



University
of Glasgow

Kelly, Robert Noel (2013) *Towards the absolute quantification of protein isoforms through the use of stable-isotope dilution mass spectrometry.*

PhD thesis

<http://theses.gla.ac.uk/4401/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

**Towards the Absolute Quantification of Protein Isoforms Through
the Use of Stable-Isotope Dilution Mass Spectrometry**

Robert Noel Kelly

M.Res. Proteomic Technologies,
M.Sc. Biotechnology,
B.Sc. Hons. Biology

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

College of Medical, Veterinary and Life Sciences
Institute of Molecular, Cell and Systems Biology
University of Glasgow

June 2013

Abstract

While the existence of protein was first described by Berzelius and Mulder back in 1838 and a single empirical formula noted ($C_{400}H_{620}N_{100}O_{120}P_1S_1$) (Vickery, 1950, Brand, 1946), early protein-based research was limited to the analysis of proteins which could be easily purified in large quantities, such as those obtained from blood, egg whites and those obtainable from slaughterhouses, such as digestive and metabolic enzymes (Chapman, 2005). Indeed, despite the development of recombinant deoxyribonucleic acid technologies in the 1970s (enabling protein expression) and the increasing sensitivity of techniques which enable the identification and sequencing of proteins separated by gel electrophoresis (Patterson and Aebersold, 2003), it was not until the late 1980s, with the description of soft biomolecule ionisation that large scale proteomic analyses were undertaken, based upon the use of mass spectrometry (Guerrera and Kleiner, 2005).

While early mass spectrometry-based proteomic analyses focussed on the systematic identification of a great number of proteins within a single organism, the field of proteomics is now becoming increasingly quantitative (Baak et al., 2005), enabling the relative comparison of protein expression patterns between phenotypes, but also the targeted absolute quantification of specific proteins.

During this project, a stable isotopically labelled internal standard based absolute quantitative technique, first described by Gerber and co-workers in 2003 (S. A. Gerber et al., 2003), was applied to the absolute quantification of three families of multiple protein isoforms. This area of research is of particular scientific interest as it is thought that up to 95% of human multi-exon genes may be subject to alternative splicing, making alternative splicing the rule, not the exception (Pan et al., 2008a). Indeed alternative splicing has also been implicated as both a cause and a consequence of disease. This technique should therefore enable both the confirmation of disease, based upon the identification of a set of phenotype specific protein biomarkers, but also the mapping of a disease's progression (Venables, 2004).

During this study, stable isotopically labelled internal standard peptides were selected for the absolute quantification of 11 confirmed protein isoforms, and two predicted protein isoforms. In addition, a separate MRM based LC-MS acquisition method was developed for the absolute quantification of each of the three families of protein isoforms (A-Raf, PDE4B and SERCA2) within a single analysis, and finally, these acquisition methods were applied to the absolute quantification of a range of immunoprecipitated, exogenously expressed protein isoforms. This project was, however, hindered by the sensitivity of the mass spectrometers available for use, preventing these acquisition methods from being applied to the absolute quantification of the endogenous levels of protein expression.

While beyond the scope of this project, the further development of this quantitative technique should enable future researchers to: (i) Quantify each endogenously expressed protein isoform within a family of multiple protein isoforms. (ii) Assess any changes in the expression of each isoform in a range of cellular states, and (iii) Assess how a targeted drug treatment may affect the expression ratio of these protein isoforms.

Table of Contents

1	Introduction	20
1.1	The Birth of Proteomics	20
1.2	Protein Separation	22
1.2.1	2DGE	22
1.2.2	Liquid Chromatography	25
1.3	MS Based Protein Identification	27
1.3.1	Top-down Proteomics	27
1.3.2	Bottom-up Proteomics	29
1.4	Protein identification	30
1.5	Mass Spectrometry	32
1.5.1	Biomolecule Ionisation	32
1.5.1.1	Matrix Assisted Laser Desorption/Ionisation	33
1.5.1.2	Electrospray Ionisation	35
1.5.2	Mass spectrometers	37
1.5.2.1	Time-of-Flight Mass Spectrometry	37
1.5.2.2	Quadrupole Mass Spectrometry	40
1.5.2.3	Fourier Transform-ion Cyclotron Resonance Mass Spectrometry	43
1.5.2.4	Orbitrap Mass Spectrometry	46
1.6	Protein Quantitation	49
1.6.1	Difference gel Electrophoresis	50
1.6.2	Stable Isotope Labelling by Amino Acids in Cell Culture	53
1.6.3	¹⁸ O Incorporation	55
1.6.4	Isotope-Coded Affinity Tags	57
1.6.5	Tandem Mass Tags	58
1.6.6	Label-Free Quantification	62
1.6.7	Labelled Internal Standard Based Quantification	64
1.6.7.1	Protein Standard Absolute Quantification	65
1.6.7.1.1	Proteotypic and Quantotypic Peptides	67
1.6.7.2	AQUA Based Peptide Quantification	67
1.6.7.3	QconCAT based Peptide Quantification	69
1.7	Protein Isoforms	70
1.7.1	Quantifying Protein Isoforms	72
1.8	Conclusion	74
1.9	Aims	75
2	AQUA Peptide Selection	77
2.1	General AQUA Peptide Selection Criteria	77
2.2	AQUA Peptide Selection for the Quantification of a Protein Isoform ...	78
2.2.1	Protease Selection	79
2.2.2	Proteolytic Digest Optimisation	81
2.2.3	Reactive Amino Acid-Containing AQUA Peptides	83
2.3	The AQUA Peptide Selection Process	85
2.3.1	Initial Candidate Peptide Screening	86
2.3.2	BLASTP Based Alignment	86
2.3.3	Screening Each Remaining Candidate Peptide for Sequence Flaws	89
2.4	Residue Selection for Stable Isotopic Labelling	89

3	Methods	90
3.1	Cell Culture and Protein Production	90
3.1.1	Plasmid Amplification	90
3.1.2	Genejuice Based Transfection of HEK293	90
3.1.3	Anti-FLAG Based Immunoprecipitation	91
3.1.4	Gel Electrophoresis	91
3.1.5	Coomassie Blue Staining	91
3.1.6	Western Blotting	92
3.2	Recombinant DNA Techniques.....	92
3.2.1	Agarose Gel Electrophoresis Based DNA Purification	92
3.2.2	Agarose Gel Based DNA Extraction	92
3.2.3	T4 DNA Ligation	93
3.2.4	Miniprep Plasmid Screening	93
3.2.5	Plasmid Sequencing and Alignment	93
3.3	AQUA Peptide Characterisation	93
3.3.1	AQUA Peptide Re-Suspension.....	93
3.3.2	AQUA Peptide Alkylation.....	94
3.3.3	Zip Tip Based Sample Cleanup	94
3.3.4	Determining the Limit of Detection	94
3.4	Single Shot Based Peptide Analysis	94
3.4.1	Peptide Preparation for MALDI-ToF	94
3.4.2	Data Capture on a 4700 MALDI-ToF Based MS.....	95
3.4.3	Data Capture on an Ultraflex II MALDI ToF Based MS.....	95
3.4.4	Direct Injection on an API 2000 MS.....	95
3.4.5	Direct Injection on a QSTAR Pulsar MS.....	96
3.4.6	Direct Injection on a QTrap 5500 MS.....	96
3.4.7	Direct Injection on a TripleToF 5600 MS.....	97
3.5	LC-MS Based Peptide Analysis	97
3.5.1	In-Gel Digest Preparation	97
3.5.2	In-Gel Digestion	97
3.5.3	Spiked Digestion	97
3.5.4	HPLC on an Ultimate 3000	98
3.5.5	LC-MS on an API 2000 MS.....	99
4	The Absolute Quantification of Four A-Raf Isoforms	100
4.1	Introduction	100
4.1.1	Raf-1	103
4.1.2	B-Raf	104
4.1.3	A-Raf	106
4.1.3.1	A-Raf WT	108
4.1.3.2	A-Raf Short	109
4.1.3.3	DA-Raf-1 and DA-Raf-2	111
4.1.4	A-Raf Literature Overview	113
4.2	Project Aims	114
4.3	Methods Specific to A-Raf.....	115
4.3.1	Proteolytic Digest Optimisation.....	115
4.3.2	Met Oxidation and Reduction	116
4.3.2.1	DMSO Based Met Oxidation	116
4.3.2.2	DMS Based Met Sulfoxide Reduction.....	116
4.3.2.3	Sodium Periodate Based Met Oxidation.....	116

4.3.2.4	B-mercaptoethanol Based Met Sulfoxide Reduction	116
4.3.3	MALDI Based Analysis of AQUA Peptide VPTV*CVDMSTNRQQ	117
4.3.3.1	Characterisation of AQUA Peptide VPTV*CVDMSTNRQQ.....	117
4.3.3.2	AnchorChip Based Sample Concentration	117
4.3.3.3	LC-MALDI Based AQUA Quantitation.....	117
4.3.4	A-Raf TNT Based Protein Production.....	118
4.3.5	A-Raf Short Plasmid Manipulation	118
4.3.5.1	A-Raf Short Primer Design for SOE PCR.....	118
4.3.5.2	SOE PCR Overlap Production	119
4.3.5.3	SOE PCR Overlap Extension.....	120
4.3.5.4	Restriction Digest of the Final PCR Product.....	121
4.3.6	Non-Immunoprecipitated Exogenously Expressed A-Raf WT.....	121
4.3.7	Endogenously Expressed A-Raf WT	122
4.4	Results and Discussion	122
4.4.1	AQUA Peptide Selection.....	122
4.4.1.1	Sequence Unique to Each A-Raf Isoform.....	123
4.4.1.2	AQUA Peptide Selection for the Quantification of A-Raf WT..	124
4.4.1.3	A-Raf AQUA Peptide Selection Overview	128
4.4.2	AQUA Peptide Optimisation	130
4.4.2.1	The Oxidation and Reduction of Met.....	130
4.4.2.1.1	The Modification of Met Within Three Intact Proteins.....	130
4.4.2.1.2	Met Modification within AQUA Peptide VPTV*CVDMSTNRQQ132	
4.4.2.2	Proteolytic Digest Optimisation.....	135
4.4.3	MRM Acquisition Method Design.....	137
4.4.3.1	MRM Selection and Optimisation.....	137
4.4.3.2	Detection of AQUA Peptide VPTV*CVDMSTNRQQ.....	141
4.4.3.2.1	AnchorChip Based Sample Concentration.....	143
4.4.3.3	MALDI Based Characterisation of AQUA Peptide VPTV*CVDMSTNRQQ.....	145
4.4.3.4	MRM Linear Response and LoD on an API 2000.....	146
4.4.3.5	LoD on a QTrap 4000.....	152
4.4.3.6	LoD on a QTrap 5500.....	153
4.4.3.7	LoD for AQUA peptide VPTV*CVDMSTNRQQ	155
4.4.4	A-Raf Protein Production	156
4.4.4.1	TnT Based Protein Production.....	157
4.4.4.2	HEK293 Based Protein Production	159
4.4.4.3	A-Raf Short Incorrect Splicing.....	162
4.4.4.3.1	Evidence Supporting Incorrect Splicing	162
4.4.4.3.2	Modifying the Splice Site of Intron Two via Site Directed Mutagenesis	164
4.4.4.3.3	Confirming the Retention of Intron Two	168
4.4.5	Spiked Digest Based Quantitative Analyses	169
4.4.5.1	Single Isoform Based Spiked Digests.....	169
4.4.5.1.1	A-Raf WT Single Isoform Spiked Digest.....	170
4.4.5.1.2	A-Raf Short Single Isoform Spiked Digest	174
4.4.5.1.3	DA-Raf-1 Single Isoform Spiked Digest	175
4.4.5.1.4	DA-Raf-2 Single Isoform Spiked Digest	179
4.4.5.1.5	Single Isoform Spiked Digest Overview.....	182
4.4.5.2	Optimisation of the Single Isoform Based Spiked Digests	184

4.4.5.2.1	A-Raf WT Single Isoform Spiked Digest Optimisation	184
4.4.5.2.2	DA-Raf-1 Single Isoform Spiked Digest Optimisation	188
4.4.5.2.3	DA-Raf-2 QTrap 4000 and Ultraflex II Based Analysis	191
4.4.5.3	Multiple Isoform Spiked Digest Based Analysis.....	194
4.4.5.4	Non-Immunoprecipitated Exogenous HEK293 Spiked Digest ...	199
4.4.5.5	Endogenously Expressed A-Raf WT Spiked Digest	206
4.5	Conclusion	206
5	Absolute Quantification of the PDE4B Group of Protein Isoforms.....	209
5.1	Introduction	209
5.1.1	cAMP and PKA Activation	209
5.1.2	cAMP Specific Phosphodiesterases.....	214
5.1.3	The Physical Properties of the PDE Family	215
5.1.4	PDE4.....	217
5.1.4.1	PDE4B	220
5.1.4.1.1	PDE4B1	221
5.1.4.1.2	PDE4B2	221
5.1.4.1.3	PDE4B3	222
5.1.4.1.4	PDE4B4	223
5.1.4.1.5	PDE4B5	224
5.1.5	PDE4B Literature Overview	224
5.2	Project Aims	226
5.3	PDE Specific Methods	227
5.3.1	Proteolytic Digest Optimisation.....	227
5.3.1.1	Arg-C Digest Optimisation	227
5.3.1.2	Probing for Residual Glu-C Catalysed Asp Cleavage in AMBIC .	227
5.3.2	PDE4B Plasmid Preparation	228
5.3.2.1	Sequencing PDE4B in pEE7	228
5.3.2.2	Transferring the PDE4B Inserts from pEE7	228
5.3.2.3	FLAG-tag Insertion Into PDE4B Containing pcDNA3 or pcDNA3.1(+)	229
5.4	Results and Discussion	230
5.4.1	AQUA Peptide Selection.....	230
5.4.1.1	Sequence Unique to Each PDE4B Isoform.....	230
5.4.1.2	AQUA Peptides Selected for the Quantification of PDE4B	233
5.4.2	AQUA Peptide Optimisation	234
5.4.2.1	Proteolytic Digest Optimisation.....	235
5.4.2.2	Residual Glu-C Catalysed Cleavage at Asp	236
5.4.3	MRM Based Acquisition	240
5.4.3.1	MRM Method Development.....	240
5.4.3.2	MRM Linear Response and LoD on an API 2000.....	241
5.4.3.3	LoD on a QTrap 5500.....	243
5.4.4	PDE4B Protein Production.....	245
5.4.4.1	Sequencing PDE4B in a pEE7 Vector	246
5.4.4.2	Transfer of the PDE4B Inserts to pcDNA3 and pcDNA3.1(+) ...	246
5.4.4.3	Incorporation of FLAG-tags into the PDE4B Plasmids.....	247
5.4.5	Spiked Digest Based Quantitative Analyses	250
5.4.5.1	PDE4B1 Single Isoform Spiked Digest.....	251
5.4.5.2	PDE4B2 Single Isoform Spiked Digest.....	254
5.4.5.3	PDE4B3 Single Isoform Spiked Digest.....	259

5.5	Conclusion	264
6	Absolute Quantification of the SERCA 2 Group of Protein Isoforms	267
6.1	Introduction	267
6.1.1	P-Type ATPase	267
6.1.2	Sarcoplasmic/Endoplasmic Reticulum Calcium ATPase	269
6.1.2.1	SERCA2	271
6.1.2.1.1	SERCA2a	273
6.1.2.1.2	SERCA2b	274
6.1.2.1.3	SERCA2c	275
6.1.2.1.4	SERCA2d	275
6.1.2.1.5	SERCA2 Oxidation as a Disease Biomarker	276
6.2	Project Aims	277
6.3	SERCA2 Specific Methods	277
6.3.1	3-Nitrotyrosine Generation	277
6.4	Results and Discussion	278
6.4.1	AQUA Peptide Selection	278
6.4.2	SERCA2 Nitration	279
6.4.2.1	Optimisation of Tyr Nitration	280
6.5	Conclusion	283
7	Conclusion	285
7.1	Significance of this Research	289
7.2	Further Research	291
7.3	Final Conclusion	293
8	References	294

List of Tables

Table 4-1: The reaction products required for the first PCR stage of SOE PCR.	120
Table 4-2: The PCR program parameters for the first stage of SOE PCR.	120
Table 4-3: The reaction products required for the second PCR stage of SOE PCR.	121
Table 4-4: A theoretical digestion performed on the sequence unique to A-Raf WT.	125
Table 4-5: An NCBI BLASTP based alignment of each A-Raf WT candidate peptide against the <i>Homo sapiens</i> proteome.	126
Table 4-6: Further screening of the A-Raf WT candidate peptides against the AQUA peptide selection criteria.	127
Table 4-7: The AQUA peptides selected for the quantification of each of the A-Raf isoforms.	129
Table 4-8: The optimisation of trypsin, Lys-C and Glu-C based digestions of BSA.	135
Table 4-9: The MRM transitions selected for the detection of each A-Raf target/AQUA peptide.	141
Table 4-10: The LoD's achieved for each of the A-Raf MRM transitions on an API 2000.	151
Table 4-11: The LoD achieved for AQUA peptide GL*NQDCCVVYR on a QTrap 4000 MS.	152
Table 4-12: The LoD achieved for AQUA peptide GL*NQDCCVVYR on a QTrap 5500 MS.	154
Table 4-13: The quantity and variation data obtained from the spiked digestion of 25 µL of immunoprecipitated A-Raf WT on an API 2000.	172
Table 4-14: An ANOVA based statistical analysis performed on the data obtained from the spiked digest of A-Raf WT transfected HEK293 IP.	174
Table 4-15: The quantity and variation data obtained from the spiked digestion of 25 µL of immunoprecipitated A-Raf Short on an API 2000.	174
Table 4-16: The quantity and variation data obtained from the spiked digestion of 25 µL of immunoprecipitated DA-Raf-1 on an API 2000.	176
Table 4-17: An ANOVA based statistical analysis performed on the data obtained from the spiked digest of DA-Raf-1 transfected HEK293 IP.	178
Table 4-18: A T-test based statistical analysis performed on the data obtained from the spiked digest of DA-Raf-1 transfected HEK293 IP.	178
Table 4-19: The quantity and variation data obtained from the spiked digestion of 25 µL of immunoprecipitated DA-Raf-2 on an API 2000.	180
Table 4-20: A T-test based statistical analysis performed on the data obtained from the spiked digest of DA-Raf-2.	182
Table 4-21: The quantity and variation data obtained from the optimised spiked digestion of 25 µL of immunoprecipitated A-Raf WT on an API 2000.	185
Table 4-22: An ANOVA based statistical analysis performed on the data obtained from the optimised spiked digest of A-Raf WT transfected HEK293 IP.	187
Table 4-23: A T-test based statistical analysis performed on the data obtained from the optimised spiked digestion of A-Raf WT.	187

Table 4-24: The quantity and variation data obtained from the spiked digestion of 5 μ L of immunoprecipitated DA-Raf-1 on an API 2000.....	189
Table 4-25: An ANOVA based statistical analysis performed on the data obtained during the optimised spiked digest of DA-Raf-1 transfected HEK293 IP.	190
Table 4-26: A T-test based statistical analysis performed on the data obtained from the optimised spiked digestion of DA-Raf-1.....	191
Table 4-27: The quantity and variation data obtained from the spiked digestion of 12.5 μ L of immunoprecipitated DA-Raf-2 on a QTrap 4000. ...	192
Table 4-28: An ANOVA based statistical analysis performed on the data obtained during the optimised spiked digest of DA-Raf-2 transfected HEK293 IP.	194
Table 4-29: The quantity and variation data obtained from the spiked digestion of 1 pmol of each of the four immunoprecipitated A-Raf isoforms on a QTrap 5500.....	197
Table 4-30: The amount of each A-Raf isoform detected during a mixed isoform based spiked digestion.....	199
Table 4-31: The quantity and variation data obtained for each A-Raf WT target peptide, detected during the spiked digestion of an A-Raf WT transfected, un-enriched, HEK293 cell lysate.	203
Table 5-1: The AQUA peptides selected for the quantification of each of the PDE4B isoforms.	233
Table 5-2: The MRM transitions selected for the detection of each of the PDE4B peptides.	241
Table 5-3: The LoD's achieved for each of the PDE4B MRM transitions, as detected on an API 2000.	242
Table 5-4: The LoD's achieved for each NSP*CFFR transition, as detected on a QTrap 5500 MS.....	244
Table 5-5: The quantity and variation data obtained from the spiked digestion of 25 μ L of immunoprecipitated PDE4B1 on a QTrap 4000.....	252
Table 5-6: The quantity and variation data obtained during the spiked digestion of 25 μ L of immunoprecipitated PDE4B2 on a QTrap 4000.	255
Table 5-7: A T-test based statistical analysis performed on the data obtained from the spiked digest of PDE4B2.	257
Table 5-8: The amount of PDE4B2 detected, based upon the quantitation of for each TDIDIATE transition, during the spiked digestion of PDE4B2 transfected HEK293 IP on a QTrap 4000.	257
Table 5-9: The quantity and variation data obtained during the spiked digestion of 25 μ L of immunoprecipitated PDE4B3 on a QTrap 4000.	260
Table 6-1: The AQUA peptides selected for the quantification of each of the SERCA2 isoforms and nitrated SERCA2.....	278

List of Figures

Figure 1-1: A theoretical 2D gel based separation of five proteins.....	23
Figure 1-2: A theoretical LC based separation of five proteins.	26
Figure 1-3: The route through which analyte molecules are ionised during MALDI.	34
Figure 1-4: The route through which an analyte molecule is ionised during ESI.	36
Figure 1-5: A diagram depicting the workings of a quadrupole based MS.	42
Figure 1-6: The components contained within an FT-ICR MS.	44
Figure 1-7: The components contained within a hybrid quadrupole / Orbitrap MS.	48
Figure 1-8: The workflow through which two proteomes can be compared based on DIGE.	51
Figure 1-9: The stages involved in a SILAC based proteomic analysis.	54
Figure 1-10: The principle behind ¹⁸ O labelling and comparative protein quantitation.	56
Figure 1-11: The theory behind a duplex TMT based quantitative proteomic analysis.	60
Figure 1-12: The theory behind an iTRAQ based quantitative proteomic analysis.	61
Figure 1-13: The techniques through which PSAQ, AQUA and QconCAT enable an absolute quantitative proteomic analysis.	66
Figure 2-1: The three stages of selection employed when choosing an AQUA peptide.	85
Figure 2-2: The NCBI BLASTP search settings utilised when aligning candidate proteolytic peptides against the <i>Homo sapiens</i> proteome.	87
Figure 2-3: An example of an NCBI BLASTP based alignment; a technique performed so as to identify which peptides are unique within their target proteome.	88
Figure 2-4: An example of an NCBI BLASTP based alignment where the peptide is not unique to a single protein.	88
Figure 4-1: The structural organisation of Raf, showing the regions of conserved sequence.	101
Figure 4-2: The activation and subsequent signalling cascade of the MAPK pathway.	102
Figure 4-3: The cellular function and interaction of hnRNP H, A-Raf WT and A-Raf Short.	109
Figure 4-4: The cellular functions and signalling events associated with A-Raf WT and DA-Raf-2, as described by Nekhoroshkova <i>et al.</i>	112
Figure 4-5: The steps involved in the SOE PCR based site directed mutagenesis of A-Raf Short.	119
Figure 4-6: An alignment of each of the four A-Raf isoforms, highlighting the regions of sequence unique to each.	123
Figure 4-7: Each of the AQUA peptides selected for the quantification of the four A-Raf isoforms.	129
Figure 4-8: The Met containing tryptic peptides identified in Myoglobin. ...	130
Figure 4-9: The DMS and DMSO based reduction and oxidation of Met containing Myoglobin peptide HPGDFGADAQGAMTK.	131

Figure 4-10: The Sodium Periodate based oxidation and β -mercaptoethanol based reduction of AQUA peptide VPTV*CVDMSTNRQQ.	133
Figure 4-11: Both the MS and MS/MS spectra obtained for AQUA peptide VSQP*TAEQAQAFK on an API 2000.....	138
Figure 4-12: The predicted fragmentation spectrum for both synthetic and endogenous peptide VSQPTAEQAQAFK.	139
Figure 4-13: A MALDI-ToF based characterisation of AQUA peptide VPTVCVDMSTNRQQ, based upon the use of both MS and MS/MS.....	142
Figure 4-14: AQUA peptide VPTV*CVDMSTNRQQ spotted to an AnchorChip PAC II MALDI plate.	144
Figure 4-15: An MS spectrum of AQUA peptide VPTV*CVDMSTNRQQ, detected on an Ultraflex III MS.	145
Figure 4-16: The predicted fragmentation pattern of both endogenous and synthetic VPTVCVDMSTNRQQ, as predicted by Protein Prospector: MS-Product.	146
Figure 4-17: The signal response and LoD automatically/incorrectly assigned to AQUA peptide TV*VTVR.	148
Figure 4-18: How target and synthetic peaks should be manually integrated for each MRM transition.....	149
Figure 4-19: The signal response and LoD manually/correctly assigned to A-Raf AQUA peptide TV*VTVR.	150
Figure 4-20: The signal response obtained for AQUA peptide VPTV*CVDMSTNRQQ on an Ultraflex II MS, when eight different quantities were injected.	155
Figure 4-21: A 1D-SDS-PAGE based separation of TNT transfected A-Raf WT and A-Raf Short.	157
Figure 4-22: A western blot of TNT transfected FLAG-tagged A-Raf WT and A-Raf Short.	158
Figure 4-23: A 1D-SDS-PAGE based separation of HEK293 transfected A-Raf WT and A-Raf Short.	160
Figure 4-24: A western blot of HEK293 transfected A-Raf WT and A-Raf Short.	161
Figure 4-25: An MRM based detection of incorrectly spliced A-Raf Short, obtained from a HEK293 based transfection.	163
Figure 4-26: The 40 mer oligonucleotide selected for the site directed mutagenesis of A-Raf Short.	167
Figure 4-27: The genetic sequence obtained from the A-Raf Short plasmid, both before and after site directed mutagenesis.....	168
Figure 4-28: A western blot showing the MW difference between HEK293 transfected A-Raf short, before and after the completion of site directed mutagenesis.	168
Figure 4-29: An XIC obtained from the spiked digestion of A-Raf WT on an API 2000.	171
Figure 4-30: The quantities of each A-Raf WT target peptide detected during the spiked digestion of A-Raf WT transfected HEK293 IP on an API 2000. ..	173
Figure 4-31: The quantity of peptide TVVTVR, detected during the spiked digestion of A-Raf Short transfected HEK293 IP on an API 2000.	175
Figure 4-32: The quantities of each DA-Raf-1 target peptide detected during the spiked digestion of DA-Raf-1 transfected HEK293 IP on an API 2000. ..	177

Figure 4-33: The quantities of each DA-Raf-2 target peptide detected during the spiked digestion of DA-Raf-2 transfected HEK293 IP on an API 2000. ..	181
Figure 4-34: An investigation performed so as to determine the source of the variation detected during the spiked digestion of AQUA peptide GL*NQDCCVVYR.	183
Figure 4-35: The quantities of each A-Raf WT target peptide detected during the optimised spiked digestion of A-Raf WT transfected HEK293 IP on an API 2000.	186
Figure 4-36: The quantities of each DA-Raf-1 target peptide detected during the optimised spiked digestion of DA-Raf-1 transfected HEK293 IP on an API 2000.	189
Figure 4-37: The quantities of each DA-Raf-2 target peptide detected during the spiked digestion of DA-Raf-2 transfected HEK293 IP on a QTrap 4000, and an Ultraflex II.	193
Figure 4-38: The complex XIC obtained during the trypsin based spiked digestion of each of the four A-Raf isoforms, analysed on a QTrap 5500. ..	196
Figure 4-39: The quantities of each A-Raf target peptide detected during the spiked digestion of 1 pmol of each of the four A-Raf isoforms on a QTrap 5500.	198
Figure 4-40: The XIC obtained when 25 μ L of transfected, un-enriched, A-Raf WT lysate was subjected to a spiked digest based analysis, detecting on a QTrap 5500, utilising either low or unit Q1 resolution MRM transition windows.	202
Figure 4-41: The MRM spectra obtained from both the synthetic and endogenous peaks detected during the analysis of 25 μ L of A-Raf WT transfected, non-enriched HEK293 cell lysate.	205
Figure 5-1: The signalling cascade required for the production of intercellular cAMP.	210
Figure 5-2: The signalling cascade required for the activation of PKA and subsequent protein phosphorylation.	212
Figure 5-3: cAMP gradient compartmentalisation, based upon the expression of PDE.	213
Figure 5-4: The sequence orientation of typical long, short and super-short PDE isoforms.	218
Figure 5-5: A Clustal X based alignment of PDE4B1, PDE4B2, PDE4B3 and PDE4B5, highlighting the areas of sequence unique to each of the isoforms.	232
Figure 5-6: The AQUA peptides selected for the quantification of each of the five PDE4B isoforms.	234
Figure 5-7: Proteolytic peptides obtained from BSA cleaved with Glu-C at Glu, but which contain Asp.	237
Figure 5-8: The peak intensities obtained for three PepMix II calibration standards.	238
Figure 5-9: The ion intensity detected for Glu-C cleaved peptide AKDAFLGSFLYE.	239
Figure 5-10: The signal intensity and LoD achieved for PDE4B AQUA peptide TDI*DIATE.	243
Figure 5-11: The design of the N-terminal FLAG-tag to be inserted into PDE4B1 in pcDNA3.	248

Figure 5-12: The design of the N-terminal FLAG-tag to be inserted into PDE4B3 in pcDNA3.	249
Figure 5-13: The design of the N-terminal FLAG-tag to be inserted into PDE4B2 in pcDNA3.1(+).	250
Figure 5-14: The quantities of each PDE4B1 target peptide detected during the spiked digestion of PDE4B1 transfected HEK293 IP on a QTrap 4000. ...	253
Figure 5-15: The increasing ion intensity detected over the course of an MRM based LC-MS/MS analysis of Glu-C cleaved PDE4B1.	254
Figure 5-16: The quantities of each PDE4B2 target peptide detected during the spiked digestion of PDE4B2 transfected HEK293 IP on a QTrap 4000. ...	256
Figure 5-17: The XIC spectra obtained for TDIDIATE transitions B ₅ and B ₆ , during the spiked digestion of 25 µL of PDE4B2 transfected HEK293 IP.	258
Figure 5-18: The quantity of each PDE4B3 target peptide detected during the spiked digestion of PDE4B3 transfected HEK293 IP on a QTrap 4000.	261
Figure 5-19: A typical peak detected for peptide VNPQEESYQK, fragment KQYS, during the spiked digestion of PDE4B3 on a QTrap 4000.	262
Figure 5-20: The MS signal intensity detected for both the TIQTYRSVSE IS and target peptide, over the duration of the PDE4B3 batch.	264
Figure 6-1: The cycle of changes in conformation undergone by P-type ATPases, enabling the transportation of ions.	268
Figure 6-2: The exon expression patterns of each of the SERCA2 isoforms.	272
Figure 6-3: The nitration of alkylated AQUA peptide GLNQDCCVVYR, through the acidification of NaNO ₂	282
Figure 6-4: The possible nitration products of AQUA peptide GLNQDCCVVYR, when exposed to acidified NaNO ₂	283

Acknowledgements

I would like to express my gratitude to the various supervisors I have had over the course of this project, including Professor Andy Pitt, based at Aston University, whose advice and numerous mass spectrometers have been instrumental to my research, Doctor Nick Morrice, based at the Beatson Institute of Cancer Research, for providing me with access to a QTrap 4000 during the final year of my research, Professor Walter Kolch, based at University College Dublin (previously of the Beatson Institute of Cancer Research), for access to his research staff and consumables during my A-Raf chapter, and particularly to Doctor George Baillie, based at the University of Glasgow, for his assistance with my PDE4B research, but also for his support during the writing of my thesis.

I am particularly indebted to Sarah Cumming, based at the University of Glasgow, for her support, advice and training over the course of my research, and indeed during the writing of my thesis, without whom none of this would be possible. Likewise I would like to thank Jens Rauch based at University College Dublin (previously of the Beatson Institute of Cancer Research), for his training and advice during my analysis of A-Raf, but also during the writing of my thesis.

I would also like to thank the various academic staff who have helped me during my research, including, in no particular order, David Sumpton, Ekaterina McKenna, Richard Goodwin, Kathryn Gilroy, Mairi Sandison, Susan Gannon, Karl Burgess, Bo Wang, Jillian Bryce, Angela Woolton and Richard Burchmore.

Furthermore, I would like to thank some past and present PhD students, including, in no particular order, Heather Allingham, Michael Lang, Kit-Yee Tan, Alexia Koletsou, Anne-Marie Reid, Roby Urcia, Judith Nicholson, Lewis Ross, Anna Robson and James Schofield for their assistance and support.

I would particularly like to thank Paul Getty, Dominic Ketley and Kshama Pansare for getting me out of the lab on occasions and keeping me sane during my research.

I would personally like to thank Andrew McFarlane for our various discussions, and to Donna, Aiden, and my entire family, for their support during my write-up.

Author's Declaration

I hereby declare that the thesis which follows is of my own composition, that it is a record of the work done by myself, and that it has not been presented in any previous application for a higher degree.

A handwritten signature in black ink that reads "R. Kelly". The signature is written in a cursive style with a long, sweeping flourish extending upwards and to the right from the end of the name.

Robert Kelly

Definitions / Abbreviations

%	Percent
°C	Degrees Celsius
2D-LC	2D-liquid chromatography
2DGE	2D-SDS-PAGE
amu	Atomic mass unit
A-Raf WT	A-Raf wild type
AAA	Amino acid analysis
AKAP	A-kinase anchoring protein
ALS	Acid labile surfactants
AMBIC	Ammonium bicarbonate
ANOVA	Analysis of variance
AQUA	Absolute quantification
ARF6	ADP ribosylation Factor six GTPase
ASK-1	Apoptosis signal-regulating kinase 1
ATP	Adenosine-5'-triphosphate
BLAST	Basic local alignment search tool
cAMP	3'-5'-cyclic adenosine monophosphate
C-terminal	Carboxy-terminal
Ca ²⁺	Calcium
cDNA	Complementary DNA
CoV	Coefficient of variation
CID	Collision induced dissociation
CO ₂	Carbon dioxide
Cps	Counts per second
CR1	Conserved region one (Raf)
CR2	Conserved region two (Raf)
CR3	Conserved region three (Raf)
CRD	Cys-rich domain (Raf)
dNTP	Deoxyribonucleotide triphosphate
DA-Raf-1	Deleted A-Raf 1
DA-Raf-2	Deleted A-Raf 2
DC	Direct-current
DEX	Glucocorticoid dexamethasone
DIGE	Difference gel electrophoresis
DISC	Disrupted-in-Schizophrenia-1
DMS	Dimethyl sulfide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
eV	Electron volt
ECD	Electron capture dissociation
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EPI	Enhanced product ion
ER	Endoplasmic reticulum
ERK	Extracellular signal-regulated kinases
ESI	Electrospray ionisation

Fmoc	Fluorenylmethyloxycarbonyl
FA	Formic acid
FT-ICR	Fourier transform-ion cyclotron resonance
GDP	Guanosine-5'-triphosphate
GTP	Guanosine-5'-triphosphate
hnRNP H	Heterogeneous nuclear ribonucleoprotein H
H ₀	Null hypothesis
H ₂ O	Water
HCCA	α-cyano-4-hydroxycinaminic acid
HCl	Hydrochloric acid
HPLC	High pressure liquid chromatography
iTRAQ	Isobaric tags for relative and absolute quantitation
IAA	Iodoacetamide
ICAT	Isotope-coded affinity tags
IP	Immunoprecipitate
IPG	Immobilized pH gradient
IRMPD	Infrared multiphoton dissociation
IS	Internal standard
LoD	Limit of detection
LC	Liquid chromatography
LIT	Linear ion trap
LLOQ	Lower limit of quantification
<i>m/z</i>	Mass to charge ratio
min	Minute
mRNA	Messenger ribonucleic acid
M	Molar
MeCN	Acetonitrile
MeOH	Methanol
MALDI	Matrix-assisted laser desorption / ionisation
MAPK	Mitogen activated protein kinase
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem MS
MST2	Mammalian sterile 20-like kinase
MW	Molecular weight
N-terminal	Amino-terminal
NaNO ₂	Sodium nitrite
NO	Nitric oxide
PCR	Polymerase chain reaction
PDE	Phosphodiesterase
PDGF	Platelet-derived growth factor
PKA	Protein kinase A
PMF	Peptide mass fingerprinting
PSAQ	Protein standard absolute quantification
PTM	Post-translational modifications
QconCAT	Quantitative concatamer
Raf	Rapidly accelerated fibrosarcoma
Ras	Rat sarcoma
RBD	Ras-binding domain

RF	Radio-frequency
RNA	Ribonucleic acid
RP	Reversed phase
siRNA	Small interfering RNA
Src	Sarcoma (MAPK signal transduction protein)
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-poly-acrylamide gel electrophoresis
SERCA	Sarcoplasmic /endoplasmic reticulum calcium ATPase
SILAC	Stable isotope labelling by amino acids in cell culture
SIN-1	3-Morpholinosydnonimine hydrochloride
SISCAPA	Stable isotope standards with capture by anti-peptide antibodies
SNP	Single nucleotide polymorphism
SnR	Signal to noise ratio
SOE-PCR	Splice overlap extension PCR
SORI-CID	Sustained off-resonance irradiation collision-induced dissociation
SPPS	Solid-phase peptide synthesis
ToF	Time-of-flight
ToF/ToF	Tandem ToF MS
TBST	Tris-buffered saline with 0.05% Tween (v/v)
TFA	Trifluoroacetic acid
TMT	Tandem mass tags
TNM	Tetranitromethane
UCR1	Upstream conserved region one (PDE4)
UCR2	Upstream conserved region two (PDE4)
v	Volume
V	Volt
w	Weight
XIC	Extracted ion chromatogram

1 Introduction

As was briefly described in the abstract, this project details the development of a stable isotopically labelled internal standard based quantitative technique for the absolute quantification of three families of protein isoforms, through the use of mass spectrometry (MS). This chapter thus details the literature surrounding the development of proteomics as a field of scientific research, the various methods through which a protein may be quantified, the various types of mass spectrometers which are utilised in proteomics, and the methods which have been developed for the absolute quantification of proteins.

1.1 The Birth of Proteomics

While genomic based studies, such as the Human Genome Project (Venter et al., 2001), have yielded vast amounts of sequence data, the “blueprints” of life (Mundy, 2001), these data have provided only a basis for evaluating what is possible within the cell, yet has yielded no information as to the biological activity, or phenotype, of a cell in a specific environment at a given point in time (Montpetit, 2003, Wasinger et al., 1995). Indeed, it is now understood that the levels of messenger ribonucleic acid (RNA) (mRNA) expressed within a cell, can give only a rough indication as to the rate at which a protein is produced (Duggan et al., 1999, Anderson and Seilhamer, 1997), which yields no information as to the final concentration of the protein, due to the varying rates at which a protein may be degraded. Furthermore, mRNA stability within the cell can be variable and the rate at which mRNA is translated into protein may also vary. mRNA expression yields no information as to the presence of post-translational modifications (PTMs), which can be suggestive of cellular function (Unwin and Whetton, 2006) and current bioinformatic techniques cannot differentiate between genes and pseudogenes (Montpetit, 2003), cannot detect alternatively spliced proteins, nor do they take into account for the fact that many amino acid chains can be incorporated into a single protein, in the form of subunits (Godovac-Zimmermann et al., 2005).

It is thus understandable that the field of proteomics has received increased interest over the last decade (Dunn, 2007), initially being envisioned as a technique capable of characterising each of the proteins within an organism simultaneously (Zhang et al., 2010). This assumption presented proteomics with its first major challenge, addressing the dynamic nature of the proteome through independently analysing each organism under a range of conditions (Zhang et al., 2010).

Further to the environment from which the organism was sampled, which will determine the protein expression pattern, it has been predicted that up to 74 percent (%) of the 20,000 to 30,000 genes in the human genome may be subject to alternative splicing (Johnson et al., 2003), yielding an average of two or three transcripts per gene (M. Nakao et al., 2005), thereby greatly increasing the complexity of the proteome. Furthermore, it is also common for proteins to be post-translationally modified with one or more of some 200 PTMs, resulting in the number of modified and unmodified proteins within an organism generally being much larger than the number of genes identified (Zhang et al., 2010, Anderson et al., 2004b).

The highly variable concentrations at which proteins are expressed are also a problem for proteomic based analyses. For example, the dynamic range of protein expression in yeast can vary by up to four orders of magnitude (Anderson and Anderson, 2002), and by up to 12 orders of magnitude in extra-cellular fluids, such as blood (Corthals et al., 2000). By way of an example of the scale of the problem, just 12 housekeeping proteins (required for cell survival, integrity or duplication, and thus are thought to be expressed in every cell at approximately the same level (Thorrez et al., 2008)) in human plasma have been found to account for up to 95% of the total protein mass (Zhang et al., 2010), and a single extracellular protein, albumin, being recorded so as to account for up to 50% of the total protein mass (Corthals et al., 2000). Were such a sample to be proteolytically digested and analysed via MS, the proteolytic peptides generated from those 12 most abundant proteins would compete during ionisation with those peptides generated from lower abundance proteins,

possibly masking their detection. Furthermore, during an information-dependent acquisition (IDA) based MS scan (a scan which monitors the MS spectrum until an ion which satisfies user specified MS/MS selection criteria is detected), an eluting peptide may be missed while another peptide undergoes tandem MS (MS/MS). This limitation can be reduced, however, through utilising an instrument with a shorter cycle time. It is thus clear that only after several stages of separation would it be possible to detect and analyse lower abundance proteins, or to assess the differences between cellular phenotypes (Klose, 1975, O'Farrell, 1975).

1.2 Protein Separation

It is indeed because of the vast range of concentrations at which proteins are expressed, the overwhelming number of modified, unmodified and alternatively spliced proteins within any given sample that the field of proteomics has required high sensitivity, high resolution, and high throughput techniques and technologies capable of delivering high-confidence protein identifications.

Classically protein separations have been performed through two-dimensional sodium dodecyl sulphate (SDS) poly-acrylamide gel electrophoresis (SDS-PAGE) (2DGE) (Corthals et al., 2000, O'Farrell, 1975), coupled with spot picking, proteolytic digestion, and a matrix-assisted laser desorption/ionisation (MALDI) coupled with time-of-flight (ToF) based protein identification. In recent years, however, liquid chromatography (LC) based separations have offered scientists an automated, MS compatible, alternative (Link et al., 1999).

1.2.1 2DGE

While poly-acrylamide based gel electrophoresis was first employed to separate protein back in 1959 (Kwietny et al., 1959), it was not until 1974 that isoelectric focusing (IEF) was combined with SDS-PAGE to yield 2DGE, a development which heralded a new age for analytical protein separation (Corthals et al., 2000, O'Farrell, 1975).

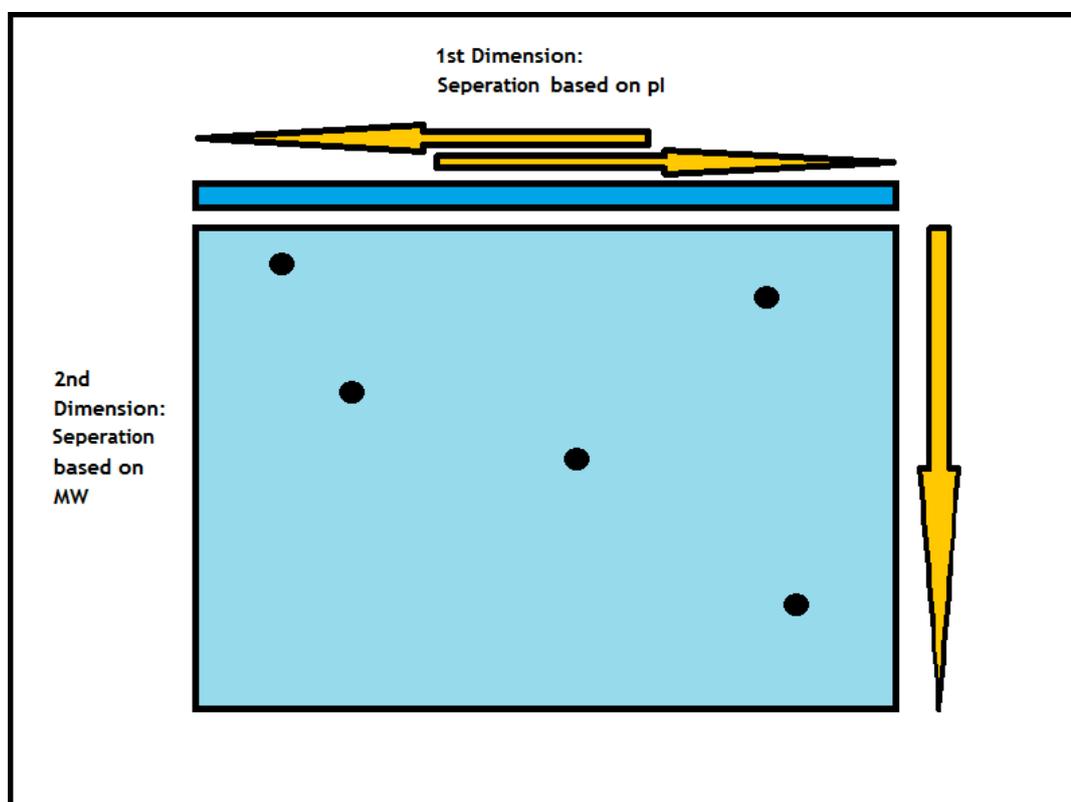


Figure 1-1: A theoretical 2D gel based separation of five proteins. The image above depicts the steps required to separate a proteome based on 2DGE, the first step of which involves each protein migrating across an IPG strip, based on its isoelectric point (pI), until a neutral charge is achieved, at which point the proteins are further separated based on the number of bound SDS molecules, and thus their MW.

The first dimension of 2DGE involves the proteome being separated based on the pI of each protein, where each protein will migrate across an immobilized pH gradient (IPG) until a neutral charge is achieved (Corthals et al., 2000, O'Farrell, 1975). The second dimension separation step is then conducted, applying SDS to the sample, so as to impart a uniform negative charge to each protein. As 1.4 grams of SDS binds to one gram of protein, each protein can be separated via electrophoresis at a rate approximate to its molecular weight (MW) (O'Farrell, 1975), as is shown on Figure 1-1. At this stage, the separated proteins can be visualised through the application of a number of protein stains or dyes. Options include Coomassie brilliant blue, Sypro and silver stain. Which dye is selected will depend on the sensitivity required (0.05 - 500 nanograms (ng) per spot), the

requirement for reproducibility, and the optional requirement for compatibility with downstream protein characterization techniques (Klose, 1975, O'Farrell, 1975), such as MS.

While 2DGE has previously been seen as the “gold standard” in protein separation, this approach has many well-documented drawbacks. In terms of protein chemistry, proteins have an average pI range of between pH 3 and pH 12. IPG strips can however only offer a pI range of between pH 3 and pH 10, with the majority of commercial IPG strips covering much smaller pI ranges. Should a protein fall outside the pI range it will migrate to the end of the IPG strip and precipitate. Likewise in the second dimension, SDS-PAGE gels are limited by their MW ranges, causing those proteins which are either too large or small to either fail to enter the gel or to proceed through and pass beyond the gel, respectively (Herbert et al., 2001). In terms of single analysis reproducibility, inter-gel variability has been well documented, preventing the direct comparison of varying proteomes on separate gels (Rabilloud, 1990). Furthermore, dissimilarities in protein migration, inefficiency during the loading of the IPG strip, limited sensitivity, limited dynamic range (Shen and Smith, 2002) and problems with the transfer of protein between the first and second dimensions have limited the use of 2DGE, especially in comparative proteomic studies outside of difference gel electrophoresis (DIGE) (Rabilloud, 1990). By way of an example of the scale of the problem, it has recently been estimated that more than 50% of proteins in yeast are not suited to detection based upon the use of 2DGE (Gygi et al., 2000).

Despite the shortcomings associated with 2DGE, this technique continues to provide analysts with a tool capable of separating protein isoforms. For example, in 1990 Kojima and co-workers utilised 2DGE and immunostaining to identify three C-protein isoforms, and to map their expression in both neonatal breast muscle and posterior latissimus dorsi muscle (Kojima et al., 1990). Furthermore, in 2006 Raikos and co-workers identified several ovalbumin and conalbumin isoforms in egg white, based upon the use of 2DGE and MALDI-ToF (Raikos et al.,

2006). Finally, in 1994 Nakamura and co-workers detected seven α - and four β -tubulin isoforms within the axoneme of *Tetrahymena* cilia, based upon the use of 2DGE and immunostaining (Nakamura et al., 1994).

In addition to the separation of protein isoforms, 2DGE offers analysts a tool capable of separating hydrophobic membrane proteins, through solubilising the proteins in a range of non-MS compatible detergents, prior to separation, in-gel digestion and an MS based detection (Babu et al., 2004).

1.2.2 Liquid Chromatography

In seeking an alternative technique to separate the proteome, researchers have recently turned to LC (Shi et al., 2004). In addition to its compatibility with electrospray ionisation (ESI) based MS (Shen and Smith, 2002, Fenn et al., 1989), LC systems have been fitted with a wide array of stationary and mobile phases, so as to enable both high resolving power, and specificity, assuming a unique sub-set of the proteome is to be targeted (Shi et al., 2004). While currently most proteomic LC based separations are performed via reversed phase (RP), ion-exchange, affinity and size-exclusion based chromatography may all be utilised (Zhang et al., 2010).

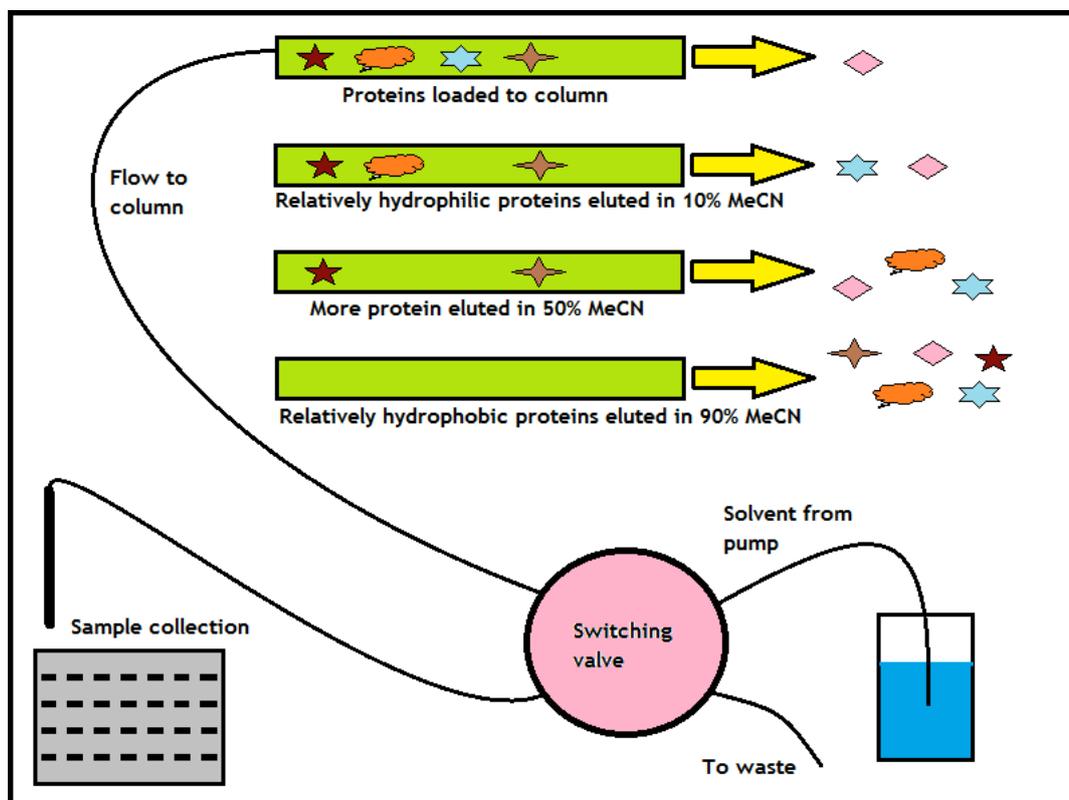


Figure 1-2: A theoretical LC based separation of five proteins. The image above depicts the steps required to separate a proteome based on RP-LC, the first step of which involves the sample being injected to the LC column in a highly polar solvent, causing the hydrophobic proteins to adsorb to the silica bound C18 chains. As the ratio of polar to organic solvent is gradually increased, the proteins or peptides are eluted based upon their hydrophobicity.

In regards to the fundamentals of RP-LC, the stationary phase is generally composed of a silica bead packed column, derivatised with a layer of aliphatic carbon (most commonly C18, though C8 and C4 may be utilised for less hydrophobic samples). The sample is loaded onto the stationary phase in a polar solvent (water (H₂O)), acidified with 0.1% (volume/volume (v/v)) formic acid (FA) (employed as both a source of protons and to prevent the carboxyl groups from dissociating, thereby preventing secondary interactions with the stationary phase), under which conditions the peptide or protein will adsorb to the stationary phase. The ratio of acidified organic solvent (most commonly Acetonitrile (MeCN) or Methanol (MeOH) acidified with 0.1% (v/v) formic acid) to polar solvent is then increased, and each peptide or protein eluted when the strength of organic solvent in the mobile phase is sufficient to compete with the

hydrophobic forces retaining the analyte (Shi et al., 2004), as is shown on Figure 1-2. The term RP is applied to this technique as historically the sample was loaded to the silica stationary phase in an organic solvent based buffer, and the sample components eluted based upon a gradient of increasingly polar solvent.

Where greater resolving power is required, two-dimensional liquid chromatography (2D-LC) may be utilised (Giddings, 1984). 2D-LC employs two orthogonal LC based separation steps, performed either on-line (with the eluant from the first column flowing directly into the second column, in series) or off-line (utilising fraction collecting and subsequent re-injection), so as to offer the analyst unparalleled resolving power. To better explain, 2D-LC enables a complex proteome to be separated based upon the exploitation of two separate biochemical parameters, such as hydrophobicity, size, or the number of basic or acidic residues a protein may contain.

1.3 MS Based Protein Identification

While 2DGE based experiments have predominantly involved the spot or protein of interest being excised from the poly-acrylamide gel, proteolytically digested, and ionised/detected through the application of MALDI-ToF, LC based separation coupled with an ESI based MS detection can be conducted in one of two ways; top-down or bottom-up.

1.3.1 Top-down Proteomics

Top-down based proteomic analyses involve the separation of a complex mixture of proteins via LC and the detection of each intact protein molecule via MS (Wu et al., 2012, Loo et al., 1990, Kelleher et al., 1999). This technique offers several advantages over bottom-up based proteomic techniques, the first of which is the ability of top-down proteomics to detect which PTMs are present on a protein, or indeed if several PTMs are mutually exclusive (Jeffrey, 2012). Furthermore, top-down based proteomics offers the analyst reduced sample

preparation time and reduced sample complexity, thereby offering better separation.

In terms of applications, top down based proteomics has been utilised in the verification of translational start and stop sites identified during genome based bioinformatic analyses (Ryan et al., 2010). In this application, however, accurate mass based detection alone can be complicated by the presence of PTMs and splice isoforms, significantly altering the mass of the protein from that predicted (Cui et al., 2011), indeed it has been noted that identifying a protein based upon intact mass alone is sufficient for only a limited number of proteins in an LC fraction (Wu et al., 2009), MS/MS is therefore a requirement of top-down based proteomics.

Furthermore, top-down based proteomics offers researchers a tool for the detection and identification of protein isoforms. Through analysing the intact structure and fragmentation pattern of each protein, it is possible to obtain sequence data which may otherwise be lost (assuming the use of top-down based proteomics). One such top-down based proteomic analysis, recently performed by Tran and co-workers, identified 1,043 unique gene products from a cultured human cell lysate, which together yielded more than 3,000 unique protein species, including protein isoforms, unique proteins generated through the addition of PTMs and protein fragments generated via proteolysis, based upon the use of a novel four-dimensional separation system (Tran et al., 2011).

Top-down based proteomics does suffer from several technological limitations though, for example, only proteins containing 500 or fewer amino acids (50 kilodaltons (kDa) or less) can be detected, due to the limitations of current mass spectrometers (Cui et al., 2011). Furthermore, high resolution mass spectrometers are required to differentiate between multiple-charge states, instruments which are expensive to both purchase and operate. In addition, top-down based proteomics has been shown to struggle with the detection of very hydrophobic or hydrophilic proteins, in addition to low copy number proteins within complex samples (Motoyama and Yates, 2008). Finally, the technique

through which MS/MS fragmentation is most commonly employed during top-down based proteomics, electron capture dissociation (ECD) (Sleno and Volmer, 2004), requires long ion accumulation, activation and detection times, limiting the applications of this technique to samples which are relatively pure.

1.3.2 Bottom-up Proteomics

In performing a bottom-up based proteomic analysis the complex mix of proteins is proteolytically digested prior to LC based separation and MS based detection (Wu et al., 2012, Fournier et al., 2007). In pre-digesting the target proteins, many of the problems associated with top-down based proteomic analyses are avoided, enabling the detection of a much larger range of protein species and PTMs. Bottom-up based proteomics has therefore become the technique of choice for large-scale protein studies (Zhang et al., 2010).

In proteolytically digesting an already complex proteome however, hundreds of thousands of peptides are created, which to date, no single LC based separative technique has been capable of resolving (Zhang et al., 2010). This can result in multiple peptide ions eluting from the LC column simultaneously, overwhelming the detector and reducing the number of peptides potentially identified during an IDA based MS/MS scan. In addition to simply limiting the results obtained for a given sample, the limited resolution offered by single dimension LC in bottom-up based proteomics has also limited the reproducibility of this technique (Zhang et al., 2010). 2D-LC is therefore often employed in bottom-up based proteomics so as to increase the resolving power of the LC separation and therefore, through increasing the number of peptides detected, increasing the protein sequence coverages achieved.

A fundamental flaw exists within bottom-up based proteomics, however, in that identical peptides can be generated through the digestion of several protein isoforms, or indeed multiple distinct proteins, leading to confusion over which protein is actually expressed, and indeed the accuracy of the protein sequence coverages or quantifications achieved (Cui et al., 2011). Because of this, top-

down based proteomics has again begun to attract more attention in recent years (Collier et al., 2008, Ferguson et al., 2009, Liu et al., 2002, Tran et al., 2011).

1.4 Protein identification

Prior to 1990 it was thought possible to sequence only the first 20 - 30 amino acids of a single, highly purified peptide or protein over a matter of hours or days. This was achieved based on the sequential cleavage and ultraviolet (UV) based identification of the peptide or proteins amino-terminal (N-terminal) amino acid through a technique known as Edman degradation (Steen and Mann, 2004, Edman, 1949). So as to avoid the sequence length limitations of this technique, proteins were often proteolytically digested and purified via RP-LC prior to the sequencing of each peptide fraction. With the advent of soft ionisation techniques however, including ESI and MALDI, MS has become the method of choice for protein identification.

Early MS based protein identifications were conducted following the completion of 2DGE; where each protein spot was picked, proteolytically digested, and the species identified based on MALDI-TOF MS coupled with peptide mass fingerprinting (PMF) (Pappin et al., 1993). In brief, PMF assumes that when a protein is digested it will yield a unique set of proteolytic peptides, thus when the masses of several such peptides are detected during an analysis, an identity for the target protein can be predicted.

PMF does, however, suffer from several limitations, the first of which is the requirement for the digested protein to be relatively pure. Considering the complexity of the proteome this can be difficult to achieve, even when 2DGE has been utilised. Furthermore, for a proteins identity to be predicted, the target protein must already exist within the database being searched. As each database is species specific, however, and the data generated through the bioinformatic analysis of a sequenced genome, for a protein identification to be made, the target organism's genome must thus first have been sequenced (Steen and Mann,

2004). Finally, many proteins will yield peptides with similar mass to charge ratios (m/z), which can lead to erroneous data, in addition to which, where a peptide contains a PTM the final mass of the peptide will be significantly altered (Mann and Wilm, 1994), possibly resulting in an incorrectly predicted identity being assigned to the protein.

Alternative approaches have therefore been developed to enable the accurate identification of a protein based upon IDA scanning and MS/MS. The first of these techniques, termed *de novo* sequencing, involves the collection and interpretation of MS/MS spectra (Steen and Mann, 2004), assessing the distance between MS/MS fragmentation peaks, so as to predict the loss of a specific amino acid. Through assessing the m/z ratio of the MS/MS fragments, it is possible to piece together this sequence information, so as to determine the structure of the target peptide. While it is possible to determine the sequence of a peptide via *de novo* sequencing, this technique is often complicated by incomplete fragmentation and the presence of interfering peaks, originating both from the target peptide and external sources (Steen and Mann, 2004).

At the beginning of the 1990s, however, it was realised that only an infinitesimally small fraction of all possible amino acid sequences actually occur in nature, and therefore, if even a limited MS/MS spectrum was obtained, in conjunction with the peptides' m/z and the specificity of the protease utilised, that the sequence of the peptide/protein could accurately be predicted (Mann and Wilm, 1994, Rappsilber and Mann, 2002).

As such, the most common mass spectrometry based technique for identifying a protein now involves: (i) proteolytically digesting the protein of interest and injecting the resulting peptide mix onto an LC-MS/MS system, (ii) ionising the target peptides, so as to obtain a m/z , or parent ion, for each peptide, (iii) having each parent ion selected for MS/MS, and (iv) deriving an MS/MS spectrum for each peptide. Finally, (v) each MS/MS spectra is compared to a genomically derived sequence database, so as to predict the identity of the target protein.

1.5 Mass Spectrometry

The origins of mass spectrometry can be traced back to the late 1800s when J. J. Thomson (Thomas, 1910), in researching the transmission of electricity through gas, built a device capable of measuring the mass of charged particles (Griffiths, 2008). This device utilised gas discharge tubes to generate ions, which were passed through parallel electric and magnetic fields, deflecting the ions into parabolic trajectories which were then detected on a photographic plate (Griffiths, 2008).

1.5.1 Biomolecule Ionisation

For a long time, however, many of the ion sources used in MS, for the chemical analysis of small molecules, employed energy rich chemical and electrical ionisation techniques which were not suited to the ionisation of large intact biomolecules, as they were found to decompose in an unpredictable manner (Sauer and Kliem, 2010). The application of MS was thus limited to the analysis of volatile compounds, or those which could be easily derivatised to become volatile (a property necessary for their conversion into the gas phase) (Kicman et al., 2006). While many “sudden energy” techniques were proposed for the ionisation of small biomolecules (Fenn, 2002), including fast ion bombardment (Benninghoven et al., 1976), fast atom bombardment (Devienne and Roustan, 1982), plasma desorption (Torgerson et al., 1974), pyrolysis mass spectrometry (Meuzelaar et al., 1973), and laser desorption (Posthumus et al., 1978), it was not until the realisation that if non-volatile solutes were dissolved in volatile solvents, and highly charged droplets of these solutions produced, that the evaporation of the solvent would leave intact gaseous ions, that the ionisation of labile molecules was thought possible (Griffiths, 2008, Hillenkamp and Karas, 2000, Fenn et al., 1989). Indeed, it was not until 1988 with the description of two soft ionization techniques, ESI and MALDI, that MS was seen to enter the biological arena (Falk et al., 2006).

1.5.1.1 Matrix Assisted Laser Desorption/Ionisation

The development of MALDI can be traced back to an experiment conducted by Hillenkamp and Karas into the spatial distribution of Calcium ions (Ca^{2+}) in heart muscle cells based on laser microprobe mass analysis (Griffiths, 2008, Karas et al., 1985). Hillenkamp noted that background signals made the spectrum hard to decipher, and that this background noise appeared to form a general pattern. It was later suggested that these signals may be fragment ions originating from the organic matrix, leading Karas and Hillenkamp to study the desorption of small organic molecules, and eventually proteins (Griffiths, 2008, Karas et al., 1985).

The breakthrough which enabled the ionisation of large biomolecules, however, which involved the analyte, dissolved in glycerol, being combined with cobalt particles, was first described by Koichi Tanaka in 1987 (Tanaka et al., 1988), for which he shared the Nobel Prize in Chemistry in 2002.

In terms of an Ultraflex II MALDI based analysis, the target protein mix is dissolved in a suitable solvent and mixed with an excess amount of matrix (generally α -cyano-4-hydroxycinnamic acid (HCCA) for peptides of less than 5 kilodaltons (kDa) and 3,5-dimethoxy-4-hydroxy-cinnamic acid for larger peptides and proteins). The sample is spotted to an instrument specific steel MALDI plate, air dried, and bombarded by a photon beam from a smartbeam laser at a wavelength of 355 nanometers (nm) (Chait and Kent, 1992, Kicman et al., 2006). Upon entering the gas phase a voltage is applied between a grid or ring electrode and the inlet orifice of the MS, accelerating the ions into the mass analyser, as is shown on Figure 1-3.

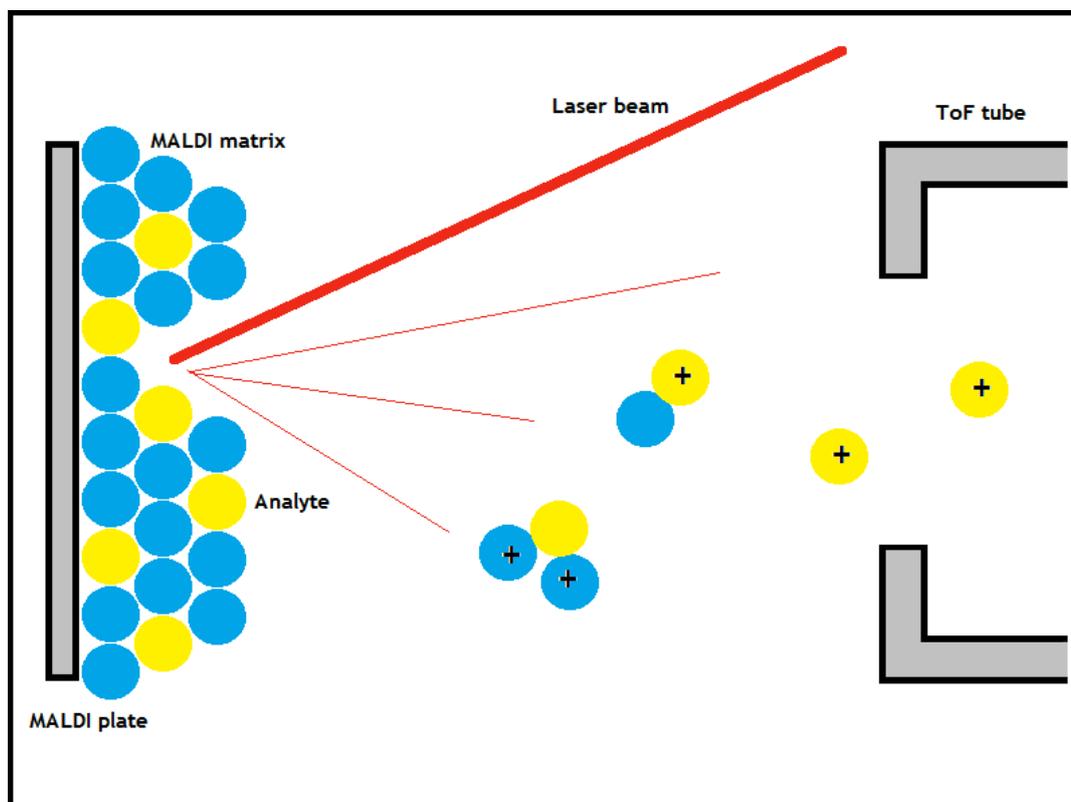


Figure 1-3: The route through which analyte molecules are ionised during MALDI. Shown above is a diagrammatic representation of how an analyte molecule is ionised during MALDI. A laser shot is fired at the spotted analyte sample, ionising any matrix molecules in its path. Subsequently an ionised matrix molecule can transfer a proton to the analyte molecule, enabling its acceleration into the ToF tube.

The advantages of MALDI include its large working mass range (capable of ionising molecules of between 100 Da and 300 kDa), enabling the rapid ionisation of both peptides and intact proteins. Furthermore, in terms of data interpretation, the predominance of singly charged ions lends itself to a relatively straightforward data interpretation (Hillenkamp and Karas, 2000).

MALDI based ionisation is generally coupled with ToF based MS, due to the large working mass range of the ToF detector, and the fact that the pulsed nature of MALDI ionisation yields spatially distinct packets of ions suited for ToF based detection (Brown and Lennon, 1995a).

1.5.1.2 Electrospray Ionisation

The development of ESI can be traced back to Dole and his co-workers, who in the late 1960s, successfully introduced a polystyrene polymer into the gas phase as a charged species (Dole et al., 1968). Unlike MALDI based ionisation where the exact mechanism of proton transfer has yet to be described, ESI based ionisation is seen as being simplistic in nature. The protein species to be analysed, dissolved in a polar solvent, are sprayed from the tip of a highly charged needle (3-4 kilovolts (kV)) towards the inlet of the MS (Fenn et al., 1989, Kicman et al., 2006, Falk et al., 2006). These highly charged droplets are subjected to a warm flow of nitrogen, vaporising the solvent, and in doing so increasing the repulsive forces encountered between the ions on the surface of the shrinking droplet. Eventually these repulsive forces exceed the surface tension of the solvent (the Rayleigh limit), ripping the droplet apart (Coulombic explosion), a process which is repeated until the intact ions are desorbed into the gas phase (Fenn et al., 1989, Jonsson, 2001), as shown on Figure 1-4.

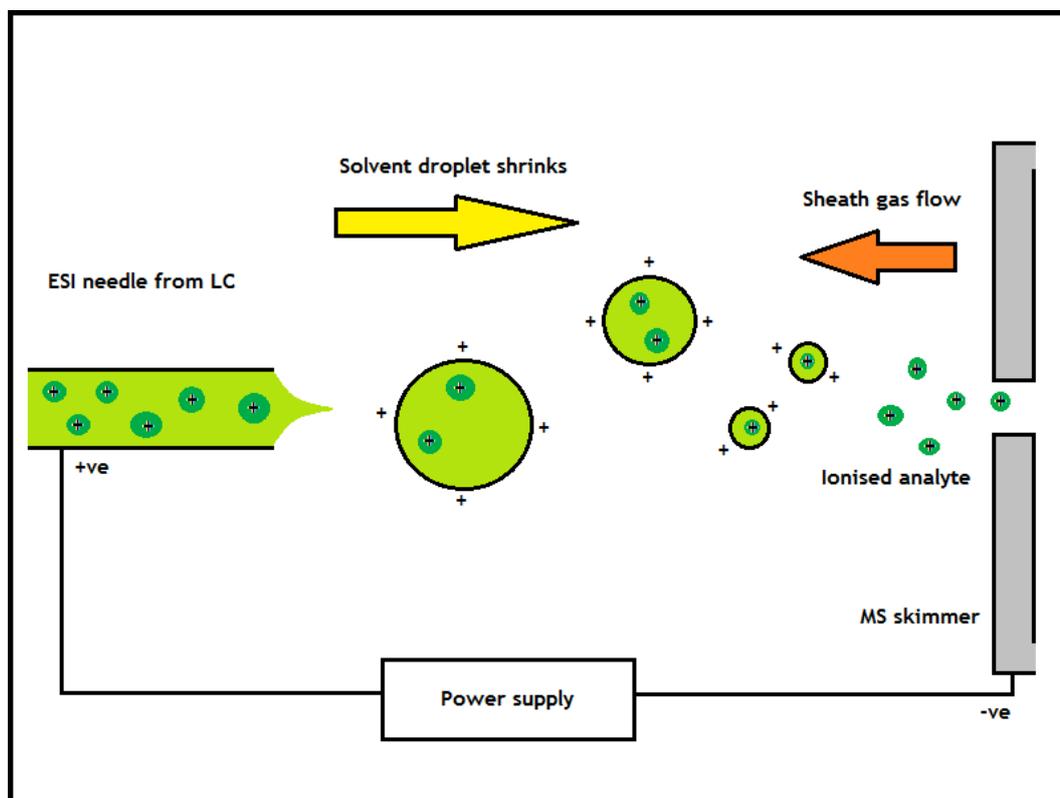


Figure 1-4: The route through which an analyte molecule is ionised during ESI. Upon the analyte molecules being eluted from the LC they are carried to the ESI needle before being sprayed towards the orifice of the MS through applying a voltage between the needle and the orifice. As the solvent droplets move towards the orifice they are dried due to the warm flow of a sheath gas (generally nitrogen). As the droplets shrink the surface charge grows, leading to a coulombic explosion, splitting the solvent droplet into several smaller droplets, a process which repeats until only the intact analyte ions remain.

Despite the presence of multiply charged ion species adding to the complexity of the data to be analysed (first requiring the charge state of each ion to be determined), these multiply charged species are one of the big advantage of ESI. In carrying multiple ions, the m/z of the target biomolecule is lowered, enabling its detection, and thus analysis, on a greater range of mass spectrometers (Bakhtiar and Tse, 2000). Furthermore, as the target biomolecule leaves the ESI needle in the liquid phase, RP-LC can be coupled directly to a mass spectrometer, without the need for manual sample manipulation and plate loading, as would be the case with LC-MALDI.

Due to the constant beam of ions created during ESI, this ion source can be coupled with the vast majority of modern mass spectrometers, including quadrupoles, orbitraps and fourier transform-ion cyclotron resonance (FT-ICR) based mass spectrometers. Indeed ESI can even be coupled with a ToF based mass analyser, though this particular setup will require the installation of an ion pusher electrode, pulsing well defined ion packets into the ToF tube (Kicman et al., 2006, Loboda and Chernushevich, 2009).

1.5.2 Mass spectrometers

Single stage mass spectrometers can be seen to perform two distinct tasks; to filter or separate the ions which enter the instrument, and to determine the m/z of those ions which are permitted to the detector. Indeed, while the development of ESI and MALDI in the late 1980s enabled the ionisation of large biomolecules, the data which could be obtained based upon the existing range of single stage mass spectrometers was limited (El-Aneed et al., 2009). It was thus the demand for peptide and protein sequence data, and the ability to identify PTM sites which led to the development of a new range of mass analyzers and complex multistage instruments, designed to tackle the challenges of the proteome (Domon and Aebersold, 2006).

Described over the following few pages are some of the most common types of mass analysers and modern instrument setups.

1.5.2.1 Time-of-Flight Mass Spectrometry

Upon the target biomolecule being ionised in ToF based MS (most commonly through the application of MALDI); a high voltage is applied between either a grid or ring electrode and the inlet of the ToF tube, accelerating the ion plume towards the detector. As each ion is accelerated with the same force, the same kinetic energy will be imparted. Each ion should therefore move towards the detector at a different velocity based on its m/z (Merchant and Weinberger, 2000, Khalsa-Moyers and McDonald, 2006). In practice, however, the ionisation

and immediate acceleration of the ion packet into the ToF tube can lead to each ion possessing one of a range of kinetic energies, broadening the final peak detected and reducing the resolution achieved (El-Aneed et al., 2009). In an attempt to correct for this inconsistency, two improvements have been introduced to modern ToF based mass spectrometers. The first, termed delayed extraction, causes the sample to be ionised under only a weak electrical field (Vestal et al., 1995). Under the influence of this weak field, each ion moves slowly towards the ToF tube, based on its initial velocity, later being subjected to high voltage based acceleration. In delaying the high energy acceleration step by approximately 100 nanoseconds (ns), the lower velocity ions are affected more by the acceleration than those ions present at a higher velocity, enabling the lower velocity ions to “catch up” (Vestal et al., 1995). The second improvement introduced to ToF based mass spectrometers has been the application of an electrostatic ion mirror, or reflectron, which acts to change the flight path of the ions towards an offset detector (El-Aneed et al., 2009, Mamyrin et al., 1973). In addition to increasing the flight path of the ions, yielding an increase in resolution (Clauser et al., 1999), the reflectron is also penetrated further by higher kinetic energy ions than low, helping to better focus the ions based on their m/z .

ToF based mass analysers have several advantages over quadrupole based mass analysers. Firstly, as a result of the ToF based mass spectrometer separating the ions rather than filtering, each ion introduced into the MS will reach the detector. Furthermore, ToF based mass analysers have a higher resolution, accuracy and mass range than quadrupoles. By way of an example, when a ToF based mass detector is coupled with a MALDI source, the upper mass range will be limited to approximately 300 kDa (Hillenkamp and Karas, 1990). When the same ToF based mass spectrometer is coupled with an ESI source however, the upper mass limit can extend into the MDa (mega Daltons) (El-Aneed et al., 2009).

In addition to simply determining the accurate mass of an ion, MALDI-ToF based instrumentation can also be used to probe the structure of a biomolecule based

on one of several techniques. One such technique, termed post-source decay (Purcell and Gorman, 2001), fragments the target ions in one of two ways after the ions have been subjected to high voltage acceleration (Suckau et al., 2003). Fragmentation is achieved, firstly, through colliding the analyte ions with the ionised matrix while the ions are being “ripped” from the ion plume into the ToF tube, in addition to which, while transversing the ToF tube, vibrationally excited metastable ions can decay, yielding a-, b- and y- ions (Kaufmann et al., 1994).

The names of these various fragment ions denotes the position at which the peptide backbone is fragmented, based upon Roepstorff-Fohlmann-Biemann nomenclature (Roepstorff and Fohlman, 1984). In brief, the terms a_n , b_n and y_n indicate both the distance at which the fragmentation occurred from the N-terminal (a/b/y), and the number of incorporated amino acids (n). While the terms z_n , y_n and x_n are used to indicate both the distance the fragmentation occurred from the C-terminal, and the number of incorporated amino acid residues (Steen and Mann, 2004).

An alternative MALDI-ToF based fragmentation method can also be sought in in-source decay (Brown and Lennon, 1995b), which differs from post-source decay in that the fragmentation occurs prior to the application of the high voltage acceleration. In in-source decay, a combination of laser induced fragmentation and rapid metastable decay result in the production of both c- and z-ions (Köcher et al., 2004).

Finally, tandem ToF based instruments (ToF/ToF) are also capable of performing collision induced dissociation (CID) (Medzihradszky et al., 1999), fragmenting the analyte across the peptide bond so as to yield both b- and y-ions, immonium fragment ions and internal d- and w-fragment ions (Domon and Aebersold, 2006). Unlike both post-source decay and in-source decay, CID involves accelerating the isolated ions and colliding them with inert gas molecules (generally nitrogen) (Shukla and Futrell, 2000).

In terms of modern hybrid ToF based instrument setups, both ToF/ToF and Q-q-ToF instruments exist (El-Aneed et al., 2009, Domon and Aebersold, 2006). Through combining two ToF based instruments, the first ToF tube can be applied solely to the separation of the ion plume, while a Bradbury-Nielsen velocity filter is placed between the two ToF stages, so as to permit only a single ion of interest into the second ToF tube. The second ToF tube can then be utilised to perform both CID, and to separate the product ions generated, thereby providing true MS/MS capability. In terms of applications, ToF/ToF based instruments are used regularly in quantitative proteomics, based upon the application of LC-MALDI, in addition to being utilised in 2DGE based proteomics. As has previously been discussed (Section 1.4), the possession of MS/MS fragmentation data can offer a vast improvement in the precision in the identity assigned to a protein, compared to a PMF (Khalsa-Moyers and McDonald, 2006, Mann and Wilm, 1994).

ToF based MS/MS may also be achieved through the use of ESI. In a typical Q-ToF based instrument, the first quadrupole will be tasked with filtering the ion beam, permitting only a target m/z to the second quadrupole. At this stage the second quadrupole can be operated as a collision cell, fragmenting the target ions and directing the product ions into the ToF based mass analyser (El-Aneed et al., 2009). In addition to utilising ESI based ionisation, and thus having an increased mass range when compared to a standard MALDI-ToF, Q-ToF based instruments also benefit from the high resolving power of the ToF mass analyser, making them ideal for the study of PTMs (Domon and Aebersold, 2006).

1.5.2.2 Quadrupole Mass Spectrometry

The quadrupole based mass analyzer was first described by Wolfgang in 1953 (El-Aneed et al., 2009, Paul, 1990) and consists of two sets of linked parallel circular metal rods, through which a cycled direct-current (DC) voltage is applied. Further to the imparted DC voltage, an oscillating radio-frequency (RF) is also applied to each rod; the combination of which enables the manipulation of the ion beam in such a way as to either stabilise or destabilise certain m/z 's passing through the quadrupole, as is shown on Figure 1-5. Indeed, by increasing

the DC voltage applied to the quadrupole, the transmission window can be shrunk so as to permit only a specific m/z range to reach the detector (operationally termed Q) (Domon and Aebersold, 2006). It is therefore possible to utilise a quadrupole to either scan through a specified mass range, or to permit only a specific m/z to constantly reach the detector (Khalsa-Moyers and McDonald, 2006). The opposite can also be achieved, however, decreasing the DC voltage applied to the quadrupole (operationally termed RF only) results in a much wider beam of ions being permitted to reach the detector (Henchman and Steel, 1998). By introduction of a stream of inert gas to a RF only quadrupole, the instrument can be operated as a CID based collision cell (operationally termed q). Finally, the quadrupole may also function as a linear ion trap (LIT) (Khalsa-Moyers and McDonald, 2006). Through applying a static DC current to each end of the four rods, the stream of ions can be trapped in the quadrupole's oscillating RF field. Slowly increasing the DC voltage of the quadrupole gradually releases the ions based on their m/z (Hager, 2002).

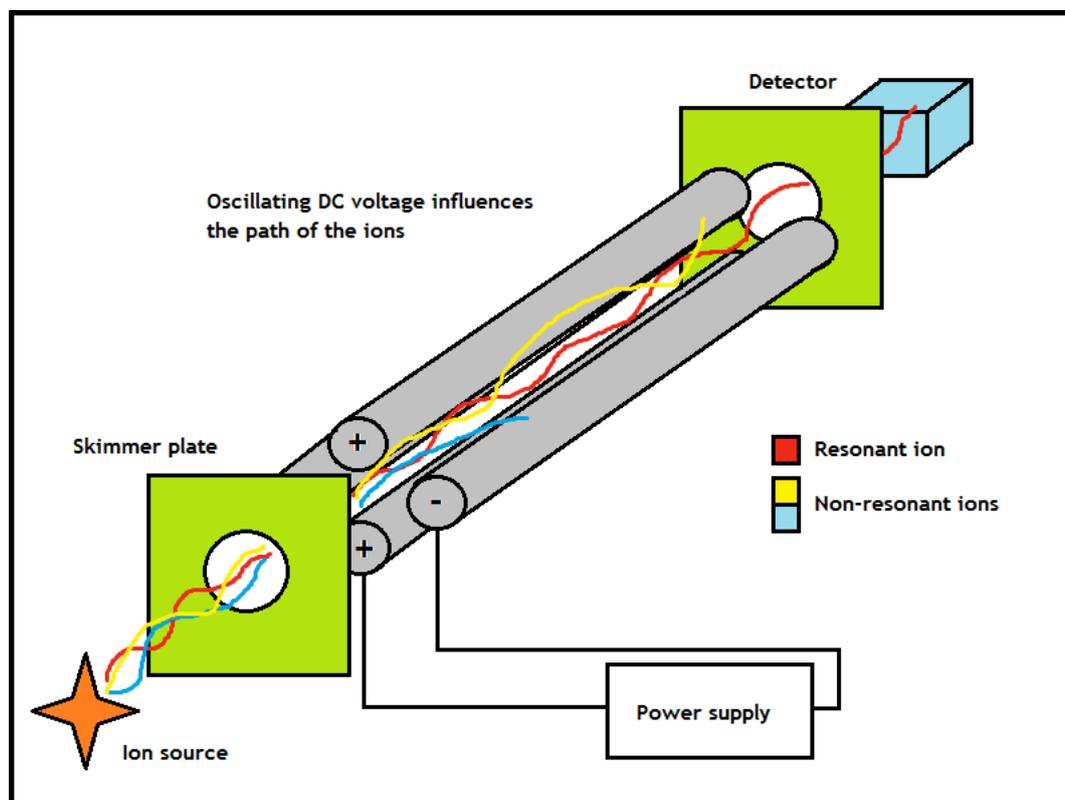


Figure 1-5: A diagram depicting the workings of a quadrupole based MS. The above diagram depicts the workings of a quadrupole based MS, where the ion beam travels through a skimmer and between four metal rods imparted with both an RF frequency and an oscillating DC voltage. Through modifying both the DC and RF parameters it is possible to permit only a single m/z range to the detector, to scan through a range of m/z 's or to stabilise all ions entering the MS. Those ions which do not resonate at a frequency similar to that of the quadrupole will be destabilised, striking one of the rods, and in doing so becoming neutralised.

While single stage quadrupole based MS are relatively cheap, small, robust and easy to use, they have limited applications in the field of proteomics (El-Aneid et al., 2009). Quadrupole based mass spectrometers are therefore often found in tandem, with the most common instrument setups being the triple quadrupole based MS (Q-q-LIT), and the Q-q-ToF based MS (Domon and Aebersold, 2006).

In performing a general MS/MS based analysis on a triple quadrupole MS, the first two quadrupole mass filters operate in much the same way as they would in a Q-ToF based instrument, where the first quadrupole is used to filter the ion beam, while the second acts as a collision cell. While the final MS stage of the Q-ToF

based MS involves each Q2 fragment ion being separated within the ToF tube based on its m/z , enabling a high resolution analysis, the final mass filter in a triple quadrupole based instrument can be set to either trap and detect each fragment ion at a low resolution, or to permit only a single ion of interest to the detector. Indeed, it is this versatility, and the variety of scan modes offered by triple quadrupole mass spectrometers which has resulted in them being one of the most abundant MS in the world (Wang and Stout, 2007).

One such scan mode, termed selected reaction monitoring, or multiple reaction monitoring (MRM) when analysing more than one transition, utilises both quadrupoles one and three as mass filters, permitting only a pre-defined parent ion (Q1) and pre-defined product ion (Q3) to react the detector, while quadrupole two is utilised as a collision cell. Through utilising this configuration it is possible to perform a highly selective analysis on a target analyte, or analytes, mapping both retention time (assuming LC is utilised) and the signal intensity detected over the duration of a run (Lange et al., 2008). As this MS acquisition method filters all but those ions which share both a similar parent and product ion m/z , the MRM spectrum contains very little noise, this type of analysis is therefore ideal for the detection of low concentration biomarkers, present within complex samples, as is commonly required in proteomics.

1.5.2.3 Fourier Transform-ion Cyclotron Resonance Mass Spectrometry

FT-ICR was first conducted by Comisarow and Marshall in 1974 (Comisarow and Marshall, 1974), who in taking inspiration from the recent description of fourier transform-nuclear magnetic resonance, applied the same fast fourier transform calculation to ion cyclotron resonance spectrometry, and in doing so reducing the time taken to complete an analysis by several orders of magnitude.

An FT-ICR functions by trapping an ion packet in a Penning trap, formed by six metal plates in a box configuration, placed within a super-conducting magnet under a high vacuum, as is shown on Figure 1-6. In addition to the strong

magnetic field acting to stabilise the ion packet in two of three dimensions, two opposing metal plates running perpendicular to the IRC cell (termed trapping plates) are activated upon the ion packet entering the cell, forming an electrostatic field which traps the ions in the remaining dimension (Steen and Mann, 2004).

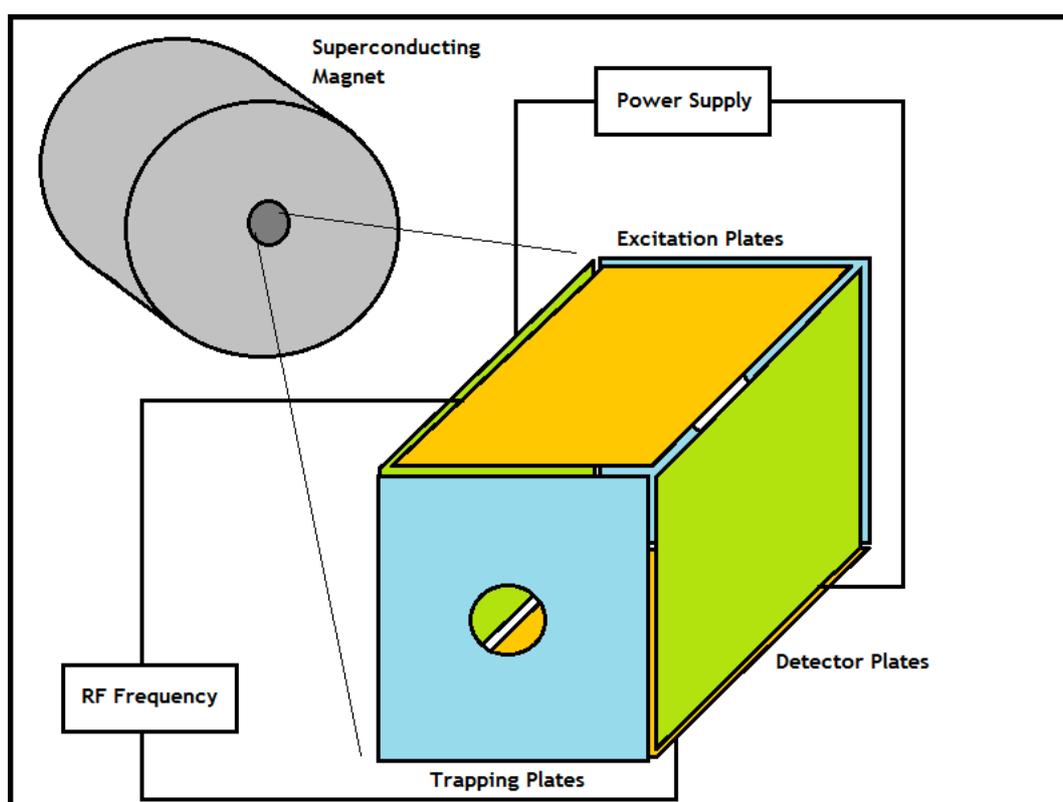


Figure 1-6: The components contained within an FT-ICR MS. Depicted above are the basic components contained within an FT-IRC based MS, including two trapping plates (marked in blue), preventing the ion packet from escaping the IRC cell, two excitation plates (marked in orange), capable of amplifying the ions oscillation path, and two detector plates (marked in green), capable of detecting the ions m/z based on a change in the alternating current.

The technique by which the m/z of the ions is measured in FT-ICR differs from that of both ToF based mass spectrometers and quadrupole based instruments, in that, the mass is determined indirectly as a measure of the frequency at which the ions oscillate in proximity to the two opposing detection plates within a magnetic field (Khalsa-Moyers and McDonald, 2006). In order for a detection to be made, the ions must thus be cycled within the IRC cell. This is achieved

through two distinct actions; firstly, when the ion packet is being impelled into the ICR cell, through the application of a weak DC voltage, the ions are subjected to a rapid DC pulse, imparting an initial gyrating motion. Secondly, the radius at which the ions are oscillating is amplified through the application of an RF voltage, applied to the two remaining opposing metal plates (termed RF plates). As the ions move both towards and away from the pair of detection plates, the m/z is measured in the form of an alternating current. When plotted, this current constitutes a sine wave, with the amplitude being proportional to the number of ions detected and the wavelength equal to the frequency at which ions are oscillating. Finally the data is deconvoluted through the application of the fast fourier transform, so as to yield a mass spectrum.

FT-ICR has been coupled with both MALDI (Cornett et al., 2008) and ESI (Irungu et al., 2008), and can be found in tandem with both a single LIT, or indeed with triple quadrupole mass filters, enabling both ultra-high resolution and high mass accuracy MS/MS and MS_n scanning (Domon and Aebersold, 2006). Both top-down and bottom-up based proteomics have been conducted through the FT-ICR (Dodds et al., 2007). It should be noted, however, that as the ions are trapped within the ICR cell at a low voltage, the kinetic energy of the ions must be kept to a minimum. This has led to the description of several new fragmentation techniques; including ECD (Sleno and Volmer, 2004), infrared multiphoton dissociation (IRMPD) (Little et al., 1994), sustained off-resonance irradiation collision-induced dissociation (SORI-CID) (Herrmann et al., 2005), and a combination of the latter, combination of infrared and collisional activation (Dodds et al., 2007).

Of these new fragmentation techniques both IRMPD and SORI-CID are considered “gentle” ion activation techniques, yielding fragment ions from the application of low-energy processes (Dodds et al., 2007). In regard to SORI-CID, the ions oscillating within the IRC cell are excited by the application of an “off-resonance” RF pulse, decelerating the ions prior to re-acceleration in the presence of a low-mass collision gas (Herrmann et al., 2005). IRMPD, meanwhile, involves the firing an infrared laser pulse into the IRC cell, slowly increasing the

energy of the ions until a bond is broken (Little et al., 1994). Both SORI-CID and IRMPD, however, have been found to fragment the biomolecules only at the sites, which require the lowest amount of energy. This has not therefore proven particularly useful for the fragmentation of singly protonated tryptic peptides, such as those obtained during MALDI. Both SORI-CID and IRMPD have therefore been combined to yield more informative fragmentation spectra than could be achieved using either IRMPD or SORI-CID alone (Dodds et al., 2007).

Standing as an alternative to both IRMPD and SORI-CID, each of which involve the excitation of an ion up until the point of fragmentation, ECD involves colliding an ionised biomolecule (which must possess multiple positive charges) with an electron, neutralising a single positive charge, and in doing so extensively cleaving the backbone of the biomolecule (Sleno and Volmer, 2004). In practice ECD is achieved by heating a metallic filament within the IRC cell so as to produce free electrons, which prior to collision, will oscillate in a direction opposite to those ions possessing a positive charge (Khalsa-Moyers and McDonald, 2006). As the target ion is stripped of a positive charge through ECD, it is applicable only to multiply charged ESI based ions.

Due to the fact that FT-ICR instruments acquire each spectrum as a frequency, and that a frequency may be measured with extreme accuracy, FT-ICR instruments offer a resolution of greater than 100,000 (the m/z value divided by the peak width at the half the peak height) (Steen and Mann, 2004). FT-ICR instruments are, however, expensive and bulky (requiring a large superconducting magnet and liquid helium and nitrogen storage), and thus have so far been limited to expert-only laboratories (Khalsa-Moyers and McDonald, 2006).

