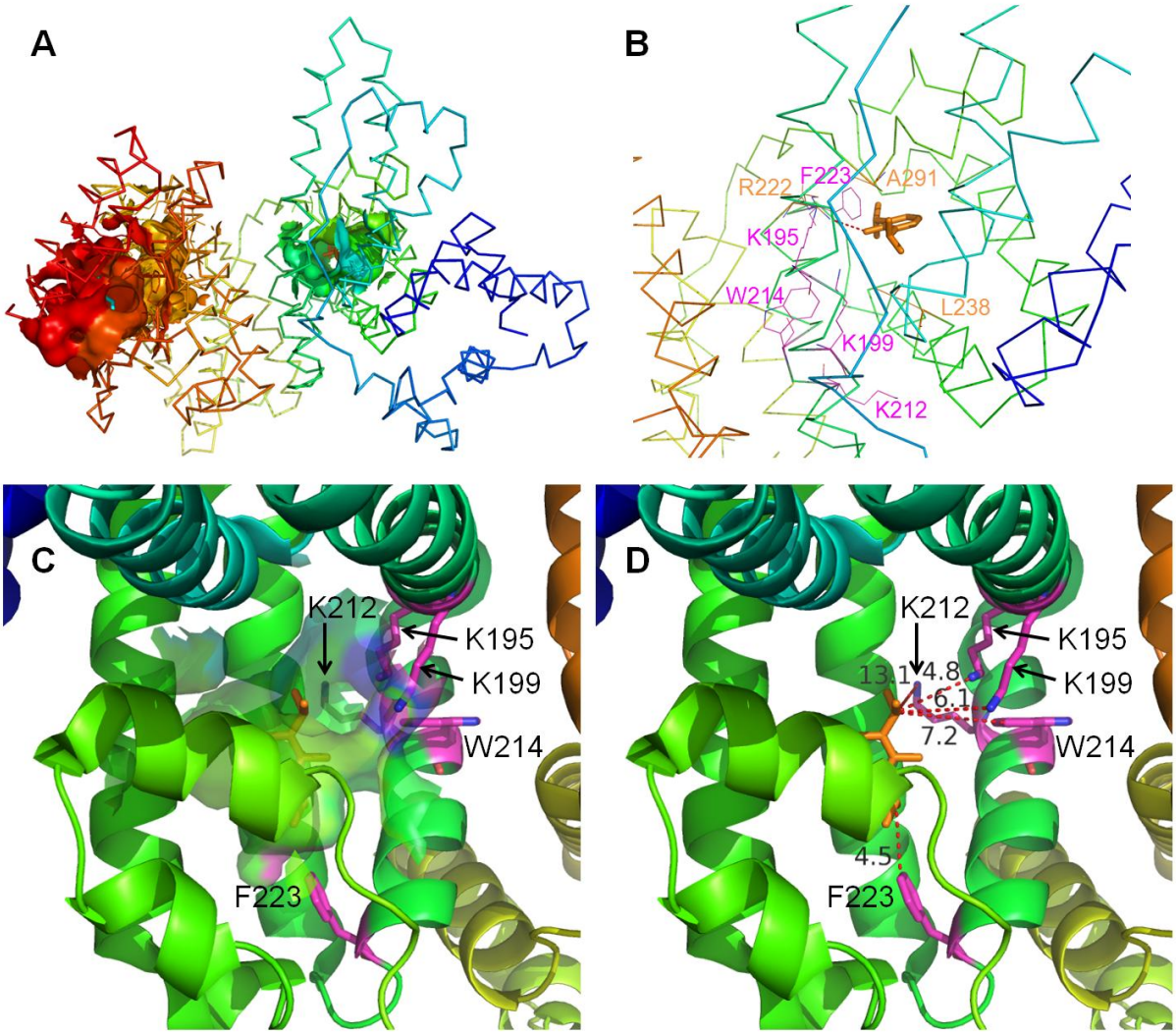


Figure 4.36: Propofol site 3 of HSA showing the proximities of the amino acids lysine-195, lysine-199, lysine-212, tryptophan-214 and phenylalanine-223—all of which have been shown to be labelled by photofol—to the bound propofol in the X-ray structure. (A) The surface of the binding cavity of all three propofol sites superimposed onto a ribbon representation of HSA viewed from the top in relation to the cartoon representation in Figure 4.33. (B) Enlarged top view of propofol site 3 viewed from same orientation as A with propofol represented in orange stick, leucine-238, alanine-291 and arginine-222 represented by orange lines and the five aforementioned labelled amino acids in magenta. The dashed line shows hydrogen bonding interaction to arginine-222. Panels (C) and (D) Views from the back of HSA in relation to cartoon representation in Figure 4.33 with (C) showing the surface of the binding cavity superimposed on to a cartoon representation of subdomain IIA and (D) showing the distances in angstroms between the middle carbon of one of the isopropyl groups of propofol and anticlockwise from the bottom the benzene ring of phenylalanine, a carbon from the aromatic ring of tryptophan, and the terminal nitrogen atoms in the R groups of lysine-199, lysine-195 and lysine-212. Note that the coordinates of the gamma, delta and epsilon carbons and the terminal nitrogens of lysine-195 and lysine-212 were manipulated on the propofol-bound HSA PDB file—accession code 1E7A—to determine the minimal distance between the terminal nitrogen and the isopropyl arm of propofol.

orientation as the propofol molecule in the X-ray structure, then it is not surprising that the amino acid residues lysine-195, lysine-199 and tryptophan-214, which line the wall of the binding cavity closest to subdomain IIIA, have been identified to be labelled. The benzene ring of phenylalanine-223 sits at the loop linking helices 2 and 3 of subdomain IIA, and is 4.5 Å from one of the isopropyl arms of propofol in the crystallographic model (Figure 4.36D). A labelled fragment, a_1^, of the labelled peptide $_{223}$FPK$_{225}$ was observed in its EPI mass spectra fragmenting the doubly charged ions of this peptide (Figure 4.24), suggesting that the photolabel must be bound to either the nitrogen of the peptide bond between phenylalanine-223 and arginine-222 (8.4 Å from the isopropyl group), or is inserted into a C–H bond. The latter means that labelling would have occurred through carbene insertion rather than through nucleophilic attacks by quinone methide. The lysine residue at position 212 is much further away from the bound propofol in the structural model being 13.1 Å away. The distance to an isopropyl group of propofol could be reduced if the anaesthetic is able to move deeper into the binding cavity to where, amongst other drugs, indomethacin, a drug that is much bulkier than propofol, binds (Ghuman *et al.*, 2005). However, the X-ray structure of HSA shows that this amino acid is likely to not line the wall of the binding cavity as it is on the side of the helix that faces the surrounding solvent.

4.9 Concluding remarks for the photoaffinity labelling of HSA

The photoaffinity labelling of HSA was successful in that there is clear evidence for the labelling of amino acids within the two propofol sites previously determined by crystallography (Bhattacharya *et al.*, 2000). The labelling of five amino acids within drug site I of HSA (Sudlow *et al.*, 1975; Sudlow *et al.*, 1976) prompted revisiting of the X-ray data, which showed propofol electron density within this site that was missed in the initial analysis conducted over a decade ago. The binding of propofol to additional sites other than propofol sites 1 and 2 is supported by Liu *et al.* (2005) showing that the binding of propofol was reduced by two-fold when a truncated HSA protein, consisting of only domain III, was used.

Tyrosine-411 was the only amino acid in the vicinity of propofol site 1 that had been labelled. This residue sits at the entrance of the pocket. The absence of labelling deep in the binding cavity is consistent with this site being relatively small; compared to propofol sites 2 (Liu *et al.*, 2005) and 3

(Curry, 2009), thus favouring the less bulky isopropyl group to enter the hydrophobic binding site rather than the alkyldiazirinyl group. In addition, the molecular volume of an alkyldiazirinyl group is more similar to that of a tertiary butyl group than to an isopropyl group, where the former's present at both the 2- and 6- position of propofol renders the drug unable to bind to a binding cavity of limited dimensions at the $GABA_A$ receptor, whilst a propofol analogue with a tertiary butyl group and a methyl group at the 2- and 6- position remains active (Krasowski *et al.*, 2001a), thus suggesting that at least in the $GABA_A$ receptor, potentially only one bulky arm can fit inside the binding cavity. The solvent-exposed isopropyl arm of the propofol at this site in the X-ray structure showed well-defined electron density (Bhattacharya *et al.*, 2000), so it is not surprising that only a single amino acid was labelled here.

Three lysine residues and a histidine residue were labelled in subdomain IIIB, which contains propofol site 2. Two of these lysines, at amino acid positions 536 and 541, sit on the same loop-secondary structure at the entrance of the cavity. The third labelled lysine at position 525 lies deep inside the pocket, with its terminal nitrogen being 4.9 Å from the inner isopropyl arms of propofol. In the myristate-bound HSA X-ray structure, the binding site in this subdomain occurs as a tunnel open on both sides (Curry *et al.*, 1998). However, in the HSA-bound structure, the opening at the back appears to be not as wide as the anaesthetic drug, capping off the binding tunnel at one end. Finally, histidine-535 lines the cavity wall at the entrance, with one of the nitrogens of its R group approximately 5 Å from the potential location of the carbene of an activated photolabel. More labelled amino acids being identified near the isopropyl arm at the binding cavity entrance rather than near the isopropyl arm that is buried within the hydrophobic pocket is consistent with the labelling results from propofol site 1.

Labelled amino acids within site 3 were almost completely confined to the side of the binding cavity containing arginine-222, in which hydrogen bond interaction (3.4 Å) occurs between one of its terminal nitrogens and the hydroxyl group of the photolabel. The presence of a handful of labelled amino acids within the same site provided information about the orientation of the photolabel within the binding cavity. The labelled phenylalanine-223 has its aromatic hydrocarbon separated from the potential carbene of photofol by only 4.5 Å, whilst its shared peptide bond with arginine-222 is about 8.5 Å from the isopropyl group of propofol in the X-ray

structure. This suggests that labelling occurred via insertion into the C–H bond of the amino acid, indicating that the reaction was via carbene addition and not quinone methide, at least in this case. However, the labelled lysine-212 is approximately 13 Å from the propofol molecule best fitted to its electron density at site 3. It was noted that this distance could be reduced if propofol is able to move deeper into the binding cavity, away from the main chamber where drugs such as indomethacin bind (Ghuman *et al.*, 2005). Nevertheless, the vast distance between the labelled amino acid and the photolabel suggests that labelling is more likely to have occurred after rearrangement of the activated photolabel to a quinone methide, which has a longer half-life than carbene.

The labelling by photofol being specific to amino acids within binding cavities was demonstrated by the photoaffinity labelling of tryptic peptides of HSA not yielding any adducted peptides (data not shown). However, it appears that lysine residues are more prone to labelling than other residues, as suggested by six of the ten identified labelled amino acids being lysines, constituting 10 % of the total number of lysines within HSA. A previous photolabel synthesised from our group, an azyl azide etomidate, labelled only two lysine residues of HSA (Bright *et al.*, 2007) as part of the peptides $_{198}$LKCASLQK$_{205}$ and $_{525}$KQTALVELVK$_{534}$. Both of these peptides have also been shown to be labelled by photofol at their missed-cleavage lysine residues, demonstrating that the binding of both of these photolabels at a lysine residue interferes with trypsin digestion.

Photolabelling with two other propofol diazirines was carried out (data not presented). These photolabels had their diazirine groups at the *meta* and *para* positions instead of replacing one of the isopropyl arms at the *ortho* position of propofol. The *para*-photolabel was shown to have a weaker potency in potentiating $GABA_A$ receptors than the *ortho*-photolabel by ten-fold, whilst the *meta*-photolabel was not shown to have anaesthetic effects from electrophysiology experiments conducted by Dr Robert Dickinson. Labelling experiments with the *para*-compound showed no labelling at propofol sites 1 and 2 but this bulkier drug was able to label amino acids within the more spacious site 3. Photoaffinity labelling experiments with the *meta*-photolabel on HSA was carried out by Erasmus student, Damian Droste. The *meta*-photolabel, which was suggested by the electrophysiology data to not bind to the propofol-binding site(s) of $GABA_A$ receptors, may also

not bind to any of the three known propofol-binding sites of HSA, as no labelled peptides were identified from the photoaffinity labelling experiments.

The usefulness of using the pPD and dPD samples in parallel was demonstrated by the ease of assigning mass spectra peaks as arising from adducted peptides based on the distinctive isotopic distribution of the dPD-labelled peptides compared to the pPD-labelled peptides observed (sections 4.3.2 and 4.7). Additionally, the simultaneous analysis of EPI mass spectra from both the pPD- and dPD-labelled samples allowed the differentiation of peaks arising from labelled fragments that were isobaric to an unlabelled common CID fragment of the same peptide. An example of this was discussed in section 4.6.2.

In conclusion, the photofol photolabel adducted to amino acids within all three crystallographically-confirmed binding sites of propofol in HSA. Five different amino acids were labelled, including a hydrophobic phenylalanine residue. The photolabelling was demonstrated to be specific in that labelling did not occur on digested peptides lacking binding sites. The novel use of two samples of photolabel, pPD and dPD demonstrated its aid in analysing the proteomic mass spectrometer data post-labelling.

5 Photoaffinity labelling of *Gloeobacter violaceus* ligand-gated ion channel

The success of photolabelling the soluble protein HSA was discussed in chapter 4. In this chapter, results from the photoaffinity labelling of a transmembrane protein are presented. It is now widely accepted that ion channels are important targets of general anaesthetics due to their roles in neuronal signal transmissions. The ligand-gated ion channel from *Gloeobacter violaceus,* GLIC, has been shown to be inhibited by propofol (Weng *et al.*, 2010). However, the main reason for using GLIC as a model for testing our novel propofol analogue photolabel photofol is that the X-ray structure of propofol-bound GLIC was solved last year (Nury *et al.*, 2011). The X-ray diffraction data of the crystals, grown under acidic conditions, show five propofol molecules, one in each of the transmembrane domains of the pentameric subunits (Figure 1.13 in section 1.4.2). The propofol molecules in orientations best fitted to their density maps show the drugs in contact with isoleucine-202, tyrosine-254 and threonine-255, whilst mutagenesis of residues within the binding site suggests that threonine-255 and valine-242 are determinants of propofol (Nury *et al.*, 2011). Note that the mutations of isoleucine-202 to alanine, tryptophan and tyrosine had no effect on propofol sensitivity and that the mutation of tyrosine-254 was not discussed by the authors.

Photofol was used to label GLIC to try to replicate the observed binding sites in the published crystallographic data. The main steps outlined in Figure 4.1 of a typical photoaffinity labelling experiment still hold, but the protocol differs from that used for labelling HSA. This is mainly due to the low quantity of protein available: HSA is commercially available whilst GLIC was kindly supplied by collaborator Professor Pierre-Jean Corringer. Other factors include the redundancy in modifying the single cysteine residue present in GLIC, and the lack of lysine and arginine residues in the transmembrane domain renders trypsin inappropriate for digesting the protein into suitably-sized peptides for HPLC-MS analysis.

Furthermore, obtaining high sequence coverage of integral membrane proteins is known to be a challenging task owing to difficulties in solubilising and digesting these proteins and the subsequent liquid chromatography separation of their peptides (Speers and Wu, 2007). Over the course of the project, four different protocols for denaturing and digesting GLIC were trialled to

optimise a methodology, which may in future be utilised for photoaffinity labelling experiments on mammalian $GABA_A$ receptors.

5.1 Identification of peptides from digested GLIC by HPLC-MS

As with the HSA labelling studies, the characterisation of the enzymatic digest of the protein from HPLC-MS is an essential initial step for using mass spectrometry to analyse samples from the photoaffinity labelling experiments. The *in silico* digestion of GLIC with trypsin yields 27 peptides but eight of these are very large peptides with molecular weights of greater than 1500 Da. One of these alone is 56 amino acids in length and covers 18 % of the complete sequence of the 317-amino acid homomeric subunit of GLIC. The identification of retention times for these long peptides will have to be reliant on observing peaks from multiply charged ions and if these peaks exist, they will lack the isotopic distribution characteristics that are associated with the charged states of the ions due to the mass spectrometer used having a range of 150–1500 m/z and a peak width at half-height of 0.6 Da.

The amino acids isoleucine-202 and leucine-206 from TM1; histidine-235, phenylalanine-238 and valine-242 from TM2; tyrosine-254, threonine-255 and isoleucine-258 from TM3; and phenylalanine-303 and asparagine-307 from TM4 have a high probability of being labelled as they are shown to be on the side of the α-helices lining the binding cavity shown in the X-ray crystallography data. These amino acids are part of three tryptic peptides at amino acid positions, aa184–248, aa249–280 and aa281–317, which have molecular weights of 7610.1, 3836.0 and 4054.3 Da respectively. It would be difficult to determine the labelled amino acids by MS-MS should one of these long peptides become labelled as the triple quadrupole mass spectrometer that was used only has a detector range of up to 1600 m/z.

As demonstrated in section 4.3.2, the use of a pair of photofol samples, pPD and dPD, has much enhanced the identification of the labelled peptides of HSA by comparing the isotopic distribution patterns of the mass spectra peaks for singly charged ions. However, for ions with two or more charges, the resolution of the mass spectrometer used is only able to show the averaged m/z of its isotopic peaks, not individual isotopic peaks separated by $^1/_z$ Da, where z is the charge valency of the ion. To capitalise on the use of the pPD and dPD samples, much effort was made to ensure

that as many peptides as possible with a mass range of about 300–1250 Da were produced from cleavages by the chosen enzyme so that their labelled counterparts would be detectable on the mass spectrometer used with their singly charged ions.

Chymotrypsin was chosen to digest GLIC. This serine protease has been reported to cleave at the carboxyl end of the aromatic residues tyrosine, tryptophan and phenylalanine with high specificity, and leucine, histidine and methionine with lower specificity (SIB ExPASy Bioinformatics Resource Portal, 2011). *In silico* chymotrypsin digestion of GLIC carried out with the Peptide Tools function of the software program *ChemStation* assuming cleavage only at high specificity sites yielded 42 peptides with lengths ranging from one to 27 amino acids. Figure 5.1 summaries the population of peptides with varying lengths produced from high specificity cleavages alone or from high and lower specificity cleavages as described above. Peptides having molecular weights of between 300 and 1250 Da, accounting for only 44 % of amino acid coverage generated from high specificity cleavages alone, are expected to be detectable by their singly charged ions upon the addition of the mass of a photolabel. Increasing the duration of chymotrypsin-digestion increases the probability of some of the lower specificity cleavages occurring. Peptides with molecular weights of between 300 and 1250 Da generated from the complete chymotrypsin cleavages after the six aforementioned amino acids account for 60 % amino acid coverage. This could be

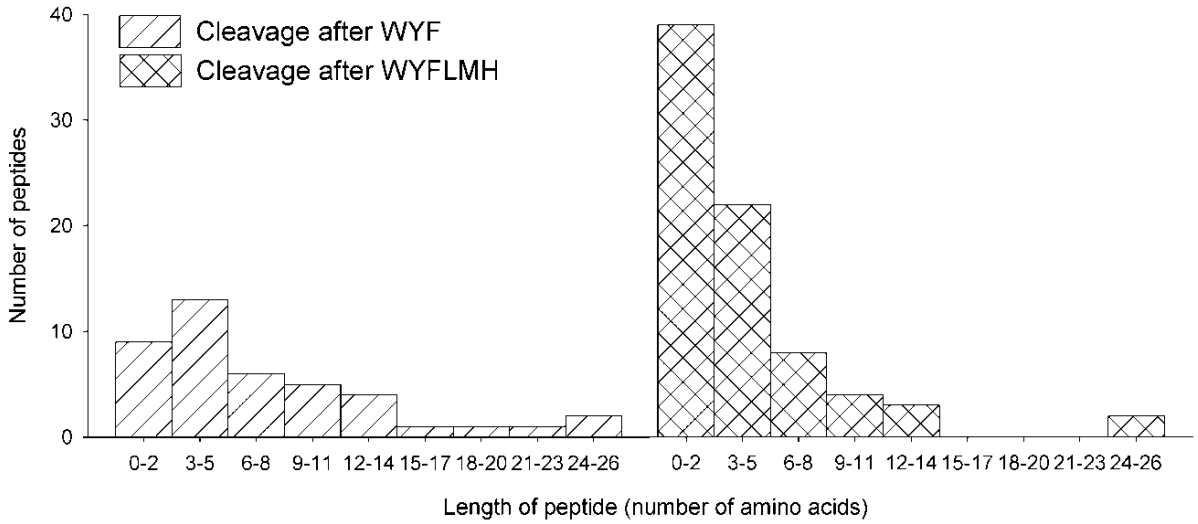


Figure 5.1: Bar charts showing the frequencies of peptides of various amino acid lengths from the hypothetical cleavage of GLIC with chymotrypsin, assuming cleavage after all high specificity sites at the C-terminal to tryptophan (W), tyrosine (Y) and phenylalanine (F), (left), or at all high specificity sites and all C-terminals to leucine (L), methionine (M) and histidine (H), (right).

improved by not allowing cleavages after all six amino acids to reach completion, which would result in the generation of less standalone amino acids and dipeptides. A time-course of chymotrypsin-digestion of GLIC was carried out and peptides generated were monitored by HPLC-MS to determine the optimal chymotrypsin-digestion duration.

5.1.1 Treatment A with guanidine hydrochloride and chymotrypsin

In the initial experiment, GLIC was denatured in 6 M guanidine hydrochloride (GnHCl) and just prior to chymotrypsin-digestion, the reaction mixture was diluted with 100 mM Tris to 1 M GnHCl. Aliquots of the sample were analysed by HPLC-MS at 280-minute intervals after adding chymotrypsin for up to 35 hours. By extracting ions for the expected molecular weights of GLIC peptides beginning after and ending in one of the six amino acids where cleavages are likely to occur and accounting for up to four missed cleavages, the retention times of some peptides were identified. Note that examples of how the Extract Ions tool of *ChemStation* was used to identify the retention times of two HSA peptides are shown in Figure 4.3 and Figure 4.4.

Figure 5.2 summarises the peptides that had been identified through extract ions analysis for all of the different durations of digestion combined, whilst a breakdown of these results for peptides in just the extracellular domain is tabulated in the appendix. The table in Figure A.6 shows that at the time points around nine to thirteen hours of chymotrypsin digestion, almost all of the peptides that had been observed over the 35-h experiment were present, consisting of peptides from cleavages at mainly aromatic residues and leucine.

Nearly all of the amino acids in the extracellular domain (aa1–195) are part of at least one peptide that had its retention time determined, with the exceptions of aa1–4, aa127–129 and aa195. However, only five peptides, spanning 26 of the 122 amino acids in the transmembrane domain, showed an alignment of peaks from their extracted ion chromatograms of their differently charged ions. This gives a total amino acid coverage for GLIC of 67 %, which includes 96 % of the extracellular domain and 21 % of the transmembrane domain.

5.1.2 Treatment B with sodium deoxycholate and chymotrypsin

A new denaturation protocol was tested to improve amino acid coverage of GLIC, which included the use of the bile salt sodium deoxycholate (SDC) to better solubilise the membrane protein



Figure 5.2: Schematic showing the amino acid coverage of GLIC from treatment A. Each yellow bar represents the elution of a peptide that matches the MW of the peptide sequence above it and all amino acid residues that are part of an identified eluted peptide are highlighted in red.

during the enzymatic digestion. The protocol is outlined in section 2.4.2.2. In brief, GLIC was incubated in 2 % SDC for one hour at 37 °C, then the detergent concentration was diluted to 1 % by 100 mM TRIS, pH 7.8 before the sample was digested with chymotrypsin for ten hours. The acidified detergent was subsequently removed by phase transfer separation (PTS) (Masuda *et al.*, 2008) before the peptide samples were injected into the HPLC-coupled mass spectrometer. Figure 5.3 shows the 214-nm absorbance spectra and the TIC chromatograms of GLIC peptides produced from chymotrypsin-digestion after the denaturation of the membrane protein with GnHCl (treatment A) and with sodium deoxycholate (treatment B). The deoxycholate-treated sample appears to have a greater number of peaks in both the UV absorbance and TIC chromatograms compared to the GnHCl-treated sample.

Extract ion analysis was then carried out for the GLIC peptides processed by treatment B and the identified peptides are summarised in Figure 5.4. For the extracellular domain, only the retention times of peptides containing aa1–4 and phenylalanine-195 were not identified. This is an improvement over treatment A, having additionally identified the retention times of $_{127}$HIY$_{129}$ and $_{127}$HIYL$_{130}$ by both their singly and doubly charged ions at 29.4 and 31.0 minutes respectively.

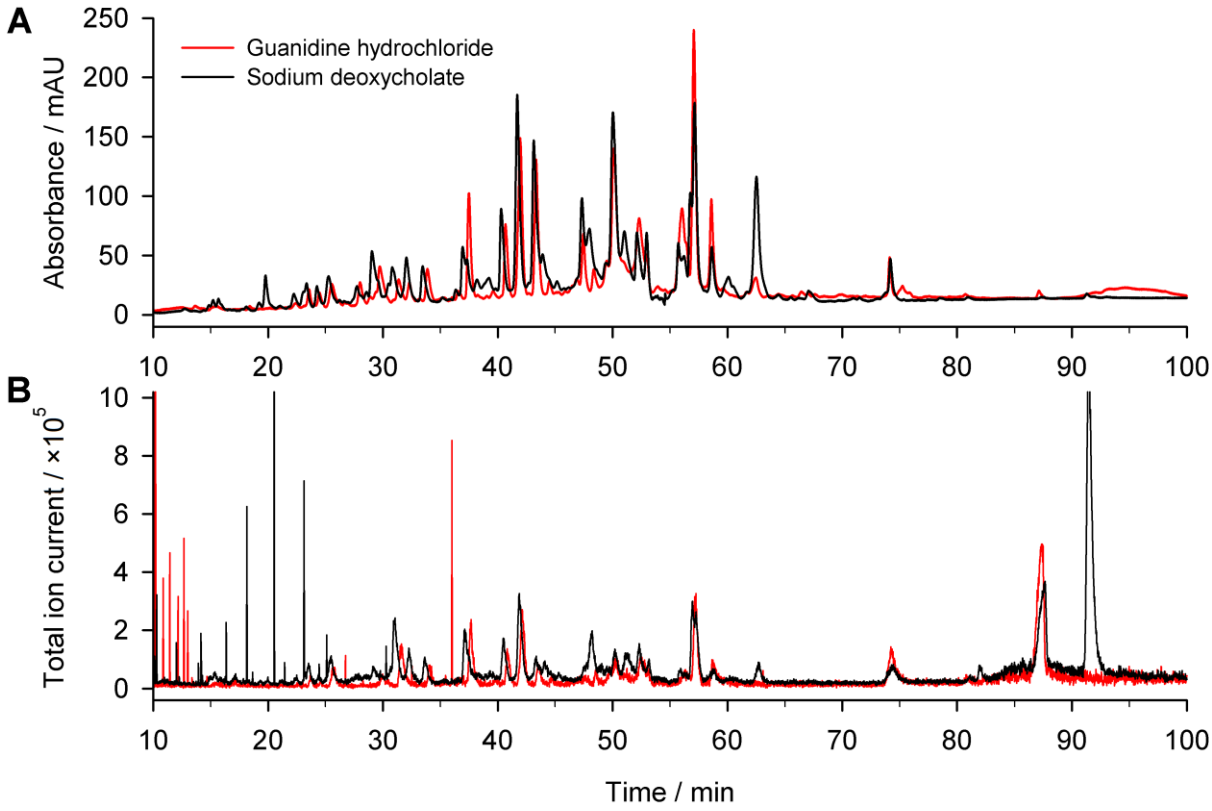


Figure 5.3: 214-nm absorbance spectra (A) and total ion current chromatograms (B) of GLIC peptides from chymotrypsin-digestion after denaturation of the protein with either guanidine hydrochloride (red) or sodium deoxycholate (black).

The deoxycholate protocol also identified more peptides from the transmembrane domain than the GnHCl protocol, with 65 amino acids identified instead of just 26. This equates to a total amino acid coverage for GLIC of 80 %, which includes 97 % of the extracellular domain and 53 % of the transmembrane domain. This was deemed to be an adequate coverage of GLIC so the photolabelling experiments proceeded with the deoxycholate denaturation protocol.

However, the peptide with a MW of 1015.5 Da that was predicted to be $_{252}$MTYTGAIIF$_{260}$ or $_{253}$TYTGAIIFM$_{261}$ and the peptide with a MW of 1450.8 Da that was predicted to be $_{255}$TGAIIFMIYLFY$_{266}$ were later sequenced by MS-MS, presented in the next section, and confirmed to be $_{134}$SVDTRNIVL$_{142}$ and $_{175}$ALEDRLESKLDY$_{186}$ respectively. With this information in mind, the amino acid coverage was recalculated to be 77 %, including only 45 % of the transmembrane region. However, this could still be an overestimate of the amino acid coverage as other peptides might also have been mismatched to a MW identified from single-stage mass spectra.



Figure 5.4: Schematic showing the amino acid coverage of GLIC from treatment B. Each yellow bar represents the elution of a peptide that matches the MW of the peptide sequence above it and all amino acid residues that are part of an identified eluted peptide are highlighted in red.

5.2 Photolabelling GLIC

Due to the low quantity of GLIC available, instead of using 10 mg of protein per sample for the labelling experiment and having redundant reaction mixtures as in the HSA experiments, between 100 µg and 150 µg of GLIC was used per sample and experiments where the samples were not exposed to light were not performed. The three main samples are listed below; in some experiments, samples containing just ethanol or the photolabels in the absence of GLIC were also used.

1. GLIC + ethanol + hν
2. GLIC + pPD + hν
3. GLIC + dPD + hν

The set plan was to run half of the peptide samples from each GLIC photolabelling experiment on the HPLC-coupled single-quadrupole mass spectrometer to compare the UV spectra and total ion

currents between the unlabelled and labelled samples. The presence of any new peaks or the reduction of native peptide peaks upon labelling was then examined by studying the averaged mass spectra spanning these peaks to determine the *m*/*z* of the ions that had contributed to the changes in UV absorbance or ionisation intensities. The other half of the samples were run only on the HPLC, bypassing the single-quadrupole mass spectrometer and the labelled peptides were collected with the automated fraction collector for further analysis with MS-MS to sequence them. However, due to blockage of the chromatography column on a few occasions, the full sets of data for the unlabelled and two labelled samples were not always obtained.

Three different post-labelling protein treatments were carried out: treatments B and C involved the denaturation of the protein by the surfactant SDC in the absence of any chaotropic agents such as urea and GnHCl having been exposed to light with or without the photolabel at pH 8 and pH 4 respectively, followed by digestion with chymotrypsin. Treatment D involved the digestion of GLIC at pH 8 post-light exposure by the endoproteinase Lys-C in 8 M urea and 0.1 % RapiGest followed by dilution to 1M urea before further digestion with chymotrypsin. The following subsections discuss the results from the labelling experiments of GLIC from these different post-labelling treatments of the proteins and subsequently their peptides.

5.2.1 Treatment B with sodium deoxycholate and chymotrypsin post-photolabelling at pH 8

GLIC post-labelling treatment B consists of denaturing the protein in 2 % SDC, followed by diluting the detergent concentration two-fold to allow digestion of the protein with chymotrypsin. Figure 5.5 shows the TIC chromatograms and absorbance spectra at 214 and 280 nm of chymotryptic peptides of GLIC from samples that were exposed to light for two minutes in the presence of ethanol, pPD and dPD.

Differences between the unlabelled and labelled GLIC samples in the TIC chromatograms can only be seen at 51.5 minutes, where reductions in intensity currents are seen for both labelled samples, and at 63.5 minutes, where increases are seen instead. However, the increases at 63.5 minutes are also seen in the control sample of photolysing the photolabel in the absence of GLIC (data not shown), therefore the increases in ionisation at that time are likely to correspond to

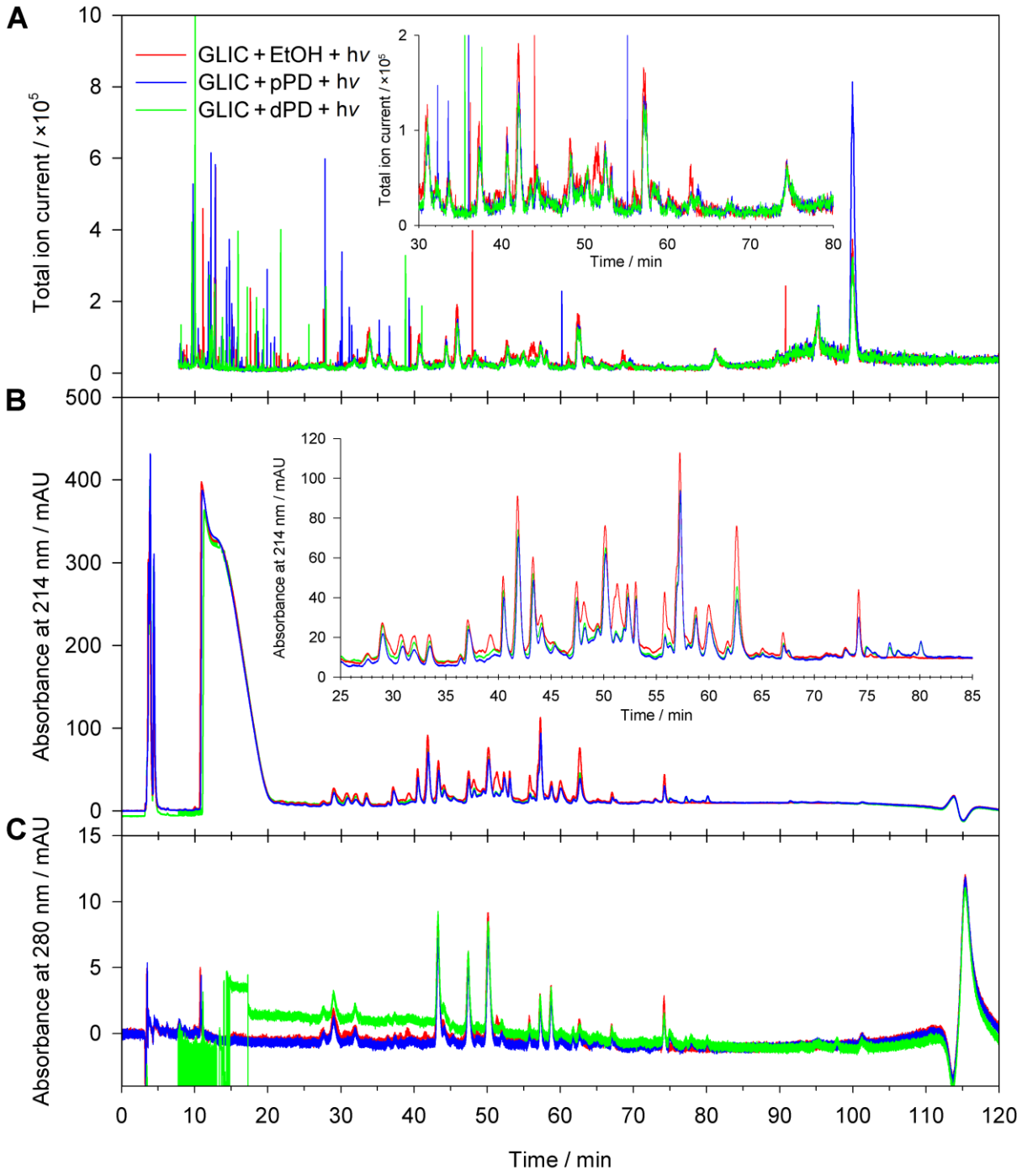


Figure 5.5: Total ion current (A), 214-nm (B) and 280-nm absorbance (C) chromatograms of GLIC that was exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green) prior to post-labelling treatment B as outlined in section 2.4.2.2. The peptides were separated by eluting from a reverse-phase liquid chromatography column with a gradient elution of ACN in eluent containing TFA. Insets in panels A and B show enlarged chromatograms.

photolabel by-products rather than labelled peptides. Despite the lack of new peaks in the TIC graphs of the labelled samples, a comparison of the absorbance spectra at both 214 and 280 nm between the labelled and unlabelled GLIC samples suggests that labelling of the protein has

occurred. New peaks in the labelled samples can clearly be seen between 74.5 and 81.0 minutes in the 214-nm (Figure 5.5B) and 280-nm spectra (Figure 5.5C).

Through analysing the major peptide components in all the TIC peaks of the pPD- and dPD-labelled samples using the Deconvolution tool of *ChemStation*, peptides with molecular weights that differed by 1 ± 0.1 Da between the two labelled samples were investigated further. Three such labelled peptides were identified with molecular weights of 1161.1 and 1162.1, 1166.8 and 1167.8 and 1908.0 and 1909.0 Da (Figure 5.6). These were predicted to be $_{166}$TAVVKPANF$_{174}$, $_{175}$ALEDRLESKLDY$_{186}$ or $_{255}$TGAIIFMITLFY$_{266}$, and $_{175}$ALEDRLESKLDYQL$_{188}$ respectively, as these are the four chymotryptic peptides identified from *in silico* digestion of GLIC that match the MW requirement of the labelled peptides less the MW of the bound photolabel (216.1 Da for the D_0 photofol and 217.1 Da for the D_1 photofol).

The labelled peptides whose D_0-labelled isotopes have molecular weights of 1161.1, 1666.8 and 1908.0 Da were collected at their corresponding retention times using the automated fraction collector coupled to the HPLC column and were confirmed to be $_{166}$TAVVKPANF$_{174}$, $_{175}$ALEDRLESKLDY$_{186}$ and $_{175}$ALEDRLESKLDYQL$_{188}$ respectively by MS-MS. The positions of the bound photolabels in all three peptides were identified to be on their lysine residues. The evidence is shown in Figure 5.7 for $_{166}$TAVVKPANF$_{174}$ by the presence of labelled b_5 and y_5 ions, denoted b_5^ and y_5^, and in Figure 5.8 for $_{175}$ALEDRLESKLDY$_{186}$ by the presence of b_9^ and y_4^ ions. Additional evidence showing the lysine residues of these peptides being labelled is provided by peaks representing labelled internal lysine ions, K_{im}*^ and K_{aa}^ at 300 and 345 Da respectively (see Figure 4.18 for structures). The EPI spectra for the pPD- and dPD-labelled $_{175}$ALEDRLESKLDYQL$_{188}$ are not shown but the peptide sequence was confirmed by the unlabelled ion fragments $b_{5–12}$ and $y_{2–5}$ and the location of the label was deduced from the observation of both b_9^ and y_6^ ions as well as the characteristic labelled internal lysine ions.

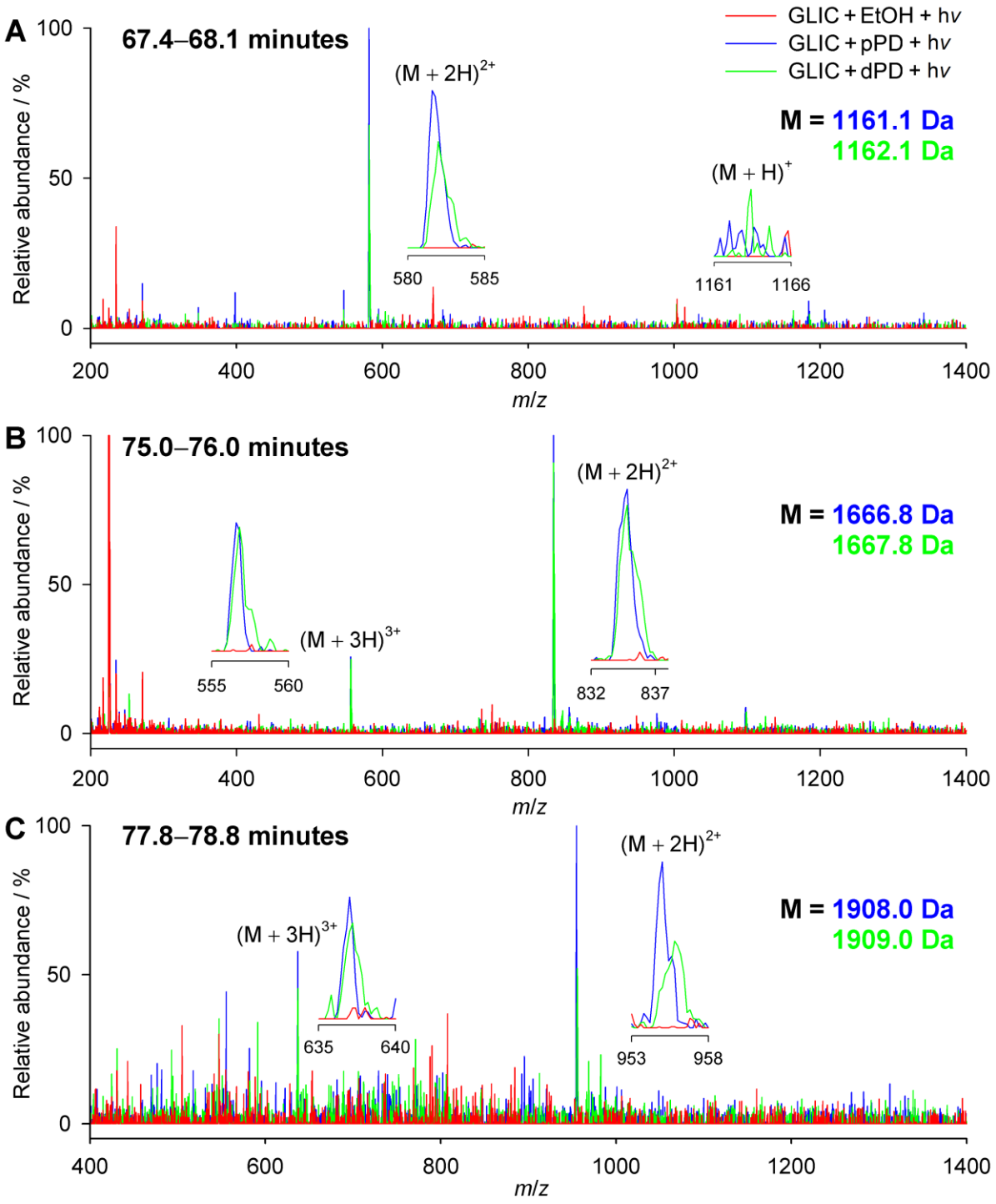


Figure 5.6: Averaged mass spectra overlays at three retention time ranges, when labelled GLIC peptides from the photolabelling treatment B experiment were eluted for the GLIC proteins exposed to light in the presence of ethanol (red), pPD (blue) and dPD (green). The most abundant labelled peptides in the pPD- and dPD-labelled samples have molecular weights of 1161.1 and 1162.1 Da (A), 1166.8 and 1167.8 Da (B) and 1908.0 and 1909.0 Da (C) respectively. The ionisation current is normalised to the most ionising peak from the three traces within the *m/z* range displayed for each elution time range. The molecular weights (M) and ion charged states are noted by their corresponding peaks.

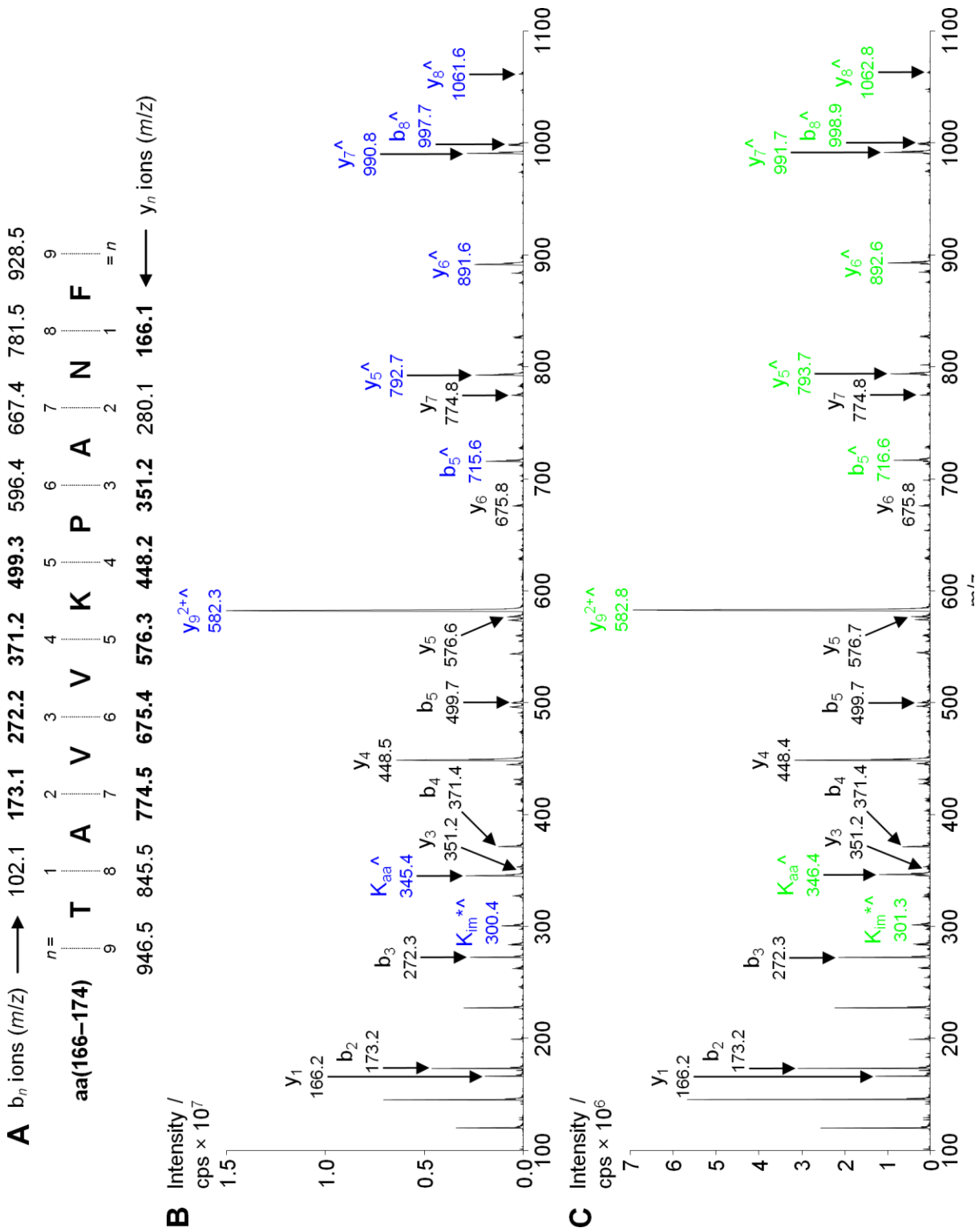


Figure 5.7: Enhanced product ion mass spectra for the collision-induced dissociation of parent ions of 582.3 and 582.8 *m/z* confirming the sequence of this labelled peptide as being $_{166}$TAVVKPANF$_{174}$. Schematic (A) summarises the b_n and y_n ions observed (in bold). Spectra for the peptides labelled with pPD (B) and dPD (C) both indicate labelling on lysine-170 due to the presence of b_5^ and y_5^ ions. Peaks corresponding to labelled fragments are marked with an additional caret (blue for fragments labelled with pPD and green for those labelled with dPD). K_{im}*^ denotes the labelled immonium-less-ammonia ion of lysine-170. K_{aa}^ denotes the labelled amino acylium ion of lysine-170.

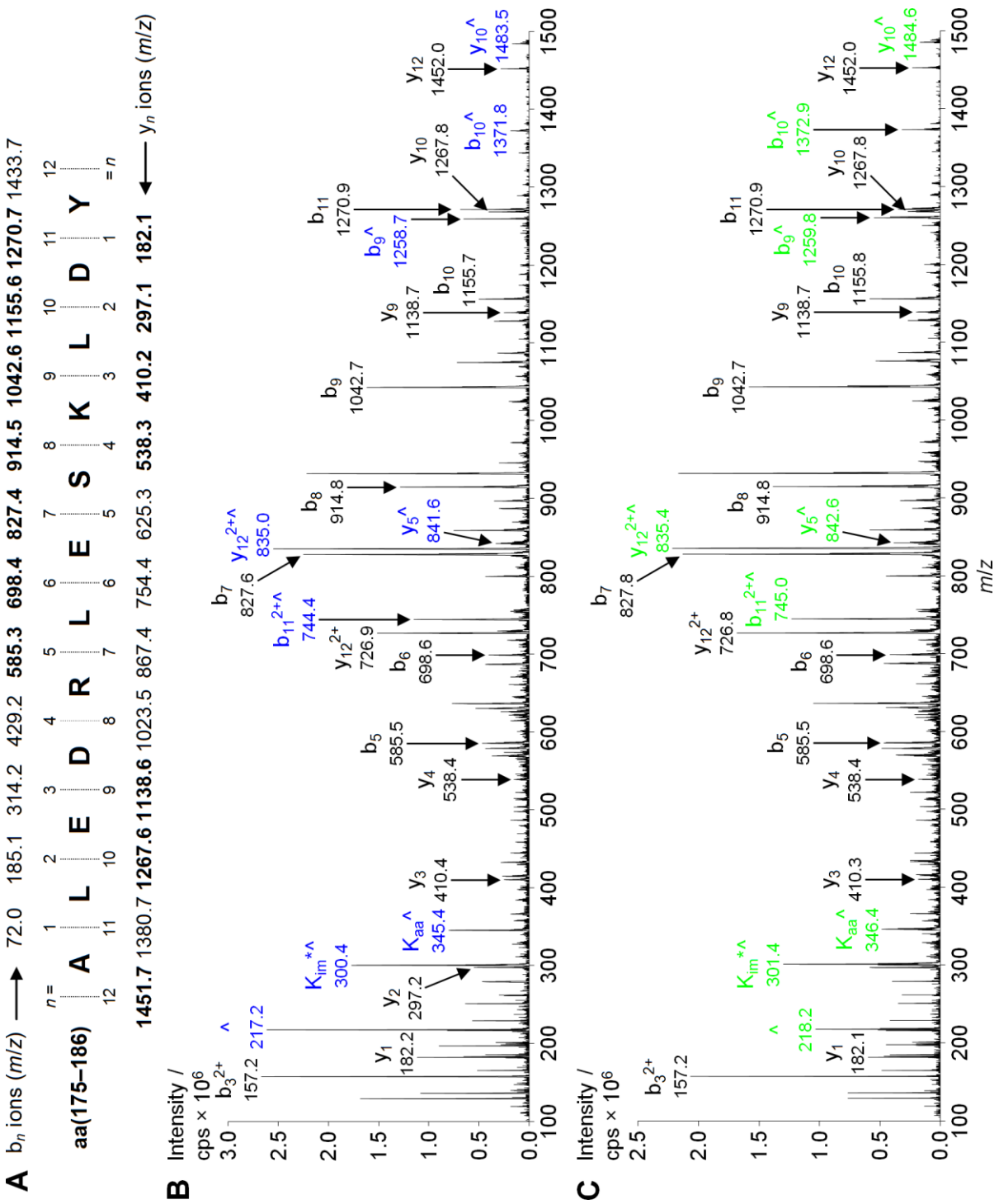


Figure 5.8: Enhanced product ion mass spectra for the fragmentation of the parent ions 834.4 and 834.7 *m/z* confirming the sequence of this labelled peptide as being $_{175}$ALEDRLESKLDY$_{186}$. Schematic (A) summarises the b_n and y_n ions observed (in bold). Spectra for the peptides labelled with pPD (B) and dPD (C) both indicate labelling on lysine-183 due to the presence of b_9^, y_4^, K_{im}*^ and K_{aa}^ ions. The singly charged photolabel products of 217.2 and 218.2 *m/z* detached from the pPD- and dPD-labelled peptides respectively are marked by a caret and peaks corresponding to labelled fragments are marked with an additional caret (blue for fragments labelled with D_0 photofol in the pPD sample and green for those labelled D_1 photofol in the dPD sample).

Sequencing two unlabelled peptides

As an aside, two fractions in the unlabelled sample (GLIC + EtOH) predicted to contain the peptides $_{252}$MTYTGAIIF$_{260}$ or $_{253}$TYTGAIIFM$_{261}$ at 48.3 minutes and $_{175}$ALEDRLESKLDY$_{186}$ or $_{255}$TGAIIFMITLFY$_{266}$ eluted at 51.5 minutes were also collected for sequencing with MS-MS as these peptides have the same molecular weights. One of the peptides of the latter unlabelled pair has contributed to the reduction in TIC and 214-nm absorbance at the retention time of 51.5 minutes (Figure 5.5). The EPI mass spectra of the unlabelled peptides of molecular weights 1015.5 Da at the retention time of 48.3 minutes and 1450.8 Da at 51.5 minutes are shown in Figure 5.9. The sequences were identified to be $_{134}$SVDTRNIVL$_{142}$ and $_{175}$ALEDRLESKLDY$_{186}$ respectively based on the presence of many of their b_n and y_n ions. $_{175}$ALEDRLESKLDY$_{186}$ is one of the two chymotryptic peptides that were predicted based on their molecular weights and being cleaved after either an aromatic residue, leucine, methionine or histidine (SIB ExPASy Bioinformatics Resource Portal, 2011).

The sequence for the peptide eluted at 48.3 minutes (Figure 5.9B) was expected to be $_{252}$MTYTGAIIF$_{260}$ or $_{253}$TYTGAIIFM$_{261}$. The peptide was sequenced due to the inability to distinguish between these sequences with the same monoisotopic MW using single-stage mass spectrometry data and because these peptides were of interest as they are both good candidates for being labelled based on the crystallographic data of GLIC with propofol. However, a selection of their common CID ions could not be located on the mass spectrum in Figure 5.9B. Through identifying differences between substantial peaks on the mass spectrum that match the weights of amino acids and searching for strings of consecutive amino acids in the GLIC sequence, the peptide was confirmed to be $_{134}$SVDTRNIVL$_{142}$ that arose from cleavage after arginine and leucine. This shows that within ten hours of chymotrypsin-digestion at 25 °C, enzymatic cleavage is not restricted to just aromatic residues, leucines, histidines and methionines. This finding suggests that the originally calculated amino acid coverage percentage of 80 % was incorrect due to predicting the wrong peptide sequences to match the molecular weights identified from the single-stage mass spectrometry data (section 5.1).

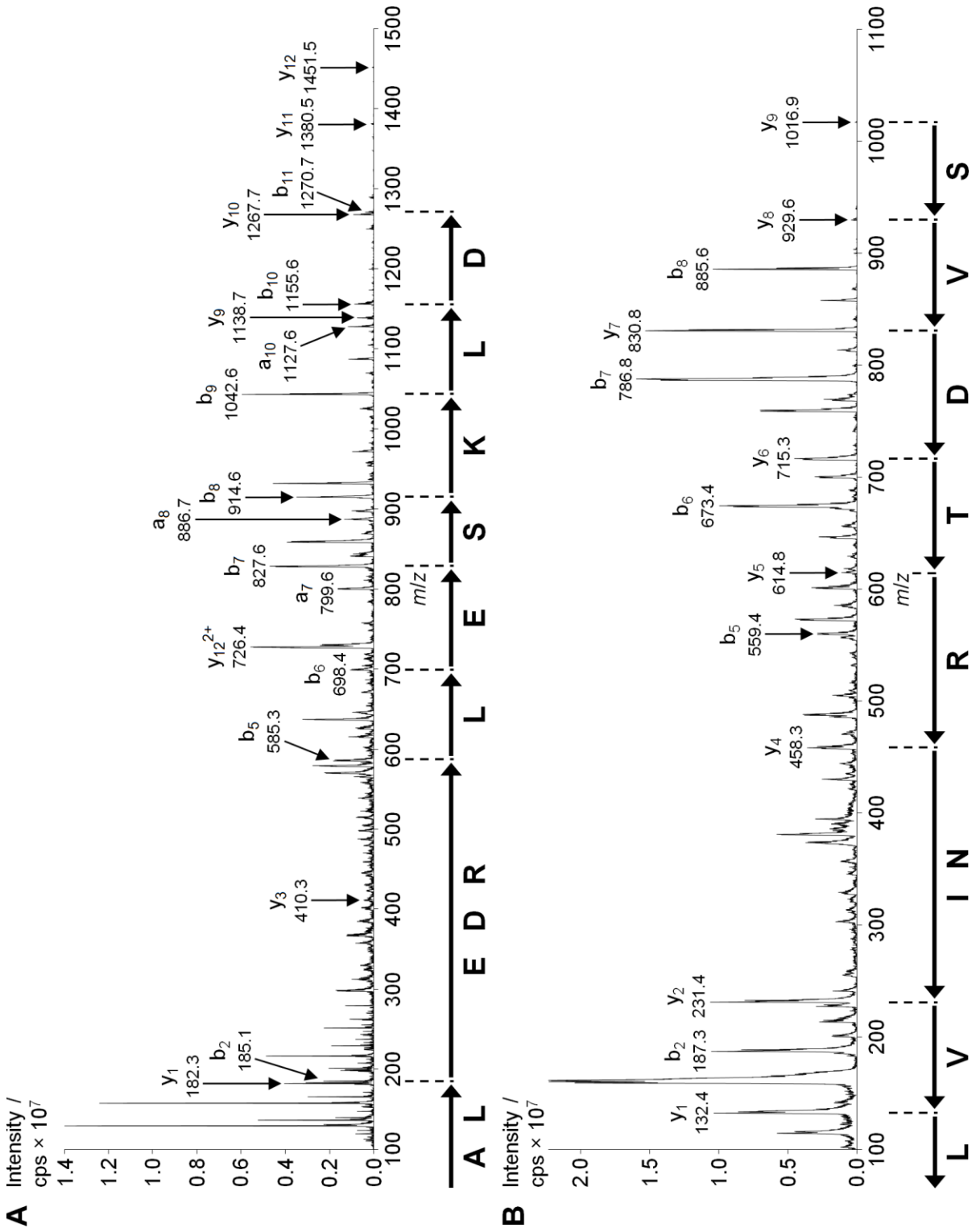


Figure 5.9: Enhanced product ion mass spectra of the unlabelled peptides $_{134}$SVDTRNIVL$_{142}$ and $_{175}$ALEDRLESKLDY$_{186}$ with their sequences confirmed by presence of many b_n and y_n-series CID ions.

Separating peptides in solvents containing formic acid

When this post-labelling protein denaturation and digestion protocol, treatment B, was repeated twice, the peptide samples obtained were separated by solvents containing formic acid (FA) instead of with TFA, as the latter suppresses ionisation of the peptides (García *et al.*, 2002), reducing the signal-to-noise ratio from the mass spectrometer. Figure 5.10 shows the TIC, 214- and 280-nm absorbance chromatograms of the unlabelled and labelled peptide samples eluted by solvents containing FA after treatment B. The TIC signals of the GLIC peptides were increased by about ten-fold by replacing TFA in the HPLC solvent with FA. Five clear peaks occurring only in the labelled samples, marked A–E, are seen across all three monitored parameters at 84.0, 94.5, 96.5, 98.3 and 100.4 minutes.

The averaged mass spectra around peaks C and D, with retention times of 96.0–97.0 and 97.8–98.8 minutes respectively, are displayed in Figure 5.11. The peptides with a MW of 1450.7 Da have been shown to be labelled by D_0 photolabels (blue peaks) and by a mixture of D_0, D_1 and D_2 photolabels (green peaks) in Figure 5.11A. These increases in ionisation intensities contribute to the TIC peaks of the labelled samples marked C in Figure 5.10.

Figure 5.11B shows three long labelled peptides with molecular weights of 2943.2, 2965.1 and 2981.0 Da being eluted at 97.8–98.8 minutes when labelled by D_0 photofol of the pPD-labelled sample and with molecular weights of one additional Dalton when labelled by the predominant D_1 photolabel of the dPD-labelled sample. All of these singly and doubly charged-ion peaks contribute to the new TIC peak marked D upon photolabelling GLIC in Figure 5.10.

The averaged mass spectra around the other three retention times of 84.0, 94.5 and 100.4 minutes were also studied to identify the m/z of the most abundant ions of the labelled peptides so that they could be selected for fragmentation during MS-MS. At 84.0 minutes (peak A in Figure 5.10), besides the photolabel products that were observed in the control experiment of photolysing the photolabel in the absence of GLIC and proceeding with the protein denaturation and digestion steps, labelled peptides with molecular weights of 1161.1 Da for the pPD-labelled sample and 1162.1 Da for the dPD-labelled sample were also observed. This peptide was confirmed to be $_{166}$TAVVKPANF$_{174}$ with the photolabel bound to lysine-170. Representative EPI spectra for this labelled peptide were shown in Figure 5.7. Peaks C and E that appeared upon labelling arose

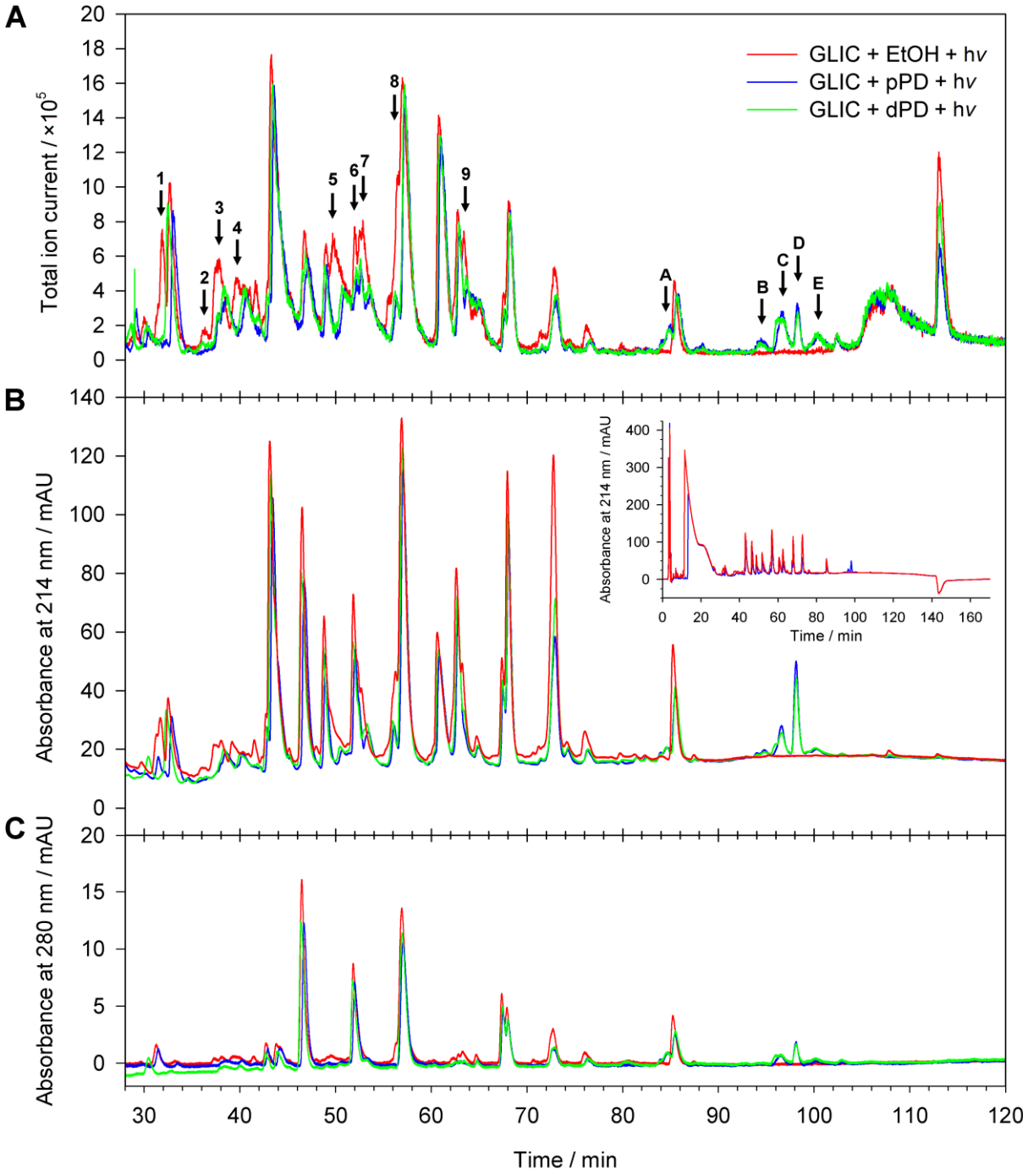


Figure 5.10: Total ion current (A), and 214- (B) and 280-nm absorbance (C) chromatograms of GLIC from the second treatment B photolabelling experiment with the peptide samples separated by eluents containing FA instead of TFA for GLIC protein that was exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green). Arrows 1–9 mark peaks that have disappeared or reduced in intensity, and arrows A–E mark the peaks that have appeared or increased in intensity, at both the TIC and absorbance at 214 nm chromatograms, when GLIC was exposed to light in the presence of photolabels.