1.5.2.4 Orbitrap Mass Spectrometry

The orbitrap has been hailed as the first truly novel mass spectrometer to be introduced in over 25 years (Domon and Aebersold, 2006), with the term orbitrap describing how an ion rotates around an electrode with a harmonic oscillation

pattern indicative of its m/z (Makarov, 2000). The way in which a mass spectrum is generated on the orbitrap is similar to that of the FT-ICR MS, in that, upon a change in current being detected, the recorded data is subjected to a FFT so as to produce the final mass spectrum. As a result of this broadband method of ion detection, and the application of a fast fourier transform, the orbitrap boasts a resolution and mass accuracy similar to that of the FT-ICR (Makarov, 2000), without the burden of an expensive superconducting magnet.

The orbitrap itself consists of a split barrel-like outer electrode and a spindle-like centre axial electrode, each of which are held at constant DC voltage so as to form an electrostatic ion trap, as is shown on Figure 1-7. As the DC voltage within the trap is constant and no ion excitation applied, the ions are instead introduced to the trap at a position offset from the equator of the barrel, with each ion eventually falling into a radial oscillation with a frequency proportional to its m/z (Makarov and Scigelova, 2004).

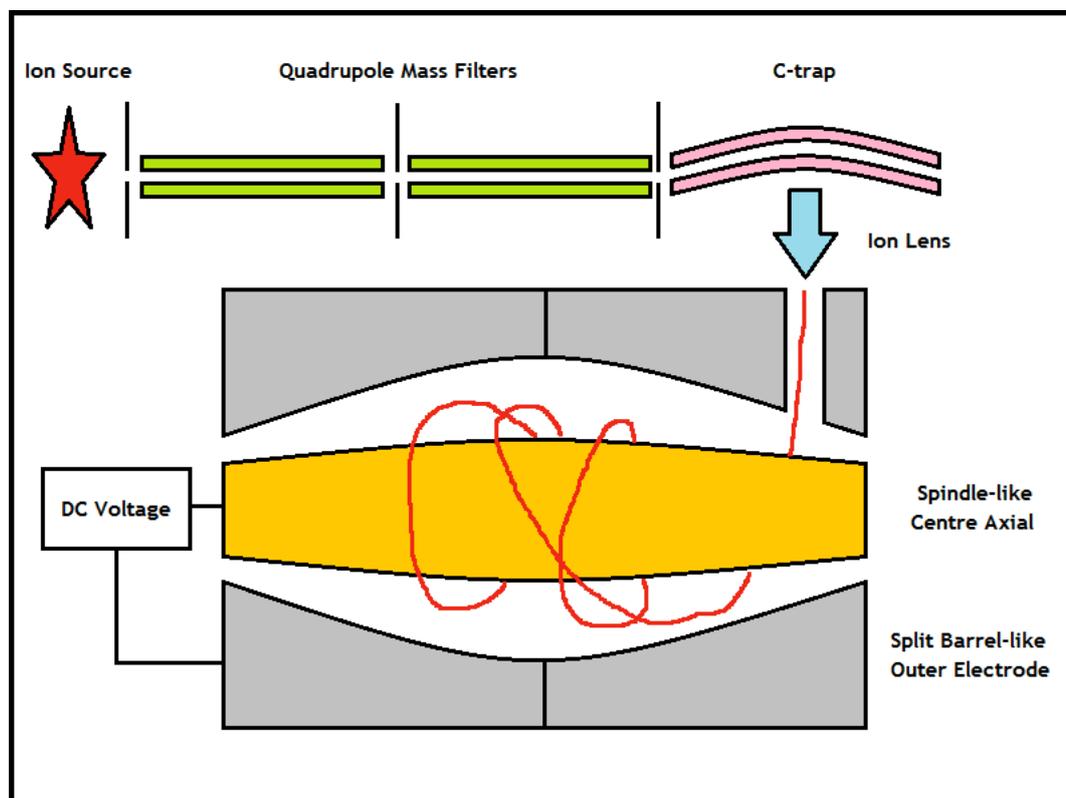


Figure 1-7: The components contained within a hybrid quadrupole / Orbitrap MS. Depicted above are the main components utilised within a hybrid quadrupole orbitrap instrument, including two quadrupoles, enabling the scanning, trapping or fragmentation of the ion beam, a C-trap, acting as an LIT and an orbitrap based detector, measuring the m/z of each ion based on its oscillation path.

Both MALDI (Luo et al., 2010), and ESI ion sources (Hu et al., 2005) have been applied to orbitrap based mass spectrometry, and the orbitrap is generally utilised in tandem with several quadrupole based mass filters. In addition to each quadrupole enabling a specific m/z to be scanned, trapped or stabilised, each distinct instrument stage also enables the atmospheric pressure to be further reduced (Makarov and Scigelova, 2004).

The final component of many orbitrap instrument setups, prior to the orbitrap itself, is a unique curved LIT, termed a “C-trap”. By trapping the ion beam before it enters the detector, the ions can be introduced into the orbitrap in a pulsed fashion, similar to that required for both ToF and FT-ICR based mass spectrometry. In addition to trapping the ion stream, the C-trap can also be

utilised as a collision cell, where-upon raising the RF voltage of the quadrupole, can be seen to result in the “triple quadrupole-like” fragmentation of the trapped ions (Olsen et al., 2007).

1.6 Protein Quantitation

Up until the last decade, the field of proteomics was largely seen as a qualitative discipline, with a typical proteomic experiment consisting of a tissue or culture of interest being digested, assessed, and a list of proteins generated, without any interest having being paid to their abundance or cellular distribution (Baak et al., 2005). Indeed, this was in stark contrast to the field of genomics, which at the time had widely adopted quantitative strategies for the analysis of gene expression through the use of both microarrays and quantitative polymerase chain reaction (PCR) (qPCR) (Schulze and Usadel, 2010, Regnier and Julka, 2006). As has previously been discussed, however, the level of mRNA detected during such an analysis does not necessarily reflect the number of proteins within a cell, which, ultimately are the workhorses, driving the enzymatic reactions required to sustain life.

Based upon the success of both qualitative proteomics and quantitative genomics, the proteomics community soon began to realise that similar comparative protein based analyses could be achieved (Regnier and Julka, 2006), enabling the comparison of disturbed and undisturbed proteomes, cells grown in different media, diseased cells compared with healthy, and indeed the tracking of a disease’s progression.

Having the capacity to visualise over 10,000 distinct protein spots, equating to over 1000 unique proteins, 2DGE was applied quite early on in the field of comparative proteomics (Schulze and Usadel, 2010, O’Farrell, 1975). Problems, however, were encountered due to both inter-gel variability, and the chemistry through which a variety of protein stains and dyes bind each biomolecule, preventing both inter- and intra-gel spot quantitation (Marouga et al., 2005). High-confidence 2DGE based comparative proteomics was thus delayed until

1997, when DIGE was first described. This features two differentially labelled samples being combined and co-analysed via 2DGE (M. Unlu et al., 1997).

While comparative proteomics was also eventually conducted via LC-MS, its application was delayed due to the variable ionisation efficiency detected for a range of biomolecules, in addition to which varying sample components and hardware performance at the time of use can limit the reproducibility of this technique (Schulze and Usadel, 2010). So as to better explain, when a protein is proteolytically digested, an equal number of each proteolytic peptide should be liberated, the signal intensity achieved for each peptide during an MS scan can, however, still vary by several orders of magnitude, assuming each peptide is detected (Steen and Mann, 2004). This variation results from several independent factors, including the accessibility of the digest site to the protease, the solubility of the peptide (having been dried post-extraction) and the efficiency with which each peptide ionises within the MS source. It was thus realised that the only way in which a quantitative proteomic comparison could be made via MS would be if one of the samples were to be labelled with a stable isotope (^{13}C , ^{15}N , ^2H or ^{18}O), and both samples co-analysed via LC-MS (Vaudel et al., 2010). Under such conditions each biomolecule should elute from the LC at the same time, both molecules should ionise under the same conditions, and each peptide should thus be directly comparable based on a slightly differing m/z (Schulze and Usadel, 2010).

Discussed over the following few pages are some of the most commonly used quantitative proteomic techniques, based on both 2DGE and LC-MS.

1.6.1 Difference gel Electrophoresis

Due to the varying conditions under which a poly-acrylamide gel is cast, the slightly varying components within each sample, and the way in which each sample traverses the 2D gel, 2DGE has been documented to contain a great deal of variation, making it unsuitable for use in comparative proteomics. In an effort to overcome these problems, Minden and co-workers in 1997 devised a technique

termed DIGE, which features the use of two molecular weight and pI-matched fluorescent cyanine dyes (Cy3, and Cy5), each covalently labelling approximately 3% of the Lys residues within a given sample (minimal labelling). Upon completion of labelling, each sample is combined and co-separated on a single 2D SDS-PAGE gel, with the samples being compared through the application of fluorescent imaging, with a sensitivity equal to that of silver staining (~1 ng) (M. Unlu et al., 1997), as depicted on Figure 1-8. A third cyanine dye (Cy2) can also be employed to label an additional pooled sample, so as to act as an internal standard, accounting for any variation in labelling efficiency (Lilley and Friedman, 2004).

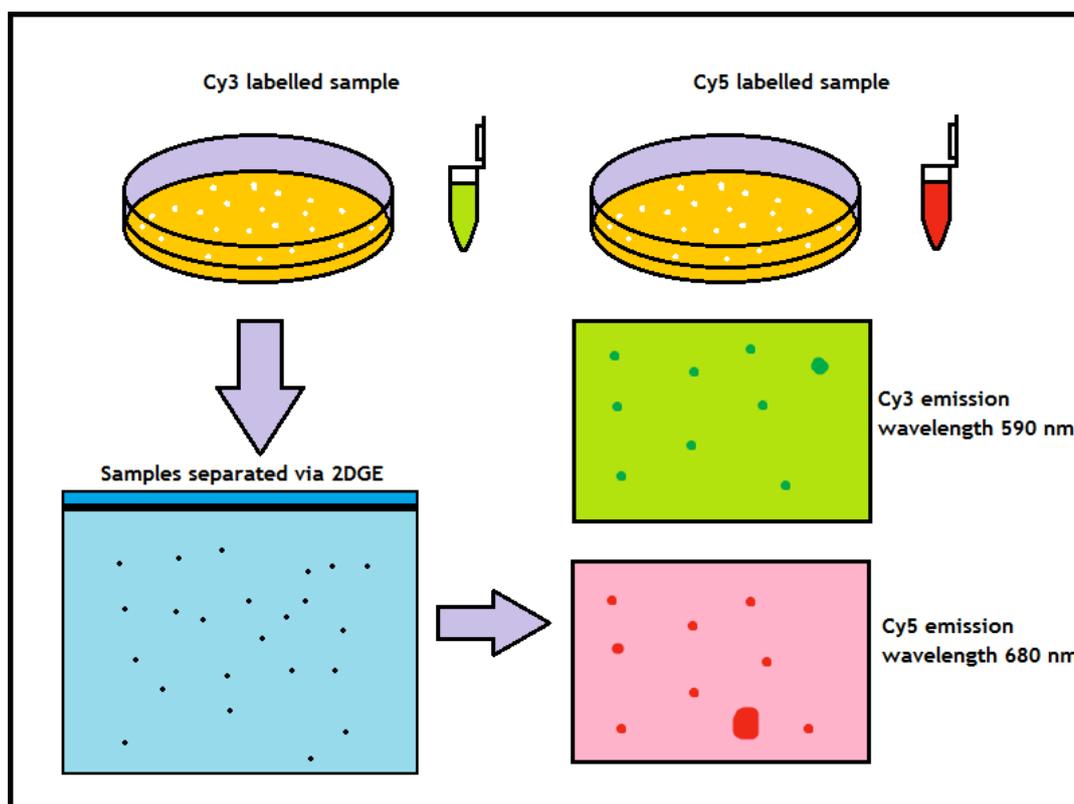


Figure 1-8: The workflow through which two proteomes can be compared based on DIGE. Depicted above is the workflow used in a DIGE based proteomic analysis, including the growth, preparation and labelling of the two samples (with Cy3 and Cy5), their mixing, and separation based on 2DGE. Upon the samples being separated each is then compared based on a differing excitation and emission wavelength, enabling the protein expression patterns to be assessed.

Despite minimal labelling being applied to this technique with the best of intentions, those being to label only one Lys residue per protein (preventing multiple gel spots per protein), both to reduce the volume of dye required per analysis and to prevent sample solubility being compromised, the minimal labelling strategy actually became the greatest failing of this technique (Wu et al., 2006, M. Unlu et al., 1997). While a slight alteration in protein mobility within the second dimension has indeed been detected in labelled samples, the minimal labelling strategy labels only a small percentage of the total protein content. This becomes a particular problem when a protein identification is required for a specific protein spot as the location of the majority of the protein and the visualised protein spot do not match up (termed dual spot migration) (Tonge et al., 2001).

In an attempt to improve upon the minimal labelling strategy, a variation of DIGE was described by Shaw and co-workers in 2003 which aimed to completely label a less prevalent amino acid, Cys, which accounts for just 2.47% of amino acids across eight random proteins (in comparison to the 10.1% accounted for by Lys) (Shaw et al., 2003). In addition to preventing dual spot migration, the improved fluorophores used in this reaction have been documented so as to offer the technique a 100- to 200-fold increase in sensitivity.

Several disadvantages have been noted, however, in the application of saturation DIGE. Firstly, the electrophoretic mobility of the labelled proteins does appear to be significantly reduced. This was one of the main reasons the minimal labelling strategy was initially employed (Kondo and Hirohashi, 2007). It should also be noted that any given spot on a 2D gel can contain more than one protein, making quantification difficult (Zhu et al., 2010). Finally, it has been noted that approximately 5% of the proteins within the human proteome do not contain a Cys residue, further restricting the number of proteins to which 2DGE can be applied.

1.6.2 Stable Isotope Labelling by Amino Acids in Cell Culture

As was discussed in Section 1.6., co-eluting sample components and instrument performance at the time of use can affect the reproducibility of the ion intensity achieved for any given biomolecule accessed via MS. Indeed the only way in which a biomolecule can be accurately compared between samples or instruments based upon the use of MS is through the introduction of one of a number of stable isotopes, including ^{13}C , ^{15}N , ^2H or ^{18}O . Several technical limitations do exist with the utilisation of these isotopes within comparative proteomics, including the difficulty in substituting atoms within a protein (^{13}C and ^{15}N). While hydrogen can be readily substituted by deuterium, this reaction has been noted to cause a small, but significant shift in retention time during RP-LC based separation (Bantscheff et al., 2007).

In an attempt to overcome these problems both Lahm and Langen (Lahm and Langen, 2000), and Chait and co-workers (Oda et al., 1999), utilised ^{15}N -substituted media in the growth of a range of autotrophic microorganisms so as to assess any proteomic changes occurring during different states of growth. While the application of this technique appears to have been successful, the degree of ^{15}N incorporated was not necessarily complete. Additionally, as each peptide contained a varying number of nitrogen atoms, the subsequent mass shift was often hard to predict (Ong et al., 2002). Lastly, while the production of the ^{15}N media for autotrophic organisms proved costly, its application in higher organisms would certainly prove even more expensive, restricting the use of this technique (Ong et al., 2002).

In an attempt to avoid the pitfalls of ^{15}N labelling, Ong and co-workers in 2002 described a comparative proteomic technique termed stable isotope labelling by amino acids in cell culture (SILAC), which utilised a growth media containing two ^{13}C labelled essential amino acids (Arg and Lys) so as to introduce a small mass shift between each protein suitable for detection on a MS (Ong et al., 2002), as is depicted on Figure 1-9. By selecting to label both Lys and Arg, and assuming the use of trypsin, each proteolytic peptide should in theory be labelled only once, ensuring a constant mass shift and enabling a rapid analysis (Vaudel et al.,

2010). Indeed, one of the big advantages of SILAC is that each sample is combined at the cellular level, prior to the application of any biochemical processing or the variability of MS. In theory this should help to ensure that any sample losses encountered should affect both samples equally (Unwin and Whetton, 2007).

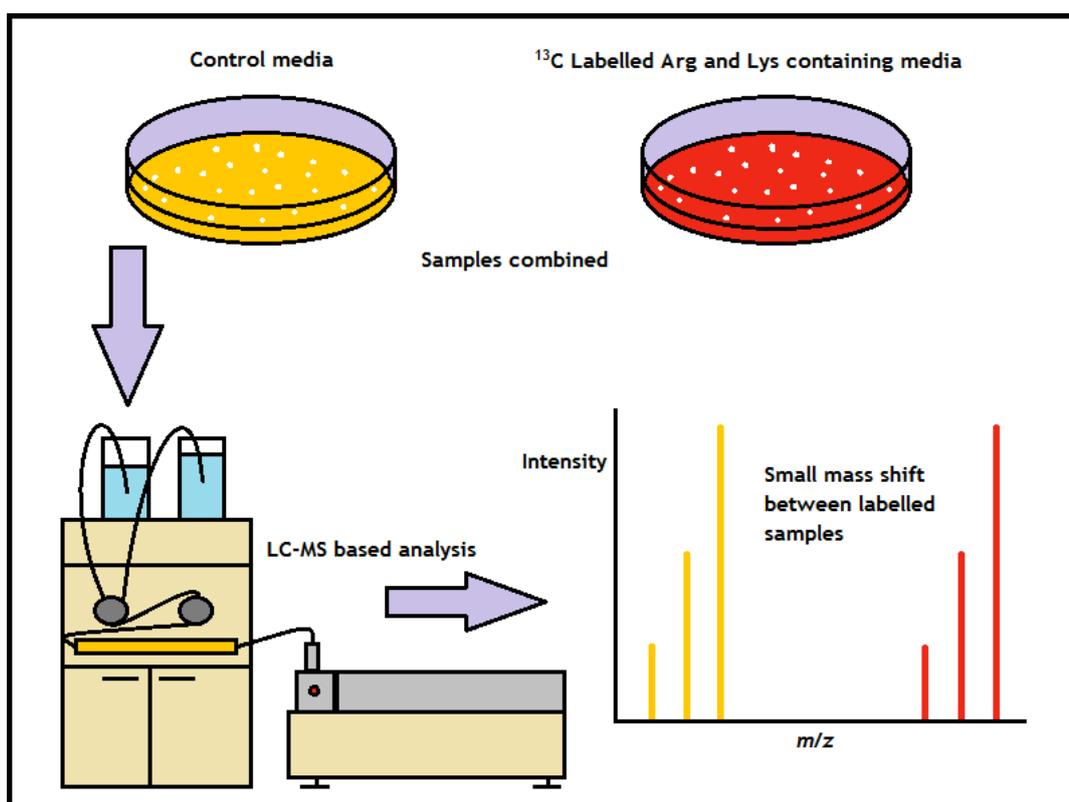


Figure 1-9: The stages involved in a SILAC based proteomic analysis. Depicted above are the stages involved in a SILAC based proteomic comparison of two samples. While one sample is grown on a control (light) media, the second sample is grown on a media containing both ¹³C labelled Arg and Lys (heavy). Upon cell lysis the samples are combined, digested, separated via LC and analysed via MS. Due to the small MW difference between the two samples, introduced by the labelled amino acids, the two peptides can be differentiated based on their parent ion *m/z*, and a quantitative comparison made.

In practice, however, SILAC has been shown to require some 6-8 passages in modified media to achieve a labelling efficiency of just 90% (Ong et al., 2002). Furthermore, it has been noted that an excess of Arg within the labelled media can lead to the production of labelled Pro, a process which further complicates the data analysis (Schulze and Usadel, 2010). Finally, only a limited number of

cell lines are suited to growth on dialyzed serum, limiting the application of SILAC (Fenselau, 2007).

1.6.3 ^{18}O Incorporation

Chemical-based labelling techniques can also be applied to quantitative proteomics, achieving the same mass shifts associated with cells grown on heavy media but in a much shorter time frame and at a fraction of the cost. One such technique was described by Yao and co-workers in 2001 and involves the tryptic digestion of a protein sample in the presence of either H_2O or H_2^{18}O . As the cleavage of a protein is a form of hydrolysis, then peptides/proteins digested in the H_2^{18}O will gain a 4 Da increase in mass (a 2 Da increase per ^{18}O , two of which are incorporated), while those digested in the presence of H_2O will not, enabling their differentiation and relative quantification via MS (Yao et al., 2001), as is shown on Figure 1-10.

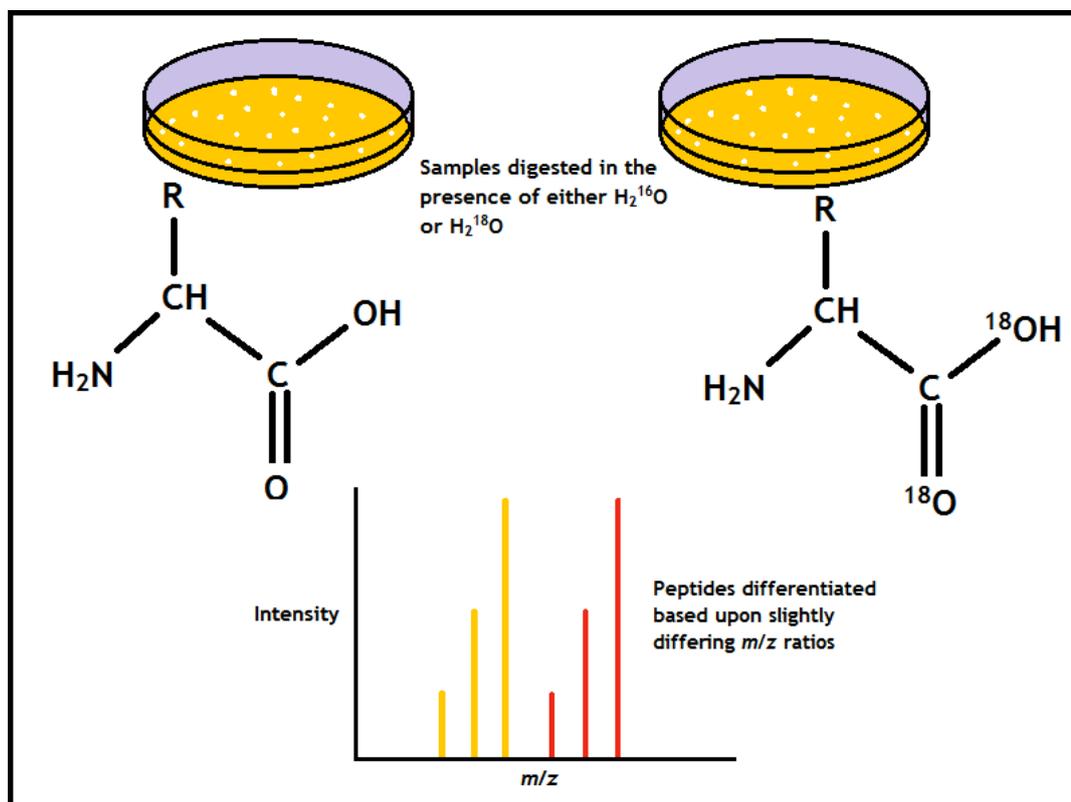


Figure 1-10: The principle behind ^{18}O labelling and comparative protein quantitation. Depicted above is the workflow utilised during an $^{18}\text{O}/^{16}\text{O}$ comparative protein quantitative analysis, including the growth of both organisms on a normal media, their digestion in the presence of either H_2^{16}O or H_2^{18}O , their mixing and their analysis via LC-MS. Because of the 4 Da mass shifts introduced to each peptide, the molecules can be differentiated on the mass spectrum and quantified.

There are many advantages conferred by the use of ^{18}O labelling in comparison to a media based quantitative method, including the fast reaction time, the simplicity of the technique and its low cost. Perhaps one of the biggest advantages is that, in theory, the technique works equally as well in a complex pre-fractionated protein sample as it would with a single purified protein, giving it a much greater appeal in the field of proteomics (Vaudel et al., 2010).

Many disadvantages have however been reported in the application of ^{18}O labelling; the most worrying of which are the description of both incomplete labelling, and the occurrence of back exchange (reintroducing ^{16}O back into a previously labelled sample). Furthermore, as ^{18}O labelling yields only a 4 Da

mass increase, peak separation can be incomplete (Makarov and Scigelova, 2004).

While the ^{18}O labelling reaction is relatively efficient, when a protein is digested, one ^{16}O atom is released into the digestion buffer for every two ^{18}O atoms incorporated, which over time can build up in the digestion buffer reducing the efficiency of the reaction. Upon the completion of labelling, the presence of either residual enzyme activity in water, or acid/base-catalysis can cause the bound ^{18}O to be replaced by an ^{16}O molecule, thereby complicating data analysis through the introduction of overlapping (+2 Da) peaks (Schnölzer et al., 1996).

1.6.4 Isotope-Coded Affinity Tags

A second chemical-based labelling technique designed for the relative quantification of peptides, which yields an increase in mass (between samples) greater than that obtained through the use of ^{18}O labelling, was first described by Gygi and co-workers in 1999, termed isotope-coded affinity tags (ICAT) (Gygi et al., 1999). In brief, this technique features the differential labelling of two protein samples (with either a heavy or light ICAT tag), their co-digestion, avidin based affinity purification, separation and analysis based upon LC-MS/MS (Patton et al., 2002).

The ICAT tags employed during this technique can be seen to be composed of three specific regions; a thiol specific (Cys binding) reactive group, a polyether linker region (containing either eight deuterium (heavy) or hydrogen (light) atoms), and a biotin affinity tag (enabling avidin based affinity chromatography) (Gygi et al., 1999).

This technique offers several advantages over the use of ^{18}O labelling, the first of which is the 8 Da mass difference imparted between the two labelled samples, double that obtained through the use of ^{18}O labelling (Gygi et al., 1999, Yao et al., 2001). This is particularly important during proteomics as it better

enables the differentiation and quantification of multiply charged peptide species, the peaks from which may otherwise become overlapped. Furthermore, as affinity chromatography is utilised during the implementation of ICAT, the final sample should contain approximately 10-fold fewer peptides than it may otherwise have (Ong et al., 2002, Patton et al., 2002), increasing the precision of the quantification through increasing the protein sequence coverages achieved (by ensuring that only those peptides which are labelled will be selected for MS/MS).

There are however several disadvantages to the utilisation of ICAT over ^{18}O labelling, for example, not every protein contains a Cys residue. Indeed it has been documented that more than 5% of human proteins do not contain a Cys residue, a figure which increases when peptides are considered (Karlin et al., 1991). In addition, as the heavy ICAT tag contains deuterium, the RPLC based retention time differs slightly between the heavy and light labelled peptides, introducing a potential source of error as a result of non-uniform peptide ionisation (varying retention times may result in different co-eluting sample components, each competing for ionisation) (Zhang et al., 2001, Patton et al., 2002). Finally, the ICAT tags have been reported to interfere with the fragmentation spectra, complicating data interpretation (Haynes and Yates III, 2000, Patton et al., 2002).

Recently an attempt has been made to improve upon the first generation of ICAT tags, firstly through making the biotin moiety acid-labile, thereby reducing the size of the ICAT tag post-purification (with the aim of reducing MS/MS interference). Furthermore, ^{13}C may be utilised instead of deuterium, which should enable peptide co-elute, enabling a more accurate relative quantitation (Yi et al., 2005).

1.6.5 Tandem Mass Tags

While SILAC, ICAT and ^{18}O labelling, among countless other techniques, rely upon a peak area comparison being made between the parent ions of both a labelled

and an unlabelled peptide, it is worth considering that this form of analysis can in some cases become a hindrance. As was discussed in Section 1.3.2, the proteome of an organism at any given point in time will contain a vast number of proteins, which when subjected to a bottom-up based proteomic analysis, such as SILAC or ^{18}O labelling, will yield an overwhelming number of proteolytic peptides. It is then worth considering that during a comparative proteomic analysis that at least two proteomes will be assessed, and therefore two parent ions will be present for each peptide, not to mention those which may result from the presence of PTMs and indeed the presence of isotopic peaks. It is therefore understandable that when it comes to analysing the data gathered during a so-called MS^1 experiment, that peaks can become merged, preventing an accurate peak area from being obtained (Vaudel et al., 2010).

In an attempt to overcome these problems, while further reducing the number of peptides present in a comparative proteomic analysis, a technique was described by Thompson and co-workers in 2003, termed tandem mass tags (TMT) (Thompson et al., 2003), which, in brief, features the harvesting, purification and digestion of two individual samples, at which point, each is differentially labelled with one of two isobaric tags. Each sample can then be pooled and analysed via LC-MS/MS, utilising an IDA scan, so as to fragment only those MS peaks which match a list of pre-defined selection criteria (Thompson et al., 2003). As the TMT technique features the use of isobaric tags, only a single MS peak should be present for each peptide. A relative quantification is performed based upon the MS/MS spectra obtained, comparing the peak areas detected for each reporter ion (Ross et al., 2004).

The duplex TMT tags utilised in this reaction can be seen to be composed of three distinct regions; a reporter group (with a mass of either 126.1 or 127.1 Da), a reactive group (so as to enable the tag to be bound to a primary amine (both on the N-terminal of each peptide in addition to the side-chain of each Lys residue)), and a balancer group (increasing the mass of each tag to a total of 225 Da) (Thompson et al., 2003, Byers et al., 2009), as is shown on Figure 1-11. More

recently TMT has been extended to enable the relative quantification of up to six biological samples (Dayon et al., 2008).

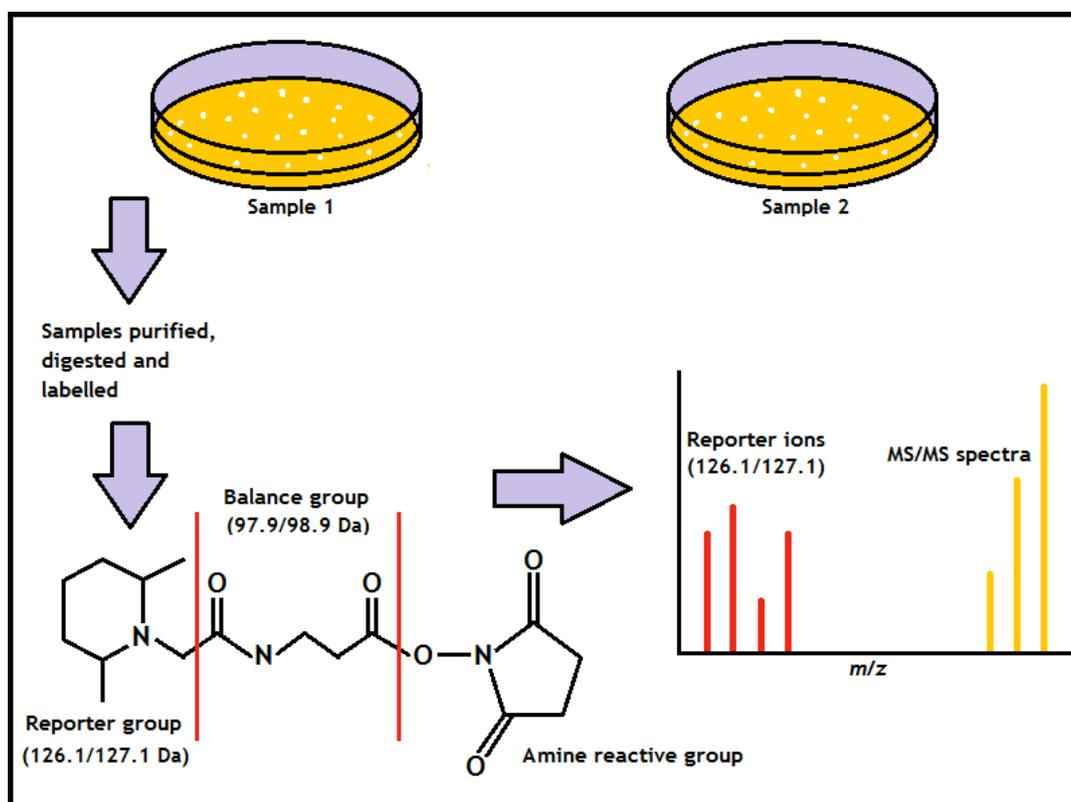


Figure 1-11: The theory behind a duplex TMT based quantitative proteomic analysis. Depicted above are the steps involved in a duplex TMT based quantitative proteomic analysis, including the growth of multiple samples, their labelling, and their detection via LC-MS. The duplex TMT tags feature a reporter group (with a mass of either 126.1 or 127.1 Da) and a balance group (with a mass of either 97.9 or 98.9 Da), which when bound to the amine group of a Lys or N-terminal of a peptide, results in an overall mass increase of 225 Da for each tag.

Following the description of duplex TMT in 2003, a second isobaric tag based quantitative technique was described by Ross and co-workers in 2004, termed isobaric tags for relative and absolute quantitation (iTRAQ), capable of comparing up to four biological samples. More recently this technique has also been extended to enable the labelling and comparison of up to eight unique biological samples; however it has been suggested that the peptide identification rate may be lower when using the eight-plex technique over the four-plex (Pichler et al., 2010).

Like TMT, each of the four iTRAQ tags are composed of three distinct regions; a reporter group (with a varying mass of between 114.1 and 117.1 Da), a reactive group (so as to enable the tag to be bound to a primary amine (both on the N-terminal of each peptide in addition to the side-chain of each Lys residue)), and a varying balancer group (increasing the mass of each tag to a total of 144.1 Da) (Ross et al., 2004), as is shown on Figure 1-12.

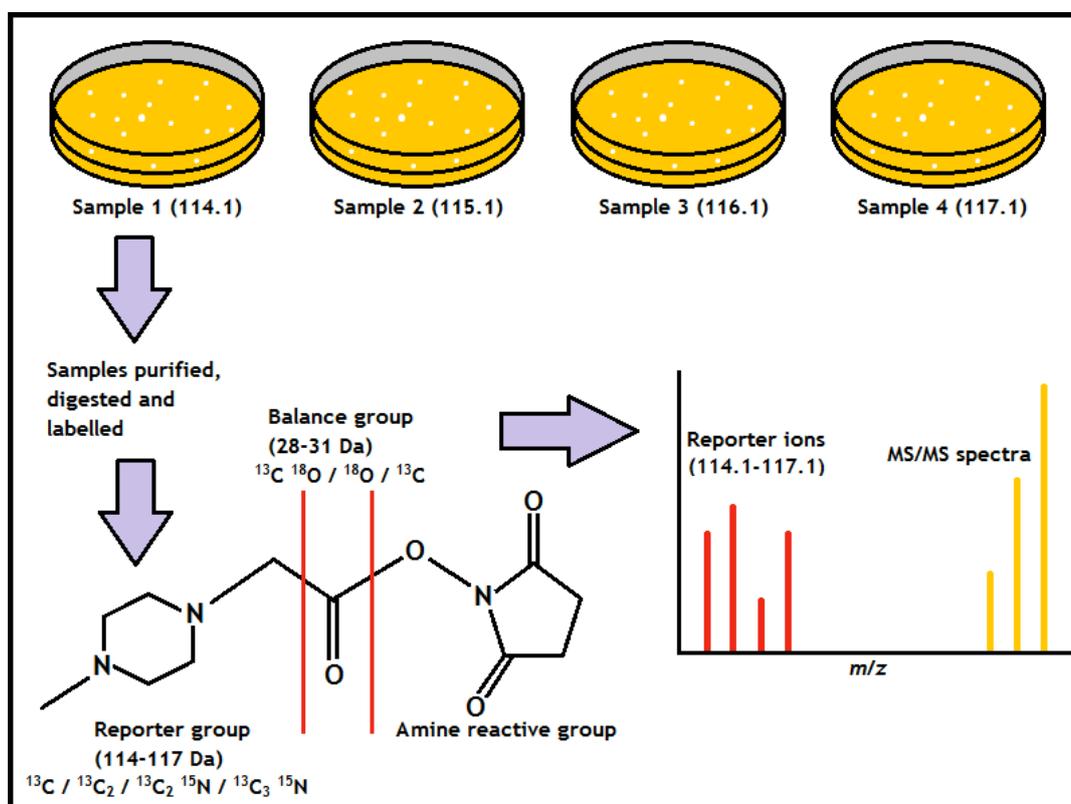


Figure 1-12: The theory behind an iTRAQ based quantitative proteomic analysis. Depicted above are the steps involved in an iTRAQ based quantitative proteomic analysis, including the growth of multiple samples, their labelling, and their detection via LC-MS. The iTRAQ tags feature a reporter group (with a mass ranging from 114-117 Da) and a balance group (with a mass ranging from 28-31 Da), which when bound to the amine group of a Lys or N-terminal of a peptide results in an overall mass increase of 144.1 Da for each tag.

Both TMT and iTRAQ have many advantages over the use of other quantitative proteomic techniques, including their ability to reduce the complexity of multiplexed MS spectra, even those resulting from up to eight unique biological samples (iTRAQ).

While the utilisation of MS/MS enables the co-identification of each peptide (based upon the peptide fragmentation spectrum (discussed in Section 1.4)), the application of these isobaric techniques has several drawbacks. Firstly, a requirement of TMT/iTRAQ is that an MS/MS scan must be performed on each peptide if a quantitation is to be achieved. This can be particularly problematic depending on both the resolution of the HPLC based separation and the cycle time of the MS utilised. As an example, if a short, single dimension LC gradient is utilised in combination with a long cycle time mass spectrometer (which will depend on the age of the instrument and mass range being scanned), then many peptides may be rapidly eluted from the column and only a limited number of MS/MS scans performed. Furthermore, as both iTRAQ and TMT tags are applied to peptides, then it is possible for inconsistencies during cell lysis, fractionation and sample labelling to effect the precision of this technique (Unwin et al., 2006). Finally, in the instance in which multiple peptides with similar m/z ratios co-elute, and thus are all subjected to MS/MS, chimeric spectra (overlapping MS/MS spectra which can complicate peptide identification) can present, summing reporter ion signals from multiple peptides, and thus yielding incorrect quantitation data (Houel et al., 2010).

1.6.6 Label-Free Quantification

Serving as an alternative to both cell-culture based labelling and chemical labelling techniques, several label-free quantitative strategies have also recently been developed with the goal of achieving a faster (requiring no additional sample processing), cleaner and simpler quantification (Zhu et al., 2010).

While an accurate MS-based comparative quantification between samples, instruments, and indeed laboratories is reliant on the use of stable isotopic labelling and the co-separation and analysis of each biomolecule (as was discussed in Section 1.6) (Schulze and Usadel, 2010), it is also true that these techniques can increase both the time and the complexity of sample preparation. They also require the use of high cost reagents and can risk reduced quantitative accuracy through incomplete labelling. Furthermore, aside from

iTRAQ, few quantitative techniques enable the direct comparison of more than two samples, something which is not a limiting factor during label-free analyses (Zhu et al., 2010).

Label-free quantitation can be conducted in one of two ways, the first of which is based on the premise that the observed MS signal intensity correlates well with the quantity of ion injected (S. Pan et al., 2009). This technique is therefore based on the comparison of a peptide's peak area between samples. This technique does however have several drawbacks, including the fact that the ion intensity of a peptide can vary between samples due both to the gradual deterioration of the nano-spray ESI needle and the reduced binding capacity of the LC column. Likewise, where a shift in the LC retention time has occurred, varying co-eluting peaks, and thus competition during ionisation, can lead to a change in the intensity of the ion (Zhu et al., 2010). These variations were foreseen and can be corrected for through the use of one of several suitable software packages (Shaw et al., 2003).

In comparison to the quantitative accuracies of both ^{15}N and SILAC, each of which regularly achieve coefficient of variation (CoV) values of less than 10%; and iTRAQ, which is described as being of a "medium" quantitative accuracy, with CoV values regularly between 10 and 30%, label-free quantification, based on the comparison of peptide ion intensities between samples is also described as being of a "medium" accuracy with CoV values regularly between 10 and 30% (Shaw et al., 2003). In brief, the CoV can be calculated by dividing the standard deviation by the mean. This is particularly useful in quantitation as it is a dimensionless measure of variation, comparable between techniques, laboratories and indeed data sets.

The second label-free approach, termed spectral counting, is based on the premise that when the amount of a protein entering the MS increases, so too should the sequence coverage, and therefore the number of proteolytic peptides identified. The amount of protein within a sample is thus quantified through comparing the number of MS/MS spectra identified for each protein, offering an

advantage over the measurement of peak area alone, which is suited only to the comparison of a single peptide between samples (Zhu et al., 2010). Indeed, this form of label-free quantitation has also been applied to absolute quantitation using a technique termed protein abundance index, which can be calculated by dividing the total number of peptides identified from a protein by the number of peptides which can theoretically be identified (Schulze and Usadel, 2010, Ishihama et al., 2005).

In applying spectral counting to the quantification of two ideal samples (samples which possess less than twofold variation in abundance and which contain at least four peptide peaks per protein), this label-free technique has been shown to display an accuracy greater than 95% (Schulze and Usadel, 2010). In contrast, when the same technique is applied to the quantification of lower abundance proteins, spectral counting has been found to provide only a “poor” level of quantitative accuracy, as a result of the reduced number of spectra available (Old et al., 2005), and will generally yield CoVs greater than 30% (Kondo and Hirohashi, 2007).

1.6.7 Labelled Internal Standard Based Quantification

While each of the previously discussed quantitative techniques has advantages and disadvantages when applied to a specific biological system, each can provide only a relative level of quantitation for peptides, between two or more samples. While useful for assessing the varying concentration at which a protein is expressed between a range of cellular states (generally expressed as a ratio), further interpretation of this data is difficult (Brownridge et al., 2011). There is growing demand for the absolute quantification of proteins within biological systems, where the exact number of protein molecules per cell is quantified, enabling both the quantification of samples where no reference material exists and the tracking of a disease’s progression (Drews, 2000, Brownridge et al., 2011).

The idea of absolutely quantifying a molecule is not a new concept, indeed the absolute quantification of drug molecules has been performed since 1932, based upon isotope dilution analysis; with a more recent adaption utilising stable isotopes, in a technique termed stable-isotope dilution (Fassett, 1995). In brief, as the stable isotope has an identical structure to the analyte, it should behave in an identical manner when subjected to LC-MS/MS. The stable isotope should therefore co-elute with the analyte, and as such, be subjected to same matrix based ion suppression or enhancement, thereby enabling the peak areas of the two molecules to be directly compared.

1.6.7.1 Protein Standard Absolute Quantification

The production of isotopically labelled full-length proteins to act as standards for absolute quantification was first described by Brun and co-workers in 2007 and termed protein standard absolute quantification (PSAQ) (Brun et al., 2007). PSAQ consists of a protein analogue being synthesised in the presence of uniformly labelled Lys and Arg residues (^{13}C and ^{15}N), which when cleaved with trypsin, will yield proteolytic peptides each containing one labelled amino acid residue. These peptides can then be compared with their non-labelled counterparts, and assuming the amount of the PSAQ-labelled protein added is known, an absolute concentration can be determined for each target protein (Brun et al., 2007), as is shown on Figure 1-13.

The main advantage of PSAQ, over AQUA (described in Section 1.6.7.2) and quantitative concatamer (QconCAT) (described in Section 1.6.7.3), is the fact that an intact protein standard is utilised as an internal standard, and thus both the protein target and the internal standard should separate and digest under identical conditions, accounting for any potential sample loss.

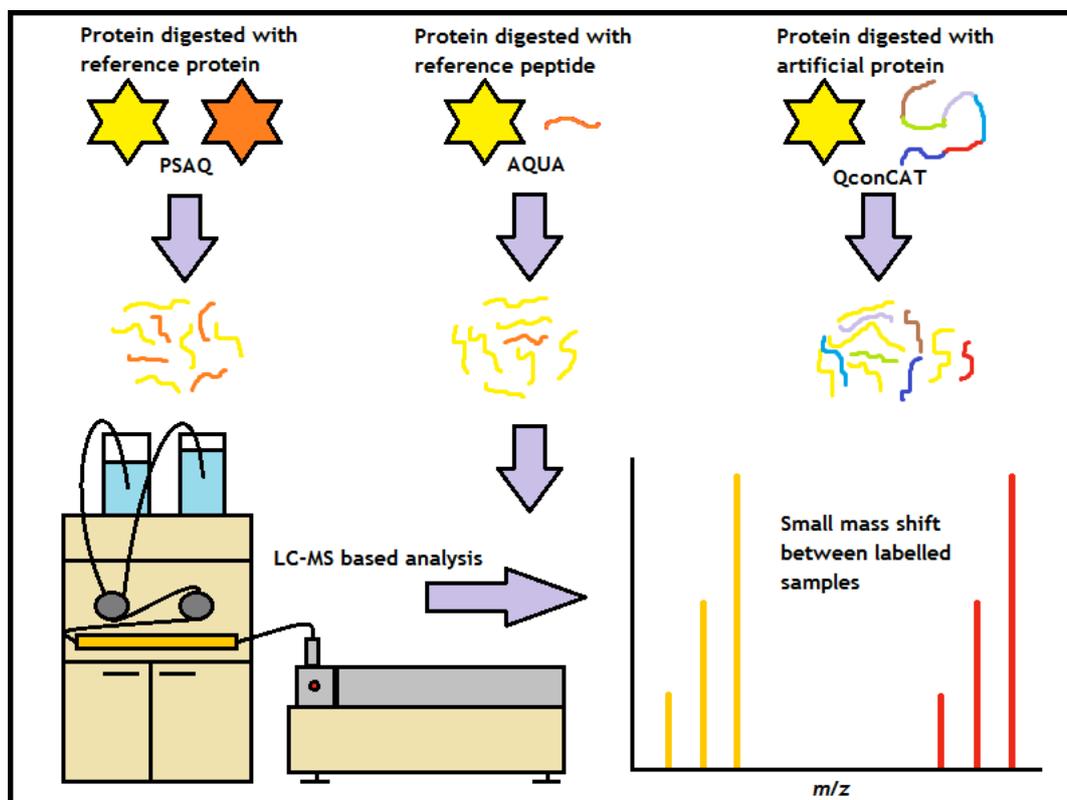


Figure 1-13: The techniques through which PSAQ, AQUA and QconCAT enable an absolute quantitative proteomic analysis. While PSAQ enables an absolute quantification to be performed based on the digestion of the target protein with a labelled intact protein standard, AQUA enables a quantification to be performed based upon the quantification of a single proteotypic target peptide. While QconCAT is performed in much the same way as AQUA, it enables the quantification of multiple proteins per sample through the construction, and subsequent co-digestion, of an artificial protein.

While very little data has been published detailing the application of PSAQ in a biological system, sensitivity levels down to 44 femtomoles (fmol) (1.2 ng of protein in 100 microlitres (μL) of sample) have been described in human urine (Brun et al., 2007), while a more recent publication from the same group has detailed the technique to yield CoV values of between 0.6 and 30.7%, with non-proteotypic peptides (the term proteotypic is discussed further in Section 1.6.7.1.1) containing greater levels of variation (Jaquinod et al., 2012).

PSAQ is, however, a time consuming and complicated technique, requiring the continued synthesis of a standard isotope-labelled protein via cell culture in labelled media, its purification, and finally its quantification by amino acid

analysis (AAA) (this point also applies to QconCAT however, while AQUA peptides are pre-quantified, enabling immediate use), prior to the protein being mixed with the target cell lysate and processed (Brun et al., 2007). Furthermore, this technique assumes that both the synthetic and target proteins will yield the same PTMs, as is required for the biomolecules to be separated and analysed under identical conditions (Brownridge et al., 2011).

1.6.7.1.1 Proteotypic and Quantotypic Peptides

In the field of quantitative proteomics, the terms proteotypic and quantotypic are used to indicate that the sequence of a peptide is ideal, and therefore the quantitation achieved for that cleaved peptide should be representative of the quantity of the parent protein present before digestion.

On the topic of proteotypic peptides, Craig *et al.* state that proteotypic peptides should be compatible with all common proteomic sample handling steps, including enzymatic digestion, gel extraction, RP-LC and ionisation (Craig et al., 2005). Bislev *et al.* further define the term proteotypic, stating that proteotypic peptides should have a length of between 7-20 amino acids and should be unique within their target proteome (with a sequence which can be mapped to single genomic location). Furthermore, Bislev *et al.* state that proteotypic peptides should not contain amino acid residues which are prone to modification (including Cys, Met, Try, Ser and N-terminal Gln residues), should not contain N-terminal acidic residues (assuming the peptide is cleaved at a basic residue) and should not be cleaved from a sequence susceptible to missed cleavages (sites which are flanked by additional cleavage sites) (Bislev et al., 2012).

1.6.7.2 AQUA Based Peptide Quantification

During a bottom-up PSAQ based quantitative analysis of a protein, only a few proteotypic peptides will ultimately be quantified, so as to avoid any erroneous data which may be obtained from the quantitation of peptides not suited to spiked digest based analyses, such as those peptides which are susceptible to

missed cleavages (PSAQ should, however, still correct for these sequence flaws during quantitation). A quicker and cheaper approach may therefore be to produce synthetic labelled copies of those peptides which are eventually quantified, and to spike those into the sample during the proteolytic digest. This technique, termed AQUA, was first described by Gerber and co-workers in 2003 (Gerber et al., 2003b) and is based upon the premise that for every protein molecule digested, a single copy of each proteolytic peptide will be generated, thus, through quantifying even a single peptide it should be possible to determine the amount of parent molecule present. AQUA peptides are now commercially available from both Sigma and Thermo, increasing their availability and application in the wider scientific community.

During a typical AQUA experiment, a known amount of AQUA peptide (a synthetic internal standard peptide containing a uniformly ^{13}C and ^{15}N labelled amino acid residue) will be combined with the targeted protein, generally after 1DGE based separation, but before the protease has been added to the gel pieces (Kirkpatrick et al., 2005a). In adding the AQUA peptide solution to the dehydrated gel pieces, the AQUA peptide is absorbed into the gel, ensuring an equal extraction efficiency for both the target and synthetic peptides post-digestion (Gerber et al., 2003b). Following the completion of the proteolytic digest, the peptide extracts are pooled and dried, re-suspended in a polar solution, separated based upon RP-LC and quantified through the application of MRM based MS (first performing a relative quantification between the synthetic and target peptides, then through knowing the exact amount of internal standard added to the sample, calculating the absolute amount of target peptide within the sample analysed) (Kirkpatrick et al., 2005b).

This technique has several advantages over the use of PSAQ, the most compelling of which is the fact that AQUA uses a labelled peptide instead of a full-length labelled protein. As peptides can be rapidly synthesised based upon solid-phase peptide synthesis (SPPS), this technique is both more rapid, and cheaper than PSAQ, therefore it is perhaps more applicable to large scale sample analysis (Brownridge et al., 2011). Furthermore, as AQUA peptides can

be synthesised so as to containing a range of PTMs, it thus enables the quantification of both modified and unmodified protein species (V. Mayya et al., 2006).

There are, however, several disadvantages to selecting AQUA over PSAQ, the most prominent of which is the assumption that during an AQUA based quantitation, the digest efficiency of the targeted protein will be complete (discussed further in Section 2.1), while PSAQ is capable of correcting for any inefficiency. In addition, another advantage of PSAQ is the fact that the intact labelled protein is added to the sample at an early stage, and thus can correct for any sample loss encountered during separation (Brun et al., 2007). During the storage of the AQUA peptide, and indeed during sample preparation, peptide loss through binding can also pose a problem, where it has been documented that up to 90% of the peptide may be lost over a period of 24 hours through the adsorption of the peptide to the sample vial (Zhang et al., 2010).

1.6.7.3 QconCAT based Peptide Quantification

A third stable isotope based quantitative technique also exists, offering an alternative to both AQUA and PSAQ, designed for the quantification of multiple (up to 100) peptides within a proteome. This technique, termed QconCAT was first described by Beynon and co-workers in 2005 (Beynon et al., 2005) and involves the synthesis of a stable isotopically labelled, artificial protein, composed of multiple peptide sequences for the absolute quantification of multiple proteins within a biological system. To produce this artificial protein, a *de novo* gene is designed *in silico* and synthesised commercially, the amplified vector is transfected into *Escherichia coli* and the protein expressed. Upon the protein having been purified, and quantified via AAA, a quantity of the artificial protein is added to the target proteome and co-digested, separated based on RP-LC and quantified through the application of MRM based mass spectrometry (first performing a relative quantification between the synthetic and target peptide peak areas, then, through knowing the amount of internal standard

added to the sample, calculating the absolute amount of target peptide, and thus parent protein, within the sample) (Beynon et al., 2005).

QconCAT has several advantages over the application of AQUA. Firstly, the QconCAT derived peptides (Qpeptides) are the product of a proteolytic digest, removing much of the bias associated with incomplete digestion. Furthermore if more than ten proteins are to be quantified (Beynon et al., 2005), then it is more economical to utilise QconCAT rather than synthesise and purify multiple AQUA peptides. QconCAT also has an advantage over PSAQ in terms of the time required to prepare multiple standards, where upon designing the QconCAT gene, only one protein is to be expressed.

QconCAT does however have several disadvantages over the application of PSAQ. Firstly, while QconCAT does involve a co-digestion step, accounting for much of the variation associated with incomplete digestion, the protease employed will bind to several amino acids both up and downstream of the cleavage site, and thus the digestion of both the target and artificial proteins may proceed with different digest efficiencies (Brownridge and Beynon, 2011, Beynon et al., 2005). Furthermore, for each protease utilised, a separate QconCAT may be required, increasing the cost and complexity of the method. Finally, as the Qpeptides are expressed through the use of cell culture, rather than chemical synthesised, they cannot be imparted with the same range of PTMs as an AQUA based analysis.

1.7 Protein Isoforms

When analysing the human genome it was noted that only 1.1% of the deoxyribonucleic acid (DNA) sequenced accounted for protein coding exons, while 24% was found to consist of intronic sequence and some 75% consisted of intergenic DNA, with no known function in the coding of either RNA or protein. This was particularly surprising as it indicated humans to have only two to three times as many genes as lower organisms such as the mustard plant or fruit fly,

indicating that functional complexity, rather than absolute gene number, was key in the complexity of the human proteome (Baak et al., 2005).

RNA splicing has been documented to be ubiquitous in eukaryotes, enabling the production of a variety of related but distinct protein isoforms through the differential incorporation of a range of exons in the mature mRNA (Andreadis et al., 1987). It has been predicted that up to 75% of the 20,000 to 30,000 genes within the human genome may be subject to some form of alternative splicing (Johnson et al., 2003), yielding an average of two or three transcripts per gene (M. Nakao et al., 2005). Furthermore, it has since been estimated that up to 95% of human multi-exon genes may be subject to alternative splicing, making alternative splicing the rule, not the exception (Pan et al., 2008a). It should, however, be noted that while some of these sequences yield novel protein isoforms, many mRNAs, even if detectable at the transcriptional level, do not encode a functional protein (Leoni et al., 2011, Tress et al., 2007).

In brief, transcribed RNA consists of both exons and introns, with the exons containing the translated protein coding sequence and the introns containing both regions of regulatory importance and intergenic sequence. Upon the removal of the introns, the exons can be ligated in different combinations, giving rise to various protein isoforms.

These introns can be identified through locating the conserved 3' and 5' intronic splice sites, where the 5' splice site is marked by conserved sequence MAG|GTRAGT (where M represents either an A or a C and R represents either an A or a G), and the 3' splice site is marked by conserved sequence CAG|G. In addition, each intron should contain a splicing branch point, usually marked by a single adenosine residue, some 18-40 bases downstream of the 5' intronic splice site (Reed, 1989).

Where an alternative protein product is produced, the modified sequence of the polypeptide chain can affect the intra- and extracellular location of the protein, its regulation, and its efficiency in binding a specific substrate (Gunning, 2001).

Indeed alternative splicing can also result in the truncation of a protein, synthesising a protein isoform which may contain a binding domain but not a catalytic domain, resulting in a protein isoform with a function differing to that of the wild-type (Rauch et al., 2011).

In addition to the application of alternative splicing in increasing the complexity of the proteome, alternative splicing has also been linked to defective mRNA splicing in diseased cells. When the synthesis of a protein isoform is linked to a disease such as cancer, the most common cause of alternative splicing is a mutation of the intronic splice site, usually leading to the exclusion of the adjacent exon, and in more than half of all cases a truncation (Venables, 2004).

1.7.1 Quantifying Protein Isoforms

A major challenge facing the field of proteomics is the highly variable concentration at which a range of proteins can be expressed, where often it is those proteins which are present at a low copy number which are the most biologically interesting, yet can be masked by higher abundance proteins (Anderson and Anderson, 2002, Klose, 1975, O'Farrell, 1975).

The analysis of protein isoforms is therefore particularly challenging due both to the high degree of sequence homology shared by many of these isoforms, and their low copy numbers. While these proteins differ structurally, this variation may have only a limited effect on the pI and MW of the final protein product, making their differentiation based upon the use of 2DGE next to impossible (S. P. Gygi et al., 2000, Galeva and Altermann, 2002).

In regards to the application of top-down based proteomics for the detection of protein isoforms, long acquisition times are required to generate the MS/MS spectra necessary to determine the structure of a protein, based upon the use of ECD. Off-line HPLC based separative techniques are therefore frequently employed in top-down based proteomics, so as to provide the levels of protein separation required. Furthermore, due to the range of ion species (charge

states) which may be generated from an intact protein, sensitivity has also proven to be an issue during the application of top-down based proteomics. This technique, therefore, is not best suited for the analysis of protein isoforms within complex cell lysates.

Furthermore, when a bottom-up based proteomic technique is applied to the analysis of a protein isoform, the protein inference problem may be encountered. In brief, when a protein is proteolytically digested, the resulting peptides may have sequences identical to those of peptides cleaved from multiple proteins, leading to ambiguities in the determination of the proteins identity (Nesvizhskii and Aebersold, 2005). If a protein isoform is to be identified based upon the application of bottom-up based proteomics, therefore, only those peptides which have a sequence unique within the target proteome should be targeted. If the resolution of the HPLC based separation is not adequate, however, the isoform specific peptides of interest may co-elute with other proteolytic peptides. When this occurs, and assuming each of the peptides meets the IDA selection criteria, the most intense ions will be selected for MS/MS. As such, the isoform specific peptide may become masked from detection based.

Based upon the targeting of only those proteolytic peptides which are unique to an isoform within the host proteome, an absolute quantification may also be performed. As AQUA, QconCAT and PSAQ are all capable of yielding labelled proteolytic peptides, the peak areas of both the labelled and un-labelled peptides may be compared, prior to the absolute amount of target protein being determined. Furthermore, as only two peptides from the complex digested cell lysate are of interest to the analyst, MRM based scanning may be utilised to reduce the complexity of the MS data obtained, while increasing the sensitivity of the MS towards those analytes which are targeted.

Several labelled peptide based quantitative analyses of protein isoforms have previously been described, including the quantification of three of five known sucrose synthase isoforms in the root nodules of *Medicago truncatula*, a technique designed to improve upon an existing western-blot based semi-

quantitative technique (Wienkoop et al., 2008). Likewise, a technique was designed to quantify several members of the Cytochrome P450 superfamily in human liver microsomes, enabling the specific quantification of isoforms CYP3A4 and CYP3A5, and the quantification of the entire CYP3A family, but not isoform CYP3A43 (Wang et al., 2008b).

1.8 Conclusion

While gel electrophoresis was first utilised for the separation of proteins back in 1959 (Kwietny et al., 1959), its use in comparative proteomics was delayed until the description of DIGE in 1997 (M. Unlu et al., 1997). Likewise, while mass spectrometry was first described in the late 1800s (Griffiths, 2008), it was not possible to ionise large intact biomolecules until the late 1980s (Falk et al., 2006), at which point a whole range of new hybrid instruments had to be developed to enable the analysis of the proteome (El-Aneed et al., 2009). The field of proteomics can thus be seen as a relatively new area of scientific research, which having been applied to the qualitative analysis of many proteins in a range of biological systems, is now becoming increasingly quantitative (Baak et al., 2005).

While new quantitative proteomic techniques are still regularly being described, the field of comparative proteomics has recently been likened to a minefield (Vaudel et al., 2010), with every advantage offered by one technique being offset by several disadvantages over another. Some of the most desirable traits in quantitative proteomics are thus the ability to multiplex (iTRAQ/TMT), which is offset by reagent costs and the inaccuracy of processing each sample separately. Additionally, the desire to combine samples early in sample preparation (^{15}N or SILAC), is offset by the cost of the labelled media, and the fact that only a limited number of cell lines are suitable for growth on dialyzed serum.

When an absolute quantification is to be performed, as appose to a relative quantitation (such as that performed by each of those techniques previously

mentioned), very few techniques exist. While AQUA offers a relatively cheap and sensitive technique for the analysis of a single peptide, the precision of the data achieved may be compromised by the efficiency of the proteolytic digest. In comparison, while both QconCAT and PSAQ can correct for any inefficiency encountered during proteolytic digestion, each requires the synthesis of a labelled protein, and subsequently its quantification prior to use, which may be both expensive and time consuming, perhaps not applicable to large scale sample analysis.

Finally, when a protein isoform is to be quantified, the high sequence homology means that a proteolytic peptide must be selected which is unique within the host proteome. As relatively few sample components are targeted during such an analysis, MRM based scanning may be utilised, reducing the complexity of the MS data obtained, while increasing the sensitivity of the MS towards the targeted analyte.

1.9 Aims

While several stable isotopic labelling based techniques have been reported which are capable of quantifying multiple protein isoforms, to the author's knowledge no technique has yet been described which enables the absolute quantification of every individual isoform within a particular family of protein isoforms. Indeed, this may be of particular interest in assessing how the ratio of a group of protein isoforms varies between a range of tissues or in a diseased state.

Furthermore, while many publications detail the sequential and physicochemical properties desirable in a proteotypic peptide, none have explored the strategies which may be employed to enable the selection of a suitable AQUA peptide for the quantification of multiple protein isoforms, where the high sequence homology of these isoforms can drastically limit the number of peptides available for selection. The use of several proteases per analysis and different enzyme based proteolytic double digests may be required, in addition to the

selection of peptides which contain reactive amino acids, so as to enable the selection of a single peptide.

In the light of this, the author aims to develop a range of techniques, which may be utilised for the selection of proteolytic peptides suitable for the absolute quantification of multiple protein isoforms. Indeed, in doing so, the author aims to explore the suitability of a range of proteases and peptides containing reactive amino acids so as to determine what measurable effects these non-ideal components have on the quantitative accuracy of the technique.

Ultimately, when fully established and optimised, this type of analysis should enable the wider scientific community to: (i) Identify which isoforms of a particular protein family are expressed within a given tissue (through detecting isoform specific target peptides). (ii) Accurately quantify each isoform detected (through detecting both the isoform specific target and synthetic AQUA peptide), and (iii) Detect changes in the expression ratio of these isoforms under different physiological conditions.

In order to achieve these ultimate technical aims, which are beyond the scope of a PhD project, the author intends to: (i) Identify several families which consist of multiple protein isoforms and screen each with the aim of identifying suitable AQUA peptides for the absolute quantification of each protein isoform. (ii) Characterise each AQUA peptide and intrinsic target peptide so as to develop a suitable triple-quadrupole based MRM acquisition method for the identification of each peptide within a single LC-MS run. (iii) Express suitable amounts of each protein isoform, enabling the proteolytic digests to be optimised using a range of commercially available proteases. (iv) Assess the accuracy and reproducibility of the AQUA technique using both a suitable number of replicates, and (v) Test these optimised AQUA based quantitative techniques on both non-immunoprecipitated transfected protein species and endogenously expressed samples.

2 AQUA Peptide Selection

2.1 General AQUA Peptide Selection Criteria

When selecting an AQUA peptide for the quantification of a target protein, there are several physical, chemical and sequence specific factors which must be considered. For example, based upon both manufacturing constraints, and the requirement for a peptide which is suited to an analysis based upon LC-MS/MS, the peptide must; (i) Be between 6 and 15 amino acids in length; (ii) Contain a residue suitable for stable isotopic labelling (either Ala, Arg, Ile, Leu, Lys, Phe, Pro or Val), and; (iii) Be water soluble, being neither too hydrophilic (ideally being composed of at least 10% hydrophobic residues (including Ile, Leu, Val, Phe, Trp and Met) nor hydrophobic (ideally being composed of less than 50% hydrophobic residues) (Brun et al., 2009, Brun et al., 2007, Kettenbach et al., 2011, Kirkpatrick et al., 2005a, Pratt et al., 2006, Pan et al., 2008b).

In addition to the above list of requirements, Kirkpatrick *et al.* make several further recommendations, stating that a candidate peptide should be selected which; (i) Resolves well via high pressure LC (HPLC); (ii) Ionises well via ESI, and; (iii) Is devoid of any chemically reactive residues, with the term chemically reactive covering Trp, Met and Cys (each of which oxidises easily), in addition to the more sequence specific examples of having an Asp residue flanking a Gly or where a peptide possesses an N-terminal Gln or Asn residue (due to the Fluorenylmethyloxycarbonyl (Fmoc) based chemistry used during peptide synthesis (Gerber et al., 2007)) (Kaufmann et al., 1994, Pan et al., 2008b).

In addition, Kettenbach *et al.* suggested that two or three peptides should be selected per protein, and that peptides generated through the cleavage of the target protein at sites susceptible to missed cleavages should be avoided (Kettenbach et al., 2011). This can be further broken down into two groups of digest site: (i) Those where the cleavage site is flanked by “ragged ends” (a cleavage site which is pre- or proceeded to either the C- or N-terminal by a similar cleavage site), and; (ii) Those where the digest site is flanked by an

amino acid carrying an opposite charge (such as a tryptic digest site (Arg or Lys) being flanked to either the C- or N-terminal by an Asp or Glu residue).

To better explain the problem with flanking oppositely charged residues, when an acidic residue is located next to a basic residue, a salt bridge can form between the β -carboxyl group on the acidic residue and the ϵ -amino group on the basic residue. This is particularly problematic when trypsin is utilised, as trypsin binds to either Arg or Lys, cleaving via the formation of a salt bridge with Asp 189 (Siepen et al., 2006). Further to the above, flanking basic residues are even more problematic when trypsin is utilised. Proteases have been shown to recognise not only a single amino acid, but to bind to approximately three residues both up and downstream of the digest site (Zabłotna et al., 2004, Brownridge and Beynon, 2011). Indeed, where two basic residues are located within close proximity, only one will be cleaved, while the other cleavage is missed, which when performing an AQUA based quantification would be disastrous (Brownridge and Beynon, 2011).

In addition to the advice offered by Kirkpatrick *et al.*, Gerber *et al.* suggests that AQUA peptides should be selected based upon the results obtained from previous LC-MS experiments, reasoning that any data obtained from a previous analysis will show which peptides present with a high signal intensity, and as such, which peptides separate well via HPLC and ionise well via ESI (Gerber et al., 2007).

2.2 AQUA Peptide Selection for the Quantification of a Protein Isoform

When screening the proteolytic digest products of a protein with the aim of identifying a proteotypic peptide suitable for an AQUA based quantitative analysis, the number of candidate peptides is often limited (Brownridge et al., 2011). Furthermore, if a PTM on a specific amino acid is to be quantified, just a single peptide per protease may be selected (Gerber et al., 2003b). Indeed this can also be the case when highly homologous protein isoforms are considered,

which may differ by as little as a single amino acid, or have identical sequences, but contain a truncation, making the selection of C-terminal peptides the only viable solution.

A technique to increase the range of peptides available for selection, and thus increasing the chances of a suitable AQUA peptide being identified, may therefore be to review the proteolytic digest products of several proteases. In some cases, however, even when multiple proteases are utilised, it may not be possible to identify a peptide which fully conforms to the AQUA peptide selection criteria. In these situations it may be necessary to select a peptide which is not considered proteotypic, including those which contain chemically reactive amino acid residues, or those which are flanked by non-ideal residues (oppositely charged).

2.2.1 Protease Selection

In bottom-up based proteomics, trypsin is by far the most widely used and documented proteolytic enzyme (Olsen et al., 2004). This is a result of both the wide availability of this high quality sequencing grade enzyme, and its digest specificity, cleaving after Arg and Lys residues, except when preceded by Pro. Furthermore, the sites at which trypsin cleaves account for approximately 10% of the residues identified within an average mix of proteins (Tsuji et al., 2010), yielding tryptic peptides which are rarely larger than 10 to 15 amino acids in length, with masses of between 1 and 3 kDa. Additionally these peptides generally exist in a doubly or triply charged state during LC-MS, and thus are suitable for detection on a range of mass spectrometers (Brownridge and Beynon, 2011).

There may be cases, however, where trypsin will not yield a suitable proteotypic peptide. In these cases, an alternative protease must be utilised. While many exist, which is utilised will depend on which enzyme yields a suitable proteotypic peptide, and which is the most applicable to a quantitative LC-MS based analysis.

On the topic of protease selection, Kettenbach *et al.* suggests that chymotrypsin, Glu-C or Lys-C may be utilised (Kettenbach *et al.*, 2011). While in theory chymotrypsin cleaves at Tyr, Trp, and Phe, general amide bond cleavage has also been documented at a slower rate. If indeed this is the case, then these non-specific cleavages may lead to missed cleavages at the intended digest site, making chymotrypsin unsuitable for AQUA based absolute quantification.

In comparison, both Lys-C and Arg-C have been documented to cleave specifically at the C-terminal of either Lys or Arg, respectively. Due to the lower frequency at which both of these enzymes cleave, however, (with both Arg and Lys each accounting for only 5% of the residues identified within a random protein mix (Tsuji *et al.*, 2010)), larger proteolytic peptides have been reported. These sometimes with m/z ratios unsuitable for detection on modern triple quadrupole based mass spectrometers (Kelleher *et al.*, 1999).

Another protease suggested by Kettenbach *et al.* is Glu-C, which, like trypsin, cleaves at two amino acid residues, Asp and Glu, both of which are present at a high frequency within an average mix of proteins (Tsuji *et al.*, 2010). It should be noted, however, that the digest specificity of Glu-C is dependent on the buffer in which the digest is conducted, with Asp based cleavages occurring at a 3,000 fold lower rate than Glu based cleavages, when conducted in a phosphate based buffer, while Glu-C is thought to cleave exclusively at Glu in an ammonium bicarbonate (AMBIC) based buffer (Houmard and Drapeau, 1972). When applying Glu-C to the absolute quantitation of a target protein, it is thus best to avoid Asp cleaved proteolytic peptides entirely, due to the low digest efficiency of Glu-C at Asp. Furthermore, as Glu-C cleaves specifically after acidic residues, the C-terminal of Glu-C cleaved proteolytic peptides is generally devoid of a C-terminal charge. This is of particular importance during ESI based mass spectrometry as it may limit the detectable fragmentation spectra exclusively to B-ions (as the only charge will be present on the peptides N-terminal), reducing the charge state of the peptide. Where a peptide's charge state is reduced, the m/z ratio will increase, affecting the ability of some mass spectrometers to detect the peptide.

In conclusion, while several proteases may serve as an alternative to trypsin, in cases where a suitable tryptic peptide cannot be identified, each has several disadvantages over the use of trypsin. Indeed the specificity, digest frequency and C-terminal charge imparted by trypsin makes it ideal for proteomics, justifying its widespread usage.

2.2.2 Proteolytic Digest Optimisation

Despite how important the process of complete proteolytic digestion is to the field of quantitative proteomics, no standardised in-solution or in-gel based digest techniques have as yet been described. Indeed most modern in-gel based proteolytic digests still follow the same basic workflows first documented by Shevchenko *et al.* back in 1996, over a decade before the description of absolute quantification based proteomics (Shevchenko *et al.*, 1996).

One of the factors most commonly optimised to achieve a complete proteolytic digestion is the enzyme:substrate ratio at which a proteolytic digest is performed. This generally ranges from 1:10 to 1:100 (enzyme:substrate, weight:weight (w/w)), as per the manufacturer's instructions (Havliš and Shevchenko, 2004, Porter *et al.*, Norrgran *et al.*, 2009). Proteolytic double digests have also been documented, adding an equal volume of the protease after eight hours and digesting for a further ten hours (Mayya *et al.*, 2006).

In-solution proteolytic digests have also been performed in buffers containing a high levels of organic solvent, where the organic solvent is thought to expose the hydrophobic core of the protein, aiding in complete digestion (Hervey *et al.*, 2007). Digests containing levels of organic solvent as high as 80% (v/v) have been reported, apparently yielding faster digestions and higher sequence coverages (Wall *et al.*, 2011). Conflicting reports, however, have suggested that proteolytic activity is reduced above 50% (v/v) organic solvent (Khmelnitsky *et al.*, 1991), and that protein precipitation can occur in solutions containing 80% organic solvent (v/v) (Polson *et al.*, 2003). A recent publication aiming to address the phenomenon, concluded that while an increase in sequence coverage may

present, the overall efficiency of the digest is greatly reduced (Wall et al., 2011).

An alternative strategy, based upon the same premise, employs the use of chaotropic agents to denature the hydrophobic core of the protein, and as such to increase the efficiency of the proteolytic digest. These agents surpass organic solvents in that they are suitable for use with mass spectrometry and have no effect on the activity of the protease (GORDON and JENCKS, 1963). Of these chaotropic agents, the most commonly documented is urea, capable of denaturing the hydrogen bonds within a protein. However, when urea is heated it has been reported to yield isocyanic acid, carbamylating free amines within the target protein, and thus blocking Lys based cleavages (Rajagopalan et al., 1961).

A more recent technique utilised to increase the digest efficiency of a protein has been in the application of acid labile surfactants (ALS) (Norrgran et al., 2009). While ALS were originally employed as a replacement for SDS during in-solution digests, it was noted that they considerably increased the efficiency of the digest (Siepen et al., 2006). While ALS are not directly compatible with mass spectrometry, they are easily degraded through acidification, a process which is already performed during in-gel digestion as a method to increase the extraction efficiency of the proteolytic peptides from the gel pieces. A recent investigation into the use of various digest additives found ALS to yield the highest average increase in digest efficiency, with other experimental parameters yielding little further benefit (Yu et al., 2003, Brownridge and Beynon, 2011).

An interesting technique through which the digest efficiency of a protein can be assessed was recently described by Norrgran *et al.* who utilised several proteotypic (optimal peptides for use in quantification) AQUA peptides (as discussed in Section 1.6.7.2) to monitor the increasing quantities of multiple proteolytic peptides cleaved from ricin over the course of a digest. This technique was applied to digests performed in the presence of ALS, in the presence or 20% MeCN in AMBIC (v/v), in AMBIC alone, and in AMBIC with the

ricin having previously been reduced and alkylated. The results from this study suggested that the ALS based technique yielded peptide levels between 36 and 73% higher than each of the other techniques (Norrgran et al., 2009).

2.2.3 Reactive Amino Acid-Containing AQUA Peptides

As this project involves the quantification of protein isoforms, which may be highly homologous, it may be necessary to develop methods to cope with the use of those peptides which contain reactive amino acids, so as to enable the selection of an AQUA peptide suitable for an absolute quantification.

Within the field of proteomics, Cys residues are commonly alkylated through the use of iodoacetamide (IAA), so as to prevent protein folding through disulfide bond formation or oxidation during sample preparation. Both Kirkpatrick *et al.* and Kettenbach *et al.* recommend that Cys residues within AQUA peptides are reduced with dithiothreitol (DTT) and alkylated with IAA, so as to block the sulfhydryl groups (Kettenbach et al., 2011, Kirkpatrick et al., 2005a). Through the use of IAA, it should therefore be possible to utilise Cys containing AQUA peptides and enable a greater range of proteolytic peptides to be selected.

Unlike the relatively simple process of alkylating a Cys residue, no single solution exists for the complete modification of Met. Should complete oxidation be attempted then it is important other oxidation-prone residues are considered, including Cys, Trp, Tyr and His (Kim et al., 2001). If a complete reduction were attempted then it is conceivable that re-oxidation may occur during downstream processing. Finally, while it is possible to alkylate Met through the use of iodoacetic acid, the reaction is reversible under reduction conditions (Goverman and Pierce, 1981, Kleanthous and Coggins, 1990).

While several reagents have been reported as being capable of oxidising Met, including dimethyl sulfoxide (DMSO), *tert*-Butyl hydroperoxide, sodium periodate, sodium perborate, hydrogen peroxide, chloramine T, tribromocresol and N-chlorosuccinimide, few are suitable for the complete oxidation of Met

alone. For example, the final three chemicals listed were all found to oxidise small amounts of Cys, His and Tyr, in addition to Met (Keck, 1995), while Try based cleavage has been reported with the latter (Fujii et al., 1978). Of the remaining mild oxidative agents, hydrogen peroxide has been identified as being non-Met specific (Keck, 1995), yielded small amounts of stable Met sulphone (Fujii et al., 1978), while *tert*-Butyl hydroperoxide has been found to modify only a limited number of exposed Met residues within an intact protein (Keck, 1995).

DMSO has also been reported as being capable of oxidising Met, yielding complete modification of free Met in less than three hours. Furthermore, the author also explains how though the use of dimethyl sulfide (DMS) it is possible to completely reduce the oxidised Met in under four minutes with no observable side reactions (Shechter, 1984).

Despite the promising results documented for DMSO, the oxidative chemical of choice is sodium periodate. In addition to both Kettenbach *et al.* and Kirkpatrick *et al.* recommending its use for oxidising Met containing AQUA peptides (Kettenbach et al., 2011, Kirkpatrick et al., 2005a), Fujii *et al.* found sodium periodate to completely oxidise Met in under seven hours, with negligible levels of sulphone (Fujii et al., 1978), while Yamasaki *et al.* reported nearly quantitative formation of Met sulfoxide when treating free Met with equimolar amounts of sodium periodate (Yamasaki et al., 1982). An alternative to sodium periodate may also exist in sodium perborate, which it is detailed can achieve the same results as sodium periodate but over much longer periods of time (Fujii et al., 1978).

While both Cys and Met containing AQUA peptides may therefore be utilised, the majority of the AQUA peptide selection criteria (detailed in Section 2.1) must be strictly adhered to (Brun et al., 2009, Kettenbach et al., 2011, Kirkpatrick et al., 2005a, Pratt et al., 2006). By way of an example, if a peptide is selected which is above the 15 amino acid limit, difficulties can occur during SPPS. Likewise, if a peptide is selected which is below six amino acids in length, problems may be

encountered in the retention of the peptide on the C18 column. Indeed, while Trp, like Cys and Met, can be easily oxidised during sample preparation, no single solution as yet exists so as to enable the use of Trp containing peptides.

2.3 The AQUA Peptide Selection Process

The AQUA peptide selection process employed during this study can be broken down into three specific stages: (i) The generation and initial screening of the candidate peptides (during which a theoretical digestion of the target protein is performed with a range of proteases and each candidate peptide screened); (ii) The alignment of each candidate peptide against the host proteome (through the use of a protein basic local alignment search tool (BLASTP)), and; (iii) Screening each remaining candidate peptide for sequence flaws, through comparing it to the AQUA peptide selection criteria (Figure 2-1).

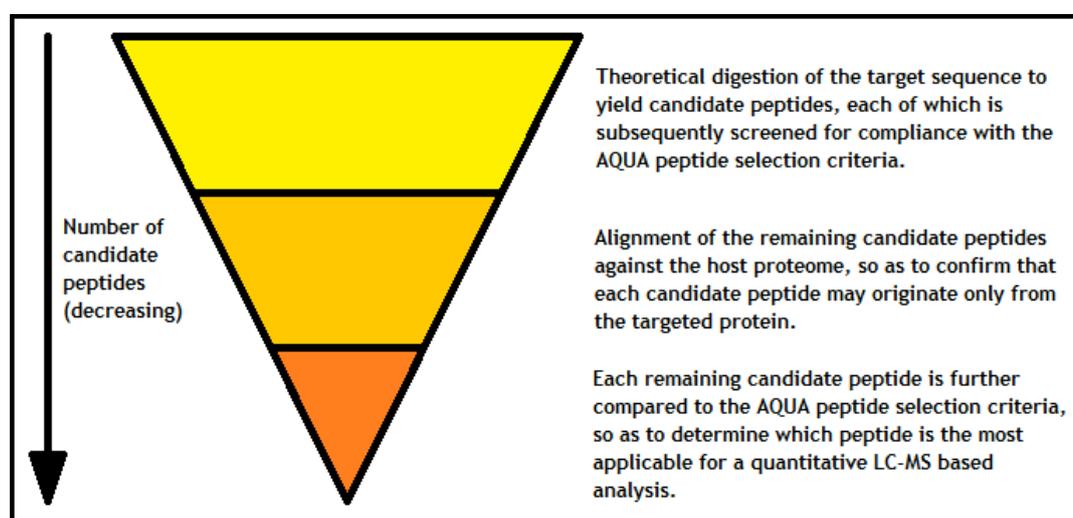


Figure 2-1: The three stages of selection employed when choosing an AQUA peptide. Described are the three specific stages of peptide selection employed during the selection of an AQUA peptide, including the theoretical digestion of the target protein sequence and the screening of each candidate peptide against the AQUA peptide selection criteria, the BLASTP based alignment of each remaining candidate peptide against the host proteome, and the comparison of each remaining candidate peptide, so as to determine which is the most applicable for an LC-MS based quantitative analysis.

2.3.1 Initial Candidate Peptide Screening

Following the theoretical digestion of the sequence unique to the target protein with a range of proteases, each candidate peptide was screened, so as to determine its length (with a length of between 6-15 amino acids being mandatory), to ensure it contained an amino acid residue suitable for stable isotopic labelling (either Ala, Arg, Ile, Leu, Lys, Phe, Pro or Val), and so as to determine the percentage of the candidate peptides sequence which was composed of hydrophobic amino acid residues (with a value of greater than 10%, but less than 50% being optimal). Those candidate peptides which fell outside these selection criteria were eliminated, while those peptides which conformed were subjected to a BLASTP based alignment.

2.3.2 BLASTP Based Alignment

So as to determine if the sequence of each candidate peptide was unique within its host proteome, each peptide was subjected to a BLASTP based alignment. Those peptides which were found to originate from several protein species were eliminated, while those candidate peptides which were found to have a unique sequence were screened for sequence flaws, through comparing each to the AQUA peptide selection criteria.

In setting up the BLASTP based search the “Non-redundant protein sequences (nr)” database was selected and the specific organism search restricted to “*Homo sapiens* (taxid:9606)”, finally the search algorithm was retained at its default setting “blastp (protein-protein BLAST)”, as shown on Figure 2-2.

BLASTP programs search protein databases using a protein query. [more...](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#) [Query subrange](#) [?](#)

From

To

Or, upload file [Browse...](#) [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

[Align two or more sequences](#) [?](#)

Choose Search Set

Database [?](#)

Organism [Exclude](#) [+](#)

Optional Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown. [?](#)

Exclude [Models \(XM/XP\)](#) [Uncultured/environmental sample sequences](#)

Optional

Entrez Query

Optional Enter an Entrez query to limit search [?](#)

Program Selection

Algorithm

[blastp \(protein-protein BLAST\)](#)

[PSI-BLAST \(Position-Specific Iterated BLAST\)](#)

[PHI-BLAST \(Pattern Hit Initiated BLAST\)](#)

Choose a BLAST algorithm [?](#)

Figure 2-2: The NCBI BLASTP search settings utilised when aligning candidate proteolytic peptides against the *Homo sapiens* proteome. So as to test if a candidate AQUA peptide was unique to the *Homo sapiens* proteome, the NCBI BLASTP based search tool was utilised, screening the non-redundant database, with the search restricted to *Homo sapiens* (taxid:9606) and the BLASTP search algorithm selected.

By way of an example, Figures 2-3 and 2-4 show the Blast output data resulting from a unique peptide and a non-unique peptide respectively.

Accession	Description	Max score	Total score	Query coverage
BAG62397.1	unnamed protein product [Homo sapiens]	36.7	36.7	100%
AAH07514.1	ARAF protein [Homo sapiens]	36.7	36.7	100%
AA08754.1	raf related protein [Homo sapiens]	36.7	36.7	100%
NP_001645.1	serine/threonine-protein kinase A-Raf [Homo sapiens] >sp P10398.2 ARA	36.7	36.7	100%
CAA28476.1	unnamed protein product [Homo sapiens]	36.7	36.7	100%
3Q96_A	Chain A, B-Raf Kinase Domain In Complex With A Tetrahydronaphthalene	27.4	27.4	90%
3QG7_A	Chain A, B-Raf Kinase V600e Oncogenic Mutant In Complex With Plx4032	27.4	27.4	90%


```

>|db|BAG62397.1| C unnamed protein product [Homo sapiens]
Length=472
GENE ID: 369 ARAF | v-raf murine sarcoma 3611 viral oncogene homolog
[Homo sapiens] (Over 10 PubMed links)
Score = 36.7 bits (79), Expect = 0.007
Identities = 11/11 (100%), Positives = 11/11 (100%), Gaps = 0/11 (0%)
Query 1 IGTGSFGTVFR 11
          IGTGSFGTVFR
Sbjct 182 IGTGSFGTVFR 192

>|db|AAH07514.1| C ARAF protein [Homo sapiens]
Length=609
GENE ID: 369 ARAF | v-raf murine sarcoma 3611 viral oncogene homolog
[Homo sapiens] (Over 10 PubMed links)
Score = 36.7 bits (79), Expect = 0.007
Identities = 11/11 (100%), Positives = 11/11 (100%), Gaps = 0/11 (0%)
Query 1 IGTGSFGTVFR 11
          IGTGSFGTVFR
Sbjct 319 IGTGSFGTVFR 329

>|db|AA08754.1| C raf related protein [Homo sapiens]
Length=298
GENE ID: 369 ARAF | v-raf murine sarcoma 3611 viral oncogene homolog
[Homo sapiens] (Over 10 PubMed links)
Score = 36.7 bits (79), Expect = 0.007
Identities = 11/11 (100%), Positives = 11/11 (100%), Gaps = 0/11 (0%)
Query 1 IGTGSFGTVFR 11
          IGTGSFGTVFR
Sbjct 25 IGTGSFGTVFR 35

>|ref|NP_001645.1| UGM serine/threonine-protein kinase A-Raf [Homo sapiens]
sp|P10398.2|ARAF HUMAN G RecName: Full=Serine/threonine-protein kinase A-Raf; AltName:
Full=Proto-oncogene A-Raf; AltName: Full=Proto-oncogene A-Raf-1;
AltName: Full=Proto-oncogene Fks
gb|AAA65219.1| C ARAF1 [Homo sapiens]
>|db|CAA28476.1| C unnamed protein product [Homo sapiens]
Length=606
GENE ID: 369 ARAF | v-raf murine sarcoma 3611 viral oncogene homolog
[Homo sapiens] (Over 10 PubMed links)
Score = 36.7 bits (79), Expect = 0.007
Identities = 11/11 (100%), Positives = 11/11 (100%), Gaps = 0/11 (0%)
Query 1 IGTGSFGTVFR 11
          IGTGSFGTVFR
Sbjct 316 IGTGSFGTVFR 326

>|emb|CAA28476.1| C unnamed protein product [Homo sapiens]
Length=606
GENE ID: 369 ARAF | v-raf murine sarcoma 3611 viral oncogene homolog
[Homo sapiens] (Over 10 PubMed links)
Score = 36.7 bits (79), Expect = 0.007
Identities = 11/11 (100%), Positives = 11/11 (100%), Gaps = 0/11 (0%)
Query 1 IGTGSFGTVFR 11
          IGTGSFGTVFR
Sbjct 316 IGTGSFGTVFR 326

>|pdb|3Q96|A Chain A, B-Raf Kinase Domain In Complex With A Tetrahydronaphthalene
Inhibitor
>|pdb|3Q96|B Chain B, B-Raf Kinase Domain In Complex With A Tetrahydronaphthalene
Inhibitor
Length=282
Score = 27.4 bits (57), Expect = 4.8
Identities = 8/10 (80%), Positives = 9/10 (90%), Gaps = 0/10 (0%)
Query 1 IGTGSFGTVF 10
          IG GSGFIV+
Sbjct 18 IGTGSFGTVV 27

```

Figure 2-3: An example of an NCBI BLASTP based alignment; a technique performed so as to identify which peptides are unique within their target proteome. An NCBI BLASTP based alignment of A-Raf WT tryptic peptide IGTGSFGTVFR. Circled in green are the search results obtained from the BLASTP based alignment. From the alignment, five positive matches were identified; these were quickly screened by identifying which proteins displayed a 100% “query coverage”. Those proteins which are circled in blue achieved a 100% query coverage, while those proteins which are marked in red had a sequence coverage of less than 100%. Those candidate peptides which achieved a 100% sequence coverage from only their parent protein were marked as unique within the *Homo sapiens* proteome.

Accession	Description	Max score	Total score	Query coverage
3Q96_A	Chain A, B-Raf Kinase Domain In Complex With A Tetrahydronaphthalene	29.9	29.9	100%
3QG7_A	Chain A, B-Raf Kinase V600e Oncogenic Mutant In Complex With Plx4032	29.9	29.9	100%
ADH51735.1	SLC45A3-BRAF fusion protein [Homo sapiens]	29.9	29.9	100%
ADH51734.1	ESRP1/RAF1 fusion protein [Homo sapiens]	29.9	29.9	100%
ADH51547.1	AGTRAP-BRAF fusion protein [Homo sapiens]	29.9	29.9	100%
3LB7_A	Chain A, Crystal Structure Of C-Raf (Raf-1) >pdb 3LB7 B Chain B, Crystal	29.9	29.9	100%
3IIS_A	Chain A, The Complex Of Wild-Type B-Raf With Pyrazolo Pyrimidine Inhibi	29.9	29.9	100%
3IDP_A	Chain A, B-Raf V600e Kinase Domain In Complex With An Aminoisoquinoli	29.9	29.9	100%
CAR64526.1	SRGAP3:RAF1 fusion protein [Homo sapiens]	29.9	29.9	100%
CAQ43116.1	S-protein KIAA1549:BRF 15_9 [Homo sapiens]	29.9	29.9	100%

Figure 2-4: An example of an NCBI BLASTP based alignment where the peptide is not unique to a single protein. An NCBI BLASTP based alignment of A-Raf WT tryptic peptide WHGDVAVK. The brief search results shown on this image detail the top 15 hits for this peptide, when aligned against the *Homo sapiens* proteome. Due to the large number of proteins identified with a 100% query coverage, this peptide was seen as not being unique, and thus was eliminated from the search.

2.3.3 Screening Each Remaining Candidate Peptide for Sequence Flaws

Having eliminated any peptides which did not possess a sequence unique within the host proteome, each remaining candidate peptide was compared to the AQUA peptide selection criteria, as was detailed in Section 2.1, so as to determine which candidate peptide was most applicable to a quantitative LC-MS based analysis.

2.4 Residue Selection for Stable Isotopic Labelling

Upon a suitable AQUA peptide having been selected, one final factor must be considered prior to ordering, the position of the stable isotopically labelled amino acid residue, assuming that more than one suitable residue exists.

Due to the varying ion fragmentation techniques utilised in each mass spectrometer, the fragmentation spectrum achieved for a peptide will differ between instruments. By way of an example, on a QToF based instrument, such as a QSTAR Pulsar, a range of Y-ions would be expected, with relatively few low series B-ions. In comparison, triple quadrupole based instruments, such as the API 2000, yield a range of both Y- and B-ions, with further internal fragmentations possible, thus presenting with a more complex MS/MS spectrum (Sherwood et al., 2009, Paizs and Suhai, 2005).

So as to enable the stable isotopically labelled fragment ions to be detected during MS/MS scans conducted on both QToF and triple quadrupole based instruments, labelled residues in close proximity to the N-terminal were preferentially selected, so as to enable the labelled residue to be detected in both low B- and high Y-ions.

3 Methods

3.1 Cell Culture and Protein Production

Each of the A-Raf plasmids used during this study were donated by Dr. Jens Rauch, a post doctorate researcher based at Systems Biology Ireland, University College Dublin (previously of the Beatson Institute of Cancer Research, University of Glasgow). Each of the A-Raf plasmids features a human complementary DNA (cDNA) A-Raf insert, pcDNA3.1(+) vector and an N-terminal start codon/FLAG tag “ATGGATTACAAGGATGACGACGATAAG” which results in the additional N-terminal peptide sequence “MDYKDDDDK”.

Each of the PDE4B plasmids (PDE4B1, 2 and 3) used during this study were donated by Dr. George Baillie, a reader based at the Institute of Cardiovascular and Medical Sciences, University of Glasgow. Each plasmid contains a human PDE4B cDNA insert within a pEE7 vector (Celltech, UK).

3.1.1 Plasmid Amplification

A vial of Subcloning Efficiency DH5 α cells (Invitrogen, USA) was transformed with 1 ng of plasmid, as per the manufacturer’s instructions. A single colony was picked from the agar plate, cultured and the amplified plasmid extracted via Midiprep (Promega, USA), as per the manufacturer’s instructions. The amplified plasmid was eluted from the Midiprep cartridge in 600 μ L of nuclease-free water and further purified via ethanol precipitation, reconstituting in 400 μ L of nuclease-free water.

3.1.2 Genejuice Based Transfection of HEK293

The HEK293 cells used during this study were donated by Dr. Sarah Cumming, a research assistant based at the Institute of Molecular, Cell and Systems Biology, University of Glasgow.

A total of 1.7×10^7 HEK293 cells were plated to a 150 mm tissue culture dish and incubated for 24 hours at 37 degrees Celsius ($^{\circ}\text{C}$), 5% (v/v) carbon dioxide (CO_2). The plated cells were transfected with Genejuice, as per the manufacturer's instructions, and incubated for a further 72 hours at 37 $^{\circ}\text{C}$, 5% (v/v) CO_2 . Cell lysis was performed on ice, replacing the spent media with 1 mL of lysis buffer. Following a 45 minute period of cell lysis, the solution was centrifuged at 12,000 rpm for 10 minutes, so as to pellet any cellular debris, while retaining the soluble proteome.

3.1.3 Anti-FLAG Based Immunoprecipitation

Anti-FLAG M2 agarose beads (Sigma-Aldrich, Germany) were added to the lysed cell solution and the FLAG-tagged transfected protein enriched, as per the manufacturer's instructions. The agarose beads were washed in cell lysis buffer three times and the enriched transfected protein eluted in protein loading buffer, vortexing for 30 seconds and incubating at 99 $^{\circ}\text{C}$ for 15 minutes.

3.1.4 Gel Electrophoresis

The enriched protein was loaded to a 4-12% Bis-Tris Gel (Invitrogen, USA) and electrophoresed in an X-cell mini electrophoresis system (Invitrogen, USA) at 50 volts (V) for 15 minutes, increasing to 150 V for 50 minutes.