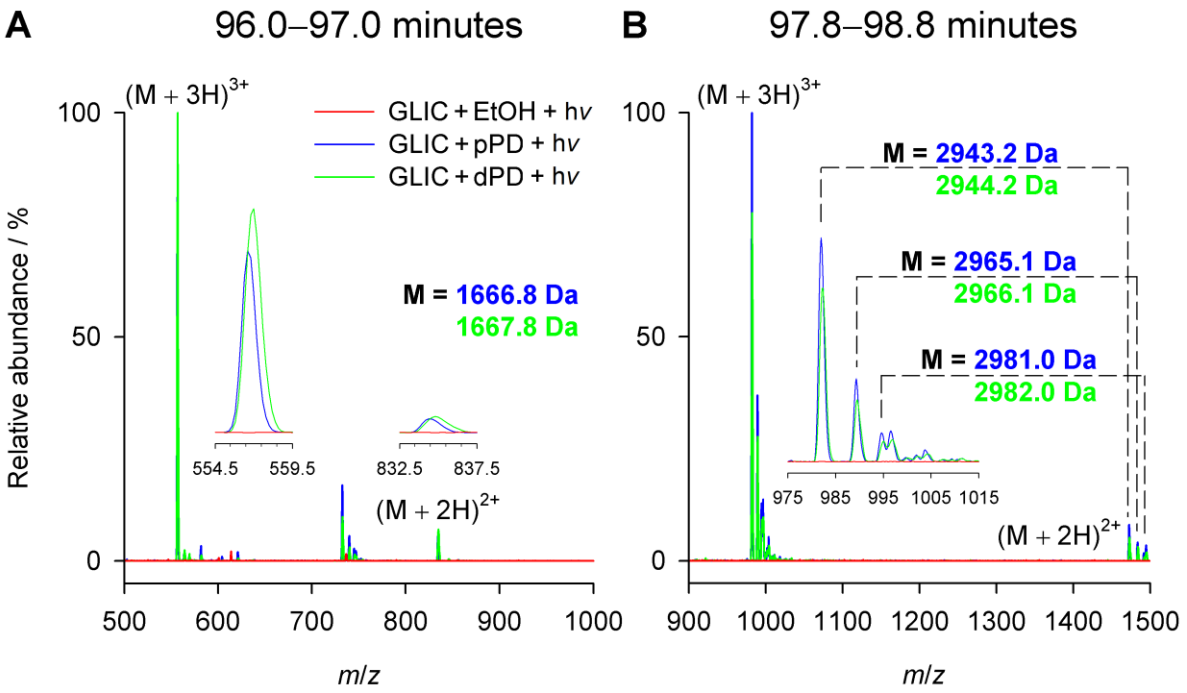


Figure 5.11: Averaged mass spectra overlays at 96.0–97.0 minutes (A) and 97.8–98.8 minutes (B) for peaks C and D of Figure 5.10 respectively, for the GLIC samples that was exposed to light in the presence of ethanol (red), pPD (blue) and dPD (green), when labelled peptides were eluted. (A) Labelled peptides with molecular weights (M) of 1666.8 or 1667.8 Da were identified based on presence of their doubly and triply charged ions. (B) Three sets of labelled peptides (labelled with pPD- or dPD sample) each with molecular weights of over 2900 Da were identified by their doubly and triply charged ions. The ionisation current is normalised to the most ionising peak from the three traces within the *m/z* range displayed for each of the two elution time ranges.

from the other two labelled peptides that have been previously identified, $_{175}$ALEDRLESKLDY$_{186}$ and $_{175}$ALEDRLESKLDYQL$_{188}$. Representative EPI spectra of $_{175}$ALEDRLESKLDY$_{186}$ were shown in Figure 5.8 and the spectra of both peptides show that the photolabel is bound to the lysine at their ninth amino acids.

The averaged mass spectra across peak B in Figure 5.10 suggests that a peptide with a MW of 1172.6 Da had been labelled. Analysis of the EPI spectra fragmenting the doubly charged pPD- and dPD-labelled ions—that were collected over the duration of peak B—indicated once again that lysine-183 had been labelled. The presence of some b_n and y_n ions representative of the peptide $_{175}$ALEDRLESKL$_{184}$ was observed in the EPI spectra along with two internal lysine ions, K_{im}* and K_{aa}, both with a photolabel attached.

The overlays of mass spectra across the new peak D of Figure 5.10 for the unlabelled and labelled samples are shown in Figure 5.11B. EPI spectra were obtained for the doubly charged ions of each of the long labelled peptides that were identified but the quality of the spectra was too poor to determine their sequences. The reasons for this are two-fold. Firstly, the ionisation of these labelled peptides was not of high intensity, therefore they did not 'fly' well in the mass spectrometer; and secondly, the longer the peptide, the higher the possibility of an amino acid within the sequence undergoing modification, thus resulting in *m*/*z* peaks that deviate from the expected common CID ions, making it more difficult to carry out *de novo* sequencing.

The chromatogram overlays for the TIC and absorbance at 214 nm in Figure 5.10 also show numerous time points at which reductions of intensities are seen in the labelled samples compared to the unlabelled sample, these were marked as peaks 1–9. The reductions in peak intensities, which correspond to the decrease in concentration of peptides eluted at those times, are likely to be due to these peptides being labelled, resulting in heavier molecular weights and a change in polarity, thus causing their elution at different retention times.

The averaged mass spectra of the control sample spanning these peaks were investigated to identify the MW of the peptide that had reduced in concentration at each of these time points. Fractions containing these peptides were then collected and sequenced by MS-MS. Confirmed sequences along with their amino acid positions are shown beneath the bar chart in Figure 5.12 and the annotated EPI spectra with their observed common CID ions marked are displayed in Figure A.7–15 of the appendix. Figure 5.12 shows the percentage changes in ionisation of the peptides—in which their molecular weights had been identified—in the labelled peptide samples compared to the unlabelled sample. This was calculated by extracting ions having up to four charges of each of these peptides in the unlabelled sample and comparing the combined area under the extracted ion current (XIC) chromatograms with that from the labelled sample using the following equation:

$$\text{Percentage change in ionisation (\%)} = 100 \times \frac{\text{XIC area}_{\text{labelled sample}} - \text{XIC area}_{\text{control sample}}}{\text{XIC area}_{\text{control sample}}}$$

This analysis was also carried out on three randomly selected molecular weights at different time points. The percentage changes seen in the XIC areas of the labelled compared to the unlabelled

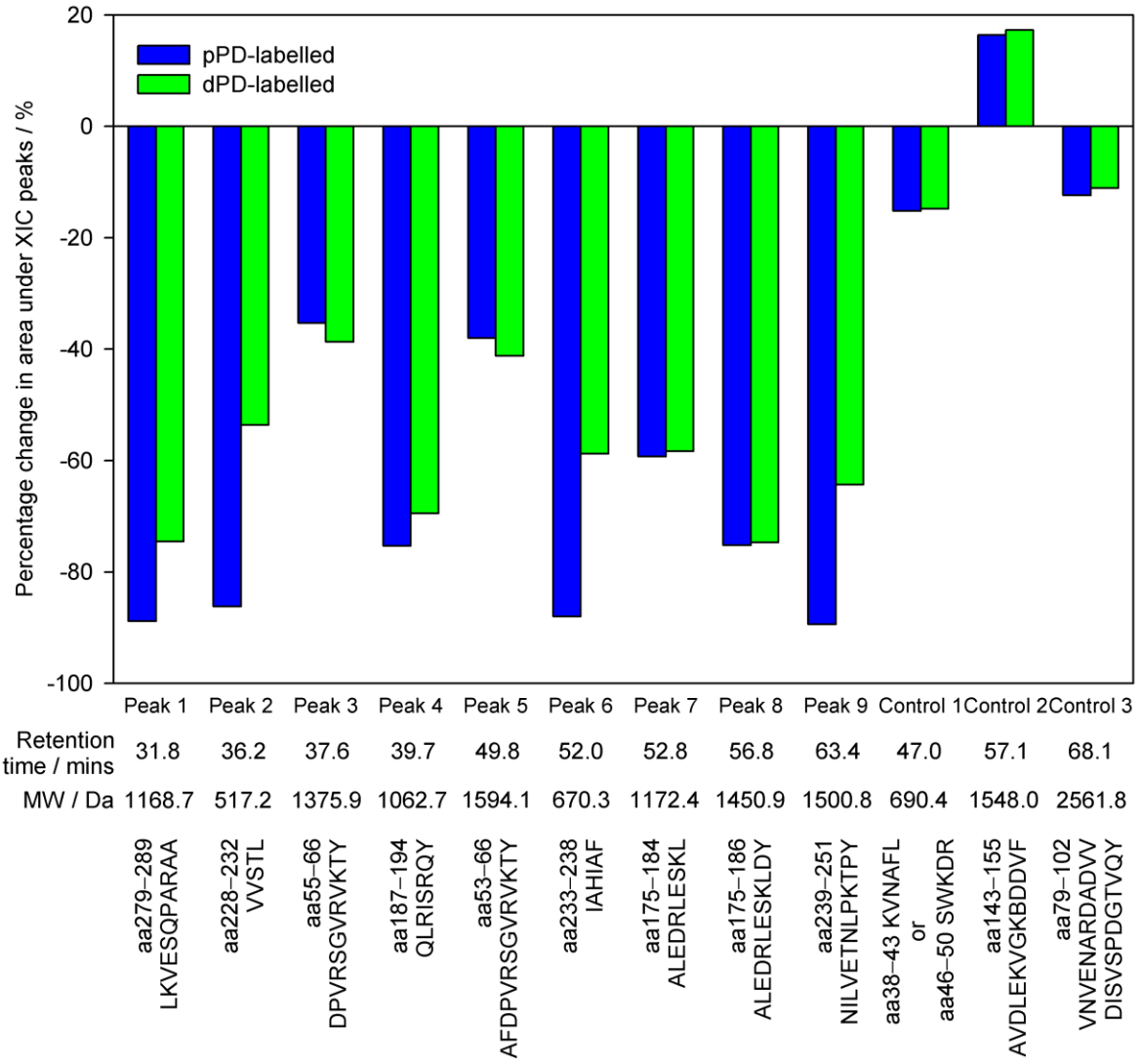


Figure 5.12: Bar chart showing the percentage changes in ionisation of twelve peptides from the second GLIC treatment B photolabelling experiment. The area under the extracted ion current chromatograms of these peptides, eluted during peaks 1–9 as indicated on Figure 5.10 and three randomly selected peptides with molecular weights of 690.4, 1548.0 and 2561.8 Da were compared between their control sample to those of the pPD- (blue) and dPD-labelled samples (green)

samples of these three peptides were within ±20 %, whilst for those molecular weights that were observed to decrease in ionisation intensity upon labelling, the percentage decrease in the XIC areas was quantified to be up to 90 %.

Although it is clear that these peptides have decreased in concentration upon labelling, it was not possible to locate most of their labelled counterparts by extracting the ions of their photolabel-bound molecular weights. The only labelled counterparts of these unlabelled peptides—with a reduction in concentration upon photolabelling—identified were the labelled peptides of

$_{175}$ALEDRLESKLDY$_{186}$ eluted at peak 7 and $_{175}$ALEDRLESKLDYQL$_{188}$ at peak 8. The ionisation of their unlabelled parent peptides reduced by 60 % and 75 % respectively in the labelled samples compared to the unlabelled sample.

The TIC and absorbance spectra from two separate experiments with the same GLIC treatment protocol differ slightly. Figure 5.13 shows the TIC data from these separate experiments. The major difference that can be observed is that only two new peaks, equivalent to peaks C and

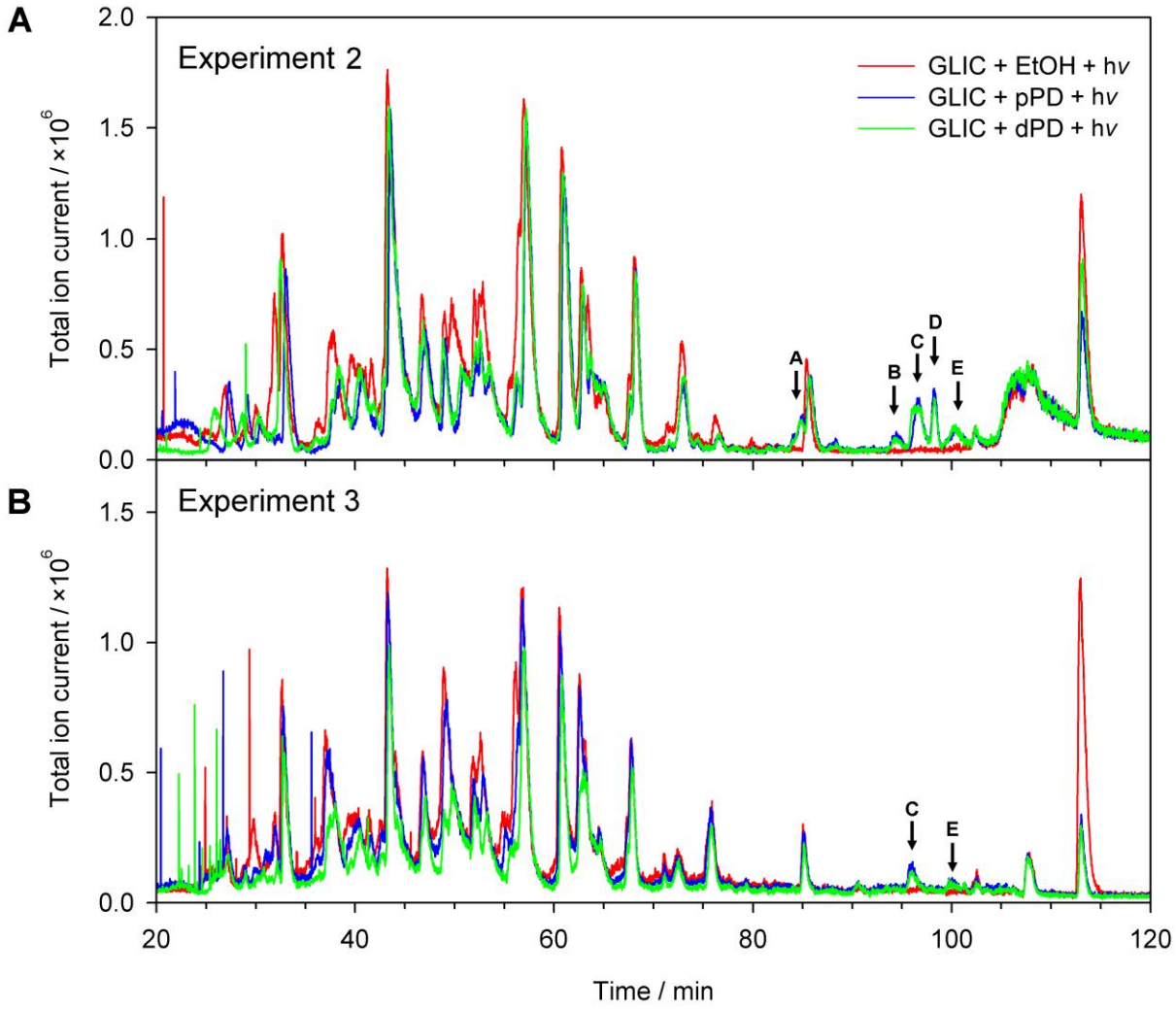


Figure 5.13: Total ion current of GLIC peptide samples from the second (A) and third (B) GLIC treatment B photolabelling experiments. Digested GLIC peptides from intact proteins that were exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green) were separated by HPLC-MS with an increasing ACN gradient elution of eluents containing 0.06 % formic acid. Arrows A–E in panel A mark the peaks that have appeared or increased in total ion current intensity, when GLIC was exposed to light in the presence of photolabels in experiment 2. Arrows C and E in panel B mark the appearance of two peaks—eluting at the same retention times as their corresponding marked peaks from panel A—when GLIC was photolabelled.

E of experiment 2 (Figure 5.10 and Figure 5.13A), were seen in the latter experiment (Figure 5.13B). These correspond to lysine-183 being labelled in the peptides $_{175}$ALEDRLESKLDY$_{186}$ and $_{175}$ALEDRLESKLDYQL$_{188}$ respectively. The long labelled peptides eluted at 97.8–98.8 minutes (peak D) of experiment 2, for which good quality EPI spectra were not obtained, appear to be absent in the peptide mixture of experiment 3, as with the labelled peptides of $_{166}$TAVVKPANF$_{174}$ and $_{175}$ALEDRLESKL$_{184}$. It should be noted that the TIC and UV absorbance of the dPD-labelled sample in experiment 3 are less intense than the pPD sample (Figure 5.13B). Comparing the sizes of the 214-nm absorbance peaks between the two labelled samples—which should overlap each other perfectly—suggests that the peptide concentration of the dPD-labelled sample is 25 % less than that of the pPD-labelled sample.

The nine previously identified peaks that have a reduction in TIC and 214-nm absorbance intensities upon labelling from experiment 2 were investigated further for the experiment 3 data. The areas under their XIC graphs were compared to their respective areas for the labelled samples and the results are summarised in Figure 5.14. As mentioned above, the peptide concentration of the dPD-labelled sample appears to be about 25 % less than the pPD-labelled sample, so to compensate for the more diluted peptide sample, the percentage change in XIC peak area for the dPD-labelled sample against the unlabelled sample was calculated by scaling up the combined XIC peaks area of the dPD-labelled sample by 1⅓-fold. With the scale-up, the percentage changes for the two labelled samples differ by less.

Ignoring the non-normalised dPD-labelled sample data in Figure 5.14 provides similar results to those from experiment 2 that are summarised in Figure 5.12. The three randomly selected molecular weights of 690.4, 1548.0 and 2561.8 Da all have a combined XIC area of their ions with between one and four charges in the labelled samples differing by ±20 % to that of the unlabelled sample. Peaks 3 and 5, where decreases of about 37 % compared to their respective control samples were seen in experiment 2 (Figure 5.12), only have a reduction of between five and 15 % from the XIC area of the unlabelled sample in experiment 3. This prompts the questioning of whether, upon labelling, the decrease in concentration of these two peptides, $_{55}$DPVRSGVRVKTY$_{66}$ and $_{53}$AFDPVRSGVRVKTY$_{66}$, observed in experiment 2 is significant.

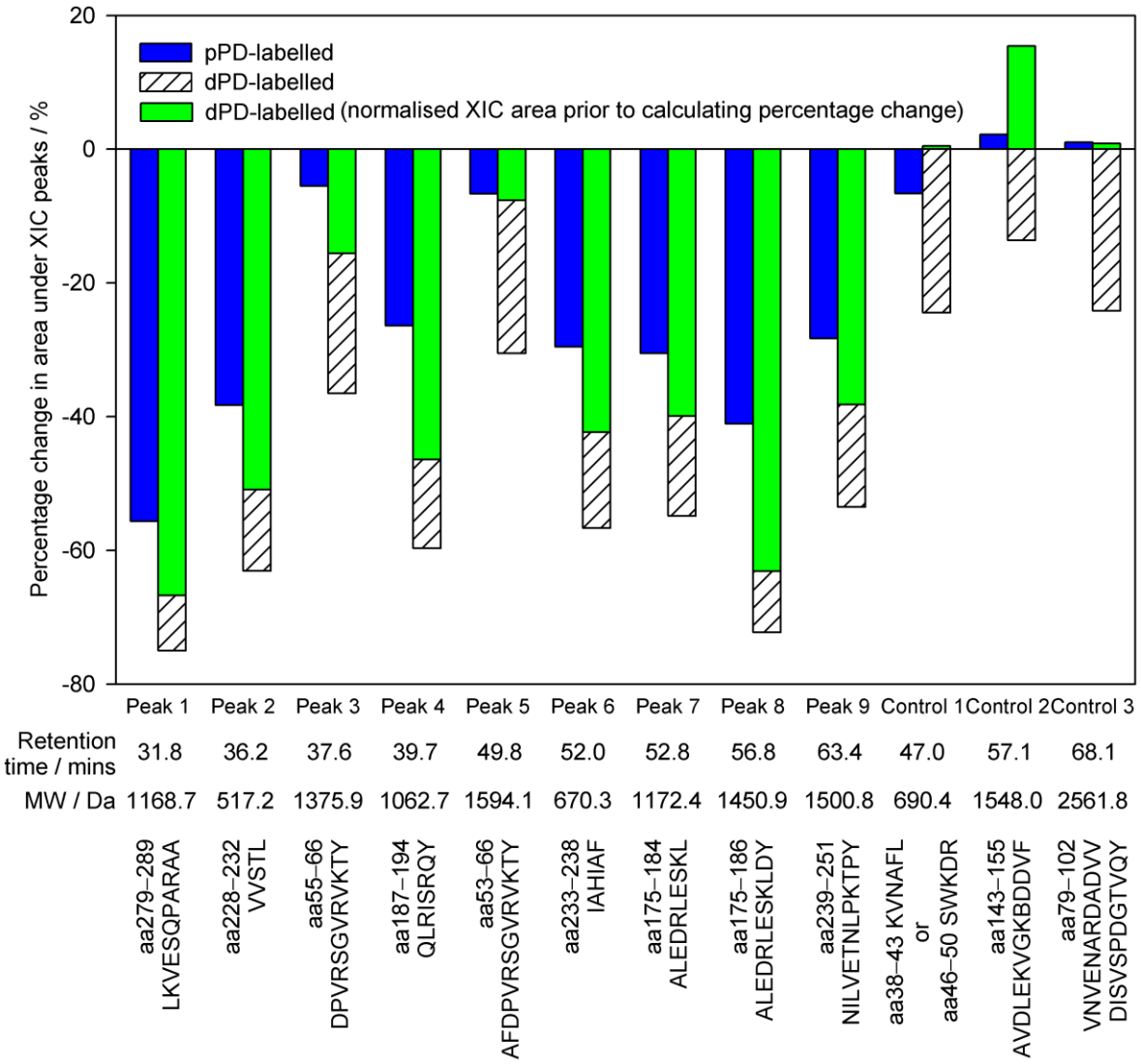


Figure 5.14: Bar chart showing percentage changes in ionisation of twelve peptides from the third GLIC treatment B photolabelling experiment. The percentage changes of the area under the XIC chromatograms from the control sample to that of the labelled samples of these peptides (eluted during peaks 1–9 as indicated on Figure 5.10 and also of the three selected peptides for the same analysis of the data of experiment two) are shown in blue for the pPD-labelled sample. The percentage changes against the dPD-labelled sample, calculated from non-normalised and normalised XIC areas, are shown by striped and green bars respectively.

However, other peptides that have undergone this analysis show a decrease in XIC peak area of 42.7 ± 3.4 (mean ± SE) % compared to their respective areas from the unlabelled sample. Peak 8, which consists of the peptide $_{175}$ALEDRLESKLDY$_{186}$ and has been confirmed to be labelled by sequencing data of its labelled peptide, showed reductions of 41 and 63 % for the pPD- and normalised dPD-labelled sample data respectively.

5.2.2 Treatment C with sodium deoxycholate and chymotrypsin post-photolabelling at pH 4

GLIC was photolabelled at pH 4 using a buffer system with citric acid and sodium hydrophosphate. Preliminary experiments were carried out to determine the composition of the acid and base to reach the required acidity of pH 4 and the quantity of sodium hydrophosphate to neutralise the mixture prior to adding the acid-cleaved detergent SDC. After incubating with SDC, the detergent concentration was reduced to allow chymotrypsin to have significant activity so as to digest the GLIC protein into manageable-sized peptides for HPLC-MS analysis. The full protocol is detailed in section 2.4.

Figure 5.15 shows the three parameters—TIC and UV absorbance at 214 and 280 nm—that were monitored as peptides were eluted from the HPLC column. The traces from the three samples overlap well with each other with only small differences observed between the unlabelled and labelled GLIC samples at a few time points. These differences are mainly reductions in TIC and 214-nm absorbance peak sizes in the labelled samples with the greatest difference seen at 50.3 minutes. Analysis of the averaged mass spectra around that time point showed that the main *m*/*z* values that decrease in intensity upon labelling are 399.6 and 532.4 *m*/*z*, corresponding to the decrease in concentration of a peptide with a MW of 1594.3 Da by assuming these *m*/*z* values are of triply and quadruply charged ions respectively. A peptide with a similar MW of 1594.1 Da was collected at a similar retention time of 49.8 minutes in the pH 8 GLIC labelling experiment and when subjected to MS-MS yielded the sequence $_{53}$AFDPVRSGVRVKTY$_{66}$ (see appendix Figure A.11 for the EPI mass spectrum).

The only increase in absorbance intensity in both the labelled samples compared to that of the unlabelled sample is observed at 90–93 minutes, where two clear peaks with absorbance of 0.8 mAU can be seen in the 280-nm absorbance spectra (Figure 5.15C) of the labelled samples. The averaged mass spectra at 90–93 minutes in Figure 5.16 shows the emergence of four *m*/*z* peaks in the labelled samples that were absent in the control sample. These four *m*/*z* are likely to be from the labelled peptides with a MW of 853.2 Da (for D_0-labelled or +1 Da for D_1-labelled), for which 427.6 and 854.2 *m*/*z* are its doubly and singly charged ions respectively, and with a MW

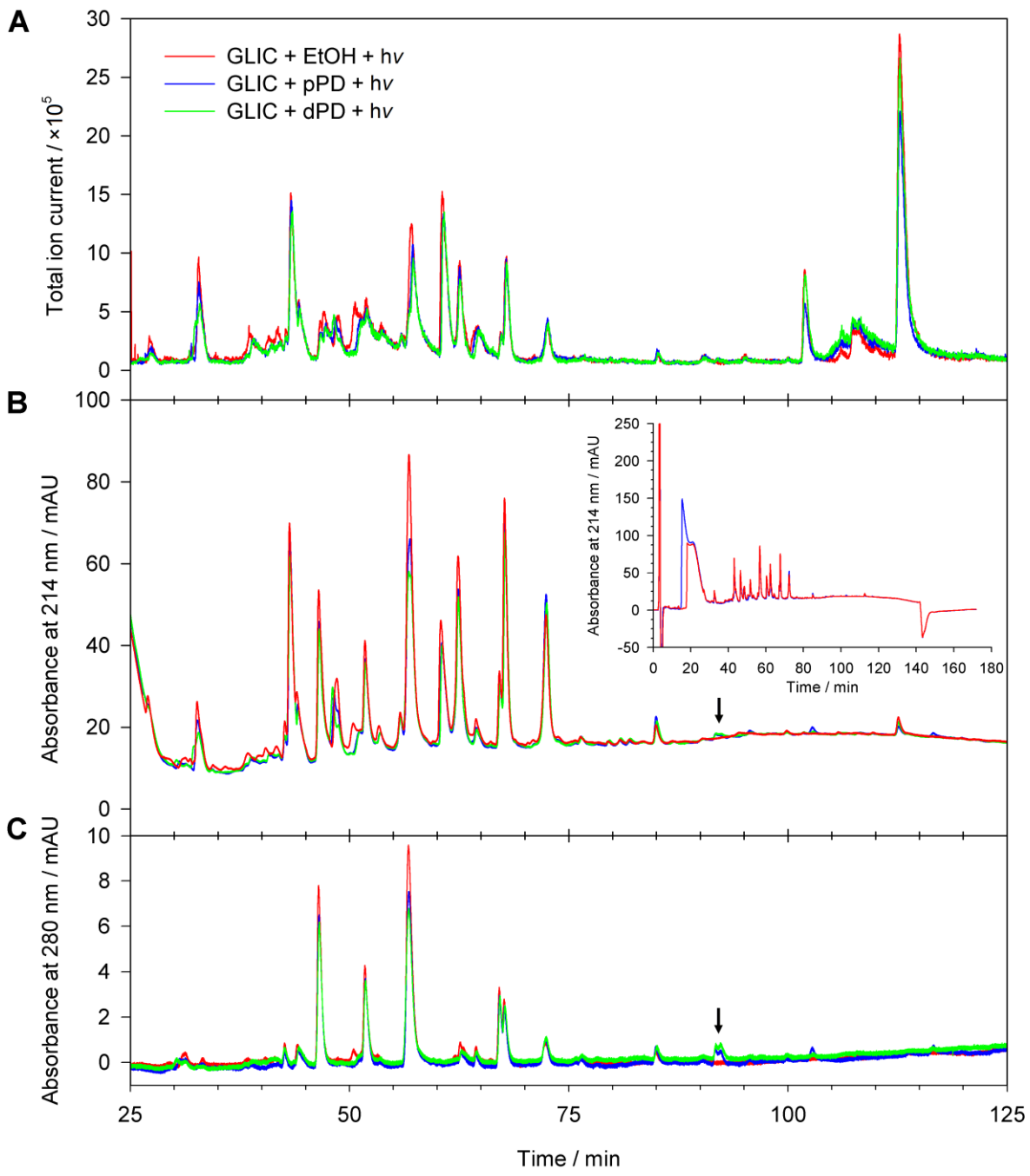


Figure 5.15: Chromatograms overlay of the TIC and absorbance at 214 and 280 nm of GLIC peptide samples from the GLIC treatment C photolabelling experiment. GLIC peptides from the intact protein that was exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green) at pH 4 prior to its denaturation and chymotrypsin digestion as outlined in section 2.4 were separated with a gradient elution of ACN from a reverse phase chromatography column using eluent containing 0.06 % formic acid. Inset in panel B shows the full 214-nm spectra for the total duration of the HPLC run. Arrows mark two adjacent peaks on the 214- and 280-nm chromatograms in the labelled samples that are absent in the unlabelled sample.

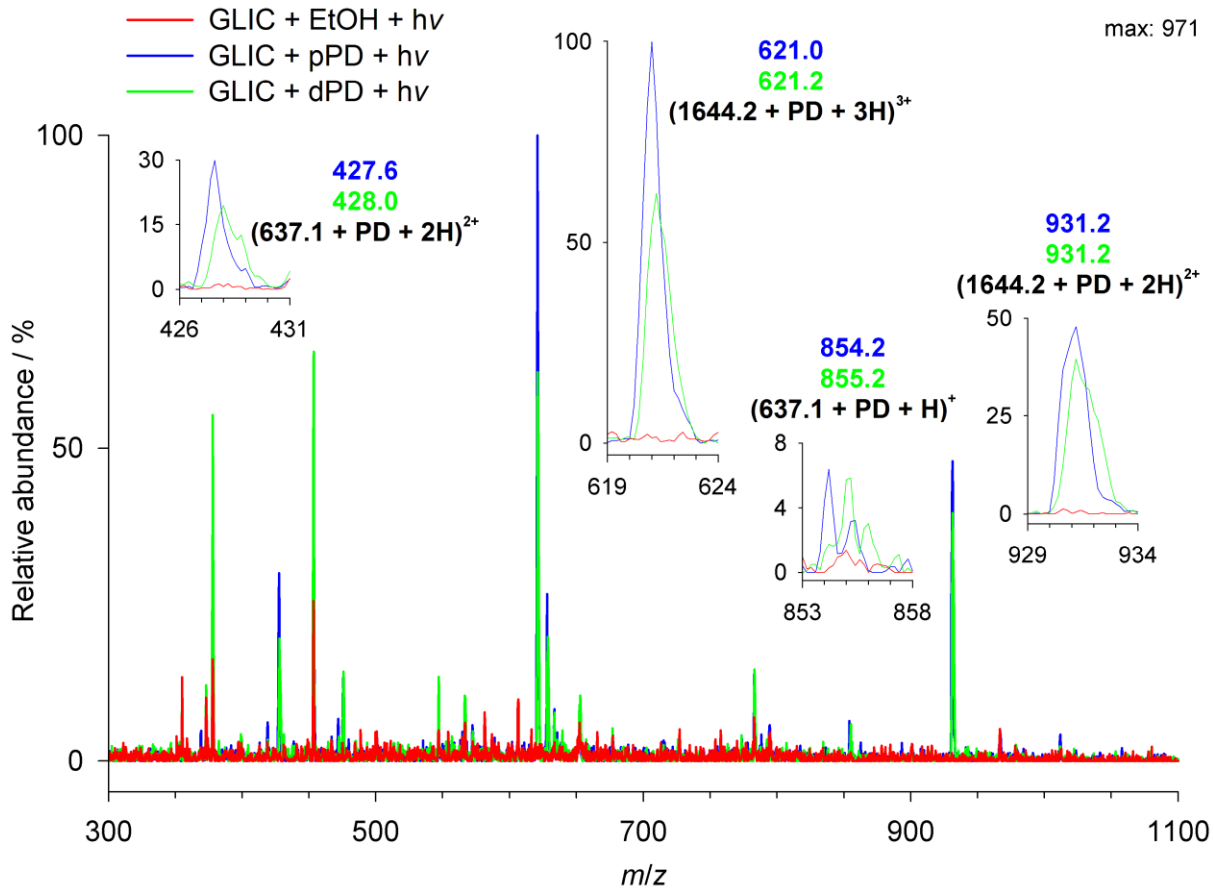


Figure 5.16: Averaged mass spectra overlay at 90–93 minutes for GLIC samples from the photolabelling treatment C experiment when GLIC was exposed to light in the presence of ethanol (red), pPD (blue) and dPD (green). The labelled peptides whose parent peptides have molecular weights of 637.1 and 1644.2 Da were eluted during this time. The ionisation current is normalised to the most ionising peak from the three traces within the *m/z* range displayed. The *m/z* of the peak apices of substantial peaks are noted, as are the molecular weights and charged states of their corresponding ions.

of 1860.3 Da (for D_0-labelled or with an additional 1 Da for D_1-labelled), for which 621.0 and 931.2 *m/z* are its triply and doubly charged ions respectively.