3.1.5 Coomassie Blue Staining

The electrophoresed 4-12% Bis-Tris Gel was washed three times in water for five minutes before being stained in Coomassie Blue staining solution for 30 minutes. Furthermore, the gel was washed in water twice, each for a period of 60 minutes. An image of the stained gel was captured on a G:Box (Syngene, India), running GeneSnap version 7.08.

3.1.6 Western Blotting

The electrophoresed 4-12% Bis-Tris Gel was incubated in NuPAGE Transfer Buffer (Invitrogen, USA), containing 20% (v/v) MeOH, for 15 minutes before being transferred to an XCell II blotting module (Invitrogen, USA) and electrophoresed at 30 V for a period of 60 minutes, so as to transfer the protein to an Immobilon-P Membrane (Millipore, USA). Post-electrophoresis, the PVDF membrane was washed for five minutes in Tris-buffered saline with 0.05% Tween (v/v) (TBST), three times, before being incubated in Western Blocking Reagent (Roche, Switzerland), as per the manufacturer's instructions. The blocked membrane was again washed for five minutes in TBST, three times, and incubated for a further 60 minutes in primary antibody solution, containing 1:1000 anti-FLAG HRP conjugate antibody (Sigma-Aldrich, Germany) and 0.5% (v/v) Western Blocking Reagent in TBST. The blocked PVDF membrane was washed for five minutes in TBST, three times, and the gel imaged on a G:Box, running GeneSnap version 7.08, with BM Chemiluminescence solution (Roche, Switzerland).

3.2 Recombinant DNA Techniques

3.2.1 Agarose Gel Electrophoresis Based DNA Purification

DNA purification was performed on a 1.5% (w/v) agarose gel, containing SYBR Green (Sigma-Aldrich, Germany), in a Mini Horizontal gel electrophoresis unit (Sigma-Aldrich, Germany) and submerged in TBE Buffer. The gel was electrophoresed at 90 V for 35 minutes and imaged on a G:Box, running GeneSnap version 7.08.

3.2.2 Agarose Gel Based DNA Extraction

The DNA gel band of interest was excised and the DNA was extracted via a QIAquick gel extraction kit (Qiagen, Netherlands), as per the manufacturer's instructions. The concentration of DNA was determined through the use of a Nanodrop 1000 Spectrophotometer (Thermo Fisher, USA), as per the manufacturer's instructions.

3.2.3 T4 DNA Ligation

DNA ligation was performed with T4 DNA Ligase, as per the manufacturer's instructions.

3.2.4 Miniprep Plasmid Screening

Upon a new plasmid having been developed and Subcloning Efficiency DH5 α cells transformed, as detailed previously, 20 colonies per agar plate were picked, cultured and the amplified plasmid extracted via Miniprep (Promega, USA), as per the manufacturer's instructions. The purified plasmid was restriction digested with an appropriate enzyme, purified via agarose gel based electrophoresis and imaged on a G:Box, so as to screen the colonies for the presence of a plasmid containing an insert of the correct size and in the intended orientation.

3.2.5 Plasmid Sequencing and Alignment

Plasmids to be sequenced were sent to DNA Sequencing & Services (Dundee, U.K.) and the resulting data files exported from Applied Biosystems Sequence Scanner Software, version 1.0, to Microsoft Notepad. Each sequence was stored in a FASTA file format and saved as a text file. The FASTA files were opened with Clustal X, version 2.1, and the sequence data aligned.

3.3 AQUA Peptide Characterisation

3.3.1 AQUA Peptide Re-Suspension

AQUA peptides TV*VTVR and GL*NQDCCVVYR (*denoting the uniform stable isotopic amino acid residues (^{13}C , ^{15}N)) were ordered from Thermo Fisher and arrived suspended in 2% (v/v) MeCN at 5 picomole (pmol)/ μL . Each remaining AQUA peptide was ordered from Sigma and arrived lyophilised in 1 nanomole (nmol) vials. Each vial was reconstituted in 200 μL of 2% (v/v) MeCN and re-suspended via sonicating for 30 seconds and vortexing for a further 30 seconds.

3.3.2 AQUA Peptide Alkylation

Immediately prior to use, each AQUA peptide was dried and re-suspended in 500 mM AMBIC, reducing in 45 mM DTT at 60 °C for 30 minutes and alkylating in 100 mM IAA added, incubating in darkness at room temperature for 30 minutes.

3.3.3 Zip Tip Based Sample Cleanup

The alkylated AQUA peptides were purified via Zip Tip, as per the manufacturer's instructions, and eluted in 50% (v/v) MeCN, 0.1% (v/v) Trifluoroacetic acid (TFA). The purified peptides were subsequently dried, re-suspended in 2% (v/v) MeCN and characterised on a range of mass spectrometers.

3.3.4 Determining the Limit of Detection

Three solutions (100 fmol, 10 fmol and 100 attomole (amol)) were prepared for each AQUA peptide, through diluting the 5 pmol/ μ L stocks in 2% (v/v) MeCN. These solutions were used to prepare eight AQUA peptide sample dilutions (5 pmol, 1 pmol, 200 fmol, 50 fmol, 10 fmol, 2 fmol, 500 amol and 100 amol), each of which was diluted to 20 μ L in 2% (v/v) MeCN, dried, re-suspended in AMBIC, reduced and alkylated.

The alkylated peptides were Zip Tip purified and re-suspended in 20 μ L of 2% (v/v) MeCN, 0.1% (v/v) FA. The limit of detection (LoD) for each peptide was determined through LC-MS/MS.

3.4 Single Shot Based Peptide Analysis

3.4.1 Peptide Preparation for MALDI-ToF

The purified alkylated AQUA peptides were spotted onto either a 192 well 4700 MALDI plate or a 384 well Ultraflex II MALDI plate and supplemented with HCCA matrix.

3.4.2 Data Capture on a 4700 MALDI-ToF Based MS

The 192 well MALDI plate containing the peptide samples was loaded to an Applied Biosystems 4700 Proteomics Analyzer MALDI-TOF-TOF-MS (AB SCIEX, USA), equipped with an Nd:YAG laser (355 nm) running in linear mode. The 4700 was run in positive ion mode with an acceleration voltage of +15 kV. The MS was set to detect ions of between 400 m/z and 2,000 m/z and final MS spectra was generated through summing 500 shots, each with an acquisition time of 0.5 ns. To better assess the samples distribution on the MALDI plate, MS data were gathered while manually moving the MALDI plate. Data analysis was performed on DataExplorer version 4.0.

3.4.3 Data Capture on an Ultraflex II MALDI ToF Based MS

The 384 well MALDI plate containing the peptide samples was loaded to a Bruker-Daltonics Ultraflex II TOF/TOF mass spectrometer (Bruker, USA) equipped with a nitrogen laser (337 nm). The Ultraflex II was calibrated before each batch against the monoisotopic $[M+H]^+$ peaks of the PepMix II calibration standard (Bruker, USA). The Ultraflex II MS was run in positive ion reflector mode, with an acceleration voltage of +25 kV and a delayed extraction time of 150 ns. The MS was set to detect ions of between 400 m/z and 2,000 m/z and a final MS spectra generated through summing 500 laser shots, each fired with a repetition rate of 50 Hertz (Hz). The laser power of the Ultraflex II was manually adjusted for each sample spot, setting it to a point just above the minimum power required for the detection of a recognisable spectrum. To better assess the samples distribution on the MALDI plate, MS data were gathered while manually moving the MALDI plate. Data analysis was performed on FlexImaging, version 2.0.

3.4.4 Direct Injection on an API 2000 MS

A borosilicate emitter was loaded with the alkylated peptide, re-suspended in 50% MeCN (v/v), 0.1 % TFA (v/v), and the emitter fitted to an API 2000 ESI triple quadrupole mass spectrometer (AB SCIEX, USA). MS scanning was performed in positive ion enhanced MS (EMS) mode with a voltage potential of 900 V applied

between the emitter and the inlet orifice. The MS was set to detect ions of between 400 and 2,000 m/z , in profile mode, with a scan time of one second and the final MS spectrum obtained through summing the MS data collected over a five minute period. The source temperature was set to 150 °C.

Upon the parent ion peak of interest having been identified in enhanced MS (EMS) mode, the API 2000 was switched to enhanced product ion (EPI) mode and the MS set to fragment the parent ion. The MS was set to detect product ions of between 50 and 1,000 m/z , with a scan time of one second. The collision energy was initially set to 5 electron volts (eV) and increased in steps of 5 until no fragment peaks could be detected above 400 m/z or until the maximum collision energy of 60 eV was reached. An EPI spectrum was obtained at each collision energy setting for a period of two minutes and data analysis was performed on Analyst, version 4.2.

3.4.5 Direct Injection on a QSTAR Pulsar MS

Direct sample injection performed on the QSTAR Pulsar MS (AB SCIEX, USA) utilised the same settings as those detailed for the API 2000 MS. MS scanning was performed, however, in “Q1 MS” mode, while product ion analysis was performed in product ion mode.

3.4.6 Direct Injection on a QTrap 5500 MS

Direct infusion was performed on a 5500 QTrap LC/MS/MS System (AB SCIEX, USA) equipped with a DuoSpray source, featuring both turboionspray and Atmospheric-pressure chemical ionization (APCI) apertures. The syringe pump flow rate was set at 20 $\mu\text{L}/\text{minute}$ (min) through the turboionspray source and an ESI voltage of 6 kV applied between the source and MS orifice. All other settings were retained as detailed for the API 2000 MS. Data analysis was performed on Analyst, version 1.5.1.

3.4.7 Direct Injection on a TripleTOF 5600 MS

Direct infusion was performed on a TripleTOF 5600 MS (AB SCIEX, USA) under the same conditions as those detailed for the 5500 QTrap, utilising the QToF scan types detailed for the QSTAR Pulsar. Data analysis was performed on Analyst TF, version 1.5.1.

3.5 LC-MS Based Peptide Analysis

3.5.1 In-Gel Digest Preparation

The Coomassie staining gel band of interest was excised, retaining 2 millimetre (mm) to each side of the gel lane and cutting at a fixed length of 10 mm, prior to dicing the excised gel band into 1 mm cubes. The gel pieces were washed in 100 mM AMBIC for 30 minutes, partially dehydrated the gel pieces through diluting to a ratio of 1:1 (v/v) in MeCN, and completely dehydrating in MeCN. The gel pieces were re-hydrated in 100 mM AMBIC, prior to reducing in 45 mM DTT, incubating at 60 °C for 30 minutes, and alkylating via the addition of 100 mM IAA, incubating in darkness for 30 minutes. Excess IAA was removed through partially dehydrating the gel pieces by diluting to a ratio of 1:1 (v/v) in MeCN, and the complete dehydrating the gel pieces in MeCN.

3.5.2 In-Gel Digestion

The protein within the gel pieces was digested with one of several proteases, as per the manufacturer's instructions. Upon completion of the digest, the digest buffer was diluted to a ratio of 1:1 (v/v) with MeCN. The gel pieces were incubated twice in 1% (v/v) FA for a period of 20 minutes before dehydrating the gel pieces, three times, in MeCN. All supernatants were pooled and dried.

3.5.3 Spiked Digestion

The reduced and alkylated gel pieces were re-suspended in 0.1% RapiGest SF (w/v) and incubated at 37 °C for ten minutes, prior to drying. Five picomoles of each AQUA peptide were transferred to one of two microcentrifuge tubes, those

which contained Cys, and those without. The Cys containing peptides were dried, re-suspended in 500 mM AMBIC, reduced with DTT and alkylated with IAA, before being transferred to the dried gel pieces, while those devoid of Cys were transferred directly. The in-gel digestion and extraction steps proceeded as described previously. Upon completion of the final in-gel extraction step, the pooled extract was diluted in 0.5% (v/v) TFA and the solution incubated at 37 °C for 45 minutes. The solution was centrifuged at 13,000 rpm for ten minutes and the supernatant decanted to a sterile microcentrifuge tube and dried to completion.

3.5.4 HPLC on an Ultimate 3000

The in-gel digest was re-suspended in 2% MeCN (v/v), 0.1% FA (v/v) and loaded onto an Ultimate 3000 HPLC autosampler (Dionex, The Netherlands), maintained at 4 °C. The HPLC was equipped with a μ -Precolumn Cartridge (300 μ m \times 5 mm, 5 μ m particle size) and a C18 capillary column (75 μ m \times 15 cm, 3 μ m particle size), both packed with PepMAP 100 C18 stationary phase (Dionex, The Netherlands), maintained at 30 °C via an Ultimate 3000 column oven (Dionex, The Netherlands). A 20 μ L sample loop was fitted to the autosampler and the sample injected via the user defined injection mode, controlled through the Chromeleon HPLC software package.

The HPLC micropump flow rate was maintained at 0.3 μ L/min, 2% (v/v) Buffer B (90% MeCN (v/v), 0.1% FA (v/v)), and the loading pump maintained at 20 μ L/min, 100% (v/v) Buffer A (2% MeCN (v/v), 0.1% FA (v/v)). A 60 minute gradient was programmed on the HPLC, increasing the ratio of Buffer B from 2-40% (v/v) Buffer B over 30 minutes, the column was washed for 10 minutes at 90% (v/v) Buffer B, and returned to 2% Buffer B for 20 minutes. The MS trigger was set to activate after 0.1 minutes.

3.5.5 LC-MS on an API 2000 MS

The API 2000 was equipped with a NanoSpray II ESI ion source, fitted with a PicoTip emitter. A voltage of 2.2 kV was applied between the ESI needle and the inlet orifice. The MS was set to detect ions of between 400 and 1,500 m/z with both the Q1 and Q3 transmission windows set to low resolution. The source temperature was set to 150 °C and a rolling collision energy was utilised. A 60 minute IDA scan was performed in positive ion mode, scanning in the MS spectrum and selecting the four most abundant peaks for MS/MS, assuming the IDA selection criteria were met (the ions were between 400 - 1,500 m/z , with a charge state of between $+1$ and $+4$). Ions were excluded from IDA selection after two MS/MS spectra had been acquired for the parent ion.

Protein identification was performed through submitting the MS/MS data to an internal Mascot server, version 1.9 (Matrix Science, UK), aligning the scan data against the NCBI nr (latest version at the time of processing) protein sequence database. Peptide tolerance was set to ± 1.2 Da and the MS/MS tolerance was set to ± 0.6 Da. The search was set to allow for one missed cleavage and for the variable modification of Met (oxidation), while the modification of Cys was set to fixed (carbamidomethylation).

4 The Absolute Quantification of Four A-Raf Isoforms

4.1 Introduction

Oncogene rapidly accelerated fibrosarcoma (Raf) was first cloned and characterised from a mouse with lymphoma and lung adenocarcinoma by Ulf Rapp back in 1983. Rapp identified the cause of the carcinoma to be an acutely transforming murine sarcoma virus (3611-MSV), which was later re-named viral Raf (v-Raf) (Baccarini, 2005, Rapp et al., 1983). Following on from this Sutrave and co-workers later isolated a second gene (v-mill), this time within an avian retrovirus (Mill Hill no. 2 (MH2)) from a spontaneous ovarian tumour within a chicken, which shared an 80% sequence homology with the nucleotide sequence of v-Raf, and a 94% sequence homology with the predicted amino acid sequence of v-Raf (Jansen et al., 1984, Moelling et al., 1984, Sutrave et al., 1984). Jansen later went on to confirm that both these strains of retrovirus (3611-MSV and MH2) contained orthologous protein sequence, while Moelling and Rapp showed v-Raf/v-Mill to be the first identified oncoproteins with both Ser and Thr kinase activity (Jansen et al., 1984, Moelling et al., 1984).

A total of three human Raf kinase genes have since been identified (A-Raf, B-Raf and Raf-1), each displaying a unique cellular expression pattern, regulatory mechanisms, and potency when functioning within the context of the mitogen activated protein kinase (MAPK) pathway (Wojnowski et al., 2000). Despite these differences, each member of the Raf family shares the same general structure (Figure 4-1), with several highly conserved domains, including conserved region one ((CR1), which contains the Rat sarcoma (Ras)-binding domain (RBD), required for Raf to bind with Ras and for membrane recruitment, and the Cys-rich domain (CRD), which functions as a secondary Ras-binding site, yet is important for Raf auto-inhibition), conserved region two ((CR2) which is rich in Ser and Thr residues, important for the inhibitory phosphorylation of Raf) and conserved region three ((CR3) which consists of the kinase domain, but which contains an activation segment whose phosphorylation is crucial for kinase

activity), each of which is separated by a region of more variable sequence (Wellbrock et al., 2004a, Matallanas et al., 2011).

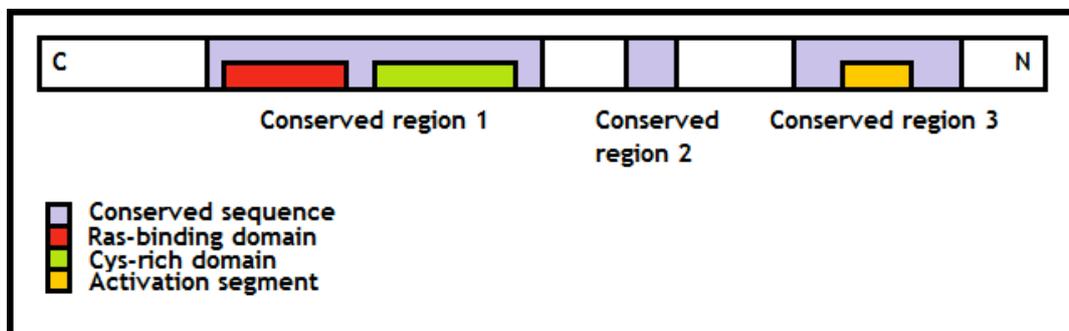


Figure 4-1: The structural organisation of Raf, showing the regions of conserved sequence. Depicted above is the structure of a typical full-length Raf protein, including CR1, which contains both the RBD and the CRD, CR2, which is rich in Ser and Thr residues, and CR3, which consists of the kinase domain, but which also contains the kinase activation segment, important for the activation of Raf.

The MAPK cascade, in which each of the Raf family members are involved, is a complex signal transduction network, controlling a range of cellular processes including cell growth, differentiation, proliferation and apoptosis (Kolch, 2000). These signalling events are triggered when extracellular growth factors, such as epidermal growth factor (EGF) (Grandis and Sok, 2004), or platelet-derived growth factor (PDGF) (Morrison et al., 1989), bind their corresponding membrane bound Tyr kinase receptors (EGFR and PDGFR, respectively). Cytoplasmic signalling proteins such as Src (sarcoma) are then recruited to the cellular membrane and bound to the receptor via the phosphotyrosine binding domain. This binding results in the phosphorylation of Src which in turn creates a binding site for the Src homology 2 domain of protein Grb2 (Cohen et al., 1995, Pawson and Nash, 2000, Schlessinger, 2000). Src binding the mitogen receptor also results in the recruitment of protein SOS to the plasma membrane, causing protein G (Ras) to release its bound guanosine-5'-triphosphate (GDP) and uptake a more abundant guanosine-5'-triphosphate (GTP) molecule, thereby taking its activated configuration (Schlessinger and Bar-Sagi, 1994, Kolch, 2000).

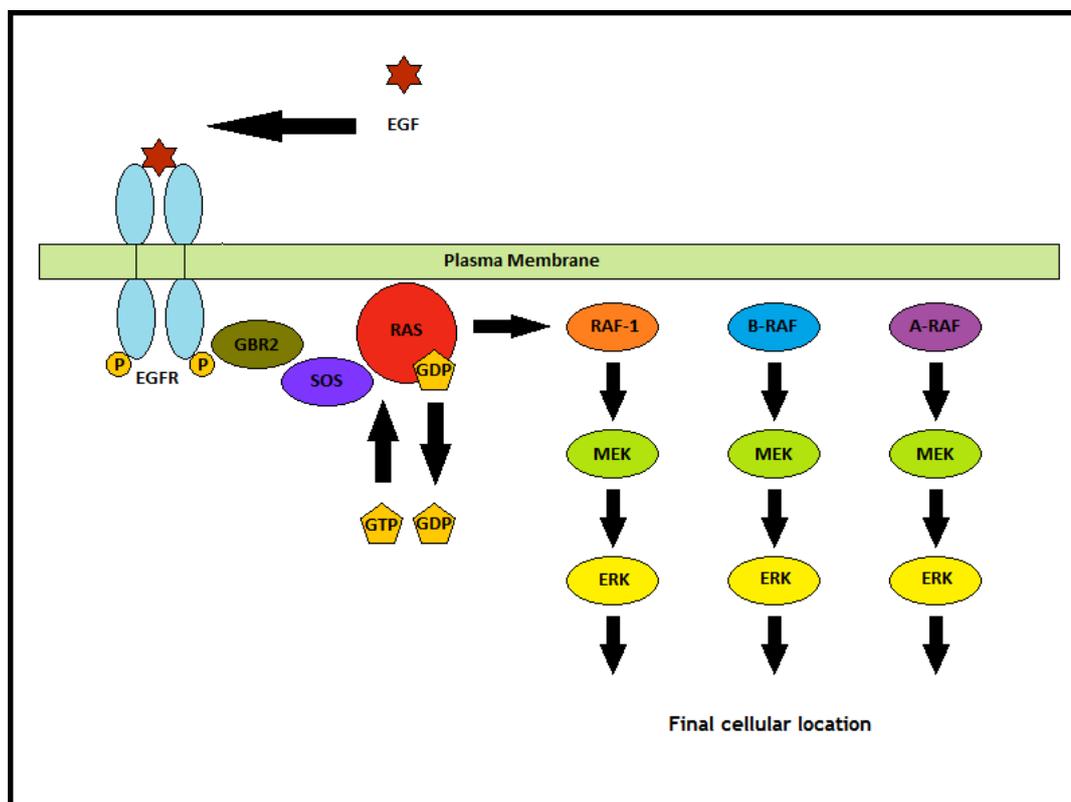


Figure 4-2: The activation and subsequent signalling cascade of the MAPK pathway. A depiction of a mitogen, in this case EGF, activating the MAPK pathway through the stimulation of EGFR. This cascade results in the activation of a variety of transcription factors depending on the initial extracellular stimulus and which MAPKKK is activated.

The activation of Ras thus begins the main kinase cascade of the MAPK pathway, with the binding of Raf (MAPKKK) (Hallberg et al., 1994, Leever et al., 1994, Stokoe et al., 1994), phosphorylation of MEK1/2 (MAPKK) and subsequent phosphorylation of extracellular signal-regulated kinase (ERK) (MAPK) (Gardner et al., 1994, Marais and Marshall, 1996, Yan and Templeton, 1994). Finally the phosphorylated ERK can be translocated throughout the cell, where it can activate various transcription factors, resulting in a unique gene expression pattern depending on the specific stimulus the cascade was activated in response to (Kolch, 2000, Schlessinger and Bar-Sagi, 1994) (Figure 4-2).

While each of the Raf kinase species has its own specific regulatory mechanisms, which will be discussed later, it has been noted that each species is capable of binding 14-3-3 proteins (Kolch, 2000), which are believed to be responsible for

stabilising Raf in each of its conformational states; including the inactive, partially active and fully active. All three Raf family members are also reported as being capable of binding lipids, which is thought to play a role in the translocation (membrane recruitment) and/or regulation of Raf kinase activity (Daub et al., 1998, Ghosh et al., 1996, Kuroda et al., 1996).

While each of the Raf kinases are expressed within human tissue, the expression of A-Raf and B-Raf was initially thought to vary between tissues, while Raf-1 was thought to be expressed ubiquitously, providing a basal level of Raf activity in all human cells (Wojnowski et al., 2000, Storm et al., 1990). More recently however A-Raf has also been shown to be expressed ubiquitously (Storm et al., 1990), and while the tissue specific expression pattern of B-Raf still appears restricted, it has been shown to be the most powerful activator of the MAPK pathway (Araujo et al., 2012), suggesting each isoform may fulfil a specific function in cellular signalling.

4.1.1 Raf-1

Raf-1 was initially thought to be the only member of the Raf family to be ubiquitously expressed, with similar levels of transcripts identified in most human tissue types. This initially led scientists to believe that the function of Raf-1 may be to provide a basal level of Raf kinase activity within all cells (Storm et al., 1990). Further to this, Raf-1 was found to be regulated by a housekeeping type promoter, further supporting the view that A-Raf and B-Raf may have played more specific roles in specialised tissues (Hagemann and Rapp, 1999). More recently Huser *et al.* (Huser et al., 2001) and Mikula *et al.* (Mikula et al., 2001) explored the function of Raf-1 in fibroblast cells isolated from Raf-1 *-/-* knockout mice, identifying Raf-1 deficient mice to suffer widespread apoptosis and to die *in utero*. While no alteration in ERK activation were detected (as B-Raf can fully compensate for the loss of Raf-1 in MAPK signalling), the loss of Raf-1 did increase the susceptibility of the cells toward apoptosis, as Raf-1 is required for the suppression of both proapoptotic mammalian sterile 20-like kinase (MST2) and apoptosis signal-regulating kinase 1 (ASK-1) (Hurst et al.,

1996, Mikula et al., 2001). In brief, MST kinases are cleaved by caspases during cell apoptosis, releasing a constitutively active MST kinase domain which translocates to the nucleus and phosphorylates histone 2B, eventually inducing DNA fragmentation (Romano et al., 2010). ASK-1 meanwhile is activated in response to various types of stress, and in its constitutively active form can phosphorylate both c-Jun N-terminus kinase (JNK) and p38. RAF2 and TRAF6 are subsequently recruited to the ASK1 signalosome where they promote ASK1-dependent cell death and inflammatory cytokine production (Soga et al., 2012).

An alternatively spliced Raf-1 isoform has recently been reported. This isoform differs from the wild type in that it lacks exon three, located within the Raf-like Ras-binding domain. Any functional consequences which result from this variation remain as yet unknown (He et al., 2009).

4.1.2 B-Raf

B-Raf was first identified by Marx *et al.* due to its transduction into the genome of an acute mitogenic retrovirus species (IC10), capable of transforming chicken embryonic neuroretina cells in primary culture (Marx et al., 1988). Simultaneously its human ortholog was also identified in NIH 3T3 cells, transfected with DNA from Ewing sarcoma (Ikawa et al., 1988).

Until recently B-Raf was thought to play a niche roll in cellular signalling (Storm et al., 1990, Wojnowski et al., 2000), due in part to its somewhat restricted expression pattern (mainly localised to within the central nervous system (Barnier et al., 1995, Storm et al., 1990)), at least in comparison to Raf-1. In the last decade however it has been reported that immunoprecipitated (IP) B-Raf has a much stronger affinity for MEK than that of Raf-1, suggesting that it may yet play a major role in MAPK signalling (Huser et al., 2001). This increase in kinase activity is thought to result from the activation of B-Raf requiring fewer phosphorylation events than either Raf-1 or A-Raf (Wellbrock et al., 2004b). An *in vivo* experiment exploring this variation in Raf-MEK binding affinity led to the preparation of several knockout mice, including B-Raf *-/-*, A-Raf *-/-* and A-Raf -

/Y (where A-Raf is located on the X chromosome). The fibroblast cells isolated from the B-Raf $-/-$ knockout mice were found to yield approximately a 30% decrease in ERK phosphorylation when compared to wild-type cell lines (Pritchard et al., 2004, Wojnowski et al., 2000). This is in contrast to A-Raf $-/-$ and $-/Y$ fibroblast experiments, where no significant changes in phosphorylation were recorded (Huser et al., 2001, Mercer et al., 2002, Mikula et al., 2001). Due to this high kinase activity, B-Raf is highly regulated within its inactive state. Inactivated B-Raf is maintained in an auto inhibitory conformation with the carboxy-terminus (C-terminus) kinase domain and the N-terminus regulatory region being involved in an intramolecular interaction (Terai and Matsuda, 2005). In addition, B-Raf is phosphorylated at both Ser 365 (located within the CR2 domain) and Ser 429 residues, further down regulating the kinase activity of the B-Raf.

In the case of B-Raf activation, the auto inhibitory conformation is released upon B-Raf binding a GTPase (Ras in the case of the MAPK pathway), ultimately resulting in the phosphorylation of residues Thr 599 and Ser 602 within the active site (Mason et al., 1999, Peyssonnaud and Eychène, 2001, Wan et al., 2004, Zhang and Guan, 2000). Despite this regulation, B-Raf is still seen as being the least regulated of the Raf kinases. Perhaps this is why B-Raf is so commonly mutated to a constitutively active state in human carcinoma (Abraham et al., 2000).

The most common mutation leading to the activation of B-Raf in human cancer results from a point mutation within the highly conserved Glycine-rich loop, a component of the activation segment located within the kinase domain. This substitution is detectable in approximately 70% of primary melanomas (Davies et al., 2002), in 10% of colorectal cancers (Di Nicolantonio et al., 2008), and in 30-70% of papillary thyroid carcinoma (Kimura et al., 2003), and involves Val 600 being replaced with a Glu; a change which is thought to contribute to tumorigenesis by markedly increasing the basal kinase activity of B-Raf and stimulating constitutive ERK phosphorylation (Prahallad et al., 2012).

An additional route of B-Raf tumorigenesis was recently identified in a human thyroid papillary carcinoma (Ciampi and Nikiforov, 2005). This route of activation involved the proteins N-terminus (responsible for the regulation of the kinase domain) being truncated and the kinase domain fused to staphylococcal nuclease and tudor domain containing 1 (SND1) (Dillon et al., 2011), leading to the constitutive activation of B-Raf (Ikawa et al., 1988, Marx et al., 1988).

In the case of B-Raf isoforms, several reports suggest that complex alternative splicing is indeed undertaken, adding an additional level of regulation to the kinase. These isoforms have been shown to arise from the alternative splicing of exons 8b and 9b, which are unique to B-Raf and located between conserved regions two (cysteine rich domain) and three (kinase active site). This alternative splicing has been shown to yield at least ten tissue specific B-Raf isoforms (Barnier et al., 1995, Hingorani et al., 2003). While the mechanism regulating the alternative splicing of B-Raf remains unclear, a function has been proposed for these structural abnormalities; modifications resulting from the alternative splicing of exon 9b have been shown to increase both MEK kinase activity and the transformation efficiency of B-Raf. However, isoforms which result from the alternative splicing of exon 8b have been shown to have the opposite effect (Hmitou et al., 2007). A proposal has also been made as to the route through which the kinase activity of B-Raf is altered by these splice isoforms. It is believed that the presence of these exons modulate the ability of the protein's N-terminus to interact with the C-terminus kinase domain, a mechanism otherwise used to inactivate B-Raf (Hmitou et al., 2007).

4.1.3 A-Raf

The function of A-Raf has been misinterpreted since its accidental discovery back in 1987 (Ishikawa et al., 1987). A-Raf was firstly thought to be tissue specific, a result of both the ubiquitous expression of Raf-1 and the variable levels of A-Raf expression within different tissue types (Storm et al., 1990, Wojnowski et al., 2000). Furthermore, being a member of the Raf family it was

assumed that A-Raf would exhibit some form of MEK kinase activity, while in fact any kinase activity exhibited by A-Raf is hard to detect (Huser et al., 2001).

While A-Raf expression is ubiquitous, the concentrations expressed within a range of cells are indeed highly variable. The highest concentrations of A-Raf have been identified in urogenital tissues (kidney, testis, epididymus and ovary) while the lowest concentrations have been identified in neuronal tissue (Storm et al., 1990). The protein expression pattern exhibited by A-Raf suggests that it may be in-part regulated by steroid hormone receptors, which are also expressed at high levels in steroid hormone-responsive urogenital tissues (Lee et al., 1996).

In comparison to the sequences of both Raf-1 and B-Raf, several single residue N-terminus substitutions have been identified in A-Raf. It is thought that these substitutions may be partially responsible for the limited kinase activity of the isoform, with Tyr 296 in particular playing a central role (Baljuls et al., 2007). This low kinase activity suggests that A-Raf may fulfil an alternative function within the cell (Huser et al., 2001), a hypothesis which is supported by A-Raf *-/-* knockout experiments, where the removal of A-Raf was found to have no significant effect on the levels of phosphorylated MEK (Pritchard et al., 2004, Wojnowski et al., 2000). Furthermore, the ERK based feedback phosphorylation of A-Raf (Thr-253/Ser-257/Ser-259) appears to positively regulate A-RAF activity (Baljuls et al., 2008), suggesting a possible role in post signalling protein recovery (Nekhoroshkova et al., 2009).

Recent publications regarding A-Raf suggest the existence of several isoform species (Rauch et al., 2011). This research suggests that high levels of heterogeneous nuclear ribonucleoprotein H (hnRNP H) (known to alternatively splice *c-src*, *bcl-x*, *plp/dm20*, *Drosophila nanos*, HIV-1 splicing substrates and rodent tropomyosin) are required for the expression of full length A-Raf (hereby referred to as A-Raf wild type. (A-Raf WT)), whereas no effects have been documented on the expression of either full-length Raf-1 or B-Raf (Rauch et al., 2010). When hnRNP H is expressed at a low level, or indeed in the absence of

hnRNP H, an alternatively spliced 171 amino acid isoform can be detected (Rauch et al., 2011), this new isoform, termed A-Raf Short, is discussed further in Section 4.1.3.2.

4.1.3.1 A-Raf WT

Further to the mouse embryonic fibroblast knockout experiments described in Section 4.1.3, Pritchard *et al.* designed a second set of knockout experiments so as to determine the function of A-Raf WT (Pritchard et al., 1996). From this work both A-Raf $-/-$ and A-Raf $-/Y$ knockout mice appeared to die between 7 and 21 days after birth. Further analysis revealed these mice to display colon organogenesis abnormalities and neurologic defects which resulted in abnormal movement and proprioception. In review Rauch *et al.* commented that the pathological phenotypes observed in these A-Raf deficient mice seemed to suggest an increase in apoptosis, which may be due to a lack of control over the MST2 pathway (Rauch et al., 2010). Rauch *et al.* subsequently performed a small interfering RNA (siRNA)-mediated knockdown of either hnRNP H or A-Raf WT, each of which lead to MST2-dependent cell apoptosis. Further to the conclusion drawn from the A-Raf knockout experiments, this siRNA experiment helped to cement the important role that hnRNP H plays in the splicing of A-Raf WT (Rauch et al., 2010). This research suggests that A-Raf WT is necessary for the inactivation of MST2, and thus the prevention of apoptosis (Figure 4-3).

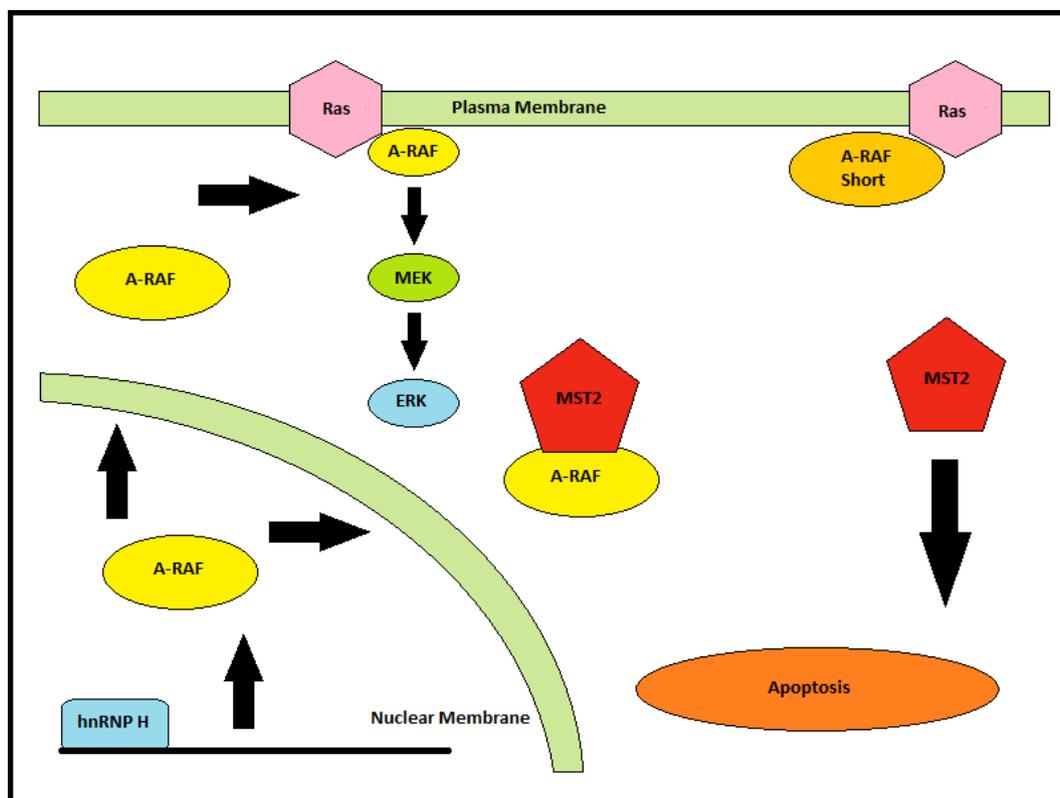


Figure 4-3: The cellular function and interaction of hnRNP H, A-Raf WT and A-Raf Short. In the presence of high levels of hnRNP H A-Raf WT is generated, binding to both MST2, thereby preventing apoptosis, and G-protein Ras, potentially regulating the MAPK pathway. In the presence of low levels of hnRNP H, or indeed in the absence of hnRNP H, A-Raf Short is generated, binding to active Ras, and exerting a dominant negative effect on the MAPK pathway yet having no influence on MST2, which if activated will signal cell apoptosis.

4.1.3.2 A-Raf Short

It was recently discovered that high concentrations of hnRNP H are required for the expression of A-Raf WT. In cells where hnRNP H levels are low, or indeed depleted, as is the case in most cell types, a new short A-Raf isoform is expressed (Rauch et al., 2011). This new isoform, termed A-Raf Short, has a MW of approximately 16 kDa and consists of some 171 amino acid residues. The structure of this new protein lacks approximately $\frac{2}{3}$ of the A-Raf WT C-terminus, including CR3 (the Raf kinase domain) and CR2 while retaining the majority of CR1, including the Ras-binding domain, located on the N-terminus of Raf (Rauch et al., 2011). The generation of this alternative A-Raf Short mRNA results from the inclusion of introns two and four while correctly splicing out

introns one and three. The termination of mRNA translation in A-Raf Short results from the inclusion of intron four, where nucleotide 716 begins the generation of a premature stop codon (Rauch et al., 2011). In a recent publication detailing the expression of A-Raf short (Rauch et al., 2011), Rauch *et al.* comment that “intron inclusion is a rare event in alternative splicing and, in combination with preterminal stop codons, these transcripts are commonly prone to nonsense-mediated decay”. However both Northern and Western blotting against A-Raf Short showed both the mRNA and truncated isoform were detectable in cultured cells, in addition to human tissues (Rauch et al., 2011). As was the case with the A-Raf WT, tissue specific expression patterns were identified for A-Raf Short, with the placenta, kidney, pancreas, lung, and spleen exhibiting high levels of expression, while skeletal muscles, heart, liver, and colon contained only low levels of expression (Rauch et al., 2011).

Due to the structure of A-Raf Short, it has been suggested that this isoform may play an opposing role to that of A-Raf WT in cell signalling. Firstly, due to the presence of CR1, including the Ras binding domain, this isoform has the ability to compete with A-Raf WT, and Raf's in general, for activated Ras, and indeed displays a similar binding affinity to that of the wild type. This functionality was confirmed by Rauch *et al.* who found that A-Raf Short behaved as a dominant negative mutant, suppressing both Ras signalling and ERK activation. Secondly, due to the truncated C-terminus, A-Raf Short lacks the ability to prevent MST2 associated proapoptotic signalling (Rauch et al., 2011).

Further research into the biological significance of A-Raf Short by Rauch *et al.* found that head and neck carcinomas and colon carcinomas which over expressed hnRNP H were also those which expressed lower levels of A-Raf Short. Indeed the expression of A-Raf WT could enable uninhibited MAPK signalling, a possible pathophysiologic mechanism used by tumours to evade apoptosis (Rauch et al., 2011).

4.1.3.3 DA-Raf-1 and DA-Raf-2

During a recent study by Yokoyama *et al.* into the molecular mechanisms through which G-protein M-Ras functions within the cell, the existence of two additional A-Raf isoforms were uncovered. The first of these isoforms, termed Deleted A-Raf 1 (DA-Raf-1), constituted a 186 amino acid protein resulting from the alternative splicing of the A-Raf gene, where intron six had been retained, giving rise to a premature stop codon starting at the second nucleotide of the intron (Yokoyama *et al.*, 2007). DA-Raf-1 was found to contain the entire sequence from CR1, including the Ras-binding domain and Cys-rich domain, but lacked conserved regions 2 and 3, containing the Ser and Thr rich domain and the Raf kinase domain respectively (Wellbrock *et al.*, 2004b, Yokoyama *et al.*, 2007).

The function of DA-Raf-1 was found to be similar to that of A-Raf Short, binding activated Ras and acting as a dominant negative antagonist to the phosphorylation of ERK through the MAPK pathway. As a result DA-Raf-1 is thought to positively regulate myogenic differentiation by inducing cell cycle arrest, muscle-specific protein expression, and myotube formation (Yokoyama *et al.*, 2007).

The second isoform identified during the study was Deleted A-Raf 2 (DA-Raf-2), constituting a 153 amino acid protein, also resulting from the alternative splicing of the A-Raf gene. In this case intron five was retained, giving rise to a premature stop codon starting at the second nucleotide of the intron (Yokoyama *et al.*, 2007).

Besides confirming the ubiquitous expression of the DA-Raf species in a variety of mouse tissues, including the brain and heart, very little additional information was published by Yokoyama *et al.* (Yokoyama *et al.*, 2007). Instead a subsequent publication by Nekhoroshkova *et al.* further addressed the function of these DA-Raf isoforms. In this study the authors suggest that A-Raf WT and DA-Raf-2 are specifically localised to the recycling endosome and that DA-Raf-2 did not necessarily function as an inhibitor of mitogenic signalling (Nekhoroshkova *et al.*,

2009). Instead Nekhoroshkova *et al.* proposed a new model of functionality for the entire A-Raf family (Figure 4-4), where Ras activates both Raf-1 and B-Raf, which in turn are responsible for the activation of the MAPK pathway. It is only then; upon the phosphorylation of ERK, that A-Raf WT is activated, interacting with ADP ribosylation Factor six GTPase (ARF6), possibly through EFA6, ultimately resulting in the recycling of endosome bound receptors (Nekhoroshkova *et al.*, 2009).

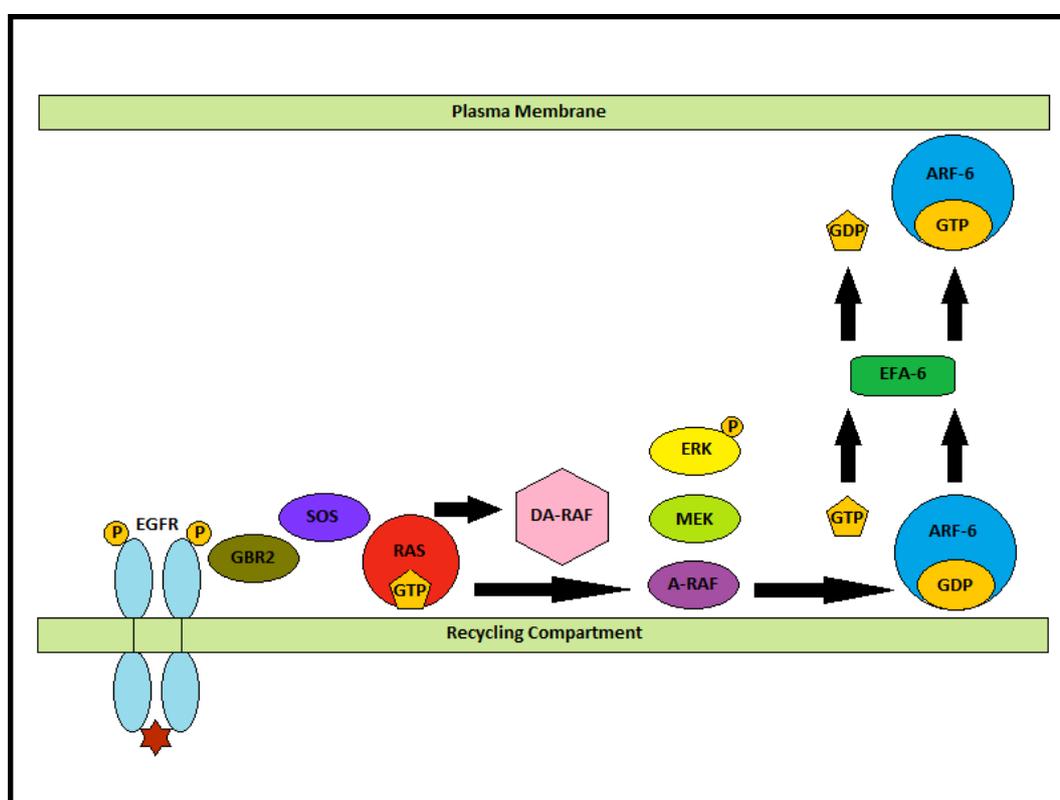


Figure 4-4: The cellular functions and signalling events associated with A-Raf WT and DA-Raf-2, as described by Nekhoroshkova *et al.* Following the activation of the MAPK pathway (not fully depicted), A-Raf WT is activated via ERK based feedback phosphorylation. Activated A-Raf binds G-protein Ras where it activates ARF-6 GTPase, possibly through interacting with EFA6. In the presence of a high concentration of DA-Raf-2, A-Raf WT is prevented from binding Ras and thus ARF-6 activation is blocked (Nekhoroshkova *et al.*, 2009).

Regarding the function of the additional A-Raf isoforms, Nekhoroshkova *et al.* further suggest that DA-Raf-2 may act as a dominant negative antagonist to A-Raf WT, much as was previously suggested for both A-Raf Short and DA-Raf-1 in the MAPK pathway (Nekhoroshkova *et al.*, 2009, Rauch *et al.*, 2011, Yokoyama *et*

al., 2007). In support of the model proposed by Nekhoroshkova *et al.*, A-Raf, ARF6 and EFA6 were all found to be expressed at high levels within the Purkinje cells of the mouse cerebellum, while DA-RAF-2 was found to be expressed at a high level throughout the brain. Nekhoroshkova *et al.* viewed this pattern of A-Raf expression to be of critical importance for rapid endocytosis and synaptic vesicle recycling within the brain, as required for the physiological functioning of neurons (Nekhoroshkova *et al.*, 2009).

So as to confirm the expression of these new DA-Raf isoforms within human tissue, Rauch *et al.* recently probed a total of 89 specimens collected from head and neck and colon biopsies, identifying the transcripts in only one of the samples. In conclusion the author commented that this profile of A-Raf expression fitted our current understanding of alternative splicing; where in general, only two of a given number of potential isoforms will be expressed within a tissue at the same time (Rauch *et al.*, 2011, Wang *et al.*, 2008a).

4.1.4 A-Raf Literature Overview

While the existence of A-Raf WT was first documented some 25 years ago (Ishikawa *et al.*, 1987), the true function of this kinase in molecular signalling is only beginning to be realised, differentiating it from the standard MAPKKK role filled by the other members of the Raf family (Wellbrock *et al.*, 2004b).

Further to that of A-Raf WT, it has only been within the last few years that three A-Raf isoforms have been reported. These isoforms, while differing in sequence, all retain CR1, containing the Ras-binding domain and Cys-rich domain, yet lack both conserved regions two and three, which enable phosphorylation based regulation and Raf kinase activity respectively. Due to the structural qualities of these isoforms many researchers have concluded that they may act as dominant negative regulator of the MAPK pathway through acting as an antagonist to Raf (Nekhoroshkova *et al.*, 2009, Yokoyama *et al.*, 2007). As yet there are very few publications exploring the expression and function of these A-Raf isoforms, and those which do at times appear contradictory. For example, Yokoyama *et al.*

was the first to publish the sequences of both DA-Raf-1 and 2, stating that these isoforms may be involved in the regulation of Ras within the MAPK pathway (Yokoyama et al., 2007). Nekhoroshkova *et al.* later published a paper on the same isoforms, suggesting that DA-Raf-2 may instead be involved in the regulation of receptor recycling within endosomes (Nekhoroshkova et al., 2009). Both of these studies were however conducted within different cell lines, and indeed in different species, which may account for much of the functional variation observed.

Most recently, Rauch *et al.* published a paper on the existence of A-Raf Short, showing this isoform to be involved in the regulation of Ras and the MAPK pathway, while in addition tested 89 tissue samples for DA-Raf 1 and 2, finding only one sample which expressed the isoforms (Rauch et al., 2011). As way of an explanation Rauch *et al.* noted that intronic sequence retention within mRNA was rare and when considered in combination with the detected premature stop codons, it was likely that these transcripts would suffer nonsense-mediated decay (Rauch et al., 2011).

Our understanding of A-Raf and its splice isoforms is, therefore, still in its infancy. It may yet be many years until the true function of these proteins are uncovered.

4.2 Project Aims

Despite there being very few peer reviewed publications confirming the structures and functions of each of these A-Raf isoforms, they make excellent candidates for this project, in which it is hoped the absolute amount of each A-Raf isoform within a sample can be calculated, based on the use of unique AQUA peptides.

Ultimately, when fully established and optimised, this type of analysis should enable researchers to: (i) Identify which A-Raf isoforms are expressed within a tissue (through detecting isoform specific target peptides). (ii) Accurately

quantify each A-Raf isoform detected (through detecting both the isoform specific target and synthetic AQUA peptide), and (iii) Detect changes in expression ratio of the A-Raf isoforms under different physiological conditions.

In order to achieve these ultimate technical aims, which are beyond the scope of a PhD project, a more specific set of project aims were devised. As such, during this project I aim to: (i) Screen each of the previously detailed A-Raf isoform sequences, identifying and selecting suitable AQUA peptides for the absolute quantification of each protein. (ii) Characterise each AQUA peptide/intrinsic target peptide and develop a suitable triple-quadrupole based MRM acquisition method for the identification of each peptide within a single LC-MS run. (iii) Express suitable amounts of each A-Raf isoform, enabling the proteolytic digestions to be optimised using a range of commercially available proteases. (iv) Assess the accuracy and reproducibility of the AQUA technique through the use of a suitable number of digest replicates, and (v) Test the optimised analytical technique on both non-immunoprecipitated exogenously and endogenously expressed A-Raf WT.

4.3 Methods Specific to A-Raf

4.3.1 Proteolytic Digest Optimisation

Lyophilised BSA was re-suspended in either 500 mM AMBIC, or 500 mM AMBIC with 0.1% (w/v) RapiGest SF surfactant. The BSA was reduced and alkylated with DTT and IAA, prior to the IAA being quenched through the addition of Cys. The alkylated BSA was digested with one of three proteases, trypsin, Glu-C or Lys-C, at an enzyme:substrate ratio of 1:10, 1:20, 1:50, 1:100 or 1:200 (w/w), as per the manufacturer's instructions. Following digestion, the solutions were vacuum centrifuged to completion and re-suspended in 2% MeCN, 0.1% FA (v/v). Each digest was separated on an Ultimate 3000 HPLC and analysed on an API 2000 MS. The MS/MS data was submitted to MASCOT and the percentage sequence coverage recorded.

4.3.2 Met Oxidation and Reduction

4.3.2.1 DMSO Based Met Oxidation

Myoglobin, Lysozyme and Lactoglobulin were re-suspended in 500 mM AMBIC, reduced with DTT and alkylated with IAA, prior to quenching with 100 mM Cys. The alkylated proteins were digested via the addition of 1:20 (enzyme:substrate (w/w)) trypsin, as per the manufacturer's instructions. The digested peptide solution was dried, prior to re-suspension in 500 mM Hydrochloric acid (HCl), 100 mM DMSO and oxidised as described by Shechter (Shechter, 1984). The oxidised peptides were spotted to a 192 well MALDI plate and analysed on an AB SCIEX 4700 MALDI-ToF MS.

4.3.2.2 DMS Based Met Sulfoxide Reduction

The DMS based reduction proceeded as described for the DMSO based oxidation, however, the reduction was performed in 100 mM HCl, 300 mM of DMS, as described by Shechter (Shechter, 1984).

4.3.2.3 Sodium Periodate Based Met Oxidation

AQUA peptide VPTV*CVDMSTNRQQ was reduced with DTT and alkylated with IAA, prior to being combined with 125 pmol of sodium periodate in 10 mM 2-(N-morpholino)ethanesulfonic acid (Mes), pH 6.0. The oxidation proceeded as described by Wolschner *et al.* (Wolschner *et al.*, 2009). The oxidised peptides were Zip Tip purified, spotted to a 384 well MALDI plate and analysed via Ultraflex II MALDI-ToF MS.

4.3.2.4 β -mercaptoethanol Based Met Sulfoxide Reduction

Alkylated, oxidised AQUA peptide VPTV*CVDMSTNRQQ, prepared as described in Section 4.3.2.3, was dried and re-suspended in 1%, 10% or 20% (v/v) β -mercaptoethanol, incubating at 99 °C for 60 minutes. The sample was vacuum centrifuged to completion, re-suspended in 0.1% (v/v) TFA and Zip Tip purified, spotting to a 384 well MALDI plate and analysing on an Ultraflex II MS.

4.3.3 MALDI Based Analysis of AQUA Peptide VPTV*CVDMSTNRQQ

4.3.3.1 Characterisation of AQUA Peptide VPTV*CVDMSTNRQQ

Lyophilised AQUA peptide VPTV*CVDMSTNRQQ was re-suspended in either 20 μ L of DMSO or 20 μ L of 10% FA (v/v), sonicated for 30 seconds and vortexed for a further 30 seconds, so as to ensure the complete re-suspension. Each was subsequently diluted to 200 μ L in 0.1% FA (v/v). Peptide VPTV*CVDMSTNRQQ was vacuum centrifuged, re-suspended in 500 mM AMBIC, reduced with DTT and alkylated with IAA, prior to Zip Tip purification and analysis on an Ultraflex II MS.

4.3.3.2 AnchorChip Based Sample Concentration

AQUA peptide VPTV*CVDMSTNRQQ, reconstituted in either 20 μ L of 10% DMSO (v/v), diluted to 200 μ L in 0.1% FA (v/v), or 20 μ L of 10% FA (v/v) diluted to 200 μ L in 0.1% FA (v/v), was spotted to a PAC II MALDI plate and washed *in situ* in 10 mM ammonium phosphate, as per the manufacturer's instructions, prior to analysis on an Ultraflex II MS.

4.3.3.3 LC-MALDI Based AQUA Quantitation

DA-Raf-2 transfected HEK293 IP was loaded to an SDS-PAGE gel and separated, stained, excised, reduced and alkylated. The protein bands were digested with either trypsin or Lys-C, as per the manufacturer's instructions, in the presence of each DA-Raf-2 AQUA peptide. The digested peptide products were extracted and vacuum centrifuging to completion, re-suspended in 10% (v/v) β -mercaptoethanol, in 100 mM AMBIC, and incubating at 99 °C for 60 minutes.

The reduced DA-Raf-2 peptides were injected onto an Ultimate 3000 HPLC and spotted to a 384 well Ultraflex II MALDI plate via a Dionex Probot MALDI Spotter. The Probot spotter was controlled with the Dionex μ Carrier software package, spotting to each MALDI well for a period of five minutes, so as to yield a total of 12 MALDI spots per sample. The 0.3 μ L/min HPLC flow rate was supplemented with a 0.3 μ L/min flow of HCCA matrix solution from the Probot, giving a total

plated droplet size of 3 μ L. The 384 well MALDI plate was analysed on an Ultraflex II MS, summing a total of 500 laser shots and comparing the peak heights for isotopic peaks M, M⁺¹, M⁺² and M⁺³.

4.3.4 A-Raf TNT Based Protein Production

A-Raf WT and A-Raf Short protein stocks were synthesised using a Quick Coupled Transcription/Translation kit, as per the manufacturer's instructions. The spent TNT solution was immunoprecipitated, separated via SDS-PAGE and western blotted, so as to enable the yield of the TNT reaction to be assessed.

4.3.5 A-Raf Short Plasmid Manipulation

4.3.5.1 A-Raf Short Primer Design for SOE PCR

The 3' splice site, 5' splice site and branch point (located 18-40 nucleotides upstream of the 3' splice site (Rogozin et al., 2005)) were identified for intron two of the A-Raf Short cDNA plasmid insert, each of the three were then screened so as to identify a suitable site directed mutagenesis point, identifying a location where; (i) A single base pair could be changed and yet the triplet code of the exon would be maintained intact (retaining the exact protein sequence of A-Raf Short); (ii) A single base pair could be changed, and in doing so modifying the consensus sequence for the splice site (preventing it from being spliced); (iii) The base is located within 1 kb of a unique restriction digest site (enabling the final splice overlap extension (SOE) product to be accurately replicated, cleaved and purified) (Vallejo et al., 2008), and (iv) The creation of a site directed mutagenesis probe would adhere to the general rules of PCR primer selection. Further to the above rules for selecting a suitable site directed mutagenesis site, the final primer must also adhere to the general rules for PCR primer design, as detailed by Dieffenbach *et al.* (Dieffenbach et al., 1993), with the exception of primer length, which in the case of SOE PCR, must be long enough to allow for sufficient annealing to either side of the point mutation (between 18 and 24 bases (Vallejo et al., 2008)).

4.3.5.2 SOE PCR Overlap Production

The SOE PCR performed in this study, as is shown on Figure 4-5, was adapted from the protocol detailed by Heckman and Pease (Heckman and Pease, 2007).

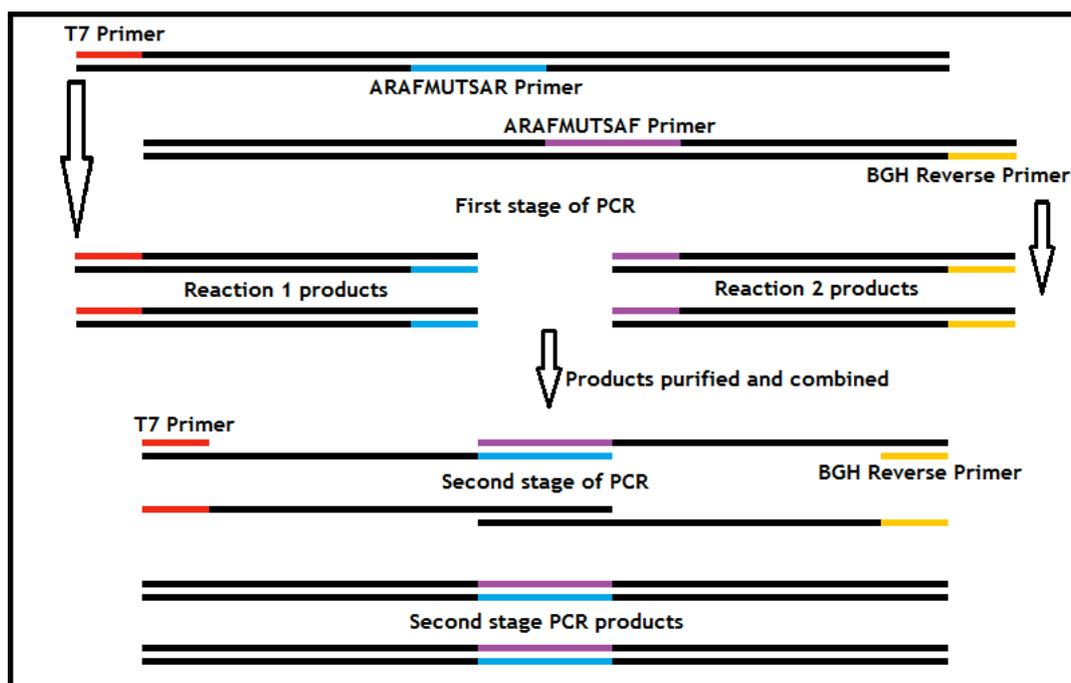


Figure 4-5: The steps involved in the SOE PCR based site directed mutagenesis of A-Raf Short. As to conduct the first step of SOE PCR, two reactions are run, the first of which features the use of both T7 and ARAFMUTSAR primers, while the second features the use of both BGH reverse and ARAFMUTSAF primers, thereby creating two small PCR products which cross the mutagenic target site. The amplified PCR products are purified and combined, serving as templates for the second stage of SOE PCR, which when combined with T7 and BGH reverse primers, extended only the mutated sequences.

Two 32 cycle PCR reactions were performed on an Eppendorf Mastercycler, the first of which contained both the T7 and ARAFMUTSAR (GCCATCCCGGACAGTCACTTGTGTGTATGTGCAGATGTAGG) primers, while the second contained both BGHrev and ARAFMUTSAF (CCTACATCTGCACATACACACAAGTGACTGTCCGGGATGGC) primers. The components added to each reaction and the PCR program settings are detailed in Tables 4-1 and 4-2 respectively.

Table 4-1: The reaction products required for the first PCR stage of SOE PCR. The components required for the production of the overlapping SOE PCR fragments; the first ranging from the T7 promoter to the end of the ARAFMUTSAR primer, while the second ranges from the start of the ARAFMUTSAF primer to the BGHrev promoter. The component volumes listed in this table are those which are required per reaction. A control reaction was also run where the PCR template was replaced by water.

Reagent:	Volume per Tube (μ L)	Final Concentration
10x <i>Pfu</i> Polymerase Buffer	5	1x
2 mmol/ μ L Deoxyribonucleotide triphosphate (dNTP) Mix	5	0.2 mmol/ μ L
100 pmol/ μ L ARAFMUTSAR/ARAFMUTSAF Primer	0.5	50 pmol/ μ L
100 pmol/ μ L T7/BGHrev Primer	0.5	50 pmol/ μ L
300 ng/ μ L A-Raf Short Plasmid	0.3	100 ng
3.5 U/ μ L <i>Pfu</i> Polymerase	1	3.5 U
Nuclease-free Water	37.7	-
Total	50	-

Table 4-2: The PCR program parameters for the first stage of SOE PCR. The PCR cycle details used during the first stage of SOE PCR. As *Pfu* polymerase was utilised, an extension temperature of 72 °C was selected. Meanwhile an annealing temperature of 50 °C was selected, so as to enable the binding of the primers which contained a mismatch.

Cycle Number	Denature	Anneal	Extend
1	95 °C for 3 minutes	-	-
2 - 31	95 °C for 1 minute	50 °C for 1 minute	72 °C for 2 minutes
32	-	-	72 °C for 10 minutes

The completed PCR reaction mixes were supplemented with DNA Gel Loading Buffer, subjected to agarose gel based purification, and extracted with a QIAquick spin column.

4.3.5.3 SOE PCR Overlap Extension

The purified PCR products from the T7/ARAFMUTSAR and BGHrev/ARAFMUTSAF reactions were combined at equal quantities (w/w) and heated to 65 °C for 10 minutes. The water bath was switched off and the temperature slowly returned to that of the room, so as to aid with the annealing of the complementary sequence around the modified base.

The second SOE PCR extension reaction was prepared as is detailed in Table 4-3 and run using the same PCR program as was detailed in Table 4-2.

Table 4-3: The reaction products required for the second PCR stage of SOE PCR. The components required for the extension of the overlapping SOE PCR fragments. The volumes listed on this table are those required per reaction. A control reaction was also run where the PCR template was replaced by water.

Reagent:	Volume per Tube (μ L)	Final Concentration
10x <i>Pfu</i> Polymerase Buffer	5	1x
2 mmol/ μ L dNTP Mix	5	0.2 mmol/ μ L
100 ng/ μ L Purified PCR Mix	2	200 ng
100 pmol/ μ L T7 Primer	0.5	50 pmol/ μ L
100 pmol/ μ L BGHrev Primer	0.5	50 pmol/ μ L
3.5 U/ μ L <i>Pfu</i> Polymerase	1	3.5 U
Nuclease-free Water	36	-
Total	50	-

4.3.5.4 Restriction Digest of the Final PCR Product

The final PCR reaction mix (which should be identical to the insert from A-Raf Short with the exception of the G→A substitution) was double restriction digested with *Bam*HI and *Nhe*I, as per the manufacturer's instructions. Likewise, pcDNA3.1(+) was double restriction digested with *Bam*HI and *Nhe*I, as per the manufacturer's instructions. The restriction digested fragments were separated via agarose gel based electrophoresis, visualised, excised, purified and quantified. The insert and vector were ligated with T4 DNA Ligase, as per the manufacturer's instructions, and transformed into library efficiency DH5 α cells. The transformed cells were plated and approximately 20 colonies per plate screened for the presence of the modified plasmid. In brief, each colony was cultured and the amplified plasmid extracted via Miniprep. The purified plasmid was double restriction digested with *Bam*HI and *Nhe*I, with the aim of identifying an insert with a length of approximately 555 bp. Where a suitable plasmid was identified, it was sent for sequencing, as to confirm the presence of the single base mutation.

4.3.6 Non-Immunoprecipitated Exogenously Expressed A-Raf WT

1.7×10^7 HEK293 cells were plated to a 150 mm tissue culture dish and transfected with the FLAG-tagged A-Raf WT plasmid, through the application of Genejuice transfection reagent, as per the manufacturer's instructions. After 72 hours the spent media was drained and the cells lysed through the addition of cell lysis buffer. The lysed cells were centrifuged at 12,000 rpm for 10 minutes

and the supernatant diluted in Protein Loading Buffer, incubating at 99°C for 15 minutes.

The exogenously expressed A-Raf WT cell lysate was loaded to an SDS-PAGE gel and electrophoresed, stained and the A-Raf WT gel band excised, reduced with DTT and alkylated with IAA, prior to digestion with trypsin in the presence of each of the A-Raf WT AQUA peptides. The proteolytic peptides were extracted, vacuum centrifuged and re-suspended in 2% MeCN, 0.1% FA (v/v), injecting onto an Ultimate 3000 and analysing via a QTrap 5500, running the A-Raf MRM acquisition method previously devised.

4.3.7 Endogenously Expressed A-Raf WT

Endogenous A-Raf WT expression was assessed in HEK293 cells incubated for 96 hours, without the application of transfection.

4.4 Results and Discussion

4.4.1 AQUA Peptide Selection

When searching for an AQUA peptide suitable for the quantification of a protein isoform, regions of sequence unique to that isoform must first be identified. These regions of sequence are theoretically digested, so as to yield a number of candidate peptides. Each of these candidate peptides is screened, so as to determine its length, if it contains an amino acid residue suitable for stable isotopic labelling, and to assess the percentage of hydrophobic residues from which the peptides sequence is composed. Each candidate peptide which is deemed initially suitable is aligned against the host proteome, using the BLASTP search tool, to determine if its sequence is unique within the host proteome, and finally each peptide is compared to the AQUA peptide selection criteria, so as to determine if it is suitable for an LC-MS based quantitative analysis.

Due to the repetitive nature of the AQUA peptide selection process, only the selection of an AQUA peptide suitable for the quantification of A-Raf WT will be detailed, in addition to which, each AQUA peptide selected will be listed.

4.4.1.1 Sequence Unique to Each A-Raf Isoform

The areas of sequence unique to each A-Raf protein isoform are shown on Figure 4-6. Of the A-Raf isoforms, A-Raf WT was found to be the least homologous A-Raf isoform, containing some 421 residues of unique sequence.

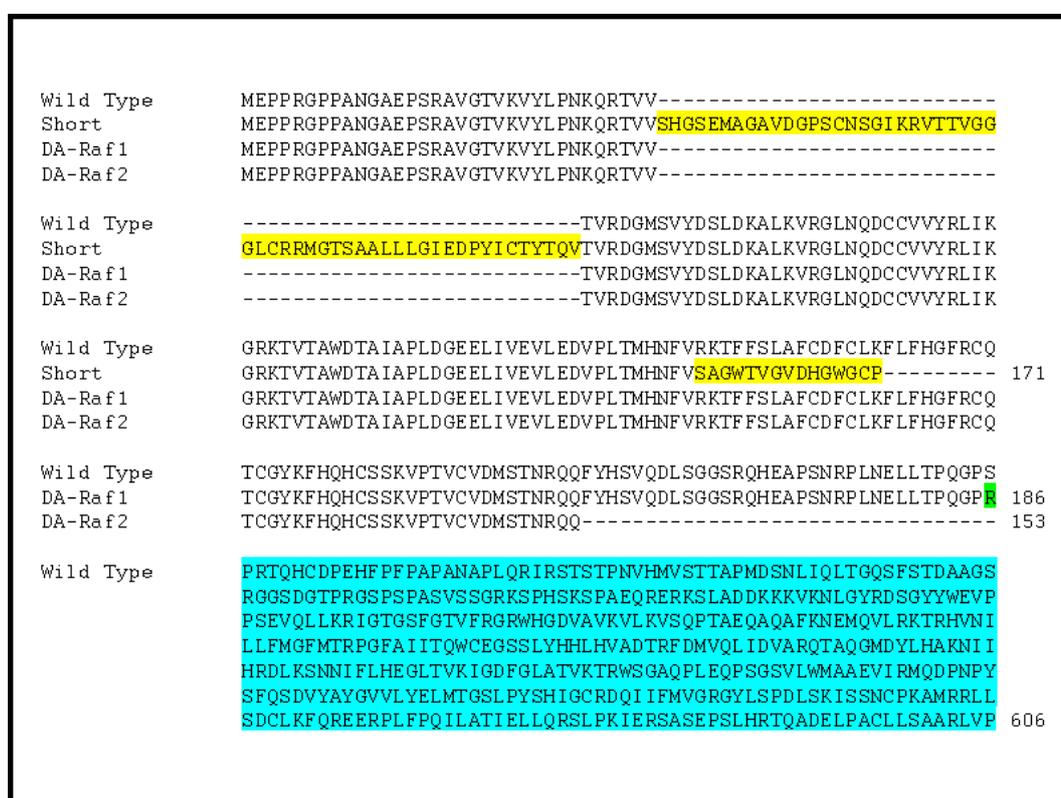


Figure 4-6: An alignment of each of the four A-Raf isoforms, highlighting the regions of sequence unique to each. An alignment of each of the four A-Raf isoforms, based on Clustal X. Highlighted in yellow, green and blue are the regions of sequence unique to A-Raf Short, DA-Raf-1 and A-Raf WT respectively. In being unique these areas are the first targeted in the search for suitable AQUA peptides.

From Figure 4-6 it was also clear that A-Raf Short has a reasonable amount of unique sequence, totalling some 69 unique residues, of which some 54 result from the inclusion of intron two, while the remaining 15 residues resulted from

the inclusion of intron four, prior to termination. DA-Raf-1 meanwhile has only one residue of unique sequence, resulting from the inclusion of intron six, while DA-Raf-2, which results from the inclusion of intron five, appears to be completely homologous to A-Raf WT.

4.4.1.2 AQUA Peptide Selection for the Quantification of A-Raf WT

Of the four known A-Raf protein isoforms, A-Raf WT contained the most unique sequence, and as such selecting an AQUA peptide suitable for the quantification of A-Raf WT should prove the least problematic.

The first step in the search for a suitable AQUA peptide was thus to take the sequence unique to A-Raf WT (Figure 4-6) and to perform a trypsin based theoretical digest, screening each peptide so as to determine its length and amino acid composition (in relation to the number of hydrophobic residues), while searching for the presence of an amino acid suitable for stable isotopic labelling. For a peptide to be deemed suitable it should thus: be between 6 and 15 amino acids in length, be composed of between 10 and 50% hydrophobic residues, and contain at least one residue suitable for stable isotopic labelling. Any peptide which did not meet these criteria was eliminated from the search.

Table 4-4: A theoretical digestion performed on the sequence unique to A-Raf WT. A theoretical digestion performed on the sequence unique to A-Raf WT, based upon the use of trypsin. Peptides which resulted from this theoretical digestion were subsequently screened, and only those with a length of between 6-15 amino acids, which were composed of between 10 and 50% hydrophobic residues and which contained an amino acid suitable for stable isotopic labelling were deemed initially suitable.

Tryptic Peptide Fragment:	Length (AA)	Hydrophobic Residues (%)	Contains AQUA Suitable AA	Initially Suitable
R-QHEAPSNRPLNELLTPQGPSR-T	22	14%	TRUE	No
R-TQHCDPEHFPPAPANAPLQR-I	21	14%	TRUE	No
R-IR-S	2	50%	TRUE	No
R-STSTPNVHMVSTTAPMDSNLIQLTGQSFSTDAAGSR-G	36	22%	TRUE	No
R-GGSDGTPR-G	8	0%	TRUE	No
R-GSPSPASVSSGR-K	12	8%	TRUE	No
R-K-S	1	0%	TRUE	No
K-SPHSK-S	5	0%	TRUE	No
K-SPAEQR-E	6	0%	TRUE	No
R-ER-K	2	0%	TRUE	No
R-K-S	1	0%	TRUE	No
K-SLADDK-K	6	17%	TRUE	Yes
K-K-K	1	0%	TRUE	No
K-K-V	1	0%	TRUE	No
K-VKNLGYR-D	7	29%	TRUE	Yes
R-DSGYWVEVPPSEVQLL-K	17	29%	TRUE	No
K-R-I	1	0%	TRUE	No
R-IGTGSFGTVFR-G	11	36%	TRUE	Yes
R-GR-W	2	0%	TRUE	No
R-WHGDVAVK-V	8	38%	TRUE	Yes
K-VLK-V	3	67%	TRUE	No
K-VSQPTAEQAQAFK-N	13	15%	TRUE	Yes
K-NEMQVLR-K	7	43%	TRUE	Yes
R-K-T	1	0%	TRUE	No
K-TRHVNILLFMGFMRPGFAITQWCEGSSLYHHLHVADTR-F	40	38%	TRUE	No
R-FDMVQLIDVAR-Q	11	55%	TRUE	No
R-QTAQGMDYLHAK-N	12	17%	TRUE	Yes
K-NIIHR-D	5	40%	TRUE	No
R-DLK-S	3	33%	TRUE	No
K-SNNIFLHEGLTVK-I	13	38%	TRUE	Yes
K-IGDFGLATVK-T	10	40%	TRUE	Yes
K-TR-W	2	0%	TRUE	No
R-WSGAQPLEQPSGSLWMAAEVIR-M	23	35%	TRUE	No
R-MQDPNPYSFQSDVYAYGVVLYELMTGSLPYSHIGCR-D	36	28%	TRUE	No
R-DQIIFMVGR-G	9	56%	TRUE	No
R-GYLSPLSK-I	9	22%	TRUE	Yes
K-ISSNCPK-A	7	14%	TRUE	Yes
K-AMR-R	3	33%	TRUE	No
R-R-L	1	0%	TRUE	No
R-LLSDCLK-F	7	43%	TRUE	Yes
K-FQR-E	3	33%	TRUE	No
R-EERPLFPQILATIELLQR-S	18	39%	TRUE	No
R-SLPK-I	4	25%	TRUE	No
K-IER-S	3	33%	TRUE	No
R-SASEPSLHR-T	9	11%	TRUE	Yes
R-TQADELPACLLSAAR-L	15	20%	TRUE	Yes
R-LVP	3	67%	TRUE	No

From Table 4-4 it is clear that even when large regions of unique sequence are theoretically digested with trypsin, the majority of the candidate peptides identified are still unsuitable. Indeed, from the 47 peptides identified during this theoretical digest, only 14 were seen as being initially suitable; less than 30%.

Each remaining candidate peptide was subjected to an NCBI BLASTP based alignment, identifying which of the 14 remaining peptides were unique to A-Raf WT within the *Homo sapiens* proteome. Where a peptide was identified as originating exclusively from the target protein, it was deemed unique. Otherwise the peptide was eliminated from the search.

Table 4-5: An NCBI BLASTP based alignment of each A-Raf WT candidate peptide against the *Homo sapiens* proteome. The results obtained from an NCBI BLASTP based alignment of 14 A-Raf WT candidate peptides against the *Homo sapiens* proteome. Peptides found to originate only from A-Raf WT within the *Homo sapiens* proteome were labelled as unique.

Tryptic Peptide Fragment:	Unique
K-SLADDK-K	No
K-VKNLGYR-D	Yes
R-IGTGSFGTVFR-G	Yes
R-WHGDVAVK-V	No
K-VSQPTAEQAQAFK-N	Yes
K-NEMQVLR-K	Yes
R-QTAQGMDYLHAK-N	No
K-SNNIFLHEGLTVK-I	No
K-IGDFGLATVK-T	No
R-GYLSPLSK-I	No
K-ISSNCPK-A	Yes
R-LLSDCLK-F	No
R-SASEPSLHR-T	No
R-TQADELPACLLSAAR-L	Yes

Following the completion of the NCBI BLASTP based alignment, only six of the initial 47 candidate peptides identified in Table 4-4 were found to be unique to A-Raf WT within the *Homo sapiens* proteome (Table 4-5), less than 13%.

The next step in the identification of an optimal A-Raf WT candidate peptide was to further analyse each of the six remaining candidate peptides, through comparing each to the AQUA peptide selection criteria, so as to identify any sequence flaws which may affect the efficiency of the proteolytic digest. Shown

in Table 4-6 are the results obtained from the screening of each of the candidate peptides.

Table 4-6: Further screening of the A-Raf WT candidate peptides against the AQUA peptide selection criteria. A more thorough analysis of each A-Raf WT candidate peptide, identifying which of the six remaining peptides complied with each of the AQUA peptide selection criteria.

Tryptic Peptide Fragment:	Comments on Peptide Suitability:
K-VKNLGYR-D	C-terminus of peptide succeeded by Asp.
KR-IGTGSFGTVFR-G	N-terminus of peptide preceded by Lys.
K-VSQPTAEQAQAFK-N	Suitable AQUA peptide.
K-NEMQVLR-K	C-terminus of peptide succeeded by Lys.
K-ISSNCPK-A	Suitable AQUA peptide.
R-TQADELPACLLSAAR-L	Suitable AQUA peptide.

From Table 4-6 it is clear that only three of the six remaining candidate peptides were somewhat suitable. Of the three peptides which were rejected, two were eliminated due to the presence of flanking basic residues, while the third was eliminated due to the presence of a flanking acidic residue.

The final step in the selection a suitable AQUA peptide was thus to compare each of the three remaining candidate peptides, so as to determine which was most suited to an LC-MS based analysis, assessing both length and sequence, the results from which are shown below.

Peptide VSQPTAEQAQAFK:

Peptide VSQPTAEQAQAFK was cleaved with trypsin, and thus a minimum of two charges per peptide would be expected. Furthermore, the initial characterisation of this peptide (Table 4-4) suggests it to be composed of just 15% hydrophobic residues, therefore no problems were foreseen with the use of this peptide during LC-MS/MS.

Peptide ISSNCPK:

Upon reviewing the sequence of peptide ISSNCPK, two consecutive Ser residues were detected, which in the presence of a low pH solution, can undergo dehydration. This reaction converts the Ser residues to dehydroalanine, and in

doing so yielding multiple MS peaks per peptide. As such, this peptide was eliminated from the search.

Peptide TQADELPACLLSAAR:

On reviewing the sequence of peptide TQADELPACLLSAAR, it was found to be composed of 15 amino acid residues, which while permitted, increases the chance of this peptide presenting with multiple charged states. So as to better explain the problem, when a peptide may exist in one of a number of charged states, the amount of peptide detected for any specific MS parent ion peak would be reduced, therefore increasing the amount of sample required for the LoD to be reached. As such this peptide was eliminated from the search.

In conclusion, of the three candidate peptides, only one was deemed suitable for an AQUA based quantification, peptide VSQPTAEQAQAFK. This peptide was found to be of a reasonable length, while lacking any obvious sequence flaws and was found to conform to all of the AQUA peptide selection criteria.

A-Raf WT serves also as an example of the problems associated with selecting a proteotypic peptide for the quantification of a protein isoform. From a 58 kDa protein, unique sequence totalling 421 amino acids was identified, from which, only 47 candidate peptides were identified, only one of which was found to be ideal.

4.4.1.3 A-Raf AQUA Peptide Selection Overview

Through the use of the AQUA peptide selection workflow, as was detailed in Chapter 2, it has been possible to select a single AQUA peptide for the quantification of all but one A-Raf isoform, A-Raf Short, which instead was quantified through the use of two AQUA peptides, each of which is listed on Table 4-7, and shown on Figure 4-7.

Table 4-7: The AQUA peptides selected for the quantification of each of the A-Raf isoforms. Each of the AQUA peptides selected for the quantification of each of the four A-Raf isoforms.

Peptide	Protease	Present in:	Comments:
VSQPTAEQAQAFK	Trypsin	A-Raf WT	Located within CR3, and thus unique to A-Raf WT
TVVTVR	Trypsin	A-Raf WT, DA-Raf-1 & DA-Raf-2	Selected for the quantification of A-Raf Short, through subtracting the quantitation value obtained for this peptide from that obtained for peptide GLNQDCCVVYR
GLNQDCCVVYR	Trypsin	All four A-Raf isoforms	Selected for the quantification of all four A-Raf isoforms.
LLTPQGPR	Glu-C	DA-Raf-1	The C-terminus peptide of DA-Raf-1
VPTVCVDMSTNRQQ	Lys-C	DA-Raf-2	The C-terminus peptide of DA-Raf-2

Wild Type	ME PPRGP PANGAE PSRAVGTVKVYLPNKQR	TVV	-----
Short	ME PPRGP PANGAE PSRAVGTVKVYLPNKQRTVV	SHGSEMAGAVDGP SCNSGIKRVTTVGG	
DA-Raf1	ME PPRGP PANGAE PSRAVGTVKVYLPNKQR	TVV	-----
DA-Raf2	ME PPRGP PANGAE PSRAVGTVKVYLPNKQR	TVV	-----
Wild Type	-----	TVRDGMSVYD SLDKALKVR	GLNQDCCVVYRLIK
Short	GLCRRMGTS AALLGIEDPY ICTYTQVT	VRD GMSVYD SLDKALKVR	GLNQDCCVVYRLIK
DA-Raf1	-----	TVRDGMSVYD SLDKALKVR	GLNQDCCVVYRLIK
DA-Raf2	-----	TVRDGMSVYD SLDKALKVR	GLNQDCCVVYRLIK
Wild Type	GRKTVTAWDTA IAPLDGEEEL IVEVLE DVPLTMHNFVRKT	FFS LA FCFDFCLKFEL FHGFRCC	
Short	GRKTVTAWDTA IAPLDGEEEL IVEVLE DVPLTMHNFV	SAGWTVGV DHGWGCP	-----
DA-Raf1	GRKTVTAWDTA IAPLDGEEEL IVEVLE DVPLTMHNFVRKT	FFS LA FCFDFCLKFEL FHGFRCC	
DA-Raf2	GRKTVTAWDTA IAPLDGEEEL IVEVLE DVPLTMHNFVRKT	FFS LA FCFDFCLKFEL FHGFRCC	
Wild Type	TCGYKFHQHCS SKVPTVCVDMSTNRQQ	FYHSVQDL SGGSRQHEA PSNR PLNELLT PQGPS	
DA-Raf1	TCGYKFHQHCS SKVPTVCVDMSTNRQQ	FYHSVQDL SGGSRQHEA PSNR PLNELLT PQGPS	
DA-Raf2	TCGYKFHQHCS SK	VPTVCVDMSTNRQQ	-----
Wild Type	PRTQHCDPEHF PFPAPANAPLQRI RSTSTPNVHMVSTTA	FMDSNLIQLTGQSFSTDAAGS	
	RGGSDGT PRGS PS PA SVS SGRKSPHSKS PAEQRRERKS	LADDDKKVKNLGYRDSGYWEVP	
	PSEVQLLKRIGTGSFGT VFRGRWHGDVAVKVLK	VSQPTAEQAQAFK	NEMQVLRKTRHVN I
	LLFMGEMTRPGFA IITQWCEGS SLYHHLHVADTRF	DMVQLIDVARQTAQGM DYLHAKNI I	
	HRDLKSNNI FLHEGLTVKIGDFGLATVKTRW SGAQPLEQ	PSG SVLWMAAEVIRMQDPNPY	
	SFQSDVYAYGVVLYELMTGSLPYSHIGCRDQ IIFMVGRG	YLS PDLSKI SSNC PKAMRRL L	
	SDCLKFQREERPLFPQILAT IELLQRSLPKIERSASE	PSLHRTQADELPA CLLSAARLVP	

Figure 4-7: Each of the AQUA peptides selected for the quantification of the four A-Raf isoforms. A Clustal X alignment of each of the four A-Raf isoforms. Highlighted in yellow is peptide TVVTVR, present in A-Raf-WT, DA-Raf-1 and DA-Raf-2. Highlighted in green is peptide GLNQDCCVVYR, present in each of the four A-Raf isoforms. Highlighted in blue is peptide LLTPQGPR, present only in DA-Raf-1. Highlighted in red is peptide VPTVCVDMSTNRQQ, present only in DA-Raf-2. Finally, highlighted in purple is peptide VSQPTAEQAQAFK, present only in A-Raf WT.

4.4.2 AQUA Peptide Optimisation

4.4.2.1 The Oxidation and Reduction of Met

4.4.2.1.1 The Modification of Met Within Three Intact Proteins

So as to assess the efficiency with which DMSO and DMS oxidise and reduce Met, respectively; Lactoglobulin, Lysozyme and Myoglobin were reduced, alkylated and digested with trypsin, prior to being oxidised with DMSO, or reduced with DMS, and analysed via MALDI-ToF. Due to the repetitive nature of this experiment, only those results obtained for Myoglobin are detailed.

Myoglobin			
Swiss-Prot: P68083.2			
Exact Molecular Weight = 17185.0586			
MGLSDGEWQQ VLNWVGKVEA DIAGHGQEV LIRLFTGHPET LEKFDKFKHL KTEAEMKASE	60		
DLKKHGTIVL TALGGILKKK GHHEAELKPL AQSHATKHKI PIKYLEFISD AIIHVLHSHK	120		
HPGDFGADAQG AMTKALELFR NDIAAKYKEL GFQG	154		
	MW	M/Z (+1)	M/Z Oxidised
TEAEMK	707.3148	708.3148	724.3148
HPGDFGADAQGAMTK	1501.6598	1502.6598	1518.6598

Figure 4-8: The Met containing tryptic peptides identified in Myoglobin. A theoretical tryptic digestion was performed on Myoglobin, so as to identify any Met containing peptides, following which the mass to charge ratios for both the oxidised and reduced species were calculated.

From Figure 4-8 it is clear that just two Met containing peptides were identified during the tryptic digestion of Myoglobin, and while both peptides were detected via MS, only peptide HPGDFGADAQGAMTK appears to have been affected by either the DMS based reduction or the DMSO based oxidation (Figure 4-9).

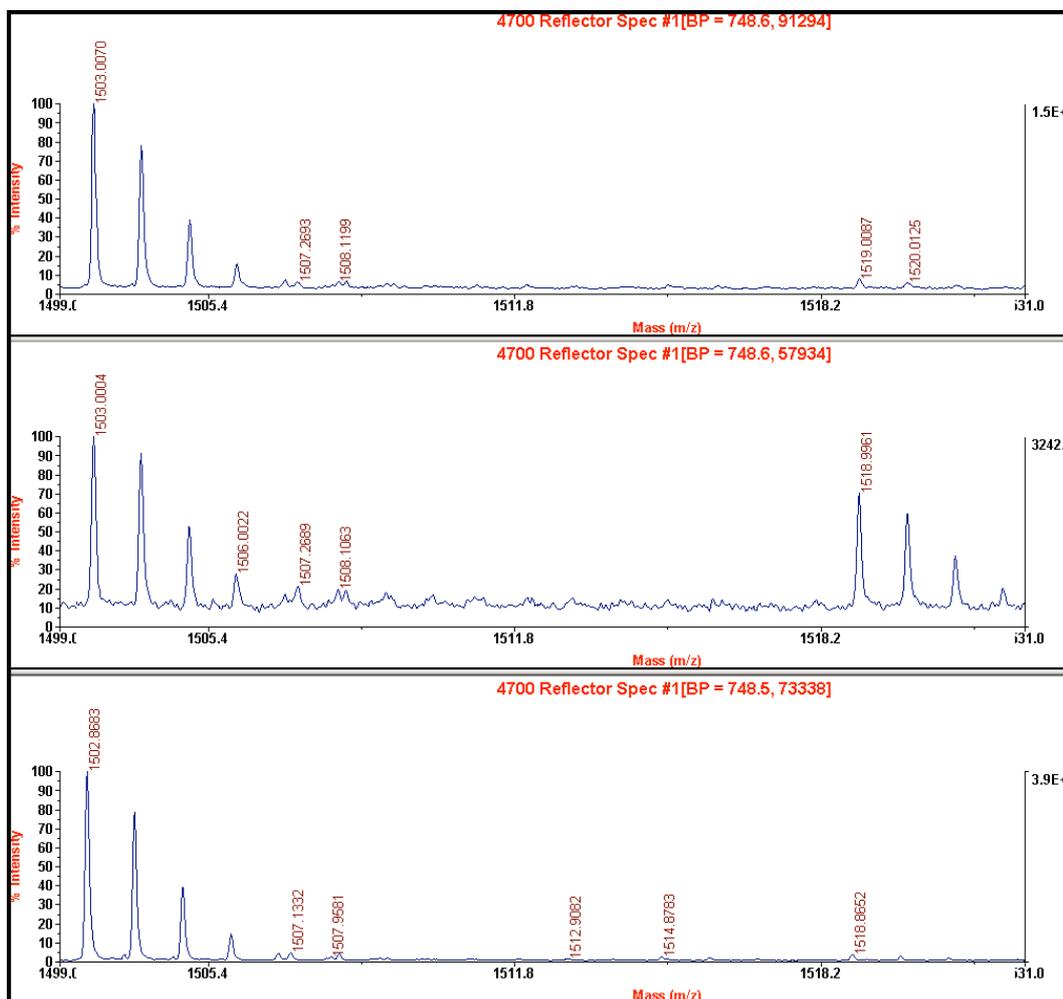


Figure 4-9: The DMS and DMSO based reduction and oxidation of Met containing Myoglobin peptide HPGDFGADAQGAMTK. Following the tryptic digestion of Myoglobin, peptide HPGDFGADAQGAMTK was oxidised with DMSO and reduced with DMS, prior to analysis based upon MALDI-ToF. Peptide HPGDFGADAQGAMTK is shown in an unmodified state (top spectra), following DMSO based oxidation (middle spectra) and after DMS based reduction (bottom spectra).

From Figure 4-9 it is clear that peptide HPGDFGADAQGAMTK, in an unmodified state, is almost completely reduced. While this is good for gauging the efficiency with which DMSO oxidises this peptide, it makes assessing the DMS based reduction impossible. When oxidised with DMSO, peptide HPGDFGADAQGAMTK presented with a +16 Da peak, approximately 70% the height of reduced peak 1,503 m/z , therefore, while the oxidation of peptide HPGDFGADAQGAMTK was at least partially successful, it is far from being complete.

Upon the completion of the DMS and DMSO based reduction and oxidation of Met, it was clear that both of these reactions were not as efficient as is required for a successful AQUA analysis, and as such were abandoned. When unmodified Met containing peptides presented in a partial oxidised state, the DMS and DMSO based reactions were reasonably efficient. In comparison, in those situations where the unmodified peptide was almost completely reduced, the efficiency of the DMSO based oxidation was poor. This would suggest that the reactions may only proceed in situations where the Met is susceptible to modification, perhaps somewhat influenced by the sequence of the surrounding amino acid residues.

4.4.2.1.2 Met Modification within AQUA Peptide VPTV*CVDMSTNRQQ

Met containing AQUA peptide VPTV*CVDMSTNRQQ was treated with DTT and IAA prior to the addition of sodium periodate, so as to completely oxidise the Met. Upon the completion of this reaction, a complete reduction was attempted, based on the use of β -mercaptoethanol.

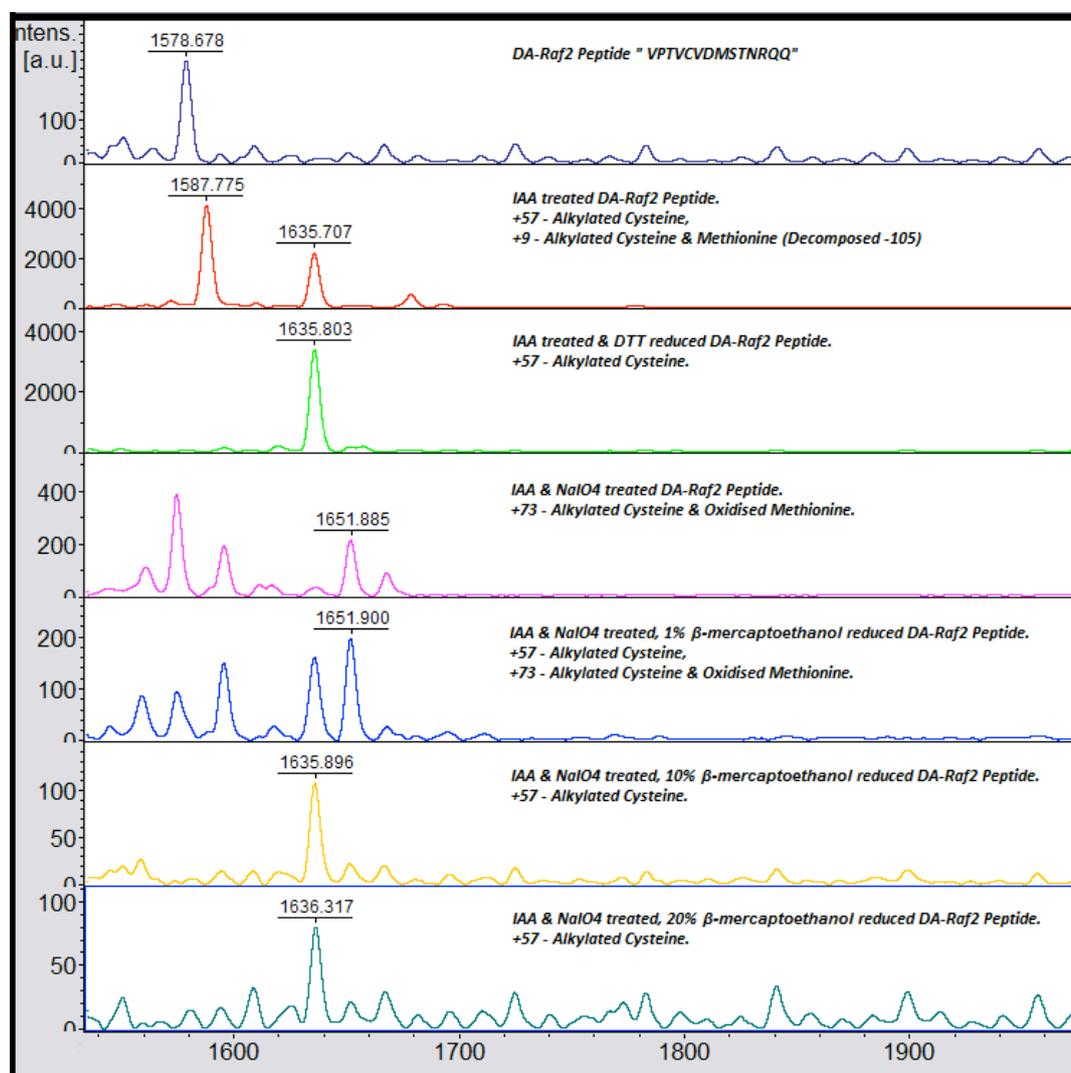


Figure 4-10: The Sodium Periodate based oxidation and β -mercaptoethanol based reduction of AQUA peptide VPTV*CVDMSTNRQQ. A screenshot of multiple spectra obtained from an Ultraflex II MS, in which the complete sodium periodate based oxidation and β -mercaptoethanol based reduction of AQUA peptide VPTV*CVDMSTNRQQ were attempted. The reduction of the Met sulfoxide back to Met was performed under three conditions, in 1%, 10% and 20% (v/v) β -mercaptoethanol. All intended modifications are detailed in comparison to the singly charged, reduced parent ion peak, 1,578.7 m/z .

From Figure 4-10 it is evident that the Met within AQUA peptide VPTV*CVDMSTNRQQ is almost completely reduced. When reacted with DTT and IAA, the parent ion peak (1,578.7 m/z) is replaced by two new peaks, the first with a mass shift of +57 Da, suggestive of IAA based alkylation, while the second has a mass shift of just +9 Da, explained by Kruger *et al.* to result from IAA based non-specific alkylation of Met. In situations where Met alkylation occurs,

two 57 Da modifications are incorporated, however, during ionisation, in-source fragmentation occurs, causing a loss of 105 Da (Kruger et al., 2005). It is therefore recommended that all future alkylation based reactions should be preceded by an additional DTT reduction step, reducing any alkylated Met, so as to yield a single +57 Da MS peak (Kleanthous and Coggins, 1990).

The alkylated peptide was subsequently oxidised with sodium periodate, while the literature suggested a single +16 Da (+73 Da from unmodified) peak should present, instead approximately nine peaks were detected, each at a different intensity. While it is conceivable that oxidation based damage may have degraded various elements within the peptide, it has proven possible to reduce each of these peaks back to a single +57 Da parent ion. Instead, as was the case during alkylation, these peaks may result from the sodium periodate causing several reversible side reactions, each of which is fragmented during ionisation.

When a β -mercaptoethanol based reduction of the oxidised peptide was attempted, 1% (v/v) β -mercaptoethanol proved somewhat inefficient. When the level of β -mercaptoethanol was increased to 10% (v/v), however, almost all of the Met sulfoxide was reduced to Met. Increasing the level of β -mercaptoethanol further, to 20% (v/v), the same results were seen as those obtained in 10% (v/v) β -mercaptoethanol, but the reduced Met peak presented at a lower intensity. The loss of peak intensity at 20% (v/v) β -mercaptoethanol may result from the high levels of β -mercaptoethanol affecting the ability of the peptide to bind the Zip Tip resin, resulting in peptide loss prior to MALDI spotting. Similarly if Zip Tip purification was failing to remove the β -mercaptoethanol, then matrix crystallisation and ionisation may be affected.

While the ability of sodium periodate to oxidise Met was clearly greater than that of the acidified DMSO, complete modification though oxidation was still seen to be the least attractive option. Through selecting the correct oxidative chemical, it was thought that undesirable side reactions could be avoided, yet non-specific oxidation persisted. Instead complete reduction in 10% (v/v) β -

mercaptoethanol, immediately prior to LC-MS, may offer this experiment the greatest chance of success.

4.4.2.2 Proteolytic Digest Optimisation

When an AQUA based quantitative analysis is conducted, the digest efficiency is of critical importance. So as to identify the reaction conditions under which the maximum digest efficiency could be achieved for each utilised protease (trypsin, Glu-C and Lys-C), BSA was digested with each proteases at a range of enzyme:substrate ratios ((w/w) from 1:10 to 1:200), with or without the addition of RapiGest SF ALS, the results of which are shown on Table 4-8.

Table 4-8: The optimisation of trypsin, Lys-C and Glu-C based digestions of BSA. The results obtained when the digest efficiencies of trypsin, Lys-C and Glu-C were assessed through the digestion of BSA at a range of enzyme:substrate ratios, with and without the addition of 0.1% (w/v) RapiGest SF surfactant. The results shown detail the percentage sequence coverages achieved.

Surfactant	Protease	Protease Concentration (Enzyme:Protein, (w/w))					Average
		1:10	1:20	1:50	1:100	1:200	
-	trypsin	17%	10%	14%	15%	13%	14%
0.10%	trypsin	14%	16%	11%	14%	14%	14%
-	Glu-C	19%	15%	15%	11%	13%	15%
0.10%	Glu-C	33%	40%	30%	34%	27%	33%
-	Lys-C	4%	5%	6%	4%	7%	5%
0.10%	Lys-C	15%	16%	11%	8%	6%	11%
Average		17%	17%	15%	14%	13%	

It should first be noted that the sequence coverages detailed on Table 4-8, while low, do not necessarily indicate the digest efficiencies of the proteases to be particularly low. These numbers instead indicate the portion of the BSA peptideome which was available for detection on an API 2000 MS. Ideally this would include all peptides between 8 and 25 residues in length (Brownridge and Beynon, 2011), which yield a *m/z* of between 400 and 1,500, with a single to quadruple charge state, which are suitable for retention on a C18 column, and elute within the 60 minute RP-LC gradient. Further to these limitations, the API 2000 has a long cycle time and a relatively high LoD, when compared to modern MS instruments; as a result, only the most abundant peptides were detected.

Each set of protease digests were also run at different times, with slightly different instrument setups, and as such the sequence coverages achieved between proteases cannot be compared. Indeed, as an alternative strategy, the number of peptides detected during each analysis may also have been compared. The aim of this experiment was, however, simply to assess if the RapiGest SF ALS surfactant had a positive impact on protein digest efficiency, as would be indicated through an increase in the BSA sequence coverage, and thus the exact method utilised was not critical. In addition, the BSA digest efficiencies achieved at a range of enzyme:substrate ratios (w/w) were also assessed.

While the addition of the ALS surfactant yielded very little improvement in the sequence coverage achieved for BSA digested with trypsin, it appears to have made a substantial difference to the sequence coverages achieved for the BSA digested with either Glu-C or Lys-C (Table 4-8). In regard to detecting the optimum enzyme:substrate ratio (w/w), the results achieved for both the trypsin and Glu-C based digestions, without the addition of ALS, would suggest the highest sequence coverages were achieved with an enzyme:substrate ratio of 1:10. These results, however, conflict with those obtained through the digestion of BSA in the presence of ALS, for which an enzyme:substrate ratio of 1:20 yielded the highest sequence coverages. This conflicting data may be traced to several factors; firstly, in the absence of the ALS, more protease may have been required to obtain a higher sequence coverage. Alternatively, the initial high sequence coverages achieved (Trypsin and Glu-C 1:10 enzyme:substrate ratio) may have resulted from the optimised ESI spray and the presence of a new picotip emitter at the beginning of the batch, temporarily boosting the performance of the API 2000.

From this brief experiment it was concluded that an enzyme:substrate ratio of at least 1:20 (w/w), should be achieved during each proteolytic digest, and that RapiGest SF should be utilised.

4.4.3 MRM Acquisition Method Design

When utilising a triple quadrupole based MS, MRM scanning can be employed, thereby increasing the sensitivity of the MS towards the specific target ions, while reducing the background noise detected. Indeed this type of acquisition method is ideal for use during an AQUA based quantitation where only a limited number of target peptides are to be detected. MRM based acquisition methods do, however, require a great deal of manual optimisation, including the selection of a dwell time low enough to achieve several data points per eluting peak, yet high enough so as provide a reasonable sensitivity (Keshishian et al., 2007), while requiring the selection of suitable fragment ions which are reasonably abundant and indeed the optimisation of the collision energy for each fragment ion (Kettenbach et al., 2011).

4.4.3.1 MRM Selection and Optimisation

So as to identify both the exact Q1 mass of each AQUA peptide, and to enable the selection of three Q3 fragment ion peaks per peptide, each AQUA peptide was reduced, alkylated, Zip Tip purified and loaded onto a Proxeon offline Borosilicate emitter and infused on an API 2000 MS. In brief, the m/z ratio of the parent ion was detected through the use of an EMS scan, collecting ions between 300-1000 m/z over one second. So as to obtain a higher quality spectrum, the MS data collected over five minutes was summed, prior to analysis. Upon the MS parent ion having been detected, the API 2000 scan mode was switched to EPI, collecting ions between 100-1500 m/z over one second. Again the EPI data was summed for a total of five minutes, prior to analysis.

In this chapter, only the characterisation of peptide VSQP*TAEQAQAFK will be detailed, a peptide which most commonly fragments at a single amino acid (Pro), and as such is one of the more difficult peptides for which to identify multiple fragment ions.

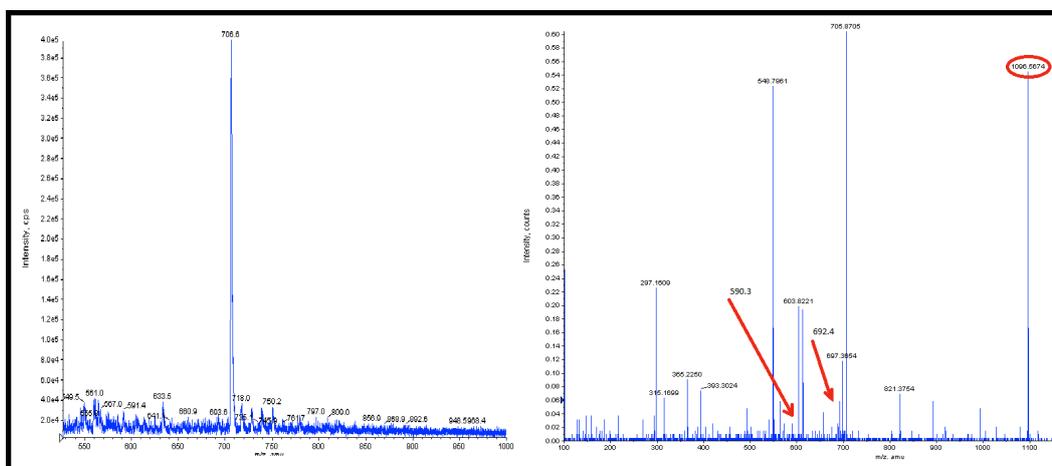


Figure 4-11: Both the MS and MS/MS spectra obtained for AQUA peptide VSQP*TAEQAQAFK on an API 2000. (Left), The MS spectrum obtained for AQUA peptide VSQP*TAEQAQAFK during a direct injection, showing a single doubly charged peak at 706.6 m/z , which was subsequently selected as the Q1 mass of interest for all future A-Raf MRM based analyses. (Right) The product ion spectra obtained from the fragmentation of peak 706.6 m/z , with a collision energy of 25 eV. Marked in red are the three Q3 fragments manually selected for use in the A-Raf MRM acquisition method.

Figure 4-11 shows the intense doubly charged peak detected for peptide VSQP*TAEQAQAFK at 706.6 m/z . As no signal was detected for the singly or triply charged peptide species (data not displayed), peak 706.6 m/z was selected as the Q1 target and therefore subjected to EPI based fragmentation.

VSQPTAEQAQAFK				VSQP(+6)TAEQAQAFK			
MH ⁺ (av)		MH ⁺ (mono)		MH ⁺ (av)		MH ⁺ (mono)	
1405.5598		1404.7118		1411.4951		1410.7118	
[-] Main Sequence Ions				[-] Main Sequence Ions			
b		y		b		y	
---	1	V	13	---	1	V	13
187.1077	2	S	12	1305.6434	187.1077	2	S
315.1663	3	Q	11	1218.6113	315.1663	3	Q
412.2191	4	P	10	1090.5527	418.2191	4	P(+6)
513.2667	5	T	9	993.5000	519.2667	5	T
584.3039	6	A	8	892.4523	590.3039	6	A
713.3464	7	E	7	821.4152	719.3464	7	E
841.4050	8	Q	6	692.3726	847.4050	8	Q
912.4421	9	A	5	564.3140	918.4421	9	A
1040.5007	10	Q	4	493.2769	1046.5007	10	Q
1111.5378	11	A	3	365.2183	1117.5378	11	A
1258.6062	12	F	2	294.1812	1264.6062	12	F
---	13	K	1	147.1128	---	13	K

Figure 4-12: The predicted fragmentation spectrum for both synthetic and endogenous peptide VSQPTAEQAQAFK. A predicted fragmentation spectrum for the synthetic (right) and endogenous (left) versions of peptide VSQP*TAEQAQAFK, as obtained from Protein Prospector: MS-Product, enabling a more rapid identification of suitable Q3 fragment ions.

When Q1 ion 706.6 m/z was subjected to an EPI based scan, several intense fragment ions were detected. One of the most abundant peaks identified was 1,096.6 m/z , which in comparison to a predicted fragmentation spectrum (Protein Prospector MS-Product, Figure 4-12) was identified as being the Y₁₀ ion. As this peptide contained a Pro residue, the most likely residue for a fragmentation to occur, few other intense Y or B-ion peaks were detected. Further analysis however identified both the Y₆ and B₆ ions, which presented at 692.4 m/z and 590.3 m/z , respectively, both of which were selected as suitable Q3 ions.

The optimum collision energy for each fragment ion was determined through increasing the Q2 collision energy from 5 eV, in multiples of five, from a minimum of five to a maximum of 60.

While selecting the most intense fragment ion for an MRM based acquisition method may seem sensible, thereby lowering the amount of peptide required per detection, Kettenbach *et al.* suggests that preference should be given to those fragments with a m/z ratio higher than their parent ion. As internal fragmentation ions are generally detected in the lower spectrum (below 400

m/z). Selecting a fragment with a higher m/z should therefore ensure less noise is detected for the transition (Kettenbach et al., 2011). Furthermore, ion trap instruments such as the API 2000 are less efficient at trapping ions with m/z 's below approximately $1/3$ of the parent ion m/z , therefore selecting low MW fragment ions is best avoided (Hopfgartner et al., 2004).

Upon selecting a total of three complete MRM transitions per peptide, the corresponding target peptide values were calculated based on the MW of the AQUA peptide minus the additional weight of the ^{13}C and ^{15}N labelling. An equal dwell time was selected for each A-Raf MRM transition, resulting in a total scan time of two seconds per cycle. This approach enabled the collection of 15 data points per transition, assuming an LC-MS peak width of 30 seconds.

Shown on Table 4-9 is the complete A-Raf MRM acquisition method, detailing the transitions selected for each target and synthetic peptide, with the exception of Met containing AQUA peptide VPTV*CVDMSTNRQQ.

Table 4-9: The MRM transitions selected for the detection of each A-Raf target/AQUA peptide. The MRM based acquisition method developed for the detection of each of the four A-Raf isoforms; including the Q1 and Q3 MRM transitions, the Q2 optimised collision energy and the dwell time selected for each transition (in milliseconds (ms)).

Peptide (Q1)	Fragment (Q3)	Q1 m/z	Q3 m/z	Dwell Time (ms)	Collision Energy (eV)
TV*VTVR	B ₂	340.7	207.1	75	5
TV*VTVR	Y ₃	340.7	375.2	75	15
TV*VTVR	Y ₄	340.7	474.3	75	10
TVVTVR	B ₂	337.7	201.1	75	5
TVVTVR	Y ₃	337.7	375.2	75	15
TVVTVR	Y ₄	337.7	474.3	75	10
VSQP*TAEQAQAFK	Y ₁₀	705.9	1096.6	75	25
VSQP*TAEQAQAFK	Y ₆	705.9	692.4	75	30
VSQP*TAEQAQAFK	B ₆	705.9	590.3	75	25
VSQPTAEQAQAFK	Y ₁₀	702.9	1090.6	75	25
VSQPTAEQAQAFK	Y ₆	702.9	692.4	75	30
VSQPTAEQAQAFK	B ₆	702.9	584.3	75	25
GL*NQDCCVVYR	B ₃	695.8	292.2	75	30
GL*NQDCCVVYR	Y ₂	695.8	338.2	75	30
GL*NQDCCVVYR	Y ₃	695.8	437.3	75	30
GLNQDCCVVYR	B ₃	692.3	285.2	75	30
GLNQDCCVVYR	Y ₂	692.3	338.2	75	30
GLNQDCCVVYR	Y ₃	692.3	437.3	75	30
LLTP*QGPR	Y ₇	444.3	774.4	75	15
LLTP*QGPR	Y ₆	444.3	661.4	75	15
LLTP*QGPR	Y ₅	444.3	560.3	75	25
LLTPQGPR	Y ₇	441.3	768.4	75	15
LLTPQGPR	Y ₆	441.3	655.4	75	15
LLTPQGPR	Y ₅	441.3	554.3	75	25

4.4.3.2 Detection of AQUA Peptide VPTV*CVDMSTNRQQ

When 5 pmol of AQUA peptide VPTV*CVDMSTNRQQ was analysed on an API 2000 MS, no MS peak was detected. So as to test if this was a fault with the MS, a second sample was prepared and injected onto a QSTAR Pulsar MS, which also failed to detect the peptide. Both machines were later injected with a range of other AQUA peptides, each of which was detected at a reasonable intensity (data not shown).

So as to test if the AQUA vials were empty, or if the peptide had failed to re-suspend, two unopened vials of AQUA peptide VPTVCVDMSTNRQQ were reconstituted in either 20 µL of 10% DMSO (v/v), or 20 µL of 10% FA (v/v), each of which was then further diluted to 200 µL in 0.1% FA (v/v), and analysed on an Ultraflex II MS.

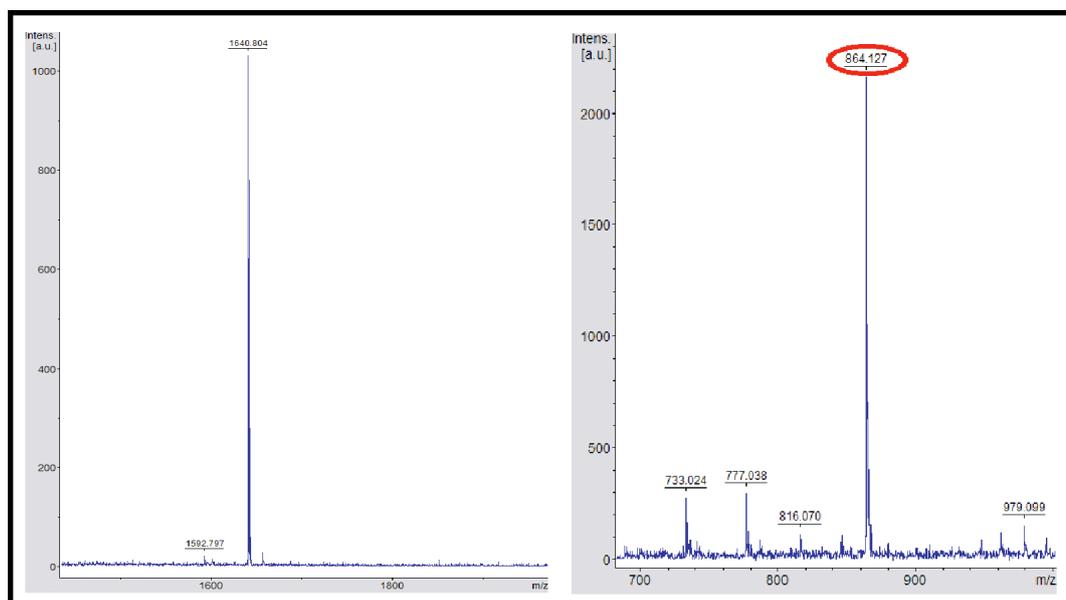


Figure 4-13: A MALDI-ToF based characterisation of AQUA peptide VPTVCVDMSTNRQQ, based upon the use of both MS and MS/MS. (Left) The MS spectrum obtained from the ionisation of AQUA peptide VPTVCVDMSTNRQQ on an Ultraflex II MALDI-ToF MS. Through summing 500 shots it was possible to detect a singly charged peak at 1,640.8 m/z . (Right) The MS/MS spectrum obtained from the fragmentation of singly charged peak 1,640.8 m/z , marked in red is the fragment ion transition selected.

Though a singly charged parent ion peak was detected in both solutions, with a mass to charge ratio of 1,640.8 m/z , the ion intensity was low, achieving only 1,000 counts per 500 summed pulses (Figure 4-13). While not directly comparable, AQUA peptides VSQP*TAEQAQAFK and LLTP*QGPR, analysed at a similar concentration levels and under similar conditions, achieved ion intensities of 5×10^4 and 3.5×10^5 , respectively.

As peptide VPTV*CVDMSTNRQQ could be detected based upon MALDI ToF, an LC-MALDI based method was developed for the quantification of this particular peptide, separating the target sample via SDS-PAGE, reducing, alkylating, digesting in the presence of each of the AQUA peptides, further reducing the Met residue in 10% 2-Mercaptoethanol (v/v), injecting onto an Ultimate 3000 HPLC, and spotting the eluting solution to a 384 well MALDI plate via a Dionex Probot MALDI Spotter, prior to sample analysis on an Ultraflex II MS.

4.4.3.2.1 AnchorChip Based Sample Concentration

In an attempt to increase the MS and MS/MS peak intensities obtained for AQUA peptide VPTV*CVDMSTNRQQ, an AnchorChip PAC II 384 well MALDI plate was tested, which is claimed to offer an increase in sensitivity of 10 and 100-fold over conventional ground steel MALDI plates. The PAC II MALDI plate offers this increase in sensitivity through the application of 800 nm hydrophilic “Anchors” set in a hydrophobic surround. When the sample is spotted to the PAC II MALDI plate, the combination of hydrophilic and hydrophobic surfaces concentrates the drying sample. Furthermore, the HCCA matrix is pre-applied to the PAC II MALDI plate, ensuring ideal analyte crystallisation conditions, while reducing sample preparation time. *In situ* sample purification may also be performed, eliminating Zip Tip based sample loss.

AQUA peptide VPTV*CVDMSTNRQQ, reconstituted in either 20 μ L of 10% DMSO (v/v), diluted to 200 μ L in 0.1% FA (v/v), or 20 μ L of 10% FA (v/v) diluted to 200 μ L in 0.1% FA (v/v), was spotted to a PAC II MALDI plate and washed *in situ* in 10 mM ammonium phosphate. The DMSO containing peptide sample, however, failed to crystallise, instead remaining on the MALDI plate as a dense brown droplet, as is shown on Figure 4-14. The DMSO containing peptide sample was therefore not subjected to MALDI-ToF based analysis.

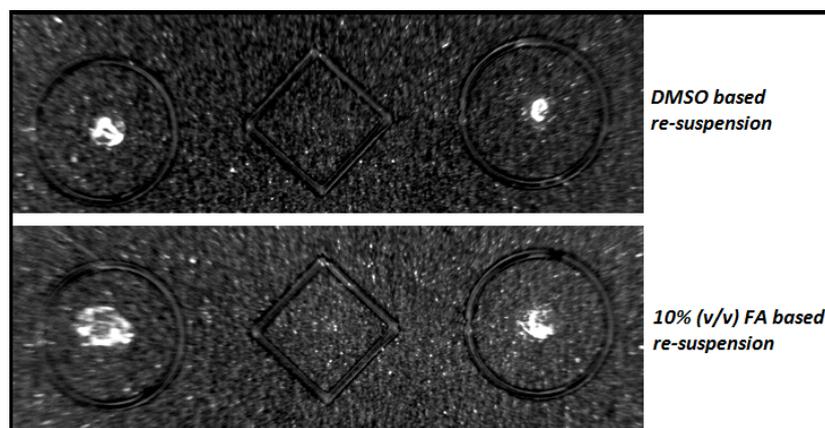


Figure 4-14: AQUA peptide VPTV*CVD MSTNRQQ spotted to an AnchorChip PAC II MALDI plate. AQUA peptide VPTV*CVD MSTNRQQ was spotted to an AnchorChip PAC II MALDI plate, reconstituted in either 20 μL of 10% DMSO (v/v), diluted to 200 μL in 0.1% FA (v/v), or 20 μL of 10% FA (v/v) diluted to 200 μL in 0.1% FA (v/v). While peptide VPTV*CVD MSTNRQQ reconstituted in 10% FA (v/v), diluted to 200 μL in 0.1% FA (v/v) crystallised on the PAC II MALDI plate, the DMSO containing samples failed to crystallise.

When the FA containing peptide sample spot was subjected to MALDI-ToF based analysis, relatively intense HCCA matrix peaks were detected at 832 m/z and 892 m/z (Smirnov et al., 2004), however, the VPTV*CVD MSTNRQQ parent ion peak (1,640.8 m/z) could not be detected (Figure 4-15). As the only difference between this analysis and those run previously was the use of the PAC II MALDI plate, then the application of this plate and *in situ* washing must have been responsible for the loss of the parent ion peak.



Figure 4-15: An MS spectrum of AQUA peptide VPTV*CVDMSTNRQQ, detected on an Ultraflex III MS. The spectrum obtained when detecting AQUA peptide VPTV*CVDMSTNRQQ, re-suspended in 10% (v/v) FA, diluted to 200 μ L in 0.1% (v/v) FA, on an Ultraflex II MS featuring an AnchorChip PAC II MALDI plate. AQUA peptide VPTV*CVDMSTNRQQ MS peak 1,640.8 m/z was not detected.

As the PAC II MALDI plate is intended for use in proteomics, the components contained within this peptide sample should not affect the retention of the peptide VPTV*CVDMSTNRQQ. Indeed the only chemical the manufacturer cautions the use of is MeCN, which was not present in this sample. As such the use of the PAC II MALDI plate was abandoned and a conventional ground steel MALDI plate utilised.

4.4.3.3 MALDI Based Characterisation of AQUA Peptide VPTV*CVDMSTNRQQ

AQUA peptide VPTV*CVDMSTNRQQ was reduced, alkylated, Zip Tip purified and spotted to a 384 well MALDI plate prior to analysis on an Ultraflex II MS.

As a singly charged peak was previously detected for peptide VPTV*CVDMSTNRQQ, at 1,640.8 m/z , the mass spectrometer was switched to MS/MS mode and CID performed at 6×10^{-6} Torr of Argon. During the MS/MS scan a single peak was found to dominate the spectrum at 864.1 m/z , which in

comparison to a predicted peptide fragmentation pattern (Protein Prospector: Ms-Product), was identified as being the Y₇ ion (Figure 4-16).

VPTVC(Carbamidomethyl) VDMSTNRQQ					VPTV(+6)C(Carbamidomethyl) VDMSTNRQQ				
MH ⁺ (av)		MH ⁺ (mono)			MH ⁺ (av)		MH ⁺ (mono)		
1635.7044		1634.7625			1641.6394		1640.7625		
[-] Main Sequence Ions					[-] Main Sequence Ions				
b		y			b		y		
---	1	V	14	---	---	1	V	14	---
197.1285	2	P	13	1535.6941	197.1285	2	P	13	1541.6941
298.1761	3	T	12	1438.6413	298.1761	3	T	12	1444.6413
397.2445	4	V	11	1337.5936	403.2445	4	V(+6)	11	1343.5936
557.2752	5	C(Carbamidomethyl)	10	1238.5252	563.2752	5	C(Carbamidomethyl)	10	1238.5252
656.3436	6	V	9	1078.4946	662.3436	6	V	9	1078.4946
771.3706	7	D	8	979.4262	777.3706	7	D	8	979.4262
902.4110	8	M	7	864.3992	908.4110	8	M	7	864.3992
989.4431	9	S	6	733.3587	995.4431	9	S	6	733.3587
1090.4907	10	T	5	646.3267	1096.4907	10	T	5	646.3267
1204.5337	11	N	4	545.2790	1210.5337	11	N	4	545.2790
1360.6348	12	R	3	431.2361	1366.6348	12	R	3	431.2361
1488.6934	13	Q	2	275.1350	1494.6934	13	Q	2	275.1350
---	14	Q	1	147.0764	---	14	Q	1	147.0764

Figure 4-16: The predicted fragmentation pattern of both endogenous and synthetic VPTVCVDMSTNRQQ, as predicted by Protein Prospector: MS-Product. The predicted fragmentation pattern of peptide VPTVCVDMSTNRQQ, showing (left) the endogenous peptide and (right) synthetic AQUA peptide fragment pattern.

While it was possible to detect both MS and MS/MS peaks for AQUA peptide VPTV*CVDMSTNRQQ, both were detected with low signal intensities, even when relatively high amounts of sample were analysed (5 pmol). As many endogenous proteins are expressed at femtomolar to atomolar levels within cell lysates (Seibert et al., 2005, He and Chiu, 2003), this MS/MS detection method may struggle to detect anything cruder than an enriched IP. The requirement for MS/MS fragmentation was thus abandoned for peptide VPTV*CVDMSTNRQQ, instead comparing only the parent ion peak heights.

4.4.3.4 MRM Linear Response and LoD on an API 2000

A total of eight dilutions of each AQUA peptide were prepared at amounts ranging from 5 pmol to 100 amol, each of which were reduced and alkylated prior to being injected onto an Ultimate 3000 LC and analysed on an API 2000 MS, using the A-Raf MRM method previously devised.

Contained within the AB SCIEX Analyst 4.2 software bundle is an MRM quantitation algorithm capable of automatically detecting and integrating a peak area for each MRM transition, based on a peak elution time defined by the user. While this may seem useful, in reality this automation can be quite problematic. For example, if the defined peak elution time is off by a matter of seconds, an alternative peak will be selected within the specified time frame, yielding an incorrect quantification. Furthermore, when the peak area is automatically integrated, the baseline alignment is often incorrect, comparing the peak area of half the synthetic peak to the entire endogenous peak, based upon the selection of sub-optimal noise percentage and peak grouping settings. Finally, if even a single ion is detected within the peak elution window, this peak will automatically be selected as the target transition, and as a result, data will always be obtained for an MRM transition, even when no target peak was present.

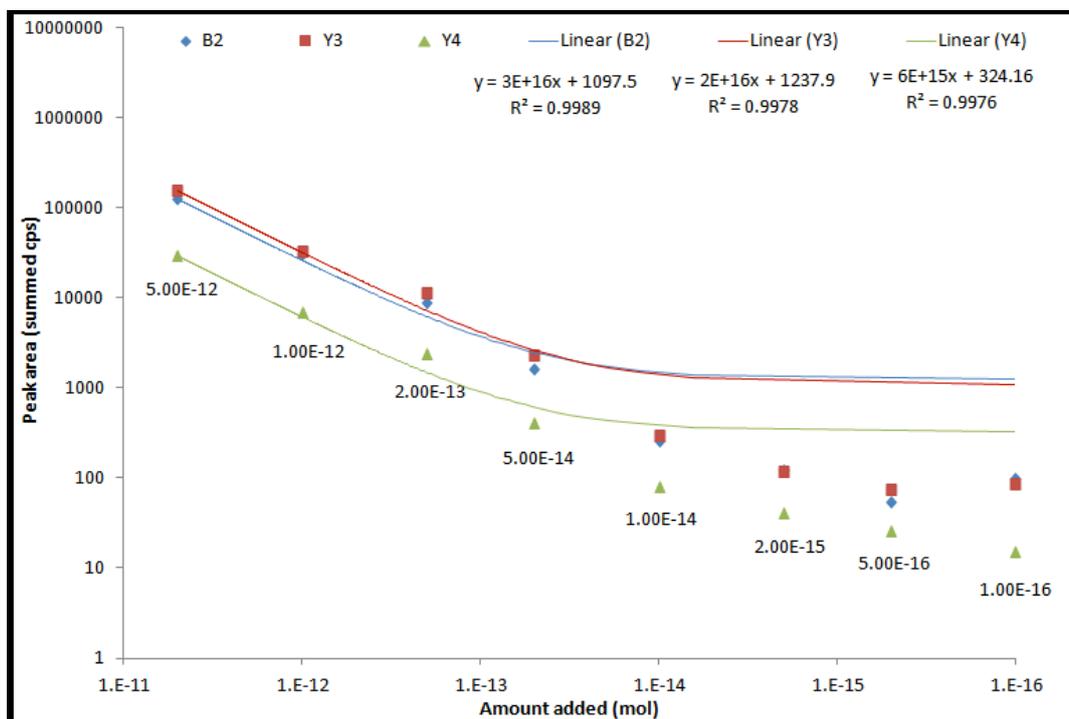


Figure 4-17: The signal response and LoD automatically/incorrectly assigned to AQUA peptide TV*VTVR. The peak areas/signal response obtained when three TV*VTVR transitions were automatically integrated with the IntelliQuan quantitation algorithm, present in Analyst 4.2, at eight different quantities, ranging from 5 pmol to 100 amol, on an API 2000 MS. The data are presented in a logarithm format, plotting the amount of peptide injected against the peak area detected. A linear trend line was fitted to each transition, showing the R^2 values for each equation. In this case the automatic integration has been performed incorrectly, as background noise has continued to be integrated, even when no analyte peak is present.

Figure 4-17 shows the results obtained when eight dilutions of AQUA peptide TV*VTVR were automatically integrated through Analyst 4.2. The peak areas obtain can be seen to level out when only a small amount of peptide was injected, leading to an intercept greater than zero, whereas under a linear response, with an intercept of zero, the peptide should simply fail to be detected. This was due to a minimum peak height not being specified in Analyst, and thus any noise detected within the specified peak elution timeframe was quantified. It is, therefore, important that manual peak integration is performed after each experiment (Figure 4-18), as was recommended by Keshishian *et al.* (Keshishian *et al.*, 2007).

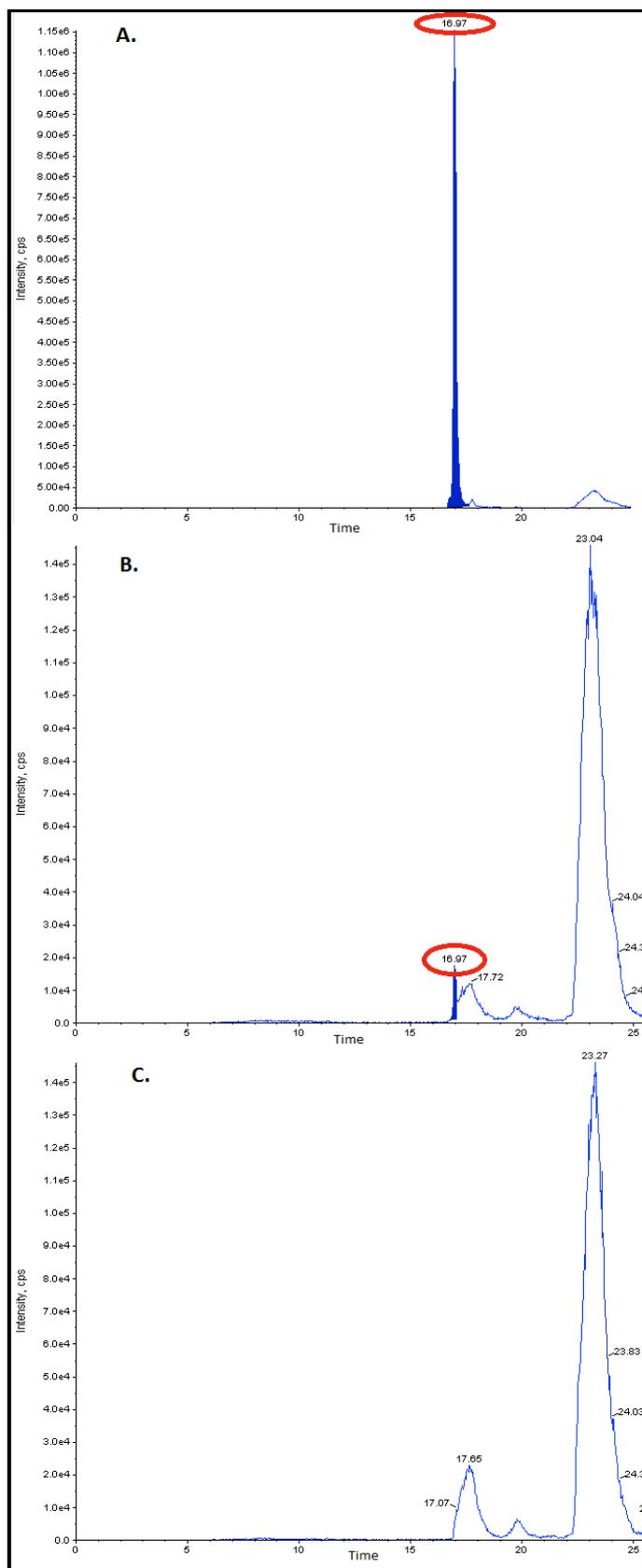


Figure 4-18: How target and synthetic peaks should be manually integrated for each MRM transition. Each MRM peak should be manually integrated, so as to ensure an accurate LoD is assigned. In this case A-Raf AQUA peptide GL*NQDCCVVYR, transition Y₃ is shown, being identified at: **A.** 5 pmol, **B.** 1 pmol, and **C.** 200 fmol on a 5500 Qtrap MS. As the transition was not identified at 200 fmol, then the limit of detection for A-Raf AQUA peptide GL*NQDCCVVYR should be recorded as 1 pmol.

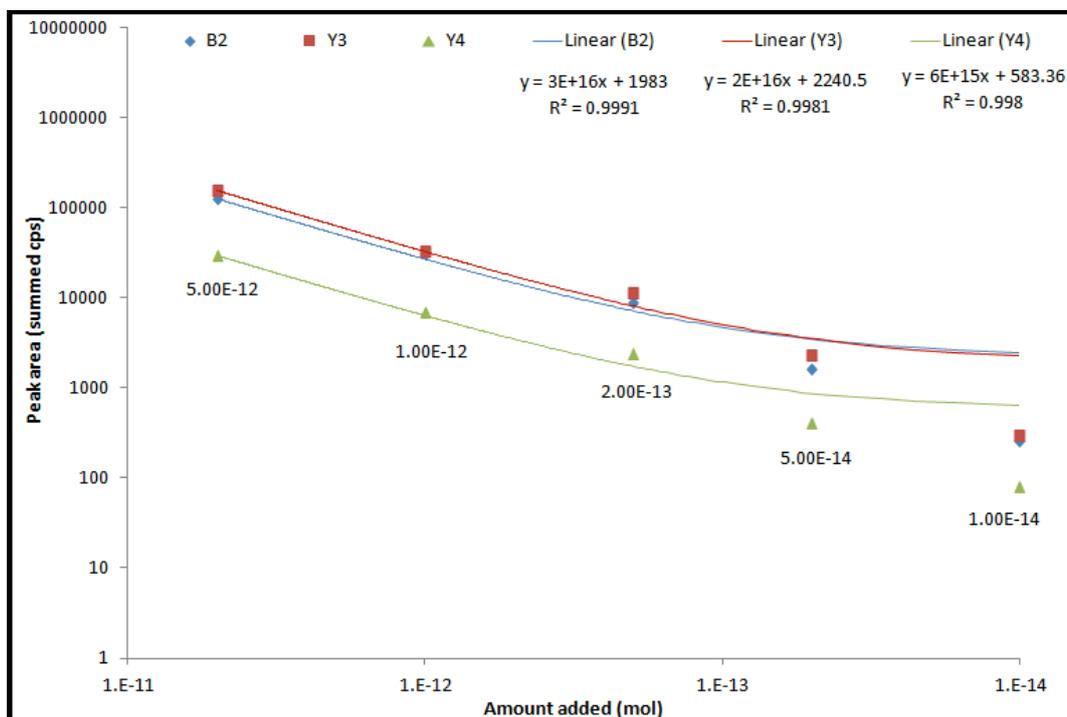


Figure 4-19: The signal response and LoD manually/correctly assigned to A-Raf AQUA peptide TV*VTVR. The peak areas/signal response obtained when three TV*VTVR transitions were manually integrated with the IntelliQuan quantitation algorithm, present in Analyst 4.2, at eight different sample quantities, ranging from 5 pmol to 100 amol, on an API 2000 MS. The data are presented in a logarithm format, plotting the amount of peptide added, against the peak area achieved. A linear trend line was fitted to each transition, showing the R^2 values for each equation. In this case only five of the eight data points were plotted, as no analyte peak was detected for the remaining three.

Through removing the 2 fmol, 500 amol and 100 amol data points, the R^2 values for each remaining linear alignment were seen to increase (Figure 4-19). Despite the data for both tables being integrated differently, the same conclusions can be drawn. A strong linear relationship was detected for the calibration line, down to 50 fmol, after which, the 10 fmol data point was less strongly correlated. This is due to the lower signal to noise ratio (SNR) obtained when only a small quantity of sample is injected, where obtaining an accurate peak area can prove difficult, as small peak integration errors can have a larger effect on the final quantity. As a linear response was confirmed for each A-Raf AQUA MRM transition, the LoD for each peptide was determined to be the lowest quantity at which each of the three Q3 transitions could be detected. For peptide TV*VTVR the LoD was thus set at 10 fmol, for GL*NQDCCVVYR the LoD

was set at 5 pmol, for LLTP*QGPR the LoD was set at 200 fmol and for VSQP*TAEQAQAFK the LoD was set at 1 pmol, assuming detection is performed on an API 2000 MS (Table 4-10).

In situations where a non-linear response is detected, this may offer an indication of the peptide binding the sample vial, a process which will have a greater effect at lower peptide quantities.

Table 4-10: The LoD's achieved for each of the A-Raf MRM transitions on an API 2000. The peak areas obtained when eight different quantities of each A-Raf AQUA peptide, ranging from 5 pmol to 100 amol, were injected. Detailed are the peak areas, where the detected counts per second (cps) values were summed over the width of the peak.

Peptide	Fragment	5 pmol	1 pmol	200 fmol	50 fmol
TV*VTVR	B ₂	7.96E+05	1.52E+05	3.92E+04	8.37E+03
TV*VTVR	Y ₃	3.42E+05	6.71E+04	1.64E+04	3.96E+03
TV*VTVR	Y ₄	1.01E+06	1.88E+05	4.55E+04	9.95E+03
GL*NQDCCVYR	B ₃	7.61E+02	0.00E+00	0.00E+00	0.00E+00
GL*NQDCCVYR	Y ₂	6.33E+03	0.00E+00	0.00E+00	0.00E+00
GL*NQDCCVYR	Y ₃	3.80E+03	1.20E+02	0.00E+00	0.00E+00
LLTP*QGPR	Y ₇	1.48E+04	1.64E+03	6.01E+02	0.00E+00
LLTP*QGPR	Y ₆	1.47E+05	1.84E+04	4.20E+03	0.00E+00
LLTP*QGPR	Y ₅	2.11E+05	2.86E+04	6.58E+03	0.00E+00
VSQP*TAEQAQAFK	Y ₁₀	5.15E+04	1.42E+04	2.36E+03	7.61E+02
VSQP*TAEQAQAFK	Y ₆	1.11E+04	2.88E+03	4.40E+02	0.00E+00
VSQP*TAEQAQAFK	B ₆	3.92E+03	1.60E+03	0.00E+00	0.00E+00
Peptide	Fragment	10 fmol	2 fmol	500 amol	100 amol
TV*VTVR	B ₂	2.72E+03	0.00E+00	0.00E+00	0.00E+00
TV*VTVR	Y ₃	8.01E+02	0.00E+00	0.00E+00	0.00E+00
TV*VTVR	Y ₄	2.76E+03	0.00E+00	0.00E+00	0.00E+00
GL*NQDCCVYR	B ₃	0.00E+00	0.00E+00	0.00E+00	0.00E+00
GL*NQDCCVYR	Y ₂	0.00E+00	0.00E+00	0.00E+00	0.00E+00
GL*NQDCCVYR	Y ₃	0.00E+00	0.00E+00	0.00E+00	0.00E+00
LLTP*QGPR	Y ₇	0.00E+00	0.00E+00	0.00E+00	0.00E+00
LLTP*QGPR	Y ₆	0.00E+00	0.00E+00	0.00E+00	0.00E+00
LLTP*QGPR	Y ₅	0.00E+00	0.00E+00	0.00E+00	0.00E+00
VSQP*TAEQAQAFK	Y ₁₀	0.00E+00	0.00E+00	0.00E+00	0.00E+00
VSQP*TAEQAQAFK	Y ₆	0.00E+00	0.00E+00	0.00E+00	0.00E+00
VSQP*TAEQAQAFK	B ₆	0.00E+00	0.00E+00	0.00E+00	0.00E+00

During a review of the precision and accuracy of MRM based analyses, Addona *et al.* recommended that a LoD is selected at which the faintest of peptide peaks can be detected. Addona *et al.* does however recommend that quantifications are performed at the lower limit of quantification (LLOQ), rather than the LoD, a value approximately three times greater than the LoD (Addona *et al.*, 2009). Based on the MRM data obtained, the lowest level at which each A-Raf isoform

can be quantified is 5 pmol, as set by the LoD for peptide GL*NQDCCVVYR. In regard to the LLOQ, a LoD of 5 pmol is already relatively high, and as such the LoD was retained as the single limiting factor.

4.4.3.5 LoD on a QTrap 4000

When access to a more sensitive triple quadrupole based mass spectrometer became available (QTrap 4000), the LoD for peptide GL*NQDCCVVYR was again tested, injecting the same amount of peptide (Table 4-11).

Table 4-11: The LoD achieved for AQUA peptide GL*NQDCCVVYR on a QTrap 4000 MS. Eight different quantities of alkylated AQUA peptide GL*NQDCCVVYR were injected onto an Ultimate 3000 HPLC, and the peak area detected on a QTrap 4000, using the previously devised A-Raf MRM based acquisition method. Both the peak height and area are measured in cps.

Fragment	Amount	Analyte Peak Area (cps)	Analyte Peak Height (cps)	Analyte SNR
B ₃	1.5 pmol	2.85E+06	1.58E+05	6.72E+04
Y ₂	1.5 pmol	1.23E+07	6.77E+05	3.06E+05
Y ₃	1.5 pmol	9.53E+06	4.98E+05	3.39E+05
B ₃	300 fmol	6.84E+04	3.74E+03	3.04E+03
Y ₂	300 fmol	2.80E+05	1.65E+04	1.52E+04
Y ₃	300 fmol	2.03E+05	1.04E+04	8.72E+03
B ₃	60 fmol	1.47E+04	8.36E+02	1.08E+03
Y ₂	60 fmol	6.38E+04	3.57E+03	2.57E+03
Y ₃	60 fmol	5.21E+04	2.83E+03	1.47E+03
B ₃	15 fmol	2.68E+03	1.05E+02	1.12E+02
Y ₂	15 fmol	4.54E+03	2.25E+02	1.96E+02
Y ₃	15 fmol	3.24E+03	1.99E+02	7.19E+01
B ₃	3 fmol	0.00E+00	0.00E+00	0.00E+00
Y ₂	3 fmol	0.00E+00	0.00E+00	0.00E+00
Y ₃	3 fmol	0.00E+00	0.00E+00	0.00E+00
B ₃	600 amol	0.00E+00	0.00E+00	0.00E+00
Y ₂	600 amol	0.00E+00	0.00E+00	0.00E+00
Y ₃	600 amol	0.00E+00	0.00E+00	0.00E+00
B ₃	150 amol	0.00E+00	0.00E+00	0.00E+00
Y ₂	150 amol	0.00E+00	0.00E+00	0.00E+00
Y ₃	150 amol	0.00E+00	0.00E+00	0.00E+00
B ₃	30 amol	0.00E+00	0.00E+00	0.00E+00
Y ₂	30 amol	0.00E+00	0.00E+00	0.00E+00
Y ₃	30 amol	0.00E+00	0.00E+00	0.00E+00

In terms of instrument specification, the API 2000 is detailed as having a 2,400 Da/sec quadrupole scan time, and an MRM LoD for Reserpine of 50 picograms (pg), with a SNR of less than 800. This equates to a LoD for Reserpine of approximately 82 fmol under ideal conditions. In comparison, the QTrap 4000 is detailed as having a 24,000 Da/sec quadrupole scan time and an MRM LoD for

Reserpine of 200 femtogram (fg), with a SNR of less than 1,200 and a CoV of less than 5%. This equates to a LoD for Reserpine of 330 amol under ideal conditions, 250 fold lower than on the API 2000.

From Table 4-11 it is clear that there is a substantial difference in the LoD achieved for AQUA peptide GL*NQDCCVVYR on the API 2000 and QTrap 4000. While the LoD was originally set to 5 pmol on the API 2000, the smallest quantity of peptide which could be injected and all three GL*NQDCCVVYR transitions detected, the LoD for the QTrap 4000 was set to 15 fmol, over a 300 fold difference, and indeed similar to the LoD changes stated in the instruments specifications. This instrument was, however, not available for use on a regular basis, and as such method development continued on the API 2000.

4.4.3.6 LoD on a QTrap 5500

Following the reduced LoD achieved for peptide GL*NQDCCVVYR on the QTrap 4000 instrument, a QTrap 5500 instrument was tested under the same conditions, injecting the same amount of peptide.

The QTrap 5500 is specified as to have a quadrupole scan speed of up to 20,000 Da/sec and an MRM LoD for Reserpine of 50 fg, with a SNR of less than 2,000, and a CoV of less than 5%. The LoD of Reserpine was thus calculated at approximately 80 amol under ideal conditions, four fold lower than on the QTrap 4000 instrument, and 1000 fold lower than on the API 2000 MS. If this increase in sensitivity were to translate directly to the LoD achieved for AQUA peptide GL*NQDCCVVYR, then the LoD for this A-Raf peptide should fall to just 5 fmol.

Table 4-12: The LoD achieved for AQUA peptide GL*NQDCCVVYR on a QTrap 5500 MS. Eight different quantities of alkylated AQUA peptide GL*NQDCCVVYR were injected onto an Ultimate 3000 HPLC and detected on a QTrap 5500, using the previously devised A-Raf MRM based acquisition method. Both the peak height and area are measured in cps.

Fragment	Amount	Analyte Peak Area (cps)	Analyte Peak Height (cps)	Analyte SNR
B ₃	5 pmol	1.02E+07	1.15E+06	4.65E+03
Y ₂	5 pmol	5.39E+07	5.50E+06	4.21E+05
Y ₃	5 pmol	4.14E+07	4.41E+06	1.24E+05
B ₃	1 pmol	1.21E+05	1.68E+04	5.96E+01
Y ₂	1 pmol	5.95E+05	7.37E+04	4.86E+03
Y ₃	1 pmol	4.52E+05	5.59E+04	1.51E+03
B ₃	200 fmol	0.00E+00	0.00E+00	N/A
Y ₂	200 fmol	2.01E+05	2.10E+04	1.57E+03
Y ₃	200 fmol	1.28E+05	1.35E+04	3.79E+02
B ₃	50 fmol	0.00E+00	0.00E+00	N/A
Y ₂	50 fmol	2.00E+04	2.28E+03	1.62E+02
Y ₃	50 fmol	1.22E+04	1.72E+03	4.59E+01
B ₃	10 fmol	0.00E+00	0.00E+00	N/A
Y ₂	10 fmol	0.00E+00	0.00E+00	N/A
Y ₃	10 fmol	5.96E+03	6.47E+02	1.76E+01
B ₃	2 fmol	0.00E+00	0.00E+00	N/A
Y ₂	2 fmol	0.00E+00	0.00E+00	N/A
Y ₃	2 fmol	0.00E+00	0.00E+00	N/A
B ₃	500 amol	0.00E+00	0.00E+00	N/A
Y ₂	500 amol	0.00E+00	0.00E+00	N/A
Y ₃	500 amol	0.00E+00	0.00E+00	N/A
B ₃	100 amol	0.00E+00	0.00E+00	N/A
Y ₂	100 amol	0.00E+00	0.00E+00	N/A
Y ₃	100 amol	0.00E+00	0.00E+00	N/A

From Table 4-12 it is clear that the improvement in LoD which was detected on the QTrap 4000 did not directly translate to an improvement in LoD on the QTrap 5500, which instead yielded a LoD for peptide GL*NQDCCVVYR of 1 pmol. This increased LoD was, however, most likely due to sub-optimal acquisition method settings, originally designed for use on the API 2000, an older instrument with a significantly different hardware setup. If the declustering potential, collision cell exit potential, ESI voltage and gas flow settings were further optimised on the QTrap 5500, then an increase in sensitivity over the QTrap 4000 may have been achieved. Access to the QTrap 5500 however was limited and thus acquisition method re-optimisation was kept to a minimum during this study.

4.4.3.7 LoD for AQUA peptide VPTV*CVDMSTNRQQ

A total of eight dilutions of AQUA peptide VPTV*CVDMSTNRQQ were prepared at levels ranging from 5 pmol to 100 amol, each sample was reduced and alkylated, reducing the sample a second time, so as to reverse non-specific Met alkylation. The peptide was further reduced in 10% 2-Mercaptoethanol (v/v), so as to completely reduce the Met residue, prior to the sample being spotted onto a 388 well MALDI plate, and the sample analysed on an Ultraflex II MS.

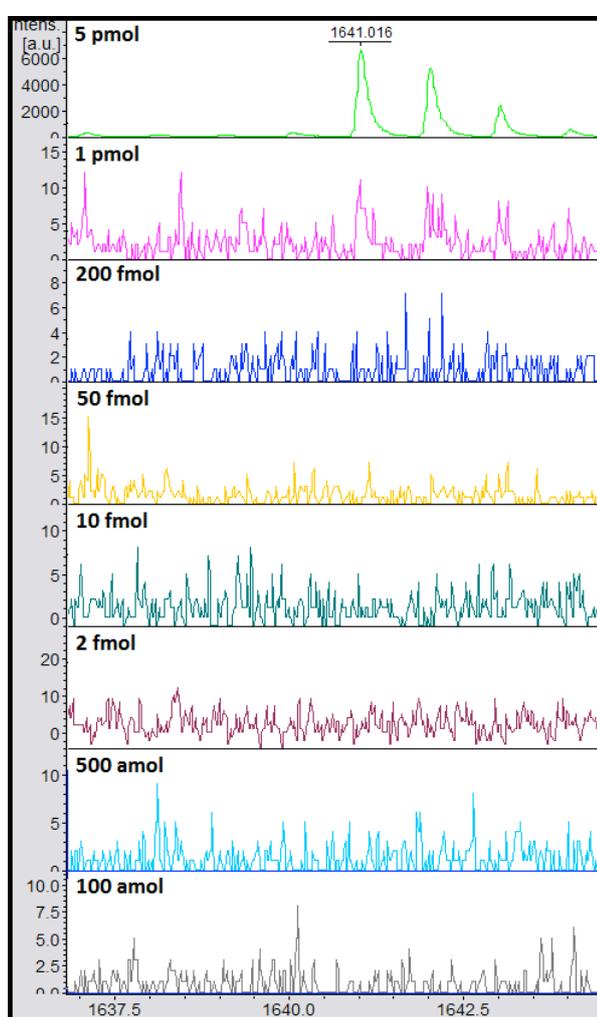


Figure 4-20: The signal response obtained for AQUA peptide VPTV*CVDMSTNRQQ on an Ultraflex II MS, when eight different quantities were injected. Eight dilutions of peptide VPTV*CVDMSTNRQQ, ranging from 5 pmol to 100 amol, were ionised on an Ultraflex II MS, the spectra shown detail the intensities achieved (summed counts) plotted against the m/z .

From Figure 4-20 it is clear that only when 5 pmol of peptide VPTV*CVDMSTNRQQ was ionised, was a strong MS signal response elicited. Peak 1,641 m/z was detected with a signal intensity of 6,600 counts, summed over the 500 laser shots fired, and while a signal response was obtained from 1 pmol of VPTV*CVDMSTNRQQ, the signal intensity reached only 11 counts. The LoD for peptide VPTV*CVDMSTNRQQ was thus set to 1 pmol, which while presenting with a weak signal response, was still the limit at which a detection was possible. Due to the lack of data points below 5 pmol of peptide VPTV*CVDMSTNRQQ, a linear signal response could not be confirmed.

4.4.4 A-Raf Protein Production

While an AQUA peptide was selected for the quantification of each of the four A-Raf isoforms, each AQUA peptide was characterised, and a single MRM based MS acquisition method developed, a stock solution was required for each A-Raf protein isoform, so as to enable the analysis of both the proteolytic digest efficiency of the spiked digest technique, and the suitability of the MRM based acquisition method for the detection and quantification of each of the four A-Raf isoforms, within increasingly complex samples. In terms of the quantity of protein required, the A-Raf peptide with the highest LoD was AQUA peptide GL*NQDCCVYR, which requires some 5 pmol of each isoform, so as to enable a single quantification to be performed. Furthermore, as each quantitation utilises in-gel spiked digestion, then a maximum gel loading volume of 25 μL applied. Thus, for an MS sample to be viable, the quantity of protein within the sample must be between 250 fmol/ μL , assuming 20 μL of sample and 5 μL of protein loading buffer, and 200 fmol/ μL , assuming the sample is pre-suspended in protein loading buffer. While no endogenous concentration data exists for A-Raf, it is clear that a sufficient amount of sample may not be achievable based upon endogenous expression alone. Instead a simple transient transfection was utilised, capable of yielding intracellular protein concentrations as high as 50 mg/L (Durocher et al., 2002).

4.4.4.1 TnT Based Protein Production

As a way of obtaining protein expression levels similar to that of transient transfection, in a solution containing only a limited number of protein species, a TNT Quick Coupled Transcription/Translation reaction was performed. This kit, it is claimed, is capable of producing significantly more protein (2 to 6-fold) in a 60 to 90 minute reaction, than a standard *in vitro* rabbit reticulocyte lysate reaction using RNA templates (Hurst et al., 1996). A TNT based reaction was thus performed for both A-Raf WT and A-Raf Short, the target proteins immunoprecipitated against anti-FLAG, and each isoform sample analysed via SDS-PAGE, first staining the gel with Coomassie, then verifying the expression through the use of a western blot (Figures 4-21 and 4-22).

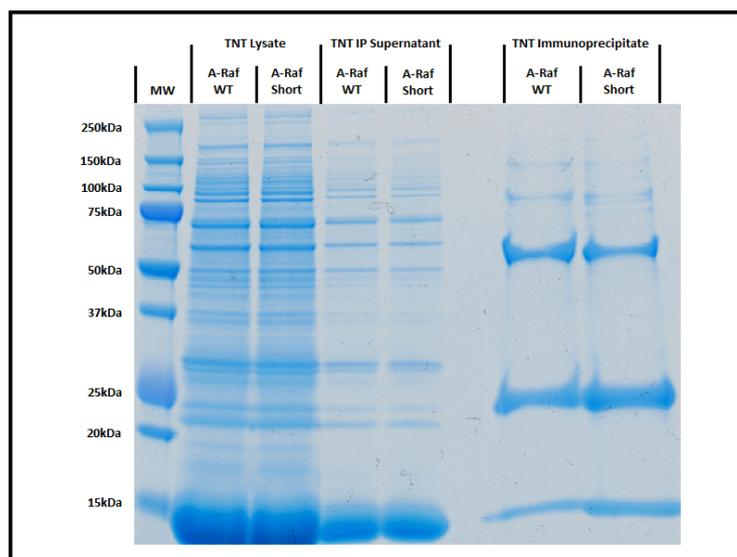


Figure 4-21: A 1D-SDS-PAGE based separation of TNT transfected A-Raf WT and A-Raf Short. A-Raf WT and Short plasmids were subjected to a TNT based transfection, the spent reaction mix was immunoprecipitated against anti-FLAG agarose beads and the immunoprecipitate run on a 1D-SDS-PAGE gel. Shown on the above image is a gel loaded with 2 μ L of each TNT spent reaction mix, 25 μ L of each IP and 25 μ L of each IP supernatant, each visualised through the use of Coomassie Blue protein stain.

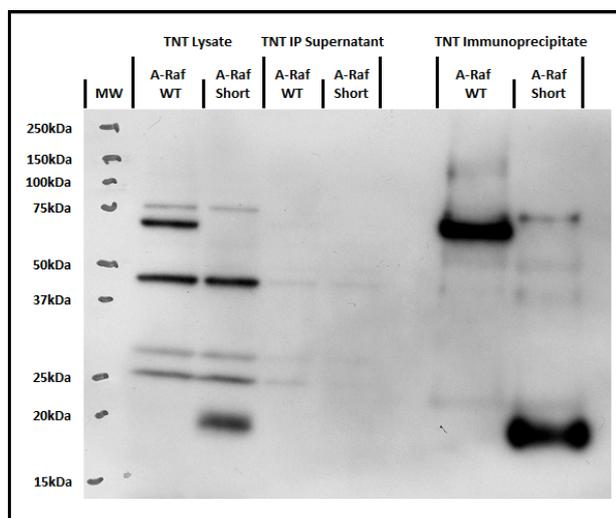


Figure 4-22: A western blot of TNT transfected FLAG-tagged A-Raf WT and A-Raf Short. A-Raf WT and A-Raf Short plasmids were subjected to a TNT based transfection, the spent reaction mix was immunoprecipitated through the use of anti-FLAG agarose beads and the immunoprecipitate run on a 1D-SDS-PAGE gel, prior to being transferred to a PVDF membrane. Shown on this image is a western blot containing 2 μ L of each TNT spent reaction mix, 25 μ L of each IP and 25 μ L of each IP supernatant, each of which were visualised through the use of an anti-FLAG HRP conjugate antibody and BM HRP Chemiluminescence reagent.

From Figure 4-21 it is clear that there were still a high number of proteins present within the spent TNT reaction mix, with no particularly bright bands indicating the presence of the transfected target. The IP supernatant gel lanes also yield the same protein expression patterns as the spent reaction mix, with no Coomassie stained bands obviously extracted. Furthermore, when the IP gel lanes were assessed, no differences could be detected between the A-Raf WT and A-Raf Short lanes, with only the anti-FLAG Ig heavy and light antibody chains detected at a high abundance. This would indicate that either the amount of protein produced was lower than the LoD for Coomassie Blue (50-100 ng (Shevchenko et al., 1996)), even upon IP based enrichment, or that the transfections had failed.

Western blotting, however, did confirm the expression of the A-Raf target proteins, yielding protein bands with a MW of approximately 58 kDa (the average MW for A-Raf WT & FLAG tag) and 19 kDa (the average MW for A-Raf Short & FLAG tag) (Figure 4-22). Furthermore, as the intensity of the

immunoprecipitated protein bands were several times brighter than those of the TNT reaction mix, it can be assumed that the IP was successful, indeed no excess protein was detected in the IP supernatant, suggesting the volume of anti-FLAG beads added to be sufficient.

In conclusion, while it was possible to produce limited amounts of the A-Raf WT and A-Raf Short with the TNT reaction kit, the final protein concentrations were so low that the protein bands could not be visualised through Coomassie staining alone. If a LoD of 100 ng for Coomassie is assumed, then a maximum quantity of 1.4 pmol can be estimated for A-Raf WT, and 5 pmol estimated for A-Raf Short. While this may seem reasonable, this is per reaction, providing enough solution to run a single replicate. As such, when several isoforms and replicates are considered, the reagent costs and labour are increased significantly.

4.4.4.2 HEK293 Based Protein Production

Due to the low expression levels achieved during the TNT reactions, HEK293 based transfections were performed, the protein immunoprecipitated using anti-FLAG, separated via SDS-PAGE and visualised through both Coomassie Blue and western blotting.

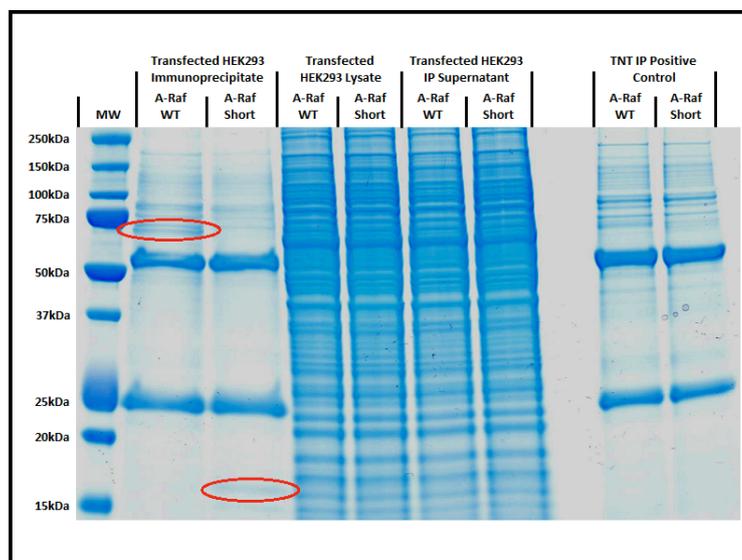


Figure 4-23: A 1D-SDS-PAGE based separation of HEK293 transfected A-Raf WT and A-Raf Short. A-Raf WT and Short plasmids were transfected into HEK293 through the use of Genejuice transfection reagent. The transfected cells were lysed after 72 hours and the transfected protein immunoprecipitated through the use of anti-FLAG M2 agarose beads. A 1D-SDS-PAGE gel was subsequently loaded with 25 μ L of IP, 25 μ L of the IP supernatant and 25 μ L of the TNT IP, which acted as an A-Raf positive control. The protein bands were visualised with Coomassie Blue protein stain. Circled in red are the two transfected A-Raf proteins.

From Figure 4-23 we can see that both A-Raf WT and A-Raf Short were expressed at levels suitable for detection through the use of Coomassie Blue, and thus both were expressed at a level of at least 50 ng/ 25 μ L (approximately 860 fmol/ 25 μ L for A-Raf WT, and 3 pmol/ 25 μ L for A-Raf Short). While this is still below the level required for an accurate detection on an API 2000, it does offer a massive improvement over that achieved with the TNT kit.

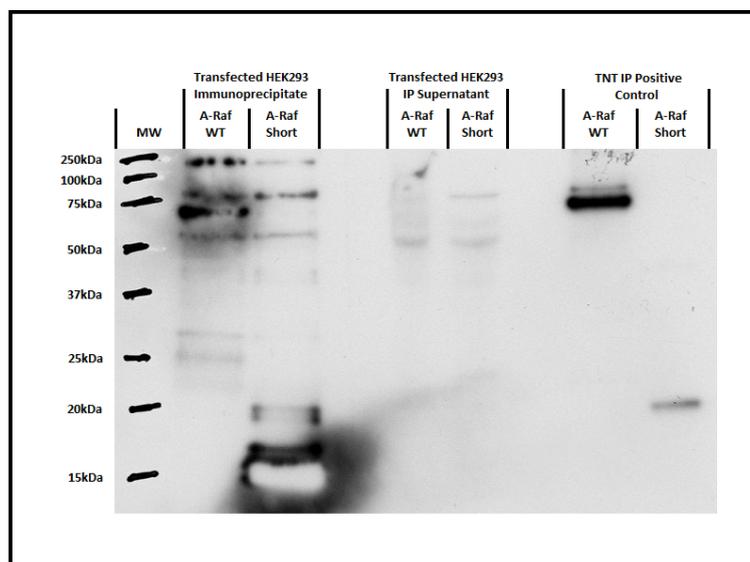


Figure 4-24: A western blot of HEK293 transfected A-Raf WT and A-Raf Short. A-Raf WT and A-Raf Short plasmids were transfected into HEK293 through the use of Genejuice transfection reagent. The HEK293 cells were lysed after 72 hours, and the transfected protein immunoprecipitated with anti-FLAG M2 agarose beads. A 1D-SDS-PAGE gel was loaded with 25 μ L of the IP, 25 μ L of the IP supernatant and 25 μ L of the TNT IP, which acted as an A-Raf positive control. The separated protein was transferred to a PVDF membrane, and the transfected protein bound to an anti-FLAG HRP conjugate antibody. The complex was visualised with BM Chemiluminescence reagent.

Figure 4-24 provides confirmation that the protein bands visualised on the Coomassie Blue stained gel were indeed the FLAG tagged transfected protein targets, expressed at much higher quantity than those achieved with the TNT kit. Also worth noting is the level at which A-Raf Short has been expressed, which is so high that photobleaching has occurred. Photobleaching will occur when all of the HRP substrate has been consumed from the core of the protein band, thus providing no further illumination. Furthermore, the MW of the A-Raf Short protein band appears to have shifted. While it was previously detected at 19 kDa (Figure 4-22), it now has a MW of approximately 15 kDa (Figure 4-24).

As the isoform appeared to have been expressed correctly during the TNT based transfection, it can be assumed that the HEK293 based transfection is responsible, perhaps as a result of incorrect splicing. If the sequence from intron two is being incorrectly spliced, then a MW difference of approximately 4.6 kDa would be expected, while if it were intron four which was being spliced, then a

MW difference of just 1.6 kDa would be expected. Furthermore, if intron four was being spliced out, then the premature stop codon would be missed, producing a much longer isoform, further suggesting intron two.

In conclusion, while there appears to have been an issue with the expression of A-Raf Short, the levels of protein achieved during the HEK293 based transfections were of a suitable level, and far exceed those achieved during the TNT based reactions. Furthermore, the volumes of IP produced mean that many SDS-PAGE replicates can be run per transfection.

4.4.4.3 A-Raf Short Incorrect Splicing

4.4.4.3.1 Evidence Supporting Incorrect Splicing

To further investigate the possible incorrect splicing of A-Raf Short, the A-Raf Short transfected HEK293 IP was loaded to, and separated on, an SDS-PAGE gel, the appropriate band excised, and in-gel digested with trypsin. The peptide solution was injected onto an Ultimate 3000 LC and analysed on an API 2000 MS, running the previously devised A-Raf MRM acquisition method.

As peptide TVVTVR crossed the splice site of intron two, it was hoped that through assessing the A-Raf Short IP peptideome for the presence of this sequence, that its detection would indicate that intron two was being incorrectly spliced. Based on the MW difference detected during the A-Raf Short transfections (Figures 4-22 and 4-24), which accounts for almost 25% of the proteins total MW, this was the only logical variation in sequence.

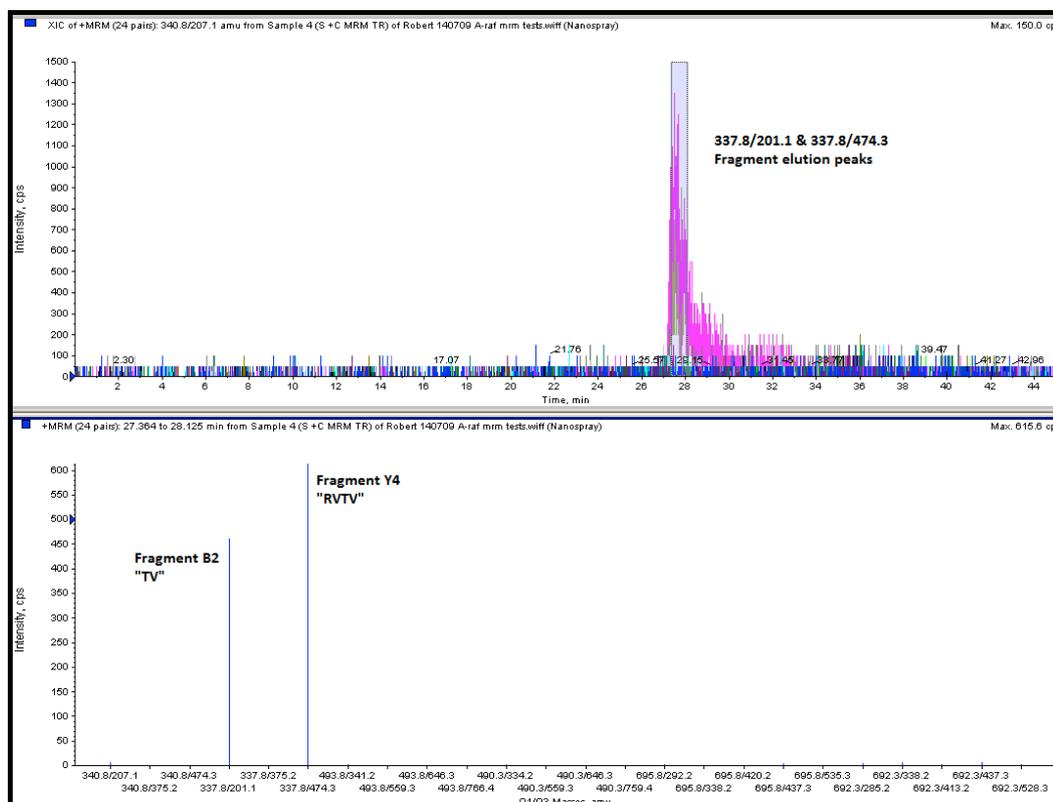


Figure 4-25: An MRM based detection of incorrectly spliced A-Raf Short, obtained from a HEK293 based transfection. The spectrum obtained from an MRM based detection of incorrectly spliced A-Raf Short, obtained from a HEK293 based transfection and analysed on an API 2000 MS, using the previously devised A-Raf MRM acquisition method. Shown on the top window is an XIC spectrum detailing the time at which peptide TVVTVR was detected. Shown on the bottom MRM spectrum meanwhile are the two TVVTVR Q3 fragment ions (B₂ and Y₄), thereby confirming the detection of peptide TVVTVR.

Figure 4-25 indicates that the main protein product produced during the HEK293 transfection of A-Raf Short was indeed missing the sequence from intron two, as is indicated by the presence of intron two crossing peptide TVVTVR. These MRM transitions were of a high intensity, discounting the detection of noise and spectral artefacts, and as two of the three TVVTVR MRM transitions were detected, the presence of peak TVVTVR was confirmed.

4.4.4.3.2 Modifying the Splice Site of Intron Two via Site Directed Mutagenesis

As the HEK293 based transfections were found to yield adequate levels of A-Raf protein expression, it was decided that the A-Raf Short plasmid should be modified in such a way so as to prevent the incorrect splicing of intron two. The full details of site directed mutagenesis, via splice overlap extension PCR (SOE), have previously been described by Ho *et al.* (Ho *et al.*, 1989), in brief however, a splice site was selected where a single base pair could be modified, maintaining the sequence of A-Raf Short, but disrupting the splice sites bordering intron two, preventing the intron from being removed.

The A-Raf Short plasmid was sequenced with CMV and T7 forward primers, located upstream of the initiation of transcription, and the BGH reverse primer, located downstream of the termination of transcription. The sequence data obtained for each primer was aligned using Clustal X and screened for the presence of the 3' and 5' splice sites and the internal branch point. While the 3' and 5' splice sites were easily identified, with the introns 5' beginning GT, of conserved sequence MAG|GTRAGT (where M represents either an A or a C and R representing either an A or a G), and the introns 3' ending AG, of conserved sequence CAG|G (Reed, 1989), the branch point was harder to identify. While the branch point is usually marked by a single adenosine residue, located some 18-40 bases downstream of the 5' DNA splice site, several adenosine residues were identified, any of which may have been acting as the branch point and as such a branch point based mutation was rejected. Of the two remaining mutation sites, the 5' splice site was also found to be unsuitable. If either the G or T from consensus sequence GT were modified, to prevent splicing, then this would ultimately lead to a change in the proteins sequence, replacing the Val with one of several other residues. The modification of the 3' splice site appeared more promising however, where changing the G from consensus sequence AG retained the Gln residue, while preventing splicing (Crick, 1968).

So as to enable site directed mutagenesis to be performed on the last base of intron two, 40 mer complementary oligonucleotides were created, where base G

was replaced with an A, while providing ample sequence to either side of the mutation point, so as to enable annealing (Figure 4-26).

```

                                T7 promoter                                NheII
>CMV For      CTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGCC
>BGH Rev      CTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGCC
                Initiation of transcription

>CMV For      GCCACCATTGGATTACAAGGATGACGACGATAAAGGAGCCACCACGGGGCCCCCTGCCAAT
>T7           -----AGGATGACGACGATAAAGGAGCCACCACGGGGCCCCCTGCCAAT
>BGH Rev      GCCACCATTGGATTACAAGGATGACGACGATAAAGGAGCCACCACGGGGCCCCCTGCCAAT
>Protein      M D Y K D D D M/K E P P R G P P A N

>CMV For      GGGGCCGAGCCATCCCGGGCAGTGGGCACCGTCAAAGTATACTGCCCAACAAGCAACGC
>T7           GGGGCCGAGCCATCCCGGGCAGTGGGCACCGTCAAAGTATACTGCCCAACAAGCAACGC
>BGH Rev      GGGGCCGAGCCATCCCGGGCAGTGGGCACCGTCAAAGTATACTGCCCAACAAGCAACGC
>Protein      G A E P S R A V G T V K V Y L P N K Q R
                Border or exon 2/intron 2

>CMV For      ACGGTGTGAGTCATGGAAGCGAAATGGCAGGGGCTGTGGATGGACCCAGTTGTAACCTCT
>T7           ACGGTGTGAGTCATGGAAGCGAAATGGCAGGGGCTGTGGATGGACCCAGTTGTAACCTCT
>BGH Rev      ACGGTGTGAGTCATGGAAGCGAAATGGCAGGGGCTGTGGATGGACCCAGTTGTAACCTCT
>Protein      T V V S H G S E M A G A V D G P S C N S

>CMV For      GGGATCAAAAGGGTGACAACGGTTGGGGGAGGCCTTTCAGAAAGGATGGGAACATCAGCT
>T7           GGGATCAAAAGGGTGACAACGGTTGGGGGAGGCCTTTCAGAAAGGATGGGAACATCAGCT
>BGH Rev      GGGATCAAAAGGGTGACAACGGTTGGGGGAGGCCTTTCAGAAAGGATGGGAACATCAGCT
>Protein      G I K R V T T V G G G L C R R M G T S A
                Border of intron 2/exon 3

>ARAFMUTSAF  CCTACATCTGCACATACACACAAGTGACTGTCCGG
>CMV For      CCGCTTCTGTGGGCATTGAGGACCCCTACATCTGCACATACACACAGGTGACTGTCCGG
>T7           CCGCTTCTGTGGGCATTGAGGACCCCTACATCTGCACATACACACAGGTGACTGTCCGG
>BGH Rev      CCGCTTCTGTGGGCATTGAGGACCCCTACATCTGCACATACACACAGGTGACTGTCCGG
>Protein      A L L L G I E D P Y I C T Y T Q V T V R

>ARAFMUTSAF  GATGCC
>CMV For      GATGGCATGAGTGTCTACGACTCTCTAGACAAGGCCCTGAAGGTGCGGGGTCTAAATCAG
>T7           GATGGCATGAGTGTCTACGACTCTCTAGACAAGGCCCTGAAGGTGCGGGGTCTAAATCAG
>BGH Rev      GATGGCATGAGTGTCTACGACTCTCTAGACAAGGCCCTGAAGGTGCGGGGTCTAAATCAG
>Protein      D G M S V Y D S L D K A L K V R G L N Q

>CMV For      GACTGCTGTGTGGTCTACCGACTCATCAAGGGACGAAAGACGGTCACTGCCTGGGACACA
>T7           GACTGCTGTGTGGTCTACCGACTCATCAAGGGACGAAAGACGGTCACTGCCTGGGACACA
>BGH Rev      GACTGCTGTGTGGTCTACCGACTCATCAAGGGACGAAAGACGGTCACTGCCTGGGACACA
>Protein      D C C V V Y R L I K G R K T V T A W D T

>CMV For      GCCATTGCTCCCCTGGATGGCGAGGAGCTCATTGTCGAGGTCTTGAAGATGTCCCCTG
>T7           GCCATTGCTCCCCTGGATGGCGAGGAGCTCATTGTCGAGGTCTTGAAGATGTCCCCTG
>BGH Rev      GCCATTGCTCCCCTGGATGGCGAGGAGCTCATTGTCGAGGTCTTGAAGATGTCCCCTG
>Protein      A I A P L D G E E L I V E V L E D V P L
                Border of exon 4/intron 4

>CMV For      ACCATGCACAATTTGTGAGTGCAGGGTGGACGGTGGGGGTGGACCATGGTTGGGGGTGT
>T7           ACCATGCACAATTTGTGAGTGCAGGGTGGACGGTGGGGGTGGACCATGGTTGGGGGTGT
>BGH Rev      ACCATGCACAATTTGTGAGTGCAGGGTGGACGGTGGGGGTGGACCATGGTTGGGGGTGT
>Protein      T M H N F V S A G W T V G V D H G W G C

                BamHI
>CMV For      CCTTAGGGATCCACTAGTCCAGTGTGGTGAATTCTGCAGATATCCAGCACAGTGGCGGC
>T7           CCTTAGGGATCCACTAGTCCAGTGTGGTGAATTCTGCAGATATCCAGCACAGTGGCGGC
>BGH Rev      CCTTAGGGATCCACTAGTCCAGTGTGGTGAATTCTGCAGATATCCAGCAC-----
>Protein      P *

                BGH Reverse promoter
>CMV For      CGCTCGAGTCTAGAGGGCCCGTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTT
>T7           CGCTCGAGTCTAGAGGGCCCGTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTT

```

Figure 4-26: The 40 mer oligonucleotide selected for the site directed mutagenesis of A-Raf Short. A Clustal X based alignment of the A-Raf Short coding sequence from the A-Raf Short plasmid, showing the intronic sequence unique to A-Raf Short (highlighted in blue) and the sequence common to all A-Raf isoforms (highlighted in green). This alignment was constructed from the sequence obtained when the A-Raf Short plasmid was sequenced with T7 and CMV forward promoters and BGH reverse promoter (highlighted in red). Based upon this genetic map, a 40 mer site directed mutagenesis primer (highlighted in purple) was designed to change a single base pair within the A-Raf Short plasmid sequence (highlighted in brown), and in doing so preventing intron two from being spliced out. Also shown on the alignment are the restriction digest sites (highlighted in yellow) used to cleave the SOE PCR strands, and the consensus protein sequence for A-Raf Short.

Having performed site directed mutagenesis on A-Raf Short, through SOE based PCR, the mutated double stranded DNA insert was restriction digested with *NheII* (upstream) and *BamHI* (downstream), prior to agarose gel based purification. The digested insert was ligated into pcDNA3.1(+) with T4 DNA ligase, where the vector had been pre-digested with the same two restriction enzymes. The ligated plasmid was cloned into library efficiency DH5a cells and plated on agar. Subsequently each colony was picked and amplified, prior to a miniprep based plasmid purification. The amplified DNA was restriction digested with *NheII* and *BamHI*, separated via agarose gel based electrophoresis, and screened for the presence of a suitably sized plasmid insert (555 bp).

Where a colony was identified as containing a mutated insert, some 500 ng of plasmid was sent for sequencing, so as to ensure that only the intended base was modified, the results from which are shown on Figure 4-27.

```

                                T7 20mer                                Nhe11
mutant      CTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGCC
unmodified  CTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGCC
*****

mutant      GCCACCATGGATTACAAGGATGACGACGATAAAGGAGCCACACGGGGCCCCCCTGCCAAT
unmodified  GCCACCATGGATTACAAGGATGACGACGATAAAGGAGCCACACGGGGCCCCCCTGCCAAT
*****

mutant      GGGGCAGAGCCATCCCGGGCAGTGGGCACCGTCAAAGTATACCTGCCCAACAAGCAACGC
unmodified  GGGGCAGAGCCATCCCGGGCAGTGGGCACCGTCAAAGTATACCTGCCCAACAAGCAACGC
*****

mutant      ACGGTGGTGAGTCATGGAAGCGAAATGGCAGGGGCTGTGGATGGACCCAGTTGTAACCTCT
unmodified  ACGGTGGTGAGTCATGGAAGCGAAATGGCAGGGGCTGTGGATGGACCCAGTTGTAACCTCT
*****

mutant      GGGATCAAAAGGGTGACAACGGTTGGGGGAGGCCTTGCAGAAGGATGGGAACATCAGCT
unmodified  GGGATCAAAAGGGTGACAACGGTTGGGGGAGGCCTTGCAGAAGGATGGGAACATCAGCT
*****

                                A-Raf short intron 2 5' border
mutant      GCGCTTCTGTGGGCATTGAGGACCCTACATCTGCACATACACACAAGTGACTGTCCGG
unmodified  GCGCTTCTGTGGGCATTGAGGACCCTACATCTGCACATACACACAAGTGACTGTCCGG
*****

mutant      GATGGCATGAGTGTCTACGACTCTCTAGACAAGGCCCTGAAGGTGCGGGGTCTAAATCAG
unmodified  GATGGCATGAGTGTCTACGACTCTCTAGACAAGGCCCTGAAGGTGCGGGGTCTAAATCAG
*****

```

Figure 4-27: The genetic sequence obtained from the A-Raf Short plasmid, both before and after site directed mutagenesis. An alignment performed on the sequence obtained from the A-Raf Short plasmid both before and after site directed mutagenesis, based upon the use of CMV forward. Highlighted in Yellow is the T7 forward sequence, while in red is the upstream Nhe1 restriction digest site. Finally, in green is the 40 mer site directed mutagenesis oligonucleotide, with the single modified nucleotide highlighted in red.

4.4.4.3.3 Confirming the Retention of Intron Two

Both the mutated and non-mutated A-Raf Short plasmids were amplified via Midiprep and transfected into HEK293 with Genejuice transfection reagent. The A-Raf Short proteins were immunoprecipitated, separated via SDS-PAGE, and visualised via western blot.

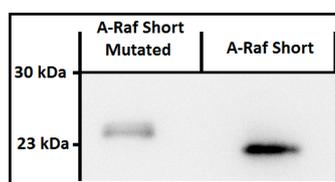


Figure 4-28: A western blot showing the MW difference between HEK293 transfected A-Raf short, before and after the completion of site directed mutagenesis. Shown on this WB are the protein bands obtained from transfected A-Raf Short, before and after site directed mutagenesis. The protein bands shown here were visualised through the use of an anti-FLAG HRP conjugate antibody and BM Chemiluminescence reagent.

From Figure 4-28 we can see that the MW of the mutated A-Raf Short protein increased, in comparison to that produced by the non-mutated plasmid. As to further prove the success of this experiment, a spiked digest was performed on A-Raf Short, identifying the expression of peptide GLNQDCCVVYR alone.

4.4.5 Spiked Digest Based Quantitative Analyses

In interfacing the separative capacity of SDS-PAGE with the analytical competence of mass spectrometry, the extraction of large polypeptides and proteins has proven problematic. The method of choice for protein recovery has instead been to perform a proteolytic digestion within the gel, eluting the peptide fragments for a bottom-up based MS analysis (Cohen and Chait, 1997). Problems have however been described during peptide extraction, with only 70-80% of the total proteolytic peptides generally being recovered. Furthermore post-recovery sample handling has been known to further reduce the amount of peptide by 10% to 15%, assuming minimal handling, and up to 50% with vacuum centrifugation based techniques (Speicher et al., 2000).

While sample losses are inevitable, AQUA aims to minimise the effect of these losses on the precision of the quantitation through adding the peptide standards directly to the dehydrated gel pieces, before the protease has been applied. In doing so, the synthetic peptides should be fully absorbed into the gel pieces and thus will be subject to the same extraction bias as the target peptides. Furthermore, any sample loss encountered during downstream processing should affect both peptides equally (Gerber et al., 2003a).

4.4.5.1 Single Isoform Based Spiked Digests

So as to assess the efficiency with which each of the target peptides were digested (through comparing peak area ratios for each target peptide/protease, each of which should be present at an equal ratio), to identify any problems resulting from the re-suspension and modification of each AQUA peptide (which would result in a specific peptide presenting with a different peak area ratio),

and to test the AQUA based absolute quantification workflow; 25 μL of each A-Raf IP (obtained from HEK293 immunoprecipitates) was loaded onto an SDS-PAGE gel and electrophoresed. The A-Raf isoform gel bands were excised and in-gel digested in the presence of each AQUA peptide. Following peptide extraction the samples were loaded onto an Ultimate 3000 LC and analysed on an API 2000 MS, running the A-Raf MRM acquisition method previously devised.

If we consider that each transfected plate contains some 1.1×10^8 HEK293 cells, and that six plates were transfected per isoform, then a total of 6.6×10^8 cells were lysed, and the transfected protein immunoprecipitated. If we then consider that the anti-FLAG beads were re-suspended in 360 μL of protein loading buffer, then the 25 μL of IP loaded to each SDS-PAGE gel well should contain the transfected, immunoprecipitated contents of approximately 4.6×10^7 HEP293 cells.

4.4.5.1.1 A-Raf WT Single Isoform Spiked Digest

Figure 4-29 shows the extracted ion chromatogram (XIC) obtained from the spiked digestion based analysis of A-Raf WT via LC-MS on an API 2000. From the spectrum four main peaks can be seen to elute, the first of which was peptide TVVTVR, eluting after approximately 17 minutes with a high signal intensity. The second peptide to elute was GLNQDCCVVYR, after approximately 21 minutes, presenting with a low signal intensity, almost entirely masked by the elution of peptide LLTP*QGPR, which also elutes after 21 minutes. Finally peptide VSQP*TAEQAQAFK eluted after 22 minutes, again almost entirely masked by the high signal intensity of peptide LLTP*QGPR.

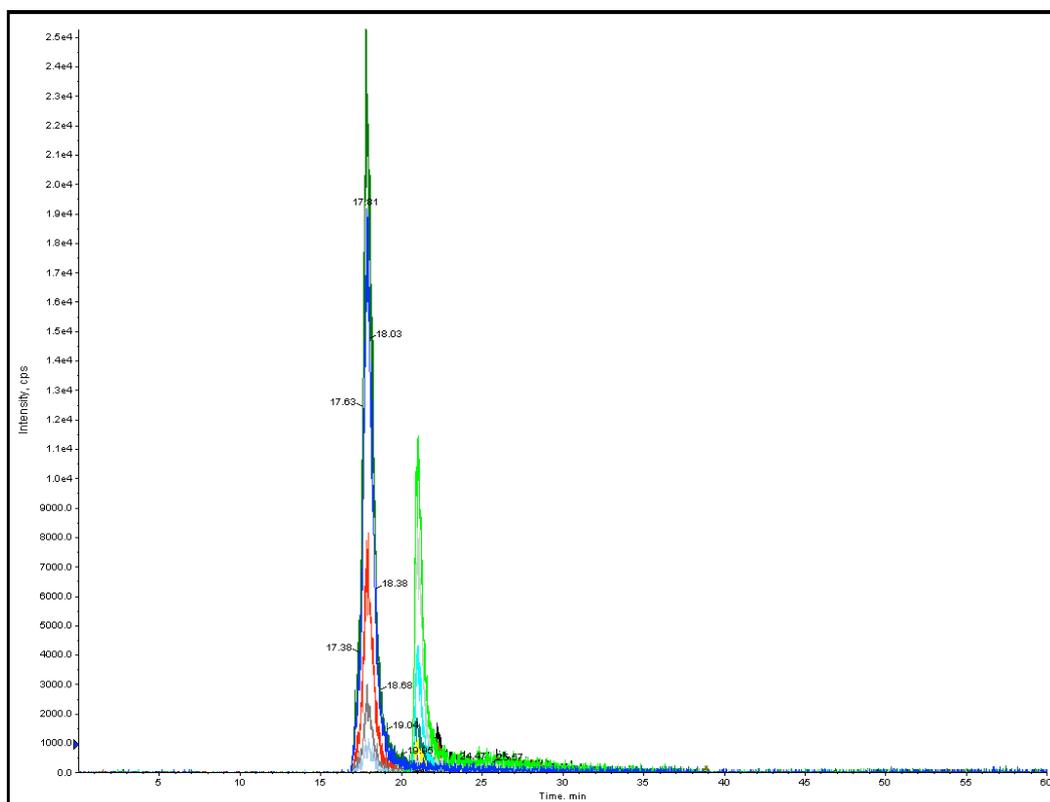


Figure 4-29: An XIC obtained from the spiked digestion of A-Raf WT on an API 2000. Through digesting 25 μ L of immunoprecipitated A-Raf WT with trypsin, in the presence of 5 pmol of TV*VTVR, GL*NQDCCVVYR, VSQP*TAEQAQAFK and LLTP*QGPR, a spiked digest solution was obtained which was injected onto an Ultimate 3000 and analysed on an API 2000 MS.

It is also worth noting the shape of the eluting peaks, each of which suffers from tailing. This would indicate that the peptides were being retained based upon several modes of retention. To better explain, while RP-LC retains and separates molecules based upon their hydrophobic interactions with the stationary phase (C18), residual polar interactions between the peptides and the silanol groups on the PepMap silica based particles may also have been occurring.

The quantification data for the nine A-Raf WT replicates was manually integrated with Analyst 4.2, both the average quantity of peptide detected (averaging three MRM transitions per peptide) and a CoV value were calculated for each peptide (Table 4-13).

Table 4-13: The quantity and variation data obtained from the spiked digestion of 25 μL of immunoprecipitated A-Raf WT on an API 2000. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

Replicate	TVVTVR - Present in all but A-Raf Short		VSQPTAEQAQAFK - Unique to A-Raf WT		GLNQDCCVVYR - Present in all isoforms	
	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV
1	0.65	2.97%	1.06	34.25%	1.38	15.44%
2	0.67	1.65%	0.97	30.72%	1.14	18.48%
3	0.66	7.36%	1.13	55.99%	0.99	16.02%
4	0.69	2.31%	0.95	24.13%	1.27	16.73%
5	0.64	7.46%	0.86	20.59%	1.19	16.89%
6	0.72	3.45%	1.06	31.78%	1.39	29.60%
7	0.66	0.71%	1.08	15.79%	1.24	18.37%
8	0.71	0.81%	1.21	17.20%	0.83	14.20%
9	0.71	11.02%	1.28	8.47%	1.33	12.24%
Average	0.68	5.62%	1.07	27.39%	1.2	21.69%

While the amount of TVVTVR appears to remain constant throughout the run (Table 4-13 & Figure 4-30), with a averaged CoV of just 5.62%, the amount of peptides GLNQDCCVVYR and VSQPTAEQAQAFK detected per sample appear more variable, with averaged CoVs of 21.69% and 27.39%, respectively. This variation may have been due to the estimated amount of A-Raf WT within the 25 μL of HEK293 IP being between 600 fmol (based upon the quantification data obtained for peptide TVVTVR) and 1.2 pmol (based upon the quantification data obtained for peptide GLNQDCCVVYR), below the 5 pmol LoD assigned to peptide GLNQDCCVVYR, and just achieving the 1 pmol LoD assigned to peptide VSQPTAEQAQAFK.