These labelled peptides were sequenced by MS-MS after collecting the eluate between 90 and 93 minutes. The EPI mass spectra in Figure 5.17A shows the overlay of the spectra for the pPD- and dPD-labelled samples obtained by fragmenting the doubly charged ions of peptides with 1860.3 and 1861.3 Da respectively. The presence of the y_1, y_2, and $y_{4–10}$ ions of the peptide $_{24}$LIECYSLDDKAETF$_{37}$ suggests that this is the peptide being labelled. Fragmentation ions with the charge on the N-terminal of this peptide are also observed at 199.3 and 358.4 *m/z* for the a_2 and b_3 ions respectively. Peaks that differ by 1 ± 0.2 *m/z* between the pPD- and dPD-labelled

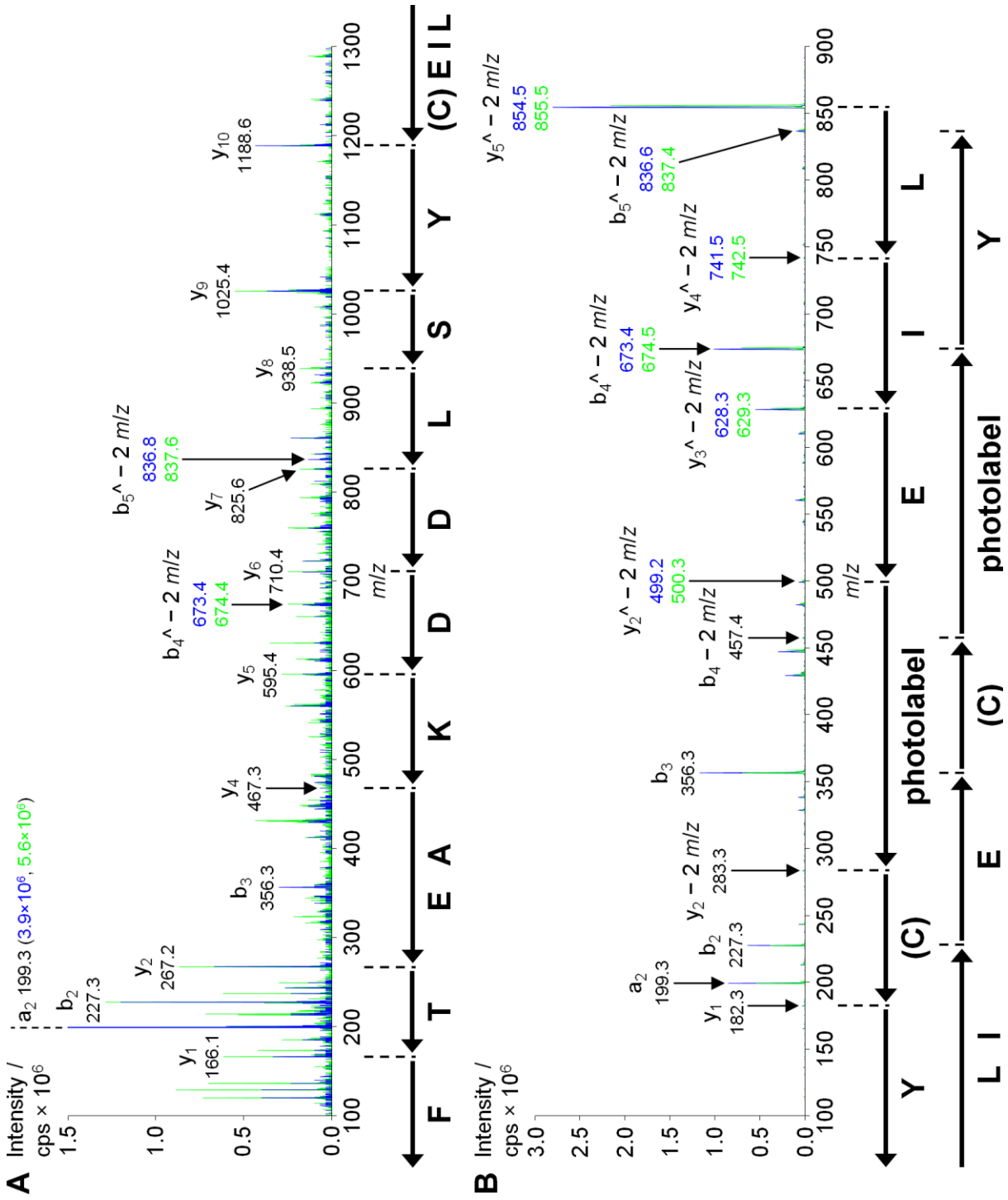


Figure 5.17: Enhanced product ion mass spectra overlays for the fragmentation of the parent ions 931.2 and 931.7 *m/z* (A) and 854.2 and 855.2 *m/z* (B) of GLIC treatment C samples labelled with pPD (blue) and dPD (green) respectively. The labelled peptides are identified as $_{24}$LIE(C)YSLDDKAETF$_{37}$ (A) and $_{24}$LIE(C)Y$_{28}$ (B) based on the presence of b_n and y_n ions associated with these peptides. (C) denotes a modified cysteine residue that is 2 Da less than the native cysteine. Peaks corresponding to labelled fragments are marked with an additional caret (blue for fragments labelled with D_0 photofol and green for those labelled with D_1 photofol).

samples are observed at 673.4 and 674.4 *m/z*, and at 836.8 and 837.6 *m/z*. These pairs of ions are best fitted to the *m/z* of labelled-b_4 and labelled-b_5 ions for the peptide $_{24}$LIECYSLDDKAETF$_{37}$, but with a 2-Da difference, thus assuming that the cysteine residue had undergone a modification resulting in a 2-Da loss.

Evidence supporting this cysteine-modification hypothesis came from the spectra in Figure 5.17B, which showed the overlay EPI spectra of fragmenting the parent ions 854.2 and 855.2 *m/z* for the pPD- and dPD-labelled samples respectively. The peaks at 182.3, 199.3, 227.3 and 356.4 *m/z* correlate to the y_1, a_2, b_2 and b_3 ions of the peptide $_{24}$LIECY$_{28}$ respectively. Neither the b_4 peak at 459.4 *m/z* nor the y_2 peak at 285.3 *m/z* are present on the spectra but peaks that are two units less than the expected *m/z* of these two common CID ions that contain the cysteine residue are instead observed. These peaks, marked b_4^ – 2 *m/z* and y_2^ – 2 *m/z* in Figure 5.17B, indicate that the photolabel was bound to a modified cysteine residue. All of the other expected b_n and y_n fragments containing the photolabel with the loss of 2 Da are also observed. The annotations beneath the spectra clearly show that the *m/z* differences between the dominant peaks of the spectra are equivalent to the residue masses of the amino acids of these labelled peptides that contain a modified cysteine residue. One explanation for the −2-Da deviation could be due to the combined effects of the oxidation (+16 Da) of cysteine coupled with the loss of water (−18 Da) to yield a net loss of 2 Da.

In a repeat of the pH 4 photolabelling experiment, it was not possible to identify peaks that were only present in the labelled samples or peaks that reduced in ionisation or absorbance intensity upon labelling due to the poor alignment of the TIC, 214- and 280-nm chromatograms obtained from running half of each sample on the HPLC-coupled mass spectrometer. In order to identify the presence of any labelled peptides, the remaining half of each sample was not subjected to fraction collection of the HPLC system, but was injected straight into the HPLC-coupled tandem mass spectrometer.

An information-dependent acquisition (IDA) method was used for these MS-MS runs. The IDA criteria selected ions from an 'include list' precedent over up to four of the most intense ions at 10-second intervals over the 90-minute duration of the run for carrying out EPI experiments. The 'include list' of ions consisted of the *m/z* of three groups of peptides. Firstly, the singly and doubly

charged ions of chymotryptic peptides (assuming cleavage occurred after only the six amino acids that are most prone to cleavage by chymotrypsin) around threonine-255—in which evidence for interaction with propofol was suggested by crystallographic and mutagenesis data (Nury *et al.*, 2011); secondly, the *m/z* of the most ionising ions of previously identified peptides, namely $_{24}$LIECY$_{28}$, $_{24}$LIECYSLDDKAETF$_{37}$, $_{166}$TAVVKPANF$_{174}$, $_{175}$ALEDRLESLDY$_{186}$ and $_{175}$ALEDRLESLDYQL$_{188}$; and finally, the calculated *m/z* of the doubly charged ions of potentially labelled peptides from the list of peptides whose concentration was observed to decrease upon labelling (Figure 5.12 and Figure 5.14).

The IDA EPI data were then analysed by the software *ProteinPilot*. This software allows its users to create their own database of peptide sequences and add user-defined amino acid modifications to its sequencing algorithm. The search database that was used consisted of the sequence of GLIC, chymotrypsin and maltoporin, the maltosaccharide-binding *Escherichia. coli* membrane protein (Ranquin and Van Gelder, 2004) that was co-purified with GLIC. The user-defined modification added to the search algorithm incorporated the possibility of all amino acids being modified with the addition of $C_{11}H_{11}F_3O$, the molecular formula of the D_0 photolabel, or with the addition of $C_{11}H_{10}DF_3O$, the molecular formula of the D_1 photolabel.

The results show that on average, 42.6 % amino acid coverage of GLIC was achieved from spectra in which there was a confidence score of greater than 95; 49.2 % amino acid coverage for a score greater than 50 and coverage of 98.4 % for all the sequenced mass spectra. However, the spectra with low confidence scores tended to ignore dominant peaks on the mass spectra and included peaks that have intensities similar to background noise peaks. Percentages of amino acid coverage were calculated for the extracellular (aa1–195) and transmembrane domains (aa196–317) separately. The amino acid coverage of the extracellular domain of the protein from spectra with sequence confidence scores of over 50 was calculated to be 71.8 % (140 out of 195 amino acids); whilst for the transmembrane domain, only 20 amino acids were identified in peptides with sequencing confidence scores of over 50, thus giving an amino coverage value of just 16.4 % (Figure 5.18). Even if only spectra with confidence scores of less than 10 are ignored, the percentage amino acid coverage of the transmembrane region is still 16.4 %, whilst the value for the extracellular domain increased to 78.5 %.

```
  1 GQDMVSPPPP IADEPLTVNT GIYLIECYSL DDKAETFKVN  40
 41 AFLSLSWKDR RLAFDPVRSG VRVKTYEPEA IWIPEIRFVN  80
 81 VENARDADVV DISVSPDGTV QYLERFSARV LSPLDFRRYP 120
121 FDSQTLHIYL IVRSVDTRNI VLAVDLEKVG KNDDVFLTGW 160
161 DIESFTAVVK PANFALEDRL ESKLDYQLRI SRQYFSYIPN 200
201 IILPMLFILF ISWTAFWSTS YEANVTLVVS TLIAHIAFNI 240
241 LVETNLPKTP YMTYTGAIIF MIYLFYFVAV IEVTVQHYLK 280
281 VESQPARAAS ITRASRIAFP VVFLLANIIL AFLFFGF    317
```

Figure 5.18: Schematic showing the amino acids of GLIC that are part of peptides sequenced from EPI spectra with a confidence score of greater than 50 (marked in red). The EPI spectra were obtained from tandem mass spectrometry runs of GLIC samples from the second treatment C photolabelling experiment using an IDA method.

Good quality spectra (confidence score of greater than 50) indicating photolabel binding to cysteine-27, lysine-170 and lysine-183 were observed, but there were no good quality spectra suggesting the labelling of other peptides.

5.2.3 Treatment D with urea, RapiGest, Lys-C and chymotrypsin post-photolabelling at pH 8

In the photolabelling experiments with the remaining protein samples, a protocol involving post-labelling treatment of GLIC that mimics the protocol used in the sequencing of $GABA_A$ α_1 subunit (Chen *et al.*, 2012a) by Professor Alex Evers' group was used. This protocol differs from the ones discussed in previous section in that urea and RapiGest were used as the denaturant and both endoproteinase Lys-C and chymotrypsin were used to digest GLIC. The detailed method is described in Section 2.4.2.4.

The use of the endoproteinase Lys-C as the primary enzyme prior to using chymotrypsin was to promote cleavage of the transmembrane region. The amino acid coverage analysis with *ProteinPilot* presented at the end of section 5.2.3 revealed that only 16 % (20 / 122 amino acids) of the transmembrane domain of GLIC was sequenced with a confidence score of above ten. There are two lysines within the transmembrane domain of GLIC, presenting two cleavage sites for Lys-C in that region. The endoproteinase Lys-C has been reported to retain 60 % of its enzyme activity in the presence of 8M urea (Klammer and MacCoss, 2006). The resultant Lys-C cleaved peptides

should provide more accessible chymotrypsin cleavage sites than when these sites are hidden within the intact transmembrane domain.

Figure 5.19 shows the separate alignments of the three monitored-parameters—TIC and UV absorbance at 214 and 280 nm—for the peptides generated from Lys-C digestion over 48 hours followed by 5-hour chymotrypsin digestion for the unlabelled and pPD- and dPD-labelled samples. The 20-hour chymotrypsin digestion chromatograms are not shown here, the results of the comparison of the labelled sample traces with their respective unlabelled samples are similar for both durations of chymotrypsin digestion. At both durations—five and twenty hours, the chromatograms for the first 75 minutes of the HPLC-MS runs are identical between the three samples. A few new peaks (absent in the GLIC + EtOH sample) are observed in the UV absorbance at 214 and 280 nm between 80 and 140 minutes (marked by arrows on the inset of Figure 5.19C). The most noticeable of these new absorbance peaks are at 95.5, 102.0 and 102.9 minutes. However, the only differences between the unlabelled and labelled samples on their TIC chromatograms within the first 120 minutes of the HPLC-MS run are at 84.8 and 93.9 minutes (marked by arrows on the inset of Figure 5.19D), where increases in intensities are seen upon labelling GLIC.

As preparation for MS-MS to sequence any labelled peptides, the averaged mass spectra around the two new TIC peaks were investigated to identify the *m*/*z* of the ions that have contributed to the increase in ionisation current of the labelled samples at these time points. The mass spectra at retention times where new peaks were observed only in the UV absorbance chromatogram were also examined. Two of these averaged mass spectra overlay are shown in Figure 5.20, whilst the results from the analysis of the averaged mass spectra at other time points are summarised in Figure 5.21.

The mass spectra in Figure 5.20A show that between 92.5 and 96.0 minutes, ions with *m*/*z* of 603.1 and 904.2, corresponding to the peptide with a MW of 1806.4 Da being eluted. Note that for both these two peaks, the ionisation abundances between the three samples are similar. However, there are peaks at 556.3, 834.2 and 856.2 *m*/*z* in the pPD-labelled sample, 556.9, 834.9 and 857.2 *m*/*z* in the dPD-labelled sample that are absent in the unlabelled sample. These *m*/*z* peaks correspond to the ions of molecular weights 639.1 and 1450.7 Da, labelled with a D_0 photolabel (GLIC + pPD

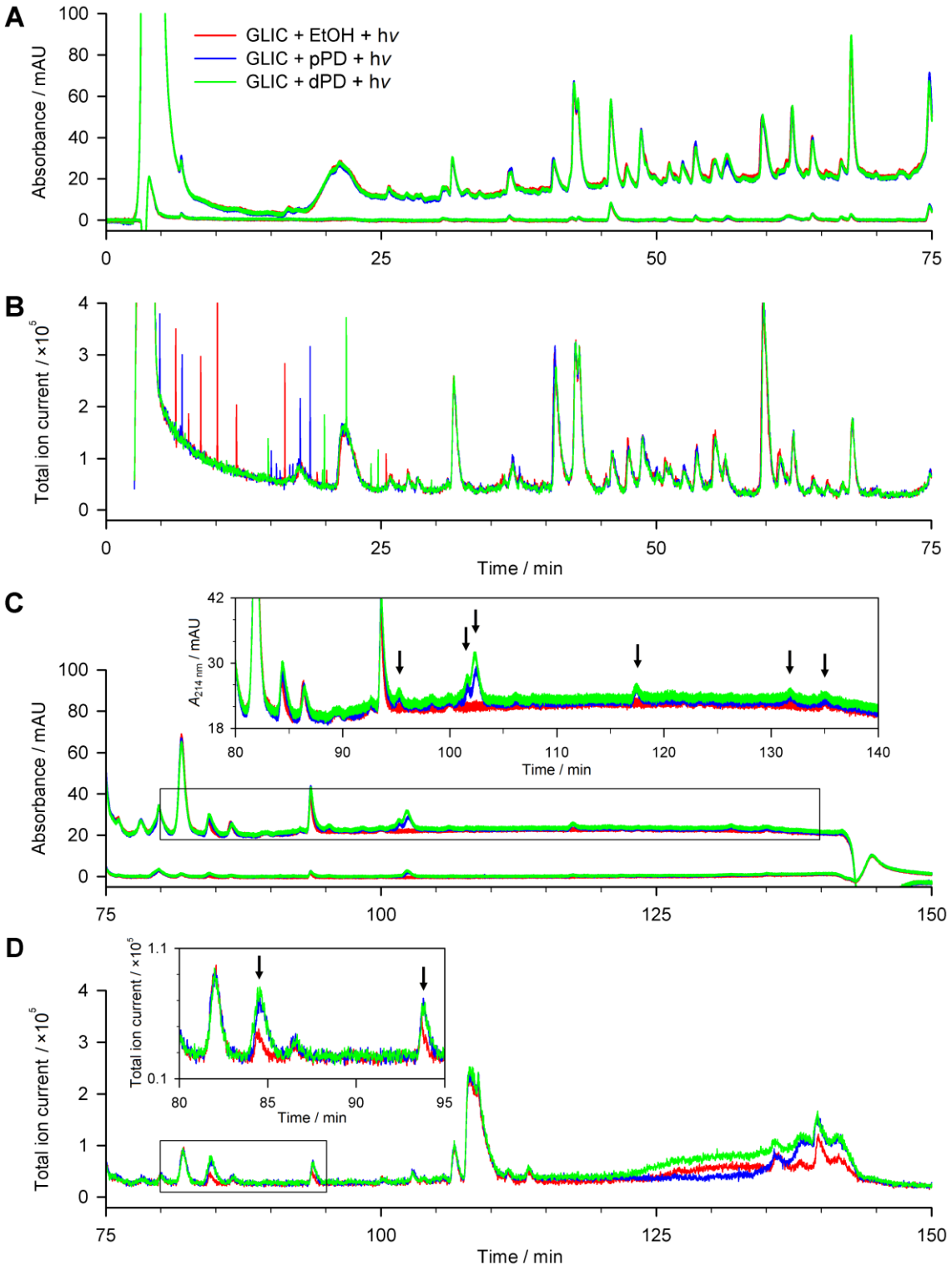


Figure 5.19: Chromatograms overlay of the absorbance at 214 and 280 nm (A and C) and TIC (B and D) of GLIC that were exposed to UV light in the presence of ethanol (red), pPD (blue) and dPD (green) at pH 7.8 prior to post-labelling treatment D as outlined in section 2.4.2.4. The peptides were separated with a gradient elution of ACN from a reverse phase HPLC column with eluent containing 0.06 % formic acid. Inset in C and D show enlarged chromatograms highlighting the presence of peaks appearing or increasing in intensity when GLIC was photolabelled with protofol. Arrows in inset C and D mark the appearance or increased intensity of peaks upon labelling.

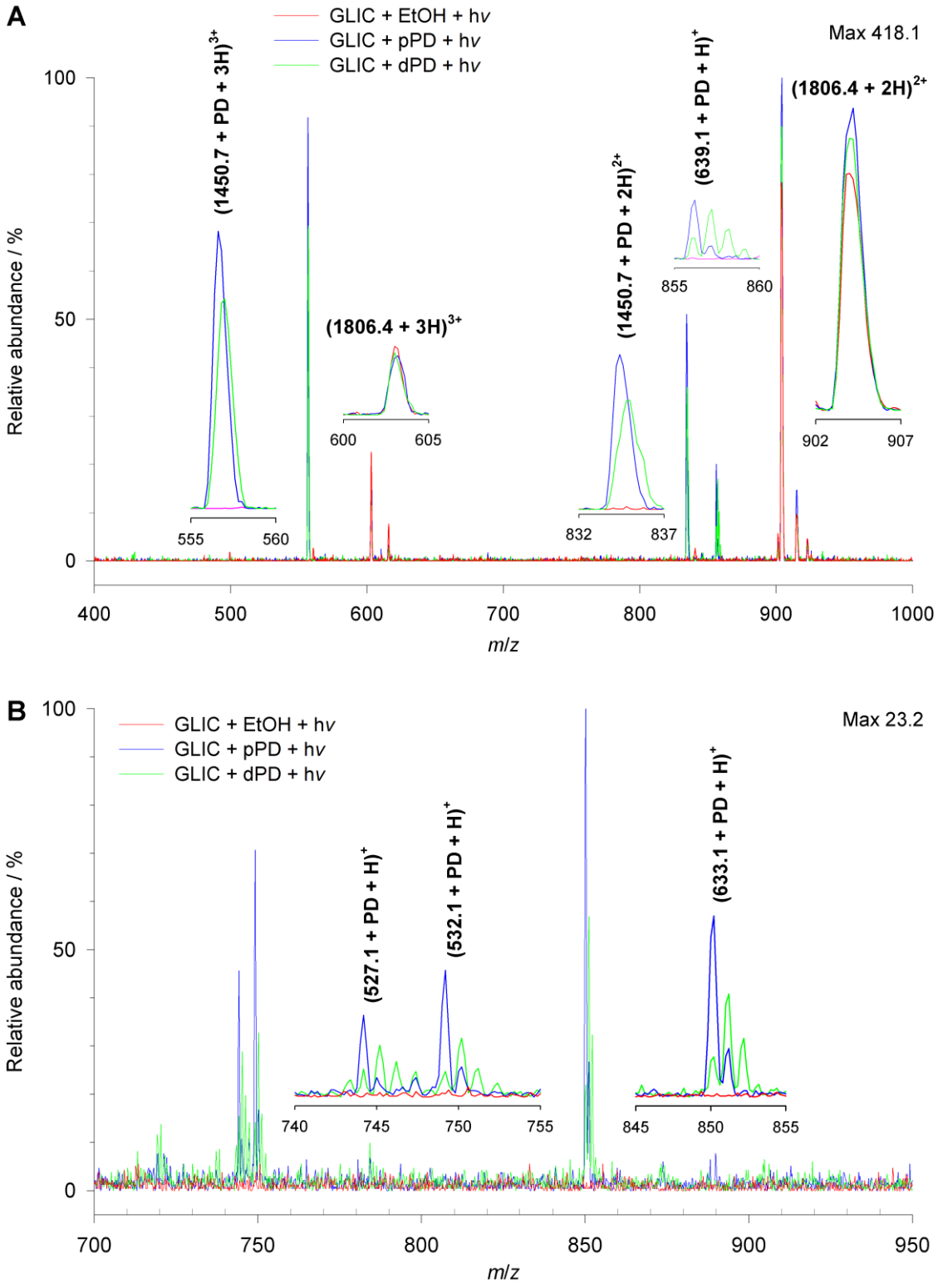


Figure 5.20: Overlays of averaged mass spectra for samples from GLIC treatment D experiments being eluted at 92.5–96.0 minutes (A) and at 120–145 minutes (B). The insets of enlarged peaks corresponding to ions of one unlabelled peptide (1806.4 Da) and two labelled peptides (with molecular weights of photolabel plus 639.1 and 1450.7 Da) in (A) and of three labelled peptides (with molecular weights of photolabel plus 527.1, 532.1 and 633.1 Da) in (B) are shown.

sample) or with a D_1 photolabel (which is the most abundant photolabel isotope in the GLIC + dPD sample). Figure 5.20B shows the overlay of the averaged mass spectra over 120–145 minutes when the TIC chromatograms of the three samples did not align well, but two new peaks at 132 and 135 minutes can been seen in the UV absorbance spectra. Three singly charged peptides with molecular weights of 527.1, 532.1 and 633.1 Da were found to be labelled by D_0 photolabel of the pPD-labelled sample and by a mixture of D_0, D_1 and D_2 photolabels of the dPD-labelled sample. The characteristic isotopic peak pattern for the specific ratio of the different photolabel isotopes used (D_0: 34 %, D_1: 43 % and D_2: 23 %) can be seen in the dPD-labelled peptides.

The more precise elution times of the ions with *m/z* peaks that only occur in the labelled samples presented in Figure 5.21 were determined by the Extract Ions tool of *ChemStation*. Fractions containing these labelled peptides were collected and these samples were subjected to MS-MS. However, the EPI spectra generated from the fragmentation of these labelled peptides were not of decent quality for *de novo* sequencing, so one can only predict the peptides being labelled based on their molecular weights that were determined by the single-stage mass spectrometer.

Averaged mass spectra period / mins	*m/z* peaks identified (charged states of ions)[a]	Molecular weight / Da	
		Labelled peptide[a]	Parent peptide
83.0–85.8	581.9 (2+)	1161.7	945.6
92.5–96.0	834.4 (2+), 556.6 (3+)	1666.8	1450.7
	856.2 (1+)	855.2	639.1
97.0–99.0	955.2 (2+), 637.1 (3+)	1908.3	1692.2
99.0–104.5	1090.2 (2+), 727.1 (3+)	2178.4	1962.3
	1464.6 (2+)[b], 976.7 (3+)[b]	2927.2	2711.1
120.0–145.0	744.2 (1+)	743.2	527.1
	749.2 (1+)	748.2	532.1
	850.2 (1+)	849.2	633.1

Figure 5.21: Table summarising the *m/z* peaks that are present in the labelled GLIC samples but absent in the unlabelled sample from the analysis of averaged mass spectra at various retention time ranges when increases in UV absorbance and/or TIC are seen in the labelled samples. The molecular weights corresponding to these charged labelled ions and of the native peptides are noted. [a] Molecular weights of the peptides with a D_0 label attached. [b] *m/z* that are only present in the 5-h chymotrypsin-digestion sample.

Some of the molecular weights of the labelled peptides listed in Figure 5.21 are the same as the molecular weights of previous identified labelled peptides, namely TAVVKPANF, ALEDRLESKLDY and ALEDRLESKLDYQL for the molecular weights of 1161.7, 1666.8 and

1908.3 Da respectively and these have similar retention times as the labelled peptides identified before, thus they are likely to be the same peptides being labelled. A list of all the GLIC peptide sequences with molecular weights that match those of the parent peptides presented in Figure 5.21 with a tolerance of ±1 Da are presented in Figure A.16 in the appendix. It is interesting to note that the peptide LIECY—which was previously identified to be labelled, but with a loss of 2 Da on all CID fragments containing the cysteine residue—has a MW that matches the labelled peptide collected at 95–96 minutes and that also another peptide fragment containing cysteine-27, IECY matches the MW of the 527.1-Da parent peptide that was subjected to labelling, thus providing some evidence that the cysteine residue might also be labelled under pH 8 condition and not just at pH 4.

Furthermore, subsequent use of the same batch of Lys-C enzyme used for this experiment to digest a protein in the presence of 8 M urea show that the batch of enzyme used was unfortunately not active at 8 M urea. Upon incubating the protein with Lys-C at 8 M urea for 2 days and then diluting the urea concentration to 1 M before analysing the mixture by HPLC-MS, the UV absorbance and TIC chromatograms were exactly the same as those obtained from a reaction mixture containing just Lys-C and 1 M urea (data not shown). This result implies that the sole attempt of the post-photolabelling treatment D on GLIC carried out during this project could not have improved the amino acid coverage of the protein—compared to the other post-photolabelling treatments of GLIC—based on the additional usage of the enzyme Lys-C.

5.3 The labelled amino acids of GLIC

To summarise the results of photolabelling GLIC presented in section 5.2: the lysine residue at amino acid position 183 was labelled in all of the experiments undertaken, appearing in the peptides $_{175}$ALEDRLESKL$_{184}$, $_{175}$ALEDRLESKLDY$_{186}$ and $_{175}$ALEDRLESKLDYQL$_{188}$. These peptide sequences and the locations of the bound photolabels were then confirmed by MS-MS. Another lysine that had been labelled in almost all of the experiments was lysine-170 of $_{166}$TAVVKPANF$_{174}$. In the pH 4 labelling experiments, the sole cysteine residue of GLIC at amino acid position 27 was shown to be labelled in the EPI spectra of both the labelled $_{24}$LIECY$_{28}$ and $_{24}$LIECYSLDDKAETF$_{37}$ peptides but with an unusual −2-Da modification of the cysteine residue. The single-stage mass spectra data from the pH 8 labelling experiment that was followed by urea

denaturation and two-stage digestion with Lys-C and chymotrypsin (treatment D) suggests that the peptide $_{24}$LIECY$_{28}$ could also be labelled in non-acidic conditions (Figure 5.20A) but the sequence of this labelled peptide, with a MW of 855.2 Da when labelled with pPD, was not confirmed by MS-MS.

The locations of the labelled amino acids cysteine-27, lysine-170 and lysine-183 were investigated on molecular models of the X-ray structural data of GLIC with propofol bound (Nury *et al.*, 2011). Figure 5.22 shows that lysine-183 sits at the opening of a groove at the interface between subunits. Closer inspection of this region shows the presence of an aromatic amino acid, phenylalanine-174, lining the top wall of this potential binding cavity, which may interact with the aromatic anaesthetic propofol; and also the presence of proton accepters, such as tyrosine-23 and asparagine-152, within this site, which may form a hydrogen bond with the drug to facilitate its binding at the site.

A propofol molecule was superimposed onto the structural model file of propofol-bound GLIC—PDB accession code 3P50—using the software *Coot* (Emsley and Cowtan, 2004). The propofol molecule was placed such that its benzene ring was parallel to and directly underneath the benzene ring of phenylalanine-174, with a distance of 3.9 Å between the π-bond layers, and such that the hydroxyl group of the drug faced towards one of the proton acceptors within the site that were within 5 Å apart, namely tyrosine-23. The distance between the terminal nitrogen of lysine-183 and the middle carbon of one of the isopropyl arms of the superimposed propofol molecule (equivalent to the location of the carbene of photoactivated photofol) in the structural model was measured to be 5.7 Å. This is comparable to the distances between labelled amino acids and the propofol molecule within their corresponding propofol-binding sites in HSA.