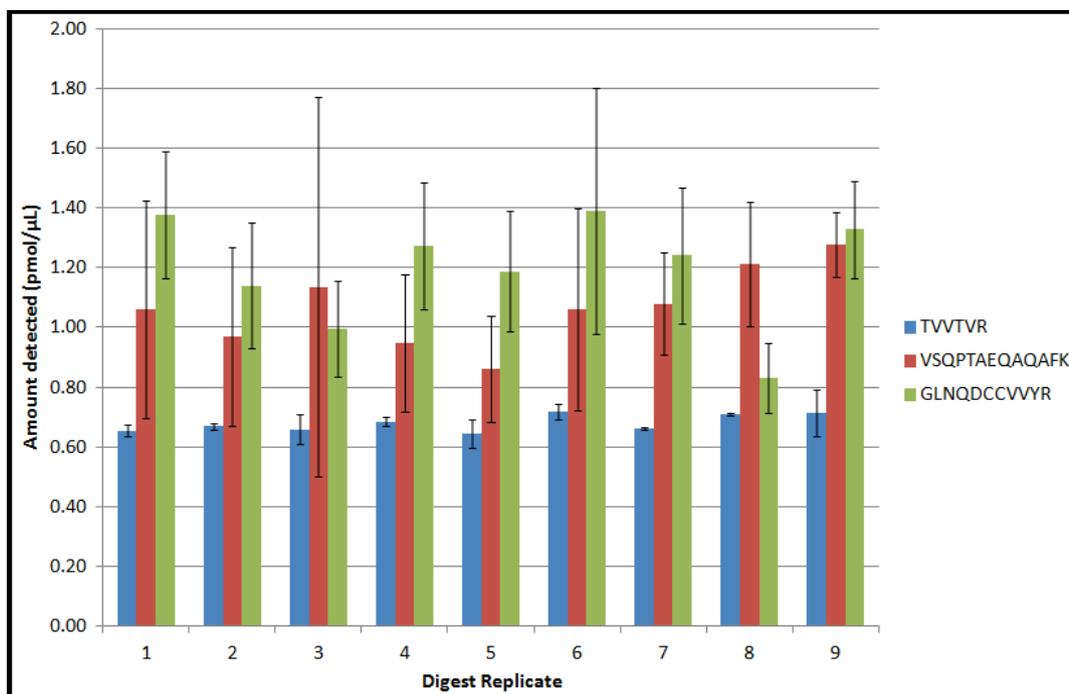


Figure 4-30: The quantities of each A-Raf WT target peptide detected during the spiked digestion of A-Raf WT transfected HEK293 IP on an API 2000. The amount of each A-Raf WT target peptide detected during the spiked digestion of 25 μ L of A-Raf WT transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between each peptide and replicate.

An analysis of variance (ANOVA) based statistical analysis was performed on this data, so as to assess if the means of the quantitations achieved for each peptide were equal, assuming a normal distribution. Therefore, during this test, the null hypothesis (H_0) stated that the means of these groups were equal. This analysis was performed based upon an alpha level of 0.05 (95% confidence level), giving a critical F-value (the variance between groups/the overall variance) of 3.4, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-14, an F-value of 37.5 was achieved, therefore the H_0 was rejected, indicating the mean quantification values achieved for these peptides to be significantly different.

Table 4-14: An ANOVA based statistical analysis performed on the data obtained from the spiked digest of A-Raf WT transfected HEK293 IP. The quantification data obtained during the spiked digestion of 25 μ L of A-Raf WT transfected HEK293 IP was subjected to ANOVA. During this statistical analysis, the data obtained for each targeted A-Raf WT peptide, and each digest replicate were tested, so as to assess if the means obtained for each peptide were equal. Highlighted in yellow is the F-value, which if above 37.5, would indicate the data to be significantly different, in 95% of cases.

Groups	Count	Sum	Average	Variance	F-value
TVVTVR	9	6.110681	0.678965	0.000783	37.45343
VSQPTAEQAQAFK	9	9.607307	1.067479	0.017092	
GLNQDCCVVYR	9	10.76608	1.196231	0.0344	

In conclusion, the data obtained from the spiked digestion of 25 μ L of A-Raf WT transfected HEK293 IP was found to be inaccurate, containing significant levels of variation between peptides. This may have resulted from two of the three A-Raf WT peptides being below their LoDs within the IP.

4.4.5.1.2 A-Raf Short Single Isoform Spiked Digest

Twenty five microlitres of A-Raf Short transfected HEK293 IP was digested with trypsin in the presence of each of the A-Raf AQUA peptides. Three transitions were compared, per peptide, and the data averaged. A total of nine replicates were run, the data from which is shown on Table 4-15.

Table 4-15: The quantity and variation data obtained from the spiked digestion of 25 μ L of immunoprecipitated A-Raf Short on an API 2000. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

GLNQDCCVVYR - Present in all isoforms		
Replicate	Average per Replicate (pmol/ μ L)	CoV
1	3.47	16.22%
2	3.84	14.42%
3	4.03	6.98%
4	3.09	15.21%
5	4.54	2.89%
6	3.40	4.06%
7	3.42	11.78%
8	3.36	12.28%
9	5.25	7.23%
Average	3.82	19.42%

As only one target peptide was present in A-Raf Short (peptide GLNQDCCVVYR), a comparison could not be made between peptide quantitation values. From Table 4-15, it is clear that the estimated amount of peptide GLNQDCCVVYR detected in the 25 μ L of A-Raf Short transfected HEK293 IP was variable between samples, despite the level of A-Raf Short being above the LoD. This would suggest that the variation detected for peptide GLNQDCCVVYR during the spiked digestion of A-Raf WT transfected HEK293 IP may not merely have been due to the low protein expression levels (Figure 4-31), but perhaps was due to the way in which alkylated peptide GLNQDCCVVYR was prepared.

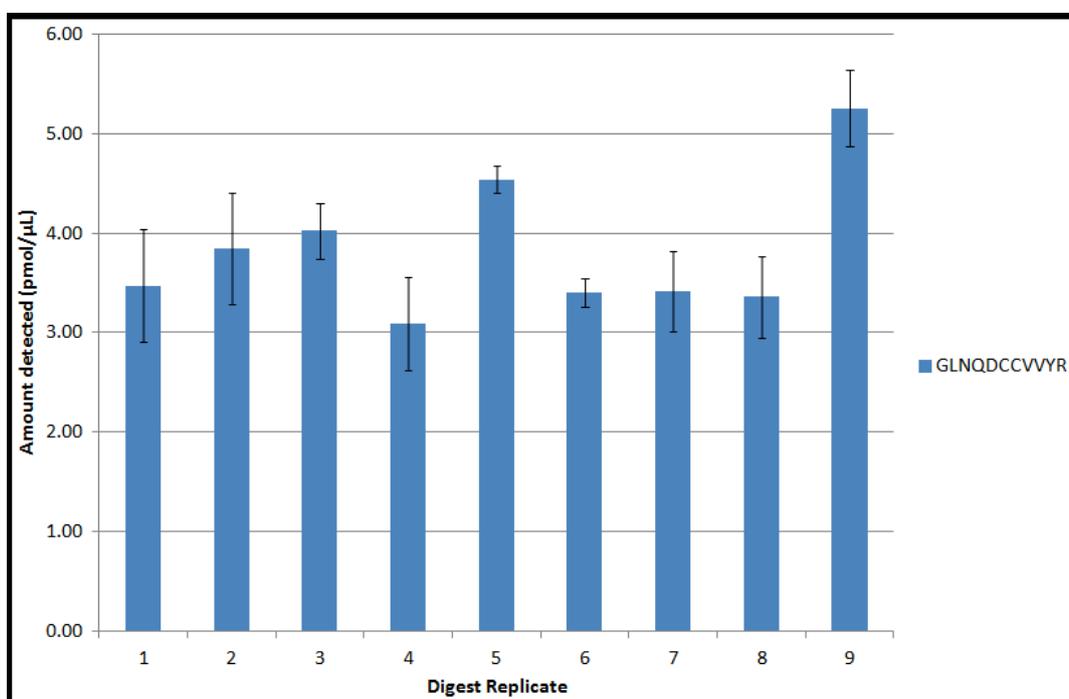


Figure 4-31: The quantity of peptide TVVTVR, detected during the spiked digestion of A-Raf Short transfected HEK293 IP on an API 2000. The amount of peptide TVVTVR detected during the spiked digestion of 25 μ L of A-Raf Short transfected HEK293 IP is shown. In brief, this single quantity was calculated through comparing the peak areas for three MRM transitions per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between each peptide and replicate.

4.4.5.1.3 DA-Raf-1 Single Isoform Spiked Digest

The quantification of DA-Raf-1 required the digestion of this truncated A-Raf isoform with either trypsin or Glu-C, separately, presenting an opportunity for

the digest efficiencies of these proteases to be compared. The data obtained from each of the digests is shown on Table 4-16.

Table 4-16: The quantity and variation data obtained from the spiked digestion of 25 μ L of immunoprecipitated DA-Raf-1 on an API 2000. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

Replicate	LLTPQGPR - Unique to DA-Raf-1		TVVTVR - Present in all but A-Raf Short		GLNQDCCVYR - Present in all isoforms	
	Average per Replicate (pmol/ μ L)	CoV	Average per Replicate (pmol/ μ L)	CoV	Average per Replicate (pmol/ μ L)	CoV
1	41.26	5.27%	31.56	2.51%	34.02	5.17%
2	39.75	2.26%	27.48	7.20%	28.04	1.38%
3	44.34	3.24%	28.06	0.09%	29.6	1.81%
4	42.28	3.76%	33.54	5.37%	38.1	3.24%
5	39.49	1.44%	33.29	7.98%	36.58	2.19%
6	43.73	4.08%	39.04	6.03%	42.05	1.95%
7	40.22	6.29%	36.49	1.72%	25.4	5.80%
8	45.82	2.96%	31.35	3.34%	36.65	5.34%
9	41.55	11.84%	36.09	0.29%	33.71	8.30%
Average	42.05	6.72%	32.99	11.82%	33.79	15.48%

From Table 4-16, and indeed Figure 4-32, it is clear that the amount of DA-Raf-1 detected per 25 μ L of DA-Raf-1 transfected HEK293 IP was higher than that estimated for either A-Raf WT or A-Raf Short. Peptide TVVTVR, which was previously found to yield the most reproducible data, estimated there to be 32.99 pmol of DA-Raf-1 per 25 μ L of IP. These high expression levels may have been due to the smaller MW of this isoform, yielding higher levels of protein expression within the 72 hours, or that more copies of this smaller protein can be maintained within the cell before becoming toxic.

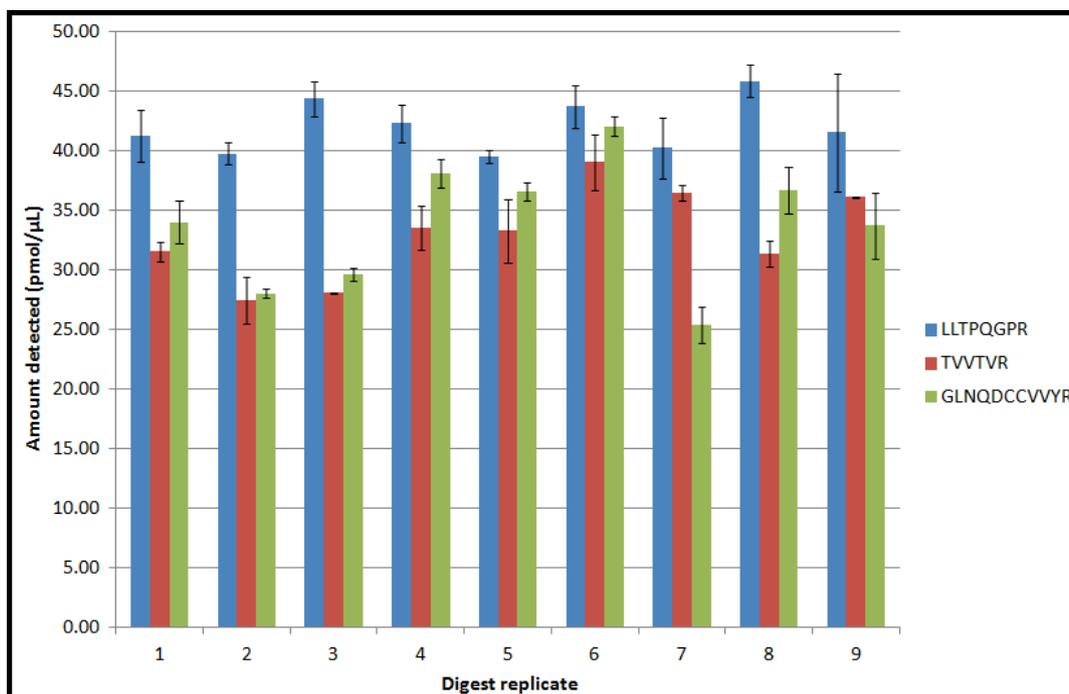


Figure 4-32: The quantities of each DA-Raf-1 target peptide detected during the spiked digestion of DA-Raf-1 transfected HEK293 IP on an API 2000. The amount of each DA-Raf-1 target peptide detected during the spiked digestion of 25 μ L of DA-Raf-1 transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between each peptide and replicate.

An ANOVA based statistical analysis was performed on this data, so as to assess if the means of the quantitations achieved for each peptide were equal, assuming a normal distribution. Therefore, during this test, the H_0 stated that the means of these groups were equal. This analysis was performed based upon an alpha level of 0.05, giving a critical F-value of 3.4, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-17, an F-value of 14.3 was achieved, therefore the H_0 was rejected, indicating the mean quantification values achieved for these peptides to be significantly different.

Table 4-17: An ANOVA based statistical analysis performed on the data obtained from the spiked digest of DA-Raf-1 transfected HEK293 IP. The quantification data obtained during the spiked digestion of 25 μ L of DA-Raf-1 transfected HEK293 IP was subjected to ANOVA. During this statistical analysis, the data obtained for each targeted DA-Raf-1 peptide, and each digest replicate were tested, so as to assess if the means obtained for each peptide were equal. Highlighted in yellow is the F-value, which if above 3.4, would indicate the data to be significantly different, in 95% of cases.

Groups	Replicates	Sum	Average	Variance	F-value
LLTPQGPR	9	378.4426	42.04918	4.806482	14.27344
TVVTVR	9	296.8988	32.98876	14.78117	
GLNQDCCVYR	9	304.1512	33.79458	27.98013	

From viewing the averaged quantitation values presented on Table 4-16, the data obtained for peptides TVVTVR and GLNQDCCVYR appeared more closely grouped than those obtained for peptide LLTPQGPR. As the level of peptide LLTPQGPR was higher than the level of peptide TVVTVR, then this difference may have resulted from the incomplete re-suspension of AQUA peptide in 2% MeCN (v/v) (Figure 4-32).

So as to confirm there was no variation detected between the means of peptides GLNQDCCVYR and TVVTVR, assuming a normal distribution, a two-tailed (two-tailed as either an increase, or a decrease in mean would be considered variation) T-test was performed (Table 4-18). During this test, a H_0 , that the means of these groups were equal, was stated. This analysis was performed based upon an alpha level of 0.05, giving a critical P-value of 2.3, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-18, a P-value of 0.64 was achieved, therefore the H_0 was accepted, indicating the mean quantification values achieved for these peptides was equal.

Table 4-18: A T-test based statistical analysis performed on the data obtained from the spiked digest of DA-Raf-1 transfected HEK293 IP. A comparison of the quantitation values obtained for peptides TVVTVR and GLNQDCCVYR, so as to determine if the means of each set of data were significantly different. Highlighted in yellow is the two-tailed P-value, which if above 2.3, would suggest that there was significant variation between the groups in 95% of cases.

	TVVTVR	GLNQDCCVYR
Mean	32.98875937	33.79457691
Variance	14.7811722	27.98012665
Replicates	9	9
P-value	0.642086961	

In conclusion, when 25 μ L of DA-Raf-1 transfected HEK293 IP was digested with trypsin, no variation was detected between the quantitation values obtained for peptides TVVTVR and GLNQDCCVVYR, based upon a two-tailed t-test. However, when a Glu-C based spiked digest was performed and the quantification values compared for both the trypsin and Glu-C cleaved peptides via ANOVA, significant variation was detected.

Large CoV values were also obtained between each replicate, which may have resulted from small variations in pipetting volume made during the loading of the SDS-PAGE gel or indeed the accuracy of the gel band excision pre-digestion. While these small variations would not usually have a noticeable effect on the data obtained, the high expression level at which the DA-Raf-1 was detected within the transfected HEK293 IP may have exacerbated this problem. Indeed, the signal response of the internal standards (IS) on the API 2000 (data not shown) appeared constant between injections, further suggesting the cause of the variation to originate before the internal standard was added.

4.4.5.1.4 DA-Raf-2 Single Isoform Spiked Digest

While DA-Raf-2 contains three AQUA peptides, GLNQDCCVVYR, TVVTVR and VPTVCVDMSTNRQQ, peptide VPTVCVDMSTNRQQ failed to be detected via ESI-MS, instead requiring a MALDI-ToF based analysis. When running this set of spiked digest based analyses, no MALDI MS was available for use, resulting in a comparison only being made between peptides TVVTVR and GLNQDCCVVYR.

Table 4-19: The quantity and variation data obtained from the spiked digestion of 25 μL of immunoprecipitated DA-Raf-2 on an API 2000. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

Replicate	TVVTVR - Present in all but A-Raf Short		GLNQDCCVVYR - Present in all isoforms	
	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV
1	11.39	7.46%	7.27	9.57%
2	10.86	2.23%	8.25	2.71%
3	11.27	6.21%	6.80	3.85%
4	11.60	7.57%	13.91	4.56%
5	11.04	1.29%	6.50	3.42%
6	11.52	5.73%	13.29	8.49%
7	12.14	1.04%	12.96	5.56%
8	11.57	1.57%	12.85	10.54%
9	11.42	5.36%	6.93	1.65%
Average	11.42	4.77%	9.86	32.38%

From Table 4-19, and Figure 4-33, peptide TVVTVR again appears to yield the most reproducible data, estimating there to be to 11.42 pmol of DA-Raf-2 per 25 μL of DA-Raf-2 transfected HEK293 IP, with a CoV of 4.77%. When the quantity of DA-Raf-2 is assessed through the use of peptide GLNQDCCVVYR, however, 25 μL of transfected IP is estimated to contain between 6.5 and 13.91 pmol of DA-Raf-2, with an average CoV of 32.38%.

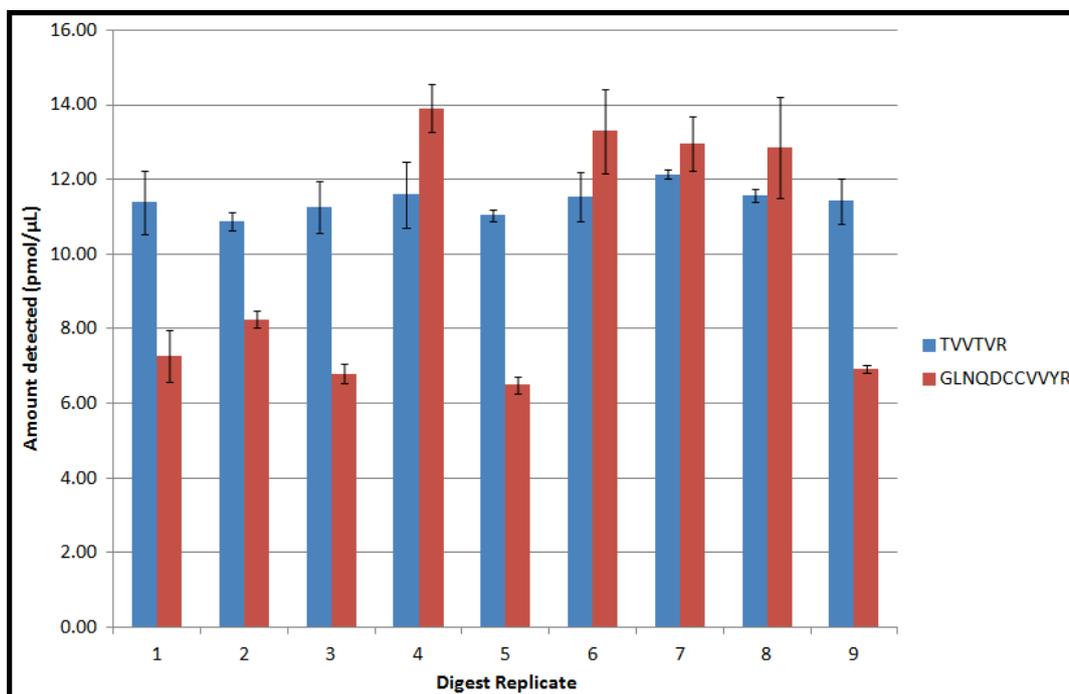


Figure 4-33: The quantities of each DA-Raf-2 target peptide detected during the spiked digestion of DA-Raf-2 transfected HEK293 IP on an API 2000. The amount of each DA-Raf-2 target peptide detected during the spiked digestion of 25 μL of DA-Raf-2 transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between each peptide and replicate.

So as to assess if there was any variation between the means of peptides GLNQDCCVVYR and TVVTVR, assuming a normal distribution, a two-tailed T-test was performed. During this test, a H_0 , that the means of these groups were equal, was stated. This analysis was performed based upon an alpha level of 0.05, giving a critical P-value of 2.3, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-20, a P-value of 0.16 was achieved, therefore the H_0 was accepted, indicating the mean quantification values achieved for these peptides was equal.

Table 4-20: A T-test based statistical analysis performed on the data obtained from the spiked digest of DA-Raf-2. A T-test based statistical analysis assessing the quantitation data obtained for peptides TVVTVR and GLNQDCCVVYR. This test was performed so as to assess if each data set contained an equal level of variation. Highlighted in yellow is the two-tailed P value, where a figure above 2.3 would suggest there to be significant variation between the two groups in 95% of cases.

	TVVTVR	GLNQDCCVVYR
Mean	11.42455233	9.861816048
Variance	0.132878544	10.65610144
Observations	9	9
P-value	0.160234094	

During the analysis of the DA-Raf-2 transfected HEK293 IP, the level of DA-Raf-2 expressed was above the LoD for peptide GLNQDCCVVYR, and thus an accurate quantification should be possible. The high CoV values detected during the quantification of peptide GLNQDCCVVYR, however, may indicate the alkylation of both the target protein (when the amount of GLNQDCCVVYR detected was lower than the amount of TVVTVR) and the IS (when the amount of GLNQDCCVVYR detected was higher than the amount of TVVTVR) to be incomplete.

4.4.5.1.5 Single Isoform Spiked Digest Overview

Several problems were identified during the spiked digestion of each of the four A-Raf isoforms on an API 2000 MS. Firstly, the high quantitation value achieved for peptide LLTPQGPR during the spiked digestion of DA-Raf-1 indicated that there may be less peptide LLTP*QGPR within the sample than was expected. If this was the case, then it may be due to the incomplete re-suspension of this AQUA peptide in 2% MeCN (v/v). So as to test this theory, an unopened vial of AQUA peptide LLTPQGPR was re-suspended in 20 µL of 10% FA (v/v), diluting further to 200 µL in 0.1% FA (v/v) and the spiked digestions re-running. Secondly, peptide GLNQDCCVVYR presented with both high CoV and quantitation values, significantly different to those detected for other peptides quantified for each A-Raf isoform. So as to ensure peptide GLNQDCCVVYR was alkylated correctly, peptide GL*NQDCCVVYR was alkylated in-solution, and the peptide further reduced with DTT or quenched with excess Cys. The results from which are shown on Figure 4-34.

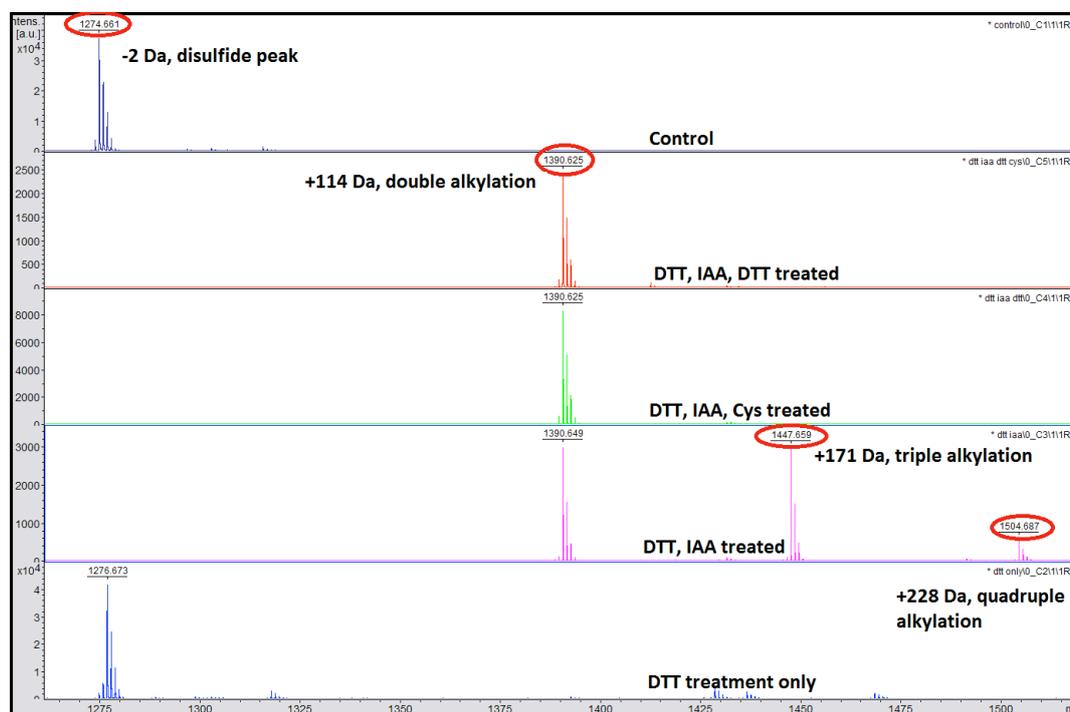


Figure 4-34: An investigation performed so as to determine the source of the variation detected during the spiked digestion of AQUA peptide GL*NQDCCVVYR. An analysis performed on 5 pmol of AQUA peptide GL*NQDCCVVYR, in an unmodified state (where the peptide exists in a partially di-sulfide bonded state), when reduced with DTT (breaking the di-sulfide bonds), when reduced with DTT and alkylated with IAA, when reduced with DTT, alkylated with IAA and reduced again with DTT, and finally when reduced with DTT, alkylated with IAA and the IAA quenched with Cys. Circled in red are the various states in which AQUA peptide GL*NQDCCVVYR was identified. This data was collected on an Ultraflex II MS.

From Figure 4-34 it appears that when peptide GL*NQDCCVVYR was treated with DTT and IAA that multiple alkylation peaks were detected. Similar results have previously been reported, including the alkylation of the amino-group on Lys (Galvani et al., 2001), it is therefore possible that in this situation the amino-group on either the peptides N-terminus, or the C-terminal Arg residue may have been alkylated. These unintended alkylations could, however, be reversed or prevented through the addition of excess DTT or Cys, post-alkylation. This experiment helps to explain some of the variation detected during the spiked digest based analyses, and as a result all future in-gel and in-solution alkylation reactions will be quenched with one of the above reagents.

4.4.5.2 Optimisation of the Single Isoform Based Spiked Digests

4.4.5.2.1 A-Raf WT Single Isoform Spiked Digest Optimisation

As peptides GLNQDCCVVYR and VSQPTAEQAQAFK yielded highly variable CoV values during the previous A-Raf WT spiked digestion, perhaps due in part to the low level of A-Raf WT detected within the transfected HEK293 IP, a second transfected HEK293 IP was prepared, in this case re-suspending the washed anti-FLAG beads in half the volume of protein loading buffer (180 μ L). The 25 μ L of IP loaded into each SDS-PAGE gel well, therefore, should contain the transfected, immunoprecipitated contents of approximately 9.2×10^7 HEK293 cells. This change, it was hoped, would double the amount of the A-Raf WT within the IP, from 600 fmol per 25 μ L to approximately 1.2 pmol per 25 μ L, enabling larger MRM peak areas to be obtained for peptides VSQPTAEQAQAFK and GLNQDCCVVYR. In addition to this change, a second DTT based reduction step was incorporated into the in-gel spiked digest based protocol, so as to ensure any undesirable alkylations were reversed, prior to LC-MS.

Table 4-21: The quantity and variation data obtained from the optimised spiked digestion of 25 μL of immunoprecipitated A-Raf WT on an API 2000. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

Replicate	TVVTVR - Present in all but A-Raf Short		VSQPTAEQAQAFK - Unique to A-Raf WT		GLNQDCCVVYR - Present in all isoforms	
	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV
1	1.24	16.04%	1.6	8.36%	2.2	13.50%
2	1.44	17.62%	1.8	3.53%	2.56	9.60%
3	1.01	13.00%	1.14	1.34%	2.31	5.58%
Average	1.23	20.61%	1.53	18.90%	2.36	10.88%

From reviewing Table 4-21 and Figure 4-35, it would appear that the amount of A-Raf WT detected within the transfected HEK293 IP increased from 600 fmol per 25 μL , to 1.23 pmol per 25 μL , as quantified with peptide TVVTVR, over three sample replicates, with a CoV of 20.6%.

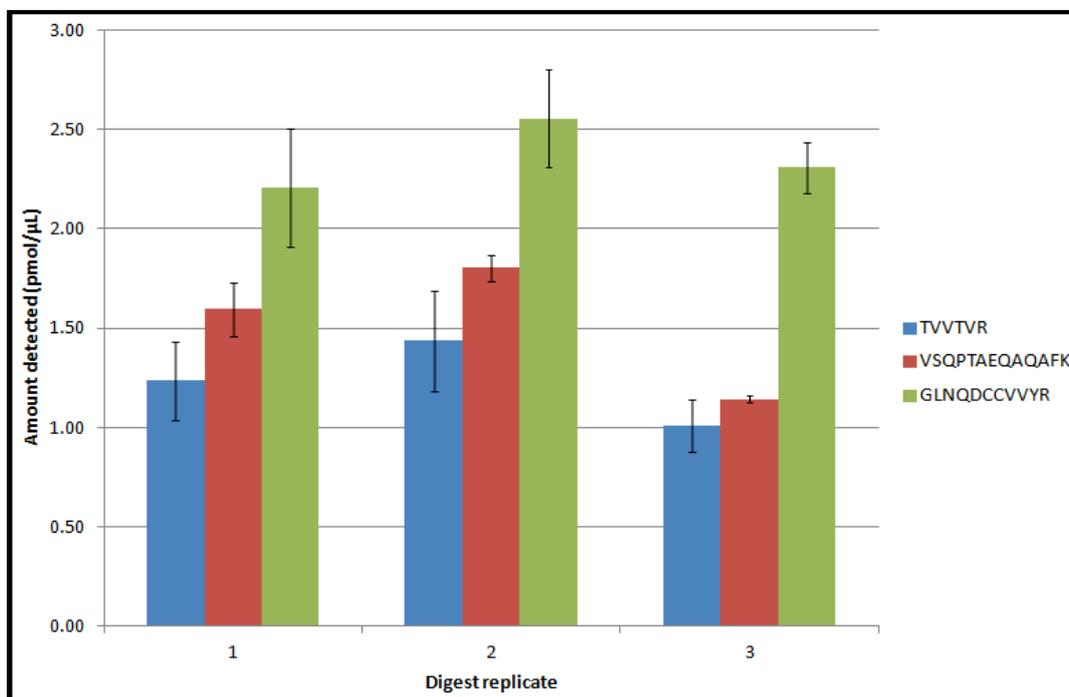


Figure 4-35: The quantities of each A-Raf WT target peptide detected during the optimised spiked digestion of A-Raf WT transfected HEK293 IP on an API 2000. The amount of each A-Raf WT target peptide detected during the optimised spiked digestion of 25 μL of A-Raf WT transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between each peptide and replicate.

An ANOVA based statistical analysis was performed on this data, so as to assess if the means of the quantitations achieved for each peptide were equal, assuming a normal distribution. Therefore, during this test, the H_0 stated that the means of these groups were equal. This analysis was performed based upon an alpha level of 0.05, giving a critical F-value of 5.1, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-22, an F-value of 15.9 was achieved, therefore the H_0 was rejected, indicating the mean quantification values achieved for these peptides to be significantly different.

Table 4-22: An ANOVA based statistical analysis performed on the data obtained from the optimised spiked digest of A-Raf WT transfected HEK293 IP. The quantification data obtained during the optimised spiked digestion of 25 μ L of A-Raf WT transfected HEK293 IP was subjected to ANOVA. During this statistical analysis, the data obtained for each targeted A-Raf WT peptide, and each digest replicate were tested, so as to assess if the means obtained for each peptide were equal. Highlighted in yellow is the F-value, which if above 5.1, would indicate the data to be significantly different, in 95% of cases.

Groups	Count	Sum	Average	Variance	F-value
TVVTVR	3	3.69	1.23	0.0463	15.86367
VSQPTAEQAQAFK	3	4.54	1.513333	0.114533	
GLNQDCCVVYR	3	7.07	2.356667	0.034033	

While significant variation was detected between the means of the three A-Raf WT peptides, quantified through the spiked digestion of 25 μ L of A-Raf WT transfected HEK293 IP, the data obtained for peptides VSQPTAEQAQAFK and TVVTVR appeared more closely grouped (Table 4-21). So as to assess if there was any variation between the means of peptides TVVTVR and VSQPTAEQAQAFK, assuming a normal distribution, a two-tailed T-test was performed. During this test, a H_0 , that the means of these groups were equal, was stated. This analysis was performed based upon an alpha level of 0.05, giving a critical P-value of 2.8, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-23, a P-value of 0.28 was achieved, therefore the H_0 was accepted, indicating the mean quantification values achieved for these peptides was equal.

Table 4-23: A T-test based statistical analysis performed on the data obtained from the optimised spiked digestion of A-Raf WT. A T-test based statistical analysis performed on the quantitation data obtained for peptides TVVTVR and VSQPTAEQAQAFK. This test was performed so as to assess if the data obtained for each of the peptide contained an equal level of variation. Highlighted in yellow is the two-tailed P-value, where a value above 2.8 would indicate that there was significant variation between the groups in 95% of cases.

	TVVTVR	VSQPTAEQAQAFK
Mean	1.227833452	1.514465864
Variance	0.045132721	0.113404655
Replicates	3	3
P-value	0.280471873	

As the CoV values obtained for peptide GLNQDCCVVYR remained low over the three replicates, and with three transitions per replicate, this would suggest that the in-gel digestion, recovery and LC-MS/MS methods were reproducible. Variation was detected between the quantitation values obtained for peptide

GLNQDCCVVYR and those obtained for peptides TVVTVR and VSQPTAEQAQAFK, however, which would indicate a problem with the quantification of this peptide (Figure 4-35). As the in-gel digest and in-solution AQUA preparation protocols were previously modified to ensure excess alkylation was not an issue (which eliminated quantification ratios which presented at a rate lower than expected), the only logical conclusion was that the in-solution processing of the AQUA peptide, prior to spiking the gel pieces, was responsible for the sample loss.

Sample loss occurring during downstream processing has previously been documented, with vacuum centrifugation based protocols, such as this, causing losses of up to 50% (Speicher et al., 2000). The only solution would therefore be to order each of the Cys containing AQUA peptides pre-alkylated, requiring no further down-stream processing. While unmodified peptides were intentionally selected during this study to correct for any inefficiency encountered during the alkylation of Cys, this reaction has since been shown to be complete.

4.4.5.2.2 DA-Raf-1 Single Isoform Spiked Digest Optimisation

When a spiked digestion was first performed on the DA-Raf-1 transfected HEK293 IP, 32.99 pmol of DA-Raf-1 was detected, per 25 μ L of sample, with a CoV value of 11.82% (based upon the use of peptide TVVTVR). In comparison, when peptide LLTPQGPR was quantified, the amount of DA-Raf-1 within 25 μ L of DA-Raf-1 transfected HEK293 IP was calculated to be 42.05 pmol, with a CoV of 6.72%. Therefore, the incomplete re-suspension of AQUA peptide LLTPQGPR was suspected. So as to test this theory, an unopened vial of peptide LLTPQGPR was re-suspended in 20 μ L of 10% FA (v/v), further diluting to 200 μ L in 0.1% FA (v/v). The spiked digest was then re-run, loading 5 μ L of the DA-Raf-1 transfected HEK293 IP to each SDS-PAGE gel lane, equal to the transfected, immunoprecipitated contents of approximately 9×10^6 HEK293 cells.

Table 4-24: The quantity and variation data obtained from the spiked digestion of 5 μL of immunoprecipitated DA-Raf-1 on an API 2000. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

Replicate	LLTPQGPR - Unique to DA-Raf-1		TVVTVR - Present in all but A-Raf Short		GLNQDCCVVYR - Present in all isoforms	
	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV
1	7.54	9.65%	7.37	3.29%	5.46	35.67%
2	7.81	16.41%	7.64	13.48%	4.69	9.64%
Average	7.68	12.30%	7.51	9.12%	5.08	26.29%

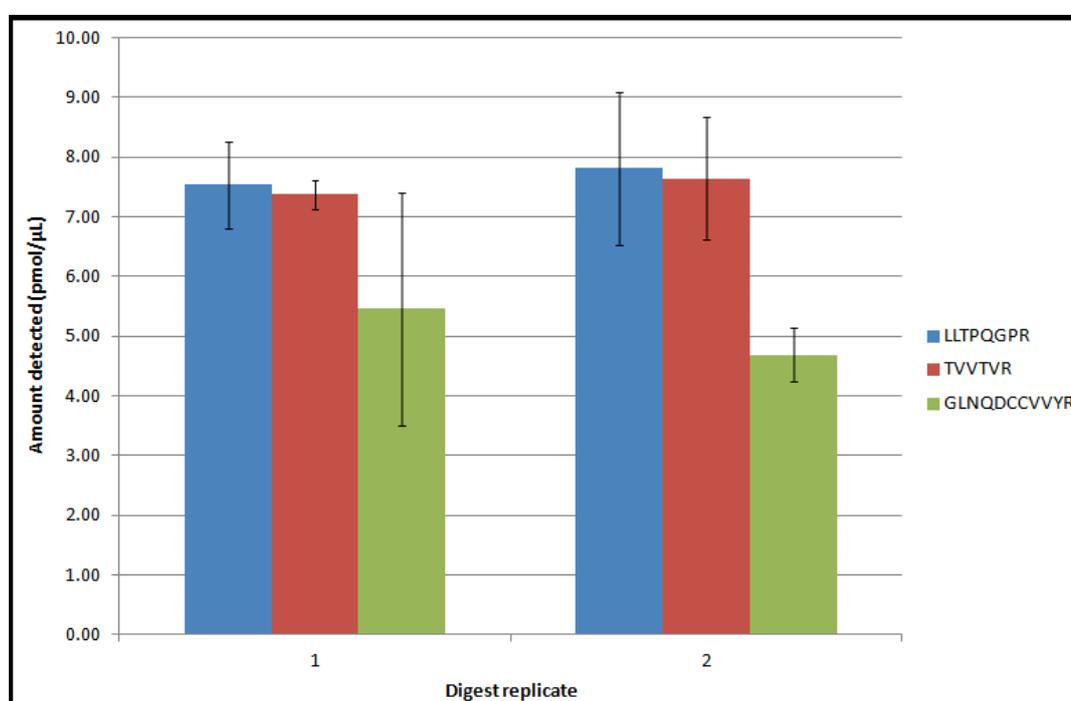


Figure 4-36: The quantities of each DA-Raf-1 target peptide detected during the optimised spiked digestion of DA-Raf-1 transfected HEK293 IP on an API 2000. The amount of each DA-Raf-1 target peptide detected during the optimised spiked digestion of 5 μL of DA-Raf-1 transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between each peptide and replicate.

Through assessing Table 4-24 and Figure 4-36, the precision of the quantifications achieved through the use of peptide LLTPQGPR were seen to improve, at least in comparison to those achieved through the use of peptide TVVTVR. An ANOVA based statistical analysis was performed on this data, so as

to assess if the means of the quantitations achieved for each peptide were equal, assuming a normal distribution. Therefore, during this test, the H_0 stated that the means of these groups were equal. This analysis was performed based upon an alpha level of 0.05, giving a critical F-value of 9.6, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-25, an F-value of 34.4 was achieved, therefore the H_0 was rejected, indicating the mean quantification values achieved for these peptides to be significantly different.

Table 4-25: An ANOVA based statistical analysis performed on the data obtained during the optimised spiked digest of DA-Raf-1 transfected HEK293 IP. The quantification data obtained during the optimised spiked digestion of 5 μ L of DA-Raf-1 transfected HEK293 IP was subjected to ANOVA. During this statistical analysis, the data obtained for each targeted A-Raf WT peptide, and each digest replicate were tested, so as to assess if the means obtained for each peptide were equal. Highlighted in yellow is the F-value, which if above 9.6, would indicate the data to be significantly different, in 95% of cases.

Groups	Count	Sum	Average	Variance	F-value
LLTPQGPR	2	15.35	7.675	0.03645	34.36794
TVVTVR	2	15.01	7.505	0.03645	
GLNQDCCVVYR	2	10.15	5.075	0.29645	

While significant variation was detected between the quantification values obtained for each of the three DA-Raf-1 peptides, during the spiked digestion of the DA-Raf-1 transfected HEK293 IP, the averaged quantitation values obtained for peptides TVVTVR and LLTPQGPR appeared more closely grouped (Table 4-24). So as to assess if there was any variation between the means of peptides TVVTVR and LLTPQGPR, assuming a normal distribution, a two-tailed T-test was performed. During this test, a H_0 , that the means of these groups were equal, was stated. This analysis was performed based upon an alpha level of 0.05, giving a critical P-value of 4.3, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-26, a P-value of 0.5 was achieved, therefore the H_0 was accepted, indicating the mean quantification values achieved for these peptides to be equal. This is particularly impressive as each of the peptides originated from a different SDS-PAGE gel, and indeed were both peptides were digested with different proteases.

Table 4-26: A T-test based statistical analysis performed on the data obtained from the optimised spiked digestion of DA-Raf-1. A T-test based statistical analysis performed on the quantitation data obtained from peptides TVVTVR and LLTPQGPR. This statistical test was performed so as to determine if the variation between the data sets was equal. Highlighted in yellow is the two-tailed P value, where a figure above 4.3 would indicate there to be significant variation between the groups in 95% of cases.

	LLTPQGPR	TVVTVR
Mean	7.67510408	7.505279247
Variance	0.037150395	0.035091527
Observations	2	2
P-value	0.465849755	

As peptides TVVTVR and LLTPQGPR were found to have equal means, while significant variation was identified between the means of the three DA-Raf-1 peptides, this would indicate peptide GLNQDCCVVYR to be the cause of the variation. As this experiment was conducted before the in-gel digest protocol was modified, then this variation most likely resulted from peptide GLNQDCCVVYR receiving excess alkylation.

4.4.5.2.3 DA-Raf-2 QTrap 4000 and Ultraflex II Based Analysis

When a spiked digestion was last performed on the DA-Raf-2 transfected HEK293 IP, no MALDI-ToF MS was available for the quantification of Lys-C cleaved peptide VPTVCVDMSTNRQQ. When running this spiked digestion however, an Ultraflex II was available, enabling all three DA-Raf-2 peptides to be quantified. Furthermore, as 11.42 pmol of DA-Raf-2 was detected per 25 μ L of IP (based upon the data obtained for peptide TVVTVR), the SDS-PAGE gel loading volume was reduced by 50% (to 12.5 μ L), equal to the transfected, immunoprecipitated contents of approximately 2.3×10^7 HEP293 cells.

Regarding the MALDI based data analysis; the peak heights for isotopic peaks M, M^{+1} , M^{+2} and M^{+3} were assessed, the data averaged, and a CoV value calculated.

Table 4-27: The quantity and variation data obtained from the spiked digestion of 12.5 μL of immunoprecipitated DA-Raf-2 on a QTrap 4000. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

Replicate	TVVTVR - Present in all but A-Raf Short		GLNQDCCVYR - Present in all A-Raf isoforms		VPTVCVDMSTNRQQ - Unique to DA-Raf-2	
	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV
1	6.67	2.07%	9.66	9.40%	10.75	14.47%
2	6.97	3.26%	9.7	9.69%	16.69	6.88%
3	6.14	4.49%	9.16	8.90%	14.91	5.06%
4	6.18	3.11%	8.67	9.99%	17.39	7.48%
5	5.78	4.61%	8.12	10.35%	18.57	7.98%
Average	6.41	7.03%	9.17	10.69%	15.66	8.37%

Table 4-27 and Figure 4-37 show the quantitation data obtained from the spiked digestion of 12.5 μL of DA-Raf-2 IP. Of the results obtained, peptide TVVTVR estimated there to be 6.41 pmol of DA-Raf-2 per 12.5 μL of DA-Raf-2 transfected HEK293 IP, with an averaged CoV of just 7.03%.

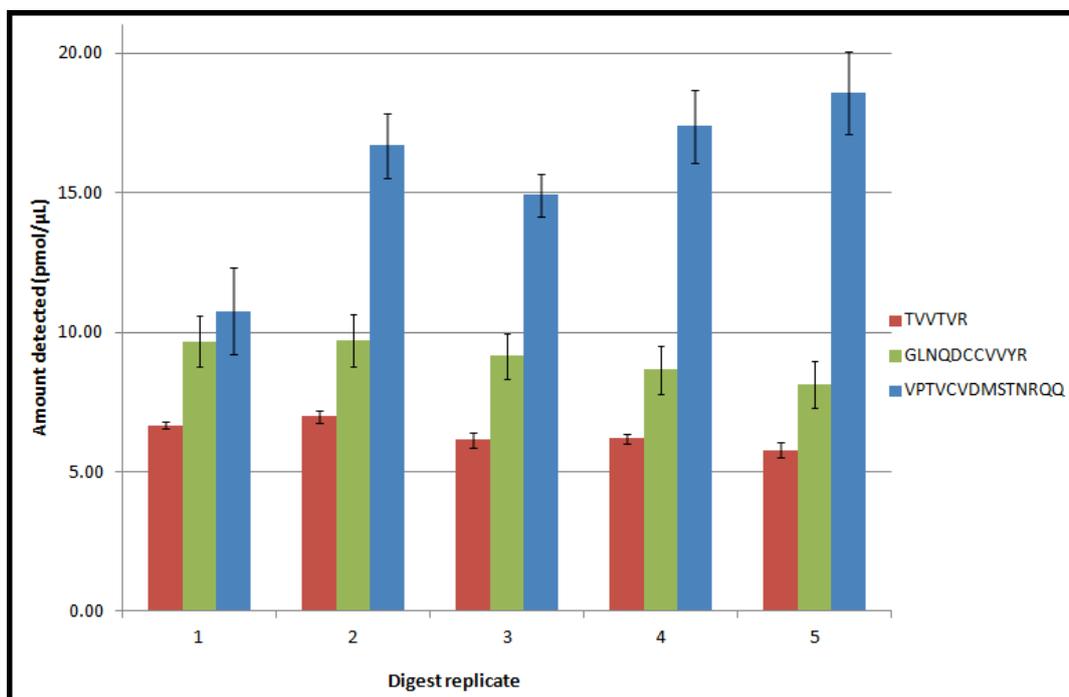


Figure 4-37: The quantities of each DA-Raf-2 target peptide detected during the spiked digestion of DA-Raf-2 transfected HEK293 IP on a QTrap 4000, and an Ultraflex II. The amount of each DA-Raf-2 target peptide detected during the spiked digestion of 12.5 μL of DA-Raf-2 transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between each peptide and replicate.

An ANOVA based statistical analysis was performed on this data, so as to assess if the means of the quantitations achieved for each peptide were equal, assuming a normal distribution. Therefore, during this test, the H_0 stated that the means of these groups were equal. This analysis was performed based upon an alpha level of 0.05, giving a critical F-value of 3.9, which if exceeded, would require the H_0 to be rejected. As is shown on Table 4-28, an F-value of 34.5 was achieved, therefore the H_0 was rejected, indicating the mean quantification values achieved for these peptides to be significantly different.

Table 4-28: An ANOVA based statistical analysis performed on the data obtained during the optimised spiked digest of DA-Raf-2 transfected HEK293 IP. The quantification data obtained during the optimised spiked digestion of 5 μ L of DA-Raf-2 transfected HEK293 IP was subjected to ANOVA. During this statistical analysis, the data obtained for each targeted A-Raf WT peptide, and each digest replicate were tested, so as to assess if the means obtained for each peptide were equal. Highlighted in yellow is the F-value, which if above 3.9, would indicate the data to be significantly different, in 95% of cases.

Groups	Count	Sum	Average	Variance	F-value
TVVTVR	5	31.74	6.348	0.22117	34.5119
GLNQDCCVVYR	5	45.31	9.062	0.45382	
VPTVCVDMSTNRQQ	5	78.31	15.662	9.29812	

In comparison to the amount of DA-Raf-2 detected in 12.5 μ L of DA-Raf-2 transfected HEK293 IP through the quantitation of peptide TVVTVR, peptide GLNQDCCVVYR estimated there to be 9.17 pmol of DA-Raf-2 per 12.5 μ L of sample, a value approximately 25% higher. This phenomenon was previously been explained to result from the loss of IS during the alkylation of peptide GLNQDCCVVYR. In comparison to both peptides TVVTVR and GLNQDCCVVYR, however, peptide VPTVCVDMSTNRQQ predicted the amount of DA-Raf-2 per 12.5 μ L of sample to be even higher (15.66 pmol/ 12.5 μ L with a CoV of only 8.37%). Therefore, there most likely was also a problem with the preparation of AQUA peptide VPTVCVDMSTNRQQ, rather than the incomplete digestion of DA-Raf-2 with Lys-C. As peptide VPTVCVDMSTNRQQ was already re-suspended in 10% FA (v/v), and diluted to 200 μ L in 0.1% FA (v/v), then there was no reason to suspect incomplete re-suspension. Instead, as with peptide GLNQDCCVVYR, AQUA peptide VPTVCVDMSTNRQQ may suffer from sample loss as a result of down-stream processing. Some redundancy does, however, exist in the quantification of the four A-Raf peptides, enabling the amount of DA-Raf-2 to be determined when each of the four A-Raf isoforms are quantified.

4.4.5.3 Multiple Isoform Spiked Digest Based Analysis

Having performed single isoform based spiked digestions to quantify the amount of A-Raf expressed in a range of transfected HEK293 cell IP, and having obtained statistically comparable quantification values for peptides TVVTVR, VSQPTAEQAQAFK and LLTPQGPR, it should be possible to combine known

amounts of each A-Raf isoform, and to re-quantify each, based upon the amount of each target peptide detected.

Approximately 1 pmol of each A-Raf isoform (each based upon a quantification performed with peptide TVVTVR, with the exception of A-Raf Short, which was instead quantified with peptide GLNQDCCVVYR) was separated via SDS-PAGE, the target gel bands excised, reduced and alkylated before being transferred to a single vial and digested in the presence of 5 pmol of each AQUA peptide. The peptides were extracted and vacuum centrifuged, injecting onto an Ultimate 3000 HPLC and analysing on a QTrap 5500 MS, running the MRM acquisition method previously devised.

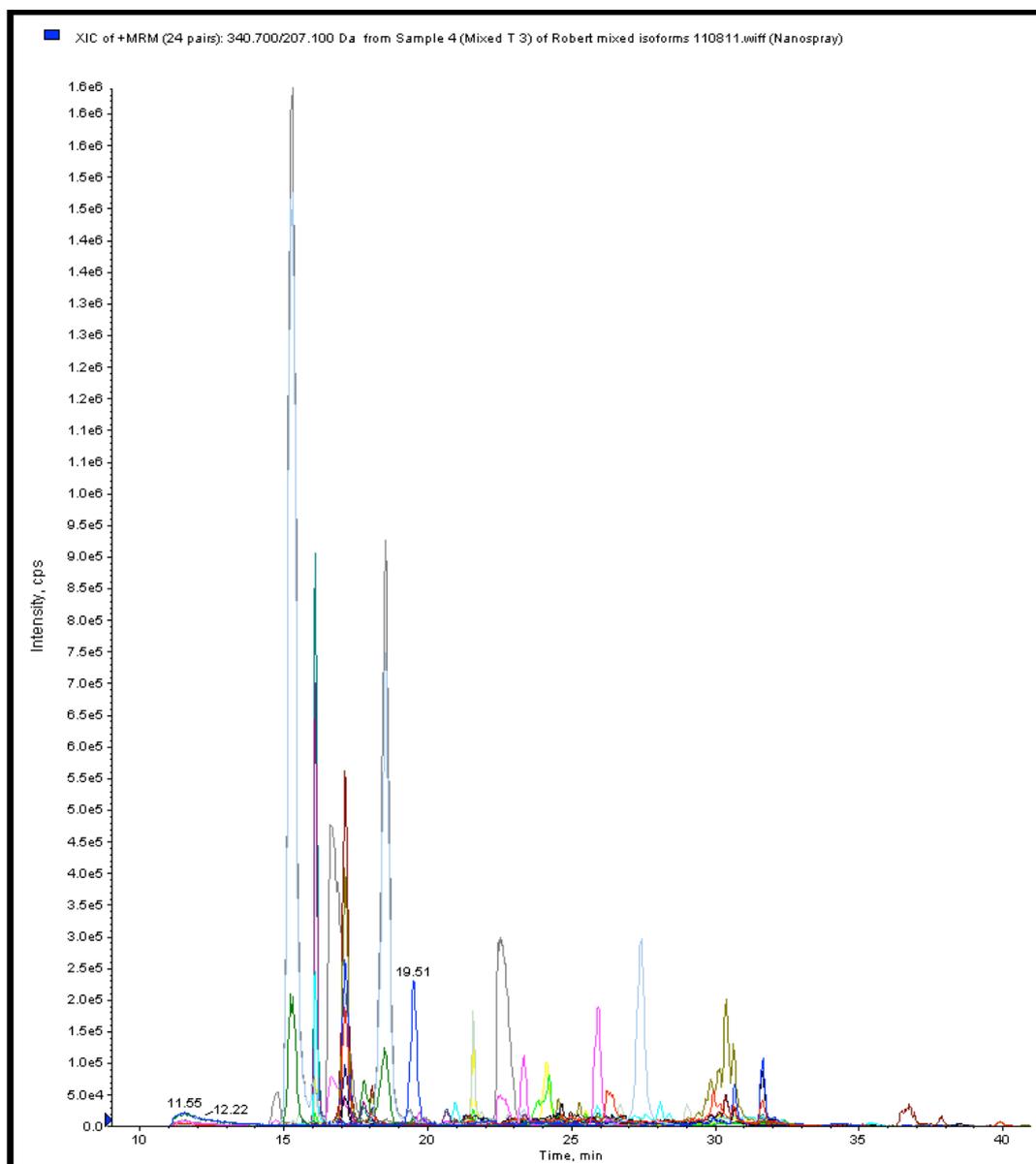


Figure 4-38: The complex XIC obtained during the trypsin based spiked digestion of each of the four A-Raf isoforms, analysed on a QTrap 5500. Through digesting immunoprecipitated A-Raf WT, A-Raf Short, DA-Raf-1 and DA-Raf-2 with trypsin, in the presence of 5 pmol of AQUA peptides TV*VTVR, GL*NQDCCVVYR, VSQP*TAEQAQAFK and LLTP*QGPR, a spiked digest solution was obtained which was injected onto an Ultimate 3000 HPLC and analysed on an QTrap 5500 MS.

From Figure 4-38 it is clear that the XIC obtained from the trypsin based digestion of the four A-Raf isoforms is a lot more complex than that obtained for A-Raf WT on the API 2000 (Figure 4-29). This will, in part, be due to the increased number of isoforms being analysed, but also due to the increase in

sensitivity offered by the QTrap 5500 over the API 2000, detecting more ions with m/z ratios similar to those targeted by the MRM acquisition method within the sample. It is thus important that when each peptide is quantified, that the correct peak is selected (selecting a peak which is present in each MRM transition, three transitions per peptide, and is present in both the synthetic and target XICs). The data obtained from the mixed isoform based analyses are listed on Table 4-29 and presented on Figure 4-39.

Table 4-29: The quantity and variation data obtained from the spiked digestion of 1 pmol of each of the four immunoprecipitated A-Raf isoforms on a QTrap 5500. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

	TVVTVR - Present in all but A-Raf Short		GLNQDCCVVYR - Present in all A-Raf isoforms	
Replicate	Average per Replicate (pmol/ μ L)	CoV	Average per Replicate (pmol/ μ L)	CoV
1	2.00	5.73%	2.42	7.71%
2	1.93	1.77%	2.19	7.93%
3	2.07	4.27%	2.52	2.97%
Average	2.00	4.81%	2.37	8.26%
	LLTPQGPR - Unique to DA-Raf-1		VSQPTAEQAQAFK - Unique to A-Raf WT	
Replicate	Average per Replicate (pmol/ μ L)	CoV	Average per Replicate (pmol/ μ L)	CoV
1	0.75	2.71%	0.33	18.21%
2	0.69	4.48%	0.33	14.15%
3	0.72	3.42%	0.44	2.03%
Average	0.72	4.69%	0.37	18.40%

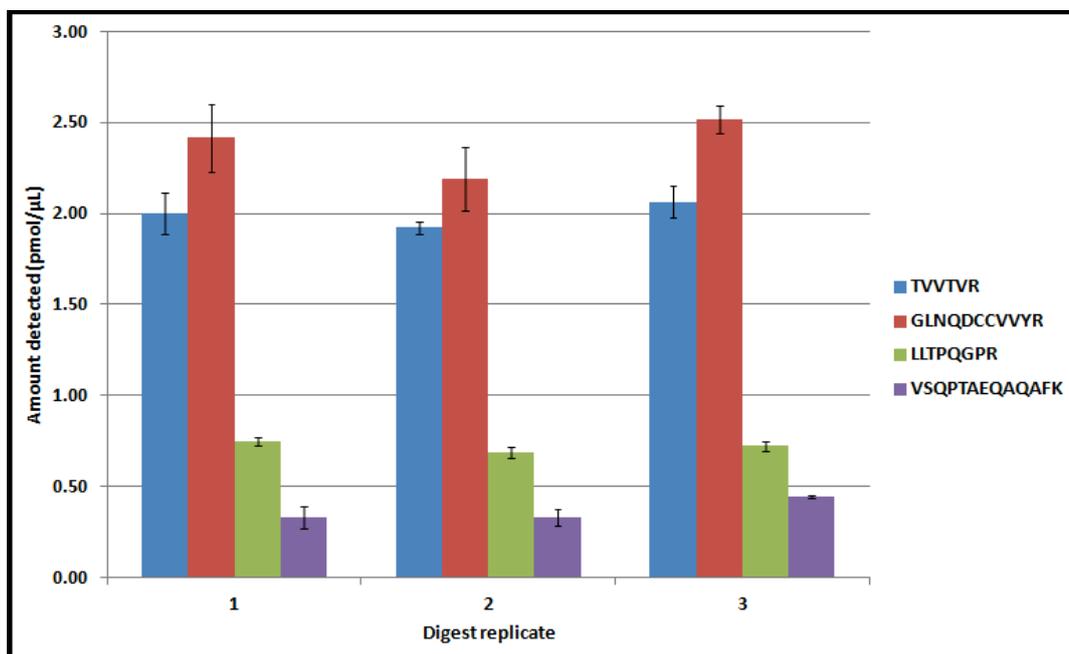


Figure 4-39: The quantities of each A-Raf target peptide detected during the spiked digestion of 1 pmol of each of the four A-Raf isoforms on a QTrap 5500. The amount of each A-Raf target peptide detected during the spiked digestion of 1 pmol of each of the four A-Raf isoforms. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per synthetic or endogenous A-Raf peptide, per replicate, and the values averaged. A CoV value was determined for this value. This bar chart helps to visualise the variation detected between each peptide and replicate.

In theory, as 1 pmol of each of the four A-Raf isoforms is digested, 4 pmol of peptide GLNQDCCVVYR would be expected (present in each of the four isoforms), 3 pmol of peptide TVVTVR would be expected (present in A-Raf WT, DA-Raf-1 and DA-Raf-2), 1 pmol of peptide LLTPQGPR would be expected (present only in DA-Raf-1) and 1 pmol of peptide VSQPTAEQAQAFK would be expected (present only in A-Raf WT). Instead, however, 2.37 pmol of peptide GLNQDCCVVYR was detected, with a CoV of 8.26%, 2 pmol of peptide TVVTVR was detected, with a CoV of 4.81%, 0.72 pmol of peptide LLTPQGPR was detected, with a CoV of 4.69%, and 0.37 pmol of peptide VSQPTAEQAQAFK was detected, with a CoV of 18.40% (Table 4-29). This peptide quantitation data was subsequently used to calculate the amount of each A-Raf isoform present within the sample. In brief, A-Raf Short was quantified through subtracting the amount of peptide TVVTVR detected, from the amount of peptide GLNQDCCVVYR

detected. Furthermore, DA-Raf-2 was quantified through subtracting the amount of peptide LLTPQGPR and VSQPTAEQAQAFK detected, from the amount of peptide TVVTVR detected (Table 4-30).

Table 4-30: The amount of each A-Raf isoform detected during a mixed isoform based spiked digestion. A set of spiked digests were performed, in which 1 pmol of each of the four A-Raf isoforms were digested. Through comparing the peak areas obtained for each peptide, a quantitation was attempted for each A-Raf isoform.

Peptide	Present in:	Average quantity (pmol)
GLNQDCCVVYR	A-Raf WT, A-Raf Short, DA-Raf-1, DA-Raf-2	2.37
TVVTVR	A-Raf WT, DA-Raf-1, DA-Raf-2	2.00
LLTPQGPR	DA-Raf-1	0.72
VSQPTAEQAQAFK	A-Raf WT	0.37
Isoform	Calculating the quantity of each isoform:	Quantity (pmol)
A-Raf WT	VSQPTAEQAQAFK	0.37
A-Raf Short	GLNQDCCVVYR - TVVTVR	0.38
DA-Raf-1	LLTPQGPR	0.72
DA-Raf-2	TVVTVR - VSQPTAEQAQAFK - LLTPQGPR	0.91

The data presented in Table 4-30 suggests that each of the four A-Raf isoforms were present at a level below 1 pmol. There may be several reasons why the peptide and protein levels detected during this experiment differed from those expected. Firstly, the amount of A-Raf Short detected within the A-Raf Short transfected HEK293 IP was quantified through the use of peptide GLNQDCCVVYR, meaning the amount of A-Raf Short, and thus the amount of peptide GLNQDCCVVYR, added to the gel, may have been incorrect. Secondly, as A-Raf Short was again quantified during this experiment with peptide GLNQDCCVVYR, the quantity detected may also have been incorrect.

In conclusion, the mixed isoform based spiked digestion may have been more successful if an accurate quantitation could have been achieved for peptides GLNQDCCVVYR and VPTVCVDMSTNRQQ.

4.4.5.4 Non-Immunoprecipitated Exogenous HEK293 Spiked Digest

While further optimisation may have improved the results obtained from the mixed isoform based spiked digestions, and indeed the single isoform based spiked digestions; a few final experiments were performed, so as to assess if it

was possible to detect the exogenous level of A-Raf WT within a transfected cell lysate, without the use of immunoprecipitation based sample enrichment. As the spiked digestions featured 1D-SDS-PAGE, LC based separation and MRM based acquisition, then in theory the majority of the non-target protein and peptide species within the cell lysate should be isolated from the point at which the target peptides elute. Should any non-target peptides co-elute with the target, it is likely they would have different Q1 and Q3 m/z ratios, and therefore should not be detected.

A 150 mm dish of HEK293 cells was transfected with A-Raf WT, and lysed after 72 hours. The cell lysate was centrifuged to remove any un-lysed cells and cellular debris and diluted to 1 mL with Protein Loading Buffer. Twenty five microlitres of cell lysate was loaded to an SDS-PAGE gel, separated, and digested with trypsin in the presence of each of the A-Raf AQUA peptides. The resulting solution was injected onto an Ultimate 3000 HPLC and detected via a QTrap 5500 MS, using the MRM acquisition method previously devised.

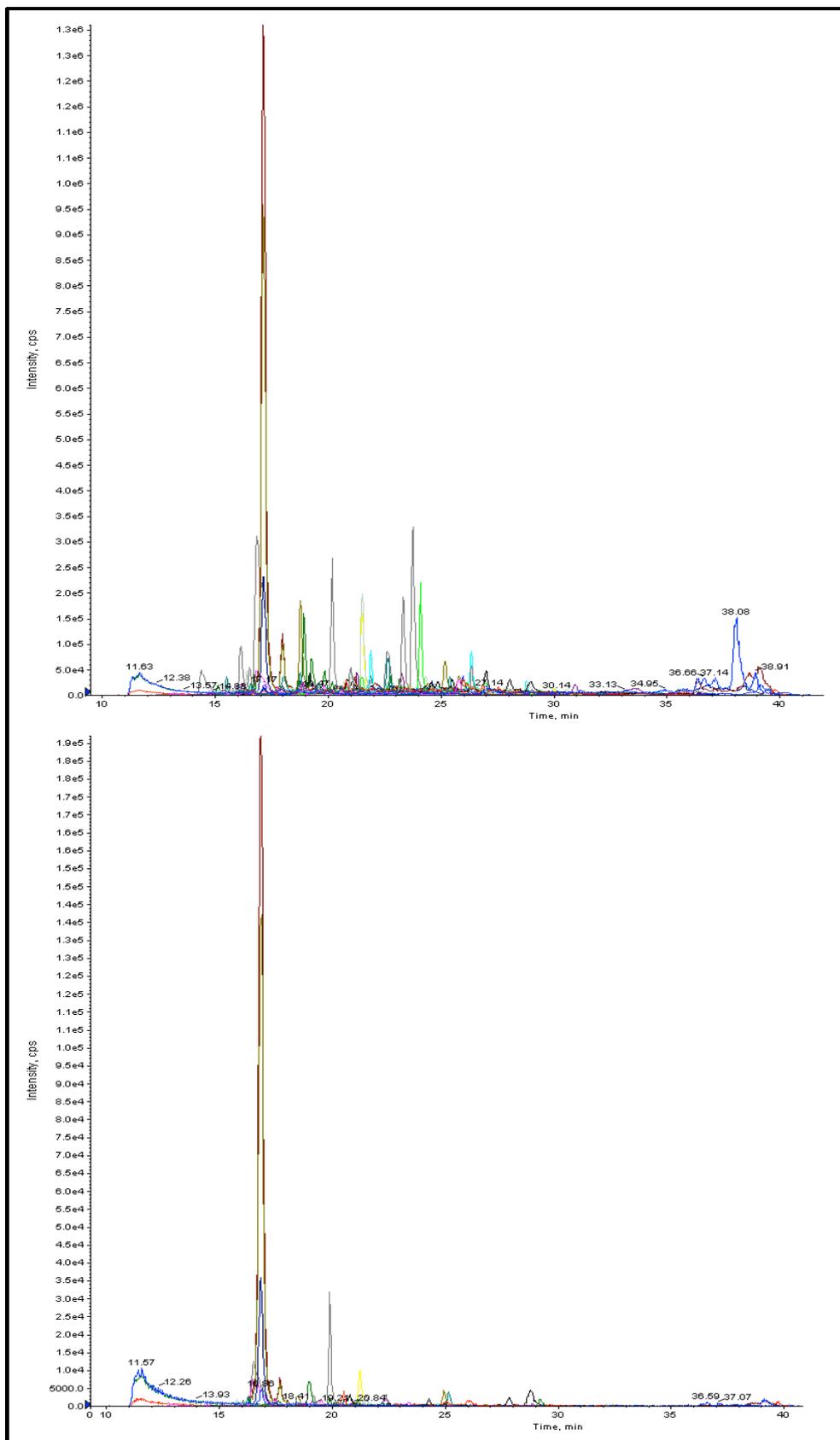


Figure 4-40: The XIC obtained when 25 μL of transfected, un-enriched, A-Raf WT lysate was subjected to a spiked digest based analysis, detecting on a QTrap 5500, utilising either low or unit Q1 resolution MRM transition windows. Twenty five μL of non-enriched, HEK 293 transfected A-Raf WT was analysed on a QTrap 5500 MS using the A-Raf MRM based MS acquisition method. The XIC spectra shown were performed with a Q1 resolution set to either low (top) or unit (bottom), utilising mass transmission windows of 1-1.4 atomic mass units (amu) and 0.6-0.8 amu, respectively, so as to reduce the number of non-targeted peaks detected.

From Figure 4-40, we can see that the XIC spectra obtained with a low resolution Q1 transition window (permitting ions within 1.2 ± 0.2 amu of the target m/z) was a lot more complex than the XICs obtained for either the QTrap 5500 based mixed isoform digests (Figure 4-38) or the API 2000 based A-Raf WT single isoform digests (Figure 4-29). In addition to the many non-target peaks detected, the spectra background was also found to contain excessive noise. This was most likely due to the absence of the IP enrichment step, adding the cell lysate directly to the SDS-PAGE gel, which resulted in more proteins being digested and thus more ions entering the mass spectrometer with similar Q1 and Q3 m/z ratios to those of the target peptides.

While the use of low Q1 and Q3 transition windows were required for the detection of each of the A-Raf AQUA peptides on the low sensitivity API 2000 MS, ensuring that all target ions were permitted to the detector, a more restrictive transmission window could be selected for the high sensitivity QTrap 5500, so as to reduce excessive noise.

The Q1 transmission window was thus changed from low resolution, permitting ions within 1.2 ± 0.2 amu of the target m/z , to unit resolution, permitting only ions within 0.7 ± 0.1 amu of the target m/z to the collision cell, while the Q3 transmission window was retained with a low resolution transition window. At the expense of slightly reducing the sensitivity of the QTrap 5500, this change in the Q1 transmission window should filter out the majority of the noise being permitted to the collision cell, and in doing so reducing the noise entering Q3. Indeed Figure 4-40 appears to confirm that using unit resolution in Q3 has

removed a number of non-target peaks from the XIC, further reducing the background noise against which the target peaks are quantified.

An additional route through which the background noise could be reduced would be to apply stable isotope standards with capture by anti-peptide antibodies (SISCAPA), post-proteolytic digestion. While the IP enriched A-Raf WT solution was separated via 1D-SDS-PAGE and only a segment of the gel excised (relating to the MW of the protein of interest), this gel band may still contain hundreds of proteins, in addition to the high amount of protease added to the gel. Therefore post-digestion many thousands of peptides may yet exist within the extracted sample, competing with the target peptides for ionisation, and possibly introducing interfering peaks to the MRM spectrum. So as to reduce the complexity of the extracted sample, SISCAPA utilises a 100 nanolitre (nL) nanoaffinity column, which contains immobilised antibodies specific to the target peptides, enabling any unbound peptides to be washed off the column and only those peptides of interest eluted (Anderson et al., 2004a).

Table 4-31: The quantity and variation data obtained for each A-Raf WT target peptide, detected during the spiked digestion of an A-Raf WT transfected, un-enriched, HEK293 cell lysate. Through comparing the peak areas for both the synthetic and endogenous A-Raf peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

Replicate	TVVTVR - Present in all but A-Raf Short		VSQPTAEQAQAFK - Unique to A-Raf WT		GLNQDCCVVYR - Present in all A-Raf isoforms	
	Average Replicate (pmol/ μ L)	per CoV	Average Replicate (pmol/ μ L)	per CoV	Average Replicate (pmol/ μ L)	per CoV
1	0.07	71.88%	6.12	149.01%	0.1	99.18%
2	0.03	136.02%	0.91	67.71%	0.1	106.65%
3	0.02	119.22%	6.02	134.43%	0.1	97.16%
4	0.1	83.35%	5.28	94.63%	0.09	88.95%
5	0.07	112.22%	8.44	145.54%	0.09	87.03%
Average	0.06	98.27%	5.35	135.84%	0.1	81.70%

From Table 4-31 it is clear that while high levels of peptide VSQPTAEQAQAFK were detected, 5.35 pmol of A-Raf WT per 25 μ L of transfected HEK293 cell lysate, large CoV values were also achieved (135.84%). Large CoV values were also achieved for both peptides TVVTVR and GLNQDCCVVYR, achieving some 98.27% and 81.70%, respectively. The quantitation values obtained for peptides

TVVTVR and GLNQDCCVVYR, however, were considerably lower, detecting just 60 fmol of A-Raf WT per 25 μ L of cell lysate, and 100 fmol of A-Raf WT per 25 μ L of cell lysate, respectively. When the MRM transitions were reviewed, the cause of the erratic quantitation values was determined (Figure 4-41). While AQUA peptides TVVTVR and GLNQDCCVVYR were detected at a high intensity, AQUA peptide VSQPTAEQAQAFK was missing. As such, while 5 pmol of AQUA peptides TV*VTVR and GL*NQDCCVVYR was being compared to background noise, resulting in low quantitation values, background noise was being compared to background noise for VSQPTAEQAQAFK, resulting in a much higher quantitation value.

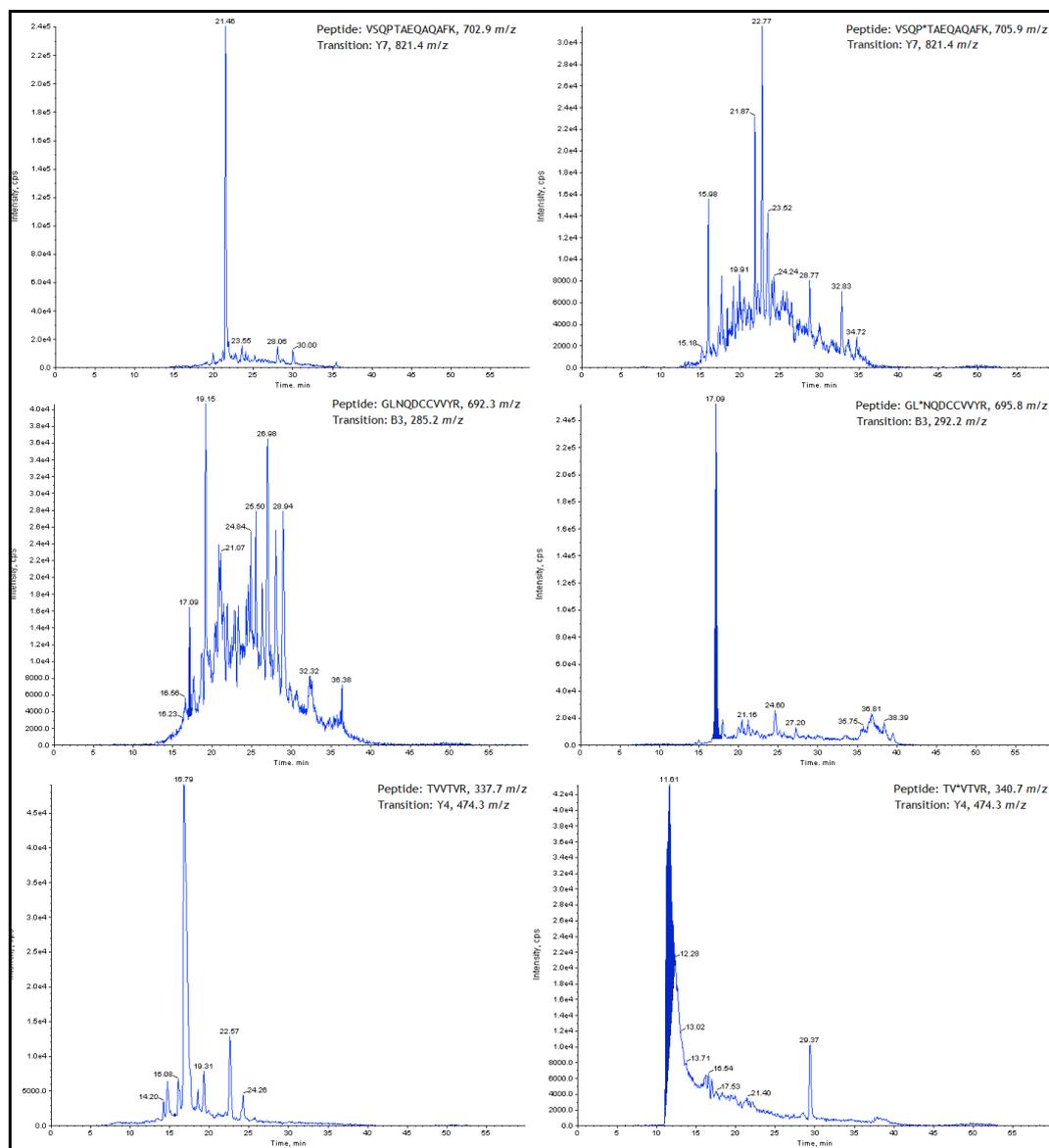


Figure 4-41: The MRM spectra obtained from both the synthetic and endogenous peaks detected during the analysis of 25 μ L of A-Raf WT transfected, non-enriched HEK293 cell lysate. The MRM peaks obtained for target and synthetic peptides VSQPTAEQAQAFK, GLNQDCCVVYR and TVVTVR following the injection of 25 μ L of A-Raf WT transfected, un-enriched, HEK293 cell lysate, on a QTrap 5500.

In conclusion, no A-Raf WT was detected during the quantification of the A-Raf WT transfected HEK293 cell lysate, suggesting the level of A-Raf WT within the non-immunoprecipitated solution to be below the LoD for this isoform on a QTrap 5500.

If we consider that the previous A-Raf WT transfected HEK293 IP contained 1.2 pmol of A-Raf WT per 25 μ L of sample, and that approximately 360 μ L of IP was produced from a total of six cell culture plates, then the total amount of transfected A-Raf WT protein, per plate, was 2.88 pmol. If the same quantity is assumed for this non-enriched sample, diluted to 1 mL in protein loading buffer, then the 25 μ L added to the SDS-PAGE gel should contain approximately 72 fmol of the A-Raf WT. While this is below the LoD for peptide GLNQDCCVVYR on the QTrap 5500, a quantification should be possible on a QTrap 4000 MS.

4.4.5.5 Endogenously Expressed A-Raf WT Spiked Digest

Further to the detection of exogenously expressed A-Raf WT from a non-immunoprecipitated cell lysate, the detection of endogenously expressed A-Raf WT was attempted from a HEK293 cell lysate.

A 150 mm plate of HEK293 cells was incubated for 96 hours, lysed, diluted to 1 mL in Protein Loading Buffer, and 25 μ L loaded to an SDS-PAGE gel. A gel band, centred on a MW of 58 kDa, was excised from the gel and digested with trypsin, in the presence of each of the four A-Raf AQUA peptides. The tryptic peptides were extracted and vacuum centrifuged, re-suspended in 20 μ L of Buffer A and injected onto an Ultimate 3000 HPLC, coupled to a QTrap 5500 MS.

As was the case with the non-immunoprecipitated exogenously expressed cell lysate, no A-Raf WT was detected during this analysis (data not shown).

4.5 Conclusion

During this project it has been possible to develop a set of AQUA peptide selection criteria, and to apply these to the selection of AQUA peptides suitable for the quantification of each of the four A-Raf isoforms. In the case of A-Raf WT this was as simple as selecting a peptide located in the catalytic domain, unique to this isoform, while for DA-Rafs 1 and 2, non-ideal C-terminus peptides were the only solution. These non-ideal peptides contained Cys and Met residues, or

were cleaved with non-ideal proteases, each challenge enabled the author to test how detrimental these sequence flaws were to an AQUA based quantification. Finally, a subtraction based quantification was performed for A-Raf Short, subtracting the estimated amount of a peptide present in A-Raf WT, DA-Raf-1 and DA-Raf-2 from that achieved for a peptide present in each of the four A-Raf isoforms.

Through characterising and quantifying these AQUA peptides it was possible to detect each Q1 ion, and to select several Q3 transitions, per peptide, in doing so creating an A-Raf MRM acquisition method capable of detecting each targeted peptide within a complex sample. Protocols were also developed for the complete modification of Cys and Met, while proteolytic digest parameters were optimised, so as to maximise the peptide yield per proteolytic digestion.

Upon development of the MS based analysis, the LoD was determined for each A-Raf isoform on a range of triple quadrupole based mass spectrometers, and each A-Raf isoform expressed at a level suitable for spiked digest based quantitative analyses.

Several problems were encountered during this project however. The first of which was the MS instrumentation available for use. As the LoD test showed, the API 2000 MS was 300 times less sensitive than the QTrap 4000 MS. Should a more sensitive instrument have been available, then the detection of each of the endogenously expressed A-Raf isoform should have been possible. Secondly, problems were encountered with AQUA peptide GL*NQDCCVVYR. While peptide GLNQDCCVVYR was intentionally ordered in a reduced state, so as to enable a comparable IAA based alkylation to be performed on both the target and synthetic peptides, this additional processing led to a 25% sample loss.

While several problems were encountered during this project, slowing the development of this quantitative technique drastically, the theory behind this analysis remains sound, and indeed some promising results were obtained during the spiked digest based analyses. Through simply re-ordering the Cys containing

AQUA peptides pre-alkylated, utilising SISCAPA base peptide purification and performing the quantitative analyses on a sensitive triple quadrupole based MS, the detection of endogenous A-Raf within a HEK293 cell lysate should be possible.

5 Absolute Quantification of the PDE4B Group of Protein Isoforms

5.1 Introduction

5.1.1 cAMP and PKA Activation

Adenine ribonucleotide, later named 3'-5'-cyclic adenosine monophosphate (cAMP), was first discovered in 1957 when Sutherland and Rall subjected dog liver fractions to homogenisation in the presence of adenosine-5'-triphosphate (ATP), Mg^{++} and epinephrine or glucagon (Sutherland and Rall, 1957). By way of an explanation the author suggested that the cAMP may be acting as a secondary messenger within the cell (Sutherland and Rall, 1958).

Several discoveries have since been made in the field of intra-cellular signalling; one of which was the discovery that cAMP is generated through the conversion of cytoplasmic ATP by adenylyl cyclase (Gary M, 1988). In turn, it was discovered that distinct ligands were capable of elevating the production of cAMP, suggesting the existence of hormone specific receptors (such as for epinephrine or norepinephrine) (Orly and Schramm, 1976). Previously each hormone had been thought to bind adenylyl cyclase directly. It was also noted that GTP was required for the activation of adenylyl cyclase, leading to the development of a new signalling model in which an adenylyl cyclase-stimulatory GTP-binding protein (G-protein) was required for the production of cAMP (Dohlman et al., 1991). This G-protein was first purified and cloned in 1987 by Gilman (Gilman, 1987), who noted it to be composed of three distinct subunits, the α , β , and γ , each of which were present in a 1:1:1 ratio and named in order of decreasing mass.

In the most recent model of cAMP production, the primary signalling molecule binds to activate its specific extracellular receptor, creating a receptor based binding site for protein G. By binding with the receptor, a conformational change occurs within the G-protein, causing the α subunit to exchange a bound GDP for a cytoplasmic GTP, and in doing so disassociating from both the β and γ

subunits (Patel et al., 2001). Having activated the G-protein, the free α subunit binds adenylyl cyclase, enabling the conversion of ATP to cAMP (Dohlman et al., 1991), as is shown on Figure 5-1.

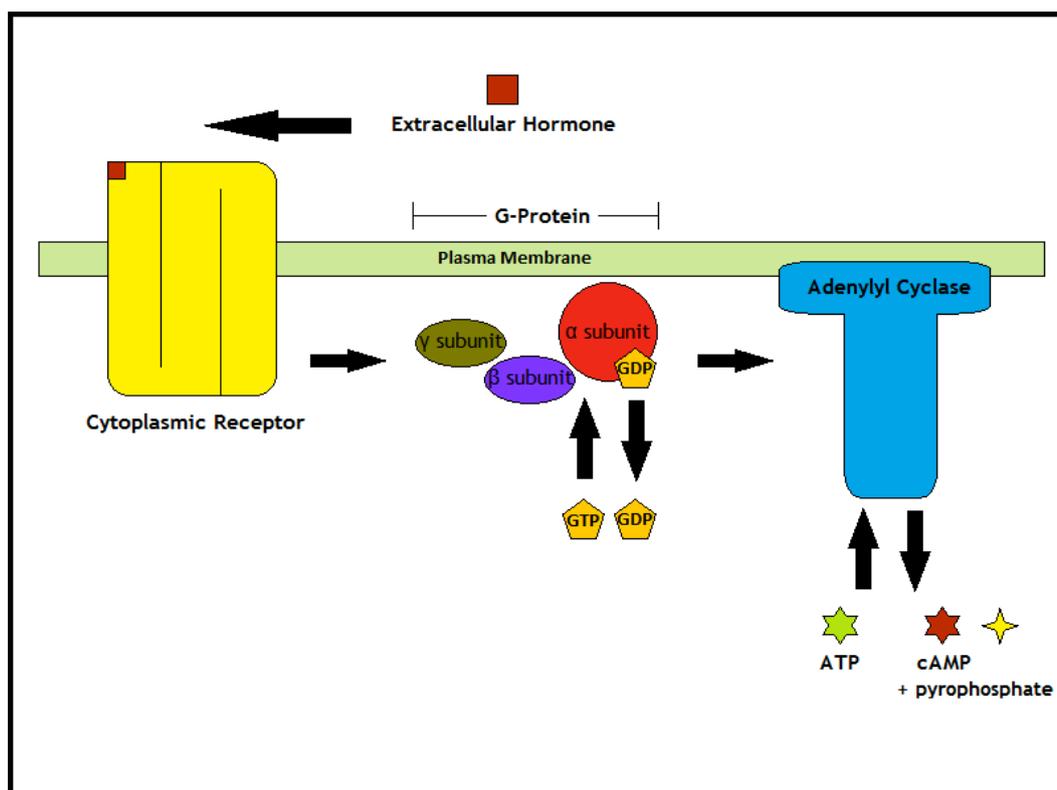


Figure 5-1: **The signalling cascade required for the production of intercellular cAMP.** This image depicts how an extracellular hormone can lead to the production of intercellular cAMP, which requires the extracellular hormone binding its specific receptor. This binding causing the α subunit of the G-protein to exchange a bound GDP molecule for a GTP, and in doing so disassociating from both the β and γ subunits. The α subunit of the activated G-protein can then bind the adenylyl cyclase, initiating the conversion of cytoplasmic ATP to cAMP.

Upon uncovering the mechanism of cAMP generation, and unhappy with a theory in which cAMP dependent protein kinase A (PKA) was “swimming about, happily phosphorylating a variety of cellular constituents” (Beavo and Brunton, 2002), it was suggested that cAMP based signalling may be compartmentalised. Were this proven to be the case, then it was theorised that the hormone specific receptors, nine of which had been identified, would be located within close proximity to the final effector (TASKÉN and AANDAHL, 2004). This was later

confirmed through experimentation with prostaglandin E and the β -adrenergic receptor agonist isoprenaline, both of which resulted in equivalent elevation of cardiac cAMP levels, but only isoprenaline activated glycogen phosphorylase (Beavo and Brunton, 2002).

Upon the production of cAMP, PKA is the classical cAMP effector protein, responsible for transmitting the secondary signal (Houslay and Adams, 2003). PKA exists within the cell as a tetramer, composed of two regulatory (R) subunits, held together by N-terminal dimerisation domains, each of which is independently bound to a catalytic (C) subunit (Francis et al., 2011). Several R and C subunit isoforms have been identified, with some four genes encoding subunit R (RI- α , RI- β , RII- α and RII- β), and three encoding subunit C (C- α , C- β and C- γ). Due to the range of possible isozymes, PKA can be classified as either a Type I or Type II enzyme; based on which R subunit it contains. Type I enzymes (RI) are predominantly located within the cytoplasm, while Type II enzymes (RII) have been identified on the cell membrane and organelles (Kim et al., 2007). This range of PKA isoforms helps to further increase the substrate specificity of the kinase (Francis et al., 2011).

The final method by which the specificity of PKA is increased is via binding to an A-kinase anchoring protein (AKAP), a scaffolding protein which maintains PKA within close proximity to its final substrate pool (Edwards and Scott, 2000). Type II PKA tetramers are bound strongly to AKAPs, with a nM affinity. In contrast, Type I PKA tetramers only display μ M levels of affinity towards the ligand, and have not as yet been shown to bind AKAP *in vitro* (Beavo and Brunton, 2002, Houslay and Adams, 2003, Kim et al., 2007).

In terms of PKA activation, each tetramer binds four cAMP molecules (two cAMP molecules per R subunit), causing the catalytic subunits to dissociate from the regulatory subunits. This conformational change activates the PKA, enabling phosphorylation, as is shown on Figure 5-2 (Kim et al., 2007, Lugnier, 2006).

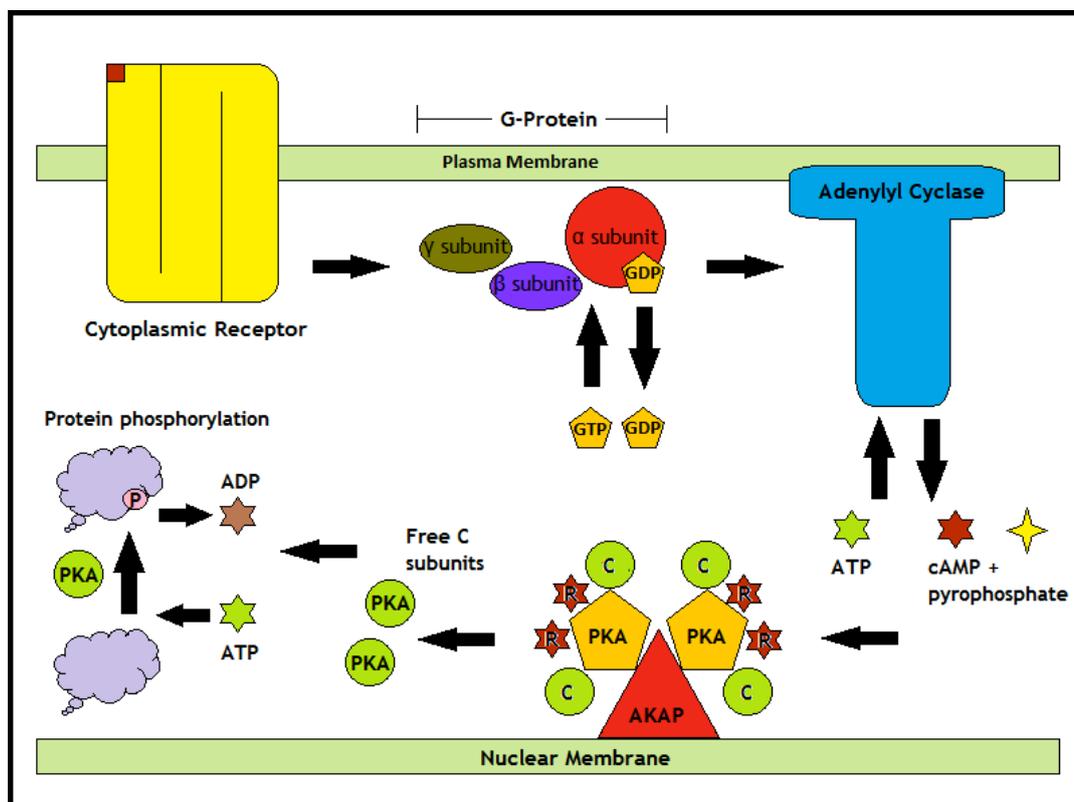


Figure 5-2: The signalling cascade required for the activation of PKA and subsequent protein phosphorylation. This image depicts how the production of cAMP, can activate AKAP-bound PKA, causing the catalytic subunits to dissociate from the regulatory, enabling protein phosphorylation to proceed. In brief, the extracellular hormone binds its specific receptor, causing the α subunit of the G-protein to exchange a bound GDP molecule for a GTP, and in doing so disassociating from both the β and γ subunits. The α subunit of the activated G-protein can then bind the adenylyl cyclase, initiating the conversion of cytoplasmic ATP to cAMP. Upon the release of cAMP into the cytoplasm it can bind an AKAP-bound PKA tetramer, causing the catalytic subunits (C) to dissociate from the regulatory (R), activating PKA and enabling it to phosphorylate protein substrates through the conversion of ATP to ADP.

Once created, the only route through which cAMP can be degraded, and thus the signalling event terminated, is through the conversion of the cAMP to 5'-AMP by cAMP degrading phosphodiesterases (PDE) (Houslay and Adams, 2003). These PDE proteins are anchored throughout the cytosol and nuclear and plasma membranes, degrading cAMP in such a way so as to create an asymmetric gradient of cAMP localised within the vicinity of the final effector, as is shown on Figure 5-3. These cAMP gradients are “read” by AKAP-bound PKA molecules,

leading to gene expression and phenotypic changes which mark the completion of the signalling cascade (Houslay et al., 1998, Miles D, 2001).

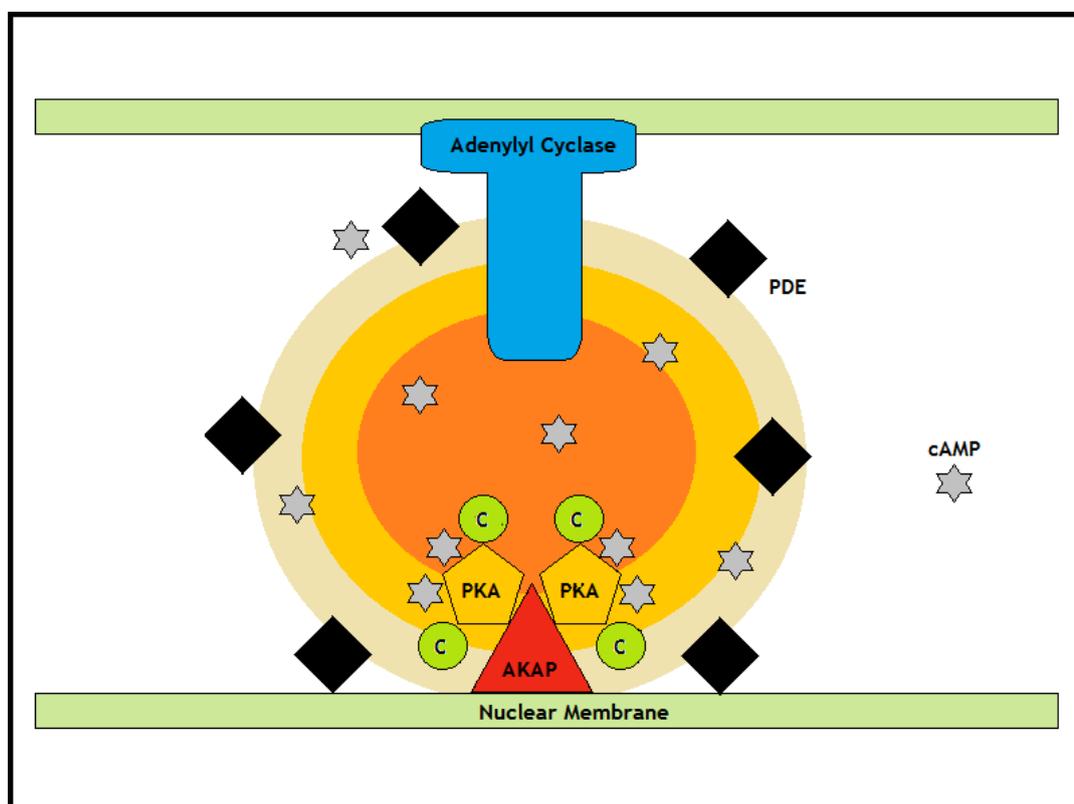


Figure 5-3: cAMP gradient compartmentalisation, based upon the expression of PDE. Depicted above is the route through which cAMP is compartmentalised exclusively to the region of the final effector. Through the expression of PDE, and thus the degradation of cAMP, cAMP gradients can be established within the cell.

By the late 1970s and early 1980s the possibility of stimulating or inhibiting these distinct PDE isoforms through the application of small organic molecules had been proposed (LEVIN and WEISS, 1976, Wells et al., 1981). This was of particular interest to the therapeutic drugs industry as PKA signalling and PDE catalysed signal degradation had been known to influence a wide range of pharmacological processes, including the production and action of proinflammatory mediators, ion channel functioning, muscle contraction, learning, cell differentiation, apoptosis, lipogenesis, glycogenolysis and gluconeogenesis (Perry and Higgs, 1998). By targeting the actions of these PDE enzymes it appeared possible to develop drugs for the treatment of heart

failure, depression, asthma, inflammation and erectile dysfunction (TORPHY, 1998). Furthermore, as each PDE isozyme appeared to exhibit a unique tissue expression pattern, then these therapeutic agents should be organ and tissue specific (TORPHY, 1998).

5.1.2 cAMP Specific Phosphodiesterases

The first research to be carried out on a PDE was undertaken by Butcher and Sutherland in 1962, identifying a magnesium dependent phosphodiesterase capable of cAMP degradation (Butcher and Sutherland, 1962). Furthermore, the authors identified this specific PDE species to be inhibited by methyl xanthenes, and stimulated by imidazole (Butcher and Sutherland, 1962). Following this discovery, it was hypothesised that there may be a number of different isoforms (Boswell-Smith et al., 2006), which was confirmed in 1970 when Beavo *at al.* characterised three PDE isoforms from bovine and rat tissues based on anion-exchange chromatography (BEAVO et al., 1970). These isozymes were calcium-calmodulin (CaM)-PDE, cAMP-PDE and cGMP-PDE, each named in regard to their substrate specificity and sensitivity to CaM, and numbered based on the order in which they eluted from the LC column (Boswell-Smith et al., 2006). By performing similar PDE based separations on a range of tissues (BEAVO et al., 1970, Wells et al., 1975, Butcher and Sutherland, 1962), and with the advent of the molecular age (Boswell-Smith et al., 2006), the number of PDE isoforms identified grew extensively. By 1995 this amounted to seven different gene families having been identified in mammalian tissues, with most families presenting several distinct genes and isoforms (Beavo, 1995, Boswell-Smith et al., 2006). This increasingly complex set of data and its associated naming convention led to several problems. For example, often when PDE isoforms eluted they didn't display the biochemical characteristics with which they were associated. While this led to some confusion, it was found to be due to the varying complement of PDE isoforms present in each tissue and species (TORPHY, 1998). Further to this, it was also common for a PDE isozyme to be named based on its inhibitor, activator or the ligand to which it preferentially bound; some examples of which include: ROI-PDE (rolipram-inhibited PDE), CaM-

PDE (calmodulin-activated PDE), cGS-PDE (cGMP-stimulated PDE), and cGB-PDE (cGMP-binding PDE) (Beavo, 1995, Lugnier, 2006). Therefore it was decided that an official PDE naming convention was required, a process which was initiated in 1995 by Beavo (Beavo, 1995). Under this convention, the first two letters of a protein's name indicated the species from which the gene was extracted, HS, for example, indicating a *Homo sapiens* origin. Further to this, the next three letters and one or two Arabic numerals indicated the gene family, and thus the elution order previously assigned to the protein, an example of which is PDE4 (HSPDE4 thus far). The penultimate letter in this nomenclature related to the specific PDE gene, for example gene B within the PDE4 family (HSPDE4B thus far), while the final Arabic numeral indicated the specific PDE isoform, for example isoform splice pattern 1, giving a final designation of HSPDE4B1 (Beavo, 1995, Lugnier, 2006).

In regards to this project, all further PDE isoforms discussed are of a *Homo sapiens* origin, unless otherwise stated, and as such the naming convention will be shortened through the removal of the first two letters, indicative of species.

5.1.3 The Physical Properties of the PDE Family

As of 2011, a total of 11 mammalian PDE families had been described, giving rise to over 20 unique genes and some 50 protein isoforms (Jeon et al., 2005). Each of these isoforms has been sequenced, biochemically analysed and pharmacologically characterised. Furthermore, an NCBI gene reference for PDE12 (NCBI:201626) was created on the 9th October 2011, suggesting there may yet be many more unknown PDE families, genes and isoforms.

Each of the 11 PDE families has been shown to share a basic common structure, which includes a catalytic domain, and one or more regulatory domains, located between the N-terminus and the catalytic domain (Lugnier, 2006). The PDE catalytic domain has been shown to consist of some 270 amino acids, homologous in nature and highly conserved throughout evolution (Lugnier, 2006, Zoraghi et al., 2004). Indeed, it is because of this distinct yet conserved

catalytic domain that each of the PDE isoforms are thought to have diverged from a common ancestor some 940 million years ago, before the separation of sponges and eumetozoans (Koyanagi et al., 1998). Research into the structure of the *Homo sapiens* PDE catalytic sub-units found only a 50% sequence identity across each family (David M, 1999). Similarly, research conducted in vertebrate and insect species found only a 30% sequence identity across each PDE family, but a 60-90% sequence identity when comparing any single PDE family between species (Graeme B, 1994). This shows just how important the conservation of each PDE family has been in evolution, while confirming just how great the differences are between each of the PDE species (Lugnier, 2006).

Of particular interest to PDE based research are the structural sub-domains of the catalytic core, frequently targeted by therapeutic agents, so as to allosterically regulate targeted PDE isozymes (TORPHY, 1998). Over the last decade this research has unveiled three helical sub-domains, including an N-terminal cyclin-folding region, a linker region, and a C-terminal helical bundle. Where these three regions interface, a deep hydrophobic pocket is formed. This pocket is composed of four sub-sites; including a metal-binding site (M site), a core pocket (Q pocket), a hydrophobic pocket (H pocket) and a lid region (L region) (Sung et al., 2003, Francis et al., 2000).

In comparison to the highly conserved catalytic domain, each of the 11 PDE families have been shown to contain unique regulatory regions, including a calmodulin binding site, unique to PDE1, an allosteric cGMP binding site, present in PDE species PDE2, PDE5, PDE6, PDE10, and PDE11 (Zoraghi et al., 2004), a phosphatidic binding site, unique to PDE4, a PAS domain, unique to PDE8, autoinhibitory sequence, unique to PDE species PDE1 and PDE4, and a membrane association domain, present on PDE species PDE2, PDE3 and PDE4. Furthermore, each PDE family has been found to contain various unique phosphorylation sites and dimerization motifs (Bolger, 1994).

5.1.4 PDE4

PDE4, previously termed cAMP-PDE, is characterised by its specificity towards cAMP. It is also potently and specifically inhibited by rolipram (Francis et al., 2000). PDE4 expression has been identified in the smooth muscle cells of the airway, within the brain and within the tissues of the cardiovascular system (Muller et al., 1996). Furthermore, PDE4 has been shown to be the predominant PDE species expressed within inflammatory cells, and has been implicated in inflammatory airway disease (Wang et al., 1999, Houslay et al., 1998).

The PDE4 family of protein isoforms are the most widely characterised of the PDE families, though it was not until the advent of molecular cloning during the 1990s that the extraordinary diversity of this family was revealed (Perry and Higgs, 1998). Four genetically distinct PDE4 genes have so far been identified, termed PDE4A-D, each spanning some 50 kb and containing at least 18 exons (Houslay and Adams, 2003). These genes are thought to give rise to at least 35 different isoforms through the alternative splicing of each protein's N-terminal (Boswell-Smith et al., 2006, Muller et al., 1996, Bolger, 1994).

In terms of structure, each PDE4 isozyme has been shown to contain the same three functional domains; the PDE catalytic domain, which accounts for between 18 and 46% of the protein and is located close to the protein's C-terminal (Jeon et al., 2005), and two regulatory sub-domains. The first of which, upstream conserved region one (UCR1), consisting of some 60 amino acids and located close to the proteins N-terminal, while the second, upstream conserved region two (UCR2), consists of some 80 amino acids and is located between the first regulatory domain and the catalytic core (Thompson, 1991, Bolger, 1994), as is depicted on Figure 5-4.

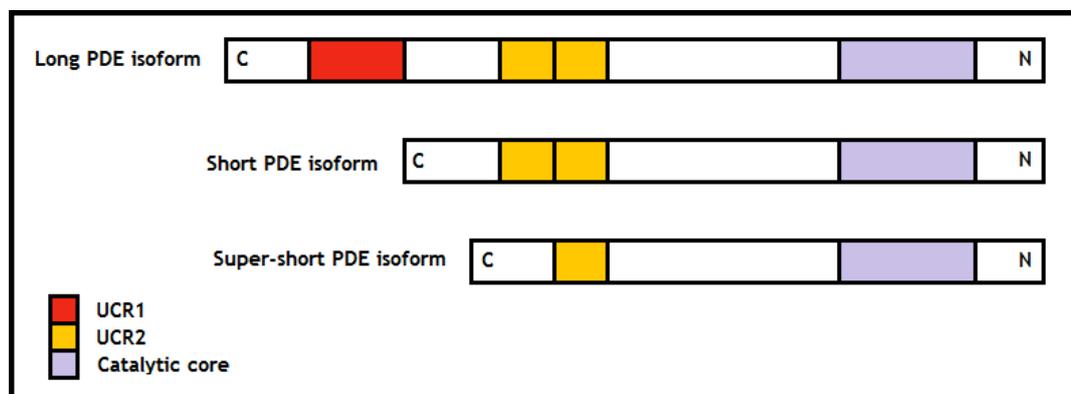


Figure 5-4: The sequence orientation of typical long, short and super-short PDE isoforms. Depicted above are the typical structures of long, short and super-short PDE4 isoforms, detailing how long PDE4 isoforms contain both UCR's 1 and 2, whereas short PDE4 isoforms contain only UCR2. Finally super-short PDE isoforms contain only half of UCR2, while each of the PDE4 isoforms contains the conserved catalytic core.

PDE4 isoforms found to contain both UCRs 1 and 2 have been termed long isoforms, whereas those found to contain only UCR2 have been termed short-isoforms. Super-short PDE4 isoforms have also been documented, containing only a truncated version of UCR2 (Houslay et al., 1998). Of these conserved regulatory domains, it has been suggested that it is UCR1 which is the most heavily regulated, involving allosteric regulation, membrane targeting and catalytic inhibition. UCR2, meanwhile, is thought to be involved in the formation of PDE4 dimers (Graeme B, 1994), a theory which complements the work of Zoraghi *et al.*, who suggests PDE to be dimeric, but concluded that the function of this dimerisation was, as yet, poorly understood (Zoraghi et al., 2004).

Research by Oki *et al.* has shown long PDE4 isoforms to be stimulated by the PKA catalysed phosphorylation of Ser-54 within UCR1 of PDE4D3. This conformational change has been shown to increase the kinase activity of all long PDE isoforms by approximately 60%. This increase in kinase activity is thought to act as a short-term feedback mechanism, helping to restore basal levels of cAMP within the cell following a signalling event (Oki et al., 2000). Richter and Conti have also demonstrated PKA catalysed phosphorylation to be dependent on long PDE isoform dimerisation. This dimeric state, it is thought, may help to stabilise the PDE4 isoforms in a high-affinity rolipram binding state (Richter and Conti, 2004).

In the early 1970s the characteristic PDE4 inhibitor rolipram was tested as a potential therapeutic agent (Scott et al., 1991). This research was conducted around the premise that elevated levels of cAMP may enhance noradrenergic neurotransmission within the central nervous system. Rolipram was thus seen as a potential treatment for depression, which, while proven effective, was shown to cause nausea and disturbance of the gastrointestinal tract long before any statistically significant therapeutic effect could be established (Scott et al., 1991). These side effects were attributed to “off target” events, which characterise general PDE4 inhibitors which inhibit each of the 35 isoforms equally, due to their action on the highly conserved catalytic site. Further research into PDE4 inhibition found it to block cell trafficking, cell proliferation and to attenuate the production of inflammatory mediators, cytokines and reactive oxygen species (Souness et al., 2000), while suppressing the immune system (Boswell-Smith et al., 2006).

One physical property of the PDE4 family which does lend itself toward further drug development is the distinct cellular and tissue specific distribution pattern displayed by each of the four PDE4 isoforms. Therefore, should one specific PDE isoform be targeted by a drug treatment, any side effects should be minimised. For example, PDE4B appears to be expressed in the heart, brain, skeletal muscle, and lung, but not in the placenta, liver, kidney, or pancreas (McLaughlin et al., 1993). PDE4C meanwhile appears to be expressed in neuronal tissues but not in immune or inflammatory cells (Engels et al., 1995, Engels et al., 1994), while PDE4D appears to be expressed in the blood fractions and skeletal muscles (Cheung et al., 2007). The exception in this case is PDE4A which appears to be expressed ubiquitously, though slightly higher levels of expression have been noted in monocytes, the skeletal muscles, the testis, and the pons (Bolger et al., 1993, Obernolte et al., 1993, Livi et al., 1990).

It should be noted that isoform specific PDE4 inhibitors which compete with cAMP for the active site have proven elusive due to the conserved nature of the PDE4 catalytic unit. Several PDE4 inhibitors have, however, entered clinical trials more recently, aiming to treat reactive airway diseases, autoimmune

diseases, B-cell malignancies, chronic lymphocytic leukemia and acute lymphoblastic leukemia (Smith et al., 2005).

5.1.4.1 PDE4B

The sequence which encodes the PDE4B family of protein isoforms is thought to be located on chromosome 1p31.2-32.1, spanning some 580 kb, and comprising at least 17 exons (Otowa et al., 2011). Alternative splicing of this gene has been shown to give rise to at least five protein isoforms (PDE4B1-5), each varying in length and regulation, as a result of a unique N-termini. The transcription of each of the five PDE4B isoforms is thought to be initiated through unique upstream promoters, each transcribing a unique 5' exon before incorporating a shared catalytic C-terminal (Wallace et al., 2005).

As was discussed previously, PDE4B appears to be expressed in the heart, brain, skeletal muscle, lung (McLaughlin et al., 1993) and blood fractions (Wallace et al., 2005). In terms of expression within the central nervous system, PDE4B has been detected in the amygdala, hypothalamus and frontal cortex (Cherry and Davis, 1999), regions which are thought to be key in the mediation of stress and anxiety (Charney and Deutch, 1996). While many members of the PDE4 family are expressed throughout the brain, including those areas associated with reward and affect (Iona et al., 1998, Cherry and Davis, 1999), the PDE4B family is thought to be of particular importance for the transduction of complex signals and brain integration (Menniti et al., 2006). Furthermore, PDE4 has been found to account for approximately 33% of the total cAMP hydrolytic activity within mouse heart (Leroy et al., 2011).

Recent publications have further defined the role of PDE4B within the brain to include the regulation of anxiety and depression (Pandey et al., 2005), with PDE4 inhibitor rolipram exhibiting antidepressant (O'Donnell and Frith, 1999) and anxiolytic-like behaviours in animals (Silvestre et al., 1999). As a way of confirming this function Rutten *et al.* prepared PDE4B *-/-* knockout mice, commenting that the animals showed “enhanced basal postsynaptic responses to

stimulation and long-term depression” (Rutten et al., 2011). PDE4B has also been linked to multifunctional scaffolding protein Disrupted-in-Schizophrenia-1 (DISC1), which is thought to be required for the parallel processing of signals from GSK-3 and PDE4. DISC1 is thought to play a role in the regulation of neurodevelopment, signal transduction and intracellular transport and exocytosis within the brain (Jaaro-Peled et al., 2010). Furthermore, DISC1 has been seen as a candidate susceptibility factor for psychiatric disorders such as schizophrenia, schizoaffective disorder, bipolar disorder and depression (Millar et al., 2005). Indeed post-mortem brain samples obtained from patients with schizophrenia and bipolar disorder have suggested that each of these conditions may be linked to the expression pattern of PDE4B, which appears to differ from that of healthy individuals (Fatemi et al., 2008, Kähler et al., 2010, Numata et al., 2009a).

5.1.4.1.1 PDE4B1

PDE4B1 is a long PDE isoform, incorporating both UCRs 1 and 2. PDE4B1 is thought to span some 736 amino acids (Bolger et al., 1993), and to present with a single SDS-PAGE gel band with a MW of approximately 104 kDa (Shepherd et al., 2003). Unlike the majority of the PDE4B family, PDE4B1 appears not to be expressed at a high concentration within human brain (Dwivedi, 2010). Instead PDE4B1 has been implicated in the transduction of cAMP based signalling within osteoblast cells; where glucocorticoid dexamethasone (DEX) based immunosuppressive treatments have been shown to reduce the levels of PDE4B1 transcription. Following long-term treatment with DEX, the metabolic activity of these osteoblast cells appears modified and the life span of the cells reduced through inhibiting bone formation (Ahlström et al., 2005).

5.1.4.1.2 PDE4B2

PDE4B2 is a short-PDE isoform, and thus is regulated solely through UCR2 and cannot be phosphorylated by PKA. PDE4B2 is thought to span some 564 amino acids (Bolger et al., 1993), and to present with a single SDS-PAGE gel band with a MW of approximately 78 kDa (Shepherd et al., 2003). The *Homo sapiens*

consensus sequence for PDE4B2, appears to be of a length identical to, and to display 90% sequence identity with, its counterpart in *Rattus norvegicus* (Bolger, 1994).

Unlike PDE4B1, PDE4B2 has been shown to be expressed at a high concentration within the human brain (Dwivedi, 2010). However, PDE4B2 expression within the prefrontal cortex and hippocampus appears to be decreased in patients suffering from depression (Dwivedi, 2010) and schizophrenia (Fatemi et al., 2008). Several single nucleotide polymorphisms (SNPs) linked to schizophrenia have also been identified in the introns bracketing the PDE4B2 coding locus in Caucasian and African American populations (Fatemi et al., 2008), again proving this isoform to be of critical importance to mental health.

PDE4B2 inhibition has also been seen as a potential treatment for diffuse large B-cell lymphoma (DLBCL), where PDE4B2 has been identified as the predominant PDE4B isoform in normal B cells, which when unabated prevents cAMP induced apoptosis (Smith et al., 2005).

5.1.4.1.3 PDE4B3

PDE4B3, much like PDE4B1, is a long PDE isoform, incorporating both UCRs 1 and 2. PDE4B3 is thought to span some 721 amino acids (Huston et al., 1997) and to yield a single SDS-PAGE gel band with a MW of approximately 103 kDa (Shepherd et al., 2003). Like PDE4B2, the consensus sequence for *Homo sapiens* PDE4B3 appears to have a length identical to, and to display a 90% sequence identity with, its counterpart in *Rattus norvegicus* (Bolger, 1994).

While PDE4B2 has been shown to be highly expressed within the human brain, PDE4B3 has been shown to be highly expressed within the brain, heart, lung and liver (Dwivedi, 2010). PDE4B3 expression has been shown to be decreased in patients suffering from bipolar disorder, but not from depression or schizophrenia, suggesting isoform specificity in mental illness (Fatemi et al., 2008). PDE4B3 has also been shown to be of importance for hippocampal long-

term potentiation (LTP) during learning, with PDE4B3 expression being up-regulated just two hours after tetanisation, peaking after six and rapidly degrading after eight (Fatemi et al., 2008).

5.1.4.1.4 PDE4B4

Whether PDE4B4 is expressed in *Homo sapiens* is a hotly debated topic first addressed by Shepherd *et al.* in 2003. While Shepherd failed to identify PDE4B4 in human tissues, a previously unpublished 659 amino acid long-PDE4B isoform was successfully identified in *Rattus norvegicus*. This new PDE4B isoform was found to express a unique 17 amino acid N-terminal, while incorporating the same regulatory and catalytic domains as are present in PDE4B1 and PDE4B3 (Shepherd et al., 2003). Upon analysis, this novel PDE4B3 coding mRNA sequence in *Rattus norvegicus* was found to be expressed in multiple tissues, including the liver and brain. BAC screening of *Mus musculus* chromosome four meanwhile, identified a gene with a sequence identity of greater than 95% to that of the *Rattus norvegicus* PDE4B4 gene. Indeed, while several *Homo sapiens* BACs were also found to contain a similar sequence, every possible reading frame was found to contain tightly clustered stop codons. Shepherd *et al.* thus concluded that the gene coding for PDE4B4 in *Homo sapiens* amounted to a dead exon (or pseudogene), coding a protein of function in rodents, but not humans (Shepherd et al., 2003). Several authors however have reported the detection of PDE4B4 in *Homo sapiens* tissues (Farooqui et al., 2000, Takahashi et al., 1999), including Fatemi *et al.* who detected three long PDE4B isoforms (PDE4B1, PDE4B3, and PDE4B4) and one short-PDE4B isoform (PDE4B2) through western blotting (Fatemi et al., 2008), and Braun *et al.* who reported the detection of a distinct 85 kDa protein band in several human tissues when western blotting with two well characterized antibodies (Braun et al., 2007).

In terms of function, Fatemi *et al.* observed a significant decrease in the levels of PDE4B4 transcription in patients suffering from schizophrenia (Fatemi et al., 2008), while Shepherd *et al.* noted PDE4B4 expression to be restricted to the cytosol of transfected COS7 cells. This was seen as being particularly interesting

as each of the previously identified long PDE4B isoforms were all membrane associated. Shepherd *et al.* did acknowledge however that this change in localisation may be due to the COS7 cells not expressing the correct PDE anchoring protein (Shepherd *et al.*, 2003).

5.1.4.1.5 PDE4B5

PDE4B5 is a super-short PDE4 isoform discovered by Cheung *et al.* in 2007. PDE4B5 is thought to span some 484 amino acids, and to include a unique 16 amino acid N-terminal region and a truncated UCR2, while maintaining the same catalytic subunit present in each of the PDE4B isoforms (Cheung *et al.*, 2007). High levels of PDE4B5 expression have been noted in brain (fetal brain, cerebellum, frontal lobe, pons, putamen, thalamus, and hippocampus), and weaker levels of expression have also been noted in the retina, spinal cord, pituitary, fetal kidney, jejunum, ileum, lung carcinoma A549 cells, testis, HeLa cells, and G361 melanoma cells. PDE4B5 has also been shown to be conserved in vertebrae, chicken, frog, zebrafish, and fugu; suggesting it may play a strong functional role. Furthermore, the N-terminal region unique to PDE4B5 was also identified in PDE4D6, suggesting this sequence to be both functional and highly conserved (Cheung *et al.*, 2007).

In regard to the function of PDE4B5, the uneven distribution pattern of this isoform within the cytosol suggests that PDE4B5 may play a role within the cytosolic vesicles and complexes (Cheung *et al.*, 2007).

5.1.5 PDE4B Literature Overview

While the first PDE isoform was identified some 50 years ago (Butcher and Sutherland, 1962), it was not until the advent of molecular cloning during the 1990s that the extraordinary diversity of PDE isoforms was revealed (Perry and Higgs, 1998). As of 2011 this has resulted in the description of some 11 mammalian PDE families, giving rise to over 20 unique genes and at least 50 PDE isoforms (Jeon *et al.*, 2005). While many PDE isoforms have been identified, and

their expression patterns mapped to a range of tissues, very little information exists about their specific function, aside from degrading cAMP. For example, PDE4B1 is not highly expressed within the brain, but has instead been linked to cell signalling within osteoblast cells (Ahlström et al., 2005).

The focus of PDE based research appears aimed at entire PDE families or genes, with PDE4 having been characterised by its potent inhibition through therapeutic agent rolipram, and thus PDE4 has been implicated in depression (Scott et al., 1991). Likewise PDE4B has been linked to multifunctional scaffolding protein DISC1, and as such has been implicated in both schizophrenia and bipolar disorder (Fatemi et al., 2008, Kähler et al., 2010, Numata et al., 2009a).

In keeping with the functions assigned to PDE4 and PDE4B, PDE4B2 has been linked to depression and schizophrenia, and PDE4B3 linked to bipolar disorder (Fatemi et al., 2008, Dwivedi, 2010). It should be noted however that the function of these isoforms was determined based on the probing of tissues obtained from patients with mental disorders specifically, and so other cellular roles may exist.

The existence of long-isoform PDE4B4 has proven controversial, with several authors having identified the isoform (Braun et al., 2007, Fatemi et al., 2008, Farooqui et al., 2000, Takahashi et al., 1999), despite all possible N-terminal reading frames being noted as to contain tightly clustered stop codons (Shepherd et al., 2003). Where PDE4B4 expression has been identified, the isoform has, like short-isoform PDE4B2, been implicated in schizophrenia (Fatemi et al., 2008), which is interesting as PDE4B4 appears to be localised exclusively to the cytoplasm, a location generally assigned to the short-PDE isoforms (Shepherd et al., 2003).

Finally, PDE4B5 has only recently been identified, and as yet no specific function has been assigned to this super-short isoform. However, the uneven distribution pattern identified for PDE4B5 within the cytosol, suggesting the isoform may be

involved in signalling within the cytosolic vesicles or complexes (Cheung et al., 2007).

While the discovery of PDE4B5 as recently as 2007 showed that new PDE isoforms continue to be discovered, the registering of a new PDE family, PDE12, on 9th October 2011 suggests our knowledge of these protein families may yet be in its infancy.

5.2 Project Aims

Despite there being very few peer reviewed publications confirming the function and tissue specific distribution pattern of each of the five PDE4B isoforms, they should make excellent candidates for this project, enabling a range of human tissues to be probed, so as to identify: (i) Which PDE4B isoforms are expressed within a tissue, and indeed whether PDE4B4 is expressed at all (through detecting isoform specific target peptides). (ii) The concentration at which each PDE4B isoform is expressed (through detecting both the isoform specific target and synthetic AQUA peptide), and (iii) How the PDE4B isoform expression ratio varies under different physiological conditions (including in patients suffering from mental disorders).

As the goals listed above are beyond the scope of a PhD project, a more achievable set of project aims were devised. As such, during this project I aim to: (i) Screen for each of the PDE4B isoforms, identifying and selecting suitable AQUA peptides for the absolute quantification of each PDE4B protein. (ii) Characterise each AQUA peptide/intrinsic target peptide so as to develop a suitable triple-quadrupole based MRM acquisition method for the identification of each peptide within a single LC-MS run. (iii) Express suitable amounts of each PDE4B isoform, enabling the proteolytic digestion of each isoform to be optimised using a range of commercially available proteases, and (iv) Assess the accuracy and reproducibility of the AQUA technique through quantifying each protein solution expressed with a suitable number of replicates.

5.3 PDE Specific Methods

5.3.1 Proteolytic Digest Optimisation

5.3.1.1 Arg-C Digest Optimisation

Lyophilised BSA was re-suspended in 500 mM AMBIC, containing 0.1% RapiGest SF surfactant (w/v), reduced with DTT and alkylated with IAA, prior to the IAA being quenched with excess Cys. Following vacuum centrifugation, the BSA was re-suspended and digested with Arg-C at an enzyme:substrate ratio of 1:10, 1:20, 1:50, 1:100 and 1:200 (w/w), as per the manufacturer's instructions. Each digested BSA sample was vacuum centrifuged to completion and re-suspended in 2% MeCN, 0.1% FA (v/v), prior to analysis on an API 2000 MS. The resulting data files were analysed via MASCOT. The percentage sequence coverage of BSA was recorded.

5.3.1.2 Probing for Residual Glu-C Catalysed Asp Cleavage in AMBIC

Lyophilised BSA was re-suspended in 500 mM AMBIC, containing 0.1% RapiGest SF surfactant (w/v), reduced with DTT and alkylated with IAA, prior to the IAA being quenched with excess Cys. Following vacuum centrifugation, the BSA was re-suspended in 25 mM AMBIC and digested with Glu-C at an enzyme:substrate ratio of 1:20 (w/w), as per the manufacturer's instructions. The digest proceeded for between 10 and 20 hours, with a vial being removed every hour and the digest terminated through the addition of FA. Each digested BSA sample was vacuum centrifuged to completion and re-suspended in 2% MeCN, 0.1% FA (v/v), spotting each solution to a 384 well MALDI plate. PepMix II calibration standard and HCCA matrix was added to each MALDI spot, prior to analysis on an Ultraflex II MS.

5.3.2 PDE4B Plasmid Preparation

5.3.2.1 Sequencing PDE4B in pEE7

Homo sapiens PDE4B1, PDE4B2 or PDE4B3 coding cDNA in pEE7 was sequenced with HCMV and SV40 and the sequence data subjected to a standard NCBI nucleotide based BLAST search, so as to identify any significant regions of known sequence, in this case seeking an alignment with the *Homo sapiens* PDE4B gene set. Because of the size of both the PDE4B cDNA inserts (accounting for between 1.8 and 2.2 kb), and the pEE7 vector (approximately 7 kb), however, the sequence of the entire plasmid could not be determined based upon the use of HCMV and SV40 alone. The pEE7 upstream regulatory domains were therefore sequenced with a custom 21 mer reverse primer “PDE4B_SEQ” (GTTTCAGCATTCTTTTGAAGTT). Upon receipt of this sequencing data, the sequences were again subjected to a standard NCBI nucleotide based BLAST search.

With both the up and downstream pEE7 regulatory domains having been sequenced, the plasmid was screened with software package NEBcutter, version 2.0, so as to identify suitable restriction digest sites applicable for the removal of the intact PDE4B inserts, yet ensuring that the enzymes presented in an order suitable for the forward ligation of the fragments into pcDNA3, enabling a higher copy number amplification of the plasmid.

5.3.2.2 Transferring the PDE4B Inserts from pEE7

PDE4B1 in pEE7, PDE4B3 in pEE7 and pcDNA3 were restriction digested with *EcoRI*, as per the manufacturer’s instructions. Likewise, PDE4B2 in pEE7 and pcDNA3.1(+) were restriction digested with *BamHI* and *HindIII*, as per the manufacturer’s instructions. The restriction digest products were separated via agarose gel based electrophoresis, visualised with SYBR green, and excised based upon expected insert lengths of 2.3 kb for PDE4B1, 1.7 kb for PDE4B2, 2.2 kb for PDE4B3 and 5.4 kb for pcDNA3 and pcDNA3.1(+). The purified DNA fragments

were recovered via a QIAquick gel extraction kit and quantified on a Nanodrop 1000 Spectrophotometer.

Each PDE4B cDNA insert and complementary pcDNA vector were ligated with T4 DNA Ligase, as per the manufacturer's instructions, and transformed into library efficiency DH5 α cells, as per the manufacturer's instructions. The transformed cells were plated and approximately 20 colonies per plate screened for the presence of the modified plasmids. In brief, each viable colony was cultured and the amplified plasmid extracted via Miniprep, prior to restriction digestion with either *EcoRI* (for PDE4B1 and PDE4B3) or *BamHI* and *HindIII* (for PDE4B2). Each digested sample was separated via agarose gel based electrophoresis and visualised through the use of SYBR green, with the aim of identifying an appropriate restriction digest pattern, based on the previously detailed cDNA insert sizes. Where a plasmid appeared compliant it was sent for sequencing, so as to confirm the presence of the correct insert, and indeed that the insert was ligated in the correct orientation.

5.3.2.3 FLAG-tag Insertion Into PDE4B Containing pcDNA3 or pcDNA3.1(+)

PDE4B1 and PDE4B3 cDNA inserts in pcDNA3 were restriction digested with *HindIII* and *BamHI*, as per the manufacturer's instructions. Likewise, PDE4B2 in pcDNA3.1(+) was restriction digested with *NheI* and *HindIII*, as per the manufacturer's instructions. The restriction digested DNA was separated via agarose gel based electrophoresis and visualised with SYBR green. The gel bands of interest were excised and recovered via a QIAquick gel extraction kit, as per the manufacturer's instructions, quantifying each sample with a Nanodrop 1000 Spectrophotometer.

Custom N-terminal FLAG-tag forward and reverse primers PDE1&3FLAGFor (AGCTTGGTACCTATGGATTACAAGGATGACGACGATAAGACG) and PDE1&3FLAGRev (GATCCGTCTTATCGTCGTCATCCTTGTAATCCATAGGTACCA) for long PDE4B isoforms PDE4B1 and PDE4B3, and custom N-terminal FLAG-tag forward and

reverse primers B2-PC3.1Forward (CTAGCGCGGCCGCATGGATTACAAGGATGAC GACGATAAGA) and B2-PC3.1Reverse (AGCTTCTTATCGTCGTCATCCTTGTAATCCA TGCGGCCGCG) for short-PDE4B isoform PDE4B2 were combined at an equal concentration (w/w) and heated to 65 °C for 10 minutes. The water bath was switched off and the temperature slowly equilibrated to that of the room, enabling the single stranded primers to anneal to form a double stranded insert, ready for ligation.

The double stranded FLAG-tag inserts were combined with their complementary plasmid and ligated with T4 DNA Ligase, as per the manufacturer's instructions. The ligated plasmids were transformed into library efficiency DH5α cells, as per the manufacturer's instructions. The transformed cells were plated and approximately 20 colonies per plate screened for the presence of the modified plasmids. In brief, each colony was cultured and the amplified plasmid recovered via Miniprep. The plasmids were restriction digested with either *KpnI* (for PDE4B1 and PDE4B3) or *NotI* (for PDE4B2) and the restriction digest products separated via agarose gel based electrophoresis, visualising through the application of SYBR green. Where a plasmid appeared to contain a FLAG tagged insert, with a length greater than the unmodified PDE4B cDNA containing plasmid, it was sent for sequencing.

5.4 Results and Discussion

5.4.1 AQUA Peptide Selection

5.4.1.1 Sequence Unique to Each PDE4B Isoform

While the four A-Raf isoforms presented with a shared N-terminal, and each was differentiated by the intron at which the isoform terminated (each encountering a premature stop codon), each of the PDE4B isoforms were found to contain a shared C-terminal and were instead differentiated on a short run of unique N-terminal sequence. In some ways, this made the quantification of the PDE4B isoforms more difficult, as the selection of unique C-terminal peptides was impossible. Instead, the quantification of PDE4B focused on the selection of

unique N-terminal proteolytic peptides, generated by a range of proteases. In those cases where a suitable peptide could not be identified, subtraction based quantification was again employed.

Shown on Figure 5-5 is the sequence unique to each of the four PDE4B isoforms, the first of which, long-isoform PDE4B1, was found to contain 93 residues of unique sequence before the inclusion of shared domain UCR1. PDE4B2 meanwhile, being a short-isoform, was found to contain only 39 residues of unique sequence before the inclusion of shared domain UCR2, while long-isoform PDE4B3 was found to contain 78 residues of unique sequence before the inclusion of shared domain UCR1.

```

CLUSTAL 2.0.12 multiple sequence alignment

B3      MTAKDSSKELTASEPEVCIKTFKEQMHLELELPRLPGNRPTSPKISPRSSPRNSPCFFRK
B5      MPEANYLLSVSWGVI-----
B2      MKEHGGTFSSSTGISGGSGDSAMDSLQPLQPNYMFVCLFA-----
B1      MKKSR SVMIVMADDNVKDYFECSLSKSYSSSNLGLDLWRGRCCSGNLQLPPLSQRQS

B3      LLVNKSIRQRRRFTVAHT-----CFDVENG PSPGR SPLD PQASS SAGLVL
B5      -----
B2      -----
B1      ERART PEGDGISRPTTLPLTLPSIAITTVSQE CF DVENG PSPGR SPLD PQASS SAGLVL

      <--UCR1-----
B3      HATFPGHSQRRESFLYRSDSDYDLS PKAMSRNSSLPSEQHGDDLI VTPFAQVLA SLRSVR
B5      -----
B2      -----
B1      HATFPGHSQRRESFLYRSDSDYDLS PKAMSRNSSLPSEQHGDDLI VTPFAQVLA SLRSVR

      ----->          <--UCR2-----
B3      NNFTILTNLHGTSNKRSPAASQPPVSRVNPQEESYQKLAMETLEE LDWCLDQLE TIQTYR
B5      -----
B2      -----EESYQKLAMETLEE LDWCLDQLE TIQTYR
B1      NNFTILTNLHGTSNKRSPAASQPPVSRVNPQEESYQKLAMETLEE LDWCLDQLE TIQTYR

      ----->
B3      SVSEMASNKFKRMLNRELTHLSEMSRSGNQVSEYI SNTFLDKQNDVEIP SPTQKDREKKK
B5      -----KFKRMLNRELTHLSEMSRSGNQVSEYI SNTFLDKQNDVEIP SPTQKDREKKK
B2      SVSEMASNKFKRMLNRELTHLSEMSRSGNQVSEYI SNTFLDKQNDVEIP SPTQKDREKKK
B1      SVSEMASNKFKRMLNRELTHLSEMSRSGNQVSEYI SNTFLDKQNDVEIP SPTQKDREKKK

      <--Catalytic region-----
All     KQQLMTQISGVKMLHSSSLNNTSISRFGVNTENE DHLAKELEDLNKWLNI FNVAG...

```

Figure 5-5: A Clustal X based alignment of PDE4B1, PDE4B2, PDE4B3 and PDE4B5, highlighting the areas of sequence unique to each of the isoforms. Each of the PDE4B isoforms was aligned via Clustal X, with the exception of PDE4B4 which has no confirmed sequence. Highlighted in purple are the areas of sequence unique to PDE4B1, highlighted in green are the areas of sequence unique to PDE4B2, highlighted in yellow are the areas of sequence unique to PDE4B3 and highlighted in blue are the areas of sequence unique to PDE4B5.

As no protein sequence has as yet been described for long-PDE4B isoform PDE4B4, it was excluded from this particular alignment. It was imperative, however, that the sequence of PDE4B4 was considered when peptides were selected for the quantification of multiple PDE4B isoforms.

Finally, super-short isoform PDE4B5 was found to contain only 15 residues of unique sequence, before the inclusion of a truncated version of shared domain UCR2, making it the most difficult to quantify.

5.4.1.2 AQUA Peptides Selected for the Quantification of PDE4B

Through the use of the AQUA peptide selection workflow, as was detailed in Chapter 2, five AQUA peptides have been selected for the quantification of each of the five PDE4B isoforms (Table 5-1).

Table 5-1: The AQUA peptides selected for the quantification of each of the PDE4B isoforms. Each of the AQUA peptides selected for the quantification of each of the four PDE4B isoforms.

Peptide	Protease	Present in:	Comments:
DYFECSLSK	Trypsin	PDE4B1	Located within the unique C-terminal of PDE4B1.
NSPCFFR	Arg-C	PDE4B3	Located within the unique C-terminal of PDE4B3.
VNPQEESYQK	Trypsin	PDE4B1, PDE4B3 and PDE4B4	Bridges the UCR2 N-terminal splice site, and therefore is present only in the long PDE4B isoforms.
TIQTYRSVSE	Glu-C	PDE4B1, PDE4B2, PDE4B3 and PDE4B4	Located within UCR2, and this is present in both long and short PDE4B isoforms.
TDIDIATE	Glu-C	All five PDE4B isoforms	Located in close proximity to the C-terminal of the PDE4B isoforms, and therefore is present in each of the PDE4B isoforms.

Each of the peptides selected for the quantification of the five PDE4B isoforms is also shown on Figure 5-6.

```

CLUSTAL 2.0.12 multiple sequence alignment

B3      MTAKDSSKELTASEPEVCIKTFKEQMHELELPRLPGNRPTSPKI SPRS SPRNSPCFFRK
B5      MPEANYLLSVSWGVI-----
B2      MKEHGGTFSSTGISGGSGDSAMDLSLQPLQPNYMPVCLFA-----
B1      MKKSR SVMTVMADDNVKDYFECSLSKSYSS SSNTLGIDLWRGRCCSGNLQLPPLSQRQS

B3      LLVNKSIQRRRFTVAHT-----CFDVENG PSPGR SPLD PQASS SAGLVL
B5      -----
B2      -----
B1      ERART PEGDGI SRPT TLPLT TLPSI AITTVSQECF DVENG PSPGR SPLD PQASS SAGLVL

      <--UCR1-----
B3      HATFPGHSQRRESFLYRSDS DYDLS PKAMSRNSSL PSEQHGDDLI VTPFAQVLA SLRSVR
B5      -----
B2      -----
B1      HATFPGHSQRRESFLYRSDS DYDLS PKAMSRNSSL PSEQHGDDLI VTPFAQVLA SLRSVR

      ----->
B3      NNFTI LTNLHGTSNKRSPAA SQPPVSR VNPQEESYQK LAMETLEE LDWCLDQLE TIQTYR
B5      -----
B2      -----EESYQK LAMETLEE LDWCLDQLE TIQTYR
B1      NNFTI LTNLHGTSNKRSPAA SQPPVSR VNPQEESYQK LAMETLEE LDWCLDQLE TIQTYR

      ----->
B3      SVSEMASNKF KRMLNRELTH LSEMSRSGNQVSEYI SNTFLDKQNDVEIP SPTQKDREKKK
B5      -----KF KRMLNRELTH LSEMSRSGNQVSEYI SNTFLDKQNDVEIP SPTQKDREKKK
B2      SVSEMASNKF KRMLNRELTH LSEMSRSGNQVSEYI SNTFLDKQNDVEIP SPTQKDREKKK
B1      SVSEMASNKF KRMLNRELTH LSEMSRSGNQVSEYI SNTFLDKQNDVEIP SPTQKDREKKK

A11     KQQLMTQISGVKMLHSSSLNNTSI SRFGVNTENE DHLAKELEDLNKWLNI FNVAGYSH
      NRPLTCIMYAI FQERDLLKT FRISSDTFIT YMTLEDHYH SDVAYHNSL HAADVAQSTHV
      LLSTPALDAVFTDLE ILAAI FAAA I HDVDH PGVSNQFLINTNSELALMYNDESVLENHHL
      AVGFKLLQEEHCDIFMNLTKKQRQT LRRMVIDMVLATDMS KHMSL LADL KTMVE TKKVT S
      SGVLLLDNYT DRIQVLRNMVHCADL SNPTKSLELYRQWTD RIMEE FFQQGDKER ERGME I
      S PMCDKHTASVEKSQVGFIDYIVHPLWETWADLVQ PDAQD ILDTLEDNRN WYQSMI PQS P
      S PPLDEQNRDCQGLMEKFQF ELTLDEEDSE GPEKE GEGHS YFSST KTL CVIDPENRDSL G
      E TDIDIATE DKSPVDT

```

Figure 5-6: The AQUA peptides selected for the quantification of each of the five PDE4B isoforms. A Clustal X based alignment of the four known PDE4B consensus sequences. Highlighted in blue is peptide NSPCFFR, selected for the quantification of PDE4B3 exclusively, highlighted in purple is peptide DYFECSLSK, selected for the quantification of PDE4B1 exclusively, highlighted in red is peptide VNPQEESYQK, selected for the quantification of PDE4B1, PDE4B3 and PDE4B4, highlighted in green is peptide TIQTYRSVSE, selected for the quantification of PDE4B1, PDE4B2, PDE4B3 and PDE4B4, and highlighted in brown is peptide TDIDIATE, selected for the quantification of each of the five PDE4B isoforms.

5.4.2 AQUA Peptide Optimisation

As each of the techniques developed for optimising both the target and synthetic peptides have previously been discussed, then the optimisation of each of the PDE4B peptides will not be discussed here. This includes the selection of suitable

Q1 and Q3 masses, collision energy optimisation, Cys alkylation, and the optimisation of tryptic, Glu-C and Lys-C catalysed digests.

As Proteolytic enzyme Arg-C has not yet been optimised however, a BSA based digest optimisation was conducted for this protease. Furthermore, while Glu-C has been optimised, AQUA peptide TDIDIATE was found to be cleaved at Glu, yet was found to contain two Asp residues, and as such this digest was further assessed for any residual Asp based cleavage upon the depletion of Glu-X peptide bonds.

5.4.2.1 Proteolytic Digest Optimisation

When performing an AQUA based absolute quantification, the digest is of critical importance (Kirkpatrick et al., 2005a). Indeed, the only protease used within this research to have not been optimised is Arg-C, as it is required for the digestion of long-PDE4B isoform PDE4B3, and thus the production of target peptide NSPCFFR.

As was performed with trypsin, Lys-C and Glu-C, 25 µg vials of BSA were digested at a range of enzyme:substrate ratios (w/w), ranging from 1:10 to 1:200, in the presence of 0.1% (w/v) RapiGest SF surfactant, as per the manufacturer's instructions. From these digests (the data for which is not shown) proteolytic peptides were detected from only the 1:20 enzyme:substrate ratio (w/w) digest sample, which even then achieved a sequence coverage for BSA of only 5%. If we consider that trypsin cleaves at both Lys and Arg, each of which occur at a frequency of 5% within an average mix of proteins (Tsuji et al., 2010), and that tryptic peptides are rarely larger than 10 to 15 amino acids in length, with a mass of between 1,000 and 3,000 Da, then it is clear that a high sequence coverage is possible. In comparison, if we consider that Arg-C cleaves exclusively at Arg, which occurs at a frequency of only 5%, then the expected peptide length doubles to approximately 20 amino acids, with a mass of between 2,000 and 6,000 Da, somewhat unsuitable for detection on modern triple quadrupole based mass spectrometers (Kelleher et al., 1999). By way of an example, when

BSA was theoretically digested with trypsin, some 47 peptides were identified, presenting with an average peptide length of 11.6 amino acids. When the same protein was digested with Arg-C however, only 21 peptides were identified, presenting with an average peptide length of 25 amino acids.

In conclusion, perhaps it was poor experimental design which has limited this Arg-C catalysed digest optimisation, where Arg-C may be more suitable for specific cleavages or for use in combination with a second protease, rather than being the sole enzyme responsible for yielding peptides suitable for MS. This result does not however necessarily signify the elimination of either peptide NSPCFFR or the use of Arg-C, as low sequence coverages were also obtained during the digestion of BSA with trypsin, Lys-C and Glu-C, yet the spiked digestion of A-Raf in no way suggested either protease based digestion to be incomplete.

5.4.2.2 Residual Glu-C Catalysed Cleavage at Asp

As has previously been discussed, in the presence of a phosphate based buffer, Glu-C is thought to cleave specifically at Glu and Asp, with Asp based cleavage occurring at a 3000 fold lower rate. When the same digest is performed in AMBIC however, the digest is thought to be restricted exclusively to Glu (Houmard and Drapeau, 1972). As AQUA peptide TDI*DIATE, selected for the quantification of PDE4B isoform PDE4B5, was found to be cleaved to both the C and N-termini at Glu, but to contain two Asp residues, an experiment was devised to detect any residual Asp based cleavage. This focussed on the final few hours of a typical 18 hour proteolytic digest, the point at which the majority of the Glu-X peptide bonds will have been depleted (based on the manufacturer's recommended digest time).

The first step in this experiment was thus to identify Glu-C cleaved BSA peptides which are digested at Glu, but which contain one or more Asp residues, the results of which are shown on Figure 5-7.

AAN17824.1 serum albumin [Bos taurus]

ORIGIN

```

1 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA FSQYLQCCPF
61 DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK VASLRETYGD MADCCEKQEP
121 ERNECFLSHK DDSPDLPLK PDPNTLCDEF KADEKKFWGK YLYEIARRHP YFYAPELLYY
181 ANKYNGVVFQE CCQAE DKGAC LLPKIE TMRE KVLTSARQR LRCASIQKFG ERALKAWSVA
241 RLSQKFPKAE FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
301 CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAE DKDVCK NYQEA KDAFL GSFLYEYSRR
361 HPEYAVSVLL RLAKYEATL EECCAKDDPH ACYSTVFDKL KHLVDEPQNL IKQNCDDQFEK
421 LGEYGFQNAL IVRYTRKVPQ VSTPILVEVS RSLGKVGTRC CTKPESERMP CTEDYLSLIL
481 NRLCVLHEKT PVSEKVTKCC TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP
541 DTEKQIKKQT ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV
601 STQTALA

```

Peptide	Length	MW (Da)	+1 M/Z
TE.LTEFAKTCVADE.S	12	1382.6	1383.6
AE.DKGACLLPKIE.T	11	1242.6	1243.6
AE.DKDVCKNYQE.A	10	1297.5	1298.5
QE.AKDAFLGSFLYE.Y	12	1359.7	1360.7

Figure 5-7: Proteolytic peptides obtained from BSA cleaved with Glu-C at Glu, but which contain Asp. A theoretical Glu-C catalysed digestion was performed on BSA, identifying any peptides which were cleaved at Glu exclusively, but which contained Asp. Highlighted in yellow are those Glu-C cleaved peptides, while the residue marked in red is a Glu-C cleavage site between two such peptides.

From Figure 5-7, we can see that some four Glu cleaved, Asp containing BSA peptides were identified. These digest solutions were spotted to a 384 well MALDI plate and ionised in an Ultraflex II MS. Analysis of the resulting MS data revealed that only one peptide was detected at a high intensity, peptide AKDAFLGSFLYE, while peptides DKGACLLPKIE and LTEFAKTCVADE were barely detectable above the spectrum background noise and no signal was detected for peptide DKDVCKNYQE. Therefore, the intensity of peptide AKDAFLGSFLYE alone was recorded at each of the ten digest time points, as to detect any residual Asp based cleavage.

So as to increase the reliability of the results obtained, PepMix II MALDI calibration standard was added to each of the digest spots. By assessing any change in the ionisation response to these known calibrants, the intensity of peptide AKDAFLGSFLYE could be adjusted so as to account for any variation in analyte-matrix co-crystallisation.

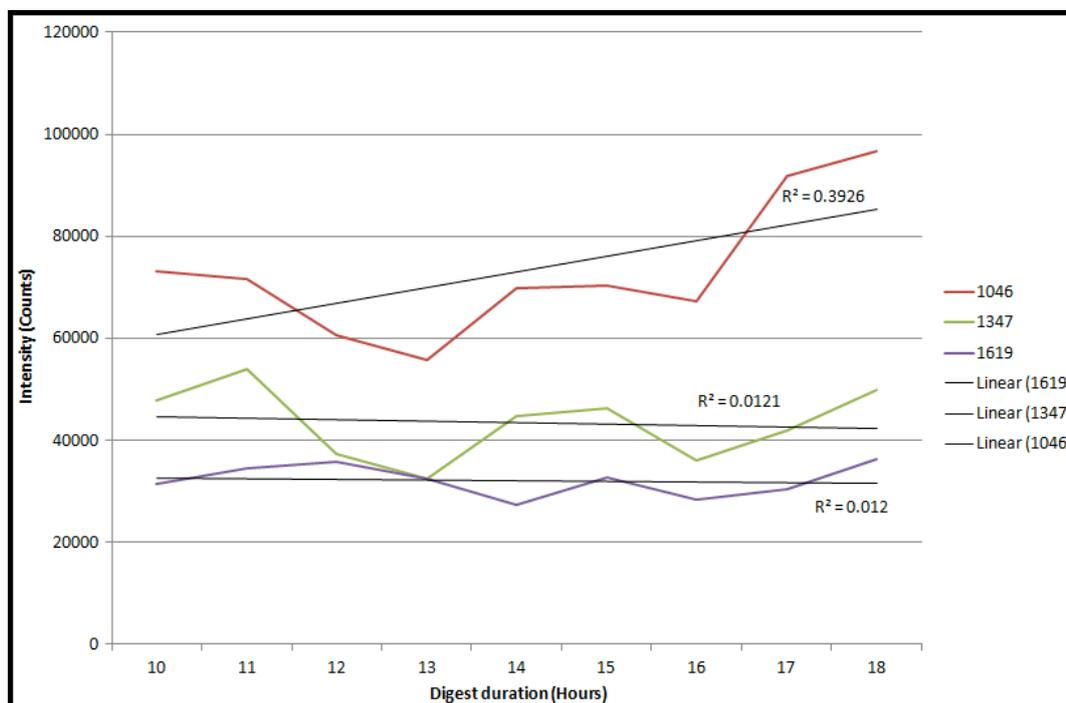


Figure 5-8: The peak intensities obtained for three PepMix II calibration standards. The peak intensities detected for three molecules from the PepMix II calibration standard, ionised on an Ultraflex II MS. 1 μ L of the PepMix II calibration standard was added to each of the BSA digest MALDI spots, enabling any variations in signal intensity detected to be used to normalise the ion intensity detected for peptide AKDAFLGSFLYE.

Figure 5-8 shows the ion intensity/MS response detected for the three small molecules from the PepMix II calibration standard, and while the ion intensity for peaks 1347 m/z and 1619 m/z appears stable over the time course, each achieving an R^2 value of 0.01, peak 1046 m/z has presented with a slight increase in intensity over the course of the digest, achieving an R^2 value of 0.39. While an actual change in quantity is not possible, this increase in intensity may instead be due to Angiotension II (peak 1046 m/z) being the smallest of the molecules detected, and as such may have been more prone to ion suppression and analyte-matrix co-crystallisation conditions.

The data collected from these PepMix II molecules was averaged for each time point and this data set compared to an overall average value for the PepMix II batch. Any variations detected in the MS response to the PepMix II molecules was

used to adjust the ion intensity detected for peptide AKDAFLGSFLYE at each time point, as is shown on Figure 5-9.

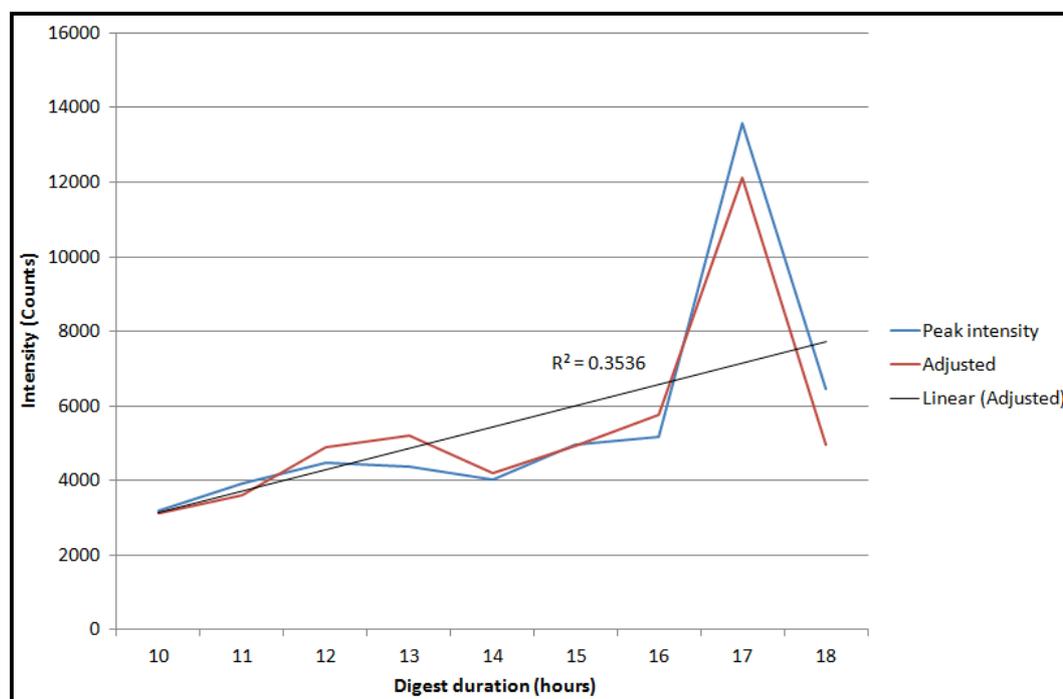


Figure 5-9: The ion intensity detected for Glu-C cleaved peptide AKDAFLGSFLYE. The ion intensity detected for Glu-C cleaved BSA peptide AKDAFLGSFLYE, as digested over a ten hour time period (between 10 and 20 hours). This peptide was selected due to its cleavage at Glu exclusively, while containing an Asp residue, enabling the detection of any residual Asp based cleavage. The ion intensity detected for peptide AKDAFLGSFLYE during this analysis was normalised based on the signal response detected for three PepMix II calibrants, so as to account for any variation in ionisation efficiency/analyte-matrix co-crystallisation conditions.

From Figure 5-9 it is clear that there is a slight increase in the amount of peptide AKDAFLGSFLYE detected over the course of the experiment, achieving an R^2 value of 0.35. This would suggest that the Glu-C catalysed digestion of BSA has continued slowly, right up until the experiment was terminated. If any residual Asp based cleavage was present, then the quantity of peptide AKDAFLGSFLYE would have been expected to fall, however this appeared to not be the case. Indeed even if a small number of Asp bonds were cleaved upon the depletion of Glu catalysed digest sites during an in-gel digestion, this residual activity should affect both the target and synthetic peptides equally.

In conclusion; as no residual Asp based cleavage was detected during the Glu-C catalysed digestion of BSA, therefore the selection and use of peptides cleaved at Glu, but which contain Asp, would appear to be a suitable approach.

5.4.3 MRM Based Acquisition

5.4.3.1 MRM Method Development

Having previously detailed the technique through which suitable Q1 and Q3 masses were selected, the optimisation of the collision energy for each transition and the selection of a suitable dwell time for each of the MRM transitions, none of these techniques will be discussed here. Instead, shown on Table 5-2 is the complete MRM acquisition method designed for the detection of the five AQUA peptides, and thus capable of quantifying each of the four known PDE4B isoforms.

Table 5-2: The MRM transitions selected for the detection of each of the PDE4B peptides. The acquisition method developed for the detection of each of the five PDE4B isoforms; detailing the Q1 and Q3 MRM transitions, the Q2 optimised collision energies and the dwell times selected for each transition.

Peptide Sequence (Q1)	Fragment Ion (Q3)	Q1 <i>m/z</i>	Q3 <i>m/z</i>	Dwell Time (ms)	Collision Energy (eV)
NSPCFFR	B ₂	464.2	202.1	60	20
NSPCFFR	Y ₄	464.2	629.3	60	25
NSPCFFR	Y ₅	464.2	726.3	60	20
NSP*CFRR	B ₂	467.2	202.1	60	20
NSP*CFRR	Y ₄	467.2	629.3	60	25
NSP*CFRR	Y ₅	467.2	732.3	60	20
TIQTYRSVSE	Y ₂	592.3	235.1	60	25
TIQTYRSVSE	Y ₇	592.3	841.4	60	25
TIQTYRSVSE	Y ₈	592.3	969.5	60	20
TIQTYR*SVSE	Y ₂	597.3	235.1	60	25
TIQTYR*SVSE	Y ₇	597.3	851.4	60	25
TIQTYR*SVSE	Y ₈	597.3	979.5	60	20
VNPQEESYQK	Y ₄	611.3	525.3	60	35
VNPQEESYQK	Y ₆	611.3	783.4	60	30
VNPQEESYQK	Y ₈	611.3	1008.5	60	30
VNP*QEESYQK	Y ₄	614.3	525.3	60	35
VNP*QEESYQK	Y ₆	614.3	783.4	60	30
VNP*QEESYQK	Y ₈	614.3	1014.5	60	30
TDIDIATE	B ₅	877.4	558.3	60	30
TDIDIATE	B ₆	877.4	629.3	60	30
TDIDIATE	B ₇	877.4	730.4	60	30
TDI*DIATE	B ₅	884.4	565.3	60	30
TDI*DIATE	B ₆	884.4	636.3	60	30
TDI*DIATE	B ₇	884.4	737.4	60	30
DYFECSLSK	B ₂	574.8	279.1	60	20
DYFECSLSK	Y ₆	574.8	723.3	60	20
DYFECSLSK	Y ₇	574.8	870.4	60	20
DYF*ECSLSK	B ₂	579.8	279.1	60	20
DYF*ECSLSK	Y ₆	579.8	723.3	60	20
DYF*ECSLSK	Y ₇	579.8	880.4	60	20

5.4.3.2 MRM Linear Response and LoD on an API 2000

As was performed for each of the A-Raf peptides, each PDE4B AQUA peptide was injected onto an Ultimate 3000 HPLC and detected on an API 2000 MS using the previously devised PDE4B MRM acquisition method. Each peptide was injected at eight different quantities, ranging from 5 pmol to 100 amol, and analysed through Analyst software package, version 4.2. Each peak automatically selected by the quantitation software was reviewed manually, so as to avoid erroneous peak integration. The results from this experiment are shown on Table 5-3.

Table 5-3: The LoD's achieved for each of the PDE4B MRM transitions, as detected on an API 2000. The data obtained when eight different quantities of each AQUA peptide, ranging from 5 pmol to 100 amol, were injected. Shown are the peak areas detected, in counts per second (cps), summed over the width of the peak. As no peaks were detected below 50 fmol, the results from these experiments were omitted.

Peptide	Q3 Ion	5 pmol	1 pmol	200 fmol	50 fmol
NSP*CFFR	B ₂	2.72E+005	1.82E+004	8.41E+002	0.00E+000
NSP*CFFR	Y ₄	4.81E+004	3.27E+003	2.60E+002	0.00E+000
NSP*CFFR	Y ₅	1.10E+005	7.75E+003	4.20E+002	0.00E+000
TIQTYR*SVSE	Y ₂	1.35E+005	3.09E+004	9.68E+003	4.05E+002
TIQTYR*SVSE	Y ₇	2.79E+004	6.33E+003	1.80E+003	2.58E+002
TIQTYR*SVSE	Y ₈	4.45E+004	9.63E+003	3.24E+003	4.30E+002
VNP*QEESYQK	B ₅	3.77E+004	8.11E+003	1.00E+003	0.00E+000
VNP*QEESYQK	B ₆	3.01E+004	5.53E+003	1.40E+003	0.00E+000
VNP*QEESYQK	B ₇	1.15E+005	2.18E+004	4.27E+003	0.00E+000
TDI*DIATE	B ₂	7.18E+004	1.90E+004	3.78E+003	1.07E+003
TDI*DIATE	Y ₆	5.89E+004	1.62E+004	3.49E+003	7.71E+002
TDI*DIATE	Y ₇	2.32E+004	7.22E+003	1.49E+003	3.68E+002
DYF*ECSLSK	B ₂	1.16E+005	7.01E+003	1.45E+003	0.00E+000
DYF*ECSLSK	Y ₄	3.30E+004	1.15E+003	4.23E+002	0.00E+000
DYF*ECSLSK	Y ₅	4.41E+004	1.43E+003	5.85E+002	0.00E+000

From Table 5-3 it is clear that there are varying LoDs for each of the PDE4B isoforms, for example, AQUA peptides TDI*DIATE and TIQTYR*SVSE could be detected on the API 2000 when 50 fmol of either peptide was injected. It should, therefore, be possible to quantify both short and super-short PDE4B isoforms, PDE4B2 and PDE4B5, in a sample containing 50 fmol or more of each protein and to achieve a linear MS response (Figure 5-10). As AQUA peptides NSP*CFFR, DYF*ECSLSK and VNP*QEESYQK were only detected when 200 fmol of peptide was injected, however, then the LoD for long-PDE4B isoforms PDE4B1, PDE4B3 and PDE4B4 on an API 2000 was set to 200 fmol.

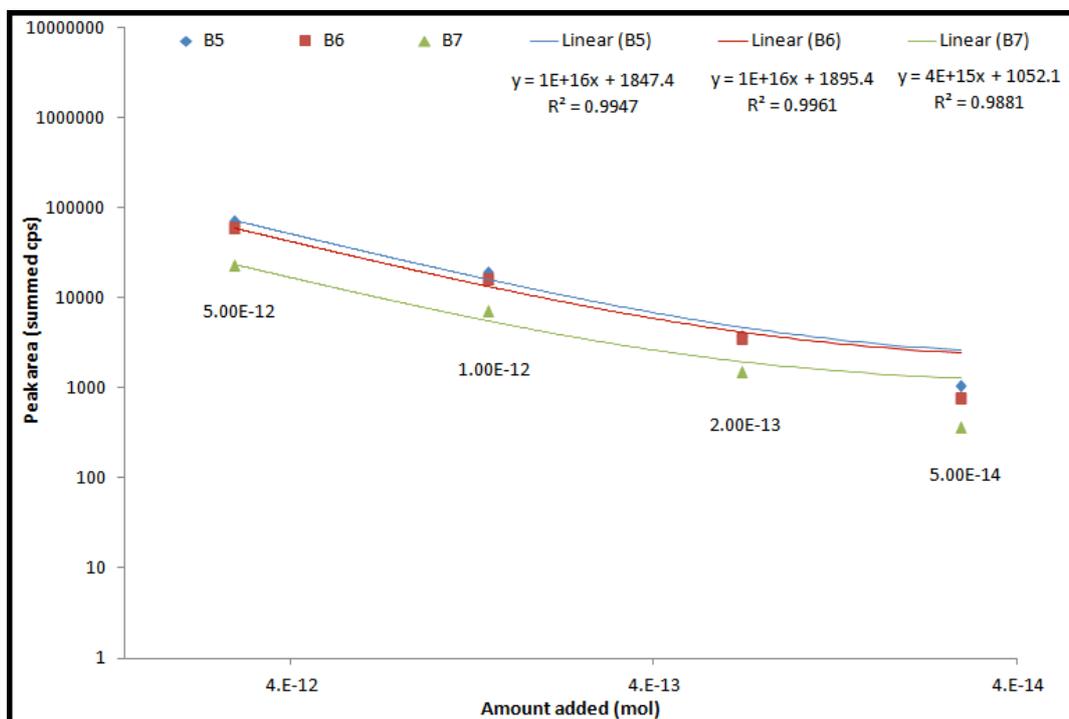


Figure 5-10: The signal intensity and LoD achieved for PDE4B AQUA peptide TDI*DIATE. The peak areas obtained during the analysis of three TDI*DIATE transitions, when injected at eight quantities, ranging from 5 pmol to 100 amol, on an API 2000 MS. This data is presented in a logarithm format, plotting the amount of peptide added against the summed signal intensity. Linear trend lines are included for each transition, showing the R^2 values for each equation.

5.4.3.3 LoD on a QTrap 5500

As was the case when quantifying each of the A-Raf protein isoforms, when instrument time became available on a more powerful triple quadrupole based MS, the LoD for the lowest performing PDE4B AQUA peptide was re-analysed. The reason for this switch in instrument was to increase the sensitivity of the MRM based acquisition method, and in doing so, lowering the LoD for each of the PDE4B peptides, thereby reducing the quantity of sample required for an analysis to be performed.

While previously the LoD for each of the long-PDE4B isoforms was found to be 200 fmol, due to the low ionisation efficiencies detected for peptides NSP*CFFR, DYF*ECLSK and VNP*QEESYQK, the lowest ion intensity was detected for PDE4B3 peptide NSP*CFFR (Table 5-3). AQUA peptide NSP*CFFR was thus injected onto

an Ultimate 3000 HPLC and analysed on a QTrap 5500 at eight different quantities, the results of which are shown on Table 5-4.

Table 5-4: The LoD's achieved for each NSP*CFFR transition, as detected on a QTrap 5500 MS. Eight quantities of alkylated AQUA peptide NSP*CFFR were injected on an Ultimate 3000 HPLC and detected on a QTrap 5500, using the previously devised PDE4B MRM based acquisition method. Both the peak height and peak area were measured and summed, with cps denoting counts per second.

Fragment	Quantity	Analyte Peak Area (cps)	Analyte Peak Height (cps)	Analyte SNR
B ₂	5 pmol	7.84E+07	4.81E+06	2.11E+05
Y ₄	5 pmol	5.66E+07	3.54E+06	3.07E+05
Y ₅	5 pmol	1.10E+08	6.45E+06	5.54E+05
B ₂	1 pmol	2.01E+06	1.35E+05	5.08E+03
Y ₄	1 pmol	1.45E+06	9.94E+04	7.67E+03
Y ₅	1 pmol	3.00E+06	1.95E+05	2.19E+04
B ₂	200 fmol	3.83E+05	2.70E+04	9.48E+02
Y ₄	200 fmol	2.74E+05	1.86E+04	1.35E+03
Y ₅	200 fmol	5.82E+05	4.00E+04	4.08E+03
B ₂	50 fmol	6.04E+04	4.01E+03	1.47E+02
Y ₄	50 fmol	4.25E+04	2.81E+03	1.92E+02
Y ₅	50 fmol	9.10E+04	6.93E+03	8.22E+02
B ₂	10 fmol	2.71E+04	1.77E+03	6.01E+01
Y ₄	10 fmol	2.24E+04	1.36E+03	8.13E+01
Y ₅	10 fmol	4.19E+04	2.81E+03	2.63E+02
B ₂	2 fmol	2.01E+03	2.44E+02	8.76E+00
Y ₄	2 fmol	1.20E+03	2.33E+02	1.82E+01
Y ₅	2 fmol	2.86E+03	2.86E+02	3.37E+01
B ₂	500 amol	0.00E+00	0.00E+00	N/A
Y ₄	500 amol	0.00E+00	0.00E+00	N/A
Y ₅	500 amol	0.00E+00	0.00E+00	N/A
B ₂	100 amol	0.00E+00	0.00E+00	N/A
Y ₄	100 amol	0.00E+00	0.00E+00	N/A
Y ₅	100 amol	0.00E+00	0.00E+00	N/A

From Table 5-4 it is clear that the LoD for PDE4B3 peptide NSP*CFFR on a QTrap 5500 is 2 fmol, the lowest quantity at which all three peptide transitions could be detected above the background noise. As the LoD for peptide NSP*CFFR on an API 2000 was 200 fmol, this change in instrument has presented a 100 fold increase in sensitivity. While the manufacturer has documented the LoD for Reserpine to have increased 1000 fold from the API 2000 to the QTrap 5500, this increase in sensitivity failed to be achieved during the analysis of A-Raf peptide GL*NQDCCVVYR, which showed only a 5 fold increase. This low sensitivity may have in part been due to sub-optimal instrument parameters, where altering the

declustering potential, collision cell exit potential and gas flow rates may have lowered the LoD.

5.4.4 PDE4B Protein Production

With a LoD similar to that of A-Raf, and an effective technique for protein production already developed, it was the author's intention for the PDE4B plasmids to be transfected into HEK293 cells using Genejuice transfection reagent, expressed exogenously, and purified via IP. Unlike the A-Raf plasmids however, the PDE4B genes were supplied in a pEE7 vector designed by Celltech in 1989 (Stephens and Cockett, 1989). While the use of an alternative plasmid was in itself not a problem, no sequence information could be obtained as to the up and downstream regulatory domains of the pEE7 plasmid, preventing the PDE4B isoforms from being FLAG tagged. Furthermore, when the PDE4B plasmids were amplified and purified via Midiprep, the plasmid yields were on average 10 fold lower than those obtained for A-Raf (data not shown). By way of an example of the scale of the problem, a typical A-Raf transfection batch would utilise six 150 x 25 mm tissue culture plates and require a total of 120 ng of plasmid. As each 50 mL midiprep reaction yielded only 20 ng of plasmid on average, protein production on this scale became unviable. Instead it was decided that the PDE4B inserts should be transferred to pcDNA3, and each isoform supplemented with an N-terminal FLAG tag so as to enable purification. In order to achieve this aim, the pEE7 based PDE4B plasmids were first partially sequenced to identify suitable up and downstream restriction digest sites, the inserts ligated into pcDNA3, and suitable N-terminal FLAG tags designed for each species. In brief, a FLAG tag consists of an artificial 8-amino acid peptide sequence, which in theory does not occur in nature and thus an anti-flag antibody based IP should be specific to the FLAG tagged target protein. Furthermore, as the FLAG tag is designed to be hydrophilic, it should be located on the surface of the protein, ensuring the sequence is available for anti-FLAG binding.

5.4.4.1 Sequencing PDE4B in a pEE7 Vector

As the only source of literature to be identified detailing the construction of the pEE7 plasmid noted the presence of a HCMV promoter and an SV40 poly(A) (Stephens and Cockett, 1989), each of the PDE4B plasmids were sequenced with both of the above primers. While the HCMV promoter appeared to be present within the pEE7 plasmid (the sequence data obtained is not shown), the data in no way aligned with any published *Homo sapiens* PDE4B sequence, though many non-PDE4B alignments were achieved, suggesting the sequence may instead have originated from the pEE7 plasmid. When the SV40 sequence data was aligned via a nucleotide BLAST search, however, a stretch of shared PDE4B sequence was revealed, from the catalytic domain of the PDE4B isoforms, to just downstream of each insert, showing both *EcoRI* and *BamHI* restriction digest sites.

Based on the sequence data obtained from the SV40 primer, a custom 21 mer reverse primer was created “GTTTCAGCATTCTTTTGAAGCTT” (PDE4B_SEQ), capable of sequencing the UCRs and upstream regulatory domains of the pEE7 plasmid. In addition to confirming the sequence of each PDE4B isozyme, the sequence obtained through the use of PDE4B_SEQ primer revealed the PDE4B1 and PDE4B3 inserts to be preceded by both *EcoRI* and *HindIII* restriction digest sites, while PDE4B2 was found to be preceded by *HindIII*.

5.4.4.2 Transfer of the PDE4B Inserts to pcDNA3 and pcDNA3.1(+)

In terms of the selection of suitable restriction enzymes for the transfer of the PDE4B inserts from pEE7 to pcDNA3, both long-PDE4B isozymes PDE4B1 and PDE4B3 were found to be preceded by both *EcoRI* and *HindIII* and proceeded by both *EcoRI* and *BamHI*. During the sequencing of the PDE4B1 and PDE4B3 UCRs with PDE4B_SEQ however, both isozymes were found to contain an additional *BamHI* restriction digest site, discounting the use of this restriction enzyme.

With *EcoRI* thus selected as the 3' restriction enzyme, both *EcoRI* and *HindIII* were identified as suitable 5' restriction enzymes for the forward insertion of PDE4B1 and PDE4B3 into pcDNA3, though neither was completely ideal. In the

case of utilising *EcoRI* as both the 3 and 5' restriction digest site, this risked the ligation of multiple inserts per plasmid or the insert being incorrectly orientated. While the use of *HindIII* as the 5' digest site appeared more trivial, it was discovered that *HindIII* was the first restriction digest site after the pcDNA3 T7 promoter, making the insertion of a FLAG-tag more complicated. As such, *EcoRI* was selected as the single restriction enzyme for the cleavage of both PDE4B1 and PDE4B3 from pEE7.

In contrast, short-PDE4B isoform, PDE4B2, was found to be preceded by *HindIII* exclusively, making it the only choice, while being preceded by both *EcoRI* and *BamHI*, each of which were suitable for the forward ligation of PDE4B2 into pcDNA3. Of these two enzymes, *BamHI* was finally selected as it yielded the shorter insert of the two.

One problem which was encountered with the transfer of PDE4B2 to pcDNA3 resulted from the use of non-optional 5' restriction digest site *HindIII*, which as previously discussed, was found to immediately precede the T7 promoter in pcDNA3. As to prevent further problems arising during the insertion of a FLAG-tag, PDE4B2 was instead ligated into pcDNA3.1(+), which featured several additional restriction digest sites upstream of *HindIII*.

5.4.4.3 Incorporation of FLAG-tags into the PDE4B Plasmids

Several components are required for the incorporation of a successful N-terminal FLAG tag, these include: (i) The identification of several restriction digest sites between the promoter of the vector and the start codon of the insert, and (ii) The incorporation of a double stranded insert which: (a) Contains restriction digest sites complementary to those used to cleave from the plasmid, featuring an overhang ready for ligation. (b) Contains an appropriate sequence for the expression of a FLAG-tag. (c) Contains an additional restriction digest site upstream of the FLAG-tag, ideally not present within the vector, for the purpose of screening colonies post cloning. (d) Contains a start codon, and (e) Maintains the correct reading frame through the incorporation of an appropriate number of

additional nucleotides, while ensuring the additional sequence does not inadvertently contain a stop codon, which would terminate translation.

In the case of long-PDE4B isoforms PDE4B1 and PDE4B3, both of which were ligated into pcDNA3 with *EcoRI*, a shared double stranded insert was designed to replace a stretch of sequence cleaved from the pcDNA3 vector, stretching from the *HindIII* restriction digest site (immediately downstream of the T7 promoter) to the *BamHI* restriction digest site (immediately upstream of the *EcoRI* restriction digest site). The final sequences for these single stranded primers are “AGCTTGGTACCTATGGATTACAAGGATGACGACGATAAGACG” for PDE1&3FLAGFor, and “GATCCGTCTTATCGTCGTCATCCTTGTAATCCATAGGTACCA” for PDE1&3FLAGRev, the components of which are more fully explained on Figures 5-11 and 5-12, for PDE4B isoforms PDE4B1 and PDE4B3, respectively.

```

PDE4B1, Hind3 / BamH1 insert:

Sequence from Hind3 -> Start codon:
5' A AGC TTG GTA CCG AGC TCG GAT CCA CTA GTA ACG GCC GCC AGT GTG CTG GAA TTC ATA
ATG 3'
3' T TCG AAC CAT GGC TCG AGC CTA GGT GAT CAT TGC CGG CGG TCA CAC GAC CTT AAG TAT
TAC 5'

Primer requirements:
5' A) AGC TTG GTA CCT ATG GAT TAC AAG GAT GAC GAC GAT AAG ACG (GAT CCA CTA GTA ACG
GCC GCC AGT GTG CTG GAA TTC ATA ATG 3'
3' T TCG A)AC CAT GGA TAC CTA ATG TTC CTA CTG CTG CTA TTC TGC CTA G(GT GAT CAT TGC
CGG CGG TCA CAC GAC CTT AAG TAT TAC 5'

Final primers:
Forward 5' AGC TTG GTA CCT ATG GAT TAC AAG GAT GAC GAC GAT AAG ACG 3'
Reverse 3' AC CAT GGA TAC CTA ATG TTC CTA CTG CTG CTA TTC TGC CTA G 5'
Reverse 5' G ATC CGT CTT ATC GTC GTC ATC CTT GTA ATC CAT AGG TAC CA 3'

Additional NH2 protein sequence:
NH2 M D Y K D D D D K T D P L V T A A S V L E F I M COOH

```

Figure 5-11: The design of the N-terminal FLAG-tag to be inserted into PDE4B1 in pcDNA3. The components required for a successful FLAG-tag capable of spanning the gap between *HindIII* (marked in red) and *BamHI* (marked in purple) in pcDNA3, enabling the rapid purification of the PDE4B protein. Included within the FLAG-tag insert are a *KpnI* restriction digest site (marked in orange), an ATG (Met) start codon (marked in green), the sequence for the FLAG-tag (marked in blue) and some additional sequence intended to maintain the correct reading frame (retained in black).

```

PDE4B3, Hind3 / BamH1 insert:

Sequence from Hind3 -> Start codon:
5' A AGC TTG GTA CCG AGC TCG GAT CCA CTA GTA ACG GCC GCC AGT GTG CTG GAA TTC ATG 3'
3' T TCG AAC CAT GGC TCG AGC CTA GGT GAT CAT TGC CGG CGG TCA CAC GAC CTT AAG TAC 5'

Primer requirements:
5' (A) AGC TTG GTA CCT ATG GAT TAC AAG GAT GAC GAC GAT AAG ACG (GAT CCA CTA GTA ACG
GCC GCC AGT GTG CTG GAA TTC ATG) 3'
3' (T TCG A)AC CAT GGA TAC CTA ATG TTC CTA CTG CTG CTA TTC TGC CTA G(GT GAT CAT TGC
CGG CGG TCA CAC GAC CTT AAG TAC) 5'

Final primers:
Forward 5' AGC TTG GTA CCT ATG GAT TAC AAG GAT GAC GAC GAT AAG ACG 3'
Reverse 3' AC CAT GGA TAC CTA ATG TTC CTA CTG CTG CTA TTC TGC CTA G 5'
Reverse 5' G ATC CGT CTT ATC GTC ATC CTT GTA ATC CAT AGG TAC CA 3'

Additional NH2 protein sequence:
NH2 M D Y K D D D K T D P L V T A A S V L E F M COOH

```

Figure 5-12: The design of the N-terminal FLAG-tag to be inserted into PDE4B3 in pcDNA3. The components required for a successful FLAG-tag capable of spanning the gap between *HindIII* (marked in red) and *BamHI* (marked in purple) in pcDNA3, enabling the rapid purification of the PDE4B protein. Included within the FLAG-tag insert are a *KpnI* restriction digest site (marked in orange), an ATG (Met) start codon (marked in green), the sequence for the FLAG-tag (marked in blue) and some additional sequence intended to maintain the correct reading frame (retained in black).

In the case of short-PDE4B isozyme PDE4B2, ligated into pcDNA3.1(+) with both *HindIII* and *BamHI*, a double stranded insert was designed to replace a stretch of sequence spanning from the *NheI* restriction digest site (immediately downstream of the T7 promoter), to the *HindIII* restriction digest site (the 3' site at which PDE4B2 was ligated). The final sequence for these single stranded inserts was “CTAGCGCGGCCGCATGGATTACAAGGATGACGACGATAAGA” for B2-PC3.1Forward and “AGCTTCTTATCGTCGTCATCCTTGTAATCCATGCGGCCGCG” for B2-PC3.1Reverse, the structures of which are more fully explained on Figure 5-13.

```

PDE4B2, Nhe1 / Hind3 insert:

Sequence from Nhe1 -> Start codon:
5' G' CTA GCG TTT AAA CTT A'AG CTT AGC GTG CAA ATA ATG 3'
3' C GAT C'GC AAA TTT GAA TTC GA'A TCG CAC GTT TAT TAC 5'

Primer requirements:
5' NG)C TAG CGC GGC CGC ATG GAT TAC AAG GAT GAC GAC GAT AAG A(AG CTT AGC GTG CAA
3' NCG ATC) GCG CCG GCG TAC CTA ATG TTC CTA CTG CTG CTA TTC TTC GA(A TCG CAC GTT

ATA ATG 3'
TAT TAC 5'

Final primers:
Forward 5' C TAG CGC GGC CGC ATG GAT TAC AAG GAT GAC GAC GAT AAG A 3'
Reverse 3' GCG CCG GCG TAC CTA ATG TTC CTA CTG CTG CTA TTC TTC GA 5'
Reverse 5' AG CTT CTT ATC GTC GTC ATC CTT GTA ATC CAT GCG GCC GCG 3'

Additional NH2 protein sequence:
NH2 M D Y K D D D D K K L S V Q I M COOH

```

Figure 5-13: The design of the N-terminal FLAG-tag to be inserted into PDE4B2 in pcDNA3.1(+). The components required for a successful FLAG-tag capable of spanning the gap between *NheI* (marked in orange) and *HindIII* (marked in red) in pcDNA3.1(+), enabling the rapid purification of the PDE4B protein. Included within the FLAG-tag insert are a *NotI* restriction digest site (marked in scarlet), an ATG (Met) start codon (marked in green), the sequence for the FLAG-tag (marked in blue) and some additional sequence intended to maintain the correct reading frame (retained in black).

5.4.5 Spiked Digest Based Quantitative Analyses

So as to assess the efficiency with which each of the target peptides were digested (through comparing peak area ratios for each target peptide/protease, each of which should be present at an equal ratio), to identify any problems resulting from the re-suspension and modification of each AQUA peptide (which would result in a specific peptide presenting with a different peak area ratio), and to test the AQUA based absolute quantification workflow; 25 μ L of each PDE4B transfected HEK93 IP was loaded onto an SDS-PAGE gel and electrophoresed. The PDE4B isoform containing gel bands were excised and in-gel digested in the presence of each AQUA peptide. Following peptide extraction the samples were loaded onto an Ultimate 3000 LC and analysed on a QTrap 4000 MS, running the PDE4B MRM acquisition method previously detailed.

If we consider that each transfected plate contains some 1.1×10^8 HEK293 cells, and that six plates were transfected per isoform, then a total of 6.6×10^8 cells

were lysed, and the transfected protein immunoprecipitated. If we then consider that the anti-FLAG beads were re-suspended in 360 μL of protein loading buffer, then the 25 μL of IP loaded to each SDS-PAGE gel well should contain the transfected, immunoprecipitated contents of approximately 4.6×10^7 HEP293 cells.

While only three of the five PDE4B plasmids were transfected (PDE4B1, PDE4B2 and PDE4B3), together these three PDE4B isoforms contained each of the five PDE4B AQUA peptides, thus enabling each to be analysed.

5.4.5.1 PDE4B1 Single Isoform Spiked Digest

As was performed during the spiked digestion of A-Raf, 25 μL of PDE4B1 transfected HEK293 IP was digested with either trypsin, or Glu-C, in the presence of each of the PDE4B AQUA peptides. Three transitions per peptide were analysed and the data averaged. A total of five digest replicates were performed and each replicate injected twice, the data from which is shown on Table 5-5.

Table 5-5: The quantity and variation data obtained from the spiked digestion of 25 μ L of immunoprecipitated PDE4B1 on a QTrap 4000. Through comparing the peak areas for both the synthetic and endogenous PDE4B peptides, three sets of quantities were obtained per peptide, per replicate. These values were averaged, so as to obtain a single quantity per peptide, per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

	TIQTYRSVSE - Present in all but PDE4B5		TDIDIATE - Present in all isoforms	
Replicate	Average per Replicate (pmol/ μ L)	CoV	Average per Replicate (pmol/ μ L)	CoV
1-1	2.28	1.09%	1.62	6.62%
1-2	2.25	6.03%	1.76	15.07%
2-1	2.03	1.84%	2.83	27.81%
2-2	1.99	2.71%	2.67	35.93%
3-1	2.18	2.69%	2.29	23.37%
3-2	2.21	2.52%	2.26	14.44%
4-1	2.10	1.19%	2.22	27.05%
4-2	2.10	0.02%	2.17	25.37%
5-1	2.51	2.69%	2.56	1.27%
5-2	2.51	2.80%	2.53	2.19%
Average	2.21	8.36%	2.29	24.74%
	VNPQEESYQK - Present in PDE4B1, 3 and 4		DYFECLSK - Unique to PDE4B1	
Replicate	Average per Replicate (pmol/ μ L)	CoV	Average per Replicate (pmol/ μ L)	CoV
1-1	2.99	4.91%	3.63	3.25%
1-2	3.03	4.76%	3.61	3.05%
2-1	2.98	6.11%	3.51	2.93%
2-2	2.98	5.63%	3.56	3.24%
3-1	2.72	4.68%	3.20	2.48%
3-2	2.74	5.99%	3.23	3.44%
4-1	3.04	5.42%	3.41	2.56%
4-2	3.03	5.36%	3.41	3.30%
5-1	2.66	5.55%	3.35	2.41%
5-2	2.71	6.17%	3.40	3.74%
Average	2.89	6.94%	3.43	4.91%

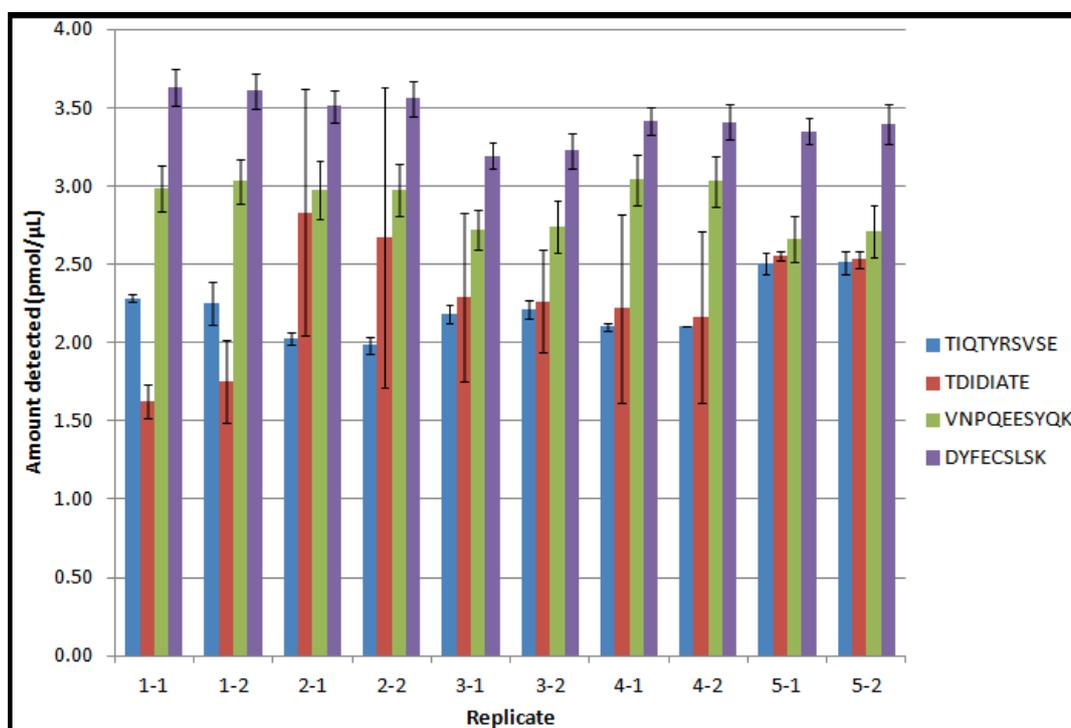


Figure 5-14: The quantities of each PDE4B1 target peptide detected during the spiked digestion of PDE4B1 transfected HEK293 IP on a QTrap 4000. The amount of each PDE4B1 target peptide detected during the spiked digestion of 25 μ L of PDE4B1 transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between both peptides and replicates.

The data presented on Table 5-5 and Figure 5-14 is interesting as each peptide initially appears to be detected at a different level. Throughout the batch, however, the peptides cleaved with Glu-C (TIQTYRSVSE and TDIDIATE) appear to be detected at levels closer to those detected for tryptic peptide VNPQEESYQK. As each digest replicate was performed on the same PDE4B1 transfected IP and each digest conducted under the same conditions, this variation can only be attribute to either the HPLC column or the ESI needle, either of which may have required further equilibration so as to enable optimal peptide binding and elution, or for a more stable ESI spray. Figure 5-15 provides evidence of this increase in ion intensity over the course of the batch.

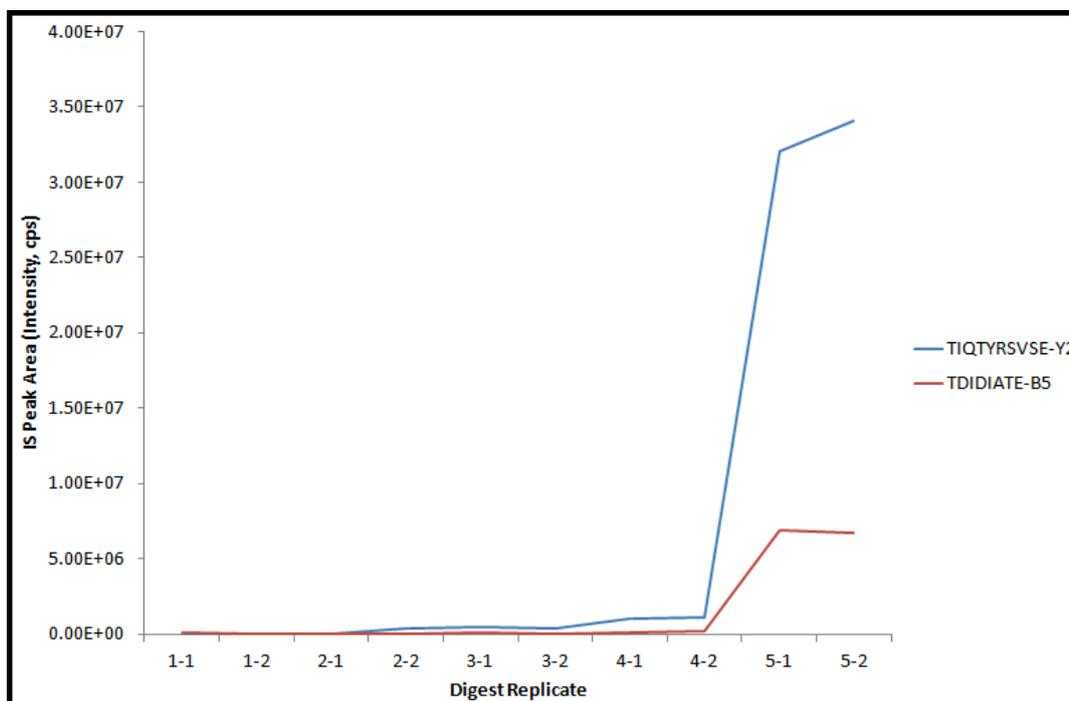


Figure 5-15: The increasing ion intensity detected over the course of an MRM based LC-MS/MS analysis of Glu-C cleaved PDE4B1. The increase in ion intensity detected for AQUA peptide TIQTYRSVSE, fragment Y₂, and AQUA peptide TDIDIATE, fragment B₅, over the course of a batch of five Glu-C digested PDE4B1 transfected HEK293 IP replicates.