The other two amino acids that have been identified to be labelled, lysine-170 and cysteine-27, are both further away from this small potential extracellular domain binding cavity of propofol. The terminal nitrogen of lysine-170 and the sulphur of cysteine-27 are both 15.3 Å from one of the isopropyl arms of the propofol molecule that was modelled into the cavity. It is important to note that lysine-170 is predicted to be in a flexible loop flanked by two β-sheet secondary structures and therefore has the potential to move closer towards the binding groove. Conversely, the X-ray structure shows that cysteine-27 does not even lie on the surface of the protein. However, an

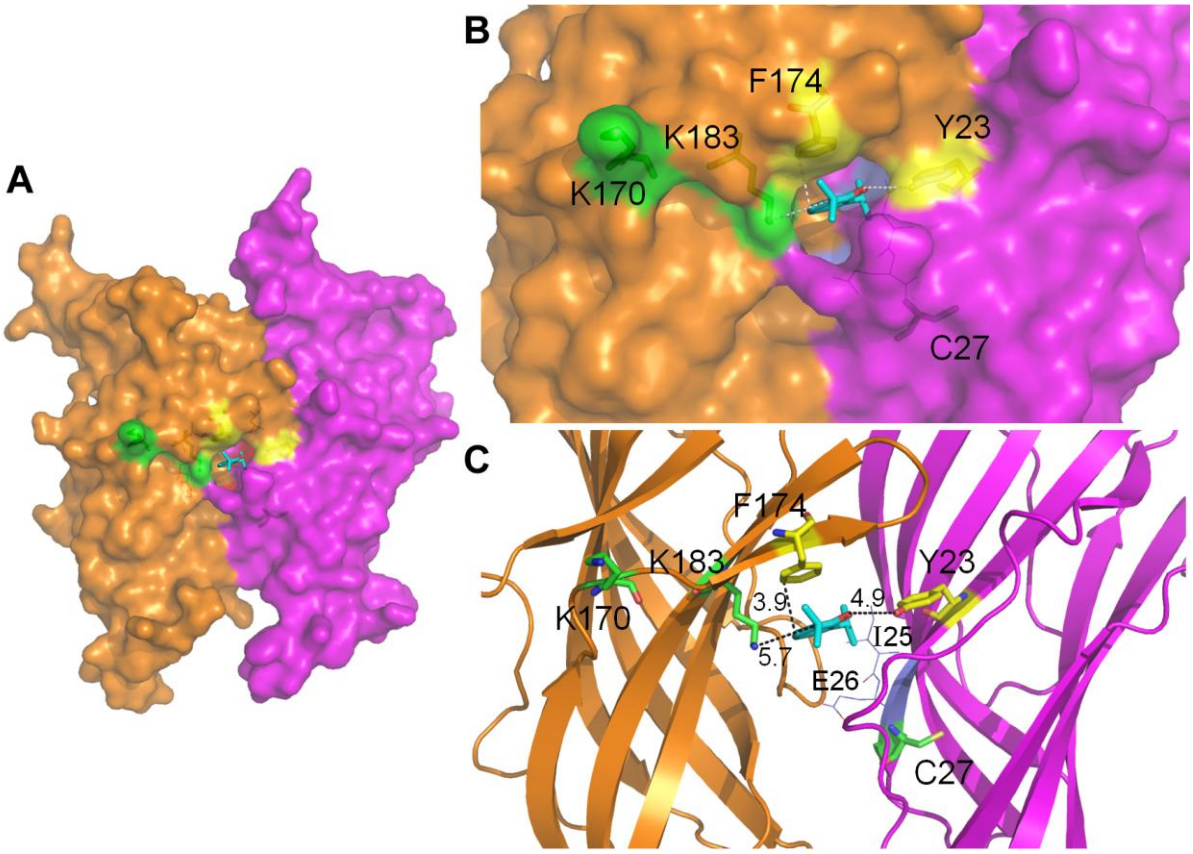


Figure 5.22: Labelled amino acids cysteine-27, lysine-170 and lysine-183 in the extracellular domain of GLIC superimposed onto the coordinates of a propofol molecule placed at a potential extracellular propofol-binding cavity. The surfaces of two adjacent subunits of the extracellular domain of GLIC (amino acids 1–199) are separately coloured orange and magenta (A), and enlarged pictures show the surface (B) and cartoon representation (C) of the groove at the interface between subunits. The amino acids phenylalanine-174 of the orange subunit and tyrosine-23 of the magenta subunit are displayed as yellow stick models. The three aforementioned labelled amino acids are displayed in green sticks. The propofol molecule, shown by cyan sticks, is superimposed onto the groove at the interface between subunits such that the aromatic ring of the propofol phenol stacks underneath the aromatic ring of phenylalanine-174 with the distance between the π-bond layers being 3.9 Å. The hydroxyl of the propofol is 4.9 Å from the hydroxyl of tyrosine-23, indicating the presence of a nucleophile in the binding vicinity with which propofol could potentially interact. The terminal nitrogen of lysine-183 is 5.7 Å from one of the isopropyl arms of the superimposed propofol molecule, whilst lysine-170—also appearing on the surface of GLIC—and cysteine-27 are both 15.3 Å from one of the isopropyl arms of the modelled propofol (distances not marked). The amino acids isoleucine-25 and glutamic acid-26—both of which are part of the labelled peptides of $_{24}$LIECY$_{28}$ and $_{24}$LIECYSLDDKAETF$_{37}$ (EPI mass spectra of Figure 5.17) and the 743.2-Da potentially labelled peptide $_{25}$IECY$_{28}$ (mass spectra of Figure 5.20)—are shown in pale blue lines. These two amino acids can be seen to line the inner surface of the groove in panel B.

interesting observation is that both amino acids proceeding the cysteine, isoleucine-25 and glutamic acid-26, which are both part of the labelled peptides $_{24}$LIECY$_{28}$ and $_{24}$LIECYSLDDKAETF$_{37}$ that were identified, line the surface within the extracellular binding cavity (coloured pale blue in Figure 5.22).

The EPI spectrum for labelled $_{24}$LIECY$_{28}$ in Figure 5.20B showed a small peak corresponding to a photolabel-bound C-terminal fragment of the last two amino acids and a larger peak corresponding to the N-terminal fragment of the first four amino acids, suggesting that the photolabel was bound to the cysteine residue. On the other hand, the EPI spectrum for labelled $_{24}$LIECYSLDDKAETF$_{37}$ in Figure 5.20A only provides the information that the label is bound to one of the first four amino acids of this peptide based on the presence of b_4^ ions, so the photolabel could potentially be bound to isoleucine-25 or glutamic acid-26.

The hypothesis that this groove at the interface between the extracellular domains of subunits being a site for propofol-binding is strengthened by the apo X-ray structure of the bacterial homologue of GLIC from *Erwinia chrysanthemi* known as ELIC. The ELIC protein was crystallised at pH 6.5 in an apparent closed conformation (Hilf and Dutzler, 2008) whilst the X-ray structure of apo GLIC showed the channel in an open-pore conformation (Bocquet *et al.*, 2009). Figure 5.23 shows that the surfaces of the propofol-bound (Nury *et al.*, 2011) and apo GLIC proteins appear to be similar, whereas the potential binding cavity in ELIC is significantly greater in size. The considerable structural rearrangements that occurs between these homologous proteins in this region has been noted by a few authors and has been linked to the gating of these channels (Hilf and Dutzler, 2009; Maksay, 2009; Zimmermann and Dutzler, 2011; Pan *et al.*, 2012b). Furthermore, the photolabelled extracellular domain site overlaps with the binding location of large antagonists of acetylcholine binding protein (AChBP) and is less than 10 Å below the agonist binding sites of AChBP (Figure 5.24), glycine receptors (Hibbs and Gouaux, 2011) and ELIC (Zimmermann and Dutzler, 2011).

More striking evidence supporting the labelled sites at the subunit interfaces of the extracellular domain being binding sites for modulators of GLIC came to light very recently. During the preparation of these results for this chapter, the crystallographic data of ketamine-bound GLIC was published (Pan *et al.*, 2012a) showing the same residues being implemented in the binding of

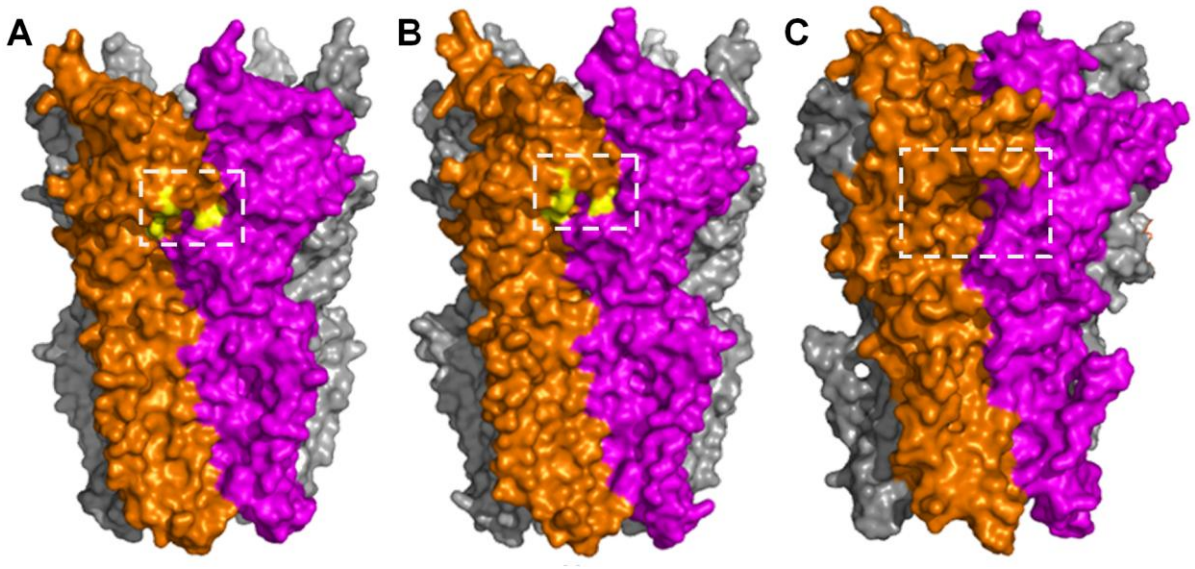


Figure 5.23: The potential extracellular binding cavities at the interface between subunits for propofol-bound GLIC (A), native GLIC (B) and native ELIC (C). The front two subunits are coloured orange and magenta, whilst the back three are coloured grey. The dashed white boxes highlight the locations of the cavities. In A and B, the amino acids phenylalanine-174 and lysine-183 of the orange subunit and tyrosine-23 of the magenta subunit are coloured yellow. These residues mark the top, left and right of the GLIC extracellular domain cavity respectively, as shown in more detail in Figure 5.22.

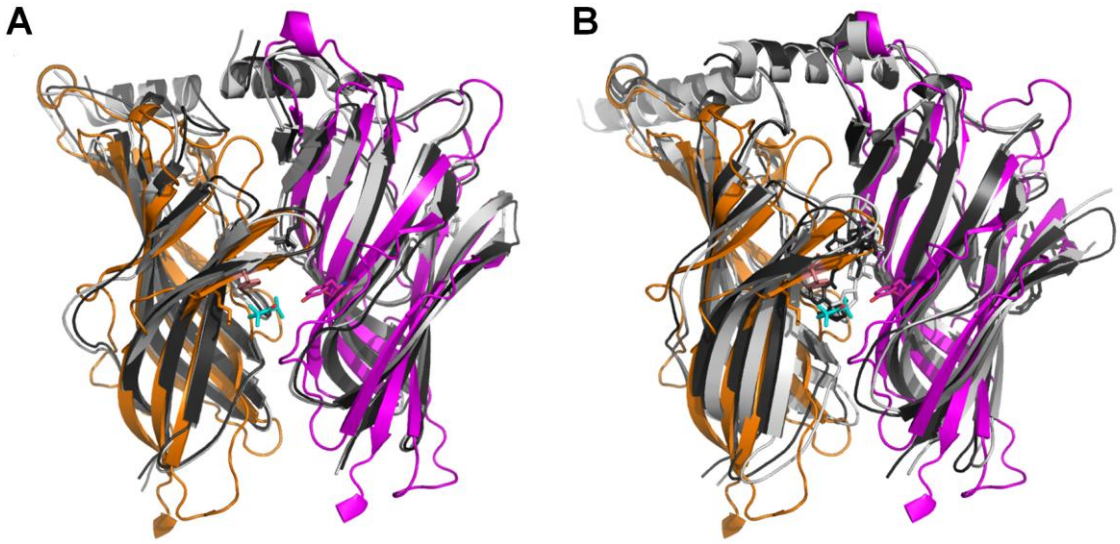


Figure 5.24: Alignment of two extracellular domains of GLIC subunits with acetylcholine binding protein (AChBP) complexed with two agonists (A) and two antagonists (B). The PDB files of AChBP with the agonists nicotine (accession code 1UW6, light grey) and carbachol (accession code 1UV6, dark grey) overlaid on GLIC (accession code 3P50) show that the predicted binding location of propofol (cyan sticks) is about 7 Å below that of the agonist-binding site of AChBP. The PDB files of AChBP with the antagonist gymnodimine A (accession code 2X00, light grey) and methyllycaconitine (accession code 2BYR, dark grey) overlaid on GLIC show that the predicted binding location of propofol overlaps the bottom of the antagonist-binding site of AChBP.

ketamine to the extracellular domain of GLIC (Figure 5.25) as those that were predicted to be involved with propofol binding there (Figure 5.22). Cysteine-27 was noted earlier to be not located on the surface of the protein in the propofol-bound GLIC structure. However, the ketamine-bound structure consisting of the photolabelled site in the extracellular domain being occupied by a ligand shows that this residue does indeed sit on the surface at the bottom of the binding cavity, but approximately 12 Å from the ligand (Figure 5.26). Pan *et al.* (2012b) conducted electrophysiology experiments on N152C/C27A double mutant GLIC and confirmed that the presence of a ligand at this ketamine site, as well as the potential propofol site, inhibits the channel. The labelling of cysteine-152 with dibenzofuransulfonic acid that mimics the presence of a ligand at the site inhibits the current from an EC_{20} proton concentration, which was restored upon the removal of dibenzofuransulfonic acid by DTT. However, the same group has shown that whether

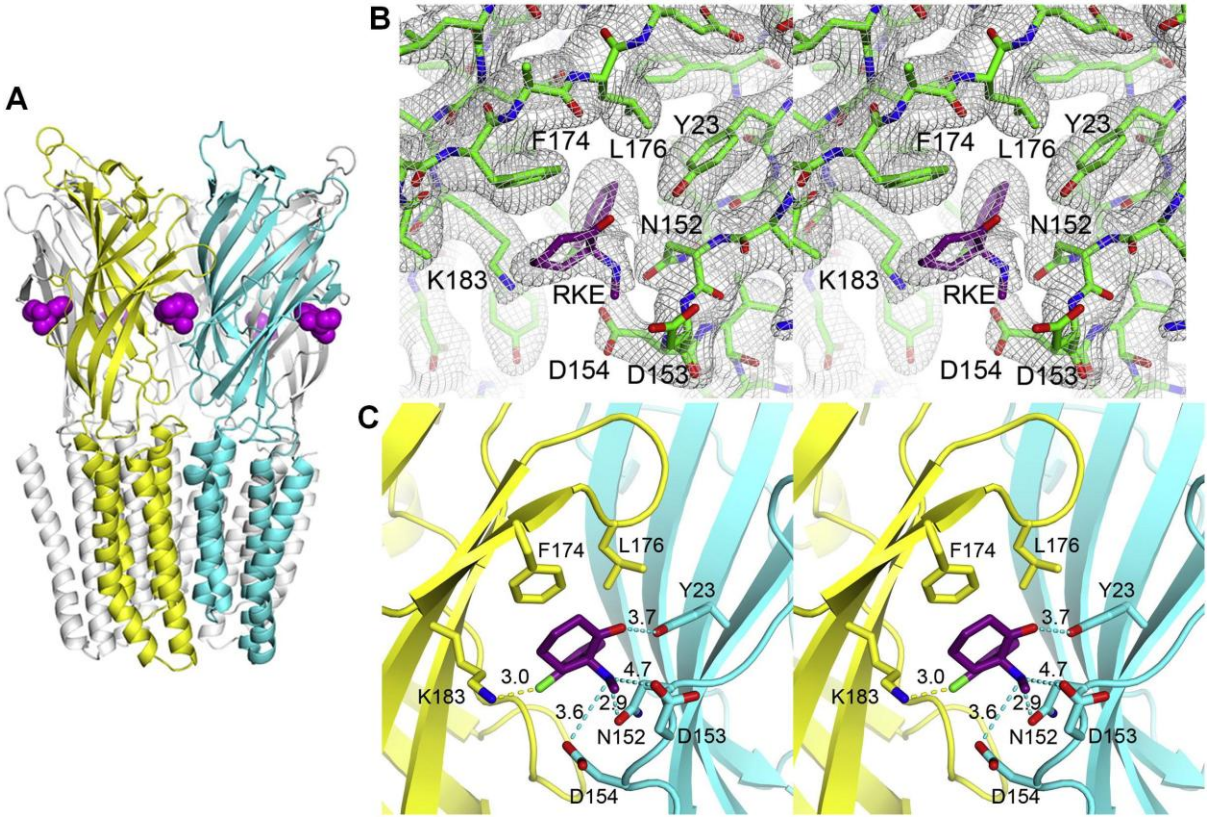


Figure 5.25: Crystal structure of ketamine-bound GLIC. The side view showing five ketamine molecules (purple spheres) bound to the interfaces between a pair of principal (yellow) and complementary (cyan) subunits is depicted in cartoon representation (A). Stereo views of the ketamine-binding site with the electron density map contoured at 1.0 σ level (B) and stick models of residues around ketamine with dashed lines indicating distances in angstroms (C) are shown. Figure taken from Pan *et al.* (2012).

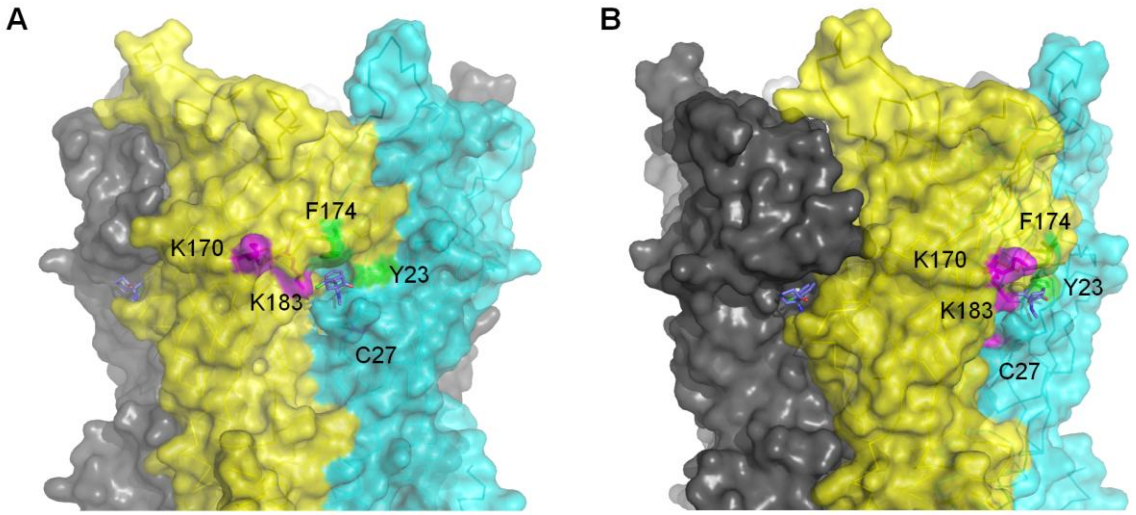


Figure 5.26: Molecular models showing the three photolabelled amino acids K170, K183 and C27 on the surface of the ketamine binding site of the ketamine-bound GLIC structure. (A) The front view of one pair of principal (yellow) and complementary (cyan) subunits and (B) a rotated view in the plane of the pore axis are shown. Ketamine molecules are depicted in purple sticks, labelled amino acids in magenta and other residues predicted to interact with propofol, namely F174 and Y23, are shown in green.

a ligand elicits agonistic or antagonistic effects at the orthosteric site depends on the interaction of the ligand with the residues at the site. This was demonstrated by acetylcholine acting as a competitive antagonist at ELIC despite producing similar conformational changes of the channel to those produced by agonists in AChBP, and its derivative 2-dimethylaminoethylacetate—lacking one methyl group in the choline group—activating ELIC (Pan *et al.*, 2012b). The inhibition of GLIC due to the binding of propofol and ketamine at the subunit interfaces of the extracellular domain may also be due to interference with the activation of the channel by its endogenous ligand, protons, for which the proton-sensing site of the channel had been attributed to the extracellular domain (Dureta *et al.*, 2011).

5.4 The potential labelled peptides of GLIC

The labelled amino acids illustrated in section 5.3 are all located within the extracellular domain, contrasting the crystallographic data for propofol-bound GLIC that show the propofol-binding site at the top of the transmembrane domain (Nury *et al.*, 2011). However, in section 5.2, the typical TIC and UV absorbance chromatograms showed that a handful of peptides reduced in intensity when GLIC was exposed to light in the presence of the photolabel. These peptides are

therefore potentially labelled in GLIC. The positions of these peptides are highlighted in Figure 5.27. Note that the peptides $_{53}$AFDPVRSGVRVKTY$_{66}$ and $_{55}$DPVRSGVRVKTY$_{66}$ at the top of the extracellular domain, highlighted in cyan in Figure 5.27A only demonstrated a reduction in concentration upon labelling from one of the two experiments for which this analysis was carried out (Figure 5.14 compared to Figure 5.12). Figure 5.27 shows that the $_{233}$IAHIAF$_{238}$ peptide, the end of $_{228}$VVSTL$_{232}$ and the beginning of $_{239}$NILVETNLPKTPY$_{251}$ are situated in close proximity to the propofol molecule in the X-ray structure.

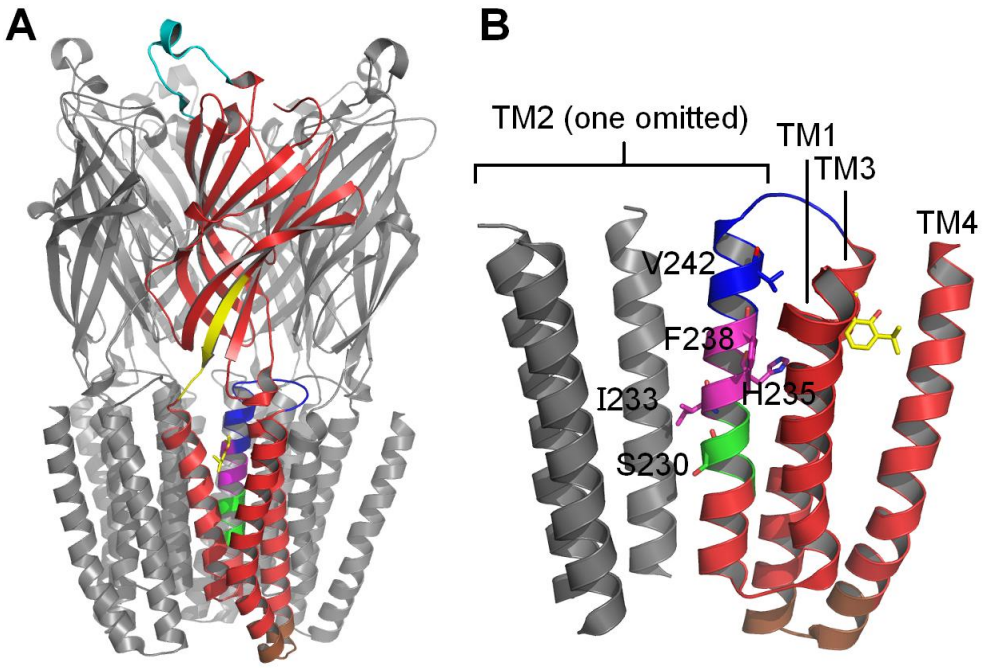


Figure 5.27: Potential labelled peptides of GLIC that have a reduction in ionisation intensity upon photolabelling with photofol as illustrated in Figure 5.12 and Figure 5.14. (A) A cartoon representation of the crystallography data of propofol-bound GLIC with the front subunit predominantly coloured red for both the extracellular and transmembrane domains, and potentially labelled peptides coloured differently. (B) Enlarged selection of panel A showing a partial transmembrane region consisting of one subunit and the TM2 helices of three other subunits. The TM2 of one subunit has been omitted for clarity of the location of isoleucine-233 and serine-230 that line the channel pore. A propofol molecule in the coloured subunit is represented in yellow sticks. The extracellular domain peptides $_{187}$QLRISRQY$_{194}$ and $_{53}$AFDPVRSGVRVKTY$_{66}$ are coloured yellow and cyan respectively. The peptides within the transmembrane domain that have been potentially labelled, $_{228}$VVSTL$_{232}$, $_{233}$IAHIAF$_{238}$, $_{239}$NILVETNLPKTPY$_{251}$ and $_{279}$LKVESQPARAA$_{289}$ are coloured green, magenta, blue and brown respectively. Residues facing the pore or towards the intra-subunit binding cavity are depicted with stick models.

Valine-242 is 6 Å away from the propofol molecule when it is best fitted to its density map within the binding site but electrophysiology data showed that its mutation to methionine increases the sensitivity of the channel to propofol. This coupled with the fact that no significant changes in propofol sensitivity were seen in the mutations of isoleucine-202 despite this amino acid being in contact with propofol suggest that propofol has significant mobility within this binding site (Nury *et al.*, 2011) and thus it is possible to photolabel the more distant valine-242 over amino acids in closer contact. Valine-242 is part of the peptide $_{239}$NILVETNLPKTPY$_{251}$ that reduced in concentration when GLIC was exposed to light in the presence of the photolabel. Two of its preceding peptides, $_{228}$VVSTL$_{232}$ and $_{233}$IAHIAF$_{238}$, also showed reductions in concentration when GLIC was labelled. The latter has two of its amino acids, histidine-235 and phenylalanine-238, on the side of TM2 facing the binding cavity, whilst both peptides have a residue—serine-230 and isoleucine-233 respectively—that lines the channel pore (Figure 5.27B) in a region where propofol had been predicted to bind from computational studies (LeBard *et al.*, 2012). $_{228}$VVSTL$_{232}$, $_{233}$IAHIAF$_{238}$ and $_{239}$NILVETNLPKTPY$_{251}$ have decreased in ionisation by a degree similar to the peptides whose labelled counterparts were identified and confirmed by MS-MS sequencing (Figure 5.12 and Figure 5.14).

However, other peptides, namely $_{187}$QLRISRQY$_{194}$ and $_{279}$LKVESQPARAA$_{289}$, also showed reductions in concentration when exposed to light in the presence of the photolabel similar to peptides $_{175}$ALEDRLESKL$_{184}$ and $_{175}$ALEDRLESKLDY$_{186}$ that had been confirmed to be labelled. $_{187}$QLRISRQY$_{194}$ and $_{279}$LKVESQPARAA$_{289}$ are neither part of the crystallography-confirmed transmembrane domain binding cavities, the potential extracellular domain binding cavities, nor sites predicted from computational studies, therefore the reductions in their concentration suggest that either GLIC has other minor binding cavities for propofol or the photolabel has labelled non-specifically. One argument for the reduction in the concentration of the peptide $_{187}$QLRISRQY$_{194}$ could be that its adjacent peptide, $_{175}$ALEDRLESKLDY$_{186}$, has a bound photolabel that blocks the cleavage site after tyrosine-186.

5.5 Concluding remarks for the photoaffinity labelling of GLIC

Photolabelling GLIC with photofol labelled three amino acids in the extracellular domain, lysine-170 (Figure 5.7), lysine-183 (Figure 5.8) and cysteine-27 (Figure 5.17), contrasting the X-

ray data showing propofol binding to the transmembrane domain from crystals grown at pH 4.0 (Nury *et al.*, 2011). However, propofol modulates GLIC by inhibiting its proton-activated cation current (Weng *et al.*, 2010) but the X-ray structures of the drug-bound GLIC and apo GLIC both showed the channel in an open state, suggesting that the propofol site identified by crystallography—at a pH that is at the EC_{100} of GLIC activation—is not the only site of the anaesthetic.

Studying the X-ray structures of propofol-bound GLIC (Nury *et al.*, 2011), native GLIC (Bocquet *et al.*, 2009; Hilf and Dutzler, 2009) and native ELIC (Hilf and Dutzler, 2008)—a bacterial homologue from *Erwinia chrysanthemi* with a similar structural alignment to GLIC that was crystallised in a closed-pore conformation—suggests that the three labelled amino acids, cysteine-27, lysine-170 and lysine-183, may be situated at or near a potential binding pocket of propofol that overlaps with the antagonist site of AChBP and just below the orthosteric agonist site of homologous ion channels (section 5.3). The surface representations of the native GLIC crystallised at pH 4.6 and the ELIC structure from crystals grown at pH 6.5 were examined on molecular viewers. These showed that the size of the putative extracellular domain binding sites were much smaller in GLIC, which was crystallised in an apparent open conformation, than in ELIC (Figure 5.23). The constriction of the orthosteric agonist site, which is less than 10 Å from the labelled site, has been associated with conformational changes that lead to the gating of the channel initiating from the movement of loop C, a loop connecting the β9 and β10 strands (Pan *et al.*, 2012b). The significant finding from crystallographic studies that ketamine also binds to the same site as the predicted propofol site based on photolabelling results and that the interactions were shown or predicted to implicate the same residues greatly support this extracellular domain site being a binding site of propofol that had not been resolved in the propofol-GLIC structure.

Labelled peptides containing amino acids from the transmembrane domain were not identified. This may be due to their unlabelled native peptides not initially being detected by HPLC-MS. The percentage amino acid coverage of the transmembrane domain from processing GLIC into peptides using treatment A was predicted to be 21 %. Using treatment B, this increased to 43 % from matching expected chymotryptic peptides with the molecular weights of peptides identified from mass spectra. The reasons for peptides from a large proportion of the transmembrane

domain not being identified could be due to inefficient chymotrypsin cleavage of the transmembrane region or over-digestion by the enzyme, both of which generate peptides that are not detectable by the 100–1500-*m*/*z* range of the quadrupole mass spectrometer used. Alternatively, the hydrophobic nature of these transmembrane peptides brings about their precipitation out of solution and thus could have been discarded along with the breakdown products of the acid-labile detergent used.

Attempts were made to improve the amino acid coverage of GLIC, particularly of the transmembrane domain, by using different surfactants during digestion as well as using multiple digestive enzymes. The use of SDC instead of chaotropic agents such as urea or GnHCl was adapted from the protocol by Masuda *et al.* (2008). This adapted protocol, referred in this text as treament B, increased the number of GLIC peptides detected by HPLC-MS compared to treatment A, in which GnHCl was used without the presence of any surfactants (Figure 5.3). However, the sole attempt of the protocol described here as treatment D, which involved the use of the surfactant RapiGest and the additional enzyme Lys-C, mimicking a protocol that obtained 80 % amino acid sequence coverage of the α_1 subunit of $GABA_A$ receptors (Chen *et al.*, 2012a; Chen *et al.*, 2012b), was not successful. This can be attributed to the fresh batch of Lys-C enzyme used for the experiment that was later determined to be inactive, so the potential of this protocol in increasing amino acid coverage of GLIC has not yet been tested adequately.

Despite labelled peptides within the transmembrane region not being identified, there is still evidence pointing towards the labelling of amino acids in that region from the reduced concentration of native peptides within the transmembrane domain when GLIC was exposed to light in the presence of photofol (section 5.4). One such peptide, $_{239}$NILVETNLPKTPY$_{251}$, contains the residue valine-242 that had been identified to be a determinant of propofol sensitivity, whilst $_{233}$IAHIAF$_{238}$ may be labelled at histidine-235 or phenylalanine-238 facing the intra-subunit binding cavity, and this peptide, along with $_{228}$VVSTL$_{232}$ contain residues such as serine-230, isoleucine-233 and isoleucine-236 facing the channel pore, which had been predicted to bind propofol through computational studies (LeBard *et al.*, 2012).