As only a single spiked digest replicate was considered accurate, then an ANOVA based statistical analysis was not performed. From reviewing the data presented on Table 5-5 however, the quantity of peptide DYFECSLSK detected during the spiked digestion of 25 μ L of PDE4B1 transfected HEK293 IP, was approximately 40% higher than that obtained for each of the other peptides. As was the case with AQUA peptide GL*NQDCCVVYR, this variation may be due to the additional processing (in-solution alkylation and vacuum centrifugation) which this Cys containing AQUA peptide was subjected to, prior to in-gel digestion.

5.4.5.2 PDE4B2 Single Isoform Spiked Digest

Twenty-five micro litres of PDE4B2 transfected HEK293 cell IP was digested with Glu-C in the presence of each of the PDE4B AQUA peptides, three transitions per peptide were analysed, and the data averaged. A total of five digest replicates

were performed, injecting each replicate twice, the data from which is shown on Table 5-6.

Table 5-6: The quantity and variation data obtained during the spiked digestion of 25 μL of immunoprecipitated PDE4B2 on a QTrap 4000. Through comparing the peak areas for both the synthetic and endogenous PDE4B peptides, three sets of quantitation values were obtained, per peptide, per replicate. These values were averaged, so as to obtain a single quantity, per peptide, per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

Replicate	TIQTYRSVSE - Present in all but PDE4B5		TDIDIATE - Present in all isoforms	
	Average per Replicate (pmol/ μL)	CoV	Average per Replicate (pmol/ μL)	CoV
1-1	0.38	4.88%	0.81	59.67%
1-2	0.38	1.41%	0.78	60.39%
2-1	0.42	2.85%	0.80	54.99%
2-2	0.42	5.03%	0.84	59.27%
3-1	0.35	2.65%	0.67	68.18%
3-2	0.35	1.73%	0.72	67.58%
4-1	0.34	1.93%	0.78	66.76%
4-2	0.35	2.37%	0.67	61.74%
5-1	0.35	5.84%	0.67	69.69%
5-2	0.35	5.82%	0.66	59.34%
Average	0.37	9.01%	0.74	52.90%

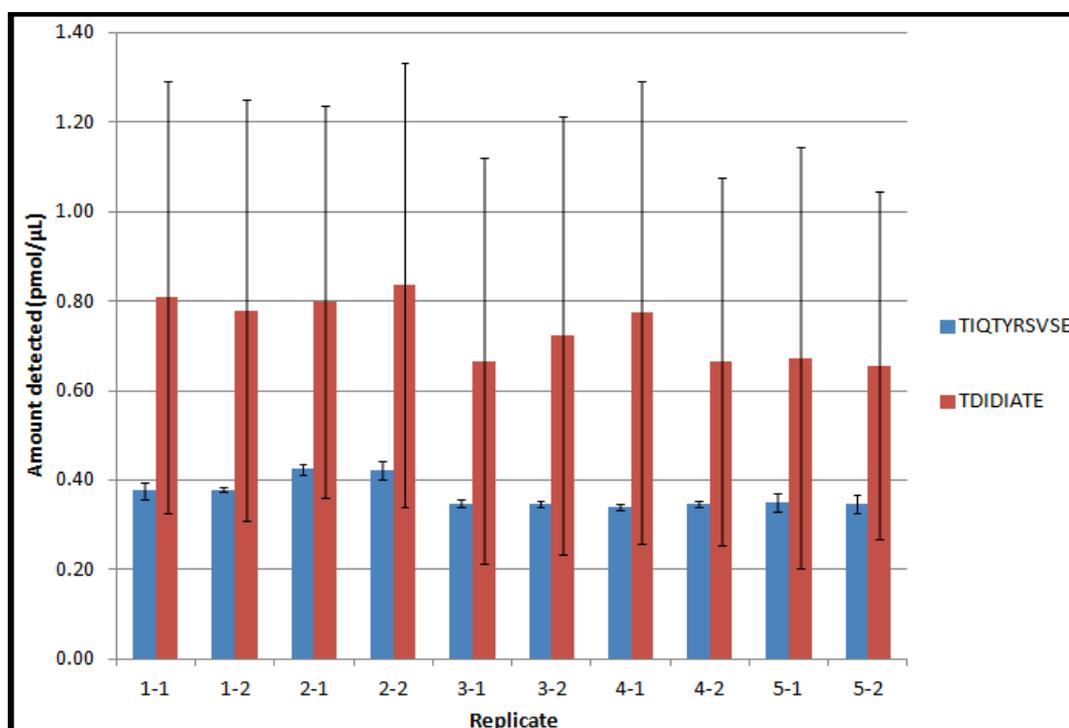


Figure 5-16: The quantities of each PDE4B2 target peptide detected during the spiked digestion of PDE4B2 transfected HEK293 IP on a QTrap 4000. The amount of each PDE4B2 target peptide detected during the spiked digestion of 25 μL of PDE4B2 transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between both each peptide and each replicate.

As is shown on Table 5-6 and Figure 5-16, despite PDE4B2 being digested with only one protease, and containing just two peptides, the results obtained for this isozyme appeared more variable than those achieved during the analysis of PDE4B1 (Table 5-5). While peptide TIQTYRSVSE achieved an average CoV of less than 10%, peptide TDIDIATE achieved a CoV of 52.9%, warranting further investigation.

So as to assess if there was any variation between the means of peptides TIQTYRSVSE and TDIDIATE, assuming a normal distribution, a two-tailed T-test was performed. During this test, a H_0 , that the means of these groups were equal, was stated. This analysis was performed based upon an alpha level of 0.05, giving a critical P-value of 2.1, which if exceeded, would require the H_0 to

be rejected. As is shown on Table 5-7, a P-value of 7×10^{12} was achieved, therefore the H_0 was rejected, indicating that the mean quantification values achieved for these peptides was different.

Table 5-7: A T-test based statistical analysis performed on the data obtained from the spiked digest of PDE4B2. A T-test based statistical analysis assessing the quantitation data obtained for peptides TIQTYRSVSE and TDIDIATE. This test was performed so as to assess if each data set contained an equal level of variation. Highlighted in yellow is the two-tailed P value, where a figure above 2.1 would suggest there to be significant variation between the two groups in 95% of cases.

	TIQTYRSVSE	TDIDIATE
Mean	0.369	0.74
Variance	0.000898889	0.0048
Observations	10	10
P(T<=t) two-tail	7.11469E-12	

Upon further investigating the data obtained for peptide TDIDIATE, fragment B₅ was found to be present at a level more than three times higher than that achieved for peptide TIQTYRSVSE (Table 5-6). However, when TDIDIATE fragment B₆ alone was analysed, the amount of peptide TDIDIATE detected in the sample was identical to the amount of peptide TIQTYRSVSE detected (Table 5-8).

Table 5-8: The amount of PDE4B2 detected, based upon the quantitation of for each TDIDIATE transition, during the spiked digestion of PDE4B2 transfected HEK293 IP on a QTrap 4000. The amount of peptide TDIDIATE detected within 25 μ L of PDE4B2 transfected HEK293 IP, based upon a quantitation performed with each of the three TDIDIATE transitions.

Replicate	TDIDIATE transition (pmol/ μ L)		
	B ₅	B ₆	B ₇
1-1	1.34	0.39	0.70
1-2	1.30	0.38	0.66
2-1	1.27	0.40	0.73
2-2	1.38	0.42	0.71
3-1	1.18	0.33	0.49
3-2	1.28	0.38	0.50
4-1	1.36	0.39	0.57
4-2	1.13	0.34	0.52
5-1	1.20	0.29	0.53
5-2	1.09	0.34	0.54
Average	1.25	0.37	0.60

When the MRM spectra obtained for fragment B₅ were analysed, it was found that several highly intense co-eluting peaks were present. In addition, fragment

B₅ was only the second most intense peak detected in the spectrum. It is therefore possible that beneath the target peak, there may also be a co-eluting matrix peak, which certainly would explain the higher than expected quantitation data obtained (Figure 5-17).

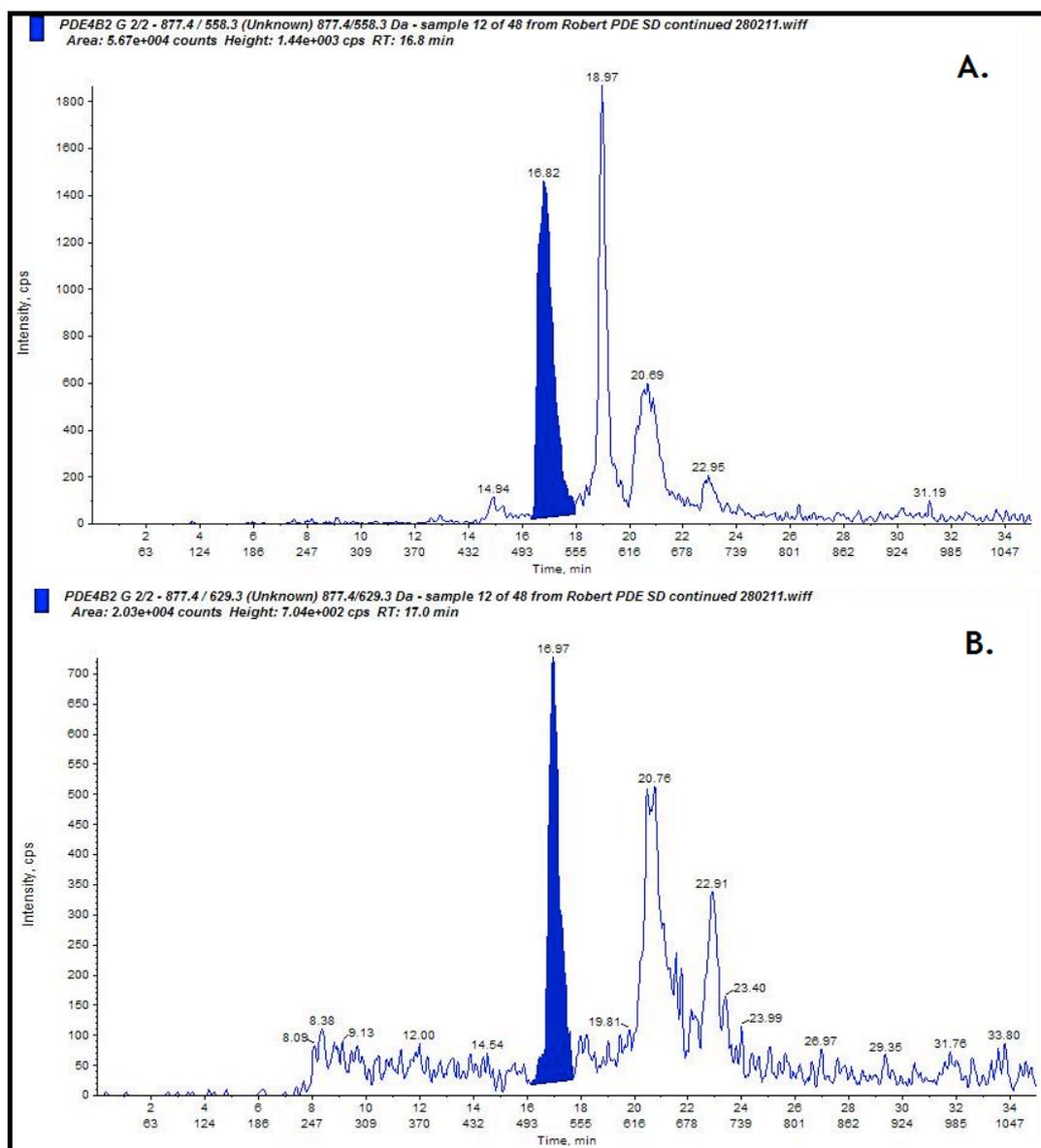


Figure 5-17: The XIC spectra obtained for TDIDIATE transitions B₅ and B₆, during the spiked digestion of 25 μ L of PDE4B2 transfected HEK293 IP. The XIC spectra obtained for TDIDIATE transitions B₅ (A) and B₆ (B) during the analysis of 25 μ L of PDE4B2 transfected HEK293 IP. Shown on spectrum A are the high intensity non-target peaks detected, eluting at a time similar to that of peptide TDIDIATE. While similar peaks were detected in spectrum B, these were of a lower intensity.

As the majority of the work conducted during this study was performed on a relatively insensitive API 2000, then both the Q1 and Q3 transmission windows were set to a low resolution, permitting ions within 1.2 ± 0.2 amu of the target m/z , to obtain the best possible limits of detection. This hardware profile, however, also permits more noise to reach the detector, including co-eluting peaks, which may interfere with an accurate quantification of the target and synthetic peptides. On transferring this acquisition method to a more sensitive MS, such as the QTrap 4000, it may be more prudent to select a more restrictive transmission window as to reduce excessive noise, such as that detected during the quantitative analysis of PDE4B2.

5.4.5.3 PDE4B3 Single Isoform Spiked Digest

Twenty-five micro litres of PDE4B3 transfected HEK293 cell IP was digested with either trypsin, Glu-C or Arg-C, in the presence of each of the PDE4B AQUA peptides. Three fragment ions were analysed per peptide, and the data averaged. A total of five digest replicates were performed and each replicate injected twice, the data from which is shown on Table 5-9.

Table 5-9: The quantity and variation data obtained during the spiked digestion of 25 μ L of immunoprecipitated PDE4B3 on a QTrap 4000. Through comparing the peak areas for both the synthetic and endogenous PDE4B peptides, three sets of quantities were obtained, per peptide, per replicate. These values were averaged, so as to obtain a single quantity, per peptide, per replicate, and a CoV determined for this value. Further to the above, an overall average was calculated for each value.

	TIQTYRSVSE - Present in all but PDE4B5		TDIDIATE - Present in all isoforms	
Replicate	Average per Replicate (pmol/ μ L)	CoV	Average per Replicate (pmol/ μ L)	CoV
1-1	1.99	1.31%	2.59	21.63%
1-2	2.01	2.63%	2.82	23.56%
2-1	5.10	0.88%	2.38	22.58%
2-2	5.11	2.14%	2.57	27.07%
3-1	3.20	3.88%	2.79	17.83%
3-2	3.13	2.12%	2.96	21.68%
4-1	3.03	0.97%	3.01	15.28%
4-2	3.02	0.83%	3.09	14.94%
5-1	5.76	2.80%	2.66	26.81%
5-2	5.76	1.67%	2.41	22.98%
Average	3.81	37.37%	2.73	19.82%
	VNPQEESYQK - Present in PDE4B1, 3 and 4		NSPCFFR - Unique to PDE4B3	
Replicate	Average per Replicate (pmol/ μ L)	CoV	Average per Replicate (pmol/ μ L)	CoV
1-1	4.40	38.63%	0.05	56.73%
1-2	5.45	17.81%	0.03	45.54%
2-1	4.02	4.56%	0.02	58.03%
2-2	3.81	10.66%	0.02	60.36%
3-1	5.29	14.43%	0.03	56.89%
3-2	4.99	14.94%	0.03	68.71%
4-1	6.43	2.24%	0.03	78.22%
4-2	6.21	4.43%	0.02	56.77%
5-1	4.80	9.24%	0.02	57.79%
5-2	5.06	6.85%	0.02	49.71%
Average	5.05	20.43%	0.02	65.15%

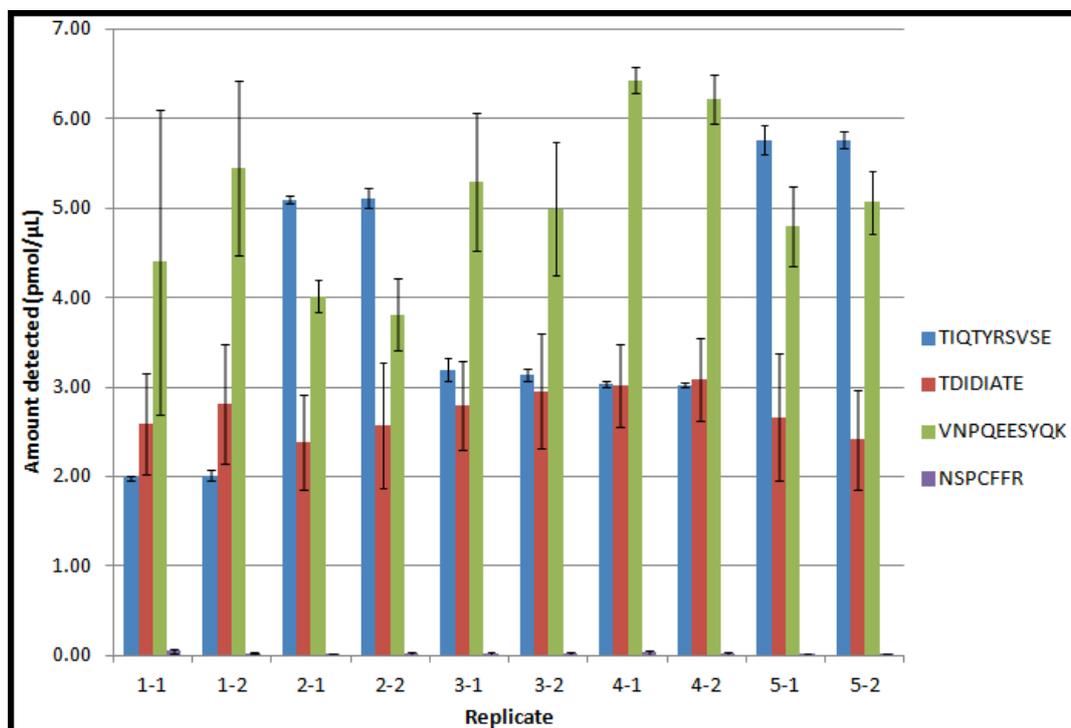


Figure 5-18: The quantity of each PDE4B3 target peptide detected during the spiked digestion of PDE4B3 transfected HEK293 IP on a QTrap 4000. The amount of each PDE4B3 target peptide detected during the spiked digestion of 25 μ L of PDE4B3 transfected HEK293 IP is shown. In brief, this single quantity, per peptide, was calculated through comparing the peak areas for three MRM transitions per peptide, per replicate, and the values averaged. A CoV value was determined for this averaged value. This bar chart helps to visualise the variation detected between each peptide and replicate.

From assessing the data shown on Table 5-9 and Figure 5-18, the levels of peptide NSPCFFR detected (digested with Arg-C) were seen to be extremely low. While this peptide contained a Cys residue, the amount of peptide detected was lower than expected, suggesting an inefficient digestion may have been responsible. Further Arg-C based digest optimisations should therefore be performed prior to the re-analysis of PDE4B3. Indeed, when the raw data obtained for peptide NSPCFFR was investigated (data not shown), no analyte peaks were detected, and background noise had instead been integrated, explaining the large CoV values obtained (with an average CoV value of 65.15%). Due to peptide NSPCFFR not being detected during this spiked digestion, and thus variation existing between the quantities of each peptide detected, an ANOVA based statistical analysis was not performed.

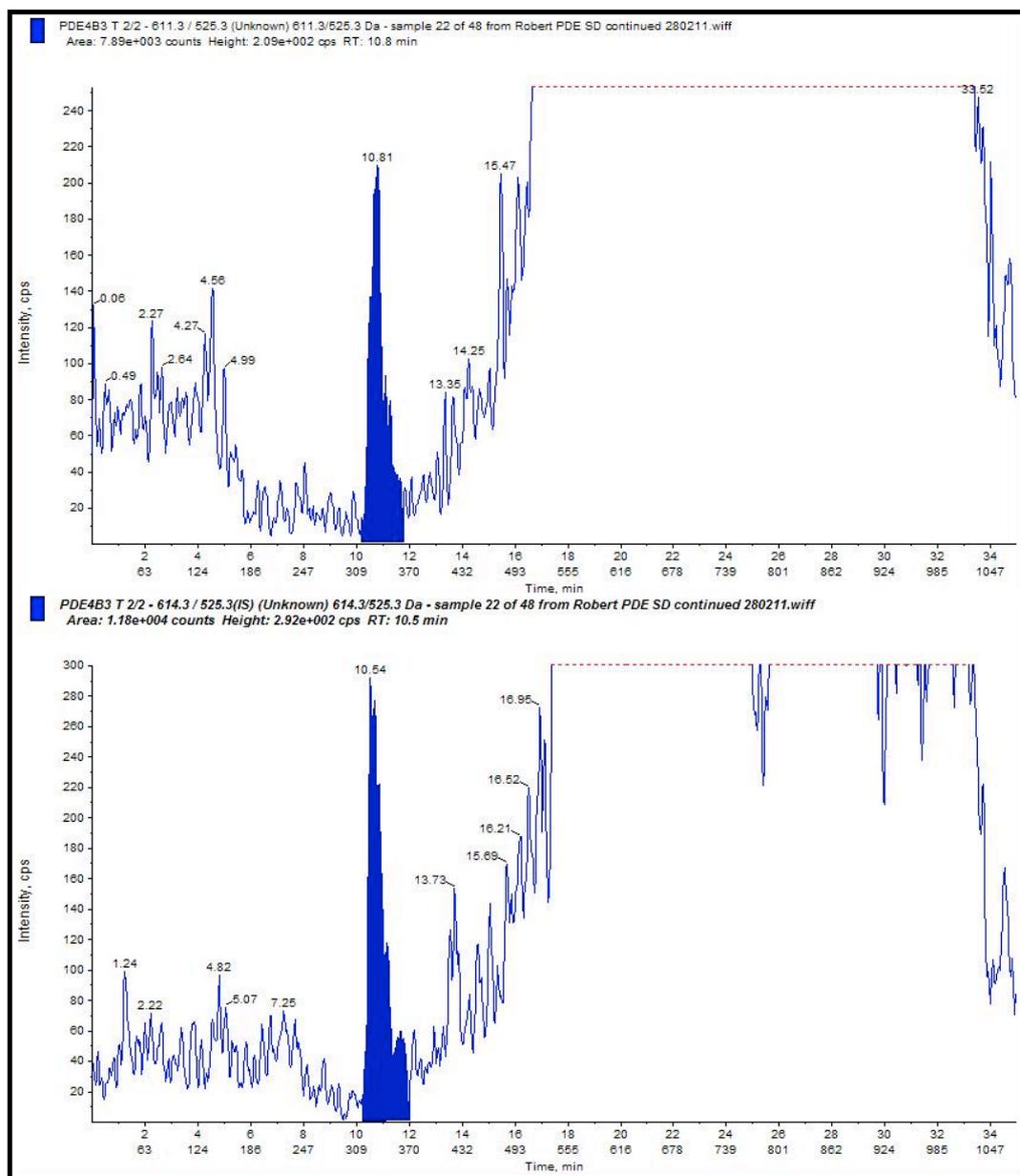


Figure 5-19: A typical peak detected for peptide VNPQEESYQK, fragment KQYS, during the spiked digestion of PDE4B3 on a QTrap 4000. A typical XIC spectra obtained for each VNPQEESYQK transition, in this case showing fragment KQYS. In addition to the low peak heights detected (<300 counts per second), the peaks were dwarfed by the background noise detected during these LC-MRM experiments.

In an attempt to further explain the variability of the data obtained during the analysis of PDE4B3 (Table 5-9), the individual XIC spectra were reviewed for each peptide and each replicate. From this review, the peak heights obtained for peptide VNPQEESYQK were noted to be considerably lower than those

achieved for the same peptide during the analysis of PDE4B isoform PDE4B1 (data not shown). While each transition of peptide VNPQEESYQK achieved a peak height greater than 1×10^5 during the analysis of PDE4B1, each achieved a peak height of less than 300 cps during the analysis of PDE4B3 (Figure 5-19). As low peak heights were achieved for both the IS and the target peptide, it would appear that an inefficient proteolytic digestion was not at fault. A more likely explanation was the inefficient recovery of peptide VNPQEESYQK from the gel, or an incomplete re-suspension of peptide VNPQEESYQK after drying.

When the data obtained from the Glu-C catalysed digestion of PDE4B3 alone was analysed, the data appeared just as variable (Table 5-9). While the levels of peptides TDIDIATE and TIQTYRSVSE were similar for replicates one, three and four, replicates two and five identified levels of peptide TIQTYRSVSE 87% and 110% higher, respectively, than the average quantity detected for peptide TDIDIATE. As the XIC spectra for each of the replicates appeared normal (data not shown), it may have been easy to attribute this quantitation pattern to the amount of AQUA peptide present within the sample being less than intended. The MS signal response to the AQUA peptide however was constant throughout this batch, if slightly decreasing (Figure 5-20), making this variation hard to explain.

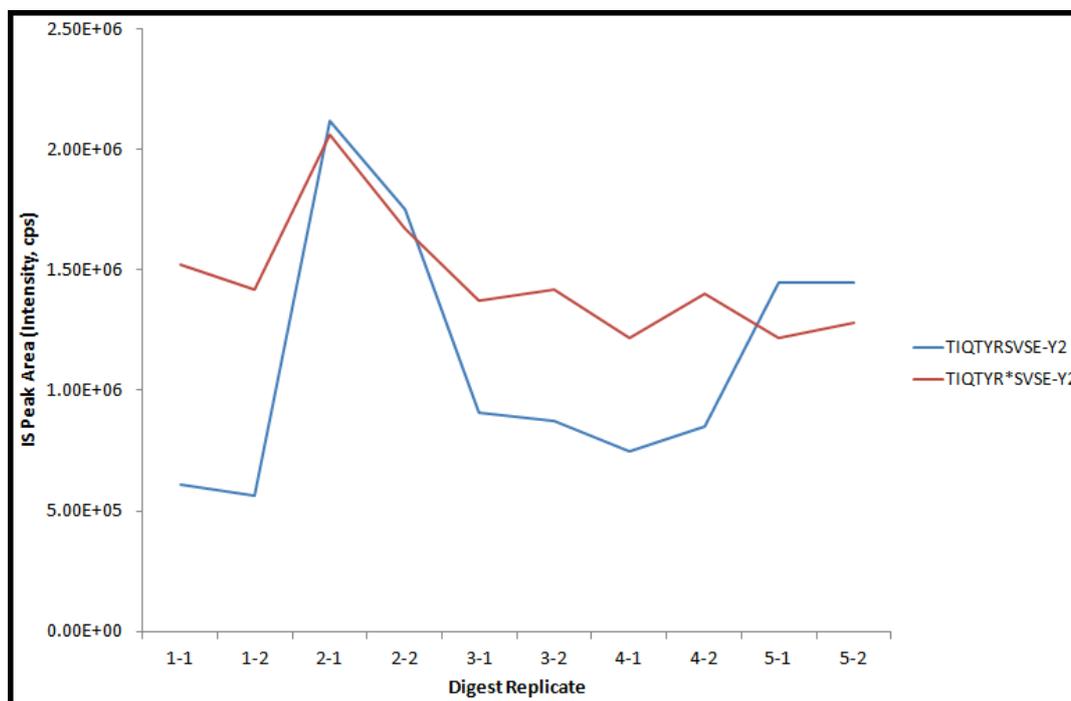


Figure 5-20: The MS signal intensity detected for both the TIQTYRSVSE IS and target peptide, over the duration of the PDE4B3 batch. Shown on this graph are the peak areas detected for both TIQTYRSVSE and TIQTYR*SVSE, fragment ion Y₂, over the duration of a PDE4B3 LC-MS/MS run. Interestingly, the IS signal intensity appears consistent, if slightly decreasing, while the target peptide response detected appears to increase during the analysis of replicates two and five.

5.5 Conclusion

Unlike the four A-Raf isoforms, where each contained a shared stretch of N-terminal sequence and the protein isoforms were differentiated on unique C-terminal sequence, each of the PDE4B isoforms contained a shared C-terminal, and were differentiated on a small amount of unique N-terminal sequence. In some ways this has made the quantification of the PDE4B isoforms more complicated, as a C-terminal peptide could not be selected as a last resort.

As long-PDE4B isoforms PDE4B1 and PDE4B3 PDE4B had the most unique sequence, it was possible to select unique AQUA peptides suitable for their quantification. While the literature would also suggest PDE4B4 to be a long-PDE4B isoform, no consensus sequence has as yet been published for this isozyyme, and indeed its expression in *Homo sapiens* had as yet to be proven.

Assuming PDE4B4 was indeed expressed, then it would certainly contain both UCRs 1, 2 and the PDE4B shared catalytic domain, requiring its quantification so as to enable a subtraction based quantification to be performed on short and super-short PDE4B isoforms PDE4B2 and PDE4B5, respectively. A peptide was thus selected from UCR1, present in long-PDE4B isoforms PDE4B1, PDE4B3 and PDE4B4, which through the subtraction of the quantitation values obtained for PDE4B1 and PDE4B3, should enable PDE4B4 alone to be quantified. Likewise, when the selection of a peptide unique to PDE4B2 could not be identified, a peptide was instead selected from the N-terminal of UCR2, present in each of the PDE4B isoforms but super-short PDE4B isoform PDE4B5, and thus enabling PDE4B2 alone to be quantified through the subtraction of the quantitation value obtained for PDE4B isoforms PDE4B1, PDE4B3 and PDE4B4. Finally, through selecting a peptide located in the C-terminal of each of the five PDE4B isoforms, the quantification of super-short PDE4B isoform PDE4B5 was possible, again based on the subtraction of the known quantitation values for each of the four other PDE4B isoforms.

In order to allow a spiked digest based analysis on each of the PDE4B isoforms, PDE4B1, PDE4B2 and PDE4B3 were transferred from Celltech pEE7 plasmids to higher copy number pcDNA3 and pcDNA3.1(+) vectors. The PDE4B plasmids were then FLAG-tagged so as to enable sample enrichment. Each plasmid was then amplified, transfected into HEK293 cells, over-expressed, and immunoprecipitated, ready for quantitation.

Each of the five PDE4B AQUA peptides were characterised and an MRM based MS acquisition method developed. During the spiked digest based analysis of the three PDE4B isozymes, several points of interest were noted. Firstly, the LC-MS/MS equipment took a significantly longer amount of time to equilibrate than was previously thought; suggesting that in future more blanks and system suitability tests should be run prior to the batch being submitted. Furthermore, with the improvements in sensitivity offered by the QTrap 4000 and QTrap 5500 over the API 2000, an increase in co-eluting peaks was also realised. In future it

may be prudent to reduce the MS background through restricting the Q1 and Q3 transmission windows to unit resolution.

Finally, the Arg-C catalysed digestion of PDE4B3 was discovered to be incomplete. Further optimisation of this critical step should therefore be undertaken prior to the analysis being repeated.

Several problems were encountered during the development of this quantitative technique, however. Firstly, despite the limited regions of unique sequence expressed within each of the PDE4B isoforms, it was possible to select an AQUA peptide suitable for the quantification of each. Secondly, each of these AQUA peptides selected was successfully characterised, a single MRM based acquisition method developed, and each target peptide and IS detected during the spiked digestion of the PDE4B isoforms. Finally, when the flawed and erroneous data had been explained, and thus could be removed, the spiked digest data obtained was accurate, both between peptides and indeed between proteases (trypsin and Glu-C).

In conclusion, further to additional optimisation, the acquisition method developed here should be suitable for the detection and quantification of endogenous PDE4B populations; assuming a sensitive triple quadrupole based MS is available and that further steps are taken to purify the sample, so as to reduce the presence of interfering peaks which led to inaccurate quantitative data throughout this project.

6 Absolute Quantification of the SERCA 2 Group of Protein Isoforms

6.1 Introduction

6.1.1 P-Type ATPase

P-type pumps constitute a superfamily of integrated membrane ATPases which are responsible for transporting ions and lipids across the cellular membrane based upon the hydrolysis of ATP (Palmgren and Nissen, 2011, Olesen et al., 2007). Indeed, the activity of P-type ATPases are thought to account for approximately one-third of the energy usage in humans (Rolfe and Brown, 1997). The action of such pumps are thus thought to be of critical importance within the body, enabling homeostasis of heavy metals and the asymmetric distribution of lipid within the membrane (Olesen et al., 2007). Early work conducted on P-type ATPases by Skou, found that by adding radiolabelled ADP to a Na^+/K^+ pump assay, that small amounts of radiolabelled ATP could be obtained. Skou thus rationalised that the intermediate step in the breakdown of ATP must be “the formation of a phosphorylated enzyme in which the phosphate is bound” (Skou, 1960, Palmgren and Nissen, 2011). Two distinct conformational states of phosphorylated ATPase intermediate were later identified, Enzyme1-P ($\text{E}_1\text{-P}$) which can be dephosphorylated through the addition of ADP, but which is insensitive to K^+ , and Enzyme2-P ($\text{E}_2\text{-P}$) which is sensitive to K^+ but insensitive to ADP (Whittam and Wheeler, 1970, Palmgren and Nissen, 2011). It is now known that the pumps E_1 state can be phosphorylated to yield $\text{E}_1\text{-P}$, and that following a change in conformation from $\text{E}_1\text{-P}$ to $\text{E}_2\text{-P}$, that a simultaneous movement of ions across the cell membrane will occur. Likewise, when the $\text{E}_2\text{-P}$ state is dephosphorylated back to E_2 , the change in conformation from E_2 to E_1 , will cause the simultaneous movement of ions (Albers, 1967), as is depicted on Figure 6-1.

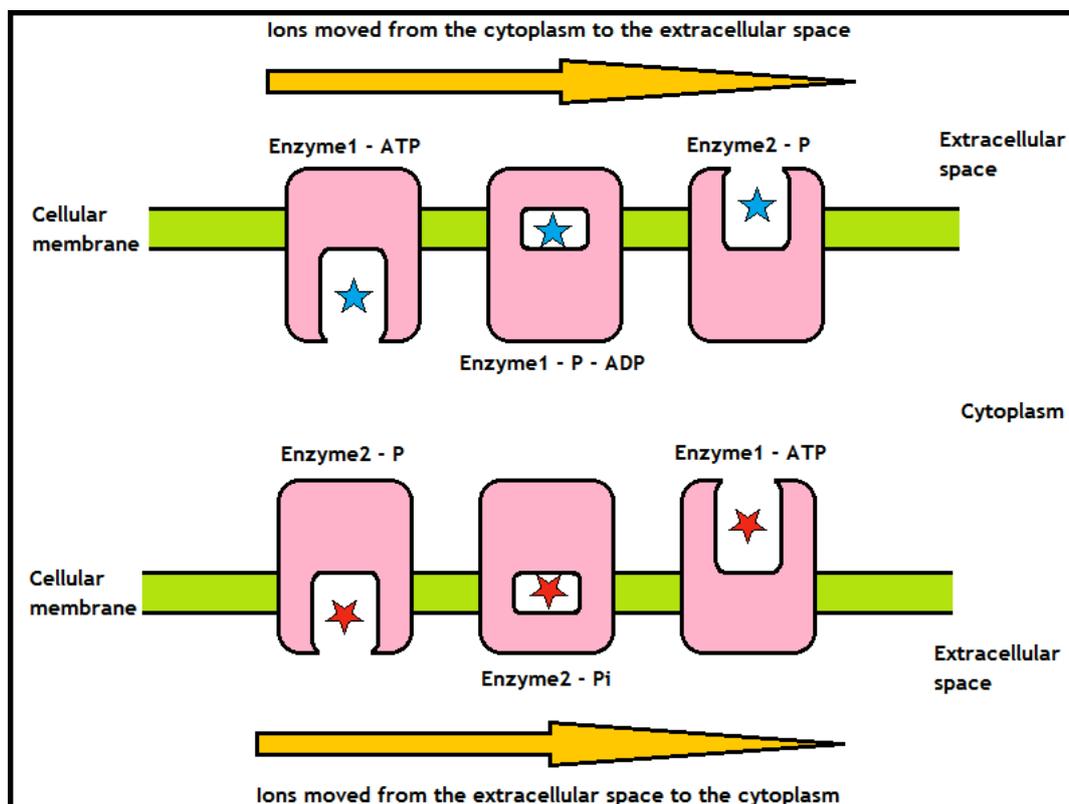


Figure 6-1: The cycle of changes in conformation undergone by P-type ATPases, enabling the transportation of ions. Depicted are the various conformations through which P-type ATPases will cycle while transporting ions both in and out of the cytoplasm. Conformational state “Enzyme1” initially accepts an ATP molecule, with which the nucleotide binding domain auto-phosphorylates the phosphorylation domain of the enzyme, yielding an ADP molecule and through changing in conformation, transporting ions from the cytoplasm to the extracellular space. Conformational state “Enzyme2” is then dephosphorylated by the actuator domain, yielding an inorganic phosphorous (Pi), and in doing so changing in conformation back to “Enzyme1” while transporting ions from the extracellular space to the cytoplasm. It should be noted that Enzyme1 and Enzyme2 are simply names given to the various conformational states of the P-type ATPase, not separate biomolecules.

In terms of structure, P-type ATPases have been shown to be elongated, with a length approximately three times that of their width. While the ATPases are embedded in the cell membrane, the majority of the protein forms a large cytoplasmic headpiece, with only a fraction of the protein exposed outside of the cell (Palmgren and Nissen, 2011). P-type ATPases have been shown to contain five functionally and structurally distinct domains, three of which are located within the exposed cytoplasmic headpiece, including the actuator, nucleotide binding and phosphorylation domains, while two are embedded in the

membrane, including the transport and class specific support domains. The final domain contained within P-type ATPases is the regulatory domain, which depending on the specific ATPase, can be located on either the C- or N-terminal and is capable of acting as an auto inhibitor to the ATPase pump molecule (Rolfe and Brown, 1997).

During each cycle of the P-type ATPase pump mechanism, it is the nucleotide binding domain which is responsible for phosphorylating the ATPase at the phosphorylation domain, while subsequent dephosphorylation is performed by the actuator domain (Rolfe and Brown, 1997).

6.1.2 Sarcoplasmic/Endoplasmic Reticulum Calcium ATPase

One such group of P-type ATPases are the P2A-ATPase family of Ca^{2+} ion transporters (Olesen et al., 2007), an evolutionarily conserved group of P-type ATPases which are thought to be present in all living organisms, from yeast to mammalian systems (Brandl et al., 1987). While these pumps most likely evolved as a mechanism for exporting Ca^{2+} ions against steep transmembrane gradients (Devés and Brodie, 1981), their main function in mammalian cells is in the relaxation of muscle cells by Ca^{2+} ion transport from the sarcoplasm, post-action potential, back into the sarcoplasmic reticulum. Muscle cells contract by releasing Ca^{2+} from the sarcoplasmic reticulum into the sarcoplasm, where they combine with troponin, so as to establish a cross-link between actin and myosin. For relaxation to occur, those same Ca^{2+} ions must dissociate from the myofibrils and be pumped back into the sarcoplasmic reticulum (Berchtold et al., 2000).

P2A-ATPase, or sarcoplasmic /endoplasmic reticulum calcium ATPase (SERCA), can also found in the endoplasmic reticulum (ER) of both muscle and non-muscle cells where it is responsible for maintaining calcium homeostasis by exporting two Ca^{2+} ions into the lumen of the endoplasmic reticulum per hydrolysed ATP molecule, counter-transporting four H^+ ions into the cytoplasm in the process (Skou, 1960, Whittam and Wheeler, 1970).

While just one SERCA gene has been identified in invertebrates (Periasamy and Kalyanasundaram, 2007), three have been described in mammalian systems. Each of these three genes are highly conserved and are located on different chromosomes, giving rise to three separate SERCA families (SERCA1-3) and some 11 protein isoforms (Pegoraro et al., 2011). Each isoform is thought to vary in its affinity towards Ca^{2+} , its Ca^{2+} turnover rate, and its tissue expression pattern, suggesting each may contribute to the unique physiological state of the tissue in which it is expressed (Periasamy and Kalyanasundaram, 2007).

SERCA1 has two known isoforms, SERCA1a, which encodes a 994 amino acid protein in adult fast-twitch skeletal muscle cells, and SERCA1b, which encodes a 1011 amino acid protein involved in the differentiation and growth of skeletal myogenic tissue in foetal populations (Brandl et al., 1987, Periasamy and Kalyanasundaram, 2007).

SERCA2, meanwhile, has three known protein isoforms and one recently described mRNA transcript. The first of these, SERCA2a, encodes a 997 amino acid protein which is expressed predominantly in both cardiac and slow-twitch skeletal muscle cells (MacLennan et al., 1985), while SERCA2b encodes a 1042 amino acid protein which is expressed in all tissues at a low level (Delabastie et al., 1990). SERCA2c encodes a 999 amino acid protein which is thought to be expressed in both heart and skeletal muscle cells (Dally et al., 2006, Periasamy and Kalyanasundaram, 2007). A SERCA2d transcript has also recently been described, which while having not yet been shown to yield a functional protein is thought to encode a 1007 amino acid protein (Kimura et al., 2005). Indeed, SERCA2 is thought to be the most studied of the P2A-ATPase families, especially in regard to its regulation (Albers, 1967).

SERCA3 has six known isoforms, SERCA3a, which encodes a 999 amino acid protein, SERCA3b, which encodes a 1042 amino acid protein, SERCA3c, which encodes a 1029 amino acid protein, SERCA3d, which encodes a 1044 amino acid protein, SERCA3e, which encodes a 1052 amino acid protein, and SERCA3f, which encodes a 1033 amino acid protein (Dally et al., 2010). Unlike the isoforms of

both SERCA1 and SERCA2, the function and tissue specific distribution pattern of each SERCA3 isoform has yet to be assessed. High levels of SERCA3 expression have however been documented in several non-muscle cells, including platelets, mast cells, T cells, epithelial cells, fibroblast cells, and endothelial cells (Pegoraro et al., 2011).

A small number of transmembrane-bound proteins capable of regulating the activity of SERCA have previously been identified; these including phospholamban, capable of regulating cardiac P2A-ATPase SERCA2a, and sarcolipin, capable of regulating fast-twitch skeletal muscle P2A-ATPase SERCA1a (MacLennan et al., 2003). These proteins are thought to act by interfering with a membrane-spanning segment (M2) of the transport domain within SERCA, stabilising the ATPase in its low affinity Ca^{2+} binding conformational state (E_2) (Palmgren and Nissen, 2011).

A recent study into the activity of sarcolipin on SERCA1a within rat muscle, found that the over expression of this protein caused a significant reduction in both twitch and tetanic peak force amplitude, a reduction in the maximum rate of contraction and relaxation, and increased fatigability with repeated electrical stimulation. This suggested that the over-expression of sarcolipin impaired the actions of SERCA1a, which as a result led to the depletion of the sarcoplasmic reticulum Ca^{2+} stores (Devés and Brodie, 1981).

6.1.2.1 SERCA2

SERCA2 is thought to be oldest of the SERCA P2A-ATPase Ca^{2+} pumps, and as such is widely distributed throughout nature (Devés and Brodie, 1981). Indeed, this may partially explain why it has the greatest number of modulator factors of all the SERCA ATPase families.

The gene which encodes SERCA2 in humans is thought to contain some 26 exons, with each of the four SERCA2 isoforms containing the first 20 exons prior to exhibiting a unique C-terminal transcription pattern. SERCA2a has been shown to

splice out exons 21, 23, 24 and 25, combining exon 20 and exon 22 before terminating during the translation of exon 26. SERCA2b meanwhile has been found to be coded for by three different mRNAs, each of which has been shown to remove only exon 21, therefore combining exons 20 and 22 before terminating during the translation of exon 23 (Brini and Carafoli, 2009), while SERCA2c has been shown to splice out exons 21 and 23, combining exons 20 and 22 before terminating during the translation of exon 24. Finally, assuming SERCA2d is expressed (Albers, 1967), this recently discovered isoform has been predicted to contain exon 21, and to terminate just six amino acids into exon 22 (as a result of the alternative reading frame utilised during the translation of this exon) (Dally et al., 2010, Kimura et al., 2005), as is shown on Figure 6-2.

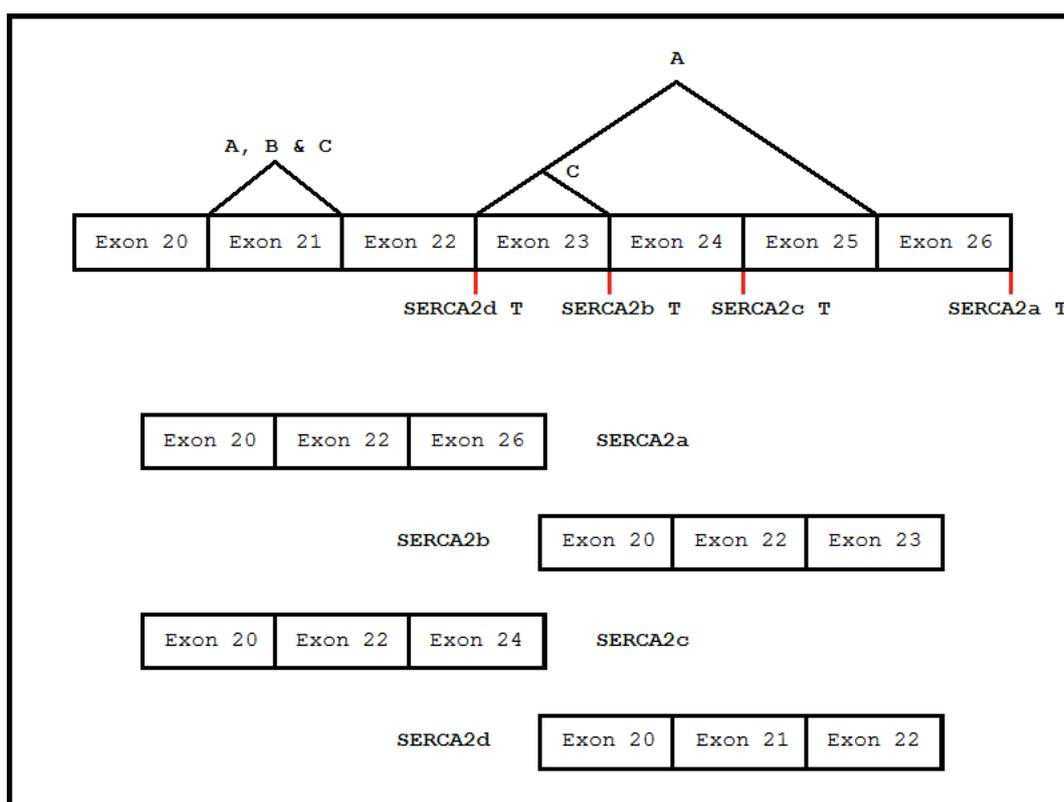


Figure 6-2: The exon expression patterns of each of the SERCA2 isoforms. The pattern of exons expressed by each of the SERCA2 isoforms, where SERCA2a contains exons 20, 22 and 26, splicing out exons 21, 23, 24 and 25, while SERCA2b contains exons 20, 22 and 23, splicing out exon 21. SERCA2c contains exons 20, 22 and 24, splicing out exons 21 and 23, while SERCA2d contains exons 20, 21 and 22. The letter T is used to indicate the introns termination site.

It is because of this mRNA sequence homology that the first 993 amino acids of each SERCA2 isoform are identical, each differentiated based solely upon a unique C-terminal. SERCA2a has been detailed to contain a unique four amino acid C-terminal, resulting from the incorporation of exon 26, while SERCA2b has been found to contain a 49 amino acid unique C-terminal, resulting from the inclusion of exon 23. SERCA2c had been found to contain a six amino acid unique C-terminal resulting from the inclusion of exon 24, while SERCA2d is predicted to express a 54 amino acid unique C-terminal, resulting from the inclusion of exons 21 and 22, with exon 22 being translated based on an alternative reading frame (Delabastie et al., 1990).

6.1.2.1.1 SERCA2a

SERCA2a is thought to be expressed in both cardiac and slow-twitch skeletal muscle cells (MacLennan et al., 1985). It is also thought to be the major Ca^{2+} ATPase within the heart, accounting for more than 50% of the total membrane protein within the sarcoplasmic reticulum and thus is responsible ensuring normal contractility in the working myocardium (Dally et al., 2010).

A recent study into cardiac dysfunction led to an experiment being conducted in transgenic mice in which either SERCA1a or SERCA2a was over-expressed. It was found that even when the mRNA encoding SERCA2a was increased 2-4 fold that only a 20 - 30% increase in endogenous SERCA2a expression was detected, leading to both an increase in the velocity at which Ca^{2+} was transported into the sarcoplasmic reticulum and an increase in maximum rate of contraction and relaxation within the heart. When SERCA1a was over-expressed by 2-2.5 fold however, the expression of SERCA2a was seen to drop by 50%, also leading to a decrease in L-type Ca^{2+} channel expression. This would suggest that SERCA2a can be substituted by SERCA1a within the heart, and that isoform specificity is not absolute (Periasamy and Huke, 2001). Furthermore, a study conducted into the intracellular location of the SERCA2 isoforms within human cardiomyocytes found SERCA2a to be present in both transverse and longitudinal immunostaining assays. This would suggest that SERCA2a is widely expressed throughout the

sarcoplasmic reticulum, including those locations which are in close proximity to the T-tubules (Whittam and Wheeler, 1970).

6.1.2.1.2 SERCA2b

SERCA2b is thought to be expressed at a low level in both muscle and non-muscle cells and as such is thought to fulfil a house-keeping role in maintaining cellular Ca^{2+} levels (Lyttton et al., 1992, Delabastie et al., 1990). It has been suggested however that the expression of SERCA2b may be induced in response to cellular stress (Wang et al., 2011).

When both SERCA2a and SERCA2b were compared in terms of their ability to transport Ca^{2+} , SERCA2b was found to have a twofold higher affinity towards Ca^{2+} than SERCA2a, but also to have a 50% lower Ca^{2+} turnover rate (Wang et al., 2011). This variation in both affinity and Ca^{2+} turnover is thought to result from the varying C-terminal sequences expressed by each isoform, where the 49 amino acids unique to SERCA2b are mostly hydrophobic and have been shown to create an extra transmembrane segment within the structure of the protein, shifting the C-terminal of the isoform from the cytoplasm to the lumen. This change in conformation is thought to stabilise SERCA2b within its Ca^{2+} binding conformation (E_1), accounting for the high Ca^{2+} binding affinity of this isoform (Pegoraro et al., 2011).

The ability of SERCA2b to act as a substitute for SERCA2a within cardiac muscle cells has also been investigated. On assessing the expression levels of SERCA2 within SERCA2a $-/-$ transgenic mice, the expression levels of SERCA2b were seen to increase, however the overall expression of SERCA2 within the cell was only 50% of that of the wild type. In 40% of cases, this resulted in embryonic or neonatal mortality, while those specimens which did survive showed mild concentric hypertrophy and impaired relaxation and contraction kinetics (Ver Heyen et al., 2001).

In regard to the intracellular expression pattern of SERCA2b, SERCA2b was detected only during the transverse immunostaining of human cardiomyocytes, suggesting the expression of this protein may be restricted to those areas of the sarcoplasmic reticulum which run parallel to both the actin and myosin, but not those areas which are in close proximity to the T-tubules (Whittam and Wheeler, 1970).

6.1.2.1.3 SERCA2c

SERCA2c is thought to be expressed in both heart and skeletal muscle cells and is thought to exist only in *Homo sapiens* (Whittam and Wheeler, 1970, Dally et al., 2006, Periasamy and Kalyanasundaram, 2007). While very little data has been published detailing the specific function of SERCA2c, SERCA2c is thought to have a lower binding affinity for Ca^{2+} than either SERCA2a or SERCA2b, but to have a high Ca^{2+} turnover rate, similar to that of SERCA2b (Whittam and Wheeler, 1970).

In regard to the intracellular expression pattern of SERCA2c, SERCA2c was found to be restricted to the intercalated discs and the subplasmalemmal areas of human cardiomyocytes. Indeed, these are the areas of the cell where the cytoplasmic Ca^{2+} concentrations are highest, which helps to explain both the expression of SERCA2c in those areas and the lower binding affinity of this isoform towards Ca^{2+} (Skou, 1960, Whittam and Wheeler, 1970).

Proceeding from the end of exon 20 (the point at which SERCA2d differentiates through the inclusion of exon 21) SERCA2c incorporates exon 22 (present in SERCA2a, b & c) before terminating during the translation of exon 24 (which provided SERCA2c with its unique six amino acid c-terminal).

6.1.2.1.4 SERCA2d

A new SERCA2 mRNA has recently been described, SERCA2d, which is predicted to code for a 1007 amino acid protein. Unlike SERCA2a, b and c, SERCA2d

contains exon 21, leading to a frame shift in the translation of exon 22, transcribing a unique C-terminal before encountering a premature stop codon (Kimura et al., 2005). The predicted protein sequence for SERCA2d has been described by both (Berchtold et al., 2000) and (Albers, 1967), however, no protein product has as yet been identified for this isoform.

6.1.2.1.5 SERCA2 Oxidation as a Disease Biomarker

Nitric oxide (NO) is produced within the body through the degradation of Arg to L-citrulline based upon the action of nitric oxide synthase (NOS) (Knyushko et al., 2005). Indeed both NO and L-citrulline are of great biological importance, with L-citrulline being involved in the efficient removal of ammonia from the body and NO being involved in the modulation of blood flow, thrombosis, and neural activity (Luiking et al., 2010). NO can however react with a mitochondrial by-product superoxide ($O_2^{\cdot-}$) to form a powerful oxidant, peroxynitrite ($ONOO^-$), capable of nitrating tyrosine residues to 3-nitrotyrosine. Indeed 3-nitrotyrosine modified proteins have been observed in over 80 different pathologies in a variety of tissues (Ischiropoulos, 1998).

The levels of 3-nitrotyrosine modified SERCA2a (nitrated at both Tyr²⁹⁴ and Tyr²⁹⁵) within skeletal muscle has recently been shown to increase with age, with young adults expressing 1.0 ± 0.5 mol of nitrotyrosine/mol of SERCA2a, in comparison to the elderly where 3.5 ± 0.7 mol of nitrotyrosine/mol of SERCA2a has been measured. This correlates to a 40% loss in Ca-ATPase activity (Knyushko et al., 2005). The nitration of SERCA2a within cardiac muscle has also been investigated because mitochondria account for approximately $1/3$ of the volume of a cardiomyocyte, and the potential for peroxynitrate to modify SERCA2a (which accounts for up to 50% of the total membrane protein within the sarcoplasmic reticulum (Dally et al., 2010)) is increased (Lokuta et al., 2005). This study found that nitrated SERCA2a was increased in dilated cardiomyopathy (DCM) related heart failure as a result of Ca^{2+} pump failure (Lokuta et al., 2005).

6.2 Project Aims

In addition to developing a technique capable of quantifying each of the SERCA2 isoforms in a range of tissues, a peptide will also be selected which is capable of quantifying the frequency at which SERCA2 is nitrated. Therefore, the development of a SERCA2 based quantitative technique should allow for the assessment of: (i) Which SERCA2a isoforms are expressed within a tissue, and whether SERCA2d is translated into a functional protein (by detecting isoform specific target peptides). (ii) The concentration at which each SERCA2 isoform is expressed (by detecting both the isoform specific target and synthetic AQUA peptide), and (iii) How both the SERCA isoform expression ratios, and the frequency at which SERCA2 is nitrated, vary under different physiological conditions. As such, during this part of the project I aim to: (i) Screen each of the previously detailed SERCA2 isoforms, identifying and selecting suitable AQUA peptides for the absolute quantification of each protein, and (ii) Screen SERCA2 with the aim of identifying a suitable AQUA peptide for the absolute quantification of nitrated Tyr²⁹⁴ and Tyr²⁹⁵.

6.3 SERCA2 Specific Methods

6.3.1 3-Nitrotyrosine Generation

Some 25 pmol of AQUA peptide GLNQDCCVVYR was reduced and alkylated, with DTT and IAA, respectively. The alkylated peptide was nitrated with either: (i) 3-Morpholinopyridone hydrochloride (SIN-1) for one hour at 37 °C, as has been described by (Richards et al., 2006); (ii) Tetranitromethane (TNM) for one hour at 37 °C, as has been described by (Yamada et al., 1990), or with (iii) Sodium nitrite (NaNO₂) for one hour at 37 °C, as has been described by (Ohshima et al., 1990). Each reaction was performed at a 1:1 ratio, a 1:500 ratio and a 1:1000 ratio of peptide to nitrating agent (mol:mol), prior to undergoing a Zip-Tip based cleanup and a MALDI-ToF based analysis on an Ultraflex II.

6.4 Results and Discussion

6.4.1 AQUA Peptide Selection

While each of the five PDE4B isoforms had unique N-terminals, with PDE4B1 containing 93 residues of unique sequence, PDE4B2 containing 39 residues of unique sequence, PDE4B3 containing 78 residues of unique sequence and PDE4B5 containing 15 residues of unique sequence, the SERCA2 isoforms are much more homologous, with SERCA2a containing only four residues of unique sequence, SERCA2b containing 49 residues of unique sequence, SERCA2c containing only 6 residues of unique sequence and SERCA2d predicted to contain 54 residues of unique sequence, making the selection of suitable AQUA peptides increasingly difficult.

While the A-Raf family of protein isoforms were also differentiated based upon unique C-terminal sequence, A-Raf WT contained an extensive unique C-terminal sequence and A-Raf Short contained intronic sequence, each reducing the homology of those particular isoforms, and thus increasing the likelihood of a suitable AQUA peptide being identified.

Through the use of the AQUA peptide selection workflow, four AQUA peptides were selected for the quantification of each of the four SERCA2 isoforms and a single additional AQUA peptide selected for the quantification of nitrated SERCA2 (Table 6-1).

Table 6-1: The AQUA peptides selected for the quantification of each of the SERCA2 isoforms and nitrated SERCA2. Each of the AQUA peptides selected for the quantification of each of the four PDE4B isoforms.

Peptide	Protease	Present in:	Comments:
NYLEPAILE	Trypsin	SERCA2a	As this peptide has an N-terminal Asn residue, then it should be ordered with extended sequence to the N-terminal and cleaved with trypsin during sample preparation.
FVARNYLEPGK	Lys-C	SERCA2b	Located within the unique C-terminal of SERCA2b.
NYLEPVLSSSEL	Trypsin	SERCA2c	As this peptide has an N-terminal Asn residue, then it should be ordered with extended sequence to the N-terminal and cleaved with trypsin during sample preparation.
CVSAHLP	Asp-N	SERCA2d	Located within the unique C-terminal of SERCA2d.
IRGAIYYF	Glu-C	All four SERCA2 isoforms	Cleaved with Chymotrypsin, which has documented to undergo non-specific proteolytic cleavage.

6.4.2 SERCA2 Nitration

As pre-nitrated Fmoc SPPS Tyr residues are commercially available, it should be possible to chemically synthesise a pre-nitrated version of AQUA peptide IRGAIYYF. If an accurate quantification is to be undertaken on the frequency at which each Tyr residue is nitrated *in vivo* however, several nitrated AQUA peptide species must be synthesised, one of which is only nitrated at Tyr²⁹⁴, one which is only nitrated at Tyr²⁹⁵, one which is nitrated at both Tyr²⁹⁴ and Tyr²⁹⁵, and one which is nitrated at neither Tyr²⁹⁴ or Tyr²⁹⁵. For this quantitative technique to be validated however, an intact nitrated SERCA2 protein must also be produced.

Due to the various pathologies in which 3-nitrotyrosine modified proteins have been observed (Ischiropoulos, 1998), several techniques already exist which enable the nitration of intact proteins under relatively mild conditions (Gow et al., 1997), these include the application of SIN-1, TNM or NaNO₂. Of these three nitration methods, the process through which SIN-1 is applied to the nitration of Tyr best mimics that which occurs *in vivo*. In brief, when SIN-1 is introduced to a sample with a pH greater than pH 7 (Bohn and Schönafinger, 1989), SIN-1 spontaneously decomposes. In doing so, it consumes oxygen and releases equimolar volumes of both NO and O₂^{·-} (Richards et al., 2006, Ashki et al., 2008). SIN-1 based nitration has, however, been found to increase the level of 3-nitrotyrosine in HepG2 cells from 23.2% to just 59.2% (Yasuda et al., 2007) and thus is somewhat inefficient.

While TNM is used in the manufacture of explosives, the technique through which TNM is applied to the nitration of Tyr is also considered mild, in that the reaction is conducted under mild pHs, ionic strengths and temperatures (Subbarao and Kenkare, 1977). Indeed TNM is seen as being the method of choice for protein nitration (Ischiropoulos, 1998). While TNM is thought to react specifically with Tyr, oxidation of Cys has also been documented (Sokolovsky et al., 1966). It is therefore imperative that Cys residues are blocked with IAA prior to the application of TNM. Indeed, while Trp and Met can also readily undergo oxidation, their oxidation has not been recorded during TNM based Tyr nitration.

Furthermore, as none of the AQUA peptides selected for the quantification of SERCA2 contain either Trp or Met, the use of this nitrating agent would appear ideal.

As TNM has been documented as being carcinogenic, any un-reacted TNM will be quenched through the application of 2-mercaptoethanol (Klevan and Tse, 1983), which is capable of quenching TNM, but will not to reverse the nitration (Subbarao and Kenkare, 1977).

Finally, the acidification of NaNO_2 can also be applied to the nitration of Tyr. This technique had been found to yield efficiencies of up to 90% when the two components are combined at an equimolar ratio (Radabaugh et al., 2008). While this reaction is conducted at a pH lower than pH 1, and thus is not particularly mild, this method of nitration is both less toxic than TNM and more efficient than SIN-1 (Oldreive et al., 1998).

6.4.2.1 Optimisation of Tyr Nitration

In order to test the relative efficiency with which SIN-1, TNM and NaNO_2 are capable of nitrating a Tyr containing peptide, A-Raf AQUA peptide GLNQDCCVYR was subjected to nitration under three different ratios of peptide:nitrating agent (mol:mol), including 1:1, 1:500 and 1:1000.

While only an unmodified parent ion peak was detected during the SIN-1 based nitration (data not shown), confirming this technique to be too inefficient for the nitration of SERCA2, the spectra obtained from the TNM based nitration contained no peaks (data not shown), not even those of the unmodified parent ion. This would suggest the TNM was either too harsh for use in the nitration of the Tyr containing peptide, possibly degrading the peptide into smaller constituents (though none were identified on the mass spectrum), or that the TNM interfered with the Zip Tip based purification of the nitrated peptide, explaining the blank spectrum.

The NaNO₂ based nitration of AQUA peptide GLNQDCCVYR appeared more suitable however, yielding three previously unidentified peaks with mass increases of +14 Da, +29 Da and +45 Da, as is shown on Figure 6-3. While the +45 Da and +29 Da peaks are indicative of nitration, with the +49 Da increase equating to the addition of NO₂ and the loss of a hydrogen, and the +29 Da increase equating to the addition of NO, and the loss of a hydrogen, the +14 Da peak is somewhat unexpected. While +14 Da would suggest the addition of nitrogen, nitrogen has a valency of three and thus its binding to a single carbon atom on the Tyr residue would be unfavourable. Instead, the +14 Da increase may indicate an *ortho*-quinone like oxidation product of the Tyr residue (Steinmann et al., 2012), as is shown on Figure 6-4. An alternative theory has also been proposed in which both the +29 and +14 Da peaks actually related to the decomposition of the +45 Da nitrated peak during the MALDI-ToF based analysis (Lee et al., 2007).

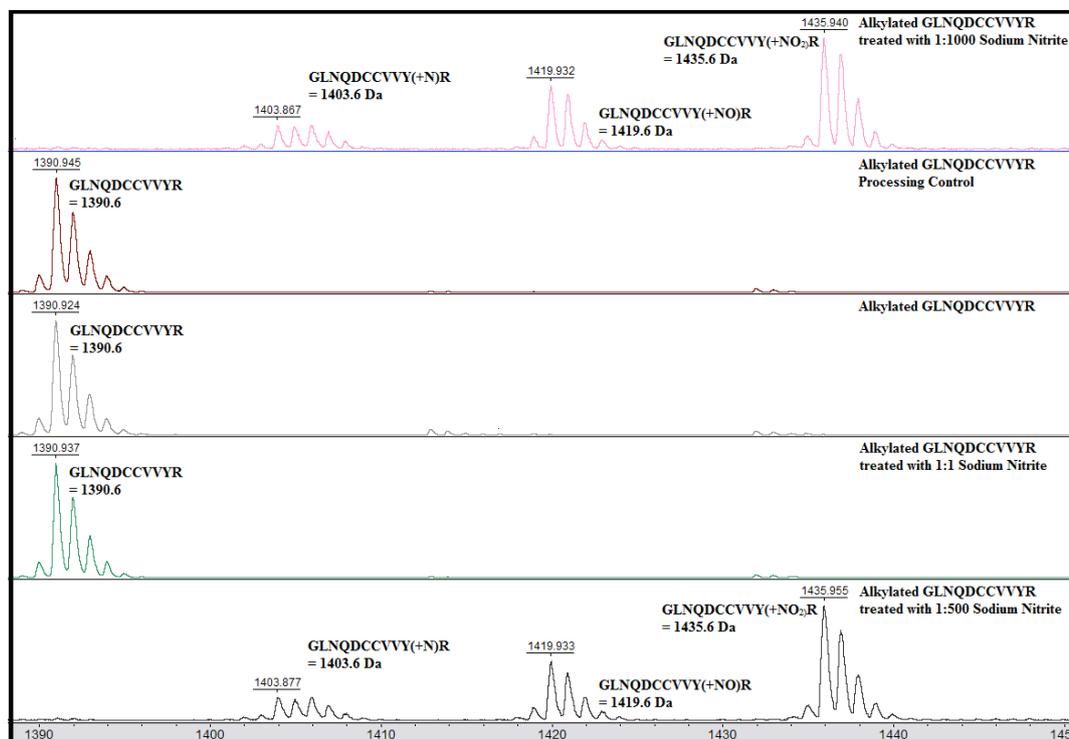


Figure 6-3: The nitration of alkylated AQUA peptide GLNQDCCVVYR, through the acidification of NaNO_2 . So as to test the efficiency with which NaNO_2 nitrates Tyr residues within a peptide, an existing Tyr containing AQUA peptide, GLNQDCCVVYR, originally selected for the quantification of A-Raf isoforms A-Raf WT, DA-Raf-1 and DA-Raf-2, was modified. As can be seen above, when the unmodified peptide, the processing control (exposed to TFA, vacuum centrifuged to completion and purified via Zip Tip) and the 1:1 (mol:mol) NaNO_2 nitrated peptide were analysed via MALDI-ToF, each was found to yield only the GLNQDCCVVYR parent ion. When both the 1:500 and 1:1000 NaNO_2 nitrated peptides were analysed however, the parent ion was no longer detectable, instead being replaced with three peaks, one at +14 Da, one at +29 Da and one at +45 Da.

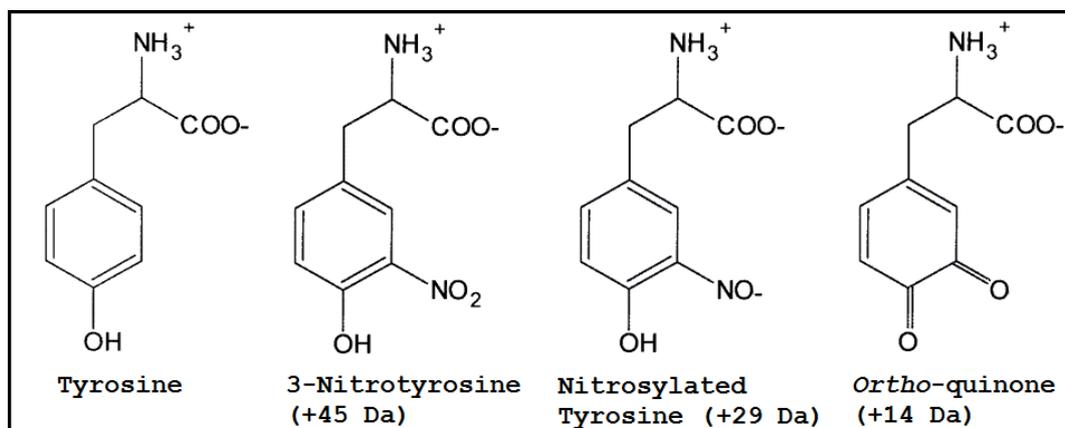


Figure 6-4: The possible nitration products of AQUA peptide GLNQDCCVVYR, when exposed to acidified NaNO_2 . When AQUA peptide GLNQDCCVVYR was nitrated with NaNO_2 , three previously undetected peaks were identified, including one at +45 Da, most likely resulting from the production of 3-nitrotyrosine, one at +29 Da, most likely from the introduction of an NO group and the loss of hydrogen. Finally a +14 Da peak was identified, which may result from either the production of *ortho*-quinone or from the MALDI-ToF based degradation of either the +29 or +45 Da peaks.

In conclusion, of the three nitrating agents applied to the nitration of AQUA peptide GLNQDCCVVYR, including SIN-1, TNM and NaNO_2 , only NaNO_2 has been shown to yield nitrated tyrosine. While this reaction may not be complete, or may be complete but undergoes degradation during MALDI-ToF, this technique should be capable of creating a small amount of nitrated SERCA2, suitable for use in the optimisation of the AQUA based quantification.

6.5 Conclusion

As the SERCA2 family of proteins was the last of the three groups of isoforms to be analysed, much less work has been completed for SERCA2 than either A-Raf or PDE4B. This group of protein isoforms has, however, proven both interesting and unique, requiring the selection of longer than necessary AQUA peptides, which are co-digested with the target protein, due to the presence of N-terminal Asn residues, the use of Chymotrypsin, a protease which has yet to be utilised during this study, and the selection of an AQUA peptide unique to SERCA2d, a predicted protein isoform which as yet has not been proven to yield a functional protein.

Furthermore, due to the description of nitrated Tyr²⁹⁴ and Tyr²⁹⁵ in various physiological conditions, a peptide sequence was also selected which is capable of quantifying the various nitrated and un-nitrated states at which SERCA2 isoforms may present, enabling the ratio of SERCA2 nitration to be assessed under a range of physiological conditions.

The development of this AQUA based technique for the quantification of the SERCA2 isoforms can thus be seen as successful, with a single AQUA peptide having been selected for the quantification of each protein species.

7 Conclusion

From the outset, the aims of this project have been to identify several families of protein isoforms, and to select suitable AQUA peptides for the quantification of each. Following the development of an MRM based LC-MS acquisition method capable of quantifying each protein isoform within a single LC-MS run, it was intended that a HEK293 cell lysate would be analysed, enabling the detection and quantification of each endogenous protein isoform.

While three families of protein isoforms have been analysed during this project, the rate at which work has progressed on each has been variable. The first set of protein isoforms to be analysed were the four A-Raf protein isoforms, as described in Chapter 4, for which the majority of the project aims were fulfilled. Indeed, in addition to aligning the sequence of each A-Raf protein isoform and selecting an AQUA peptide suitable for the quantification of each, each AQUA peptide was also characterised and optimised; so as to enable the development of an MRM based MS acquisition method. Furthermore, upon performing SOE PCR on A-Raf Short, each A-Raf isoform was expressed in transfected HEK293 at a quantity suitable for detection during this project. Each transfected HEK293 cell lysate was subjected to FLAG-tag immunoprecipitation, and the enriched protein used as a basis for single isoform based spiked digests, so as to assess the reproducibility of this quantitative technique. Finally, the optimised A-Raf MS acquisition method was applied to the quantification of both non-immunoprecipitated exogenous A-Raf expression in HEK293, and endogenous A-Raf expression in HEK293, though both of these analyses failed to detect either A-Raf isoform, due to the limitations of the mass spectrometer utilised.

The second set of protein isoforms to be analysed meanwhile were the four confirmed and one predicted PDE4B protein isoforms, as described in Chapter 5, and while the same level of progress was not achieved, the protein sequence for each PDE4B isoform was aligned, and a suitable AQUA peptide selected for the quantification of each. Each of the PDE4B AQUA peptides were then characterised and optimised; so as to enable the development of an MRM based

MS acquisition method. Furthermore, upon transferring each PDE4B insert to pcDNA3 or pcDNA3.1(+), and FLAG-tagging each insert, so as to enable a rapid enrichment, each PDE4B isoform was expressed in transfected HEK293 at a level suitable for detection during this project. Finally, each transfected HEK293 cell lysate was FLAG-tag immunoprecipitated, using the enriched protein as a basis for single isoform based spiked digests, so as to assess the reproducibility of this quantitative technique. Indeed, while the MS acquisition method developed for the quantification of PDE4B was complete, this technique was not applied to the absolute quantification of endogenous PDE4B expression, due both to the negative results obtained from the endogenous A-Raf quantification, but also as a result of time constraints.

Finally, the last set of protein isoforms to be analysed were the three confirmed and one predicted SERCA2 protein isoforms, as detailed in Chapter 6, and while this set of protein isoforms was explored only in brief, due to time constraints, the sequence for each SERCA2 isoform was aligned, and a suitable AQUA peptide selected for the quantification of each. Furthermore, an AQUA peptide was also selected for the quantification of nitrated SERCA2, a PTM which has been linked to more than 80 different pathologies in a variety of *Homo sapiens* tissues (Ischiropoulos, 1998).

Indeed, of the 11 confirmed protein isoforms and two predicted protein isoforms explored during this study, it has been possible to select a suitable AQUA peptide for the quantification of each, although each family of protein isoforms presented the author with unique challenges. For example, when the four A-Raf protein isoforms were explored, each of which exhibited a homologous N-terminal, and three of which were translated from a truncated version of the A-Raf WT mRNA, it was necessary to explore the quantification of unique C-terminal peptides, subtraction based quantification, the selection of peptides which included reactive amino acids and the selection of peptides cleaved through the use of non-ideal proteases. To the author's knowledge no publications have yet explored the quantification of an entire family of protein isoforms, nor the use of such intricate peptide selection techniques, such as

subtraction based quantification, thus this work significantly advancing the field of absolute protein quantification.

Furthermore, when assessing the four confirmed PDE4B protein isoforms, and the one predicted PDE4B protein isoform, each of which exhibited a homologous C-terminal and was differentiated based upon a unique N-terminal, it was necessary to explore the use of several proteases, so as to maximise the number of candidate peptides which could be generated from the limited amounts of unique sequence. In addition, where a suitable AQUA peptide was not identified, subtraction based quantification was again employed. As was previously mentioned, to the author's knowledge the use of subtraction based quantification has not yet been explored in an AQUA based quantification. Furthermore, this specific set of protein isoforms provided the author with an opportunity to develop an AQUA based quantitative technique which should enable the confirmation of a predicted protein, a process which is more commonly based upon either immunoblotting (Shepherd et al., 2003), or LC-MS (through analysing a cell lysate with the aim of identifying each of the proteins contained within) (Lamontagne et al., 2010). Indeed while the MS sensitivity achieved during this project did not allow for the exploration of endogenous protein expression levels, this technique should enable others to both quantify endogenous PDE4B expression, but also to confirm the existence of PDE4B4.

Indeed, utilising an AQUA based approach to confirm a predicted protein has several advantages over both immunoblotting and LC-MS. For example, if an antibody is to be raised against a predicted protein, a biomolecule which shares a certain amount of sequence homology must serve as an antigen (such as a protein isoform generated from the same pre-mRNA), restricting the use of this technique. Furthermore, any proteins which do share a degree of homology must have a significantly different MW to that of the predicted protein, if a differentiation is to be achieved. Finally, LC-MS based proteomic analyses are most commonly performed through the use of bottom-up based proteomics (Zhang et al., 2010), a technique through which only those peptides which are

present in each of the protein isoforms will be detected, rendering highly homologous proteins indistinguishable (Wu et al., 2009).

Finally, when the three confirmed SERCA2 protein isoforms, and one predicted SERCA2 protein isoform were explored, each of which exhibited a homologous 993 amino acid N-terminal and was differentiated based upon a region of unique C-terminal sequence (ranging from 4 to 54 amino acids in length), it was again necessary to explore the selection of unique C-terminal peptides, the use of several non-ideal proteases, and indeed the selection of longer than necessary AQUA peptides which require a proteolytic digestion step to yield the final AQUA product. To the author's knowledge no data has yet been published which explores the use of AQUA peptides which require a proteolytic digest step, however, this technique is somewhat similar to both PSAQ and QconCAT.

While QconCAT utilises an artificial protein, constructed of several Qpeptides, and thus also requires a proteolytic digest step, this technique would be non-ideal for use on C-terminal peptides as C-terminal peptides require only a single digest step, while each Qpeptide would require two, reducing the accuracy of this technique. In the case of PSAQ, while the digest efficiency of both the target and labelled proteins would be identical, the *in vitro* synthesis and subsequent quantification of the labelled protein would require considerable effort so as to enable the quantification of a single target peptide.

As such, while the quantitative techniques explored during this study have been based upon existing technologies, the further development of these methods has enabled their application in a completely different field of research, the exploration and quantitation of protein isoforms. Indeed while previously the selection of proteotypic peptides has been the cornerstone of a successful AQUA quantitation, this project has shown that both non-ideal peptides and non-ideal proteases may be utilised so as to enable a quantitative analysis. Finally, while the overall aim of this project, the absolute quantification of endogenous protein isoforms, was not achieved, the factor limiting this achievement was the MS instrumentation available for use.

7.1 Significance of this Research

Increasingly complex sets of proteins have been quantified through the use of synthetic, isotopically-labelled, internal standards since the first description of AQUA in 2003 (Gerber et al., 2003a). Initially this form of quantitative analysis was restricted to the quantification of single protein species (Gerber et al., 2007, Gerber et al., 2003a). Two such low-abundance proteins were quantified in an unspecified whole yeast lysate (Sir2 and Sir4), for which protein expression levels of 1,750 molecules per cell and 1,150 molecules per cell were identified, respectively (Gerber et al., 2003a).

More recently QconCAT was applied to the absolute quantitation of 27 proteins involved in the glycolytic pathway of *Saccharomyces cerevisiae* (Carroll et al., 2011). From this study, protein expression levels of between 14,000 and 10 million molecules per cell were identified (Carroll et al., 2011). Furthermore, work is underway to quantify at least 4,000 proteins from *Saccharomyces cerevisiae*, again based upon the use of QconCAT (Brownridge et al., 2011), suggesting the field of quantitative proteomics may still be in its infancy.

Fewer sets of protein isoforms have thus far been quantified however, and to the author's knowledge, the absolute quantitation of a complete family of protein isoforms has yet to be explored. Those families of protein isoforms which have been quantified include: (i) Three of five known sucrose synthase isoforms from the root nodules of *Medicago truncatula*, for which concentrations of between 0.3 and 160 fmol/ μ g of protein were identified (Wienkoop et al., 2008), and (ii) Several members of the Cytochrome P450 superfamily of protein isoforms in human liver microsomes, from which a LLOQ of approximately 20 fmol was achieved (a value which is defined as being three times the LoD (Addona et al., 2009)) (Wang et al., 2008b).

During this study three complete sets of protein isoforms have been explored, A-Raf, PDE4B and SERCA2. However, when a HEK293 cell lysate was probed to assess its endogenous expression of A-Raf, none of the four A-Raf isoforms was identified. The failure of this experiment can be traced to two main problems.

Firstly, the majority of the work completed during this project was conducted on an API 2000 MS, an instrument with a LoD for Reserpine of 50 pg, relatively insensitive in comparison to most modern triple quadrupole based mass spectrometers. Secondly, when a limited amount of instrument time became available on a QTrap 5500, the sub-optimal instrument parameters set in the A-Raf MRM acquisition method failed to fully utilise the sensitivity offered by this device, which has a LoD for Reserpine of just 50 fg.

To better explain, when the API 2000 A-Raf MRM based acquisition method was imported to an optimised QTrap 4000 MS, the LoD for AQUA peptide GLNQDCCVVYR fell from 5 pmol to just 15 fmol, comparable to the femtomolar levels of protein expression detected during other published quantitative analyses (Wang et al., 2008b, Wienkoop et al., 2008). Based upon this increase in sensitivity, if the QTrap 5500 were fully optimised, the LoD for GLNQDCCVVYR should fall to just 5 fmol, increasing the likelihood of endogenous A-Raf being detected. Indeed AB SCIEX has just recently released a new QTrap instrument, the QTrap 6500, which offers a 10 x increase in sensitivity over the QTrap 5500. This would suggest that a LoD as low as 500 amol could be achieved for A-Raf limiting peptide GLNQDCCVVYR. Indeed the only factor limiting the quantification of endogenous A-Raf expression is the sensitivity offered by the MS instrument utilised, which if not detectable on either the QTrap 5500 or QTrap 6500, should be detectable upon the release of new instrumentation in the future.

Despite the poor LoDs achieved during this study, this AQUA based quantitative technique still offers researchers many advantages over other methods of isoform quantitation, including antibody based techniques such as western blotting, gene expression based techniques such as qPCR, and general LC-MS based quantitative techniques. Regarding quantitation based on the use of an antibody, each targeted protein isoform must present with significantly different MW's, if each isoform is to be differentiated. Furthermore, if mRNA expression is to be used as an indicator of protein concentration, it should be noted that not all mRNAs are translated into functional proteins, and that those which are may

be translated at different rates. Finally, if a non-isoform specific LC-MS based quantitative method is used, only those peptides which are present in multiple protein isoforms will be detected, rendering highly homologous proteins indistinguishable (Wu et al., 2009).

The description of this AQUA based technique for the quantitation of protein isoforms thus offers researchers a new tool to both confirm the expression of predicted proteins, and to quantify known protein isoforms.

7.2 Further Research

While the mass spectrometers available for use during the development of this quantitative method have hindered progress, it is the author's opinion that several other aspects of this project could be improved upon.

Firstly, when Cys containing AQUA peptides GLNQDCCVVYR, VPTVCVDMSTNRQQ, DYFECSLSK and NSPCFFR were ordered, each was intentionally selected with a reduced sulfhydryl group, so as to ensure both the synthetic peptide and the target protein were alkylated with an equal efficiency. Upon reviewing the efficiency of the IAA based alkylation however, the reaction appeared complete. Instead, through choosing to reduce and alkylate each of the Cys containing AQUA peptide separately, considerable sample losses were incurred, most likely as a result of non specific peptide binding (Speicher et al., 2000). If each Cys containing AQUA peptide was to be re-ordered pre-alkylated, each should enable a more reliable quantitation.

Secondly, when an AQUA based protein quantification is conducted on a complex protein sample such as a cell lysate, an alternative method of protein purification may be more appropriate. While 1D-SDS-PAGE enables the number of proteins permitted to the proteolytic digest step to be reduced considerably, it relies upon the excision of a single gel band for each protein isoform, which are then combined and digested in a single tube, a process which is prone to both inter-gel variability and human error. Furthermore, only a relatively small

number of samples can be run on a single 1D-SDS-PAGE gel, where a 4-12% Bis-Tris gel contains 10 wells, in which only five samples were run, leaving a blank well between each sample so as to prevent mixing. A suitable sample cleanup technique may instead be solid-phase extraction, a method which enables sample multiplexing, orthogonal sample purification and faster sample cleanup and elution (Nissum et al., 2004).

Thirdly, when complex samples such as mixed isoform spiked digests and HEK293 cell lysates were analysed, additional MRM peaks were detected which often interfered with the integration of either the IS or analyte peaks. These peaks most likely resulted from the detection of proteolytic peptides with m/z 's similar to those of targeted peptides, where the Q1 and Q3 MS transmission windows used during this study were set to permit ions within 1.2 ± 0.2 amu of the target m/z to reach the collision cell/detector. A method for reducing the number of interfering peaks may be to modify the Q1 and Q3 transmission window parameters, so as to permit only those ions within 0.7 ± 0.1 amu of the target m/z to reach the collision cell/detector. Alternatively SISCAPA based peptide purification could also be applied to reduce the complexity of the lysate, post-proteolytic digestion. Through reducing the complexity of the peptide sample, fewer undesirable MRM peaks should be detected, while lowering the background MS noise detected.

Aside from the improvements detailed above, the overall aims behind the development of this quantitative method included (i) The identification of which protein isoforms from a particular protein family are expressed within a given tissue. (ii) The accurate quantification each isoform detected, and (iii) The detection of changes in the isoform expression ratio under different physiological conditions. These aims thus set the scene for any future work which may be conducted upon the completion of this project.

7.3 Final Conclusion

While comparative and quantitative proteomic technologies have not yet progressed to the same extent as genomic and transcriptomic technologies in their ability to access differential gene expression (Roulhac et al., 2011), new mass spectrometers and analytical quantitative techniques are being released regularly, offering increased sensitivity, dynamic range, resolution and specificity to the field of proteomics.

Indeed proteomics only progressed beyond the limitations of 2DGE in the early 1990s with the description of soft biomolecule ionisation, enabling proteomic analyses to be performed based upon the use of mass spectrometry (Falk et al., 2006). Even then, proteomic research has required the development of a whole new range of complex multi-stage mass spectrometers, so as to enable peptide fragmentation, peptide/protein sequencing and PTM identification (Domon and Aebersold, 2006).

In regard to quantitative proteomics, the first MS based quantitative technique, ^{15}N labelling, was described as recently as 1999 (Oda et al., 1999), while the first stable isotopically labelled internal standard based absolute quantitative technique was described less than a decade ago (Gerber et al., 2003a).

Building upon these successes, this project has explored the use of stable isotopically labelled internal standard peptides so as to enable the quantification of complete families of protein isoforms, and in doing so has investigated the use of novel non-ideal peptide selection criteria.

8 References

- ABRAHAM, D., PODAR, K., PACHER, M., KUBICEK, M., WELZEL, N., HEMMINGS, B. A., DILWORTH, S. M., MISCHAK, H., KOLCH, W. & BACCARINI, M. 2000. Raf-1-associated Protein Phosphatase 2A as a Positive Regulator of Kinase Activation. *Journal of Biological Chemistry*, 275, 22300-22304.
- ADDONA, T. A., ABBATIello, S. E., SCHILLING, B., SKATES, S. J., MANI, D. R., BUNK, D. M., SPIEGELMAN, C. H., ZIMMERMAN, L. J., HAM, A.-J. L., KESHISHIAN, H., HALL, S. C., ALLEN, S., BLACKMAN, R. K., BORCHERS, C. H., BUCK, C., CARDASIS, H. L., CUSACK, M. P., DODDER, N. G., GIBSON, B. W., HELD, J. M., HILTKE, T., JACKSON, A., JOHANSEN, E. B., KINSINGER, C. R., LI, J., MESRI, M., NEUBERT, T. A., NILES, R. K., PULSIPHER, T. C., RANSOHOFF, D., RODRIGUEZ, H., RUDNICK, P. A., SMITH, D., TABB, D. L., TEGELER, T. J., VARIYATH, A. M., VEGA-MONTOTO, L. J., WAHLANDER, A., WALDEMARSON, S., WANG, M., WHITEAKER, J. R., ZHAO, L., ANDERSON, N. L., FISHER, S. J., LIEBLER, D. C., PAULOVICH, A. G., REGNIER, F. E., TEMPST, P. & CARR, S. A. 2009. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat Biotech*, 27, 633-641.
- AHLSTRÖM, M., PEKKINEN, M., HUTTUNEN, M. & LAMBERG-ALLARDT, C. 2005. Dexamethasone down-regulates cAMP-phosphodiesterase in human osteosarcoma cells. *Biochemical Pharmacology*, 69, 267-275.
- ALBERS, R. W. 1967. Biochemical Aspects of Active Transport. *Annual Review of Biochemistry*, 36, 727-756.
- ANDERSON, L. & SEILHAMER, J. 1997. A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis*, 18, 533-7.
- ANDERSON, N. L. & ANDERSON, N. G. 2002. The Human Plasma Proteome. *Molecular & Cellular Proteomics*, 1, 845-867.
- ANDERSON, N. L., ANDERSON, N. G., HAINES, L. R., HARDIE, D. B., OLAFSON, R. W. & PEARSON, T. W. 2004a. Mass Spectrometric Quantitation of Peptides and Proteins Using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *Journal of Proteome Research*, 3, 235-244.
- ANDERSON, N. L., POLANSKI, M., PIEPER, R., GATLIN, T., TIRUMALAI, R. S., CONRADS, T. P., VEENSTRA, T. D., ADKINS, J. N., POUNDS, J. G., FAGAN, R. & LOBLEY, A. 2004b. The Human Plasma Proteome. *Molecular & Cellular Proteomics*, 3, 311-326.
- ANDREADIS, A., GALLEGO, M. E. & NADAL-GINARD, B. 1987. Generation of protein isoform diversity by alternative splicing: mechanistic and biological implications. *Annu Rev Cell Biol*, 3, 207-42.
- ARAUJO, P. P. C., MARCELLO, M. A., TINCANI, A. J., GUILHEN, A. C. T., MORARI, E. C. & WARD, L. S. 2012. mRNA BRAF expression helps to identify papillary thyroid carcinomas in thyroid nodules independently of the presence of BRAFV600E mutation. *Pathology - Research and Practice*, 208, 489-492.
- ASHKI, N., HAYES, K. C. & BAO, F. 2008. The peroxynitrite donor 3-morpholinopyridone induces reversible changes in electrophysiological properties of neurons of the guinea-pig spinal cord. *Neuroscience*, 156, 107-117.
- BAAK, J. P. A., JANSSEN, E. A. M., SOREIDE, K. & HEIKKILÄ, R. 2005. Genomics and proteomics—the way forward. *Annals of Oncology*, 16, ii30-ii44.
- BABU, G. J., WHEELER, D., ALZATE, O. & PERIASAMY, M. 2004. Solubilization of membrane proteins for two-dimensional gel electrophoresis: identification of sarcoplasmic reticulum membrane proteins. *Analytical Biochemistry*, 325, 121-125.
- BACCARINI, M. 2005. Second nature: Biological functions of the Raf-1 "kinase". *FEBS Letters*, 579, 3271-3277.

- BAKHTIAR, R. & TSE, F. L. S. 2000. Biological mass spectrometry: a primer. *Mutagenesis*, 15, 415-430.
- BALJULS, A., MUELLER, T., DREXLER, H. C. A., HEKMAN, M. & RAPP, U. R. 2007. Unique N-region Determines Low Basal Activity and Limited Inducibility of A-RAF Kinase. *Journal of Biological Chemistry*, 282, 26575-26590.
- BALJULS, A., SCHMITZ, W., MUELLER, T., ZAHEDI, R. P., SICKMANN, A., HEKMAN, M. & RAPP, U. R. 2008. Positive Regulation of A-RAF by Phosphorylation of Isoform-specific Hinge Segment and Identification of Novel Phosphorylation Sites. *Journal of Biological Chemistry*, 283, 27239-27254.
- BANTSCHIEFF, M., SCHIRLE, M., SWEETMAN, G., RICK, J. & KUSTER, B. 2007. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.*, 3899, 1017-31.
- BARNIER, J. V., PAPIN, C., EYCHÈNE, A., LECOQ, O. & CALOTHY, G. 1995. The Mouse B-raf Gene Encodes Multiple Protein Isoforms with Tissue-specific Expression. *Journal of Biological Chemistry*, 270, 23381-23389.
- BEAVO, J. A. 1995. Cyclic nucleotide phosphodiesterases: functional implications of multiple isoforms. *Physiological Reviews*, 75, 725-748.
- BEAVO, J. A. & BRUNTON, L. L. 2002. Cyclic nucleotide research [mdash] still expanding after half a century. *Nat Rev Mol Cell Biol*, 3, 710-718.
- BEAVO, J. A., HARDMAN, J. G. & SUTHERLAND, E. W. 1970. Hydrolysis of cyclic guanosine and adenosine 3',5'-monophosphates by rat and bovine tissues. *J.Biol.Chem.*, 245, 5649-5655.
- BENNINGHOVEN, A., JASPERS, D. & SICHTERMANN, W. 1976. Secondary-ion emission of amino acids. *Applied Physics A: Materials Science & Processing*, 11, 35-39.
- BERCHTOLD, M. W., BRINKMEIER, H. & MÜNTENER, M. 2000. Calcium ion in skeletal muscle: its crucial role for muscle function, plasticity, and disease. *Physiological Reviews*, 80, 1215-1265.
- BEYNON, R. J., DOHERTY, M. K., PRATT, J. M. & GASKELL, S. J. 2005. Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat Meth*, 2, 587-589.
- BISLEV, S. L., KUSEBAUCH, U., CODREA, M. C., BEYNON, R. J., HARMAN, V. M., RØNTVED, C. M., AEBERSOLD, R., MORITZ, R. L. & BENDIXEN, E. 2012. Quantotypic Properties of QconCAT Peptides Targeting Bovine Host Response to *Streptococcus uberis*. *Journal of Proteome Research*, 11, 1832-1843.
- BOHN, H. & SCHÖNAFINGER, K. 1989. Oxygen and oxidation promote the release of nitric oxide from sydnonimines. *Journal of cardiovascular pharmacology*, 14 Suppl 11, S6-12.
- BOLGER, G., MICHAELI, T., MARTINS, T., ST JOHN, T., STEINER, B., RODGERS, L., RIGGS, M., WIGLER, M. & FERGUSON, K. 1993. A family of human phosphodiesterases homologous to the dunce learning and memory gene product of *Drosophila melanogaster* are potential targets for antidepressant drugs. *Molecular and Cellular Biology*, 13, 6558-6571.
- BOLGER, G. B. 1994. Molecular biology of the cyclic AMP-specific cyclic nucleotide phosphodiesterases: A diverse family of regulatory enzymes. *Cellular Signalling*, 6, 851-859.
- BOSWELL-SMITH, V., SPINA, D. & PAGE, C. P. 2006. Phosphodiesterase inhibitors. *British Journal of Pharmacology*, 147, S252-S257.
- BRAND, E. 1946. AMINO ACID COMPOSITION OF SIMPLE PROTEINS. *Annals of the New York Academy of Sciences*, 47, 187-228.
- BRANDL, C. J., DELEON, S., MARTIN, D. R. & MACLENNAN, D. H. 1987. ADULT FORMS OF THE CA²⁺ ATPASE OF SARCOPLASMIC-RETICULUM - EXPRESSION IN DEVELOPING SKELETAL-MUSCLE. *Journal of Biological Chemistry*, 262, 3768-3774.

- BRAUN, N. N., REUTIMAN, T. J., LEE, S., FOLSOM, T. D. & FATEMI, S. H. 2007. Expression of phosphodiesterase 4 is altered in the brains of subjects with autism. *NeuroReport*, 18, 1841-1844.
- BRINI, M. & CARAFOLI, E. 2009. Calcium Pumps in Health and Disease. *Physiological Reviews*, 89, 1341-1378.
- BROWN, R. S. & LENNON, J. J. 1995a. Mass Resolution Improvement by Incorporation of Pulsed Ion Extraction in a Matrix-Assisted Laser Desorption/Ionization Linear Time-of-Flight Mass Spectrometer. *Analytical Chemistry*, 67, 1998-2003.
- BROWN, R. S. & LENNON, J. J. 1995b. Sequence-Specific Fragmentation of Matrix-Assisted Laser-Desorbed Protein/Peptide Ions. *Analytical Chemistry*, 67, 3990-3999.
- BROWNRIDGE, P. & BEYNON, R. J. 2011. The importance of the digest: Proteolysis and absolute quantification in proteomics. *Methods*, 54, 351-360.
- BROWNRIDGE, P., HOLMAN, S. W., GASKELL, S. J., GRANT, C. M., HARMAN, V. M., HUBBARD, S. J., LANTHALER, K., LAWLESS, C., O'CUALAIN, R., SIMS, P., WATKINS, R. & BEYNON, R. J. 2011. Global absolute quantification of a proteome: Challenges in the deployment of a QconCAT strategy. *Proteomics*, 11, 2957-2970.
- BRUN, V., DUPUIS, A., ADRAIT, A., MARCELLIN, M., THOMAS, D., COURT, M., VANDENESCH, F. & GARIN, J. 2007. Isotope-labeled Protein Standards. *Molecular & Cellular Proteomics*, 6, 2139-2149.
- BRUN, V., MASSELO, C., GARIN, J. & DUPUIS, A. 2009. Isotope dilution strategies for absolute quantitative proteomics. *Journal of Proteomics*, 72, 740-749.
- BUTCHER, R. W. & SUTHERLAND, E. W. 1962. Adenosine 3',5'-phosphate in biological materials. I. Purification and properties of cyclic 3',5'-nucleotide phosphodiesterase and use of this enzyme to characterize adenosine 3',5'-phosphate in human urine. *The Journal of biological chemistry*, 237, 1244-50.
- BYERS, H. L., CAMPBELL, J., VAN ULSEN, P., TOMMASSEN, J., WARD, M. A., SCHULZ-KNAPPE, P., PRINZ, T. & KUHN, K. 2009. Candidate verification of iron-regulated *Neisseria meningitidis* proteins using isotopic versions of tandem mass tags (TMT) and single reaction monitoring. *Journal of Proteomics*, 73, 231-239.
- CARROLL, K. M., SIMPSON, D. M., EYERS, C. E., KNIGHT, C. G., BROWNRIDGE, P., DUNN, W. B., WINDER, C. L., LANTHALER, K., PIR, P., MALYS, N., KELL, D. B., OLIVER, S. G., GASKELL, S. J. & BEYNON, R. J. 2011. Absolute Quantification of the Glycolytic Pathway in Yeast. *Molecular & Cellular Proteomics*, 10.
- CHAIT, B. T. & KENT, S. B. H. 1992. Weighing naked proteins: practical, high accuracy mass measurement of peptides and proteins. *Science*, 257, 1885-94.
- CHAPMAN, T. 2005. Protein purification: pure but not simple. *Nature*, 424, 795-8.
- CHARNEY, D. S. & DEUTCH, A. 1996. A functional neuroanatomy of anxiety and fear: implications for the pathophysiology and treatment of anxiety disorders. *Crit Rev Neurobiol*, 10, 419-446.
- CHERRY, J. A. & DAVIS, R. L. 1999. Cyclic AMP phosphodiesterases are localized in regions of the mouse brain associated with reinforcement, movement, and affect. *J Comp Neurol*, 407, 287-301.
- CHEUNG, Y.-F., KAN, Z., GARRETT-ENGELE, P., GALL, I., MURDOCH, H., BAILLIE, G. S., CAMARGO, L. M., JOHNSON, J. M., HOUSLAY, M. D. & CASTLE, J. C. 2007. PDE4B5, a Novel, Super-Short, Brain-Specific cAMP Phosphodiesterase-4 Variant Whose Isoform-Specifying N-Terminal Region Is Identical to That of cAMP Phosphodiesterase-4D6 (PDE4D6). *Journal of Pharmacology and Experimental Therapeutics*, 322, 600-609.
- CIAMPI, R. & NIKIFOROV, Y. E. 2005. Alterations of the BRAF gene in thyroid tumors. *Endocr. Pathol.*, 16, 163-172.

- CLAUSER, K. R., BAKER, P. & BURLINGAME, A. L. 1999. Role of Accurate Mass Measurement (± 10 ppm) in Protein Identification Strategies Employing MS or MS/MS and Database Searching. *Analytical Chemistry*, 71, 2871-2882.
- COHEN, G. B., REN, R. & BALTIMORE, D. 1995. Modular binding domains in signal transduction proteins. *Cell*, 80, 237-248.
- COHEN, S. L. & CHAIT, B. T. 1997. Mass Spectrometry of Whole Proteins Eluted from Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis Gels. *Analytical Biochemistry*, 247, 257-267.
- COLLIER, T. S., HAWKRIDGE, A. M., GEORGIANNA, D. R., PAYNE, G. A. & MUDDIMAN, D. C. 2008. Top-Down Identification and Quantification of Stable Isotope Labeled Proteins from *Aspergillus flavus* Using Online Nano-Flow Reversed-Phase Liquid Chromatography Coupled to a LTQ-FTICR Mass Spectrometer. *Analytical Chemistry*, 80, 4994-5001.
- COMISAROW, M. B. & MARSHALL, A. G. 1974. Fourier transform ion cyclotron resonance spectroscopy. *Chemical Physics Letters*, 25, 282-283.
- CORNETT, D. S., FRAPPIER, S. L. & CAPRIOLI, R. M. 2008. MALDI-FTICR Imaging Mass Spectrometry of Drugs and Metabolites in Tissue. *Analytical Chemistry*, 80, 5648-5653.
- CORTHALS, G. L., WASINGER, V. C., HOCHSTRASSER, D. F. & SANCHEZ, J.-C. 2000. The dynamic range of protein expression: A challenge for proteomic research. *Electrophoresis*, 21, 1104-1115.
- CRAIG, R., CORTENS, J. P. & BEAVIS, R. C. 2005. The use of proteotypic peptide libraries for protein identification. *Rapid Communications in Mass Spectrometry*, 19, 1844-1850.
- CRICK, F. H. C. 1968. The origin of the genetic code. *Journal of Molecular Biology*, 38, 367-379.
- CUI, W., ROHRS, H. W. & GROSS, M. L. 2011. Top-down mass spectrometry: Recent developments, applications and perspectives. *Analyst*, 136, 3854-3864.
- DALLY, S., BREDOUX, R., CORVAZIER, E., ANDERSEN, J. P., CLAUSEN, J. D., DODE, L., FANCHAOUY, M., GELEBART, P., MONCEAU, V., DEL MONTE, F., GWATHMEY, J. K., HAJJAR, R., CHAABANE, C., BOBE, R., RAIES, A. & ENOUF, J. 2006. Ca²⁺-ATPases in non-failing and failing heart: evidence for a novel cardiac sarco/endoplasmic reticulum Ca²⁺-ATPase 2 isoform (SERCA2c) *Biochem. J.*, 394, 249-258.
- DALLY, S., CORVAZIER, E., BREDOUX, R., BOBE, R. & ENOUF, J. 2010. Multiple and diverse coexpression, location, and regulation of additional SERCA2 and SERCA3 isoforms in nonfailing and failing human heart. *Journal of Molecular and Cellular Cardiology*, 48, 633-644.
- DAUB, M., JOCKEL, J., QUACK, T., WEBER, C. K., SCHMITZ, F., RAPP, U. R., WITTINGHOFER, A. & BLOCK, C. 1998. The RafC1 Cysteine-Rich Domain Contains Multiple Distinct Regulatory Epitopes Which Control Ras-Dependent Raf Activation. *Mol. Cell. Biol.*, 18, 6698-6710.
- DAVID M, E. 1999. Cyclic nucleotide phosphodiesterase (PDE) inhibitors and immunomodulation. *Biochemical Pharmacology*, 57, 965-973.
- DAVIES, H., BIGNELL, G. R., COX, C., STEPHENS, P., EDKINS, S., CLEGG, S., TEAGUE, J., WOFFENDIN, H., GARNETT, M. J., BOTTOMLEY, W., DAVIS, N., DICKS, E., EWING, R., FLOYD, Y., GRAY, K., HALL, S., HAWES, R., HUGHES, J., KOSMIDOU, V., MENZIES, A., MOULD, C., PARKER, A., STEVENS, C., WATT, S., HOOPER, S., WILSON, R., JAYATILAKE, H., GUSTERSON, B. A., COOPER, C., SHIPLEY, J., HARGRAVE, D., PRITCHARD-JONES, K., MAITLAND, N., CHENEVIX-TRENCH, G., RIGGINS, G. J., BIGNER, D. D., PALMIERI, G., COSSU, A., FLANAGAN, A., NICHOLSON, A., HO, J. W. C., LEUNG, S. Y., YUEN, S. T., WEBER, B. L., SEIGLER, H. F., DARROW, T. L., PATERSON, H., MARAIS, R., MARSHALL, C. J., WOOSTER, R., STRATTON, M. R. & FUTREAL, P. A. 2002. Mutations of the BRAF gene in human cancer. *Nature*, 417, 949-954.
- DAYON, L., HAINARD, A., LICKER, V., TURCK, N., KUHN, K., HOCHSTRASSER, D. F., BURKHARD, P. R. & SANCHEZ, J.-C. 2008. Relative Quantification of Proteins in Human Cerebrospinal Fluids by MS/MS Using 6-Plex Isobaric Tags. *Analytical Chemistry*, 80, 2921-2931.

- DELABASTIE, D., LEVITSKY, D., RAPPAPORT, L., MERCADIER, J. J., MAROTTE, F., WISNEWSKY, C., BROVKOVICH, V., SCHWARTZ, K. & LOMPRES, A. M. 1990. FUNCTION OF THE SARCOPLASMIC-RETICULUM AND EXPRESSION OF ITS CA²⁺-ATPASE GENE IN PRESSURE OVERLOAD-INDUCED CARDIAC-HYPERTROPHY IN THE RAT. *Circulation Research*, 66, 554-564.
- DEVÉS, R. & BRODIE, A. F. 1981. Active transport of Ca²⁺ in bacteria: bioenergetics and function. *Molecular and cellular biochemistry*, 36, 65-84.
- DEVIENCE, F. M. & ROUSTAN, J.-C. 1982. 'Fast atom bombardment'—A rediscovered method for mass spectrometry. *Organic Mass Spectrometry*, 17, 173-181.
- DI NICOLANTONIO, F., MARTINI, M., MOLINARI, F., SARTORE-BIANCHI, A., ARENA, S., SALETTI, P., DE DOSSO, S., MAZZUCHELLI, L., FRATTINI, M., SIENA, S. & BARDELLI, A. 2008. Wild-Type BRAF Is Required for Response to Panitumumab or Cetuximab in Metastatic Colorectal Cancer. *Journal of Clinical Oncology*, 26, 5705-5712.
- DIEFFENBACH, C. W., LOWE, T. M. & DVEKSLER, G. S. 1993. General concepts for PCR primer design. *Genome Res*, 3, S30-S37.
- DILLON, R., NILSSON, C. L., SHI, S. D. H., LEE, N. V., KRSTINS, B. & GREIG, M. J. 2011. Discovery of a Novel B-Raf Fusion Protein Related to c-Met Drug Resistance. *Journal of Proteome Research*, 10, 5084-5094.
- DODDS, E. D., GERMAN, J. B. & LEBRILLA, C. B. 2007. Enabling MALDI-FTICR-MS/MS for High-Performance Proteomics through Combination of Infrared and Collisional Activation. *Analytical Chemistry*, 79, 9547-9556.
- DOHLMAN, H. G., THORNER, J., CARON, M. G. & LEFKOWITZ, R. J. 1991. Model Systems for the Study of Seven-Transmembrane-Segment Receptors. *Annu. Rev. Biochem.*, 60, 653-688.
- DOLE, M., MACK, L. L., HINES, R. L., MOBLEY, R. C., FERGUSON, L. D. & ALICE, M. B. 1968. Molecular Beams of Macroions. *The Journal of Chemical Physics*, 49, 2240-2249.
- DOMON, B. & AEBERSOLD, R. 2006. Mass Spectrometry and Protein Analysis. *Science*, 312, 212-217.
- DREWS, J. 2000. Drug discovery: a historical perspective. *Science*, 287, 1960-4.
- DUGGAN, D. J., BITTNER, M., CHEN, Y., MELTZER, P. & TRENT, J. M. 1999. Expression profiling using cDNA microarrays. *Nat. Genet.*, 21, 10-4.
- DUNN, M. J. 2007. PROTEOMICS: The Birth of the PROTEOMICS Family of Journals. *Proteomics*, 7, 1-3.
- DUROCHER, Y., PERRET, S. & KAMEN, A. 2002. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Research*, 30, e9.
- DWIVEDI, Y. 2010. S24-02 - Phosphodiesterase 4 (PDE4) variants: role in depression and treatment. *European Psychiatry*, 25, Supplement 1, 30.
- EDMAN, P. 1949. Method for determination of the amino acid sequence in peptides. *Arch Biochem.*, 22, 475.
- EDWARDS, A. S. & SCOTT, J. D. 2000. A-kinase anchoring proteins: protein kinase A and beyond. *Current Opinion in Cell Biology*, 12, 217-221.
- EL-ANEED, A., COHEN, A. & BANOUB, J. 2009. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Applied Spectroscopy Reviews*, 44, 210-230.
- ENGELS, P., FICHTEL, K. & LÜBBERT, H. 1994. Expression and regulation of human and rat phosphodiesterase type IV isogenes. *FEBS letters*, 350, 291-295.
- ENGELS, P., SULLIVAN, M., MÜLLER, T. & LÜBBERT, H. 1995. Molecular cloning and functional expression in yeast of a human cAMP-specific phosphodiesterase subtype (PDE IV-C). *FEBS letters*, 358, 305-310.
- FALK, R., RAMSTROM, M., STAHL, S. & HOBER, S. 2006. Approaches for systematic proteome exploration. *Biomolecular Engineering*, 24, 13.

- FAROOQUI, S. M., ZHANG, K., MAKHAY, M., JACKSON, K., FAROOQUI, S. Q., CHERRY, J. A. & O'DONNELL, J. M. 2000. Noradrenergic lesions differentially alter the expression of two subtypes of low Km cAMP-sensitive phosphodiesterase type 4 (PDE4A and PDE4B) in rat brain. *Brain Research*, 867, 52-61.
- FASSETT, J. D. 1995. Isotopic and nuclear analytical techniques in biological systems: A critical study-X. Elemental isotope dilution analysis with radioactive and stable isotopes (Technical Report). *Pure Appl. Chem.*, 67, 1943-1949.
- FATEMI, S. H., KING, D. P., REUTIMAN, T. J., FOLSOM, T. D., LAURENCE, J. A., LEE, S., FAN, Y.-T., PACIGA, S. A., CONTI, M. & MENNITI, F. S. 2008. PDE4B polymorphisms and decreased PDE4B expression are associated with schizophrenia. *Schizophrenia Research*, 101, 36-49.
- FENN, J. B. 2002. Electrospray ionization mass spectrometry: How it all began. *Journal of biomolecular techniques : JBT*, 13, 101-118.
- FENN, J. B., MANN, M., MENG, C. K., WONG, S. F. & WHITEHOUSE, C. M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246, 64-71.
- FENSELAU, C. 2007. A review of quantitative methods for proteomic studies. *J Chromatogr B Analyt Technol Biomed Life Sci.*, 855, 14-20.
- FERGUSON, J. T., WENGER, C. D., METCALF, W. W. & KELLEHER, N. L. 2009. Top-Down Proteomics Reveals Novel Protein Forms Expressed in *Methanosarcina acetivorans*. *Journal of the American Society for Mass Spectrometry*, 20, 1743-1750.
- FOURNIER, M. L., GILMORE, J. M., MARTIN-BROWN, S. A. & WASHBURN, M. P. 2007. Multidimensional Separations-Based Shotgun Proteomics. *Chemical Reviews*, 107, 3654-3686.
- FRANCIS, S. H., BLOUNT, M. A. & CORBIN, J. D. 2011. Mammalian Cyclic Nucleotide Phosphodiesterases: Molecular Mechanisms and Physiological Functions. *Physiological Reviews*, 91, 651-690.
- FRANCIS, S. H., TURKO, I. V. & CORBIN, J. D. 2000. Cyclic nucleotide phosphodiesterases: Relating structure and function. *Progress in Nucleic Acid Research and Molecular Biology*. Academic Press.
- FUJII, N., SASAKI, T., FUNAKOSHI, S., IRIE, H. & YAJIMA, H. 1978. Studies on Peptides. LXXIV. Convenient Procedure for the Preparation of Methionine Sulphoxide Derivatives. *Chem. Pharm. Bull.*, 26, 650-653.
- GALEVA, N. & ALTERMANN, M. 2002. Comparison of one-dimensional and two-dimensional gel electrophoresis as a separation tool for proteomic analysis of rat liver microsomes: cytochromes P450 and other membrane proteins. *Proteomics*, 2, 713-22.
- GALVANI, M., HAMDAN, M., HERBERT, B. & RIGHETTI, P. G. 2001. Alkylation kinetics of proteins in preparation for two-dimensional maps: A matrix assisted laser desorption/ionization-mass spectrometry investigation. *Electrophoresis*, 22, 2058-2065.
- GARDNER, A. M., VAILLANCOURT, R. R., LANGE-CARTER, C. A. & JOHNSON, G. L. 1994. MEK-1 phosphorylation by MEK kinase, Raf, and mitogen-activated protein kinase: analysis of phosphopeptides and regulation of activity. *Molecular Biology of the Cell*, 5, 193-201.
- GARY M, K. 1988. The adenylate cyclase-cAMP-protein kinase A pathway and regulation of the immune response. *Immunology Today*, 9, 222-229.
- GERBER, S. A., KETTENBACH, A. N., RUSH, J. & GYGI, S. P. 2007. The Absolute Quantification Strategy.
- GERBER, S. A., RUSH, J., STEMMAN, O., KIRSCHNER, M. W. & GYGI, S. P. 2003a. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences*, 100, 6940-6945.
- GERBER, S. A., RUSH, J., STEMMAN, O., KIRSCHNER, M. W. & GYGI, S. P. 2003b. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Nat. Acad. Sci. U.S.A.*, 100, 6940-5.

- GHOSH, S., STRUM, J. C., SCIORRA, V. A., DANIEL, L. & BELL, R. M. 1996. Raf-1 Kinase Possesses Distinct Binding Domains for Phosphatidylserine and Phosphatidic Acid. *Journal of Biological Chemistry*, 271, 8472-8480.
- GIDDINGS, J. C. 1984. Two-dimensional separations: concept and promise. *Analytical Chemistry*, 56, 1258A-1270A.
- GILMAN, A. G. 1987. G Proteins: Transducers of Receptor-Generated Signals. *Annual Review of Biochemistry*, 56, 615-649.
- GODOVAC-ZIMMERMANN, J., KLEINER, O., BROWN, L. R. & DRUKIER, A. K. 2005. Perspectives in splicing up proteomics with splicing. *Proteomics*, 5, 699-709.
- GORDON, J. A. & JENCKS, W. P. 1963. The relationship of structure to the effectiveness of denaturing agents for proteins. *Biochemistry*, 2, 47-57.
- GOVERMAN, J. M. & PIERCE, J. G. 1981. Differential Effects of Alkylation of Methionine Residues on the Activities of Pituitary Thyrotropin and Lutropin*. *The Journal of Biological Chemistry*, 256, 9431-9435.
- GOW, A. J., MCCLELLAND, M., GARNER, S. E., MALCOLM, S. & ISCHIROPOULOS, H. 1997. The Determination of Nitrotyrosine Residues in Proteins.
- GRAEME B, B. 1994. Molecular biology of the cyclic AMP-specific cyclic nucleotide phosphodiesterases: A diverse family of regulatory enzymes. *Cellular Signalling*, 6, 851-859.
- GRANDIS, J. R. & SOK, J. C. 2004. Signaling through the epidermal growth factor receptor during the development of malignancy. *Pharmacology & Therapeutics*, 102, 37-46.
- GRIFFITHS, J. 2008. A Brief History of Mass Spectrometry. *Analytical Chemistry*, 80, 5678-5683.
- GUERRERA, I. C. & KLEINER, O. 2005. Application of mass spectrometry in proteomics. *Biosci Rep.*, 25, 71-93.
- GUNNING, P. W. 2001. Protein Isoforms and Isozymes. *eLS*. John Wiley & Sons, Ltd.
- GYGI, S. P., CORTHALS, G. L., ZHANG, Y., ROCHON, Y. & AEBERSOLD, R. 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proceedings of the National Academy of Sciences*, 97, 9390-9395.
- GYGI, S. P., RIST, B., GERBER, S. A., TURECEK, F., GELB, M. H. & AEBERSOLD, R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotech*, 17, 994-999.
- HAGEMANN, C. & RAPP, U. R. 1999. Isozyme-Specific Functions of Raf Kinases. *Experimental Cell Research*, 253, 34-46.
- HAGER, J. W. 2002. A new linear ion trap mass spectrometer. *Rapid Communications in Mass Spectrometry*, 16, 512-526.
- HALLBERG, B., RAYTER, S. I. & DOWNWARD, J. 1994. Interaction of Ras and Raf in intact mammalian cells upon extracellular stimulation. *Journal of Biological Chemistry*, 269, 3913-3916.
- HAVLIŠ, J. & SHEVCHENKO, A. 2004. Absolute Quantification of Proteins in Solutions and in Polyacrylamide Gels by Mass Spectrometry. *Analytical Chemistry*, 76, 3029-3036.
- HAYNES, P. A. & YATES III, J. R. 2000. Proteome Profiling-Pitfalls and Progress. *Yeast*, 1, 81-87.
- HE, C., ZHOU, F., ZUO, Z., CHENG, H. & ZHOU, R. 2009. A Global View of Cancer-Specific Transcript Variants by Subtractive Transcriptome-Wide Analysis. *PLoS ONE*, 4, e4732.
- HE, Q.-Y. & CHIU, J.-F. 2003. Proteomics in biomarker discovery and drug development. *Journal of Cellular Biochemistry*, 89, 868-886.
- HECKMAN, K. L. & PEASE, L. R. 2007. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protocols*, 2, 924-932.
- HENCHMAN, M. & STEEL, C. 1998. Understanding the Quadrupole Mass Filter through Computer Simulation. *Journal of Chemical Education*, 75, 1049.
- HERBERT, B. R., HARRY, J. L., PACKER, N. H., GOOLEY, A. A., PEDERSEN, S. K. & WILLIAMS, K. L. 2001. What place for polyacrylamide in proteomics? *Trends in Biotechnology*, 19, S3-S9.

- HERRMANN, K. A., SOMOGYI, Á., WYSOCKI, V. H., DRAHOS, L. & VÉKEY, K. 2005. Combination of Sustained Off-Resonance Irradiation and On-Resonance Excitation in FT-ICR. *Analytical Chemistry*, 77, 7626-7638.
- HERVEY, STRADER, M. B. & HURST, G. B. 2007. Comparison of Digestion Protocols for Microgram Quantities of Enriched Protein Samples. *Journal of Proteome Research*, 6, 3054-3061.
- HILLENKAMP, F. & KARAS, M. 1990. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. In: JAMES, A. M. (ed.) *Methods in Enzymology*. Academic Press.
- HILLENKAMP, F. & KARAS, M. 2000. Matrix-assisted laser desorption/ionisation, an experience. *International Journal of Mass Spectrometry*, 200, 71-77.
- HINGORANI, S. R., JACOBETZ, M. A., ROBERTSON, G. P., HERLYN, M. & TUVESON, D. A. 2003. Suppression of BRAFV599E in Human Melanoma Abrogates Transformation. *Cancer Research*, 63, 5198-5202.
- HMITOU, I., DRUILLENNEC, S., VALLUET, A., PEYSSONNAUX, C. & EYCHENE, A. 2007. Differential Regulation of B-Raf Isoforms by Phosphorylation and Autoinhibitory Mechanisms. *Mol. Cell. Biol.*, 27, 31-43.
- HO, S. N., HUNT, H. D., HORTON, R. M., PULLEN, J. K. & PEASE, L. R. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, 77, 51-59.
- HOPFGARTNER, G., VAREGIO, E., TSCHÄPPÄT, V., GRIVET, C., BOURGOGNE, E. & LEUTHOLD, L. A. 2004. Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. *Journal of Mass Spectrometry*, 39, 845-855.
- HOUËL, S., ABERNATHY, R., RENGANATHAN, K., MEYER-ARENDET, K., AHN, N. G. & OLD, W. M. 2010. Quantifying the Impact of Chimera MS/MS Spectra on Peptide Identification in Large-Scale Proteomics Studies. *Journal of Proteome Research*, 9, 4152-4160.
- HOUMARD, J. & DRAPEAU, G. R. 1972. Staphylococcal Protease: A Proteolytic Enzyme Specific for Glutamoyl Bonds. *Proceedings of the National Academy of Sciences*, 69, 3506-3509.
- HOUSLAY, M. D. & ADAMS, D. R. 2003. PDE4 cAMP phosphodiesterases: modular enzymes that orchestrate signalling cross-talk, desensitization and compartmentalization. *Biochemical Journal*, 370, 1-18.
- HOUSLAY, M. D., SULLIVAN, M. & BOLGERZ, G. B. 1998. The Multienzyme PDE4 Cyclic Adenosine Monophosphate-Specific Phosphodiesterase Family: Intracellular Targeting, Regulation, and Selective Inhibition by Compounds Exerting Anti-inflammatory and Antidepressant Actions. In: J. THOMAS AUGUST, M. W. A. F. M. & JOSEPH, T. C. (eds.) *Advances in Pharmacology*. Academic Press.
- HU, Q., NOLL, R. J., LI, H., MAKAROV, A., HARDMAN, M. & GRAHAM COOKS, R. 2005. The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry*, 40, 430-443.
- HURST, R., MAFFITT, M., MURRAY, E., KAPPELMAN, J., XU, Q., MENDEZ, J., BUTLER, B., RYAN, A. & BECKLER, G. S. 1996. The TNT® T7 Quick Coupled Transcription/Translation System. *Promega Notes*, 58, 68.
- HUSER, M., LUCKETT, J., CHILOECHES, A., MERCER, K., IWOBI, M., GIBLETT, S., SUN, X.-M., BROWN, J., MARAIS, R. & PRITCHARD, C. 2001. MEK kinase activity is not necessary for Raf-1 function. *EMBO J*, 20, 1940-1951.
- HUSTON, E., LUMB, S., RUSSELL, A., CATTERALL, C., ROSS, A. H., STEELE, M. R., BOLGER, G. B., PERRY, M. J., OWENS, R. J. & HOUSLAY, M. D. 1997. Molecular cloning and transient expression in COS7 cells of a novel human PDE4B cAMP-specific phosphodiesterase, HSPDE4B3. *Biochem. J.*, 328, 549-58.
- IKAWA, S., FUKUI, M., UEYAMA, Y., TAMAOKI, N., YAMAMOTO, T. & TOYOSHIMA, K. 1988. B-raf, a new member of the raf family, is activated by DNA rearrangement. *Mol. Cell. Biol.*, 8, 2651-2654.
- IONA, S., CUOMO, M., BUSHNIK, T., NARO, F., SETTE, C., HESS, M., SHELTON, E. R. & CONTI, M. 1998. Characterization of the rolipram-sensitive, cyclic AMP-specific phosphodiesterases:

- identification and differential expression of immunologically distinct forms in the rat brain. *Mol. Pharmacol.*, 53, 23-32.
- IRUNGU, J., GO, E., ZHANG, Y., DALPATHADO, D., LIAO, H.-X., HAYNES, B. & DESAIRE, H. 2008. Comparison of HPLC/ESI-FTICR MS versus MALDI-TOF/TOF MS for glycopeptide analysis of a highly glycosylated HIV envelope glycoprotein. *Journal of the American Society for Mass Spectrometry*, 19, 1209-1220.
- ISCHIROPOULOS, H. 1998. Biological Tyrosine Nitration: A Pathophysiological Function of Nitric Oxide and Reactive Oxygen Species. *Archives of Biochemistry and Biophysics*, 356, 1-11.
- ISHIHAMA, Y., ODA, Y., TABATA, T., SATO, T., NAGASU, T., RAPPSILBER, J. & MANN, M. 2005. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics*, 4, 1265-72.
- ISHIKAWA, F., TAKAKU, F., NAGAO, M. & SUGIMURA, T. 1987. The complete primary structure of the rat A-raf cDNA coding region: Conservation of the putative regulatory regions present in rat c-raf. *Oncogene Research*, 1, 243-253.
- JAARO-PELED, H., AYHAN, Y., PLETNIKOV, M. V. & SAWA, A. 2010. Review of Pathological Hallmarks of Schizophrenia: Comparison of Genetic Models With Patients and Nongenetic Models. *Schizophrenia Bulletin*, 36, 301-313.
- JANSEN, H. W., LURZ, R., BISTER, K., BONNER, T. I., MARK, G. E. & RAPP, U. R. 1984. Homologous cell-derived oncogenes in avian carcinoma virus MH2 and murine sarcoma virus 3611. *Nature*, 307, 281-4.
- JAQUINOD, M., TRAUCHESSEC, M., HUILLET, C., LOUWAGIE, M., LEBERT, D., PICARD, G., ADRAIT, A., DUPUIS, A., GARIN, J., BRUN, V. & BRULEY, C. 2012. Mass spectrometry-based absolute protein quantification: PSAQ™ strategy makes use of “noncanonical” proteotypic peptides. *Proteomics*, 12, 1217-1221.
- JEFFREY, M. J. M. P. 2012. Tearing the top off 'Top-Down' Proteomics. *Biotechniques*, 53, 75-78.
- JEON, Y. H., HEO, Y. S., KIM, C. M., HYUN, Y. L., LEE, T. G., RO, S. & CHO, J. M. 2005. Phosphodiesterase: overview of protein structures, potential therapeutic applications and recent progress in drug development. *Cellular and Molecular Life Sciences*, 62, 1198-1220.
- JOHNSON, J. M., CASTLE, J. C., GARRETT-ENGELE, P., KAN, Z., LOERCH, P. M., ARMOUR, C. D., SANTOS, R., SCHADT, E. E., STOUGHTON, R. & SHOEMAKER, D. D. 2003. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, 302, 2141-4.
- JONSSON, A. P. 2001. Mass spectrometry for protein and peptide characterization. *Cell. Mol. Life Sci.*, 58, 868-84.
- KÄHLER, A. K., OTNAESS, M. K., WIRGENES, K. V., HANSEN, T., JÖNSSON, E. G., AGARTZ, I. & AL., E. 2010. Association study of PDE4B gene variants in Scandinavian schizophrenia and bipolar disorder multicenter case-control samples. *Am J Med Genet B Neuropsychiatr Genet*, 153b, 86-96.
- KARAS, M., BACHMANN, D. & HILLENKAMP, F. 1985. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry*, 57, 2935-2939.
- KARLIN, S., BUCHER, P., BRENDDEL, V. & ALTSCHUL, S. F. 1991. Statistical methods and insights for protein and DNA sequences. *Annu Rev Biophys Biophys Chem*, 20, 175-203.
- KAUFMANN, R., KIRSCH, D. & SPENGLER, B. 1994. Sequencing of Peptides in a Time-of-Flight Mass Spectrometer - Evaluation of Postsource Decay Following Matrix-Assisted Laser Desorption Ionisation (MALDI). *International Journal of Mass Spectrometry and Ion Processes*, 131, 355.
- KECK, R. G. 1995. The Use of t-Butyl Hydroperoxide as a Probe for Methionine Oxidation in Proteins. *Analytical Biochemistry*, 236, 56-62.

- KELLEHER, N. L., LIN, H. Y., VALASKOVIC, G. A., AASERUD, D. J., FRIDRIKSSON, E. K. & MCLAFFERTY, F. W. 1999. Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *Journal of the American Chemical Society*, 121, 806-812.
- KESHISHIAN, H., ADDONA, T., BURGESS, M., KUHN, E. & CARR, S. A. 2007. Quantitative, Multiplexed Assays for Low Abundance Proteins in Plasma by Targeted Mass Spectrometry and Stable Isotope Dilution. *Molecular & Cellular Proteomics*, 6, 2212-2229.
- KETTENBACH, A. N., RUSH, J. & GERBER, S. A. 2011. Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides. *Nat. Protocols*, 6, 175-186.
- KHALSA-MOYERS, G. & MCDONALD, W. H. 2006. Developments in mass spectrometry for the analysis of complex protein mixtures. *Briefings in Functional Genomics & Proteomics*, 5, 98-111.
- KHMELNITSKY, Y. L., MOZHAEV, V. V., BELOVA, A. B., SERGEEVA, M. V. & MARTINEK, K. 1991. Denaturation capacity: a new quantitative criterion for selection of organic solvents as reaction media in biocatalysis. *European Journal of Biochemistry*, 198, 31-41.
- KICMAN, A. T., PARKIN, M. C. & ILES, R. K. 2006. An introduction to mass spectrometry based proteomics—Detection and characterization of gonadotropins and related molecules. *Molecular and Cellular Endocrinology*, 260 - 262, 212-27.
- KIM, C., CHENG, C. Y., SALDANHA, S. A. & TAYLOR, S. S. 2007. PKA-I Holoenzyme Structure Reveals a Mechanism for cAMP-Dependent Activation. *Cell*, 130, 1032-1043.
- KIM, Y. H., BERRY, A. H., SPENCER, D. S. & STITES, W. E. 2001. Comparing the effect on protein stability of methionine oxidation versus mutagenesis: steps toward engineering oxidative resistance in proteins. *Protein Engineering*, 14, 343-347.
- KIMURA, E. T., NIKIFOROVA, M. N., ZHU, Z., KNAUF, J. A., NIKIFOROV, Y. E. & FAGIN, J. A. 2003. High Prevalence of BRAF Mutations in Thyroid Cancer. *Cancer Research*, 63, 1454-1457.
- KIMURA, T., NAKAMORI, M., LUECK, J. D., POULIQUIN, P., AOIKE, F., FUJIMURA, H., DIRKSEN, R. T., TAKAHASHI, M. P., DULHUNTY, A. F. & SAKODA, S. 2005. Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca²⁺-ATPase in myotonic dystrophy type 1. *Human Molecular Genetics*, 14, 2189-2200.
- KIRKPATRICK, D. S., GERBER, S. A. & GYGI, S. P. 2005a. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods*, 35, 265-273.
- KIRKPATRICK, D. S., GERBER, S. A. & GYGI, S. P. 2005b. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Mass Spectrometry in Proteomics*, 35, 265-73.
- KLEANTHOUS, C. & COGGINS, J. R. 1990. Reversible Alkylation of an Active Site Methionine Residue in Dehydroquinase*. *The Journal of Biological Chemistry*, 265, 10935-10939.
- KLEVAN, L. & TSE, Y.-C. 1983. Chemical modification of essential tyrosine residues in DNA topoisomerases. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 745, 175-180.
- KLOSE, J. 1975. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik*, 26, 231-43.
- KNYUSHKO, T. V., SHAROV, V. S., WILLIAMS, T. D., SCHÖNEICH, C. & BIGELOW, D. J. 2005. 3-Nitrotyrosine Modification of SERCA2a in the Aging Heart: A Distinct Signature of the Cellular Redox Environment†. *Biochemistry*, 44, 13071-13081.
- KÖCHER, T., ENGSTRÖM, Å. & ZUBAREV, R. A. 2004. Fragmentation of Peptides in MALDI In-Source Decay Mediated by Hydrogen Radicals. *Analytical Chemistry*, 77, 172-177.
- KOJIMA, T., SANO, K., HIRABAYASHI, T. & OBINATA, T. 1990. Characterization of C-Protein Isoforms Expressed in Developing, Denervated, and Dystrophic Chicken Skeletal Muscles by Two-Dimensional Gel Electrophoresis. *Journal of Biochemistry*, 107, 470-475.

- KOLCH, W. 2000. Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. *Biochem J.*, 351, 289-305.
- KONDO, T. & HIROHASHI, S. 2007. Application of highly sensitive fluorescent dyes (CyDye DIGE Fluor saturation dyes) to laser microdissection and two-dimensional difference gel electrophoresis (2D-DIGE) for cancer proteomics. *Nat. Protocols*, 1, 2940-2956.
- KOYANAGI, M., SUGA, H., HOSHIYAMA, D., ONO, K., IWABE, N., KUMA, K.-I. & MIYATA, T. 1998. Ancient gene duplication and domain shuffling in the animal cyclic nucleotide phosphodiesterase family1 The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases with accession numbers AB017021-AB017024.1. *FEBS letters*, 436, 323-328.
- KRUGER, R., HUGH, C. W., EDELSON-AVERBUKH, M. & LEHMANN, W. D. 2005. Iodoacetamide-alkylated methionine can mimic neutral loss of phosphoric acid from phosphopeptides as exemplified by nano-electrospray ionisation quadrupole time-of-flight parent ion scanning. *Rapid Communications in Mass Spectrometry*, 19, 1709-1716.
- KURODA, S., OHTSUKA, T., YAMAMORI, B., FUKUI, K., SHIMIZU, K. & TAKAI, Y. 1996. Different Effects of Various Phospholipids on Ki-Ras-, Ha-Ras-, and Rap1B-induced B-Raf Activation. *Journal of Biological Chemistry*, 271, 14680-14683.
- KWIETNY, H., LEVIN, G., BERGMANN, F. & BROWN, D. J. 1959. Mechanism of Enzymatic Oxidation of Purines. *Science*, 130, 711-712.
- LAHM, H.-W. & LANGEN, H. 2000. Mass spectrometry: A tool for the identification of proteins separated by gels. *Electrophoresis*, 21, 2105-2114.
- LAMONTAGNE, J., BELAND, M., FOREST, A., COTE-MARTIN, A., NASSIF, N., TOMAKI, F., MORIYON, I., MORENO, E. & PARAMITHIOTIS, E. 2010. Proteomics-based confirmation of protein expression and correction of annotation errors in the *Brucella abortus* genome. *BMC Genomics*, 11, 300.
- LANGE, V., PICOTTI, P., DOMON, B. & AEBERSOLD, R. 2008. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.*, 4, 222.
- LEE, J. E., BECK, T. W., WOJNOWSKI, L. & RAPP, U. R. 1996. Regulation of A-raf expression. *Oncogene*, 12, 1669-77.
- LEE, S. J., LEE, J. R., KIM, Y. H., PARK, Y. S., PARK, S. I., PARK, H. S. & KIM, K. P. 2007. Investigation of tyrosine nitration and nitrosylation of angiotensin II and bovine serum albumin with electrospray ionization mass spectrometry. *Rapid Communications in Mass Spectrometry*, 21, 2797-2804.
- LEEVERS, S. J., PATERSON, H. F. & MARSHALL, C. J. 1994. Requirement for Ras in Raf activation is overcome by targeting Raf to the plasma membrane. *Nature*, 369, 411-414.
- LEONI, G., LE PERA, L., FERRÈ, F., RAIMONDO, D. & TRAMONTANO, A. 2011. Coding potential of the products of alternative splicing in human. *Genome Biology*, 12, 1-10.
- LEROY, J., RICHTER, W., MIKA, D., CASTRO, L. R. V., ABI-GERGES, A., XIE, M., SCHEITRUM, C., LEFEBVRE, F., SCHITTL, J., MATEO, P., WESTENBROEK, R., CATTERALL, W. A., CHARPENTIER, F., CONTI, M., FISCHMEISTER, R. & VANDECASTEELE, G. 2011. Phosphodiesterase 4B in the cardiac L-type Ca²⁺ channel complex regulates Ca²⁺ current and protects against ventricular arrhythmias in mice. *The Journal of Clinical Investigation*, 121, 2651-2661.
- LEVIN, R. M. & WEISS, B. 1976. Mechanism by Which Psychotropic Drugs Inhibit Adenosine Cyclic 3',5'-Monophosphate Phosphodiesterase of Brain. *Molecular Pharmacology*, 12, 581-589.
- LILLEY, K. S. & FRIEDMAN, D. B. 2004. All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Review of Proteomics*, 1, 401-409.
- LINK, A. J., ENG, J., SCHIELTZ, D. M., CARMACK, E., MIZE, G. J., MORRIS, D. R., GARVIK, B. M. & YATES, J. R. 1999. Direct analysis of protein complexes using mass spectrometry. *Nat Biotech*, 17, 676-682.

- LITTLE, D. P., SPEIR, J. P., SENKO, M. W., O'CONNOR, P. B. & MCLAFFERTY, F. W. 1994. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Analytical Chemistry*, 66, 2809-2815.
- LIU, H., LIN, D. & YATES, J. R. R. 2002. Multidimensional separations for protein/peptide analysis in the post-genomic era. *Biotechniques*, 32, 898-911.
- LIVI, G. P., KMETZ, P., MCHALE, M. M., CIESLINSKI, L. B., SATHE, G. M., TAYLOR, D. P., DAVIS, R. L., TORPHY, T. J. & BALCAREK, J. M. 1990. Cloning and expression of cDNA for a human low-Km, rolipram-sensitive cyclic AMP phosphodiesterase. *Molecular and Cellular Biology*, 10, 2678-2686.
- LOBODA, A. V. & CHERNUSHEVICH, I. V. 2009. A Novel Ion Trap That Enables High Duty Cycle and Wide m/z Range on an Orthogonal Injection TOF Mass Spectrometer. *Journal of the American Society for Mass Spectrometry*, 20, 1342-1348.
- LOKUTA, A. J., MAERTZ, N. A., MEETHAL, S. V., POTTER, K. T., KAMP, T. J., VALDIVIA, H. H. & HAWORTH, R. A. 2005. Increased Nitration of Sarcoplasmic Reticulum Ca²⁺-ATPase in Human Heart Failure. *Circulation*, 111, 988-995.
- LOO, J., EDMONDS, C. & SMITH, R. 1990. Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry. *Science*, 248, 201-204.
- LUGNIER, C. 2006. Cyclic nucleotide phosphodiesterase (PDE) superfamily: A new target for the development of specific therapeutic agents. *Pharmacology & Therapeutics*, 109, 366-398.
- LUIKING, Y. C., ENGELEN, M. P. & DEUTZ, N. E. 2010. Regulation of nitric oxide production in health and disease. *Curr Opin Clin Nutr Metab Care.*, 1, 97-104.
- LUO, Y., LI, T., YU, F., KRAMER, T. & CRISTEA, I. 2010. Resolving the composition of protein complexes using a MALDI LTQ orbitrap. *Journal of the American Society for Mass Spectrometry*, 21, 34-46.
- LYTTON, J., WESTLIN, M., BURK, S. E., SHULL, G. E. & MACLENNAN, D. H. 1992. Functional comparisons between isoforms of the sarcoplasmic or endoplasmic reticulum family of calcium pumps. *Journal of Biological Chemistry*, 267, 14483-14489.
- M. NAKAO, R. A. BARRERO, Y. MUKAI, C. MOTONO, M. SUWA & K. NAKAI 2005. Large-scale analysis of human alternative protein isoforms: pattern classification and correlation with subcellular localization signals. *Nucleic Acids Res.*, 33, 8.
- M. UNLU, M. MORGAN & J. MINDEN 1997. Difference gel electrophoresis. A single gel method for detecting changes in protein extracts. *Electrophoresis*, 18, 2071-2077.
- MACLENNAN, D. H., ASAH, M. & TUPLING, A. R. 2003. The Regulation of SERCA-Type Pumps by Phospholamban and Sarcolipin. *Annals of the New York Academy of Sciences*, 986, 472-480.
- MACLENNAN, D. H., BRANDL, C. J., KORCZAK, B. & GREEN, N. M. 1985. Amino-acid sequence of a Ca²⁺ + Mg²⁺ -dependent ATPase from rabbit muscle sarcoplasmic reticulum, deduced from its complementary DNA sequence. *Nature*, 316, 696-700.
- MAKAROV, A. 2000. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry*, 72, 1156-1162.
- MAKAROV, A. & SCIGELOVA, M. 2004. Orbitrap mass analyzer. *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*. John Wiley & Sons, Ltd.
- MAMYRIN, B. A., KARATAEV, V. I., SHMIKK, D. V. & ZAGULIN, V. A. 1973. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *TrAC, Trends in Anal. Chem.*, 37, 45-48.
- MANN, M. & WILM, M. 1994. Error-Tolerant Identification of Peptides in Sequence Databases by Peptide Sequence Tags. *Analytical Chemistry*, 66, 4390-4399.
- MARAIS, R. & MARSHALL, C. J. 1996. Control of the ERK MAP kinase cascade by Ras and Raf. *Cancer Surv.*, 27, 101-25.

- MAROUGA, R., DAVID, S. & HAWKINS, E. 2005. The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Analytical and Bioanalytical Chemistry*, 382, 669-678.
- MARX, M., EYCHÈNE, A., LAUGIER, D., BÉCHADE, C., CRISANTI, P., DEZÉLÉE, P., PESSAC, B. & CALOTHY, G. 1988. A novel oncogene related to c-mil is transduced in chicken neuroretina cells induced to proliferate by infection with an avian lymphomatosis virus. *EMBO J*, 7, 3369-3373.
- MASON, C. S., SPRINGER, C. J., COOPER, R. G., SUPERTI-FURGA, G., MARSHALL, C. J. & MARAIS, R. 1999. Serine and tyrosine phosphorylations cooperate in Raf-1, but not B-Raf activation. *EMBO J*, 18, 2137-2148.
- MATALLANAS, D., BIRTWISTLE, M., ROMANO, D., ZEBISCH, A., RAUCH, J., VON KRIEGSHEIM, A. & KOLCH, W. 2011. Raf Family Kinases: Old Dogs Have Learned New Tricks. *Genes Cancer*, 2, 232-260.
- MAYYA, V., REZUAL, K., WU, L., FONG, M. B. & HAN, D. K. 2006. Absolute Quantification of Multisite Phosphorylation by Selective Reaction Monitoring Mass Spectrometry. *Molecular & Cellular Proteomics*, 5, 1146-1157.
- MCLAUGHLIN, M. M., CIESLINSKI, L. B., BURMAN, M., TORPHY, T. J. & LIVI, G. P. 1993. A low-K_m, rolipram-sensitive, cAMP-specific phosphodiesterase from human brain. Cloning and expression of cDNA, biochemical characterization of recombinant protein, and tissue distribution of mRNA. *Journal of Biological Chemistry*, 268, 6470-6476.
- MEDZIHRADESKY, K. F., CAMPBELL, J. M., BALDWIN, M. A., FALICK, A. M., JUHASZ, P., VESTAL, M. L. & BURLINGAME, A. L. 1999. The Characteristics of Peptide Collision-Induced Dissociation Using a High-Performance MALDI-TOF/TOF Tandem Mass Spectrometer. *Analytical Chemistry*, 72, 552-558.
- MENNITI, F. S., FARACI, W. S. & SCHMIDT, C. J. 2006. Phosphodiesterases in the CNS: targets for drug development. *Nat Rev Drug Discov*, 5, 660-670.
- MERCER, K., CHILOECHES, A., HÜSER, M., KIERNAN, M., MARAIS, R. & PRITCHARD, C. 2002. ERK signalling and oncogene transformation are not impaired in cells lacking A-Raf. *Oncogene*, 21, 347-55.
- MERCHANT, M. & WEINBERGER, S. R. 2000. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. *Electrophoresis*, 21, 1164-1177.
- MEUZELAAR, H. L. C., POSTHUMUS, M. A., KISTEMAKER, P. G. & KISTEMAKER, J. 1973. Curie point pyrolysis in direct combination with low voltage electron impact ionization mass spectrometry. New method for the analysis of nonvolatile organic materials. *Analytical Chemistry*, 45, 1546-1549.
- MIKULA, M., SCHREIBER, M., HUSAK, Z., KUCEROVA, L., RUTH, J., WIESER, R., ZATLOUKAL, K., BEUG, H., WAGNER, E. F. & BACCARINI, M. 2001. Embryonic lethality and fetal liver apoptosis in mice lacking the c-raf-1 gene. *EMBO J*, 20, 1952-1962.
- MILES D, H. 2001. PDE4 cAMP-specific phosphodiesterases. *Progress in Nucleic Acid Research and Molecular Biology*. Academic Press.
- MILLAR, J. K., PICKARD, B. S., MACKIE, S., JAMES, R., CHRISTIE, S., BUCHANAN, S. R., MALLOY, M. P., CHUBB, J. E., HUSTON, E., BAILLIE, G. S., THOMSON, P. A., HILL, E. V., BRANDON, N. J., RAIN, J.-C., CAMARGO, L. M., WHITING, P. J., HOUSLAY, M. D., BLACKWOOD, D. H. R., MUIR, W. J. & PORTEOUS, D. J. 2005. DISC1 and PDE4B Are Interacting Genetic Factors in Schizophrenia That Regulate cAMP Signaling. *Science*, 310, 1187-1191.
- MOELLING, K., HEIMANN, B., BEIMLING, P., RAPP, U. R. & SANDER, T. 1984. Serine- and threonine-specific protein kinase activities of purified gag-mil and gag-raf proteins. *Nature*, 312, 558-61.
- MONTPETIT, B. 2003. Proteomics in the Post-Genomics Era: Weighing in the Data. *BioTeach Reviews and Readings*, 1, 9-12.

- MORRISON, D. K., KAPLAN, D. R., ESCOBEDO, J. A., RAPP, U. R., ROBERTS, T. M. & WILLIAMS, L. T. 1989. Direct activation of the serine/threonine kinase activity of raf-1 through tyrosine phosphorylation by the PDGF β -receptor. *Cell*, 58, 649-657.
- MOTOYAMA, A. & YATES, J. R. 2008. Multidimensional LC Separations in Shotgun Proteomics. *Analytical Chemistry*, 80, 7187-7193.
- MULLER, T., ENGELS, P. & FOZARD, J. R. 1996. Subtypes of the type 4 cAMP phosphodiesterases: Structure, regulation and selective inhibition. *Trends in Pharmacological Sciences*, 17, 294-298.
- MUNDY, C. 2001. The human genome project: a historical perspective. *Pharmacogenomics*, 2, 37-49.
- NAKAMURA, K.-I., OKUYA, Y., KATAHIRA, M., YOSHIDA, S., WADA, S. & OKUNO, M. 1994. Analysis of tubulin isoforms by two-dimensional gel electrophoresis using SDS-polyacrylamide gel electrophoresis in the first dimension. *Journal of Biochemical and Biophysical Methods*, 24, 195-203.
- NEKHOROSHKOVA, E., ALBERT, S., BECKER, M. & RAPP, U. R. 2009. A-RAF Kinase Functions in ARF6 Regulated Endocytic Membrane Traffic. *PLoS ONE*, 4, e4647.
- NESVIZHSHKII, A. I. & AEBERSOLD, R. 2005. Interpretation of Shotgun Proteomic Data: The Protein Inference Problem. *Molecular & Cellular Proteomics*, 4, 1419-1440.
- NISSUM, M., SCHNEIDER, U., KUHFUSS, S., OBERMAIER, C., WILDGRUBER, R., POSCH, A. & ECKERSKORN, C. 2004. In-Gel Digestion of Proteins Using a Solid-Phase Extraction Microplate. *Analytical Chemistry*, 76, 2040-2045.
- NORRGRAN, J., WILLIAMS, T. L., WOOLFITT, A. R., SOLANO, M. I., PIRKLE, J. L. & BARR, J. R. 2009. Optimization of digestion parameters for protein quantification. *Analytical Biochemistry*, 393, 48-55.
- NUMATA, S., UENO, S., IGA, J., SONG, H., NAKATAKI, M., TAYOSHI, S. & AL., E. 2009a. Positive association of the PDE4B (phosphodiesterase 4B) gene with schizophrenia in the Japanese population. *J Psychiatr Res*, 43, 7-12.
- O'DONNELL, J. M. & FRITH, S. 1999. Behavioral Effects of Family-Selective Inhibitors of Cyclic Nucleotide Phosphodiesterases. *Pharmacology Biochemistry and Behavior*, 63, 185-192.
- O'FARRELL, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.*, 250, 4007-21.
- OBERNOLTE, R., BHAKTA, S., ALVAREZ, R., BACH, C., ZUPPAN, P., MULLEINS, M., JARNAGIN, K. & SHELTON, E. R. 1993. The cDNA of a human lymphocyte cyclic-AMP phosphodiesterase (PDE IV) reveals a multigene family. *Gene*, 129, 239-247.
- ODA, Y., HUANG, K., CROSS, F. R., COWBURN, D. & CHAIT, B. T. 1999. Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences*, 96, 6591-6596.
- OHSHIMA, H., FRIESEN, M., BROUET, I. & BARTSCH, H. 1990. Nitrotyrosine as a new marker for endogenous nitrosation and nitration of proteins. *Food and Chemical Toxicology*, 28, 647-652.
- OKI, N., TAKAHASHI, S.-I., HIDAKA, H. & CONTI, M. 2000. Short Term Feedback Regulation of cAMP in FRTL-5 Thyroid Cells. *Journal of Biological Chemistry*, 275, 10831-10837.
- OLD, W. M., MEYER-ARENDT, K., AVELINE-WOLF, L., PIERCE, K. G., MENDOZA, A., SEVINSKY, J. R., RESING, K. A. & AHN, N. G. 2005. Comparison of Label-free Methods for Quantifying Human Proteins by Shotgun Proteomics. *Molecular & Cellular Proteomics*, 4, 1487-1502.
- OLDREIVE, C., ZHAO, K., PAGANGA, G., HALLIWELL, B. & RICE-EVANS, C. 1998. Inhibition of Nitrous Acid-Dependent Tyrosine Nitration and DNA Base Deamination by Flavonoids and Other Phenolic Compounds. *Chemical Research in Toxicology*, 11, 1574-1579.
- OLESEN, C., PICARD, M., WINTHER, A.-M. L., GYRUP, C., MORTH, J. P., OXVIG, C., MOLLER, J. V. & NISSEN, P. 2007. The structural basis of calcium transport by the calcium pump. *Nature*, 450, 1036-1042.

- OLSEN, J. V., MACEK, B., LANGE, O., MAKAROV, A., HORNING, S. & MANN, M. 2007. Higher-energy C-trap dissociation for peptide modification analysis. *Nat Meth*, 4, 709-712.
- OLSEN, J. V., ONG, S.-E. & MANN, M. 2004. Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. *Molecular & Cellular Proteomics*, 3, 608-614.
- ONG, S.-E., BLAGOEV, B., KRATCHMAROVA, I., KRISTENSEN, D. B., STEEN, H., PANDEY, A. & MANN, M. 2002. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Molecular & Cellular Proteomics*, 1, 376-386.
- ORLY, J. & SCHRAMM, M. 1976. Coupling of catecholamine receptor from one cell with adenylate cyclase from another cell by cell fusion. *PNAS*, 73, 4410-4414.
- OTOWA, T., KAWAMURA, Y., SUGAYA, N., YOSHIDA, E., SHIMADA, T., LIU, X., TOCHIGI, M., UMEKAGE, T., MIYAGAWA, T., NISHIDA, N., KAIYA, H., OKAZAKI, Y., TOKUNAGA, K. & SASAKI, T. 2011. Association study of PDE4B with panic disorder in the Japanese population. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 35, 545-549.
- PAIZS, B. & SUHAI, S. 2005. Fragmentation pathways of protonated peptides. *Mass Spectrometry Reviews*, 24, 508-548.
- PALMGREN, M. G. & NISSEN, P. 2011. P-Type ATPases. *Annual Review of Biophysics*, 40, 243-266.
- PAN, Q., SHAI, O., LEE, L. J., FREY, B. J. & BLENCOWE, B. J. 2008a. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet*, 40, 1413-1415.
- PAN, S., AEBERSOLD, R., CHEN, R., RUSH, J., GOODLETT, D. R., MCINTOSH, M. W., ZHANG, J. & BRENTNALL, T. A. 2008b. Mass Spectrometry Based Targeted Protein Quantification: Methods and Applications. *Journal of Proteome Research*, 8, 787-797.
- PANDEY, S. C., ZHANG, H., ROY, A. & XU, T. 2005. Deficits in amygdaloid cAMP-responsive element-binding protein signaling play a role in genetic predisposition to anxiety and alcoholism. *The Journal of Clinical Investigation*, 115, 2762-2773.
- PAPPIN, D. J. C., HOJRUP, P. & BLEASBY, A. J. 1993. Rapid identification of proteins by peptide-mass fingerprinting. *Current Biology*, 3, 327-332.
- PATEL, T. B., DU, Z., PIERRE, S., CARTIN, L. & SCHOLICH, K. 2001. Molecular biological approaches to unravel adenylyl cyclase signaling and function. *Gene*, 269, 13-25.
- PATTERSON, S. D. & AEBERSOLD, R. H. 2003. Proteomics: the first decade and beyond. *Nat Genet*, 33, 311-23.
- PATTON, W. F., SCHULENBERG, B. & STEINBERG, T. H. 2002. Two-dimensional gel electrophoresis; better than a poke in the ICAT? *Current Opinion in Biotechnology*, 13, 321-328.
- PAUL, W. 1990. Electromagnetic traps for charged and neutral particles. *Reviews of Modern Physics*, 62, 531-540.
- PAWSON, T. & NASH, P. 2000. Protein-protein interactions define specificity in signal transduction. *Genes & Development*, 14, 1027-1047.
- PEGORARO, C., POLLET, N. & MONSORO-BURQ, A. H. 2011. Tissue-specific expression of Sarcoplasmic/Endoplasmic Reticulum Calcium ATPases (ATP2A/SERCA) 1, 2, 3 during *Xenopus laevis* development. *Gene Expression Patterns*, 11, 122-128.
- PERIASAMY, M. & HUKU, S. 2001. SERCA Pump Level is a Critical Determinant of Ca²⁺ Homeostasis and Cardiac Contractility. *Journal of Molecular and Cellular Cardiology*, 33, 1053-1063.
- PERIASAMY, M. & KALYANASUNDARAM, A. 2007. SERCA pump isoforms: Their role in calcium transport and disease. *Muscle & Nerve*, 35, 430-442.
- PERRY, M. J. & HIGGS, G. A. 1998. Chemotherapeutic potential of phosphodiesterase inhibitors. *Current Opinion in Chemical Biology*, 2, 472-481.
- PEYSSONNAUX, C. & EYCHÈNE, A. 2001. The Raf/MEK/ERK pathway: new concepts of activation. *Biology of the Cell*, 93, 53-62.

- PICHLER, P., KÖCHER, T., HOLZMANN, J., MAZANEK, M., TAUS, T., AMMERER, G. & MECHTLER, K. 2010. Peptide Labeling with Isobaric Tags Yields Higher Identification Rates Using iTRAQ 4-Plex Compared to TMT 6-Plex and iTRAQ 8-Plex on LTQ Orbitrap. *Analytical Chemistry*, 82, 6549-6558.
- POLSON, C., SARKAR, P., INCLEDON, B., RAGUVARAN, V. & GRANT, R. 2003. Optimization of protein precipitation based upon effectiveness of protein removal and ionization effect in liquid chromatography–tandem mass spectrometry. *Journal of Chromatography B*, 785, 263-275.
- PORTER, J. J., MEHIGH, R. J., WILDSMITH, J., HARVEY, M. C., BOYLE, J. A., EWING, B. K., HEUERMANN, K. E., POLAND, K. A., KAPPEL, W. K. & SCOTT, G. B. I. Absolute Quantification as an Effective Tool for Measuring Protein Expression in Targeted Gene Knockdown Experiments. *Sigma-Aldrich Biotechnology*.
- POSTHUMUS, M. A., KISTEMAKER, P. G., MEUZELAAR, H. L. C. & TEN NOEVER DE BRAUW, M. C. 1978. Laser desorption-mass spectrometry of polar nonvolatile bio-organic molecules. *Analytical Chemistry*, 50, 985-991.
- PRAHALLAD, A., SUN, C., HUANG, S., DI NICOLANTONIO, F., SALAZAR, R., ZECCHIN, D., BEIJERSBERGEN, R. L., BARDELLI, A. & BERNARDS, R. 2012. Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature*, 483, 100-103.
- PRATT, J. M., SIMPSON, D. M., DOHERTY, M. K., RIVERS, J., GASKELL, S. J. & BEYNON, R. J. 2006. Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat. Protocols*, 1, 1029-1043.
- PRITCHARD, C. A., BOLIN, L., SLATTERY, R., MURRAY, R. & MCMAHON, M. 1996. Post-natal lethality and neurological and gastrointestinal defects in mice with targeted disruption of the A-Raf protein kinase gene. *Current Biology*, 6, 614-617.
- PRITCHARD, C. A., HAYES, L., WOJNOWSKI, L., ZIMMER, A., MARAIS, R. M. & NORMAN, J. C. 2004. B-Raf Acts via the ROCKII/LIMK/Cofilin Pathway To Maintain Actin Stress Fibers in Fibroblasts. *Mol. Cell. Biol.*, 24, 5937-5952.
- PURCELL, A. W. & GORMAN, J. J. 2001. The use of post-source decay in matrix-assisted laser desorption/ionisation mass spectrometry to delineate T cell determinants. *Journal of Immunological Methods*, 249, 17-31.
- RABILLOUD, T. 1990. Mechanisms of protein silver staining in polyacrylamide gels: a 10 year synthesis. *Electrophoresis*, 11, 785-94.
- RADABAUGH, M. R., NEMIROVSKIY, O. V., MISKO, T. P., AGGARWAL, P. & MATHEWS, W. R. 2008. Immunoaffinity liquid chromatography–tandem mass spectrometry detection of nitrotyrosine in biological fluids: Development of a clinically translatable biomarker. *Analytical Biochemistry*, 380, 68-76.
- RAIKOS, V., HANSEN, R., CAMPBELL, L. & EUSTON, S. R. 2006. Separation and identification of hen egg protein isoforms using SDS–PAGE and 2D gel electrophoresis with MALDI-TOF mass spectrometry. *Food Chemistry*, 99, 702-710.
- RAJAGOPALAN, K. V., FRIDOVICH, I. & HANDLER, P. 1961. Competitive Inhibition of Enzyme Activity by Urea. *Journal of Biological Chemistry*, 236, 1059-1065.
- RAPP, U. R., GOLDSBOROUGH, M. D., MARK, G. E., BONNER, T. I., GROFFEN, J., REYNOLDS, F. H. & STEPHENSON, J. R. 1983. Structure and biological activity of v-raf, a unique oncogene transduced by a retrovirus. *Proceedings of the National Academy of Sciences*, 80, 4218-4222.
- RAPPSILBER, J. & MANN, M. 2002. What does it mean to identify a protein in proteomics? *Trends in Biochemical Sciences*, 27, 74-78.
- RAUCH, J., MORAN-JONES, K., ALBRECHT, V., SCHWARZL, T., HUNTER, K., GIRES, O. & KOLCH, W. 2011. c-Myc Regulates RNA Splicing of the A-Raf Kinase and Its Activation of the ERK Pathway. *Cancer Research*, 71, 4664-4674.

- RAUCH, J., O'NEILL, E., MACK, B., MATTHIAS, C., MUNZ, M., KOLCH, W. & GIRES, O. 2010. Heterogeneous Nuclear Ribonucleoprotein H Blocks MST2-Mediated Apoptosis in Cancer Cells by Regulating a-raf Transcription. *Cancer Research*, 70, 1679-1688.
- REED, R. 1989. The organization of 3' splice-site sequences in mammalian introns. *Genes & Development*, 3, 2113-2123.
- REGNIER, F. E. & JULKA, S. 2006. Primary amine coding as a path to comparative proteomics. *Proteomics*, 6, 3968-3979.
- RICHARDS, D. A., SILVA, M. A. & DEVAL, A. J. 2006. Electrochemical detection of free 3-nitrotyrosine: Application to microdialysis studies. *Analytical Biochemistry*, 351, 77-83.
- RICHTER, W. & CONTI, M. 2004. The Oligomerization State Determines Regulatory Properties and Inhibitor Sensitivity of Type 4 cAMP-specific Phosphodiesterases. *Journal of Biological Chemistry*, 279, 30338-30348.
- ROEPSTORFF, P. & FOHLMAN, J. 1984. Letter to the editors. *Biological Mass Spectrometry*, 11, 601-601.
- ROGOZIN, I. B., SVERDLOV, A. V., BABENKO, V. N. & KOONIN, E. V. 2005. Analysis of evolution of exon-intron structure of eukaryotic genes. *Briefings in Bioinformatics*, 6, 118-134.
- ROLFE, D. F. & BROWN, G. C. 1997. Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiological Reviews*, 77, 731-758.
- ROMANO, D., MATALLANAS, D., WEITSMAN, G., PREISINGER, C., NG, T. & KOLCH, W. 2010. Proapoptotic Kinase MST2 Coordinates Signaling Crosstalk between RASSF1A, Raf-1, and Akt. *Cancer Research*, 70, 1195-1203.
- ROSS, P. L., HUANG, Y. N., MARCHESE, J. N., WILLIAMSON, B., PARKER, K., HATTAN, S., KHAINOVSKI, N., PILLAI, S., DEY, S., DANIELS, S., PURKAYASTHA, S., JUHASZ, P., MARTIN, S., BARTLET-JONES, M., HE, F., JACOBSON, A. & PAPPIN, D. J. 2004. Multiplexed Protein Quantitation in *Saccharomyces cerevisiae* Using Amine-reactive Isobaric Tagging Reagents. *Molecular & Cellular Proteomics*, 3, 1154-1169.
- ROULHAC, P. L., WARD, J. M., THOMPSON, J. W., SODERBLOM, E. J., SILVA, M., MOSELEY, M. A. & JARVIS, E. D. 2011. Microproteomics: Quantitative Proteomic Profiling of Small Numbers of Laser-Captured Cells. *Cold Spring Harbor Protocols*, 2011, pdb.prot5573.
- RUTTEN, K., WALLACE, T. L., WORKS, M., PRICKAERTS, J., BLOKLAND, A., NOVAK, T. J., SANTARELLI, L. & MISNER, D. L. 2011. Enhanced long-term depression and impaired reversal learning in phosphodiesterase 4B-knockout (PDE4B^{-/-}) mice. *Neuropharmacology*, 61, 138-147.
- RYAN, C. M., SOUDA, P., BASSILIAN, S., UJWAL, R., ZHANG, J., ABRAMSON, J., PING, P., DURAZO, A., BOWIE, J. U., HASAN, S. S., BANIULIS, D., CRAMER, W. A., FAULL, K. F. & WHITELEGGE, J. P. 2010. Post-translational Modifications of Integral Membrane Proteins Resolved by Top-down Fourier Transform Mass Spectrometry with Collisionally Activated Dissociation. *Molecular & Cellular Proteomics*, 9, 791-803.
- S. A. GERBER, J. RUSH, O. STEMMAN, M. W. KIRSCHNER & S. P. GYGI 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Nat. Acad. Sci. U.S.A.*, 100, 5.
- S. P. GYGI, G. L. CORTHALS, Y. ZHANG, Y. ROCHON & R. AEBERSOLD 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Nat. Acad. Sci.*, 97, 5.
- S. PAN, R. AEBERSOLD, R. CHEN, J. RUSH, D. R. GOODLETT, M. W. MCINTOSH, J. ZHANG & T. A. BRENTNALL 2009. Mass Spectrometry Based Targeted Protein Quantification: Methods and Applications. *J. Proteome. Res.*, 8, 10.
- SAUER, S. & KLIEM, M. 2010. Mass spectrometry tools for the classification and identification of bacteria. *Nat Rev Micro*, 8, 74-82.
- SCHLESSINGER, J. 2000. Cell Signaling by Receptor Tyrosine Kinases. *Cell*, 103, 211-225.

- SCHLESSINGER, J. & BAR-SAGI, D. 1994. Activation of Ras and Other Signaling Pathways by Receptor Tyrosine Kinases. *Cold Spring Harbor Symposia on Quantitative Biology*, 59, 173-179.
- SCHNÖLZER, M., JEDRZEJEWSKI, P. & LEHMANN, W. D. 1996. Protease-catalyzed incorporation of ¹⁸O into peptide fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. *Electrophoresis*, 17, 945-953.
- SCHULZE, W. X. & USADEL, B. 2010. Quantitation in Mass-Spectrometry-Based Proteomics. *Annual Review of Plant Biology*, 61, 491-516.
- SCOTT, A. I. F., PERINI, A. F., SHERING, P. A. & WHALLEY, L. J. 1991. INPATIENT MAJOR DEPRESSION - IS ROLIPRAM AS EFFECTIVE AS AMITRIPTYLINE. *European Journal of Clinical Pharmacology*, 40, 127-129.
- SEIBERT, V., EBERT, M. P. A. & BUSCHMANN, T. 2005. Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and biomarker discovery. *Briefings in Functional Genomics & Proteomics*, 4, 16-26.
- SHAW, J., ROWLINSON, R., NICKSON, J., STONE, T., SWEET, A., WILLIAMS, K. & TONGE, R. 2003. Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. *Proteomics*, 3, 1181-95.
- SHECHTER, Y. 1984. Selective Oxidation and Reduction of Methionine Residues in Peptides and Proteins by Oxygen Exchange between Sulfoxide and Sulfide*. *The Journal of Biological Chemistry*, 261, 66-70.
- SHEN, Y. & SMITH, R. D. 2002. Proteomics based on high-efficiency capillary separations. *Electrophoresis*, 23, 3106-3124.
- SHEPHERD, M., MCSORLEY, T., OLSEN, A. E., JOHNSTON, L. A., THOMSON, N. C., BAILLIE, G. S., HOUSLAY, M. D. & BOLGER, G. B. 2003. Molecular cloning and subcellular distribution of the novel PDE4B4 cAMP-specific phosphodiesterase isoform. *Biochem J*, 370, 429-438.
- SHERWOOD, C. A., EASTHAM, A., LEE, L. W., RISLER, J., VITEK, O. & MARTIN, D. B. 2009. Correlation between γ -Type Ions Observed in Ion Trap and Triple Quadrupole Mass Spectrometers. *Journal of Proteome Research*, 8, 4243-4251.
- SHEVCHENKO, A., WILM, M., VORM, O. & MANN, M. 1996. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Analytical Chemistry*, 68, 850-858.
- SHI, Y., XIANG, R., HORVÁTH, C. & WILKINS, J. A. 2004. The role of liquid chromatography in proteomics. *Journal of Chromatography A*, 1053, 27-36.
- SHUKLA, A. K. & FUTRELL, J. H. 2000. Tandem mass spectrometry: dissociation of ions by collisional activation. *Journal of Mass Spectrometry*, 35, 1069-1090.
- SIEPEN, J. A., KEEVIL, E.-J., KNIGHT, D. & HUBBARD, S. J. 2006. Prediction of Missed Cleavage Sites in Tryptic Peptides Aids Protein Identification in Proteomics. *Journal of Proteome Research*, 6, 399-408.
- SILVESTRE, J. S., FERNÁNDEZ, A. G. & PALACIOS, J. M. 1999. Effects of rolipram on the elevated plus-maze test in rats: a preliminary study. *Journal of Psychopharmacology*, 13, 274-277.
- SKOU, J. C. 1960. Further investigations on a Mg⁺⁺ + Na⁺-activated adenosinetriphosphatase, possibly related to the active, linked transport of Na⁺ and K⁺ across the nerve membrane. *Biochimica et Biophysica Acta*, 42, 6-23.
- SLENO, L. & VOLMER, D. A. 2004. Ion activation methods for tandem mass spectrometry. *Journal of Mass Spectrometry*, 39, 1091-1112.
- SMIRNOV, I. P., ZHU, X., TAYLOR, T., HUANG, Y., ROSS, P., PAPAYANOPOULOS, I. A., MARTIN, S. A. & PAPPIN, D. J. 2004. Suppression of α -Cyano-4-hydroxycinnamic Acid Matrix Clusters and Reduction of Chemical Noise in MALDI-TOF Mass Spectrometry. *Analytical Chemistry*, 76, 2958-2965.
- SMITH, P. G., WANG, F., WILKINSON, K. N., SAVAGE, K. J., KLEIN, U., NEUBERG, D. S., BOLLAG, G., SHIPP, M. A. & AGUIAR, R. C. T. 2005. The phosphodiesterase PDE4B limits cAMP-

- associated PI3K/AKT-dependent apoptosis in diffuse large B-cell lymphoma. *Blood*, 105, 308-316.
- SOGA, M., MATSUZAWA, A. & ICHIJO, H. 2012. Oxidative Stress-Induced Diseases via the ASK1 Signaling Pathway. *International Journal of Cell Biology*, 2012, 5.
- SOKOLOVSKY, M., RIORDAN, J. F. & VALLEE, B. L. 1966. Tetranitromethane. A Reagent for the Nitration of Tyrosyl Residues in Proteins*. *Biochemistry*, 5, 3582-3589.
- SOUNESS, J. E., ALDOUS, D. & SARGENT, C. 2000. Immunosuppressive and anti-inflammatory effects of cyclic AMP phosphodiesterase (PDE) type 4 inhibitors. *Immunopharmacology*, 47, 127-162.
- SPEICHER, K., KOLBAS, O., HARPER, S. & SPEICHER, D. 2000. Systematic analysis of peptide recoveries from in-gel digestions for protein identifications in proteome studies. *J Biomol Tech.*, 11, 74-86.
- STEEN, H. & MANN, M. 2004. The abc's (and xyz's) of peptide sequencing. *Nat Rev Mol Cell Biol*, 5, 699-711.
- STEINMANN, D., JI, J. A., WANG, Y. J. & SCHÖNEICH, C. 2012. Oxidation of Human Growth Hormone by Oxygen-Centered Radicals: Formation of Leu-101 Hydroperoxide and Tyr-103 Oxidation Products. *Molecular Pharmaceutics*, 9, 803-814.
- STEPHENS, P. E. & COCKETT, M. I. 1989. The construction of a highly efficient and versatile set of mammalian expression vectors. *Nucleic Acids Res.*, 17, 7110.
- STOKOE, D., MACDONALD, S., CADWALLADER, K., SYMONS, M. & HANCOCK, J. 1994. Activation of Raf as a result of recruitment to the plasma membrane. *Science*, 264, 1463-1467.
- STORM, S. M., CLEVELAND, J. L. & RAPP, U. R. 1990. Expression of raf family proto-oncogenes in normal mouse tissues. *Oncogene*, 5, 345-51.
- SUBBARAO, B. & KENKARE, U. W. 1977. Reaction of brain hexokinase with tetranitromethane: Oxidation of essential thiol groups. *Archives of Biochemistry and Biophysics*, 181, 8-18.
- SUCKAU, D., RESEMANN, A., SCHUERENBERG, M., HUFNAGEL, P., FRANZEN, J. & HOLLE, A. 2003. A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Analytical and Bioanalytical Chemistry*, 376, 952-965.
- SUNG, B.-J., YEON HWANG, K., HO JEON, Y., LEE, J. I., HEO, Y.-S., HWAN KIM, J., MOON, J., MIN YOON, J., HYUN, Y.-L., KIM, E., JIN EUM, S., PARK, S.-Y., LEE, J.-O., GYU LEE, T., RO, S. & MYUNG CHO, J. 2003. Structure of the catalytic domain of human phosphodiesterase 5 with bound drug molecules. *Nature*, 425, 98-102.
- SUTHERLAND, E. W. & RALL, T. W. 1957. THE PROPERTIES OF AN ADENINE RIBONUCLEOTIDE PRODUCED WITH CELLULAR PARTICLES, ATP, Mg⁺⁺, AND EPINEPHRINE OR GLUCAGON. *Journal of the American Chemical Society*, 79, 3608-3608.
- SUTHERLAND, E. W. & RALL, T. W. 1958. FRACTIONATION AND CHARACTERIZATION OF A CYCLIC ADENINE RIBONUCLEOTIDE FORMED BY TISSUE PARTICLES *J. Biol. Chem.*, 232, 1077-1092.
- SUTRAVE, P., BONNER, T. I., RAPP, U. R., JANSEN, H. W., PATSCHINSKY, T. & BISTER, K. 1984. Nucleotide sequence of avian retroviral oncogene v-mil: homologue of murine retroviral oncogene v-raf. *Nature*, 309, 85-88.
- TAKAHASHI, M., TERWILLIGER, R., LANE, C., MEZES, P. S., CONTI, M. & DUMAN, R. S. 1999. Chronic antidepressant administration increases the expression of cAMP-specific phosphodiesterase 4A and 4B isoforms. *J. Neurosci.*, 19, 610-618.
- TANAKA, K., WAKI, H., IDO, Y., AKITA, S., YOSHIDA, Y., YOSHIDA, T. & MATSUO, T. 1988. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 2, 151-153.
- TASKÉN, K. & AANDAHL, E. M. 2004. Localized Effects of cAMP Mediated by Distinct Routes of Protein Kinase A. *Physiological Reviews*, 84, 137-167.
- TERAI, K. & MATSUDA, M. 2005. Ras binding opens c-Raf to expose the docking site for mitogen-activated protein kinase kinase. *EMBO Rep*, 6, 251-255.
- THOMAS, J. J. 1910. LXXXIII. Rays of positive electricity. *Philosophical Magazine Series 6*, 20.

- THOMPSON, A., SCHÄFER, J., KUHN, K., KIENLE, S., SCHWARZ, J., SCHMIDT, G., NEUMANN, T. & HAMON, C. 2003. Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Analytical Chemistry*, 75, 1895-1904.
- THOMPSON, W. J. 1991. Cyclic nucleotide phosphodiesterases: Pharmacology, biochemistry and function. *Pharmacology & Therapeutics*, 51, 13-33.
- THORREZ, L., VAN DEUN, K., TRANCHEVENT, L.-C., VAN LOMMEL, L., ENGELEN, K., MARCHAL, K., MOREAU, Y., VAN MECHELEN, I. & SCHUIT, F. 2008. Using Ribosomal Protein Genes as Reference: A Tale of Caution. *PLoS ONE*, 3, e1854.
- TONGE, R., SHAW, J., MIDDLETON, B., ROWLINSON, R., RAYNER, S., YOUNG, J., POGNAN, F., HAWKINS, E., CURRIE, I. & DAVIDSON, M. 2001. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics*, 1, 377-96.
- TORGERSON, D. F., SKOWRONSKI, R. P. & MACFARLANE, R. D. 1974. New approach to the mass spectroscopy of non-volatile compounds. *Biochemical and Biophysical Research Communications*, 60, 616-621.
- TORPHY, T. J. 1998. Phosphodiesterase Isozymes. *American Journal of Respiratory and Critical Care Medicine*, 157, 351-370.
- TRAN, J. C., ZAMDBORG, L., AHLF, D. R., LEE, J. E., CATHERMAN, A. D., DURBIN, K. R., TIPTON, J. D., VELLAICHAMY, A., KELLIE, J. F., LI, M., WU, C., SWEET, S. M. M., EARLY, B. P., SIUTI, N., LEDUC, R. D., COMPTON, P. D., THOMAS, P. M. & KELLEHER, N. L. 2011. Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature*, 480, 254-258.
- TRESS, M. L., MARTELLI, P. L., FRANKISH, A., REEVES, G. A., WESSELINK, J. J., YEATS, C., ÓLASON, P. Ó., ALBRECHT, M., HEGYI, H., GIORGETTI, A., RAIMONDO, D., LAGARDE, J., LASKOWSKI, R. A., LÓPEZ, G., SADOWSKI, M. I., WATSON, J. D., FARISELLI, P., ROSSI, I., NAGY, A., KAI, W., STØRLING, Z., ORSINI, M., ASSENOV, Y., BLANKENBURG, H., HUTHMACHER, C., RAMÍREZ, F., SCHLICKER, A., DENOEUDE, F., JONES, P., KERRIEN, S., ORCHARD, S., ANTONARAKIS, S. E., REYMOND, A., BIRNEY, E., BRUNAK, S., CASADIO, R., GUIGO, R., HARROW, J., HERMIAKOB, H., JONES, D. T., LENGAUER, T., A. ORENGO, C., PATTHY, L., THORNTON, J. M., TRAMONTANO, A. & VALENCIA, A. 2007. The implications of alternative splicing in the ENCODE protein complement. *Proceedings of the National Academy of Sciences*, 104, 5495-5500.
- TSUJI, J., NYDZA, R., WOLCOTT, E., MANNOR, E., MORAN, B., HESSON, G., ARVIDSON, T., HOWE, K., HAYES, R., RAMIREZ, M. & WAY, M. 2010. The Frequencies of Amino Acids Encoded by Genomes that Utilize Standard and Nonstandard Genetic Codes. *BIOS*, 81, 22-31.
- UNWIN, R. D., EVANS, C. A. & WHETTON, A. D. 2006. Relative quantification in proteomics: new approaches for biochemistry. *Trends Biochem. Sci.*, 31, 473-84.
- UNWIN, R. D. & WHETTON, A. D. 2006. Systematic proteome and transcriptome analysis of stem cell populations. *Cell Cycle*, 5, 1587-91.
- UNWIN, R. D. & WHETTON, A. D. 2007. How Will Haematologists Use Proteomics? *Blood Reviews*, 21, 315-26.
- V. MAYYA, K. REZUAL, L. WU, M. B. FONG & D. K. HAN 2006. Absolute Quantification of Multisite Phosphorylation by Selective Reaction Monitoring Mass Spectrometry. *Mol. Cell Prot.*, 5, 11.
- VALLEJO, A. N., POGULIS, R. J. & PEASE, L. R. 2008. PCR Mutagenesis by Overlap Extension and Gene SOE. *Cold Spring Harbor Protocols*, 2008, pdb.prot4861.
- VAUDEL, M., SICKMANN, A. & MARTENS, L. 2010. Peptide and protein quantification: A map of the minefield. *Proteomics*, 10, 650-670.
- VENABLES, J. P. 2004. Aberrant and Alternative Splicing in Cancer. *Cancer Research*, 64, 7647-7654.
- VENTER, J. C., ADAMS, M. D., MYERS, E. W., LI, P. W., MURAL, R. J., SUTTON, G. G., SMITH, H. O., YANDELL, M., EVANS, C. A., HOLT, R. A., GOCAYNE, J. D., AMANATIDES, P., BALLEW, R. M.,

- HUSON, D. H., WORTMAN, J. R., ZHANG, Q., KODIRA, C. D., ZHENG, X. H., CHEN, L., SKUPSKI, M., SUBRAMANIAN, G., THOMAS, P. D., ZHANG, J., GABOR MIKLOS, G. L., NELSON, C., BRODER, S., CLARK, A. G., NADEAU, J., MCKUSICK, V. A., ZINDER, N., LEVINE, A. J., ROBERTS, R. J., SIMON, M., SLAYMAN, C., HUNKAPILLER, M., BOLANOS, R., DELCHER, A., DEW, I., FASULO, D., FLANIGAN, M., FLOREA, L., HALPERN, A., HANNENHALLI, S., KRAVITZ, S., LEVY, S., MOBARRY, C., REINERT, K., REMINGTON, K., ABU-THREIDEH, J., BEASLEY, E., BIDDICK, K., BONAZZI, V., BRANDON, R., CARGILL, M., CHANDRAMOULISWARAN, I., CHARLAB, R., CHATURVEDI, K., DENG, Z., FRANCESCO, V. D., DUNN, P., EILBECK, K., EVANGELISTA, C., GABRIELIAN, A. E., GAN, W., GE, W., GONG, F., GU, Z., GUAN, P., HEIMAN, T. J., HIGGINS, M. E., JI, R.-R., KE, Z., KETCHUM, K. A., LAI, Z., LEI, Y., LI, Z., LI, J., LIANG, Y., LIN, X., LU, F., MERKULOV, G. V., MILSHINA, N., MOORE, H. M., NAIK, A. K., NARAYAN, V. A., NEELAM, B., NUSSKERN, D., RUSCH, D. B., SALZBERG, S., SHAO, W., SHUE, B., SUN, J., WANG, Z. Y., WANG, A., WANG, X., WANG, J., WEI, M.-H., WIDES, R., XIAO, C., YAN, C., et al. 2001. The Sequence of the Human Genome. *Science*, 291, 1304-1351.
- VER HEYEN, M., HEYMANS, S., ANTOONS, G., REED, T., PERIASAMY, M., AWEDE, B., LEBACQ, J., VANGHELuwe, P., DEWERCHIN, M., COLLEN, D., SIPIDO, K., CARMELIET, P. & WUYTACK, F. 2001. Replacement of the Muscle-Specific Sarcoplasmic Reticulum Ca²⁺-ATPase Isoform SERCA2a by the Nonmuscle SERCA2b Homologue Causes Mild Concentric Hypertrophy and Impairs Contraction-Relaxation of the Heart. *Circulation Research*, 89, 838-846.
- VESTAL, M. L., JUHASZ, P. & MARTIN, S. A. 1995. Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 9, 1044-1050.
- VICKERY, H. B. 1950. The Origin of the Word Protein. *Yale J Biol Med.*, 22, 387-393.
- WALL, M. J., CROWELL, A. M. J., SIMMS, G. A., LIU, F. & DOUCETTE, A. A. 2011. Implications of partial tryptic digestion in organic-aqueous solvent systems for bottom-up proteome analysis. *Analytica Chimica Acta*, 703, 194-203.
- WALLACE, D. A., JOHNSTON, L. A., HUSTON, E., MACMASTER, D., HOUSLAY, T. M., CHEUNG, Y.-F., CAMPBELL, L., MILLEN, J. E., SMITH, R. A., GALL, I., KNOWLES, R. G., SULLIVAN, M. & HOUSLAY, M. D. 2005. Identification and Characterization of PDE4A11, a Novel, Widely Expressed Long Isoform Encoded by the Human PDE4A cAMP Phosphodiesterase Gene. *Molecular Pharmacology*, 67, 1920-1934.
- WAN, P. T. C., GARNETT, M. J., ROE, S. M., LEE, S., NICULESCU-DUVAZ, D., GOOD, V. M., PROJECT, C. G., JONES, C. M., MARSHALL, C. J., SPRINGER, C. J., BARFORD, D. & MARAIS, R. 2004. Mechanism of Activation of the RAF-ERK Signaling Pathway by Oncogenic Mutations of B-RAF. *Cell*, 116, 855-867.
- WANG, E. T., SANDBERG, R., LUO, S., KHREBTUKOVA, I., ZHANG, L., MAYR, C., KINGSMORE, S. F., SCHROTH, G. P. & BURGE, C. B. 2008a. Alternative isoform regulation in human tissue transcriptomes. *Nature*, 456, 470-476.
- WANG, M. Z., WU, J. Q., DENNISON, J. B., BRIDGES, A. S., HALL, S. D., KORNBLOTH, S., TIDWELL, R. R., SMITH, P. C., VOYKSNER, R. D., PAINE, M. F. & HALL, J. E. 2008b. A gel-free MS-based quantitative proteomic approach accurately measures cytochrome P450 protein concentrations in human liver microsomes. *Proteomics*, 8, 4186-4196.
- WANG, P., WU, P., OHLETH, K. M., EGAN, R. W. & BILLAH, M. M. 1999. Phosphodiesterase 4B2 Is the Predominant Phosphodiesterase Species and Undergoes Differential Regulation of Gene Expression in Human Monocytes and Neutrophils. *Molecular Pharmacology*, 56, 170-174.
- WANG, Y., BRUCE, A. T., TU, C., MA, K., ZENG, L., ZHENG, P., LIU, Y. & LIU, Y. 2011. Protein aggregation of SERCA2 mutants associated with Darier disease elicits ER stress and apoptosis in keratinocytes. *Journal of Cell Science*, 124, 3568-3580.

- WANG, Z. & STOUT, S. A. 2007. *Oil spill environmental forensics: fingerprinting and source identification*, Academic Press.
- WASINGER, V. C., CORDWELL, S. J., POLJAK, C., YAN, J. X., GOOLEY, A. A., WILKINS, M. R., DUNCAN, M. W., HARRIS, R., WILLIAMS, K. L. & SMITH, H. 1995. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis*, 16, 1090-1094.
- WELLBROCK, C., KARASARIDES, M. & MARAIS, R. 2004a. The RAF proteins take centre stage. *Nat. Rev. Mol. Cell Biol.*, 5, 875-885.
- WELLBROCK, C., KARASARIDES, M. & MARAIS, R. 2004b. The RAF proteins take centre stage. *Nat Rev Mol Cell Biol*, 5, 875-885.
- WELLS, J. N., BAIRD, C. E., WU, Y. J. & HARDMAN, J. G. 1975. Cyclic nucleotide phosphodiesterase activities of pig coronary arteries. *Biochimica et Biophysica Acta (BBA) - Enzymology*, 384, 430-442.
- WELLS, J. N., GARST, J. E. & KRAMER, G. L. 1981. Inhibition of separated forms of cyclic nucleotide phosphodiesterase from pig coronary arteries by 1,3-disubstituted and 1,3,8-trisubstituted xanthenes. *Journal of Medicinal Chemistry*, 24, 954-958.
- WHITTAM, R. & WHEELER, K. P. 1970. Transport Across Cell Membranes. *Annual Review of Physiology*, 32, 21-60.
- WIENKOOP, S., LARRAINZAR, E., GLINSKI, M., GONZÁLEZ, E. M., ARRESE-IGOR, C. & WECKWERTH, W. 2008. Absolute quantification of *Medicago truncatula* sucrose synthase isoforms and N-metabolism enzymes in symbiotic root nodules and the detection of novel nodule phosphoproteins by mass spectrometry. *Journal of experimental botany*, 59, 3307-3315.
- WOJNOWSKI, L., STANCATO, L. F., LARNER, A. C., RAPP, U. R. & ZIMMER, A. 2000. Overlapping and specific functions of Braf and Craf-1 proto-oncogenes during mouse embryogenesis. *Mechanisms of Development*, 91, 97-104.
- WOLSCHNER, C., GIESE, A., KRETZSCHMAR, H. A., HUBER, R., MORODER, L. & BUDISA, N. 2009. Design of anti- and pro-aggregation variants to assess the effects of methionine oxidation in human prion protein. *Proceedings of the National Academy of Sciences*, 106, 7756-7761.
- WU, Q., YUAN, H., ZHANG, L. & ZHANG, Y. 2012. Recent advances on multidimensional liquid chromatography–mass spectrometry for proteomics: From qualitative to quantitative analysis—A review. *Analytica Chimica Acta*, 731, 1-10.
- WU, S., LOURETTE, N. M., TOLIĆ, N., ZHAO, R., ROBINSON, E. W., TOLMACHEV, A. V., SMITH, R. D. & PAŠA-TOLIĆ, L. 2009. An Integrated Top-Down and Bottom-Up Strategy for Broadly Characterizing Protein Isoforms and Modifications. *Journal of Proteome Research*, 8, 1347-1357.
- WU, W. W., WANG, G., BAEK, S. J. & SHEN, R. F. 2006. Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. *J. Proteome. Res.*, 5, 651-8.
- YAMADA, H., YAMASHITA, T., DOMOTO, H. & IMOTO, T. 1990. Reaction of Hen Egg-White Lysozyme with Tetranitromethane: A New Side Reaction, Oxidative Bond Cleavage at Glycine 104, and Sequential Nitration of Three Tyrosine Residues. *Journal of Biochemistry*, 108, 432-440.
- YAMASAKI, R. B., OSUGA, D. T. & FEENEY, R. E. 1982. Periodate oxidation of methionine in proteins. *Analytical Biochemistry*, 126, 183-189.
- YAN, M. & TEMPLETON, D. J. 1994. Identification of 2 serine residues of MEK-1 that are differentially phosphorylated during activation by raf and MEK kinase. *Journal of Biological Chemistry*, 269, 19067-19073.
- YAO, X., FREAS, A., RAMIREZ, J., DEMIREV, P. A. & FENSELAU, C. 2001. Proteolytic 18O Labeling for Comparative Proteomics: Model Studies with Two Serotypes of Adenovirus. *Analytical Chemistry*, 73, 2836-2842.

- YASUDA, S., IDELL, S. & LIU, M. C. 2007. Generation and release of nitrotyrosine O-sulfate by HepG2 human hepatoma cells upon SIN-1 stimulation: identification of SULT1A3 as the enzyme responsible. *Biochem. J.*, 401, 497–503.
- YI, E. C., LI, X.-J., COOKE, K., LEE, H., RAUGHT, B., PAGE, A., ANELIUNAS, V., HIETER, P., GOODLETT, D. R. & AEBERSOLD, R. 2005. Increased quantitative proteome coverage with ¹³C/¹²C-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. *Proteomics*, 5, 380-387.
- YOKOYAMA, T., TAKANO, K., YOSHIDA, A., KATADA, F., SUN, P., TAKENAWA, T., ANDOH, T. & ENDO, T. 2007. DA-Raf1, a competent intrinsic dominant-negative antagonist of the Ras–ERK pathway, is required for myogenic differentiation. *The Journal of Cell Biology*, 177, 781-793.
- YU, Y.-Q., GILAR, M., LEE, P. J., BOUVIER, E. S. P. & GEBLER, J. C. 2003. Enzyme-Friendly, Mass Spectrometry-Compatible Surfactant for In-Solution Enzymatic Digestion of Proteins. *Analytical Chemistry*, 75, 6023-6028.
- ZABŁOTNA, E., DYSASZ, H., LESNER, A., JAŚKIEWICZ, A., KAŻMIERCZAK, K., MIECZNIKOWSKA, H. & ROLKA, K. 2004. A simple method for selection of trypsin chromogenic substrates using combinatorial chemistry approach. *Biochemical and Biophysical Research Communications*, 319, 185-188.
- ZHANG, B.-H. & GUAN, K.-L. 2000. Activation of B-Raf kinase requires phosphorylation of the conserved residues Thr598 and Ser601. *EMBO J*, 19, 5429-5439.
- ZHANG, R., SIOMA, C. S., WANG, S. & REGNIER, F. E. 2001. Fractionation of Isotopically Labeled Peptides in Quantitative Proteomics. *Analytical Chemistry*, 73, 5142-5149.
- ZHANG, X., FANG, A., RILEY, C. P., WANG, M., REGNIER, F. E. & BUCK, C. 2010. Multi-dimensional liquid chromatography in proteomics—A review. *Analytica Chimica Acta*, 664, 101-113.
- ZHU, W., SMITH, J. W. & HUANG, C.-M. 2010. Mass Spectrometry-Based Label-Free Quantitative Proteomics. *Journal of Biomedicine and Biotechnology*, 2010.
- ZORAGHI, R., CORBIN, J. D. & FRANCIS, S. H. 2004. Properties and Functions of GAF Domains in Cyclic Nucleotide Phosphodiesterases and Other Proteins. *Molecular Pharmacology*, 65, 267-278.