In conclusion, only weak evidence for the photoaffinity labelling of the crystallography-confirmed propofol-binding site in the transmembrane domain was provided in this study through the

reduction in ionisation current of peptides located in the binding site vicinity, when GLIC was exposed to light in the presence of photofol. There is reproducible labelling of the amino acids lysine-170, lysine-183 and cysteine-27 in the extracellular domain. These amino acids, particularly lysine-183, are hypothesised in this report to be at or near a potential binding pocket of propofol at the interface between subunits, which can be seen in the X-ray structure of opened-pore GLIC and is of a greater volume in the closed-pore conformation ELIC. The hypothesis of this binding groove at the extracellular subunit interface being a binding site for propofol is currently being tested using patch clamp experiments on oocytes expressing the following single amino acid mutants of GLIC: F174A, L175W, A172W, Y129A and I131W by Professor Andrew Jenkins. Recently, the photolabelled extracellular domain site has been confirmed to bind another intravenous anaesthetic, ketamine, from a high-resolution crystal structure. Ketamine also inhibits the proton-activated current of GLIC, therefore the binding of propofol at the same site may act via the same mechanism.

6 Overall conclusion

The aim of this project was to verify that the novel propofol diazirine analogue photofol was a suitable photolabel for labelling amino acids in known binding sites of propofol. Photoaffinity labelling is a technique that is increasingly being used in the study of structural-function relationships in biological systems for many reasons. Firstly, it is more accessible to membrane proteins where crystallography study is difficult. It is also advantageous over mutagenesis studies, where the determinants of the ligand sensitivity of ion channels cannot be differentiated between interfering with channel-binding or –gating. Furthermore, it does not rely on preconceived hypotheses about which amino acids are important, unlike mutagenesis. Finally, photoaffinity labelling is a versatile technique, allowing the kinetics of ligand-binding to different physiological states of proteins to be studied.

During this project, the anaesthetic action of photofol was confirmed *in vitro* via potentiating GABA-evoked currents in HEK cells expressing $\alpha_1\beta_2\gamma_{2S}$ $GABA_A$ receptors and *in vivo* via loss of righting reflex assays in rats. There is reduced modulation of GABA-evoked currents by 1.9-fold for the concentration of photofol tested and a 3.1-fold decrease in potency in inducing loss of righting reflex compared to propofol. These could be attributed to weaker binding of photofol to propofol-binding sites due to the presence of a bulkier photoactivatable group compared to an isopropyl arm in propofol at the *ortho* position to the hydroxyl group of the anaesthetic, a group which plays a crucial role in forming hydrogen bond interactions with their targets.

Through photoaffinity labelling experiments with HSA and GLIC, photofol has labelled six different amino acids, including phenylalanine, demonstrating the ability of the photoactivatable moiety of photofol to insert into unreactive C–H bonds. The novelty use of two samples of photofol, one with a mixture of a bespoke ratio of different isotopes, has aided the analysis of mass spectra, allowing unambiguous assignment of peaks corresponding to unlabelled and labelled peptides or CID fragments.

Photofol has successfully labelled previously identified propofol-binding sites of HSA and identified an extra binding site that was initially missed in viewing the electron density data. A potential binding site of propofol in the extracellular domain of GLIC that is not in the published

crystallographic data was also identified along with weak evidence supporting the labelling of the propofol-binding site in the transmembrane domain of its X-ray structure. The extracellular domain site has recently been shown to bind ketamine, an anaesthetic that also inhibits proton-activated currents of GLIC, supporting the hypothesis that propofol may bind there despite its presence not being resolved in the crystal structure. Further study is required to confirm the potential of the extracellular site as being a site of propofol binding via mutagenesis studies.

One of the reasons for not locating labelled peptides of the transmembrane region of GLIC is due to its low amino acid coverage as a result of difficulties with sample preparation and the relatively low sensitivity of the mass spectrometer used. A direction that this project has taken is to collaborate with Professor Alex Evers and his colleagues, who have successfully sequenced $GABA_A$ α_1 receptors with 80 % amino acid coverage, in providing them with the pPD and dPD samples to label $GABA_A$ receptors.

The ultimate aim of this work was to identify unknown general anaesthetic-binding sites in established general anaesthesia targets. In this project, photofol has been shown to act and bind in a similar manner to propofol, labelling various amino acids within crystallographically-confirmed anaesthetic-binding sites of HSA and GLIC, demonstrating that it is a successful photolabel. Future work using photofol to photolabel $GABA_A$ should indicate the binding sites of propofol on these well-established general anaesthetic targets, therefore advancing our understanding of the binding sites and mechanisms of these drugs.

References

Antkowiak, B. (2001) How do general anaesthetics work? *Naturwissenschaften,* 88(5), 201–213.

Baenziger, J.E. and Corringer, P.J. (2011) 3D structure and allosteric modulation of the transmembrane domain of pentameric ligand-gated ion channels. *Neuropharmacology*, 60(1), 116–25.

Bali, M. and Akabas, M.H. (2004) Defining the propofol binding site location on the $GABA_A$ receptor. *Mol Pharmacol,* 65(1), 68–76.

Bali, M., Jansen, M. and Akabas, M.H. (2009) GABA-induced intersubunit conformational movement in the $GABA_A$ receptor α1M1-β2M3 transmembrane subunit interface: experimental basis for homology modeling of an intravenous anesthetic binding site. *J Neurosci,* 29(10), 3083–3092.

Barber, M., Bordoli, R.S., Sedgwick, R.D., Tyler, A.N. and Bycroft, B.W. (1981) Fast atom bombardment mass spectrometry of bleomycin A_2 and B_2 and their metal complexes. *Biochem Biophys Res Commun,* 101(2), 632–638.

Belelli, D., Callachan, H., Hill-Venning, C., Peters, J.A. and Lambert, J.J. (1996) Interaction of positive allosteric modulators with human and *Drosophila* recombinant GABA receptors expressed in *Xenopus laevis* oocytes. *Br J Pharmacol,* 118(3), 563–576.

Belelli, D., Lambert, J.J., Peters, J.A., Wafford, K. and Whiting, P.J. (1997) The interaction of the general anesthetic etomidate with the γ-aminobutyric acid type A receptor is influenced by a single amino acid. *Proc Natl Acad Sci U S A,* 94(20), 11031–11036.

Berezhnoy, D., Gibbs, T.T. and Farb, D.H. (2009) Docking of 1,4-benzodiazepines in the α_1/γ_2 $GABA_A$ receptor modulator site. *Mol Pharmacol,* 76(2), 440–450.

Bhattacharya, A.A., Curry, S. and Franks, N.P. (2000) Binding of the general anesthetics propofol and halothane to human serum albumin. High resolution crystal structures. *J Biol Chem,* 275(49), 38731–38738.

Biemann, K. (1992) Mass spectrometry of peptides and proteins. *Annu Rev Biochem,* Volume 61, 977–1010.

Bocquet, N., Nury, H., Baaden, M., Le Poupon, C., Changeux, J.P., Delarue, M. and Corringer, P.J. (2009) X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature,* 457(7225), 111–114.

Bocquet, N., Prado de Carvalho, L., Cartaud, J., Neyton, J., Le Poupon, C., Taly, A., Grutter, T., Changeux, J.P. and Corringer, P.J. (2007) A prokaryotic proton-gated ion channel from the nicotinic acetylcholine receptor family. *Nature,* 445(7123), 116–119.

Bright, D.P., Adham, S.D., Lemaire, L.C.J.M., Benavides, R., Gruss, M., Taylor, G.W., Smith, E.H. and Franks, N.P. (2007) Identification of anaesthetic binding sites on human serum albumin using a novel etomidate photolabel. *J Biol Chem,* 282(16), 12038–12047.

Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol,* 7(8), 2745-2752.

Chen, R., Belelli, D., Lambert, J.J., Peters, J.A., Reyes, A. and Lan, N.C. (1994) Cloning and functional expression of a *Drosophila* γ-aminobutyric acid receptor. *Proc Natl Acad Sci USA,* 91(13), 6069–6073.

Chen, Z.W., Fuchs, K., Sieghart, W., Townsend, R.R. and Evers, A.S. (2012a) Deep amino acid sequencing of native brain $GABA_A$ receptors using high-resolution mass spectrometry. *Mol Cell Proteomics,* Volume 11, 1–18.

Chen, Z.W., Manion, B., Townsend, R.R., Reichert, D.E., Covey, D.F., Steinbach, J.H., Sieghart, W., Fuchs, K. and Evers, A.S. (2012b) Neurosteroid analog photolabeling of a site in the third transmembrane domain of the β3 Subunit of the $GABA_A$ receptor. *Mol Pharmacol,* 82(3), 408–419.

Cheng, X., Wang, H., Grant, B., Sine, S.M. and McCammon, J.A. (2006) Targeted molecular dynamics study of C-loop closure and channel gating in nicotinic receptors. *PLoS Comput Biol,* 2(9), 1173–1184.

Chowdhry, V. and Westheimer, F.H. (1979) Photoaffinity labeling of biological systems. *Annu Rev Biochem,* Volume 48, 293–325.

Cohen, S.L. and Chait, B.T. (1996) Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal Chem,* 68(1), 31–37.

Corringer, P.J., Baaden, M., Bocquet, N., Delarue, M., Dufresne, V., Nury, H., Prevost, M. and Van Renterghem, C. (2010) Atomic structure and dynamics of pentameric ligand-gated ion channels: new insight from bacterial homologues. *J Physiol,* 588(Pt 4), 565–572.

Curry, S. (2009) Lessons from the crystallographic analysis of small molecule binding to human serum albumin. *Drug Metab Pharmacokinet,* 24(4), 342–357.

Curry, S. (2011) X-ray Crystallography of Albumin. In: M. Otagiri, ed. *Human Serum Albumin – New Insights on its Structural Dynamics, Functional Impacts and Pharmaceutical Applications.* Kumamoto, Japan: Sojo Univ. Publishing Center, 1–29.

Curry, S., Mandelkow, H., Brick, P. and Franks, N. (1998) Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat Struct Biol,* 5(9), 827–835.

Davies, P.A., Kirkness, E.F., Hales, T.G., (1997) Modulation by general anaesthetics of rat $GABA_A$ receptors comprised of $\alpha 1\beta 3$ and $\beta 3$ subunits expressed in human embryonic kidney 293 cells. *Br J Pharmacol,* 120(5), 899–909.

Dickinson, R., de Sousa, S.L., Lieb, W.R. and Franks, N.P. (2002) Selective synaptic actions of thiopental and its enantiomers. *Anesthesiology,* 96(4), 884–892.

Doering, W.v.E. and Knox, L.H. (1956) The reaction of carbalkoxycarbene with saturated hydrocarbons. *J. Am. Chem. Soc,* 78(19), 4947–4950.

Downie, D.L., Hall, A.C., Lieb, W.R. and Franks, N.P. (1996) Effects of inhalational general anaesthetics on native glycine receptors in rat medullary neurones and recombinant glycine receptors in *Xenopus* oocytes. *Br J Pharmacol,* 118(3), 493–502.

Duncalfe, L.L., Carpenter, M.R., Smillie, L.B., Martin, I.L. and Dunn, S.M. (1996) The major site of photoaffinity labeling of the γ-aminobutyric acid type A receptor by [$_3$H]flunitrazepam is histidine 102 of the α subunit. *J Biol Chem,* 271(16), 9209–9214.

Dureta, G., Van Renterghema, C., Weng, Y., Prevosta, M., Moraga-Cid, G., Huona, Sonner, J.M. and Corringer, P.-J. (2011) Functional prokaryotic–eukaryotic chimera from the pentameric ligand-gated ion channel family. *PNAS,* 108(29), 12143–12148.

Duret, G., Van Renterghem, C., Weng, Y., Prevost, M., Moraga-Cid, G., Huon, C., Sonner, J.M. and Corringer, P.J. (2011) Functional prokaryotic-eukaryotic chimera from the pentameric ligand-gated ion channel family. *Proc Natl Acad Sci USA,* 108(29), 12143–12148.

Eckenhoff, R.G. (1996a) Amino acid resolution of halothane binding sites in serum albumin. *J Biol Chem,* 271(26), 15521–15526.

Eckenhoff, R.G. (1996b) An inhalational anesthetic binding domain in the nicotinic acetylcholine receptor. *Proc Natl Acad Sci USA,* 93(7), 2807–2810.

Eckenhoff, R.G. and Shuman, H. (1993) Halothane binding to soluble proteins determined by photoaffinity labeling. *Anesthesiology,* 79(1), 96–106.

Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60(Pt12 Pt1), 2126–2132.

Ernst, M., Bruckner, S., Boresch, S. and Sieghart, W. (2005) Comparative models of $GABA_A$ receptor extracellular and transmembrane domains: important insights in pharmacology and function. *Mol Pharmacol,* 68(5), 1291–1300.

Falick, A.M., Hines, W.M., Medzihradszky, K.F., Baldwin, M.A. and Gibson, B.W. (1993) Low-mass ions produced from peptides by high-energy collision-induced dissociation in tandem mass spectrometry. *J Am Soc Mass Spctrom,* 4(11), 882–893.

Farrant, M. and Nusser, Z. (2005) Variations on an inhibitory theme: phasic and tonic activation of $GABA_A$ receptors. *Nat Rev Neurosci,* 6(3), 215–229.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science,* 246(4926), 64–71.

Franks, N.P. (2008) General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. *Nat Rev Neurosci,* 9(5), 370–386.

Franks, N.P. and Lieb, W. (1984) Do general anaesthetics act by competitive binding to specific receptors? *Nature,* 310(5978), 599–601.

Franks, N.P. and Lieb, W. (1991) Stereospecific effects of inhalational general anesthetic optical isomers on nerve ion channels. *Science,* 254(5030), 427–430.

Franks, N.P. and Lieb, W.R. (1985) Mapping of general anaesthetic target sites provides a molecular basis for cutoff effects. *Nature,* 316(6026), 349–351.

Franks, N.P. and Lieb, W.R. (1994) Molecular and cellular mechanisms of general anaesthesia. *Nature,* 367(6464), 607–614.

Friederich, P. and Urban, B.W. (1999) Interaction of intravenous anesthetics with human neuronal potassium currents in relation to clinical concentrations. *Anesthesiology,* 91(6), 1853–1860.

Fukami, S., Uchida, I., Takenoshita, M., Mashimo, T. and Yoshiya, I. (1999) The effects of a point mutation of the β_2 subunit of $GABA_A$ receptor on direct and modulatory actions of general anesthetics. *Eur J Pharmacol,* 368(2–3), 269–276.

Ganem, B., Li, Y.T. and Henion, J.D. (1991) Detection of noncovalent receptor-ligand complexes by mass spectrometry. *J Am Chem Soc,* 113(16), 6296–6298.

García, M.C., Hogenboom, A.C., Zappey, H. and Irth, H. (2002) Effect of the mobile phase composition on the separation and detection of intact proteins by reversed-phase liquid chromatography-electrospray mass spectrometry. *J Chromatogr A,* 957(2), 187–199.

Ghuman, J., Zunszain, P.A., Petitpas, I., Bhattacharya, A.A., Otagiri, M. and Curry, S. (2005) Structural basis of the drug-binding specificity of human serum albumin. *J Mol Biol,* 353(1), 38–52.

Gonzalez-Gutierrez, G. and Grosman, C. (2010) Bridging the gap between structural models of nicotinic receptor superfamily ion channels and their corresponding functional states. *J Mol Biol,* 403(5), 693–705.

Guertin, P.A. and Hounsgaard, J. (1999) Non-volatile general anaesthetics reduce spinal activity by suppressing plateau potentials. *Neuroscience,* 88(2), 353–358.

Günther, U., Benson, J., Benke, D., Fritschy, J.M., Reyes, G., Knoflach, F., Crestani, F., Aguzzi, A., Arigoni, M., Lang, Y., Bluethmann, H., Möhler, H. and Luscher, B. (1995) Benzodiazepine-insensitive mice generated by targeted disruption of the γ_2 subunit gene of γ-aminobutyric acid type A receptors. *Proc Natl Acad Sci USA,* 92(17), 7749–7753.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H. and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol,* 17(10), 994–999.

Hales, T.G., Deeb, T.Z., Tang, H., Bollan, K.A., King, D.P., Johnson, S.J. and Connolly, C.N. (2006) An asymmetric contribution to γ-aminobutyric type A receptor function of a conserved lysine within TM2-3 of α1, β2, and γ2 subunits. *J Biol Chem,* 281(25), 17034–17043.

Hales, T.G. and Lambert, J.J. (1991) The actions of propofol on inhibitory amino acid receptors of bovine adrenomedullary chromaffin cells and rodent central neurones. *Br J Pharmacol,* 104(3), 619–628.

Hall, M.A., Xi, J., Lor, C., Dai, S., Pearce, R., Dailey, W.P. and Eckenhoff, R.G. (2010) *m*-Azipropofol (AziP*m*) a photoactive analogue of the intravenous general anesthetic propofol. *J Med Chem,* 53(15), 5667–5675.

Hansen, S.B., Sulzenbacher, G., Huxford, T., Marchot, P., Taylor, P. and Bourne, Y. (2005) Structures of *Aplysia* AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *EMBO J,* 24(20), 3635–3646.

Herbert, C.G. and Johnstone, R.A.W. (2003) *Mass spectrometry basics.* Boca Raton, Florida: CRC Press LLC.

He, X.M. and Carter, D.C. (1992) Atomic structure and chemistry of human serum albumin. *Nature,* 358(6383), 209–15.

Heyne, B., Tfibel, F., Hoebeke, M., Hans, P., Maurel, V. and Fontaine-Aupart, M.P. (2006) Photochemistry of 2,6-diisopropylphenol (propofol). *Photochem Photobiol Sci,* 5(11), 1059–1067.

Hibbs, R.E. and Gouaux, E. (2011) Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature,* 474(7349), 54–60.

Hilf, R.J., Bertozzi, C., Zimmermann, I., Reiter, A., Trauner, D. and Dutzler, R. (2010) Structural basis of open channel block in a prokaryotic pentameric ligand-gated ion channel. *Nat Struct Mol Biol,* 17(11), 1330–1336.

Hilf, R.J. and Dutzler, R. (2008) X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature,* 452(7185), 375–379.

Hilf, R.J. and Dutzler, R. (2009) Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature,* 457(7225), 115–119.

Hill-Venning, C., Belelli, D., Peters, J.A. and Lambert, J.J. (1997) Subunit-dependent interaction of the general anaesthetic etomidate with the γ-aminobutyric acid type A receptor. *Br J Pharmacol,* 120(5), 749–756.

Jones, M.V., Harrison, N.L., Pritchett, D.B. and Hales, T.G. (1995) Modulation of the $GABA_A$ receptor by propofol is independent of the γ subunit. *J Pharmacol Exp Ther,* 274(2), 962–968.

Jurd, R., Arras, M., Lambert, S., Drexler, B., Siegwart, R., Crestani, F., Zaugg, M., Vogt, K.E., Ledermann, B., Antkowiak, B. and Rudolph, U. (2003) General anesthetic actions *in vivo* strongly attenuated by a point mutation in the $GABA_A$ receptor β3 subunit. *FASEB J,* 17(2), 250–252.

Kawahara, K., Kuniyasu, A., Masuda, K., Ishiguro, M. and Nakayama, H. (2002) Efficient identification of photolabelled amino acid residues by combining immunoaffinity purification with MS: revealing the semotiadil-binding site and its relevance to binding sites for myristates in domain III of human serum albumin. *Biochem J,* 336(Pt 2), 223–232.

Klammer, A.A. and MacCoss, M.J. (2006) Effects of modified digestion schemes on the identification of proteins from complex mixtures. *J Proteome Res,* 5(3), 695–700.

Koblin, D.D., Chortkoff, B.S., Laster, M.J., Eger, E.I., 2nd, Halsey, M.J. and Ionescu, P. (1994) Polyhalogenated and perfluorinated compounds that disobey the Meyer-Overton hypothesis. *Anesth Analg,* 79(6), 1043–1048.

Kopp Lugli, A., Yost, C.S. and Kindler, C.H. (2009) Anaesthetic mechanisms: update on the challenge of unravelling the mystery of anaesthesia. *Eur J Anaesthesiol,* 26(10), 807–820.

Krasowski, M.D., Jenkins, A., Flood, P., Kung, A.Y., Hopfinger, A. and Harrison, N.L. (2001a) General anesthetic potencies of a series of propofol analogs correlate with potency for potentiation of γ-aminobutyric acid (GABA) current at the $GABA_A$ receptor but not with lipid solubility. *J Pharmacol Exp Ther,* 297(1), 338–351.

Krasowski, M.D., Koltchine, V.V., Rick, C.E.Y.Q., Finn, S.E. and Harrison, N.L. (1998) Propofol and other intravenous anesthetics have sites of action on the γ-aminobutyric acid type A receptor distinct from that for isoflurane. *Mol Pharmacol,* 53(3), 530–538.

Krasowski, M.D., Nishikawa, K., Nikolaeva, N., A, L. and Harrison, N.L. (2001b) Methionine 286 in transmembrane domain 3 of the $GABA_A$ receptor β subunit controls a binding cavity for propofol and other alkylphenol general anesthetics. *Neuropharmacology,* 41(8), 952–964.

Krasowski, M.D., O'Shea, S.M., Rick, C.E., Whiting, P.J., Hadingham, K.L., Czajkowski, C. and Harrison, N.L. (1997) α subunit isoform influences $GABA_A$ receptor modulation by propofol. *Neuropharmacology,* 36(7), 941–949.

LaBella, F.S. (1981) Is there a general anesthesia receptor. *Can J Physiol Pharmacol,* 59(5), 432–442.

Larrabee, M.G. and Posternak, J.M. (1952) Selective action of anesthetics on synapses and axons in mammalian sympathetic ganglia. *J Neurophysiol,* 15(2), 91–114.

LeBard, D.N., Hénin, J., Eckenhoff, R.G., Klein, M.L. and Brannigan, G. (2012) General anesthetics predicted to block the GLIC pore with micromolar affinity. *PLoS Comput Biol,* 8(5), e1002532.

Leon, I., Cocinero, E.J., Millán, J., Jaeqx, S., Rijs, A.M., A, L., Castaño, F. and Fernández, J.A. (2012) Exploring microsolvation of the anesthetic propofol. *Phys Chem Chem Phys,* 14(13), 4398–4409.

Li, G.D., Chiara, D.C., Cohen, J.B. and Olsen, R.W. (2010) Numerous classes of general anesthetics inhibit etomidate binding to γ-aminobutyric acid type A ($GABA_A$) receptors. *J Biol Chem,* 285(12), 8615–8620.

Li, G.D., Chiara, D.C., Sawyer, G.W., Husain, S.S., Olsen, R.W. and Cohen, J.B. (2006) Identification of a $GABA_A$ receptor anesthetic binding site at subunit interfaces by photolabeling with an etomidate analog. *J Neurosci,* 26(45), 11599–11605.

Lingamaneni, R., Birch, M.L. and Hemmings, H.C., Jr. (2001) Widespread inhibition of sodium channel-dependent glutamate release from isolated nerve terminals by isoflurane and propofol. *Anesthesiology,* 95(6), 1460–1466.

Liu, R., Meng, Q., Xi, J., Yang, J., Ha, C.E., Bhagavan, N.V. and Eckenhoff, R.G. (2004) Comparative binding character of two general anaesthetics for sites on human serum albumin. *Biochem J,* 380(Pt 1), 147–152.

Liu, R., Perez-Aguilar, J.M., Liang, D. and Saven, J.G. (2012) Binding site and affinity prediction of general anesthetics to protein targets using docking. *Anesth Analg,* 114(5), 947–955.

Liu, R., Yang, J., Ha, C.-E., Bhagavan, N.V. and Eckenhoff, R.G. (2005) Truncated human serum albumin retains general anaesthetic binding activity. *Biochem J,* 388(Pt 1), 39–45.

Luscher, B., Fuchs, T., Kilpatrick and L, C. (2011) $GABA_A$ receptor trafficking-mediated plasticity of inhibitory synapses. *Neuron,* 70(3), 385–409.

Machu, T.K. and Harris, R.A. (1994) Alcohols and anesthetics enhance the function of 5-hydroxytryptamine$_3$ receptors expressed in *Xenopus laevis* oocytes. *J Pharmacol Exp Ther,* 271(2), 898–905.

Maksay, G. (2009) Ligand-gated pentameric ion channels, from binding to gating. *Curr Mol Pharmacol,* 2(3), 253–262.

Masuda, T., Tomita, M. and Ishihama, Y. (2008) Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J Proteome Res,* 7(2), 731–740.

Matthews, E.K. and Quilliam, J.P. (1964) Effects of central depressant drugs upon acetylcholine release. *Br J Pharmacol Chemother,* Volume 22, 415–440.

Mazoit, J.X. and Samii, K. (1999) Binding of propofol to blood components: implications for pharmacokinetics and for pharmacodynamics. *Br J Clin Pharmacol,* 47(1), 35–42.

McKernan, R.M., Farrar, S., Collins, I., Emms, F., Asuni, A., Quirk, K. and Broughton, H. (1998) Photoaffinity labeling of the benzodiazepine binding site of $\alpha 1\beta 2\gamma 2$ γ-aminobutyric $acid_A$ receptors with flunitrazepam identifies a subset of ligands that interact directly with His102 of the α Subunit and predicts orientation of these within the benzodiazepine pharmacophore. *Mol Pharmacol,* 54(1), 33–43.

McKernan, R.M. and Whiting, P.J. (1996) Which $GABA_A$-receptor subtypes really occur in the brain? *Trends Neurosci,* 19(4), 139–143.

Meyer, H. (1899) Welche Eigenschaft der Anästhetica bedingt ihre narkotische Wirkung? [in German] *Arch Exp Pathol Pharmakol (Naunyn-Schmiedebergs),* Volume 42, 109–118.

Miyazawa, A., Fujiyoshi, Y. and Unwin, N. (2003) Structure and gating mechanism of the acetylcholine receptor pore. *Nature,* 423(6943), 949–955.

Möhler, H., Fritschy, J.M. and Rudolph, U. (2002) A new benzodiazepine pharmacology. *J Pharmacol Exp Ther,* 300(1), 2–8.

Nury, H., Van Renterghem, C., Weng, Y., Tran, A., Blaaden, M., Dufresne, V., Changeux, J.P., Sonner, J.M., Delarue, M. and Corringer, P.J. (2011) X-ray structures of general anaesthetics bound to a pentameric ligand-gated ion channel. *Nature,* 469(7330), 428–431.

Overton, E. (1901) *Studien über die Narkose, zugleich ein Beitrag zur allgemeinen Pharmakologie* [in German]. Jena: Fischer, G.

Pan, J., Chen, Q., Willenbring, D., Mowrey, D., Kong, X.-P., Cohen, A., Divito, C.B., Xu, Y. and Tang, P. (2012a) Structure of the pentameric ligand-gated ion channel GLIC bound with anesthetic ketamine. *Structure,* 20(9), 1463–1469.

Pan, J., Chen, Q., Willenbring, D., Yoshida, K., Tillman, T., Kashlan, O.B., Cohen, A., Kong, X.P., Xu, Y. and Tang, P. (2012b) Structure of the pentameric ligand-gated ion channel ELIC cocrystallized with its competitive antagonist acetylcholine. *Nat Commun,* 3(714), doi: 10.1038/ncomms1703.

Papayannopoulos, I.A. (1995) The interpretation of collision induced dissociation tandem mass spectra of peptides. *Mass Spectrometry Reviews,* 14(1), 49–73.

Parikh, R.B., Bali, M. and Akabas, M.H. (2011) Structure of the M2 transmembrane segment of GLIC, a prokaryotic Cys loop receptor homologue from *Gloeobacter violaceus*, probed by substituted cysteine accessibility. *J Biol Chem,* 286(16), 14098–14109.

Pistis, M., Belelli, D., McGurk, K., Peters, J.A. and Lambert, J.J. (1999) Complementary regulation of anaesthetic activation of human ($\alpha_6\beta_3\gamma_{2L}$) and *Drosophila* (RDL) GABA receptors by a single amino acid residue. *J Physiol,* 515(Pt 1), 3–18.

Plate, M.S. (1997) Photoaffinity labeling introduction. *Photochemistry and Photobiology,* 65(2), 193–194.

Ranquin, A. and Van Gelder, P. (2004) Maltoporin: sugar for physics and biology. *Res Microbiol,* 155(8), 661–666.

Richardson, J.E., Garcia, P.S., O'Toole, K.K., Derry, J.M., Bell, S.V. and Jenkins, A. (2007) A conserved tyrosine in the β2 subunit M4 segment is a determinant of γ-aminobutyric acid type A receptor sensitivity to propofol. *Anesthesiology,* 107(3), 412–418.

Robinson, D.H. and Toledo, A.H. (2012) Historical development of modern anesthesia. *J Invest Surg,* 25(3), 141–149.

Sanna, E., Murgia, A., Casula, A. and Biggio, G. (1997) Differential subunit dependence of the actions of the general anesthetics alphaxalone and etomidate at γ-aminobutyric acid type A receptors expressed in *Xenopus laevis* oocytes. *Mol Pharmacol,* 51(3), 484–490.

Schmidt, A., Kellermann, J. and Lottspeich, F. (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics,* 5(1), 4–15.

SIB ExPASy Bioinformatics Resource Portal, (2011) *PeptideCutter.* [Online] Available at: http://web.expasy.org/peptide_cutter/ [Accessed 25 July 2012].

Siegwart, R., Jurd, R. and Rudolph, U. (2002) Molecular determinants for the action of general anesthetics at recombinant $\alpha_2\beta_3\gamma_2$ γ-aminobutyric $acid_A$ receptors. *J Neurochem,* 80(1), 140–148.

Sigma-Aldrich, n.d. *Buffer Reference Centre.* [Online] Available at: http://www.sigmaaldrich.com/life-science/core-bioreagents/biological-buffers/learning-center/buffer-reference-center.html [Accessed 27 July 2012].

Singh, A., Thornton, E.R. and Westheimer, F.H. (1962) The photolysis of diazoacetylchymotrypsin. *J Biol Chem,* Volume 237, 3006–3008.

Smith, R.A. and Knowles, J.R. (1973) Letter: Aryldiazirines. Potential reagents for photolabeling of biological receptor sites. *J Am Chem Soc,* 95(15), 5072–5073.

Smith, R.D., Loo, J.A., Edmonds, C.G., Barinaga, C.J. and Udseth, H.R. (1990) New developments in biochemical mass spectrometry: electrospray ionization. *Anal Chem,* 62(9), 882–899.

Speers, A.E. and Wu, C.C. (2007) Proteomics of integral membrane proteins—theory and application. *Chem Rev,* 107(8), 3687–3714.

Stewart, D.S., Savechenkov, P.Y., Dostalova, Z., Chiara, D.C., Ge, R., Raines, D.E., Cohen, J.B., Forman, S.A., Bruzik, K.S. and Miller, K.W. (2011) *p*-(4-Azipentyl)propofol: a potent photoreactive general anesthetic derivative of propofol. *J Med Chem,* 54(23), 8124–8135.

Sudlow, G., Birkett, D.J. and Wade, D.N. (1975) The characterization of two specific drug binding sites on human serum albumin. *Mol Pharmacol,* 11(6), 824–835.

Sudlow, G., Birkett, D.J. and Wade, D.N. (1976) Further characterization of specific drug binding sites on human serum albumin. *Mol Pharmacol,* 12(6), 1052–1061.

Sugio, S., Kashima, A., Mochizuki, S., Noda, M. and Kobayashi, K. (1999) Crystal structure of human serum albumin at 2.5 Å resolution. *Protein Eng,* 12(6), 439–446.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y. and Yoshida, T. (1988) Protein and polymer analyses up to *m/z* 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom,* 2(8), 151–153.

Tanelian, D.L., Kosek, P., Mody, I. and MacIver, M.B. (1993) The role of the $GABA_A$ receptor/chloride channel complex in anesthesia. *Anesthesiology,* 75(4), 757–776.

Tomlin, S.L., Jenkins, A., Lieb, W.R. and Franks, N.P. (1999) Preparation of barbiturate optical isomers and their effects on $GABA_A$ receptors. *Anesthesiology,* 90(6), 1714–1722.

Uchida, I., Li, L. and Yang, J. (1997) The role of the $GABA_A$ receptor α1 subunit N-terminal extracellular domain in propofol potentiation of chloride current. *Neuropharmacology,* 36(11-12), 1611–1627.

Vanlersberghe, C. and Camu, F. (2008) Etomidate and other non-barbiturates. *Handb Exp Pharmacol,* Volume 182, 267–282.

Vaughan, R.J. and Westheimer, F.H. (1969) A method for marking the hydrophobic binding sites of enzymes. An insertion into the methyl group of an alanine residue of trypsin. *J Am Chem Soc,* 91(1), 217–218.

Vedula, L.S., Brannigan, G., Economou, N.J., Xi, J., Hall, M.A., Liu, R., Rossi, M.J., Dailey, W.P., Grasty, K.C., Klein, M.L., Eckenhoff, R.G. and Loll, P.J. (2009) A unitary anesthetic binding site at high resolution. *J Biol Chem,* 284(36), 24176–24184.

Violet, J.M., Downie, D.L., Nakisa, R.C., Lieb, W.R. and Franks, N.P. (1997) Differential sensitivities of mammalian neuronal and muscle nicotinic acetylcholine receptors to general anesthetics. *Anesthesiology,* 86(4), 866–874.

Wang, H.-L., Cheng, X. and Sine, S.M. (2012) Intra-membrane proton binding site linked to activation of a bacterial pentameric ion channel. *J Biol Chem,* 287(9), 6482–6489.

Waud, D.R. (1972) On biological assays involving quantal responses. *J Pharmacol Exp Ther,* 183(3), 577–607.

Weng, Y., Yang, L., Corringer, P.J. and Sonner, J.M. (2010) Anesthetic sensitivity of the *Gloeobacter violaceus* proton-gated ion channel. *Anesth Analg,* 110(1), 59–63.

Yamakura, T. and Harris, R.A. (2000) Effects of gaseous anesthetics nitrous oxide and xenon on ligand-gated ion channels. Comparison with isoflurane and ethanol. *Anesthesiology,* 93(4), 1095–1101.

Yamasaki, K., Maruyama, T., Kragh-Hansen, U. and Otagiri, M. (1996) Characterization of site I on human serum albumin: concept about the structure of a drug binding site. *Biochim Biophys Acta,* 1295(2), 147–157.

Zimmermann, I. and Dutzler, R. (2011) Ligand activation of the prokaryotic pentameric ligand-gated ion channel ELIC. *PLoS Biol,* 9(6), e1001101.

Appendix

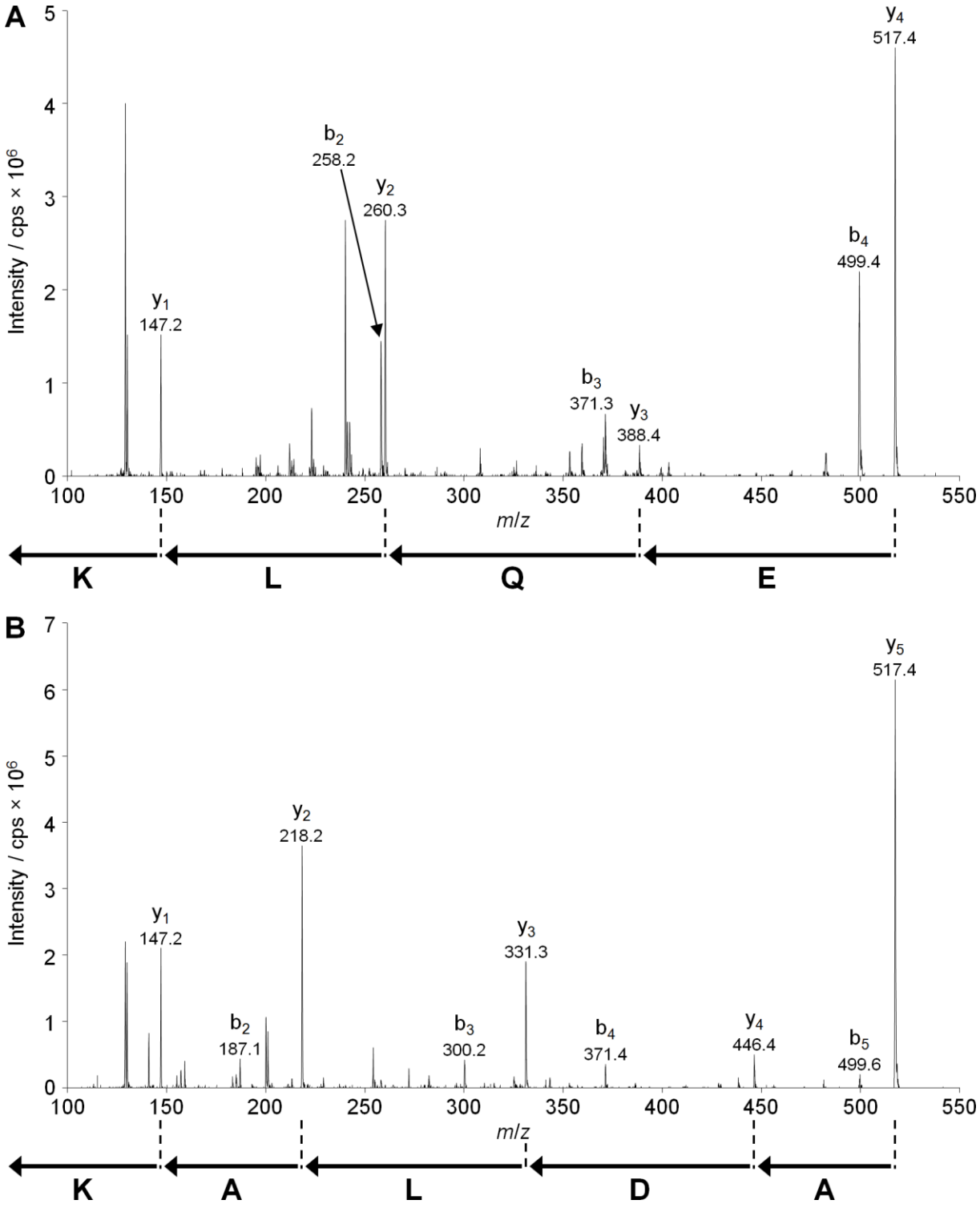


Figure A.1: Enhanced product ion spectra for two peptides with molecular weights of 516.3 Da eluted at 16.8 (A) and 18.3 minutes (B) respectively. The presence of b_n and y_n ions associated with collision-induced dissociation of the peptides $_{542}$EQLK$_{545}$ and $_{258}$ADLAK$_{262}$ confirms their sequences.

	Peptide sequence	MW / Da	m/z of ions (M + H)+	(M + 2H)2+	(M + 3H)3+	(M + 4H)4+	Retention time / mins
1–4	DAHK	469.2	*	*	*		29.6
5–10	SEVAHR	697.4	*	*	*		17.0
11–12	FK	293.2	*	*			13.0 / 16.3
13–20	DLGEENFK	950.4	*	*	*	*	37.0
21–41	ALVLIAFAQYLQQCPFEDHVK	2490.3		*	*	*	78.9
42–51	LVNEVTEFAK	1148.6	*	*	*		46.8
52–64	TCVADESAENCDK	1497.6		*	*		25.9
65–73	SLHTLFGDK	1016.5	*	*	*		48.1
74–81	LCTVATLR	932.5	*	*	*		40.4
82–93	ETYGEMADCCAK	1433.5	*	*	*		34.7
94–98	QEPER	657.3	*	*			15.3
99–106	NECFLQHK	1074.5	*	*	*		35.8
107–114	DDNPNLPR	939.4	*	*	*		30.4
115–136	LVRPEVDVMCTAFHDNEETFLK	2650.3		*	*		65.1
137	K	146.1					ND
138–144	YLYEIAR	926.5	*	*	*		45.2
145	R	174.1					ND
146–159	HPYFYAPELLFFAK	1742.9		*	*		72.6
160	R	174.1					ND
161–162	YK	309.2	*				4.4 / 7.0
163–174	AAFTECCQAADK	1370.6	*	*	*		32.1
175–181	AACLLPK	771.4	*	*	*		39.3
182–186	LDELR	644.3	*	*			31.2
187–190	DEGK	447.2	*	*			4.0
191–195	ASSAK	462.2	*	*			4.0
196–197	QR	302.2					ND
198–199	LK	259.2	*				4.4 / 7.0
200–205	CASLQK	705.3	*	*			21.9
206–209	FGER	507.2	*	*			16.3 / 17.8
210–212	AFK	364.2	*	*			36.4
213–218	AWAVAR	672.4	*	*			14.1
219–222	LSQR	502.3	*	*			24.9
223–225	FPK	390.2	*	*			33.6
226–233	AEFAEVSK	879.4	*	*	*		36.5
234–240	LVTDLTK	788.5	*	*	*		42.6
241–257	VHTECCHGDLLECADDR	2086.8		*	*	*	42.6
258–262	ADLAK	516.3	*	*			18.3
263–274	YICENQDSISSK	1442.6	*	*	*		32.7
275–276	LK	259.2	*				4.4 / 7.0
277–286	ECCEKPLLEK	1304.6	*	*	*		35.4
287–313	SHCIAEVENDEMPADLPSLAADFVESK	2974.3	*	*	*	*	70.8
314–317	DVCK	520.2	*	*			9.0
318–323	NYAEAK	694.3	*	*			20.1
324–336	DVFLGMFLYEYAR	1622.8		*	*		77.1
337	R	174.1					ND
338–348	HPDYSVVLLLR	1310.7	*	*	*		63.0
349–351	LAK	330.2	*	*			11.1
352–359	TYETTLEK	983.5	*	*			33.2
360–372	CCAADPHECYAK	1551.6		*	*		29.0
373–389	VFDEFKPLVEEPQNLIK	2045.1		*	*	*	63.5
390–402	QNCELFEQLGEYK	1656.7		*	*		56.9
403–410	FQNALLVR	959.6	*	*	*		47.6
411–413	YTK	410.2	*	*			10.5
414	K	146.1					ND
415–428	VPQVSTPTLVEVSR	1510.8		*	*		49.6
429–432	NLGK	430.3	*	*			14.7
433–436	VGSK	389.2	*	*			4.3
437–439	CCK	466.2					ND
440–444	HPEAK	580.3	*	*			12.9
445	R	174.1					ND
446–466	MPCAEDYLSVVLNQLCVLHEK	2518.2		*	*	*	78.5
467–472	TPVSDR	673.3	*	*			19.7
473–475	VTK	346.2	*	*			4.3
476–484	CCTESLVNR	1137.5	*	*	*		33.0
485–500	RPCFSALEVDETYVPK	1910.9		*	*	*	56.0
501–519	EFNAETFTFHADICTLSEK	2260.0		*	*	*	62.0
520–521	ER	303.2	*				3.9
522–524	QIK	387.2	*				9.6
525	K	146.1					ND
526–534	QTALVELVK	999.6	*	*	*		50.9
535–538	HKPK	508.3	*	*			5.3
539–541	ATK	318.2	*				3.8
542–545	EQLK	516.3	*	*			16.8
546–557	AVMDDFAAFVEK	1341.6	*	*	*		62.1
558–560	CCK	466.2					ND
561–564	ADDK	447.2	*	*			4.0
565–573	ETCFAEEGK	1069.4	*	*			29.9
574	K	146.1					ND
575–585	LVAASQAALGL	1012.6	*	*	*		52.5

Figure A.2: Retention times of HSA tryptic peptides eluted from the HPLC column with eluent containing TFA and increasing ACN concentration. Asterisks denote the presence of peaks corresponding to a particular charged ion. ND denotes peptides for which retention times were not determined.

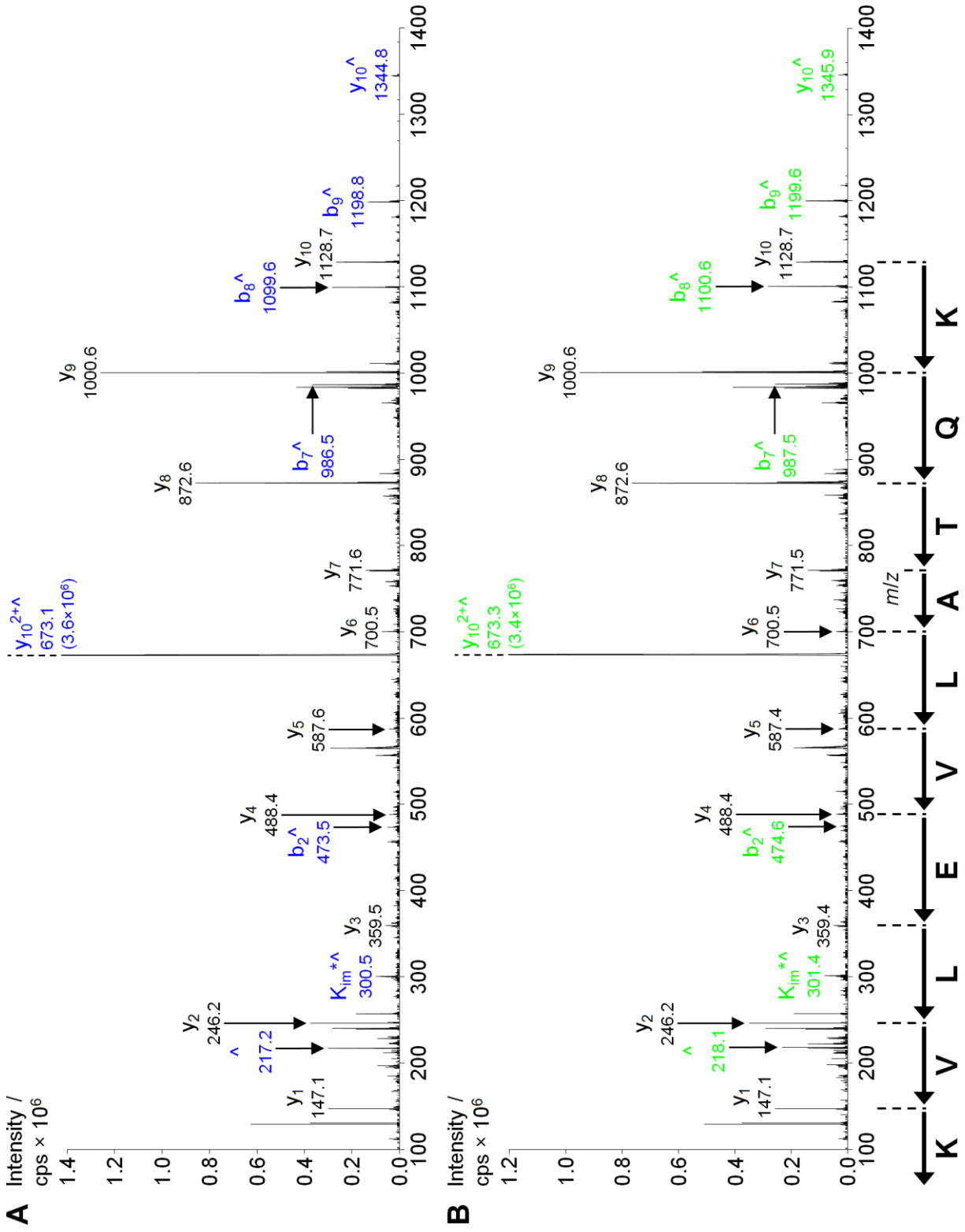


Figure A.3: Enhanced product ion spectra for labelled peptide $_{525}$KQTALVELVK$_{534}$ for the pPD- (A) and dPD-labelled (B) samples respectively showing that the label is attached to the first amino acid of this peptide. Carets indicate fragments containing the photolabel. K_{im}* and K_{aa} are internal lysine CID ions of 84 and 129 *m/z* respectively that were acknowledged in Falick *et al.*, 1993 and Papayannopoulos, 1995.

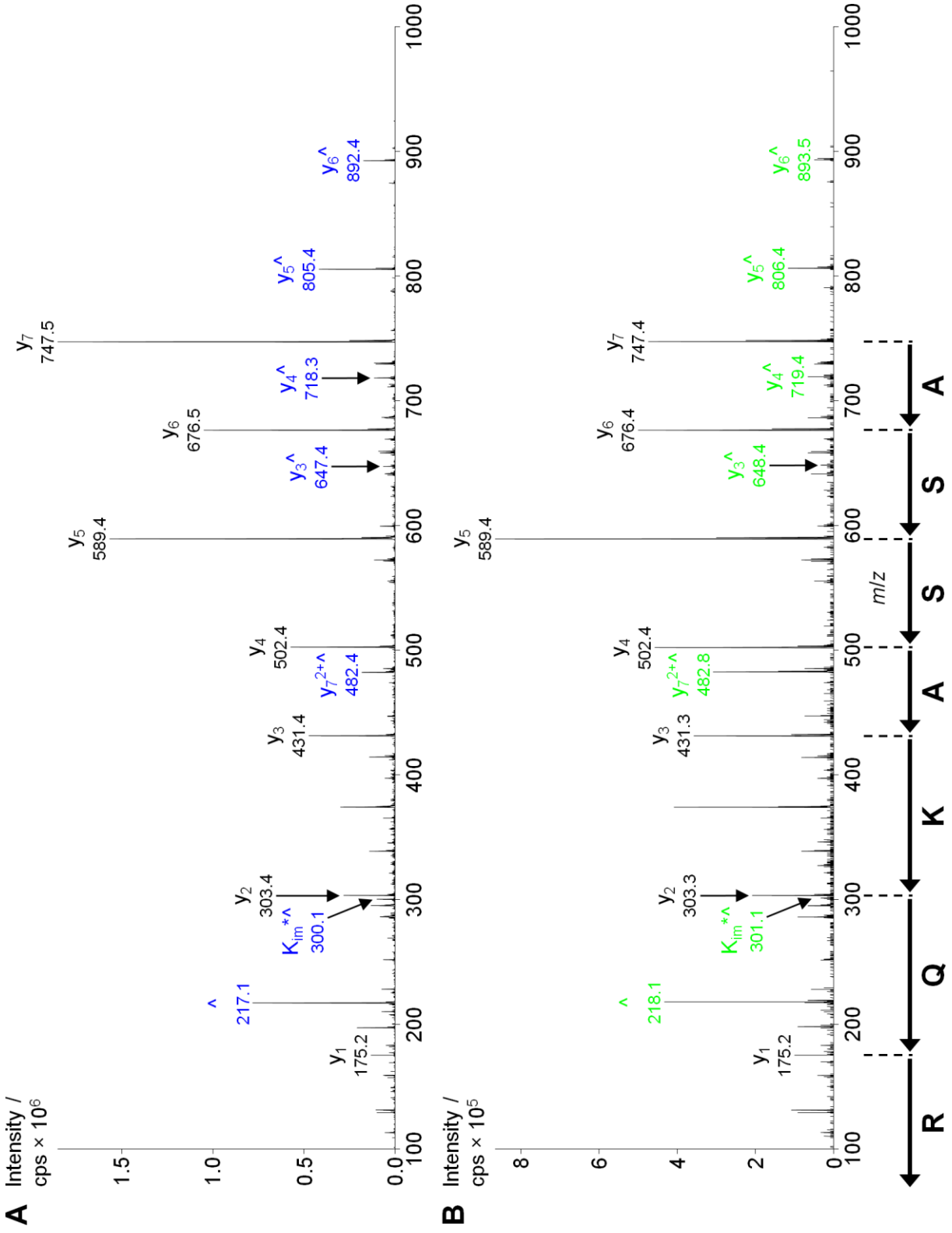


Figure A.4: Enhanced product ion spectra for labelled peptide $_{199}$ASSAKQR$_{205}$ for the pPD- (A) and dPD-labelled (B) samples respectively showing that the label is attached to the third amino acid of this peptide. Carets indicate fragments containing the photolabel. K_{im}* and K_{aa} are internal lysine CID ions of 84 and 129 *m/z* respectively that were acknowledged in Falick *et al.*, 1993 and Papayannopoulos, 1995.

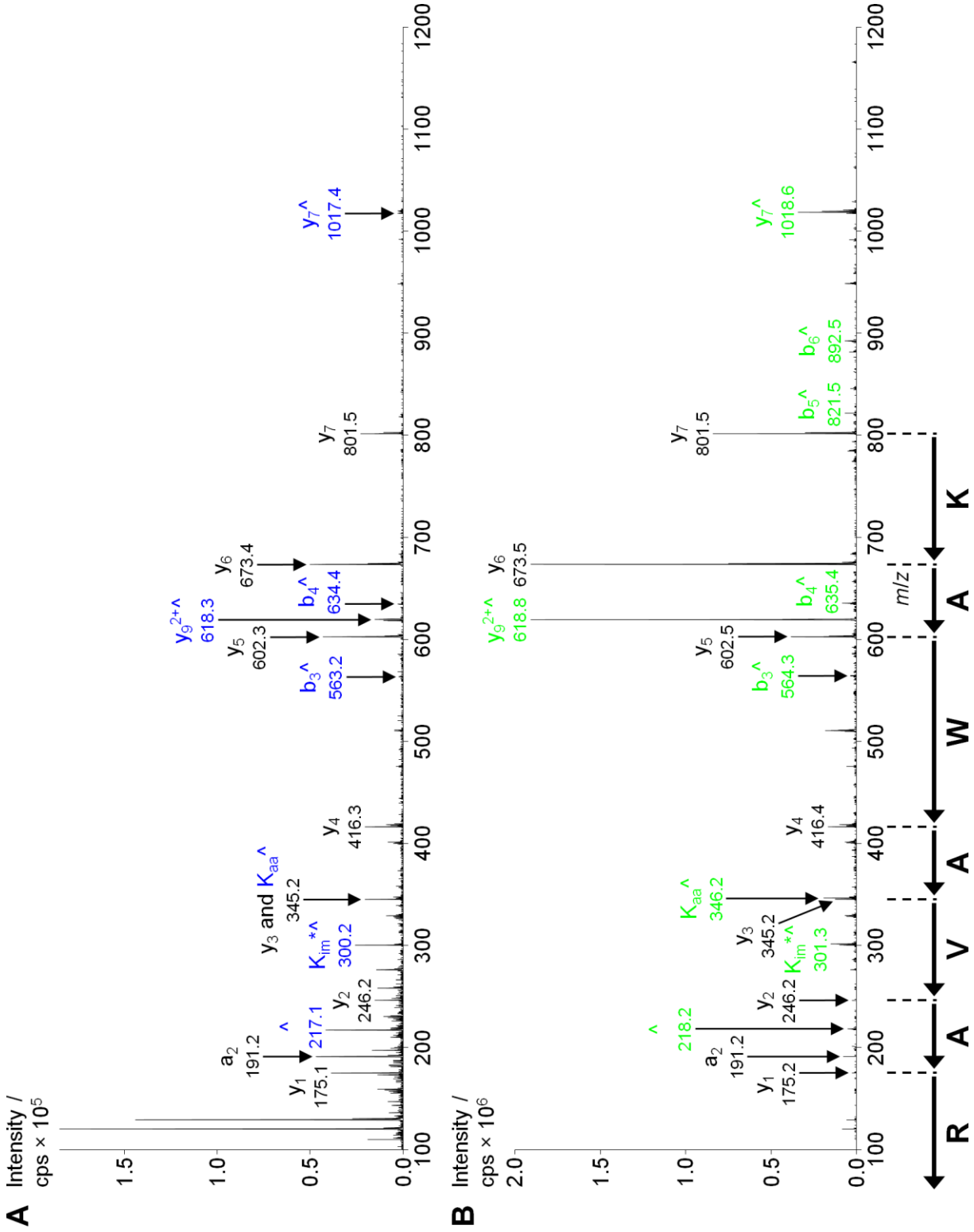


Figure A.5: Enhanced product ion spectra for labelled peptide $_{210}$AFKAWAVAR$_{218}$ for the pPD- (A) and dPD-labelled (B) samples respectively showing that the label is attached to the third amino acid of this peptide. Carets indicate fragments containing the photolabel. K_{im}* and K_{aa} are internal lysine CID ions of 84 and 129 *m/z* respectively that were acknowledged in Falick *et al.*, 1993 and Papayannopoulos, 1995.

		m/z of ions			Retention time / mins	Presence of peptide at different durations of chymotrypsin digestion													
	Peptide sequence	MH^+	MH^{2+}	MH^{3+}		4h20	6h40	9h	11h20	13h40	16h	18h20	20h40	23h	25h20	27h40	30h	32h20	34h40
1–23	**GQDMVSPPPPIADEPLTVNTGIY**		1206.1	804.4	—														
1–4	GQDM	450.2	225.6	150.7	—														
5–16	VSPPPPIADEPL	1231.7	**616.3**	*411.2*	51.2								*	*	*	*	*		
17–23	TVNTGIY	**767.4**	**384.2**	256.5	37.7	*	*	*	*	*	*	*	*	*	*	*	*	*	*
24–28	**LIECY**	**640.3**	**320.7**	214.1	43.8	*	*	*	*	*	*	*							
24	L	132.1			—														
25–28	IECY	527.2	264.1	176.4	—														
29–37	**SLDDKAETF**	**1025.5**	**513.2**	342.5	40.9	*	*	*	*	*	*	*	*	*	*	*	*	*	*
29–30	SL	219.1	110.1		—														
31–37	DDKAETF	825.4	413.2	275.8	—														
38–42	**KVNAF**	**578.3**	**289.7**	193.4	34.5	*	*	*	*	*	*	*	*	*	*	*	*	*	*
43–47	**LSLSW**	**605.3**	**303.2**	202.4	58.8	*	*	*	*	*	*	*	*	*	*	*	*	*	*
43	L	132.1			—														
44–45	SL	219.1	110.1		—														
46–47	SW	**292.1**	146.6		28.3	*	*	*	*	*	*	*	*	*	*	*	*	*	*
48–54	**KDRRLAF**	905.5	**453.3**	**302.5**	51.7	*	*	*	*	*	*	*	*	*	*	*	*	*	*
48–52	KDRRL	687.4	**344.2**	**229.8**	36.9	*	*	*	*	*	*	*	*	*	*	*	*	*	*
53–54	AF	237.1	119.1		—														
55–66	**DPVRSGVRVKTY**	1376.8	688.9	459.6	42.4	*	*	*	*	*	*	*	*	*	*	*	*	*	*
67–72	**EPEAIW**	**744.4**	**372.7**	248.8	50.3	*	*	*	*	*	*	*	*	*	*	*	*	*	*
73–78	**IPEIRF**	**774.5**	**387.7**	258.8	57.2	*	*	*	*	*	*	*	*	*	*	*	*	*	*
79–102	**VNVENARDADVVDISVSPDGTVQY**		**1281.6**	**854.7**	57.3	*	*	*	*	*	*	*	*	*	*	*	*	*	*
103–106	**LERF**	**564.3**	**282.7**	188.8	37.8	*	*	*	*	*	*	*	*	*	*	*	*	*	*
103	L	132.1			—														
104–106	ERF	451.2	226.1	151.1	—														
107–116	**SARVLSPLDF**	1104.6	**552.8**	**368.9**	40.5	*	*	*	*	*	*	*	*	*	*	*	*	*	*
107–111	SARVL	**545.3**	**273.2**	182.4	31.7	*	*	*	*	*	*	*	*	*	*	*	*	*	*
112–114	SPL	**316.2**	158.6	106.1	25.7	*	*	*	*	*	*	*	*	*	*	*	*	*	*
115–116	DF	**281.1**	141.1		22.3												*	*	*
117–121	**RRYPF**	**738.4**	**369.7**	246.8	45.8	*	*	*	*	*	*	*	*	*	*	*	*	*	*
122–129	**DSQTLHIY**	976.5	488.7	326.2	—														
122–126	DSQTL	**563.3**	**282.1**	188.4	23.6	*	*	*	*	*	*	*	*	*	*	*	*	*	*
127	H	156.1			—														
128–129	IY	295.2	148.1		—														
130–156	**LIVRSVDTRNIVLAVDLEKVGKNDDVF**	**1009.9 (MH^{3+})**	**757.1 (MH^{4+})**	**605.9 (MH^{5+})**	79.9	*	*	*	*	*									
130	L	132.1			—														
131–142	IVRSVDTRNIVL	1384.8	**692.9**	**462.3**	52.4			*	*	*	*	*	*	*	*	*	*	*	*
143–146	AVDL	**417.2**	209.1	139.7	48.3												*	*	*
147–156	EKVGKNDDVF	1150.6	575.8	384.2	—														
157–160	**LTGW**	**476.3**	238.6	159.4	43.4	*	*	*	*	*	*	*	*	*	*	*	*	*	*
157	L	132.1			—														
158–160	TGW	363.2	182.1	121.7	32.5			*	*	*	*	*	*	*	*	*	*	*	*
161–165	**DIESF**	**610.3**	**305.6**	204.1	42.0	*	*	*	*	*	*	*	*	*	*	*	*	*	*
166–174	**TAVVKPANF**	**946.5**	**473.8**	316.2	42.0	*	*	*	*	*	*	*	*	*	*	*	*	*	*
175–186	**ALEDRLESKLDY**	1451.7	**726.4**	**484.6**	52.5	*	*	*	*	*	*	*	*	*	*	*	*	*	*
175–176	AL	203.1	102.1		—														
177–180	EDRL	532.3	266.6	178.1	—														
181–184	ESKL	**476.1**	238.5	159.4	43.5	*	*	*	*	*	*	*	*	*	*	*	*	*	*
184–185	DY	**297.1**	149.1		13.8				*	*	*	*	*	*	*	*	*	*	*
187–194	**QLRISRQY**	1063.6	**532.3**	**355.2**	44.3	*	*	*	*	*	*	*	*	*	*	*	*	*	*
187–188	QL	260.2	130.6		—														
189–194	RISRQY	822.5	411.7	274.8	—														
195	**F**	166.1			—														

Figure A.6: Table showing the extract ions analysis of GLIC extracellular domain peptides from treatment A at various durations of chymotrypsin digestion. Peptide sequences arising from cleavages of only aromatic residues are in bold and those arising from further cleavages at leucine, methionine and histidine are listed underneath. The calculated *m/z* of the ions of these peptides within the range of 100–1500 *m/z* are listed and those that are observed in bold. Asterisks mark the presence of the peptide in the samples of a particular duration of digestion.

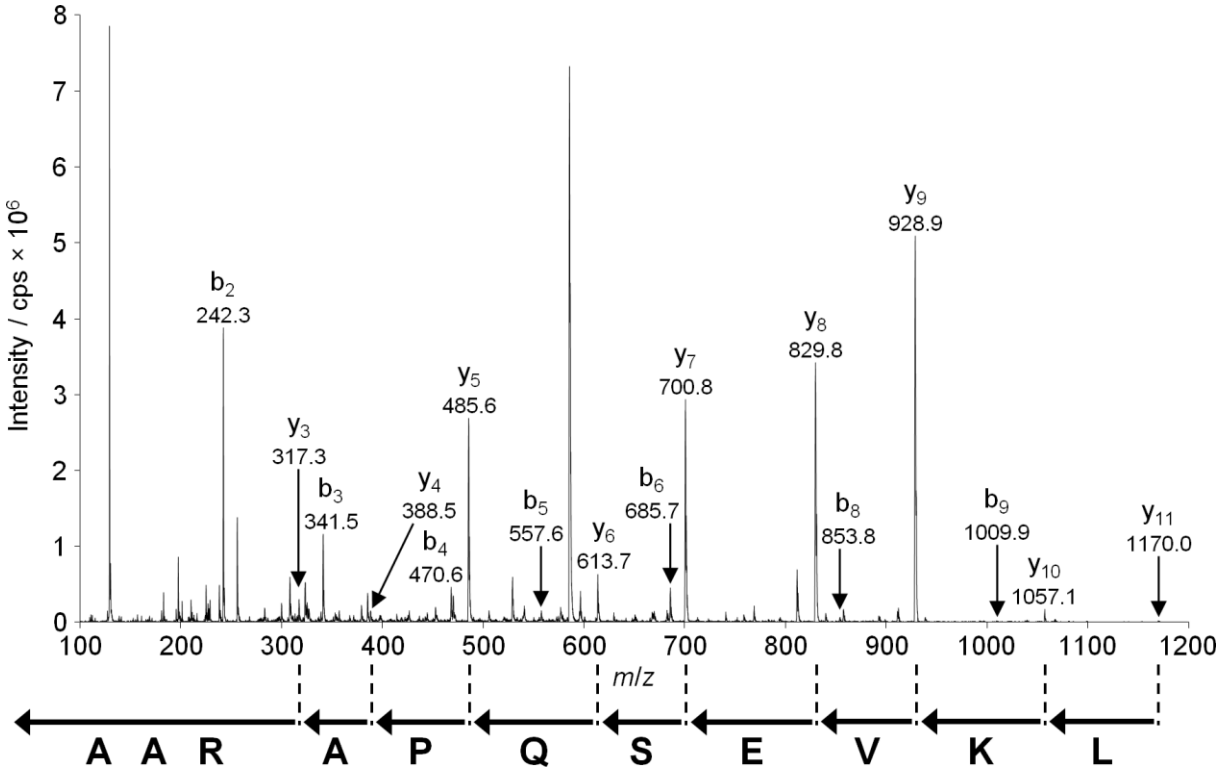


Figure A.7: Enhanced product ion spectra for the peptide with a MW of 1168.7 Da eluted at 31.8 minutes (peak 1 of Figure 5.10) confirms the peptide sequence as being $_{279}$LKVESQPARAA$_{289}$.

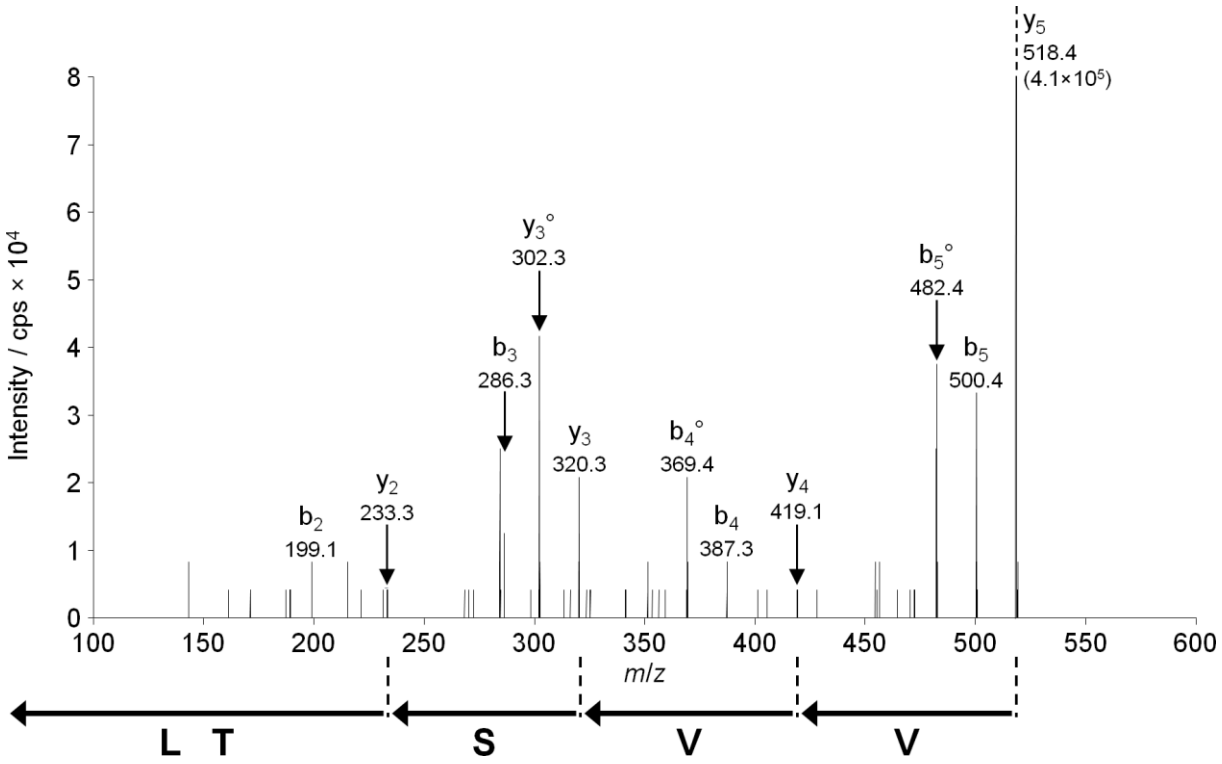


Figure A.8: Enhanced product ion spectra for the peptide with a MW of 517.2 Da eluted at 36.2 minutes (peak 2 of Figure 5.10) confirms the peptide sequence as being $_{228}$VVSTL$_{232}$.

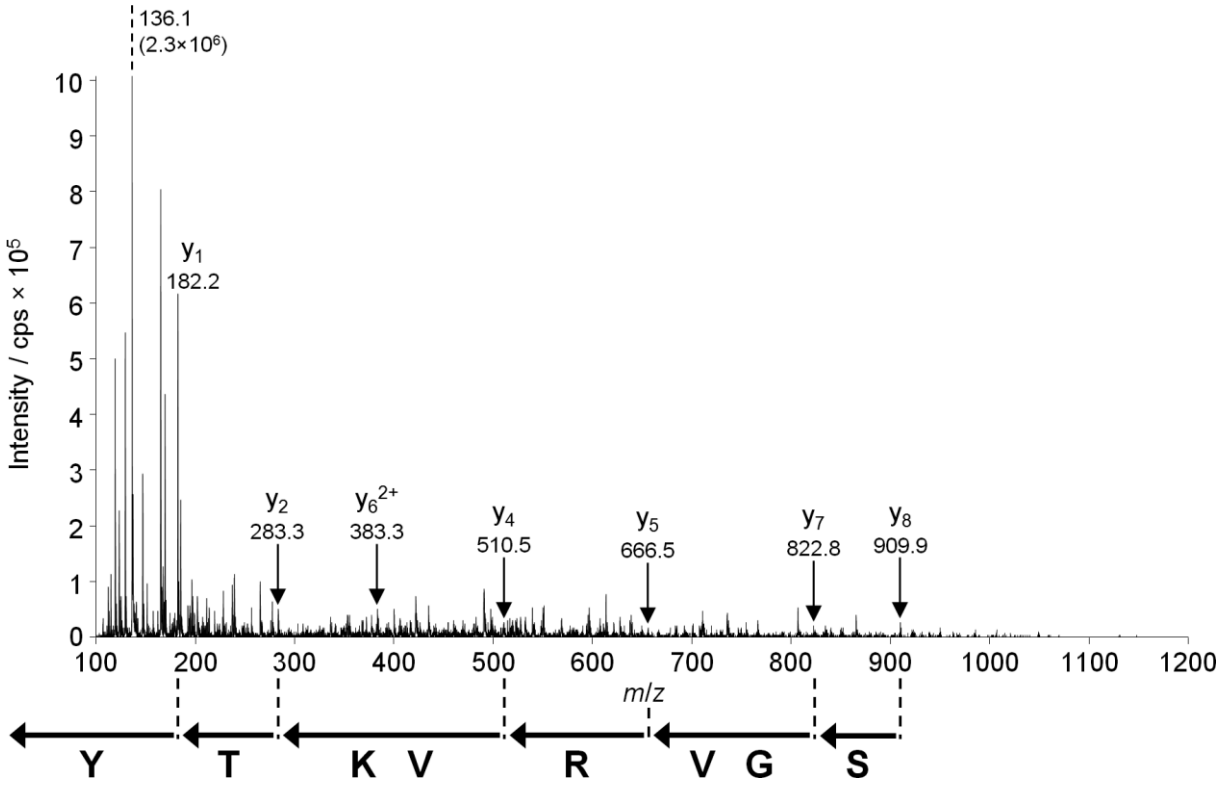


Figure A.9: Enhanced product ion spectra for the peptide with a MW of 1375.9 Da eluted at 37.6 minutes (peak 3 of Figure 5.10) confirms the peptide sequence as being $_{55}$DPVRSGVRVKTY$_{66}$.

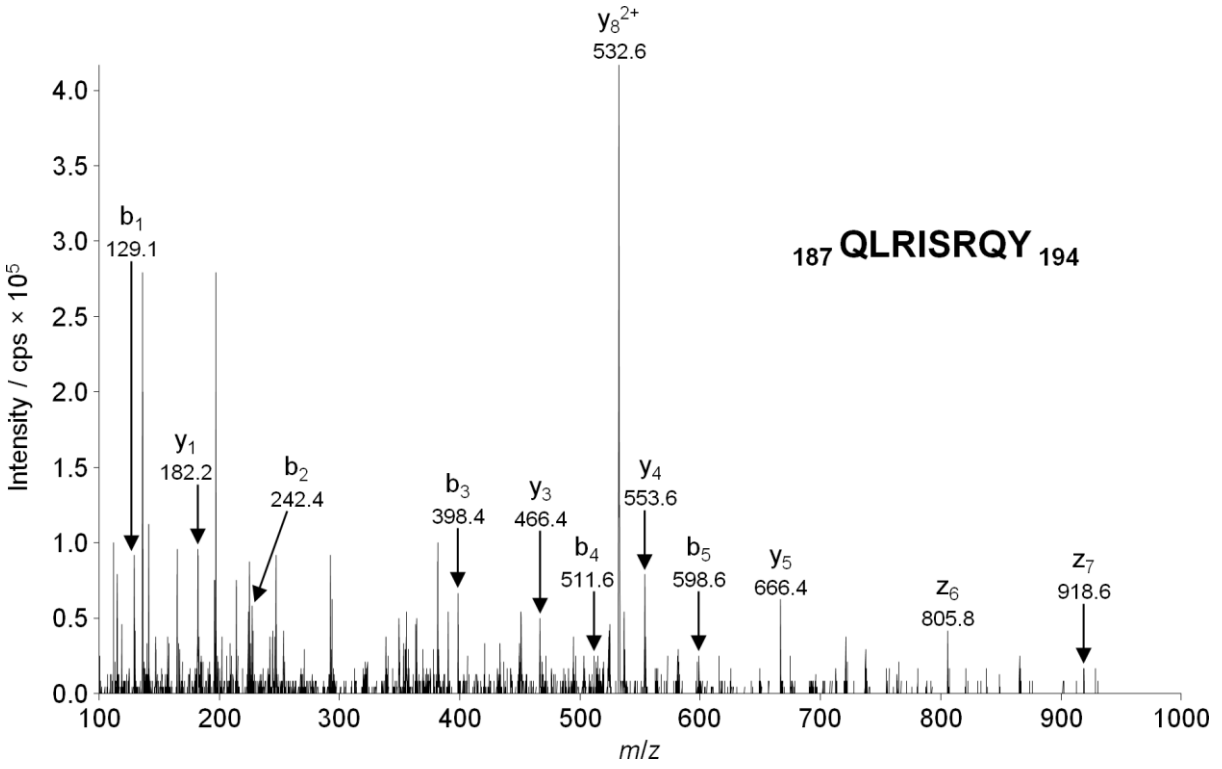


Figure A.10: Enhanced product ion spectra for the peptide with a MW of 1062.7 Da eluted at 39.7 minutes (peak 4 of Figure 5.10) confirms the peptide sequence as being $_{187}$QLRISRQY$_{194}$.

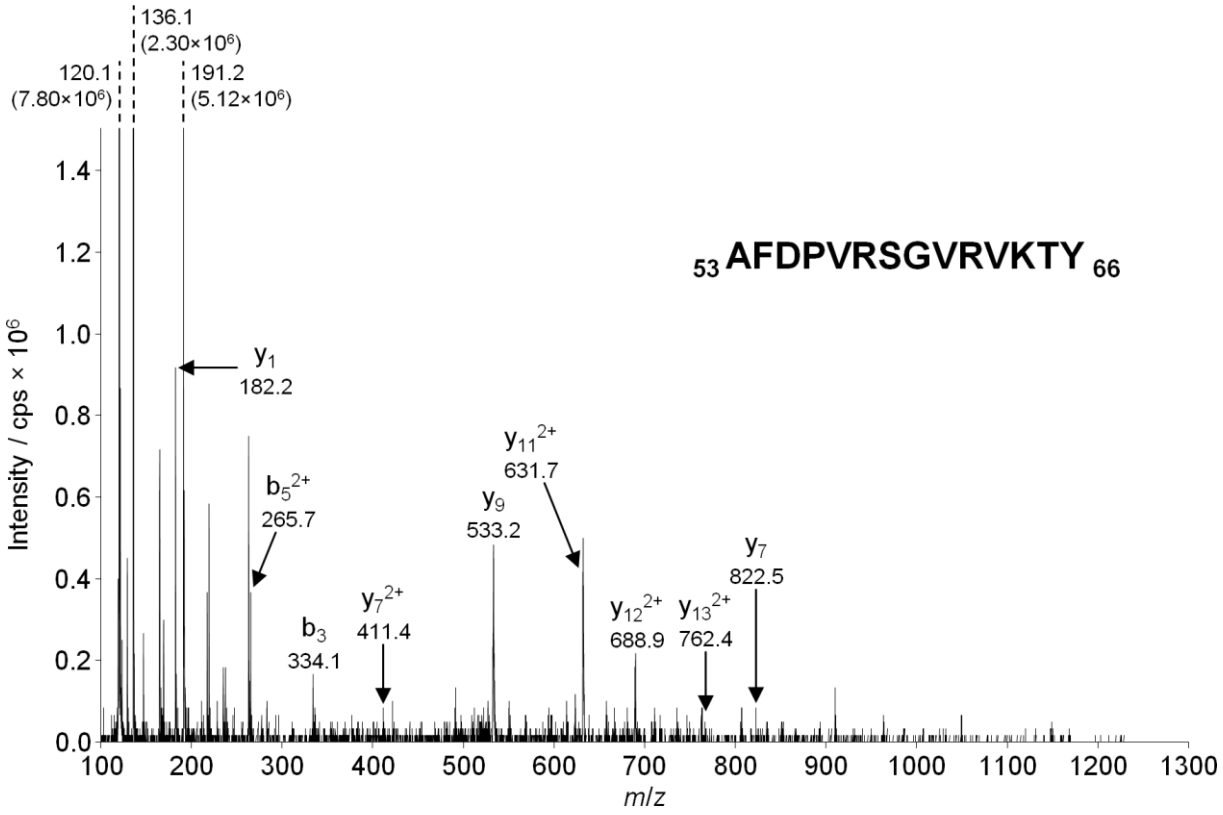


Figure A.11: Enhanced product ion spectra for the peptide with a MW of 1594.1 Da eluted at 49.8 minutes (peak 5 of Figure 5.10) confirms the peptide sequence as being $_{53}$AFDPVRSGVRVKTY$_{66}$.

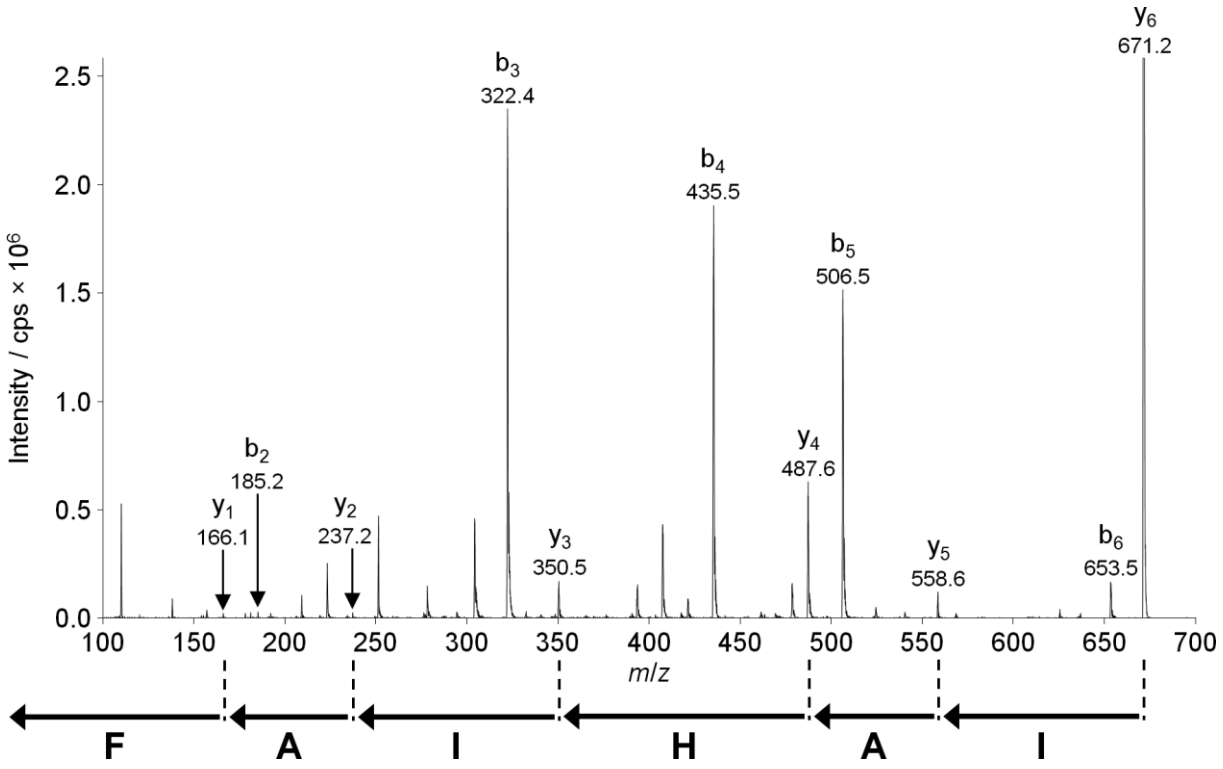


Figure A.12: Enhanced product ion spectra for the peptide with a MW of 670.3 Da eluted at 52.0 minutes (peak 6 of Figure 5.10) confirms the peptide sequence as being $_{233}$IAHIAF$_{238}$.

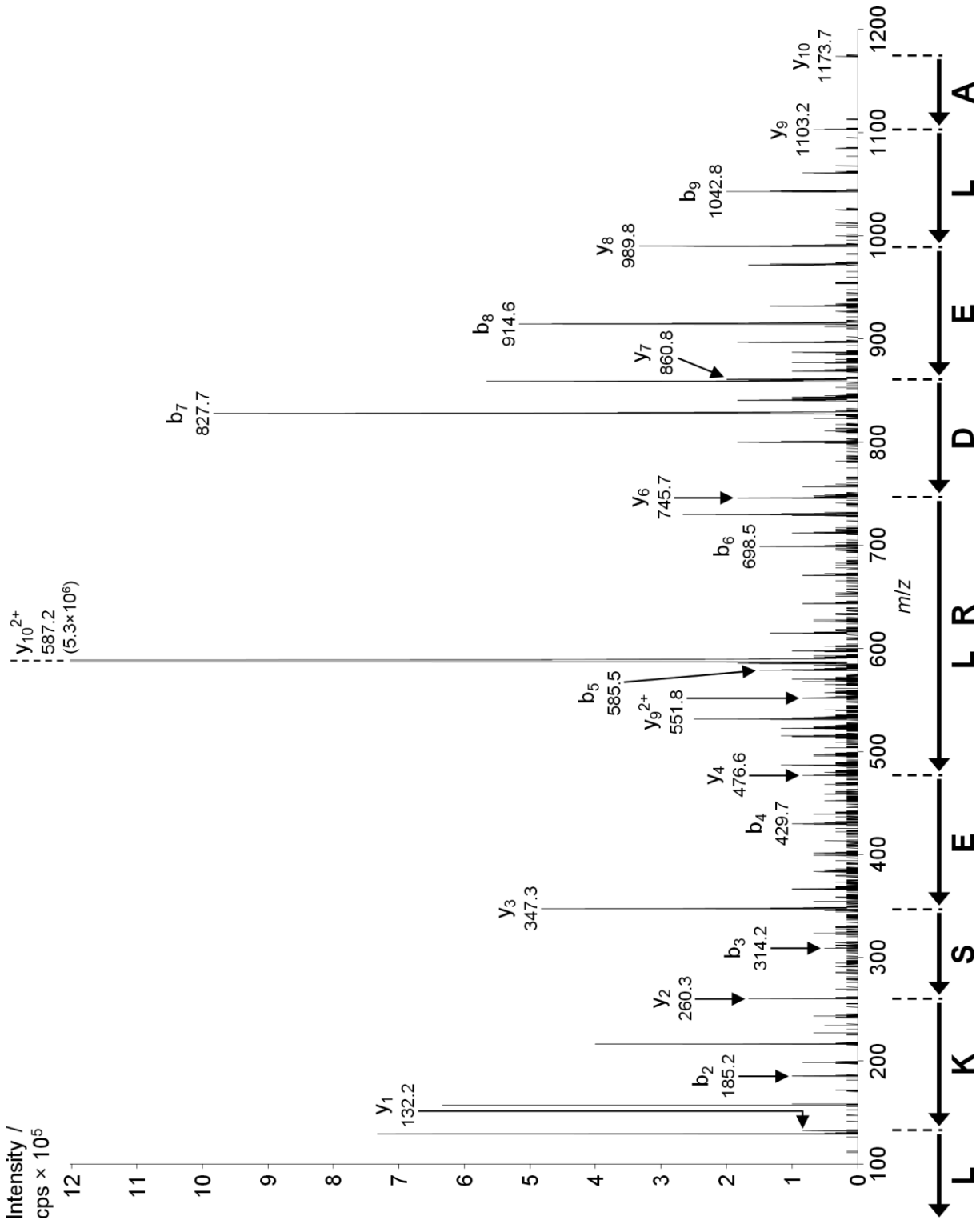


Figure A.13: Enhanced product ion spectra for the peptide with a MW of 1172.4 Da eluted at 52.8 minutes (peak 7 of Figure 5.10) confirms the peptide sequence to as being $_{175}$ALEDRLESKL$_{184}$.

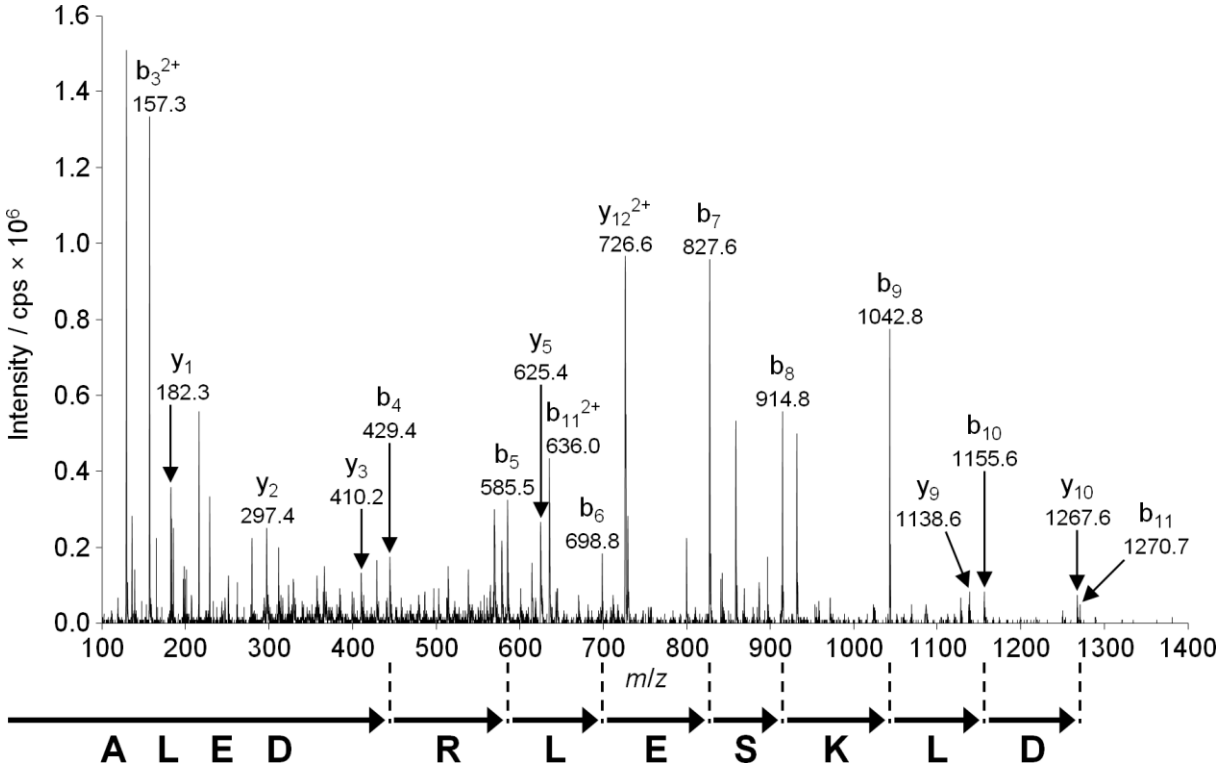


Figure A.14: Enhanced product ion spectra for the peptide with a MW of 1450.9 Da eluted at 56.8 minutes (peak 8 of Figure 5.10) confirms the peptide sequence as being $_{175}$ALEDRLESKLDY$_{186}$.

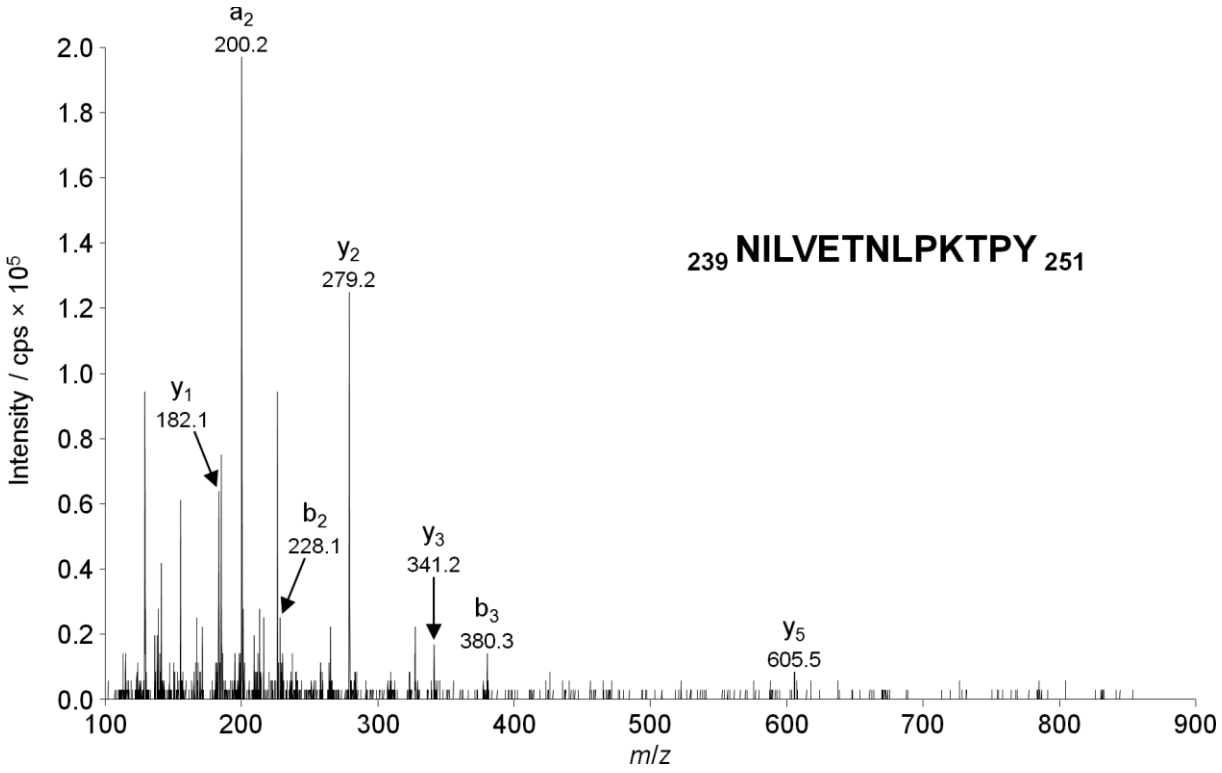


Figure A.15: Enhanced product ion spectra for the peptide with a MW of 1500.8 Da eluted at 63.4 minutes (peak 9 of Figure 5.10) confirms the peptide sequence as being $_{239}$NILVETNLPKTPY$_{251}$.

Actual MW / Da	Calculated MW / Da	Potential peptide
527.1	526.2	(L) $_{25}$ IECY $_{28}$
	527.3	(A) $_{71}$ IWIP $_{74}$
	527.3	(R) $_{110}$ VLSPL $_{114}$
	527.3	(V) $_{169}$ VKPAN $_{173}$
532.1	532.3	(S) $_{45}$ LSWK $_{48}$
	531.3	(D) $_{89}$ VVDIS $_{93}$
	531.3	(V) $_{90}$ VDISV $_{94}$
	531.3	(K) $_{149}$ VGKND $_{153}$
	531.3	(A) $_{176}$ LEDR $_{179}$
	531.3	(L) $_{177}$ EDRL $_{180}$
	531.3	(E) $_{178}$ DRLE $_{181}$
	532.3	(Y) $_{222}$ EANVT $_{226}$
	531.3	(V) $_{229}$ VSTLI $_{233}$
	531.3	(I) $_{298}$ AFPVV $_{302}$
633.1	632.3	(A) $_{54}$ FDPVR $_{58}$
	633.4	(I) $_{77}$ RFVNV $_{81}$
	632.3	(R) $_{86}$ DADVVD $_{91}$
	633.3	(A) $_{290}$ SITRAS $_{295}$
639.1	639.3	(I) $_{23}$ YLIEC $_{27}$
	639.3	(Y) $_{24}$ LIECY $_{28}$
	638.3	(R) $_{63}$ VKTYE $_{67}$
	638.3	(P) $_{248}$ KTPYM $_{252}$
945.6	946.5	(L) $_{53}$ AFDPVRSGV $_{61}$
	945.5	(A) $_{109}$ RVLSPLDF $_{116}$
	945.5	(R) $_{110}$ VLSPLDFR $_{117}$
	945.5	(I) $_{132}$ VRSVDTRN $_{139}$
	945.5	(S) $_{165}$ FTAVVKPAN $_{173}$
	945.5	(F) $_{166}$ TAVVKPANF $_{174}$
	945.5	(I) $_{259}$ IFMIYLF $_{265}$
1450.7	1451.7	$_{1}$ GQDMVSPPPPIADE $_{14}$
	1449.7	(I) $_{93}$ SVSPDGTVQYLER $_{105}$
	1450.7	(F) $_{175}$ ALEDRLESKLDY $_{186}$
	1449.9	(E) $_{223}$ ANVTLVVSTLIAHI $_{236}$
	1449.9	(A) $_{224}$ NVTLVVSTLIAHIA $_{237}$
	1450.8	(T) $_{254}$ YTGAIIFMIYLF $_{265}$
	1450.8	(Y) $_{255}$ TGAIIFMIYLFY $_{266}$
	1450.9	(F) $_{304}$ LLANIILAFLFFG $_{316}$
1692.2	1693.0	(L) $_{104}$ ERFSARVLSPLDFR $_{117}$
	1692.0	(P) $_{121}$ FDSQTLHIYLIVRS $_{134}$
	1691.9	(A) $_{144}$ VDLEKVGKNDDVFLT $_{158}$
	1692.9	(F) $_{175}$ ALEDRLESKLDYQL $_{188}$
1962.2	1962.3	(Y) $_{103}$ LERFSARVLSPLDFRR $_{118}$
	1962.2	(F) $_{175}$ ALEDRLESKLDYQLRI $_{190}$
	1962.2	(I) $_{212}$ SWTAFWSTSYEANVTLV $_{228}$
	1961.4	(T) $_{245}$ NLPKTPYMTYTGAIIFM $_{261}$
	1962.4	(M) $_{253}$ TYTGAIIFMI YLFYFV $_{268}$
2711.1	2710.1	(G) $_{2}$ QDMVSPPPPIADEPLTVNTGIYLIE $_{26}$
	2711.1	(L) $_{31}$ DDKAETFKVNAFLSLSWKDRRLA $_{53}$
	2711.1	(F) $_{55}$ DPVRSGVRVKTYEPEAIWIPEIR $_{77}$
	2711.0	(N) $_{84}$ ARDADVVDISVSPDGTVQYLERFSA $_{108}$
	2711.0	(V) $_{142}$ LAVDLEKVGKNDDVFLTGWDIESF $_{165}$
	2711.4	(Q) $_{188}$ LRISRQYFSYIPNIILPMLFIL $_{209}$
	2711.3	(Y) $_{198}$ IPNIILPMLFILFISWTAFWSTS $_{220}$
	2710.1	(S) $_{213}$ WTAFWSTSYEANVTLVVSTLIAHI $_{236}$
	2710.1	(F) $_{268}$ VAVIEVTVQHYLKVESQPARAASIT $_{292}$

Figure A.16: A list of peptides that match the molecular weights within 1.0 Da of the peaks that correspond to ions of labelled peptides from the single-stage mass spectrometry data obtained from the GLIC treatment D photoaffinity labelling experiment as summarised in Figure 5.21. Each peptide was generated from cleaving at the C-terminals of the amino acid shown in brackets and the last amino acid of its sequence.